TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Flugsystemdynamik

Optimal Control Based Clearance of Flight Control Laws

Johannes Wolfgang Maximilian Diepolder

Vollständiger Abdruck der von der Fakultät für Luftfahrt, Raumfahrt und Geodäsie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Vorsitzender: apl. Prof. Dr.-Ing. habil. Christian W. M. Breitsamter

Prüfer der Dissertation:
1. Prof. Dr.-Ing. Florian Holzapfel
2. Prof. Joseph Z. Ben-Asher, Ph.D.

Die Dissertation wurde am 30.09.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Luftfahrt, Raumfahrt und Geodäsie am 24.02.2021 angenommen.

Acknowledgement

First of all, I would like to thank my family members Luise Diepolder, Wolfgang Diepolder, Sophie Stettner, and Elisabeth Diepolder for their support over the years. With their help I was able to obtain the opportunities and make the experiences that have encouraged me to finish this thesis.

I really enjoyed my work as a research associate at the Institute of Flight System Dynamics at the Technical University of Munich. In this time my doctoral supervisor Prof. Florian Holzapfel has provided me all possible opportunities for my professional growth. His ideas regarding the application of the methods developed in this thesis constitute an integral part of this work. Moreover, I would like to extend my sincere gratitude to Prof. Joseph Z. Ben-Asher for his guidance throughout my years as a doctoral candidate. In particular, his longer visits at the Institute of Flight system Dynamics in Munich and my time at the Technion in Israel have been the most crucial periods for the completion of my doctoral thesis. I am immensely grateful for his kind encouragement and help.

Special thanks go to all of my closest colleagues and friends Mattias Bittner, Rainer Matthias Rieck, Benedikt Grüter, Christopher Schropp, Patrick Piprek, David Seiferth, Tuğba Akman, Felix Schweighofer, Rubens Afonso, Barzin Hosseini, Varvara Turova, and Nikolai Botkin. The good atmosphere in the office and the numerous discussions have kept my spirits up and contributed a lot to this work.

Last but not least I want to thank the general public, the Deutsche Forschungsgemeinschaft (grants HO4190/8-1, HO4190/8-2), the Leibniz Supercomputing Centre (grant pr74lu) as well as the Deutsche Akademische Austauschdienst (DAAD) for their support regarding my work as a research associate and the opportunity to complete my doctoral thesis.

Munich, September 2020 Johannes Diepolder

Abstract

The application of optimal control theory for flight control law clearance is investigated. In this context a classification scheme based on the well-known categorization for aircraft pilot in the loop oscillations is proposed for different model alternatives (linear, quasi-linear, nonlinear). The nonlinear category is further refined based on the respective solution methodology, namely non-intrusive (black box), intrusive (white box), and hybrid (grey box) methods. The main concept for the approaches investigated in this thesis is the formulation of the clearance problem as a parameter dependent, state constrained, optimal control problem. For the linear and quasi-linear case the worst-case control is characterized based on the Minimum Principle. It is shown that the nominal clearance problem for linear models with state constraints and fixed terminal time yields a convex parameter optimization problem. The global solution for the discretized problem is achieved based on Linear Programming methods. In addition, parameter dependent methods (deterministic and stochastic) and methods for multi-criteria analysis are proposed. These approaches are tested using linearized aircraft models for the longitudinal and lateral motion. Multiple inputs are considered, such as pilot commands and wind gusts. Moreover, state constraints for rate and position limits of the servomechanism are introduced. In the nonlinear case novel approaches for hybrid and intrusive methods are presented. The hybrid approach is based on results from the linear domain and formulates the problem as an optimization problem in the switching points of the worst-case control function. Intrusive methods cover the solution of the parameter dependent clearance problem using direct optimal control methods. Formulations for the simultaneous determination of worst-case pilot control functions and parameter combinations are presented. Furthermore, post-optimal sensitivity analysis is performed to obtain the sensitivity of the worst-case solution with respect to fixed parameters. The approaches are applied to test the inner-loop controller of a nonlinear closed-loop model. Additionally, a method for computing reachable sets is proposed for the clearance task and illustrated by estimating the reachable set in the error subspace for a model reference adaptive controller.

Kurzfassung

Diese Arbeit befasst sich mit der Anwendung von Optimalsteuerungsmethoden für das Testen von Flugregelungssystemen. Die entwickelten Methoden sind in ein Klassifizierungsschema für die verschiedenen Modellierungsansätze (linear, quasi-linear, nichtlinear) eingebettet, welches auf dem bekannten Klassifizierungsschema für pilotenindizierte Oszillationen beruht. Des Weiteren werden die Ansätze für nichtlinearen Methoden in nicht-intrusive (black box), intrusive (white box) und hybride (grey box) Methoden unterteilt. Die grundlegende Modellierungsstrategie welche in dieser Arbeit verwendet wird, ist die Formulierung des Problems als parameterabhängiges Optimalsteuerungsproblem mit reinen Zustandsbeschränkungen. Im linearen sowie quasilinearen Fall wird die optimale Steuerfunktion basierend auf dem Minimumprinzip charakterisiert. Zudem werden Ansätze für die numerische Lösung des nominellen Problems mit fester Endzeit vorgestellt. Es wird gezeigt, dass dieses Problem als konvexes Optimierungsproblem mit Hilfe von Methoden der linearen Programmierung global gelöst werden kann. Weitere Verfahren ermöglichen die Lösung von parameterabhängigen (deterministisch und stochastisch), sowie multikriteriellen Problemen. Diese Methoden werden anhand der Untersuchung eines geregelten Systems für die Längs- und Seitenbewegung illustriert, welches durch die Linearisierung um einen Referenzzustand abgeleitet wird. Das verwendete Modell weist mehrere Eingänge für Pilotenkommandos sowie Windstörungen auf. Zudem sind Raten- und Positionsbeschränkungen für die Aktuatorik in der Problemformulierung berücksichtigt. Im nichtlinearen Fall werden hybride und intrusive Methoden vorgestellt. Der hybride Ansatz beruht auf einer Optimierung der Schaltstruktur für das nichtlineare Modell, welche basierend auf linearen Modellen bestimmt wird. Die intrusiven Verfahren hingegen basieren auf Methoden der direkten Optimalsteuerung. Die gemeinsame Optimierung der Kontrollfunktionen in Kombination mit Parametervariationen ist hierbei möglich. Zudem kann mit Hilfe von post-optimalen Sensitivitäten im Anschluss an die Optimierung der Einfluss von Parametern auf die optimale Lösung berechnet werden. Die Methoden werden für das Testen der inneren Regelschleife eines nichtlinearen Flugmodells verwendet. Weiterhin findet eine Methode zur Approximation von Erreichbarkeitsmengen für die Abschätzung des Fehlers im Folgeverhalten eines adaptiven Modellfolgereglers Anwendung.

Contents

Chapter 1

Introduction

1.1 Motivation

Testing can be considered one of the most important tasks in the development of engineering solutions. Especially for safety critical applications it is mandatory to show that systems can only be operated under safe conditions. Flight control laws, which are investigated in the this thesis, belong to this class of safety critical systems. As such, all relevant internal and external influences have to be considered during testing and the safety assessment of the entire flight system needs to be performed in the development process using quantifiable criteria. In this context, internal influences for flight control systems are, for example, quantities related to the structural properties, such as mass and mass distribution, as well as aerodynamic properties of the vehicle. An important external influence is the atmospheric impact due to the varying density of the surrounding air or wind gusts. Especially the latter can lead to hazardous flight conditions due to the strong effects of wind on the aerodynamic forces. Moreover, the pilot steering the aircraft needs to be considered due to the high degree of influence on the motion of the aircraft. Ideally, the control system should be designed in such as way that it is not possible for the pilot to bring the vehicle into an undesired flight condition. It is important to note that some of the mentioned quantities may be deterministic with a fixed range, known for the particular system, whereas others can be considered uncertain, i.e. subject to a distribution. Typical criteria for flight control systems involve bounds on aerodynamic quantities, structural loads, or tracking performance, to name a few. To illustrate the clearance task, consider the maximum Angle-of-Attack (AoA) exceeding criterion for wing-borne flight. The lift force, which is mainly produced by the wing of the aircraft, is heavily influenced by the aerodynamic AoA. An increased AoA leads to an elevated lift force but only until a certain limit is reached. If this limit is exceeded the lift force breaks down to a large extent due to the altered flow conditions on the wing and the aircraft stalls. This effect can lead

to dangerous conditions if the vehicle cannot be recovered. Typically, the flight control system ensures by means of built-in protection laws that the stall limit may not be breached and the vehicle is only operated within the safe limits of the AoA. However, note that numerous quantities have an influence on this particular criterion. Besides the structural quantities, such as those related to mass distribution, the aerodynamic properties of the aircraft which are, in most cases, subject to a considerable uncertainty need to be taken into account. Furthermore, disturbances such as wind gusts may produce very sudden changes in the aerodynamic AoA and are not known a-priori due the unpredictable nature of the atmospheric motion. The combination of these effects with any possible pilot command history and other unknowns, such as measurement noise, needs to be considered in order to ensure that the system can be kept within the prescribed limits. It should be clear from this example that the number of unknowns and the numerous combinations of all possible parameter values, envelope points, and disturbances adds to the complexity of the clearance task. In addition, one of the major challenges for a rigorous clearance of flight control systems is the high degree of nonlinearity associated with some of these effects. Besides the inherent nonlinearity of the aircraft dynamic model, a typical example for strongly nonlinear effects are internal limits for the servomechanism. For example, the actuators' maximum position and rate represent natural limits constraining the motion of the primary control surfaces and thus need to be considered in the dynamic model. Note that as soon as the actuators are driven into saturation the maneuverability of the aerial vehicle is impaired. This situation needs to be handled adequately by the control law, e.g. by means of suitable control allocation schemes. The illustrated complexity of the clearance task has led to the development of various advanced approaches for flight control law testing in the last decades. In particular, optimization based approaches have shown a high potential to efficiently identify worst-case scenarios. The state of the art for these approaches is discussed in the following section.

1.2 State of the Art

In Ref. [1] collected results for modern flight control law clearance methods are summarized. The main objective behind the approaches in this reference is the application of advanced flight control law clearance methods in industry practice. This is motivated by the fact that, in many cases, still the most popular approaches for flight control law testing are typically based on Gridding and standard Monte Carlo. For Gridding approaches the criterion under investigation is tested for a finite set of points in the space of all influential parameters. It is important to mention that the corner cases are usually of particular interest as worst-case combinations are often found at

extremal values of the parameters. However, in general, it is by no means certain that a violation of the criterion may not occur in between grid points and would thus remain undetected by this approach. Furthermore, the number of grid points grows rapidly with an increasing dimension of the parameter space due to the combinatorial nature of the problem (*curse of dimensionality*). For Monte Carlo methods the criterion is evaluated using a sampling-based approach and statistical information can be obtained from the results. Nevertheless, typically an enormous amount of realizations needs to be performed until the required confidence level is reached. This is true, in particular, for the standard Monte Carlo approach if very small failure domains are investigated. More advanced methods such as subset simulation methods (cf. Ref. [2]) tackle this problem by a successive construction of sets advancing towards the failure domain. Note that by moving towards this failure domain the algorithm concentrates not on the whole parameter space but on the region where further violations are expected. Note further that in some sense the idea to increase efficiency by advancing towards the failure domain naturally leads to the use of optimization assisted search methods for worst-case analysis. These approaches employ optimization methods to search the admissible parameter space for a violation of the criterion. In Ref. [3] several optimization methods are considered for the application to worst-case parameter search. The main categories are gradient-based methods (Sequential Quadratic Programming methods [4] and limited-memory BFGS with bound constraints [5]), gradient free methods (pattern search [6], constrained optimization by linear approximations [7], derivative free trust region [8]), and global methods (simulated annealing [9], genetic algorithms [10], multilevel coordinate search [11]). It is observed that the main challenges for the application of optimization based methods are expensive function evaluations, multiple local minima, noisy functions, and discontinuous derivatives. The application of these optimization methods to linear and nonlinear criteria is investigated in Refs. [12, 13]. For linear criteria the application of optimization based methods is benchmarked in Ref. [12] against the standard Gridding approach. Regarding the efficiency, it is noted that the Gridding approach exhibits an exponential computational complexity with respect to the number of parameter dimensions. This is not the case for the optimization methods employed in this reference. For the nonlinear benchmark in Ref. [13] the AoA exceeding criterion and the maximum load factor criterion are investigated. Besides aerodynamic derivatives, other parameters related to the mass distribution of the aircraft (center of gravity position and moment of inertia around the pitch axis) are used for the clearance task. For this study the optimization based methods outperform the Gridding methods. Particularly noteworthy is the fact that optimization based methods are able to identify considerable degradations in between grid points, that are regions not even explored by the Gridding approach. As for the linear benchmark, it is found that the (local) gradient-based methods perform more

efficiently, whereas global algorithms seem to yield more reliable worst-case results in this study. The authors suggest to combine local and global optimization schemes to harvest the effectiveness of local gradient-based methods together with the reliability of global methods. This line of research is continued in Refs. [14, 15] where two global schemes, namely differential evolution (DE) [16] and genetic algorithms (GA) are combined with a local search algorithm. The study considers the AoA limit exceeding criterion for the ADMIRE (Aerodata Model in Research Environment, cf. Ref. [17]) fighter aircraft model on a time horizon of ten seconds in a pull-up maneuver. Parameters related to aerodynamic coefficients and mass distribution are used. The algorithms with local search outperform the basic algorithms (without local search) both in computational time and global convergence. Overall, DE with local search shows superior performance for this study. In Ref. [18] the authors further investigate the effectiveness of the DIRECT optimization method [19] combined with a local search algorithm. The results are compared to the DE algorithm with local search in a similar setting as in Refs. [14, 15]. The results indicate that the DIRECT method with local search outperforms DE with local search. The study in Ref. [20] investigates the clearance problem using the ADMIRE model for the maximum load factor criterion by means of discretized pilot command inputs. The pitch and roll signals are discretized with ten amplitude points each. Three optimization algorithms, DE, GA, and a gradient-based search are compared. The gradient-based search performs well initially, but then gets trapped in a local solution. The DE method outperforms GA in this study. Following the developments in Ref. [1] the application of optimization based methods for civil aircraft is studied in Ref. [21]. Optimization based worst-case parameter search (cf. Refs. [22, 23]) and worst-case pilot input determination (cf. Refs. [24, 25]) are investigated using local and global optimization algorithms. The optimization based methods show excellent performance for the clearance task. For example in Ref. [25] optimization based methods easily find pilot inputs violating the maximum AoA exceeding criterion. Additionally, multi-criteria analysis is performed in Ref. [26] for a combined AoA exceeding and dive speed criterion. This combination is interesting from the point of view that the AoA exceeding criterion is typically more relevant at low velocities and the dive speed criterion for high velocities. Optimization methods and a Monte Carlo approach are compared using parametrized pilot commands (longitudinal stick, lateral stick, pedal, air-brakes, and throttle position) in combination with other parameters (altitude, speed, mass, and center of gravity position). The overall proposed validation workflow for the implementation of optimization based methods in industry practice is to first perform a Monte Carlo analysis and then start the optimization based search using global algorithms from identified worst-cases. The results from the global algorithms are then refined by a local algorithm and complemented by a sensitivity analysis based on simulations. In Ref. [27] the carefree handling clearance

problem for the Eurofighter Typhoon is investigated. The pitch and roll stick command with 20 parameters each (ten time points and corresponding amplitude parameters) in combination with other parameters is used to test the control system. The global optimization algorithm MAGO (Multi Dynamics Algorithm for Global Optimization) [28] is applied in combination with a local search algorithm and compared to two global algorithms (particle swarm optimization (PSO) [29] and SPS-LSHADE-EIG [30]). The proposed algorithm shows good performance compared to the other global schemes. It is important to observe that the discretization of the control signal adds a dominant amount of parameters to the clearance problem which is undesirable for global optimization algorithms that perform typically well in low dimensional parameter spaces. One key observation in this context is that the clearance problem with input commands can be viewed as an optimal control problem. In other words, a worst-case control function is searched for which minimizes (or maximizes) the criterion under investigation. Recently, the authors in Ref. [31] proposed to state the problem of finding the worst-case pilot inputs based on optimal control methods. This approach is illustrated using a short period approximation of the ADMIRE model. The advantages to regard the clearance problem from this point of view are on the one side a rich theoretical background in optimal control theory and on the other side powerful numerical tools which have been developed in the past decades for the solution of dynamic optimization problems. Due to the appealing potential of the idea this line of research is followed for the approaches developed in this thesis.

1.3 Goals and Contributions

The primary goal of this thesis is to lay ground for the application of optimal control based clearance methods in industry practice. This includes advancements of the theoretical and practical basis, which can be used for the implementation of novel test procedures for realistic flight control systems. As such, a framework for the application of optimal control based methods for typical flight control law clearance tasks is developed. The contributions can be summarized as follows:

I Development of a model classification scheme for optimal control based clearance approaches for different modeling alternatives (linear, quasi-linear, nonlinear) and corresponding solution strategies (non-intrusive, intrusive, hybrid).

II Development of clearance methods for linear and quasi-linear systems including approaches for

- deterministic, parameter dependent systems,
- uncertain, parameter dependent systems,

- multi-criteria and reachability analysis.

III Development of approaches for nonlinear clearance tasks including

- clearance methods based on direct optimal control methods,
- the application of post-optimal sensitivity analysis, and
- a hybrid clearance method based on the optimization of switching times.

IV Implementation of a counter optimization library (*COLIBRY*) with high level interfaces to MATLAB® and Simulink® for the application of the approaches developed in this thesis as well as contributions to the implementation of the FSD optimal control toolbox *FALCON.m*[1] for the solution of nonlinear clearance problems.

These contributions are described in detail in the following sections.

Contribution I: Model classifications for optimal control based clearance

Due to the novelty of the optimal control approach in the context of flight control law clearance it is useful to introduce a classification scheme. To the knowledge of the author a classification in this form has not been proposed in public literature. Optimal control based clearance approaches may be classified based on the following model categories:

- **Category I**: Completely linear system analysis.
- **Category II**: Quasi-linear system analysis including internal limits in the servomechanism.
- **Category III**: General nonlinear system analysis.

This categorization is based on the model categories for the analysis of pilot induced oscillations (cf. Ref. [32]). It is noteworthy that, using the tools developed in this thesis, several problems from Cat I and Cat II can be solved to global optimality. This is not the case for Cat III models as the theoretical and algorithmic development has not yet lead to global optimal control methods which can operate in high dimensional state-spaces. As such, the goal for Cat III models is to develop effective optimal control based methods which are comprehensive for the application in industry practice and can efficiently be applied to large-scale models. For Cat III models it is helpful to increase the classification granularity by the introduction of the following three solution methodologies:

[1] www.falcon-m.com

- **Non-intrusive**: The dynamic nature of the problem is regarded as a black box, typically using the combination of parameter optimization and simulation methods in a single sweep.
- **Intrusive**: The dynamic system is exposed to the algorithm as a white box, e.g. in the form of a single point execution model.
- **Hybrid**: An intermediate modeling alternative is used in which worst-case inputs are constructed based on a reduced white box model and applied to the full simulation model.

Note that this classification is only introduced for Cat III models as Cat I and Cat II models, that are linear and quasi-linear models, can for the most part be treated in an intrusive manner. The majority of the approaches discussed in Sec. 1.2 belong to Cat III non-intrusive methods. The optimal control approach from Ref. [31] however may be viewed as a Cat II intrusive method due to the fact that it approaches the problem based on the characteristics of the dynamic model at a single time point.

Contribution II: Development of clearance methods for parameter dependent (quasi-)linear systems

Following the ideas in [31] the theoretical analysis in this thesis characterizes the solution of Cat I and Cat II models based on the Minimum Principle. This analysis includes the state constrained case which is used to model internal limits of the servomechanism. Four formulations are developed for the application to flight control law testing under Cat I and Cat II:

- **Nominal Case**:
  It is shown that the dynamic optimization problem for Cat I and Cat II models with fixed terminal time can be transcribed into a linear programming problem. Worst-case disturbance inputs (e.g. wind gusts and pilot commands) and state constraints related to actuator position and rate limits are considered under this approach. It is particularly noteworthy that this formulation allows for an efficient solution of the nominal, discretized problem to global optimality.
- **Parameter Dependent Case (Deterministic)**:
  Based on the nominal formulation, a bi-level method is proposed which can be used to solve dynamic optimization problems for linear parameter varying problems. On the lower level a discretized optimal control problem is solved for fixed values of the parameters using linear programming. On the upper level the typically low dimensional search is performed by a parameter optimization

algorithm. For this upper level problem gradient-based algorithms using post-optimal sensitivities and gradient-free, global methods are proposed.

- **Parameter Dependent Case (Stochastic)**:
  In the context of uncertainty quantification a generalized polynomial chaos approach with stochastic collocation is applied to the clearance task. Statistical information about the clearance criterion is obtained using the solution of deterministic optimal control problems at stochastic collocation points and the projection into the space of orthogonal polynomial basis functions. The fast convergence of the spectral method is verified based on the comparison to a Monte Carlo analysis.

- **Multi-Criteria and Reachability Analysis**:
  A numerical scheme for the simultaneous treatment of multiple criteria is developed. This scheme approximates the boundary of the reachable set in the space of criteria up to a given tolerance. The proposed method is applicable to very high dimensional state spaces and is able to consider input and state constraints.

Contribution III: Development of clearance methods for general nonlinear systems

For the nonlinear domain both intrusive and hybrid Cat III approaches are developed. As for the linear case, the CAT III intrusive approaches include the nominal and parameter dependent case as well as multi-criteria and reachability analysis:

- **Nominal and Parameter Dependent Case**:
  The problem formulation for the nonlinear case is directly able to treat the parameter dependent case. As for the linear case, the numerical solution of the nonlinear optimization problem is achieved through the discretization of the problem using direct transcription methods. Numerical evidence suggests that due to the resulting sparse optimization problem the solution of the clearance problem is efficient and produces plausible results from an engineering point of view. Moreover, results from post-optimal sensitivity analysis are used to study the effect of parameters on the *worst-case* optimal solution. This information is deemed useful to determine influential parameters. The nonlinear closed-loop model used for the study in this thesis is a full six-degree-of freedom flight dynamic model including aeroelastic effects, sensor models, the inner-loop controller, and an actuation system with rate and position limits. As such, the optimal control problem solved for the clearance task has 82 states, six controls modeling pilot commands and wind disturbances and depends on parameters related to aerodynamic derivatives and mass distribution.

- **Multi-Criteria and Reachability Analysis**:
  Reachability analysis for the nonlinear problem is treated through the application of the distance fields on grids (DFOG) method [33]. Besides the general applicability of the approach it is observed (in congruence with Ref. [33]) that the local nature of the solution for nonlinear optimal control problems is mitigated to a certain extent through the solution of a large number of optimal control problems on a grid. However, an obvious downside of this approach is the considerable computational effort which requires an efficient implementation. The approach is illustrated using an example problem for the approximation of the reachable set in the tracking error subspace of a nonlinear model reference adaptive controller.

Besides the Cat III intrusive methods a hybrid approach is developed which extends the tools developed for linear models to the nonlinear domain. First, a linear model is employed to construct the worst-case command. It is observed that the switching structure obtained from the linear model present a good parametrization of the control function which may also be applied to the nonlinear model. As such, the problem is reformulated for the Cat III hybrid approach as an optimization in the switching times of the worst-case control obtained from the linear model. In addition to the good initial guess from the linear model, the number of switches is typically very low which greatly reduces the parameter space to be searched. The appealing simplicity paired with its effectiveness makes this approach directly applicable in industry practice. Note that the additional modeling effort for the application is very low as linearization results for closed-loop models and a simulation model are typically available during the development process.

It is important to mention that the approaches related to Cat III models are computationally quite demanding. Consider for example the reachability analysis using the DFOG method which requires the solution of potentially several thousands of nonlinear optimal control problems. As such, the exploitation of problem characteristics in the numerical implementation, e.g. the exploitation of sparse structures for the dynamic systems or the discretization methods, is of paramount importance. This leads to the last contribution, that is the software development for clearance methods.

Contribution IV: Software development for clearance methods

The developed approaches for Cat I and Cat II models as well as the hybrid method for Cat III models are implemented by the author in a counter optimization library (*COLIBRY*) with an interface to MATLAB® and Simulink®. The main objective of this software package is to allow for the direct application of the approaches developed in this thesis to realistic clearance problems. For Cat III intrusive methods the FSD Optimal control software *FALCON.m* is employed which has since the initial version been

co-developed by the author with major contributions from the authors of Refs. [34, 35]. This tool allows for an efficient numerical solution of nonlinear optimal control problems using interfaces to third-party solvers. Besides the general development for the core capabilities of the tool the author has implemented a Simulink® interface and contributed to the implementation of post-optimal sensitivities.

1.4 Outline of the Thesis

This thesis is structured as follows. Firstly, the mathematical background is outlined in chapter 2. This theoretical background summarizes well established results from function interpolation and approximation, quadrature approximations, and the numerical solution of initial value problems including sensitivity analysis. The chapters 3 and 4 present the theoretical background for unconstrained and constrained parameter optimization problems. Besides the necessary and sufficient conditions of optimality, numerical methods for linear, quadratic, and general nonlinear optimization problems are presented. Moreover, important results from post-optimal sensitivity analysis are discussed. In chapter 5 the step is taken from finite dimensional parameter optimization problems to infinite dimensional optimal control problems in function space. As such, chapter 5 revisits first-order necessary conditions for differential and integral formulations and summarizes theoretical results for control affine and state dependent optimal control problems. The following chapter 6 discusses direct solution approaches, namely shooting and collocation methods, for the numerical treatment of optimal control problems. The connection between the continuous time optimal control problem and the discretized versions using direct transcription schemes is established based on the derivation of a co-state estimation schemes for segmented Legendre-Gauss methods in differential form. Chapter 7 presents the proposed framework for optimal control based worst-case analysis developed in this thesis. Besides the classification of optimal control based testing, particular modeling choices related to the servomechanism of the system are put into context. The proposed classification is the basis for the approaches presented in the following two chapters. Chapter 8 discusses the first two categories, namely the linear and quasi-linear case (Cat I and Cat II) with and without parameter dependencies. First, the optimal control is characterized based on the Minimum Principle and the developed testing procedures for the nominal and parameter dependent case are presented. Furthermore, the application of generalized polynomial chaos for uncertainty quantification in the context of optimal control based testing is proposed. Moreover, an efficient reachable set algorithm for Cat I and Cat II models is developed. Chapter 9 discusses the general nonlinear case from the last modeling category (Cat III). Both intrusive and hybrid methods are

presented. These methods are benchmarked using a nonlinear closed-loop model. In addition, an intrusive method for performing reachable set approximations is applied to a F16 fighter model with a nonlinear model reference adaptive controller.

Chapter 2

Mathematical Preliminaries

The flight control law clearance approaches presented in this thesis are based on the application of direct optimal control methods. These methods transcribe the clearance optimal control problem into a discretized form. Transcription schemes of this type primarily rely on mathematical tools related to function parametrization and the solution of initial value problems. On the one side, B-splines (cf. Sec. 2.6) are widely used in this context for the parametrization of control functions due to their local support property. On the other side, the solution of initial value problems (cf. Sec. 2.3 and Sec. 2.4) is, in most cases, achieved numerically using Runge-Kutta methods. In addition, segmented collocation methods (cf. Sec. 2.5) are commonly employed for this task. As such, two different forms of collocation methods, namely the differential and integral form, are derived in this chapter and their connection to implicit Runge-Kutta methods is established. Lagrange interpolation (cf. Sec. 2.1) and Legendre-Gauss quadrature methods (cf. Sec. 2.2) represent the mathematical basis for these methods and are, as such, discussed in the beginning of this chapter.

2.1 Lagrange Interpolation

One of the basic tools for function interpolation is the use of a Lagrange interpolation polynomial [36, 37]. In the following the interpolation of a function $f : \mathbb{R} \to \mathbb{R}$ at knots $t_i, i = 0, \ldots, n-1$ is considered. These knots are defined on the Lagrange interpolation grid $\mathbb{G}_n^L$ of cardinality n:

$$\mathbb{G}_n^L := \{t_i : i = 0, \ldots, n-1, t_{i+1} > t_i\} \tag{2.1}$$

The Lagrange interpolation for n values $f(t_i), i = 0, \ldots, n-1$ can be written as

$$f(t) \approx \sum_{i=0}^{n-1} l_i(t) f(t_i), \tag{2.2}$$

using the basis functions $l_i : \mathbb{R} \to \mathbb{R}$ defined by

$$l_i(t) := \prod_{\substack{j=0 \\ j \neq i}}^{n-1} \frac{t - t_j}{t_i - t_j}. \tag{2.3}$$

It is easily verified that each of the Lagrange basis functions has value one at a single knot point t_i and zero at all others, i.e.

$$l_i(t_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

Please refer to Fig. 2.1 for an illustration of these basis functions.

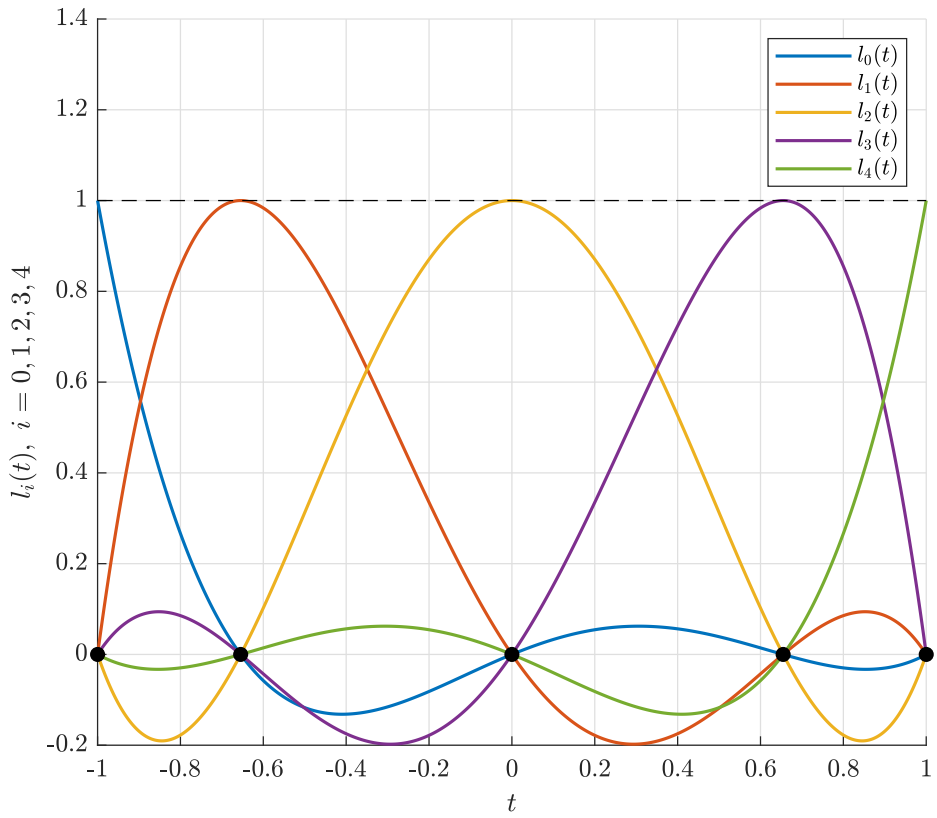


Figure 2.1: Illustration of the first five Lagrange basis function values $l_0(t), \ldots, l_4(t)$ on the interval $t \in [-1, 1]$. Abscissas for the interpolation points are highlighted with black markers.

From a practical point of view it is important to mention that the representation in Eq. (2.3) is not the most favorable form for numerical implementations. Among other reasons, this is contributed to the fact that $\mathcal{O}(n^2)$ additions and multiplications are required for each evaluation. Thus, in practice it is preferable to implement Eq. (2.2) together with the basis functions defined in Eq. (2.3) in a different form, termed the *barycentric* form [37]

$$f(t) \approx \sum_{i=0}^{n-1} \prod_{\substack{j=0 \\ j\neq i}}^{n-1} \frac{t-t_j}{t_i-t_j} f(t_i) = \frac{\sum_{i=0}^{n-1} \frac{w_i}{t-t_i} f(t_i)}{\sum_{i=0}^{n-1} \frac{w_i}{t-t_i}}, \tag{2.5}$$

with the barycentric weights $w_i, i = 0, \ldots, n-1$

$$w_i := \frac{1}{\prod_{\substack{j=0 \\ j\neq i}}^{n-1} (t_i - t_j)}. \tag{2.6}$$

If the interpolation grid $\mathbb{G}_n^L$ does not change, the barycentric weights w_i can be precomputed resulting in $\mathcal{O}(n)$ operations for a single evaluation of the polynomial. Moreover, it is important to mention that the selection of the knots t_i is crucial for the error between the true function value $f(t)$ and the interpolating polynomial. The intuitive choice of equidistant points in fact leads to *Runge's phenomenon* (cf. Ref. [38]) close to both ends of the interpolation grid $\mathbb{G}_n^L$. This phenomenon becomes increasingly visible for higher orders of the polynomial. To counteract this phenomenon the grid can be chosen in such a way that the density of points is higher close to the ends of the interpolation interval. Particularly, the sets of Legendre-Gauss (LG), Legendre-Gauss-Radau (LGR), and Legendre-Gauss-Lobatto (LGL) points (cf. Sec. 2.2) alleviate the negative effect of this phenomenon. The classical example for Runge's phenomenon is illustrated in Fig. 2.2. In this figure, the interpolation of function values from

$$f(t) = \frac{1}{1+t^2}, \tag{2.7}$$

is shown for equally spaced knots and abscissas based on LGL points. Note that the interpolation using equally spaced points leads to large oscillations at both ends of the interval. This phenomenon is not observed for the interpolation using the LGL points.

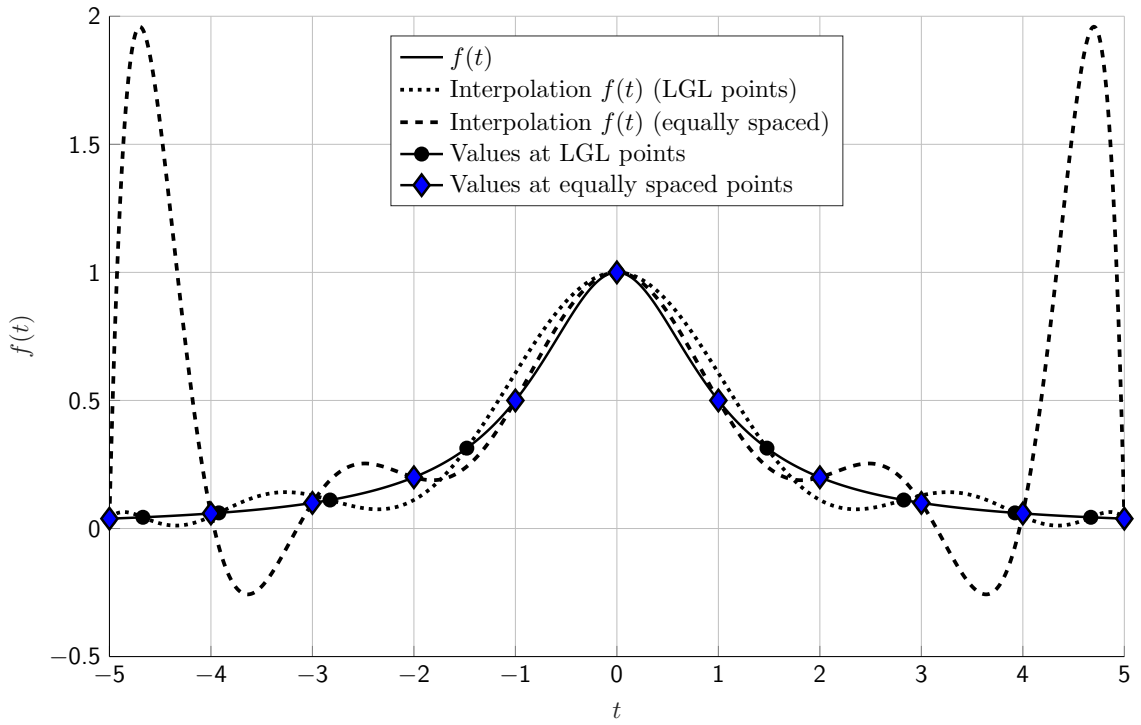


Figure 2.2: Illustration of the classical example for Runge's phenomenon with $f(t) = 1/(1+t^2)$ on the the interval $t \in [-5, 5]$. Equally spaced interpolation points lead to oscillations towards both ends of the interval. For LGL points this effect is not observed.

2.2 Legendre-Gauss Quadrature

In the following numerical integration methods are discussed. These integration methods can be used to approximate the integral of the function $f : \mathbb{R} \to \mathbb{R}$ on a closed interval $t \in [a, b]$. An important class of numerical methods in this context is based on quadrature rules [39, 40]. For schemes of this class the integrand f is evaluated on a quadrature grid $\mathbb{G}_n^Q$ of cardinality n

$$\mathbb{G}_n^Q := \{t_i : i = 0, \ldots, n-1, t_{i+1} > t_i, t_0 \geq a, t_{n-1} \leq b\}, \tag{2.8}$$

and a weight w_i is associated with each of the function values. The sum over these weighted function values yields an approximation to the integral of the form:

$$\int_a^b f(t)\, dt \approx \sum_{i=0}^{n-1} w_i f(t_i) \tag{2.9}$$

The approximation of an integral using a quadrature of this form is illustrated in Fig. 2.3. Without loss of generality, the function f is in the following considered on the normalized interval $[-1, 1]$. Note that the function f can always be mapped from the

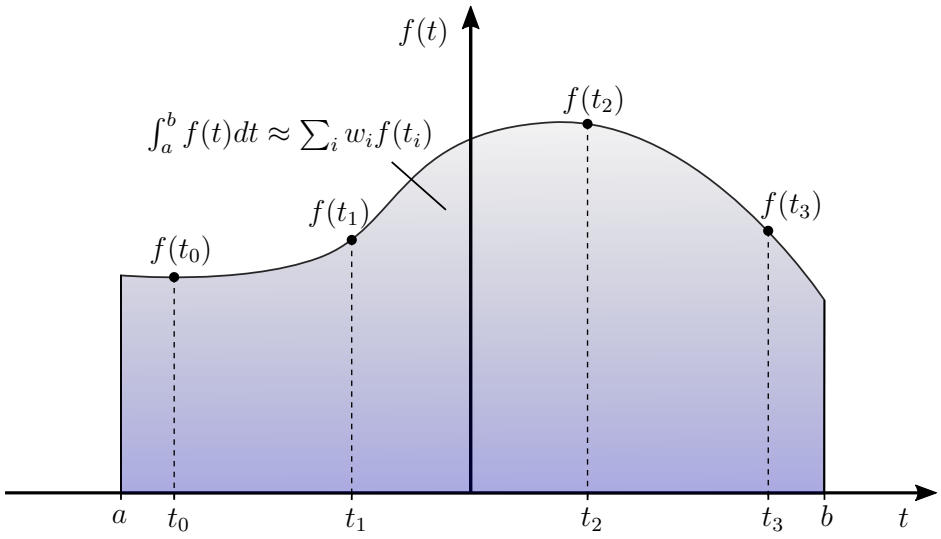


Figure 2.3: Illustration of the numerical approximation for the definite integral of the function f on the interval $[a, b]$ by a weighted sum of function values at specific locations.

interval $[a, b]$ to this normalized interval using an affine transformation for t. Observe that for the quadrature approximation (2.9) there are n abscissas $t_i,\ i = 0, \ldots, n-1$ and n weights $w_i,\ i = 0, \ldots, n-1$ to be defined. The primary motivation of how to select these points and weights is to obtain a quadrature method with a high accuracy given n points. In this work, the class of Legendre-Gauss quadrature rules are considered due to their fast convergence to the true value of the integral. The points $t_i,\ i = 0, \ldots, n-1$ for these quadrature rules can be obtained from the roots of the Jacobi polynomials $P_p^{(\alpha,\beta)} : \mathbb{R} \to \mathbb{R},\ \alpha, \beta > -1$ of degree p. Jacobi polynomials belong to the class of orthogonal polynomials and can be defined in terms of their recurrence relation [41]

$$\begin{aligned} P_p^{(\alpha,\beta)}(t) :=& \frac{(2p+\alpha+\beta-1)\left((2p+\alpha+\beta)(2p+\alpha+\beta-2)\,t+\alpha^2-\beta^2\right)}{2p\,(p+\alpha+\beta)(2p+\alpha+\beta-2)} P_{p-1}^{(\alpha,\beta)}(t) \\ &- \frac{2(p+\alpha-1)(p+\beta-1)(2p+\alpha+\beta)}{2p\,(p+\alpha+\beta)(2p+\alpha+\beta-2)} P_{p-2}^{(\alpha,\beta)}(t), \end{aligned} \tag{2.10}$$

for $p > 1$ together with the first two Jacobi polynomials

$$P_0^{(\alpha,\beta)}(t) := 1, \tag{2.11}$$

$$P_1^{(\alpha,\beta)}(t) := \frac{1}{2}(\alpha+\beta+2)\,t + \frac{1}{2}(\alpha-\beta). \tag{2.12}$$

Quadrature methods based on the abscissas obtained from the roots of the Jacobi polynomials are exact for polynomials up to a degree of $2n-1-\alpha-\beta$. The following special cases are of particular interest [40]:

Legendre-Gauss $(\alpha = 0, \beta = 0)$
The quadrature grid $\mathbb{G}_n^{LG}$ of cardinality n contains the abscissas of the LG quadrature rule based on the roots of $P_n^{(0,0)}$:

$$\mathbb{G}_n^{LG} := \{t_i^{LG} : P_n^{(0,0)}\left(t_i^{LG}\right) = 0, i = 0, \ldots, n-1\},\ n > 0. \tag{2.13}$$

All LG quadrature points defined in $\mathbb{G}_n^{LG}$ are internal, i.e. neither of the interval endpoints is included in the set of quadrature points. This quadrature method provides an exact integration of polynomials up to a degree of $2n-1$.

Legendre-Gauss-Radau $(\alpha = 0, \beta = 1)$**,** $(\alpha = 1, \beta = 0)$
For the LGR points either the initial point or the endpoint of the interval is included in the set of quadrature points. The quadrature grid $\mathbb{G}_n^{LGR}$ for the abscissas of the LGR quadrature rule is defined as

$$\mathbb{G}_n^{LGR} := \{t_i^{LGR} : t_0^{LGR} = -1, P_{n-1}^{(0,1)}\left(t_j^{LGR}\right) = 0,\ j = 1, \ldots, n-1\},\ n > 0. \tag{2.14}$$

The distribution of LGR points is not symmetric with respect to the origin. Thus, it is possible to use a *flipped* set of points which includes the end-point instead of the initial point. In this case, the set of points are usually referred to as the flipped Legendre-Gauss-Radau (LGRF) points and the corresponding quadrature grid $\mathbb{G}_n^{LGRF}$ is defined as

$$\mathbb{G}_n^{LGRF} := \{t_i^{LGRF} : P_{n-1}^{(0,1)}\left(t_j^{LGRF}\right) = 0,\ j = 0, \ldots, n-2, t_{n-1}^{LGRF} = 1, \},\ n > 0. \tag{2.15}$$

Radau quadrature methods are exact for polynomials up to a degree $2n-2$, i.e. one degree less compared to the LG quadrature rule.

Legendre-Gauss-Lobatto $(\alpha = 1, \beta = 1)$
For the LGL points both the initial point and the endpoint of the interval are included in the set of quadrature points. This leads to the following definition of the LGL quadrature grid $\mathbb{G}_n^{LGL}$ based on the Jacobi polynomials with $\alpha = 1$ and $\beta = 1$:

$$\mathbb{G}_n^{LGL} := \{t_i^{LGL} : t_0^{LGL} = -1, P_{n-2}^{(1,1)}\left(t_j^{LGL}\right) = 0,\ j = 1, \ldots, n-2, t_{n-1}^{LGL} = 1\},\ n > 1 \tag{2.16}$$

The LGL quadrature is exact for polynomials up to a degree of $2n-3$, i.e. two degrees less than the LG method.

Based on the location of the quadrature points $t_i, i = 0, \ldots, n-1$ the corresponding weights $w_i, i = 0, \ldots, n-1$ can be easily constructed.

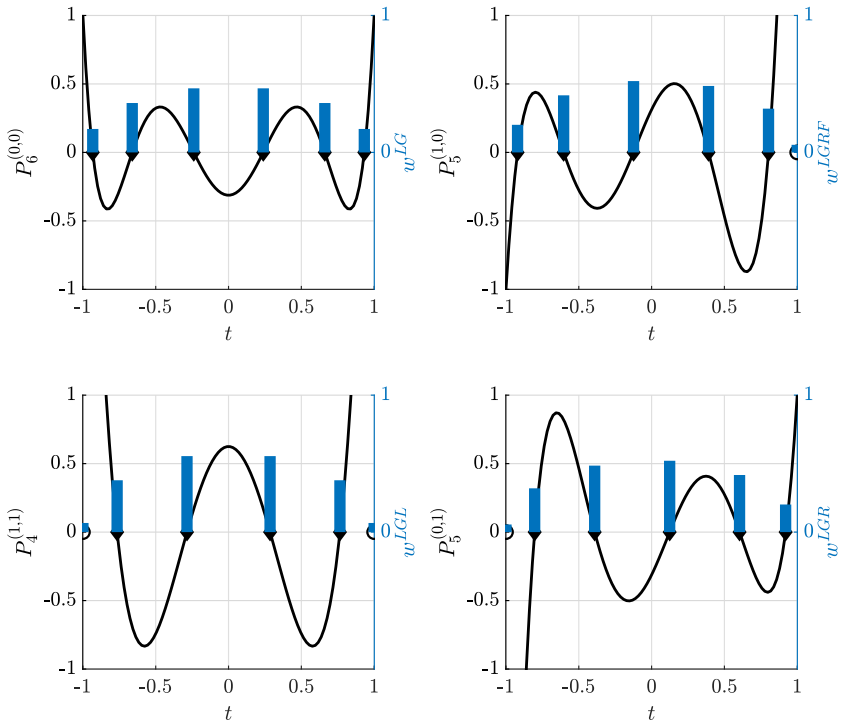


Figure 2.4: Illustration of the LG (top left), LGL (bottom left), LGRF (top right), and LGR (bottom right) quadrature weights and abscissas. The points in the strict interior of the interval (diamond markers) are obtained from the roots of the respective Jacobi polynomial. Fixed quadrature points at the beginning and end of the intervals for the set of LGR, LGRF, and LGL points are marked by a circle. Moreover, the blue bars illustrate the weights associated with each of the quadrature points.

Consider a Lagrange interpolation polynomial for the n values of the integrand $f(t_i)$ at the quadrature abscissas t_i:

$$f(t) \approx \sum_{i=0}^{n-1} l_i(t) f(t_i) \tag{2.17}$$

Integration on both sides

$$\int_{-1}^{1} f(t)\, dt \approx \sum_{i=0}^{n-1} \int_{-1}^{1} l_i(t)\, dt\, f(t_i) \tag{2.18}$$

and matching the coefficients with Eq. (2.9) yields

$$w_i := \int_{-1}^{1} l_i(t)\, dt,\ i = 0, \ldots, n-1. \tag{2.19}$$

The quadrature points and corresponding weights associated with each of the discussed quadrature methods are depicted in Fig. 2.4 for $n = 6$. To illustrate the convergence of the four quadrature methods presented in this section consider the integral

$$v_E = \int_{-1}^{1} e^t dt, \tag{2.20}$$

i.e. $f(t) = e^t$. The absolute errors between the exact value v_E of the integral and the values obtained from a quadrature approximation of the form

$$v_Q = \sum_{i=0}^{n-1} w_i e^{t_i}, \tag{2.21}$$

for all four methods (LG, LGR, LGRF, and LGL) is shown in Fig. 2.5. Note the exponential decay of the error for all methods which illustrates the fast convergence of the Legendre-Gauss quadrature schemes.

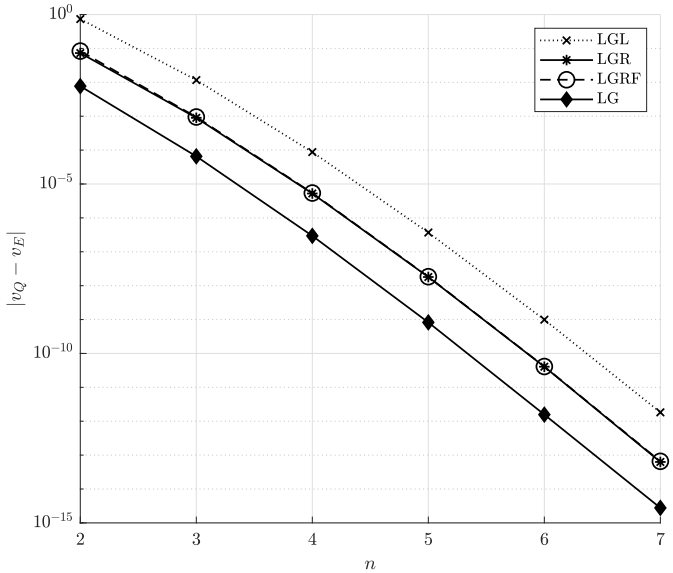


Figure 2.5: Convergence plot of the error on a logarithmic scale for $n = 2, \ldots, 7$ points regarding the numerical approximation of the integral of e^t on the interval $t \in [-1, 1]$ for the LG, LGR, LGRF, and LGL quadrature methods.

2.3 Numerical Methods for Initial Value Problems

Initial value problems may be stated as follows: Find a function $\mathbf{x} : I^t \to \mathbb{R}^{n_x}$ on the interval

$$I^t := [t_0, t_f] \subset \mathbb{R}, t_f > t_0, \tag{2.22}$$

starting at the initial value $\mathbf{x}_0 \in \mathbb{R}^{n_x}$, which fulfills the differential equation $\dot{\mathbf{x}}(t) := \mathbf{f}(\mathbf{x}(t), t)$ with $\mathbf{f} : \mathbb{R}^{n_x} \times I^t \to \mathbb{R}^{n_x}$, i.e. :

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t), \qquad \mathbf{x}(t_0) = \mathbf{x}_0, \qquad t \in I^t \tag{2.23}$$

In the following Runge-Kutta methods [38, 42], which belong to the class of one-step methods, are considered for the numerical solution of initial value problems. These methods approximate the solution of the initial value problem (2.23) on the integration grid $\mathbb{G}^I_{N+1}$ of cardinality $N + 1$

$$\mathbb{G}^I_{N+1} := \{t^{[i]} : t_0 = t^{[0]} < t^{[1]} < \ldots < t^{[N-1]} < t^{[N]} = t_f\}, \tag{2.24}$$

with mesh lengths $h^{[i]} = t^{[i+1]} - t^{[i]}, i = 0, \ldots, N-1$ by a grid function $\mathbf{x}^{[i]} : \mathbb{G}^I_{N+1} \to \mathbb{R}^{n_x}$. In particular, methods of this class construct the solution for a single step from $t^{[i]}$ to $t^{[i+1]}$ without using information from previous steps $(t^{[i-1]}, t^{[i-2]}, ...)$ or next steps $(t^{[i+2]}, t^{[i+3]}, ...)$. For the following description, it is useful to introduce the short notation $f_k^{[i]}(\mathbf{x}(t), t) = h^{[i]} f_k(\mathbf{x}(t), t), k = 0, \ldots, n_x - 1$ for the scaled form of the right-hand side function of each component. For Runge-Kutta methods n_s intermediate stages $\mathbf{s}_j^{[i]} \in \mathbb{R}^{n_x}, j = 0, \ldots, n_s - 1$ for each integration interval $[t_i, t_{i+1}]$ are defined by the stage equations

$$s_{j,k}^{[i]} = x_k^{[i]} + \sum_{q=0}^{n_s-1} a_{j,q} f_k^{[i]} \left(\mathbf{s}_q^{[i]}, t^{[i]} + c_q h^{[i]}\right). \tag{2.25}$$

Equivalently, these stage equations (2.25) can be written in matrix form:

$$\begin{bmatrix} s_{0,k}^{[i]} \\ \vdots \\ s_{n_s-1,k}^{[i]} \end{bmatrix} = \mathbf{1} x_k^{[i]} + \begin{bmatrix} a_{0,0} & \ldots & a_{0,n_s-1} \\ \vdots & \ddots & \vdots \\ a_{n_s-1,0} & \ldots & a_{n_s-1,n_s-1} \end{bmatrix} \begin{bmatrix} f_k^{[i]} \left(\mathbf{s}_0^{[i]}, t^{[i]} + c_0 h^{[i]}\right) \\ \vdots \\ f_k^{[i]} \left(\mathbf{s}_{n_s-1}^{[i]}, t^{[i]} + c_{n_s-1} h^{[i]}\right) \end{bmatrix} \tag{2.26}$$

The values $x_k^{[i+1]}, k = 0, \ldots, n_x - 1$ at the next integration grid point $t^{[i+1]}$ are obtained by the following quadrature rule with weights $b_j, j = 0, \ldots, n_s - 1$:

$$x_k^{[i+1]} = x_k^{[i]} + \sum_{j=0}^{n_s-1} b_j f_k^{[i]} \left(\mathbf{s}_j^{[i]}, t^{[i]} + c_j h^{[i]}\right) \tag{2.27}$$

The coefficients $a_{j,q}$ and c_q in Eq. (2.25), as well as b_j in Eq. (2.27), are usually represented in a tabular form termed *Butcher-tableau*

$$\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b} \end{array} \tag{2.28}$$

with $\mathbf{A} \in \mathbb{R}^{n_s \times n_s}$, $\mathbf{b} \in \mathbb{R}^{1 \times n_s}$, and $\mathbf{c} \in \mathbb{R}^{n_s \times 1}$. By normalization the sum of the weights $\sum_{j=0}^{n_s-1} b_j, j = 0, \ldots, n_s - 1$ is equal to one. In addition, for each row (indices $j = 0, \ldots, n_s - 1$) in $\mathbf{A}$ the sum over the columns, i.e. $\sum_{k=0}^{n_s-1} a_{j,k}$, is equal to c_j. Moreover, the coefficients defined in a Butcher-tableau (2.28) have to fulfill certain conditions depending on the order of the Runge-Kutta scheme. Table 2.1 summarizes the conditions up to order three. Note that for a method to have order p all lower order conditions must be fulfilled as well.

Table 2.1: Runge-Kutta order conditions up to order three [43]

Order	Conditions
1	$\sum_{i=0}^{n_s-1} b_i = 1$
2	$\sum_{i=0}^{n_s-1} b_i c_i = \frac{1}{2}$
3	$\sum_{i=0}^{n_s-1} b_i c_i^2 = \frac{1}{3}, \quad \sum_{i,j=0}^{n_s-1} b_i a_{i,j} c_j = \frac{1}{6}$

Furthermore, in case the stage integration matrix $\mathbf{A}$ is strictly lower triangular, i.e. $a_{i,j} = 0$ for all $i < j$, the method is called *explicit* and the system (2.26) can be solved by a sequential computation of the stages. Contrary, if the matrix $\mathbf{A}$ cannot be represented in a strictly lower triangular form, nonzero entries on the diagonal or above imply that it is not possible to directly solve for the stages. Runge-Kutta schemes where this is the case are termed *implicit*. These methods usually need to be solved numerically by employing root finding schemes within each integration step, e.g. using the Newton method. One of the most commonly used explicit methods is the fourth-order Runge-Kutta method, also called the *classical Runge-Kutta method*. This method can be represented by the following Butcher-tableau:

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array} \tag{2.29}$$

Note that the stage integration matrix in the Butcher-tableau (2.29) is strictly lower triangular, implying that the method is explicit and the stages can be computed se-

quentially starting from the first stages $s_{0,k}^{[i]} = x_k^{[i]}, k = 0, \ldots, n_x - 1$. For this particular method the sequence of computation within one integration step $t^{[i]} \rightarrow t^{[i+1]}$ can be summarized as follows:

- $s_{1,k}^{[i]}$: Forward Euler with a half step $h^{[i]}/2$ (Predictor)
- $s_{2,k}^{[i]}$: Backward Euler with a half step $h^{[i]}/2$ (Corrector)
- $s_{3,k}^{[i]}$: Midpoint rule with a full step $h^{[i]}$
- $x^{[i+1]}$: Quadrature using all stage derivatives

This sequential construction for the classical Runge-Kutta method is illustrated in Fig. 2.6.

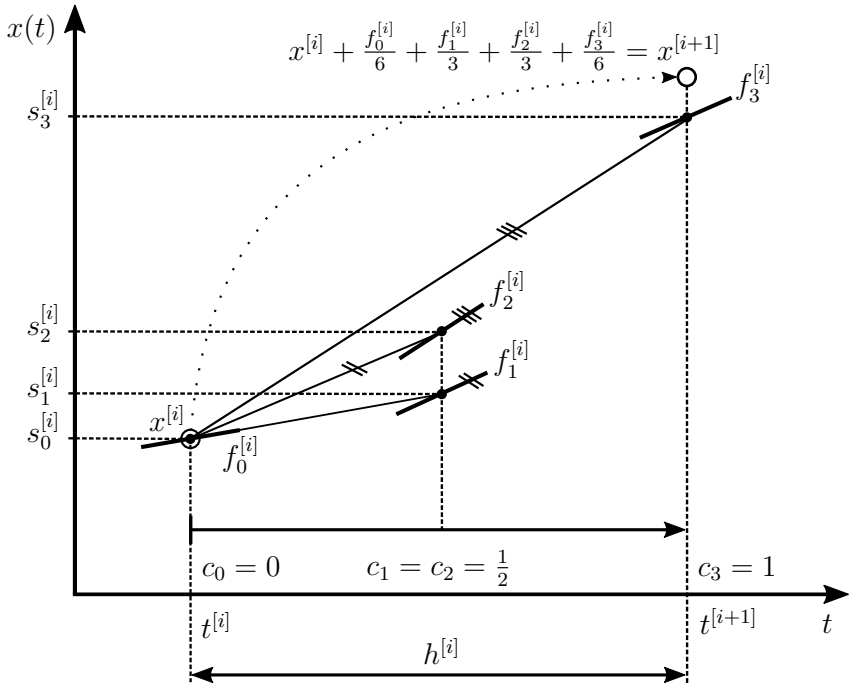


Figure 2.6: Illustration of the sequential stage computation for the classical, fourth-order, Runge-Kutta method in case of a single state. The derivatives corresponding to the stages are denoted using the short notation $f_j^{[i]} = h^{[i]} f(s_j^{[i]}, t^{[i]} + c_j h^{[i]}), j = 0, \ldots, 3$ for the scaled form of the right-handside.

2.4 Sensitivity Analysis for Initial Value Problems

Consider the parameter dependent form of an initial value problem (cf. Eq. (2.23))

$$\dot{\mathbf{x}}(t; \mathbf{p}) = \mathbf{f}\left(\mathbf{x}(t; \mathbf{p}), \mathbf{p}, t\right), \qquad \mathbf{x}(t_0) = \mathbf{x}_0(\mathbf{p}), \qquad t \in I^t, \tag{2.30}$$

depending on the parameter vector $\mathbf{p} \in \mathbb{R}^{n_p}$. The solution of the matrix-valued initial value problem

$$\dot{\mathbf{S}}(t) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\mathbf{S}(t) + \frac{\partial \mathbf{f}}{\partial \mathbf{p}}, \qquad \mathbf{S}(t_0) = \frac{\partial \mathbf{x}_0}{\partial \mathbf{p}}, \qquad t \in I^t, \tag{2.31}$$

for the sensitivities $\mathbf{S} \in \mathbb{R}^{n_x \times n_p}$ defined as

$$\mathbf{S}(t) := \frac{\partial \mathbf{x}(t;\mathbf{p})}{\partial \mathbf{p}}, \tag{2.32}$$

yields the derivatives of $\mathbf{x}(t;\mathbf{p})$ with respect to the parameters $\mathbf{p}$ [44]. The approach for obtaining the sensitivities $\mathbf{S}(t)$ from the solution of problem (2.31) is called *internal* differentiation method. Contrary, the *external* differentiation method relies on the numerical approximation of the derivatives through the application of a finite difference scheme, such as

$$\frac{\partial \mathbf{x}(t;\mathbf{p})}{\partial p_i} \approx \frac{\mathbf{x}\left(t;[p_0,\ldots,p_i+\Delta p_i,\ldots,p_{n_p-1}]^T\right) - \mathbf{x}(t;\mathbf{p})}{\Delta p_i},\ i = 0,\ldots,n_p-1. \tag{2.33}$$

Despite the appealing simplicity of the external differentiation method the internal differentiation method is typically preferred due the approximate nature of the finite difference scheme.

2.5 Segmented Collocation Methods

The solution for an initial value problem of the form (2.23) can also be found by a piecewise polynomial approximation on N subintervals

$$I^{t,[i]} := \left[t_S^{[i]}, t_S^{[i+1]}\right] \subset \mathbb{R}, t_S^{[i+1]} > t_S^{[i]},\ i = 0,\ldots,N-1, \tag{2.34}$$

termed segments in the following [45, 44]. These segments of length $h^{[i]} = t_S^{[i+1]} - t_S^{[i]}$ are defined based on the segment grid $\mathbb{G}_{N+1}^S$:

$$\mathbb{G}_{N+1}^S := \left\{t_S^{[i]} : i = 0\ldots,N, t_S^{[i+1]} > t_S^{[i]}, t_S^{[0]} = t_0, t_S^{[N]} = t_f\right\} \tag{2.35}$$

The piecewise polynomials $\mathbf{s}^{[i]}(t)$ of order $p^{[i]}$ which are used for the approximation of the solution within each segment need to satisfy the initial conditions

$$\mathbf{s}^{[i]}\left(t_S^{[i]}\right) = \mathbf{x}\left(t_S^{[i]}\right), \tag{2.36}$$

as well as the collocation conditions

$$\dot{\mathbf{s}}^{[i]}\left(t_S^{[i]}+c_j^{[i]}h^{[i]}\right)=\mathbf{f}\left(\mathbf{s}^{[i]}\left(t_S^{[i]}+c_j^{[i]}h^{[i]}\right),t_S^{[i]}+c_j^{[i]}h^{[i]}\right),\ j=0,...,p^{[i]}-1. \tag{2.37}$$

In Eq. (2.37) the collocation points $t_S^{[i]}+c_j^{[i]}h^{[i]}$ are defined by the collocation grid $\mathbb{G}_{p^{[i]}}^{C,[i]}$ containing the constants $c_j^{[i]}\in[0,1]$ normalized to the respective segment:

$$\mathbb{G}_{p^{[i]}}^{C,[i]}:=\left\{c_j^{[i]}:0\leq c_0^{[i]},\ldots,c_{p^{[i]}-1}^{[i]}\leq 1\right\},\ i=0,\ldots,N-1 \tag{2.38}$$

The conditions (2.36) and (2.37) thus impose that the solution matches the initial values at the beginning $t_S^{[i]}, i=0,\ldots,N-1$ of each segment and the derivatives at the respective collocation points. Furthermore, observe that these conditions require the piecewise polynomial to be continuous, but not necessarily continuously differentiable at the boundaries between segments. Only in the special case where the initial and final point in each segment are collocation points, i.e.

$$c_0^{[i]}=0,\qquad c_{p^{[i]}-1}^{[i]}=1,\qquad i=0,\ldots,N-1, \tag{2.39}$$

the polynomial approximation is continuously differentiable on the whole interval and the derivatives at the boundaries between segments satisfy for $N>1$:

$$\dot{\mathbf{s}}^{[i-1]}\left(t_S^{[i]}\right)=\dot{\mathbf{s}}^{[i]}\left(t_S^{[i]}\right),\ i=1,\ldots,N-1. \tag{2.40}$$

For the description in the following sections a single variable $x(t)$ is considered and the focus is on a single segment. As such, the index $\square^{[i]}$ is omitted and, additionally, the segment interval is treated in normalized form. This implies that the polynomial approximation $s(t)\approx x(t)$ is constructed for $t\in[0,1]$ and the normalized collocation points defined on the grid $\mathbb{G}_p^C$ for a polynomial approximation of order p are denoted with $c_j, j=0,\ldots,p-1$. This construction is general in the sense that the right-hand side can always be scaled by the segment length h, i.e. for each collocation point the following scaled form may be used:

$$f_j:=hf\left(s\left(c_j\right),hc_j\right),\ j=0,\ldots,p-1 \tag{2.41}$$

Moreover, the initial value of the segment is denoted by x_0 and the discussion is restricted to LG, LGRF, and LGL methods (cf. Sec. 2.2) which are particularly important for the application in direct optimal control methods (cf. chapter 6).

Differential Formulation

In the following a matrix formulation for the collocation conditions (2.36) and (2.37) is derived. For this purpose the polynomial $s(t)$ is formulated as a Lagrange interpolation polynomial on the grid $\mathbb{G}_{n_D}^D$

$$\mathbb{G}_{n_D}^D := \left\{ t_j^D : j = 0, \ldots, n_D - 1 \right\}, \tag{2.42}$$

with Lagrange basis functions

$$l_j^D(t) := \prod_{\substack{k=0 \\ k \neq j}}^{n_D} \frac{t - t_k^D}{t_j^D - t_k^D}, \qquad j = 0, \ldots, n_D - 1. \tag{2.43}$$

Using these Lagrange basis functions the interpolating polynomial for the corresponding values $s\left(t_j^D\right), j = 0, \ldots, n_D - 1$ can be written as follows:

$$s(t) = \sum_{j=0}^{n_D - 1} l_j^D(t)\, s\left(t_j^D\right) \tag{2.44}$$

It is important to mention that the initial condition (2.36) is taken into account by the additional inclusion of the initial point $t = 0$ of the segment in the interpolation grid. This is necessary for the set of LG and LGRF points which do not have a collocation point at $c_0 = 0$. The interpolation points in $\mathbb{G}_{n_D}^D$ are defined for this case as the union of $t = 0$ and the set of collocation points defined in $\mathbb{G}_p^C$ (cf. Eq. (2.38)):

$$\mathbb{G}_{n_D}^D = \{0\} \cup \mathbb{G}_p^C \tag{2.45}$$

For an illustration of the polynomial for LG and LGRF collocation methods in differential form please refer to Fig. 2.7. Note that for LG collocation continuity between segment boundaries is not automatically ensured by construction as merely the initial point is included in the definition of the interpolating polynomial (2.44). Therefore, it is necessary in this case to explicitly impose continuity of the profile between segments. See chapter 6 for a related discussion in the context of optimal control methods.

The p collocation conditions in (2.37) can be written in matrix form by differentiation of (2.44) and the evaluation at the collocation points defined by $\mathbb{G}_p^C$:

$$\begin{bmatrix} \dot{l}_0^D(c_0) & \ldots & \dot{l}_{n_D-1}^D(c_0) \\ \vdots & \ddots & \vdots \\ \dot{l}_0^D(c_{p-1}) & \ldots & \dot{l}_{n_D-1}^D(c_{p-1}) \end{bmatrix} \begin{bmatrix} s\left(t_0^D\right) \\ \vdots \\ s\left(t_{n_D-1}^D\right) \end{bmatrix} = \begin{bmatrix} f_0 \\ \vdots \\ f_{c_{p-1}} \end{bmatrix} \tag{2.46}$$

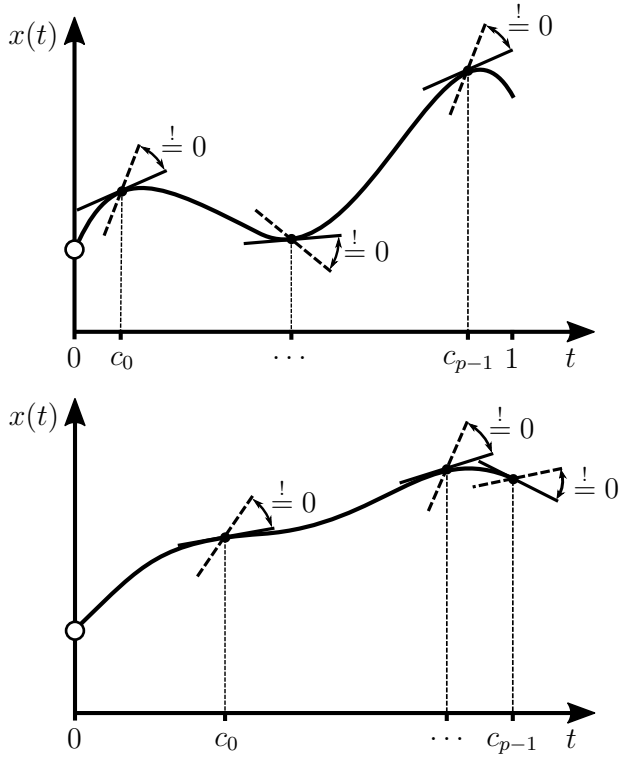


Figure 2.7: Illustration of the polynomial approximation for LG (upper) and LGRF (lower) collocation methods in differential form.

This form is particularly useful in order to establish the connection between the differential formulation and the integral formulation of collocation methods discussed in the following sections.

Integral Formulation

For the integral formulation the Lagrange interpolation polynomial is constructed for the derivative values $\dot{s}(c_j), j = 0, \ldots, p-1$ at the p collocation points. The Lagrange basis functions for the collocation grid $\mathbb{G}_p^C$ of cardinality p are defined as

$$l_j^I(t) := \prod_{\substack{k=0 \\ k \neq j}}^{p-1} \frac{t - c_k}{c_j - c_k}, \qquad j = 0, \ldots, p-1, \tag{2.47}$$

and the interpolating polynomial for the derivatives is obtained similar to (2.44) as follows:

$$\dot{s}(t) = \sum_{j=0}^{p-1} l_j^I(t)\, \dot{s}(c_j) \tag{2.48}$$

Integration of this interpolating polynomial with the initial value $s(0) = x_0$ yields:

$$\begin{bmatrix} s\left(c_0\right) \\ \vdots \\ s\left(c_{p-1}\right) \end{bmatrix} = \mathbf{1}x_0 + \begin{bmatrix} \int_0^{c_0} l_0^I\left(t\right) dt & \dots & \int_0^{c_0} l_{p-1}^I\left(t\right) dt \\ \vdots & \ddots & \vdots \\ \int_0^{c_{p-1}} l_0^I\left(t\right) dt & \dots & \int_0^{c_{p-1}} l_{p-1}^I\left(t\right) dt \end{bmatrix} \begin{bmatrix} \dot{s}\left(c_0\right) \\ \vdots \\ \dot{s}\left(c_{p-1}\right) \end{bmatrix} \tag{2.49}$$

Note that by including the initial value $s(0) = x_0$ in the integral the collocation condition (2.36) is automatically fulfilled. It can be shown that collocation methods of this form belong to the class of implicit Runge-Kutta methods with p stages [46]. This can be motivated by defining

$$a_{i,j} = \int_0^{c_i} l_j^I(t)\, dt, \qquad b_j = \int_0^1 l_j^I(t)\, dt, \qquad i, j = 0, \dots, p-1. \tag{2.50}$$

Furthermore, denoting the stage values as

$$s_j = s(c_j),\ j = 0, \dots, p-1, \tag{2.51}$$

and inserting the collocation conditions (2.37), i.e. $f_j = \dot{s}\left(c_j\right) j = 0, \dots, p-1$, in Eq. (2.49) the following form is obtained:

$$\begin{bmatrix} s_0 \\ \vdots \\ s_{p-1} \end{bmatrix} = \mathbf{1}x_0 + \begin{bmatrix} a_{1,1} & \dots & a_{1,p-1} \\ \vdots & \ddots & \vdots \\ a_{p-1,1} & \dots & a_{p-1,p-1} \end{bmatrix} \begin{bmatrix} f_0 \\ \vdots \\ f_{p-1} \end{bmatrix} \tag{2.52}$$

Observe that Eq. (2.52) represents exactly the form of Eq. (2.26) with the number of stages equal to the order p of the method.

Connection between the Differential and Integral Form

There exists an interesting connection between the differential and integral form of LG collocation methods, which has important implications for the relationship between both formulations. First, the following identity for the sum of the Lagrange basis function values at the non-collocated point for LG and LGRF methods are introduced [47]

$$l_0^D(c_j) = -\sum_{q=0}^{p-1} l_{q+1}^D(c_j),\ j = 0, \dots, p-1. \tag{2.53}$$

In order to illustrate the connection between the differential and integral form the non-collocated variable (node) is denoted with n, the collocated variables (stages) are denoted with $s_j, j = 0, \dots, p-1$, and the corresponding right-hand side function values

with $f_j, j = 0, \ldots, p-1$. With these quantities the collocation conditions for the differential formulation for LG and LGRF methods can be written as:

$$-\sum_{q=0}^{p-1} \dot{l}_{q+1}^{D}(c_j)n + \sum_{q=0}^{p-1} \dot{l}_{q+1}^{D}(c_j)s_q - f_j = 0. \tag{2.54}$$

The connection between the differential and integral form can be easily seen when the equality conditions of the differential and integral form are compared in matrix form

$$-\mathbf{D1}n + \mathbf{Ds} - \mathbf{f} = \mathbf{0}, \tag{2.55}$$

$$-\mathbf{1}n + \mathbf{s} - \mathbf{Af} = \mathbf{0}, \tag{2.56}$$

with the stage vector $\mathbf{s} = [s_0, \ldots, s_{p-1}]^T$, the vector of right-hand side function values $\mathbf{f} = [f_0, \ldots, f_{p-1}]^T$, as well as the differentiation matrix $\mathbf{D} \in \mathbb{R}^{p \times p}$ and integration matrix $\mathbf{A} \in \mathbb{R}^{p \times p}$ defined as

$$\mathbf{D} := \begin{bmatrix} \dot{l}_1^D(c_0) & \ldots & \dot{l}_p^D(c_0) \\ \vdots & \ddots & \vdots \\ \dot{l}_1^D(c_{p-1}) & \ldots & \dot{l}_p^D(c_{p-1}) \end{bmatrix}, \qquad \mathbf{A} := \begin{bmatrix} \int_0^{c_0} l_0^I(t)\,dt & \ldots & \int_0^{c_0} l_{p-1}^I(t)\,dt \\ \vdots & \ddots & \vdots \\ \int_0^{c_{p-1}} l_0^I(t)\,dt & \ldots & \int_0^{c_{p-1}} l_{p-1}^I(t)\,dt \end{bmatrix}. \tag{2.57}$$

It can be shown that the relationship [48, 49]

$$\mathbf{A} = \mathbf{D}^{-1}, \tag{2.58}$$

holds between the differentiation and integration matrix for Gauss- and Radau-type collocation methods which do not include both endpoints of the segment interval in the collocation grid. This implies that the differential and integral forms are equivalent which is not true for Lobatto collocation methods and the differentiation and integration matrices within a single segment are rank deficient. Observe that the first row of the integration matrix $\mathbf{A}$ based the set of LGL collocation points

$$\int_0^{c_0} l_j^I(t)\,dt = 0,\ j = 0, \ldots, p-1, \tag{2.59}$$

contains the integrated Lagrange basis functions with $c_0 = 0$. As the first row in $\mathbf{A}$ is zero regardless of the order p the integration matrix cannot have full rank. Furthermore, it is interesting to note that for the last row the integrated Lagrange basis functions with $c_{p-1} = 1$ yield

$$\int_0^{c_{p-1}} l_j^I(t)\,dt = b_j,\ j = 0, \ldots, p-1. \tag{2.60}$$

This implies that the last row contains the weights of the quadrature rule with p points associated with the LGL method. Both properties, (2.59) and (2.60), are of interest when deriving important special cases, which are particularly useful in the context of direct optimal control methods, in the following section.

Discussion of Special Cases for Legendre-Gauss Collocation Methods

In this section, important special cases of collocation methods are discussed which are the basis for popular transcription schemes in optimal control applications. Note that the equations for the differential and integral formulations implicitly define the stage values at collocation points. In most cases these equations are solved numerically. For example, if a collocation method is used inside an optimization problem, which is the case in the context of this thesis, the variables at the stages may be introduced as optimization variables and the collocation conditions can be imposed as equality constraints. In the following important special cases are derived for integral collocation methods based on the LG, LGR, LGRF, and LGL quadrature abscissas, normalized to the interval $[0, 1]$. Note that the segment index $\square^{[i]}$ is re-introduced in order to distinguish between variables of the current segment ($\square^{[i]}$) and the following segment ($\square^{[i+1]}$).

Order $\mathbf{p = 1}$: The approximating polynomial $s^{[i]}(t)$ for $p = 1$ is easily obtained from the semi-definite integral $\int_0^t l_0^I(s)\,ds = t$ with $l_0^I(t) = 1$ [45]:

$$s^{[i]}(t) = x_0^{[i]} + t f_0^{[i]}, \tag{2.61}$$

$$x_0^{[i+1]} = x_0^{[i]} + f_0^{[i]}. \tag{2.62}$$

From this representation the Forward and Backward Euler method, as well as the implicit midpoint rule can be directly derived by inserting the respective collocation point (see Tab. 2.2).

Table 2.2: Resulting methods for $p = 1$ in case of LG, LGR, LGRF points

Type	c_0	Butcher Tableaus	Methods
LGR	0	$\begin{array}{c\|c} 0 & 0 \\ \hline & 1 \end{array}$	Forward Euler
LG	$\frac{1}{2}$	$\begin{array}{c\|c} \frac{1}{2} & \frac{1}{2} \\ \hline & 1 \end{array}$	Midpoint Rule
LGRF	1	$\begin{array}{c\|c} 1 & 1 \\ \hline & 1 \end{array}$	Backward Euler

Order $\mathbf{p = 2}$: For $p = 2$ the Lagrange basis functions are

$$l_0^I(t) = \frac{t - c_1}{c_0 - c_1}, \qquad l_1^I(t) = \frac{t - c_0}{c_1 - c_0}, \tag{2.63}$$

and the corresponding semi-definite integrals yield

$$\begin{aligned} \int_0^t l_0^I(s)\, ds &= \frac{-t\ (2\,c_1 - t)}{2\ (c_0 - c_1)}, \\ \int_0^t l_1^I(s)\, ds &= \frac{t\ (2\,c_0 - t)}{2\ (c_0 - c_1)}. \end{aligned} \tag{2.64}$$

Thus, the following expression for the polynomial approximation $s^{[i]}(t; c_0, c_1)$, parameterized by the collocation constants c_0 and c_1 is obtained:

$$s^{[i]}(t; c_0, c_1) = x_0^{[i]} + \frac{-t\ (2\,c_1 - t)}{2\ (c_0 - c_1)} f_0^{[i]} + \frac{t\ (2\,c_0 - t)}{2\ (c_0 - c_1)} f_1^{[i]} \tag{2.65}$$

Inserting the LGL points for $p = 2$, i.e. $c_0 = 0$ and $c_1 = 1$, yields the Butcher tableau

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \tag{2.66}$$

This Butcher tableau represents the second-order LobattoIIIA method, or Trapezoidal method. [45] Note that the first row of Eq. (2.66) contains only zero entries (in accordance with Eq. (2.59)) and simply states that the first stage coincides with the initial value of the segment, i.e.

$$s^{[i]}(0; 0, 1) = x_0^{[i]}. \tag{2.67}$$

Furthermore the integrated stage $s(1; 0, 0)$ at the end of the interval is equal to the value obtained from the quadrature rule $x_0^{[i+1]}$ (in accordance with Eq. (2.60)). Thus, the two conditions

$$s^{[i]}(1; 0, 1) = x_0^{[i]} + \frac{1}{2}\left(f_0^{[i]} + f_1^{[i]}\right), \tag{2.68}$$

$$x_0^{[i+1]} = x_0^{[i]} + \frac{1}{2}\left(f_0^{[i]} + f_1^{[i]}\right), \tag{2.69}$$

together with $s^{[i]}(1; 0, 1) = x_0^{[i+1]}$ are equivalent. Therefore, only one condition is required to solve the implicit equations for $x_0^{[i+1]}$, which is the common representation for the Trapezoidal collocation method [50].

Order p = 3: Another important case which is often used in practice can be derived for $p = 3$ using the basis functions

$$l_0^I(t) = \frac{(c_1 - t)(c_2 - t)}{(c_0 - c_1)(c_0 - c_2)}, \quad l_1^I(t) = -\frac{(c_0 - t)(c_2 - t)}{(c_0 - c_1)(c_1 - c_2)}, \quad l_2^I(t) = \frac{(c_0 - t)(c_1 - t)}{(c_0 - c_2)(c_1 - c_2)}. \tag{2.70}$$

The semi-definite integrals for the parametrized polynomial approximation yield in this case:

$$\begin{aligned} s^{[i]}(t; c_0, c_1, c_2) = x_0^{[i]} &+ \frac{t\,(6\,c_1\,c_2 - 3\,c_1\,t - 3\,c_2\,t + 2\,t^2)}{6\,(c_0 - c_1)\,(c_0 - c_2)} f_0^{[i]} \\ &- \frac{t\,(6\,c_0\,c_2 - 3\,c_0\,t - 3\,c_2\,t + 2\,t^2)}{6\,(c_0 - c_1)\,(c_1 - c_2)} f_1^{[i]} \\ &+ \frac{t\,(6\,c_0\,c_1 - 3\,c_0\,t - 3\,c_1\,t + 2\,t^2)}{6\,(c_0 - c_2)\,(c_1 - c_2)} f_2^{[i]}. \end{aligned} \tag{2.71}$$

For the Lobatto points $c_0 = 0$, $c_1 = \frac{1}{2}$, and $c_2 = 1$ the following Butcher-tableau is obtained:

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{5}{24} & \frac{1}{3} & \frac{-1}{24} \\ 1 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ \hline & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{array} \tag{2.72}$$

By the same argument as for the Trapezoidal method the first stage $s^{[i]}\left(0; 0, \frac{1}{2}, 1\right)$ coincides with the initial value $x_0^{[i]}$ and the last row and the quadrature equation are redundant, meaning that the two relevant equations are:

$$s^{[i]}\left(\frac{1}{2}; 0, \frac{1}{2}, 1\right) = x_0^{[i]} + \frac{1}{24}\left(5 f_0^{[i]} + 8 f_1^{[i]} - f_2^{[i]}\right), \tag{2.73}$$

$$x_0^{[i+1]} = x_0^{[i]} + \frac{1}{24}\left(4 f_0^{[i]} + 16 f_1^{[i]} + 4 f_2^{[i]}\right) \tag{2.74}$$

Note that using these two equations it is possible to explicitly solve for the stage $s^{[i]}(\frac{1}{2}; 0, \frac{1}{2}, 1)$ by subtracting two times the first equation (2.73) from the second equation (2.74). This yields the following conditions for the Hermite-Simpson collocation method in compressed form [50]:

$$s^{[i]}\left(\frac{1}{2}; 0, \frac{1}{2}, 1\right) = \frac{\left(x_0^{[i]} + x_0^{[i+1]}\right)}{2} + \frac{f_0^{[i]} - f_2^{[i]}}{8}, \tag{2.75}$$

$$x_0^{[i+1]} = x_0^{[i]} + \frac{1}{24}\left(4 f_0^{[i]} + 16 f_1^{[i]} + 4 f_2^{[i]}\right) \tag{2.76}$$

2.6 B-spline Parametrization

Function parametrizations based on B-splines [51, 50] are particularly useful in the context of direct optimal control methods for the discretization of control functions. A B-spline parametrization of degree p is defined based on a break-point grid $\mathbb{G}_M^B$ of cardinality M

$$\mathbb{G}_M^B := \left\{ t_i^B : i = 0, \ldots, M-1, t_{i+1}^B \geq t_i^B \right\}. \tag{2.77}$$

For this break-point grid an auxiliary *knot* grid $\mathbb{G}_m^K$ is constructed. In the following this knot grid is assumed to be a *clamped* grid of cardinality $m = M + 2p$ of the form:

$$\mathbb{G}_m^K := \left\{ t_i^K : i = 0, \ldots, m-1 \right\} = \{ \underbrace{t_0^B, \ldots, t_0^B}_{p+1}, \underbrace{t_1^B, \ldots, t_{M-2}^B}_{M-2}, \underbrace{t_{M-1}^B, \ldots, t_{M-1}^B}_{p+1} \} \tag{2.78}$$

It is important to mention that the B-spline knot grid points do not need to be unique. In particular, for clamped knot grids the points in the beginning and the end have multiplicity $p+1$. The B-spline basis functions $B_{p,i}(t)$ can be defined recursively by the relation

$$B_{p,i}(t) = \frac{t - t_i^K}{t_{i+p}^K - t_i^K} B_{p-1,i}(t) + \frac{t_{i+p+1}^K - t}{t_{i+p+1}^K - t_{i+1}^K} B_{p-1,i+1}(t), \tag{2.79}$$

with

$$B_{0,i}(t) := \begin{cases} 1, & \text{if } t \in \left[t_i^K, t_{i+1}^K \right), \\ 0, & \text{otherwise.} \end{cases} \tag{2.80}$$

and the definition of the quotient $0/0 = 0$. The $n_w = M + p - 1$ non-zero basis functions have *local support*. This means that the B-spline basis functions satisfy for $p > 0$:

$$B_{p,i}(t) \begin{cases} > 0, & \text{if } t \in \left(t_i^K, t_{i+p+1}^K \right), \\ = 0, & \text{otherwise.} \end{cases} \tag{2.81}$$

For an illustration for these B-spline basis functions up to degree $p = 3$ please refer to Fig. 2.8.

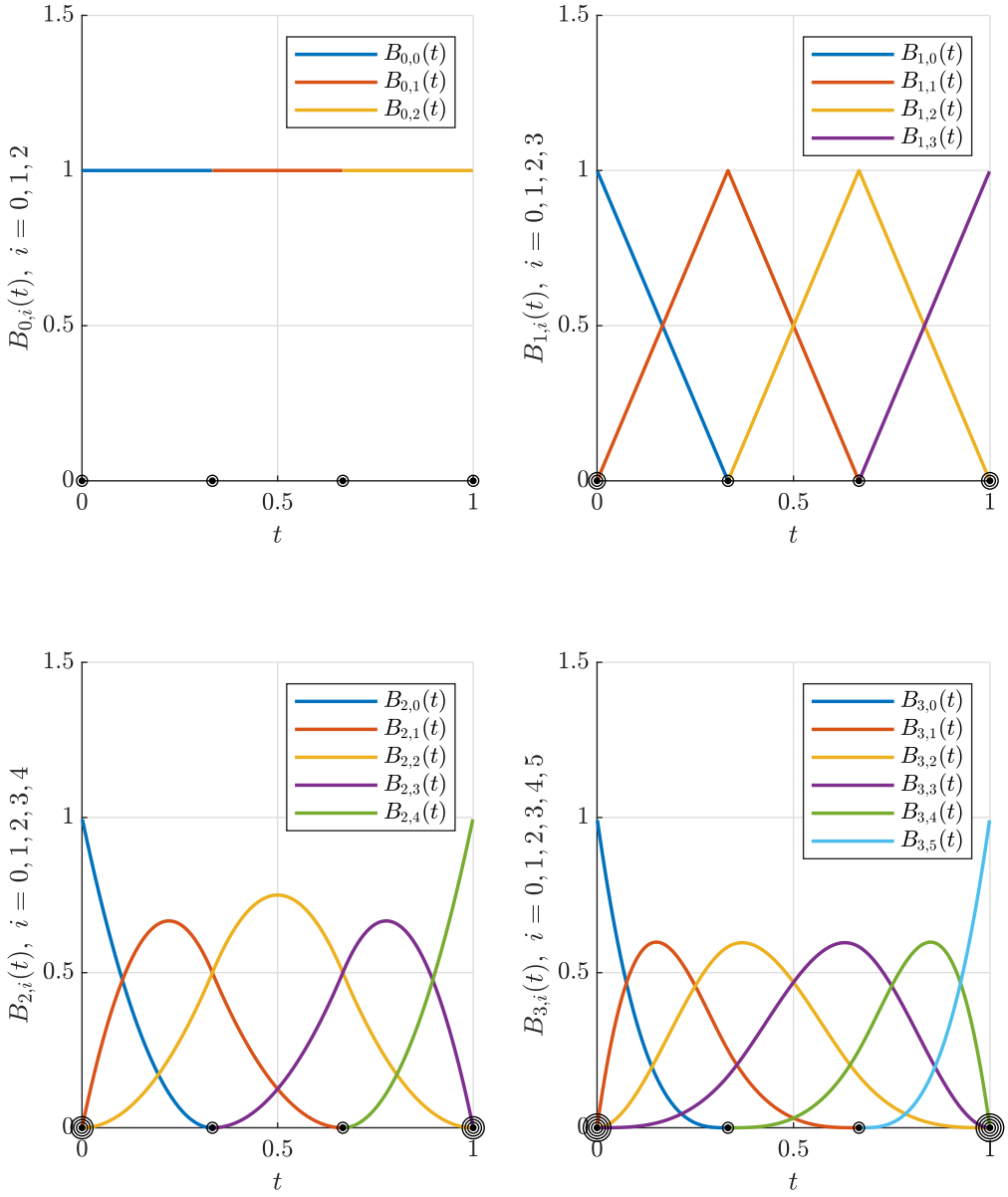


Figure 2.8: B-spline basis functions (non-zero values) of degree $p = 0$ (top left), $p = 1$ (top right), $p = 2$ (bottom left), and $p = 3$ (bottom right) for a break point grid $\mathbb{G}_4^B = \{0, 1/3, 2/3, 1\}$ (black dots). The knot locations, including multiplicity at the first and last grid point, are illustrated with black circles.

The function $f(t)$ based on the B-spline parametrization with a parameter vector $\mathbf{w} = [w_0, \ldots, w_{n_w-1}]^T$ is defined as

$$f(t) = \sum_{i=0}^{n_w-1} B_{p,i}(t) w_i. \tag{2.82}$$

For computational implementations it is usually not efficient to evaluate each B-spline basis function $B_{p,i}(t)$ based on Eq. (2.79) as common terms can be saved and re-used in the recursion. This fact can be easily seen when visualizing the B-spline basis functions needed to construct the recursion from Eq. (2.79) in triangular table form:

$$\begin{array}{ccccccc}
B_{0,0} & & & & & & \\
& \searrow & & & & & \\
& & B_{1,0} & & & & \\
& \nearrow & & \searrow & & & \\
B_{0,1} & & & & B_{2,0} & & \\
& \searrow & & \nearrow & & \ddots & \\
& & B_{1,1} & & & & \\
& \nearrow & & \searrow & & & \\
B_{0,2} & & & & B_{2,1} & & \\
& \searrow & & \nearrow & \vdots & \ddots & \\
& & B_{1,2} & & & & \\
& \nearrow & \vdots & \ddots & & & \\
B_{0,3} & & & & & & \\
\vdots & \ddots & & & & &
\end{array} \tag{2.83}$$

For example, consider the construction of the basis functions $B_{2,0}(t)$ and $B_{2,1}(t)$. For the basis function $B_{2,0}(t)$ the functions $B_{1,0}(t)$ and $B_{1,1}(t)$ need to be constructed. Similarly, the construction of $B_{2,1}(t)$ is based on $B_{1,1}(t)$ and $B_{1,2}(t)$. Thus, the function $B_{1,1}(t)$ is shared among $B_{2,0}(t)$ and $B_{2,1}(t)$ which can be exploited in the implementation [51]. One of the major advantages regarding the use of B-spline parametrizations for the discretization of control functions in optimal control applications is that the interpolation Jacobians are typically very sparse. To illustrate this point consider a query grid $\mathbb{G}_{n_q}^Q$ of cardinality n_q

$$\mathbb{G}_{n_q}^Q := \{t_j : j = 0, \ldots, n_q - 1\}, \tag{2.84}$$

and the corresponding value vector $\mathbf{v} = \left[f\left(t_0\right), \ldots, f\left(t_{n_q-1}\right)\right]^T$ with

$$f\left(t_j\right) = \sum_{i=0}^{n_w-1} B_{p,i}\left(t_j\right) w_i,\ j = 0, \ldots, n_q - 1. \tag{2.85}$$

First, note that the Jacobian

$$\frac{\partial \mathbf{v}}{\partial \mathbf{w}} = \begin{bmatrix} B_{p,0}(t_0) & \ldots & B_{p,n_w-1}(t_0) \\ \vdots & \ddots & \vdots \\ B_{p,0}(t_{n_q-1}) & \ldots & B_{p,n_w-1}(t_{n_q-1}) \end{bmatrix}, \tag{2.86}$$

of $\mathbf{v}$ with respect to the parameter vector $\mathbf{w}$ merely contains the B-spline basis functions evaluated at the query grid points. As such, this Jacobian remains constant if the breakpoint and query grids do not change. Moreover, due to the local support property of the B-spline functions (cf. Eq. (2.81)) most of the entries are usually zero, i.e. the matrix is sparse. This property plays an important role in numerical optimal control methods as it reduces coupling within the problem [52].

Chapter 3

Unconstrained Optimization

The numerical treatment of optimal control problems in the context of direct methods (cf. chapter 6) is based on the solution of constrained parameter optimization problems. In order to streamline the discussion for the constrained case (cf. chapter 4), first the unconstrained case is presented in this chapter. Many of the concepts introduced in this chapter, such as the statement of the necessary and sufficient conditions of optimality and the application of iterative numerical methods, can then be extended to the constrained case later on.

3.1 Problem Statement

For unconstrained optimization problems [4, 50] the goal is to find an optimization variable vector $\mathbf{z} \in \mathbb{R}^{n_z}$ which minimizes the cost function $j : \mathbb{R}^{n_z} \to \mathbb{R}$:

$$\underset{\mathbf{z} \in \mathbb{R}^{n_z}}{\text{minimize}} \quad j(\mathbf{z}) \tag{3.1}$$

The solution of this problem is termed *globally optimal* if the optimal point $\hat{\mathbf{z}}$ satisfies

$$j(\hat{\mathbf{z}}) \leq j(\mathbf{z}), \qquad \mathbf{z} \in \mathbb{R}^{n_z}, \tag{3.2}$$

and *locally optimal* if

$$j(\hat{\mathbf{z}}) \leq j(\mathbf{z}), \qquad \mathbf{z} \in \mathcal{N} \tag{3.3}$$

holds in a neighborhood $\mathcal{N}$ of $\hat{\mathbf{z}}$. Furthermore, the solution can be characterized as a *strict* global or local optimum if strict inequalities hold for conditions (3.2) and (3.3) for all $\mathbf{z} \neq \hat{\mathbf{z}}$.

3.2 Necessary and Sufficient Conditions

The first-order necessary conditions for a minimizer $\hat{\mathbf{z}}$ of $j(\mathbf{z})$ require the cost function to be continuously differentiable in an open neighborhood $\mathcal{N}$ of $\hat{\mathbf{z}}$. These conditions state that the gradient of the cost function has to satisfy:

$$\frac{\partial j}{\partial \mathbf{z}}(\hat{\mathbf{z}}) = \mathbf{0} \tag{3.4}$$

A second-order sufficient condition can be formulated under the assumption of a twice continuously differentiable cost function. This sufficiency condition states that in addition to the first-order necessary conditions (3.4) the Hessian $\frac{\partial^2 j}{\partial \mathbf{z}^2}(\hat{\mathbf{z}})$ is required to be positive definite:

$$\mathbf{d}^T \frac{\partial^2 j}{\partial \mathbf{z}^2}(\hat{\mathbf{z}})\mathbf{d} > 0,\ \forall\, \mathbf{d} : \hat{\mathbf{z}} + \mathbf{d} \in \mathcal{N}, \mathbf{d} \neq \mathbf{0} \tag{3.5}$$

Furthermore, if the cost function j is convex, the stationary point of the cost function (cf. Eq. 3.4) is a global minimizer. There exists a number of algorithms to solve the unconstrained optimization problem in an iterative process. This means that the algorithms generate a sequence of iterates $\{\mathbf{z}_k\}$ which converges to a local minimizer $\hat{\mathbf{z}}$. The numerical methods differ in how a step from one iterate $\mathbf{z}_k$ to the next iterate $\mathbf{z}_{k+1}$ is taken. Two important approaches in this context are Trust-Region and Line-Search methods. The major difference between these two approaches is that for Trust-Region algorithms the step to the next iterate is computed by solving a constrained model problem which approximates the function $j(\mathbf{z})$ locally around the current iterate. The length and direction of the step are chosen simultaneously for this case by solving the model optimization problem within the so-called Trust-Region. For Line-Search algorithms the step to the next iterate is determined in a sequential manner. First, a search direction is found based on the solution to an unconstrained optimization problem. Second, a step-size in this direction is computed. Both Trust-Region and Line-Search methods are discussed in the following.

3.3 Trust-Region Methods

For Trust-Region methods the step to the next iterate is determined from the solution of a simplified minimization problem. This simplified problem uses a model $\tilde{j}_k(\mathbf{z})$, which locally approximates the cost function $j(\mathbf{z})$ around the current iterate $\mathbf{z}_k$. The region in which this model is "trusted" to provide a good approximation and produce sufficient improvements to the actual minimization problem is called the Trust-Region.

The shape of the Trust-Region with radius $\Delta_k > 0$ is often defined as

$$\|\mathbf{M}\mathbf{s}_k\|_p \leq \Delta_k, \tag{3.6}$$

with the matrix $\mathbf{M} \in \mathbb{R}^{n_z \times n_z}$, the vector $\mathbf{s}_k = \mathbf{z} - \mathbf{z}_k$ and using a suitable vector p-norm $\|\cdot\|_p$. Typically, a quadratic model is used for the local approximation of the cost function around $\mathbf{z}_k$

$$j(\mathbf{z}) \approx \tilde{j}_k(\mathbf{z}) = a_k + \mathbf{c}_k^T \mathbf{s}_k + \frac{1}{2}\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k, \tag{3.7}$$

If the model in Eq. (3.7) is based on a Taylor expansion around the current iterate $\mathbf{z}_k$, the constant a_k represents the cost function value $j(\mathbf{z}_k)$ at the current point, $\mathbf{c}_k$ is the gradient $\frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k)$, and $\mathbf{B}_k$ the Hessian $\frac{\partial^2 j}{\partial \mathbf{z}^2}(\mathbf{z}_k)$. The approximate minimization problem with the model function (3.7) and the Trust-Region constraint (3.6) is thus defined as:

$$\begin{aligned} \underset{\mathbf{s}_k \in \mathbb{R}^{n_z}}{\text{minimize}} \quad & \mathbf{c}_k^T \mathbf{s}_k + \frac{1}{2}\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k, \\ \text{subject to} \quad & \|\mathbf{M}\mathbf{s}_k\|_p \leq \Delta_k. \end{aligned} \tag{3.8}$$

For a diagonal matrix $\mathbf{M} = diag(m_1, \ldots, m_n)$ with positive entries on the diagonal and $p = 2$, Eq. (3.6) defines an ellipsoidal and for $\mathbf{M} = \mathbf{I}$ with $p = 2$ a spherical Trust-Region. An elliptical region can be helpful if the optimization variables exhibit different rates of change in different directions. In this case, the entries on the diagonal of $\mathbf{M}$ represent scaling factors for the minimization problem. This can be easily seen by applying the variable transformation $\tilde{\mathbf{s}}_k = \mathbf{M}\mathbf{s}_\mathbf{k}$ to (3.8), and defining $\tilde{\mathbf{B}}_k = \mathbf{M}^{-1}\mathbf{B}_k\mathbf{M}^{-1}$, as well as $\tilde{\mathbf{c}}_k = \mathbf{M}^{-1}\mathbf{c}_k$ which yields:

$$\begin{aligned} \underset{\tilde{\mathbf{s}} \in \mathbb{R}^{n_z}}{\text{minimize}} \quad & \tilde{\mathbf{c}}_k^T \tilde{\mathbf{s}}_k + \frac{1}{2}\tilde{\mathbf{s}}_k^T \tilde{\mathbf{B}}_k \tilde{\mathbf{s}}_k, \\ \text{subject to} \quad & \|\tilde{\mathbf{s}}_k\|_p \leq \Delta_k. \end{aligned} \tag{3.9}$$

From the solution $\hat{\mathbf{s}}_k$ of the model problem (3.8) the next iterate $\mathbf{z}_{k+1}$ is readily obtained:

$$\mathbf{z}_{k+1} = \mathbf{z}_k + \hat{\mathbf{s}}_k \tag{3.10}$$

In order to measure the quality of the approximation and to adapt the Trust-Region radius Δ_k, the quotient

$$\tilde{\rho}_k := \frac{j(\mathbf{z}_k) - j(\mathbf{z}_k + \hat{\mathbf{s}}_k)}{\tilde{j}(\mathbf{z}_k) - \tilde{j}(\mathbf{z}_k + \hat{\mathbf{s}}_k)}, \tag{3.11}$$

can be used. For example, if $\tilde{\rho}_k$ has a value close to one the approximation within the Trust-Region seems to be trustworthy and the step can be accepted. In addition, the Trust-Region radius Δ_k may be increased for the next iteration. Other values for $\tilde{\rho}_k$ exist in which the trust region should be reduced or kept unchanged (cf. Ref. [4]).

3.4 Line-Search Methods

Recall that the solution of the approximate minimization problem for Trust-Region methods simultaneously yields a direction and magnitude of the step from the current iterate $\mathbf{z}_k$ to the next iterate $\mathbf{z}_{k+1}$. However, for line search methods the direction and magnitude are computed sequentially. In the following two possibilities to obtain the search direction, namely Newton and Quasi-Newton methods, as well as two conditions to determine the step size (Armijo rule and Wolfe conditions) which are commonly used for Line-Search methods are discussed.

3.4.1 Search Direction Computation

Newton Method

In case the model function is chosen to be the truncated Taylor expansion of $j(\mathbf{z})$ around the current point $\mathbf{z}_k$

$$\tilde{j}\left(\mathbf{s}_k\right) := j(\mathbf{z}_k) + \frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k)\mathbf{s}_k + \frac{1}{2}\mathbf{s}_k^T \frac{\partial^2 j}{\partial \mathbf{z}^2}(\mathbf{z}_k)\mathbf{s}_k, \tag{3.12}$$

with $\mathbf{s}_k = \mathbf{z} - \mathbf{z}_k$, the solution $\hat{\mathbf{s}}_k$ to the unconstrained minimization problem

$$\underset{\mathbf{s}_k \in \mathbb{R}^{n_z}}{\text{minimize}} \quad \frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k)\mathbf{s}_k + \frac{1}{2}\mathbf{s}_k^T \frac{\partial^2 j}{\partial \mathbf{z}^2}(\mathbf{z}_k)\mathbf{s}_k, \tag{3.13}$$

can be obtained simply by setting the gradient to zero, i.e. :

$$\frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k) + \frac{\partial^2 j}{\partial \mathbf{z}^2}(\mathbf{z}_k)\mathbf{s}_k = \mathbf{0} \tag{3.14}$$

This yields the Newton-direction:

$$\hat{\mathbf{s}}_k = -\left(\frac{\partial^2 j}{\partial \mathbf{z}^2}(\mathbf{z}_k)\right)^{-1} \frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k) \tag{3.15}$$

It can be shown, that close to a minimizer $\hat{\mathbf{z}}$ where the Hessian is positive definite, the Newton-direction leads to a fast, quadratic convergence rate. Note that the inverse of the Hessian needs to exist in order to compute the search direction $\hat{\mathbf{s}}_k$. This may not be the case further away from the minimizer $\hat{\mathbf{z}}$ where the Hessian can be indefinite. This poses a major problem, as in this case the Hessian needs to be modified such that all eigenvalues are strictly positive. Moreover, the evaluation of the Hessian can be computationally quite expensive. One possible remedy for the mentioned difficulties, which partially preserves the good local convergence properties of the Newton-direction, is the application of the so-called Quasi-Newton method.

Quasi-Newton Method

For the Quasi-Newton method the Hessian in the model (3.12) is replaced by an approximation $\mathbf{B}_k$

$$\tilde{j}\left(\mathbf{s}_k\right) = j(\mathbf{z}_k) + \frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k)\mathbf{s}_k + \frac{1}{2}\mathbf{s}_k^T\mathbf{B}_k\mathbf{s}_k, \tag{3.16}$$

and the unconstrained minimization problem (3.13) becomes:

$$\underset{\mathbf{s}_k \,\in\, \mathbb{R}^{n_z}}{\text{minimize}} \quad \frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k)\mathbf{s}_k + \frac{1}{2}\mathbf{s}_k^T\mathbf{B}_k\mathbf{s}_k \tag{3.17}$$

The solution of this problem with $\mathbf{H}_k = \mathbf{B}_k^{-1}$ yields the direction

$$\hat{\mathbf{s}}_k = -\mathbf{H}_k\frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k), \tag{3.18}$$

similar to Eq. (3.15). One possibility for the construction of a suitable sequence of matrices $\{\mathbf{B}_k\}$ is the Broyden-Fletcher-Goldfarb-Shanno (BFGS) update rule [4]

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \frac{\mathbf{B}_k\mathbf{s}\mathbf{s}_k^T\mathbf{B}_k}{\mathbf{s}_k^T\mathbf{B}_k\mathbf{s}_k} + \rho_k\mathbf{y}_k\mathbf{y}_k^T, \qquad \rho_k = \left(\mathbf{y}_k^T\mathbf{s}_k\right)^{-1}, \tag{3.19}$$

paired with a suitable Line-Search method, such as the Wolfe-rules (cf. Sec. 3.4.2). The idea behind this method is to apply a rank-2 update based on curvature information from the change in the gradient $\mathbf{y}_k = \frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_{k+1}) - \frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k)$ between two iterations. The sequence of matrices $\{\mathbf{B}_k\}$ generated by this update rule maintains a symmetric and positive definite matrix $\mathbf{B}_k$. In addition to Eq. (3.19), the following update formula which iterates on the inverse $\mathbf{H}_k = \mathbf{B}_k^{-1}$ can be applied:

$$\mathbf{H}_{k+1} = \left(\mathbf{I} - \rho_k\mathbf{s}_k\mathbf{y}_k^T\right)\mathbf{H}_k\left(\mathbf{I} - \rho_k\mathbf{y}_k\mathbf{s}_k^T\right) + \rho_k\mathbf{s}_k\mathbf{s}_k^T \tag{3.20}$$

This form is typically preferred as the matrix $\mathbf{H}_k$ in Eq. (3.18) is directly available to compute the search direction without the additional solution of a linear system.

Steepest Descent

Further simplifying the unconstrained minimization problem (3.16) and setting the Hessian approximation to the identity $\mathbf{B}_k = \mathbf{I}$ yields the minimization problem

$$\underset{\mathbf{s}_k \,\in\, \mathbb{R}^{n_z}}{\text{minimize}} \quad \frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k)\mathbf{s}_k + \frac{1}{2}\mathbf{s}_k^T\mathbf{s}_k, \tag{3.21}$$

with the closed solution

$$\hat{\mathbf{s}}_k^T = -\frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k). \tag{3.22}$$

This is the direction of steepest descent at the current iterate $\mathbf{z}_k$ as the gradient points in the direction of steepest ascent.

3.4.2 Step Size Selection

The step size α_k is used as a scaling factor for the search direction $\hat{\mathbf{s}}_k$ to obtain the next iterate $\mathbf{z}_{k+1}$:

$$\mathbf{z}_{k+1} = \mathbf{z}_k + \alpha_k \hat{\mathbf{s}}_k \tag{3.23}$$

In this context, it is reasonable to impose a requirement that the step size improves the cost function in the direction of $\hat{\mathbf{s}}_k$. This requirement can be expressed by the *sufficient decrease* inequality condition

$$j(\mathbf{z}_k + \alpha_k \hat{\mathbf{s}}_k) \leq j(\mathbf{z}_k) + \beta \frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k) \alpha_k \hat{\mathbf{s}}_k, \tag{3.24}$$

with the parameter $0 < \beta < 1$. This condition, sometimes termed *Armijo rule*, is illustrated in Fig. 3.1. Despite the fact, that the condition (3.24) ensures non-increasing values of the cost function it may result in very small improvements from one iteration to the next. Note that the step size $\alpha_k = 0$ trivially satisfies the inequality. A stronger condition for selecting the step size α_k can be formulated by imposing a second condition, termed *curvature condition*, of the form

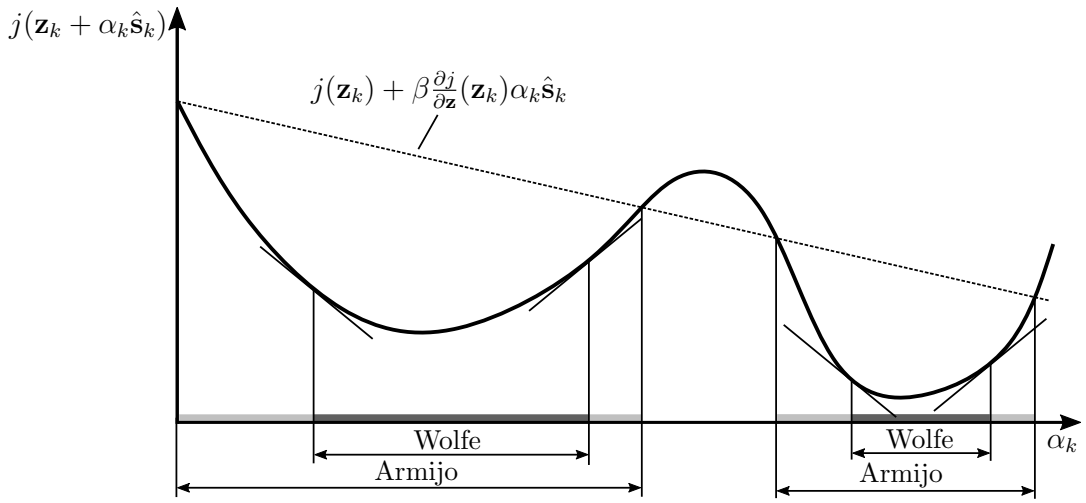


Figure 3.1: Acceptable step sizes using the Armijo rule and the strong Wolfe conditions.

$$\frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k + \alpha_k \mathbf{s}_k) \hat{\mathbf{s}}_k \geq \gamma \frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k) \hat{\mathbf{s}}_k \tag{3.25}$$

where the constant γ satisfies $\beta < \gamma < 1$. Essentially, this condition imposes a requirement for a decrease in the cost function gradient. The sufficient decrease inequality

(3.24) together with the curvature condition (3.25) are usually referred to as *Wolfe* conditions. Furthermore, the *strong Wolfe* conditions additionally impose

$$\left|\frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k + \alpha_k \mathbf{s}_k)\hat{\mathbf{s}}_k\right| \leq \left|\gamma \frac{\partial j}{\partial \mathbf{z}}(\mathbf{z}_k)\mathbf{s}_k\right|, \tag{3.26}$$

instead of Eq. (3.25). It can be shown that for a continuously differentiable cost function which is bounded from below in the direction $\hat{\mathbf{s}}_k$ there exists a step length α_k which satisfies the strong Wolfe conditions for $0 < \beta < \gamma < 1$. [4]

Chapter 4

Constrained Optimization

This chapter extends the theoretical background regarding parameter optimization problems from the unconstrained case, presented in chapter 3, to the constrained case. First, the general, nonlinear case with equality and inequality constraints is revisited. Afterwards, numerical methods for this nonlinear case as well as other relevant special cases, namely linear and quadratic programming, are discussed. These methods are employed in this thesis to solve the discretized form of the clearance optimal control problem. Finally, methods related to the determination of post-optimal sensitivities for parametric optimization problems are presented which can be used to investigate the sensitivity of optimal solutions, that are worst-case solutions in the context of this thesis, with respect to parameters of interest.

4.1 Problem Statement

The goal in constrained optimization [4, 50, 53] is to find the optimization variable vector $\mathbf{z} \in \mathbb{R}^{n_z}$, which minimizes the cost function $j : \mathbb{R}^{n_z} \to \mathbb{R}$ subject to the equality constraints $\mathbf{h} : \mathbb{R}^{n_z} \to \mathbb{R}^{n_h}$ and the inequality constraints $\mathbf{g} : \mathbb{R}^{n_z} \to \mathbb{R}^{n_g}$:

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{R}^{n_z}}{\text{minimize}} \quad & j(\mathbf{z}) \\ \text{subject to} \quad & \mathbf{h}(\mathbf{z}) = \mathbf{0}, \\ & \mathbf{g}(\mathbf{z}) \leq \mathbf{0} \end{aligned} \tag{4.1}$$

All functions in problem (4.1) are assumed to be at least twice continuously differentiable. The feasible region $\mathcal{F}$ for this optimization problem is represented by the set of points satisfying the constraints:

$$\mathcal{F} := \{\mathbf{z} \in \mathbb{R}^{n_z} : \mathbf{h}(\mathbf{z}) = \mathbf{0}, \mathbf{g}(\mathbf{z}) \leq \mathbf{0}\} \tag{4.2}$$

A solution $\hat{\mathbf{z}} \in \mathcal{F}$ to the optimization problem (4.1) is termed a *global solution* if

$$j(\hat{\mathbf{z}}) \leq j(\mathbf{z}), \qquad \mathbf{z} \in \mathcal{F}, \tag{4.3}$$

is satisfied. For a *local solution* $\hat{\mathbf{z}} \in \mathcal{F}$ there has to exist a neighborhood $\mathcal{N}$ of $\hat{\mathbf{z}}$ in which

$$j(\hat{\mathbf{z}}) \leq j(\mathbf{z}), \qquad \mathbf{z} \in \mathcal{N} \cap \mathcal{F} \tag{4.4}$$

holds. For a *strict* global, respectively local, solution strict inequalities need to hold for all $\mathbf{z} \neq \hat{\mathbf{z}}$ in Eq. (4.3), respectively Eq. (4.4). In addition, the index set $\mathcal{A}(\mathbf{z})$ of the active inequality constraints is introduced

$$\mathcal{A}(\mathbf{z}) := \{i \in \mathbb{N}_0 : g_i(\mathbf{z}) = 0\}, \tag{4.5}$$

which is used in the following discussion regarding the conditions of optimality.

4.2 Necessary and Sufficient Conditions

In order to state the optimality conditions for the optimization problem (4.1) it is useful to introduce multipliers $l_0 \in \mathbb{R}$, $\boldsymbol{\lambda} \in \mathbb{R}^{n_h}$, and $\boldsymbol{\mu} \in \mathbb{R}^{n_g}$ and to define the Lagrangian function $\mathcal{L} : \mathbb{R} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_h} \times \mathbb{R}^{n_g} \to \mathbb{R}$:

$$\mathcal{L}(l_0, \mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}) := l_0 j(\mathbf{z}) + \boldsymbol{\lambda}^T \mathbf{h}(\mathbf{z}) + \boldsymbol{\mu}^T \mathbf{g}(\mathbf{z}) \tag{4.6}$$

In the following the normal case is assumed for which the multiplier l_0 can be set to one [54]. A local solution $\hat{\mathbf{z}}$ with corresponding multipliers $\hat{\boldsymbol{\lambda}}$ and $\hat{\boldsymbol{\mu}}$ of the constrained optimization problem (4.1) has to satisfy the necessary conditions:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{z}}\left(\hat{\mathbf{z}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}\right) = \frac{\partial j}{\partial \mathbf{z}}(\hat{\mathbf{z}}) + \hat{\boldsymbol{\lambda}}^T \frac{\partial \mathbf{h}}{\partial \mathbf{z}}(\hat{\mathbf{z}}) + \hat{\boldsymbol{\mu}}^T \frac{\partial \mathbf{g}}{\partial \mathbf{z}}(\hat{\mathbf{z}}) = \mathbf{0}, \tag{4.7}$$

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda}}\left(\hat{\mathbf{z}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}\right) = \mathbf{h}^T(\hat{\mathbf{z}}) = \mathbf{0}, \tag{4.8}$$

$$\mathbf{g}(\hat{\mathbf{z}}) \leq \mathbf{0}, \tag{4.9}$$

$$\hat{\mu}_i \geq 0, \qquad i = 0, \ldots, n_g - 1, \tag{4.10}$$

$$g_i(\hat{\mathbf{z}})\hat{\mu}_i = 0, \qquad i = 0, \ldots, n_g - 1 \tag{4.11}$$

Additionally, the *linear independence constraint qualification* (LICQ) needs to hold, implying that all gradients of the active constraints have to be linearly independent. These first-order necessary conditions are also called Karush-Kuhn-Tucker (KKT) conditions. Note that the *stationarity condition* (4.7) requires the negative gradient of the cost func-

tion to be a linear combination of the constraint Jacobians:

$$-\frac{\partial j}{\partial \mathbf{z}}(\hat{\mathbf{z}}) = \hat{\boldsymbol{\lambda}}^T \frac{\partial \mathbf{h}}{\partial \mathbf{z}}(\hat{\mathbf{z}}) + \hat{\boldsymbol{\mu}}^T \frac{\partial \mathbf{g}}{\partial \mathbf{z}}(\hat{\mathbf{z}}) \tag{4.12}$$

In some sense, this linear combination can be compared to a vectorial force equation in which the Lagrange multipliers play the role of scaling factors maintaining an equilibrium (see Fig. 4.1). Moreover, the *primal feasibility conditions* (4.8) and (4.9) simply

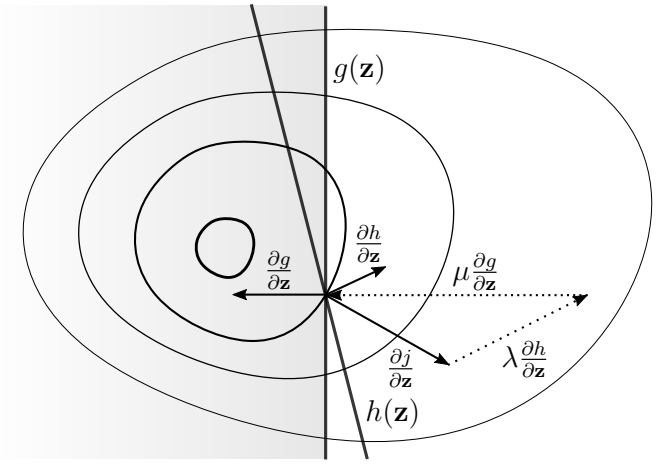


Figure 4.1: Two dimensional illustration of the vectorial sum of the cost function gradient and constraint Jacobians (scaled by the corresponding multipliers) at the optimal solution. A scalar equality constraint $h(\mathbf{z})$ as well as a scalar inequality constraint $g(\mathbf{z})$ are considered.

state that the constraints need to be satisfied at the local solution $\hat{\mathbf{z}}$. Additionally, the *dual feasibility conditions* (4.10) require the multipliers of the inequality constraints to be non-negative. Finally, the *complementarity conditions* (4.11) ensure that for inactive constraints the corresponding multipliers, or both the constraint values and the multipliers, have to be zero. In the latter case a constraint is said to be *weakly active*. A stronger condition is the *strict* complementarity condition which requires the multipliers corresponding to active constraints to be strictly positive:

$$\mu_i > 0, \qquad g_i(\hat{\mathbf{z}})\hat{\mu}_i = 0, \qquad \forall i \in \mathcal{A}(\hat{\mathbf{z}}) \tag{4.13}$$

Note that the first-order conditions do not specify the type of extremal point $\hat{\mathbf{z}}$ for which second-order information is required. For the statement of the second-order necessary conditions it is useful to introduce the critical cone $\mathcal{C}(\hat{\mathbf{z}})$

$$\mathcal{C}(\hat{\mathbf{z}}) := \left\{ \mathbf{d} \in \mathbb{R}^{n_z} : \begin{cases} \frac{\partial \mathbf{h}}{\partial \mathbf{z}}(\hat{\mathbf{z}})\mathbf{d} = \mathbf{0} \\ \frac{\partial g_i}{\partial \mathbf{z}}(\hat{\mathbf{z}})\mathbf{d} = 0, & \forall i \in \mathcal{A}(\hat{\mathbf{z}}), \hat{\mu}_i > 0 \\ \frac{\partial g_i}{\partial \mathbf{z}}(\hat{\mathbf{z}})\mathbf{d} \leq 0, & \forall i \in \mathcal{A}(\hat{\mathbf{z}}), \hat{\mu}_i = 0 \end{cases} \right\}, \tag{4.14}$$

which represents a conical approximation of the feasible set. The second-order necessary conditions additionally require the projection of the Hessian of the Lagrangian in all admissibly directions $\mathbf{d}$, as defined by the critical cone $\mathcal{C}(\hat{\mathbf{z}})$, to be greater than or equal to zero:

$$\mathbf{d}^T \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}^2} \mathbf{d} \geq 0, \qquad \mathbf{d} \in \mathcal{C}(\hat{\mathbf{z}}) \tag{4.15}$$

For the second-order condition to be sufficient the inequality in condition (4.15) is required to be strict.

Note that for the general case it is by no means easy to find analytical solutions satisfying the KKT conditions. As such, numerical methods are vital for practical applications. These methods are the topic of the following sections. Important special cases, namely Linear Programming (LP) and Quadratic Programming (QP), as well as the general case of Nonlinear Programming (NLP) are discussed.

4.3 Linear Programming

Consider a special case of the optimization problem (4.1) where the objective function, the equality constraints, as well as inequality constraints are linear. This optimization problem can be written as:

$$\begin{aligned} \underset{\bar{\mathbf{z}} \in \mathbb{R}^{n_{\bar{z}}}}{\text{minimize}} \quad & \bar{\mathbf{c}}^T \bar{\mathbf{z}} \\ \text{subject to} \quad & \mathbf{M}_h \bar{\mathbf{z}} = \mathbf{b}_h, \\ & \mathbf{M}_g \bar{\mathbf{z}} \leq \mathbf{b}_g \end{aligned} \tag{4.16}$$

The linear cost function is modeled using the vector $\bar{\mathbf{c}} \in \mathbb{R}^{n_{\bar{z}}}$. Furthermore, the matrix $\mathbf{M}_h \in \mathbb{R}^{n_h \times n_{\bar{z}}}$ and corresponding vector $\mathbf{b}_h \in \mathbb{R}^{n_h}$ define the equality constraints. Similarly, the matrix $\mathbf{M}_g \in \mathbb{R}^{n_g \times n_{\bar{z}}}$ and corresponding vector $\mathbf{b}_g \in \mathbb{R}^{n_h}$ define the inequality constraints. Usually the problem is stated in *standard form*

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{R}^{n_z}}{\text{minimize}} \quad & \mathbf{c}^T \mathbf{z} \\ \text{subject to} \quad & \mathbf{M}_s \mathbf{z} = \mathbf{b}_s, \\ & \mathbf{z} \geq \mathbf{0}, \end{aligned} \tag{4.17}$$

with $\mathbf{c} \in \mathbb{R}^{n_z}$, $\mathbf{M}_s \in \mathbb{R}^{n_c \times n_z}$, and $\mathbf{b}_s \in \mathbb{R}^{n_z}$. The statement of the LP (4.16) in the form (4.17) can always be achieved by means of suitable standard transformations (cf. Ref. [4]). It is assumed in the following that the feasible set is non-empty, $n_c < n_z$ holds, and the rows in matrix $\mathbf{M}_s$ are linearly independent (directly implying that $\mathbf{M}_h$ is required to have full row rank). The Lagrangian for the standard form with

multiplier vectors $\boldsymbol{\lambda} \in \mathbb{R}^{n_c}$ and $\boldsymbol{\mu} \in \mathbb{R}^{n_z}$ may be written as

$$\mathcal{L}\left(\mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}\right) := \mathbf{c}^T\mathbf{z} + \boldsymbol{\lambda}^T\left(\mathbf{b}_s - \mathbf{M}_s\mathbf{z}\right) - \boldsymbol{\mu}^T\mathbf{z}, \tag{4.18}$$

and the optimality conditions for the KKT triple $\left(\hat{\mathbf{z}}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\lambda}}\right)$ are [4]

$$\mathbf{M}_s^T\hat{\boldsymbol{\lambda}} + \hat{\boldsymbol{\mu}} = \mathbf{c}, \tag{4.19}$$
$$\mathbf{M}_s\hat{\mathbf{z}} = \mathbf{b}_s, \tag{4.20}$$
$$\hat{\boldsymbol{\mu}} \geq 0, \tag{4.21}$$
$$\hat{\mathbf{z}} \geq 0, \tag{4.22}$$
$$\hat{\mathbf{Z}}\hat{\boldsymbol{\mu}} = 0, \tag{4.23}$$
$$\tag{4.24}$$

with $\hat{\mathbf{Z}} = diag(\hat{\mathbf{z}})$. A globally optimal solution of the LP problem is guaranteed to be on a vertex of the convex polytope defining the feasible set. An effective method for finding such an optimal vertex is the Simplex method [55]. Essentially, this algorithm visits vertices of the feasible set until an optimal solution is found. Geometrically, this can be interpreted as moving along the edges of the polytope in directions of non-increasing cost until the cost cannot be improved further. All vertices which are visited during the iterations of the Simplex algorithm are *basic feasible*. This means that they are vertices of the feasible polytope and at each point there exists an index set $\mathcal{B} \subseteq \{0, \ldots, n_z - 1\}$ of cardinality n_c with $z_i = 0, \forall i \notin \mathcal{B}$. Moreover, the *basis matrix* $\mathbf{M}_B$ comprised of the columns $\mathbf{m}_i, i \in \mathcal{B}$ of $\mathbf{M}_s$ has full rank and and is invertible. Furthermore, the complementary *non-basic* index set $\mathcal{N} = \{0, \ldots, n_z - 1\} \setminus \mathcal{B}$ with cardinality $n_z - n_c$ may be defined. Using the definition of the basic and non-basic index sets the variables $\mathbf{z}$ and $\boldsymbol{\mu}$ may be rearranged into the *basic variables* $\mathbf{z}_b \in \mathbb{R}^{n_c}$ and $\boldsymbol{\mu}_b \in \mathbb{R}^{n_c}$ which contain all $z_i, \forall i \in \mathcal{B}$ and $\mu_i, \forall i \in \mathcal{B}$ as well as the *non-basic variables* $\mathbf{z}_n \in \mathbb{R}^{n_z - n_c}$ and $\boldsymbol{\mu}_n \in \mathbb{R}^{n_z - n_c}$ collecting all $z_i, \forall i \in \mathcal{N}$, respectively $\mu_i, \forall i \in \mathcal{N}$. Furthermore, the columns of the matrix $\mathbf{M}_s$ can be rearranged to $[\mathbf{M}_B, \mathbf{M}_N]$ with the invertible basis matrix $\mathbf{M}_B$ and the matrix $\mathbf{M}_N$ collecting all columns of $\mathbf{M}_s$ not indexed by $\mathcal{B}$. A *basic feasible optimal* point is a basic feasible point which is a solution for the LP (4.17). A basic feasible optimal solution is guaranteed to exist under relatively mild assumptions for the LP problem (see Theorem 13.2 in Ref. [4]):

- "If [(4.17)] has a nonempty feasible region, then there is at least one basic feasible point"
- "If [(4.17)] has solutions, then at least one such solution is a basic optimal point"
- "If [(4.17)] is feasible and bounded, then it has an optimal solution"

Moreover, if the strict complementarity condition optimal solution, which may be expressed as

$$\hat{\mathbf{z}} + \hat{\boldsymbol{\mu}} > \mathbf{0}, \tag{4.25}$$

then the optimal solution is unique [56] and the basic and non-basic variables satisfy $\mathbf{z}_b > 0, \mathbf{z}_n = 0$, respectively $\boldsymbol{\mu}_b = 0, \boldsymbol{\mu}_n > 0$. Efficient implementations of LP methods which are able to handle large scale problems are available in software packages such as CPLEX [57], GUROBI [58], and CLP [59].

4.4 Quadratic Programming

QP methods solve an optimization problem with a quadratic cost function and linear constraints of the form:

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{R}^{n_z}}{\text{minimize}} \quad & \mathbf{c}^T\mathbf{z} + \frac{1}{2}\mathbf{z}^T\mathbf{B}\mathbf{z} \\ \text{subject to} \quad & \mathbf{M}_h\mathbf{z} = \mathbf{b}_h, \\ & \mathbf{M}_g\mathbf{z} \leq \mathbf{b}_g \end{aligned} \tag{4.26}$$

The quadratic cost function contains a linear term with vector $\mathbf{c} \in \mathbb{R}^{n_z}$ and a quadratic term with symmetric matrix $\mathbf{B} \in \mathbb{R}^{n_z \times n_z}$. The equality and inequality constraints are defined as for LPs (cf. Sec. 4.3). The Lagrangian of the QP problem is defined as follows:

$$\mathcal{L}\left(\mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}\right) := \mathbf{c}^T\mathbf{z} + \frac{1}{2}\mathbf{z}^T\mathbf{B}\mathbf{z} + \boldsymbol{\lambda}^T\left(\mathbf{M}_h\mathbf{z} - \mathbf{b}_h\right) + \boldsymbol{\mu}^T\left(\mathbf{M}_g\mathbf{z} - \mathbf{b}_g\right) \tag{4.27}$$

An important special case arises if the matrix $\mathbf{B}$ is positive definite. In this case the problem is convex and there exists a unique global minimizer. This is not true for non-convex problems and there may exist several local solutions. Moreover, in case the quadratic problem has only equality constraints the first-order optimality conditions can be stated as a solution for the optimization variables $\hat{\mathbf{z}}$ and the Lagrange multipliers $\hat{\boldsymbol{\lambda}}$ of the system:

$$\begin{bmatrix} \mathbf{B} & \mathbf{M}_h^T \\ \mathbf{M}_h & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{z}} \\ \hat{\boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} -\mathbf{c} \\ \mathbf{b}_h \end{bmatrix} \tag{4.28}$$

Two commonly applied approaches to solve the QP problem with inequality constraints are active set and interior point methods. Among others, a popular implementation of an active set QP algorithm is qpOASES [60]. This solver is widely used in optimal control applications, especially in the context of model predictive control.

4.5 Nonlinear Programming

For problems of the general nonlinear type two methods, namely Sequential Quadratic Programming (SQP) and Interior Point (IP) methods, are discussed in the following. They both differ in the way how iterates are computed and how inequality constraints are treated. SQP methods solve the constrained optimization problem by a series of quadratic model problems with linear constraints. The index set for the active inequality constraints is updated in each iteration based on the active constraints from the solution of the QP. For solving the quadratic model problem in each iteration methods discussed in Sec. 4.4 are used. A popular implementation of a SQP method is the nonlinear solver SNOPT [61]. Besides SQP methods, IP methods are commonly used for solving NLPs. These methods can be seen as homotopy methods for solving a sequence of perturbed optimization problems. For this purpose the nonlinear optimization problem (4.1) is first transformed to the equivalent form:

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{R}^{n_z}, \mathbf{s} \in \mathbb{R}^{n_g}}{\text{minimize}} \quad & j(\mathbf{z}) \\ \text{subject to} \quad & \mathbf{h}(\mathbf{z}) = \mathbf{0}, \\ & \mathbf{g}(\mathbf{z}) + \mathbf{s} = \mathbf{0}, \\ & \mathbf{s} \geq \mathbf{0} \end{aligned} \tag{4.29}$$

Note that the inequality constraints are replaced by equality constraints $\mathbf{g}(\mathbf{z}) + \mathbf{s} = \mathbf{0}$ using additional slack variables $\mathbf{s} \in \mathbb{R}^{n_g}, \mathbf{s} \geq 0$. Defining the Lagrangian of this problem based on the multipliers $\boldsymbol{\lambda} \in \mathbb{R}^{n_h}$ and $\boldsymbol{\mu} \in \mathbb{R}^{n_g}$

$$\mathcal{L}(\mathbf{z}, \mathbf{s}, \boldsymbol{\lambda}, \boldsymbol{\mu}) := j(\mathbf{z}) + \boldsymbol{\lambda}^T \mathbf{h}(\mathbf{z}) + \boldsymbol{\mu}^T (\mathbf{g}(\mathbf{z}) + \mathbf{s}), \tag{4.30}$$

together with the non-negativity inequality constraints $\mathbf{s} \geq \mathbf{0}$ the KKT conditions at a local solution with $\hat{\mathbf{z}}$, $\hat{\mathbf{s}}$, $\hat{\boldsymbol{\lambda}}$, and $\hat{\boldsymbol{\mu}}$ can be stated as [4]

$$\frac{\partial \mathcal{L}}{\partial \mathbf{z}} \left(\hat{\mathbf{z}}, \hat{\mathbf{s}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}} \right) = \frac{\partial j}{\partial \mathbf{z}} (\hat{\mathbf{z}}) + \hat{\boldsymbol{\lambda}}^T \frac{\partial \mathbf{h}}{\partial \mathbf{z}} (\hat{\mathbf{z}}) + \hat{\boldsymbol{\mu}}^T \frac{\partial \mathbf{g}}{\partial \mathbf{z}} (\hat{\mathbf{z}}) = \mathbf{0}, \tag{4.31}$$

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda}} \left(\hat{\mathbf{z}}, \hat{\mathbf{s}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}} \right) = \mathbf{h}^T (\hat{\mathbf{z}}) = \mathbf{0}, \tag{4.32}$$

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\mu}} \left(\hat{\mathbf{z}}, \hat{\mathbf{s}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}} \right) = \mathbf{g}^T (\hat{\mathbf{z}}) + \hat{\mathbf{s}}^T = \mathbf{0}, \tag{4.33}$$

$$\hat{\mu}_i \geq 0, \qquad i = 0, \ldots, n_g - 1, \tag{4.34}$$

$$\hat{s}_i \geq 0, \qquad i = 0, \ldots, n_g - 1, \tag{4.35}$$

$$\hat{s}_i \hat{\mu}_i = \tau, \qquad i = 0, \ldots, n_g - 1, \tag{4.36}$$

for $\tau = 0$. The solution of the constrained optimization problem is now determined by solving a sequence of perturbed problems with a positive perturbation parameter $\tau > 0$. As such, IP methods can be seen as continuation methods with homotopy parameter τ. In this context, it is interesting to observe that the condition (4.36) approximates the complementarity condition using a "smoothed corner" as illustrated in Fig. 4.2. Hence, the concept of inactive constraints essentially looses its meaning and, there are merely "less active" or "more active" constraints. Note further that for $\tau > 0$ the slack variables $\mathbf{s}$ and the multipliers $\boldsymbol{\mu}$ are forced to be strictly positive by the relaxed complementarity condition and hence $\boldsymbol{\mu} > \mathbf{0}$ as well as $\mathbf{s} > \mathbf{0}$ needs to be ensured in each iteration. This implies that strict complementarity holds for all iterates with $\tau > 0$ in the interior of the feasible set. The Newton directions $\mathbf{d}_z$, $\mathbf{d}_s$, $\mathbf{d}_\lambda$, and $\mathbf{d}_\mu$ for solving

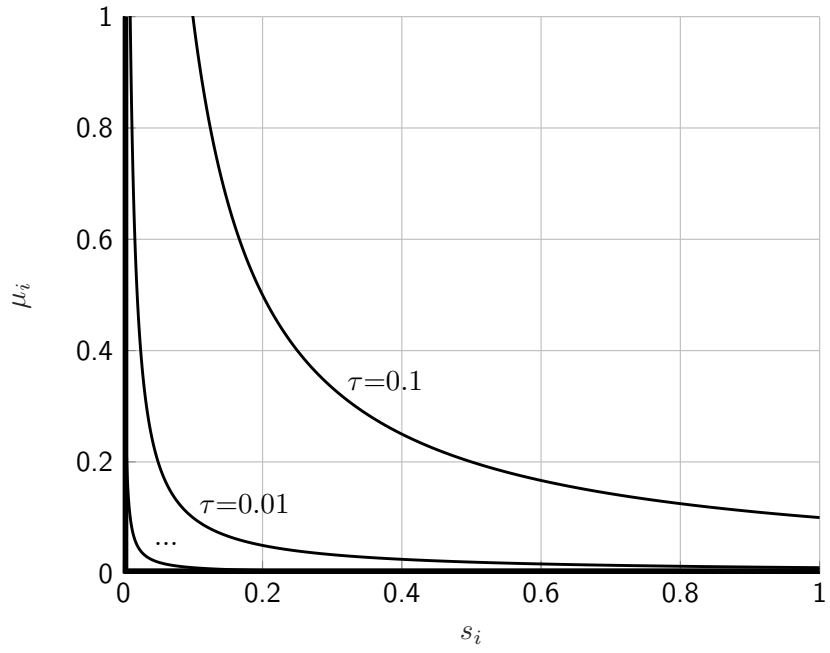


Figure 4.2: Illustration of the relaxed complementarity condition for one multiplier μ_i and corresponding slack s_i in case of interior point problems with $\tau > 0$.

the KKT equations in each iteration are determined from the solution of

$$\begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}^2} & \left(\frac{\partial \mathbf{h}}{\partial \mathbf{z}}\right)^T & \left(\frac{\partial \mathbf{g}}{\partial \mathbf{z}}\right)^T & \mathbf{0} \\ \frac{\partial \mathbf{h}}{\partial \mathbf{z}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{\partial \mathbf{g}}{\partial \mathbf{z}} & \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{S} & \mathbf{M} \end{bmatrix} \begin{bmatrix} \mathbf{d}_z \\ \mathbf{d}_\lambda \\ \mathbf{d}_\mu \\ \mathbf{d}_s \end{bmatrix} = - \begin{bmatrix} \frac{\partial j}{\partial \mathbf{z}} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{h}}{\partial \mathbf{z}} + \boldsymbol{\mu}^T \frac{\partial \mathbf{g}}{\partial \mathbf{z}} \\ \mathbf{h}(\mathbf{z}) \\ \mathbf{g}(\mathbf{z}) + \mathbf{s} \\ \mathbf{S}\boldsymbol{\mu} - \mathbf{1}\tau \end{bmatrix}, \tag{4.37}$$

with $\mathbf{S} = diag(\mathbf{s})$ and $\mathbf{M} = diag(\boldsymbol{\mu})$. In order to measure progress in each iteration when employing a Line-Search method a so-called *merit function* is defined. This merit

function takes the cost function value as well as constraint violations into account. An example for such a merit function based on a suitable vector norm $\|\cdot\|$ is [4]:

$$m(\mathbf{z}, \mathbf{s}, \boldsymbol{\lambda}, \boldsymbol{\mu}; \tau) := \max \left\{ \left\| \frac{\partial j}{\partial \mathbf{z}} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{h}}{\partial \mathbf{z}} + \boldsymbol{\mu}^T \frac{\partial \mathbf{g}}{\partial \mathbf{z}} \right\|, \|\mathbf{S}\boldsymbol{\mu} - \mathbf{1}\tau\|, \|\mathbf{h}(\mathbf{z})\|, \|\mathbf{g}(\mathbf{z}) + \mathbf{s}\| \right\} \tag{4.38}$$

Popular interior point solvers are IPOPT [62], KNITRO [63], and LOQO [64].

4.6 Sensitivity Analysis for Parametric Optimization Problems

Post-optimal sensitivity analysis [53, 65, 52, 66] can be used to study the effect of parameters on the optimal solution of optimization problems. For this purpose, consider the parameter dependent case of the constrained minimization problem (4.1):

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{R}^{n_z}}{\text{minimize}} \quad & j(\mathbf{z}; \mathbf{p}), \\ \text{subject to} \quad & \mathbf{h}(\mathbf{z}; \mathbf{p}) = \mathbf{0}, \\ & \mathbf{g}(\mathbf{z}; \mathbf{p}) \leq \mathbf{0} \end{aligned} \tag{4.39}$$

Note that the parameters $\mathbf{p} \in \mathbb{R}^{n_p}$ in this optimization problem are not subject to optimization. The Lagrangian of this parametric optimization problem is defined as:

$$\mathcal{L}(\mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{p}) := j(\mathbf{z}; \mathbf{p}) + \boldsymbol{\lambda}^T \mathbf{h}(\mathbf{z}; \mathbf{p}) + \boldsymbol{\mu}^T \mathbf{g}(\mathbf{z}; \mathbf{p}) \tag{4.40}$$

It is assumed that the second-order sufficient conditions with strict complementarity and LICQ (cf. Sec. 4.2) hold at the nominal solution $\hat{\mathbf{z}}(\mathbf{p}_0)$. In this case there exists a neighborhood of $\mathbf{p}_0$ where the set of active constraints remains unchanged and the trajectories $\hat{\mathbf{z}}(\mathbf{p})$, $\hat{\boldsymbol{\lambda}}(\mathbf{p})$ and $\hat{\boldsymbol{\mu}}(\mathbf{p})$ are continuously differentiable [52]. In the following function arguments are dropped in order to improve readability. In addition, all quantities are assumed to be evaluated at a KKT point $\left(\hat{\mathbf{z}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}\right)$ for a nominal parameter value $\mathbf{p}_0$. First, the vector function $\mathbf{v}(\mathbf{p})$

$$\mathbf{v}(\mathbf{p}) := \left[\mathbf{z}^T(\mathbf{p}), \boldsymbol{\lambda}^T(\mathbf{p}), \boldsymbol{\mu}^T(\mathbf{p})\right]^T, \tag{4.41}$$

is introduced which expresses the implicit dependence of the optimal optimization variables and Lagrange multipliers on the parameter vector $\mathbf{p}$. The following condition

holds at an optimal solution of problem (4.39) (cf. Eqs. (4.7), (4.8), and (4.11))

$$\mathbf{c}(\mathbf{p}) := \begin{bmatrix} \left(\frac{\partial \mathcal{L}}{\partial \mathbf{z}}\right)^T \\ \mathbf{h} \\ \mathbf{M}\mathbf{g} \end{bmatrix} = \mathbf{0}, \tag{4.42}$$

with $\mathbf{M} = diag(\boldsymbol{\mu})$. Note that the total derivative of this condition is not allowed to change for the perturbed solution to remain optimal, i.e. :

$$\frac{\partial \mathbf{c}}{\partial \mathbf{v}} \frac{d\mathbf{v}}{d\mathbf{p}} + \frac{\partial \mathbf{c}}{\partial \mathbf{p}} = \mathbf{0} \tag{4.43}$$

The Jacobian

$$\frac{\partial \mathbf{c}}{\partial \mathbf{v}} = \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}^2} & \left(\frac{\partial \mathbf{h}}{\partial \mathbf{z}}\right)^T & \left(\frac{\partial \mathbf{g}}{\partial \mathbf{z}}\right)^T \\ \frac{\partial \mathbf{h}}{\partial \mathbf{z}} & \mathbf{0} & \mathbf{0} \\ \mathbf{M}\frac{\partial \mathbf{g}}{\partial \mathbf{z}} & \mathbf{0} & \mathbf{G} \end{bmatrix}, \tag{4.44}$$

with $\mathbf{G} = diag(\mathbf{g})$ exhibits a particular structure which can be exploited to simplify Eq. (4.43). Let the active inequality constraints and corresponding Lagrange multipliers be denoted with $\mathbf{g}^a \in \mathbb{R}^{n_a}$, respectively $\boldsymbol{\mu}^a \in \mathbb{R}^{n_a}$. Similarly, the inactive constraints and multipliers are denoted with $\mathbf{g}^i \in \mathbb{R}^{n_i}$ and $\boldsymbol{\mu}^i \in \mathbb{R}^{n_i}$. Note that due to the assumption of a strict complementary solution (cf. Eq. (4.13)) the multipliers $\boldsymbol{\mu}^i$ and constraints $\mathbf{g}^a$ are zero and the multipliers $\boldsymbol{\mu}^a$ and constraints $\mathbf{g}^i$ are non-zero. Defining the matrices $\mathbf{G}^i = diag(\mathbf{g}^i)$ and $\mathbf{G}^a = diag(\mathbf{g}^a)$ as well as $\mathbf{M}^i = diag(\boldsymbol{\mu}^i)$ and $\mathbf{M}^a = diag(\boldsymbol{\mu}^a)$ the block structures

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}^i & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^a \end{bmatrix} = \begin{bmatrix} \mathbf{G}^i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{4.45}$$

and

$$\mathbf{M}\frac{\partial \mathbf{g}}{\partial \mathbf{z}} = \begin{bmatrix} \mathbf{M}^i \frac{\partial \mathbf{g}^i}{\partial \mathbf{z}} \\ \mathbf{M}^a \frac{\partial \mathbf{g}^a}{\partial \mathbf{z}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{M}^a \frac{\partial \mathbf{g}^a}{\partial \mathbf{z}} \end{bmatrix}, \tag{4.46}$$

can be exposed under the assumption that the constraints are arranged as follows:

$$\mathbf{g} = \begin{bmatrix} \mathbf{g}^i \\ \mathbf{g}^a \end{bmatrix} \tag{4.47}$$

Exploiting these structures in Eq. (4.44), the following linear system is obtained from Eq. (4.43):

$$\begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}^2} & \left(\frac{\partial \mathbf{h}}{\partial \mathbf{z}}\right)^T & \left(\frac{\partial \mathbf{g}^i}{\partial \mathbf{z}}\right)^T & \left(\frac{\partial \mathbf{g}^a}{\partial \mathbf{z}}\right)^T \\ \frac{\partial \mathbf{h}}{\partial \mathbf{z}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}^i & \mathbf{0} \\ \mathbf{M}^a \frac{\partial \mathbf{g}^a}{\partial \mathbf{z}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \frac{d\mathbf{z}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\lambda}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\mu}^i}{d\mathbf{p}} \\ \frac{d\boldsymbol{\mu}^a}{d\mathbf{p}} \end{bmatrix} = - \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z} \partial \mathbf{p}} \\ \frac{\partial \mathbf{h}}{\partial \mathbf{p}} \\ \mathbf{0} \\ \mathbf{M}^a \frac{\partial \mathbf{g}^a}{\partial \mathbf{p}} \end{bmatrix} \tag{4.48}$$

As the third row block equation of Eq. (4.48) implies

$$\mathbf{G}^i \frac{d\boldsymbol{\mu}^i}{d\mathbf{p}} = \mathbf{0}, \tag{4.49}$$

it immediately follows that

$$\frac{d\boldsymbol{\mu}^i}{d\mathbf{p}} = \mathbf{0}, \tag{4.50}$$

due to the fact that $\mathbf{G}^i$ is a diagonal matrix with non-zero entries on the diagonal. Note that this is consistent with the requirement that there are no active set changes in the vicinity of the nominal solution (inactive constraints stay inactive). Note further that $\mathbf{M}^a$ is a diagonal matrix with strictly positive entries on the main diagonal. Thus, after multiplying the fourth row with the inverse of $\mathbf{M}^a$ and removing the rows and columns corresponding to the result from Eq. (4.50), the following system is obtained:

$$\begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}^2} & \left(\frac{\partial \mathbf{h}}{\partial \mathbf{z}}\right)^T & \left(\frac{\partial \mathbf{g}^a}{\partial \mathbf{z}}\right)^T \\ \frac{\partial \mathbf{h}}{\partial \mathbf{z}} & \mathbf{0} & \mathbf{0} \\ \frac{\partial \mathbf{g}^a}{\partial \mathbf{z}} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \frac{d\mathbf{z}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\lambda}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\mu}^a}{d\mathbf{p}} \end{bmatrix} = - \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z} \partial \mathbf{p}} \\ \frac{\partial \mathbf{h}}{\partial \mathbf{p}} \\ \frac{\partial \mathbf{g}^a}{\partial \mathbf{p}} \end{bmatrix} \tag{4.51}$$

Due to the assumption that second-order sufficient conditions with strict complementarity and LICQ hold, the sensitivity matrix in Eq. (4.51) has full rank and is invertible (cf. Ref. [52]). Hence, the post-optimal sensitivities w.r.t. the parameters $\mathbf{p}$ yield:

$$\begin{bmatrix} \frac{d\mathbf{z}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\lambda}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\mu}^a}{d\mathbf{p}} \end{bmatrix} = - \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}^2} & \left(\frac{\partial \mathbf{h}}{\partial \mathbf{z}}\right)^T & \left(\frac{\partial \mathbf{g}^a}{\partial \mathbf{z}}\right)^T \\ \frac{\partial \mathbf{h}}{\partial \mathbf{z}} & \mathbf{0} & \mathbf{0} \\ \frac{\partial \mathbf{g}^a}{\partial \mathbf{z}} & \mathbf{0} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z} \partial \mathbf{p}} \\ \frac{\partial \mathbf{h}}{\partial \mathbf{p}} \\ \frac{\partial \mathbf{g}^a}{\partial \mathbf{p}} \end{bmatrix} \tag{4.52}$$

These sensitivities are termed *post-optimal* sensitivities as they reveal how the optimal solution, i.e. the optimal optimization variables and multipliers, changes in a neighborhood around the nominal solution. Based on the post-optimal sensitivities of the optimization variables and multipliers it is possible to obtain the sensitivities of the

optimal cost function:

$$\frac{dj}{d\mathbf{p}} = \frac{\partial j}{\partial \mathbf{z}} \frac{d\mathbf{z}}{d\mathbf{p}} + \frac{\partial j}{\partial \mathbf{p}} \tag{4.53}$$

In addition to expression (4.53), it can be shown that the post-optimal sensitivity of the cost function is equal to the partial derivatives of the Lagrangian with respect to the parameters [66]:

$$\frac{dj}{d\mathbf{p}} = \frac{\partial \mathcal{L}}{\partial \mathbf{p}} \tag{4.54}$$

Note that this expression is considerably cheaper to evaluate as it does not require the sensitivity of the optimization variables with respect to the parameters. Interestingly, also second-order sensitivities can be derived for the cost function [65]:

$$\frac{d^2 j}{d\mathbf{p}^2} = \frac{\partial^2 \mathcal{L}}{\partial \mathbf{p} \partial \mathbf{z}} \frac{d\mathbf{z}}{d\mathbf{p}} + \frac{\partial^2 \mathcal{L}}{\partial \mathbf{p}^2} + \left(\frac{\partial \mathbf{g}^a}{\partial \mathbf{p}} \right)^T \frac{d\boldsymbol{\mu}^a}{d\mathbf{p}} + \left(\frac{\partial \mathbf{h}}{\partial \mathbf{p}} \right)^T \frac{d\boldsymbol{\lambda}}{d\mathbf{p}} \tag{4.55}$$

Note that the cost function sensitivity deserves particular attention considering the application investigated in this thesis. This is due to the fact that the cost function models the clearance criterion under investigation. Thus, the post-optimal sensitivities of the cost function with respect to parameters can provide valuable information regarding the influence of these parameters on the worst-case solution in the context of flight control law clearance.

Chapter 5

Optimal Control

Optimal control theory is the basis of the clearance methods developed in this thesis. Two approaches for solving optimal control problems are commonly employed, termed the *direct* and the *indirect* approach. In the direct approach the problem is first discretized in the temporal dimension and the resulting parameter optimization problem is then solved using the methods discussed in the previous chapter 4. This sequence may be described as *first discretize, then optimize*. The *indirect* approach relies on the derivation of the necessary optimality conditions for the continuous time problem. A solution of these necessary conditions can then be obtained analytically or numerically. If a numerical method is used to solve the necessary conditions, the indirect approach can be described as *first optimize, then discretize* [52]. Both options for the indirect approach are by no means foolproof and in most cases quite involved. On the one hand, the analytical solution is only possible for very few, often relatively simple, optimal control problems. On the other hand, the numerical solution of the necessary conditions, e.g. using shooting techniques [67], requires a good initial guess and knowledge about the structure of the problem (such as the sequence of active and inactive boundary arcs, cf. Sec. 5.4). In this thesis the direct approach is used for the solution of the clearance optimal control problem due to its general applicability. Nevertheless, the concepts presented in this chapter are useful to characterize the worst-case control function and to verify if the solution obtained from the discretized problem in fact represents a valid approximation of the solution expected from the continuous time problem. Especially the connection between the discretized problem and the continuous time problem, which is discussed in chapter 6, is essential in order to perform these sanity checks. In this chapter results from the continuous case are highlighted, which are deemed relevant for the clearance problem investigated in this thesis.

5.1 Problem Statement

Optimal control problems [52, 54, 50] are optimization problems in function space. In order to state a commonly used formulation involving ordinary differential equations consider the dynamic constraint

$$\mathbf{f}\left(\mathbf{x}(t), \mathbf{u}(t), t\right) - \dot{\mathbf{x}}(t) = \mathbf{0}, \tag{5.1}$$

on the time interval I^t

$$I^t := [t_0, t_f] \subset \mathbb{R}, t_f > t_0. \tag{5.2}$$

The non-autonomous dynamic model function $\mathbf{f} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times I^t \to \mathbb{R}^{n_x}$ in Eq. (5.1) depends on the state function $\mathbf{x}(t) \in W_{1,\infty}^{n_x}(I^t)$, the control function $\mathbf{u}(t) \in L_{\infty}^{n_u}(I^t)$, and the monotonically increasing time variable $t \in I^t$. For an exact definition of the function spaces please refer to Ref. [52]. In particular, these function spaces allow the control to exhibit discontinuities, but require the state history to be continuous, i.e. without "jumps". The optimal control problem formulation considered in the following can be stated as follows: Find the optimal control function $\hat{\mathbf{u}}(t) \in \mathcal{U}$ within the control set $\mathcal{U} \subseteq \mathbb{R}^{n_u}$ and the corresponding optimal state history $\hat{\mathbf{x}}(t)$ which minimize the Bolza cost function:

$$j^B := j^M\left(\mathbf{x}(t_0), t_0, \mathbf{x}(t_f), t_f\right) + \int_{t_0}^{t_f} j^L\left(\mathbf{x}(t), \mathbf{u}(t), t\right)\, dt \tag{5.3}$$

On the one hand, the Mayer cost function $j^M : \mathbb{R}^{n_x} \times \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}$ in (5.3) depends on the values of the initial time t_0, final time t_f, and the state values at these time points, i.e. $\mathbf{x}(t_0)$, respectively $\mathbf{x}(t_f)$. On the other hand, the Lagrange cost function $j^L : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times I^t \to \mathbb{R}$ is integrated on the time interval I^t. Additionally, the path-constraints $\mathbf{c} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times I^t \to \mathbb{R}^{n_c}$,

$$\mathbf{c}\left(\mathbf{x}(t), \mathbf{u}(t), t\right) \leq \mathbf{0}, \tag{5.4}$$

are imposed along the trajectory and boundary conditions $\boldsymbol{\phi} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_\phi}$,

$$\boldsymbol{\phi}\left(\mathbf{x}(t_0), \mathbf{x}(t_f)\right) = \mathbf{0}, \tag{5.5}$$

depending on the state values at the initial and final time point constrain the problem at the beginning and end of the time interval. Note that the formulation above also includes the case with parameters $\mathbf{p} \in \mathbb{R}^{n_p}$ which can be formally introduced by adding additional state variables $x_{p_i}(t), i = 0, \ldots, n_p - 1$ to the formulation with a zero rate and

an initial (or final) boundary condition:

$$\dot{x}_{p_i}(t) = 0, \qquad x_{p_i}(t_0) = p_i \tag{5.6}$$

This transformation can be used to introduce the parameters for the initial time t_0 and final time t_f as states $x_{t_0}(t)$, respectively $x_{t_f}(t)$, with

$$\begin{aligned} \dot{x}_{t_0}(t) &= 0, \qquad x_{t_0}(t_0) = t_0, \\ \dot{x}_{t_f}(t) &= 0, \qquad x_{t_f}(t_f) = t_f. \end{aligned} \tag{5.7}$$

A similar idea can be used to transform the problem to an autonomous formulation with no explicit dependence on the independent variable t by the introduction of an additional state $x_t(t)$:

$$\dot{x}_t(t) = 1, \qquad x_t(t_0) = t_0 \tag{5.8}$$

Furthermore, the Bolza cost function can be transformed into a pure Mayer-type cost function by the introduction of an additional state $x_{j^L}(t)$:

$$\dot{x}_{j^L}(t) = j^L\left(\mathbf{x}(t), \mathbf{u}(t), t\right), \qquad x_{j^L}(t_0) = 0 \tag{5.9}$$

The augmented state vector $\bar{\mathbf{x}}(t)$ using these transformations thus becomes

$$\bar{\mathbf{x}}(t) = \left[\mathbf{x}^T(t), x_{t_0}(t), x_{t_f}(t), x_t(t), x_{j^L}(t)\right]^T. \tag{5.10}$$

Moreover, the problem can be formulated in time normalized form, which is preferred in most derivations throughout this thesis. This means that the problem is stated using the normalized time variable $\tau \in I^\tau$ on the interval

$$I^\tau := [0, 1]. \tag{5.11}$$

The transformations between the time t and the normalized time τ are

$$t = x_{t_0}(t) + \tau\left(x_{t_f}(t) - x_{t_0}(t)\right), \qquad \tau = \frac{t - x_{t_0}(t)}{x_{t_f}(t) - x_{t_0}(t)}, \tag{5.12}$$

Using these transformations the time scaled form of the dynamic equation involving the augmented state vector $\bar{\mathbf{x}}(t)$ can be given as follows:

$$\bar{\mathbf{x}}'(\tau) = \frac{d\bar{\mathbf{x}}}{dt}\frac{dt}{d\tau} = \frac{d\bar{\mathbf{x}}}{dt}\left(x_{t_f}(t) - x_{t_0}(t)\right) = \bar{\mathbf{f}}\left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right) \tag{5.13}$$

This leads to the following time normalized Mayer-type problem:

$$\begin{aligned} \underset{\mathbf{u}(\tau) \in \mathcal{U}}{\text{minimize}} \quad & j^M\left(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(1)\right) \\ \text{subject to} \quad & \bar{\mathbf{f}}\left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right) - \bar{\mathbf{x}}'(\tau) = \mathbf{0}, \\ & \mathbf{c}\left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right) \leq \mathbf{0}, \\ & \boldsymbol{\phi}\left(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(1)\right) = \mathbf{0}, \\ & \tau \in I^\tau \end{aligned} \tag{5.14}$$

5.2 First-order Necessary Conditions

In the following a reduced problem formulation without state constraints and control bounds is considered for the derivation of the first-order necessary conditions. This simplified case is useful to later show the connection between the discretized form and the continuous time problem in chapter 6. Moreover, the theoretical results are extended in the sequel of this chapter to the case with control constraints and the case with purely state dependent path-constraints, which are relevant for the application considered in this thesis. The optimal control formulation used for the derivation of the first-order necessary conditions is of the form:

$$\begin{aligned} \underset{\mathbf{u}(\tau) \in \mathbb{R}^{n_u}}{\text{minimize}} \quad & j^M\left(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(1)\right) \\ \text{subject to} \quad & \bar{\mathbf{f}}\left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right) - \bar{\mathbf{x}}'(\tau) = \mathbf{0}, \\ & \boldsymbol{\phi}\left(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(1)\right) = \mathbf{0}, \\ & \tau \in I^\tau. \end{aligned} \tag{5.15}$$

First, the augmented cost j^A is introduced

$$j^A := l_0 j^M\left(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(1)\right) + \boldsymbol{\gamma}^T \boldsymbol{\phi}\left(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(1)\right) + \int_0^1 \boldsymbol{\lambda}^T(\tau) \left[\bar{\mathbf{f}}\left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right) - \bar{\mathbf{x}}'(\tau)\right] \, d\tau, \tag{5.16}$$

by adjoining the boundary constraints with multipliers $\boldsymbol{\gamma} \in \mathbb{R}^{n_\phi}$ and the dynamic constraints in differential form using the co-states $\boldsymbol{\lambda}(\tau) \in W_{1,\infty}^{n_x}(I^\tau)$ to the cost with corresponding multiplier $l_0 \in \mathbb{R}$. In the following arguments are omitted for better readability and, in particular, the short notations $\bar{\mathbf{x}}_0 = \bar{\mathbf{x}}(0)$, $\bar{\mathbf{x}}_f = \bar{\mathbf{x}}(1)$, and $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}(0)$, $\boldsymbol{\lambda}_f = \boldsymbol{\lambda}(1)$ are used. The first variation of the augmented cost function (5.16) for the

time normalized form with no path-constraints (5.15) can be written as

$$\begin{aligned}\delta j^A = &\left[l_0 \frac{\partial j^M}{\partial \bar{\mathbf{x}}_0} + \boldsymbol{\gamma}^T \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}_0}\right] \delta \bar{\mathbf{x}}_0 + \left[l_0 \frac{\partial j^M}{\partial \bar{\mathbf{x}}_f} + \boldsymbol{\gamma}^T \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}_f}\right] \delta \bar{\mathbf{x}}_f + \\ &\int_0^1 \left(\boldsymbol{\lambda}^T \frac{\partial \bar{\mathbf{f}}}{\partial \mathbf{u}} \delta \mathbf{u} + \bar{\mathbf{f}}^T \delta \boldsymbol{\lambda} - \left(\bar{\mathbf{x}}'\right)^T \delta \boldsymbol{\lambda} - \boldsymbol{\lambda}^T \delta \bar{\mathbf{x}}' + \boldsymbol{\lambda}^T \frac{\partial \bar{\mathbf{f}}}{\partial \bar{\mathbf{x}}} \delta \bar{\mathbf{x}} \right) \, d\tau.\end{aligned} \tag{5.17}$$

Integration by parts for the term $\boldsymbol{\lambda}^T \delta \bar{\mathbf{x}}'$

$$\int_0^1 \boldsymbol{\lambda}^T \delta \bar{\mathbf{x}}' d\tau = \boldsymbol{\lambda}_f^T \delta \bar{\mathbf{x}}_f - \boldsymbol{\lambda}_0^T \delta \bar{\mathbf{x}}_0 - \int_0^1 \left(\boldsymbol{\lambda}'\right)^T \delta \bar{\mathbf{x}} d\tau, \tag{5.18}$$

yields after insertion of Eq. (5.18) and collection of the terms:

$$\begin{aligned}\delta j^A = &\left[l_0 \frac{\partial j^M}{\partial \bar{\mathbf{x}}_0} + \boldsymbol{\gamma}^T \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}_0} + \boldsymbol{\lambda}_0^T\right] \delta \bar{\mathbf{x}}_0 + \left[l_0 \frac{\partial j^M}{\partial \bar{\mathbf{x}}_f} + \boldsymbol{\gamma}^T \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}_f} - \boldsymbol{\lambda}_f^T\right] \delta \bar{\mathbf{x}}_f + \\ &\int_0^1 \left(\boldsymbol{\lambda}^T \frac{\partial \bar{\mathbf{f}}}{\partial \mathbf{u}} \delta \mathbf{u} + \left[\bar{\mathbf{f}}^T - \left(\bar{\mathbf{x}}'\right)^T\right] \delta \boldsymbol{\lambda} + \left[\boldsymbol{\lambda}^T \frac{\partial \bar{\mathbf{f}}}{\partial \bar{\mathbf{x}}} + \left(\boldsymbol{\lambda}'\right)^T\right] \delta \bar{\mathbf{x}} \right) d\tau\end{aligned} \tag{5.19}$$

Introducing the Hamiltonian

$$\mathcal{H}(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)) := \boldsymbol{\lambda}^T(\tau) \bar{\mathbf{f}} \left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right), \tag{5.20}$$

with

$$\frac{\partial \mathcal{H}}{\partial \boldsymbol{\lambda}} = \bar{\mathbf{f}}^T, \qquad \frac{\partial \mathcal{H}}{\partial \bar{\mathbf{x}}} = \boldsymbol{\lambda}^T \frac{\partial \bar{\mathbf{f}}}{\partial \bar{\mathbf{x}}}, \qquad \frac{\partial \mathcal{H}}{\partial \mathbf{u}} = \boldsymbol{\lambda}^T \frac{\partial \bar{\mathbf{f}}}{\partial \mathbf{u}}, \tag{5.21}$$

Eq. (5.19) can be written as

$$\begin{aligned}\delta j^A = &\left[l_0 \frac{\partial j^M}{\partial \bar{\mathbf{x}}_0} + \boldsymbol{\gamma}^T \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}_0} + \boldsymbol{\lambda}_0^T\right] \delta \bar{\mathbf{x}}_0 + \left[l_0 \frac{\partial j^M}{\partial \bar{\mathbf{x}}_f} + \boldsymbol{\gamma}^T \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}_f} - \boldsymbol{\lambda}_f^T\right] \delta \bar{\mathbf{x}}_f + \\ &\int_0^1 \left(\frac{\partial \mathcal{H}}{\partial \mathbf{u}} \delta \mathbf{u} + \delta \boldsymbol{\lambda}^T \left[\frac{\partial \mathcal{H}}{\partial \boldsymbol{\lambda}} - \left(\bar{\mathbf{x}}'\right)^T\right] + \left[\frac{\partial \mathcal{H}}{\partial \bar{\mathbf{x}}} + \left(\boldsymbol{\lambda}'\right)^T\right] \delta \bar{\mathbf{x}} \right) d\tau.\end{aligned} \tag{5.22}$$

Equating the first variation of the augmented cost to zero, i.e. $\delta j^A = 0$, yields the following conditions for each single term:

- **State Equation** (Variation $\delta \boldsymbol{\lambda}$):

$$\left(\bar{\mathbf{x}}'\right)^T = \frac{\partial \mathcal{H}}{\partial \boldsymbol{\lambda}} \tag{5.23}$$

- **Adjoint Equation** (Variation $\delta \bar{\mathbf{x}}$):

$$\left(\boldsymbol{\lambda}'\right)^T = -\frac{\partial \mathcal{H}}{\partial \bar{\mathbf{x}}} \tag{5.24}$$

- **Transversality Conditions** (Variations $\delta\bar{\mathbf{x}}_0$, $\delta\bar{\mathbf{x}}_f$):

$$\boldsymbol{\lambda}_0^T = -l_0 \frac{\partial j^M}{\partial \bar{\mathbf{x}}_0} - \boldsymbol{\gamma}^T \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}_0}, \tag{5.25}$$

$$\boldsymbol{\lambda}_f^T = l_0 \frac{\partial j^M}{\partial \bar{\mathbf{x}}_f} + \boldsymbol{\gamma}^T \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}_f} \tag{5.26}$$

- **Control Equation** (Variation $\delta\mathbf{u}$):

$$\frac{\partial \mathcal{H}}{\partial \mathbf{u}} = \mathbf{0} \tag{5.27}$$

Based on these conditions, an important result for the Hamiltonian can be derived:

$$\frac{d\mathcal{H}}{d\tau} = 0 \tag{5.28}$$

This result can be obtained by taking the total derivative of the Hamiltonian with respect to time:

$$\frac{d\mathcal{H}}{d\tau} = \frac{\partial \mathcal{H}}{\partial \boldsymbol{\lambda}} \frac{d\boldsymbol{\lambda}}{d\tau} + \frac{\partial \mathcal{H}}{\partial \bar{\mathbf{x}}} \frac{d\bar{\mathbf{x}}}{d\tau} + \frac{\partial \mathcal{H}}{\partial \mathbf{u}} \frac{d\mathbf{u}}{d\tau} \tag{5.29}$$

Inserting the state equation (5.23) and co-state equation (5.24) yields

$$\frac{d\mathcal{H}}{d\tau} = \bar{\mathbf{f}} \frac{d\boldsymbol{\lambda}}{d\tau} - \boldsymbol{\lambda}' \frac{d\bar{\mathbf{x}}}{d\tau} + \frac{\partial \mathcal{H}}{\partial \mathbf{u}} \frac{d\mathbf{u}}{d\tau} = \frac{\partial \mathcal{H}}{\partial \mathbf{u}} \frac{d\mathbf{u}}{d\tau}. \tag{5.30}$$

Together with the control equation $\frac{\partial \mathcal{H}}{\partial \mathbf{u}} = \mathbf{0}$, Eq. (5.28) is readily obtained. Note that according to Eq. (5.27) the control $\mathbf{u}(\tau)$ for the optimal control problem (5.15) is assumed to be uniquely determined by the condition

$$\frac{\partial \mathcal{H}}{\partial \mathbf{u}} = \boldsymbol{\lambda}^T \frac{\partial \bar{\mathbf{f}}}{\partial \mathbf{u}} = \mathbf{0}. \tag{5.31}$$

In case the control is bounded, i.e. $\mathbf{u}(\tau) \in \mathcal{U}$, not all variations $\delta\mathbf{u}$ in the control are feasible. For this case the *Minimum Principle* provides the necessary conditions. It states that for any admissible control $\mathbf{u}(\tau) \in \mathcal{U}$ the Hamiltonian $\mathcal{H}\left(\hat{\bar{\mathbf{x}}}(\tau), \hat{\mathbf{u}}(\tau), \hat{\boldsymbol{\lambda}}(\tau)\right)$ at an optimal solution needs to satisfy

$$\frac{\partial \mathcal{H}}{\partial \mathbf{u}} \left(\mathbf{u}(\tau) - \hat{\mathbf{u}}(\tau)\right) \geq 0, \tag{5.32}$$

together with the transversality conditions (5.25) and (5.26) as well as the co-state equation (5.24). Note that the condition (5.32) implies that the optimal control $\hat{\mathbf{u}}(\tau)$

is an admissible control which minimizes the Hamiltonian

$$\hat{\mathbf{u}}(\tau) = \arg\min_{\mathbf{u}(t)\in\mathcal{U}} \mathcal{H}\left(\hat{\bar{\mathbf{x}}}(\tau), \mathbf{u}(\tau), \hat{\boldsymbol{\lambda}}(\tau)\right). \tag{5.33}$$

5.3 Control Affine Systems

Consider the control affine case [52] for a dynamic constraint of the form

$$\sum_{i=0}^{n_u-1} \bar{\mathbf{f}}_i\left(\bar{\mathbf{x}}(\tau)\right) u_i(\tau) + \bar{\mathbf{f}}_{n_u}\left(\bar{\mathbf{x}}(\tau)\right) - \bar{\mathbf{x}}'(\tau) = \mathbf{0}, \tag{5.34}$$

with state dependent functions $\bar{\mathbf{f}}_i : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}, i = 0, \ldots, n_u$. In this case, the differential form of the time normalized Mayer-type problem without path-constraints reads

$$\begin{aligned} \underset{\mathbf{u}(\tau)\in\mathcal{U}}{\text{minimize}} \quad & j^M\left(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(1)\right) \\ \text{subject to} \quad & \sum_{i=0}^{n_u-1} \bar{\mathbf{f}}_i\left(\bar{\mathbf{x}}(\tau)\right) u_i(\tau) + \bar{\mathbf{f}}_{n_u}\left(\bar{\mathbf{x}}(\tau)\right) - \bar{\mathbf{x}}'(\tau) = \mathbf{0}, \\ & \boldsymbol{\phi}\left(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(1)\right) = \mathbf{0}, \\ & \tau \in I^\tau, \end{aligned} \tag{5.35}$$

and the Hamiltonian is defined as

$$\mathcal{H}(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)) := \boldsymbol{\lambda}^T(\tau) \sum_{i=0}^{n_u-1} \bar{\mathbf{f}}_i\left(\bar{\mathbf{x}}(\tau)\right) u_i(\tau) + \boldsymbol{\lambda}^T(\tau) \bar{\mathbf{f}}_{n_u}\left(\bar{\mathbf{x}}(\tau)\right). \tag{5.36}$$

Introducing the so-called *switching function*

$$\boldsymbol{\mathcal{S}}^T\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right) := \frac{\partial \mathcal{H}}{\partial \mathbf{u}} = \boldsymbol{\lambda}^T\left[\bar{\mathbf{f}}_0\bar{\mathbf{x}}(\tau), \ldots, \bar{\mathbf{f}}_{n_u-1}\bar{\mathbf{x}}(\tau)\right], \tag{5.37}$$

the Hamiltonian can be expressed as

$$\mathcal{H}(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)) = \boldsymbol{\lambda}^T(\tau) \bar{\mathbf{f}}_{n_u}\left(\bar{\mathbf{x}}(\tau)\right) + \boldsymbol{\mathcal{S}}^T\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right) \mathbf{u}(\tau). \tag{5.38}$$

Recall that according to the Minimum Principle the optimal control is an admissible control which minimizes the Hamiltonian (cf. Eq. (5.33)). Therefore, using the definition of the switching function in Eq. (5.37) the product of $\boldsymbol{\mathcal{S}}^T\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right)$ with $\mathbf{u}(\tau) \in \mathcal{U}$ needs to be minimized in order to determine the optimal control. For a scalar, box bounded, control $u(\tau) \in [u_{lb}, u_{ub}] \subset \mathbb{R}, u_{ub} > u_{lb}$ this implies that merely the sign of the

switching function determines the value of the control

$$u(\tau) = \begin{cases} u_{lb}, & \text{if } \mathcal{S}\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right) > 0, \\ u_{ub}, & \text{if } \mathcal{S}\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right) < 0, \\ u_s & \text{otherwise.} \end{cases} \tag{5.39}$$

In case the switching function remains zero, i.e. $\mathcal{S}\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right) = 0$, for a non-instantaneous time interval $\tau \in I^s := [\tau_1, \tau_2], \tau_1 < \tau_2$ the control cannot be determined from the value of the switching function alone and the resulting control u_s is termed *singular*. Observe that the condition $\mathcal{S}\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right) = 0, \tau \in I^s$ implies that all time derivatives of the switching function remain zero on I^s. The i-th time derivative of the switching function can be defined as follows:

$$\begin{aligned} \mathcal{S}^{(0)} &:= \mathcal{S}\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right), \\ \mathcal{S}^{(i)} &:= \frac{\partial \mathcal{S}^{(i-1)}}{\partial \bar{\mathbf{x}}} \frac{d\bar{\mathbf{x}}}{d\tau} + \frac{\partial \mathcal{S}^{(i-1)}}{\partial \boldsymbol{\lambda}} \frac{d\boldsymbol{\lambda}}{d\tau}. \end{aligned} \tag{5.40}$$

If the p-th derivative ($p < \infty$) satisfies

$$\frac{\partial \mathcal{S}^{(p-1)}}{\partial u} = 0, \tag{5.41}$$

$$\frac{\partial \mathcal{S}^{(p)}}{\partial u} \neq 0, \tag{5.42}$$

the switching function $\mathcal{S}^{(p)}$ can be decomposed as follows

$$\mathcal{S}^{(p)} = \mathcal{S}_1^{(p)}\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right) + \mathcal{S}_2^{(p)}\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right) u_s(\tau) = 0. \tag{5.43}$$

and the expression can directly be solved for the singular control $u_s(\tau)$

$$u_s(\tau) = -\frac{\mathcal{S}_1^{(p)}\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right)}{\mathcal{S}_2^{(p)}\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right)}. \tag{5.44}$$

It can be shown that p is an even integer i.e. the control appears first in an even order time derivative of $\mathcal{S}\left(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau)\right)$ [68].

5.4 Purely State Dependent Path-constraints

In the following the case is analyzed in which the optimal control problem is subject to a purely state dependent path-constraint $c\left(\bar{\mathbf{x}}(\tau)\right) \leq 0$ with $c : \mathbb{R}^{n_x} \to \mathbb{R}$ [69, 52]. This case is not trivial from a theoretical perspective as the control does not appear explicitly in the constraint. In particular, the time points where the constraint becomes active,

respectively in-active, are not known a-priori and in general difficult to determine. The following cases can occur when a state dependent constraint $c(\bar{\mathbf{x}}(\tau))$ is part of the optimal control problem [52] (see Fig. 5.1):

- **Free arc**: The constraint is inactive on $\tau \in I^f := [\tau_1, \tau_2], \tau_2 > \tau_1$
- **Boundary arc**: The constraint is active on $\tau \in I^b := [\tau_1, \tau_2], \tau_2 > \tau_1$
- **Contact point**: The constraint is active at a single time point
- **Touch point**: The time derivative of the constraint is continuous at a contact point

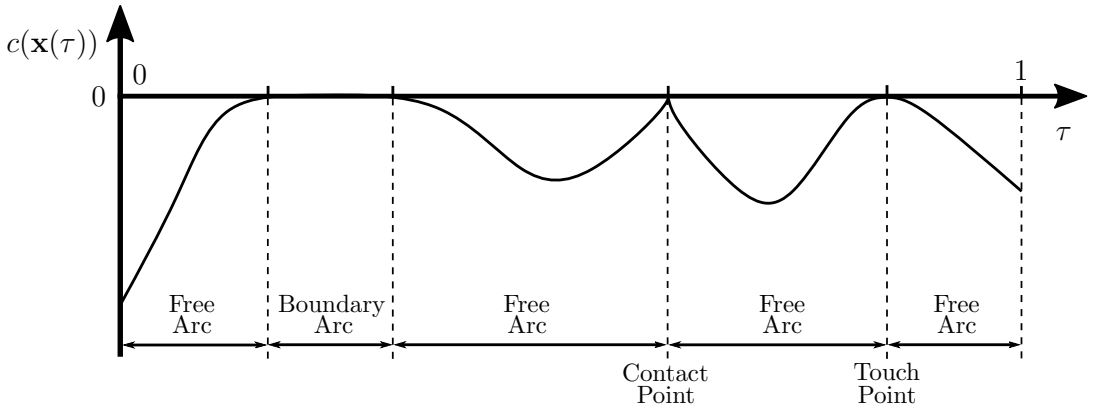


Figure 5.1: Illustration of free and boundary arcs as well as contact and touch points.

In addition, the order of the constraint is defined as the number of time derivatives required until the control appears explicitly. The i-th time derivative of the constraint can be defined recursively

$$c^{(0)} := c\left(\bar{\mathbf{x}}(\tau)\right), \tag{5.45}$$

$$c^{(i)} := \frac{\partial c^{(i-1)}}{\partial \bar{\mathbf{x}}} \frac{d\bar{\mathbf{x}}}{d\tau}, \tag{5.46}$$

and the path-constraint is said to be of order p if the control appears explicitly for the first time in $c^{(p)}, p < \infty$. Two commonly used approaches to treat purely state dependent path-constraints in optimal control theory are the indirect and the direct adjoining approach [69]. In the indirect adjoining approach [70] an index reduction technique is applied to the constraint. This means that the constraint is derived p times until the direct dependence of the constraint on the control is revealed. In the direct adjoining approach [71], which is used in this thesis for the characterization of the worst-case control, the state constraint is directly adjoined to the Hamiltonian using a piecewise continuous multiplier $\mu(\tau)$. This yields the so-called *augmented* Hamiltonian [71, 69]

$$\mathcal{H}^A(\boldsymbol{\lambda}(\tau), \bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)) := \boldsymbol{\lambda}^T(\tau)\bar{\mathbf{f}}\left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right) + \mu(\tau)c\left(\bar{\mathbf{x}}(\tau)\right), \tag{5.47}$$

and the adjoint equation evolves according to

$$(\boldsymbol{\lambda}')^T = -\frac{\partial \mathcal{H}^A}{\partial \bar{\mathbf{x}}}. \tag{5.48}$$

In addition, the following complementarity condition needs to hold:

$$\mu(\tau) \geq 0, \qquad \mu(\tau) c\left(\bar{\mathbf{x}}(\tau)\right) = 0 \tag{5.49}$$

It is important to mention that in this case the co-states may exhibit "jumps". These jumps can take place when the state constraint becomes active, respectively inactive. At a point of discontinuity τ_J of the co-state such a jump is of the form

$$\boldsymbol{\lambda}^T\big|_{\tau=\tau_J^-} = \boldsymbol{\lambda}^T\big|_{\tau=\tau_J^+} + \eta \left.\frac{\partial c}{\partial \bar{\mathbf{x}}}\right|_{\tau=\tau_J}, \tag{5.50}$$

with a non-negative jump height η. In Eq. (5.50) τ_J^- denotes the left sided limit for the jump time point and τ_J^+ the right sided limit, respectively.

Chapter 6

Direct Optimal Control Methods

In this chapter *direct* optimal control methods [52, 50, 40] are discussed. These methods rely on discretization schemes (cf. Sec. 6.1) which transcribe the continuous time optimal control problem introduced in chapter 5 into a discretized form. Popular choices for these transcription schemes are shooting and collocation methods. The discretized problem represents a finite dimensional optimization problem which can be solved using NLP methods (cf. chapter 4). It is important to mention that direct methods, as opposed to most indirect methods, do not require a-priori knowledge regarding the co-state histories or the switching structure of the problem. This makes these approaches particularly useful to obtain approximate solutions in practical applications. In this context, the relationship between the KKT conditions of the discretized problem and the optimality conditions of the continuous time optimal control problem is of interest in order to check the validity of the approximate solution obtained from a direct method (cf. Sec. 6.3).

6.1 Discretization Methods

In the following two important discretization methods, namely shooting and collocation, for direct optimal control methods are presented. Both discretization methods are discussed for the Mayer-type optimal control problem formulation (5.14) in time normalized form which is re-stated here for the reader's convenience:

$$\begin{aligned} \underset{\mathbf{u}(\tau) \in \mathcal{U}}{\text{minimize}} \quad & j^M\left(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(1)\right) \\ \text{subject to} \quad & \bar{\mathbf{f}}\left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right) - \bar{\mathbf{x}}'(\tau) = \mathbf{0}, \\ & \mathbf{c}\left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right) \leq \mathbf{0}, \\ & \boldsymbol{\phi}\left(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(1)\right) = \mathbf{0}, \\ & \tau \in I^\tau \end{aligned} \tag{6.1}$$

Note that this formulation is general in the sense that it also covers the Bolza-type cost function (5.3), non-autonomous systems with a direct dependence on the independent variable, free and fixed final time formulations, as well as parameter dependent systems through the transformations introduced in Sec. 5. It is important to mention that for direct methods no distinction is made between purely state dependent, purely control dependent, or mixed state-control path-constraints.

Both shooting methods and collocation methods are discussed in their segmented form. For this purpose, it is useful to introduce the normalized segment time grid of cardinality $N+1$

$$\mathbb{G}_{N+1}^{S} := \left\{ \tau_S^{[i]} : i = 0, \ldots, N, \tau_S^{[i+1]} > \tau_S^{[i]}, \tau_S^{[0]} = 0, \tau_S^{[N]} = 1 \right\}, \tag{6.2}$$

which defines N time intervals (segments)

$$I^{\tau,[i]} := \left[\tau_S^{[i]}, \tau_S^{[i+1]} \right] \subset \mathbb{R}, \tau_S^{[i+1]} > \tau_S^{[i]},\ i = 0, \ldots, N-1, \tag{6.3}$$

each of length $h^{[i]} = \tau_S^{[i+1]} - \tau_S^{[i]}$. To simplify notation the scaled form of the dynamic equation for each segment, $\bar{\mathbf{f}}^{[i]}\left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right) := h^{[i]}\bar{\mathbf{f}}\left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right)$, is used in most derivations throughout this chapter. The transcription of the continuous problem (5.14) is achieved using a particular state and control discretization for each segment. It is noteworthy that a common extension to the optimal control problem (5.14) are so-called *interior point* constraints. These constraints are imposed at some time point in the interior of the time interval $[0, 1]$. This type of constraint is usually introduced by splitting the problem into so-called *phases*. Each phase may be viewed as an optimal control problem of the type (5.14) and continuity between phases is enforced by additional constraints. Note that for some applications it may not be required to ensure continuity for all quantities. This can occur if different dynamic models with different state vectors are used for each phase or if discontinuities are desired due to the specific problem formulation. As the notion of multi-phase problems does not add to the discussion of discretization methods in the context of this thesis, the following descriptions are limited to a single phase problem of the type (5.14).

6.1.1 Shooting Methods

The idea of shooting methods is based on the solution of several initial value problems for each segment $I^{\tau,[i]}$. This shooting discretization in segmented form is illustrated in Fig. 6.1. Three methods are usually distinguished for shooting methods. The first extreme case, termed *single shooting*, uses a single segment for the discretization of the optimal control problem. This implies that the integration on the whole time interval

of the problem is performed in a single sweep. Despite the appealing simplicity of this method the long integration interval often leads to difficulties in the numerical solution process. In the second case, termed *multiple shooting*, more than one segment is used which reduces the length of the integration intervals. Splitting the problem into several segments is achieved by the introduction of so-called *shooting nodes* and additional defect constraints which link the segments together. Decoupling the problem in this way often facilitates the numerical solution. Another extreme case is the *full discretization* approach. For this method a shooting node is introduced at each step of the integration scheme. This transcription method results in a high dimensional, but also very loosely coupled, parameter optimization problem. The trade-off between the length of the integration intervals and the number of optimization variables, respectively constraints, usually needs to be made according to the specific problem.

State and Control Discretization

The states are discretized at each time point $\tau_S^{[i]}, i = 0, \ldots, N$ of the segment grid defined in Eq. (6.2) by introducing so-called shooting nodes $\bar{\mathbf{x}}^{[i]}, i = 0, \ldots, N$ as optimization variables. The state trajectories on each segment $I^{\tau,[i]}$ are then obtained from the solution of an initial value problem:

$$\bar{\mathbf{x}}(\tau) = \bar{\mathbf{x}}^{[i]} + \int_{\tau_S^{[i]}}^{\tau} \bar{\mathbf{f}}\left(\bar{\mathbf{x}}(s), \mathbf{u}(s)\right)\, ds, \qquad \tau \in I^{\tau,[i]} \tag{6.4}$$

The numerical solution on an integration grid $\mathbb{G}_{N^{[i]}+1}^{I,[i]}$

$$\mathbb{G}_{N^{[i]}+1}^{I,[i]} := \left\{ \tau_{I,j}^{[i]} : j = 0, \ldots, N^{[i]}, \tau_S^{[i]} = \tau_{I,0}^{[i]} \leq \tau_{I,1}^{[i]} \leq \ldots \leq \tau_{I,N^{[i]}-1}^{[i]} \leq \tau_{I,N^{[i]}}^{[i]} = \tau_S^{[i+1]} \right\} \tag{6.5}$$

of cardinality $N^{[i]} + 1$ can be obtained for example using Runge-Kutta one-step methods (cf. Sec. 2.3). Moreover, the control parametrization is often based on B-splines (cf. Sec. 2.6) of degree p with a control break point grid $\mathbb{G}_M^B$ of the form:

$$\mathbb{G}_M^B := \left\{ \tau_{B,j} : j = 0, \ldots, M-1, 0 = \tau_{B,0} \leq \tau_{B,1} \leq \ldots \leq \tau_{B,M-2} \leq \tau_{B,M-1} = 1 \right\}. \tag{6.6}$$

As discussed in Sec. 2.6, B-spline parametrizations are a popular choice due to their local support property, which decreases coupling within the problem. For the B-spline control parametrization of degree p, $n_w = M + p - 1$ variables $w_{q,k}, q = 0, \ldots, n_u - 1, k = 0, \ldots, n_w - 1$ are introduced and the interpolated values for each component $u_q(\tau)$ of the control vector $\mathbf{u}(\tau)$ can be obtained from:

$$u_q(\tau) = \sum_{k=0}^{n_w-1} B_{p,k}(\tau)\, w_{q,k} \tag{6.7}$$

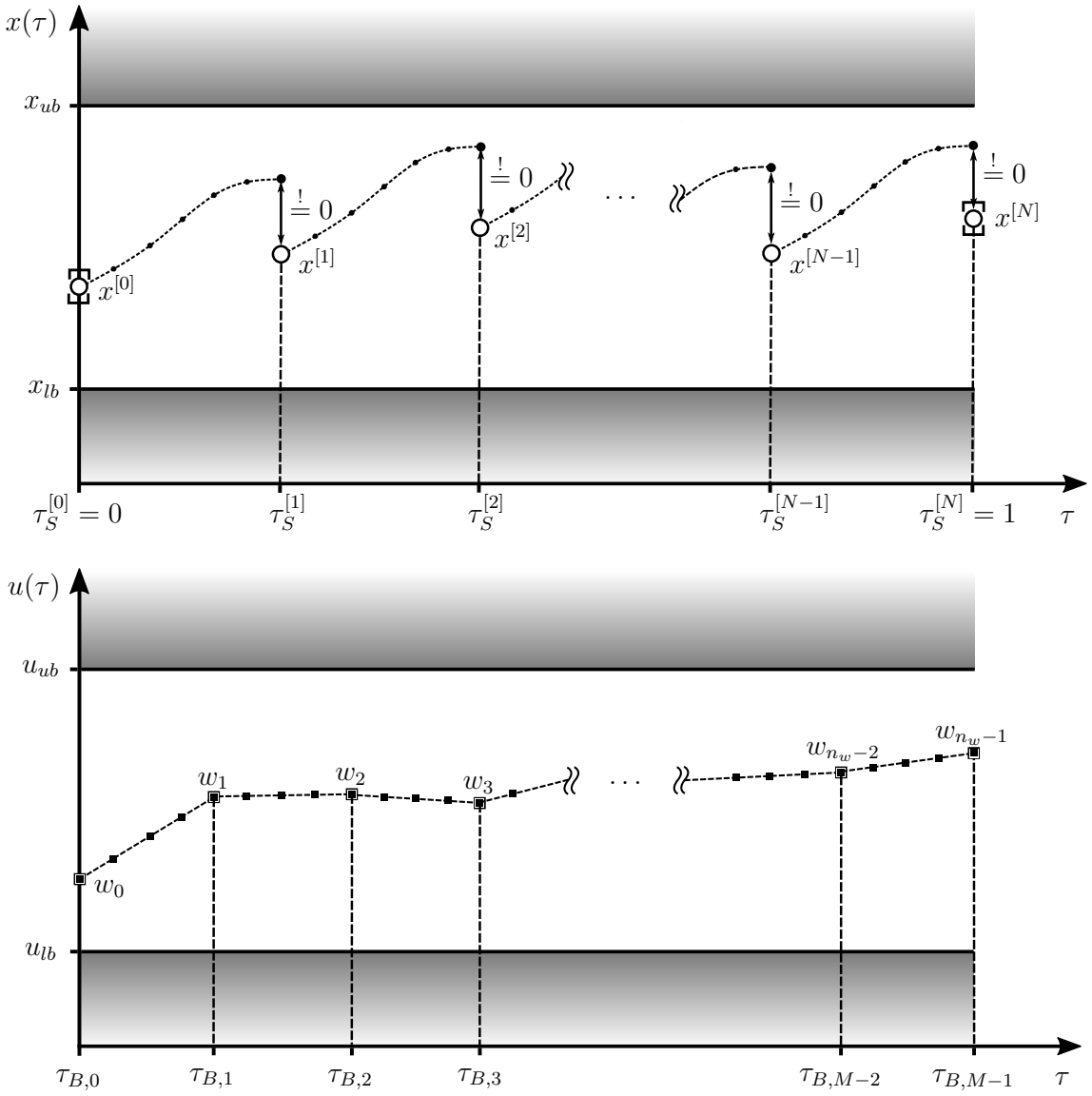


Figure 6.1: Illustration of a segmented shooting discretization for a scalar state $x_{lb} \leq x(\tau) \leq x_{ub}$ and scalar control $u_{lb} \leq u(\tau) \leq u_{ub}$ with control parameters $w_{0,\ldots,n_w-1}$. Initial and final boundary conditions are imposed on the first, respectively final state (brackets) and state continuity is ensured by enforcing equality between the state values at segment boundaries.

Cost and Constraints

The cost function in Mayer-form and the boundary constraints can directly be evaluated using the first and last shooting node, i.e. $j^M\left(\bar{\mathbf{x}}^{[0]}, \bar{\mathbf{x}}^{[N]}\right)$ and $\boldsymbol{\phi}\left(\bar{\mathbf{x}}^{[0]}, \bar{\mathbf{x}}^{[N]}\right)$. Additionally, defect equality constraints $\mathbf{d}^{[i,i+1]}, i = 0, \ldots, N-1$ are introduced

$$\mathbf{d}^{[i,i+1]} := \underbrace{\bar{\mathbf{x}}^{[i]} + \int_{\tau_S^{[i]}}^{\tau_S^{[i+1]}} \bar{\mathbf{f}}\left(\bar{\mathbf{x}}(\tau), \mathbf{u}(\tau)\right)\, d\tau}_{\bar{\mathbf{x}}\left(\tau_S^{[i+1]}\right)} - \bar{\mathbf{x}}^{[i+1]} = \mathbf{0}, \tag{6.8}$$

which enforce state continuity across the respective segment boundaries. The path-constraints

$$\mathbf{c}_k := \mathbf{c}\left(\bar{\mathbf{x}}\left(\tau_{C,k}\right), \mathbf{u}\left(\tau_{C,k}\right)\right) \leq \mathbf{0},\ k = 0, \ldots, P-1 \tag{6.9}$$

are usually imposed on a grid $\mathbb{G}_P^C$ of cardinality P

$$\mathbb{G}_P^C := \left\{\tau_{C,k} : k = 0, \ldots, P-1\right\}, \tag{6.10}$$

which contains a subset of the grid points from the union of all state integration grids (6.5). The state values at these grid points can be obtained from

$$\bar{\mathbf{x}}\left(\tau_{C,k}\right) = \bar{\mathbf{x}}^{[i]} + \int_{\tau_S^{[i]}}^{\tau_{C,k}} \bar{\mathbf{f}}\left(\bar{\mathbf{x}}(s), \mathbf{u}(s)\right)\, ds,\ k \in \mathbb{G}_P^C, \tag{6.11}$$

and the control values $\mathbf{u}\left(\tau_{C,k}\right)$ in Eq. (6.9) are determined using the control interpolation in Eq. (6.7). It is important to mention that typically gradient-based optimization methods are employed for the numerical solution of the discretized problem. These methods require the derivatives of the cost and constraint functions with respect to the optimization variables. As such, sensitivity analysis plays an important role in the context of shooting methods to determine the derivatives of the constraints w.r.t. shooting nodes and control parameters. Based on the state and control discretization introduced in the last section the shooting nodes and control parameters can be collected in the parameter vector $\mathbf{z} \in \mathbb{R}^{n_z}$

$$\mathbf{z} := \left[\left(\bar{\mathbf{x}}^{[0]}\right)^T, \ldots, \left(\bar{\mathbf{x}}^{[N]}\right)^T, \left(\mathbf{w}_0\right)^T, \ldots, \left(\mathbf{w}_{n_w-1}\right)^T\right]^T, \tag{6.12}$$

with $n_z = n_x(N+1) + n_u n_w$. The major challenge for the efficient computation of derivative information originates from the state dependency of the constraint functions. One key concept for obtaining derivative information for the states with respect to the parameter vector $\mathbf{z}$ is to use internal numerical differentiation (cf. Sec. 2.4). The

initial value problem for the sensitivity differential equation of the state sensitivities

$$\mathbf{S}_x(\tau) := \frac{\partial \bar{\mathbf{x}}}{\partial \mathbf{z}}(\tau), \tag{6.13}$$

which needs to be solved on each shooting segment is defined as

$$\mathbf{S}'_x(\tau) = \frac{\partial \bar{\mathbf{f}}}{\partial \bar{\mathbf{x}}}\mathbf{S}_x(\tau) + \frac{\partial \bar{\mathbf{f}}}{\partial \mathbf{u}}\mathbf{S}_u(\tau),\ \tau \in I^{\tau,[i]}, \tag{6.14}$$

$$\mathbf{S}_x\left(\tau_S^{[i]}\right) = \frac{\partial \bar{\mathbf{x}}^{[i]}}{\partial \mathbf{z}}, \tag{6.15}$$

with the control sensitivity matrix

$$\mathbf{S}_u(\tau) := \frac{\partial \mathbf{u}}{\partial \mathbf{z}}(\tau). \tag{6.16}$$

Note that for a B-spline control parametrization the matrix $\frac{\partial \mathbf{u}}{\partial \mathbf{z}}(\tau), j = 0, \ldots, n_u - 1$ merely contains the values of the B-spline basis functions at fixed normalized time points. Furthermore, due to the local support property of the B-spline interpolation method this matrix is usually very sparse. This initial value problem for the sensitivities is solved with the same numerical integration method and the same discretization grid (cf. Eq. 6.5) as the initial value problem for the states. Let $\bar{\mathbf{x}}_j^{[i]} = \mathbf{x}\left(\tau_{I,j}^{[i]}\right)$ and $\mathbf{u}_j^{[i]} = \mathbf{u}\left(\tau_{I,j}^{[i]}\right)$ denote values of the states and controls at time points of the integration grids. Given the state derivatives (6.13) the Jacobian matrices for all quantities in the parameter optimization problem, that is for the cost

$$\frac{\partial j^M}{\partial \mathbf{z}} = \frac{\partial j^M}{\partial \bar{\mathbf{x}}^{[0]}}\frac{\partial \bar{\mathbf{x}}^{[0]}}{\partial \mathbf{z}} + \frac{\partial j^M}{\partial \bar{\mathbf{x}}^{[N]}}\frac{\partial \bar{\mathbf{x}}^{[N]}}{\partial \mathbf{z}}, \tag{6.17}$$

the boundary conditions

$$\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{z}} = \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}^{[0]}}\frac{\partial \bar{\mathbf{x}}^{[0]}}{\partial \mathbf{z}} + \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}^{[N]}}\frac{\partial \bar{\mathbf{x}}^{[N]}}{\partial \mathbf{z}}, \tag{6.18}$$

the defect-constraints

$$\frac{\partial \mathbf{d}^{[i,i+1]}}{\partial \mathbf{z}} = \frac{\partial \mathbf{d}^{[i,i+1]}}{\partial \bar{\mathbf{x}}_{N^{[i]}}^{[i]}}\frac{\partial \bar{\mathbf{x}}_{N^{[i]}}^{[i]}}{\partial \mathbf{z}} + \frac{\partial \mathbf{d}^{[i,i+1]}}{\partial \bar{\mathbf{x}}^{[i+1]}}\frac{\partial \bar{\mathbf{x}}^{[i+1]}}{\partial \mathbf{z}}, \tag{6.19}$$

and the path-constraints

$$\frac{\partial \mathbf{c}_k}{\partial \mathbf{z}} = \frac{\partial \mathbf{c}_k}{\partial \bar{\mathbf{x}}_k}\frac{\partial \bar{\mathbf{x}}_k}{\partial \mathbf{z}} + \frac{\partial \mathbf{c}_k}{\partial \mathbf{u}_k}\frac{\partial \mathbf{u}_k}{\partial \mathbf{z}}, \tag{6.20}$$

with $\bar{\mathbf{x}}_k = \mathbf{x}\left(\tau_{C,k}\right)$ and $\mathbf{u}_k = \mathbf{u}\left(\tau_{C,k}\right)$ can be derived from the chain rule. It is important to note that the sensitivity differential equation only requires a single step in case a full

discretization scheme is chosen. Besides the very high degree of decoupling for this discretization type it is interesting to observe that certain implicit integration schemes can easily be used. For example, the defect constraint of the implicit Backward Euler method

$$\bar{\mathbf{x}}^{[i+1]} - \bar{\mathbf{x}}^{[i]} - \bar{\mathbf{f}}^{[i]}\left(\bar{\mathbf{x}}^{[i+1]}, \mathbf{u}^{[i+1]}\right) = \mathbf{0}, \tag{6.21}$$

can be directly evaluated as all quantities ($\bar{\mathbf{x}}^{[i]}$,$\bar{\mathbf{x}}^{[i+1]}$, and $\mathbf{u}^{[i+1]}$) are available. For implicit Runge-Kutta methods (cf. Sec. 2.3) the stages may also be introduced as additional optimization variables and the stage equations (cf. Eq. (2.25)) can be added as equality constraints [50]. Recall that as discussed in Sec. 2.5 collocation methods can be interpreted as implicit Runge-Kutta methods. This point of view leads to collocation methods discussed in the following section.

6.1.2 Collocation Methods

Different types of collocation methods based on Legendre-Gauss quadrature points (LG, LGRF, LGL) are discussed in the following. As outlined in Sec. 2.3 collocation methods can be formulated in a differential as well as integral form. Both flavors and their application to direct optimal control methods are presented in the sequel.

State and Control Discretization

For segmented collocation schemes (cf. Sec. 2.5) the dynamic equation is discretized on the segments $I^{\tau,[i]}, i = 0, \ldots, N-1$ using a method of order $p^{[i]}$. For each collocation time point $c_j^{[i]}, j = 0, \ldots, p^{[i]} - 1$ one control variable vector $\mathbf{u}_j^{[i]}$ and one state variable vector $\mathbf{s}_j^{[i]}$ (collocated *stages*) is introduced. Additionally, a state variable vector $\mathbf{n}^{[i]}$ (non-collocated *nodes*) is used at all segment grid points, which do not coincide with collocation time points. The collocation methods discussed in the following are summarized in Tab. 6.1. All Lagrange basis functions for the differential and integral methods are assumed to be constructed on a normalized time interval $[0, 1]$ (cf. Sec. 2.5). To simplify notation, let $d_{j,q}^{[i]}$ denote values of the derivatives of the Lagrange basis functions $\frac{d}{d\tau} l_{q+1}^{D,[i]}\left(c_j^{[i]}\right)$ for the differential methods. Similarly, for the integral methods $a_{j,q}^{[i]}$ denotes the integral of the Lagrange basis functions $\int_0^{c_j^{[i]}} l_q^{I,[i]}(\tau)\, d\tau$. In addition, $b_q^{[i]} = \int_0^1 l_q^{I,[i]}(\tau)\, d\tau$ are the corresponding quadrature weights. Using these definitions, the collocation constraints for the differential LG (DLG) collocation method at the collocation time points $c_j^{[i]}$ for all segments and states $k = 0, \ldots, n_x - 1$ can be expressed

Table 6.1: Overview of Legendre-Gauss collocation methods (differential and integral forms).

Method	Description	Collocated end-points (single segment)
DLG-L	Differential Legendre-Gauss (linear state continuity constraint)	-
DLG-N	Differential Legendre-Gauss (nonlinear state continuity constraint)	-
ILG	Integral Legendre-Gauss (nonlinear state continuity constraint)	-
DLGRF	Differential Legendre-Gauss-Radau (flipped)	Final point
ILGRF	Integral Legendre-Gauss-Radau (flipped)	Final point
ILGL	Integral Legendre-Gauss-Lobatto	Initial and final point

as (cf. Eq. (2.54)):

$$\bar{f}_k^{[i]}\left(\mathbf{s}_j^{[i]}, \mathbf{u}_j^{[i]}\right) + \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} n_k^{[i]} - \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} s_{q,k}^{[i]} = 0. \tag{6.22}$$

Note that the interpolating polynomial for the DLG method only includes the initial point of the segment and the collocation time points (cf. Sec. 2.5). This implies that state continuity is not automatically ensured as all LG collocation time points are internal (see Fig. 6.2). As such, an additional constraint must be imposed which ensures a

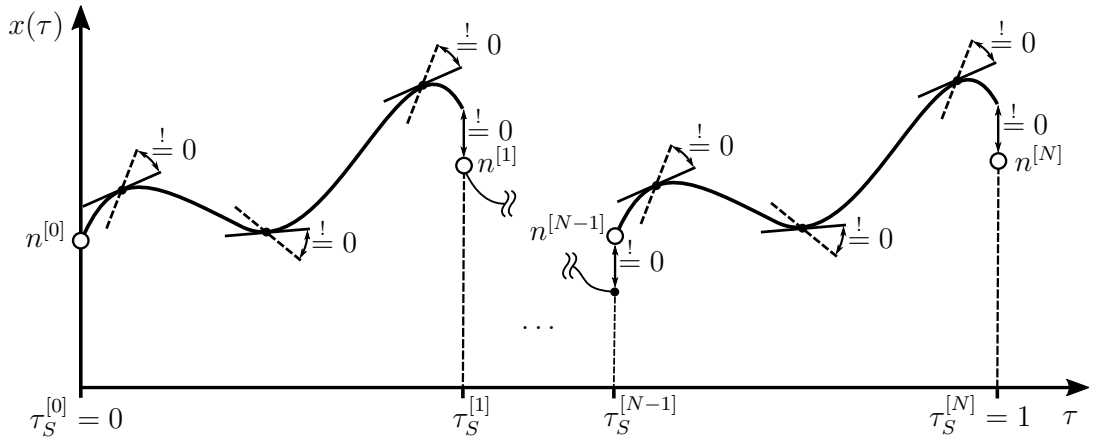


Figure 6.2: Illustration of the differential Legendre Gauss collocation method with collocation constraints strictly in the interior of the segments and state defect constraints at the interior segment boundaries and the final time point.

continuous trajectory of the states. At this point, it is interesting to observe that due to the collocation conditions (6.22) the following relationship holds on each segment

using the quadrature weights $b_j^{[i]}$ (cf. Eq. (2.50)):

$$\begin{aligned} & n_k^{[i]} + \sum_{j=0}^{p^{[i]}-1} b_j^{[i]} \bar{f}_k^{[i]} \left(\mathbf{s}_j^{[i]}, \mathbf{u}_j^{[i]} \right) \\ = \; & n_k^{[i]} + \sum_{j=0}^{p^{[i]}-1} b_j^{[i]} \left[- \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} n_k^{[i]} + \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} s_{q,k}^{[i]} \right] \end{aligned} \tag{6.23}$$

This relation implies that the defect constraint for the differential LG method can be formulated either using a quadrature with the right-hand side function

$$n_k^{[i]} + \sum_{j=0}^{p^{[i]}-1} b_j^{[i]} \bar{f}_k^{[i]} \left(\mathbf{s}_j^{[i]}, \mathbf{u}_j^{[i]} \right) - n_k^{[i+1]} = 0, \tag{6.24}$$

which yields a nonlinear constraint (DLG-N) or the derivative values of the Lagrange basis functions

$$n_k^{[i]} + \sum_{j=0}^{p^{[i]}-1} b_j^{[i]} \left[- \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} n_k^{[i]} + \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} s_{q,k}^{[i]} \right] - n_k^{[i+1]} = 0, \tag{6.25}$$

which yields a linear constraint (DLG-L). Even though both formulations of the state continuity constraint seem equivalent it is clear that they differ considering the nonlinearity of the defect constraint. The linearity of the constraint formulation (6.25) has an interesting implication when considering Newton-type iterations for the solution of the KKT conditions. If the linear constraints are fulfilled at some iterate during the solution process each following Newton step will exactly respect the linearized constraints. Moreover, the linear constraint is fulfilled from the first iterate on in case the initial guess for the state approximation is continuous. This effect is illustrated in Sec. 6.3. Furthermore, note that the constraint Jacobian for (6.25) is constant which implies that it needs to be computed only once before solving the problem.

Next, the differential LGRF (DLGRF) collocation method is discussed. The Lagrange basis for the state approximation is constructed in this case similar to the DLG collocation method based on the initial point of the segment and the collocation time points (cf. Sec. 2.5). For the first segment the initial point is not a collocation point (see Fig. 6.3). Hence, the collocation conditions for the first segment ($i = 0$) are exactly of the form (6.22) with the flipped Radau collocation time points $c_j^{[0]}, j = 0, \ldots, p^{[0]} - 1$. Note however, that the endpoint of the segment is included in the set of DLGRF collocation time points. This means that for the remaining segments $I^{\tau,[i]}, i = 1, \ldots, N-1$ the last stage of the preceding segment coincides with the node at the beginning of the

next segment

$$\mathbf{n}^{[i]} = \mathbf{s}_{p^{[i]}-1}^{[i-1]},\ 0 < i < N. \tag{6.26}$$

Thus, these stages can be re-used for the state approximation (see Fig. 6.3). Note that by re-using the stage at the end of the preceding segment state continuity is automatically ensured by construction of the piecewise polynomials. This implies that a state continuity constraint as for the DLG method does not need to be enforced. Re-using the state variables at internal segment boundaries the collocation conditions of the DLGRF method on the remaining segments $i = 1, \ldots, N-1$ thus read:

$$\bar{f}_k^{[i]}\left(\mathbf{s}_j^{[i]}, \mathbf{u}_j^{[i]}\right) + \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} s_{p^{[i]}-1,k}^{[i-1]} - \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} s_{q,k}^{[i]} = 0,\ 0 < i < N-1 \tag{6.27}$$

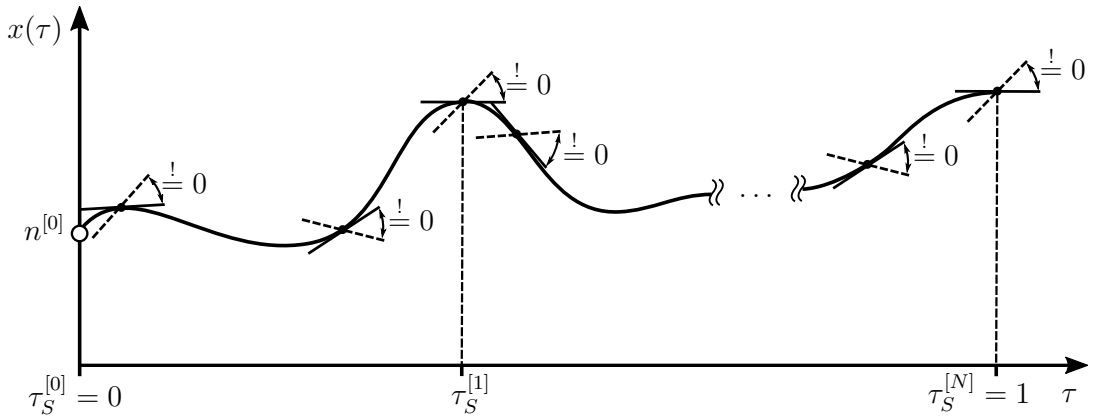


Figure 6.3: Illustration of the differential LGRF collocation method with collocation defect constraints.

For both the DLG and DLGRF there exists an equivalent integral form. In case of the integral LG (ILG) method the collocation conditions (cf. Eq. (2.49)) may be stated as

$$n_k^{[i]} + \sum_{q=0}^{p^{[i]}-1} a_{j,q}^{[i]} \bar{f}_k^{[i]}\left(\mathbf{s}_q^{[i]}, \mathbf{u}_q^{[i]}\right) - s_{j,k}^{[i]} = 0. \tag{6.28}$$

In addition, state continuity must be ensured for the ILG method which can be achieved by introducing constraints of the form (6.24). In case of the integral LGRF (ILGRF) method the collocation conditions for the first segment are of the form (6.28) and for all following segments the last stage of the preceding segment may be re-used as for the differential form:

$$s_{p^{[i]}-1,k}^{[i-1]} + \sum_{q=0}^{p^{[i]}-1} a_{j,q}^{[i]}\, \bar{f}_k^{[i]}\left(\mathbf{s}_q^{[i]}, \mathbf{u}_q^{[i]}\right) - s_{j,k}^{[i]} = 0,\ 0 < i < N-1 \tag{6.29}$$

For integral LGL (ILGL) methods the collocation conditions are also of the form (6.28). However, as shown in Sec. 2.5 the first collocation condition simply implies

$$\mathbf{n}^{[i]} = \mathbf{s}_0^{[i]},\ i = 0, \ldots, N-1. \tag{6.30}$$

Recall that this can be directly seen from the corresponding Butcher Tableau which for all ILGL methods have only zero entries in their first row (cf. Sec. 2.5). Hence, only the second until the last collocation conditions $j = 1, ..., p^{[i]}$ need to be imposed explicitly for the ILGL method

$$s_{0,k}^{[i]} + \sum_{q=0}^{p^{[i]}-1} a_{q,j}^{[i]}\ \bar{f}_k^{[i]}\left(\mathbf{s}_j^{[i]}, \mathbf{u}_j^{[i]}\right) - s_{j,k}^{[i]} = 0. \tag{6.31}$$

Note that all segment grid points correspond to collocation time points as the first and last time point of the segments are included in the set of LGL points. Two popular special cases for ILGL collocation methods derived in Sec. 2.5 which are commonly employed in optimal control applications are the Trapezoidal method ($p^{[i]} = 2$) and the compressed Hermite-Simpson method ($p^{[i]} = 3$) [50]. For these methods only the segment grid needs to be considered for the discretization and all time points represent collocation time points. As all state discretization points correspond to collocation time points and additionally coincide with the segment grid points the state variables may simply be designated with $\bar{\mathbf{x}}^{[i]}, i = 0, \ldots, N$ in the following. For the Trapezoidal collocation scheme one control variable $\mathbf{u}^{[i]}$ is introduced for each segment grid time point $\tau_S^{[i]}, i = 0, \ldots, N$. The defect constraint for this case reads:

$$\bar{\mathbf{x}}^{[i+1]} - \bar{\mathbf{x}}^{[i]} - \frac{\bar{\mathbf{f}}^{[i]}\left(\bar{\mathbf{x}}^{[i]}, \mathbf{u}^{[i]}\right) + \bar{\mathbf{f}}^{[i]}\left(\bar{\mathbf{x}}^{[i+1]}, \mathbf{u}^{[i+1]}\right)}{2} = \mathbf{0} \tag{6.32}$$

For the Hermite-Simpson collocation method an additional control variable $\mathbf{u}^{[i,i+1]}$ at the midpoint of each segment is introduced. In addition, the state at this segment midpoint can be calculated from:

$$\bar{\mathbf{x}}^{[i,i+1]} = \frac{\bar{\mathbf{x}}^{[i]} + \bar{\mathbf{x}}^{[i+1]}}{2} + \frac{\bar{\mathbf{f}}^{[i]}\left(\bar{\mathbf{x}}^{[i]}, \mathbf{u}^{[i]}\right) - \bar{\mathbf{f}}^{[i]}\left(\bar{\mathbf{x}}^{[i+1]}, \mathbf{u}^{[i+1]}\right)}{8} \tag{6.33}$$

The defect constraint for this special case reads:

$$\bar{\mathbf{x}}^{[i+1]} - \bar{\mathbf{x}}^{[i]} - \frac{4\bar{\mathbf{f}}^{[i]}\left(\bar{\mathbf{x}}^{[i]}, \mathbf{u}^{[i]}\right) + 16\bar{\mathbf{f}}^{[i]}\left(\bar{\mathbf{x}}^{[i,i+1]}, \mathbf{u}^{[i,i+1]}\right) + 4\bar{\mathbf{f}}^{[i]}\left(\bar{\mathbf{x}}^{[i+1]}, \mathbf{u}^{[i+1]}\right)}{24} = \mathbf{0} \tag{6.34}$$

Using the affine transformation $\tau^{[i]}(\tau) = \left(\tau - \tau_S^{[i]}\right)/h^{[i]}, \tau \in I^{\tau,[i]}$ the approximating

polynomial is directly obtained as

$$\begin{aligned}\mathbf{s}_T\left(\tau^{[i]}(\tau)\right) = \bar{\mathbf{x}}^{[i]} &+ \left[\tau^{[i]}(\tau)\right]\bar{\mathbf{f}}^{[i]}\left(\bar{\mathbf{x}}^{[i]},\mathbf{u}^{[i]}\right)\\ &+ \left[\tau^{[i]}(\tau)\right]^2\frac{\bar{\mathbf{f}}^{[i]}(\bar{\mathbf{x}}^{[i+1]},\mathbf{u}^{[i+1]}) - \bar{\mathbf{f}}^{[i]}(\bar{\mathbf{x}}^{[i]},\mathbf{u}^{[i]})}{2},\end{aligned} \tag{6.35}$$

by substituting $c_0^{[i]} = 0$ and $c_1^{[i]} = 1$ in Eq. (2.65). For the Hermite-Simpson method the state $\mathbf{s}_H(\tau^{[i]})$ is approximated with a cubic function

$$\begin{aligned}\mathbf{s}_H\left(\tau^{[i]}(\tau)\right) = \bar{\mathbf{x}}^{[i]} &+ \left[\tau^{[i]}(\tau)\right]\bar{\mathbf{f}}^{[i]}\left(\bar{\mathbf{x}}^{[i]},\mathbf{u}^{[i]}\right)\\ &+ \left[\tau^{[i]}(\tau)\right]^2\frac{-3\bar{\mathbf{f}}^{[i]}(\bar{\mathbf{x}}^{[i+1]},\mathbf{u}^{[i+1]}) + 4\bar{\mathbf{f}}^{[i]}(\bar{\mathbf{x}}^{[i,i+1]},\mathbf{u}^{[i,i+1]}) - \bar{\mathbf{f}}^{[i]}(\bar{\mathbf{x}}^{[i+1]},\mathbf{u}^{[i+1]})}{2}\\ &+ \left[\tau^{[i]}(\tau)\right]^3\frac{2\bar{\mathbf{f}}^{[i]}(\bar{\mathbf{x}}^{[i+1]},\mathbf{u}^{[i+1]}) - 4\bar{\mathbf{f}}^{[i]}(\bar{\mathbf{x}}^{[i,i+1]},\mathbf{u}^{[i,i+1]}) + 2\bar{\mathbf{f}}^{[i]}(\bar{\mathbf{x}}^{[i+1]},\mathbf{u}^{[i+1]})}{3},\end{aligned} \tag{6.36}$$

which is obtained from Eq. (2.71) with $c_0^{[i]} = 0$, $c_1^{[i]} = 1/2$ and $c_2^{[i]} = 1$. It is noteworthy that in practical applications the control function is often defined on a coarser grid and using function parametrizations with different smoothness compared to the standard control discretization (e.g. using B-splines).

Cost and Constraints

The collocation methods for the state and control discretization presented in the last section have slight differences in how the state and control discretization is performed. In the following the evaluation of the cost and constraint functions for the optimal control problem (5.14) are illustrated for the ILGL methods (Trapezoidal, Hermite-Simpson) using a B-spline control parametrization. As for the shooting methods presented in Sec. 6.1.1 B-splines are a popular choice for this purpose. The definitions of the break-point grid (6.6) and the control interpolation in Eq. (6.7) directly apply to collocation methods. The optimization parameter vector for this case has the same elements as for shooting methods (cf. Eq. (6.12)) and the Mayer-type cost function as well as the boundary constraints can directly be computed, i.e. $j^M\left(\bar{\mathbf{x}}^{[0]},\bar{\mathbf{x}}^{[N]}\right)$ and $\boldsymbol{\phi}\left(\bar{\mathbf{x}}^{[0]},\bar{\mathbf{x}}^{[N]}\right)$. Moreover, the defect equality constraints $\mathbf{d}^{[i,i+1]}, i = 0,\ldots,N-1$ as defined in Eqs. (6.32) and (6.34)) can be evaluated based on the discretized state variables and the interpolated controls. The values of the control components at segment grid points $u_q^{[i]},\ q = 0,\ldots,n_u-1$ and segment mid-points $u_q^{[i,i+1]},\ q = 0,\ldots,n_u-1$ can be obtained from

$$u_q^{[i]} = \sum_{k=0}^{n_w-1} B_{p,k}\left(\tau_S^{[i]}\right) w_{q,k}, \tag{6.37}$$

$$u_q^{[i,i+1]} = \sum_{k=0}^{n_w-1} B_{p,k}\left(\tau_S^{[i,i+1]}\right) w_{q,k}, \tag{6.38}$$

with $\tau_S^{[i,i+1]} = \frac{1}{2}\left(\tau_S^{[i]} + \tau_S^{[i+1]}\right)$. Furthermore, the path-constraints

$$\mathbf{c}_k := \mathbf{c}\left(\bar{\mathbf{x}}^{[k]}, \mathbf{u}^{[k]}\right) \leq \mathbf{0}, \tag{6.39}$$

are evaluated on the constraint grid $\mathbb{G}_P^C$ of cardinality P

$$\mathbb{G}_P^C := \left\{\tau_{C,k} : \tau_S^{[k]} \in \mathbb{G}_{N+1}^S\right\} \tag{6.40}$$

using the state variables $\bar{\mathbf{x}}^{[k]} = \bar{\mathbf{x}}\left(\tau_{C,k}\right)$ and the interpolated controls $\mathbf{u}^{[k]} = \left(\tau_{C,k}\right)$. Note that the calculation of derivatives for collocation methods is considerably easier compared to shooting methods. This is contributed to the fact that there is no implicit dependence on the optimization variables through an integration procedure. The derivatives of the cost function $\frac{\partial j^M}{\partial \mathbf{z}}$ and boundary constraints $\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{z}}$ are equivalent to shooting methods (cf. Eqs. (6.17) and (6.18)) and the defect constraint derivatives are

$$\begin{aligned}\frac{\partial \mathbf{d}^{[i,i+1]}}{\partial \mathbf{z}} = &\frac{\partial \mathbf{d}^{[i,i+1]}}{\partial \bar{\mathbf{x}}^{[i]}}\frac{\partial \bar{\mathbf{x}}^{[i]}}{\partial \mathbf{z}} + \frac{\partial \mathbf{d}^{[i,i+1]}}{\partial \bar{\mathbf{x}}^{[i+1]}}\frac{\partial \bar{\mathbf{x}}^{[i+1]}}{\partial \mathbf{z}} \\ &+ \frac{\partial \mathbf{d}^{[i,i+1]}}{\partial \mathbf{u}^{[i]}}\frac{\partial \mathbf{u}^{[i]}}{\partial \mathbf{z}} + \frac{\partial \mathbf{d}^{[i,i+1]}}{\partial \mathbf{u}^{[i+1]}}\frac{\partial \mathbf{u}^{[i+1]}}{\partial \mathbf{z}} + \frac{\partial \mathbf{d}^{[i,i+1]}}{\partial \mathbf{u}^{[i,i+1]}}\frac{\partial \mathbf{u}^{[i,i+1]}}{\partial \mathbf{z}}.\end{aligned} \tag{6.41}$$

Note that the last term can be dropped for the Trapezoidal collocation method as no control variable at the segment midpoint is required. Moreover, the derivatives for the path-constraints can be evaluated as follows:

$$\frac{\partial \mathbf{c}_k}{\partial \mathbf{z}} = \frac{\partial \mathbf{c}_k}{\partial \bar{\mathbf{x}}^{[k]}}\frac{\partial \bar{\mathbf{x}}^{[k]}}{\partial \mathbf{z}} + \frac{\partial \mathbf{c}_k}{\partial \mathbf{u}^{[k]}}\frac{\partial \mathbf{u}^{[k]}}{\partial \mathbf{z}} \tag{6.42}$$

6.2 The Transcribed Problem

Using the cost and constraint functions introduced in the last sections, the discretized optimal control problem can be stated in form of a nonlinear parameter optimization problem (cf. chapter 4). On the one hand, the vector-valued function of equality constraints $\mathbf{h} : \mathbb{R}^{n_z} \to \mathbb{R}^{n_h}$ with $n_h = Nn_x + n_\phi$ conatains the boundary constraints and defect constraints for all segments

$$\mathbf{h}\left(\mathbf{z}\right) := \left[\boldsymbol{\phi}^T\left(\bar{\mathbf{x}}^{[0]}, \bar{\mathbf{x}}^{[N]}\right), \left(\mathbf{d}^{[0,1]}\right)^T, \ldots, \left(\mathbf{d}^{N-1,N}\right)^T\right]^T = \mathbf{0}. \tag{6.43}$$

On the other hand, the vector-valued function of inequality constraints $\mathbf{g} : \mathbb{R}^{n_z} \to \mathbb{R}^{n_g}$ contains all $n_g = Pn_c$ discretized path-constraints:

$$\mathbf{g}(\mathbf{z}) := \left[(\mathbf{c}_0)^T, \ldots, (\mathbf{c}_{P-1})^T \right]^T \leq \mathbf{0} \tag{6.44}$$

Thus, the discretized version of the optimal control problem (5.14) using a shooting transcription method or one of the ILGL methods discussed in the last section (Trapezoidal or Hermite-Simpson) can be stated as a nonlinear parameter optimization problem of the form (4.1):

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{R}^{n_z}}{\text{minimize}} \quad & j^M \left(\bar{\mathbf{x}}^{[0]}, \bar{\mathbf{x}}^{[N]} \right), \\ \text{subject to} \quad & \mathbf{h}(\mathbf{z}) = \mathbf{0}, \\ & \mathbf{g}(\mathbf{z}) \leq \mathbf{0} \end{aligned} \tag{6.45}$$

Introducing the multiplier vectors $\boldsymbol{\alpha}$ for the boundary constraints, $\boldsymbol{\kappa}^{[i,i+1]}, i = 0, \ldots, N-1$ for the defect constraints, and $\boldsymbol{\beta}_j^{[i]}, j = 0, \ldots, P-1$ corresponding to the path-inequality constraints, the Lagrangian $\mathcal{L} : \mathbb{R}^{n_z} \times \mathbb{R}^{n_h} \times \mathbb{R}^{n_g} \to \mathbb{R}$ is defined as

$$\mathcal{L}(\mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}) := j^M \left(\bar{\mathbf{x}}^{[0]}, \bar{\mathbf{x}}^{[N]} \right) + \boldsymbol{\lambda}^T \mathbf{h}(\mathbf{z}) + \boldsymbol{\mu}^T \mathbf{g}(\mathbf{z}), \tag{6.46}$$

with the multiplier vector $\boldsymbol{\lambda} \in \mathbb{R}^{n_h}$

$$\boldsymbol{\lambda} := \left[\boldsymbol{\alpha}^T, \left(\boldsymbol{\kappa}^{[0,1]} \right)^T, \ldots, \left(\boldsymbol{\kappa}^{[N-1,N]} \right)^T \right]^T, \tag{6.47}$$

and the multiplier vector $\boldsymbol{\mu} \in \mathbb{R}^{n_g}$

$$\boldsymbol{\mu} := \left[(\boldsymbol{\beta}_0)^T, \ldots, \left(\boldsymbol{\beta}_{P-1} \right)^T \right]^T. \tag{6.48}$$

The necessary and sufficient conditions of a parameter optimization problem in this form are discussed in Sec. 4.2. Moreover, the numerical methods presented in chapter 4 can be employed to obtain a solution of the discretized problem which approximates the solution of the continuous time problem.

6.3 Connection Between the Direct and Indirect Approach

In the following the connection between the necessary conditions from the direct and indirect approach is discussed. In particular, the interpretation of the Lagrange multipliers from the direct approach has been subject to extensive research in the past years in order to establish co-state estimation schemes. For example, in Ref. [52] a discrete

Minimum Principle is presented for the Backward Euler discretization method. It is shown how the multipliers from the discretized problem can be matched with the multipliers from the continuous time problem. Moreover, in Refs. [72, 73] the authors relate the direct and indirect approach for segmented LG and LGRF collocation methods and derive co-state estimates.

Co-state Estimation

The discussion in the following is based on the DLG-L method. This means that the differential LG method is formulated using the linear state continuity constraint (cf. Eq. (6.25)). With the constraints introduced in Sec. 6.1.2 the Lagrangian can be formulated by adjoining the boundary constraints $\phi_r\left(\mathbf{n}^{[0]}, \mathbf{n}^{[N]}\right), r = 0, \ldots, n_\phi - 1$ using multipliers $\alpha_r, r = 0, \ldots, n_\phi - 1$, the collocation constraints (cf. Eq. (6.22)) using multipliers $\nu_{j,k}^{[i]}, j = 0, \ldots, p^{[i]} - 1, k = 0, \ldots, n_x - 1$, and the linear state continuity constraint (cf. Eq. (6.25)) using multipliers $\kappa_k^{[i,i+1]}$ to the Mayer-cost function $j^M\left(\mathbf{n}^{[0]}, \mathbf{n}^{[N]}\right)$:

$$\begin{aligned}\mathcal{L} := & j^M\left(\mathbf{n}^{[0]}, \mathbf{n}^{[N]}\right) + \sum_{r=0}^{n_\phi-1} \alpha_r \phi_r\left(\mathbf{n}^{[0]}, \mathbf{n}^{[N]}\right) + \\ & \sum_{i=0}^{N-1} \sum_{j=0}^{p^{[i]}-1} \sum_{k=0}^{n_x-1} b_j^{[i]} \nu_{j,k}^{[i]} \left[\bar{f}_k^{[i]}\left(\mathbf{s}_j^{[i]}, \mathbf{u}_j^{[i]}\right) + \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} n_k^{[i]} - \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} s_{q,k}^{[i]} \right] + \\ & \sum_{i=0}^{N-1} \sum_{k=0}^{n_x-1} \kappa_k^{[i,i+1]} \left[n_k^{[i]} + \sum_{j=0}^{p^{[i]}-1} b_j^{[i]} \left[- \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} n_k^{[i]} + \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{[i]} s_{q,k}^{[i]} \right] - n_k^{[i+1]} \right]\end{aligned} \tag{6.49}$$

Additionally, the defect equality constraints in the Lagrangian are scaled by the quadrature weights $b_j^{[i]}$. Note that this is valid, as scaling a constraint by the constant factors $b_j^{[i]} > 0$ has no effect on the extreme points, i.e. the location of stationary points of the Lagrangian[1].

The necessary conditions of optimality for constrained optimization problems require the partial derivatives of the Lagrangian with respect to the optimization variables to be zero (cf. Eq. (4.7)). The derivative of the Lagrangian (6.49) with respect to the final node $n_k^{[N]}$ yields

$$\frac{\partial \mathcal{L}}{\partial n_k^{[N]}} = \frac{\partial j^M}{\partial n_k^{[N]}} + \sum_{r=0}^{n_\phi-1} \alpha_r \frac{\partial \phi_r}{\partial n_k^{[N]}} - \kappa_k^{[N-1,N]} = 0, \; k = 0, \ldots, n_x - 1. \tag{6.50}$$

Re-arranging this relation and comparing the equation to the transversality condition

[1] In fact, this is the form one would obtain from a discretization of the augmented performance index (5.16) using the piecewise polynomial approximation corresponding to the collocation method. The quadrature approximation of the integral automatically scales each discretized multiplier by the respective quadrature weight.

(cf. Eq. (5.26)) for $l_0 = 1$ at the final time point

$$\kappa_k^{[N-1,N]} = \frac{\partial j^M}{\partial n_k^{[N]}} + \sum_{r=0}^{n_\phi - 1} \alpha_r \frac{\partial \phi_r}{\partial n_k^{[N]}},\ k = 0, \dots, n_x - 1 \tag{6.51}$$

$$\boldsymbol{\lambda}_f^T = \frac{\partial j^M}{\partial \bar{\mathbf{x}}_f} + \boldsymbol{\gamma}^T \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}_f}, \tag{6.52}$$

it can be deduced that $\boldsymbol{\alpha} \approx \boldsymbol{\gamma}$ and that the last defect multiplier $\boldsymbol{\kappa}^{[N-1,N]}$ appears to match the co-state at the final time-point:

$$\boldsymbol{\kappa}^{[N-1,N]} \approx \boldsymbol{\lambda}_f. \tag{6.53}$$

The partial derivatives of the Lagrangian (6.49) with respect to a stage $s_{j,k}^{[i]}$ in the last segment ($i = N - 1$) yields:

$$\frac{\partial \mathcal{L}}{\partial s_{j,k}^{[i]}} = \sum_{k=0}^{n_x - 1} b_j^{[i]} \nu_{j,k}^{[i]} \frac{\partial \bar{f}_k^{[i]} \left(\mathbf{s}_j^{[i]}, \mathbf{u}_j^{[i]} \right)}{\partial s_{j,k}^{[i]}} + \sum_{q=0}^{p^{[i]}-1} b_q^{[i]} \kappa_k^{[i,i+1]} d_{q,j}^{[i]} - \sum_{q=0}^{p^{[i]}-1} b_q^{[i]} \nu_{q,k}^{[i]} d_{q,j}^{[i]} = 0 \tag{6.54}$$

After division by the weight factor $b_j^{[i]}$, Eq. (6.54) can be written as

$$\sum_{k=0}^{n_x - 1} \nu_{j,k}^{[i]} \frac{\partial \bar{f}_k^{[i]} \left(\mathbf{s}_j^{[i]}, \mathbf{u}_j^{[i]} \right)}{\partial s_{j,k}^{[i]}} + \sum_{q=0}^{p^{[i]}-1} \frac{b_q^{[i]}}{b_j^{[i]}} \kappa_k^{[i,i+1]} d_{q,j}^{[i]} - \sum_{q=0}^{p^{[i]}-1} \frac{b_q^{[i]}}{b_j^{[i]}} \nu_{q,k}^{[i]} d_{q,j}^{[i]} = 0. \tag{6.55}$$

Here, it is useful to introduce the so-called *adjoint basis* for the LG collocation method. Recall that the polynomial approximation with basis functions $l_q^{D,[i]}(\tau), q = 0, \dots, p^{[i]}$ for LG methods includes the initial point of the segment and all collocation points (cf. Sec. 2.5). This basis is commonly termed the *forward basis*. Besides the construction of this forward basis, it is possible to construct a different basis for the collocation method where the polynomial approximation includes the final point of the segment together with the collocation points. This basis is termed the *adjoint basis* and the basis functions are denoted with $l_q^{\dagger,[i]}(\tau), q = 0, \dots, p^{[i]}$. Here, $l_0^{\dagger,[i]}(\tau)$ is the basis function associated with the last point of the segment. Introducing the short notation $d_{j,q}^{\dagger,[i]} = \frac{d}{d\tau} l_{q+1}^{\dagger,[i]} \left(c_j^{[i]} \right)$ as for the forward basis it can be shown that the following relationship holds between the derivative values of the forward and adjoint basis functions [74]:

$$d_{j,q}^{\dagger,[i]} = -\frac{b_q^{[i]}}{b_j^{[i]}} d_{q,j}^{[i]} \tag{6.56}$$

In addition, equivalently to the relation for the forward basis (cf. Eq. (2.54)) the derivatives at collocation time points for a quantity discretized with the adjoint basis can be

represented by

$$-\sum_{q=0}^{p^{[i]}-1} d_{j,q}^{\dagger,[i]} \blacksquare_k^{[i]} + \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{\dagger,[i]} \square_{q,k}^{[i]} = (\square')_{j,k}^{[i]}\,, \tag{6.57}$$

where $\blacksquare_k^{[i]}$ denotes the quantity at the (non-collocated) final point of the segment and $\square_{q,k}^{[i]}$ are quantities corresponding to the collocation time points. Inserting the identity (6.56) between the forward and the adjoint basis and rearranging terms, Eq. (6.55) becomes

$$-\sum_{q=0}^{p^{[i]}-1} d_{j,q}^{\dagger,[i]} \kappa_k^{[i,i+1]} + \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{\dagger,[i]} \nu_{q,k}^{[i]} = -\sum_{k=0}^{n_x-1} \nu_{j,k}^{[i]} \frac{\partial \bar{f}_k^{[i]}\left(\mathbf{s}_j^{[i]}, \mathbf{u}_j^{[i]}\right)}{\partial s_{j,k}^{[i]}}. \tag{6.58}$$

If the multipliers $\nu_{q,k}^{[i]}$ are identified as discrete values $\lambda_{q,k}^{[i]}$ of the co-states at collocation time points, i.e.

$$\nu_{q,k}^{[i]} \approx \lambda_{q,k}^{[i]}, \tag{6.59}$$

Eq. (6.58) for the final segment $i = N-1$ has the form of a collocation condition for the adjoint system (cf. Eq. (5.24)). This can be seen by inserting (6.59) into the left side of Eq. (6.58)

$$-\sum_{q=0}^{p^{[i]}-1} d_{j,q}^{\dagger,[i]} \kappa_k^{[i,i+1]} + \sum_{q=0}^{p^{[i]}-1} d_{j,q}^{\dagger,[i]} \lambda_{q,k}^{[i]} = (\lambda')_{j,k}^{[i]}\,, \tag{6.60}$$

which yields:

$$(\lambda')_{j,k}^{[i]} = -\sum_{k=0}^{n_x-1} \nu_{j,k}^{[i]} \frac{\partial \bar{f}_k^{[i]}\left(\mathbf{s}_j^{[i]}, \mathbf{u}_j^{[i]}\right)}{\partial s_{j,k}^{[i]}} \tag{6.61}$$

The derivative of the Lagrangian with respect to the node $n_k^{[i]}$ is

$$\frac{\partial \mathcal{L}}{\partial n_k^{[i]}} = \kappa_k^{[i,i+1]} - \sum_{j=0}^{p^{[i]}-1} \sum_{q=0}^{p^{[i]}-1} b_j^{[i]} d_{j,q}^{[i]} \kappa_k^{[i,i+1]} + \sum_{j=0}^{p^{[i]}-1} \sum_{q=0}^{p^{[i]}-1} b_j^{[i]} d_{j,q}^{[i]} \nu_{j,k}^{[i]} - \kappa_k^{[i-1,i]} = 0. \tag{6.62}$$

Inserting the identity (6.56) the following form is obtained:

$$\kappa_k^{[i,i+1]} + \sum_{j=0}^{p^{[i]}-1} \sum_{q=0}^{p^{[i]}-1} b_q^{[i]} d_{q,j}^{\dagger,[i]} \kappa_k^{[i,i+1]} - \sum_{j=0}^{p^{[i]}-1} \sum_{q=0}^{p^{[i]}-1} b_q^{[i]} d_{q,j}^{\dagger,[i]} \nu_{j,k}^{[i]} - \kappa_k^{[i-1,i]} = 0 \tag{6.63}$$

Under the assumption that the multipliers $\kappa_k^{[i,i+1]}$ and $\nu_{j,k}^{[i]}$ can be identified as discrete

values of the co-states the two double sums can be written for $i = N-1$ as:

$$\sum_{j=0}^{p^{[i]}-1}\sum_{q=0}^{p^{[i]}-1} b_q^{[i]} d_{q,j}^{\dagger,[i]} \kappa_k^{[i,i+1]} - \sum_{j=0}^{p^{[i]}-1}\sum_{q=0}^{p^{[i]}-1} b_q^{[i]} d_{q,j}^{\dagger,[i]} \lambda_{j,k}^{[i]} = \tag{6.64}$$

$$-\sum_{q=0}^{p^{[i]}-1} b_q^{[i]} \left[-\sum_{j=0}^{p^{[i]}-1} d_{q,j}^{\dagger,[i]} \kappa_k^{[i,i+1]} + \sum_{j=0}^{p^{[i]}-1} d_{q,j}^{\dagger,[i]} \lambda_{j,k}^{[i]} \right] = \tag{6.65}$$

$$-\sum_{q=0}^{p^{[i]}-1} b_q^{[i]} \left(\lambda'\right)_{q,k}^{[i]} = -\Delta_{\lambda,k}^{[i]} \tag{6.66}$$

Inserting this relationship in Eq. (6.63)

$$\kappa_k^{[i-1,i]} = \kappa_k^{[i,i+1]} - \Delta_{\lambda,k}^{[i]}, \tag{6.67}$$

it can be seen that, if $\kappa_k^{[i,i+1]}$ represents the value of the co-state at the final point of the last segment $i = N-1$ and $\Delta_{\lambda,k}^{[i]}$ is the change of the co-state value within this segment, the defect multipliers $\kappa_k^{[i-1,i]}$ represent the values of the co-state at the initial point of this segment. The arguments presented for matching the multipliers between the necessary conditions of the direct and indirect approach for the last segment are valid for all other segments. The derivations starting from Eq. (6.54) just need to be repeated for $i = N-2, \ldots, 0$ in exactly the same manner. Note that the equation related to the defect multipliers $\kappa_k^{[i-1,i]}$ (cf. Eq. (6.67)) essentially propagates the information regarding the value of the co-state across the segments, whereas Eq. (6.58) (respectively Eq. (6.61)) ensures that the co-state differential equation is fulfilled at the collocation time points. Finally, it remains to investigate if the derivation yields a consistent result for the transversality condition at the initial point (cf. Eq. (5.25)). The derivative of the Lagrangian with respect to the initial node $n_k^{[0]}$

$$\frac{\partial \mathcal{L}}{\partial n_k^{[0]}} = \frac{\partial j^M}{\partial n_k^{[0]}} + \sum_{r=0}^{n_\phi-1} \alpha_r \frac{\partial \phi_r}{\partial n_k^{[0]}} + \kappa_k^{[0,1]} - \sum_{j=0}^{p^{[0]}-1}\sum_{q=0}^{p^{[0]}-1} b_j^{[0]} d_{j,q}^{[0]} \kappa_k^{[0,1]} + \sum_{j=0}^{p^{[0]}-1}\sum_{q=0}^{p^{[0]}-1} b_j^{[0]} d_{j,q}^{[0]} \nu_{j,k}^{[0]} = 0, \tag{6.68}$$

can be re-arranged and written using the relation for the adjoint basis:

$$\kappa_k^{[0,1]} + \sum_{j=0}^{p^{[0]}-1}\sum_{q=0}^{p^{[0]}-1} b_q^{[0]} d_{j,q}^{\dagger,[0]} \kappa_k^{[0,1]} - \sum_{j=0}^{p^{[0]}-1}\sum_{q=0}^{p^{[0]}-1} b_q^{[0]} d_{j,q}^{\dagger,[0]} \nu_{j,k}^{[0]} = -\frac{\partial j^M}{\partial n_k^{[0]}} - \sum_{r=0}^{n_\phi-1} \alpha_r \frac{\partial \phi_r}{\partial n_k^{[0]}} \tag{6.69}$$

With the multiplier $\kappa_k^{[0,1]}$ representing the co-state at the beginning of the second segment and exploiting the relation (6.67) the result can be compared to the initial transver-

sality condition (cf. Eq. (5.25)) for $l_0 = 1$:

$$\kappa_k^{[0,1]} - \Delta_{\lambda,k}^{[0]} = -\frac{\partial j^M}{\partial n_k^{[0]}} - \sum_{r=0}^{n_\phi - 1} \alpha_r \frac{\partial \phi_r}{\partial n_k^{[0]}},\ k = 0, \ldots, n_x - 1 \tag{6.70}$$

$$\boldsymbol{\lambda}_0^T = -\frac{\partial j^M}{\partial \bar{\mathbf{x}}_0} - \boldsymbol{\gamma}^T \frac{\partial \boldsymbol{\phi}}{\partial \bar{\mathbf{x}}_0} \tag{6.71}$$

Note that exactly the same form is obtained, i.e. the derivations in this section are consistent regarding the transversality condition at the initial time point. Furthermore, the partial derivative of the Lagrangian with respect to the control vector $\mathbf{u}_j^{[i]}$ yields

$$\frac{\partial \mathcal{L}}{\partial \mathbf{u}_j^{[i]}} = b_j^{[i]} \left(\boldsymbol{\nu}_j^{[i]} \right)^T \frac{\partial \bar{\mathbf{f}}^{[i]}}{\partial \mathbf{u}_j^{[i]}} = \mathbf{0}. \tag{6.72}$$

After division by the weight factor $b_j^{[i]}$ and with the identified relation $\boldsymbol{\nu}_j^{[i]} \approx \boldsymbol{\lambda}_j^{[i]}$ a discretized form of the control equation (5.27) for each collocation point is obtained:

$$\left(\boldsymbol{\lambda}_j^{[i]} \right)^T \frac{\partial \bar{\mathbf{f}}^{[i]}}{\partial \mathbf{u}_j^{[i]}} = \mathbf{0} \tag{6.73}$$

Summarizing, for all segments $i = 0, \ldots, N-1$ the state continuity multipliers $\boldsymbol{\kappa}^{[i,i+1]}$ and the collocation defect multipliers $\boldsymbol{\nu}_j^{[i]}, j = 0, \ldots, p^{[i]} - 1$ can be interpreted as the co-states from the necessary conditions of the continuous problem. In addition, the control equation and the co-state equation can be identified in a discretized form at the collocation time points.

Numerical Examples

The DLG-L, DLG-N, ILG, and DLGRF collocation methods are illustrated and discussed using an example problem from Ref. [48]. This problem is stated in Mayer form and involves a scalar state $x(t)$ and a scalar control $u(t)$:

$$\begin{aligned} \underset{u(t)}{\text{minimize}} \quad & -x(t_f) \\ \text{subject to} \quad & \dot{x}(t) = -x(t) + x(t)u(t) - u^2(t), \\ & x(0) = 1, \\ & t_f = 5 \end{aligned} \tag{6.74}$$

The optimal solutions for the state $\hat{x}(t)$, corresponding co-state $\hat{\lambda}(t)$, and the control $\hat{u}(t)$ are:

$$\begin{aligned} \hat{x}(t) &:= \frac{4}{1+3e^t}, \\ \hat{\lambda}(t) &:= \frac{-e^{\left(2\ln\left(1+3e^t\right)-t\right)}}{e^{-5}+6+9e^5}, \\ \hat{u}(t) &:= \frac{\hat{x}(t)}{2} \end{aligned} \tag{6.75}$$

In order to transcribe this continuous time problem into a parameter optimization problem the DLG-L, DLG-N, ILG, and DLGRF methods are used. Moreover, for the numerical solution of the resulting NLP the SQP solver SNOPT [61] is employed with default settings. In all graphs the solutions for the state and co-state trajectories are depicted with dense output, i.e. using the interpolating polynomials. For all methods two segments with ten collocation points each are used. The analytical solution and the solution obtained from the DLG-L method is shown in Fig. 6.4. Note that the numerical solution closely matches the analytical solution for all three quantities. Moreover, the deviations between the polynomial approximation of all collocation schemes (DLG-L, DLG-N, ILG, DLGRF) for the state and co-state are depicted in Fig. 6.5. It can be seen that using relatively few discretization points the optimal solution is accurately approximated for all methods.

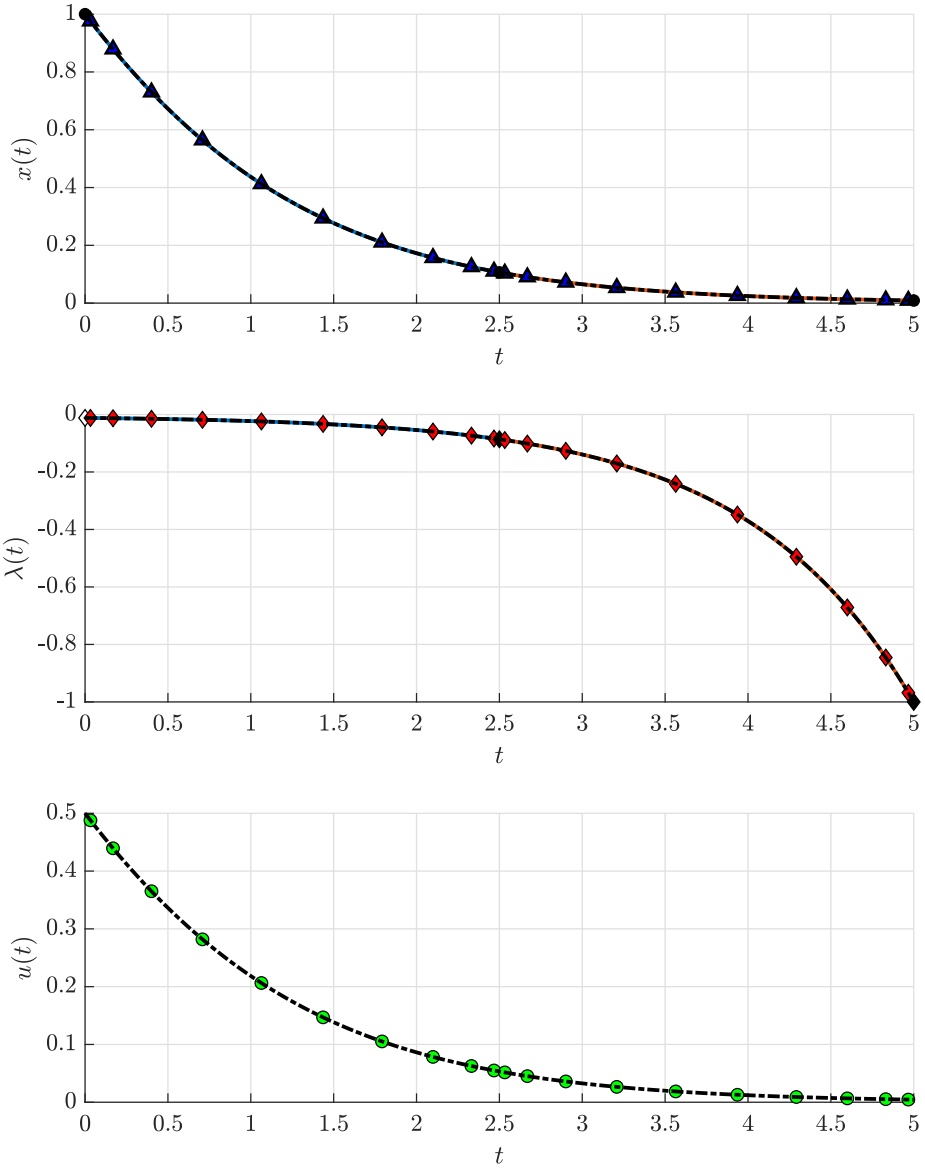


Figure 6.4: Comparison between the analytical and numerical solution (DLG-L) of the example problem. The analytical solution is represented by a black dash-dotted line. Dense outputs for the state and co-state trajectories are shown in different colors for different segments. For the state trajectory blue triangles represent stage variables and black dots show the nodes at non-collocated time points. For the co-state trajectory black markers represent the co-state estimates from the state continuity constraint multipliers, red markers the co-state estimates at collocation points, and the white marker depicts the extrapolated co-state estimate at the beginning of the time interval.

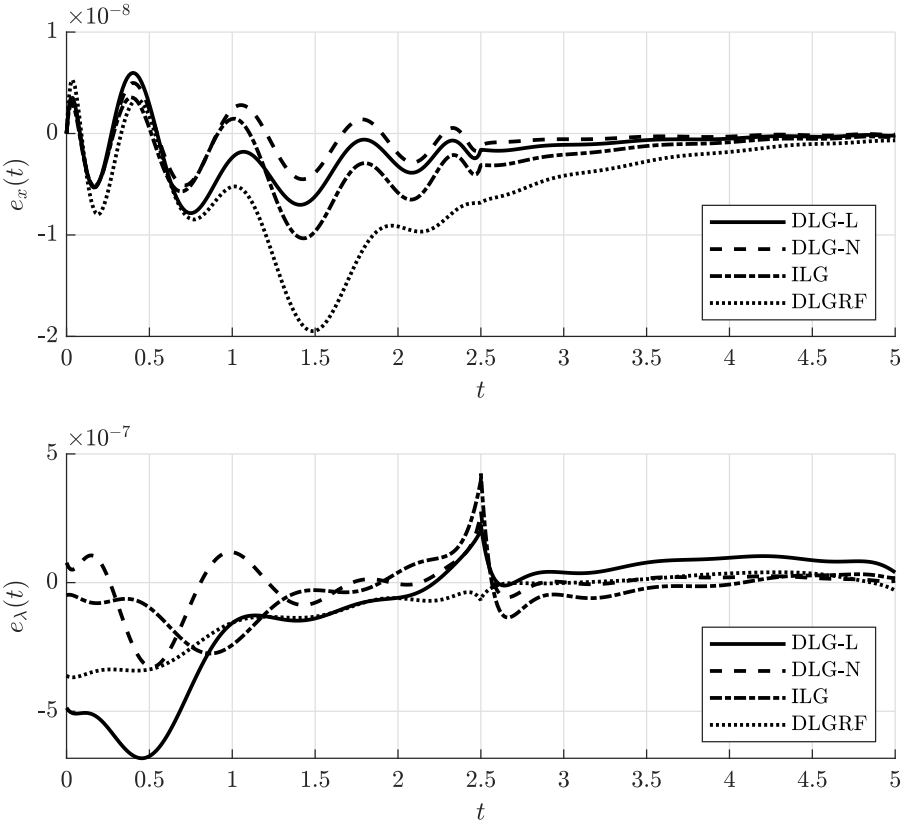


Figure 6.5: Comparison of the errors $e_x(t)$ and $e_\lambda(t)$ between the dense outputs of state $x(t)$ and co-state $\lambda(t)$ from the numerical solution with the analytical solution using the DLG-L, DLG-N, ILG, and DLGRF collocation methods.

Furthermore, Fig. 6.6 presents a comparison between the DLG-L and DLG-N method regarding the state continuity constraint value at the end of the first segment. The first 26 major iterations are shown. The optimization parameter vector is initialized with all components set to one which implies that the linear defect constraint is trivially satisfied at the initial point. This is not the case for the nonlinear defect constraint. As can be seen in Fig. 6.6 the linear constraint is not violated in any of the iterations as expected from the exactness of the Newton-step for linear constraints.

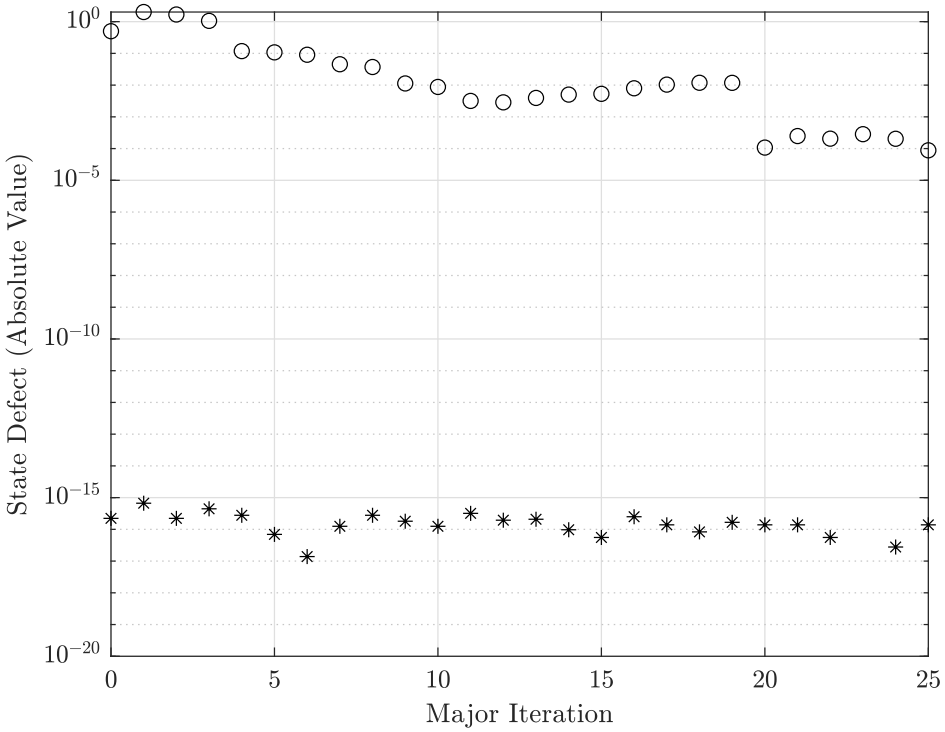


Figure 6.6: Comparison of the constraint values of the state continuity constraint between the first and the second segment for the Legendre Gauss method with linear defect constraint (star markers) and with nonlinear defect constraint (round markers) for the first 26 major iterations.

Chapter 7

Models for Optimal Control Based Clearance of Flight Control Systems

This chapter discusses the basic modeling concepts for the optimal control based clearance approach. First, a system classification for the clearance task is proposed which considers the nonlinearity of the system and the presence of internal bounds (cf. Sec. 7.1). The remainder of this chapter describes models for aircraft closed-loop systems which are under consideration regarding the approaches developed in this thesis. First, basic flight mechanical quantities for aircraft plant models are introduced (cf. Sec. 7.2.1). Following, the modeling concept of the optimal control based approach regarding actuator limits is discussed from a phase plane perspective (cf. Sec. 7.2.2). Finally, a basic overview of flight control systems is given and relevant clearance criteria for worst-case pilot inputs are summarized (cf. Sec. 7.2.3).

7.1 Modeling Philosophy and System Classifications

The layout of the closed-loop system considered for the clearance task is depicted in Fig. 7.1. Basic building blocks of this system are models for the aircraft plant, the servomechanism, sensors, and the controller. In addition, external influences considered for testing the flight control law are pilot commands and disturbances, such as wind. Moreover, the system may depend on parameters with known limits or subject to a distribution. Furthermore, internal bounds of the servomechanism such as rate and deflection limits are considered for the clearance task. The modeling philosophy for optimal control based worst-case analysis with respect to pilot commands and disturbances follows the categorization for pilot induced oscillations (PIO) from Ref. [32]. As such, the proposed classification considers the nonlinearity of the closed-loop model on the one side and the presence of internal limits of the actuators on the other.

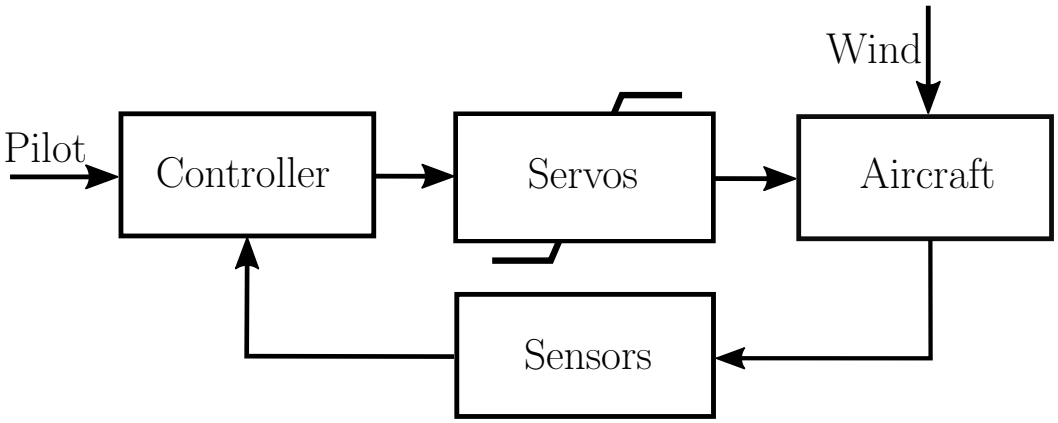


Figure 7.1: Illustration of the basic layout of the closed-loop system with actuator limits and external inputs (pilot, wind).

The following categories are introduced to classify the clearance task under the optimal control based approach investigated in this thesis:

- **Category I**: Completely linear system analysis. The dynamic model is linear and only subject to limits for the external pilot commands and disturbance inputs. Nominal and parameter dependent systems are considered.
- **Category II**: Quasi-linear system analysis with nonlinear influences originating from rate and deflection limits introduced from the servomechanism of the system. The pilot commands and disturbances are bounded and the limits for the servomechanism are modeled using state constraints. Nominal and parameter dependent systems are considered.
- **Category III**: Nonlinear system analysis. All other cases which cannot be modeled as a Cat I or Cat II model are collected in this category. This category is further refined based on the respective solution methodologies (non-intrusive, intrusive, hybrid, cf. chapter 9).

Note that the investigation of PIO effects is not considered within the scope of this thesis and merely the modeling categories are defined similar to Ref. [32]. Nevertheless, it is interesting that the excitement of oscillatory modes due to pilot-vehicle interactions naturally arises for certain types of clearance tasks.

7.2 Aircraft Closed-loop Systems

7.2.1 Flight Mechanical Model

In the following a basic overview of the flight mechanical quantities [75, 76, 35] for the aircraft plant model is given. First, it is useful to introduce the reference frames illus-

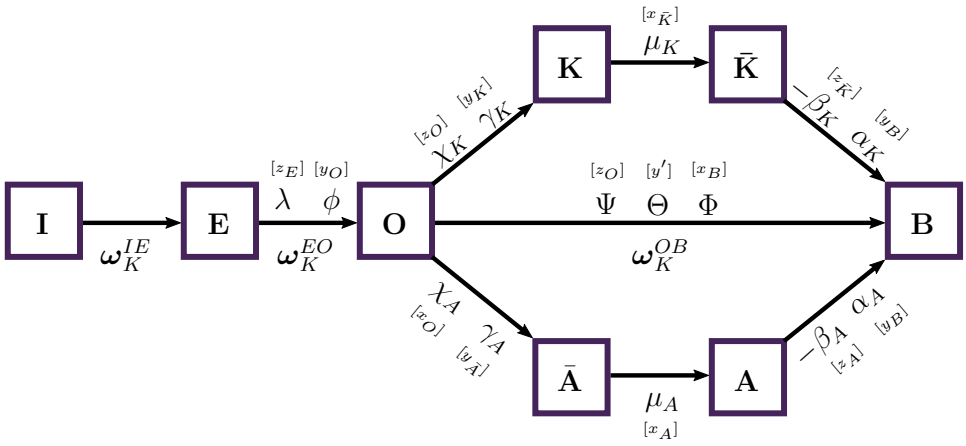


Figure 7.2: Relevant reference frames and their relationship including angles with rotation axis and angular rotation rates between the I-, E-, O- and B-frame.

trated in Fig. 7.2. The earth centered, inertial (ECI) frame (index I) is a non-rotating, inertial frame with its origin in the center of the earth. The x-axis of the ECI-frame is oriented towards a fix-point (*vernal equinox*) and the z-axis points in the direction of the earth's rotation axis towards the north pole. The earth centered, earth fixed (ECEF) frame (index E) also has its origin in the center of the earth and shares the z-axis with the I-frame. The kinematic (K) angular velocity $\boldsymbol{\omega}_K^{IE} \in \mathbb{R}^3$ between the I- and the E-frame is approximately

$$\boldsymbol{\omega}_K^{IE} \approx \begin{bmatrix} 0 \\ 0 \\ 2\pi/24h \end{bmatrix}. \tag{7.1}$$

All velocity vectors at a reference point R, that are the kinematic velocity $\left(\mathbf{V}_K^R\right)^E \in \mathbb{R}^3$, the aerodynamic velocity $\left(\mathbf{V}_A^R\right)^E \in \mathbb{R}^3$, and the wind velocity $\left(\mathbf{V}_W^R\right)^E \in \mathbb{R}^3$ are described relative to the earth (E), which is a natural description for the velocity. Another important frame for aircraft orientation, the so-called North-East-Down (NED) frame or Orientation frame (index O), is centered at the reference point R of the aircraft and can be obtained from the E-frame by a rotation around the z_E-axis with angle λ (geodetic longitude) and a rotation around the y_O-axis with the negative angle of geodetic latitude ($\phi + \pi/2$). The z_O-axis is pointing downwards, perpendicular to the

surface of the earth. The x_O-axis and y_O-axis are oriented towards the north and east. The O-frame is rotating relative to the E-frame with the so-called kinematic (K) transport rate $\boldsymbol{\omega}_K^{EO} \in \mathbb{R}^3$.

The kinematic frame (index K) represents a trajectory frame with the x_K-axis oriented in the direction of the kinematic velocity $\left(\mathbf{V}_K^R\right)^E$. This frame is obtained from the NED-frame by a rotation around the z_O-axis with the kinematic course angle χ_K, and a rotation around the y_K-axis with the kinematic inclination angle γ_K. Rotating the kinematic frame around the x_K-axis with the bank angle μ_K yields the rotated kinematic frame (index $\bar{K}$). From the rotated kinematic frame the body fixed frame (index B) is obtained by a rotation around the $z_{\bar{K}}$-axis with the negative kinematic Angle-of-Sideslip (AoS) β_K and a rotation around the y_B-axis with the Angle-of-Attack (AoA) α_K. From the O-frame, the B-frame is obtained through a rotation around the z_O-axis with yaw angle Ψ, a rotation around the intermediate y-axis with pitch angle Θ, and a rotation around the x_B-axis with roll angle Φ. The aerodynamic frame (index A) is defined with its x_A-axis pointing in the direction of the aerodynamic velocity vector $\left(\mathbf{V}_A^R\right)^E$. The rotated aerodynamic frame (index $\bar{A}$) is obtained from the A-frame through a rotation around the x_A-axis with the negative aerodynamic bank angle μ_A. Note that if the wind velocity is zero, i.e. $\left(\mathbf{V}_W^R\right)^E = \mathbf{0}$, the kinematic velocity vector $\left(\mathbf{V}_K^R\right)^E$ is equal to the aerodynamic velocity vector $\left(\mathbf{V}_A^R\right)^E$ according to the vectorial wind equation:

$$\left(\mathbf{V}_A^R\right)^E = \left(\mathbf{V}_K^R\right)^E - \left(\mathbf{V}_W^R\right)^E = \left(\mathbf{V}_K^R\right)^E \tag{7.2}$$

For this case, the A-frame coincides with the $\bar{K}$-frame and the $\bar{A}$-frame coincides with the K-frame. In the following all rotation matrices are described based on the elementary rotation matrices for a generic angle δ around the respective x-axis, y-axis, and z-axis

$$\mathbf{R}_x(\delta) := \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\delta & -s_\delta \\ 0 & s_\delta & c_\delta \end{bmatrix}, \qquad \mathbf{R}_y(\delta) := \begin{bmatrix} c_\delta & 0 & s_\delta \\ 0 & 1 & 0 \\ -s_\delta & 0 & c_\delta \end{bmatrix}, \qquad \mathbf{R}_z(\delta) := \begin{bmatrix} c_\delta & -s_\delta & 0 \\ s_\delta & c_\delta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{7.3}$$

with $c_\delta = \cos(\delta)$ and $s_\delta = \sin(\delta)$. In order to introduce the basic flight mechanical quantities which are of interest in the context of this thesis it is sufficient to discuss the aircraft dynamic equations with the following simplifying assumptions:

- The aircraft is modeled as a rigid-body.
- The change of mass of the aircraft in the dynamic equations is assumed quasi-stationary
- Effects of earth rotation are neglected ($\boldsymbol{\omega}_K^{IE} := \mathbf{0}$).

- Effects of earth curvature are neglected ($\boldsymbol{\omega}_K^{EO} := \mathbf{0}$).

Additionally, the reference point R is set to the center of gravity ($R := G$). The states describing the motion of the rigid-body aircraft model are position states, attitude states, as well as states for the translational and rotational velocity dynamics. There are many possibilities to describe the motion of the aircraft, due to the non-uniqueness of a state space representation. In the following the position states λ^G, ϕ^G and the geodetic height h^G describe the position of the aircraft's center of gravity (G) using WGS84 coordinates. The description of the aircraft's attitude relative to the O-frame can be based on the Euler angles Ψ, Θ, and Φ. The kinematic (K) rotational velocity vector $\left(\boldsymbol{\omega}_K^{OB}\right)_B \in \mathbb{R}^3$ describing the rotational velocity vector of the B-frame w.r.t. the O-frame denoted in the B-frame has the components

$$\left(\boldsymbol{\omega}_K^{OB}\right)_B := \begin{bmatrix} p_K \\ q_K \\ r_K \end{bmatrix}_B . \tag{7.4}$$

Furthermore, the translational states can be described by the velocity vector $\left(\mathbf{V}_K^G\right)_B^E$ denoted in the B-frame:

$$\left(\mathbf{V}_K^G\right)_B^E := \begin{bmatrix} u_K^G \\ v_K^G \\ w_K^G \end{bmatrix}_B^E \tag{7.5}$$

As such, the state vector for the rigid-body model has twelve components:

$$\mathbf{x} := \left[\lambda^G, \phi^G, h^G, \left(u_K^G\right)_B^E, \left(v_K^G\right)_B^E, \left(w_K^G\right)_B^E, \Psi, \Theta, \Phi, p_K, q_K, r_K\right]^T \tag{7.6}$$

Other relevant quantities may be derived from these states. For example, another set of states describing the translational dynamics, which is useful in the context of this thesis, contains the absolute velocity V_K^G, the kinematic AoA α_K, and the kinematic AoS β_K which can be obtained from the rigid-body velocities as follows:

$$V_K^G = \left\| \left(\mathbf{V}_K^G\right)_B^E \right\|_2 , \tag{7.7}$$

$$\alpha_K = \arctan\left(\frac{\left(w_K^G\right)_B^E}{\left(u_K^G\right)_B^E}\right), \tag{7.8}$$

$$\beta_K = \arctan\left(\frac{\left(v_K^G\right)_B^E}{\sqrt{\left[\left(u_K^G\right)_B^E\right]^2 + \left[\left(w_K^G\right)_B^E\right]^2}}\right) \tag{7.9}$$

In addition, the kinematic inclination angle γ_K and the kinematic course angle χ_K can be obtained from

$$\chi_K = \arctan\left(\frac{\left(v_K^G\right)_O^E}{\left(u_K^G\right)_O^E}\right), \tag{7.10}$$

$$\gamma_K = -\arctan\left(\frac{\left(w_K^G\right)_O^E}{\sqrt{\left[\left(u_K^G\right)_O^E\right]^2 + \left[\left(v_K^G\right)_O^E\right]^2}}\right), \tag{7.11}$$

with $\mathbf{R}_{BO} = \mathbf{R}_x(\Phi)\mathbf{R}_y(\Theta)\mathbf{R}_z(\Psi)$ as

$$\left(\mathbf{V}_K^G\right)_O^E := \begin{bmatrix} u_K^G \\ v_K^G \\ w_K^G \end{bmatrix}_O^E = \mathbf{R}_{BO}^T \left(\mathbf{V}_K^G\right)_B^E. \tag{7.12}$$

The same relations (7.8)-(7.12) hold for the corresponding aerodynamic quantities and merely the index K needs to be replaced by the index A in these equations. The dynamic equations for the propagation of the position states

$$\dot{\lambda}^G = \frac{V_K^G \cos(\gamma_K)\sin(\chi_K)}{(N_\phi + h^G)\cos(\phi)} \tag{7.13}$$

$$\dot{\phi}^G = \frac{V_K^G \cos(\gamma_K)\cos(\chi_K)}{M_\phi + h^G} \tag{7.14}$$

$$\dot{h}^G = V_K^G \sin(\gamma_K) \tag{7.15}$$

are defined based on the WGS84 reference ellipsoid with semi-major axis a and semi-minor axis b as well as the flattening f and first eccentricity e

$$a = 6378137\ m, \qquad b = 6356752.3142\ m, \tag{7.16}$$

$$f = (a-b)/a, \qquad e = \sqrt{f(2-f)}, \tag{7.17}$$

as well as the curvature radii in the prime vertical section and the meridian:

$$N_\phi = \frac{a}{\sqrt{1 - e^2\sin(\phi)^2}}, \tag{7.18}$$

$$M_\phi = N_\phi \frac{1-e^2}{1-e^2\sin(\phi)^2} \tag{7.19}$$

Moreover, the attitude propagation for the Euler angles can be stated as follows:

$$\begin{bmatrix} \dot{\Phi} \\ \dot{\Theta} \\ \dot{\Psi} \end{bmatrix} = \begin{bmatrix} 1 & \sin(\Phi)\tan(\Theta) & \cos(\Phi)\tan(\Theta) \\ 0 & \cos(\Phi) & -\sin(\Phi) \\ 0 & \sin(\Phi)/\cos(\Theta) & \cos(\Phi)/\cos(\Theta) \end{bmatrix} \begin{bmatrix} p_K \\ q_K \\ r_K \end{bmatrix}_B \tag{7.20}$$

The translational dynamics denoted in the B-frame may be derived based on the conservation law for the linear momentum which yields

$$\left(\dot{\mathbf{V}}_K^G\right)_B^{EB} = \frac{\left(\mathbf{F}_T^G\right)_B}{m} - \left(\boldsymbol{\omega}_K^{OB}\right)_B \times \left(\mathbf{V}_K^G\right)_B^E, \tag{7.21}$$

with the aircraft's mass m and the total (T) force $\left(\mathbf{F}_T^G\right)_B$ in the center of gravity (G). This force is the sum of propulsion (P), gravitation (G), and aerodynamic (A) forces denoted in the B-frame

$$\left(\mathbf{F}_T^G\right)_B = \left(\mathbf{F}_P^G\right)_B + \left(\mathbf{F}_G^G\right)_B + \left(\mathbf{F}_A^G\right)_B, \tag{7.22}$$

and is often given in a normalized form as the load factor $\mathbf{n}^G = \left[n_x^G, n_y^G, n_z^G\right]^T$

$$\begin{bmatrix} n_x^G \\ n_y^G \\ -n_z^G \end{bmatrix} = \frac{\left(\mathbf{F}_T^G\right)_B - \left(\mathbf{F}_G^G\right)_B}{mg}. \tag{7.23}$$

Moreover, the rotational dynamics are derived based on the conservation law of the angular momentum

$$\left(\dot{\boldsymbol{\omega}}_K^{OB}\right)_B^B = \left(\mathbf{I}^G\right)_{BB}^{-1}\left[\left(\mathbf{M}_T^G\right)_B - \left(\boldsymbol{\omega}_K^{OB}\right)_B \times \left(\mathbf{I}^G\right)_{BB}\left(\boldsymbol{\omega}_K^{OB}\right)_B\right], \tag{7.24}$$

with the inertia tensor $\left(\mathbf{I}^G\right)_{BB} \in \mathbb{R}^{3\times 3}$ and the vector of total (T) moments $\left(\mathbf{M}_T^G\right)_B$ at the center of gravity G denoted in the B-frame. Similar to the forces, the total moments can be described as the sum of the propulsion (P) and aerodynamic moments (A)

$$\left(\mathbf{M}_T^G\right)_B = \left(\mathbf{M}_P^G\right)_B + \left(\mathbf{M}_A^G\right)_B. \tag{7.25}$$

Assuming a constant gravitational force in the direction of the z_O-axis $\left(\mathbf{F}_G^G\right)_O = [0, 0, mg]^T$ with the constant acceleration g, the gravitational force $\left(\mathbf{F}_G^G\right)_B$ in the B-frame is obtained from:

$$\left(\mathbf{F}_G^G\right)_B = \mathbf{R}_{BO}\left(\mathbf{F}_G^G\right)_O \tag{7.26}$$

The aerodynamic force vector $\left(\mathbf{F}_A^A\right)_A = \left(\mathbf{F}_A^G\right)_A$ at the aerodynamic reference point (A)

is stated in the A-frame

$$\left(\mathbf{F}_A^A\right)_A := \bar{q}S \begin{bmatrix} -C_D \\ C_Q \\ -C_L \end{bmatrix}_A , \tag{7.27}$$

with the drag-coefficient C_D, the side-force coefficient C_Q, the lift-force coefficient C_L, the reference wing area S, and the dynamic pressure $\bar{q} = \frac{1}{2}\rho \left(V_A^G\right)^2$ which is defined using the air density ρ and the absolute value of the aerodynamic velocity V_A^G. The aerodynamic coefficients are usually modeled in linear form. For example, the lift coefficient

$$C_L := C_{L,0} + C_{L,\alpha}\alpha_A + C_{L,\eta}\eta + \dots , \tag{7.28}$$

is the sum of the zero lift coefficient $C_{L,0}$, motion-induced influences (e.g. $C_{L,\alpha}$ for the aerodynamic AoA) as well as control-induced influences (e.g. $C_{L,\eta}$ for the elevator deflection η). The aerodynamic moment at the center of gravity (G) is obtained from

$$\left(\mathbf{M}_A^G\right)_B = \left(\mathbf{M}_A^G\right)_B + \left(\mathbf{r}^{GA}\right)_B \times \left(\mathbf{F}_A^A\right)_B , \tag{7.29}$$

with the lever arm $\left(\mathbf{r}^{GA}\right)_B$ between the center of gravity (G) and the aerodynamic reference point (A). The aerodynamic moment vector $\left(\mathbf{M}_A^A\right)_B = \left(\mathbf{M}_A^G\right)_B$ at the aerodynamic reference point (A) is defined in the B-frame

$$\left(\mathbf{M}_A^A\right)_B := \bar{q}S \begin{bmatrix} sC_l \\ \bar{c}C_m \\ sC_n \end{bmatrix}_B , \tag{7.30}$$

with the mean aerodynamic chord $\bar{c}$ and the wing semi-span s. The coefficients are, as for the aerodynamic forces, often based on linear models of the form

$$\begin{aligned} C_l &:= C_{l,\beta}\beta_A + C_{l,p}p^* + C_{l,r}r^* + C_{l,\zeta}\zeta + C_{l,\xi}\xi + \dots , \\ C_m &:= C_{m,0} + C_{m,\alpha}\alpha_A + C_{m,q}q^* + C_{m,\eta}\eta + \dots , \\ C_n &:= C_{n,\beta}\beta_A + C_{n,p}p^* + C_{n,r}r^* + C_{n,\zeta}\zeta + C_{n,\xi}\xi + \dots , \end{aligned} \tag{7.31}$$

with the normalized rotation rates

$$p^* := \frac{p_A b}{2V_K^G}, \qquad q^* := \frac{q_A \bar{c}}{2V_K^G}, \qquad r^* := \frac{r_A b}{2V_K^G}, \qquad b = 2s. \tag{7.32}$$

where p_A, q_A, and r_A are the aerodynamic rotation rates. The moment coefficients include among others: $C_{n,\beta}$ related to the weathercock stability, $C_{l,\beta}$ related to the Dihedral stability, $C_{m,\alpha}$ related to pitch stiffness, $C_{m,q}$ related to pitch damping, and control

surface coefficients ($C_{l,\zeta}, C_{l,\xi}$, $C_{m,\eta}$, $C_{n,\zeta}, C_{n,\xi}$) related to the control effectiveness of the aileron deflection (ξ), elevator deflection (η) and rudder deflection (ζ). The primary control surfaces are in many cases moved by a servomechanism. Related modeling concepts in the context of the optimal control based clearance approach are discussed in the following section.

7.2.2 Actuator Modeling Alternatives for the Clearance Task

A dynamic model of the servomechanism, which is often used in practice, can be based on a second-order model (damping $\bar{\zeta}$, natural frequency ω_n) of the form

$$\begin{bmatrix} \dot{x}_p(t) \\ \dot{x}_v(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\bar{\zeta}\omega_n \end{bmatrix} \begin{bmatrix} x_p(t) \\ x_v(t) \end{bmatrix} + \begin{bmatrix} 0 \\ k \end{bmatrix} \delta(t), \tag{7.33}$$

with the actuator position state $x_p(t)$, the actuator velocity (or rate) state $x_v(t)$ as well as the commanded value $\delta(t)$ and corresponding input gain k. One of the key concepts to model the clearance problem as an optimal control problem is to introduce the actuator rate and position limits as purely state dependent constraints. This concept is discussed in the following. It is assumed that both the actuator position state $x_p(t)$ and the actuator rate state $x_v(t)$ are subject to box constraints of the form

$$x_{p,lb} \leq x_p(t) \leq x_{p,ub}, \tag{7.34}$$

$$x_{v,lb} \leq x_v(t) \leq x_{v,ub}, \tag{7.35}$$

and that the commanded value $\delta(t)$ is depending on the control input which is used for the clearance task. These constraints are the only constraints considered under the optimal control based clearance approach investigated in this thesis. The actuator modeling approaches for the actuator position and rate limits are discussed from a phase plane perspective as illustrated in Fig. 7.3.

Consider the case where the commanded input $\delta(t)$ is at the maximum value. A typical physical behavior of a servomechanism for the primary control surfaces of an aircraft subject to rate and position limits is illustrated by the dashed trajectory in Fig. 7.3. If the command is at the maximum value, the servo controller increases the rate until it gradually saturates, e.g. due to the maximum moment which can be generated by the motor of the actuator. As soon as the rate saturates, it stays at the same value until the position of the actuator approaches its maximum value. When reaching the position limit, the rate starts to drop until arriving at the maximum position. At this point the rate is zero and, obviously, the actuator position cannot be increased further. Note that the control input was unchanged during the process, and therefore external dis-

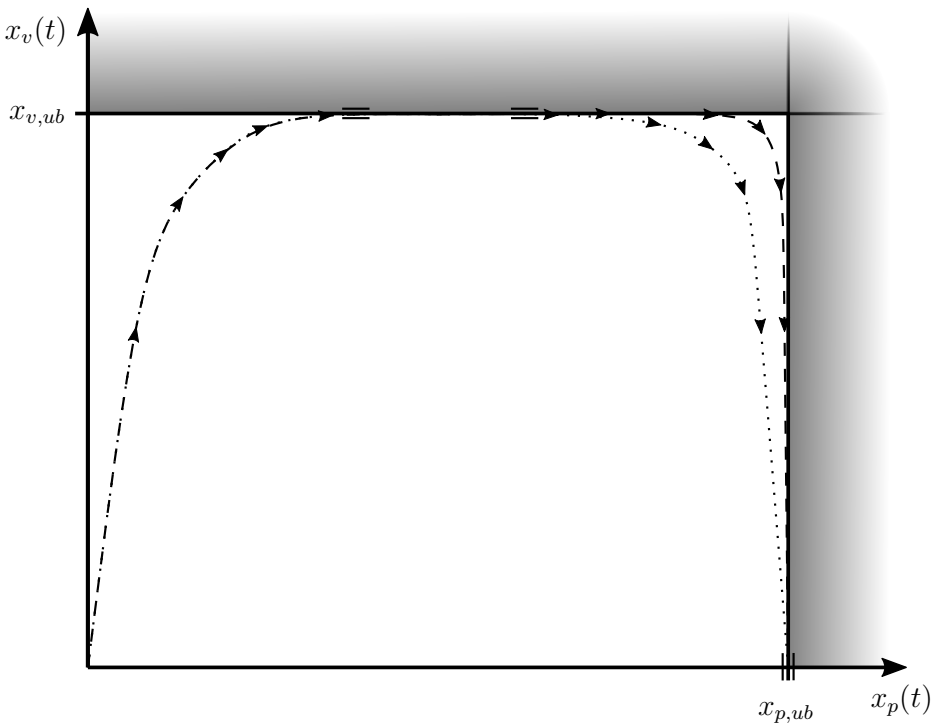


Figure 7.3: Schematic view of paths in the phase plane for two different actuator modeling approaches (optimal control with state constraints dotted line, physical behavior dashed line).

turbance inputs are assumed to exist to satisfy the state equations (e.g. reaction forces). Note also that, under this assumption, the control input is not uniquely defined.

Next, consider the optimal control based modeling approach using purely state dependent constraints. In this case the actuator rate and position limits are introduced as constraints along the trajectory and the responsibility for not exceeding the limits is transferred from the system to the control. Note that under this modeling approach the internal limits are essentially exposed and the task of the optimal control is to bring the system to the bound (if it is optimal to go there), to hold the system on the bound (if it is optimal to remain there), and to bring the system away from the bound (if it is optimal not to stay there). Such a trajectory in the phase plane is illustrated in Fig. 7.3 by the dotted line. The system is first steered to the rate bound and held there using a particular choice for the control (*boundary control*). Before the system can reach the maximum position state it needs to leave the rate bound and control the rate to zero at the point where the position limit is reached. In order to remain on the position limit the rate needs to be controlled to zero as for the rate bound. It is important to observe that in the optimal control approach the control is uniquely determined by the mentioned conditions for the boundary control. Under certain assumptions this allows for a unique characterization of the worst-case control based on the Minimum Principle as shown in chapter 8.

7.2.3 Flight Control Systems

Flight control systems [76, 77] are often designed in a cascaded control architecture. The main motivation for this layout is contributed to the time scale separation between the different dynamic effects of the aircraft dynamics. To illustrate this point, consider for example a desired change the height of an aircraft flying in a stationary, straight and level flight condition. From Newton's laws it is known that in order to change the direction of motion, a force needs to be exerted perpendicular to this direction. In the context of a height change for the aircraft this implies that the climb angle and, as such, the lift force needs to be increased. In order to increase the lift force, the AoA needs to be changed which can be achieved by rotating the aircraft around the y_B-axis. For conventional aircraft this rotation is taking place on a considerably smaller timescale than changes in the translational dynamics. Because of this, the control system is in many cases split into an inner-loop responsible for controlling the moment dynamics (rotation) and an outer-loop responsible for controlling the force dynamics (translation). One of the most basic tasks of aircraft control systems is to alter the dynamic characteristics of the aircraft. Note that in a direct law full authority may be given to the pilot over the aircraft control, e.g. by means of a rod connection to the primary control surfaces. In order to change the dynamic behavior of the aircraft a Stability Augmentation System (SAS) can be employed. This system introduces an artificial feedback which can be used to adequately change quantities such as Eigenfrequencies and damping depending on the flight condition. A classical example for this application is to counteract the decreased aerodynamic damping in higher altitudes due to the lower density of the surrounding air. In order to artificially change the damping characteristics of the aircraft for example a feedback of the pitch rate of the aircraft to the elevator may be embedded in the control law. The internal parameters of the control system, such as gain values, are typically scheduled depending on the flight condition, e.g. based on the static and dynamic pressure. This scheduling takes into account the altered aerodynamic characteristics of the aircraft dynamics when flying at different speeds or different altitudes. Gain design (e.g. using the LQR method) in industry practice is primarily based on linearized systems of the aircraft model over a grid of flight conditions and load cases. Besides the SAS, a Control and Stability Augmentation System (CSAS) or a Fly-by-Wire (FBW) system may be used. For CSAS systems the control system has an additional electrical component which improves the handling qualities. Here, the pilot does not anymore give direct commands but uses a higher level command, such as the angle of attack or the load factor. For FBW systems the mechanical link between the pilot and the aircraft control is completely removed giving full authority to the flight control system. The basic internal tasks of a flight control system include:

- The translation of the commanded value (such as a roll angle) into outputs (e.g. the required roll rate) and the allocation of the necessary commands to the actuators (such as control surface deflections for the ailerons).
- The conditioning of the sensory inputs measuring the states of the aircraft, e.g. by using filters to improve the quality of the signals.
- The conditioning of the outputs, that are the commanded values from the control law fed to the actuation system.
- The implementation of protections, such as stall protections or protections against internal saturations.

It is noteworthy, that most commercial aircraft are equipped with autopilot systems. These systems provide high-level interfaces for the pilot to accomplish tasks such as speed and attitude control, as well as flight path and trajectory control. Advanced components of autopilot systems may include capabilities for automatic take-off and landing as well as a flight management system for mission control. There exist several requirements for flight control systems which define certification specifications and acceptable means of compliance (cf. e.g. MIL-F-8785C [78], CS-23 [79], CS-25 [80]). An excellent overview of typical criteria in civil aviation can be found in Ref. [81]. Here, it is important to mention that the formulation considered in this thesis is general in the sense that a combination of control inputs and parameters can be used with a generic cost function modeling the criterion. For the investigation of worst-case pilot inputs, which is the main focus of the approaches developed in chapters 8 and 9, the following criteria are of particular interest (cf. Ref. [81]):

- Criteria for the maximum allowable AoA (α_A) and AoS (β_A).
- Limits on the load factors $n_{x,B}$, $n_{y,B}$, and $n_{z,B}$.
- Maximum allowable limits for other aircraft states (such as angular body rates).

In Ref. [81] these criteria are considered over a fixed time horizon $[0, t_f]$ where the final time t_f is 5-10 times the short period time scale of the closed-loop system. It should be noted that the classical methodology to aircraft control based on linear controllers is still the prevailing approach in industry practice. As such, the approaches developed in chapters 8 and 9 are shown primarily for a conventional fixed-wing aircraft with a control system of this type. However, rapid advancements in the field of nonlinear aircraft control have led to increased interest in the application of modern and advanced control systems. Thus, besides the classical example, a modern nonlinear model reference adaptive control law is investigated in chapter 9 using combination of pilot inputs and parametric uncertainties.

Chapter 8

Optimal Control Based Clearance for Linear and Quasi-linear Systems

In the following chapter, the clearance problem for linear dynamic systems with and without actuator limits is investigated. The worst-case control input is characterized based on the Minimum Principle for both cases, that is the linear case (Cat I) as well as the quasi-linear case where the actuators are subject to rate and position limits (Cat II). For the numerical analysis a closed-loop aircraft system with actuator dynamics, linearized around a stationary, straight and level flight condition is used. Note that for this case the linearized systems in the longitudinal and lateral plane decouple. Therefore, both cases can be treated separately. Besides the investigation of the nominal case, approaches related to parameter dependent systems with known limits and subject to a distribution as well as multi-criteria analysis are presented.

8.1 Problem Definition

Consider a linear dynamic system with state vector $\mathbf{x}(t) \in \mathbb{R}^{n_x}$, control vector $\mathbf{u}(t) \in \mathbb{R}^{n_u}$, as well as the system matrix $\mathbf{A} \in \mathbb{R}^{n_x \times n_x}$, and the control matrix $\mathbf{B} \in \mathbb{R}^{n_x \times n_u}$:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{8.1}$$

The output equation for the scalar output $y(t)$ with the corresponding output vector $\mathbf{c} \in \mathbb{R}^{n_x}$ is given as follows

$$y(t) = \mathbf{c}^T\mathbf{x}(t), \tag{8.2}$$

and models the criterion under investigation. Moreover, a subset of the state variables, indexed by the set $\mathcal{I}_b$, is assumed to be constrained:

$$x_{j,lb} \leq x_j(t) \leq x_{j,ub}, \qquad \forall j \in \mathcal{I}_b \tag{8.3}$$

In particular, these constraints are introduced to model actuator rate and position limits of the system (cf. Sec. 7.2.2). For testing the closed-loop system the problem is formulated as a Mayer-type optimal control problem on a time interval $I^t := [t_0, t_f], t_0 < t_f$ starting from the initial equilibrium point $\mathbf{x}(t_0) = \mathbf{x}_0$. Under this Mayer-type formulation the optimal value of the cost function $j^M(\mathbf{x}(t_f))$ depending on the value of the states at the final time point is sought for. For the clearance problem under investigation this cost function represents the scalar criterion $y(t_f)$ at the final time point t_f

$$j^M(\mathbf{x}(t_f)) := y(t_f) = \mathbf{c}^T\mathbf{x}(t_f). \tag{8.4}$$

Note that the general case of a Bolza-type cost function can, without loss of generality, be transformed to the Mayer-type cost function by introducing an additional state for the Lagrange term (cf. Sec. 5.1). The control input $\mathbf{u}(t)$ is assumed to be bounded by box bounds $\mathbf{u}(t) \in \mathcal{U}$ where $\mathcal{U} := [\mathbf{u}_{lb}, \mathbf{u}_{ub}], \mathbf{u}_{ub} > \mathbf{u}_{lb}$. Summarizing, the quasi-linear, Cat II clearance problem takes the following form including the purely state dependent path constraints for the actuator rate and position states:

$$\begin{aligned} \underset{\mathbf{u}(t) \in \mathcal{U}}{\text{minimize}} \quad & \mathbf{c}^T\mathbf{x}(t_f) \\ \text{subject to} \quad & \dot{\mathbf{x}}(t) - \mathbf{A}\mathbf{x}(t) - \mathbf{B}\mathbf{u}(t) = \mathbf{0}, \\ & \mathbf{x}(t_0) = \mathbf{x}_0, \\ & x_{j,lb} \leq x_j(t) \leq x_{j,ub}, \ \forall j \in \mathcal{I}_b, \\ & t \in I^t \end{aligned} \tag{8.5}$$

Similarly, the linear problem for Cat I type models can be stated directly based on the Cat II type problem formulation with $\mathcal{I}_b = \emptyset$:

$$\begin{aligned} \underset{\mathbf{u}(t) \in \mathcal{U}}{\text{minimize}} \quad & \mathbf{c}^T\mathbf{x}(t_f) \\ \text{subject to} \quad & \dot{\mathbf{x}}(t) - \mathbf{A}\mathbf{x}(t) - \mathbf{B}\mathbf{u}(t) = \mathbf{0}, \\ & \mathbf{x}(t_0) = \mathbf{x}_0, \\ & t \in I^t \end{aligned} \tag{8.6}$$

From an optimal control perspective, the characterization of the worst-case control differs for both formulations and is thus considered separately in the following section. Moreover, for both formulations it is interesting to observe that the cost function is non-decreasing for increasing final times t_f, if the system starts in an equilibrium point $\mathbf{x}(t_0)$. This fact can be illustrated as follows: suppose the optimal cost function value for a given t_f^0 is $j_{t_f^0}$. For a new final time $t_f^1 = t_f^0 + \Delta t$ it is always possible to keep the system in the initial equilibrium point for the time interval Δt and then apply

the previous controls, obtaining the previous cost function value $j_{t_f^0}$. Thus, the cost function value is non-decreasing with the final time t_f. As a direct consequence, the maximal reasonable (fixed) final time $t_f = t_{f,max}$ should be considered for the clearance problem. From a practical perspective the length of this time interval can be based for example on characteristic quantities of the aircraft dynamics, such as the short period time-scale (cf. Ref. [81]). The cases with and without state constraints, i.e. using a model of type Cat II or Cat I, are analyzed separately in the theoretical analysis. Following the ideas in Refs. [31, 82] the optimal (worst-case) control is characterized based on the Minimum Principle. Moreover, the orders of the state constraints are determined and explicit formulas for the Lagrange's multipliers corresponding to both types of state constraints (actuator rate and actuator deflection) are derived.

8.1.1 Cat I: Optimal Control Analysis without State Constraints

For the theoretical analysis considering Cat I models a scalar, box bounded, control input $u(t) \in \mathcal{U} = [u_{lb}, u_{ub}], u_{ub} > u_{lb}$ with the corresponding input vector $\mathbf{b} \in \mathbb{R}^{n_x}$ is considered. This leads to the following problem formulation:

$$\begin{aligned} \underset{u(t)\,\in\,\mathcal{U}}{\text{minimize}} \quad & \mathbf{c}^T\mathbf{x}(t_f) \\ \text{subject to} \quad & \dot{\mathbf{x}}(t) - \mathbf{A}\mathbf{x}(t) - \mathbf{b}u(t) = \mathbf{0}, \\ & \mathbf{x}(t_0) = \mathbf{x}_0, \\ & t \in I^t \end{aligned} \tag{8.7}$$

The Hamiltonian $\mathcal{H}\left(\boldsymbol{\lambda}(t), \mathbf{x}(t), u(t)\right)$ for this problem can be written using the co-states $\boldsymbol{\lambda}(t)$ as follows:

$$\mathcal{H}\left(\boldsymbol{\lambda}(t), \mathbf{x}(t), u(t)\right) := \boldsymbol{\lambda}^T(t)\left(\mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)\right) \tag{8.8}$$

It is assumed that the system is controllable with respect to the control $u(t)$ which can be expressed by the condition that the controllability matrix

$$\mathbf{U} := \begin{bmatrix} \mathbf{b} & \mathbf{A}\mathbf{b} & \dots & \mathbf{A}^{n_x-1}\mathbf{b} \end{bmatrix}, \tag{8.9}$$

has full rank. The following arguments for characterizing the control follow the derivation for showing the non-singularity of linear time invariant systems with a single input. In case of more than one control the analysis remains valid if the system is controllable with respect to each control. [54]

From the Minimum Principle the differential equations governing the evolution of

the co-states $\boldsymbol{\lambda}(t)$ is obtained as follows

$$\frac{\partial \mathcal{H}}{\partial \mathbf{x}} = -\dot{\boldsymbol{\lambda}}^T(t) = \boldsymbol{\lambda}^T(t)\mathbf{A}, \tag{8.10}$$

with non-trivial multipliers $\boldsymbol{\lambda}(t) \neq \mathbf{0}$. Denoting the switching function as

$$\mathcal{S}(t) := \frac{\partial \mathcal{H}}{\partial u} = \boldsymbol{\lambda}^T(t)\mathbf{b}, \tag{8.11}$$

the following cases characterize the optimal control for a non-instantaneous time interval depending on the sign of the switching function:

$$\hat{u}(t) = \begin{cases} u_{ub}, & \mathcal{S}(t) < 0 \\ u_{lb}, & \mathcal{S}(t) > 0 \\ u_s(t), & \mathcal{S}(t) = 0 \end{cases} \tag{8.12}$$

In the first two cases the control is regular and the sign of the switching function determines if the optimal control is on the upper bound u_{ub} or on the lower bound u_{lb}. In the third case, the control $u_s(t)$ is singular and cannot be directly obtained as the switching function is zero on an interval $[t_1, t_2], t_2 > t_1$. By taking the time derivatives of the switching function

$$\mathcal{S}^{(k)}(t) = \boldsymbol{\lambda}^T(t)\mathbf{A}^k\mathbf{b} = 0,\ k = 0, \ldots, n_x - 1, \tag{8.13}$$

it is important to observe that for a controllable system only the trivial solution $\boldsymbol{\lambda}(t) = \mathbf{0}$ satisfies the equation

$$\boldsymbol{\lambda}^T(t)\begin{bmatrix}\mathbf{b} & \mathbf{Ab} & \ldots & \mathbf{A}^{n_x-1}\mathbf{b}\end{bmatrix} = \boldsymbol{\lambda}^T(t)\mathbf{U} = \mathbf{0}, \tag{8.14}$$

as the controllability matrix $\mathbf{U}$ has full rank. This contradicts the non-triviality of the multipliers from the Minimum Principle. Thus, the switching function only exhibits isolated zeros and the optimal control is of pure bang-bang type. In particular, the sign of the switching function determines if the control is on the lower or upper bound (see cases in Eq. (8.12)).

8.1.2 Cat II: Optimal Control Analysis with State Constraints

For Cat II models (with state constraints) a second-order actuator model is considered where the control $u(t)$ directly influences the actuator deflection rate $x_v(t)$

$$\dot{x}_v(t) = \mathbf{a}_v^T\mathbf{x}(t) + b_v u(t), \tag{8.15}$$

with the system vector $\mathbf{a}_v \in \mathbb{R}^{n_x}$ and the scalar control influence $b_v \neq 0$. The time derivative of the position state of the actuator $x_p(t)$ is equal to the actuator rate $x_v(t)$ and may therefore be expressed as:

$$\dot{x}_p(t) = x_v(t) \tag{8.16}$$

In order to simplify notation, it is assumed that the state vector $\mathbf{x}(t)$ is arranged in such a way that the position state $x_p(t)$ appears in the first and the rate state $x_v(t)$ in the second element. Two types of constraints are considered in the following. The first type models symmetric rate limits $x_{v,b} > 0$ on the actuator of the form

$$|x_v(t)| \leq x_{v,b}. \tag{8.17}$$

The second type models symmetric limits with $x_{p,b} > 0$ for the actuator's position state

$$|x_p(t)| \leq x_{p,b}. \tag{8.18}$$

Adding these constraints to the Cat I problem (8.7) yields the following Cat II problem:

$$\begin{aligned} \underset{u(t) \in \mathcal{U}}{\text{minimize}} \quad & \mathbf{c}^T \mathbf{x}(t_f) \\ \text{subject to} \quad & \dot{\mathbf{x}}(t) - \mathbf{A}\mathbf{x}(t) - \mathbf{b}u(t) = \mathbf{0}, \\ & |x_v(t)| \leq x_{v,b}, \\ & |x_p(t)| \leq x_{p,b}, \\ & \mathbf{x}(t_0) = \mathbf{x}_0, \\ & t \in I^t \end{aligned} \tag{8.19}$$

It is noteworthy that for the following analysis it is useful to express the constraints for the actuator rate and position states in Eqs. (8.17) and (8.18) in the equivalent form

$$c_v(t) = \frac{1}{2}\left(x_v^2(t) - x_{v,b}^2\right) \leq 0, \tag{8.20}$$

$$c_p(t) = \frac{1}{2}\left(x_p^2(t) - x_{p,b}^2\right) \leq 0. \tag{8.21}$$

First, the order of the state constraints is determined. Recall that the order of a state constraint is defined as the number of times the constraint needs to be derived with respect to time until the control appears explicitly (cf. Sec. 5.4). On a state constrained arc of the rate state with $c_v(t) = 0, t \in [t_1, t_2], t_2 > t_1$ the boundary control $u_{b,v}(t)$

appears explicitly in first time derivative:

$$\dot{c}_v(t) = x_v(t)\dot{x}_v(t) = x_v(t)\left(\mathbf{a}_v^T\mathbf{x}(t) + b_v u_{b,v}(t)\right) = 0 \tag{8.22}$$

due to $x_v(t) = \pm x_{v,b} \neq 0$ and $b_v \neq 0$. Thus, the rate constraint $c_v(t)$ is of first-order. On a state constrained arc of the position state with $c_p(t) = 0, t \in [t_1, t_2], t_2 > t_1$ the first time derivative yields:

$$\dot{c}_p(t) = x_p(t)\dot{x}_p(t) = x_p(t)x_v(t) = 0 \tag{8.23}$$

Due to $x_p(t) = \pm x_{p,b} \neq 0$ it follows that the rate state needs to satisfy

$$x_v(t) = 0. \tag{8.24}$$

Note that this directly implies with $x_{v,b} > 0$ that it is not possible for the rate and position constraint to be active simultaneously. The second time derivative equates to

$$\ddot{c}_p(t) = \dot{x}_p(t)x_v(t) + x_p(t)\dot{x}_v(t) = x_p(t)\left(\mathbf{a}_v^T\mathbf{x}(t) + b_v u_{b,p}(t)\right) = 0, \tag{8.25}$$

as the rate satisfies Eq. (8.24). Moreover, due to $x_p(t) = \pm x_{p,b} \neq 0$ and $b_v \neq 0$ the boundary control $u_{b,p}(t)$ appears explicitly. Thus, the actuator position constraint $c_p(t)$ is of second-order. A single expression for the boundary control $u_b(t)$ on state constrained arcs for the rate or position can be directly derived from Eqs. (8.22) and (8.25). Note that for the second-order actuator model the condition

$$\mathbf{a}_v^T\mathbf{x}(t) + b_v u_b(t) = 0, \tag{8.26}$$

defines the boundary control $u_b(t)$ for both the position and rate constrained arc. This expression can be solved to obtain a single expression for the boundary control in both cases:

$$u_b(t) = -\frac{\mathbf{a}_v^T\mathbf{x}(t)}{b_v} \tag{8.27}$$

Next, expressions for the multiplier functions associated with both the rate and position state constraint are derived. For the analysis with purely state dependent constraints the augmented Hamiltonian $\mathcal{H}^A\left(\boldsymbol{\lambda}(t), \mathbf{x}(t), u(t), \mu_v(t), \mu_p(t)\right)$ is formed (cf. Sec. 5.4) using the multiplier functions $\mu_v(t)$ for the rate constraint and the multiplier function $\mu_p(t)$ for the position constraint

$$\mathcal{H}^A\left(\boldsymbol{\lambda}(t), \mathbf{x}(t), u(t), \mu_v(t), \mu_p(t)\right) := \boldsymbol{\lambda}^T\left(\mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)\right) + \mu_v(t)c_v(t) + \mu_p(t)c_p(t), \tag{8.28}$$

by directly adjoining the state constraints to the Hamiltonian. These multiplier func-

tions satisfy:

$$\mu_{p/v}(t) = \begin{cases} = 0, & c_{p/v}(t) < 0 \\ \geq 0, & c_{p/v}(t) = 0 \end{cases} . \tag{8.29}$$

Recall that either the rate or the position constraint can be active. Thus, on a rate constrained arc $t \in [t_1, t_2], t_2 > t_1$ the multiplier for the position constraint satisfies $\mu_p(t) = 0$. Similarly, on a position constrained arc $\mu_v(t) = 0$ holds. As in the unconstrained case the switching function is defined as

$$\mathcal{S}(t) = \boldsymbol{\lambda}^T(t)\mathbf{b}, \tag{8.30}$$

and the co-state equation evolves according to

$$\frac{\partial \mathcal{H}^A}{\partial \mathbf{x}} = -\dot{\boldsymbol{\lambda}}^T(t) = \boldsymbol{\lambda}^T(t)\mathbf{A} + \mu_v(t)\frac{\partial c_v}{\partial \mathbf{x}} + \mu_p(t)\frac{\partial c_p}{\partial \mathbf{x}}, \tag{8.31}$$

with

$$\frac{\partial c_v}{\partial \mathbf{x}} = [0, x_v(t), \mathbf{0}], \tag{8.32}$$

$$\frac{\partial c_p}{\partial \mathbf{x}} = [x_p(t), 0, \mathbf{0}]. \tag{8.33}$$

On a state constrained arc $t \in [t_1, t_2], t_2 > t_1$ the switching function satisfies $\mathcal{S}(t) = 0$. Taking the first time derivative yields:

$$\dot{S}(t) = \dot{\boldsymbol{\lambda}}^T(t)\mathbf{b} = -\boldsymbol{\lambda}^T(t)\mathbf{A}\mathbf{b} - \mu_v(t)\frac{\partial c_v}{\partial \mathbf{x}}\mathbf{b} + \mu_p(t)\frac{\partial c_p}{\partial \mathbf{x}}\mathbf{b} = 0 \tag{8.34}$$

On a rate constrained arc $\mu_p(t) = 0$ holds and the product $\frac{\partial c_v}{\partial \mathbf{x}}\mathbf{b}$ evaluates to

$$\frac{\partial c_v}{\partial \mathbf{x}}\mathbf{b} = [0, x_v(t), \mathbf{0}] \begin{bmatrix} 0 \\ b_v \\ \mathbf{0} \end{bmatrix} = x_v(t)b_v. \tag{8.35}$$

After insertion into Eq. (8.34)

$$-\boldsymbol{\lambda}^T(t)\mathbf{A}\mathbf{b} - \mu_v(t)\frac{\partial c_v}{\partial \mathbf{x}}\mathbf{b} = -\boldsymbol{\lambda}^T(t)\mathbf{A}\mathbf{b} - \mu_v(t)x_v(t)b_v = 0, \tag{8.36}$$

it directly follows that the multiplier $\mu_v(t)$ satisfies

$$\mu_v(t) = -\frac{\boldsymbol{\lambda}^T(t)\mathbf{A}\mathbf{b}}{x_v(t)b_v}, \tag{8.37}$$

with $x_v(t) = \pm x_{v,b}$. Note that the sign of $x_v(t)$ expresses if the constraint for the lower or

upper bound is active. In order to derive an explicit expression for the multiplier $\mu_p(t)$ on a position constrained arc the switching function needs to be derived a second time. On a position constrained arc the rate constraint is in-active, that is $\mu_v(t) = \dot{\mu}_v(t) = 0$, and the product $\frac{\partial c_p}{\partial \mathbf{x}}\mathbf{b}$ evaluates to zero:

$$\frac{\partial c_p}{\partial \mathbf{x}}\mathbf{b} = [x_p(t), 0, \mathbf{0}] \begin{bmatrix} 0 \\ b_v \\ \mathbf{0} \end{bmatrix} = 0 \tag{8.38}$$

Thus, based on Eq. (8.34) the second time derivative of the switching function can be expressed as

$$\ddot{S}(t) = -\dot{\boldsymbol{\lambda}}^T(t)\mathbf{A}\mathbf{b} = \boldsymbol{\lambda}^T(t)\mathbf{A}^2\mathbf{b} + \mu_p(t)\frac{\partial c_p}{\partial \mathbf{x}}\mathbf{A}\mathbf{b} = 0. \tag{8.39}$$

Note that the product $\frac{\partial c_p}{\partial \mathbf{x}}\mathbf{A}\mathbf{b}$ simplifies to

$$\frac{\partial c_p}{\partial \mathbf{x}}\mathbf{A}\mathbf{b} = [0, x_p(t), \mathbf{0}] \begin{bmatrix} 0 \\ b_v \\ \mathbf{0} \end{bmatrix} = x_p(t)b_v, \tag{8.40}$$

as the product $\frac{\partial c_p}{\partial \mathbf{x}}\mathbf{A}$ yields

$$\frac{\partial c_p}{\partial \mathbf{x}}\mathbf{A} = [x_p(t), \mathbf{0}]\,\mathbf{A} = [0, x_p(t), \mathbf{0}]\,, \tag{8.41}$$

due to the fact that the row corresponding to the position differential equation only has a single entry in the column corresponding to the rate state according to Eq. (8.16). After insertion into Eq. (8.39)

$$\boldsymbol{\lambda}^T(t)\mathbf{A}^2\mathbf{b} + \mu_p(t)x_p(t)b_v = 0, \tag{8.42}$$

it directly follows that the multiplier $\mu_p(t)$ satisfies

$$\mu_p(t) = -\frac{\boldsymbol{\lambda}^T(t)\mathbf{A}^2\mathbf{b}}{x_p(t)b_v}, \tag{8.43}$$

with $x_p(t) = \pm x_{p,b}$. Summarizing, for the case with actuator rate and position limits the control $u(t)$ is either regular, i.e. of bang-bang type, if both constraints are inactive and takes the value of the respective boundary control if one of the state constraints is active. Note that for the case of more than one control the analysis is considerably more involved as the control functions may exhibit different combinations of regular and singular controls.

8.1.3 Discretized Form for Cat I and Cat II Models

In order to obtain a form for the linear analysis which can be solved numerically the problem is transcribed into a finite dimensional parameter optimization problem (cf. chapter 6). In order to illustrate the transcription for Cat I and Cat II models the Backward Euler method with $N+1$ discretization points and discretization step length $h = (t_f - t_0)/N$ is considered, which is the basis of the discrete Minimum Principle presented in Ref. [52]. The results from this discrete Minimum Principle are used in the following sections to estimate quantities such as the co-states in order to verify the validity of the theoretical results from Secs. 8.1.1 and 8.1.2.

For the Backward Euler method state variable vectors $\mathbf{x}^{[i]} \in \mathbb{R}^{n_x}, i = 0, \ldots, N$ and control variable vectors $\mathbf{u}^{[i]} \in \mathbb{R}^{n_u}, i = 1, \ldots, N$ are introduced and the dynamic constraints are expressed as

$$\mathbf{x}^{[i]} - h\mathbf{A}\mathbf{x}^{[i]} - h\mathbf{B}\mathbf{u}^{[i]} - \mathbf{x}^{[i-1]} = \mathbf{0},\ i = 1, \ldots, N. \tag{8.44}$$

As such, the discretized problem statement (8.5) for Cat II models may be written as

$$\begin{aligned} \underset{\mathbf{x}^{[0]}, \mathbf{x}^{[i]}, \mathbf{u}^{[i]}, i \in 1 \ldots N}{\text{minimize}} \quad & \mathbf{c}^T \mathbf{x}^{[N]} \\ \text{subject to} \quad & \mathbf{A}_d \mathbf{x}^{[i]} + \mathbf{B}_d \mathbf{u}^{[i]} - \mathbf{x}^{[i-1]} = \mathbf{0},\ i = 1, \ldots, N, \\ & \mathbf{x}^{[0]} = \mathbf{x}(t_0), \\ & x_{j,lb} \leq x_j^{[i]} \leq x_{j,ub},\ i = 1, \ldots, N, \forall j \in \mathcal{I}_b, \\ & \mathbf{u}_{lb} \leq \mathbf{u}^{[i]} \leq \mathbf{u}_{ub},\ i = 1, \ldots, N, \end{aligned} \tag{8.45}$$

with the matrices $\mathbf{A}_d := \mathbf{I} - h\mathbf{A}$ and $\mathbf{B}_d := -h\mathbf{B}$ and the vector $\mathbf{c} \in \mathbb{R}^{n_x}$. Note that this discretized Cat II type problem reduces to a Cat I type problem for $\mathcal{I}_b = \emptyset$. Note further that the discretized problem can be stated as a Linear Program (LP) of the form

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{R}^{n_z}}{\text{minimize}} \quad & \mathbf{v}^T \mathbf{z} \\ \text{subject to} \quad & \mathbf{M}_h \mathbf{z} = \mathbf{b}_h, \\ & \mathbf{M}_g \mathbf{z} \leq \mathbf{b}_g, \end{aligned} \tag{8.46}$$

with the $n_z = (N+1)n_x + Nn_u$ dimensional parameter vector $\mathbf{z} \in \mathbb{R}^{n_z}$

$$\mathbf{z} := \left[\left(\mathbf{x}^{[0]}\right)^T, \ldots, \left(\mathbf{x}^{[N]}\right)^T, \left(\mathbf{u}^{[1]}\right)^T, \ldots, \left(\mathbf{u}^{[N]}\right)^T \right]^T, \tag{8.47}$$

the cost vector $\mathbf{v} \in \mathbb{R}^{n_z}$

$$\mathbf{v} := \left[\mathbf{0}, \ldots, \mathbf{c}^T, \mathbf{0}, \ldots, \mathbf{0} \right]^T, \tag{8.48}$$

the equality constraint matrix $\mathbf{M}_h \in \mathbb{R}^{n_h \times n_z}, n_h = (N+1)n_x$ and the corresponding right-hand side vector $\mathbf{b}_h \in \mathbb{R}^{n_h}$

$$\mathbf{M}_h := \begin{bmatrix} \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ -\mathbf{I} & \mathbf{A}_d & \dots & \mathbf{0} & \mathbf{0} & \mathbf{B}_d & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & -\mathbf{I} & \mathbf{A}_d & \mathbf{0} & \dots & \mathbf{B}_d \end{bmatrix}, \qquad \mathbf{b}_h := \begin{bmatrix} \mathbf{x}(t_0) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}, \tag{8.49}$$

as well as the inequality constraint matrix $\mathbf{M}_g \in \mathbb{R}^{n_g \times n_z}, n_g = (N-1)(n_x + n_u)$ with the right-hand side vector $\mathbf{b}_g \in \mathbb{R}^{n_g}$:

$$\mathbf{M}_g := \begin{bmatrix} \mathbf{0} & -\mathbf{I} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & -\mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & -\mathbf{I} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & -\mathbf{I} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{I} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} \end{bmatrix}, \qquad \mathbf{b}_g := \begin{bmatrix} -\mathbf{x}_{lb} \\ \vdots \\ -\mathbf{x}_{lb} \\ \mathbf{x}_{ub} \\ \vdots \\ \mathbf{x}_{ub} \\ -\mathbf{u}_{lb} \\ \vdots \\ -\mathbf{u}_{lb} \\ \mathbf{u}_{ub} \\ \vdots \\ \mathbf{u}_{ub} \end{bmatrix} \tag{8.50}$$

Observe that the initial state is already uniquely determined by the equality constraint corresponding to the first block row in $\mathbf{M}_h$ and $\mathbf{b}_h$ and each of the vectors $\mathbf{x}_{lb}, \mathbf{x}_{ub} \in \mathbb{R}^{n_x}$ for the state bounds as well as the vectors $\mathbf{u}_{lb}$ and $\mathbf{u}_{ub}$ for the control bounds is repeated N times in $\mathbf{b}_g$. In order to simplify notation the lower and upper bounds in $\mathbf{x}_{lb}$ and $\mathbf{x}_{ub}$ for the unbounded states are defined as $-\infty$, respectively $+\infty$

$$x_{k,lb} := -\infty, \qquad k \in \mathcal{I}_n, \tag{8.51}$$

$$x_{k,ub} := +\infty, \qquad k \in \mathcal{I}_n, \tag{8.52}$$

with $\mathcal{I}_n = \{0, \dots, n_x - 1\} \setminus \mathcal{I}_b$. Other discretization schemes with a higher order may

be employed, such as the Trapezoidal method with the defect constraint

$$\mathbf{x}^{[i]} - h\mathbf{A}\frac{\mathbf{x}^{[i]} + \mathbf{x}^{[i-1]}}{2} - h\mathbf{B}\frac{\mathbf{u}^{[i]} + \mathbf{u}^{[i-1]}}{2} - \mathbf{x}^{[i-1]} = \mathbf{0},\ i = 1, \ldots, N, \tag{8.53}$$

and the additional control vector $\mathbf{u}^{[0]}$. The transcribed problem has a similar form in this case.

8.2 Nominal Problems for Linearized Aircraft Dynamics

8.2.1 Nominal Linear Benchmark Problems

For the following examples the longitudinal and lateral channel of the inner-loop controller for a conventional fixed-wing aircraft is investigated. The model is obtained from linearization around a steady state, straight and level flight condition at a reference altitude h_{ref} and reference velocity V_{ref}. For this linearization point with trim state vector $\mathbf{x}_0 \in \mathbb{R}^{n_x}$ and trim control vector $\mathbf{u}_0 \in \mathbb{R}^{n_u}$ the linear model is of the form

$$\delta\dot{\mathbf{x}}(t) = \mathbf{A}\delta\mathbf{x}(t) + \mathbf{B}\delta\mathbf{u}(t),\ \mathbf{A} \in \mathbb{R}^{n_x \times n_x},\ \mathbf{B} \in \mathbb{R}^{n_x \times n_u}, \tag{8.54}$$

$$\delta y(t) = \mathbf{c}^T\delta\mathbf{x}(t),\ \mathbf{c} \in \mathbb{R}^{n_x}, \tag{8.55}$$

with

$$\delta\mathbf{x}(t) := \mathbf{x}(t) - \mathbf{x}_0, \tag{8.56}$$

$$\delta\mathbf{u}(t) := \mathbf{u}(t) - \mathbf{u}_0. \tag{8.57}$$

In the following the δ-notation is dropped for better readability, bearing in mind that all quantities are relative to the specific trim values. In addition, unless otherwise specified, all quantities are given at the reference point $R := G$ and the superscript G for the center of gravity is omitted. Moreover, the normal load factor and the lateral factor in the B-frame are denoted with n_z and n_y.

It is important to observe, that under the chosen flight condition the aircraft's linearized dynamic equations decouple for the longitudinal and lateral plane. This implies that both channels can be investigated separately. The inputs to the inner-loop controller are the normal load factor command $n_{z,c}(t)$, the roll angle command $\Phi_c(t)$, and the lateral factor command $n_{y,c}(t)$. Note that the lateral factor is controlled to zero for most maneuvers.

In addition to the command inputs from the pilot, the influence of wind is investigated. The wind velocities $u_W(t)$, $v_W(t)$, and $w_W(t)$ relative to the earth (E) at the

aerodynamic reference point (A) are denoted in the O-frame and, as such, describe the wind components corresponding to the x_O-, y_O-, and z_O-axis. The first and second time derivatives of these wind velocities are modeled using linear second-order models (damping $\bar{\zeta}_W$, natural frequency ω_W). As such, besides the wind velocity states $u_W(t)$, $v_W(t)$, and $w_W(t)$, additionally the wind acceleration states $u_{W,a}(t)$, $v_{W,a}(t)$, and $w_{W,a}(t)$ are added to the state vector of the system. For all three wind velocities dynamic models of the same form are used which, for example, in case of the vertical wind velocity reads:

$$\begin{bmatrix} \dot{w}_W(t) \\ \dot{w}_{W,a}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_W^2 & -2\bar{\zeta}_W\omega_W \end{bmatrix} \begin{bmatrix} w_W(t) \\ w_{W,a}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \omega_W^2 \end{bmatrix} w_{W,c}(t) \tag{8.58}$$

The reasoning behind using second-order models for the wind dynamics is that in addition to the translational wind velocities the effect of wind rotation is considered in the analysis. Here, the rotational components $q_W(t)$ and $r_W(t)$ are approximated as

$$q_W(t) \approx \frac{w_{W,a}(t)}{V_{ref}}, \tag{8.59}$$

$$r_W(t) \approx \frac{v_{W,a}(t)}{V_{ref}}. \tag{8.60}$$

Note that as the control-like variables introduced for modeling wind merely influence the time derivative of the wind acceleration states ($u_{W,a}(t)$, $v_{W,a}(t)$, $w_{W,a}(t)$) in the second-order models the wind controls are not able to instantaneously produce a wind rotation.

The minimum and maximum wind gust velocities which are used for the lower and upper limits of the artificial wind controls $u_{W,c}(t)$, $v_{W,c}(t)$, and $w_{W,c}(t)$ are set to $\pm 7.62\ m/s$ ($25\ ft/s$) which represent severe gust velocities [83]. It is important to mention that the gust modeling approach using second-order models does not allow to directly impose specific gust lengths. In fact, the exact shape of the gust is implicitly determined by the optimization. Thus, it is important to ensure that the optimal gust is capable of exciting relevant Eigenmodes of the system under investigation which implies that the parameters $\bar{\zeta}$ and ω_W are application specific. For the system investigated in this thesis these parameters are set to $\bar{\zeta}_W = 1$ as well as $\omega_W = 1$ and the artificial wind controls are bounded by

$$-7.62\ m/s \leq u_W(t) \leq 7.62\ m/s, \tag{8.61}$$

$$-7.62\ m/s \leq v_W(t) \leq 7.62\ m/s, \tag{8.62}$$

$$-7.62\ m/s \leq w_W(t) \leq 7.62\ m/s. \tag{8.63}$$

The aircraft's plant model with the state vector $\mathbf{x}_{ac}(t) \in \mathbb{R}^{10}$

$$\mathbf{x}_{ac}(t) := [h(t), V_K(t), \alpha_K(t), \beta_K(t), \Phi(t), \Theta(t), \Psi(t), p_K(t), q_K(t), r_K(t)]^T \tag{8.64}$$

and the system matrix $\mathbf{A}_{ac} \in \mathbb{R}^{10\times 10}$ is of the form

$$\dot{\mathbf{x}}_{ac}(t) = \mathbf{A}_{ac}\mathbf{x}_{ac}(t) + \mathbf{B}_{act}\mathbf{x}_{act}(t) + \mathbf{B}_{dis}\begin{bmatrix}\mathbf{x}_{dis}(t)\\ \dot{\mathbf{x}}_{dis}(t)\end{bmatrix}, \tag{8.65}$$

with the actuator states

$$\mathbf{x}_{act}(t) := [x_{e,p}(t), x_{a,p}(t), x_{r,p}(t)]^T, \tag{8.66}$$

for the primary control surfaces (elevator position $x_{e,p}(t)$, aileron position $x_{a,p}(t)$, rudder position $x_{r,p}(t)$), the corresponding input matrix $\mathbf{B}_{act} \in \mathbb{R}^{10\times 3}$, as well as the wind disturbance states

$$\mathbf{x}_{dis}(t) := [u_W(t), v_W(t), w_W(t)]^T, \tag{8.67}$$

with input matrix $\mathbf{B}_{dis} \in \mathbb{R}^{10\times 6}$. The dynamic equations for the actuation system are modeled by linear second-order models (damping $\bar{\zeta}_k$, natural frequency $\omega_{n,k}$) of the form

$$\begin{bmatrix}\dot{x}_{k,p}(t)\\ \dot{x}_{k,v}(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -\omega_{n,k}^2 & -2\bar{\zeta}_k\omega_{n,k}\end{bmatrix}\begin{bmatrix}x_{k,p}(t)\\ x_{k,v}(t)\end{bmatrix} + \begin{bmatrix}0\\ b_k\end{bmatrix}x_{k,c}(t), \tag{8.68}$$

with $k \in \{e, a, r\}$ (elevator, aileron, and rudder) and the actuator rate states $x_{e,v}(t)$, $x_{a,v}(t)$, and $x_{r,v}(t)$. Moreover, the commanded values from the controller fed to the actuator are denoted by $x_{k,c}(t)$. Due to the decoupled nature of the complete closed-loop dynamic model for the flight condition in the linearization point the model is separated in the following into a longitudinal and lateral model. For the control system under investigation the basic internal structure of the inner-loop controller has both proportional and integral parts. In particular, the states of the basic a PI-type control law in this controller are denoted with $e_{n_z}(t)$ for the longitudinal channel and $e_\Phi(t)$ as well as $e_\Psi(t)$ for the lateral channel. For illustration purposes the step responses and pole-zero plots for the basic closed-loop systems in the longitudinal and lateral channel are presented. In case of the longitudinal channel the reduced state vector of the close-loop system is defined as

$$\mathbf{x}_{lon,r} = [x_{e,p}(t), x_{e,v}(t), V_K(t), \alpha_K(t), \Theta(t), q_K(t), e_{n_z}(t)], \tag{8.69}$$

and in case of the lateral model:

$$\mathbf{x}_{lat,r} = [x_{a,p}(t), x_{r,p}(t), x_{a,v}(t), x_{r,v}(t), \beta_K(t), \Phi(t), \Psi(t), r_K(t), p_K(t), e_\Phi(t), e_\Psi(t)] \quad (8.70)$$

The step responses of the basic control law for the longitudinal and lateral inputs at a typical envelope point are illustrated in Fig. 8.1. Notice the non-minimum phase behavior of the normal load factor $n_z(t)$ and $n_y(t)$ as well as the high damping of the roll angle response $\Phi(t)$. Additionally, the response of the second-order wind model for $w_W(t)$ is presented in Fig. 8.1. Moreover, the pole-zero plots for the normal load factor, lateral load factor, and the roll angle are depicted in Figs. 8.2-8.4.

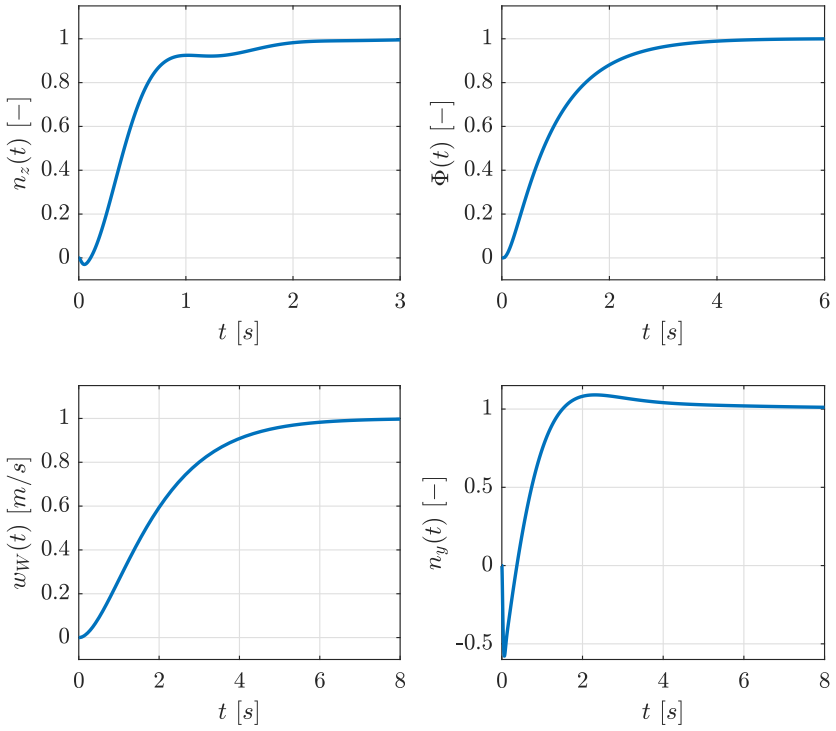


Figure 8.1: Responses of the normal load factor $n_z(t)$ (top left), the roll-angle $\Phi(t)$ (top right), the normal wind $w_W(t)$ (bottom left), and the lateral load factor $n_y(t)$ (bottom right) to a unit step.

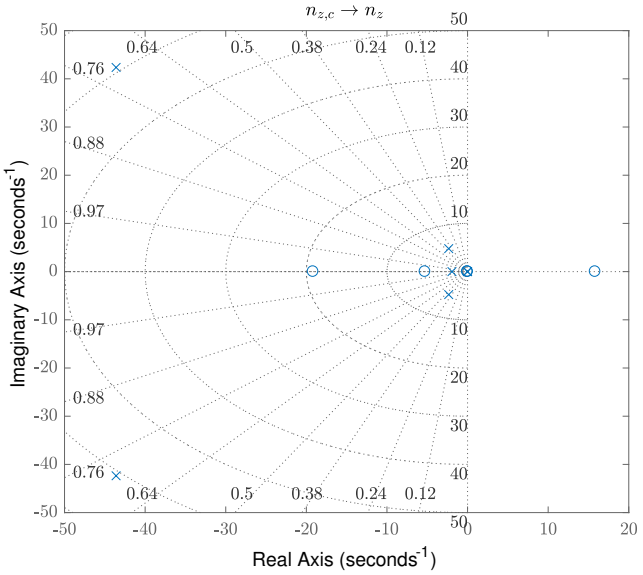


Figure 8.2: Pole-zero plot for the normal load factor n_z.

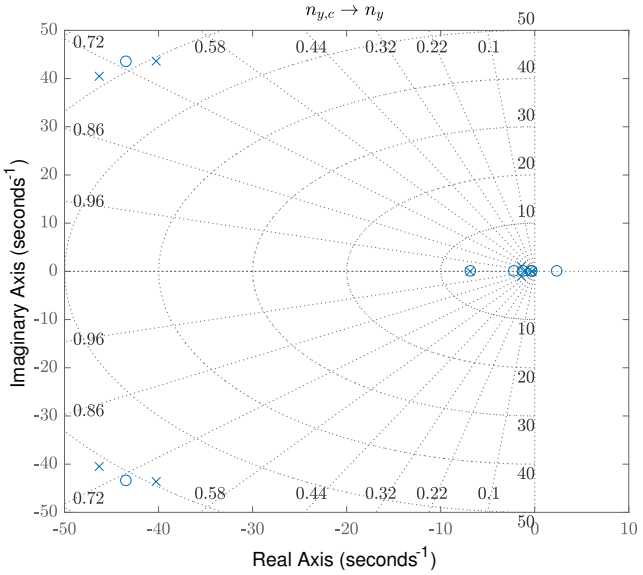


Figure 8.3: Pole-zero plot for the lateral load factor n_y.

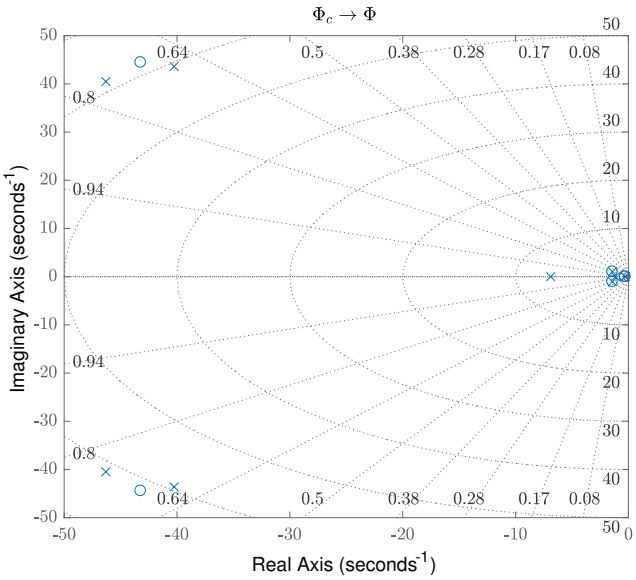


Figure 8.4: Pole-zero plot for the roll angle Φ.

For the following analysis of the longitudinal and lateral channel the basic inner-loop control law is considered with all available states corresponding to sensor models, aeroelasticity models, and other internal states, e.g. for notch filters in the inner-loop controller.

Longitudinal Model

For the longitudinal dynamics the state vector $\mathbf{x}_{lon}(t) \in \mathbb{R}^{46}$ contains the wind velocities $u_W(t)$ and $w_W(t)$, the actuator position and rate states of the elevator $x_{e,p}(t)$ and $x_{e,v}(t)$, the geodetic height $h(t)$, the kinematic velocity $V_K(t)$, the kinematic AoA $\alpha_K(t)$, the pitch angle $\Theta(t)$, the pitch rate $q(t)$, the integral state of the PI-type error controller $e_{n_z}(t)$, as well as sensor model and additional internal controller states collected in the vector $\mathbf{x}_{F,lon}(t) \in \mathbb{R}^{34}$:

$$\begin{aligned}\mathbf{x}_{lon}(t) := [&u_W(t), u_{W,a}(t), w_W(t), w_{W,a}(t), x_{e,p}(t), x_{e,v}(t),\\ &h(t), V_K(t), \alpha_K(t), \Theta(t), q_K(t), e_{n_z}(t), \mathbf{x}_{F,lon}^T(t)]^T\end{aligned} \tag{8.71}$$

The control vector $\mathbf{u}_{lon}(t) := [u_{W,c}(t), w_{W,c}(t), n_{z,c}(t)]^T$ contains the wind commands subject to general box bounds (cf. constraints (8.61) and (8.63)) and the normal load factor command $n_{z,c}(t)$ relative to the trim value bounded by $-0.5 \leq n_{z,c}(t) \leq 0.5$. These constraints define the admissible control set $\mathcal{U}_{lon}$ for the longitudinal plane. Let

$$\tilde{\mathbf{x}}_{lon}(t) := \left[h(t), V_K(t), \alpha_K(t), \Theta(t), q_K(t), e_{n_z}(t), \mathbf{x}_{F,lon}^T(t)\right]^T, \tag{8.72}$$

collect all states besides the elevator and wind states. The dynamic equations for the

elevator position state $x_{e,p}(t)$ is expressed as

$$\dot{x}_{e,p}(t) = \mathbf{a}_{e,p}^T \left[x_{e,p}(t), x_{e,v}(t), \tilde{\mathbf{x}}_{lon}^T(t)\right]^T = x_{e,v}(t), \qquad \mathbf{a}_{e,p}^T := [0, 1, \mathbf{0}], \tag{8.73}$$

and the servo rate $x_{e,v}(t)$ with the scalar control influence $b_{e,v}$ evolves according to

$$\dot{x}_{e,v}(t) = \mathbf{a}_{e,v}^T \left[x_{e,p}(t), x_{e,v}(t), \tilde{\mathbf{x}}_{lon}^T(t)\right]^T + b_{e,v} n_{z,c}(t), \qquad \mathbf{a}_{e,v} \in \mathbb{R}^{42}. \tag{8.74}$$

Thus, the dynamic equations for the longitudinal channel are of the form

$$\begin{aligned}
\frac{d}{dt} \underbrace{\begin{bmatrix} u_W(t) \\ u_{W,a}(t) \\ w_W(t) \\ w_{W,a}(t) \\ x_{e,p}(t) \\ x_{e,v}(t) \\ \tilde{\mathbf{x}}_{lon}(t) \end{bmatrix}}_{\mathbf{x}_{lon}(t)} &= \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 & \mathbf{0} \\ -\omega_W^2 & -2\bar{\zeta}_W\omega_W & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 1 & \mathbf{0} \\ 0 & 0 & -\omega_W^2 & -2\bar{\zeta}_W\omega_W & \mathbf{0} \\ 0 & 0 & 0 & 0 & \mathbf{a}_{e,p}^T \\ 0 & 0 & 0 & 0 & \mathbf{a}_{e,v}^T \\ \tilde{\mathbf{a}}_{lon,u_W} & \tilde{\mathbf{a}}_{lon,u_{W,a}} & \tilde{\mathbf{a}}_{lon,w_W} & \tilde{\mathbf{a}}_{lon,w_{W,a}} & \tilde{\mathbf{A}}_{lon} \end{bmatrix}}_{\mathbf{A}_{lon}} \mathbf{x}_{lon}(t) + \\
&\quad \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ \omega_W^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \omega_W^2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & b_{e,v} \\ \tilde{\mathbf{b}}_{lon,u_{W,c}} & \tilde{\mathbf{b}}_{lon,w_{W,c}} & \tilde{\mathbf{b}}_{lon,n_{z,c}} \end{bmatrix}}_{\mathbf{B}_{lon}} \mathbf{u}_{lon}(t),
\end{aligned} \tag{8.75}$$

where the system matrix $\mathbf{A}_{lon} \in \mathbb{R}^{46\times46}$ is split into the column vectors $\tilde{\mathbf{a}}_{lon,u_W} \in \mathbb{R}^{40}$, $\tilde{\mathbf{a}}_{lon,u_{W,a}} \in \mathbb{R}^{40}$, $\tilde{\mathbf{a}}_{lon,w_W} \in \mathbb{R}^{40}$, and $\tilde{\mathbf{a}}_{lon,w_{W,a}} \in \mathbb{R}^{40}$, the matrix $\tilde{\mathbf{A}}_{lon} \in \mathbb{R}^{40\times42}$, as well as the entries from the actuator and wind models. Similarly, the input matrix $\mathbf{B}_{lon} \in \mathbb{R}^{46\times3}$ is split into the input column vectors $\tilde{\mathbf{b}}_{lon,u_{W,c}} \in \mathbb{R}^{40}$, $\tilde{\mathbf{b}}_{lon,w_{W,c}} \in \mathbb{R}^{40}$, and $\tilde{\mathbf{b}}_{lon,n_{z,c}} \in \mathbb{R}^{40}$ as well as other entries from the actuator and wind models. Note that the control influences originating from the normal load factor command $n_{z,c}(t)$ and the wind commands $u_{W,c}(t)$ and $w_{W,c}(t)$ are decoupled regarding the influence on the elevator actuator model and the wind dynamics. For the following analysis two different criteria are investigated. These criteria are the normal load factor $j_{n_z}(\mathbf{x}_{lon}(t_f)) = -n_z(t_f)$ and the aerodynamic AoA $j_{\alpha_A}(\mathbf{x}_{lon}(t_f)) = -\alpha_A(t_f)$. The normal load factor may be written

in the form

$$n_z(t_f) = \mathbf{c}_{n_z,A}^T \mathbf{x}_{lon}(t_f), \qquad \mathbf{c}_{n_z,A} \in \mathbb{R}^{46}, \tag{8.76}$$

if the wind is non-zero and in the form

$$n_z(t_f) = \mathbf{c}_{n_z,K}^T \left[x_{e,p}(t), x_{e,v}(t), \tilde{\mathbf{x}}_{lon}^T(t)\right]^T, \qquad \mathbf{c}_{n_z,K} \in \mathbb{R}^{42}, \tag{8.77}$$

in case of no wind. Moreover, the aerodynamic AoA is expressed as

$$\alpha_A(t_f) = \mathbf{c}_{\alpha_A}^T \mathbf{x}_{lon}(t_f), \qquad \mathbf{c}_{\alpha_A} \in \mathbb{R}^{46}. \tag{8.78}$$

In case the actuator position and rate states are bounded by the box constraints

$$x_{e,p,lb} \leq x_{e,p}(t) \leq x_{e,p,ub}, \tag{8.79}$$

$$x_{e,v,lb} \leq x_{e,v}(t) \leq x_{e,v,ub}, \tag{8.80}$$

the following Cat II type clearance problems are obtained:

$$\begin{aligned} \underset{\mathbf{u}_{lon}(t) \in \mathcal{U}_{lon}}{\text{minimize}} \quad & -n_z(t_f)/-\alpha_A(t_f) \\ \text{subject to} \quad & \dot{\mathbf{x}}_{lon}(t) - \mathbf{A}_{lon}\mathbf{x}_{lon}(t) - \mathbf{B}_{lon}\mathbf{u}_{lon}(t) = \mathbf{0}, \\ & \mathbf{x}_{lon}(t_0) = \mathbf{0}, \\ & x_{e,p,lb} \leq x_{e,p}(t) \leq x_{e,p,ub}, \\ & x_{e,v,lb} \leq x_{e,v}(t) \leq x_{e,v,ub}, \\ & t \in I^t \end{aligned} \tag{8.81}$$

Lateral Model

For the lateral dynamic equations the state vector $\mathbf{x}_{lat} \in \mathbb{R}^{36}$ contains the lateral wind velocity $v_W(t)$, the aileron position state $x_{a,p}(t)$ and rate state $x_{a,v}(t)$, the rudder position state $x_{r,p}(t)$ and rate state $x_{r,v}(t)$, the kinematic AoS $\beta_K(t)$, the roll angle $\Phi(t)$, the yaw angle $\Psi(t)$, the yaw rate $r(t)$, the roll rate $p(t)$, the error controller states $e_\Phi(t)$ and $e_\Psi(t)$, as well as other internal controller and sensor model states collected in the vector $\mathbf{x}_{F,lat}(t) \in \mathbb{R}^{23}$:

$$\begin{aligned} \mathbf{x}_{lat}(t) := [&v_W(t), v_{W,a}(t), x_{a,p}(t), x_{r,p}(t), x_{a,v}(t), x_{r,v}(t), \\ &\beta_K(t), \Phi(t), \Psi(t), r_K(t), p_K(t), e_\Phi(t), e_\Psi(t), \mathbf{x}_{F,lat}^T(t)]^T \end{aligned} \tag{8.82}$$

The control vector $\mathbf{u}_{lat}(t) \in \mathbb{R}^3$

$$\mathbf{u}_{lat}(t) := \left[v_{W,c}(t), n_{y,c}(t), \Phi_c(t)\right]^T, \tag{8.83}$$

contains the lateral wind command $v_{W,c}(t)$ constrained by box bounds (cf. constraint (8.62)), the lateral load factor command $n_{y,c}(t)$ bounded by

$$-0.1 \leq n_{y,c}(t) \leq 0.1, \tag{8.84}$$

as well as the roll angle command $\Phi_c(t)$ limited to

$$-45\,\mathrm{deg} \leq \Phi_c(t) \leq 45\,\mathrm{deg}\,, \tag{8.85}$$

which define the control set $\mathcal{U}_{lat}$ for the lateral plane. Similar to the longitudinal plane, let

$$\tilde{\mathbf{x}}_{lat}(t) := \left[\beta_K(t), \Phi(t), \Psi(t), r_K(t), p_K(t), e_\Phi(t), e_\Psi(t), \mathbf{x}_{F,lat}^T(t)\right]^T, \tag{8.86}$$

collect all states besides the actuator rate and position states and the wind state. The dynamic equations for the aileron and rudder servo positions, $x_{a,p}(t)$ and $x_{r,p}(t)$, are:

$$\dot{x}_{a,p}(t) = \mathbf{a}_{a,p}^T \left[x_{a,p}(t), x_{r,p}(t), x_{a,v}(t), x_{r,v}(t), \tilde{\mathbf{x}}_{lat}^T(t)\right]^T = x_{a,v}(t), \qquad \mathbf{a}_{a,p}^T := [0,0,1,0,\mathbf{0}]\,, \tag{8.87}$$

$$\dot{x}_{r,p}(t) = \mathbf{a}_{r,p}^T \left[x_{a,p}(t), x_{r,p}(t), x_{a,v}(t), x_{r,v}(t), \tilde{\mathbf{x}}_{lat}^T(t)\right]^T = x_{r,v}(t), \qquad \mathbf{a}_{r,p}^T := [0,0,0,1,\mathbf{0}] \tag{8.88}$$

For the aileron and rudder servo rate states, i.e. $x_{a,v}(t)$ and $x_{r,v}(t)$, the dynamic equations are of the form

$$\dot{x}_{a,v}(t) = \mathbf{a}_{a,v}^T \left[x_{a,p}(t), x_{r,p}(t), x_{a,v}(t), x_{r,v}(t), \tilde{\mathbf{x}}_{lat}^T(t)\right]^T + b_{a,n_{y,c}} n_{y,c}(t) + b_{a,\Phi_c}\Phi_c(t),\ \mathbf{a}_{a,v} \in \mathbb{R}^{34}, \tag{8.89}$$

$$\dot{x}_{r,v}(t) = \mathbf{a}_{r,v}^T \left[x_{a,p}(t), x_{r,p}(t), x_{a,v}(t), x_{r,v}(t), \tilde{\mathbf{x}}_{lat}^T(t)\right]^T + b_{r,n_{y,c}} n_{y,c}(t) + b_{r,\Phi_c}\Phi_c(t),\ \mathbf{a}_{r,v} \in \mathbb{R}^{34}, \tag{8.90}$$

with the scalar control influences $b_{a,n_{y,c}}$, b_{a,Φ_c}, $b_{r,n_{y,c}}$, and b_{r,Φ_c}. Together with the equation for the lateral wind acceleration these dynamic equations may be written jointly as:

$$\begin{bmatrix} \dot{v}_{W,a}(t) \\ \dot{x}_{a,v}(t) \\ \dot{x}_{r,v}(t) \end{bmatrix} = \begin{bmatrix} -\omega_W^2 & -2\bar{\zeta}_W\omega_W & \mathbf{0} \\ 0 & 0 & \mathbf{a}_{a,v}^T \\ 0 & 0 & \mathbf{a}_{r,v}^T \end{bmatrix} \mathbf{x}_{lat}(t) + \begin{bmatrix} \omega_W^2 & 0 & 0 \\ 0 & b_{a,n_{y,c}} & b_{a,\Phi_c} \\ 0 & b_{r,n_{y,c}} & b_{r,\Phi_c} \end{bmatrix} \begin{bmatrix} v_{W,c}(t) \\ n_{y,c}(t) \\ \Phi_c(t) \end{bmatrix} \tag{8.91}$$

Note that the input equations for the aileron and rudder states are coupled and the full rank matrix

$$\mathbf{B}_{a,v} := \begin{bmatrix} b_{a,n_{y,c}} & b_{a,\Phi_c} \\ b_{r,n_{y,c}} & b_{r,\Phi_c} \end{bmatrix}, \tag{8.92}$$

expresses the control allocation in the lateral channel. Thus, the equations governing the lateral motion are of the form

$$\frac{d}{dt}\underbrace{\begin{bmatrix} v_W(t) \\ v_{W,a}(t) \\ x_{a,p}(t) \\ x_{r,p}(t) \\ x_{a,v}(t) \\ x_{r,v}(t) \\ \tilde{\mathbf{x}}_{lat}(t) \end{bmatrix}}_{\mathbf{x}_{lat}(t)} = \underbrace{\begin{bmatrix} 0 & 1 & \mathbf{0} \\ -\omega_W^2 & -2\bar{\zeta}_W\omega_W & \mathbf{0} \\ 0 & 0 & \mathbf{a}_{a,p}^T \\ 0 & 0 & \mathbf{a}_{r,p}^T \\ 0 & 0 & \mathbf{a}_{a,v}^T \\ 0 & 0 & \mathbf{a}_{r,v}^T \\ \tilde{\mathbf{a}}_{lat,v_W} & \tilde{\mathbf{a}}_{lat,v_{W,a}} & \tilde{\mathbf{A}}_{lat} \end{bmatrix}}_{\mathbf{A}_{lat}} \mathbf{x}_{lat}(t) + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ \omega_W^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & b_{a,n_{y,c}} & b_{a,\Phi_c} \\ 0 & b_{r,n_{y,c}} & b_{r,\Phi_c} \\ \tilde{\mathbf{b}}_{lat,v_{W,c}} & \tilde{\mathbf{b}}_{lat,n_{y,c}} & \tilde{\mathbf{b}}_{lat,\Phi_c} \end{bmatrix}}_{\mathbf{B}_{lat}} \mathbf{u}_{lat}(t), \tag{8.93}$$

where the system matrix $\mathbf{A}_{lat} \in \mathbb{R}^{36\times 36}$ is split into the column vectors $\tilde{\mathbf{a}}_{lat,v_W} \in \mathbb{R}^{30}$ and $\tilde{\mathbf{a}}_{lat,v_{W,a}} \in \mathbb{R}^{30}$, the matrix $\tilde{\mathbf{A}}_{lat} \in \mathbb{R}^{30\times 34}$, as well as the entries from the actuator and wind models. Similarly, the input matrix $\mathbf{B}_{lat} \in \mathbb{R}^{36\times 3}$ is split into the input column vectors $\tilde{\mathbf{b}}_{lat,v_{W,c}} \in \mathbb{R}^{30}$, $\tilde{\mathbf{b}}_{lat,n_{y,c}} \in \mathbb{R}^{30}$, and $\tilde{\mathbf{b}}_{lat,\Phi_c} \in \mathbb{R}^{30}$ as well as the contributions from the actuator and wind models. As for the longitudinal model two different criteria are investigated, namely the lateral load factor $j_{n_y}(\mathbf{x}_{lat}(t_f)) = -n_y(t_f)$ and the aerodynamic AoS $j_{\beta_A}(\mathbf{x}_{lat}(t_f)) = -\beta_A(t_f)$. The load factor may be written in the form

$$n_y(t_f) = \mathbf{c}_{n_y,A}^T \mathbf{x}_{lat}(t_f), \qquad \mathbf{c}_{n_y,A} \in \mathbb{R}^{35}, \tag{8.94}$$

if the wind is non-zero and in the form

$$n_y(t_f) = \mathbf{c}_{n_y,K}^T \left[x_{a,p}(t), x_{r,p}(t), x_{a,v}(t), x_{r,v}(t), \tilde{\mathbf{x}}_{lat}^T(t)\right]^T, \qquad \mathbf{c}_{n_y,K} \in \mathbb{R}^{34}, \tag{8.95}$$

in case of no wind. Furthermore, the aerodynamic AoS is given as

$$\beta_A(t_f) = \mathbf{c}_{\beta_A}^T \mathbf{x}_{lat}(t_f), \qquad \mathbf{c}_{\beta_A} \in \mathbb{R}^{35}. \tag{8.96}$$

For a Cat II type model where the actuator rate and position state are bounded by the box constraints

$$x_{a,p,lb} \leq x_{a,p}(t) \leq x_{a,p,ub}, \tag{8.97}$$

$$x_{r,p,lb} \leq x_{r,p}(t) \leq x_{r,p,ub}, \tag{8.98}$$

$$x_{a,v,lb} \leq x_{a,v}(t) \leq x_{a,v,ub}, \tag{8.99}$$

$$x_{r,v,lb} \leq x_{r,v}(t) \leq x_{r,v,ub}, \tag{8.100}$$

the problem formulations thus reads as follows:

$$
\begin{aligned}
\underset{\mathbf{u}_{lat}(t) \in \mathcal{U}_{lat}}{\text{minimize}} \quad & -n_y(t_f)/-\beta_A(t_f) \\
\text{subject to} \quad & \dot{\mathbf{x}}_{lat}(t) - \mathbf{A}_{lat}\mathbf{x}_{lat}(t) - \mathbf{B}_{lat}\mathbf{u}_{lat}(t) = \mathbf{0}, \\
& \mathbf{x}_{lat}(t_0) = \mathbf{0}, \\
& x_{a,p,lb} \leq x_{a,p}(t) \leq x_{a,p,ub}, \\
& x_{r,p,lb} \leq x_{r,p}(t) \leq x_{r,p,ub}, \\
& x_{a,v,lb} \leq x_{a,v}(t) \leq x_{a,v,ub}, \\
& x_{r,v,lb} \leq x_{r,v}(t) \leq x_{r,v,ub}, \\
& t \in I^t
\end{aligned}
\tag{8.101}
$$

8.2.2 Numerical Examples

Here and in what follows, the numerical examples are computed using the Counter Optimization Library (COLIBRY) developed by the author (Contribution IV) and the LP solver employed for all studies is CPLEX 12.9 [57] (Primal Simplex, Dual Simplex, Interior Point) with standard settings. For these numerical examples the linearized models for the longitudinal and lateral plane are obtained from a trim point with reference height $h_{ref} \approx 2500\ m$ and reference velocity $V_{ref} \approx 67\ m/s$. Recall that all quantities which are presented in the following are relative to the respective trim values. The fixed final time t_f for the analysis is chosen as approximately five times the closed-loop short period time constant $T_{SP} \approx 1.19\ s$, i.e. $t_f \approx 6\ s$. A Backward Euler discretization with a constant discretization step size of $h = 0.005\ s$ is used for all examples. Moreover, the switching functions are estimated based on the results from the discrete Minimum Principle in Ref. [52]. In all numerical results the signs of the switching function values are indicated in red for negative values ($< 10^{-10}$) and green for positive values ($> 10^{-10}$). All other values, which are regarded as zero, are colored black. Moreover, the full set of rigid body states and error controller states are provided as additional material in App. A. Note that in this chapter only the linear case is presented. The validity of the linearized models regarding the qualitative and quantitative behavior is discussed in Sec. 9.1 using the Cat III intrusive approach. First, the longitudinal model without wind is under investigation. The Cat II type problems for maximizing the normal load factor or the kinematic AoA are stated as follows:

$$
\begin{aligned}
\underset{n_{z,c}(t)}{\text{minimize}} \quad & -n_z(t_f)/-\alpha_K(t_f) \\
\text{subject to} \quad & \frac{d}{dt}\begin{bmatrix} x_{e,p}(t) \\ x_{e,v}(t) \\ \tilde{\mathbf{x}}_{lon}(t) \end{bmatrix} - \begin{bmatrix} \mathbf{a}_{e,p}^T \\ \mathbf{a}_{e,v}^T \\ \tilde{\mathbf{A}}_{lon} \end{bmatrix} \begin{bmatrix} x_{e,p}(t) \\ x_{e,v}(t) \\ \tilde{\mathbf{x}}_{lon}(t) \end{bmatrix} - \begin{bmatrix} 0 \\ b_{e,v} \\ \tilde{\mathbf{b}}_{lon,n_{z,c}} \end{bmatrix} n_{z,c}(t) = \mathbf{0}, \\
& \begin{bmatrix} x_{e,p}(0) \\ x_{e,v}(0) \\ \tilde{\mathbf{x}}_{lon}(0) \end{bmatrix} = \mathbf{0}, \\
& -0.5 \leq n_{z,c}(t) \leq 0.5, \\
& x_{e,p,lb} \leq x_{e,p}(t) \leq x_{e,p,ub}, \\
& x_{e,v,lb} \leq x_{e,v}(t) \leq x_{e,v,ub}, \\
& t \in [0, t_f]
\end{aligned}
\tag{8.102}
$$

At first, the actuator limits are set to $x_{e,p,lb}, x_{e,v,lb} = -\infty$ and $x_{e,p,ub}, x_{e,v,ub} = \infty$ which yields a Cat I type problem. The optimal solutions for the trajectory of the normal load

factor $n_z(t)$, as well as the load factor command history and corresponding switching function are depicted in Fig. 8.5. As predicted by the theoretical analysis for Cat I problems (cf. Sec. 8.1.1) the worst-case command is non-singular, bang-bang. In particular, the switching function only exhibits isolated zeros and the time points where it switches sign correspond to the time points where the control changes from one bound to the other. These switching time points are:

$$\mathbf{t}_s^{n_z} \approx [3.32\ s, 3.74\ s, 4.62\ s, 5.06\ s, 5.95\ s] \tag{8.103}$$

Each pulse has a total length ($t_{s,0}^{n_z} \rightarrow t_{s,2}^{n_z}$ and $t_{s,2}^{n_z} \rightarrow t_{s,4}^{n_z}$) which approximately matches the short period time constant T_{SP}. Moreover, the width of the pulses ($t_{s,0}^{n_z} \rightarrow t_{s,1}^{n_z}$ and $t_{s,2}^{n_z} \rightarrow t_{s,3}^{n_z}$) is $\approx T_{SP}/3$. For the kinematic AoA a similar structure is obtained. The control is again of bang-bang type and the switching time points of the control match the location of the isolated zeros of the corresponding switching function. These switching time points are:

$$\mathbf{t}_s^{\alpha} \approx [4.7\ s, 5.04\ s] \tag{8.104}$$

Only one pulse is observed for which the total length ($t_{s,0}^{\alpha} \rightarrow t_f$) approximately matches the short period time constant T_{SP} and the width of the pulse ($t_{s,0}^{\alpha} \rightarrow t_{s,1}^{\alpha}$) is $\approx T_{SP}/3$ as for the maximization of the normal load factor. The cases with actuator limits for both cost functions are depicted in Fig. 8.7 and Fig. 8.8. Note that the elevator rate constraint becomes active (the position limit is not reached for both cases). As predicted by the theoretical analysis, the control is bang-bang with singular parts on time intervals where the state constraint on the elevator rate state is active. On these intervals the control takes values in the interior of the admissible control set, i.e. the values of the boundary control $n_{z,c,b}(t)$. This boundary control of the normal load factor command can be expressed based on Eq. (8.27) as follows:

$$n_{z,c,b}(t) := -\frac{1}{b_{e,v}} \mathbf{a}_{e,v}^T \begin{bmatrix} x_{e,p}(t) \\ x_{e,v}(t) \\ \tilde{\mathbf{x}}_{lon}(t) \end{bmatrix} \tag{8.105}$$

A detailed view comparing the boundary control for maximizing the load factor and AoA obtained from the solution of the discretized problem and the control predicted from the theoretical analysis is presented in Fig. 8.9.

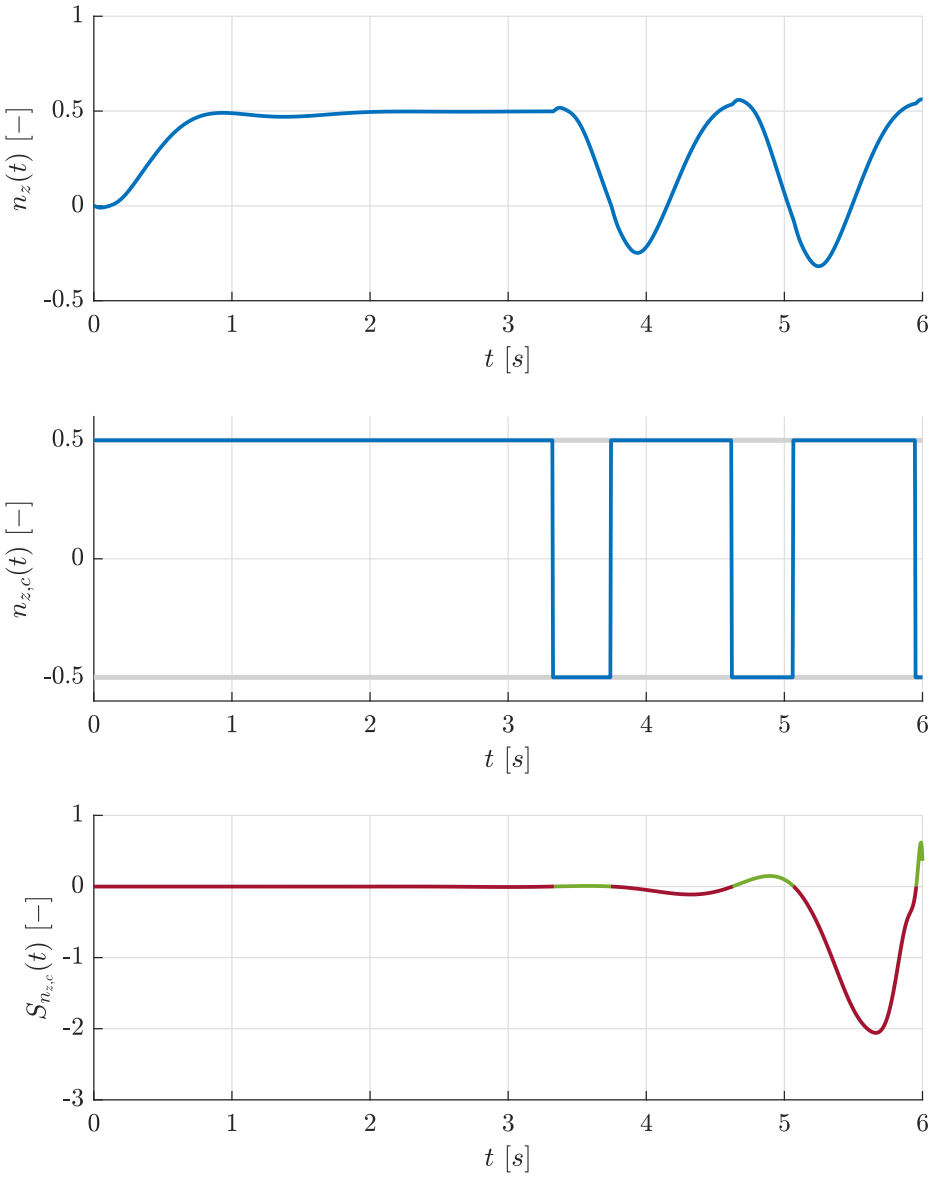


Figure 8.5: Optimal results for the longitudinal plane using a Cat I model for maximizing $n_z(t_f), t_f = 6\ s$ including the normal load factor $n_z(t)$, the normal load factor command $n_{z,c}(t)$, and the switching function corresponding to the normal load factor command $\mathcal{S}_{n_{z,c}}(t)$.

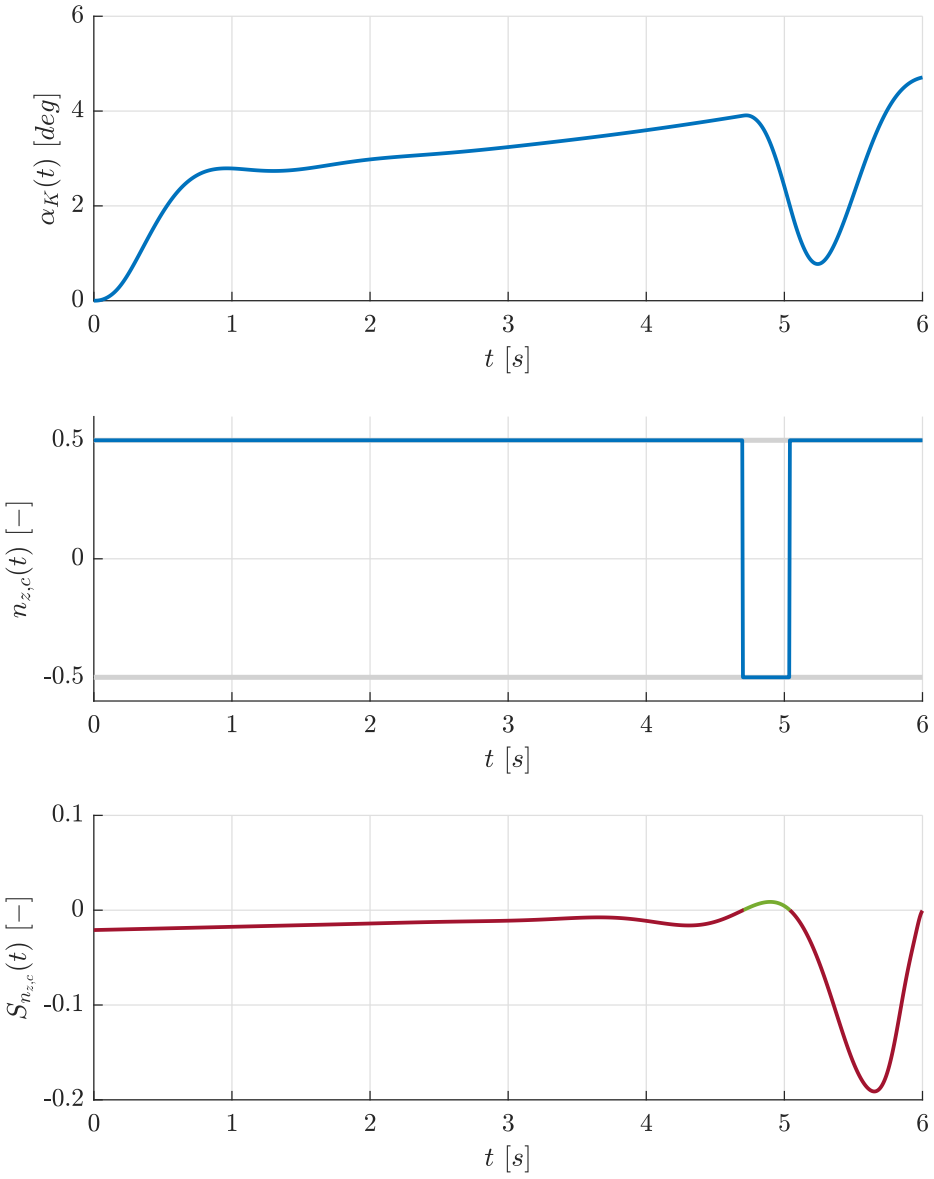


Figure 8.6: Optimal results for the longitudinal plane using a Cat I model for maximizing $\alpha_K(t_f), t_f = 6\ s$ including the kinematic AoA $\alpha_K(t)$, the normal load factor command $n_{z,c}(t)$, and the switching function corresponding to the normal load factor command $\mathcal{S}_{n_{z,c}}(t)$.

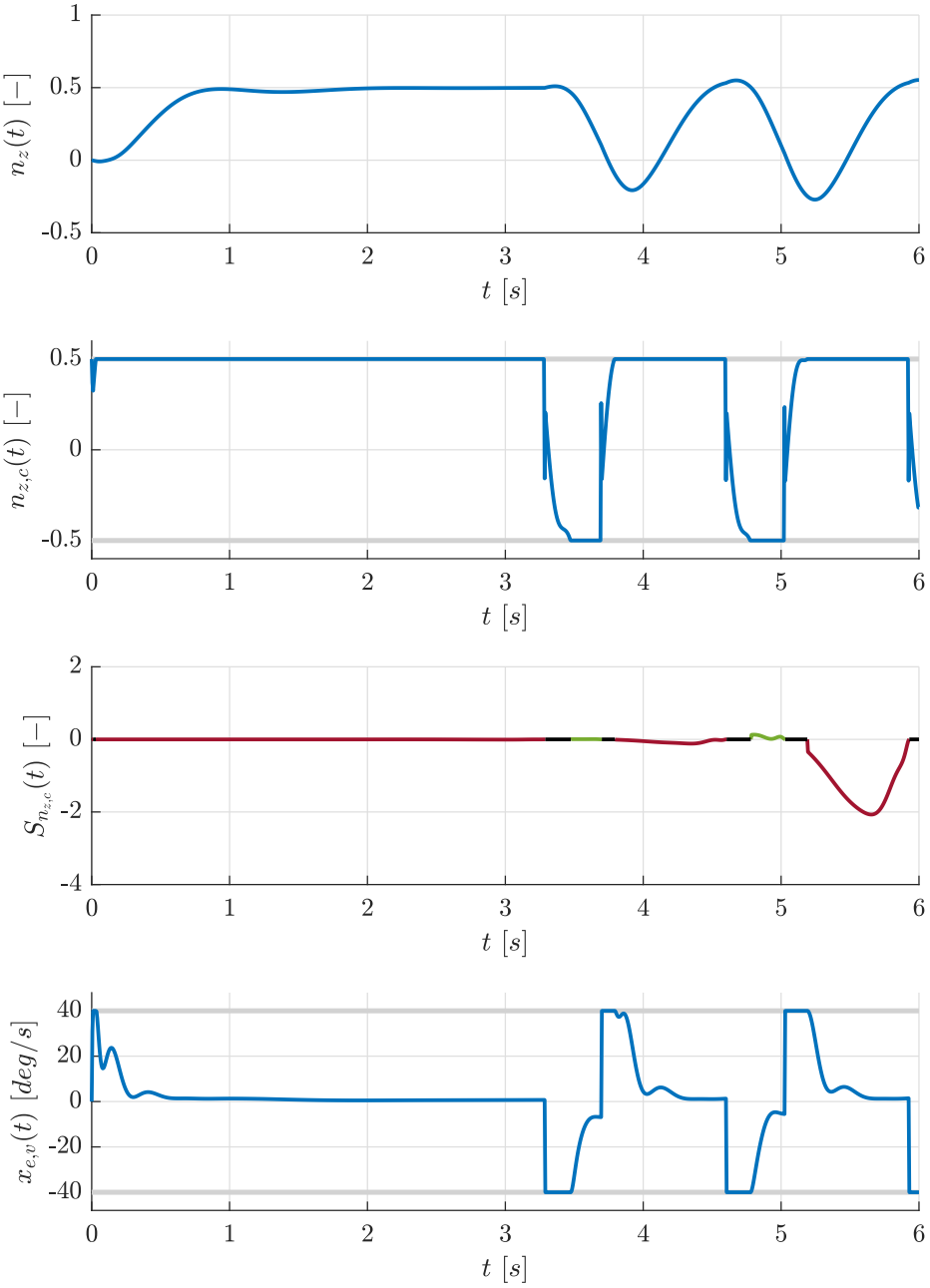


Figure 8.7: Optimal results for the longitudinal plane using a Cat II model for maximizing $n_z(t_f), t_f = 6\ s$ including the normal load factor $n_z(t)$, the normal load factor command $n_{z,c}(t)$, the switching function corresponding to the normal load factor command $\mathcal{S}_{n_{z,c}}(t)$, and the elevator rate $x_{e,v}(t)$.

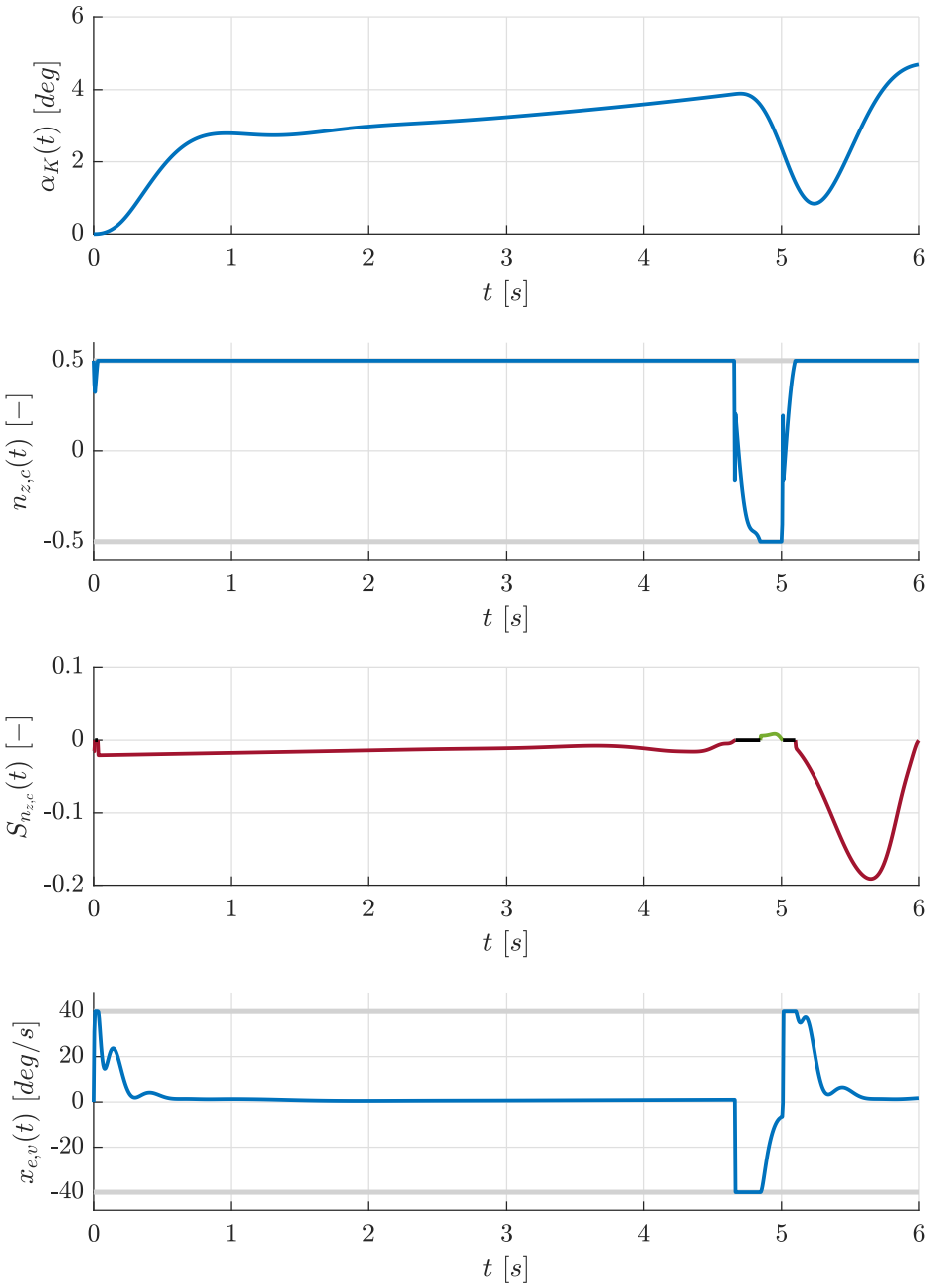


Figure 8.8: Optimal results for the longitudinal plane using a Cat II model for maximizing $\alpha_K(t_f), t_f = 6\ s$ including the kinematic AoA $\alpha_K(t)$, the normal load factor command $n_{z,c}(t)$, the switching function corresponding to the normal load factor command $\mathcal{S}_{n_{z,c}}(t)$, and the elevator rate $x_{e,v}(t)$.

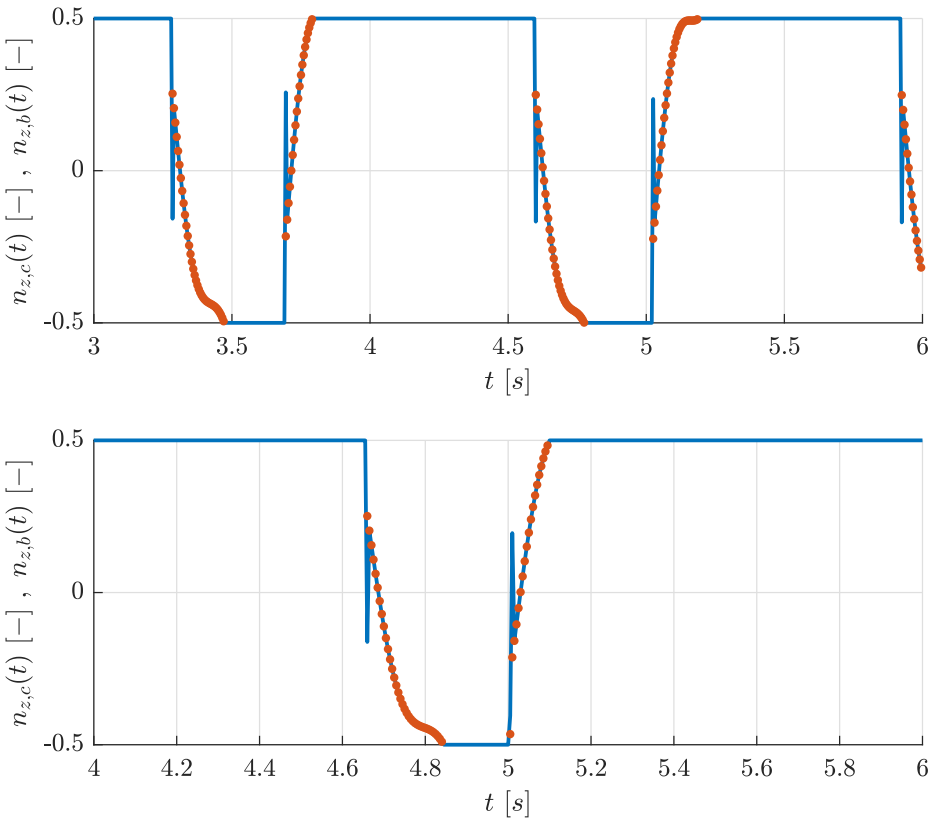


Figure 8.9: Detailed view of the boundary control on rate constrained arcs of the elevator for the maximization of the normal load factor (top) and AoA (bottom). The control values obtained from the discretized problem ($n_{z,c}(t)$, blue) are compared to the boundary control ($n_{z,b}(t)$, red dots) according to the theoretical analysis.

The numerical results appear to be in coherence with the theoretical analysis. Note that in all Cat II results (cf. Figs. 8.7 and 8.8) there are overshoots visible in the control at the entry time points of the state constrained arcs. This is an inherent effect originating from the discretization of the problem as the entry time points of the state constrained arcs where the control jumps are not exactly resolved by the time grid. Note further that the values of the AoA and normal load factor for the Cat II model are lower compared to the Cat I model. This is expected as the feasible region for Cat II type problems is reduced by the addition of state constraints corresponding to the limits of the servomechanism.

For the lateral model without wind the Cat II problems for maximizing the lateral load factor or the AoS are defined as follows:

$$
\begin{aligned}
&\underset{n_{y,c}(t),\,\Phi_c(t)}{\text{minimize}} && -n_y(t_f)/-\beta_K(t_f)\\
&\text{subject to} && \frac{d}{dt}\begin{bmatrix} x_{a,p}(t)\\ x_{r,p}(t)\\ x_{a,v}(t)\\ x_{r,v}(t)\\ \tilde{\mathbf{x}}_{lat}\end{bmatrix} - \begin{bmatrix} \mathbf{a}_{a,p}^T\\ \mathbf{a}_{r,p}^T\\ \mathbf{a}_{a,v}^T\\ \mathbf{a}_{r,v}^T\\ \tilde{\mathbf{A}}_{lat}\end{bmatrix}\begin{bmatrix} x_{a,p}(t)\\ x_{r,p}(t)\\ x_{a,v}(t)\\ x_{r,v}(t)\\ \tilde{\mathbf{x}}_{lat}\end{bmatrix} - \begin{bmatrix} 0 & 0\\ 0 & 0\\ b_{a,n_{y,c}} & b_{a,\Phi_c}\\ b_{r,n_{y,c}} & b_{r,\Phi_c}\\ \tilde{\mathbf{b}}_{lat,n_{y,c}} & \tilde{\mathbf{b}}_{lat,\Phi_c}\end{bmatrix}\begin{bmatrix} n_{y,c}(t)\\ \Phi_c(t)\end{bmatrix} = \mathbf{0},\\
&&& \begin{bmatrix} x_{a,p}(0)\\ x_{r,p}(0)\\ x_{a,v}(0)\\ x_{r,v}(0)\\ \tilde{\mathbf{x}}_{lat}(0)\end{bmatrix} = \mathbf{0},\\
&&& -0.1 \le n_{y,c}(t) \le 0.1,\\
&&& -45\ deg \le \Phi_c(t) \le 45\ deg,\\
&&& x_{a,p,lb} \le x_{a,p}(t) \le x_{a,p,ub},\\
&&& x_{r,p,lb} \le x_{r,p}(t) \le x_{r,p,ub},\\
&&& x_{a,v,lb} \le x_{a,v}(t) \le x_{a,v,ub},\\
&&& x_{r,v,lb} \le x_{r,v}(t) \le x_{r,v,ub},\\
&&& t \in [0, t_f]
\end{aligned}
\tag{8.106}
$$

First, only the roll angle command $\Phi_c(t)$ is used to solve the clearance problem for the lateral plane. It is found that the maximization of the lateral load factor is effective whereas the values obtained from the maximization of the AoS are fairly small. This is illustrated for the Cat I version of the problem, i.e. without state constraints on the aileron and rudder states ($x_{a,p,lb}, x_{a,v,lb}, x_{r,p,lb}, x_{r,v,lb} = -\infty$, $x_{a,p,ub}, x_{a,v,ub}, x_{r,p,ub}, x_{r,v,ub} = \infty$) in Fig. 8.10 and Fig. 8.11. In both cases the control is regular, i.e. of bang-bang type. Despite the in-effectiveness of the roll angle command for the maximization of the AoS this case is interesting to illustrate typical numerical effects regarding the optimization with fixed final time. It is observed that a weak dependence of the cost function with respect to control actions in the beginning of the time interval can results in high frequency switches of the control.

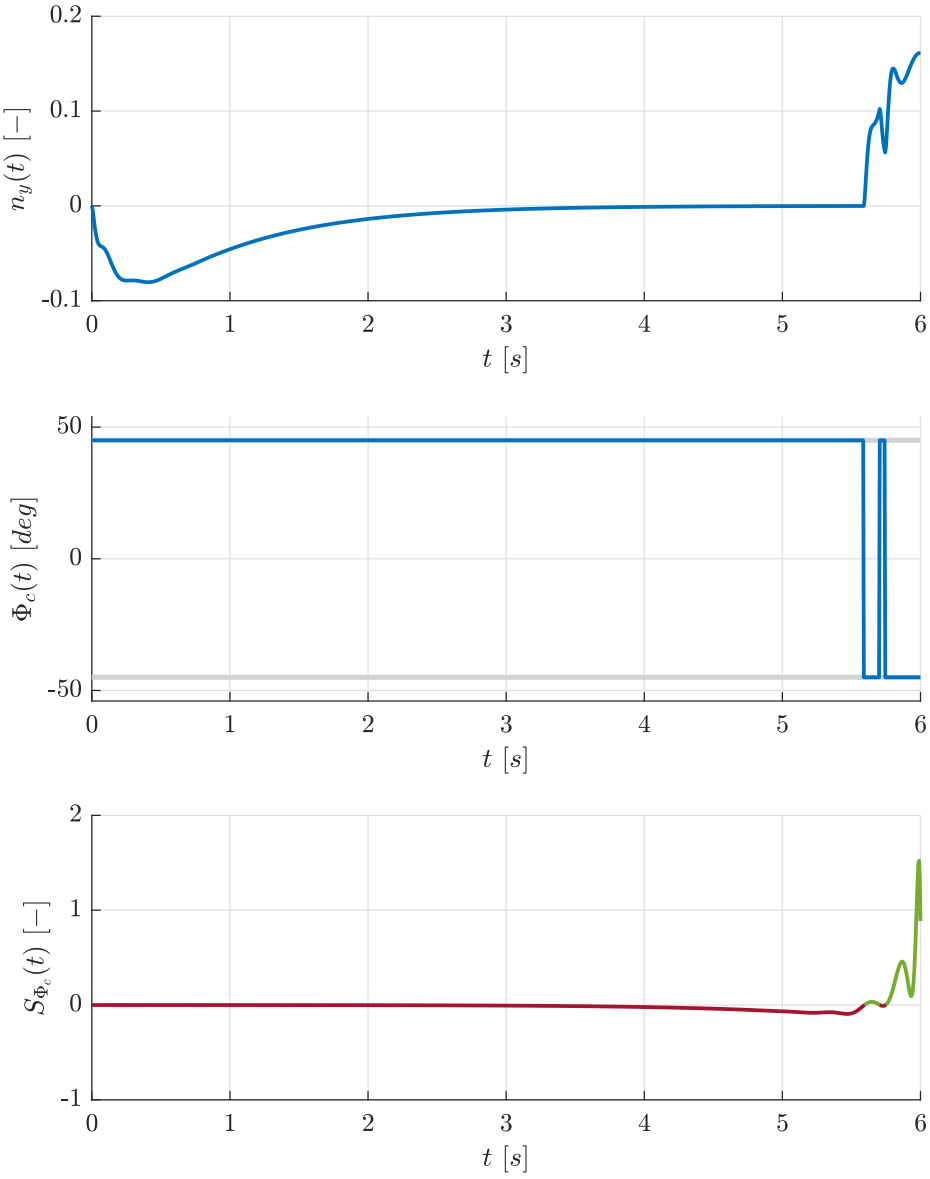


Figure 8.10: Optimal results for the lateral plane using a Cat I model for maximizing $n_y(t_f), t_f = 6\ s$ including the lateral load factor $n_y(t)$, the roll angle command $\Phi_c(t)$, and the switching function corresponding to the roll angle command $\mathcal{S}_{\Phi_c}(t)$.

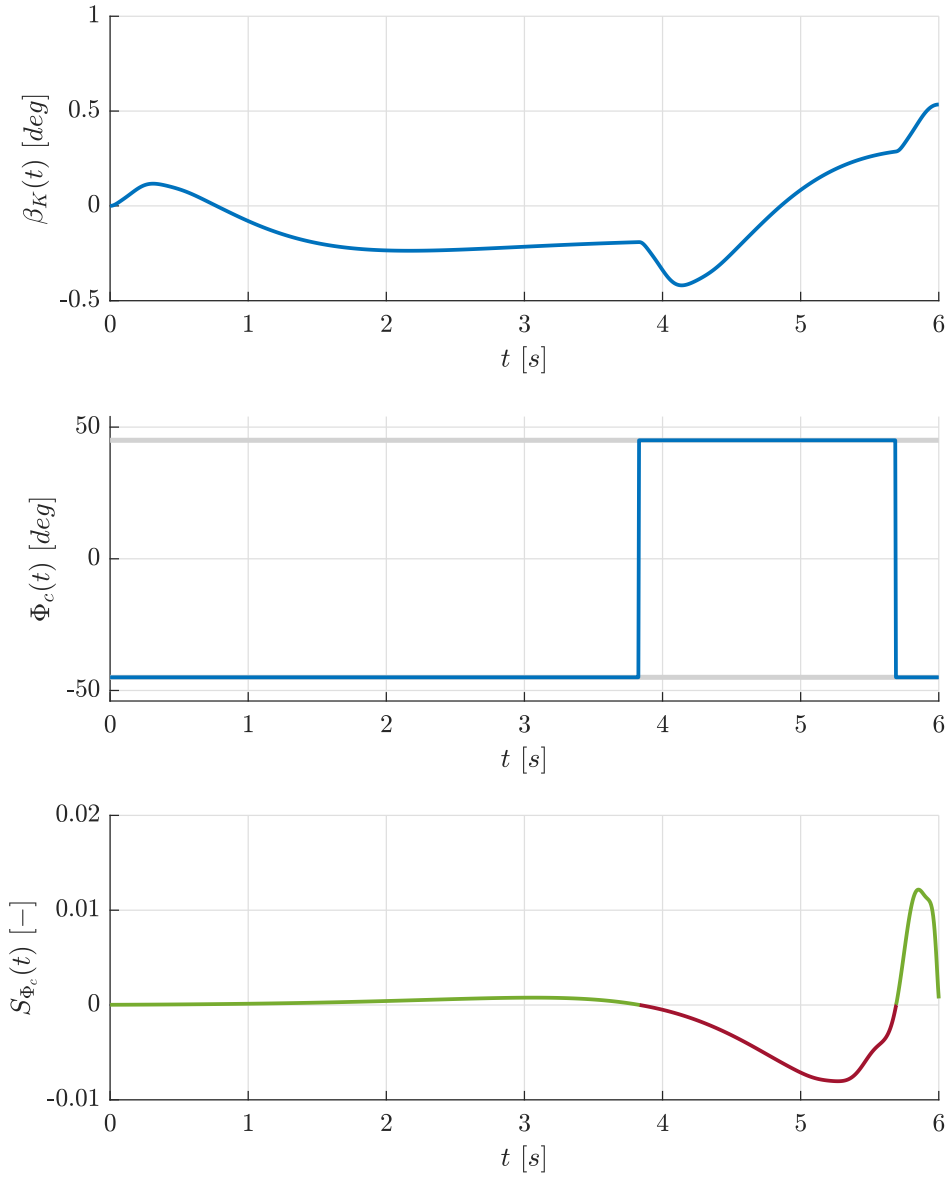


Figure 8.11: Optimal results for the lateral plane using a Cat I model for maximizing $\beta_K(t_f), t_f = 6\ s$ including the kinematic AoS $\beta_K(t)$, the roll angle command $\Phi_c(t)$, and the switching function corresponding to the roll angle command $\mathcal{S}_{\Phi_c}(t)$.

This chattering effect becomes apparent if the length of the fixed time interval $[0, t_f]$ is increased. Recall that as discussed in Sec. 8.1 the cost function is non-decreasing with the final time, if the system is starting from an equilibrium point. However, one direct result of fixing the final time is that the problem becomes very in-sensitive with respect to control actions at the beginning of the time interval if the interval is too long. For the value of the cost function this is not detrimental because the solution of the LP is global. Loosely speaking - it simply matters very little for the value of the cost function at the final time point what exactly the control is doing in the beginning of the time interval. The cost and control histories for maximizing $\beta_K(t_f)$ with $t_f = 7\ s$ are presented in Fig. 8.12. Observe the chattering effect of the control (blue line). If this effect is undesired, e.g. if not solely the value of the clearance criterion itself but a clean control history is of interest, it is possible to force the control to the trim value on intervals where the sensitivity of the control is low. This can be achieved by the introduction of a penalty term in the cost function of the form

$$\begin{aligned} \underset{\mathbf{x}^{[0]}, \mathbf{x}^{[i]}, \mathbf{u}^{[i]}, i=1,\dots,N}{\text{minimize}} \quad & \mathbf{c}^T \mathbf{x}^{[N]} + \epsilon \sum_{i=1}^{N} \sum_{q=0}^{n_u - 1} |u_q^{[i]}| \\ \text{subject to} \quad & \mathbf{A}_d \mathbf{x}^{[i]} + \mathbf{B}_d \mathbf{u}^{[i]} - \mathbf{x}^{[i-1]} = \mathbf{0},\ i = 1, \dots, N, \\ & \mathbf{x}^{[0]} = \mathbf{0}, \\ & x_{j,lb} \leq x_j^{[i]} \leq x_{j,ub},\ i = 1, \dots, N, \forall j \in \mathcal{I}_b, \\ & \mathbf{u}_{lb} \leq \mathbf{u}^{[i]} \leq \mathbf{u}_{ub},\ i = 1, \dots, N, \end{aligned} \tag{8.107}$$

with a small constant $\epsilon << 1$. This reformulation is not covered in the theoretical analysis and as such introduces inaccuracies in the numerical results compared to the theoretical results. In fact, the control may take singular values in this case. Note that despite the absolute value function in the cost this optimization problem can still be solved using LP methods. For this purpose the control is split into its positive part $\mathbf{u}_+^{[i]}$ and its negative part $\mathbf{u}_-^{[i]}$. As the control is assumed bounded and given relative to its trim value $\mathbf{u}_0$ this transformation is readily accomplished. Based on its positive and negative part the control $\mathbf{u}^{[i]}$ can be expressed as

$$\mathbf{u}^{[i]} = -\mathbf{u}_-^{[i]} + \mathbf{u}_+^{[i]}, \tag{8.108}$$

and the problem formulation reads:

$$\begin{aligned}
\underset{\mathbf{x}^{[0]},\,\mathbf{x}^{[i]},\,\mathbf{u}^{[i]}_{+/-},\,i=1,\ldots,N}{\text{minimize}} \quad & \mathbf{c}^T\mathbf{x}^{[N]} + \epsilon \sum_{i=1}^{N} \mathbf{1}\left(\mathbf{u}^{[i]}_{+} + \mathbf{u}^{[i]}_{-}\right) \\
\text{subject to} \quad & \mathbf{A}_d\mathbf{x}^{[i]} + \mathbf{B}_d\mathbf{u}^{[i]} - \mathbf{x}^{[i-1]} = \mathbf{0},\, i = 1,\ldots,N, \\
& \mathbf{x}^{[0]} = \mathbf{0}, \\
& x_{j,lb} \leq x_j^{[i]} \leq x_{j,ub},\, i = 1,\ldots,N, \forall j \in \mathcal{I}_b, \\
& \mathbf{0} \leq \mathbf{u}^{[i]}_{+} \leq \mathbf{u}_{ub},\; i = 1,\ldots,N, \\
& \mathbf{0} \leq \mathbf{u}^{[i]}_{-} \leq -\mathbf{u}_{lb},\; i = 1,\ldots,N, \\
& \mathbf{u}^{[i]} = -\mathbf{u}^{[i]}_{-} + \mathbf{u}^{[i]}_{+},\; i = 1,\ldots,N
\end{aligned} \tag{8.109}$$

It is noteworthy, that from a clearance perspective the reduction ϵ_0 of the clearance criterion value when removing the penalty has a conservative bound of

$$\epsilon_0 \leq \epsilon \sum_{i=1}^{N} \mathbf{1}\left(\mathbf{u}^{[i]}_{+} + \mathbf{u}^{[i]}_{-}\right), \tag{8.110}$$

which for small ϵ is typically negligible compared to the value of the clearance criterion itself. The other way round, if ϵ is chosen based on an acceptable error tolerance ϵ_0 for the respective quantity under investigation the reformulation presents a valid extension to the problem formulation from a practical perspective. Moreover, the penalty may only be introduced for the variables corresponding to a time interval $t \in [t_0, t_2], t_0 \leq t_2 \leq t_f$. In Fig. 8.12 the numerical result obtained from following this strategy with $t_2 = 2T_{SP} \approx 2.38\ s$ are presented (black lines). Note that the control takes the (singular) values of the trim control for the first interval $t \in [t_0, t_1], t_1 < t_2$ and is regular on the remaining time interval $t \in [t_1, t_f]$. This result solidifies the argument presented in Sec. 8.1 regarding the non-decreasing value of the cost function with increasing values of the final time t_f. By employing the penalty approach the solver is essentially pushed towards a particular solution which is to hold the system in the trim condition for the first t_1 seconds and then applies the worst-case control to minimize the cost function.

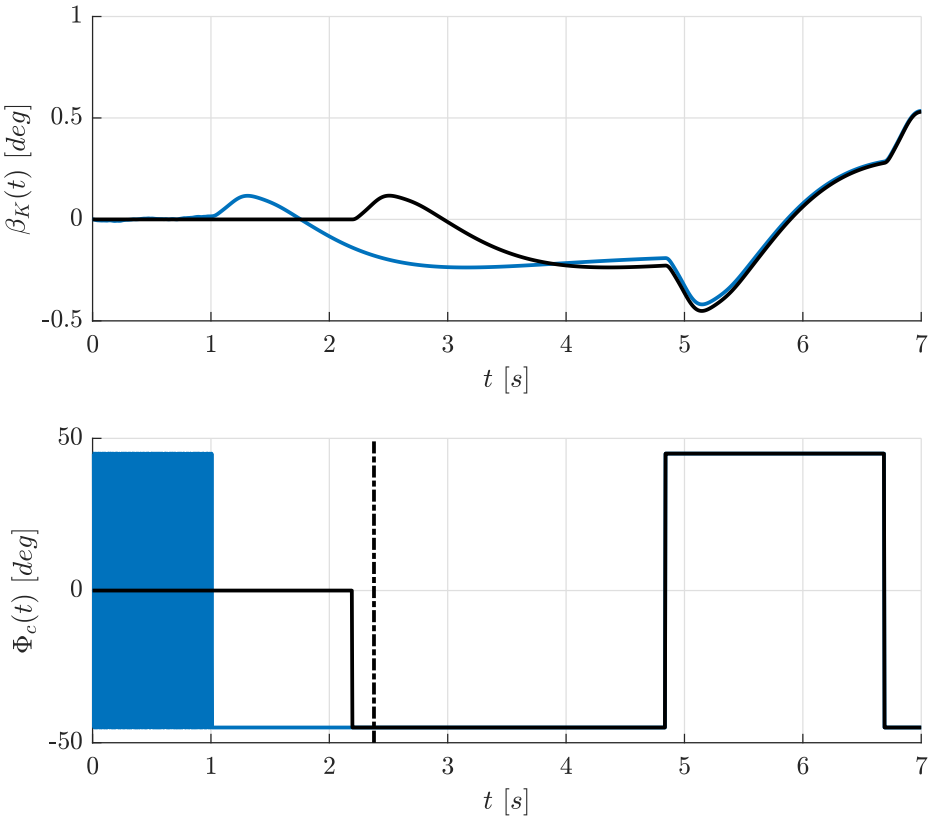


Figure 8.12: Optimal values for maximizing $\beta_K(t_f), t_f = 7\ s$, including the cost function $\beta_K(t)$ as well as the control $\Phi_c(t)$ (without penalty: blue; with penalty: black). The control penalty is imposed on the control variables for the first 2.38 s and has a contribution to the cost function of 0.0011735 deg.

To investigate the Cat II type problem using solely the roll angle command, the position and rate limits for the aileron and rudder with finite bounds are reintroduced in the problem formulation. The command is, as expected, less effective for both cases, i.e. when maximizing $\beta_K(t_f)$ (cf. Fig 8.14) or $n_y(t_f)$ (cf. Fig 8.13). In particular, the considerable control effort required to attain a negligible value for the AoS at the final time point, as illustrated in Fig 8.14, forces the servomechanism of the aileron both in the rate and position limits. Particularly noteworthy is the fact that solution exhibits a sequence of regular and singular parts where the singular parts are of first- and second-order. These singular parts correspond to constrained arcs for the aileron rate constraints (first-order) and the aileron position constraints (second-order). Note that as soon as the position limit of the aileron is reached the rate rapidly drops to zero and remains at zero while the position bound is active. This result confirms the condition (8.24), i.e. $x_{a,v}(t) = 0$ for the aileron rate state on an aileron position constrained arc.

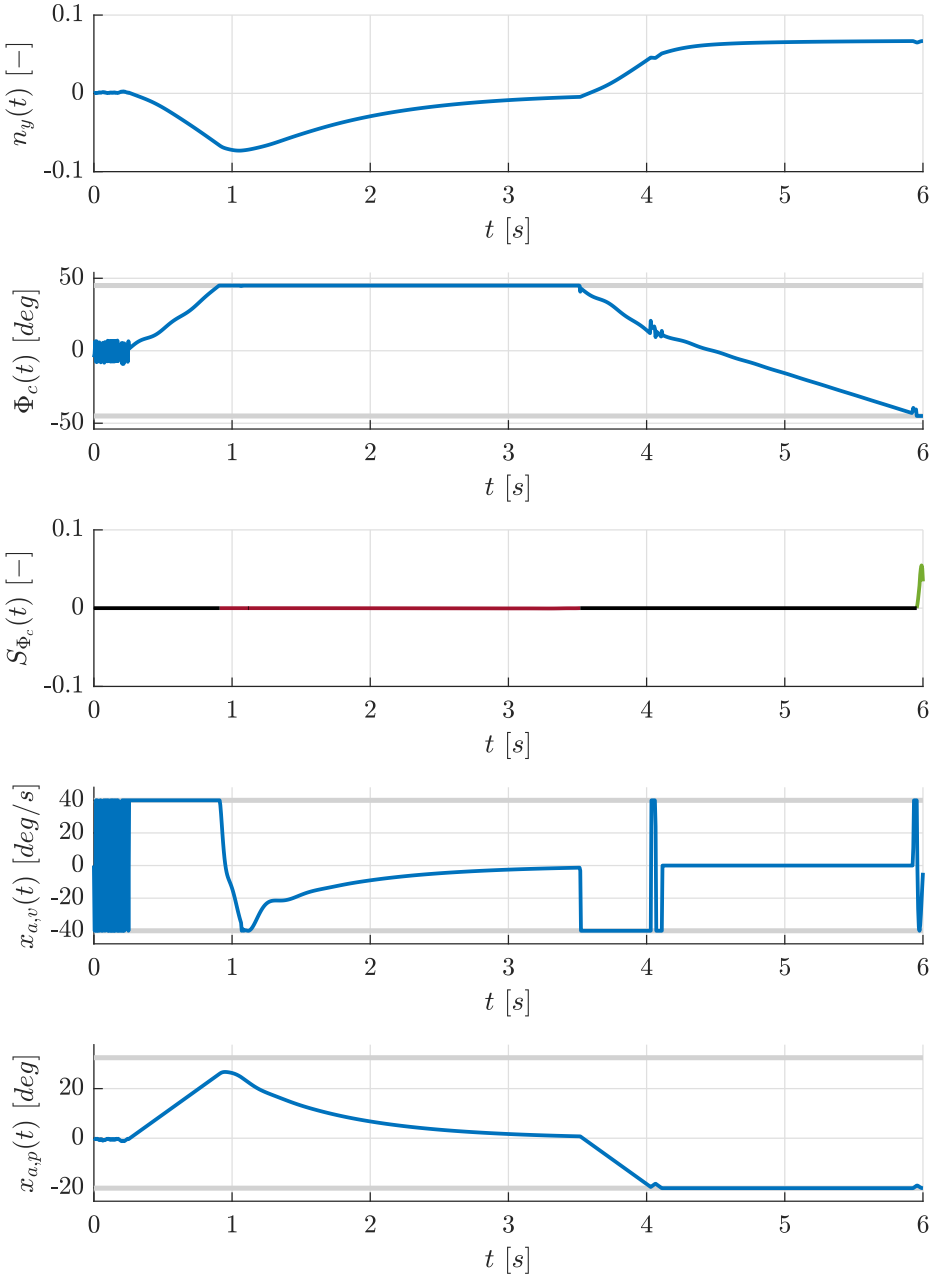


Figure 8.13: Optimal results for the lateral plane using a Cat II model for maximizing $n_y(t_f), t_f = 6\ s$ including the lateral load factor $n_y(t)$, the roll angle command $\Phi_c(t)$, the switching function corresponding to the roll angle command $\mathcal{S}_{\Phi_c}(t)$, the aileron rate $x_{a,v}(t)$, and the aileron position $x_{a,p}(t)$.

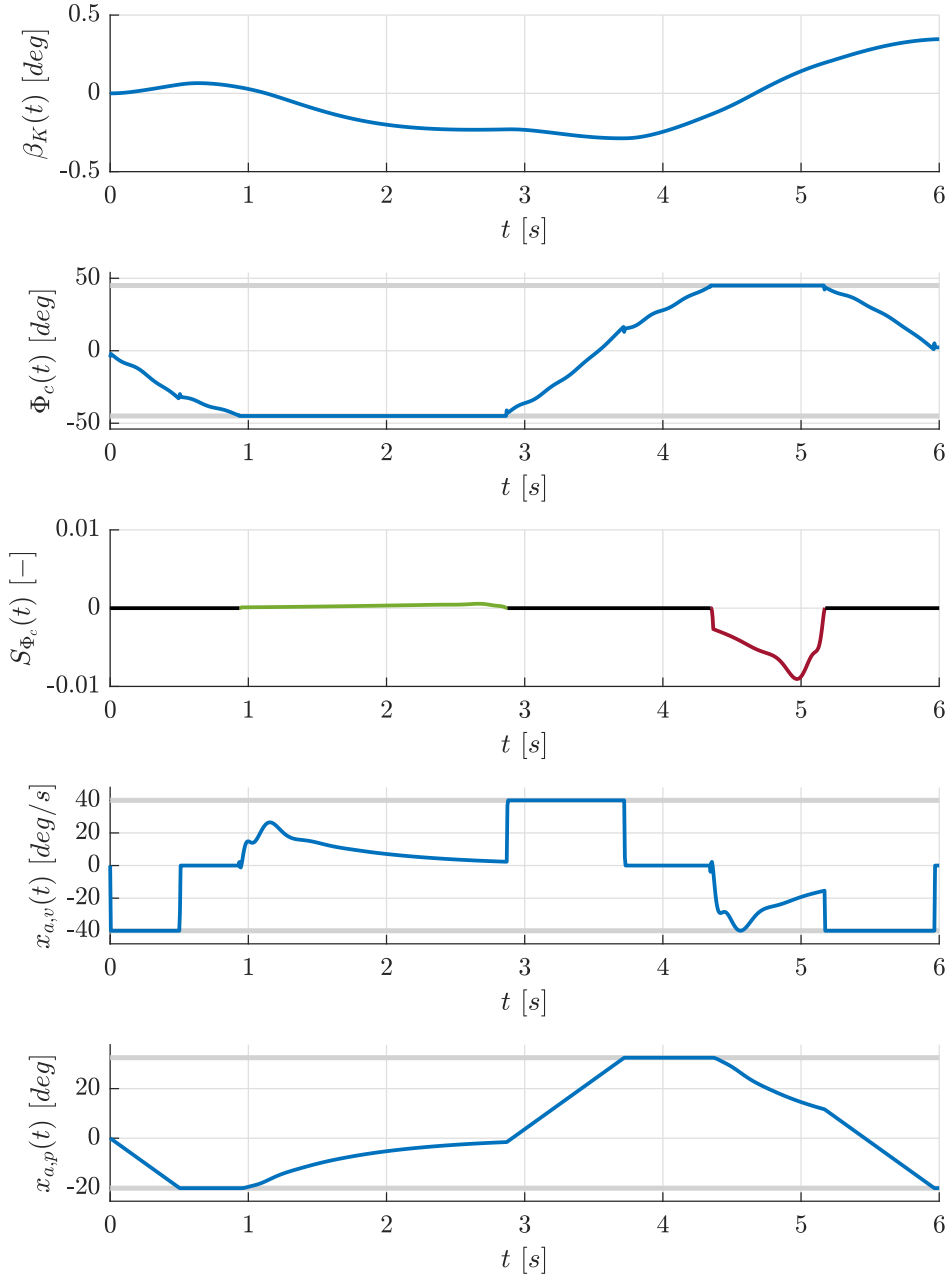


Figure 8.14: Optimal results for the lateral plane using a Cat II model for maximizing $\beta_K(t_f), t_f = 6\ s$ including the kinematic AoS $\beta_K(t)$, the roll angle command $\Phi_c(t)$, the switching function corresponding to the roll angle command $\mathcal{S}_{\Phi_c}(t)$, the aileron rate $x_{a,v}(t)$, and the aileron position $x_{a,p}(t)$.

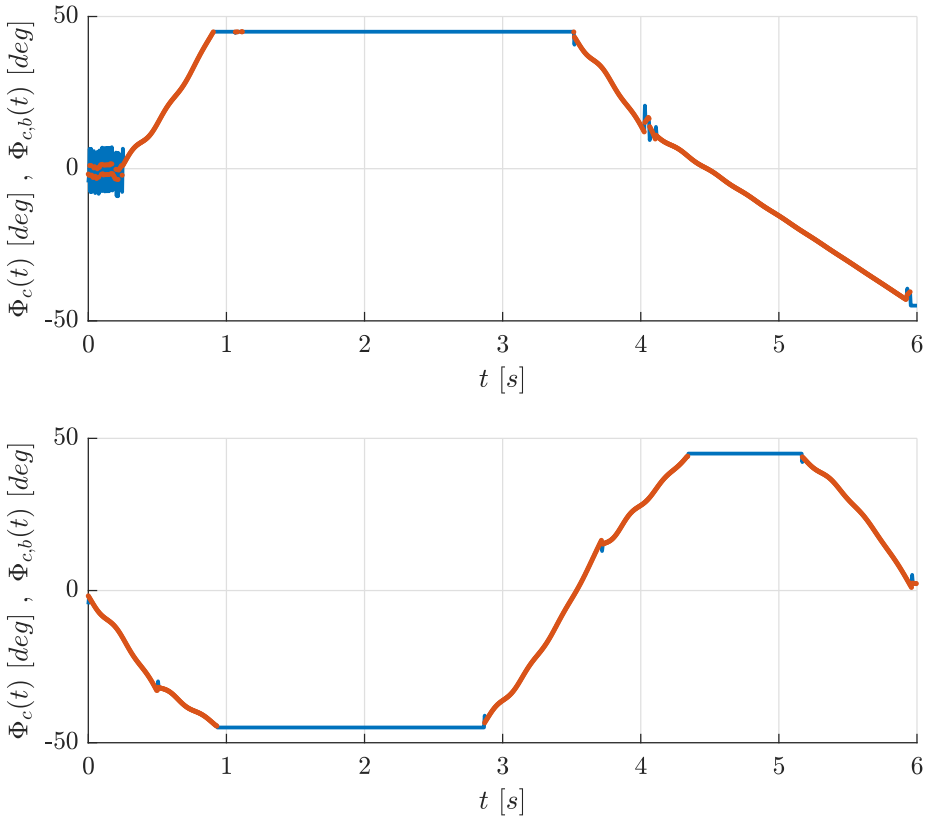


Figure 8.15: Boundary control on rate and position constrained arcs of the aileron for the maximization of the lateral load factor (top) and AoS (bottom). The control values obtained from the discretized problem ($\Phi_c(t)$, blue) are compared to the boundary control ($\Phi_{c,b}(t)$, red dots) according to the theoretical analysis.

Moreover, the boundary control can be estimated based on the condition for first- and second-order state constraints of the aileron actuator states:

$$\Phi_{c,b}(t) := -\frac{1}{b_{a,v}} \mathbf{a}_{a,v}^T \begin{bmatrix} x_{a,p}(t) \\ x_{r,p}(t) \\ x_{a,v}(t) \\ x_{r,v}(t) \\ \tilde{\mathbf{x}}_{lat}(t) \end{bmatrix} \tag{8.111}$$

The comparison of the boundary controls for maximizing the load factor and AoS between the solution of the discretized problem and the control predicted from the theoretical analysis is depicted in Fig. 8.15. Note that there is fast chattering visible on the state constrained arc of first-order in the beginning of the time interval when maximizing the lateral load factor. A possible explanation may be again the low sensitivity of control actions in the beginning of the time interval with respect to the cost function. On all other parts the results regarding the values of the boundary controls match very

well on both first- and second-order state constrained arcs.

Next, only the commanded lateral load factor $n_{y,c}(t)$ is used to maximize the lateral load factor $n_y(t_f)$ and the AoS $\beta_K(t_f)$. For the maximization of $n_y(t_f)$ (cf. Fig. 8.16) under the Cat I model the worst-case command represents a pulse in the second half of the time interval. In case of the maximization of $\beta_K(t_f)$ (cf. Fig. 8.17) the control switches only once. This switch takes place approximately in the middle of the time interval. Note that in both cases the controls are completely regular. In case of the Cat II models (cf. Fig. 8.18 and Fig. 8.19) the structure of the controls is similar. However, there are singular parts close to the time points where the control switches in case of the Cat I model. Here, the rudder rate state is on the limit and the control takes the values of the boundary control. Similar to the other cases (cf. Eq. 8.105 and Eq. 8.111) the boundary control corresponding to state constrained arcs of the rudder actuator states can be computed from the condition

$$n_{y,c,b}(t) := -\frac{1}{b_{r,v}} \mathbf{a}_{r,v}^T \begin{bmatrix} x_{a,p}(t) \\ x_{r,p}(t) \\ x_{a,v}(t) \\ x_{r,v}(t) \\ \tilde{\mathbf{x}}_{lat}(t) \end{bmatrix}, \tag{8.112}$$

The comparison between the numerical values and the boundary control obtained from Eq. (8.112) is depicted in Fig. 8.20 and appears to confirm the theoretical analysis.

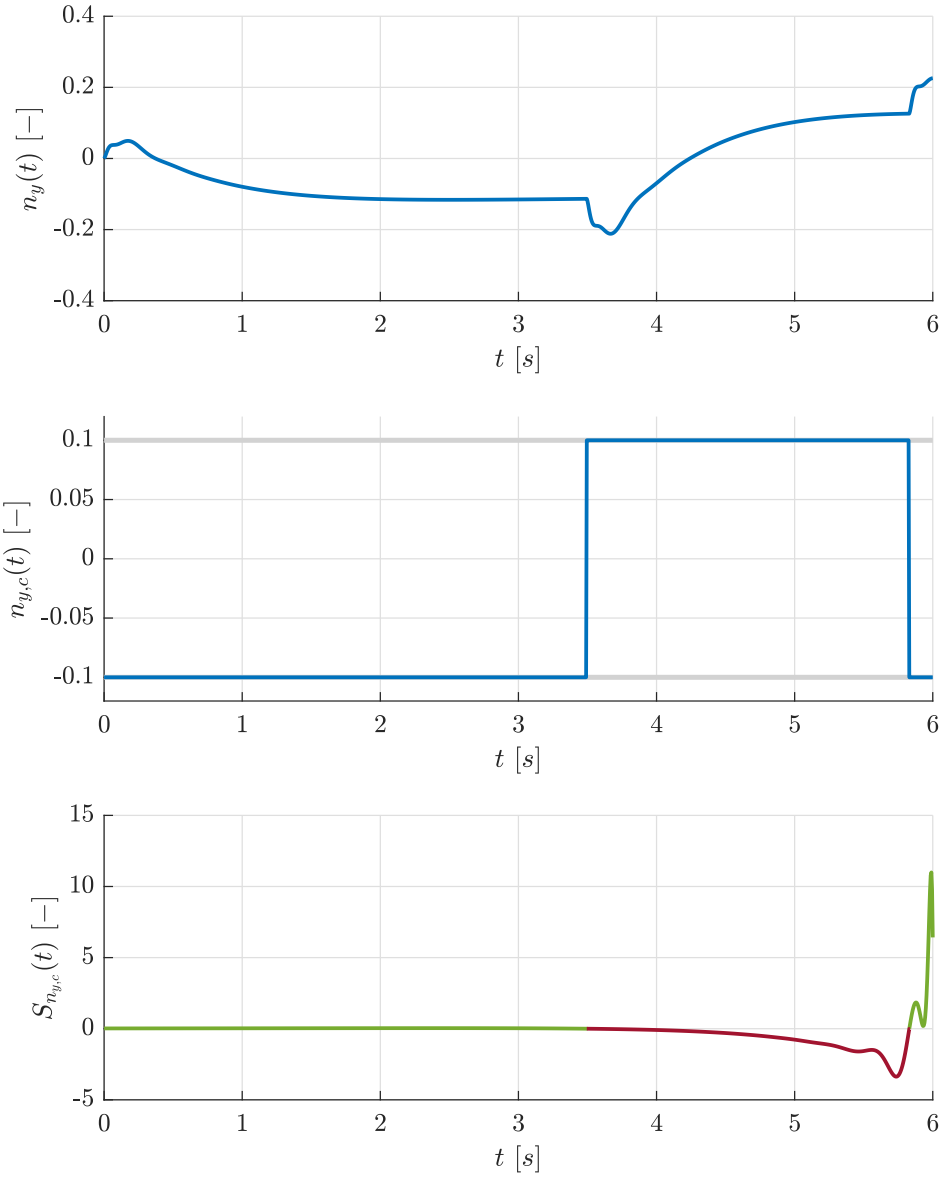


Figure 8.16: Optimal results for the lateral plane using a Cat I model for maximizing $n_y(t_f), t_f = 6\ s$ including the lateral load factor $n_y(t)$, the lateral load factor command $n_{y,c}(t)$, and the switching function corresponding to the lateral load factor command $\mathcal{S}_{n_{y,c}}(t)$.

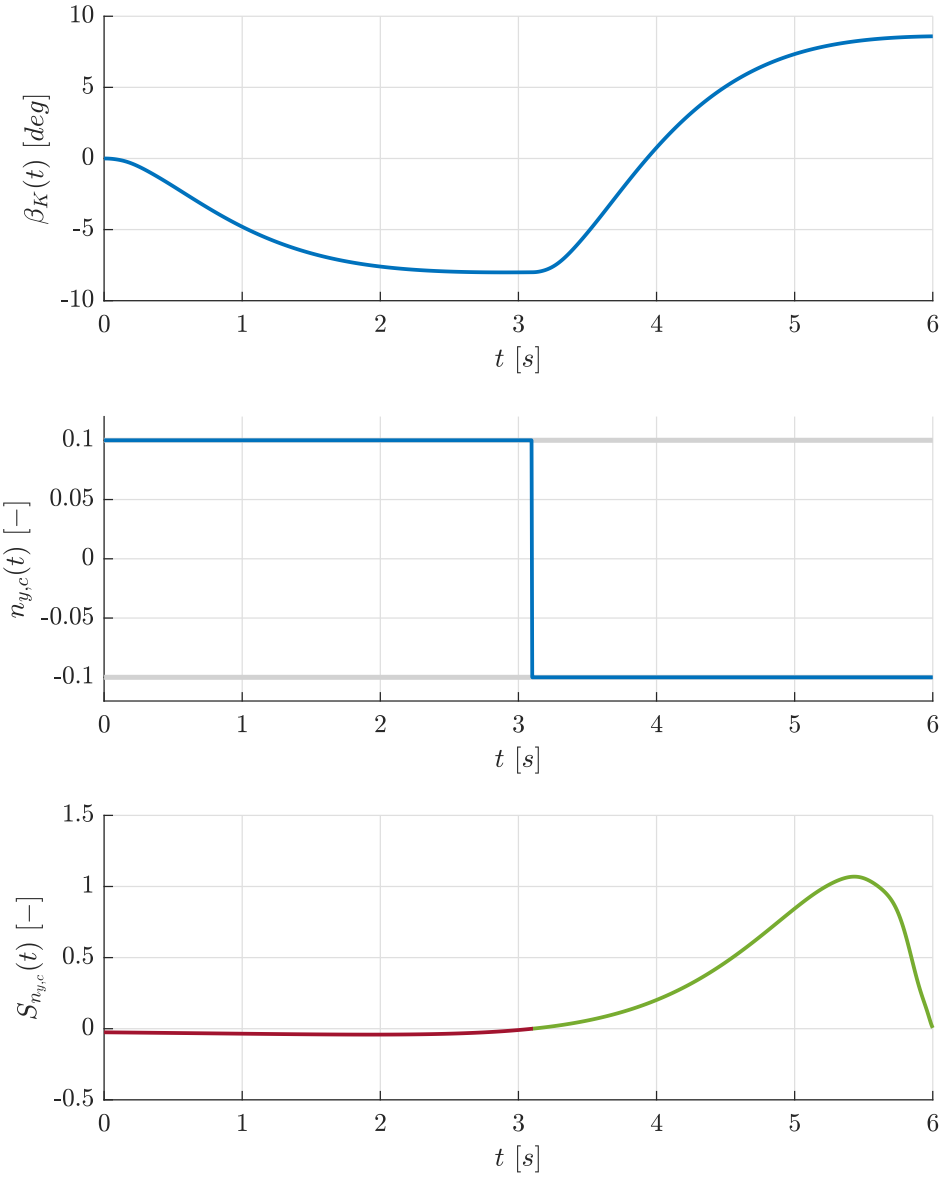


Figure 8.17: Optimal results for the lateral plane using a Cat I model for maximizing $\beta_K(t_f), t_f = 6\ s$ including the kinematic AoS $\beta_K(t)$, the lateral load factor command $n_{y,c}(t)$, and the switching function corresponding to the lateral load factor command $\mathcal{S}_{n_{y,c}}(t)$.

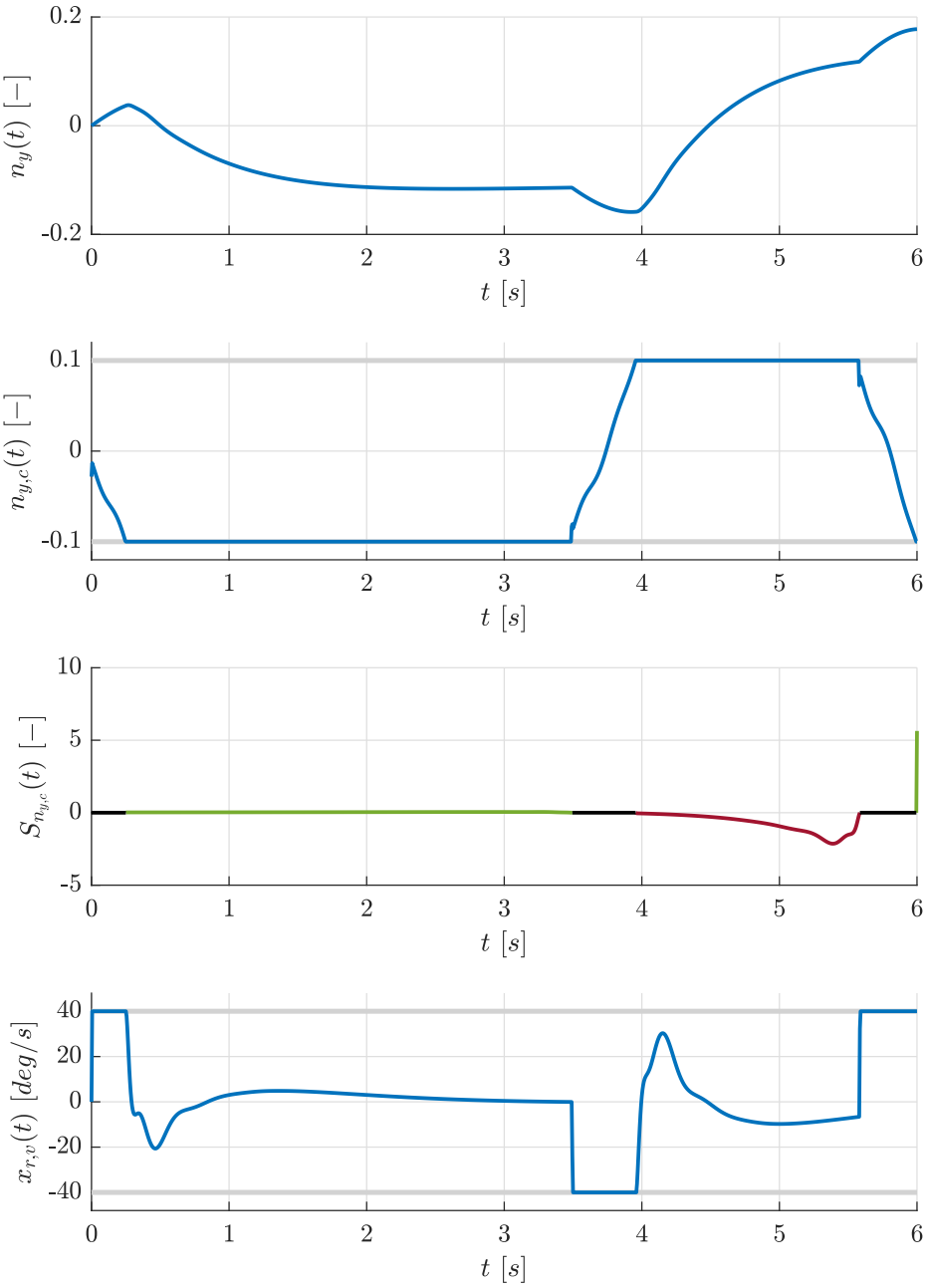


Figure 8.18: Optimal results for the lateral plane using a Cat II model for maximizing $n_y(t_f), t_f = 6\ s$ including the lateral load factor $n_y(t)$, the lateral load factor command $n_{y,c}(t)$, the switching function corresponding to the lateral load factor command $\mathcal{S}_{n_{y,c}}(t)$, and the rudder rate $x_{r,v}(t)$.

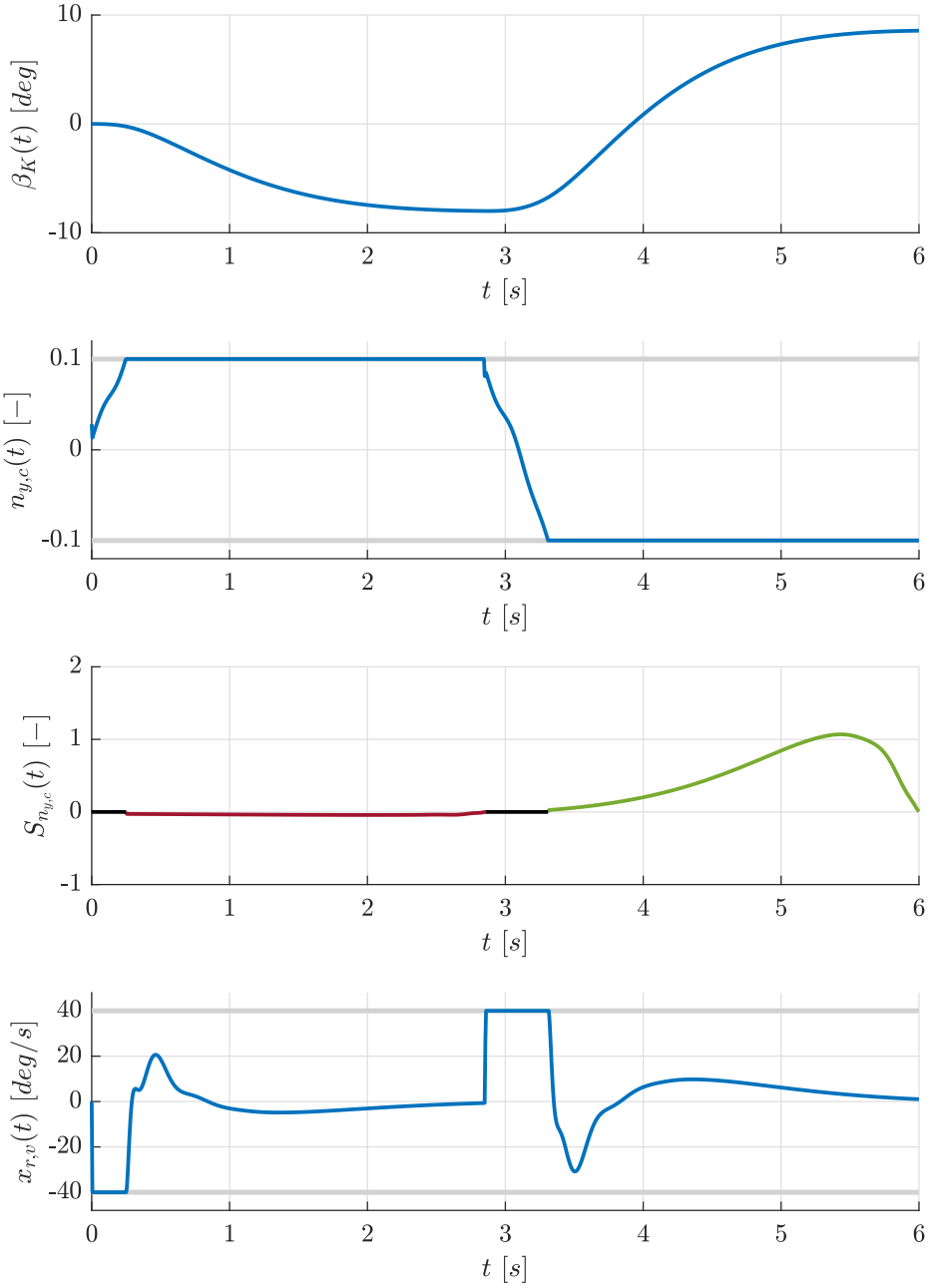


Figure 8.19: Optimal results for the lateral plane using a Cat II model for maximizing $\beta_K(t_f), t_f = 6\ s$ including the kinematic AoS $\beta_K(t)$, the lateral load factor command $n_{y,c}(t)$, the switching function corresponding to the lateral load factor command $\mathcal{S}_{n_{y,c}}(t)$, and the rudder rate $x_{r,v}(t)$.

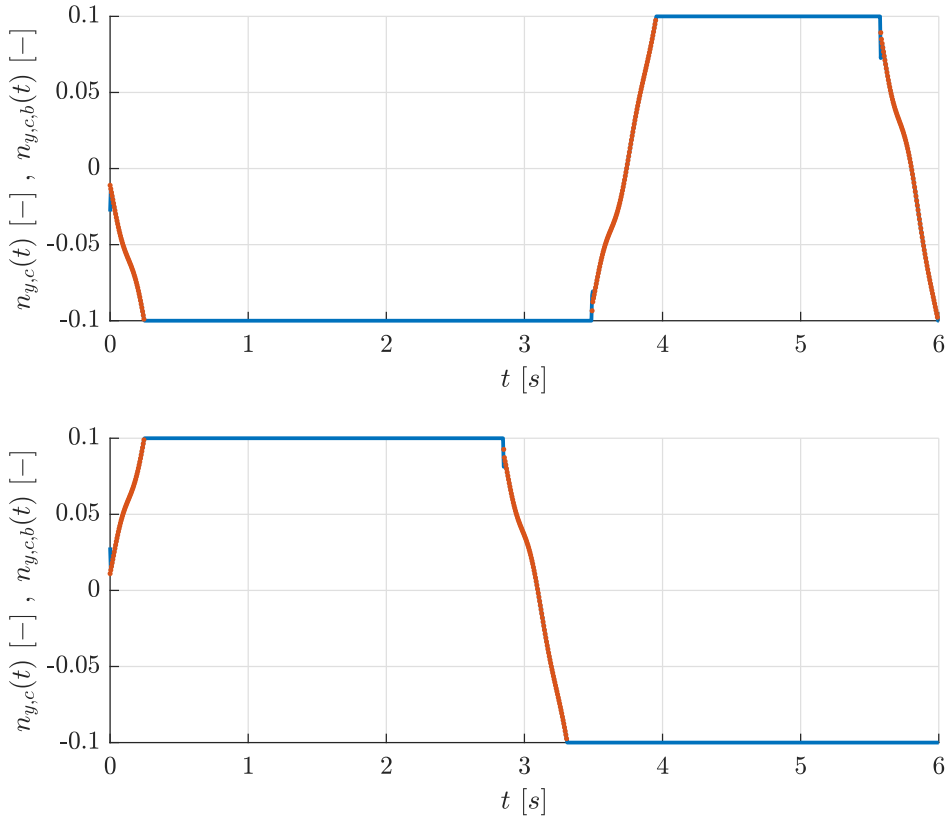


Figure 8.20: Boundary control on rate constrained arcs of the rudder for the maximization of the lateral load factor (top) and AoS (bottom). The control values obtained from the discretized problem ($n_{y,c}(t)$, blue) are compared to the boundary control ($n_{y,c,b}(t)$, red dots) according to the theoretical analysis.

Finally, the optimal solutions for combined control actions are presented and discussed. It is found that the (optimal) worst-case gust has a detrimental effect on all clearance criteria. This is illustrated for the Cat I and Cat II clearance problems in the longitudinal plane under the combined load factor command $n_{z,c}(t)$ and wind commands ($u_{W,c}(t)$ and $w_{W,c}(t)$) in Figs. 8.21-8.24. Observe that the optimal values of the criteria including wind are considerably worse compared to the case without wind (see Figs. 8.5-8.8). Note further that the worst-case wind commands are bang-bang for all examples. The load factor command is bang-bang for the Cat I models (see Fig. 8.21 and Fig. 8.22) and exhibits singular parts corresponding to state constrained arcs of the elevator rate state (see Fig. 8.22 and Fig. 8.23). It is interesting that the structure of the worst-case load factor command essentially remains unchanged compared to the results without wind and the wind commands primarily seem to support the control actions taken by the load factor command. In particular, some switches of the command inputs appear to be very close (see for example the last switches of the wind commands as well as the first switch of the normal wind command $w_{W,c}(t)$ and the load factor command $n_{z,c}(t)$ in Fig. 8.21). Observe that for the Cat II examples in Fig. 8.23 and Fig. 8.24 the wind command is completely regular, i.e. of bang-bang type, whereas the normal load factor command is singular when the elevator rate state is on one of the bounds.

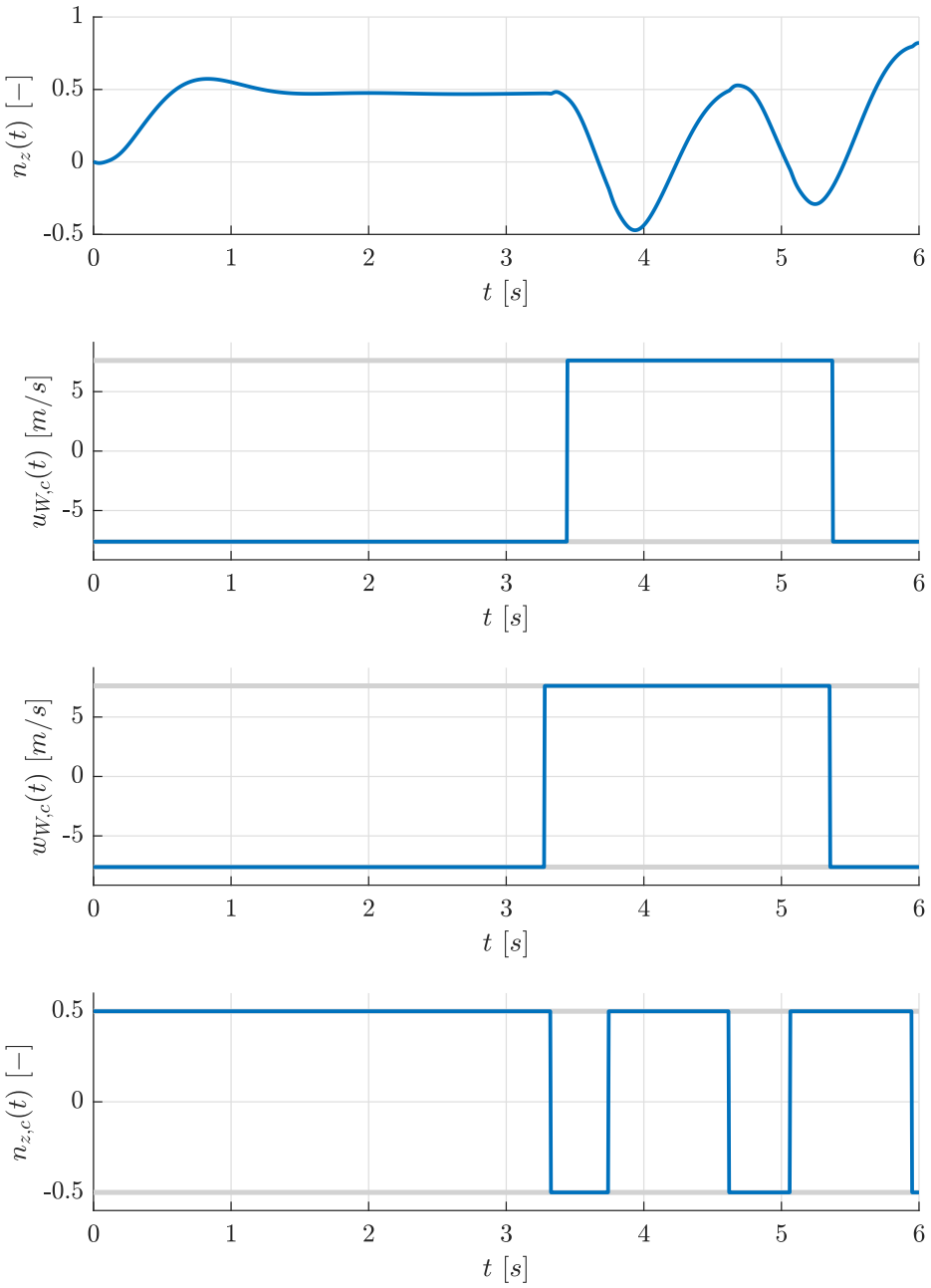


Figure 8.21: Optimal results for the longitudinal plane using a Cat I model for maximizing $n_z(t_f), t_f = 6\ s$ with wind including the normal load factor $n_z(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, and the normal load factor command $n_{z,c}(t)$.

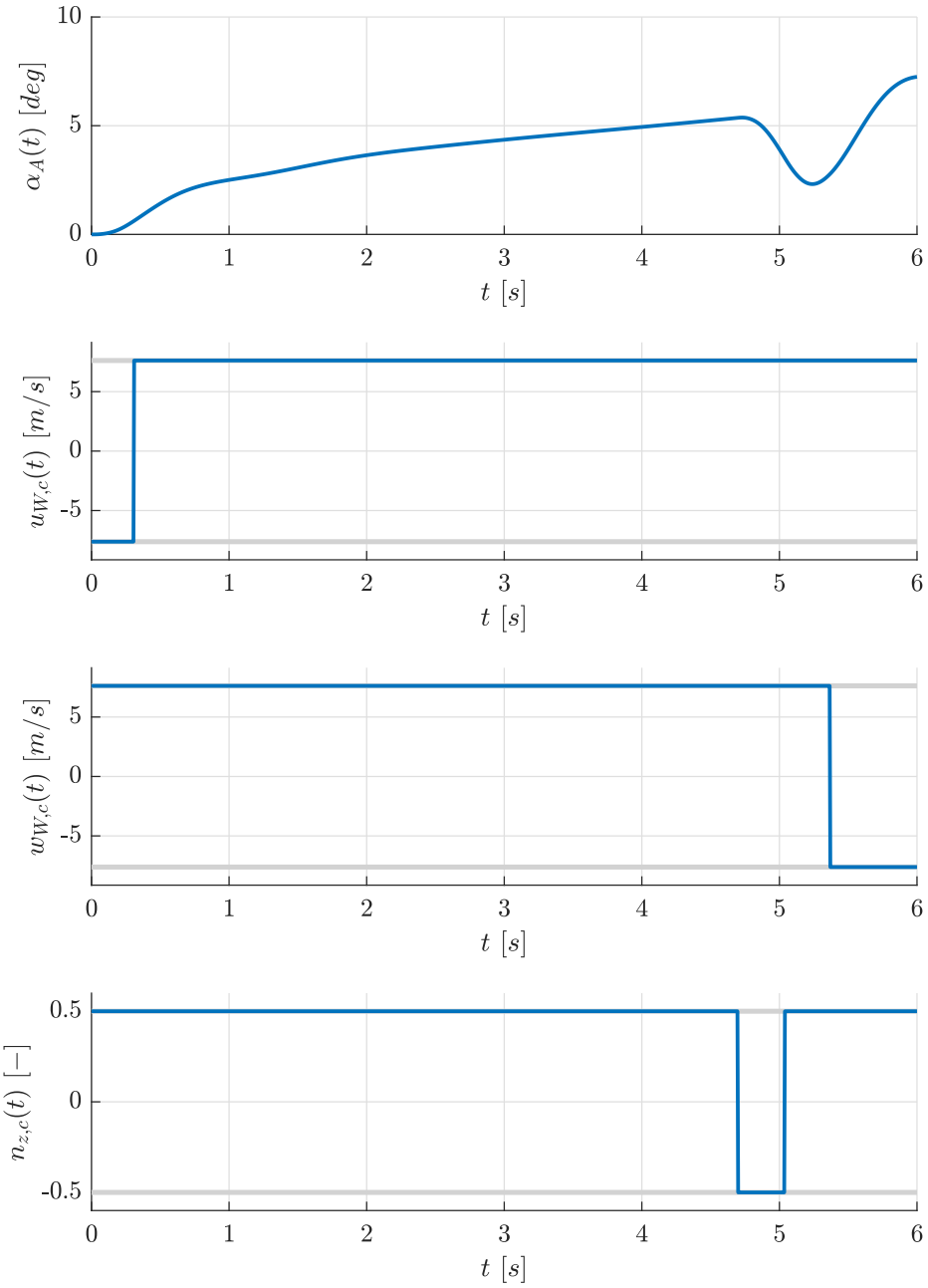


Figure 8.22: Optimal results for the longitudinal plane using a Cat I model for maximizing $\alpha_A(t_f), t_f = 6\ s$ with wind including the AoA $\alpha_A(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, and the normal load factor command $n_{z,c}(t)$.

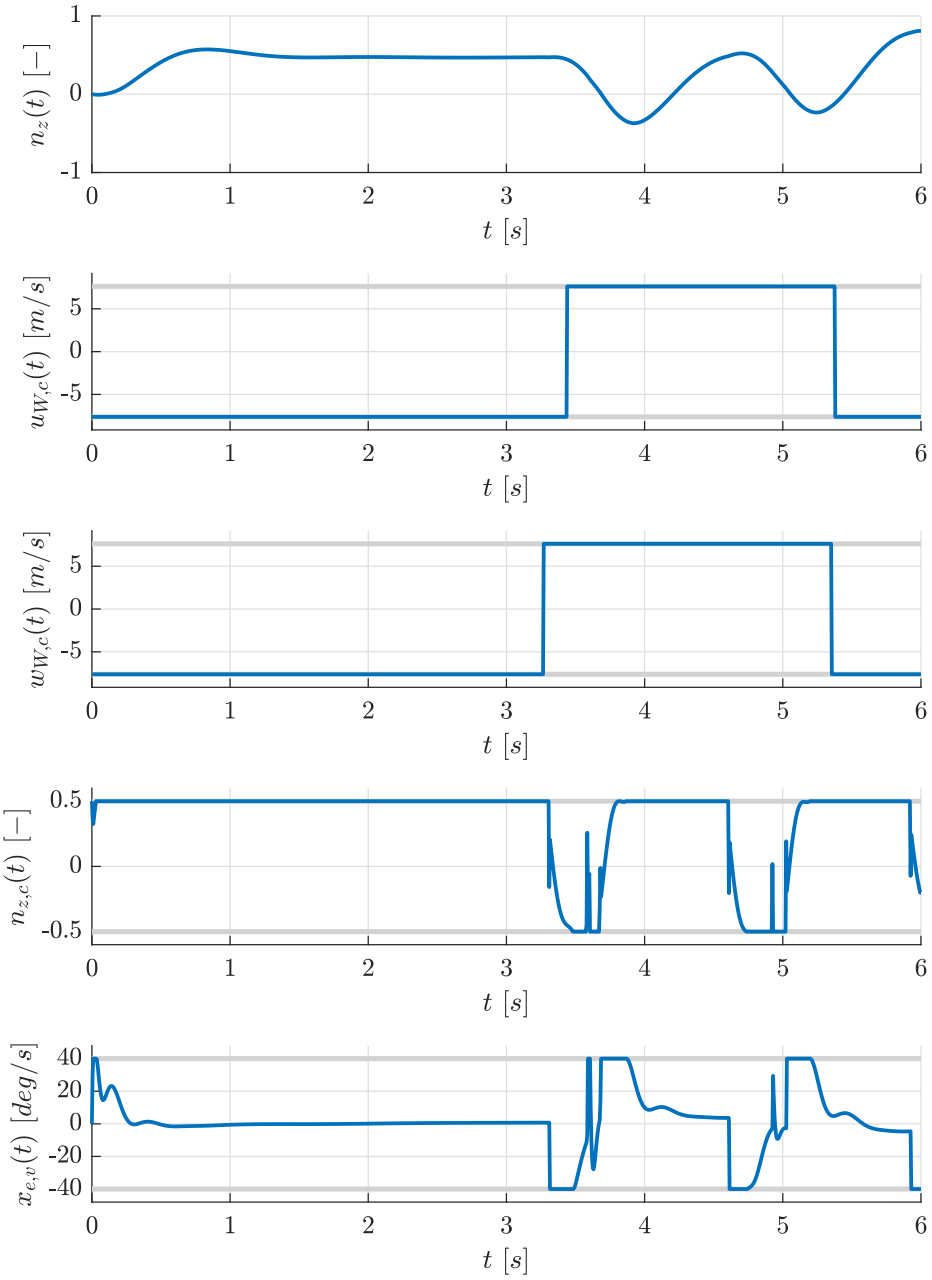


Figure 8.23: Optimal results for the longitudinal plane using a Cat II model for maximizing $n_z(t_f), t_f = 6\ s$ with wind including the normal load factor $n_z(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, the normal load factor command $n_{z,c}(t)$, and the elevator rate $x_{e,v}(t)$.

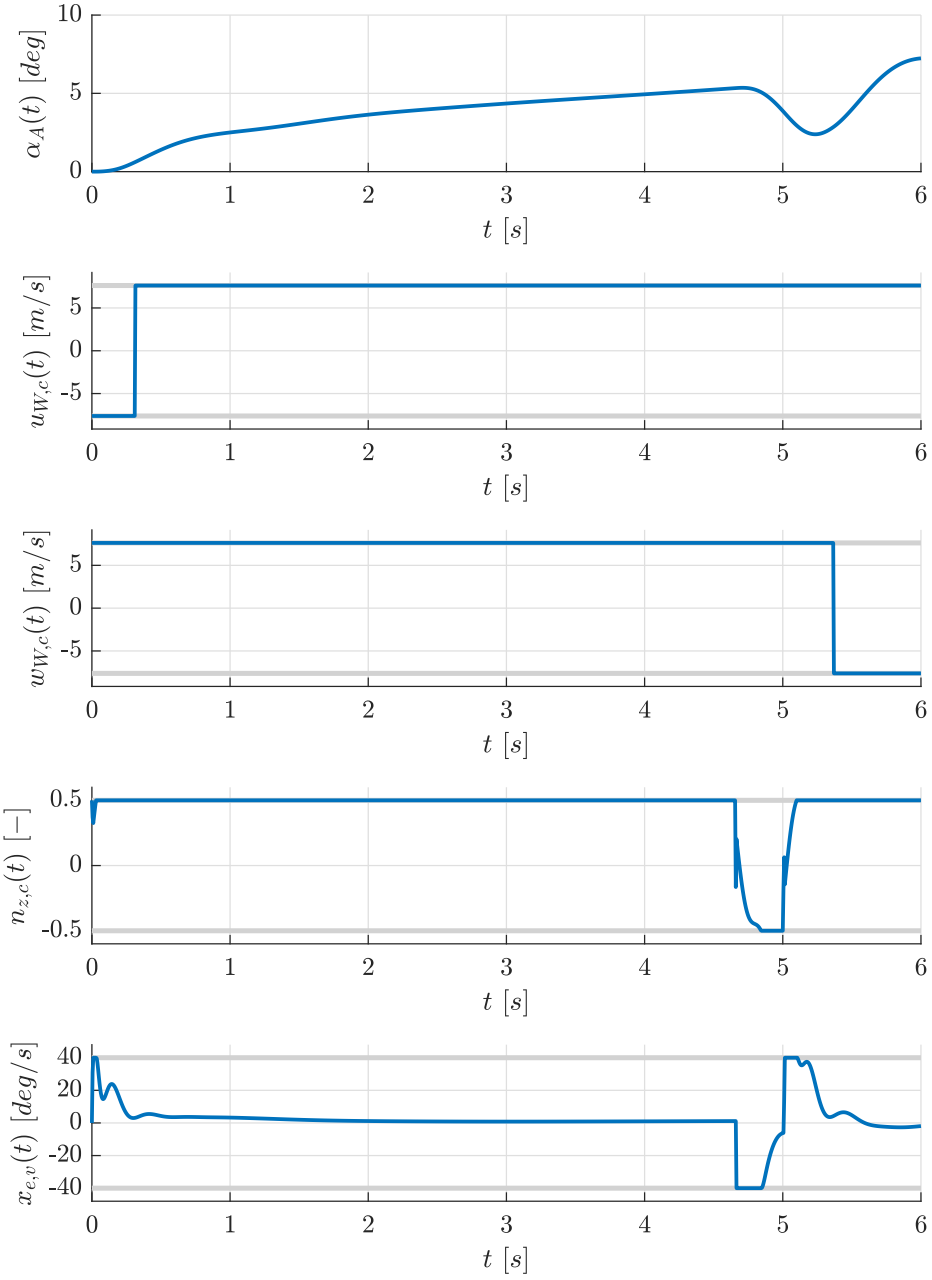


Figure 8.24: Optimal results for the longitudinal plane using a Cat II model for maximizing $\alpha_A(t_f), t_f = 6\ s$ with wind including the AoA $\alpha_A(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, the normal load factor command $n_{z,c}(t)$, and the elevator rate $x_{e,v}(t)$.

For the lateral plane the optimal solutions for a combination of the lateral load factor command $n_{y,c}(t)$, the roll angle command $\Phi_c(t)$, and the lateral gust $v_{W,c}(t)$ are investigated. As for the longitudinal plane the basic structure of the load factor command and the roll angle command is quite similar to the cases when using only a single command input. This can be seen as an indicator for a high roll/yaw decoupling in the closed-loop system. Note that for the state constrained cases (see Fig. 8.27 and Fig. 8.28) both controls exhibit a rather complicated sequence of regular and singular parts with state constraint arcs of first- and second-order.

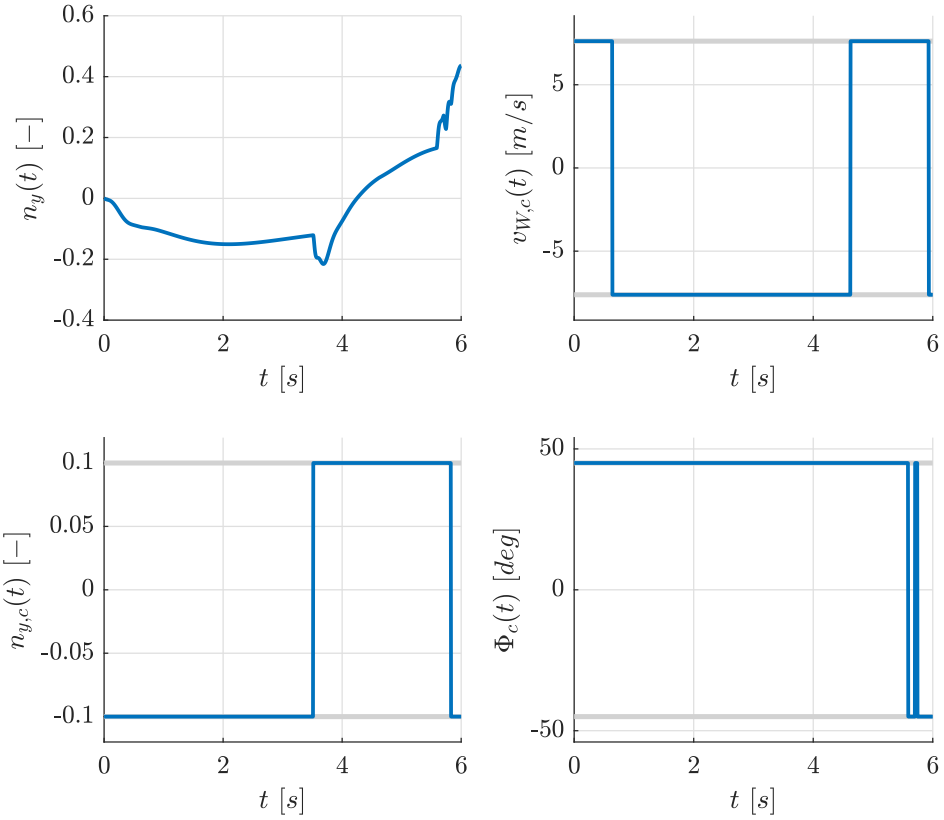


Figure 8.25: Optimal results for the lateral plane using a Cat I model for maximizing $n_y(t_f), t_f = 6\ s$ with wind including the lateral load factor $n_y(t)$, the lateral wind velocity command $v_{W,c}(t)$, the lateral load factor command $n_{y,c}(t)$, and the roll angle command $\Phi_c(t)$.

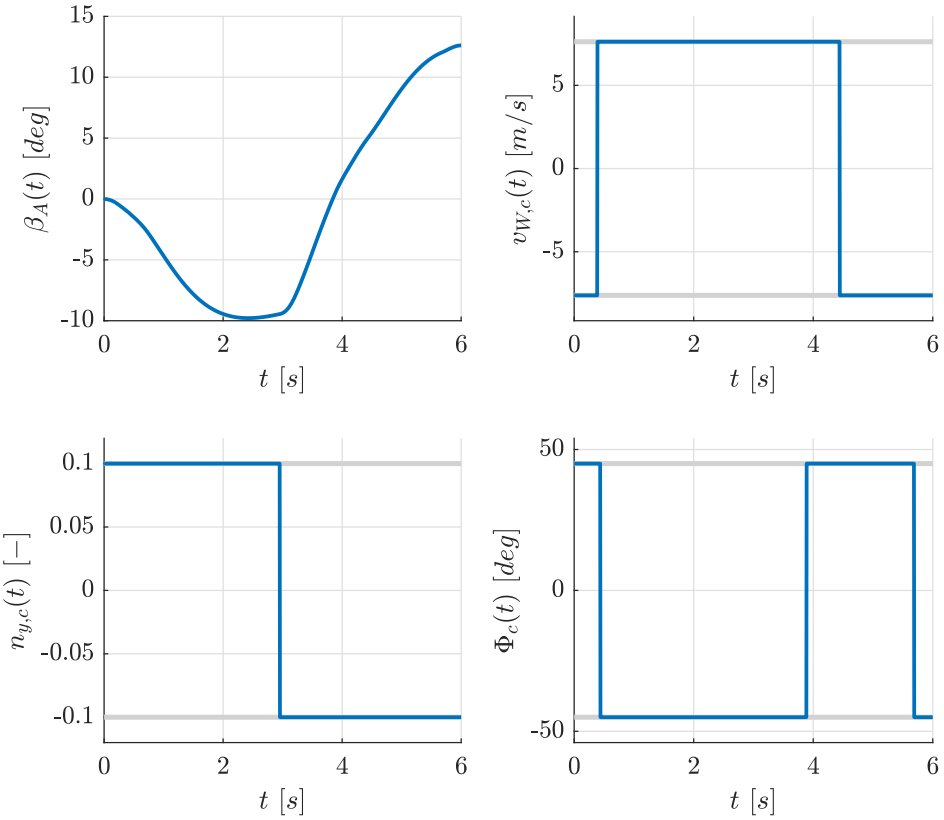


Figure 8.26: Optimal results for the lateral plane using a Cat I model for maximizing $\beta_A(t_f), t_f = 6\ s$ with wind including the AoS $\beta_A(t)$, the lateral wind velocity command $v_{W,c}(t)$, the lateral load factor command $n_{y,c}(t)$, and the roll angle command $\Phi_c(t)$.

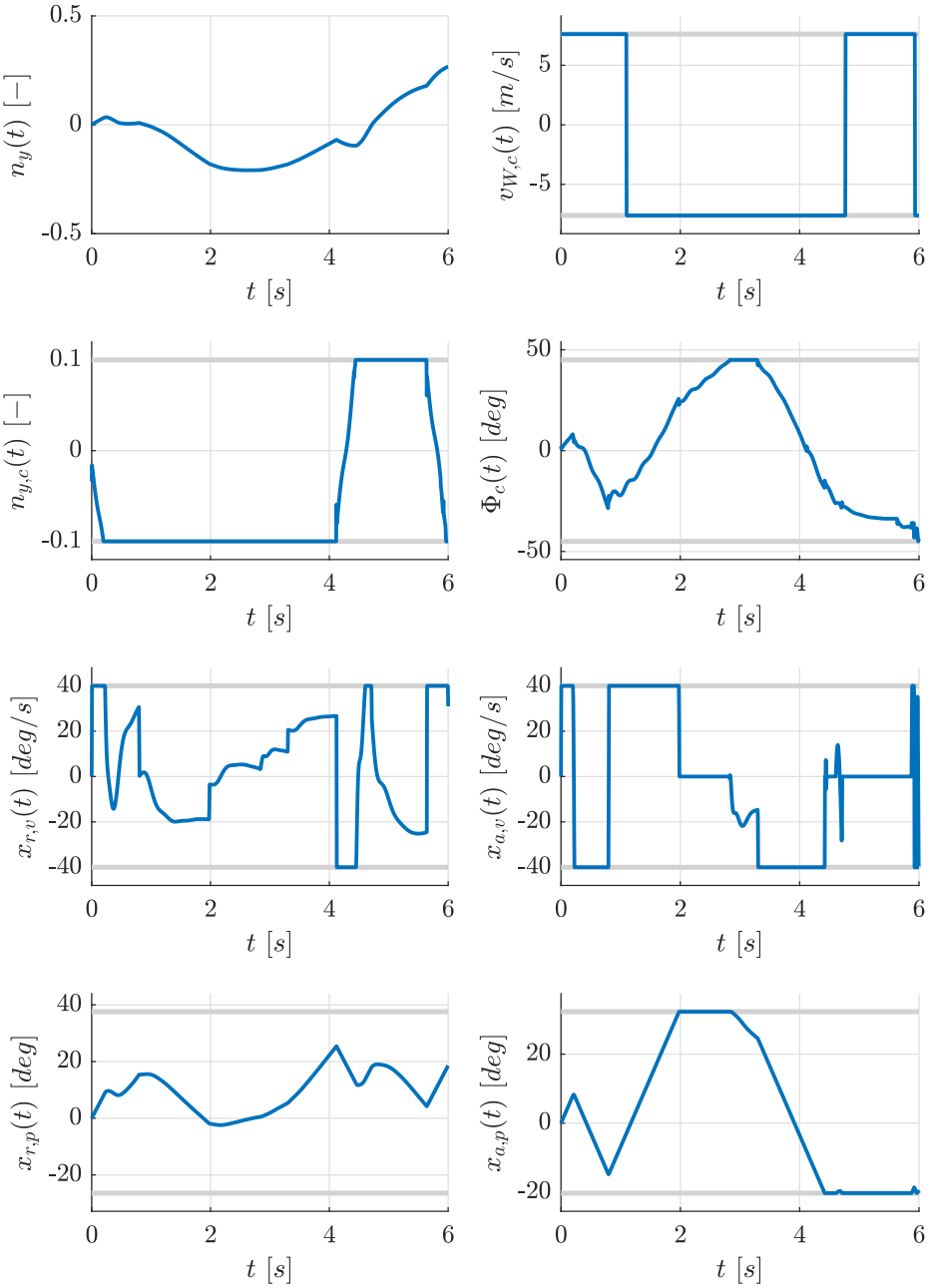


Figure 8.27: Optimal results for the lateral plane using a Cat II model for maximizing $n_y(t_f), t_f = 6\ s$ with wind including the lateral load factor $n_y(t)$, the lateral wind velocity command $v_{W,c}(t)$, the lateral load factor command $n_{y,c}(t)$, the roll angle command $\Phi_c(t)$, the rudder rate $x_{r,v}(t)$, the aileron rate $x_{a,v}(t)$, the rudder position $x_{r,p}(t)$, and the aileron position $x_{a,p}(t)$.

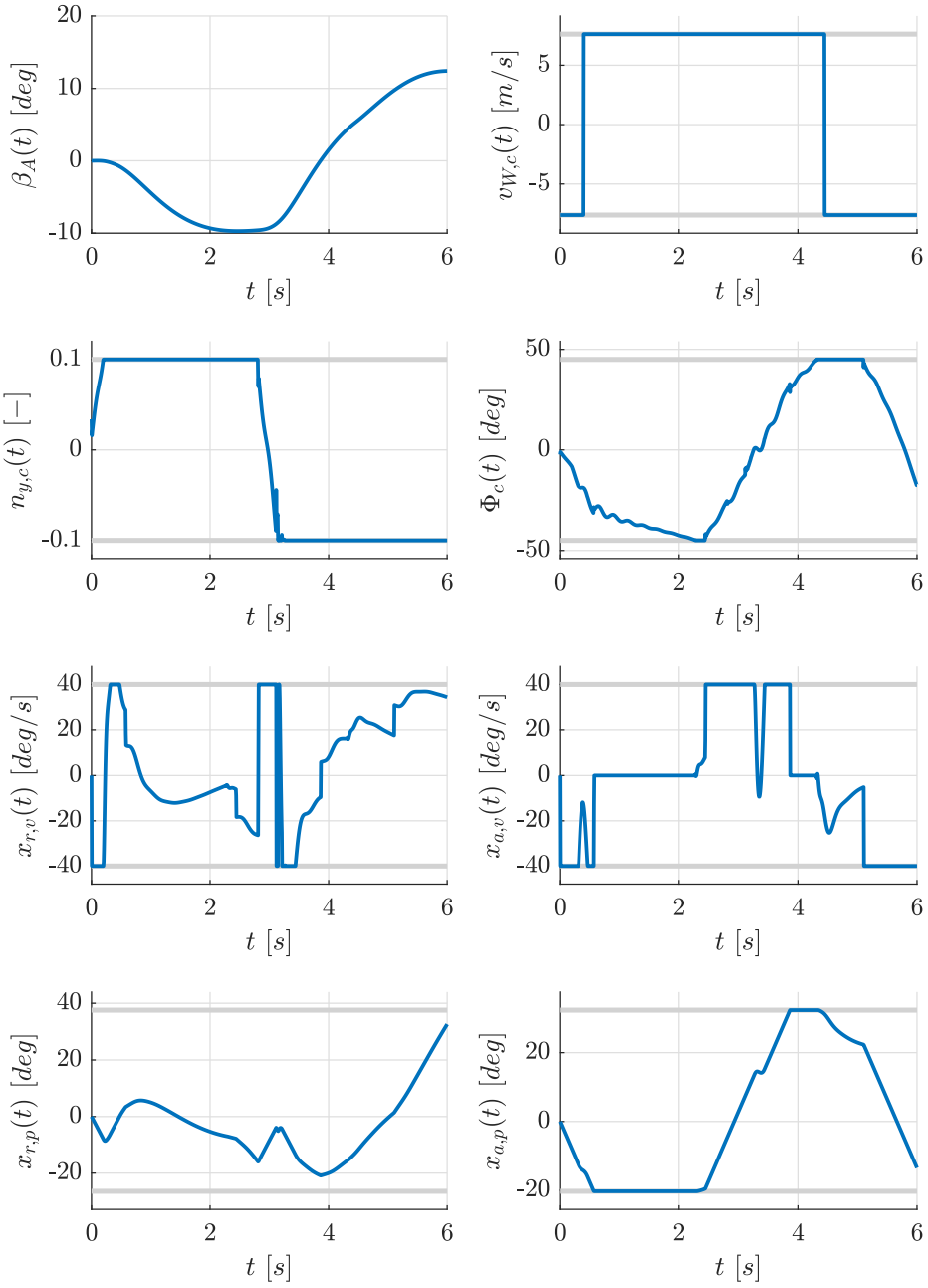


Figure 8.28: Optimal results for the lateral plane using a Cat II model for maximizing $\beta_A(t_f), t_f = 6\ s$ with wind including the AoS $\beta_A(t)$, the lateral wind velocity command $v_{W,c}(t)$, the lateral load factor command $n_{y,c}(t)$, the roll angle command $\Phi_c(t)$, the rudder rate $x_{r,v}(t)$, the aileron rate $x_{a,v}(t)$, the rudder position $x_{r,p}(t)$, and the aileron position $x_{a,p}(t)$.

Recall that throughout this section the Backward Euler method is employed for all examples and the discrete Minimum Principle (cf. Ref. [52]) is used to verify the theoretical results. As mentioned before, this transcription method is found to produce very clean results regarding the control functions. This is mainly observed to be the case under Cat II type models (i.e. with state constraints) and clearly facilitates the interpretation of the numerical results in the light of the theoretical developments from Sec. 8.1.1 and Sec. 8.1.2. However, for example the Trapezoidal method shows to be rather prone to the effect of control chattering (high frequency switches of the optimal control) on state constrained arcs. This effect is illustrated for the longitudinal plane in Fig. 8.29. In this figure, a detailed view of the optimal solution to the Cat II clearance

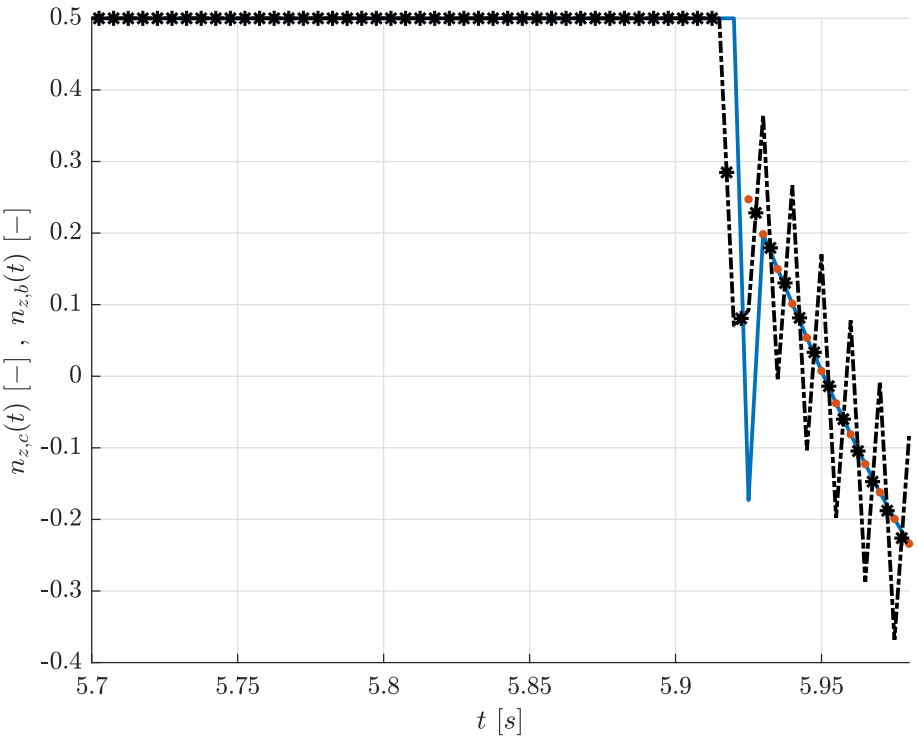


Figure 8.29: Optimal solutions of the Cat II problem for maximizing $n_z(t_f), t_f = 6\ s$ using the Backward Euler method (blue line) and the Trapezoidal method (black dashed line). Black star markers represent the mean control values between two control discretization points of the Trapezoidal method and the red dots are the values of the boundary control based on the theoretical analysis.

problem for maximizing $n_z(t_f), t_f = 6\ s$ using the normal load factor command $n_{z,c}(t)$ is presented. The depicted detail shows the entry to a first-order state constrained arc (elevator rate constraint) coming from a regular part of the control. Due to the finite discretization the entry point is not exactly resolved and for both the Backward Euler method (blue line) and the Trapezoidal method (black dashed line) the controls appear to overshoot the entry of the arc. This may be interpreted as a compensatory action in order to enter the arc caused by the restricted input space of the discretized control.

Note that the control from the Backward Euler method (blue line) exhibits only a single overshooting control variable and shows a clean discretized version of the boundary control on the state constrained arc (red dots, estimated based on the theoretical results using values from the Backward Euler solution). The control history corresponding to the Trapezoidal method however heavily oscillates on the state constrained interval. A possible explanation for this behavior is that for the Trapezoidal method the mean state vector $\mathbf{x}_m^{[i]} := \frac{\mathbf{x}^{[i]}+\mathbf{x}^{[i-1]}}{2}$ and mean control vector $\mathbf{u}_m^{[i]} := \frac{\mathbf{u}^{[i]}+\mathbf{u}^{[i-1]}}{2}$ are in fact the quantities entering the dynamic constraint (cf. Eq. (8.53)):

$$\mathbf{x}^{[i]} - h\mathbf{A}\mathbf{x}_m^{[i]} - h\mathbf{B}\mathbf{u}_m^{[i]} - \mathbf{x}^{[i-1]} = \mathbf{0},\ i = 1, \ldots, N \tag{8.113}$$

These mean controls for the Trapezoidal method are depicted in Fig. 8.29 (black star markers). Note that these control values appear to be very close to the boundary control found from the Backward Euler method (blue line) and the boundary control estimated by the theoretical results using the values of the Backward Euler solution (red dots). Despite the observation that the lower order Backward Euler method produces cleaner results for the control compared to the higher order Trapezoidal method does not mean that higher order methods should not be used for the clearance task. At first, it is important to mention that this chattering effect is not as harmful as it may seem at first glance. The solution may not be as pleasing to the eye as a clean control history, however the main figure obtained from the solution of the flight control law clearance problem is the value of the criterion itself (which is essentially unaffected by this phenomenon). Moreover, higher order schemes are expected to be considerably more efficient than lower order schemes as the time required to solve the clearance problem depends on the number of variables and constraints in the discretized problem. Computational experience indicates that the solution time for a relatively coarse discretization may be within the range of a few milliseconds whereas the solution times for a very fine discretization spreads from several seconds to over minutes. The number of variables and constraints is obviously lower if a higher order discretization scheme is used as the time step length may be chosen larger which yields a fewer number of variables and constraints. As such, it is found that the solvers find worst-case values considerably faster if a higher order scheme is used which requires less discretization points. As the approaches presented in the following sections are solely focused on the value of the criterion a higher order scheme is applied instead of the Backward Euler method. Here, the Trapezoidal scheme is used for simplicity as the focus in the following is not on the choice of a particular discretization method but on the discussion of the clearance approaches. However, it is important to mention that all schemes presented in Sec. 6 yield LP problems under the Cat I and Cat II modeling approach and may be employed to transcribe the optimal control problem.

8.3 Parameter Dependent Clearance Problems

In the following section the Cat I and Cat II clearance problems including parameter variations are investigated (cf. Sec. 8.3.1). On the one side, a bi-level algorithm (cf. Sec. 8.3.2) is presented which can be employed to identify worst-case solutions (parameters in combination with optimal control actions) using global and local optimization methods. In this context, results from post-optimal sensitivity analysis are used to determine influential parameters regarding the worst-case solution and to obtain derivative information for gradient-based methods. On the other side, the application of generalized Polynomial Chaos (gPC) for the clearance task is proposed (cf. Sec. 8.3.3) in order to efficiently determine statistical information concerning the clearance criterion. (Contribution II, [84, 85])

8.3.1 Problem Statement

Consider the case where the dynamic system under investigation depends on parameters $\mathbf{p} \in \mathbb{R}^{n_p}$. For the parameter dependent case it is useful to introduce the nonlinear dynamic system $\mathbf{f} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \to \mathbb{R}^{n_x}$ and the purely state dependent nonlinear output equation modeling the criterion $j : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \to \mathbb{R}$

$$\dot{\mathbf{x}}(t) := \mathbf{f}\left(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}\right), \tag{8.114}$$

$$y(t) := j\left(\mathbf{x}(t), \mathbf{p}\right) \tag{8.115}$$

This system is assumed to be in a stationary condition, that is $\dot{\mathbf{x}}(t) = \mathbf{0}$. The state and control vectors at the reference flight condition $\mathbf{x}_0 \in \mathbb{R}^{n_x}$ and $\mathbf{u}_0 \in \mathbb{R}^{n_u}$ are thus defined by the n_t parameter dependent trim conditions $\mathbf{r} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \to \mathbb{R}^{n_t}$ of the general form

$$\mathbf{r}\left(\mathbf{x}_0, \mathbf{u}_0, \mathbf{p}\right) = \mathbf{0}, \tag{8.116}$$

and the trim state vector $\mathbf{x}_0$ and control vector $\mathbf{u}_0$ implicitly depend on the parameters. The behavior of the system around such trim points may be approximated as:

$$\dot{\mathbf{x}}(t) \approx \underbrace{\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\Big|_{\mathbf{x}_0,\mathbf{u}_0,\mathbf{p}}}_{:=\mathbf{A}(\mathbf{x}_0,\mathbf{u}_0,\mathbf{p})} \left(\mathbf{x}(t) - \mathbf{x}_0\right) + \underbrace{\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\Big|_{\mathbf{x}_0,\mathbf{u}_0,\mathbf{p}}}_{:=\mathbf{B}(\mathbf{x}_0,\mathbf{u}_0,\mathbf{p})} \left(\mathbf{u}(t) - \mathbf{u}_0\right), \tag{8.117}$$

$$y(t) \approx j\left(\mathbf{x}_0, \mathbf{p}\right) + \underbrace{\frac{\partial j}{\partial \mathbf{x}}\Big|_{\mathbf{x}_0,\mathbf{p}}}_{:=\mathbf{c}^T(\mathbf{x}_0,\mathbf{p})} \left(\mathbf{x}(t) - \mathbf{x}_0\right) \tag{8.118}$$

Defining the admissible region for the parameters by a general inequality constraint function $\mathbf{g}_p : \mathbb{R}^{n_p} \to \mathbb{R}^{n_g}$, i.e.

$$\mathbf{g}_p(\mathbf{p}) \leq \mathbf{0}, \tag{8.119}$$

and under the linearized dynamic equation (8.117) and output equation (8.118) the parameter dependent Cat II problem is stated as follows

$$\begin{aligned}
\underset{\mathbf{u}(t),\,\mathbf{p},\,\mathbf{x}_0,\,\mathbf{u}_0}{\text{minimize}} \quad & j\left(\mathbf{x}_0, \mathbf{p}\right) + \mathbf{c}^T(\mathbf{x}_0, \mathbf{p})\left(\mathbf{x}(t_f) - \mathbf{x}_0\right) \\
\text{subject to} \quad & \dot{\mathbf{x}}(t) - \mathbf{A}\left(\mathbf{x}_0, \mathbf{u}_0, \mathbf{p}\right)\left(\mathbf{x}(t) - \mathbf{x}_0\right) - \mathbf{B}\left(\mathbf{x}_0, \mathbf{u}_0, \mathbf{p}\right)\left(\mathbf{u}(t) - \mathbf{u}_0\right) = \mathbf{0}, \\
& x_{j,lb} \leq x_j(t) \leq x_{j,lb},\ \forall j \in \mathcal{I}_b, \\
& \mathbf{x}(t_0) = \mathbf{x}_0, \\
& \mathbf{g}_p(\mathbf{p}) \leq \mathbf{0}, \\
& \mathbf{r}\left(\mathbf{x}_0, \mathbf{u}_0, \mathbf{p}\right) = \mathbf{0}, \\
& \mathbf{u}(t) \in \mathcal{U}, \\
& t \in I^t.
\end{aligned} \tag{8.120}$$

Note that this problem cannot be treated anymore as linear due to the nonlinearity introduced through the parameter dependency. In particular, a discretization of this problem using direct transcription methods yields a Nonlinear Programming (NLP) problem of high dimension (optimization variables and constraints). The nonlinearity on the one side and the dimension of the problem (typically several thousands) on the other represents the two main challenges for solving this parameter dependent version of the Cat II type clearance problem. Note that as in the nominal case, the Cat II type problem reduces to a Cat I type problem for $\mathcal{I}_b = \emptyset$. On the one hand, gradient-based NLP methods may be employed for the numerical solution in order to efficiently solve the high dimensional problem. However, the optimization problem (8.120) is by no means trivial due to the simultaneous determination of the worst-case parameter combination with the worst-case controls under the trim constraint. As such, these methods can easily get trapped in local solutions and it is not possible to guarantee that the global worst-case as for the LP approach in the nominal case is found. On the other hand, it may be tempting to employ global optimization algorithms to increase the chance of finding a global minimizer. Unfortunately, these methods only operate efficiently for low dimensional parameter spaces (typically several tens). In order to resolve this conflict, a decomposition approach is proposed for the parameter dependent clearance task which separates the search in the parameter space from the optimal control problem. For this decomposition approach the problem is split into two levels. The higher level treats the parameter dependency and the lower level solves LP problems for fixed parameter values. This lower level is of a similar form as the one used

for the nominal clearance problem. In particular, the trim condition is evaluated internally in the lower level and does not need to be exposed to the optimization algorithm. On the one side, this approach directly leverages the efficiency of LP algorithms for finding global solutions to the nominal, high dimensional lower level problem. On the other side, the upper level problem is of low dimension for most clearance problems which enables the use of global parameter optimization methods. Besides the use of global algorithms it is observed that gradient-based schemes are highly effective for the upper level but merely require a fraction of the computational cost associated with global methods.

In addition, it is noteworthy that the decomposition approach is modular in the sense that it may be used for applications where the parameters are uncertain, that is subject to a distribution. In the sequel of this chapter a generalized Polynomial Chaos (gPC) approach is employed to obtain statistical information regarding the clearance criterion under investigation. Here, the upper level is replaced by a gPC-solver which requires the solution of deterministic problems in the lower level, i.e. the same two-level setup can be used.

8.3.2 Bi-level Worst-Case and Post-optimal Sensitivity Analysis

Under the proposed decomposition approach, the parameters $\mathbf{p}$ are optimized within the admissible region defined by $\mathbf{g}_p(\mathbf{p}) \leq \mathbf{0}$ in the upper level

$$\begin{aligned} \underset{\mathbf{p}}{\text{minimize}} \quad & \hat{j}^{LL}(\mathbf{p}) \\ \text{subject to} \quad & \mathbf{g}_p(\mathbf{p}) \leq \mathbf{0}, \end{aligned} \tag{8.121}$$

where $\hat{j}^{LL}(\mathbf{p})$ represents the optimal cost function value of the lower level problem depending on the values of the parameter $\mathbf{p}$. In the lower level, the trim constraint is imposed by explicitly solving the trim condition for the trim state $\hat{\mathbf{x}}_0(\mathbf{p})$ and trim control $\hat{\mathbf{u}}_0(\mathbf{p})$ based on the parameter vector $\mathbf{p} = \mathbf{p}_0$ provided from the upper level:

$$\mathbf{r}(\mathbf{x}_0(\mathbf{p}), \mathbf{u}_0(\mathbf{p}), \mathbf{p}) = \mathbf{0} \rightarrow \hat{\mathbf{x}}_0(\mathbf{p}), \hat{\mathbf{u}}_0(\mathbf{p}) \tag{8.122}$$

Hence, the implicit parameter dependencies of the system matrix $\mathbf{A}(\hat{\mathbf{x}}_0(\mathbf{p}), \hat{\mathbf{u}}_0(\mathbf{p}), \mathbf{p})$, control matrix $\mathbf{B}(\hat{\mathbf{x}}_0(\mathbf{p}), \hat{\mathbf{u}}_0(\mathbf{p}), \mathbf{p})$, as well as the cost vector $\mathbf{c}(\hat{\mathbf{x}}_0(\mathbf{p}), \mathbf{p})$ are exposed and

the parameter dependent LP problem in the lower level is

$$\begin{aligned}
\underset{\mathbf{u}(t)}{\text{minimize}} \quad & \mathbf{c}^T(\hat{\mathbf{x}}_0(\mathbf{p}),\mathbf{p})\mathbf{x}(t_f) \\
\text{subject to} \quad & \dot{\mathbf{x}}(t) = \mathbf{A}\left(\hat{\mathbf{x}}_0(\mathbf{p}),\hat{\mathbf{u}}_0(\mathbf{p}),\mathbf{p}\right)\left(\mathbf{x}(t)-\hat{\mathbf{x}}_0(\mathbf{p})\right) + \mathbf{B}\left(\hat{\mathbf{x}}_0(\mathbf{p}),\hat{\mathbf{u}}_0(\mathbf{p}),\mathbf{p}\right)\left(\mathbf{u}(t)-\hat{\mathbf{u}}_0(\mathbf{p})\right), \\
& \mathbf{x}(t_0) = \hat{\mathbf{x}}_0(\mathbf{p}), \\
& x_{j,lb} \leq x_j(t) \leq x_{j,lb},\ \forall j \in \mathcal{I}_b, \\
& \mathbf{u}(t) \in \mathcal{U}, \\
& t \in I^t,
\end{aligned} \tag{8.123}$$

which yields the optimal cost function value $\hat{j}^{LP}(\mathbf{p}) := \mathbf{c}^T(\hat{\mathbf{x}}_0(\mathbf{p}),\mathbf{p})\hat{\mathbf{x}}(t_f)$ depending on the parameter $\mathbf{p}$ and returns

$$\hat{j}^{LL}(\mathbf{p}) := \hat{j}^{LP}(\mathbf{p}) + j\left(\hat{\mathbf{x}}_0(\mathbf{p}),\mathbf{p}\right) - \mathbf{c}^T(\hat{\mathbf{x}}_0(\mathbf{p}),\mathbf{p})\hat{\mathbf{x}}_0(\mathbf{p}), \tag{8.124}$$

to the upper level. Note that using this decomposition approach the dynamic optimization problem (8.123) in the lower level is reduced to a Cat II type clearance problem. This problem can be solved to global optimality using the transcription process presented in Sec. 8.1.3. In the following the Trapezoidal method using a step size h is employed for which the discretized problem can be written as

$$\begin{aligned}
\underset{\mathbf{x}^{[i]},\,\mathbf{u}^{[i]},\,i=0,\ldots,N}{\text{minimize}} \quad & \mathbf{c}^T(\hat{\mathbf{x}}_0(\mathbf{p}),\mathbf{p})\mathbf{x}^{[N]} \\
\text{subject to} \quad & \mathbf{r}_0(\mathbf{p}) = \mathbf{A}_d^-(\mathbf{p})\mathbf{x}^{[i-1]} + \mathbf{A}_d^+(\mathbf{p})\mathbf{x}^{[i]} + \mathbf{B}_d(\mathbf{p})\left(\mathbf{u}^{[i]}+\mathbf{u}^{[i-1]}\right),\ i = 1,\ldots,N, \\
& \mathbf{x}^{[0]} = \hat{\mathbf{x}}_0(\mathbf{p}), \\
& x_{j,lb} \leq x_j^{[i]} \leq x_{j,ub},\ i = 1,\ldots,N, \forall j \in \mathcal{I}_b, \\
& \mathbf{u}_{lb} \leq \mathbf{u}^{[i]} \leq \mathbf{u}_{ub},\ i = 0,\ldots,N,
\end{aligned} \tag{8.125}$$

with

$$\mathbf{A}_d^+(\mathbf{p}) := \mathbf{I} - \frac{h}{2}\mathbf{A}(\hat{\mathbf{x}}_0(\mathbf{p}),\hat{\mathbf{u}}_0(\mathbf{p}),\mathbf{p}), \tag{8.126}$$

$$\mathbf{A}_d^-(\mathbf{p}) := -\mathbf{I} - \frac{h}{2}\mathbf{A}(\hat{\mathbf{x}}_0(\mathbf{p}),\hat{\mathbf{u}}_0(\mathbf{p}),\mathbf{p}), \tag{8.127}$$

$$\mathbf{B}_d(\mathbf{p}) := -\frac{h}{2}\mathbf{B}(\hat{\mathbf{x}}_0(\mathbf{p}),\hat{\mathbf{u}}_0(\mathbf{p}),\mathbf{p}), \tag{8.128}$$

$$\mathbf{r}_0(\mathbf{p}) := -h\left[\mathbf{A}(\hat{\mathbf{x}}_0(\mathbf{p}),\hat{\mathbf{u}}_0(\mathbf{p}),\mathbf{p})\hat{\mathbf{x}}_0(\mathbf{p}) + \mathbf{B}(\hat{\mathbf{x}}_0(\mathbf{p}),\hat{\mathbf{u}}_0(\mathbf{p}),\mathbf{p})\hat{\mathbf{u}}_0(\mathbf{p})\right]. \tag{8.129}$$

It is noteworthy, that the control and state variables subject to optimization are not relative to the trim condition and, as such, the state and control bounds represent absolute limits. This problem formulation can be written in the form of a parameter dependent

LP problem

$$\begin{aligned} \underset{\bar{\mathbf{z}} \in \mathbb{R}^{n_{\bar{z}}}}{\text{minimize}} \quad & \bar{\mathbf{v}}^T(\mathbf{p})\bar{\mathbf{z}} \\ \text{subject to} \quad & \mathbf{M}_h(\mathbf{p})\,\bar{\mathbf{z}} = \mathbf{b}_h(\mathbf{p}), \\ & \mathbf{M}_g\bar{\mathbf{z}} \leq \mathbf{b}_g, \end{aligned} \tag{8.130}$$

with the $n_{\bar{z}} = (N+1)n_x + (N+1)n_u$ dimensional parameter vector $\bar{\mathbf{z}} \in \mathbb{R}^{n_{\bar{z}}}$ defined as

$$\bar{\mathbf{z}} := \left[\left(\mathbf{x}^{[0]}\right)^T, \ldots, \left(\mathbf{x}^{[N]}\right)^T, \left(\mathbf{u}^{[0]}\right)^T, \ldots, \left(\mathbf{u}^{[N]}\right)^T\right]^T, \tag{8.131}$$

the cost vector $\bar{\mathbf{v}}(\mathbf{p}) \in \mathbb{R}^{n_{\bar{z}}}$ of the form

$$\bar{\mathbf{v}}(\mathbf{p}) := \left[\mathbf{0}, \ldots, \mathbf{c}^T(\hat{\mathbf{x}}_0(\mathbf{p}), \mathbf{p}), \mathbf{0}, \ldots, \mathbf{0}\right]^T, \tag{8.132}$$

as well as the equality constraint matrix $\mathbf{M}_h(\mathbf{p}) \in \mathbb{R}^{n_h \times n_{\bar{z}}}$, $n_h = (N+1)n_x$

$$\mathbf{M}_h(\mathbf{p}) := \begin{bmatrix} \mathbf{I} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_d^-(\mathbf{p}) & \mathbf{A}_d^+(\mathbf{p}) & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{B}_d(\mathbf{p}) & \mathbf{B}_d(\mathbf{p}) & \ldots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{A}_d^-(\mathbf{p}) & \mathbf{A}_d^+(\mathbf{p}) & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{B}_d(\mathbf{p}) & \mathbf{B}_d(\mathbf{p}) \end{bmatrix}, \tag{8.133}$$

and the right-hand side vector $\mathbf{b}_h(\mathbf{p}) \in \mathbb{R}^{n_h}$

$$\mathbf{b}_h(\mathbf{p}) := \begin{bmatrix} \hat{\mathbf{x}}_0(\mathbf{p}) \\ \mathbf{r}_0(\mathbf{p}) \\ \vdots \\ \mathbf{r}_0(\mathbf{p}) \end{bmatrix}. \tag{8.134}$$

The inequality constraint matrix $\mathbf{M}_g \in \mathbb{R}^{n_g \times n_{\bar{z}}}$, $n_g = (N-1)n_x + Nn_u$ and the right-hand side vector $\mathbf{b}_g \in \mathbb{R}^{n_g}$ have exactly the form in Eq. (8.50) with the only difference being that each of the vectors $\mathbf{x}_{lb}$ and $\mathbf{x}_{ub}$ is repeated $N-1$ times and the vectors $\mathbf{u}_{lb}$ and $\mathbf{u}_{ub}$ are repeated N times in the right-hand side vector $\mathbf{b}_g$. Note that in case a control

penalty is introduced, the problem formulation may be extended to the LP problem

$$\begin{aligned}
\underset{\mathbf{x}^{[i]},\,\mathbf{u}^{[i]}_{+/-},\,i=0,\ldots,N}{\text{minimize}} \quad & \mathbf{c}^T(\hat{\mathbf{x}}_0(\mathbf{p}),\mathbf{p})\mathbf{x}^{[N]} + \epsilon \sum_{i=0}^{N} \mathbf{1}\left(\mathbf{u}^{[i]}_{+} + \mathbf{u}^{[i]}_{-}\right) \\
\text{subject to} \quad & \mathbf{r}_0(\mathbf{p}) = \mathbf{A}^{-}_d(\mathbf{p})\mathbf{x}^{[i-1]} + \mathbf{A}^{+}_d(\mathbf{p})\mathbf{x}^{[i]} + \mathbf{B}_d(\mathbf{p})\left(\mathbf{u}^{[i]} + \mathbf{u}^{[i-1]}\right), i = 1,\ldots,N, \\
& \mathbf{x}^{[0]} = \mathbf{x}_0(\mathbf{p}), \\
& x_{j,lb} \leq x^{[i]}_j \leq x_{j,ub},\ i = 1,\ldots,N, \forall j \in \mathcal{I}_b, \\
& \mathbf{0} \leq \mathbf{u}^{[i]}_{+} \leq \mathbf{u}^{+}_{ub},\ i = 0,\ldots,N, \\
& \mathbf{0} \leq \mathbf{u}^{[i]}_{-} \leq \mathbf{u}^{-}_{lb},\ i = 0,\ldots,N, \\
& \mathbf{u}^{[i]} = -\mathbf{u}^{[i]}_{-} + \mathbf{u}^{[i]}_{+},\ i = 0,\ldots,N,
\end{aligned} \tag{8.135}$$

similar to the problem formulation (8.109) with suitable limits $\mathbf{u}^{+}_{ub}$ and $\mathbf{u}^{-}_{lb}$. The penalty constant ϵ is defined as

$$\epsilon := \frac{\epsilon_0}{(N+1)\mathbf{1}\left(\mathbf{u}_{ub} - \mathbf{u}_{lb}\right)}, \tag{8.136}$$

where ϵ_0 represents a tolerance for the clearance criterion which may be accepted from a practical perspective. The potential reduction of the cost function when removing the penalty is strictly smaller or equal to ϵ_0 in this case.

For typical clearance applications the dimension in the high level search, being the number of parameters n_p, is low compared to the lower level LP problem. Consequently, even a gradient-free, global algorithm can be employed. Global optimization methods which utilize values of the cost function at certain points in the parameter space are easily applied as the lower level merely needs to be solved and the cost function returned to the upper level. It is noteworthy that the operations required to evaluate the lower level are computationally cheap. The solution of the trim condition is typically achieved in the range of a few milliseconds and for the solution of the LP problem there exist highly efficient solvers (cf. Sec. 4.3). Besides the use of global algorithms, local gradient-based schemes may be employed which are very efficient in finding extremal points. The effectiveness of these methods can additionally be improved if the problem is started from randomized initial values which mitigates the local nature of these methods. The numerical examples presented at the end of this section indicate that the bi-level approach using gradient-based methods often converge in a few iterations which allows for the initialization of the optimization scheme from several tens or even hundreds of points. For gradient-based methods first and possibly second-order derivative information is required. The computation of derivatives for the cost function is more involved for the bi-level setup compared to a standard optimization problem due to the implicit dependence of the upper level on the optimal solution of the lower level problem. In particular, the upper level problem requires

derivative information of the cost function with respect to the parameters in the lower level problem. In the following the definitions from Sec. 4.3 for the standard form are used to streamline the discussion and the derivations. This standard form depending on the parameter vector $\mathbf{p}$ reads

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{R}^{n_z}}{\text{minimize}} \quad & \mathbf{c}_s^T(\mathbf{p})\mathbf{z} \\ \text{subject to} \quad & \mathbf{M}_s(\mathbf{p})\mathbf{z} = \mathbf{b}_s(\mathbf{p}), \\ & \mathbf{z} \geq \mathbf{0}, \end{aligned} \tag{8.137}$$

with the optimization variable vector $\mathbf{z} \in \mathbb{R}^{n_z}$, the cost vector, $\mathbf{c}_s \in \mathbb{R}^{n_z}$, constraint matrix $\mathbf{M}_s \in \mathbb{R}^{n_c \times n_z}$, and the right-hand side vector $\mathbf{b}_s(\mathbf{p}) \in \mathbb{R}^{n_c}$. The parameter dependent Lagrangian of this standard form is defined as:

$$\mathcal{L}\left(\mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{p}\right) := \mathbf{c}_s^T(\mathbf{p})\mathbf{z} + \boldsymbol{\lambda}^T\left(\mathbf{M}_s(\mathbf{p})\mathbf{z} - \mathbf{b}_s(\mathbf{p})\right) - \boldsymbol{\mu}^T\mathbf{z} \tag{8.138}$$

In order to provide efficient formulas for calculating the derivatives of the upper level problem, results from post-optimal sensitivity analysis are employed.

Post-optimal sensitivity analysis is well established for LP problems [86, 87, 56, 88]. In most references the sensitivities of the optimal point or the optimal cost function value are investigated separately depending on the type of influence in the different parts (right-hand side, constraint matrix, cost coefficients) of the LP problem. The theorem from Ref. [66] however provides the cost function derivatives in a very general setting, i.e. the dependency on the parameters within the LP problem can be essentially of arbitrary type. Let $J(\mathbf{z}; \mathbf{p}) = \mathbf{c}_s^T(\mathbf{p})\mathbf{z}$ denote the cost function of the parameter dependent problem and $\mathbf{h}(\mathbf{z}; \mathbf{p}) = \mathbf{M}_s(\mathbf{p})\mathbf{z} - \mathbf{b}_s(\mathbf{p})$ denote the equality constraints. The total derivative of the optimal cost function value is equal to the partial derivative of the Lagrangian with respect to the parameters $\mathbf{p}$ [66] which for LP problems can be expressed as:

$$\frac{dJ}{d\mathbf{p}} = \frac{\partial \mathcal{L}}{\partial \mathbf{p}} = \mathbf{z}^T\frac{\partial \mathbf{c}_s}{\partial \mathbf{p}} + \boldsymbol{\lambda}^T\frac{\partial \mathbf{h}}{\partial \mathbf{p}} \tag{8.139}$$

All quantities in Eq. (8.139) need to be evaluated at a KKT point $\left(\hat{\mathbf{z}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}\right)$ for a nominal parameter value $\mathbf{p} = \mathbf{p}_0$.

Using this first-order derivative information for the cost function from Eq. (8.139), a gradient method may be used in the upper level. In order to speed up convergence using Newton-type iterations it is advisable to provide not only information regarding the first derivatives but also the Hessian of the cost function, or an approximation thereof (cf. Sec. 3.4.1).

Superior convergence properties are typically obtained when utilizing the exact Hessian of the cost function, which requires the second-order post-optimal sensitiv-

ities of the cost function with respect to the parameters. This case is not covered in Ref. [66] but a formula for the general nonlinear case is provided in Ref. [65].

However, for LP problems this second-order information can be obtained very efficiently as shown in the following. Under the assumption of a basic feasible optimal solution for which the strict complementarity condition (4.25) holds, the following proposition provides a formula for the computation of second-order derivative information of the optimal cost function with respect to parameters.

Proposition: *Second-order Cost Sensitivity for Linear Programming*
Let $\left(\hat{\mathbf{z}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}\right)$ *be a basic feasible optimal point with basis* $\mathbf{M}_B$ *and basic variables* $\mathbf{z}_b > 0$ *of the parameter dependent LP* (8.137) *satisfying the strict complementarity condition. Then, the LP problem has a unique solution in a neighborhood around the nominal parameter and the second-order cost sensitivities can be expressed as:*

$$\frac{d^2 J}{d\mathbf{p}^2} = \frac{\partial^2 \mathcal{L}}{\partial \mathbf{p}^2} - \left(\mathbf{D} + \mathbf{D}^T\right), \qquad \mathbf{D} = \frac{\partial^2 \mathcal{L}}{\partial \mathbf{p} \partial \mathbf{z}_b} \mathbf{M}_B^{-1} \frac{\partial \mathbf{h}}{\partial \mathbf{p}} \tag{8.140}$$

The cost and constraint functions of the LP need to be at least twice continuously differentiable and all quantities are evaluated at the optimal point for a nominal value $\mathbf{p} = \mathbf{p}_0$.

Proof: Consider the augmented optimization problem (cf. Ref. [66])

$$\begin{aligned} \underset{\mathbf{z},\, \tilde{\mathbf{p}}}{\text{minimize}} \quad & \mathbf{c}^T(\tilde{\mathbf{p}})\mathbf{z} \\ \text{subject to} \quad & \mathbf{M}_s(\tilde{\mathbf{p}})\mathbf{z} = \mathbf{b}(\tilde{\mathbf{p}}), \\ & \tilde{\mathbf{p}} = \mathbf{p}, \\ & \mathbf{z} \geq \mathbf{0}, \end{aligned} \tag{8.141}$$

at a nominal parameter value $\mathbf{p} = \mathbf{p}_0$ with the corresponding Lagrangian $\mathcal{L}^A\left(\mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \tilde{\mathbf{p}}; \mathbf{p}\right)$:

$$\mathcal{L}^A\left(\mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \tilde{\mathbf{p}}; \mathbf{p}\right) := \mathbf{c}^T(\tilde{\mathbf{p}})\mathbf{z} + \boldsymbol{\lambda}^T\left(\mathbf{M}_s(\tilde{\mathbf{p}})\mathbf{z} - \mathbf{b}(\tilde{\mathbf{p}})\right) - \boldsymbol{\mu}^T\mathbf{z} + \boldsymbol{\eta}^T\left(\tilde{\mathbf{p}} - \mathbf{p}\right) \tag{8.142}$$

This Lagrangian of the augmented problem may be written in terms of the Lagrangian $\mathcal{L}\left(\mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \tilde{\mathbf{p}}\right)$ corresponding to the original problem (without the parameter embedding constraint) as:

$$\mathcal{L}^A\left(\mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \tilde{\mathbf{p}}; \mathbf{p}\right) = \mathcal{L}\left(\mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \tilde{\mathbf{p}}\right) + \boldsymbol{\eta}^T\left(\tilde{\mathbf{p}} - \mathbf{p}\right) \tag{8.143}$$

Moreover, the KKT conditions of the augmented problem can be expressed as follows

$$\left(\frac{\partial \mathcal{L}}{\partial \mathbf{z}}\right)^T = \mathbf{0}, \tag{8.144}$$

$$\left(\frac{\partial \mathcal{L}}{\partial \tilde{\mathbf{p}}}\right)^T + \boldsymbol{\eta} = \mathbf{0}, \tag{8.145}$$

$$\mathbf{h}(\mathbf{z}, \tilde{\mathbf{p}}) = \mathbf{0}, \tag{8.146}$$

$$\tilde{\mathbf{p}} - \mathbf{p} = \mathbf{0}, \tag{8.147}$$

$$\mathbf{Z}\boldsymbol{\mu} = \mathbf{0}, \tag{8.148}$$

$$\boldsymbol{\mu} \geq 0, \tag{8.149}$$

$$\mathbf{z} \geq 0, \tag{8.150}$$

with $\left(\frac{\partial \mathcal{L}}{\partial \mathbf{z}}\right)^T = \mathbf{M}_s^T(\tilde{\mathbf{p}})\boldsymbol{\lambda} - \boldsymbol{\mu} + \mathbf{c}(\tilde{\mathbf{p}})$, $\mathbf{h}(\mathbf{z}, \tilde{\mathbf{p}}) = \mathbf{M}_s(\tilde{\mathbf{p}})\mathbf{z} - \mathbf{b}(\tilde{\mathbf{p}})$, and $\mathbf{Z} = diag(\mathbf{z})$. Note that there are two additional conditions compared to the original problem due to the parameter embedding constraint, namely $\left(\frac{\partial \mathcal{L}}{\partial \tilde{\mathbf{p}}}\right)^T = -\boldsymbol{\eta}$ and $\tilde{\mathbf{p}} = \mathbf{p}$. Obviously, this augmented problem has the same optimal cost function value as the problem without the parameter embedding constraint. For the solution to remain optimal in a neighborhood of the nominal parameter $\mathbf{p}_0$ the total derivative $\frac{d\mathbf{F}}{d\mathbf{p}}$ of the first-order conditions (8.144)-(8.148) defined by

$$\mathbf{F}(\mathbf{v}(\mathbf{p}); \mathbf{p}) := \begin{bmatrix} \left(\frac{\partial \mathcal{L}}{\partial \mathbf{z}}\right)^T \\ \left(\frac{\partial \mathcal{L}}{\partial \tilde{\mathbf{p}}}\right)^T + \boldsymbol{\eta} \\ \mathbf{h}(\mathbf{z}, \tilde{\mathbf{p}}) \\ \tilde{\mathbf{p}} - \mathbf{p} \\ \mathbf{Z}\boldsymbol{\mu} \end{bmatrix} = \mathbf{0}, \tag{8.151}$$

with $\mathbf{v}^T := \left[\mathbf{z}^T, \tilde{\mathbf{p}}^T, \boldsymbol{\lambda}^T, \boldsymbol{\eta}^T, \boldsymbol{\mu}^T\right]$ needs to satisfy $\frac{d\mathbf{F}}{d\mathbf{p}} = \mathbf{0}$ which implies

$$\frac{\partial \mathbf{F}}{\partial \mathbf{v}} \frac{\partial \mathbf{v}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{F}}{\partial \mathbf{p}}, \tag{8.152}$$

or explicitly

$$\begin{bmatrix} \mathbf{0} & \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z} \partial \tilde{\mathbf{p}}} & \mathbf{M}_s^T & \mathbf{0} & \mathbf{I} \\ \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}} & \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}}^2} & \left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}}\right)^T & \mathbf{I} & \mathbf{0} \\ \mathbf{M}_s & \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{M} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Z} \end{bmatrix} \begin{bmatrix} \frac{d\mathbf{z}}{d\mathbf{p}} \\ \frac{d\tilde{\mathbf{p}}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\lambda}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\eta}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\mu}}{d\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{I} \\ \mathbf{0} \end{bmatrix}, \tag{8.153}$$

with $\mathbf{M} := diag(\boldsymbol{\mu})$. Due to the strict complementarity condition $\mathbf{z}+\boldsymbol{\mu} > \mathbf{0}$ the matrices $\mathbf{M}$ and $\mathbf{Z}$ can be expressed as

$$\mathbf{M} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_n \end{bmatrix}, \qquad \mathbf{Z} = \begin{bmatrix} \mathbf{Z}_b & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{8.154}$$

with $\mathbf{M}_n := diag\left(\boldsymbol{\mu}_n\right)$ and $\mathbf{Z}_b := diag\left(\mathbf{z}_b\right)$ as the variables $\boldsymbol{\mu}_b$ and $\mathbf{z}_n$ are zero at a basic feasible optimal point (see the definitions of basic and non-basic variables in Sec. 4.3). Together with the partitioning of $\mathbf{M}_s(\tilde{\mathbf{p}})$ in $[\mathbf{M}_B, \mathbf{M}_N]$, as well as $\mathbf{z}$ in $[\mathbf{z}_b, \mathbf{z}_n]$, and similarly $\boldsymbol{\mu}$ in $[\boldsymbol{\mu}_b, \boldsymbol{\mu}_n]$, Eq. (8.153) can be expanded to:

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}_b \partial \tilde{\mathbf{p}}} & \mathbf{M}_B^T & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}_n \partial \tilde{\mathbf{p}}} & \mathbf{M}_N^T & \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}_b} & \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}_n} & \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}}^2} & \left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}}\right)^T & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{M}_B & \mathbf{M}_N & \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Z}_b & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_n & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \frac{d\mathbf{z}_b}{d\mathbf{p}} \\ \frac{d\mathbf{z}_n}{d\mathbf{p}} \\ \frac{d\tilde{\mathbf{p}}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\lambda}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\eta}}{d\mathbf{p}} \\ \frac{d\boldsymbol{\mu}_b}{d\mathbf{p}} \\ \frac{d\boldsymbol{\mu}_n}{d\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{I} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{8.155}$$

The fifth row block matrix equation simply expresses the consistent result $\frac{d\tilde{\mathbf{p}}}{d\mathbf{p}} = \mathbf{I}$ for the total differential of the parameters $\tilde{\mathbf{p}}$ with respect to the parameters $\mathbf{p}$. The last two row block matrix equations yield

$$\mathbf{Z}_b \frac{d\boldsymbol{\mu}_b}{d\mathbf{p}} = \mathbf{0} \rightarrow \frac{d\boldsymbol{\mu}_b}{d\mathbf{p}} = \mathbf{0}, \tag{8.156}$$

$$\mathbf{M}_n \frac{d\mathbf{z}_n}{d\mathbf{p}} = \mathbf{0} \rightarrow \frac{d\mathbf{z}_n}{d\mathbf{p}} = \mathbf{0}, \tag{8.157}$$

as the matrices $\mathbf{M}_n$ and $\mathbf{Z}_b$ are both diagonal matrices with strictly positive entries on the main diagonal. Note that this result reflects the requirement that the variables $\boldsymbol{\mu}_b$ and $\mathbf{z}_n$ remain zero and the basis matrix $\mathbf{M}_B$ is unchanged in a neighborhood of the solution around $\mathbf{p}_0$. From the first and fourth row block matrix equations the total derivatives of the Lagrange multipliers $\boldsymbol{\lambda}$ and the basic variables $\mathbf{z}_b$ are obtained as

$$\frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}_b \partial \tilde{\mathbf{p}}} + \mathbf{M}_B^T \frac{d\boldsymbol{\lambda}}{d\mathbf{p}} = \mathbf{0} \rightarrow \frac{d\boldsymbol{\lambda}}{d\mathbf{p}} = -\left(\mathbf{M}_B^T\right)^{-1} \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}_b \partial \tilde{\mathbf{p}}}, \tag{8.158}$$

$$\mathbf{M}_B \frac{d\mathbf{z_b}}{d\mathbf{p}} + \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}} = \mathbf{0} \rightarrow \frac{d\mathbf{z_b}}{d\mathbf{p}} = -\mathbf{M}_B^{-1} \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}}. \tag{8.159}$$

Furthermore, for completeness, the second row block matrix equation yields:

$$\frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}_n \partial \tilde{\mathbf{p}}} + \mathbf{M}_N^T \frac{d\boldsymbol{\lambda}}{d\mathbf{p}} + \frac{d\boldsymbol{\mu}_n}{d\mathbf{p}} = \mathbf{0} \rightarrow \frac{d\boldsymbol{\mu}_n}{d\mathbf{p}} = -\frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}_n \partial \tilde{\mathbf{p}}} + \mathbf{M}_N^T \left(\mathbf{M}_B^T\right)^{-1} \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}_b \partial \tilde{\mathbf{p}}} \tag{8.160}$$

Finally, note that the negative multiplier $\boldsymbol{\eta}$ corresponding to the parameter embedding constraint represents the first-order cost sensitivity with respect to the parameter $\mathbf{p}$, i.e. $\frac{dJ}{d\tilde{\mathbf{p}}} = \frac{\partial \mathcal{L}}{\partial \tilde{\mathbf{p}}} = -\boldsymbol{\eta}^T$ (cf. Eq. (8.139)). Together with the third row block matrix equation

$$\frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}_b} \frac{d\mathbf{z}_b}{d\mathbf{p}} + \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}}^2} + \left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}}\right)^T \frac{d\boldsymbol{\lambda}}{d\mathbf{p}} + \frac{d\boldsymbol{\eta}}{d\mathbf{p}} = \mathbf{0}, \tag{8.161}$$

the relationship for the second-order cost sensitivity in Eq. (8.140) is readily derived

$$\begin{aligned}
\frac{d^2 J}{d\mathbf{p}^2} &= \frac{d}{d\mathbf{p}}\left(\frac{dJ}{d\mathbf{p}}\right) = -\frac{d\boldsymbol{\eta}}{d\mathbf{p}} && (8.162)\\
&= \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}_b} \frac{d\mathbf{z}_b}{d\mathbf{p}} + \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}}^2} - \left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}}\right)^T \frac{d\boldsymbol{\lambda}}{d\mathbf{p}} && (8.163)\\
&= -\frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}_b} \mathbf{M}_B^{-1} \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}} + \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}}^2} - \left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}}\right)^T \left(\mathbf{M}_B^T\right)^{-1} \frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}_b \partial \tilde{\mathbf{p}}} && (8.164)\\
&= -\frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}_b} \mathbf{M}_B^{-1} \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}} + \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}}^2} - \left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}}\right)^T \left(\mathbf{M}_B^{-1}\right)^T \left(\frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}_b}\right)^T && (8.165)\\
&= -\frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}_b} \mathbf{M}_B^{-1} \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}} + \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}}^2} - \left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}}\right)^T \left(\frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}_b} \mathbf{M}_B^{-1}\right)^T && (8.166)\\
&= -\frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}_b} \mathbf{M}_B^{-1} \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}} + \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}}^2} - \left(\frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{p}} \partial \mathbf{z}_b} \mathbf{M}_B^{-1} \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{p}}}\right)^T, && (8.167)
\end{aligned}$$

completing the proof.
Observe that the formula for the post-optimal second-order cost function sensitivity is efficient as the sparse linear system

$$\mathbf{M}_B \frac{d\mathbf{z_b}}{d\mathbf{p}} = -\frac{\partial \mathbf{h}}{\partial \mathbf{p}}, \tag{8.168}$$

only needs to be computed once and the remaining operations merely involve (typically sparse) matrix multiplications and additions.

The bi-level method is illustrated in the following using both models introduced in Sec. 8.2.1, that are the models for the lateral and the longitudinal plane. First, the post-optimal sensitivities with respect to parameters for nominal solutions from Sec. 8.2.2 are presented. These nominal solutions are obtained from the Cat I and Cat II problems using combined control inputs, i.e. the clearance problem (8.81) for the longitudinal plane and (8.101) for the lateral plane with $t_f = 6\ s$. Additionally, the trim condition is considered according to Eq. (8.124) from which the absolute value of the

worst-case cost function is obtained. It is noteworthy that besides the utilization of the post-optimal sensitivities in the bi-level setup they can be seen as a tool on its own in the context of optimal control based clearance to identify influential parameters. The author would like to emphasize that the post-optimal sensitivities discussed in the following represent the sensitivities for the *optimal solutions* of the nominal clearance problems. This means that, as opposed to the sensitivities regarding the solution to initial value problems (i.e. simulations), the post-optimal sensitivities indicate how the *worst-case* may change under parameter variations. These sensitivities can be obtained efficiently based on Eq. (8.139) for arbitrary parameters in the closed-loop system.

In the following the parameters for the post-optimal sensitivity analysis in the longitudinal plane represent uncertainties related to the pitch stiffness ΔM_α, pitch damping ΔM_q, elevator effectiveness ΔM_η, mass Δm, moment of inertia around the y_B-axis ΔI_{yy}, and the distance Δx^{RG} of the neutral point (N) from a reference point (R). All of these parameters are introduced as multiplicative uncertainties for the respective quantities with nominal value zero and the parameter Δx^{RG} is scaled such that for $\Delta x^{RG} = -1$ the center of gravity would coincide with the neutral point. Additionally, the reference velocity V_{ref} and reference height h_{ref} at the linearization point are considered as parameters. The sensitivities corresponding to a maximization of the normal load factor $n_z(t_f)$ and the aerodynamic AoA $\alpha_A(t_f)$ using the normal load factor command $n_{z,c}(t)$ in combination with the wind commands $u_{W,c}(t)$ and $w_{W,c}(t)$ are depicted in Fig. 8.30 and Fig. 8.31. In each chart the sensitivities are shown both for the Cat I (blue) and Cat II (black) type model. In order to facilitate the interpretation of the sensitivities the values are provided as normalized and absolute values. Observe that the indications regarding the parameter sensitivities related to the Cat I and Cat II type clearance problems are similar. This is most likely contributed to the fact that despite the singular intervals in the Cat II cases the worst-case inputs exhibit a comparable structure for the optimal control actions under Cat I and Cat II models (cf. Figs. 8.21-8.28). For both the maximization of the load factor (Fig. 8.30) and AoA (Fig. 8.31) high influences are identified for the parameters Δx^{RG} and ΔM_α. This result is not unexpected as both quantities are indicators for the static stability and in particular robustness with respect to disturbances in the longitudinal plane. Additionally, the parameters associated with the mass Δm and the moment of inertia ΔI_{yy} show high sensitivities. The reference velocity V_{ref} and height h_{ref} however appear to have little to no influence on the worst-case cost function values.

For the lateral plane the parameters under investigation are uncertainties related to the Dihedral stability ΔL_β, Weathercock stability ΔN_β, roll and yaw damping ΔL_p, ΔL_r, ΔN_r, aileron and rudder effectiveness ΔL_ξ, ΔL_ζ, ΔN_ζ as well as the moments of inertia around the x_B-axis ΔI_{xx} and the z_B-axis ΔI_{zz}. As for the longitudinal model all of these parameters represent multiplicative uncertainties centered at zero. The

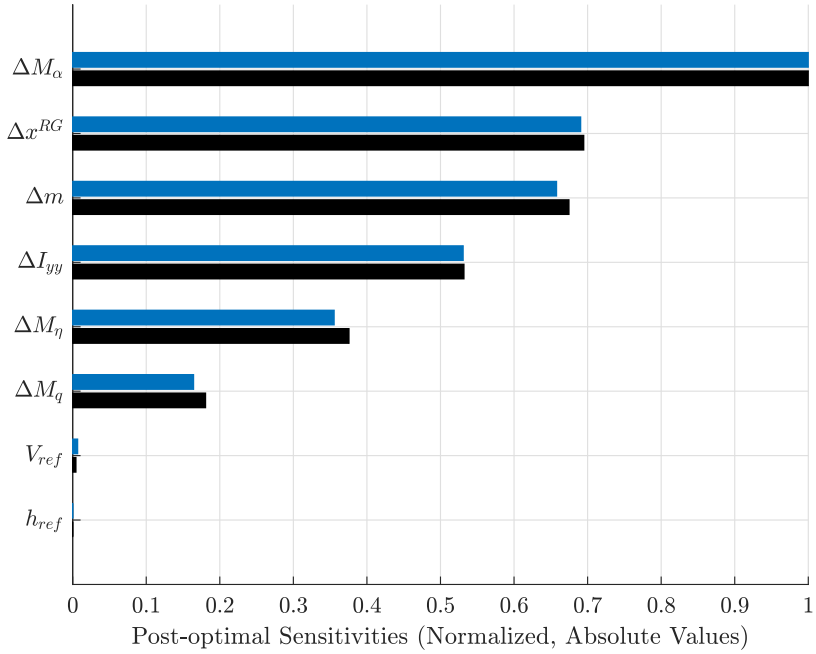


Figure 8.30: Post-optimal sensitivities for maximizing $n_z(t_f), t_f = 6\ s$ under a Cat I (blue) and Cat II (black) type model using the normal load factor command $n_{z,c}(t)$ as well as the wind commands $u_{W,c}(t)$ and $w_{W,c}(t)$.

sensitivities for the maximization of the lateral load factor $n_y(t_f)$ and the aerodynamic AoS $\beta_A(t_f)$ using the lateral load factor command $n_{y,c}(t)$ in combination with the roll angle command $\Phi_c(t)$ and wind command $v_{W,c}(t)$ are depicted in Fig. 8.32 and Fig. 8.33. For the lateral plane the uncertainties corresponding to the Weathercock stability ΔN_β and the moment of inertia ΔI_{zz} around the yaw axis are influential parameters for both the maximization of the lateral load factor (see Fig. 8.32) and the aerodynamic AoS (see Fig. 8.33). Moreover, the parameter related to the rudder effectiveness ΔN_ζ appears to play a crucial role for the maximization of the AoS (cf. Fig. 8.33). Note that the control actions primarily excite the yaw motion in the worst-case solutions (see Figs. 8.25-8.28). As such, it is reasonable that these quantities exhibit a high sensitivity regarding the optimal cost function values. Conversely, the quantities corresponding to the roll motion (roll moment coefficients and aileron effectiveness) only seem to play a minor role. As for the longitudinal model the reference velocity and height have a negligible effect on the optimal values of the criteria.

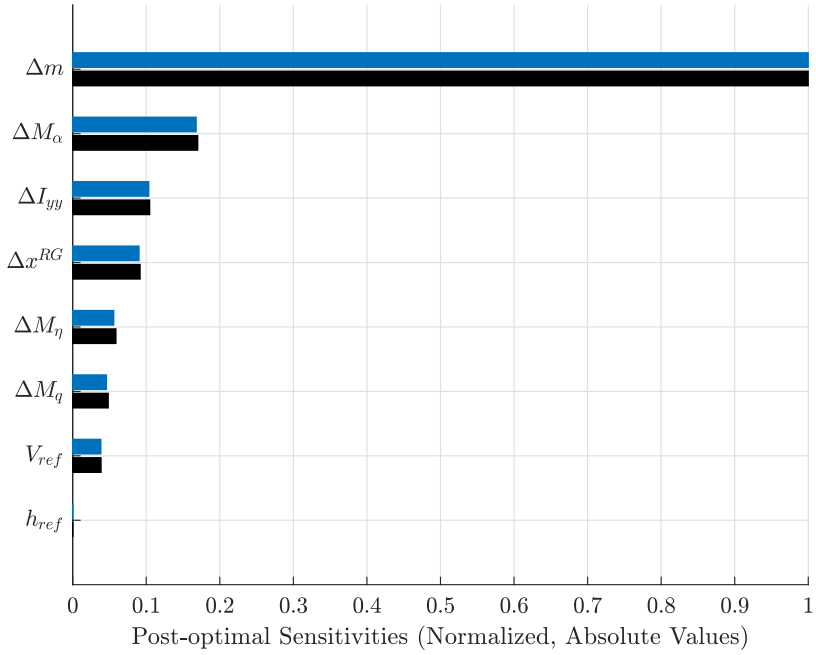


Figure 8.31: Post-optimal sensitivities for maximizing $\alpha_A(t_f), t_f = 6\ s$ under a Cat I (blue) and Cat II (black) type model using the normal load factor command $n_{z,c}(t)$ as well as the wind commands $u_{W,c}(t)$ and $w_{W,c}(t)$.

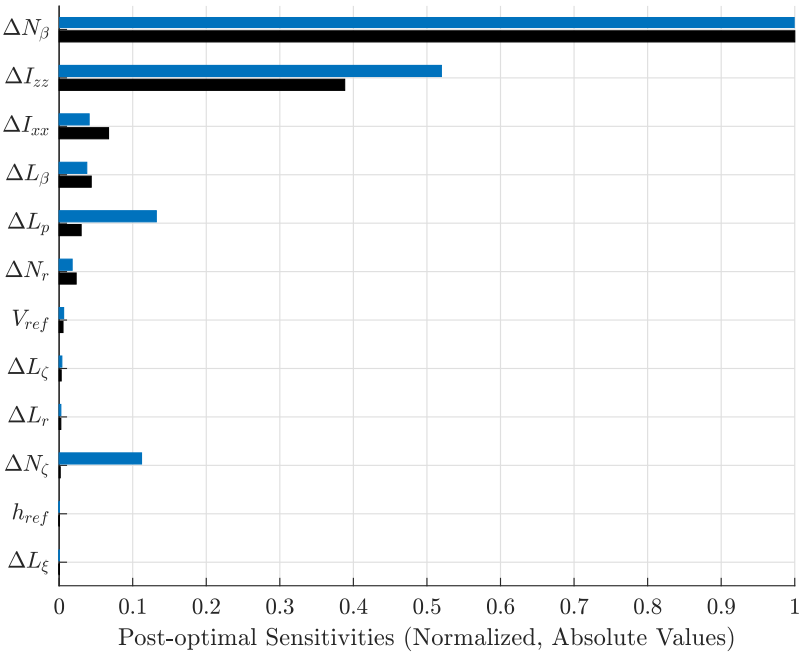


Figure 8.32: Post-optimal sensitivities for maximizing $n_y(t_f), t_f = 6\ s$ under a Cat I (blue) and Cat II (black) using the lateral roll angle command $\Phi_c(t)$, the lateral load factor command $n_{y,c}(t)$ as well as the wind command $v_{W,c}(t)$.

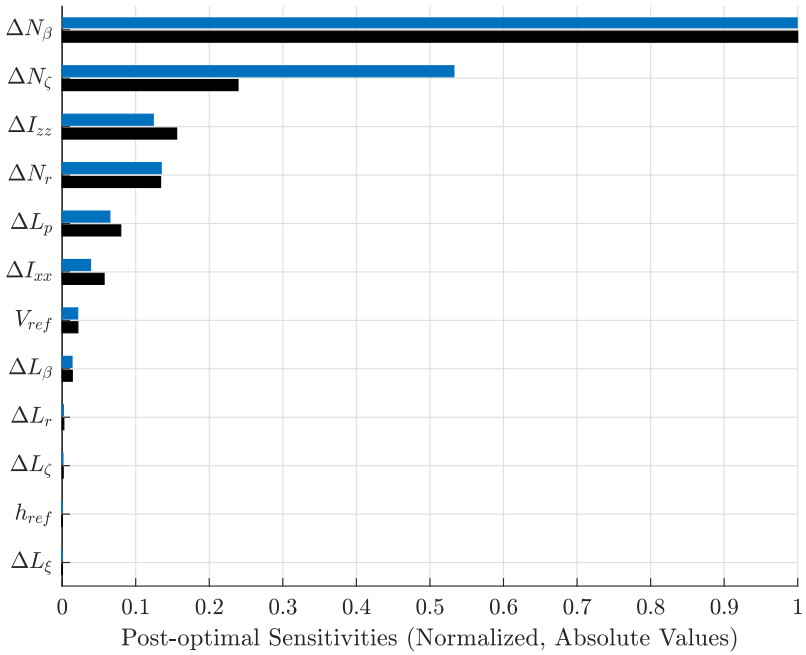


Figure 8.33: Post-optimal sensitivities for maximizing $\beta_A(t_f), t_f = 6\ s$ under a Cat I (blue) and Cat II (black) type model using the lateral roll angle command $\Phi_c(t)$, the lateral load factor command $n_{y,c}(t)$ as well as the wind command $v_{W,c}(t)$.

To illustrate the bi-level optimization approach a subset of influential parameters is considered. In case of the longitudinal plane these parameters are collected in the vector $\mathbf{p}_{lon} \in \mathbb{R}^6$:

$$\mathbf{p}_{lon} = \left[\Delta M_\alpha, \Delta M_q, \Delta M_\eta, \Delta m, \Delta I_{yy}, \Delta x^{RG}\right]^T \tag{8.169}$$

Similarly, for the lateral model the parameter vector $\mathbf{p}_{lat} \in \mathbb{R}^6$ is introduced:

$$\mathbf{p}_{lat} = \left[\Delta L_\beta, \Delta L_p, \Delta N_\beta, \Delta N_\zeta, \Delta I_{xx}, \Delta I_{zz}\right]^T \tag{8.170}$$

The acceptable accuracy of the clearance criteria is set to $\epsilon_0 = 0.001$ for all criteria which yields $\epsilon_0 \cdot 180 deg/\pi \approx 0.0573\ deg$ for the aerodynamic angles $\alpha_A(t_f)$ and $\beta_A(t_f)$. For all numerical results the Trapezoidal transcription (cf. LP problem (8.135)) is employed and the tolerance ϵ_0 is used to impose penalties on all controls according to Eq. (8.136). Moreover, the final time is set to $t_f = 8\ s \approx 7T_{SP}$ which is slightly higher compared to the nominal cases presented in Sec. 8.2.2 where a final time of $t_f = 6\ s$ is used. It is noteworthy that the approaches in the sequel of this chapter purely focus on the optimal value of the cost function. The convergence of the optimal criteria values for the longitudinal and lateral plane using the Cat II models with $h = 0.005\ s, \ldots, 0.25\ s$ are depicted in Figs. 8.34-8.37. Additionally, the normalized times required to numerically solve the LP problems are provided for all four criteria ($\alpha_A(t_f)$, $n_z(t_f)$, $\beta_A(t_f)$, $n_y(t_f)$, $t_f = 8\ s$). Moreover, the respective tolerance intervals are illustrated relative to the solution with the smallest discretization step length $h = 0.005\ s$ (green boxes). First, observe that the values of all criteria appear to exhibit values within the tolerances for relatively coarse discretization step lengths. In addition, for smaller step sizes $h \lesssim 0.05\ s$ the solution of the LPs appears to require a rapidly increasing amount of computational work with little effect on the worst-case values from a practical perspective. In the following the step length for all analyses is chosen as $h = 0.025\ s$ which based on the convergence studies is expected to yield sufficiently accurate approximations of the optimal cost function values for all criteria below the selected tolerances. Regarding the efficiency of the numerical solution of the LPs it is important to mention that the approaches discussed in the following rely on the solution of a high number of LP problems (> 1000 for some methods). In many cases these problems differ only slightly depending on the value of the parameter vector. Moreover, the structure of the discretized problem is unaltered if a constant discretization step length is used. To improve the efficiency for the solution of a large number of these similar clearance problems the optimal basis matrix returned after each solution of a LP problem is used to warm-start the next LP problem. Moreover, each warm-started LP problem is solved with the *Concurrent* strategy in CPLEX 12.9 for which several algorithms are run in an opportunistic parallel mode [57]. Computational experience suggests that these two strategies considerably speed up the numerical solution.

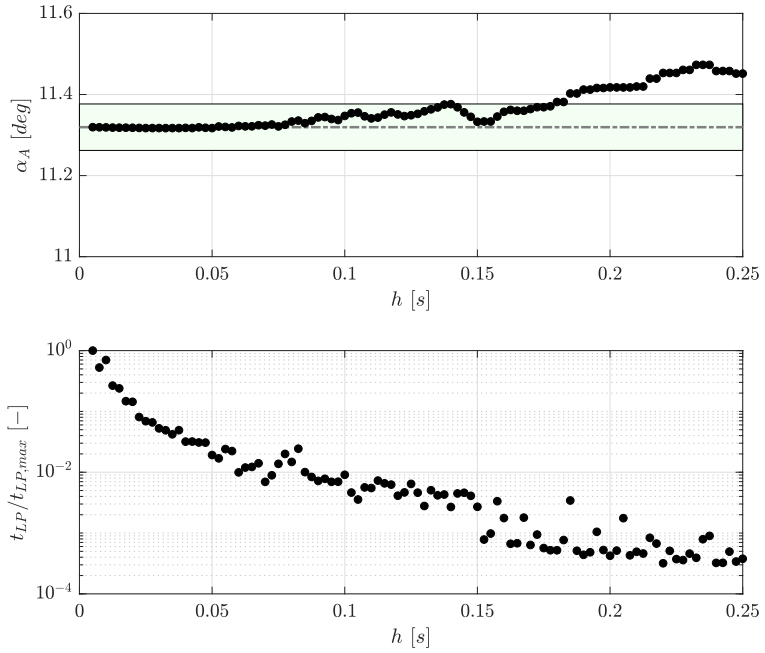


Figure 8.34: Maximization of $\alpha_A(t_f), t_f = 8\ s$ (Cat II): Optimal cost function values (top) and solution time (bottom) for discretization step sizes $h = 0.005, \ldots, 0.25\ s$.

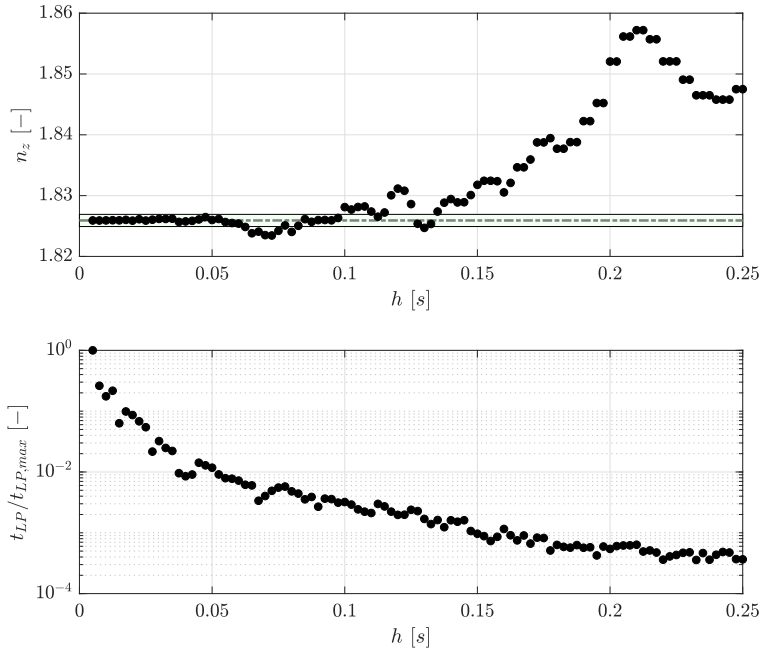


Figure 8.35: Maximization of $n_z(t_f), t_f = 8\ s$ (Cat II): Optimal cost function values (top) and solution time (bottom) for discretization step sizes $h = 0.005, \ldots, 0.25\ s$.

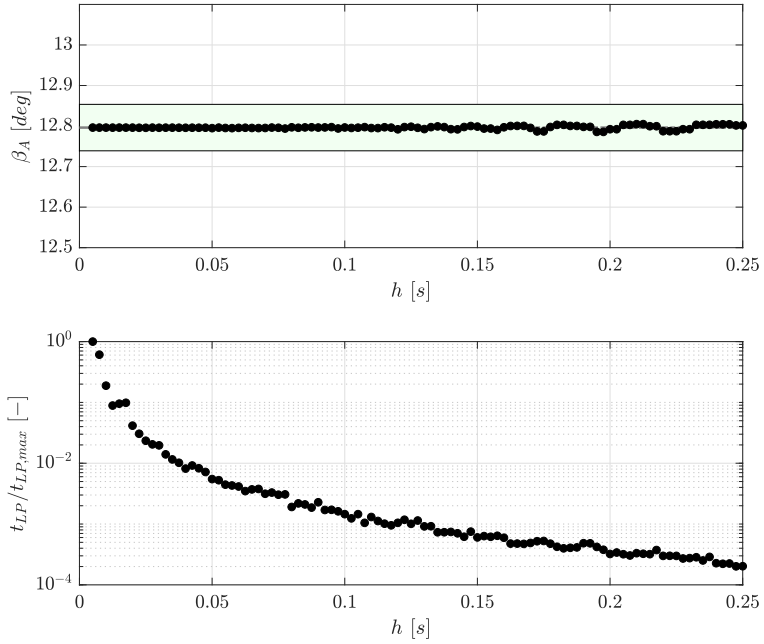


Figure 8.36: Maximization of $\beta_A(t_f), t_f = 8\ s$ (Cat II): Optimal cost function values (top) and solution time (bottom) for discretization step sizes $h = 0.005, \ldots, 0.25\ s$.

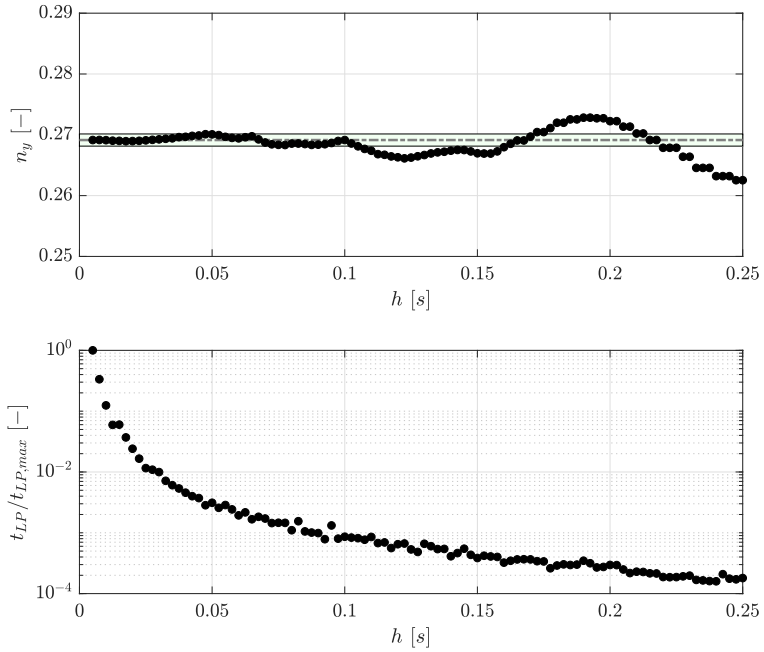


Figure 8.37: Maximization of $n_y(t_f), t_f = 8\ s$ (Cat II): Optimal cost function values (top) and solution time (bottom) for discretization step sizes $h = 0.005, \ldots, 0.25\ s$.

The bi-level optimization of the parameter dependent longitudinal and lateral Cat II models is performed using two different kinds of limits regarding the admissible parameter region. The first kind allows all parameters to vary between $\pm 10\%$, i.e. simple box bounds are used. For the second kind a subset of uncertain parameters (aerodynamic coefficients) are constrained to a region which contains a $1-\alpha$ portion of the probability mass. The motivation behind the second type of constraint is as follows: If no worst-case value in the admissible region is found which exceeds a certain critical limit and if the solution is global then the exceedance probability of this criterion is at most α.

For the box-bounded case the following optimization algorithms for the higher level are used to benchmark the approach:

- **Gradient-based (Jacobian)**: The Sequential Quadratic Programming (SQP) solver from the $fmincon()$-function in MATLAB® with Quasi-Newton Hessian approximation is used. This solver is able to handle nonlinear inequality constraints as well as linear equality and inequality constraints. The optimality tolerance for this solver is set to 10^{-4}.

- **Gradient-based (Jacobian and Hessian)**: The Trust-Region-reflective (TR) solver from the $fmincon()$-function in MATLAB® is employed. This solver can handle simple box bounds and linear equality constraints. As for the SQP solver the optimality tolerance is set to 10^{-4}.

- **Global (Stochastic)**: The surrogate optimization (SR) algorithm $surrogateopt()$ introduced in MATLAB® R2018B is used. This solver is designed for the global optimization of time-consuming objective functions subject to box-bounds and nonlinear inequality constraints. The algorithm constructs a surrogate model at random points using radial basis functions and searches for a minimum with a sampling based approach. It is found that this solver converges quickly for the clearance problems under investigation. As such, the maximum number of function evaluations is set to 50 as a termination criterion.

- **Global (Deterministic)**: The DIRECT algorithm (DR, cf. Ref. [19]) using the implementation from Ref. [89] is employed. This solver divides the search space with a particular strategy that balances local and global search subject to box-bounds. In particular, this algorithm "*is guaranteed to converge to the globally optimal function value if the objective function is continuous - or at least continuous in the neighborhood of a global optimum*" [19]. This arguments holds if the number of iterations goes to infinity - it is clear that in practice the search needs to be stopped at some point. Here, the maximum number of iterations is chosen as 50.

The results for the box-bounded case of the longitudinal Cat II model are summarized in Tab. 8.1 and Tab. 8.2. Moreover, the results for the lateral Cat II model are presented in Tab. 8.3 and Tab. 8.4. The reference values for the nominal problems are $\hat{\alpha}_{A,0}(t_f) = 11.3172\ deg$, $\hat{n}_{z,0}(t_f) = 1.8259$, $\hat{\beta}_{A,0}(t_f) = 12.7957\ deg$, and $\hat{n}_{y,0}(t_f) = 0.2691$. It is noteworthy that, as opposed to the results in Sec. 8.2.2, the values of the criteria presented here are not relative to the trim condition but absolute.

Observe that the worst-case criterion values at the optimal parameter combinations presented in Tabs. 8.1-8.4 are considerably higher compared to the nominal values. Moreover, the solvers seem to yield very consistent results regarding both the optimal criterion values and the worst-case parameter combinations. The parameters almost exclusively converge to one of the bounds. In some cases intermediate values are observed (see for example $\Delta\hat{L}_p$ in Tab. 8.3). However, for these cases the criterion values seem to be of comparable magnitude. This may be seen as an indication that the criterion is only depending weakly on these parameters and the solvers simply converge to one particular point based on the termination criterion. The gradient-based schemes appear to work very efficiently and converge within a few iterations. This is true for both the SQP solver with Quasi-Newton Hessian update and the TR solver with the exact Hessian. For the maximization of $\alpha_A(t_f)$ both solvers (SQP, TR) require the solution of a comparable amount of LPs until convergence. For the maximization of $n_z(t_f)$ the SQP method requires less solutions than the TR method and for the maximization of $\beta_A(t_f)$ and $n_y(t_f)$ the TR method appears to be more efficient. In all cases (cf. Tabs. 8.1-8.4) the worst-case parameter combinations of the Cat I and Cat II type clearance problems are very similar.

Note that for the longitudinal Cat I and Cat II models the pitch stiffness and pitch damping are decreased and the center of gravity is moved towards the neutral point. These parameter changes decrease the stability of the aircraft which seems a reasonable result from a flight mechanical perspective. Moreover, the elevator effectiveness is increased which essentially augments the effect of the worst-case control actions. Also the moment of inertia around the y_B-axis is increased, whereas the mass is on the upper bound for the maximization of $\alpha_A(t_f)$ and on the lower bound for the maximization of $n_z(t_f)$. In the lateral cases, i.e. the maximization of $\beta_A(t_f)$ (see Tab. 8.3) and $n_y(t_f)$ (see Tab. 8.4) the Weathercock and Dihedral stability is lowered and the moments of inertia for both axes (roll and yaw) are increased. As mentioned above the roll damping $\Delta\hat{L}_p$ appears to have little influence the worst-case value for the maximization of $\beta_A(t_f)$.

Table 8.1: Bi-level optimzation results (box-bounded parameters) for the maximization of $\alpha_A(t_f)$ using the normal load factor command $n_{z,c}(t)$ as well as the wind commands $u_W(t)$ and $w_W(t)$.

Solver	$\hat{\alpha}_A(t_f)$	n_{LP}	$\Delta\hat{M}_\alpha$	$\Delta\hat{M}_q$	$\Delta\hat{M}_\eta$	$\Delta\hat{m}$	$\Delta\hat{I}_{yy}$	$\Delta\hat{x}^{RG}$
SQP	14.166 $[deg]$	7	-0.100	-0.100	0.100	0.100	0.100	-0.100
TR	14.165 $[deg]$	6	-0.100	-0.100	0.100	0.100	0.100	-0.100
SG	14.164 $[deg]$	50	-0.100	-0.099	0.100	0.100	0.100	-0.100
DR	14.147 $[deg]$	1473	-0.100	-0.099	0.100	0.100	0.100	-0.099

Table 8.2: Bi-level optimzation results (Cat II, box-bounded parameters) for the maximization of $n_z(t_f)$ using the normal load factor command $n_{z,c}(t)$ as well as the wind commands $u_W(t)$ and $w_W(t)$.

Solver	$\hat{n}_z(t_f)$	n_{LP}	$\Delta\hat{M}_\alpha$	$\Delta\hat{M}_q$	$\Delta\hat{M}_\eta$	$\Delta\hat{m}$	$\Delta\hat{I}_{yy}$	$\Delta\hat{x}^{RG}$
SQP	2.222	3	-0.100	-0.100	0.100	-0.100	0.100	-0.100
TR	2.222	31	-0.100	-0.100	0.100	-0.100	0.100	-0.100
SG	2.222	50	-0.100	-0.100	0.100	-0.100	0.100	-0.100
DR	2.222	1041	-0.100	-0.100	0.100	-0.100	0.100	-0.100

Table 8.3: Bi-level optimzation results (Cat II, box-bounded parameters) for the maximization of $\beta_A(t_f)$ using the roll angle command $\Phi_c(t)$, lateral load factor command $n_y(t)$, and the wind command $v_W(t)$.

Solver	$\hat{\beta}_A(t_f)$	n_{LP}	$\Delta\hat{L}_\beta$	$\Delta\hat{L}_p$	$\Delta\hat{N}_\beta$	$\Delta\hat{N}_\zeta$	$\Delta\hat{I}_{xx}$	$\Delta\hat{I}_{zz}$
SQP	16.325 $[deg]$	13	-0.100	0.007	-0.100	0.100	0.100	0.100
TR	16.332 $[deg]$	7	-0.100	-0.100	-0.100	0.100	0.100	0.100
SG	16.343 $[deg]$	50	-0.091	0.100	-0.100	0.100	0.100	0.100
DR	16.326 $[deg]$	1053	-0.100	-0.100	-0.100	0.100	0.100	0.100

Table 8.4: Bi-level optimzation results (Cat II, box-bounded parameters) for the maximization of $n_y(t_f)$ using the roll angle command $\Phi_c(t)$, lateral load factor command $n_y(t)$, and the wind command $v_W(t)$

Solver	$\hat{n}_y(t_f)$	n_{LP}	$\Delta\hat{L}_\beta$	$\Delta\hat{L}_p$	$\Delta\hat{N}_\beta$	$\Delta\hat{N}_\zeta$	$\Delta\hat{I}_{xx}$	$\Delta\hat{I}_{zz}$
SQP	0.323	11	-0.100	-0.100	-0.100	0.100	0.100	0.100
TR	0.323	6	-0.100	-0.100	-0.100	0.100	0.100	0.100
SG	0.323	50	-0.100	-0.100	-0.100	0.100	0.100	0.100
DR	0.323	1115	-0.100	-0.100	-0.100	0.100	0.100	0.100

For the second type of parameter region the parameters Δx^{RG}, Δm, ΔI_{xx}, ΔI_{yy}, and ΔI_{zz} are again allowed to vary within $\pm 10\%$ as for the box-bounded case. However, all aerodynamic parameters that are ΔM_α,ΔM_q, and ΔM_η for the longitudinal model as well as ΔL_β, ΔL_p, ΔN_β, and ΔN_ζ for the lateral model are now subject to an uni-variate normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 0.05$. For these Gaussian distributions the admissible parameter space is defined as the region containing a probability of $1 - \alpha$. This region may be for example defined as an ellipsoidal-shaped region of the form [90]

$$g_c(\mathbf{p}; \alpha) := (\mathbf{p} - \mathbf{p}_0)^T \, \mathbf{\Sigma}^{-1} \, (\mathbf{p} - \mathbf{p}_0) - \chi^2_{n_p, 1-\alpha} \leq 0, \tag{8.171}$$

with the n_p-dimensional parameter vector $\mathbf{p} \in \mathbb{R}^{n_p}$, the vector of mean values $\mathbf{p}_0 \in \mathbb{R}^{n_p}$, and the positive definite covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{n_p \times n_p}$. Moreover, $\chi^2_{n_p,1-\alpha}$ represents the $(1 - \alpha)$-quantile of the χ^2 distribution with n_p degrees of freedom. In order to provide a general approach that can easily be used for other distribution types an over-approximation of this region is used which is defined as a convex polytope of the form $\mathbf{A}_c\mathbf{p} \leq \mathbf{b}_c$. In the following the three-dimensional polytope for the longitudinal plane is constructed based on the tangent planes at 14 support points on the surface defined by the level-set $g_c(\mathbf{p}; \alpha) = 0$ with $\alpha = 10^{-3}$ (see Fig. 8.38).

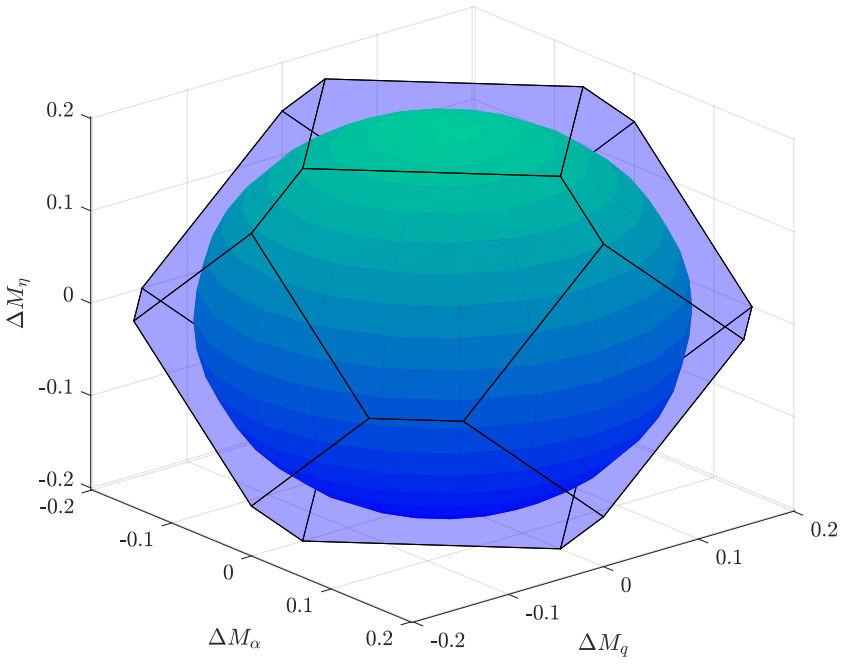


Figure 8.38: Illustration of the polytopal over-approximation of the admissible parameter region for ΔM_α, ΔM_q, and ΔM_η in the longitudinal plane.

Similarly, for the lateral plane the four-dimensional polytope is determined using 24 support points with the same definition of the level set and value for α. These conservative regions contain approximately a probability mass of $1 - \tilde{\alpha} \approx 1 - 3.67 \cdot 10^{-4}$ for the longitudinal and $1 - \tilde{\alpha} \approx 1 - 2.01 \cdot 10^{-4}$ for the lateral plane. Both figures are estimated based on 10^7 random samples. The numerical solution with these parameter regions are illustrated using the SQP method which inherently supports linear inequality constraints. The results are summarized in Tab. 8.5 for the longitudinal plane and Tab. 8.6 for the lateral plane. As for the box-bounded case the solver only requires the evaluation of a few lower level problems to converge. Moreover, observe that the criteria exhibit higher values compared to the box-bounded case. This can be attributed to the fact that the admissible parameter region allows for higher (absolute) parameter values at the vertices of the polytopal regions compared to the box-bounded case.

Table 8.5: Bi-level optimzation results (Cat II, probability-bounded parameters) for the maximization of $\alpha_A(t_f)$ and $n_z(t_f)$ with the SQP method using the normal load factor command $n_{z,c}(t)$ as well as the wind commands $u_W(t)$ and $w_W(t)$.

Criterion	$\hat{j}^{LL}$	n_{LP}	$\Delta\hat{M}_\alpha$	$\Delta\hat{M}_q$	$\Delta\hat{M}_\eta$	$\Delta\hat{m}$	$\Delta\hat{I}_{yy}$	$\Delta\hat{x}^{RG}$
$\alpha_A(t_f)$	14.867 $[deg]$	15	-0.202	0.000	0.148	0.100	0.100	-0.100
$n_z(t_f)$	2.402	7	-0.202	-0.148	0.000	-0.100	0.100	-0.100

Table 8.6: Bi-level optimzation results (Cat II, probability-bounded parameters) for the maximization of $\beta_A(t_f)$ and $n_y(t_f)$ with the SQP method using the roll angle command $\Phi_c(t)$, lateral load factor command $n_y(t)$, and the wind command $v_W(t)$.

Criterion	$\hat{j}^{LL}$	n_{LP}	$\Delta\hat{L}_\beta$	$\Delta\hat{L}_p$	$\Delta\hat{N}_\beta$	$\Delta\hat{N}_\zeta$	$\Delta\hat{I}_{xx}$	$\Delta\hat{I}_{zz}$
$\beta_A(t_f)$	20.949 $[deg]$	11	0.000	-0.000	-0.215	0.215	0.100	0.100
$n_y(t_f)$	0.393	35	0.000	0.000	-0.215	0.215	-0.100	0.100

The results in this section indicate that the proposed decomposition approach seems to be suitable to solve the parameter dependent Cat I and Cat II type clearance problems. In particular, gradient-based methods appear to be not only efficient as they converge within a few steps but, most importantly, attain essentially the same optimal values as the global optimization algorithms. Moreover, these methods have well defined termination criteria based on the KKT conditions (cf. Sec. 4). Note that using the proposed decomposition approach also global optimization algorithms can be used in the upper level. However, for these methods it is difficult to define meaningful termination criteria for the clearance task. From a practical perspective a simple yet most likely effective method for the bi-level approach is to start the local gradient-based solvers from several randomized initial points in order to increase the chance of finding the global worst-case.

8.3.3 Uncertainty Quantification using Generalized Polynomial Chaos

Note that the bi-level approach presented in the last section is able to determine a single worst-case solution in the parameter region. In the following the quantification of the uncertainty related to a particular criterion for distributed parameters is under investigation (Contribution II, [85]). Uncertainty quantification based on gPC [91, 92] has gained rapidly increasing popularity in the last decades. In particular, gPC using the stochastic collocation approach is a popular choice mainly due to its general applicability and relatively straight-forward implementation. For this approach the user merely needs to specify a set of nodes and run deterministic realizations. After a single post-processing step, that is the calculation of expansion coefficients, statistical moments are readily available. It is particularly noteworthy that "*the applicability of stochastic collocation is not affected by the complexity or nonlinearity of the original problem so long as one can develop a reliable deterministic solver*" [93]. Note that with the transcription of Cat I and Cat II type clearance problems to LP problems very reliable and efficient solvers are at hand. In particular, the global worst-case value of the criterion under investigation for the nominal Cat I and II type models can be obtained.

This section extends a preliminary study concerning gPC based uncertainty quantification under Cat I and Cat II models previously published by the author in Ref. [85]. First, the basic theoretical background for uncertainty quantification using gPC is summarized. For extended introductions regarding this topic please refer to Refs. [91, 92, 93]. In the following the uncertain parameters $\mathbf{p} \in \mathbb{R}^{n_p}$ for the clearance problem are considered as mutually independent, random variables with probability density functions $\rho_i(p_i), i = 0, \ldots, n_p - 1$. This implies that the joint probability density function $\rho(\mathbf{p})$ can be expressed as:

$$\rho(\mathbf{p}) := \prod_{i=0}^{n_p-1} \rho_i\left(p_i\right). \tag{8.172}$$

The random spaces of the parameters $p_i, i = 0, \ldots, n_p - 1$ are denoted with Ω_i and

$$\Omega := \prod_{i=0}^{n_p-1} \Omega_i, \tag{8.173}$$

represents the joint random space. For each random variable an uni-variate polynomial space $W^{i,q_i}, i = 0, \ldots, n_p - 1$ of highest degree q_i is defined

$$W^{i,q_i} := \left\{ w : \Omega_i \to \mathbb{R} : w \in span\left\{\psi_{i,j}(p_i)\right\}_{j=0}^{q_i} \right\}. \tag{8.174}$$

The polynomials $\psi_{i,j}(p_i), j = 0, \ldots, q_i$ need to satisfy the orthogonality conditions

$$\langle \psi_{i,l}, \psi_{i,r} \rangle = \int_{\Omega_i} \psi_{i,l}(p_i)\psi_{i,r}\left(p_i\right)\rho_i\left(p_i\right)\mathrm{d}p_i = h_l^2\delta_{l,r}, \qquad l, r = 0, \ldots, q_i, \tag{8.175}$$

with the Kronecker delta

$$\delta_{l,r} := \begin{cases} 1, \ l = r \\ 0, \ \text{otherwise} \end{cases}, \tag{8.176}$$

and are selected based on the distribution type (see Tab. 8.7). Without loss of generality it is assumed in the following that the basis functions are normalized with respect to the normalization factors:

$$h_l^2 := \int_{\Omega_i} \psi_{i,l}(p_i)\psi_{i,l}\left(p_i\right)\rho_i\left(p_i\right)\mathrm{d}p_i, \ l = 0, \ldots, q_i \tag{8.177}$$

Table 8.7: Polynomial basis for different continuous distribution types [92]

Distribution	Polynomial Basis Type	Support
Normal	Hermite	$(-\infty, \infty)$
Gamma	Laguerre	$[0, \infty)$
Beta	Jacobi	$[-1, 1]$
Uniform	Legendre	$[-1, 1]$

The n_p-variate orthogonal polynomial space $W_{n_p}^q$ is defined as the full tensor product space of the uni-variate polynomial spaces $W^{i,q_i}, i = 0, \ldots, n_p - 1$. Let the polynomials from $W_{n_p}^q$ be denoted with $\phi_k(\mathbf{p}), k = 0, \ldots, n_\gamma - 1$. The approximation of the criterion function $y : \Omega \to \mathbb{R}$ can be constructed using an expansion of the form

$$y(\mathbf{p}) := \sum_{k=0}^{n_\gamma - 1} \gamma_k \phi_k(\mathbf{p}), \tag{8.178}$$

with n_γ expansion coefficients $\gamma_k, k = 0, \ldots, n_\gamma - 1$. Using a spectral projection, the expansion coefficients are defined as follows:

$$\gamma_k := \int_\Omega y(\mathbf{p})\phi_k(\mathbf{p})\rho\left(\mathbf{p}\right)\mathrm{d}\mathbf{p} \tag{8.179}$$

This integral can be approximated using a quadrature

$$\gamma_k = \sum_{j=0}^{n_b - 1} y\left(\mathbf{p}^{[j]}\right)\phi_k\left(\mathbf{p}^{[j]}\right)b_j, \tag{8.180}$$

with weights $b_j, j = 0, \ldots, n_b - 1$ and quadrature nodes $\mathbf{p}^{[j]}, j = 0, \ldots, n_b - 1$. These weights and nodes need to be computed depending on the distribution type, i.e. based on Gauss-Hermite, Gauss-Laguerre, Gauss-Jacobi, and Gauss-Legendre quadrature methods for the types provided in Tab. 8.7. From the expansion coefficients the moments can be easily derived. In particular, the first expansion coefficient γ_0 represents the mean and the variance can be obtained from the sum of the squared second to last expansion coefficients:

$$\mathrm{Var}(y(\mathbf{p})) = \sum_{i=1}^{n_\gamma - 1} \gamma_k^2 \tag{8.181}$$

It is noteworthy, that the highest degree of the joint polynomial space $W_{n_p}^q$ is often truncated. Moreover, for higher dimensional parameter spaces it is advisable to use sparse grids in order to alleviate the curse of dimensionality associated with the full tensor product space.

To illustrate the application of the gPC approach the mean and standard deviation of the criteria for the longitudinal Cat II model ($\alpha_A(t_f)$,$n_z(t_f)$) and lateral Cat II model ($\beta_A(t_f)$,$n_y(t_f)$) are investigated using the same lower level problems as in the numerical examples of the bi-level approach (cf. Sec. 8.3.2). As such, merely the upper level needs to be replaced by the gPC solver and the optimal control problems are solved for each stochastic collocation point in the parameter space. The uncertain parameters considered in the following are the aerodynamic uncertainties ΔM_α,ΔM_q, and ΔM_η for the longitudinal model and the uncertainties ΔL_β, ΔL_p, ΔN_β, ΔN_ζ for the lateral model. All uncertain parameters are assumed to have a normal distribution with zero mean and a standard deviation $\sigma = 0.05$. Hence, uni-variate Hermite-polynomials (cf. Tab. 8.7) are used for the stochastic collocation approach. Two different orders, that are $q_i = 3, i = 0, \ldots, n_p - 1$ and $q_i = 4, i = 0, \ldots, n_p - 1$ are used for all uni-variate parameter spaces. For the longitudinal model ($n_p = 3$) the gPC approach thus requires the solution of $3^3 = 27$ LPs for the third-order approximation and $3^4 = 64$ LPs for the fourth-order approximation. For the lateral model ($n_p = 4$) $4^3 = 64$ LP solutions for the third-order approximation and $4^4 = 256$ LP solutions for the fourth-order approximation are necessary to compute the expansion coefficients. The results from the gPC approach are compared to a Monte Carlo analysis with $n_{MC} = 10^3$ Latin-Hypercube samples (MATLAB®-function *lhsdesign()*) in Tab. 8.8 for the mean and Tab. 8.9 for the standard deviation. Additionally, $(1 - \alpha)$-confidence intervals for the mean $\mathcal{C}_{\mu,MC}^{[1-\alpha]}$ [94]

$$\mathcal{C}_{\mu,MC}^{[1-\alpha]} := \left[\mu_{MC} - t_{n_{MC}-1,\alpha/2} \frac{\sigma_{MC}}{\sqrt{n_{MC}}}, \mu_{MC} + t_{n_{MC}-1,\alpha/2} \frac{\sigma_{MC}}{\sqrt{n_{MC}}} \right], \tag{8.182}$$

with the upper half tail percentage of Student's t-distribution $t_{n_{MC}-1,\alpha/2}$ with $n_{MC} - 1$

degrees of freedom and the standard deviation $\mathcal{C}_{\sigma,MC}^{[1-\alpha]}$ [95]

$$\mathcal{C}_{\sigma,MC}^{[1-\alpha]} := \left[\sqrt{\frac{n_{MC}-1}{\chi^2_{n_{MC}-1,1-\alpha/2}}}\sigma_{MC}, \sqrt{\frac{n_{MC}-1}{\chi^2_{n_{MC}-1,\alpha/2}}}\sigma_{MC} \right], \tag{8.183}$$

with the quantiles $1-\alpha/2$ and $\alpha/2$ of the χ^2 distribution with $n_{MC}-1$ degrees of freedom are provided.

Table 8.8: Comparison of the mean between gPC and MC including confidence intervals.

Criterion	$\mu_{gPC}^{[3]}$	$\mu_{gPC}^{[4]}$	μ_{MC}	$\mathcal{C}_{\mu,MC}^{[0.95]}$
$\alpha_A(t_f)$ $[deg]$	11.2024	11.2012	11.2054	$[11.1755, 11.2353]$
$n_z(t_f)$ $[-]$	1.8304	1.8304	1.8305	$[1.8282, 1.8328]$
$\beta_A(t_f)$ $[deg]$	13.1136	13.1136	13.1087	$[13.0570, 13.1604]$
$n_y(t_f)$ $[-]$	0.2722	0.2722	0.2721	$[0.2713, 0.2729]$

Table 8.9: Comparison of the standard deviation between gPC and MC including confidence intervals.

Criterion	$\sigma_{gPC}^{[3]}$	$\sigma_{gPC}^{[4]}$	σ_{MC}	$\mathcal{C}_{\sigma,MC}^{[0.95]}$
$\alpha_A(t_f)$ $[deg]$	0.4862	0.4873	0.4811	$[0.4609, 0.5031]$
$n_z(t_f)$ $[-]$	0.0369	0.0369	0.0369	$[0.0354, 0.0386]$
$\beta_A(t_f)$ $[deg]$	0.8189	0.8189	0.8330	$[0.7980, 0.8712]$
$n_y(t_f)$ $[-]$	0.0125	0.0125	0.0123	$[0.0118, 0.0129]$

Note that the results from the gPC approach match well with the results obtained from the MC analysis and, in particular, are within the 95% confidence intervals. The gPC approach is expected to be efficient in lower dimensional parameter spaces as the number of stochastic collocation points grows exponentially with the number of dimensions. In higher dimensions (typically $n_p > 5$) sparse grids can be used instead of the full tensor grid. However, for very large parameter dimensions the MC approach should be preferred due to the weak dependence on the number of parameters. Note that both the gPC and MC approach allow for a considerable parallelization as all samples can be computed independently. Another interesting observation is that the gPC approach directly yields an approximation of the criterion in the parameter space (cf. Eq. (8.178)) which can be used as a surrogate model of the response.

It is important to mention that in the context of flight control law clearance the exceedance probability $\mathcal{P}(y \geq y_{max})$ of a criterion y with limit y_{max} is of particular interest. Under the assumption that there exist a finite mean μ and a finite, non-zero variance $Var(y) = \sigma^2$ a bound for this probability is given by the one-sided Chebyshev-

Cantelli inequality:

$$\mathcal{P}\left(y \geq y_{max}\right) \leq \frac{Var(y)}{Var(y) + \left(y_{max} - \mu\right)^2},\ y_{max} > \mu \tag{8.184}$$

Besides the Chebyshev-Cantelli inequality, there exist other bounds based on higher moments (cf. Ref. [96]). To provide an example regarding the clearance criteria investigated in this section consider the maximum normal load factor $n_z(t_f)$ with an upper limit of $n_{z,max}(t_f) = 3.8$. Using the results for the mean and standard deviation of the fourth-order gPC method (cf. Tab. 8.8 and Tab. 8.9) the exceedance probability for this criterion is bounded by

$$\mathcal{P}\left(n_z(t_f) \geq n_{z,max}(t_f)\right) \leq \frac{\left(\sigma_{gPC}^{[4]}\right)^2}{\left(\sigma_{gPC}^{[4]}\right)^2 + \left(n_{z,max}(t_f) - \mu_{gPC}^{[4]}\right)^2} \approx 3.5 \cdot 10^{-4}, \tag{8.185}$$

according to the inequality (8.184).

8.4 Multi-Criteria and Reachability Analysis

All the approaches presented in the last sections are focused on a single criterion. However, it may be of interest for clearance applications which combination of criterion values may occur. Here, reachable set methods can provide valuable information regarding the attainable set in the output space of a closed-loop system. In the following an algorithm tailored to the Cat I and Cat II type clearance problems is presented which can be used to compute a tight over-approximation of the forward reachable set at a time point t_f starting from the trim state $\mathbf{x}(t_0) = \mathbf{0}$ (Contribution II). All quantities are in the following regarded relative to the trim condition which is general in the sense that the origin merely needs to be shifted in order to obtain the set in absolute coordinates (the shape of the set remains unchanged).

For linear systems it is possible to construct reachable sets using basic set operations such as sums, differences, and over-approximations of convex sets [97, 98]. However, it is important to mention that the computational effort for these methods becomes increasingly costly for higher dimensions as the number of facets grows very quickly. Often, inner and outer approximations are obtained using simple geometrical bodies such as ellipsoids (cf. Refs. [99, 100]), zonotopes (cf. Refs. [101, 102]), and parallelotopes (cf. Refs. [103, 104]). Moreover, optimal control based methods are presented in Refs. [105, 106] for reachable set approximations under input constraints. In the context of the optimal control based approaches investigated in this thesis these methods naturally fit into the clearance framework. Note that for Cat II type problems the reachable set needs to be computed considering the convex control constraint set $\mathcal{U}$

$$\mathbf{u}_{lb} \leq \mathbf{u}(t) \leq \mathbf{u}_{lb}, \tag{8.186}$$

and convex state constraint set $\mathcal{X}$

$$x_{j,lb} \leq x_j(t) \leq x_{j,ub}, \forall j \in \mathcal{I}_b. \tag{8.187}$$

It is noteworthy that the reachable set of linear systems with convex input and state constraints is convex and compact [107]. As such, it is sufficient to compute a conservative outer approximation of the convex boundary of the set. Clearly, it is desirable to provide bounds for the accuracy of this approximation and to refine the set until a certain threshold is reached. For this purpose an algorithm considering the Cat I and Cat II type clearance problems is presented in the following based on the ideas in Refs. [105, 106]. In particular, this algorithm computes an inner approximation through the successive optimization along rays in the output space. The outer approximation is constructed using the collection of tangent planes at the vertices of the inner approximation (see Fig. 8.39). These tangent planes are obtained using post-optimal sensitivity

analysis. Note that the boundary of the true reachable set is guaranteed to lie within the inner and outer approximation. Hence, the strategy of the algorithm refines both approximations until the maximum distance between the two is below a specified tolerance ϵ_0.

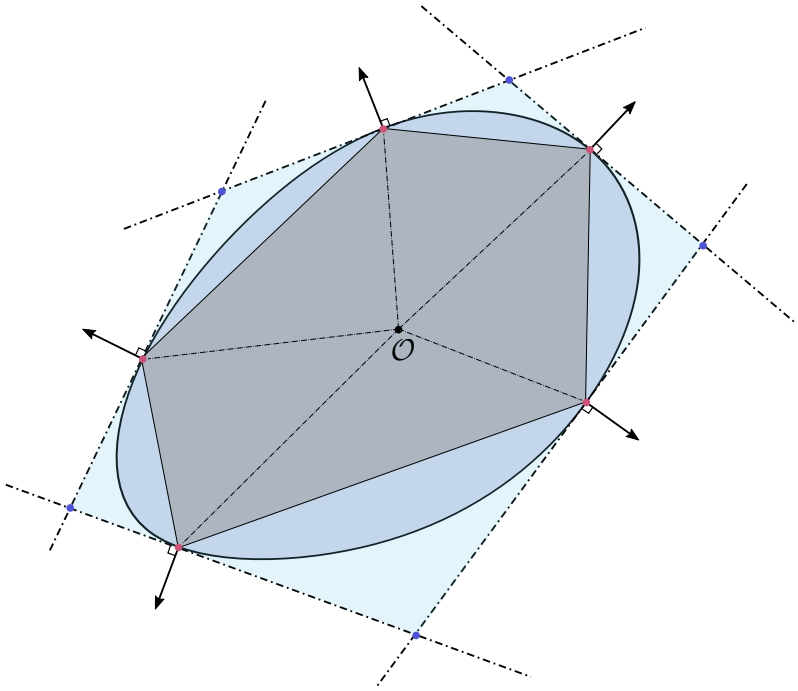


Figure 8.39: Illustration of the inner approximation (grey region/red vertices) and outer approximation (union of the light blue region, dark blue region, and the grey region / blue vertices) of a convex reachable set (union of the dark blue region and the grey region) centered at the origin $\mathcal{O}$.

The LP for the maximization of the distance $r \in \mathbb{R}$ from the origin along a ray defined by angles $\boldsymbol{\phi} \in \mathbb{R}^{n_y-1}$ is performed with the optimization problem LPMRAY($\boldsymbol{\phi}, \mathcal{U}, \mathcal{X}$):

$$\begin{aligned} \underset{r, \mathbf{x}^{[i]}, \mathbf{u}^{[i]}, i=0,\dots,N}{\text{minimize}} \quad & -r \\ \text{subject to} \quad & \mathbf{0} = \mathbf{A}_d^- \mathbf{x}^{[i-1]} + \mathbf{A}_d^+ \mathbf{x}^{[i]} + \mathbf{B}_d \left(\mathbf{u}^{[i]} + \mathbf{u}^{[i-1]} \right), \ i = 1, \dots, N, \\ & \mathbf{x}^{[0]} = \mathbf{0}, \\ & \mathbf{h}\left(r, \boldsymbol{\phi}, \mathbf{x}^{[N]} \right) = \mathbf{0}, \\ & \mathbf{u}^{[i]} \in \mathcal{U}, \ i = 0, \dots, N, \\ & \mathbf{x}^{[i]} \in \mathcal{X}, \ i = 1, \dots, N \end{aligned} \tag{8.188}$$

In alignment with the last sections a Trapezoidal scheme with step length h is used for the discretization of the dynamic constraints and the matrices for the Trapezoidal

discretization $\mathbf{A}_d^+$, $\mathbf{A}_d^-$, and $\mathbf{B}_d$ are (cf. Eqs. (8.126)-(8.128)):

$$\mathbf{A}_d^+ := \mathbf{I} - \frac{h}{2}\mathbf{A}, \tag{8.189}$$

$$\mathbf{A}_d^- := -\mathbf{I} - \frac{h}{2}\mathbf{A}, \tag{8.190}$$

$$\mathbf{B}_d := -\frac{h}{2}\mathbf{B} \tag{8.191}$$

The linear equality constraint $\mathbf{h}\left(r, \boldsymbol{\phi}, \mathbf{x}^{[N]}\right) = \mathbf{0}$ defining the direction of the ray in the output space is parameterized with hyper-spherical coordinates, that are the magnitude r and the angles $\boldsymbol{\phi}$ [108]

$$\mathbf{h}\left(r, \boldsymbol{\phi}, \mathbf{x}^{[N]}\right) := \begin{bmatrix} \mathbf{c}_0^T\mathbf{x}^{[N]} - r\cos(\phi_0) \\ \mathbf{c}_1^T\mathbf{x}^{[N]} - r\sin(\phi_0)\cos(\phi_1) \\ \vdots \\ \mathbf{c}_{n_y-2}^T\mathbf{x}^{[N]} - r\sin(\phi_0)\ldots\sin(\phi_{n_y-3})\cos(\phi_{n_y-2}) \\ \mathbf{c}_{n_y-1}^T\mathbf{x}^{[N]} - r\sin(\phi_0)\ldots\sin(\phi_{n_y-3})\sin(\phi_{n_y-2}) \end{bmatrix} = \mathbf{0}, \tag{8.192}$$

with $0 \leq \phi_{0,\ldots,ny-3} \leq \pi$ and $0 \leq \phi_{ny-2} < 2\pi$. Note that this equality constraint is the only quantity in problem (8.188) depending on $\boldsymbol{\phi}$. As such, based on the relation for the post-optimal sensitivities of the cost function from Eq. (8.139) the total derivative of the negative cost function for this LP problem is expressed as

$$\frac{d\hat{r}}{d\boldsymbol{\phi}} = -\hat{\boldsymbol{\lambda}}_\phi^T \frac{\partial \mathbf{h}}{\partial \boldsymbol{\phi}}, \tag{8.193}$$

where $\hat{\boldsymbol{\lambda}}_\phi$ represents the Lagrange multiplier vector corresponding to the equality constraint (8.192) and $\hat{r}(\boldsymbol{\phi})$ is the negative, optimal cost function value of problem (8.188). Consider the parametric surface:

$$\hat{\mathbf{y}}\left(\boldsymbol{\phi}, \hat{r}(\boldsymbol{\phi})\right) := \begin{bmatrix} \hat{y}_0(\boldsymbol{\phi}) \\ \hat{y}_1(\boldsymbol{\phi}) \\ \vdots \\ \hat{y}_{n_y-2}(\boldsymbol{\phi}) \\ \hat{y}_{n_y-1}(\boldsymbol{\phi}) \end{bmatrix} = \begin{bmatrix} \hat{r}(\boldsymbol{\phi})\cos(\phi_0) \\ \hat{r}(\boldsymbol{\phi})\sin(\phi_0)\cos(\phi_1) \\ \vdots \\ \hat{r}(\boldsymbol{\phi})\sin(\phi_0)\ldots\sin(\phi_{n_y-3})\cos(\phi_{n_y-2}) \\ \hat{r}(\boldsymbol{\phi})\sin(\phi_0)\ldots\sin(\phi_{n_y-3})\sin(\phi_{n_y-2}) \end{bmatrix} \tag{8.194}$$

The tangent plane of this surface at $\boldsymbol{\phi} = \boldsymbol{\phi}^{(j)}$ can be represented by the column space of the matrix $\hat{\mathbf{T}}(\boldsymbol{\phi})$

$$\hat{\mathbf{T}}(\boldsymbol{\phi}) := \frac{\partial \hat{\mathbf{y}}}{\partial \hat{r}}\frac{d\hat{r}}{d\boldsymbol{\phi}} + \frac{\partial \hat{\mathbf{y}}}{\partial \boldsymbol{\phi}}. \tag{8.195}$$

Let $\hat{\mathbf{n}}(\boldsymbol{\phi})$ denote a non-zero unit vector in the null-space of the column space of $\hat{\mathbf{T}}(\boldsymbol{\phi})$ pointing outward and the distance from the origin to the tangent hyper plane $d(\boldsymbol{\phi}) \geq 0$ be defined as $d(\boldsymbol{\phi}) := \hat{\mathbf{n}}^T(\boldsymbol{\phi})\hat{\mathbf{y}}(\boldsymbol{\phi})$. One of the inequalities corresponding to a particular $\boldsymbol{\phi} = \boldsymbol{\phi}^{(j)}$ defining the outer approximation can be expressed as

$$\mathbf{a}_j^O \mathbf{y} \leq b_j^O, \tag{8.196}$$

with

$$\mathbf{a}_j^O := \hat{\mathbf{n}}^T \left(\boldsymbol{\phi}^{(j)}\right), \qquad b_j^O := d\left(\boldsymbol{\phi}^{(j)}\right). \tag{8.197}$$

Consider a set of points $\mathcal{V}^I$

$$\mathcal{V}^I := \left\{\hat{\mathbf{y}}\left(\boldsymbol{\phi}_0\right), \ldots, \hat{\mathbf{y}}\left(\boldsymbol{\phi}_{n_I-1}\right)\right\}, \tag{8.198}$$

obtained from the solutions of LPMRAY($\boldsymbol{\phi}_0, \mathcal{U}, \mathcal{X}$),. . ., LPMRAY($\boldsymbol{\phi}_{n_I-1}, \mathcal{U}, \mathcal{X}$). The convex hull of these points which can be described by the inequalities

$$\mathbf{A}^I \mathbf{y} \leq \mathbf{b}^I, \tag{8.199}$$

represents an inner approximation of the true set. Now, consider the collection of all inequalities from Eq. (8.196)

$$\mathbf{A}^O \mathbf{y} \leq \mathbf{b}^O, \tag{8.200}$$

with

$$\mathbf{A}^O := \begin{bmatrix} \mathbf{a}_0^O \\ \vdots \\ \mathbf{a}_{n_I-1}^O \end{bmatrix}, \qquad \mathbf{b}^O := \begin{bmatrix} b_0^O \\ \vdots \\ b_{n_I-1}^O \end{bmatrix}. \tag{8.201}$$

For an appropriate selection of the angles $\boldsymbol{\phi}_0, \ldots, \boldsymbol{\phi}_{n_I-1}$ these inequalities defined by all tangent planes form a closed and convex polytope representing the outer approximation. Note that the boundary of the true set is guaranteed to lie in between the inner approximation (8.199) and outer approximation (8.200). Using a suitable refinement strategy the approximation of the true set up to a tolerance ϵ_0 is achieved as described in Alg. 1.

Algorithm 1 Algorithm for Output Reachable Set Approximation with Tolerance ϵ_0

procedure APPROXIMATEOUTPUTREACHABLESET($\epsilon_0, \mathcal{U}, \mathcal{X}, \mathbf{c}_0, \ldots, \mathbf{c}_{n_y-1}$)
  $\Delta\hat{\mathbf{y}} \leftarrow \mathbf{0}$
  for $j \in \{0, \ldots, n_y - 1\}$ **do**
    $\hat{\mathbf{x}}_{max}^{[N]} \leftarrow \arg_{\mathbf{x}^{[N]}} \mathrm{LPM}\left(-\mathbf{c}_j, \mathcal{U}, \mathcal{X}\right)$ ▷ Maximize output y_j
    $\hat{\mathbf{x}}_{min}^{[N]} \leftarrow \arg_{\mathbf{x}^{[N]}} \mathrm{LPM}\left(+\mathbf{c}_j, \mathcal{U}, \mathcal{X}\right)$ ▷ Minimize output y_j
    for $k \in \{0, \ldots, n_y - 1\}$ **do**
      $\hat{y}_{k,max} \leftarrow \mathbf{c}_k^T \hat{\mathbf{x}}_{max}^{[N]}$
      $\hat{y}_{k,min} \leftarrow \mathbf{c}_k^T \hat{\mathbf{x}}_{min}^{[N]}$
    end for
    $\mathbf{A}^O \leftarrow \left[\left(\mathbf{A}^O\right)^T, \mathbf{e}_j, -\mathbf{e}_j\right]^T$
    $\mathbf{b}^O \leftarrow \left[\left(\mathbf{b}^O\right)^T, \hat{y}_{j,max}, -\hat{y}_{j,min}\right]^T$
    $\mathcal{V}^I \leftarrow \mathcal{V}^I \cup \left\{\left[\hat{y}_{0,max}, \ldots, \hat{y}_{n_y-1,max}\right]^T, \left[\hat{y}_{0,min}, \ldots, \hat{y}_{n_y-1,min}\right]^T\right\}$
    $\Delta\hat{\mathbf{y}} \leftarrow \Delta\hat{\mathbf{y}} + \mathbf{e}_j\left(\hat{y}_{j,max} - \hat{y}_{j,min}\right)$
  end for
  repeat
    $\mathcal{V}^R \leftarrow \emptyset$ ▷ Initialize refinement set
    $\mathcal{V}^O \leftarrow \mathrm{I2V}\left(\mathbf{A}^O, \mathbf{b}^O\right)$ ▷ Vertices from inequalities
    $\left[\mathbf{A}^I, \mathbf{b}^I\right] \leftarrow \mathrm{V2I}\left(\mathcal{V}^I\right)$ ▷ Inequalities from vertices
    $e_\infty = \epsilon_0$
    for all $\mathbf{y}^O \in \mathcal{V}^O$ **do**
      $e^C \leftarrow \mathrm{LLS}\left(\mathbf{y}^O, \Delta\hat{\mathbf{y}}, \mathbf{A}^I, \mathbf{b}^I\right)$ ▷ Compute scaled distance
      if $e \geq \epsilon_0$ **then** ▷ Select refinement points
        $\mathcal{V}^R \leftarrow \mathcal{V}^R \cup \mathbf{y}^O$
        $e_\infty \leftarrow \max\{e_\infty, e^C\}$ ▷ Determine maximum error
      end if
    end for
    for all $\mathbf{y}^R \in \mathcal{V}^R$ **do**
      $\boldsymbol{\phi} \leftarrow \mathrm{C2S}\left(\mathbf{y}^R\right)$ ▷ Spherical angles from Cartesian vector
      $\hat{r}(\boldsymbol{\phi}) \leftarrow \mathrm{LPMRAY}\left(\boldsymbol{\phi}, \mathcal{U}, \mathcal{X}\right)$ ▷ Maximize in direction defined by $\boldsymbol{\phi}$
      $\hat{\mathbf{y}}(\hat{r}(\boldsymbol{\phi}), \boldsymbol{\phi}) \leftarrow \mathrm{S2C}\left(\hat{r}(\boldsymbol{\phi}), \boldsymbol{\phi}\right)$ ▷ Cartesian vector from spherical coordinates
      $\hat{\mathbf{T}} = \frac{\partial\hat{\mathbf{y}}}{\partial\hat{r}}\frac{d\hat{r}}{d\boldsymbol{\phi}} + \frac{\partial\hat{\mathbf{y}}}{\partial\boldsymbol{\phi}}$ ▷ Compute tangent space matrix
      $\hat{\mathbf{n}} \leftarrow \mathrm{null}\left(\hat{\mathbf{T}}\right)$ ▷ Compute normalized null-space vector
      $\hat{d} \leftarrow \hat{\mathbf{n}}^T\hat{\mathbf{y}}$
      $\mathbf{A}^O \leftarrow \left[\left(\mathbf{A}^O\right)^T, \hat{\mathbf{n}}\right]^T$
      $\mathbf{b}^O \leftarrow \left[\left(\mathbf{b}^O\right)^T, \hat{d}\right]^T$
      $\mathcal{V}^I \leftarrow \mathcal{V}^I \cup \hat{\mathbf{y}}$
    end for
  until $e_\infty < \epsilon_0$
  return $\left[\mathbf{A}^O, \mathbf{b}^O\right]$
end procedure

This algorithm first solves $2n_y$ LP problems which determine the maximal and minimal possible values $\hat{\mathbf{y}}_{j,max}, j = 0, \ldots, n_y - 1$ and $\hat{\mathbf{y}}_{j,min}, j = 0, \ldots, n_y - 1$ for each individual output.

These values are obtained from the solution of the following optimization problem LPM($\mathbf{c}, \mathcal{U}, \mathcal{X}$) depending on one of the output vectors $\mathbf{c}$:

$$\begin{aligned} \underset{\mathbf{x}^{[i]}, \mathbf{u}^{[i]}, i=0,\ldots,N}{\text{minimize}} \quad & \mathbf{c}^T \mathbf{x}^{[N]} \\ \text{subject to} \quad & \mathbf{0} = \mathbf{A}_d^- \mathbf{x}^{[i-1]} + \mathbf{A}_d^+ \mathbf{x}^{[i]} + \mathbf{B}_d \left(\mathbf{u}^{[i]} + \mathbf{u}^{[i-1]} \right), \ i = 1, \ldots, N, \\ & \mathbf{x}^{[0]} = \mathbf{0}, \\ & \mathbf{u}^{[i]} \in \mathcal{U}, \ i = 0, \ldots, N, \\ & \mathbf{x}^{[i]} \in \mathcal{X}, \ i = 1, \ldots, N \end{aligned} \tag{8.202}$$

Let $\hat{\mathbf{x}}_{j,max}^{[N]}, \ j = 0, \ldots, n_y - 1$ be obtained from solutions of LPM($-\mathbf{c}_j, \mathcal{U}, \mathcal{X}$), i.e. a maximization of the $j-$th output. The upper extremal points of the true set are then:

$$\hat{\mathbf{y}}_{j,max} := \left[\mathbf{c}_0 \hat{\mathbf{x}}_{j,max}^{[N]}, \ldots, \mathbf{c}_{n_y-1} \hat{\mathbf{x}}_{j,max}^{[N]} \right]^T, \ j = 0, \ldots, n_y - 1 \tag{8.203}$$

Similarly, let $\hat{\mathbf{x}}_{j,min}^{[N]}, \ j = 0, \ldots, n_y - 1$ be obtained from the solutions of LPM($\mathbf{c}_j, \mathcal{U}, \mathcal{X}$), i.e. a minimization of the $j-$th output. The points

$$\hat{\mathbf{y}}_{j.min} := \left[\mathbf{c}_0 \hat{\mathbf{x}}_{j,min}^{[N]}, \ldots, \mathbf{c}_{n_y-1} \hat{\mathbf{x}}_{j,min}^{[N]} \right]^T, \ j = 0, \ldots, n_y - 1, \tag{8.204}$$

then represent lower extremal points of the set. From these extremal points the initial set for the inner approximation is thus defined as:

$$\mathcal{V}^I := \left\{ \hat{\mathbf{y}}_{0,max}, \hat{\mathbf{y}}_{0,min}, \ldots, \hat{\mathbf{y}}_{n_y-1,max}, \hat{\mathbf{y}}_{n_y-1,min} \right\} \tag{8.205}$$

On the one side, the corresponding normal vector for each point $\hat{\mathbf{y}}_{j,max}$ is chosen as the basis vector $\mathbf{e}_j$, which is a zero vector with a single entry 1 at the index j. On the other side, the normal vector for each point $\hat{\mathbf{y}}_{j,min}$ is chosen as the negative basis vector $-\mathbf{e}_j$. Note that this choice has the property that the outer approximation is guaranteed to be closed from the beginning as the inequalities

$$\begin{bmatrix} \mathbf{e}_0^T \\ -\mathbf{e}_0^T \\ \vdots \\ \mathbf{e}_{n_y-1}^T \\ -\mathbf{e}_{n_y-1}^T \end{bmatrix} \leq \begin{bmatrix} \mathbf{e}_0^T \hat{\mathbf{y}}_{0,max} \\ -\mathbf{e}_0^T \hat{\mathbf{y}}_{0,min} \\ \vdots \\ \mathbf{e}_{n_y-1}^T \hat{\mathbf{y}}_{n_y-1,max} \\ -\mathbf{e}_{n_y-1}^T \hat{\mathbf{y}}_{n_y-1,min} \end{bmatrix}, \tag{8.206}$$

defining the initial outer approximation (cf. Eq. (8.200)) represent a hyper-rectangle (see Fig. 8.40). Furthermore, scaling factors which are collected in the vector $\Delta\hat{\mathbf{y}} \in \mathbb{R}^{n_y}$

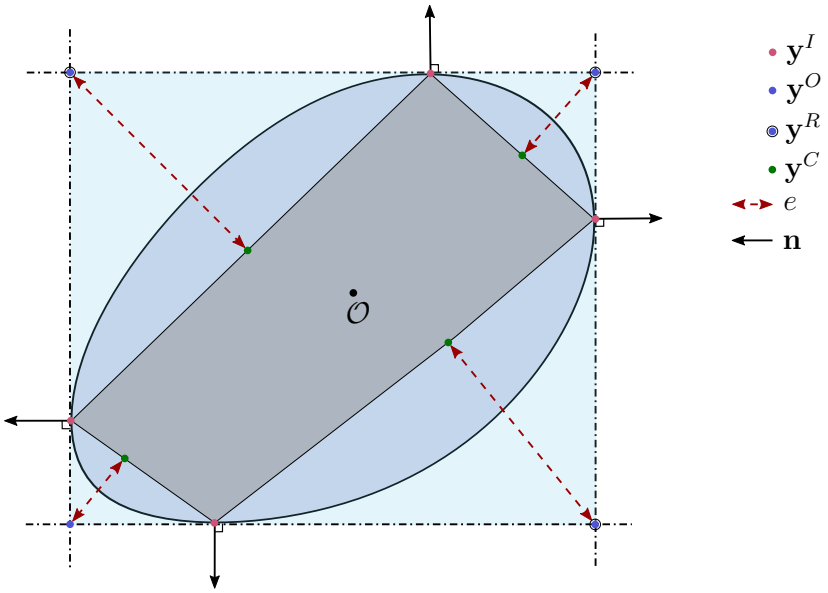


Figure 8.40: Illustration of the initial set approximations.

are determined from the differences of the maximum and minimum values

$$\Delta\hat{\mathbf{y}} := \left[\mathbf{e}_0^T\left(\hat{\mathbf{y}}_{0,max} - \hat{\mathbf{y}}_{0,min}\right), \ldots, \mathbf{e}_{n_y-1}^T\left(\hat{\mathbf{y}}_{n_y-1,max} - \hat{\mathbf{y}}_{n_y-1,min}\right)\right]^T \tag{8.207}$$

These scaling factors are used to quantify the scaled distance of the points from the outer approximation to the inner approximation. The following steps are repeated until the maximum of all scaled distances of the vertices in the outer approximation from the inner approximation is less than ϵ_0. In the beginning of each refinement cycle these scaled distances are determined as follows: First, the vertices of the outer approximation in the current iteration are computed from the inequalities (8.200) and collected in the set $\mathcal{V}^O$:

$$\mathcal{V}^O := \left\{\hat{\mathbf{y}}_j^O : j = 0, \ldots, n_O - 1\right\} \tag{8.208}$$

Note that in the first iteration these points are simply the vertices of a hyper-rectangle. Next, the matrix $\mathbf{A}^I$ and the vector $\mathbf{b}^I$ of the inequalities (8.199) defining the convex polytope of the inner approximation are computed using the vertices in $\mathcal{V}^I$. The minimum distances between all points in $\mathcal{V}^O$ from the convex hull of the inner approximation can be quantified by computing the closest point $\hat{\mathbf{y}}_j^C$ for each $\hat{\mathbf{y}}_j^O \in \mathcal{V}^O$ to the boundary of the polytope $\mathbf{A}^I\mathbf{y} \leq \mathbf{b}^I$ under consideration of the scaling factors $\Delta\hat{\mathbf{y}}$. These points are obtained from the following linear least-squares problem

LLS $\left(\hat{\mathbf{y}}_j^O, \Delta\hat{\mathbf{y}}, \mathbf{A}^I, \mathbf{b}^I\right)$:

$$\begin{aligned} \underset{\mathbf{y}_j^C}{\text{minimize}} \quad & \left\| \mathbf{y}_j^C - \operatorname{diag}\left(\Delta\hat{\mathbf{y}}\right)^{-1} \hat{\mathbf{y}}_j^O \right\|_2 \\ \text{subject to} \quad & \mathbf{A}^I \operatorname{diag}\left(\Delta\hat{\mathbf{y}}\right) \mathbf{y}_j^C \leq \mathbf{b}^I \end{aligned} \tag{8.209}$$

The optimal cost function of this problem yields the minimum scaled distance e_j^C for each point $\hat{\mathbf{y}}_j^O \in \mathcal{V}^O$ to the convex boundary of the inner approximation. The maximum scaled distance is then determined as $e_\infty = \max\{e_0^C, \ldots, e_{n_O-1}^C\}$. Subsequently, all points $\hat{\mathbf{y}}_j^O$ satisfying $e_j^C > \epsilon_0$ are collected in the refinement set $\mathcal{V}^R$

$$\mathcal{V}^R := \left\{ \mathbf{y}_j^R : j = 0, \ldots, n_R - 1, \mathbf{y}_j^R \in \mathcal{V}^O, e_j^C > \epsilon_0 \right\}. \tag{8.210}$$

For each point $\mathbf{y}_j^R$ the vector of angles $\boldsymbol{\phi}^{(j)}$ is computed using a transformation to spherical coordinates and the optimal control problem LPMRAY($\boldsymbol{\phi}_j, \mathcal{U}, \mathcal{X}$) is solved, i.e. the distance to the origin along the ray defined by the angles in $\boldsymbol{\phi}_j$ is maximized. At the optimal solution of each problem the output vector $\hat{\mathbf{y}}(\boldsymbol{\phi}_j)$ is added to the set of points $\mathcal{V}^I$ for the inner approximation and the inequality (8.196) is added to the set of inequalities for the outer approximation. These steps represent the major iterations of Alg. 1 and are repeated until convergence. In this algorithm the transformation S2C($r, \boldsymbol{\phi}$) is defined as in Eq. (8.194) and C2S($\mathbf{v}$) computes the hyper-spherical angles $\boldsymbol{\phi}$ from a general Cartesian vector $\mathbf{v} \in \mathbb{R}^{n_v}$. Moreover, the function I2V $[\mathbf{A}, \mathbf{b}]$ determines the vertices of a convex polytope defined by $\mathbf{A}\mathbf{v} \leq \mathbf{b}$. Similarly, the function V2I $[\mathcal{V}]$ computes inequalities $\mathbf{A}\mathbf{v} \leq \mathbf{b}$ of the convex polytope defined by vertices $\mathcal{V}$. For both functions the implementation from Ref. [109] is used.

The application of the algorithm is illustrated for the nominal models in the longitudinal and lateral plane (cf. Sec. 8.2.1) with $t_f = 8\ s$ and $h = 0.025\ s$. For all set approximations a tolerance of $\epsilon_0 = 0.01$ is used. First, two-dimensional examples for the Cat I and Cat II type models are presented. The resulting outer approximations for the longitudinal model are depicted in Fig. 8.41 and Fig. 8.42. Moreover, a comparison of these two sets is shown in Fig. 8.43. In this figure the set obtained for the Cat II model is colored in dark blue whereas the set for the Cat I model is colored in light blue. Note that the set for the Cat II model is fully contained within the set for the Cat I model as the addition of state constraints limits the attainable values of the criteria. The difference between the two sets is not significantly large. This may be explained based on the control histories in Figs. 8.21-8.24 for the maximization of the individual criteria. Herein, the solutions for the longitudinal Cat II model merely exhibits short constrained arcs for the elevator rate state (the position limit is not reached) which appear to have only a small effect on the optimal cost function values.

The outer approximations for the Cat I and Cat II models in the lateral plane are

presented in Fig. 8.44 and Fig. 8.45. Moreover, the comparison of the combined sets is shown in Fig. 8.46. It is apparent that the state constraints considerably constrain the attainable set in the output space. Note that contrary to the Cat II results for the longitudinal model the solutions for the maximization of the individual criteria in the lateral plane under the Cat II models exhibits both rate and position constrained arcs on most parts of the time interval (see Fig. 8.27 and Fig. 8.28).

Finally, the sequence of construction for the inner and outer approximations in each major iteration of the algorithm are depicted in Figs. 8.47-8.50 for all cases. The vertices of the outer approximation (blue dots) which are selected to define the rays in the next refinement cycle are circled in black. Note that a fairly low number of LPs (n_{LP}) is required to obtain over-approximations at a reasonable accuracy.

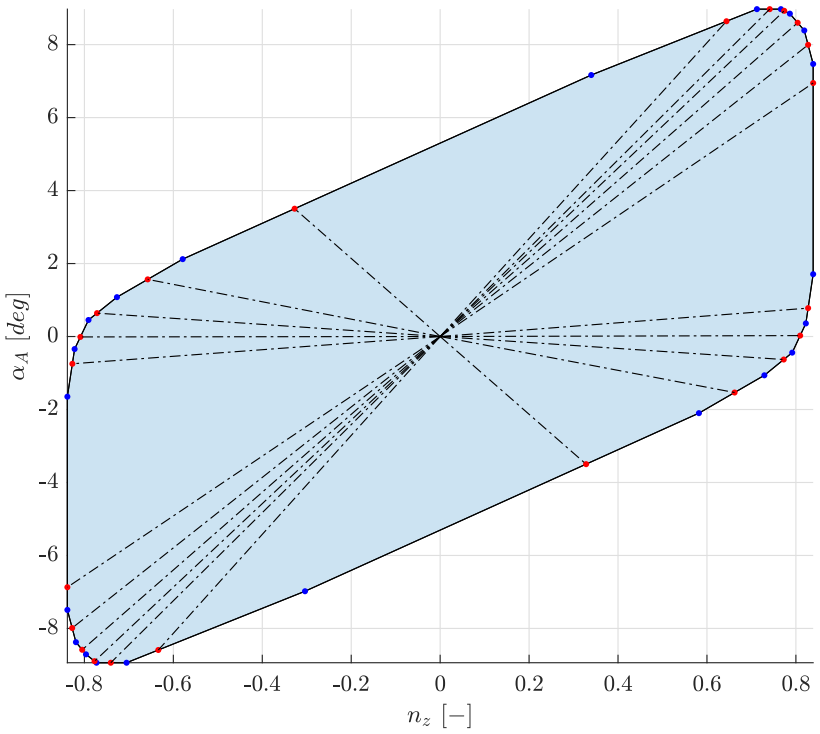


Figure 8.41: Final iteration of the reachable set approximation (Cat I) for $n_z(t_f)$ and $\alpha_A(t_f)$ ($e = 0.00682 < \epsilon_0, n_{LP} = 22$)

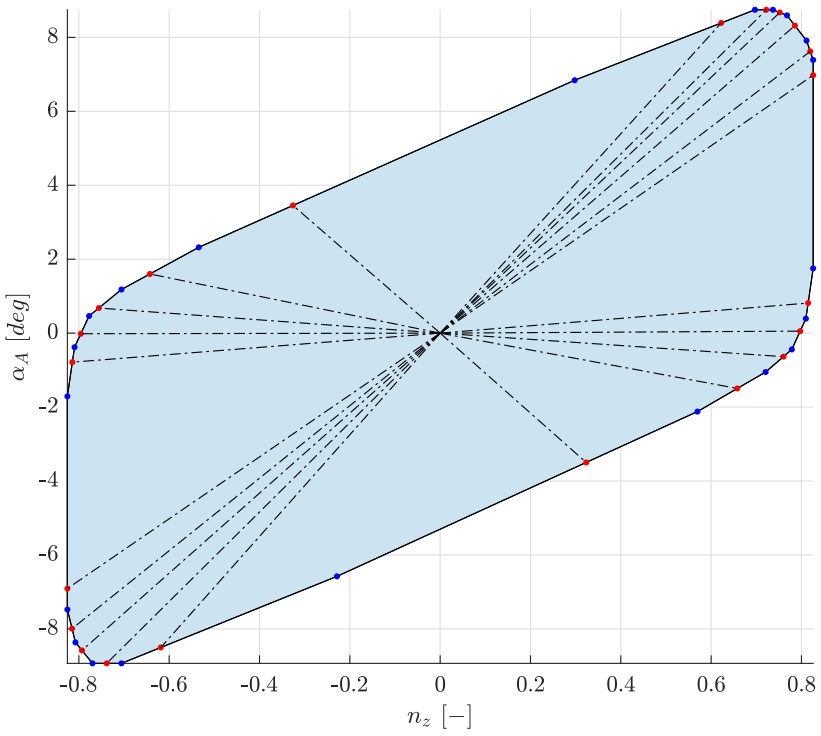


Figure 8.42: Final iteration of the reachable set approximation (Cat II) for $n_z(t_f)$ and $\alpha_A(t_f)$ ($e = 0.00978 < \epsilon_0, n_{LP} = 21$)

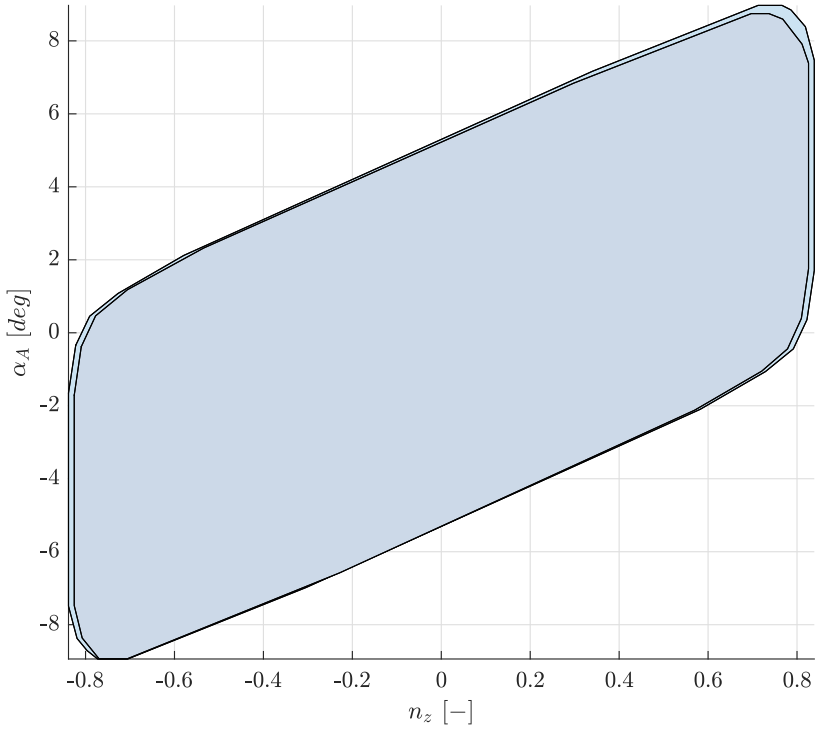


Figure 8.43: Comparsion between the reachable set approximations (Cat I light blue, Cat II dark blue) for $n_z(t_f)$ and $\alpha_A(t_f)$.

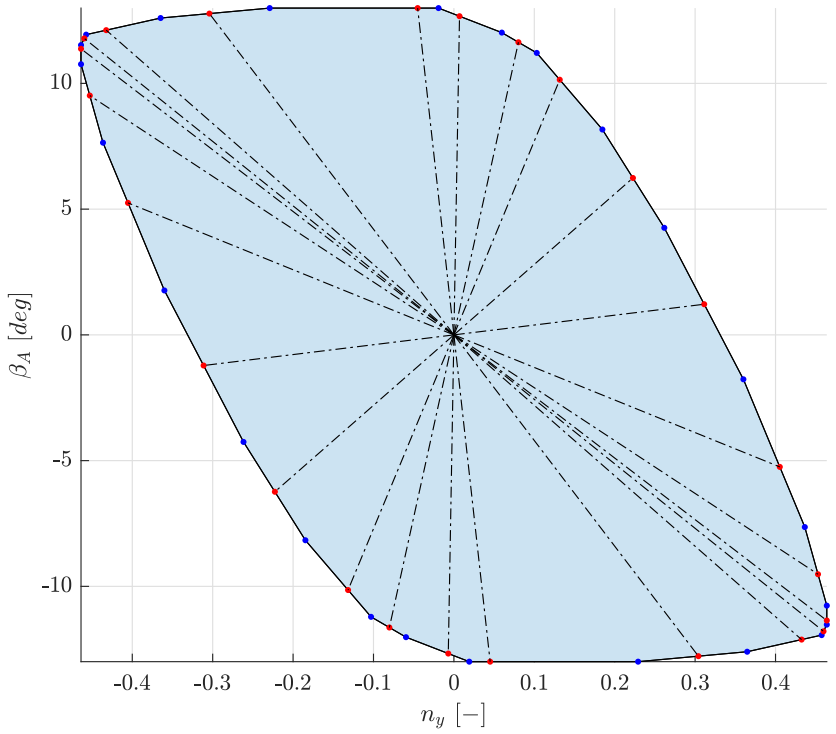


Figure 8.44: Final iteration of the reachable set approximation (Cat I) for $n_y(t_f)$ and $\beta_A(t_f)$ ($e = 0.00636 < \epsilon_0, n_{LP} = 24$)

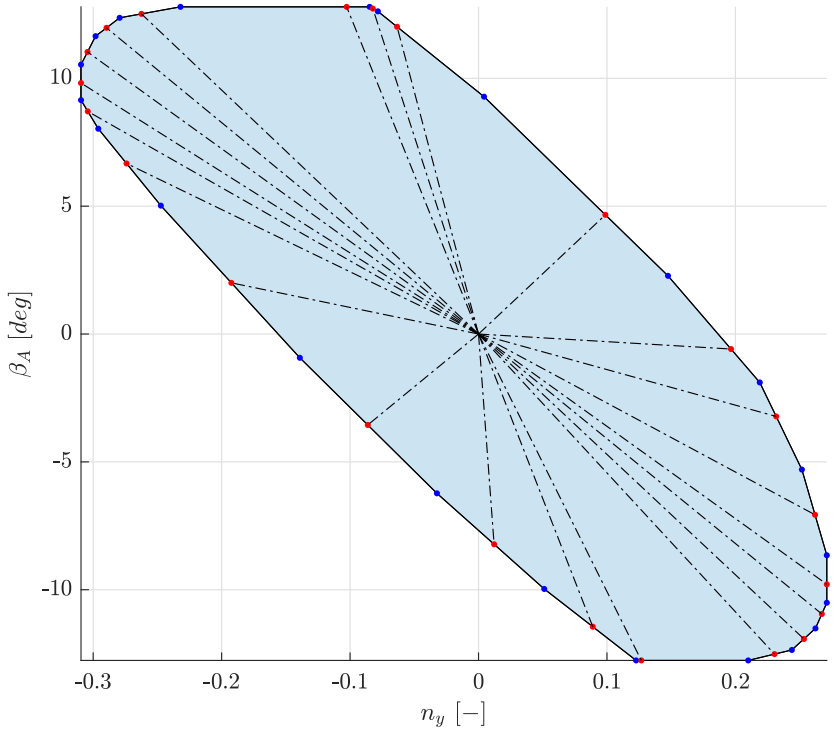


Figure 8.45: Final iteration of the reachable set approximation (Cat II) for $n_y(t_f)$ and $\beta_A(t_f)$ ($e = 0.00920 < \epsilon_0, n_{LP} = 22$)

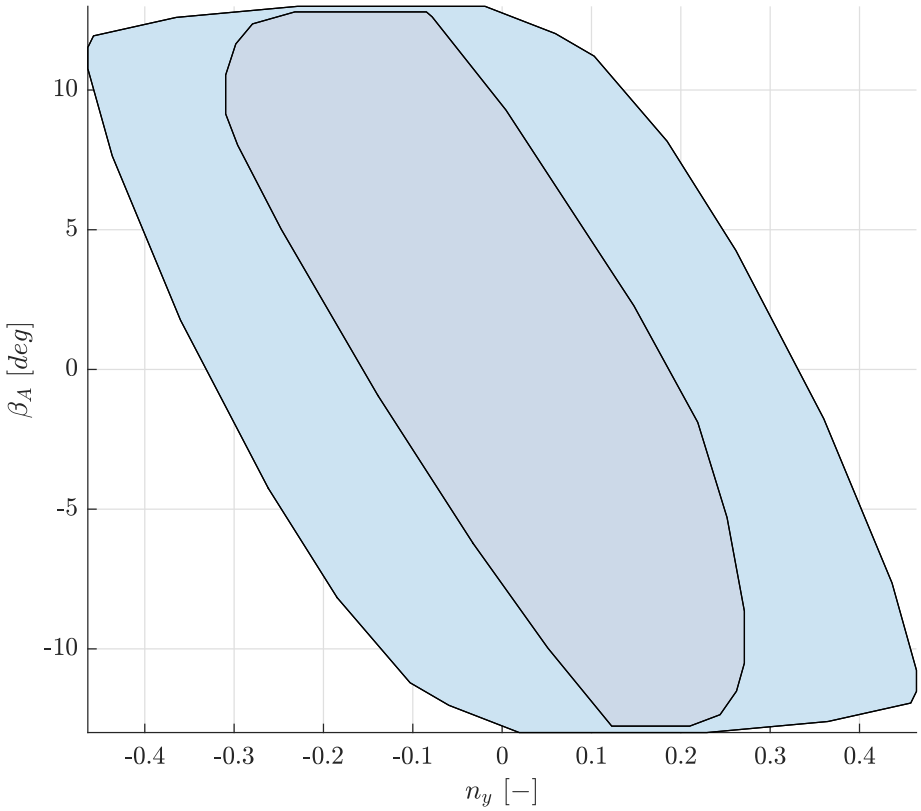
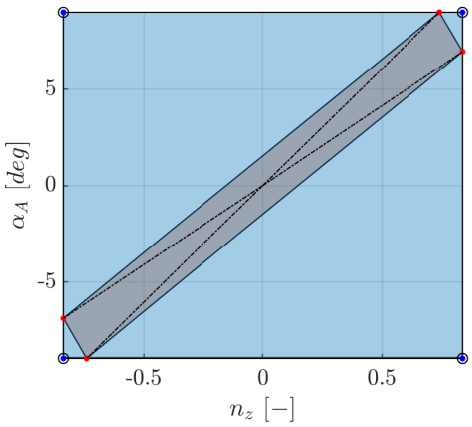
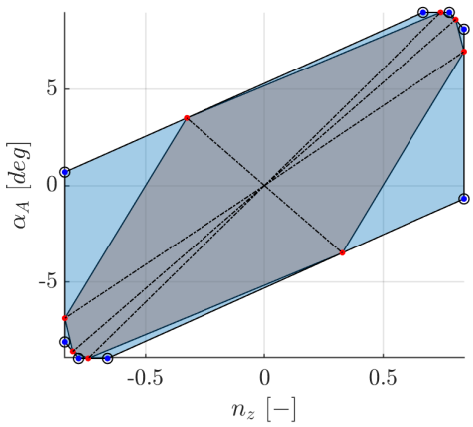


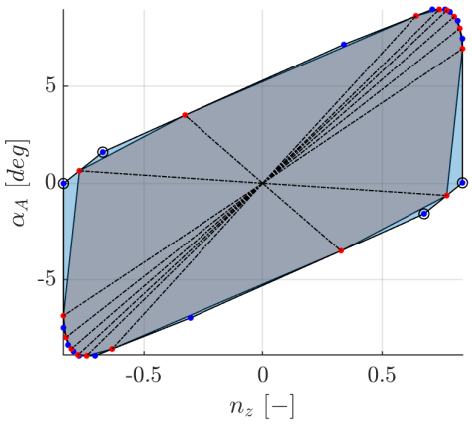
Figure 8.46: Comparsion between the reachable set approximations (Cat I light blue, Cat II dark blue) for $n_y(t_f)$ and $\beta_A(t_f)$.



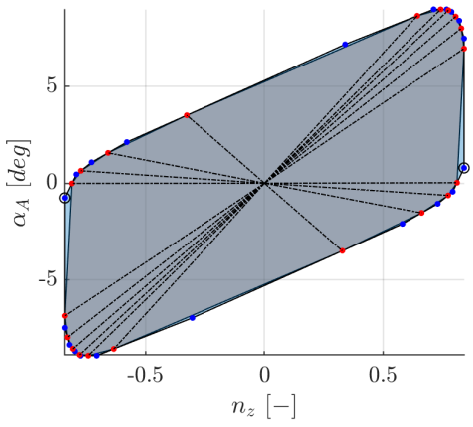
Iteration 0 ($e = 0.64695, n_{LP} = 4$)



Iteration 1 ($e = 0.19742, n_{LP} = 8$)

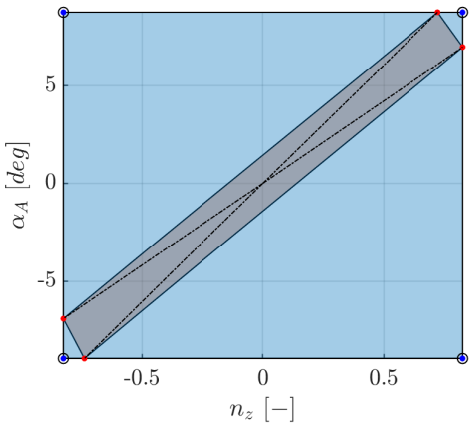


Iteration 2 ($e = 0.03619, n_{LP} = 16$)

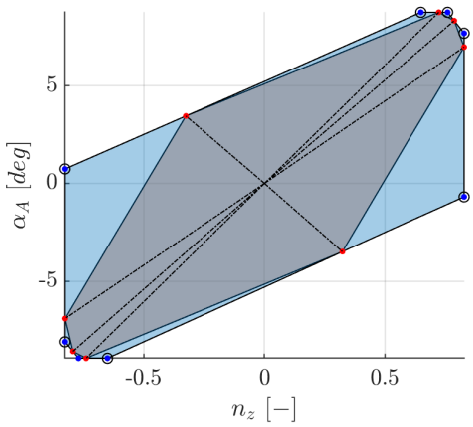


Iteration 3 ($e = 0.01555, n_{LP} = 20$)

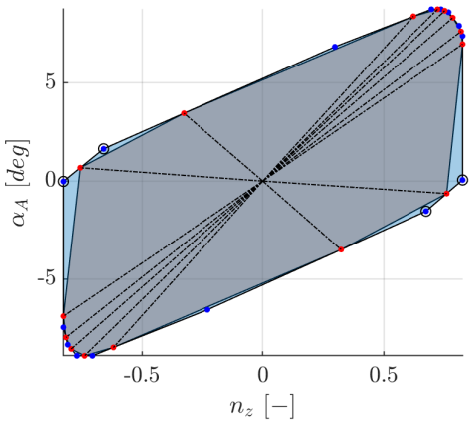
Figure 8.47: Illustration of the iterations 0,...,3 for the reachable set approximation (Cat I) in the output space for the normal load factor n_z and the aerodynamic AoA α_A.



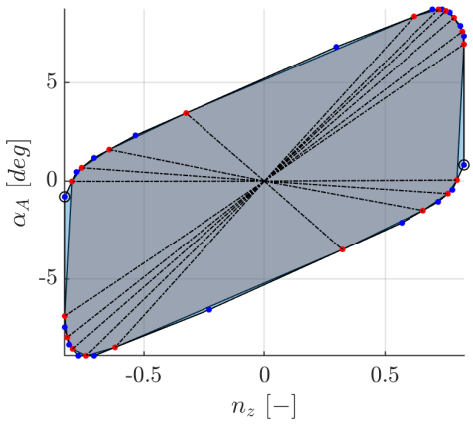
Iteration 0 ($e = 0.65370, n_{LP} = 4$)



Iteration 1 ($e = 0.19866, n_{LP} = 8$)

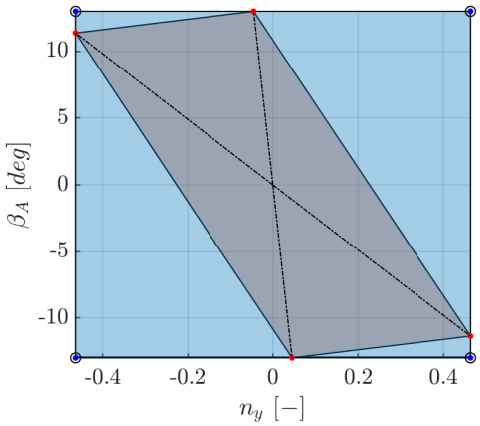


Iteration 2 ($e = 0.03810, n_{LP} = 15$)

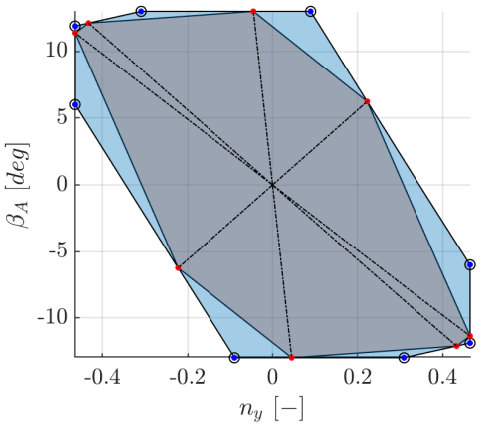


Iteration 3 ($e = 0.01596, n_{LP} = 19$)

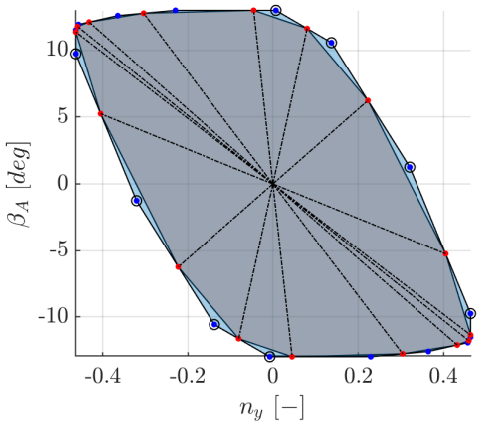
Figure 8.48: Illustration of the iterations 0,...,3 for the reachable set approximation (Cat II) in the output space for the normal load factor n_z and the aerodynamic AoA α_A.



Iteration 0 ($e = 0.47414, n_{LP} = 4$)

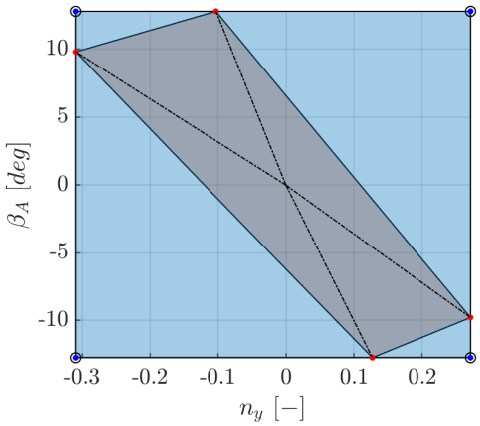


Iteration 1 ($e = 0.09734, n_{LP} = 8$)

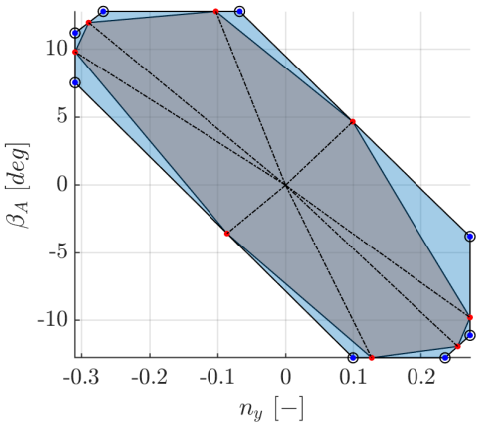


Iteration 2 ($e = 0.02521, n_{LP} = 16$)

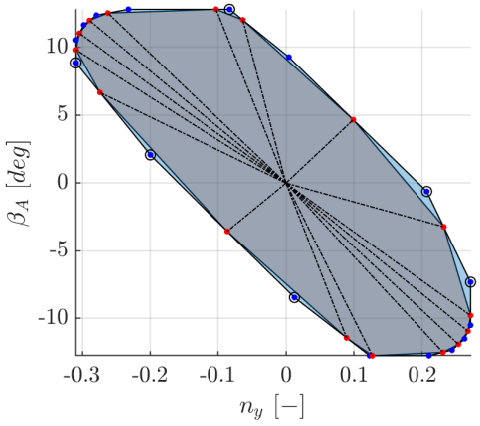
Figure 8.49: Illustration of the iterations 0,...,2 for the reachable set approximation (Cat I) in the output space for the lateral load factor n_y and the aerodynamic AoS β_A.



Iteration 0 ($e = 0.57337, n_{LP} = 4$)



Iteration 1 ($e = 0.10964, n_{LP} = 8$)



Iteration 2 ($e = 0.02615, n_{LP} = 16$)

Figure 8.50: Illustration of the iterations 0,...,2 for the reachable set approximation (Cat II) in the output space for the lateral load factor n_y and the aerodynamic AoS β_A.

To illustrate the application of the algorithm in three dimensions the aerodynamic pitch rate $q_A(t_f)$ is added to the criteria for the longitudinal model. Similarly, for the lateral model the aerodynamic yaw rate $r_A(t_f)$ is added. The resulting sets for the Cat I and Cat II models are depicted in Fig. 8.51 and Fig. 8.52 for the longitudinal as well as Fig. 8.53 and Fig. 8.54 for the lateral model. The sequence of construction for the inner and outer approximations in each major iteration are shown in Figs. 8.55-8.58. As in the two-dimensional case the resulting over-approximations for the longitudinal Cat I and Cat II models appear to be very similar (compare Fig. 8.51 and Fig. 8.52). Contrary, in the lateral plane the set in the Cat II case is considerable smaller in comparison to the Cat I case (see Fig. 8.51 and Fig. 8.52). Note further that the strategy in Alg. 1 automatically allocates a higher density of points in regions where the set requires a refinement (observe for example the refinement of the rim-like feature in Fig. 8.55). The number of LPs is obviously higher in the three-dimensional case and is expected to grow rapidly for higher dimensions. However, it is noteworthy that the algorithm is essentially applicable to arbitrary state and control dimensions. In the examples presented here the number of variables and constraints for the discretized problem are in the range of several thousands. As the solution of LP problems is a highly developed field, state of the art solvers are able to efficiently handle considerably higher numbers of variables and constraints.

Several aspects may help to alleviate the increasing computational cost for higher-dimensional criteria spaces. For example, symmetry can be exploited if the set is point-symmetric. This is for example expected in case all limits are symmetric as a minimization and maximization in the direction of a particular ray should yield opposite responses. In addition, it is clear that the algorithm allows for a massive parallelization as all LPs within one refinement cycle can be solved independently. Moreover, note that a closed outer approximation is maintained from the first iteration of the algorithm onward and that the largest improvement in the error is observed in the first iterations. If a coarse, however conservative, approximation is sufficient for a particular application the algorithm returns an outer approximation of the set including the expected accuracy after the first iteration with a maximum number of $n_{LP} = 2n_y + 2^{n_y}$ LPs to be solved.

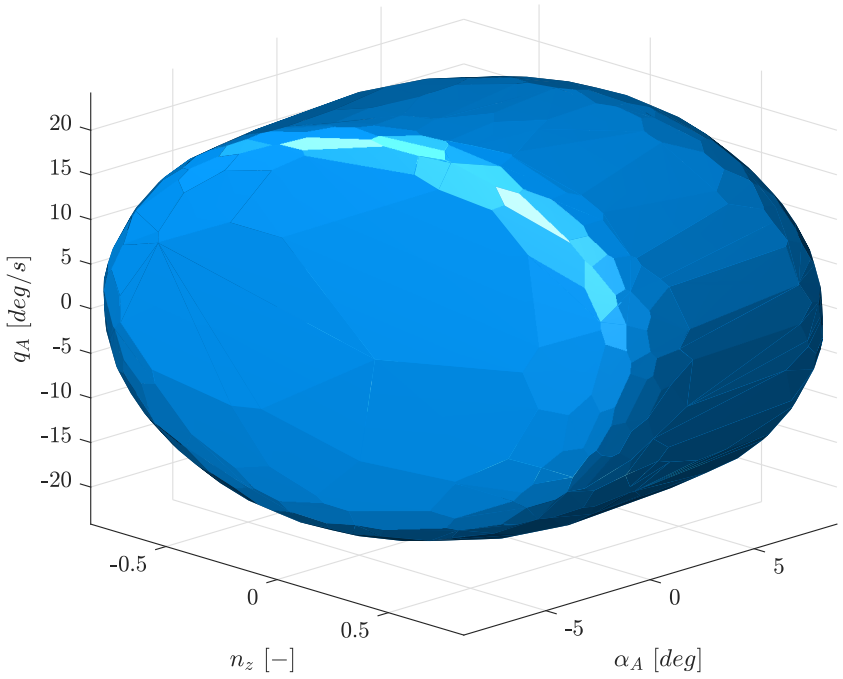


Figure 8.51: Final iteration of the reachable set approximation (Cat I) for $n_z(t_f)$, $\alpha_A(t_f)$, and $q_A(t_f)$ ($e = 0.00957 < \epsilon_0, n_{LP} = 332$)

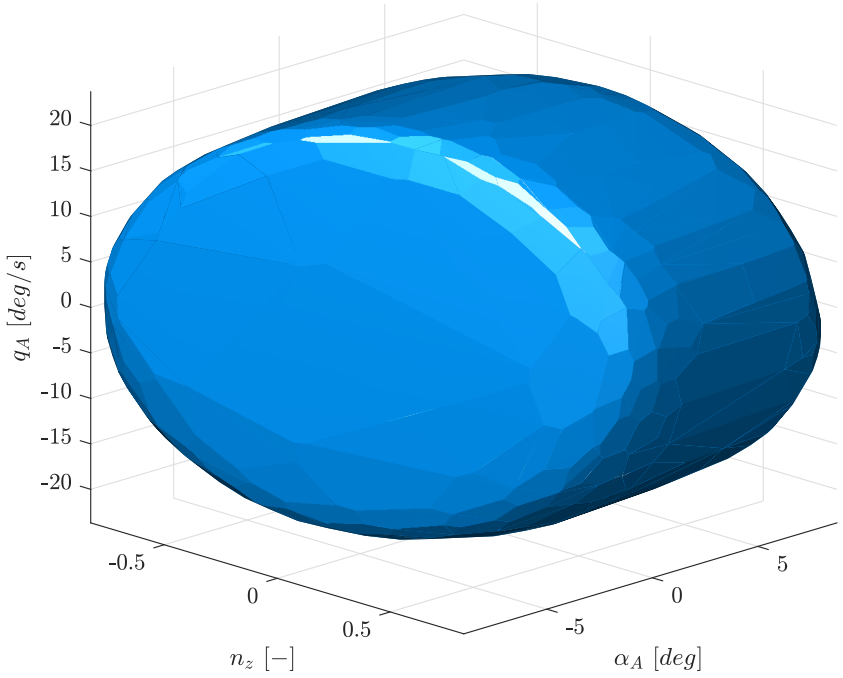


Figure 8.52: Final iteration of the reachable set approximation (Cat II) for $n_z(t_f)$, $\alpha_A(t_f)$, and $q_A(t_f)$ ($e = 0.00989 < \epsilon_0, n_{LP} = 349$)

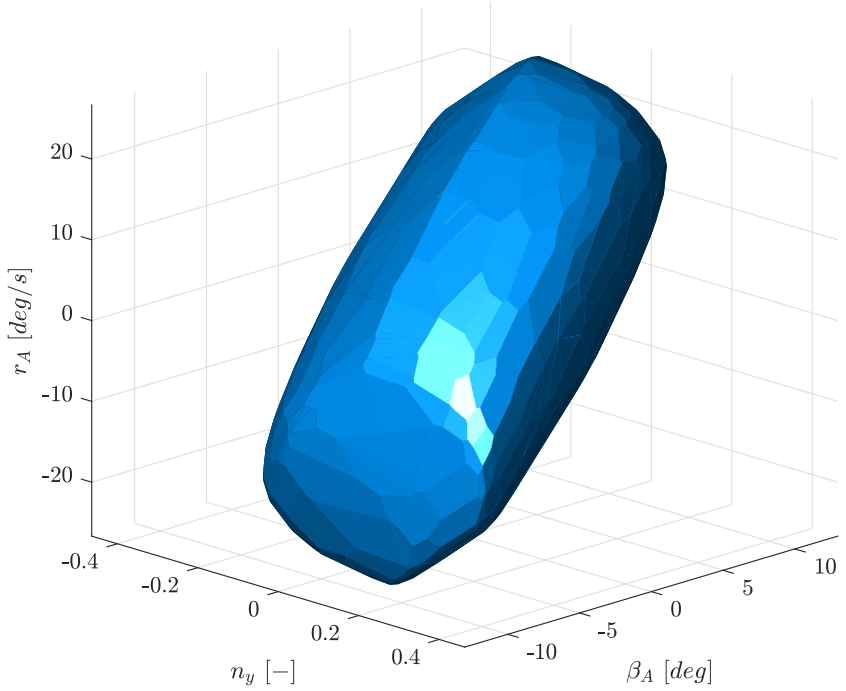


Figure 8.53: Final iteration of the reachable set approximation (Cat I) for $n_y(t_f)$, $\beta_A(t_f)$, and $r_A(t_f)$ ($e = 0.00955 < \epsilon_0, n_{LP} = 348$)

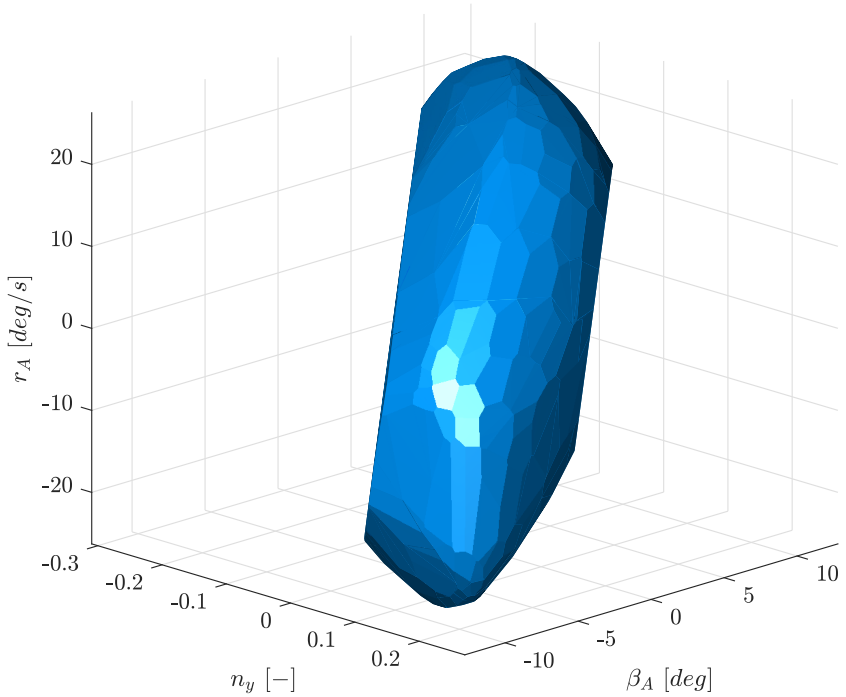
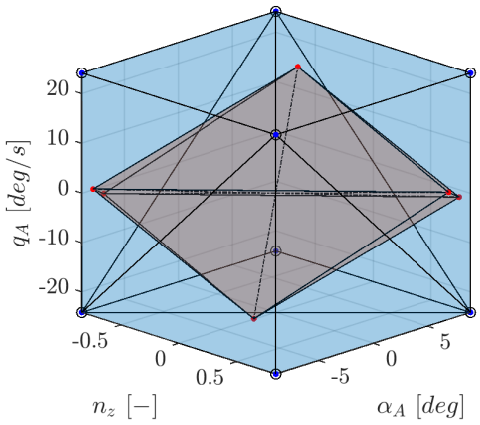
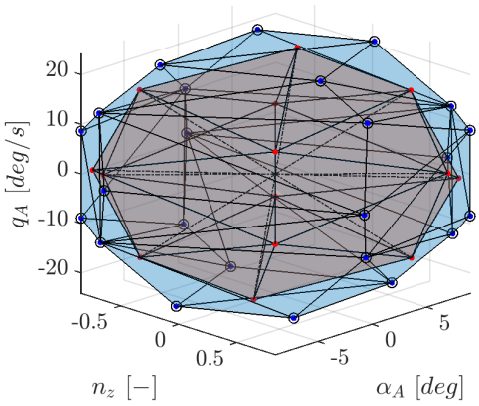


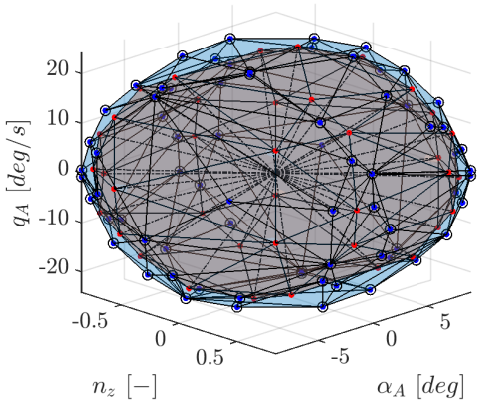
Figure 8.54: Final iteration of the reachable set approximation (Cat II) for $n_y(t_f)$, $\beta_A(t_f)$, and $r_A(t_f)$ ($e = 0.00999 < \epsilon_0, n_{LP} = 300$)



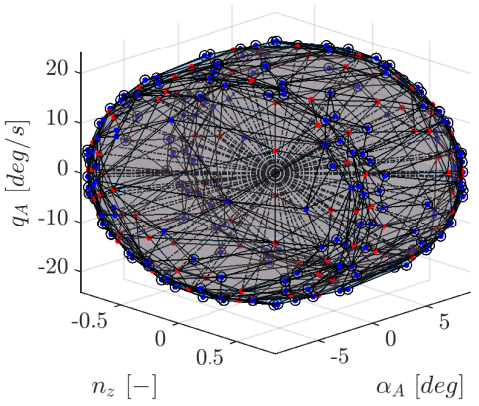
Iteration 0 ($e = 0.75201, n_{LP} = 6$)



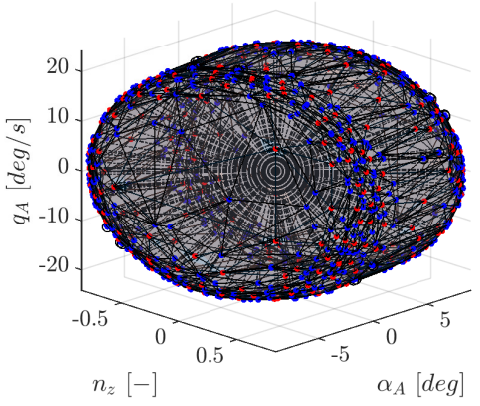
Iteration 1 ($e = 0.27299, n_{LP} = 14$)



Iteration 2 ($e = 0.09371, n_{LP} = 38$)

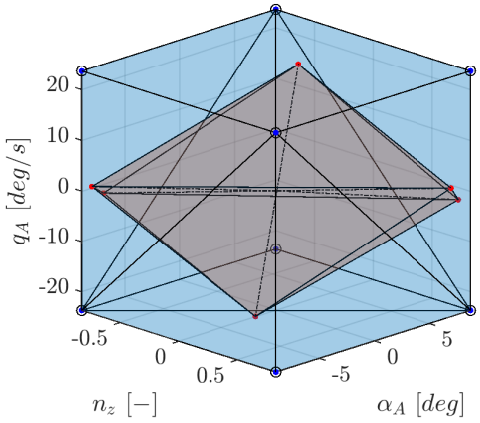


Iteration 3 ($e = 0.04801, n_{LP} = 108$)

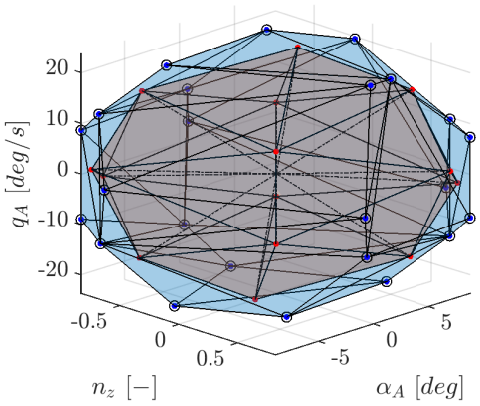


Iteration 4 ($e = 0.01476, n_{LP} = 287$)

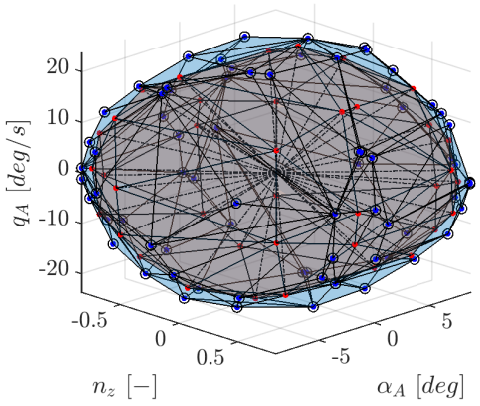
Figure 8.55: Illustration of the iterations 0...4 for the reachable set approximation (Cat I) in the output space for the normal load factor n_z, the aerodynamic AoA α_A, and the aerodynamic pitch rate q_A.



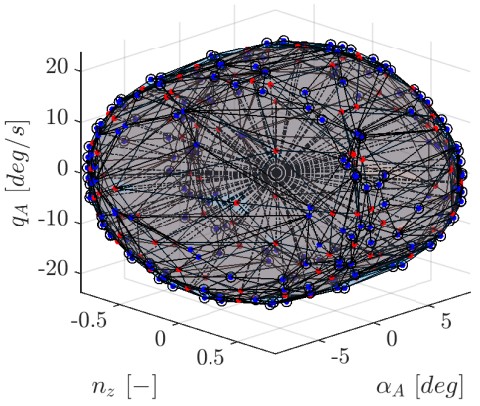
Iteration 0 ($e = 0.75763, n_{LP} = 6$)



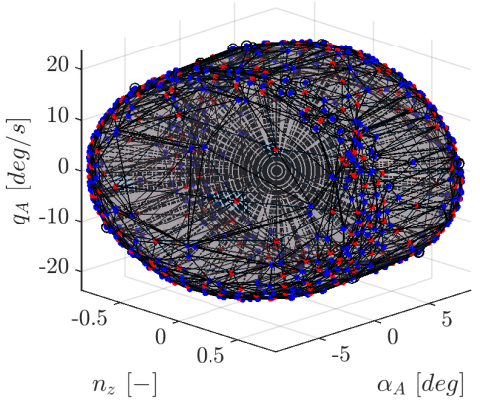
Iteration 1 ($e = 0.33521, n_{LP} = 14$)



Iteration 2 ($e = 0.12775, n_{LP} = 38$)

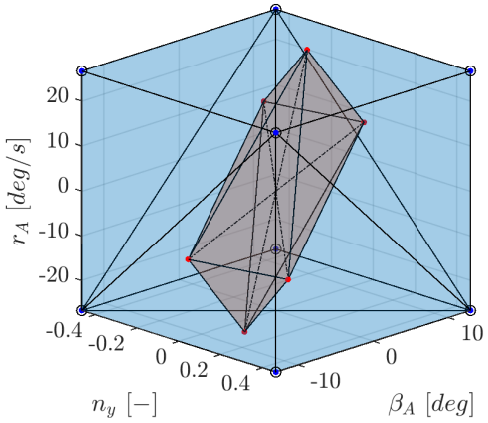


Iteration 3 ($e = 0.04765, n_{LP} = 110$)

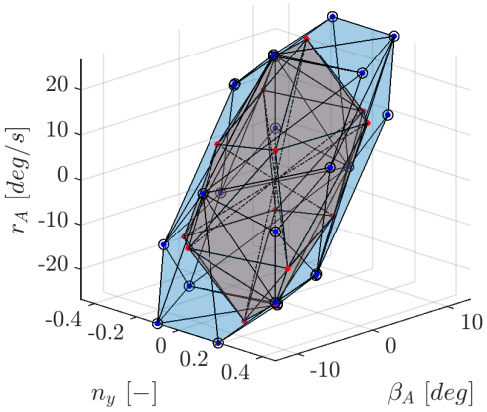


Iteration 4 ($e = 0.01677, n_{LP} = 288$)

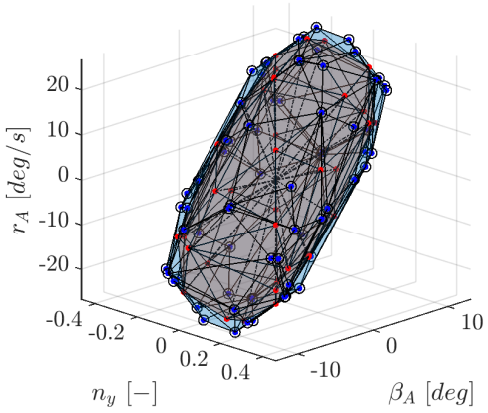
Figure 8.56: Illustration of the iterations 0,...,4 for the reachable set approximation (Cat II) in the output space for the normal load factor n_z, the aerodynamic AoA α_A, and the aerodynamic pitch rate q_A.



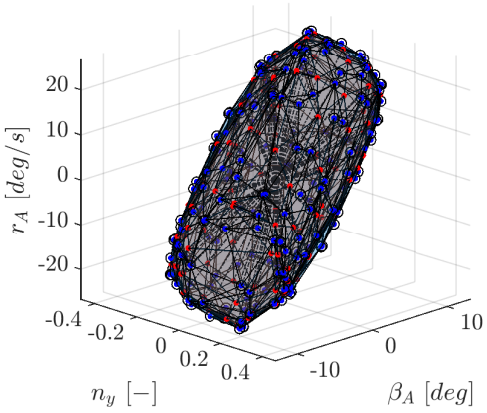
Iteration 0 ($e = 0.68941, n_{LP} = 6$)



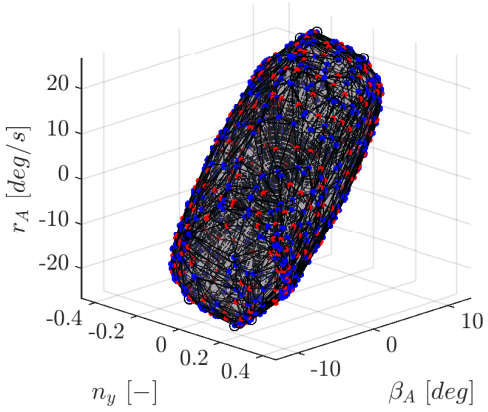
Iteration 1 ($e = 0.26707, n_{LP} = 14$)



Iteration 2 ($e = 0.09262, n_{LP} = 38$)

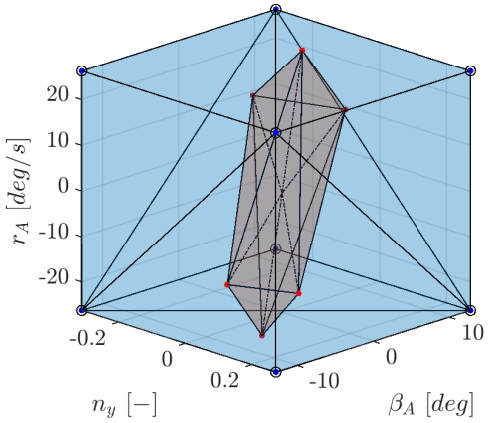


Iteration 3 ($e = 0.03188, n_{LP} = 110$)

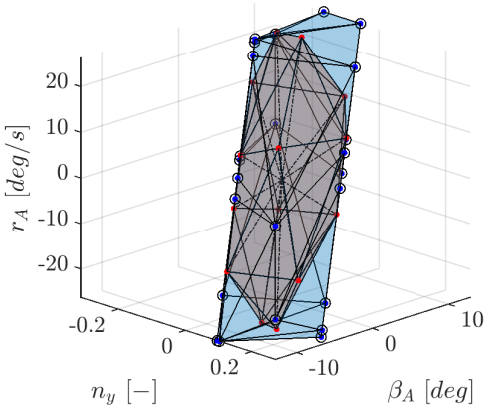


Iteration 4 ($e = 0.01511, n_{LP} = 302$)

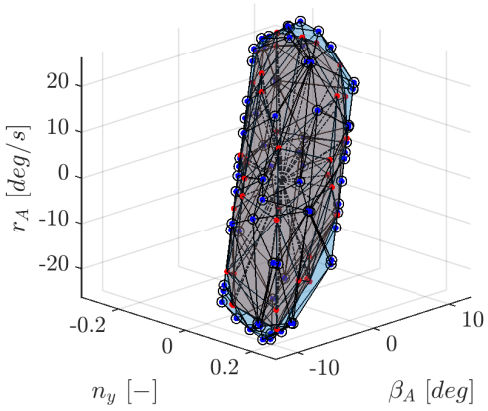
Figure 8.57: Illustration of the iterations 0,...,4 for the reachable set approximation (Cat I) in the output space for the lateral load factor n_y, the aerodynamic AoS β_A, and the aerodynamic yaw rate r_A.



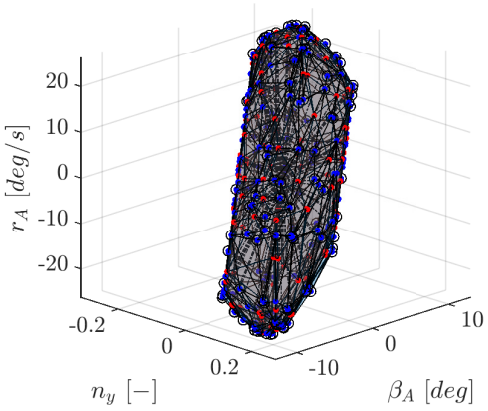
Iteration 0 ($e = 0.74304, n_{LP} = 6$)



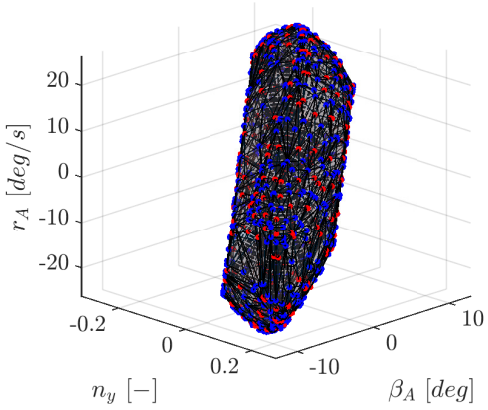
Iteration 1 ($e = 0.25435, n_{LP} = 14$)



Iteration 2 ($e = 0.10478, n_{LP} = 38$)



Iteration 3 ($e = 0.04976, n_{LP} = 110$)



Iteration 4 ($e = 0.01832, n_{LP} = 263$)

Figure 8.58: Illustration of the iterations 0,...,4 for the reachable set approximation (Cat II) in the output space for the lateral load factor n_y, the aerodynamic AoS β_A, and the aerodynamic yaw rate r_A.

Chapter 9

Optimal Control Based Clearance for Nonlinear Systems

According to the classification scheme proposed in Sec. 7.1 all cases which cannot be treated under Cat I and Cat II belong to the general Cat III modeling class. The solution of the clearance problem for these models is considerably more involved due to the potential nonlinearity of the closed-loop system. However, for practical clearance applications the development of a framework for the solution of Cat III models is of high importance due to the restricted validity of Cat I and Cat II models close to the linearization point. Due to the challenges encountered for the solution of Cat III type models this category is further refined into:

- Non-intrusive methods
- Intrusive methods
- Hybrid methods

Non-intrusive methods (cf. Refs. [3, 21]) treat the problem as a black-box function $j : \mathbb{R}^{n_p} \rightarrow \mathbb{R}$ which depends on parameters $\mathbf{p} \in \mathbb{R}^{n_p}$ and returns the value of the criterion under investigation. Here, the vector $\mathbf{p}$ collects parameters of the aircraft closed-loop model and, additionally, a parametrization of the control functions. In particular, under the non-intrusive approach the dynamic character of the underlying problem is not exposed to the optimization method. This implies, that the dynamic behavior is resolved internally through the solution of initial value problems, i.e. by performing simulations using the closed-loop model. Note that the non-intrusive approach is closely related to the single shooting concept (cf. Sec. 6.1.1) and, as such, comes with all its advantages and disadvantages. One of the major advantages is the simplicity of its application as, in most cases, the aircraft closed-loop model is provided in form of a parametrized simulation model. Hence, this model can directly be used to

evaluate the criterion based on the parameters $\mathbf{p}$. However, the disadvantages include the high degree of coupling in the problem and the potential nonlinearity of the criterion with respect to the parameters through the long simulation intervals. These two aspects render a numerical solution challenging. Additionally, most aircraft closed-loop models feature a variety of non-differentiable functions, such as table data with non-smooth interpolation methods or internal limiters which may violate the differentiability requirements for gradient-based methods. As such, typically gradient-free methods are employed if the search space is restricted to low-dimensional parameter spaces. Moreover, note that the information which is made available to the worst-case search is minimal in the sense that merely the admissible parameter space and the resulting value of the criterion is visible to the algorithm. Hence, a considerable amount of information regarding the internal characteristics of the black-box function is hidden under this approach which may otherwise be exploited for the solution of the clearance problem.

Intrusive methods do not purely rely on a simulation over the whole time interval and the dynamic character of the underlying problem is partially exposed under this approach, e.g. through a segmentation using direct optimal control methods (cf. Sec. 6). One typical requirement for this type of model is that the system needs to be available in state-space form. This implies that the model provides the state time derivatives and outputs depending on the current states, inputs, and parameters. Moreover, differentiability requirements need to be met in case direct optimal control methods with gradient-based schemes are employed. In practice, models are rarely given in this form but are rather embedded in a simulation framework (such as Simulink®) and additionally contain non-differentiable features. As such, some modeling effort is required to bring the closed-loop model in a form which is suitable for the application of direct optimal control methods. In particular, the model needs be transformed into a single-point execution model of the form

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}), \tag{9.1}$$

which fulfills the differentiability requirements and allows for an efficient computation of first- and (desirably) second-order derivative information. This transformation can be considered one of the major challenges when dealing with intrusive methods as the process of transcribing the model into the required form has to ensure that the primary characteristics are not altered. However, it is clear that the amount of information provided to the solver is considerably higher compared to the non-intrusive (black-box) approach. First of all, using a segmented transcription method (cf. chapter 6) under the direct optimal control approach the problem can be highly decoupled which is preferable for many optimization algorithms. Moreover, modeling concepts such as

state constraints can be employed to expose internal limits to the worst-case search. Finally, the worst-case inputs can be determined with a high resolution, meaning that the control functions can be parameterized with a large number of parameters.

Hybrid methods can be seen as a middle-ground between the intrusive and non-intrusive approach. The idea behind this approach is to employ a reduced model for the computation of the worst-case solution with a Cat I, Cat II, or Cat III model under the intrusive approach. This solution is then used to increase the information content provided to a Cat III non-intrusive method, e.g. by determining a suitable control parametrization and a meaningful initial guess for the parameters involved in the worst-case search. In this thesis the proposed hybrid method uses Cat I and Cat II models discussed in chapter 8 as reduced models. Clearly, the problem specific information content used under this approach is higher compared to the Cat III non-intrusive (black-box) method. At the same time, the potentially restrictive modeling requirements of the Cat III intrusive approach can be avoided if the reduced models are based on Cat I and Cat II models. The two requirements for the hybrid approach to be effective are that the reduced model at least shares the same tendencies as the full model regarding the worst-case parameter combinations and that the structure, e.g. the number and location of the switches for the worst-case control input, is comparable to a worst-case solution for the full model.

9.1 A Cat III Intrusive Approach for Worst-Case Analysis

In the following the Cat III intrusive approach for optimal control based clearance is introduced which is based on the application of direct optimal control methods (Contribution III, cf. Refs. [110, 111, 112, 113]). It is noteworthy, that similar to global parameter optimization methods which are often employed for non-intrusive methods there exist global algorithms for solving optimal control problems. For example, Dynamic Programming (DP) [114] approaches guarantee that the discretized solution is in fact globally optimal. Unfortunately, DP methods for nonlinear systems are only able to operate in low-dimensional state-spaces as not only the temporal dimension but also the whole state-space is discretized. However, in practice aircraft closed-loop models feature several tens or often even more than a hundred states. In particular, care must be taken if a reduced model is used in order to preserve the essential dynamic characteristics for performing the clearance task. As such, the following section focus on the direct approach which given the current state of the art for these methods allow for the solution of the clearance problem in a fairly general setting.

9.1.1 Nonlinear Optimal Control Formulation

Under the Cat III intrusive approach the nonlinear dynamic model $\mathbf{f} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \to \mathbb{R}^{n_x}$ in state-space form

$$\dot{\mathbf{x}}(t) = \mathbf{f}\left(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}\right), \tag{9.2}$$

with the state vector $\mathbf{x} \in \mathbb{R}^{n_x}$, the control vector $\mathbf{u} \in \mathbb{R}^{n_u}$, and the parameter vector $\mathbf{p} \in \mathbb{R}^{n_p}$ is considered. Furthermore, the nonlinear output function $j : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \to \mathbb{R}$

$$y(t) := j\left(\mathbf{x}(t), \mathbf{p}\right), \tag{9.3}$$

for the scalar output $y(t) \in \mathbb{R}$ represents the criterion under investigation. As in the linear case a subset of the state variables defined by the index set $\mathcal{I}_b$ is assumed to be bounded

$$x_{i,lb} \leq x(t) \leq x_{i,ub}, i \in \mathcal{I}_b \tag{9.4}$$

and the control set $\mathcal{U}$ is defined as:

$$\mathbf{u}_{lb} \leq \mathbf{u}(t) \leq \mathbf{u}_{ub} \tag{9.5}$$

Furthermore, the parameters are constrained to an admissible region defined by

$$\mathbf{g}(\mathbf{p}) \leq \mathbf{0}, \tag{9.6}$$

with the constraint function $\mathbf{g} : \mathbb{R}^{n_p} \to \mathbb{R}^{n_g}$. In the following all functions for the Cat III clearance problem under the intrusive approach are assumed to be at least twice continuously differentiable. The general Mayer-form of this problem on a time interval $I^t := [t_0, t_f], t_f > t_0$ for a suitable fixed final time t_f can be formalized as follows:

$$\begin{aligned}
\underset{\mathbf{u}(t),\,\mathbf{p}}{\text{minimize}} \quad & j\left(\mathbf{x}(t_f), \mathbf{p}\right) \\
\text{subject to} \quad & \dot{\mathbf{x}}(t) - \mathbf{f}\left(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}\right) = \mathbf{0}, \\
& \mathbf{x}(t_0) = \mathbf{x}_0, \\
& x_{j,lb} \leq x_j(t) \leq x_{j,ub},\ j \in \mathcal{I}_b, \\
& \mathbf{u}(t) \in \mathcal{U}, \\
& \mathbf{g}(\mathbf{p}) \leq \mathbf{0}, \\
& t \in I^t
\end{aligned} \tag{9.7}$$

The direct optimal control approach (cf. Sec. 6) can be used to transcribe the optimal control problem (9.7) into a parameter optimization problem. As for the linear case

(cf. Sec. 8.1.3) a Backward Euler method with step size h is used in the following if the focus is on the investigation of the control functions as this discretization approach is observed to yield rather "clean" control histories (cf. Sec. 8.2.2):

$$\begin{aligned}
\underset{\mathbf{p},\mathbf{x}^{[0]},\mathbf{x}^{[i]},\mathbf{u}^{[i]},i=1,\ldots,N}{\text{minimize}} \quad & j\left(\mathbf{x}^{[N]},\mathbf{p}\right) \\
\text{subject to} \quad & \mathbf{x}^{[i]} = \mathbf{x}^{[i-1]} + h\mathbf{f}\left(\mathbf{x}^{[i]},\mathbf{u}^{[i]},\mathbf{p}\right),\ i=1,\ldots,N, \\
& \mathbf{x}^{[0]} = \mathbf{x}_0, \\
& x_{j,lb}^{[i]} \leq x_j^{[i]} \leq x_{j,ub}^{[i]},\ \forall j \in \mathcal{I}_b, i = 1,\ldots,N, \\
& \mathbf{u}_{lb}^{[i]} \leq \mathbf{u}^{[i]} \leq \mathbf{u}_{ub}^{[i]}, i = 0,\ldots,N, \\
& \mathbf{g}(\mathbf{p}) \leq \mathbf{0}, \\
& t \in I^t
\end{aligned} \tag{9.8}$$

Similarly, if the focus is primarily on the value of the cost function a Trapezoidal collocation approach is used for which the Cat III clearance problem reads as follows:

$$\begin{aligned}
\underset{\mathbf{p},\mathbf{x}^{[i]},\mathbf{u}^{[i]},i=0,\ldots,N}{\text{minimize}} \quad & j\left(\mathbf{x}^{[N]},\mathbf{p}\right) \\
\text{subject to} \quad & \mathbf{x}^{[i]} = \mathbf{x}^{[i-1]} + h\frac{\mathbf{f}\left(\mathbf{x}^{[i-1]},\mathbf{u}^{[i-1]},\mathbf{p}\right) + \mathbf{f}\left(\mathbf{x}^{[i]},\mathbf{u}^{[i]},\mathbf{p}\right)}{2},\ i=1,\ldots,N, \\
& \mathbf{x}^{[0]} = \mathbf{x}_0, \\
& x_{j,lb}^{[i]} \leq x_j^{[i]} \leq x_{j,ub}^{[i]},\ \forall j \in \mathcal{I}_b, i = 1,\ldots,N, \\
& \mathbf{u}_{lb}^{[i]} \leq \mathbf{u}^{[i]} \leq \mathbf{u}_{ub}^{[i]},\ i = 0,\ldots,N, \\
& \mathbf{g}(\mathbf{p}) \leq \mathbf{0}, \\
& t \in I^t
\end{aligned} \tag{9.9}$$

Note that in the linear case the initial condition is changed depending on the value of the parameter $\mathbf{p}$ (cf. Sec. 8.3.2) in order to linearize the model around an equilibrium point. This requirement is dropped under the Cat III type model as a nonlinear model can directly be used under this approach. As such, the system may originate from a general initial point $\mathbf{x}_0$ in the state-space which is suitable for the particular clearance problem under investigation. Moreover, as illustrated in the linear case (cf. Sec. 8.2.2) the solution can become highly in-sensitive to control actions in the beginning of the time interval if the final time t_f is too large. For the nonlinear case it is possible to let the optimizer choose the final time t_f, i.e. a free final time can be used. Thus, the time point where the worst-case occurs becomes part of the optimization problem. However, in order to constrain the solution to a reasonable time interval it is advisable to either impose limits on the free final time $t_f \in [t_0, t_{f,ub}], t_{f,ub} > t_0$, or to add a (small) penalty depending on t_f to the cost function which monotonically increases towards

higher values of the terminal time. As described in chapter 6, optimal control problems which are discretized using a segmented method represent high-dimensional but very sparse nonlinear programming problems (NLPs). These NLPs can be efficiently solved by one of the methods described in Sec. 4.5. In particular, computational experience suggests that the interior point solver IPOPT [62] is highly efficient and effective for this purpose. For all numerical solutions which are presented in the following the optimal control toolbox *FALCON.m*[1] is used which is co-developed by the author (Contribution IV). In particular, a transcription tool for Simulink® models for internal use is developed to obtain a model in state-space form which can be used by *FALCON.m* for the dynamic optimization. Moreover, post-optimal sensitivity analysis can be performed to identify influential parameters regarding the *worst-case* cost function value (cf. Sec. 8.3.2). One practical method to obtain the post-optimal sensitivities of the optimal cost function for nonlinear clearance problems is to use a parameter embedding approach. Under this approach an additional equality constraint $\mathbf{p} - \mathbf{p}_0 = \mathbf{0}$ with the nominal parameter vector $\mathbf{p}_0$ is introduced in the clearance problem which fixes the parameters to their nominal values. The negative Lagrange multipliers associated with this parameter embedding constraint then represent the post-optimal sensitivities of the optimal cost function with respect to the parameters (cf. Ref. [66]).

9.1.2 The Nonlinear Benchmark Problem

The nonlinear benchmark model which is used for the illustration of the Cat III intrusive approach represents a nonlinear version of the closed-loop system introduced in Sec. 8.2.1 for the Cat I and Cat II type models. It is noteworthy that in the nonlinear case the model is not separated into a longitudinal and lateral model but for all numerical experiments the full dynamic model is considered. For this nonlinear closed-loop model only the differentiable part is used and other components such as internal limiters are dropped in order to enable the numerical solution using gradient-based solvers (the full model is then treated under the Cat III hybrid approach, cf. Sec. 9.2). In particular, the subsystems responsible for not exceeding the limits of the servomechanism are removed and replaced by state constraints for the rate and position states. The only exception are table data (e.g. for gain scheduling) which are interpolated using a multi-linear method. Typically, it is advisable to use differentiable interpolation methods for direct optimal control approaches. However, computational experience indicates that the numerical solution for the clearance problems presented in the following section can be obtained even using multi-linear interpolation. Hence, in order to keep the dynamic model used for the Cat III intrusive approach as close as possible to the full model the table data interpolation method is left unaltered.

[1] www.falcon-m.com

The nonlinear model has 82 states and includes the nonlinear plant model with aeroelastic effects, sensor models, the gain scheduled controller, and the dynamic model for the servomechanisms of all primary control surfaces including rate and position limits (introduced as state constraints). Moreover, six control inputs, namely the normal load factor command, the lateral load factor command, the roll angle command and the wind commands for the longitudinal, normal, and lateral wind gust are used in different combinations to test the system. Inputs which are not used for the particular clearance problem are fixed to the trim condition. For example, in the case of the criteria associated with the longitudinal plane this implies that all controls for the lateral motion, i.e. the lateral load factor command, the roll angle command, and the lateral wind gust command are set to the trim values. The same holds for the investigation of the criteria in the lateral plane regarding the inputs associated with the longitudinal plane, i.e. the normal load factor command as well as the longitudinal and normal wind commands. Additionally, the model depends on the parameters introduced in Sec. 8.3.2 (cf. Eq. 8.169 and Eq. 8.170) under the bi-level optimization approach.

9.1.3 Numerical Results

Both the nominal and the parameter dependent clearance problems are investigated in the following. All results under the Cat III intrusive method are compared to the results obtained in the linear cases using the same transcription methods and discretization step lengths (cf. Sec. 8.2.2 and Sec. 8.3.2). As such, for the solution of the nominal problems and the investigation of the worst-case control functions the Backward Euler method is used. Regarding the determination of the worst-case parameter combination a Trapezoidal transcription is employed. Additional states (rigid body states, error controller states, and wind states) which are not presented this section are provided for reference in App. B. The initial guess for all states and controls of the nonlinear model is set to values corresponding to the trim condition (cf. Sec. 8.2.2). All four criteria, i.e. the aerodynamic AoA $\alpha_A(t_f)$, the normal load factor $n_z(t_f)$, the aerodynamic AoS $\beta_A(t_f)$, and the lateral load factor $n_y(t_f)$ at $t_f = 6\ s \approx 5T_{SP}$, are investigated under the Cat III intrusive approach. For the optimization of the criteria corresponding to the longitudinal plane ($\alpha_A(t_f)$, $n_z(t_f)$) the inputs for the normal load factor $n_{z,c}(t)$ as well as the normal wind command $w_W(t)$ and the longitudinal wind command $u_W(t)$ are used. Similarly, for the optimization of the criteria corresponding to the lateral plane ($\beta_A(t_f)$, $n_y(t_f)$) the lateral load factor command $n_{y,c}(t)$, the roll angle command $\Phi_c(t)$, and the lateral wind input $v_W(t)$ are used. Moreover, state constraints for the rate and position limits are imposed for all actuators (elevator, aileron, rudder). The numerical results for the nominal clearance problems are presented in Figs. 9.1-9.3. In all figures the results using the nonlinear model (black lines) are compared to the results

obtained from the linear models (blue lines). The author would like to emphasize that the solution of the nonlinear optimal control problems posed considerable challenges regarding the convergence of the NLPs. Here, for all problems except the maximization of $n_y(t_f)$ (nominal case) the solver converged to an optimal solution. In case of the maximization of $n_y(t_f)$ the solver appeared to get stuck despite the fact that the optimization was performed using different combinations of scaling values, derivative options (BFGS, Hessian), discretization methods (Backward Euler / Trapezoidal), discretization step lengths, and initial guesses. Considering the complexity of the nonlinear closed-loop model it is deemed practically impossible to narrow down the root cause to a single element (or a combination of events) in the optimization problem which inhibits convergence in this particular case.

Regarding the numerical results for the other nominal cases, i.e. the maximization of $n_z(t_f)$ and $\alpha_A(t_f)$ in the longitudinal plane as well as $\beta_A(t_f)$ in the lateral plane, it is noteworthy that, as for the linear case, the worst-case controls obtained from the solutions of the nonlinear clearance problems are of bang-bang type in case of the wind commands and bang-bang with singular parts for the other command inputs. Here, the singular parts of the command inputs correspond to the time intervals where the actuators reach rate or position limits. For the maximization of the normal load factor $n_z(t_f)$ (cf. Fig. 8.23) the structure obtained for the nonlinear model shares common features with the linear results. It is interesting, that in the nonlinear case the load factor command $n_{z,c}(t)$ and the longitudinal wind command $u_W(t)$ in the first half of the time interval are on the opposite bound compared to the linear case. In the second half all three controls appear to exhibit a comparable structure and the response $n_z(t)$ for both models yields almost the same final value. Merely, the last switching time point of the longitudinal wind command $u_W(t)$ appears to be shifted and the entry, respectively exit time points of the elevator rate constrained arcs do not match perfectly. The normal wind command $w_W(t)$, however, is almost identical for both the Cat II and Cat III model. Regarding the maximization of $\alpha_A(t_f)$ (cf. Fig. 9.2) it is observed that the nonlinear model reaches the upper position limit of the elevator, which is not the case for the linear model. As such, the load factor command $n_{z,c}(t)$ for the nonlinear model exhibits additional arcs on which the elevator rate $x_{e,v}(t)$ is controlled to zero. In fact, the elevator is saturated almost on the entire second half of the time interval (excluding a short regular interval where the normal load factor command is on the bound). Note that in the linear case the elevator position gets very close to the upper limit but does not quite touch it. As the control inputs for the first half of the time interval are very similar this is probably contributed to the fact that the linear and nonlinear dynamics propagate in a slightly different manner further away from the linearization point. Moreover, it is interesting to observe that the switching structures regarding the wind commands ($u_W(t)$,$w_W(t)$) appear to be very similar.

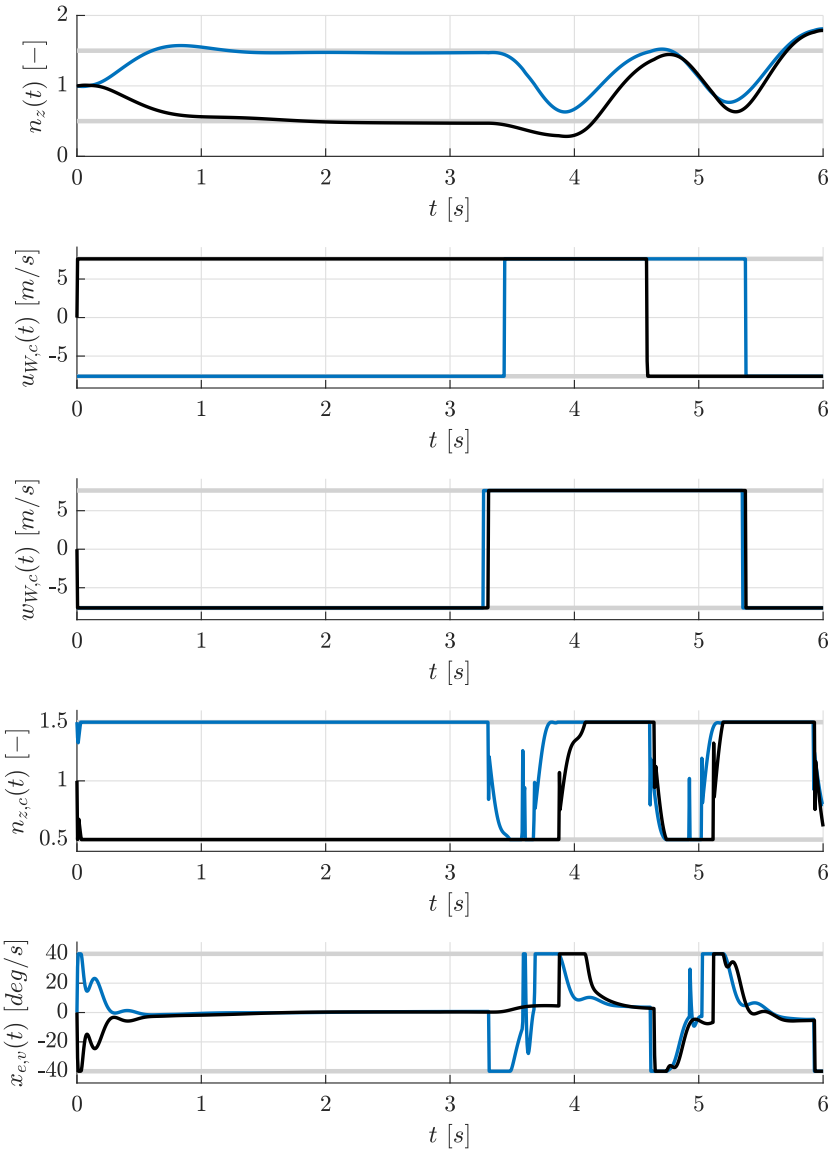


Figure 9.1: Optimal results for the longitudinal plane using a Cat III model (black lines) and a Cat II model (blue lines) for maximizing $n_z(t_f), t_f = 6\ s$ with wind including the normal load factor $n_z(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, the normal load factor command $n_{z,c}(t)$, and the elevator rate $x_{e,v}(t)$.

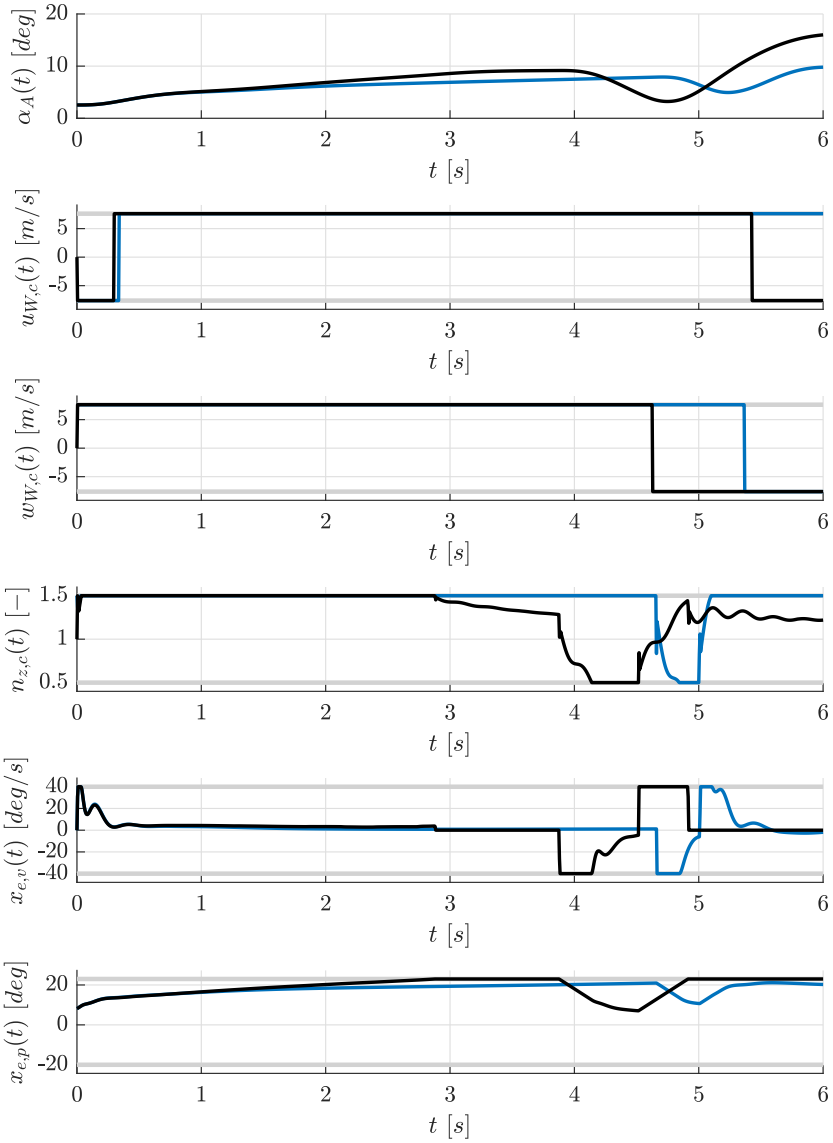


Figure 9.2: Optimal results for the longitudinal plane using a Cat III model (black lines) and a Cat II model (blue lines) for maximizing $\alpha_A(t_f), t_f = 6\ s$ with wind including the AoA $\alpha_A(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, the normal load factor command $n_{z,c}(t)$, the elevator rate $x_{e,v}(t)$, and the elevator position $x_{e,p}(t)$.

For the maximization of $\beta_A(t_f)$ (cf. Fig. 9.3) the response and the structure of the control functions from the linear and nonlinear analysis appear to be comparable. In particular, the lateral wind command $v_W(t)$ and the lateral load factor command $n_y(t)$ match very closely. However, the roll angle command $\Phi_c(t)$ appears to react differently and exhibits rather complicated switches between different state constrained arcs.

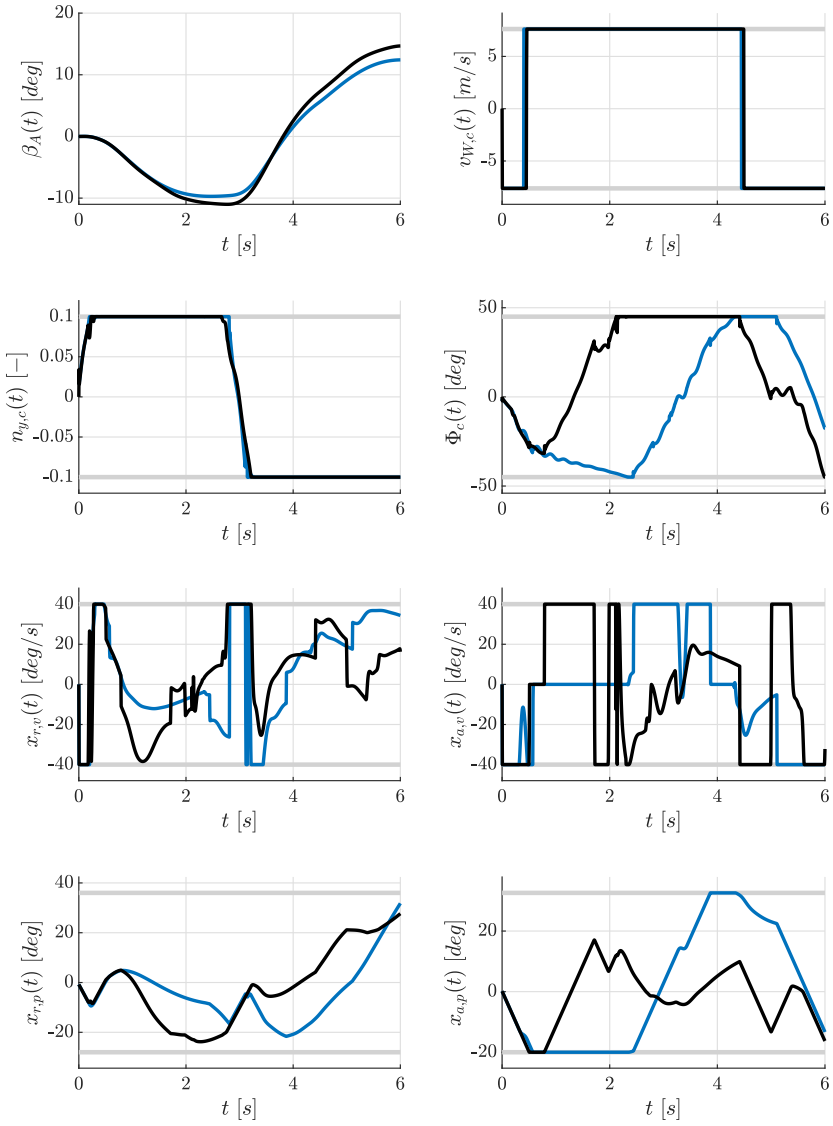


Figure 9.3: Optimal results for the lateral plane using a Cat III model (black lines) and a Cat II model (blue lines) for maximizing $\beta_A(t_f), t_f = 6\ s$ with wind including the AoS $\beta_A(t)$, the lateral wind velocity command $v_{W,c}(t)$, the lateral load factor command $n_{y,c}(t)$, the roll angle command $\Phi_c(t)$, the rudder rate $x_{r,v}(t)$, the aileron rate $x_{a,v}(t)$, the rudder position $x_{r,p}(t)$, and the aileron position $x_{a,p}(t)$.

As for the linear analysis (cf. Sec. 8.3.2, Figs. 8.30-Fig. 8.33) the post-optimal sensitivities at the optimal solutions are presented. The indications regarding the degree of influence on the worst-case solution are similar to the results obtained from the linear analysis for the maximization of $n_z(t_f)$ (cf. Fig. 9.4) and the maximization of $\beta_A(t_f)$ (cf. Fig. 9.6). Note that the structure of the solution for the worst-case control histories (cf. Fig. 8.23 and Fig. 9.3) is comparable to the linear case. However, the indications for the maximization of $\alpha_A(t_f)$ (cf. Fig. 9.5) differ significantly. This may be contributed to the fact that in the nonlinear case the elevator position limit is reached. Observe that in the linear results merely two short state constraint arcs for the elevator rate state are visible. As the sensitivity information is local there is no information contained in the linear results that the position of the actuator is close to being saturated. In the nonlinear results the actuator is saturated (rate and position) for the most part of the second half of the time interval. As such, the post-optimal sensitivities need to be understood under the following conditions: in the linear case the elevator position is not saturated and does not get saturated under small parameter variations whereas in the nonlinear case the elevator position is saturated and stays saturated under small parameter variations. It is not un-reasonable that this fact has a considerable effect on the optimal solution of the problem and is, as such, reflected in the post-optimal sensitivities of the optimal cost function value regarding the local influence of the parameters.

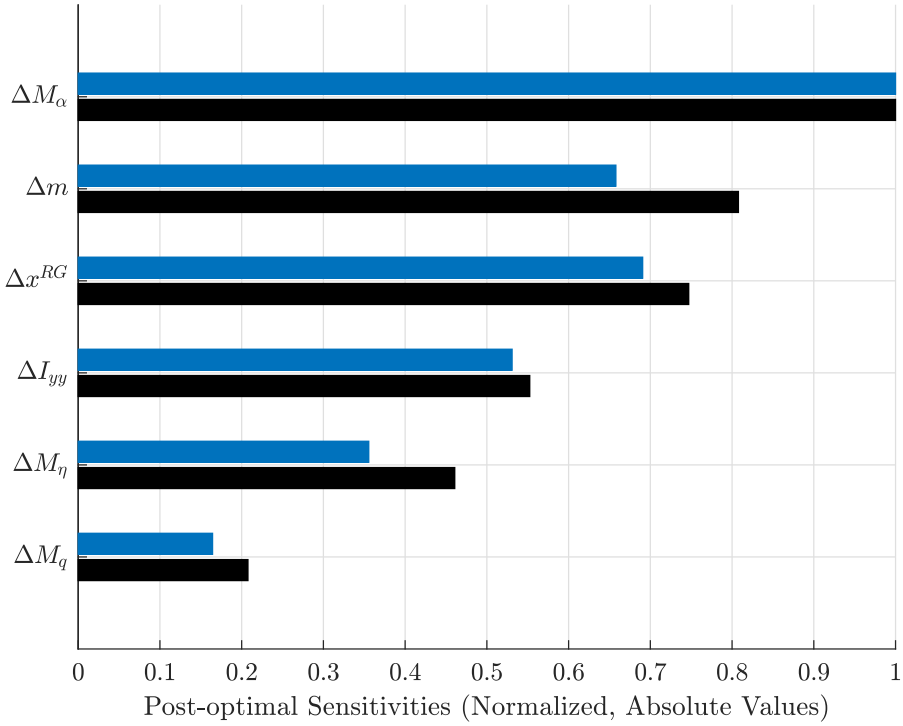


Figure 9.4: Post-optimal sensitivities for maximizing $n_z(t_f), t_f = 6\ s$ under a Cat II (blue) and Cat III (black) type model using the normal load factor command $n_{z,c}(t)$ as well as the wind commands $u_{W,c}(t)$ and $w_{W,c}(t)$.

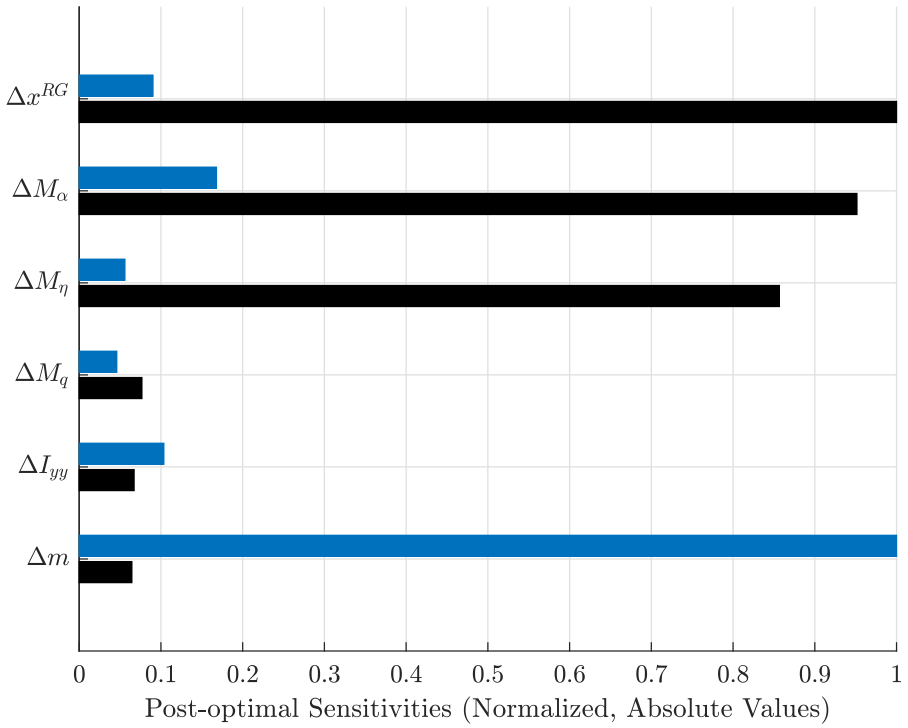


Figure 9.5: Post-optimal sensitivities for maximizing $\alpha_A(t_f), t_f = 6\ s$ under a Cat II (blue) and Cat III (black) type model using the normal load factor command $n_{z,c}(t)$ as well as the wind commands $u_{W,c}(t)$ and $w_{W,c}(t)$.

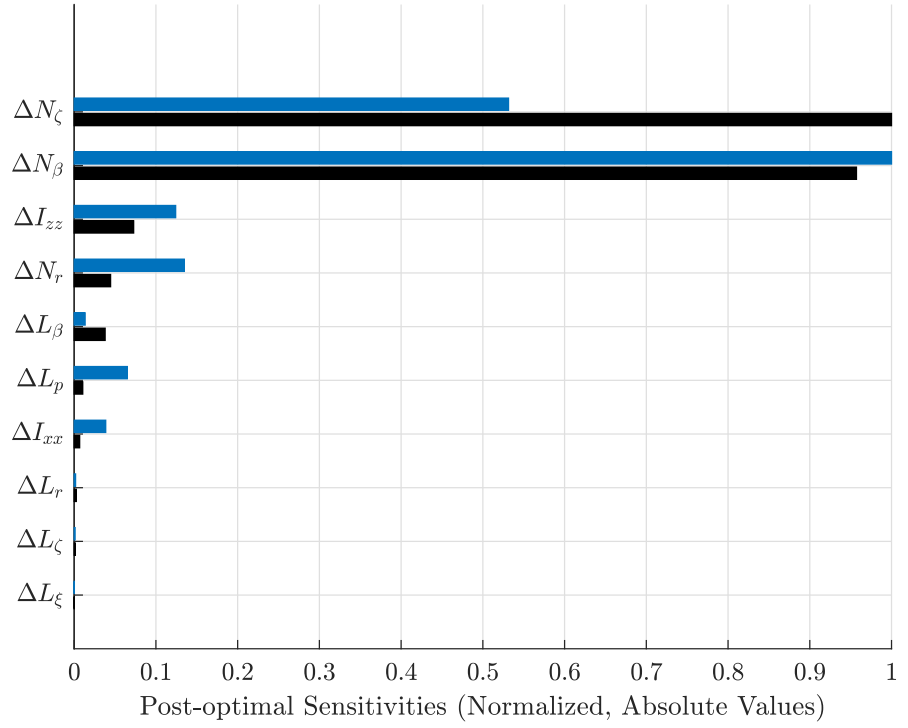


Figure 9.6: Post-optimal sensitivities for maximizing $\beta_A(t_f), t_f = 6\ s$ under a Cat II (blue) and Cat III (black) type model using the lateral roll angle command $\Phi_c(t)$, the lateral load factor command $n_{y,c}(t)$ as well as the wind command $v_{W,c}(t)$.

For the determination of the worst-case parameter combination the final time is increased to $t_f = 8\ s$ as for the parameter dependent cases in Sec. 8.3.2. The results presented in Tab. 9.1-Tab. 9.4 compare the solution obtained from the bi-level approach using a gradient-based method (SQP, cf. Tabs. 8.1-8.4) to the solutions obtained from the nonlinear analysis. The optimal cost function values are similar for the load factors $\hat{n}_z(t_f)$ and $\hat{n}_y(t_f)$ but differ for the aerodynamic angles $\hat{\alpha}_A(t_f)$ and $\hat{\beta}_A(t_f)$. In particular, the values obtained for the AoA are clearly non-physical as the aircraft would have stalled far below the optimal value.

It is remarkable, that the worst-case parameter combinations of the linear and nonlinear results are very similar (merely, the values for $\Delta\hat{L}_p$ and $\Delta\hat{I}_{xx}$ in Tab. 9.3 and Tab. 9.4 show different values in the comparison). This may be seen as an indicator that the Cat II and Cat III models qualitatively behave in a similar way. Hence, Cat II models may be used as an indicator for the direction in the parameter space where the worst-case solution is expected. This observation together with the indications that the worst-case structure of the optimal control histories share similarities is the motivation behind the Cat III hybrid approach presented in Sec. 9.2 which uses information from a reduced model to improve the worst-case search under a Cat III non-intrusive (black-box) method.

Table 9.1: Comparison of the bi-level optimization results of the Cat II model with the results obtained from the Cat III intrusive method (box-bounded parameters) for the maximization of $\alpha_A(t_f)$ using the normal load factor command $n_{z,c}(t)$ as well as the wind commands $u_W(t)$ and $w_W(t)$.

Case	$\hat{\alpha}_A(t_f)$	$\Delta\hat{M}_\alpha$	$\Delta\hat{M}_q$	$\Delta\hat{M}_\eta$	$\Delta\hat{m}$	$\Delta\hat{I}_{yy}$	$\Delta\hat{x}^{GN}$
CAT II	14.166 $[deg]$	-0.100	-0.100	0.100	0.100	0.100	-0.100
CAT III	29.877 $[deg]$	-0.100	-0.100	0.100	0.100	0.100	-0.100

Table 9.2: Comparison of the bi-level optimization results of the Cat II model with the results obtained from the Cat III intrusive method (box-bounded parameters) for the maximization of $n_z(t_f)$ using the normal load factor command $n_{z,c}(t)$ as well as the wind commands $u_W(t)$ and $w_W(t)$.

Case	$\hat{n}_z(t_f)$	$\Delta\hat{M}_\alpha$	$\Delta\hat{M}_q$	$\Delta\hat{M}_\eta$	$\Delta\hat{m}$	$\Delta\hat{I}_{yy}$	$\Delta\hat{x}^{GN}$
CAT II	2.222	-0.100	-0.100	0.100	-0.100	0.100	-0.100
CAT III	2.349	-0.100	-0.100	0.100	-0.100	0.100	-0.100

Table 9.3: Comparison of the bi-level optimization results of the Cat II model with the results obtained from the Cat III intrusive method (box-bounded parameters) for the maximization of $\beta_A(t_f)$ using the roll angle command $\Phi_c(t)$, lateral load factor command $n_y(t)$, and the wind command $v_W(t)$.

Solver	$\hat{\beta}_A(t_f)$	$\Delta\hat{L}_\beta$	$\Delta\hat{L}_p$	$\Delta\hat{N}_\beta$	$\Delta\hat{N}_\zeta$	$\Delta\hat{I}_{xx}$	$\Delta\hat{I}_{zz}$
CAT II	16.325 $[deg]$	-0.100	**0.007**	-0.100	0.100	**0.100**	0.100
CAT III	20.532 $[deg]$	-0.100	**0.100**	-0.100	0.100	**-0.100**	0.100

Table 9.4: Comparison of the bi-level optimization results of the Cat II model with the results obtained from the Cat III intrusive method (box-bounded parameters) for the maximization of $n_y(t_f)$ using the roll angle command $\Phi_c(t)$, lateral load factor command $n_y(t)$, and the wind command $v_W(t)$.

Solver	$\hat{n}_y(t_f)$	$\Delta\hat{L}_\beta$	$\Delta\hat{L}_p$	$\Delta\hat{N}_\beta$	$\Delta\hat{N}_\zeta$	$\Delta\hat{I}_{xx}$	$\Delta\hat{I}_{zz}$
Cat II	0.323	-0.100	**-0.100**	-0.100	0.100	**0.100**	0.100
Cat III	0.344	-0.100	**0.100**	-0.100	0.100	**-0.100**	0.100

9.1.4 Multi-Criteria and Reachability Analysis

As for Cat I and Cat II type models (cf. Sec. 8.4) it can be of interest for Cat III type models which values of combined criteria may occur. Reachability analysis for nonlinear models is considerably more involved compared to the linear case. In particular, the reachable set is not guaranteed to be convex in the general case. The computation of forward and backward reachable sets in the context of safety analysis has been of particular interest in the last years (see for example Ref. [115]). A popular approach for this purpose are level-set methods which require the formulation of the reachability problem based on a suitable Hamilton-Jacobi equation (cf. Refs. [116, 117]). However, the applicability of these approaches in higher dimensional state-spaces is computationally challenging. Nevertheless, these methods have been proven useful in aerospace applications such as safety analysis during landing (cf. Ref. [118]) or maneuvering envelope estimation (cf. Ref. [119]). As for the linear case there also exist approaches based on direct optimal control methods such as the Distance Fields on Grids (DFOG) method (cf. Ref. [33]). This method relies on the solution of multiple optimal control problems which is conceptually close to the formulation of the clearance problems investigated in this thesis. Moreover, the requirements on the type of dynamic system are fairly non-restrictive and the consideration of input and state constraints is supported. As such, this method is considered in the following for the general Cat III models under the intrusive approach in order to approximate the forward reachable set at a time point t_f. For the DFOG method the space of interest is discretized using a grid $\mathbb{G}_{n_g}^{DFOG}$

with constant mesh size containing n_g grid points:

$$\mathbb{G}_{n_g}^{DFOG} := \{\mathbf{y}_k : k = 0, \ldots, n_g - 1, \mathbf{y}_k \in \mathbb{R}^{n_y}\} \tag{9.10}$$

The reachable set can then be approximated by solving an appropriate optimal control problem for each of the grid points in $\mathbb{G}_{n_g}^{DFOG}$. This optimal control problem solved for each of the points $\mathbf{y}_k, k = 0, \ldots, n_g - 1$ using a Cat III model reads

$$\begin{aligned} \underset{\mathbf{u}(t),\,\mathbf{p}}{\text{minimize}} \quad & \|\mathbf{j}\left(\mathbf{x}(t_f), \mathbf{p}\right) - \mathbf{y}_k\|_2 \\ \text{subject to} \quad & \dot{\mathbf{x}}(t) - \mathbf{f}\left(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}\right) = \mathbf{0}, \\ & \mathbf{x}(t_0) = \mathbf{x}_0, \\ & \mathbf{x}(t) \in \mathcal{X}, \\ & \mathbf{u}(t) \in \mathcal{U}, \\ & \mathbf{g}(\mathbf{p}) \leq \mathbf{0}, \\ & t \in I^t, \end{aligned} \tag{9.11}$$

with the vector-valued output function $\mathbf{j} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \to \mathbb{R}^{n_y}$:

$$\mathbf{y}(t_f) = \mathbf{j}\left(\mathbf{x}(t_f), \mathbf{p}\right) \tag{9.12}$$

Here, the problem formulation (9.11) is adapted from Ref. [33] for the clearance application considered in this thesis. In particular, the formulation is extended by the parameter dependency and the states are assumed to be constrained to the set $\mathcal{X}$ defined by:

$$x_{j,lb} \leq x_j(t) \leq x_{j,ub}, \ \forall j \in \mathcal{I}_b \tag{9.13}$$

The discretization of this problem using the Trapezoidal collocation method with step size h yields:

$$\begin{aligned} \underset{\mathbf{p},\,\mathbf{x}^{[i]},\,\mathbf{u}^{[i]},\,i=0,\ldots,N}{\text{minimize}} \quad & \left\|\mathbf{j}\left(\mathbf{x}^{[N]}, \mathbf{p}\right) - \mathbf{y}_k\right\|_2 \\ \text{subject to} \quad & \mathbf{x}^{[i]} = \mathbf{x}^{[i-1]} + h\frac{\mathbf{f}\left(\mathbf{x}^{[i-1]}, \mathbf{u}^{[i-1]}, \mathbf{p}\right) + \mathbf{f}\left(\mathbf{x}^{[i]}, \mathbf{u}^{[i]}, \mathbf{p}\right)}{2}, \ i{=}1, \ldots, N, \\ & \mathbf{x}^{[0]} = \mathbf{x}_0, \\ & \mathbf{u}^{[i]} \in \mathcal{U}, \ i = 0, \ldots, N, \\ & \mathbf{x}^{[i]} \in \mathcal{X}, \ i = 1, \ldots, N \end{aligned} \tag{9.14}$$

Obviously, the reachable set approximation requires the solution of a potentially large number of optimal control problems depending on the number of grid points n_g (see Fig. 9.7). Moreover, it is important to mention that the nonlinear optimal control prob-

lems associated with each grid point can usually only be solved for a local minimizer. However, it is observed for the DFOG method "*that the lack of global optimality is often cured by considering many grid points*" [33]. Hence, the high number of optimal control problems which need to be solved for this approach are expected to mitigate the local nature of the solution using direct methods to a certain extent.

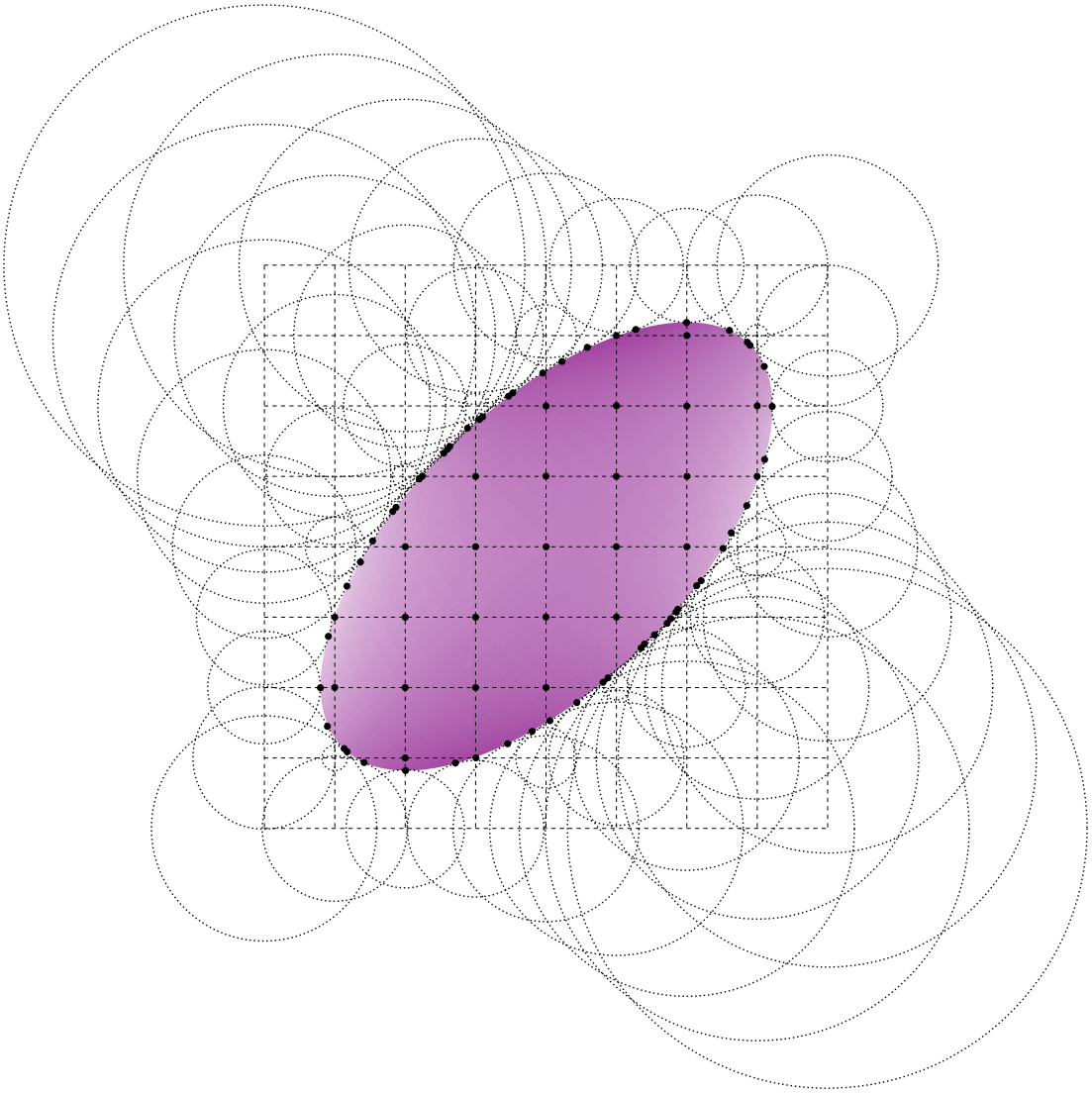


Figure 9.7: Illustration of the reachable set approximation using a distance function on a rectangular grid with constant mesh size. The black markers represent the points with the smallest distances to the respective grid points.

The application of the DFOG method is applied in Ref. [110] by the author for the estimation of the tracking errors of a nonlinear aircraft closed-loop system. For illustration purposes this analysis is re-produced in the following. The system investigated in this reference represents a reduced model for a F-16 fighter aircraft with a Model Reference Adaptive Controller (MRAC). For this system, the plant is modeled by a short

period approximation of the form:

$$\dot{\mathbf{x}}_P(t) = \mathbf{A}_P \mathbf{x}_P(t) + \mathbf{B}_P \lambda_{M_\eta} u(t) \tag{9.15}$$

Here, the state vector of the plant $\mathbf{x}_P(t) = [\alpha_P(t), q_P(t)]^T$ collects the AoA $\alpha_P(t)$ as well as the pitch rate $q_P(t)$. Moreover, the input is denoted with $u(t)$ and the system is subject to the parametric uncertainty $\lambda_{M_\eta} \in [0.75, 1.25]$ for the elevator effectiveness. In addition, the uncertain system matrix $\mathbf{A}_P$ is of the form

$$\mathbf{A}_P := \begin{bmatrix} Z_\alpha & 1 + Z_q \\ M_\alpha \lambda_{M_\alpha} & M_q \lambda_{M_q} \end{bmatrix}, \tag{9.16}$$

and contains the entries $M_\alpha = -30.79$, $M_q = -3.75$, $Z_\alpha = -1.84$, and $Z_q = -0.09$ as well as the multiplicative uncertainties $\lambda_{M_\alpha} \in [0.75, 1.25]$ and $\lambda_{M_q} \in [0.75, 1.25]$ for the pitch stiffness and pitch damping. In addition, the uncertain input matrix $\mathbf{B}_P$

$$\mathbf{B}_P := \begin{bmatrix} 0 \\ M_\eta \end{bmatrix}, \tag{9.17}$$

contains the coefficient $M_\eta = -12.75$. The reference model for the MRAC

$$\dot{\mathbf{x}}_M(t) = \mathbf{A}_M \mathbf{x}_M(t) + \mathbf{B}_M K_r r(t), \tag{9.18}$$

with the feed-forward gain $K_r = -1.55$ has two states $\mathbf{x}_M(t) = [\alpha_M(t), q_M(t)]^T$, namely the reference state for the AoA $\alpha_M(t)$ and the pitch rate $q_M(t)$. These states should be tracked by the plant states $\alpha_P(t)$ and $q_P(t)$. The system matrix of the reference model $\mathbf{A}_M$ is of the form

$$\mathbf{A}_M := \begin{bmatrix} Z_\alpha & 1 + Z_q \\ M_{\alpha,M} & M_{q,M} \end{bmatrix}, \tag{9.19}$$

with the desired values $M_{\alpha,M} = -11.38$ and $M_{q,M} = -4.16$. The reference command $r(t)$ for the AoA is subject to box bounds $r(t) \in [-5\,\mathrm{deg}, +5\,\mathrm{deg}]$ and the input matrix $\mathbf{B}_M$ is defined as $\mathbf{B}_M := \mathbf{B}_P$. The reference model (9.18) exhibits desirable properties as it ensures stationary accuracy and has an ideal damping of $\frac{\sqrt{2}}{2}$. The task of the controller is to let the plant states follow the reference model states, i.e. $\mathbf{x}_P(t) \approx \mathbf{x}_M(t)$, as in this case these properties of the reference dynamics are transferred to the plant.

The matching conditions

$$\mathbf{A}_M = \mathbf{A}_P + \mathbf{B}_P \lambda_{M_\eta} \boldsymbol{\Theta}_x^*(t), \tag{9.20}$$

$$\mathbf{B}_P K_r = \mathbf{B}_P \lambda_{M_\eta} \Theta_r^*(t), \tag{9.21}$$

relating the matrices of the plant (9.15) to those of the reference model (9.18) guarantee that this control objective is achievable. In these conditions, the so-called ideal parameters are denoted with $\mathbf{\Theta}_x^* \in \mathbb{R}^{1\times 2}$ and $\Theta_r^* \in \mathbb{R}$. Observe that these ideal parameters are uncertain as they depend on the multiplicative uncertainties λ_{M_α}, and λ_{M_q} through the plant matrix $\mathbf{A}_P$ and the uncertainty λ_{M_η}.

The following adaptive control law (cf. Ref. [120])

$$u(t) = \mathbf{\Theta}_x(t)\mathbf{x}_P(t) + \Theta_r(t)r(t) = \mathbf{\Theta}(t)\boldsymbol{\omega}(t), \tag{9.22}$$

$$\dot{\mathbf{\Theta}}^T(t) = -\mathbf{\Gamma}\boldsymbol{\omega}(t)\mathbf{e}_C^T(t)\mathbf{P}\mathbf{B}_P, \tag{9.23}$$

with the regressor vector $\boldsymbol{\omega}^T(t) = [\mathbf{x}_P^T(t), r(t)]$, the parameter vector $\mathbf{\Theta}(t) = [\mathbf{\Theta}_x(t), \Theta_r(t)]$ as well as the error vector $\mathbf{e}_C(t) = [e_\alpha(t), e_q(t)]^T = \mathbf{x}_P(t) - \mathbf{x}_M(t)$ ensures that the plant asymptotically tracks the reference model, i.e. $\lim\limits_{t\to\infty} \mathbf{e}_C(t) = \mathbf{0}$.

In Eq. (9.23) the matrix $\mathbf{\Gamma} := diag([\Gamma, \Gamma, \Gamma])$ with the learning rate $\Gamma \in \mathbb{R}$ and the symmetric, positive definite matrix $\mathbf{P} \in \mathbb{R}^{2\times 2}$ obtained from the solution of the Lyapunov equation

$$\mathbf{A}_M^T\mathbf{P} + \mathbf{P}\mathbf{A}_M = -\mathbf{Q}, \tag{9.24}$$

with $\mathbf{Q} := \mathbf{I}$ represent design parameters of the control law.

By adding and subtracting $\mathbf{B}_P K_r$ to the dynamic equations of the plant (9.15) as well as using the matching conditions (9.20) and (9.21) yields:

$$\dot{\mathbf{x}}_P(t) = \mathbf{A}_M\mathbf{x}_P(t) + \mathbf{B}_P K_r r(t) + \mathbf{B}_P\lambda_{M_\eta}(u(t) - \mathbf{\Theta}^*(t)\boldsymbol{\omega}(t)), \qquad \mathbf{\Theta}^*(t) := [\mathbf{\Theta}_x^*(t), \Theta_r^*(t)] \tag{9.25}$$

Moreover, inserting the adaptive control law (9.22) one obtains:

$$\dot{\mathbf{x}}_P(t) = \mathbf{A}_M\mathbf{x}_P(t) + \mathbf{B}_P K_r r(t) + \mathbf{B}_P\lambda_{M_\eta}(\mathbf{\Theta}(t) - \mathbf{\Theta}^*(t))\boldsymbol{\omega}(t). \tag{9.26}$$

Thus, the tracking error dynamics can be expressed as:

$$\dot{\mathbf{e}}_C(t) = \dot{\mathbf{x}}_P(t) - \dot{\mathbf{x}}_M(t) = \mathbf{A}_M\mathbf{e}_C(t) + \mathbf{B}_P\lambda_{M_\eta}(\mathbf{\Theta}(t) - \mathbf{\Theta}^*(t))\boldsymbol{\omega}(t) \tag{9.27}$$

Hence, the complete set of dynamic equations for the nonlinear closed-loop system can be written as:

$$\begin{aligned}
\dot{\mathbf{e}}_C(t) &= \mathbf{A}_M\mathbf{e}_C(t) + \mathbf{B}_P\lambda_{M_\eta}(\mathbf{\Theta}(t) - \mathbf{\Theta}^*(t))\boldsymbol{\omega}(t),\\
\dot{\mathbf{x}}_M(t) &= \mathbf{A}_M\mathbf{x}_M(t) + \mathbf{B}_P K_r r(t),\\
\dot{\mathbf{\Theta}}^T(t) &= -\mathbf{\Gamma}\boldsymbol{\omega}(t)\mathbf{e}_C^T(t)\mathbf{P}\mathbf{B}_P
\end{aligned} \tag{9.28}$$

The quantification of the attainable set in the tracking error subspace for $e_\alpha(t)$ and $e_q(t)$

represents an important figure of merit to quantify the deviation between the states of the plant and the reference model. Note that in the context of flight control law clearance typically the boundary of the set is of primary interest. If a point-wise estimate of this boundary is sufficient for a particular application it is reasonable to consider only those optimal control problems which minimize the distance to the points in the outer boundary of the grid. For an extended investigation which employs this idea the reader is referred to Ref. [110]. Here, for illustration purposes the reachable set estimation from Ref. [110] is reproduced for $\Gamma = 10$ at $t_f = 1\ s$ using the DFOG-method. The resulting set including the worst-case parameters corresponding to each grid point are depicted in Figs. 9.8-9.10. Note that the DFOG method for the estimation of reachable sets in the clearance framework is particularly appealing as it

- enables the estimation of the reachable set in a general setting (control constraints, state constraints, parameters, ...),
- allows for a parallel solution of all optimal control problems associated with each grid point, and
- mitigates the local nature for direct optimal control problems by considering the solution of a large number of problems.

Thus, it can be expected that the application of this methods represents be a valuable extension for the Cat III intrusive approach for optimal control based clearance.

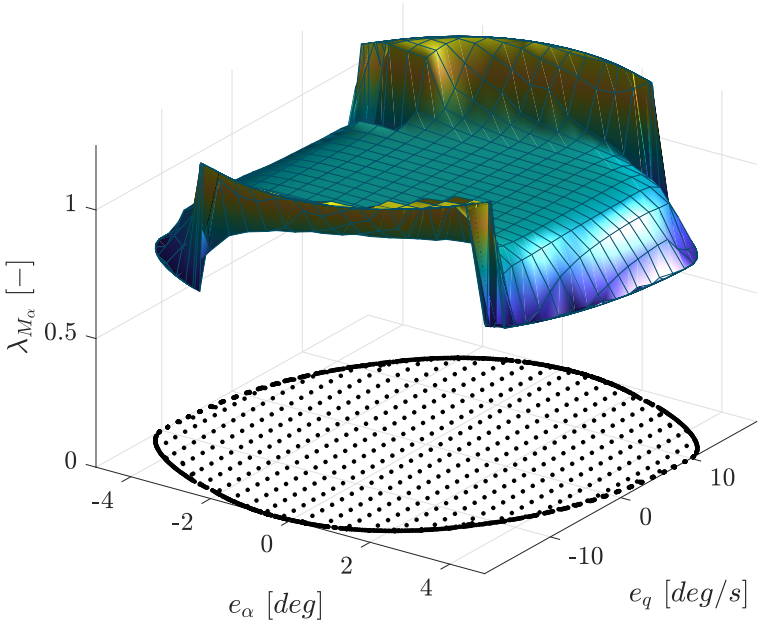


Figure 9.8: Reachable set approximation and optimal parameter values for λ_{M_α} corresponding to each grid point.

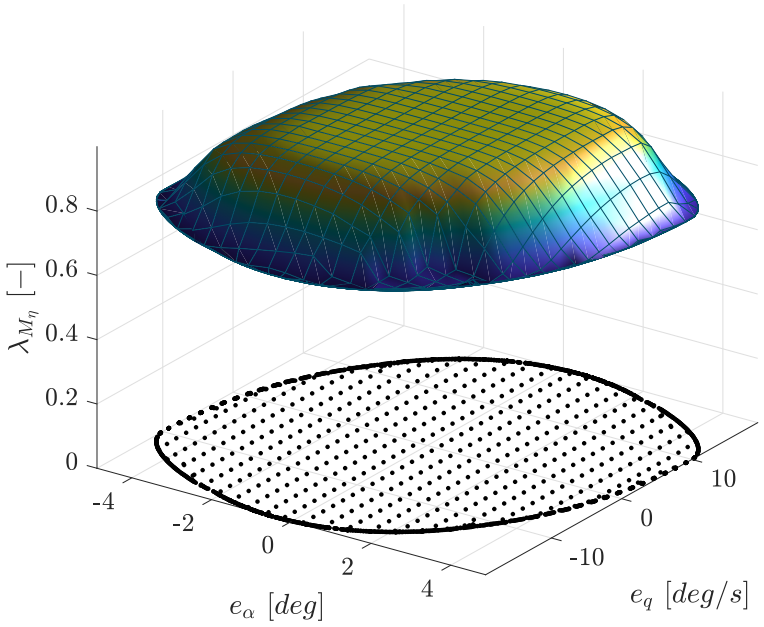


Figure 9.9: Reachable set approximation and optimal parameter values for λ_{M_η} corresponding to each grid point.

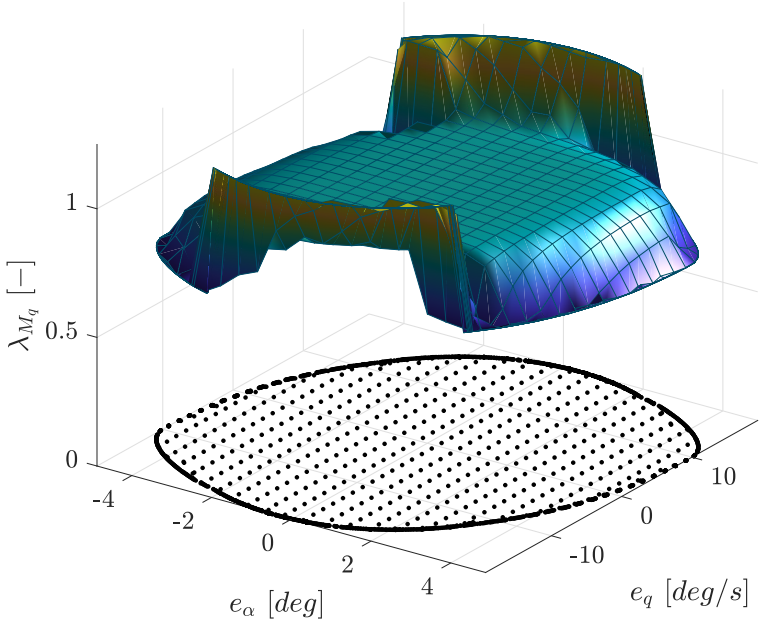


Figure 9.10: Reachable set approximation and optimal parameter values for λ_{M_q} corresponding to each grid point.

9.2 A Cat III Hybrid Approach for Practical Worst-case Analysis

The main motivation for the Cat III hybrid approach presented in the following is to introduce problem-specific information extracted from the solution of an intrusive method in the non-intrusive (black-box) approach. For this purpose a reduced model is employed which shares basic characteristics with the full model and allows for a solution of the clearance problem using a Cat I, Cat II, or Cat III type model under the optimal control based approach. This solution from the reduced model is used to determine the worst-case parameter combination and the structure of the worst-case control functions. The parametrization obtained in this way is then applied to test the full model under the Cat III non-intrusive method. Here, a simultaneous refinement of the worst-case control functions in combination with the parameters is possible or, to simplify the approach, the parameters may be fixed and solely the control function is refined. Note that an initial guess for the starting point of the black-box approach is readily available from the solution of the reduced model.

Obviously, the results obtained under this approach depend on the validity of the reduced model. If the information extracted from the solution of the reduced model is in-accurate the search space may not be parametrized appropriately and the initial guess is expected to be poor. As such, the two underlying assumptions for the application of this approach are:

- **Assumption I**: The reduced model shares the same tendency regarding the combination of worst-case parameter values and optimal control actions with the full model.
- **Assumption II**: The switching structure, e.g. the sequence of bang-bang arcs, of the worst-case control functions from the reduced model is representative for the full model.

Assumption I ensures that even if the results obtained from the reduced model do not match the full model quantitatively at least the reduced model provides an indicator regarding the direction in the parameter space where the worst-case parameter combination is expected. Under this assumption the reduced model may be used to search the parameter space in order to determine the parameter combination which leads to the most detrimental value of the criterion. At the worst-case parameter combination identified by this approach the optimal control functions can be analyzed and the structure parametrized. In case of a bang-bang type control this parametrization can be for example chosen as the switching time points with the amplitude being fixed to the respective bounds. If assumption II holds it can be expected that a fine-tuning of

this parametrized control function, e.g. by adapting the switching time points using a suitable optimization algorithm, improves the worst-case solution for the full model. Note that under this approach the number of parameters is greatly reduced compared to a search of the optimal control function "from scratch" and an initial guess is readily available from the reduced model. Note further that, alternatively, also a simultaneous refinement of the worst-case parameter combination and control parametrization obtained by the reduced model is possible under this approach.

In the following Cat I and Cat II models are used as reduced models. The Cat II model is preferred for the parameter search under the bi-level approach presented in Sec. 8.3.2 in order to determine the worst-case parameter combination. The reasoning behind this approach is that the degree of realism for the worst-case parameter combination is expected to be higher if actuator limits are included in the analysis. For determining a suitable control parametrization the Cat I model is then solved at the worst-case parameter combination. Cat I models are observed to represent an excellent choice for this purpose as the bang-bang commands seem to be the most hazardous type of input signals for the full model and a parametrization based on the switching time points is easily determined. Note that the full model is expected to handle constraints internally, e.g. by featuring an actuator model which takes saturations into account. Hence, if the full model handles the actuator limits internally the responsibility for not exceeding the limits of the servomechanism is transferred from the control to the dynamic model. It is observed that the number of switches for bang-bang type worst-case control functions is typically low. As such, the number of parameters (number of switching time points) is expected to be low if the amplitudes are fixed to the lower, respectively upper control bounds. This implies that also gradient-free optimization algorithms can be easily applied to refine the switching time points for the full model.

Let the switching time points identified from the reduced model be collected in the vector $\mathbf{t} \in \mathbb{R}^{n_t}$. A simple method to determine these switching time points is to find the zero-crossings of the switching functions in case of Cat I models. Computational experience suggests that this approach works very robustly even if the control function exhibits numerical inaccuracies (such as chattering effects). The evaluation of the criterion function $j : \mathbb{R}^{n_t} \to \mathbb{R}$ is then performed based on a simulation of the full model. As the values of the function $j(\mathbf{t})$ from the nonlinear model can be noisy due to the potentially complex structure of aircraft closed-loop systems a gradient-free optimization method is used in the following. In particular, the Nelder-Mead Simplex method (cf. Ref. [121, 122]) has shown to be effective for this task. This method represents a derivative-free optimization algorithm for the solution of unconstrained optimization problems which merely requires the values of the cost function at particular points in the parameter space. The name "Simplex"-method originates from the fact that the method performs a set of operations (reflection, expansion, contraction)

on a simplex body such that the "*simplex adapts itself to the local landscape, and contracts on to the final minimum*" [121]. Note that this method is not to be confused with Dantzig's Simplex method (cf. Ref. [55]) for solving LPs. An implementation is for example available in the MATLAB®-function *fminsearch()* designed for unconstrained minimization. Clearly, it is desirable for the control function parametrization using the switching time points to impose bounds of the form $0 \leq t_k \leq t_f, k = 0, \ldots, n_t - 1$. In order to include these simple bounds in the unconstrained algorithm from *fminsearch()* the version from Ref. [123] is used which applies simple variable transformations to handle box-bounds. Moreover, it is required that the sequence of the switching structure, meaning the sequence of bang-bang arcs, is maintained which implies $t_{i+1} \geq t_i$ (recall that the amplitudes are fixed to the lower or upper bounds). This can be ensured by sorting the vector $\mathbf{t}$ in ascending order before evaluating the cost function.

Regarding the cost-function it would be possible to use the infinity norm of the criterion under investigation over the whole time interval which is a common approach for optimization based clearance methods (see for example Ref. [21]). Here, this approach is not followed but the problem is stated as a Mayer-type problem regarding the value of the criterion at the final time point t_f. The reasoning behind this strategy is to avoid that the algorithm concentrates on some extremal point (e.g. right in the beginning of the time interval) which may not exploit the full potential of the control function parametrization over the whole time interval. Consider for example a control function with two switches (see Fig. 9.11). If in some iteration the two switches are not

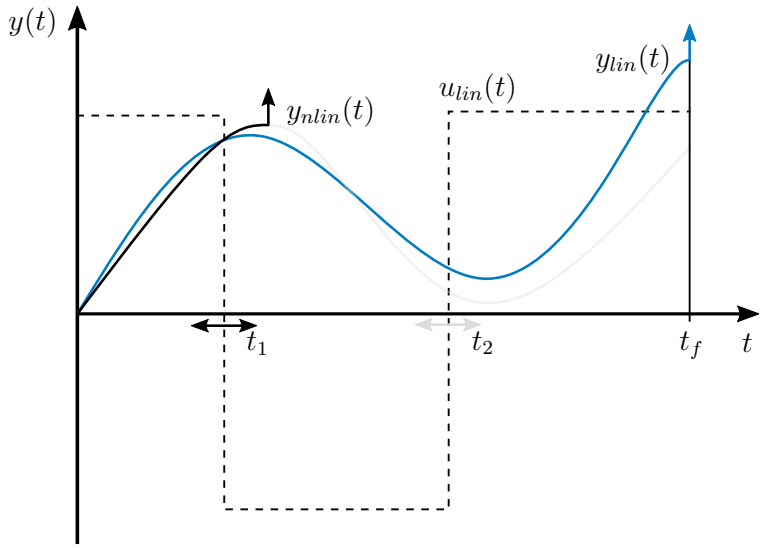


Figure 9.11: Linear response ($y_{lin}(t)$, blue line) and nonlinear response ($y_{nlin}(t)$, black line / grey line) to the control input ($u_{lin}(t)$, black dashed line) parametrized by the time points t_1 and t_2. Components which are greyed out indicate that the maximum value of the nonlinear response is not influenced by these quantities due to the causality of the system.

well tuned the maximum may occur before the second switch takes place. If the algorithm focuses on the highest value of the cost function history the second parameter is essentially disregarded due to the causality of the system (future control actions do not affect the past). For example, when using a gradient-based method the gradient of the maximum value with respect to the second parameter (switching time point) would be zero in this case. However, by maximizing the criterion value at the final time point the algorithm is "motivated" to use the full parameter space in all iterates as all parameters should have an effect on the final value of the criterion (past control actions do affect the future). Moreover, the algorithm concentrates on the maximization of the criterion on a single, defined point and a jumping between extremal points or even cycling between several extremals at different locations is avoided.

Both the longitudinal and the lateral models are used to illustrate the Cat III hybrid approach. First, the nominal cases, i.e. without considering parameter uncertainties, are presented and the Cat I solutions using the combined controls from Sec. 8.2.2 are parametrized based on the switching time points. Figs. 9.12 and 9.13 present the results for the maximization of $n_z(t_f)$ and $\alpha_A(t_f)$ using the normal load factor command $n_{z,c}(t)$ including the wind commands $u_W(t)$ as well as $w_W(t)$. Similarly, Fig. 9.14 and Fig. 9.15 present the results for the maximization of $n_y(t_f)$ and $\beta_A(t_f)$ using the lateral load factor command $n_{y,c}(t)$ including the wind command $v_W(t)$. In all results (top sub-figures) the initial response in the first iteration (blue lines) is compared to the worst-case solution identified by the optimization algorithm (black lines). Note that the results in the initial iteration indicate that merely applying the commands obtained from the Cat I models is not effective for the final values of the criteria as the switching time points are not tuned to the nonlinear model. Only for the maximization of $n_y(t_f)$ the initial and final value at the terminal time point appear to be similar (cf. Fig. 9.14). However, for all cases the optimization algorithm behaves as expected and adjusts the switching time points in order to maximize the criteria at the terminal time point. Observe that in some cases the optimizer chooses to collapse some of the switching time intervals for the nonlinear model (see for example Fig. 9.12, Fig. 9.16, and Fig. 9.17).

Next, the parameter dependent cases are investigated. Here, the results from Tabs. 8.1-8.4 (SQP) are used as starting points and the optimal control functions at the worst-case parameter combinations are first determined using a Cat I model and then parametrized using the zero-crossings of the switching functions. Note that as for the nominal case simply applying the worst-case control from the reduced model is not effective. However, through the refinement using the Nelder-Mead Simplex algorithm the results are considerably improved. Observe that for the maximization of the load factors $n_z(t_f)$ and $n_y(t_f)$ the identified worst-case solutions exceed the maximum commanded values by a noticeable amount.

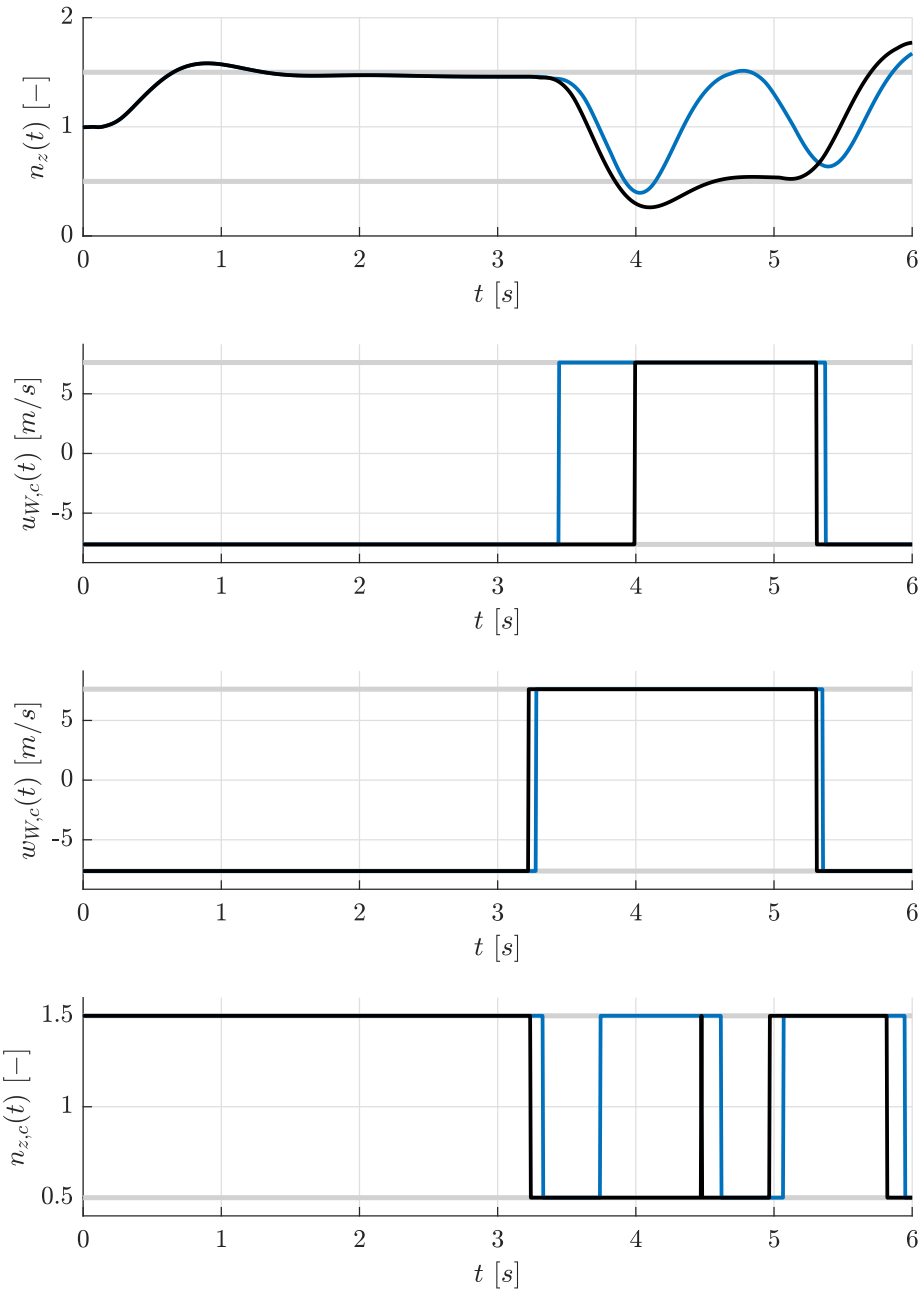


Figure 9.12: Optimal results for the longitudinal plane using the Cat III hybrid approach (nominal case) for maximizing $n_z(t_f), t_f = 6\ s$ with wind including the normal load factor $n_z(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, and the normal load factor command $n_{z,c}(t)$. The blue lines represent the initial control and response histories whereas the black lines are the final results after the optimization of the switching time points requiring 269 simulations.

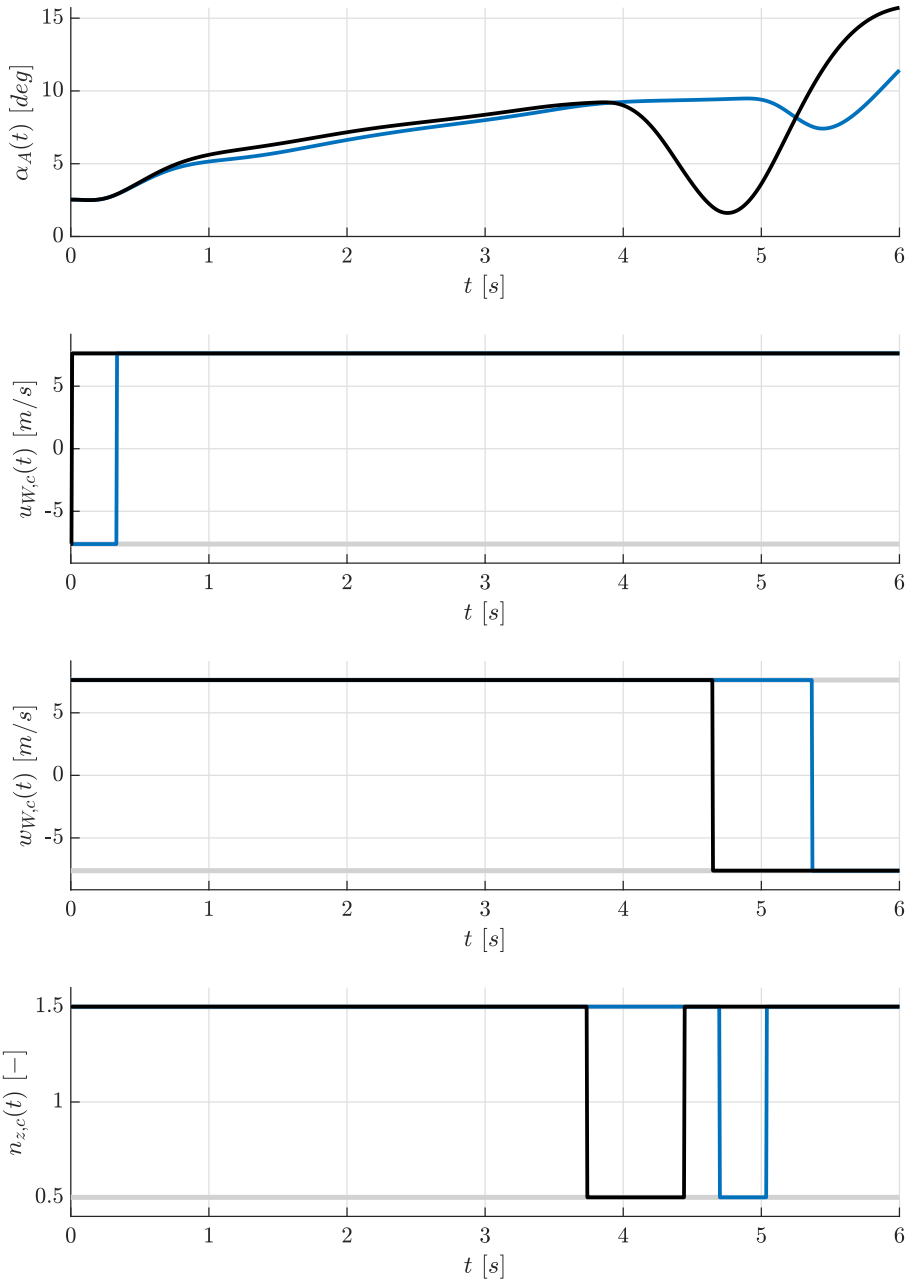


Figure 9.13: Optimal results for the longitudinal plane using the Cat III hybrid approach (nominal case) for maximizing $\alpha_A(t_f), t_f = 6\ s$ with wind including the AoA $\alpha_A(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, and the normal load factor command $n_{z,c}(t)$. The blue lines represent the initial control and response histories whereas the black lines are the final results after the optimization of the switching time points requiring 223 simulations

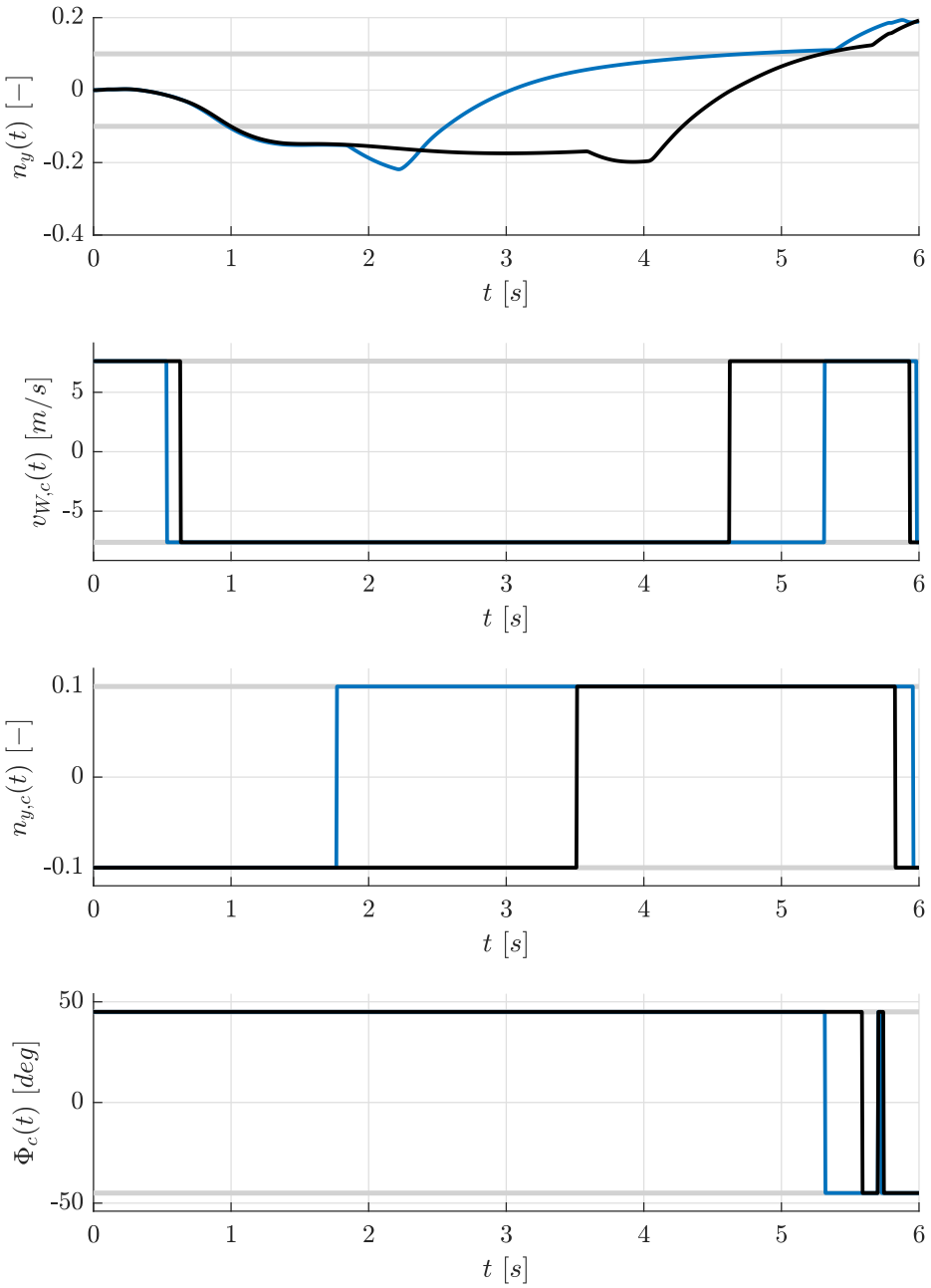


Figure 9.14: Optimal results for the lateral plane using the Cat III hybrid approach (nominal case) for maximizing $n_y(t_f), t_f = 6\ s$ with wind including the lateral load factor $n_y(t)$, the lateral wind velocity command $v_{W,c}(t)$, the lateral load factor command $n_{y,c}(t)$, and the roll angle command $\Phi_c(t)$. The blue lines represent the initial control and response histories whereas the black lines are the final results after the optimization of the switching time points requiring 338 simulations.

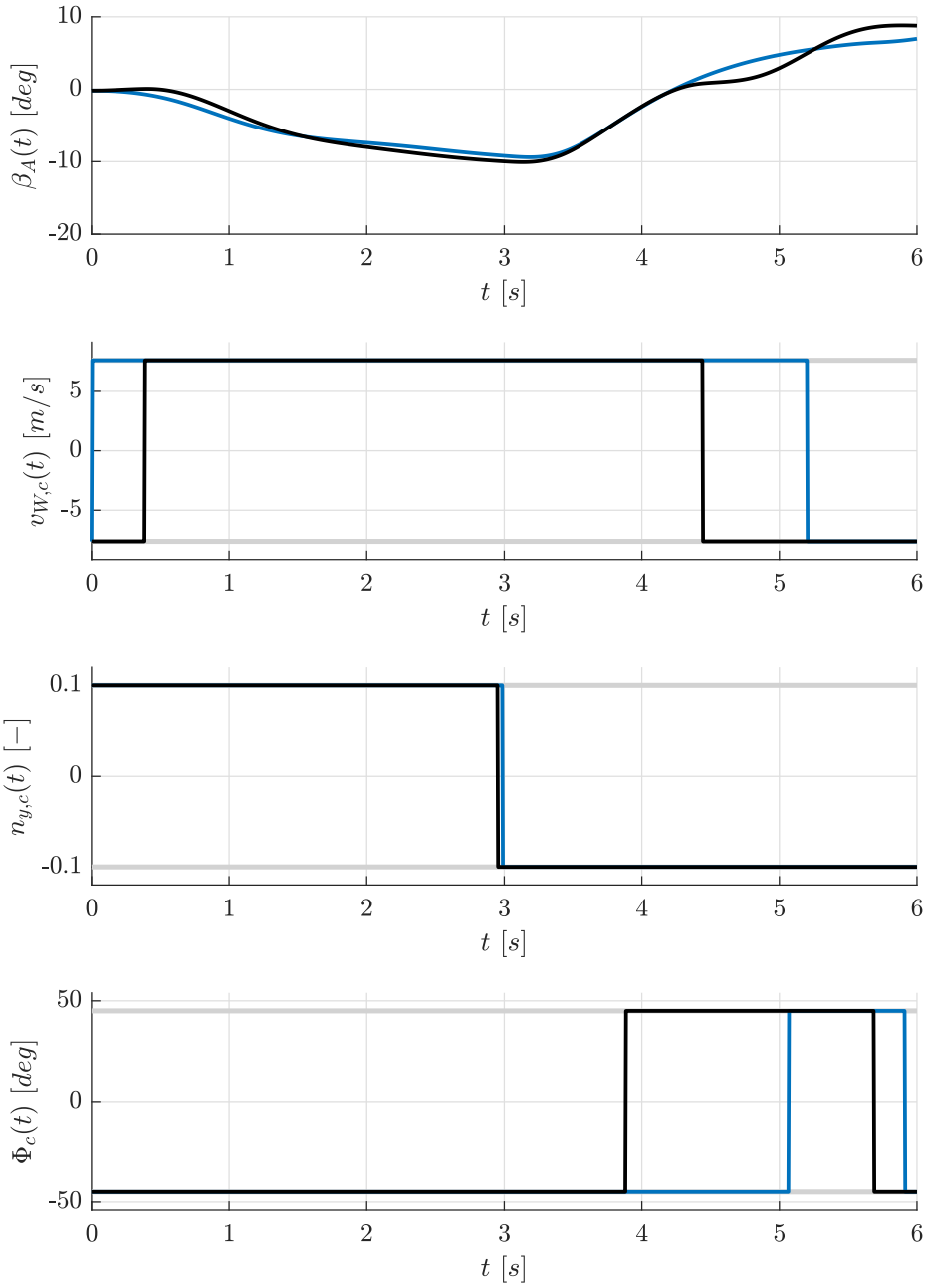


Figure 9.15: Optimal results for the lateral plane using the Cat III hybrid approach (nominal case) for maximizing $\beta_A(t_f), t_f = 6\ s$ with wind including the AoS $\beta_A(t)$, the lateral wind velocity command $v_{W,c}(t)$, the lateral load factor command $n_{y,c}(t)$, and the roll angle command $\Phi_c(t)$. The blue lines represent the initial control and response histories whereas the black lines are the final results after the optimization of the switching time points requiring 213 simulations.

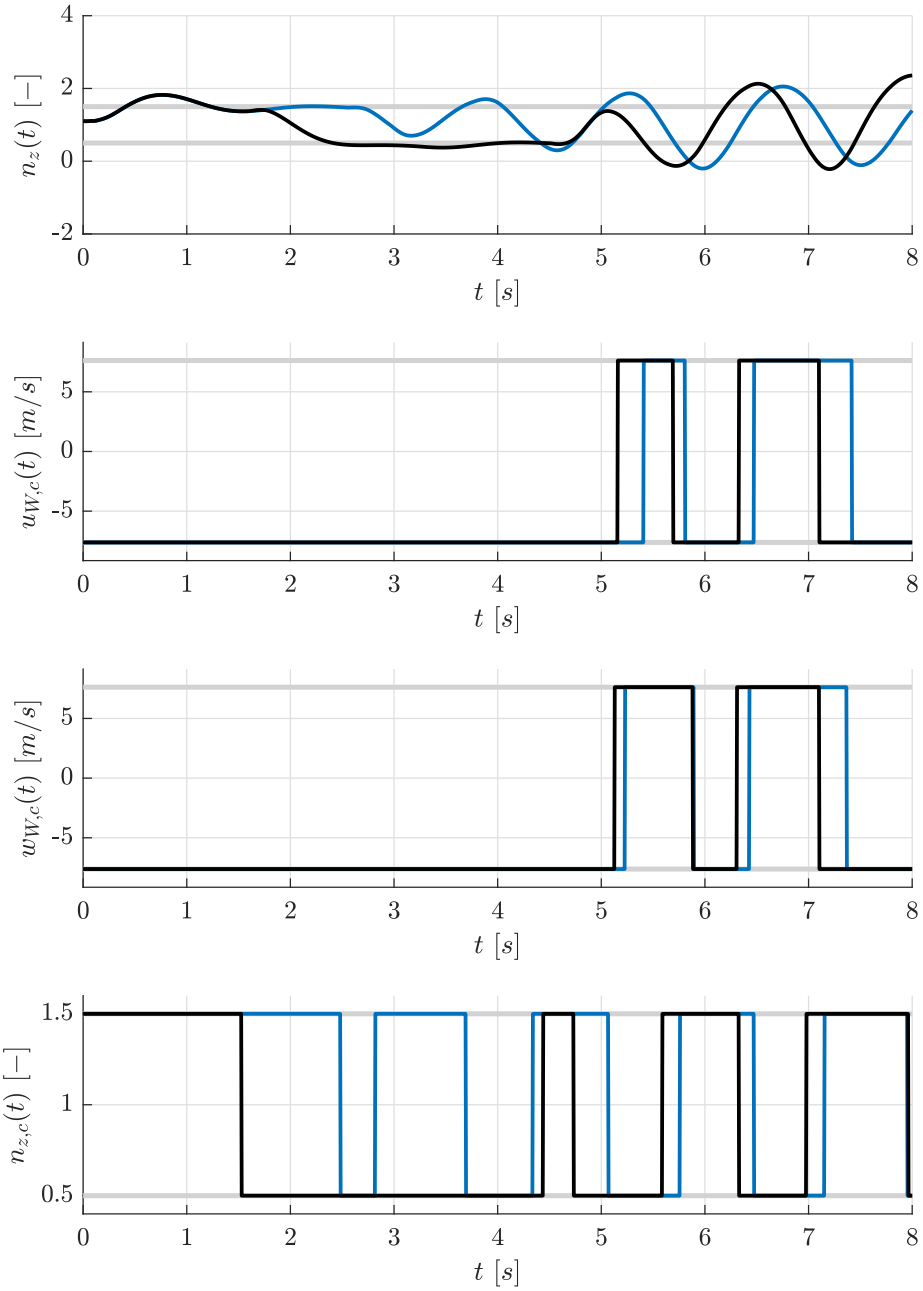


Figure 9.16: Optimal results for the longitudinal plane using the Cat III hybrid approach (parameter dependent case) for maximizing $n_z(t_f), t_f = 8\ s$ with wind including the normal load factor $n_z(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, and the normal load factor command $n_{z,c}(t)$. The blue lines represent the initial control and response histories whereas the black lines are the final results after the optimization of the switching time points requiring 657 simulations.

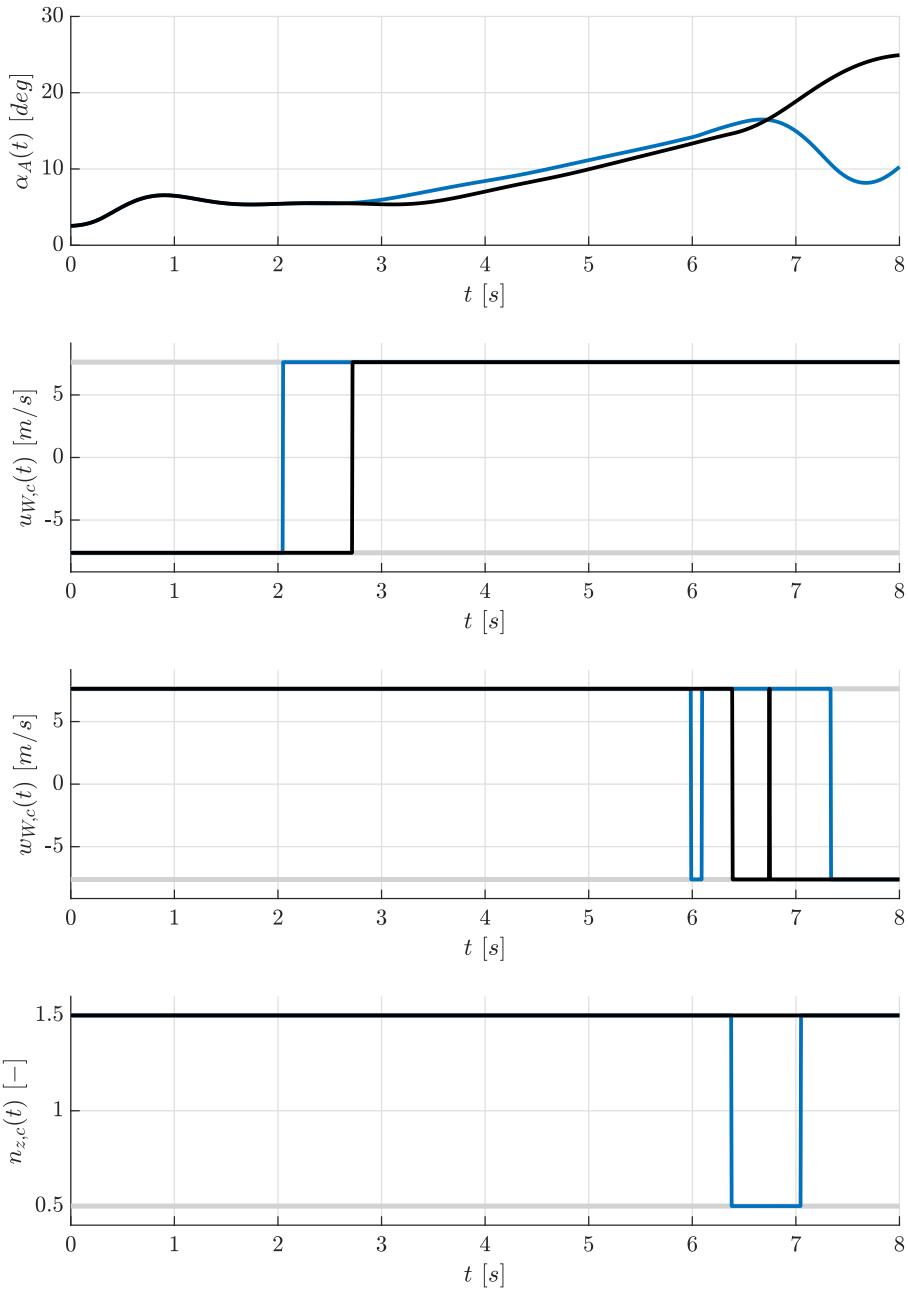


Figure 9.17: Optimal results for the longitudinal plane using the Cat III hybrid approach (parameter dependent case) for maximizing $\alpha_A(t_f), t_f = 8\ s$ with wind including the AoA $\alpha_A(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, and the normal load factor command $n_{z,c}(t)$. The blue lines represent the initial control and response histories whereas the black lines are the final results after the optimization of the switching time points requiring 174 simulations.

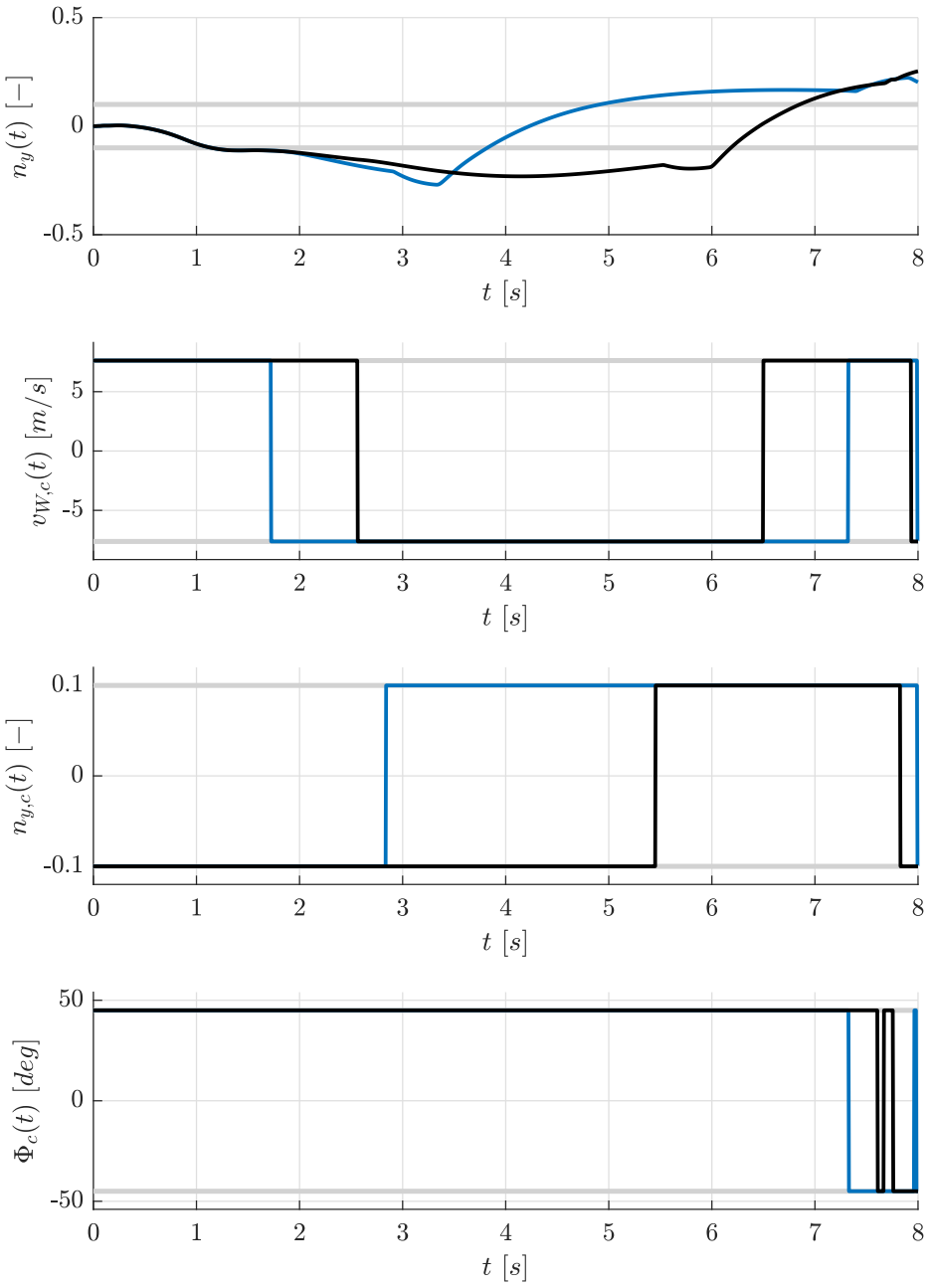


Figure 9.18: Optimal results for the lateral plane using the Cat III hybrid approach (parameter dependent case) for maximizing $n_y(t_f), t_f = 8\ s$ with wind including the lateral load factor $n_y(t)$, the lateral wind velocity command $v_{W,c}(t)$, the lateral load factor command $n_{y,c}(t)$, and the roll angle command $\Phi_c(t)$. The blue lines represent the initial control and response histories whereas the black lines are the final results after the optimization of the switching time points requiring 321 simulations.

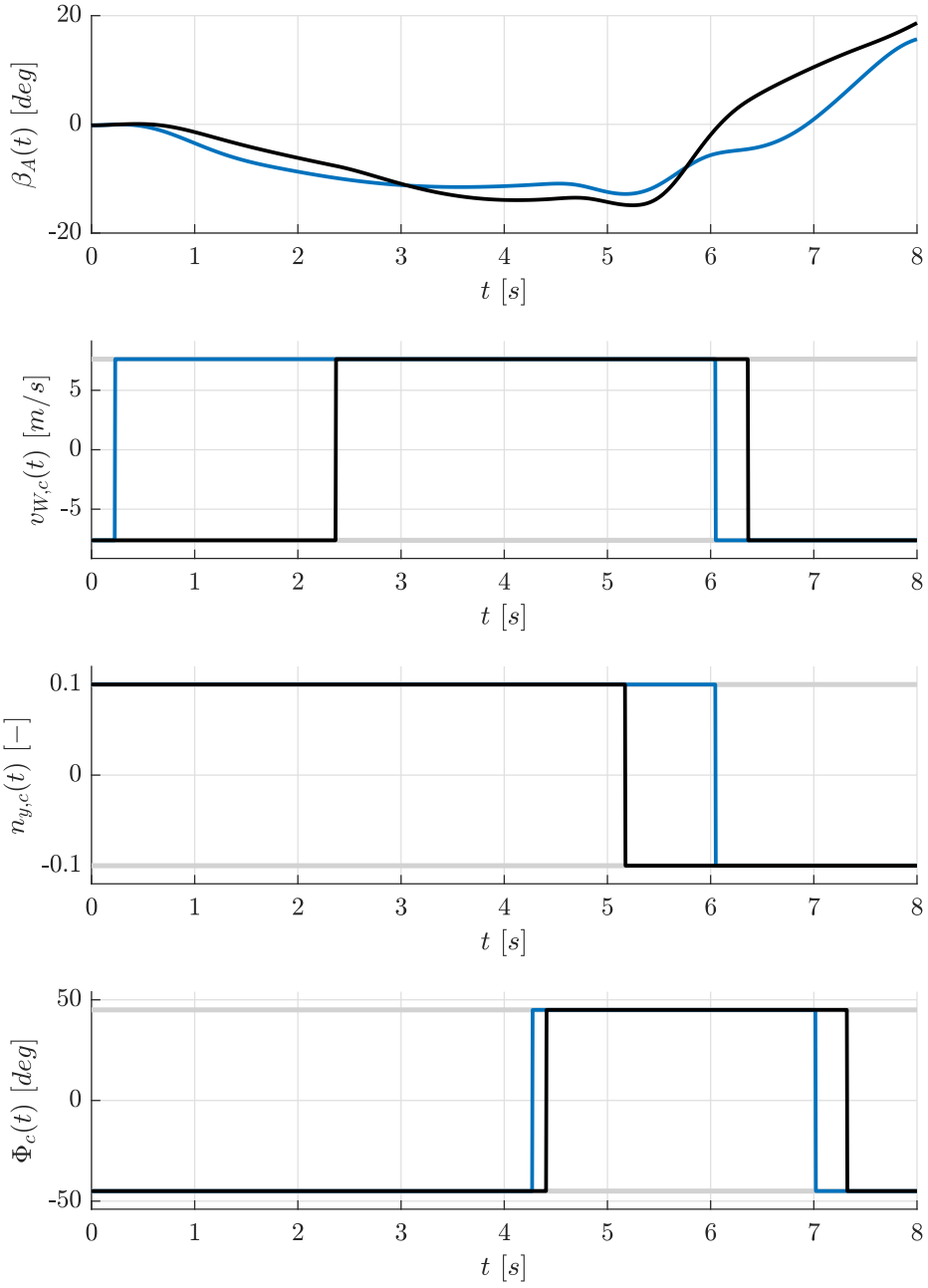


Figure 9.19: Optimal results for the lateral plane using the Cat III hybrid approach (parameter dependent case) for maximizing $\beta_A(t_f), t_f = 8\ s$ with wind including the AoS $\beta_A(t)$, the lateral wind velocity command $v_{W,c}(t)$, the lateral load factor command $n_{y,c}(t)$, and the roll angle command $\Phi_c(t)$. The blue lines represent the initial control and response histories whereas the black lines are the final results after the optimization of the switching time points requiring 187 simulations.

Regarding the maximization of $n_z(t_f)$ for the parameter dependent case (see Fig. 9.16) it appears as if the controls are exciting an oscillatory mode. Here, it is worthwhile to investigate if increasing the final time of the analysis is expected to yield a higher value of the criterion. A comparison of the parametrizations obtained from the solution using the Cat I model at the worst-case parameter combinations are presented in Figs. 9.20-9.22 for $t_f \in \{8\ s, 12\ s, 16\ s\}$. Note that the structure of the solution in the final $6\ s$ remains essentially unchanged and merely the longitudinal wind command seems to produce an additional switch in the beginning of the time interval for the cases with $t_f = 12\ s$ and $t_f = 16\ s$. This indicates that increasing the time interval is not expected to yield a significantly higher value by refining the switches using a non-intrusive method.

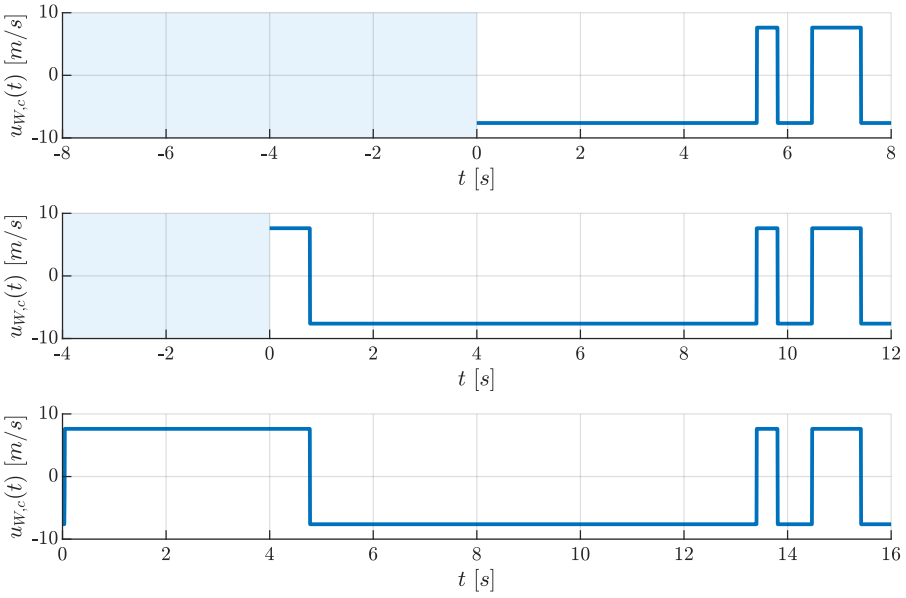


Figure 9.20: Comparison of the parametrized control functions for $u_{W,c}(t)$ obtained from the solution using a Cat I model at the worst-case parameter combinations for $t_f = 8\ s$ (top), $t_f = 12\ s$ (middle), and $t_f = 16\ s$ (bottom).

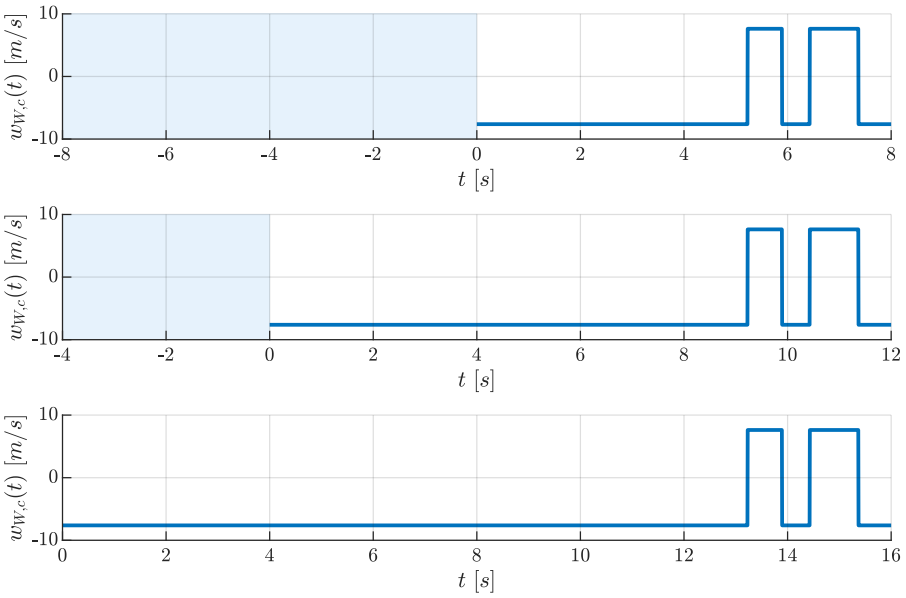


Figure 9.21: Comparison of the parametrized control functions for $w_{W,c}(t)$ obtained from the solution using a Cat I model at the worst-case parameter combinations for $t_f = 8$ s (top), $t_f = 12$ s (middle), and $t_f = 16$ s (bottom).

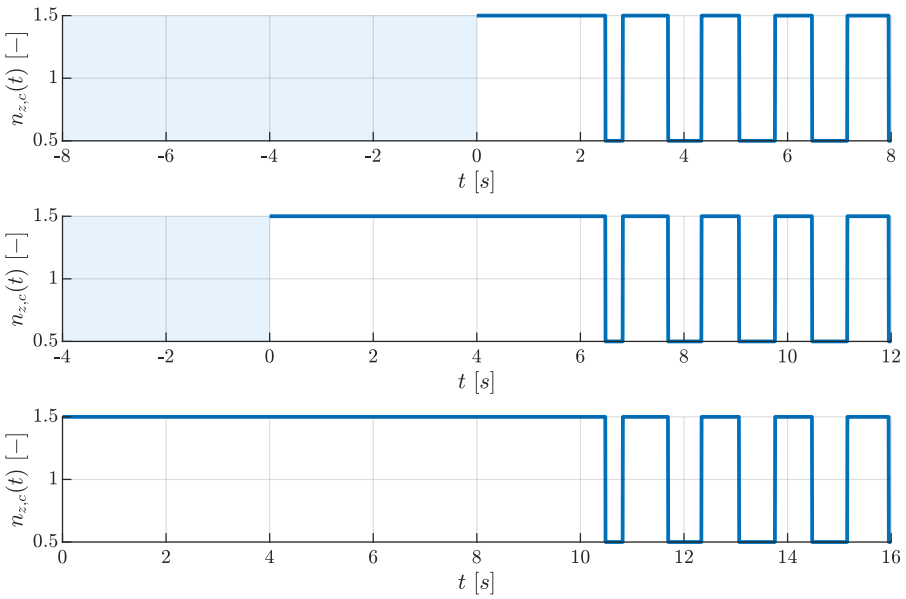


Figure 9.22: Comparison of the parametrized control functions for $n_{z,c}(t)$ obtained from the solution using a Cat I model at the worst-case parameter combinations for $t_f = 8$ s (top), $t_f = 12$ s (middle), and $t_f = 16$ s (bottom).

It is clear that the application of sanity checks is advisable in order to improve the confidence in the worst-case results obtained from the Cat III hybrid approach. The following checks may be performed to check if the assumptions on the reduced model are valid and if the refinement using the optimization algorithm under the non-intrusive approach converged to a meaningful value:

- **Check I**: Investigate the flatness of the slope for the criterion at the final time.
- **Check II**: Check if the most detrimental value of the criterion occurs at the final time and not in the interior of the time interval.
- **Check III**: Check how the optimization algorithm under the non-intrusive approach exploits the parametrization obtained from the intrusive approach (e.g. check if intervals are collapsed after the refinement).
- **Check IV**: Apply different optimization algorithms for the refinement of the solution under the non-intrusive approach and check the consistency of the results.
- **Check V**: Perform the analysis for increasing values of the final time t_f (e.g. in multiples of a characteristic time constant such as the one corresponding to the closed-loop short period mode) and check if the results obtained from the analyses yield consistent worst-case solutions.

If the above checks indicate that the solution is not satisfactory the process may be repeated for a new final time, the analysis may be re-started from a randomized initial parametrization around the nominal values obtained from the non-intrusive approach or a different optimization algorithm can be employed (e.g. the DIRECT method or the surrogate optimization method, cf. Sec. 8.3.2). These strategies may also be applied for some of the cases presented in this section to further improve the worst-case solutions.

However, the examples presented in this section indicate that the Cat III hybrid approach is effective for the fully nonlinear model and can be easily integrated in existing tool-chains. Merely, linearized models and a simulation model are required which are essentially available at no cost - typically these models are required anyway for the development of flight control laws. Thus, the Cat III hybrid approach should be directly applicable for the solution of clearance problems in industry practice. In particular, it is observed that the Cat III hybrid approach allows for a considerable automatization. On the one side, the solution using Cat I and Cat II models is reliable and efficient. On the other side, it is found that the automatic detection of the switching time points using the values of the switching functions works very robustly. These aspects, in particular, allow for the application of this approach for example in the form of an automatic sanity check whenever the model is altered (such as in a continuous integration environment).

Chapter 10

Summary, Conclusion, and Outlook

10.1 Summary and Conclusion

In this thesis the application of optimal control methods for flight control law clearance is investigated. The optimal control formulation considers bounded inputs as well as state constraints which are introduced to expose internal limits (actuator rate and deflection) to the optimization algorithm. For all categories from the proposed modeling classification scheme (cf. Sec. 7.1, Contribution I: Model classifications for optimal control based clearance) approaches tailored to linear (Cat I), quasi-linear (Cat II), and general nonlinear (Cat III) models are developed. These approaches consider the solution of the clearance problem in the nominal and parameter dependent cases.

Contribution II: Development of clearance methods for parameter dependent (quasi-)linear systems

For Cat I and Cat II type models the optimal control is characterized based on the Minimum Principle (cf. Sec. 8.1.1 and Sec. 8.1.2). Under the assumption that the system is controllable it is shown that the worst-case control is of bang-bang type if the states are unbounded and potentially exhibits singular intervals in the bounded case. On these singular intervals the control takes the values of the boundary control which allows the system to ride the state constrained arcs. Regarding the numerical solution using Cat I and Cat II type models it is shown that based on direct transcription methods the continuous time optimal control problem can be transcribed into a finite dimensional Linear Program (LP), cf. Sec. 8.1.3. This implies that the global worst-case can be efficiently determined using LP solvers even under the consideration of actuator rate and deflection limits. A benchmark model is used to illustrate the application of this approach to a realistic aircraft closed-loop model. This model is linearized around a steady-state flight condition and the longitudinal and lateral channels are investigated

separately (cf. Sec. 8.2.1). Four different criteria namely the load factors in normal and lateral direction as well as the aerodynamic Angle-of-Attack and Angle-of-Sideslip are considered for testing the flight control law. Besides the commands of the closed-loop model (normal and lateral load factor commands, roll angle command) additional artificial control-like variables are used to model the influence of wind. The numerical solutions show that the approach is effective not only for the single input case but also when considering combined control actions. Moreover, the numerical results appear to be in alignment with the theoretical analysis (cf. Sec. 8.2.2).

Regarding the parameter dependent case for Cat I and Cat II models (cf. Sec. 8.3) a decomposition approach is proposed which separates the search in the parameter space from the solution of the optimal control problem. Under this approach a bi-level optimization setup is used where the upper level searches the parameter space and is subject to the lower level which solves for the worst-case control functions (cf. Sec. 8.3.2). For most practical applications the number of parameters in the upper level is expected to be low which implies that global optimization methods can be efficiently employed. Note that the lower level solves for the worst-case control functions through the solution of LP problems at nominal parameter values. As such, the lower level can be evaluated efficiently and is always guaranteed to yield the global worst-case for the particular parameter values provided from the upper level. Besides the use of global optimization algorithms in the upper level, gradient-based methods are shown to be efficient as convergence is achieved within a few steps. Moreover, the illustrative examples demonstrate that gradient-based methods yield worst-case solutions close to the ones obtained from global methods (at a considerably lower computational cost). For the application of gradient-based methods it is shown that first- and second-order derivative information of the cost function in the lower level with respect to parameters in the upper level can be obtained efficiently using post-optimal sensitivity analysis. It is noteworthy that this kind of sensitivity analysis can also be seen as a valuable extension for the clearance framework developed in this thesis in order to determine influential parameters regarding the worst-case solution at nominal parameter values. Here, it should be mentioned that essentially arbitrary parameters in the aircraft closed-loop system can be considered for this type of sensitivity analysis.

Besides the proposed worst-case search algorithm for parameter dependent Cat I and Cat II models the case is investigated where the parameters are uncertain, i.e. subject to a distribution (cf. Sec. 8.3.3). Uncertainty quantification is performed under the optimal control based clearance approach using generalized Polynomial Chaos (gPC) and Monte Carlo (MC) analysis. In the numerical examples criteria for the longitudinal and lateral channel are investigated including uncertainties corresponding the the aerodynamic coefficients. The gPC and MC approaches are used to estimate the first two moments of the criteria. It is shown that the results regarding both methods are in

good alignment and that the gPC approach can be used to efficiently estimate the moments and the response surface for the criterion under investigation. Moreover, these results can be used to determine bounds on the exceedance probability for a particular criterion. It is noteworthy that both approaches allow for a considerable parallelization as all LPs corresponding to the sampling points can be solved independently.

For the investigation of multi-criteria problems using Cat I and Cat II type models reachability analysis is considered (cf. Sec. 8.4). An algorithm is developed based on the successive solution of LPs along rays in the output space in combination with post-optimal sensitivity analysis. This algorithm constructs inner and outer approximations and adaptively refines these approximations until a given accuracy is achieved. In particular, this algorithm is able to handle input and state constraints and provides a bound on the expected error in each major iteration. Two- and three-dimensional examples for different criteria regarding the longitudinal and lateral plane are presented to illustrate the approach.

Contribution III: Development of clearance methods for general nonlinear systems

Cat III type models for optimal control based clearance approaches have shown to be the most challenging type regarding the numerical solution. Three sub-classes are presented for this category namely non-intrusive (black-box), intrusive (white-box), and hybrid (grey-box) approaches (cf. chapter 9). For the intrusive approach a nonlinear optimal control problem is formulated which simultaneously determines the worst-case parameters in combination with the worst-case control functions. It is shown that given the current state of the art regarding direct optimal control methods, solutions of the discretized problems can be obtained using nonlinear programming solvers tailored to large-scale and sparse problems. However, it is also found that the solution of the nonlinear optimal control problems using the direct approach are considerably more challenging compared to the linear case. Furthermore, the results indicate that the solution structures obtained from Cat II type models share similarities with the solutions obtained from the general Cat III type model. In particular, the worst-case parameter combination obtained under the simultaneous optimization of parameters and control functions yields almost identical results (cf. Sec. 9.1.3).

Regarding multi-criteria analysis for Cat III type models the application of the Distance Fields on Grids (DFOG) method is proposed (cf. Sec. 9.1.4). This method is based on a particular optimal control formulation for which a distance function is used to obtain a point-wise representation of the reachable set in the space of criteria. An inherent characteristic of this method is that the solution of a potentially large number of optimal control problems is required. Even though this property obviously represents

a challenge regarding the efficient solution of the associated optimal control problems it can in fact be seen as a good practical strategy to globalize the approximation. Moreover, the DFOG method allows for a considerable parallelization as all optimal control problems which need to be solved are independent. The method is illustrated using a F-16 model with nonlinear model reference adaptive controller (MRAC). For this example the forward reachable set in the error subspace is estimated including the consideration of parametric uncertainties.

Finally, the Cat III hybrid approach is presented which employs a reduced model to determine the worst-case parameter combination and the worst-case inputs (cf. Sec. 9.2). This result is then used to parametrize the control functions and to refine the solution using a Cat III non-intrusive (black-box) method. Here, it is found that the maximization of the criterion at the final time point represents an effective way to formulate the clearance problem. Moreover, the refinement of the control functions based on the Nelder-Mead Simplex method is shown to be effective for this task. In all cases the final worst-case control functions identified by this optimization algorithm appear to increase the value of the criterion at the final time point compared to the initial guess provided from the reduced model. In particular, for the maximization of the normal and lateral load factors it is shown that the response of the system considerably exceeds the commanded values. From a practical perspective the Cat III hybrid approach appears to be a promising candidate for the direct integration into state of the art clearance procedures. These state of the art methods in optimization based clearance are currently primarily based on black-box methods using global optimization algorithms (see Ref. [21]). Note that the efficiency and effectiveness of these algorithms under the Cat III hybrid approach is leveraged by the optimal control based methods. As such, this approach allows for a gradual integration of optimal control based methods in optimization based clearance frameworks (cf. Fig. 10.1).

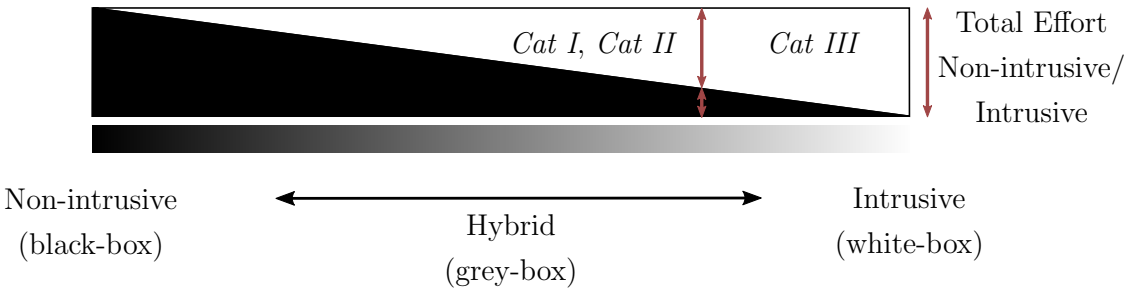


Figure 10.1: Combination of state of the art (black-box) methods with the optimal control based approaches developed in this thesis. The work share regarding the optimization/optimal control based solution of clearance problems between the non-intrusive (back-box) methods and the intrusive (white box) methods is illustrated by the black and white triangles.

Due to the fact that computational optimal control methods are currently evolving rapidly it is expected that besides the use of Cat I and Cat II models also nonlinear models may be used for this task in the near future. Obviously, the closer the reduced model gets to the full model the less work is required for the final refinement of the solution using a non-intrusive (black-box) approach.

Contribution IV: Software development for clearance methods

All approaches presented in this thesis regarding the Cat I and Cat II models as well as the Cat III hybrid approach are implemented by the author in a Counter Optimization Library (COLIBRY) which has a defined interface for the closed-loop model under investigation. From a practical perspective, the user merely needs to provide a trim, linearization, and simulation method in order to apply the approaches. For the Cat III intrusive approaches the optimal control toolbox *FALCON.m* has been co-developed by the author and applied for the solution of nonlinear clearance problems. Note that the modeling effort for the application of this approach is higher compared to the first two categories as the model needs to be reduced first and converted into a model in state-space form which which fulfills the model requirements in *FALCON.m*.

However, the Simulink® transcription tool developed by the author appears to be very helpful in reducing the workload for the conversion task. Regarding the computational efficiency it is important to mention that care was taken in the selection and design of the methods to allow for a considerable parallelization. This is desirable considering the fact that current trends in computing architectures are focused on parallel computing, e.g. by increasing the number of threads which can be executed simultaneously on multi-core platforms. Hence, it is expected that this development increases the efficiency of the proposed methods in the near future.

As suggested in Ref. [21] the application of optimization based approaches is not meant to completely replace the existing methods which are currently employed for flight control law clearance. The same holds for the optimal control based methods. However, the effectiveness and efficiency of the developed approaches indicate that an introduction of these methods in industry practice can help to add an additional safety layer to the clearance task. In particular, the application of the proposed methods in form of sanity checks during the development process are expected to present a valuable extension of the available tools in order to increase the confidence in the safety assessment of flight control laws.

10.2 Outlook

In this thesis numerical solutions for nonlinear clearance problems under the Cat III intrusive approach were obtained based on the direct optimal control approach. Besides direct methods, the application of Dynamic Programming [124] seems appealing as a global worst-case is obtained under these methods. Unfortunately, as discussed in chapter 9, the computational requirements are very restrictive regarding the number of state dimensions which can be considered under this approach. Nevertheless, there also exist approximate DP methods [124] which can be applied in higher dimensional spaces. It would be worthwhile to benchmark the efficiency and effectiveness of these approaches compared to the direct optimal control approach under the proposed Cat III hybrid method. However, regardless of the optimal control method employed for the clearance task it is important to mention that models usually need to be available in state-space form. In this work the nonlinear simulation model was reduced manually and then transcribed by the Simulink® transcription tool into a form which can be used by *FALCON.m*. Obviously, a suitable modeling framework which further automatizes this process would greatly facilitate the application of Cat III intrusive methods. It is clear, that for such a framework additional effort during the creation of the simulation model would be required, e.g. to tag certain subsystems in the dynamic model which need to be removed or replaced using an appropriate replacement library. Nevertheless, in the opinion of the author such an integrated modeling approach appears to be technically feasible and would help to significantly decrease the amount of manual work for the application of nonlinear optimal control based clearance methods.

Appendix A

Results Numerical Experiments (Linear)

This appendix chapter contains additional numerical results from Sec. 8.2.2.

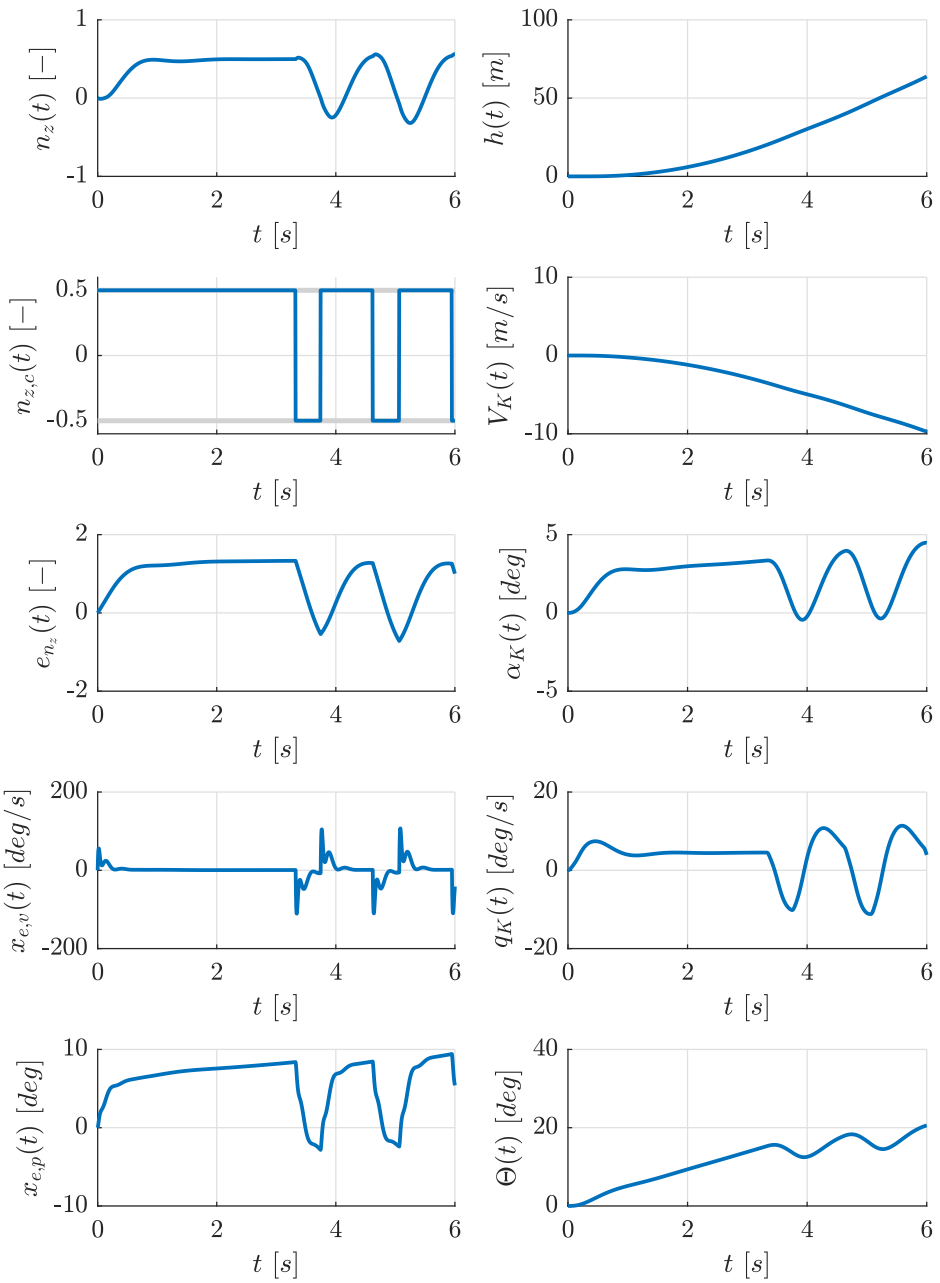


Figure A.1: Optimal results for the longitudinal plane using a Cat I model for maximizing $n_z(t_f), t_f = 6\ s$ including the normal load factor $n_z(t)$, the geodetic height $h(t)$, the normal load factor command $n_{z,c}(t)$, the kinematic velocity $V_K(t)$, the error controller state $e_{n_z}(t)$, the kinematic AoA $\alpha_K(t)$, the elevator rate $x_{e,v}(t)$, the pitch rate $q_K(t)$, the elevator position $x_{e,p}(t)$, and the pitch angle $\Theta(t)$.

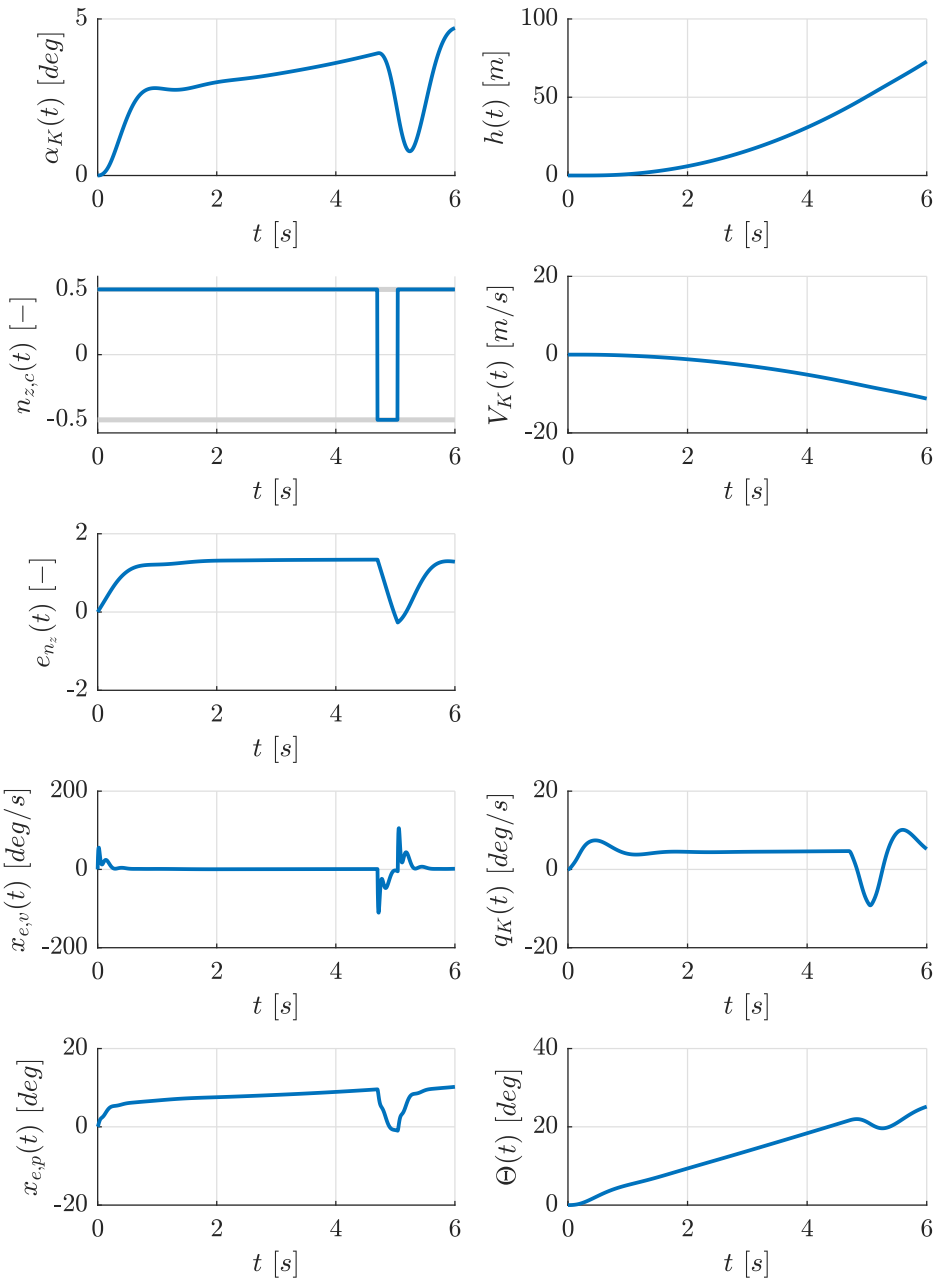


Figure A.2: Optimal results for the longitudinal plane using a Cat I model for maximizing $\alpha_K(t_f), t_f = 6\ s$ including the kinematic AoA $\alpha_K(t)$, the geodetic height $h(t)$, the normal load factor command $n_{z,c}(t)$, the kinematic velocity $V_K(t)$, the error controller state $e_{n_z}(t)$, the elevator rate $x_{e,v}(t)$, the pitch rate $q_K(t)$, the elevator position $x_{e,p}(t)$, and the pitch angle $\Theta(t)$.

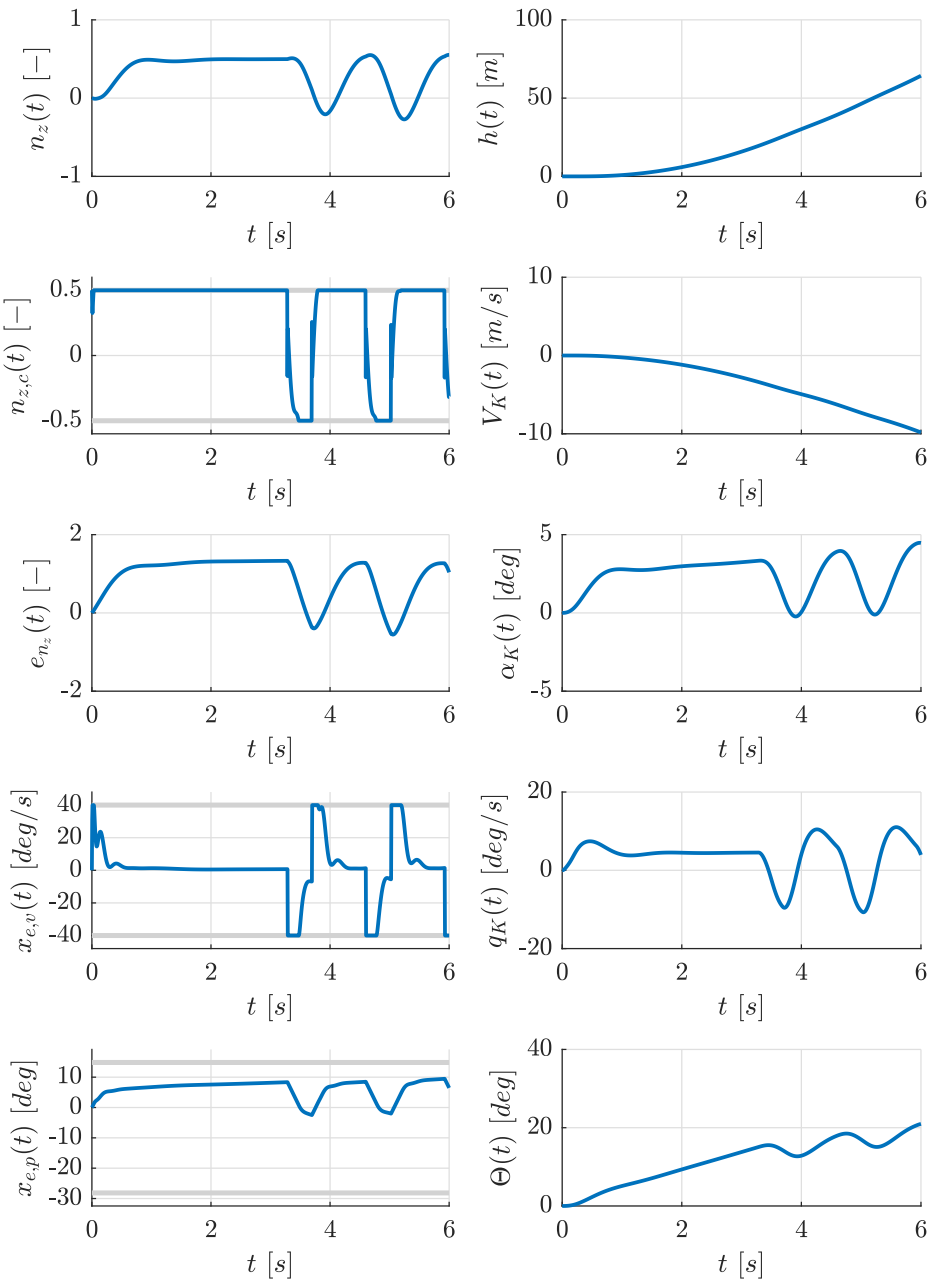


Figure A.3: Optimal results for the longitudinal plane using a Cat II model for maximizing $n_z(t_f), t_f = 6\ s$ including the normal load factor $n_z(t)$, the geodetic height $h(t)$, the normal load factor command $n_{z,c}(t)$, the kinematic velocity $V_K(t)$, the error controller state $e_{n_z}(t)$, the kinematic AoA $\alpha_K(t)$, the elevator rate $x_{e,v}(t)$, the pitch rate $q_K(t)$, the elevator position $x_{e,p}(t)$, and the pitch angle $\Theta(t)$.

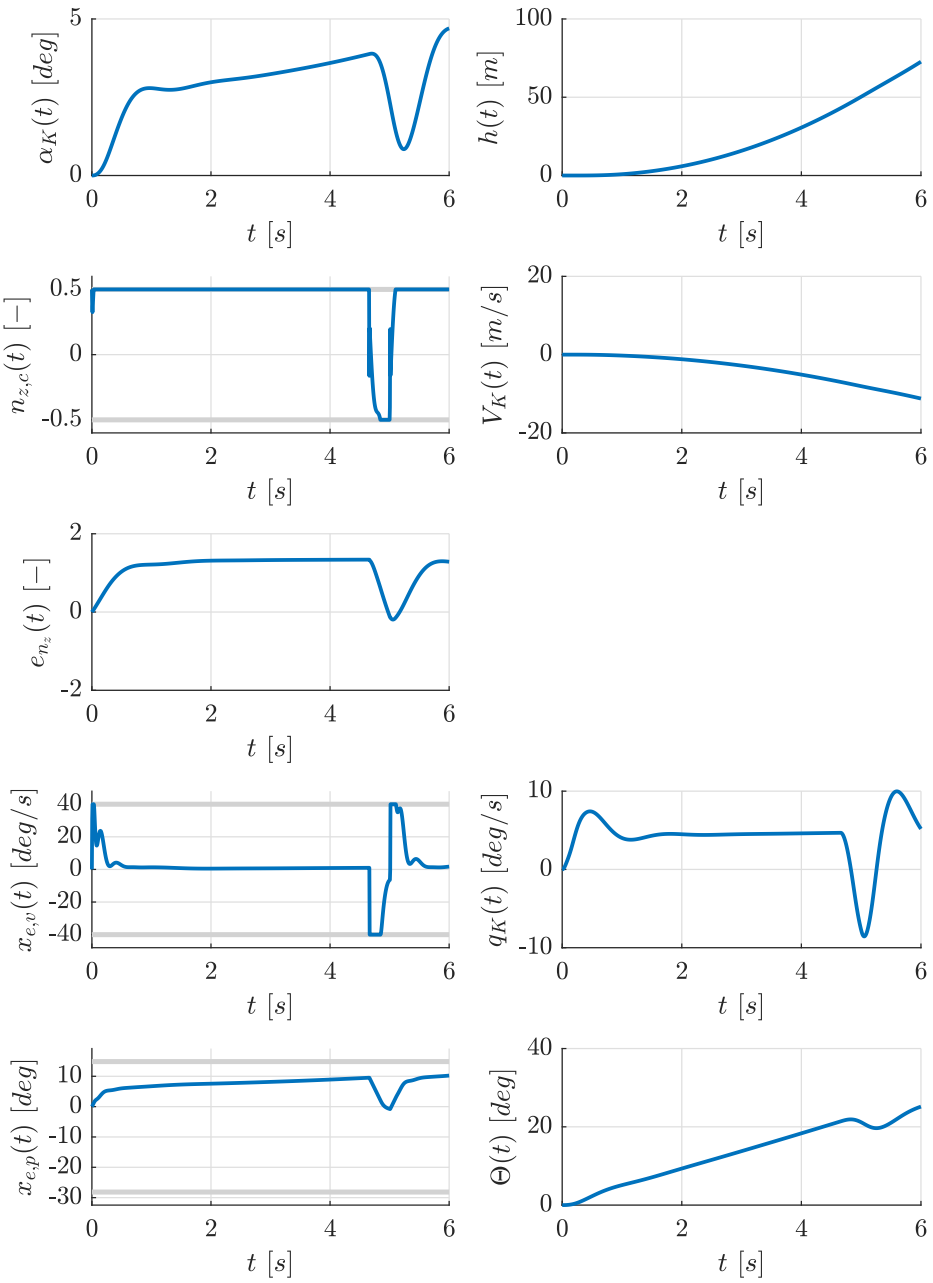


Figure A.4: Optimal results for the longitudinal plane using a Cat II model for maximizing $\alpha_K(t_f), t_f = 6\ s$ including the kinematic AoA $\alpha_K(t)$, the geodetic height $h(t)$, the normal load factor command $n_{z,c}(t)$, the kinematic velocity $V_K(t)$, the error controller state $e_{n_z}(t)$, the elevator rate $x_{e,v}(t)$, the pitch rate $q_K(t)$, the elevator position $x_{e,p}(t)$, and the pitch angle $\Theta(t)$.

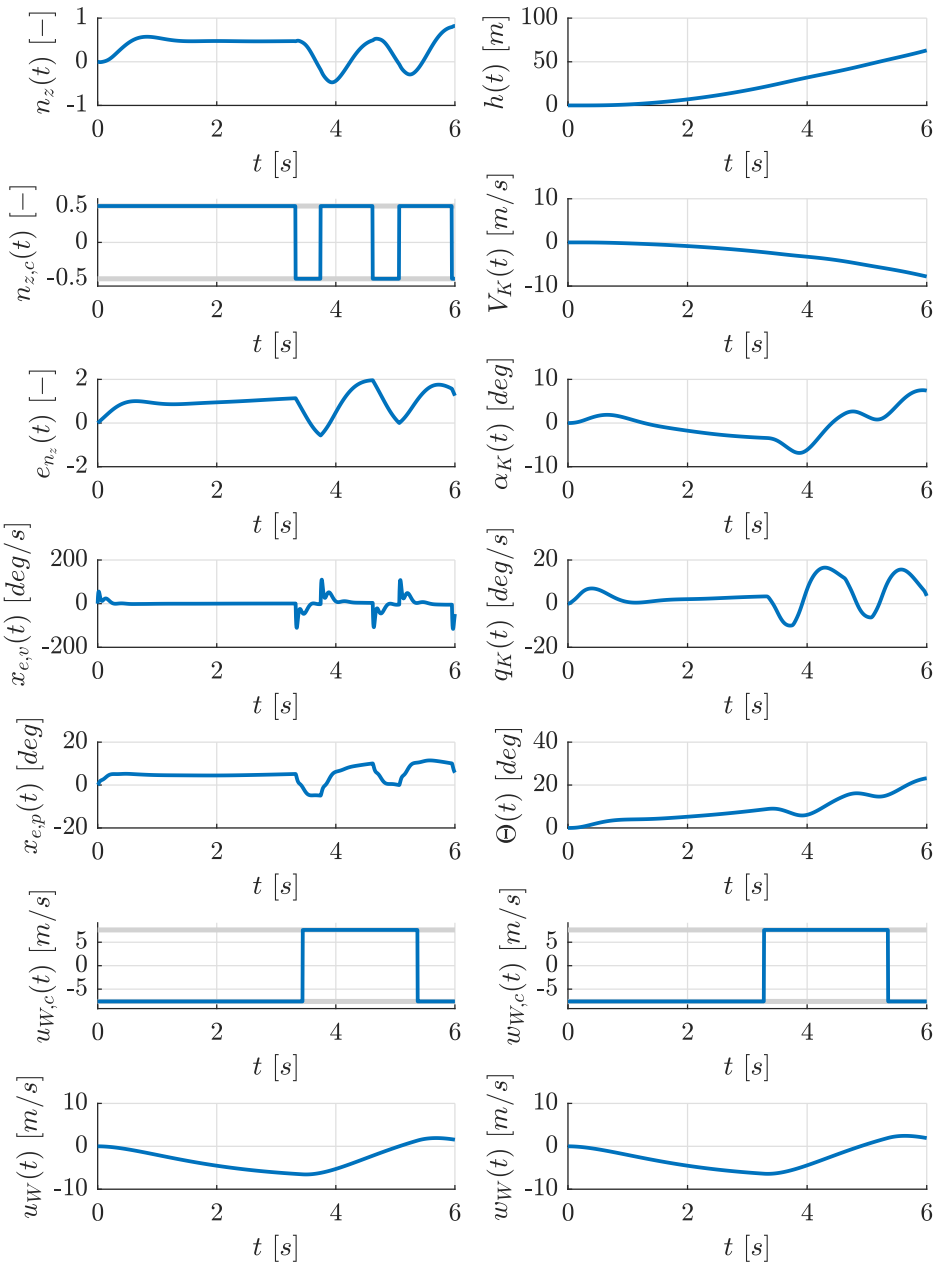


Figure A.5: Optimal results for the longitudinal plane using a Cat III model (black lines) and a Cat I model (blue lines) for maximizing $n_z(t_f), t_f = 6\ s$ with wind including the normal load factor $n_z(t)$, the geodetic height $h(t)$, the normal load factor command $n_{z,c}(t)$, the kinematic velocity $V_K(t)$, the error controller state $e_{n_z}(t)$, the kinematic AoA $\alpha_K(t)$, the elevator rate $x_{e,v}(t)$, the pitch rate $q_K(t)$, the elevator position $x_{e,p}(t)$, the pitch angle $\Theta(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, the longitudinal wind velocity $u_W(t)$, and the normal wind velocity $w_W(t)$.

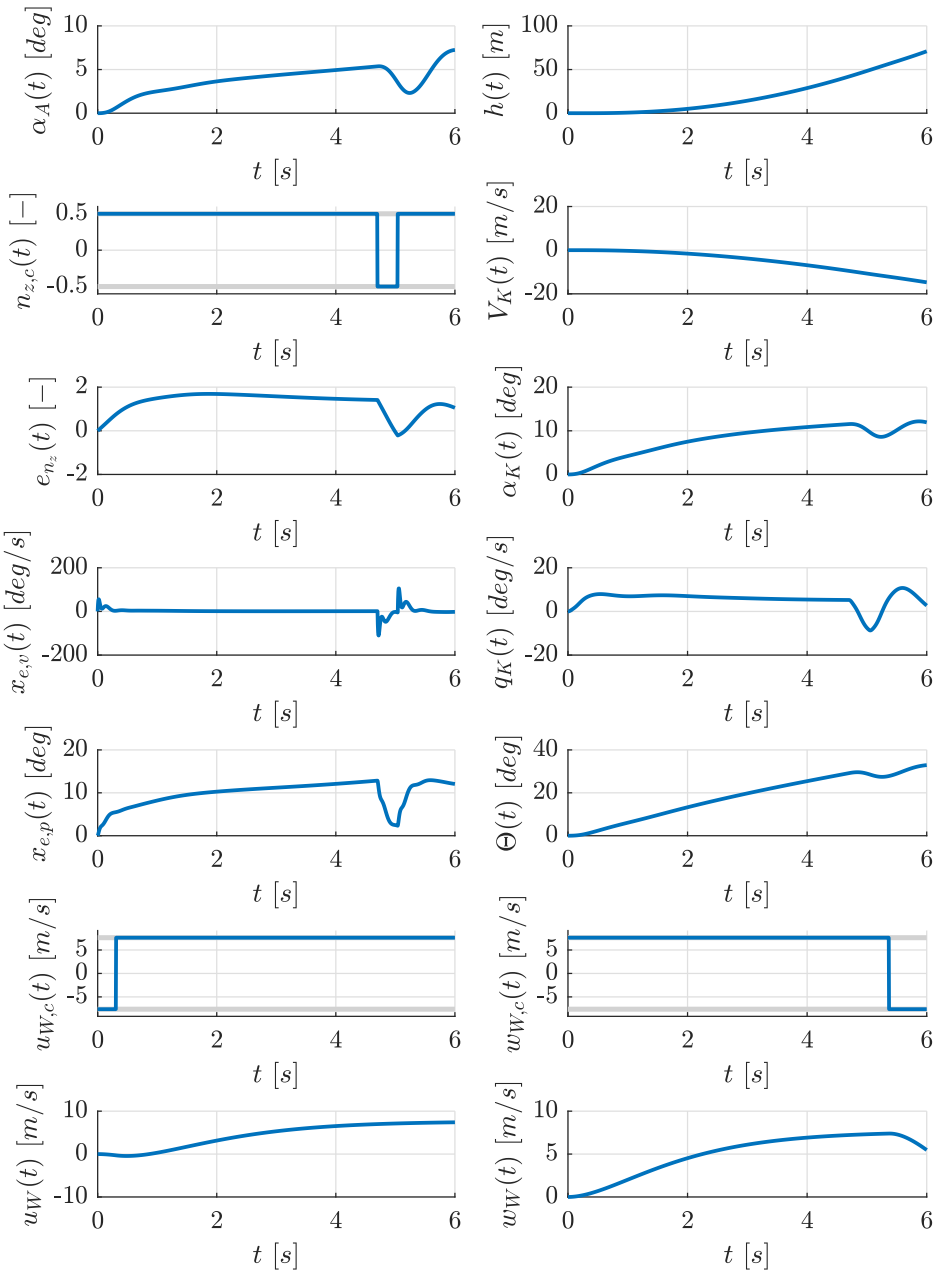


Figure A.6: Optimal results for the longitudinal plane using a Cat III model (black lines) and a Cat I model (blue lines) for maximizing $\alpha_A(t_f), t_f = 6$ s with wind including the AoA $\alpha_A(t)$, the geodetic height $h(t)$, the normal load factor command $n_{z,c}(t)$, the kinematic velocity $V_K(t)$, the error controller state $e_{n_z}(t)$, the kinematic AoA $\alpha_K(t)$, the elevator rate $x_{e,v}(t)$, the pitch rate $q_K(t)$, the elevator position $x_{e,p}(t)$, the pitch angle $\Theta(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, the longitudinal wind velocity $u_W(t)$, and the normal wind velocity $w_W(t)$.

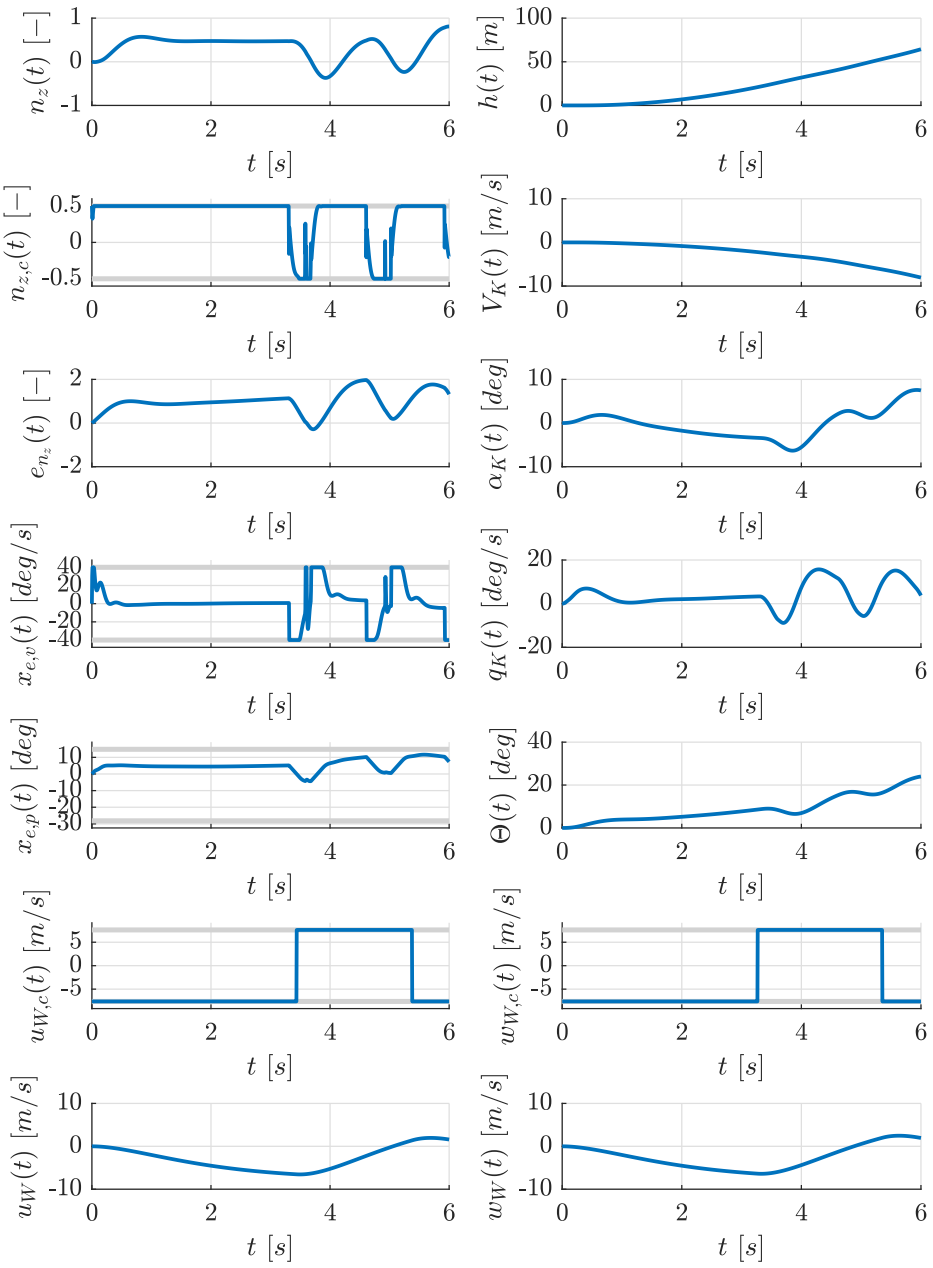


Figure A.7: Optimal results for the longitudinal plane using a Cat III model (black lines) and a Cat II model (blue lines) for maximizing $n_z(t_f), t_f = 6\ s$ with wind including the normal load factor $n_z(t)$, the geodetic height $h(t)$, the normal load factor command $n_{z,c}(t)$, the kinematic velocity $V_K(t)$, the error controller state $e_{n_z}(t)$, the kinematic AoA $\alpha_K(t)$, the elevator rate $x_{e,v}(t)$, the pitch rate $q_K(t)$, the elevator position $x_{e,p}(t)$, the pitch angle $\Theta(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, the longitudinal wind velocity $u_W(t)$, and the normal wind velocity $w_W(t)$.

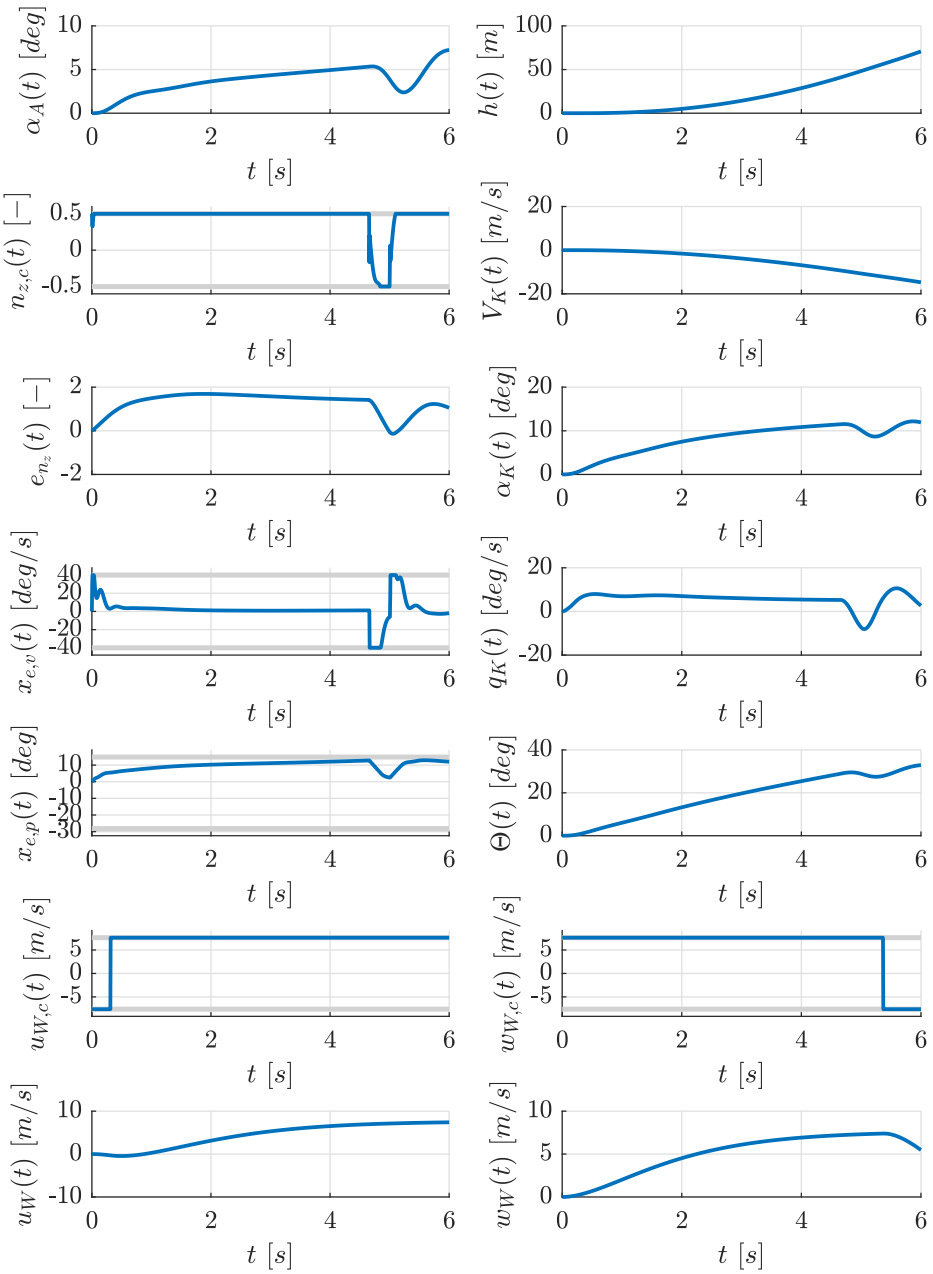


Figure A.8: Optimal results for the longitudinal plane using a Cat III model (black lines) and a Cat II model (blue lines) for maximizing $\alpha_A(t_f), t_f = 6$ s with wind including the AoA $\alpha_A(t)$, the geodetic height $h(t)$, the normal load factor command $n_{z,c}(t)$, the kinematic velocity $V_K(t)$, the error controller state $e_{n_z}(t)$, the kinematic AoA $\alpha_K(t)$, the elevator rate $x_{e,v}(t)$, the pitch rate $q_K(t)$, the elevator position $x_{e,p}(t)$, the pitch angle $\Theta(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, the longitudinal wind velocity $u_W(t)$, and the normal wind velocity $w_W(t)$.

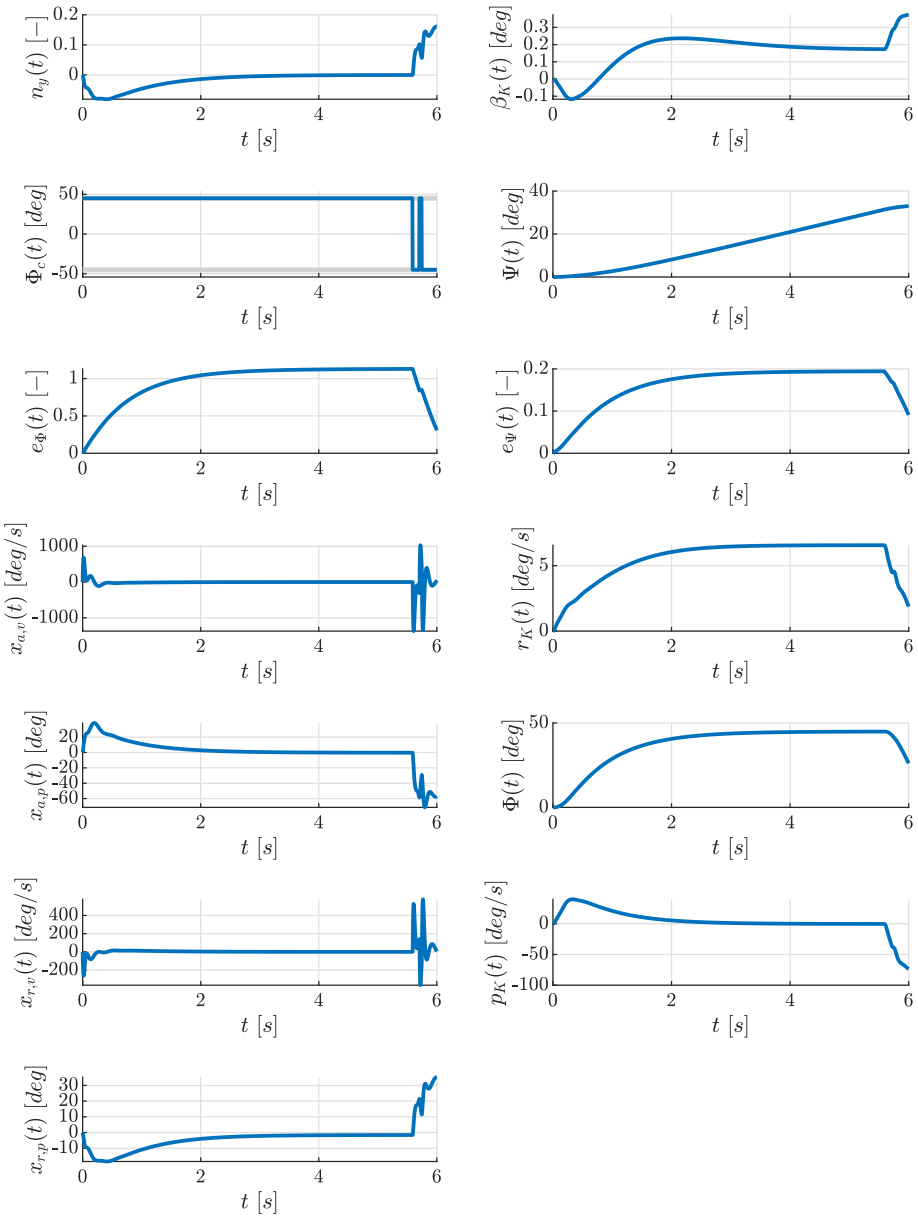


Figure A.9: Optimal results for the lateral plane using a Cat I model for maximizing $n_y(t_f), t_f = 6\ s$ including the lateral load factor $n_y(t)$, the kinematic AoS $\beta_K(t)$, the roll angle command $\Phi_c(t)$, the yaw angle $\Psi(t)$, the error controller state $e_\Phi(t)$, the error controller state $e_\Psi(t)$, the aileron rate $x_{a,v}(t)$, the pitch rate $r_K(t)$, the aileron position $x_{a,p}(t)$, the roll angle $\Phi(t)$, the rudder rate $x_{r,v}(t)$, the roll rate $p_K(t)$, the rudder position $x_{r,p}(t)$,

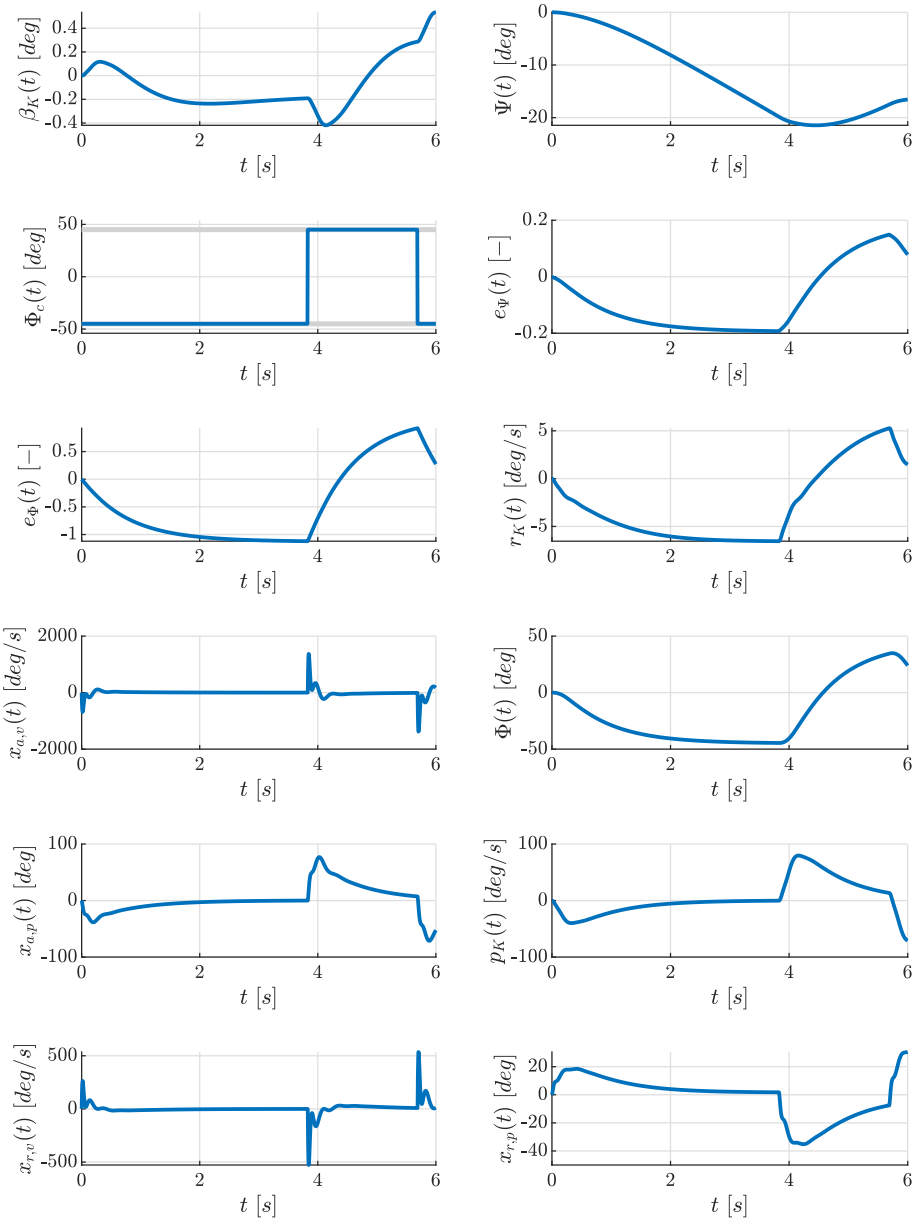


Figure A.10: Optimal results for the lateral plane using a Cat I model for maximizing $\beta_K(t_f), t_f = 6\ s$ including the kinematic AoS $\beta_K(t)$, the yaw angle $\Psi(t)$, the roll angle command $\Phi_c(t)$, the error controller state $e_\Psi(t)$, the error controller state $e_\Phi(t)$, the pitch rate $r_K(t)$, the aileron rate $x_{a,v}(t)$, the roll angle $\Phi(t)$, the aileron position $x_{a,p}(t)$, the roll rate $p_K(t)$, the rudder rate $x_{r,v}(t)$, and the rudder position $x_{r,p}(t)$.

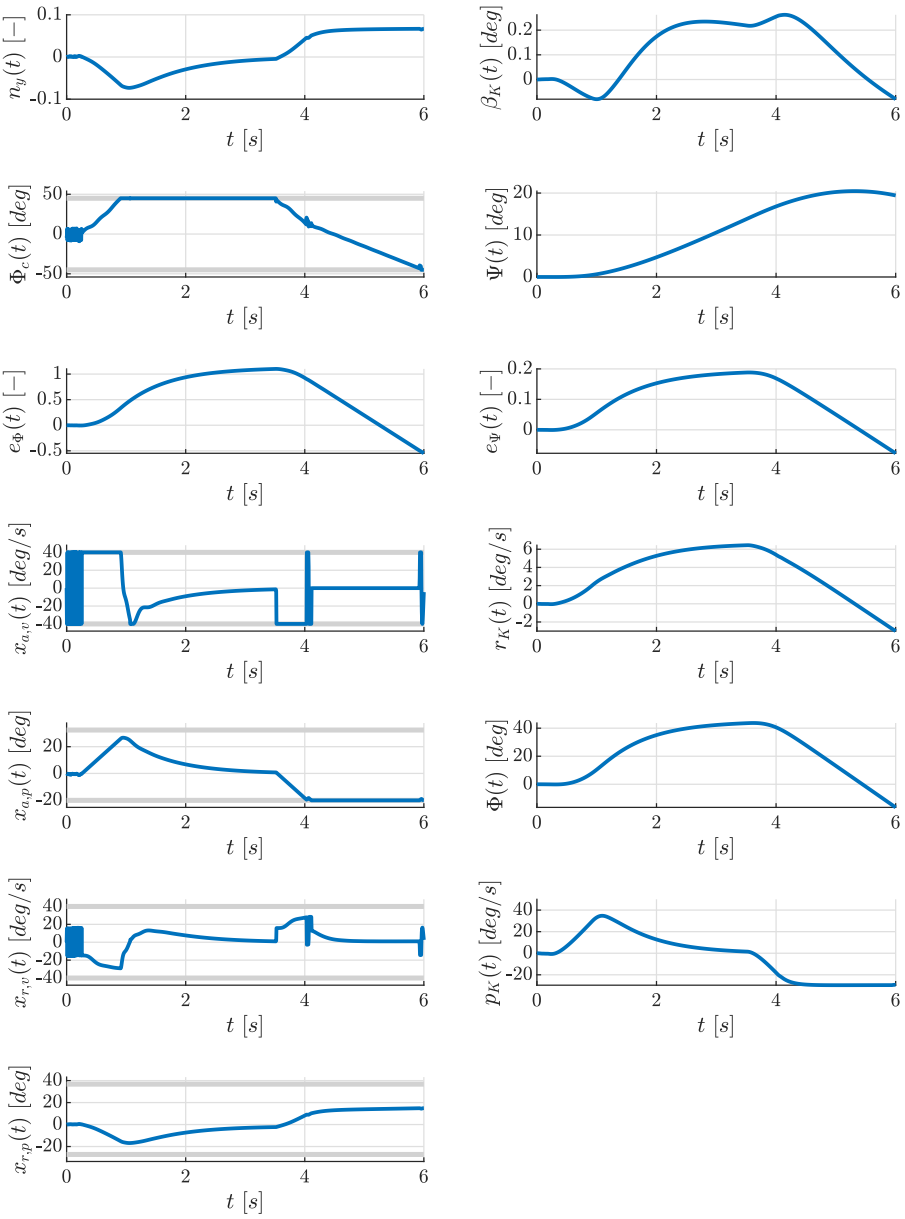


Figure A.11: Optimal results for the lateral plane using a Cat II model for maximizing $n_y(t_f), t_f = 6\ s$ including the lateral load factor $n_y(t)$, the kinematic AoS $\beta_K(t)$, the roll angle command $\Phi_c(t)$, the yaw angle $\Psi(t)$, the error controller state $e_\Phi(t)$, the error controller state $e_\Psi(t)$, the aileron rate $x_{a,v}(t)$, the pitch rate $r_K(t)$, the aileron position $x_{a,p}(t)$, the roll angle $\Phi(t)$, the rudder rate $x_{r,v}(t)$, the roll rate $p_K(t)$, the rudder position $x_{r,p}(t)$,

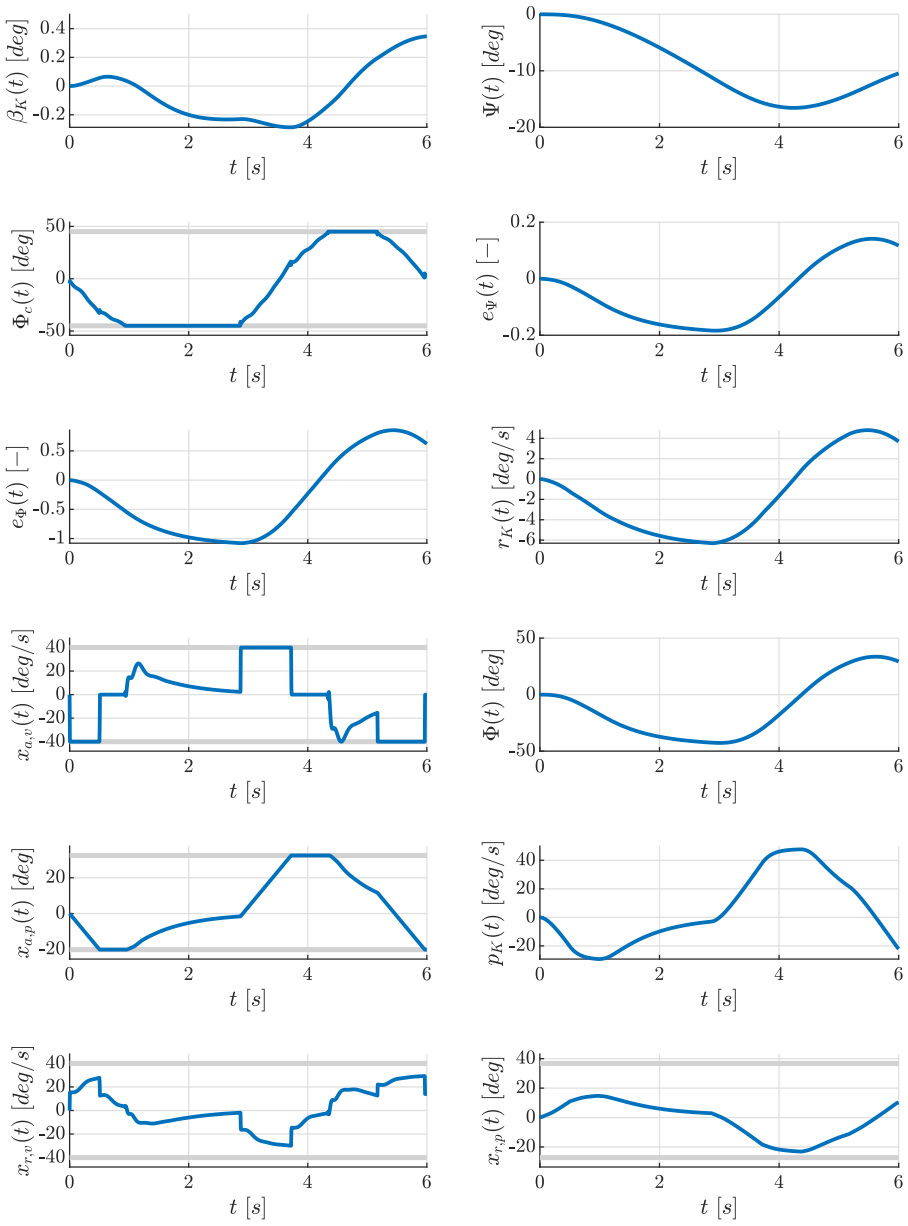


Figure A.12: Optimal results for the lateral plane using a Cat II model for maximizing $\beta_K(t_f), t_f = 6\ s$ including the kinematic AoS $\beta_K(t)$, the yaw angle $\Psi(t)$, the roll angle command $\Phi_c(t)$, the error controller state $e_\Psi(t)$, the error controller state $e_\Phi(t)$, the pitch rate $r_K(t)$, the aileron rate $x_{a,v}(t)$, the roll angle $\Phi(t)$, the aileron position $x_{a,p}(t)$, the roll rate $p_K(t)$, the rudder rate $x_{r,v}(t)$, and the rudder position $x_{r,p}(t)$.

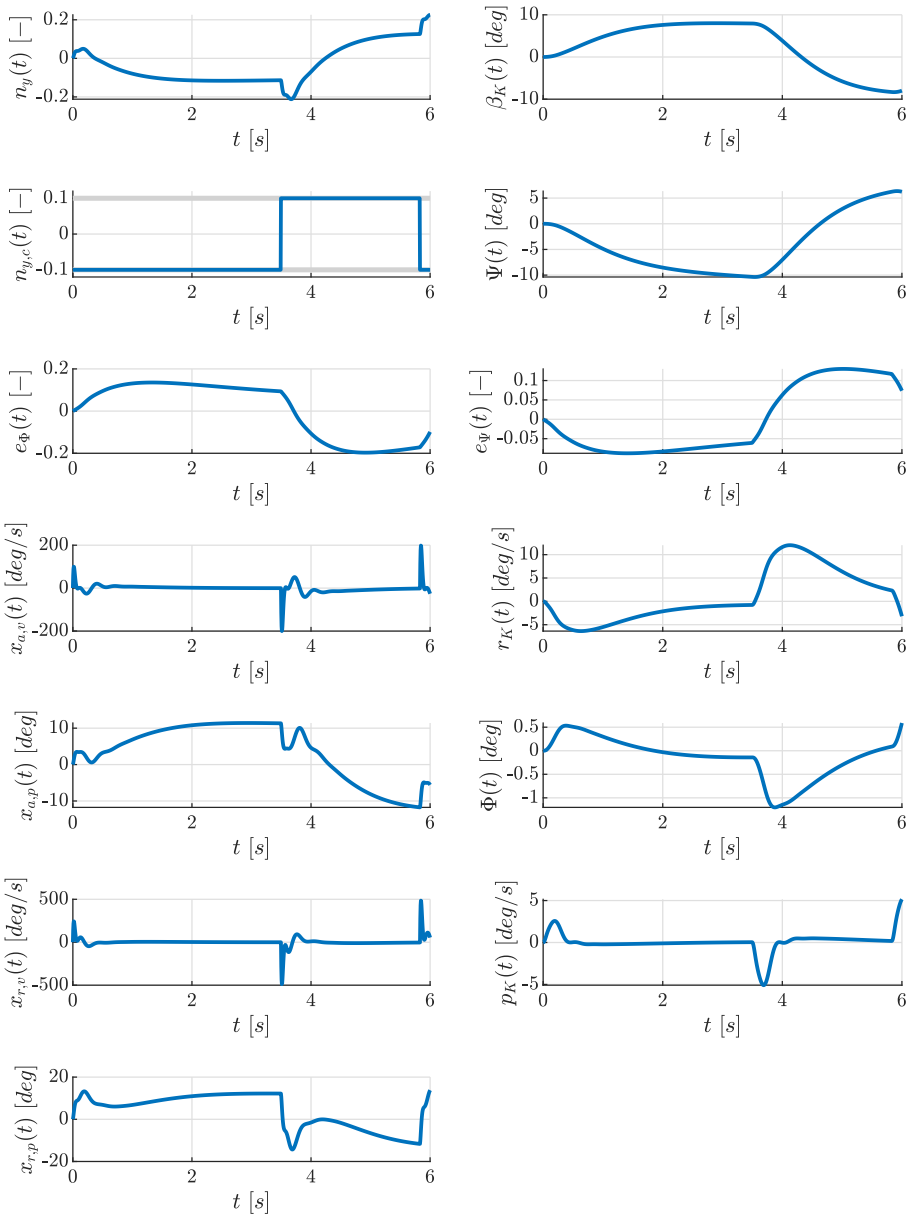


Figure A.13: Optimal results for the lateral plane using a Cat I model for maximizing $n_y(t_f), t_f = 6\ s$ including the lateral load factor $n_y(t)$, the kinematic AoS $\beta_K(t)$, the lateral load factor command $n_{y,c}(t)$, the yaw angle $\Psi(t)$, the error controller state $e_\Phi(t)$, the error controller state $e_\Psi(t)$, the aileron rate $x_{a,v}(t)$, the pitch rate $r_K(t)$, the aileron position $x_{a,p}(t)$, the roll angle $\Phi(t)$, the rudder rate $x_{r,v}(t)$, the roll rate $p_K(t)$, the rudder position $x_{r,p}(t)$,

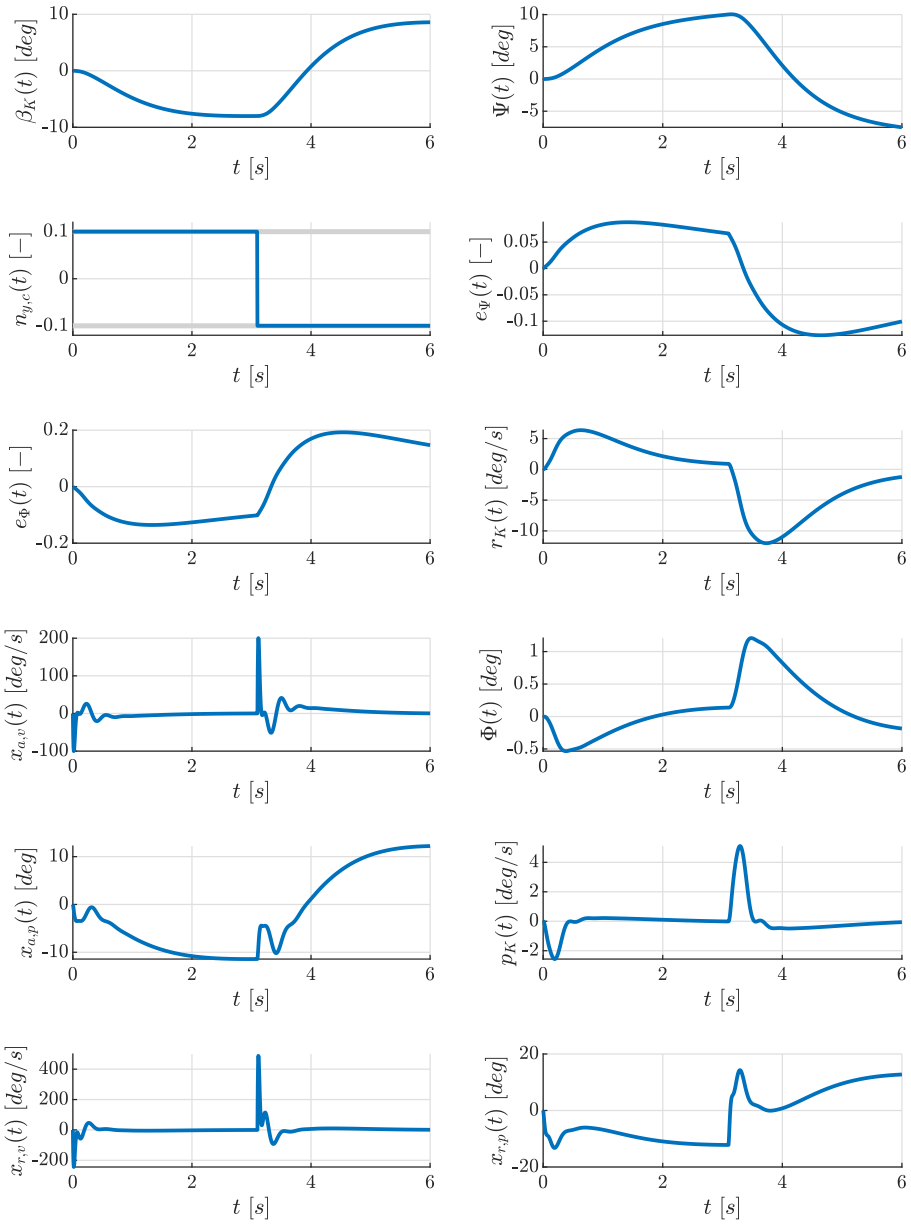


Figure A.14: Optimal results for the lateral plane using a Cat I model for maximizing $\beta_K(t_f), t_f = 6\ s$ including the kinematic AoS $\beta_K(t)$, the yaw angle $\Psi(t)$, the lateral load factor command $n_{y,c}(t)$, the error controller state $e_\Psi(t)$, the error controller state $e_\Phi(t)$, the pitch rate $r_K(t)$, the aileron rate $x_{a,v}(t)$, the roll angle $\Phi(t)$, the aileron position $x_{a,p}(t)$, the roll rate $p_K(t)$, the rudder rate $x_{r,v}(t)$, and the rudder position $x_{r,p}(t)$.

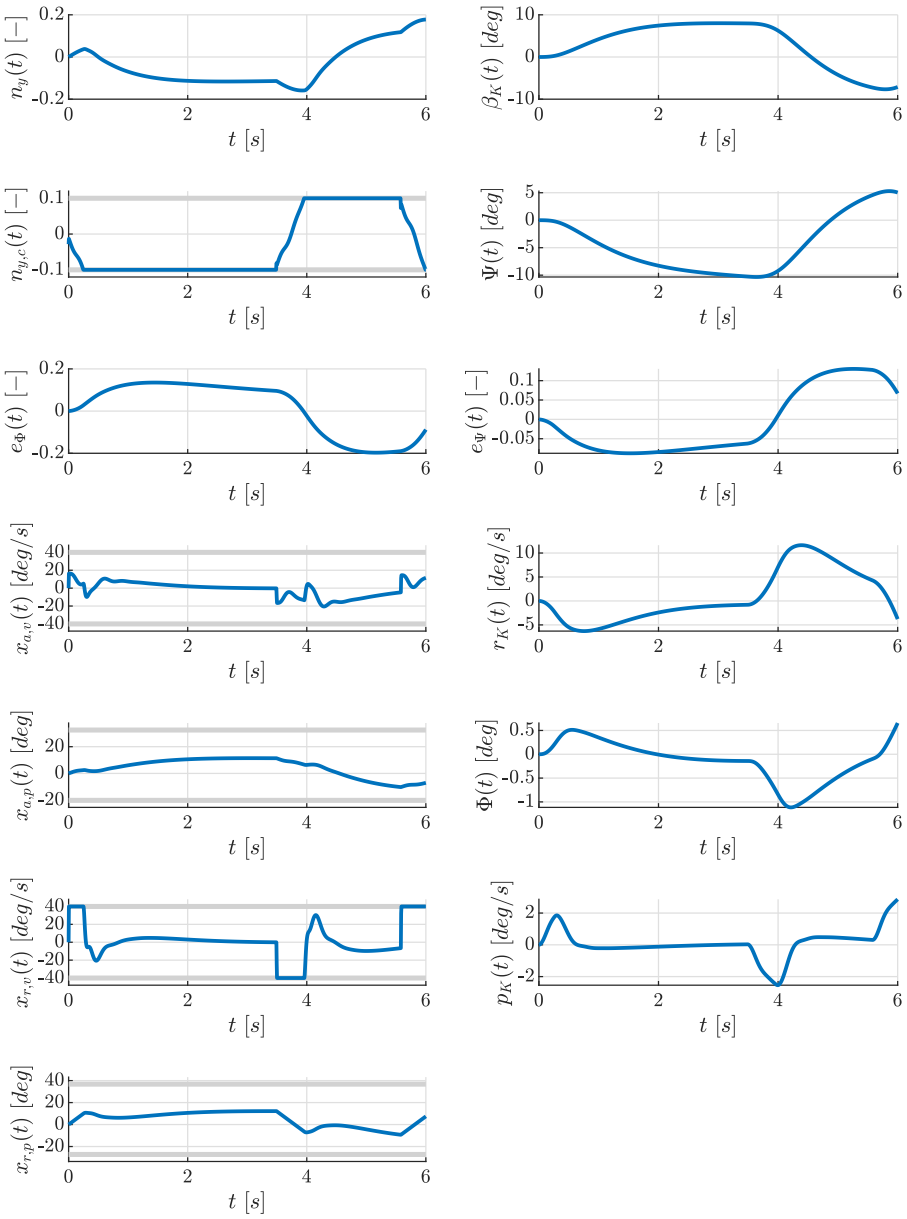


Figure A.15: Optimal results for the lateral plane using a Cat II model for maximizing $n_y(t_f), t_f = 6\ s$ including the lateral load factor $n_y(t)$, the kinematic AoS $\beta_K(t)$, the lateral load factor command $n_{y,c}(t)$, the yaw angle $\Psi(t)$, the error controller state $e_\Phi(t)$, the error controller state $e_\Psi(t)$, the aileron rate $x_{a,v}(t)$, the pitch rate $r_K(t)$, the aileron position $x_{a,p}(t)$, the roll angle $\Phi(t)$, the rudder rate $x_{r,v}(t)$, the roll rate $p_K(t)$, the rudder position $x_{r,p}(t)$,

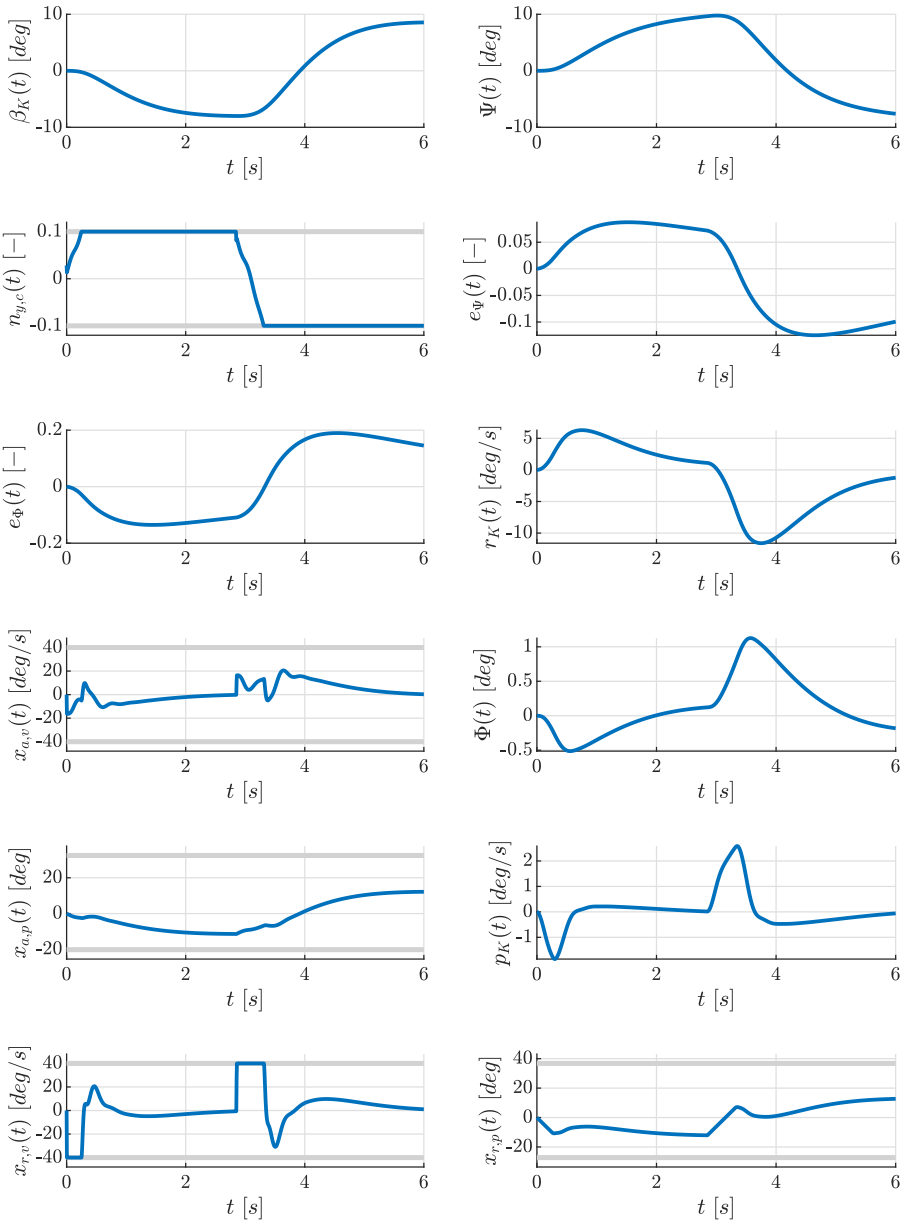


Figure A.16: Optimal results for the lateral plane using a Cat II model for maximizing $\beta_K(t_f), t_f = 6\ s$ including the kinematic AoS $\beta_K(t)$, the yaw angle $\Psi(t)$, the lateral load factor command $n_{y,c}(t)$, the error controller state $e_\Psi(t)$, the error controller state $e_\Phi(t)$, the pitch rate $r_K(t)$, the aileron rate $x_{a,v}(t)$, the roll angle $\Phi(t)$, the aileron position $x_{a,p}(t)$, the roll rate $p_K(t)$, the rudder rate $x_{r,v}(t)$, and the rudder position $x_{r,p}(t)$.

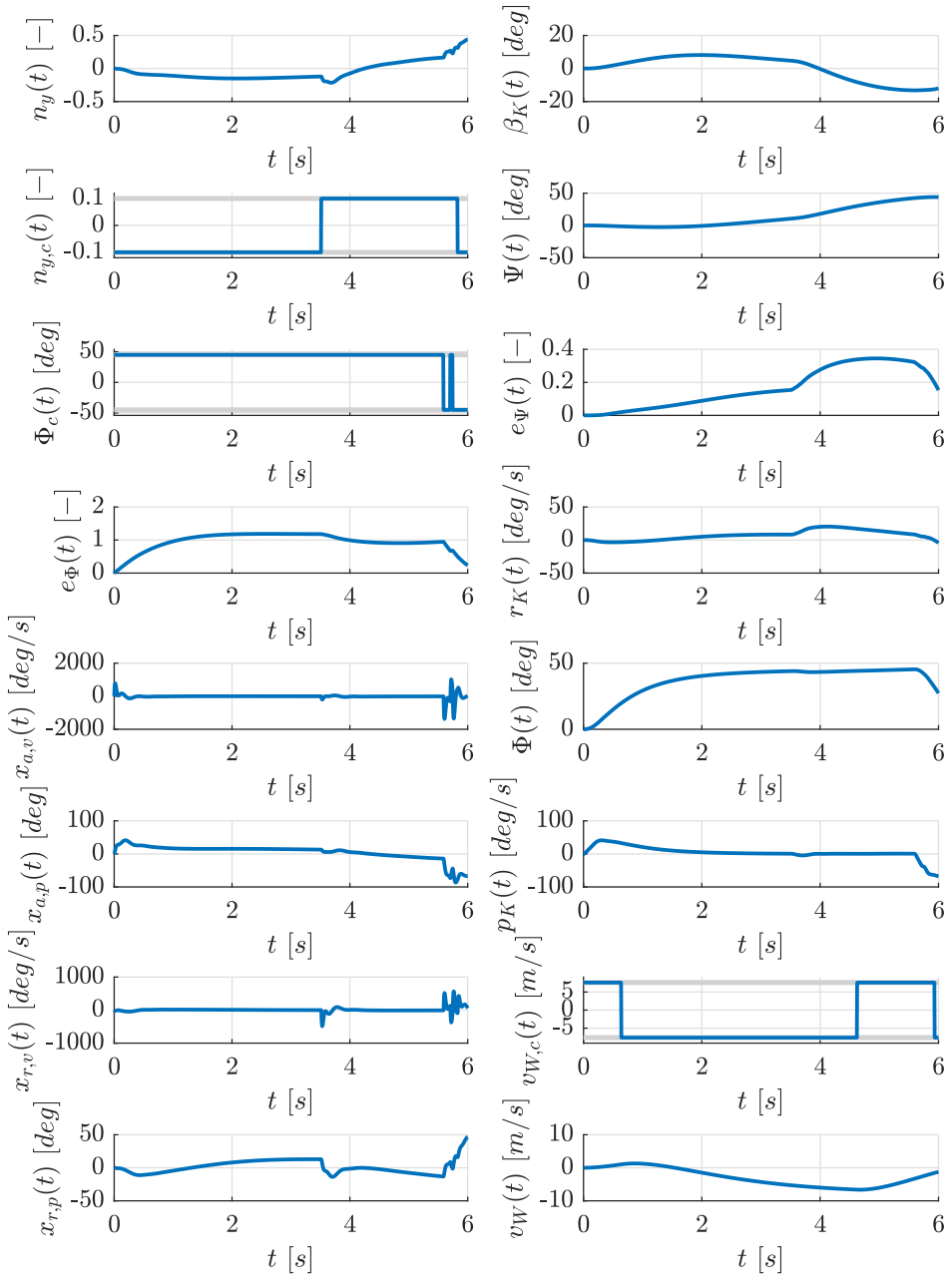


Figure A.17: Optimal results for the lateral plane using a Cat III model (black lines) and a Cat I model (blue lines) for maximizing $n_y(t_f), t_f = 6\ s$ with wind including the lateral load factor $n_y(t)$, the kinematic AoS $\beta_K(t)$, the lateral load factor command $n_{y,c}(t)$, the yaw angle $\Psi(t)$, the roll angle command $\Phi_c(t)$, the error controller state $e_\Psi(t)$, the error controller state $e_\Phi(t)$, the pitch rate $r_K(t)$, the aileron rate $x_{a,v}(t)$, the roll angle $\Phi(t)$, the aileron position $x_{a,p}(t)$, the roll rate $p_K(t)$, the rudder rate $x_{r,v}(t)$, the lateral wind velocity command $v_{W,c}(t)$, the rudder position $x_{r,p}(t)$, and the lateral wind velocity $v_W(t)$.

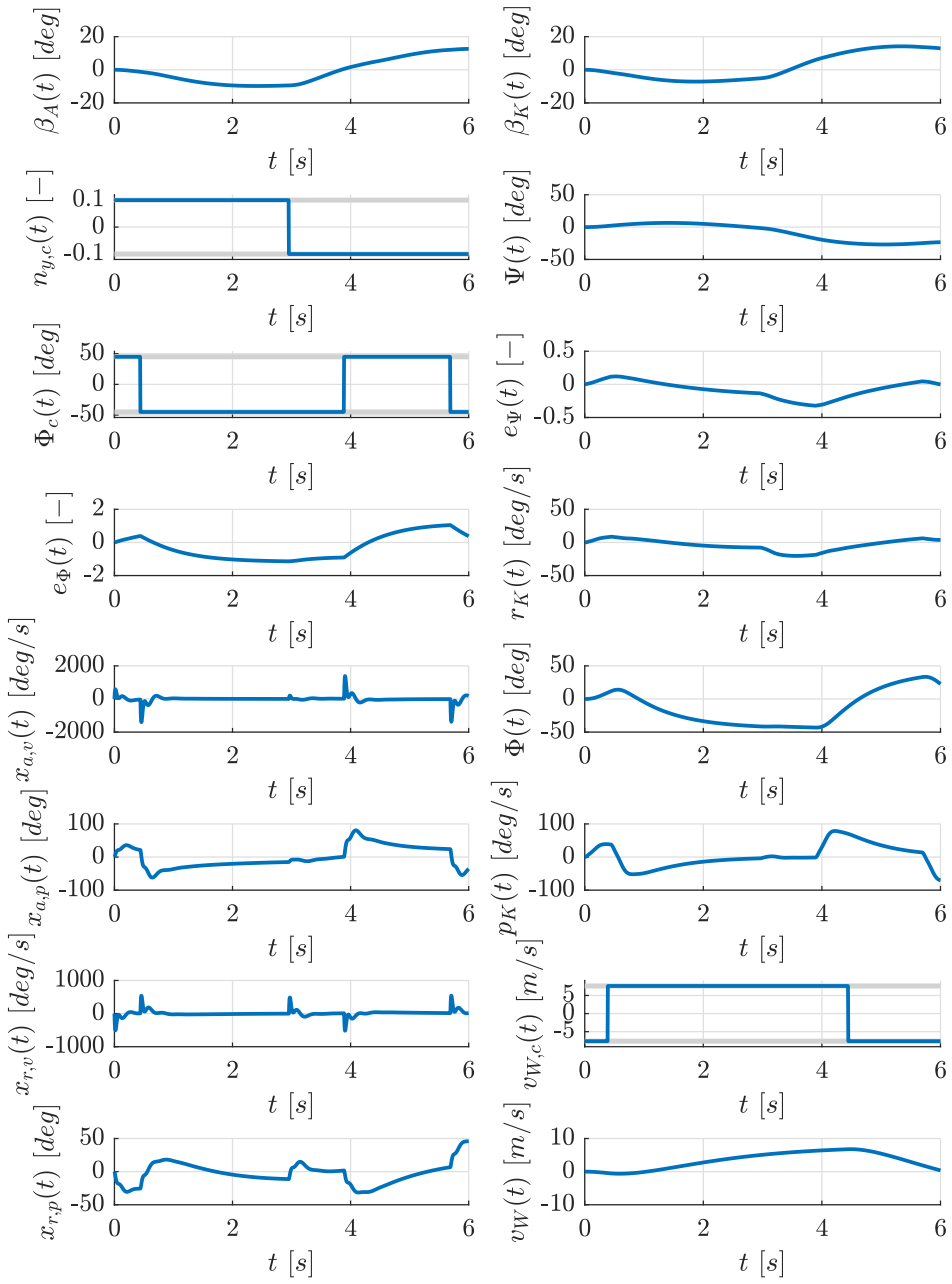


Figure A.18: Optimal results for the lateral plane using a Cat III model (black lines) and a Cat I model (blue lines) for maximizing $\beta_A(t_f), t_f = 6\ s$ with wind including the AoS $\beta_A(t)$, the kinematic AoS $\beta_K(t)$, the lateral load factor command $n_{y,c}(t)$, the yaw angle $\Psi(t)$, the roll angle command $\Phi_c(t)$, the error controller state $e_\Psi(t)$, the error controller state $e_\Phi(t)$, the pitch rate $r_K(t)$, the aileron rate $x_{a,v}(t)$, the roll angle $\Phi(t)$, the aileron position $x_{a,p}(t)$, the roll rate $p_K(t)$, the rudder rate $x_{r,v}(t)$, the lateral wind velocity command $v_{W,c}(t)$, the rudder position $x_{r,p}(t)$, and the lateral wind velocity $v_W(t)$.

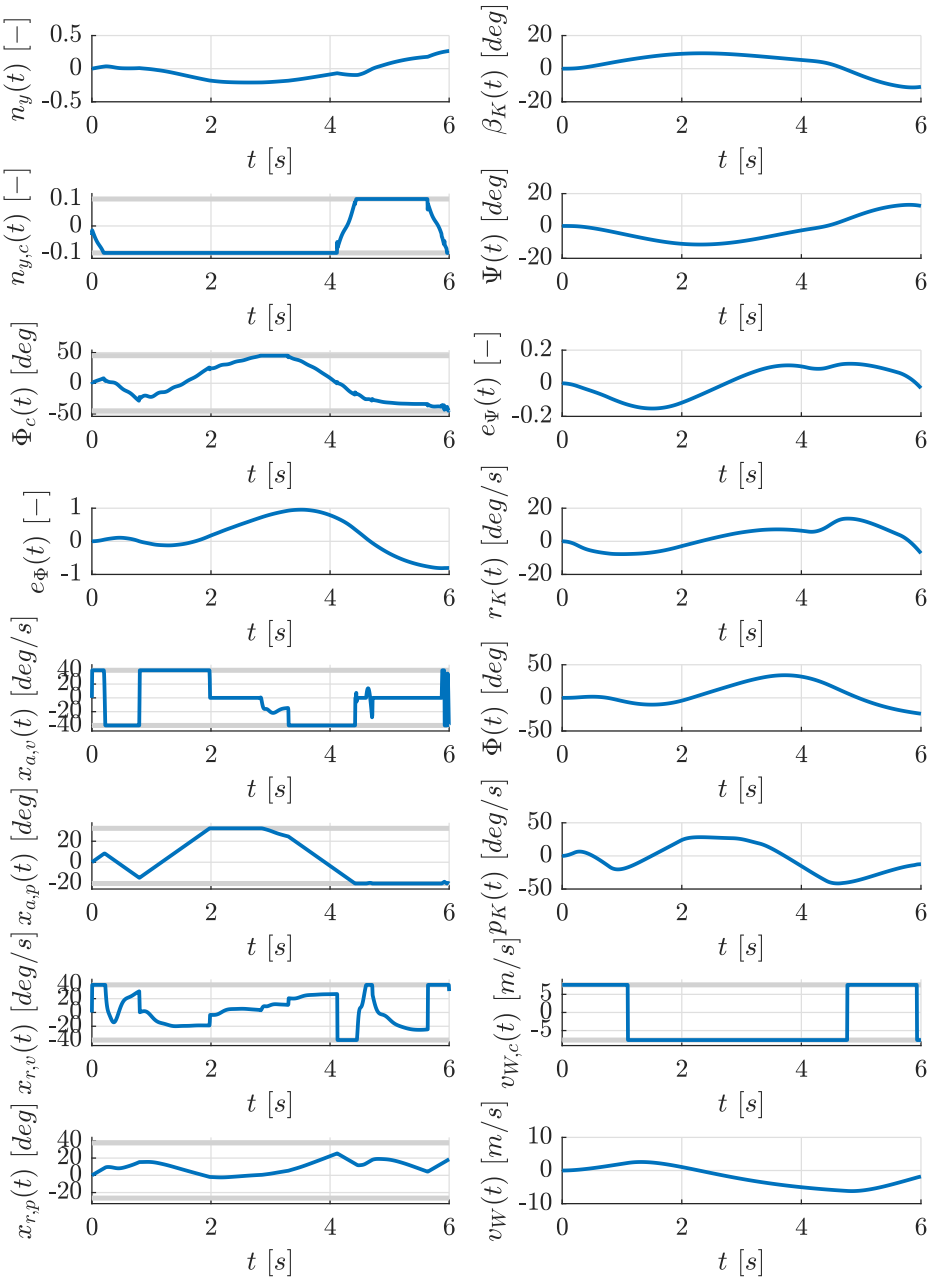


Figure A.19: Optimal results for the lateral plane using a Cat III model (black lines) and a Cat II model (blue lines) for maximizing $n_y(t_f), t_f = 6\ s$ with wind including the lateral load factor $n_y(t)$, the kinematic AoS $\beta_K(t)$, the lateral load factor command $n_{y,c}(t)$, the yaw angle $\Psi(t)$, the roll angle command $\Phi_c(t)$, the error controller state $e_\Psi(t)$, the error controller state $e_\Phi(t)$, the pitch rate $r_K(t)$, the aileron rate $x_{a,v}(t)$, the roll angle $\Phi(t)$, the aileron position $x_{a,p}(t)$, the roll rate $p_K(t)$, the rudder rate $x_{r,v}(t)$, the lateral wind velocity command $v_{W,c}(t)$, the rudder position $x_{r,p}(t)$, and the lateral wind velocity $v_W(t)$.

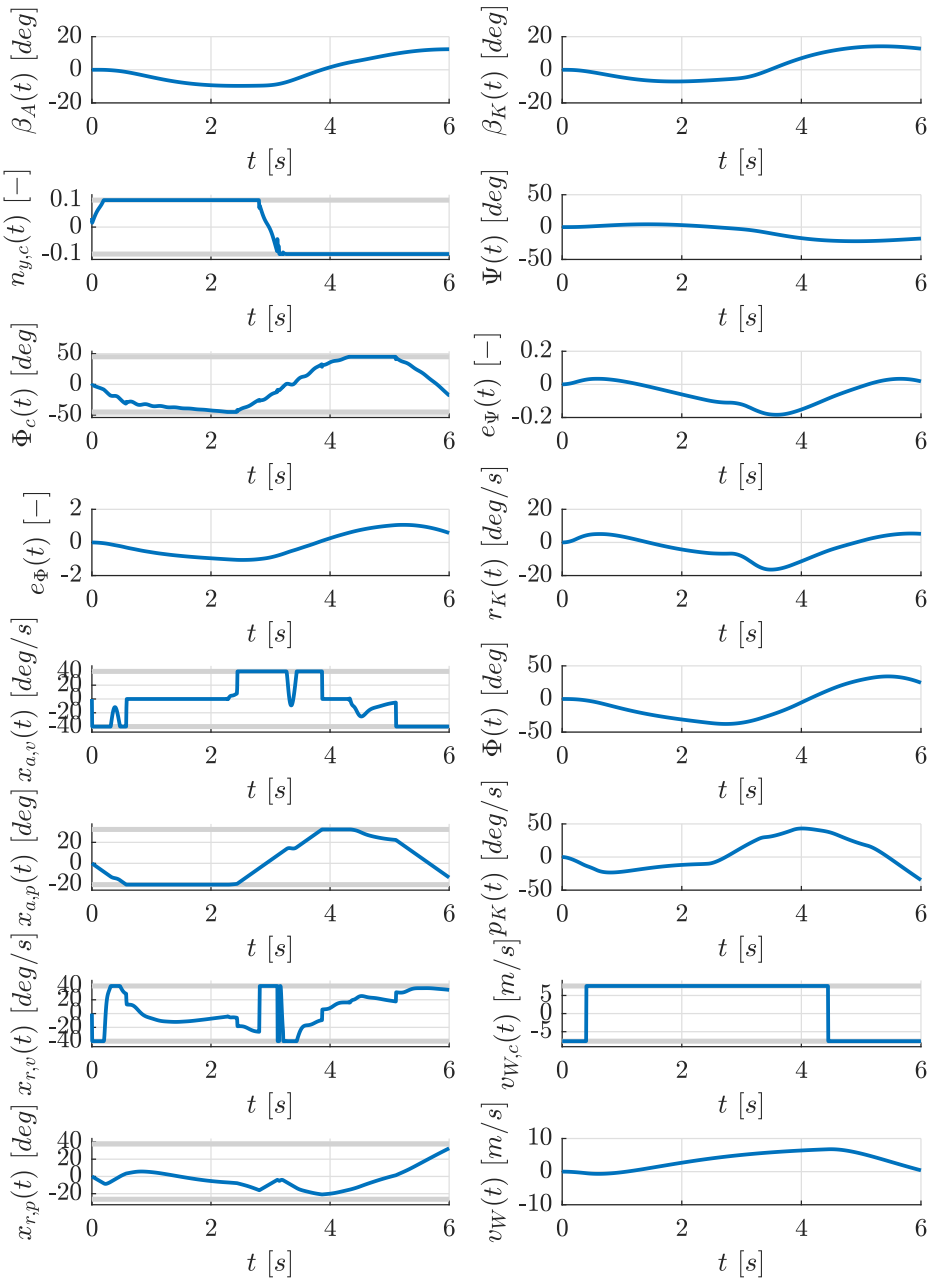


Figure A.20: Optimal results for the lateral plane using a Cat III model (black lines) and a Cat II model (blue lines) for maximizing $\beta_A(t_f), t_f = 6\ s$ with wind including the AoS $\beta_A(t)$, the kinematic AoS $\beta_K(t)$, the lateral load factor command $n_{y,c}(t)$, the yaw angle $\Psi(t)$, the roll angle command $\Phi_c(t)$, the error controller state $e_\Psi(t)$, the error controller state $e_\Phi(t)$, the pitch rate $r_K(t)$, the aileron rate $x_{a,v}(t)$, the roll angle $\Phi(t)$, the aileron position $x_{a,p}(t)$, the roll rate $p_K(t)$, the rudder rate $x_{r,v}(t)$, the lateral wind velocity command $v_{W,c}(t)$, the rudder position $x_{r,p}(t)$, and the lateral wind velocity $v_W(t)$.

Appendix B

Results Numerical Experiments (Nonlinear)

This appendix chapter contains additional numerical results from Sec. 9.1.3.

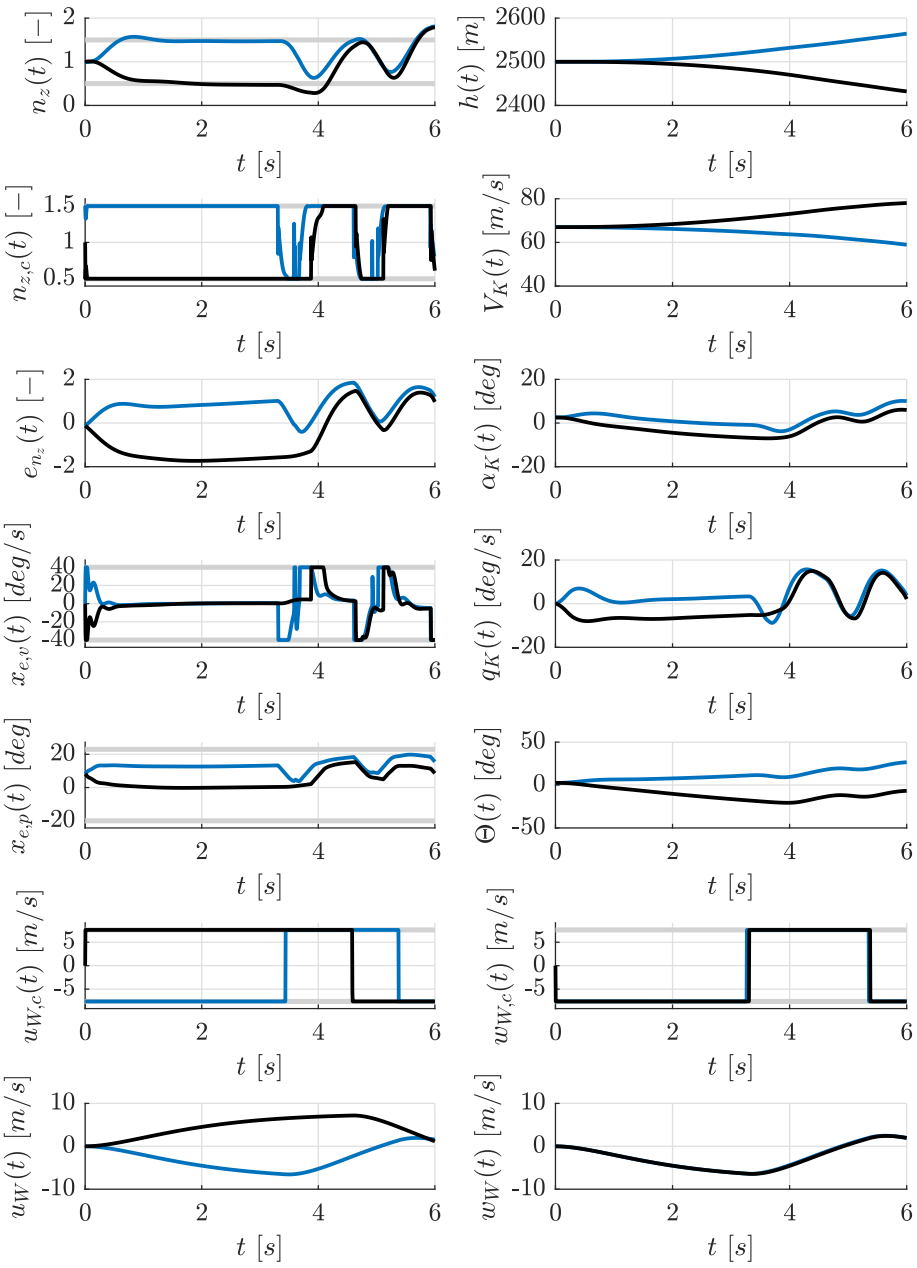


Figure B.1: Optimal results for the longitudinal plane using a Cat III model (black lines) and a Cat II model (blue lines) for maximizing $n_z(t_f), t_f = 6\ s$ with wind including the normal load factor $n_z(t)$, the geodetic height $h(t)$, the normal load factor command $n_{z,c}(t)$, the kinematic velocity $V_K(t)$, the error controller state $e_{n_z}(t)$, the kinematic AoA $\alpha_K(t)$, the elevator rate $x_{e,v}(t)$, the pitch rate $q_K(t)$, the elevator position $x_{e,p}(t)$, the pitch angle $\Theta(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, the longitudinal wind velocity $u_W(t)$, and the normal wind velocity $w_W(t)$.

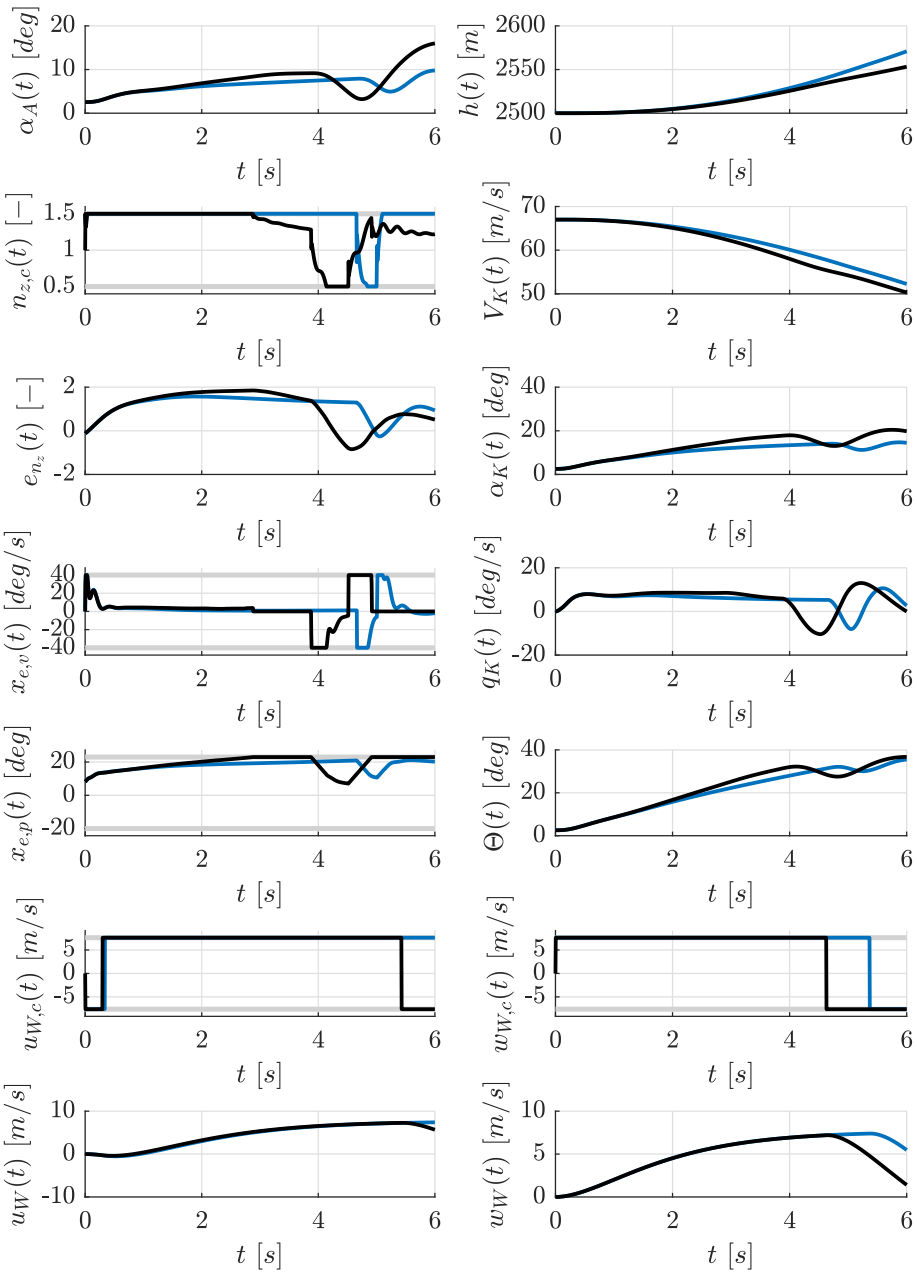


Figure B.2: Optimal results for the longitudinal plane using a Cat III model (black lines) and a Cat II model (blue lines) for maximizing $\alpha_A(t_f), t_f = 6$ s with wind including the AoA $\alpha_A(t)$, the geodetic height $h(t)$, the normal load factor command $n_{z,c}(t)$, the kinematic velocity $V_K(t)$, the error controller state $e_{n_z}(t)$, the kinematic AoA $\alpha_K(t)$, the elevator rate $x_{e,v}(t)$, the pitch rate $q_K(t)$, the elevator position $x_{e,p}(t)$, the pitch angle $\Theta(t)$, the longitudinal wind velocity command $u_{W,c}(t)$, the normal wind velocity command $w_{W,c}(t)$, the longitudinal wind velocity $u_W(t)$, and the normal wind velocity $w_W(t)$.

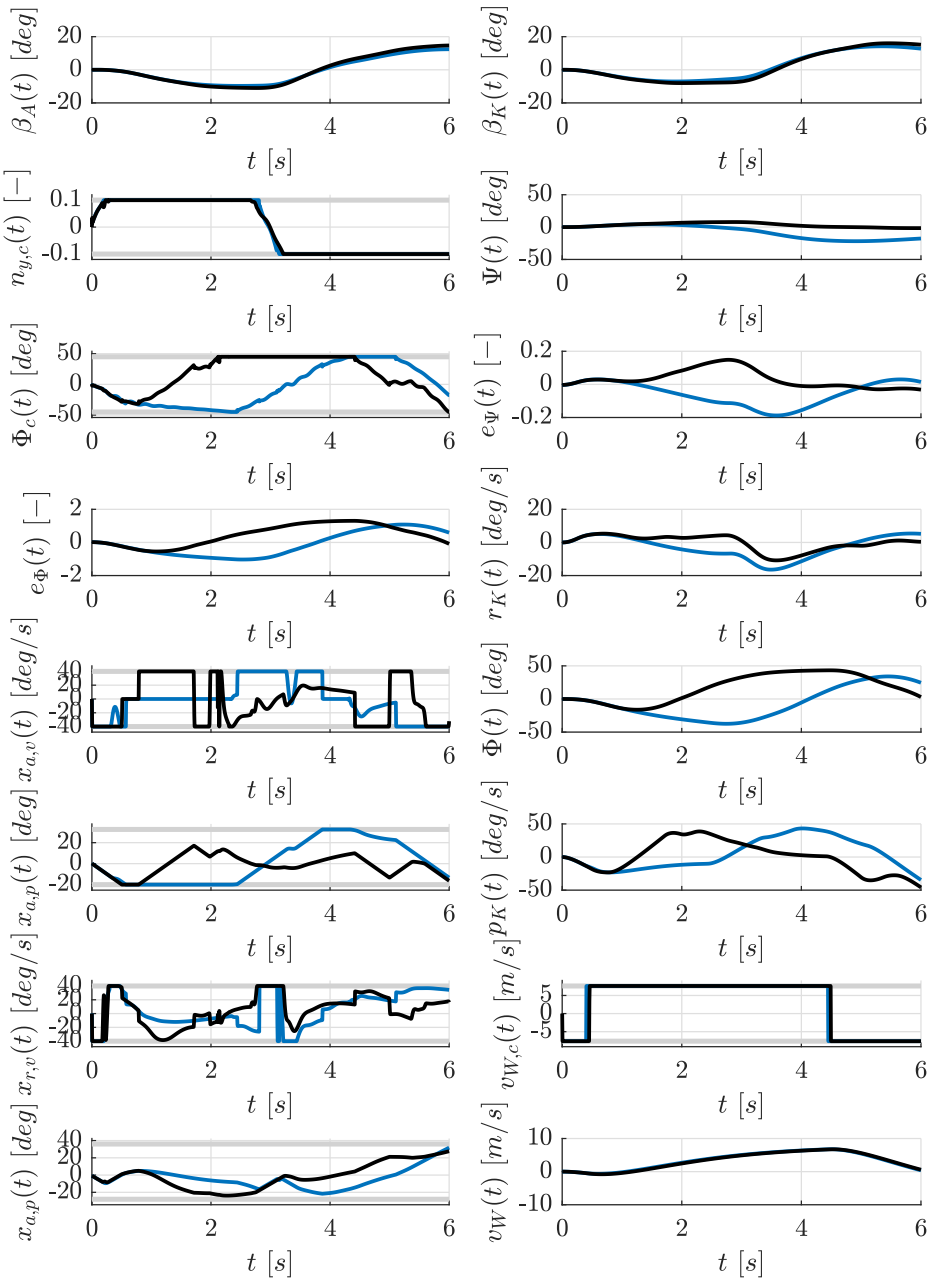


Figure B.3: Optimal results for the lateral plane using a Cat III model (black lines) and a Cat II model (blue lines) for maximizing $\beta_A(t_f), t_f = 6\ s$ with wind including the AoS $\beta_A(t)$, the kinematic AoS $\beta_K(t)$, the lateral load factor command $n_{y,c}(t)$, the yaw angle $\Psi(t)$, the roll angle command $\Phi_c(t)$, the error controller state $e_\Psi(t)$, the error controller state $e_\Phi(t)$, the pitch rate $r_K(t)$, the aileron rate $x_{a,v}(t)$, the roll angle $\Phi(t)$, the aileron position $x_{a,p}(t)$, the roll rate $p_K(t)$, the rudder rate $x_{r,v}(t)$, the lateral wind velocity command $v_{W,c}(t)$, the rudder position $x_{r,p}(t)$, and the lateral wind velocity $v_W(t)$.

Appendix C

Practical Considerations and Best Practices (Summary)

In the following practical aspects as well as best practices for the optimal control based clearance approaches presented in this thesis are summarized:

- **Models**: For Cat I (linear) and Cat II (quasi-linear) type models the dynamic system is obtained from a linearization around a steady-state flight condition. Depending on the reference flight condition the longitudinal and lateral channel can be investigated separately. In case of Cat III (nonlinear) models under the intrusive approach the model needs to be available in state-space form in order to meet the modeling requirements of state of the art direct optimal control software. Moreover, the models should fulfill the differentiability requirements if gradient-based methods are used. As such, it is advisable to perform a model reduction under this approach and to remove non-differentiable features such as internal limiters which may otherwise cause difficulties regarding the numerical solution. Furthermore, it is desirable to use smooth interpolation schemes for table data (such as for aerodynamic derivatives or gain scheduling) even though it is observed that optimization solvers appear to cope well with multi-linear interpolation methods. Regarding the simulation models under the Cat III non-intrusive approach the computational load is observed to be *considerably* lower if the model is executed in compiled from (e.g. as MATLAB®-Executable for Simulink®-models).

- **Wind**: Optimal wind disturbance inputs are used in this work to test the flight control system. In order to produce smooth wind gusts a linear second order system is introduced to model the wind velocity and rotation. It is important to mention that under this modeling approach the exact shape of the wind gust is not prescribed but determined by the solution of the optimal control problem.

- **Discretization**: Under the direct optimal control approach the discretization can be performed using shooting, full discretization, and collocation methods. In this thesis the Backward Euler method is used if the focus lies on the control functions as this method appears to produce few numerical artifacts in the optimal control histories. However, it is important to mention that also discretization methods of higher order can be used. For the Trapezoidal collocation method it is observed that this method is prone to chattering effects on state constrained arcs.

- **Constraints**: Inputs (commands, wind) are subject to box bounds. Actuator rate and position limits are introduced as purely state dependent constraints.

- **Numerical Solution (Linear Optimal Control Problems)**: Linear Programming (LP) solvers can be used to solve the discretized form of the optimal control problems. Throughout this thesis the LP solver CPLEX is used with different methods (Primal Simplex, Dual Simplex, Interior Point, Concurrent). It is observed that depending on the particular problem a considerable speed-up can be achieved simply by switching to a different method. In addition, some methods seem to produce cleaner results which is most likely contributed to internal numerical tolerances used by the solver. Moreover, if a large number of similar problems needs to be solved it appears to be beneficial for the solution time to provide a basis matrix to the solver in order to warm-start the solution.

- **Numerical Solution (Parameter Optimization)**: Regarding the parameter optimization methods under the bi-level approach both global and local (gradient-based) methods can be used. Regarding the local optimization methods both the Trust-Region and the SQP method used in this thesis appear to work very efficiently. A simple globalization strategy for these methods is to start the solution from several initial points.

- **Numerical Solution (Nonlinear Optimal Control Problems)**: For the numerical solution of the nonlinear optimal control problems the interior point solver IPOPT is used in this thesis which is well suited for large scale problems.

- **Control Parametrization (Cat III Hybrid)**: For the Cat III hybrid approach the switching time points of the optimal control functions obtained from the solution using a Cat I model need to be determined. It is found that the identification of theses switching time points based on the zeros of the switching functions works very robustly regarding the numerical solution. A determination purely depending on the control functions is observed to be challenging as in some cases these function appear to be more prone to numerical artifacts such as chattering effects.

Appendix D

Publications

The following references which are accepted for publication or have been published by the author during the time as a research associate at the Institute of Flight System Dynamics (TUM):

- J. Diepolder, J. Z. Ben-Asher, and F. Holzapfel, "Flight Control Law Clearance Using Worst-Case Inputs Under Parameter Uncertainty," *AIAA Journal*, Vol. 43, No. 10, pp. 1967-1974, 2020, doi: 10.2514/1.G005236.
- M. Wang, S. Zhang, J. Diepolder, and F. Holzapfel, "Battery package design optimization for small electric aircraft," *Chinese Journal of Aeronautics (2020)*, accessed 16 Aug 2020, doi: 10.1016/j.cja.2020.04.021.
- S. Hosseini , N. D. Botkin , J. Diepolder, and F. Holzapfel, "Robust optimal input design for flight vehicle system identification," *AIAA Scitech Forum*, AIAA 2020-0290, Orlando (US-FL), 2020, doi: 10.2514/6.2020-0290.
- J. Diepolder, N. D. Botkin, and F. Holzapfel, "Optimal Disturbance Generation for Flight Control Law Testing," *IFAC-PapersOnLine*, Vol. 52-16, pp. 730-734, 2019, doi: 10.1016/j.ifacol.2019.12.049.
- J. Z. Ben-Asher, J. Wetzler, E. Rimon, and J. Diepolder, "Optimal Trajectories for a Mobile Robot with Bounded Accelerations in the Presence of a Wall or a Bounded Obstacle," *27th Mediterranean Conference on Control and Automation (MED)*, Akko (Israel), pp. 481-488, 2019, doi: 10.1109/MED.2019.8798561.
- N. D. Botkin, K. Martynov, V. L. Turova, and J. Diepolder, "Generation of dangerous disturbances for flight systems," *Dynamic Games and Applications 9*, pp. 628–651, 2018, doi: 10.1007/s13235-018-0259-5.
- N. D. Botkin, J. Diepolder, and V. L. Turova, "Viability Approach to Simulation of an Adaptive Controller," *Radon Series on Computational and Applied Mathematics, 21*, Hamilton-Jacobi-Bellman Equations, pp. 43–60, 2018, doi: 10.1515/9783110543599-003.

- N. D. Botkin, V. L. Turova, J. Diepolder, and F. Holzapfel, "Computation of Viability Kernels on Grid Computers for Aircraft Control in Windshear," *Advances in Science, Technology and Engineering Systems Journal*, Vol. 3, No. 1, pp. 502-510, 2018, doi: 10.25046/aj030161.
- N. D. Botkin, V. L. Turova, J. Diepolder, M. Bittner, and F. Holzapfel, "Viability Approach to Aircraft Control in Windshear Conditions," *Advances in Dynamic and Mean Field Games*, Annals of the International Society of Dynamic Games, Vol. 15, pp. 325-343, 2017, doi: 10.1007/978-3-319-70619-1_15.
- B. Grüter, F. Sevilla, J. Diepolder, and F. Holzapfel, *"Parameter Envelope for Safe Emergency Landing Using Pareto Optimal Control," AIAA Scitech Forum*, AIAA 2019-0371, San Diego (US-CA), 2019, doi: 10.2514/6.2019-0371.
- B. Grüter, J. Diepolder, P. Piprek, T. Akman, and F. Holzapfel, "Nonlinear model predictive control for optimal aircraft sequencing," *31st Congress of the International Council of the Aeronautical Science (ICAS)*, Belo Horizonte (Brazil), 2018.
- T. Akman, J. Diepolder, B. Grüter, and F. Holzapfel, "Using Sensitivity Penalties to Robustify the Optimal Reentry Trajectory of a Hypersonic Vehicle," *31st Congress of the International Council of the Aeronautical Science (ICAS)*, Belo Horizonte (Brazil), 2018.
- K. Martynov, N. Botkin, V. Turova, and J. Diepolder, "Real-Time Control of Aircraft Take-Off in Windshear. Part II: Simulations and Model Enhancement," *25th Mediterranean Conference on Control and Automation (MED)*, pp. 285-290, 2017, doi: 10.1109/MED.2017.7984132.
- K. Martynov, N. Botkin, V. Turova, and J. Diepolder, "Real-Time Control of Aircraft Take-Off in Windshear. Part I: Aircraft Model and Control Schemes," *25th Mediterranean Conference on Control and Automation (MED)*, pp. 277-284, 2017, doi: 10.1109/MED.2017.7984131.
- B. Grüter, J. Diepolder, T. Akman, and F. Holzapfel, "Integrated Bi-level Arrival and Departure Time Assignment for Optimal Trajectories," *ENRI Int. Workshop on ATM/CNS (EIWAC)*, Nakano (Tokyo), 2017.
- B. Grüter, J. Diepolder, M. Bittner, M. Rieck, and F. Holzapfel, "Optimal Control-Based Altitude Profile Envelope for Emergency Landing," *4th CEAS Specialist Conference on Guidance, Navigation and Control (EuroGNC)*, Warsaw (Poland), 2017.
- J. Diepolder, S. Saboo, V. S. Akkinapalli, S. Raab, J. Zhang, P. Bhardwaj, M. Krenmayr, B. Grüter, and F. Holzapfel, "Flight Control Law Testing using Optimal Control and Postoptimal Sensitivity Analysis," *Advances in Aerospace Guidance, Navigation and Control*, pp. 25-45. doi: 10.1007/978-3-319-65283-2_2.

- B. Grüter, M. Bittner, M. Rieck, J. Diepolder, and F. Holzapfel, "Bi-level Homotopic Aircraft Sequencing Using Gradient-Based Arrival Time Assignment and Direct Optimal Control," *57th Israel Annual Conference on Aerospace Sciences (IACAS)*, Tel-Aviv and Haifa (Israel), 2017.
- J. Diepolder, C. D. Heise, M. Bittner, M. Rieck, B. Grüter, F. Holzapfel, and J. Z. Ben-Asher, "Optimal control based tracking error estimation for model reference adaptive control," *57th Israel Annual Conference on Aerospace Sciences (IACAS)*, Tel-Aviv and Haifa (Israel), 2017.
- M. Bittner, B. Grüter, J. Diepolder, F. Holzapfel, and J. Z. Ben-Asher, "High fidelity trajectory optimization using FALCON.m," *57th Israel Annual Conference on Aerospace Sciences (IACAS)*, Tel-Aviv and Haifa (Israel), 2017.
- J. Diepolder, P. Piprek, B. Grüter, T. Akman, and F. Holzapfel, "Aircraft Safety Analysis using Generalized Polynomial Chaos," *ENRI International Workshop on ATM/CNS 2017*, EIWAC, 2017, doi: 10.1007/978-981-13-7086-1_5.
- J. Diepolder, P. Piprek, N. Botkin, V. Turova, and F. Holzapfel, "A Robust Aircraft Control Approach in the Presence of Wind," *Australian and New Zealand Control Conference (ANZCC)*, Gold Coast (QLD, Australia), 2017, doi: 10.1109/ANZCC.2017.8298503.
- J. Diepolder, C. Göttlicher, B. Grüter, T. Akman, F. Holzapfel, and J. Z. Ben-Asher, "Optimal Control Based Flight Control Law Testing with Parameter Uncertainties," *IEEE Conference on Control Technology and Applications (CCTA)*, Mauna Lani (HI-US), 2017, doi: 10.1109/CCTA.2017.8062660.
- J. Diepolder, M. Bittner, P. Piprek, B. Grüter, and F. Holzapfel, "Facilitating aircraft optimal control based on numerical nonlinear dynamic inversion," *25th Mediterranean Conference on Control and Automation (MED)*, Valletta (Malta), pp. 141-146, 2017, doi: 10.1109/MED.2017.7984108.
- N. Botkin, V. Turova, J. Diepolder, M. Bittner, and F. Holzapfel, "Aircraft Control During Cruise Flight in Windshear Conditions: Viability Approach," *Dynamic Games and Applications*, Vol. 7, No. 4, pp. 594-608, 2017, doi: 10.1007/s13235-017-0215-9.
- M. Bittner, B. Grüter, J. Diepolder, and F. Holzapfel, "A Direct Solution Approach for Multi Timescale Optimal Control Problems," *Coupled Problems*, pp. 1121-1132, ISBN 978-84-946909-2-1, 2017.
- J. Diepolder, J. Z. Ben-Asher, A. C. Gabrys, S. P. Schatz, M. Bittner, M. Rieck, B. Grüter, and F. Holzapfel, "Flight Control Law Clearance Using Worst-Case Inputs," *30th International Congress of the International Council of the Aeronautical Sciences (ICAS)*, Daejeon (Korea), 2016.

- B. Grüter, M. Bittner, M. Rieck, J. Diepolder, and F. Holzapfel, "Optimal Sequencing in ATM Combining Genetic Algorithms and Gradient Based Methods to a Bilevel Approach," *30th International Congress of the International Council of the Aeronautical Sciences (ICAS)*, Daejeon (Korea), 2016.

List of Figures

List of Tables

Acronyms

ADMIRE	Aerodata Model in Research Environment
AoA	Angle-of-Attack
AoS	Angle-of-Sideslip
BFGS	Broyden-Fletcher–Goldfarb-Shanno
CG	Center of Gravity
CLP	Coin-or Linear Programming
COLIBRY	Counter Optimization Library
CPLEX	IBM ILOG CPLEX Optimization Studio
CSAS	Control and Stability Augmentation System
DE	Differential Evolution
DFOG	Distance Fields On Grids
DIRECT	Dividing Rectangles
DLG	Differential Legendre-Gauss
DLG-L	Differential Legendre-Gauss (Linear Defect)
DLG-N	Differential Legendre-Gauss (Nonlinear Defect)
DLGRF	Differential Legendre-Gauss-Radau (flipped)
DP	Dynamic Programming
ECEF	Earth Centered Earth Fixed
ECI	Earth Centered Inertial
FALCON.m	FSD Optimal Control Toolbox for MATLAB®
FBW	Fly-by-Wire
FSD	Flight System Dynamics
GA	Genetic Algorithm
gPC	Generalized Polynomial Chaos

GUROBI	Linear Programming Optimizer (Zonghao Gu, Edward Rothberg und Robert Bixby)
ILG	Integral Legendre-Gauss
ILGL	Integral Legendre-Gauss-Lobatto
IP	Interior Point
IPOPT	Interior Point Optimizer
KKT	Karush-Kuhn-Tucker
KNITRO	Nonlinear Interior Point Trust Region Optimization
LG	Legendre-Gauss
LGL	Legendre-Gauss-Lobatto
LGR	Legendre-Gauss-Radau
LGRF	Legendre-Gauss-Radau (flipped)
LICQ	Linear Independence Constraint Qualification
LOQO	Nonlinear Interior Point Solver
LP	Linear Program
LQR	Linear Quadratic Regulator
MAGO	Multi Dynamics Algorithm for Global Optimization
MATLAB®	MATrix LABoratory
MC	Monte Carlo
MRAC	Model Reference Adaptive Controller
NED	North-East-Down
NLP	Nonlinear Program
PIO	Pilot-induced Oscillation
PSO	Particle Swarm Optimization
QP	Quadratic Program
qpOASIS	Quadratic Programming Optimizer (Active Set)
SAS	Stability Augmentation System
SG	Surrogate Optimizer
SNOPT	Sparse Nonlinear Optimizer
SQP	Sequential Quadratic Programming
TR	Trust-Region
WGS84	World Geodetic System 1984

Symbols and Indices

Mathematical Preliminaries

B	B-spline basis function
I^t	Interval for independent variable
$P_p^{(\alpha,\beta)}$	Jacobi polynomial of degree p
$\mathbb{G}^B$	B-spline break-point grid
$\mathbb{G}^C$	Collocation grid
$\mathbb{G}^K$	B-spline knot grid
$\mathbb{G}^L$	Lagrange interpolation grid
$\mathbb{G}^Q$	B-spline query grid
$\mathbb{G}^Q$	Quadrature grid
$\mathbb{G}^S$	Segment grid
$\mathbb{G}^D$	Lagrange interpolation grid (collocation, differential form)
$\mathbb{G}^I$	Integration grid
$\mathbb{G}^{LGL}$	Legendre-Gauss-Lobatto quadrature grid
$\mathbb{G}^{LGRF}$	Flipped Legendre-Gauss-Radau quadrature grid
$\mathbb{G}^{LGR}$	Legendre-Gauss-Radau quadrature grid
$\mathbb{G}^{LG}$	Legendre-Gauss quadrature grid
$\mathbf{A}$	Integration matrix (Butcher-tableau)
$\mathbf{D}$	Differentiation matrix (collocation)
$\mathbf{S}$	Sensitivity matrix
$\mathbf{b}$	Quadrature weights (Butcher-tableau)
$\mathbf{c}$	Time points (Butcher-tableau)
$\mathbf{p}$	Parameter vector
$\mathbf{s}(t)$	Polynomial approximation

$\mathbf{s}$	Stage vector
$\mathbf{v}$	Value vector (B-splines)
$\mathbf{w}$	B-spline parameter vector
f	General function
h	Step length
l^D	Lagrange basis function (collocation, differential form)
l^I	Lagrange basis function (collocation, integral form)
l	Lagrange basis function
t_0	Initial time
t_f	Final time
t	Independent variable
w	Weight
x_0	Variable at initial point

Unconstrained Optimization

Δ	Trust-region radius
α	Line-search step length
β	Parameter in Armijo rule
$\mathbf{B}$	Quadratic model matrix
$\mathbf{H}$	Hessian approximation inverse
$\mathbf{M}$	Scaling matrix
$\mathbf{c}$	Quadratic model vector
$\mathbf{s}$	General direction
$\mathbf{y}$	Gradient difference
$\mathbf{z}$	Optimization variable vector
γ	Parameter in Wolf conditions
ρ	Quantity in BFGS update
$\tilde{\rho}$	Trust-region quotient
a	Quadratic model scalar
j	Cost function

Constrained Optimization

$\mathbf{G}$	Diagonal matrix of inequality constraints
$\mathbf{M}^a$	Diagonal matrix of actice inequality multipliers
$\mathbf{M}^i$	Diagonal matrix of inactice inequality multipliers
$\mathbf{M}_B$	Basis matrix
$\mathbf{M}_N$	Non-basis matrix
$\mathbf{M}_g$	Inequality constraint matrix
$\mathbf{M}_h$	Equality constraint matrix
$\mathbf{M}_s$	Equality constraint matrix standard form
$\mathbf{M}$	Diagonal matrix of inequality multipliers
$\mathbf{Z}$	Diagonal matrix of optimization variables
$\boldsymbol{\lambda}$	Lagrangian multipliers (equalities)
$\boldsymbol{\mu}^a$	Active inequality mutlipliers
$\boldsymbol{\mu}^i$	Inactive inequality mutlipliers
$\boldsymbol{\mu}_b$	Basic inequality multipliers
$\boldsymbol{\mu}_n$	Non-basic inequality multipliers
$\boldsymbol{\mu}$	Lagrangian multipliers (inequalities)
$\mathbf{b}_g$	Right hand-side (inequality constraints)
$\mathbf{b}_h$	Right hand-side (equality constraints)
$\mathbf{c}$	Cost vector
$\mathbf{d}$	General direction
$\mathbf{g}^a$	Active inequality constraints
$\mathbf{g}^i$	Inactive inequality constraints
$\mathbf{g}$	Inequality constraint vector
$\mathbf{h}$	Equality constraint vector
$\mathbf{s}$	Slack variables (inequality constraints)
$\mathbf{v}$	Vector of optmization variables and multipliers
$\mathbf{z}_b$	Basic optimization variables
$\mathbf{z}_n$	Non-basic optimization variables
$\mathbf{z}$	Optimization variable vector
$\mathcal{A}$	Active set
$\mathcal{B}$	Basic index set

$\mathcal{C}$	Critical cone
$\mathcal{F}$	Feasible region
$\mathcal{L}$	Lagrangian
$\mathcal{N}$	Non-basic index set
τ	Barrier parameter
j	Cost function
m	Merit function

Optimal Control

I^s	Singular interval
I^t	Time interval
I^τ	Normalized interval
$\bar{\mathbf{f}}$	Augmented and time scaled dynamic model
η	Jump height
$\boldsymbol{\gamma}$	Multiplier vector (boundary conditions)
$\boldsymbol{\lambda}_0$	Initial co-state vector
$\boldsymbol{\lambda}_f$	Final co-state vector
$\boldsymbol{\lambda}$	Co-state vector
$\boldsymbol{\phi}$	Boundary conditions
$\mathbf{c}$	Path-constraints
$\mathbf{f}$	Dynamic model
$\mathbf{p}$	Parameter vector
$\mathbf{u}$	Control vector
$\mathbf{x}_0$	Initial state vector
$\mathbf{x}_f$	Final state vector
$\mathbf{x}$	State vector
$\mathcal{H}^A$	Augmented Hamiltonian
$\mathcal{H}$	Hamiltonian
$\mathcal{S}$	Switching function
$\mathcal{U}$	Control set
μ	Multiplier (purely state dependent constraint)
τ_J	Jump location

τ	Normalized time
j^A	Augmented cost function
j^B	Bolza cost function
j^L	Lagrange cost function
j^M	Mayer cost function
l_0	Multiplier (cost)
t_0	Initial time
t_f	Final time
t	Time
x_{j^L}	Additional state for Lagrange cost function
x_{t_0}	Additional state for initial time parameter
x_{t_f}	Additional state for final time parameter

Direct Optimal Control

$\mathbb{G}^S$	Segment grid
$\mathbb{G}^B$	B-spline grid
$\mathbb{G}^C$	Path-constraint grid
$\mathbb{G}^I$	Integration grid
$\mathbf{S}_u$	Control sensitivity
$\mathbf{S}_x$	State sensitivity
$\boldsymbol{\alpha}$	Multiplier vector (boundary conditions)
$\boldsymbol{\kappa}$	Multiplier vector (defect constraints)
$\mathbf{d}$	Defect equality constraints
$\mathbf{n}$	Nodes (non-collocated states)
$\mathbf{s}_H$	State approximation (Hermite Simpson)
$\mathbf{s}_T$	State approximation (Trapezoidal)
$\mathbf{s}$	Stages (collocated states)
$\mathbf{z}$	Optimization variable vector
ν	Multiplier (collocation constraint)
τ_S	Normalized segment time grid point
h	Discretization step length
$l^\dagger$	Adjoint basis function

Models for Optimal Control Based Clearance of Flight Control Systems

C Aerodynamic coefficient

M_ϕ WGS84 curvature radius (meridian)

N_ϕ WGS84 curvature radius (prime vertical)

S Wing reference area

V Absolute velocity

Φ Roll angle

Ψ Yaw angle

Θ Pitch angle

α Angle-of-Attack

$\bar{\zeta}$ Damping coefficient

$\bar{c}$ Mean aerodynamic chord

$\bar{q}$ Dynamic pressure

β Angle-of-Sideslip

χ Course angle

δ Actuator command

η Elevator deflection

$\mathbf{F}$ Force vector

$\mathbf{I}$ Moment of inertia

$\mathbf{M}$ Moment vector

$\mathbf{R}$ Rotation matrix

$\mathbf{V}$ Velocity vector

$\boldsymbol{\omega}$ Rotational velocity vector

$\mathbf{n}$ Load factors

$\mathbf{r}$ Position vector

γ Climb angle

λ Longitude

ω_n Natural frequency

ϕ Latitude

ξ Aileron deflection

ζ Rudder deflection

a WGS84 semi-major axis

b	WGS84 semi-minor axis
b	Semi wing-span
e	WGS84 first eccentricity
f	WGS84 flattening
g	Gravitational constant
h	Geodetic height
k	Actuator input gain
m	Mass
p^*	Normalized roll rate
p	Roll rate
q^*	Normalized pitch rate
q	Pitch rate
r^*	Normalized yaw rate
r	Yaw rate
s	Wing-span
u	Velocity (x-component)
v	Velocity (y-component)
w	Velocity (z-component)
x_p	Actuator position
x_v	Actuator rate

Optimal Control Based Clearance for Linear and Quasi-linear Systems

I^t	Time interval
J	Cost function (Linear Program in standard form)
T_{SP}	Short-period time constant
V_K	Kinematic velocity
V_{ref}	Reference velocity
W	Polynomial space
ΔI_{xx}	Uncertainty related to moment of inertia around the x-axis (Body-fixed frame)
ΔI_{yy}	Uncertainty related to moment of inertia around the y-axis (Body-fixed frame)

ΔI_{zz}	Uncertainty related to moment of inertia around the z-axis (Body-fixed frame)
ΔL_β	Uncertainty related to Dihedral stability
ΔL_ξ	Uncertainty related to roll influence due to aileron deflection
ΔL_ζ	Uncertainty related to roll influence due to rudder deflection
ΔL_p	Uncertainty related to roll influence due to roll rate
ΔL_r	Uncertainty related to roll influence due to yaw rate
ΔM_α	Uncertainty related to pitch influence due to Angle-of-Attack
ΔM_η	Uncertainty related to pitch influence due to elevator deflection
ΔM_q	Uncertainty related to pitch influence due to pitch rate
ΔN_β	Uncertainty related to Weathercock stability
ΔN_ζ	Uncertainty related to yaw influence due to rudder deflection
ΔN_r	Uncertainty related to yaw influence due to yaw rate
$\Delta\hat{\mathbf{y}}$	Vector of scaling factors
Δm	Uncertainty related to mass
Δx^{RG}	Uncertainty related to center of gravity position
Ω	Random space
Φ_c	Roll angle command
$\Phi_{c,b}$	Boundary control (roll angle command)
Φ	Roll angle
Ψ	Yaw angle
Θ	Pitch angle
α_A	Aerodynamic Angle-of-Attack
α_K	Kinematic Angle-of-Attack
$\bar{\mathbf{v}}$	Cost vector (Linear Program in lower level)
$\bar{\mathbf{z}}$	Optimization variable vector (Linear Program in lower level)

$\bar{\zeta}_W$	Damping coefficient (wind dynamics)
$\bar{\zeta}$	Damping coefficient (actuator dynamics)
β_A	Aerodynamic Angle-of-Sideslip
β_K	Kinematic Angle-of-Sideslip
$\delta_{l,r}$	Kronecker delta
ϵ_0	Criterion tolerance
ϵ	Control penalty constant
$\mathbf{A}_d^+$	Discrete system matrix (positive identity matrix)
$\mathbf{A}_d^-$	Discrete system matrix (negative identity matrix)
$\mathbf{A}^I$	Inequality constraint matrix (inner approximation)
$\mathbf{A}^O$	Inequality constraint matrix (outer approximation)
$\mathbf{A}_c$	Constraint matrix (polytope)
$\mathbf{A}_d$	Discrete system matrix
$\mathbf{A}_{lat}$	System matrix (lateral)
$\mathbf{A}_{lon}$	System matrix (longitudinal)
$\mathbf{A}$	System matrix
$\mathbf{B}_d$	Discrete input matrix
$\mathbf{B}_{lat}$	Input matrix (lateral)
$\mathbf{B}_{lon}$	Input matrix (longitudinal)
$\mathbf{B}$	Input matrix
$\mathbf{F}$	Vector of optimality conditions
$\mathbf{M}_g$	Linear Programming inequality constraint matrix
$\mathbf{M}_h$	Linear Programming equality constraint matrix
$\mathbf{M}_n$	Diagonal matrix of non-basic inequality constraint multipliers
$\mathbf{M}_s$	Equality constraint matrix (Linear Program in standard form)
$\mathbf{M}$	Diagonal matrix of inequality constraint multipliers
$\mathbf{U}$	Controllability matrix
$\mathbf{Z}_b$	Diagonal matrix of basic optimization variables
$\mathbf{Z}$	Diagonal matrix of optimization variables
$\mathbf{\Sigma}$	Covariance matrix
$\boldsymbol{\eta}$	Multipliers of parameter embedding constraints

$\hat{\boldsymbol{\lambda}}_{\phi}$	Lagrangian multiplier vector corresponding to a ray constraint
$\boldsymbol{\lambda}$	Multiplier vector (equality constraints) / co-state vector
$\boldsymbol{\mu}_b$	Basic inequality constraint multipliers
$\boldsymbol{\mu}_n$	Non-basic inequality constraint multipliers
$\boldsymbol{\mu}$	Inequality constraint multipliers (Linear Program in standard form)
$\boldsymbol{\phi}$	Angles (spherical coordinates)
$\mathbf{a}_v$	System vector (actuator rate)
$\mathbf{a}_{a,p}$	System vector (aileron position)
$\mathbf{a}_{a,v}$	System vector (aileron rate)
$\mathbf{a}_{e,p}$	System vector (elevator position)
$\mathbf{a}_{e,v}$	System vector (elevator rate)
$\mathbf{a}_{r,p}$	System vector (rudder position)
$\mathbf{a}_{r,v}$	System vector (rudder rate)
$\mathbf{b}^I$	Right-handside vector (inner approximation)
$\mathbf{b}^O$	Right-handside vector (outer approximation)
$\mathbf{b}_c$	Right hand-side vector (polytope)
$\mathbf{b}_g$	Linear Programming right hand-side vector (inequality constraints)
$\mathbf{b}_h$	Linear Programming right hand-side vector (equality constraints)
$\mathbf{b}_s$	Equality constraint right hand-side vector (Linear Program in standard form)
$\mathbf{b}$	Column of input matrix
$\mathbf{c}_s$	Cost vector (Linear Program in standard form)
$\mathbf{c}_{\alpha_A}$	Output vector (aerodynamic Angle-of-Attack)
$\mathbf{c}_{\alpha_K}$	Output vector (kinematic Angle-of-Attack
$\mathbf{c}_{\beta_A}$	Output vector (aerodynamic Angle-of-Sideslip)
$\mathbf{c}_{\beta_K}$	Output vector (kinematic Angle-of-Sideslip)
$\mathbf{c}_{n_y,A}$	Output vector (lateral load factor, with wind)
$\mathbf{c}_{n_y,K}$	Output vector (lateral load factor, without wind)

$\mathbf{c}_{n_z,A}$	Output vector (normal load factor, with wind)
$\mathbf{c}_{n_z,K}$	Output vector (normal load factor, without wind)
$\mathbf{c}$	Output vector
$\mathbf{e}_i$	Basis vector (i-th)
$\mathbf{g}_p$	Parameter constraints
$\mathbf{h}$	Equality constraints (Linear Program in standard form)
$\mathbf{p}_0$	Nominal/mean parameter values
$\mathbf{p}$	Closed-loop system parameter vector
$\mathbf{r}$	Trim residuum vector
$\mathbf{u}^+$	Control vector (positive part)
$\mathbf{u}^-$	Control vector (negative part)
$\mathbf{u}_0$	Trim controls
$\mathbf{u}_{lat}$	Control vector (lateral dynamics)
$\mathbf{u}_{lon}$	Control vector (longitudinal dynamics)
$\mathbf{u}_m$	Mid-point control vector
$\mathbf{u}$	Closed-loop system input vector
$\mathbf{v}$	Cost vector
$\mathbf{x}_0$	Trim states
$\mathbf{x}_{lat}$	State vector (lateral dynamics)
$\mathbf{x}_{lon}$	State vector (longitudinal dynamics)
$\mathbf{x}_m$	Mid-point state vector
$\mathbf{x}$	Closed-loop system state vector
$\mathbf{y}^C$	Closest point
$\mathbf{y}^I$	Vertex of inner approximation
$\mathbf{y}^O$	Vertex of outer approximation
$\mathbf{y}^R$	Refinement point
$\mathbf{z}_b$	Basic optimization variables
$\mathbf{z}_n$	Non-basic optimization variables
$\mathbf{z}$	Optimization variable vector
γ	Expansion coefficient
$\hat{\mathbf{T}}$	Tangent space matrix

$\hat{\mathbf{n}}$	Normal vector
$\mathcal{C}$	Confidence interval
$\mathcal{H}^A$	Augmented Hamiltonian
$\mathcal{H}$	Hamiltonian
$\mathcal{I}_b$	Index set (bounded states)
$\mathcal{I}_n$	Index set (unbounded states)
$\mathcal{L}^A$	Augmented Lagrangian (Linear Program in standard form)
$\mathcal{L}$	Lagrangian (Linear Program in standard form)
$\mathcal{S}$	Switching function
$\mathcal{U}_{lat}$	Control set (lateral dynamics)
$\mathcal{U}_{lon}$	Control set (longitudinal dynamics)
$\mathcal{U}$	Admissible input set
$\mathcal{V}^I$	Set of points (inner approximation)
$\mathcal{V}^O$	Set of points (outer approximation)
$\mathcal{V}^R$	Set of refinement points
$\mathcal{X}$	State constraint set
Var	Variance
μ_p	Multiplier (actuator position constraint)
μ_v	Multiplier (actuator rate constraint)
μ	Mean
ω_W	Natural frequency (wind dynamics)
ω_n	Natural frequency (actuator dynamics)
ϕ	Joint polynomial basis function
ψ	Uni-variate polynomial basis function
ρ	Probability density function
σ	Standard deviation
$\tilde{\mathbf{p}}$	Embedded parameter vector
b_v	Input gain actuator rate
$b_{a,v}$	Input matrix entry (aileron rate)
$b_{e,v}$	Input matrix entry (elevator rate)
$b_{r,v}$	Input matrix entry (rudder rate)

b	Quadrature weight
c_p	Path-constraint for actuator position
c_v	Path-constraint for actuator rate
e_∞	Maximum scaled distance
e_Φ	Error controller state (Φ)
e_Ψ	Error controller state (Ψ)
e_{n_z}	Error controller state (n_z)
g_c	Constraint function (polytope)
$h(t)$	Geodetic height
h_l	Orthogonality constant
h_{ref}	Reference height
h	Discretization step-length
j	Output/criterion function
$n_{y,c,b}$	Boundary control (lateral load factor command)
n_y	Lateral load factor
$n_{z,c,b}$	Boundary control (normal load factor command)
$n_{z,c}$	Normal load factor command
n_z	Normal load factor
p_K	Roll rate
q_K	Pitch rate
r_K	Yaw rate
r	Magnitude (spherical coordinates)
t_0	Initial time
t_f	Final time
u_W	Wind velocity (x-component)
$u_{W,a}$	Wind acceleration (x-component)
u_b	Boundary control
u_s	Singular control
v_W	Wind velocity (y-component)
$v_{W,a}$	Wind acceleration (y-component)
w_W	Wind velocity (z-component)
$w_{W,a}$	Wind acceleration (z-component)

x_p	Actuator position state
x_v	Actuator rate state
$x_{a,p}$	Actuator position (Aileron)
$x_{a,v}$	Actuator rate (Aileron)
$x_{e,p}$	Actuator position (Elevator)
$x_{e,v}$	Actuator rate (Elevator)
$x_{r,p}$	Actuator position (Rudder)
$x_{r,v}$	Actuator rate (Rudder)
y	Criterion
j^{LL}	Lower level cost function
j^{LP}	Cost function of Linear Program

Optimal Control Based Clearance for Non-linear Systems

I^t	Time interval
K_r	Feed-forward gain
Γ	Learning rate
α_M	Angle-of-Attack (model)
α_P	Angle-of-Attack (plant)
$\mathbb{G}^{DFOG}$	Grid points DFOG method
$\mathbf{A}_M$	System matrix (model)
$\mathbf{A}_P$	System matrix (plant)
$\mathbf{B}_M$	Input matrix (model)
$\mathbf{B}_P$	Input matrix (plant)
$\boldsymbol{\Theta}^*$	MRAC ideal parameter vector
$\boldsymbol{\Theta}$	MRAC parameter vector
$\boldsymbol{\omega}$	Regressor vector
$\mathbf{e}_C$	Tracking error state vector
$\mathbf{j}$	Nonlinear output functions
$\mathbf{p}$	Nonlinear closed-loop system parameter vector
$\mathbf{t}$	Vector of switching time points
$\mathbf{u}$	Nonlinear closed-loop system input vector
$\mathbf{x}_0$	Initial state

$\mathbf{x}_M$	State vector (model)
$\mathbf{x}_P$	State vector (plant)
$\mathbf{x}$	Nonlinear closed-loop system state vector
λ_{M_α}	Multiplicative uncertainty (Pitch stiffness)
λ_{M_η}	Multiplicative uncertainty (Elevator effectiveness)
λ_{M_q}	Multiplicative uncertainty (Pitch damping)
$\mathcal{I}_b$	Index set (bounded states)
$\mathcal{U}$	Admissible input set
$\mathcal{X}$	State constraint set
e_α	Tracking error (Angle-of-Attack)
e_q	Tracking error (pitch rate)
j	Nonlinear output function
q_M	Pitch rate (model)
q_P	Pitch rate (plant)
r	Angle-of-Attack command
t_0	Initial time
t_f	Final time
u	Input (plant)
y	Nonlinear criterion

Indices

0	Initial, nominal
A	Aerodynamik
G	Gravitational
K	Kinematic
P	Propulsion
T	Total
f	Final
lb	Lower bound
max	Maximum
min	Minimum
ub	Upper bound

Notation

$\delta\square$	Variation; deviation from reference value
$\dot{\square}$	Derivative w.r.t. independent variable t
$\mathbf{0}$	Zero matrix of appropriate size
$\mathbf{1}$	All-ones matrix of appropriate size
$\mathbf{I}$	Identity matrix of appropriate size
$\hat{\square}$	Optimal quantitiy
$\square'$	Derivative w.r.t. normalized time τ
$\square, \blacksquare$	Placeholders
$\square^{[i]}$	Quantitiy associated to segment with index i
$diag(\square)$	Zero matrix with the components of $\square$ on the main diagonal

Bibliography

[1] C. Fielding and A. Varga, eds., *Advanced Techniques for Clearance of Flight Control Laws*. Lecture Notes in Control and Information Sciences, Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.

[2] K. M. Zuev, J. L. Beck, S.-K. Au, and L. S. Katafygiotis, "Bayesian post-processor and other enhancements of subset simulation for estimating failure probabilities in high dimensions," *Computers & Structures*, vol. 92-93, pp. 283–296, 2012.

[3] A. Varga, "Optimisation-based clearance," in *Advanced Techniques for Clearance of Flight Control Laws* (C. Fielding and A. Varga, eds.), vol. 283 of *Lecture Notes in Control and Information Sciences*, pp. 107–117, Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.

[4] J. Nocedal and S. J. Wright, *Numerical optimization: Chapter 2, Chapter 3, Chapter 4, Chapter 6, Chapter 13, Chapter 14, Chapter 16, Chapter 18, Chapter 19*. Springer series in operations research and financial engineering, New York: Springer, 2. ed., 2006.

[5] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization," *ACM Transactions on Mathematical Software*, vol. 23, no. 4, pp. 550–560, 1997.

[6] R. M. Lewis and V. Torczon, "Pattern search algorithms for bound constrained minimization," *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 1082–1099, 1999.

[7] M. J. D. Powell, "A direct search optimization method that models the objective and constraint functions by linear interpolation," in *Gomez S., Hennart JP. (eds) Advances in Optimization and Numerical Analysis. Mathematics and Its Applications*, vol. 7, pp. 51–67, 1994.

[8] A. Conn, K. Scheinberg, and P. Toint, "A derivative free optimization algorithm in practice," in *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, (Reston, Virigina), p. 83, American Institute of Aeronautics and Astronautics, 1998.

[9] P. J. M. van Laarhoven and E. H. L. Aarts, *Simulated annealing: Theory and applications*, vol. MA 37 of *Mathematics and its applications*. Dordrecht: Kluwer, 3. ed., 1992.

[10] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Boston: Addison-Wesley, 30 ed., 2012.

[11] W. Huyer and A. Neumaier, "Global optimization by multilevel coordinate search," *Journal of Global Optimization*, vol. 14, no. 4, pp. 331–355, 1999.

[12] A. Varga, "Optimisation-based clearance: The linear analysis," in *Advanced Techniques for Clearance of Flight Control Laws* (C. Fielding and A. Varga, eds.), vol. 283 of *Lecture Notes in Control and Information Sciences*, pp. 385–413, Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.

[13] L. Forssell and A. Sandblom, "Optimisation-based clearance: The nonlinear analysis," in *Advanced Techniques for Clearance of Flight Control Laws* (C. Fielding and A. Varga, eds.), vol. 283 of *Lecture Notes in Control and Information Sciences*, pp. 415–430, Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.

[14] P. P. Menon, J. Kim, D. G. Bates, and I. Postlethwaite, "Improved clearance of flight control laws using hybrid optimisation," in *2004 IEEE Conference on Cybernetics and Intelligent Systems*, (Piscataway, N.J.), pp. 677–682, IEEE, 2004.

[15] P. P. Menon, D. G. Bates, and I. Postlethwaite, "A deterministic hybrid optimisation algorithm for nonlinear flight control systems analysis," in *American Control Conference, 2006*, (Piscataway, NJ), pp. 333–338, IEEE Operations Center, 2006.

[16] K. V. Price, R. M. Storn, and J. A. Lampinen, *Differential evolution: A practical approach to global optimization*. Natural computing series, Berlin and New York: Springer, 2005.

[17] L. Forssell and U. Nilsson, "Admire the aero-data model in a research environment version 4.0, model description." 2005.

[18] P. P. Menon, J. Kim, D. G. Bates, and I. Postlethwaite, "Clearance of nonlinear flight control laws using hybrid evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 6, pp. 689–699, 2006.

[19] D. R. Jones, C. D. Perttunen, and B. E. Stuckman, "Lipschitzian optimization without the lipschitz constant," *Journal of Optimization Theory and Applications*, vol. 79, no. 1, pp. 157–181, 1993.

[20] D. Skoogh, P. Eliasson, F. Berefelt, R. Amiree, D. Tourde, and L. Forssell, "Clearance of flight control laws for time varying pilot input signals," *IFAC Proceedings Volumes*, vol. 42, no. 6, pp. 343–348, 2009.

[21] A. Varga, A. Hansson, and G. Puyou, eds., *Optimization based clearance of flight control laws: A civil aircraft application*, vol. 416 of *Lecture Notes in Control and Information Sciences*. Berlin and New York: Springer, 2012.

[22] H.-D. Joos, "Worst-case parameter search based clearance using parallel nonlinear programming methods," in *Optimization based clearance of flight control laws* (A. Varga, A. Hansson, and G. Puyou, eds.), vol. 416 of *Lecture Notes in Control and Information Sciences*, pp. 149–159, Berlin and New York: Springer, 2012.

[23] H.-D. Joos, "Application of parallel nonlinear programming methods for worst-case parameter search," in *Optimization based clearance of flight control laws* (A. Varga, A. Hansson, and G. Puyou, eds.), vol. 416 of *Lecture Notes in Control and Information Sciences*, pp. 233–252, Berlin and New York: Springer, 2012.

[24] D. Skoogh and F. Berefelt, "Application of nonlinear programming methods for determination of worst-case pilot inputs," in *Optimization based clearance of flight control laws* (A. Varga, A. Hansson, and G. Puyou, eds.), vol. 416 of *Lecture Notes in Control and Information Sciences*, pp. 299–315, Berlin and New York: Springer, 2012.

[25] D. Skoogh and F. Berefelt, "Nonlinear programming methods for worst-case pilot input determination," in *Optimization based clearance of flight control laws* (A. Varga, A. Hansson, and G. Puyou, eds.), vol. 416 of *Lecture Notes in Control and Information Sciences*, pp. 203–217, Berlin and New York: Springer, 2012.

[26] G. Puyou, R. F. de Oliveira, and A. Berard, "Evaluation of clearance techniques in an industrial context," in *Optimization based clearance of flight control laws* (A. Varga, A. Hansson, and G. Puyou, eds.), vol. 416 of *Lecture Notes in Control and Information Sciences*, pp. 319–357, Berlin and New York: Springer, 2012.

[27] R. Rodriguez Robles, M. Sabaris Boullosa, and F. Asensio Nieto, "Flight control laws carefree handling clearance of a highly manoeuvrable aircraft using multi-strategy adaptive global optimization." 2017.

[28] J. A. Hernández and J. D. Ospina, "A multi dynamics algorithm for global optimization," *Mathematical and Computer Modelling*, vol. 52, no. 7-8, pp. 1271–1278, 2010.

[29] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *1995 IEEE International Conference on Neural Networks Proceedings*, pp. 1942–1948, IEEE, 1995.

[30] S.-M. Guo, J. S.-H. Tsai, C.-C. Yang, and P.-H. Hsu, "A self-optimization approach for l-shade incorporated with eigenvector-based crossover and successful-parent-selecting framework on cec 2015 benchmark set," in *IVCNZ 2015* (I. C. o. I. Zealand and V. C. New, eds.), (Piscataway, NJ), pp. 1003–1010, IEEE, 2015.

[31] A. A. Herrmann and J. Z. Ben-Asher, "Flight control law clearance using optimal control theory," *Journal of Aircraft*, vol. 53, no. 2, pp. 515–529, 2016.

[32] D. T. McRuer, "Pilot-induced oscillations and human dynamic behavior: Nasa contractor report 4683." NASA Contractor Report 4683, 1995.

[33] R. Baier, M. Gerdts, and I. Xausa, "Approximation of reachable sets using optimal control algorithms," *Numerical Algebra, Control and Optimization*, vol. 3, no. 3, pp. 519–548, 2013.

[34] R. M. Rieck, *Discrete Controls and Constraints in Optimal Control Problems*. Dissertation, Technische Universität München, München, 2017.

[35] M. Bittner, *Utilization of Problem and Dynamic Characteristics for Solving Large Scale Optimal Control Problems*. Dissertation, Technische Universität München, München, 2017.

[36] H. Rutishauser, *Lectures on Numerical Mathematics*. Boston, MA: Birkhäuser Boston, 1990.

[37] J.-P. Berrut and L. N. Trefethen, "Barycentric lagrange interpolation," *SIAM Review*, vol. 46, no. 3, pp. 501–517, 2004.

[38] C. Runge, "Über empirische funktionen und die interpolation zwischen äquidistanten ordinaten," in *Zeitschrift für Mathematik und Physik*, vol. 46, pp. 224–243.

[39] C. F. Gauß, "Methodus nova integraliam valores per approximationem inveniendi werke 3," *163-196*, vol. 1814.

[40] L. T. Biegler, *Nonlinear programming: Concepts, algorithms, and applications to chemical processes*. Philadelphia, Pa.: Society for Industrial and Applied Mathematics (SIAM 3600 Market Street Floor 6 Philadelphia PA 19104), 2010.

[41] G. Szegö, *Orthogonal polynomials*, vol. 23 of *Colloquium Publications*. Providence, Rhode Island: American Mathematical Society, 4. ed., 1939.

[42] W. Kutta, *Beitrag zur näherungsweisen Integration totaler Differentialgleichungen*. B.G Teubner, 1901.

[43] J. C. Butcher, *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons, 2. ed., 2016.

[44] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations I. Nonstiff Problems*, vol. 8. Berlin, Heidelberg: Springer Berlin Heidelberg, 1993.

[45] H. Brunner, *Collocation methods for Volterra integral and related functional differential equations*, vol. 15 of *Cambridge monographs on applied and computational mathematics*. Cambridge UK and New York: Cambridge University Press, 2004.

[46] K. Wright, "Some relationships between implicit runge-kutta, collocation and lanczos τ methods, and their stability properties," *BIT*, vol. 10, no. 2, pp. 217–227, 1970.

[47] D. Garg, W. Hager, and A. Rao, "Gauss pseudospectral method for solving infinite-horizon optimal control problems," in *AIAA Guidance, Navigation and Control Conference*, (Reston, VA), p. 269, American Institute of Aeronautics and Astronautics, 2010.

[48] D. Garg, M. Patterson, W. Hager, A. Rao, D. Benson, and G. Huntington, "An overview of three pseudospectral methods for the numerical solution of optimal control problems," 2017.

[49] D. Garg, M. A. Patterson, C. Francolin, C. L. Darby, G. T. Huntington, W. W. Hager, and A. V. Rao, "Direct trajectory optimization and costate estimation of finite-horizon and infinite-horizon optimal control problems using a radau pseudospectral method," *Computational Optimization and Applications*, vol. 49, no. 2, pp. 335–358, 2011.

[50] J. T. Betts, *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*. Society for Industrial and Applied Mathematics, 2010.

[51] L. Piegl and W. Tiller, *The NURBS Book*. Monographs in Visual Communications, Berlin and Heidelberg: Springer, 1995.

[52] M. Gerdts, *Optimal Control of ODEs and DAEs*. De Gruyter textbook, Berlin: De Gruyter, 2012.

[53] A. V. Fiacco and G. P. McCormick, *Nonlinear programming: Sequential unconstrained minimization techniques*, vol. 4 of *Classics in applied mathematics*. Philadelphia, Pa: Society for Industrial and Applied Mathematics (SIAM 3600 Market Street Floor 6 Philadelphia PA 19104), 1990.

[54] J. Z. Ben-Asher, *Optimal control theory with aerospace applications*. AIAA education series, Reston, VA: American Institute of Aeronautics and Astronautics, 2010.

[55] G. Dantzig, *Linear Programming and Extensions*. RAND Corporation, 1963.

[56] H. J. Greenberg, "The use of the optimal partition in a linear programming solution for postoptimal analysis," *Operations Research Letters*, vol. 15, no. 4, pp. 179–185, 1994.

[57] IBM, "Ilog cplex optimization studio: Cplex user's manual 12.8," 2017.

[58] L. L. Gurobi Optimization, "Gurobi optimizer reference manual," 2020.

[59] J. Forrest, D. de La Nuez, and R. Lougee-Heimer, "Clp user guide," 2004.

[60] H. J. Ferreau, C. Kirches, A. Potschka, H. G. Bock, and M. Diehl, "qpoases: a parametric active-set algorithm for quadratic programming," *Mathematical Programming Computation*, vol. 6, no. 4, pp. 327–363, 2014.

[61] P. E. Gill, W. Murray, and M. A. Saunders, "Snopt: An sqp algorithm for large-scale constrained optimization," *SIAM Journal on Optimization*, vol. 12, no. 4, pp. 979–1006, 2002.

[62] A. Wächter, *An Interior Point Algorithm for Large-Scale Nonlinear Optimization with Applications in Process Engineering*. PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania, 2002.

[63] R. H. Byrd, J. Nocedal, and R. A. Waltz, "Knitro: An integrated package for nonlinear optimization," in *Large-scale nonlinear optimization* (G. Di Pillo, ed.), vol. 83 of *Nonconvex optimization & its applications*, pp. 35–59, New York: Springer, 2006.

[64] R. J. Vanderbei, "Loqo:an interior point code for quadratic programming," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 451–484, 1999.

[65] C. Büskens, *Optimierungmethoden und Sensitivitätsarialyse ftir optimale Steuerprozesse mit Steuer- und Zustands-Beschränkungen*. PhD thesis, Westfälische Wilhelms-Universität, Münster, 1998.

[66] E. Castillo, A. J. Conejo, R. Mínguez, and C. Castillo, "A closed formula for local sensitivity analysis in mathematical programming," *Engineering Optimization*, vol. 38, no. 1, pp. 93–112, 2006.

[67] H. Oberle and W. Grimm, "Bndsco: A program for the numerical solution of optimal control problems," 1989.

[68] D. J. Bell, "Lie brackets and singular optimal control," *IMA Journal of Mathematical Control and Information*, vol. 1, no. 1, pp. 83–94, 1984.

[69] R. F. Hartl, S. P. Sethi, and R. G. Vickson, "A survey of the maximum principles for optimal control problems with state constraints," *SIAM Review*, vol. 37, no. 2, pp. 181–218, 1995.

[70] A. E. Bryson and Y.-C. Ho, *Applied optimal control*. Hemisphere, rev. print ed., 1975.

[71] S. Chang, "Optimal control in bounded phase space," *Automatica*, vol. 1, no. 1, pp. 55–67, 1963.

[72] C. L. Darby, W. W. Hager, and A. V. Rao, "An hp-adaptive pseudospectral method for solving optimal control problems," *Optimal Control Applications and Methods*, vol. 32, no. 4, pp. 476–502, 2011.

[73] Jianli Wei, Xiaojun Tang, and Jie Yan, "Costate estimation for a multiple-interval pseudospectral method using collocation at the flipped legendre-gauss-radau points," *IEEE/CAA Journal of Automatica Sinica*, pp. 1–15, 2017.

[74] D. Benson, *A Gauss pseudospectral transcription for optimal control*. Phd thesis, Massachusetts Institute of Technology, 2005.

[75] F. Holzapfel, "Lecture notes flight controls 1," 2019.

[76] F. Holzapfel, *Nichtlineare adaptive Regelung eines unbemannten Fluggerätes*. Dissertation, Technische Universität München, München, 2004.

[77] R. Brockhaus, W. Alles, and R. Luckner, *Flugregelung*. Berlin: Springer Berlin, 3., neu bearbeitete aufl. ed., 2010.

[78] "Mil-f-8785c: Military specification, flying qualities of piloted airplanes," 1980.

[79] "Cs-23: Certification specifications for normal-category aeroplanes: Amendment 5," 2017.

[80] "Cs-25: Certification specifications and acceptable means of compliance for large aeroplanes: Amendment 24," 2020.

[81] Andreas Varga, "Clearance criteria for civil aircraft - ast5-ct-2006-030768," *Technical Report*, vol. 2010.

[82] J. Diepolder, J. Z. Ben-Asher, A. C. Gabrys, S. P. Schatz, M. Bittner, M. Rieck, B. Grüter, and F. Holzapfel, "Flight control law clearance using worst-case inputs," in *ICAS 30th International Congress of the International Council of the Aeronautical Sciences*, 2016.

[83] "Mil-std-1797a: Department of defense handbook, flying qualities of piloted aircraft," 1995.

[84] J. Diepolder, J. Z. Ben-Asher, and F. Holzapfel, "Flight control law clearance using worst-case inputs under parameter uncertainty," *AIAA Journal*, vol. 43, no. 10, pp. 1967–1974, 2020.

[85] J. Diepolder, J. Z. Ben-Asher, P. Piprek, and F. Holzapfel, "Optimal control based flight control law clearance using generalized polynomial chaos," in *60th Israel Annual Conference 2020*.

[86] T. Gal, "Linear parametric programming—a brief survey," in *Sensitivity, Stability and Parametric Analysis* (A. V. Fiacco, ed.), Mathematical Programming Studies, pp. 43–68, Berlin and Heidelberg: Springer, 1984.

[87] I. Adler and R. D. C. Monteiro, "A geometric view of parametric linear programming," *Algorithmica*, vol. 8, no. 1, pp. 161–176, 1992.

[88] R. D. C. Monteiro and S. Mehrotra, "A general parametric analysis approach and its implication to sensitivity analysis in interior point methods," *Mathematical Programming*, vol. 72, no. 1, pp. 65–82, 1996.

[89] M. Bjorkman and K. Holmström, "Global optimization using the direct algorithm in matlab," *Advanced Modeling and Optimization*, vol. 1, 2004.

[90] M. Slotani, "Tolerance regions for a multivariate normal population," *Annals of the Institute of Statistical Mathematics*, vol. 16, no. 1, pp. 135–153, 1964.

[91] D. Xiu and G. E. Karniadakis, "The wiener–askey polynomial chaos for stochastic differential equations," *SIAM Journal on Scientific Computing*, vol. 24, no. 2, pp. 619–644, 2002.

[92] D. Xiu, "Fast numerical methods for stochastic computations: A review," *Communications in Computational Physics*, vol. 5, pp. 242–272, 2009.

[93] D. Xiu, *Numerical methods for stochastic computations: A spectral method approach*. Princeton, N.J.: Princeton University Press, 2010.

[94] N. T. Thomopoulos, *Essentials of Monte Carlo Simulation: Statistical Methods for Building Simulation Models*. New York, NY: Springer New York and Imprint: Springer, 1. ed., 2013.

[95] D. J. Sheskin, *Handbook of parametric and nonparametric statistical procedures*. Boca Raton: Chapman & Hall/CRC, 4. ed., 2009.

[96] S. He, J. Zhang, and S. Zhang, "Bounding probability of small deviation: A fourth moment approach," *Math. Oper. Res.*, vol. 35, pp. 208–232, 2010.

[97] M. Kvasnica, P. Grieder, M. Baotić, and M. Morari, "Multi-parametric toolbox (mpt)," in *Hybrid systems* (R. Alur and G. J. Pappas, eds.), vol. 2993 of *Lecture Notes in Computer Science*, pp. 448–462, Berlin and New York: Springer-Verlag, 2004.

[98] M. Kvasnica, P. Grieder, and M. Baotić, "Multi-parametric toolbox (mpt)," 2004.

[99] A. B. Kurzhanski and P. Varaiya, "Ellipsoidal techniques for reachability analysis," in *Hybrid systems* (N. Lynch, ed.), vol. 1790 of *Lecture Notes in Computer Science*, pp. 202–214, Berlin: Springer, 2000.

[100] A. A. Kurzhanskiy and P. Varaiya, "Ellipsoidal toolbox." 2006.

[101] A. Girard, "Reachability of uncertain linear systems using zonotopes," in *Hybrid Systems: Computation and Control* (M. Morari, F. Rossi, and L. Thiele, eds.), vol. 3414 of *Lecture Notes in Computer Science*, pp. 291–305, New York: Springer-Verlag Berlin/Heidelberg, 2005.

[102] D. Althoff, M. Althoff, and S. Scherer, "Online safety verification of trajectories for unmanned flight with offline computed robust invariant sets," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (W. Burgard, ed.), (Piscataway, NJ), pp. 3470–3477, IEEE, 2015.

[103] E. K. Kostousova, "Control synthesis via parallelotopes: optimzation and parallel compuations," *Optimization Methods and Software*, vol. 14, no. 4, pp. 267–310, 2001.

[104] E. K. Kostousova, "On tight polyhedral estimates for reachable sets of linear differential systems," in *AIP Conference Proceedings*, vol. 1493, pp. 579–586, 2012.

[105] P. Varaiya, "Reach set computation using optimal control," in *Verification of digital and hybrid systems* (S. P. Davies, ed.), pp. 323–331, Springer, 2012.

[106] R. Baier, C. Büskens, I. A. Chahma, and M. Gerdts, "Approximation of reachable sets by direct solution methods for optimal control problems," *Optimization Methods and Software*, vol. 22, no. 3, pp. 433–452, 2007.

[107] A. B. Kurzhanski and P. Varaiya, "Ellipsoidal techniques for reachability under state constraints," *SIAM Journal on Control and Optimization*, vol. 45, no. 4, pp. 1369–1394, 2006.

[108] L. E. Blumenson, "A derivation of n-dimensional spherical coordinates," *The American Mathematical Monthly*, vol. 67, no. 1, pp. 63–66, 1960.

[109] J. Matt, "Matlab central: Analyze n-dimensional polyhedra in terms of vertices or (in)equalities," Retrieved May 19, 2020.

[110] J. Diepolder, P. Piprek, B. Grüter, T. Akman, and F. Holzapfel, "Aircraft safety analysis using generalized polynomial chaos," in *ENRI International Workshop on ATM/CNS*, pp. 67–81, 2017.

[111] J. Diepolder, C. Gottlicher, B. Gruter, T. Akman, F. Holzapfel, and J. Z. Ben-Asher, "Optimal control based flight control law testing with parameter uncertainties," in *2017 IEEE Conference on Control Technology and Applications (CCTA)*, pp. 1432–1437, IEEE, 2017.

[112] J. Diepolder, C. Heise, M. Bittner, M. Rieck, B. Grüter, F. Holzapfel, and J. Z. Ben-Asher, "Optimal control based tracking error estimation for model reference adaptive control," in *57th Israel Annual Conference on Aerospace Sciences, IACAS 2017*, 2017.

[113] J. Diepolder, S. Saboo, V. S. Akkinapalli, S. Raab, J. Zhang, P. Bhardwaj, M. Krenmayr, B. Grüter, and F. Holzapfel, "Flight control law testing using optimal control and postoptimal sensitivity analysis," in *Dolega, Glebocki et al. (Hg.) – Advances in Aerospace Guidance*, pp. 25–45, 2017.

[114] D. P. Bertsekas, *Dynamic programming and optimal control - Vol I*. Belmont, Mass.: Athena Scientific, 4. ed., 2017.

[115] I. M. Mitchell, "Comparing forward and backward reachability as tools for safety analysis," in *HSCC*, 2007.

[116] O. Bokanowski, N. Forcadel, and H. Zidani, "Reachability and minimal times for state constrained nonlinear problems without any controllability assumption," *SIAM Journal on Control and Optimization*, vol. 48, no. 7, pp. 4292–4316, 2010.

[117] I. M. Mitchell, "The flexible, extensible and efficient toolbox of level set methods," *Journal of Scientific Computing*, vol. 35, no. 2-3, pp. 300–329, 2008.

[118] A. M. Bayen, I. M. Mitchell, M. M. K. Oishi, and C. J. Tomlin, "Aircraft autolander safety analysis through optimal control-based reach set computation," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 1, pp. 68–77, 2007.

[119] T. Lombaerts, S. Schuet, K. Wheeler, D. M. Acosta, and J. Kaneshige, "Safe maneuvering envelope estimation based on a physical approach," in *Guidance, Navigation, and Control and Co-located Conferences*, 2013.

[120] C. D. Heise, *Survivable Flight Control with Guaranteed Stability and Performance Characteristics*. Dissertation, Technische Universität München, München, 2017.

[121] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Computer J*, vol. 7, pp. 308–313, 1965.

[122] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the nelder–mead simplex method in low dimensions," *SIAM Journal on Optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[123] J. D'Errico, "Matlab central: fminsearchbnd, fminsearchcon," Retrieved May 31, 2020.

[124] D. P. Bertsekas, *Dynamic programming and optimal control - Vol II*. Belmont, Mass.: Athena Scientific, 4. ed., 2017.