Interview Article   **Open Access**

Fabrizio Durante, Giovanni Puccetti*, and Matthias Scherer

A Journey from Statistics and Probability to Risk Theory

An interview with Ludger Rüschendorf

Ludger Rüschendorf is Professor of Mathematics at the University of Freiburg. He completed his PhD in 1974 at the University of Hamburg, where he specialized in asymptotic statistics; he then obtained his habilitation in 1979 at TH Aachen, based on work related to stochastic orderings. He held positions in Freiburg and Münster before succeeding Herman Witting as Statistics Chair in Freiburg in 1993. Besides asymptotic statistics and stochastic orderings, he is interested in dependence modeling and its connections with mass transportation problems, the stochastic analysis of algorithms, mathematical risk analysis, mathematical finance and various areas in applied probability, including optimal stopping and asymptotics for random graphs and networks. He wrote several monographs and textbooks in these areas, including a two-volume book on mass transportation in 1998 with S.T. Rachev, a research monograph on mathematical risk analysis in 2013, and textbooks in statistics and asymptotic statistics. He has published around 200 scientific papers in various areas and has collaborated with many leading scientists internationally. He is an elected fellow of the Institute of Mathematical Statistics and was the first president of the German Stochastics Association from 1994 to 1996; he was also Pauli Fellow in 2009–10 at the Pauli Institute of the University of Vienna.

1 Mathematical education

With this interview, *Dependence Modeling* continues its series of *Interview Articles*. These articles are published on a semiannual basis; they feature discussions with personalities who have played a pivotal role in the field of dependence modeling. We believe that students at all levels, teachers, scientists as well as practitioners can benefit from the following virtual meeting with Ludger Rüschendorf, a well known and influential personality in the fields of mass transportation, mathematical finance and stochastic analysis, among many others. In the following, our questions to Ludger Rüschendorf are typed in bold-face.

Let's start from the very beginning. Please share with us some memories from your youth.

I grew up in a small rural village named Rüschendorf, in a farmer's family which resided there for hundreds of years. I attended primary school in the village — a "dwarf school". Four grades were being taught in the same room by a single teacher. While one group was being taught, the other three were given exercises to work on. Pupils who finished their exercises early were allowed to participate in the ongoing class, and I enjoyed doing that. Next, I attended a middle school or *Realschule* in Damme, a slightly larger town located nearby; I was there until Grade 8. I did quite well and hence was recommended for higher studies at

*Corresponding Author: Giovanni Puccetti: Department of Economics, Management and Quantitative Methods, University of Milan, Italy. E-mail: giovanni.puccetti@unimi.it
Fabrizio Durante: Faculty of Economics & Management, Free University of Bozen/Bolzano, Italy.
Matthias Scherer: Chair of Mathematical Finance, Technische Universität München, Germany.

a *Gymnasium* (high-school) which was located about 30 km away from home; as a result, I ended up living in a catholic residence. I had to catch up two years of latin, which I liked a lot along with mathematics, but I had more trouble with other topics such as German literature and foreign languages. I enjoyed being together with other pupils, had an intensive religious education, learned discipline and practiced a lot of sports. In particular we played soccer nearly every day.

If you recall your high school education, what aspects were most influential in your choice to study mathematics in university?

I was always quite good at mathematics. I appreciated the clear and precise way in which problems were formulated. I found that in real life, and particularly in rural life, the cause of conflicts was often the imprecise formulation of problems. This was also the reason why I later preferred mathematics to physics. In my physics classes, we often could not see clearly the distinction between assumptions and conclusions. At school I was not particularly attracted or stimulated to read math textbooks, although I had a good and dedicated math teacher.

As a father of five children, you witnessed many changes in the German education system (which varies across federal states). How would you comment on these changes, especially as regards mathematics and natural sciences?

It is clear that nowadays, teachers generally get a better education than in former times, but this doesn't necessarily mean that they are more effective teachers than those of the past. Today, school teachers are confronted with a large number of pupils who are sometimes lacking a solid education. Many students are not so much interested in school topics; they are often distracted by all kinds of amusement, mobile phones, and the internet. So today's education brings with it great challenges. The old system of more practically oriented *Realschulen*, *Gymnasiums* and *Hauptschulen*, seemed better suited for most pupils. It is rather doubtful that the changes that were introduced in pedagogic approaches (e.g., the teacher being viewed not as an instructor but rather as a companion of pupils, pupil-oriented teaching, awakening of interest by organizing math events, and math-science teaching through experimental methods, or small-group teaching incorporating media, etc.) will lead to deeper understanding. Four of my children attended a Waldorf school; some of them enjoyed its free atmosphere, but some of them also missed the benefits of strict rules and regular exercises, competition, grading, etc.

At what point in your life did you decide to pursue an academic career?

I was lucky enough that at the end of my studies, after presenting my diploma thesis in 1971–72 on "Nonparametric tests for stochastic processes", Herman Witting, a nice man with a strong personality who was also a very influential statistician, proposed a thesis topic in asymptotic statistics. This connection with Witting, who was a major figure in building up the statistics program within the mathematics faculties in Germany, greatly influenced my development and my career. He had many good students; several of them had academic careers. Through his connections, I was fortunate to get regular invitations to Oberwolfach meetings in probability and statistics, typically once or twice a year; there, I had the opportunity to meet leading probabilists and statisticians and to get acquainted with the most recent developments. I was asked by Olaf Krafft in early 1972 to join him on an Assistant Position at his new Chair in Hamburg, where he moved from Münster.

I was happy to work in asymptotic statistics and also that my thesis adviser, who moved to Freiburg a little later, was not too close; this way, I had more freedom to follow my own interests. The subject of experimental design which Olaf Krafft started to explore did not arouse my interest all that much, even if it gave me the impulse later on to write a few papers on combinatorial topics in connection with balanced incomplete block designs (BIBD's). I found, for example, a generalization of Hall's Theorem (also called the "Marriage Theorem"; see [17]) to "harem-like" situations, and continuous versions of the Marriage Theorem using Lyapounov's Convexity Theorem, as well as relations to multi-matroids and to partition problems.

My PhD, which I completed in 1974, was on empirical processes for mixing (i.e., weakly dependent) random sequences and their applications in asymptotic statistics. Empirical processes were a new and very popular method at that time for determining the asymptotic distribution of test statistics by approximating them using functionals of the empirical process (e.g., the Pyke–Shorack approach or Hoeffding projection). For the PhD, one had to undergo an exam in one's broad area of research (in my case in applied mathematics, more precisely probability), but also one exam in a pure math subject. I had arranged with Professor Helene

Braun, a well known algebraist, to be examined on the three volumes of Choquet's analysis, with a focus on integral representation theory. During my exam I was very surprised that she did not know these topics but instead insisted on how and when Herrmann Minkowski formulated his early, specific representation result, even though this was merely glossed over in the book. Nevertheless, it seems that I did not make too bad an impression on her.

In 1976, Olaf Krafft gave me the opportunity to teach my first course on a topic of my choice: the application of stochastic models in biology and medicine. A great source for that was the recent book of Iosifescu and Tautu [20] and I learned many interesting methods and results such as the Karlin–McGregor analysis on continuous time Markov chains (see, e.g., [21]), branching processes and many more. I drew a lot of motivation from this class. It showed me that probability has important applications and I learned a lot about biological and medical concepts, in particular on growing cell models, haematopoese, gene formation processes, etc. It is a good experience to do self determined teaching early on and to be backed up by one's mentor in face of critical remarks by students who complain about the lack of pedagogical or didactic concepts in the class.

After my PhD, I worked on some applications and joined Olaf Krafft at the TH Aachen. I wanted to do research in a new area, because I thought I knew enough by then about empirical processes, so I wanted to explore a new field. After spending much time reading various sources like Grenander's book on probability on algebraic structures [16] or Linnik's very interesting book [24] on analytical statistics, I focused on stochastic integration theory, which was just being developed at the time. I was mainly motivated by connections with the study of statistical tests for stochastic processes arising in stochastic exponentials or in Girsanov's Theorem. Introductory texts were not available, so I had to read articles for nearly one year before I even began to form an idea of the broad picture. In a mood of frustration, I applied to the DAAD (German Exchange Program) for a position in Brazil and started learning Portuguese. When it became clear to me that my contract with the university in Fortaleza would be for five years, and that a subsequent return to the German university system was improbable, I declined the position and decided to try to pursue a university career in Germany.

Although the study of stochastic analysis was still very much under development at the time, it gave me a good basis later on for understanding the development of continuous time models and analysis, especially in connection with financial mathematics.

In 1979 I submitted my habilitation thesis, which had to do with stochastic orderings of probability measures induced by function classes. As part of this work, I found a basic dual representation of an optimization problem in marginal classes. While my motivation for this result came from the description of the possible influence of dependence, it turned out that this result also was an extension of the result of Kantorovich for the mass-transportation problem. My paper motivated Hans Kellerer to develop an important further extension in 1984; see [22]. After I earned my first professorship in 1981 in Freiburg, I moved to Münster in 1987 and in 1993 got a professorship in Freiburg again, as Hermann Witting's successor.

Who influenced your scientific life the most?

I was impressed by Hans Kellerer's mathematical ideas and work, particularly in connection with marginal problems, the Fréchet problem and on the general form of the duality theorem. I was very impressed and learned a lot from Albert Shiryaev and Marc Yor, especially about semi-martingales, limit theory for stochastic processes, Brownian motion and many other topics. They visited our institute quite frequently and gave inspiring courses and talks; we also had regular discussions about science and life in general. Albert Shiryaev started visiting us in the 1980s, Marc Yor in more recent years. Their inspiring and generous personalities enriched the scientific life at our institute.

I was also impressed by the mathematical and personal qualities of Hans Föllmer, by his ability to explain complex notions in an understandable way. His great mathematical work and understanding of random fields and early work on financial mathematics influenced me a lot. As a young researcher I had the great privilege and pleasure to organize with him an Oberwolfach workshop; I enjoyed his charming personality and respectful way of approaching colleagues and students.

Over a long time — in fact, during most of my scientific career — I was in close contact with Helmut Strasser. We had many fruitful discussions about asymptotic statistics; it is through him and his great book [43] in 1985, that I came to understand much of the basic developments in LeCam's theory. He gave a crystal-clear exposition of the basic principles. This influenced my work in asymptotic statistics as for example on

Figure 1: Ludger Rüschendorf with his wife Gabi, his children and Zari Rachev at the Aasee in Münster, 1987.

a general functional δ-method in Banach spaces with applications to statistical problems in point process models, which Holtrode [19] and I investigated, along with applications to the efficiency of Kaplan–Meyer or Nelsen–Aalen estimators. A similar theory was developed nearly at the same time by Andersen, Borgan, Gill and Keiding in their very influential book, published in 1993; see [1].

The greatest influence on my scientific life came from my frequent contact, friendship and cooperation with Svetloszar (Zari) Rachev (see Figure 1). I learned from him the method of probability metrics, much about stable models and their use in finance. We developed at an early stage pricing formulas for interesting financial models like stable or hyperbolic models. I was impressed by his mathematical talent, by the way in which he worked and pursued his goals, and by his organizational skills. In the late 1990s, he set up a consulting firm called *Bravo*, which had to be sold in early 2000 (to an Italian investor) following the "dot-com" crisis. This start-up continues to enjoy success to this day. It is represented in big financial places like London and New York with Zari as Chief Scientist. There are not many scientists who have his energy and ability to combine strong mathematical skill and science with real practical applications.

At a dinner with friends, people ask your wife what you do at work, besides teaching. What does *she* answer?

While I was studying mathematics, other students from pure math were often asked the same thing and were absolutely unable, even at the end of their studies, to explain why they were studying mathematics except to say that they liked abstract structures and so on. In my field, this question is (and was) much easier to answer, due to the wealth of applications. On a personal level, mathematics helps to develop many important qualities such as perseverance, frustration tolerance, the ability to find new angles and perspectives, clearness in mind and many more. Mathematics should also be important for developing a country and society, and it should help to solve the great and important problems in the world. My wife knows me well and probably would also answer along these lines. Then she would smile and say "but beyond that, there always remains something unknown like a secret".

Every good mathematician has a side activity which runs as a hobby. Is it true that you were a professional table tennis player?

Except in the last few years, I have always done a lot of sports, including soccer, volleyball, tennis and table tennis, karate, swimming, and surfing. In my student days, I played in fact a lot of table tennis, often the whole night, accompanied by some beer. I also played table tennis at many conferences with experienced colleagues. This being said, I remember that in our dorm, there was a professional table tennis player from the first *Bundesliga* team of Osnabrück who made me aware of my limits in a somewhat frustrating way. My wife and I enjoy music a lot, classic and modern; we often go to concerts or to operas and take part in the rich cultural life in and around our city like exhibitions, public lectures and theater. My wife is an active painter.
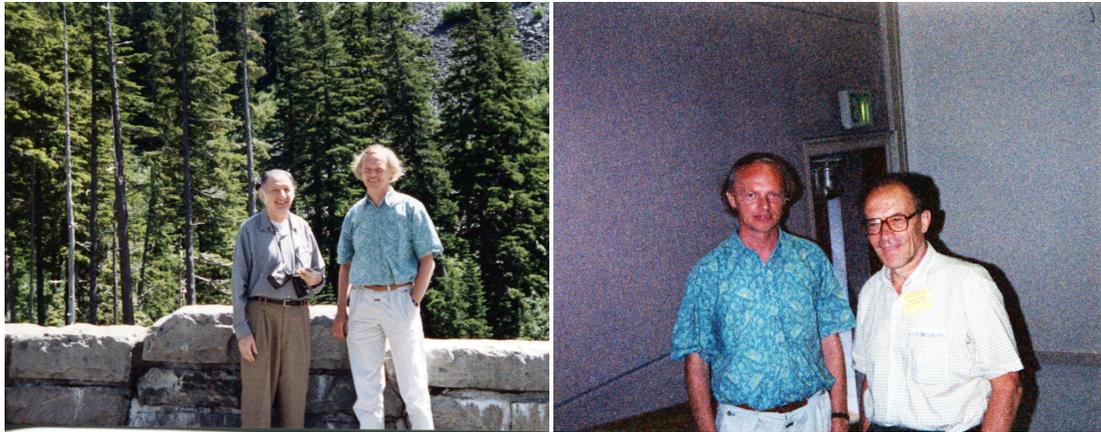
Figure 2: Ludger Rüschendorf with Abe Sklar (left picture) and Giorgio Dall'Aglio (right) in Rome (Italy), 1990, during the conference *Advances in Probability Distributions with Given Marginals* (see the proceedings in [5]).

For 30 years I have been singing in choirs, typically classical programs. These kinds of activities allow me to relax and refresh my mind.

2 Dependence Modeling

When (and how) did you first become aware of the term "copula"?

I became aware of the notion of a copula in the early or mid 1980s but I must confess that its importance as an essential organizing principle for dependence modeling, as described in Sklar's Theorem [42] (i.e., to separate the effects of the marginals from the effects of dependence) only became clear to me much later. However, I was already well aware of the notion of Fréchet class as the class of all dependence models with the given marginals in the late 1970s. As consequence of ample experience with the construction of randomized tests the standardization to uniform marginals was not new to me even before hearing or reading of Sklar's Theorem the first time. For example in my paper [33] on stochastic ordering in Fréchet classes, published in 1981 in *Mathematische Operationsforschung und Statistik* (later called *Statistics*), I used this standardization to prove a stochastic ordering result between two different Fréchet classes. I also gave in that paper a simple proof of Sklar's Theorem without being aware of its existence. Much later, in 2013, reading the review article [9] of Durante and Sempi, I became aware of the fact that Moore and Spruill had already used the same idea of proof in a statistics paper; see [27]. In fact this reference appeared already in the survey paper of Schweizer in the proceedings (see [5]) of the conference held in 1990 in Rome, but it escaped my attention.

What do you think of the way in which the field of dependence modeling evolved over these past 50 years?

Beyond the classical multivariate modeling approaches and multivariate standard models such as Gaussian models or extreme-value models or in dynamic form by the corresponding time series models, this field has been greatly impacted in its theoretical development by the series of conferences on *Probabilities With Given Marginals* that started in Rome in 1990 and was later on combined with conferences on general multivariate modelling as in the Tartu 2007 conference. However, the big breakthrough for dependence modeling and copula models came only at the end of the 1990s with the need felt in finance and insurance to cope with the obvious problems of the classical models, which ignored the influence of dependence on risks. From one moment to the next, all the early work on copula (dependence) models was picked up by practitioners, along with the corresponding statistical tools (documented, e.g., in conference proceedings). Nowadays, we have access to a series of good books and comprehensive presentations on copula model constructions and simu-

lation methods (as, e.g., pair copula models) and on various applications in quantitative risk management, in finance and insurance, but also in reliability engineering, hydrology, medicine and in many other fields.

You have participated in one of the first conferences devoted to distributions with fixed marginals, held in Rome in 1990. Why did you go there? What was the "climate" like?

The conference in Rome in 1990 (see Figure 2) was the first in a series of conferences held in Seattle, Prague, Barcelona, Québec, Tartu and São Paulo devoted to distributions with given margins. The initial Rome conference gave an essential impetus to the general field of dependence modeling. During that conference there was a feeling that this conference was a historical event. It was touching to experience the great and humble personality of Abe Sklar, but also to learn from Giorgio Dall'Aglio about the relevant and important contributions of Italian probabilists to the subject of Fréchet classes. Berthold Schweizer gave an impressive historical review on thirty years of copulas. The conference was inspiring and the atmosphere was very nice. It offered a variety of new perspectives on the field as well as the chance to meet with a lot of eminent scientists and personalities. I was invited to give a talk, possibly because I had given a presentation on probabilities with higher dimensional marginals in a conference on stochastic orderings in 1989 in Hamburg, organized by Karl Mosler and Marco Scarsini. A very interesting and valuable IMS proceedings volume of the latter conference had been published and many researchers who participated in it also went to Rome. The subsequent conference in Seattle, held in 1993, was equally inspiring and drew many more participants (see Figure 3). Together with Berthold Schweizer and Mike Taylor, I edited the proceedings of this conference; see [40].

What change have you observed from the first conferences devoted to copulas and dependence to more current events?

I think that the more recent workshops and conferences on the subject of dependence modeling include to a much larger degree real and interesting applications; they are also more concerned with statistical issues and related simulation techniques. It is a great achievement that the consciousness of new tools and methods has now reached areas that did not seem open to change and new developments earlier on. It is great when scientists like Paul Embrechts are now asked for advice and consulting at important institutions and are invited to address large audiences from the world of finance and insurance.

Various disciplines contribute to the field of dependence modeling, e.g., probability, statistics, analysis, and computer science. Where do you position yourself and how can these disciplines profit from one another in the best possible way?

As I mentioned earlier, it is clear that in most quantitative scientific disciplines, describing and modeling the effect of dependence is a key issue. Consequently, the experiences and achievements in the various fields are of relevance and interest for the general development of the methods, as well as for applications in other fields. We should keep an open mind with respect to the different and varying conditions and circumstances that prevail in these fields. Just to give one or two examples drawn from my own experience, I am — and was considered — a relatively theoretically orientated mathematician, doing research mainly motivated by its inherent mathematical logic or motivation. It also seems that a few of these undertakings were successful and led to reasonable and interesting results.

In my recent years I have become more and more interested in real applications of research. I like relevant, practical results even when their mathematical contents is not so sophisticated. For example with my former PhD student Georg Mainik, I recently developed a portfolio optimization method based on a new extremal risk index that we derived from extreme-value theory; see [26]. This index measures in a scientific way the influence of dependence on the portfolio. Together with Georgi Mitov we implemented this approach in a practical system and obtained encouraging results in a large-scale empirical study on S&P data comparing our solution to the classical Markowitz method. I was quite happy when we managed to publish these results in the *Journal of Empirical Finance* in 2015; see [25]. A more elaborate practical implementation remains to be done.

In recent years I have been cooperating a lot with several colleagues from mathematics and Economics and also from practice on finding improved bounds for the aggregated risk of portfolios in banks and in insurance companies. This is a highly relevant topic with great consequences. Our main goal is to describe, besides the usually available information on the individual risks (marginals), the basic available dependence

Figure 3: Ludger Rüschendorf at the conference *Distributions with Fixed Marginals, Doubly Stochastic Measures, and Markov Operators* held in Seattle (see the proceedings in [40]), 1993, with (left to right) Marco Scarsini, Svetloszar Rachev, Ingram Olkin, Albert Marshall and Moshe Shaked.

information and its consequences for the evaluation of risk. It is a great pleasure to work on this interesting type of relevant problems with excellent researchers like Giovanni Puccetti, Paul Embrechts, Steven Vanduffel, Carole Bernard, Ruodu Wang and others; we expect some relevant progress by including practical experience and information.

What do you consider the most important challenges in dependence modeling?

As there are so many different areas of application, one can offer many different answers to this question. In the previously mentioned risk bounds for aggregated risks, for example, it is of high interest to determine the kind of dependence information available in companies or from experts. This may require the development of statistical tools to determine this information. One may also need to model specific causes or dynamic structure. The determination of best bounds based on this information poses new challenges, too. I think in applications from different fields, one faces similar issues, which points to the need for cooperation between disciplines.

Copulas have become classical tools in finance, insurance, geostatistics and hydrology. Do you see other fields where they might soon emerge? Big Data analysis?

As I said before, dependence modeling is a relevant part of modeling in essentially any field of research. The analysis of Big Data seems to be a hot topic currently and so it is worthwhile to be concerned more with its structure, components and temporal development. Probably large networks will play a major role in the study of its dynamic behavior. Also the statistical part will be relevant, but my knowledge of the field is too restricted to be more precise at this time.

You and Makarov have proved an outstanding result on the sum of random variables that now has relevant applied consequences. How did you arrive at the problem in 1982, and how do you think of the problem nowadays?

In my habilitation thesis, back in 1979, I derived a dual representation for a class of extremal problems on Fréchet classes. As an application, I showed in a 1981 paper [32] the sharpness of the Fréchet upper and lower bounds in a more general framework for maximizing (resp. minimizing) the probability of product sets. This was, in particular, the first proof that the lower Fréchet bound is sharp. As a second application, this result also implies sharp bounds for the distribution function of the sum in dimension 2; in fact, I derived this result from Strassen's classical result. My paper [34] was published in 1982; Makarov published the same result in 1981, answering a question put to him by Kolmogorov. The two papers were completely independent from each other, and they used different techniques. In recent years the great relevance of this kind of results for estimating the value-at-risk of aggregated portfolios has come to light, in particular through papers of

Embrechts and Puccetti, see [10], who, based on the duality approach, established some interesting results. This led to a very active research area which, at this time, is still ongoing; there are many open and interesting problems to be solved.

You contributed a lot to the field of mass transportation problems / optimal couplings by writing your famous two-volume book with S.T. Rachev. This theory was then further developed by many mathematicians. In particular Cédric Villani, who won the Fields Medal, showed interesting connections between optimal transport and curvature. How do you feel now about the investigations in this area?

As mentioned before, I found in my habilitation thesis a dual representation for extremal problems on the class of distributions with given marginals. My main motivation at that time was to describe bounds for the possible influence of dependence on integral functionals. In the case of two marginals where the functional is a distance (or, more generally, a cost function), this problem reduces to the classical formulation of the mass transportation problem as introduced by Kantorovich in 1942. I was not aware of this work and got to know about it only in the mid 1980s from S.T. Rachev, with whom I started to cooperate at that time on probability metrics and their application to establish Central Limit Theorems with convergence rates.

Zari Rachev had written his Dr. Science PhD thesis with Zolotarev on probability metrics and Kantorovich was on the thesis committee. Our cooperation was very fruitful and lasted over many years. In 1990 we found, on an indication by Uwe Rösler on his convergence proof of the Quicksort algorithm, that probability metrics and the techniques to prove Central Limit Theorems are also a suitable tool to prove convergence results for recursive algorithms. This led to the birth of the "contraction method", which to this day is perhaps to most general method for this area at the frontier between mathematics and computer science; it originated in the work of Uwe Rösler and my own with Zari Rachev. We also started nearly at the same time to work together on problems of mathematical finance, specifically on establishing valuation formulas. This field was just beginning to develop and I was motivated by papers of Harrison and Kreps, and by Föllmer and Sondermann. This later became Rachev's main subject area. Our collaboration culminated in our book [30] on mass transportation in 1998, which includes applications from many different areas.

While mass transportation is related to dependence modeling, many researchers in copula theory are not aware of it. Would you briefly introduce this field and sketch possible applications?

The basic problem of mass transportation is to find an optimal transport (a mapping), i.e., one with minimal transport cost from one mass distribution to another. In the classical case it was formulated in $\mathbb{R}^2$ and $\mathbb{R}^3$ for uniform distributions. In a joint paper in 1990 (see [39]), Rachev and I found a complete characterization of optimal transports for general distributions in $\mathbb{R}^k$ w.r.t. L^2 costs. The same result was found in 1991 independently for a slightly more special class of distributions by Brenier (see [3]), who recognized the importance of this result for the solution of the Monge–Ampère and other nonlinear partial differential equations. This led to a strong development of these connections in analysis and geometry up to Villani's achievements. In analysis, only Brenier's paper tends to be referenced. In a subsequent paper, published in 1991, I extended the characterization of optimal transports to general cost functions [35]. The solution is based on the notion of c-subgradient which was introduced in that paper.

The optimal transport of mass distributions is essentially equivalent to the construction of optimal couplings, i.e., the construction of two random variables in $\mathbb{R}^k$ that have the given distributions and are positively dependent in the strongest possible way. In the case of real random variables, one gets a simple general solution: the comonotonic vector gives the strongest possible positive dependence. But in higher dimensions, the problem is not so simple and depends strongly on the cost function; it has connections with Voronoi diagrams and several interesting algorithms (discrete and continuous) have been developed to solve these kinds of problems.

You started investigating the convergence of the empirical copula process in 1976. What have been the subsequent achievements that you like most? Are there any further steps to be taken?

I introduced the empirical copula process in my thesis; the result was published two years later, in 1976, in *The Annals of Statistics*; see [31]. I was unaware of the notion of a copula at that time and used the term "reduced empirical process" for the copula process. It was a very natural idea to replace the standardization transformation of the components by their distribution functions by a transformation of the components by

Figure 4: Ludger Rüschendorf with four members of the board of *Dependence Modeling* in Brussels, 2014: (from left to right) Fabrizio Durante, Carole Bernard, Ruodu Wang and Giovanni Puccetti.

their empirical distribution functions. I called the resulting "empirical copula process" multivariate rank-order process and gave a functional convergence result for this process.

The applications I considered had to do with the asymptotic distribution of statistics for tests of independence such as the Kolmogorov or Cramér–von Mises statistics; I also looked at the asymptotics of dependence measures like Spearman's ρ and Kendall's τ. The empirical copula process was reintroduced and further studied by other authors [6, 12, 44] and used for the construction and analysis of statistical tests. Efficiency considerations were given in [2] and many other developments ensued through the work of Christian Genest and his collaborators; see, e.g., [13]. In several papers in the mid 1980s, I considered the question of how to improve estimators and test procedures when additional information on the marginals is available. In particular I determined the tangent space of marginal models which is necessary to the determination of asymptotic efficient estimators and I used this information to improve estimators. As mentioned earlier, the statistical analysis of marginal models based on realistic dependence information is a relevant and worthwhile topic of further study.

In recent years, your research profile somewhat changed from very theoretical research to more applied topics. Would you explain this change?

As I already explained before, I have always been interested in applied research topics as well. For example in my main area of statistical research, I was very much interested for some years in nonparametric estimation techniques like neural nets, radial basis functions and others with application to image reconstruction, pattern classification and survival analysis. I liked the interesting mathematical problems in these areas, e.g., the development of new functional approximation results as well as nonparametric consistency rates arising from connections with Vapnik–Cervonenkis theory, as in my joint work [8] with my former PhD student Sebastian Döhler. As mentioned earlier, an important motivation for switching to more applied topics was my growing interest in real applications.

Your theoretical contributions on "rearrangements" have — many years later — turned out to be of great practical value. Any more hidden theoretical pearls (of yours) we should look out for?

Rearrangements are in fact a nice way to look at statistical dependencies. They are very natural and easily understandable and have been developed in the classical text of Hardy, Littlewood and Pólya [18] but also later in interesting functional analytic literature. This connection led to a better understanding of extremal problems in Fréchet classes, of the structure of solutions and of the connection of these problems with (convex) ordering properties. A nice and related point in the application to dependence modeling was the invention of the Rearrangement Algorithm (RA) in a paper with Giovanni Puccetti, see [29], extended further

in [11]. This led to a really practical tool in the context of extremal problems in Fréchet classes. Several relevant extensions of this algorithm have been given in the meantime.

What do you consider your biggest contribution to the field of dependence modeling?

If I had to mention only one paper (or line of research), it would be the dual representation of the optimization problems in Fréchet classes from my habilitation paper. This work generated a lot of work and motivated, e.g., the extension of the duality result by Kellerer [22] and joint work with my friend Doraiswamy (Chandra) Ramachandran (1995–99) where we found that the duality result in a strong sense only holds true in perfect measure spaces (so it typically will not be true in many function spaces) and also discussed its relevance for assignment problems in relevant economic models. But the duality result also led, due to the connection with the aggregation of risk problem in [10], to the treatment of a very interesting class of risk aggregation problems in recent years.

Doing mathematical research is truly hard work and turning an idea into a published paper can take years. You have written a lot of scientific papers. What motivates you each time you start a new project?

In fact I had the pleasure to celebrate my 200th scientific paper recently. I know some colleagues who have published more and, of course, not all of my work is of the same quality. At the root of all this work is curiosity in many different areas; this led me to switch focus several times to different, but in some sense, related subjects. I started with asymptotic statistics, which I continued to work on throughout my scientific way. Then I investigated stochastic orderings, mass transportation, and Fréchet classes. I then dealt with probability metric applications to Central Limit Theorems, to the analysis of algorithms (from the mid 1990s until now) and to problems in applied probability like general forms of the Poincaré Theorem, propagation of chaos or in probabilistic algorithms and random graphs and networks. A result of particular charm, and also interest for dependence modeling, is the proof given in [36] that the iterative proportional fitting procedure converges. This algorithm goes back to classical work of Deming and Stephan in 1940 (see [7]), but a proof of its convergence was known only in the finite, discrete case. The algorithm enables one to approximate the projection of a probability measure on a Fréchet class and can thus be used, among other things, as the basis for goodness-of-fit tests.

I have also been working in several other areas of applied probability such as optimal stopping theory, where I, together with my PhD students Robert Kühne and Sebastian Faller, worked out a general method to determine approximate optimal stopping and selection rules. An area that I followed since the mid 1990s is financial mathematics and the related topic of stochastic analysis. In several papers with Jan Bergenthum, I used methods from stochastic analysis to derive comparison results for option prices in semi-martingale models. A particularly nice result is a characterization of optimal portfolio strategies in a paper with Thomas Goll (see [14]) that was based on the characterization of projections of probability measures with respect to ϕ-divergence distances which I found in 1984. All in all by lucky circumstances I could explore so many different areas and was not confined to a single topic.

Being an adviser of many PhD students, what do you teach them, and vice versa?

I had many very good students and I learned a lot from them and from cooperating with them. I always found it of primary importance that they follow their own way, that they have broad interests, and that they be independent enough from their thesis adviser. So far this worked out quite well. I experienced painfully a few times that extremely talented PhD students with very good perspectives for a university career chose to leave academia and settled for an industry position for security or private reasons.

Do you have an unfinished research project that keeps you busy at night?

I still have several projects to be finished. In the last few years I also enjoyed writing longer texts on broader areas. In 2013 I published my Springer text on Mathematical Risk Analysis, see [37], one of my main research areas in the last years. In 2014 I wrote a textbook on statistics (advanced level) in German, see [38], and at the moment I am writing a German textbook on probability with an advanced and applied touch. I have not yet decided what to do afterwards.

3 Development of the scientific community

You started doing research in the "pre-Google days". How did the advent of microcomputers, the internet and statistical software change your personal life as a researcher and how did it change our field of research in general?

I still remember that when I was a postdoctoral fellow, I spent about 30–40% of my time looking through the library and trying to get some relevant information. This task is so much easier today; it is unimaginable what difference this makes for scientific research. It may be, though, that the abundance of information one can get this way sometimes prevents one from starting a project under the impression that so much is known and that so many people are working on it.

What aspects of this change do you appreciate and which ones would you sometimes like to "undo"?

I don't think I would like to do away with any of these developments. One down-side is that it is now difficult for young people to turn off their iPhone and to concentrate and work in a calm environment.

Do you personally typeset your papers?

No. I usually rely on the competent help of my secretary. The initial draft of this interview was typed by my youngest son.

Students often ask about how it was like to write academic papers (or to be an editor) in the pre-LaTeX days. Any anecdotes?

Typing papers and in particular doing corrections was very difficult in the pre-TeX times and in particular in the pre-computer times. Systems eventually emerged that could produce simple formulas but for the more complicated formulas one had to use handwritten symbols. In those days, the proportion of handwritten symbols in a paper was a good indication of the level of technical advance in the author's country and institution.

What do you think the current scientific editorial system needs most?

I know that some editors took the initiative to improve the editorial system. In some journals the review times are much too long and very disappointing for a researcher who submits a paper and is waiting for the result for so long. This problem seems to be particularly acute in economics and finance. It seems that some journals have recently introduced measures to fix this problem. Also a younger and in some sense less saturated Editorial Board may help to improve things.

The style of mathematical talks has changed dramatically from classical blackboard "Oberwolfach style" to beamer slide presentations. What are the pros and cons?

In my view, some talks focusing on new concepts can still convey the message best if given in the Oberwolfach style. For many talks with a lot of material, the beamer style is better. I never liked lectures that did not challenge the audience to a sufficient degree.

You wrote papers in a variety of different fields. Which paper of yours do you consider as the farthest away from probability and statistics?

Probably a paper with Thomas Bruss on the "perception of time", in which we found a logarithmic law for the thinning out of the feeling of time with increasing age; see [4]. A nice joint paper with Olaf Krafft and Wilhelm Plesken from 1980 addressed the problem of whether it is possible to move a pointed cone, generated by n vectors in $\mathbb{R}^n$ with pairwise angles smaller than 90 degrees by an (orthogonal) movement into the positive orthant. Given that this is quite obvious for dimensions $n \leq 3$, we wanted to prove it in generality. This result was needed for my strategy to prove that a multivariate normal vector is associated (i.e., strongly positive dependent) if and only if its covariance matrix is non-negative. After a lot of work, however, we found counterexamples in dimension $n = 5$; the result holds in dimension $n = 4$ though. Our paper was never published. When we submitted it to *Linear Algebra and its Applications*, the Editor Olga Taussky-Todd wrote back almost immediately that the journal had just accepted a paper proving the same result; see [15]. The authors of that paper had a completely different motivation. As for the dependence result for normal vectors, it was proved a little later by Pitt using analytic tools; see [28].

A very nice paper in a remote area with Wolfgang Thomsen, see [41], concerned the extension of the famous result of Kolmogorov [23] with which he (in part with Arnold) disproved conjecture No. 13 of Hilbert's famous list of 23 mathematical problems. Kolmogorov showed that any continuous function of n variables has a (minimal) representation in terms of a superposition of two layers of continuous functions each of which depends on a single variable. As a result, one needs only continuous functions of one variable and addition to describe any continuous function of n variables. This result was considered as a theoretical justification for neural networks (with a single hidden layer). Based on a somewhat intricate closeness result, we found an extension of this representation result to general locally bounded measurable functions, such as the indicator functions which typically appear in classification applications of neural networks.

4 Final questions

You constantly carry with you a bag load of papers. Are there some in the lot that you *always* carry around?

No, the contents of my bag varies, depending on my concerns at the time. Some of them are used more often than others.

You are close to retiring. What are your plans for the future?

I plan to keep on working on some interesting projects with my colleagues and students. I hope to remain in good health and active, and maybe to get involved in some unusual social engagement and to spend more time with my family and friends.

Had you not been a mathematician, what would you have done in life?

I always enjoyed my work as a mathematician and I could happily work anywhere. Several of my papers have been written on holidays or by a lake. Some parts of my most recent statistics textbook have been written at the beach in Hammamet, where I was staying in a very nice place. My wife likes to solve sudokus for relaxation. I had my way of relaxation and the opportunity to solve problems with my mathematics. Of course there was not only sunshine in my scientific life, but we as humans are fortunate enough to forget this over time.

Give three pieces of advice to a young mathematician starting his/her carrier.

Recommendations for a young mathematician before or after PhD would depend heavily, I think, on the individual and his/her preferences. Some general good and leading rules from my point of view might be:

- Choose a research topic that stimulates your interest to a sufficiently high degree. If it becomes too much of a routine, skip it and switch if possible.
- Keep independent enough from your advisor or research group to be able to take your own free decisions.
- Don't always choose the easiest route: don't hesitate to take risks!
- Sometimes it is better to let the problems rest and to relax

References

[1] Andersen, P. K., O. Borgan, R. D. Gill, and N. Keiding (2012). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.

[2] Bickel, P. J., Y. Ritov, and J. A. Wellner (1991). Efficient estimation of linear functionals of a probability measure P with known marginal distributions. *Ann. Statist. 19*(3), 1316–1346.

[3] Brenier, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* 44(4), 375–417.

[4] Bruss, F. T. and L. Rüschendorf (2010). On the perception of time. *Gerontology* 56(4), 361–370.

[5] Dall'Aglio, G., S. Kotz, and G. Salinetti (Eds.) (1991). *Advances in Probability Distributions with Given Marginals*. Kluwer Academic Publishers Group, Dordrecht.

[6] Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. *Acad. Roy. Belg. Bull. Cl. Sci.(5) 65*(6), 274–292.

[7] Deming, W. E. and F. F. Stephan (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Stat. 11*(4), 427–444.

[8] Döhler, S. and L. Rüschendorf (2003). Nonparametric estimation of regression functions in point process models. *Stat. Inference Stoch. Process. 6*(3), 291–307.

[9] Durante, F. and C. Sempi (2010). Copula theory: an introduction. In *Copula Theory and Its Applications*, Volume 198 of *Lecture Notes in Statistics*, pp. 3–31. Springer, Berlin.

[10] Embrechts, P. and G. Puccetti (2006). Bounds for functions of dependent risks. *Finance Stoch. 10*(3), 341–352.

[11] Embrechts, P., G. Puccetti, and L. Rüschendorf (2013). Model uncertainty and VaR aggregation. *J. Bank. Financ. 37*(8), 2750–2764.

[12] Fermanian, J.-D., D. Radulović, M. Wegkamp (2004). Weak convergence of empirical copula processes. *Bernoulli 10*(5), 847–860.

[13] Genest, C., J.-F. Quessy, B. Rémillard (2007). Asymptotic local efficiency of Cramér-von Mises tests for multivariate independence. *Ann. Statist. 35*(1), 166–191.

[14] Goll, T. and L. Rüschendorf (2001). Minimax and minimal distance martingale measures and their relationship to portfolio optimization. *Finance Stoch. 5*(4), 557–581.

[15] Gray, L. and D. Wilson (1980). Nonnegative factorization of positive semidefinite nonnegative matrices. *Linear Algebra Appl. 31*, 119 – 127.

[16] Grenander, U. (1968). *Probabilities on Algebraic Structures* (2nd edition). Almqvist & Wiksell, Stockholm and John Wiley, New York.

[17] Hall, P. (1935). On representatives of subsets. *J. London Math. Soc. s1-10*(1), 26–30.

[18] Hardy, G. H., J. E. Littlewood, and G. Pólya (1952). *Inequalities* (2nd edition). Cambridge University Press, Camdridge.

[19] Holtrode, R. and L. Rüschendorf (1993). Differentiablity of point process models and asymptotic efficiency of differentiable functionals. *Statistics 24*(1), 17–42.

[20] Iosifescu, M. and P. Tăutu (1973). *Stochastic Processes and Applications in Biology and Medicine.* Springer-Verlag, Berlin-New York.

[21] Karlin, S. and J. McGregor (1964). Direct product branching processes and related Markov chains. *Proc. Nat. Acad. Sci. U.S.A. 51*, 598–602.

[22] Kellerer, H. G. (1984). Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete 67*(4), 399–432.

[23] Kolmogorov, A. N. (1957). On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Dokl. Akad. Nauk SSSR 114*, 953–956.

[24] Linnik, Y. V. (1975). *Problems of Analytical Statistics*. Statistical Publishing Society, Calcutta.

[25] Mainik, G., G. Mitov, and L. Rüschendorf (2015). Portfolio optimization for heavy-tailed assets: Extreme risk index vs. Markowitz. *J. Empirical Finance 32*, 115–134.

[26] Mainik, G. and L. Rüschendorf (2010). On optimal portfolio diversification with respect to extreme risks. *Finance Stoch. 14*(4), 593–623.

[27] Moore, D. S. and M. C. Spruill (1975). Unified large-sample theory of general chi-squared statistics for tests of fit. *Ann. Statist. 3*, 599–616.

[28] Pitt, L. D. (1982). Positively correlated normal variables are associated. *Ann. Probab. 10*, 496–499.

[29] Puccetti, G. and L. Rüschendorf (2012). Computation of sharp bounds on the distribution of a function of dependent risks. *J. Comput. Appl. Math. 236*(7), 1833–1840.

[30] Rachev, S. T. and L. Rüschendorf (1998). *Mass Transportation Problems. Vol. I–II*. Springer, New York.

[31] Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Ann. Statist. 4*, 912–923.

[32] Rüschendorf, L. (1981a). Sharpness of Fréchet bounds. *Z. Wahrsch. Verw. Gebiete 57*(2), 293–302.

[33] Rüschendorf, L. (1981b). Stochastically ordered distributions and monotonicity of the OC-function of sequential probability ratio tests. *Math. Operationsforsch. Statist. Ser. Statist. 12*(3), 327–338.

[34] Rüschendorf, L. (1982). Random variables with maximum sums. *Adv. Appl. Probab. 14*, 623–632.

[35] Rüschendorf, L. (1991). Fréchet-bounds and their applications. In *Advances in Probability Distributions with Given Marginals*, Volume 67, pp. 151–187. Dordrecht: Kluwer Acad. Publ.

[36] Rüschendorf, L. (1995). Convergence of the iterative proportional fitting procedure. *Ann. Statist. 23*, 1160–1174.

[37] Rüschendorf, L. (2013). *Mathematical Risk Analysis. Dependence, Risk Bounds, Optimal Allocations and Portfolios*. Springer, Heidelberg.

[38] Rüschendorf, L. (2014). *Mathematische Statistik*. Springer, Berlin.

[39] Rüschendorf, L. and S. T. Rachev (1990). A characterization of random variables with minimum L^2-distance. *J. Multivariate Anal. 32*(1), 48–54.

[40] Rüschendorf, L., B. Schweizer, and M. Taylor (Eds.) (1996). *Distributions with Fixed Marginals and Related Topics*, Hayward, CA. Inst. Math. Statist.

[41] Rüschendorf, L. and W. Thomsen (1998). Closedness of sum spaces and the generalized 'Schrödinger problem'. *Theory Probab. Appl. 42*(3), 483–494.

[42] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris 8*, 229–231.

[43] Strasser, H. (1985). *Mathematical Theory of Statistics: Statistical Experiments and Asymptotic Decision Theory*. Walter de Gruyter & Co., Berlin.

[44] Stute, W. (1984). The oscillation behavior of empirical processes: the multivariate case. *Ann. Probab. 12*, 361–379.