Institut für Photogrammetrie und Kartographie
der Technischen Universität München
Lehrstuhl für Methodik der Fernerkundung

Very High Resolution Tomographic SAR Inversion for Urban Infrastructure Monitoring —
A Sparse and Nonlinear Tour

Xiaoxiang Zhu, M.Sc.

Abstract

Synthetic aperture radar (SAR) is the only way to assess deformation of the Earth's surface from space on the order of centimeters and millimeters due to its coherent nature and short wavelengths (typically 3-25 cm). In particular, with the launches of new SAR sensors, such as the German TerraSAR-X/TanDEM-X and the Italian COSMO-Skymed satellites, SAR remote sensing from space has made a big leap forward. These satellites deliver SAR data with a very high spatial resolution of up to 1 m, and hence open up for the first time opportunities to use SAR for 2-D, 3-D, 4-D (space-time) or even higher dimensional imaging of urban structures and individual buildings from space. That means the 3-D shape and deformation or subsidence of the individual buildings can be retrieved.

A single SAR image can only provide cartographic information in the two native coordinates "azimuth" and "range". In order to retrieve the 3-D position, i.e. including the "elevation" coordinate, as well as motion information of the scattering objects, advanced interferometric SAR techniques are required that exploit stacks of complex-valued SAR images with diversity in space and time. Among them, *tomographic SAR inversion*, including SAR tomography and differential SAR tomography, provides the most advanced means for 4-D SAR imaging to date. It is a relatively new technique and is not yet exploited with very high resolution SAR data over urban areas. The intention of this thesis is to further develop this technique, and hence, to explore the potential of very high resolution SAR data for urban infrastructure mapping. The work presented in this thesis contributes to the field by addressing the following four new aspects:

Very high resolution tomographic SAR inversion is demonstrated using TerraSAR-X spotlight data to provide 3-D and 4-D maps of an entire high rise city area including layover separation. For individual buildings, a high proportion of double scatterers — up to 20% — is detected by using a modified version of the conventional singular value decomposition inversion method followed by model order selection.

Due to the tight orbital tube of modern SAR sensors the elevation aperture is small, i.e. the inherent resolution in elevation is about 50 times worse than in azimuth or range. This extreme anisotropy calls for super-resolution algorithms in the elevation direction while maintaining the meter azimuth-range resolution. On the other hand, the high anisotropy of the 3-D tomographic resolution element renders the signals sparse in the elevation direction; only a few point-like reflections are expected per azimuth-range cell. *A compressive sensing based algorithm* tailored to very high resolution SAR data is developed for tomographic SAR inversion by exploiting the sparsity of the signal in elevation. It is named "$\underline{S}$cale-down by $\underline{L}_1$ norm $\underline{M}$inimization, $\underline{M}$odel selection, and $\underline{E}$stimation $\underline{R}$econstruction" (SL1MMER, pronounced "slimmer"). SL1MMER combines the advantage of compressive sensing sparse reconstruction (e.g. super-resolution properties and high point localization accuracy) and amplitude and phase estimation accuracy of linear estimation, and hence gives reliable estimation of the number of scatterers, elevation, motion parameters, amplitude and phase of each scatterer. Furthermore, a practical demonstration of the super-resolution of SL1MMER for SAR tomographic reconstruction is provided with a tremendously increased proportion of detected double scatterers of up to 38%.

A systematic performance assessment of the proposed SL1MMER algorithm is performed regarding the elevation estimation accuracy, super-resolution power and robustness. Compared to the Cramér-Rao lower bound, both numeric results and an analytic approximation of the elevation estimation accuracy are provided. It is shown that SL1MMER is an *efficient* estimator. The super-resolution factors are found by extensive simulations. These establish fundamental bounds for super-resolution of spectral estimators. The achievable super-resolution factors of SL1MMER in the typical parameter range of tomographic SAR are found to be promising and

are on the order 1.5~25. The minimal number of acquisitions required for a robust estimation is derived and given by explicit formulas.

Conventional tomographic inversion allows only for the retrieval of linear motion, although motion or deformation of buildings is often nonlinear (periodic, accelerating, stepwise, etc.). *A generalized time warp method* is developed which enables tomographic SAR to estimate multi-component nonlinear motion by a nonlinear warping of the time axis.

All developed methods are validated with both simulated and extensive processing of large volumes of real data from TerraSAR-X.

I hope the work presented in this thesis constitutes a substantial contribution to the vision of *"a dynamic city model showing the shape and the deformation of each building"*.

Zusammenfassung

Synthetik Apertur Radar (SAR) ist ein kohärentes Abbildungsverfahren und arbeitet mit Wellenlängen von typisch 3 – 25 cm. Daher ist es das einzige Weltraum gestützte Verfahren, das es erlaubt, Deformationen der Erdoberfläche mit Zentimeter- und sogar Millimeter-Genauigkeit zu erfassen. Mit dem Start neuer SAR-Systeme, wie den deutschen TerraSAR-X und TanDEM-X, sowie den italienischen COSMO-Skymed-Satelliten, hat die SAR-Fernerkundung einen wichtigen Schritt in die Zukunft getan. Diese Satelliten liefern SAR-Daten mit sehr hoher räumlicher Auflösung (VHR für englisch „very high resolution") von bis zu 1m. Sie eröffnen deshalb erstmals die Möglichkeit, SAR zur 2-D, 3-D, 4-D (raum-zeitlichen) oder sogar höherdimensionalen Abbildung urbaner Infrastruktur und einzelner Gebäude vom Weltraum aus einzusetzen, d.h. die 3-D Form von Gebäuden, sowie deren Deformationen oder Absenkungen abzuleiten.

Eine einzelne SAR-Aufnahme liefert nur kartographische Information in den beiden nativen Koordinaten „Azimut" und „Range". Um die tatsächliche 3-D Lokalisation (also einschließlich der dritten Koordinate „Elevation") und die Bewegungsparameter der streuenden Objekte zu schätzen, werden moderne interferometrische SAR-Verfahren benötigt. Diese nutzen Stapel komplexwertiger SAR-Daten, aufgenommen von unterschiedlichen Orbitpositionen und zu unterschiedlichen Zeiten. *Tomographische SAR-Inversion*, also SAR-Tomographie und differentielle SAR-Tomographie, stellt das derzeit fortschrittlichste 4-D SAR-Abbildungskonzept dar. Diese relativ neue Methode wird allerdings noch kaum mit den nun verfügbaren VHR-Daten zur Abbildung urbaner Gebiete genutzt. Diese Dissertation hat zum Ziel, die tomographische Technik weiter zu entwickeln, um das Potential der VHR SAR-Systeme zur raum-zeitlichen Abbildung urbaner Infrastruktur auszuschöpfen. Die Arbeit trägt dazu vier neue Aspekte und Lösungen bei:

VHR tomographische SAR-Inversion wird mit TerraSAR-X Spotlight-Daten demonstriert. Es wird gezeigt, dass eine 3-D und 4-D Karte eines gesamten Innenstadtgebiets mit Hochhäusern abgeleitet werden kann, einschließlich der Auflösung von Layover-Bereichen, d.h. der Trennung von mehrfachen Streuern in einem Bildpunkt. Für einzelne Gebäude werden mit einer modifizierten Version des bekannten Einzelwert-Zerlegungs-Algorithmus und nachgeschalteter Modelselektion hohe Zweifachstreuer-Dichten – bis zu 20% – detektiert.

Wegen des sehr engen Orbitschlauchs moderner SAR-Satelliten ist die Elevationsapertur klein und damit die Auflösung in Elevationsrichtung ca. 50 mal schlechter als in Azimut und Range. Diese starke Anisotropie des 3-D Auflösungselements verlangt nach Überauflösungsverfahren in Elevation, ohne die Azimut- und Range-Auflösung zu verschlechtern. Andererseits garantiert diese anisotrope Punktantwort auch, dass das Signal als „sparse" in Elevation angesehen werden kann. Somit werden nur wenige diskrete Streubeiträge pro Azimut-Range-Pixel erwartet. Dies legt die Nutzung von Compressive Sensing Verfahren nahe. Zur tomographischen Inversion von VHR SAR-Daten wurde ein solcher *Compressive Sensing basierter Algorithmus* entwickelt mit dem Namen „S̲cale-down by $\underline{L_1}$ norm M̲inimization, M̲odel selection, and E̲stimation R̲econstruction" (SL1MMER, ausgesprochen wie „slimmer"). SL1MMER kombiniert die Vorteile der Compressive Sensing Rekonstruktion, also Überauflösungsfähigkeit und genaue Positionsbestimmung, mit der Amplituden- und Phasengenauigkeit linearer Schätzer. Der Algorithmus gibt daher verlässliche Schätzungen der Anzahl der Streuer sowie deren Elevationspositionen, Amplituden und Phasen. Die Überauflösungsfähigkeit von SL1MMER wird mit realen TerraSAR-X-Daten demonstriert und zeigt eine beachtliche Steigerung der Anzahl der Zweifachstreuer auf bis zu 38%.

Eine *systematische Analyse der Leistungsfähigkeit* des vorgeschlagenen SL1MMER-Algorithmus zeigt dessen Elevations-Schätzgenauigkeit, Überauflösungsfähigkeit und Robustheit auf. Sowohl numerische wie auch analytische Näherungen der Elevations-Genauigkeit werden mit der Cramér-Rao-Grenze verglichen. Es wird gezeigt, dass SL1MMER ein *effizienter* Schätzer ist.

Die erreichbare Überauflösung wird durch extensive Simulationen bestimmt. Diese Überauflösungsfaktoren stellen zugleich absolute und grundsätzliche Grenzen für Spektralschätzer dar. Für den typischen Parameterbereich in der SAR-Tomographie werden Überauflösungsfaktoren von 1.5 – 25 erreicht. Schließlich wird die Mindestanzahl von Datensätzen zur robusten topographischen Rekonstruktion hergeleitet.

Bisherige tomographische Inversionsverfahren berücksichtigen nur lineare Bewegung, obwohl Deformationen von Gebäuden oder der Erdoberfläche meist nichtlinear sind, z.B. periodisch, beschleunigt, oder abrupt. Daher wurde eine *generalisierte „Time Warp" Methode* entwickelt, mit der die Aufnahmezeiten auf beliebige Basisfunktionen abgebildet werden, so dass nichtlineare multi-modale Bewegungen geschätzt werden können.

Alle im hier entwickelten Verfahren und Algorithmen wurden sowohl in Simulationen wie auch durch Verarbeitung großer realer TerraSAR-X-Datenstapel validiert.

Die mit dieser Dissertation präsentierte Arbeit möge ein substantieller Beitrag sein zu der Vision *„Ein dynamisches Stadtmodell, das Form und Deformation jedes Gebäudes zeigt"*.

Acknowledgments

In the past three years, I indulged in the development of TomoSAR with the hope to make it more mature. While doing so, my personal growth has been significant, which is attributable to many people who helped me.

In this period, Prof. Richard Bamler has been my supervisor, mentor and friend. As my supervisor, Richard is always there to listen and to give advice. He promotes me and provides me the best opportunities for the further development of my academic carrier. His perfectionism strongly infected me and made me much better organized. As a mentor, he guided me through when I was wondering about my life and future path. He is like a father who taught me the importance of "Haftpflichtversicherung", "Stapler"... Among my friends, he belongs to the ones who understand well my black humor. Overall, he deserves my deeply gratitude appreciation for many reasons.

I am especially grateful to my second supervisor Prof. Uwe Sörgel from Leibniz Universität Hannover. He is an active figure in the society of radar remote sensing. And his contribution in this field, especially in urban areas, has been remarkable. I must thank for his constant attention and valuable support to my work. I greatly appreciate Prof. Moreira, from Deutsches Zentrum für Luft- und Raumfahrt (DLR). He is a pioneer in TomoSAR, who presented the first demonstration of airborne SAR tomography. My work has been greatly inspired by his innovation and creativity. It is really my honor to invite him as my second external reviewer.

Besides my supervisors, Prof. Liqiu Meng has always been my idol, not only because of her profound knowledge and great wisdom, but also her superior cooking. I appreciate a lot her constructive suggestions and selflessly help. I highly appreciate my previous teachers who pay continuous attention to my academic development including Prof. Rummel who is the father of the ESPACE master course that eventually brought me here to Germany, Prof. Stefan Hinz who inspired my initial interest in remote sensing, Prof. Hogentobler, Prof. Liu Weidong and Prof. Xi Xiaoning. A special mention goes to Dr. Ramon Brcic for the valuable discussions and proof-reading of my dissertation during the weekend.

I am deeply grateful to Dr. Gianfranco Fornaro who hosted me during my stay in Naples and took me to explore the best espresso and mozzarella of the world.

My warm gratitude goes to all the people of TUM-LMF and DLR-IMF for the friendly working atmosphere: Dr. Michael Eineder who is the team leader of my PhD project, Dr. Matthias Butenuth for his patience to my wrongly formatted "Reiseabrechnung", Nico Adam who is an expert in PSI and the developer of the GENESIS system, Stefan Gernhardt who invited me for the first western wedding, Alessandro Parizzi for teaching me Italian, Stefan Auer for his cooperation, Archy Wang who trusts me unconditionally, Ke Zhu for organizing Feuertopf activities and brought laughter to our office, Daniel Frey who is always willing to help, Xiaoying Cong who accompanied me along, Marie Lachaise as almost my best friend ... the list goes on and on.

Last but not least, I would like to thank my family: my parents, for given birth to me in the first place, for educating me and for their unconditional support; my sister, for sharing her life with me; my boyfriend Yilei, for his love, for cheering me up and for bringing me courage all the time. This thesis is dedicated to all those who love me, in particular to my mother and Yilei.

Contents

List of abbreviations

2-D	two-dimensional
3-D	three-dimensional
AIC	Akaike information criterion
APS	atmospheric phase screen
BIC	Bayesian information criterion
CAT	computed axial tomography
COSMO-SkyMed	COnstellation of small Satellites for the Mediterranean basin Observation
CRLB	Cramér-Rao lower bound
CS	compressive sensing
D-InSAR	differential interferometric SAR
D-TomoSAR	differential SAR tomography
DEM	digital elevation model
DLR	German Aerospace Center
E-SAR	Experimental SAR of DLR
ERS	Earth Resources Satellite
GENESIS	generic system for interferometric SAR
GLRT	generalized likelihood ratio test
InSAR	interferometric SAR
LASSO	least absolute shrinkage and selection operator
LiDAR	light detection and ranging
LOS	line-of-sight
MAP	maximum a posteriori
MB	multi-baseline
MD	maxima detection
MDL	minimum description length
MLE	maximum likelihood estimator
MUSIC	multiple signal classification
NLS	nonlinear least squares
PCT	polarization coherence tomography
PS	persistent scatterer

PSI   persistent scatterer interferometry

RIP   restricted isometry property

SAR   synthetic aperture radar

SNR   signal-to-noise ratio

SR   super-resolution

SV   singular value

SVD   singular value decomposition

TanDEM-X   TerraSAR-X add-on for Digital Elevation Measurement

TDBP   time-domain back-projection

TomoSAR   SAR tomography

TSVD   truncated singular value decomposition

VHR   very high resolution

List of Symbols

$(\cdot)^{\mathrm{H}}$	transpose conjugate
$(\cdot)^{\dagger}$	pseudoinverse
$C(K)$	complexity penalty
K	sparsity (number of scatterers inside an azimuth-range pixel)
M	order of the motion model
N	number of measurements
N_L	number of looks
$N_{min,90\%}$	minimal required number of acquisitions at a probability of detection of 90%
P_D	probability of detection
SNR_q	SNR of the qth scatterer
$V(s)$	LOS linear deformation velocity profile along elevation s
W	chirp bandwidth
$\Delta\varphi$	phase difference between two scatterers
Δb	elevation aperture size
Δs	extent of scattering objects in elevation
Δv	range of possible velocities
Δx	azimuth synthetic aperture length
α	normalized distance between two scatterers (distance devided by ρ_s)
$\mathbf{J}$	Fisher information matrix
$\mathbf{\Phi}$	N possibly random basis functions (the sensing matrix)
$\mathbf{\Psi}$	orthogonal basis
$\boldsymbol{\gamma}$	discrete reflectivity vector with L elements (sparse signal)
$\boldsymbol{\nu}$	any vector having K non-zero coefficients at the same positions as $\boldsymbol{\gamma}$
$\boldsymbol{\theta}$	unknown parameter vecotor, including elevation positions s_1 and s_2, amplitudes a_1 and a_2 and phases φ_1 and φ_2
$\boldsymbol{\varepsilon}$	measurement noise
δ_s	a small positive number
δr	range migration caused by the different viewing angles
η_n	velocity frequency

$\eta_{m,n}$	m^{th} temporal frequency component at t_n
$\gamma(s)$	continuous reflectivity function along elevation
$\hat{K}$	estimated sparsity (number of scatterers)
$\hat{\mathbf{C}}_{\mathbf{gg}}$	estimated covariance matrix from multi-looking measurements, also called data or sample covariance matrix
$\hat{\boldsymbol{\gamma}}$	estimated discrete reflectivity
$\hat{\boldsymbol{\gamma}}(\hat{\mathbf{s}})$	estimated sparse reflectivity profile, i.e. estimated complex-valued reflectivity of discrete scatterers
κ_{P_D}	super-resolution factor at a probability of detection of P_D
λ	wavelength
λ_K	Lagrange multiplier
μ	coherence
ρ_r	range resolution
ρ_s	elevation resolution
ρ_x	azimuth resolution
ρ_{P_D}	minimal separable distance at a probability of detection of P_D
σ_ε	noise level
σ_n	n^{th} non-negative singular value of R
$\sigma_{s_q,0}$	CRLB of the elevation estimates of the qth scatterer in the absence of other scatterers
σ_{s_q}	CRLB of the elevation estimates of the qth scatterer in the presence of a second scatterer
$\tau_m(t_n)$	m^{th} motion base function
θ	elevation angle (incidence angle)
φ_q	phase of qth scatterer
ξ_n	spatial (elevation) frequency
a	amplitude of the periodic motion
a_q	amplitude of qth scatterer
b_n	n^{th} elevation aperture position
c	speed of light
c_0	σ_{s_q} normalized to $\sigma_{s_q,0}$
$c_{SR,i}$	coefficient for the ith order term of $N \cdot SNR$
g_n	focused complex-valued measurement of an azimuth-range pixel for the nth acquisition

k	number of parameters of a model, describing the model complexity
$p_m(s)$	m^{th} motion coefficient
r	range coordinate
s	elevation coordinate
s_q	elevation of qth scatterer
t_0	initial time offset of a seasonal motion
t_n	n^{th} acquisition time
v	LOS linear deformation velocity
x	azimuth coordinate
y	ground range coordinate
z	height coordinate
$\mathbf{C_{gg}}$	covariance matrix
$\mathbf{R}$	mapping matrix (sparse mapping operator)
$\mathbf{e}$	combined measurement noise and the model error
$\mathbf{g}$	measurement vector
$\mathbf{g}_{n_l}$	measurement vector of the $n_l{}^{th}$ look ($n_l = 1, \dots, N_L$)
$\mathbf{r}_l$	l^{th} column of R
$\mathbf{u}_n$ and $\mathbf{v}_n$	n^{th} left and right singular vector of R, respectively
$\mathbf{x}$	signal of interest with a length of L
$\mathbf{C}_{\gamma\gamma}$	covariance matrix of the prior
$\mathbf{C}_{\varepsilon\varepsilon}$	noise covariance matrix
$\hat{\mathbf{s}}$	estimated elevation positions
$\boldsymbol{\theta}(K)$	vector of the unknown amplitudes, phases, and elevations for K scatterers
$\hat{\mathbf{G}}$	noise subspace of $\mathbf{C_{gg}}$
$d(s, t_n)$	LOS displacement as a function of elevation and time
$f_{(\cdot)}$	frequency, corresponding subscripts represents the directions
$p(\mathbf{g}|\boldsymbol{\theta}(K), K)$	likelihood of parameter vector $\boldsymbol{\theta}(K)$ given measurements $\mathbf{g}$
$|\hat{\gamma}_l|^2$	estimated reflected power at elevation position $s_l (l = 1, \dots, L)$
$|\hat{\gamma}_l|^2_{MU}$	estimated "pseudospectrum" at elevation position $s_l (l = 1, \dots, L)$

1   Motivation and objectives

In Earth observation, satellite remote sensing plays an important role since it enables us to recover contact-free large-scale information about the physical properties of Earth's surface. In particular, unique capabilities are associated with the use of remote sensing via synthetic aperture radars (SARs) which generate a microwave reflectivity map of an illuminated area through transmission and reception of electromagnetic energy. As an active sensor, SAR functions independently of solar illumination (Curlander and McDonough, 1991) (Cumming and Wong, 2004). It is capable of penetrating clouds, (partially) vegetation canopy, soil and snow (Franceschetti and Lanari, 1999) (Massonnet and Souyris, 2008). Compared to images obtained by optical and infrared sensors, SAR images contain different information. In the optical image, molecular resonances on the object surfaces are mainly responsible for the spectral characteristics of the object reflectivity; in the microwave region, dielectric and geometrical properties become more relevant for backscattering (Elachi, 1988). As a consequence, radar images emphasize the humidity, relief and morphological structure of the observed terrain.

SAR is the only remote sensing method to assess deformation on the order of centimeters and millimeters due to its coherent nature and short wavelengths (typically 3-25 cm). Space-borne SAR systems are particularly suited for long-term monitoring of such dynamic processes (Ferretti et al., 2000) (Kampes and Adam, 2005). A single SAR image can only provide cartographic information. In order to retrieve the real three-dimensional (3-D) localization and motion information of the scattering objects, *advanced interferometric methods* are required which exploit stacks of complex-valued SAR images with diversity in space and time.

The topic of this thesis is *very high resolution tomographic SAR inversion for urban infrastructure monitoring*. To this end, very high resolution (VHR) SAR tomography (TomoSAR) and differential SAR tomography (D-TomoSAR) are demonstrated using TerraSAR-X spotlight data for providing 3-D and 4-D map of an entire high rise city area including layover separation examples. A compressive sensing (CS) based algorithm (SL1MMER) tailored to VHR SAR data is proposed for tomographic SAR inversion by exploiting the sparsity of the signal in elevation. A generalized time warp method is proposed which enables D-TomoSAR to estimate multi-component nonlinear motion.

1.1   Scientific relevance of the topic

Conventional SAR imaging provides a projection of the 3-D object reflection to the two-dimensional (2-D) azimuth-range (x-r) plane (Curlander and McDonough, 1991). Due to the side-looking geometry, this projection introduces typical foreshortening, layover and shadowing problems which handicap the interpretation of SAR images, especially of 1) volumetric scatterers and 2) of urban areas and man-made objects, i.e. objects with surface elements oriented at steeper angles than the local incidence angle (Hanssen, 2001). Advanced interferometric SAR techniques enable 2.5-D, 3-D, 4-D or even higher dimensional SAR imaging (Bamler et al., 2009). That means the 3-D shape and deformation or subsidence of the individual buildings can be retrieved. In the following, the term "pixel" means an azimuth-range pixel and the term "multiple scatterers" means the presence of several scattering objects at different elevation positions but at the same pixel.

Across-track *Interferometric SAR (InSAR)* techniques combine two or more complex-valued SAR images acquired from slightly different positions to determine geometric information about the imaged objects by exploiting the phase differences (Zebker and Goldstein, 1986) (Massonnet et al., 1993) (Gabriel and Goldstein, 1988) (Bamler and Hartl, 1998) (Gatelli et al., 1994) (Ruf et al., 2002) (Eineder et al., 2009). This phase difference is related to the terrain topography

and can be used to generate digital elevation models (DEMs) (Li and Goldstein, 2002). There-fore, InSAR enables us to access the third dimension, elevation s, which is perpendicular to the azimuth and range plane. Still, with only two acquisitions, multiple scatterers along elevation within one pixel cannot be separated. Besides the topographic mapping, an extended version of InSAR, called *differential interferometric SAR (D-InSAR)* can be used for precise mapping of movement in the range direction (Zebker and Rosen, 1994) (Goldstein and Werner, 1998) (Massonnet and Rabaute, 2002) (Berardino et al., 2003) (Bamler and Eineder, 2005). The accuracy of InSAR and D-InSAR is limited by temporal decorrelation of the surface and by electromagnetic path delay variations in the troposphere (Hanssen, 2001). The latter distortions can be reduced by temporal averaging of multiple interferograms which in turn reduces the temporal resolution.

Persistent scatterer interferometry (PSI) was introduced in 1999 (Ferretti et al., 2000) (Ferretti et al., 2001) as a methodology for long-term monitoring of subsidence, preferably in urban environments. Typically 20 - 100 interferometric data sets of the same area taken from approximately the same orbit are stacked for PSI analysis. Each pixel in each interferogram is characterized by its range and azimuth coordinates as well as by the temporal and the spatial baseline of the interferogram it belongs to. These are four coordinates which allow a much better data analysis than in a single interferogram where temporal and spatial baselines are fixed. By exploring different scales of the physical phenomena, PSI attempts to separate the following phase contributions: elevation of the point, deformation parameter (e.g. deformation rate, amplitude of seasonal motion), orbit errors and tropospheric water vapor delay. (Colesanti et al., 2002) (Hooper et al., 2004) (Adam et al., 2005) (Kampes, 2006) (Meyer et al., 2007) (Perissin and Ferretti, 2007) (Rocca, 2007) (Vasco et al., 2008) (Gernhardt et al., 2010). PSI assumes the presence of only a single scatterer in the pixel, however, in a high rise urban environment, it is rather likely that multiple scatterers are mapped onto one pixel.

TomoSAR aims at real and unambiguous 3-D SAR imaging, i.e. imaging also in elevation (Homer et al., 1996) (Pasquali et al., 1995) (Reigber and Moreira, 2000) (Cheney, 2000) (Fornaro et al., 2003) (Lombardini et al., 2003) (Zhu et al., 2008) (Nannini et al., 2009) (Baselice et al., 2009) (Reale et al., 2011). TomoSAR uses SAR data stacks, like PSI does, to establish a synthetic aperture in the elevation direction (see section 2.1.2). While in PSI the coordinates of single points are retrieved, TomoSAR tries to derive the full scattering density, i.e. the reflectivity profile, in elevation by spectral analysis with special consideration of the difficulties caused by sparse and irregular sampling. From this reconstructed profile in elevation multiple scatterers in any pixel are separated, and hence the full 3-D (azimuth, range and elevation) reflectivity distribution is obtained. Therefore, TomoSAR is the strictest way of 3-D SAR imaging while classical InSAR can be regarded as the limiting case of parametric TomoSAR.

D-TomoSAR, also referred to as *4-D focusing* (Lombardini, 2005) (Fornaro et al., 2009) (Zhu et al., 2010), exploits the strength of both TomoSAR and PSI. It uses the fact that the different acquisitions are taken at different times and introduces a new dimension to the TomoSAR system model attributing to the motion of the scatterers (Lombardini, 2005). By means of higher dimensional spectral analysis, it is capable of retrieving elevation *and* deformation information even of multiple scatterers inside a single SAR pixel. D-TomoSAR provides the most advanced means for space-time 4-D SAR imaging.

Along with the launch of new SAR sensors, like the German TerraSAR-X/TanDEM-X and the Italian COSMO-Skymed satellites, SAR remote sensing from space made a big leap forward. Among all other advantages, such as high absolute geometric accuracy, precise orbit determination and short revisit time, this new class of SAR sensors deliver SAR data with a very high spatial resolution of up to 1 m compared to the medium (10 ∼ 30 m)- and high (3 ∼ 10 m)-resolution SAR systems available so far (Bamler et al., 2009) (Eineder et al., 2009) (Werninghaus and Buckreuss, 2010). In Fig. 1 and 2 the mean intensity maps of a TerraSAR-X stack

Figure 1. TerraSAR-X mean intensity map from 30 images of a high rise urban area over downtown Las Vegas with a spatial resolution of 1.1 m × 0.6 m in azimuth and range.
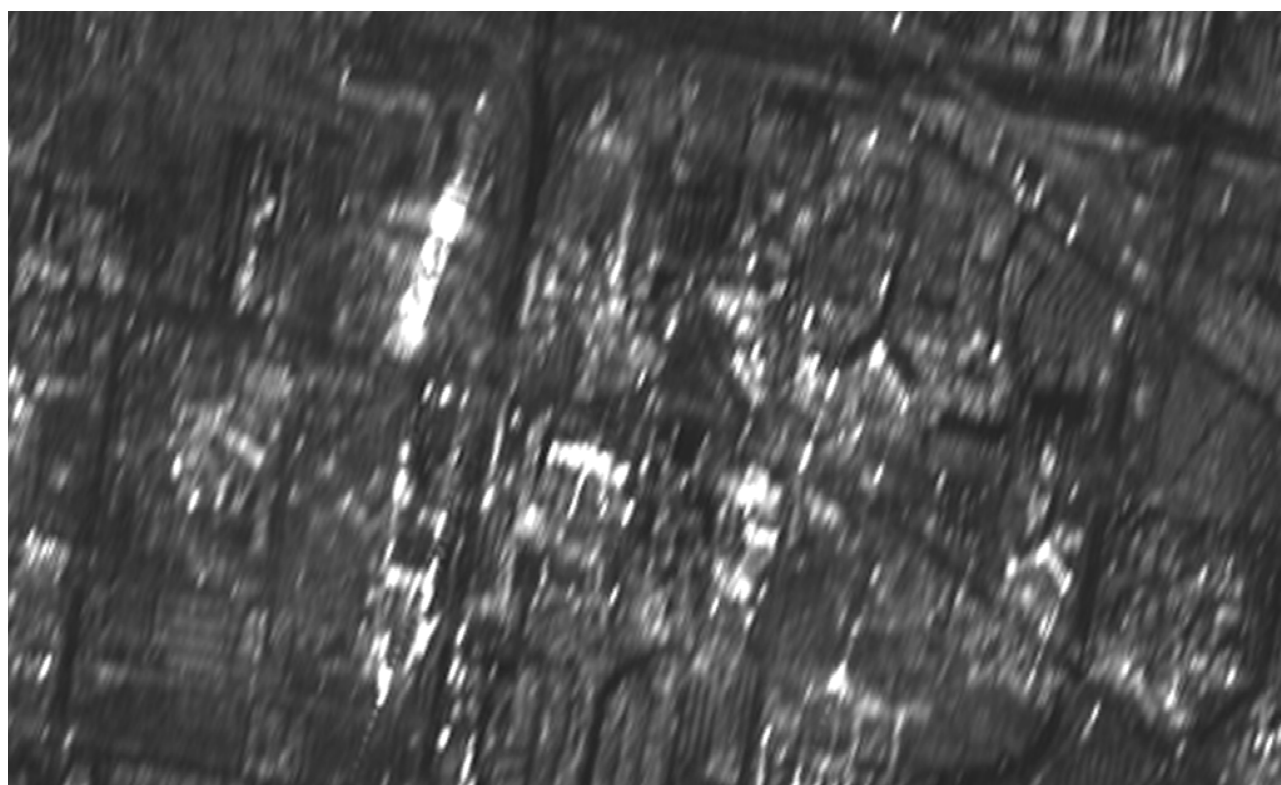

Figure 2. ERS mean intensity map from 51 images of the same area shown in Fig. 1 with a spatial resolution of 5 m × 25 m in azimuth and range.

(30 images) and an Earth Resources Satellite (ERS) stack (51 images) of the same area are presented, respectively.

The 1 m resolution opens up for the first time the opportunity to use SAR for 2-D, 3-D, 4-D (space-time) or even higher dimensional imaging of urban structures and individual buildings from space. The inherent spatial scales of buildings are dominated by the typical height of floors of $3 \sim 3.5$ m (i.e. in slant range (at 30°) $2.6 \sim 3.0$ m) and the spacing between windows. Hence, in the context of urban structure imaging, a tremendous improvement in information content is observed when going from high resolution (3 m) to VHR (1 m). The advantage of VHR imagery for cartographic applications is obvious (Soergel et al., 2006). Yet the real potential of this class of SAR data for urban mapping lies in applications where the coherent nature of SAR data is exploited, such as in interferometry or tomography (Bamler et al., 2009) (Soergel, 2010).

As a consequence, further developing the above mentioned advanced interferometric techniques to explore the potentials of VHR SAR data for urban infrastructure mapping is a new and exciting field.

1.2   Objective and focus of the thesis

1.2.1   General goal

The general goal of the thesis is the optimum exploitation of interferometric VHR SAR data stacks for urban infrastructure mapping and monitoring with the vision of

"A dynamic city model derived from TomoSAR and D-TomoSAR showing the shape and the deformation of each building".

To this end, VHR TomoSAR has to be demonstrated by providing a layover separation example using TerraSAR-X spotlight data. Together with its extension to D-TomoSAR, multiple scatterers inside an azimuth-range cell have to be identified and resolved from sparsely and irregularly sampled apertures considering different motion histories. Moreover, a 3-D and 4-D map of an entire building complex and even an entire high rise city area has to be reconstructed from a tomographic stack of SAR images.

1.2.2   Methodology goals

Compared to computed axial tomography (CAT), known from medical imaging, TomoSAR uses only a small angular diversity. Hence, instead of back projection, spectral estimation (Jakowatz, 1996) (Stoica and Moses, 2005) is sufficient for TomoSAR if the range migration caused by the different viewing angles is much smaller than the range resolution (Reigber, 2002) (Zhu and Bamler, 2010d). Many spectral estimation methods are implemented for tomographic SAR imaging. However, for urban infrastructure monitoring, the following requirements must be met:

  - Maintaining the meter azimuth-range resolution, i.e. the method should require no multi-looking;
  - Improving the elevation resolution, i.e. it has to provide super-resolution (SR) properties;
  - High 3-D localization accuracy, i.e. it is required to provide robust performance and deliver high elevation estimation accuracy;
  - Retrieval of nonlinear motion.

Aiming at the development of tomographic algorithms fulfilling the above mentioned requirements, three methodology objectives must be achieved:

Objective 1: New tomographic method
Tomographic SAR inversion is a relatively new technique and has not yet been exploited with VHR SAR data over urban areas. The goal is to develop a robust TomoSAR and D-TomoSAR algorithm that can go far beyond the Rayleigh resolution limit and gives reliable estimation of the number of scatterers and the elevation, motion parameters, complex-valued reflectivity of each scatterer from a few sparse and irregular spectral samples while requiring no azimuth-range averaging.

Objective 2: Performance assessment
Systematic performance assessment of the tomographic method should be performed for deriving confidence measures for the final space-time estimates. The elevation estimation accuracy of two closely spaced interfering scatterers needs to be investigated, e.g. by comparison to the Cramér-Rao lower bound (CRLB); the SR power, i.e. the minimal separable distance between two scatterers under certain conditions, must be assessed and compared to the Rayleigh resolution unit; the robustness of algorithm, i.e. the minimal number of acquisitions required for a robust reconstruction given a certain signal-to-noise ratio (SNR) has to be estimated. Finally, the limits of the algorithm, especially with the application to tomographic SAR inversion, should be determined.

Objective 3: Multi-component motion estimation
The motion history appears as a phase term in the measurements. This phase term, however, is often nonlinear (periodic, accelerating, stepwise, etc.), and hence cannot be retrieved by spectral estimation within the framework of conventional D-TomoSAR. This is particularly true if very high resolution mapping of urban infrastructure is to be performed. Complicated motion, e.g. linear subsidence caused by long term geodynamic processes coupled with seasonal motion caused by thermal dilation, must be taken into account. Unmodelled motion not only leads to wrong motion estimates, but also interferes with elevation estimation. The goal is to extend D-TomoSAR for multi-component nonlinear motion separation and estimation.

1.3   Reader's guide

The remainder of this thesis is organized as follows. Chapter 2 gives a brief introduction to the basics of tomographic SAR inversion and CS. An overview on the state of the art, followed by a discussion of the new aspects of this thesis regarding to most related work is presented in Chapter 3. Chapter 4 summarizes the four main contributions of the attached relevant publications from the author including 1) the first demonstration of VHR TomoSAR and D-TomoSAR; 2) the SL1MMER algorithm for sparse tomographic SAR reconstruction; 3) a systematic performance assessment of elevation estimation accuracy, SR power and robustness; 4) the time warp method for multi-component nonlinear motion estimation in D-TomoSAR. Furthermore, a practical demonstration of the SR of SL1MMER for SAR tomographic reconstruction is provided in Chapter 5 (not yet published). The results of this thesis and an outlook on further work are discussed in the last chapter.

2   Basics

2.1   Basics of SAR tomography

2.1.1   SAR geometry

A normal monostatic imaging SAR system consists of a side-looking transmitter and receiver operated on a moving platform such as an airplane or satellite. The native 3-D reference frame of a SAR sensor is defined as (see Fig. 3):

x: *azimuth*, the flight direction of the sensor, also called the *along-track* direction;

r: *range*, line-of-sight (LOS) direction of the antenna, also referred to as *slant range*;

θ: *elevation angle*, i.e. perpendicular to the azimuth-range plane. Due to the large range distance on the order of hundreds kilometer and small angular diversities, it can be assumed to be a straight line. It is often called *elevation* or *cross range s*

SAR maps in the three directions with different principles:

- Azimuth: synthetic aperture principle enabled by the movement of the sensor.
- Range: radar ranging principle, i.e. the transmitter emits pulses, e.g. chirps. These pulses are reflected from objects on the ground and part of the energy reaches again the receiver. Different objects can be resolved if they have a different distance to the sensor, i.e. their echoes show a different two-way travel time.
- Elevation: SAR does not provide imaging in this direction, and hence targets sharing the same azimuth-range $(x - r)$ coordinates but with different elevation positions can not be distinguished.

The complex-valued pixels of a conventional two-dimensional backscattering map contain amplitude (brightness) and phase information. The non-ambiguous phase ϕ of a pixel with a range distance r and at zero Doppler position x can be written as (omit an additional phase caused by the scattering process):

$$\phi = -\frac{4\pi}{\lambda}r \tag{1}$$

where λ is the wavelength. Since λ is typically several centimeters, the phase is very sensitive to range.

The 3-D resolution element of a conventional SAR is illustrated in Fig. 4. The azimuth resolution ρ_x of the SAR image is determined by the azimuth synthetic aperture length Δx (Massonnet and Souyris, 2008):

$$\rho_x = \frac{\lambda\, r}{2\Delta x} \tag{2}$$

The TerraSAR-X spotlight data used in this thesis have an azimuth resolution of 1.1 m.

The achievable range resolution ρ_r depends on the chirp bandwidth W (Cumming and Wong, 2004):

$$\rho_r = \frac{c}{2W} \tag{3}$$

where c denotes the speed of light. For instance, depending on the transmitted bandwidth, the range resolutions provided by TerraSAR-X Spotlight data are 1.2 m at 150 MHz signal
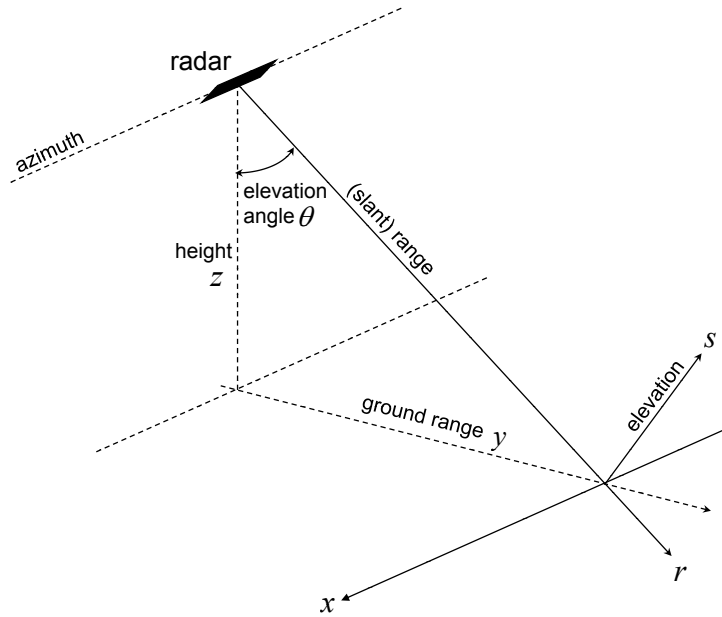
Figure 3. Radar side-looking imaging geometry. The flight direction of the sensor is normally called the *along-track* direction or *azimuth x*; the LOS direction of the antenna is defined by the elevation angle θ, normally called *range* or *slant range r*; the *cross range* direction, i.e. perpendicular to the azimuth-range $(x - r)$ plane, is frequently referred to as *elevation s*. Strictly speaking, the third dimension should be the elevation angle θ, however, due to the large range distance on the order of hundreds kilometer, it can be assumed to be a straight line for small angular diversities.
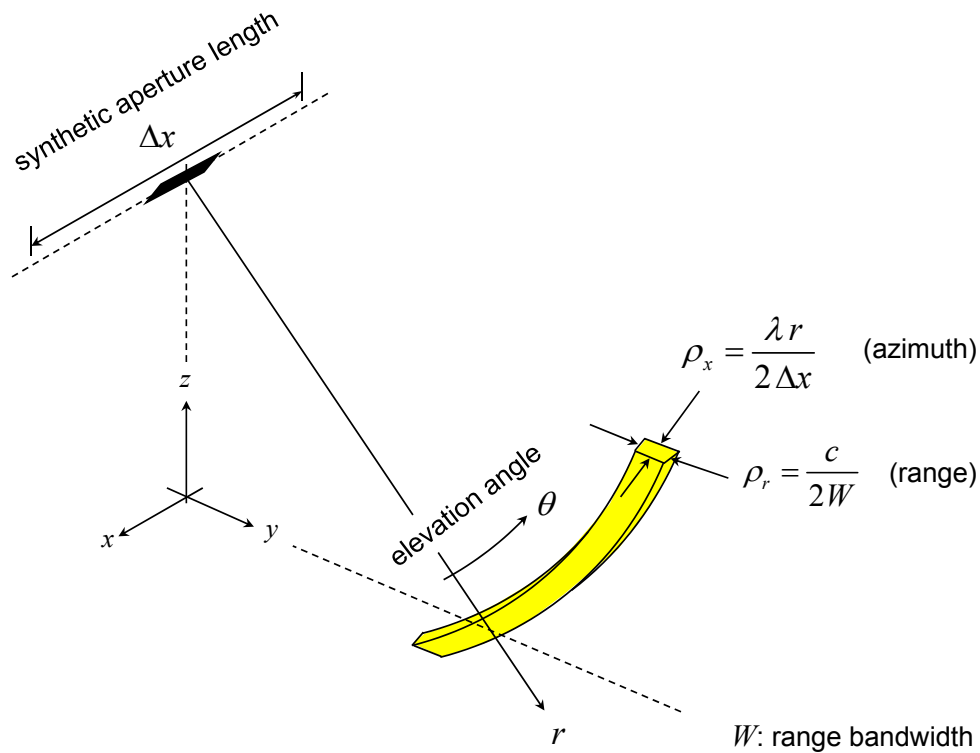

Figure 4. 2-D resolution element of a conventional SAR. The azimuth resolution ρ_x is determined by the azimuth synthetic aperture length Δx while the range resolution ρ_s is determined by the bandwidth W of the chirp emitted by the antenna (Bamler, 2010).

bandwidth or 0.6 m at 300 MHz (Werninghaus and Buckreuss, 2010). The small deviation from the theoretical values 1 m and 0.5 m is due to spectral weighting for sidelobe reduction. As mentioned before, SAR does not provide imaging in the third dimension, therefore, as illustrated in Fig. 4, conventional SAR has a hoop-shaped 3-D resolution element, i.e. no resolution in elevation.

2.1.2   TomoSAR imaging model

TomoSAR builds up a synthetic elevation aperture from a stack of N complex SAR datasets of the same area taken at different times and slightly different orbit positions (see Fig. 5). Based on the Fraunhofer approximation of free-space propagation, it is well known that far-field diffraction acts like a Fourier transform. The focused complex value g_n of an azimuth-range pixel for the nth acquisition at aperture position b_n ($n = 1, ..., N$) is therefore the integral of the reflected signal along the elevation direction weighted by a linear phase term (note that a possible deformation term is ignored here for simplicity):

$$g_n = \int_{\Delta s} \gamma(s) exp(-j2\pi\xi_n s) \mathrm{d}s \tag{4}$$

where $\gamma(s)$ represents the reflectivity function along elevation s. $\xi_n = -2b_n/(\lambda r)$ is the spatial (elevation) frequency. A detailed derivation of eq. (4) can be found overall in the literature, e.g. (Fornaro et al., 2003) (Zhu, 2008). The imaging model eq. (4) can be approximated by discretizing the continuous reflectivity function along s within its extent Δs, in the presence of noise ε, and can be written as (ignoring an inconsequential constant):

$$\mathbf{g} = \mathbf{R}\,\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{5}$$

where $\mathbf{g} = (g_1, g_2, ..., g_N)^T$ is the measurements vector with N elements, $\mathbf{R}$ is an $N \times L$ mapping matrix with $R_{nl} = \exp(-j2\pi\xi_n s_l)$, $\boldsymbol{\gamma}$ is the discrete reflectivity vector with L elements, $\gamma_l = \gamma(s_l)$, and s_l ($l = 1, ..., L$) are the discrete elevation positions.
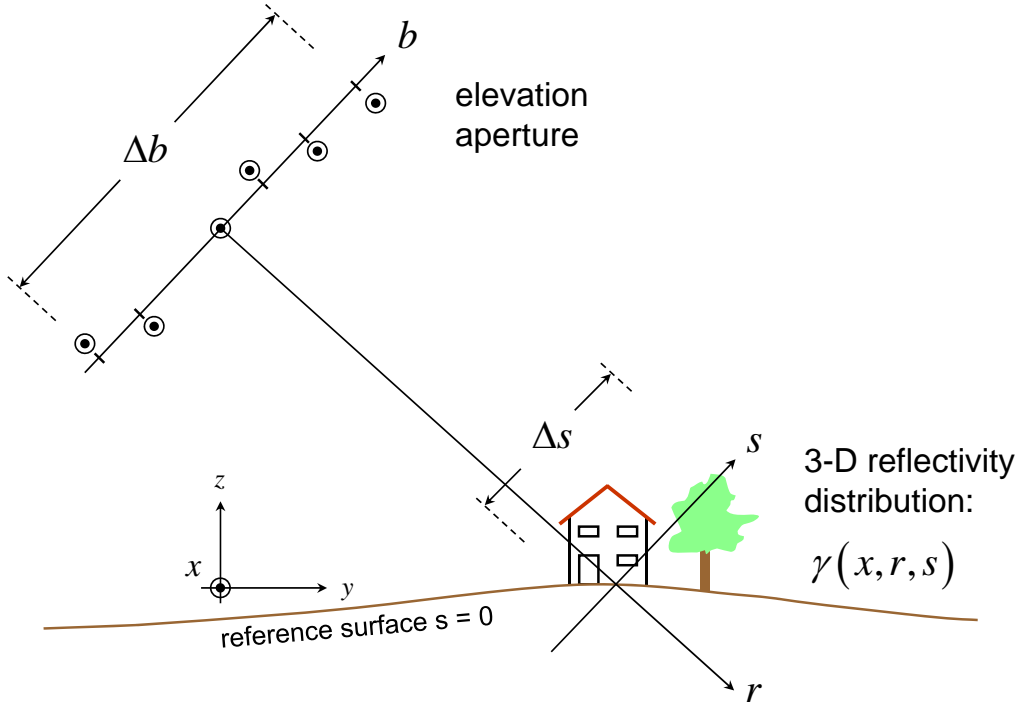

Figure 5. TomoSAR imaging geometry. The elevation synthetic aperture is built up by multi-pass SAR data from slightly different viewing angles. Flight direction is orthogonal out of the plane.

Eq. (5) is an irregularly sampled discrete Fourier transform of the elevation profile $\gamma(s)$. Therefore, one SAR acquisition may be considered to be one spectral coefficient of the complex reflectivity of the object along elevation. Fig. 6 illustrates the Fourier domain TomoSAR imaging model showing the slices (blue bold) of the object spectrum in the $(f_y - f_z)$ plane that are
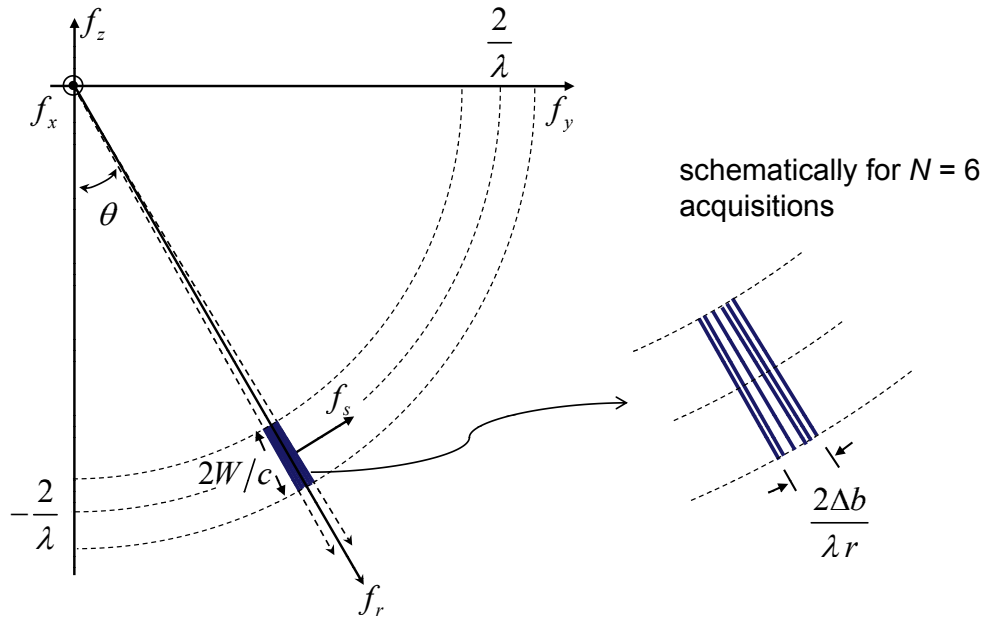
Figure 6. Fourier domain TomoSAR imaging model showing the slices (blue bold) of the object spectrum in the $(f_y - f_z)$ plane that are transferred to the SAR image. A rectangular shape range system transfer function of bandwidth W is assumed. The shaded area shows the maximum extent of the range-elevation frequency $(f_r - f_s)$ support of the tomographic transfer function, determined by chirp bandwidth W and elevation aperture size Δb (see the zoom in on the right).
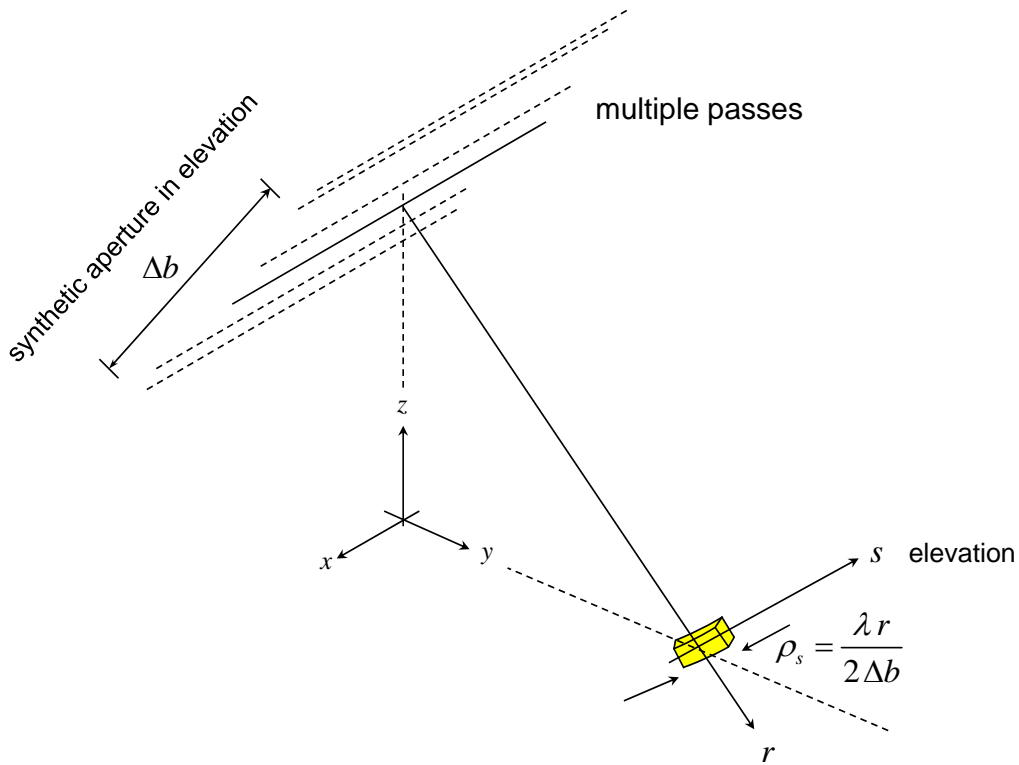

Figure 7. 3-D resolution element of a tomographic SAR, similar to the azimuth direction, the elevation resolution ρ_s is determined by the elevation aperture size Δb (Bamler, 2010).

transferred to the SAR image, where $f_{(\cdot)}$ stands for frequency and the corresponding subscripts represents the directions defined in Fig. 3. TomoSAR uses several of those slices obtained from multi-pass measurements to retrieve the elevation reflectivity profile for each pixel. A special limiting case of TomoSAR is InSAR where only two slices are used.

Similar to the azimuth direction, for non-parametric spectral analysis, the expected elevation resolution ρ_s, i.e. the width of the elevation point response function, depends on the elevation aperture size Δb and is approximately (provided there is sufficiently dense sampling of the elevation aperture) (see Fig. 7):

$$\rho_s = \frac{\lambda\, r}{2\Delta b} \tag{6}$$

As illustrated in Fig. 6, due to the tight orbit control of modern VHR SAR sensors, the angular diversity between the slices is very small, and hence renders the 3-D SAR resolution element extremely anisotropic. This will be further discussed in Chapter 4.

As mentioned in Chapter 1, TomoSAR can be treated as a spectral estimation problem, only if the range migration δr caused by the different viewing angles is much smaller than the range resolution ρ_r. This gives a limitation to the extent Δs of the illuminated objects (Zhu and Bamler, 2010d):

$$\Delta s \ll \frac{\rho_r r}{\Delta b} \tag{7}$$

The condition (7) is mostly true for the space-borne case due to the large range distance and small angle diversity. However, for the airborne case, instead of spectral estimation, back projection might be required as in CAT (Reigber and Moreira, 2000).

2.1.3   Differential TomoSAR system model

For the space-borne case, the multi-pass acquisitions are taken over a time span of several weeks to years (depending on the revisit time of the satellite and the number of stacked images). Hence, the long-term motion of the scattering object during the acquisition period must be considered. Taking a possible motion term into account, the system model eq. (4) can be extended to (Lombardini, 2005):

$$g_n = \int\limits_{\Delta s} \gamma(s) exp\left(-j2\pi\left(\xi_n s + 2d(s,t_n)/\lambda\right)\right) \mathrm{d}s \tag{8}$$

where $d(s,t_n)$ is the LOS displacement as a function of elevation and time. By assuming a *linear* motion with elevation dependent velocity $V(s)$ but possibly several scatterers per pixel, eq. (8) simplifies to:

$$g_n = \int\limits_{\Delta s} \gamma(s) exp\left(-j2\pi\left(\xi_n s + \eta_n V(s)\right)\right) \mathrm{d}s \tag{9}$$

$\eta_n = 2t_n/\lambda$ is the so-called "velocity frequency". Eq. (8) can be rewritten as (Zhu and Bamler, 2010d) (Zhu and Bamler, 2011a):

$$g_n = \int\limits_{\Delta v}\int\limits_{\Delta s} \gamma(s)\delta\left(v - V(s)\right) exp\left(-j2\pi\left(\xi_n s + \eta_n v\right)\right) \mathrm{d}s \mathrm{d}v \tag{10}$$

where Δv is the range of possible velocities. An alternative, but equivalent, formulation of the D-TomoSAR system model with the so called "velocity spectrum" can be found in (Fornaro et al., 2009).

Eq. (10) is a 2-D Fourier transform of $\gamma(s)\delta\left(v - V(s)\right)$ which is a delta-line in the elevation-velocity $(s - v)$ plane along $v = V(s)$. Its projection onto the elevation axis is the reflectivity profile $\gamma(s)$.

If only K discrete scatterers are present, $\gamma(s)\delta\left(v - V(s)\right)$ consists of K δ-points in the 2-D $(s-v)$ plane at the respective elevations and velocities. The tomographic reconstruction will be a blurred and noisy version of these peaks from which elevations and velocities are estimated. Of course, eq. (10) can also be approximated by a discrete version sharing the same expression as eq. (5).

2.2 Introduction to compressive sensing

2.2.1 Compressive sensing

The well-known Shannon sampling theorem states that for a band-limited signal to be reconstructed exactly from its samples, the signal needs to be sampled at of least its (mathematical) bandwidth (the so-called Nyquist rate). In fact, this principle underlies nearly all signal acquisition methods used. However, such a sampling scheme excludes many signals of interest that are not necessarily band-limited but that can still be represented by a model with a small number of degrees of freedom (Starck et al., 2010). Compressive sensing theory completes the sampling scheme and it allows us to go beyond the Shannon limit by exploiting the sparse property of the signal (Candès, 2006) (Donoho, 2006) (Baraniuk, 2007) (Candès and Wakin, 2008). It offers an aesthetic non-parametric realization of a parametric estimator by assigning some pre-conditions to the signals and the sensing systems.

A signal of interest $\mathbf{x}$ with length L is said to be K-sparse in an orthogonal basis $\mathbf{\Psi}$ if the projection coefficient vector $\boldsymbol{\gamma} = \mathbf{\Psi x}$ has only K non-zero or significant elements. $\mathbf{x}$ is represented by $\mathbf{\Psi}^{\mathrm{H}}\boldsymbol{\gamma}$. N measurements $\mathbf{g}$ can be obtained by projecting the signal onto N possibly random basis functions $\mathbf{\Phi}$ (the sensing matrix):

$$\mathbf{g} = \mathbf{\Phi x} \tag{11}$$

Let us define $\mathbf{R} = \mathbf{\Phi\Psi}^{\mathrm{H}}$ as the *sparse mapping operator*, the measurement vector can be rewritten as:

$$\mathbf{g} = \mathbf{\Psi\Phi}^{\mathrm{H}}\boldsymbol{\gamma} = \mathbf{R}\boldsymbol{\gamma} \tag{12}$$

Note that we chose the symbols $\mathbf{g}$, $\mathbf{R}$ and $\boldsymbol{\gamma}$ such that they reflect our TomoSAR problem statement of eq. (5).

CS theory tells us that in the absence of noise the exact solution of $\boldsymbol{\gamma}$ is the one satisfying eq. (12) and employing the *least number of coefficients*, i.e. having the minimal L_0 norm:

$$\min_{\boldsymbol{\gamma}} \|\boldsymbol{\gamma}\|_0 \quad \text{s.t.} \quad \mathbf{g} = \mathbf{R}\boldsymbol{\gamma} \tag{13}$$

Theoretically speaking, this is the correct solution, but unfortunately, the L_0 norm minimization problem is NP-hard. For $N \gg O(K\log(L/K))$, i.e. with $\mathbf{x}$ sufficiently sparse, it can be shown that the L_1 norm minimization leads to probably the same result as the L_0 norm minimization (Gan et al., 2009) (Candès et al., 2006):

$$\min_{\boldsymbol{\gamma}} \|\boldsymbol{\gamma}\|_1 \quad \text{s.t.} \quad \mathbf{g} = \mathbf{R}\boldsymbol{\gamma} \tag{14}$$

The use of the L_1 norm as a sparsity-promoting function traces back several decades. A leading early application was reflection seismology, in which a sparse reflection function (indicating meaningful changes between subsurface layers) was sought from bandlimited data. The basic idea is illustrated graphically in Fig. 8 with a simple example of a two-dimensional coefficient vector. In this case, $K = 1$. Green lines represent the solutions fulfilling the under-determined system model. Blue curves illustrate the points sharing the same L_0, L_2 or L_1 norms, lie on the axes, a circle and a diamond, respectively. Red points are the corresponding retrieved solutions. The red point in the left sketch indicates the exact solution found by L_0 norm minimization. This solution is the sparsest fulfilling the under-determined system model since it locates on the axis, and hence has only one non-zero coefficient, i.e. γ_2. The conventional L_2 norm always gives non-sparse solutions, i.e. both coordinates γ_1 and γ_2 have non-zero values. The L_1 norm retrieves very likely the same sparse solution as the L_0 norm.

In case there is no prior knowledge about K and in the presence of measurement noise, eq. (14) can be approximated by:

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} \left\{ \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_2^2 + \lambda_K \|\boldsymbol{\gamma}\|_1 \right\} \tag{15}$$

where λ_K is a Lagrange multiplier depending on the number of samples N and the noise level σ_ε. Eq. (15) consists of an L_2 norm residual and an L_1 norm regularizer. Eq. (15) can be interpreted as a Bayesian estimate with an exponential prior favoring sparse solutions (Seeger and Wipf, 2010).

For a given K, quadratic relaxation can provide a good approximation of the sparsest solution:

$$\hat{\boldsymbol{\gamma}} = \min_{\boldsymbol{\gamma}} \|\boldsymbol{\gamma}\|_1 \quad \text{s.t.} \quad \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_2 < \sigma_\varepsilon \tag{16}$$

Eq. (15) and eq. (16) are different forms of the Least Absolute Shrinkage and Selection Operator (LASSO) (Boyd and Vandenberghe, 2004) which was first proposed in statistics in order to solve the inverse problem by minimizing the residual sum of squares (i.e. L_2 norm) subject to the sum of the absolute values of the coefficients (i.e. L_1 norm) being smaller than a constant (Tibshirani, 1996). By appropriately choosing λ_k, eq. (15) is identical to eq. (16).

The statistical roles of different regularizers (e.g. the L_0 norm, the conventional L_2 norm and the L_1 norm) are understood by inspecting the priors (see Fig. 9) and the posteriors (see Fig. 10) they lead to. The priors introduced by the $L_{0.5}$, L_2 and L_0 norm are compared graphically in Fig. 9. It is obvious that the L_p norm with p closer to 0 approximates the prior introduced by the L_0 norm better. Here the two green lines indicate the 3 dB width of the likelihood. The corresponding posteriors (i.e. essentially the product of likelihood and prior) are visualized in Fig. 10. The green line here represents the noise free solutions fulfilling the under-determined system model as in Fig. 8 while the red point indicates the true sparsest solution of the system. By comparing the posterior distributions, it can been seen that:

- The $L_{0.5}$ norm finds the sparsest solutions located on the axes γ_1 or γ_2.
- The conventional L_2 norm prior (i.e. Tikhonov regularization) gives again with a high likelihood non-sparse solution because the maxima is not located on the axis.
- The L_1 norm finds very probably the sparsest solution though with a biased amplitude, i.e. the L_1 norm favors slightly smaller amplitudes.

In contrary to the NP-hard L_0 norm and the non-convex L_p norm ($0 < p < 1$), the L_1 norm minimization (as the limiting case of a convex problem) subject to linear equality constraints can easily be recast as a linear programming problem making available a host of ever more efficient solution algorithms. This fact makes CS a charming sampling paradigm.
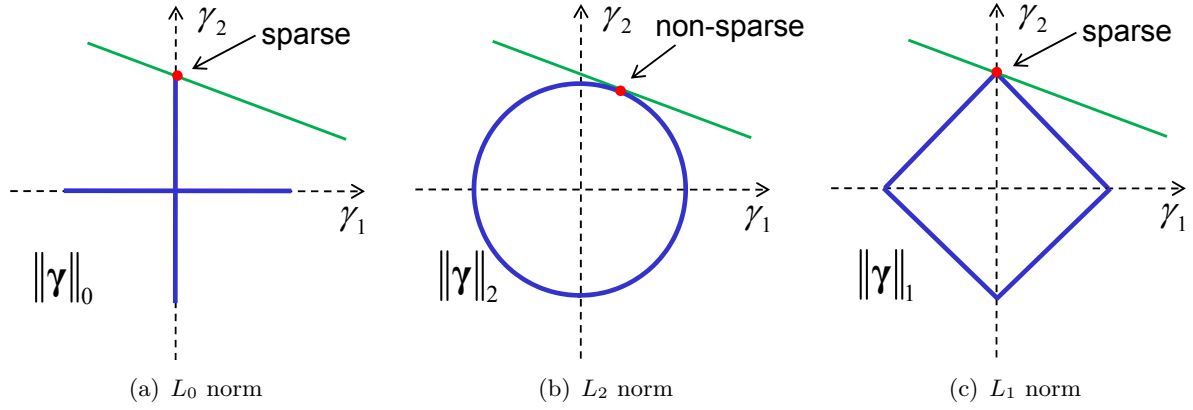
(a) L_0 norm   (b) L_2 norm   (c) L_1 norm

Figure 8. Graphical comparison between L_0 norm, L_2 norm and L_1 norm minimization for sparse reconstruction. Green lines represent the solutions of the under-determined system model; blue curves illustrate the points sharing the same L_0, L_2 and L_1 norms respectively; red points are the corresponding retrieved solutions. The NP-hard L_0 norm always gives the sparsest solution; The conventional L_2 norm gives a non-sparse solution; The L_1 norm gives very likely the same sparse solution as the L_0 norm.

(a) L_p norm   (b) L_2 norm   (c) L_1 norm
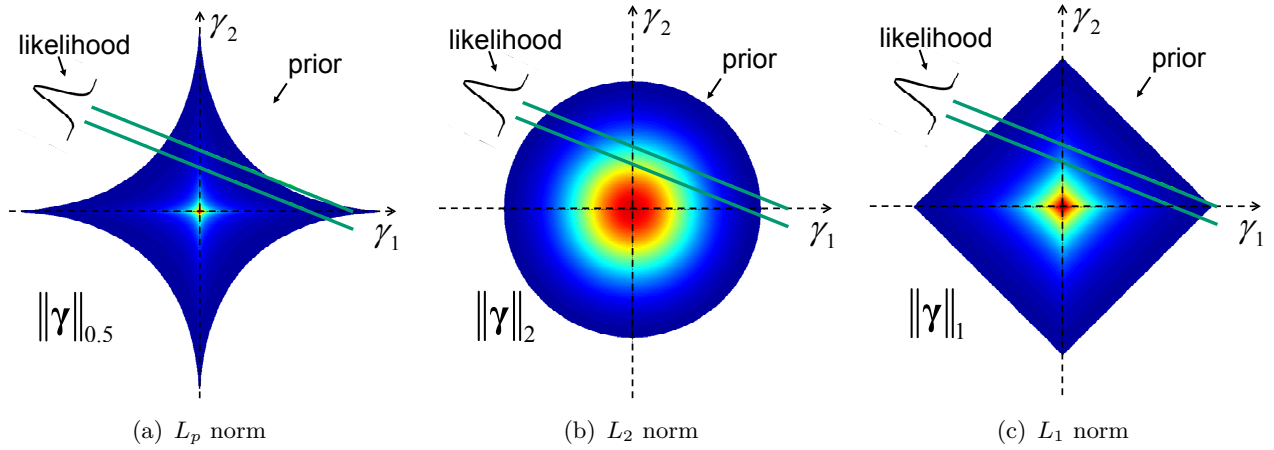
Figure 9. Graphical comparison between the L_p norm ($p = 0.5$), the L_2 norm and the L_1 norm minimization for sparse reconstruction in the presence of noise: likelihood and prior. The L_p norm minimization ($p \leq 1$) can be interpreted as MAP estimator with a prior favoring sparse solutions. Green lines: 3 dB width of the likelihood.

(a) L_p norm   (b) L_2 norm   (c) L_1 norm
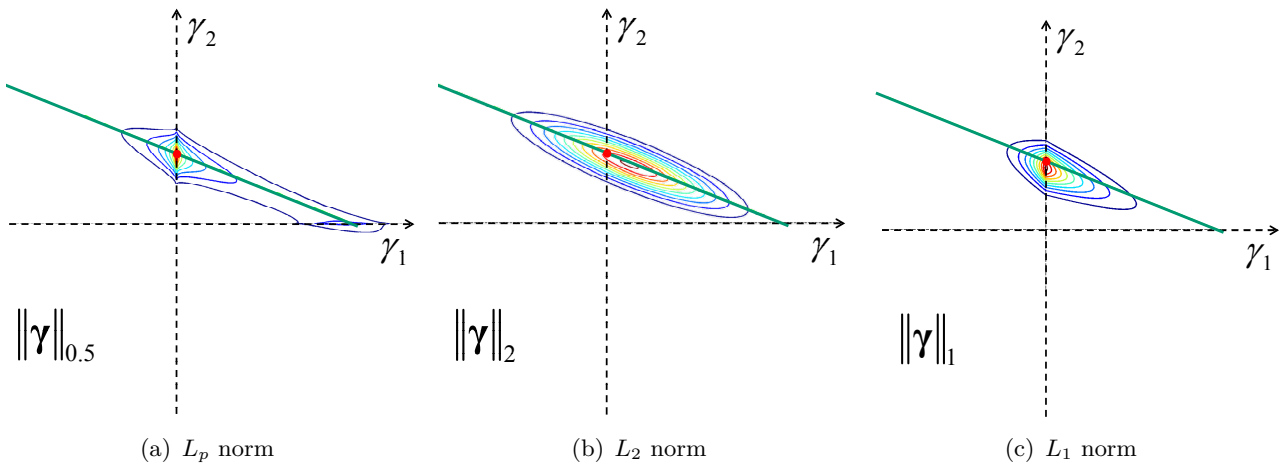
Figure 10. Graphical comparison between the L_p norm ($p = 0.5$), the L_2 norm and the L_1 norm minimization for sparse reconstruction: *posteriors*. Green lines represent the noise-free solutions of the under-determined system model; red points: the true sparsest solutions. The $L_{0.5}$ and the L_1 norms give very likely the sparsest solution with a biased amplitude (i.e. amplitude is often underestimated); The conventional L_2 norm gives a non-sparse solution.

To make this possible, beside the *sparsity* of the signal, CS relies on two principles: *incoherence*, which pertains to the sensing modality (Candès and Wakin, 2008), and *restricted isometry property (RIP)* which guarantees sufficiently sparse reconstruction in the presence of noise.

2.2.2   Incoherence and restricted isometry property (RIP)

The incoherence and RIP properties are the bargaining chip to find the unique sparse solution.

- **Incoherence**

  The coherence between $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ is defined by:

  $$\mu(\mathbf{R}) = \mu(\boldsymbol{\Phi}, \boldsymbol{\Psi}) = \sqrt{N} \max_{1 \leq k,j \leq N} \frac{|\langle \boldsymbol{\varphi_k}, \boldsymbol{\psi_j} \rangle|}{\|\boldsymbol{\varphi_k}\|_2 \|\boldsymbol{\psi_j}\|_2} \tag{17}$$

  where $\boldsymbol{\varphi_k}$ is the kth row of the sensing matrix $\boldsymbol{\Phi}$ while $\boldsymbol{\psi_j}$ is the jth row of the orthogonal basis $\boldsymbol{\Psi}$. μ has a value between 1 and $\sqrt{N}$. The incoherence property requires that rows of $\boldsymbol{\Phi}$ (sensing matrix) cannot sparsely represent the columns of $\boldsymbol{\Psi}^{\mathrm{H}}$ (orthogonal basis) and vice versa. A small μ means that the sparse mapping operator $\mathbf{R}$ will spread out information of sparse (highly localized) coefficients over the entire measurement space (like a Fourier transform does) and thus make them insensitive to random undersampling, otherwise the reconstruction of non-zero coefficients will be biased towards certain positions. For instance, let us randomly choose N rows from $\boldsymbol{\Psi}$ as our sensing matrix $\boldsymbol{\Phi}$. Then what we are actually sensing with $\boldsymbol{\Phi}$ is only the N elements of $\boldsymbol{\gamma}$ located at the positions of the N randomly chosen columns. As a consequence of this counter-example, the sparse reconstruction would only be possible if the K non-zero coefficients were located by chance within the N chosen positions.

  CS is mainly concerned with low coherence pairs, and the well acknowledged $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ pairs are (Candès and Wakin, 2008):

  - ⬦ *Fourier matrix & identity matrix*
    Corresponds to the classical sampling scheme in time or space. The time-frequency pair obeys $\mu = 1$ and, therefore, has maximal incoherence. Further, spikes and sinusoids are maximally incoherent not just in one dimension but in any number of dimensions. If $\boldsymbol{\Psi}$ is the identity matrix, coefficients are identical to the signal. This allows us to employ the numerous powerful sparse reconstruction techniques for spectral estimation when the signal itself is sparse, i.e. sparse in an identity matrix basis in the language of CS. Typical examples are ISAR (Zhang et al., 2009) and TomoSAR (Budillon et al., 2009) (Zhu and Bamler, 2010c).

  - ⬦ *Noiselets & wavelets basis*
    The coherence between noiselets and Haar wavelets is $\sqrt{2}$ and that between noiselets and Daubechies D4 and D8 wavelets is, respectively, about 2.2 and 2.9 across a wide range of sample sizes N (Candès and Wakin, 2008). Since most image data are sparse or compressible in a wavelet basis, this reveals great potential for using CS for image processing such as image sharpening (Yang et al., 2010) (Yang et al., 2008), inpainting (Elad et al., 2005) (Fadili et al., 2009) (Shen et al., 2009), denoising (Elad and Aharon, 2006) (Malgouyres and Zeng, 2009), medical Imaging (Lustig et al., 2007) (Lustig et al., 2008) (Yu and Wang, 2009), Hyperspectral Imaging (Willett et al., 2007) and radar imaging (Ender, 2010) (Baraniuk and Steeghs, 2007). In addition, there are tremendously well-developed noiselet transform algorithms that run in $O(N)$ time. This fact is practically important since it allows numerically efficient CS implementations.

  - ⬦ *Random matrices & any fixed basis*
    Random matrices are matrices consisting of random vectors that have a flat power
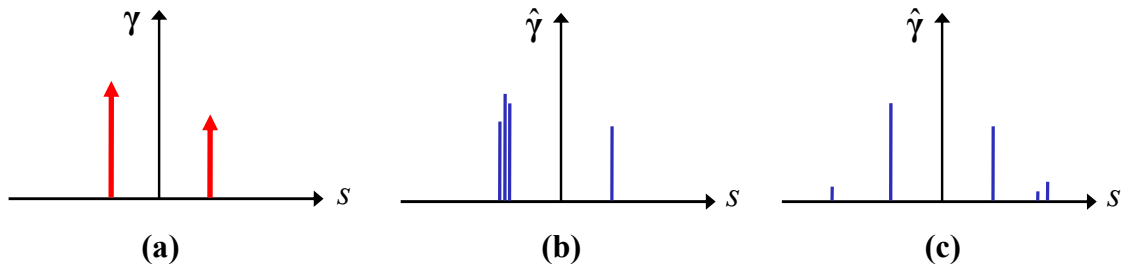
Figure 11. The violation of RIP and incoherence properties may cause artifacts in the reconstruction. (a): true reflectivity profile; (b) reconstruction of a scatterer by several adjacent lines instead of a single peak; (c) reconstruction with noise-like outliers.

spectral density. They are largely incoherent with any fixed basis with a coherence of about $\sqrt{2 \log(N)}$. The most popular examples of such kind of matrices are random waveforms with independent identically distributed (i.i.d.) Gaussian entries or ± 1 binary entries. Surprisingly, the efficient mechanism ought to acquire signals with random vectors, e.g. white noise. This feature inspired scientists to conceive novel measuring systems (see the one pixel camera in (Duarte et al., 2008)).

- **Restricted isometry property (RIP)**

  RIP guarantees the robust sparse reconstruction in the presence of noise. The RIP requires:

  $$(1 - \delta_s) \, \|\boldsymbol{\nu}\|_2^2 \ \leq \ \|\mathbf{R} \, \boldsymbol{\nu}\|_2^2 \ \leq \ (1 + \delta_s) \, \|\boldsymbol{\nu}\|_2^2 \tag{18}$$

  where $\boldsymbol{\nu}$ is any vector having K non-zero coefficients at the same positions as $\boldsymbol{\gamma}$ and δ_s is a small positive number. The smaller δ_s is the better the sparse signal can be reconstructed in the presence of noise. Eq. (18) essentially says that all sub-matrices of $\mathbf{R}$ composed of K columns should be mutually nearly orthogonal and hence, preserve the energy of all vectors having only non-zero elements at the same K positions as $\boldsymbol{\gamma}$. There are a couple of well-known stable measurement matrices (in the sense of RIP), e.g. random Gaussian, random binary and randomly selected Fourier samples (Candès and Wakin, 2008).

As mentioned before, CS can be seen as the non-parametric realization of a parametric estimator. To make this possible, first of all, the signal must be sparse, i.e. can be described by a model with a small number of degrees of freedom. Let us assume the sparsity K and the positions of non-zero elements are given. Then with $N \gg K$ measurements, the system is over-determined and it is well-known that an optimum solution can be obtained if the sparse mapping operator $\mathbf{R}$ has full rank. Now the situation is that the positions of non-zero elements, i.e. the exact model, are unknown. This would require any randomly chosen K rows of $\mathbf{R}$ to have full rank. This is essentially the concept of RIP and incoherence.

Any violation of RIP and incoherence will introduce artifacts in the sparse reconstruction, i.e. the solution $\hat{\boldsymbol{\gamma}}$ is no longer the sparsest one $\boldsymbol{\gamma}$. There are two classes of artifacts shown schematically in Fig. 11 (Zhu and Bamler, 2011b):

1) Single significant elements may be broadened to a few adjacent lines (Fig. 11, middle) and with a certain probability small spurious impulses may show up (Fig. 11, right).
2) Eq. (15) will slightly, yet systematically, underestimate the amplitudes of the coefficients because the L_1 norm minimization favors smaller amplitudes. This artifact can be clearly observed in Fig. 12 which shows a zoom into the region around the maximum of Fig. 10.c. The posterior shows that biased amplitude estimates (i.e. under-estimated amplitudes, see the position of the extreme) are preferred.

A new algorithm will be introduced in Chapter 4 to improve the CS estimator and correct for these two deficiencies.
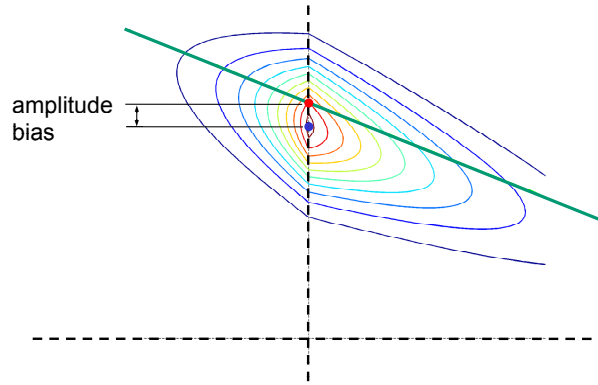
Figure 12. A zoom into the posterior shown in Fig. 10.c. the sparsest solutions with biased amplitude estimates (i.e. under-estimated amplitudes, see the position of the extreme) are preferred.

2.2.3   Sparse reconstruction resources

Various sparse reconstruction toolboxes are available for solving the L_1 norm minimization problem. According to the experiences of the author, an incomplete list of the toolboxes which provide solvers for eq. (15) and/or eq (16) is:

- *SparseLab*
  Provided by: D. Donoho, V. Stodden and Y. Tsaig (Donoho et al., 2005) (Donoho et al., 2009)
  Implemented algorithms: Stagewise orthogonal matching pursuit algorithm for solving eq (16) and least angle regression algorithm for solving eq. (15)
  Download link: http://sparselab.stanford.edu/

- *l1_ls*
  Provided by: K. Koh, S. Kim and S. Boyd (Koh et al., 2007)
  Implemented algorithms: interior-point method algorithm for solving eq. (15)
  Download link: http://www.stanford.edu/ boyd/l1_ls/

- *YALL1*
  Provided by : Y. Zhang, J. Yang and W. Yin (Yang and Zhang, 2009)
  Implemented algorithms: the truncated Newton interior-point method for solving eq. (15)
  Download link: http://yall1.blogs.rice.edu/

- *cvx*
  Provided by M. Grant and S. Boyd (Grant and Boyd, 2011) (Grant and Boyd, 2008).
  Implemented algorithms: SDPT convex optimization method for solving eq. (15)
  Download link: http://cvxr.com/cvx/

- *FPC_AS (fixed-point continuation and active set)*
  Provided by Z. Wen and W. Yin (Hale et al., 2007) (Wen et al., 2010).
  Implemented algorithms: Algorithm based based upon the active-set algorithm with a continuation strategy for solving eq. (15)
  Download link: http://www.caam.rice.edu/ optimization/L1/FPC_AS/

- *l1-MAGIC*
  Provided by: E. Candes and J. Romberg (Candès and Tao, 2005) (Candès et al., 2006)
  Implemented algorithms: Algorithms are based on standard interior-point methods for

solving eq. (16)
Download link: http://www.acm.caltech.edu/l1magic/

The above mentioned toolboxes are all written in Matlab©. Based on the experimental comparison with application to pan-sharpening, the toolboxes provide comparable estimation accuracy (Wang, 2011). cvx and SparseLab are recommended by the author. cvx is implemented for complex numbers and has been used for tomographic reconstruction in this thesis. SparseLab provides so far the best computational efficiency. Note that the above mentioned evaluation of the toolboxes is based on the versions published in 2010.

3 State of the art in tomographic SAR

The idea of tomographic imaging in radar was proposed in (Chan and Farhat, 1981) (Farhat et al., 1984). This concept was introduced to SAR in (Piau, 1994) (Jakowatz and Thompson, 1995) (Homer et al., 1996) as a way of overcoming limitations of standard 2-D imaging by achieving, similar to CAT, focused 3-D images. The first experiments were carried out in a laboratory with ideal experimental conditions (Pasquali et al., 1995), by using spaceborne systems (Gatelli et al., 1994) (She et al., 1999) and airborne systems (Reigber and Moreira, 2000). Since then, this technique has developed considerably. Research is mainly carried out 1) to exploit the application field of TomoSAR and 2) to explore methodologies for tomographic reconstruction. In this Chapter, the state of the art in tomographic SAR will be discussed by following the aforementioned two tracks. In addition, current research on CS applied to radar is discussed.

3.1 Applications of TomoSAR

TomoSAR intends to resolve the reflected power distribution along the elevation direction. The potential future applications of TomoSAR span geology, glaciology, biomass estimation, forestry, and detection of buried structures for archeology and civil applications. According to the scattering mechanism, the coherent targets, i.e. the signal, to be resolved can be categorized as *discrete scatterers* and *volumetric scatterers*. The reflected power of discrete scatterers can be characterized by several δ-functions, i.e. the signal can be described by a deterministic model with a few parameters. Volumetric scatterers have a continuous backscatter profile associated with completely random scattering phases, i.e. the signal can only be described by stochastic models. Furthermore, for space-borne SAR systems, the data stacks are built up by collecting multi-pass SAR images, and hence any LOS motion of the scatterers has influence on the phase and is coupled with the phase term caused by the elevation of the point. Therefore, another important issue is to resolve *discrete scatterers with motion*.

3.1.1 TomoSAR for resolving discrete scatterers

The major interest of resolving discrete scatterers rises from the need of layover scatterer separation in urban environment. For this purpose, TomoSAR has been applied to C-band ERS data of extended scenes over the Bay of Naples in (Fornaro et al., 2005), to TerraSAR-X data over downtown Las Vegas in (Zhu et al., 2008) and to L-band airborne SAR data acquired by the DLR's experimental SAR (E-SAR) system over the city of Dresden in Germany in (Guillaso and Reigber, 2005) (Sauer et al., 2009). In (Fornaro and Serafino, 2006) single and double scatterer cases were separated at the same test-site as in (Fornaro et al., 2005). Another attempt to separate layover scatterers can be found in (Ferretti et al., 2005) which extended the concept of permanent scatterers to two scatterers inside a pixel. Due to the multi-pass nature of the data acquisition, the application of pure 3-D tomographic SAR imaging (i.e. without taking the motion into account) at spaceborne systems has been limited, however, the first attempts of spaceborne TomoSAR could pave the way for resolving discrete scatterers with motion (see section 3.1.4).

3.1.2 TomoSAR for imaging volumetric scatterers

Another important field of TomoSAR is to image volumetric scatterers. A typical target application is sensing forest layer structure, height and biomass. Other potential applications are

sensing ground topography, soil humidity and ice thickness. TomoSAR for imaging of the vertical structure of forested terrain was investigation in (Reigber et al., 2000) (Lombardini and Reigber, 2003). Imaging of targets hidden beneath the vegetation was demonstrated in (Nannini et al., 2008). In (Frey et al., 2008), P-band tomographic SAR images of a larger forested area (400 m × 1000 m) were presented for the first time. The sector interpolation method was proposed for TomoSAR in (Lombardini et al., 2007) by exploiting a priori information about the height sector containing the scatterers. An investigation into the minimal required number of acquisitions required by using subspace methods was carried out in (Nannini et al., 2009). Based on the Van Cittert-Zernike theorem (Van Cittert, 1934), a simple way to parameterize imaging of the volumetric scatterers, the so-called coherence tomography, was proposed in (Cloude, 2006) (Cloude, 2007). The basic idea is to localize volumetric scatterers based on the estimated coherence.

3.1.3   TomoSAR for scatterer discrimination

Polarimetric SAR interferometry was proposed for scatterer discrimination in (Cloude and Papathanassiou, 1998) where a full polarimetric signal is decomposed into different orthogonal scattering mechanisms within the SAR resolution pixel (e.g. double bounce at ground-trunk interface and volume scattering in the canopy). Although interesting results have been achieved in forested areas, elevation imaging is not possible in the presence of either more scatterers with the same scattering mechanism or with more general distributed scattering mechanisms. In (Reigber and Moreira, 2000) (Reigber et al., 2000), airborne polarimetric TomoSAR was demonstrated using fully polarimetric L-band SAR data on the test-site of Oberpfaffenhofen acquired by E-SAR where the Pauli-decomposition result was presented showing fully polarimetric information from different scattering objects such as spruce forest, road, buildings, cars and corner reflector. In (Cloude, 2006), polarization coherence tomography (PCT) was introduced, which employs variation of the interferometric coherence with polarization to reconstruct a vertical profile function (e.g. modeled as superposition of Legendre polynomials) in penetrable volume scattering. More investigations of polarimetric TomoSAR can be found in (Guillaso et al., 2003) (Guillaso and Reigber, 2005) (Sauer et al., 2009) (Tebaldini, 2010).

3.1.4   TomoSAR for resolving scatterers with motion

Resolving discrete scatterers with possible motion is also referred to as D-TomoSAR or 4-D SAR focusing. The goal is not only to separate targets interfering in the same azimuth-range pixel, but also to estimate their possible relative motion. It exploits the strength of both TomoSAR (3-D SAR imaging) and PSI (long term motion monitoring). This concept was proposed in (Lombardini, 2005) and applied to ERS real data in (Fornaro et al., 2009). The extraction of time series of displacement for single and double interfering scatterers based on the so-called velocity spectrum was presented in (Fornaro et al., 2009). In (Fornaro et al., 2010) the application of 4-D imaging to a real data case using ERS data over the city of Rome was presented. In (De Maio et al., 2009a) and (De Maio et al., 2009b) the problem of the detection of single and double scatterers was addressed by presenting detectors based on the generalized likelihood ratio test (GLRT). However, due to the lack of VHR SAR data, only PSI was able to give some reasonable reconstruction of the structures and motion of buildings by selecting the pixels containing only single scatterers. So far, tomographic reconstruction of the building complex hasn't been presented in the literature.

3.2   TomoSAR algorithms

As depicted in section 2, since TomoSAR uses only a small angular diversity, instead of back projection spectral estimation is sufficient in most cases. It makes numerous standard spectral estimation resources applicable for tomographic SAR reconstruction. This section mainly presents the spectral estimation methods with application to TomoSAR. In addition, back projection methods are also addressed.

3.2.1   Spectral estimation methods

This section mainly discusses the classification of spectral estimators and gives an overview on the most relevant methods with application to TomoSAR.

The standard spectral estimation methods can be classified as *non-parametric* methods (most of the classic methods) and *parametric* methods (most of the modern methods): Non-parametric methods make no assumption about the statistical properties of the received signals and simply pass the data through a set of band-pass filters and measure the filter output powers; Parametric methods model the data as a sum of a few possibly damped sinusoids and estimate their parameters (Stoica and Moses, 2005) (Li and Stoica, 2008). In general, parametric methods may offer better estimates if the data closely agrees with the assumed model, otherwise nonparametric methods may be better.

With application to TomoSAR, the techniques can be further categorized as *single-looking* and *multi-looking* methods. Single-looking methods are based on the stacked measurements of single azimuth-range pixels and do not explore the correlation between the surrounding pixels. Multi-looking methods require an estimation of the data covariance matrix $\mathbf{C_{gg}}$ which they take from an ensemble average of the measurements at the surrounding pixels. Generally, multi-looking methods obtain higher SNR and SR in elevation by sacrificing resolution in azimuth and range. Note that all single-looking methods can be extended to multi-looks, e.g. by simply averaging estimates of N_L looks.

The most relevant spectral estimation algorithms, including the conventional beamforming (BF), singular value decomposition (SVD), adaptive beamforming, multiple signal classification (MUSIC), nonlinear least squares (NLS) and M-RELAX, are introduced in the following. The description of the methods are based on single polarization TomoSAR. Their corresponding extension to polarimetric TomoSAR was detailed in (Sauer, 2008).

- **Conventional beamforming (BF)**

  Beamforming (Stoica and Moses, 1997) was the first algorithm to be used for solving the layover problem (Pasquali et al., 1995). TomoSAR can be cast as an array signal processing problem. In eq. (5), columns of $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_L]$ can be treated as steering vectors along the elevation dimension and used for beamforming. To this end, the reflectivity along elevation can be obtained by:

  $$\hat{\boldsymbol{\gamma}} = \mathbf{R}^{\mathrm{H}} \mathbf{g} \tag{19}$$

  It is simply an inverse Fourier transform applied to the irregularly sampled signals and its magnitude is also named as periodogram. The name "beamforming" stands for the fact that the reflected power at elevation position $s_l(l = 1, ..., L)$ can be estimated as:

  $$|\hat{\gamma}_l|^2 = |\mathbf{r}_L^{\mathrm{H}} \mathbf{g}|^2 \tag{20}$$

  Frequency (elevation) estimates correspond to the locations of the maxima of the $|\hat{\boldsymbol{\gamma}}|^2$. As a single-looking method, conventional beamforming is computational very efficient. However, it has no SR capability. In addition, it has a severe and anomalous sidelobe

problem caused by the irregular sampling. Further discussions of beamforming applied to TomoSAR can be found in (She et al., 1999) (Lombardini et al., 2003) (Fornaro et al., 2009).

- **Singular value decomposition (SVD)**

  The SVD inversion framework has been described in (Fornaro et al., 2003). The discrete reflectivity signal $\boldsymbol{\gamma}$ can be reconstructed from $\mathbf{g}$ through pseudo inversion of $\mathbf{R}$, see eq. (5). However, the solution may include significant noise propagation due to the ill-conditioned nature of the problem. In particular, non-uniform track distribution exacerbates the ill-conditioness. The SVD is a simple and valuable tool for analyzing image quality and the amount of independent information about the unknowns which can be reliably retrieved from observations in the presence of noise (Hansen, 1994). Consider now an estimate of $\boldsymbol{\gamma}$ is obtained via the pseudo-inverse $\mathbf{R}^\dagger$. Using the SVD, we get:

  $$\hat{\gamma} = \mathbf{R}^\dagger \mathbf{g} = \sum_{n=1}^{N} \sigma_n^{-1} \left(\mathbf{u}_n^{\mathrm{H}} \mathbf{g} \right) \mathbf{v}_n \tag{21}$$

  where the σ_n are the non-negative singular values (SVs) of $\mathbf{R}$ while the vectors $\mathbf{u}_n$ and $\mathbf{v}_n$ are the left and right singular vectors of $\mathbf{R}$ respectively. Due to the reciprocal of σ_n , noise propagation caused by small SVs will compromise this solution and regularization tools are required.

  A well-known method to deal with an ill-conditioned matrix in eq. (21) is the truncated singular value decomposition (TSVD) (Hansen, 1987). The basic idea of TSVD as well as other regularization methods is to impose additional requirements on the solution, thus hopefully damping the contributions from the errors of the right hand side of eq. (21). In the case of TSVD this is achieved by discarding the components of the solution corresponding to the smallest $N - Q$ SVs. Here, Q is called "numerical rank" or "effective rank" of $\mathbf{R}$ defined by the number of SVs greater than some noise level σ_ε. The TSVD was implemented in (Fornaro et al., 2003) for a TomoSAR experiment with 44 ERS acquisitions and robust performance was achieved. However, from the perturbation theory for the TSVD (Hansen, 1987), it is a stable method only for problems with a well-determined numerical rank.

  For Gaussian stationary white measurement noise, i.e. $\mathbf{C}_{\varepsilon\varepsilon} = \sigma_\varepsilon^2 \mathbf{I}$, and a white prior, i.e. $\mathbf{C}_{\gamma\gamma} = \mathbf{I}$, transforming the maximum a posteriori (MAP) estimator of eq. (5) to the SV space results in a soft thresholding, e.g. weighting the SVs according to their magnitudes, also referred to as a Tikhonov regularization:

  $$\hat{\gamma} = \left(\mathbf{R}^{\mathrm{H}} \mathbf{C}_{\varepsilon\varepsilon}^{-1} \mathbf{R} + \mathbf{C}_{\gamma\gamma}^{-1} \right)^{-1} \mathbf{R}^{\mathrm{H}} \mathbf{C}_{\varepsilon\varepsilon}^{-1} \mathbf{g} = \sum_{i=0}^{\mathrm{N}} \frac{\sigma_i}{\sigma_i^2 + \sigma_\varepsilon^2} \left(\mathbf{u}_i^{\mathrm{H}} \mathbf{g} \right) \mathbf{v}_i \tag{22}$$

  This type of weighting resembles the Wiener filter under white noise, and hence the method is named SVD-Wiener in (Zhu, 2008). An experiment with 8 TerraSAR-X acquisitions was carried out in (Zhu, 2008) where SVD-Wiener was demonstrated to provide more stable performance especially for the case of few acquisitions.

  In general, as a single-looking method, SVD based regularized inversion is computationally efficient and is not sensitive to irregular sampling. Compared to beamforming, it shows better sidelobe suppression and can provide slight SR.

- **Adaptive beamforming (Capon)**

  Adaptive beamforming, also named Capon, was first proposed in (Capon, 1969) for array processing of geo-seismic signals. Compared to conventional beamforming, Capon intro-

duces self interference cancellation by weighting the steering vectors according to the co-variance matrix $\mathbf{C_{gg}}$. In practice, $\mathbf{C_{gg}}$ is estimated from multi-looking measurements with N_L looks. Let us assume the measurement vector for the $n_l{}^{th}$ look is $\mathbf{g}_{n_l}, n_l = 1, \cdots N_L$. $\mathbf{C_{gg}}$ is estimated by:

$$\hat{\mathbf{C}}_{\mathbf{gg}} = \frac{1}{N_L} \sum_{n_l=1}^{N_L} (\mathbf{g}_{n_l} \mathbf{g}_{n_l}^{\mathrm{H}}) \tag{23}$$

Then, the reflected power at elevation position s_l ($l = 1, ..., L$) can be estimated as:

$$|\hat{\gamma}_l|^2 = \frac{1}{\mathbf{r}_l^{\mathrm{H}} \hat{\mathbf{C}}_{\mathbf{gg}}^{-1} \mathbf{r}_l} \tag{24}$$

Capon was first applied to TomoSAR in (Lombardini et al., 2001). Thanks to self-cancellation, Capon has better performance regarding elevation resolution and sidelobe suppression compared to beamforming. However, the self-cancellation may also be critical when the steering vectors mismatch to the position range of the scatterers, phase noise is too severe (this is particularly critical with application to TomoSAR since the unmodelled motion often leads to severe phase errors) or the number of looks is too small (like all other multi-looking methods) (Stoica and Moses, 1997). In addition, Capon is not radiometrically correct.

- **Multiple Signal Classification (MUSIC)**

  MUSIC is a model-based spectral estimator introduced in (Bienvenu, 1979) (Schmidt, 1979) for signals buried in white noise. It is a subspace-based technique which decomposes the covariance matrix $\mathbf{C_{gg}}$ into its eigenvalues and eigenvectors and analyzes their proper-ties (Stoica and Arye, 1989) (Stoica and Cedervall, 2002). As a first step, $\mathbf{C_{gg}}$ is estimated by (23) and the eigenvalue decomposition of $\hat{\mathbf{C}}_{\mathbf{gg}}$ is performed. The number of scatterers K is either given as priori or can be roughly estimated from the number of significant eigenvalues. The noise subspace $\hat{\mathbf{G}}$, i.e. orthogonal to the signal space, can be determined by the $N - K$ eigenvectors corresponding to the $N - K$ smallest eigenvalues. Therefore, the "pseudospectrum" at elevation position s_l ($l = 1, ..., L$) can be estimated as:

  $$|\hat{\gamma}_l|^2_{MU} = \frac{1}{\mathbf{r}_l^{\mathrm{H}} \hat{\mathbf{G}} \hat{\mathbf{G}}^{\mathrm{H}} \mathbf{r}_l} \tag{25}$$

  The K elevations are determined as the locations of the K highest peaks of the MUSIC spectrum $|\hat{\gamma}_l|^2_{MU}$. Note the MUSIC spectrum is not the real reflected power, hence, MUSIC is not radiometrically correct.

  MUSIC was first introduced for TomoSAR in (Gini and Lombardini, 2002). It has been further addressed in (Lombardini et al., 2001) (Gini and Lombardini, 2005) (Guillaso and Reigber, 2005) (Nannini et al., 2008) (Sauer et al., 2009). In general, MUSIC provides better resolution and lower sidelobes than beamforming and Capon. It was pointed out in (Stoica and Arye, 1989) that MUSIC possesses good performance for uncorrelated signals, i.e. the rank of the signal subspace is equal to the number of scatterers K, whereas the estimation precision deteriorates significantly when the signals are highly correlated. For correlated signals, a modified version of MUSIC can be found in (Stoica and Moses, 1990).

- **Nonlinear least squares (NLS)**

  Assuming the presence of K scatterers inside a pixel with elevations of $\mathbf{s} = [s_1, ..., s_K]$ and

complex-valued reflectivity $\hat{\gamma}(\mathbf{s})$, the under-determined system model (5) reduces to the following over-determined problem:

$$\mathbf{g} = \mathbf{R}(\mathbf{s})\,\boldsymbol{\gamma}(\mathbf{s}) + \varepsilon \tag{26}$$

where the $N \times K$ matrix $\mathbf{R}(\mathbf{s})$ depends on the unknown elevations of the scatterers $\mathbf{s}$ with the element $R_{nk} = \exp\left(-j2\pi\xi_n s_k\right)$. Eq. (26) is linear in amplitude and nonlinear in elevation. The least squares error can be minimized with respect to a given $\boldsymbol{\gamma}(\mathbf{s})$ in a closed analytic form. Thus, the main issue is to determine the unknown elevations where a K dimensional search is needed (for differential TomoSAR, an additional search for velocity is required). The optimum estimate $\hat{\boldsymbol{\gamma}}(\hat{\mathbf{s}})$ is found to be (Stoica and Moses, 2005):

$$\hat{\boldsymbol{\gamma}}(\hat{\mathbf{s}}) = \arg\min_{\boldsymbol{\gamma}(\mathbf{s})} \left\{ \|\mathbf{g} - \mathbf{R}(\mathbf{s})\boldsymbol{\gamma}(\mathbf{s})\|_2^2 \right\}. \tag{27}$$

For Gaussian white noise, NLS is identical to the maximum likelihood estimator (MLE). It is therefore theoretically the best estimator for our application if and only if the data closely agree with the assumed model. Therefore, in order to avoid model error, the NLS estimator has been associated to a model selection step in (Zhu, 2008) for determining the most likely true model. NLS was first applied for layover separation in (Ferretti et al., 2005) and the result of NLS (together with model selection) was considered as a performance reference for other estimators in (Zhu, 2008). Due to the large computational effort to the multi-dimensional search, the NP-hard NLS is not recommended for practical data processing.

- **M-RELAX**

  M-RELAX was proposed in (Li and Stoica, 1996) as a more computationally efficient approximation of NLS. The basic idea is, instead of a K-dimensional search, a one-dimensional search is performed K times. After each iteration step, the dominating component is subtracted from the measurements and new measurements are obtained. It can provide generally higher estimation precision than MUSIC, especially in low SNR. However, it will break down if there are two closely spacing scatterers sharing comparable amplitudes. In this case, a one-dimensional search may lead to a wrongly located maxima. M-RELAX for TomoSAR are addressed in (Gini and Lombardini, 2002) (Lombardini et al., 2003).

Method	Computati--onal cost	Parame--tric	Multi--looking	Accuracy for multi-scatterers	Elevation resolution	Risk of incor--rect elevation estimation
BF	low	no	no	low	low	medium
SVD	low-medium	no	no	medium	low-medium	low
Capon	medium-high	no	yes	high	medium	medium
MUSIC	high	yes	yes	high	high	medium
NLS+MS	Very high	yes	no	very high	very high	low
M-RELAX	medium	yes	no	high	high	high

Table 1. Summary of spectral estimation methods with application to TomoSAR

The characteristics of the above mentioned spectral estimation methods with application to

TomoSAR are summarized in Table 1 regarding the computational cost, parametric or non-parametric, single-look or multi-looking, estimation accuracy in case of multiple scatterers (in the presence of only single scatterers, all above mentioned methods can give an accurate parameter estimation), elevation resolution and the risk of incorrect elevation estimates.

In general, depending on the application, different methods are recommended, e.g. for resolving discrete scatterers with or without motion in urban environment, it is very important to maintain the azimuth-range resolution to observe the inherent scale of urban infrastructure, and hence single-looking methods are recommended. It is recommended to use SVD or M-RELAX for medium resolution applications. For high resolution applications, NLS can provide excellent performance, however, the large computational cost renders it impractical. This confirms that new TomoSAR algorithms are required for high resolution applications in urban environments and this is one of the objectives of this thesis.

3.2.2   Back projection methods

As mentioned, spectral estimation is only applicable for large range distances and small angular diversities. If this is not the case back projection should be applied. Frequency domain back projection was applied to TomoSAR in (Reigber and Moreira, 2000) (Nannini and Scheiber, 2007). A time domain back projection (TDBP) was proposed in (Frey et al., 2008) which can cope with irregular sampling without introducing any approximation with respect to the geometry.

3.3   State of the art on compressive sensing applied to radar

The potential of compressive sensing in radar lies in applications where signals are compressible or sparse. Since CS is a relatively new technique, there is little literature that addresses compressive sensing applied to radar. (Baraniuk and Steeghs, 2007), one of the first publications on CS-radar, demonstrated that one can eliminate the matched filter in the radar receiver and lower the receiver A/D converter bandwidth using CS principles. (Bhattacharya et al., 2007) proposed the use of the CS framework for fast compression of SAR raw data to ease the computational requirements of satellite onboard processing. In (Herman and Strohmer, 2009), a "stylized" compressed sensing radar, which transmits a sufficiently incoherent pulse and employs the techniques of compressive sensing to reconstruct the target scene was proposed. With application to wide-angle SAR, (Varshney et al., 2008) addressed a graph-structured sparse signal representation algorithm for over-complete dictionaries that can be decomposed into sub-dictionaries with dictionary elements arranged in a hierarchy. (Berger et al., 2008a), (Xie and Zhang, 2010), (Berger et al., 2008b) and (Shastry et al., 2010) discussed the possible use of CS in passive radar, ground-based radar and MIMO radar, and ultra-wide band stochastic waveform radar respectively. (Ender, 2010) is an interesting review article about CS applied to radar, including potential applications of pulse compression, radar imaging, and airspace surveillance with array antennas. (Smith et al., 2010) presented a study of how the analogue-to-digital converter sampling rate in a digital radar can be reduced—without a reduction in waveform bandwidth—through the use of CS. In (Zhang et al., 2009), CS was applied to inverse SAR and CS was demonstrated to outperform the conventional range-Doppler approach in resolution. A first attempt to use CS for TomoSAR was presented in (Budillon et al., 2009) using simulated data and followed by a practical experiment using ERS data in (Budillon et al., 2011). However, since medium resolution SAR data does not possess a sparse representation in elevation, the achievable SR is strongly limited.

4   Summary of the work

The four objectives, including one general and three methodology objectives, mentioned in Chapter 1 are addressed in six peer-reviewed journal articles by the author (four as the first author, two as co-author). This chapter gives a brief summary of the articles (see the Appendices) including the following four main contributions:

- A first demonstration of VHR TomoSAR and D-TomoSAR with 3-D and 4-D reconstructions of an entire city area using TerraSAR-X spotlight data is presented in section 4.1 (see also A.1 – A.3);
- Considering the sparsity of the signal in the elevation direction, a CS based algorithm (SL1MMER) is proposed in section 4.2 for sparse tomographic SAR reconstruction (see also A.4);
- A detailed systematic performance assessment is presented in section 4.3 regarding to elevation estimation accuracy, SR power and the required minimum number of acquisitions for robust reconstruction (see also A.5);
- To cope with the nonlinear motion problem, the time warp method is proposed in section 4.4 for multi-component nonlinear motion estimation and a real data example is provided where coupled linear and seasonal motion are separated (see also A.6).

Beyond what is presented in these journal papers, new tomographic SAR inversion processing results of the entire high rise city area will be presented in this Chapter.

4.1   VHR TomoSAR and D-TomoSAR

In A.1 – A.3, the tomographic potential of VHR SARs and the achievable quality on the basis of TerraSAR-X spotlight data of urban environments are demonstrated by presenting the first 3-D and 4-D tomographic reconstructions of an entire building complex (incl. its radar reflectivity) with a very high level of detail using a modified version of the conventional SVD method, called SVD-Wiener (equivalent to a MAP linear estimator). Beyond the first demonstration presented in A.2, a reconstruction of the entire city area shown in Fig. 1 will be presented in this section.

The stack used in this thesis consists of 60 TerraSAR-X spotlight images with a slant-range resolution of 0.6 m and an azimuth resolution of 1.1 m covering a time period of more than 2 years, from February 2008 to June 2010, over downtown Las Vegas. The pre-processing including atmospheric phase screen (APS) correction is performed by German Aerospace Center (DLR) PSI-GENESIS system on a PS network of high-SNR pixels containing only single scatterers (Adam et al., 2008). Fig. 13 illustrates the elevation aperture positions and temporal baselines relative to the master image acquired on April 17, 2009. During different time periods this test site undergoes different deformations which will be detailed in section 4.4. Therefore, the whole stack is divided into an "asterisk" sub-stack and a "diamond" sub-stack as shown in Fig. 13. In this experiment, only the "asterisk" sub-stack consisting of 30 images is used with an elevation aperture size of about 269.5 m. For non-parametric spectral analysis, this results in an elevation resolution of $\rho_s = 40.5$ m, i.e. approximately 20 m resolution in height with the elevation-to-height factor $\sin\theta$, where θ is the incidence angle of 31.8 degrees here.

The whole area is processed using the non-parametric SVD-Wiener algorithm described in A.2 followed by model order selection. The model order is limited to zero, one or two scatterers inside one azimuth-range pixel. As detailed in A.2, by using the conventional D-TomoSAR system model, this would result in a number of scatterers map and elevation and LOS linear deformation velocity for each of the detected (single and double) scatterers, respectively. However, according to the evidence in A.3, within the period of monitoring, rather than *linear*, a *seasonal* motion
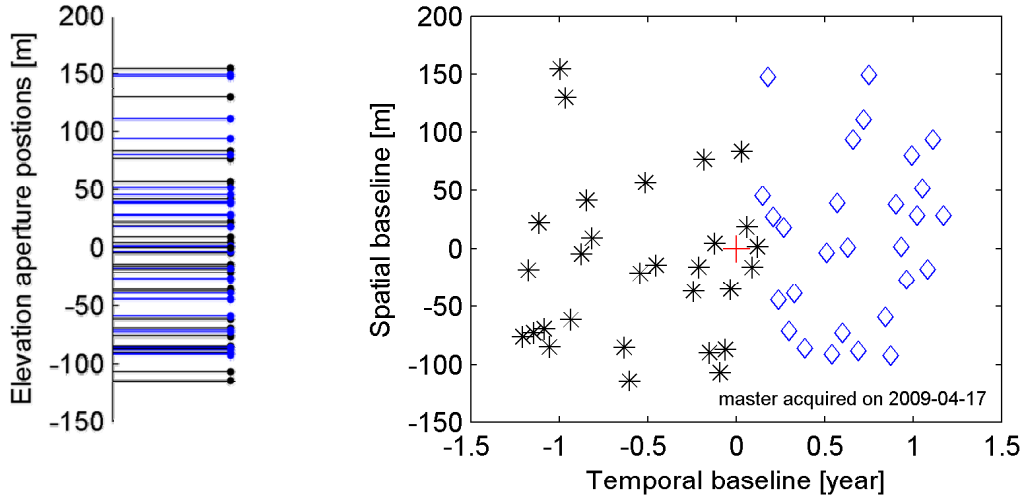
Figure 13. Left: Elevation aperture sampling positions of the 60 acquisitions. Right: spatial-temporal baseline distribution (asterisks: sub-stack with seasonal motion only; diamonds: sub-stack with seasonal and linear motion.).
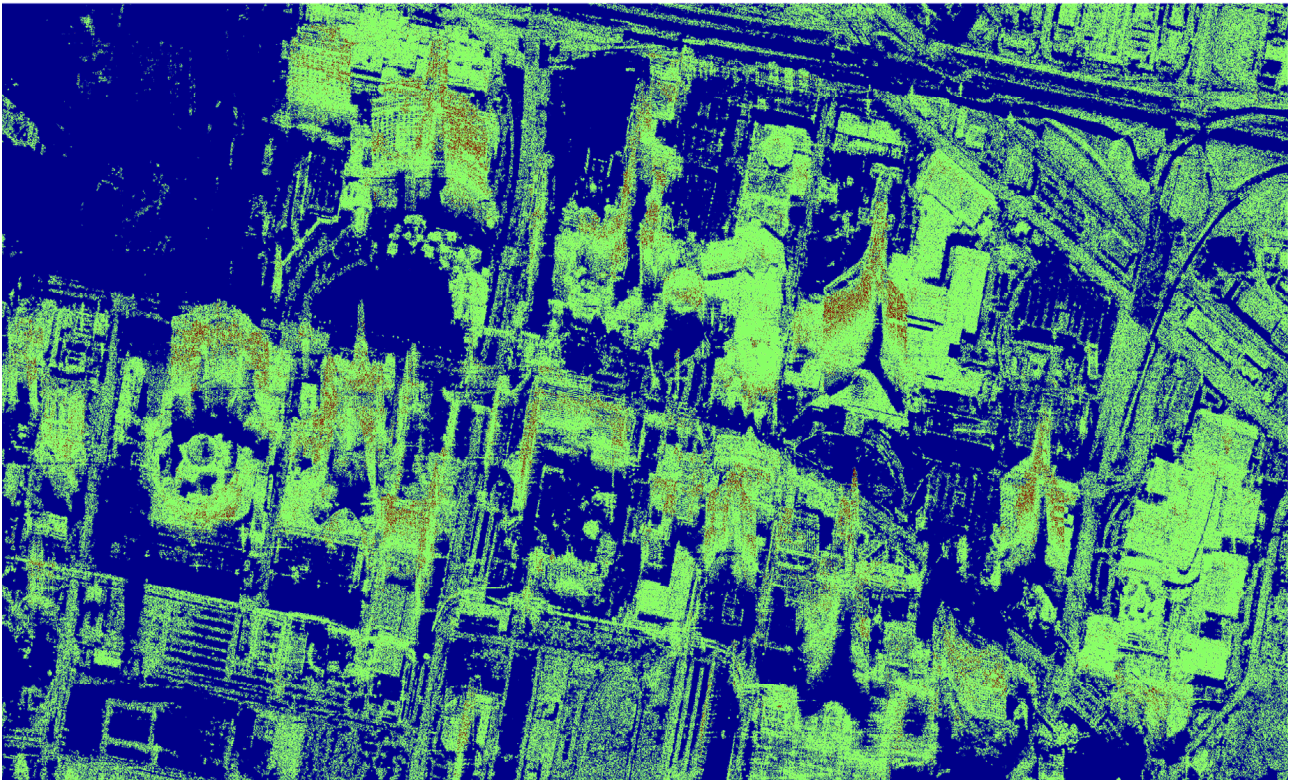
Figure 14. Results of model order selection: Number of scatterers map of the test area. Blue: zero scatterer inside the azimuth-range pixel; green: one scatterer inside the pixel; red: two scatterers inside the pixel. For the mean intensity image, see Fig. 1.

caused by thermal dilation dominates here, and hence the mean velocity is only in part able to explain this movement. To cope with the single-, and even multi-component nonlinear motion estimation problem, a time warp method is proposed in section 4.4 where the corresponding motion estimates will be detailed. In this section, only the following results are presented:

- **Number of scatterers map**

    Fig. 14 is the estimated number of scatterers map of the test area where blue indicates that no scatterer is detected in the analyzed azimuth-range pixel, green indicates that one (the so called "single scatterer") and red indicates that two scatterers (the so called "double scatterers", i.e. layover) are detected. It is obvious that the denser layovered pixels (colored
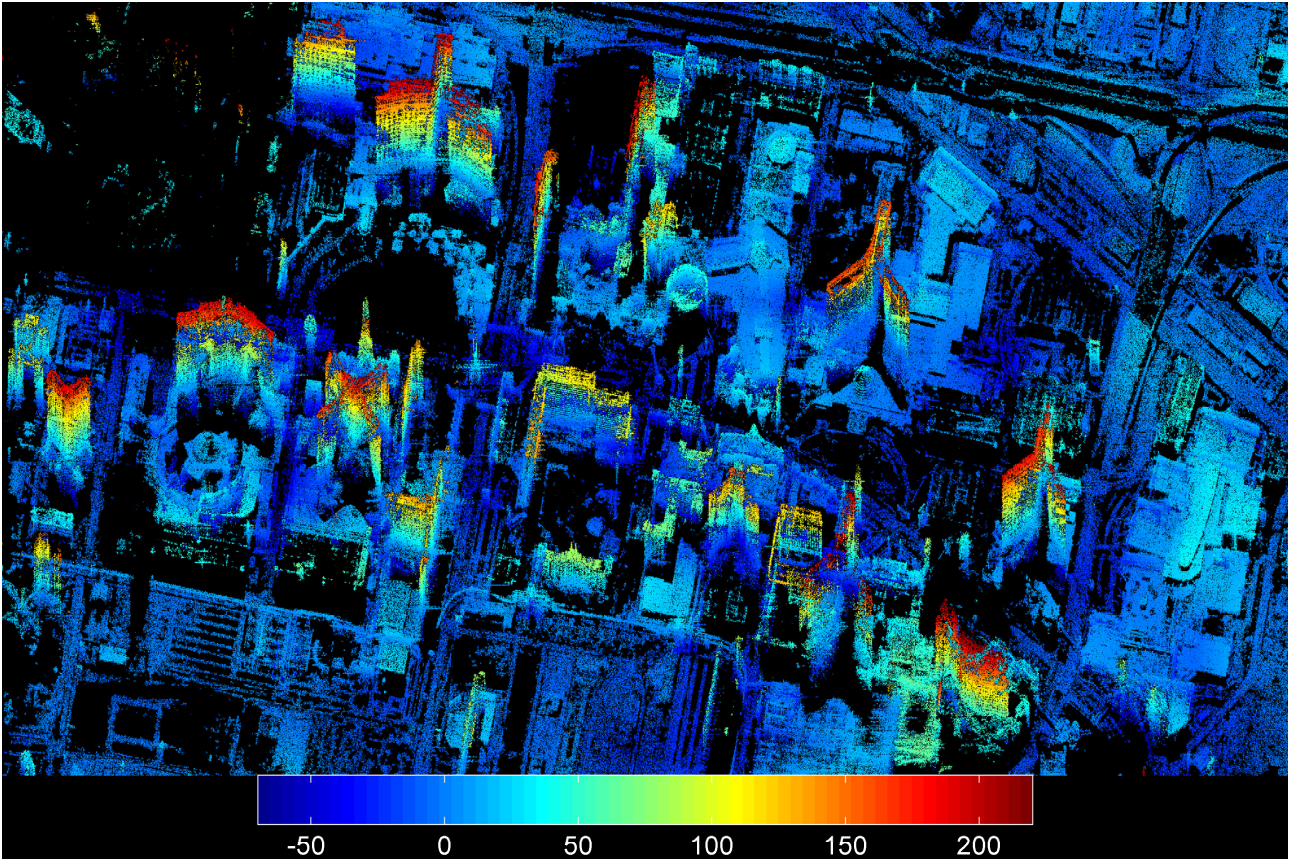
Figure 15. Topography estimates of the detected single scatterers [unit: meters].

by red) are located in the high rise areas.

- **Single scatterer analysis**

  Fig. 15 presents the topography estimates (in meters) of the detected single scatterers. Except for the layover areas, the full structure of the high rise city area is captured at a very detailed level. Besides the buildings, more detail such as the roads which have weak but correlated returns are well resolved.

- **Double scatterers analysis**

  Fig. 16 and 17 are associated with the topography estimates of the ground and top layer extracted from double scatterer results, respectively. These figures demonstrate the capability of the tomographic approach to separate the interfering layers associated with ground infrastructure and the façade of the building. Its effectiveness in resolving layover resolving is particularly evident in the homogeneity of the ground scatterers layer colored blue, and the smooth gradation of colors on the layer corresponding to the building façade. The high density of detected double scatterers should be noticed which complete the structures of individual high rise buildings shown in Fig. 15.

Fig. 18 gives an overview of the ultimate information exploration achieved by TomoSAR by combining the topographic estimates of the detected single and double scatterers. The full structure of the layovered high rise buildings is now retrieved. A zoom into the Bellagio hotel of the aforementioned results is presented in Fig. 19 – 20. For an individual building, e.g. the skyscraper on the right, a considerable amount of double scatterers of up to 20% is detected. This confirms the results with medium resolution SAR data in (Adam et al., 2005) and demonstrates the necessity of tomographic reconstruction for urban infrastructure monitoring (for more detail, see A.3).
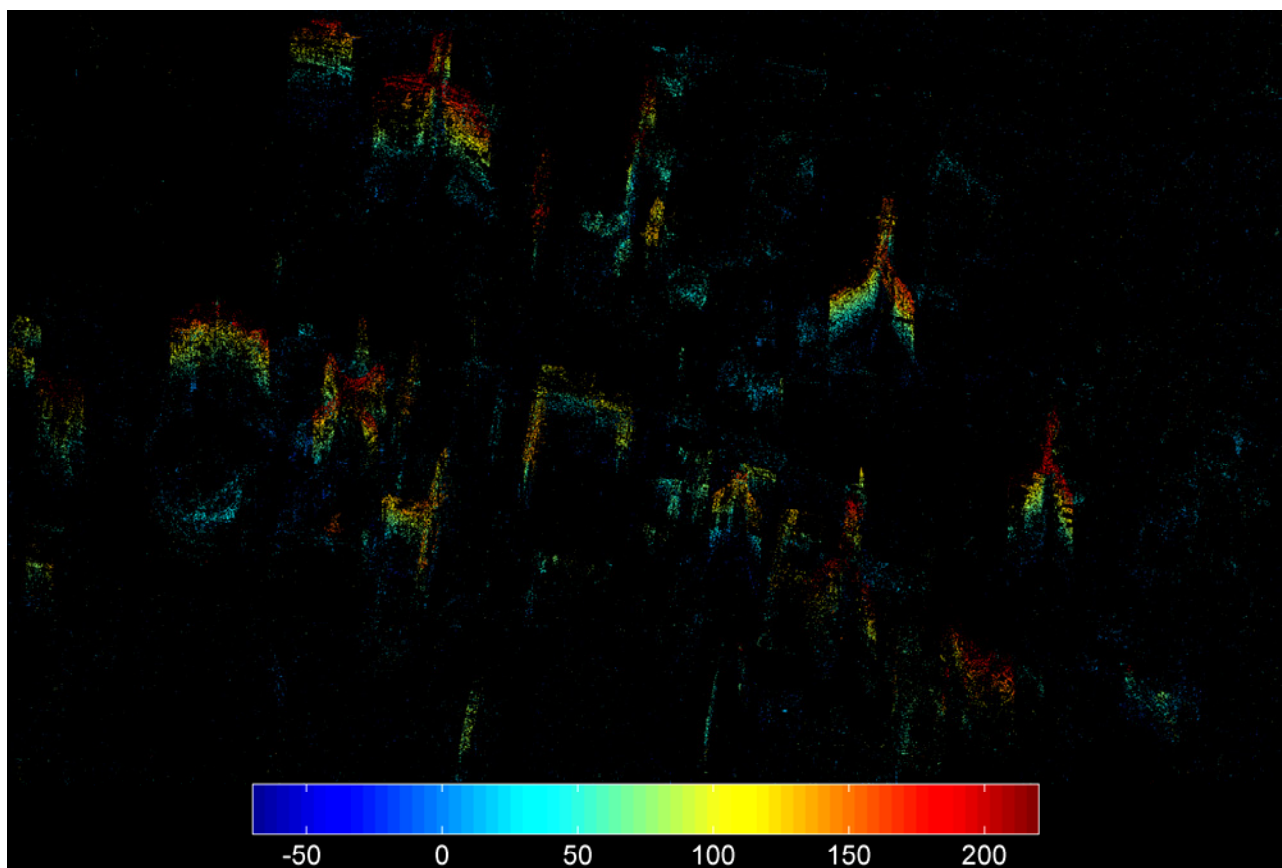
Figure 16. Topography estimates of the separated double scatterers - one of the two from the building façade [unit: meters].
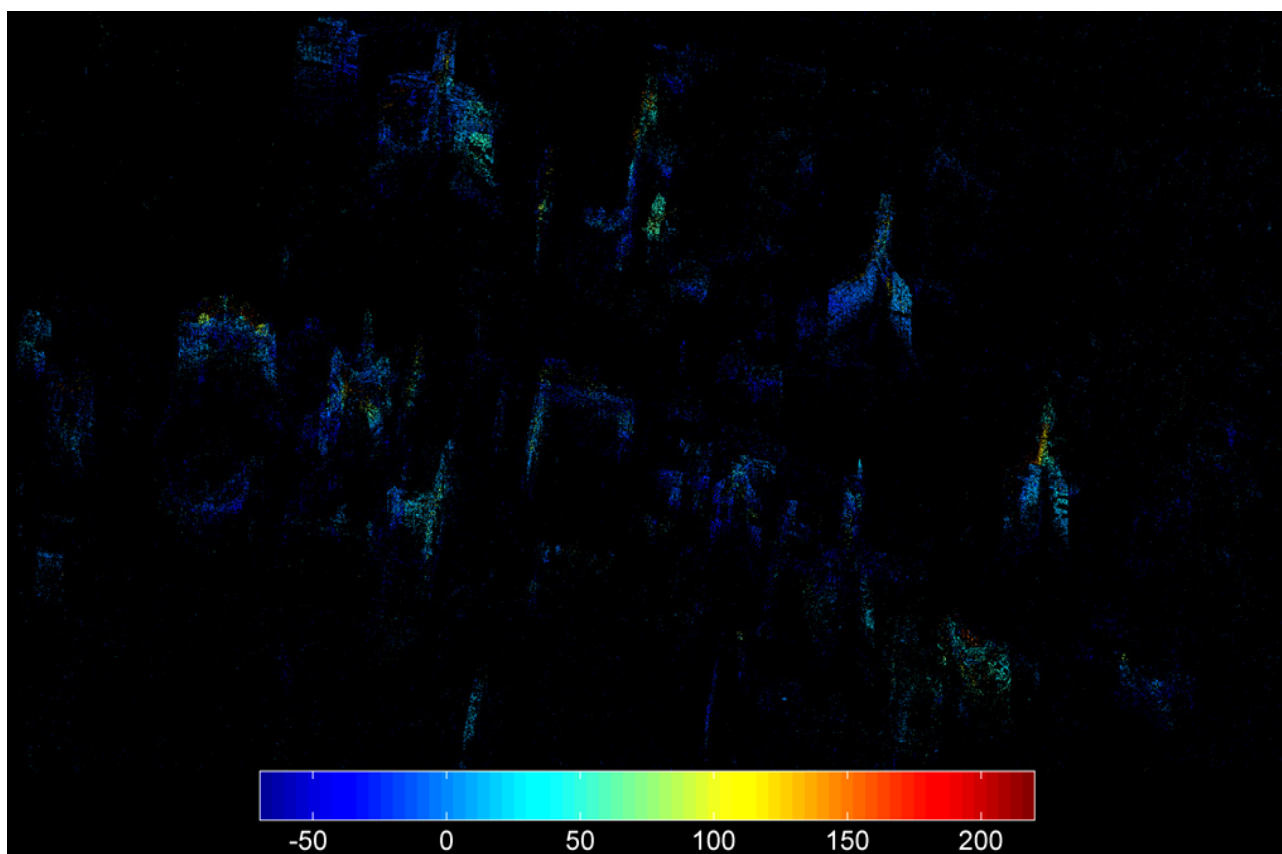
Figure 17. Topography estimates of the separated double scatterers - one of the two from the ground structures [unit: meters].
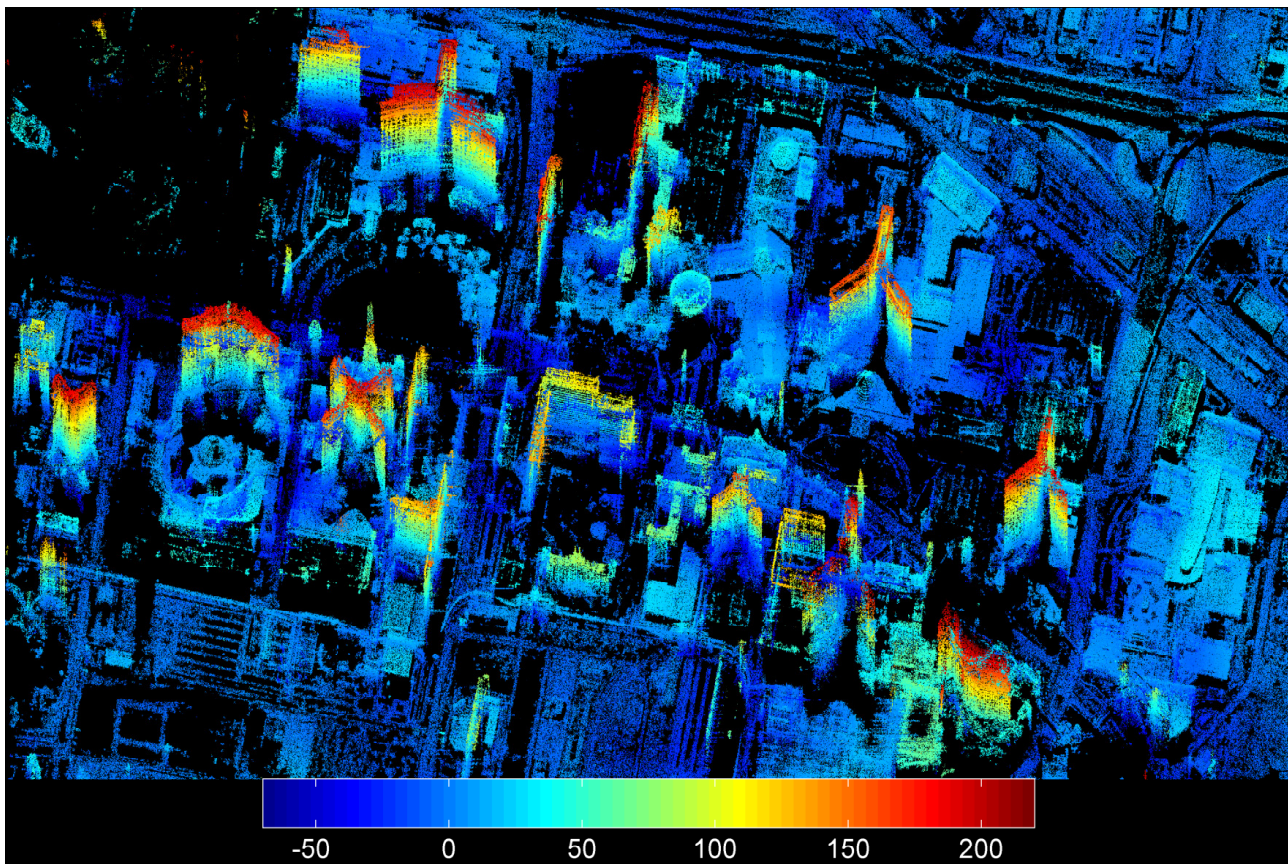
Figure 18. Completed information by adding the information of the layover area: topography estimates of the detected single and double scatterers [unit: meters]. Cf. Fig. 15. For double scatterers the higher one (façade) is displayed.
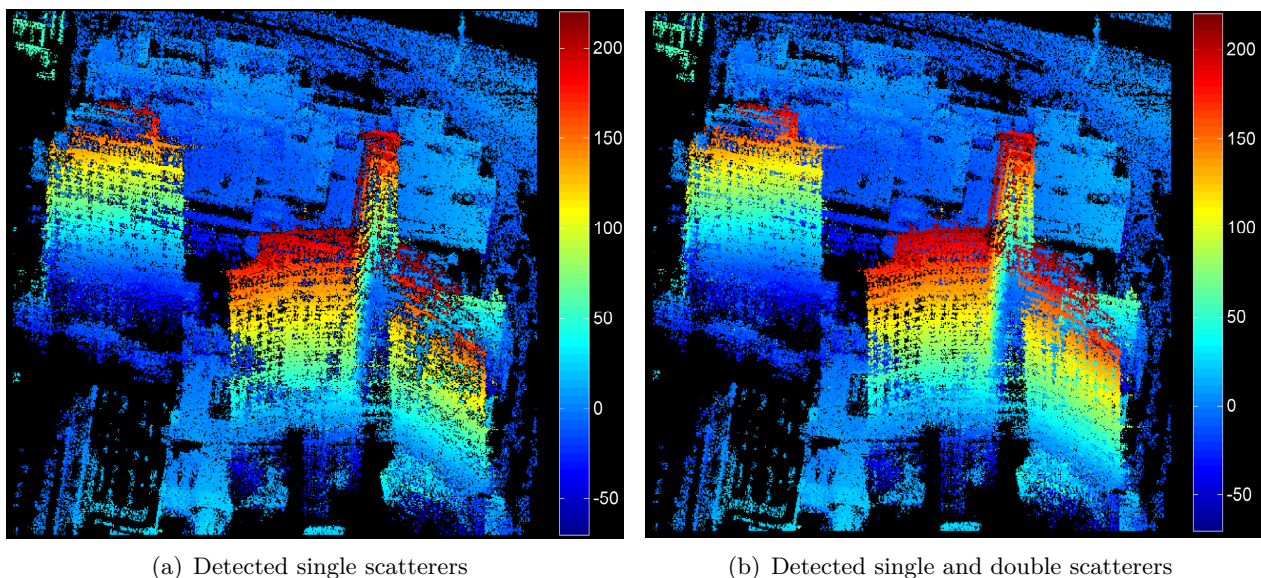
(a) Detected single scatterers

(b) Detected single and double scatterers

Figure 19. Information increment by adding the information of the layover area: Zoom into the Bellagio hotel area [unit: meters].

4.2 Sparse tomographic SAR reconstruction — the SL1MMER algorithm

The previous section demonstrated that VHR TomoSAR and D-TomoSAR can provide a very detailed view of individual buildings and therefore enable 3-D and 4-D SAR imaging in urban

(a) Building façade   (b) Ground structures
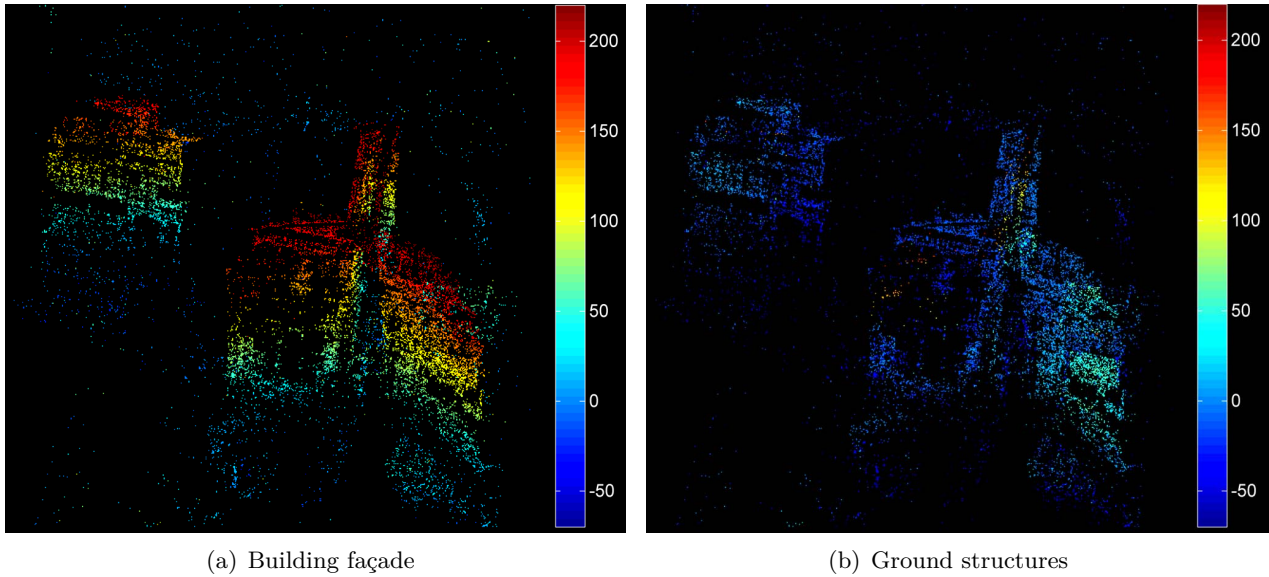
Figure 20. Topography estimates of the separated double scatterers: Zoom into the Bellagio hotel area [unit: meters].

areas. However, due to the tight orbit tube of TerraSAR-X the elevation aperture is small, i.e. the inherent resolution in elevation is about 50 times worse than in azimuth or range. This extreme anisotropy calls for SR algorithms in the elevation direction. In order to maintain the azimuth-range resolution for urban infrastructure imaging, these algorithms should not require averaging in azimuth and range.

This section summarizes the work in A.4 and partially in A.5, in which:

- The sparsity of the VHR SAR data in elevation is explored;
- CS for TomoSAR is outlined. By comparing it with parametric (NLS) and non-parametric (SVD-Wiener) reconstruction methods, the pros (e.g. SR properties and high point localization accuracy) and cons (e.g. outliers and biased reflectivity estimates) of CS for TomoSAR are discussed using simulations and real data;
- A CS based algorithm (SL1MMER) is proposed which improves the CS algorithm and gives reliable estimation of the number of scatterers, elevation, motion parameter, amplitude and phase of each scatterer.

4.2.1   Sparse representation of the signal in elevation

In VHR X-band data, for one azimuth-range cell, the following signal contributions are expected (see Fig. 21) (Zhu and Bamler, 2009):

- Weak diffuse scattering from —mostly horizontal or vertical —rough surfaces (roads, building walls). These contributions have an elevation extent of $\rho_r / \tan\theta$ for horizontal and $\rho_r \cdot \tan\theta$ for vertical surfaces. In both cases these extents are much smaller than the elevation resolution ρ_s, and hence these surfaces can be treated as *discrete* scatterers in elevation (delta-functions).
- Strong returns from metallic structures or specular and dihedral or trihedral reflections. These are points that would also be used in PSI. They are the dominating signal contributions. With VHR SAR data the density of these points can be very high (Bamler et al., 2009) (Gernhardt et al., 2010).
- Returns from volumetric scatterers, e.g. from vegetation. These may result in a continuous signal background in elevation. These ensembles of scatterers, however, often decorrelate in time, and their response is therefore treated as noise.

The noise sources are the following:
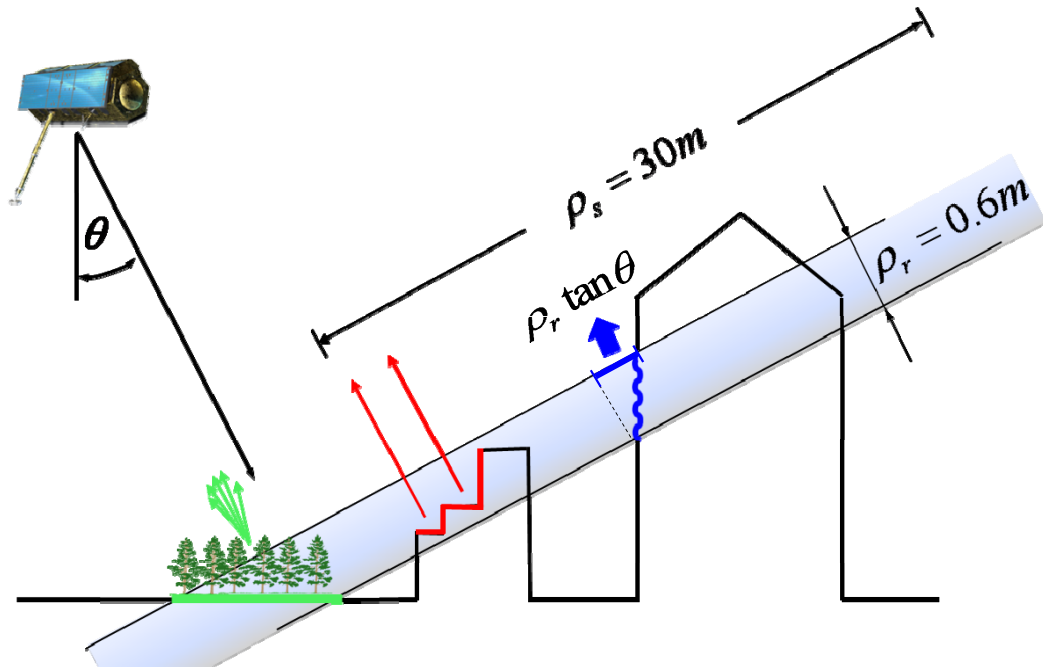
Figure 21. Possible signal contributions within one azimuth - range cell: The sparse representation of signal in elevation.

- Gaussian noise which is caused by thermal noise and the above mentioned temporal decorrelation.
- Calibration errors in amplitude. According to an unpublished DLR-internal calibration report (Buckreuss, 2009), the radiometric stability of TerraSAR-X, i.e. the amplitude variations within one stack, is 0.14 dB and is therefore negligible compared to the typical SNR.
- Phase errors caused by atmospheric delay and unmodelled motion. They require robust and phase error tolerant estimation methods.

As a consequence, the high anisotropy of the 3-D tomographic resolution element together with the fact that thick volumetric scatterers tend to temporarily decorrelate in X-band renders the signals *sparse* in the elevation direction, i.e. only a few (typically $0-4$) point-like reflections along elevation of unknown positions and unknown amplitudes and phases are expected per azimuth-range pixel. As the central concept of and a prerequisite for CS this property suggests using sparse reconstruction methods for tomographic SAR inversion.

4.2.2   SL1MMER —   a compressive sensing based algorithm

The proposed CS-based TomoSAR spectral estimation algorithm is named "Scale-down by L_1 norm Minimization, Model selection, and Estimation Reconstruction" (SL1MMER, pronounced "slimmer"). It consists of three main steps: 1) a dimensionality scale-down by L_1 norm minimization, 2) model selection and 3) linear parameter estimation (see Fig. 22). Note that although the SL1MMER algorithm was originally implemented for differential TomoSAR (D-TomoSAR), in order to simplify the mathematics, the motion is neglected in the following introduction to the algorithm.

1. Scale-down by L_1 norm minimization

Since $L \gg N$, the system model eq. (5)

$$\mathbf{g} = \mathbf{R}\,\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{28}$$
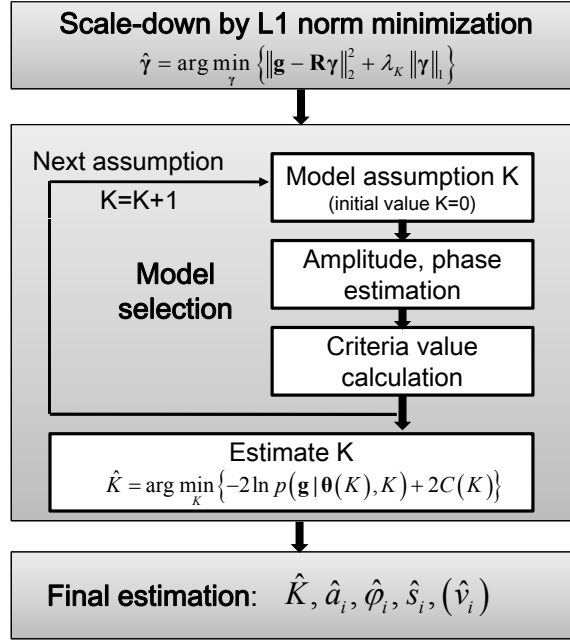
Figure 22. Flow chart of the SL1MMER algorithm.

is severely under-determined. Hence, there are infinitely many solutions. Considering the sparsity of the signal in elevation ($K = 0 \sim 4$), according to the CS theory described in section 2.2, $\boldsymbol{\gamma}$ can be very well approximated by the solution with the least number of scatterers (non-zero elements of $\boldsymbol{\gamma}$), i.e. the minimal L_0 norm. The required number of measurements depends mainly on the sparsity K instead of the length of the signal L, i.e. $N \gg O(K \log(L/K))$. In case there is no prior knowledge about K and in the presence of measurement noise, the reconstruction of $\boldsymbol{\gamma}$ with CS boils down to the L_1- L_2 norm minimization of eq. (15):

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} \left\{ \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_2^2 + \lambda_K \|\boldsymbol{\gamma}\|_1 \right\} \tag{29}$$

As described in section 2.2, eq. (29) gives the sparsest estimation of $\boldsymbol{\gamma}$ if (and only if) our sparse mapping matrix $\mathbf{R}$ fulfills the RIP and incoherence properties. However, for our application TomoSAR, RIP and incoherence are violated for several reasons. First the mapping matrix $\mathbf{R}$ is pre-determined by the measurement system (the elevation aperture sampling pattern) and may not be optimum. Second, the profile $\gamma(s)$ to be reconstructed is often sampled much more densely than the elevation resolution unit ρ_s in order to allow for good resolution and scatterer positioning accuracy. The small sampling distance renders $\mathbf{R}$ over-complete, reduces RIP and increases coherence. This fact may introduce outliers in the estimated $\hat{\boldsymbol{\gamma}}$ (see section 2.2, Fig. 11). In addition, as detailed in section 2.2, the L_1 approximation introduces systematic amplitude biases. Those artifacts are not critical when the aim is only to reconstruct a reasonable reflectivity profile. However, high-quality sparse tomographic SAR inversion requires the estimation of the number of scatterers, as well as the amplitude, phase, and elevation of each scatterer. Hence, special care must be taken of these nuisance artifacts.

Nonetheless, CS plays a perfect role for dimensionality reduction as it gives a robust estimate of the plausible positions of the scatterers. By only selecting columns corresponding to the non-zero elements of $\hat{\boldsymbol{\gamma}}$, $\mathbf{R}$ is scaled down significantly. It renders the severely under-determined system model finally over-determined (see Fig. 23.a).

2. **Model selection**

Step 1 shrinks $\mathbf{R}$ dramatically and gives a first sparse estimate of $\boldsymbol{\gamma}$. This estimate, though, may
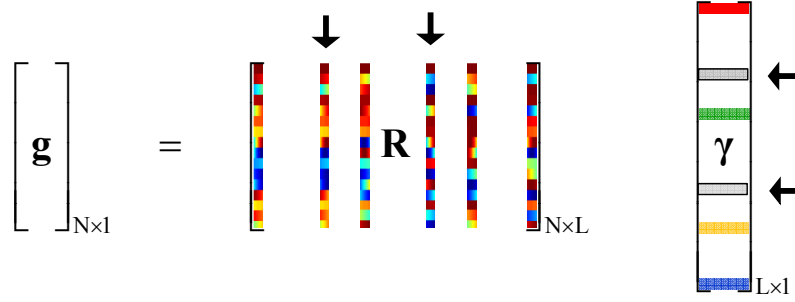
(a) Step 1: Scale-down of the underdetermined linear system by L_1 norm minimization. A sparse estimate $\hat{\boldsymbol{\gamma}}$ is obtained by eq. (29)(colored elements). Among them there might be outliers (two gray elements with arrows). The sparse mapping matrix $\mathbf{R}$ is scaled down dramatically by choosing its columns corresponding to the non-zero elements of $\hat{\boldsymbol{\gamma}}$, i.e. the colored columns.
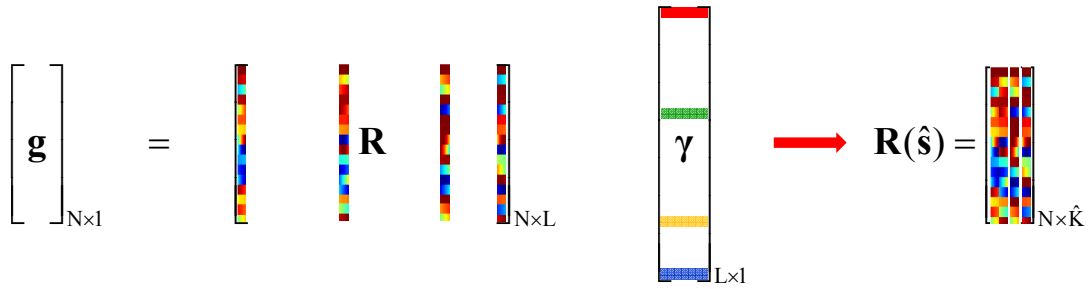
(b) Step 2: Model selection. The outliers are removed and the number of the scatterers K is estimated. By choosing those $\hat{K}$ columns from $\mathbf{R}$, the final mapping matrix $\mathbf{R}(\hat{\mathbf{s}})$ is obtained.

(c) Step 3: Final parameter estimation. The amplitude and phase, i.e. complex-valued reflectivity of the $\hat{K}$ scatterers are estimated from the aforementioned system.

Figure 23. Graphic illustration: The basic idea of SL1MMER

still contain the aforementioned outliers (Fig. 11) and, hence, the sparsity K (i.e. the number of the scatterers) is often overestimated. Model selection is used to clean the $\boldsymbol{\gamma}$ estimate of spurious, non-significant scatterers and to finally obtain the most likely number of scatterers $\hat{K}$ inside the azimuth-range pixel (Zhu et al., 2009). It provides a trade-off between the model complexity and the model fit.

The model complexity can be described by the number of parameters k. In our case $k = 3K$ because each scatterer is characterized by three parameters (amplitude, phase, and elevation) (or $4K$ for D-TomoSAR by adding an additional motion parameter). Let further $\boldsymbol{\theta}(K)$ be the vector of the unknown amplitudes, phases, and elevations for all K scatterers. The goodness of model fit can be described by the likelihood $p(\mathbf{g}|\boldsymbol{\theta}(K), K)$. A more complex model always fits the observations better, hence, for selecting the true model, the model complexity must be penalized to avoid over-fitting of the data. This is the main concept of penalized likelihood criteria. They have the following general form (Bozdogan, 1987) (Hansen and Yu, 2001):

$$\hat{K} = \arg \min_K \left\{ -2 \ln p \left(\mathbf{g}|\hat{\boldsymbol{\theta}}(K), K \right) + 2C(K) \right\} \tag{30}$$

where $C(K)$ is the complexity penalty. Note that for stationary white Gaussian measurement noise, with covariance matrix $\mathbf{C}_{\varepsilon\varepsilon} = \sigma_\varepsilon^2 \mathbf{I}$, the log-likelihood term is simply:

$$-2 \ln p \left(\mathbf{g} | \hat{\boldsymbol{\theta}}(K), K \right) = \sigma_\varepsilon^2 \left\| \mathbf{g} - \mathbf{R} \hat{\boldsymbol{\gamma}} \right\|_2^2 \tag{31}$$

Many types of penalty terms are proposed in the literature, e.g. Bayesian information criterion (BIC) (Burnham and Anderson, 2004), Akaike information criterion (AIC) (Schwarz, 1978), and minimum description length (MDL) (Rissanen, 1978). As illustrated in Fig. 22, for each model assumption $K = 0, \ldots, 4$, the elevations are given by the previous scale-down step, thus, the amplitudes and phases can be easily obtained by linear least-squares estimation. The preferred number of scatterers $\hat{K}$ inside an azimuth-range pixel is finally the one with the lowest penalized likelihood criterion value according to (30) (see Fig. 23.b). Implementation details of different model selection schemes for VHR TomoSAR are addressed in (Zhu et al., 2009).

3. Parameter estimation

As a last refinement, a much slimmer mapping matrix $\mathbf{R}(\hat{\mathbf{s}})$, i.e. the $N \times \hat{K}$ matrix with $\mathbf{R}_{n,\hat{k}}(\hat{\mathbf{s}}) = \exp(-j2\pi\xi_n \hat{s}_k)$, is built up, and the final complex-valued reflectivity $\boldsymbol{\gamma}(\hat{\mathbf{s}})$ for the $\hat{K}$ scatterer is obtained by solving the following over-determined linear system equation:

$$\mathbf{g} = \mathbf{R}(\hat{\mathbf{s}}) \, \boldsymbol{\gamma}(\hat{\mathbf{s}}) + \mathbf{e} \tag{32}$$

where $\mathbf{e}$ combines the measurement noise and the model error, i.e. the deviation from sparsity or unmodelled motion-induced phase terms (in the D-TomoSAR case). The sparse reflectivity profile $\hat{\boldsymbol{\gamma}}(\hat{\mathbf{s}})$ is finally found by standard least-squares estimation (see Fig. 23.c):

$$\hat{\boldsymbol{\gamma}}(\hat{\mathbf{s}}) = \left(\mathbf{R}^{\mathrm{H}}(\hat{\mathbf{s}}) \, \mathbf{R}(\hat{\mathbf{s}}) \right)^{-1} \mathbf{R}^{\mathrm{H}}(\hat{\mathbf{s}}) \, \mathbf{g} \tag{33}$$

By introducing model selection (30) and refinement of the parameter estimation eq. (33), SL1MMER combines the advantages of CS with the amplitude and phase accuracy of linear estimators. As demonstrated in A.4 and A.5, compared to non-parametric and fully parametric L_2-norm methods, SL1MMER (Zhu and Bamler, 2010b) (Zhu and Bamler, 2010a):

- Is more robust than NLS parametric methods with respect to unmodelled phase errors and does not suffer from self-cancellation artifacts, like CAPON (Capon, 1969) (Lombardini and Reigber, 2003);
- Approaches the elevation estimation accuracy of NLS (i.e. the CRLB) under Gaussian noise for two interfering scatterers;
- Is computationally more efficient than NLS;
- Requires no prior knowledge about the number of scatterers K;
- Provides amplitude and phase estimation accuracy approaching the CRLB;
- Can particularly achieve substantial SR in elevation while maintaining the full azimuth-range resolution.

Here, an inductive simulation example is provided to show the SR of SL1MMER. Since multiple scatterers inside an azimuth-range pixel most likely occurs in high rise urban areas, the situation that two scatterers inside one pixel (one point scatter from the building façade and another rough surface reflection from the ground, see the simulation scenario in Fig. 24) is simulated using the elevation aperture distribution of the asterisk sub-stack in Fig. 13 (elevation resolution $\rho_s = 40.5$ m, $N = 30$). The building is assumed to have a height of 40 m which results in a maximum elevation extent of $\Delta s = 80$ m $\approx 2\rho_s$ by assuming an incidence angle of 30°. The decorrelation effect is introduced by adding Gaussian noise with a SNR of 10 dB. A perfect TomoSAR algorithm should always be able to separate the contributions from the building façade (elevation varying from 0 to 80 m) and the ground (zero elevation) even with a very small distance. Assuming the two scatterers have the same amplitude and phase, Fig. 25 shows
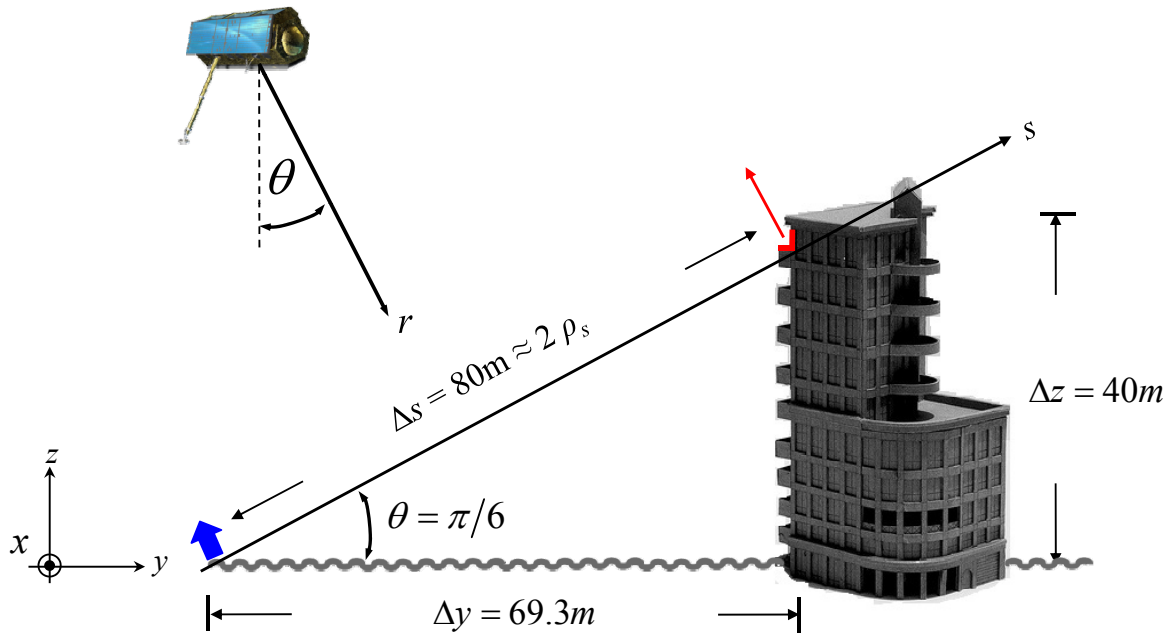
Figure 24. Simulation scenario: Two scatterers inside one resolution cell (one from the building façade and another from the ground).

a comparison of the reconstructed reflectivity profiles along the elevation direction using SVD-Wiener (red) and SL1MMER (blue). Red lines represent the reconstruction using the non-parametric method SVD-Wiener. Blue lines show the same result using SL1MMER. Starting with two scatterers with elevation of 0 and 72.9 m, i.e. the distance $\delta s = 1.8\rho_s$, as shown in Fig. 25.a, SL1MMER reconstructs spectral lines instead of sinc-like point response functions. Both methods can distinguish the two scatterers well. However, with increasing range, once the scatterers move closer into one elevation resolution cell with elevations of 0 and 32.4 m (Fig. 25.b), i.e. $\delta s = 0.8\rho_s$, SVD-Wiener is no longer able to distinguish them while CS detects very clearly two spectral lines with even closer scatterers, e.g. with a distance of $0.4\rho_s$ (see Fig. 25.c). In A.4, the reflectivity profile in elevation is reconstructed for each "range pixel" along the LOS direction using SL1MMER, SVD-Wiener and NLS. The comparison reveals that SL1MMER provides a SR capability comparable to the parametric NLS estimator (see A.4 for more details).

4.3   Systematic performance assessment of SL1MMER

The results shown in Fig. 25 lead naturally to the question of what the limits of the technique are. Therefore, a systematic performance assessment of the proposed SL1MMER algorithm is performed in A.5 concerning the following questions:

- How accurately can the positions of two closely spaced scatterers be estimated?
- What is the closest distance of two scatterers, such that they can be separated with a detection rate of 50% at given $N \cdot SNR$ and amplitude ratio with a uniformly distributed phase difference?
- How many acquisitions N are required for robust estimation? We define "robustness" as the capability of separating two scatterers spaced by one Rayleigh resolution unit with a probability of 90%.

Without going too much into detail, this section summarizes the findings in A.5.

- Compared to the CRLB, both numerical results and an analytic approximation of the
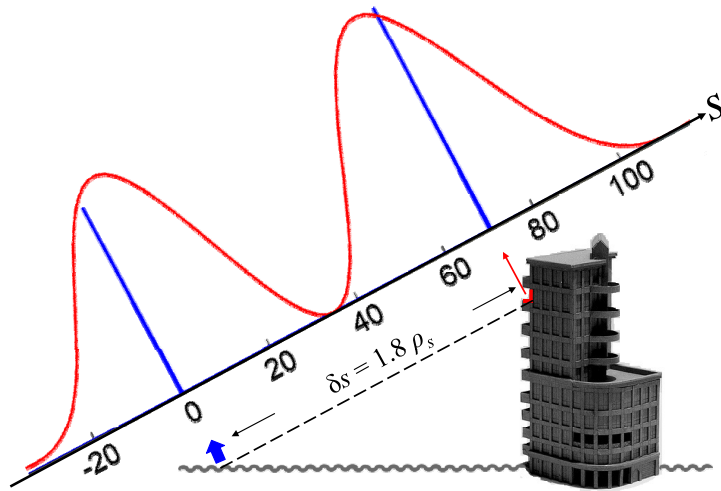
(a) $\delta s = 1.8\rho_s$
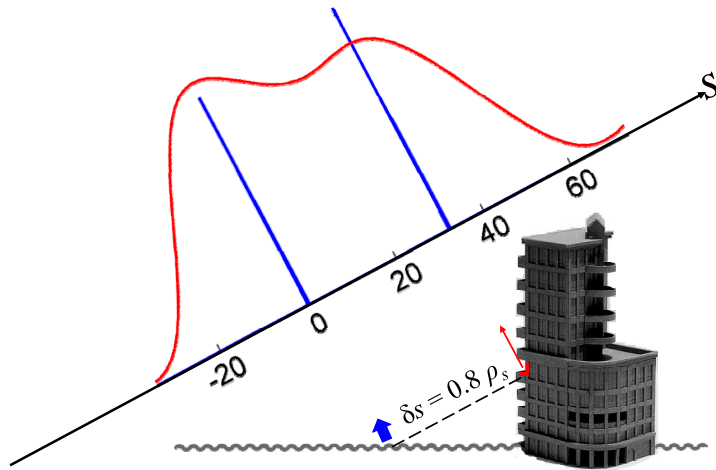
(b) $\delta s = 0.8\rho_s$
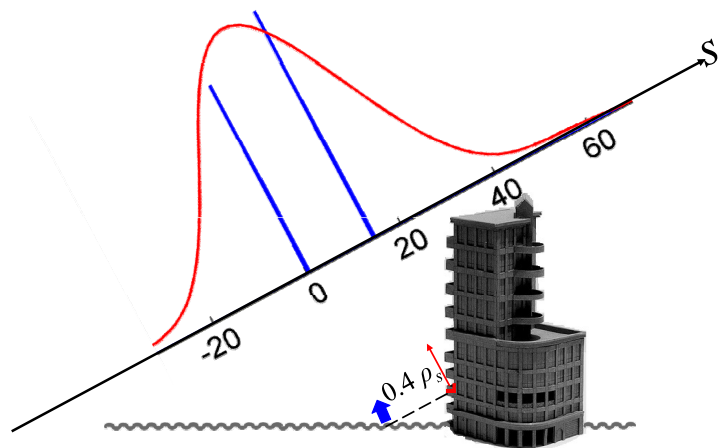
(c) $\delta s = 0.4\rho_s$

Figure 25. Comparison of the reconstructed reflectivity profiles along elevation direction between SVD-Wiener and SL1MMER under a SNR of 10dB. Red solid line: SVD-Wiener; Blue solid line: SL1MMER. (a): $\delta s = 1.8\rho_s$; (b): $\delta s = 0.8\rho_s$; (c): $\delta s = 0.4\rho_s$.

elevation estimation accuracy are provided in section 4.3.1.
- The SR factors are found by extensive simulations, and an easy-to-use polynomial approximation is provided in section 4.3.2. Although TomoSAR has been taken as the preferred application, the findings on SR in A.5 are generally applicable to sparse spectral estimation, including SR SAR focusing of point-like objects. In addition, the results in A.5 are approximately applicable to NLS estimation, and hence establish a fundamental bound for SR of spectral estimators.
- The conditions for ensuring 90% detection rate of two scatterers with a distance of $\rho_s(\alpha = 1)$ are given in section 4.3.3 by explicit formulas.

Note that the BIC penalized likelihood criterion with a penalty term $C(K) = 0.5\ k \ln N = 1.5K \ln N$ is used throughout the investigation in all algorithms, i.e. the subsequent comparison is independent of the model selection scheme.

4.3.1   Location estimation accuracy

In the case of single scatterers, the location estimation accuracy of the SL1MMER algorithm is investigated in A.4 under circular Gaussian noise and non-Gaussian phase noise. It is demonstrated that the location (elevation) estimation accuracy of SL1MMER approaches the CRLB under Gaussian noise and is superior than NLS under non-Gaussian phase noise. In this section, the investigation will be focused on the case that two scatterers within one (azimuth-range) pixel interfere with each other.

Assuming two scatterers at elevation positions s_1 and s_2 with amplitudes a_1 and a_2 and phases φ_1 and φ_2, the unknown parameter vector can be written as $\boldsymbol{\theta} = [\,a_1\ \varphi_1\ s_1\ a_2\ \varphi_2\ s_2\,]$. In A.5, the CRLB for location estimation errors, i.e. σ_{s_1} and σ_{s_2}, are derived by taking the inverse of the *Fisher information matrix* $\mathbf{J}$. For the sequel the elevation estimation error is split into two contributions:

$$\sigma_{s_q} = c_0 \cdot \sigma_{s_q,0} \tag{34}$$

where $\sigma_{s_q,0}$ is the CRLB of the elevation estimates of the qth scatterer in the absence of the other one and c_0 is the essential interference correction factor for closely spaced scatterers. Among them:

$$\sigma_{s_q,0} = \frac{\lambda r_0}{4\pi\sqrt{2} \cdot \sqrt{N \cdot SNR_q} \cdot \sigma_b} \tag{35}$$

σ_b is the standard deviation of the elevation aperture sample positions. For uniformly distributed baselines (randomly or equidistantly spaced) b_n and $\sigma_{s_q,0}$ can be related to the elevation resolution ρ_s by:

$$\sigma_{s_q,0} = \frac{\sqrt{3/2}}{\pi} \frac{\rho_s}{N \cdot SNR_q} \approx \frac{0.39}{N \cdot SNR_q}\rho_s \tag{36}$$

In Fig. 26 we plot $\sigma_{s_q,0}$ normalized to the Rayleigh resolution unit ρ_s, as a function of $N \cdot SNR_q$ to explore its impact on estimation accuracy where $\sigma_{s_q,0}$ follows exactly a -0.5 power law. In the most relevant parameter range, i.e. $N = 10, \ldots, 100$ and $SNR = 0, \ldots, 10dB$, the elevation estimation accuracy of a single scatterer is on the order of roughly $1/100$ to $1/10$ of the resolution unit.

For the subsequent discussion only the impact of the normalized CRLB will be considered:
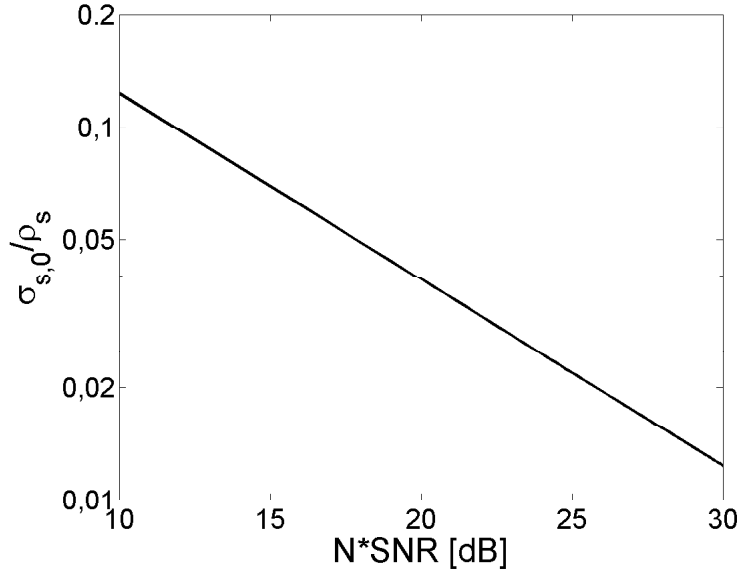
Figure 26. CRLB $\sigma_{s,0}$ of the elevation estimation accuracy of a single scatterer, normalized to the elevation resolution, as a function of $N \cdot SNR$.

$$c_0 = \frac{\sigma_s}{\sigma_{s,0}} \tag{37}$$

The dependency factors on c_0 are systematically investigated in A.5. For a fixed elevation aperture, c_0 is almost independent of N and SNR, instead, it depends on the normalized distance α (distance divided by ρ_s), the number of samples N (for very small N only) and the phase difference $\Delta\varphi$.

- **Dependency on α**

  If the two scatterers are separated by much more than one Rayleigh resolution unit, i.e. $\alpha \gg 1$, their elevation estimation accuracy is the one of individual scatterers, i.e. c_0 approaches 1. The closer they move, the more they interfere and the larger their elevation estimation error becomes, i.e. $c_0 > 1$.

- **Dependency on N**

  Besides the $N \cdot SNR_q$ term in eq. (35), for small N, c_0 has a non-negligible additional dependency on N. For the elevation estimation accuracy and super-resolution study, we will neglect this effect and all experiments are based on sufficiently large N, e.g. $N > 15$. The situation of small N will be discussed in section 4.3.3.

- **Dependency on $\Delta\varphi$**

  The phase difference $\Delta\varphi$ affects c_0 with a period of π. The closer the two scatterers are, i.e. the smaller α, the more pronounced is the dependence on their phase difference. For $\alpha < 1$, the effect of $\Delta\varphi$ is very significant. For $\Delta\varphi = 0$, and π, the two scatterers interfere with each other the most.

Since for the SR study the case of small α is particularly of interest, the impact of $\Delta\varphi$ must be considered. For TomoSAR, the phase difference $\Delta\varphi$ is totally random because it depends on the unknown geometric configuration. Hence, the CRLB for the elevation estimation error is calculated by integrating the variances over $\Delta\varphi$. The resulting dependence can be approximated by the following analytic expression:
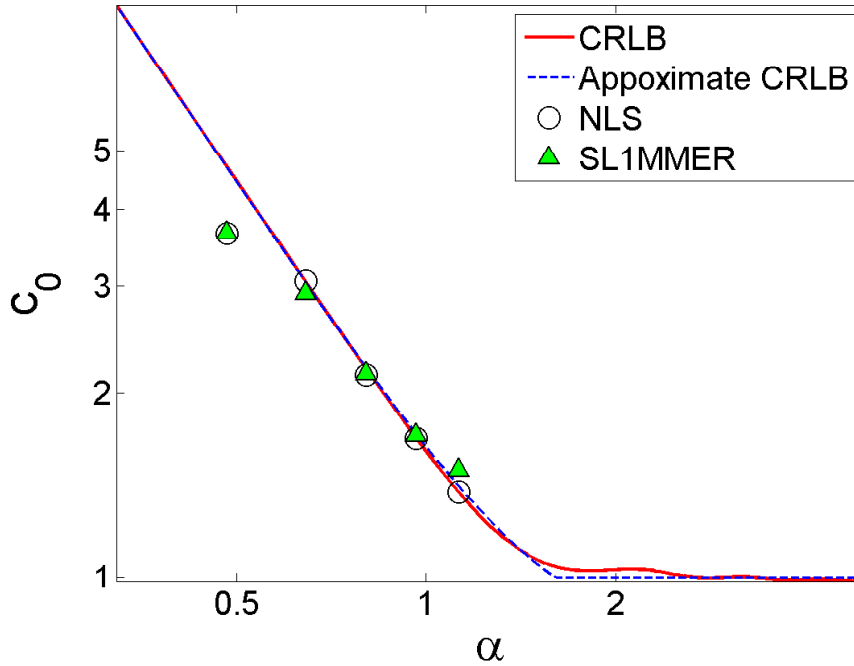
Figure 27. Theory vs. approximation vs. experiments: Normalized CRLB c_0 of the elevation estimation accuracy of two interfering scatterers integrated over $\Delta\varphi$ as a function of their normalized distance α.

$$c_0^2 = \frac{\sigma_s^2}{\sigma_{s,0}^2} \approx \max\left\{2.57(\alpha^{-1.5} - 0.11)^2 + 0.62, 1\right\} \tag{38}$$

It fits the CRLB in this range of α slightly better than the approximation derived in (Swingler, 1993), $c_0^2 = \max\left\{(15/\pi^2)^{2.5}\alpha^{-3}, 1\right\}$. The square-root of the normalized CRLB variance integrated over $\Delta\varphi$ as a function of α is plotted in Fig. 27 together with the approximation eq. (38) from the author. The dependency on $\alpha^{-1.5}$ can be clearly observed. The normalized standard deviation of elevation estimates for two interfering scatterers using NLS and SL1MMER as a function of α is also shown in Fig. 27. These have been obtained by simulations with randomly distributed phase differences. The elevation estimation accuracy of SL1MMER approaches the CRLB and is comparable to NLS. For the leftmost experiment point, i.e. for $\alpha = 0.47$, NLS and SL1MMER appear to have a better accuracy than the CRLB. This is caused by the fact that we always assume that the two closely spaced scatterers never exchange their positions which is not valid when moving the two scatterers closer such that the standard deviation increases and reaches the order of the distance between them.

4.3.2   Super-resolution power

Nonlinear and parametric spectral estimation methods yield reconstructions with much sharper point responses than traditional non-parametric linear algorithms. One might be tempted to infer a very high resolution power from this narrow point response width. However, the needle-like responses tell us neither the location accuracy of the individual points nor the ability of the algorithm to resolve two close scatterers.

In A.5, the (elevation) resolution is defined by the minimum distance ρ_{P_D} between two δ-functions (scatterers) that are separable at a given SNR with a certain N, and at a pre-specified probability of detection P_D. Throughout this section ρ_{P_D} is normalized to the Rayleigh resolution unit ρ_s. The SR factor κ_{P_D} is defined by:

$$\kappa_{P_D} = \frac{\rho_s}{\rho_{P_D}} \tag{39}$$
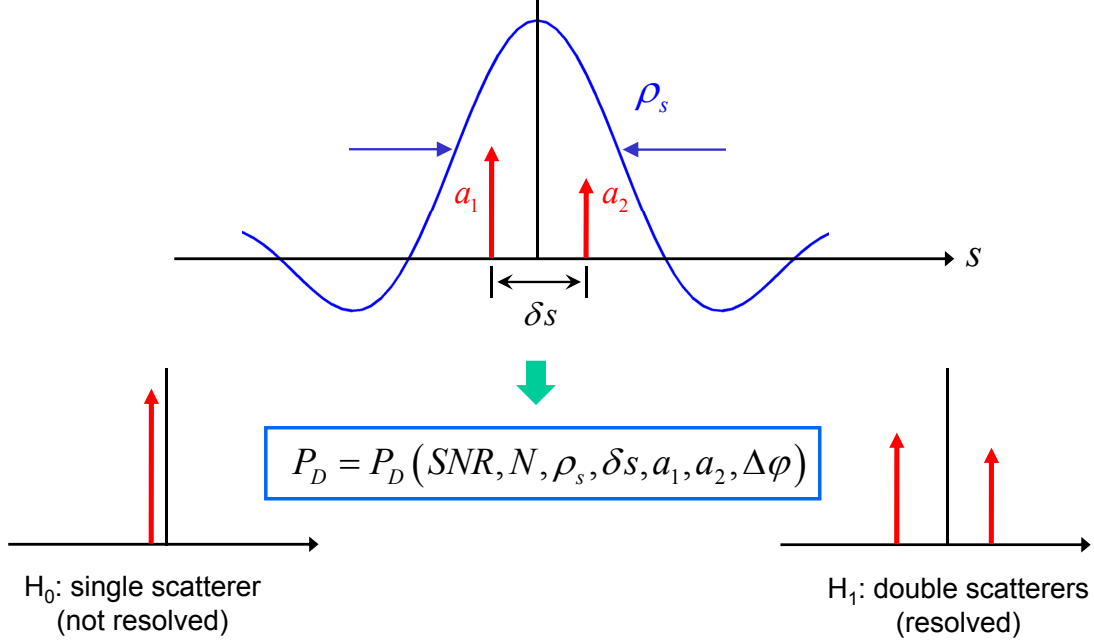
Figure 28. Super resolution is a detection problem. H_0: zero or only one scatterer inside the given azimuth-range cell; H_1: two scatterers inside the given azimuth-range cell. The detection rate depends on SNR, number of samples N, elevation resolution ρ_s, the distance between the two scatterers δ_s, the amplitude ratio a_1/a_2 and the phase difference between the two scatterers $\Delta\varphi$.

It depends on the required detection rate P_D and is larger than unity for SR. In this section, the SR power for different N, SNR, as well as amplitude ratios and relative phase differences $\Delta\varphi$ of the two scatterers are investigated.

The problem of resolution can be regarded as a hypothesis test with the two hypotheses (Fig. 28):

H_0: zero or only one scatterer inside the cell;
H_1: two scatterers inside the cell.

The goal is to derive the SR power of an estimator that provides location, amplitude and phase of each scatterer. Since it is very complicated to find a theory supporting all possible situations, the problem is approached experimentally. Starting with an example of detecting two scatterers with the same amplitudes and phases by using $N = 25$ regularly sampled acquisitions, Fig. 28 shows the detection rate P_D as a function of normalized distance α at different SNR levels using SL1MMER (red), NLS (green) and maxima detection (blue). Maxima detection (MD) means SVD-Wiener linear reconstruction followed by model order selection (see A.2). The SNR of each scatterer for the two sets of curves is 0 and 6 dB, respectively. It gives us several perspectives of the problem:

- The Rayleigh resolution unit is rather a definition in the absence of noise since P_D for MD increases with SNR.
- NLS is identical to a MLE under Gaussian noise with large N and high SNR. Hence, it should result in the highest detection rate if there is no a priori knowledge about the true hypothesis. In Fig. 29, NLS gives an improved P_D compared to MD. The detection rate depends strongly on the $N \cdot SNR$.
- The SR power of SL1MMER is slightly better than that of NLS. The better detection rate is due to the implicit model selection of the algorithm which gives a prior knowledge about the true hypothesis.

Due to the similarity between NLS and SL1MMER, all the following results are approximately applicable to NLS as well and, hence, establish a fundamental limit for SR. It is pointed out in A.5 that:
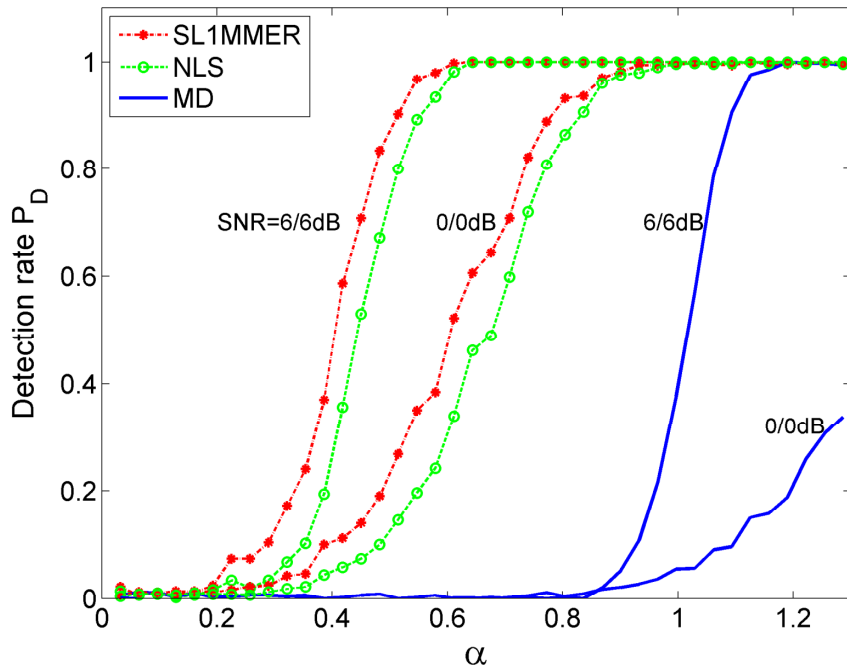
Figure 29. Detection rate as a function of α using SL1MMER (dashed star), NLS (dashed circle) and MD (solid) with $SNR = 0/0$dB and $6/6$dB, $N = 25$ and $\Delta\varphi = 0$ (worst case).
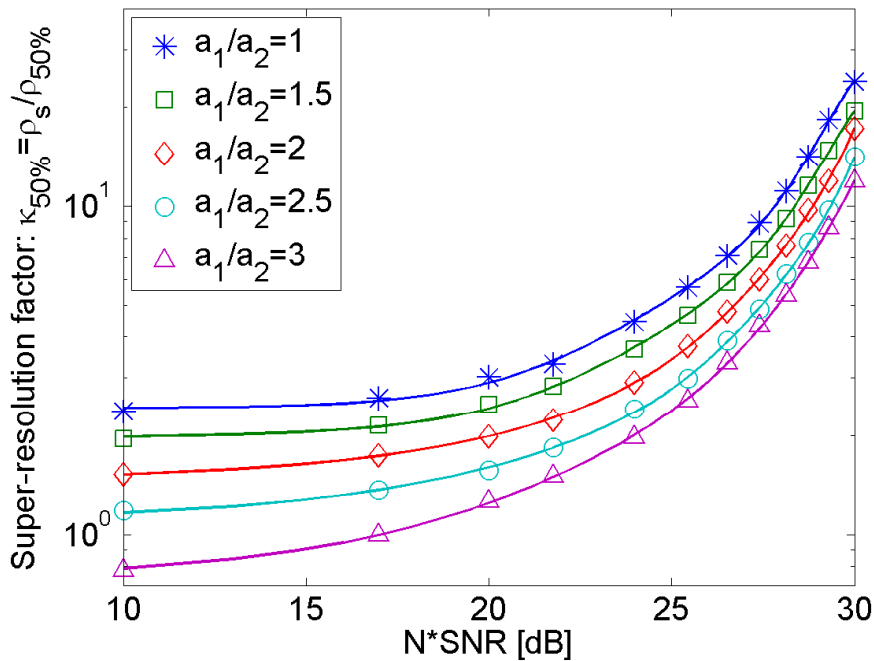
Figure 30. Super-resolution factor averaged over $\Delta\varphi$ as a function of $N \cdot SNR$ under different amplitude ratios a_1/a_2: Experimental results (markers) vs. polynomial fitting (solid lines).

- The SR power of SL1MMER depends asymptotically on the product of N and SNR;
- Irregular sampling does not have a large impact on SR;
- The detection rate varies dramatically with $\Delta\varphi$.

Without loss of generality, a regular sampling in the elevation aperture $[-\Delta b, \Delta b]/2$ is used in the following examples.

Following the above mentioned dependency analysis, the SR factor $\kappa_{50\%}$ is analyzed, which indicates the center point of the detection curve. Since the phase difference $\Delta\varphi$ is a random variable and has significant impact on detection rate, the detection rate curves are averaged over $\Delta\varphi$ before finding the 50% point. The markers in Fig. 30 stand for $\kappa_{50\%} = \rho_s/\rho_{50\%}$ from

the simulations for varying $N \cdot SNR$ within the range of interest (10-30 dB) under different amplitude ratios. It says the achievable super-resolution factors in this range typical for TomoSAR are promising and on the order of $1.5 - 25$. For the readers' convenience the experimental results are fitted to the following polynomial expansion:

$$\kappa_{50\%}(N \cdot SNR) = \sum_{i=1}^{5} c_{SR,i}(N \cdot SNR)^i \tag{40}$$

where $c_{SR,i}$ is the coefficient for the ith order term of $N \cdot SNR$.

	$c_{SR,0}$	$c_{SR,1}$	$c_{SR,2}(10^{-4})$	$c_{SR,3}(10^{-6})$	$c_{SR,4}(10^{-9})$	$c_{SR,5}(10^{-12})$
$a_1/a_2 = 1$	2.4392	−0.0007	0.7116	−0.2013	0.2671	−0.1148
$a_1/a_2 = 1.5$	1.9717	0.0013	0.4374	−0.1197	0.1616	−0.0694
$a_1/a_2 = 2$	1.4691	0.0056	−0.0687	0.0392	−0.0444	0.0223
$a_1/a_2 = 2.5$	1.1108	0.0057	−0.1137	0.0463	−0.0531	0.0256
$a_1/a_2 = 3$	0.7343	0.0055	−0.0496	0.0147	−0.0064	0.0023

Table 2. Polynomial approximation of super-resolution factor as a function of $N \cdot SNR$: coefficients for different amplitude ratios a_1/a_2

Table 2 lists the coefficients for different amplitude ratios a_1/a_2 and the solid lines in Fig. 30 illustrate the results of polynomial fitting. The results in Table 2 can be used as general fundamental bounds for SR of any imaging system.

4.3.3   Required minimum number of acquisitions

As demonstrated in A.5 the performance of tomographic reconstruction deviates from the $N \cdot SNR$ dependency for small N. Yet the required minimal number of acquisitions for robust reconstruction at a given SNR is of great interest since each SAR acquisition is expensive. In (Nannini et al., 2009), the minimal number of tracks for subspace methods for TomoSAR is proposed by assuming a sufficient SNR which is rarely true for the spaceborne case. In A.5, the *robustness* of the algorithm is defined by the minimal required number of acquisitions $N_{min,90\%}$ at a given SNR, which still allows detecting two scatterers with a distance of one Rayleigh resolution unit with $\Delta\varphi = 0$ (the worst case) at a high probability ($P_D \geq 90\%$). At a given SNR, it is found in A.5 that 90% detection rate of two scatterers with a distance of $\rho_s(\alpha = 1)$ can be achieved when the following condition holds:

$$\begin{cases} N_{min,90\%} \cdot SNR \left(2.5 - 1.5\cos\left(2\Delta\varphi + 2\pi\left(1 - \frac{1}{N_{min,90\%}}\right)\right)\right)^2 = c_1 & N_{min,90\%} < 15 \\ N_{min,90\%} \cdot SNR = c_2 & N_{min,90\%} \geq 15 \end{cases} \tag{41}$$

c_1 and c_2 are constants which depend on the amplitude ratio and phase difference between two scatterers.

In Fig. 31, the required minimal number of acquisitions $N_{min,90\%}$ obtained from the experiments for different amplitude ratios ($a_1 = a_2$ (left), $a_1 = 2a_2$ (right)) and for $\Delta\varphi = 0$ is presented. The circles show the 90% detection rate positions with $N \in [8, 20]$. E.g. for a total SNR of 6 dB, only 11 acquisitions are required for $a_1 = a_2$ ($SNR = 3/3$ dB respectively) while 17 acquisitions are needed for $a_1 = 2a_2$ ($SNR = 5/ - 1$ dB respectively). The dashed line ($N < 15$) & solid line ($N \geq 15$) in Fig. 31 show the estimated $\hat{N}_{min,90\%}$ from eq. (41) with estimated $\hat{c}_1$ and $\hat{c}_2$

for $a_1 = a_2$ and $a_1 = 2a_2$ listed in Table 3. For $N \geq 15$ the N and SNR can be substituted by $N \cdot SNR = const.$

(a) $a_1 = a_2$

(b) $a_1 = 2a_2$
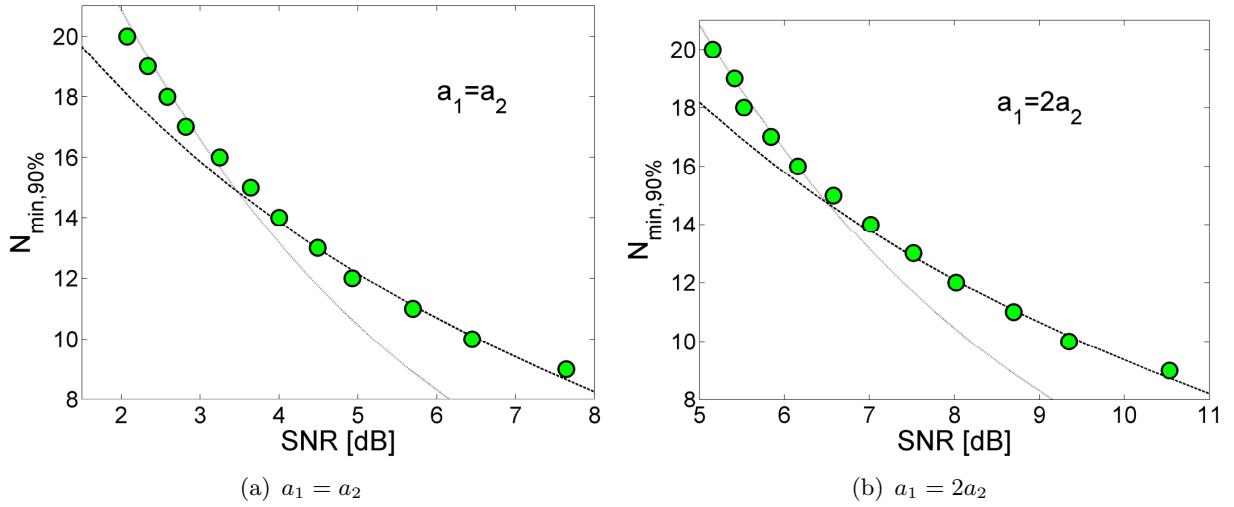
Figure 31. Required minimal number of acquisitions $N_{min,90\%}$ for robust reconstruction (i.e. well detecting two scatterers ($P_D \geq 90\%$) with a distance of one Rayleigh resolution unit) at a given SNR: theory (dashed line ($N < 15$) and solid line ($N \geq 15$)) vs. experiment (circles). The worst case $\Delta\varphi = 0$ is assumed.

	c_1	c_2
$a_1/a_2 = 1$	54.2	33.1
$a_1/a_2 = 2$	107.3	65.9

Table 3. Value of c_1 and c_2 under different amplitude ratios between the two scatterers ($\Delta\varphi = 0$, the worst case)

4.4   Multi-component nonlinear motion estimation — the time warp method

Conventional D-TomoSAR uses multi-baseline, multi-temporal SAR acquisitions for reconstructing the 3-D distribution of scatterers and their motion (section 2.1.3). The elevation estimation of a scatterer exploits the linear relationship between the measured phase and the product of baseline and elevation, and hence the full arsenal of spectral estimation methods can be applied. Motion, however, is often nonlinear (periodic, accelerating, stepwise, etc.), and hence does not fit well into the spectral analysis framework.

This section summarizes the work in A.6 in which a nonlinear warping of the time axis is introduced to accommodate nonlinear motion models. The method is extended to multi-component motion and validated with both simulated and real data. As hinted in section 4.1, the TerraSAR-X stack used in this thesis undergoes different motions in the two sub-stacks (Fig. 13):

- *The "asterisk" sub-stack*: No long-term motion has been observed in the test area, i.e. motion-induced phase is only induced by periodic thermal dilation.
- *The "diamond" sub-stack*: Since July 2009, the area shown in Fig. 32 has undergone a pronounced subsidence centered at the convention center which can be seen from the differential interferogram in Fig. 33 generated by two images taken on April 17, 2009 (master) and April 4, 2010 (slave). Hence, the "diamond" sub-stack is characterized by a multi-component nonlinear motion, i.e. a combined linear and thermal dilation induced seasonal motion.

Figure 32. Mean intensity map of the TerraSAR-X spotlight data stack of the test area surrounding the convention center, Las Vegas, Nevada, USA.

This motion is used to validate the generalized time warp method for single-component and multi-component nonlinear motion estimation, respectively. Note that since the demonstration of the time warp method is independent of the tomographic (spectral estimation) method itself, the simple linear reconstruction SVD-Wiener is used throughout this section.

4.4.1   Generalized D-TomoSAR system model

The displacement relative to the master acquisition $d(s, t_n)$ in eq. (8) may be modeled using a linear combination of M basis functions $\tau_m(t_n)$:

$$d(s, t_n) = \sum_{m=1}^{M} p_m(s)\tau_m(t_n) \tag{42}$$

where $p_m(s)$ is the corresponding motion coefficient to be estimated. The choice of the basis functions depends on the underlying physical motion processes. For example, long-term geo-dynamic processes may lead to a linear, accelerating or decelerating motion; instantaneous geodynamic events, e.g. earthquake and volcano eruption, result in a stepwise motion; thermal expansion causes a motion correlated to the temperature etc. Linear, seasonal and thermal motion will be addressed in the next section. Alternatively, without a priori information about the deformation mechanisms, the motion may be approximated by a sum of polynomials.

4.4.2   The time warp method

An approach is briefly introduced by the author in (Zhu and Bamler, 2009) to cope with the problem of D-TomoSAR reconstruction with single-component nonlinear motion by rearranging the acquisitions on the time axis (the so-called "time warp").
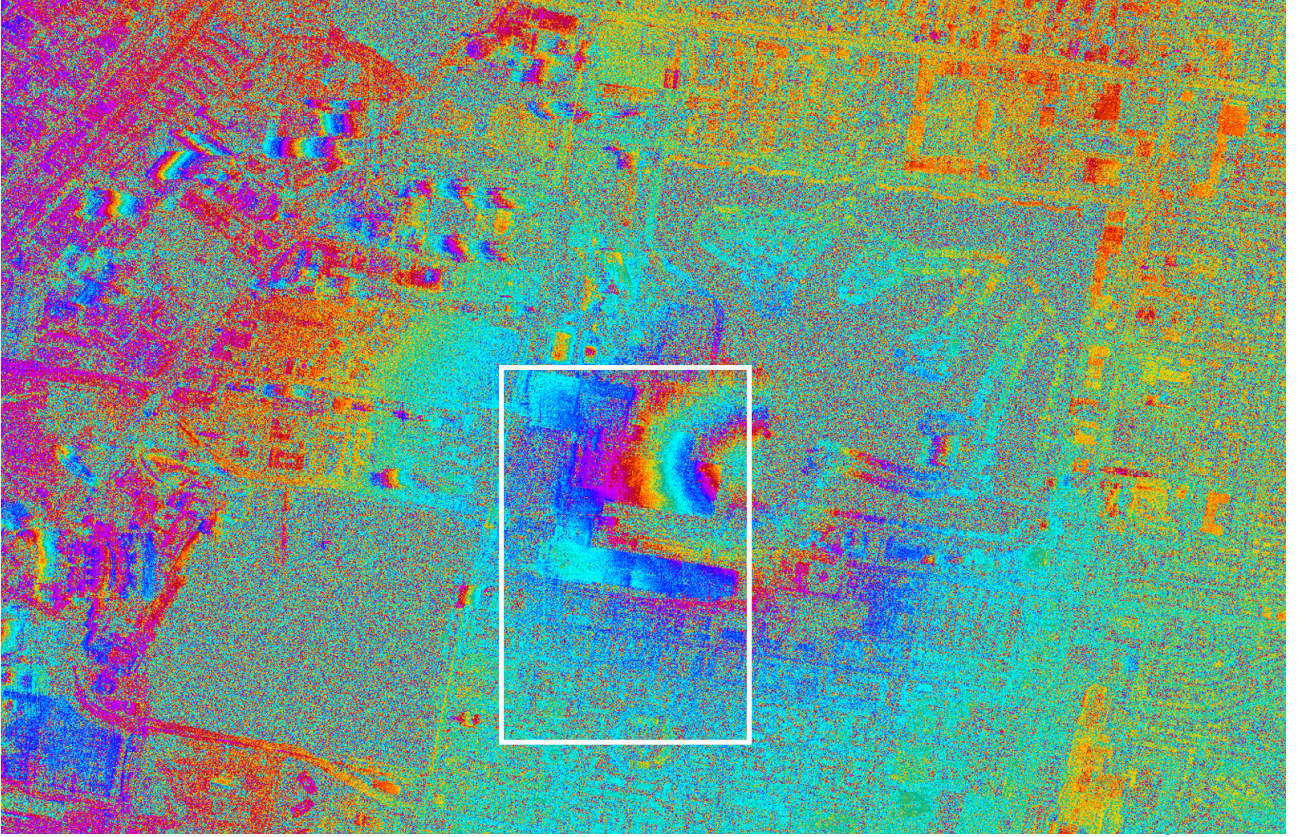
Figure 33. Interferogram of the test site showing a pronounced circular subsidence pattern. The interferogram has been generated by two TerraSAR-X images taken from the "diamond" sub-stack of Fig. 13. white box indicated the investigation area of Fig. 37.

For $M = 1$, by introducing the temporal frequency $\eta_{1,n} = 2\tau_{1,n}(t_n)/\lambda$ as a function of a motion base function $\tau_{1,n}(t_n)$, and the motion coefficient p_1, the proposed time warp method leads to a generalized system description which can be adapted for different nonlinear motion models:

$$g_n = \int\limits_{\Delta p_1} \int\limits_{\Delta s} \gamma(s)\, \delta\left(p_1 - p_1(s)\right) exp\left(-j2\pi\left(\xi_n s + \eta_{1,n}p_1\right)\right) \mathrm{d}s\mathrm{d}p_1 \qquad (43)$$

The temporal frequency $\eta_{1,n}$ requires a warping, or resorting of the acquisition time from t_n to $\tau_{1,n}$. The time warp operation rewrites the D-TomoSAR model with single-component motion (linear or nonlinear) as a standard 2-D spectral estimation problem which makes all spectral estimators applicable. This principle is illustrated for a periodic motion model in Fig. 34. The most common motion base functions are:

- Linear motion: $\tau_{1,n} = t_n$ and the coefficient $p_1(s)$ stands for the LOS velocity (v) as a function of s.
- Seasonal motion: $\tau_{1,n} = sin\left(2\pi(t_n - t_0)\right)$ and $p_1(s)$ stands for the amplitude (a) of the periodic motion; t_0 is the initial phase offset.
- Thermal dilation: $\tau_{1,n} = T(t_n)$, i.e. the temperature, and $p_1(s)$ is the scaling factor of the thermal dilation along s.

The "asterisk" sub-stack is used to estimate the thermal dilation induced motion. Theoretically speaking, the temperature history at the acquisition times should be chosen as the motion basis function and the scaling factor should be estimated. However, due to the lack of ground truth, a simple sine-function with a period of one year is used here. The initial offset is estimated to be $t_0 = 0.013y$ by fitting the sine model to the monthly average temperature of the test site in the years of 2008 to 2010. By applying the time warp method to this stack, the number of scatterers map, topography and amplitude of seasonal motion can be obtained for each azimuth-range
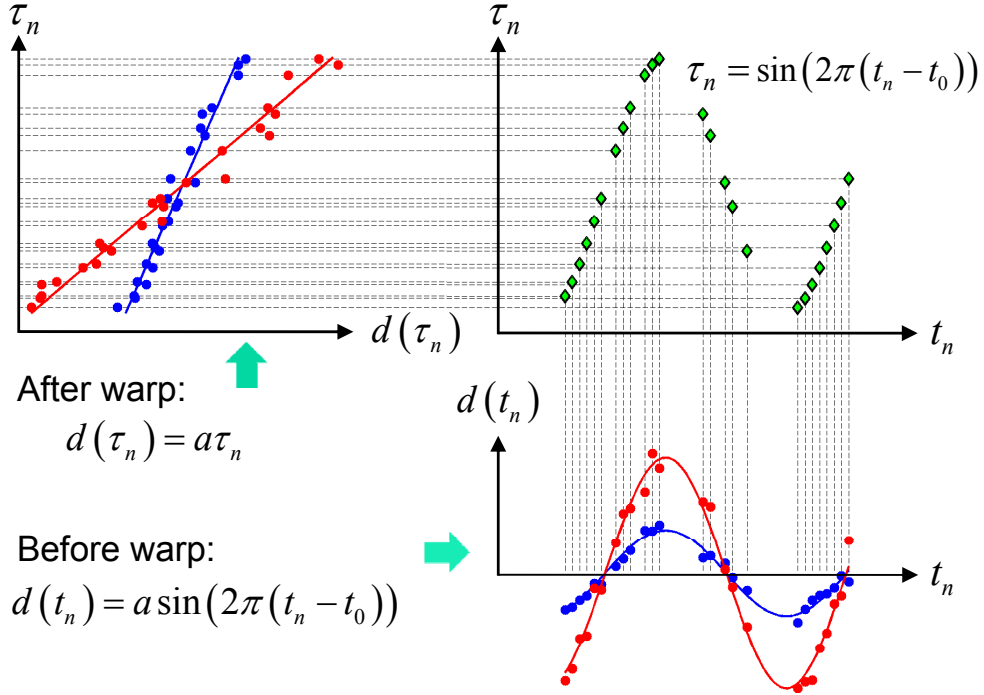
Figure 34. The time warp converts seasonal periodic motions with different amplitudes (red and blue) into linear functions of different slopes.

pixel among them the topography estimates of the test area in Fig. 1 have been presented in section 4.1. Fig. 35 exhibits the amplitude of the seasonal motion for the detected single and double scatterers (see Fig. 14). Some properties of the thermal dilation can be observed:

- For the high rise buildings, the amplitude of the seasonal motion increases with increasing height;
- The ground structures shows almost no seasonal motion;
- Some low buildings have a metallic structure, e.g. the shopping mall on the lower right corner, and are more severely affected by thermal dilation than the surrounding infrastructure.

4.4.3   "Let's do the time warp again"

The generalization of the time warp method for $M > 1$ is straightforward. Let us define the mth temporal frequency component at t_n as $\eta_{m,n} = 2\tau_m(t_n)/\lambda$. Then eq. (8) can be rewritten as an $M + 1$ dimensional Fourier transform of $\gamma(s)\delta\left(p_1 - p_1(s), p_2 - p_2(s), \cdots, p_M - p_M(s)\right)$ which is a delta-line in the $M + 1$ elevation-motion parameter space:

$$
g_n = \int\limits_{\Delta p_M} \cdots \int\limits_{\Delta p_1} \int\limits_{\Delta s} \gamma(s)\ \delta\left(p_1 - p_1(s), p_2 - p_2(s), \cdots, p_M - p_M(s)\right) \cdot
$$
$$
exp\left(-j2\pi(\xi_n s + \eta_{1,n}p_1 + \cdots + \eta_{M,n}p_M)\right) \mathrm{d}s\mathrm{d}p_1\cdots\mathrm{d}p_M \qquad (44)
$$

This extension is a general solution to the multi-component nonlinear motion estimation problem for D-TomoSAR, and hence completes the time warp concept for all possible complex motion models.

A multi-component motion reconstruction example using SVD-Wiener, i.e. combined linear and seasonal motion, with two scatterers inside a SAR pixel, is presented here to illustrate this concept. A realistic SNR of 3 dB and a baseline configuration of the TerraSAR-X "asterisk"-stack with 30 acquisitions used in chapter 4.1 (see Fig. 13) are simulated. The two scatterers
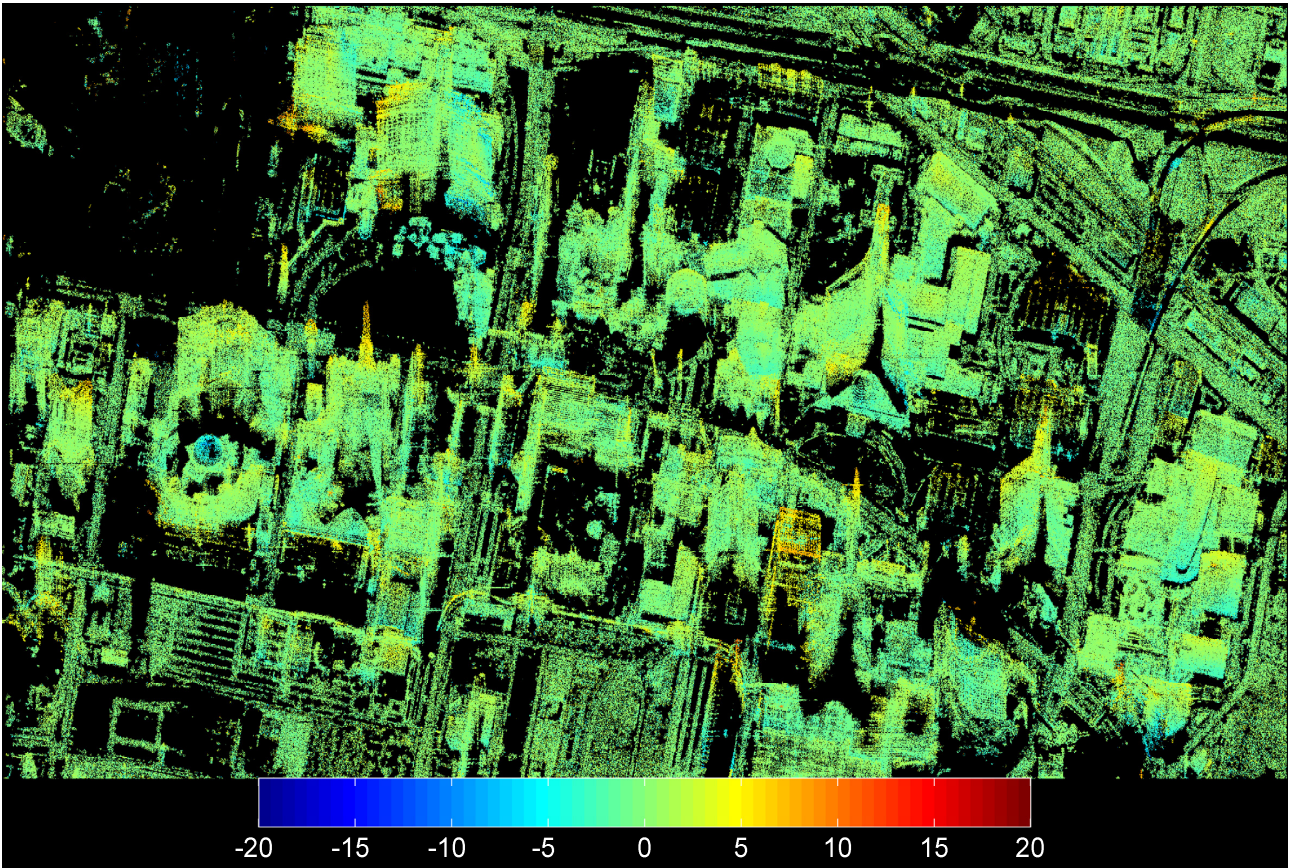
Figure 35. Estimated amplitude of thermal dilation caused seasonal motion of the detected single and double scatterers [unit: millimeters].
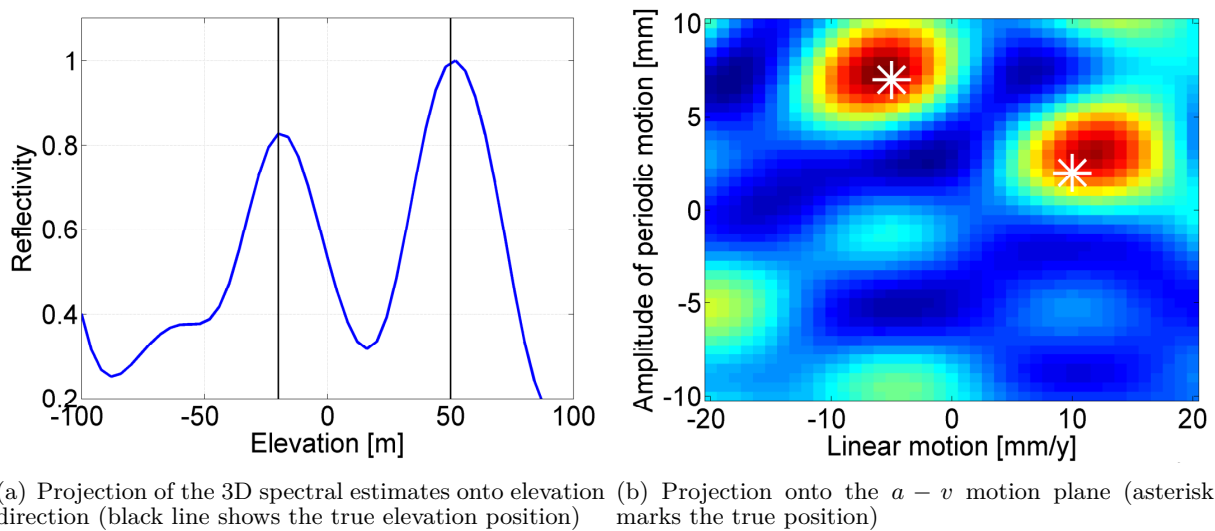
(a) Projection of the 3D spectral estimates onto elevation direction (black line shows the true elevation position)

(b) Projection onto the $a - v$ motion plane (asterisk marks the true position)

Figure 36. Multi-component motion reconstruction example with linear and seasonal periodic motion using simulated data: two scatterers inside the cell with elevations of $-20/50$ m, linear motion of $10/-5$ mm/y and seasonal motion amplitudes of $2/7$ mm ($SNR = 3$ dB).

are located at -20 and 50 m (ca. 2 Rayleigh elevation resolution), with linear motion of 10 and -5 mm/y and seasonal motion with amplitudes of 2 and 7 mm, respectively. In this case, $M = 2$, therefore, it is a 3-D spectral estimation problem. Fig. 36 presents a projection of the 3-D spectral estimates onto the elevation direction, i.e. a sum over the linear and periodic motion plane. The right plot shows the corresponding projection in the amplitude-velocity ($a-v$) plane and the white asterisk indicates the true amplitude and velocity. Comparing the estimates to the simulation truth, it demonstrates that the generalized time warp method can give robust
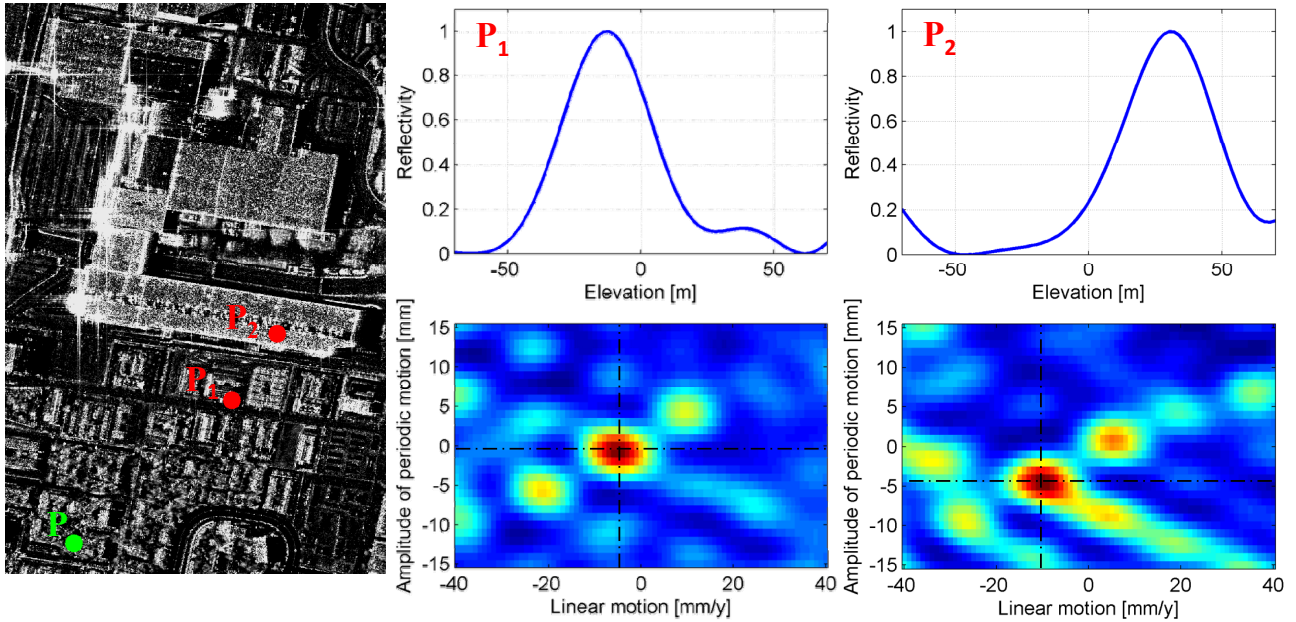
Figure 37. Validation examples of the generalized time warp approach. Left: TerraSAR-X intensity map. Middle: reflectivity profile (upper) and scatterer distribution in $(a - v)$ motion plane (bottom) of analyzed ground pixel P_1; Right: the same plots of analyzed roof pixel P_2. P is the reference point. Dashed lines mark the location of the maxima in the spectral estimates.

estimation of multi-component nonlinear motion even under moderate SNR.

In a real data experiment, the generalized time warp method for $M = 2$ is applied to the "diamond" sub-stack of Fig. 13 by choosing linear and seasonal motion as the basis functions. Fig. 37, left, is the TerraSAR-X intensity map of the region of interest (marked by a box in Fig. 33). According to Fig. 33, the center of the subsidence pattern, i.e. the "epicenter", is located on the right upper part of the intensity map. Therefore, together with the seasonal motion results shown in Fig. 35, one can expect that: 1) only the building structures suffer from thermal dilation; 2) the closer to the "epicenter" the point is, the bigger is the linear subsidence.

To validate the proposed generalized time warp method, the two pixels P_1 and P_2 marked by red points are selected and will be analyzed in the following (see the left image of Fig. 37: P is the reference point; all the estimates are relative to P). As P_1 is located outside the region of the convention center, it is expected to contain only a single component linear motion. P_2 is located on the roof and is closer to the "epicenter". Hence, a combined seasonal and more significant linear motion are expected. The middle and right image pairs in Fig. 37 are the corresponding estimated reflectivity profiles (upper) and scatterer distributions in $(a - v)$ motion plane (bottom) of the analyzed ground pixel P_1 and roof pixel P_2, respectively. The estimates reveal a single scatterer in the pixel P_1 with an elevation of ≈ 12 m (≈ 6.3 m in height relative to reference pixel), almost no seasonal motion and a subsidence of -5 mm/y. P_2 is also estimated to be a single scatterer, with an elevation of 32 m (≈ 16.9 m in height), a seasonal motion with absolute amplitude of 4.5 mm and a subsidence of -10 mm/y. These results are consistent with the expectation and, hence, demonstrate the capability of the time warp method for multi-component nonlinear motion estimation.

5 A practical demonstration of the super-resolution of SL1MMER

In section 4.2.2, the SR capability of the proposed SL1MMER algorithm has been illustrated by an example using simulated data. Further practical super-resolving examples by reconstructing some selected pixels are provided in A.4 and A.5. So far, no substantial real data example has been presented. This chapter presents a practical demonstration of the SR of SL1MMER for SAR tomographic reconstruction.

5.1   Test building: Bellagio hotel

For the same reason explained in section 4.1, the "asterisk" sub-stack (Fig. 13) consisting of 30 spotlight TerraSAR-X images with an elevation resolution ρ_s of 40.5 m, i.e. ca. 23.7 m in height, is used in this chapter. In particular, the Las Vegas Bellagio hotel has been chosen as a test building to demonstrate the SR of SL1MMER since 1) From the tomographic result of the non-parametric SVR-Wiener shown in Fig. 20, this building has severe layover effects; 2) The surrounding infrastructure has a metallic structure, therefore the reflections from both the building façade and ground structures are strong. It is expected that the SR of the algorithm will be observed when reconstructing the layover area where the lower part of the building and the ground infrastructures are mapped together.

(a) Optical image (©Google)   (b) TerraSAR-X mean intensity map
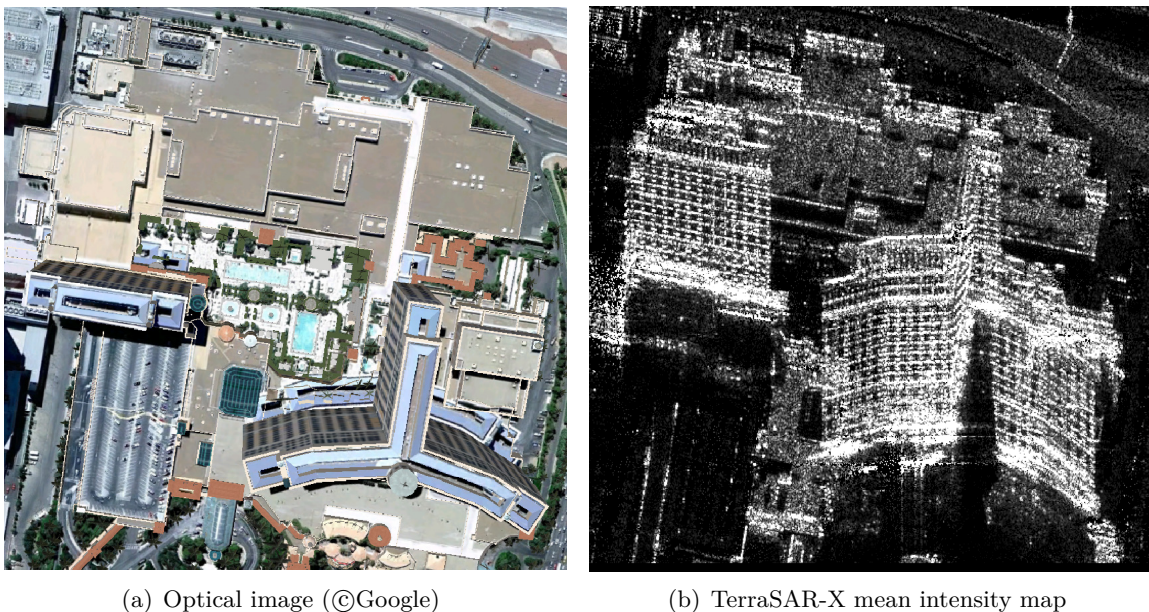
Figure 38. Test building: Bellagio hotel

The left image of Fig. 38 shows the optical view of the Bellagio hotel in Las Vegas with a height close to 125 m, corresponding to an elevation range of 237 m. The right image in Fig. 38 is the corresponding TerraSAR-X mean intensity image. Compared to the optical image, many noteworthy features of the SAR image can be pointed out: 1) The folding of the building towards the sensor due to the layover phenomenon (the base of the Bellagio hotel is almost horizontally aligned in the optical and TerraSAR images) and 2) Due to the weaker backscattering, ground infrastructure visible in the optical image is completely "hidden" by the strong returns from the building façade in the SAR image.

(a) Detected single scatterers   (b) Detected single and double scatterers
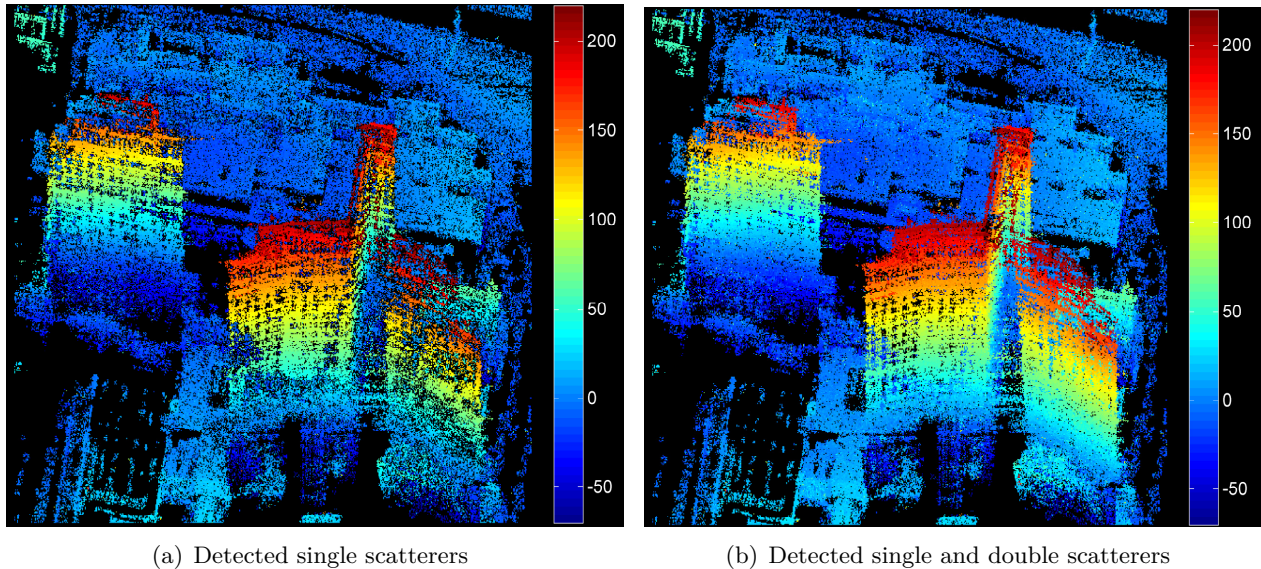
Figure 39. The same result as Fig. 19 but estimated by SL1MMER, i.e. topography estimates in meters. Left: detected single scatterers; right: detected single and double scatterers (for double scatterers the higher one (façade) is displayed).

The goal is to separate the interfering layers associated with the ground and the façade of the building and reconstruct the 3-D and 4-D maps of this area.

5.2   Experimental results

As depicted in section 4.4.2, within the acquisition period of this sub-stack, no long-term motion has been observed in the test area, i.e. motion-induced phase is only caused by periodic thermal dilation. The D-TomoSAR system model with a time warp assuming a seasonal motion ($t_0 = 0.013$ year) is used. The SL1MMER algorithm is applied to each pixel of the test area and the same penalized likelihood criterion as applied in MD (section 4.1) is used, i.e. BIC with a penalty term $C(K) = 0.5\ k \ln N = 2K \ln N$. The number of scatterers map, elevation and amplitude of seasonal motion for each of the detected single and double scatterers are then obtained. Since the goal is to demonstrate the SR of SL1MMER, the discussion will be focused on the topography estimates since the shape of the building can provide a plausible "ground truth" for the estimates.

The left image of Fig. 39 presents the topography estimates, i.e. estimated elevations, of the detected single scatterers while the corresponding fused topography estimates with both of the detected single scatterers and the top layer of the detected double scatterers are shown in the right image. The information increment is more significant than with MD shown in Fig. 19.

Fig. 40 presents the topography estimates of the two layers of the detected double scatterers, i.e. a top layer mostly caused by reflections from the façade of the high rise building and a ground layer caused by reflections from lower buildings and ground infrastructures. Again, the gradation of elevation estimates on the top layer (see left image of Fig. 40) and the homogeneity on the ground layer (see right image of Fig. 40) suggest the correctness of the elevation estimation and the layover separation capability. It is interesting to observe that the full structure of the high rise building is almost captured even only with the detected double scatterers. In addition, "hidden" ground infrastructures are now "visible" again.

In Fig. 41 the reconstructed topography of the ground infrastructures (right) is interpreted with the assistance of the optical image (left):

- There are two blocks on the ground layer (left top and right bottom, respectively) that show brighter blue. This indicates higher topography at that area and this is consistent

(a) Building façade   (b) Ground structures
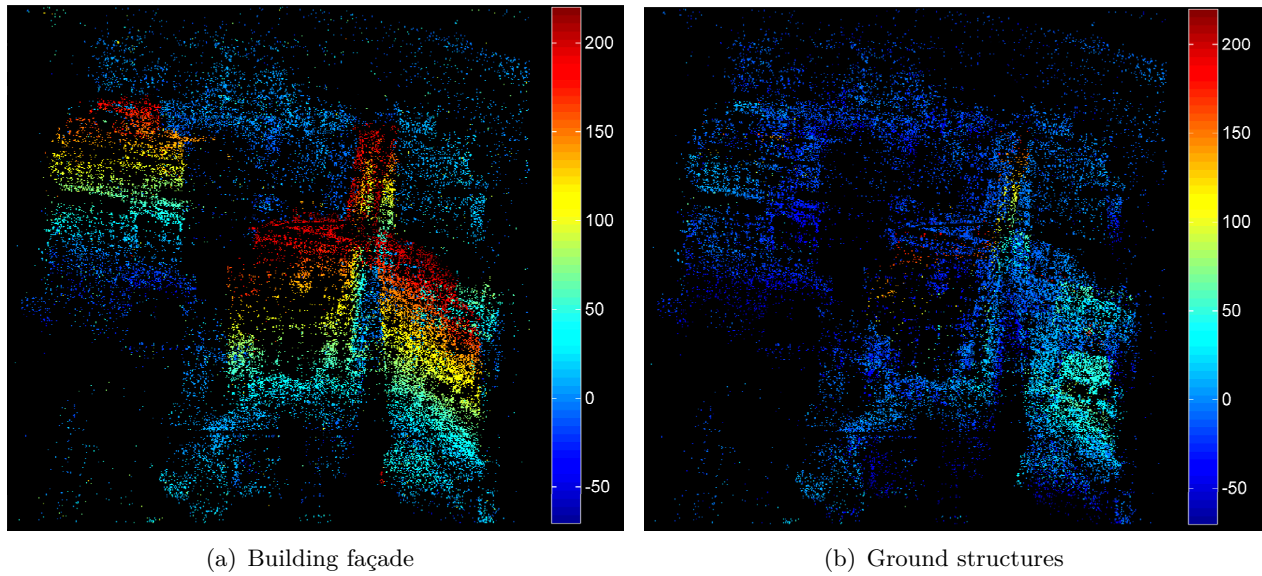
Figure 40. Topography estimates of the separated double scatterers - the same result as Fig. 20 but estimated by SL1MMER.

(a) Optical image (©Google)   (b) Layovered ground infrastructure reconstructed by SL1MMER (lower layer - one of the detected double scatterers)
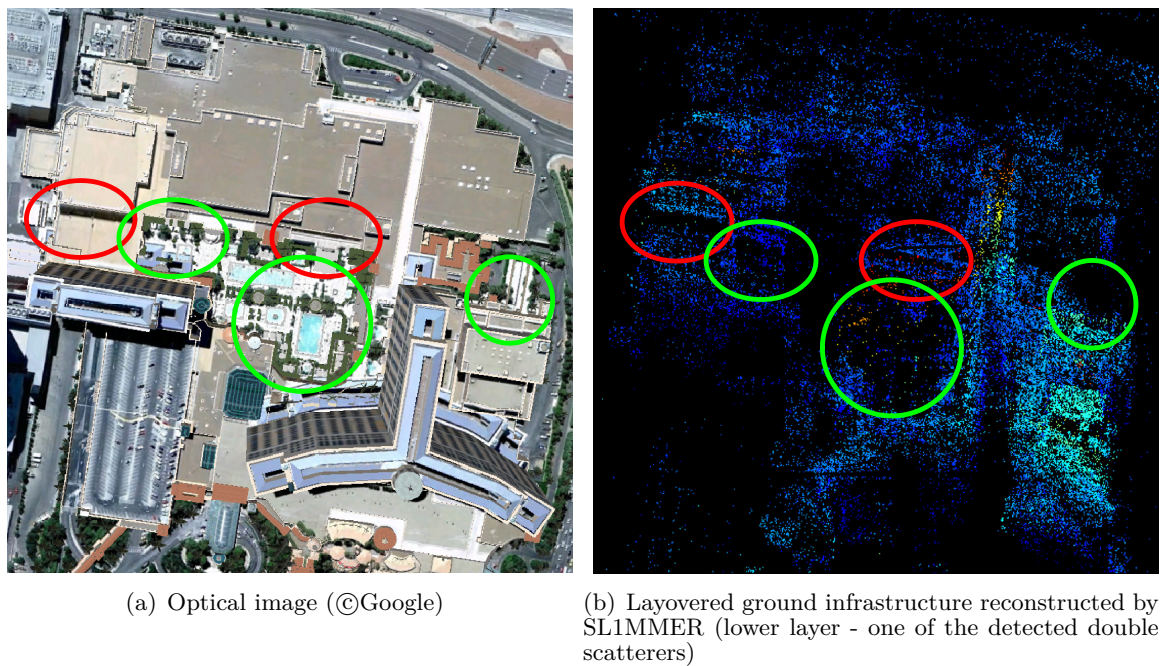
Figure 41. Layovered ground infrastructure: Optical image (©Google) vs. reconstruction from layover separation obtained by SL1MMER. The red circles: shadowing areas; the green circle: vegetation areas.

with the 3-D building model provided by Google-earth.

- Although most of the "hidden" ground infrastructure is retrieved, there are still some areas (e.g. the ones marked by circles) showing homogenous black color, i.e. no layovered coherent object on the ground layer. By comparison with the optical image, it can be easily found that the areas marked by red circles are shadowing areas while the areas marked by green circles are vegetation areas, i.e. no coherent return. This fact confirms our estimates.

- Some pixels at the area near the top of the building have very large topography estimates showing some regular structures even for the lower layer. It seems that both of the detected scatterers are located on the building façade. This might be caused by the complicated structures on the top of the Bellagio hotel. However, to verify it, a more precise 3-D model of the building is required.

5.3 Comparison with linear detector — maxima detection

In this section, the results obtained by using the SL1MMER algorithm will be compared to the result of maxima detection (MD) shown in section 4.1 (i.e. SVD-Wiener linear reconstruction followed by model order selection).

Fig. 42 presents the number of scatterers map obtained by MD (left) and SL1MMER (right) over the test area where blue indicates zero scatterers inside the azimuth-range pixel, green stands for one and red for two. Non-parametric estimators can only detect two scatterers with an elevation distance larger than approximately the Rayleigh elevation resolution unit ρ_s (i.e. 40.5 m in elevation, ca. 23.7 m in height). Therefore, it is not surprising that the double scatterers detected by MD are mainly located on the upper part of the building area. The result of SL1MMER shows a much denser red color which indicates a larger amount of detected double scatterers. In the following discussion, *the percentage of detected double scatterers* is with respect to the sum of all detected scatterers, i.e. single plus double (note a pixel including double scatterers is counted twice). For the whole area, 29.9% and 13.1% of the scatterers detected by SL1MMER and MD, respectively, are double scatterers. In particular, for an individual building, i.e. the skyscraper on the right, SL1MMER increases the percentage of the detected double scatterers from 20.4% (MD, see section 4.1) to 37.8%.
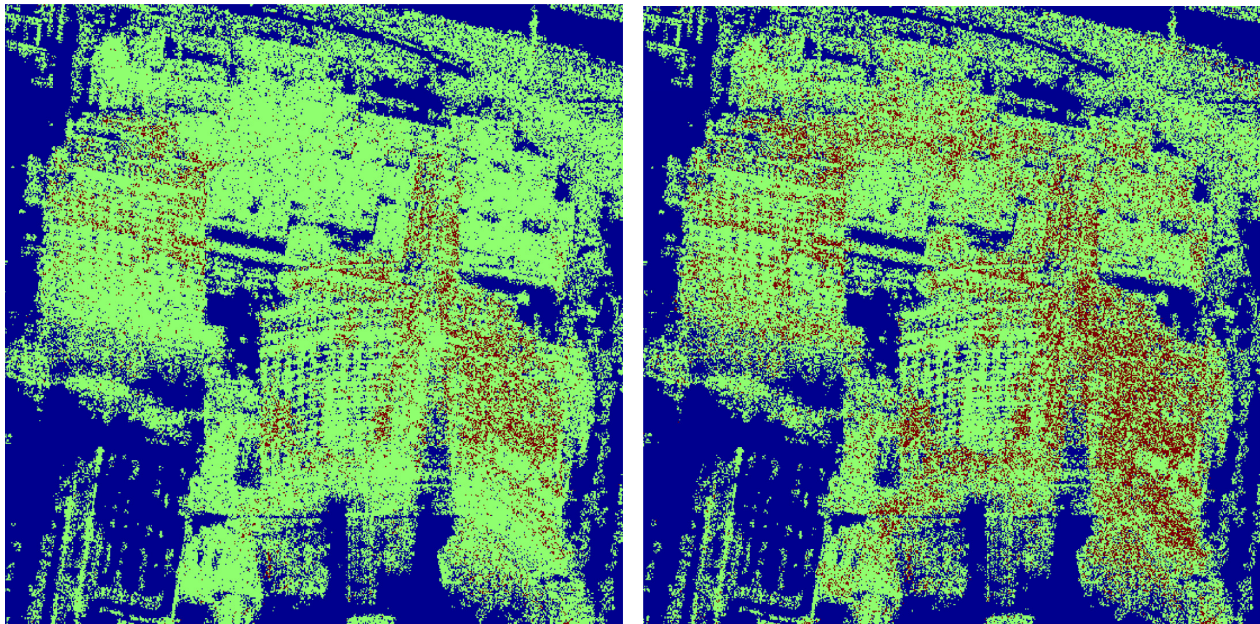
(a) MD (SVD-Wiener plus model selection)   (b) SL1MMER

Figure 42. Number of scatterers map obtained by MD and SL1MMER at the test area. Blue: zero scatterers inside the azimuth-range pixel; green: one scatterer inside the pixel; red: two scatterers inside the pixel. For the whole area, 29.9% and 13.1% of the scatterers detected by SL1MMER and MD, respectively, are double scatterers. In particular, for an individual building, i.e. the skyscraper on the right, SL1MMER increases the percentage of the detected double scatterers from 20.4% (MD, see section 4.1) to 37.8%.

The dramatically improved layover separation capability is mainly associated with the SR power of SL1MMER. E.g. in Fig. 43, the top layer (building façade) of the double scatterers detected by MD and SL1MMER is illustrated. Besides the aforementioned information increment within the layover area in general, the marked areas are worthy of special attention. The magenta boxes mark the areas where the lower parts of the hotel façade and the ground structures are mapped together while the yellow box marks an area where the lower ground structures (of different height) are layovered. Those are the typical areas where SR is required for layover

(a) Building façade: MD   (b) Building façade: SL1MMER
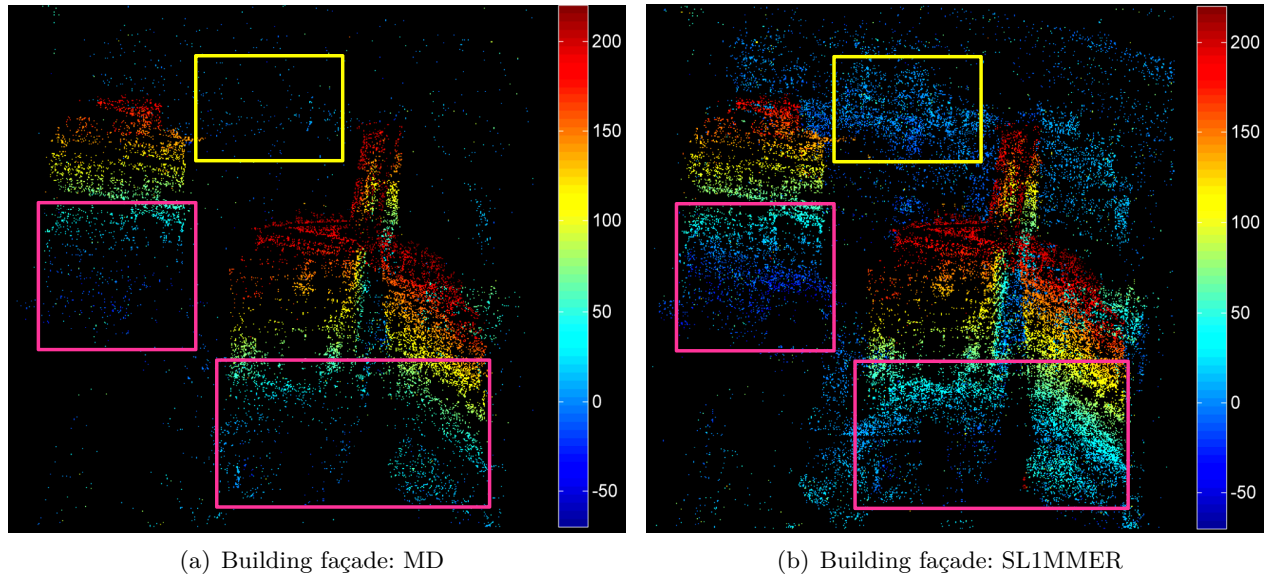
Figure 43. MD vs. SL1MMER: topography estimates of the separated double scatterers - one of the two from the building façade. Magenta boxes: areas where the low part of the hotel and the ground structures are mapped together; Yellow box: area where the ground structures (of different height) are mapped together. In both cases, SL1MMER detects a big amount of double scatterers while non-parametric MD can not well separate them.

separation since the elevation distances between the two scatterers are small. In all marked areas, SL1MMER detects a large amount of double scatterers while non-parametric MD can not separate them well.

In order to further quantify the SR capability of SL1MMER, the histogram of the elevation distances between each of the detected double scatterer pairs is provided in Fig. 44. The blue curve represents the result of SL1MMER while the red curve shows the results of MD. In this plot, 135 bins with a elevation spacing of 2 m are used for producing the histogram. The horizontal axis, i.e. the elevation distance between the detected double scatterers is normalized to the Rayleigh resolution unit ρ_s. Theoretically, the case of two scatterers within one Rayleigh resolution, i.e. the normalized elevation distance $\alpha < 1$, is relevant for SR. In our case, the L_2 norm regularization introduced in MD gives also SR though non-significant. Therefore, for a fair comparison, the 3 dB width of the point response function (PRF) in elevation (the black line) is chosen as the watershed. I.e. only on the left hand side of the black line, the two scatterers are super-resolved.

Fig. 44 demonstrates:

- SL1MMER has impressive SR capability, i.e. many of double scatterers with $\alpha < 1$ are detected;
- The layover separation capability of SL1MMER is much higher compared to MD, i.e. much more double scatterers are detected by SL1MMER compared to MD, mainly contributed to SR (green zone);
- For $\alpha > 1$, SL1MMER also provides remarkably better layover separation performance (pink zone).

5.4   Discussion — the role of super-resolution for urban infrastructure monitoring

The histogram in Fig. 44 gives us a strong hint that there are *much more* double scatterers with *smaller* elevation distances and the amount of double scatterers decays strongly with the distance. This leads eventually to a discussion of the role of SR in tomographic SAR inversion
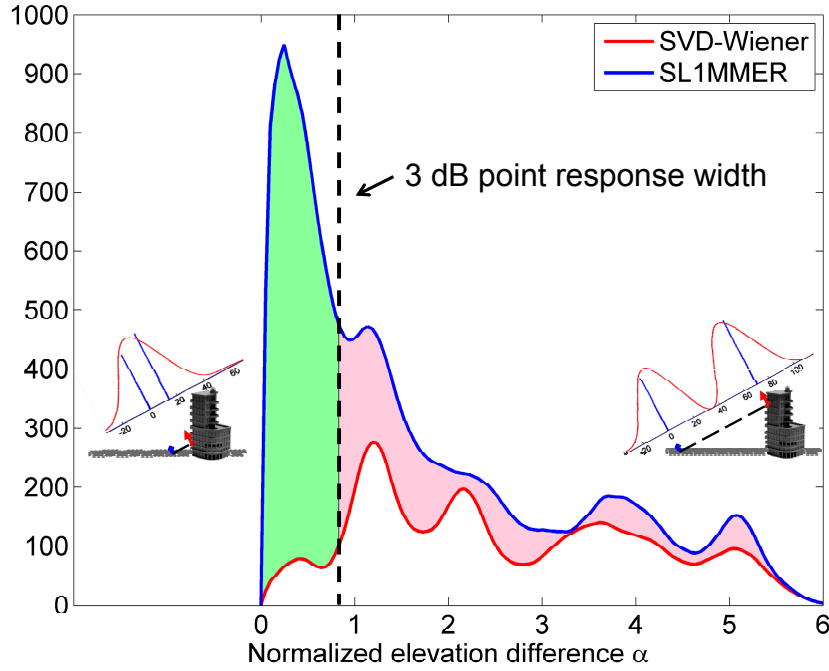
Figure 44. SL1MMER detects much more double scatterers than MD: histogram of the distance between the detected double scatterers using MD (red) and SL1MMER (blue) (135 bins, distance spacing: 2 m). The elevation distance is normalized to the Rayleigh resolution unit ρ_s. Note since the L_2 norm regularization gives also slight SR, the black line marks the 3 dB width of the point response function in elevation. The green area is the information increment contributed by SR of SL1MMER.

for urban infrastructure monitoring by addressing the following two questions:

1) Why are there much more double scatterers with smaller elevation distances?
   The layover phenomena in a SAR image of an urban area is mainly caused by the following two scenarios:

   - Buildings with different heights layovered with the ground:
     As sketched in Fig. 45, the layover caused by the taller building and the lower building both covers the range with smaller elevation distances. However, only the layover area of the taller building cover the range with larger elevation distances.
   - Taller building layovered with the ground and the roof of the lower building:
     As sketched in Fig. 46, the layover of the taller building and lower building also leads to smaller elevation distances.

   From these purely geometric considerations, double scatterer pairs with smaller elevation distances are more frequent than with larger distances. If we assume that the heights of the buildings follow a uniform distribution, the elevation distance between the double scatterer pairs should follow approximately a *logarithmic* law:

$$p(\delta s) = \frac{\sin \theta}{h_{max}} \left(\ln(h_{max}) - \ln(\delta s \ \sin \theta) \right) \quad \text{for} \ \ \delta s = 0, \ldots, \frac{h_{max}}{\sin \theta} \tag{45}$$

   where h_{max} is maximum building height and $p(\delta s)$ is the probability density function of the elevation distance between two scatterers δs

2) How much more information can be obtained in practice?
   In the test area shown in Fig. 38, the SL1MMER algorithm which can provide SR detected 29.9% double scatterers pairs while 45.5% of them are within the Rayleigh resolution unit, even in this test site, where most of the buildings are high rise buildings. Therefore, according to eq. (45), it is expect that more than half of the double scatterer pairs have an elevation distance smaller than the Rayleigh resolution unit. The above discussion makes

Figure 45. Buildings with different heights layovered with the ground. Double scatterer pairs with small distances are present at both buildings, and hence are more frequent.
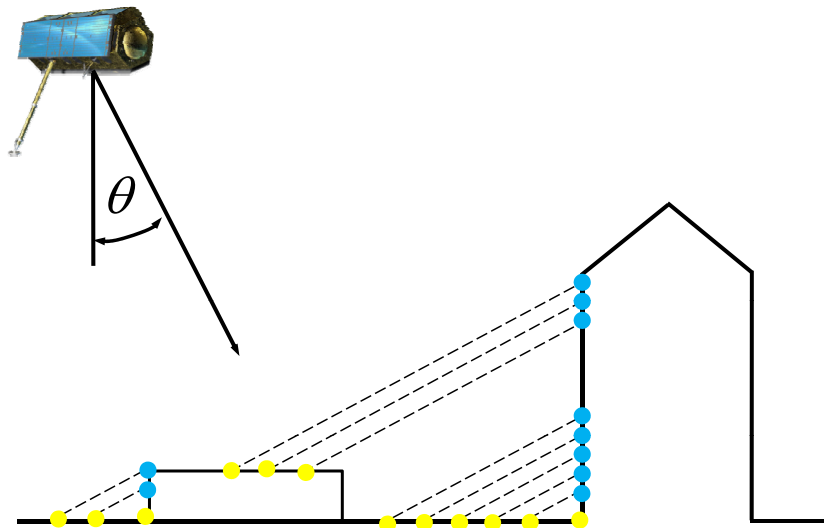
Figure 46. Higher building layovered with the ground and the roof of the lower building. As in Fig. 45, double scatterer pairs are more probable with small elevation distances.

it obvious that SR is even more important than already assumed.

The practical demonstration of SR of SL1MMER presented in this Chapter completes the methodological investigation in this work concerning new tomographic SAR inversion algorithms tailored to VHR SAR data. Moreover, it also demonstrates that super-resolution is a very crucial requirement for VHR tomographic SAR inversion for urban infrastructure monitoring.

6   Conclusion and outlook

6.1   Conclusion

The aim of this thesis is the optimum exploitation of tomographic VHR SAR data stacks for urban infrastructure mapping and monitoring with the vision of

"A dynamic city model derived from TomoSAR and D-TomoSAR showing the shape and the deformation of each building".

In order to realize this vision, four objectives, including one general and three methodology objectives, summarized in Chapter 1, have been pursued. With reference to the work presented in this thesis, the following conclusions can be drawn:

- VHR tomographic SAR inversion is able to reconstruct the shape and motion of individual buildings and city areas by providing:

  Number of scatterers. The incoherent pixels (zero scatterers) and layover areas (two scatterers) of a SAR image can be clearly identified and distinguished from the coherent but non-layover areas. This also assists those applications requiring a single dominating scatterer inside a pixel, like PSI.

  Topography and motion maps of single scatterers. Except from the layover areas, the full 4-D (3-D location and motion) structure of the high rise city area is captured at a very detailed level. Those are the standard products expected from PSI processing. On the one hand, tomographic SAR inversion possesses the capability of PSI, i.e. accurate 3-D localization of PSs and long-term motion monitoring. On the other hand, tomographic SAR inversion does not require the pre-selection of the long-term stable pixels, i.e. the PSs, and hence it can provide reference results for evaluating PS selection algorithms.

  Reflectivity, topography and motion maps of detected double scatterers. Information for layover areas is retrieved. Promising layover separation results are obtained. For individual buildings, a high proportion of double scatterers — up to 38% — are detected. The amount of information increment for individual buildings demonstrates the necessity of tomographic reconstruction for urban infrastructure monitoring.

- The motion or deformation of buildings is often nonlinear (periodic, accelerating, stepwise, etc.). This is particularly true with VHR SAR data. Multi-component nonlinear motion of multiple scatterers can be separated and further estimated by tomographic reconstruction using the newly developed generalized time warp method.

- Super-resolution is crucial for VHR tomographic SAR inversion for urban infrastructure. Without super-resolution, large parts of the layover area are not accessible or are wrongly estimated.

  This attributes to the purely geometric consideration. The layover phenomena in a SAR image of an urban area is mainly caused by the following two scenarios: 1) Buildings with different heights layovered with the ground; 2) Taller building layovered with the ground and the roof of the lower building. Both scenarios favor double scatterer pairs with smaller elevation distances than with larger distances. If we assume that the heights of the buildings follow a uniform distribution, the elevation distance between the double scatterer pairs should follow approximately a *logarithmic* law.

- Super-resolution is possible. According to the derived fundamental bounds, the achievable super-resolution factors in the typical parameter range of tomographic SAR are found to be promising and are on the order 1.5~25. E.g. in the TerraSAR-X case, ca. 4 m elevation resolution (i.e. ca. 2 m in height) can be achieved with 25 acquisitions at a high SNR of

13dB. Those bounds of SR can be achieved with the newly proposed SL1MMER algorithm. Experiments using TerraSAR-X demonstrate that SL1MMER increases tremendously the proportion of the detected double scatterers from 20% to 38%. This dramatically improved layover separation capability is proven to be mainly associated with the SR power of SL1MMER.

- It is important to understand the fact that the high anisotropic 3-D tomographic resolution element of modern SAR sensors renders the signals sparse in the elevation direction. The proposed CS based SL1MMER algorithm exploits this sparse properties. Besides the mentioned SR capability, it is an *efficient* estimator which provides reliable elevation, motion, amplitude and phase estimates approaching the CRLB, and hence the accuracy only depends on the data quality.

- A robust reconstruction is defined by detecting two scatterers with a distance of one Rayleigh resolution unit at a high probability. Based on this definition, from the derived formulas, it is evident that the minimal number of acquisitions required by the algorithm for robust reconstruction is about 10∼20 as a function of SNR.

- Even if the L_1-L_2 norm minimization is computationally more efficient than the NP-hard NLS, it is still by far more demanding than PSI processing or the simpler linear estimators.

6.2   Outlook

Based on the current status of VHR tomographic SAR inversion presented in this work, a few topics for further study are outlined which mainly concern 1) Data based fusion, 2) Combination of tomographic SAR inversion and PSI, 3) New visualization methods, and 4) Exploitation of sparse signals in the field of remote sensing.

6.2.1   Data based fusion

The data used in this thesis is a single stack of TerraSAR-X spotlight data. Merging measurements of the same or different SAR sensors from different views can increase dramatically the information content:

- The results obtained from a single pass direction, i.e. ascending or descending, represent only the 3-D position and motion of the scattering objects for the side illuminated by the sensor, e.g. only one or two façades of a building. The information of the shadowed side is not measured. Fusion of the results from different pass directions, i.e. ascending *and* descending, can provide the full shape of buildings.

- Tomographic reconstruction from a single viewing angle, can only provide motion parameters in the LOS direction. Merging motion measurements from *different viewing angles* allows the retrieval of motion parameters in the range-elevation plane. With acquisitions from different heading angles, even the motion information in azimuth can be estimated though the estimation accuracy is limited by the geometric configuration. Since the difference in heading angles is usually small, we possess relatively worse estimation accuracy in azimuth. This would render the 3-D error ellipse anisotropic.

- Another possibility is to fuse two or more stacks sharing *comparable wavelengths and comparable incidence angles*, such that the range-elevation frequency support of the tomographic transfer function is as compact as possible (see Fig. 6). The fusion process can be implemented *before* tomographic SAR inversion with special consideration to the possibly oscillating point response function. Such a straightforward data fusion renders the SNR higher, and hence leads to more robust parameter estimation.

- With TanDEM-X, for the first time there is a real multi-antenna array system in space, even though only with single baseline. It enables us to acquire data pairs simultaneously and repeatedly in time. The fusion of TerraSAR-X *and* TanDEM-X data, i.e. adding a couple of TanDEM-X acquisition pairs to the TerraSAR-X data stacks, can be used to improve the result of tomographic SAR inversion on the one hand, and to explore the limits of tomographic reconstruction on the other hand.

  The TanDEM-X data pairs are free of motion, atmosphere and temporal decorrelation, and hence possesses much higher data quality. An elegant way to realize the fusion is introducing the noise covariance matrix $\mathbf{C}_{\varepsilon\varepsilon}$ to all the linear estimators mentioned in Chapter 3 and 4. Taking the L_1-L_2 minimization of eq. (15) as an example, introducing a weighting according to $\mathbf{C}_{\varepsilon\varepsilon}$, yield:

  $$\hat{\boldsymbol{\gamma}} = \arg \min_{\boldsymbol{\gamma}} \left\{ (\mathbf{g} - \mathbf{R}\boldsymbol{\gamma})^{\mathrm{H}} \, \mathbf{C}_{\varepsilon\varepsilon}^{-1} \, (\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}) + \lambda_K \|\boldsymbol{\gamma}\|_1 \right\} \tag{46}$$

  Or alternatively, it can be written as:

  $$\hat{\boldsymbol{\gamma}} = \arg \min_{\boldsymbol{\gamma}} \left\{ \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_{\mathbf{C}_{\varepsilon\varepsilon}}^2 + \lambda_K \|\boldsymbol{\gamma}\|_1 \right\} \tag{47}$$

  Note that since so far TanDEM-X only offers stripmap products, *multi-resolution tomographic processing*, i.e TerraSAR-X spotlight data stack plus a few TanDEM-X stripmap pairs, are required. Also for this task eq. (46) is the basis.

- Modern spaceborne SAR satellite constellation concepts, like the COSMO-SkyMed constellation, allow for acquiring SAR data of the same area from different viewing angles within very short time period. Distinct from aforementioned spectral estimation approach, their realization will open up the opportunity for real nominal SAR tomography, i.e. *spaceborne* tomographic reconstruction *via back projection*.

Besides fusion based on SAR data only, fusion of data from multiple sensors, such as SAR, stereo-optical images and Light Detection and Ranging (LiDAR), is very much of interest. The different sensors complement each other: Stereo-optical data have the best visual interpretability, LiDAR provides very accurate surface models and tomographic SAR inversion as well as the related PSI are the only methods to provide the dynamic component of buildings, e.g. seasonal thermal dilation, structural deformation, or subsidence due to groundwater extraction or underground construction. Their fusion leads directly towards the vision of "a 4-D city".

6.2.2   Combination of tomographic SAR inversion and PSI

PSI is a special case of TomoSAR and has advantages of 1) Computational efficiency and 2) Direct access to phase histories independent of the motion models. However, it is restricted to single scatterers. Tomographic SAR inversion delivers more information such as the number of scatterers, topography and motion information of layover areas even *with very high elevation resolution.* However, it is computationally more expensive. For practical data processing, the combination of the efficiency of PSI and the information increment of tomographic SAR could be very helpful. For instance, as the pre-processing for tomographic reconstruction, PSI can be implemented at the first stage to pixels where single dominating scatterers are ensured, e.g. by examining the pixel intensities. This pre-processing on PS points builds up a reference network for tomographic reconstruction. In addition, the APS can be estimated and corrected in this pre-processing step. Subsequently, tomographic reconstruction can be implemented to the remaining pixels.

6.2.3   New visualization methods

Tomographic SAR inversion creates large volumetric data sets containing a huge amount of information. For a better visualization of the tomographic SAR results, a challenge is to develop new approaches which can: 1) Provide meaningful visualization of the *directional dependencies of the information* since the expected fused TomoSAR results will contain 3-D positioning and 3-D motion information; 2) Handle the *large differences in scale*, because the data sets may cover several kilometers in each direction, while the deformations are measured in millimeters per year; 3) Visualize anisotropic *uncertainties*, e.g. by visualizing boundaries using different transparencies and/or fuzzy visualization.

6.2.4   Exploitation of sparse signals in remote sensing

The SL1MMER algorithm and some findings (e.g. SR bounds) presented in this thesis are generally applicable to sparse reconstruction. Although TomoSAR is taken as the preferred application here, there are other potential applications which can be further exploited. For instance, SL1MMER has been implemented for pan-sharpening in (Zhu, 2011) by exploring the sparsity of the multi-spectral image patches in a dictionary pair learnt from the panchromatic image. It is demonstrated to give higher spatial and spectral resolution with less spectral distortion compared to the conventional methods. In remote sensing, other potential applications are hyperspectral image sharpening and spectral unmixing, ship detection, or radar focusing of targets containing only a few scattering centers which becomes particularly relevant for the even higher resolution SAR systems of the future.

A   Relevant publications as part of the thesis

A.1   Bamler, R., Eineder, M., Adam, N., Zhu, X., Gernhardt, S., 2009. Interferometric potential of high resolution spaceborne SAR. Photogrammetrie-Fernerkundung-Geoinformation 2009 (5): 407-419

Interferometric Potential of High Resolution Spaceborne SAR

RICHARD BAMLER, MICHAEL EINEDER, NICO ADAM, XIAOXIANG ZHU, STEFAN GERNHARDT, München, Oberpfaffenhofen

Summary: The new class of high resolution spaceborne SAR systems, like TerraSAR-X and COSMO-Skymed opens new possibilities for SAR interferometry. The 1m resolution is particularly helpful when 2D, 2.5D, 3D, or 4D (space-time) imaging of buildings and urban infrastructure is required, where the non-interferometric interpretation of SAR imagery is difficult. Structure and deformation of individual buildings can be mapped, rather than only coarse deformation patterns of areas. The paper demonstrates several new developments in high resolution SAR interferometry using TerraSAR-X as an example. Of particular interest is the very high resolution spotlight mode, which requires some care in interferometric processing. Results from interferometry, Persistent Scatterer Interferometry (PSI), and tomographic SAR in urban environment are presented. The high resolution of TerraSAR-X also supports accurate speckle and feature tracking. An example of glacier monitoring is shown and discussed.

Zusammenfassung: *Neue Möglichkeiten der SAR-Interferometrie durch hochauflösende Weltraum gestützte SAR-Systeme*
Die neue Generation hochauflösender SAR-Satelliten, wie TerraSAR-X und COSMO-Skymed, eröffnen neue Möglichkeiten in der SAR-Interferometrie. Die Auflösung von ca. 1m wird vor allem benötigt, wenn Gebäude und urbane Infrastruktur 2-, 2,5-, 3- oder 4-dimensional abgebildet werden sollen. Gerade diese Objekte sind in nicht-interferometrischen SAR-Bildern schlecht zu interpretieren. So können nun Struktur und Deformation einzelner Gebäude interferometrisch aus dem Weltraum vermessen werden, wo bisher nur grobe Deformationsmuster einer Stadt erfassbar waren. In diesem Aufsatz werden mehrere neue Entwicklungen der hochauflösenden SAR-Interferometrie am Beispiel TerraSAR-X vorgestellt. Von besonderem Interesse sind dabei Daten aus dem Spotlight-Modus, deren interferometrische Verarbeitung erläutert wird. Ergebnisse aus Interferometrie, Persistent Scatterer Interferometrie (PSI) und SAR-Tomographie werden präsentiert. Die Anwendung von Speckle Tracking und Feature Tracking zur Erfassung von Gletscherbewegungen wird demonstriert.

1 Introduction

In 2007 Synthetic Aperture Radar (SAR) remote sensing from space (BAMLER & HARTL 1998) made a big leap forward. With the German TerraSAR-X and the Italian COSMO-Skymed satellites have been launched that deliver SAR data with a spatial resolution of up to 1 m compared to typically 10–25 m available so far. The advantage of very high resolution (VHR) imagery for cartographic applications is obvious. The real potential of this class of SAR data, however, lies in applications, where the coherent nature of SAR data is exploited, like interferometry or tomography. The 1m resolution is particularly helpful when it comes to 2D, 2.5D, 3D, or 4D imaging of buildings and urban structures. The inherent spatial scales of buildings are dominated by the typical height between floors of 3–4 m. Hence, for imaging of urban structures we can expect a tremendous improvement in information content with the new VHR data.

This paper presents recent results of TerraSAR-X interferometric data exploitation. The results are representative for the potential of all of the new VHR SAR systems.

2 TERRASAR-X DATA CHARACTERISTICS

TerraSAR-X can be operated in three basic resolution modes (cf. Fig. 1, BUCKREUSS 2003):

Stripmap is the standard mode. It allows imaging of long strips at a ground resolution of about 3m. The swath width is about 40 km.

Spotlight mode uses electronic antenna beam sweeping to increase the coherent integration time, and, hence, the resolution. The beam steering range is ±0.75°. Up to 249 different azimuth patterns are sequentially activated to approximate a continuous beam sweep. 1m resolution in azimuth can be achieved. Depending on the transmitted bandwidth slant range resolutions are 1.2m (@ 150MHz signal bandwidth) or 0.6m (@ 300MHz). The downside is that in spotlight mode only short scene lengths and narrow swaths of 5–10km can be acquired, which are, however, sufficient for most investigations in urban environments.

ScanSAR is a mode of an increased swath width of 100km at the expense of resolution (ScanSAR product resolution: 17m).

These resolution modes are available at different look angles (20° – 55°) and different polarizations, which makes TerraSAR-X an extremely flexible SAR instument.

An important advantage of TerraSAR-X compared to competitors is the high absolute geometric accuracy of the final data products. Every pixel is absolutely georeferenced to within 0.5 – 1 m, provided that an accurate digital surface model is available. This has been achieved by a precise orbit determination and a careful calibration of all instrumental and physical timing error sources, e.g., tropospheric signal delay.

Another advantage of TerraSAR-X is its short revisit time of 11 days. It allows generating interferometric data stacks three times as fast as with ENVISAT/ASAR with its 35 days revisit cycle.

For interferometric applications the short X-band wavelength of 3.1cm is not optimum, since many objects that remain coherent at longer wavelengths will decorrelate rapidly. Even low vegetation is subject to strong temporal decorrelation at this wavelength. On the other hand, the short wavelength lets surfaces appear rough that would be smooth at longer wavelengths. E.g., roads or flat roofs, which show up as totally black in longer wavelength images, have proven to give sufficient backscatter with TerraSAR-X to be exploited as interferometrically useful objects.
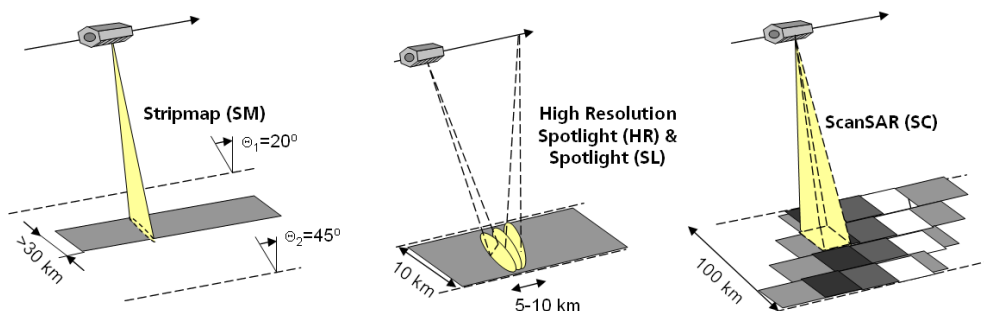
Fig. 1: Basic imaging modes of TerraSAR-X (BUCKREUSS 2003)

3   SPOTLIGHT INTERFEROMETRY

The focused complex TerraSAR-X spotlight images are represented in zero-Doppler coordinates. Due to the quais-continuous beam steering during data acquisition, however, there is a systematic Doppler centroid drift in azimuth direction which must be accounted for during all subsequent InSAR processing steps (EINEDER et al. 2009). Fig. 2 shows the variation of the processed azimuth spectrum of a typical TerraSAR-X product. Although the sampling rate is sufficient to avoid aliasing, the linear drift and the wraps to the principal sampling band are significant. In consequence, the center frequency of the interpolation kernels, used for interferometric image coregistration, must be adjusted in azimuth (band-pass interpolator of varying center frequency). Also the azimuth common band spectral filtering needs to be updated accordingly. These required operations complicate processing so that standard interferometric processing systems can not be used. We have integrated spotlight capabilities into DLR's InSAR processing software GENESIS early enough to be ready for launch of

TerraSAR-X. Therefore, we were able to process spotlight interferograms already during the commissioning phase (ADAM et al. 2007). The interferogram of the city of Paris (cf. Fig. 3) and the detail views shown in Fig. 4 – Fig. 5 demonstrate the wealth of information in this new class of data.
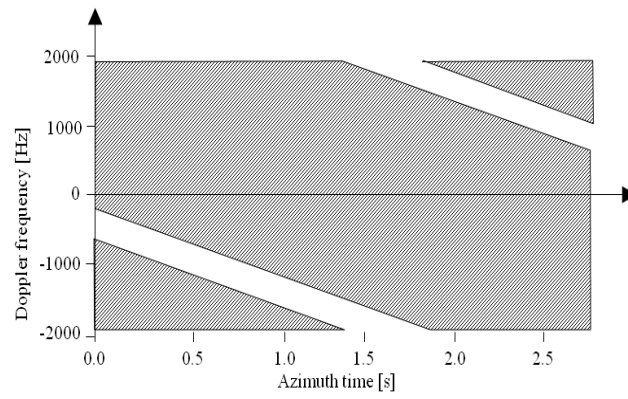
Fig. 2: Support (grey) of time-varying Doppler spectrum of a complex TerraSAR-X spotlight image.
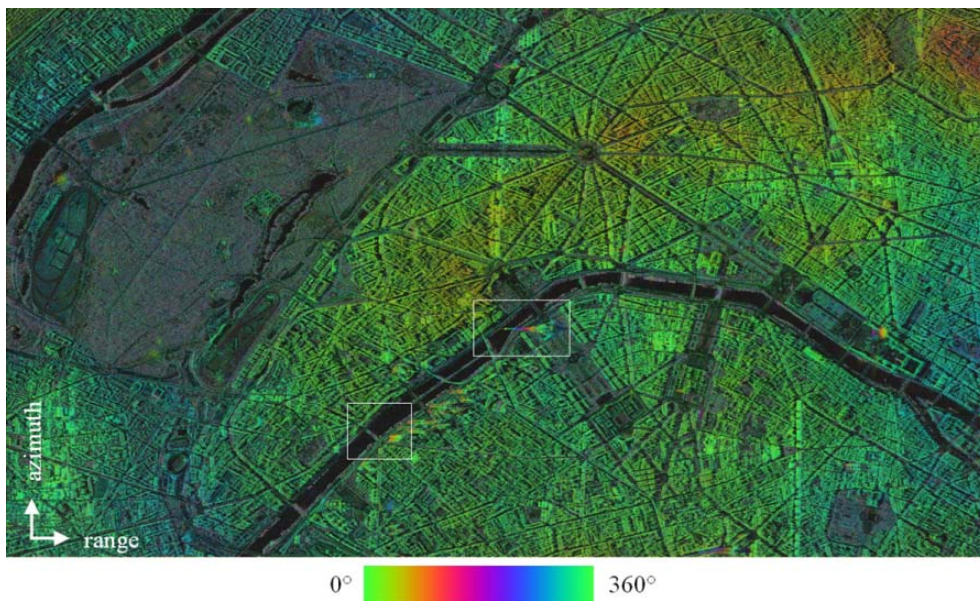
Fig. 3: Flat earth phase corrected TerraSAR-X spotlight interferogram (city of Paris).
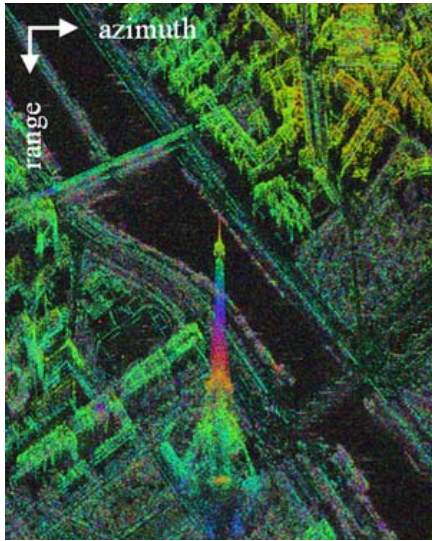
Fig. 4: Zoom of Fig. 3: TerraSAR-X spotlight interferogram of the Eiffel tower. One phase cycle corresponds to 321 m height.
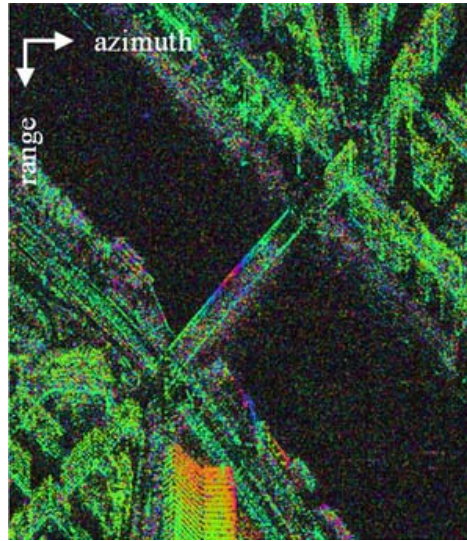
Fig. 5: Zoom of Fig. 3: TerraSAR-X spotlight interferogram of the Mirabeau bridge in Paris. It shows phase changes of 270°. This can be interpreted as a bump of 240 m or (realistically) as a deformation between acquisitions of 12 mm in the radar line of sight direction.

4   PERSISTENT SCATTERER INTERFEROMETRY (PSI)

Persistent Scatterer Interferometry (PSI) has been introduced in 1999 (FERRETTI et al. 2001) as a methodology for long-term monitoring of subsidence, preferably in urban environment. Typically 20 – 100 interferometric data sets of the same area taken from repeat orbit cycles are stacked for PSI analysis. Each pixel in each interferogram is characterized by its range and azimuth coordinates as well as by the temporal and the spatial baselines of the interferogram it belongs to. These four co-ordinates allow for a much better data analysis than possible from a single interferogram, where temporal and spatial baselines are fixed. The goal of the PSI analysis is to separate the following contributions to the interferometric phase:

Elevation of the point: Elevation is the spatial co-ordinate orthogonal to range and azimuth. Its phase contribution is proportional to the spatial baseline.

Deformation rate: If linear deformation, e.g., subsidence, can be assumed, its phase contribution is proportional to the temporal baseline. If the deformation is non-linear often some other parameterized temporal model is assumed whose parameters are adjusted.

Orbit errors and tropospheric water vapor delay: These are spatially long wavelength patterns and and temporally uncorrelated.

By exploiting these proportionalities and correlation properties, elevation and deformation rate can be estimated. Fig. 6 shows the PS location in 3D on Hotel Bellagio in Las Vegas and provides an example for the DEM update and Fig. 7 illustrates an example (Las Vegas Convention Center) for the deformation estimation.
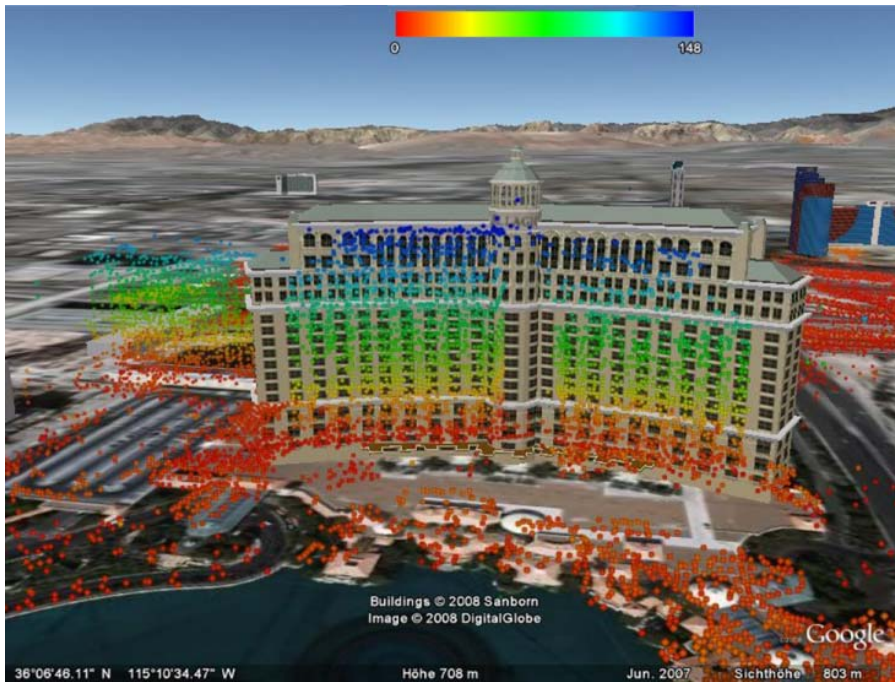
Fig. 6: The color coded elevation estimated by the PSI processing facilitates a geocoding of the PSs [unit: m]. This example presents the 3D location precision and the PS density which perfectly provides the shape of the building's front (Hotel Bellagio, Las Vegas). Note that the Google-Earth building model is too broad.
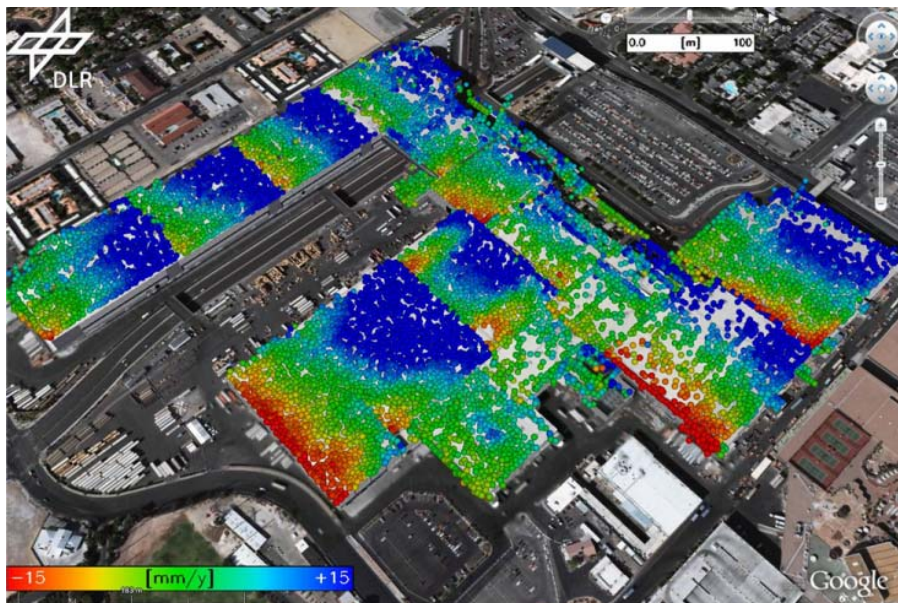
Fig. 7: The PS density of TerraSAR-X allows to monitor the structural stress of buildings. This example presents the deformation (red to blue: -15mm/y to +15mm/y) which results from thermal delation of the building's roof (Las Vegas Convention Center)

Since the data stacks cover a time span of months or years, i.e., much longer than the coherence times of distributed objects (e.g., vegetation), PSI performs all the analysis steps only on bright temporarily stable points, the so-called persistent, or permanent, scatterers (PSs). They are identified in the data stacks by metrics like signal-to-clutter-ratio. In pactise, the SCR allows an ensemble estimate and a prediction of the phase stability. Typical PSs are metallic structures (gratings, poles), facade or roof elements that act as dihedral or trihedral corner reflectors, etc.

PSI has been applied very successfully with medium resolution SAR data from ERS-1/2 and ENVISAT/ASAR. Subsidence rate accuracies of better than 1mm/a have been reported (FERRETTI et al. 2007, ADAM et al. 2009) However, the physical interpretation of these estimates has proven difficult. The low resolution does not give access to details of the buildings. The PSs appear to be quite randomly distributed at a density of about 100–500 PS/km^2, i.e., one PS per block of 100m x 100m to 50m x 50m. There is no guarantee that a particular building of interest is represented by a PS. It is also difficult to differentiate between the subsidence of the building itself or of the pavement surrounding the building.

Many PS are structures acting as dihedral or trihedral reflectors. Assuming a background clutter of -4 dB and a PS detection threshold of 6 dB above clutter, then in ENVISAT/ASAR-type data an ideal full trihedral structure of about 30 cm side length is required to be detected as a PS. This estimation is based on the radar cross section (RCS) of a full trihedral $RCS = 12\pi a^4/\lambda^2$ a ground resolution of 24.4 m x 4.8 m. With TerraSAR-X high resolution spotlight mode and the same assumptions (i.e., a 23° look angle which results in a ground resolution of 1.5 m x 1.1 m) any full trihedral structure of 8 cm side length will be detected as a PS. These types of structures are typically plenty on modern building facades and can be well resolved due to the high resolution. Therefore, we can detect many PSs at a single building and are able to estimate deformation of the building itself. Our first experiences with TerraSAR-X high resolution spotlight date show typical PS densities of 124,000 PS/km^2, i.e., one PS per area of 2 m x 2 m. The PS density depends on resolution (cf. Fig. 8), incidence angle (cf. Fig. 9), and polarization. The increase of PSs with resolution is dramatic because the very high resolution fits well with the typical spatial scales of constructive elements at buildings. The loss of PS density at shallower incidence angles is due to partial shadowing of facades from adjacent buildings. The effect depends on the height of the building, its orientation, and its distance to other buildings.
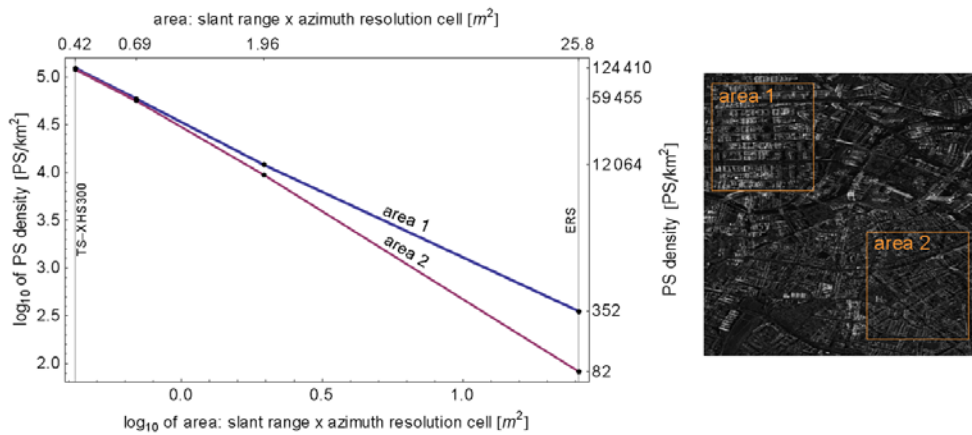
Fig. 8: PS density as a function of spatial resolution. The test site is the city of Berlin. Two areas with different urban development are selected. Area 1 has a high building density and area 2 is a typical urban area.
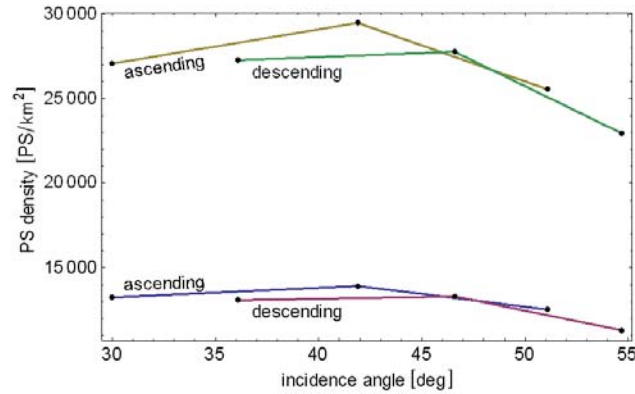
Fig. 9: PS density as a function of incidence angle. The PS density is measured in ascending and descending stacks and two different thresholds on the SCR for the PS detection are applied. The two upper graphs correspond to the low threshold resulting in a high PS density, while the lower ones show the distribution for the high threshold. Both thresholds are higher than the one used for Fig. 8. In princpile, flat look angles result in a loss of PS density because of shadowing. However, 42-47 deg look angle result in the highest PS density.

5   SAR TOMOGRAPHY

Conventional SAR maps the 3D radar reflectivity distribution on ground into the 2D range-azimuth radar coordinate system. Consequences of this geometry are imaging ambiguities like layover, which are particularily pronounced in urban areas. SAR Tomography (TomoSAR) aims at accessing the third dimension, elevation (perpendicular to the range-azimuth plane), resolving layover, and mapping any scatterer in 3D coordinates (REIGBER & MOREIRA 2000). TomoSAR uses stacks of several acquisitions from repeat orbits (and slightly different viewing angles) to establish a synthetic aperture in the elevation direction. In contrast to the – otherwise quite similar – synthetic aperture principle employed in the azimuth direction, TomoSAR must deal with sparse (typically only a few tens of acquisitions) and irregular sampling of the elevation aperture. The reflectivity distribution in the elevation direction for each range-azimuth pixel, and therefore the full 3D reflectivity distribution, is estimated by spectral analysis with special consideration of the mentioned difficulties caused by sparse and irregular sampling. The achievable resolution ρ_s in the elevation direction depends on the extent B_s of the elevation aperture, i.e., the maximum spatial baseline spread, via $\rho_s = \lambda r / 2B_s$ where λ is the wavelengh and r is range (In classical InSAR with baseline B_s, ρ_s is the elevation that causes an interferometric phase of 2π). Since the different acquisitions are taken at different times, possible motion of the sactterers has to be accounted for. Classical InSAR and PSI can be regarded as special cases of parametric TomoSAR.

From the reconstructed reflectivity profile in elevation, multiple layovered objects in any pixel are separated and the following information can be retrieved:

Number of scattering objects: As essential prior knowledge for higher order PSI, the number of scatterers in a resolution cell can be estimated by applying model order selection schemes to the estimated reflectivity profile.

Reflectivity and elevation of the scattering objects: It is obtained by implementing parameter estimation to the reflectivity profile estimates and leads to better localization and understanding of the scattering objects in three dimensions.

Like PSI, TomoSAR benefits greatly from the high resolution of TerraSAR-X data, as the density of coherent pixels and the signal-to-clutter ratio increase significantly with resolution.

In the following examples, TomoSAR is applied to TerraSAR-X high-resolution spotlight data acquired over the city of Las Vegas, USA. 16 scenes are used. The elevation resolution is about $\rho_s =$ 40.5m due to the limited baseline range of the stack of 269.5m. This, however, does not mean that individual scatterers can only be located to within this poor elevation resolution. The Cramér-Rao Lower Bound (CRLB) on elevation estimates can be shown to be:

$$\sigma_{\hat{s}} = \frac{\lambda r}{4\pi\sqrt{NOA} \cdot \sqrt{2\,SNR} \cdot \sigma_{Bs}}$$

where *NOA* is the number of acquisitions, SNR is the signal-to-noise ratio, and σ_{Bs} is the standard deviation of the baseline distribution. For instance, the stack used in this example has $\sigma_{Bs} =$ 78.4m, $\lambda = 0.031m$ and r = 704km. Let us assume the SNR=10dB, the CRLB on elevation estimation is 1.24m. With regular baseline distribution samples, the location accuracy can be related to the elevation resolution by:

$$\sigma_{\hat{s}} = \frac{0.39}{\sqrt{NOA} \cdot \sqrt{SNR}}\rho_s \quad .$$

As input data to our tomographic algorithm we use TerraSAR-X data stacks generated by DLR's GENESIS PSI processor mentioned before. Fig. 10, left, shows the Wynn hotel in Las Vegas with a height close to 200m, corresponding to an elevation range of 380m. The middle image in Fig. 10 is the corresponding TerraSAR-X intensity image where green dots refer to selected PS points. The mean intensity image indicates that pixels containing multiple scatterers are mainly located at the intersection of the bright texture of the building with structures near ground. To exemplify the potential of the TomoSAR method, two pixels marked by red stars and a reference point marked by a yellow star have been selected and will be analyzed in the following. As P1 is located outside of the region of the high-rise building, it is expected that it only contains a single scatterer situated near the ground. By contrast, as P2 is located at the intersection area, we expect two scatterers inside this pixel, among them a weaker ground reflection and a stronger reflection from the building facade. The corresponding reflectivity estimates for those two pixels are shown in the right image of Fig. 10. Ground and building contributions can be well separated, even with such a small number of acquisitions. The spectral estimation method used for this elevation profiles was a singular value decomposition (SVD)-based approach (FORNARO et al. 2003) with singular value weighting according to the Wiener criteron (ZHU et al. 2009). In this example, we can see the potential of the tomographic approach to separate multiple scatterers in one cell using TerraSAR-X data.
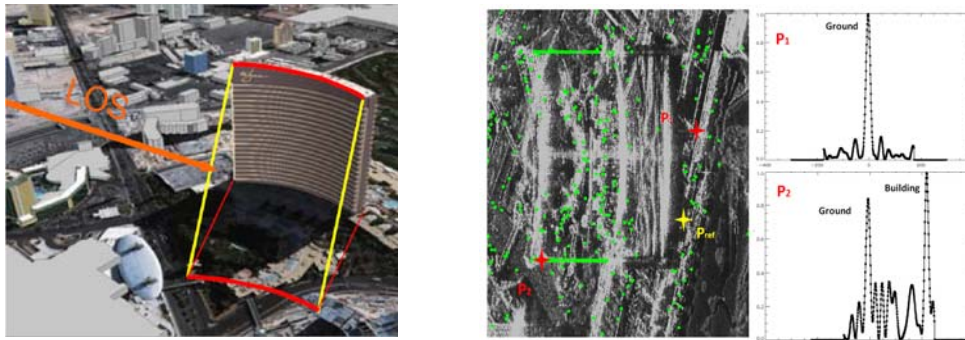
Fig. 6: Wynn Hotel, Las Vegas: Optical image (left) with viewing direction of SAR (LOS = line-of-sight) and two iso-elevation lines (yellow). Mean TerraSAR-X intensity image (center) with reference (yellow) and analysis points (red) P$_1$ and P$_2$. Reconstructed elevation profiles (right) for analysis points P$_1$ (single scatterer) and P$_2$ (two scatterers).

With this convincing result of PS based tomography, a pixel based 3D focusing procedure is now applied to the same data stack. The Las Vegas convention center with a height of about 20m as shown in the left image of Fig. 11 is a very interesting test building for 3D focusing as it is very large and has a regular shape. Therefore, we expect strong and stable returns which provide a very good starting point for pixel-by-pixel 3D focusing. The right image of Fig. 11 shows the corresponding mean TerraSAR-X intensity image. Due to the limited extent of the investigated area phase errors caused by water vapor disturbances can be neglected. The reflectivity profile in elevation direction is reconstructed pixel by pixel. First the number of scatterers is estimated by model order selection based on the Akaike information criterion (SAKAMOTO et al. 1986), then reflectivity and elevation of each scatterer are estimated. Since the different acquisitions have been taken at different times, deformation must be considered, i.e., an additional velocity parameter is estimated in our SVD-Wiener algorithm, leading to a full 4D reconstruction. The left image of Fig. 12 shows the DEM generated from the elevation estimates (height relative to the reference point). The black cross marks the position of the selected reference point. Compared to the InSAR DEM generation procedure, TomoSAR overcomes layover and phase unwrapping problems. The full structure of the convention center has been captured at a very detailed level. For instance, different parts of the building have different heights, which is not visible in the Google Earth building model. Besides the building, more details such as the roads surrounding the convention center and even two bridges above the roads have been captured. There are still some distortions remaining in the middle of the image (red part) where a smooth roof is expected. It may be due to the incorrect linear deformation model assumption. The right image of Fig. 12 shows the extracted deformation velocity corresponding to the reference point. The deformation velocity map is consistent with the PSI processing result of Fig. 7. By checking the deformation of the distorted part mentioned above, unexpected significant subsidence appears which confirms again the incorrect linear deformation model assumption.

Fig. 7: Optical image of Las Vegas Convention Center (left) and corresponding mean TerraSAR-X intensity image (right).
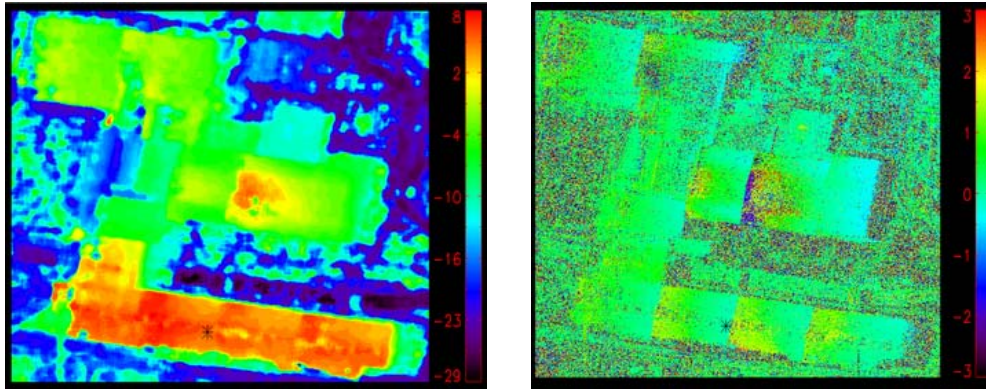
Fig. 8: Left: Generated DEM [unit: m]. Right: Extracted linear deformation velocity w.r.t. the reference point [unit: cm/y] (Black cross: reference point).

6   FEATURE AND SPECKLE TRACKING

The shrinkage or growth of glaciers and especially their flow velocity are indicators for subtle climatic changes. In the past InSAR techniques have proved to be useful tools for monitoring the remote ice sheets of Antarctica and Greenland (GRAY et al. 1998, JOUGHIN 2002), where conventional ground based measurements are not available. Today, high resolution systems such as TerraSAR-X have the potential to further improve the robustness of glacier velocity measurements and they even allow two derive accurate two dimensional motion fields.

SAR interferometry allows the measurement of glacier motion in the radar line of sight with sub-wavelength accuracy, i.e., millimeters to centimeters, depending on the radar wavelength. Typically, the interferometric phase difference between two images taken in two consecutive orbital repeat cycles is exploited, but in many practical cases this method is much too sensitive for typical glacier velocities between 0.1 and 5 meters per day. A further problem arises if spatial velocity gradients lead to more than one fringe per sample which can not be resolved anymore. Assuming an interferometric phase determination accuracy of $\lambda/10$, velocity sensitivities of 0.8 mm/day are achieved with ENVISAT or 1.4 mm/day for TerraSAR-X. Using multiple repeat cycles further increases the sensitivity. Furthermore the InSAR method requires the surface to stay coherent between the times of the repeated observations. This assumption usually holds if the time between acquisitions is short, like the one day during the ERS-1/ERS-2 tandem campaign 1996-2000, or the 3-day repeat orbital period during the ERS-1 ice orbit phase 1993/1994. However, it is seldom the case in the normal repeat cycles of SARs such as the 24-day repeat orbit of RADARSAT-1 and 2 or the 35 days of ERS and ENVISAT/ASAR. The TerraSAR-X orbital cycle of 11 days and the high resolution seem to be a good combination for coherent (speckle) and incoherent (feature) tracking.
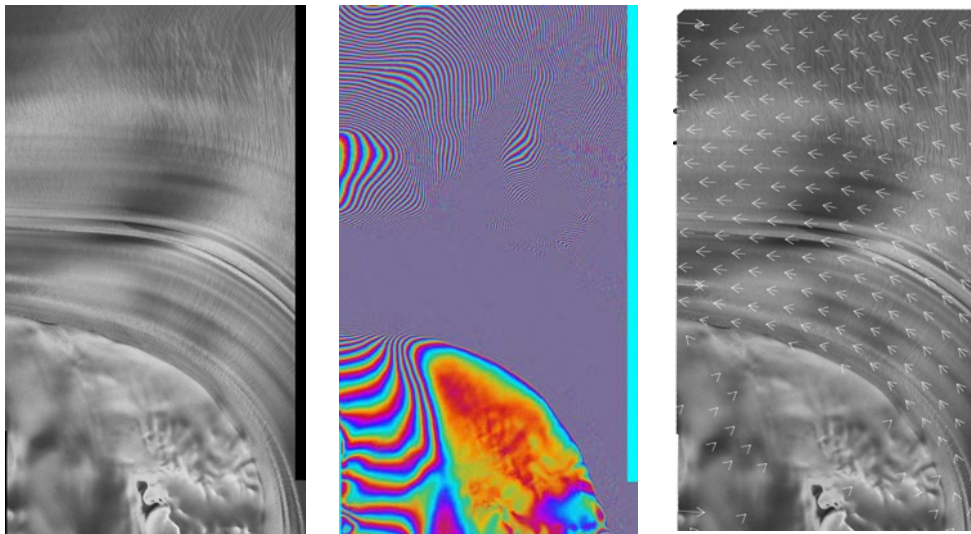
Fig. 13: Left: TerraSAR-X intensity image of the Antarctic Recovery glacier (30.10.2008). Center: 11 day interferogram (29 meter baseline). Right: Motion field derived from Speckle Tracking. Average coherence: 0.3. Horizontal direction (right): range. Vertical direction (up): azimuth.

Fig. 13 shows the TerraSAR-X intensity image of the Antarctic Recovery glacier and the interferogram generated from two images separated by 11 days. The great advantage of the method is evident: even in homogenous areas where the human eye can not identify any texture, the interferometric phase signal reveals surface motion. But a number of constranits impede the InSAR method: First, due to the 2pi-phase ambiguity in each SAR image only relative measurements are available between two points in the interferogram. Therefore, phase unwrapping and a reference point with known (e.g., zero) motion is required. Second, only the line-of-sight component of the surface motion is sensed by the radar, and the three-dimensional surface velocity vector must be derived by either using a DEM and downhill flow assumptions or, by combining satellite passes with different aspect angles. Third, InSAR suffers from decorrelation caused by all surface changes in the size of the wavelength. Such changes occur frequently on glaciers due to melting, rain, snowfall or structural deformation.

High resolution SAR sensors can overcome the above obstacles. Firstly, if image correlation techniques are applied (ignoring the interferometric phase information), ambiguity problems are obsolete. Both, coherent speckle (GRAY et al. 2001) or incoherent features can be tracked with similar correlation methods. Especially coherent speckle tracking requires both images taken from the same position in space, i.e., from orbital repeat cycles. The major error sources for correlation techniques are: 1) atmospheric water vapor of max. 50 cm (range) depending on the humidity, 2) orbital errors on the order of 10 cm (range & azimuth), and 3) the accuracy of the correlation approach. The latter can be brought down to 10 centimeters and below by increasing the correlation chip size accordingly. The overall absolute accuracy of e.g., 50 cm / 11 days = 4.5 cm/day is sufficient for many applications and can be increased easily by using local zero-motion-tie-points, increasing the correlation window or the time lag (increasing the time lag generally leads to decorrelation errors that overcompensates the gain in accuracy). Secondly, image correlation can be performed in both range and azimuth directions and therefore provides a two-dimensional motion vector. Thirdly, geometric phase decorrelation will impede the InSAR phase derivation but not image correlation as long as some object contrast is present, e.g., from crevasses or other surface structures.

Fig. 14 shows the correlation function between pairs of TerraSAR-X images with different coherence and surface characteristics. For convenience, TerraSAR-X EEC-SE products (EINEDER 2005, FRITZ &

EINEDER 2009) were used, i.e., spatially enhanced geocoded detected images with 1.25 m pixel spacing. In these products only about 1.3 looks are averaged to keep the highest quadratic spatial resolution. The small number of looks leaves a good part of the coherent speckle – wide band pseudo-noise that allows accurate correlation. Using these products all correlation can be easily performed in geocoded ground range pixels since co-registration and terrain compensation has been done in the TerraSAR-X SAR processor before.

The left image shows a patch of an Antarctic glacier (Recovery) with high coherence. In consequence the speckle signal delivers a well defined peak even on ice surfaces without object contrast. The correlation peak width is 4 pixels (6 meters). With the used correlation block size of about 2562 indepenndent resolution cells (64 x 64 pixels) and the coherence of about 0.3 an accuracy of 0.029 pixels (0.037 meter) is estimated according to (BAMLER & EINEDER 2005).

The right image shows a patch of a glacier with significant surface changes within 11 days which lead to strong phase decorrelation. In consequence, coherent speckle tracking is no more possible and features have to be used to derive the motion field. The object contrast in Fig. 14 right is caused by crevasses. The achievable resolution now depends on the spatial structure of the features. High contrast and sharp features allow higher accuracy than slowly undulating patterns. In our experiments it turned out that about 4x4=16 looks may to be averaged to sufficiently reduce the amplitude of incoherent speckle maintaining the signal of the lower resolution features. The width of the correlation peak is much wider than that of the speckle tracking result: 17 pixels in range and 12 in azimuth (85 m x 72m). Correlation coefficient of the object features is much lower, in the order of 0.08. With the used correlation block size of 2562 resolution cells this transforms (BAMLER & EINEDER 2005) to an accuracy of 0.47 (East) and 0.33 (North) pixels corresponding to 2.3 m and 1.6 m.

In conclusion, relative accuracies of 0.3 cm/day and, depending on orbit errors and atmospheric conditions, absolute accuracies of 5 cm/day can be achieved by speckle correlation. 20 cm/day are achievable by feature correlation with rather simple methods. More details of the work performed by the authors can be found in (ROTT et al. 2008) and (FLORICIOIU et al. 2008).

a) Recovery Ice Stream 11 days (coherent)   b) Drygalski Glacier 11 days (incoherent)
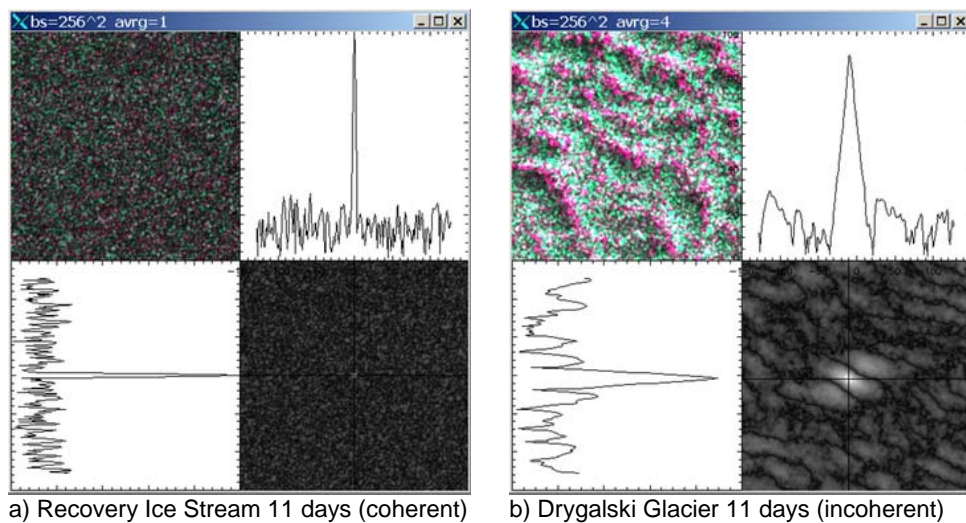
Fig. 14: Correlation functions of TerraSAR-X images with 11 day interval. Upper left: multi-temporal color composite. Lower right: 2D-correlation function. Upper right: range cut through correlation maximum. Lower left: azimuth cut through maximum. a) Homeogeneous coherent glacier surface. Correlation peak width: 6 m. b) Decorrelated surface with crevasses. Resolution: 85 m and 72 m in range and azimuth, respectively. Window size: 256 x 256 pixels. Right image was averaged and decimated by a factor of 4 before correlation.

7 Outlook

Using different interferometric techniques a number of encouraging results were achieved within short time after the launch of TerraSAR-X. Some of them, such as glacier monitoring may be transferred to applications soon while more novel techniques such as TomoSAR will need more experiments to prove their applicability. In 2009 the launch of the co-operative sister satellite TanDEM-X will open even more opportunities for interferometric imaging by adding simultaneously acquired image pairs with high coherence.

References

ADAM, N., EINEDER, M., SCHÄTTLER, B. & YAGUE-MARTINEZ, N., 2007: First TerraSAR-X Interferometry Evaluation. – Proceedings of ESA FRINGE Workshop, Frascati.

ADAM, N., EINEDER, M., YAGUE-MARTINEZ, N. & BAMLER, R., 2008: High Resolution Interferometric Stacking with TerraSAR-X. – Proceedings of IEEE International Geoscience and Remote Sensing Symposium 2008, Boston, USA.

ADAM N., PARIZZI, A., EINEDER, M. & CROSETTO, M., 2008: Practical Persistent Scatterer Processing Validation in the Course of the TERRAFIRMA Project. – Journal of Applied Geophysics, in print 2009.

BAMLER, R. & HARTL, P., 1998: Synthetic Aperture Radar Interferometry. – Inverse Problems **14**, R1–R54.

BAMLER, R. & EINEDER, M., 2005: Accuracy of Differential Shift Estimation by Correlation and Split Bandwidth Interferometry for Wideband and Delta-k SAR Systems. – IEEE Geoscience and Remote Sensing Letters **2** (2), 151–155.

BUCKREUSS, S., BALZER, W., MÜHLBAUER, P., WERNINGHAUS, R. & PITZ, W., 2003: The TerraSAR-X Satellite Project. – Proceedings of IEEE International Geoscience and Remote Sensing Symposium 2003, Toulouse, France.

EINEDER, M., ADAM, N., BAMLER, R., YAGUE-MARTINEZ, N. & BREIT, H., 2009: Spaceborne Spotlight SAR Interferometry with TerraSAR-X. – IEEE Transactions on Geoscience and Remote Sensing **47** (5): 1524–1535.

EINEDER, M., SCHÄTTLER, B., BREIT, H., FRITZ, T. & ROTH, A., 2005: TerraSAR-X SAR Products and Processing Algorithms. – Proceedings of IEEE International Geoscience and Remote Sensing Symposium 2005, Seoul, South Korea.

FERRETTI, A., PRATI, C. & ROCCA, F., 2001: Permanent Scatterers in SAR Interferometry- – IEEE Transactions on Geoscience and Remote Sensing **39** (1): 8–20.

FERRETTI, A., SAVIO, G., BARZAGHI, R., BORGHI, A., MUSAZZI, S., NOVALI, F., PRATI, C. & ROCCA, F., 2007: Submillimeter Accuracy of InSAR Time Series: Experimental Validation. – IEEE Transactions on Geoscience and Remote Sensing **45** (5): 1142–1153.

FLORICIOIU, D., EINEDER, M., ROTT, H. & NAGLER, T., 2008: Velocities of Major Outlet Glaciers of the Patagonia Icefield Observed by TerraSAR-X. – Proceedings of IEEE International Geoscience and Remote Sensing Symposium 2008, Boston, USA.

FORNARO, G., SERAFINO, F. & SOLDOVIERI, F., 2003: Three-dimensional focusing with multipass SAR data. – IEEE Transactions on Geoscience and Remote Sensing **41** (3): 507–517.

FRITZ, T. & EINEDER, M., 2009: TerraSAR-X Ground Segment – Basic Product Specification Document. – Doc.: TX-GS-DD-3302, Revision 1.6, available online.

GRAY, A. L., MATTAR, K.E. & VACHON, P. W., 1998: InSAR results from the RADARSAT Antarctic Mapping Mission Data: Estimation of Glacier Motion using a Simple Registration Procedure. – Proceedings of IEEE International Geoscience and Remote Sensing Symposium 1998, Seattle, USA.

GRAY, A. L., SHORT, N., MATTAR, K. E. & JEZEK, K. C., 2001: Velocities and flux of the Filchner ice shelf and its tributaries determined from speckle tracking interferometry. – Canadian Journal of Remote Sensing **27** (3): 193–206.

JOUGHIN, I., 2002: Ice-sheet velocity mapping: a combined interferometric and speckle-tracking approach. – Annals of Glaciology 34.

REIGBER, A. & MOREIRA, A., 2000: First demonstration of airborne SAR tomography using multibaseline L-band data. – IEEE Transactions on Geoscience and Remote Sensing **38** (5): 2142–2152.

ROTT, H., EINEDER, M., NAGLER, T. & FLORICIOIU, D., 2008: New results on dynamic instability of Antarctic Peninsula glaciers detected by TerraSAR-X ice motion analysis. – In: European Conference on Synthetic Aperture Radar (EUSAR), VDE Conference Services, EUSAR European Conference on Synthetic Aperture Radar 2008, Friedrichshafen, Germany.

SAKAMOTO, Y., ISHIGURO, M. & KITAGAWA, G., 1986: Akaike information criterion statistics. – Reidel, Dordrecht, The Netherlands.

ZHU, X., ADAM, N., BRCIC, R. & BAMLER, R., 2009: Space-borne High Resolution SAR Tomography: Experiments in Urban Environment Using TerraSAR-X Data. – Proceedings of Joint Urban Remote Sensing Event 2009, Shanghai, China.

Addresses of the Authors:

Prof. Dr.-Ing. habil. RICHARD BAMLER, Dr. rer. nat. MICHAEL EINEDER, NICO ADAM, Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, 82234 Wessling, Germany, Tel.: +49-8153-28-2673, -1396, -1326, Fax: -1420, e-mail: richard.bamler@dlr.de, michael.einder@dlr.de, nico.adam@dlr.de

Prof. Dr.-Ing. habil. RICHARD BAMLER, XIAOXIANG ZHU, STEFAN GERNHARDT, Technische Universität München, Lehrstuhl für Methodik der Fernerkundung, Institut für Photogrammetrie und Kartographie, Arcisstraße 21, 80333 München, Germany, Tel.: +49-8153-28-2673, Fax: -1420, e-mail: richard.bamler@dlr.de, xiaoxiang.zhu@bv.tum.de, Stefan.Gernhardt@bv.tum.de.

A.2 Zhu, X., Bamler, R., 2010d. Very High Resolution Spaceborne SAR Tomography in Urban Environment. IEEE Transactions on Geoscience and Remote Sensing 48 (12): 4296-4308

Very High Resolution Spaceborne SAR Tomography in Urban Environment

Xiao Xiang Zhu, *Student Member, IEEE*, and Richard Bamler, *Fellow, IEEE*

Abstract—Synthetic aperture radar tomography (TomoSAR) extends the synthetic aperture principle into the elevation direction for 3-D imaging. It uses stacks of several acquisitions from slightly different viewing angles (the elevation aperture) to reconstruct the reflectivity function along the elevation direction by means of spectral analysis for every azimuth–range pixel. The new class of meter-resolution spaceborne SAR systems (TerraSAR-X and COSMO-Skymed) offers a tremendous improvement in tomographic reconstruction of urban areas and man-made infrastructure. The high resolution fits well to the inherent scale of buildings (floor height, distance of windows, etc.). This paper demonstrates the tomographic potential of these SARs and the achievable quality on the basis of TerraSAR-X spotlight data of urban environment. A new Wiener-type regularization to the singular-value decomposition method—equivalent to a maximum *a posteriori* estimator—for TomoSAR is introduced and is extended to the differential case (4-D, i.e., space–time). Different model selection schemes for the estimation of the number of scatterers in a resolution cell are compared and proven to be applicable in practice. Two parametric estimation algorithms of the scatterers' elevation and their velocities are evaluated. First 3-D and 4-D reconstructions of an entire building complex (including its radar reflectivity) with very high level of detail from spaceborne SAR data by pixelwise TomoSAR are presented.

Index Terms—Differential synthetic aperture radar tomography (D-TomoSAR), spotlight SAR, TerraSAR-X, urban mapping.

I. INTRODUCTION

A CONVENTIONAL space- or airborne synthetic aperture radar (SAR) maps the 3-D reflectivity distribution of a scene to be imaged into the 2-D azimuth–range $(x-r)$ plane. This can be seen as a projection along the third radar coordinate, namely, elevation (s). x, r, and s form an orthogonal coordinate system specific to the particular SAR imaging geometry. This projection particularly handicaps the interpretation of SAR images of the following: 1) volumetric scatterers and 2) urban areas and man-made objects, i.e., objects with constructive elements oriented at steeper angles than the local incidence angle.

Fig. 1. TomoSAR imaging geometry. The coordinate s is referred to as elevation.

SAR tomography (TomoSAR), introduced to SAR in the early 1990s [1], extends the synthetic aperture principle into the elevation direction for 3-D imaging (within the first Born approximation). It uses data stacks of several acquisitions from slightly different viewing angles (the elevation aperture) to reconstruct the reflectivity function along the elevation direction by means of spectral analysis for every azimuth–range pixel, and hence obtains focused 3-D SAR images (Fig. 1). A further extension is *differential* SAR tomography (D-TomoSAR) [2], also referred to as 4-D focusing. It provides retrieval of both the elevation and the deformation information of multiple scatterers inside an azimuth–range resolution cell and therefore obtains a 4-D (space–time) map of scatterers.

Compared to computed axial tomography, known from medical imaging, TomoSAR uses only a small angular diversity. Hence, instead of back projection, spectral estimation is sufficient for TomoSAR if the range migration δr caused by the different viewing angles is much smaller than the range resolution ρ_r. This gives a limitation to the extent Δs of the illuminated objects

$$\Delta s \ll \frac{\rho_r r}{\Delta b} \tag{1}$$

where r is the range and Δb is the perpendicular (or effective) baseline range (i.e., the elevation aperture length). The term "baseline" is a heritage of interferometry. It is the spatial dimension in the elevation aperture, relative to a reference (master) track.

The first experiments in TomoSAR were carried out in the laboratory [3] under ideal experimental conditions or by using

airborne systems [1], [4]. Spaceborne TomoSAR tests were reported in [5] and [6]. It has been applied to C-band European Remote Sensing Satellite (ERS) data over extended scenes in [7] and to TerraSAR-X data in [8]. In [9], the single- and double-scatterer cases were separated. The concept of 4-D SAR imaging (D-TomoSAR) was proposed in [2] and first applied to ERS data in [10].

The major challenges in spaceborne TomoSAR are the following. First, acquisitions are unevenly distributed in baseline, i.e., in the elevation aperture space, so that appropriate regularization is required instead of a classical Fourier-based inversion [11]. Second, 3-D data cannot be collected simultaneously, at least with existing satellites, but must rather be acquired via repeated passes that are separated in time. Hence, time-dependent phase terms from motion and from the propagation medium are present in the data and must be considered. Third, the number of acquisitions may be limited.

In 2007, SAR remote sensing from space made a big leap forward; the German TerraSAR-X and the Italian COSMO-Skymed satellites have been launched. They deliver SAR data with a very high spatial resolution of up to 1 m compared to medium (10–30-m)- and high (3–10-m)-resolution SAR systems available so far. The advantage of very high resolution (VHR) imagery for cartographic applications is obvious. The real potential of this class of SAR data, however, lies in applications, where the coherent nature of SAR data is exploited, like interferometry or tomography. The 1-m resolution is particularly helpful when it comes to 2-D, 3-D, or 4-D imaging of buildings and urban structures. The inherent spatial scales of buildings are dominated by the typical height between floors of 3–3.5 m, i.e., in slant range (at 30°) 2.6–3.0 m, and the distance of windows. Hence, for imaging of urban structures, we can expect a tremendous improvement in information content when we go from high to VHR. We work with TerraSAR-X *spotlight* data. These VHR X-band spaceborne repeat-pass tomographic data stacks of urban areas have some particular properties: A very detailed view of individual buildings is possible; the density of bright (high clutter-to-noise ratio) points, like persistent scatterers, is extremely high (40 000–100 000/km^2). However, also nonlinear (e.g., thermally induced) movements of different building parts must be expected and will introduce additional phase errors and require robust inversion methods.

This paper aims at demonstrating the potential of the new class of VHR spaceborne SAR systems for TomoSAR in urban environment. In particular, we introduce a new Wiener-type regularization to the singular-value decomposition (SVD) method [7] for TomoSAR and extend it to the D-TomoSAR case (Section IV). Different model selection schemes for the estimation of the number of scatterers are evaluated and validated (Section V). Two parametric estimation algorithms of the scatterers' elevation and their velocities are evaluated (Section VI). We demonstrate first 3-D and 4-D reconstructions of an entire building from spaceborne VHR data by pixelwise TomoSAR (Section VII). We will concentrate on urban areas and man-made infrastructure. Volumetric objects, like trees, can be considered incoherent in X-band repeat pass and are hence treated as noise. A limited number (typically one to three) of scatterers is expected along every elevation profile that allows parametric estimation.
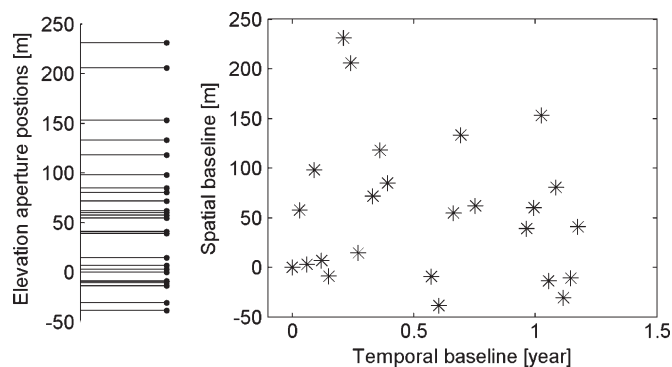
Fig. 2. (Left) Elevation aperture sampling positions of the 25 acquisitions. (Right) Spatial–temporal baseline distribution (to be used for D-TomoSAR in Section III).

II. Data Set

For the purpose of this paper, we work with TerraSAR-X "high-resolution spotlight data" (TerraSAR-X product terminology) acquired with a range bandwidth of 300 MHz. They have a slant-range resolution of 0.6 m and an azimuth resolution of 1.1 m. In this mode, image lengths of 5–10 km can be acquired, which is sufficient for our investigations in urban environments. Note that interferometric use of spotlight data requires some special care, e.g., in synchronous data acquisition, coregistration, and resampling [12]. Our test site is Las Vegas, NV, U.S. The acquisition repeat cycle is 11 days. The orbit of TerraSAR-X is controlled in a predefined tube of 500 m diameter throughout the entire mission [13]. Due to this small orbit tube, the precondition for the spectral estimation approximation mentioned in (1) is very easily fulfilled. It is worth mentioning that, unlike in airborne TomoSAR, the relatively large temporal separation of the repeated passes of spaceborne data collection introduces motion and atmospheric phase contributions that have to be accounted for—albeit as nuisance parameters. This requires a lot more data sets to get unambiguous results and resolve multiple scatterers inside an azimuth–range cell. In our experiment, a data stack of 25 scenes is used for our test site. The elevation aperture sampling positions are shown in Fig. 2. The elevation aperture size Δb is about 269.5 m. According to (1), with $r = 704$ km and $\rho_r = 0.6$ m, the elevation extent Δs of the illuminated objects must be much smaller than 1568 m. This is always true for our test site. Therefore, in the following sections, we handle TomoSAR as a spectral estimation problem, and the detailed system model will be introduced in Section III.

For nonparametric spectral analysis, the expected elevation resolution ρ_s, i.e., the width of the elevation point response function (PRF), depends on the elevation aperture length Δb and is approximately (sufficiently dense sampling of the elevation aperture provided)

$$\rho_s = \frac{\lambda r}{2\Delta b} \tag{2}$$

where λ is the wavelength. It results in 40.5-m resolution in elevation expected for our stack, which is approximately 20-m resolution in height with the elevation-to-height factor $\sin \theta$, where θ is the incidence angle and equals 31.8° here. This, however, does not mean that individual scatterers can only be
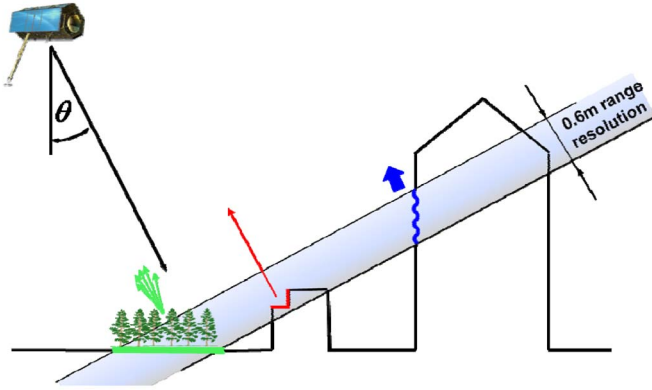
Fig. 3.   Possible signal contributions in a single SAR image azimuth–range pixel.

located to within this poor elevation resolution. The Cramér–Rao lower bound (CRLB) on elevation estimates can be shown to be [14]

$$\sigma_{\hat{s}} = \frac{\lambda r}{4\pi\sqrt{NOA} \cdot \sqrt{2SNR} \cdot \sigma_b} \qquad (3)$$

where NOA is the number of acquisitions, SNR is the signal-to-noise ratio, and σ_b is the standard deviation of the baseline distribution. For instance, the stack used in this example has $\sigma_b = 70.9$ m. For many bright points, we can assume an SNR of 10 dB; then, the CRLB on elevation estimation is 1.1 m, i.e., almost a 1/40 of the elevation resolution.

III. TomoSAR

A. System Model

For a single SAR acquisition, the focused complex-valued measurement $g_n(x_0, r_0)$ of a specific azimuth–range pixel (x_0, r_0) for the nth acquisition with aperture position b_n and temporal baseline t_n is the integral of the reflected signal along the elevation direction, as shown in Fig. 3. In VHR X-band data, we expect the following signal contributions (see Fig. 3).

1) Weak diffuse scattering from—mostly horizontal or vertical—rough surfaces (roads and building walls). They have an elevation extent of $\rho_r/\tan\theta$ for horizontal and $\rho_r \cdot \tan\theta$ for vertical surfaces. In both cases, these extents are much smaller than our elevation resolution ρ_s, and hence, these surfaces can be treated as discrete scatterers in the elevation direction (delta functions).
2) Strong returns from metallic structures or specular and dihedral or trihedral reflections. These are points that would also be used in persistent scatter interferometry. They are the dominating signal contributions. With VHR SAR data, the density of these points can be very high, as mentioned before.
3) Returns from volumetric scatterers, e.g., from vegetation. These result in a continuous signal background in elevation. These ensembles of scatterers, however, often decorrelate in time, and their response is therefore treated as noise.

The noise sources are the following.

a) Gaussian noise, which is caused by thermal noise and temporal decorrelation, as mentioned previously.

b) Calibration errors in amplitude. According to an unpublished DLR internal calibration report [15], the radiometric stability of TerraSAR-X, i.e., the amplitude variations within one stack, is 0.14 dB and is therefore negligible compared to our typical SNR.
c) Phase errors caused by atmospheric delay and unmodeled motion. They require robust and phase-error-tolerant estimation methods.

One SAR acquisition may be considered to be one tomographic projection of the complex reflectivity of the object along elevation [16] (note that the deformation term is ignored here for simplicity)

$$g_n = \int_{\Delta s} \gamma(s) \exp(-j2\pi\xi_n s) \, ds, \qquad n = 1, \dots, N \quad (4)$$

where $\gamma(s)$ represents the reflectivity function along elevation s. $\xi_n = -2b_n/(\lambda r)$ is the spatial (elevation) frequency. The continuous-space system model of (4) can be approximated by discretizing the continuous reflectivity function along s within its extent Δs by s_l $(l = 1, \dots, L)$

$$g_n \approx \delta s \cdot \sum_{l=1}^{L} \gamma(s_l) \exp(-j2\pi\xi_n s_l), \qquad n = 1, \dots, N \quad (5)$$

where L is the number of discrete elevation indices and the discretization interval is $\delta s = \Delta s/(L-1)$. After dropping the inconsequential leading constant δs, the system imaging model becomes

$$\mathbf{g} = \mathbf{R}\,\boldsymbol{\gamma} \qquad (6)$$

where $\mathbf{g}$ is the measurement vector with N elements g_n, $\mathbf{R}$ is an $N \times L$ mapping matrix with $R_{nl} = \exp(-j2\pi\xi_n s_l)$, and $\boldsymbol{\gamma}$ is the discrete reflectivity vector with L elements $\gamma_l = \gamma(s_l)$. Equation (6) is essentially an irregularly sampled discrete Fourier transform of the elevation profile $\gamma(s)$. The objective of TomoSAR is to retrieve the reflectivity profile for each azimuth–range pixel and then use it to estimate scattering parameters such as the number of scatterers present in the cell, their elevations, reflectivities, and line-of-sight (LOS) deformation velocities (see Section IV). This can be achieved from a spectral analysis of the multipass data stack with N SAR acquisitions.

B. Processing Sequence

The processing procedure, with the objective of reconstructing the 3-D scatterer distribution from measurements of the scattered field and estimating LOS deformation, is shown in Fig. 4. The preprocessing, including atmospheric phase screen correction, is performed by the German Aerospace Center (DLR)'s PSI-GENESIS system [17]–[20].

To acquire an estimate of the reflectivity profile along elevation for a certain azimuth–range pixel, *nonparametric* spectral analysis is used for the first stage of processing. Except the maximum elevation extent Δs of the object and some statistical properties of the prior and the noise, no prior
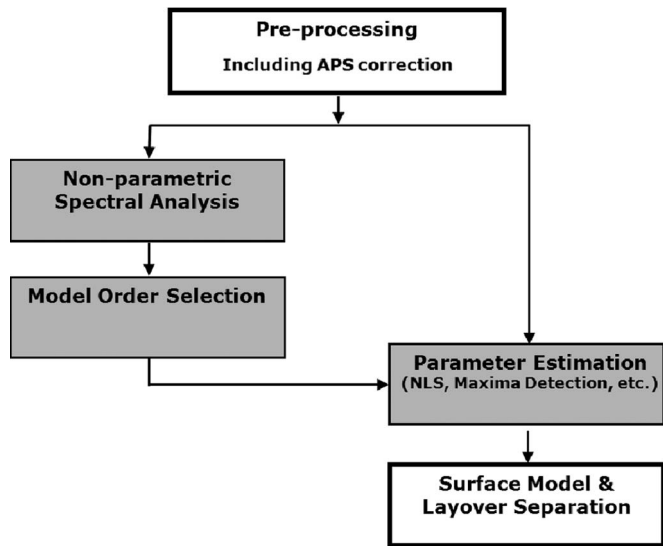
Fig. 4. Processing sequence of tomographic SAR data stacks.

knowledge, such as the number of scattering objects, and no assumption about the scattering mechanism are required. The estimated profile is then used as a prior knowledge for model order selection and parameter estimation. There are many inversion methods, such as SVD, Capon, multiple signal classification (MUSIC), etc. However, it is not our purpose to compare different spectral estimation methods, which have been discussed in detail in [21]–[23]. The high phase errors due to unmodeled motion require robust methods (see self-cancellation of Capon [24]). In addition, we want to maintain the full range and azimuth resolution since the 1-m range–azimuth resolution of the data is essential for urban applications. As mentioned in Section I, the inner scale of buildings is typically in the range of 2–3 m. Any blurring, as it is necessary for the covariance matrix estimation in MUSIC, Capon, etc., is lethal for the information content; individual bright points will be merged to bright clusters. Moreover, the typical signals of buildings are not ergodic. Taking all this into account, the SVD-based method has been chosen because of its good behavior at high noise levels without compromising azimuth–range resolution. With nonparametric estimates, the scatterers' distribution in the elevation–velocity $(v-s)$ plane, to be introduced later, and a 3-D reflectivity map of the entire illuminated scene are obtained. They are input to model order selection, i.e., to the estimation of the number of discrete (pointlike) scatterers.

With prior knowledge of the number of scatterers, a parametric spectral estimator, for instance, nonlinear least squares (NLS), which is the maximum-likelihood estimator (MLE) for Gaussian white noise, can be applied to the measurements to refine the estimates at the cost of large computational effort. Alternatively, we can simply estimate the location of the scatterers by detecting the peaks of the nonparametric SVD estimates, which is much faster but may introduce estimation bias caused by interference between multiple scatterers.

We will not go into the details of preprocessing as they are sufficiently addressed in the literature. The remaining modules (gray boxes in Fig. 4) are outlined next.

IV. MAP ESTIMATOR AND WIENER SVD

The standard MAP estimator for γ from (6) is given by

$$\hat{\gamma}_{\mathrm{MAP}} = \left(\mathbf{R}^{\mathrm{T}}\mathbf{C}_{\varepsilon\varepsilon}^{-1}\mathbf{R} + \mathbf{C}_{\gamma\gamma}^{-1}\right)^{-1}\mathbf{R}^{\mathrm{T}}\mathbf{C}_{\varepsilon\varepsilon}^{-1}\mathbf{g} \tag{7}$$

where $\mathbf{C}_{\varepsilon\varepsilon}$ is the noise covariance matrix and $\mathbf{C}_{\gamma\gamma}$ is the covariance matrix of the prior. It reduces to

$$\hat{\gamma}_{\mathrm{MAP}} = \left(\mathbf{R}^{\mathrm{T}}\mathbf{R} + |\varepsilon|^2\mathbf{I}\right)^{-1}\mathbf{R}^{\mathrm{T}}\mathbf{g} \tag{8}$$

if both the noise and the prior are assumed to be white, i.e., the noise covariance matrix $\mathbf{C}_{\varepsilon\varepsilon} = |\varepsilon|^2\mathbf{I}$, the covariance matrix of the prior $\mathbf{C}_{\gamma\gamma} = \mathbf{I}$, and the signal power is assumed to be normalized to unity. This nonparametric spectral estimation method has been chosen because of its robustness at high noise levels without sacrificing the azimuth–range resolution.

Although the MAP estimator from (8) could be implemented directly, the treatment of the problem in the singular-value (SV) space is helpful. As will be shown, the distribution of the SVs helps us understand the determinedness of the problem and estimate the noise level.

In this section, we will show that the MAP estimator is equivalent to a Wiener-type regularization of the SVs. It is a more strict solution than the original truncated SVD (TSVD) method [16]. We also give its extension to D-TomoSAR.

A. SVD method

The SVD inversion framework has been elegantly described in [16]. The discrete reflectivity signal γ can be reconstructed from $\mathbf{g}$ through pseudoinversion of the imaging system matrix $\mathbf{R}$ [(6)]. However, due to the nonuniform track distribution, the solution may include significant noise propagation due to the ill-conditioned nature of the problem. The SVD is a simple and valuable tool for analyzing image quality and the amount of independent information about the unknowns that can be reliably retrieved from observations in the presence of noise [11]. The SVD of $\mathbf{R}$ is a decomposition of the form

$$\mathbf{R} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}} = \sum_{n=1}^{N}\mathbf{u}_n\sigma_n\mathbf{v}_n^{\mathrm{T}}. \tag{9}$$

where $\mathbf{U} = (\mathbf{u_1}, \ldots, \mathbf{u_N})$ and $\mathbf{V} = (\mathbf{v_1}, \ldots, \mathbf{v_N})$ are matrices with orthonormal columns, $\mathbf{U^T U} = \mathbf{V^T V} = \mathbf{I_N}$, and $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_N)$ has nonnegative diagonal elements such that $\sigma_1 \geq \cdots \geq \sigma_N \geq 0$. σ_n denotes the SVs of $\mathbf{R}$, while the vectors $\mathbf{u}_n$ and $\mathbf{v}_n$ are the left and right singular vectors of $\mathbf{R}$, respectively. Consider now that an estimate of γ is obtained via the pseudoinverse $\mathbf{R}^{\dagger}$. Using the SVD, we get

$$\hat{\gamma} = \mathbf{R}^{\dagger}\mathbf{g} = \sum_{n=1}^{N}\sigma_n^{-1}\left(\mathbf{u}_n^{\mathrm{T}}\mathbf{g}\right)\mathbf{v}_n. \tag{10}$$

Due to the reciprocal of σ_n, noise propagation caused by small SVs will compromise this solution, and regularization tools are required.
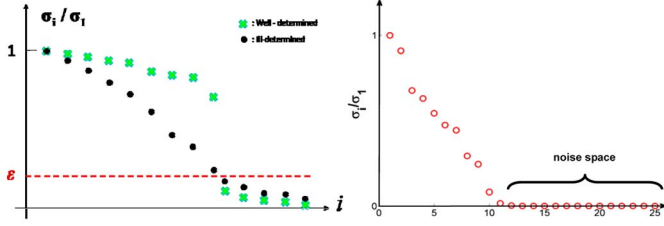
Fig. 5. (Left) SV spectra corresponding to ill-conditioned matrices with well- and ill-determined numerical ranks. (Right) Corresponding SV spectrum of our TomoSAR configuration with elevation aperture sampling from Fig. 2.

B. Wiener Regularization

A well-known method for dealing with an ill-conditioned matrix in problem (7) is TSVD [25]. The basic idea of TSVD and other regularization methods is to impose additional requirements on the solution, thus hopefully damping the contributions from the errors of the right-hand side of (7). In the case of TSVD, this is achieved by discarding the components of the solution corresponding to the smallest $N - Q$ SVs. These contributions to the solution are most likely noise and would be amplified unduely by the inverse of σ_n. Q is called "numerical rank" or "effective rank" of $\mathbf{R}$ defined by the number of SVs bigger than some noise level ε. The TSVD was implemented in [16] for an experiment with 44 acquisitions, and robust performance was achieved.

Depending on the SV spectrum, it is common to characterize an ill-conditioned matrix as either with a *well-* or an *ill-determined* numerical rank [26]. An ill-conditioned matrix with a well-determined numerical rank has a well-defined gap between the significant SVs contributing to the signal space and the small SVs contributing to the noise space (see Fig. 5, left, green crosses). Matrices with ill-determined rank degrade gradually from the signal to the noise space (see Fig. 5, left, black dots). Any distribution of SVs in between the two extremes of the left figure of Fig. 5 may, of course, be expected in practical applications. However, from the perturbation theory for the TSVD [25], TSVD is a stable method only for problems with a well-determined numerical rank.

The corresponding SVs of our Las Vegas TerraSAR-X data set with the baseline distribution of Fig. 2 are shown in the right plot of Fig. 5. It is obviously of ill-determined numerical rank. The result is then overly dependent on how a hard threshold is set. Transforming the MAP estimator of (8) to the SV space results readily in a soft thresholding, e.g., weighting the SVs according to their magnitudes, also referred to as a Tikhonov regularization

$$\hat{\gamma}_{\mathrm{MAP}} = \left(\mathbf{\Sigma}^T \mathbf{\Sigma} + |\varepsilon|^2 \mathbf{I}\right)^{-1} \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T \mathbf{g}$$
$$= \sum_{n=1}^{N} \sigma_{n,\mathrm{Wiener}}^{-1} \left(\mathbf{u}_n^T \mathbf{g}\right) \mathbf{v}_n \qquad (11)$$

where $\sigma_{n,\mathrm{Wiener}}^{-1}$ denotes the optimum weights

$$\sigma_{n,\mathrm{Wiener}}^{-1} = \frac{\sigma_n}{|\sigma_n|^2 + |\varepsilon|^2}. \qquad (12)$$

It replaces σ_n^{-1} in (10). $|\varepsilon|^2$ is the noise power level. A small ε corresponds to a high SNR. This type of weighting resembles the Wiener filter under white noise, and hence, we call the
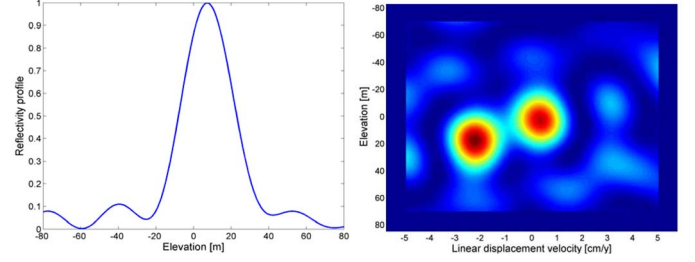
Fig. 6. Two close scatterers separable with D-TomoSAR due to the different velocities. (Left) Reconstructed reflectivity profile with TomoSAR; the scatterers are not separable. (Right) Retrieved scatterer distribution in the $s-v$ plane with D-TomoSAR.

method SVD-Wiener. It provides more stable performance, particularly for the case of few (15–30) acquisitions.

Now, we come to the problem of estimating the noise level ε. Let us define coefficients $\beta_n = \mathbf{u}_n^T \mathbf{g}$, which are the projection of measurements onto the singular vectors. The noise level can be estimated from the n_ε coefficients β_n ($n = N - n_\varepsilon + 1, \ldots, N$) corresponding to the noise space (in our experiment, n_ε can be set to 14; see Fig. 5, right). An estimate $\hat{\varepsilon}$ of the noise level ε for every azimuth–range pixel can be obtained via

$$\hat{\varepsilon} = \sqrt{\frac{N}{n_\varepsilon} \sum_{n=N-n_\varepsilon+1}^{N} |\beta_n|^2}. \qquad (13)$$

C. Extension to D-TomoSAR

Taking the motion term into account, the system model (4) can be extended to

$$g_n = \int_{\Delta s} \gamma(s) \exp\left(-j2\pi\left(\xi_n s + \eta_n V(s)\right)\right) \, ds,$$
$$n = 1, \ldots, N \quad (14)$$

where $V(s)$ is the deformation LOS velocity profile along elevation and $\eta_n = 2t_n/\lambda$ may, in analogy, be called a "velocity frequency." Formally, (14) can be rewritten as

$$g_n = \int_{\Delta v} \int_{\Delta s} \gamma(s)\delta\left(v - V(s)\right) \exp\left(-j2\pi(\xi_n s + \eta_n v)\right) \, ds \, dv,$$
$$n = 1, \ldots, N \quad (15)$$

where Δv is the velocity range, which is typically on the order of some tens of centimeters per year. Equation (12) is a 2-D Fourier transform of $\gamma(s)\delta(v - V(s))$, which is a delta line in the elevation–velocity ($s-v$) plane along $v = V(s)$. Projected onto the elevation axis, $\gamma(s)\delta(v - V(s))$ follows the reflectivity profile $\gamma(s)$. If we accept $\gamma(s)\delta(v - V(s))$ as the object to be reconstructed, the SVD-based spectral estimation methods used for TomoSAR can be easily extended to the 4-D case that includes the LOS deformation terms, i.e., D-TomoSAR [2], [10]. Instead of treating the deformation phase term as noise, D-TomoSAR can provide retrieval of the elevation and deformation information of multiple scatterers inside an azimuth–range resolution cell and thus obtain a 4-D map of

scatterers. It is required for reliable 3-D and 4-D city mapping from repeat-pass acquisitions.

Likewise, due to the ill-conditioning of the problem, regularization tools, such as TSVD and Wiener filtering, can be implemented, as described earlier. Fig. 6 shows an example. Two scatterers with an elevation distance of 20 m ($s_1 = 0$ m; $s_2 = 20$ m) are simulated. With standard 3-D TomoSAR, they are not separable with our baseline configuration since they are within a 40-m elevation resolution element, as shown in the left plot of Fig. 6. However, due to their different velocities ($V_1 = 0$ cm/year; $V_2 = -2$ cm/year), they can be easily distinguished in the $s-v$ plane, as shown in the right panel of Fig. 6.

V. MODEL ORDER SELECTION

By implementing nonparametric spectral estimation, the reflected power along the elevation direction can be extracted. Model selection schemes aim at estimating the number of point scatterers along elevation inside an azimuth–range pixel [29].

Let k be a parameter that defines the complexity of the model. In our case, k is the number of parameters to describe $\boldsymbol{\gamma}$. It depends on the number of scatterers n_p in the azimuth–range pixel. As each scatterer can be described by three parameters (amplitude, phase, and elevation), $k = 3 n_p$. Let further $\boldsymbol{\theta}(k)$ be the vector of the unknown amplitudes, phases, and elevations for all the n_p scatterers. Then, the reflectivity profile can be written as $\boldsymbol{\gamma}(\boldsymbol{\theta}(k))$. The relationship between $\boldsymbol{\gamma}(\boldsymbol{\theta}(k))$ and the observed data $\mathbf{g}$ is described by the observation model [(6)]. $p(\mathbf{g}|\boldsymbol{\theta}(k), k)$ is the likelihood function. Under the assumption that the model errors or disturbances $\mathbf{e} = \mathbf{g} - \mathbf{R} \, \boldsymbol{\gamma}(\boldsymbol{\theta}(k))$ are circular Gaussian distributed with zero mean and a covariance matrix of $\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = |\varepsilon|^2 \mathbf{I}$, the likelihood function can be written as

$$p\left(\mathbf{g}|\boldsymbol{\theta}(k), k\right) = \frac{1}{\pi^N |\varepsilon|^{2N}} \exp\left(-\frac{1}{|\varepsilon|^2} \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\left(\boldsymbol{\theta}(k)\right)\|^2\right). \quad (16)$$

It will increase with increasing k, since a more complex model fits the observations better. As a consequence, maximization of the likelihood function is not sufficient for model selection. Instead of using only the likelihood as a criterion, *penalized* likelihood criteria are used for model selection. The general form of penalized likelihood criteria is

$$\hat{\boldsymbol{\theta}}(k) = \arg \max_{\theta(k)} \left\{\ln p\left(\mathbf{g}|\boldsymbol{\theta}(k), k\right) - C\left(\boldsymbol{\theta}(k)\right)\right\}. \quad (17)$$

$\ln p(\mathbf{g}|\boldsymbol{\theta}(k), k)$ is the log-likelihood and $C(\boldsymbol{\theta}(k))$ is a complexity penalty, from which we can see that model selection is actually a tradeoff between how well the model fits the data and the complexity of the model. Note that the likelihood depends on the noise model, e.g., for Gaussian noise, the log-likelihood is essentially the sum of squared residuals. If this term only depends on the model dimension, then

$$\hat{k} = \arg \max_{k} \left\{\ln p\left(\mathbf{g}|\hat{\boldsymbol{\theta}}(k), k\right) - C(k)\right\}. \quad (18)$$

In other words, estimate the best parameters for each k, and then choose among these models. In our application, model
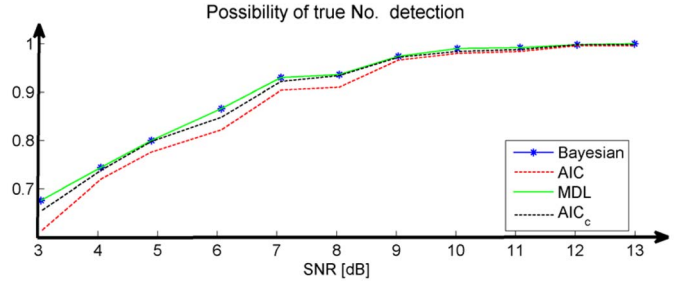
Fig. 7. Probability of correctly detecting two scatterers using different model selection schemes.

complexity only depends on the number of scatterers n_p in the azimuth–range pixel. It is common in the literature to multiply the cost function by a factor of minus two

$$\hat{k} = \arg \min_{k} \left\{-2 \ln p\left(\mathbf{g}|\hat{\boldsymbol{\theta}}(k), k\right) + 2C(k)\right\}. \quad (19)$$

For each $k = 3 n_p$ (e.g., $n_p = 1, 2$, or 3), the estimated amplitudes, phases, and elevations of the n_p scatterers are used to synthesize an estimate of $\boldsymbol{\gamma}$ and to compute $\|\mathbf{g} - \mathbf{R} \, \hat{\boldsymbol{\gamma}}(\hat{\boldsymbol{\theta}}(k))\|^2$, the exponent of the likelihood function. The preferred model is finally the one with the lowest penalized likelihood criterion value according to (15).

There are many types of penalized likelihood criteria, such as the Bayesian information criterion (BIC), the Akaike information criterion (AIC), and the minimum description length (MDL). Their basic principles are the same, and the main difference is in the penalty term. In [30], BIC, MDL, and AIC are discussed in detail for the purpose of determining the number of scatterers inside an azimuth–range pixel of multipass SAR data stack with nine acquisitions.

BIC is also called the Schwarz criterion or Schwarz information criterion (SIC). It is so named because Gideon E. Schwarz [31] gave a Bayesian argument for adopting it. The detailed derivation and performance of BIC are described in [32]. If the models are quasi-nested [33], BIC with $C(k) = 0.5 \, k \ln N$, where N refers to the number of samples (in our case, the number of acquisitions), is an approximation of the Bayesian method that says that models should be compared according to their posterior probabilities.

MDL is a formalization of Occam's razor and tries to find the hypotheses or combination of hypotheses that compress the data the most [34]. The MDL was introduced by Jorma Rissanen in 1978 [35]; it is an important concept in information and learning theory. Without prior knowledge of the model, it is identical to BIC with a penalty term of $C(k) = 0.5 \, k \ln N$.

AIC with $C(k) = k$ tries to minimize the expected relative distance between the fitted model and the unknown true mechanism that generated the observed data [36]. Rejecting a null hypothesis when it should have been accepted creates a type I error; accepting a null hypothesis when it should have been rejected creates a type II error. AIC effectively trades off those two types of errors. As a result, AIC may give less weight to simplicity than to data fit as compared to classical hypothesis testing [37]. Therefore, when the number of samples is large, AIC tends to underpenalize complexity.
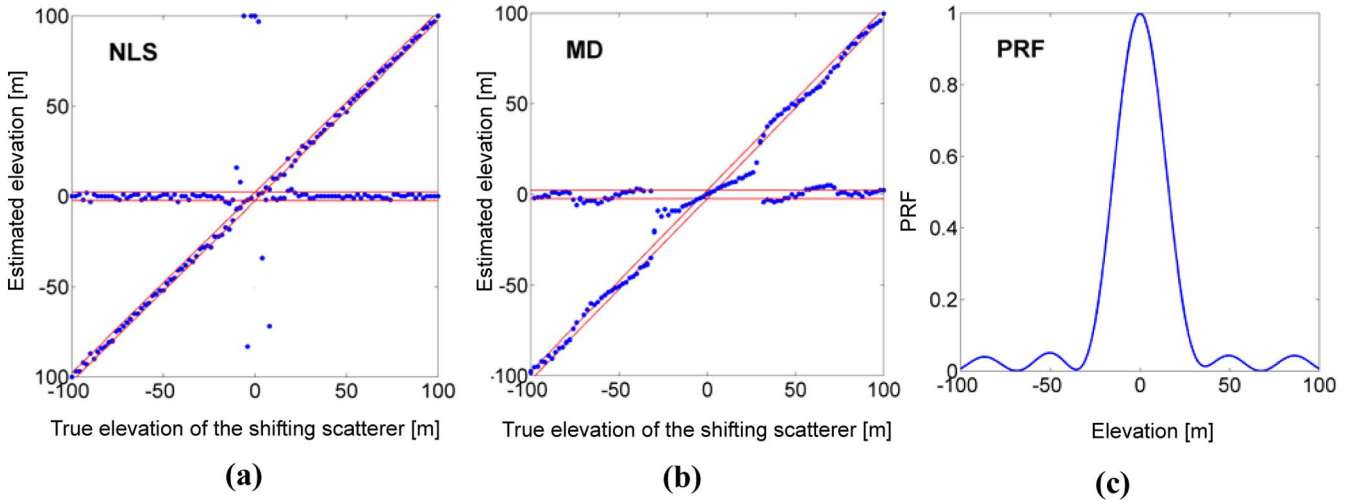
Fig. 8. Estimated elevations using (a) NLS estimation and (b) MD with SNR = 20 dB. (c) Elevation point spread function from SVD-Wiener reconstruction. The estimation truth is a horizontal line referring to the ground and a diagonal line referring to the scatterer at variable elevation. The red lines in the plots show ±3 times the CRLB on elevation estimates for the single scatterer.
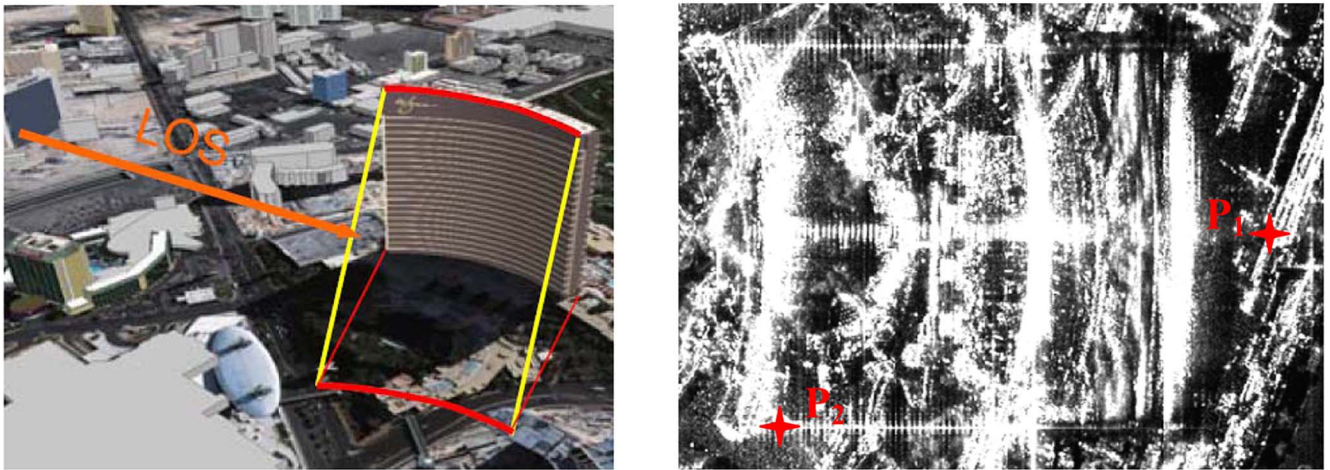

Fig. 9. Wynn hotel, Las Vegas. (Left) Optical image with viewing direction of SAR (LOS = line of sight) and (yellow) two iso-elevation lines. (Right) Mean TerraSAR-X intensity image with (red) analysis points P1 and P2.
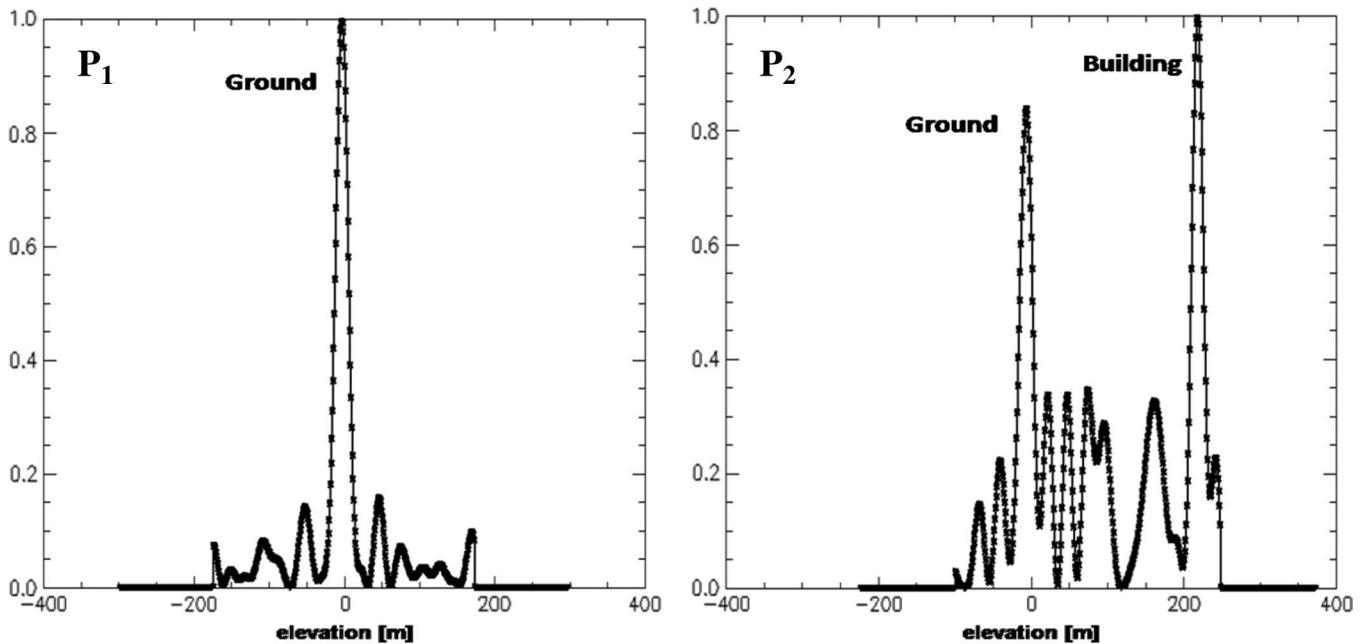
Fig. 10. Elevation profiles at analysis points P1 (single scatterer) and P2 (two scatterers) marked in Fig. 11.
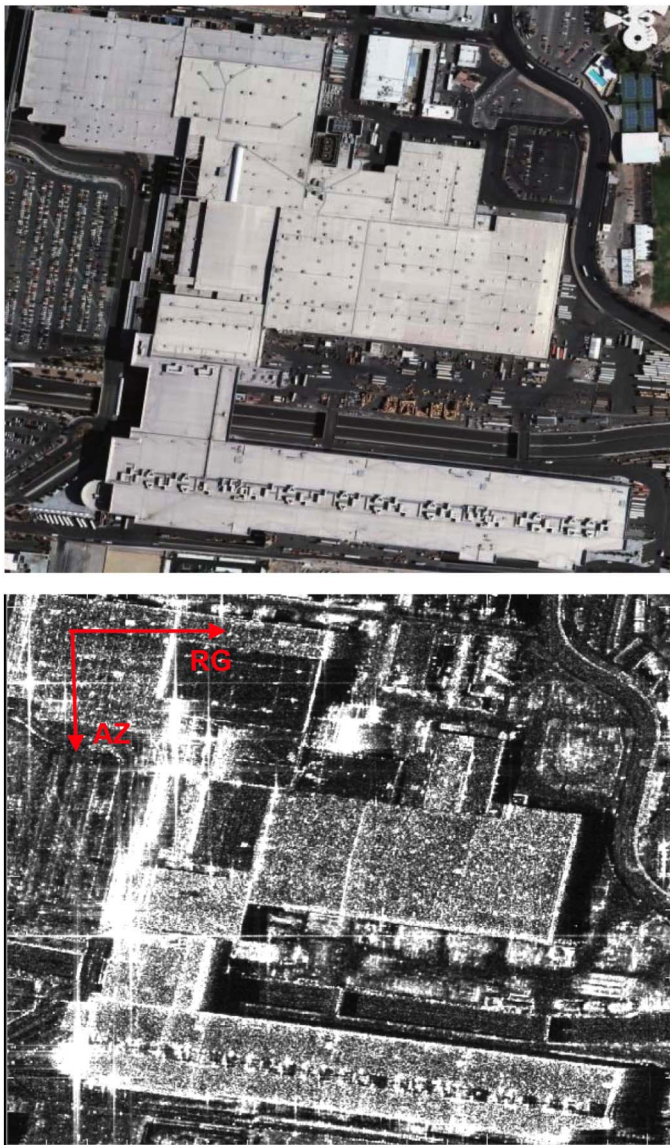
Fig. 11. (Top) Las Vegas Convention Center (Google Earth). (Bottom) TerraSAR-X intensity map.

AICc is the small-sample bias adjustment for AIC [38]

$$AICc = AIC + \frac{2k(k+1)}{N-k-1}. \qquad (20)$$

An issue with all these model selection methods is selection bias, which cannot be easily corrected. Selection bias refers to the fact that model criteria are particularly risky when a selection is made from a large number of competing models. The random fluctuation in the data will increase the scores of some models more than others. The more models there are, the greater is the risk that the optimal model is chosen at random. In [36], it was emphasized that selection criteria should not be followed blindly and that the term "selection" suggests something definite, which, in fact, has not been reached. Therefore, the selection of a criterion must be based on the experiments for the specific situation, and it is also not possible for model order selection algorithms to control the false-alarm probability.

In order to choose a favorable model selection method, different schemes are evaluated on simulated data with the elevation aperture sampling of Fig. 2. The decorrelation effect is simulated by adding Gaussian noise with a certain SNR. Phase noise due to unmodeled deformation and atmospheric effects are simulated by adding a uniformly distributed phase on $[-0.5\pi, 0.5\pi]$. The simulated "truth" is two scatterers at elevations of -20 and 40 m with reflectivities of 1 and 0.8, respectively. The distance between the two scatterers is hence 1.5 elevation resolution cells. The most important characteristic for evaluating the performance of the model selection criteria is the detection rate that refers here to the probability of correctly detecting the number of scatterers. A Monte Carlo simulation with 1000 realizations per SNR value was performed to evaluate the detection rates of different schemes. The probability of correctly detecting two scatterers for various SNRs is shown in Fig. 7.

In this example, the number of scatterers is chosen from three hypotheses, namely, one, two, or three scatterers. Since the complexity of the model only depends on k, the MDL is identical to BIC in our application, and they give the same results. All the model selection schemes appear to have similar performance. At a typical SNR of 3 dB, a threshold that is also often used for persistent scatterer identification, the probability of correctly detecting the number of scatterers is at least 60%. Overall, the MDL and BIC provide the best performance in our case and are used in the following experiments.

VI. PARAMETRIC ESTIMATION

With prior knowledge n_p, the parameters associated with individual scatterers, such as elevation, reflectivity, and LOS velocity (for D-TomoSAR only), can be estimated by either implementing a parametric estimation method, such as NLS, or simply detecting n_p peaks of the nonparametric estimates.

For the following, let us consider that the elevation reflectivity profile is composed of n_p delta functions of complex amplitudes x_i

$$\gamma(s) = \sum_{i=1}^{n_p} x_i \cdot \delta(s - s_i). \qquad (21)$$

NLS: The noise-corrupted SAR observations of the n_p scatterers is according to the observation equation (4) [39]

$$g_n = \sum_{i=1}^{n_p} x_i \exp(j2\pi\xi_n s_i) + v_n, \qquad n = 1, \ldots, N \qquad (22)$$

where v_n is the observation noise. As both the complex amplitudes x_i and the elevations s_i of the individual n_p scatterers are unknown, the spectrum estimation problem is nonlinear. Although the signal model is nonlinear, it is at least linear in the amplitude

$$\mathbf{g} = \mathbf{H}(s)\mathbf{x} + \mathbf{u} \qquad (23)$$
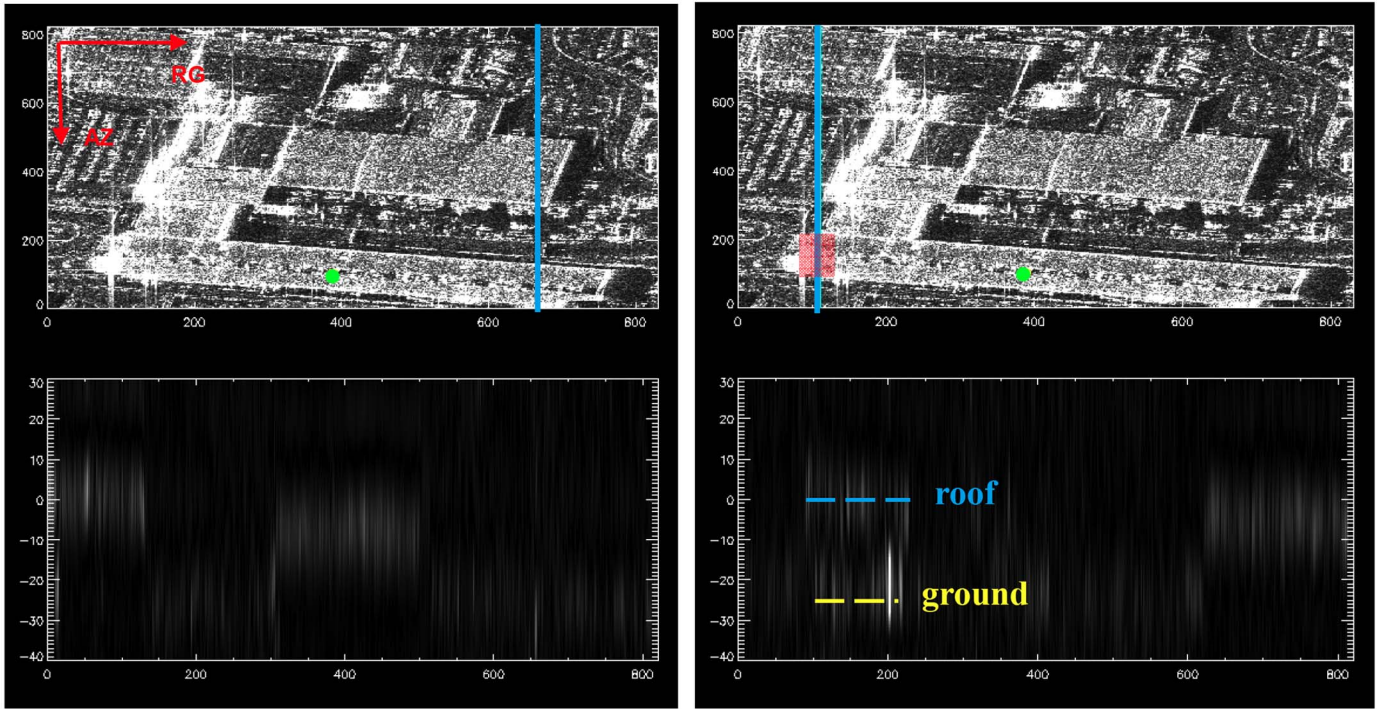
Fig. 12. (Top) TerraSAR-X radar intensity image of the Las Vegas Convention Center. The green dot is the reference point, and the blue lines are the positions of the respective slices. (Bottom) The estimated reflectivity is shown in the azimuth–elevation plane [horizontal: azimuth; vertical: elevation, converted to height (in meters)].

with

$$
\mathbf{g} = \begin{bmatrix} g_1 \\ . \\ . \\ . \\ g_N \end{bmatrix}_{N \times 1} \qquad \mathbf{H}(s) = \begin{bmatrix} e^{j 2\pi \xi_1 s_1} & \cdots & e^{j 2\pi \xi_1 s_{n_p}} \\ . & & . \\ . & & . \\ . & & . \\ e^{j 2\pi \xi_N s_1} & \cdots & e^{j 2\pi \xi_N s_{n_p}} \end{bmatrix}_{N \times n_p}
$$

$$
\mathbf{x} = \begin{bmatrix} x_1 \\ . \\ . \\ . \\ x_{n_p} \end{bmatrix}_{n_p \times 1} \qquad \mathbf{u} = \begin{bmatrix} v_1 \\ . \\ . \\ . \\ v_N \end{bmatrix}_{N \times 1}
$$

where the matrix $\mathbf{H}(s)$ depends on the unknown elevations of the scatterers. As this model is linear in amplitude and nonlinear in elevation, the least squares error may be minimized with respect to $\mathbf{x}$ in a closed analytic form and thus be reduced to a function of elevations only, which means that an n_p-dimensional search is needed (for D-TomoSAR, an additional search for velocity is required). Since the object function [40] $\mathbf{x}$, which minimizes

$$
J(s, \mathbf{x}) = (\mathbf{g} - \mathbf{H}(s)\mathbf{x})^{\mathrm{T}} (\mathbf{g} - \mathbf{H}(s)\mathbf{x}) \tag{24}
$$

for a given s, is

$$
\hat{\mathbf{x}} = \left(\mathbf{H}^{\mathrm{T}}(s)\mathbf{H}(s)\right)^{-1} \mathbf{H}^{\mathrm{T}}(s)\mathbf{g} \tag{25}
$$

the resulting error is

$$
J(s, \hat{\mathbf{x}}) = \mathbf{g}^{\mathrm{T}} \left(\mathbf{I} - \mathbf{H}(s) \left(\mathbf{H}^{\mathrm{T}}(s)\mathbf{H}(s)\right)^{-1} \mathbf{H}^{\mathrm{T}}(s)\right) \mathbf{g}. \tag{26}
$$
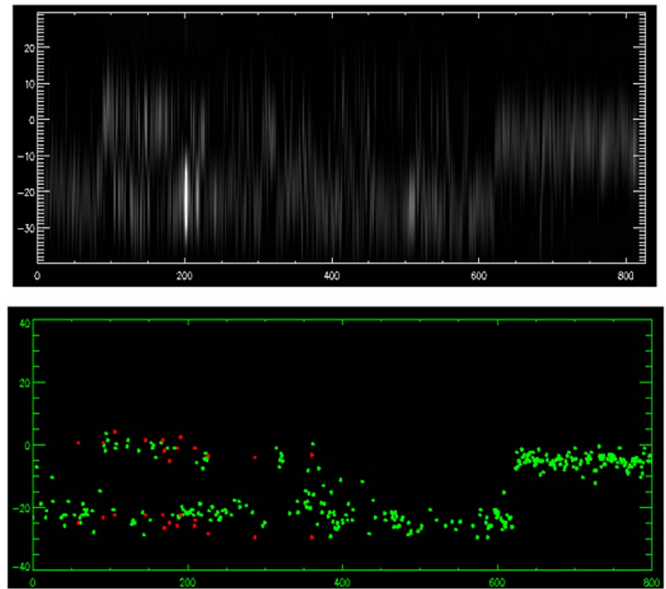
Fig. 13. (Top) Reconstructed reflectivity slice (azimuth–elevation plane). (Bottom) Estimated elevation (converted to height) and model selection results (green dots: single scatterer in an azimuth–range pixel; red dots: two scatterers).

The problem now reduces to a maximization of $\mathbf{g}^{\mathrm{T}}\mathbf{H}(s)(\mathbf{H}^{\mathrm{T}}(s)\mathbf{H}(s))^{-1}\mathbf{H}^{\mathrm{T}}(s)$ $\mathbf{g}$ over n_p values of s, and a grid search can be used. With Gaussian white noise, NLS is identical to MLE. It is therefore theoretically the best estimator for our application if and only if the data closely agree with the assumed model. However, multidimensional search leads to a large computational effort.

Maxima detection (MD): As implied by the name, MD detects n_p peaks in the nonparametric estimates. It is relatively

fast. However, it may introduce estimation bias caused by interference between multiple scatterers, i.e., the location of the peaks may be shifted by the sidelobes of other scatterer responses. In addition, it has the inherent elevation resolution limit of nonparametric spectral estimation and is therefore not capable of separating too close objects.

Two scatterers inside one resolution cell, among which one changes the elevation gradually from -100 to 100 m and the other stays at zero elevation, are simulated as an example for evaluating the performance of both methods. Fig. 8 shows the estimated elevation of the two scatterers by NLS [Fig. 8(a)] and MD [Fig. 8(b)] with SNR = 20 dB. The x-axis refers to the true elevation of the shifting scatterer. The y-axis refers to the estimated elevation. The ideal image should be two straight lines (one horizontal and one diagonal).

The red lines in the plots indicate ± 3 times the CRLB $\sigma_{\hat{s}}$ on elevation estimates for the single-scatterer case [(3)]. Fig. 8(c) shows the elevation PRF with elevation aperture sampling shown in Fig. 2. As expected, NLS gives the better performance. When the distance between the two scatterers is large enough (low interference effect), NLS is able to locate the scatterers quite well with localization accuracy within the $3\sigma_{\hat{s}}$ band. Correspondingly, MD of nonparametric estimates is limited by the elevation resolution. Moreover, even with the two scatterers further apart than the resolution cell size, MD suffers from the interference of the scatterers [Fig. 8(b)]. The elevation estimate of one scatterer is systematically biased by the side-lobes of the other and vice versa, even though the SNR is high.

VII. EXPERIMENTAL RESULTS

A. Las Vegas, Wynn Hotel

The Las Vegas Wynn hotel has been chosen as a test building to demonstrate the potentials of layover object separation since it is very high and has strong layover effect in the SAR image. Fig. 9 (left) shows the Wynn hotel in Las Vegas with a height close to 200 m, corresponding to an elevation range of 380 m. The right image in Fig. 9 is the corresponding TerraSAR-X intensity image.

Pixels containing multiple scatterers are mainly located at the intersection of the bright texture of the building with structures near the ground. To exemplify the potential of the TomoSAR method, the two pixels marked by red stars have been selected and will be analyzed in the following. As P1 is located outside the region of the high-rise building, it is expected that it only contains a single scatterer situated near the ground. P2 is located at the intersection area; we expect two scatterers inside this pixel, among them one from the ground and one from the building facade. The corresponding reflectivity estimates for those two pixels are shown in the right image of Fig. 10. In this example, we can see the potential of the tomographic approach with TerraSAR-X to separate multiple scatterers in layover areas at a VHR.

B. Las Vegas Convention Center

The Las Vegas Convention Center is a very interesting test building for 3-D focusing for two reasons. First, it is very big and has a regular shape. Therefore, we are able to check the plausibility of the results. Second, it has a height of about 20 m,
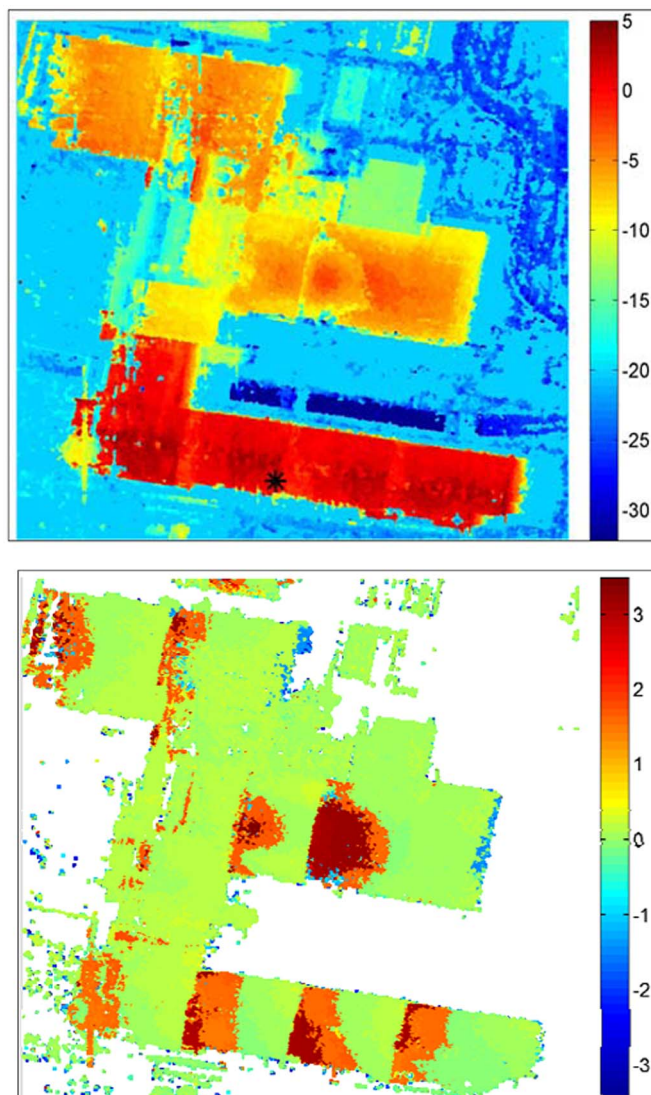
Fig. 14. (Top) Digital surface model of the Las Vegas convention center generated from D-TomoSAR (in meters) (black cross: reference point) and (Bottom) the estimated linear deformation velocity w.r.t. the reference point (in centimeters per year).

the critical distinguishable distance between two scatterers (one from the ground and the other from the building) for our elevation aperture size. The presence of two scatterers within azimuth–range pixels is expected in layover areas. The top image in Fig. 11 shows the convention center visualized in Google Earth. The bottom image is the TerraSAR-X intensity map of the area. After preprocessing, we choose a reference pixel according to [41], which most likely has only a single scatterer inside.

SVD-Wiener, as described in Section IV, is applied to each azimuth–range pixel in the area of interest. Fig. 12 shows an example of the reconstructed reflectivity map of the azimuth–elevation slices with fixed range coordinates marked by bright blue lines. The green dot on the TS-X intensity map is the selected reference that is located on the roof of the convention center. The lower images of Fig. 12 show the reconstructed reflectivity slice (bright means high reflectivity; the y-axis refers to the elevation relative to the reference point). The structure of the building can be recognized, and
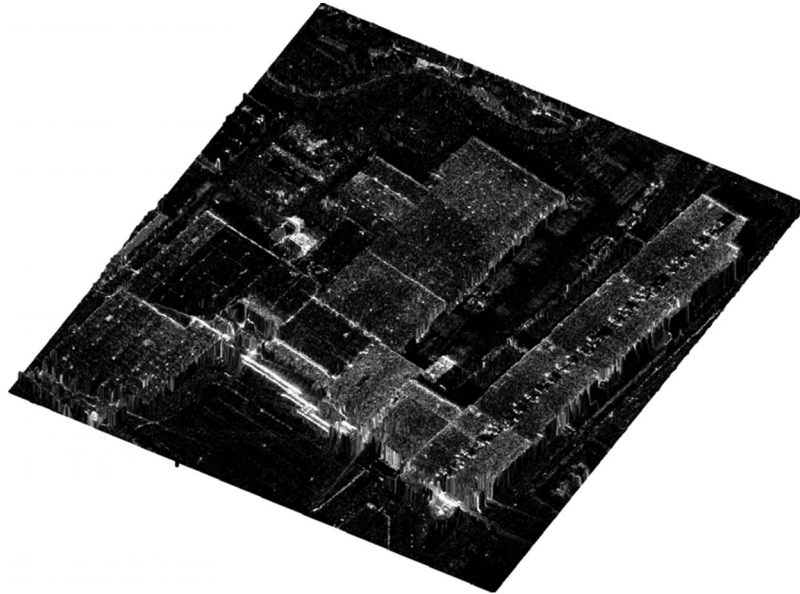
Fig. 15.   World at X-band: Tomographic surface reconstruction overlaid by the 3-D reflectivity map.

the reflectivity difference between individual parts can be seen. For instance, from the left slice, it is obvious that the returns from the roof of the building are generally stronger than the returns from the ground and the building has a different height for individual parts (which is not visible in the Google Earth building model). At the range position of the slice on the right-hand side, the lower part of the building marked by the red block (left part in the reflectivity slice) is a layover area. From the Google Earth image, we can see that there is a small triangular-shaped plaza on the ground made of the same material as the building. Thereby, multiple scatterers are expected. When we check the estimated reflectivity slice, multiple scatterers appear, even though the distance between two scatterers (one from the building marked with a blue line, and the other from the small structures on the ground marked with a yellow line) is approximately at the minimal distinguishable distance. This demonstrates the multiple-scatterer separation potential of TomoSAR and the stability of our algorithm.

After model selection, the number of scatterers inside every azimuth–range pixel is retrieved to provide the required prior knowledge for parameter estimation. The top panel of Fig. 13 shows the reflectivity slice reconstructed by the estimated elevation and reflectivity based on the nonparametric estimates on the right image in Fig. 12. The bottom plot in Fig. 13 reveals the estimated elevations of azimuth–range pixels with good signal. The green dots represent the elevation of a *single* scatterer in an azimuth–range pixel. The red dots represent the elevation estimates of the detected *double* scatterers. These results seem plausible. The deformation term is ignored for these results.

From the parametric elevation estimates, a digital surface model of the building can be reconstructed. Even ambiguities in layover areas are resolved. However, deformation has not been considered so far. Therefore, when the deformation signal is strong, the reflectivity reconstruction in the elevation will be distorted, and the elevation estimates degrade. Therefore, we implement D-TomoSAR to our data stack.

Fig. 14 shows the surface model generated from the elevation estimates (converted to height relative to the reference point). The full structure of the convention center has been captured at a very detailed level. Other than the building, more details, such as the roads surrounding the convention center, as well as two bridges above the roads that have weak but correlated returns, are well resolved. Compared to classical interferometric surface reconstruction, TomoSAR overcomes layover and phase-unwrapping problems. The height estimates are very precise compared to the 40.5-m elevation resolution due to the high SNR of TerraSAR-X data. There are still some distortions in the middle of the image where a flat roof surface is expected. It may be due to the incorrect linear deformation model assumption. Since the deformation is presumably caused by thermal dilation, it rather follows a periodic seasonal model. The bottom image of Fig. 14 shows the estimated deformation velocity relative to the reference point. Some areas exhibit a significant uplift of up to 3 cm/year.

Fig. 15 shows the final surface model with deformation correction overlaid by the reflectivity map. This, for the first time, visualizes in detail how the convention center would look like from the position of TerraSAR-X if our eyes could see X-band radiation. This may lead to a better understanding of the nature of scattering. For instance, an overview about the multiple bounce can be acquired by looking at the very bright structure in Fig. 15. Also, the very bright individual scatterers that behave as corner reflectors can be precisely located. This may help in finding natural corner reflectors.

VIII. CONCLUSION

This paper has demonstrated the potential of the new class of VHR spaceborne SAR systems, like TerraSAR-X and COSMO-Skymed, for TomoSAR in urban environment. A stack of TerraSAR-X high-resolution spotlight data over the city of Las Vegas has been used. Compared to the medium resolution SAR systems available so far, the information content

and level of detail have increased dramatically. A full tomographic high-resolution reconstruction of a building complex has been presented.

Depending on the application, nonparametric or parametric estimation methods are preferred. Examples for both classes of estimators have been demonstrated and compared to the Cramér–Rao bound. Motivated by the ill-determinedness of the problem, a MAP estimator has been proposed, which leads to a Wiener-type regularization for the nonparametric SVD method for both 3-D and 4-D (differential) tomographic reconstructions. Model selection, i.e., the estimation of the number of discrete scatterers in a resolution cell, has been shown to be a necessary prerequisite for parametric estimation.

One of the major error sources is unmodeled, e.g., nonlinear, motion. These phase errors are able to deteriorate the elevation estimates. With the launch of TanDEM-X, single-pass (motion-free) data pairs will be available. Mixed data stacks from TerraSAR-X and TanDEM-X will be an attractive option for mitigating motion errors and for retrieving profiles of temporarily decorrelated scatterers.

REFERENCES

[1] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.

[2] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," in *Proc. IGARSS*, Toulouse, France, 2003, pp. 1206–1208.

[3] P. Pasquali, C. Prati, F. Rocca, and M. Seymour, "A 3-D SAR experiment with EMSL data," in *Proc. IGARSS*, 1995, pp. 784–786.

[4] S. Sauer, L. Ferro-Famil, A. Reigber, and E. Pottier, "3D urban remote sensing using dual-baseline POL-InSAR images at L-band," in *Proc. IGARSS*, 2008, pp. IV-145–IV-148.

[5] J. Homer, D. I. Longstaff, Z. She, and D. Gray, "High resolution 3-D imaging via multi-pass SAR," *Proc. Inst. Elect. Eng.—Pt. F*, vol. 149, no. 1, pp. 45–50, Feb. 2002.

[6] Z. She, D. Gray, R. E. Bogner, J. Homer, and D. Longstaff, "Three-dimensional spaceborne synthetic aperture radar (SAR) imaging with multipass processing," *Int. J. Remote Sens.*, vol. 23, no. 20, pp. 4357–4382, Oct. 2002.

[7] F. Fornaro, F. Serafino, and F. Lombardini, "Three-dimensional multipass SAR focusing: Experiments with long-term spaceborne data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 702–714, Apr. 2005.

[8] X. Zhu, N. Adam, and R. Bamler, "First demonstration of spaceborne high resolution SAR tomography in urban environment using TerraSAR-X data," in *Proc. CEOS SAR Workshop Calibration Validation*, 2008.

[9] G. Fornaro and F. Serafino, "Imaging of single and double scatterers in urban areas via SAR tomography," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 12, pp. 3497–3505, Dec. 2006.

[10] G. Fornaro, D. Reale, and F. Serafino, "Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 12, pp. 224–237, Dec. 2009.

[11] P. C. Hansen, *Regularization Tools*, 1992. [Online]. Available: http://www.imm.dtu.dk/~pch/Regutools/

[12] M. Eineder, N. Adam, R. Bamler, and N. Yague-Martinez, "Spaceborne spotlight SAR interferometry with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1524–1535, May 2009.

[13] S. D'Amico, C. Arbinger, M. Kirschner, and S. Campagnola, "Generation of an optimum target trajectory for the TerraSAR-X repeat observation satellite," in *Proc. 18th Int. Symp. Space Flight Dynamics*, Munich, Germany, 2004, p. 137.

[14] R. Bamler, M. Eineder, N. Adam, X. Zhu, and S. Gernhardt, "Interferometric potential of high resolution spaceborne SAR," *Photogramm. Fernerkundung Geoinf.*, vol. 2009, no. 5, pp. 407–419, Nov. 2009.

[15] S. Buckreuss, private communication, Aug. 31, 2009.

[16] G. Fornaro, F. Serafino, and F. Soldovieri, "Three-dimensional focusing with multipass SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 507–517, Mar. 2003.

[17] B. Kampes, *Radar Interferometry, The Persistent Scatterer Technique,* 1st ed. New York: Springer-Verlag, 2006, ser. Remote Sensing and Digital Image Processing Series.

[18] F. Meyer, S. Gernhardt, and N. Adam, "Long-term and seasonal subsidence rates in urban areas from persistent scatterer interferometry," in *Proc. URBAN*, Paris, France, 2007.

[19] C. Cafforio, C. Prati, and F. Rocca, "SAR data focusing using seismic migration techniques," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 2, pp. 194–207, Mar. 1991.

[20] N. Adam, B. Kampes, M. Eineder, J. Worawattanamateekul, and M. Kircher, "The development of a scientific permanent scatterer system," presented at the Int. Society Photogrammetry Remote Sensing (ISPRS) Hannover Workshop, Hannover, Germany, 2003.

[21] F. Lombardini and A. Reigber, "Adaptive spectral estimation for multibaseline SAR tomography with airborne L-band data," in *Proc. IEEE IGARSS*, 2003, vol. 3, pp. 2014–2016.

[22] S. Guillaso and A. Reigber, "Polarimetric SAR Tomography (POLTOMSAR)," in *Proc. POLINSAR*, Frascati, Italy, 2005.

[23] F. Gini and F. Lombardini, "Multibaseline cross-track SAR interferometry: A signal processing perspective," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 20, no. 8, pp. 71–93, Aug. 2005.

[24] F. Lombardini and M. Pardini, "Detection of scatterer multiplicity in spaceborne SAR tomography with array errors," in *Proc. IEEE Radar Conf.*, May 4–8, 2009, pp. 1–6.

[25] P. C. Hansen, "The truncated SVD as a method for regularization," *BIT Numer. Math.*, vol. 27, no. 4, pp. 534–553, Oct. 1987.

[26] G. H. Golub, V. Klema, and G. W. Stewart, "Rank degeneracy and least squares problems," Dept. Comput. Sci., Univ. Maryland, College Park, MD, Tech. Rep. TR-456, 1976.

[27] H. P. William, T. A. Saul, V. T. William, and P. F. Brian, *Numerical Recipes in C,* 2nd ed. New York: Press Syndicate Univ. Cambridge, 1992.

[28] X. Zhu, N. Adam, and R. Bamler, "Space-borne high resolution tomographic interferometry," in *Proc. IEEE IGARSS*, Cape Town, Africa, 2009, pp. IV-869–IV-872.

[29] F. Lombardini and F. Gini, "Model order selection in multi-baseline interferometric radar systems," *EURASIP J. Appl. Signal Process.*, vol. 2005, no. 20, pp. 3206–3219, 2005.

[30] X. Zhu, "High-resolution spaceborne radar tomography," M.S. thesis, Lehrstuhl für Photogrammetrie und Fernerkundung, Tech. Univ. München, Lehrstuhlbibliothek, München, Germany, 2008.

[31] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, no. 2, pp. 461–464, Mar. 1978.

[32] K. P. Burnham and D. R. Anderson, "Multimodel inference: Understanding AIC and BIC in model selection," *Sociol. Methods Res.*, vol. 33, no. 2, pp. 261–304, Nov. 2004.

[33] M. R. Forster, "Key concepts in model selection: Performance and generalizability," *J. Math. Psychol.*, vol. 44, no. 1, pp. 205–231, Mar. 2000.

[34] J. Rissanen, "A short introduction to Model Selection," *Kolmogorov Complexity and Minimum Description Length*, 2005. [Online]. Available: www.mdl-research.org

[35] J. Rissanen, "Modelling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, Sep. 1978.

[36] M. Browne, "Cross-validation methods," *J. Math. Psychol.*, vol. 44, no. 1, pp. 108–132, Mar. 2000.

[37] Y. Sakamoto, M. Ishiguro, and G. Kitagawa, *Akaike Information Criterion Statistics.* Dordrecht, The Netherlands: Reidel, 1986.

[38] Y. Sakamoto, *Categorical Data Analysis by AIC.* Tokyo, Japan: KTK Scientific Publishers, 1991.

[39] X. Zhu, N. Adam, R. Brcic, and R. Bamler, "Space-borne high resolution SAR tomography: Experiments in urban environment using TerraSAR-X data," in *Proc. JURSE*, 2009.

[40] S. M. Kay, *Fundamentals of Statistical Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[41] N. Adam, R. Bamler, M. Eineder, and B. Kampes, "Parametric estimation and model selection based on amplitude-only data in PS-interferometry," presented at the FRINGE Workshop, Frascati, Italy, 2005. CD-ROM.

Xiao Xiang Zhu (S'10) was born in Changsha, China, on December 12, 1984. She received the B.S. degree in space engineering from the National University of Defense Technology, Changsha, in 2006 and the M.Sc. degree in earth oriented space science and technology (ESPACE) from the Technische Universität München (TUM), München, Germany, in 2008. She is curently working toward the Ph.D. degree in differential synthetic aperture radar (SAR) tomography using TerraSAR-X data at the Remote Sensing Technology, TUM.

Since May 2008, she has been a Full-Time Scientific Collaborator with Remote Sensing Technology, TUM. Her work is part of the project team "Dynamic Earth," which was established in the International Graduate School of Science and Engineering, TUM, as a result of the German Excellence Initiative in 2007. In October/November 2009, she was a Guest Scientist with the Institute for Electromagnetic Sensing of the Environment (IREA), Italian National Research Council (CNR), Naples, Italy. Her main research interests are in the signal processing field with applications to spaceborne SAR data, in particular SAR tomography, and differential SAR tomography.

Richard Bamler (M'95–SM'00–F'05) received the Diploma degree in electrical engineering, the Doctor of Engineering degree, and the "Habilitation" degree in signal and systems theory from the Technische Universität München (TUM), München, Germany, in 1980, 1986, and 1988, respectively.

During 1981 and 1989, he was with TUM, working on optical signal processing, holography, wave propagation, and tomography. In 1989, he joined the German Aerospace Center (DLR), Oberpfaffenhofen, Germany, where he is currently the Director of the Remote Sensing Technology Institute (IMF). Since then, he and his team have been working on synthetic aperture radar (SAR) signal processing algorithms (ERS, SIR-C/X-SAR, Radarsat, SRTM, ASAR, TerraSAR-X, and TanDEM-X), SAR calibration and product validation, SAR interferometry, phase unwrapping, estimation theory and model-based inversion methods for atmospheric sounding (GOME, SCIAMACHY, MIPAS, and GOME-2), and oceanography. In early 1994, he was a Visiting Scientist with the Jet Propulsion Laboratory in preparation of the SIC-C/X-SAR missions, and in 1996, he was a Guest Professor with the University of Innsbruck, Innsbruck, Austria. Since 2003, he has been a Full Professor in remote sensing technology with TUM. His current research interests are in algorithms for optimum information extraction from remote sensing data with emphasis on SAR, SAR interferometry, persistent scatterer interferometry, SAR tomography, and ground moving target indication for security-related applications. He and his team have developed and are currently developing the operational processor systems for the German missions TerraSAR-X, TanDEM-X, and EnMAP. He is the author of more than 160 scientific publications, among them about 40 journal papers, a book on multidimensional linear systems theory, and several patents on SAR signal processing.

A.3 **Reale, D., Fornaro, G., Pauciullo, A., Zhu, X., Bamler, R., 2011. Tomographic Imaging and Monitoring of Buildings With Very High Resolution SAR Data. IEEE Geoscience and Remote Sensing Letters 8 (4): 661-665**

Tomographic Imaging and Monitoring of Buildings With Very High Resolution SAR Data

D. Reale, G. Fornaro, *Senior Member, IEEE*, A. Pauciullo, X. Zhu, *Student Member, IEEE*, and R. Bamler, *Fellow, IEEE*

Abstract—Layover is frequent in imaging and monitoring with synthetic aperture radar (SAR) areas characterized by a high density of scatterers with steep topography, e.g., in urban environment. Using medium-resolution SAR data tomographic techniques has been proven to be capable of separating multiple scatterers interfering (in layover) in the same pixel. With the advent of the new generation of high-resolution sensors, the layover effect on buildings becomes more evident. In this letter, we exploit the potential of the 4-D imaging applied to a set of TerraSAR-X spotlight acquisitions. Results show that the combination of high-resolution data and advanced coherent processing techniques can lead to impressive reconstruction and monitoring capabilities of the whole 3-D structure of buildings.

Index Terms—Differential interferometric synthetic aperture radar (DInSAR), multidimensional SAR processing, SAR tomography, synthetic aperture radar (SAR), TerraSAR-X (TSX), 4-D SAR imaging.

I. INTRODUCTION

INTERFEROMETRIC synthetic aperture radar (SAR) (InSAR) and differential InSAR (DInSAR), particularly multitemporal DInSAR, have been proven to be effective for accurate scatterer localization and monitoring of displacements [1], [2]. The high accuracy and spatial density of the measurements make these techniques cost effective compared to classical geodetic techniques, typically used in environmental risk monitoring.

The increase of the spatial resolution provides a tangible improvement in the monitoring capabilities: Most of the international space agencies have hence hugely invested in the launch of large bandwidth spaceborne SAR systems. The hardware improvement must be complemented by the development of processing techniques that are able to extract the highest possible information content from the data. In this sense, SAR

D. Reale, G. Fornaro, and A. Pauciullo are with the Institute for the Electromagnetic Sensing of the Environment (IREA), National Research Council (CNR), 80124 Napoli, Italy (e-mail: reale.d@irea.cnr.it; fornaro.g@irea.cnr.it; pauciullo.a@irea.cnr.it).

X. Zhu is with the Technische Universität München, Lehrstuhl für Methodik der Fernerkundung, 80333 Munich, Germany (e-mail: xiaoxiang.zhu@bv.tum.de).

R. Bamler is with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Oberpfaffenhofen, Germany, and also with the Technische Universität München, Lehrstuhl für Methodik der Fernerkundung, 80333 Munich, Germany (e-mail: richard.bamler@dlr.de).

tomography, also known as multidimensional (3-D and 4-D) imaging SAR (MDI-SAR), is recognized as a powerful technique that extends interferometry.

DInSAR and persistent scatterer interferometry (PSI) assume the presence of only a single (dominant) scattering center in each pixel. However, SAR images of complex scenarios are affected by the interference between the responses of scatterers located at different elevations (slant heights). Standard multipass interferometric techniques "look" for the matching between the received signal and the "multipass signature" of a scatterer: The interference of responses may hence lead to misdetection of persistent scatterers and to height, velocity, and time-series measurement inaccuracies.

The layover effect causes interference between the responses of different scatterers. Layover is particularly critical in urban areas which are characterized by a high density of scatterers distributed on vertical structures.

As briefly explained next, MDI-SAR allows the overcoming of the single scatterer assumption and has opened a new scenario in the 3-D target reconstruction and monitoring with SAR systems [3], [4]. On medium-resolution systems, MDI-SAR imaging has already been proven to be effective in separating and monitoring scatterers in layover [5], [6].

The new generation of high-resolution SAR sensors, such as TerraSAR-X (TSX) and the COSMO-SkyMed constellation, allows the systematic acquisition of data with spatial resolution reaching metric/submetric values. The preliminary analysis of these images in dense urban areas has indicated that the resolution improvement brings layover of vertical structures to be more pronounced. On high-resolution SAR data, the interference between scatterers on the ground and on buildings is more frequent, and it is distributed on more pixels than on data acquired by medium-resolution satellites (e.g., European Remote Sensing (ERS) satellite or ENVISAT): The tomographic approach is a tool that allows mitigating this problem [7]. Moreover, the higher the resolution, the higher are the expectations for 3-D reconstruction on vertical structures.

In this letter, we investigate the application of SAR tomography to a real data set of TSX spotlight images over the city of Las Vegas, NV. The characteristics of this data set allow clear demonstration of the potential and the advantages offered by the SAR tomography technique.

II. LAYOVER AND TOMOGRAPHY

The imaging mechanism of radar is measuring the distances (range) of the scatterers from the sensor. If two scatterers are

Fig. 1.   Temporal multilook image of the spotlight TSX data set over the city of Las Vegas. The layover induced by the buildings is well recognizable.

Fig. 2.   Distribution of the acquisitions in the spatial/temporal baseline domain.

located at different positions but share the same range, they are imaged into the same pixel: This effect is known as layover.

In the presence of a vertical structure, such as a building, the radar signal is affected by layover between the ground, the façade, and possibly the roof. To have an idea of the effects of layover, a data set of 25 TSX spotlight images from ascending orbits over the city of Las Vegas, NV, has been considered. This celebrated city, located over a flat area, includes a high number of tall structures. The multitemporal averaged amplitude image is shown in Fig. 1: Tall buildings are well visible, although "folded" onto the ground toward the sensor; hence, their responses interfere with those of the targets located below the structures.

The tomography principle is simple: By using SAR data acquired from different elevations, an antenna along the slant height direction can be synthesized. The synthesized array brings resolving capabilities on the backscattering distribution along the elevation direction, orthogonal to the radar line of sight, and hence, it leads to the possibility to separate contributions coming from scatterers with different elevations and interfering in the same pixel [5]. The tomographic technique (3-D imaging) has been extended also to the time direction: The differential-tomography technique (also known as 4-D, i.e., space-velocity imaging) allows the separation of interfering scatterers and the measurement of their (possibly) different velocities [3] and time series [4].

MDI-SAR exploits both amplitude and phase information to reconstruct, for each pixel in the spatial (i.e., azimuth/ range) domain, the backscattering distribution along the slant height/mean deformation velocity plane. This fact already allows the improvement of performance in terms of dominant persistent scatterer detection with respect to classical PSI that uses only phase information [8]. In this letter, we limited the tomographic analysis to single (dominant) and double scatterers.
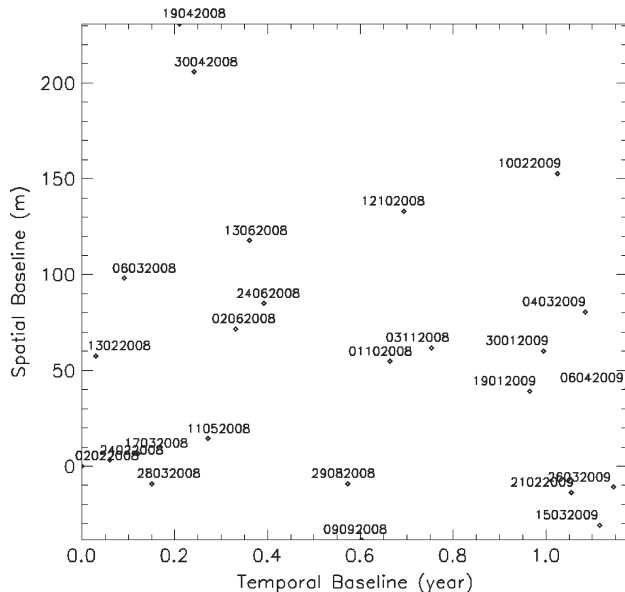
To search for single and double scatterers, we used the detection approach discussed in [9] and [10] based on the generalized likelihood ratio test. It exploits the detector for single scatterers in [9] in a sequential way and tests the energy contribution of the (possible) second scatterer after the cancellation of the dominant contribution: If this test declares the absence of the second scatterer, a second test on the presence of only one scatterer is carried out; see [10] for more details.

III. EXPERIMENTAL RESULTS

The TSX spotlight acquisition mode provides resolutions of 1.1 m in azimuth and 0.6 m in slant range. We applied the MDI technique to the area of Boulevard South, also known as "The Strip," where many of the largest hotels, casinos, and resorts are located. Almost all the images are acquired with the minimum repeat cycle of 11 days, from February 2008 to April 2009: Fig. 2 shows the baseline distribution. We note that, except for two acquisitions, the orbital tube is rather strict: The baseline span (B) is only approximately 207 m. This fact results in a poor slant height resolution of about $\delta_s = \lambda r / 2B \cong 47$ m, corresponding to a height resolution of $\delta_z = \delta_s \sin(\vartheta) \cong 27$ m, where λ, r, and ϑ are the wavelength, the distance from the scene center, and the look angle, respectively. Superresolution SAR tomography techniques could limit the effects of this poor resolution [7], [12]: In this letter, however, we limited our analysis to the classical linear tomographic approach [6].

The data set was calibrated for atmospheric phase components estimated via the low-resolution multipass DInSAR approach in [13] before the tomographic processing.

We focused our analysis on the block of the Mirage Hotel and Casino. It presents a tall (about 100 m) building surrounded by a lower flat structure (entertainment attractions) about 15–20 m over the street level.

Many features can be pointed out by comparing, in Fig. 3, the amplitude image of the area with an orthophotograph:

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

REALE *et al.*: IMAGING AND MONITORING OF BUILDINGS WITH VERY HIGH RESOLUTION SAR DATA 663
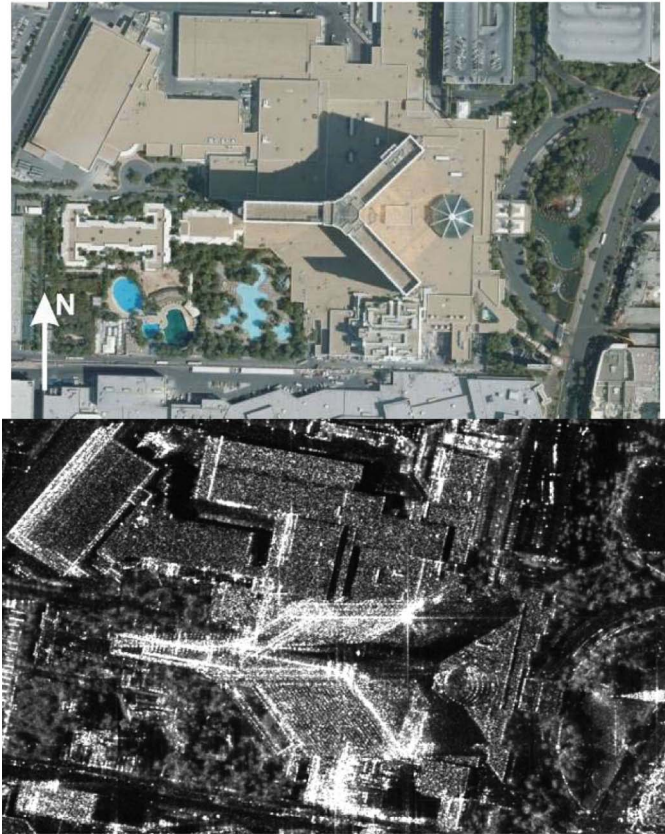
Fig. 3. (Top) Mirage Hotel image taken from Bing maps. (Bottom) TSX amplitude image.

1) the folding of the building toward the sensor due to the layover (the base of the Mirage hotel is almost vertically aligned in the Bing and TSX images); 2) the high range resolution distributes the response of the building over a large number of pixels; 3) the extremely high resolving capabilities of the TSX spotlight imaging that allows distinguishing floors on the southern façade.

A. Single-Scatterer Analysis

In Fig. 4 (top), we show the residual topography (i.e., the topography estimated after the subtraction of the external digital elevation model—in our case, Shuttle Radar Topography Mission) resulting from the MDI, followed by the single scatterer detection algorithm in [9], which tests the presence of a persistent scatterer based on energy content along the direction of the peak of the tomographic reconstruction. The building rising toward the sensor is well recognizable in the detected scatterers. As for previous analyses of TSX data [8], the density of the detected points is also impressive.

Some considerations are now in order: First, on the southern façade, many blue points corresponding to the ground are detected and are mixed to scatterers colored from green to red, corresponding to the vertical structure of the hotel. This fact testifies that the interference in the façade and ground is very likely. Second, in the upper right part of the image, two straight black strips (almost aligned to the azimuth) appear clearly. These areas correspond to two shadowing areas caused by small
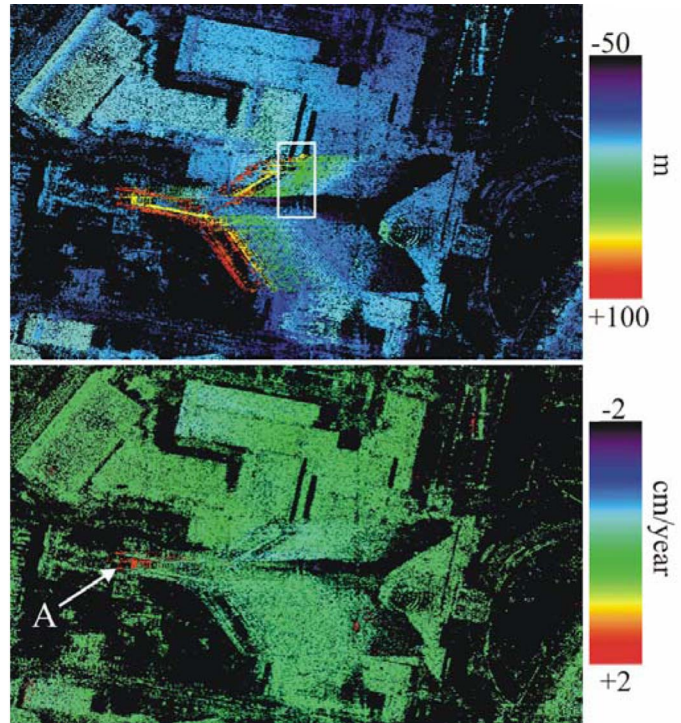
Fig. 4. (Top) Residual topography and (bottom) mean deformation velocity estimated by means of SAR tomography for the single-scatterer analysis.
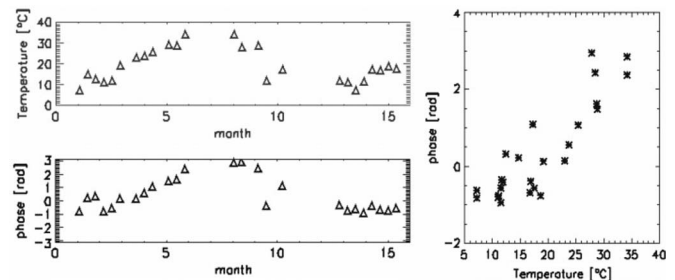
Fig. 5. (Upper left image) Daily averaged temperature of the area. (Lower left image) Residual phases after topography calibration for pixel A. (Right image) Scatter plot.

steps (a few meters high) on the roof of the surrounding structure. One of these shadow strips falls in the radar image areas under the layover of the north façade of the hotel. It is interesting to notice in this area the presence of a high density of scatterers (green pixels) on the part of the façade that falls over the shadowed strips (see the white box in Fig. 4): This high density is the result of the absence of any interference with the ground.

The deformation map presented at the bottom of Fig. 4 also shows an interesting phenomenon: While all the rest is stable, the roof appears moving toward the sensor at about 2 cm/year.

For one of these apparently inflating scatterer (A in Fig. 4), the phase signal obtained after the compensation of the topographic signature is shown in the lower left image in Fig. 5: This plot highlights the presence of a seasonal motion, and hence, the mean velocity is only in part able to explain this movement. The average daily temperatures of the area, provided by the University of Dayton database [14], are shown in the upper left image in Fig. 5. The high degree of correlation with the deformation is evident; see also the scatter plot in the image on the right. As
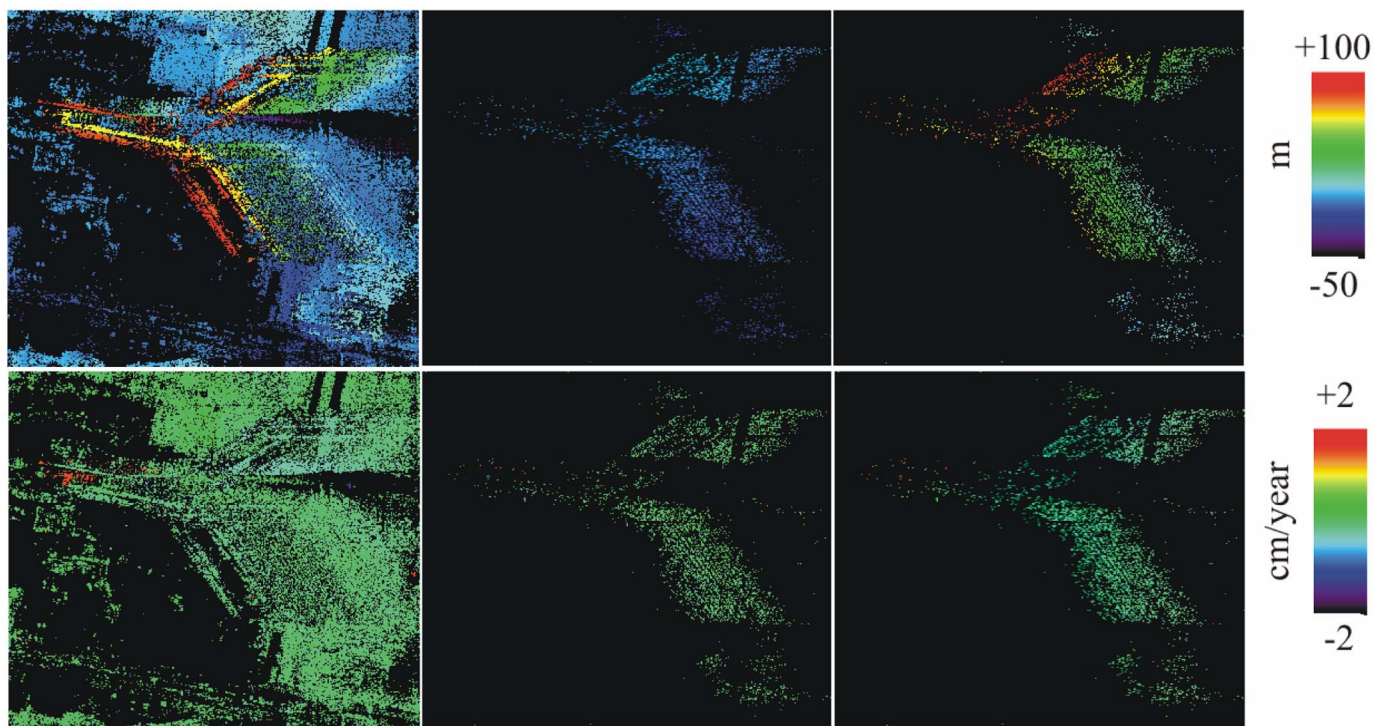
Fig. 6. (Upper row) Residual topography and (bottom row) mean deformation velocity estimated by means of the SAR tomography for (left column) single scatterers, (middle column) lower layer of double scatterers, and (right column) the upper layer of double scatterers.
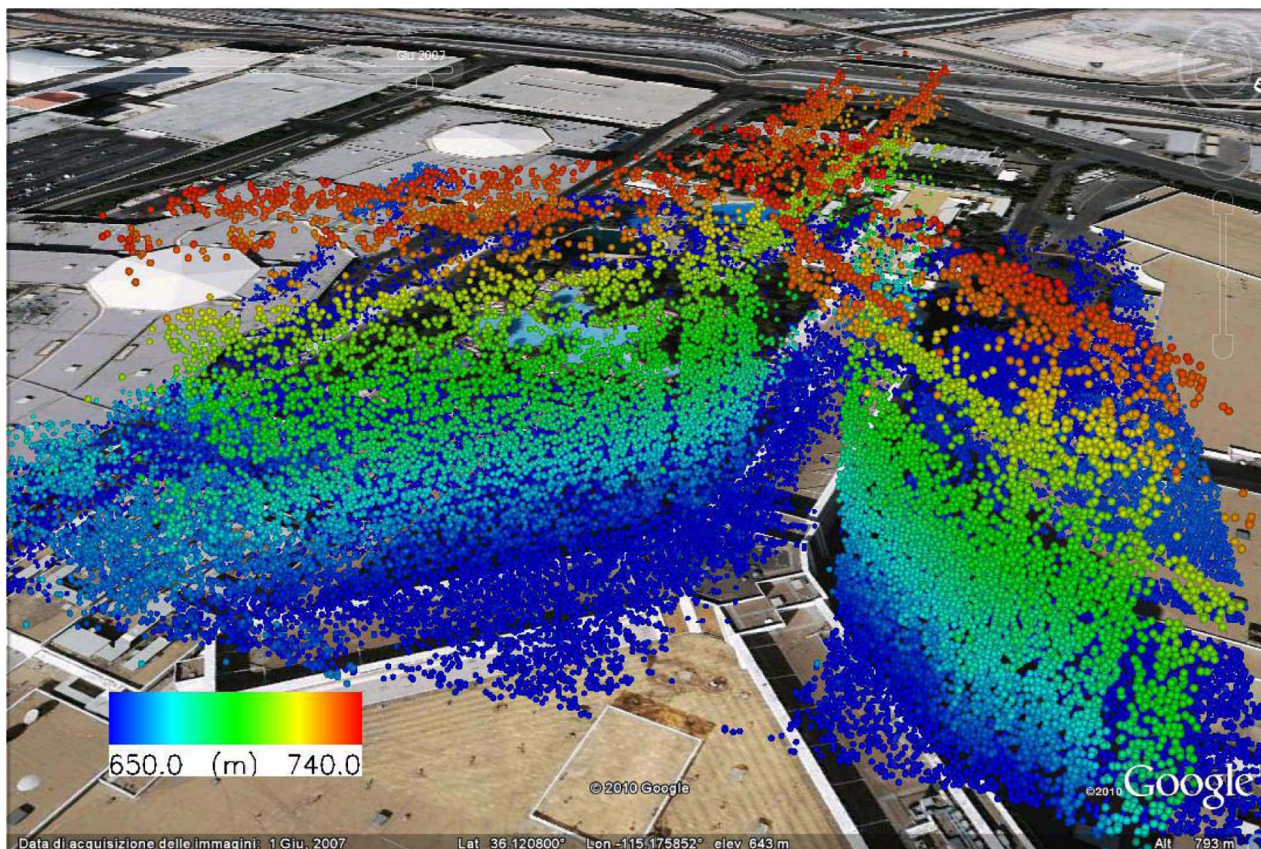
Fig. 7. Three-dimensional visualization of the single and double scatterers reconstructed with SAR tomography on Google Earth. The color is associated to the estimated height.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

REALE *et al.*: IMAGING AND MONITORING OF BUILDINGS WITH VERY HIGH RESOLUTION SAR DATA

665

can be seen, thermal dilation provides contributions leading to a mismatch with the linear displacement model that is commonly adopted in the detection of scatterers [15]. This aspect is the subject of future investigations.

B. Double-Scatterer Analysis

The assumption, made by the classical interferometric techniques, of a single scatterer per pixel neglects the interference of scatterers. We therefore applied the detection scheme described in [10] which is able to test the presence of single and double scatterers. In Fig. 6, the results of this detection algorithm are presented with the colors coded accordingly to (upper row) the estimated topography and (bottom row) mean deformation velocity. These figures show the capability of the tomographic approach to "separate" the interfering layers associated with the ground and the façade of the building. The images on the left column show the detected single scatterers, whereas in the middle and right columns, the images are associated with the ground and top layer extracted from double-scatter results, respectively. The effectiveness of tomography in scatterer pair separation on this layover (distributed over several range pixels) is particularly evident in the topography reconstruction; see the homogeneity of blue color of the ground scatterer layer and the gradation of colors on the layer corresponding to the façade. The high density of detected double scatterers that fills the lack of the single scatterers analysis should be noticed.

A further confirmation of the results is provided by the shadow stripe highlighted by the white box in Fig. 4: As expected, no double scatterers were detected in this area.

For what concerns the deformation maps shown in the right column in Fig. 6, by analyzing both the estimated mean deformation velocity and the previously estimated topography, it is interesting to note the presence of few pixels showing an estimated velocity that is fully congruent with that of the single scatterers affected by strong thermal dilation.

Finally, the 3-D view of the building is shown in Fig. 7 to demonstrate the impressive potential of the new (high-resolution) sensor generation and the potential of SAR tomography for urban area analysis. It shows a 3-D view of the Mirage Hotel in Google Earth obtained with the identified single and double scatterers and without the use of the optical Google 3-D model of the building as background: The different floors are well visible in the left façade. The results show that these SAR sensors orbiting hundreds of kilometers from the Earth can provide accurate 3-D reconstruction and monitoring of single buildings.

IV. Conclusion

High-resolution SAR systems, such as TSX and Cosmo-SkyMed, provide an obvious improvement in the imaging capabilities. However, specific problems associated with the geometry of SAR become more evident: Layover is among them, and it affects particularly the images of urban areas. By processing spotlight TSX data, in this letter, we have shown that SAR tomography can solve this problem and allow accurate 3-D reconstruction and monitoring. Layover associated to tall buildings and distributed over several pixels was successfully resolved.

Whereas layover is solvable by using, as shown, SAR tomography, no solutions are available for shadowing. Hence, small incidence angles are preferred for imaging urban areas to "pierce" areas with high density of buildings and reduce shadowing.

References

[1] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti, "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2375–2383, Nov. 2002.

[2] A. Ferretti, C. Prati, and F. Rocca, "Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2202–2212, Sep. 2000.

[3] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 37–44, Jan. 2005.

[4] G. Fornaro, D. Reale, and F. Serafino, "Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 224–237, Jan. 2009.

[5] G. Fornaro, F. Serafino, and F. Lombardini, "3D multipass SAR focusing: Experiments with long-term spaceborne data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 702–712, Apr. 2005.

[6] G. Fornaro, F. Serafino, and D. Reale, "4D SAR imaging: The case study of Rome," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 236–240, Apr. 2010.

[7] X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4296–4308, Dec. 2010.

[8] S. Gernhardt S, N. Adam, M. Eineder, and R. Bamler, "Potential of very high resolution SAR for persistent scatterer interferometry in urban areas," *Ann. GIS*, vol. 16, no. 2, pp. 103–111, Jun. 2010.

[9] A. De Maio, G. Fornaro, and A. Pauciullo, "Detection of single scatterers in multi-dimensional SAR imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2284–2297, Jul. 2009.

[10] A. De Maio, G. Fornaro, A. Pauciullo, and D. Reale, "Detection of double scatterers in SAR tomography," in *Proc. IEEE IGARSS Conf.*, Cape Town, South Africa, Jul. 13-17, 2009.

[11] X. Zhu and R. Bamler, "Tomographic SAR inversion by L1-norm regularization—The compressive sensing approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3839–3846, Oct. 2010.

[12] A. Budillon, A. Evangelista, and G. Schirinzi, "Three-dimensional SAR focusing from multipass signals using compressive sampling," *IEEE Trans. Geosci. Remote Sens.*, pp. 1–12, to be published.

[13] G. Fornaro, A. Pauciullo, and F. Serafino, "Deformation monitoring over large areas with multipass differential SAR interferometry: A new approach based on the use of spatial differences," *Int. J. Remote Sens.*, vol. 30, no. 6, pp. 1455–1478, Apr. 2009.

[14] Average Daily Temperature Archive. [Online]. Available: http://www.engr.udayton.edu/weather/

[15] X. Zhu and R. Bamler, "Very high resolution SAR tomography via compressive sensing," in *Proc. ESA FRINGE Workshop Adv. Sci. Appl. SAR Interferometry*, Frascati, Italy, 2009.

A.4 Zhu, X., Bamler, R., 2010c. Tomographic SAR Inversion by L1-Norm Regularization— The Compressive Sensing Approach. IEEE Transactions on Geoscience and Remote Sensing 48 (10): 3839-3846

Tomographic SAR Inversion by L_1-Norm Regularization—The Compressive Sensing Approach

Xiao Xiang Zhu, *Student Member, IEEE*, and Richard Bamler, *Fellow, IEEE*

Abstract—Synthetic aperture radar (SAR) tomography (TomoSAR) extends the synthetic aperture principle into the elevation direction for 3-D imaging. The resolution in the elevation direction depends on the size of the elevation aperture, i.e., on the spread of orbit tracks. Since the orbits of modern meter-resolution spaceborne SAR systems, like TerraSAR-X, are tightly controlled, the tomographic elevation resolution is at least an order of magnitude lower than in range and azimuth. Hence, super-resolution reconstruction algorithms are desired. The high anisotropy of the 3-D tomographic resolution element renders the signals *sparse* in the elevation direction; only a few pointlike reflections are expected per azimuth–range cell. This property suggests using compressive sensing (CS) methods for tomographic reconstruction. This paper presents the theory of 4-D (differential, i.e., space–time) CS TomoSAR and compares it with parametric (nonlinear least squares) and nonparametric (singular value decomposition) reconstruction methods. Super-resolution properties and point localization accuracies are demonstrated using simulations and real data. A CS reconstruction of a building complex from TerraSAR-X spotlight data is presented.

Index Terms—Compressive sensing (CS), differential synthetic aperture radar tomography (D-TomoSAR), TerraSAR-X, urban mapping.

I. INTRODUCTION

SYNTHETIC aperture radar (SAR), SAR tomography (TomoSAR) [1] extends the synthetic aperture principle of SAR into the elevation direction for 3-D imaging. It uses acquisitions from slightly different viewing angles (the elevation aperture) to reconstruct for every azimuth–range $(x-r)$ pixel the reflectivity function along the elevation direction s, i.e., the third dimension perpendicular to x and r (Fig. 1). It is essentially a spectral analysis problem. *Differential* SAR tomography (TomoSAR) [2], also referred to as 4-D focusing, obtains a 4-D (space–time) map of scatterers by estimating both the elevation and the deformation velocity of multiple scatterers inside an azimuth–range resolution cell.

The first experiments in TomoSAR were carried out in the laboratory [3] under ideal experimental conditions or by using airborne systems [1]. Spaceborne TomoSAR tests were reported in [4] and [5]. It has been applied to C-band ERS data over extended scenes in [6] and to TerraSAR-X data in [7]. In [8], the single- and double-scatterer cases were separated. The concept of 4-D SAR imaging (differential TomoSAR) was proposed in [2] and first applied to ERS data in [9].

With the German TerraSAR-X and the Italian COSMO-Skymed satellites, SAR data with a very high spatial resolution (VHR) of up to 1 m are available. This resolution is particularly helpful when it comes to interferometric and tomographic imaging of buildings and urban infrastructure. The inherent spatial scales of these objects are in the meter range (e.g., typical height between floors of 3–3.5 m).

We work with TerraSAR-X high-resolution *spotlight* data. These VHR X-band spaceborne repeat-pass tomographic data stacks of urban areas have some particular properties. A very detailed view of individual buildings is possible. The density of bright points, like persistent scatterers, is extremely high (up to $100\,000/\text{km}^2$). However, nonlinear (e.g., thermal-induced) deformations of different building parts must also be expected and will introduce additional phase errors, if not modeled. Due to the tight orbit tube of TerraSAR-X, the elevation aperture is small, i.e., the inherent resolution in elevation is about 50 times worse than that in azimuth or range. This extreme anisotropy calls for super-resolution algorithms in the elevation direction. Finally, VHR data are expensive, and hence, data stacks should be kept small. There are several super-resolving methods, such as CAPON, MUSIC, etc. They are discussed in detail in [10]–[12]. We concentrate on methods that do not require averaging in azimuth and range in order to fully exploit the potential of this class of VHR data.

Compressive sensing (CS) [13]–[16], as a favorable sparse reconstruction technique, is a new and attractive method for TomoSAR. It aims at minimizing the number of measurements to be taken from signals while still retaining the information necessary to approximate them well. It provides a good compromise between classical parametric and nonparametric spectral analysis methods. Compared to parametric spectral analysis, CS is more robust to phase noise, has lower computational effort, and does not require model selection to provide the prior knowledge about the number of scatterers in a resolution cell. Compared to nonparametric spectral estimation, CS has no interference problem and it overcomes the limitation of elevation resolution caused by the extent of elevation aperture, i.e., CS has super-resolution properties.
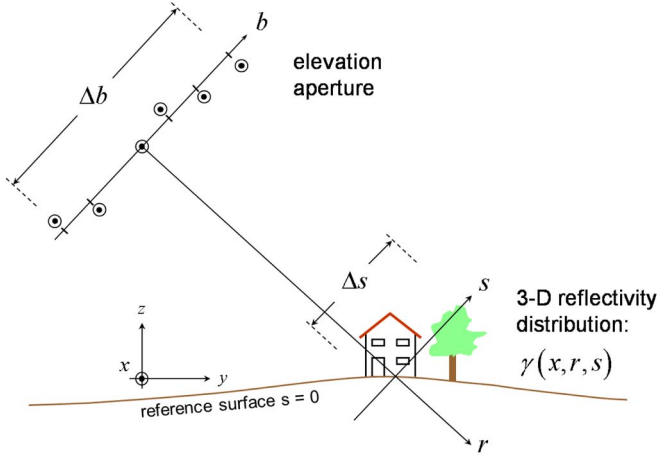
Fig. 1. TomoSAR imaging geometry. The coordinate s is referred to as elevation, and b (parallel to s) is regarded as aperture position.
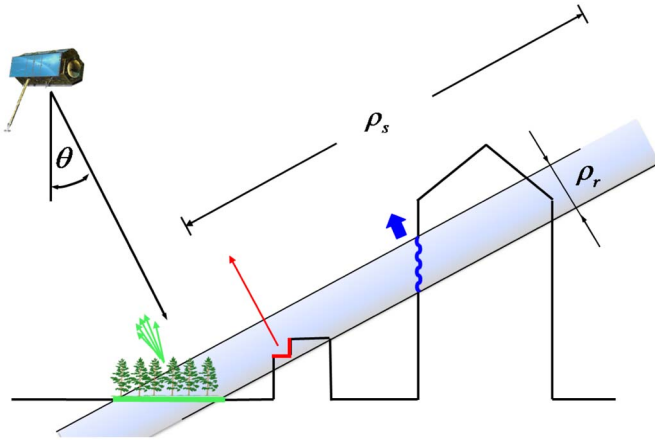
Fig. 2. Possible signal contributions in a single SAR image azimuth–range pixel. ρ_r and ρ_s: range and elevation resolutions, respectively (size of resolution cells not to scale).

Motivated by recent work on inverse SAR [17] and first TomoSAR simulations [18], the CS approach to TomoSAR is outlined in this paper. Its extension to differential (4-D) TomoSAR is introduced. Numerical simulations for realistic acquisition and noise scenarios will be presented to evaluate the potential and limits of the technique. The first CS TomoSAR results with TerraSAR-X spotlight data over urban areas will be presented.

II. SIGNAL AND NOISE MODEL

In VHR X-band data, we expect the following signal contributions (see Fig. 2) [19].

1) Weak diffuse scattering from—mostly horizontal or vertical—rough surfaces (roads and building walls). These objects have an elevation extent of $\rho_r/\tan(\theta - \alpha)$, where ρ_r is the (slant)-range resolution, θ is the local incidence angle, and α is the slope of the surface relative to horizontal. Except from large surfaces accidentally oriented along elevation, these responses are of much smaller extent than the elevation resolution ρ_s, and hence, they can be treated as discrete scatterers in the elevation direction (delta functions).

2) Strong returns from metallic structures or specular and dihedral or trihedral reflections. These are points that would also be used in persistent scatterer interferometry. They are the dominating signal contributions. With VHR SAR data, the density of these points can be very high.

3) Returns from volumetric scatterers, e.g., from vegetation. These result in a continuous-signal background in elevation. These ensembles of scatterers, however, often decorrelate in time, and their response is therefore treated as noise.

The noise sources are the following.

1) Gaussian noise, which is caused by thermal noise and temporal decorrelation, as mentioned earlier.

2) Calibration errors in amplitude. The radiometric stability of TerraSAR-X, i.e., the amplitude variations within one stack, is 0.14 dB and is therefore negligible compared to our typical signal-to-noise ratio (SNR).

3) Phase errors caused by atmospheric delay and unmodeled motion. They require robust and phase-error-tolerant estimation methods.

These considerations suggest that the elevation signal to be reconstructed is *sparse* in the object domain, i.e., it can be described by a few (typically one to three) pointlike contributions of unknown positions and unknown amplitudes and phases. Sparsity is the central concept of and a prerequisite for CS.

III. TomoSAR IMAGING MODEL

For a single SAR acquisition, the focused complex-valued measurement $g_n(x_0, r_0)$ of an azimuth–range pixel (x_0, r_0) for the nth acquisition at aperture position b_n and at time t_n is the integral (tomographic projection) of the reflected signal along the elevation direction [20], as shown in Fig. 1 (the deformation term is ignored here for simplicity) [9]

$$g_n = \int_{\Delta s} \gamma(s) \exp(-j2\pi\xi_n s)\, ds, \qquad n = 1, \ldots, N \quad (1)$$

where $\gamma(s)$ represents the reflectivity function along elevation s. $\xi_n = -2b_n/(\lambda r)$ is the spatial (elevation) frequency. The continuous-space system model of (1) can be approximated by discretizing the continuous-reflectivity function along s (ignoring an inconsequential constant)

$$\mathbf{g} = \mathbf{R}\boldsymbol{\gamma} \quad (2)$$

where $\mathbf{g}$ is the measurement vector with N elements g_n, $\mathbf{R}$ is an $N \times L$ mapping matrix with $R_{nl} = \exp(-j2\pi\xi_n s_l)$, and $\boldsymbol{\gamma}$ is the discrete reflectivity vector with L elements $\gamma_l = \gamma(s_l)$. s_l $(l = 1, \ldots, L)$ denotes the discrete elevation positions. Equation (1) is an irregularly sampled discrete Fourier transform of the elevation profile $\gamma(s)$. The objective of TomoSAR is to retrieve the reflectivity profile for each azimuth–range pixel.

The extension to the 4-D (space-time) case is straightforward [2], [19]. Taking the motion term into account, the system model (1) can be extended to

$$g_n = \int_{\Delta s} \gamma(s) \exp(-j2\pi\left(\xi_n s + \eta_n V(s)\right))\, ds, \qquad n = 1, \ldots, N$$

$$(3)$$

where $V(s)$ is the deformation line-of-sight (LOS) velocity profile along elevation, and $\eta_n = 2t_n/\lambda$ may, in analogy, be called a "velocity frequency." Formally, (3) can be rewritten as

$$g_n = \int_{\Delta v} \int_{\Delta s} \gamma(s)\delta\left(v - V(s)\right) \exp\left(-j2\pi(\xi_n s + \eta_n v)\right) ds\, dv,$$
$$n = 1, \ldots, N \quad (4)$$

where Δv is the range of possible velocities. Equation (4) is a 2-D Fourier transform of $\gamma(s)\delta(v - V(s))$, which is a delta line in the elevation–velocity $(s-v)$ plane along $v = V(s)$. Its projection onto the elevation axis $\gamma(s)\delta(v - V(s))$ is the reflectivity profile $\gamma(s)$. If we accept $\gamma(s)\delta(v - V(s))$ as the object to be reconstructed, the discretized system model of (2) is easily adopted and simply becomes a 2-D Fourier transform [2], [9]. Its inversion provides retrieval of the elevation and deformation information even of multiple scatterers inside an azimuth–range resolution cell and thus obtains a 4-D map of scatterers. It is required for reliable 3- and 4-D city mapping from repeat-pass acquisitions.

IV. TomoSAR via CS

A. CS

CS is a new and popular approach for sparse signal reconstruction. A signal of interest $\mathbf{x}$ with a length of L is said to be K-sparse in an orthogonal basis $\boldsymbol{\Psi}$ if the projection coefficient vector $\mathbf{s} = \boldsymbol{\Psi}\mathbf{x}$ has only K nonzero or significant elements. $\mathbf{x}$ is represented by $\boldsymbol{\Psi}^{\mathbf{H}}\mathbf{s}$. N measurements $\mathbf{y}$ can be obtained by projecting the signal onto N random basis functions $\boldsymbol{\Phi}$ (the sensing matrix)

$$\mathbf{y} = \boldsymbol{\Phi}\mathbf{x}. \quad (5)$$

The measurement vector can be rewritten as

$$\mathbf{y} = \boldsymbol{\Phi}\boldsymbol{\Psi}^{\mathbf{H}}\mathbf{s} = \boldsymbol{\Theta}\mathbf{s}. \quad (6)$$

Within the CS framework, $\mathbf{s}$ can be reconstructed by L_0-norm minimization, i.e., by finding the solution of (6) employing the least number of coefficients

$$\min_{\mathbf{s}} \|\mathbf{s}\|_0 \quad \text{s.t.} \quad \mathbf{y} = \boldsymbol{\Theta}\mathbf{s}. \quad (7)$$

For $N = O(K\log(L/K))$, it can be shown that L_1-norm minimization leads to the same result as L_0-norm minimization [21]

$$\min_{\mathbf{s}} \|\mathbf{s}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \boldsymbol{\Theta}\mathbf{s}. \quad (8)$$

However, the following conditions must hold in order to find the unique sparse solution.

First, the sensing matrix $\boldsymbol{\Phi}$ and the orthogonal basis $\boldsymbol{\Psi}$ must be mutually *incoherent*. Incoherence means that the orthogonal projection by $\boldsymbol{\Psi}$ will spread out information of sparse (highly localized) signals in the entire projection space and thus makes them insensitive to "undersampling." Otherwise, the reconstruction of nonzero coefficients will be biased toward certain positions. For instance, let us randomly choose N columns

from $\boldsymbol{\Psi}$ as our sensing matrix $\boldsymbol{\Phi}$. Then, what we are actually sensing with $\boldsymbol{\Theta}$ is only the N elements of $\mathbf{s}$ located at the positions of the N randomly chosen columns. As a consequence of this counterexample, the sparse reconstruction would only be possible if the K nonzero coefficients were located within the N chosen positions.

Second, the mapping matrix $\boldsymbol{\Theta}$ must follow the *restricted isometry property* (RIP) to guarantee the sufficiently sparse reconstruction in the presence of noise. The RIP requires that

$$(1 - \delta_s)\|\mathbf{v}\|_2^2 \le \|\boldsymbol{\Theta}\mathbf{v}\|_2^2 \le (1 + \delta_s)\|\mathbf{v}\|_2^2 \quad (9)$$

where $\mathbf{v}$ is any vector having K nonzero coefficiencs at the same positions as $\mathbf{s}$ and δ_s is a small number. The smaller the δ_s is, the better the sparse signal can be reconstructed in the presence of noise. Equation (9) essentially says that all submatrices of $\boldsymbol{\Theta}$ composed of K significant columns should be nearly orthogonal and hence preserve the length of the vectors sharing the same K nonzero coefficients as $\mathbf{s}$. There are some well-known pairs of incoherent bases, such as randomly selected Fourier samples as the sensing matrix and the identity matrix as the orthogonal basis, as well as the Gaussian sensing matrix and any other basis.

Therefore, in order to understand whether CS is applicable for a specific problem or not, one should check first the sparsity of signal, the required minimum number of measurements, the incoherence, and the RIP.

B. TomoSAR via CS

As described in Section II, for VHR spaceborne X-band TomoSAR, the elevation signal $\boldsymbol{\gamma}$ to be reconstructed is *sparse* in the object domain with typically one to three pointlike contributions of unknown positions and unknown amplitudes and phases, i.e., $\boldsymbol{\gamma}$ is sparse in the identity orthogonal basis ($\boldsymbol{\Psi} = \mathbf{I}$). According to (1) and (2), the sensing matrix $\boldsymbol{\Phi} = \mathbf{R}$ is a randomly distributed Fourier sampling matrix that has the best incoherence property with our identity orthogonal basis $\mathbf{I}$. Due to the small K, $N = O(K\log(L/K))$ can be very easily fulfilled; hence, the K-sparse signal $\boldsymbol{\gamma}$ can be exactly recovered in the absence of noise by L_1 minimization

$$\min_{\boldsymbol{\gamma}} \|\boldsymbol{\gamma}\|_1 \quad \text{s.t.} \quad \mathbf{g} = \mathbf{R}\boldsymbol{\gamma}. \quad (10)$$

In case there is no prior knowledge about K and in the presence of measurement noise, it can be approximated by

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} \left\{ \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_2^2 + \lambda_K\|\boldsymbol{\gamma}\|_1 \right\} \quad (11)$$

where λ_K is a factor adjusted according to the noise level. The choice of λ_K is described in great detail in [22]. This equation can be solved by basis pursuit methods [23]. Instead of detecting K most significant coefficients, it tries to minimize the residual by employing an L_1-norm regularization. By providing the overcompleteness of $\boldsymbol{\gamma}$ (i.e., several close spectral lines for one scatterer instead of single spectral line), it can provide more robust solutions.

With the given elevation aperture positions, the RIP properties can be checked.
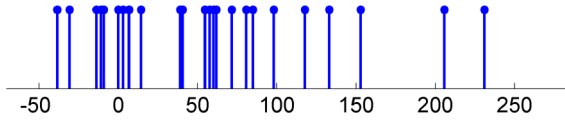
Fig. 3. Elevation aperture positions (in meters).

V. EXPERIMENTS

A. Data Set

For the purpose of this paper, we work with TerraSAR-X spotlight data with a slant-range resolution of 0.6 m and an azimuth resolution of 1 m [24]. Our test site is Las Vegas, NV, U.S. The acquisition repeat cycle is 11 days. The orbit of TerraSAR-X is controlled in a predefined tube of 500-m diameter throughout the entire mission [25]. A data stack of 25 scenes is used for our experiment. The elevation aperture sampling positions are shown in Fig. 3. The elevation aperture size Δb is about 269.5 m.

For nonparametric linear spectral analysis, the expected elevation resolution ρ_s, i.e., the width of the elevation point response function, depends on the elevation aperture length Δb and is approximately (sufficiently dense and regular sampling of the elevation aperture provided)

$$\rho_s = \frac{\lambda r}{2\Delta b} = 40.5 \text{ m} \tag{12}$$

or about 20 m in height (z) at an incidence angle of $31.8°$ [19]. This, however, does not mean that individual scatterers can only be located to within this poor elevation resolution. The Cramér–Rao lower bound (CRLB) on elevation estimates can be shown to be asymptotically (high-SNR approximation) [26]

$$\sigma_{\hat{s}} = \frac{\lambda r}{4\pi\sqrt{NOA} \cdot \sqrt{2SNR} \cdot \sigma_b} \tag{13}$$

where NOA is the number of acquisitions, SNR is the signal-to-noise ratio of the individual scatterer, and σ_b is the standard deviation of the elevation aperture sampling distribution. The stack used in this example has $\sigma_b = 70.9$ m. For many bright points, we can assume a SNR of 10 dB; then, the CRLB on elevation estimation is 1.1 m, i.e., almost 1/40 of the elevation resolution.

With our elevation aperture position distribution, the RIP requirement of (9) is perfectly met in the case of a single scatterer. That is trivial. For two scatterers of equal reflectivity and phase, the quantity δ_s of (9) is a function of the scatterers' distance (Fig. 4). It increases once the scatterers come closer than the resolution ρ_s. In other words, the closer the two scatterers are, the more sensitive the reconstruction becomes to noise, which leads to a lower probability of separating close scatterers. Distinguishing closer scatterers robustly requires higher SNR, which will be discussed in more detail later.

B. Simulated Data

In this section, the CS approach is compared to conventional nonparametric and parametric methods using simulated data. The data are simulated using the elevation aperture distribution of Fig. 3 (elevation resolution $\rho_s = 40.5$ m). The decorrelation effect is introduced by adding Gaussian noise
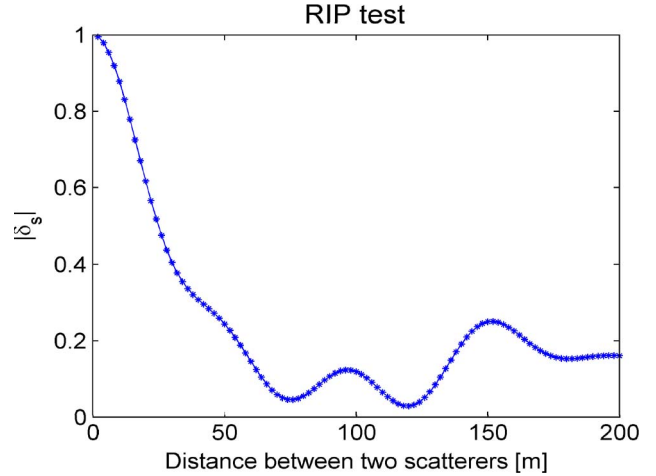
Fig. 4. RIP property of two scatterers with the same reflectivity (i.e., amplitude and phase). δ_s refers to (9).
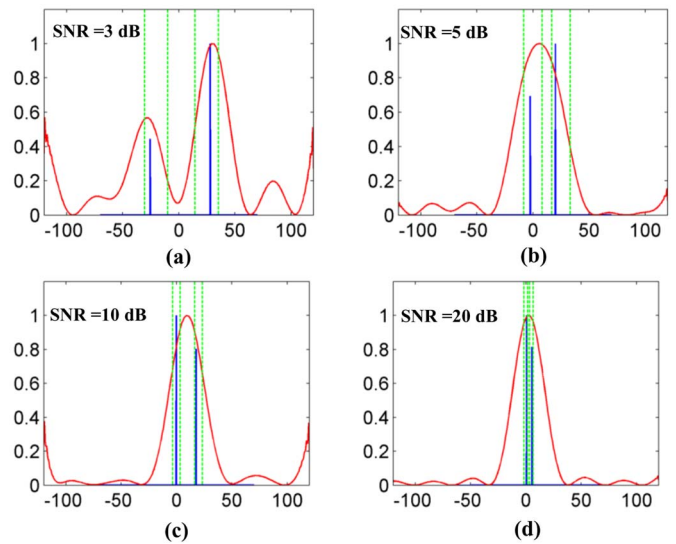
Fig. 5. Comparison of the reconstructed reflectivity profiles along the elevation direction between SVD-Wiener and CS. (Red solid lines) SVD-Wiener. (Blue solid lines) CS. (Green dashed lines) ± 3 times CRLB. (a) Two scatterers with $s_1 = -20$ m and $s_2 = 25$ m (total SNR = 3 dB). (b) Two scatterers with $s_1 = 0$ m and $s_2 = 25$ m (SNR = 5 dB). (c) Two scatterers with $s_1 = 0$ m and $s_2 = 20$ m (SNR = 10 dB). (d) Two scatterers with $s_1 = 0$ m and $s_2 = 5$ m (SNR = 20 dB).

with different SNR. Phase noise due to unmodeled deformation and atmospheric effects are simulated by adding a uniformly distributed phase.

Fig. 5 shows the comparison of the reconstructed reflectivity profiles along the elevation direction between a singular value decomposition (SVD) reconstruction with Wiener-type regularization (SVD-Wiener) [19] and CS. The red lines represent the reconstruction using the nonparametric method SVD-Wiener. The blue lines show the same result using CS. We start with two scatterers with elevations of -25 and 20 m [Fig. 5(a)] with total SNR = 3 dB (i.e., for each of the two scatterers: SNR = 0 dB). CS reconstructs spectral lines instead of sinclike point response functions. Both methods can distinguish the two scatterers well. However, once they move close into one elevation resolution cell with elevations of 0 and 25 m, SVD-Wiener is no longer able to distinguish them. CS detects very
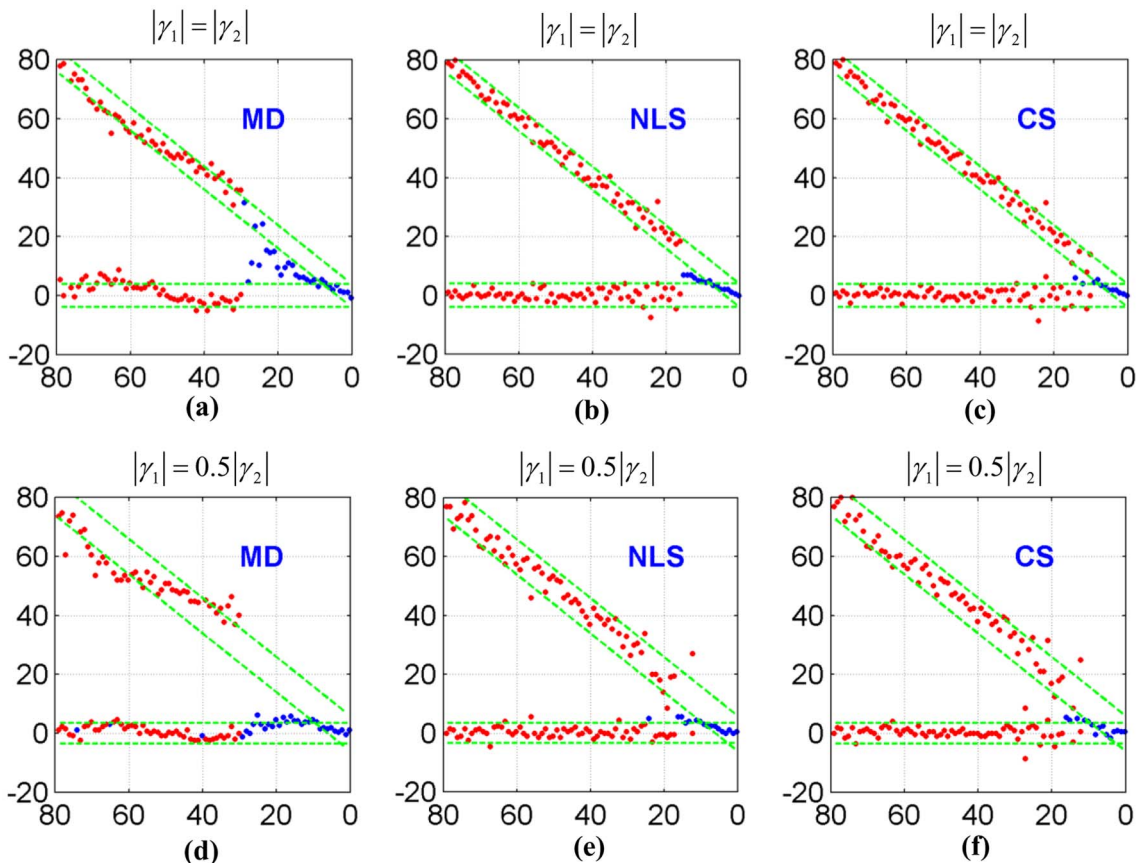
Fig. 6. Estimated elevations (in meters) of two scatterers of equal phase of increasing elevation distance. Shown are the results of (a) and (d) MD, (b) and (e) NLS estimation, and (c) and (f) CS, with total SNR = 10 dB under different reflectivity ratio. (Upper plots) The two scatterers have the same reflectivity. (Lower plots) The reflectivity of the scatterer on the building facade is 0.5 times of the one on the ground. The true positions are a horizontal line referring to the ground and a diagonal line referring to the scatterer at variable elevation. The green dashed lines show ± 3 times the CRLB of elevation estimates for single scatterers (blue: detected single scatterer; red: detected two scatterers).
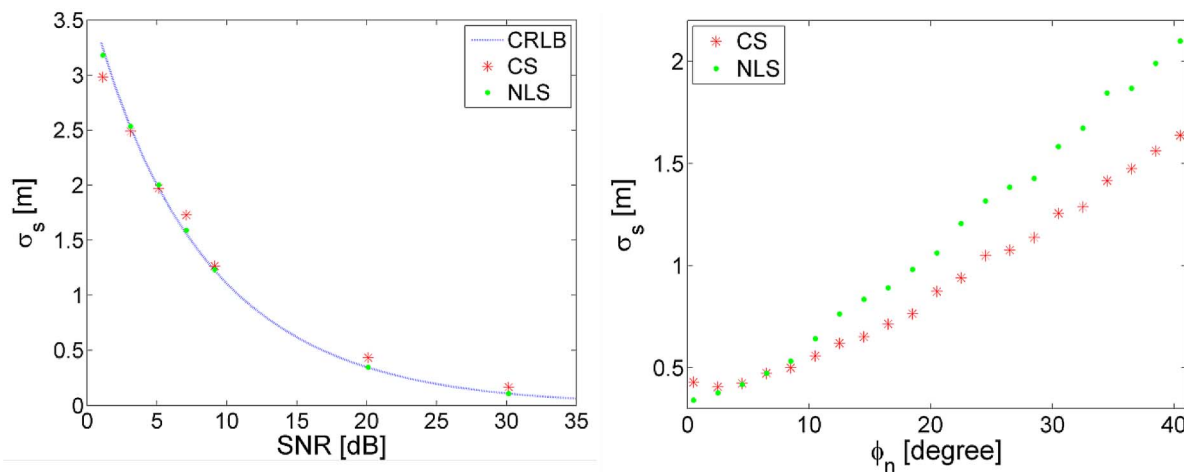
Fig. 7. Single scatterer elevation estimation accuracy of NLS and CS compared to the Cramér–Rao Lower Bound as a function of SNR (left) and as a function of different phase noise levels $[-\phi_n, \phi_n]$ (SNR = 20 dB) (right).

clearly two spectral lines with an accuracy within ± 3 times the CRLB under total SNR = 5 dB [Fig. 5(b)]. With higher SNR, CS is even able to separate closer scatterers. For instance, with SNR = 10 dB, two scatterers with elevations of 0 and 20 m can be well separated [Fig. 5(c)], and with SNR = 20 dB, two scatterers even with elevations of 0 and 5 m can be well separated [Fig. 5(d)].

As multiple scatterers inside one resolution cell most likely occur in high-rise urban areas, the situation in which there are two scatterers inside one resolution cell (one from the building facade and another from the ground) is simulated as an another example to evaluate the performance of the spectral estimation methods. The building is assumed to have an elevation of 80 m, where ground is at zero elevation.

Fig. 8. (Left) Las Vegas Convention Center (Google Earth). (Right) TerraSAR-X intensity map.
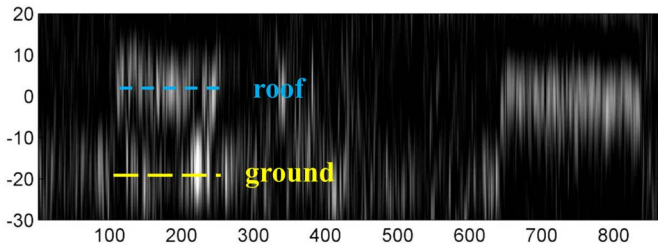
Fig. 9. Estimated reflectivity with SVD-Wiener shown in the azimuth–elevation plane [horizontal: azimuth; vertical: elevation, converted to height (in meters)].
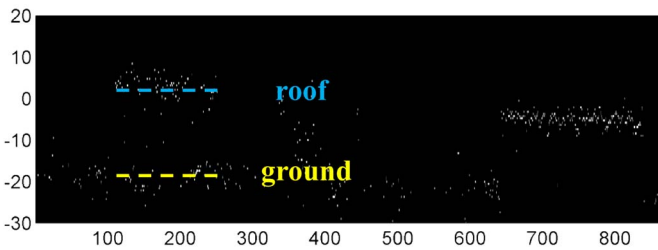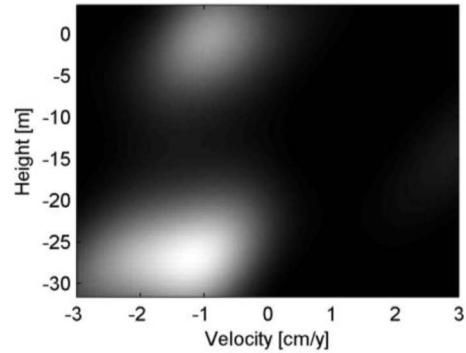
Fig. 10. Same slice as Fig. 9 but estimated by CS.
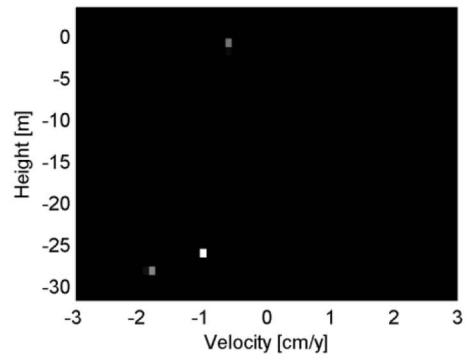
Fig. 11. Four-dimensional reconstruction example. CS versus SVD-Wiener. (Top) TS-X intensity map and analysis point P. (Middle) SVD-Wiener reconstruction results. (Bottom) CS result.

Fig. 6 shows the estimated elevation values of the two scatterers with maximum detection (MD) [22] (left), nonlinear least squares (NLS) [25] (middle), and CS with different reflectivity ratio (i.e., for the upper plots, the two scatterers have the same reflectivity, while for the lower plots, the (amplitude) reflectivity of the scatterer on the building facade is 0.5 times of the one on the ground). MD simply uses the maxima of the nonparametric SVD-Wiener reconstruction as estimates. The x-axis refers to the true elevation of scatterers on the building facade. The y-axis shows their estimated elevations. The ideal image would be two straight lines (one horizontal and another one diagonal). The green dashed lines in the plots show again ±3 times the CRLB on elevation estimates for single scatterers.

From Fig. 6, MD shows the resolution limit of classical nonparametric methods. Once the two scatterers are closer than the elevation resolution, only a single maximum in-between the two true positions is detected. Even if the scatterers are farther apart, their sidelobes mutually distort the location of their maxima [e.g., Fig. 6(a)]. The elevation estimates are biased and follow the sidelobe structure of the elevation point response function. NLS is identical to a maximum-likelihood estimator (MLE) and is the theoretically best solution under Gaussian noise and for single scatterers. NLS requires high

computational effort due to the multidimensional search in elevation. As a parametric method (model-based), we need the *a priori* information about the number of scatterers, i.e., we require model selection and multiple runs of the algorithm for different model orders. This will further increase the computational effort for different hypothesis tests. Compared to MD, CS dramatically improves the elevation resolution under low noise level. For instance, CS is able to reach an elevation separability of about 10 m with SNR = 10 dB. In addition, CS shows no sidelobe interference problem, even in the case where one of the two scatterers has a large reflectivity while the other one is much weaker [e.g., Fig. 6(f)]. Compared to NLS, CS is relatively fast and does not require the number of scatterers as *a priori* knowledge. From visual comparison, CS shows similar elevation estimation accuracy as NLS. The estimates of NLS and CS agree well with the CRLB.

In Fig. 7, the left plot shows the elevation estimation accuracy of a single scatterer in the phase-noise-free case using NLS and CS compared to the CRLB under different SNR levels.

NLS as the MLE with Gaussian noise shows an accuracy that is consistent to the CRLB. Moreover, the elevation estimation accuracy of CS for single scatterers is almost identical to that of NLS. The right plot shows the elevation estimation accuracy of a single scatterer using NLS and CS by adding a phase noise uniformly distributed in $[-\phi_n, \phi_n)$ under SNR = 20 dB. The estimation accuracy obviously highly depends on the phase noise, and CS is more robust against this type of non-Gaussian noise.

Taking all those aspects into account, CS provides the best of both worlds of nonparametric and parametric spectral estimation methods and is hence proven very attractive for TomoSAR.

C. Real Data

a) CS TomoSAR: The Las Vegas Convention Center is a very interesting test building for 3-D focusing for two reasons. First, it is very big and has a regular shape. Therefore, we are able to check the plausibility of the results. Second, it has a height of about 20 m, the critical distinguishable distance between two scatterers (one from the ground and the other from the building) by using SVD-Wiener for our elevation aperture size. The presence of two scatterers within azimuth–range pixels is expected in layover areas and has been validated by using SVD-Wiener in [22]. Thus, we are able to compare the performance of CS at the layover areas to that of the SVD-Wiener method. The left image in Fig. 8 shows the convention center visualized in Google Earth. The right image is the TerraSAR-X intensity map of the area. We choose a reference pixel according to Adam *et al.* [28], which has most likely only a single scatterer inside. The bright blue line shows the position of the analysis slice, and the area marked by the red block is a layover area. From the Google Earth image, we can see that there is a small triangular-shaped plaza on the ground made of the same material as the building. Thereby, multiple scatterers are expected.

Fig. 9 shows the estimated reflectivity with SVD-Wiener in the azimuth–elevation plane [horizontal: azimuth; vertical: elevation (converted to height)]. Multiple scatterers with marginally distinguishable distance appear (one from the building marked with blue line, and the other from the small structures on the ground marked with yellow line). Even though it demonstrates the stability of SVD-Wiener, the resolution limitation blurs the reflectivity profile for each pixel. In contrast, Fig. 10 shows the same plot as Fig. 9 estimated by the CS approach. Compared to Fig. 9, not only the layover area can be separated but also the elevation positions can be easily located in the reflectivity slice.

b) Differential CS TomoSAR: We have implemented the CS approach to *differential* TomoSAR as well. The top image in Fig. 11 shows again the TerraSAR-X intensity map of the convention center. The pixel P marked by the red dot that locates at the layover area and a reference point (green) on the roof of the building have been taken as an example. Again, two scatterers with slightly different linear deformation in LOS (one from the ground and one from the roof) have been detected in the elevation–velocity plane by SVD-Wiener (middle image). However, the scatterer on the ground appears much brighter and

wider. It is very likely to have two scatterers together that are not separable. The bottom image of Fig. 11 shows the result using the CS approach. Two very close scatterers with slightly different heights and velocities (about 2 m in height and 1 cm/y in velocity) have been detected. This signal may be caused by the semicircular structure of the convention center overlaid with the plaza, which can be seen from the Google Earth image.

VI. CONCLUSION

TomoSAR with very high-resolution spaceborne systems like TerraSAR-X and COSMO-Skymed requires robust inversion algorithms with super-resolution capabilities. Since the elevation profiles to be reconstructed can often be assumed sparse, i.e., they consist only of a small number of pointlike scatterers, the CS framework is applicable. These algorithms use L_1-norm minimization and regularization. Compared to nonparametric and fully parametric L_2-norm methods, they have the following several advantages.

1) CS is more robust than NLS parametric methods with respect to unmodeled phase errors. It does not suffer from self-cancellation artifacts, like CAPON [10].
2) In the single-scatterer case and under Gaussian noise, CS approaches the accuracy of NLS, i.e., the CRLB.
3) CS is computationally more efficient than NLS.
4) For multiple scatterers, CS exhibits a much better resolution than linear nonparametric methods.
5) CS does not need model selection, i.e., it "automatically" chooses the number of scatterers that can be resolved.
6) CS can achieve super resolution in elevation while maintaining the full azimuth–range resolution.

Further work will focus on evaluating the super-resolution power and robustness of this technique, i.e., trying to find its limit.

REFERENCES

[1] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.
[2] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," in *Proc. IGARSS*, Toulouse, France, 2003, pp. 1206–1208.
[3] P. Pasquali, C. Prati, F. Rocca, and M. Seymour, "A 3-D SAR experiment with EMSL data," in *Proc. IGARSS*, 1995, pp. 784–786.
[4] J. Homer, D. I. Longstaff, Z. She, and D. Gray, "High resolution 3-D imaging via multi-pass SAR," *Proc. Inst. Elect. Eng.*, vol. 149, no. 1, pt. F, pp. 45–50, Feb. 2002.
[5] Z. She, D. Gray, R. E. Bogner, J. Homer, and I. D. Longstaff, "Three-dimensional spaceborne synthetic aperture radar (SAR) imaging with multipass processing," *Int. J. Remote Sens.*, vol. 23, no. 20, pp. 4357–4382, Oct. 2002.
[6] G. Fornaro, F. Serafino, and F. Lombardini, "Three-dimensional multipass SAR focusing: Experiments with long-term spaceborne data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 702–714, Apr. 2005.
[7] X. Zhu, N. Adam, and R. Bamler, "First demonstration of space-borne high resolution SAR tomography in urban environment using TerraSAR-X data," in *Proc. CEOS SAR Workshop Calibration Validation*, 2008.
[8] G. Fornaro and F. Serafino, "Imaging of single and double scatterers in urban areas via SAR tomography," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 12, pp. 3497–3505, Dec. 2006.
[9] G. Fornaro, D. Reale, and F. Serafino, "Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 224–237, Jan. 2009.

[10] F. Lombardini and A. Reigber, "Adaptive spectral estimation for multi-baseline SAR tomography with airborne L-band data," in *Proc. IEEE IGARSS*, 2003, vol. 3, pp. 2014–2016.

[11] S. Guillaso and A. Reigber, "Polarimetric SAR tomography (POLTOMSAR)," in *Proc. POLINSAR*, Frascati, Italy, 2005.

[12] F. Gini and F. Lombardini, "Multibaseline cross-track SAR interferometry: A signal processing perspective," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 20, no. 8, pp. 71–93, Aug. 2005.

[13] E. Candès, "Compressive sampling," in *Proc. Int. Congr. Math.*, Madrid, Spain, 2006, vol. 3, pp. 1433–1452.

[14] R. Baraniuk, "Compressive sensing," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.

[15] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[16] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[17] L. Zhang, M. Xing, C. Qiu, J. Li, and Z. Bao, "Achieving higher resolution ISAR imaging with limited pulses via compressed sampling," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 567–571, Jul. 2009.

[18] A. Budillon, A. Evangelista, and G. Schirinzi, "SAR tomography from sparse samples," in *Proc. IEEE IGARSS*, Cape Town, Africa, 2009, pp. IV-865–IV-868.

[19] X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *IEEE Trans. Geosci. Remote Sens.*, to be published.

[20] G. Fornaro, F. Serafino, and F. Soldovieri, "Three-dimensional focusing with multipass SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 507–517, Mar. 2003.

[21] L. Gan, C. Ling, T. T. Do, and T. D. Tran, *Analysis of the Statistical Restricted Isometry Property for Deterministic Sensing Matrices Using Stein's Method*, 2009, (Preprint).

[22] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.

[23] D. K. Hammond, L. Jacques, M. J. Fadili, G. Puy, and P. Vandergheynst, The Basis Pursuit DeQuantizer (BPDQ) Toolbox, Signal Processing Laboratory (LTS2), EPFL, Lausanne, Switzerland, Jul. 2009. [Online]. Available: http://wiki.epfl.ch/bpdq

[24] M. Eineder, N. Adam, R. Bamler, N. Yague-Martinez, and H. Breit, "Spaceborne spotlight SAR interferometry with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1524–1535, May 2009.

[25] S. D'Amico, C. Arbinger, M. Kirschner, and S. Campagnola, "Generation of an optimum target trajectory for the TerraSAR-X repeat observation satellite," in *Proc. 18th Int. Symp. Space Flight Dyn.*, Munich, Germany, 2004.

[26] R. Bamler, M. Eineder, N. Adam, X. Zhu, and S. Gernhardt, "Interferometric potential of high resolution spaceborne SAR," *Photogramm. Fernerkundung Geoinf.*, vol. 2009, no. 5, pp. 407–419, Nov. 2009.

[27] X. Zhu, N. Adam, R. Brcic, and R. Bamler, "Space-borne high resolution SAR tomography: Experiments in urban environment using TerraSAR-X data," in *Proc. JURSE*, 2009.

[28] N. Adam, R. Bamler, M. Eineder, and B. Kampes, "Parametric estimation and model selection based on amplitude-only data in PS-interferometry," in *Proc. FRINGE-Workshop*, Frascati, Italy, 2005.

Xiao Xiang Zhu (S'10) was born in Changsha, China, on December 12, 1984. She received the B.S. degree in space engineering from the National University of Defense Technology, Changsha, in 2006 and the M.Sc. degree in earth oriented space science and technology (ESPACE) from the Technische Universität München (TUM), München, Germany, in 2008. She is curently working toward the Ph.D. degree in differential synthetic aperture radar (SAR) tomography using TerraSAR-X data at the Remote Sensing Technology, TUM.

Since May 2008, she has been a Full-Time Scientific Collaborator with Remote Sensing Technology, TUM. Her work is part of the project team "Dynamic Earth," which was established in the International Graduate School of Science and Engineering, TUM, as a result of the German Excellence Initiative in 2007. In October/November 2009, she was a Guest Scientist with the Institute for Electromagnetic Sensing of the Environment (IREA), Italian National Research Council (CNR), Naples, Italy. Her main research interests are in the signal processing field with applications to spaceborne SAR data, in particular SAR tomography, and differential SAR tomography.

Richard Bamler (M'95–SM'00–F'05) received the Diploma degree in electrical engineering, the Doctor of Engineering degree, and the "Habilitation" degree in signal and systems theory from the Technische Universität München (TUM), München, Germany, in 1980, 1986, and 1988, respectively.

During 1981 and 1989, he was with TUM, working on optical signal processing, holography, wave propagation, and tomography. In 1989, he joined the German Aerospace Center (DLR), Oberpfaffenhofen, Germany, where he is currently the Director of the Remote Sensing Technology Institute (IMF). Since then, he and his team have been working on synthetic aperture radar (SAR) signal processing algorithms (ERS, SIR-C/X-SAR, Radarsat, SRTM, ASAR, TerraSAR-X, and TanDEM-X), SAR calibration and product validation, SAR interferometry, phase unwrapping, estimation theory and model-based inversion methods for atmospheric sounding (GOME, SCIAMACHY, MIPAS, and GOME-2), and oceanography. In early 1994, he was a Visiting Scientist with the Jet Propulsion Laboratory in preparation of the SIC-C/X-SAR missions, and in 1996, he was a Guest Professor with the University of Innsbruck, Innsbruck, Austria. Since 2003, he has been a Full Professor in remote sensing technology with TUM. His current research interests are in algorithms for optimum information extraction from remote sensing data with emphasis on SAR, SAR interferometry, persistent scatterer interferometry, SAR tomography, and ground moving target indication for security-related applications. He and his team have developed and are currently developing the operational processor systems for the German missions TerraSAR-X, TanDEM-X, and EnMAP. He is the author of more than 160 scientific publications, among them about 40 journal papers, a book on multidimensional linear systems theory, and several patents on SAR signal processing.

A.5 Zhu, X., Bamler, R., 2011b. Super-Resolution Power and Robustness of Compressive Sensing for Spectral Estimation with Application to Spaceborne Tomographic SAR. IEEE Transactions on Geoscience and Remote Sensing, in press

Super-Resolution Power and Robustness of Compressive Sensing for Spectral Estimation with Application to Spaceborne Tomographic SAR

Xiao Xiang Zhu [1], Student Member, IEEE, Richard Bamler [1, 2], Fellow, IEEE

(1) Technische Universität München, Lehrstuhl für Methodik der Fernerkundung, Arcisstraße 21, 80333 Munich, Germany

Email: xiaoxiang.zhu@bv.tum.de, Tel: +49-8153-3531, Fax: +49-8153-1420

(2) German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), Oberpfaffenhofen, 82234 Wessling,

Germany

Abstract

We address the problem of resolving two closely spaced complex-valued points from N irregular Fourier domain samples. Although this is a generic super-resolution (SR) problem, our target application is SAR tomography (TomoSAR) where typically the number of acquisitions is $N = 10\dots100$ and SNR $= 0 \dots 10$dB.

As the TomoSAR algorithm we introduce "Scale-down by L1 norm Minimization, Model selection, and Estimation Reconstruction" (SL1MMER), a spectral estimation algorithm based on compressive sensing, model order selection and final maximum likelihood parameter estimation.

We investigate the limits of SL1MMER concerning the following questions:

- How accurately can the positions of two closely spaced scatterers be estimated?

2

- What is the closest distance of two scatterers, such that they can be separated with a detection

rate of 50% assuming a uniformly distributed phase difference?

- How many acquisitions N are required for a robust estimation, i.e. for separating two scatterers

spaced by one Rayleigh resolution unit with a probability of 90%?

For all these questions we provide numeric results, simulations and analytic approximations.

Although we take TomoSAR as the preferred application, the SL1MMER algorithm and our results

on SR are generally applicable to sparse spectral estimation, including SR SAR focusing of point-like

objects. Our results are approximately applicable to nonlinear least-squares estimation, and hence,

although being derived experimentally, can be considered as a fundamental bound for SR of spectral

estimators. We show that SR factors are in the range of 1.5 to 25 for the aforementioned parameter

ranges of N and SNR.

Keywords: super-resolution, spectral estimation, compressive sensing, synthetic aperture

radar (SAR), SAR tomography

3

I.   INTRODUCTION

SAR tomography (TomoSAR) uses stacks of repeat-pass SAR acquisitions to reconstruct the reflectivity of the scattering objects along elevation *s* by means of spectral analysis for every azimuth-range (*x-r*) pixel [1], [2], [3]. With the German TerraSAR-X (TS-X) and the Italian COSMO-Skymed satellites, SAR data with a very high spatial resolution (VHR) of up to 1m are available. This resolution is particularly helpful when it comes to interferometric and tomographic imaging of buildings and urban infrastructure. The inherent spatial scales of these objects are in the meter range, e.g. typical height between floors of 3–3.5 m [4].

There are several challenges to fully exploit the potential of this class of VHR data. Firstly, compared to the airborne case, spaceborne data are collected over a longer time period. Temporal decorrelation thus reduces the SNR of coherent objects – for TS-X the signal to noise ratio (SNR) caused by decorrelation is typically 0-10dB [5] – calling for a larger number of acquisitions for accurate reconstruction. On the other hand, VHR data are expensive and, hence, the data stacks should be kept small. Secondly, nonlinear, e.g. thermal-induced, deformations of different building parts must be expected and this will introduce additional phase errors if not modeled. In this case Differential-TomoSAR (D-TomoSAR) provides a solution [1] [6] [3]. Finally, in TomoSAR imaging, it would be favorable to have an isotropic tomographic spatial resolution element, i.e. an elevation resolution in the order of that in azimuth and range. This would require a huge elevation aperture of several kilometers. However, the tight orbit control of modern meter-resolution satellites leads to a tomographic elevation resolution ρ_s of about

10-50 times less than that in azimuth ρ_x or range ρ_r [3]. Those particular challenges call for robust super-resolution (SR) TomoSAR algorithms. In order to maintain the azimuth-range resolution for urban infrastructure imaging, these algorithms should not require averaging in azimuth and range.

The high anisotropy of the 3D tomographic resolution element together with the fact that thick volumetric scatterers tend to temporarily decorrelate in X-band renders the signals *sparse* in the elevation direction, i.e. only a few point-like reflections along elevation are expected per azimuth-range cell. For instance, TS-X in its high resolution spotlight mode has a typical resolution of 1.1 m in azimuth, 0.6m in range and 30m in elevation. Possible signal contributions in a single azimuth-range pixel could be point scatterers and rough surfaces. Their projections onto the elevation direction have a size of about 1/30 of the elevation resolution and, hence, can be treated as a few individual delta functions (typically 1-4) [10]. These kind of signals are refered to as *sparse*. Considering the *sparsity* of the signals in the elevation direction, compressive sensing (CS) [7] [8] [9], as a favorable sparse reconstruction technique, has been recently introduced to radar [10] [11] [12] [13] and TomoSAR reconstruction: The first CS TomoSAR simulations were presented in [14] and the SR capability of CS for TomoSAR reconstruction and its robustness on elevation estimation against phase noise have been proven in [15] using TS-X high resolution spotlight data. However, the better performance of CS for elevation estimates over nonlinear least squares (NLS) estimation is paid by the generation of occasional outliers which leads to a reduced accuracy in complex-valued reflectivity estimation. Also, CS is known to slightly underestimate amplitudes of scatterers. Therefore, we combine the CS reconstruction with model selection and a final linear estimation of the complex amplitudes. Note that

the CS TomoSAR algorithm has been recently introduced by the authors in [15] but is presented here in more detail, because the different algorithm steps are essential for the SR power. We name our CS-based TomoSAR spectral estimation algorithm "Scale-down by L1 norm Minimization, Model selection, and Estimation Reconstruction" (*SL1MMER*, pronounced "slimmer"). It combines the advantages of CS with the amplitude and phase accuracy of linear estimators.

In this paper, we investigate the limits of the technique by addressing the following three questions:

- How accurately can the positions of two closely spaced scatterers be estimated? We give both numeric results and an analytic approximation.

- What is the closest distance of two scatterers, such that they can be separated with a detection rate of 50% at given N·SNR and amplitude ratio with a uniformly distributed phase difference? We find the SR factors by extensive simulations and provide an easy-to-use polynomial approximation.

- How many acquisitions *N* are required for a *robust* estimation, i.e. for separating two scatterers spaced by one Rayleigh resolution unit with a probability of 90%?

Although we take TomoSAR as the preferred application, the SL1MMER algorithm and our results on SR are generally applicable to sparse spectral estimation, including SAR focusing of point-like objects. We will also show that our proposed SL1MMER algorithm provides an aesthetic approximation of the NP-hard L_0 norm regularization which is equivalent to an NLS plus model selection (see Section III), and hence, our results are approximately applicable to NLS estimation. They are asymptotically optimal with an estimation accuracy approaching the Cramér-Rao Lower Bound (CRLB). As a consequence,

although being derived experimentally our results can be considered as fundamental bounds for the SR

capability of spectral estimators.

The paper is organized as follows. Chapter II gives a brief introduction to TomoSAR imaging model.

Chapter III provides an overview on inversion methods and regularizers. The basic idea of the

SL1MMER algorithm is presented in Chapter IV. Chapter V derives the CRLB of elevation estimation

for two interfering scatterers with a uniformly distributed phase difference and further analyzes the

elevation estimation accuracy of the algorithm. Chapter VI and Chapter VII investigate the SR power

and robustness of the proposed algorithm for TomoSAR and some conclusions are drawn in the last

Chapter. For all our simulations we assume a range of $N = 10\ldots100$ and SNR $= 0 \ldots 10$dB, i.e. $N \cdot$SNR $=$

$10 \ldots 30$dB.

II.   TOMOGRAPHIC SAR IMAGING MODEL

TomoSAR inversion starts from a stack of N complex SAR data sets of the same area taken at different

times and slightly different orbit positions (the elevation aperture). Let the elevation aperture positions,

i.e. the perpendicular baselines with respect to a master track, be b_n. After some trivial phase

corrections, the focused complex value g_n of an azimuth-range pixel (x_0, r_0) of the nth acquisition is

[2], [16]:

$$g_n = \int\limits_{\Delta s} \gamma(s) \exp(-j2\pi\xi_n s) ds. \tag{1}$$

where $\gamma(s)$ represents the reflectivity function along elevation s, and Δs describes the range of

possible elevations. $\xi_n = -2b_n/(\lambda r_0)$ is the spatial (elevation) frequency depending on the (more or less

random) elevation aperture position b_n where λ stands for the wavelength. Equation (1) is a

randomly sampled Fourier transform of $\gamma(s)$. From this we expect an inherent Rayleigh resolution in

elevation of $\rho_s = \lambda r / (2\Delta b)$ where Δb is the elevation aperture size. The system model (1) can be

approximated by discretizing the object along s, in presence of the noise ε, it can be written as

(ignoring an inconsequential constant):

$$\mathbf{g} = \mathbf{R}\,\gamma + \varepsilon \tag{2}$$

where $\mathbf{g}$ is the measurement vector of length N, $\mathbf{R}$ with $R_{n \times l} = \exp(-j2\pi\xi_n s_l)$ is an $N \times L$

mapping (Fourier transform) matrix and γ is $\gamma(s)$ uniformly sampled in elevation at s_l $(l = 1, ..., L)$.

Taking motion terms into account, the system model (1) can be extended to differential TomoSAR

(D-TomoSAR) [1]. Its 2D inversion provides retrieval of the elevation and deformation information

even of multiple scatterers inside an azimuth-range resolution cell and thus obtains a 4D map of

scatterers. Extensions to multi-modal motion models are possible, but require an even higher

dimensional spectral estimation inversion [6]. For the sake of simplicity we restrict ourselves to the 1D

spectral estimation case here.

III.   INVERSION METHODS AND REGULARIZERS

Since $L \gg N$, the system model (2) is severely under-determined. Hence, there are infinitely many

solutions. Depending on the type of prior (i.e. regularizer) different solutions can be obtained:

A. L_2 Norm Regularization

The conventional TomoSAR algorithms are based on the L_2 norm regularization, also referred to as

Tikhonov regularization:

$$\hat{\boldsymbol{\gamma}} = \arg \min_{\boldsymbol{\gamma}} \left\{ \left\| \mathbf{g} - \mathbf{R}\boldsymbol{\gamma} \right\|_2^2 + \left\| \boldsymbol{\Gamma} \, \boldsymbol{\gamma} \right\|_2^2 \right\} \tag{3}$$

for some suitably chosen Tikhonov matrix $\boldsymbol{\Gamma}$. For Gaussian stationary white measurement noise, i.e.

$\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = \sigma_\varepsilon^2 \mathbf{I}$, and Gaussian prior with covariance matrix $\boldsymbol{\Gamma}^{-\frac{1}{2}}$ this is the standard MAP estimator. The

special case of white prior, i.e. $\boldsymbol{\Gamma} = \mathbf{I}$, renders (3) to:

$$\hat{\boldsymbol{\gamma}} = \arg \min_{\boldsymbol{\gamma}} \left\{ \left\| \mathbf{g} - \mathbf{R}\boldsymbol{\gamma} \right\|_2^2 + \sigma_\varepsilon^2 \left\| \boldsymbol{\gamma} \right\|_2^2 \right\} \tag{4}$$

This estimator has been named "SVD Wiener" in [3]. Another choice of $\boldsymbol{\Gamma}$ leads to the well-known

truncated SVD estimator [16].

As a non-parametric estimator, eq. (4) is computational efficient and does not require any assumption

about the number of scatterers. However, it suffers from the sidelobe problem and does not provide

super resolving capability.

B. L_0 Norm Regularization

As described in the introduction and outlined in detail in [15], $\boldsymbol{\gamma}$ is *sparse* in the object (elevation)

domain for VHR space-borne X-band TomoSAR with typically 1-4 point-like contributions of unknown

positions, amplitudes and phases, i.e. $\boldsymbol{\gamma}$ is assumed to contain only $K = 1 \dots 4$ non-zero elements.

This *sparsity* property of $\boldsymbol{\gamma}$ suggests using the CS approach. It says, if the mapping matrix fulfills certain

requirements, i.e. the *restricted isometry property* (*RIP*) and the *incoherence* property [9], with number

of measurements depending on the sparsity K instead of the length of the signal L, $\boldsymbol{\gamma}$ can be very well

approximated by the solution with the least number of scatterers (non-zero elements of $\boldsymbol{\gamma}$), i.e. the

minimal L_0 norm, which satisfies the measurements with noise:

9

$$\min_{\boldsymbol{\gamma}} \|\boldsymbol{\gamma}\|_0 \quad \text{s.t.} \quad \mathbf{g} = \mathbf{R}\boldsymbol{\gamma} \tag{5}$$

In the presence of measurement noise, the estimator can be written as:

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} \left\{ \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_2^2 + \lambda_{MS} \|\boldsymbol{\gamma}\|_0 \right\} \tag{6}$$

This is a NLS parametric estimator, i.e. a maximum likelihood estimator (MLE) under Gaussian noise, plus a model complexity penalty term. By choosing $\lambda_{MS} = 3\sigma_\varepsilon^2 \ln N$, the penalty term is equivalent to the Bayesian Information Criterion (BIC) model selection scheme which will be addressed in the next Chapter.

Due to the sparse nature of the signal in elevation direction, minimization of L_0 norm should be the correct way to find the exact solution, even with two scatterers closer than a resolution cell. However, it is an N-P hard problem. Therefore, except we limit ourselves to the cases with only 1 or 2 scatterers and in a *single* dimension – elevation – (i.e. no extension to D-TomoSAR), we need a more efficient algorithm.

C. L_1 Norm Regularization (Compressive Sensing)

Compressive sensing theory tells us that, if the mapping matrix $\mathbf{R}$ fulfills the RIP and incoherence properties and $N = \mathrm{O}\left(K \log\left(L/K\right)\right)$, the *convex* L_1 norm minimization gives the same solution as the N-P hard L_0 norm minimization. In case there is no prior knowledge about K and in the presence of measurement noise, it can be approximated by the following L_1- L_2 norm minimization [15]:

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} \left\{ \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_2^2 + \lambda_K \|\boldsymbol{\gamma}\|_1 \right\} \tag{7}$$

where λ_K is the Lagrange multiplier as a function of N and ε [22] [27]. The CS estimator of (7) has approximately the same SR capability as (6), but it is more robust to phase noise, needs less

computational effort, and does not require model selection to provide the prior knowledge about the

number of scatterers K.

D. Drawbacks of the CS Estimator

The simple CS estimator (7) has two drawbacks that will be discussed in the following. Both are

corrected in the proposed SL1MMER algorithm.

-　***RIP and Incoherence***

RIP guarantees the sufficiently sparse reconstruction in the presence of noise. It essentially says that all

sub-matrices of $\mathbf{R}$ composed of K columns should be nearly orthogonal and, hence, preserve the

energy of all vectors having only non-zero elements at the same K positions as $\boldsymbol{\gamma}$. *Incoherence* means

that the mapping operator $\mathbf{R}$ will spread out information of sparse (highly localized) signals in the

entire measurement space (like a Fourier transform does) and thus makes them insensitive to random

undersampling.

However, for many applications like TomoSAR, RIP and incoherence are violated for several reasons.

First the mapping matrix $\mathbf{R}$ is pre-determined by the measurement system (the elevation aperture

sampling pattern) and may be not optimum. Second, the profile $\gamma(s)$ to be reconstructed is often

sampled much more densely than the elevation resolution unit in order to allow for good resolution and

scatterer positioning accuracy. This small sampling distance renders $\mathbf{R}$ over-complete, reduces RIP

[15] and increases coherence. As a consequence the CS estimator (7) will not always give the solution

with the correct number of scatterers K but rather may produce spurious artifacts. There are two classes

of artifacts shown schematically in Fig. 1. If the sampling raster is very dense, single scatterers may be

broadened to a few adjacent lines (Fig. 1.b), however, well within a Rayleigh resolution unit. With a

certain probability small spurious impulses may show up (Fig. 1.c).
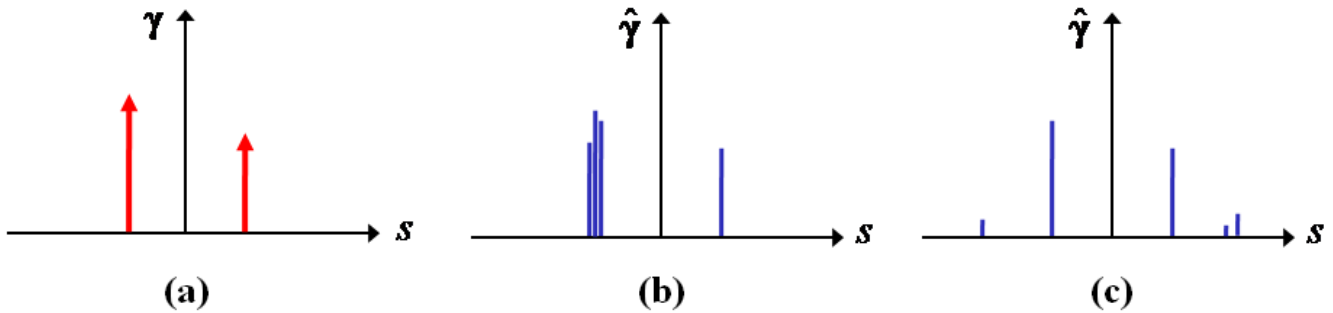
Fig. 1. The violation of RIP and incoherence properties may cause artifacts in the reconstruction. (a):

true reflectivity profile; (b) reconstruction of a scatterer by several adjacent lines instead of a single

peak; (c) reconstruction with noise-like outliers.

- ***Amplitude Bias***

Although the CS estimator (7) is radiometrically much more reliable than MUSIC, CAPON, AR etc., it

will slightly, yet systematically, underestimate the amplitudes of the scatterers, because the L_1 prior

favors smaller amplitudes.

IV.   THE SL1MMER ALGORITHM

The artifacts described in last Chapter are not critical when the aim is only to reconstruct a reasonable

reflectivity profile. However, high-quality sparse tomographic SAR inversion requires the estimation of

the number of scatterers, as well as the amplitude, phase, and elevation of each scatterer. Hence, special

care must be taken to those nuisance artifacts. As will be shown, the SL1MMER algorithm improves the

CS algorithm and estimates these parameters in a very accurate and robust way. It consists of three main

steps: 1) a dimensionality scale-down by L_1 norm minimization, 2) model selection and 3) parameter

estimation (see Fig. 2).

A. Scale-down by L_1 norm minimization

For dimensionality reduction, the L_1- L_2 norm minimization from (7) has been implemented at the first

stage which gives a robust estimate of the plausible positions of the scatterers, among which there might

be a few outliers contributed by noise. By only selecting its columns corresponding to the non-zero

elements of $\hat{\gamma}$, the mapping matrix **R** is scaled down significantly. It renders the severely

under-determined system model finally over-determined.
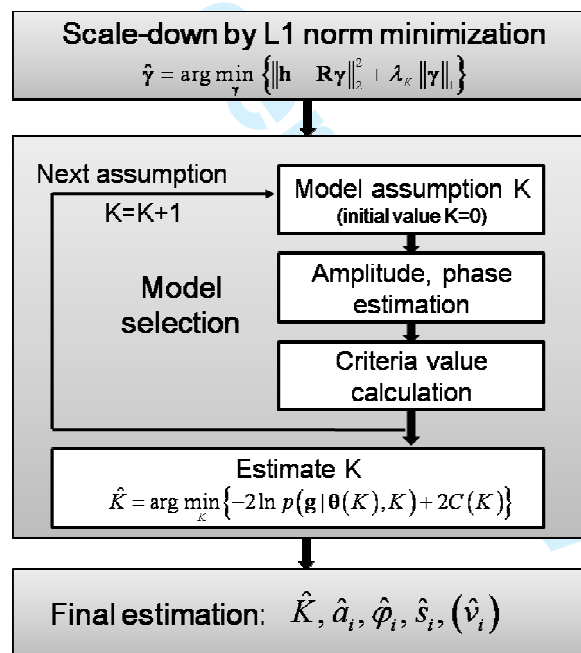
Fig. 2. The flow chart of the SL1MMER algorithm. Note that in the final estimation step, a motion parameter $\hat{v}$ shows up since the SL1MMER algorithm has originally been implemented for *differential* TomoSAR (D-TomoSAR). However, to keep the mathematics of this paper simpler, we neglect motion for our study.

B. Model selection

The L_1-L_2 norm minimization step shrinks **R** dramatically and gives a first sparse estimate of **γ**. This

estimate, though, may still contain the aforementioned outliers (Fig. 1) and, hence, the sparsity K (i.e.

the number of the scatterers) is often overestimated. Model selection is used to clean the $\boldsymbol{\gamma}$ estimate of

spurious, non-significant scatterers and to finally obtain the most likely number $\hat{K}$ of scatterers inside

an azimuth-range cell [17]. It provides a trade-off between the model complexity and the model fit.

The model complexity can be described by the number of parameters k. In our case $k = 3K$, since each

scatterer is characterized by three parameters (amplitude, phase, and elevation). Let further $\boldsymbol{\theta}(K)$ be the

vector of the unknown amplitudes, phases, and elevations for all the K scatterers. The goodness of

model fit can be described by the likelihood $p\left(\mathbf{g} \mid \hat{\boldsymbol{\theta}}(K), K\right)$. A more complex model always fits the

observations better, hence, for the purpose of selecting the true model, the complexity of the model

must be penalized to avoid over-fitting of the data. This reveals the main concept of penalized

likelihood criteria. The general form of penalized likelihood criteria is [18]:

$$\hat{K} = \arg \min_{K} \left\{ -2 \ln p\left(\mathbf{g} \mid \hat{\boldsymbol{\theta}}(K), K\right) + 2C(K) \right\} \tag{8}$$

where $C(K)$ is the complexity penalty. Note that for stationary white Gaussian measurement noise

$-2 \ln p\left(\mathbf{g} \mid \hat{\boldsymbol{\theta}}(K), K\right) = \sigma_{\varepsilon}^2 \left\| \mathbf{g} - \mathbf{R}\hat{\boldsymbol{\gamma}} \right\|_2^2$. Many types of penalty terms are proposed in the literature, e.g. the

Bayesian Information Criterion (BIC) [19], the Akaike information criterion (AIC) [20], and the

Minimum Description Length (MDL) [21]. As illustrated in Fig. 2, for each model assumption

$K = 0...4$, the elevations are given by the previous scale-down step, thus, the amplitudes and phases can

be easily obtained by linear least-squares estimation. The preferred number of scatterers $\hat{K}$ inside an

azimuth-range cell is finally the one with the lowest penalized likelihood criterion value according to

equation (8). Implementation details of different model selection schemes for VHR TomoSAR are

addressed in [3]. We use the BIC penalized likelihood criterion with a penalty term $C(K) = 0.5\,k\ln N = 1.5K\ln N$ throughout the paper with the same parameter setting, i.e. the subsequent comparison between different algorithms is independent on the model selection schemes.

C. Parameter estimation

As a last refinement, a much slimmer mapping matrix $\mathbf{R}(\hat{\mathbf{s}})$, i.e. the $N \times \hat{K}$ matrix with $\mathbf{R}_{n,\hat{k}}(\hat{\mathbf{s}}) = \exp(-j2\pi\xi_n\hat{s}_k)$, is built up, and the final complex-valued reflectivity $\boldsymbol{\gamma}(\hat{\mathbf{s}})$ for the $\hat{K}$ scatterer is obtained by solving the following over-determined linear system equation:

$$\mathbf{g} = \mathbf{R}(\hat{\mathbf{s}})\,\boldsymbol{\gamma}(\hat{\mathbf{s}}) + \mathbf{e} \tag{9}$$

where $\mathbf{e}$ combines the measurement noise and the model error, i.e. the deviation from sparsity or unmodeled motion-induced phase terms (in the D-TomoSAR case). The sparse reflectivity profile $\hat{\boldsymbol{\gamma}}(\hat{\mathbf{s}})$ is finally found by:

$$\hat{\boldsymbol{\gamma}}(\hat{\mathbf{s}}) = \left(\mathbf{R}^H(\hat{\mathbf{s}})\mathbf{R}(\hat{\mathbf{s}})\right)^{-1}\mathbf{R}^H(\hat{\mathbf{s}})\,\mathbf{g} \tag{10}$$

By introducing model selection (8) and refinement of the parameter estimation (10), SL1MMER improves the result of CS by means of removing possible outliers and providing more accurate amplitude and phase estimates.

V.   ESTIMATION ACCURACY

A. The CRLB for Two Scatterers

Assuming two scatterers at elevation positions s_1 and s_2 with amplitudes a_1 and a_2 and phases φ_1 and φ_2, the (noise-free) measurements according to the system model (2) are:

$$\bar{\mathbf{g}}(\boldsymbol{\theta}) = \left[\bar{g}_1(\boldsymbol{\theta}), \bar{g}_2(\boldsymbol{\theta}), \cdots, \bar{g}_N(\boldsymbol{\theta})\right]^T$$

where

$$\bar{g}_n(\boldsymbol{\theta}) = \sum_{q=1}^{2} a_q \exp\left(j\left(2\pi\xi_n s_q + \varphi_q\right)\right) \tag{11}$$

and the parameter vector

$$\boldsymbol{\theta} = [\theta_1, \ \theta_2, \ \cdots, \ \theta_6]^T = [a_1, \ \varphi_1, \ s_1, a_2, \ \varphi_2, \ s_2]^T \tag{12}$$

Then the data vector according to (2), including zero-mean circular Gaussian noise is:

$$\mathbf{g} = \left[g_1, \ldots, g_N\right]^T \tag{13}$$

with mean

$$\bar{g}(\boldsymbol{\theta}) = E\{\mathbf{g}\} = \begin{bmatrix} \exp(j2\pi\xi_1 s_1) & \exp(j2\pi\xi_1 s_2) \\ \vdots & \vdots \\ \exp(j2\pi\xi_N s_1) & \exp(j2\pi\xi_N s_2) \end{bmatrix} \begin{bmatrix} a_1 \exp(j\varphi_1) \\ a_2 \exp(j\varphi_2) \end{bmatrix} \tag{14}$$

and covariance matrix

$$\mathbf{C}_{\varepsilon\varepsilon} = \sigma_\varepsilon^2 \mathbf{I} \tag{15}$$

The likelihood function is given by:

$$p(\mathbf{g}|\boldsymbol{\theta}) = \frac{1}{(2\pi)^N |\mathbf{C}_{\varepsilon\varepsilon}|} \exp\left(-\left(\mathbf{g} - \bar{\mathbf{g}}(\boldsymbol{\theta})\right)^T \mathbf{C}_{\varepsilon\varepsilon}^{-1}\left(\mathbf{g} - \bar{\mathbf{g}}(\boldsymbol{\theta})\right)\right) \tag{16}$$

From the *Fisher information matrix*

$$\mathbf{J} = -E\left\{\frac{\partial^2 \ln p(\mathbf{g}|\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^T}\right\} \tag{17}$$

we derive the CRLB $\mathbf{P}_{CR} = \mathbf{J}^{-1}$. For our idealized data statistics (12-14) [22]:

$$J_{ij} = 2\,\text{Re}\left[\bar{\mathbf{g}}_i'^H \mathbf{C}_{\varepsilon\varepsilon}^{-1} \bar{\mathbf{g}}_j'\right] \tag{18}$$

Where $\bar{\mathbf{g}}_i'$ denotes the derivative of $\bar{\mathbf{g}}$ with respect to the ith element of $\boldsymbol{\theta}$:

$$
\begin{aligned}
\overline{g}_n^{'} &= \left[\frac{\partial \overline{g}_n}{\partial \theta_1}, \quad \cdots, \quad \frac{\partial \overline{g}_n}{\partial \theta_6} \right]^{\mathrm{T}} \\
&= \left[\exp\left(j\left(2\pi\xi_n s_1 + \varphi_1 \right) \right) \begin{bmatrix} 1 & ja_1 & j2\pi\xi_n a_1 \end{bmatrix}, \right. \\
&\qquad \left. \exp\left(j\left(2\pi\xi_n s_2 + \varphi_2 \right) \right) \begin{bmatrix} 1 & ja_2 & j2\pi\xi_n a_2 \end{bmatrix} \right]^{\mathrm{T}}
\end{aligned}
\tag{19}
$$

The Fisher information matrix for our two-scatterer case can be written as:

$$
\mathbf{J} = \begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{12}^{\mathrm{T}} & \mathbf{J}_{22} \end{bmatrix}
\tag{20}
$$

where $\mathbf{J}_{11}$ and $\mathbf{J}_{22}$ describe the Fisher information matrices of the individual isolated scatterers (i.e. in absence of the other one):

$$
\mathbf{J}_{qq} = \frac{2}{\sigma_\varepsilon^2} \cdot \begin{bmatrix} N & 0 & 0 \\ 0 & a_q^2 N & 2\pi a_q^2 \sum_N \xi_n \\ 0 & 2\pi a_q^2 \sum_N \xi_n & 4\pi^2 a_q^2 \sum_N \xi_n^2 \end{bmatrix}, \quad q = 1 \, or \, 2
$$

$\mathbf{J}_{12}$ contributes to the interference between them. Let us define $\Delta =: 2\pi\xi_n\delta_s + \Delta\varphi$, where $\delta_s = s_2 - s_1$ and $\Delta\varphi = \varphi_2 - \varphi_1$ are, respectively, the distance and phase difference between the two scatterers, $\mathbf{J}_{12}$ is given by:

$$
\mathbf{J}_{12} = \frac{2}{\sigma_\varepsilon^2} \cdot \begin{bmatrix} \sum_N \cos\Delta & -a_2 \sum_N \sin\Delta & -2\pi a_2 \sum_N \xi_n \sin\Delta \\ a_1 \sum_N \sin\Delta & a_1 a_2 \sum_N \cos\Delta & 2\pi a_1 a_2 \sum_N \xi_n \cos\Delta \\ 2\pi a_1 \sum_N \xi_n \sin\Delta & 2\pi a_1 a_2 \sum_N \xi_n \cos\Delta & 4\pi^2 a_1 a_2 \sum_N \xi_n^2 \cos\Delta \end{bmatrix}
$$

Note that the elevation aperture positions ξ_n show up in the Fisher information matrix only in the form of moments of their distribution. Hence, for sufficiently large N there will be no difference between a regularily sampled aperture and a uniformly randomly sampled one.

Since the analytic inversion of $\mathbf{J}$ – though possible – leads to a complicated and lengthy expression,

we retrieve the relevant elements of the CRLB matrix by solving the inversion numerically. From the CRLB matrix $\mathbf{P}_{CR} = \mathbf{J}^{-1}$, the elements $P_{CR,3,3} = \sigma_{s1}^2$ and $P_{CR,6,6} = \sigma_{s2}^2$ are most of interest, i.e. the CRLBs of the location estimation errors for the two scatterers.

For the sequel we split the elevation estimation error into two contributions:

$$\sigma_{s_q} = c_0 \cdot \sigma_{s_q,0} \tag{21}$$

where

$$\sigma_{s_q,0} = \frac{\lambda r_0}{4\pi\sqrt{2} \cdot \sqrt{N \cdot SNR_q} \cdot \sigma_b} \tag{22}$$

is the CRLB of the elevation estimates of the qth scatterer in the absence of the other one. σ_b is the standard deviation of the elevation aperture sample positions b_n. For uniformly distributed baselines (randomly or equidistantly spaced) $\sigma_b = \rho_s/\sqrt{12}$ and $\sigma_{s_q,0}$ can be related to the elevation resolution ρ_s by [4]:

$$\sigma_{s_q,0} = \frac{\sqrt{3/2}}{\pi} \frac{\rho_s}{\sqrt{N \cdot SNR_q}} \approx \frac{0.39}{\sqrt{N \cdot SNR_q}} \rho_s \tag{23}$$

and c_0 is the essential interference correction factor for closely spaced scatterers. It is almost independent on N and SNR. As a reference, in Fig. 3 we plot $\sigma_{s_q,0}$, normalized to the Rayleigh resolution unit ρ_s, as a function of $N \cdot SNR_q$ to explore its impact on estimation accuracy where $\sigma_{s_q,0}$ follows exactly a -0.5 power law. In the parameter range we consider here, i.e. $N = 10 \ldots 100$ and SNR $= 0 \ldots 10$ dB, the elevation estimation accuracy of a single scatterer is in the order of roughly 1/100 to 1/10 of the resolution unit.
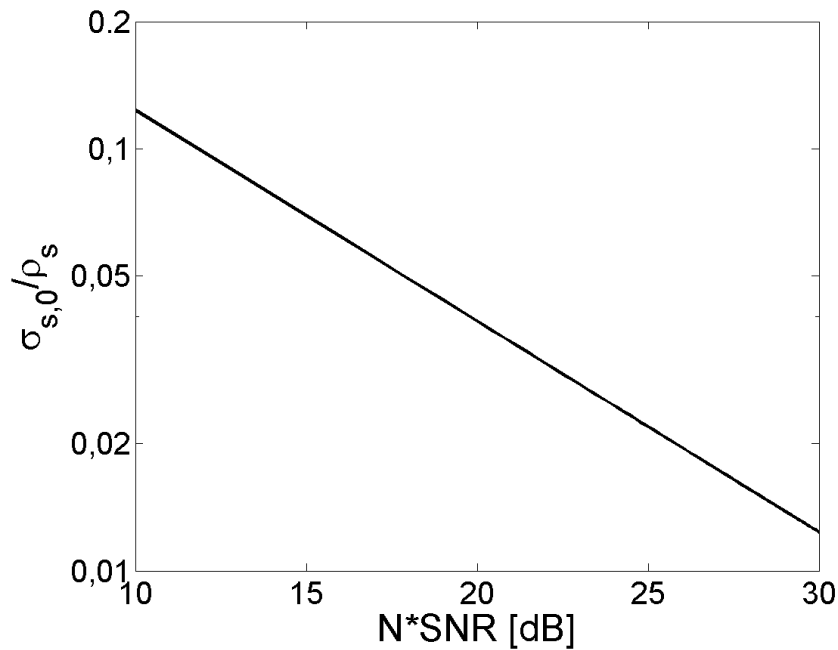
Fig. 3. CRLB $\sigma_{s,0}$ of the elevation estimation accuracy of a single scatterer, normalized to the elevation resolution, as a function of $N \cdot \text{SNR}$

For the subsequent discussion we only consider the impact of the normalized CRLB $c_0 = \dfrac{\sigma_s}{\sigma_{s,0}}$.

If one wants to avoid the numerical inversion of the Fisher information matrix, an experimental approximation of c_0 with a relative error within 2%, was proposed in [23].

Modified after [23]:

$$c_0 = \sqrt{\max\left\{\frac{40\alpha^{-2}\left(1-\alpha/3\right)}{9-6\left(3-2\alpha\right)\cos\left(2\Delta\varphi + 2\pi\alpha\left(\dfrac{2\delta_b}{\Delta b}-\dfrac{1}{N}\right)\right)+\left(3-2\alpha\right)^2}, \, 1\right\}} \tag{24}$$

where $\alpha = \delta_s / \rho_s$ is the distance between the two scatterers, normalized to the Rayleigh resolution unit, and δ_b is the center position of the elevation aperture Δb. Without loss of generality, the elevation aperture may to be set to $\left[-\Delta b, \Delta b\right]/2$, i.e. $\delta_b = 0$. Eq. (24) says, that for a fixed elevation aperture (i.e. Δb and δ_b), σ_{s_q} depends asymptotically on the normalized distance α, $N \cdot \text{SNR}_q$ (for sufficiently large N) and the phase difference $\Delta\varphi$. Fig. 4 ~ Fig. 6 illustrate the normalized CRLB standard

deviation $c_0 = \sigma_s / \sigma_{s,0}$ of the elevation estimation error obtained by numerical inversion of (20).

Fig. 4 describes c_0 as a function of α with different values of $\Delta\varphi$. If the two scatteres are separated by much more than one Rayleigh resolution unit, i.e. $\alpha \gg 1$, their elevation estimation accuracy is the one of individual scatterers, i.e. c_0 approaches 1. The closer they move, the more they interfere and the larger their elevation estimation error becomes, i.e. $c_0 > 1$. The worst case is $\Delta\varphi = 0$ or π, where c_0 follows approximately a -2 power law. The least interference between the scatterers occurs at $\Delta\varphi = \pi/2$ with a c_0 power law of approximately -1. For our SR study the case of $\alpha < 1$ is of interest. Besides the $N \cdot SNR_q$ term in $\sigma_{s,0}$, $1/N$ appears in the argument of the cosine of eq. (24), i.e. for small N, c_0 has a non-negligible additional dependency on N. Fig. 5 shows c_0 as a function of N under different $\Delta\varphi$ with $\alpha = 0.5$, SNR=6dB. Except for the most unfavorable phase differences $\Delta\varphi = 0$ and π, the effect of N is negligible once $N > 15$. Hereafter, we will neglect this effect and all experiments are based on sufficiently large N. The situation of small N will be discussed in Chapter VII.
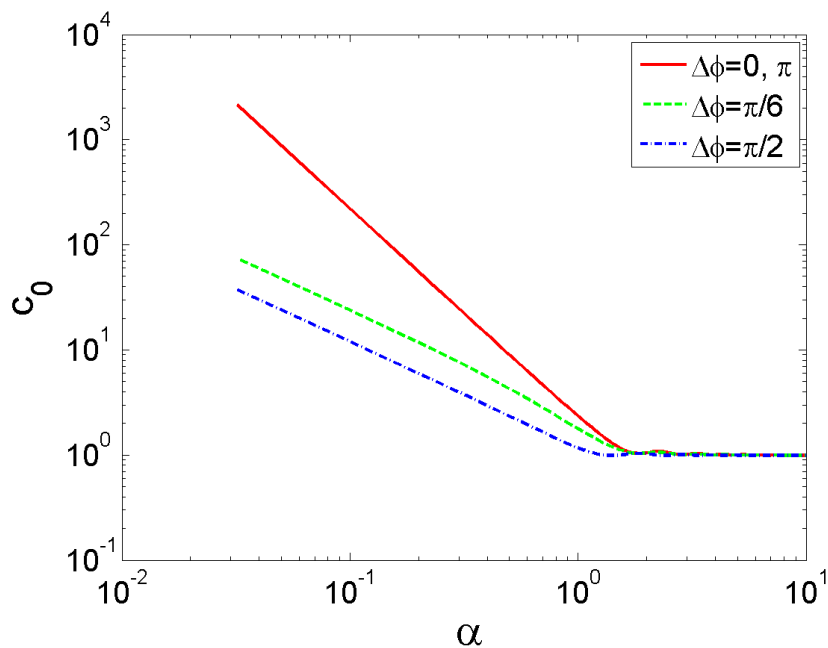
Fig. 4. Normalized CRLB $c_0 = \sigma_s / \sigma_{s,0}$ of the elevation estimation accuracy of two close scatterers, normalized to the one of a single scatterer, as a function of normalized distance α between two scatterers of the same amplitude and with different values of the phase difference $\Delta\varphi$
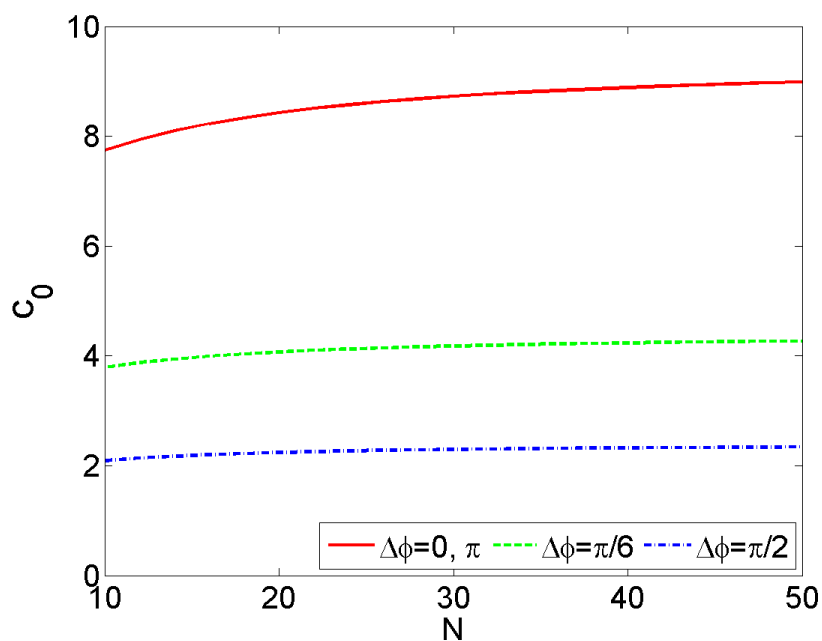
Fig. 5. Normalized CRLB c_0 as function of N with different values of $\Delta\varphi$ (α=0.5, SNR=6dB)
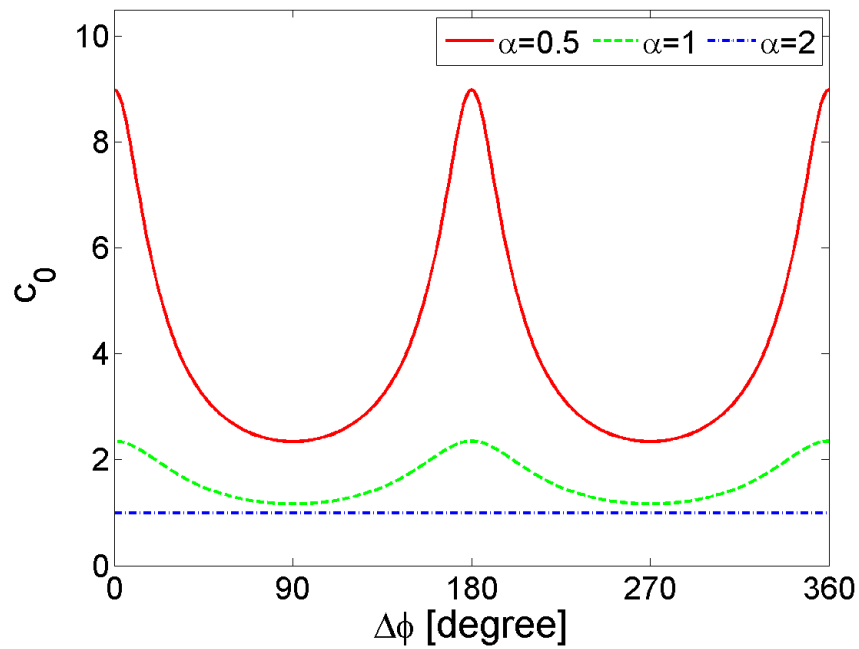
Fig. 6. Normalized CRLB c_0 as a function of the phase difference $\Delta\varphi$ between two scatterers with different values of α

The phase difference $\Delta\varphi$ appears multiplied by two in the argument of the cosine of eq. (24). Therefore, in Fig. 4 and Fig. 5 the cases $\Delta\varphi = 0$ and $\Delta\varphi = \pi$ are identical. Fig. 6 analyzes c_0 as a function of $\Delta\varphi$ with different values of α. Intuitively this phase behavior can be understood as follows: Two scatterers with orthogonal complex-valued reflectivities (i.e. $\Delta\varphi = \pi/2$) use the two "channels" of complex measurements most efficiently and, hence can be better localized and identified than those with non-orthogonal ones. In particular, $\Delta\varphi = 0$ and $\Delta\varphi = \pi$ are the worst cases, i.e. the reflectivities of the two scatterers are collinear in the complex plane. Not surprisingly, the closer the two scatterers are, i.e. the smaller α, the more pronounced is the dependence on their phase difference. Since for our SR study the case of small α is particularly of interest, the impact of $\Delta\varphi$ must be considered. For TomoSAR, the phase difference $\Delta\varphi$ is a random variable because it depends on the unknown geometric configuration [24]. Hence, we calculate the CRLB for the elevation estimation error by integrating the variances over $\Delta\varphi$. The resulting dependence can be approximated by the following

analytic expression:

$$c_0^2 = \frac{\sigma_s^2}{\sigma_{s,0}^2} \approx \max\left\{2.57\left(\alpha^{-1.5} - 0.11\right)^2 + 0.62,\ 1\right\} \tag{25}$$

It fits the CRLB in this range of α better than the approximation derived in [23], $c_0^2 = \max\left\{\left(15/\pi^2\right)^{2.5} \alpha^{-3},\ 1\right\}$. The square-root of the normalized CRLB variance integrated over $\Delta\varphi$ as a function of α is plotted in Fig. 7 together with our approximation. The dependency on $\alpha^{-1.5}$ can be clearly observed.
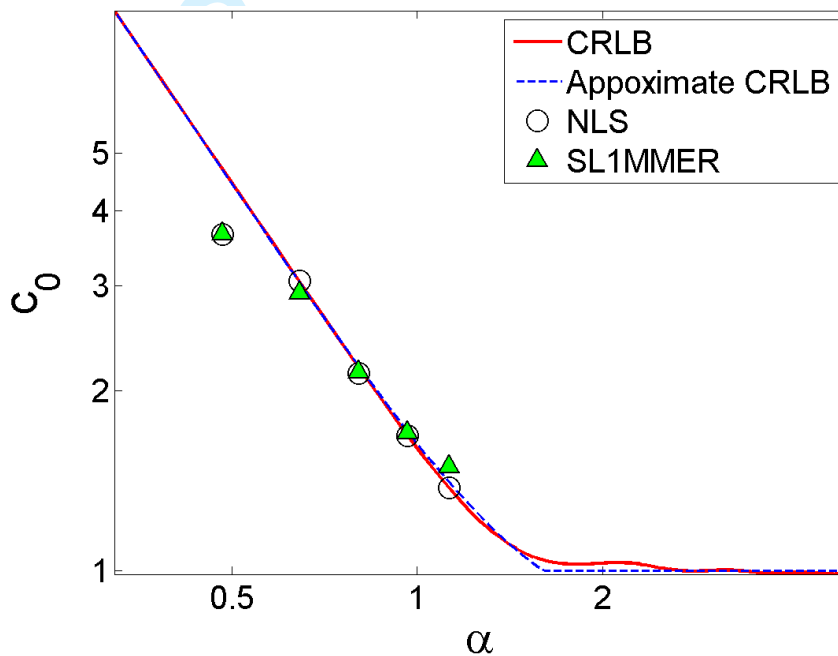
Fig. 7. Theory vs. approximation vs. experiments: Normalized CRLB c_0 integrated over $\Delta\varphi$ as a function of α

B. Elevation Estimation Accuracy

The normalized standard deviation of elevation estimates for two interfering scatterers using NLS and SL1MMER as a function of α is overplotted in Fig. 7. These have been obtained by simulations with randomly distributed phase differences. The elevation estimation accuracy of SL1MMER approaches the CRLB and is comparable to NLS. It is worth mentioning that for the leftmost experiment point, i.e.

for $\alpha = 0.47$, the NLS and SL1MMER pretend to have a better accuracy than the CRLB. This is caused by the fact that we always assume that the two closely spaced scatterers never exchange their positions which is not valid while moving the two scatterers closer such that the standard deviation increases and reaches the order of the distance between them.

VI.   SUPER-RESOLUTION POWER

Non-linear and parametric spectral estimation methods yield reconstructions with much sharper point responses than traditional non-parametric linear algorithms. One might be tempted to infer a very high resolution power from this narrow point response width. However, the needle-like responses tell us neither the location accuracy of the individual points nor the ability of the algorithm to resolve two close scatterers. In this study, the (elevation) resolution is defined by the *minimum distance* ρ_{P_D} between two δ-functions (scatterers) that are separable at a given SNR with a certain number of measurements N, and at a pre-specified probability of detection P_D. As in the preceding chapters we normalize resolution to the Rayleigh resolution unit ρ_s. The *SR factor* κ_{P_D} is defined by:

$$\kappa_{P_d} = \frac{\rho_s}{\rho_{P_D}} \tag{26}$$

It depends on the required detection rate P_D and is larger than unity for SR. In this Chapter, the SR power for different N, SNR, as well as amplitude ratios and phase differences of the two scatterers will be investigated.

A. Detection Rate Study

The problem of resolution can be regarded as a hypothesis test:

H$_0$:   null or only one scatterer inside the given azimuth-range cell

H$_1$:   two scatterers inside the given azimuth-range cell

In [25], the detection probability of two real-valued sinusoids with closely spaced frequencies is studied

and a sophisticated *detector* is proposed. Our goal, however, is to derive the super-resolution power of an

estimator that provides also location, amplitude and phase of each scatterer.

Since it is very complicated to find a theory supporting all possible situations, we approach the problem

experimentally. We start with an example of detecting two scatterers with the same amplitudes and

phases by using N=25 regularly sampled acquisitions. Fig. 8 shows the detection rate P_D as a function of

normalized distance α at different SNR levels using SL1MMER, NLS and linear reconstruction

followed by the detection of the maxima (MD) [3]. The phase difference in this simulation is $\Delta\varphi = 0$, i.e.

the worst case. The SNR of each scatterer for the two sets of curves is 0dB and 6dB, respectively. This

plot gives us several perspectives of the problem:

- The detection curve of conventional MD tells us that the Rayleigh resolution unit is rather a

  definition in absence of noise since the detection rate of MD increases with SNR. And at a

  SNR=0/0dB, even at the distances larger than the resolution unit, e.g. $\alpha = 1.2$, the detection rate

  is still unsatisfactory.

- NLS, as a parametric spectral analysis method, is identical to a MLE under Gaussian noise with

  large N and high SNR. Hence, it should result in the highest detection rate if there is no a prior

  knowledge about the true hypothesis. In Fig. 8, NLS gives a significantly improved probability

of detection compared to MD. It is trivial but important to observe that the detection rate

depends strongly on the SNR·N.

The SR power of SL1MMER is slightly better than NLS. The reason for the "better"

performance of SL1MMER is that the L_1 minimization gives us e.g. two possible positions of

scatterers. The following model selection only decides between taking 0, 1 or 2 of those. NLS,

on the other hand, estimates a new position for the 1-scattrerer case (probably located between

the positions of the 2-scatterer case). This renders the likelihood a little bit higher in NLS and

makes NLS accept the 1-scatterer case a little bit more often.
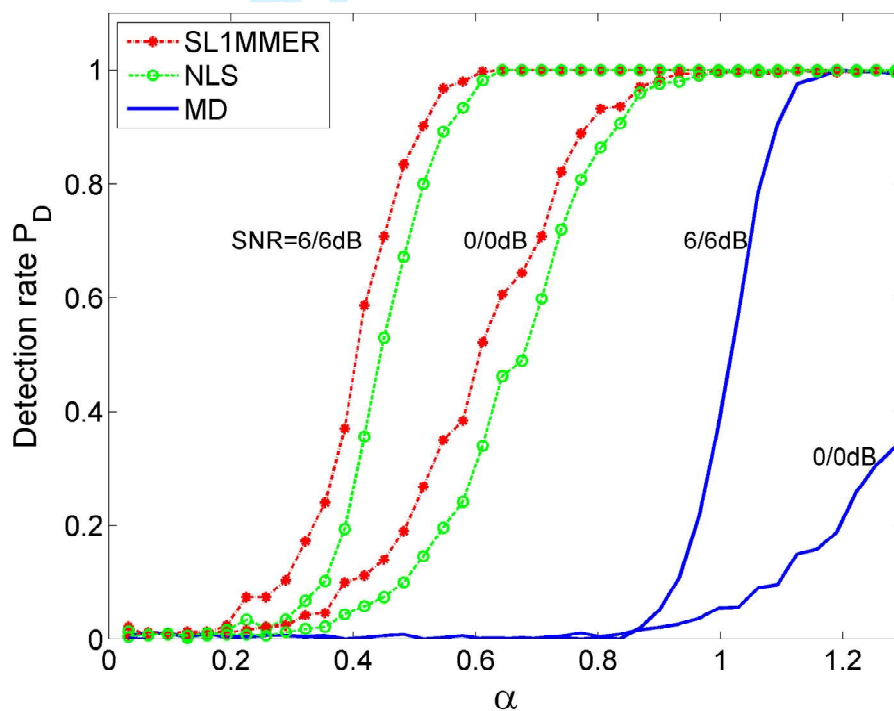
Fig. 8. Detection rate as a function of α using SL1MMER (dashed star), NLS (dashed circle) and MD (solid) with SNR =0/0 dB and 6/6dB, N=25 and $\Delta\varphi$ = 0 (worst case).

Due to the similarity between NLS and SL1MMER, all the following results are approximately

applicable to NLS as well which is asymptotically optimal (see Fig. 7), and hence establish the absolute

limit for SR.

B. Dependency on N·SNR

We expect the SR power of SL1MMER to depend asymptotically on the product N·SNR i.e. the number

of acquisitions can be traded-off against SNR. Fig. 9 shows the detection rates for different N vs. SNR

combinations for N·SNR fixed to 400. As expected, the detection rate curves are consistent with each

other in most cases. Only when the N is very small, e.g. N=10, the detection rate decrease slightly. In

chapter VII we will show that the limit for N·SNR to be an invariant is at N≈15. Of course, besides

N·SNR, the detection rate would also depend on the phase difference of the two scatterers $\Delta\varphi$ which is

assumed to be zero here.
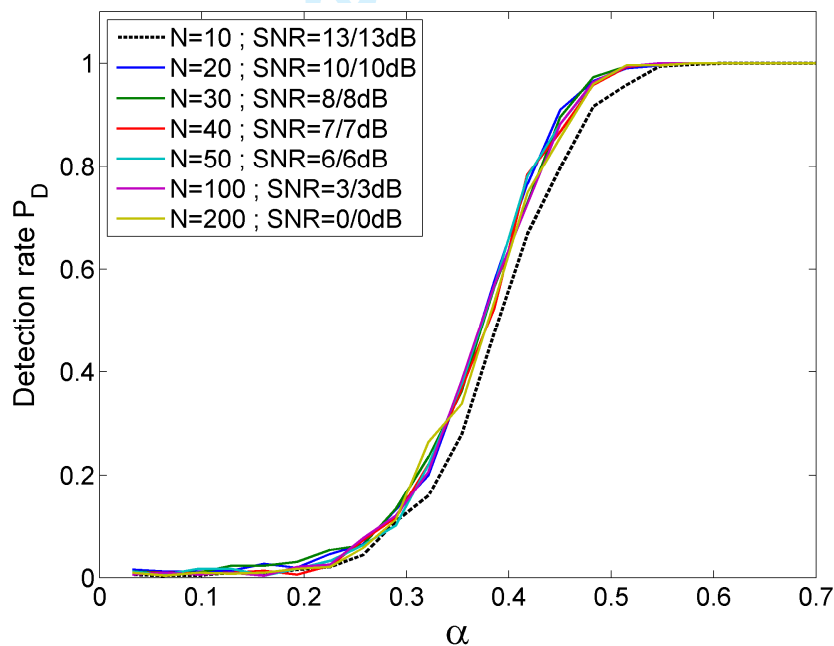
Fig. 9. Detection rate as a function of *α* with fixed *N*·SNR (aperture is regularly sampled) and $\Delta\varphi = 0$

C. Dependency on Irregular Sampling

To give an impression about the impact of irregular sampling on detection rate, Fig. 10 shows an

example of several simulations with N=50 uniformly distributed randomly sampled acquisitions at

SNR=6/6dB. The solid circle curve is the detection rate with regular samples and can be regarded as a

reference, the solid square curve shows the same result using the particular irregularly distributed

sample positions of the real data stack used in [15]. The solid black curve corresponds to irregular

sampling where the sampling distribution is uniform. From this plot, we get two important messages:

Firstly, the curves do not differ much from each other and, hence, irregular sampling does not have a

large impact on SR (we have tried other sampling distributions and they show very little effect.);

Secondly, irregular sampling does not act as an additional difficulty, on the contrary, it can give slightly

better detection rate compared to regular sampling. Based on these observations, without loss of

generality, we will use regular sampling in the range $\left[-\Delta b, \Delta b\right]/2$ for the following examples.
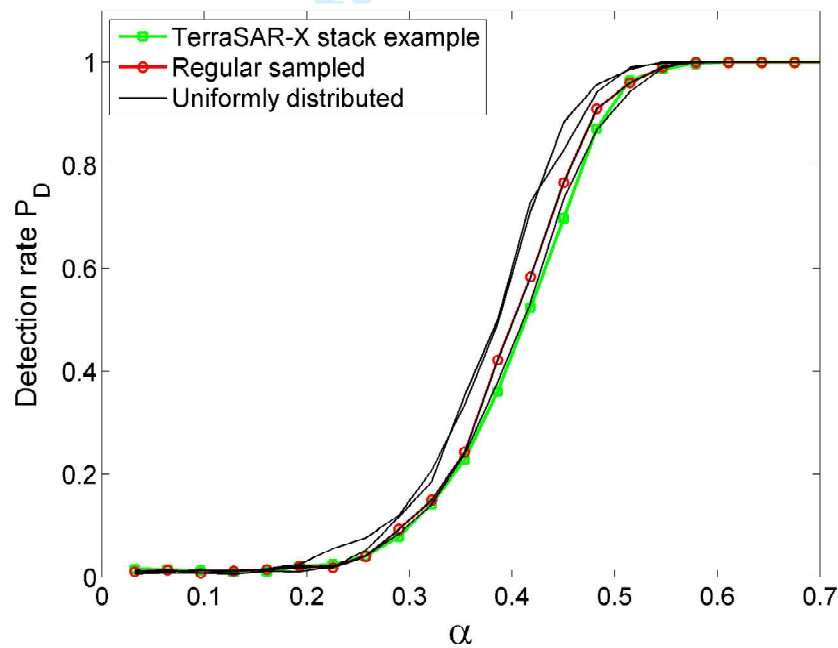
Fig. 10. Detection rate as a function of α at N=50, SNR=6/6dB and $\Delta\varphi$ = 0 with regular sampling, uniformly distributed random sampling, and a particular real-world aperture sample distribution from a TS-X data stack.

D. Dependency on Phase Difference

For $\alpha < 1$, the elevation estimation accuracy depends strongly on the phase difference $\Delta\varphi$ as shown

in last Chapter. Fig. 11 gives an example for the detection rate as a function of $\Delta\varphi$ under different

$N \cdot$SNR with $\alpha = 0.15$. It is not surprising that the detection rate varies dramatically with $\Delta\varphi$ and the

worst case is $\Delta\varphi = 0$.
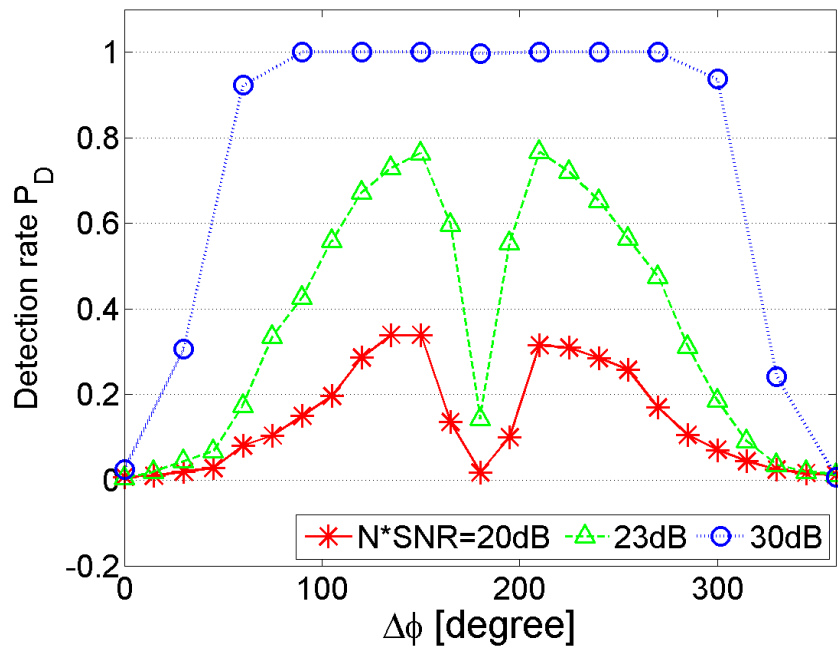
Fig. 11. Detection rate for α=0.15 as a function of $\Delta\varphi$ under different $N \cdot$SNR.

E. A Fundamental Bound for Super-Resolution

Followed by the above dependency analysis, in this section we analyze the super-resolution factor

$\kappa_{50\%} = \rho_s / \rho_{50\%}$, which indicates the center point of the detection curve. Since the phase difference

$\Delta\varphi$ is a random variable and has significant impact on detection rate, we average the detecting curves

and then find the 50% point. The markers in Fig. 12 stand for $\kappa_{50\%}$ from the *simulations* for varying

$N \cdot$SNR within the range of interest (10dB~30dB) under different amplitude ratios. It says the achievable

super-resolution factors in this range typical for TomoSAR are promising and in the order of 1.5 – 25.

For the readers' convenience to further use our findings, we fit the experimental results to the following

polynomial expansion:

$$\kappa_{50\%}\left(N \cdot SNR\right) = \sum_{i=0}^{5} c_{SR,i}\left(N \cdot SNR\right)^{i} \tag{27}$$

where $c_{SR,i}$ is the coefficient for the ith order term of N·SNR. Table 1 summarizes the coefficients for different amplitude ratios a_1/a_2 and the solid lines in Fig. 12 present the results of polynomial fitting. Although being derived experimentally, the results in this section can be considered as general fundamental bounds for SR of imaging system.
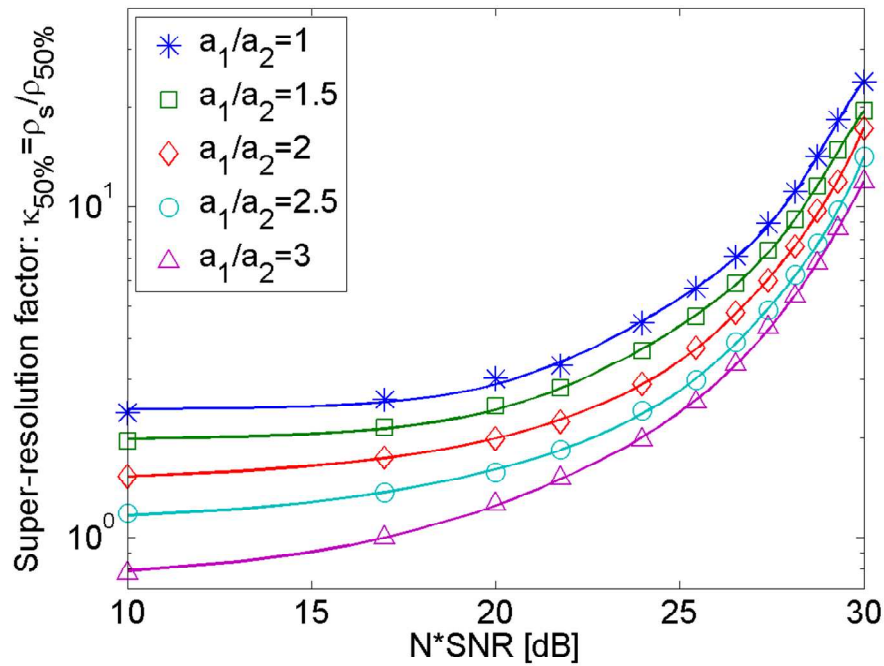
Fig. 12. Super-resolution factor averaged over $\Delta\varphi$ as a function of N·SNR under different amplitude ratios a_1/a_2: Experimental results (markers) vs. polynomial fitting (solid lines).

Table 1. Polynomial approximation of super-resolution factor as a function of N·SNR: coefficients for different amplitude ratios a_1/a_2

	$c_{SR,0}$	$c_{SR,1}$	$c_{SR,2}\left(\times 10^{-4}\right)$	$c_{SR,3}\left(\times 10^{-6}\right)$	$c_{SR,4}\left(\times 10^{-9}\right)$	$c_{SR,5}\left(\times 10^{-12}\right)$
$a_1/a_2=1$	2.4392	-0.0007	0.7116	-0.2013	0.2671	-0.1148
$a_1/a_2=1.5$	1.9717	0.0013	0.4374	-0.1197	0.1616	-0.0694
$a_1/a_2=2$	1.4691	0.0056	-0.0687	0.0392	-0.0444	0.0223
$a_1/a_2=2.5$	1.1108	0.0057	-0.1137	0.0463	-0.0531	0.0256
$a_1/a_2=3$	0.7343	0.0055	-0.0496	0.0147	-0.0064	0.0023

VII.   MINIMUM NUMBER OF ACQUISITIONS

The above shown examples demonstrate that the performance of tomographic reconstruction deviates

from the $N \cdot SNR$ dependency for small N. Yet the required minimal number of acquisitions for robust

reconstruction at a given SNR is of great interest since each SAR acquisition is expensive. In [29], the

minimal number of tracks for subspace methods for TomoSAR is proposed by assuming a sufficient

SNR which is rarely true for the spaceborne case. In this Chapter, the robustness of the algorithm is

defined by the minimal required number of acquisitions $N_{min,90\%}$ at a given SNR, which still allows

detecting two scatterers with a distance of one Rayleigh resolution unit with $\Delta\varphi = 0$ (the worst case)

at a high probability ($P_D \geq 90\%$). At a given SNR, we will provide an explicit expression for $N_{min,90\%}$.

In the preceding chapter we have seen that the detection rate has the same dependent factors as the

CRLB. There is a correction $1/N$ term in c_0 to compensate the small N case. Introducing the same

compensation factor which is N-dependent as in (24), inserting $\alpha = 1$ and making some trivial

normalization, we find that 90% detection rate of two scatterers with a distance of ρ_s (α=1) can be

achieved when the following condition holds:

$$
\begin{cases}
N_{min,90\%} \cdot SNR \cdot \left(2.5 - 1.5\cos\left(2\Delta\varphi + 2\pi\left(1 - \dfrac{1}{N_{min,90\%}} \right) \right) \right)^2 = c_1, & N_{min,90\%} < 15 \\
N_{min,90\%} \cdot SNR = c_2, & N_{min,90\%} \geq 15
\end{cases}
\tag{28}
$$

c_1 and c_2 are constants which depend on the amplitude ratio and phase difference between two

scatterers.

In Fig. 13, the required minimal number of acquisitions $N_{min,90}$ for different amplitude ratios ($a_1=a_2$ (left),

$a_1=2a_2$ (right)) and for $\Delta\varphi = 0$ is presented. The circles show the 90% detection rate positions with

$N \in [8, 20]$. E.g. for a total SNR of 6dB, only 11 acquisitions are required for $a_1=a_2$ (SNR=3/3db respectively) while 17 acquisitions are needed for $a_1=2a_2$ (SNR= 5/−1dB respectively). From the above experimental results, according to eq. (28), we get c_1= 54.2 and c_2= 33.1 for $a_1 = a_2, \Delta\varphi = 0$; and c_1 =107.3 and c_2=65.9 for $a_1 = 2a_2, \Delta\varphi = 0$ The dashed line (N<15) & solid line ($N \geq 15$) in Fig. 13 show the estimated $N_{min,90\%}$ from eq. (28) with the above estimated c_1 and c_2. For $N \geq 15$ the N and SNR can be traded for $N \cdot$SNR = const.
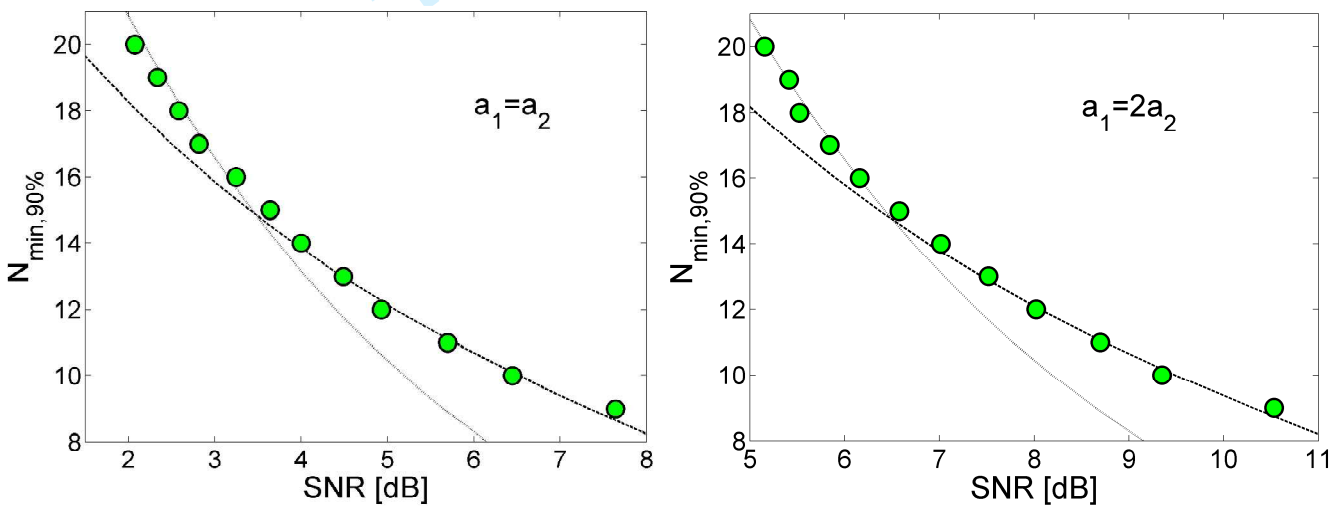
Fig. 13. The required minimal number of acquisitions $N_{min,90\%}$ for robust reconstruction (i.e. well detecting two scatterers (P_D≥90%) with a distance of one Rayleigh resolution unit) at a given SNR: theory (dashed line (N<15) & solid line (N≥15)) vs. experiment (circles). (Left) $a_1=a_2$; (right) $a_1=2a_2$. The worst case $\Delta\varphi$=0 is assumed.

VIII.   CONCLUSION AND FINAL STATEMENTS

In this paper we have investigated the elevation estimation accuracy, super-resolution power and the robustness of sparse signal reconstruction from random and regular spectral samples for the application of TomoSAR. The results are obtained by the compressive sensing based algorithm SL1MMER and are

compared to NLS. They establish thus absolute bounds for super-resolution. In particular, with eq. (25) and (27) we provide easy-to-use analytic expressions for location accuracy and SR factors.

A few additional remarks might be helpful for further use of our results:

- The minimal separable distance between two interfering scatterers depends significantly on the phase difference $\Delta\varphi$ which is random. Hence, we investigated the super-resolution power in a more general way, i.e. assuming $\Delta\varphi$ uniformly distributed in $[-\pi, \pi]$ instead of given a specific value. The theoretically achievable super-resolution factors in the low N and low SNR situation of TomoSAR are promising and in the order of $1.5 - 25$.

- We have shown that regular and irregular aperture sample distributions yield approximately the same results. This is only true for the super-resolution problem, i.e. when two scatterers are *close*. The irregular distribution, though, has an additional advantage if it comes to imaging of reflectivity profiles of *large* support in elevation s. Compared to regular aperture sampling it does not suffer from strong and discrete ambiguities.

- If the sample distribution is not rectangular with width Δb but e.g. Gaussian, then similar results apply as if a rectangular distribution with the same standard deviation σ_b were used. In our equations only Δb has to be replaced by $\sqrt{12}\,\sigma_b$.

- We have restricted ourselves to 1D spectral estimation. Although the results are easily extendable to the 2D case, i.e. motion retrieval by *differential* TomoSAR, there are a few subtleties to be considered. E.g. a possible correlation between the temporal and the spatial coordinates of the 2D

aperture sampling pattern has some influence on the 2D super-resolution power. This will be addressed in a further study.

- Sparsity is the prerequisite of CS based algorithms such as SL1MMER. This property is always ensured for point like scatterers, independent of the SAR system's range or elevation resolutions. Rough surfaces, assumed to be mostly either horizontal or vertical, require $\rho_r << \min\{\rho_s \tan\theta, \quad \rho_s/\sin\theta\}$ for appearing sparse in elevation, where θ is the incidence angle. Experimental results [13] [30], however, show that the CS based estimators degrade gracefully with the loss of sparsity. Even in the case of ERS with an almost isotropic resolution cell moderate improvement over linear estimators can be achieved.

ACKNOWLEDGEMENT

REFERENCES

[1] F. Lombardini, *Differential Tomography: A New Framework for SAR Interferometry*, in Proc. IGARSS, Toulouse, France, pp.1206–1208, 2003.

[2] G. Fornaro, D. Reale, F. Serafino, *Four-Dimensional SAR Imaging for Height Estimation and Monitoring of Single and Double Scatterers*, IEEE Trans. Geosci. Remote Sensing, Vol. 47, pp: 224 – 237, 2009.

[3] X. Zhu, R. Bamler, *Very High Resolution Spaceborne SAR Tomography in Urban Environment,* IEEE Transactions on Geoscience and Remote Sensing 48(12), pp. 4296 - 4308

[4] R. Bamler, M. Eineder, N. Adam, X. Zhu, S. Gernhardt, *Interferometric Potential of High Resolution Spaceborne SAR,* Photogrammetrie Fernerkundung Geoinformation 5/2009, pp. 407-419.

[5] N. Adam, M. Eineder, N. Yague-Martinez and R. Bamler, *High-resolution interferometric stacking with TerraSAR-X*. In: Proc. International Geoscience and Remote Sensing Symposium (IGARSS), Boston, USA, 2008.

[6] X. Zhu, R. Bamler, *Let's Do the Time Warp: Multi-Component Nonlinear Motion Estimation in Differential SAR Tomography*, IEEE Geoscience and Remote Sensing Letters, in press.

[7] E. Candès, *Compressive sampling*, Int. Congress of Mathematics, 3, pp. 1433-1452, Madrid, 2006.

[8] R. Baraniuk, *Compressive sensing*, IEEE Signal Processing Magazine, 24(4), pp. 118-121, 2007.

[9] D. Donoho, *Compressed sensing*, IEEE Trans. on Information Theory, 52(4), pp. 1289 - 1306, 2006.

[10] J. H.G. Ender, *On compressive sensing applied to radar*, Signal Processing, Volume 90, Issue 5, Special Section on Statistical Signal & Array Processing, Pages 1402-1414, 2010.

[11] L. Zhang, M. Xing, C. Qiu, J. Li, and Zh. Bao, *Achieving Higher Resolution ISAR Imaging With Limited Pulses via Compressed Sampling*, IEEE Geoscience and Remote Sensing Letters, vol. 6, no. 3, July 2009.

[12] M. Tello Alonso, P. López-Dekker, J. J. Mallorquí, *A Novel Strategy for Radar Imaging Based on Compressive Sensing*, IEEE Transactions on Geoscience and Remote Sensing 48 (12), pp. 4285 -4295, 2011

[13] A. Budillon, A. Evangelista, G. Schirinzi, *Three Dimensional SAR Focusing from Multi-Pass Signals Using Compressive Sampling*, IEEE Transactions on Geoscience and Remote Sensing 49 (1), pp. 488 – 499, 2011.

[14] A. Budillon, A. Evangelista, G. Schirinzi, *SAR Tomography from sparse samples.* Proceedings of the IEEE International Geoscience & Remote Sensing Symposium (IGARSS), Cape Town, Africa.

[15] X. Zhu, R. Bamler, *Tomographic SAR Inversion by L1 Norm Regularization – The Compressive Sensing Approach*, IEEE Transactions on Geoscience and Remote Sensing, 48(10), pp. 3839-3846.

[16] G. Fornaro, F. Serafino, and F. Lombardini, *Three-dimensional multipass SAR focusing: Experiments with long-term spaceborne data*, IEEE Trans. Geosci. Remote Sens., vol. 43, no. 4, pp. 702–714, Apr. 2005.

[17] X. Zhu, *High-Resolution Spaceborne Radar Tomography*, Master thesis, Lehrstuhl für Methodik der Fernerkundung, Technische Universität München, Lehrstuhlbibliothek, 2008.

[18] J. Rissanen, *A short introduction to Model Selection, Kolmogorov Complexity and Minimum Description Length*, Available online at www.mdl-research.org, 2005.

[19] K. P. Burnham and D. R. Anderson, *Multimodel Inference: Understanding AIC and BIC in Model Selection*. Sociological Methods & Research, Vol. 33, No. 2, 261-304 (2004).

[20] G. Schwarz, *Estimating the dimension of a model*. Annals of Statistics 6(2):461-464, 1978.

[21] J. Rissanen, *Modelling by shortest data description*, Automatica, 14:465-471,1978.

[22] P. Stoica & R. Moses, *Spectral analysis of signals*, Pearson Prentice Hall, 2005.

[23] D. N. Swingler, *Frequency estimation for closely spaced sinsoids: Simple approximations to the Cramér-Rao lower bound*, IEEE Transactions on Signal Processing, Institute of Electrical and Electronics Engineers, Inc, 445 Hoes Ln, Piscataway, NJ, 08854-1331, USA, 1993.

[24] S. Tebaldini, A. M. Guarnieri, *On the Role of Phase Stability in SAR Multibaseline Applications*, Geoscience and Remote Sensing, IEEE Transactions on , vol.48, no.7, pp.2953-2966, 2010

[25] M. Shahram, and P. Milanfar, *Improved Spectral Analysis of Nearby Tones Using Local Detectors*, Proceedings of the 2005 International Conference on Acoustic, Speech, and Signal Processing, Philadelphia, Pennsylvania, March 2005.

[26] A. M. Bruckstein, D. L. Donoho and M. Elad, *From sparse solutions of systems of equations to sparse modeling of signals,* SIAM Rev. 51 34–81, 2009.

[27] S. S. Chen, D. L. Donoho and M. A. Saunders: *Atomic decomposition by basis pursuit,* SIAM Rev. 43 129–59, 2001.

[28] D. Wilkes & J. Cadzow, *The effects of phase on high-resolution frequency estimators, IEEE Transactions on Signal Processing,* 1993, 41, 1319 -1330

[29] M. Nannini, R. Scheiber & A. Moreira, *Estimation of the minimum number of tracks for SAR tomography*, IEEE Transactions on Geoscience and Remote Sensing, 2009, 47, 531-543

[30] X. Zhu, G. Fornaro, unpublished results with ERS data stacks over the city of Roma, 2009

A.6 Zhu, X., Bamler, R., 2011a. Let's Do the Time Warp: Multi-Component Nonlinear Motion Estimation in Differential SAR Tomography. IEEE Geoscience and Remote Sensing Letters 8 (4): 735-739.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 8, NO. 4, JULY 2011
735

Let's Do the Time Warp: Multicomponent Nonlinear Motion Estimation in Differential SAR Tomography

Xiao Xiang Zhu, *Student Member, IEEE*, and Richard Bamler, *Fellow, IEEE*

Abstract—In the differential synthetic aperture radar tomography (D-TomoSAR) system model, the motion history appears as a phase term. In the case of nonlinear motion, this phase term is no longer linear and, hence, cannot be retrieved by spectral estimation. We propose the "time warp" method that rearranges the acquisition dates such that a linear motion is pretended. The multicomponent generalization of time warp rewrites the D-TomoSAR system model to an $(M + 1)$-dimensional standard spectral estimation problem, where M indicates the user-defined motion model order and, hence, enables the motion estimation for all possible complex motion models. Both simulations and real data (from TerraSAR-X spotlight) examples demonstrate the applicability of the method and show that linear and periodic (seasonal) motion components can be separated and retrieved.

Index Terms—Differential synthetic aperture radar tomography (D-TomoSAR), multicomponent nonlinear motion, TerraSAR-X (TS-X), synthetic aperture radar (SAR), time warp.

I. INTRODUCTION

DIFFERENTIAL synthetic aperture radar tomography (D-TomoSAR) uses multibaseline multitemporal SAR acquisitions for reconstructing the 3-D distribution of scatterers and their motion [1]–[3]. For each azimuth–range $(x–r)$ pixel, a scattering profile along the third dimension, elevation (s), is estimated. Ideally, a motion history (e.g., subsidence or building deformation) can be assigned to each elevation position. In many cases, e.g., in urban environment, these scattering profiles are dominated by a few individual scatterers only, typically, zero, one, or two.

The *elevation* estimation of a scatterer exploits the linear relationship between the measured phase and the product of baseline and elevation, and hence, the full arsenal of spectral estimation methods can be applied. *Motion*, however, is often nonlinear (periodic, accelerating, stepwise, etc.) and, hence, does not fit well into the spectral analysis framework. This

is particularly true if very high-resolution mapping of urban infrastructure is to be performed. In this letter, we introduce nonlinear warping of the time axis to accommodate nonlinear motion models. The method is extended to multicomponent motion.

II. IMAGING MODEL

A. General $(3 + M)$-Dimensional SAR Imaging Model

The focused complex-valued measurement g_n at an azimuth–range pixel for the nth acquisition at time t_n ($n = 1, \ldots, N$) is [2]

$$g_n = \int_{\Delta s} \gamma(s) \exp\left(-j2\pi\left(\xi_n s + 2d(s, t_n)/\lambda\right)\right) ds \quad (1)$$

where $\gamma(s)$ represents the reflectivity function along elevation s with an extent of Δs and $\xi_n = -2b_n/(\lambda r)$ is the spatial (elevation) frequency proportional to the respective aperture position (baseline) b_n. $d(s, t_n)$ is the line-of-sight (LOS) motion as a function of elevation and time. The motion relative to the master acquisition may be modeled using a linear combination of M base functions $\tau_m(t_n)$

$$d(s, t_n) = \sum_{m=1}^{M} p_m(s)\tau_m(t_n) \quad (2)$$

where $p_m(s)$ is the corresponding motion coefficient to be estimated. In Section III, we will show that $\tau_m(t_n)$ can be interpreted as a warped time variable if we choose the units of the coefficients appropriately. The choice of the base functions depends on the underlying physical motion processes.

B. Special Case I: PSI System Model

Persistent scatterer (PS) interferometry (PSI) assumes the presence of only a single dominant scatterer in each pixel, and only the phase of each acquisition is used

$$\arg(g_n) = 2\pi\left(\xi_n s_0 + 2d(t_n)/\lambda\right). \quad (3)$$

This restriction means that the motion depends only on t_n, i.e., $d(t_n) = \sum_{m=1}^{M} p_m \tau_m(t_n)$, and the motion parameters and the elevation of each scatterer can be estimated by model fitting and phase unwrapping, e.g., by the LAMBDA algorithm [4].

C. Special Case II: D-TomoSAR With Linear Motion

By assuming a *linear* motion with (elevation-dependent) velocity $V(s)$ (but possibly several scatterers per pixel), (1) simplifies to

$$g_n = \int_{\Delta s} \gamma(s) \exp\left(-j2\pi\left(\xi_n s + \eta_n V(s)\right)\right) ds \qquad (4)$$

where $\eta_n = 2t_n/\lambda$ is the so-called "velocity frequency." Equation (4) can be rewritten as [5]

$$g_n = \int_{\Delta s} \int_{\Delta v} \gamma(s)\delta\left(v - V(s)\right) \exp(-j2\pi(\xi_n s + \eta_n v))\, dv\, ds,$$
$$n = 1, \ldots, N \quad (5)$$

where Δv is the range of possible velocities. Equation (5) is a 2-D Fourier transform of $\gamma(s)\delta(v - V(s))$ which is a delta line in the elevation–velocity (s–v) plane along $v = V(s)$. Its projection onto the elevation axis $\gamma(s)\delta(v - V(s))$ is the reflectivity profile $\gamma(s)$.

If only K discrete scatterers are present, $\gamma(s)\delta(v - V(s))$ consists of K δ-points in the 2-D (s–v) plane at the respective elevations and velocities. The tomographic reconstruction will be a blurred and noisy version of these peaks from which elevations and velocities are estimated.

D. Velocity Spectrum

In the original D-TomoSAR paper [2], the concept of the velocity spectrum has been introduced. It generalizes the system model (5) by replacing $\delta(v - V(s))$ by an arbitrary motion term $a(s, v)$

$$g_n = \int_{\Delta s} \int_{\Delta v} \gamma(s)a(s,v) \exp(-j2\pi(\xi_n s + \eta_n v))\, dv\, ds \quad (6)$$

where

$$\int_{\Delta v} a(s,v) \exp(-j2\pi\eta_n v)\, dv = \exp\left(-j\frac{4\pi}{\lambda}d(s, t_n)\right). \quad (7)$$

Estimates of $\gamma(s)a(s, v)$ are derived from the measurements g_n by 2-D spectral analysis. Individual scatterers with *linear* motion will show up as peaks in the (s–v) plane at their respective elevations and velocities. In the case of moderately nonlinear motion and sufficient SNR, these peaks are blurred in the v-direction, but their elevations are still detectable [2]. Then, for each of these elevations, a Fourier transform in the v-direction leads to an estimate of $\exp(-j4\pi d(s, t_n)/\lambda)$ and, hence, of the motion $d(s, t_n)$ [2], [6].

This method works well with single scatterers and for multiple scatterers as long as they can be separated in the (s–v) plane. There are, however, situations where it fails.

1) If the motion is too nonlinear, e.g., periodic, and the signal-to-noise ratio (SNR) is not high enough, the scatterer cannot be detected in the (s–v) plane.
2) If two scatterers are closer in elevation and mean velocity than the respective resolution elements, their velocity spectra will influence each other with an unpredictable effect on the reconstruction of $d(s, t_n)$.
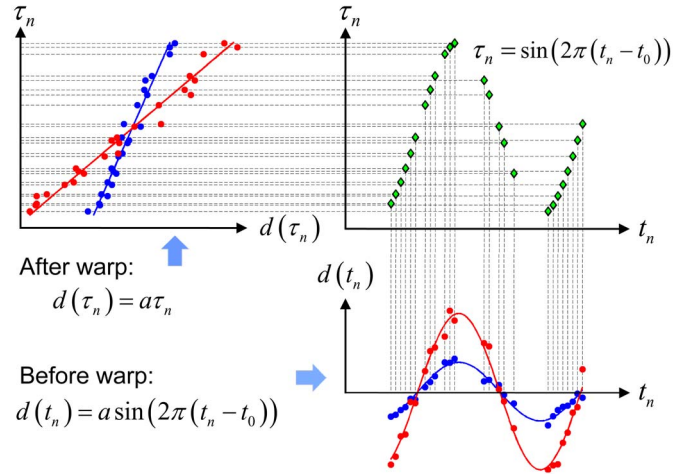
Fig. 1.   Time warp converts seasonal periodic motions with different amplitudes (red and blue) into linear functions of different slopes.

From this discussion, it is evident that we need a transformation that is able to generate peaks of high SNR in some motion coefficient space just like spectral estimation works for linear motion.

III. GENERALIZED TIME WARP METHOD

We have briefly introduced an approach in [7] to cope with the problem of D-TomoSAR reconstruction with single-component nonlinear motion by rearranging the acquisitions on the time axis (the so-called "time warp"). In this letter, we generalize it to the multicomponent model.

A. Time Warp for $M = 1$

In the case of $M = 1$, by introducing the temporal frequency $\eta_{1,n} = 2\tau_{1,n}(t_n)/\lambda$ as a function of a *motion base function* $\tau_{1,n}(t_n)$, and the *motion coefficient* $p_1(s)$, the proposed time warp method leads to a generalized system description which can be adapted for different nonlinear motion models

$$g_n = \int_{\Delta p} \int_{\Delta s} \gamma(s)\delta\left(p_1 - p_1(s)\right) \exp\left(-j2\pi(\xi_n s + \eta_{1,n} p_1)\right) ds\, dp_1,$$
$$n = 1, \ldots, N. \quad (8)$$

The time warp rewrites the D-TomoSAR model with single-component motion (linear or nonlinear) to a standard 2-D spectral estimation problem which makes all spectral estimators applicable. This principle is shown for a periodic motion model in Fig. 1. The most common motion base functions are as follows.

1) *Linear motion*: $\eta_{1,n} = 2t_n/\lambda$, and the coefficient $p_1(s)$ stands for the LOS velocity (v) as a function of s.
2) *Seasonal motion*: $\tau_1(t_n) = \sin(2\pi(t_n - t_0))$, and $p_1(s)$ stands for the amplitude (a) of the periodic motion; t_0 is the initial phase offset.
3) *Thermal dilation*: $\tau_1(t_n) = T(t_n)$, i.e., the temperature, and $p_1(s)$ is the scaling factor of the thermal dilation along s.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

ZHU AND BAMLER: MULTICOMPONENT NONLINEAR MOTION ESTIMATION

737
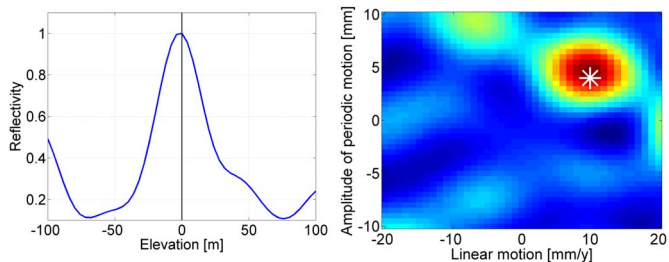
Fig. 2. Multicomponent motion reconstruction example with linear and seasonal periodic motion. Single scatterer with an elevation of 0 m, a linear motion of 10 mm/year, and a seasonal motion with an amplitude of 4 mm (SNR = 3 dB). (Left) Projection of the 3-D spectral estimates to elevation direction (black line shows the true elevation position). (Right) Projection to the $a–v$ motion plane (asterisk marks the true position).
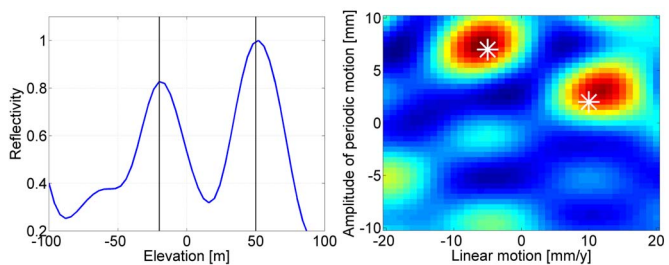
Fig. 3. Similar simulation as in Fig. 2 but with *two* scatterers inside the cell with elevations of $-20/50$ m, linear motion of $10/-5$ mm/year, and seasonal motion amplitudes of $2/7$ mm.

B. Let's Do the Time Warp Again

The generalization of the time warp method for $M > 1$ is straightforward. Let us define the mth temporal frequency component at t_n as $\eta_{m,n} = 2\tau_m(t_n)/\lambda$.

Then, (1) can be rewritten as an $M + 1$-dimensional Fourier transform of $\gamma(s)\delta(p_1 - p_1(s), \ldots, p_M - p_M(s))$ which is a delta line in the $M + 1$ elevation–motion parameter space

$$g_n = \int_{\Delta p_M} \cdots \int_{\Delta p_1} \int_{\Delta s} \gamma(s)\delta\left(p_1 - p_1(s), \ldots, p_M - p_M(s)\right)$$
$$\cdot \exp\left(-j2\pi(\xi_n s + \eta_{1,n}p_1 + \cdots + \eta_{M,n}p_M)\right)$$
$$\times ds\,dp_1 \ldots dp_M, \quad n = 1, \ldots, N. \quad (9)$$

This extension is a general solution to the multicomponent nonlinear motion estimation problem for D-TomoSAR and, hence, completes the time warp concept for complex motion models.

Multicomponent motion reconstruction examples, i.e., combined linear and seasonal motion, with single (see Fig. 2) and double scatterers (see Fig. 3) inside a resolution cell, respectively, are presented here to illustrate this concept. A realistic SNR of 3 dB and a baseline configuration of the TS-X "asterisk" stack with 30 acquisitions used in Section IV (see Fig. 5) are simulated in both examples. We start with a single scatterer located at zero elevation undergoing a linear motion of 10 mm/year and a seasonal motion whose amplitude is 4 mm. In this case, $M = 2$; therefore, we are dealing with a 3-D spectral estimation problem. The left plot in Fig. 2 shows a projection of the 3-D spectral estimates to the elevation direction, i.e., a sum over the linear and periodic motion
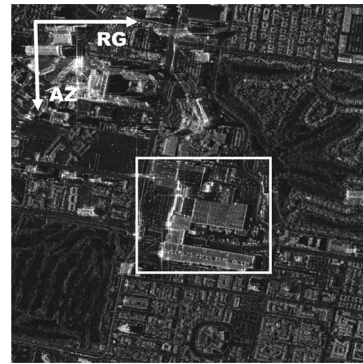
Fig. 4. Mean intensity map of the TS-X spotlight data stack over the test site in Las Vegas, Nevada. White box indicates the investigation area in Fig. 8.
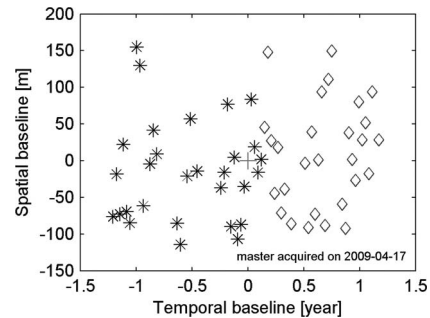
Fig. 5. Spatial–temporal baseline distribution. (Asterisks) Substack with seasonal motion only and (diamonds) substack with seasonal and linear motion.

plane. The right plot shows the corresponding projection in the amplitude–velocity $(a–v)$ plane, and the white asterisk indicates the true amplitude and velocity. Fig. 3 uses the same configuration as Fig. 2 but with two scatterers, located at -20 and 50 m (ca. 2× Rayleigh elevation resolution), with a linear motion of 10 and -5 mm/year and a seasonal motion with amplitudes of 2 and 7 mm, respectively. Comparing the estimates to the simulation truth, it demonstrates that the generalized time warp method can give a robust estimation of the multicomponent nonlinear motion even under a moderate SNR.

IV. EXPERIMENTS WITH REAL DATA

A. Test Site

The proposed method has been applied to TerraSAR-X high-resolution spotlight data with a slant-range resolution of 0.6 m and an azimuth resolution of 1.1 m, over the city of Las Vegas. A stack of 60 SAR images covering a time period of more than two years, from February 2008 to June 2010, is used. The preprocessing, including atmospheric phase screen correction, is performed by the German Aerospace Center's PSI-GENESIS system on a sparse PS network of high-SNR pixels containing only single scatterers. Fig. 4 shows a temporally averaged intensity map of the whole stack over the test area centered at the convention center in Las Vegas. Fig. 5 shows the elevation aperture positions and temporal baselines relative to the master image acquired on April 17, 2009.

This stack is particularly interesting for our purpose: No long-term motion has been observed in the test area over the first year of acquisition (the "asterisk" substack in Fig. 5), i.e., the motion-induced phase is only contributed by thermal dilation. Since July 2009, however, this area has suddenly been
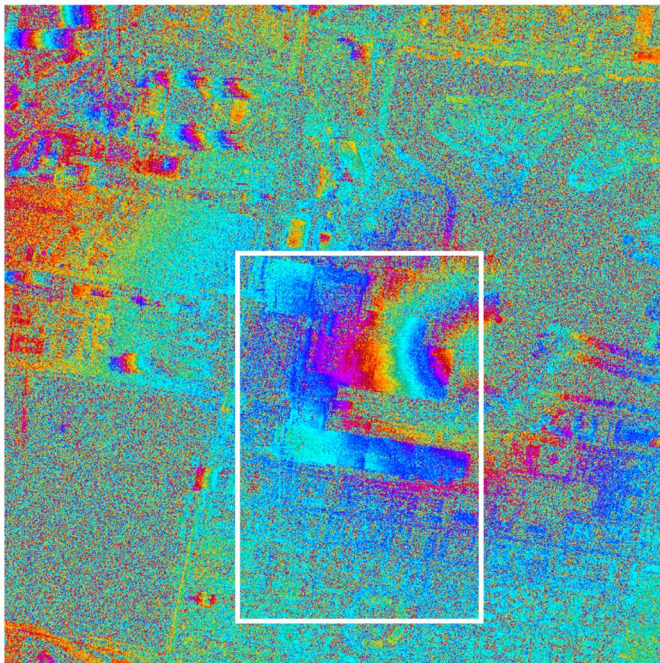
Fig. 6. Differential interferogram of the test site showing a pronounced circular subsidence pattern. The interferogram has been generated by two TS-X images taken from the "diamond" substack in Fig. 5. White box indicates the investigation area in Fig. 9.
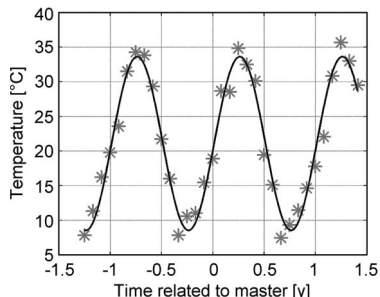
Fig. 7. (Asterisk) Monthly average temperature of Las Vegas, Nevada, in the years of 2008–2010 versus (solid line) modeled seasonal motion with initial offset $t_0 = 0.013$ years.

undergoing a pronounced subsidence centered at the convention center which can be seen from the differential interferogram in Fig. 6 generated by two images taken on April 17, 2009 (master), and April 4, 2010 (slave). Hence, the second-half acquisitions (the "diamond" substack in Fig. 5) are characterized by a multicomponent nonlinear motion, i.e., a combined linear and thermal-dilation-induced seasonal motion. We will use these favorable motion properties to validate the generalized time warp method for single-component and multicomponent nonlinear motion estimation, respectively.

B. Thermal-Dilation-Induced Seasonal Motion Estimation

In this section, the "asterisk" substack is used to estimate the thermal-dilation-induced motion. Theoretically speaking, we should use the temperature history at the acquisition times as the motion base function and find out the scaling factor. However, due to the lack of ground truth, we use a simple sine function with a period of one year. The initial offset is estimated by fitting the sine model to the monthly average temperature of

the test site in the years of 2008–2010 as shown in Fig. 7. The asterisks mark the monthly average temperatures [8], and the solid line drafts the modeled seasonal motion with initial offset $t_0 = 0.013$ years. Likewise, the initial offset can be obtained from the PSI-based preprocessing mentioned in Section IV-A.

The surface model and amplitude map of seasonal motion have been obtained by the proposed time warp method for the whole building complex of the convention center (marked by a box in Fig. 4). Fig. 8 shows the surface model generated from the elevation estimates (converted to height). The shapes of the building and the surrounding infrastructure, e.g., roads and bridges, have been captured at a very detailed level. The right image in Fig. 8 shows the amplitude map of the seasonal motion caused by thermal dilation. Since the building has a metallic structure, it has been more severely affected by the thermal dilation than the surrounding infrastructure. The building shows a smooth amplitude variation for individual structural blocks. There are sudden changes of the seasonal motion amplitude between the adjacent blocks of up to 8 mm. The surrounding area shows almost no seasonal motion.

C. Linear Motion + Seasonal Motion Estimation

The generalized time warp method for $M = 2$ has been applied to the "diamond" substack in Fig. 5 by choosing linear and seasonal motion as the base functions. The left image in Fig. 9 shows the TS-X intensity map of the region of interest (marked by a box in Fig. 6). According to Fig. 6, the center of the subsidence pattern, i.e., the "epicenter," is located on the right upper part of the intensity map. Therefore, together with the seasonal motion results shown in Fig. 8, we can expect the following: 1) only the building structures suffer from thermal dilation, and 2) the closer to the "epicenter" the point is, the bigger is the linear subsidence.

To validate the proposed generalized time warp method, the two pixels P_1 and P_2 marked by red points have been selected and will be analyzed hereinafter (see the left image in Fig. 9; P is the reference point). As P_1 is located outside the region of the convention center, it is expected to contain only a single-component linear motion. P_2 is located on the roof and is closer to the "epicenter." Hence, we expect a combined seasonal and more significant linear motion. The middle and right image pairs in Fig. 9 shows the corresponding estimated reflectivity profiles (upper) and scatterer distributions in the $(a$–$v)$ motion plane (bottom) of the analyzed ground pixel P_1 and roof pixel P_2, respectively. The estimates reveal a single scatterer in the pixel P_1 with an elevation of -12 m (≈ -6.3 m in height relative to reference pixel), almost no seasonal motion, and a subsidence of -5 mm/year. Also, P_2 is estimated to be a single scatterer, with an elevation of 32 m (≈ 16.9 m in height), a seasonal motion with an absolute amplitude of 4.5 mm, and a subsidence of -10 mm/year. These results are consistent with our expectation and, hence, validate the capability of the time warp method for multicomponent nonlinear motion estimation.

V. CONCLUSION

The proposed time warp method for D-TomoSAR converts any nonlinear multiple (M) component motion history into a linear one. It renders D-TomoSAR an ($M + 1$)-dimensional

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

ZHU AND BAMLER: MULTICOMPONENT NONLINEAR MOTION ESTIMATION

739
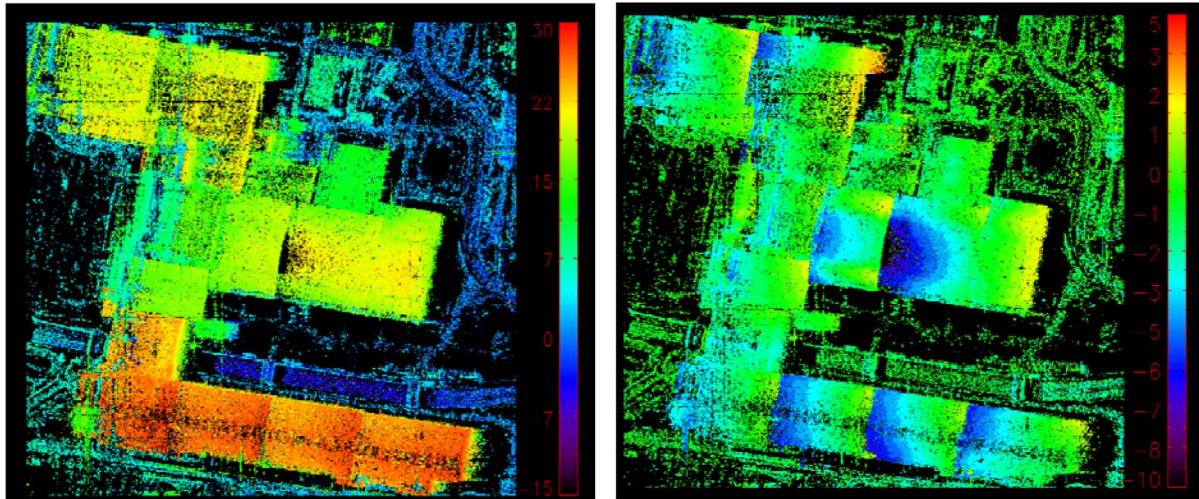
Fig. 8. (Left) Reconstructed digital surface model from differential TomoSAR (unit: meters) and (right) estimated amplitude of seasonal motion (unit: millimeters).
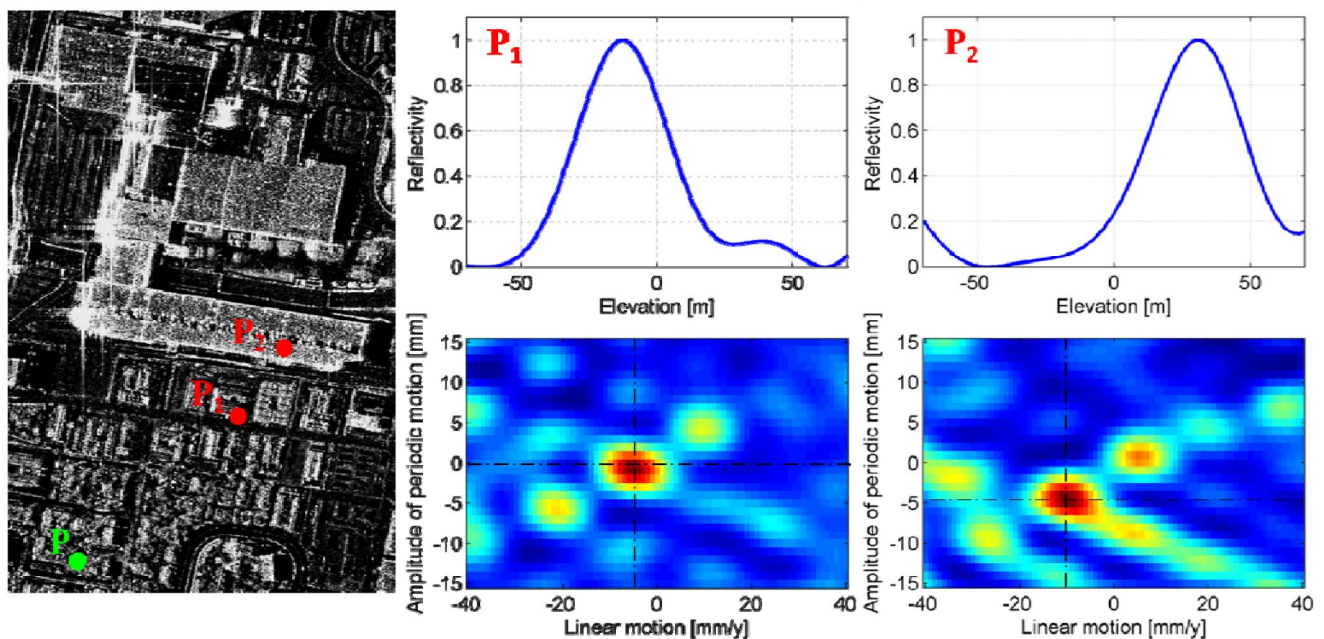
Fig. 9. Validation examples of the generalized time warp approach. (Left) TS-X intensity map. (Middle) (Upper) Reflectivity profile and (bottom) scatterer distribution in $(a–v)$ motion plane of analyzed ground pixel P_1. (Right) Same plots of analyzed roof pixel P_2. P is the reference point. Dashed lines mark the location of the maxima in the spectral estimates.

spectral estimation problem. The choice of the motion base functions reflects our prior knowledge about the underlying motion process. An interesting option is to use GPS measurements in the vicinity of the objects of interest for constructing the base functions. These measurements can provide important features of the motion process, e.g., start time and acceleration rate of a subsidence.

References

[1] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," in *Proc. IGARSS*, Toulouse, France, 2003, pp. 1206–1208.

[2] G. Fornaro, D. Reale, and F. Serafino, "Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 224–237, Jan. 2009.

[3] X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4296–4308, Dec. 2010.

[4] B. M. Kampes, "Radar interferometry: The persistent scatterer technique," in *Remote Sensing and Digital Image Processing Series*, 1st ed. New York: Springer-Verlag, 2006.

[5] X. Zhu and R. Bamler, "Tomographic SAR inversion by L1 norm regularization—The compressive sensing approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3839–3846, Oct. 2010.

[6] G. Fornaro, D. Reale, and F. Serafino, "4D SAR focusing: A tool for improved imaging and monitoring of urban areas," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2008, pp. V-475–V-478.

[7] X. Zhu and R. Bamler, "Very high resolution SAR tomography via compressive sensing," in *Proc. ESA FRINGE Workshop, Adv. Sci. Appl. SAR Interferometry*, Frascati, Italy, 2009.

[8] *Average Daily Temperature Archive*. [Online]. Available: http://www.wrh. noaa.gov/vef/motemps.php

References

Adam, N., Bamler, R., Eineder, M., Kampes, B., 2005. Parametric estimation and model selection based on amplitude-only data in ps-interferometry. In: Proceedings of FRINGE.

Adam, N., Eineder, M., Yague-Martinez, N., Bamler, R., 2008. High resolution interferometric stacking with TerraSAR-X. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 2, II–117.

Bamler, R., 2010. Introduction to SAR Interferometry. SAR Principles and Application, Carl-Cranz-Gesellschaft (CCG) course, Oberpfaffenhofen.

Bamler, R., Eineder, M., 2005. Accuracy of differential shift estimation by correlation and split-bandwidth interferometry for wideband and delta-k SAR systems. IEEE Geoscience and Remote Sensing Letters 2 (2): 151–155.

Bamler, R., Eineder, M., Adam, N., Zhu, X., Gernhardt, S., 2009. Interferometric potential of high resolution spaceborne SAR. Photogrammetrie-Fernerkundung-Geoinformation 2009 (5): 407–419.

Bamler, R., Hartl, P., 1998. Synthetic aperture radar interferometry. Inverse problems 14: R1–R54.

Baraniuk, R., 2007. Compressive sensing. IEEE Signal Processing Magazine 24 (4): 118.

Baraniuk, R., Steeghs, P., 2007. Compressive radar imaging. In: Proceedings of IEEE Radar Conference, 128–133.

Baselice, F., Budillon, A., Ferraioli, G., Pascazio, V., 2009. Layover Solution in SAR Imaging: A Statistical Approach. IEEE Geoscience and Remote Sensing Letters 6 (3): 577 –581.

Berardino, P., Fornaro, G., Lanari, R., Sansosti, E., 2003. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. IEEE Transactions on Geoscience and Remote Sensing 40 (11): 2375–2383.

Berger, C., Zhou, S., Willett, P., 2008a. Signal extraction using compressed sensing for passive radar with OFDM signals. In: Proceedings of IEEE International Conference on Information Fusion, 1–6.

Berger, C., Zhou, S., Willett, P., Demissie, B., Heckenbach, J., 2008b. Compressed sensing for OFDM/MIMO radar. In: Proceedings of IEEE 42nd Asilomar Conference on Signals, Systems and Computers, 213–217.

Bhattacharya, S., Blumensath, T., Mulgrew, B., Davies, M., 2007. Fast encoding of synthetic aperture radar raw data using compressed sensing. In: Proceedings of IEEE/SP 14th Workshop on Statistical Signal Processing, IEEE, 448–452.

Bienvenu, G., 1979. Influence of the spatial coherence of the background noise on high resolution passive methods. In: Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Vol. 4, 306–309.

Boyd, S. P., Vandenberghe, L., 2004. Convex optimization. Cambridge Univ Pr.

Bozdogan, H., 1987. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika 52 (3): 345–370.

Buckreuss, S., 2009. personal communication.

Budillon, A., Evangelista, A., Schirinzi, G., 2009. SAR tomography from sparse samples. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 4, 865–868.

Budillon, A., Evangelista, A., Schirinzi, G., 2011. Three-Dimensional SAR Focusing From Multipass Signals Using Compressive Sampling. IEEE Transactions on Geoscience and Remote Sensing 49 (1): 488 –499.

Burnham, K., Anderson, D., 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociological Methods & Research 33 (2): 261.

Candès, E., 2006. Compressive sampling. In: Proceedings of the International Congress of Mathematicians, Vol. 3, Citeseer, 1433–1452.

Candès, E., Romberg, J., Tao, T., 2006. Robust uncertainty principles: Exact signal recon-

struction from highly incomplete frequency information. IEEE Transactions on Information Theory 52 (2): 489–509.

Candès, E., Tao, T., 2005. Decoding by linear programming. Information Theory, IEEE Transactions on 51 (12): 4203–4215.

Candès, E., Wakin, M., 2008. An introduction to compressive sampling. IEEE Signal Processing Magazine 25 (2): 21–30.

Capon, J., 1969. High-resolution frequency-wavenumber spectrum analysis. Proceedings of the IEEE 57 (8): 1408–1418.

Chan, C., Farhat, N., 1981. Frequency swept tomographic imaging of three-dimensional perfectly conducting objects. IEEE Transactions on Antennas and Propagation 29 (2): 312–319.

Cheney, M., 2000. Tomography problems arising in synthetic aperture radar. In: Proceedings of AMS-IMS-SIAM Joint Summer Research Conference on Radon Transforms and Tomography, Vol. 278, Amer Mathematical Society, 15.

Cloude, S., 2006. Polarization coherence tomography. Radio Science 41 (4): RS4017.

Cloude, S., 2007. Dual-baseline coherence tomography. IEEE Geoscience and Remote Sensing Letters 4 (1): 127–131.

Cloude, S., Papathanassiou, K., 1998. Polarimetric SAR interferometry. IEEE Transactions on Geoscience and Remote Sensing 36 (5): 1551–1565.

Colesanti, C., Locatelli, R., Novali, F., 2002. Ground deformation monitoring exploiting SAR permanent scatterers. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 2, 1219–1221.

Cumming, I., Wong, F., 2004. Digital Signal Processing of Synthetic Aperture Radar Data: Algorithms and Implementation. Artech House.

Curlander, J., McDonough, R., 1991. Synthetic aperture radar: Systems and signal processing. John Wiley and Sons, Inc, New York, 1991. .

De Maio, A., Fornaro, G., Pauciullo, A., 2009a. Detection of single scatterers in multidimensional SAR imaging. IEEE Transactions on Geoscience and Remote Sensing 47 (7): 2284–2297.

De Maio, A., Fornaro, G., Pauciullo, A., Reale, D., 2009b. Detection of double scatterers in SAR Tomography. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 3, III–172.

Donoho, D., 2006. Compressed sensing. IEEE Transactions on Information Theory vol.52 (4): pp.1289–1306.

Donoho, D., Stodden, V., Tsaig, Y., 2005. Sparselab. Software: http://sparselab. stanford. edu.

Donoho, D., Stodden, V., Tsaig, Y., 2009. Sparselab Architecture. Software: http://sparselab. stanford. edu .

Duarte, M., Davenport, M., Takhar, D., Laska, J., Sun, T., Kelly, K., Baraniuk, R., 2008. Single-pixel imaging via compressive sampling. IEEE Signal Processing Magazine 25 (2): 83–91.

Eineder, M., Adam, N., Bamler, R., Yague-Martinez, N., Breit, H., 2009. Spaceborne spotlight SAR interferometry with TerraSAR-X. IEEE Transactions on Geoscience and Remote Sensing 47 (5): 1524–1535.

Elachi, C., 1988. Spaceborne radar remote sensing: applications and techniques. New York, IEEE Press, 1988, 285 p. 1.

Elad, M., Aharon, M., 2006. Image denoising via sparse and redundant representations over learned dictionaries. IEEE Transactions on Image Processing 15 (12): 3736–3745.

Elad, M., Starck, J., Querre, P., Donoho, D., 2005. Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). Applied and Computational Harmonic Analysis 19 (3): 340–358.

Ender, J., 2010. On compressive sensing applied to radar. Signal Processing: Special Section on Statistical Signal & Array Processing 90: 1402–1414.

Fadili, M., Starck, J., Murtagh, F., 2009. Inpainting and zooming using sparse representations.

The Computer Journal 52 (1): 64.

Farhat, N., Werner, C., Chu, T., 1984. Prospects for three-dimensional projective and tomographic imaging radar networks. Radio Science 19 (5): 1347–1355.

Ferretti, A., Bianchi, M., Prati, C., Rocca, F., 2005. Higher-order permanent scatterers analysis. EURASIP Journal on Applied Signal Processing 2005 (20): 3231–3242.

Ferretti, A., Prati, C., Rocca, F., 2000. Analysis of permanent scatterers in SAR interferometry. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 2, 761–763.

Ferretti, A., Prati, C., Rocca, F., et al., 2001. Permanent scatterers in SAR interferometry. IEEE Transactions on Geoscience and Remote Sensing 39 (1): 8–20.

Fornaro, G., Lombardini, F., Serafino, F., 2005. Three-dimensional multipass SAR focusing: Experiments with long-term spaceborne data. IEEE Transactions on Geoscience and Remote Sensing 43 (4): 702–714.

Fornaro, G., Reale, D., Serafino, F., 2009. Four-Dimensional SAR Imaging for Height Estimation and Monitoring of Single and Double Scatterers. IEEE Transactions on Geoscience and Remote Sensing 47 (1 Part 2): 224–237.

Fornaro, G., Serafino, F., 2006. Imaging of Single and Double Scatterers in Urban Areas via SAR Tomography. IEEE Transactions on Geoscience and Remote Sensing 44 (12): 3497–3505.

Fornaro, G., Serafino, F., Reale, D., 2010. 4-D SAR Imaging: The Case Study of Rome. IEEE Geoscience and Remote Sensing Letters 7 (2): 236 –240.

Fornaro, G., Serafino, F., Soldovieri, F., 2003. Three-dimensional focusing with multipass SAR data. IEEE Transactions on Geoscience and Remote Sensing 41 (3): 507–517.

Franceschetti, G., Lanari, R., 1999. Synthetic aperture radar processing. CRC.

Frey, O., Morsdorf, F., Meier, E., 2008. Tomographic imaging of a forested area by airborne multi-baseline P-band SAR. Sensors 8 (9): 5884–5896.

Gabriel, A., Goldstein, R., 1988. Crossed orbit interferometry: theory and experimental results from SIR-B. International Journal of Remote Sensing 9 (5): 857–872.

Gan, L., Ling, C., Do, T., Tran, T., 2009. Analysis of the statistical restricted isometry property for deterministic sensing matrices using Stein's method. Preprint .

Gatelli, F., Guamieri, A., Parizzi, F., Pasquali, P., Prati, C., Rocca, F., 1994. The wavenumber shift in SAR interferometry. IEEE Transactions on Geoscience and Remote Sensing 32 (4): 855–865.

Gernhardt, S., Adam, N., Eineder, M., Bamler, R., 2010. Potential of very high resolution SAR for persistent scatterer interferometry in urban areas. Annals of GIS 16 (2): 103–111.

Gini, F., Lombardini, F., 2002. Multilook APES for multibaseline SAR interferometry. IEEE Transactions on Signal Processing 50 (7): 1800–1803.

Gini, F., Lombardini, F., 2005. Multibaseline cross-track SAR interferometry: a signal processing perspective. IEEE Aerospace and Electronic Systems Magazine 20 (8): 71–93.

Goldstein, R., Werner, C., 1998. Radar interferogram filtering for geophysical applications. Geophysical Research Letters 25 (21): 4035–4038.

Grant, M., Boyd, S., 2008. Graph implementations for nonsmooth convex programs. In: Blondel, V., Boyd, S., Kimura, H. (Hrsg.), Recent Advances in Learning and Control. Lecture Notes in Control and Information Sciences. Springer-Verlag Limited, 95–110.

Grant, M., Boyd, S., 2011. CVX: Matlab software for disciplined convex programming, version 1.21. http://cvxr.com/cvx.

Guillaso, S., Ferro-Famil, L., Reigber, A., Pottier, E., 2003. Analysis of built-up areas from polarimetric interferometric SAR images. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 3, 1727 – 1729.

Guillaso, S., Reigber, A., 2005. Scatterer characterisation using polarimetric SAR tomography. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 4, 2685.

Hale, E., Yin, W., Zhang, Y., 2007. A fixed-point continuation method for L1-regularized min-

imization with applications to compressed sensing. Rice University Department of Computational and Applied Mathematics, Tech. Rep .

Hansen, M., Yu, B., 2001. Model selection and the principle of minimum description length. Journal of the American Statistical Association 96 (454): 746–774.

Hansen, P., 1987. The truncated SVD as a method for regularization. BIT Numerical Mathematics 27 (4): 534–553.

Hansen, P., 1994. Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems. Numerical Algorithms 6 (1): 1–35.

Hanssen, R., 2001. Radar interferometry: data interpretation and error analysis. Kluwer Academic Pub.

Herman, M., Strohmer, T., 2009. High-resolution radar via compressed sensing. IEEE Transactions on Signal Processing 57 (6): 2275–2284.

Homer, J., Longstaff, I., Callaghan, G., 1996. High resolution 3-D SAR via multi-baseline interferometry. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 1, 796–798.

Hooper, A., Zebker, H., Segall, P., Kampes, B., 2004. A new method for measuring deformation on volcanoes and other natural terrains using InSAR persistent scatterers. Geophysical Research Letters 31 (23): 5.

Jakowatz, C., 1996. Spotlight-mode synthetic aperture radar: a signal processing approach. Kluwer Academic Pub.

Jakowatz, C., Thompson, P., 1995. A new look at spotlight mode synthetic aperture radar as tomography: imaging 3-D targets. IEEE Transactions on Image Processing 4 (5): 699–703.

Kampes, B., 2006. Radar interferometry: persistent scatterer technique. Springer.

Kampes, B., Adam, N., 2005. The STUN Algorithm for Persistent Scatterer Interferometry. In: Proceedings of FRINGE.

Koh, K., Kim, S., Boyd, S., 2007. An interior-point method for large-scale l1-regularized logistic regression. Journal of Machine learning research 8 (8): 1519–1555.

Li, F., Goldstein, R., 2002. Studies of multibaseline spaceborne interferometric synthetic aperture radars. IEEE Transactions on Geoscience and Remote Sensing 28 (1): 88–97.

Li, J., Stoica, P., 1996. Efficient mixed-spectrum estimation with applications to target feature extraction. IEEE Transactions on Signal Processing 44 (2): 281–295.

Li, J., Stoica, P., 2008. MIMO radar signal processing. J. Wiley & Sons.

Lombardini, F., 2005. Differential tomography: a new framework for SAR interferometry. IEEE Transactions on Geoscience and Remote Sensing 43 (1): 37–44.

Lombardini, F., Gini, F., Matteucci, P., 2001. Application of array processing techniques to multibaseline InSAR for layover solution. In: Proceedings of the 2001 IEEE Radar Conference, 210–215.

Lombardini, F., Montanari, M., Gini, F., 2003. Reflectivity estimation for multibaseline interferometric radar imaging of layover extended sources. IEEE Transactions on Signal Processing 51 (6): 1508–1519.

Lombardini, F., Pardini, M., Gini, F., 2007. Sector interpolation for 3D SAR imaging with baseline diversity data. In: Proceedings of International Waveform Diversity and Design Conference, 297 –301.

Lombardini, F., Reigber, A., 2003. Adaptive spectral estimation for multibaseline SAR tomography with airborne L-band data. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 3, 2014–2016.

Lustig, M., Donoho, D., Pauly, J., 2007. Sparse MRI: The application of compressed sensing for rapid MR imaging. Magnetic Resonance in Medicine 58 (6): 1182–1195.

Lustig, M., Donoho, D., Santos, J., Pauly, J., 2008. Compressed sensing MRI. IEEE Signal Processing Magazine 25 (2): 72–82.

Malgouyres, F., Zeng, T., 2009. A predual proximal point algorithm solving a non negative basis pursuit denoising model. International journal of computer vision 83 (3): 294–311.

Massonnet, D., Rabaute, T., 2002. Radar interferometry: limits and potential. IEEE Transactions on Geoscience and Remote Sensing 31 (2): 455–464.

Massonnet, D., Rossi, M., Carmona, C., Adragna, F., Peltzer, G., Feigl, K., Rabaute, T., 1993. The displacement field of the Landers earthquake mapped by radar interferometry. Nature 364 (6433): 138–142.

Massonnet, D., Souyris, J., 2008. Imaging with synthetic aperture radar. EFPL Press.

Meyer, F., Gernhardt, S., Adam, N., 2007. Long-Term and Seasonal Subsidence Rates in Urban Areas from Persistent Scatterer Interferometry. In: Proceedings of URBAN.

Nannini, M., Scheiber, R., 2007. Height dependent motion compensation and coregistration for airborne sar tomography. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 5041–5044.

Nannini, M., Scheiber, R., Horn, R., 2008. Imaging of targets beneath foliage with SAR tomography. In: Proceedings of EUSAR, VDE VERLAG GmbH.

Nannini, M., Scheiber, R., Moreira, A., 2009. Estimation of the minimum number of tracks for SAR tomography. IEEE Transactions on Geoscience and Remote Sensing 47 (2): 531–543.

Pasquali, P., Prati, C., Rocca, F., Seymour, M., Fortuny, J., Ohlmer, E., Sieber, A., 1995. A 3-D SAR experiment with EMSL data. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 1, 784–786 vol.1.

Perissin, D., Ferretti, A., 2007. Urban-target recognition by means of repeated spaceborne SAR images. IEEE Transactions on Geoscience and Remote Sensing 45 (12): 4043–4058.

Piau, P., 1994. Performances of the 3D-SAR imagery. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 4, 2267–2271.

Reale, D., Fornaro, G., Pauciullo, A., Zhu, X., Bamler, R., 2011. Tomographic Imaging and Monitoring of Buildings With Very High Resolution SAR Data. IEEE Geoscience and Remote Sensing Letters 8 (4): 661–665.

Reigber, A., 2002. Airborne polarimetric sar tomography. Deutsches Zentrum für Luft- und Raumfahrt. Forschungsberichte .

Reigber, A., Moreira, A., 2000. First demonstration of airborne SAR tomography using multibaseline L-band data. IEEE Transactions on Geoscience and Remote Sensing 38 (5): 2142–2152.

Reigber, A., Moreira, A., Papathanassiou, K., Cloude, S., 2000. SAR tomography and interferometry for the remote sensing of forested terrain. In: Proceedings of EUSAR, 137–140.

Rissanen, J., 1978. Modeling by shortest data description. Automatica 14 (5): 465–471.

Rocca, F., 2007. Modeling interferogram stacks. IEEE Transactions on Geoscience and Remote Sensing 45 (10): 3289–3299.

Ruf, C., Swift, C., Tanner, A., Le Vine, D., 2002. Interferometric synthetic aperture microwave radiometry for the remote sensing of the earth. IEEE Transactions on Geoscience and Remote Sensing 26 (5): 597–611.

Sauer, S., 2008. Interferometric sar remote sensing of urban areas at l-band using multibaseline and polarimetric spectral analysis techniques. Deutsches Zentrum für Luft- und Raumfahrt. Forschungsberichte .

Sauer, S., Ferro-Famil, L., Reigber, A., Pottier, E., 2009. Polarimetric dual-baseline InSAR building height estimation at L-band. IEEE Geoscience and Remote Sensing Letters 6 (3): 408–412.

Schmidt, R., 1979. Multiple emitter location and signal parameter estimation. IEEE Transactions on Antennas and Propagation 34 (3): 276–280.

Schwarz, G., 1978. Estimating the dimension of a model. The annals of statistics 6 (2): 461–464.

Seeger, M., Wipf, D., 2010. Variational Bayesian Inference Techniques. IEEE Signal Processing Magazine 27 (6): 81 –91.

Shastry, M., Narayanan, R., Rangaswamy, M., 2010. Compressive radar imaging using white stochastic waveforms. In: Proceedings of International Waveform Diversity and Design Conference (WDD), 90–94.

She, Z., Gray, D., Bogner, R., Homer, J., 1999. Three-dimensional SAR imaging via multiple pass processing. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 5, 2389 –2391 vol.5.

Shen, B., Hu, W., Zhang, Y., Zhang, Y., 2009. Image inpainting via sparse representation. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE Computer Society, Washington, DC, USA, 697–700.

Smith, G., Diethe, T., Hussain, Z., Shawe-Taylor, J., Hardoon, D., 2010. Compressed sampling for pulse doppler radar. In: Proceedings of IEEE Radar Conference, 887 –892.

Soergel, U., 2010. Review of Radar Remote Sensing on Urban Areas. Radar Remote Sensing of Urban Areas : 1–47.

Soergel, U., Thoennessen, U., Brenner, A., Stilla, U., 2006. High-resolution SAR data: new opportunities and challenges for the analysis of urban areas. IEE Proceedings-Radar, Sonar and Navigation 153: 294.

Starck, J., Murtagh, F., Fadili, J., 2010. Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity. Cambridge Univ Pr.

Stoica, P., Arye, N., 1989. MUSIC, maximum likelihood, and Cramer-Rao bound. IEEE Transactions on Acoustics, Speech and Signal Processing 37 (5): 720–741.

Stoica, P., Cedervall, M., 2002. On LP-MUSIC. IEEE Transactions on Signal Processing 43 (2): 552–555.

Stoica, P., Moses, R., 1990. On biased estimators and the unbiased Cramér-Rao lower bound. Signal Processing 21 (4): 349–350.

Stoica, P., Moses, R., 1997. Introduction to spectral analysis. Vol. 57. Prentice Hall Upper Saddle River, New Jersey.

Stoica, P., Moses, R., 2005. Spectral analysis of signals. Pearson Prentice Hall.

Swingler, D., 1993. Frequency estimation for closely spaced sinsoids: Simple approximations to the Cramér-Rao lower bound. IEEE Transactions on Signal Processing 41 (1): 489.

Tebaldini, S., 2010. Single and multipolarimetric SAR tomography of forested areas: A parametric approach. IEEE Transactions on Geoscience and Remote Sensing 48 (5): 2375–2387.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological) 58 (1): 267–288.

Van Cittert, P., 1934. Die wahrscheinliche Schwingungsverteilung in einer von einer Lichtquelle direkt oder mittels einer Linse beleuchteten Ebene. Physica 1 (1-6): 201–210.

Varshney, K., Çetin, M., Fisher, J., Willsky, A., 2008. Sparse representation in structured dictionaries with application to synthetic aperture radar. IEEE Transactions on Signal Processing 56 (8): 3548–3561.

Vasco, D., Ferretti, A., Novali, F., 2008. Estimating permeability from quasi-static deformation: temporal variations and arrival-time inversion. Geophysics 73: O37.

Wang, X., 2011. Compressive Sensing for Pan-sharpening. Master's thesis.

Wen, Z., Yin, W., Goldfarb, D., Zhang, Y., 2010. A fast algorithm for sparse reconstruction based on shrinkage, subspace optimization and continuation. SIAM Journal on Scientific Computing 32 (4): 1832–1857.

Werninghaus, R., Buckreuss, S., 2010. The TerraSAR-X mission and system design. IEEE Transactions on Geoscience and Remote Sensing 48 (2): 606–614.

Willett, R., Gehm, M., Brady, D., 2007. Multiscale reconstruction for computational spectral imaging. Computational Imaging V .

Xie, X., Zhang, Y., 2010. High-resolution imaging of moving train by ground-based radar with compressive sensing. Electronics letters 46 (7): 529–531.

Yang, J., Wright, J., Huang, T., Ma, Y., 2008. Image super-resolution as sparse representation of raw image patches .

Yang, J., Wright, J., Huang, T., Ma, Y., 2010. Image super-resolution via sparse representation. IEEE Transactions on Image Processing 19 (11): 2861–2873.

Yang, J., Zhang, Y., 2009. Alternating Direction Algorithms for l1-Problems in Compressive

Sensing. Arxiv preprint arXiv:0912.1185 .

Yu, H., Wang, G., 2009. Compressed sensing based interior tomography. Physics in medicine and biology 54: 2791.

Zebker, H., Goldstein, R., 1986. Topographic mapping from interferometric synthetic aperture radar observations. Journal of Geophysical Research 91 (B5): 4993–4999.

Zebker, H., Rosen, P., 1994. On the derivation of coseismic displacement fields using differential radar interferometry: The Landers earthquake. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 1, 286–288.

Zhang, L., Xing, M., Qiu, C., Li, J., 2009. Achieving Higher Resolution ISAR Imaging With Limited Pulses via Compressed Sampling. IEEE Geoscience and Remote Sensing Letters 6: 567–571.

Zhu, X., 2008. High-Resolution Spaceborne Radar Tomography. Master's thesis.

Zhu, X., Adam, N., Bamler, R., 2008. First Demonstration of Space-borne High Resolution SAR Tomography in Urban Environment Using TerraSAR-X Data. In: Proceedings of CEOS SAR Workshop on Calibration and Validation.

Zhu, X., Adam, N., Bamler, R., 2010. Space-borne high resolution tomographic interferometry. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 4.

Zhu, X., Adam, N. B. R., Bamler, R., 2009. Space-borne High Resolution SAR Tomography: Experiments in Urban Environment Using TerraSAR-X Data. In: Proceedings of Joint Urban Remote Sensing Event (JURSE).

Zhu, X., Bamler, R., 2009. Very high Resolution SAR tomography via Compressive Sensing. In: Proceedings of FRINGE.

Zhu, X., Bamler, R., 2010a. Compressive sensing for high resolution differential SAR tomography-the SL1MMER algorithm. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 17–20.

Zhu, X., Bamler, R., 2010b. Super-resolution for 4-D SAR Tomography via Compressive Sensing. In: Proceedings of EUSAR, VDE VERLAG GmbH.

Zhu, X., Bamler, R., 2010c. Tomographic SAR Inversion by L_1-Norm Regularization—The Compressive Sensing Approach. IEEE Transactions on Geoscience and Remote Sensing 48 (10): 3839–3846.

Zhu, X., Bamler, R., 2010d. Very High Resolution Spaceborne SAR Tomography in Urban Environment. IEEE Transactions on Geoscience and Remote Sensing 48 (12): 4296–4308.

Zhu, X., Bamler, R., 2011a. Let's Do the Time Warp: Multi-Component Nonlinear Motion Estimation in Differential SAR Tomography. IEEE Geoscience and Remote Sensing Letters 8 (4): 735–739.

Zhu, X., Bamler, R., 2011b. Super-Resolution Power and Robustness of Compressive Sensing for Spectral Estimation with Application to Spaceborne Tomographic SAR. IEEE Transactions on Geoscience and Remote Sensing ,in press.

Zhu, X. Wang, X. B. R., 2011. Compressive sensing for image fusion - with application to pan-sharpening. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS).

List of Figures

List of Tables