The $\mathbb{Z}\pi M$ Algorithm: A Method for Interferometric Image Reconstruction in SAR/SAS

José M. B. Dias, *Member, IEEE,* and José M. N. Leitão, *Member, IEEE,*

***Abstract*—This paper presents an effective algorithm for absolute phase (not simply modulo-2π) estimation from incomplete, noisy and modulo-2π observations in interferometric aperture radar and sonar (InSAR/InSAS). The adopted framework is also representative of other applications such as optical interferometry, magnetic resonance imaging and diffraction tomography. The Bayesian viewpoint is adopted; the observation density is 2π-periodic and accounts for the interferometric pair decorrelation and system noise; the *a priori* probability of the absolute phase is modeled by a *compound Gauss–Markov random field* (CGMRF) tailored to piecewise smooth absolute phase images. We propose an iterative scheme for the computation of the *maximum a posteriori probability* (MAP) absolute phase estimate. Each iteration embodies a discrete optimization step ($\mathbb{Z}$-step), implemented by network programming techniques and an *iterative conditional modes* (ICM) step (π-step). Accordingly, the algorithm is termed $\mathbb{Z}\pi M$, where the letter M stands for maximization. An important contribution of the paper is the simultaneous implementation of phase unwrapping (inference of the 2π-multiples) and smoothing (denoising of the observations). This improves considerably the accuracy of the absolute phase estimates compared to methods in which the data is low-pass filtered prior to unwrapping. A set of experimental results, comparing the proposed algorithm with alternative methods, illustrates the effectiveness of our approach.**

***Index Terms*—Bayesian estimation, compound Gauss–Markov random, interferometry, iterative conditioonal modes (ICM), network programming, phase unwrapping.**

I. Introduction

THE need for estimating phase[1] from incomplete, noisy and modulo-2π observations appears in many classes of imaging techniques. Some relevant examples are as follows.

1) *Interferometric synthetic aperture radar* (InSAR) [1] and *interferometric synthetic aperture sonar* (InSAS) [2]. Synthetic aperture radar (SAR) and synthetic aperture sonar (SAS) systems produce high resolution images of the coherent fields backscatterd by the surface being illuminated. SAR and SAS images are typically acquired by a single antenna. By using two antennas (actually two antennas in single-pass mode or one antenna in repeat-pass mode) separated by a baseline, it is possible to *interfere* the two images in such a way that the common scene reflectivity is cancelled out and the geometric information contained in the scene topography is retained in the phase difference.
2) *Magnetic resonance imaging* systems, where the phase estimation is a necessary tool for determining the magnetic field deviation maps. These deviation maps are then used to correct geometric distortions in echo-planar images [3]. Other applications areas are water and fat separation [4] and dynamic range improvement of phase contrast measurements.
3) *Optical interferometry*, where phase differences are used to obtain information such as shape, displacement, or vibration of a surface [5].
4) *Diffraction tomography* (e.g., geophysical tomography or acoustic tomography), where the Rythov approximation yields a mapping between the observed object and the phase of the measured field [6].

In all these applications, the observed data relates to the phase in a nonlinear and noisy way; the nonlinearity, closely related with the wave propagation phenomena involved in the acquisition process, is sinusoidal; the noise is introduced both by the acquisition mechanism and by the electronic equipment. Therefore, the phase should be inferred from noisy modulo-2π observations, (the so-called *principal phase values* or *interferogram*).

The mainstream of phase estimation research in InSAR/InSAS usually takes a two step approach: in the first step a filtered interferogram is obtained from noisy InSAR/InSAS image pairs; in the second step the phase is unwrapped by determining 2π-multiples consistent with the filtered interferogram (*unwrapping* in the interferometric jargon). The book [7] and the algorithms therein presented and compared are representative of this approach. Throughout this paper we use the term *unwrapping* to designate the latter step.

Broadly speaking, phase estimation methods can be classified into four major classes: path following methods, minimum-norm methods, Bayesian and regularization methods and methods based on parametric models. Thesis [8] and paper [9] provide a comprehensive account of the mentioned methods. We stress that while the methods from the two first classes implement only the phase unwrapping step mentioned in the paragraph above, some Bayesian and regularization methods might implement both smoothing and unwrapping steps. Therefore, some care must me taken in comparing phase unwrapping methods with Bayesian and regularization methods.

Manuscript received March 30, 2001; revised January 8, 2002. This work was supported by the Fundação para a Ciência e Tecnologia under Projects POSI/34071/CPS/2000 and 2/2.1.TIT/1580/95. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Mark A. Richards.

J. Dias is with the Instituto Superior Técnico, Instituto de Telecomunicações, 1049-001 Lisboa, Portugal (e-mail: bioucas@lx.it.pt; http://maxwell.lx.it.pt/~dias/).

J. Leitão is with the Instituto Superior Técnico, Instituto de Telecomunicações, 1049-001 Lisboa, Portugal (e-mail: jleitao@lx.it.pt).

Publisher Item Identifier S 1057-7149(02)02858-0.

[1]In this paper, the word *phase* means absolute phase (not simply modulo-2π).

In the path following schemes [7], [9], [10] phase is unwrapped along selected image paths. In the presence of discontinuities, noise, or undersampling, different paths between two points may lead to different phase values. To resolve or mitigate theses inconsistencies, heuristic rules are applied to provide path-independent integration.

Minimum-norm phase unwrapping methods cast the unwrapping problem as the minimization of an L^p norm [7], [11]. L^2 norm (least-squares) has long been used [12], [13]; the least-squares solution can be computed efficiently by using fast cosine or Fourier transforms [14]. Works [15] and [16] have proposed (independently) network programming based algorithms that minimize the L^1 norm; this criterion is able to preserve sharp transitions without modeling them explicitly [17]. This ability is further enhanced by using L^p norms with $0 \leq p < 1$. However, these norms lead to hard nonconvex optimization problems, with unbearable computational load. A suboptimal L^0 solution is proposed in [7, ch. 5].

Due to decorrelation (temporal and spatial), no-return or low return areas (e.g., due to layover phenomena in InSAR/InSAS), the modulo-2π phase estimates corresponding to those areas might be extremely biased and/or noisy. To handle this problem, both the path following and the minimum-norm procedures have incorporated quality maps as a measure of confidence on the observed data at each site (see, e.g., [18] for path following, [19] for weighted least-squares and [15] for weighted L^1 norm).

In a quite different vein and recognizing that phase estimation is an ill-posed problem, papers [20]–[23] have adopted the regularization framework to impose smoothness on the solution. The same objective has been pursued in papers [24]–[29] by adopting a Bayesian viewpoint. Papers [24] and [25] propose a nonlinear recursive filtering solution to the phase reconstruction. Paper [26] considers the InSAR observation model taking into account not only the image phase, but also the *backscattering coefficient* and the *correlation factor* images, which are jointly recovered from InSAR image pairs. Paper [27] proposes a fractal-based prior and a simulated annealing scheme to compute the phase image. Works [28]–[30], although proposing a phase unwrapping approach to phase estimation, can be classified as Bayesian, since the differences between neighboring 2π-multiples of the phase are modeled as random variables; the phase is unwrapped using mean field inference in [29], probability propagation in graphical models in [28] and network-flow techniques [31] to approximate the MAP solution in [30].

Methods based on parametric models constrain the phase to belong to a given parametric model. Works [32] and [33] have adopted low-order polynomials. These approaches yield good results if the low-order polynomials represent accurately the phase. However, in practical applications the entire phase function cannot be approximated by a single two-dimensional (2-D) polynomial model. To circumvent model mismatches, work [32] proposes a partition of the observed field where each partition element has its own parametric model.

In one way or another, most phase estimation algorithms assume that the phase difference between two neighboring sites varies smoothly (less than π in a deterministic or stochastic sense, depending on the paradigm); based on this assumption, it is possible, by exploiting the neighboring observed phases, to infer the 2π-multiple component of the phase of a given site. However, in situations such as undersampling (e.g., topography inducing high phase rates), abrupt features/objects, or the *layover* phenomenon), the smoothness assumption can not be made. In this case, the principal phase values are inconsistent in the sense that they do not uniquely determine the phase. Discontinuities or inconsistencies may also appear as a consequence of the smoothness step applied by most phase unwrapping algorithms; typically, this step assumes that phase is practically constant within small windows, which is not true for high phase rate regions. Independently of their origin, phase discontinuities/inconsistencies are the principal source of error in any phase estimation algorithm that does not take them into account. To further complicate the problem, the discontinuity field can not be uniquely determined from the observed data, even in the absence of noise, due to the periodic structure of the observation mechanism.

A. Proposed Approach

We adopt the Bayesian viewpoint. The likelihood function, which models the observation mechanism, is 2π-periodic and accounts for the interferometric pair decorrelation and the system noise. The *a priori* probability of the phase is modeled by a first-order compound Gauss–Markov random field (CGMRF) [34] tailored to piecewise smooth phase fields.

Due to the periodic structure of the likelihood function, the discontinuity field in InSAR/InSAS applications can not be uniquely determined from the observed data. However, it can be inferred using information external to the phase estimation framework. The major source of discontinuities in InSAR/InSAS applications is the layover phenomenon (see, e.g., [7, ch. 3]). The layover areas can be separated from the nonlayover ones as proposed in [35]. This work exploits spectral shift that exists between the signal read by the two sensors as function of the along-range local slope. An alternative approach to handle discontinuities/inconsistencies is to segment the observed data into a phase-consistent region and its complement and use only the observed data in the former region to estimate the whole phase. Our approach accepts both: a discontinuity field and a region of consistency. Detailed procedures aiming at discontinuity detection or data segmentation are, however, beyond the scope of this paper.

In papers [24]–[26], following a Bayesian approach, the prior is a first-order causal GMRF. Taking advantage of this prior and using the *reduced-order model* (ROM) approximation of the GMRF [36], the phase is estimated with a nonlinear recursive filtering technique. Compared with the present approach, the main differences concern the prior and the estimation algorithm: we use a first-order noncausal CGMRF prior. In terms of estimation, the noncausal prior implies a batch perspective, where the phase estimate at each site is based on the complete observed image. This is in contrast with the filtering technique implemented in the referenced papers, where the phase estimate of a given site is inferred only from past (in the lexicographic sense) observed data.

To compute the MAP estimate, we derive an iterative procedure with two steps per iteration: the first step, termed the

$\mathbb{Z}$-step, maximizes the posterior density with respect to the field of 2π phase multiples; the second step, termed the π-step, maximizes the posterior density with respect to the phase principal values. The $\mathbb{Z}$-step is a discrete optimization problem solved by network programming techniques inspired by Flyn's minimum weighted discontinuity algorithm [15]. The π-step is a continuous optimization problem solved approximately by the *iterated conditional modes* (ICM) [37] scheme. Accordingly, we term our algorithm $\mathbb{Z}\pi M$, where the letter M stands for maximization.

As previously mentioned, the smoothing step that many phase estimation schemes apply prior to unwrapping jeopardizes the phase unwrapping step in areas of high phase rate. This problem is minimized in the proposed methodology, as it does not split the phase estimation into independent smoothing and unwrapping steps; these steps are instead implemented simultaneously and implicitly in computing the MAP estimate. The accuracy of the proposed scheme is in this way considerably improved compared to the phase unwrapping approaches.

This paper is organized as follows. Section II introduces the observation model, the CGMRF prior and the posteriori density. Section III elaborates on the estimation procedure; namely, we derive solutions for the $\mathbb{Z}$-step and for the π-step. Section IV presents results based on synthetic data and on data generated by a simulator fed with real elevation measurements.

II. Proposed Models

A. Observation Model

Fig. 1 shows a typical SAR/SAS geometry. The trajectories of sensors s_1 and s_2 are parallel and separated by the baseline B. The height h of a given terrain element is a function $g(\phi)$ of the phase $\phi = \phi_1 - \phi_2$, where ϕ_1 and ϕ_2 are the propagation path phases associated to sensor 1 and sensor 2, respectively. Phase ϕ is to be inferred from x_1 and x_2, the complex amplitudes of the backscattered field read by each sensor from a given site. These amplitudes are given by

$$x_1 = z_1 e^{-j\phi_1} + n_1 \tag{1}$$

$$x_2 = z_2 e^{-j\phi_2} + n_2 \tag{2}$$

where z_1 and z_2 are the complex amplitudes originated by the scatterers illuminated by apertures 1 and 2, respectively, and n_1 and n_2 are the electronic noise of sensor 1 and sensor 2, respectively.

Assuming that the surface being illuminated is rough compared to the wavelength, that there are no strong specular reflectors and that there are a *large* number of scatterers per resolution cell, then the complex amplitude z_1 is complex zero-mean circular Gaussian distributed [38, ch. 5]. Noises n_1 and n_2 are also independent (corresponding to different sensors) and complex zero-mean circular Gaussian distributed. Furthermore, we assume that n_1 and n_2 are independent of z_1 and z_2.

Complex amplitudes z_1 and z_2 are different due to the following reasons:

1) spatial decorrelation originated by image misregistration;
2) temporal decorrelation originated by scatterer displacements (only in repeated-pass mode);

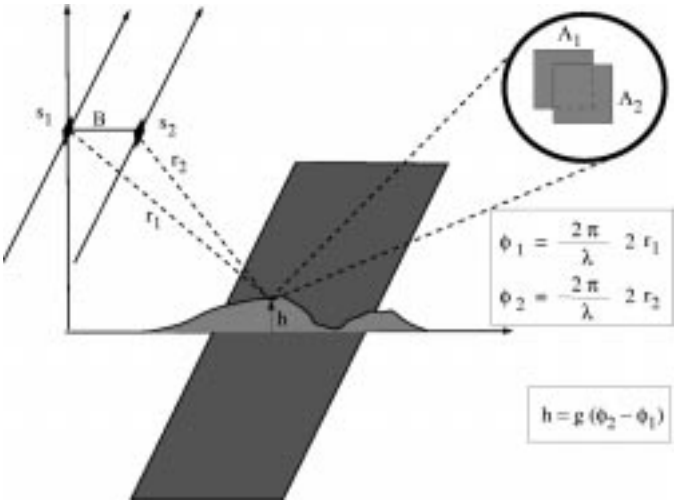

Fig. 1. Typical interferometric SAR/SAS geometry. The trajectories of sensors s_1 and s_2 are parallel and separated by the baseline B. The height h of a given terrain element is a known function of the phase $\phi = \phi_1 - \phi_2$, i.e., $h = g(\phi)$.

3) focusing errors originated by the imaging algorithm or by platform displacements with respect to the nominal trajectory;
4) geometric decorrelation originated by the different geometries of each sensor.

We assume that $\theta^2 \equiv E[|z_1|^2] = E[|z_2|^2]$ and that the correlation factor $\alpha \equiv (E[z_1 z_2^*])/\theta^2$ (also termed the *change parameter* or *degree of coherence* [38, ch. 5]) is real. A sufficient condition for α to be real is that the difference between the phase induced by each scatterer in z_1 and z_2 has an even density.

Defining $x \equiv [x_1\ x_2]^T$, $\sigma_n^2 \equiv E[|n_1|^2]$ and assuming that $E[|n_1|^2] = E[|n_2|^2]$, the probability density function[2] of x is [38, ch. 5]

$$p_{x|\phi}(x|\phi) = \frac{1}{\pi^2 |\mathbf{Q}|} e^{-x^H \mathbf{Q}^{-1} x} \tag{3}$$

where $\mathbf{Q} \equiv E[xx^H]$ is given by

$$\mathbf{Q} = \begin{bmatrix} \theta^2 + \sigma_n^2 & \alpha\theta^2 e^{j\phi} \\ \alpha\theta^2 e^{-j\phi} & \theta^2 + \sigma_n^2 \end{bmatrix}. \tag{4}$$

Developing the quadratic form in (3), one is led to

$$p_{x|\phi}(x|\phi) = c e^{\lambda \cos(\phi - \eta)} \tag{5}$$

where $c = c(x, \theta, \alpha)$ and

$$\eta = \arg(x_1 x_2^*) \tag{6}$$

$$\lambda = \frac{2\alpha\theta^2 |x_1 x_2|}{|\mathbf{Q}|}. \tag{7}$$

The likelihood function $p_{x|\phi}(x|\phi)$ is 2π-periodic with respect to ϕ with maxima at $\phi = 2\pi k + \eta$, for $k \in \mathbb{Z}$ ($\mathbb{Z}$ denotes the integer set). Parameter η is a maximum likelihood estimate of ϕ. The peakiness of (5) about $2\pi k + \eta$, controlled by the parameter λ, is an indication, in a statistical sense, of how trustworthy the data is.

[2]For compactness, lowercase letters will denote random variables and their values as well.

Often, the parameter λ is unknown and must be jointly estimated with the phase ϕ. Herein, when necessary, we take the maximum likelihood estimate $\widehat{\lambda}_{ML}$ (see [38, ch. 5]).

Let $\boldsymbol{\phi} \equiv \{\phi_{ij} | (i,j) \in Z\}$ and $\mathbf{x} \equiv \{x_{ij} | (i,j) \in Z_o\}$ denote the phase and complex amplitude associated with sites $Z \equiv \{(i,j) | \, i,j = 1,\ldots,N\}$ (we assume without lack of generality that the images are squared) and $Z_o \subset Z$, respectively. Assuming that the components of $\mathbf{x}$ are conditionally independent, then

$$p_{\mathbf{x}|\phi}(\mathbf{X}|\phi) = \prod_{ij \in Z_o} p_{x_{ij}|\phi_{ij}}(x_{ij}|\phi_{ij}). \tag{8}$$

The conditional independence assumption is valid if the resolution cells associated with any pair of pixels are disjoint. Usually this is a good approximation, since the *point spread function* of the imaging systems is only slightly larger than the corresponding inter-pixel distance [39].

As stated before, we assume that the observation set of sites Z_o is a subset of the phase sites Z. Sites $Z - Z_o$ are either not observed or belong to inconsistent phase regions.

B. Prior Model

Image $\boldsymbol{\phi}$ is assumed to be *piecewise smooth*, with abrupt variations between neighboring regions. These variations are due to undersampling in areas with high fringe rates, mainly due to the presence of *layover* phenomena and/or abrupt feature or objects. Whatever their origin, discontinuities of the phase ϕ are the principal source of error in any unwrapping algorithm that does not take them into account.

Gauss-Markov random fields [40] are both mathematically and computationally suitable for representing local interactions and particularly continuity between neighboring pixels. However, the continuity constraint must be discarded for those pixels in the neighborhood of discontinuities. For this purpose we take a *first-order* noncausal CGMRF [34] with density

$$p_{\phi|\mathbf{l}}(\boldsymbol{\phi}|\boldsymbol{l}) \propto \exp\left\{-\frac{\mu}{2}\sum_{ij \in Z_1} \left(\Delta\phi_{ij}^h\right)^2 \overline{v}_{ij} + \left(\Delta\phi_{ij}^v\right)^2 \overline{h}_{ij}\right\} \tag{9}$$

where $\mathbf{l} \equiv \{v_{ij}, h_{ij} | \, (i,j) \in Z\}$ is the so-called *line field* process, $\overline{v}_{ij} \equiv (1 - v_{ij})$, $\overline{h}_{ij} \equiv (1 - h_{ij})$, $\Delta\phi_{ij}^h \equiv (\phi_{ij} - \phi_{i,j-1})$, $\Delta\phi_{ij}^v \equiv (\phi_{ij} - \phi_{i-1,j})$, $Z_1 \equiv \{(i,j) | \, i,j = 2,\ldots,N\}$ and μ^{-1} means the variance of increments $\Delta\phi_{ij}^h$ and $\Delta\phi_{ij}^v$. Variables $v_{ij}, h_{ij} \in \{0,1\}$ serve the purpose of signaling discontinuities. Notice that the continuity constraint between sites (i,j) and $(i,j-1)$ is removed if variable v_{ij} is set to one; similarly, the continuity constraint between sites (i,j) and $(i-1,j)$ is removed if variable v_{ij} is set to one. Fig. 2 shows the site (i,j) and its four first-order neighbors. A line field variable between each pair of neighboring sites represents a possible discontinuity.

Parameter μ of (9) controls the smoothness of the phase field: as μ gets large, the random phase fields generated by (9) become smoother. If μ is not *a priori* known it should be dealt with as a random variable and either estimated jointly with the phase surface or integrated out from the posterior distribution. In the former case and assuming a uniform prior on μ, the estimate of μ, given $\boldsymbol{\phi}$, is

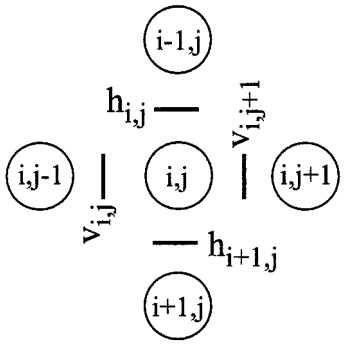

Fig. 2. Representation of the site (i,j) and its first-order neighbors along with the line field variables affecting each pair of neighboring sites.

$$\widehat{\mu} = \frac{\sum_{ij \in Z_1} \overline{v}_{ij} + \overline{h}_{ij}}{\sum_{ij \in Z_1} \left(\Delta\phi_{ij}^h\right)^2 \overline{v}_{ij} + \left(\Delta\phi_{ij}^v\right)^2 \overline{h}_{ij}} \tag{10}$$

which is also the maximum likelihood estimate of this parameter.

C. Posterior Density

Consider that the line field process $\mathbf{l}$ is known. Invoking the Bayes rule and noting that $p_{\mathbf{x}|\phi,\mathbf{l}}(\mathbf{x}|\boldsymbol{\phi},\mathbf{l}) = p_{\mathbf{x}|\boldsymbol{\phi}}(\mathbf{x}|\boldsymbol{\phi})$, we obtain the posterior probability density function of $\boldsymbol{\phi}$, given $(\mathbf{x},\mathbf{l})$, as

$$p_{\phi|\mathbf{x},\mathbf{l}}(\boldsymbol{\phi}|\mathbf{x},\mathbf{l}) \propto p_{\mathbf{x}|\boldsymbol{\phi}}(\mathbf{x}|\boldsymbol{\phi})\, p_{\phi|\mathbf{l}}(\boldsymbol{\phi}|\mathbf{l}) \tag{11}$$

where the factors not depending on $\boldsymbol{\phi}$ were discarded. Introducing (8) and (9) into (11), we obtain

$$p_{\boldsymbol{\phi}|\mathbf{x},\mathbf{l}}(\boldsymbol{\phi}|\mathbf{x},\mathbf{l}) \propto e^{\sum_{ij \in Z_o} \lambda_{ij} \cos(\phi_{ij} - \eta_{ij})} \times e^{-\frac{\mu}{2}\sum_{ij \in Z_1}\left(\left(\Delta\phi_{ij}^h\right)^2 \overline{v}_{ij} + \left(\Delta\phi_{ij}^v\right)^2 \overline{h}_{ij}\right)}. \tag{12}$$

The next section is devoted to maximization of the posterior density (12) with respect to the phase image $\boldsymbol{\phi}$. We stress that, contrary to the optimization schemes implemented by phase unwrapping algorithms, we do not constrain explicitly the modulo-2π phase derivatives to be irrotational [7,ch. 2]. Nevertheless, this constraint is indirectly enforced by the prior: if the referred rotational is not zero at some point this implies the presence of large phase differences $\Delta\phi_{ij}^h$ or $\Delta\phi_{ij}^v$ which are penalized by the prior.

III. Estimation Procedure

The MAP criterion is adopted for computing $\widehat{\boldsymbol{\phi}}$. Accordingly

$$\widehat{\boldsymbol{\phi}}_{MAP} = \arg\max_{\boldsymbol{\phi}} p_{\phi|\mathbf{x},\mathbf{l}}(\boldsymbol{\phi}|\mathbf{x},\mathbf{l}). \tag{13}$$

Due to the periodic structure of $p_{x|\phi}(x|\phi)$, computing the MAP solution leads to a huge nonconvex optimization problem, with unbearable computation burden. Instead of computing the exact estimate $\widehat{\boldsymbol{\phi}}_{MAP}$, we resort to a suboptimal scheme that delivers nearly optimal estimates, with a far lower computational load.

Initialization: $\widehat{\boldsymbol{\psi}}^{(0)} = \boldsymbol{\eta}$

for $t = 1, 2, \ldots,$

Unwrapping step:

$$\widehat{\mathbf{k}}^{(t)} = \arg\max_{\mathbf{k}} p_{\phi|\mathbf{x},\mathbf{l}}(\boldsymbol{\psi}^{(t-1)} + 2\pi\mathbf{k}|\mathbf{x},\mathbf{l})$$

Smoothing step:

$$\widehat{\boldsymbol{\psi}}^{(t)} = \arg\max_{\boldsymbol{\psi}} p_{\boldsymbol{\phi}|\mathbf{x},\mathbf{l}}(\boldsymbol{\psi} + 2\pi\mathbf{k}^{(t)}|\mathbf{x},\mathbf{l})$$

Termination test:

if (stop test == true)

break loop for

end for

Fig. 3. Pseudo-code for the $\mathbb{Z}\pi M$ algorithm.

Suppose initially that $Z_o = Z$, i.e., all sites are observed. Let the image ϕ_{ij} be uniquely decomposed as

$$\phi_{ij} = \psi_{ij} + 2\pi k_{ij} \tag{14}$$

where $k_{ij} \equiv \lfloor(\phi_{ij} + \pi)/(2\pi)\rfloor \in \mathbb{Z}$ ($\lfloor x\rfloor$ denotes the largest integer less than or equal to x) is the so-called wrap-count component of ϕ_{ij} and $\psi_{ij} \in [-\pi, \pi[$ is the principal value of ϕ_{ij}. The MAP estimate (13) can be rewritten in terms of $\boldsymbol{\psi} \equiv \{\psi_{ij} \,|\, (i,j) \in Z\}$ and $\mathbf{k} \equiv \{k_{ij} \,|\, (i,j) \in Z\}$ as

$$\left(\widehat{\boldsymbol{\psi}}_{MAP}, \widehat{\mathbf{k}}_{MAP}\right) = \arg\max_{\boldsymbol{\psi},\mathbf{k}} p_{\phi|\mathbf{x},\mathbf{l}}(\boldsymbol{\psi} + 2\pi\mathbf{k}|\mathbf{x},\mathbf{l}) \tag{15}$$

$$= \arg\left\{\max_{\boldsymbol{\psi}}\left\{\max_{\mathbf{k}} p_{\phi|\mathbf{x},\mathbf{l}}(\boldsymbol{\psi} + 2\pi\mathbf{k}|\mathbf{x},\mathbf{l})\right\}\right\}. \tag{16}$$

Instead of computing (16), we propose a procedure that successively and iteratively maximizes $p_{\phi|\mathbf{x},\mathbf{l}}(\boldsymbol{\psi} + 2\pi\mathbf{k}|\mathbf{x},\mathbf{l})$ with respect to $\mathbf{k} \in \mathbb{Z}^{N^2}$ and $\boldsymbol{\psi} \in [-\pi, \pi[^{N^2}$. We term this maximization on sets $\mathbb{Z}$ and $[-\pi, \pi[$ as the $\mathbb{Z}\pi M$ algorithm; Fig. 3 shows the corresponding high-level pseudo-code.

The $\mathbb{Z}\pi M$ algorithm is greedy since the posterior density $p_{\phi|\mathbf{x},\mathbf{l}}(\boldsymbol{\phi}|\mathbf{x},\mathbf{l})$ can not decrease in each step of each iteration. Thus, the stationary points of the unwrapping and smoothing steeps correspond to local maxima of $p_{\phi|\mathbf{x},\mathbf{l}}(\boldsymbol{\phi}|\mathbf{x},\mathbf{l})$. Nevertheless, the proposed method yields systematically good results, as we will document in next section.

The unwrapping step finds the maximum of the posterior density $p_{\phi|\mathbf{x},\mathbf{l}}(\boldsymbol{\phi}|\mathbf{x},\mathbf{l})$ on a mesh obtained by discretizing each coordinate ϕ_{ij} according to (14). The first estimate $\widehat{\mathbf{k}}^{(1)}$ delivered by the unwrapping step is based on the maximum likelihood estimate $\boldsymbol{\eta} \equiv \{\eta_{ij} \,|\, (i,j) \in Z\}$. Smoothing is next implemented. This is in contrast with the scheme followed by most phase unwrapping algorithms, where the phase is estimated from a smoothed version of $\boldsymbol{\eta}$ given, for example, by the ML estimate, under the assumption that the phase ϕ is constant within windows of given size. This assumption leads to large errors in areas of high phase rate.

A. $\mathbb{Z}$-Step

Since the logarithm is strictly increasing and $\cos(\psi_{ij} + 2\pi k_{ij} - \eta_{ij})$ does not depend on k_{ij}, solving the maximization step is equivalent to solving

$$\widehat{\mathbf{k}} = \arg\min_{\mathbf{k}} E(\mathbf{k}|\boldsymbol{\psi}) \tag{17}$$

where the energy $E(\mathbf{k}|\boldsymbol{\psi})$ is given by

$$E(\mathbf{k}|\boldsymbol{\psi}) \equiv \sum_{ij\in Z_1} \left(\Delta\phi^h_{ij}\right)^2 \bar{v}_{ij} + \left(\Delta\phi^v_{ij}\right)^2 \bar{h}_{ij} \tag{18}$$

with

$$\Delta\phi^h_{ij} = \left[2\pi\left(k_{ij} - k_{i,j-1}\right) - \Delta\psi^h_{ij}\right] \tag{19}$$

$$\Delta\phi^v_{ij} = \left[2\pi\left(k_{ij} - k_{i-1,j}\right) - \Delta\psi^v_{ij}\right] \tag{20}$$

and $\Delta\psi^h_{ij} \equiv \psi_{i,j-1} - \psi_{ij}$ and $\Delta\psi^v_{ij} \equiv \psi_{i-1,j} - \psi_{ij}$.

The energy $E(\mathbf{k}|\boldsymbol{\psi})$ is a sum of quadratic functions of $(k_{ij} - k_{i-1,j})$ and $(k_{ij} - k_{i,j-1})$. This is a special case of the so-called *nearest lattice vector problem*, which, for general positive definite quadratic forms of k_{ij}, is known to be NP-hard [41]. For the problem at hand, we propose a network programming algorithm that finds the exact solution in polynomial time. The algorithm is inspired by Flyn's minimum discontinuity approach [15], which minimizes the sum of $\left|\lfloor\Delta\phi^h_{ij} + \pi\rfloor\right|$ and $\left|\lfloor\Delta\phi^v_{ij} + \pi\rfloor\right|$. Flyn's objective function is, in fact, quite different from ours. However, both objective functions are the sum of first-order clique potentials depending only on $\Delta\phi^h_{ij}$ and $\Delta\phi^v_{ij}$. This structural similarity allows us to adapt the above ideas to our problem.

The following lemma assures that if the minimum of $E(\mathbf{k}|\boldsymbol{\psi})$ is not yet reached, then there exists a binary image $\delta\mathbf{k}$ (i.e., the elements of $\delta\mathbf{k}$ are all 0 or 1) such that $E(\mathbf{k}+\delta\mathbf{k}|\boldsymbol{\psi}) < E(\mathbf{k}|\boldsymbol{\psi})$.

Lemma 1: Let $\mathbf{k}_1$ and $\mathbf{k}_2$ be two wrap-counts images such that

$$E(\mathbf{k}_2|\boldsymbol{\psi}) < E(\mathbf{k}_1|\boldsymbol{\psi}). \tag{21}$$

Then there exists a binary image $\delta\mathbf{k}$ such that

$$E(\mathbf{k}_1 + \delta\mathbf{k}|\boldsymbol{\psi}) < E(\mathbf{k}_1|\boldsymbol{\psi}). \tag{22}$$

Proof: See Appendix A. ■

According to Lemma 1, we can iteratively compute $\mathbf{k}_i = \mathbf{k}_{i-1} + \delta\mathbf{k}$, where $\delta\mathbf{k} \in \{0,1\}^{N^2}$ minimizes $E(\mathbf{k}_{i-1} + \delta\mathbf{k}|\boldsymbol{\psi})$, until the the minimum energy is reached. Each minimization is a discrete optimization problem that can be exactly solved in polynomial time by using network programming techniques such as maximum flow [42] or minimum cut [43]. We note, however, that in the iterative scheme just described, it is not necessary to compute the exact minimizer of $E(\mathbf{k}_{i-1} + \delta\mathbf{k}|\boldsymbol{\psi})$ with respect to $\delta\mathbf{k}$, but only a binary image $\delta\mathbf{k}$ that decreases $E(\mathbf{k}_{i-1} + \delta\mathbf{k}|\boldsymbol{\psi})$. Based on this fact, we propose an efficient algorithm that iteratively searches for improving binary images $\delta\mathbf{k}$.

The following lemma, presented and proved in [15, Appendix], assures that if there exists an improving binary image $\delta\mathbf{k}$ [i.e., $E(\mathbf{k} + \delta\mathbf{k}|\boldsymbol{\psi}) < E(\mathbf{k}|\boldsymbol{\psi})$], then there exists another improving binary image $\delta\mathbf{l}$ such that the set $S_1(\delta\mathbf{l}) \equiv \{(i,j) \in Z \,|\, \delta l_{ij} = 1\}$ is connected in the first-order

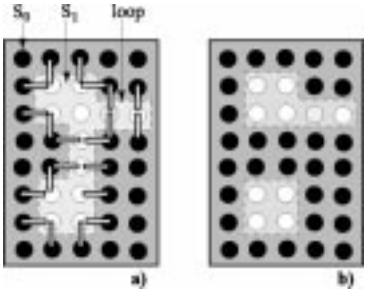

Fig. 4. Illustration of a binary partition. Set S_1 in part (a) defines a binary partition; small rectangles connect first-order neighbors with one element in S_1 an another in $Z - S_1$. The boundary of S_1 defines a loop. Set S_1 in part (b) does not define a binary partition.

neighborhood sense, i.e., given two sites s_1 and s_n of S_1 there exists a sequence of first-order neighbors, all in S_1, that begins in s_1 and ends in s_n. We call images $\delta\mathbf{l}$ with this property binary partitions of Z. Fig. 4 illustrates this concept. Set S_1 in part (a) defines a binary partition; small rectangles connect first-order neighbors with one element in S_1 an another in S_0. The boundary of S_1 defines a loop. Set S_1 in part (b) does not define a binary partition.

Lemma 2: Suppose that there exits a binary image $\delta\mathbf{k}$ such that

$$E(\mathbf{k}+\delta\mathbf{k}|\boldsymbol{\psi}) < E(\mathbf{k}|\boldsymbol{\psi}).$$

Then, there exists a binary partition of Z, $\delta\mathbf{l}$, such that

$$E(\mathbf{k}+\delta\mathbf{l}|\boldsymbol{\psi}) < E(\mathbf{k}|\boldsymbol{\psi}).$$

Proof: See Lemma 2 in [15, Appendix]. ■

Flyn's central idea is to search for improving binary partitions $\delta\mathbf{l}$, termed in [15] elementary operations (EO's). Once $\delta\mathbf{l}$ is found the wrap-count image $\mathbf{k}$ is updated to $\mathbf{k}+\delta\mathbf{l}$. If no EO is possible, then, according to Lemma 2, the energy $E(\mathbf{k}|\boldsymbol{\psi})$ can not be decreased by any binary image increment of the actual argument $\mathbf{k}$. Thus, by Lemma 1, $E(\mathbf{k}|\boldsymbol{\psi})$ has reached its minimum.

Given that that the clique potentials $(\Delta\phi^v_{ij})$ and $(\Delta\phi^h_{ij})$ are functions of phase differences computed between first-order neighbors, to check if a given binary partition $\delta\mathbf{l}$ improves the energy, one has to compute only the variation of those clique potentials containing sites on set $S_1(\delta\mathbf{l})$ and on its complement S_0 [sites marked with a small rectangle on Fig. 4(a)], i.e., one has to compute clique potentials of $E(\mathbf{k}|\boldsymbol{\psi})$ only along loops (this is still true on the boundary of Z, by taking zero potentials). Flyn's algorithm uses graph theory techniques to represent and generate EOs. The details of the implementation are presented in Appendix B.

B. Smoothing Step

The smoothing step amounts to computeing $\widehat{\boldsymbol{\psi}}$ given by

$$\widehat{\boldsymbol{\psi}} = \arg \max_{\boldsymbol{\psi}\in[-\pi,\pi[^{N^2}} \sum_{ij\in Z} \lambda_{ij}\cos(\phi_{ij}-\eta_{ij}) - \frac{\mu}{2}\sum_{ij\in Z_1} \left(\Delta\phi^h_{ij}\right)^2 \bar{v}_{ij} + \left(\Delta\phi^v_{ij}\right)^2 \bar{h}_{ij} \tag{23}$$

where $\phi_{ij} = 2\pi k_{ij} + \psi_{ij}$. The function to be maximized in (23) is not convex due to the terms $\lambda_{ij}\cos(\phi_{ij}-\eta_{ij})$. Computing $\widehat{\boldsymbol{\psi}}$ is therefore a hard problem. Herein, we adopt the ICM approach [40], which, in spite of being suboptimal, yields good results for the problem at hand.

ICM is a coordinatewise ascent technique where all coordinates are visited according to a given schedule. After some simple algebraic manipulation of the objective function (23), we conclude that its maximum with respect to ψ_{ij} is given by

$$\widehat{\psi}_{ij} = \arg \max_{\psi_{ij}\in[-\pi,\pi[} \left\{\beta_{ij}\cos(\psi_{ij}-\eta_{ij}) - \left(\psi_{ij}-\bar{\psi}_{ij}\right)^2\right\} \tag{24}$$

where

$$\beta_{ij} = 2\frac{\lambda_{ij}}{\mu\bar{l}_{ij}} \tag{25}$$

$$\bar{l}_{ij} = \bar{h}_{ij} + \bar{h}_{i+1,j} + \bar{v}_{ij} + \bar{v}_{i,j+1} \tag{26}$$

$$\bar{\psi}_{ij} = \bar{\phi}_{ij} - 2\pi k_{ij} \tag{27}$$

$$\bar{\phi}_{ij} = \frac{\phi_{i-1,j}\bar{h}_{ij} + \phi_{i,j-1}\bar{v}_{ij} + \phi_{i+1,j}\bar{h}_{i+1,j} + \phi_{i,j+1}\bar{v}_{i,j+1}}{\bar{l}_{ij}}. \tag{28}$$

There are no closed form solutions for maximization of (24), since it involves transcendental and power functions. We compute $\widehat{\psi}_{ij}$ using a simple two-resolution numeric method. First, we search $\widehat{\psi}_{ij}$ in the set $\{\pi i/M \,|\, i = -M,\ldots,M-1\}$. Next, we refine the search by using the set $\{\pi i_0/M + \pi i/M^2 \,|\, i = -M,\ldots,M-1\}$, where $\pi i_0/M$ is the result of the first search. We have used $M = 20$, which leads to the maximum error of $\pi/(20)^2$.

The phase estimate $\widehat{\psi}_{ij}$ depends in a nonlinear way of data η_{ij} and on the mean weighted phase $\bar{\psi}_{ij}$. The balance between these two components is controlled by parameter β_{ij}. Fig. 5 displays solutions of (24) as function of $\bar{\psi}_{ij}$, parameterized by β_{ij}. The principal phase value is $\eta = 0$. Assuming that $|\widehat{\psi}_{ij} - \eta_{ij}| \ll \pi$, then $\cos(\psi_{ij}-\eta_{ij})$ is well-approximated by the quadratic form $1-(\psi_{ij}-\eta_{ij})^2/2$, thus, leading to the linear approximation

$$\widehat{\psi}_{ij} \simeq \frac{\beta_{ij}\eta_{ij} + 2\bar{\psi}_{ij}}{\beta_{ij}+2}. \tag{29}$$

Reintroducing (29) in this condition, one gets $|\bar{\psi}_{ij} - \eta_{ij}| \ll 2\pi/(\beta_{ij}+2)$. If this condition is not met, the solution becomes highly nonlinear on η_{ij} and $\bar{\psi}_{ij}$. This is illustrated by Fig. 5: as $|\bar{\psi}_{ij} - \eta_{ij}|$ increases, at some point the phase $\widehat{\psi}_{ij}$ becomes clipped at $\pm\pi$, being therefore independent of the observed data η_{ij}.

In computing the ICM solution, we have updated sites column by column and each site was updated four times. We have noticed that practically no improvement is obtained by taking a large number of updates by site. This will be illustrated in the next section.

C. Incomplete Observations

Now suppose that data $\mathbf{x}$ is partially observed, i.e., data components x_{ij} are not observed on sites $Z - Z_o$. The $\mathbb{Z}\pi M$ algorithm presented in Fig. 3 still works, provided that we supply values for the initial principal phase values $\widehat{\psi}^{(0)}_{ij}$ in the set $Z - Z_o$

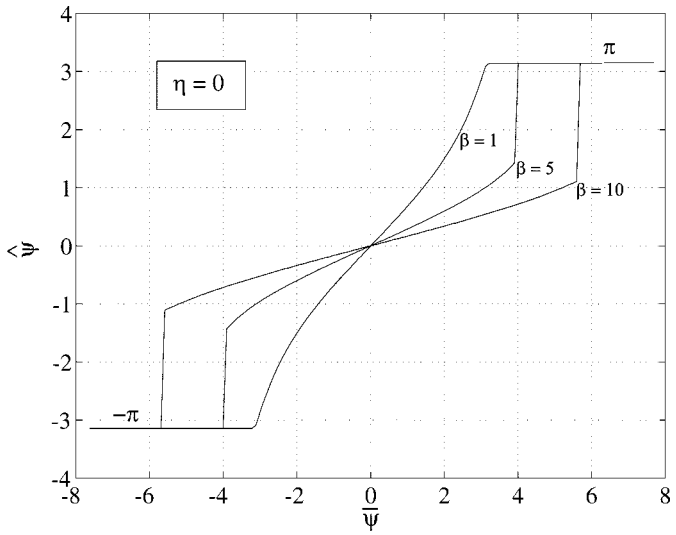

Fig. 5. Solution of maximization (25) as function of $\bar{\psi}$ (the mean principal phase), parameterized with β and for $\eta = 0$.

and take $\lambda_{ij} = 0$ also in this set. The drawback of this is that, due to almost certain inconsistence of observed phases η_{ij} for $(i,j) \in Z_o$ and $\widehat{\psi}_{ij}^{(0)}$ for $Z - Z_o$, the first $\mathbb{Z}$-step might produce poor results implying slower converge of the algorithm. To overcome this drawback, first we estimate phases on Z_0 as if the sites in $Z - Z_0$ were disconnected from sites in Z_0. Next we maximize the posterior density on $Z - Z_0$ given the phase estimates on Z_0.

The maximization of the posterior density on $Z - Z_0$ given the phase estimates on Z_0 is a quadratic problem that we solve again using ICM. A simple manipulation of $p_{\phi|\mathbf{x},\mathbf{l}}(\widehat{\boldsymbol{\phi}}|\mathbf{x},\mathbf{l})$ leads to the conclusion that its maximum with respect to ϕ_{ij}, $(i,j) \in (Z - Z_o)$ is given by (28) with $\bar{l}_{ij}$ given by (26).

With the purpose of disconnecting the nonobserved sites from the observed sites, we introduce the line field $\mathbf{l}'(Z_o)$ that signals a discontinuity between every site in Z_o that has a first-order neighbor in $Z - Z_o$. We introduce also the line field $\mathbf{l} \oplus \mathbf{l}'(Z_o)$, meaning the discontinuities denoted by $\mathbf{l}$ plus those of $\mathbf{l}'(Z_o)$.

Algorithm 1 $\mathbb{Z}\pi M$ procedure.
Input: $\boldsymbol{\eta}$, $\mathbf{l}$, Z_o, $t_{\max}$
1) $\widehat{\boldsymbol{\psi}}^{(0)} := \boldsymbol{\eta}$, $\mathbf{l} := \mathbf{l} \oplus \mathbf{l}'(Z_o)$ {*initialization*} {Estimate phase on the observed sites Z_o}
2) **for** $t := 1$ to $t := t_{\max}$ **do**
3) compute $\widehat{\mathbf{k}}^{(t)}$ ($\mathbb{Z}$-step) given by **(17)** {*$\mathbb{Z}$-step is implemented by steps 1,2 and 3 of the network programming procedure presented in Appendix B*}
4) compute $\widehat{\mu}^{(t)}$ given by **(10)**
5) **for** $n := 1$ to 4 **do**
6) for each site $(i,j) \in Z_o$ update $\widehat{\psi}_{ij}^{(t)}$ according to **(24)** (π-step) using a two-resolution numeric method
7) **end for**
8) **if** (stop test == true) **then**
9) **break loop for**
10) **end if**
11) **end for**
{Estimate phase on the nonobserved sites $Z - Z_o$}
12) **for** $t := 1$ to 10 **do**
13) for each site $(i,j) \in Z - Z_o$ update $\widehat{\psi}_{ij}^{(t)}$ given by **(28)**
14) **end for**

Algorithm 1 shows the complete $\mathbb{Z}\pi M$ procedure including situations where some sites are not observed. Steps 1–11 compute the phase $\widehat{\boldsymbol{\psi}}$ on sites Z_o by setting the line field to $\mathbf{l} \oplus \mathbf{l}'(Z_o)$. Steps 12–14 compute the phase $\widehat{\boldsymbol{\psi}}$ on sites $Z - Z_o$.

Concerning computational complexity, the $\mathbb{Z}$-step is, by far, the most demanding one, using a number of floating point operations very close to that required by Flyn's minimum discontinuity algorithm. Since the proposed scheme needs roughly four $\mathbb{Z}$-steps, it has approximately four times the complexity of Flyn's minimum discontinuity algorithm. To our knowledge there is no formula for the complexity of Flyn's algorithm (see Flyn's remarks about complexity in [15]). Nevertheless, we have found, empirically, a complexity of approximately $O(N^3)$ for the $\mathbb{Z}$-step.

IV. Experimental Results

The algorithm derived in the previous sections is now applied to synthetic InSAR pairs generated from both synthetic and real elevation data. The results are divided into two parts: (a) continuous surfaces and (b) discontinuous surfaces with unknown discontinuities. In part (a), we take the line field to be $\mathbf{l} = \mathbf{0}$, i.e., there are no discontinuities; in part (b) we assume the discontinuity locations are unknown, but belong to the nonobserved data set Z_0. Part (a) deals with InSAR pairs generated from synthetic elevation data, whereas part (b) deals with an InSAR pair generated from real elevation data.

Concerning the parameter λ, we use the maximum likelihood estimate (see [38, ch. 5]), assuming that image parameters are constant within 10×10 rectangular windows. When this assumption does not hold, the maximum likelihood estimates are poor, particularly the coherence α on which λ depends. If the phase variation inside 10×10 windows is greater than $\pi/4$, the coherence estimate becomes strongly biased. In these cases, we estimate the coherence based on 2×2 windows and then apply a 10×10 rectangular low-pass filter. In any event, there is no need to have precise estimates of λ since, in our experience, the algorithm is not sensitive to local fluctuations of this parameter.

Step 4 of the $\mathbb{Z}\pi M$ algorithm estimates the smoothness parameter μ iteratively according to (10). The algorithm is therefore adaptive with respect to this parameter. This scheme yields good results if the correlation factor is, roughly, greater than 0.7, i.e., the interferogram is not too noisy. Otherwise, phase estimates tend to be overly smooth, as the data term of the posterior density does not have enough strength to dominate the smoothing term. A solution is to impose an adequate prior on μ that prevents its estimates from growing to much. In this section, we push the idea of prior to an extreme and set μ to a constant. In a large set of experiments we observed that any value of $\mu \in [1,2]$ leads to good results. Having in mind that

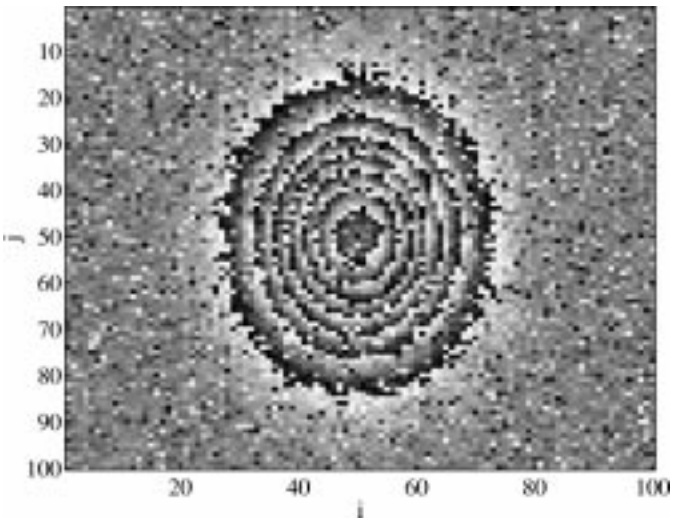

Fig. 6. Interfergram ($\boldsymbol{\eta}$-image) of a Gaussian elevation of height 14π rad and standard deviations $\sigma_i = 10$ and $\sigma_j = 15$ pixels. The correlation coefficient of the associated InSAR pair is $\alpha = 0.8$.

$\mu = \sigma^{-2}$, where σ^2 is the variance of the first-order phase differences $\Delta\phi^h_{ij}$ and $\Delta\phi^v_{ij}$ for $(i,j) \in Z_o$, then the interval $\mu \in [1,2]$ corresponds to the interval $\sigma \in [1/\sqrt{2}, 1]$. With σ in this interval, it is very unlike to have first-order phase differences larger in magnitude than π (approximately three standard deviations). This is therefore consistent with the fact that phase estimation is meaningless when first-order phase differences are larger than π in magnitude in a large number of sites. In this section we set $\mu = 1.5$.

A. Continuous Surfaces

Fig. 6 displays the interferogram ($\boldsymbol{\eta} = \{\eta_{ij}\}$ image) obtained with parameters $\theta = 1$, $\alpha = 0.8$ and $\sigma_n = 0$. The phase image $\boldsymbol{\phi}$ is a Gaussian elevation of height 14π rad and standard deviations $\sigma_i = 10$ and $\sigma_j = 15$ pixels. The magnitude of the phase difference $\Delta\phi^h_{ij}$ takes the maximum value of 2.5 and is greater than 2 in many sites. On the other hand, a correlation coefficient of $\alpha = 0.8$ implies a standard deviation of the maximum likelihood estimate η_{ij} of 0.91. This figure is computed based on the density of η obtained from the joint density (3). In these conditions, the task of phase estimation is extremely hard, as the interferogram exhibits a large number of inconsistencies, i.e., the observed image $\boldsymbol{\eta}$ is not consistent with the assumption of phase differences less than π at a large number of sites. In the unwrapping jargon we say that the interferogram has a lot of residues.

The $\mathbb{Z}\pi M$ estimates are presented in Fig. 7; part (a) shows the phase estimate $\widehat{\boldsymbol{\phi}}^{(1)}$ and part (b) shows the phase estimate $\widehat{\boldsymbol{\phi}}^{(10)}$.

Fig. 8 plots the logarithm of the posterior density $\ln p_{\phi|\mathbf{x},\mathbf{l}}(\widehat{\boldsymbol{\phi}}^{(t)}|\mathbf{x},\mathbf{l})$ and the L^2 norm of the estimation error $||\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}||^2 \equiv N^{-2}\sum_{ij}(\widehat{\phi}_{ij} - \phi_{ij})^2$ as function of the iteration t. The four nonintegers ticked between two consecutive integers refer to four consecutive ICM sweeps, implementing the π-step of the $\mathbb{Z}\pi M$ algorithm. Notice that the larger increment in

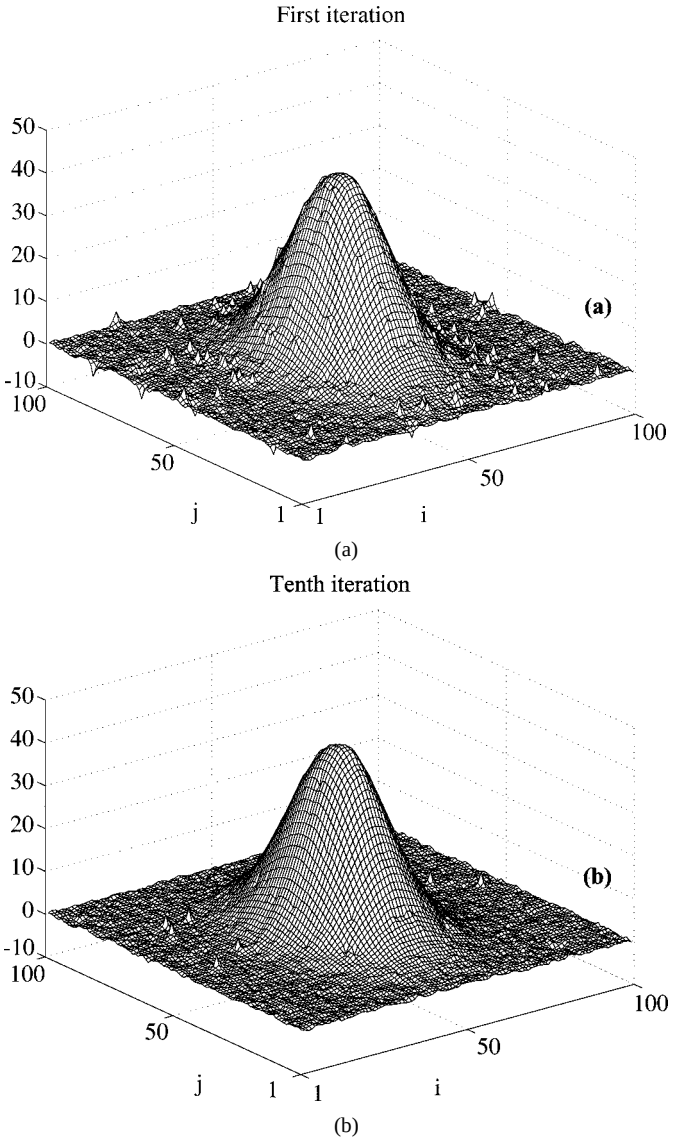

Fig. 7. Phase estimate $\widehat{\boldsymbol{\phi}}^{(t)}$; (a) $t = 1$ and (b) $t = 10$.

$\ln p_{\phi|\mathbf{x},\mathbf{l}}(\widehat{\phi}^{(t)}|\mathbf{x},\mathbf{l})$ happens in both steps of the first iteration. For $t \geq 2$ only the $\mathbb{Z}$-step produces noticeable increments in the posterior density. These increments are, however, possible due to the very small increments produced by the smoothing step. For $t > 4$ there is practically no improvement in the estimates.

Fig. 9 shows the histogram of the error $\{\widehat{\phi}^{(10)}_{ij} - \phi_{ij}\}$. The sample mean is -0.038 and the sample variance is 0.1. Notice that the estimator did not produce any error on the $\mathbf{k}$ component of $\boldsymbol{\phi}$ and reduced the initial variance of the interferogram from 0.91 to 0.1.

To rank the $\mathbb{Z}\pi M$ algorithm, we applied the following algorithms to the present problem:

- **Path following type:** Goldstein's branch cut (GBC) [44]; quality guided (QG) [18], [45]; and mask cut (MC) [46].
- **Minimum norm type:** Flyn's minimum discontinuity (FMD) [15]; weighted least-square (WLS) [19], [47]; and L^0 norm (L0N) (see [7, ch. 5.5]).
- **Bayesian type:** recursive nonlinear filters [25] and [26] (NLF).

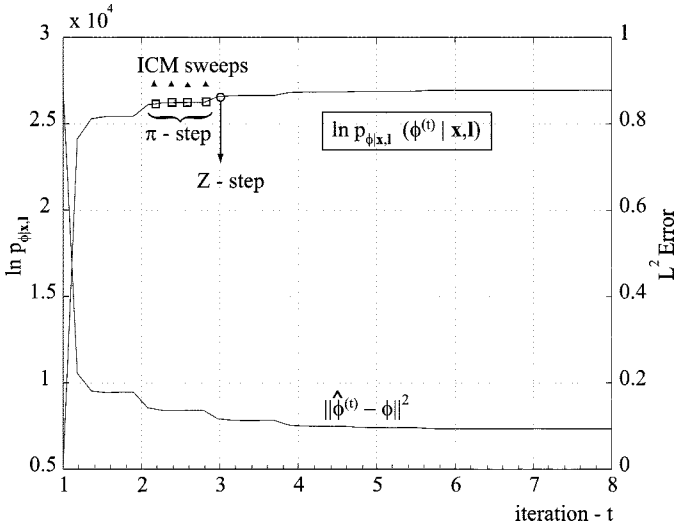

Fig. 8. Evolution of the logarithm of the posterior density $\ln p_{\boldsymbol{\phi}|\mathbf{x},\mathbf{l}}(\widehat{\boldsymbol{\phi}}^{(t)}|\mathbf{x},\mathbf{l})$ and of the L^2 norm of the estimation error as function of the iteration t. $\mathbb{Z}$-steps coincide with integers, whereas ICM sweeps implementing π-step are assigned to the noninteger part of t.

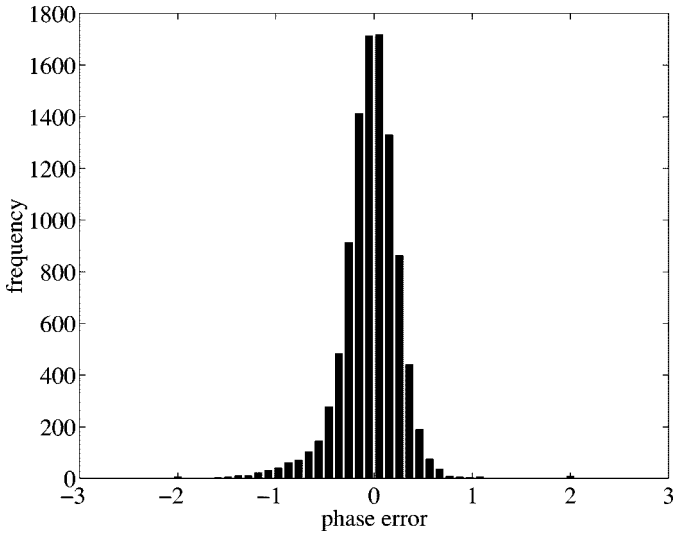

Fig. 9. Histogram of the error $\widehat{\boldsymbol{\phi}}^{(10)} - \boldsymbol{\phi}$.

The path following and minimum norm algorithms were implemented with the code supplied in the book [7], using the following settings: GBC (-dipole yes); QG, MC, (-mode min_var -tsize 3); and WLS (-mode min_var -tsize 3, -thresh yes). The unweighted versions of the FMD and L0N algorithms have been used in this subsection whereas the weighted version will be used in the following two subsections.

With the exception of NLF, all the algorithms referred to above are of the phase unwrapping type. We should be careful, therefore, in comparing $\mathbb{Z}\pi M$ method with those algorithms, as the former implements simultaneous smoothing and phase unwrapping, whereas the latter ones implement only the unwrapping step. This means that the estimates produced by the phase unwrapping methods depend on the type of smoothing applied previously to the interferogram. For this reason we considered two different scenarios: (a) nonsmooth interferogram; and (b) smooth interferogram produced by a maximum likelihood estimate of $\boldsymbol{\eta}$, using a 3×3 rectangular window (see [38, ch. 5]).

TABLE I
L^2 NORM OF THE ESTIMATION ERRORS OF $\mathbb{Z}\pi M$ AND OF CLASSICAL UNWRAPPING ALGORITHMS. THE LEFT COLUMN PLOTS RESULTS BASED ON THE MAXIMUM LIKELIHOOD ESTIMATE OF $\boldsymbol{\eta}$ USING A 3×3 RECTANGULAR WINDOW; THE RIGHT COLUMN PLOTS RESULTS BASED ON THE NON-SMOOTH $\boldsymbol{\eta}$ GIVEN BY (7)

	$\|\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\|^2$	
Algorithm	Smooth $\boldsymbol{\eta}$	Non-smooth $\boldsymbol{\eta}$
$\mathbb{Z}\pi$M	–	0.1
GBC	48.0	7.0
QG	10.0	2.2
MC	40.8	28.6
FMD	22.4	3.4
WLS	8.8	3.5
L0N	24.1	2.6
NLF	–	40.1

Table I displays the L^2 norm of the estimation error $\|\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\|^2$ for each of the classic algorithms referred to above. Apart from the proposed $\mathbb{Z}\pi M$ scheme, all the algorithms produced poor results, some of them catastrophic. The reasons depend on the class of algorithms and are basically the following.

- In the path following and minimum norm methods the noise filtering is the first processing step and is disconnected from the phase unwrapping process. The noise filtering assumes the phase to be constant within given windows. In data sets such as the one at hand, this assumption is catastrophic, even using small windows. On the other hand, if the smoothing step is not applied, even if the algorithm is able to infer most of the 2π-multiples, the observation noise is fully present in the estimated phase.
- The recursive nonlinear filtering solutions [25] and [26] fail because they use only the past observed data, in the lexicographic sense, to infer the phase.

As a final note with respect to this data set, we call attention to the estimation errors of the first $\mathbb{Z}$-step and of the FMD algorithm: these errors have L^2 norms of approximately 0.9 and 3.4, respectively. This difference illustrates what was said in the previous section: despite the structural similarity between the FMD algorithm and the $\mathbb{Z}$-step, the objective functions they minimize are different, leading to different estimates. The former basically minimizes an L^1 norm, while the latter minimizes an L^2 norm.

In this data set the FMD algorithm, contrary to the first $\mathbb{Z}$-step, was not able to unwrap part of the top of the Gaussian elevation. We do not have, however, any evidence that $\mathbb{Z}$-step always produces better results; for example, it is well known that the L^1 norm is more robust to outliers than the L^2 norm. Hence, the FMD algorithm should yield better results in phase surfaces with discontinuities not signaled. Notice, however, that the $\mathbb{Z}$-step can easily be modified to minimize any L^p norm with $p \geq 1$. This is a topic for future research.

Fig. 10 displays the interferogram obtained from an InSAR pair with parameters $\theta = 1$, $\alpha = 0.5$ and $\sigma_n = 0$. The phase

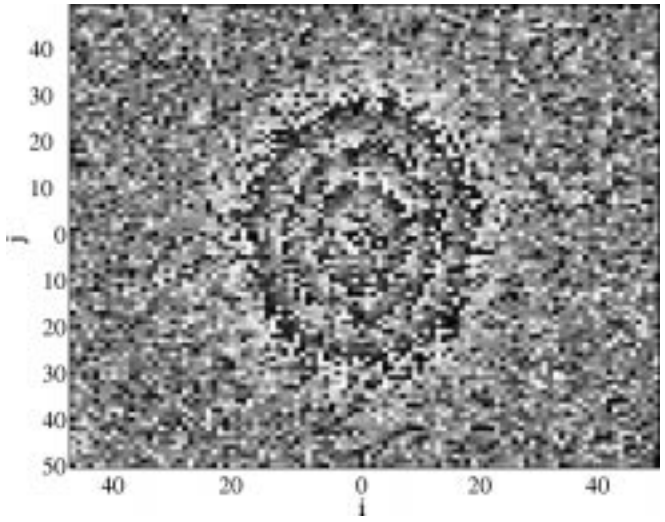

Fig. 10. Interferogram ($\boldsymbol{\eta}$-image) of a Gaussian elevation of height 7π rad and standard deviations $\sigma_i = 10$ and $\sigma_j = 15$. The correlation coefficient of the associated InSAR pair is $\alpha = 0.5$.

image $\boldsymbol{\phi}$ is a Gaussian elevation of height 7π rad and standard deviations $\sigma_i = 10$ and $\sigma_j = 15$. Compared with the previous example, the maximum phase difference has been reduced from 2.5 rad to 1.25 and the standard deviation of the error η_{ij} increased from 0.91 to 1.33. The smaller phase rate allows now filtering the interferogram in small windows without destroying the phase information as happened in the previous data set. Thus, phase estimation from this data set is not as hard as the previous case.

The $\mathbb{Z}\pi M$ estimates are displayed in Fig. 11; part (a) shows the phase estimate $\widehat{\boldsymbol{\phi}}^{(1)}$ and part (b) shows the phase estimate $\widehat{\boldsymbol{\phi}}^{(10)}$. Table II presents the L^2 norm $\|\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\|^2$ of the estimation error for each of the classic algorithm referred above. We consider however two smoothing filters: one with support of 4×4 pixels and the other with 3×3 pixels. The best phase estimate given by FMD and L0N methods exhibit an L^2 error approximately four times larger than the L^2 error of the proposed estimate. Notice the high sensitivity of the phase unwrapping methods to dimension of the filter. What happens is that the 4×4 smoothing filter destroys the phase information in areas of high phase rate, jeopardizing the success of phase unwrapping methods. On the other side, filters smaller than (3×3) do not enforce enough smoothness on the interferogram.

B. Discontinuous Surfaces With Unknown Discontinuities

In this section, we use a simulated InSAR example supplied in the book [7]. The data set was generated based on a real digital elevation model of mountainous terrain around Long's Peak, CO, using a high-fidelity InSAR simulator that models the SAR point spread function, the InSAR geometry, the speckle noise and the layover and shadow phenomena. For a detailed description of the simulator, see [7, ch. 3] and the references therein.

Fig. 12 shows a contour plot of the terrain used to generate the InSAR data. The size of the image in pixels is 458(azimuth) $\times$ 152(range). To compare the estimated surfaces directly with the "ground truth," the surface has been resampled in the SAR slant plane. Figs. 13 and 14 show the maximum likelihood estimate of the interferogram and the correlation factor, using a rectangular window of size 4(azimuth) $\times$ 4(range) and

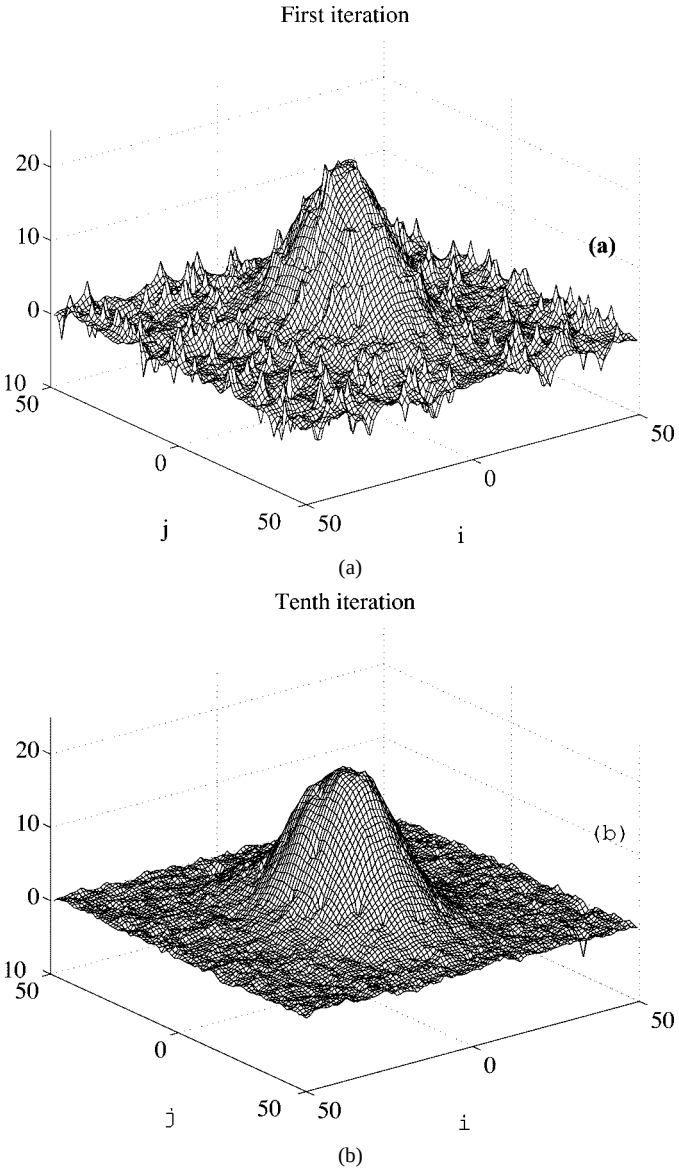

Fig. 11. Phase estimate $\widehat{\boldsymbol{\phi}}^{(t)}$: (a) $t = 1$ and (b) $t = 10$.

TABLE II
AS IN FOR A GAUSSIAN ELEVATION OF HEIGHT 7π rad AND STANDARD DEVIATIONS $\sigma_i = 10$ AND $\sigma_j = 15$. THE CORRELATION COEFFICIENT OF THE ASSOCIATED INSAR PAIR IS $\alpha = 0.5$

	$\|\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\|^2$		
Algorithm	Smooth $\boldsymbol{\eta}$ (4×4)	Smooth $\boldsymbol{\eta}$ (3×3)	Non-smooth $\boldsymbol{\eta}$
$\mathbb{Z}\pi$M	–	–	0.18
GBC	25	0.95	624
QG	60	1.05	5×10^3
MC	62	2.72	5×10^3
FMD	2.3	0.81	3.95
WLS	165	147.4	683
L0N	2.1	0.87	4.3
NLF	–	–	2×10^3

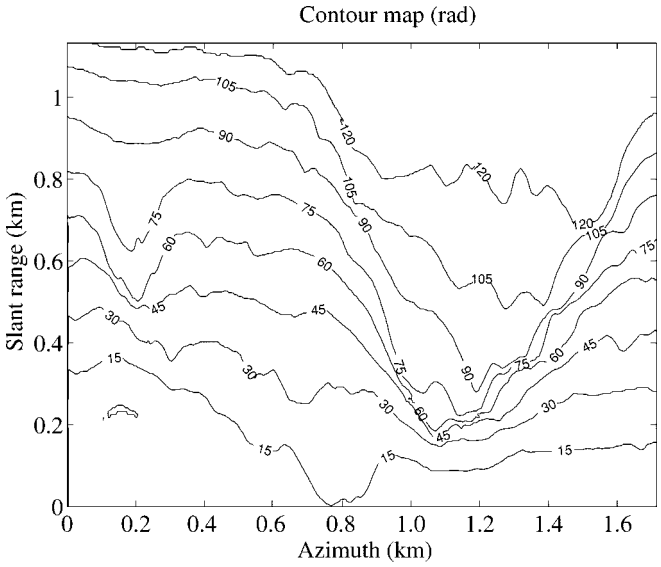

Fig. 12. Contour plot of the terrain used to generate the InSAR data. The surface, a digital terrain elevation model of mountainous terrain around Long's Peak, CO, has been resampled in the SAR slant plane; it can be therefore directly compared with the estimated surfaces. (Data distributed with [7]).

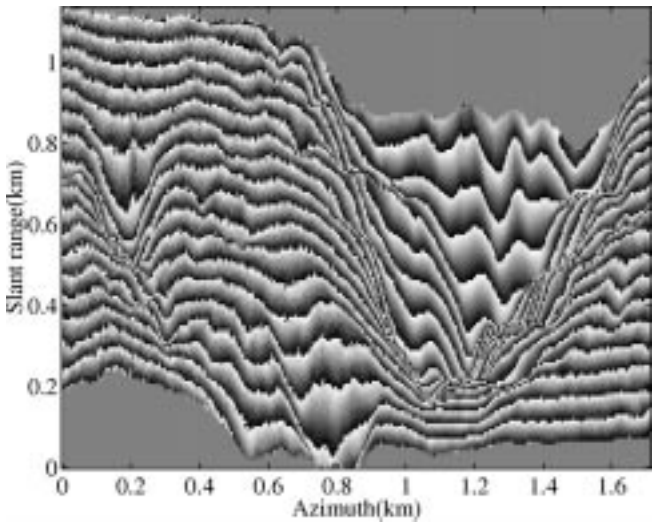

Fig. 13. Interferogram computed from the InSAR data generated based on the surface shown in Fig. 12. The layover phenomenon leads to very close fringes in some regions of the interferogram. (Data distributed with [7]).

taking $\sigma_n = 0$. The two flat regions in gray on the top and on the bottom of Fig. 13 correspond to undefined data due to the projection of the high terrain relief into the slant plane.

The SAR layover phenomenon leads to very close fringes in some regions of the interferogram, clearly visible in Fig. 13. In these regions, the assumption of constant phase within small regions is far from being true, leading to incorrect estimates of the principal phase values in the SAR layover regions. Therefore, the principal phase values in these regions should not be used, as they are inconsistent with the true phase values. In terms of the $\mathbb{Z}\pi M$ algorithm, this means that the phase should be inferred from the data observed in a subset Z_o of Z.

To infer the set Z_o we have adopted the phase derivative and thresholding procedures presented in [7, ch. 3]. Fig. 15 represents in white the set of sites Z_o. Notice that the sites in black representing the set $Z - Z_o$ are mostly in the layover regions.

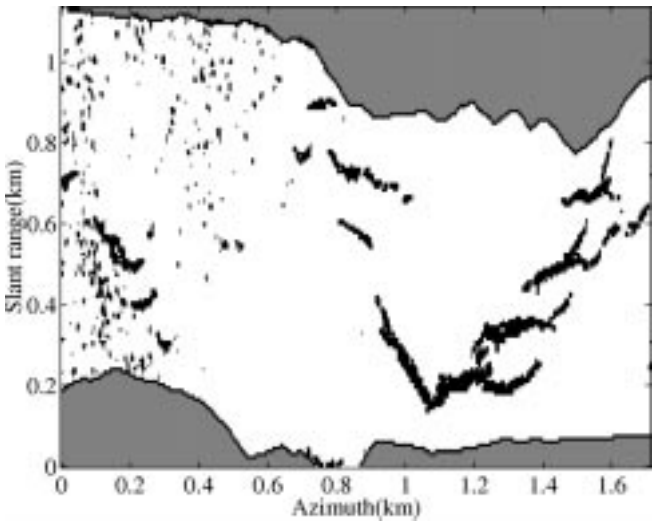

Fig. 15. Quality map relative to the interferogram shown in Fig. 13, computed using the phase derivative and thresholding procedures presented in [7, ch. 3]. Black color signals sites where the interferogram is of low quality.

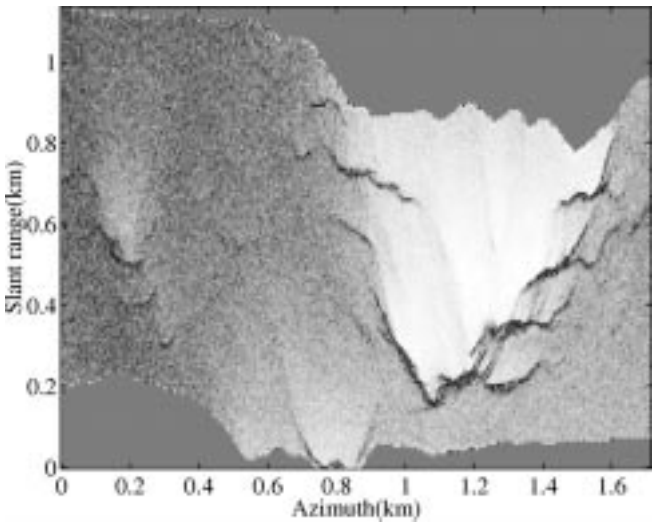

Fig. 14. Correlation factor estimated from the InSAR data generated based on the surface shown in Fig. 12. Notice the very low and erroneous estimated correlation in the regions of SAR layover. The large low-correlation region on the left of the image is due to image misregistration. (Data distributed with [7].)

To compute the image $\{\lambda_{ij}\}$ we need the amplitude $|x_1 x_2|$ for all image pixels [see definition (7)]. As this data has not been supplied in the the book [7], we have generated it under the assumption that the electronic noise σ_n is zero and that the power $\theta^2 = E[|x|^2] = 1$ for all image pixels. This choice is justified noting that if $\sigma_n = 0$ the density of λ does not depend on θ.

Fig. 16 shows the error map $\{|\widehat{\phi}_{ij}^{(2)} - \phi_{ij}|\}$ produced by the $\mathbb{Z}\pi M$ algorithm. We have only taken $t = 2$ (two iterations) as more $\mathbb{Z}\pi M$ iterations do not improve the phase estimate. Larger errors are confined to the sites signaled as low-quality in the quality map shown in Fig. 15 (i.e., sites in the set $Z - Z_o$). This was to be expected as the phase in these region is determined by the estimated phase along its boundary. The overall phase

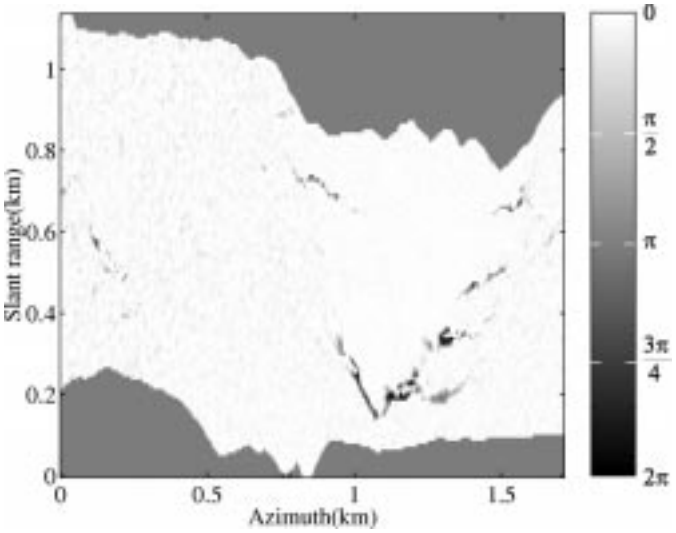

Fig. 16. Error map image $\{|\widehat{\phi}_{ij}^{(2)} - \phi_{ij}|\}$ produced by the $\mathbb{Z}\pi M$ algorithm. Black color denotes an absolute error greater or equal to 2π. Larger errors are confined to the low-quality sites signaled in the quality map shown in Fig. 15.

TABLE III
COMPUTATION TIMES FOR THE LONG'S PEAK EXAMPLE IN A PC WORKSTATION EQUIPPED WITH A 350 MHZ PENTIUM-III CPU

	Computation Time (sec.)	
Algorithm	152×458	304×916
$\mathbb{Z}\pi$M	220	1600
FMD	100	780
L0N	330	1300

estimate is of good quality as confirmed by the L^2 estimate error of 0.78. This value decreases to 0.05 if we measure the error only in the observed set Z_0.

We have applied FMD and L0N phase unwrapping algorithms to the same data set using the same quality map. These algorithms were selected because they are ranked as the best phase unwrapping techniques (see [7, ch. 6] and [48]). The L^2 estimate error is 1.53 for the FMD algorithm and 0.91 for the L0N algorithm. These values decrease to 0.1 and 0.09, respectively, if the error is measured only over the observed set Z_0. The $\mathbb{Z}\pi M$ yields the lower L^2 error. The improvement over the L0N algorithm is almost a factor of two in the observed set Z_0.

All results presented in this section were obtained with a PC workstation equipped with a 350 Mhz Pentium-III CPU. Table III shows the computation times for the Long's Peak data set; the left column refers to the example distributed with book [7]; the right column refers to data generated synthetically by upsampling the surface $\boldsymbol{\phi}$ and the correlation $\boldsymbol{\alpha}$ with a factor of two and then generation of a data set according to the observation model (3). As expected, the computation time of the $\mathbb{Z}\pi M$ algorithm is, approximately, twice the computation time of the FMD algorithm. Concerning computer complexity, both algorithms display values that conform to the $O(N^3)$ trend discussed in Section III.

V. CONCLUDING REMARKS

This paper presents an effective algorithm for absolute phase estimation in interferometric aperture radar/sonar (InSAR/InSAS) applications. The Bayesian standpoint was adopted. The likelihood function, which models the observation mechanism, is 2π-periodic and accounts for interferometric pair decorrelation and system noise. The *a priori* probability of the absolute phase is modeled by a noncausal first-order compound Gauss–Markov random field (CGMRF). This prior is suited to piecewise smooth fields, in the sense that it enforces smoothness, in a statistical way, between neighboring sites not split by discontinuities. The adopted framework also models incomplete data observations.

The CGMRF prior is parameterized by the location of the discontinuities, the so-called line field. Due to the periodic structure of the likelihood function and to the interferometric noise, the determination of the line field is an ill-posed problem, i.e., it can not be uniquely determined from the observed data. The line field might, however, be inferred using information external to the absolute phase estimation framework, as is proposed for example in [35], which aims at the detection of discontinuities existing between layover and nonlayover areas in InSAR/InSAS applications by exploiting the spectral shift that exists between the signal read by the two sensors as function of the along range local slope. Another approach for handling discontinuities/inconsistencies is to segment the observed data into a phase-consistent region and its complement and use only the observed data in the former region to estimate the whole absolute phase. Our approach accepts both: a discontinuity field and a region of consistency. These ideas were illustrated with examples in Section IV. Detailed procedures aiming at discontinuity detection or data segmentation are, however, beyond the scope of this paper.

To compute the absolute phase estimate we adopted the maximum *a posteriori* (MAP) criterion. We derived a suboptimal iterative procedure consisting of two steps per iteration: the first step, termed the $\mathbb{Z}$-step, maximizes the posterior density with respect to the 2π phase multiples; the second step, termed the π-step, maximizes the posterior density with respect to the phase principal values. The $\mathbb{Z}$-step is a discrete optimization problem solved exactly by network programing techniques inspired by Flyn's*minimum discontinuity algorithm* [15]. The π-step is a continuous optimization problem solved approximately by the *iterated conditional modes* (ICM) procedure. We named the proposed algorithm $\mathbb{Z}\pi M$, where the letter M stands for maximization.

The $\mathbb{Z}\pi M$ algorithm accounts for the observation noise in a model-based fashion. More specifically, the observation mechanism took into account the electronic noise and the decorrelation noise. This is a crucial feature that underlies the advantage of the $\mathbb{Z}\pi M$ algorithm over the classical path following and minimum-norm schemes, mainly in regions where the phase rate is close to π. In fact, these schemes approach absolute phase estimation by splitting the problem into two separate steps: in the first the noise in the interferogram is filtered out by applying low-pass filtering; in the second step, termed phase unwrapping, the 2π phase multiples are computed. For high phase rate regions, the application of first step make it impossible to recover the absolute phase,

as the principal values estimates are highly biased. This is in contrast with the $\mathbb{Z}\pi M$ algorithm, where the first step, the $\mathbb{Z}$-step, is an unwrapping applied over the observed interferogram.

To evaluate the performance of the $\mathbb{Z}\pi M$ algorithm, it was applied to two classes of problems: continuous surfaces and discontinuous surfaces with unknown discontinuities. In the latter class, the $\mathbb{Z}\pi M$ algorithm was fed with a consistency map using the phase derivative and thresholding procedures presented in [7, ch. 3]. In all examples studied we have compared the computed estimates with those provided by the best path following and minimum-norm schemes, namely the Goldstein's branch cut, the quality guided, the Flyn's minimum discontinuity, the weighted least-square and the L^0 norm. The proposed algorithm yields excellent results in all examples considered, performing, in some cases, much better than the alternative techniques just mentioned.

Concerning computer complexity, the $\mathbb{Z}\pi M$ algorithm takes, approximately, a number of floating point operations proportional to the 1.5 power of the number of pixels. By far, the $\mathbb{Z}$-step is the most demanding, using a number of floating point operations very close to the Flyn's minimum discontinuity algorithm. Since the proposed scheme needs roughly four $\mathbb{Z}$-steps, is has approximately four times the Flyn's minimum discontinuity algorithm complexity.

Concerning future developments, we foresee the integration of the principal phase values in the posterior density as a major research direction. If this goal were reached then the wrap-count image would be the only unknown of the obtained posterior density and, most importantly, there would be no need for iterations in estimating the wrap-count image. After obtaining this image, the principal phase values could be obtained using the π-step of the $\mathbb{Z}\pi M$ algorithm.

Appendix A

Proof of Lemma 1

The structure of the proof is essentially that of Appendix A of [15]. Define $\Delta k_{ij} = [\mathbf{k}_2]_{ij} - [\mathbf{k}_1]_{ij}$, for $(i,j) \in Z$. Given that the energy function $E(\mathbf{k}|\boldsymbol{\psi})$ depend only on differences between elements of $\mathbf{k}$, we take $\Delta k_{ij} \geq 0$ for $(i,j) \in Z$. Define $n = \max_{ij}(\Delta k_{ij})$ and the wrap-count image sequence $\{\mathbf{k}^{(t)}, t = 0,\ldots,n\}$, such that $\mathbf{k}^{(0)} = \mathbf{k}_1$, $\mathbf{k}^{(n)} = \mathbf{k}_2$ and

$$k_{ij}^{(t)} = k_{ij}^{(0)} + \min(t, \Delta k_{ij}), \quad t = 0,\ldots,n. \tag{30}$$

The energy variation $\Delta E \equiv E(\mathbf{k}_2|\boldsymbol{\psi}) - E(\mathbf{k}_1|\boldsymbol{\psi})$ can be decomposed as

$$\Delta E \equiv \sum_{t=1}^{n} \underbrace{[E\left(\mathbf{k}^{(t)}|\boldsymbol{\psi}\right) - E\left(\mathbf{k}^{(t-1)}|\boldsymbol{\psi}\right)]}_{\Delta E^{(t)}}.$$

Since $\Delta E < 0$ by hypothesis, then at least one of the terms $\Delta E^{(t)}$ of the above sum is negative. The lemma is proved if we show that the variation $\delta E^{(t)} \equiv E(\mathbf{k}^{(0)} + \delta\mathbf{k}^{(t)}|\boldsymbol{\psi}) - E(\mathbf{k}^{(0)}|\boldsymbol{\psi})$ satisfies $\delta E^{(t)} \leq \Delta E^{(t)}$, where $\delta\mathbf{k}^{(t)} \equiv \mathbf{k}^{(t)} - \mathbf{k}^{(t-1)}$, for any $t = 1,\ldots,n$. This condition is equivalent to

$$\begin{aligned} 0 \leq\ & E\left(\mathbf{k}^{(t)}|\boldsymbol{\psi}\right) - E\left(\mathbf{k}^{(t-1)}|\boldsymbol{\psi}\right) \\ & - E\left(\mathbf{k}^{(0)} + \mathbf{k}^{(t)} - \mathbf{k}^{(t-1)}|\boldsymbol{\psi}\right) + E\left(\mathbf{k}^{(0)}|\boldsymbol{\psi}\right) \end{aligned} \tag{31}$$

for $t = 1,\ldots,n$. Introducing (18) into (31), we obtain $0 \leq S^h + S^v$, where

$$\begin{aligned} S^h = & \sum_{ij} \left[\left(\Delta\phi_{ij}^{h(t)}\right)^2 - \left(\Delta\phi_{ij}^{h(t-1)}\right)^2 + \left(\Delta\phi_{ij}^{h(0)}\right)^2 \right. \\ & \left. - \left(\Delta\phi_{ij}^{h(0)} + \Delta\phi_{ij}^{h(t)} - \Delta\phi_{ij}^{h(t-1)}\right)^2 \right] \bar{v}_{ij} \end{aligned} \tag{32}$$

$$\begin{aligned} S^v = & \sum_{ij} \left[\left(\Delta\phi_{ij}^{v(t)}\right)^2 - \left(\Delta\phi_{ij}^{v(t-1)}\right)^2 + \left(\Delta\phi_{ij}^{v(0)}\right)^2 \right. \\ & \left. - \left(\Delta\phi_{ij}^{v(0)} + \Delta\phi_{ij}^{v(t)} - \Delta\phi_{ij}^{v(t-1)}\right)^2 \right] \bar{h}_{ij} \end{aligned} \tag{33}$$

where $\Delta\phi_{ij}^{h(t)}$ and $\Delta\phi_{ij}^{v(t)}$ are given by (19) and (20), respectively, computed at the wrap-count image $\mathbf{k}^{(t)}$. To prove (31), we now show that the terms of S^h corresponding to a given site $(i,j) \in Z$ have positive sum. The same is true concerning S^v.

The difference $k_{ij}^{(t)} - k_{i,j-1}^{(t)}$, for $t = 0,\ldots,n$, is a monotone sequence. This is a consequence of the definition (30): if $\Delta k_{ij} > \Delta k_{i,j-1}$ the sequence is monotone increasing; if $\Delta k_{ij} \leq \Delta k_{i,j-1}$ the sequence is monotone decreasing. Therefore the sequence $\left\{\Delta\phi_{ij}^{h(t)}\right\}$, for $t = 0,\ldots,n$, is also monotone. Define $a \equiv \Delta\phi_{ij}^{h(0)}$, $b \equiv \Delta\phi_{ij}^{h(t-1)}$ and $c \equiv \Delta\phi_{ij}^{h(t)}$. The sum of terms of S^h corresponding to the site (i,j) is

$$c^2 - b^2 + a^2 - (a - b + c)^2 = 2(b - a)(c - b). \tag{34}$$

Since $a \leq b \leq c$ or $a \geq b \geq c$, the right hand side of (34) is always nonnegative, as we want to prove. The same reasoning applies to S^v. ■

Appendix B

$\mathbb{Z}$-Step Implementation

Fig. 17 shows an auxiliary graph, whose nodes are interleaved with the phase sites. The edges denote which wrap-counts are to be incremented: a leftward (rightward) edge indicates an unit increment of the wrap-count below (above) the edge. A downward (upward) edge indicates a unit increment of the wrap-count right (left) of the edge. The algorithm works by creating and extending paths made of directed edges. When a path is extended to form a loop, the algorithm performs an EO, removes the loop from the collection of paths and resumes the path extension.

Assume that the array of auxiliary nodes has indices in the set $\{(i,j)\,|\,i,j = 1,\ldots,N+1\}$. Define the cost of an edge $\delta V(i,j;i',j')$ between the first-order neighbors $(i',j')$and (i,j) as $E(\mathbf{k}|\boldsymbol{\psi}) - E(\mathbf{k} + \delta\mathbf{k}|\boldsymbol{\psi})$, where $\delta\mathbf{k}$ is the wrap-count increment induced by the edge. With this definitions and having in attention the structure of $E(\mathbf{k}|\boldsymbol{\psi})$ [see (18)], we are led to

$$\begin{aligned} \delta V(i,j;i,j-1) &= -4\pi\left(\pi + \Delta\phi_{i,j-1}^{v}\right)\bar{h}_{i,j-1} \\ \delta V(i,j-1;i,j) &= -4\pi\left(\pi - \Delta\phi_{i,j-1}^{v}\right)\bar{h}_{i,j-1} \\ \delta V(i-1,j;i,j) &= -4\pi\left(\pi + \Delta\phi_{i-1,j}^{h}\right)\bar{v}_{i-1,j} \\ \delta V(i,j;i-1,j) &= -4\pi\left(\pi - \Delta\phi_{i-1,j}^{h}\right)\bar{v}_{i-1,j}. \end{aligned}$$

The values of boundary edges are defined to be zero; i.e., $\delta V(1,j) = \delta V(N+1,j) = \delta V(i,1) = \delta V(i,N+1) = 0$.

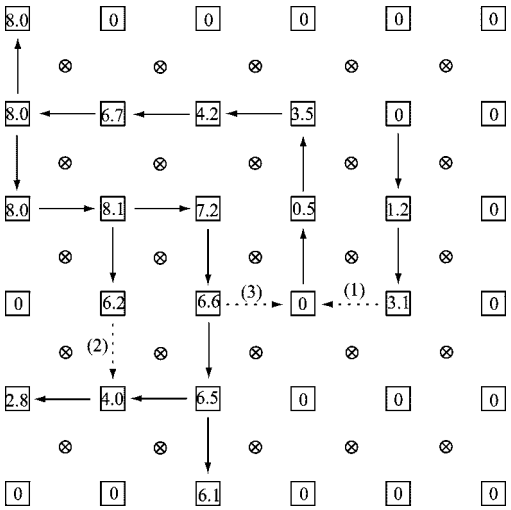

Fig. 17. Auxiliary graph to implement Flyn's algorithm (squared nodes) interleaved with phase sites (circled and crossed nodes). A leftward (rightward) edge indicates an unit increment of the wrap-count below (above) the edge. A downward (upward) edge indicates an unit increment of the wrap-count right (left) of the edge.

Fig. 17 represents the state of the graph at a given instant. Assuming that there are no loops, the set of edges defines a given number of trees. The value of each node, $V(i,j)$, is the sum of edge values corresponding to the path between the node and the tree root. In Fig. 17 there are two trees. We stress that the node values are real numbers, whereas in Flyn's algorithm they are integers. The reason is that our energy E takes values in the nonnegative reals while Flyn's energy takes values on the positive integers.

The basic step of Flyn's algorithm is to revise the set of paths by adding a new edge. An edge from (i,j) to a first-order neighbor (i',j'), if not presented, is added if

$$\Delta V \equiv V(i,j) + \delta V(i,j;i',j') - V(i',j') > 0.$$

If $\Delta V \leq 0$ then the new path to (i',j') would have a negative or zero value or would fail to improve an existing path. If the edge is added, the set of paths is revised in one of the three possible ways (a minor modification of [15]):

A. Edge Addition

If (i',j') is not a root or isolated node and the path to (i,j), if any, does start at (i',j'), then the algorithm adds the edge and leaves existing edges unchanged. If both nodes are isolated, this starts a new tree. If (i',j') is a root, the value of all nodes in the subtree is increased by ΔV.

B. Edge Replacement

If (i',j') is a branch node and the path to (i,j), if any, does not contain (i',j'), then the algorithm removes the existing edge to (i',j') and adds the new edge. The paths through (i',j') are changed to include the new edge and the values in its subtree are increased by ΔV.

C. Loop Completion

If the path to (i,j) contains (i',j'), then the new edge completes a loop. The algorithm applies the corresponding EO $\mathbf{k}+\delta\mathbf{l}$. For counterclockwise (clockwise) loops, $\delta l_{mn} = 1$ for those sites inside (outside) the loop. The edge values along the loop are updated according to the new wrap-count values. The values of some paths may become negative or zero, making then invalid. To account for this, all the paths containing loop edges are removed and the values of the resulting isolated nodes are set to zero.

The dashed edges in Fig. 17 illustrate graph revisions of type 1, 2, and 3. For a more detailed example, see Flyn's paper [15].

The algorithm alternates between type 1 and type 2 revisions until a loop is found, performing then a type 3 revision. If, for any attempt of edge addition $\Delta V \leq 0$, then no loop completion is possible and, according to Lemma 2 and Lemma 1, the algorithm terminates.

REFERENCES

[1] P. Rosen, S. Hensley, I. Joughin, F. Li, S. Madsen, E. Rodriguez, and R. Goldstein, "Synthetic aperture radar interferometry," *Proc. IEEE*, vol. 88, pp. 333–382, Mar. 2000.

[2] H. Griffiths, T. Rafik, Z. Meng, C. Cowan, H. Shafeeu, and D. Anthony, "Interferometic synthetic aperture sonar for high- resolution 3-d mapping of the seabed," *Proc. Inst. Elect. Eng.*, vol. 144, no. 2, pp. 96–103, April 1997.

[3] P. Jezzard and R. Balaban, "Correction for geometric distortion in echo-planar images from B_0 field variations," *Magn. Reson. Med.*, vol. 34, pp. 65–73, 1995.

[4] G. Glover and E. Schneider, "Three-point Dixon technique for true water/fat decomposition with B_0 inhomogeneity correction," *Magn. Reson. Med.*, vol. 18, pp. 371–383, 1991.

[5] S. Pandit, N. Jordache, and G. Joshi, "Data-dependent systems methodology for noise-insensitive phase unwrapping in laser interferometric surface characterization," *J. Opt. Soc. Amer.*, vol. 11, no. 10, pp. 2584–2592, 1994.

[6] A. Devaney, "Diffraction tomographic reconstruction from intensity data," *IEEE Trans. Image Processing*, vol. 1, pp. 221–228, Apr. 1992.

[7] D. Ghiglia and M. Pritt, *Two-Dimentional Phase Unwrapping. Theory, Algorithms and Software*. New York: Wiley, 1998.

[8] J. Strand, "Two-dimensional phase unwrapping with applications," Ph.D., Dept. Math., Fac. Math. Nat. Sci., Univ. Bergen, 1999.

[9] J. Strand, T. Taxt, and A. Jain, "Two-dimensional phase-unwrapping using a block least square method," *IEEE Trans. Image Processing*, vol. 8, pp. 375–386, Mar. 1999.

[10] S. Madsen, H. Zebker, and J. Martin, "Topographic mapping using radar interferometry: Processing techniques," *IEEE Trans. Geosci. Remote Sensing*, vol. 31, pp. 246–256, Jan. 1993.

[11] D. Ghiglia and L. Romero, "Minimum L^p norm two-dimensional phase unwrapping," *J. Opt. Soc. Amer.*, vol. 13, no. 10, pp. 1999–2013, 1996.

[12] D. Fried, "Least-squares fitting a wave-front distortion estimate to an array of phase-difference measurements," *J. Opt. Soc. Amer.*, vol. 67, no. 3, pp. 370–375, 1977.

[13] R. Hudgin, "Wave-front reconstruction for compensated imaging," *J. Opt. Soc. Amer.*, vol. 67, no. 3, pp. 375–378, 1977.

[14] R. Frost, C. Rushforth, and B. Baxter, "Fast FFT-algorithm for phase estimation in speckle imaging," *Appl. Opt.*, vol. 18, pp. 2056–2051, 1979.

[15] T. Flynn, "Two-dimensional phase unwrapping with minimum weighted discontinuity," *J. Opt. Soc. Amer. A*, vol. 14, no. 10, pp. 2692–2701, 1997.

[16] M. Costantini, "A novel phase unwrapping method based on network programing," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 813–821, May 1998.

[17] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Processing*, vol. 2, pp. 296–310, July 1993.

[18] H. Lim, W. Xu, and X. Huang, "Two new practical methods for phase unwrapping," in *Proc. 1995 Int. Geoscience Remote Sensing Symp.*, Firenze, Italy, 1995, pp. 196–198.

[19] D. Ghiglia and L. Romero, "Roboust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods," *J. Opt. Soc. Amer.*, vol. 11, no. 1, pp. 107–117, 1994.

[20] J. Marroquin and M. Rivera, "Quadratic regularization functionals for phase unwrapping," *J. Opt. Soc. Amer.*, vol. 11, no. 12, pp. 2393–2400, 1995.
[21] L. Guerriero, G. Nico, G. Pasquariello, and S. Starmaglia, "New regularization scheme for phase unwrapping," *Appl. Opt.*, vol. 37, no. 14, pp. 3053–3058, 1998.
[22] M. Rivera, J. Marroquin, and R. Rodriguez-Vera, "Fast algorithm for integrating inconsistent gradient fields," *Appl. Opt.*, vol. 36, no. 32, pp. 8381–8390, 1995.
[23] M. Servin, J. Marroquin, D. Malacara, and F. Cuevas, "Phase unwrapping with a regularized phase-tracking system," *Appl. Opt.*, vol. 37, no. 10, pp. 19 171–1923, 1998.
[24] J. Leitão and M. Figueiredo, "Interferometric image reconstruction as a nonlinear Bayesian estimation problem," in *Proc. 1st IEEE Int. Conf. Image Processing (ICIP'95)*, vol. 2, 1995, pp. 453–456.
[25] ——, "Absolute phase image reconstruction: A stochastic nonlinear filtering approach," *IEEE Trans. Image Processing*, vol. 7, pp. 868–882, June 1998.
[26] J. Dias and J. Leitão, "Simultaneous phase unwrapping and speckle smoothing in SAR images: A stochastic nonlinear filtering approach," in *EUSAR'98 Eur. Conf. Synthetic Aperture Radar*, Friedrichshafen, Germany, May 1998, pp. 373–377.
[27] G. Nico, G. Palubinskas, and M. Datcu, "Bayesian approach to phase unwrapping: Theoretical study," *IEEE Trans. Signal Processing*, vol. 48, pp. 2445–2556, Sept. 2000.
[28] Koetter, B. Frey, N. Petrivic, and D. Munson, "Unwrapping phase images by propagating probabilities across graphs," in *Proc. Int. Conf. Acoustics, Speech. Signal Processing (ICASSP'01)*, vol. 3, Salt Lake City, UT, 2001, pp. 1845–1848.
[29] K. Achan, B. Frey, R. Koetter, and D. Munson, "Unwrapping phases by relaxed mean field inference," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP'01)*, vol. 3, Salt Lake City, UT, 2001, pp. 1901–1904.
[30] C. Chen and H. Zebker, "Two-dimensional phase unwrapping with use of statistical models for cost functions in nonlinear optimization," *J. Opt. Soc. Amer. A*, vol. 18, pp. 338–351, 2001.
[31] ——, "Network approaches to two-dimensional phase unwrapping: Intractability and two new algorithms," *J. Opt. Soc. Amer. A*, vol. 17, pp. 401–414, 2000.
[32] B. Friedlander and J. Francos, "Model based phase unwrapping of 2-d signals," *IEEE Trans. Signal Processing*, vol. 44, pp. 2999–3007, Dec. 1996.
[33] Z. Liang, "A model-based for phase unwrapping," *IEEE Trans. Med. Imag.*, vol. 15, pp. 893–897, Dec. 1996.
[34] F. Jeng and J. Woods, "Compound Gauss-Markov random fields for image processing," in *Digital Image Restoration*, A. Katsaggelos, Ed. Berlin, Germany: Springer-Verlag, 1991, pp. 89–108.
[35] F. Gatelli, A. Guarnieri, P. Pasquali, C. Prati, and F. Rocca, "The wavenumber shift in SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 855–865, July 1994.
[36] D. Angwin and H. Kaufman, "Image restoration using reduced order models," *Signal Process.*, vol. 16, no. 89, pp. 21–28.
[37] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. R. Statist. Soc. B*, vol. 36, no. 2, pp. 192–225, 1974.
[38] C. Jakowatz, D. Wahl, P. Eichel, D. Ghiglia, and P. Thompson, *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*. Norwell, MA: Kluwer, 1996.
[39] E. Rignot and R. Chellappa, "Segmentation of polarimetric synthetic aperture radar data," *IEEE Trans. Image Processing*, vol. 1, pp. 281–300, July 1992.
[40] J. Besag, "On the statistical analysis of dirty pictures," *J. R. Statist. Soc. B*, vol. 48, no. 3, pp. 259–302, 1986.
[41] M. Grotschel, L. Lovasz, and A. Schrijver, *Beometric Algorithms and Combinatorial Optimization, Algorithms and Combinatorics*. New York: Springer-Verlag, 1988.
[42] D. Greig, B. Porteus, and A. Seheult, "Exact maximum a posteriory estimation for binary images," *J. R. Statist. Soc. B*, vol. 51, no. 2, pp. 271–279, 1989.
[43] Y. Boykov, O. Veksler, and R. Zabih, "A new minimization algorithm for energy minimaization with discontinuities," in *Energy Minimization Methods in Computer Vision and Pattern Recognition—EMMCVPR'99*, E. Hancock and M. Pelillo, Eds. New York: Springer, 1999, pp. 205–220.
[44] R. Goldstein, H. Zebker, and C. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," in *Proc. Symp. Ionospheric Effects on Communication and Related Systems Radio Science*, vol. 23, 1988, pp. 713–720.
[45] W. Xu and I. Cumming, "A region growing algorithm for InSAR phase unwrapping," in *Proc. 1996 Int. Geoscience and Remote Sensing Symposium*, Lincoln, NE, 1996, pp. 2044–2046.
[46] T. Flynn, "Consistent 2-D phase unwrapping guided by a quality map," in *Proc. 1996 Int. Geosci. Remote Sensing Symp.*, Lincoln, NE, 1996, pp. 2057–2059.
[47] M. Pritt, "Phase unwrapping by means of multigrid techniques for interferometric SAR," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 728–738, May 1996.
[48] J. Strand and T. Taxt, "Performance evaluation of two-dimensional phase unwrapping algorithms," *Appl. Opt.*, vol. 38, no. 20, pp. 4333–4344, July 1999.

José M. B. Dias (S'87–M'95) received the E.E., M.Sc., and Ph.D. degrees in electrical and computer engineering, all from Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal, in 1985, 1991, and 1995, respectively.

He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, IST. He is also a researcher with the Communication Theory and Pattern Recognition Group of the Institute of Telecommunications. His research interests includes remote sensing, signal and image processing, pattern recognition, and telecommunications.

José M. N. Leitao (M'95) received the E.E. and Ph.D. degrees in electrical engineering, in 1970 and 1983, respectively, both from Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal. He received the "Agregado" degree in electrical and computer engineering, also from IST, in 1992.

He was with the Laboratory of Physiology, Instituto Gulbenkian de Ciência, Oeiras, Portugal, from 1970 to 1972. After spending three years at the University of Tübingen, Germany, he joined the faculty of IST in 1976, where he is currently a full Professor with the Department of Electrical and Computer Engineering. He is also the coordinator of the Communication Theory and Pattern Recognition Group of the Institute of Telecommunications. His main research interests are communication and information theory, pattern recognition, signal and image processing.