Models and algorithms for phase estimation in differential interference contrast microscopy

Simone Rebegoldi, Lola Bautista, Laure Blanc-Féraud, Marco Prato, Luca Zanni, Arturo Plata

Models and algorithms for phase estimation in differential interference contrast microscopy

S Rebegoldi[1], L Bautista[2,3], L Blanc-Féraud[3], M Prato[1], L Zanni[1] and A Plata[2]

[1] Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio Emilia, Via Campi 213/b, 41125 Modena, Italy
[2] Escuela de Ingeniería de Sistemas y Escuela de Física, Universidad Industrial de Santander, 680002 Bucaramanga, Colombia
[3] Université Côte d'Azur, CNRS, Laboratoire I3S, UMR 7271, 06903 Sophia Antipolis, France

E-mail: `simone.rebegoldi@unimore.it`

Abstract. In this paper we address the problem of estimating the phase from color images acquired with differential-interference-contrast microscopy. In particular, we consider the nonlinear and nonconvex optimization problem obtained by regularizing a least-squares-like discrepancy term with a total variation functional, possibly smoothed with the introduction of a positive constant. We deeply investigate the analytical properties of the resulting objective function, proving the existence of minimum points, and several optimization methods able to address the minimization problem. Besides revisiting the conjugate gradient method proposed in the literature for this problem and comparing it with standard conjugate gradient approaches, we introduce more recent effective optimization tools able to obtain both in the smooth and in the nonsmooth case accurate reconstructions with a reduced computational demand.

AMS classification scheme numbers: 65K05, 90C30, 90C90

Keywords: DIC microscopy, phase estimation, nonlinear optimization methods
Submitted to: *Inverse Problems*

1. Introduction

Since their invention, microscopes have been a powerful tool in a variety of disciplines such as biology, medicine and the study of materials. In particular, the branch of optical microscopy (also referred as light microscopy) has been successfully applied in biomedical sciences and cell biology in order to study detailed structures and understand their function in biological specimens. The optical microscope uses visible light for illuminating the object and contains lenses that magnify the image of the object and focus the

light on the retina of the observer's eye [1]. Optical microscopy includes several techniques, such as bright-field, dark-field, phase contrast, differential interference contrast (DIC), fluorescence and confocal microscopy. We refer to the work of Wilson and Bacic [2] for a comparison of the advantages and limitations of these techniques.

The observation of biological structures is a challenging task, especially in live-cell imaging. In fact, optical microscopes are limited by the diffraction of light, and imaging is affected by the optical properties of the object, such as spatial variations in refractive index which introduces aberrations as the light traverses the object [3]. Consequently, since most of the cell components are transparent to visible light [4] and because of the high content of water, traditional light microscopy may suffer from a lack of contrast, reason for which staining is often used to produce contrast by light absorption [1]. Unfortunately, such a process is toxic to living cells and thus it is not suitable for in vivo applications. An alternative solution consists in reducing the condenser numerical aperture, which however worsens dramatically the resolution of the image.

The technique of interest on this paper is DIC microscopy, designed by Allen, David and Nomarski [5] to overcome the inability to image unstained transparent biological specimens, which is typical of bright-field microscopes, while avoiding at the same time the halo artifacts of other techniques designed for the same purpose, such as phase contrast. DIC microscopes are able to provide contrast to images by exploiting the phase shifts in light induced by the transparent specimens (also called phase objects) while passing through them. This phenomenon is not detected by the human eye, neither by an automatic visual system, and occurs because of the interaction of light with different refractive indexes of both the specimen and its surrounding medium. In DIC microscopy, such phase shifts are converted into artificial black and white shadows in the image, which correspond to changes in the spatial gradient of the specimen's optical path length. Furthermore, this technique has been widely recognized by its possibility to use full numerical apertures in the objective, which results in high contrast images at high lateral resolution.

One disadvantage of DIC microscopy is that the observed images cannot be easily used for topographical and morphological interpretation, because the changes in phase of the light are hidden in the intensity image. It is then of vital importance to recover the specimen's phase function from the observed DIC images. The problem of phase estimation in optical imaging has been widely studied, as shown in the review made in [6]. Previous work for reconstructing the DIC phase function has been done by Munster et al [7], who retrieve the phase information by deconvolution with a Wiener filter; line integration of DIC images is proposed by Kam in [8], supposing that the line integration along the shear angle yields a positive-definite image, which is not always the case since the intensity image is a nonlinear relation between the transmission function of the specimen and the point spread function of the microscope. Kou et al [9] introduce

the use of transport of intensity equation to retrieve the phase function; Bostan et al [10] also use this approach, including a total variation regularization term to preserve the phase transitions. Finally, in the work of Preza [11, 12, 13, 14], the phase estimation in DIC microscopy has been addressed by considering the minimization of a Tikhonov regularized discrepancy term, which is performed by means of a modified nonlinear conjugate gradient (CG) method.

In this work, we are interested in reconstructing the phase by minimization of a penalized least-squares term as proposed in [13]. Firstly, we introduce regularization penalties able to reconstruct both sharp and smooth variations of the unknown phase, as the total variation (TV) functional and its smooth generalization; secondly, we deepen the analysis of the optimization methods used. To this aim, we revisit the modified nonlinear CG algorithm described in [13, 14], highlight possible drawbacks due to the lack of the strong Wolfe conditions and compare the performances with standard nonlinear Fletcher-Reeves (FR) and Polak-Ribière (PR) methods. Moreover, we compare the CG approaches with a standard gradient method, in which suitable adaptive steplength parameters are chosen to improve the convergence rate of the algorithm. Finally, we introduce a recently proposed linesearch–based forward–backward method able to address the nonsmoothness of the TV functional [15], which reduces to a standard gradient method if a smooth TV regularizer is used. We are also interested in extending the one color acquisition to polychromatic ones, trying to improve the reconstruction.

The organization of the paper is as follows. In Section 2, the DIC system for transmitted coherent light is described, together with the corresponding polychromatic image formation model. Furthermore the nonlinear inverse problem of the phase reconstruction and its corresponding optimization problem are presented, remarking some analytical properties of the objective function, such as the existence of minimum points. In Section 3 the iterative optimization algorithms designed to address the phase reconstruction problem are detailed. In Section 4 numerical simulations on synthetic images are presented in order to evaluate efficiency and robustness of the considered approaches. Conclusions and perspectives are included in Section 5.

2. Model and problem formulation

2.1. The DIC system

DIC microscopy works under the principle of dual-beam interference of polarized light, as depicted in Figure 1. Coherent light coming from a source is passed through a polarizer lens. Every incident ray of polarized light is splitted by a Nomarski prism placed at the front focal plane of the condenser. This splitting produces two wave components – ordinary and extraordinary – such that their corresponding electromagnetic fields are orthogonal and separated at a fixed shear distance $2\Delta x$ along a specific shear direction,

whose angle τ_k formed with the x-axis is denominated shear angle. The specimen is sampled by the pair of waves; if they pass through a region where there is a gradient in the refractive index, the waves will be differentially shifted in phase. After this, they will reach a second Normarski prism placed at the back focal plane of the objective lens. This prism introduces an additional phase shift, called the bias retardation and indicated with $2\Delta\theta$, which helps to improve the contrast of the observed image and to give the shadow-cast effect characteristic of DIC images (see Figure 2). The interference of the two sheared and phase shifted waves occurs inside this prism and, thus, the two waves are recombined into a single beam that goes through a second polarizer lens called the analyzer. For further details on the DIC working principle we refer to the work of Murphy [16] and Mehta et al [17]. As explained in [12] and references therein, for a given shear angle, one acquisition is related to the directional derivative of the object along the direction τ_k. Then, in order to retrieve the optical path length of the specimen, one acquires at least two images with a shear angle difference of $\pi/2$ [13].

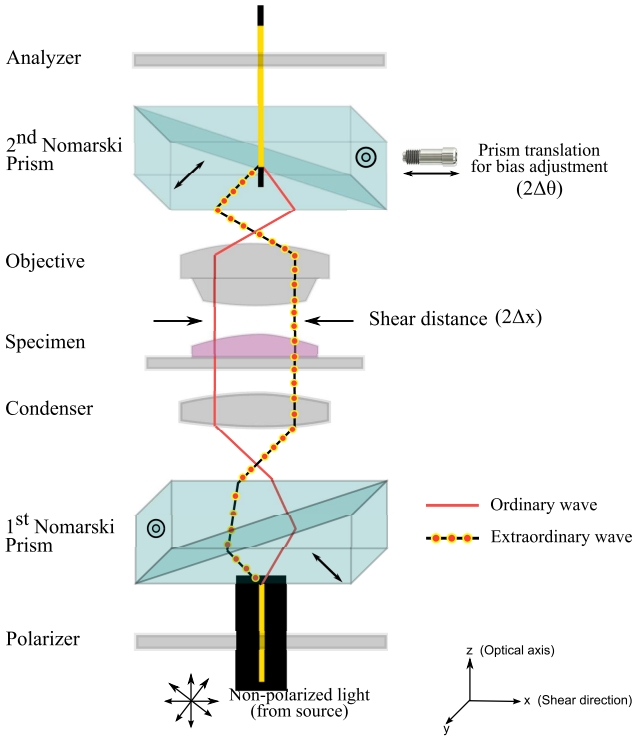
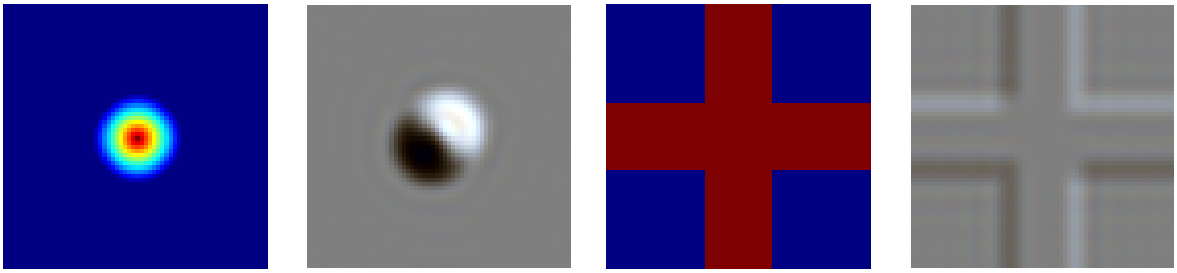


Figure 1. Transmitted-light Nomarski DIC microscope. The difference of colors of the ordinary and extraordinary waves indicates that their electromagnetic fields are orthogonal to each other.

The observed images will have a uniform gray background on regions where there are no changes in the optical path, whereas they will have dark shadows and bright highlights where there are phase gradients in the direction of shear, having a 3-D relief-like appearance (see Figure 2). It is important to note that the shadows and highlights indicate the signs and slope of phase gradients in the specimen, and not necessarily



(a) (b) (c) (d)

Figure 2. Phase functions of two phantom specimens and corresponding noiseless DIC color images: (a) phase function of the "cone" object, (b) DIC image of the cone, (c) phase function of the "cross" object, (d) DIC image of the cross. The images have been computed by using model (1) and setting the shear to $2\Delta x = 0.6$ μm, the bias to $2\Delta\theta = \pi/2$ rad and the shear angle to $\tau = \pi/4$ rad.

indicate high or low spots [5].

In this paper we consider the polychromatic rotational-diversity model [18], which is an extension of the model presented in [13] to color image acquisition. This model assumes that K color images are acquired by rotating the specimen K times with respect to the shear axis, which results in K rotations of the amplitude point spread function. In this configuration, the relation between the acquired images and the unknown true phase ϕ is given by

$$(o_{k,\lambda_\ell})_j = a_1 \left|(h_{k,\lambda_\ell} \otimes e^{-i\phi/\lambda_\ell})_j\right|^2 + (\eta_{k,\lambda_\ell})_j, \tag{1}$$

for $k = 1, \ldots, K$, $\ell = 1, 2, 3$, $j \in \chi$, where

- k is the index of the angles τ_k that the shear direction makes with the horizontal axis [13], ℓ is the index denoting one of the three RGB channels and $j = (j_1, j_2)$ is a 2D–index varying in the set $\chi = \{1, \ldots, M\} \times \{1, \ldots, P\}$;
- λ_ℓ is the $\ell-$th illumination wavelength, being $\lambda_1 = 0.65$, $\lambda_2 = 0.55$ and $\lambda_3 = 0.45$ the values for the red, green and blue wavelengths respectively;
- $o_{k,\lambda_\ell} \in \mathbb{R}^{MP}$ is the $\ell-$th color component of the $k-$th discrete observed image $o_k = (o_{k,\lambda_1}, o_{k,\lambda_2}, o_{k,\lambda_3}) \in \mathbb{R}^{MP\times 3}$;
- $\phi \in \mathbb{R}^{MP}$ is the unknown phase vector and $e^{-i\phi/\lambda_\ell} \in \mathbb{C}^{MP}$ stands for the vector defined by $(e^{-i\phi/\lambda_\ell})_j = e^{-i\phi_j/\lambda_\ell}$;
- $h_{k,\lambda_\ell} \in \mathbb{C}^{MP}$ is the discretization of the continuous DIC point spread function [12, 19] corresponding to the illumination wavelength λ_ℓ and rotated by the angle τ_k , i.e.,

$$h_{k,\lambda_\ell}(x, y) = \frac{1}{2}\left[e^{-i\Delta\theta} p_{\lambda_\ell}\left(R_k \cdot (x - \Delta x, y)^T\right) - e^{i\Delta\theta} p_{\lambda_\ell}\left(R_k \cdot (x + \Delta x, y)^T\right)\right], \tag{2}$$

where $p_{\lambda_\ell}(x, y)$ is the coherent PSF of the microscope's objective lens for the wavelength λ_ℓ, which is given by the inverse Fourier transform of the disk support

function of amplitude 1 and radius equal to the cutoff frequency $f_c = NA/\lambda_\ell$ [12], being NA the numerical aperture of the objective lens, $2\Delta\theta$ is the DIC bias retardation, $2\Delta x$ is the shear distance and R_k is the rotation matrix which rotates the coordinates according to the shear angle τ_k;

- $h_1 \otimes h_2$ denotes the 2D convolution between the two $M \times P$ images h_1, h_2, extended with periodic boundary conditions;
- $\eta_{k,\lambda_\ell} \in \mathbb{R}^{MP}$ is the noise corrupting the data, which is assumed to be a realization of a Gaussian random vector with mean $\mathbf{0} \in \mathbb{R}^{MP}$ and covariance matrix $\sigma^2 I_{(MP)^2}$, where $I_{(MP)^2}$ is the identity matrix of size $(MP)^2$;
- $a_1 \in \mathbb{R}$ is a constant which corresponds to closing the condenser aperture down to a single point.

2.2. Optimization problem

The phase reconstruction problem consists in finding an approximation of the unknown phase vector ϕ from the observed RGB images $o_1, \ldots, o_K$. Let us first address this problem by means of the maximum likelihood (ML) approach. Since the $3K$ images o_{k,λ_ℓ} are corrupted by Gaussian noise, then the negative log likelihood of each image is a least-squares measure, which is nonlinear due to the presence of the exponential in (1). If we assume white Gaussian noise, statistically independent of the data, the negative log likelihood of the problem is the sum of the negative log likelihoods of the different images, namely the following fit-to-data term

$$J_0(\phi) = \sum_{\ell=1}^{3} \sum_{k=1}^{K} \sum_{j \in \chi} \left[(o_{k,\lambda_\ell})_j - a_1 \left| (h_{k,\lambda_\ell} \otimes e^{-i\phi/\lambda_\ell})_j \right|^2 \right]^2. \tag{3}$$

Then the ML approach to the phase reconstruction problem consists in the minimization of the function in (3):

$$\min_{\phi \in \mathbb{R}^{MP}} J_0(\phi). \tag{4}$$

In the next result, we collect some properties of J_0 that will be useful hereafter.

Lemma 1 *Let $J_0 : \mathbb{R}^{MP} \to \mathbb{R}$ be defined as in (3). Then we have the following:*

(i) There exists $T > 0$ such that J_0 is periodic of period T with respect to each variable, i.e. for any $j \in \chi$, defining $\boldsymbol{e}_j = (\delta_{j,r})_{r \in \chi} = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^{MP}$ where $\delta_{j,r}$ is the Kronecker delta, it holds

$$J_0(\phi + T\boldsymbol{e}_j) = J_0(\phi), \quad \forall\ \phi \in \mathbb{R}^{MP}. \tag{5}$$

(ii) $J_0(\phi + c\mathbf{1}) = J_0(\phi)$, $\forall\ c \in \mathbb{R}$, where $\mathbf{1} \in \mathbb{R}^{MP}$ is the vector of all ones.

(iii) J_0 is an analytic function on $\mathbb{R}^{MP}$ and therefore $J_0 \in C^\infty(\mathbb{R}^{MP})$.

Proof. (i) Fix $j \in \chi$, $\ell \in \{1, 2, 3\}$ and consider the exponential in (3). Then for all $r \in \chi$

$$\left(e^{-i(\phi + 2\pi\lambda_\ell \boldsymbol{e}_j)/\lambda_\ell}\right)_r = \begin{cases} e^{-i\phi_r/\lambda_\ell} & ,\ r \neq j \\ e^{-i[(\phi_j/\lambda_\ell) + 2\pi]} = e^{-i\phi_r/\lambda_\ell} & ,\ r = j \end{cases} = (e^{-i\phi/\lambda_\ell})_r, \tag{6}$$

where the equality inside the curly bracket is due to the periodicity of the complex exponential. Then, for a fixed $\ell \in \{1,2,3\}$, the expression given in (3) without the sum in ℓ is $2\pi\lambda_\ell$ periodic w.r.t. the variable ϕ_j. This means that J_0 is the sum of three periodic functions of variable ϕ_j whose periods are $2\pi\lambda_1$, $2\pi\lambda_2$ and $2\pi\lambda_3$ respectively. By recalling that the sum of two periodic functions is periodic if the ratio of the periods is a rational number, we can conclude that J_0 is periodic, as we have $\frac{\lambda_\ell}{\lambda_{\ell'}}$ rational for all $\ell, \ell' \in \{1,2,3\}$.

(ii) Set $J_{\ell,k,j}(\phi) = \left|(h_{k,\lambda_\ell} \otimes e^{-i\phi/\lambda_\ell})_j\right|^2 = \left|\sum_{r\in\chi}(h_{k,\lambda_\ell})_r e^{-i(\phi_{j-r})/\lambda_\ell}\right|^2$. If the thesis holds for $J_{\ell,k,j}$, then it holds also for J_0. We have

$$\begin{aligned} J_{\ell,k,j}(\phi + c\mathbf{1}) &= \left|\sum_{r\in\chi}(h_{k,\lambda_\ell})_r e^{-i(\phi_{j-r}+c)/\lambda_\ell}\right|^2 = \left|e^{-ic/\lambda_\ell}\sum_{r\in\chi}(h_{k,\lambda_\ell})_r e^{-i(\phi_{j-r})/\lambda_\ell}\right|^2 \\ &= \left|e^{-ic/\lambda_\ell}\right|^2 \left|\sum_{r\in\chi}(h_{k,\lambda_\ell})_r e^{-i(\phi_{j-r})/\lambda_\ell}\right|^2 = J_{\ell,k,j}(\phi). \end{aligned} \tag{7}$$

(iii) If $J_{\ell,k,j}$ is an analytic function on $\mathbb{R}^{MP}$, then J_0 is given by sums and compositions of analytic functions and thus it is itself analytic [20, Propositions 1.6.2 and 1.6.7]. Hence we focus on $J_{\ell,k,j}$. Since $(h_{k,\lambda_\ell})_r \in \mathbb{C}$, it can be expressed in its trigonometric form $(h_{k,\lambda_\ell})_r = \rho_r e^{i\theta_r}$, with $\rho_r \in \mathbb{R}_{\geq 0}$, $\theta_r \in [0, 2\pi)$. Then we can rewrite $J_{\ell,k,j}$ as follows

$$\begin{aligned} J_{\ell,k,j}(\phi) &= \left|\sum_{r\in\chi}\rho_r e^{i[\theta_r - (\phi_{j-r}/\lambda_\ell)]}\right|^2 = \\ &= \left|\sum_{r\in\chi}\rho_r\cos(\theta_r - (\phi_{j-r}/\lambda_\ell)) + i\sum_{r\in\chi}\rho_r\sin(\theta_r - (\phi_{j-r}/\lambda_\ell))\right|^2 = \\ &= \left(\sum_{r\in\chi}\rho_r\cos(\theta_r - (\phi_{j-r}/\lambda_\ell))\right)^2 + \left(\sum_{r\in\chi}\rho_r\sin(\theta_r - (\phi_{j-r}/\lambda_\ell))\right)^2. \end{aligned}$$

We now observe that the function $J_{\ell,k,j}$ contains $\sin(\theta_r - (\phi_{j-r}/\lambda_\ell))$ and $\cos(\theta_r - (\phi_{j-r}/\lambda_\ell))$, which are both analytic functions with respect to the single variable ϕ_{j-r} and thus also with respect to ϕ, and the square function $(\cdot)^2$, which is also analytic. Since $J_{\ell,k,j}$ is given by sums and compositions of these functions, it is analytic. □

Problem (4) admits infinitely many solutions, as stated in the following theorem.

Theorem 1 *J_0 admits at least one global minimum point. Furthermore, if $\psi \in \mathbb{R}^{MP}$ is a global minimizer of J_0, then also $\{\psi + c\mathbf{1} : c \in \mathbb{R}\} \cup \{\psi + mT\boldsymbol{e}_j : j \in \chi,\ m \in \mathbb{Z}\}$ are global minimizers of J_0.*

Proof. Let $\Omega = [0,T]^{MP} \subset \mathbb{R}^{MP}$. Point (iii) of Lemma 1 ensures that J_0 is continuous on Ω, thus from the extreme value theorem J_0 admits at least one minimum point ψ on Ω. By contradiction, assume that there exists $\phi \in \mathbb{R}^{MP} \setminus \Omega$ such that $J_0(\phi) < J_0(\psi)$.

Let $I \subset \chi$ be the subset of indices such that $\{\phi_s\}_{s\in I}$ is the set of all components of ϕ which belong to $\mathbb{R} \setminus [0, T]$ and $\{m_s\}_{s\in I} \subset \mathbb{Z} \setminus \{1\}$ is the set of integers such that $\phi_s \in [(m_s - 1)T, m_s T]$. Define $\bar{\phi} = \phi - \sum_{s\in I}(m_s - 1)T\boldsymbol{e}_s \in \Omega$. By periodicity of J_0 w.r.t. the variables ϕ_s, $s \in I$, we obtain

$$J_0(\bar{\phi}) = J_0(\phi) < J_0(\psi). \tag{8}$$

Therefore, we have found a point $\bar{\phi} \in \Omega$ such that $J_0(\bar{\phi}) < J_0(\psi)$, where ψ is a minimum point on Ω. This is absurd, hence ψ is a global minimizer for J_0. The second part of the thesis follows from points (i)-(ii) of Lemma 1. □

Theorem 1 asserts that the solution to problem (4) is not unique and it may be determined only up to an unknown real constant or to multiples of the period T w.r.t. any variable ϕ_j. Furthermore, since J_0 is periodic, it is a nonconvex function of the phase ϕ, thus it may admit several local minima as well as saddle points. In the light of these considerations, we can conclude that (4) is a severely ill-posed problem, which requires regularization in order to impose some a priori knowledge on the unknown phase. In particular, we propose to solve the following regularized optimization problem

$$\min_{\phi\in\mathbb{R}^{MP}} J(\phi) \equiv J_0(\phi) + J_{TV}(\phi), \tag{9}$$

where J_0 is the least-squares distance defined in (3) and J_{TV} is the smooth total variation functional (also known as hypersurface potential - HS) defined as [21, 22]

$$J_{TV}(\phi) = \mu \sum_{j\in\chi} \sqrt{((\mathcal{D}\phi)_j)_1^2 + ((\mathcal{D}\phi)_j)_2^2 + \delta^2}, \tag{10}$$

where $\mu > 0$ is a regularization parameter, the discrete gradient operator $\mathcal{D} : \mathbb{R}^{MP} \longrightarrow \mathbb{R}^{2MP}$ is set through the standard finite difference with periodic boundary conditions

$$(\mathcal{D}\phi)_{j_1,j_2} = \begin{pmatrix} ((\mathcal{D}\phi)_{j_1,j_2})_1 \\ ((\mathcal{D}\phi)_{j_1,j_2})_2 \end{pmatrix} = \begin{pmatrix} \phi_{j_1+1,j_2} - \phi_{j_1,j_2} \\ \phi_{j_1,j_2+1} - \phi_{j_1,j_2} \end{pmatrix}, \ \phi_{M+1,j_2} = \phi_{1,j_2}, \ \phi_{j_1,P+1} = \phi_{j_1,1}$$

and the additional parameter $\delta \geq 0$ plays the role of a threshold for the gradient of the phase. Obviously J_{TV} reduces to the standard TV functional [23] by setting $\delta = 0$. The choice of this kind of regularization term instead of the first-order Tikhonov one used e.g. in [13, 14] lies in the capability of the HS regularizer to behave both as a Tikhonov-like regularization in regions where the gradient assumes small values (w.r.t. δ), and as an edge-preserving regularizer in regions where the gradient is very large, as it happens in the neighborhood of jumps in the values of the phase.

Problem (9) is still a difficult nonconvex optimization problem and, when $\delta = 0$, it is also nondifferentiable. Some properties of the objective function J are now reported.

Lemma 2 *Let $J : \mathbb{R}^{MP} \to \mathbb{R}$ be defined as in (9). Then:*

(i) $J(\phi + c\mathbf{1}) = J(\phi)$, $\forall\ c \in \mathbb{R}$.

(ii) If $\delta > 0$, then $J \in C^\infty(\mathbb{R}^{MP})$ and ∇J is Lipschitz continuous, namely there exists $L > 0$ such that

$$\|\nabla J(\phi) - \nabla J(\psi)\|_2 \leq L\|\phi - \psi\|_2, \quad \forall \phi, \psi \in \mathbb{R}^{MP}. \tag{11}$$

Proof. (i) We have already proved in point (ii) of Lemma 1 that the property holds for J_0. Since it is immediate to check that $(\mathcal{D}(\phi + c\mathbf{1}))_{j_1,j_2} = (\mathcal{D}\phi)_{j_1,j_2}$, the property is true also for J_{TV} and thus for J.
(ii) Point (iii) of Lemma 1 states that $J_0 \in C^\infty(\mathbb{R}^{MP})$ and the same property holds for J_{TV}, hence J is the sum of two $C^\infty(\mathbb{R}^{MP})$ functions.
It is known that ∇J_{TV} is $L_{TV}-$Lipschitz continuous with $L_{TV} = 8\mu/\delta^2$ [24]. We prove that also ∇J_0 is Lipschitz continuous. If we introduce the residual image $r_{k,\lambda_\ell} = \left|(h_{k,\lambda_\ell} \otimes e^{-i\phi/\lambda_\ell})\right|^2 - o_{k,\lambda_\ell}$ and fix $s \in \chi$, the partial derivative of J_0 with respect to ϕ_s is given by

$$\frac{\partial J_0(\phi)}{\partial \phi_s} = \sum_{\ell=1}^{3}\sum_{k=1}^{K}\sum_{j\in\chi} \frac{4}{\lambda_\ell}(r_{k,\lambda_\ell})_j \mathrm{Im}\Big\{e^{-i\phi_s/\lambda_\ell}(h_{k,\lambda_\ell})_{j-s}\overline{(h_{k,\lambda_\ell} \otimes e^{-i\phi/\lambda_\ell})_j}\Big\}, \tag{12}$$

where $\mathrm{Im}(\cdot)$ denotes the imaginary part of a complex number. As concerns the entries of the Hessian $\nabla^2 J_0$, the second derivative w.r.t. ϕ_s, ϕ_t $(s, t \in \chi)$ is given by

$$\frac{\partial^2 J_0(\phi)}{\partial \phi_t \partial \phi_s} = 4\sum_{\ell=1}^{3}\sum_{k=1}^{K}\sum_{j\in\chi} \frac{2}{\lambda_\ell^2}\mathrm{Im}\{\vartheta_s\}\,\mathrm{Im}\{\vartheta_t\} + \frac{(r_{k,\lambda_\ell})_j}{\lambda_\ell^2}\mathrm{Re}\Big\{e^{i(\phi_t-\phi_s)/\lambda_\ell}(h_{k,\lambda_\ell})_{j-s}\overline{(h_{k,\lambda_\ell})_{j-t}} - \delta_{s,t}\vartheta_s\Big\}, \tag{13}$$

where $\vartheta_p = e^{-i\phi_p/\lambda_\ell}(h_{k,\lambda_\ell})_{j-p}\overline{(h_{k,\lambda_\ell} \otimes e^{-i\phi/\lambda_\ell})_j}$ $(p \in \chi)$, $\mathrm{Re}(\cdot)$ denotes the real part of a complex number and $\delta_{s,t}$ is the Kronecker delta. By using the triangle inequality and the fact that $|e^{-i\phi_r/\lambda_\ell}| = 1$, the following inequality hold:

$$|\vartheta_p| \leq |(h_{k,\lambda_\ell})_{j-p}| \sum_{r\in\chi} |(h_{k,\lambda_\ell})_r|. \tag{14}$$

By applying the triangle inequality, the fact that $|e^{-i\phi_r/\lambda_\ell}| = 1$, $|\mathrm{Im}(z)| \leq |z|$ and $|\mathrm{Re}(z)| \leq |z|$ for any $z \in \mathbb{C}$ and inequality (14) to (13), we obtain the following bound on the second derivative of J_0:

$$\left|\frac{\partial^2 J_0(\phi)}{\partial \phi_t \partial \phi_s}\right| \leq 4\sum_{\ell=1}^{3}\sum_{k=1}^{K}\sum_{j\in\chi} \frac{2}{\lambda_\ell^2}|(h_{k,\lambda_\ell})_{j-s}||(h_{k,\lambda_\ell})_{j-t}| \left(\sum_{r\in\chi}|(h_{k,\lambda_\ell})_r|\right)^2 + \frac{|(r_{k,\lambda_\ell})_j|}{\lambda_\ell^2}\left\{|(h_{k,\lambda_\ell})_{j-s}||(h_{k,\lambda_\ell})_{j-t}| + |(h_{k,\lambda_\ell})_{j-s}|\sum_{r\in\chi}|(h_{k,\lambda_\ell})_r|\right\}. \tag{15}$$

Set $H_{k,\ell} = \sum_{r\in\chi}|(h_{k,\lambda_\ell})_r|$. Taking the sum of (15) over $s \in \chi$ and picking the maximum over $t \in \chi$, a bound on the $\ell_\infty-$norm of the Hessian $\nabla^2 J_0$ is obtained:

$$\|\nabla^2 J_0(\phi)\|_\infty = \max_{t\in\chi}\sum_{s\in\chi}\left|\frac{\partial^2 J_0(\phi)}{\partial \phi_t \partial \phi_s}\right|$$

$$\leq 4\sum_{\ell=1}^{3}\sum_{k=1}^{K}\sum_{j\in\chi}\frac{H_{k,\ell}}{\lambda_\ell^2}\left\{2\max_{t\in\chi}|(h_{k,\lambda_\ell})_{j-t}|H_{k,\ell}^2+|(r_{k,\lambda_\ell})_j|\left[\max_{t\in\chi}|(h_{k,\lambda_\ell})_{j-t}|+H_{k,\ell}\right]\right\}$$
$$= L_0, \quad \forall\ \phi\in\mathbb{R}^{MP}.$$

From relation $\|A\|_2 \leq \sqrt{\|A\|_1\|A\|_\infty}$ and the fact that $\|\nabla^2 J(\phi)\|_1 = \|\nabla^2 J(\phi)\|_\infty$ ($\nabla^2 J_0(\phi)$ is a symmetric matrix), it follows that $\|\nabla^2 J_0(\phi)\|_2 \leq L_0$ for all $\phi\in\mathbb{R}^{MP}$. Fix $\phi,\psi\in\mathbb{R}^{MP}$. By the mean value theorem for vector-valued functions, we have

$$\|\nabla J_0(\phi)-\nabla J_0(\psi)\|_2 \leq \sup_{\theta\in(0,1)}\|\nabla^2 J_0(\psi+\theta(\phi-\psi))\|_2\|\phi-\psi\|_2 \leq L_0\|\phi-\psi\|_2. \tag{16}$$

Then ∇J_0 is L_0-Lipschitz continuous and consequently also ∇J is Lipschitz continuous with constant $L = L_0 + L_{TV}$. □

Point (i) of Lemma 2 makes clear that, if a solution to problem (9) exists, then it is not unique and it can be determined only up to a real constant. This is a common feature shared with the unregularized problem (4). However, unlike in (4), the objective function J is not periodic and, in addition, none of the two terms J_0 and J_{TV} are coercive, therefore we can not prove the existence of a minimum point of J neither as in Theorem 1 nor by coercivity. A specific proof of existence of the solution for problem (9) is now presented.

Theorem 2 *The objective function J admits at least one global minimum point. Furthermore, if $\psi\in\mathbb{R}^{MP}$ is a global minimizer of J, then also $\{\psi + c\mathbf{1} :\ c\in\mathbb{R}\}$ are global minimizers of J.*

Proof. Let $S=\{\phi\in\mathbb{R}^{MP} :\ \phi = c\mathbf{1},\ c\in\mathbb{R}\}$ be the line in $\mathbb{R}^{MP}$ of all constant images and Π any hyperplane intersecting S in one point ϕ_S, i.e.

$$\Pi = \{\phi\in\mathbb{R}^{MP} : \sum_{r\in\chi} a_r\phi_r + b = 0\},\ \sum_{r\in\chi} a_r \neq 0,\ b\in\mathbb{R}. \tag{17}$$

Thanks to part (i) of Lemma 2, for any $\phi\in\mathbb{R}^{MP}$ the point $\phi_\Pi = \phi - \left(\frac{\sum_r a_r\phi_r + b}{\sum_r a_r}\right)\mathbf{1}\in\Pi$ is such that $J(\phi_\Pi) = J(\phi)$. Consequently, if ψ is a minimum point of J on Π, then it is also a minimum point on $\mathbb{R}^{MP}$, because $J(\psi)\leq J(\phi_\Pi) = J(\phi)$ for all $\phi\in\mathbb{R}^{MP}$. Hence we restrict the search of the minimum point on Π and we denote with $J|_\Pi$ the restriction of J to Π. Since $S = \arg\min_{\phi\in\mathbb{R}^{MP}} J_{TV}(\phi)$ and Π intersects S only in ϕ_S, J_{TV} is a convex function with a unique minimum point on Π, which implies that J_{TV} is coercive on Π. Furthermore, being J_0 periodic and continuous, it is a bounded function on Π. Then $J|_\Pi$ is the sum of a coercive term and a bounded one, therefore it is itself coercive. This allows to conclude that J admits a minimum point on Π and thus also on $\mathbb{R}^{MP}$. The second part of the thesis follows from Lemma 2, part (i). □

Note that the above proof of existence holds also for the regularized DIC problem proposed in [13, 14], in which the Tikhonov-like regularizer used instead of the TV functional is also noncoercive.

3. Optimization methods

In previous works [13, 14, 11], the problem of DIC phase reconstruction had been addressed with the nonlinear conjugate gradient method [25]. However, as better explained in Subsection 3.2, these methods require in practice several evaluations of the objective function and possibly its gradient in order to compute the linesearch parameter. What we propose instead is to tackle problem (9) with a gradient descent algorithm in the differentiable case ($\delta > 0$) and a proximal-gradient method in the nondifferentiable case ($\delta = 0$). The key ingredients of both methods are the use of an Armijo linesearch at each iteration, which ensures convergence to a stationary point of problem (9), and a clever adaptive choice of the steplength in order to improve the speed of convergence.
For sake of simplicity, from now on we assume that each monochromatic image is treated as a vector in $\mathbb{R}^N$ (being $N = MP$) obtained by a lexicographic reordering of its pixels.

3.1. Gradient and proximal-gradient methods: LMSD and ILA

In this subsection we describe the two proposed algorithms to address problem (9) for both cases $\delta > 0$ and $\delta = 0$. In the former case the objective function is differentiable and we exploit the limited memory steepest descent (LMSD) method proposed by Fletcher [26] and outlined in Algorithm 1. The LMSD method is a standard gradient method equipped with a monotone Armijo linesearch and variable steplengths approximating the inverse of some eigenvalues of the Hessian matrix $\nabla^2 J(\phi^{(n)})$ in order to improve the convergence speed. Unlike the classical Barzilai–Borwein (BB) rules [27] and its generalizations (see e.g. [28, 29, 30]) which tries to approximate $(\nabla^2 J(\phi^{(n)}))^{-1}$ with a constant diagonal matrix, the idea proposed by Fletcher for quadratic objective functions is based on a Lanczos iterative process applied to the Hessian matrix of the objective function. Some algebra shows that this can be practically performed without the explicit knowledge of the Hessian itself but exploiting only a set of back gradients and steplengths (see steps 6–10 of Algorithm 1). Generalization to nonquadratic functions can be obtained by computing the eigenvalues of the matrix $\widetilde{\Phi}$ in step 10 instead of Φ (we remark that for quadratic J the two matrices coincide).
Some practical issues have to be addressed in the implementation of Algorithm 1:

- The first loop (step 1 to 5) build a matrix

  $$G = \left[\nabla J(\phi^{(n-m)}) \;\; \nabla J(\phi^{(n-m+1)}) \ldots \nabla J(\phi^{(n-1)})\right]$$

  of size $MP \times m$. The initial values for the first m steplengths can be provided by the user (e.g. by computing the BB ones) or can be chosen with the same approach described in steps 6–10 but with smaller matrices. For example, one can fix $\alpha_0^{(0)}$, compute $G = \nabla J(\phi^{(0)})$ and use steps 6–10 to compute $\alpha_1^{(0)}$. At this point, defining $G = [\nabla J(\phi^{(0)}) \;\; \nabla J(\phi^{(1)})]$ one can compute $\alpha_2^{(0)}$ and $\alpha_3^{(0)}$ and repeat the procedure until a whole set of m back gradients is available.

Algorithm 1 Limited memory steepest descent (LMSD) method

Choose $\rho, \omega \in (0,1)$, $m \in \mathbb{N}_{>0}$, $\alpha_0^{(0)}, \ldots, \alpha_{m-1}^{(0)} > 0$, $\phi^{(0)} \in \mathbb{R}^N$ and set $n = 0$.
While True

For $l = 1, \ldots, m$

1. Define $G(:, l) = \nabla J(\phi^{(n)})$.
2. Compute the smallest non-negative integer i_n such that $\alpha_n = \alpha_n^{(0)} \rho^{i_n}$ satisfies
$$J(\phi^{(n)} - \alpha_n \nabla J(\phi^{(n)})) \leq J(\phi^{(n)}) - \omega \alpha_n \|\nabla J(\phi^{(n)})\|^2. \tag{18}$$
3. Compute $\phi^{(n+1)} = \phi^{(n)} - \alpha_n \nabla J(\phi^{(n)})$.
If "Stopping Criterion" is satisfied
4. Return
Else
5. Set $n = n + 1$.
EndIf

EndFor

6. Define the $(m+1) \times m$ matrix $\Gamma = \begin{bmatrix} \alpha_{n-m}^{-1} & & \\ -\alpha_{n-m}^{-1} & \ddots & \\ & \ddots & \alpha_{n-1}^{-1} \\ & & -\alpha_{n-1}^{-1} \end{bmatrix}$.
7. Compute the Cholesky factorization $R^T R$ of the $m \times m$ matrix $G^T G$.
8. Solve the linear system $R^T r = G^T \nabla J(\phi^{(n)})$.
9. Define the $m \times m$ matrix $\Phi = [R, r] \Gamma R^{-1}$.
10. Compute the eigenvalues $\theta_1, \ldots, \theta_m$ of the symmetric and tridiagonal approximation $\widetilde{\Phi}$ of Φ defined as
$$\widetilde{\Phi} = \mathrm{diag}(\Phi) + \mathrm{tril}(\Phi, -1) + \mathrm{tril}(\Phi, -1)^T,$$
being $\mathrm{diag}(\cdot)$ and $\mathrm{tril}(\cdot, -1)$ the diagonal and the strictly lower triangular parts of a matrix.
11. Define $\alpha_{n+i-1}^{(0)} = 1/\theta_i$, $i = 1, \ldots, m$.

EndWhile

- The same procedure can be adopted when step 10 provides only $m' < m$ positive eigenvalues. In this case, all columns of G are discarded, G becomes the empty matrix and the algorithm proceeds with m' instead of m until a whole set of m back gradients is computed. If $m' = 0$, a set of m "safeguard" steplengths, corresponding to the last set of m positive steplengths values provided by step 10, is exploited for the next m iterations.
- If $G^T G$ in step 7 is not positive definite, then the oldest gradient of G is discarded and a new matrix $G^T G$ is computed. This step is repeated until $G^T G$ becomes

positive definite.

- The stopping criterion can be chosen by the user and be related to the decrease of J or to the distance between two successive iterates. In our tests we decided to arrest the iterations when the norm of the gradient ∇J goes below a given threshold κ:

$$\|\nabla J(\phi^{(n)})\| \leq \kappa. \tag{19}$$

Concerning the computational costs of LMSD, the heaviest tasks at each iteration are the computation of $\nabla J(\phi^{(n)})$ at step 1 and $J(\phi^{(n)} - \alpha_n \nabla J(\phi^{(n)}))$ at step 2. Considering step 1, we focus on ∇J_0. As it is written in (12), due to the product between $e^{-i\phi_s/\lambda_\ell}$ and $(h_{k,\lambda_\ell})_{j-s}$, ∇J_0 can be performed with $\mathcal{O}(N^2)$ complexity; this is how the gradient is computed in [13]. However, if we take the sum over j of the residuals into the argument of $\mathrm{Im}(\cdot)$, then we can conveniently rewrite (12) as

$$\frac{\partial J_0(\phi)}{\partial \phi_s} = \sum_{\ell=1}^{3} \sum_{k=1}^{K} \frac{4}{\lambda_\ell} \mathrm{Im}\left\{ \left((r_{k,\lambda_\ell} . * (\overline{h_{k,\lambda_\ell}} \otimes e^{i\phi/\lambda_\ell})) \otimes \tilde{h}_{k,\lambda_\ell} \right)_s e^{-i\phi_s/\lambda_\ell} \right\}, \tag{20}$$

where $h_1 . * h_2$ denotes the componentwise product between two images h_1, h_2 and $(\tilde{h}_{k,\lambda_\ell})_j = (h_{k,\lambda_\ell})_{-j}$ for all $j \in \chi$. Then the heaviest operations in (20) are the two convolutions which, thanks to the assumption of periodic boundary conditions, can be performed with a FFT/IFFT pair ($\mathcal{O}(N \log N)$) complexity). Hence, since ∇J_{TV} has $\mathcal{O}(N)$ complexity, we can conclude that step 1 has an overall complexity of $\mathcal{O}(N \log N)$. Similarly, the function at step 2 is computed with complexity $\mathcal{O}(N \log N)$, due to the presence of one convolution inside the triple sum in (3).

From a practical point of view, we have already shown that the LMSD method is an effective tool for DIC imaging, especially if compared to more standard gradient methods equipped with the BB rules [18]. From a mathematical point of view, one can prove, in the same way as in [31], that every limit point of the sequence generated by Algorithm 1 is a stationary point for problem (9). In addition, the convergence of Algorithm 1 can be asserted whenever the objective function J satisfies the Kurdyka–Łojasiewicz (KL) property [32, 33] at each point of its domain. More precisely, as shown in a number of recent papers [34, 35, 36], one can prove the convergence of a sequence $\{\phi^{(n)}\}_{n\in\mathbb{N}}$ to a limit point (if any exists) which is stationary for J if the following three conditions are satisfied:

$(H1)$ $\exists\ a > 0 :\ J(\phi^{(n+1)}) + a\|\phi^{(n+1)} - \phi^{(n)}\|^2 \leq J(\phi^{(n)})$

$(H2)$ $\exists\ b > 0 :\ \|\nabla J(\phi^{(n+1)})\| \leq b\|\phi^{(n+1)} - \phi^{(n)}\|$

$(H3)$ J satisfies the KL property.

This scheme applies to the LMSD method. First of all, condition $(H3)$ is satisfied for the DIC functional defined in (9). Indeed J_0 is an analytic function (Lemma 1, part (iii)) and J_{TV} is a semialgebraic function, which means that its graph is defined by a finite sequence of polynomial equations and inequalities (see [37] for a definition). Hence J is the sum of an analytic function and a semialgebraic one and for this reason

it satisfies the KL property on $\mathbb{R}^N$ (see [37, p. 1769] and references therein). Conditions $(H1)-(H2)$ follows from step 2 and 3, combined with the fact that ∇J is Lipschitz continuous (Lemma 2, part (ii)), provided that the sequence of steplengths $\alpha_n^{(0)}$ defined at step 11 is bounded from above. Therefore we can state the following result:

Theorem 3 *Let J be defined as in (9), $\{\phi^{(n)}\}_{n\in\mathbb{N}}$ the sequence generated by Algorithm 1 and assume that $\alpha_n^{(0)} \leq \alpha_{\max}$, where $\alpha_{\max} > 0$. If ϕ^* is a limit point of $\{\phi^{(n)}\}_{n\in\mathbb{N}}$, then ϕ^* is a stationary point of J and $\phi^{(n)}$ converges to ϕ^*.*

Proof. We start by proving condition $(H1)$. Step 3 of Algorithm 1 can be rewritten in the following way:

$$-\alpha_n \nabla J(\phi^{(n)}) = \phi^{(n+1)} - \phi^{(n)} \tag{21}$$

from which we have

$$\alpha_n \|\nabla J(\phi^{(n)})\|^2 = \frac{1}{\alpha_n}\|\phi^{(n+1)} - \phi^{(n)}\|^2. \tag{22}$$

By substituting (22) in step 2 and since $\alpha_n \leq \alpha_n^{(0)} \leq \alpha_{\max}$, we obtain

$$J(\phi^{(n+1)}) \leq J(\phi^{(n)}) - \frac{\omega}{\alpha_n}\|\phi^{(n+1)} - \phi^{(n)}\|^2 \leq J(\phi^{(n)}) - \frac{\omega}{\alpha_{\max}}\|\phi^{(n+1)} - \phi^{(n)}\|^2. \tag{23}$$

Then $(H1)$ holds with $a = \omega/\alpha_{\max}$. Regarding condition $(H2)$, we can rewrite again step 3 as:

$$\nabla J(\phi^{(n)}) = \frac{1}{\alpha_n}(\phi^{(n)} - \phi^{(n+1)}). \tag{24}$$

Recall that the Lipschitz continuity of ∇J implies that there is $\alpha_{\min} > 0$ such that the linesearch parameter $\alpha_n \geq \alpha_{\min}$ (see [15, Proposition 4.2] for a proof). Then

$$\begin{aligned}
\|\nabla J(\phi^{(n+1)})\| &\leq \|\nabla J(\phi^{(n+1)}) - \nabla J(\phi^{(n)})\| + \|\nabla J(\phi^{(n)})\| \\
&\leq L\|\phi^{(n+1)} - \phi^{(n)}\| + \frac{1}{\alpha_n}\|\phi^{(n+1)} - \phi^{(n)}\| \\
&\leq \left(L + \frac{1}{\alpha_{\min}}\right)\|\phi^{(n+1)} - \phi^{(n)}\|.
\end{aligned}$$

This concludes the proof of $(H2)$ with $b = L + 1/\alpha_{\min}$. The thesis follows from [34, Theorem 2.9]. □

We now turn to the algorithm we used to address the nonsmooth case $\delta = 0$. In particular, we considered a simplified version of a recently proposed proximal-gradient method called VMILA (Variable Metric Inexact Linesearch Algorithm) [15]. In its general form, this method exploits a variable metric in the (possibly inexact) computation of the proximal point at each iteration and a backtracking loop to satisfy an Armijo–like inequality. Effective variable metrics can be designed for specific objective functions by exploiting suitable decompositions of the gradient of the smooth part of the objective function itself [31, 38, 39, 40]. However, since in the DIC problem the gradient of J_0 does not lead to a natural decomposition in the required form, in our tests we used the

Algorithm 2 Inexact Linesearch based Algorithm (ILA)

Choose $0 < \alpha_{\min} \leq \alpha_{\max}$, $\rho, \omega \in (0,1)$, $\gamma \in [0,1]$, $\tau > 0$, $\phi^{(0)} \in \mathbb{R}^N$ and set $n = 0$.
While True

1. Set $\alpha_n = \max\left\{\min\left\{\alpha_n^{(0)}, \alpha_{\max}\right\}, \alpha_{\min}\right\}$, where $\alpha_n^{(0)}$ is chosen as in Algorithm 1.
2. Let $h_\gamma^{(n)}$, $h^{(n)}$ and $\psi^{(n)}$ be defined as in (27)-(28). Compute $\tilde{\psi}^{(n)} \in \mathbb{R}^N$ and $\epsilon_n \geq 0$ such that
$$h^{(n)}(\tilde{\psi}^{(n)}) - h^{(n)}(\psi^{(n)}) \leq \epsilon_n \quad ; \quad \epsilon_n \leq -\tau h_\gamma^{(n)}(\tilde{\psi}^{(n)}). \tag{25}$$
3. Set $d^{(n)} = \tilde{\psi}^{(n)} - \phi^{(n)}$.
4. Compute the smallest non-negative integer i_n such that $\lambda_n = \rho^{i_n}$ satisfies
$$J(\phi^{(n)} + \lambda_n d^{(n)}) \leq J(\phi^{(n)}) + \omega \lambda_n h_\gamma^{(n)}(\tilde{\psi}^{(n)}). \tag{26}$$
5. Compute the new point as $\phi^{(n+1)} = \phi^{(n)} + \lambda_n d^{(n)}$.

   If "Stopping Criterion" is satisfied

   6. Return

   Else

   7. Set $n = n + 1$.

   EndIf

EndWhile

standard Euclidean distance (we will denote with ILA this simplified version of VMILA).

The main steps of ILA are detailed in Algorithm 2. At each iteration n, given the point $\phi^{(n)} \in \mathbb{R}^N$ and the parameters $\alpha_n > 0$, $\gamma \in [0,1]$, we define the function

$$h_\gamma^{(n)}(\phi) = \nabla J_0(\phi^{(n)})^T(\phi - \phi^{(n)}) + \frac{\gamma}{2\alpha_n}\|\phi - \phi^{(n)}\|^2 + J_{TV}(\phi) - J_{TV}(\phi^{(n)}). \tag{27}$$

We observe that $h_\gamma^{(n)}$ is strongly convex for any $\gamma \in (0,1]$. By setting $h^{(n)} = h_1^{(n)}$ and $z^{(n)} = \phi^{(n)} - \alpha_n \nabla J_0(\phi^{(n)})$, we define the unique proximal point

$$\psi^{(n)} := \mathrm{prox}_{\alpha_n J_{TV}}(z^{(n)}) = \arg\min_{\phi \in \mathbb{R}^N} h^{(n)}(\phi). \tag{28}$$

In step 2 of Algorithm 2, an approximation $\tilde{\psi}^{(n)}$ of the proximal point $\psi^{(n)}$ is defined by means of condition (25). Such a point can be practically computed by remarking that J_{TV} can be written as

$$J_{TV}(\phi) = g(\mathcal{D}\phi), \qquad g(t) = \mu \sum_{j=1}^{N} \left\| \begin{pmatrix} t_{2j-1} \\ t_{2j} \end{pmatrix} \right\|, \qquad t \in \mathbb{R}^{2N}.$$

Then considering the dual problem of (28)

$$\max_{v \in \mathbb{R}^{2N}} \Gamma^{(n)}(v), \tag{29}$$

the dual function $\Gamma^{(n)}$ has the following form

$$\Gamma^{(n)}(v) = -\frac{\|\alpha_n \mathcal{D}^T v - z^{(n)}\|^2}{2\alpha_n} - g^*(v) - J_{TV}(\phi^{(n)}) - \frac{\alpha_n}{2}\|\nabla J_0(\phi^{(n)})\|^2 + \frac{\|z^{(n)}\|^2}{2\alpha_n} \tag{30}$$

where g^* is the convex conjugate of g, namely the indicator function of the set $\left(B^2_{0,\mu}\right)^N$, being $B^2_{0,\mu} \subset \mathbb{R}^2$ the 2-dimensional Euclidean ball centered in 0 with radius μ.
Condition (25) is fulfilled by any point $\tilde{\psi}^{(n)} = z^{(n)} - \alpha_n A^T v$ with $v \in \mathbb{R}^{2N}$ satisfying [15]

$$h^{(n)}(\tilde{\psi}^{(n)}) \leq \eta \Gamma^{(n)}(v), \qquad \eta = 1/(1+\tau). \tag{31}$$

Such a point can be found by applying an iterative method to problem (29) and using (31) as stopping criterion.

Similarly to LMSD, any limit point of the sequence generated by ILA is stationary for problem (9) [15, Theorem 4.1] and, under the assumption that a limit point exists, the convergence of ILA to such a point holds when J satisfies the Kurdyka–Łojasiewicz property, the gradient of the smooth part ∇J_0 is Lipschitz continuous and the proximal point $\tilde{\psi}^{(n)}$ is computed exactly [36]. Whether and when ILA converges when the proximal point is computed inexactly is still an open problem, therefore all we can say for Algorithm 2 applied to the DIC problem is that all its limit points are stationary.

3.2. Nonlinear conjugate gradient methods

We compare the performances of LMSD and ILA with several nonlinear conjugate gradient methods, including some standard CG methods [25, 41] and the heuristic CG method previously used for DIC problems [11, 13]. The general scheme for a CG method is recalled in Algorithm 3 and some classical choices for the parameter β_{n+1} are shown in Table 1, namely the Fletcher-Reeves (FR), Polak-Ribière (PR), PR with nonnegative values (PR^+) and PR constrained by the FR values (FR-PR) strategies [42].

Algorithm 3 Conjugate gradient (CG) method

Choose $\phi^{(0)} \in \mathbb{R}^N$ and set $n = 0$, $p^{(0)} = -\nabla J(\phi^{(0)})$.
While True

1. Compute α_n and set $\phi^{(n+1)} = \phi^{(n)} + \alpha_n p^{(n)}$.
2. Choose the scalar parameter β_{n+1} according to the CG strategy used.
3. Define $p^{(n+1)} = -\nabla J(\phi^{(n+1)}) + \beta_{n+1} p^{(n)}$.

   If "Stopping Criterion" is satisfied

   4. Return

   Else

   5. Set $n = n + 1$.

   EndIf

EndWhile

CG algorithm	β_{n+1}
FR	$\beta_{n+1}^{\mathrm{FR}} = \dfrac{\nabla J(\phi^{(n+1)})^T \nabla J(\phi^{(n+1)})}{\nabla J(\phi^{(n)})^T \nabla J(\phi^{(n)})}$
PR	$\beta_{n+1}^{\mathrm{PR}} = \dfrac{\nabla J(\phi^{(n+1)})^T (\nabla J(\phi^{(n+1)}) - \nabla J(\phi^{(n)}))}{\nabla J(\phi^{(n)})^T \nabla J(\phi^{(n)})}$
PR^+	$\beta_{n+1}^{\mathrm{PR}^+} = \max(\beta_{n+1}^{\mathrm{PR}}, 0)$
FR-PR	$\beta_{n+1}^{\mathrm{FR-PR}} = \begin{cases} \beta_{n+1}^{\mathrm{PR}} & \text{if } \lvert\beta_{n+1}^{\mathrm{PR}}\rvert \le \beta_{n+1}^{\mathrm{FR}} \\ \beta_{n+1}^{\mathrm{FR}} & \text{otherwise} \end{cases}$

Table 1. Choice of the parameter β_{n+1} in CG methods. From top to bottom: Fletcher-Reeves (FR), Polak-Ribière (PR), Polak-Ribière with nonnegative β_{n+1} (PR^+), Polak-Ribière constrained by the FR method (FR-PR).

In order to ensure the global convergence of the FR and FR-PR methods, the steplength parameter α_n in step 1 must comply with the strong Wolfe conditions [42, 25]

$$\begin{aligned} J(\phi^{(n)} + \alpha_n p^{(n)}) &\le J(\phi^{(n)}) + c_1 \alpha_n \nabla J(\phi^{(n)})^T p^{(n)} \\ |\nabla J(\phi^{(n)} + \alpha_n p^{(n)})^T p^{(n)}| &\le c_2 |\nabla J(\phi^{(n)})^T p^{(n)}| \end{aligned} \tag{32}$$

where $0 < c_1 < c_2 < \frac{1}{2}$. Concerning the PR methods, one can prove convergence if β_{n+1} is chosen according to the PR^+ rule and α_n satisfies both (32) and the following additional condition [42, 25]

$$\nabla J(\phi^{(n)})^T p^{(n)} \le -c_3 \|\nabla J(\phi^{(n)})\|^2, \qquad 0 < c_3 \le 1. \tag{33}$$

For a practical implementation of a backtracking method to satisfy (32) see e.g. [25, Section 3.5], while for the addition of condition (33) see [42, Section 6]. In Section 4, the CG methods equipped with the FR, FR-PR, PR^+ rules for the parameter β_{n+1}, together with conditions (32) for the linesearch parameter α_n, will be denominated FR-SW, FR-PR-SW and PR^+-SW respectively, where SW stands for Strong Wolfe conditions.

Since in the DIC problem the evaluation of the gradient ∇J is computational demanding and its nonlinearity w.r.t. α requires a new computation for each step of the backtracking loop, in [11, 13] a heuristic version of the FR and PR methods is used exploiting a linesearch based on a polynomial approximation method. The resulting scheme for the choice of α_n is detailed in Algorithm 4, even if we recognize that our routines might differ from those used in [11, 13] due to the lack of several details crucial for reproducing their practical implementation. As we will see in the next Section, this linesearch is quite sensitive to the choice of the parameter t. Moreover, since the strong Wolfe conditions are not imposed, there is no guarantee that the FR or PR methods endowed with this choice for α_n converges, nor that $p^{(n+1)}$ is a descent direction for all n. In the following, the CG methods equipped with the FR and PR rule, together with the linesearch described in Algorithm 4, will be indicated as FR-PA and PR-PA respectively, where PA stands for polynomial approximation.

Algorithm 4 Linesearch based on polynomial approximation

Let $\psi(\alpha) := J(\phi^{(n)} + \alpha p^{(n)})$ and set $t > 0$, $a = 0$, $b = t$.
Compute $\psi(a)$ and $\psi(b)$.

1. Find a point $c \in [a, b]$ such that $\psi(a) > \psi(c) < \psi(b)$ as follows
   If $\psi(b) < \psi(a)$
      Set $c = 2b$ and compute $\psi(c)$.
      While $\psi(c) \leq \psi(b)$
         Set $a = b$, $b = c$, $c = 2c$ and compute $\psi(c)$.
      EndWhile
   Else
      Set $c = \frac{b}{2}$ and compute $\psi(c)$.
      While $\psi(c) \geq \psi(a)$
         Set $b = c$, $c = \frac{c}{2}$ and compute $\psi(c)$.
      EndWhile
   EndIf
2. Compute α_n as the minimum point of the parabola interpolating the points $(a, \psi(a)), (b, \psi(b)), (c, \psi(c))$.

4. Numerical experiments

In this section we test the effectiveness of the algorithms previously described in some synthetic problems.

4.1. Comparison between LMSD and CG methods

The evaluations of the various optimization methods discussed in Section 3 have been carried out on two phantom objects (see Figure 3), which have been computed by using the formula for the phase difference between two waves travelling through two different media

$$\phi_s = 2\pi(n_1 - n_2)t_s, \tag{34}$$

where n_1 and n_2 are the refractive indices of the object structure and the surrounding medium, respectively, and t_s is the thickness of the object at pixel $s \in \chi$. The first phantom, denominated "cone" and reported at the top row of Figure 3, is a 64×64 phase function representing a truncated cone of radius $r = 3.2$ μm with $n_1 = 1.33$, $n_2 = 1$ and maximum value $\phi_{\max} = 1.57$ rad attained at the cone vertex. The "cross" phantom, shown at the bottom row of Figure 3, is another 64×64 phase function of two crossing bars, each one of width 5 μm, measuring 0.114 rad inside the bars and 0 in the background. For both simulations, the DIC microscope parameters were set as follows:

- shear: $2\Delta x = 0.6$ μm;

- bias: $2\Delta\theta = \pi/2$ rad;
- numerical aperture of the objective: NA $= 0.9$.

For each phantom, a dataset consisting of $K = 2$ polychromatic DIC images acquired at shear angles $\tau_1 = -\pi/4$ rad and $\tau_2 = \pi/4$ rad was created, as in model (1), by convolving the true phase function with the accordingly rotated DIC PSFs and then by corrupting the result with white Gaussian noise at different values of the signal-to-noise ratio

$$\mathrm{SNR} = 10 \log_{10} \left(\frac{\overline{\phi^*}}{\sigma} \right) \tag{35}$$

where $\overline{\phi^*}$ is the mean value of the true object and σ is the standard deviation of noise. The SNR values chosen in the simulations were 9 dB and 4.5 dB.

As far as the regularization parameter μ and the threshold δ in (10) are concerned, these have been manually chosen from a fixed range in order to obtain a visually satisfactory reconstruction. Note that the parameters were first set in the differentiable case ($\delta > 0$) for the LMSD and the nonlinear CG methods and then the same value of the parameter μ was used also in the nondifferentiable case ($\delta = 0$) for the ILA method. The values reported below have been used for each simulation presented in this section. The resulting values have been $\mu = 10^{-2}, \delta = 10^{-2}$ for the cone and $\mu = 4 \cdot 10^{-2}, \delta = 10^{-3}$ for the cross.

Some details regarding the choice of the parameters involved in the optimization methods of Section 3 are now provided. The linesearch parameters ρ, ω of the LMSD and ILA methods have been respectively set to 0.5, 10^{-4}. These are the standard choices for the Armijo parameters, however it is known that the linesearch algorithm is not so sensible to modifications of these values [31, 43]. The parameter γ in the Armijo–like rule (26) has been fixed equal to 1, which corresponds to the mildest choice in terms of decrease of the objective function J. The parameter m in Algorithm 1 is typically a small value ($m = 3, 4, 5$), in order to avoid a significant computational cost in the calculation of the steplengths $\alpha_n^{(0)}$; here we let $m = 4$. The same choice for m is done in Algorithm 2, where the values $\alpha_n^{(0)}$ are constrained in the interval $[\alpha_{\min}, \alpha_{\max}]$ with $\alpha_{\min} = 10^{-5}$ and $\alpha_{\max} = 10^2$. The dual problem (29) is addressed, at each iteration of ILA, by means of algorithm FISTA [44] which is stopped by using criterion (31) with $\eta = 10^{-6}$. This value represents a good balance between convergence speed and computational time per iteration [15]. Concerning the nonlinear CG methods equipped with the strong Wolfe conditions, we set $c_1 = 10^{-4}$ and $c_2 = 0.1$ in (32) as done in [42] and we initialize the related backtracking procedure as suggested in [25, p. 59]. Regarding the CG methods endowed with the polynomial approximation detailed in Algorithm 4, a restart of the method is performed by setting $\beta_{n+1} = 0$, hence by taking a steepest descent step, whenever the vector $p^{(n+1)}$ fails to be a descent direction. Finally, the constant phase object $\phi^{(0)} = 0$ is chosen as initial guess for all methods.

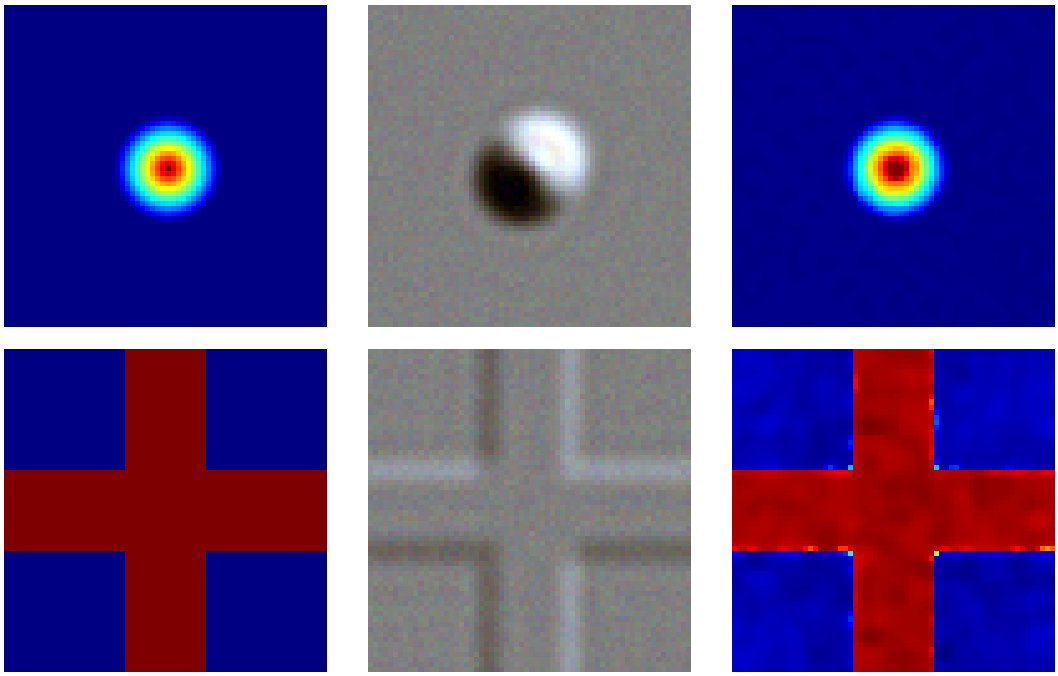


Figure 3. Data and results for the cone (top row) and cross (bottom row) objects. From left to right: true object, noisy DIC color image taken at shear angle $\frac{\pi}{4}$ rad and corrupted with white Gaussian noise at SNR = 4.5 dB, and reconstructed phase with the LMSD method from observations at shear angles equal to $-\pi/4$ rad and $\pi/4$ rad.

In order to evaluate the performance of the phase reconstruction methods proposed in Section 3, we will make use of the following error distance

$$E(\phi^{(n)}, \phi^*) = \min_{c\in\mathbb{R}} \frac{\|\phi^{(n)} - \phi^* - c\mathbf{1}\|}{\|\phi^*\|} = \frac{\|\phi^{(n)} - \phi^* - \bar{c}\mathbf{1}\|}{\|\phi^*\|} \tag{36}$$

where ϕ^* is the phase to be reconstructed and $\bar{c} = \sum_{j\in\chi} \frac{(\phi_j^{(n)} - \phi_j^*)}{N}$. Unlike the usual root mean squared error, which is recovered by setting $c = 0$ in (36), the error distance defined in (36) is invariant with respect to phase shifts, i.e.

$$E(\phi + c\mathbf{1}, \phi^*) = E(\phi, \phi^*), \quad \forall \phi \in \mathbb{R}^N, \ \forall c \in \mathbb{R}. \tag{37}$$

That makes the choice of (36) well-suited for problem (9), whose solution might be recovered only up to a real constant.

The methods have been run for the cone and cross phantoms with the parameter setting outlined in the previous subsection. The iterations of the LMSD and the CG methods have been arrested when the stopping criterion (19) was met with $\kappa = 10^{-3}$, while the ILA method has been stopped when the error up-to-a-constant between two successive iterates was lower than a prefixed $\kappa > 0$, that is

$$\frac{\left\|\phi^{(n+1)} - \phi^{(n)} - \left(\overline{\phi^{(n+1)} - \phi^{(n)}}\right)\mathbf{1}\right\|}{\|\phi^{(n+1)}\|} \leq \kappa, \tag{38}$$

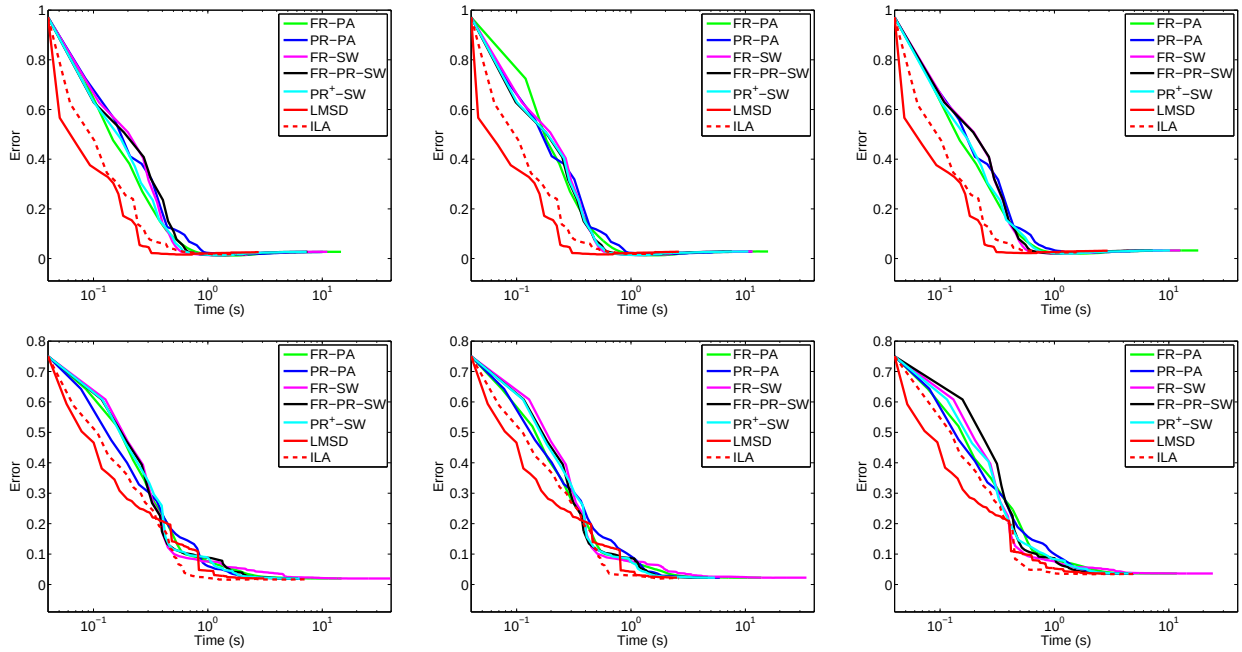


Figure 4. Error versus computational time plots for the cone (top row) and cross (bottom row) objects. From left to right: noise-free data, SNR = 9 dB and SNR = 4.5 dB.

where $\overline{\phi^{(n+1)} - \phi^{(n)}}$ is the mean value of the difference between the two objects. The tolerance κ in (38) was set equal to $5 \cdot 10^{-5}$ for the cone and 10^{-4} for the cross.

In Figure 4 we show the reconstruction error provided by the different methods as a function of the computational time. We start by comparing LMSD with the CG methods equipped with Algorithm 4 (FR-PA, PR-PA) and the CG methods equipped with the Strong Wolfe conditions (FR-SW, FR-PR-SW, PR$^+$-SW). From the plots of Figure 4, it can be drawn that each method is quite stable with respect to the noise level on the DIC images. However, in terms of time efficiency, LMSD outperforms all the CG methods in the cone tests, showing a time reduction of nearly 50% to achieve the smallest error. Furthermore, what emerges by looking at Tables 2 and 3 is that the CG methods are much more computationally demanding than LMSD. For instance, in the case of the cone (Table 1), LMSD evaluates the function on average less than 2 times per iteration. By contrast, the backtracking procedure exploited in the FR, FR-PR and PR$^+$ methods requires an average of $4-5$ evaluations per iteration of both the function and gradient to satisfy the strong Wolfe conditions, whereas the FR-PA and PR-PA methods, despite evaluating the gradient only once, need on average $10-12$ evaluations of the function before detecting the three-points-interval described in Algorithm 4. One could reduce the number of evaluations in FR-PA and PR-PA by properly tuning the parameter t in Algorithm 4. However, as it is evident from Table 4, these methods are quite sensitive to the choice of t, as little variations of this parameter might result in a great increase of the number of restarts and, eventually, in the divergence of the

SNR (dB)	Algorithm	Iterations	# f	# g	Time (s)	Obj fun	Error
	FR–PA	280	3016	280	14.60	0.89	2.72 %
	PR–PA	168	2137	168	10.10	0.89	2.66 %
	FR–SW	183	770	770	11.08	0.89	2.73 %
∞	FR-PR–SW	127	514	514	7.41	0.89	2.71 %
	PR^+–SW	129	504	504	7.32	0.89	2.71 %
	LMSD	153	212	153	2.77	0.89	2.60 %
	ILA	66	119	66	1.77	0.52	1.76 %
	FR–PA	306	3245	306	15.79	1.65	2.85 %
	PR–PA	188	2393	188	11.41	1.65	2.80 %
	FR–SW	194	804	804	11.60	1.65	2.85 %
9	FR-PR–SW	134	520	520	7.61	1.65	2.84 %
	PR^+–SW	144	734	734	10.61	1.65	2.84 %
	LMSD	149	197	149	2.61	1.65	2.75 %
	ILA	60	91	60	1.56	1.29	1.91 %
	FR–PA	347	3696	347	18.08	6.88	3.26 %
	PR–PA	146	1858	146	8.84	6.88	3.24 %
	FR–SW	204	867	867	12.58	6.88	3.26 %
4.5	FR-PR–SW	152	492	492	7.24	6.88	3.26 %
	PR^+–SW	144	701	701	10.22	6.88	3.26 %
	LMSD	163	228	163	2.90	6.88	3.17 %
	ILA	61	104	61	1.56	6.80	2.50 %

Table 2. Cone tests. From left to right: number of iterations required to meet the stopping criteria, number of function and gradient evaluations, execution time, objective function value and error achieved at the last iteration.

algorithm. In addition, it seems that the optimal value of t strictly depends on the object to be reconstructed.

4.2. Comparison between LMSD and ILA

We now compare the performances of LMSD and ILA. On one hand, ILA reconstructs the cross object slightly better than LMSD. Indeed, ILA provides the lowest reconstruction error in Table 3 for each SNR value and the corresponding phase estimates have better preserved edges, as clearly depicted in Figure 5, where we consider the following "up-to-a-constant" residual

$$R_j = \left|\phi_j - \phi_j^* - \overline{\phi - \phi^*}\right|, \quad \forall j \in \chi \tag{39}$$

to measure the quality of the reconstructions provided by the two methods. This result was expected, since ILA addresses problem (9) with the standard TV functional ($\delta = 0$ in (10)), which is more suited than HS regularization ($\delta > 0$) when the object to be reconstructed is piecewise-constant. On the other hand, ILA may be computationally

SNR (dB)	Algorithm	Iterations	# f	# g	Time (s)	Obj fun	Error
	FR–PA	412	2618	412	14.75	1.01	1.98 %
	PR–PA	138	1373	138	6.73	1.01	1.98 %
	FR–SW	411	2768	2768	39.27	1.01	1.98 %
∞	FR-PR–SW	109	423	423	6.14	1.01	1.98 %
	PR^+–SW	116	438	438	6.32	1.01	1.98 %
	LMSD	168	231	168	3.09	1.01	2.00 %
	ILA	100	176	100	7.18	0.87	1.66 %
	FR–PA	391	2490	391	13.77	1.96	2.25 %
	PR–PA	121	1209	121	5.97	1.96	2.26 %
	FR–SW	388	2417	2417	34.18	1.96	2.25 %
9	FR-PR–SW	106	323	323	4.69	1.96	2.25 %
	PR^+–SW	109	375	375	5.41	1.96	2.25 %
	LMSD	140	190	140	2.52	1.96	2.27 %
	ILA	57	106	57	2.60	1.82	1.94 %
	FR–PA	303	2164	303	11.74	8.57	3.63 %
	PR–PA	98	997	98	4.97	8.57	3.63 %
	FR–SW	299	1705	1705	24.28	8.57	3.63 %
4.5	FR-PR–SW	96	300	300	4.41	8.57	3.63 %
	PR^+–SW	98	326	326	4.74	8.57	3.63 %
	LMSD	152	221	152	2.75	8.57	3.64 %
	ILA	97	179	97	5.26	8.47	3.46 %

Table 3. Cross tests. From left to right: number of iterations required to meet the stopping criteria, number of function and gradient evaluations, execution time, objective function value and error achieved at the last iteration.

Dataset	t	Iterations	# f	Time (s)	Obj fun	Error	Restarts
	10^{-4}	500	4272	24.57	8.57	3.63 %	8
Cross	10^{-3}	500	2911	19.66	8.57	3.63 %	6
-	$5 \cdot 10^{-3}$	500	3073	19.63	8.57	3.63 %	1
SNR	10^{-2}	500	5337	28.97	8.57	3.63 %	21
4.5 dB	$5 \cdot 10^{-2}$	500	2023	15.22	8.59	3.91 %	424
	10^{-1}	500	2032	15.44	8.88	5.05 %	365
	10^{-3}	500	4788	26.13	6.88	3.27 %	0
Cone	10^{-2}	500	3260	19.84	6.88	3.27 %	0
-	10^{-1}	500	2126	15.86	6.88	3.27 %	3
SNR	$2 \cdot 10^{-1}$	500	2427	16.78	6.88	3.27 %	0
4.5 dB	$2.25 \cdot 10^{-1}$	500	1610	13.39	1507.4	130.94 %	41
	$2.5 \cdot 10^{-1}$	500	1713	13.67	2373.4	315.50 %	87

Table 4. Setting the parameter t in the PR-PA algorithm.

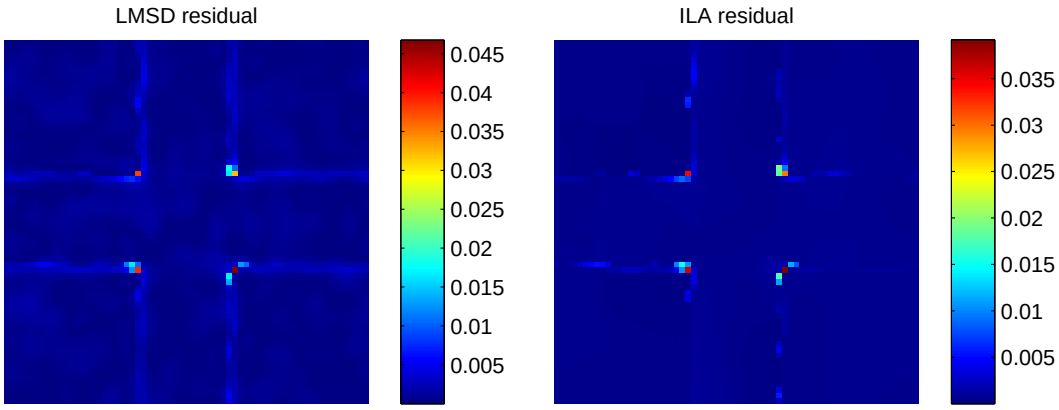


Figure 5. Cross test. The residuals defined in (39) for the reconstructions provided by LMSD and ILA, respectively, when the acquired images are corrupted with SNR = 9 dB.

more expensive since, unlike LMSD, it requires to iteratively solve the inner subproblem (29) at each outer iteration. Indeed, looking at Table 3 we notice that, although the number of function evaluations per iteration in LMSD and ILA is quite similar (on average around 1.4 for LMSD and 1.8 for ILA) and the ILA iterations are stopped way before the LMSD ones, the computational time in ILA is always higher. For instance, in the case SNR = 9 dB, the methods require approximately the same time, although the number of iterations of ILA is more than halved. This fact is explained if we look at the average number of inner iterations required by ILA to compute the approximate proximal point: 21.3, 10.11 and 13.43 for SNR $= \infty, 9, 4.5$ dB respectively. Analogous conclusions can be drawn by considering the results on the cone object (see Table 2).

In order to deepen the analysis between the differentiable TV approximation and the original nondifferentiable one, we compared the LMSD and ILA methods in one further realistic simulation. In particular, we considered the "grid" object in Figure 6, which is a 1388×1040 image emulating the phase function of a multi-area calibration artifact [45], which measures 1.212 rad inside the black regions and 2.187 rad inside the white ones. The setup of the two methods is identical to that of the previous tests (with the exception of the numerical aperture of the objective NA which has been set equal to 0.8), and the parameters μ (for both models) and δ (for the smooth TV functional) have been set equal to $2 \cdot 10^{-1}$ and 10^{-1}, respectively. Instead of three levels of noise, here we only considered a SNR equal to 9 dB. In Figure 7 we report the behaviour of the error (36) as a function of time and the number of inner iterations needed by ILA to address problem (29)–(31).

The grid dataset confirms the remarks previously done, since ILA takes almost twice the time than LMSD to provide an estimate of the phase. This is again due to the

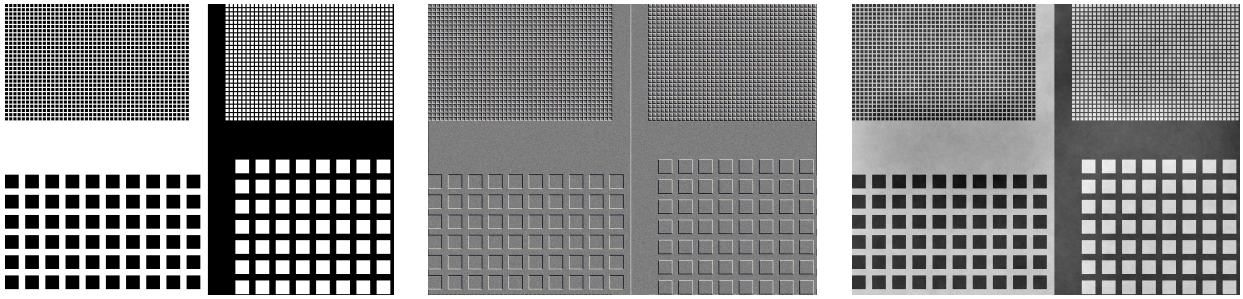


Figure 6. Data and results for the grid object. From left to right: true object, noisy DIC color image taken at shear angle $\frac{\pi}{4}$ rad and corrupted with white Gaussian noise at SNR = 9 dB, and reconstructed phase with the LMSD method from observations at shear angles equal to $-\pi/4$ rad and $\pi/4$ rad.

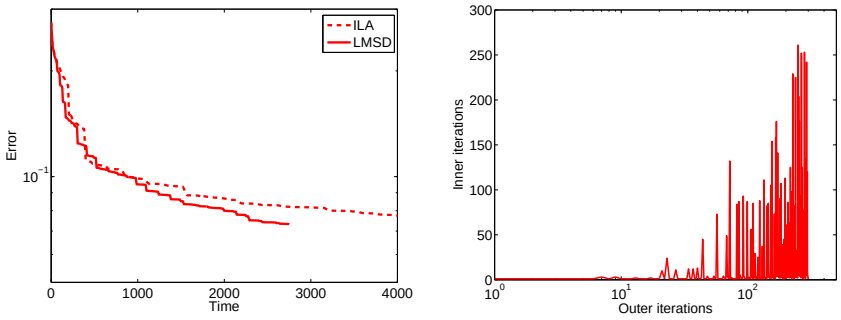


Figure 7. Grid test. From left to right: error versus time plots for LMSD and ILA and number of inner iterations versus number of outer iterations for ILA.

number of inner iterations, which starts to oscillatory increase after the first 20 iterations (see Figure 7). To conclude, we reckon that the LMSD method is generally preferable since, unlike ILA, it does not require any inner subproblem to be solved and thus it is generally less expensive from the computational point of view. However, the ILA method should be considered as a valid alternative when the object to be reconstructed is piecewise-constant.

4.3. Influence of color and bias retardation on phase reconstruction

Another analysis that was of our interest was to observe how color information and bias retardation in the observations affect the behavior of phase reconstruction. We set four scenarios for comparison: independent monochromatic observations with red, green, and blue light, and polychromatic observation where all wavelengths are combined. For each of these scenarios we used the cross object to generate 100 observations at different realizations of noise, for both SNR = 4.5 dB and SNR = 9 dB, and bias retardation of 0 rad and $\pi/2$ rad, at shear angles equal to $-\pi/4$ rad and $\pi/4$ rad. We tested the LMSD method to perform the reconstructions; results for SNR = 4.5 dB are shown in Figure 8 and for SNR = 9 dB in Figure 9.

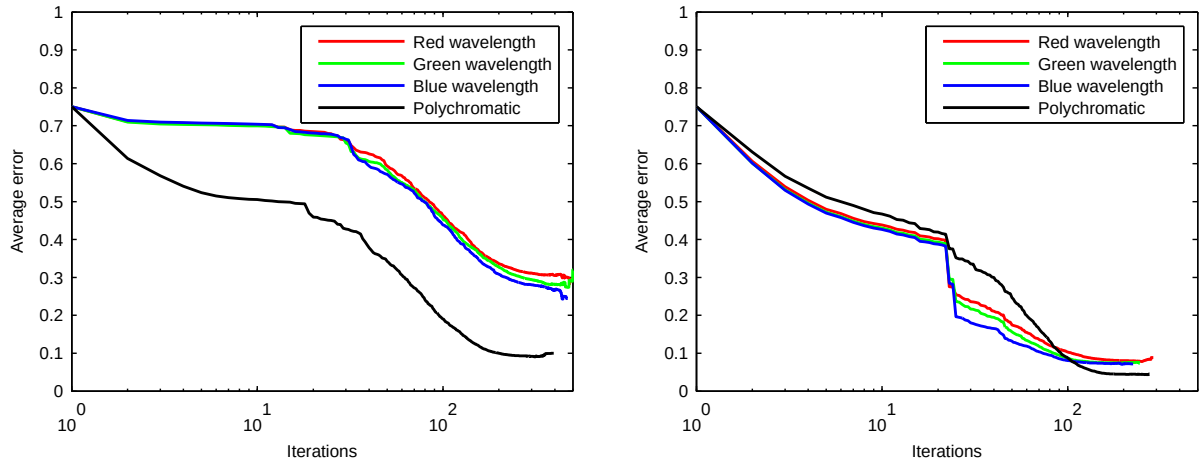


Figure 8. Average error comparison between monochromatic and polychromatic reconstructions. SNR = 4.5 dB. Left: bias 0 rad; right: bias $\pi/2$ rad.

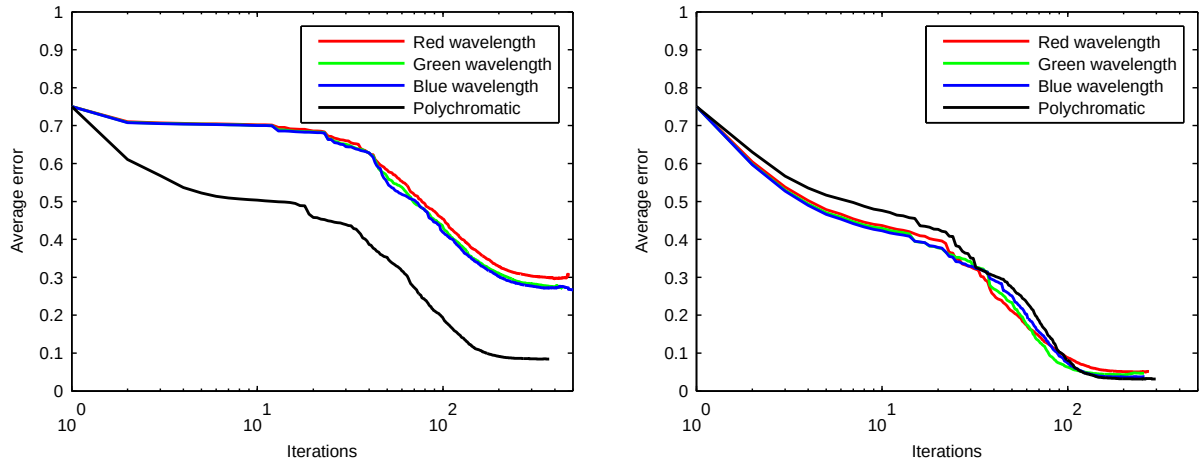


Figure 9. Average error comparison between monochromatic and polychromatic reconstructions. SNR = 9 dB. Left: bias 0 rad; right: bias $\pi/2$ rad.

The lines show the average error over the 100 observations. It is noticed that for 0 rad bias retardation, the reconstruction for polychromatic observations behave better than for the monochromatic ones, even though the amount of error is not promising of a good reconstruction. For $\pi/2$ rad bias retardation the estimation the algorithm stops before the maximum number of iterations (500) is reached. In this case, for both levels of noise, the performance of the reconstruction with polychromatic light is quite comparable with monochromatic light. Another interesting finding about the convergence for monochromatic light, is that for all cases, it happens in the order red-green-blue; this is due to the fact that the amplitude PSF for blue light has the bigger frequency support, thus provides more information for reconstruction.

5. Conclusions and future work

In this paper we provided both theoretical and practical contributions to the inverse problem of phase estimation from polychromatic DIC images. First of all, we studied the analytical properties of the data fidelity function arising from a maximum likelihood approach, showing its periodicity, shift-invariancy and analyticity. Secondly, we analyzed the minimization problem of the functional given by the sum of the discrepancy term and a (possibly smoothed version of) total variation regularizer, proving the existence of minimum points. Furthermore, we revisited the state-of-the-art optimization method for phase estimation in DIC microscopy providing implementation details, showing possible pitfalls and comparing its performances with standard conjugate gradient algorithms. Finally, we proposed two recent optimization strategies for the smooth and nonsmooth cases, showing in simulated datasets that they provide accurate reconstructions of the phase in a lower computational time.

Future work will concern the application of the proposed strategies to real acquisitions and the reformulation of the minimization problem in terms of the specimen's transmission function $e^{-i\phi}$. This would lead to a standard regularized least-squares problem restricted to a nonconvex feasible set, which would require a generalization of the (VM)ILA approach able to account for nonconvex projections and to exploit the steplength selection rule proposed by Fletcher [26] in the presence of constraints [46, 47].

Acknowledgments

This work has been partially supported by the ECOS-Nord grant C15M01, the Universidad Industrial de Santander and the Italian Ministry for University and Research (MIUR) under the projects FIRB - Futuro in Ricerca 2012, contract RBFR12M3AC, and PRIN 2012, contract 2012MTE38N. The Italian GNCS - INdAM (Gruppo Nazionale per il Calcolo Scientifico - Istituto Nazionale di Alta Matematica) is also acknowledged.

References

[1] S. P. Singh and B. S. Tomar. *Cell Biology.* Rastogi Publications, India, 2007.

[2] S. M. Wilson and A. Bacic. Preparation of plant cells for transmission electron microscopy to optimize immunogold labeling of carbohydrate and protein epitopes. *Nat. Protoc.*, 7(9):1716–1727, August 2012.

[3] M. J. Booth. Adaptive optical microscopy: the ongoing quest for a perfect image. *Light Sci. Appl.*, 3:1–7, April 2014.

[4] M. M. Frigault, J. Lacoste, J. L. Swift, and C. M. Brown. Live-cell microscopy tips and tools. *J. Cell Sci.*, 122:753–767, 2009.

[5] R. D. Allen, G. B. David, and G. Nomarski. The Zeiss-Nomarski differential interference equipment for transmitted-light microscopy. *Z. Wiss. Mikrosk.*, 69(4):193–221, 1969.

[6] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev. Phase retrieval with application to optical imaging. *IEEE Signal Process. Mag.*, 32(3):87–109, 2015.

[7] E. B. van Munster, L. J. van Vliet, and J. A. Aten. Reconstruction of optical pathlength distributions from images obtained by a wide-field differential interference contrast microscope. *J. Microsc.*, 188(2):149–157, 1997.

[8] Z. Kam. Microscopic differential interference contrast image processing by line integration (LID) and deconvolution. *Bioimaging*, 6(4):166–176, 1998.

[9] S. S. Kou, L. Waller, G. Barbastathis, and C. J. R. Sheppard. Transport-of-intensity approach to differential interference contrast (TI-DIC) microscopy for quantitative phase imaging. *Opt. Lett.*, 35(3):447–449, 2010.

[10] E. Bostan, E. Froustey, B. Rappaz, E. Shaffer, D. Sage, and M. Unser. Phase retrieval by using transport-of-intensity equation and differential interference contrast microscopy. In *Proc. 21th IEEE International Conference on Image Processing*, pages 3939–3943, 2014.

[11] C. Preza, D. L. Snyder, and J.-A. Conchello. Image reconstruction for three-dimensional transmitted-light DIC microscopy. In J.-A. Conchello, C. J. Cogswell, and T. Wilson, editors, *Three-Dimensional Microscopy: Image Acquisition and Processing IV*, volume 2984 of *Proc. SPIE*, 1997.

[12] C. Preza, D. L. Snyder, and J.-A. Conchello. Theoretical development and experimental evaluation of imaging models for differential interference contrast microscopy. *J. Opt. Soc. Am. A*, 16(9):2185–2199, 1999.

[13] C. Preza. Rotational-diversity phase estimation from differential interference contrast microscopy images. *J. Opt. Soc. Am. A*, 17(3):415–424, 2000.

[14] C. Preza, S. V. King, and C. J. Cogswell. Algorithms for extracting true phase from rotationally-diverse and phase-shifted DIC images. In J.-A. Conchello, C. J. Cogswell, and T. Wilson, editors, *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XIII*, volume 6090 of *Proc. SPIE*, 2006.

[15] S. Bonettini, I. Loris, F. Porta, and M. Prato. Variable metric inexact line–search based methods for nonsmooth optimization. *SIAM J. Optim.*, 26(2):891–921, 2016.

[16] D. B. Murphy. *Fundamentals of light microscopy and electronic imaging.* Wiley-Liss, New York, 2001.

[17] S. B. Mehta and C. J. R. Sheppard. Partially coherent image formation in differential interference contrast (DIC) microscope. *Opt. Express*, 16(24):19462–19479, November 2008.

[18] L. Bautista, S. Rebegoldi, L. Blanc-Féraud, M. Prato, L. Zanni, and A. Plata. Phase estimation in differential-interference-contrast (DIC) microscopy. In *Proc. IEEE Int. Symp. Biomed. Imaging*, pages 136–139, 2016.

[19] J. W. Goodman. *Statistical Optics.* Wiley, New York, 1984.

[20] S. Krantz and H. R. Parks. *A primer of Real Analytic Functions.* Birkhäuser, 2002.

[21] R. Acar and C. R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Probl.*, 10(6):1217–1229, June 1994.

[22] M. Bertero, P. Boccacci, G. Talenti, R. Zanella, and L. Zanni. A discrepancy principle for Poisson data. *Inverse Probl.*, 26(10), October 2010.

[23] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *J. Phys. D.*, 60(1–4):259–268, 1992.

[24] A. Cornelio, F. Porta, and M. Prato. A convergent least-squares regularized blind deconvolution approach. *Appl. Math. Comput.*, 259(12):173–186, May 2015.

[25] J. Nocedal and S. J. Wright. *Numerical optimization.* Springer, New York, 2nd edition, 2006.

[26] R. Fletcher. A limited memory steepest descent method. *Math. Program.*, 135(1–2):413–436, October 2012.

[27] J. Barzilai and J. M. Borwein. Two-point step size gradient methods. *IMA J. Numer. Anal.*, 8(1):141–148, January 1988.

[28] Y. H. Dai and Y. X. Yuan. Alternate minimization gradient method. *IMA J. Numer. Anal.*,

23(3):377–393, July 2003.

[29] R. De Asmundis, D. di Serafino, F. Riccio, and G. Toraldo. On spectral properties of steepest descent methods. *IMA J. Numer. Anal.*, 33(4):1416–1435, October 2013.

[30] B. Zhou, L. Gao, and Y. H. Dai. Gradient methods with adaptive step-sizes. *Comput. Optim. Appl.*, 35(1):69–86, September 2006.

[31] S. Bonettini, R. Zanella, and L. Zanni. A scaled gradient projection method for constrained image deblurring. *Inverse Probl.*, 25(1), January 2009.

[32] S. Łojasiewicz. Une propriété topologique des sous-ensembles analytiques réels. In *Les Équations aux Dérivées Partielles*, pages 87–89. Éditions du Centre National de la Recherche Scientifique, Paris, 1963.

[33] K. Kurdyka. On gradients of functions definable in o-minimal structures. *Ann. Inst. Fourier*, 48(3):769–783, 1998.

[34] H. Attouch, J. Bolte, and B. F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. *Math. Program.*, 137(1–2):91–129, February 2013.

[35] P. Frankel, G. Garrigos, and J. Peypouquet. Splitting methods with variable metric for Kurdyka–Łojasiewicz functions and general convergence rates. *J. Optim. Theory Appl.*, 165(3):874–900, June 2015.

[36] S. Bonettini, I. Loris, F. Porta, M. Prato, and S. Rebegoldi. On the convergence of variable metric line-search based proximal-gradient method under the Kurdyka– Łojasiewicz inequality. *ArXiv e-prints*, page 1605.03791, 2016.

[37] Y. Xu and W. Yin. A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion. *SIAM J. Imaging Sci.*, 6(3):1758–1789, 2013.

[38] H. Lantéri, M. Roche, and C. Aime. Penalized maximum likelihood image restoration with positivity constraints: multiplicative algorithms. *Inverse Probl.*, 18(5):1397–1419, October 2002.

[39] S. Bonettini, G. Landi, E. Loli Piccolomini, and L. Zanni. Scaling techniques for gradient projection-type methods in astronomical image deblurring. *Int. J. Comput. Math.*, 90(1):9–29, January 2013.

[40] S. Bonettini, A. Chiuso, and M. Prato. A scaled gradient projection method for Bayesian learning in dynamical systems. *SIAM J. Sci. Comput.*, 37(3):A1297–A1318, 2015.

[41] R. Fletcher. *Practical methods of optimization*. John Wiley and Sons, New York, 2nd edition, 2000.

[42] J. C. Gilbert and J. Nocedal. Global convergence properties of conjugate gradient methods for optimization. *SIAM J. Optim.*, 2(1):21–42, 1992.

[43] M. Prato, R. Cavicchioli, L. Zanni, P. Boccacci, and M. Bertero. Efficient deconvolution methods for astronomical imaging: algorithms and IDL-GPU codes. *Astron. Astrophys.*, 539:A133, March 2012.

[44] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, 2(1):183–202, 2009.

[45] Bruker AFM Probes. Product description APCS-0099. `http://www.brukerafmprobes.com/a-3472-apcs-0099.aspx`, July 2016.

[46] F. Porta, R. Zanella, G. Zanghirati, and L. Zanni. Limited-memory scaled gradient projection methods for real-time image deconvolution in microscopy. *Commun. Nonlinear Sci. Numer. Simul.*, 21(1–3):112–127, April 2015.

[47] F. Porta, M. Prato, and L. Zanni. A new steplength selection for scaled gradient methods with application to image deblurring. *J. Sci. Comput.*, 65(3):895–919, December 2015.