Optimization methods for phase estimation in differential-interference-contrast (DIC) microscopy

Simone Rebegoldi[1], Lola Bautista[2,3], Laure Blanc-Féraud[3], Marco Prato[1], Luca Zanni[1] and Arturo Plata[2]

[1]Università di Modena e Reggio Emilia, [2]Universidad Industrial de Santander, [3]Université Côte d'Azur

The DIC phase estimation problem

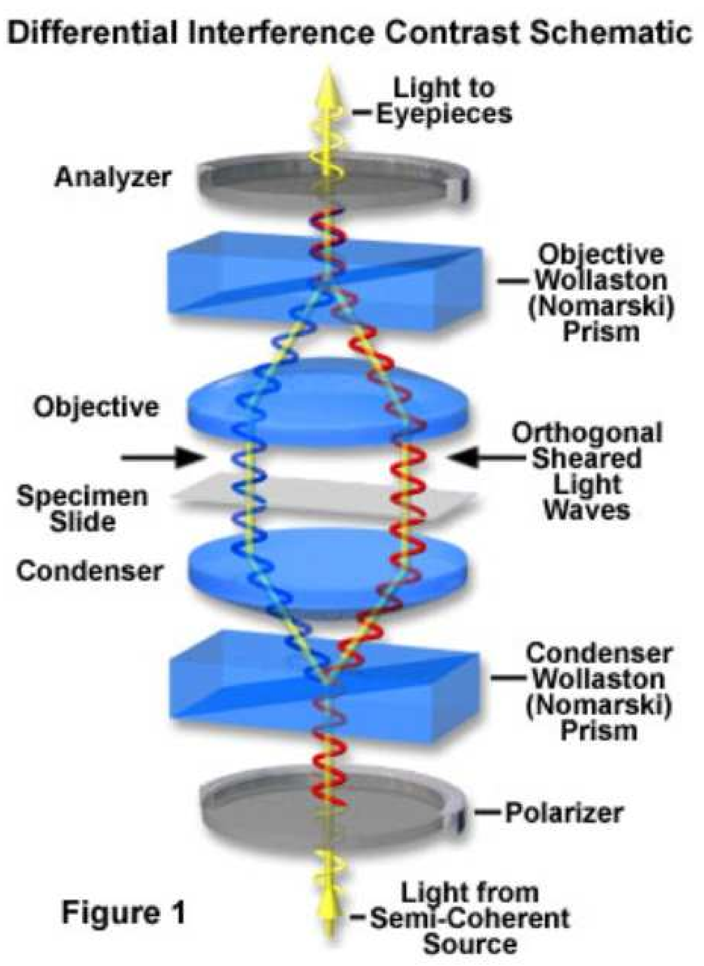
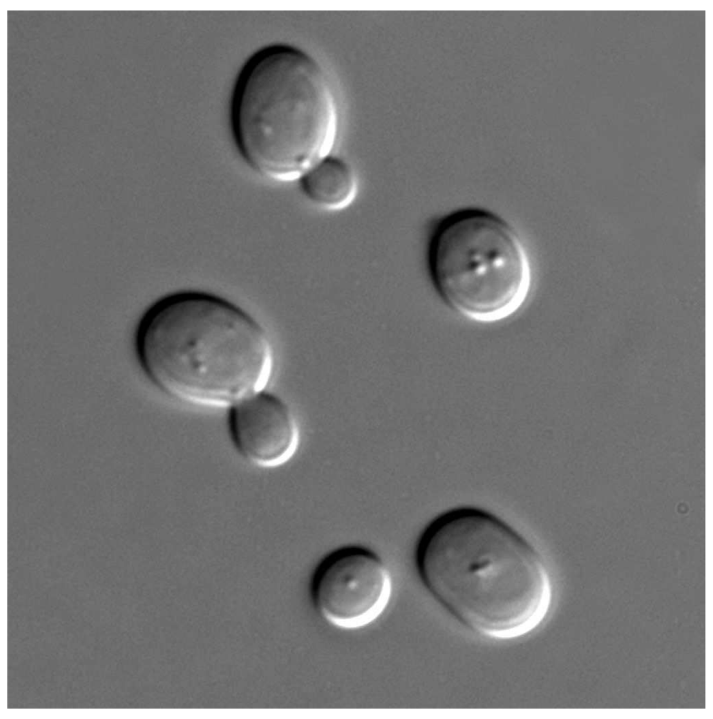


Figure 1



The DIC image is formed by the interference of two orthogonally polarized beams that have a lateral displacement (called shear) and are phase shifted relatively one to each other. The resulting image has a 3D high contrast appearance, which can be enhanced by introducing a uniform phase difference between the beams (called bias).

Model: the DIC image formation is described by the polychromatic rotational-diversity model [1,2]

$$(o_{k,\lambda_\ell})_j = \left|(h_{k,\lambda_\ell} \otimes e^{-i\phi/\lambda_\ell})_j\right|^2 + (\eta_{k,\lambda_\ell})_j, \qquad k = 1,\dots,K, \;\; \ell = 1,2,3, \;\; j \in \chi$$

- k is the index of the rotation of the specimen w.r.t. the horizontal axis, ℓ is the index denoting one of the RGB channels and $j = (j_1, j_2)$ is a 2D–index varying in the set $\chi = \{1,\dots,M\} \times \{1,\dots,P\}$
- λ_ℓ is the ℓ–th illumination wavelength
- $o_{k,\lambda_\ell} \in \mathbb{R}^{MP}$ is the ℓ–th color component of the k–th 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 $e^{-i\phi/\lambda_\ell} \in \mathbb{C}^{MP}$ is 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
- $\eta_{k,\lambda_\ell} \in \mathbb{R}^{MP}$ is the noise corrupting the data, $\eta_{k,\lambda_\ell} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_{(MP)^2})$.

Problem: given the rotationally diverse images $o_1,\dots,o_K$, retrieve the phase vector ϕ by solving

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

- $J_0(\phi) = \sum_{\ell=1}^{3}\sum_{k=1}^{K}\sum_{j\in\chi}\left[(o_{k,\lambda_\ell})_j - \left|(h_{k,\lambda_\ell} \otimes e^{-i\phi/\lambda_\ell})_j\right|^2\right]^2$ is the nonconvex least-squares term
- $J_{TV}(\phi) = \mu \sum_{j\in\chi}\sqrt{((\mathcal{D}\phi)_j)_1^2 + ((\mathcal{D}\phi)_j)_2^2 + \delta^2}$ is the total variation (TV) functional, where $\mu > 0$ is the regularization parameter and $\delta \geq 0$ is the smoothing parameter ($\delta = 0 \to$ standard TV).

Optimization methods

Case $\delta > 0$: problem (P) is differentiable $\to$ we use a gradient-descent method.

Algorithm 1 Limited Memory Steepest Descent (LMSD) method [3]

Set $\rho, \omega \in (0,1)$, $m > 0$, $\alpha_0^{(0)}, \dots, \alpha_{m-1}^{(0)} > 0$, $\phi^{(0)} \in \mathbb{R}^{MP}$, $G = [\,]$, $\Theta = [\,]$, $n = 0$.

WHILE True

FOR $l = 1,\dots,m$

1. Compute the smallest non-negative integer i_n such that $\alpha_n = \alpha_n^{(0)}\rho^{i_n}$ satisfies
$$J(\phi^{(n)} - \alpha_n \nabla J(\phi^{(n)})) \leq J(\phi^{(n)}) - \omega\alpha_n \|\nabla J(\phi^{(n)})\|^2.$$
2. Compute the new point as $\phi^{(n+1)} = \phi^{(n)} - \alpha_n \nabla J(\phi^{(n)})$.
3. Update $G = [G \;\; \nabla J(\phi^{(n)})]$ and $\Theta = [\Theta \;\; \alpha_n^{-1}]$.
4. Set $n = n+1$.

END

6. Define the $(m+1)\times m$ matrix $\Gamma = \begin{bmatrix}\text{diag}(\Theta)\\ \text{zeros}(1,m)\end{bmatrix} - \begin{bmatrix}\text{zeros}(1,m)\\ \text{diag}(\Theta)\end{bmatrix}$.
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 $\Phi = [R, r]\Gamma R^{-1}$ and its approximation
$$\widetilde{\Phi} = \text{diag}(\Phi) + \text{tril}(\Phi, -1) + \text{tril}(\Phi, -1)^T,$$
which is symmetric and tridiagonal.
10. Compute eigenvalues $\theta_1,\dots,\theta_m$ of $\widetilde{\Phi}$ and define $\alpha_{n+i-1}^{(0)} = 1/\theta_i$, $i = 1,\dots,m$.

END

Case $\delta = 0$: problem (P) is non differentiable $\to$ we use a proximal-gradient method.

Algorithm 2 Inexact Linesearch based Algorithm (ILA) [4]

Set $\rho, \omega \in (0,1)$, $0 < \alpha_{\min} \leq \alpha_{\max}$, $\tau > 0$, $\phi^{(0)} \in \mathbb{R}^{MP}$, $n = 0$.

WHILE True

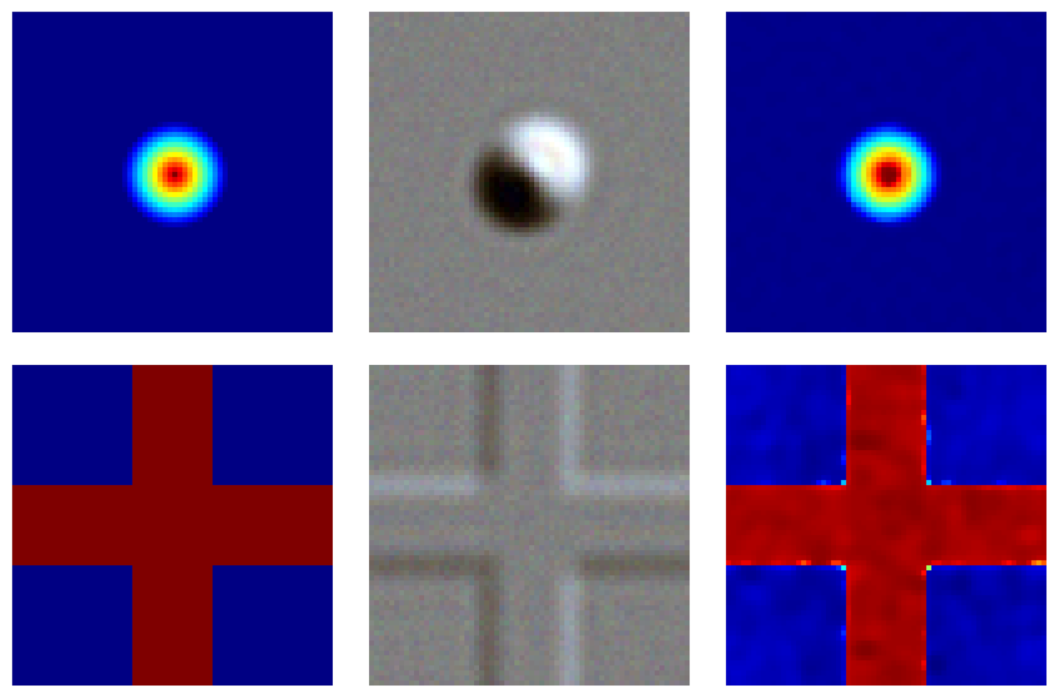
1. Set $\alpha_n = \max\left\{\min\left\{\alpha_n^{(0)}, \alpha_{\max}\right\}, \alpha_{\min}\right\}$, where $\alpha_n^{(0)}$ is chosen as in Algorithm 1.
2. Let $\psi^{(n)} = \text{prox}_{\alpha_n J_{TV}}\left(\phi^{(n)} - \alpha_n \nabla J_0(\phi^{(n)})\right) = \text{argmin}_{\phi\in\mathbb{R}^{MP}} h^{(n)}(\phi)$.
Compute $\tilde{\psi}^{(n)}$ such that $h^{(n)}(\tilde{\psi}^{(n)}) - h^{(n)}(\psi^{(n)}) \leq \epsilon_n$ and $0 \leq \epsilon_n \leq -\tau h^{(n)}(\tilde{\psi}^{(n)})$.
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
$$J(\phi^{(n)} + \lambda_n d^{(n)}) \leq J(\phi^{(n)}) + \omega\lambda_n h^{(n)}(\tilde{\psi}^{(n)}).$$
5. Compute the new point as $\phi^{(n+1)} = \phi^{(n)} + \lambda_n d^{(n)}$.
6. Set $n = n+1$.

END

Convergence and numerical results

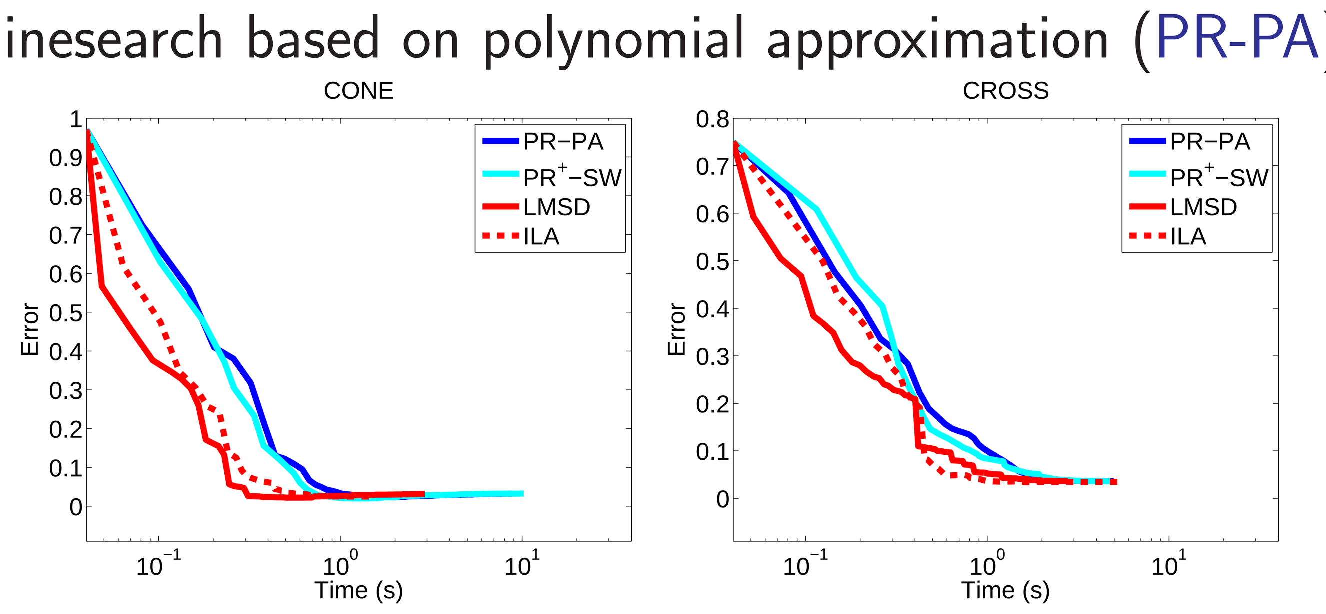
Convergence: Any limit point of Algorithm 1 and 2 is stationary for problem (P). Since J satisfies the Kurdyka-Łojasiewicz property, Algorithm 1 converges to a limit point; the same result can be proved for Algorithm 2 when the proximal point is computed exactly [4].

Results: for both objects, cone (top row) and cross (bottom row), $K = 2$ DIC images have been generated.



True phase   Noisy DIC image   Rec. phase

The parameters of the methods have been tuned as follows: $\rho = 0.5$, $\omega = 10^{-4}$, $m = 4$, $\alpha_{\min} = 10^{-5}$, $\alpha_{\max} = 10^2$, $\tau = 10^6 - 1$, $\phi^{(0)} = 0$. The methods are compared with the Polak-Ribière conjugate gradient method equipped with the strong Wolfe conditions (PR$^+$-SW) and a linesearch based on polynomial approximation (PR-PA) [1].



Object	Algorithm	Iterations	# f	# g	Error
Cross	PR–PA	98	997	98	3.63 %
	PR$^+$–SW	98	326	326	3.63 %
	LMSD	152	221	152	3.64 %
	ILA	97	179	97	3.46 %

[1] C. Preza, J. Opt. Soc. Am. A, 17(3), 415–424, 2000.
[2] L. Bautista et al., 2016 IEEE 13th Int. Symp. on Biomedical Imaging, 136–139, 2016.
[3] R. Fletcher, Math. Program., 135(1–2), 413–436, 2012.
[4] S. Bonettini et al., ArXiv: 1605.03791, 2016.