Appl Math Optim (2016) 74:303–324
DOI 10.1007/s00245-015-9315-3

Incremental projection approach of regularization for inverse problems

Innocent Souopgui[1] · **Hans E. Ngodock**[2] · **Arthur Vidard**[3] · **François-Xavier Le Dimet**[3]

Published online: 22 September 2015

Abstract This paper presents an alternative approach to the regularized least squares solution of ill-posed inverse problems. Instead of solving a minimization problem with an objective function composed of a data term and a regularization term, the regularization information is used to define a projection onto a convex subspace of regularized candidate solutions. The objective function is modified to include the projection of each iterate in the place of the regularization. Numerical experiments based on the problem of motion estimation for geophysical fluid images, show the improvement of the proposed method compared with regularization methods. For the presented test case, the incremental projection method uses 7 times less computation time than the regularization method, to reach the same error target. Moreover, at convergence, the incremental projection is two order of magnitude more accurate than the regularization method.

Keywords Regularization · Projection · Inverse problems · Motion estimation

✉ Innocent Souopgui
innocent.souopgui@usm.edu

Hans E. Ngodock
hans.ngodock@nrlssc.navy.mil

Arthur Vidard
arthur.vidard@imag.fr

François-Xavier Le Dimet
ledimet@imag.fr

1 Department of Marine Science, The University of Southern Mississippi, 1020 Balch Blvd, Stennis Space Center, MS 39529, USA

2 Naval Research Laboratory, 1009 Balch Blvd, Stennis Space Center, MS 39529-5004, USA

3 Laboratoire Jean Kuntzmann, 51 rue des Maths, 38400 Saint Martin d'Hères, France

Mathematics Subject Classification 49K40 · 49M05 · 68U10

1 Introduction

Regularization methods are usually applied to ill-posed inverse problems in order to retrieve a solution with desired properties. The solution is generally sought as the minimizer of an objective function composed of a data term and a regularization term. The data term measures the distance between a given candidate solution and an observed output of the studied system. The regularization term is a functional constraint that implicitly defines the subset of admissible solutions; it is problem dependent and is chosen to enforce the well-posedness of the inverse problem.

In the community of image processing, projections methods have been considered in the place of regularization. Based on earlier works from Natterer [35] and Louis [26] that demonstrated the regularizing effects of projection methods, Schuster and Weickert [38] used a projection method to solve the inverse problem associated with motion estimation (Appendix 1 gives an overview of motion estimation). They used an explicit basis to determine a projection subspace, but did not provide any criteria for choosing or constructing such a subspace. However, they cautioned about carefully choosing the subspace as it could affect the solution of the inverse problem. Furthermore, the regularizing effects of a projection method does not necessarily correspond to those of a given regularization method. In fact the subset of regularized admissible solutions is not necessarily a subspace (in the sense of vector space). For that reason, the authors cautioned of choosing the projection subspace carefully since it determines the properties of the solution. Because an explicit basis is required to construct the projection, the method is called explicit projection thereafter.

However, it is possible to define a projection in such a way that its regularizing effect corresponds to those of a given regularization method, i.e. solving the projection problem will yield a solution with all the properties of the regularized solution.

In this study we propose a generalized approach to define a projection method that corresponds to any regularization method. The functional constraint of the regularized objective function is used to define a projection onto the subspace of admissible solutions that is convex for linear regularization functions and for the linearized approximation of nonlinear ones. The subset of admissible solutions is implicitly defined by the projection and has all the properties imposed by the regularization constraint. Thanks to those properties, the solution of the projection method will have all the properties of the regularized solution, and the convexity of the subspace enforces the convergence towards the solution. Afterward, the objective function is modified to include the projection in the place of the regularization, and solved with a standard optimization algorithm. In the implementation, the projection is applied to each iterate, and for that reason, the method is called incremental.

The remaining part of the paper is organized as follows: Sect. 2 gives a short description of inverse problems, regularization and explicit projection methods. Section 3

which is the main contribution of the present work, introduces the proposed projection method for inverse problems. Then details on construction of such projection are given for the general case of discretized problems, followed by the example of the minimum norm and the minimum gradient regularization in the continuous case. Section 4 shows numerical examples based on the motion estimation problem, and indeed illustrate the feasibility and the effectiveness of the proposed method. Section 5 concludes the paper.

2 Inverse Problems and Approaches of Solution

2.1 Notation

Hereafter, we use the following notations unless otherwise stated: Ω is the physical space of the problem, $\mathcal{V}$ is the vector space of parameter, the generic element $\mathbf{x} \in \Omega$ is called the physical space variable, $\mathcal{V}$ is called control space and the generic element $\mathbf{v} \in \mathcal{V}$ is referred to as the control variable. The norm is denoted by $\|\cdot\|$ in a given vector space and the associated dot product is denoted $\langle \cdot, \cdot \rangle$. The symbol ∇ is the gradient with respect to the control variable and $\nabla_{\mathbf{x}}$ is the gradient with respect to the physical space variable. Without loss of generality, the details are given for the following considerations: the physical space Ω is a rectangular box in R^n, $\mathcal{V}$ is a subspace of $(\mathcal{L}^2(\Omega))^m$ for given positive integers m and n, and $\|\cdot\|$ is the Euclidean norm in R^n or the L^2 norm in $(\mathcal{L}^2(\Omega))^m$.

2.2 Inverse Problems

In most applications of science and engineering, there is a need to find the best approximation $\mathbf{v}^* \in \mathcal{V}$ that satisfies an equation of the form:

$$\mathcal{A}(\mathbf{v}) + \mathbf{e} = \mathbf{f} \tag{1}$$

where $\mathcal{A}$ is an operator (linear or nonlinear) that describes the relationship between the input ($\mathbf{v}$) and the output. In the general case, $\mathbf{v}$ might be constrained to satisfy a given model. For example, the velocity field of a geophysical fluid must satisfied the evolution model of the fluid that is in general nonlinear. This is the general case of the problem that we consider in numerical experiments. The vector $\mathbf{f}$ is the observed output and the vector $\mathbf{e}$ is the observation error. The latter is unknown and seldom sought for. If information on its distribution is available, it can be used in the definition of norms. Given the output $\mathbf{f}$, the problem of finding the input $\mathbf{v}$ that satisfies (1) is known as an inverse problem. Such problems arise in most science and engineering applications like geophysics, remote sensing, image processing, medical imaging, astronomy, etc. Inverse problems are known to be ill-posed in the general case, meaning that at least one of the three conditions of the well-posedness stated by Hadamard [23] is violated: the existence, the uniqueness and the continuity (or the stability) of the solution. Due to the ill-posedness, other information are required to solve an inverse problem. In the present work, we consider the regularization and the projection.

2.3 Regularization

A solution $\mathbf{v}^*$ to the inverse problem associated with (1) is given by:

$$\mathbf{v}^* = \arg\min_{\mathbf{v}\in\mathcal{V}} \mathrm{J}(\mathbf{v}), \tag{2}$$

where the objective (or cost) function J is defined over the control space $\mathcal{V}$ as:

$$\mathrm{J}(\mathbf{v}) = \frac{1}{2}\|\mathcal{A}(\mathbf{v}) - \mathbf{f}\|^2 + \frac{1}{2}\mathrm{J_r}(\mathbf{v}). \tag{3}$$

The first term of the right-hand side of (3) defines the closeness to the observed output; it is generally called the data term and noted $\mathrm{J_d}$ hereafter. The second term, $\mathrm{J_r}(\mathbf{v})$, is the regularization term; it is given by a functional constraint that implicitly restricts the subset of acceptable solutions to those satisfying the constraint. The regularization term is most often given under the form:

$$\mathrm{J_r}(\mathbf{v}) = \int_\Omega c_\Phi(\mathbf{x})\|\Phi(\mathbf{v}(\mathbf{x}))\|^2 d\mathbf{x}, \tag{4}$$

where Φ is a given function and c_Φ is a non negative weighting function. Φ is called regularization function hereafter and the solution given by the associated regularization method is said to be Φ-regularized. Most of the time, the weighting function c_Φ is constant and replaced by a multiplicative scalar. Multiple functional constraints can be combined into the objective function, each one associated with its own weighting function.

The commonly used functional constraints for regularization belong to the family of Tikhonov stabilizing functionals [41] defined as:

$$\mathrm{J_r}(v) = \int \sum_{s=0}^{p} c_s(x)\left(\frac{d^s v}{dx^s}\right)^2 dx, \tag{5}$$

for one dimensional problems where v is a scalar function of one variable x and c_s are the weighting functions. The extension to higher dimensions is straightforward.

A necessary condition for $\mathbf{v}^*$ to be the minimum of the objective function (3) is the Euler–Lagrange equation given by:

$$\nabla \mathrm{J}(\mathbf{v}^*) = 0, \tag{6}$$

where ∇J is the gradient of the objective function with respect to the control variable $\mathbf{v}$. It is implicitly assumed that J is differentiable. The condition (6) is sufficient if J is strictly convex. The convexity of J is improved by the regularization term. Regularization theory for ill-posed problems has been rigorously developed over the past decades. Early development can be found in the work of Tikhonov [41,42] and Nashed [34]. The state of the art on the topic is large and there is no general solution as regularization depends on the problem of interest.

2.4 Explicit Projection

Based on earlier works from Natterer [35] and Louis [26] that demonstrated the regularizing effects of projection methods, Schuster and Weickert [38] used a projection method to solve the inverse problem associated with the problem of motion estimation (see Appendix 1 for an overview of motion estimation). Instead of solving an optimization problem defined by the objective function of the form (3) composed of a data and a regularization term, they define a projected objective function of the form:

$$\mathrm{J}_{\Psi}(\mathbf{v}) = \frac{1}{2}\int_0^{\mathrm{T}}\int_{\Omega} \Psi\left(\langle \mathbf{v}, \nabla f\rangle^2\right) d\mathbf{x}dt, \tag{7}$$

for a single channel image f where $\Psi \in \mathcal{C}^2([0, +\infty])$ is a strictly increasing function and $\Psi(s^2)$ is convex in s. The subset of admissible solutions is restricted to a finite dimensional subspace that defines the projection subspace. The projected problem consists of finding the coefficients of the solution in a properly chosen basis of the projection subspace. Schuster and Weickert derived an explicit convergence rate for arbitrary order of tensor product B-splines basis functions under certain smoothness of the image function.

The explicit projection method as defined in [38] requires a finite dimensional projection space and its basis. However, the authors did not provide the criteria for choosing their subspace in particular, or the projection subspace and its basis for a general inverse problem. As mentioned in the introduction, the regularizing effects of a given projection do not necessarily correspond to those of a regularization, that is why the authors cautioned of choosing the projection subspace carefully since it determines the properties of the solution.

3 From Regularization Functions to Projection

In the present section, we introduce an incremental projection method that is completely different from the explicit projection for solving inverse problems. There is neither requirement of finite dimensional subspace nor a requirement of an explicit basis. In addition we define such projection implicitly such that its range is the subset of Φ-regularized admissible solutions, for any regularization function Φ. The regularizing effects of such a projection are the same as those of the regularization defined by Φ.

3.1 Incremental Projection Method for Inverse Problems

Let $\mathbf{f}$ be the observed output of a physical system that is governed by (1), which is rewritten here for convenience:

$$\mathcal{A}(\mathbf{v}) + \mathbf{e} = \mathbf{f}, \tag{8}$$

Let $\mathcal{V}_{\Phi} \subset \mathcal{V}$ be the subspace of admissible candidate solutions. For instance, $\mathcal{V}_{\Phi}$ is the subspace of Φ-regularized admissible solutions. Let $\mathcal{P}_{\Phi}$ be a projection defined as:

$$\begin{aligned} \mathcal{P}_\Phi : \mathcal{V} &\to \mathcal{V} \\ \mathbf{v} &\mapsto \mathcal{P}(\mathbf{v}), \end{aligned} \tag{9}$$

such that $\mathcal{V}_\Phi$ is the image of $\mathcal{V}$ by $\mathcal{P}_\Phi$. We define the projected solution of the inverse problem associated with (8) as:

$$\mathbf{v}^* = \arg\min_{\mathbf{v}\in\mathcal{V}} \frac{1}{2}\|\mathcal{A}\left(\mathcal{P}_\Phi(\mathbf{v})\right) - \mathbf{f}\|^2. \tag{10}$$

Equation (10) minimizes the the non-regularized objective function associated with the modified problem given by:

$$\mathcal{A}\left(\mathcal{P}_\Phi(\mathbf{v})\right) + \mathbf{e} = \mathbf{f}. \tag{11}$$

In the subspace $\mathcal{V}_\Phi$, the projected problem is equivalent to the original problem. In addition, the regularization condition is already satisfied and the objective function is reduced to the data term (10). In the implementation, the minimization is carried out iteratively, and the projection is applied to each iterate of the solution (see the details in Sect. 3.2), and for that reason, the method is called incremental projection.

As mentioned in the introduction, the subset of regularized admissible solutions is not necessarily a subspace, in the sence of vector space. The method proposed in this paper (see Sect. 3.2) avoids that problem by implicitly defining the projection so that its subspace has all the properties imposed by the regularization constraints. The subspace $\mathcal{V}_\Phi$ as defined in the next section is convex for linear regularization functions Φ and the linearized approximation of nonlinear ones. The convexity is easily verified under the quasi-solution formulation of the regularized cost function (3). The quasi-solution formulation of inverse problems and its relation with regularization can be found in [36] and references therein. An example of quasi-solution formulation is given by (12).

The formulation (11) turns the problem of regularization into the definition of the projection operator $\mathcal{P}_\Phi$.

If $\mathcal{V}_\Phi$ is a finite dimensional subspace, we get a formulation similar to the one in [38]. In the present work, our attention is focused on the subspace of Φ-regularized admissible solutions. However, the method can be extended to any subspace or subset of admissible solutions of the problem.

3.2 Projections Based on Regularization Functions

3.2.1 From Regularization to Projection

Let Ω, $\mathcal{V}$, Φ and φ be respectively the physical space of the system defined by a rectangular subset of R^m, the control space defined as $\mathcal{V} = (\mathcal{L}^2(\Omega))^n$, a regularization function as defined in (4) and a positive scalar function ($\forall \mathbf{x} \in \Omega, \varphi(\mathbf{x}) \geq 0$) of the physical space variable. We define the projection associated with the regularization function Φ as follows:

$$\begin{aligned} \mathcal{P}^0_\Phi : \mathcal{V} &\to \mathcal{V} \\ \mathbf{v} &\mapsto \arg\min_{\mathbf{u}\in\mathcal{V}} \int_\Omega \varphi(\mathbf{x})\|\mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x})\|^2 \mathrm{d}\mathbf{x}, \\ &\quad \text{subject to } \int_\Omega \|\Phi(\mathbf{u}(\mathbf{x}))\|^2 \mathrm{d}\mathbf{x} = 0, \end{aligned} \tag{12}$$

where the function φ describes the prior knowledge of the inverse problem at each point $\mathbf{x}$. φ will be called trust function and must be large where the problem is known to be well-posed and small where it is ill-posed. Any appropriate norm can be used in the place of the L_2 norm. Let $\mathcal{V}^0_\Phi$ be defined as follows:

$$\mathcal{V}^0_\Phi = \{\mathbf{v} \in \mathcal{V} \text{ such that} \|\Phi(\mathbf{v})\| = 0\}.$$

If $\mathcal{V}^0_\Phi$ has the structure of a subspace of $\mathcal{V}$, then $\mathcal{P}^0_\Phi$ is obviously a projection onto $\mathcal{V}^0_\Phi$. In fact, $\forall \mathbf{v} \in \mathcal{V}$, $\mathcal{P}^0_\Phi(\mathbf{v}) \in \mathcal{V}^0_\Phi$ and $\mathcal{P}^0_\Phi(\mathcal{P}^0_\Phi(\mathbf{v})) = \mathcal{P}^0_\Phi(\mathbf{v})$. For example, $\mathcal{V}^0_\Phi$ has the structure of a subspace when the goal of the regularization is to filter out the high frequencies in the solution. In the general case, $\mathcal{V}^0_\Phi$ is not a subspace, however $\mathcal{P}^0_\Phi$ still has all the properties of a projection with the exception that it is not a linear application. In any case, $\mathcal{V}^0_\Phi$ is a convex subset of $\mathcal{V}$ for linear regularization functions.

In practice, the goal of the regularization if not to find a solution which satisfies exactly $\Phi(\mathbf{v}) = 0$, but rather to find a solution $\mathbf{v}^*$ such that $\|\Phi(\mathbf{v}^*)\|$ is small enough. $\mathcal{P}^0_\Phi$ can then be approximated by:

$$\begin{aligned} \mathcal{P}_\Phi : \mathcal{V} &\to \mathcal{V} \\ \mathbf{v} &\mapsto \mathcal{P}_\Phi(\mathbf{v}) = \arg\min_{\mathbf{u}\in\mathcal{V}}(\varepsilon_{\mathbf{v},\Phi}(\mathbf{u})), \end{aligned} \tag{13}$$

with $\varepsilon_{\mathbf{v},\Phi}$ defined as:

$$\begin{aligned} \varepsilon_{\mathbf{v},\Phi} : \mathcal{V}_\Phi &\to R_+ \\ \mathbf{u} &\mapsto \varepsilon_{\mathbf{v},\Phi}(\mathbf{u}) = \frac{1}{2}\int_\Omega \lambda\|\Phi(\mathbf{u}(\mathbf{x}))\|^2 + \varphi(\mathbf{x})\|\mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x})\|^2 \mathrm{d}\mathbf{x}, \end{aligned} \tag{14}$$

where λ is a regularization parameter; If λ is large enough, then $\mathcal{P}_\Phi = \mathcal{V}^0_\Phi$. Hereafter, lambda is set to 1 for the sake of simplicity, and without loss of generality. It is preferable to solve the unconstrained problem defined by (13) and (14) than solving the constrained problem (12), see for example [11,12,14]

To better understand the behavior of $\mathcal{P}_\Phi$, let us consider the numerical iterative process to minimize the objective function (3) without the regularization term. The minimization process converges to a local minimum or converges slowly, or both. The iterate $\mathbf{v}_i$ of the solution is more accurate in the physical region where the problem is known to be well posed. The idea behind the incremental projection is to use $\mathcal{P}_\Phi(\mathbf{v}_i)$, instead of $\mathbf{v}_i$ as the starting point of the next iteration. The problem being well posed in $\mathcal{V}_\Phi$, $\mathcal{P}_\Phi(\mathbf{v}_i)$ which is a Φ-regularized approximation of $\mathbf{v}_i$ is a better approximation of the solution.

Equation (10) equiped with the projection defined in Sect. 3.2.1 gives the solution of the inversed problem as:

$$\mathbf{v}^* = \arg\min_{\mathbf{v}\in\mathcal{V}} \frac{1}{2}\|\mathcal{A}\left(\mathcal{P}_\Phi(\mathbf{v})\right) - \mathbf{f}\|^2, \tag{15}$$

where

$$\mathcal{P}_\Phi(\mathbf{v}) = \arg\min_{\mathbf{v}\in\mathcal{V}} \frac{1}{2}\int_\Omega \|\Phi(\mathbf{u}(\mathbf{x}))\|^2 + \varphi(\mathbf{x})\|\mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x})\|^2 \mathrm{d}\mathbf{x}, \tag{16}$$

which is a least squares optimization problem. The regularization parameter, λ is set to 1 in (16). The operator $\mathcal{P}_\Phi$ is also defined as the solution of another least squares optimization problem, we have an imbrication of least squares problems. There is a wide range of methods to solve optimization problems. The methods for solving the imbricated problem must be compatible with the method for solving the external problem. For example, if the external problem is solved with a gradient-based method, one must be able to provide the Jacobian Matrix of the process that solves the imbricated problem. If we focus on (13), the extended Bregman iterative regularization method [49] can be used. Moreover, if the $l1$ norm is used for the regularization, any iterative shrinkage-thresholding algorithm (ISTA) [4,10,12,16,44,47]) can be used.

In the subsequent part of the paper we focus on the class of gradient descent algorithms for solving least squares problems. The family of gradient descent algorithms is chosen because it is well understood by a broad audience and can easily be used in the case of nonlinear operator $\mathcal{A}$ as well as additional nonlinear model constraints on the solution $\mathbf{v}^*$. More importantly, gradient descent methods are widely used in geophysical applications, and the numerical section of this paper is based on a geophysical application. Since the gradient methods require the Jacobian matrix of $\mathcal{P}_\Phi$, we derive explicit expressions of $\mathcal{P}_\Phi$ for discretized problems. For the continuous cases, we derive formulae for two commonly used regularizations, namely the minimum norm and the minimum gradient.

Gradient descent methods define $\mathbf{v}^*$ as the solution of the Euler–Lagrange equation given by:

$$\nabla \mathrm{J}(\mathbf{v}) = 0, \tag{17}$$

Equation (17) is then solved using an iterative process of the form of algorithm 1.

Input: $\mathbf{v}^0, \mathbf{f}$
// $\mathrm{J}(\mathbf{v}) = \frac{1}{2}\|\mathcal{A}\left(\mathcal{P}_\Phi(\mathbf{v})\right) - \mathbf{f}\|^2$
// $\mathcal{P}_\Phi(\mathbf{v}) = \arg\min_{\mathbf{v}\in\mathcal{V}} \frac{1}{2}\int_\Omega \|\Phi(\mathbf{u}(\mathbf{x}))\|^2 + \varphi(\mathbf{x})\|\mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x})\|^2 \mathrm{d}\mathbf{x}$.
$k \leftarrow 0$;
while *not converged* **do**
  compute $\mathbf{d}^k = -\nabla \mathrm{J}(\mathbf{v}^k)$;
  compute α^k that minimizes $\mathrm{J}(\mathbf{v}^k + \alpha^k \mathbf{d}^k)$;
  $\mathbf{v}^{k+1} \leftarrow \mathbf{v}^k + \alpha^k \mathbf{d}^k$;
  $k \leftarrow k + 1$;
end
Output: $\mathbf{v}^k$

Algorithm 1: Gradient descent algorithm for the modified problem (15)

Each iteration of the algorithm requires the application of the operator $\mathcal{P}_\Phi$ in J and the transpose of its Jacobian matrix in ∇J. Since the projection (13) is defined as the solution of a least squares optimization problem, it can also be turned into the solution of the Euler–Lagrange equation given by:

$$\nabla \varepsilon_{\mathbf{v},\Phi}(\mathbf{u}) = 0. \tag{18}$$

The challenge is to evaluate $\nabla \varepsilon_{\mathbf{v},\Phi}$ for a given Φ.

Proposition 3.1 *Let Φ be a regularization function, $\|.\|$ be the norm defined by the dot product $\langle .,. \rangle$, Φ^* be the adjoint operator of the linearized approximation of Φ associated with the dot product $\langle .,. \rangle$, $\circ$ be the composition operator, and $\varepsilon_{\mathbf{v},\Phi}$ be defined as:*

$$\begin{aligned} \varepsilon_{\mathbf{v},\Phi} : \mathcal{V} &\to R_+ \\ \mathbf{u} &\mapsto \varepsilon_{\mathbf{v},\Phi}(\mathbf{u}) = \frac{1}{2}\int_\Omega \|\Phi(\mathbf{u}(\mathbf{x}))\|^2 + \varphi(\mathbf{x})\|\mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x})\|^2 \mathrm{d}\mathbf{x}. \end{aligned} \tag{19}$$

If $\varepsilon_{\mathbf{v},\Phi}$ is differentiable in $\mathbf{u}$ then $\nabla \varepsilon_{\mathbf{v},\Phi}$ is given by:

$$\nabla \varepsilon_{\mathbf{v},\Phi}(\mathbf{u}) = \Phi^* \circ \Phi(\mathbf{u}(\mathbf{x})) + \varphi(\mathbf{x})(\mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x})), \tag{20}$$

The Proposition (3.1) turns the problem of the evaluation of the gradient of $\nabla \varepsilon_{\mathbf{v},\Phi}$ into the evaluation of $\Phi^* \circ \Phi$ which depends on the regularization function. That is the topic of the next two paragraphs.

3.2.2 Application to Discrete Problems

In most applications, continuous functions are discretized, regularization functions are given by the derivatives of the control variable and can be written under the form of a matrix multiplication using numerical approximations of derivatives. Tikhonov stabilizing functions (5) fall under this category. Under the matrix multiplication form, the regularization function writes:

$$\Phi(\mathbf{v}) = \mathbf{B}\mathbf{v}, \tag{21}$$

with an appropriate matrix $\mathbf{B}$; the same symbol is used for both the continuous and the discretized representation of functions ($\mathbf{v}$ and $\mathbf{u}$). The Euler Lagrange equation (18) rewrites as:

$$\mathbf{B}^T\mathbf{B}\mathbf{u} + \mathbf{D}_\varphi(\mathbf{u} - \mathbf{v}) = 0, \tag{22}$$

where $\mathbf{D}_\varphi$ is a diagonal matrix with the discrete representation of φ on its diagonal. The solution of (22) defines the approximation of the projection $\mathcal{P}$ as:

$$\mathcal{P}(\mathbf{v}) \simeq \left(\mathbf{B}^T\mathbf{B} + \mathbf{D}_\varphi\right)^{-1} \mathbf{D}_\varphi \mathbf{v} \tag{23}$$

The trust function φ must be chosen to enforce the well conditioning of the matrix $\mathbf{B}^T\mathbf{B}+\mathbf{D}_\varphi$. An appropriate direct or iterative method can be used to solve the Eq. (22).

3.2.3 Application to Continuous Problems

In this paragraph, we derive an explicit expression of $\nabla\varepsilon_{\mathbf{v},\Phi}$ for two commonly used regularization functions in the continuous form namely the minimum norm and the minimum gradient.

Minimum norm regularization The regularization term in the minimum norm regularization is defined as:

$$\mathrm{J}_{\mathrm{norm}}(\mathbf{v}) = \int_\Omega \|\mathbf{v}\|^2 \mathrm{d}\mathbf{x}. \tag{24}$$

The associated regularization function Φ_{norm} is the identity so that $\Phi^*_{\mathrm{norm}} \circ \Phi_{\mathrm{norm}}$ is the identity too. The gradient $\nabla\varepsilon_{\mathbf{v},\mathrm{norm}}$ is then given by:

$$\nabla\varepsilon_{\mathbf{v},\mathrm{norm}}(\mathbf{u}(\mathbf{x})) = \mathbf{u}(\mathbf{x}) + \varphi(\mathbf{x})(\mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x})). \tag{25}$$

The solution of the Euler Lagrange equation defines the projection as:

$$\mathcal{P}_\Phi(\mathbf{v}(\mathbf{x})) = \frac{\varphi(\mathbf{x})}{1+\varphi(\mathbf{x})}\mathbf{v}(\mathbf{x}). \tag{26}$$

Minimum gradient regularization The gradient regularization term is given by:

$$\mathrm{J}_{\mathrm{grad}}(\mathbf{v}) = \sum_{i=1}^{n} \int_\Omega \|\nabla_{\mathbf{x}} v_i(\mathbf{x})\|^2 \mathrm{d}\mathbf{x}, \tag{27}$$

which can be considered as a multi-term regularization, a simple term being:

$$\mathrm{J}_{\mathrm{grad}}(v_i) = \int_\Omega \|\nabla_{\mathbf{x}} v_i(\mathbf{x})\|^2 \mathrm{d}\mathbf{x}, \, 1 < i < n. \tag{28}$$

The simple term regularization function is then defined as:

$$\Phi_{\mathrm{grad}}(v_i) = \nabla_{\mathbf{x}} v_i(\mathbf{x}), \, 1 < i < n. \tag{29}$$

It is possible to define a single trust function φ for all terms or a special one for each term. In the present work, we use the same trust function.

Proposition 3.2 *Let $\mathcal{V} = (L^2(\Omega))^n$ with Ω a rectangular subdomain of R^m and Φ be the linear regularization function defined by:*

$$\begin{aligned} \Phi : \mathcal{V} &\to R_+ \\ \mathbf{u} &\mapsto \Phi(\mathbf{u}) = \nabla_{\mathbf{x}} u_i, \, 1 \le i \le n, \end{aligned} \tag{30}$$

with homogeneous Neumann conditions on the boundaries, the operator $\Phi^* \circ \Phi$ *is given by:*

$$\Phi^* \circ \Phi = -\nabla^2_{\mathbf{x}}. \tag{31}$$

By Proposition 3.2, the restriction of $\Phi^*_{\text{grad}} \circ \Phi_{\text{grad}}$ to the ith component of the vector field is given by:

$$\Phi^*_{\text{grad}} \circ \Phi_{\text{grad}} = -\nabla^2_{\mathbf{x}}, \tag{32}$$

and the gradient of the energy functional $\varepsilon_{\mathbf{v},\text{grad}}$ is given by:

$$\nabla \varepsilon_{\mathbf{v},\text{grad}}(u_i) = -\nabla^2_{\mathbf{x}} u_i(\mathbf{x}) + \varphi(\mathbf{x})(u_i(\mathbf{x}) - v_i(\mathbf{x})),\ 1 < i < n. \tag{33}$$

The projection itself is the solution of the Euler–Lagrange equation given by:

$$\nabla^2 u_i - \varphi(\mathbf{x})(u_i(\mathbf{x}) - v_i(\mathbf{x})) = 0, \quad 1 \leq i \leq n. \tag{34}$$

Equation (34) can be solved by any appropriate methods. We give in the next few lines the outlines of one methods. According to the development in [48], (34) can be solved by considering u_i as a function of time as follows:

$$\begin{aligned} \frac{\partial}{\partial t} u_i(\mathbf{x}, t) &= \nabla^2 u_i(\mathbf{x}, t) - \varphi(\mathbf{x})(u_i(\mathbf{x}, t) - v_i(\mathbf{x})), \quad 1 \leq i \leq n, \\ u_i(\mathbf{x}, 0) &= v_i(\mathbf{x}). \end{aligned} \tag{35}$$

The set of equations (35) is known as the generalized diffusion equations. It gives an asymptotic solution to (34). Numerical experiments in Sect. 4 are based on that implementation for gradient regularization function. The iterative process (35) is truncated to a fixed number (50 in this paper) and algorithmic differentiation is used in the computation of the gradient of J in Algorithm 1.

4 Numerical Experiments: motion estimation for Geophysical Fluid Images

We applied the proposed incremental projection method to the problem of motion estimation with 2D fluid images. Appendix 1 gives an overview of motion estimation. The optical flow constraint is replaced by the convection diffusion equation that is well suited for fluid flows. The projection operators for the minimum norm and the minimum gradient regularization functions are implemented and compared with the associated regularization methods. The same optimization toolbox is used for all experiments. We ensure that the optimal setup is used in each case. For instance, we consider constant weighting functions and determine the optimal weighting parameter for each regularization problem, see Appendix "Selection of the Regularization Weighting Parameter" section.

4.1 Setup of Numerical Experiments

4.1.1 Trust Function for Motion Estimation

The trust function is defined by the contours map. Following [48], the contour map of a gray-level image function $\mathrm{f}(\mathbf{x})$ takes one of the following forms:

$$c^1(\mathbf{x}, \mathrm{f}) = \|\nabla_{\mathbf{x}} \mathrm{f}(\mathbf{x})\|^2 \text{ or,} \tag{36}$$

$$c^2(\mathbf{x}, \mathrm{f}) = \|\nabla_{\mathbf{x}} (\mathbf{G}_\sigma(\mathbf{x}) * \mathrm{f}(\mathbf{x}))\|^2, \tag{37}$$

where $\mathbf{G}_\sigma(\mathbf{x})$ is a Gaussian function with standard deviation σ and $*$ is the convolution operator. The Gaussian convolution is used for denoising purpose. The contour map is chosen because it accounts for the aperture problem in motion estimation which states that: only the motion along the normal to iso-contours can be inferred from an image sequence (see for example [30] and references therein). It has the effect of keeping $\mathcal{P}_\Phi(\mathbf{v})$ close to $\mathbf{v}$ on contours and forces it to satisfy the regularization condition in homogeneous areas. In the present work, the trust function is normalized for the minimum gradient projection so that $0 \leq \varphi(\mathbf{x}) \leq 1$. For the minimum norm projection, the normalized trust function is scaled by a large factor (10^3) to ensure that $\mathcal{P}_\Phi(\mathbf{v})$ is close to $\mathbf{v}$ where φ is large.

4.1.2 Experimental Data

The image sequence used in the present work comes from a laboratory experiment on the drift of an isolated vortex (see [20] for details on that experiment). Figure 1 shows the true motion field superimposed on the first image of the sequence. Only

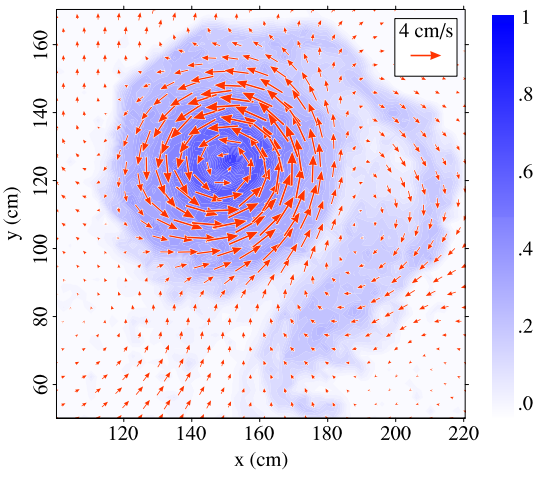

Fig. 1 True velocity field superimposed on the first image

the 60×60 pixels most active region of the image domain is shown. The full domain is 256.512cm $\times$ 256.512cm for an image of 128×128 pixels. The flow is made up of two vortices: a major one rotating counterclockwise and a minor one rotating clockwise. The true motion field is computed by direct assimilation of image sequence, a technique recently introduced for the use of images in data assimilation (see [43]). The computed motion field is an approximation of the truth. In order to have an image sequence and the associated true motion field, a second image is computed by convection-diffusion of the first image with the motion field. The first image and the computed image make up the experimental image sequence. The goal of motion estimation is to recover the velocity field from the experimental image sequence. The optimization algorithm used for numerical experiments is M1QN3, a quasi-Newton algorithm from the MODULOPT library [21].

4.2 Comparison Between Projection and Regularization

The present subsection presents a comparative study of the evolution of the minimization process. The comparison is based on the velocity error, the angular error and the vorticity error. The vorticity error is not frequently used in motion estimation. It is included in the present study because of the importance of the vorticity in fluid flows. The vorticity quantifies the local spinning of a fluid particle around a nearby point. Its computation is straightforward compared with the angular error that uses approximation. All the error functions are normalized. In the present study, the normalization is the division by the same function evaluated at the initial guess. Appendix "Diagnostic Functions" section gives the details of the computation of those error functions. The mean and the maximum are computed over the 60×60 pixels most active subdomain shown in Fig. 1.

4.2.1 Minimum Norm Regularization and Projection

Figure 2 shows the motion field estimated using the minimum norm regularization and projection as well as the true motion. An interpretation is that the minimum norm regularization is not adapted to motion estimation. The solution of the the projection method captures the main feature of the truth (the two vortices) but underestimates the order of magnitude. Figure 3 shows the evolution of the normalized error functions during the minimization. The objective function is not shown because it is completely different for both methods. The regularization methods converges after 250 iterations. At that point, the regularization term dominates the objective function and the optimization algorithm can not find a descent direction or step. After 250 iterations, the mean velocity error is reduced by 40 % for the regularization method, and 58 % for the projection method. The regularization method reduces the error by 40 % after 250 iterations while the projection method reaches the same target after 16 iterations. That is 15.6 times less iterations for the projection method. In terms of maximum velocity error, the difference is insignificant between the two methods. The vorticity error (mean and maximum) increases over the iterations for both methods with an advantage to the projection. It means that the estimation of the local spinning is inaccurate.

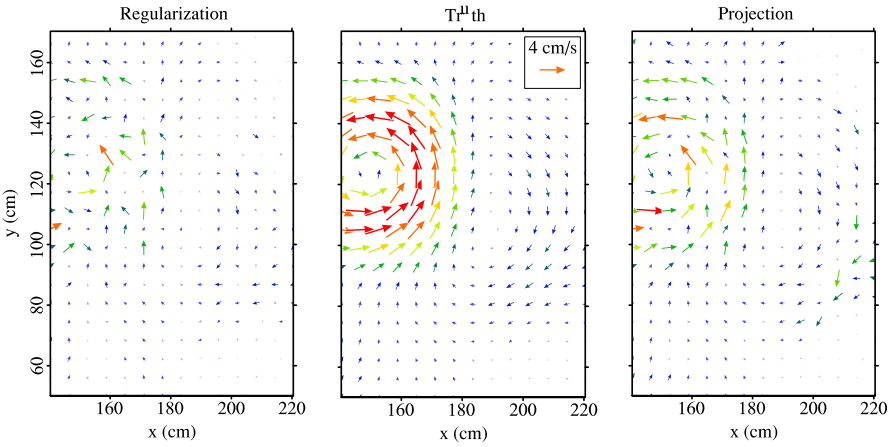

Fig. 2 Velocity field for the minimum norm regularization and projection

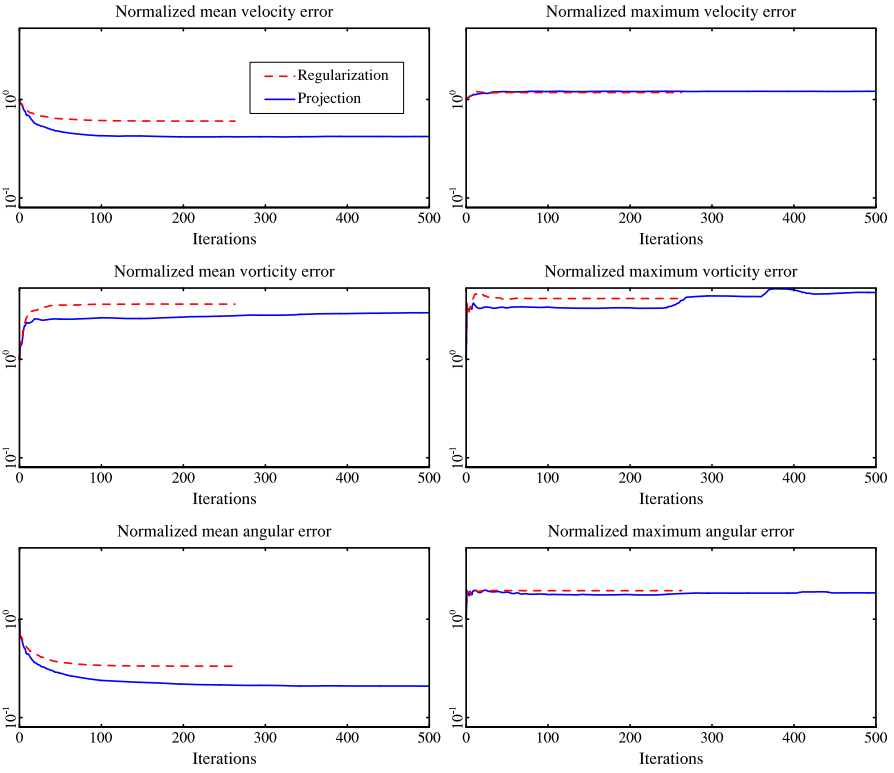

Fig. 3 Evolution of errors function during the minimization process for the minimum norm regularization and projection

Table 1 Summary of converged error

Method	$\bar{e}^{v}$	$\bar{e}^{\mathrm{vr}}$	$\bar{e}^{\mathrm{ag}}$	$e^{\mathrm{ag}}(^\circ)$
Normalized mean error (%) and mean angular error (°)				
Minimum norm regularization	60	364	33	29
Minimum norm projection	42	297	21	18
Minimum gradient regularization	9	77	3	2
Minimum gradient projection	0.1	0.3	0.1	0.1
Normalized maximum error (%) and maximum angular error (°)				
Minimum norm regularization	117	414	195	174
Minimum norm projection	121	477	185	165
Minimum gradient regularization	27	148	27	24
Minimum gradient projection	0.4	0.4	9	8

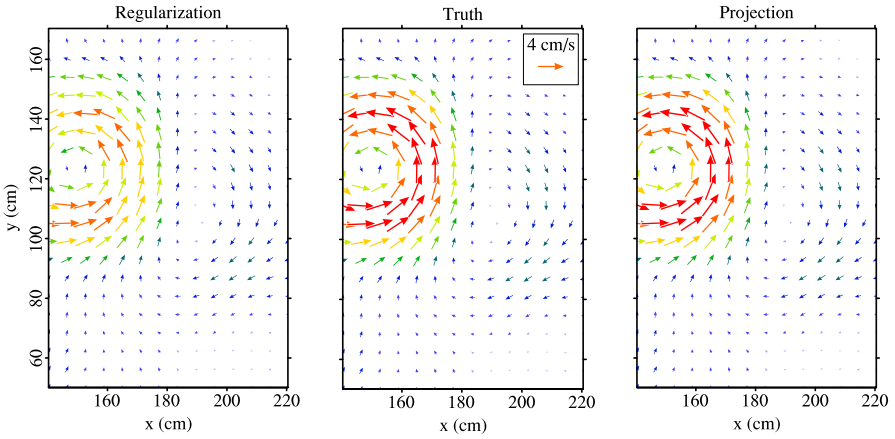

Fig. 4 Velocity field for the minimum gradient regularization and projection

The angular error (mean and maximum) shows similar behavior to the velocity error. In terms of computation time per iteration, the projection method performs slightly better than the regularization. That is not surprising; the difference between the two methods is that the projection performs a vector scaling at each iteration while the regularization computes the regularization term that is more expensive. Overall, the projection yield fast iterations, rapid convergence and better results at convergence. Table 1 gives a summary of diagnostic functions when the minimization algorithm is stopped.

4.2.2 Minimum Gradient Regularization and Projection

Figure 4 shows the motion field estimated using minimum gradient regularization and projection. The regularization method globally underestimates the magnitude because of the smoothing effect of the regularization term. The solution of the projection

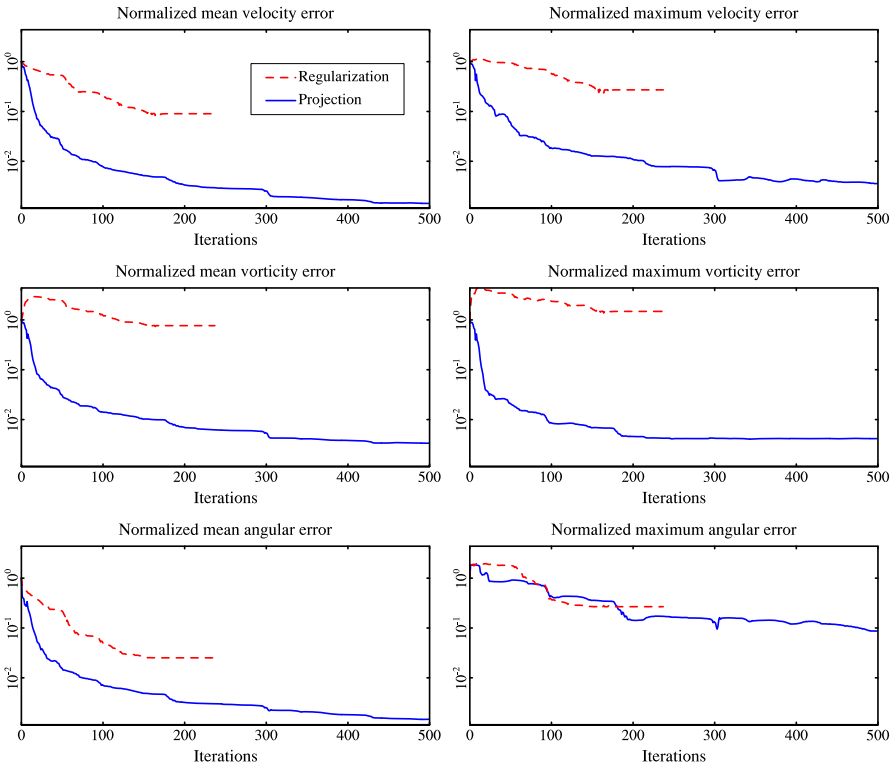

Fig. 5 Evolution of error during the minimization process for the minimum gradient regularization and projection

method is visually indistinguishable from the truth. Figure 5 shows the evolution of the normalized error functions during the minimization. The regularization methods converges after 164 iterations. At that point, the regularization term dominates the objective function and the optimization algorithm can not find a descent direction or step. After 164 iterations, the mean velocity error is reduced by ∼91 % for the regularization method, and 99.9 % for the projection method. The regularization method reduces the error by 91 % after 164 iterations while the projection method reaches the same target after 17 iterations. That corresponds to almost 10 times less iterations than the regularization method to reach the same target. After 164 iterations, the maximum velocity error is reduced by 73 % for the regularization method and 99.6 % for the projection method. In terms of vorticity, the regularization method decreases the mean error by 23 % and increases the maximum error while the projection method decreases both the mean error and the maximum error by 99.7 and 99.6 % respectively. The regularization is inaccurate in the estimation of the local spinning. In terms of angular error, the mean is reduced by 97 % and the maximum is reduced by 73 % for the regularization method; for the projection, the mean is reduced by 99.9 % and the maximum is reduced by 92 %. Both methods get the right orientation of the velocity fields, but only the projection gets the right amplitude. It is also important to notice

that the projection method globally improves each indicator from the first iteration, while the regularization method first deteriorates the vorticity over a large range of iteration, and does not improve the maximum vorticity error. In terms of computation time per iteration, the regularization method is approximately 1.5 faster than the projection and it takes 10 times more iterations than the projection method to reduce the velocity error by 91 %. Globally, the projection methods takes 6.7 times less computation time to reduce the velocity error by 91 %. Overall, there is a net advantage of the incremental projection method compared to the regularization. Table 1 gives a summary of diagnostic functions when the minimization algorithm is stopped.

Experiments with the minimum gradient projection and constant trust function ($\varphi(\mathbf{x}) = 1$), not included in this document, shows similar behavior to the projection with the trust function (36): faster convergence and final result closer to the truth than the regularization. However, the projection with the variable trust function (36) is better than the projection with constant trust function.

5 Conclusion

In this study, we introduced an incremental projection method as an alternative to the regularization of inverse problems. The proposed method is based on the projection onto the subspace of *regularized* admissible solutions. The projection replaces the use of a regularization term in the objective function and significantly improves the convergence of optimization algorithms both in terms of speed and the error in the final solution. The proposition is supported by numerical experiments that demonstrate its feasibility and the improvement compared with regularization methods. The example problem shows that for the minimum gradient regularization, the incremental projection method uses 6.7 times less computation time than the regularization method to achieve the same reduction of the error. At convergence, the incremental projection result is 90 times closer to the exact solution than the regularized solution is. It makes the proposed formulation a good alternative approach to regularization methods. Further investigations of this ongoing work include the theoretical study of convergence and the construction of trust functions.

Acknowledgments Part of this work was sponsored by the Office of Naval Research Program Element 0601153N as part of the "Variational Data Assimilation for Ocean Prediction" and "A Multiscale Approach to Assessing Predictability of ASW Environment" projects. This paper is NRL paper contribution number NRL/JA/7320-14-xxx. Francois-Xavier LeDimet and Arthur Vidard were sponsored by the French National Research Agency through projects ADDISA and VODA

Appendix 1: Optical Flow Motion Estimation

The motion estimation is an example of inverse problem that consists in determining the vector field that transports one frame to the next in a sequence of images under the *optical flow* constraint which states that the luminance function is conserved over time. Let $\Omega \subset R^2$ be a rectangular image domain and f a given image function defined from $\Omega \times [0, T]$ to R. The *optical flow* constraint defines the motion vector field $\mathbf{v}$ as the solution of the following equation:

$$\frac{\partial \mathrm{f}^{\mathrm{t}}}{\partial t} + \nabla \mathrm{f}^{\mathrm{t}} \cdot \mathbf{v} = 0, \quad (\mathbf{x}, t) \in \Omega \times [0, T]. \tag{38}$$

The control space is defined as $\mathcal{V} = (L^2(\Omega))^2$. Motion estimation is known to be subject to the *aperture problem*: it is impossible to determine the motion where the gradient of the image function is null, or in the direction orthogonal to the gradient of the image function (see [30] and references therein). This is the main source of ill-posedness that induces the necessity of regularization. The ill-posedness of inverse problems in image processing are detailed in [5]. Given the observation f of the image function, the solution $\mathbf{v}$ to (38) can be estimated through a minimization problem. To this end, the objective function J is defined as follows:

$$\mathrm{J}(\mathbf{v}) = \frac{1}{2} \int_0^{\mathrm{T}} \left\| \frac{\partial \mathrm{f}}{\partial t} + \nabla \mathrm{f} \cdot \mathbf{v} \right\|_{\Omega}^2 dt + \alpha_{\mathrm{r}} \mathrm{J}_{\mathrm{r}}(\mathbf{v}) \tag{39}$$

To turn (39) into the form of the objective function (3), one considers f as an implicit function of $\mathbf{v}$ given by the solution of the *optical flow* constraint (38). The optical flow approach of motion estimation is well addressed in the literature; the interested reader can consider [19,24,32,46] for more detailed survey. Since the work of Horn and Schunck [24], many propositions have been made for the regularization of the optical flow.

Additional techniques to improve the solution of motion estimation problems include the discretization scheme such as Helmhotz decomposition [50], the use of L_1-norm [45,51], and multi-scale strategies [1,3,6,8,9,15,22,27,31] and references therein.

Appendix 2: Diagnostic Functions and Selection of the Regularization Weighting Parameter

Diagnostic Functions

The following functions are used for the evaluation of the estimated velocity field: velocity error, vorticity error, angular error, objective function and its data term. The commonly used functions are the velocity error and the angular error. The vorticity error is introduced in the present work because of its importance in fluid flows. It quantifies the local spinning of a fluid particle around a nearby point and its computation is straightforward compared with the angular error that uses approximation.

The physical domain Ω, is a rectangular subset of R^2 and the physical space variable is given by $\mathbf{x} = (x_1, x_2)$. The control space is $\mathcal{V} = (\mathcal{L}^2(\Omega))^2$ and the control variable (transport velocity field) is of the form $\mathbf{v}(\mathbf{x}) = (v_1(\mathbf{x}), v_2(\mathbf{x}))$. Let vr denote the vorticity of $\mathbf{v}$, and the superscripts $^{\mathrm{a}}$, $^{\mathrm{t}}$ and $^{\mathrm{b}}$ stand for the analysis (estimated solution), the truth and the background (initial guess) respectively; the velocity error ($e^{\mathbf{v}}$), the vorticity error (e^{vr}) and the angular error (e^{ag}) are defined as follows:

$$e^{\mathbf{v}}(\mathbf{x}) = \|\mathbf{v}^*(\mathbf{x}) - \mathbf{v}^{\mathrm{t}}(\mathbf{x})\|; \tag{40}$$

$$e^{\mathrm{vr}}(\mathbf{x}) = |\mathrm{vr}^{\mathrm{a}}(\mathbf{x}) - \mathrm{vr}^{\mathrm{t}}(\mathbf{x})|; \tag{41}$$

$$e^{\mathrm{ag}}(\mathbf{x}) = |\mathrm{ag}\left(\mathbf{v}^{\mathrm{t}}(\mathbf{x}), \mathbf{v}^*(\mathbf{x})\right)|; \tag{42}$$

where the vorticity of a vector field $\mathbf{v}$ is defined as:

$$\mathrm{vr}(\mathbf{x}) = \frac{\partial v_2}{\partial x_1}(\mathbf{x}) - \frac{\partial v_1}{\partial x_2}(\mathbf{x}). \tag{43}$$

The angle between $\mathbf{v}^{\mathrm{t}}$ and $\mathbf{v}^*$ if defined as follows:

$$\mathrm{ag}(\mathbf{x}) = \arccos\left(\frac{\left\langle \mathbf{w}^{\mathrm{a}}(\mathbf{x}), \mathbf{w}^{\mathrm{t}}(\mathbf{x})\right\rangle}{\|\mathbf{w}^{\mathrm{a}}(\mathbf{x})\|\|\mathbf{w}^{\mathrm{t}}(\mathbf{x})\|}\right), \tag{44}$$

where $\mathbf{w}(\mathbf{x}) = (v_1(\mathbf{x}), v_2(\mathbf{x}), \epsilon)$, with $\epsilon \ll 1$, a positive scalar, chosen to be $\epsilon = \frac{1}{100}\max(\|\mathbf{v}^{\mathrm{t}}(\mathbf{x})\|)$, $\mathbf{x} \in \Omega$ in the present work. The parameter ϵ makes it possible to compute the angle if one of the vectors ($\mathbf{v}^{\mathrm{t}}$ or $\mathbf{v}^*$) is null. The Eq. (44) is an improvement of the commonly used angular error where ϵ is set to 1 [2,17,18]. Setting ϵ to 1 induces large error in the estimated angular error when vectors ($\mathbf{v}^{\mathrm{t}}$ or $\mathbf{v}^*$) have small norms.

In each case, we define the mean error e_{mean} (mean over the physical domain) and the normalized mean error $\bar{e}_{\mathrm{mean}}$ as:

$$e_{\mathrm{mean}} = \frac{1}{\mathrm{mes}(\Omega)}\left(\int_{\Omega} e(\mathbf{x})\mathrm{d}\mathbf{x}\right), \quad \bar{e}_{\mathrm{mean}} = \frac{e_{\mathrm{mean}}}{e^{\mathrm{b}}_{\mathrm{mean}}}, \tag{45}$$

where $\mathrm{mes}(\Omega)$ is the measure of Ω and e^{b} is the error on the background (initial guess). e_{mean} defines the continuous form of the root mean squares error (RMSE) and $\bar{e}_{\mathrm{mean}}$ defines the normalized RMSE (NRMSE). The maximum error e_{max} and the normalized maximum error are defined as:

$$e_{\mathrm{max}} = \max_{\mathbf{x}\in\Omega} e(\mathbf{x}) \text{ and } \bar{e}_{\mathrm{max}} = \frac{e_{\mathrm{max}}}{e^{\mathrm{b}}_{\mathrm{max}}}. \tag{46}$$

The normalized objective function is defined as:

$$\bar{\mathrm{J}} = \frac{\mathrm{J}}{\mathrm{J}^{\mathrm{b}}}, \tag{47}$$

where J^{b} is the value of the objective function associated with the background. In the present work, the background (initial guess) $\mathbf{v}^{\mathrm{b}}$ is set to zero ($\mathbf{v}^{\mathrm{b}} = 0$).

Selection of the Regularization Weighting Parameter

To select the best weighting parameter, we carried out the error analysis with varying regularization parameter, see Fig. 6 for the example of minimum gradient regularization. The optimal weighting parameter is chosen to be the point where the mean

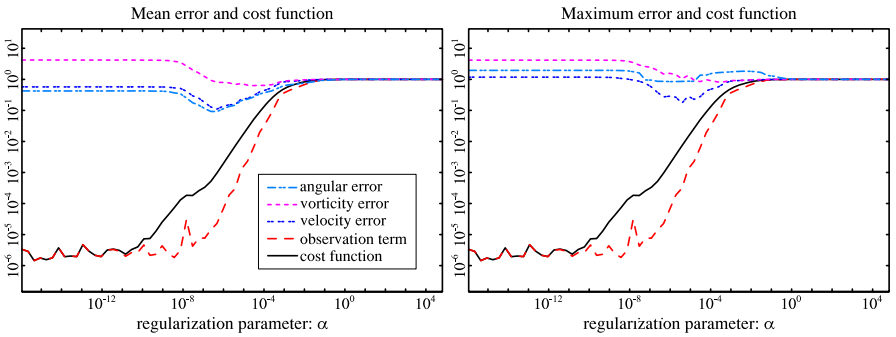

Fig. 6 Minimum gradient regularization: evolution of converged diagnostic functions with the regularization weighting parameter (logarithmic scale)

velocity error reaches its minimum value. In the case of minimum norm regularization, there is no valley as in the case of minimum gradient. The curves grow smoothly from the small values of the weighting parameter, where the data term dominates the objective function to the large values, where the regularization term dominates. Experiments with other regularizations for motion estimation (first and second-order divergence or curl regularization, isotropic image/flow-driven regularization) on the same data shows similar behavior to the minimum gradient regularization. The error analysis method was chosen because of the availability of the ground truth. A selection of scientific works on the estimation of the weighting parameter in the absence of the truth includes [7,13,25,28,29,33,37,39,40] and references therein.

References

1. Anandan, P.: A computational framework and an algorithm for the measurement of visual motion. Int. J. Comput. Vis. **2**, 283–310 (1989). doi:10.1007/BF00158167
2. Barron, J.L., Fleet, D.J., Beauchemin, S.S.: Performance of optical flow techniques. Int. J. Comput. Vis. **12**, 43–77 (1994)
3. Battiti, R., Amaldi, E., Koch, C.: Computing optical flow across multiple scales: an adaptive coarse-to-fine strategy. Int. J. Comput. Vis. **6**, 133–145 (1991). doi:10.1007/BF00128153
4. Beck, Amir, Teboulle, Marc: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**, 183–202 (2009)
5. Bertero, M., Pogcio, T.A., Torre, V.: Ill-posed problems in early vision. Proc. IEEE **76**, 869–889 (1988)
6. Black, M.J., Anandan, P.: The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. Comput. Vis. Image Underst. **63**, 75–104 (1996)
7. Bonesky, T.: Morozov's discrepancy principle and tikhonov-type functionals. Inverse Probl. **25**, 015015 (2009)
8. Bruhn, A., Weickert, J.: A multigrid platform for real-time motion computation with discontinuity-preserving variational methods. Int. J. Comput. Vis. **70**, 257–277 (2006)
9. Bruhn, A., Weickert, J., Schnörr, C.: Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods. Int. J. Comput. Vis. **61**, 211–231 (2005)
10. Chambolle, A., De Vore, R.A., Lee, N.Y., Lucier, B.J.: Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage. IEEE Trans. Image Process. **7**, 319–335 (1998)
11. Combettes, P., Pesquet, J.: Proximal thresholding algorithm for minimization over orthonormal bases. SIAM J. Optim. **18**, 1351–1376 (2008)

12. Daubechies, I., Defrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun. Pure Appl. Math. **57**, 1413–1457 (2004)
13. Eggermont, P.P.B., LaRiccia, V.N., Nashed, M.Z.: On weakly bounded noise in ill-posed problems. Inverse Probl. **25**, 115018 (2009)
14. Elad, M., Matalon, B., Zibulevsky, M.: Coordinate and subspace optimization methods for linear least squares with non-quadratic regularization. Appl. Comput. Harmon. Anal. **23**, 346–367 (2007)
15. Enkelmann, W.: Investigations of multigrid algorithms for the estimation of optical flow fieldsin image sequences. Comput. Vis. Graph. Image Process. **43**, 150–177 (1988)
16. Figueiredo, M.A.T., Nowak, R.D.: An em algorithm for wavelet-based image restoration. IEEE Trans. Image Process. **12**, 906–916 (2003)
17. Fleet, D.J.: Measurement of Image Velocity. Kluwer Academic Publishers, Dordrecht (1992)
18. Fleet, D.J., Jepson, A.D.: Computation of component image velocity from local phase information. Int. J. Comput. Vis. **5**, 77–104 (1990)
19. Fleet, D.J., Black, M.J., Yacoob, Y., Jepson, A.D.: Design and use of linear models for image motion analysis. Int. J. Comput. Vis. **36**, 171–193 (2000)
20. Flór, J.-B., Eames, I.: Dynamcis of monopolar vortices on a topographic beta-plane. J. Fluid Mech. **456**, 353–376 (2002)
21. Gilbert, J.C., Jonsson, X.: LIBOPT—an environment for testing solvers on heterogeneous collections of problems—version 1.0. Technical Report RT-0331, INRIA (2007)
22. Glazer, F., Reynolds, G., Anandan, P.: Scene matching by hierarchical correlation. In: Proceedings of CVPR, pp. 432–441 (1983)
23. Hadamard, J.: Sur les problèmes aux dérivés partielles et leur signification physique. Princet. Univ. Bull. **13**, 49–52 (1902)
24. Horn, B.K.P., Schunck, B.G.: Determining optical flow. Artif. Intell. **59**, 81–87 (1981)
25. Lepskiĭ, O.V.: A problem of adaptive estimation in gaussian white noise. Theory Probab. Appl. **35**, 454–466 (1990)
26. Louis, A.K.: Inverse und Schlecht Gestellte Probleme, Teubner Studienbücher Mathematik. B.G. Teubner, Stuttgart (1989)
27. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: Proceedings of the 7th International Joint Conference on Artificial Intelligence—Volume 2, IJCAI'81, pp. 674–679. Morgan Kaufmann Publishers Inc., San Francisco (1981)
28. Mathé, P.: The lepskii principle revisited. Inverse Probl. **22**, L11 (2006)
29. Mathé, P., Tautenhahn, U.: Regularization under general noise assumptions. Inverse Probl. **27**, 035016 (2011)
30. McDermott, J., Weiss, Y., Adelson, E.H.: Beyond junctions: nonlocal form constraints on motion interpretation. Perception **30**, 905–23 (2001)
31. Mémin, E., Pérez, P.: Hierarchical estimation and segmentation of dense motion fields. Int. J. Comput. Vis. **46**, 129–155 (2002)
32. Mitiche, A., Bouthemy, P.: Computation and analysis of image motion: a synopsis of current problems and methods. Int. J. Comput. Vis. **19**, 29–55 (1996)
33. Morozov, V.A., Nashed, Z.: Methods for Solving Incorrectly Posed Problems. Springer-Verlag, New York (1984)
34. Nashed, M.Z.: Approximate regularized solutions to improperly posed linear integral and operator equations. In: Colton, D.L., Gilbert, R.P. (eds.) Constructive and Computational Methods for Differential and Integral Equations. Lecture Notes in Mathematics, vol. 430, pp. 289–332. Springer, Berlin (1974)
35. Natterer, F.: Regularisierung schlecht gestellter Probleme durch Projektionsverfahren. Numer. Math. **28**, 329–341 (1977)
36. Poggio, T., Koch, C.: Ill-Posed problems in early vision: from computational theory to analogue networks. Proc. R. Soc. Lond. B Biol. Sci. **226**, 303–323 (1985)
37. Ramlau, R.: Morozov's discrepancy principle for tikhonov regularization of nonlinear operators. Numer. Funct. Anal. Optim. **23**, 2002 (2001)
38. Schuster, T., Weickert, J.: On the application of projection methods for computing optical flow fields. Inverse Probl. Imaging **1**, 673–690 (2007)
39. Schuster, T., Kaltenbacher, B., Hofmann, B., Kazimierski, K.S.: Tikhonov regularization of linear operators with power-type penalties. In: Regularization Methods in Banach Spaces. Radon Series on Computational and Applied Mathematics, pp. 108–142. Walter de Gruyter, Berlin (2012)

40. Simpson, I.J.A., Woolrich, M.W., Groves, A.R., Schnabel, J.A.: Longitudinal brain MRI analysis with uncertain registration. In: Proceedings of the 14th International Conference on Medical Image Computing and Computer-Assisted Intervention—Volume Part II, MICCAI'11, pp. 647–654. Springer-Verlag, Berlin (2011)
41. Tikhonov, A.N.: Regularization of incorrectly posed problems. Sov. Math. **4**, 1624–1627 (1963)
42. Tikhonov, A.N., Arsenin, V.Y.: Solutions of Ill-Posed Problems. W.H. Winston, New York (1977)
43. Titaud, O., Vidard, A., Souopgui, I., Le Dimet, F.-X.: Assimilation of image sequences in numerical models. Tellus A **62**, 30–47 (2010)
44. Vonesch, C., Unser, M.: A fast iterative thresholding algorithm for wavelet-regularized deconvolution. In: Proceedings of the SPIE Conference on Mathematical Imaging: Wavelet XII, vol. 6701, pp. 67010D-1–67010D-5, San Diego, August 26–29, 2007
45. Wedel, A., Pock, T., Zach, C., Bischof, H., Cremers, D.: An improved algorithm for TV-L1 optical flow. In: Cremers, D., Rosenhahn, B., Yuille, A.L., Schmidt, F.R. (eds.) Statistical and Geometrical Approaches to Visual Motion Analysis, pp. 23–45. Springer-Verlag, Berlin (2009)
46. Weickert, J., Bruhn, A., Brox, T., Papenberg, N.: Mathematical Models for Registration and Applications to Medical Imaging. Mathematics in Industry, p. 103. Springer, Berlin (2006)
47. Wright, Stephen J., Nowak, Robert D., Mário, A.T.: Figueiredo, Sparse reconstruction by separable approximation. Trans. Signal Process. **57**, 2479–2493 (2009)
48. Xu, C., Prince, J.L.: Snakes, shapes, and gradient vector flow. TIP **7**, 359–369 (1998)
49. Yin, W., Osher, S., Goldfarb, D., Darbon, J.: Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing. SIAM J. Imaging Sci. **1**, 143–168 (2008)
50. Yuan, J., Schnörr, C., Mémin, E.: Discrete orthogonal decomposition and variational fluid flow estimation. J. Math. Imag. Vis. **28**, 67–80 (2007)
51. Zach, C., Pock, T., Bischof, H.: A duality based approach for realtime tv-l1 optical flow. In: Proceedings of the 29th DAGM Conference on Pattern Recognition, pp. 214–223. Springer-Verlag, Berlin (2007)