MOX-Report No. 25/2020

Bayesian mesh adaptation for estimating distributed parameters

Calvetti, D.; Cosmo, A.; Perotto, S.; Somersalo, E.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it http://mox.polimi.it

Bayesian Mesh Adaptation for Estimating Distributed Parameters

D Calvetti[1] A Cosmo[2*] S Perotto[2] E Somersalo[1]

[1]Department of Mathematics, Applied Mathematics and Statistics
Case Western Reserve University
[2] MOX - Dipartimento di Matematica
Politecnico di Milano, Italy

Abstract

The problem of estimating numerically a distributed parameter from indirect measurements arises in many applications, and in that context the choice of the discretization plays an important role. In fact, to guarantee a certain level of accuracy of the forward model that maps the unknown to the observations may require a fine discretization, adding to the complexity of the problem and to the computational cost. On the other hand, reducing the complexity of the problem by adopting a coarser discretization may increase the modeling error and can be very detrimental for ill-posed inverse problems. To balance accuracy and complexity, we propose an adaptive algorithm for adjusting the discretization level automatically and dynamically while estimating the unknown distributed parameter by an iterative scheme. In the Bayesian paradigm, all unknowns, including the metric that defines the discretization, are modeled as random variables. Our approach couples the discretization with a Bayesian hierarchical hyperparameter that is estimated simultaneously with the unknown parameter of primary interest. The viability of the proposed algorithm, the Bayesian Mesh Adaptation (BMA) is assessed on two test cases, a fan-beam X-ray tomography and an inverse source problem for a Darcy flow model.

1 Introduction

In a wide range of applications in, e.g., medical imaging, non-destructive material testing, geophysical sounding or remote sensing, the estimation of a spatially distributed parameter from indirect and noisy measurements is required. Approximating the distributed parameter in terms of a mesh-based basis provides a natural discretization scheme, in particular when the connection between the observed quantity and the unknown parameter is governed by partial differential or integral equations. An example of the former is electrical impedance tomography, where the unknown conductivity distribution inside a body is to be estimated from boundary values of the solution of the second order elliptic equation for the voltage potential [13], while an example of the latter is the X-ray tomography problem where the unknown density distribution inside the body is related to the data through the Radon transform [29].

The choice of the density of the mesh where to represent the unknown is the result of a trade-off between stability and resolution on one hand, and accuracy and complexity of the inverse problem on the other. In fact, while the reduction of the degrees of freedom attained with a coarse mesh representation may improve the stability of the discretized inverse problem, this comes at the potential loss of spatial resolution. An even bigger concern, however, is the fact that a coarse mesh discretization may introduce a significant discrepancy between the presumably accurate continuous model and the discretized one. The corresponding modeling error tends to behave like correlated noise

*Current address: Microsoft

and, if not accounted for properly, may have very detrimental effects on the solution of the inverse problem which, by its ill-posed nature, is sensitive to all noise. In contrast, a fine discretization yields higher accuracy, and the small associated modeling error can be safely lumped with the measurement noise, at a higher computational cost. These considerations motivate the search for adaptive meshing strategies that refine the mesh dynamically, adding nodes, hence increasing the accuracy, where a higher resolution is needed, while coarsening the mesh in uniform regions, where a too fine discretization would only increases the computational burden and risk fitting the model to noise. For example, if the distributed parameter is a piecewise constant function, ideally a fine mesh is required only near the discontinuities. The problem with this idea is that we usually do not know a priori where the refinement is required, as it is determined by the properties of the unknown function to be estimated. Thus, the discretization of the problem is part of the inverse problem. Addressing this problem is the topic of this article.

Mesh adaptation is a standard procedure in numerical approximation of partial differential equations (PDE), in particular, when using the finite element method [19, 34]. In that context, a qualitative understanding where refinement is needed is often available: the solution of a PDE is expected to change rapidly near coefficient discontinuities corresponding to material interfaces, or near singular points of the boundary of the computational domain, all of which constitute the input of the problem. To guarantee also a quantitative control on the discretization error, rigorous mathematical tools, such as a priori and a posteriori error estimators, can be derived to control a suitable norm or a functional of the discretization error [1]. Here, we take a different approach based on the Bayesian paradigm. In that framework, hierarchical conditionally Gaussian prior models have been previously proposed as a viable way of dealing with inverse problems whose solution is expected to contain singularities such as discontinuities, the dynamically updated prior variance acting as an indicator of the presence of these singularities. In this work, we propose a novel way of combining some classical adaptive mesh strategies in FEM analysis with a dynamical updating of hierarchical models: The prior variance of the gradient, indicating the level of credibility of the gradient estimation in the current meshing, is given the additional role of signaling how the mesh should be adapted. The proposed BMA algorithm, is illustrated with computed examples involving inverse problems in X-ray tomography and a PDE based inverse source problem.

2 Statement of the problem

In this work, we consider inverse problems of estimating a distributed parameter u over a bounded domain $\Omega \subset \mathbb{R}^d$ from a finite number of noisy observations of a quantity depending on u, that is,

$$b = F^*(u) + \varepsilon, \tag{1}$$

where $b \in \mathbb{R}^m$ comprises the data, F^* is the function mapping the parameter u to the data and $\varepsilon \in \mathbb{R}^m$ represents additive exogenous noise independent of u. While the basic idea may be extended to any dimension, we restrict the discussion here to $d = 1$ and, in particular, $d = 2$.

We consider a discretized version of (1) starting from a tessellation $\mathscr{T}_h = \{K_j\}$ of the domain by non-overlapping intervals in $d = 1$, or triangles or quadrilaterals in $d = 2$, covering a polygon Ω_h approximating the domain Ω, so that $\overline{\Omega}_h = \cup_j \overline{K}_j$, where the subindex h refers to the mesh density discussed in detail later. For the time being, the subindex h is included as a reminder of the approximation. Furthermore, we assume that the tessellation is conforming, that is, $\overline{K}_j \cap \overline{K}_i$ is empty or it coincides with a vertex (for $d = 1, 2$) or and edge (for $d = 2$) [14]. This condition excludes meshes with hanging nodes.

Given a tessellation of the domain Ω with vertices $\{x_j\}_{j=1}^{n_v}$, we assume that a distributed parameter u defined over Ω can be approximated in terms of its nodal values, i.e.,

$$u(x) \approx u_h(x) = \sum_{j=1}^{n_v} u_j \psi_j(x), \quad u_j = u(x_j), \tag{2}$$

where $\{\psi_j\}_{j=1}^{n_v}$ is any Lagrange basis, such that $\psi_j(x_k) = \delta_{jk}$ [14].

We approximate the continuous problem (1) by a computationally feasible finite dimensional approximation and write the model

$$b = F(u_h, \mathscr{T}_h) + \varepsilon_h, \tag{3}$$

where $u_h \in \mathbb{R}^d$ is the unknown grid function in (2) with components u_j, and F is a mesh-dependent computed approximation of the continuous forward model F^*. Typically, in PDE based inverse problems, the computation of F requires solving a finite element problem. Observe that since the discretization introduces a modeling error, the additive noise ε_h is representing not only the exogenous noise ε in the continuous model (1), but also the modeling error. The estimation of modeling error due to discretization has been discussed extensively in the literature, see, e.g., [21, 20, 2, 4], and will be briefly revisited below.

In this article, we consider the selection of the discretization mesh to be part of the inverse problem, and we propose a dynamical updating scheme to iteratively estimate the pair $(u_h, \mathscr{T}_h)$ from the data and the a priori information about the unknown.

The method that we advocate will be illustrated in the context of the inverse problems in the next section, and the results of numerical tests will be discussed in detail in Section 5.

2.1 Example 1: X-ray tomography

. Consider the X-ray tomography problem of estimating a density function $u(x) \geq 0$ in a domain $\Omega \subset \mathbb{R}^2$ from the

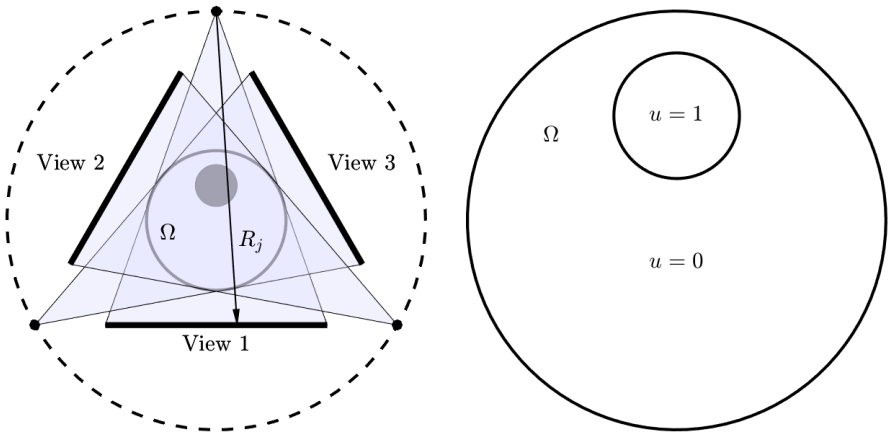

Figure 1: The geometry of the tomography problem: The circular target Ω (right) is illuminated from a given number of directions (here 3) by a fan beam source. Each view consists of a finite number of rays R_j along which the density in Ω is integrated to obtain the attenuation by the target. On the right, the target used in the numerical simulations.

attenuation of X-rays traversing the domain [29]. The setting assumes a fan beam geometry with several views, each view consisting of a finite number of rays traversing the domain Ω, as illustrated schematically in Figure 1. If the

kth ray R_k is parametrized by the arc length t,

$$R_k : \quad x = c_k(t), \quad 0 \le t \le L_k = \text{length of the ray segment},$$

the noiseless forward model consists of integrals of the form

$$b_k = F_k^*(u) = \int_0^{L_k} u(c_k(t))dt, \quad 1 \le k \le m,$$

and if we approximate $u(x)$ according to (2), the forward model becomes

$$b_k \approx \sum_{j=1}^{n_v} \big(\int_0^{L_k} \psi_j(c_k(t))dt \big) u_j = \sum_{j=1}^{n_v} \mathsf{A}(\mathscr{T}_h)_{kj} u_j = F_k(u_h, \mathscr{T}_h),$$

where $\mathsf{A}(\mathscr{T}_h) \in \mathbb{R}^{m \times n_v}$ is a discretization dependent matrix.

2.2 Example 2: Inverse source problem

In the second problem, we consider a linear PDE model over $\Omega = (0,1) \times (0,1)$,

$$\Delta f = -u,$$

with mixed boundary conditions

$$\left.\frac{\partial f}{\partial n}\right|_{\Gamma_1} = \left.\frac{\partial f}{\partial n}\right|_{\Gamma_3} = 0, \quad f\big|_{\Gamma_2} = f\big|_{\Gamma_4} = 0,$$

see Figure 2 for a schematic illustration of the problem setting. The data are assumed to consist of noisy observations of the solution f at discrete observation points $x_k \in \Omega$, i.e., the noiseless observation model is

$$b_k = f(x_k) = F_k(u), \quad 1 \le k \le m.$$

To discretize the problem, the domain Ω is divided in triangular elements, and the source term is approximated in terms of first order (piecewise linear) basis functions $\{\psi_j\}_{j=1}^{n_v}$, while in order to control the approximation error, the solution f is represented in terms of second order (piecewise quadratic) basis functions $\{\varphi_j\}_{j=1}^{N_v}$ corresponding to free second order nodes, excluding the Dirichlet nodes on $\Gamma_2 \cup \Gamma_4$. Observe that the boundary conditions for u are not yet specified. The standard FEM discretization yields the system of equations

$$\sum_{j=1}^{N_v} \underbrace{\left(\int \nabla\varphi_k \cdot \nabla\varphi_j \right)}_{(\mathsf{K}_h)_{kj}} f_j = \sum_{j=1}^{n_v} \underbrace{\left(\int \varphi_k \psi_j \right)}_{(\mathsf{M}_h)_{kj}} u_j, \quad 1 \le k \le N_v,$$

whose solution,

$$f_h = \mathsf{K}_h^{-1} \mathsf{M}_h u_h,$$

is a stable approximation of f at the second order nodal points, while the data are obtained through an interpolation matrix $\mathsf{P}_h \in \mathbb{R}^{m \times N_v}$,

$$b = \mathsf{P}_h \mathsf{K}_h^{-1} \mathsf{M}_h u_h = F(u_h, \mathscr{T}_h).$$

In the following section, we revisit the discretized inverse problem in a framework of Bayesian hierarchical models, formulate certain sparsity promoting hierarchical priors, and outline how that connects naturally with the mesh determination problem.

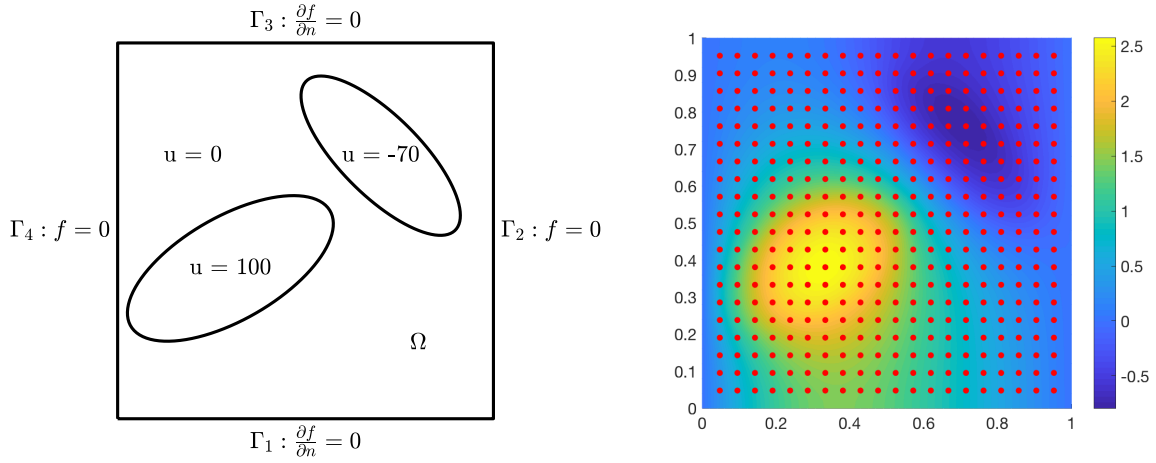

Figure 2: Left: The domain $\Omega = (0,1) \times (0,1)$ with the boundary conditions for f, and the true underlying source u used in the numerical simulations. Right: The computed solution f corresponding to the source u. The observation points are indicated by red dots.

3 Hierarchical Bayesian prior models

We begin by reviewing some results concerning Bayesian hypermodels in inverse problems [11]. We assume that the discrete forward model is linear. An extension to non-linear models is possible but beyond the scope of the present discussion.

Consider the linear discrete inverse problem of estimating $x \in \mathbb{R}^n$ from noisy indirect observations of the form

$$b = \mathsf{A}x + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \Sigma), \tag{4}$$

where $\mathsf{A} \in \mathbb{R}^{m \times n}$, and the noise covariance matrix $\Sigma \in \mathbb{R}^{m \times m}$ is symmetric positive definite. We assume a priori that x is sparse, that is $\|x\|_0 = \mathrm{card}\{\mathrm{supp}(x)\} \ll n$, or more generally, compressible, that is,

$$\mathrm{card}\{x_j \mid |x_j| > \delta\} \ll n,$$

where the threshold δ corresponds to insignificant background values of the components. If Σ is known, the additive noise can be whitened by multiplying both sides of (4) by a matrix $\mathsf{S} \in \mathbb{R}^{m \times m}$ satisfying $\mathsf{S}\Sigma\mathsf{S}^\mathsf{T} = \mathsf{I}_m$, where I_m is the identity matrix of size $m \times m$. The matrix S can be chosen, for instance, to be the Cholesky factor of the precision matrix Σ^{-1}. Hence, without loss of generality, we may assume that $\Sigma = \mathsf{I}_m$, leading to a likelihood model

$$\pi_{b|x}(b \mid x) \propto \exp\left(-\frac{1}{2}\|b - \mathsf{A}x\|^2\right).$$

Define the conditionally Gaussian prior model for x

$$\pi_{x|\theta}(x \mid \theta) \propto \frac{1}{(\theta_1\theta_2\cdots\theta_n)^{1/2}} \exp\left(-\frac{1}{2}\sum_{j=1}^{n}\frac{x_j^2}{\theta_j}\right),$$

where the quantities $\theta_j > 0$ are the a priori variances of the mutually independent components of x. Since we believe that the unknown x is sparse or compressible, in order for the prior to favor such solutions, most of the variances should be close to zero, with only few significantly large. However, since we do not know a priori which components are significantly different from zero, we model θ as a random variable. To promote the desired properties for θ, we

introduce a hyperprior that favors small values for θ_j, while allowing significantly large but rare outliers. Among the distributions with such properties, we concentrate on the three parameter family of the generalized gamma distributions, with probability density function of the form

$$\pi(\theta_j) \sim \left(\frac{\theta_j}{\vartheta_j}\right)^{r\beta-1} \exp\left(-\left(\frac{\theta_j}{\vartheta_j}\right)^r\right).$$

In our setting, the parameter $r \neq 0$ identifies a member of the generalized gamma family, the values of the scale parameters $\vartheta_j > 0$ are chosen separately for each j, while the same shape parameter $\beta > 0$ is used for all θ_j. Setting $r = 1$ gives the gamma distribution, while $r = -1$ is the inverse gamma distribution.

Combining the likelihood, conditional prior and hyperprior according to Bayes' formula we obtain the posterior distribution for the pair (x, θ),

$$\pi_{x,\theta|b}(x, \theta \mid b) \propto \exp\left(-\frac{1}{2}\|b - \mathsf{A}x\|^2 - \frac{1}{2}\sum_{j=1}^n \frac{x_j^2}{\theta_j} + \left(r\beta - \frac{3}{2}\right)\sum_{j=1}^n \log\left(\frac{\theta_j}{\vartheta_j}\right) - \sum_{j=1}^n \left(\frac{\theta_j}{\vartheta_j}\right)^r\right).$$

In this work, we consider the maximum a posteriori (MAP) estimator, defined as the maximizer of the above expression, or, equivalently, minimizer of the energy functional,

$$(x, \theta)_{\mathrm{MAP}} = \operatorname{argmin}\left\{\frac{1}{2}\|b - \mathsf{A}x\|^2 + \frac{1}{2}\sum_{j=1}^n \frac{x_j^2}{\theta_j} - \eta\sum_{j=1}^n \log\left(\frac{\theta_j}{\vartheta_j}\right) + \sum_{j=1}^n \left(\frac{\theta_j}{\vartheta_j}\right)^r\right\}, \tag{5}$$

where $\eta = r\beta - 3/2$. An iterative algorithm to effectively compute the MAP estimate, the Iterative Alternating Sequential algorithm (IAS), has been suggested and extensively studied in the literature, see [6, 9, 12, 7, 8]. The structure of the algorithm can be summarized as follows:

Given the parameters $r \neq 0$, β, $\vartheta \in \mathbb{R}^n$:

Initialize: Set $\theta^{(0)} = \vartheta$, $k = 0$.

Iterate until convergence:

(a) Update x,

$$x^{(k+1)} = \operatorname{argmin}\left\{\frac{1}{2}\|b - \mathsf{A}x\|^2 + \frac{1}{2}\sum_{j=1}^n \frac{x_j^2}{\theta_j^{(k)}}\right\};$$

(b) Update θ,

$$\theta^{(k+1)} = \operatorname{argmin}\left\{\frac{1}{2}\sum_{j=1}^n \frac{(x_j^{(k+1)})^2}{\theta_j} - \eta\sum_{j=1}^n \log\left(\frac{\theta_j}{\vartheta_j}\right) + \sum_{j=1}^n \left(\frac{\theta_j}{\vartheta_j}\right)^r\right\};$$

(c) Increase $k \leftarrow k + 1$;

As pointed out in the cited articles, the algorithm is particularly simple to implement: The update step (a) with fixed θ requires only the solution of a linear least squares problem, while in step (b), each component θ_j can be updated separately. The first order optimality condition yields a uniquely solvable condition for each θ_j, see [6, 9], and in [7], it was shown that, for all values $r \neq 0$, the updating can be done efficiently by solving a simple ordinary differential initial value problem. For particular values of r, including $r = \pm 1$, a closed form updating formula for θ_j can be found.

It was proved in [9] that, for $r = 1$, the IAS algorithm converges to a unique minimizer of the energy functional, and the convergence rate was analyzed in [12]. For other values of r, the convergence properties were further studied in [7]. The articles [12, 7] established the connection between the proposed algorithm and a number of classical sparsity-promoting regularization schemes, including the ℓ^p penalty regularization. Furthermore, in the cited articles, it was demonstrated that, if an estimate of the signal-to-noise ratio (SNR) of the problem is available, under certain natural conditions the selection of the hyperparameter ϑ can be made automatic, and it is directly related to the sensitivity of the inverse problem to the different components of x [9, 12, 7]. We summarize the results relevant for the present work in the following theorems.

Theorem 3.1 *For $r = 1$ and $\eta > 0$, the objective function (5) is strictly convex, and the IAS algorithm converges to the global minimizer $(\widehat{x}, \widehat{\theta})$, where $\widehat{\theta}$ satisfies*

$$\widehat{\theta}_j = \frac{1}{2}\vartheta_j \left(\eta + \sqrt{\eta^2 + \frac{\widehat{x}_j^2}{2\vartheta_j}} \right).$$

Moreover, when $\eta \to 0+$, the minimizer $\widehat{x}$ converges to the minimizer of the functional

$$\frac{1}{2}\|b - \mathsf{A}x\|^2 + \sqrt{2}\sum_{j=1}^{n} \frac{|x_j|}{\sqrt{\vartheta_j}}. \tag{6}$$

From this theorem it follows that, in the limit, ϑ_j represents a weight of the ℓ^1 penalty. In geophysical and biomedical literature, the weight in the penalty is usually introduced to account for the sensitivity of the data to different components in x, see, e.g. [25, 23, 24]. While such weighting is often introduced based on heuristic arguments, and considered to correspond to a prior that cannot be justified in the Bayesian paradigm (the prior would depend on A, hence on the measurement configuration and cannot be viewed as a proper prior), the following theorem establishes a proper Bayesian interpretation.

Theorem 3.2 *(a) Assuming that a support set $S \subset \{1, 2, \ldots, n\}$ is given, the SNR conditional to the unknown x being supported on S, denoted by SNR_S, is given by*

$$\mathrm{SNR}_S = \frac{\sum_{j\in S} \nu(r,\beta)\vartheta_j}{\mathrm{trace}(\Sigma)} + 1, \quad \nu(r,\beta) = \frac{\Gamma(\beta + 1/r)}{\Gamma(\beta)},$$

provided that $\beta > -1/r$.

(b) Let $p_k = \mathsf{P}\{\|x\|_0 = k\}$ denote the probability that the support of the signal has cardinality k, for $k = 1, 2, \ldots, n$. Then the exchangeability condition ($\mathscr{E}$),

$$(\mathscr{E}): \quad \mathrm{SNR}_S = \mathrm{SNR}_{S'} \text{ whenever } S \text{ and } S' \text{ are of the same cardinality,}$$

is satisfied if and only if ϑ_j is chosen as

$$\vartheta_j = \frac{C}{\|a^{(j)}\|^2}, \quad C = \frac{(\overline{\mathrm{SNR}} - 1)\mathrm{trace}(\Sigma)}{\nu(r,\beta)} \sum_{k=1}^{n} \frac{p_k}{k},$$

where $a^{(j)} = \mathsf{A}e_j$ is the jth column of the matrix A.

We point out that since $\|\mathsf{A}e_j\|$ is the classical sensitivity of the data to the jth component of x, by choosing the weights ϑ_j as in the previous theorem, the components x_j in the penalty are weighted as suggested in the literature, and the weighting is justified by a Bayesian argument rather than heuristics. The explanation for how the matrix A can enter in the prior is that the concept of SNR implicitly carries information about the measuring configuration.

We address now how to solve the least squares problem (a) in the IAS algorithm. The solution of the linear least squares problem can be restated as

$$x = \operatorname{argmin}\left\{\left\|\begin{bmatrix}\mathsf{A}\\ \mathsf{D}_\theta^{-1/2}\end{bmatrix}x - \begin{bmatrix}b\\0\end{bmatrix}\right\|\right\} = \operatorname{argmin}\left\{\left\|\begin{bmatrix}\mathsf{A}\mathsf{D}_\theta^{1/2}\\ \mathsf{I}_n\end{bmatrix}(\mathsf{D}_\theta^{-1/2}x) - \begin{bmatrix}b\\0\end{bmatrix}\right\|\right\},$$

where

$$\mathsf{D}_\theta = \operatorname{diag}(\theta) \in \mathbb{R}^{n\times n},$$

yielding a formulation that is formally similar to the standard Tikhonov regularized problem for the system

$$\mathsf{A}\mathsf{D}_\theta^{1/2}w = b, \quad D_\theta^{1/2}x = w,$$

with regularization parameter equal to one. Inspired by this observation, it has been suggested to replace the exact solution of the quadratic exact minimization problem (a) by an approximate one based on the regularization properties of the Krylov subspace iterations [10, 7]. In the approximate IAS algorithm, the update $x^{(k+1)}$ is the *reduced Krylov subspace* (RKS) solution, defined as follows. Introduce the notation $\mathsf{A}_\theta = \mathsf{A}\mathsf{D}_\theta^{1/2}$ and consider the sequence of approximate solutions

$$w_k = \operatorname{argmin}\big\{\|b - \mathsf{A}_\theta w\| \mid w \in \mathscr{K}_k(\mathsf{A}_\theta^{\mathsf{T}}b, \mathsf{A}_\theta^{\mathsf{T}}\mathsf{A}_\theta)\big\}, \quad x_k = \mathsf{D}_\theta^{1/2}w_k, \tag{7}$$

where

$$\mathscr{K}_k(\mathsf{A}_\theta^{\mathsf{T}}b, \mathsf{A}_\theta^{\mathsf{T}}\mathsf{A}_\theta) = \operatorname{span}\big\{(\mathsf{A}_\theta^{\mathsf{T}}\mathsf{A}_\theta)^\ell \mathsf{A}_\theta^{\mathsf{T}}b \mid 0 \le \ell \le k-1\big\},$$

is the kth Krylov subspace associated with the vector $\mathsf{A}_\theta^{\mathsf{T}}b$ and the matrix $\mathsf{A}_\theta^{\mathsf{T}}\mathsf{A}_\theta$. The quantity $b - \mathsf{A}_\theta w_k$ whose norm is minimized is the discrepancy vector corresponding to w_k. The RKS solution is the $k = k_0$th iterate, with k_0 chosen to be the first index k satisfying the criterion

$$(\mathscr{C}): \ \|b - \mathsf{A}_\theta w_{k+1}\| \le \sqrt{m}, \quad \text{or} \quad G(w_{k+1}) > \tau G(w_k),$$

where $\tau - 1 = \epsilon > 0$ is a small safeguard parameter, and G is the norm of the discrepancy of the original linear system,

$$G(w) = \left\|\begin{bmatrix}\mathsf{A}\mathsf{D}_\theta^{1/2}\\ \mathsf{I}_n\end{bmatrix}w - \begin{bmatrix}b\\0\end{bmatrix}\right\|^2 = \|b - \mathsf{A}\mathsf{D}_\theta^{1/2}w\|^2 + \|w\|^2. \tag{8}$$

Extensive computed examples reported in the cited articles have shown the approximate IAS solutions to be consistently similar to those obtained with the exact IAS, at a much lower computational cost.

3.1 Sparse increments

In this article we focus on the inverse problems of estimating piecewise constant distributed parameters with sparse gradients. In order to take advantage of the sparsity promotion of the IAS algorithm, the minimization problem must be expressed in terms of a discrete approximation of the function increments rather than the function itself. For one-dimensional problems, the modification is quite straightforward. Denoting by $u(t)$ the unknown distributed

parameter, $0 \leq t \leq 1$, with discretization $x_j = u(jh)$, $h = 1/n$ being the discretization parameter, we may write x in terms of the increments, i.e.,

$$x_j = x_0 + \sum_{k=1}^{j} (x_k - x_{k-1}), 1 \leq j \leq n.$$

Assuming for simplicity that $x_0 = 0$, and defining $z_j = x_j - x_{j-1}$, we have

$$\mathsf{L}x = z, \quad \mathsf{L} = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}.$$

Since the matrix L is invertible, we may reformulate the problem as estimating z from the observation model

$$b = \mathsf{A}\mathsf{L}^{-1}z + \varepsilon,$$

with the a priori belief that the solution is sparse. Accordingly, we modify the IAS update of the unknown in terms of the increment vector, namely,

$$z^{(k+1)} = \operatorname{argmin} \left\{ \frac{1}{2}\|b - \mathsf{A}\mathsf{L}^{-1}z\|^2 + \frac{1}{2}\sum_{j=1}^{n} \frac{z_j^2}{\theta_j^{(k)}} \right\}, \quad x^{(k+1)} = \mathsf{L}^{-1}z^{(k+1)},$$

followed by the update step of θ,

$$\theta^{(k+1)} = \operatorname{argmin} \left\{ \frac{1}{2}\sum_{j=1}^{n} \frac{(z_j^{(k+1)})^2}{\theta_j} - \eta \sum_{j=1}^{n} \log\left(\frac{\theta_j}{\vartheta_j}\right) + \sum_{j=1}^{n} \left(\frac{\theta_j}{\vartheta_j}\right)^r \right\}.$$

A different approach is required in dimension $d \geq 2$ because the finite difference matrix is no longer square, and there is no one-to-one correspondence between x and z.

Consider a triangular mesh as shown in Figure 3, and assume that the values of the unknown function are known to vanish at some of the nodes, referred to as bound nodes (or Dirichlet nodes), whereas we refer to the remaining ones as free nodes. In order to keep the notation simple, let n_v denote from now on the number of free nodes, and n_e the number of free edges $e_\ell = \{v_{\ell_1}, v_{\ell_2}\} \in E_{\text{free}}$, for which at least one endpoint is a free node. Let $\mathsf{L} \in \mathbb{R}^{n_e \times n_v}$ denote the matrix mapping the nodal values x at free nodes to increments z along free edges,

$$z = \mathsf{L}x. \tag{9}$$

Note that the increments along edges between bound nodes vanish. Since the nodal values at bound nodes vanish, it is straightforward to show that L has the null space $\mathscr{N}(\mathsf{L}) = \{0\}$. Let n_t denote the number of all triangular elements in the mesh, and let $c \in \mathbb{R}^{n_t}$ be a vector containing the sums of the increments in a clockwise order over the edges of each element. We denote by $\mathsf{M} \in \mathbb{R}^{n_e \times n_t}$ the matrix of entries 0 or ± 1, such that

$$c = \mathsf{M}z. \tag{10}$$

If z corresponds to a grid function x through (9), the circulations must vanish, that is, $z \in \mathscr{N}(\mathsf{M})$. Hence, the matrices L and M define a short exact chain,

$$\{0\} \longrightarrow \mathbb{R}^{n_v} \xrightarrow{\mathsf{L}} \mathbb{R}^{n_e} \xrightarrow{\mathsf{M}} \mathbb{R}^{n_t} \longrightarrow \{0\}.$$

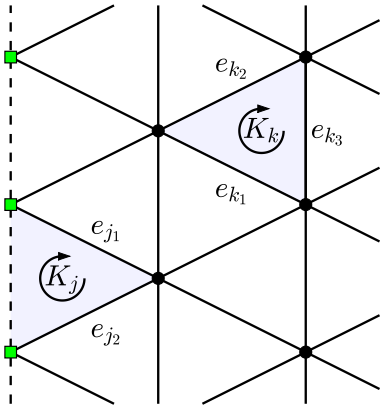

Figure 3: Schematic illustration of the circulation condition $\mathsf{M}z = 0$. In this figure, the black dots indicate the free nodes and the green squares are bound nodes, or Dirichlet nodes, where the grid function is assumed to vanish. Free edges, drawn as solid lines, have at least one free node as an end point, while bound edges, drawn with dashed lines, connect bound nodes. Bound edges give a zero contribution to the circulation. If z is the vector of increments corresponding to a grid function x, the sums of increments around edges of each element with a given orientation (here clockwise) must vanish.

To define a prior for the increments analogously to the one-dimensional case, we must secure that the prior is concentrated on the subspace $\mathscr{N}(\mathsf{M})$. To do this let

$$\mathsf{L}_\theta = \mathsf{D}_\theta^{-1/2}\mathsf{L}, \quad \text{where} \quad \mathsf{D}_\theta = \operatorname{diag}(\theta_1, \ldots, \theta_{n_e}). \tag{11}$$

Given a grid function x, we introduce the auxiliary variable

$$\beta = \mathsf{L}_\theta x. \tag{12}$$

Observe that

$$\sum_{j=1}^{n_e} \frac{(\mathsf{L}x)_j^2}{\theta_j} = \|\beta\|^2.$$

In order that β corresponds to a grid function, we must require that $\beta \in \mathscr{H} = \mathscr{R}(\mathsf{L}_\theta)$. If this is the case, x can be solved through the pseudoinverse, i.e., $x = \mathsf{L}_\theta^\dagger \beta$. We therefore define the conditional density for β, setting

$$\pi_{\beta|\theta,b}(\beta \mid \theta, b) \propto \exp\left(-\frac{1}{2}\|b - \mathsf{A}\mathsf{L}_\theta^\dagger \beta\|^2 - \frac{1}{2}\|\beta\|^2\right) \delta_{\mathscr{H}}(\beta), \tag{13}$$

where $\delta_{\mathscr{H}}(\beta)$ is the singular measure concentrated on the subspace $\mathscr{H}$. It was shown in [8] that for a fixed θ, the minimizer of the energy functional,

$$\beta = \operatorname{argmin}\left\{\|b - \mathsf{A}\mathsf{L}_\theta^\dagger \beta\|^2 + \|\beta\|^2\right\},$$

is automatically in the subspace $\mathscr{H}$, i.e., the compatibility condition needs not to be enforced. This observation leads to the following version of the IAS algorithm for the increments:

Given the parameters $r \neq 0$, β, $\vartheta \in \mathbb{R}^n$:

Initialize: Set $\theta^{(0)} = \vartheta$, $k = 0$.

Iterate until convergence:

(a) Update x,

$$x^{(k+1)} = \mathsf{L}_{\theta^{(k)}}^{\dagger}\beta^{(k+1)}, \text{ where } \beta^{(k+1)} = \operatorname{argmin}\left\{\|b - \mathsf{A}\mathsf{L}_{\theta^{(k)}}^{\dagger}\beta\|^2 + \|\beta\|^2\right\}.$$

(b) Update θ,

$$\theta^{(k+1)} = \operatorname{argmin}\left\{\frac{1}{2}\sum_{j=1}^{n_e}\frac{(\mathsf{L}x^{(k+1)})_j^2}{\theta_j} - \eta\sum_{j=1}^{n_e}\log\left(\frac{\theta_j}{\vartheta_j}\right) + \sum_{j=1}^{n_e}\left(\frac{\theta_j}{\vartheta_j}\right)^r\right\};$$

(c) Increase $k \leftarrow k+1$;

In order to update x, we need to solve in the least squares sense the linear system

$$\begin{bmatrix} \mathsf{A}\mathsf{L}_{\theta}^{\dagger} \\ \mathsf{I} \end{bmatrix}\beta = \begin{bmatrix} b \\ 0 \end{bmatrix}. \tag{14}$$

Alternatives, as proposed in [7], we can consider the reduced problem

$$\mathsf{A}\mathsf{L}_{\theta}^{\dagger}\beta = b,$$

and compute its approximate solution by the conjugate gradient method for least squares (CGLS) with an appropriate stopping criterion. More precisely, if

$$\mathscr{K}_k = \mathscr{K}_k\left((\mathsf{A}\mathsf{L}_{\theta}^{\dagger})^{\mathsf{T}}b, (\mathsf{A}\mathsf{L}_{\theta}^{\dagger})^{\mathsf{T}}\mathsf{A}\mathsf{L}_{\theta}^{\dagger}\right) = \operatorname{span}\left\{\left[(\mathsf{A}\mathsf{L}_{\theta}^{\dagger})^{\mathsf{T}}\mathsf{A}\mathsf{L}_{\theta}^{\dagger}\right]^j(\mathsf{A}\mathsf{L}_{\theta}^{\dagger})^{\mathsf{T}}b \mid j = 0, 1, \ldots, k-1\right\},$$

defines the Krylov subspace of dimension k associated with the vector $(\mathsf{A}\mathsf{L}_{\theta}^{\dagger})^{\mathsf{T}}b$ and the matrix $(\mathsf{A}\mathsf{L}_{\theta}^{\dagger})^{\mathsf{T}}\mathsf{A}\mathsf{L}_{\theta}^{\dagger}$, we approximate the solution of (14) by the kth iterate of the CGLS algorithm,

$$\beta_k = \operatorname{argmin}\left\{\|b - \mathsf{A}\mathsf{L}_{\theta}^{\dagger}\beta\|,\ \beta \in \mathscr{K}_k\right\},$$

where k is the smallest index that satisfies the discrepancy criterion

$$(\mathscr{C}):\ \|b - \mathsf{A}\mathsf{L}_{\theta}^{\dagger}\beta_{k+1}\| \le \sqrt{m}, \quad \text{or} \quad G(\beta_{k+1}) > \tau G(\beta_k), \tag{15}$$

with

$$G(\beta) = \|b - \mathsf{A}\mathsf{L}_{\theta}^{\dagger}\beta\|^2 + \|\beta\|^2.$$

It was shown in [7] that it is possible to add a positivity constraint to the IAS algorithm by a proper projection on the positive cone.

In the CGLS algorithm, the multiplication of a vector z with $\mathsf{A}\mathsf{L}_{\theta}^{\dagger}$ can be evaluated by first solving the linear system

$$\mathsf{L}_{\theta}\alpha = z$$

in the least squares sense, then multiplying the solution α by A. To implement the multiplication of b by $(\mathsf{A}\mathsf{L}_{\theta}^{\dagger})^{\mathsf{T}} = (\mathsf{L}_{\theta}^{\dagger})^{\mathsf{T}}\mathsf{A}^{\mathsf{T}}$, we recall the definition of the pseudo-inverse and its transpose as

$$\mathsf{L}_{\theta}^{\dagger} = \left(\mathsf{L}_{\theta}^{\mathsf{T}}\mathsf{L}_{\theta}\right)^{-1}\mathsf{L}_{\theta}^{\mathsf{T}}, \quad \left(\mathsf{L}_{\theta}^{\dagger}\right)^{\mathsf{T}} = \mathsf{L}_{\theta}\left(\mathsf{L}_{\theta}^{\mathsf{T}}\mathsf{L}_{\theta}\right)^{-1},$$

where the matrix $\mathsf{L}_{\theta}^{\mathsf{T}}\mathsf{L}_{\theta}$ in our application is very sparse. Therefore, for the multiplication with the transpose, we solve first

$$\left(\mathsf{L}_{\theta}^{\mathsf{T}}\mathsf{L}_{\theta}\right)w = \mathsf{A}^{\mathsf{T}}b,$$

for w, then multiply w by L_{θ}.

3.2 Modeling error

As pointed out earlier, the discretization of a continuous problem introduces a discretization error that cannot be ignored when solving inverse problems unless the discretization is so fine that the modeling error is dominated by the exogenous error in the data. The ways in which approximation errors are addressed in classical FEM applications and inverse problems have some fundamental differences. In the former, approximation error is the discrepancy between the true solution of a PDE and the computed approximation, and sophisticated methods to control the accuracy through mesh adaptation have been developed. In inverse problems, the intrinsic ill-posedness adds an extra layer of complexity, as small inaccuracies in the forward model, if not properly addressed, may propagate as huge errors in the computed solution. We briefly review here how the modeling error can be addressed in the Bayesian framework. Given the exact model (1) and the corresponding discrete approximation (3), we write

$$\begin{aligned} b &= F^*(u) + \varepsilon \\ &= F(u_h, \mathscr{T}_h) + \big[F^*(u) - F(u_h, \mathscr{T}_h)\big] + \varepsilon, \end{aligned}$$

expressing the modeling error due to the discretization as

$$m_h = m(u, \mathscr{T}_h) = F^*(u) - F(u_h, \mathscr{T}_h) = \mathscr{F}_h(u).$$

The problem with estimating the modeling error is that it depends on the solution of the inverse problem, which is unknown. In the Bayesian framework, the distribution of the modeling error can be expressed as the push-forward of the prior distribution,

$$\pi_{m_h} = \mathscr{F}_h^* \pi_u.$$

This idea, originally proposed in [21, 20], was further developed and applied to a variety of inverse problems (see, e.g., [2, 30, 31]), and generalized in [5, 4] through an iterative updating process taking into account the increasing information on u as the inverse problem is solved. In this paper, to contain the computational burden, we adopt a simplified approach and approximate only the order of magnitude of the modeling error in a straightforward manner that will be described when dealing with the computed examples.

4 Mesh adaptation strategy

Mesh adaptation is a well-established tool to improve the accuracy and the efficiency of a discretization scheme. The goal pursued by a mesh adaptation procedure is to minimize the number of the mesh elements for a fixed accuracy on the discrete solution, or, vice versa, to maximize such accuracy for a given number of tiles in the mesh.

The alternative to an adapted mesh is a uniform tessellation of the domain Ω_h, with a sufficiently small discretization size so that the complex features of the solution can be correctly identified. This last choice is clearly not optimal from a computational viewpoint, in particular in the presence of problems which are highly heterogeneous in space and/or in time.

The actual challenge when generating an adapted mesh is represented by the capability to identify the areas of the domain which deserve to be refined or coarsened. This goal can be pursued by resorting to a preliminary knowledge of the phenomenon at hand (if any), to a heuristic criterion (e.g., by tracking the gradient or the Hessian of the discrete solution), or to theoretically sound mathematical tools which are known in the literature as *a priori* or *a posteriori* error estimators [1]. The former are expressed in terms of the exact solution (e.g., via an intepolation error control), the latter involve the discrete solution, thus providing a practical (i.e., directly computable) information. Among the several a posteriori error estimators available in the literature, we focus on the recovery-based error

estimators proposed for the first time by O.C. Zienkiewicz and J.Z. Zhu in 1987 [35]. This methodology inspired us to propose a new mesh adaptation for inverse problems.

According to a recovery-based approach [35, 37], the mesh adaptation is driven by a control of the H^1-seminorm of the discretization error,

$$|e_h|^2_{H^1(\Omega)} = \int_\Omega |\nabla u(x) - \nabla u_h(x)|^2 \, dx = \sum_{K_j \in \mathscr{T}_h} \int_{K_j} |\nabla u(x) - \nabla u_h(x)|^2 \, dx, \tag{16}$$

where, for simplicity, we have assumed $\Omega \equiv \Omega_h$. Since the gradient of the exact solution is not known, we need to replace ∇u in (16) by a computable quantity, $\nabla^* u$, known as *recovered gradient* [36]. This yields an *a posteriori* error control for e_h, given by

$$|e_h|^2_{H^1(\Omega)} \approx \sum_{K_j \in \mathscr{T}_h} \int_{K_j} |\nabla^* u(x) - \nabla u_h(x)|^2 \, dx = \sum_{K_j \in \mathscr{T}_h} \eta^2_{K_j}. \tag{17}$$

Several procedures to recover the gradient have been proposed in the literature, each yielding a different estimator in (17). In the pioneering work [35], the authors proposed an estimator based on either a least squares process or an averaging step. Independently of the adopted procedure, at first one computes the values $\nabla^* u(x_i)$ at certain points, $\{x_i\}_{i=1}^{m_r}$, known as *recovery points*, starting from the values assumed by ∇u_h on a set of *sampling points*, $\{s_j\}_{j=1}^{n_s}$, located in a patch associated with x_i. Subsequently, $\nabla^* u$ is defined over elements by a piecewise polynomial interpolation of the recovered values $\nabla^* u(x_i)$. If the recovery points x_i are chosen to be vertices of the mesh and the sampling points s_j are the triangle barycenters, the averaging method seeks a recovered gradient such that

$$\nabla^* u(x_i) = \frac{1}{W^i} \sum_{j=1}^{n_s} w^i_j \nabla u_h(s_j) \quad \text{with} \quad i = 1, \ldots, m_r, \tag{18}$$

where w^i_j are suitable weights and $W^i = \sum_{j=1}^{n_s} w^i_j$. The values of the weights w^i_j can be set equal to the area, $|K_j|$, of the elements K_j in $\mathscr{T}_h$ containing x_i as a vertex, with j running over the triangles constituting the patch, $\Delta_{x_i} = \{K \in \mathscr{T}_h \, : \, x_i \in \overline{K}\}$, associated with the vertex x_i, so that $W^i = |\Delta_{x_i}|$, the area of the patch; for common choices in practical applications, see, e.g., [36, 22, 32, 3, 28]. We remark that some attention has to be paid in the definition of the recovered gradient when recovery points belong to the domain boundary (for details, see, e.g., [36]).

Once the recovered gradient has been computed, we can obtain the contribution of each element to the total a posteriori error estimate. According to the *equidistribution principle*, an ideal mesh is such that the error contribution is equally distributed among the elements. To attain this goal, define the scaled local error estimator as

$$\widehat{\eta}^2_{K_j} = \frac{1}{|K_j|} \eta^2_{K_j},$$

thus defined to even out the effect of different sizes of the elements. If N_h is the cardinality of $\mathscr{T}_h$, the equidistribution principle requires that

$$\eta^2_{K_j} = |K_j| \widehat{\eta}^2_{K_j} = \frac{\tau^2}{N_h},$$

where $\tau > 0$ is the accuracy required for the discrete solution. Assume for simplicity that the mesh is isotropic, i.e., all triangles are equilateral, and each element is a shifted, rotated and scaled version of the reference element $\widehat{K}$, an equilateral triangle with unit diameter, then $|K_j| = h_j^2 |\widehat{K}|$, where the diameter h_j of the element K_j defines the mesh density parameter $h = \{h_j = \mathrm{diam}(K_j) \mid K_j \in \mathscr{T}_h\}$. Then the equidistribution principle can be translated into a condition for the diameter, i.e.,

$$h_j^2 = \frac{\tau^2}{N_h |\widehat{K}| \widehat{\eta}^2_{K_j}}. \tag{19}$$

These target values can be used to drive the mesh refinement and coarsening. In particular, as detailed in the next section, the new adapted mesh can be described in terms of a metric defined in Ω, providing us a useful starting point for inverse problems.

4.1 Meshing and metric

Let $\widehat{K}$ denote the reference triangle defined in the previous section, and K be a generic triangle in the tessellation. We define an affine mapping

$$F_K : \widehat{K} \to K, \quad x \mapsto s_K + \mathsf{F}_K x,$$

where $\mathsf{F}_K \in \mathbb{R}^{2\times 2}$ is the Jacobian of the map and $s_K \in \mathbb{R}^2$ is the shift parameter. Let

$$\mathsf{F}_K = \mathsf{P}_K \mathsf{U}_K,$$

be the polar decomposition of F_K, where U_K is an orthogonal matrix and P_K is a symmetric positive definite matrix whose eigenvalues, $0 < \lambda_{K,1} \leq \lambda_{K,2}$ are the singular values of F_K.

We associate a piecewise constant metric in Ω_h with a tessellation $\{K_j\}$ in terms of a piecewise constant metric tensor $\mathsf{G} = [g_{ij}]$, such that

$$\mathsf{G}\big|_{K_j} = \mathsf{P}_{K_j}^{-2}. \tag{20}$$

To compute the diameter of the elements with respect to this metric, consider a line segment across the reference triangle, parametrized with respect to the arc length, $\xi = \xi(t) \in \widehat{K}$, and the corresponding parametrized curve

$$\gamma(t) = s_{K_j} + \mathsf{F}_K \xi(t).$$

In the metric G defined by (20), the arc length element is given by

$$ds = \big(\gamma'(t)^{\mathsf{T}} \mathsf{G} \gamma'(t)\big)^{1/2} dt = \big(\xi'(t)^{\mathsf{T}} \mathsf{U}_{K_j}^{\mathsf{T}} \mathsf{P}_{K_j} \mathsf{G} \mathsf{P}_{K_j} \mathsf{U}_{K_j} \xi'(t)\big)^{1/2} dt = \|\xi'(t)\| dt,$$

because of the orthogonality of U_{K_j}, hence in this metric the elements K_j have the same diameter.

Conversely, given a metric tensor G on Ω, we can look for a tessellation so that the elements have approximately the same size with respect to this metric. The meshing problem can be formally written as follows.

Problem 4.1 *Given a metric G in Ω and $0 < h_{\min} < h_{\max}$, find a tessellation $\{K_j\}$ of Ω such that $h_{\min} \leq \operatorname{diam}(K_j) \leq h_{\max}$ by minimizing the discrepancy*

$$\max_j \|\mathsf{G}\big|_{K_j} - \mathsf{P}_{K_j}^{-2}\|_\infty$$

over the tessellation.

Observe that if for an element K, $\mathsf{P}_K = \lambda \mathsf{I}_2$, the element K is, up to a translation and an orthogonal transformation, a scaled copy of $\widehat{K}$, and its diameter is equal to λ. Consequently, if $h = h(x)$ is a given function defining the target distance between the vertices in the mesh and $x \in \Omega$, to find an isotropic non-homogenous mesh corresponding to $h(x)$,we define the metric as

$$\mathsf{G}(x) = \frac{1}{h^2(x)} \mathsf{I}.$$

In the next section, we propose an automatic mesh adaptation algorithm for inverse problems, where the current approximation is used, in a predictive way, to generate the function $h(x)$, hence the metric yielding the new discretization.

4.2 Bayesian Mesh Adaptation (BMA)

In this subsection we customize the mesh adaptation procedure to inverse problems, using the hierarchical IAS algorithm as a tool to convey the information for the mesh generation. The driving idea of this approach is the interpretation of large estimated variance of the solution along an edge as an indication of a potential jump in the solution that may require a finer discretization, while a small variance allows larger elements, thus reducing the complexity of the problem.

Consider the discretized inverse problem

$$b = \mathsf{A}_h u_h + \varepsilon_h, \tag{21}$$

where A_h is the discretized forward matrix based on the current tessellation $\mathscr{T}_h$, identified with the triplet $(\mathsf{V}_h, \mathsf{E}_h, \mathsf{T}_h)$ of matrices, where $\mathsf{V}_h \in \mathbb{R}^{2\times\overline{n}_v}$ contains all $\overline{n}_v$ vertex coordinates, including the Dirichlet nodes, $\mathsf{E}_h \in \mathbb{N}^{\overline{n}_e\times 2}$ is the edge list of all $\overline{n}_e$ edges with pointers to the endpoint vertices, and $\mathsf{T}_h \in \mathbb{N}^{n_t\times 3}$ is the element topology matrix with pointers to the vertex nodes. Here u_h is the vector of the nodal values of the unknown function at the free nodes of the mesh, and ε_h is the noise vector with zero mean and covariance Σ_h estimated from the approximation error model. Let L_h denote the full-rank matrix mapping the nodal values to the increments along the free edges.

Given the current model (21), we run the IAS algorithm for sparse increments, compute the current approximation, $(x_h, \theta_h) \in \mathbb{R}^{n_v} \times \mathbb{R}^{n_e}$ for the free nodes and edges, and extend θ_h by zero to the boundary edges, i.e., the edges between the Dirichlet nodes. Observe that if $(\theta_h)_j$ is small, we expect the increment of the solution x_h along the jth edge to be small. In other words, a small $(\theta_h)_j$ is taken as an indication of unlikely discontinuity of the unknown along that edge, thus suggesting that there is no need to refine the mesh, or even that a coarser mesh would suffice. Conversely, a large value of $(\theta_h)_j$ indicates the potential for a large increment, and to better localize the point of discontinuity along the edge, a refinement is desirable. This heuristics suggests that we can use θ_h as a proxy for the metric, after interpolating the edge-bound values of θ_h to Ω, giving rise to the following remeshing algorithm.

Algorithm 1: Variance-based meshing (VBM)

Input: Triplet $(\mathsf{V}_h, \mathsf{E}_h, \mathsf{T}_h)$, a variance vector θ_h with cardinality equal to the number of edges.

(a) Add an auxiliary vertex at the midpoint of each edge,

$$\widehat{v}_\ell = \frac{1}{2}\left[\mathsf{V}_h\left(:, \mathsf{E}(\ell, 1)\right) + \mathsf{V}_h\left(:, \mathsf{E}(\ell, 2)\right)\right], \quad 1 \le \ell \le \overline{n}_e.$$

Denote by $\widehat{\mathsf{V}}_h \in \mathbb{R}^{2\times n_e}$ the matrix of all auxiliary nodes.

(b) Generate an auxiliary triplet $\widehat{\mathscr{T}_h} = (\widehat{\mathsf{V}}_h, \widehat{\mathsf{E}}_h, \widehat{\mathsf{T}}_h)$ by using Delaunay triangularization [17].

(c) Define the metric

$$\mathsf{G}(x) = \max\{\widehat{\theta}_h(x), \theta_*\}\mathsf{I}_2,$$

where $\widehat{\theta}_h(x)$ is the piecewise linear extension of the nodal values θ_h in the auxiliary mesh, and $\theta_* > 0$ is a threshold value identifying when the value of the parameter θ_h is non-negligeable.

(d) Given the metric, find a new triplet $(\mathsf{V}_{h'}, \mathsf{E}_{h'}, \mathsf{T}_{h'})$.

We point out that in the above algorithm, slight adjustments of the nodes $\widehat{\mathsf{V}}_h$ at the boundary of the domain may be necessary to ensure that the resulting polygonal domain properly approximates the domain Ω. In step (d), we

generate the new mesh using the meshing tool included in the `FreeFem++` package [18]. The algorithm allows a nodal input of the metric in the current mesh, which in this case is the midpoint mesh constructed above, and requires the specification of a range $[h_{\min}, h_{\max}]$ for the mesh size, providing a way to control the complexity and the accuracy of the forward map.

We may now proceed iteratively by repeatedly recalculating the forward model, estimating the modeling error, and applying the IAS algorithm to the new discretized problem. We summarize how to organize the calculations in the algorithm below.

Algorithm 2: Bayesian Mesh Adaptation with IAS (BMA-IAS)

Input: The data $b \in \mathbb{R}^m$, the noise covariance matrix $\Sigma \in \mathbb{R}^{m \times m}$, the initial mesh parameter h.

Initialize: Generate the initial triplet $(\mathsf{V}_h, \mathsf{E}_h, \mathsf{T}_h)$.

Iterate until stopping criterion is satisfied:

(a) Compute the forward matrices A_h and L_h on the current mesh.

(b) Estimate the modeling error, compute the covariance matrix Σ_h. Whiten the noise ε_h.

(c) Compute the sensitivity vector corresponding to $\mathsf{AL}^\dagger$; compute ϑ.

(d) Initialize the IAS iteration, setting $\theta = \vartheta$.

(e) **Run the IAS** algorithm with sparse increments by repeating until convergence:

(i) Update $\beta_h \to u_h$ by CGLS with early stopping criterion;

(ii) Update θ_h.

(d) **Update the mesh** using Algorithm 1.

end

The IAS iteration can be augmented with the projection of the solution u_h to the positive cone if the parameter u is a priori known to be positive, see [7] for details. Above, the stopping criterion for the IAS algorithm is given by the condition (15). The stopping criterion for the mesh update is discussed in the light of computed examples, where preliminary versions of a systematic criterion are considered.

5 Computed examples

In this section we demonstrate the viability of Algorithm 2 by applying it to the linear inverse problems described in Sections 2.1 and 2.2.

5.1 Tomography problem

In this example, we consider the X-ray tomography problem sketched in Figure 1. In the first test, we assume that the target is illuminated from several directions, choosing 15 equally spaced view directions, each view consisting of 300 rays, thus the total dimensionality of the data is $m = 15 \times 300 = 4\,500$. The generative model used to compute the noiseless data, shown in Figure 1, consists of a unit disc Ω of zero background density with a circular inclusion with center at $(0, 0.5)$, radius 0.3, and density $u = 1$. We compute the noiseless data by solving analytically for

each ray the length of the line segment inside the inclusion, thus avoiding any error due to discretization. In this experiment, the exogenous additive noise is white noise scaled by σ_{ex} equal to 0.5% of the maximum of the entries of the noiseless data.

To find a computationally operational model for the discretization error, let b_0 denote the exact noiseless data, and assume that the observation model is of the form

$$b = \mathsf{A}u + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma_{\text{ex}}^2 \mathsf{I}_m),$$

where A denotes an integral operator. Given a discretization level $h > 0$, we write the scaled white noise likelihood model

$$\pi(b \mid u_h) \sim \mathcal{N}(0, \sigma_h^2 \mathsf{I}_m).$$

Choosing $\sigma_h^2 = \sigma_{\text{ex}}^2$, thus ignoring the modeling error, may not be a good choice if the discretization is coarse and the exogenous noise level is low, as the IAS algorithm with the discretized model may not be able to approximate the data within the noise level σ_{ex}^2. Therefore, we increase artificially the noise level σ_h so that the modeling error, although possibly not zero mean Gaussian, is within the noise envelope. To find a suitable value for σ_h, we begin by writing

$$b = \mathsf{A}u + \varepsilon = \mathsf{A}_h u_h + (\mathsf{A}u - \mathsf{A}_h u_h) + \varepsilon = \mathsf{A}_h u_h + \varepsilon_h + \varepsilon,$$

where $\varepsilon_h = \mathsf{A}u - \mathsf{A}_h u_h$ is the modeling error due to discretization.

To estimate the level of the modeling error, we start by generating a sample of realizations u, assuming circular inclusions with random radius and at random positions, denoted by $\mathscr{S} = \{u^{(1)}, \ldots, u^{(K)}\}$. Since the amplitude of the modeling error scales linearly, we assume for simplicity that the amplitude of the circular inclusion is one. For a given mesh size $h > 0$, we discretize the domain Ω generating a uniform mesh, and approximating the densities $u^{(k)}$ by a piecewise linear function with nodal values given by the true density $u^{(k)}$. We compute the corresponding modeling error realizations

$$\varepsilon_h^{(k)} = \mathsf{A}u^{(k)} - \mathsf{A}_h u_h^{(k)}, \quad 1 \le k \le K.$$

We approximate the underlying noise distribution by scaled Gaussian white noise with variance $\widetilde{\sigma}_h^2$, obtained by

$$\widetilde{\sigma}_h^2 = \frac{1}{m}\text{trace}\left(\frac{1}{K}\sum_{k=1}^{K} \varepsilon_h^{(k)} \big(\varepsilon_h^{(k)}\big)^{\mathsf{T}}\right) = \frac{1}{mK}\sum_{k=1}^{K} \|\varepsilon_h^{(k)}\|^2.$$

Observe that the above equation is exact in the limit $K \to \infty$ if the underlying distribution is indeed a scaled white noise. To account for the modeling error within a scaled white noise framework, we set

$$\sigma_h^2 = \widetilde{\sigma}_h^2 + \sigma_{\text{ex}}^2.$$

Figure 4 shows the value of $\widetilde{\sigma}_h$, computed by drawing 200 disc inclusions for the 15 view tomography data, for different uniform mesh sizes h. Since the dependency of $\widetilde{\sigma}_h$ on h is almost linear, we can fit a linear model to the data and use $\sigma_h = 0.3\,h$ for the estimate of the noise model.

Since the modeling error scales linearly with respect to the inclusion, we need to have some a priori information about the density. To test how critical it is to know the amplitude, we run the IAS algorithm with fixed mesh, for different choices of the approximation error noise variance. In this preliminary test, we use a fine, fixed mesh of size $h = 0.02$, with $17\,813$ elements and $9\,062$ nodes. In this test, the exogenous noise level is set at zero, that is, we assume that the only error source is the discretization error in the forward model. To test how sensitive the solution is to the estimation of the noise level, we consider four different cases: $\sigma_h = 0.03\,h$ (underestimated), $\sigma_h = 0.15\,h$ (slightly underestimated), $\sigma_h = 0.3\,h$ (correctly estimated), and $\sigma_h = 0.6\,h$ (overestimated).

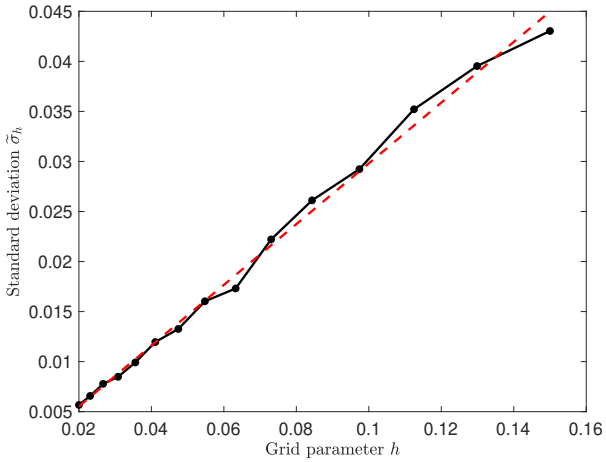

Figure 4: Estimated level of the modeling error as a function of the grid parameter given by the solid curve. The dashed line is the least squares fit, given by $\sigma_h = \alpha h + \beta$ with $\alpha = 0.3035$, $\beta = -0.0005$.

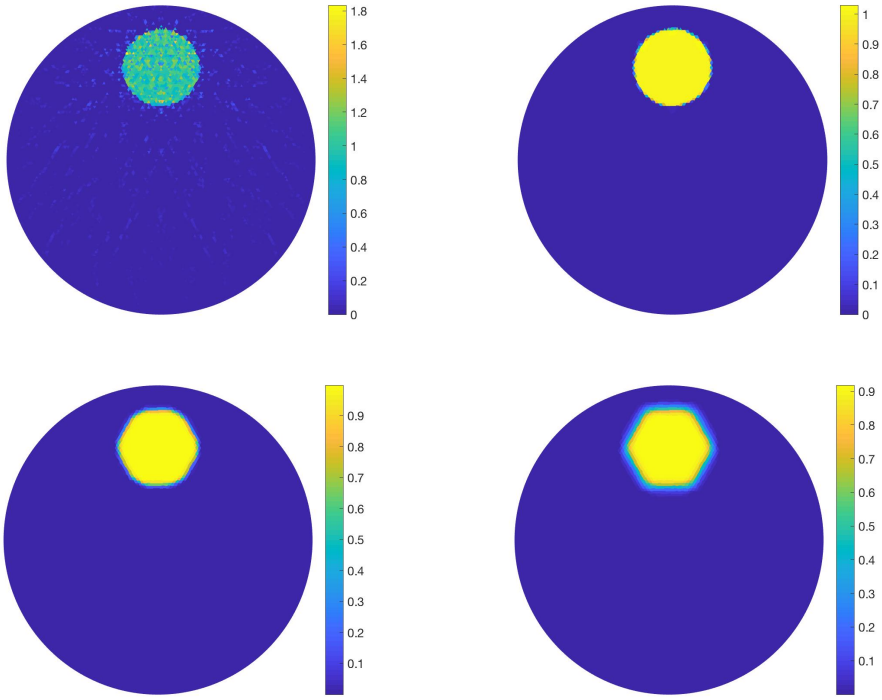

Figure 5: IAS reconstructions after two outer iteration rounds with fixed discretization mesh and with four different estimates of the discretization error, corresponding to, in lexicographical order: $\sigma_h = 0.03\,h$, $\sigma_h = 0.06\,h$, $\sigma_h = 0.3\,h$, and $\sigma_h = 0.6\,h$. The focality parameter of the IAS algorithm controlling the sharpness of the reconstruction is $\eta = 0.001$.

Figure 5 shows the results with the IAS algorithm, the hyperparameter controlling the focality of the solution was set at $\eta = 0.001$. Two observations are in order. First, when the modeling error is strongly underestimated, the edges of the target are well identified but noise clutter is clearly visible in the solution. On the other hand, when the modeling error level is overestimated, the reconstruction of the inclusion is not sharp, and the structure of the underlying discretization mesh becomes visible by making the target hexagonal instead of round. Numerical tests suggest that a rough estimate of the modeling error level is sufficient for a reasonable quality of reconstruction, with a slight preference for underestimating the error level.

Next we ran the BMA algorithm using the modeling error formula $\sigma_h = 0.3 \min(h_j)$, the minimum taken over all edges. In the `FreeFem++` code, the minimum and maximum mesh sizes were set at 0.006 and 0.1, respectively. In the metric update, we set the threshold value to

$$G_j = \min\{\theta_j, 0.5 \times \theta_{\max}\}, \quad \theta_{\max} = \max\{\theta_j\}.$$

Changes in the threshold value slightly affect the width of the region around the boundary of the inclusion discretized with the finer mesh. Figure 6 shows the results at each of the four first iterations: the left panel displays the underlying mesh, and the right one the reconstruction based on that mesh. The IAS algorithm, terminated when the relative change in the θ update falls below 10% of the ℓ_∞ norm of θ, required between 4 and 8 outer iterations. The stopping criterion for the BMA algorithm is based on the three indicators shown in Figure 7, namely, from left to right, the norm of the discrepancy $\|b - \mathsf{A}_h u_h\|$, the complexity measured in terms of the number of nodes of the mesh, and the product of these two numbers. Figure 7 shows that the discrepancy using the initial coarse mesh is high, as one would expect, and subsequently it drops to a level where it stabilizes. While the discrepancy based on the second mesh is slightly smaller than that based on the third, the second mesh is significantly more complex, indicating an initial over-refinement. An indicator taking into account both the discrepancy and the complexity such as the product of the two seems to identify the third mesh as a good compromise. A visual inspection of the meshes confirms that after the third iteration, no significant changes take place, and therefore continuing the iterations would not improve the solution.

The X-ray tomography data with 15 views represent a rather complete set of projections, and it is of interest to see how the proposed methodology works with more limited data, as it is often collected, e.g., in some industrial and medical applications, where the number of views is limited due to geometric constraints, health considerations, or constraints in data acquisition speed. Figure 8 shows the results after three mesh adaptation iterations when the number of views is limited to 3 and 7, respectively, all other parameters remaining the same. The reconstructions, very much in line with what could be expected, exhibit reduced sharpness along the boundary of the inclusion in directions not coinciding with the wave front sets.

5.2 Inverse source problem for Darcy flow

In the second example we consider the inverse source problem for the Darcy flow introduced in Section 2.2. The generative model consists of the piecewise constant source term shown in Figure 2, and the simulated data, collected in a vector of size $m = 400$, are generated by a second order FEM solver with a uniform mesh consisting of $4\,974$ triangular elements, with $9\,897$ degrees of freedom with exogenous additive scaled white noise with standard deviation σ_{ex} equal to 0.1% of the maximum of the noiseless signal.

We initialize the BAM algorithm by generating a coarse structured mesh with $n_t = 780$ triangular elements shown in Figure 9 (first row). In the IAS iterations, we use the parameter values

$$\eta = 0.001, \quad \vartheta_j = \frac{1 \times 10^{-3}}{\|\mathsf{A}_h e_j\|},$$

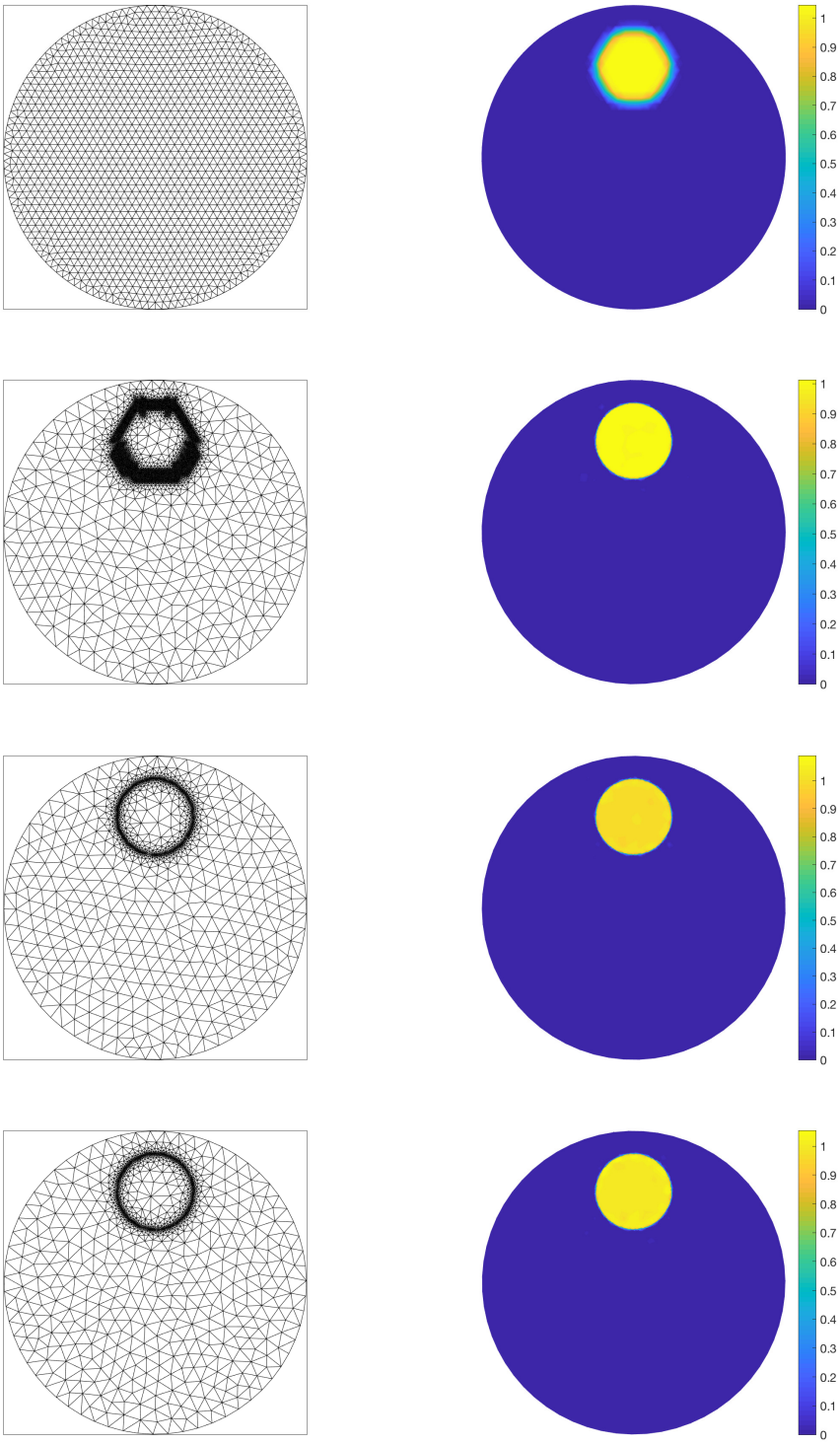

Figure 6: BMA-IAS iterations for the fan beam tomography estimation problem, with 15 views and noise level 0.5%. Each row shows the discretization mesh (left), and the corresponding IAS estimate (right). The mesh in the next row is based on the prior variance of the estimate in the previous row.

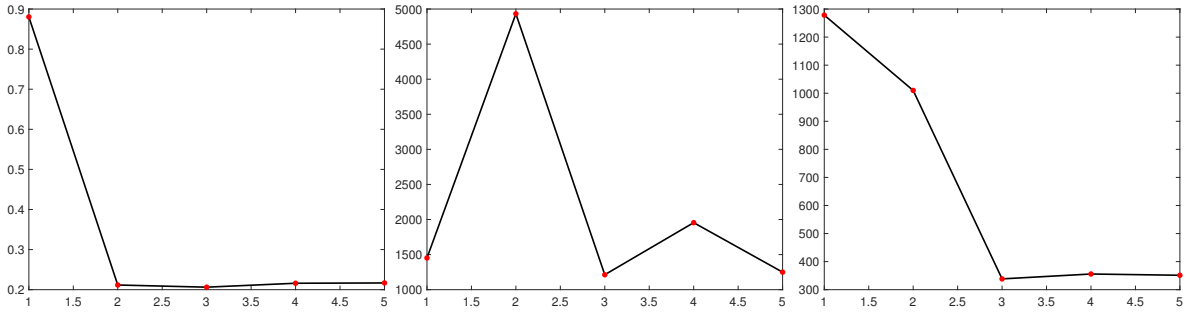

Figure 7: Example 1. Left: The norm of the discrepancy, $\|b - \mathsf{A}_h u_h\|$ (left), the complexity measured in terms of the number, $n_{v,\text{free}}$, of degrees of freedom in the problem (center), and the product $n_{v,\text{free}} \times \|b - \mathsf{A}_h u_h\|$ of the two (right), as a function of the iteration.

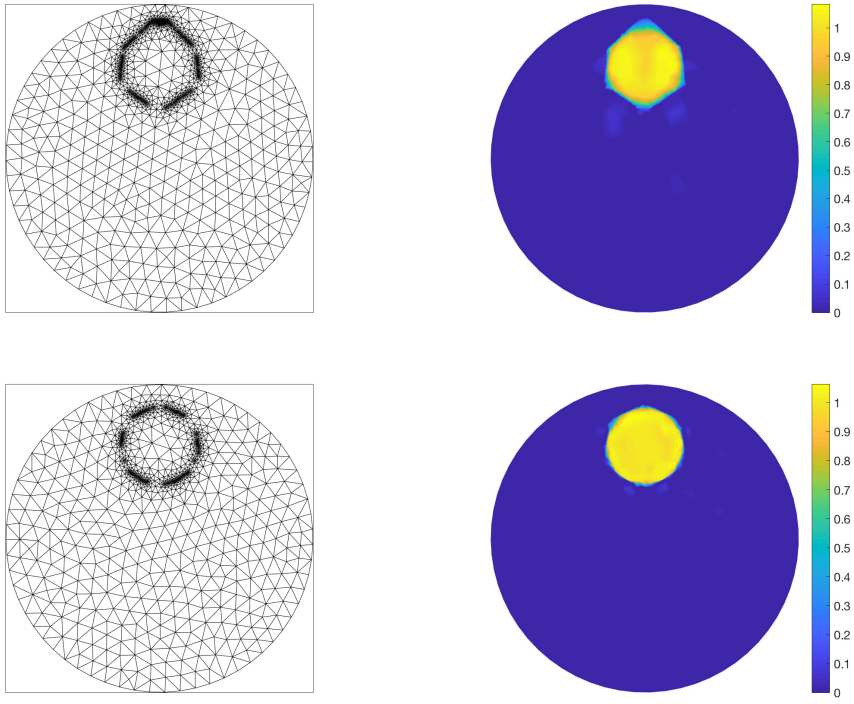

Figure 8: The third iteration of the BMA-IAS iteration mesh and corresponding reconstruction of the density, when the number of views is limited to 3 (top row) and 7 (bottom row). The noise level as well as the model parameters are the same as in Figures 6 and 7.

and stop the outer iterations when the relative change in the prior parameter θ is less than 10% of the previous value. As in the previous example, this stopping condition is satisfied after only 4 to 7 outer iterations. The modeling error is estimated by setting $\sigma_h = 0.3 \times h_{\min}$, in analogy with the previous example, where $h_{\min}$ is the minimum edge length in the current discretization. In the update of the metric, we set

$$G_j = \min\{\theta_j, 0.3 \times \theta_{\max}\}, \quad \theta_{\max} = \max\{\theta_j\}.$$

When computing the new mesh using `FreeFem++`, we specify the minimum and maximum allowed element size to be $h_{\min} = 0.015$ and $h_{\max} = 0.04$ to avoid overly dense meshes that make the FEM solver slow and badly scaled, and too coarse discretization that would make the FEM approximation error too large.

Figure 9 shows the progression of the algorithm by displaying the first four meshes and the corresponding reconstructions of the source term. As in the tomography problem, the first refinement increases significantly the complexity of the system, while the second update reduces the degrees of freedom by identifying the areas where the source term suggests a jump. After the second refinement, the iterations do not change significantly the discrepancy, and the solution of the inverse problem remains visually unaltered. Figure 10 shows the iteration history of the norm of the discrepancy $\|\mathsf{A}_h u_h - b\|$ as well as the number of nodes of the mesh are plotted, revelaing similar behavious as in the tomography problem. Again, the product of the mesh size and the norm of the discrepancy provides a useful and informative indicator, suggesting that after the second adaptation (Mesh 3), there is no reason to continue iterating.

6 Discussion

In this paper we propose a novel Bayesian method for simultaneously solving an inverse problem for a distributed parameter and updating the discretization mesh where to represent the unknown. We illustrate the performance of the proposed method on two test cases, both arising from linear inverse problems. Extending the approach to nonlinear problems is possible, but not immediate, as the underlying IAS algorithm requires a further linearization and iteration steps. In the second example, the mesh determined by the BMA-IAS algorithm is used also for solving the forward model with a higher order finite element method. The order of the FEM approximation is an important feature, as an underlying coarse mesh may be sufficient to represent accurately the distributed parameter of interest, but may fail to produce a sufficiently accurate approximation of the solution of the PDE unless higher order methods are used. To fully control the mesh adaptation process, in addition to the variance parameter of the inverse problem, the accuracy of the FEM solution via the a posteriori error estimates should be monitored. The accuracy of the forward model approximation constitutes an important part of a careful modeling error analysis that is not addressed in detail in this paper. A further extension of the modeling error analysis entails the correlation structure of the modeling error. It is well known that the modeling error due to discretization is typically strongly correlated and has normally a non-vanishing mean. A more careful analysis would require new ideas for dynamically assessing the modeling error without the need to resort to very dense discretization meshes that we aim to avoid by proposing the dynamic mesh updating.

In our discussion, it was assumed that the target mesh at each adaptation should be close to isotropic. This assumption was encoded in the algorithm by assuming that the metric guiding the mesh calculation is such that the eigenvalues of the matrix defining the metric at each point are equal. In principle, this is not necessary, and, in fact, an anisotropic metric gives rise to meshes characterized by stretched elements that are used successfully to represent strongly directional features, see, e.g., [15, 16, 27]. Extension of the present work to anisotropic meshes is a possible future direction, but out of the scope of the present paper.

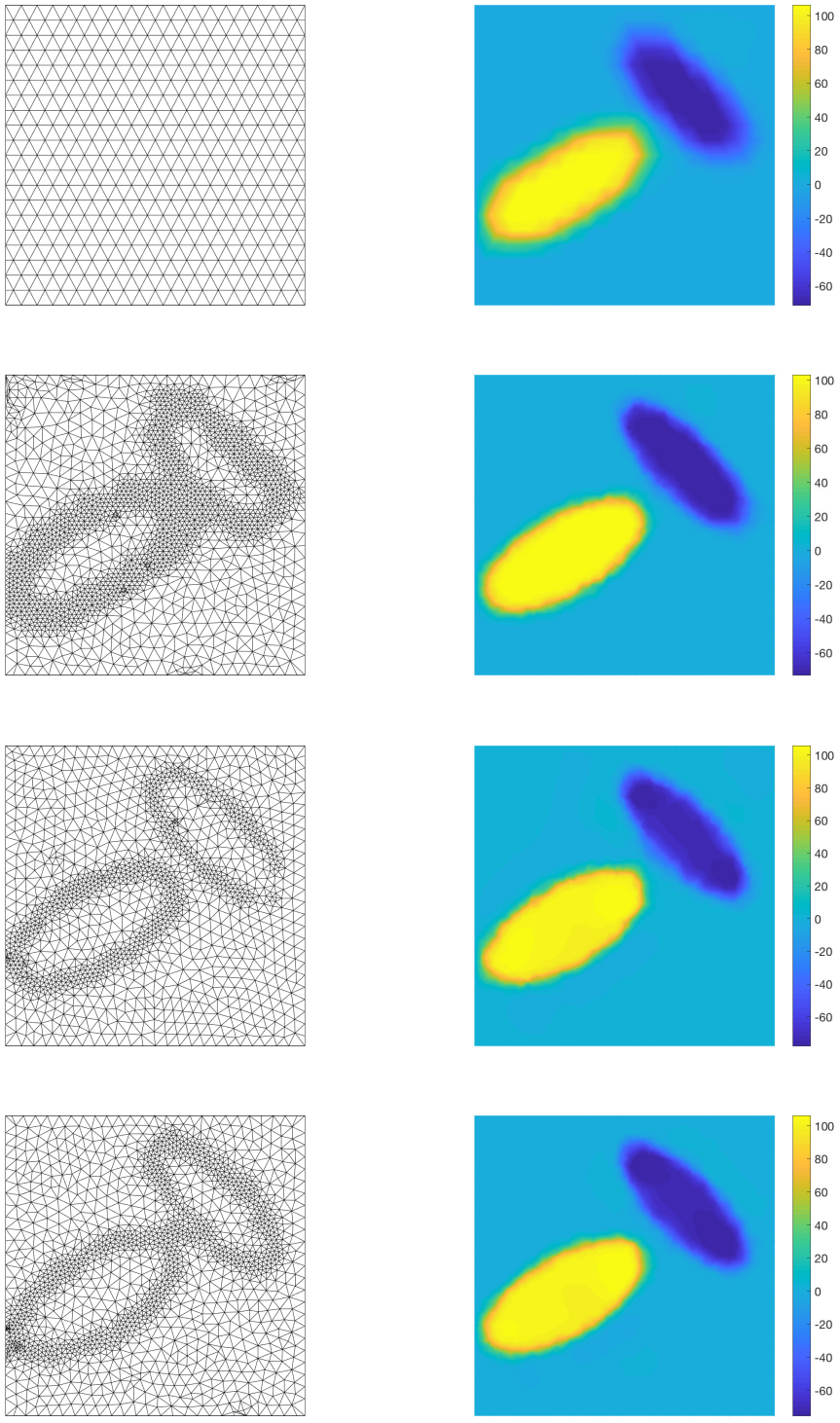

Figure 9: BAM-IAS iterations for the source estimation problem in a Darcy flow model. Each row shows the discretization mesh (left), and the corresponding IAS estimate (right). The mesh in the next row is based on the prior variance of the preceding estimate.

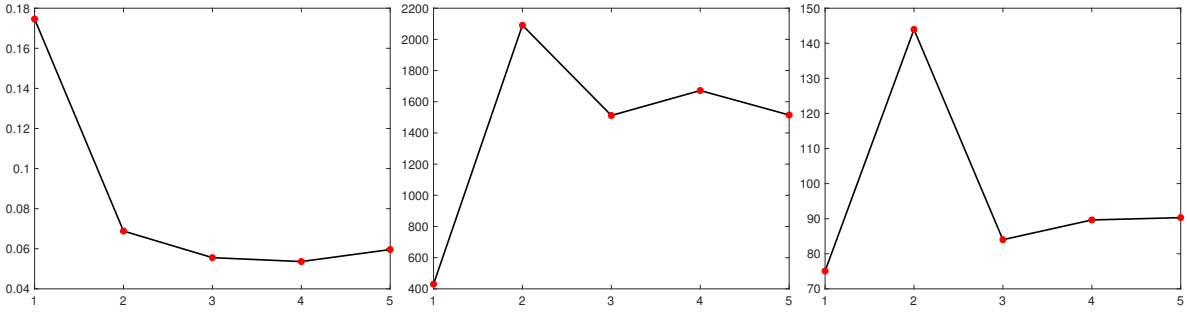

Figure 10: Left: The norm of the discrepancy, $\|b - \mathsf{A}_h u_h\|$ as a function of the iteration (left), the complexity measured in terms of the number $n_{v,\text{free}}$ of degrees of freedom in the problem (center), and the product $n_{v,\text{free}} \times \|b - \mathsf{A}_h u_h\|$ of the two (right).

Acknowledgements

The work of DC and AS was partly supported by NSF grant DMS-1522334, and of ES by NSF grant DMS-1714617. The financial support of INdAM-GNCS Project 2020 is gratefully acknowledged by SP.

References

[1] Ainsworth M and Oden J T (2000) *A Posteriori Error Estimation in Finite Element Analysis.* Wiley, New York.

[2] Arridge SR, Kaipio JP, Kolehmainen V, Schweiger M, Somersalo E, Tarvainen T and Vauhkonen M (2006) Approximation errors and model reduction with an application in optical diffusion tomography. Inverse Problems **22** 175-195.

[3] Bottasso C L, Maisano G, Micheletti S and Perotto S (2006) On some new recovery based a posteriori error estimators. Comput Meth Appl Mech Engr **195**: 4794-4815.

[4] Calvetti D, Dunlop M, Somersalo E and Stuart A (2018) Iterative updating of model error in Bayesian inversion. Inverse Problems **34**: 025008.

[5] Calvetti D, Ernst O and Somersalo E (2014) Dynamic updating of numerical model discrepancy using sequential sampling. Inverse Problems **30**: 114019.

[6] Calvetti D, Hakula H, Pursiainen S and Somersalo E (2009) Conditionally Gaussian hypermodels for cerebral source localization. SIAM J Imaging Sci **2** 879-909.

[7] Calvetti D, Pragiola M, Somersalo E and Strang A (2019) Sparse reconstructions from few noisy data: analysis of hierarchical Bayesian models with generalized gamma hyperpriors. Inverse Problems (to appear).

[8] Calvetti D, Pragiola M and Somersalo E (2020) Hybrid solver for hierarchical Bayesian inverse problems. Manuscript.

[9] Calvetti D, Pascarella A, Pitolli F, Somersalo E and Vantaggi B (2015) A hierarchical Krylov-Bayes iterative inverse solver for MEG with physiological preconditioning. Inverse Problems **31** 125005.

[10] Calvetti D and Somersalo E (2005) Priorconditioners for linear systems. Inverse Problems **21**: 1397-1418.

[11] Calvetti D and Somersalo E (2011) Subjective knowledge or objective belief? an oblique look to bayesian methods. in: *Large-Scale Inverse Problems and Quantification of Uncertainty*, pp.33-70.

[12] Calvetti D, Somersalo E and Strang A (2019) Hierarchical Bayesian models and sparsity: ℓ_2-magic. Inverse Problems **35**: 035003 (26pp).

[13] Cheney M, Isaacson D and Newell JC (1999) Electrical impedance tomography. SIAM Review **41**: 85-101.

[14] Ciarlet Ph G (1978) *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam.

[15] Farrell P E, Micheletti S and Perotto S S (2011) An anisotropic Zienkiewicz-Zhu type error estimator for 3D applications. Int J Numer Meth Engr **85**: 671-92.

[16] Formaggia L, Micheletti S and Perotto S (2002) Anisotropic mesh adaption with application to CFD problems In: Mang H, Rammerstorfer F and Eberhardsteiner J, Eds: *Proceedings of WCCM V, Fifth World Congress on Computational Mechanics* 1481-1493

[17] George P L and Borouchaki H (1998) *Delaunay Triangulation and Meshing. Application to Finite Elements*. Editions Hermès, Paris.

[18] Hecht F (2012) New developement in FreeFem++. J Numer Math **20**: 251-265.

[19] Johnson C (2012) *Numerical Solutions of Partial Differential Equations by Finite Element Method*. Dover Book on Mathematics Series, Dover Publications Inc.

[20] Kaipio JP and Somersalo E (2007) Statistical inverse problems: discretization, model reduction and inverse crimes. J Comp Appl Math **198** 493–504

[21] Kaipio J and Somersalo E (2004) *Statistical and Computational Inverse Problems.* Springer Verlag, New York.

[22] Krìzek M and Neittaanmäki P (1984) Superconvergence phenomenon in the finite element method arising from averaging gradients. Numer Math **45**: 105-116.

[23] Li Y and Oldenburg DW (1996) 3-D inversion of magnetic data. Geophysics 61(2): 394-408.

[24] Li Y and Oldenburg DW (1998) 3-D inversion of gravity data. Geophysics 63(1): 109-119.

[25] Lin FH, Witzel T, Ahlfors SP, Stufflebeam SM, Belliveau JW and Hämäläinen MS (2006) Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. Neuroimage **31**: 160-171.

[26] Micheletti S, Farrell P E and Perotto S (2010) A recovery-based error estimator for anisotropic mesh adaptation in CFD. Bol Soc Esp Mat Apl **50**: 115-138.

[27] Micheletti S and Perotto S (2010) Anisotropic adaptation via a Zienkiewicz-Zhu error estimator for 2D elliptic problems. In: Kreiss G, Lötstedt P, Målqvist A and Neytcheva M, Eds *Numerical Mathematics and Advanced Applications* (Berlin Heidelberg: Springer-Verlag) 645-653.

[28] Naga A and Zhang Z (2004) A posteriori errorestimates based on the polynomial preserving recovery. SIAM J Numer Anal **42** 1780-1800.

[29] Natterer F (2001) *The mathematics of computerized tomography*. SIAM, Philadelphia.

[30] Nissinen A, Heikkinen LM and Kaipio JP (2007) The Bayesian approximation error approach for electrical impedance tomography?experimental results. Meas Sci Technol **19**: 015501.

[31] Nissinen A, Kolehmainen V and Kaipio JP (2011) Reconstruction of domain boundary and conductivity in electrical impedance tomography using the approximation error approach. Int J Uncertainty Quantification **1**: 203-222.

[32] Rodríguez R (1994) Some remarks on Zienkiewicz-Zhu estimator. Numer Meth PDE **10**: 625-635.

[33] Yan N and Zhou A (2001) Gradient recovery type a posteriori error estimation for finite element approximations on irregular meshes. Comp Meth Appl Mech Engr **190**: 4289-4299.

[34] Zienkiewicz O C and Morgan K (2013) *Finite Elements and Approximation*. Dover Book on Engineering Series, Dover Publications Inc.

[35] Zienkiewicz O C and Zhu J Z (1987) A simple error estimator and adaptive procedure for practical engineering analysis. Int. J Numer Meth Engr **24**: 337-357.

[36] Zienkiewicz O C and Zhu J Z (1992) The superconvergent patch recovery and a posteriori error estimates. I: The recovery technique. Int. J Numer Meth Engr **33**: 1331-1364.

[37] Zienkiewicz O C and Zhu J Z (1992) The superconvergent patch recovery and a posteriori error estimates. II: Error estimates and adaptivity. Int. J Numer Meth Engr **33**: 1365-1382.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

21/2020 Benacchio, T.; Bonaventura, L.; Altenbernd, M.; Cantwell, C.D.; Düben, P.D.; Gillard, M.; Gir
Resilience and fault-tolerance in high-performance computing for numerical weather and climate prediction

22/2020 Zeni, G.; Fontana, M.; Vantini, F.
Conformal Prediction: a Unified Review of Theory and New Challenges

23/2020 Spreafico, M.; Ieva, F.
Functional modelling of recurrent events on time-to-event processes

24/2020 Formaggia, L.; Scotti, A.; Fumagalli, A.
Numerical Methods for Flow in Fractured Porous Media

20/2020 Almi, S.; Belz, S.; Micheletti, S.; Perotto, S.
A DIMENSION-REDUCTION MODEL FOR BRITTLE FRACTURES ON THIN SHELLS WITH MESH ADAPTIVITY

19/2020 Stella, S.; Vergara, C.; Maines, M.; Catanzariti, D.; Africa, P.; Dematté, C.; Centonze, M.; Nob
Integration of maps of activation times in computational cardiac electrophysiology

18/2020 Fumagalli, A.; Scotti, A.; Formaggia, L.
Performances of the mixed virtual element method on complex grids for underground flow

17/2020 Cerroni, D.; Formaggia, L.; Scotti, A.
A control problem approach to Coulomb's friction

16/2020 Paolucci, R.; Mazzieri, I.; Piunno, G.; Smerzini, C.; Vanini, M.; Ozcebe, A.G.
Earthquake ground motion modelling of induced seismicity in the Groningen gas field

15/2020 Fumagalli, I.; Fedele, M.; Vergara, C.; Dede', L.; Ippolito, S.; Nicolò, F.; Antona, C.; Scrofani,
An Image-based Computational Hemodynamics Study of the Systolic Anterior Motion of the Mitral Valve