Eddy Current Tomography Using a Binary Markov Model

Mila Nikolova and Ali Mohammad-Djafari
Laboratoire des Signaux et Systèmes (CNRS—ESE—UPS)
École Supérieure d'Électricité, Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France.

Abstract – *The non-destructive evaluation (*NDE*) problem we treat is the testing of a globally homogeneous conductive medium for anomalies such as cracks and notches. The medium is illuminated with a monochromatic electric field; the anomalies induce eddy currents and they modify the total field which can be measured. The tomographic approach, aimed to draw up an image of the medium, is recent in this area. It corresponds to an extremely difficult ill-posed inverse problem and its resolution needs the use of pertinent prior information. The considered anomalies can be represented using images whose pixels can only take the values 0 and 1.*

*Our main contribution lies in the regularization of a large-support ill-posed observation operator using a locally constant binary image Markov random field (*MRF*). The resulting high-dimensional combinatorial optimization problem is tedious: neither exact resolution nor simulated annealing are feasible. Instead, we establish an equivalent continuous-valued optimization problem. A nearly optimal solution is then calculated using a graduated non-convexity (*GNC*) algorithm adapted for this purpose.*

The proposed inversion technique surpasses the particular NDE *problem and can be applied whenever a binary image is observed using a linear system and corrupted by Gaussian noise.*

1 Introduction

Conductive medium ageing manifests itself in the appearance of *cracks* and *notches*. Their detection in very early stages and their surveillance may be of crucial importance in various fields such as nuclear power plants and aerospace engineering.

The objective of *eddy current* (EC) testing is to characterize a material – globally homogeneous, but within which some structural anomalies may be present. The material is placed inside an electric field and the resultant (anomalous) field is measured outside the material. The tomographic imaging approach in this context is aimed at determining an *image* which represents the anomaly layout. It relies on a mathematical description of the link between image and data, and allows a finer characterisation than usual inspection while presenting a greater challenge.

We consider the 2D plane geometry observation model, developed in [1] and applied to a conductive and non-magnetic material in order to check it for the presence of anomalies such as cracks and notches. In the frame of the Born approximation, the image-data relation can be reduced to a *Fourier-Laplace* transform (FLT). This *forward problem* is briefly formulated in Section 2. It comprises a formidable information loss (Section 3) which, after discretization, makes the problem severely ill-conditioned (Section 4).

The reconstruction of the image constitutes hence an extremely difficult *ill-posed inverse problem* [2]. In Bayesian terms, its resolution necessitates the introduction of judicious prior information, in order to compensate for the poor character of the data. At the same time, the numerical cost of the forward problem does not favour the use of sophisticated priors. Usual priors are Gaussians (the y act as quadratic regularizations) [3], [4] but they are not adequate. The introduction of maximum entropy priors permits an improvement [5], although it does not account for the pixel correlation.

The desired image is *zero-one binary* and *locally constant*. This constitutes an extremely strong statistical prior information which we incorporate into a *maximum a posteriori* (MAP) estimation. Such a prior can be represented by a binary MRF [6] (Section 5). However, this MAP estimation comes up against a combinatorial optimization problem of great complexity, impractical using conventional means.

In order to deal with it, we first transform the MAP criterion on binary variables into an equivalent family of criteria on continuous valued variables (Section 7).

We focused on a continuation optimization approach, namely the GNC [7] (Section 6), which has been applied with success for the calculation of solutions regularized using pairwise interaction MRFs [8]. In oredr to allow its use in the actual context, we transform the MAP criterion on binary variables into an equivalent family of criteria on continuous valued variables (Section 7). Then the solution is reached after a sequence of local descents (Section 8).

Simulation results are presented in Section 9 and concluding remarks in Section 10. Theoretical derivations are exposed in the Appendix.

2 Eddy current tomography model

We consider a globally homogeneous, linear, non-magnetic and isotropic conductive medium, which is illuminated by a known time-harmonic normally incident excitation field $\mathcal{E}_0$ with frequency ν. The *image* of the interior structure of the medium can be defined as a contrast function:

$$f(\boldsymbol{r}) = \frac{k_0^2 - k^2(\boldsymbol{r})}{k_0^2} \simeq \frac{\sigma_0 - \sigma(\boldsymbol{r})}{\sigma_0}, \tag{1}$$

where k_0 and σ_0 are respectively the wavenumber and the conductivity inside the homogeneous regions; $k(\boldsymbol{r})$ is the wavenumber as a function of the position $\boldsymbol{r}$ and similarly $\sigma(\boldsymbol{r})$ is the conductivity as a function of $\boldsymbol{r}$. The approximation in (1) is valid when the permeability μ is constant. Since the considered anomalies are *voids* (cracks and notches), inside them $\sigma(\boldsymbol{r}) = 0$. So

$$f(\boldsymbol{r}) = \begin{cases} 1 & \text{(black)} \quad \text{inside an anomaly} \\ 0 & \text{(white)} \quad \text{in the homogeneous part} \end{cases}.$$

The EC testing relies upon the fact that the anomaly modifies the total field $\mathcal{E} = \mathcal{E}_0 + \mathcal{E}_a$, where $\mathcal{E}_a$ is the anomalous field. The link between the image f and the total field $\mathcal{E}$ comes from the Helmholtz equation: $[\Delta + k_0^2]\mathcal{E}(\boldsymbol{r}) = k_0^2 f(\boldsymbol{r})\mathcal{E}(\boldsymbol{r})$. Then $\mathcal{E}_a$ is related to the image f by a Fredholm integral equation of the second type:

$$\mathcal{E}_a(\boldsymbol{r}) = k_0^2 \int_A G(\boldsymbol{r}, \boldsymbol{r}') f(\boldsymbol{r}') \mathcal{E}(\boldsymbol{r}') d\boldsymbol{r}', \tag{2}$$

where the integration is computed over the anomaly interior A and G stands for Green function. This equation is non-linear in $\mathcal{E}_a$. It is classicaly linearised by the Born approximation which makes the assumption that inside an anomaly, the anomalous field can be neglected: $\mathcal{E}(\boldsymbol{r}) = \mathcal{E}_0$ for $\boldsymbol{r} \in A$.

The resolution of (2) for the 2D plane geometry (see Fig. 1 (a)) using the Born approximation has been developped in [1]. The measured quantity is the potential variation induced into a sensing coil along a line (l) under the conductive material and parallel to it at the height $-y_0$, and is directly related to the anomalous magnetic field $\mathcal{H}_a$. For a normally incident excitation field $\mathcal{E}_0$, the magnetic and the electric anomalous fields are related by

$$\frac{\partial \mathcal{E}_a}{\partial x} = -i\mu_0\omega\mathcal{H}_a, \tag{3}$$

where $\mathcal{H}_a$ has only one component along the y-axis. The integral formulation relating the image f and the observations $\mathcal{H}_a$ reads

$$\mathcal{H}_a(x;y_0,\nu) = \underbrace{-\frac{T}{2\delta^2\pi^2\mu_0\nu}}_{K(\nu)}\int_R e^{iux}\underbrace{\frac{ue^{-i\beta_1 y_0}}{\beta_1+\beta_0}}_{C_\nu(u)}\underbrace{\iint_A e^{-iux'}e^{i(k_0+\beta_0)y'}f(x',y')dx'dy'}_{L_\nu(u)}du, \tag{4}$$

where $\delta = (\pi\nu\mu_0\sigma_0)^{-\frac{1}{2}}$ is the skin depth (μ_0 is the magnetic permeability), u is the spacial frequency corresponding to the x-axis, $\beta_1 = \sqrt{k_1^2 - u^2}$ and $\beta_0 = \sqrt{k_0^2 - u^2}$ are the propagation factors in the air and the conductive medium respectively. Furthermore, $k_1 = 2\pi\nu\sqrt{\mu_0\epsilon_0}$ is the wavenumber in the air (ϵ_0 means dielectric permittivity), while the wavenumber in the conductive medium is $k_0 = 1/\delta + i/\delta$ which means that in this medium, the propagation equals the attenuation. The transmission coefficient on the air-metal interface is $T \simeq 2k_1/k_0$.

Since f is linearly related to $\mathcal{H}_a$, the observation relation (4) may be rewritten as

$$\mathcal{H}_a(x;y_0,\nu) = \mathcal{A}f(x;y_0,\nu) \tag{5}$$

where $\mathcal{A}$ is the integral operator which, applied to f, leads to $\mathcal{H}_a$ according to (4).

Measures of $\mathcal{H}_a(x;y_0,\nu)$ are in practice collected for a set of frequences $[\nu_1,\ldots,\nu_m]$. The parameters describing the observation configuration are listed in Section 9. Simulated magnetic-field data corresponding to the anomaly in Fig. 6 (a), are shown in Fig. 1 (b); each observation curve corresponds to a frequency.

3 Analysis of the ill-posedness

Let us split the observation model (4) into:

$$\mathcal{H}_a(x;y_0,\nu) = K(\nu)\mathcal{F}_u^{-1}\{C_\nu L_\nu\}(x;y_0,\nu), \tag{6}$$

where $\mathcal{F}_u^{-1}$ means Inverse Fourier transform (FT) with respect to u. The observation equation is seen to be extremely ill-posed (which explains why data are "poor"), for three major reasons.

L_ν constitutes a Fourier-Laplace transform of the image $f(x,y)$: the transformation along its x-axis is a FT, while the exponent along the y-axis $(k_0+\beta_0)$ is complex, so it defines a Laplace transform[1](LT):

$$p_\nu(u) = -i(k_0+\beta_0) = [\frac{1}{\delta} + \Im(\beta_0)] - i[\frac{1}{\delta} + \Re(\beta_0)].$$

The real part $\Re(p_\nu)$ of the exponent performs an *attenuation* along the y-axis whose magnitude increases with ν and with u (Fig. 2 (a)). (Besides, some inversion techniques, discussed in [3], neglect this attenuation partially or totally, in order to reformulate (4) as a FT.)

The imaginary part $\Im(p_\nu)$ (Fig. 2 (b)) defines some algebraic curves in the FT2D frequency domain. Data points are located along these curves. Their repartition is very irregular which entails that $\mathcal{H}_a$ represents very incompletely the features of f. Besides, an analytical inversion of (5) is not feasible, even in an idealized situation where data are exact samples of $\mathcal{H}_a$.

Furthermore, $C_\nu(u)$ originates from a Green function and produces an extremely severe pass-band filtering (Fig. 2 (c)). Its effect can be qualitatively evaluated in Fig. 1 (b): data are extremely smooth, *i.e.* high-frequency information is lost. Due to (3), the low pass filtering effect for electric field data is stronger than for magnetic field data: the inverse problem is slightly "less difficult" in the second case.

4 The discrete inverse problem

Numerical processing requires the observation relation (5) to be discretized. This question is of crucial importance for the success of the inversion. Discretization is classically performed using the method of moments.

The image of the interior structure $\boldsymbol{f}$ is then an array of M pixels $\boldsymbol{f}$ on a $I \times J$ $(M = IJ)$ lattice $\mathcal{S}$. The N complex data values are available as a set of vectors $\boldsymbol{h}$. The integral operator becomes a matrix-type operator $\mathcal{A}$. Measured data present some

[1]The LT of $f(t)$ is defined as $F(p) = \int f(t)e^{-pt}dt$.

random deviations from theoretical equations, which is modeled as an additive white Gaussian noise $\boldsymbol{n}$ with $\Re(n_i), \Im(n_i) \sim \mathcal{N}(0, \sigma^2)$. The observation relation in matrix notation reads:

$$\boldsymbol{h} = \mathcal{A}\boldsymbol{f} + \boldsymbol{n}, \quad \boldsymbol{f} \in \mathbb{B}^M, \quad \mathbb{B} = \{0, 1\}, \tag{7}$$

where $\mathbb{B}^M \doteq \mathbb{B}^{I \times J}$ is the set of binary images. (By a slight abuse of language, we assimilate $\mathbb{R}^{I \times J}$ to $\mathbb{R}^M$.)

In the following, a binary image is noted $\boldsymbol{f}$ while a continuous valued image $\mathbf{f}$. The last relation (7) in the general case when $\mathbf{f} \in \mathbb{R}^M$, is the most common observation model in applied physics, inverse problems and image processing. However, essential particularities subsist in our context.

Firstly, the operator $\mathcal{A}$ combines several formidable numerical defaults:

(a) it comes from an ill-posed integral equation, so it is in general singular, or at least seriously ill-conditioned;

(b) it performs an *attenuation* along the y-axis which means that deep pixels are more weakly represented in the data than shallow pixels;

(c) it has a large support – a datum y_j results from the contributions of the entire image f;

(d) it has no particular structure favourable to be exploited numerically.

Secondly, the image is binary ($\boldsymbol{f} \in \mathbb{B}^M$) while the set of binary images is *non convex*[2]. There cannot be any convex (neither locally nor globally) optimality criterion defined on such a set, so working on it is difficult. The binary character of $\boldsymbol{f}$ is usually dropped in NDE applications and continuous valued images $\mathbf{f} \in \mathbb{R}^M$ are reconstructed.

The observation operator $\mathcal{A}$ is not invertible and the data are noisy, so that direct inversion of (7) is unthinkable. Back-projection algebraic techniques [3] are known to produce important artefacts when data are incomplete and noisy.

The likelihood $p(\boldsymbol{h}|\mathbf{f})$ of the data $\boldsymbol{h}$ conditionally to the image $\mathbf{f}$ evaluates the plausibility for an image to give rize to actual data $\boldsymbol{h}$. The maximum likelihood estimate

[2]The definition of convexity fails to apply: $\forall \boldsymbol{f}_1 \in \mathbb{B}^M, \quad \forall \boldsymbol{f}_2 \in \mathbb{B}^M, \quad \forall \mu \in]0,1[\quad \Rightarrow \mu \boldsymbol{f}_1 + (1-\mu)\boldsymbol{f}_2 \notin \mathbb{B}^M$.

$\hat{\mathbf{f}}_{\mathrm{ml}}$ is continuous valued and equals the least squares solution:

$$\begin{aligned}\hat{\mathbf{f}}_{\mathrm{ml}} &= \arg\max_{\mathbf{f}\in\mathbb{R}^M} p(\boldsymbol{h}|\mathbf{f}) = \arg\min_{\mathbf{f}\in\mathbb{R}^M} \mathcal{L}(\boldsymbol{h}|\mathbf{f}),\\ \mathcal{L}(\boldsymbol{h}|\mathbf{f}) &= \|\mathcal{A}\mathbf{f}-\boldsymbol{h}\|^2 \propto -\log p(\boldsymbol{h}|\mathbf{f}) + \mathrm{const}.\end{aligned}$$

It is hardly surprising that the estimate is not unique and is extremely unstable: the inverse problem (7) is seriously underdetermined and its resolution requires this lack to be compensated using priors.

The conditions stated earlier (a), (b) make it *essential* that priors be strong and suited. Besides, it is very tempting to account for the binary character of the unknown image. At the same time (c) allows only the use of optimization algorithms which do not require the observation operator to be local [9].

5 MAP inversion using a binary image prior

Bayesian estimation [10] permits to combine both information contained in the data $\mathcal{L}(\boldsymbol{h}|\boldsymbol{f})$ and probabilistic prior guesses about the reconstructed image

$$P(\boldsymbol{f}) = \frac{1}{Z}\exp\left[-U(\boldsymbol{f})\right],$$

where U means prior energy and contains prior guesses and knowledge; Z is the partition function. Inferences about $\boldsymbol{f}$ rely upon the posterior likelihood

$$\begin{aligned}P(\boldsymbol{f}|\boldsymbol{h}) &= \frac{1}{p(\boldsymbol{h})}p(\boldsymbol{h}|\boldsymbol{f})P(\boldsymbol{f}) = \frac{1}{Z^P}\exp\left[-U^P(\boldsymbol{f})\right],\\ U^P(\boldsymbol{f}) &= \mathcal{L}(\boldsymbol{h}|\boldsymbol{f}) + U(\boldsymbol{f}),\end{aligned}$$

where U^P is the posterior energy and Z^P its normalization factor. Bayesian estimators are built upon the posterior law. The *maximum a posteriori* (MAP) estimator appears quite handy for image reconstructions. It selects the image $\hat{\boldsymbol{f}}$ which is the most probable *conditionally* to observed data $\boldsymbol{h}$:

$$\hat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} P(\boldsymbol{f}|\boldsymbol{h}) = \arg\min_{\boldsymbol{f}}[\mathcal{L}(\boldsymbol{h}|\boldsymbol{f}) + U(\boldsymbol{f})]. \tag{8}$$

Recall that $U(\boldsymbol{f})$ in (8) can also be regarded as a regularization term [2]

The main priors which should be conveyed by U in our case, are:

(a) the image $\boldsymbol{f}$ is binary (the interior of the anomalies is air);

(b) it is constituted from locally constant black patches (anomalies are sizeable);

(c) the background is white (the investigated medium is globally sound).

These elements constitute an extremely *strong* probabilistic prior knowledge – since the set of the solutions which are plausible (i.e. satisfy $(a)-(c)$) is drastically restricted.

It is easy to construct probabilistic models which take (a)-(c) into account; the major difficulty resides in the calculation of the relevant estimate. In practice, U has often resulted from a compromise between prior knowledge and numerical feasability; then U (and U^P) has been chosen convex [4], [3], [5]. However, convex energies can model neither piecewise constant features, nor binary variables.

In spite of these difficulties, we can afford to satisfy $(a)-(c)$ using a *binary Markov random field* since we have some ideas about the calculation of the estimate.

We model, after [6], [11], the image $\boldsymbol{f}$ as a realization of a homogeneous pairwise interaction binary MRF with respect to a neighbourhood system $\mathcal{V} = \{\mathcal{V}_i, i \in \mathcal{S}\}$. Let $\mathcal{V}$ be the 4 nearest-neighbour system: $\mathcal{V}_i$ is the set of four pixels directly adjacent to i (Fig. 3 (a)). Neighbouring pairs, noted $(i \sim j)$, form $2IJ - I - J$ cliques; the rest of the cliques are M singletons. The prior energy cumulates theirs potentials ϕ:

$$U(\boldsymbol{f}) = \alpha \sum_{(i\sim j)} \phi(f_i - f_j) + \beta \sum_{i\in\mathcal{S}} \phi(f_i), \quad \boldsymbol{f} \in \mathbb{B}^M, \tag{9}$$

$$\phi(t) = 1 - d(t), \quad t \in \{0,1\}, \tag{10}$$

where d stands for Kronecker function ($d(0) = 1$ and $d(t) = 0,\ t \neq 0$).

Remark 1 *The potential function ϕ (10) can also be written as $\phi(t) = t^2,\ t \in \{0,1\}$; despite the quadratic power, $\phi(t)$ is not convex since t is binary. It is then impossible to build a convex function upon $\mathbb{B}$: let $s, t \in \mathbb{B}$ and $\mu \in \mathbb{B}$, then $\phi[\mu s + (1-\mu)t]$ is not defined.*

Introduce the following notations: $\mathcal{N}_{(0)}$ and $\mathcal{N}_{(1)}$ the number of respectively white (0) and black (1) pixels ($\mathcal{N}_{(0)} + \mathcal{N}_{(1)} = M$); $\mathcal{N}_{(=)}$ and $\mathcal{N}_{(\neq)}$ the number of neighbour pairs with respectively like colours and distinct colours (their difference is either 1 or

-1), ($\mathcal{N}_{(=)} + \mathcal{N}_{(\neq)} = IJ - I - J$). Then $\mathcal{N}_{(\neq)}$ measures the anomaly contour length and $\mathcal{N}_{(1)}$ measures the anomalous regions size. The prior energy (9) can be rewritten as

$$U(\boldsymbol{f}) = \alpha \mathcal{N}_{(\neq)} + \beta \mathcal{N}_{(1)}. \tag{11}$$

$\alpha > 0$ favours pixels to form patches and $\beta > 0$ favours white background. This field can also be regarded as a label field.

If $\mathcal{V}$ is the 8 nearest-neighbour system ($\mathcal{V}_i$ is the set of 8 pixels directly or diagonally adjacent to i) (see Fig. 3 (b)), the prior energy reads

$$U(\boldsymbol{f}) = \alpha \mathcal{N}_{(\neq)} + \frac{\alpha}{\sqrt{2}} \tilde{\mathcal{N}}_{(\neq)} + \beta \mathcal{N}_{(1)},$$

where $\tilde{\mathcal{N}}_{(\neq)}$ counts the diagonal neighbours with distinct colours. For simplicity, in the following we only treat the 4 nearest-neighbour case.

The MAP criterion $\mathcal{J}(\boldsymbol{f})$ is derived from the posterior energy (8), (11):

$$\mathcal{J}(\boldsymbol{f}) = \|\mathcal{A}\boldsymbol{f} - \boldsymbol{h}\|^2 + \alpha \mathcal{N}_{(\neq)} + \beta \mathcal{N}_{(1)}, \quad \boldsymbol{f} \in \mathbb{B}^M. \tag{12}$$

$\mathcal{J}(\boldsymbol{f})$ is *discrete*: it takes on continuous values "sampled" on the discrete M-dimensionnal grid $\mathbb{B}^M$.

Direct minimization of $\mathcal{J}(\boldsymbol{f})$ is prohibitive since 2^M images must be compared! The exact MAP estimate for a binary image can (in principle) be evaluated using a variant of the Ford-Fulkerson algorithm [12]; a prerequisite is that each datum h_i depends on $\boldsymbol{f}$ only through f_i. However, (7) presents just the opposite feature: the support of $\mathcal{A}$ is the entire image, so each datum h_i depends on all the pixels of the image. The posterior distribution $P(\boldsymbol{f}|\boldsymbol{h})$ exhibits *global interactions* – each pixel i is neighbour of each other pixel $j \in \mathcal{S} \backslash \{i\}$. This is also the reason why simulated annealing (SA) cannot be used either [9]. SA evolves by locally updating each pixel f_i according to its posterior conditional distribution $P(f_i | f_j, j \in \mathcal{V}_i, \boldsymbol{h})$; the latter requires an evaluation of (7) and leads to a prohibitive cost. The problem remains basically unchanged with the ICM algorithm [6], which is a deterministic approximation of SA. Another deterministic optimization technique is mean feald annealing [13], but it requires $\mathcal{A}$ diagonal and invertible and the image continuous valued. Markovian MAP inversion of (7) needs an optimization technique, independent of the range of the posterior interactions and working on binary images.

6 The GNC idea

Graduated non-convexity (GNC) was proposed by Blake & Zisserman [14] in the field of image segmentation and noise cancellation. It is aimed to find the global minimum of a *weak membrane*

$$\mathcal{G}(\mathbf{f}) = \|\mathbf{f} - \boldsymbol{h}\|^2 + \lambda \sum_{(i\sim j)} \phi(\mathrm{f}_i - \mathrm{f}_j), \quad \phi(\mathrm{t}) = \min\{\alpha \mathrm{t}^2, 1\}, \quad \mathbf{f} \in \mathbb{R}^M.$$

Two capital differences with respect to (12) are to be stressed: $\mathcal{G}$ is *continuous* since $\mathbf{f} \in \mathbb{R}^M$ and the observation operator $\mathcal{A}$ is the identity. GNC is a deterministic suboptimal technique which proved more effective than simulated annealing [14]. The GNC idea has been extended by Leclerc to potentials involving Kroneckers [15] in order to treat images. Its generalization to ill-posed inverse problems has been developped in [8]. GNC reveals a very good practical global convergence in a broad class of situations.

The underlying idea of the GNC is the following. Consider the minimization of $\mathcal{G}(\mathbf{f})$, $\mathbf{f} \in \mathbb{R}^M$. A family of *relaxed* criteria $(\mathcal{G}_r(\mathbf{f}))_{r>0}$, parameterized by r, is constucted so that

(a) these are approximations of the true criterion $\mathcal{G}(\mathbf{f})$ and $\lim_{r\to\infty} \mathcal{G}_r = \mathcal{G}$;

(b) $(\mathcal{G}_r)_{r>0}$ are twice differentiable with respect to $\mathbf{f}$;

(c) it exists r_0 such that $\mathcal{G}_{r_0}$ is *convex*.

Let $[r_0, r_1, \ldots]$ be an increasing relaxation sequence. The minimizer of $\mathcal{G}_{r_0}$ is unique:

$$\hat{\mathbf{f}}^{(0)} = \arg \min_{\mathbf{f}\in\mathbb{R}^M} \mathcal{G}_{r_0}(\mathbf{f}).$$

Starting with $\hat{\mathbf{f}}^{(0)}$, a sequence of local minima is tracked where each minimum is computed by *local descent* in the attraction valley $\mathcal{W}$ of the previously obtained minimum:

$$\text{for} \quad k := 1, 2, ..., \quad \hat{\mathbf{f}}^{(k)} = \arg \min_{\mathrm{f}\in\mathcal{W}(\hat{\mathrm{f}}^{(k-1)})} \mathcal{G}_{r_k}(\mathbf{f}).$$

7 Continuous formulation of the discrete criterion

GNC is insensitive to the posterior neighbourhood size and thus it is a good candidate for our optimization problem (12). However, it requires $\mathbf{f} \in \mathbb{R}^M$ while $\mathcal{J}$ is discrete. In order to apply a GNC type approach, we *relax* $\mathcal{J}$ to an equivalent family of continuous optimality criteria.

Consider the following r-parameterized family:

$$\begin{aligned}
\mathcal{G}_r(\mathbf{f}) &= \|\mathcal{A}\mathbf{f}+\boldsymbol{h}\|^2+\Psi_r(\mathbf{f})+\Phi_r(\mathbf{f}), \quad \mathbf{f}\in\mathbb{R}^M, \\
\Psi_r(\mathbf{f}) &= -\sum_{(i\sim j)}[\frac{\alpha+r}{r}\rho_r(\mathrm{f}_i-\mathrm{f}_j)+\rho_r(\mathrm{f}_i-\mathrm{f}_j-1)+\rho_r(\mathrm{f}_i-\mathrm{f}_j+1)], \\
\Phi_r(\mathbf{f}) &= -\sum_{i\in\mathcal{S}}[\frac{\beta+r}{r}\rho_r(\mathrm{f}_i)+\rho_r(\mathrm{f}_i-1)].
\end{aligned} \tag{13}$$

$(\rho_r)_{r>0}$ is a regularization family (up to a multiplicative factor)

$$\rho_r(\mathrm{t}) = r\rho(r\mathrm{t}). \tag{14}$$

The function $\rho(\mathrm{t})$ should be

(a) symmetric and positive $\rho(\mathrm{t})=\rho(-\mathrm{t})>0$, and monotone decreasing on $[0,\infty[$;

(b) twice differentiable which is concave on a connected interval;

(c) integrable $\int_{-\infty}^{+\infty}\rho(\mathrm{t})d\mathrm{t}=C<\infty$;

(d) $\rho(0)=1$.

It follows [16] that

$$\rho_r(\mathrm{t})\to C\delta(\mathrm{t}) \quad \text{as} \quad r\to\infty.$$

As $r\to\infty$, $\rho_r(\mathrm{t})$ becomes a narrow pulse: its height grows at the same rate as its width shrinks. The idea is that such functions, located at 0 and 1, make the continuous MAP criterion infinitely "lower" for binary variables than for any other values.

We show in the following that global optimization of $(\mathcal{G}_r)_{r>0}$ is equivalent to global optimization of $\mathcal{J}$.

Proposition 1 *The binary images ($\mathbb{B}^M$) become minima of $(\mathcal{G}_r)_{r>0}$ as $r\to\infty$ which minimize the criterion better than any local minimum having non-binary elements.*

Then the minima of $\mathcal{J}$ become minima of $(\mathcal{G}_r)_{r>0}$ as $r\to\infty$. On the other hand, $(\mathcal{G}_r)_{r>0}$ keep some shallow local minima which converge towards images having some continuous valued elements. In particular, a generalized inverse $\mathbf{f}^\dagger$ which satisfies $\|\mathcal{A}\mathbf{f}^\dagger-\boldsymbol{h}\|^2=0$ becoms the most shallow local minimum.

Proposition 2 *Let $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ be binary images. Then*

$$\mathcal{J}(\boldsymbol{f}_1) < \mathcal{J}(\boldsymbol{f}_2) \quad \Leftrightarrow \quad \mathcal{G}_r(\boldsymbol{f}_1) < \mathcal{G}_r(\boldsymbol{f}_2), \quad \forall r > 0. \tag{15}$$

Indeed, (13) was designed to keep the altitude difference between any couple of binary images $\mathcal{G}_r(\boldsymbol{f}_1) - \mathcal{G}_r(\boldsymbol{f}_2)$ *independent* of r. The proofs of the Propositions can be found in the Appendix. A direct consequence of these is that the relative weights in terms of criterion value of the minima of $\mathcal{J}$ and of the binary minima of $(\mathcal{G}_r)_{r>0}$ become similar as $r \to \infty$. In particular, $\mathcal{J}$ and $\mathcal{G}_r$ have the same global minimum as $r \to \infty$.

Various functions can serve to construct $(\rho_r)_{r>0}$. We found it suitable to use a Lorentzian shaped function (which is infinitely differentiable), Fig. 4:

$$\rho(\mathrm{t}) = \frac{1}{1+\mathrm{t}^2}, \quad \text{then} \quad \rho_r(\mathrm{t}) = \frac{r}{1+r^2\mathrm{t}^2}, \quad \mathrm{t} \in \mathbb{R}.$$

Relaxed criteria $(\mathcal{G}_r)_{r>0}$ corresponding to a scalar discrete criterion are presented in Fig. 5 and illustrate the foregoing discussion.

Remark 2 (Kronecker potentials) *Efficient segmentation of an image into a given set of levels should be performed using regularized Dirac distributions, rather then using Kronecker functions. A prior energy based on the potential $\phi(\mathrm{t}) = \alpha d(\mathrm{t}) + \lambda d(1-\mathrm{t})$ favours the apparition of "many" zeros and ones in the solution, but other values are not prohibited. For example $\mathcal{J}(\mathrm{t}) = (2.75\mathrm{t} - 1.1)^2 + d(\mathrm{t}) + d(1-\mathrm{t})$ has two local minima at 0 and 1, but the global minimum is $\hat{\mathrm{t}} = 0.4$.*

8 Initial convex criterion

$(\mathcal{G}_r)_{r>0}$ (13) is likely to be a relaxed family of criteria since it satisfies the conditions (a) and (b) of the Section 6. It remains to find an initial convex approximation.

Following [7], convexity of a criterion $\mathcal{G}_r$ is tested by checking whether its Hessian $\mathcal{H}_r$ is positive definite. The GNC approach has already been extended to the optimization of continuous valued criteria involving general observation operators and general continuous interaction functions ϕ [8]. Finding a convex criterion in the actual context is basically similar though it leads to different mathematical expressions. The following results are established in the Appendix.

Let η_r be the maximum convexity value reached by ρ_r: $\eta_r = \max \rho_r''(\mathrm{t}) = r^3/2$. Let D_h and D_v be the difference operators which furnish respectively the horizontal and vertical differences between neighbour pairs, and set $\mathcal{D}^T\mathcal{D} = D_h^T D_h + D_v^T D_v$ (where T stands for transposed). Let $\mu_{\min}$ denotes the smallest eigenvalue of $\mathcal{A}^\dagger\mathcal{A}$ while $\nu_{\max}$ is the largest eigenvalue of $\mathcal{D}^T\mathcal{D}$ (then $\nu_{\max} = 8$ whose calculation can be found in [7]).

Whenever $\mu_{\min} > 0$ ($\mathcal{A}^\dagger\mathcal{A}$ is invertible), $\mathcal{G}_r$ is convex for $r < r_0$ where r_0 is the positive solution[3] of

$$2\mu_{\min} = \eta_r \frac{3r+\alpha}{r}\nu_{\max} + \eta_r \frac{2r+\beta}{r}. \tag{16}$$

A difficulty arises when $\mu_{\min} \approx 0$ or $\mu_{\min} = 0$ ($\mathcal{A}^\dagger\mathcal{A}$ is respectively ill-conditioned or singular). In the first case, mathematical convexity is reached for $r \approx 0$ which is of no practical interest. In the second, (16) does not allow any strictly positive solution. It can be shown that at some points, each criterion $\mathcal{G}_r$ is concave along the directions in $\mathrm{Ker}(\mathcal{A}^\dagger\mathcal{A})$. The family $(\mathcal{G}_r)_{r>0}$ does not possess any globally convex element.

In order to circumvent this difficulty, we construct a family of *doubly relaxed* criteria $(\mathcal{G}_{r,s})_{r>0,s>0}$ wich admit convex elements, by appending two quadratic auxiliary terms dependant on an auxiliary relaxation parameter s:

$$\begin{aligned} \mathcal{G}_{r,s}(\mathbf{f}) &= \|\mathcal{A}\mathbf{f} + \boldsymbol{h}\|^2 + \Psi_{r,s}(\mathbf{f}) + \Phi_{r,s}(\mathbf{f}), \quad \mathbf{f} \in \mathbb{R}^M, \\ \Psi_{r,s}(\mathbf{f}) &= -\Psi_r(\mathbf{f}) + s\sum_{(i\sim j)} [\frac{\alpha+r}{r}(\mathrm{f}_i - \mathrm{f}_j)^2 + (\mathrm{f}_i - \mathrm{f}_j - 1)^2 + (\mathrm{f}_i - \mathrm{f}_j + 1)^2], \\ \Phi_{r,s}(\mathbf{f}) &= -\Phi_r(\mathbf{f}) + s\sum_{i\in\mathcal{S}} [\frac{\beta+r}{r}\mathrm{f}_i^2 + (\mathrm{f}_i - 1)^2]. \end{aligned} \tag{17}$$

$\mathcal{G}_{r,s}$ is convex when $s > r^3/4$.

Let (r_0, s_0) be such that $\mathcal{G}_{r_0,s_0}$ is convex. The solution (8) is calculated in two stages. Fix first $r = r_0$ and relax s alone from s_0 to 0. A sequence of minima is tracked as described in Section 6. The remaining criteria are $\mathcal{G}_r$ (since $\mathcal{G}_r = \mathcal{G}_{r,0}$) and r is relaxed from r_0 to a sufficiently large value, using an exponential schedule. The algorithm should stop when the image pixels become sufficiently close to 0 or to 1.

The set of minima tracked by the GNC along a family of scalar relaxed criteria are again presented in Fig. 5.

[3] The right hand side of this equation when ρ_r are Lorentzians is strictly increasing for $r > 0$, since its first derivative is of the form $c_1 r^2 + c_2 r$.

Remark 3 (8 nearest-neighbour system) *It is not difficult to see how all the preceeding developpements apply to a binary* MRF *defined over an 8 nearest-neighbour system. Let* $(i \approx j)$ *designate diagonal neighbours and* $\tilde{\Psi}_{r,s}$ *the term of the relaxed criterion which concerns diagonal differences,* $\tilde{\Psi}_{r,s} = -\sum_{(i \approx j)} \psi_{r,s}(\mathrm{f}_i - \mathrm{f}_j)$. *The total relaxed energy is of the form* $\mathcal{G}_{r,s} = \mathcal{L} + \Psi_{r,s} + (1/\sqrt{2})\tilde{\Psi}_{r,s} + \Phi_{r,s}$.

9 Numerical results

We treated the plane geometry configuration (Section 2) with the following parameters:

- $\sigma_0 = 1.96\ 10^7\ S/m$ (dur-aluminium);
- $y_0 = -1.5\ 10^{-3}\ m$;
- $\nu = [10\quad 13.8\quad 19.1\quad 26.4\quad 36.5\quad 50.5\quad 69.7\quad 96.3\quad 133\quad 184]\ kHz$ and the relevant skin depths are $\delta = [1.1\ 1.0\ 0.8\ 0.7\ 0.6\ 0.5\ 0.4\ 0.4\ 0.3\ 0.3]\ mm$;
- 55 data per frequency, sampling step $0.25\ 10^{-3}\ m$;
- the measurement line is placed beneath the metal, as shown in the Fig. 1 (a).

Two images – representing a surface (Fig. 6 (a)) and an interior anomaly (Fig. 7 (a)) – have been reconstructed from simulated noisy data with 10 dB SNR.

A classical inversion using Tikhonov regularization ($\hat{\mathbf{f}} = \arg\min[\mathcal{L}(\mathbf{f}) + \lambda\|\mathbf{f}\|^2]$) has been applied. The reconstructed images have been calculated for $\lambda = 1.7$ and are shown in Figs. 6 (b) and 7 (b). The amplitude of the anomaly is largely underestimated and the image is extremely smooth.

The reconstructions using a binary MRF prior are presented in Figs. 6 (c) and 7 (c). A MRF based on an 8 nearest-neighbour system has been introduced with parameters ($\alpha = 0.05, \beta = 0.1$); these values have been chosen experimentally.

The improvement obtained with the Markovian prior is quite notable. It is due to both the strength and the adequacy of the incorporated prior.

10 Conclusions

A formulation of the eddy current tomography as a linear inverse problem has been used. It tries to reconstruct images showing the anomalies inside the investigated medium. Data represent the image quite poorly since the observation is accompanied by grave attenuation and low-pass filtering, and is corrupted by noise. At the same time, strong priors are available from the physics of the problem: the image is binary and is constituted by small black patches over a white background.

However, their incorporation in an estimation algorithm is not straightforward at all – the set of binary images is non convex and the observation operator is non local. Usual techniques drop the binary character and reconstruct continuous valued images.

We model the image as a binary MRF and define the optimal solution as the MAP estimate. The MAP criterion is discrete and has numerous local minima; calculation of the solution using available techniques is not feasible. We reformulated the discrete MAP criterion as the limit of a family of continuous criteria which we could optimize using a specially adapted GNC algorithm. The solution thus obtained is nearly optimal while the calculation cost is reasonable.

Eddy current tomography is a good example of the situation when very limited observation capacity is "compensated" by strong physical priors about the solution. The resultant estimator is seen to be highly performant.

11 Appendix

11.1 Proof of Proposition 1

The idea of the proof is that for each non binary image $\mathbf{f}$, it exists r_1 such that for $r > r_1$ there are binary images $\boldsymbol{f}$ for which $\mathcal{G}_r(\mathbf{f}) > \mathcal{G}_r(\boldsymbol{f})$.

Consider $\mathbf{f}$ such that $\{\mathrm{f}_{n_1}, \ldots, \mathrm{f}_{n_k}\} \notin \mathbb{B}^k$ while $\{\mathrm{f}_{n_{k+1}}, \ldots, \mathrm{f}_{n_M}\} \in \mathbb{B}^{M-k}$, $k \geq 1$ (at least one pixel is not in $\mathbb{B}$). This image contains l neighbour pairs $\{(m_i \sim n_i), i = 1, \ldots, l$ such that $|\mathrm{f}_{m_i} - \mathrm{f}_{n_i}| \notin \mathbb{B}\}$.

Take $\boldsymbol{f} \in \mathbb{B}^M$ such that $\{f_{n_{k+1}} = \mathrm{f}_{n_{k+1}}, \ldots, f_{n_M} = \mathrm{f}_{n_M}\} \in \mathbb{B}^{M-k}$, while $\{f_{n_1}, \ldots, f_{n_k}\} \in \mathbb{B}^k$ is arbitrary.

Let ε be the smallest distance of a $\mathrm{f}_{n_i}, 1 \leq i \leq k$ to $\mathbb{B}$: $\varepsilon = \min_{1 \leq i \leq k}\{|\mathrm{f}_{n_i}|, |\mathrm{f}_{n_i} - 1|\}$, while ϵ is the analogous quantity concerning differences between neighbours: $\epsilon = \min_{1 \leq i \leq l}\{|\mathrm{f}_{n_i} - \mathrm{f}_{m_i}|, ||\mathrm{f}_{n_i} - \mathrm{f}_{m_i}| - 1|\}$.

Using that ρ_r is increasing on $[0, \infty[$ and symmetric, we deduce

$$\Psi_r(\mathbf{f}) - \Psi_r(\boldsymbol{f}) \quad > \quad l[r - (1 + \alpha/r)\rho_r(\epsilon)],$$

$$\Phi_r(\mathbf{f}) - \Phi_r(\boldsymbol{f}) \;>\; k[r - (1 + \beta/r)\rho_r(\varepsilon)].$$

On the other hand, let $K = \|\mathcal{A}\mathbf{f} - \boldsymbol{h}\|^2 - \|\mathcal{A}\boldsymbol{f} - \boldsymbol{h}\|^2$; K can be negative but it is independent on r.

Consider the following functions

$$\begin{aligned} g_1(r) &= l[r - (1 + \alpha/r)\rho_r(\epsilon)] + K/2, \\ g_2(r) &= k[r - (1 + \beta/r)\rho_r(\varepsilon)] + K/2. \end{aligned}$$

We seek to show the existance of r_1 and r_2 which ensure that $g_1(r) > 0$ for $r > r_1$ and $g_2(r) > 0$ for $r > r_2$, since then we would have: for $r > \max\{r_1, r_2\}$, $\mathcal{G}_r(\mathbf{f}) > \mathcal{G}_r(\boldsymbol{f})$.

We can easily take $r > 2\alpha$ and $r > 2\beta$, so

$$\begin{aligned} g_1(r) &> l[r - \rho_r(\epsilon)3/2 + K/(2l)], \\ g_2(r) &> k[r - \rho_r(\varepsilon)3/2 + K/(2k)]. \end{aligned}$$

If $K > 0$ we can drop $K/(2l)$ and $K/(2k)$. Remarking that $\forall\varepsilon, \forall\kappa > 1$ it exists r_1 such that for $r > r_1$, $\kappa\rho_r(\varepsilon) < r$ (a direct consequence of the requirements addressed to ρ), we obtain the result.

Otherwise ($K < 0$) so $K/(2l) = -|K|/(2l)$ and $K/(2k) = -|K|/(2k)$) we can always take $r > 10|K|/(2l)$ and $r > 10|K|/(2k)$ which permits to deduce

$$\begin{aligned} g_1(r) &> l[r - \rho_r(\epsilon)5/3]10/9, \\ g_2(r) &> k[r - \rho_r(\varepsilon)5/3]10/9. \end{aligned}$$

Applying the remark done previously, we deduce the existence of the required r_1, r_2.

11.2 Proof of Proposition 2

Apart from an additive constant, the MAP criterion (12) can be written as

$$\mathcal{J}(\boldsymbol{f}) = \|\mathcal{A}\boldsymbol{f} - \boldsymbol{h}\|^2 + (\alpha + 1)\mathcal{N}_{(\neq)} + \mathcal{N}_{(=)} + (\beta + 1)\mathcal{N}_{(1)} + \mathcal{N}_{(0)},$$

which transforms, again up to an additive constant, to

$$\mathcal{J}(\boldsymbol{f}) = \|\mathcal{A}\boldsymbol{f} - \boldsymbol{h}\|^2 - \sum_{(i\sim j)} [(\alpha+1)d(f_i - f_j) + d(|f_i - f_j| - 1)] - \sum_{i\in\mathcal{S}} [d(f_i) + d(f_i - 1)]. \quad (18)$$

Suppose that $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ have respectively W_1 and W_2 white pixels and B_1 and B_2 black pixels; Q_1 and Q_2 differences between neighbours which are -1, R_1 and R_2 neighbours with like colours, and S_1 and S_2 neighbours whose difference is +1.

Then, according to (18)

$$\mathcal{J}(\boldsymbol{f}_1) - \mathcal{J}(\boldsymbol{f}_2) = \|\mathcal{A}\boldsymbol{f}_1 - \boldsymbol{h}\|^2 - \|\mathcal{A}\boldsymbol{f}_2 - \boldsymbol{h}\|^2 - \alpha(R_1 - R_2) - \beta(W_1 - W_2).$$

The evaluation of (13) leads to

$$\begin{aligned}
\mathcal{G}_r(\boldsymbol{f}_1) - \mathcal{G}_r(\boldsymbol{f}_2) &= \\
\|\mathcal{A}\boldsymbol{f}_1 - \boldsymbol{h}\|^2 &- \frac{\alpha + r}{r} r R_1 - rQ_1 - rS_1 - \frac{\beta + r}{r} rW_1 - rB_1 - \\
\|\mathcal{A}\boldsymbol{f}_2 - \boldsymbol{h}\|^2 &+ \frac{\alpha + r}{r} r R_2 + rQ_2 + rS_2 + \frac{\beta + r}{r} rW_2 + rB_2.
\end{aligned}$$

Using $Q_1 + R_1 + S_1 = Q_2 + R_2 + S_2 = 2M - I - J$ (the total number of neighbour pairs) and $B_1 + W_1 = B_2 + W_2 = M$ (the total number of pixels), we deduce

$$\begin{aligned}
\mathcal{G}_r(\boldsymbol{f}_1) - \mathcal{G}_r(\boldsymbol{f}_2) &= \|\mathcal{A}\boldsymbol{f}_1 - \boldsymbol{h}\|^2 - \|\mathcal{A}\boldsymbol{f}_2 - \boldsymbol{h}\|^2 \\
&- \alpha(R_1 - R_2) - \beta(W_1 - W_2) = \mathcal{J}(\boldsymbol{f}_1) - \mathcal{J}(\boldsymbol{f}_2).
\end{aligned}$$

11.3 Derivation of the initial convexity

The Hessian of $\mathcal{G}_r$ at $\mathbf{f}$ – a $M \times M$ matrix which depends on $\mathbf{f}$ – is of the form:

$$\mathcal{H}_r(\mathbf{f}) \doteq D^2\mathcal{F}(\mathbf{f}) = 2\mathcal{A}^\dagger\mathcal{A} + D^2\Psi_r(\mathbf{f}) + D^2\Phi_r(\mathbf{f}),$$

where D^2 means second order differential operator and $\dagger$ stands for Hermitian. Let us calculate $D^2\Psi_r(\mathbf{f})$ and $D^2\Phi_r(\mathbf{f})$.

$$[D^2\Psi_r(\mathbf{f})]_{ii} = \frac{\partial^2\Psi_r(\mathbf{f})}{\partial \mathrm{f}_i^2} = -\sum_{j\in\mathcal{V}_i}[\frac{\alpha + r}{r}\rho_r''(\mathrm{f}_i - \mathrm{f}_j) + \rho_r''(\mathrm{f}_i - \mathrm{f}_j - 1) + \rho_r''(\mathrm{f}_i - \mathrm{f}_j + 1)]$$

$$[D^2\Psi_r(\mathbf{f})]_{ij} = \frac{\partial^2\Psi_r(\mathbf{f})}{\partial \mathrm{f}_i \partial \mathrm{f}_j} = -[\frac{\alpha + r}{r}\rho_r''(\mathrm{f}_i - \mathrm{f}_j) + \rho_r''(\mathrm{f}_i - \mathrm{f}_j - 1) + \rho_r''(\mathrm{f}_i - \mathrm{f}_j + 1)], \quad j \in \mathcal{V}_i.$$

$$[D^2\Psi_r(\mathbf{f})]_{ij} = 0, \quad j \notin \mathcal{V}_i.$$

It is useful to remark that

$$\frac{\partial^2\Psi_r(\mathbf{f})}{\partial \mathrm{f}_i^2} = -\sum_{j\in\mathcal{V}_i}\frac{\partial^2\Psi_r(\mathbf{f})}{\partial \mathrm{f}_i \partial \mathrm{f}_j}.$$

The Hessian of $\Phi_r(\mathbf{f})$ is diagonal:

$$[D^2\Phi_r(\mathbf{f})]_{ii} = \frac{\partial^2\Phi_r(\mathbf{f})}{\partial \mathrm{f}_i^2} = -[\frac{\beta+r}{r}\rho''(\mathrm{f}_i) + \rho''(\mathrm{f}_i - 1)]$$

$\mathcal{F}$ is convex at $\boldsymbol{x}$ along a direction $\boldsymbol{v}$ iff $\boldsymbol{v}^T\mathcal{H}\boldsymbol{v} > 0$. $\mathcal{F}$ is convex at $\boldsymbol{x}$ iff the previous inequality holds $\forall \boldsymbol{v}$, which means that its Hessian is positive definite. Furthermore, $\mathcal{F}(\boldsymbol{x})$ is globally convex iff this inequality holds $\forall \boldsymbol{v}$ and $\forall \boldsymbol{x}$.

Recall that $\mathcal{F}$ is convex at $\mathbf{f}$ *iff* its Hessian is positive definite. We check the latter by applying the definition and testing the positivity of the following quantity:

$$\boldsymbol{v}^T\mathcal{H}_r(\mathbf{f})\boldsymbol{v} = 2\boldsymbol{v}^T\mathcal{A}^\dagger\mathcal{A}\boldsymbol{v} + \boldsymbol{v}^T D^2\Psi_r(\mathbf{f})\boldsymbol{v} + \boldsymbol{v}^T D^2\Phi_r(\mathbf{f})\boldsymbol{v}, \quad \boldsymbol{v} \in \mathbb{R}^M, \quad \|\boldsymbol{v}\| = 1.$$

Furthermore, we require that $\mathcal{F}$ be globally convex, *i.e.* that the set of Hessians $\forall \mathbf{f} \in \mathbb{R}^M$ be positive definite. Remarking that $[D^2\Psi_r(\mathbf{f})]_{ij} \leq \eta_r(\alpha+3r)/r$, similarly to [9] we deduce

$$\boldsymbol{v}^T D^2\Psi_r(\mathbf{f})\boldsymbol{v} = -\sum_{(i\sim j)}[D^2\Psi_r(\mathbf{f})]_{ij}(v_i - v_j)^2 \geq -\eta_r\frac{\alpha+3r}{r}\mathcal{D}^T\mathcal{D}.$$

Recall also that a non negative definite matrix $\mathcal{M}$ whose smallest and largest eigenvalues are respectively $\xi_{\min}$ and $\xi_{\max}$, verifies $\forall \boldsymbol{v} \in \mathbb{R}^M$, $\xi_{\min}\|vb\|^2 \leq \boldsymbol{v}^T\mathcal{M}\boldsymbol{v} \leq \xi_{\max}\|vb\|^2$.

Thsi permits to establish a lower bound for the set of the Hessians:

$$\boldsymbol{v}^T\mathcal{H}_r(\mathbf{f})\boldsymbol{v} \geq 2\mu_{\min} - \eta_r\frac{\alpha+3r}{r}\nu_{\max} - \eta_r\frac{\beta+2r}{r}, \quad \forall \boldsymbol{v} \in \mathbb{R}^M, \ \|\boldsymbol{v}\|^2 = 1.$$

The convexity condition (16) is straightforward.

For the case $\mu_{\min} = 0$ (or $\mu_{\min} \approx 0$), we build a family $(\mathcal{G}_{r,s})_{r>0,s>0}$ whose Hessians are of the form:

$$\mathcal{H}_{r,s}(\mathbf{f}) = \mathcal{H}_r(\mathbf{f}) + 2s\frac{\alpha+3r}{r}\mathcal{D}^T\mathcal{D} + 2s\frac{\beta+2r}{r}\mathcal{I}.$$

Convexity clearly occurs when $2s > \eta_r$. At the same time, the family of criteria should be as close as possible to $\mathcal{J}$; this requirement led us to the doubly relaxed family (17).

Acknowledgement

The authors thank Jérôme IDIER for his critical remark on Proposition 1.

References

[1] R. Zorgati, B. Duchene, D. Lesselier, and F. Pons, "Eddy current testing of anomalies in conductive materials, part I: Qualitative imaging via diffraction tomography techniques", *IEEE Transactions on Magnetics*, vol. M-27, no. 6, pp. 4416–4437, 1991.

[2] G. Demoment, "Image reconstruction and restoration : Overview of common estimation structure and problems", *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. ASSP-37, no. 12, pp. 2024–2036, 1989.

[3] R. Zorgati, B. Duchene, D. Lesselier, and F. Pons, "Eddy current testing of anomalies in conductive materials, part II: Quantitative imaging via deterministic and stochastic inversion techniques", *IEEE Transactions on Magnetics*, vol. M-28, no. 3, pp. 1850–1862, 1992.

[4] H. Sabbagh, "An eddy-current model for three-dimensional inversion", *IEEE Transactions on Magnetics*, vol. MAG-22, no. 4, pp. 282–292, 1986.

[5] M. Nikolova and A. Mohammad-Djafari, *Maximum entropy image reconstruction in eddy current tomography*, pp. 273–278, Maximum entropy and Bayesian methods. Kluwer Academic Publ., Paris, a. Mohammad-Djafari & G. Demoment edition, 1992.

[6] J. E. Besag, "On the statistical analysis of dirty pictures (with discussion)", *Journal of Royal Statistical Society B*, vol. 48, no. 3, pp. 259–302, 1986.

[7] A. Blake and A. Zisserman, *Visual reconstruction*, The MIT Press, Cambridge, 1987.

[8] M. Nikolova, "Markovian reconstruction in computed imaging and Fourier synthesis", in *Proceedings of IEEE ICIP*, 1994, vol. 2, pp. 690–694.

[9] M. Nikolova, J. Idier, and A. Mohammad-Djafari, "Inversion of ill-posed operators using a piecewise Gaussian MRF", Tech. Rep., GPI–LSS, 1994.

[10] G. Box and G. Tiao, *Bayesian inference in statistical analysis*, Addison-Wesley publishing, 1972.

[11] J. Besag, "Digital image processing : Towards Bayesian image analysis", *Journal of Applied Statistics*, vol. 16, no. 3, pp. 395–407, 1989.

[12] D. M. Grieg, B. T. Porteous, and A. H. Seheult, "Exact maximum a posteriori estimation for binary images", *Journal of Royal Statistical Society B*, vol. 51, no. 2, pp. 271–279, 1989.

[13] D. Geiger and F. Girosi, "Parallel and deterministic algorithms from MRF's: Surface reconstruction", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-13, no. 5, pp. 401–412, 1991.

[14] A. Blake, "Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-11, no. 1, pp. 2–12, 1989.

[15] Y. Leclerc, "Constructing simple stable description for image partitioning", *International Journal of Computer Vision*, vol. 3, pp. 73–102, 1989.

[16] I. Richards and Y. Heekyung, *Theory of Distributions: a non-technical introduction*, Cambridge University Press, Cambridge, 1990.

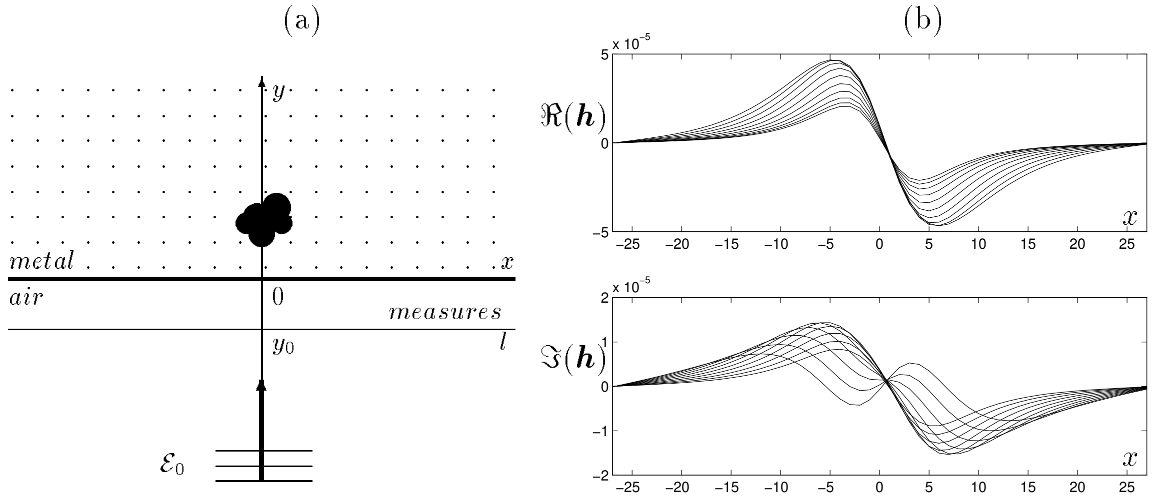

Figure 1: **(a)** Plane geometry configuration. The material is illuminated by a known time-harmonic electric field $\mathcal{E}_0$. Data are the potential variations induced into a sensing coil along a line (l) under the material at the height $-y_0$, and are directly related to the anomalous magnetic field $\mathcal{H}_a$. **(b)** Noise-free anomalous magnetical field data $\mathcal{H}_a$ for the anomaly shown in Fig. 6 (a): real part (up) and imaginary part (down). Each curve corresponds to a measurement performed at a given frequency.

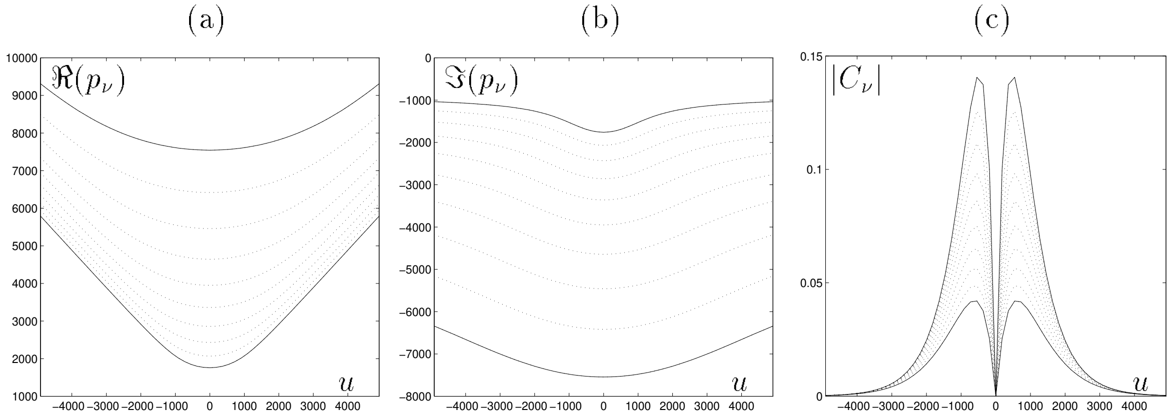

Figure 2: **(a)** $\Re(p_\nu)$ versus u for the set of frequencies given in Section 9. $\Re(p_\nu)$ is the attenuation factor which acts along the y-axis of the image; it clearly increases with $|u|$ and with ν. **(b)** $\Im(p_\nu)$ versus u for the same freqences. It gives the data point repartition in the FT2D domain, which is seen to be very irregular. **(c)** $|C_\nu(u)|$ for the same freqences; $C_\nu(u)$ originates from a Green function and produces an extremely severe pass-band filtering.

Figure 3: **(a)** Four nearest-neighbour system (the neighbours of i are the four pixels directly adjacent to it). **(b)** Eight nearest-neighbour system (the neighbours of i are the 8 pixels directly or diagonally adjacent to i). In both cases, only pairwise interactions are considered.

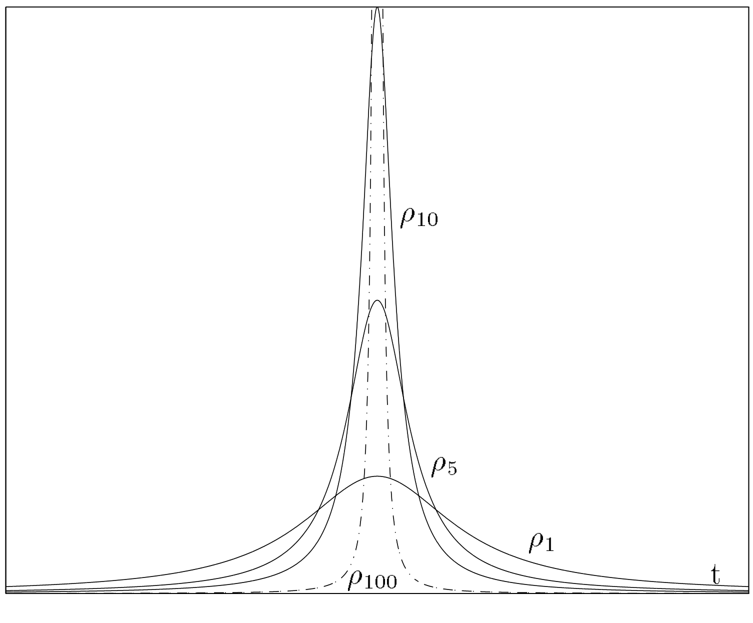

Figure 4: Lorentzian regularization family $(\rho_r)_{r>0}$, $\rho_r(\mathrm{t}) = r/(1 + r^2\mathrm{t}^2)$ for several values of r ($r = 1$, $r = 5$, $r = 10$, $r = 100$ (-.-)). This function converges towards a Dirac distribution as $r \to \infty$.

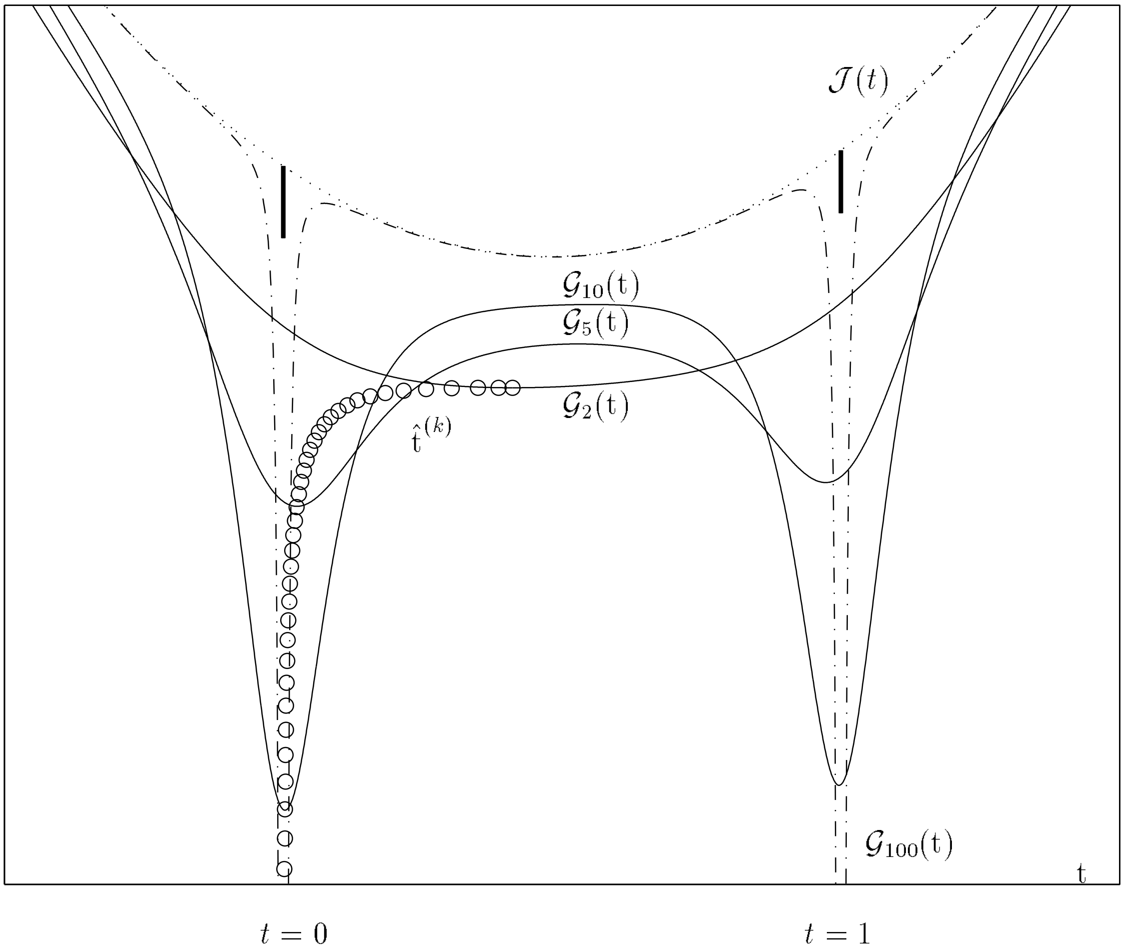

Figure 5: Relaxed criteria $(\mathcal{G}_r(\mathrm{t}))_{r>0}$ corresponding to a scalar discrete criterion $\mathcal{J}(t)$; the altitude difference between the criterion value at 0 and at 1 is kept independent of r. The minima tracked by the GNC are noted by 'o'. As far as the GNC progresses, the sequence of minima converges towards the global minimum of the discrete criterion $\mathcal{J}(t)$.

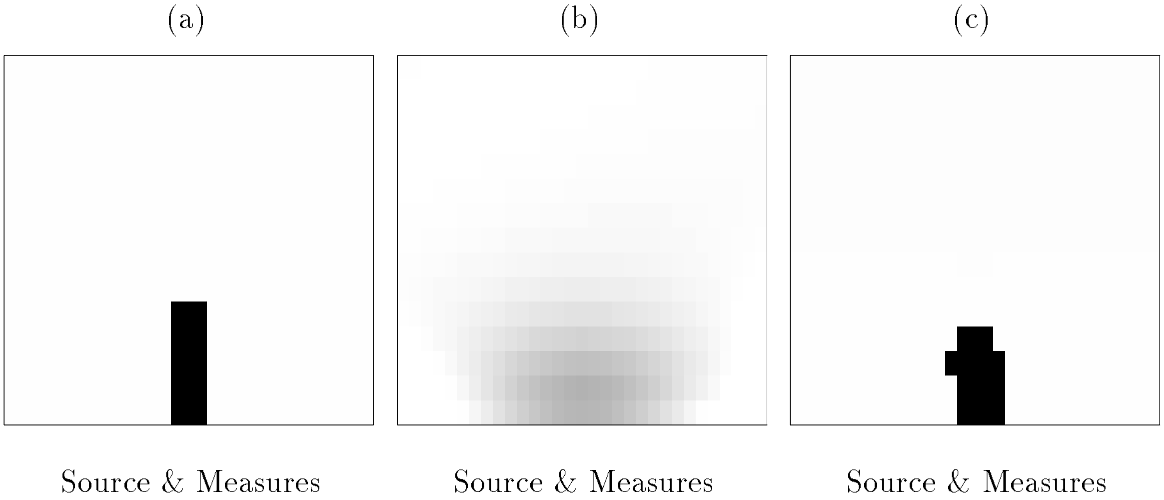

Figure 6: **(a)** Image I representing a surface anomaly. Reconstructions are performed from simulated data with 10 dB SNR. **(b)** Classical inversion using Tikhonov regularization ($\hat{\mathbf{f}} = \arg\min[\mathcal{L}(\mathbf{f}) + \lambda\|\mathbf{f}\|^2]$) for $\lambda = 1.7$. **(c)** The proposed regularization using a binary MRF prior based on an 8 nearest-neighbour system. The model parameters are $(\alpha = 0.05, \beta = 0.1)$.

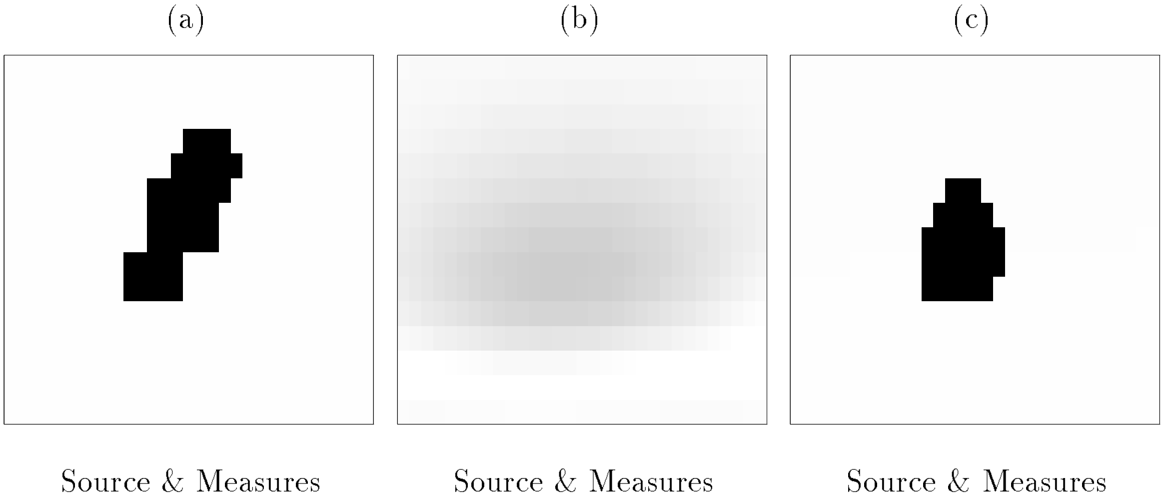

Figure 7: **(a)** Image II representing an interior anomaly. Reconstructions are performed from simulated data with 10 dB SNR. **(b)** Classical inversion using Tikhonov regularization for $\lambda = 1.7$. **(c)** The proposed Markovian regularization with $(\alpha = 0.05, \beta = 0.1)$.