A novel Bayesian strategy for the identification of spatially-varying parameters and model validation in inverse problems: an application to elastography

Koutsourelakis, Phadeon-Stelios

Published in:
International Journal for Numerical Methods in Engineering

Publication date:
2012

Journal Code:				Article ID				Dispatch: 08.02.12	CE: Sandy Cheromyl Vilan
N	M	E		4	2	6	1	No. of Pages: 20	ME:

INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng (2012)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nme.4261

A novel Bayesian strategy for the identification of spatially varying material properties and model validation: an application to static elastography

Phaedon-Stelios Koutsourelakis*,†

~~Center for Applied Mathematics, Cornell University, Ithaca, NY, USA~~

Q1

SUMMARY

The present paper proposes a novel Bayesian, a computational strategy in the context of model-based inverse problems in elastostatics. On one hand, we attempt to provide probabilistic estimates of the material properties and their spatial variability that account for the various sources of uncertainty. On the other hand, we attempt to address the question of model fidelity in relation to the experimental reality and particularly in the context of the material constitutive law adopted. This is especially important in biomedical settings when the inferred material properties will be used to make decisions/diagnoses. We propose an expanded parametrization that enables the quantification of model discrepancies in addition to the constitutive parameters. We propose scalable computational strategies for carrying out inference and learning tasks and demonstrate their effectiveness in numerical examples with noiseless and noisy synthetic data. Copyright © 2012 John Wiley & Sons, Ltd.

Received 28 February 2011; Revised 14 October 2011; Accepted 28 November 2011

KEY WORDS: Bayesian, elastography, model discrepancy, uncertainty, inverse problems

1. INTRODUCTION

The extensive use of large-scale computational models poses several challenges in parameter identification in the context of system identification or performing predictive simulations. Medical imaging represents such an application, which has attracted significant interest in recent years as the correct identification of material properties can reveal various pathologies [1, 2] as well as quantitatively assess the progress of various treatments.

Ultrasound elasticity imaging (elastography) has gained prominence in the context of performing medical diagnosis because of its accuracy and low cost. It is based on ultrasound tracking of precompression and post-compression images to obtain a map of position changes from which deformations can be inferred. The pioneering work of Ophir and coworkers [3] followed by several clinical studies [4–13] have demonstrated that the resulting strain images typically improve the diagnostic accuracy over ultrasound alone.

Broadly speaking, there are two approaches that are utilized for calculating the constitutive parameters. In the direct approach, the equations of equilibrium are interpreted as equations for the material parameters of interest, where the inferred strains and their derivatives appear as coefficients [14–16]. Although such an approach provides a computationally efficient strategy that does not require solution over the whole domain nor knowledge of the boundary conditions, it has certain drawbacks. More importantly perhaps, it does not use the raw data (i.e., noisy displacements) but

*Correspondence to: Phaedon-Stelios Koutsourelakis, ~~Center for Applied Mathematics, Cornell University, Ithaca, NY, USA~~

†E-mail: ~~pk285@cornell.edu~~

transformed versions, that is, strain fields which arise by applying sometimes ad hoc filtering and smoothing operators. ~~Although these might be plausible, in general~~ alter the informational content of the data and make difficult the quantification of the effect of observation noise. This is amplified when strain derivatives are computed, although not all such approaches require them for example, [17]. Furthermore, the smoothing employed can smear regions with sharply varying properties and hinder proper identification. Finally, it is nontrivial to determine appropriate boundary conditions in terms of the material parameters of interest.

The alternative to direct methods, that is, indirect, or iterative, as they are most commonly referred to, admit an inverse problem formulation where the discrepancy (in various norms, [18, 19]) between observed and model-predicted displacements is minimized with respect to the material fields of interest [20–30]. Although these approaches utilize directly the raw data, they generally imply an increased computational cost as the forward problem and potentially derivatives have to be solved/computed several times. This effort is amplified when stochastic/statistical formulations are employed as those arising from the Bayesian paradigm, whose cost is comparable with that of a deterministic global optimization technique [31].

Bayesian techniques are advocated in this paper because of their ability to quantify the effect of various sources of uncertainty to the hypotheses tested or the inferences made. One source of uncertainty is obviously the noise in the data, which constitutes probabilistic estimates more rational. This is particularly important when multiple hypotheses are consistent with the data or the level of confidence in the estimates produced needs to be quantified. Another source of uncertainty which is largely unaccounted for, is *model uncertainty* [32]. Namely, the parameters, whose values are estimated, are associated with a particular forward model about the behavior of the medium (in our case, a system of PDEs consisting of equilibrium and constitutive equations), but one cannot be certain about the validity of the model employed. In general, there will be deviations between the physical reality, where measurements are made and the idealized mathematical/computational description. Especially in the context of medical applications, it is crucially important to account for the model discrepancy or inadequacy in order to infer the right material properties and make accurate diagnoses.[‡] Nonintrusive Bayesian strategies, that is, those that basically make use of the forward model as a black-box, capture model discrepancy with regression models (e.g., Gaussian processes) which are not easily physically interpretable and cumbersome or impractical when they depend on a large number of input parameters [32, 33]. In contrast, our approach is intrusive. This enables us to overcome the aforementioned limitations and allows us to directly infer the stresses/pressure in the context of *elastostatics*.

The rest of the paper is organized as follows. Section 2 is devoted to the presentation of the novel Bayesian framework proposed in the context of elastostatics. Section 2.1 discusses computational aspects related to inference techniques for sampling from the posterior and learning schemes for estimating parameter values. Finally, Section 3 presents numerical results under static plane stress conditions using noiseless and noisy data with particular emphasis on quantifying model discrepancy.

2. PROPOSED METHODOLOGY

The presentation of the ideas in this paper is centered around solid mechanics, in particular elastostatics, but the framework introduced can be directly transitioned to other continua. We discuss first the formulation of the probabilistic model proposed and in subsection 2.1, the inference and learning tasks associated with this description. We adopt a physically inspired strategy that focuses on quantifying model discrepancies in the context of the *constitutive equation*. From a deterministic point of view, it resembles techniques such as constitutive relation error or error in the constitutive equation that have been developed for a posteriori error estimation and the solution of over specified inverse problems [30, 34–37]. We use the term constitutive equations to refer in general to relations

[‡]"I remember my friend Johnny von Neumann used to say, 'with four parameters, I can fit an elephant and with five, I can make him wiggle his trunk." *A meeting with Enrico Fermi*, Nature 427, 297; 2004.

DOI: 10.1002/nme

between conjugate thermodynamic variables, that is, stress and strain in solid mechanics or velocity and pressure in flow through permeable media or flux and temperature in heat diffusion.

In the formulations proposed, the constitutive relation supplements the observables and an augmented state space is used that includes all conjugate variables. As it is demonstrated in the sequence, the addition of these unknown parameters simplifies inference tasks and enables the quantification of model errors. The motivation for such an approach stems from the fact that inverse problems in the context of continuum models consist of

- a *conservation law* that arises from physical principles that are generally well-founded and trusted. In the case of single-phase flow through a porous medium, this amounts to the conservation of mass in solid mechanics to the conservation of linear momentum. In elastostatics in particular, this is written as

$$\nabla \cdot \tilde{\sigma}(\mathbf{x}) + \mathbf{b}(\mathbf{x}) = \mathbf{0}, \quad \mathbf{x} \in \Omega, \tag{1}$$

  where $\tilde{\sigma}(\mathbf{x})$ is the stress tensor, $\mathbf{b}$ the body force, and Ω the problem domain. Discretized versions of the aforementioned PDE are employed, which naturally introduce *discretization* error. This is generally well-studied in the context of linear problems and several a priori (and a posteriori) error estimates are available. In this work, we will ignore the discretization error in Equation (1), which corresponds to the verification stage and focus on the validation and calibration aspects.
- a *constitutive law* that is by-and-large phenomenological and therefore provides the primary source of *model uncertainty*. This is represented by the conductivity tensor in heat diffusion, the permeability tensor in flow through porous media or the elasticity tensor $\mathbf{D}$ in solid mechanics

$$\boldsymbol{\sigma}(\mathbf{x}) = \mathbf{D}(\mathbf{x})\boldsymbol{\epsilon}(\mathbf{x}), \quad \forall \mathbf{x} \in \Omega, \tag{2}$$

  where $\boldsymbol{\sigma}(\mathbf{x})$ is the vector of stress, and $\boldsymbol{\epsilon}(\mathbf{x})$ is the vector of strains.
- boundary/initial conditions or observables in general (which might include interior displacements). The available data are contaminated by noise and represent the main source of *observation errors*.

In the Bayesian setting advocated, the goal is to evaluate the *posterior density* for the material parameters (i.e., $\mathbf{D}(\mathbf{x})$) as well as quantitatively assess the validity of the aforementioned constitutive relation (Equation (2)).

The numerical implementation requires discretization of the aforementioned equations. For economy of notation, we consider the simplest perhaps discretization consisting of a finite element triangulation $\mathcal{T}$ of the problem domain Ω using n_{el} constant-strain/stress elements [§].

If e denotes the element number, the parameters in the formulation proposed are

- the stress vectors $\boldsymbol{\sigma}_e$, $e = 1, \ldots, n_{el}$ (3− dimensional under plane stress/strain conditions or 6-dimensional in general three-dimensional problems), which are jointly denoted by $\boldsymbol{\sigma} = \left[\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_{n_{el}}\right]^T$.
- the global displacement vector $\mathbf{u}$. If $\mathbf{u}_e$ denotes the nodal displacement vector of element e, then we represent by $\mathbf{L}_e$ the Boolean matrices that relate local and global displacement vectors, that is, $\mathbf{u}_e = \mathbf{L}_e\mathbf{u}$. We further denote by $\boldsymbol{\epsilon}_e$ the element strain vector, which is related to $\mathbf{u}_e$ as $\boldsymbol{\epsilon}_e = \mathbf{B}_e\mathbf{u}_e$ where $\mathbf{B}_e$ is the well-known strain-displacement matrix.
- the local constitutive matrices $\mathbf{D}_e$ that relate stress and strains over element e, that is, $\boldsymbol{\sigma}_e = \mathbf{D}_e\boldsymbol{\epsilon}_e$. These are assumed constant over each element, but they could be assigned different values at the nodes of the mesh or integration points of each element.

We will further assume that *noisy* displacement data (at interior or boundary points) are provided and will be denoted by $\mathbf{u}_Q \in \mathbb{R}^{n_Q}$. It is assumed that the observed nodal displacements are

[§]For more complex elements/discretizations, the ensuing formulations can be readily applied if instead we consider each integration point in the element

DOI: 10.1002/nme

given by $\mathbf{Q}\,\mathbf{u}$, where $\mathbf{Q}$ is an appropriate Boolean matrix (if all displacements are observed at all the nodes, then $\mathbf{Q}=\mathbf{I}$). Assuming Gaussian noise with variance ν^2, the *likelihood* of $\mathbf{u}_Q$ given $\mathbf{u}$ is normal and

$$p(\mathbf{u}_Q \mid \mathbf{u}) \propto \frac{1}{\nu^{n_Q}} \exp\left\{-\frac{1}{2\nu^2}(\mathbf{u}_Q-\mathbf{Q}\mathbf{u})^T(\mathbf{u}_Q-\mathbf{Q}\mathbf{u})\right\}. \tag{3}$$

The observation noise variance ν^2 can be known or unknown in which case we propose employing a conjugate $inverse-Gamma$ hyperprior with hyperparameters (α_ν,β_ν), that is,

$$p\left(\nu^2\right) \propto \left(\nu^{-2}\right)^{\alpha_\nu-1} e^{-\beta_\nu/\nu^2}. \tag{4}$$

Naturally, more complex models that can capture perhaps the spatial dependence of ν can be employed.

In general, nonessential boundary conditions might be available as well, that is, tractions might be prescribed at part of the boundary $\partial\Omega_N \subset \partial\Omega$, that is,

$$\mathbf{n}\cdot\tilde{\sigma}(\mathbf{x})\mid_{\Omega_N} = \boldsymbol{\tau}(\mathbf{x}), \quad \mathbf{x}\in\partial\Omega. \tag{5}$$

Noise in these observations could also be added, but we omit this to simplify the notation.

In the proposed framework, apart from the aforementioned observations, the *data* or likelihood consist also of model-related equations, that is, the conservation law (Equation (1)) which in the case of standard Bubnov–Galerkin finite element schemes is enforced weakly as

$$\int_\Omega \boldsymbol{\epsilon}(\mathbf{w})\cdot\boldsymbol{\sigma}\, d\mathbf{x} = \int_\Omega \mathbf{w}\cdot\mathbf{b}\, d\Omega + \int_{\partial\Omega_N} \mathbf{w}\cdot\boldsymbol{\tau}\, d\Gamma, \tag{6}$$

where $\boldsymbol{\epsilon}(\mathbf{w})$ denote the strains associated with the weighting functions $\mathbf{w}\in H_0^1(\Omega)$. It is noted that other discretization schemes such as finite volume or discontinuous Galerkin can also be used to enforce the conservation law with small alterations. In the triangulation $\mathcal{T}$ adopted for discretizing, the solution and the weighting functions $\mathbf{w}$, this reduces to

$$\hat{\mathbf{B}}^T\boldsymbol{\sigma} = \mathbf{f}, \tag{7}$$

where $\mathbf{f}$ is the force vector and

$$\hat{\mathbf{B}}^T = \sum_{e=1}^{n_{el}} (\mathbf{L}_e)^T \int_{\Omega_e} (\mathbf{B}_e)^T\, d\mathbf{x} = \sum_{e=1}^{n_{el}} V_e (\mathbf{L}_e)^T (\mathbf{B}_e)^T, \tag{8}$$

where V_e is the volume of element e.

The second model equation relates to the constitutive law which we propose enforcing for every element probabilistically. If the true constitutive law (which is unknown) is different from the one prescribed in Equation (2), then there will be a discrepancy/error $\mathbf{c}_e$ between the actual stresses $\boldsymbol{\sigma}_e$ and the model-predicted stresses $\mathbf{D}_e\boldsymbol{\epsilon}_e = \mathbf{D}_e\mathbf{B}_e\mathbf{u}_e$

$$\mathbf{c}_e = \boldsymbol{\sigma}_e - \mathbf{D}_e\mathbf{B}_e\mathbf{u}_e \tag{9}$$

Because $\mathbf{c}_e$ is unknown and in accordance with the Bayesian formulation advocated, we propose a hierarchical prior model where

$$\begin{aligned} &\mathbf{c}_e \mid \boldsymbol{\sigma}_e, \mathbf{u}_e, \Sigma_e \sim \mathcal{N}(\boldsymbol{\sigma}_e - \mathbf{D}_e\mathbf{B}_e\mathbf{u}_e, \Sigma_e) \\ &or \\ &p(c_e \mid \boldsymbol{\sigma}_e, \mathbf{u}_e, \Sigma_e) \propto \tfrac{1}{|\Sigma_e|^{1/2}} exp\left\{-\tfrac{1}{2}(\boldsymbol{\sigma}_e - \mathbf{D}_e\mathbf{B}_e\mathbf{u}_e)^T\Sigma_e^{-1}(\boldsymbol{\sigma}_e - \mathbf{D}_e\mathbf{B}_e\mathbf{u}_e)\right\}. \end{aligned} \tag{10}$$

In this work, we consider a special form of the covariances $\Sigma_e = \lambda_e^2\mathbf{I}$.

DOI: 10.1002/nme

The hyperparameters λ_e^2 express the variability of the constitutive error and their magnitude quantifies the *model discrepancy* over each element e. The inferred values λ_e^2 will reveal elements, where the model error is high and refinement/improvement is needed. Note for example that if the elastic properties vary within an element e, the corresponding λ_e^2 will be nonzero even if no noise exists in the data. When different discretization schemes are used, which might employ higher-order shape functions, distinct λ_e^2 for each integration point can be introduced. The normal prior for $\mathbf{c}_e$ (Equation (10)) is not the only option and was selected here for computational convenience because of its conjugacy with the other distributions as it will be seen in the sequel. It would certainly be worthwhile to investigate alternative prior models.

Because the hyperparameters λ_e^2 are unknown, prior models can be employed as well. In this study, we make use of a Gaussian Markov Random Field ([38, 39]) prior, which accounts for the fact that the magnitude of the model errors are expected to be spatially correlated. In particular, and because $\lambda_e^2 \geqslant 0$, we define the prior implicitly through the vector $\mathbf{Z} = \{z_e\}_{e=1}^{n_{el}}$, where $z_e = \log \lambda_e^2$

$$p(\mathbf{\Lambda}) \propto \exp\left\{-\frac{1}{2}\mathbf{Z}^T\mathbf{W}\mathbf{Z}\right\} \tag{11}$$

The precision matrix is given by $\mathbf{W} = \frac{1}{\sigma_z^2}\mathbf{H}$, where σ_z^2 is a scale parameter and $\mathbf{H} = [H_{e_1,e_2}]$

$$H_{e_1,e_2} = \begin{cases} \sum_{e_2=1}^{n_{el}} h_{e_1,e_2} & \text{if } e_1 = e_2 \\ -h_{e_1,e_2} & \text{otrherwise} \end{cases}, \tag{12}$$

where $h_{e_1,e_2} > 0$ is a measure of proximity between elements e_1 and e_2. In this work, this was defined with respect to the distance d_{e_1,e_2} between the element centroid as $h_{e_1,e_2} = e^{-d_{e_1,e_2}/d_0}$, where d_0 is a correlation-length parameter. The aforementioned model represents an intrinsic autoregressive prior [40, 41], which is an improper distribution (because $\mathbf{W}$ is semi-positive definite) that has been extensively used in spatial statistics. In particular, because $\sum_{e_2} W_{e_1,e_2} = 0\ \forall e_1$, it can be easily established that $p(\mathbf{\Lambda})$ penalizes the 'jumps' in $\mathbf{Z}$ at neighboring elements, that is,

$$p(\mathbf{\Lambda}) \propto \exp\left\{\sum_{e_1<e_2} W_{e_1,e_2}(z_{e_1} - z_{e_2})^2\right\}. \tag{13}$$

It is noted finally that values for the parameters (d_0, σ_z^2) are provided in the numerical results section.

The combination of Equations (3), (4), (7), (10), and (11) leads to the *posterior density* on the model parameters $\mathbf{\Theta} = \left(\nu^2, \boldsymbol{\sigma}, \{\mathbf{D}_e\}_{e=1}^{n_{el}}, \mathbf{u}, \mathbf{\Lambda} = \{\lambda_e^2\}_{e=1}^{n_{el}}\right)$. In addition to the observations $\mathbf{u}_Q$, the posterior on $\mathbf{\Theta}$ is *explicitly* conditioned on the model equations, that is, the discretized equation of equilibrium and the constitutive law [¶]

$$\begin{aligned} \pi(\mathbf{\Theta}) = p(\mathbf{\Theta} \mid \mathbf{u}_Q, \mathcal{M}) = p\left(\mathbf{u}_Q \mid \mathbf{u}, \nu^2\right) p(\nu^2) \\ 1_{\{\hat{\mathbf{B}}^T\boldsymbol{\sigma}=\mathbf{f}\}}(\mathbf{\Theta}) \\ \prod_{e=1}^{n_{el}} p\left(\mathbf{c}_e \mid \boldsymbol{\sigma}_e, \mathbf{u}_e, \lambda_e^2\right)\ p(\mathbf{\Lambda}) \\ p(\mathbf{u}). \end{aligned} \tag{14}$$

The indicator function $1_{\{\hat{\mathbf{B}}^T\boldsymbol{\sigma}=\mathbf{f}\}}(\mathbf{\Theta})$ implies that the support of the distribution includes only stress vectors that satisfy the (discretized) equilibrium equations in Equation (7).

[¶]This conditioning is denoted by $\mathcal{M}$ in 14

DOI: 10.1002/nme

A prior model could also be adopted with respect to the constitutive parameters $\mathbf{D}_e$. Such priors apart from improving the regularity of posterior are also physically plausible as one would expect the constitutive properties at neighboring locations to be correlated. Naturally, several such models have been proposed in the literature [31]. In this work, however, this was found unnecessary as the formulation proposed provides a natural correlation between $\mathbf{D}_e$ through the displacements $\mathbf{u}$ and stresses $\boldsymbol{\sigma}$, which are themselves spatially correlated because of the equilibrium and constitutive equations. This is evident in the conditional posteriors presented in the sequence. In contrast, a prior model was adopted for the displacement vector denoted by $p(\mathbf{u})$ in Equation (14). This can be useful when the observed displacements are sparse or restricted to a portion of the problem domain, but its primary utility in the examples contained in Section 3 was found to be the regularization of the displacement field in the presence of noise. In particular, we adopted an intrinsic autoregressive model as the one employed for $\boldsymbol{\Lambda}$ in Equation (11)

$$p(\mathbf{u}) \propto \exp\left\{-\frac{1}{2}\mathbf{u}^T\mathbf{V}\mathbf{u}\right\}, \tag{15}$$

where $\mathbf{V} = \frac{1}{\sigma_u^2}\mathbf{J}$. The matrix $\mathbf{J}$ defined exactly as $\mathbf{H}$ in Equation (12) with proximity between two arbitrary entries u_i , u_j defined with respect to the nodal distance.

It is worth emphasizing that the proposed model and associated posterior contain two sets of additional parameters as compared with traditional Bayesian formulations of the inverse problem: (1) the stress vector $\boldsymbol{\sigma}$; and (2) the model discrepancy parameters λ_e^2. The introduction of the former enables the quantification of the model discrepancy. Despite the augmented set of parameters, these additional vectors play the role of auxiliary variables that expedite the exploration of the posterior using Gibbs sampling [42] as discussed in subsection 2.1. One can readily obtain, *conditional* posterior densities for all the parameters appearing in $\boldsymbol{\Theta}$. In particular,

- For ν^2:

$$\nu^{-2} \mid \mathbf{u} \sim Gamma\left(\alpha_\nu + \frac{n_q}{2}, \beta_\nu + \frac{1}{2} \parallel \mathbf{u_Q} - \mathbf{Q}\,\mathbf{u} \parallel^2\right) \tag{16}$$

- For $\mathbf{u}$:

$$\mathbf{u} \mid \nu^2, \boldsymbol{\sigma}, \{\mathbf{D}_e, \lambda_e^2\}_{e=1}^{n_{el}} \sim \mathcal{N}(\mu_u, \mathbf{C}_u), \tag{17}$$

where

$$\begin{aligned} \mathbf{C}_u^{-1} &= \mathbf{C}^T\boldsymbol{\Lambda}^{-1}\mathbf{C} + \frac{1}{\nu^2}\mathbf{Q}^T\mathbf{Q} + \mathbf{V} \\ \mu_u &= \mathbf{C}_u\left(\mathbf{C}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{\sigma} + \frac{1}{\nu^2}\mathbf{Q}^T\mathbf{u}_Q\right). \end{aligned} \tag{18}$$

The aforementioned matrices $\mathbf{C}$ and $\boldsymbol{\Lambda}$ arise from the model discrepancy terms in Equation (14) as follows:

$$\begin{aligned} \mathbf{C} &= \begin{bmatrix} \mathbf{D}_1\mathbf{B}_1\mathbf{L}_1 \\ \mathbf{D}_2\mathbf{B}_2\mathbf{L}_2 \\ \dots \\ \mathbf{D}_{n_{el}}\mathbf{B}_{n_{el}}\mathbf{L}_{n_{el}} \end{bmatrix} \\ \boldsymbol{\Lambda} &= \begin{bmatrix} \lambda_1^2\mathbf{I} & \mathbf{0} & \dots\mathbf{0} & \\ \mathbf{0} & \lambda_2^2\mathbf{I} & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \lambda_{n_{el}}^2\mathbf{I} \end{bmatrix} \end{aligned} \tag{19}$$

DOI: 10.1002/nme

- For $\boldsymbol{\sigma}$:

$$\boldsymbol{\sigma} \mid \boldsymbol{u}, \{\mathbf{D}_e, \lambda_e^2\}_{e=1}^{n_{el}} \sim \mathcal{N}(\boldsymbol{\mu}_\sigma, \mathbf{C}_\sigma), \tag{20}$$

where

$$\begin{aligned} \mathbf{C}_\sigma &= \boldsymbol{\Lambda} + (\hat{\mathbf{B}}^T \boldsymbol{\Lambda})^T \left(\hat{\mathbf{B}}^T \boldsymbol{\Lambda} \hat{\mathbf{B}}\right)^{-1} (\hat{\mathbf{B}}^T \boldsymbol{\Lambda}) \\ \boldsymbol{\mu}_\sigma &= \mathbf{C}\mathbf{u} + (\hat{\mathbf{B}}^T \boldsymbol{\Lambda})^T \left(\hat{\mathbf{B}}^T \boldsymbol{\Lambda} \hat{\mathbf{B}}\right)^{-1} (\mathbf{f} - \hat{\mathbf{B}}^T \mathbf{C}\mathbf{u}) \end{aligned} \tag{21}$$

- For $\mathbf{D}_e$ assuming we are interested in the elastic modulus E_e such that $\mathbf{D}_e = E_e \hat{\mathbf{D}}_e$ (where $\hat{\mathbf{D}}_e$ is known):

$$E_e \mid \boldsymbol{\sigma}_e, \mathbf{u}_e, \lambda_e^2 \sim \mathcal{N}(\mu_E, \sigma_E^2), \tag{22}$$

where

$$\begin{aligned} \sigma_E^2 &= \frac{\lambda_e^2}{\| \hat{\mathbf{D}}_e \boldsymbol{\epsilon}_e \|^2} \\ \mu_E &= \frac{\boldsymbol{\epsilon}_e^T \hat{\mathbf{D}}_e^T \boldsymbol{\sigma}_e}{\| \hat{\mathbf{D}}_e \boldsymbol{\epsilon}_e \|^2} \end{aligned} \tag{23}$$

In the following, we propose a hybrid scheme based on the expectation–maximization algorithm [43] that provides maximum a posteriori point estimates for the model discrepancy parameters $\boldsymbol{\Lambda} = \{\lambda_e^2\}$ while fully sampling from the posterior of Equation (14) for the remaining parameters $\boldsymbol{\theta} = \left(\nu^2, \boldsymbol{\sigma}, \{\mathbf{D}_e\}_{e=1}^{n_{el}}, \mathbf{u}\right)$ (Figure 1). F1

2.1. *Inference and learning*

We advocate a scalable procedure for carrying out inference and learning with respect to the posterior $\pi(\boldsymbol{\Theta})$ (Equation (14)), which is a common practice in pertinent probabilistic models [44]. We compute point estimates for the vector $\boldsymbol{\Lambda} = \{\lambda_e^2\}$, which correspond to maxima $\boldsymbol{\Lambda}^*$ of the *log-posterior*.

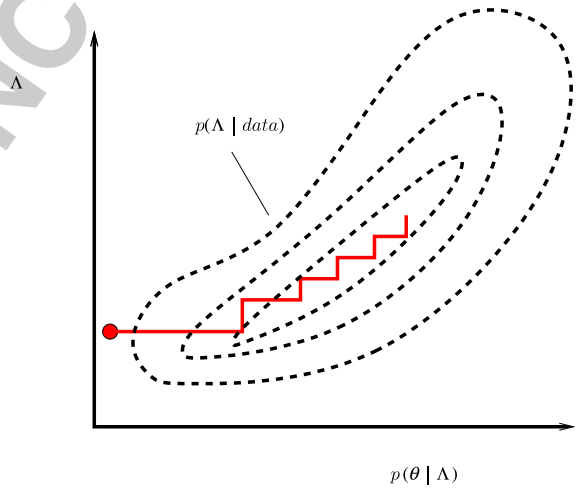


Figure 1. Schematic illustration of the expectation–maximization scheme.

Color Online, B&W in Print

DOI: 10.1002/nme

$$L(\boldsymbol{\Lambda}) = \log p(\boldsymbol{\Lambda} \mid \mathbf{u}_Q, \mathcal{M}) = \log \int \underbrace{p(\boldsymbol{\Lambda}, \boldsymbol{\theta} \mid \mathbf{u}_Q, \mathcal{M})}_{posteriorEquation\ (14)} \, d\boldsymbol{\theta}$$
$$= \log \int \pi(\boldsymbol{\Lambda}, \boldsymbol{\theta}) \, d\boldsymbol{\theta} \tag{24}$$

while the remaining parameters $\boldsymbol{\theta} = \left(\nu^2, \boldsymbol{\sigma}, \{\mathbf{D}_e\}_{e=1}^{n_{el}}, \mathbf{u}\right)$ are sampled from the full posterior $\pi(\boldsymbol{\theta}, \boldsymbol{\Lambda}*)$.

Maximization of $L(\boldsymbol{\Lambda})$ is more complex than a standard optimization task as it involves integration over the unobserved variables $\boldsymbol{\theta}$. We propose therefore, adopting an expectation–maximization framework (EM) which is an iterative, robust scheme that is guaranteed to increase the log-posterior at each iteration [43, 44]. It is based on constructing a series of increasing lower bounds of the log-posterior using auxiliary distributions $q(\boldsymbol{\theta})$

$$\begin{aligned} L(\boldsymbol{\theta}) = \quad & = \log \int \pi(\boldsymbol{\Lambda}, \boldsymbol{\theta}) \, d\boldsymbol{\theta} \\ & = \log \int q(\boldsymbol{\theta}) \frac{\pi(\boldsymbol{\Lambda}, \boldsymbol{\theta})}{q(\boldsymbol{\theta})} \, d\boldsymbol{\theta} \\ & \geqslant \int q(\boldsymbol{\theta}) \log \frac{\pi(\boldsymbol{\Lambda}, \boldsymbol{\theta})}{q(\boldsymbol{\theta})} \, d\boldsymbol{\theta} \quad \text{(Jensen's inequality)} \\ & = F(q, \boldsymbol{\theta}). \end{aligned} \tag{25}$$

It is obvious that this inequality becomes an equality when in place of the auxiliary distribution $q(\boldsymbol{\theta})$, the conditional posterior $\pi(\boldsymbol{\theta} \mid \boldsymbol{\Lambda}) = p(\boldsymbol{\theta} \mid \boldsymbol{\Lambda}, \mathbf{u}_Q, \mathcal{M})$ is selected. Given an estimate $\boldsymbol{\Lambda}^{(j)}$ at step j, this suggests iterating between an Expectation step (E-step), whereby we average with respect to $q^{(j)}(\boldsymbol{\theta}) = \pi(\boldsymbol{\theta} \mid \boldsymbol{\Lambda}^{(j)}, \mathbf{u}_Q, \mathcal{M})$ to evaluate the lower bound

$$\begin{aligned} \textbf{E-step: } F^{(j)}(q^{(j)}, \boldsymbol{\Lambda}) = & \int q^{(s)}(\boldsymbol{\theta}) \log \pi(\boldsymbol{\Lambda}, \boldsymbol{\theta}) \, d\boldsymbol{\theta} \\ & - \int q^{(j)}(\boldsymbol{\theta}) \log q^{(j)}(\boldsymbol{\theta}) \, d\boldsymbol{\theta} \end{aligned} \tag{26}$$

and a Maximization step (M-step) with respect to $F^{(j)}(q^{(j)}, \boldsymbol{\Lambda})$ (and in particular, the first part in Equation (26) because the second does not depend on $\boldsymbol{\Lambda}$)

$$\begin{aligned} \textbf{M-step: } \boldsymbol{\Lambda}^{(j+1)} & = arg \max_{\boldsymbol{\Lambda}} F^{(j)}(q^{(j)}, \boldsymbol{\Lambda}) \\ & = arg \max_{\boldsymbol{\Theta}} E_{q^{(j)}(\boldsymbol{\theta})} [\log \pi(\boldsymbol{\Lambda}, \boldsymbol{\theta})] \\ & = arg \max_{\boldsymbol{\Lambda}} Q(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda}). \end{aligned} \tag{27}$$

Given the expression of the (unormalized) posterior in Equation (14), the aforementioned objective function $Q(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda})$ becomes

$$\begin{aligned} Q(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda}) & = E_{q^{(j)}(\boldsymbol{\theta})} [\log \pi(\boldsymbol{\Lambda}, \boldsymbol{\theta})] \\ & = E_{q^{(j)}(\boldsymbol{\theta})} \left[\log \prod_{e=1}^{n_{el}} p(\mathbf{c}_e \mid \boldsymbol{\sigma}_e, \mathbf{u}_e, \lambda_e^2) \; p(\boldsymbol{\Lambda}) \right] \\ & = E_{q^{(j)}(\boldsymbol{\theta})} \left[\sum_{e=1}^{n_{el}} \log p\left(\mathbf{c}_e \mid \boldsymbol{\sigma}_e, \mathbf{u}_e, \lambda_e^2\right) \right] + E_{q^{(j)}(\boldsymbol{\theta})} [\log p(\boldsymbol{\Lambda})] \\ & = \sum_{e=1}^{n_{el}} E_{q^{(j)}(\boldsymbol{\theta})} \left[\log p\left(\mathbf{c}_e \mid \boldsymbol{\sigma}_e, \mathbf{u}_e, \lambda_e^2\right)\right] + \log p(\boldsymbol{\Lambda}). \end{aligned} \tag{28}$$

DOI: 10.1002/nme

Although the second term in the expression previously mentioned is essentially a penalty term arising from the prior on $\boldsymbol{\Lambda}$ (Equation (11)), the first term from Equation (10) leads to

$$E_{q^{(j)}(\boldsymbol{\theta})}\left[\log p\left(\mathbf{c}_e \mid \boldsymbol{\sigma}_e, \mathbf{u}_e, \lambda_e^2\right)\right] = -\frac{n_\sigma}{2}\log\lambda_e^2 - \frac{1}{\lambda_e^2} E_{q^{(j)}(\boldsymbol{\theta})}\left[\| \boldsymbol{\sigma}_e - \mathbf{D}_e\mathbf{B}_e\mathbf{u}_e \|^2\right]. \tag{29}$$

It is evident that the M-step requires computation of the sufficient statistics Φ_e

$$\Phi_e^{(j)} = E_{q^{(j)}(\boldsymbol{\theta})}\left[\| \boldsymbol{\sigma}_e - \mathbf{D}_e\mathbf{B}_e\mathbf{u}_e \|^2\right], \tag{30}$$

that is, the expected values (with respect to $q^{(j)}$) of the constitutive relation discrepancy in each of the elements $e = 1, \ldots, n_{el}$. Given the dependence amongst the components of $\boldsymbol{\Lambda}$ in the prior model, we propose an *incremental* version of the EM scheme ([45, 46]), where rather than maximizing $Q(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda})$ in the M-step, we set $\boldsymbol{\Lambda}^{(j+1)}$ such that

$$Q(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda}^{(j+1)}) \geqslant Q(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda}^{(j)}). \tag{31}$$

To that end, we propose maximizing $Q(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda})$ with respect to a single component of $\boldsymbol{\Lambda}$ (i.e., $\lambda_e^2,\ e = 1, \ldots, n_{el}$) at a time while keeping the rest fixed. At each step, all the components of $\boldsymbol{\Lambda}$ were scanned and details on the computations entailed are provided in the Appendix.

The critical task is that of inference, that is, the calculation of the expectations with respect to $q^{(j)}(\boldsymbol{\theta})$ in the E-step (Equation (26) or Equation (29)). As mentioned earlier, the optimal choice for $q^{(j)}(\boldsymbol{\theta})$ is the (conditional) posterior $\pi(\boldsymbol{\theta} \mid \boldsymbol{\Lambda}^{(j)})$, which is analytically intractable as it can readily be established from Equation (14). Although suboptimal variational approximations can be employed (e.g., [47–49]), in this work, we explore asymptotically exact approximations based on MCMC sampling from the posterior [50]. If $\{\boldsymbol{\theta}^{(i,j)}\}_{i=1}^N$ denote N samples from such a Markov chain with the (conditional) posterior $q^{(j)}(\boldsymbol{\theta}) = \pi(\boldsymbol{\theta} \mid \boldsymbol{\Lambda}^{(j)})$ at iteration j as the target, then the E-step in Equation (26) can be substituted by

$$Q(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda}) = \int q^{(j)}(\boldsymbol{\theta}) \log\pi(\boldsymbol{\Lambda}, \boldsymbol{\theta})\, d\boldsymbol{\theta} \approx \hat{Q}(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda}) = \frac{1}{N}\sum_{i=1}^{N} \log\pi(\boldsymbol{\Lambda}, \boldsymbol{\theta}^{(i,j)}). \tag{32}$$

The unavoidable noise introduced in these estimates by MCMC might necessitate an exuberant number of samples N to obtain a robust algorithm particularly close to the maximum of $L(\boldsymbol{\Lambda})$ (Equation (24)). For that purpose, we propose employing a stochastic approximation variant of the Robbins & Monro scheme [51, 52]. Rather than increasing the simulation size N in order to reduce the variance, we compute a weighted average at the current and previous iterations. By employing a decreasing sequence of weights, information from the earlier iterations gets discarded gradually and more emphasis is placed on the recent iterations. As it is shown in [53], this method converges with a fixed sample size N (even when $N = 1$). Convergence results that take into account the dependence of the Markov chains at each EM-step have been obtained by constraining the sequence of $\boldsymbol{\Lambda}^{(j)}$ to some compact set $\mathcal{C}$ by means of a reprojection onto $\mathcal{C}$ [54]. Even though this does not pose much problems in computational practice, weakened conditions have been established in [55, 56].

In particular, rather than using $\hat{Q}(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda})$ (which according to Equation (32) approximates $Q(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda})$) in the M-step (Equation (27)), we use

$$\tilde{Q}(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda}) = (1 - \gamma_j)\,\tilde{Q}(\boldsymbol{\Lambda}^{(j-1)}, \boldsymbol{\Lambda}) + \gamma_j\,\hat{Q}(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda}), \tag{33}$$

DOI: 10.1002/nme

where the sequence of weights $\{\gamma_j\}$ is such that $\sum_{j=1}^{\infty}\gamma_j = \infty$ and $\sum_{j=1}^{\infty}\gamma_j^2 < \infty$ [‖]. As it can be seen from Equations (28)–(30) in order to estimate the weighted average in Equation (33), it suffices to keep track of the weighted averages $\tilde{\Phi}_e^{(j)}$ of the sufficient statistics $\Phi_e^{(j)}$ (Equation (30))

$$\tilde{\Phi}_e^{(j)} = (1-\gamma_j)\tilde{\Phi}_e^{(j-1)} + \gamma_j \Phi_e^{(j)}. \tag{34}$$

The MCMC steps can be carried out using Gibbs sampling with respect to each of the components of $\boldsymbol{\theta}$, that is, ν^2 $\mathbf{u}$, $\boldsymbol{\sigma}$, and $\{\mathbf{D}_e\}_{e=1}^{n_{el}}$, which require the conditional distributions enumerated in the previous subsection (i.e., Equations (16), (17), (20) and (22)). It is worth pointing out that the $n \times n$ system of linear equations *does not need to be solved* (which has a cost of $O(n^3)$ operations) at any stage as in traditional inverse problems. If J is the total number of EM iterations and N is the number of MCMC steps at each iteration, then sampling from the aforementioned conditionals implies

- the inversion and Cholesky factorization of $\mathbf{C}_u$ in order to generate samples of $\mathbf{u}$. This must be repeated at *every* MCMC step because $\{\mathbf{D}_e\}$ are updated. The cost of this operation is $O(J\ N\ n^3)$.
- the Cholesky factorization of $\mathbf{C}_\sigma$ in order to generate samples of $\boldsymbol{\sigma}$. This must be repeated at *every* EM iteration and *not at every* MCMC step because $\mathbf{C}_\sigma$ solely depends on $\boldsymbol{\Lambda}$. The cost of this operation is $O(J\ (n_{el}n_\sigma)^3)$, where n_σ is the number of stress components ($n_\sigma = 6$ in three dimensions, $n_\sigma = 3$ in plane stress/strain, etc.).

In order to reduce the cost associated with these operations, one can employ block-Gibbs updates with respect to each of the components of $\mathbf{u}$ (or blocks of $\mathbf{u}$) rather than updating the whole vector at once. As it is demonstrated in the sequence, the cost of such a scheme is $O(J\ N\ n(n_{el}n_\sigma))$. The mixing is obviously slower than the full updates and as a consequence, the variance in the MCMC estimates is larger. In general, therefore, more EM iterations (assuming the same number of samples N are used at each iteration) are needed to converge. Nevertheless, the linear scaling with J constitutes such a scheme more efficient. Similar block-Gibbs updates can be carried out for $\boldsymbol{\sigma}$ reducing the cost associated with this task to $J\ N\ n(n_{el}n_\sigma))$. The conditional posteriors for performing block-Gibbs moves are described in the sequence.

Let $\begin{bmatrix} u_i \\ \mathbf{u}_{-i} \end{bmatrix}$ be a partitioning of $\mathbf{u}$ with respect to component i [**]. Let also $\mathbf{Q} = [\mathbf{Q}_i \mid \mathbf{Q}_{-i}]$, $\mathbf{C} = [\mathbf{C}_i \mid \mathbf{C}_{-i}]$ the corresponding partitioning of the matrices appearing in Equations (3) and (19). Then, the conditional posterior of u_i from Equation (14) is

$$u_i \mid \mathbf{u}_{-i}, \nu^2, \boldsymbol{\sigma}, \{\mathbf{D}_e, \lambda_e^2\}_{e=1}^{n_{el}} \sim \mathcal{N}(\mu_{u_i}, \sigma_{u_i}^2), \tag{35}$$

where

$$\sigma_{u_i}^{-2} = \mathbf{C}_i^T \boldsymbol{\Lambda}^{-1}\mathbf{C}_i + \frac{1}{\nu^2}\mathbf{Q}_i^T\mathbf{Q}_i \tag{36}$$

$$\mu_{u_i} = \sigma_{u_i}^2 \left(\mathbf{C}_i^T \boldsymbol{\Lambda}^{-1}(\boldsymbol{\sigma} - \mathbf{C}_{-i}\mathbf{u}_{-i}) + \frac{1}{\nu^2}\mathbf{Q}_i^T(\mathbf{u}_Q - \mathbf{Q}_{-i}\mathbf{u}_{-i}) \right). \tag{37}$$

It is noted that the leading order of computational operations for updating successively all components of $\mathbf{u}$ as previously mentioned is $O(n(n_{el}n_\sigma))$. This is approximately one order less than the

[‖] A family of such sequences that was used in this work is $\gamma_j = \frac{1}{j^p}$ with $1/2 < p \leqslant 1$. The value of $p = 0.51$ was employed

[**] An identical procedure can be followed when u_i corresponds to a block of $\mathbf{u}$

DOI: 10.1002/nme

$O(n^3)$ cost associated with the full update (Equation (17)), considering that the dimension of the stress vector $n_{el}n_\sigma$ is comparable with n.

3. NUMERICAL EXAMPLES

In this section, we report results on the accuracy and performance of the algorithm on two-dimensional elastography problems on synthetic data obtained for the configuration depicted in Figure 2 [14, 16], where the boundary displacements normal to the walls are prescribed. We intend to provide a clinical validation of the approach in a future study. F2

We assume an isotropic elastic material with Poisson's ratio $\nu = 0.5$ (incompressible) and employ the selective reduced integration quadrilateral elements for the forward problem [57, 58].

We examine two distributions for the elastic modulus occurring in elliptic and circular inclusions. In the first problem (Figure 3), the emphasis is on demonstrating the capabilities of the proposed method in identifying the ground truth as well as providing probabilistic confidence metrics particularly in the presence of noise. In the second case (Figure 8), the emphasis is on detecting and quantifying model discrepancies in the sense described in Section 2. It is noted that in all cases F3

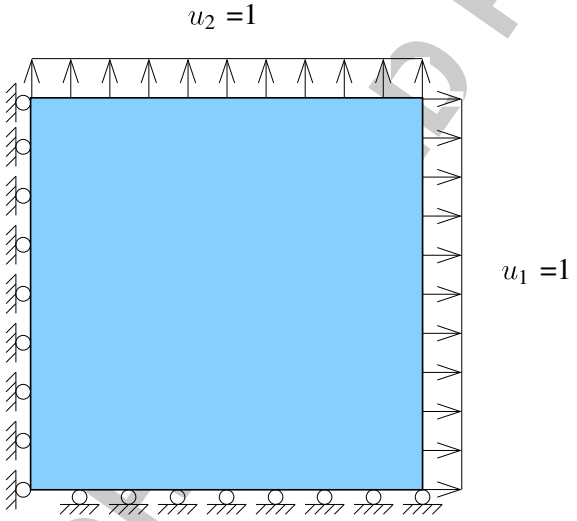


Figure 2. Problem configuration used in both examples 1 and 2 [16].

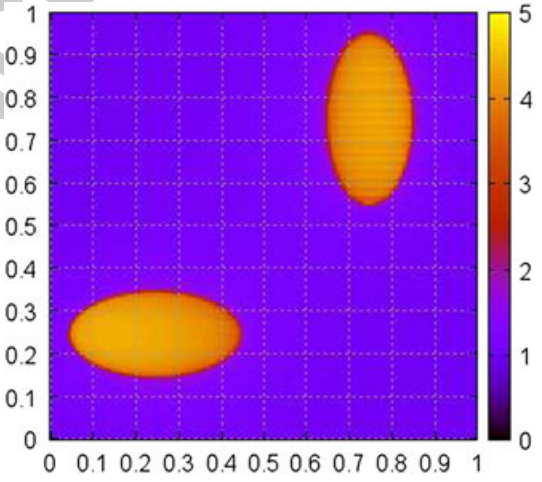


Figure 3. Example 1 - Elastic modulus E spatial distribution: In the inclusions, $E = 5$, whereas in the rest of the domain, $E = 1$.

Color Online, B&W in Print

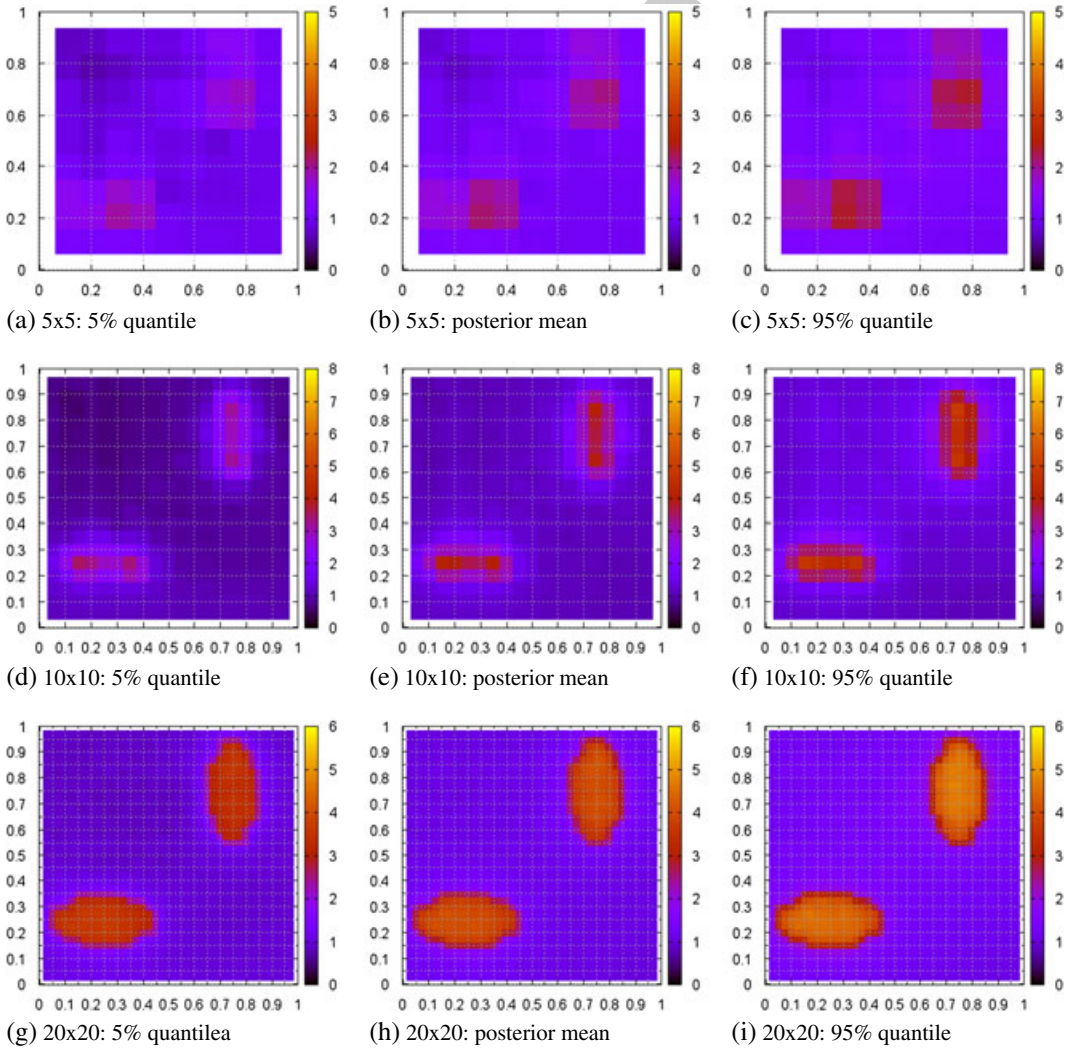
DOI: 10.1002/nme

apart from the identification of material properties, a direct output of the computations is the stress distribution. It is finally noted that in order to generate the displacement data, the forward problem was solved with a randomly generated mesh consisting of $10,000$ elements.

The following values were used for the parameters appearing in prior models described previously

- $\sigma_z^2 = 100$ (Equation (11)), which corresponds to a diffuse prior and $\sigma_u^2 = 1$ in Equation (15). The latter was selected based on the magnitude of the prescribed boundary displacements in Figure 2.
- $d_0 = 0.1$ for the correlation-length parameter appearing in the $\mathbf{H}$ (Equation (12)) and $\mathbf{V}$ (Equation (15)). Numerical evidence suggested that the effect of this parameter was minimal when varied in the range $[0.01, 0.5]$ given that the characteristic dimension of the problem domain is 1.
- an uninformative Jeffry's prior was adopted for the observation noise variance ν^2 (Equation (16)) with $\alpha_\nu = 2$ and $\beta_\nu = 0$.

With regards to the EM scheme, at each iteration $N = 10$, MCMC updates of all model parameters were performed and iterations were terminated when the *relative increase* in



(a) 5x5: 5% quantile (b) 5x5: posterior mean (c) 5x5: 95% quantile

(d) 10x10: 5% quantile (e) 10x10: posterior mean (f) 10x10: 95% quantile

(g) 20x20: 5% quantilea (h) 20x20: posterior mean (i) 20x20: 95% quantile

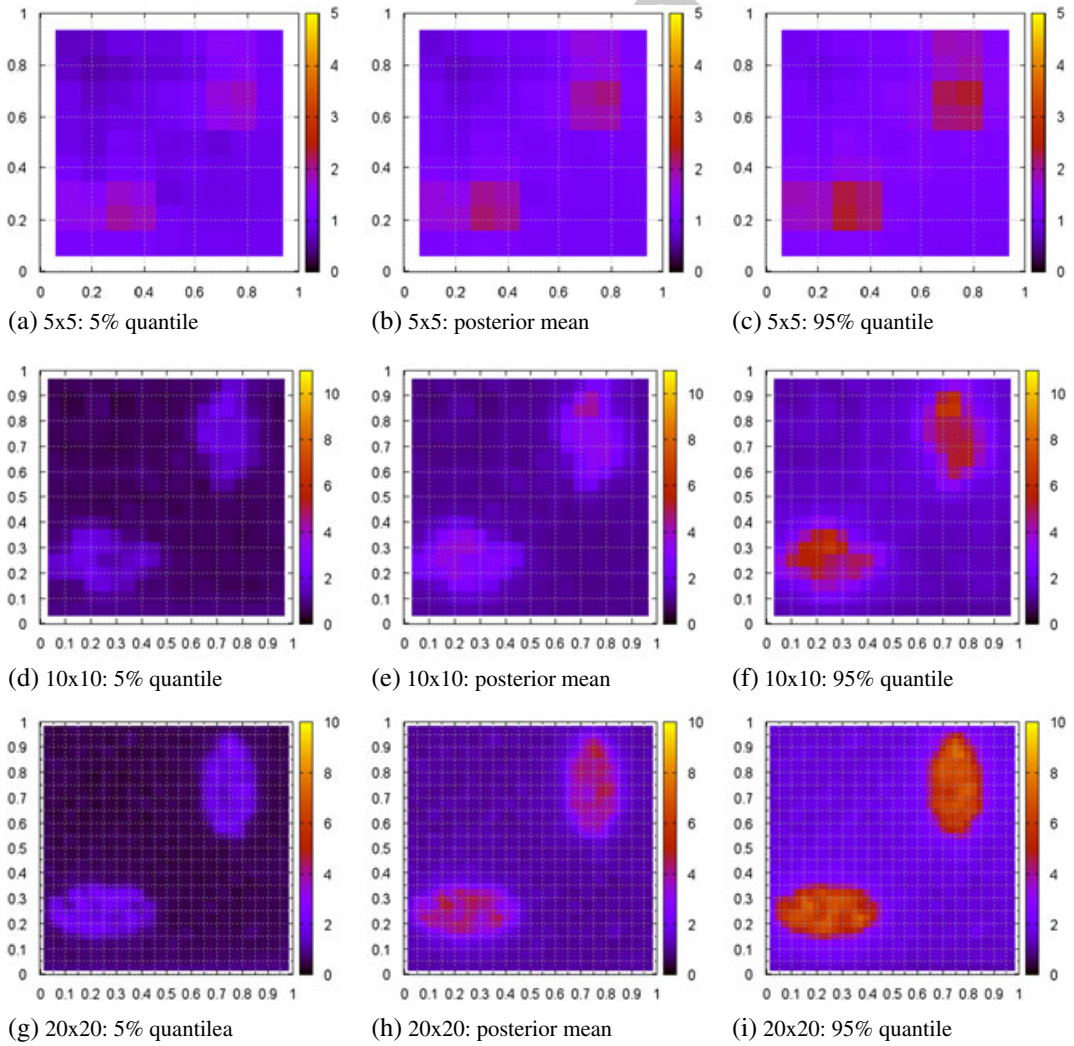
Figure 4. Example 1: Posterior statistics of the elastic modulus distribution for noiseless data.

DOI: 10.1002/nme

the objective $Q(\boldsymbol{\Lambda}^{(j)}, \boldsymbol{\Lambda})$ in Equation (28) was less than or equal to $\epsilon = 0.001$, that is, $\frac{|Q(\boldsymbol{\Lambda}^{(j)},\boldsymbol{\Lambda}) - Q(\boldsymbol{\Lambda}^{(j-1)},\boldsymbol{\Lambda})|}{|Q(\boldsymbol{\Lambda}^{(j-1)},\boldsymbol{\Lambda})|} < \epsilon$.

3.1. Example 1

The first scenario involves two elliptical inclusions centered at $(0.25, 0.25)$ and $(0.75, 0.75)$ with principal axes 0.1 and 0.2. with a contrast ratio 5 : 1 in the elastic modulus (Figure 3). A useful outcome of the numerical investigations was the fact that the overall inference and learning process can be greatly accelerated by operating on a sequence of discretizations with increased refinement. In particular, initially, a coarse mesh is adopted with few nodes and elements, where the proposed EM scheme is applied. The parameter values learned (i.e., $\boldsymbol{\Lambda}$) are used as the initial values for a refined mesh. The MCMC chains with respect to the other model parameters at the new mesh are initiated from samples drawn at the coarser mesh. It was found that this led to a reduction of the number of EM iterations needed to achieve convergence and significant acceleration because the order of operations at coarse meshes is smaller. For that purpose, we report in this problem the results obtained at three different resolutions employing a regular mesh with 5×5 , 10×10 and 20×20 elements. A potentially important implication involves the possibility of *adaptive refinement*, where the mesh



(a) 5x5: 5% quantile (b) 5x5: posterior mean (c) 5x5: 95% quantile

(d) 10x10: 5% quantile (e) 10x10: posterior mean (f) 10x10: 95% quantile

(g) 20x20: 5% quantilea (h) 20x20: posterior mean (i) 20x20: 95% quantile

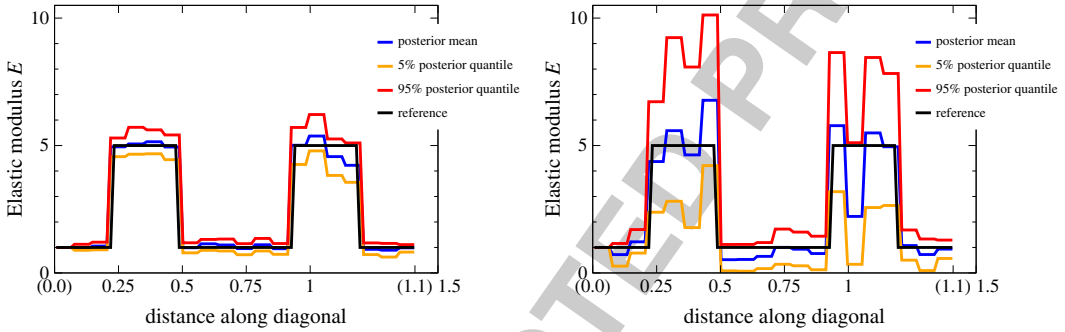
Figure 5. Example 1: Posterior statistics of the elastic modulus distribution for noisy data with SNR=40 dB.

DOI: 10.1002/nme

can be refined at selected regions of the problem domain where further information is needed as determined by the inferences produced at coarser resolutions [28].

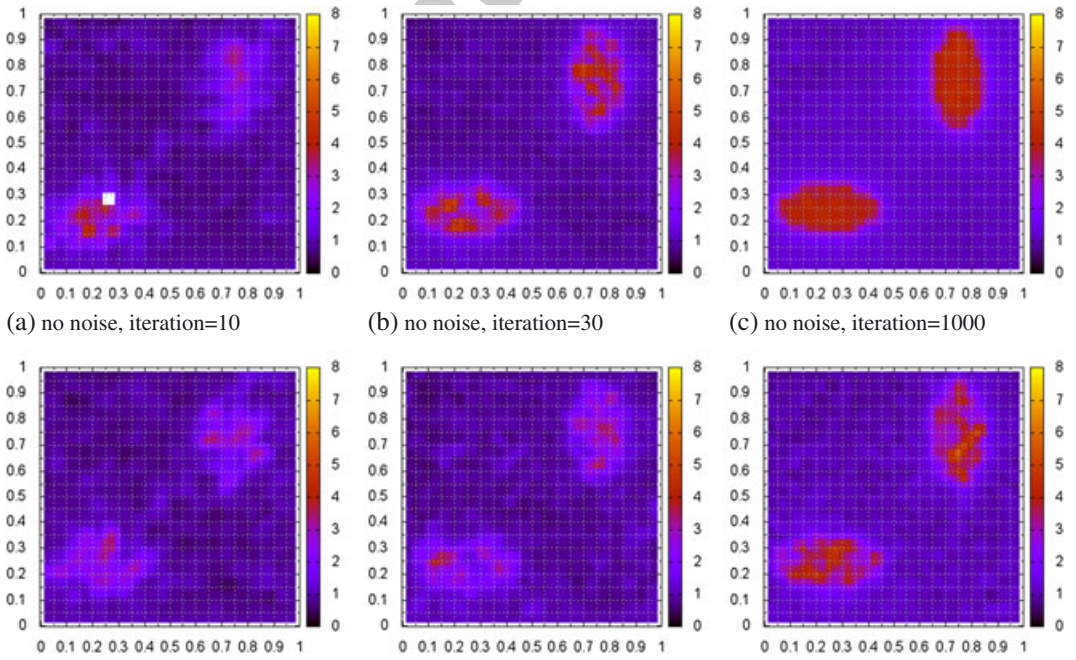
Figure 4 depicts the posterior mean as well as the posterior quantiles at 5% and 95% for the elastic modulus at these three resolutions and in the absence of noise in the data. It is readily observed that the proposed scheme can identify the ground truth as well as provide posterior credible intervals on the inferences made. These are more clearly depicted in Figure 6(a), which presents the results along the diagonal from $(0,0)$ to $(1,1)$.

We also investigated the performance of the algorithm in the presence of zero mean, Gaussian noise, and in particular with a signal-to-noise-ratio (SNR) $SNR = 40dB$, which is typical for ultrasound systems [21,23]. The results are shown in Figure 5 in terms of posterior mean and posterior quantiles. As it can also be seen in Figure 6(b), the algorithm is able to quantify the uncertainty introduced by the presence of noise and posterior bounds provided enclose the ground truth. Finally, Figure 7 depicts randomly selected samples drawn at various iterations of the EM scheme (for the



(a) no noise (b) SNR=40

Figure 6. Example 1: Posterior statistics of the elastic modulus distribution along the diagonal from $(0,0)$ to $(1,1)$.



(a) no noise, iteration=10 (b) no noise, iteration=30 (c) no noise, iteration=1000

(d) SNR=40, iteration=10 (e) SNR=40, iteration=30 (f) SNR=40, iteration=1000

Figure 7. Samples of the elastic modulus distributions obtained from the posterior at various iteration numbers of the proposed expectation–maximization framework scheme.

DOI: 10.1002/nme

finest resolution 20×20) that demonstrate the evolution of the learning algorithm proposed.

3.2. Example 2

The primary goal in the second example is to demonstrate the capability of quantifying model discrepancy in the constitutive equation. In particular, we consider the synthetic data generated by the material distribution in Figure 8. The circular inclusion centered at $(0.5, 0.5)$ with radius 0.2 is assumed to have an elastic modulus that is 5 times larger than the rest of the domain. We further assumed a square region on the top left corner $[0, 0.2]\times[0.8, 1]$, where rather than an *isotropic*, elastic material, we employed an anisotropic constitutive matrix $\mathbf{D} = \begin{bmatrix} 10 & -5 & -5 \\ -5, & 20. & -5 \\ -5 & -5 & 100 \end{bmatrix}$. Although this is a valid constitutive model (i.e., $\mathbf{D}$ is positive definite), it is obviously inconsistent with the isotropic assumption made in the model used to identify material properties. Although other inversion schemes might be able to find an elastic modulus corresponding to an isotropic material that fits adequately the observed displacements, they would be unable to identify that the model employed

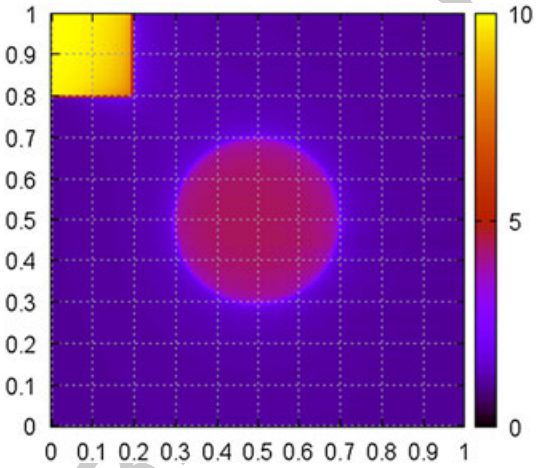
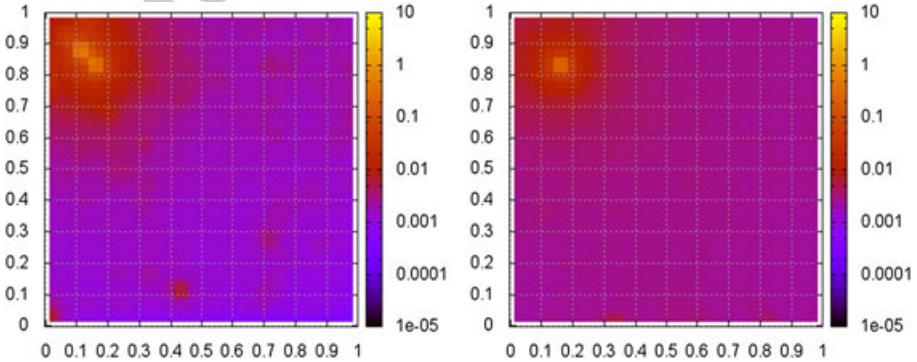


Figure 8. Example 2: Elastic modulus E spatial distribution: In the circular inclusion, $E = 5$, in the subdomain $[0, 0.2] \times [0.8, 1]$, we employed a constitutive matrix $\mathbf{D} = \begin{bmatrix} 10 & -5 & -5 \\ -5, & 10. & -5 \\ -5 & -5 & 100 \end{bmatrix}$,whereas in the rest of the domain, $E = 1$.



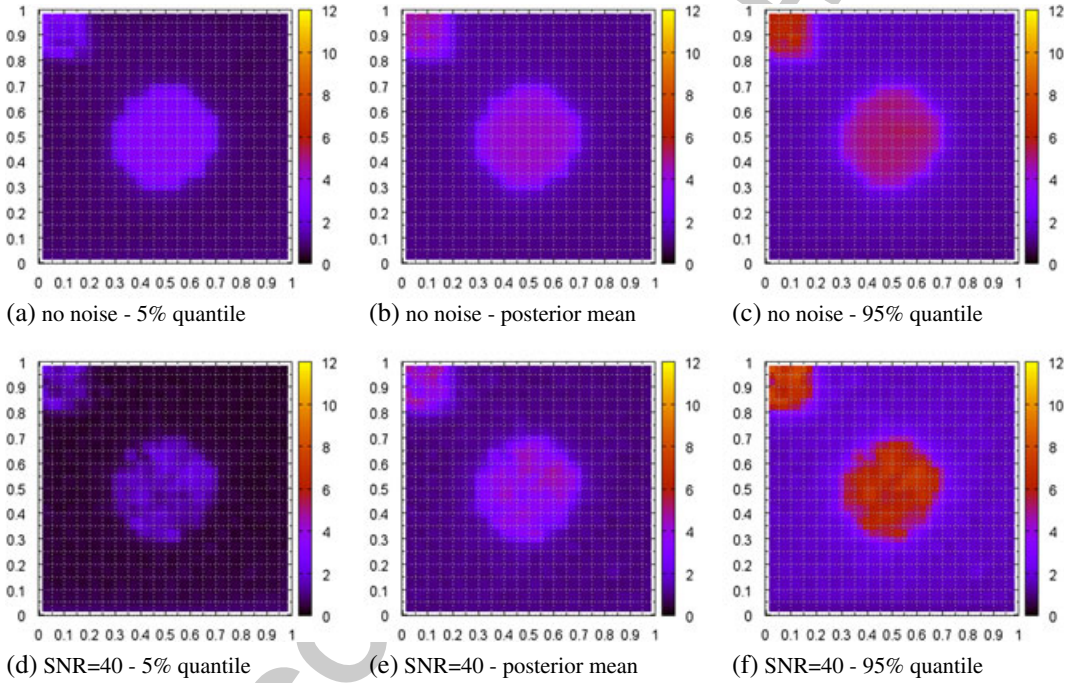
(a) no noise (b) SNR=40

Figure 9. Example 2: Model discrepancies/errors $\{\lambda_e^2\}_{e=1}^{n_{el}}$ for (a) no noise, and (b) SNR=40 dB (in log-scale).

DOI: 10.1002/nme

is inadequate. As a result, erroneous conclusions would be drawn about the state of the material at this portion of the problem domain.

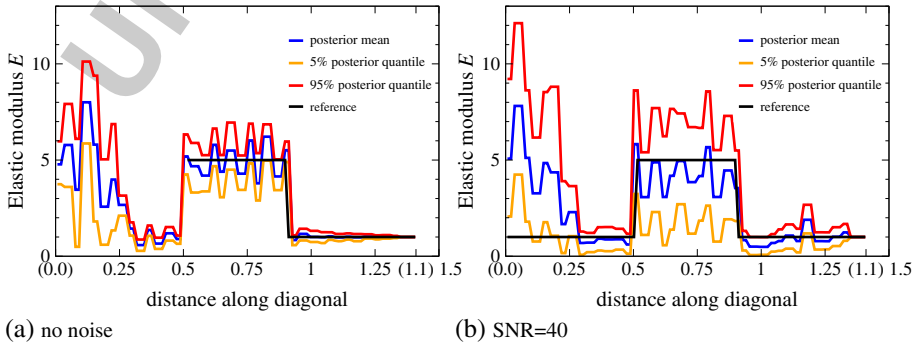
Figure 9 depicts the learned values of the parameters $\mathbf{\Lambda} = \{\lambda_e^2\}$ (Equation (10)), which express the magnitude of model error over each element of the domain. Both in the absence of noise and when SNR$= 40dB$, the algorithm clearly identifies a significant model error in the region on the top-left corner. It is noted that the λ_e^2 values in this region are from 2 to 4 orders of magnitude larger than in the rest of the problem domain. Despite the model inadequacy, the algorithm correctly identifies the presence of the circular inclusion as it can be seen in Figure 10 and more clearly in Figure 11, which shows the elastic modulus variation along the diagonal from $(0,1)$ to $(1,0)$. It is particularly interesting to note that even though the isotropic elastic constitutive model endowed in the inversion scheme is inadequate at least for a subdomain of the problem, the proposed scheme can correctly identify the stresses (pressure and shear) in the whole domain as it can be seen in Figures 12 and 13. These depict the ground truth in comparison with the posterior means obtained



(a) no noise - 5% quantile (b) no noise - posterior mean (c) no noise - 95% quantile

(d) SNR=40 - 5% quantile (e) SNR=40 - posterior mean (f) SNR=40 - 95% quantile

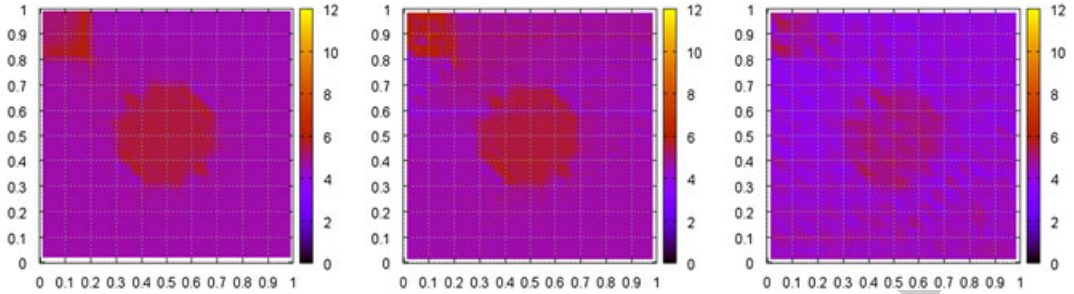
Figure 10. Example 2: Posterior statistics of the elastic modulus distribution when data have no noise and for SNR=40 dB.



(a) no noise (b) SNR=40

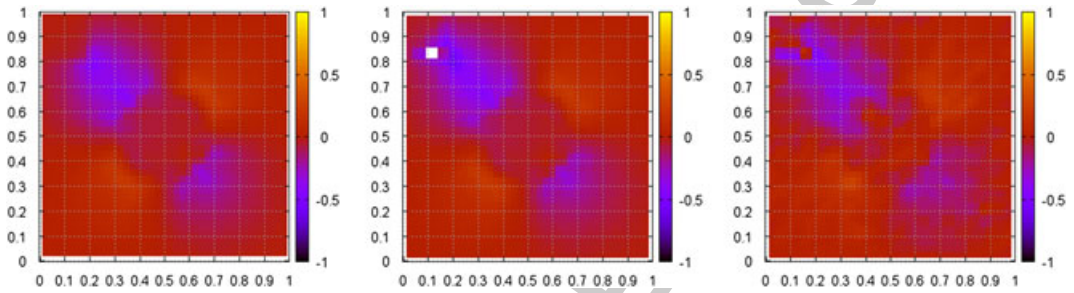
Figure 11. Example 2: Posterior statistics of the elastic modulus distribution along the diagonal from $(0,1)$ to $(1,0)$.

DOI: 10.1002/nme



(a) reference (ground truth) (b) no noise - posterior mean (c) SNR=40 - posterior mean

Figure 12. Example 2: Comparison of pressure's spatial distribution with the posterior means obtained when data have no noise and for SNR=40 dB.



(a) reference (ground truth) (b) no noise - posterior mean (c) SNR=40 - posterior mean

Figure 13. Example 2: Comparison of shear stress' σ_{xy} spatial distribution with the posterior means obtained when data have no noise and for SNR= 40 dB.

with no noise and for SNR=40 dB. The posterior quantiles (which are omitted herein for economy of space) fully envelop the ground truth.

4. CONCLUSIONS

Although existing stochastic (Bayesian) strategies for the solution of inverse problems associated with the identification of material properties in biomechanics are able to account for various sources of uncertainty in the problem, they are generally deficient in terms of assessing model fidelity. We proposed an intrusive formulation that incorporates the various model equations in the likelihood (posterior) and is capable of inferring model discrepancies from noisy displacement data. In contrast to direct methods, it does not require imputations of strains nor their derivatives. It provides probabilistic confidence metrics (credible intervals) that can be very useful to the analyst as well as probabilistic estimates of the (unobserved) stresses/pressures. We discussed a scalable computational framework which can be greatly accelerated by employing a multiresolution strategy. The latter could be utilized in order to propose adaptively, refinements of the discretized domain, which we intend to explore in the future. Current investigations also involve extending this approach to dynamic settings where the parameter vector should include velocities and accelerations in addition to displacements, and the model equations should include the time-integration scheme adopted.

APPENDIX A: MAXIMIZATION WITH RESPECT TO $\boldsymbol{\Lambda}$

This section describes the computations involved during the maximization step of the EM algorithm described in Section 2. In particular, according to Equations (28), (29), (30) and the prior model in

DOI: 10.1002/nme

Equation (11), this entails a maximization with respect to $\mathbf{\Lambda} = \{\lambda_e^2\}_e$ of

$$\begin{aligned} Q(\mathbf{\Lambda}^{(j)}, \mathbf{\Lambda}) &= -\frac{n_\sigma}{2} \sum_{e=1}^{n_{el}} \log \lambda_e^2 - \frac{1}{\lambda_e^2} \sum_{e=1}^{n_{el}} \Phi_e^{(j)} + \log p(\mathbf{\Lambda}) \\ &= -\frac{n_\sigma}{2} \sum_{e=1}^{n_{el}} \log \lambda_e^2 - \frac{1}{\lambda_e^2} \sum_{e=1}^{n_{el}} \Phi_e^{(j)} - \frac{1}{2} \mathbf{Z}^T \mathbf{W} \mathbf{Z}. \end{aligned} \tag{38}$$

It is reminded that the vector $\mathbf{Z} = \{z_e\}_{e=1}^{n_{el}}$ contains the log values of $\mathbf{\Lambda}$, that is, $z_e = \log \lambda_e^2$. Rather than solving an optimization in the n_{el}-dimensional space at each iteration j, we perform successive updates of each λ_e^2 or z_e while keeping the remaining fixed. This *incremental* version of the EM algorithm entails performing n_{el} optimizations of one-dimensional functions. We propose carrying out the latter task with respect to z_e (as they are allowed to take any value on the real axis in contrast to λ_e^2, which must be positive) and employ a standard Newton–Raphson scheme. This requires the first-order and second-order derivatives of the objective function previously mentioned which are given by

$$\frac{\partial Q(\mathbf{\Lambda}^{(j)}, \mathbf{\Lambda})}{\partial z_e} = -\frac{3}{2} + \frac{\Phi_e^{(j)}}{2} e^{-z_e} - \frac{z_e - \mu_{z_e}}{\sigma_{z_e}^2} \tag{39}$$

and

$$\frac{\partial^2 Q(\mathbf{\Lambda}^{(j)}, \mathbf{\Lambda})}{\partial z_e^2} = -\frac{\Phi_e^{(j)}}{2} e^{-z_e} - \frac{1}{\sigma_{z_e}^2}, \tag{40}$$

where

$$\begin{aligned} \sigma_{z_e}^2 &= 1/W_{e,e} \\ \mu_{z_e} &= -\frac{1}{W_{e,e}} \sum_{k \neq e} W_{e,k} z_k. \end{aligned} \tag{41}$$

It can be easily seen that the second derivative is always, strictly negative $\frac{\partial^2 Q(\mathbf{\Lambda}^{(j)}, \mathbf{\Lambda})}{\partial z_e^2} < 0$ and therefore the problem is convex.

REFERENCES

1. Oberai AA, Gokhale NH, Goenezen S, Barbone PE, Hall TJ, Sommer AM, Jiang J. Linear and nonlinear elasticity imaging of soft tissue in vivo: demonstration of feasibility. *PHYSICS IN MEDICINE AND BIOLOGY* 2009; **54**(5):1191–1207.
2. Ganne-Carrié N, Ziol M, de Ledinghen V, Douvin C, Marcellin P, Castera L, Dhumeaux D, Trinchet JC, Beaugrand M. Accuracy of liver stiffness measurement for the diagnosis of cirrhosis in patients with chronic liver diseases. *Hepatology* 2006; **44**(6):1511–1517.
3. Ophir J, Cespedes I, Ponnekanti H, Yazdi Y, Li X. Elastography - a quantitative method for imaging the elasticity of biological tissues. *ULTRASONIC IMAGING* 1991; **13**(2):111–134.
4. Garra BS, Cespedes EI, Ophir J, Spratt SR, Zuurbier RA, Magnant CM, Pennanen MF. Elastography of breast lesions: initial clinical results ~~1997; **202**(1):79–86.~~
5. Bamber JC, Barbone PE, Bush NL, Cosgrove DO, Doyely MM, Fuechsel FG, Meaney PM, Miller NR, Shiina T, Tranquart F. Progress in freehand elastography of the breast. *IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS* 2002; **E85D**(1):5–14.
6. Hall TJ, Zhu YN, Spalding CS. In vivo real-time freehand palpation imaging. *ULTRASOUND IN MEDICINE AND BIOLOGY* 2003; **29**(3):427–435.
7. Giuseppetti GM, Martegani A, Di cioccio B, Baldassarre S. Elastosonography in the diagnosis of the nodular breast lesions: preliminary report. *RADIOLOGIA MEDICA* 2005; **110**(1–2):69–76.
8. Itoh A, Ueno E, Tohno E, Kamma H, Takahashi H, Shiina T, Yamakawa M, Matsumura T. Breast disease: clinical application of us elastography for diagnosis ~~2006; **239**(2):341–350.~~
9. Thomas A, Fischer T, Frey H, Ohlinger R, Grunwald S, Blohmer JU, Winzer KJ, Weber S, Kristiansen G, Ebert B, Kummel S. Real-time elastography- an advanced method of ultrasound: first results in 108 patients with breast lesions. *ULTRASOUND IN OBSTETRICS AND GYNECOLOGY* 2006; **28**(3):335–340.

DOI: 10.1002/nme

10. Regner DM, Hesley GK, Hangiandreou NJ, Morton MJ, Nordland MR, Meixner DD, Hall TJ, Farrell MA, Mandrekar JN, Harmsen WS, Charboneau JW. Breast lesions: evaluation with us strain imaging-clinical experience of multiple observers 2006; **238**(2):425 – 437.
11. Burnside ES, Hall TJ, Sommer AM, Hesley GK, Sisney GA, Svensson WE, Fine JP, Jiang JJ, Hangiandreou NJ. Differentiating benign from malignant solid breast masses with us strain imaging 2007; **245**(2):401–410.
12. Zhi H, Ou B, Luo BM, Feng X, Wen YL, Yang HY. Comparison of ultrasound elastography, mammography, and sonography in the diagnosis of solid breast lesions. *JOURNAL OF ULTRASOUND IN MEDICINE* 2007; **26**(6):807 – 815.
13. Parker KJ, Doyley MM, Rubens DJ. Imaging the elastic properties of tissue: the 20 year perspective. *Phys. Med. Biol.* 2011. Q5 Q6
14. Barbone PE, Rivas CE, Harari I, Albocher U, Oberai AA, Zhang Y. Adjointweighted variational formulation for the direct solution of inverse problems of general linear elasticity with full interior data. *INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING* 2010; **81**(13):1713–1736.
15. McLaughlin J, Renzi D. Using level set based inversion of arrival times to recover shear wave speed in transient elastography and supersonic imaging. *INVERSE PROBLEMS* 2006; **22**(2):707–725.
16. Albocher U, Oberai AA, Barbone PE, Harai I. Adjoint-weighted equation for inverse problems of incompressible plane-stress elasticity. *Comput. Methods Appl. Mech. Engrg* 2009; **198**:2412–2420.
17. Park E, Maniatty AM. Finite element formulation for shear modulus reconstruction in transient elastography. *INVERSE PROBLEMS IN SCIENCE AND ENGINEERING* 2009; **17**(5):605–626.
18. Gockenbach MS, Khan AA. Identification of lame parameters in linear elasticity: a fixed point approach. *JOURNAL OF INDUSTRIAL AND MANAGEMENT OPTIMIZATION* 2005; **1**(4):487–497.
19. Gockenbach MS, Jadamba B, Khan AA. Equation error approach for elliptic inverse problems with an application to the identification of lame parameters. *INVERSE PROBLEMS IN SCIENCE AND ENGINEERING* 2008; **16**(3):349–367.
20. Zhang Y, Hall LO, Goldgof DB, Sarkar S. A constrained genetic approach for computing material property of elastic objects. *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION* 2006; **10**(3):341–357.
21. Oberai AA, Gokhale NH, Doyley MM, Bamber JC. Evaluation of the adjoint equation based algorithm for elasticity imaging. *PHYSICS IN MEDICINE AND BIOLOGY* 2004; **49**(13):2955–2974.
22. Fehrenbach J, Masmoudi M, Souchon R, Trompette P. Detection of small inclusions by elastography. *INVERSE PROBLEMS* 2006; **22**(3):1055–1069.
23. Doyley MM, Srinivasan S, Dimidenko E, Soni N, Ophir J. Enhancing the performance of model-based elastography by incorporating additional a priori information in the modulus image reconstruction process. *PHYSICS IN MEDICINE AND BIOLOGY* 2006; **51**(1):95–112.
24. Liu Y, Sun LZ, Wang G. Tomography-based 3-d anisotropic elastography using boundary measurements. *IEEE TRANSACTIONS ON MEDICAL IMAGING* 2005; **24**(10):1323–1333.
25. Liew HL, Pinsky PM. Recovery of shear modulus in elastography using an ajoint method with b-spline representation. *FINITE ELEMENTS IN ANALYSIS AND DESIGN* 2005; **41**(7-8):778–799.
26. Banerjee B, Roy D, Vasu RM. A pseudo-dynamic sub-optimal filter for elastography under static loading and measurements. *PHYSICS IN MEDICINE AND BIOLOGY* 2009; **54**(2):285–305.
27. Gokhale NH, Barbone PE, Oberai AA. Solution of the nonlinear elasticity imaging inverse problem: the compressible case. *INVERSE PROBLEMS* 2008; **24**(4). Q7
28. Arnold A, Reichling S, Bruhns OT, Mosler J. Efficient computation of the elastography inverse problem by combining variational mesh adaption and a clustering technique. *PHYSICS IN MEDICINE AND BIOLOGY* 2010; **55**(7):2035–2056.
29. Olson LG, Throne RD. Numerical simulation of an inverse method for tumour size and location estimation. *INVERSE PROBLEMS IN SCIENCE AND ENGINEERING* 2010; **18**(6):813–834.
30. Bonnet M, Constantinescu A. Inverse problems in elasticity. *Inverse Problems* 2005; **21**:R1–R50.
31. Kaipio JP, Somersalo E. *Computational and Statistical Methods for Inverse Problems*. Springer-Verlag: New York, 2005.
32. Higdon D, Nakhleb B, Gattiker J, Williams B. A bayesian calibration approach to the thermal problem. *COMPUTER METHODS IN APPLIED MECHANICS AND ENGINEERING* 2008; **197**(29-32):2431–2441.
33. Kennedy MC, O'Hagan A. Bayesian calibration of computer models. *JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES B-STATISTICAL METHODOLOGY* 2001; **63**:425–450.
34. Kohn RV, Vogelius M. Determining conductivity by boundary measurements commun. *Pure Appl. Math* 1984:37289. Q8
35. Ladéveze P, Chouaki A. Application of a posteriori error estimation for structural model updating. *Inverse Problems* 1999; **15**:49–58.
36. Deraemaeker A, Ladéveze P, Romeuf T. Model validation in the presence of uncertain experimental data. *Eng. Comput* 2004;21808–21833. Q9
37. Feissel P, Allix O. Modified constitutive relation error identification strategy for transient dynamics with corrupted data: the elastic case. *Comput. Methods Appl. Mech. Eng* 2007;1968–1983.
38. Besag J, York J, Mollié. Bayesian image restoration, with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics* 1991; **43**:1–59. Q10
39. Besag J, Green PJ. Spatial statistics and bayesian computation. *J. Royal Statist. Soc. Ser. B, Methodological* 1993; **55**:25–37.

DOI: 10.1002/nme

40. Kunsch HR. Intrinsic autoregressions and related models on the two-dimensional lattice. *BIOMETRIKA* 1987; **74**:517.
41. Besag J, Kooperberg C. On conditional and intrinsic autoregressions. *BIOMETRIKA* 1995; **82**(4):733–746.
42. Higdon DM. Auxiliary variable methods for markov chain monte carlo with applications. *Journal of the American Statistical Association* 1997.
43. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the em algorithm (with discussion). *J. Roy. Statist. Soc. Ser. B* 1977; **39**(1):1–38.
44. Ghahramani Z. An introduction to hidden markov models and bayesian networks. *Journal of Pattern Recognition and Artificial Intelligence* 2001; **15**(1):9–42.
45. Meng XL, Rubin DB. Maximum likelihood estimation via the ecm algorithm: a general framework. *Biometrika* 1993; **80**:267–278.
46. Neal R, Hinton GE. A view of the em algorithm that justifies incremental, sparse, and other variants. In *Learning in Graphical Models*. Kluwer Academic Publishers, 1998; 355–368.
47. Ghahramani Z, Attias H. Online variational bayesian learning. *Slides from talk presented at NIPS 2000 workshop on Online Learning*, 2000.
48. Beal MJ, Ghahramani Z. The variational bayesian em algorithm for incomplete data: with application to scoring graphical model structures. *Bayesian Statistics,* (7) 2003.
49. Wainwright MJ, Jordan M. Graphical models, exponential families, and variational inference. In *Foundations and Trends in Machine Learning*, Vol. 1, 2008; 1–305.
50. Robert CP, Casella G. *Monte Carlo Statistical Methods*, 2nd edition. Springer: New York, 2004.
51. Robbins H, Monro S. A stochastic approximation method. *The Annals of Mathematical Statistics* 1951; **22**:400–407.
52. Cappé O, Moulines E, Rydén T. *Inference in Hidden Markov Models*. Springer-Verlag, 2005.
53. Delyon B, Lavielle M, Moulines E. Convergence of a stochastic approximation version of the em algorithm. *The Annals of Statistics* 1999; **27**:94–128.
54. Kushner HJ, Yin G. *Stochastic approximation and recurcive algorithms and applications*. Springer, 2003.
55. Andrieu C, Moulines E, Priouret P. Stability of stochastic approximation under verifiable conditions. *SIAM J. Control Optim.* 2005.
56. Liang F, Liu C, Carroll RJ. Stochastic approximation in Monte Carlo computation. *J. Amer. Statist. Assoc.* 2007; **102**:305–320.
57. Hughes TJR. Generalization of selective integration procedures to anisotropic and non-linear media. *Int. J. Numer. Methods Eng.* 1980; **15**:1413.
58. Hughes TJR. *The Finite Element MethodŮLinear Static and Dynamic Finite Element Analysis*. Dover, 2000.

DOI: 10.1002/nme

Author Query Form

Journal: International Journal for Numerical Methods in Engineering

Article: nme_4261

Dear Author,

During the copyediting of your paper, the following queries arose. Please respond to these by annotating your proofs with the necessary changes/additions.

- If you intend to annotate your proof electronically, please refer to the E-annotation guidelines.
- If you intend to annotate your proof by means of hard-copy mark-up, please refer to the proof mark-up symbols guidelines. If manually writing corrections on your proof and returning it by fax, do not write too close to the edge of the paper. Please remember that illegible mark-ups may delay publication.

Whether you opt for hard-copy or electronic annotation of your proofs, we recommend that you provide additional clarification of answers to queries by entering your answers on the query sheet, in addition to the text mark-up.

Query No.	Query	Remark
Q1	AUTHOR: Changes were made in the affiliations. Please check.	
Q2	AUTHOR: 'Although these might be plausible, in general alter the informational content of the data and make difficult the quantification of the effect of observation noise.' The meaning of this sentence is not clear; please rewrite or confirm that the sentence is correct.	
Q3	AUTHOR: Please check the capturing of pages in References 2 and 16, if correct.	
Q4	AUTHOR: Please provide full journal title in References 4, 8, 10 and 11.	
Q5	AUTHOR: Please provide page range, volume and issue number in Reference 13.	
Q6	AUTHOR: Please provide full journal title in References 13, 16, 34, 36, 37, 39, 43, 55, 56 and 57.	
Q7	AUTHOR: Please provide page range in Reference 27.	
Q8	AUTHOR: Please provide volume and issue number in Reference 34.	
Q9	AUTHOR: Please provide volume and issue number in References 36 and 37.	
Q10	AUTHOR: Please check the capturing of authors name in Reference 38, if correct .	

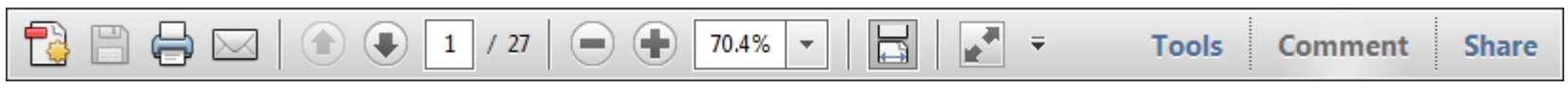
Q11	AUTHOR: If this reference has now been published online, please add relevant year/DOI information. If this reference has now been published in print, please add relevant volume/issue/page/year information.	
Q12	AUTHOR: Please provide publisher address in References 46, 52, 54 and 58.	
Q13	AUTHOR: Please provide location where the conference or proceedings was held in Reference 47.	
Q14	AUTHOR: Please provide page range and volume in Reference 48.	
Q15	AUTHOR: Please check the capturing of Reference 49 if correct.	
Q16	AUTHOR: Please provide publisher and publisher address in Reference 49.	
Q17	AUTHOR: If this reference has now been published online, please add relevant year/DOI information. If this reference has now been published in print, please add relevant volume/issue/page/year information.	

WILEY-BLACKWELL

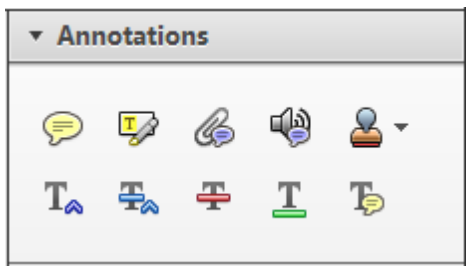
USING e-ANNOTATION TOOLS FOR ELECTRONIC PROOF CORRECTION

Required software to e-Annotate PDFs: Adobe Acrobat Professional or Adobe Reader (version 8.0 or above). (Note that this document uses screenshots from Adobe Reader X)
The latest version of Acrobat Reader can be downloaded for free at: http://get.adobe.com/reader/

Once you have Acrobat Reader open on your computer, click on the Comment tab at the right of the toolbar:



This will open up a panel down the right side of the document. The majority of tools you will use for annotating your proof will be in the Annotations section, pictured opposite. We've picked out some of these tools below:

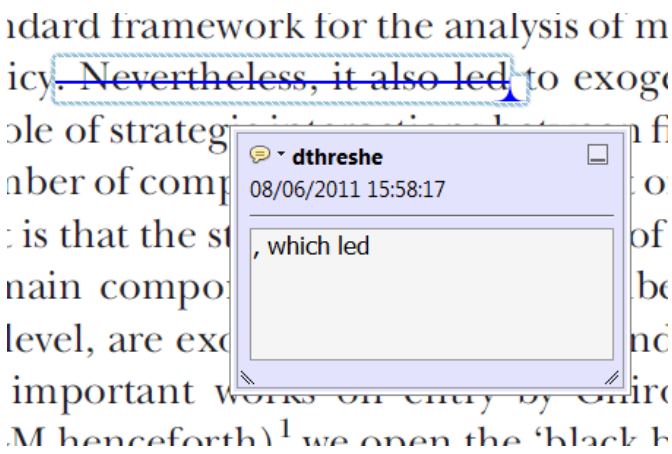


1. Replace (Ins) Tool – for replacing text.

Strikes a line through text and opens up a text box where replacement text can be entered.

How to use it

- Highlight a word or sentence.
- Click on the Replace (Ins) icon in the Annotations section.
- Type the replacement text into the blue box that appears.

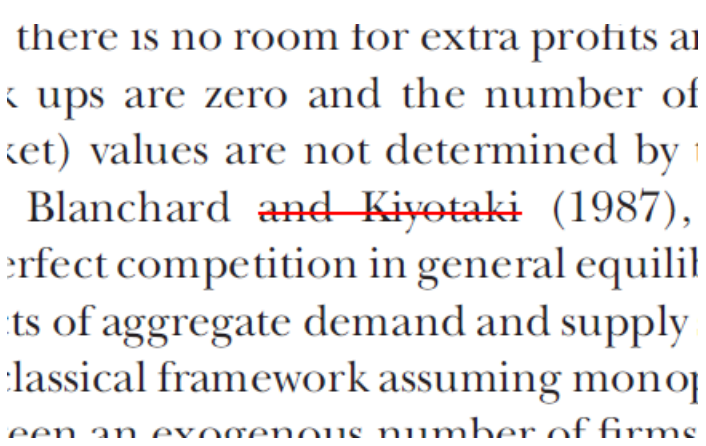


2. Strikethrough (Del) Tool – for deleting text.

Strikes a red line through text that is to be deleted.

How to use it

- Highlight a word or sentence.
- Click on the Strikethrough (Del) icon in the Annotations section.

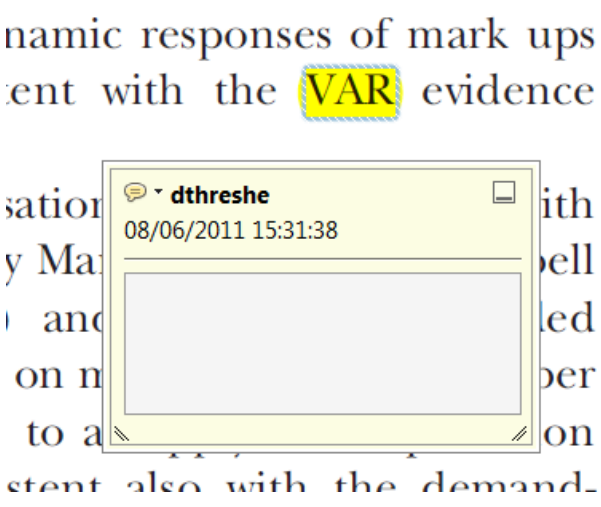


3. Add note to text Tool – for highlighting a section to be changed to bold or italic.

Highlights text in yellow and opens up a text box where comments can be entered.

How to use it

- Highlight the relevant section of text.
- Click on the Add note to text icon in the Annotations section.
- Type instruction on what should be changed regarding the text into the yellow box that appears.

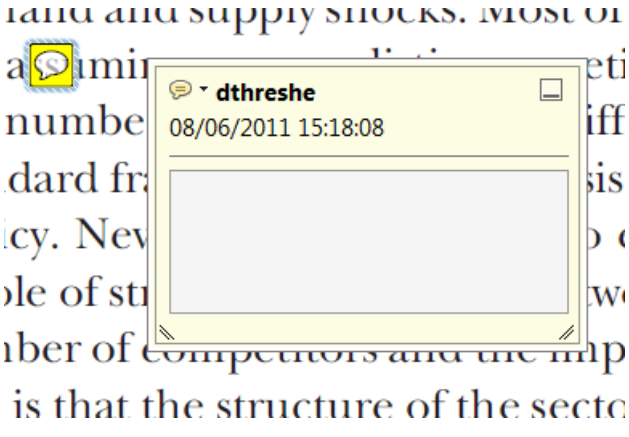


4. Add sticky note Tool – for making notes at specific points in the text.

Marks a point in the proof where a comment needs to be highlighted.

How to use it

- Click on the Add sticky note icon in the Annotations section.
- Click at the point in the proof where the comment should be inserted.
- Type the comment into the yellow box that appears.



WILEY-BLACKWELL

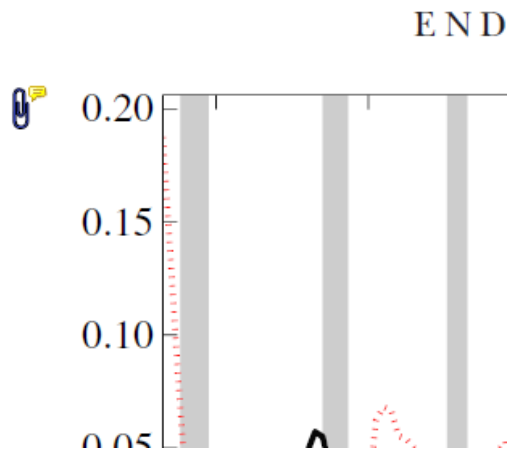
USING e-ANNOTATION TOOLS FOR ELECTRONIC PROOF CORRECTION

5. Attach File Tool – for inserting large amounts of text or replacement figures.

Inserts an icon linking to the attached file in the appropriate pace in the text.

How to use it

- Click on the Attach File icon in the Annotations section.
- Click on the proof to where you'd like the attached file to be linked.
- Select the file to be attached from your computer or network.
- Select the colour and type of icon that will appear in the proof. Click OK.

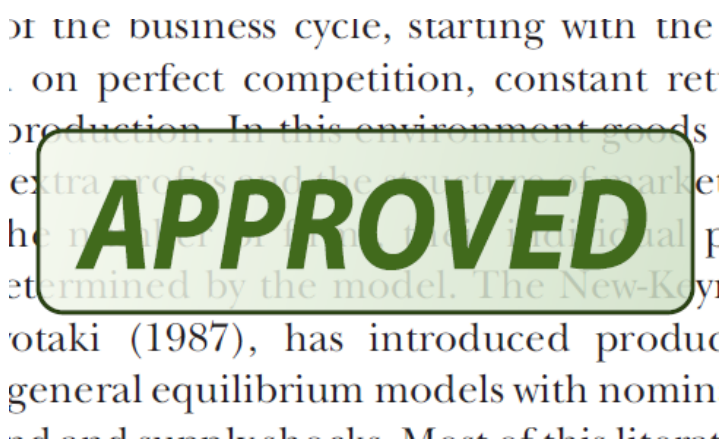


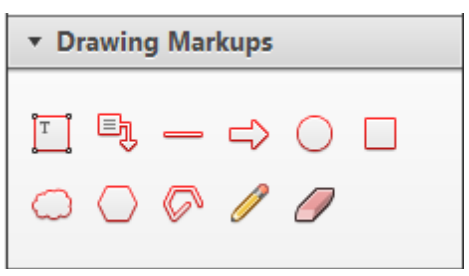
6. Add stamp Tool – for approving a proof if no corrections are required.

Inserts a selected stamp onto an appropriate place in the proof.

How to use it

- Click on the Add stamp icon in the Annotations section.
- Select the stamp you want to use. (The Approved stamp is usually available directly in the menu that appears).
- Click on the proof where you'd like the stamp to appear. (Where a proof is to be approved as it is, this would normally be on the first page).



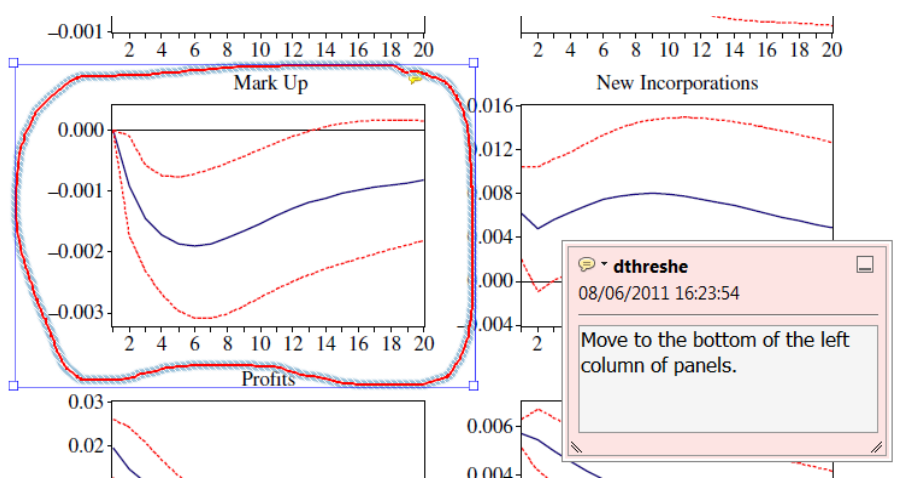


7. Drawing Markups Tools – for drawing shapes, lines and freeform annotations on proofs and commenting on these marks.

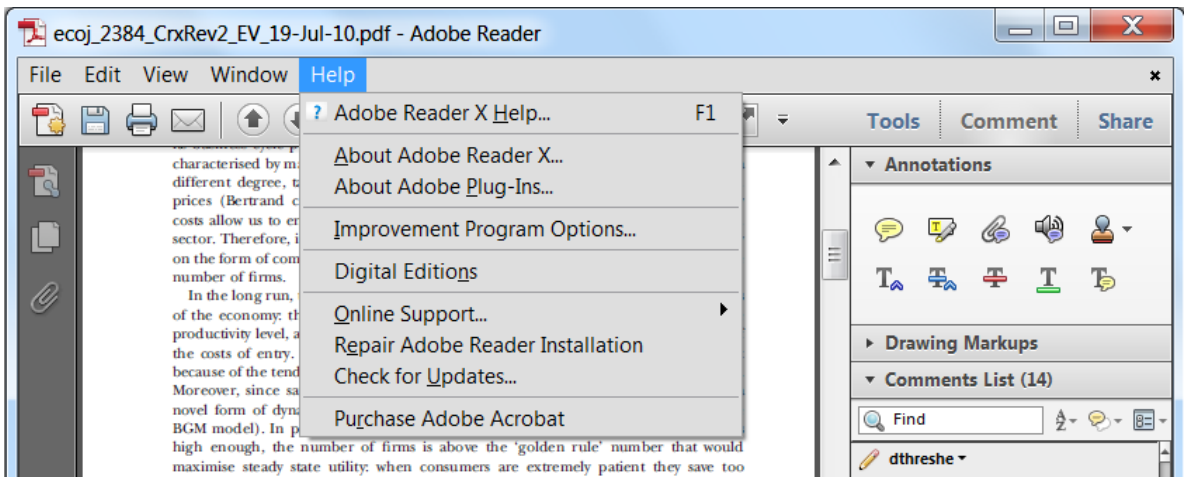
Allows shapes, lines and freeform annotations to be drawn on proofs and for comment to be made on these marks..

How to use it

- Click on one of the shapes in the Drawing Markups section.
- Click on the proof at the relevant point and draw the selected shape with the cursor.
- To add a comment to the drawn shape, move the cursor over the shape until an arrowhead appears.
- Double click on the shape and type any text in the red box that appears.



For further information on how to annotate proofs, click on the Help menu to reveal a list of further options:



WILEY AUTHOR DISCOUNT CLUB

We would like to show our appreciation to you, a highly valued contributor to Wiley's publications, by offering a **unique 25% discount** off the published price of any of our books*.

All you need to do is apply for the **Wiley Author Discount Card** by completing the attached form and returning it to us at the following address:

The Database Group (Author Club)
John Wiley & Sons Ltd
The Atrium
Southern Gate
Chichester
PO19 8SQ
UK

Alternatively, you can **register online** at www.wileyeurope.com/go/authordiscount
Please pass on details of this offer to any co-authors or fellow contributors.

After registering you will receive your Wiley Author Discount Card with a special promotion code, which you will need to quote whenever you order books direct from us.

The quickest way to order your books from us is via our European website at:

http://www.wileyeurope.com

Key benefits to using the site and ordering online include:
- Real-time SECURE on-line ordering
- Easy catalogue browsing
- Dedicated Author resource centre
- Opportunity to sign up for subject-orientated e-mail alerts

Alternatively, you can order direct through Customer Services at:
cs-books@wiley.co.uk, or call +44 (0)1243 843294, fax +44 (0)1243 843303

So take advantage of this great offer and return your completed form today.

Yours sincerely,

Verity Leaver
Group Marketing Manager
author@wiley.co.uk

*TERMS AND CONDITIONS

This offer is exclusive to Wiley Authors, Editors, Contributors and Editorial Board Members in acquiring books for their personal use. There must be no resale through any channel. The offer is subject to stock availability and cannot be applied retrospectively. This entitlement cannot be used in conjunction with any other special offer. Wiley reserves the right to amend the terms of the offer at any time.

REGISTRATION FORM
For Wiley Author Club Discount Card

To enjoy your 25% discount, tell us your areas of interest and you will receive relevant catalogues or leaflets from which to select your books. Please indicate your specific subject areas below.

Accounting [ ]
- Public [ ]
- Corporate [ ]

Chemistry [ ]
- Analytical [ ]
- Industrial/Safety [ ]
- Organic [ ]
- Inorganic [ ]
- Polymer [ ]
- Spectroscopy [ ]

Encyclopedia/Reference [ ]
- Business/Finance [ ]
- Life Sciences [ ]
- Medical Sciences [ ]
- Physical Sciences [ ]
- Technology [ ]

Earth & Environmental Science [ ]

Hospitality [ ]

Genetics [ ]
- Bioinformatics/ Computational Biology [ ]
- Proteomics [ ]
- Genomics [ ]
- Gene Mapping [ ]
- Clinical Genetics [ ]

Medical Science [ ]
- Cardiovascular [ ]
- Diabetes [ ]
- Endocrinology [ ]
- Imaging [ ]
- Obstetrics/Gynaecology [ ]
- Oncology [ ]
- Pharmacology [ ]
- Psychiatry [ ]

Non-Profit [ ]

Architecture [ ]

Business/Management [ ]

Computer Science [ ]
- Database/Data Warehouse [ ]
- Internet Business [ ]
- Networking [ ]
- Programming/Software Development [ ]
- Object Technology [ ]

Engineering [ ]
- Civil [ ]
- Communications Technology [ ]
- Electronic [ ]
- Environmental [ ]
- Industrial [ ]
- Mechanical [ ]

Finance/Investing [ ]
- Economics [ ]
- Institutional [ ]
- Personal Finance [ ]

Life Science [ ]

Landscape Architecture [ ]

Mathematics [ ]
Statistics [ ]

Manufacturing [ ]

Materials Science [ ]

Psychology [ ]
- Clinical [ ]
- Forensic [ ]
- Social & Personality [ ]
- Health & Sport [ ]
- Cognitive [ ]
- Organizational [ ]
- Developmental & Special Ed [ ]
- Child Welfare [ ]
- Self-Help [ ]

Physics/Physical Science [ ]

Please complete the next page /

I confirm that I am (*delete where not applicable):

a **Wiley** Book Author/Editor/Contributor* of the following book(s):
ISBN:
ISBN:

a **Wiley** Journal Editor/Contributor/Editorial Board Member* of the following journal(s):

SIGNATURE: …… Date: ……

PLEASE COMPLETE THE FOLLOWING DETAILS IN BLOCK CAPITALS:

TITLE: (e.g. Mr, Mrs, Dr) …… FULL NAME: ……

JOB TITLE (or Occupation): ……

DEPARTMENT: ……

COMPANY/INSTITUTION: ……

ADDRESS: ……

……

TOWN/CITY: ……

COUNTY/STATE: ……

COUNTRY: ……

POSTCODE/ZIP CODE: ……

DAYTIME TEL: ……

FAX: ……

E-MAIL: ……

YOUR PERSONAL DATA
We, John Wiley & Sons Ltd, will use the information you have provided to fulfil your request. In addition, we would like to:

1. Use your information to keep you informed by post of titles and offers of interest to you and available from us or other Wiley Group companies worldwide, and may supply your details to members of the Wiley Group for this purpose.
   [ ] Please tick the box if you do **NOT** wish to receive this information

2. Share your information with other carefully selected companies so that they may contact you by post with details of titles and offers that may be of interest to you.
   [ ] Please tick the box if you do **NOT** wish to receive this information.

E-MAIL ALERTING SERVICE
We also offer an alerting service to our author base via e-mail, with regular special offers and competitions. If you **DO** wish to receive these, please opt in by ticking the box [ ].

If, at any time, you wish to stop receiving information, please contact the Database Group (databasegroup@wiley.co.uk) at John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, PO19 8SQ, UK.

TERMS & CONDITIONS
This offer is exclusive to Wiley Authors, Editors, Contributors and Editorial Board Members in acquiring books for their personal use. There should be no resale through any channel. The offer is subject to stock availability and may not be applied retrospectively. This entitlement cannot be used in conjunction with any other special offer. Wiley reserves the right to vary the terms of the offer at any time.

PLEASE RETURN THIS FORM TO:
Database Group (Author Club), John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, PO19 8SQ, UK author@wiley.co.uk
Fax: +44 (0)1243 770154