eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Less is often more: applied inverse problems using hp-forward models

Danny Smyl[a,b,], Dong Liu[c,d,e,1]

[a]*Department of Civil and Structural Engineering, University of Sheffield, Sheffield, UK*

[b]*Integrated Civil and Infrastructure Research Centre (ICAIR), Sheffield, UK*

[c]*Hefei National Laboratory for Physical Sciences at the Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China*

[d]*Key Laboratory of Microscale Magnetic Resonance, University of Science and Technology of China, Hefei 230026, China*

[e]*Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, China*

ARTICLE INFO

ABSTRACT

To solve an applied inverse problem, a numerical forward model for the problem's physics is required. Commonly, the finite element method is employed with discretizations consisting of elements with variable size h and polynomial degree p. Solutions to hp-forward models are known to converge exponentially by simultaneously decreasing h and increasing p. On the other hand, applied inverse problems are often ill-posed and their minimization rate exhibits uncertainty. Presently, the behavior of applied inverse problems incorporating hp elements of differing p, h, and geometry is not fully understood. Nonetheless, recent research suggests that employing increasingly higher-order hp-forward models (increasing mesh density and p) decreases reconstruction errors compared to inverse regimes using lower-order hp-forward models (coarser meshes and lower p). However, an affirmative or negative answer to following question has not been provided, "Does the use of higher order hp-forward models in applied inverse problems always result in lower error reconstructions than approaches using lower order hp-forward models?"

In this article we aim to reduce the current knowledge gap and answer the open question by conducting extensive numerical investigations in the context of two contemporary applied inverse problems: elasticity imaging and hydraulic tomography – nonlinear inverse problems with a PDE describing the underlying physics. Our results support a *negative* answer to the question – i.e. decreasing h (increasing mesh density), increasing p, or simultaneously decreasing h and increasing p does not guarantee lower error reconstructions in applied inverse problems. Rather, there is complex balance between the accuracy of the hp-forward model, noise, prior knowledge (regularization), Jacobian accuracy, and ill-conditioning of the Jacobian matrix which ultimately contribute to reconstruction errors. As demonstrated herein, it is often more advantageous to use lower-order hp-forward models than higher-order hp-forward models in applied inverse problems. These realizations and other counterintuitive behavior of applied inverse problems using hp-forward models are described in detail herein.

1. Introduction

Solutions to inverse problems are required in a number of fields within engineering, medicine, and science [1]. Medical imaging is one of the most robust fields for which inverse problems are applied, where a number of tomographic modalities are used to detect, e.g. tissue abnormalities. Some common modalities include X-ray tomography [2, 3, 4], electrical impedance tomography (EIT) [5, 6, 7, 8, 9, 10], photoacoustic imaging/optical diffusion tomography [11, 12, 13], magnetic resonance imaging (MRI) [14, 15], and elasticity imaging [16, 17, 18]. In the broad field of physical sciences, compiling a complete list of applications for inverse problems would be a dubious task. As a representative picture, applied inverse problems in the physical sciences vary from implementing gravitational wave data to solve the relativistic inverse stellar problem [19] to imaging volcano cross sections using EIT and muon imaging [20] to protein structure prediction [21]. Applications of applied inverse problems in engineering are quite diverse; including use cases in structural health monitoring [22, 23, 24, 25], hydraulic tomography [26, 27, 28], multiphase flow tomography [29, 30], materials characterization [31, 32], petroleum reservoir characterization [33], and many more.

Despite the numerous optimization approaches used for solving inverse problems, e.g. adjoint methods [34], Jacobian-based methods [35], and Newton-Krylov methods [36], there is one commonality in applied inverse problems: a numerical model for the problem's physics is required. This model is referred to as a forward model. In a broad sense, we aim to match the data simulated using the forward model to the observed measurements as closely as possible when solving an inverse problem. Therefore, we require a forward model with sufficient accuracy (more on the meaning of "sufficient" later). Over the years, researchers have adopted a suite of different computational regimes for solving forward problems, such as finite difference [37], boundary element [38], and finite element methods (FEM) [39]. Some advantages of the finite difference and boundary element methods are their ease of implementation and speed, respectively. The FEM, however, is likely the most common numerical method used in forward solvers, with a possible exception of inverse problems that may arise from computational fluid dynamics – whereby utilizing well-developed finite volume platforms such as OpenFOAM [40] may be advantageous. The popularity of the FEM in inverse problems possibly results from the well-studied accuracy/convergence behavior [41, 42, 43, 44], ease of discretization [45], straightforward nature of handling boundary conditions [46], and the sparsity of the resulting systems of equations [47]. For these reasons, we focus our attention to inverse problems utilizing finite element-based forward models.

The research area encompassed by the FEM is extensive. Indeed, over the years researchers have proposed innovative regimes such as meshless methods [48, 49], adaptive methods with clouds [50, 51], smoothed

corresponding authors: d.smyl@sheffield.ac.uk (Danny Smyl), dong2016@ustc.edu.cn (Dong Liu)

methods [52], discontinuous methods [53, 53, 54], and phase-field approaches [55, 56]. Many of the afore-mentioned FEM regimes offer highly-accurate solutions by incorporating high-order elements and complex numerical approaches, which may come with a high computational price. As such, researchers have developed methods, such as parallelization, resulting in significant speed-ups [57, 58]. In some work, such as in [59], parallel FEM solvers have even been implemented as forward models in inverse problems. At the heart of much recent FEM research, work has been conducted to improve the accuracy of finite element solutions.

Due to the scarcity of work relating FEM forward solvers accuracy to inverse solutions, however, we aim to first study the effect element size h and polynomial degree p have on inverse solutions. In future work, more advanced FEM solvers will be studied. From first principles, it is well known from early works such as [60] that exponential solution convergence is achieved by simultaneously refining h and increasing p in hp finite element regimes. However, such a guarantee may not necessarily exist for inverse problems implementing hp finite element forward models. This realization results from the ill-posed nature of inverse problems, meaning that at least one of the following criteria does not hold: (i) the solution is unique, (ii) a solution exists, and/or (iii) the behavior of a solutions changes continually with the problem conditions (i.e. small input changes to the problem result in small changes in the solution) [61].

From an implementation standpoint, the solution to an inverse problem often requires an iterative approach. For this, we are required to solve the forward problem numerous times, both in evaluating some cost function and in computing a gradient, Jacobian, etc. (depending on the optimization regime). In addition to attributes contributing to an inverse problem's solution, such as regularization and measurement noise, this realization makes the task of predicting the minimization rate, number of iterations required, and inverse solution accuracy as a function of h and p difficult. Moreover, research investigating the effect of differing hp discretizations in applied inverse solutions is scarce.

Nonetheless, some light has been shed on this topic in the context of, for example, stationary EIT [62] and convection-diffusion inverse problems [63]. In a pioneering work (Ledger, 2012) [62], the author investigated triangular hp discretizations where a range of p from 1 to 5 and several different sizes of h were considered. Regardless of the choice of regularization, Ledger found superior reconstructions when $p > 1$ and that, despite the noise level, higher order (i.e. increased mesh density and p) elements generally resulted in lower reconstruction error. Importantly, the author also noted a distinct trade-off between the resolution used to represent the forward problem and the reconstructed parameter field. In the recent article (Yeo et. al, 2019) [63], the authors utilized different L_1 and L_2 regularization schemes in solving a convection-diffusion inverse problem considering p ranging from 2 to 10 on a rectangular mesh. The authors arrived at a similar conclusion expressed by Ledger, whereby increasing p led to lower error inverse solutions.

Additional relevant research has also investigated adaptive meshing schemes in inverse problems; for example, in [64] the author utilized adaptive meshing and efficient linear solvers for solving large-scale inverse problems. One significant contribution of this work was the implementation of a method bypassing finite-dimensional norm dependence on mesh size. In non-stationary contexts, as in [65], authors have compared

hp adaptive regimes for adjoint systems finding that estimations of the objective function gradient are greatly improved – resulting in increased accuracy of inverse solutions. Such works have helped close gaps with respect to the feasibility of using *hp* elements and adaptive elements in stationary and non-stationary inverse problems.

It is important to reinforce that, at the end of the day, the main purpose of including *hp*-adaptive forward models in reconstruction approaches is to improve solutions to inverse problems. As such, the aforementioned advances in applied inverse problems, among the others fields mentioned, dates back to fundamental research conducted decades ago. Important to the theme of this article is the notion of *distinguishability* proposed by Isaacson in 1986 [66]. In the seminal article, Isaacson determined a criterion, related to EIT measurement precision and the noise level, above which a small inclusion could still be reconstructed. Several years later, the notion of distinguishability was used to deepen researchers' understanding of the stability and resolution of inverse conductivity problems in [67]. By combining information from one or both of these works, authors thereafter began to use relevant concepts in designing improved discretizations used in solving inverse conductivity problems, e.g. in [68, 69, 70]. Interestingly, it was later shown that in special circumstances one may solve an inverse problem to "optimize" the inversion grid used in EIT [71]. Of particular relevance to the inverse problems solved here, researchers have also developed optimal discretization approaches for hydraulic tomography inverse problems, e.g. [72]. Meanwhile, research has also shown that the choice of discretization, itself, has a regularizing or "self-regularizing" effect [73, 74].

Notwithstanding the listed works, where researchers have largely focused on computational and theoretical aspects related to, e.g., forward modeling, discretization, regularization, and stabilizing inverse regimes – others have focused their attention on directly modeling errors intrinsic to the numerical forward models [75, 76, 13] as a means for improving reconstructions. To do this, researchers often recast the standard inverse problem relating some data $\mathbf{d}$, an "exact" (i.e. highly-accurate) forward model $U_A(\theta)$, Gaussian distributed measurement noise $\mathbf{e}$, and some estimated parameter θ:

$$\mathbf{d} = U_A(\theta) + \mathbf{e} \tag{1}$$

by rewriting the observation model in the following form

$$\mathbf{d} = \underbrace{[U_A(\theta) - U_h(\theta)]}_{\mathbf{e}_m} + U_h(\theta) + \mathbf{e} \tag{2}$$

such that a reduced-order forward model $U_h(\theta)$ may be used in generating a model error term $\mathbf{e}_m$ by substituting into the discrepancy term in Eq. 3 and rearranging as follows

$$\mathbf{d} = U_h(\theta) + \mathbf{e}_m + \mathbf{e}. \tag{3}$$

This technique and the estimation of $\mathbf{e}_m$ is called approximation error modeling (AEM) and is often formulated in a Bayesian framework [35]. Solving problems using AEM is attractive for a multitude of reasons. Firstly, it allows for the use of the reduced model U_h, which can significantly decrease computational demand

by allowing for coarser meshes and/or a lower p. Secondly, AEM provides a means for quantifying modeling errors in a statistical sense which may shed useful light onto a problem's physics or the behavior of a given inverse problem. As a whole, AEM is an effective technique that had been broadly applied to, e.g., EIT [76], ultrasound tomography [75], photoacoustic tomography [77], and optical diffusion tomography [13].

In this work, we are inspired by the aforementioned works investigating hp elements in inverse problems and the recent advances in computing allowing us to solve large-scale inverse problems. In this article we aspire to build upon previous works, where we aim to expand the contemporary understanding of inverse problems using hp discretized forward models by quantifying reconstruction sensitivity with respect to (i) discretization density, (ii) differing element shape, (iii) accuracy of derivative (Jacobian) estimates, and (iv) differing p considering a suite of different problem conditions. In addition, we aim to affirmatively or negatively answer the pervasive open research question arising from recent works such as [63, 65, 62], "Does the use of higher order hp-forward models in applied inverse problems always result in lower error reconstructions than approaches using lower order hp-forward models?"

For this, we conduct numerical studies investigating two contemporary applied inverse problems with dissimilar problem physics and ill posedness: *elasticity imaging* (typically an overdetermined problem) and *hydraulic tomography* (typically an underdetermined problem). In these studies, we conduct a large-scale analysis considering geometries with inclusions having random ellipsoidal shapes and locations as well as randomized distributions. The aim of conducting the large suite of reconstructions is to incorporate a statistical perspective on the roles of h and p in solving inverse problems with various levels of measurement noise, data interpolation error, and data reduction (in the case of hydraulic tomography).

The article is organized as follows. We begin by introducing the forward and inverse models for elasticity imaging and hydraulic tomography. Following, we detail the numerical investigation, including the hp elements tested, regularization, the computational resources used, and the numerical tomography case studies. Next, the numerical studies are conducted and discussed. Lastly, conclusions are presented.

2. Elasticity imaging

This section details the forward and inverse problems used in solving the elasticity imaging problem. From a high-level viewpoint, elasticity imaging is the inverse estimation of an inhomogeneous elastic modulus distribution from a quasi-static displacement field. Some applications of elasticity imaging include crack detection in structural health monitoring [78, 79], engineering materials characterization [31], and localization of tissue abnormalities in medical imaging [16, 80, 17, 18, 34, 81, 82]. We further discuss this problem in this section by first defining the forward problem and then formulating the inverse regime.

2.1. Forward problem

In this work, we consider a two dimensional problem in a linear elastic body Ω with a boundary $\partial\Omega$. We then define the boundary tractions $\hat{f}_i$ on $\partial\Omega_{i1}$, which is a part of the boundary and $i = 1, 2$. Following, we

denote the displacements on $\partial\Omega_{i2}$ by $\hat{u}_i$. Neglecting body forces, the elastic forward problem is written as

$$\frac{\partial\sigma_{ij}}{\partial x_j} = 0, \mathbf{x} \in \Omega \tag{4}$$

$$f_i(\mathbf{x}) \equiv \sigma_{ji}n_j \equiv \hat{f}_i, \mathbf{x} \in \partial\Omega_{i1} \tag{5}$$

$$u_i = \hat{u}_i, \in \partial\Omega_{i2} \tag{6}$$

where σ_{ij} is the stress tensor, $\mathbf{x}$ is the spatial coordinate, n_j is the outward unit normal, and $\partial\Omega_{i1} \cup \partial\Omega_{i1} \equiv \partial\Omega$.

In obtaining a discretized finite element formulation of the forward model, we use the weak-Galerkin formulation of Eq. 4 [41]. In the discretized problem, we obtain the familiar system of equations (using integration via Gaussian Quadrature) as a function of an elastic modulus vector $\mathbf{E}$

$$\sum_{i=1}^{n} K_{ij}(\mathbf{E})u_j = F_i \tag{7}$$

where K_{ij} is the stiffness matrix and F_i is the force vector. Conveniently, since we are interested in computing the displacement field, we may also write

$$u_j = \sum_{i=1}^{n} K_{ij}^{-1}(\mathbf{E})F_i \tag{8}$$

where n is the number of unknown displacements, which is dependent on hp discretization prescribed for the problem (these selections are discussed in a later section).

In this work, it is important to note that we assume deformations are infinitesimal, i.e. $||\nabla\mathbf{u}|| \ll 1$. By imparting such an assumption, we recall that the standard strain tensor is given by $\epsilon = \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^{\mathrm{T}})$ [79]. Moreover, we assume plane stress conditions and element-wise isotropy; in other words; Ω is inhomogeneous with respect to $\mathbf{E}$, but the elastic modulus is constant in each element. In the following subsection, we will discuss the implementation of this forward model in the inverse approach.

2.2. Inverse problem

The inverse elasticity imaging problem defined in this work is as follows: given a noisy displacement field $\mathbf{u}_m$, undeformed Ω, and the force vector $\mathbf{F}$, estimate $\mathbf{E}$. The observation model for this problem is written

$$\mathbf{u}_m = U(\mathbf{E}) + \varepsilon \tag{9}$$

where U is the hp forward operator described in Eq. 8 and ε is a term lumping measurement noise and modeling error. To obtain a solution to the problem described in Eq. 9 we minimize the following functional

$$\Psi_{\mathrm{EI}}(\mathbf{E} > 0) = ||\mathbf{L}_E(\mathbf{u}_m - U(\mathbf{E}))||^2 + R(\mathbf{E}) + \Psi_C(\mathbf{E}) \tag{10}$$

where $|| \cdot ||$ is the L_2 norm, $\mathbf{L}_E$ is the Cholesky factor of the inverted noise covariance matrix, explicitly written as $\mathbf{L}_E^{\mathrm{T}}\mathbf{L}_E = \mathbf{W}^{-1}$, $R(\mathbf{E})$ is a regularization functional, and $\Psi_C(\mathbf{E})$ is the constraint functional

which penalizes negative values. The necessity for regularizing the problem results from the ill-posed nature of the elasticity imaging problem, i.e. that standard least-squares solutions are non-unique, do not yield stable solutions, and generally do not meet the Hadamard conditions for a well-posed problem [35]. By incorporating regularization, particularly by incorporating prior information based on realizations of the problem's physics [78], the stability and accuracy of the solution is improved. Moreover, in obtaining a minimum of Ψ_{EI}, we may incorporate constrained optimization to ensure solutions computed with the forward operator are realistic, i.e. non-complex by setting $\mathbf{E} > 0$.

To solve the constrained minimization problem given in Eq. 10, we use a Gauss-Newton algorithm incorporating a linesearch to find the optimal step-size s_k. The general iterative regime is written as

$$\mathbf{E}_k = \mathbf{E}_{k-1} + s_k \bar{\mathbf{E}} \tag{11}$$

where $\mathbf{E}_k$ denotes the current estimate at iteration k and $\bar{\mathbf{E}}$ is the least-squares update given by

$$\bar{\mathbf{E}} = (\mathbf{J}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{J} + \Gamma_R^{-1} + \Gamma_C)^{-1}(\mathbf{J}^{\mathrm{T}}\mathbf{W}^{-1}(\mathbf{u}_m - U(\mathbf{E}_{k-1})) - \mathbf{g}_R - \mathbf{g}_C) \tag{12}$$

where $\mathbf{J} = \frac{\partial U(\mathbf{E})}{\partial \mathbf{E}}$ is the Jacobian matrix, Γ_R^{-1} is the inverse of the prior covariance matrix, $\mathbf{g}_R$ is the regularization gradient term (if the form of regularization is differentiable), and Γ_C and $\mathbf{g}_C$ are the Hessian and gradient terms related to the constraints. While selection of the regularization terms will be discussed in a later section, the handling of the non-negativity constraint was done using the interior point method with second degree polynomials following [22]. The methods used for computing $\mathbf{J}$, however, were selected with care since their calculation is the most computationally intensive and time-consuming portion of the algorithm. Based on this realization, we selected two methods for computing $\mathbf{J}$. In the first method, $\mathbf{J}$ is computed from the forward model at each iteration and is highly accurate but computationally demanding. In the second method, $\mathbf{J}$ is approximated at each iteration and is therefor less accurate, but significantly less computationally demanding. A description of both methods is provided in the following.

2.2.1. An "accurate" and "slow" method for computing $\boldsymbol{J}$

We begin by describing the first method. Since the inverse problems solved herein include geometries with many degrees of freedom, we tested several methods for computing $\mathbf{J}$ from the forward models including (i) standard forward differencing, (ii) central differencing, (iii) semi-analytical computation following [83], and (iv) complex step differentiation [84]. In preliminary studies, it was found that the use of method (i) was not sufficiently accurate and resulting in slow convergence of the Hessian $\mathbf{H} = \mathbf{J}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{J}$ requiring many iterations until the stopping criteria was reached. On the other hand, the use of method (ii) did result in sufficiently accurate computations of $\mathbf{H}$; however, estimation of $\mathbf{J}$ using central differencing is overly slow and computationally demanding, requiring $2 \times N_d$ solves of the forward model, where N_d are the degrees of freedom in $\mathbf{E}$. The use of method (iii) provided the most accurate computations of $\mathbf{H}$, which are known to be increasingly accurate near the global minimum. However, while method (iii) was excellent for problems

with low N_d, problems with $N_d > 10^3$ were infeasible on the available computing platform since the method requires the storage a gradient matrix that scales with the square of N_d. The final option, (iv), however was sufficiently accurate and fast (relative to methods (ii) and (iii)), requiring N_d forward solves per iteration, and was therefore selected as the first method for computing $\mathbf{J}$.

The advantage of complex differentiation methods over standard perturbation methods largely lies in the realization that estimates of $\mathbf{J}$ are not corrupted by subtractive canceling errors. This is because complex differentiation does not require a differencing operation. As such, it has been shown that even standard central differencing, using two data points for each first derivative estimate, is less accurate than complex step methods using only one data point [85]. Given these advantages, the complex-step method for computing $\mathbf{J}(\mathbf{E})$ was utilized and is written below in terms of the "j$^{\text{th}}$" column entry

$$\mathbf{J}_j = \frac{\text{Im}(U(\mathbf{E} + \text{i}\beta))}{\beta} \tag{13}$$

where i is an imaginary number and β is the step size, which is selected as the machine precision. Next, we describe the less accurate but faster method for computing $\mathbf{J}$.

2.2.2. A "reduced accuracy" and "fast" method for computing $\mathbf{J}$

One challenge with using Eq. 13 in an iterative scheme is the sheer amount of time, memory, and number of cores (if the loop is in parallel) required. This is especially true in inverse problems with many degrees of freedom. The reason for this lies in the fact the cost of solving $\mathbf{J}$, alone, using differentiation scales as $\mathcal{O}(N_d^3)$ [18]. This cost can be drastically reduced by updating $\mathbf{J}$ rather than re-computing it using the forward model at each iteration.

As our second method for computing $\mathbf{J}$, we use Broyden's method [86] which takes advantage of the secant method. In the context of elasticity imaging, we begin by initializing $\mathbf{J}$ by setting $\mathbf{J}_{k=1} = \mathbf{J}_0(\mathbf{E}_0)$, where $\mathbf{J}_0$ is computed using Eq. 13 and $\mathbf{E}_0$ is the initial guess detailed in a later section. For successive iterations ($k > 1$), we use

$$\mathbf{J}_k = \mathbf{J}_{k-1} + \frac{\Delta U_k - \mathbf{J}_{k-1}\Delta\mathbf{E}_k}{||\Delta\mathbf{E}_k||^2}\Delta\mathbf{E}_k^{\text{T}} \tag{14}$$

where $\Delta U_k = U(\mathbf{E}_{k-1}) - U(\mathbf{E}_{k-2})$ and $\Delta\mathbf{E}_k = \mathbf{E}_{k-1} - \mathbf{E}_{k-2}$.

The obvious advantage of Broyden's method for computing $\mathbf{J}$ is that the complexity is $\mathcal{O}(N_d^2)$. There are some well-known disadvantages of Broyden's method in optimization problems, including (i) slower convergence of the optimization scheme, (ii) propagating/cumulative errors in $\mathbf{J}$ as a result of round-off errors, and (iii) sensitivity to the initial matrix $\mathbf{J}_0$ [87]. While other, possibly more accurate methods of updating $\mathbf{J}$ exist [88, 89], we reinforce that we are interested in studying the computational performance of inverse problems using hp discretized forward models in the presence of errors in $\mathbf{J}$ and therefore $\mathbf{H}$. Further details regarding the computational environments of these regimes will be provided in section 4.

3. Hydraulic tomography

This section details the forward and inverse approaches for solving the hydraulic tomography problem. Broadly speaking, hydraulic tomography is used to characterize aquifers by measuring pressure head at discrete locations (wells) [28]. Tomograms generated from this modality are useful in characterizing the constitution, porosity, permeability/conductivity, and other properties of an aquifer [27, 90]. Hydraulic tomography, however is a wide field for which there are numerous forms of input data, using for example oscillatory or transient signals [26, 91]. In this work, however, we investigate a stationary case of hydraulic tomography, where we consider reconstructing the hydraulic conductivity in steady-state conditions. In the following, we begin by describing the forward problem and then the inverse approach.

3.1. Forward problem

We consider a two-dimensional aquifer Ω with a boundary defined by $\partial\Omega_{i1} \cup \partial\Omega_{i1} \equiv \partial\Omega$ and $i = 1, 2$. The flow equations for this forward problem are described by the following equations

$$\mathbf{S}^{-1}\mathbf{v} + \nabla\mathrm{p} = 0, \mathbf{x} \in \Omega \tag{15}$$

$$\nabla \cdot \mathbf{v} = f, \mathbf{x} \in \Omega \tag{16}$$

$$\mathrm{p} = \mathrm{p}_0, \mathbf{x} \in \partial\Omega_{i1} \tag{17}$$

$$\mathbf{v} \cdot \mathbf{n} = v_0, \mathbf{x} \in \partial\Omega_{i2} \tag{18}$$

where $\mathbf{S}$ is the spatially inhomogeneous hydraulic conductivity tensor, $\mathbf{v} = -\mathbf{S}\nabla\mathrm{p}$ is the velocity, p is the unknown pressure head, f is a source term, and p_0 and v_0 are prescribed boundary conditions. Moreover, in a discretized form, we denote the numerical velocity $\mathbf{v}_{\bar{h}}$, for which we require that the mass is locally conserved on each element e's boundary ∂e by writing

$$\int_{\partial e} \mathbf{v}_{\bar{h}} \cdot \mathbf{n} = \int_e f \tag{19}$$

and we also enforce normal flux continuity on edges γ between elements e_i such that

$$\int_\gamma \mathbf{v}_{\bar{h}|e_1} \cdot \mathbf{n}_1 + \int_\gamma \mathbf{v}_{\bar{h}|e_2} \cdot \mathbf{n}_2 = 0. \tag{20}$$

To compute p given in Eqs. 15-18, we use the mixed variational formulation described in [92, 93]. For this, we conduct numerical integration over the hp discretization using Gaussian Quadrature and solve the symmetric indefinite linear system by implementing an adaptation of the open source finite element code DarcyLite [94]. In the following subsection, we will discuss the implementation of the described forward model in the hydraulic tomography approach adopted herein.

3.2. Inverse problem

The hydraulic tomography problem solved in this work is defined as follows: given a noisy vector of pressure data $\mathbf{p}_m$ measured at discrete wells $\mathbf{w}(\mathbf{x}_w, \mathbf{y}_w)$ with coordinates $(\mathbf{x}_w, \mathbf{y}_w)$, aquifer geometry Ω, source terms p_0 and v_0, determine the inhomogeneous hydraulic conductivity $\mathbf{S}$. The observation model for this problem is then written

$$\mathbf{p}_m = U(\mathbf{S}) + \varepsilon \tag{21}$$

where U is the hp hydraulic tomography forward problem described in the previous subsection. In obtaining a solution to this problem, we aim to minimize the following functional

$$\Psi_\mathrm{S}(\mathbf{S} > 0) = ||\mathbf{L}_E(\mathbf{p}_m - U(\mathbf{S}))||^2 + R(\mathbf{S}) + \Psi_C(\mathbf{S}) \tag{22}$$

where $\mathbf{L}_E$, R, and Ψ_C are the same as described in section 2.2 and details on the choice of R will be detailed later. Moreover, the constraint $\mathbf{S} > 0$ results from the fact that hydraulic conductivity cannot be negative. To solve the hydraulic tomography problem, we again adopt a Gauss Newton regime where, by substitution into Eq. 11 we have

$$\mathbf{S}_k = \mathbf{S}_{k-1} + s_k \bar{\mathbf{S}} \tag{23}$$

where $\bar{\mathbf{S}}$ is the least-squares update computed using Eq. 12 and simply substituting the appropriate parameters for the hydraulic tomography problem. For obtaining $\mathbf{J} = \mathbf{J}(\mathbf{S})$ we also use complex differentiation described in Eq. 13 for the "accurate" computation and Eq. 14 for the fast and "reduced accuracy" computation.

4. Numerical investigation

In this section, we detail important components of the numerical investigation, where we study the efficacy of hp elements in elasticity imaging and hydraulic tomography. We begin by describing the hp elements selected for both elasticity imaging and hydraulic tomography. Following, we discuss the selection regularization techniques used in solving the inverse problems. Next, we detail the numerical testing program. Lastly, we discuss the computational resources and details relevant to implementing the inverse algorithms.

4.1. Selection of hp elements

In this paper, we consider two dimensional inverse problems. This is largely for two reasons, (i) 2D problems remain the most common types of inverse problems in contemporary literature - possibly due to the computational reduction in comparison to 3D problems; (ii) the extension of 2D problems to 3D is generally straightforward, with the addition of significant computational cost (especially when considering hierarchical hp elements) [95]. After choosing the problem dimensions, element shapes were considered. For

this, triangular and quadrilateral elements were selected since they are widely used in solutions to a variety of inverse problems.

Initially, candidate triangular and quadrilateral elements were considered having polynomial order ranging from $p =1$ to 5 following [62]. During preliminary testing, we aimed to determine a reasonable range of p satisfying both of the following criteria (i) a stopping criteria (here we used $(\Psi_{k-1} - \Psi_k)/\Psi_{k-1} < 10^{-5}$) was reached in less than less than 2 hours and (ii) sufficient RAM is available for an inverse problem with 10^3 estimation parameters on a machine with 32 cores at 2.4 GHz having a total of 128 Gb RAM. For the target parameter distributions in the trials, a randomized Gaussian blob-like distributions were considered. Moreover, criteria (i) and (ii) were chosen based on available computing resources and knowledge that hundreds of test cases are required to gain adequate simulation statistics.

Based on the preliminary trials, the following element types were selected for both elasticity imaging and hydraulic tomography inverse problems: T3 (3-node triangle), T6, Q4 (4-node quadrilateral), Q8, and Q9. These elements are classified as $p = 1$ and $p = 2$ elements. Moreover, simulated data needs to be generated by employing discretizations with a reduced h and higher degree in p relative to the inverse meshes to avoid an inverse crime [61]. For this, Q16 ($p = 3$) elements were chosen for simulating data on dense forward meshes. We would like to mention that the selection of $p = 1, 2$ due to available computation resources for the inversion meshes is a limitation in this study, especially given the prevalent use of cubic elements in inverse problems. Elements with $p > 2$ will be studied in follow-on work. A summary of the elements used in this study is provided in Fig. 1.
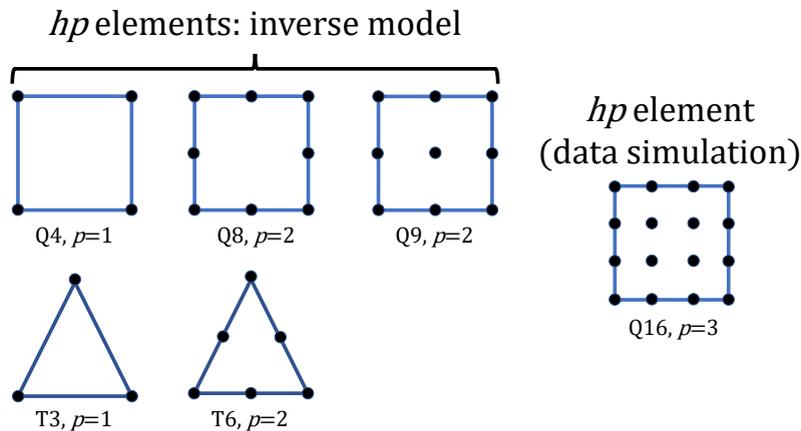
Fig. 1. Schematic showing the hp finite elements used in solving the inverse forward problems (left) and the hp element used for simulating data (right).

4.2. Regularization

Due to the ill-posed nature of elasticity imaging and hydraulic tomography, regularization is required. Mathematically speaking, regularizing these ill-posed problems restricts the domain to a more compact set which results to a conditionally well-posed problem [96]. While this feature of regularization is essential for

solving many inverse problems, regularization also has the functionality of incorporating prior information in the solution regime. In cases where, for example, solutions are known to exhibit smoothness or sparsity, a number of regularization techniques are available to promote such distributions. In cases where sparsity of the estimated parameter field is expected, Total Variation (TV) regularization is often used [97]. On the other hand, when at least some of the solution is expected to be spatially smooth, Tikhonov-type (e.g., ridge regression and elastic net) and spatially-weighted regularizers have proven successful in a number of applications [98].

In this work we solve inverse problems where discretizations range from coarse to fine, therefore we desire regularizers without mesh dependency. This omits certain gradient sparsity-promoting regimes such as TV, were decreasing mesh size results in solution smoothing [99]. For this we choose L_2 Tikhonov regularization (ridge regression) and spatially-weighted regularization. For both methods, we require terms present in the cost functional and least squares update. In the case of Tikhonov regularization, which has C^0 continuity, and therefore does not include a gradient term, we have the functional

$$R_T(\theta) = ||\Gamma_T \theta||^2 \tag{24}$$

where the subscript T denotes "Tikhonov", θ is the estimation parameter (here, $\mathbf{E}$ or $\mathbf{S}$), and the prior matrix Γ_T is defined by

$$\Gamma_T = \lambda \mathbf{I} \tag{25}$$

where $\mathbf{I}$ is the identity matrix and λ is the regularization parameter adaptively selected at each iteration using 5-point L-curve analysis following the description in [100]. In the L-curve analysis used here, the aim is to plot the tradeoff between the data fit $\rho = ||\mathbf{L}_E(\mathbf{d} - U(\theta))||^2$ and the "size" of the regularized solution $\eta = R_T(\theta)$ in estimating the "optimal tradeoff". Here, the optimal tradeoff is estimated to be the corner of the log-log L-curve given by $(\hat{\rho}/2, \hat{\eta}/2) = (\log(\rho), \log(\eta))$ – this is, in essence, a compromise between data fitting and regularization. To estimate the location of the L-curve corner, we conduct a one-dimensional search by first fitting a spline using the five trial points and then finding the minimum of the curvature inverse function $(\kappa(\lambda))^{-1}$ (i.e. finding the maximum curvature) given by

$$(\kappa(\lambda))^{-1} = \left(2 \frac{\rho\eta}{\eta'} \frac{\lambda^2 \eta' \rho + 2\lambda\eta\rho + \lambda^4 \eta\eta'}{(\lambda^2\eta^2 + \rho^2)^{3/2}} \right)^{-1} \tag{26}$$

where η' is first derivative with respect to λ.

For spatially-weighted regularization, we have

$$R_W(\theta) = ||L_W(\theta - \theta_{\exp})||^2 \tag{27}$$

where $\theta_{\exp}$ (also used for $\mathbf{E}_0$ and $\mathbf{S}_0$ in Broyden's method) is determined from the best homogeneous one-parameter estimate for θ using a basic search function, and L_W is the the Cholesky factorization of the

prior matrix Γ_W, (i.e. $\Gamma_W^{-1} = L_W{}^T L_W$). Specifically, Γ_W is determined in an element wise manner, where the matrix element (i,j) for a parameter θ at locations x_i and x_j is given by

$$\Gamma_W(i,j) = a \exp\left(- \frac{||x_i - x_j||}{2b} \right) + c\delta_{ij} \tag{28}$$

where a, b, and c are positive scalars and δ_{ij} is the Kronecker delta function. For the coefficients, parameter a controls the weighting, b is incorporates spatial correlation, and c is small positive parameter used to guarantee that inverse of Γ_W exists [61]. In selecting a, $a = V(\theta_{\exp})$ is chosen, where V is the prior variance defined as $V = (\frac{\theta_{\exp}}{6})^2$, which is based on statistical premises described in [35]. For b, a fixed value of 0.1 is used and the coefficient c was chosen as 10^{-5}. Moreover, we also require the prior gradient at each iteration k in the least square update, which is defined by

$$\mathbf{g}_W = \theta_{k-1} - \theta_{\exp}\Gamma_W^T. \tag{29}$$

Some primary advantages of both selected regularization models are that they (i) do not require computation of a spatial gradient at each iteration, which can become costly as h is refined, (ii) they are robust for reconstructing both spatially-smooth distributions and sharp inclusions, and (iii) they are widely used in solving inverse problems [101] and therefore their behavior in hp regimes may be of broad interest. Disadvantages include (i) for Tikhonov regularization, the 5-point L-curve needs to be calculated at each iteration which is time consuming, (ii) L-curve solutions may be overly smooth [102], (iii) computing the large/dense matrix Γ_W is time consuming for a small h, and (iv) fluctuations in the background values of θ are sometimes present when using weighted regularization for reconstructing inclusions in an otherwise homogeneous media.

4.3. Numerical testing: elasticity imaging

In this study, we are interested in evaluating the performance of elasticity imaging using different hp discretized forward models. For this, we consider two imaging applications, (i) localization and quantification of randomized ellipsoidal inclusions and (ii) imaging randomized blob-like distributions. These applications are effective in testing the modality's robustness for reconstructing general cases, such as sharp features (present as local damage in structural health monitoring [78] or tissue abnormalities in medical applications [34]) and smooth distributions (e.g. organs [17] or inhomogeneous material distributions [31]).

For application (i), ellipses, simulating cracks with near-zero (10^{-2}) stiffness, having random centroid locations (x_l, y_l) and dimensions $0.5 \leq (x_d, y_d) \leq 5$, are superimposed on a 10×10 fine quadrilateral mesh consisting of $65 \times 65 = 4225$ elements with a uniform elasticity distribution $E_{\text{homogeneous}} = 200$. For application (ii), randomized blob-like distributions are generated on the same mesh with a range $100 \leq \mathbf{E}_{\text{target}} \leq 200$. We note that all dimensions, distributed forces (applied nodally), and displacements are unitless in these applications. In generating the data, simulations are carried out for both applications on the fine mesh with Q16 elements. Following, the displacement data is then corrupted with Gaussian noise having

standard deviations of 1, 2, and 5% of the corresponding displacement value. For consistency, the noisy data is computed only once per application and reused in solving the reconstruction problems. Moreover, in solving the elasticity imaging problems on inverse meshes of differing h, the noisy data is interpolated onto the coarse inverse mesh nodes using bilinear interpolation. To better illustrate the applications discussed thus far, a schematic summarizing applications (i) and (ii) and their boundary conditions are shown in Fig. 2.
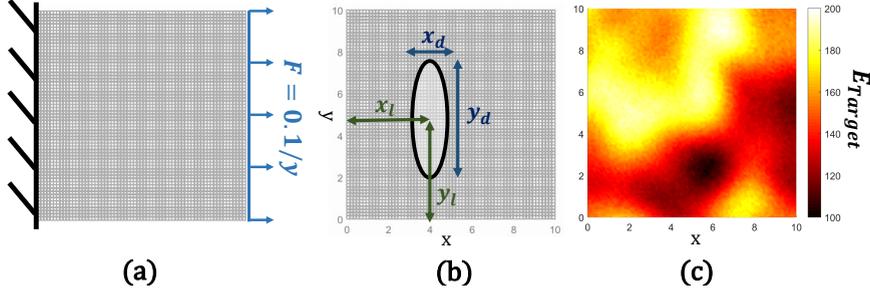
Fig. 2. Schematic showing (a) the boundary conditions for the elasticity imaging problem atop the Q16 forward model discretization, (b) parameters for the random ellipses generated for application (i), and (c) a randomized blob-like distribution used for application (ii).

For solving the inverse problems, the T3, T6, Q4, Q8, and Q9 elements described in section 4.1 are used. To ensure that the uniform quadrilateral and triangular inversion meshes are afforded the same spatial resolution in x and y directions, the triangular meshes require twice the elements compared to the quadrilateral meshes (i.e. two right triangles per quadrilateral). In addition, for solving elasticity imaging problems using complex differentiation to compute $\mathbf{J}$ at each iteration, and based on the preliminary analysis discussed in section 4.1: quadrilateral meshes with $N_d = 100, 225, 400, 625, 900$, and 1225 elements are employed, while the triangular meshes have double these values. When Broyden's method is used for computing $\mathbf{J}$, $N_d = 100, 225, 400, 625, 900, 1225, 1600$, and 2025 elements are used and the triangular meshes have double these values. The increased mesh density and degrees of freedom used in problems employing Broyden's method results from the significantly decreased computational demand.

In the elasticity imaging numerical testing program described herein, a total of 84000 elasticity imaging problems are solved. This number corresponds to both applications (i) and (ii), where we aim to reconstruct 100 different randomized ellipsoidal inclusions and 100 different random blob-like distributions per hp discretization. For this, we consider five different hp elements, two different methods for computing $\mathbf{J}$, three levels of noise, and between six to eight differing degrees of h. The large number of problems was chosen so that we may gain statistical insight into the performance of the hp regimes considering the broad suite of variables influencing reconstruction quality (i.e. accuracy of the forward model, sensitivity to noise, accuracy of $\mathbf{J}$, discretizaiton density, and hp element type). A summary of the elasticity imaging numerical program is provided in Table 1.

| Study	Targets	hp Elements	J Method	Noise (%)	Quadrilateral Meshes (N_d)	Triangular Meshes (N_d)
App. (i)	100	T3, T6, Q4, Q8, Q9	Complex	1, 2, 5	100, 225, 400, 625, 900, 1225	200, 550, 800, 1250, 1800, 2450
	100	T3, T6, Q4, Q8, Q9	Broyden	1, 2, 5	100, 225, 400, 625, 900, 1225, 1600, 2025	200, 550, 800, 1250, 1800, 2450, 3200, 4050
App. (ii)	100	T3, T6, Q4, Q8, Q9	Complex	1, 2, 5	100, 225, 400, 625, 900, 1225	200, 550, 800, 1250, 1800, 2450
	100	T3, T6, Q4, Q8, Q9	Broyden	1, 2, 5	100, 225, 400, 625, 900, 1225, 1600, 2025	200, 550, 800, 1250, 1800, 2450, 3200, 4050

Table 1. Summary of the elasticity imaging numerical program. Abbreviations "Broyden" and "Complex" refer to the use of Broyden's method and Complex differentiation for computing J .

4.4. Numerical testing: hydraulic tomography

In this numerical study, we are interested in testing the performance of the hydraulic tomography inverse problem using various hp discretizations. We must first comment on the stark differences between the hydraulic tomography inverse problem and other inverse problems (e.g. elasticity imaging), largely resulting from the data size. Unlike many inverse problems, hydraulic tomography is generally an *underdetermined* problem – meaning that the number of data points is less than the inverse degrees of freedom (for element-wise constant estimates: $N_d > n$ where n is the number of data points). This fact results from the practical limitations of hydraulic tomography measurements; namely the difficulty of collecting n data points from wells physically dug or drilled into the target aquifer [103]. As a result, the hydraulic tomography inverse problem is more ill-conditioned than if the problem was *overdetermined* ($n > N_d$) [104]. One may therefore expect hydraulic tomography inverse problems using different hp-forward models to exhibit different sensitivities than those of overdetermined problems, such as elasticity imaging - which is of particular interest in the context of the article's theme within "applied inverse problems."

To test the hydraulic tomography regimes with hp-forward models, we simulate data on a fine 10×10 mesh with $65 \times 65 = 4225$ Q16 elements. From these geometries, representing aquifers having hydraulic conductivity distributions $\mathbf{S}$, we interpolate the nodal pressure data $\mathbf{p}_d$ onto n wells using bi-linear interpolation resulting in the measurement pressure vector $\mathbf{p}_m$. For this, we selected five evenly spaced well grids consisting of $n = 25, 81, 169, 289$, and 441 wells which are shown plotted atop a sample of a ranzomized target distribution of $\mathbf{S}$ in Fig. 3. The 100 different target distributions of $\mathbf{S}$ were selected as randomized distributions varying three orders of magnitude (2.5×10^{-8} – 2.5×10^{-5}). Although unitless, these magnitudes are consistent with ranges that may be found in groundwater aquifers [105]. Consistent with sandbox experimental conditions, the left side, right side, and the top boundary were modeled as constant-head boundary conditions, while the bottom boundary was modeled a zero flux boundary [106]. As in the elasticity imaging numerical study, we again use unitless dimensions for consistency.

In the numerical study, we investigate cases where well measurements $\mathbf{p}_m$ were corrupted with 1 and 5% noise. For the meshing of the inverse problem, we utilize the same quadrilateral and triangular discretizations used in the elasticity imaging study (except the densest two meshes implemented when using Broyden's method). This reduced number of elements in applications using Broyden's method resulted from the fact that the solutions to the hydraulic tomography forward problems were notably more computationally demanding than those of the elasticity imaging forward models. As a result, only quadrilateral meshes
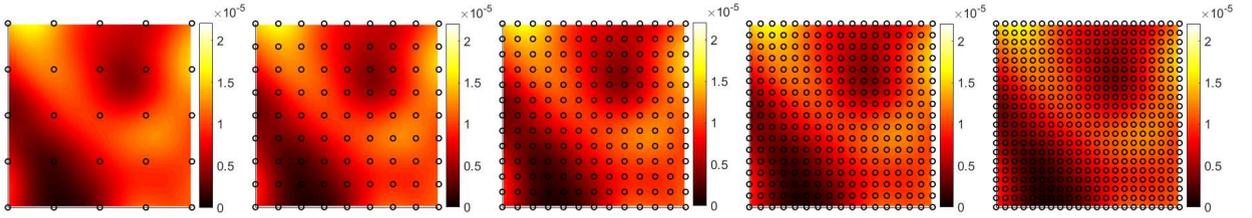
Fig. 3. Evenly spaced well grids used for hydraulic tomography reconstructions. The grids consisted of 25, 81, 169, 289, and 441 wells which are shown as black circles atop a sample of a randomized target distribution of the hydraulic conductivity S. The colorbars denote the range of S for the target aquifer.

with 100 - 1225 elements and triangular meshes with 200 - 2450 elements were used to ensure the problems were tractable using the available computing resources. In solving the inverse problems, we again utilized the weighted prior and L-curve analysis considering both complex differentiation and Broyden's method for computing $\mathbf{J}$. In total, 120000 hydraulic tomography inverse problems are solved in this study and are summarized in Table 2. The large number of problems is again selected to gain adequate statistical insight as to the performance of hydraulic tomography inverse problems using a suite of different hpforward models.

Study	Wells	Targets	hp Elements	J Method	Noise (%)	Quadrilateral Meshes (N_d)	Triangular Meshes (N_d)
Randomized **S**	25, 81, 169, 289, 441	100	T3, T6, Q4, Q8, Q9	Complex	1, 5	100, 225, 400, 625, 900, 1225	200, 550, 800, 1250, 1800, 2450
	25, 81, 169, 289, 441	100	T3, T6, Q4, Q8, Q9	Broyden	1, 5	100, 225, 400, 625, 900, 1225	200, 550, 800, 1250, 1800, 2450

Table 2. Summary of the hydraulic tomography numerical program. Abbreviations "Broyden" and "Complex" refer to the use of Broyden's method and Complex differentiation for computing J .

4.5. Computational considerations and implementation

The inverse problems solved in this work range from small ($N_d =$100 parametric degrees of freedom) to large ($N_d = 4050$). In general, however, the most demanding portion of the code is not conducting a linesearch or solving a system of equations related to the least-squares update, which scales directly with the number of parametric degrees of freedom. Rather, the computational demand lies mainly in computing $\mathbf{J}$, which is a primary disadvantage of least-squares problems. As an example, in the "worst case scenario" encountered in this article we are required compute $\mathbf{J}$ with complex differentiation using T6 elasticity imaging forward discretizations with $N_d = 4050$, where the resulting complex stiffness matrix $\mathbf{K}(\mathbf{E}) \in \mathbb{C}^{16562 \times 16562}$. Therefore, when $\mathbf{J}$ is computed using complex differencing, we are required to assemble and solve the resulting linear system $U(\mathbf{E}) = \mathbf{K}(\mathbf{E})^{-1}\mathbf{F}$ 4050 times per iteration.

In order to deal with such problems, and considering the large number of elasticity and hydraulic imaging problems solved, we found the following computing strategy to be effective. Using MATLAB as the main code wrapper, all hp-forward models were coded as C-source `mex` files using sparse representations of the stiffness matrices. For computing $\mathbf{J}$, which calls the forward solver in computing the "j^{th}" column, the `parfor` command was used, allowing for N_c CPUs to compute different columns of $\mathbf{J}$ simultaneously.

To execute the numerical programs proposed in this article, we utilized the University of Sheffield SharC

HPC facility. For each job, we utilized a system comprised of $N_c = 32$ Intel Xeon E5-2630 v3 CPUs at 2.4 GHz and a total of 128 Gb 1866 MHz DDR4 RAM. As a whole, the jobs ran between 24 - 96 hours, where only one noise level and method for computing $\mathbf{J}$ was considered per job. Moreover, all iterative solutions were considered sufficiently minimized when the objective function stopping criteria $(\Psi_{k-1} - \Psi_k)/\Psi_{k-1} < 10^{-5}$ was reached.

5. Results and discussion

This section reports and discusses the numerical results for elasticity imaging and hydraulic tomography studies. We begin by presenting the results for the elasticity imaging followed by hydraulic tomography. Discussion is provided for both studies.

5.1. Numerical results: elasticity imaging

In this subsection, we present the results for elasticity imaging applications (i) and (ii), where we aim to reconstruct 100 randomly sized and located ellipsoidal cracks and 100 random blob-like distributions, respectively. For this, we utilize different hp discretizations, prior models, two methods for computing $\mathbf{J}$, and investigate three levels of noise (1, 2, and 5%) - specific details for this study are provided in Table 1. Since a total of 84000 reconstructions are computed, providing the totality of these reconstructions visually is virtually impossible. Therefore, we begin by providing visualizations for a suite of reconstructions considering one ellipsoidal inclusion, one blob-like randomized distribution, and 2% noise. These target images are shown in Fig. 4. The first group of reconstructions, computed using complex differentiation, is shown in Fig. 5.
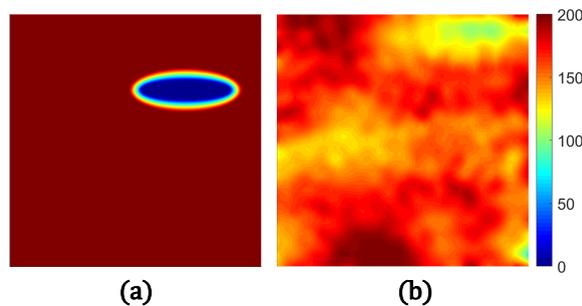
Fig. 4. True elasticity modulus distributions (E, unitless) for reconstruction demonstrations using hp discretizations: (a) ellipsoidal inclusion and (b) randomized blob-like distribution.

Upon first visual inspection of inclusion reconstructions of Fig. 5 (top half), it immediately apparent that, for any element type, the sharpness of the of the inclusion and overall reconstruction quality are heavily influenced by the mesh density. This is obvious when comparing reconstructions in column 2 (low density) to those in the far right column (high density), where the low density discretizations are far more blurry. This result is expected, since the lower density meshes obviously cannot well depict sharp ellipsoidal geometries in comparison to high density meshes. Moreover, it is well known that increases in mesh density result
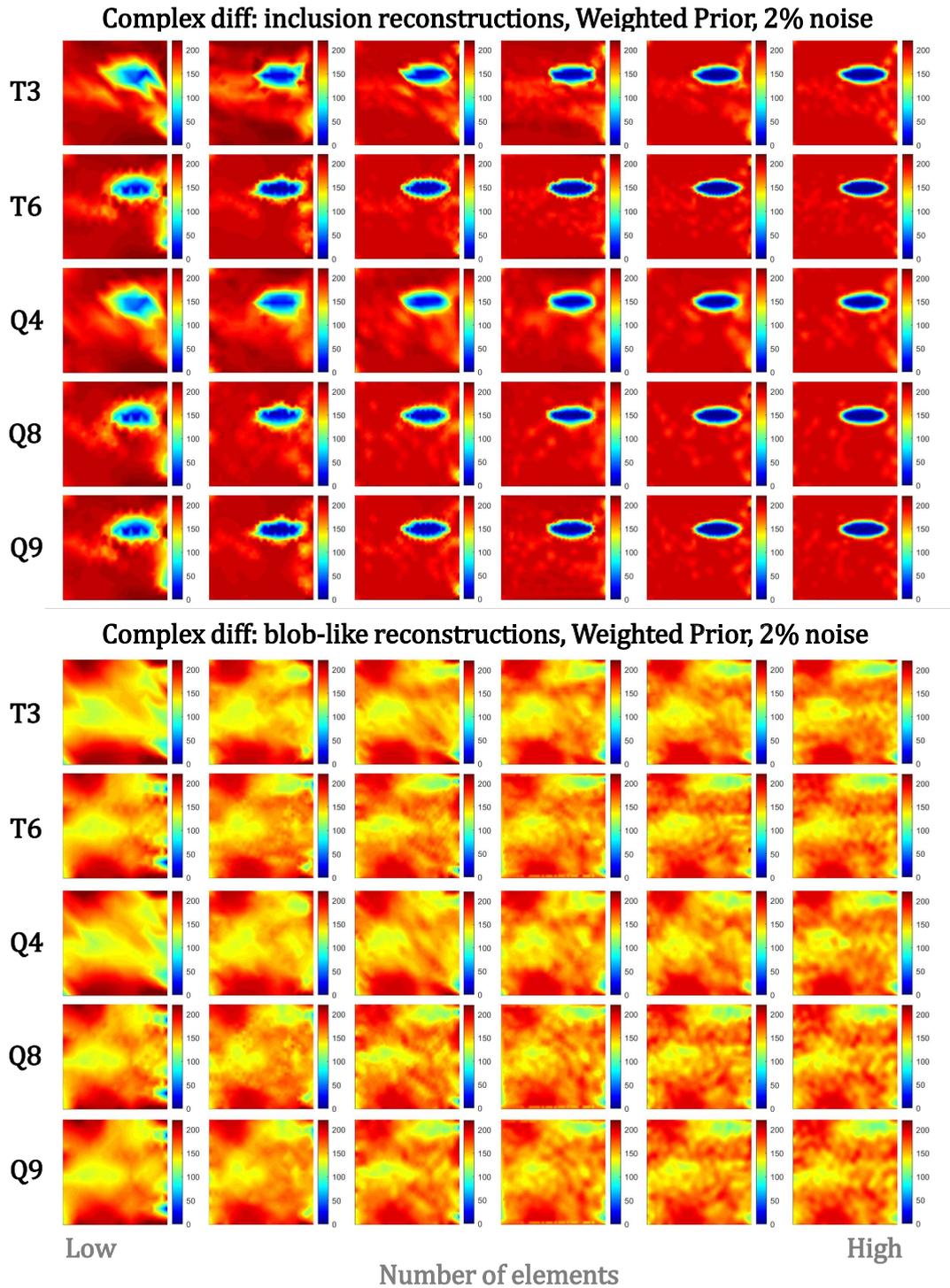
Fig. 5. 30 elasticity imaging reconstructions of an ellipsoidal inclusion and 30 reconstructions for a randomized blob-like distribution. Reconstructions are computed utilizing complex differentiation for obtaining J, 5 different element types, and input displacement data with 2% noise. The left column denotes the element type and the bottom row denotes the discretization density, which increases from left to right. For triangular elements (T3/T6) the number of elements vary between 200 and 2025 while quadrilateral element discretizations (Q4/Q8/Q9) range between 100 and 1225 elements.

in more accurate finite element solutions. We may therefore preliminary surmise the the improvements in reconstruction quality attained by increasing mesh density are a compound effect of the realizations.

Perhaps less intuitive is the effect of p on reconstruction quality. Comparing T3/T6 and Q4/Q8/Q9 elements for any fixed mesh density, we observe that increasing element order noticeably improves the inclusions' sharpness. This effect is less obvious as mesh density increases; which is especially true when comparing Q8/Q9 reconstructions at the highest mesh density - where there are miniscule differences in the inclusions' sharpness or background values. Although not an expected result, it is interesting to note that elements using the same interpolation scheme, i.e. T3/Q4 and T6/Q8/Q9, have similar visual reconstruction quality. What's more, this visual observation holds true despite the fact that the quadrilateral discretizations have half the elements relative to the triangular discretizations. This may indicate that quadrilateral elements may have a computational advantage over triangular elements in this application.

Next, we consider the reconstructions of randomized blob-like distributions. A similar trend to that observed for the ellipsoidal reconstructions is also observed here, i.e. increased image resolution with increasing discretization density and p. The primary difference here, however, is the increase in boundary artifacts apparent in reconstructions with low discretization density. These artifacts are mostly present on the right hand side of these low element density reconstructions, where the distributed load is applied to fewer nodes in areas where $\mathbf{E}$ is locally reduced. The local reduction in stiffness and increase in load results in higher localized stresses not experienced to the same degree in meshes with higher density. These artifacts are present despite the element type, indicating that increasing element order may not improve reconstruction quality for low-density discretized elasticity imaging problems.

We now analyze the reconstructions obtained using Broyden's method for computing $\mathbf{J}$; these results are shown in Fig. 6. As a whole, the ellipsoidal reconstructions obtained with Broyden's method are of far lower quality than those using complex differentiation. In all reconstructions, the ellipsoidal shape of the inclusion is not well captured; in fact most reconstructions are more rectangular than ellipsoidal. Moreover, there are significantly more reconstructions artifacts present. The lowered reconstruction quality obtained using Broyden's method results primarily from the reduced accuracy of $\mathbf{J}$, which decreases in accuracy in each successive iteration.

## Broyden's method: inclusion reconstructions, Weighted Prior, 2% noise
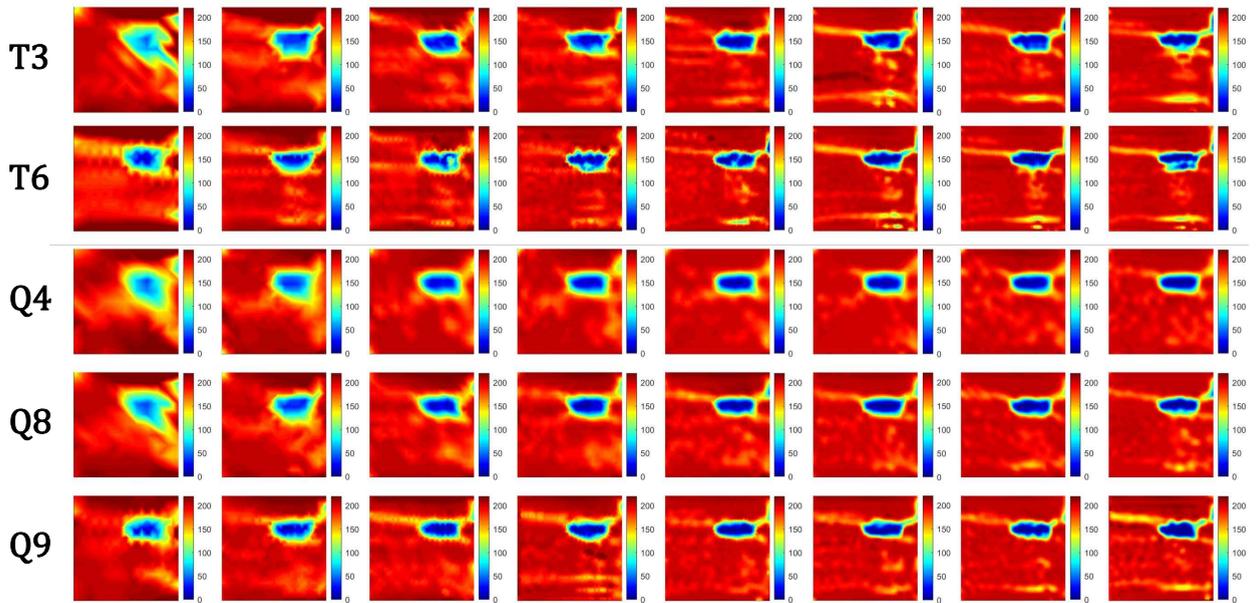
Broyden's method: blob-like reconstructions, Weighted Prior, 2% noise

Low

High

Number of elements
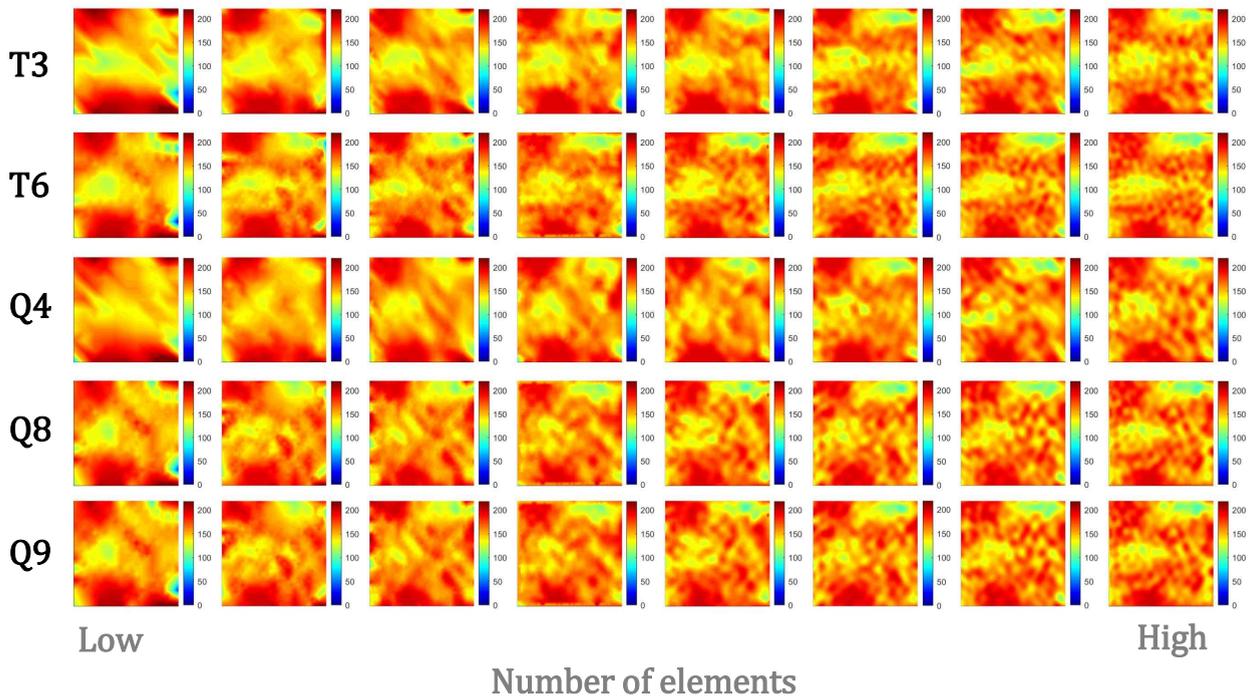
Fig. 6. 30 elasticity imaging reconstructions of an ellipsoidal inclusion and 30 reconstructions for a randomized blob-like distribution. Reconstructions are computed utilizing Broyden's mehthod for obtaining J, 5 different element types, and input displacement data with 2% noise. The left column denotes the element type and the bottom row denotes the discretization density, which increases from left to right. For triangular elements (T3/T6) the number of elements vary between 200 and 4050 while quadrilateral element discretizations (Q4/Q8/Q9) range between 100 and 2025 elements.

An unexpected result here is that the sharpness of the reconstructions using Broyden's method (although not the correct ellipsoidal shape) are better preserved in reconstructions with higher p. The influence of element order on reconstruction sharpness was also observed in reconstructions using complex differentiation. Based on the visual observations, we preliminarily conclude that the preservation of sharpness in reconstructions (with sufficiently high discretization density) using higher order elements results from two primary factors, (i) computations of $\mathbf{J}$ are more accurate and (ii) the sensitivity of $\mathbf{E}$ to noise in $\mathbf{u}_m$ is decreased. However, since these observations are qualitative, considered only one type of regularization, and is only based on two groups of reconstructions, we require a more results and a quantitative analysis to confirm or deny this.

To more comprehensively and quantitatively study the performance of hp elements in elasticity imaging, we analyze the normalized root mean square error (NRMSE) of the reconstructions. We begin by providing line plots of the mean NRMSEs of the reconstructions using complex differentiation in Fig. 7. In these plots, each data point is the mean of 100 reconstructions with the corresponding standard error bars. These line plots also include standard error bars, where the standard error (SE) is computed using $SE = \frac{\sigma}{\sqrt{T}}$ where $\sigma = \sigma(\text{NRMSE})$ is the standard deviation of the sample and $T = 100$ is the number of samples (targets to reconstruct, either inclusions or blob-like distributions) at each point (discretization density, h).
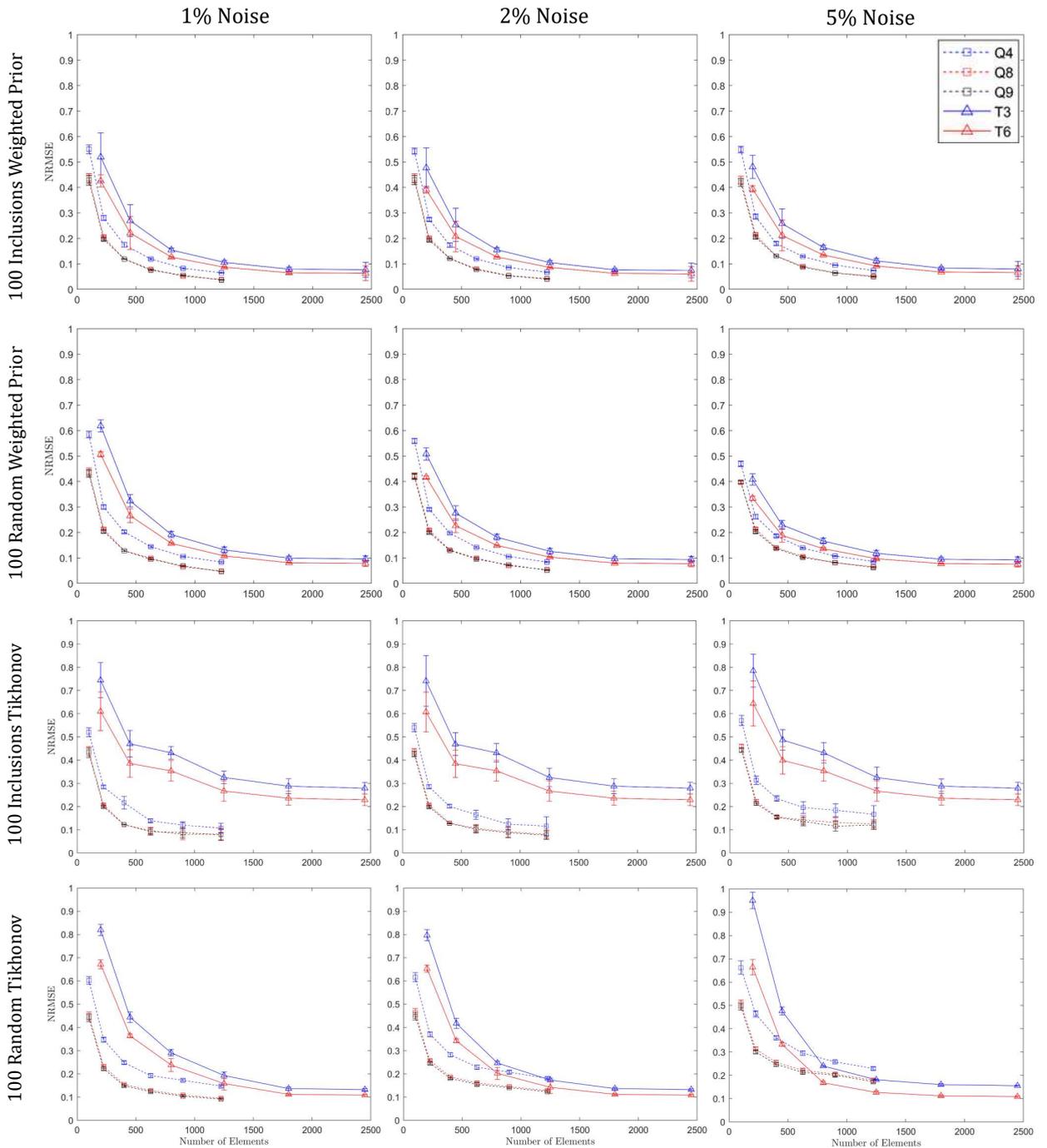
Fig. 7. NRMSEs of elasticity imaging reconstructions plotted against the number of inverse degrees of free-doms. Each data point is the mean of 100 reconstructions with the corresponding standard error bars. Two test cases are considered: simulated inclusions with random locations/ellipsoidal geometries and randomized blob-like distributions. Reconstructions are computed using complex differentiation for determining J considering 5 different hp elements. The top two rows represent reconstructions using the weighted prior while the bottom two rows use Tikhonov regularization. Each column represents a different noise level.

Results shown in Fig. 7 rows 1 and 2, reconstructions using the weighted prior, indicate an expected trend; that is, (a) NRMSEs increase with increasing noise, (b) NRMSEs decrease with both increasing discretization density and increasing p (for a given element shape), and (c) SE predominately decreases with increasing discretization density and p. Consistent with visual observations from the reconstructed images shown earlier, reconstructions using linear elements resulted in higher NRMSEs than reconstructions using quadratic elements. Interestingly, the quadratic T6 elements with the weighted prior had higher NRMSEs than the linear Q4 elements using the weighted prior, possibly indicating superior performance of quadratic elements using a weighted prior in this application.

On the other hand, reconstructions in the bottom two rows of Fig. 7, utilizing Tikhonov regularization, exhibit quite different behavior. Indeed, in the case where ellipsoidal inclusions are reconstructed with Tikhonov regularization, there is a large gap in NRMSE between triangular and quadrilateral discretizations where quadrilateral elements performed significantly better (lower NRMSE) than triangular elements despite the p level. In contrast, upon analyzing reconstructions of randomized distributions using Tikhonov regularization, results at 2 and 5% noise deviate from expectation. At 2 and 5% noise (bottom row), reconstructions using triangular meshes often outperform reconstructions using quadrilateral elements. For example, at 5% noise, for reconstructions with approximately 1250 elements, the performance of reconstructions using T3 elements is comparable to reconstructions using Q8/Q9 elements[1]. Moreover, T6 elements notably outperformed the Q8/Q9 elements at medium-high discretization densities ($N_d \approx 800 - 1250$) despite having fewer nodal degrees of freedom per element. This inconsistent behavior of reconstruction quality $w.r.t.$ element type is in contrast to the predictable behavior observed when employing the weighted prior. This realization may be due to the fact that Tikhonov regularization is inferior to the weighted prior model for reconstructing smooth blob-like distributions [107]. These quantitative results show that the use of higher-order hp discretizations – alone – does not guarantee lower NRMSEs or improved reconstruction quality relative to lower order hp discretizations in elasticity imaging. Rather, other factors (clearly the choice of prior information/regularization) play a significant role.

At this point, the quantitative question remains as to how the reduced accuracy of computing $\mathbf{J}$ using Broyden's method affects reconstruction quality and NRMSEs when employing different hp discretizations in elasticity imaging. To answer this question, we analyze the performance of of elasticity imaging using Broyden's method as a function of inverse degrees of freedom in Fig 8. We immediately notice that the NRMSEs of reconstructions using the weighted prior are roughly comparable to reconstructions using complex differentiation. This indicates that the sensitivity of elasticity imaging reconstructions to the method by which $\mathbf{J}$ is computed is significantly influenced by the choice of prior model. In the case of all elasticity imaging reconstructions computed using the weighted prior, we may surmise that the method used in com-

---

[1]In general, the performance of elasticity imaging reconstructions using Q8 and Q9 elements are shown to be essentially indistinguishable in this work.

puting **J** does not significantly affect NRMSEs. This observation is owed to the stabilization added to the inverse problem via spatial structural information incorporated in the prior.
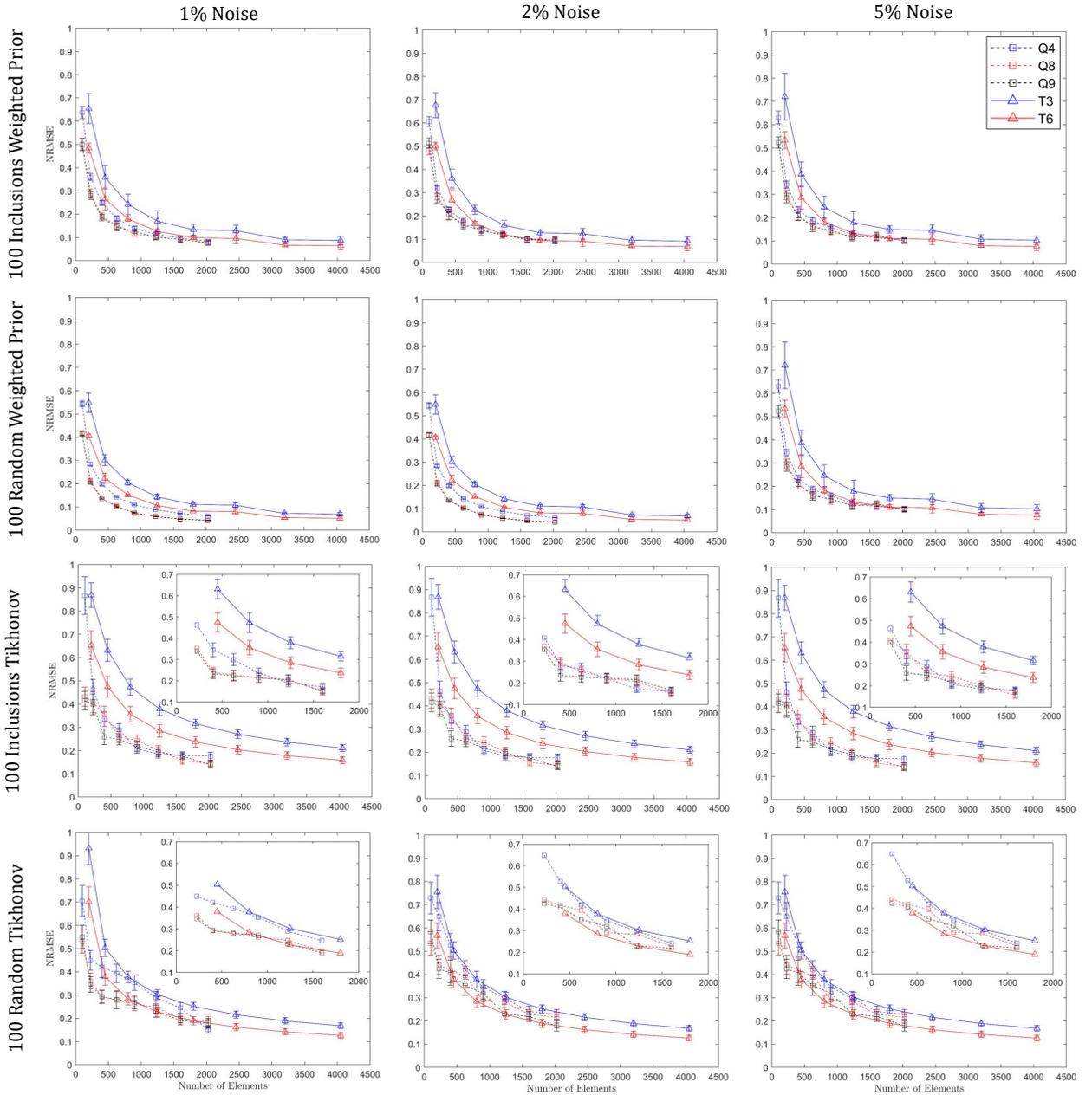
Fig. 8. NRMSEs of elasticity imaging reconstructions plotted against the number of inverse degrees of freedoms. Each data point is the mean of 100 reconstructions with the corresponding standard error bars. Two test cases are considered: simulated inclusions with random locations/ellipsoidal geometries and randomized blob-like distributions. Reconstructions are computed using Broyden's methods for determining J considering 5 different hp elements. The top two rows represent reconstructions using the weighted prior while the bottom two rows use Tikhonov regularization. Each column represents a different noise level.

On the other hand, when the chosen prior model contains less spatial prior information – as in the case of Tikhonov regularization – we notice sharply contrasting behavior to that observed with the weighted

prior. In the cases where we reconstruct inclusions using Tikhonov regularization, we again observe superior performance of the quadrilateral elements. However, unlike when complex differentiation was used, the performance of Q4, Q8, and Q9 elements is only slightly distinguishable at 1% noise and completely indistinguishable at 2% and 5% noise levels. This behavior is quite peculiar, since it is expected that, e.g., Q4 elements have more modeling error than Q8/Q9 elements. Therefore, reconstructions using Q4 elements would be expected to have noticeably higher NRMSE.

Possibly more intuitively interesting is the performance of all elements in the cases where we aim to reconstruct randomized distributions using Tikhonov regularization. In general, the behavior of regimes using triangular elements preform consistently – with the expected decease in NRMSE with increasing degrees of freedom. This was also observed in cases where complex differentiation was used in reconstructing inclusions, which indicates that the performance of regimes using triangular discretizations are not as sensitive as regimes using quadrilateral discretizations to the method used for computing $\mathbf{J}$. It is, however, unexpected that T6 elements generally outperform Q9 elements. Meanwhile, it appears that the NRMSEs of quadrilateral elements, while generally decreasing with increasing degrees of freedom, show some scatter. This observation indicates that elasticity imaging regimes using quadrilateral discretizations and Tikhonov regularization are highly sensitive to errors intrinsic in using "reduced-accuracy" methods for computing $\mathbf{J}$. This realization, among others reported in this subsection result from the joint effects of hp discretizations, regularization, and sensitivites related to how $\mathbf{J}$ is computed. Further discussion is provided in the following subsection.

5.2. Discussion of elasticity imaging results

As a whole, the results reported in the previous subsection illuminate counterintuitive behaviors of elasticity imaging inverse problems using hp-elements. It was demonstrated that the ability to accurately reconstruct an image involves a complex interplay between (a) prior modeling, (b) the hp-discretization, and (c) the Jacobian. Indeed, it was shown that the use of higher-order hp-discretizations does not universally guarantee a decrease in NRMSE or an increase reconstruction quality relative to lower-order hp-discretizations. This counterintuitive reality is largely influenced by the fact that there always exists an interplay between errors in the data (noise), errors in the hp-forward model, and the sensitivity matrix (in the case of least-squares problems, $\mathbf{J}$).

Regarding point (a), the use of a proper prior model that well represents the problem's spatial characteristics is expected to improve results and decrease reconstructions errors. This was clearly reflected by the general trend of lower NRMSEs when employing the weighted prior relative to standard Tikhonov regularization equipped with L-curve analysis. In the case of the weighted prior, the increases in performance were noted to result from improved spatial information. On the other hand, when the less informative prior model was used, the intuitive assumption that "higher-order discretizations should always lead to better inverse solutions" was deemed invalid.

Interestingly, it was found that elasticity imaging regimes employing lower order discretizations can outperform regimes with higher-order discretizations. Of course, part of this is due to a loss in stability of the inverse problem owed to the incorporation of less prior information. However, the interplay between the hp-forward model and the data also has an important role. This issue is at the very heart of this article and is discussed in the following.

As in any inverse problem, these exists errors in both the data and the forward model. The errors in the data are decreased by increasing measurement precision. On the other hand, errors in the forward model are decreased by increasing p and reducing h. However, as noted in [108], "*errors in the data should not converge faster than the errors in the operator* [forward model] ... *there is no need to measure something more precisely than we can predict it.*" In other words, employing an overly-accurate forward model may actually degrade inverse solutions. In fact, the authors noted that in applied applications, it may be useful to match the accuracy of the forward model to the data quality. This is result is highly counterintuitive; but, in the context of finite computational resources: the realization that the use of lower-order hp-discretizations may result in better results than the use of higher-order hp-discretizations may be a relief. Further discussion on this point will be provided in Section 6.

Lastly, we arrive at point (c), the influence of the Jacobian and its approximation on elasticity imaging results. The effects of this point are realized through (i) the accuracy of the Jacobian estimate at each iteration and (ii) the condition number (CN) of $\mathbf{J}$. Broadly speaking, it is well-known that the use of approximate methods for computing $\mathbf{J}$ (such as Broyden's method) generally result in less accurate estimates for $\mathbf{J}$ compared to analytical, semi-analytical or numerical (finite-difference) methods. Moreover, iterative regimes using such an approximation method are expected to suffer from a sequentially degraded estimates of $\mathbf{J}$ at each iteration. In non-linear inverse problems, such as elasticity imaging, this effect is also non-linear as the derivative of the forward operator is clearly non-linear in the parameter field. Since this non-linearity increases proportional to p, we may surmise that the sensitivity of inverse solutions to the accuracy of $\mathbf{J}$ is also directly proportional to p.

Point (ii) is subtle and is rooted in the conditioning of $\mathbf{J}$. Indeed, upon inspection of Eq. 12, it is clear that that the more ill-conditioned $\mathbf{J}$ is, the more poorly the inverse problem behaves[2]. Because $\mathbf{J}$ is computed using an hp-forward model of choice, the ill-conditioning of $\mathbf{J}$ is, of course, influenced by the hp-discretization. It is, however, presently unclear how hp-discretizations affect the CN of $\mathbf{J}$. To quantify this, we begin by writing the expression for the condition number of $\mathbf{J}$ as

$$\mathrm{CN}(\mathbf{J}) = ||\mathbf{J}||||\mathbf{J}^{-1}|| \tag{30}$$

however, since $\mathbf{J}$ is usually a non-square matrix, the following relation of $\mathbf{J}^{-1}$ with the pseudo inverse is useful: $\mathbf{J}^{-1} \equiv \mathbf{J}^{\dagger} = (\mathbf{J}^{\mathrm{T}}\mathbf{J})^{-1}\mathbf{J}^{\mathrm{T}}$. Moreover, we may also write an alternative definition for $\mathrm{CN}(\mathbf{J})$, where

---

[2]The influence of the CN of $\mathbf{J}$ is also immediately apparent by observing the linear elasticity imaging inverse problem $\mathbf{u}_m = \mathbf{J}\mathbf{E}$.

we simply take the ratio of the largest eigenvalue $\Lambda_{\max}$ to the smallest eigenvalue $\Lambda_{\min}$, i.e. $\mathrm{CN}(\mathbf{J}) = \frac{\Lambda_{\max}}{\Lambda_{\min}}$ [109]. For a given elasticity imaging problem, one can compute $\mathrm{CN}(\mathbf{J})$ readily to determine the influence of the hp-discretization. To demonstrate this, we plot $\mathrm{CN}(\mathbf{J})$ (using both definitions for the CN) for the first iteration of an inclusion localization problem considering T3, T6, Q4, Q8, and Q9 hp-discretizations with a five different of degrees of freedom in Fig. 9.
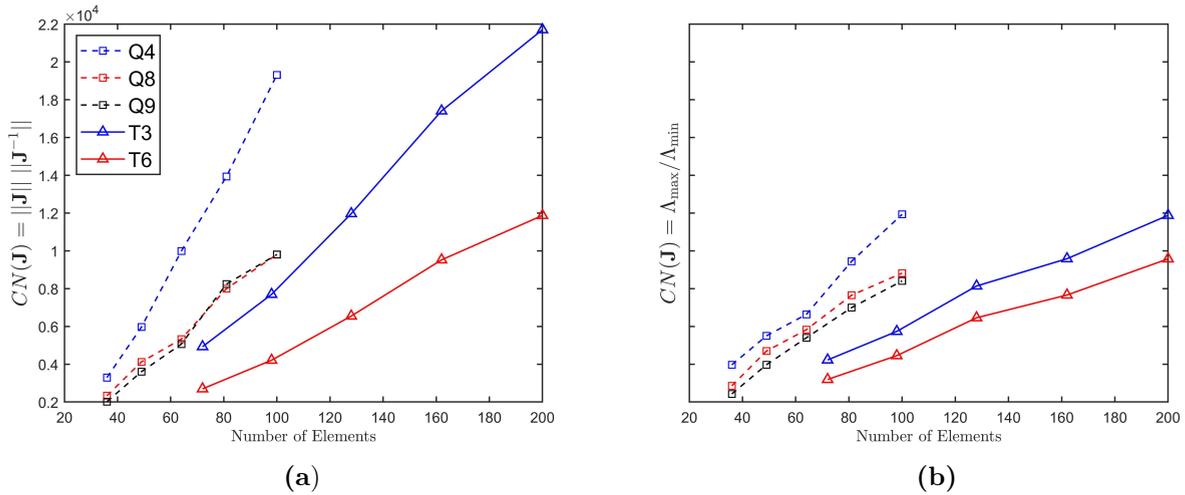
Fig. 9. Condition numbers (CN) of the Jacobian matrix considering a range of degrees of freedom and 5 hp-discretizations using (a) the pseudo inverse definition for the CN and (b) the eigenvalue ratio form of the CN.

We begin the analysis of the $\mathrm{CN}(\mathbf{J})$ results by remarking that the trends shown in Fig. 9a and Fig. 9b, using different definitions for the CN, are similar and comparable for the analysis purposes in this section. Important takeaways from Fig. 9 are that (i) for a fixed number of degrees of freedom, the $\mathrm{CN}(\mathbf{J})$ is not constant as a function of element type and (ii) the $\mathrm{CN}(\mathbf{J})$ increases proportional to the degrees of freedom in the inverse problem. Interestingly, the $\mathrm{CN}(\mathbf{J})$ is consistently higher for discretizations using quadrilateral elements compared to discretizations using triangular elements. This means that the use of quadrilateral elements, in this example, results in a more ill-conditioned Jacobian and, for a fixed prior model, a more ill-posed inverse problem. This realization is also counterintuitive, as one may intuitively believe that higher-order hp-elements should result in lower $\mathrm{CN}(\mathbf{J})$. In reality, the ill-conditioning of $\mathbf{J}$ is proportional to the linear dependence of its columns – not necessarily the hp-element type alone. In the case of hp-discretizations for elasticity problems, the $\mathrm{CN}(\mathbf{J})$ may result from the ill-conditioning of the stiffness matrix $\mathbf{K}$ [110]. Proving this, however, is out of the scope of this article; for this work, it is sufficient to note that different hp-discretizations result in varying conditions of $\mathbf{J}$.

To summarize this subsection, we have discussed the effects of three important variables contributing to the accuracy of elasticity imaging solutions. These variables included prior models, the impact of hp-discretizatized forward model accuracy, and the influence of the accuracy/conditioning of the Jacobian matrix. To this end, each variable influences the sensitivity of the the inverse solution to input parameters

(in this case **E**) in a highly non-linear way. Therefore, the performance of the elasticity imaging inverse problem can be attributed to the cumulative effects of all variables.

### 5.3. Numerical results: hydraulic tomography

In this subsection, we report the results of the hydraulic tomography study. In this study, we analyze the effects of hp discretizations, data size (number of wells), regularization, and the method of computing **J** on hydraulic tomography solutions considering randomized target distributions of hydraulic conductivity. In total, 120000 reconstructions were computed considering 5 hp elements, 5 well grids, two noise levels, and two different methods for computing **J** – a summary of the study variables is provided in Table 2.

To set the stage for more detailed investigation, we begin by presenting images for cases where reconstructions were computed using 25 and 441 wells (the least and most considered herein, respectively), 1% noise, complex differentiation, and Tikhonov regularization with L-curve analysis. The images for cases using 25 and 441 wells are shown in Fig. 10a and 10b, respectively; the columns of these figures represent mesh density while the rows represent the elements used in the forward model. We note that the true image and well distributions are provided in Fig. 3 for the readers' reference.

As a whole, results in Fig. 10 demonstrate that reconstruction quality, in this case the ability to capture large fluctuations in **S**, is most largely influenced by degrees of freedom in the inverse problems N_d. Indeed, despite the hp element type used or the number of wells, images having low N_d do not well capture large fluctuations in **S** and appear relatively homogeneous in comparison to reconstructions with high N_d. It is also interesting to note that reconstructions using triangular discretizations are visually smoother when less data (25 wells) was used than when more data was used (441 wells). Reconstructions in Fig. 10b using T3 and T6 elements have significant grainy features, which are consistent with discretization and interpolation errors. This observation is supported by the realization that the regularity of Q4/Q8/Q9 element nodes, for which the well data is interpolated onto, is more consistent with the Q16 simulation meshes than T3/T6 discretizations.

Next, we aim to determine the effect(s) that the reduced accuracy Jacobian estimation has on reconstruction quality for the same scenario just presented. To explore this visually, **S** reconstructions using Broyden's method for computing **J** considering the same randomized distributions and parameters as the previous example are shown in Fig. 11. Upon first inspection of Fig. 10, it is immediately apparent that the use of Broyden's method significantly degraded image quality relative to reconstructions using complex differentiation. Visually, the increase in reconstruction artifacts and misrepresentation of **S** appears to be amplified in comparison elasticity imaging using Broyden's method. The likely culprit for this realization is the relatively higher ill-conditioned nature of hydraulic tomography (compared with elasticity imaging) resulting from the fact that hydraulic tomography is an underdetermined inverse problem. Moreover, we note that the discretization and interpolation errors reported for reconstructions using complex differentiation are again present in triangular discretizations using 441 well data points. This indicates that these

systematic errors are of sufficient magnitude to not be displaced by reconstruction errors resulting from a highly approximative **J**.
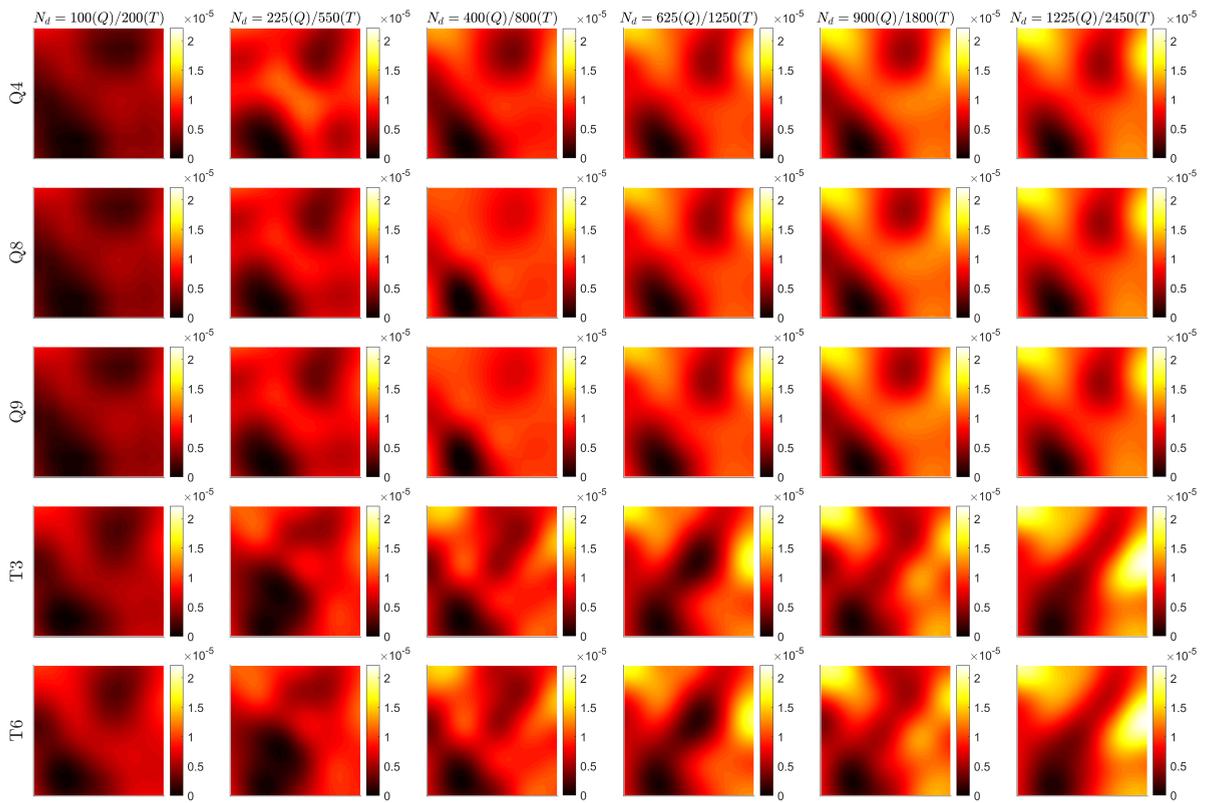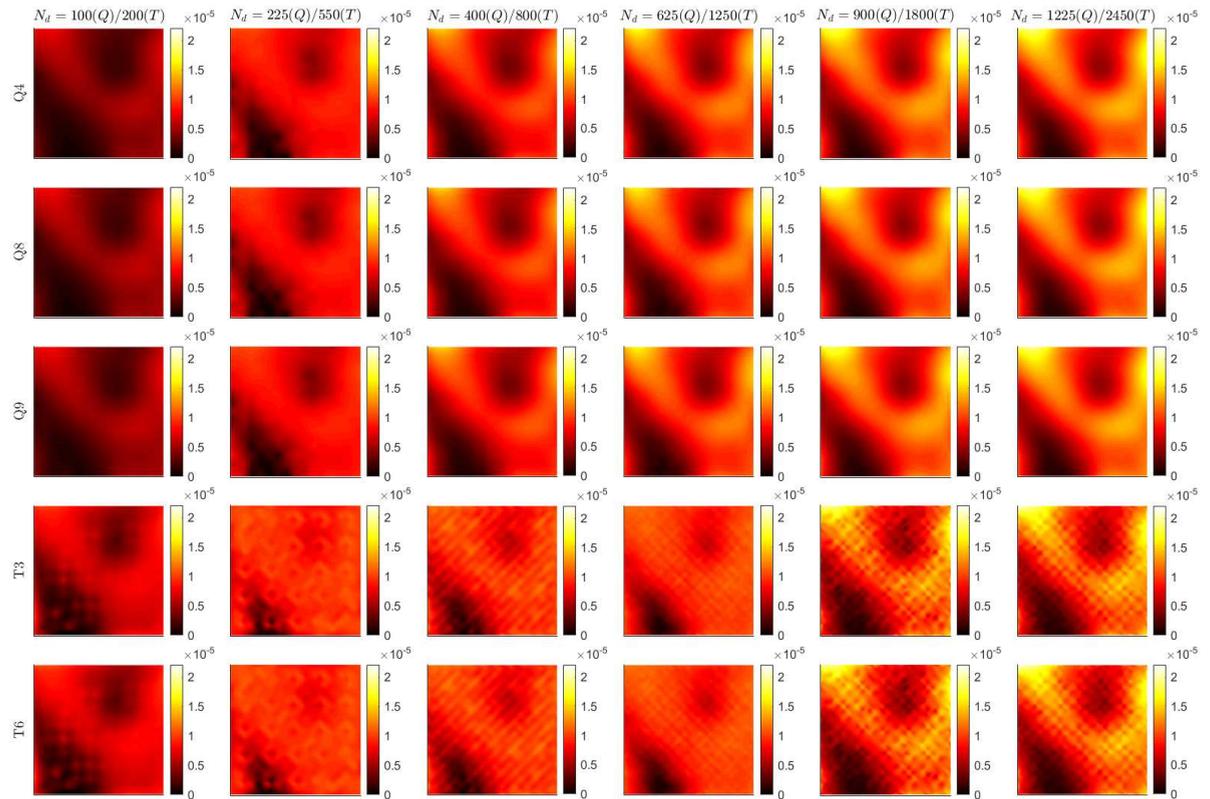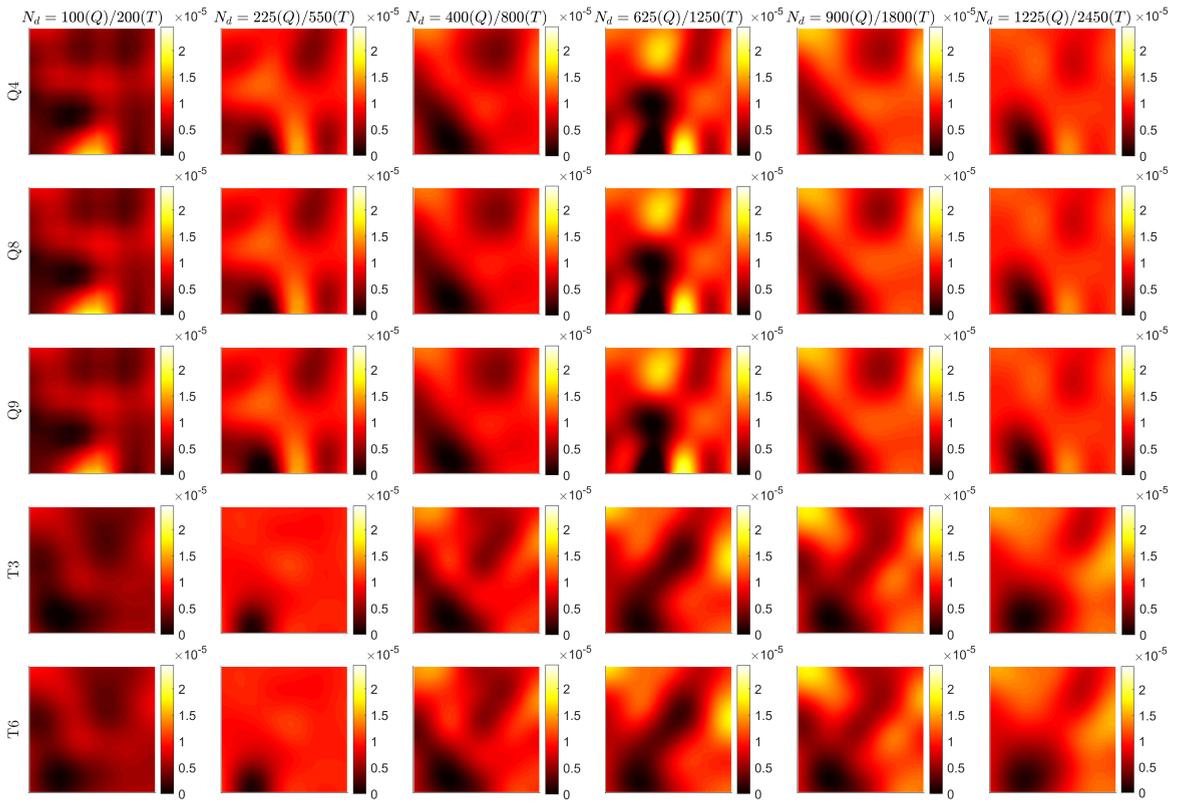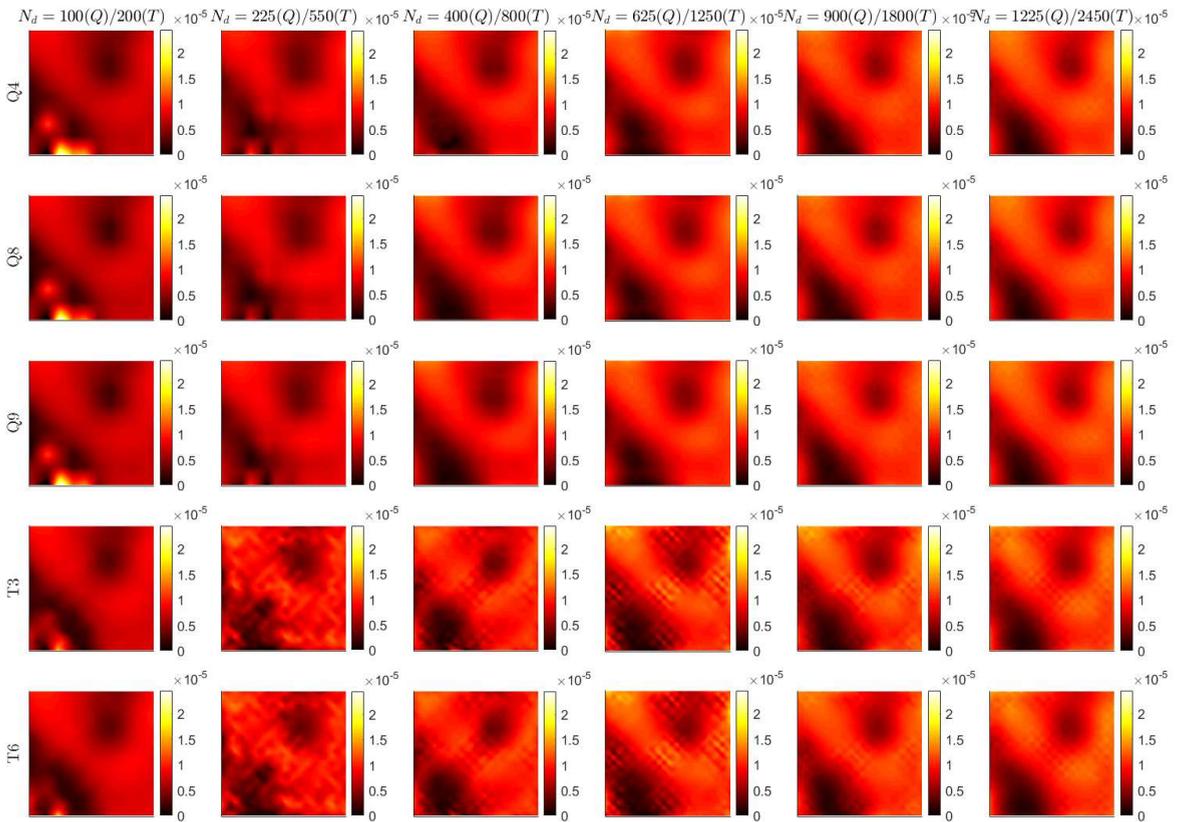
Fig. 10. Hydraulic tomography reconstructions of a randomized hydraulic conductivity distribution considering data points from **(a)** 25 well locations and **(b)** 441 well locations. Reconstructions were computed using 1% added noise, complex differentiation, and Tikhonov regularization. The columns represent the number of quadrilateral **(Q)** and triangular elements **(T)** used in the forward models while the rows indicate the element type used. The colorbar denotes the hydraulic conductivity S.

Fig. 11. Hydraulic tomography reconstructions of a randomized hydraulic conductivity distribution considering data points from (a) 25 well locations and (b) 441 well locations. Reconstructions were computed using 1% added noise, Broyden's method, and Tikhonov regularization. The columns represent the number of quadrilateral (Q) and triangular elements (T) used in the forward models while the rows indicate the element type used. The colorbar denotes the hydraulic conductivity S.

To this point, the analysis of the hydraulic tomography numerical results have been visual and therefore qualitative. In order to investigate these results quantitatively, we again plot the NRMSEs and corresponding SEs against the inverse degrees of freedom. We begin with the results computed using complex differentiation, which are reported in Fig. 12 where each data point is the mean of 100 reconstructions alongside the SE bar. Upon first inspection, we notice that NRMSE generally decreases proportional to increases in inverse degrees of freedom and the number of wells. These observations are expected since (a) the number of data points is inversely related to the ill posedness of the problem (i.e. more data points, less ill posed) and (b) when mesh density increases, interpolation and discretization errors are decreased.

On the other hand, there are some unexpected results shown in Fig. 12. Namely, (i) the relatively higher NRMSEs of regimes using quadrilateral elements in reconstructions using Tikhonov regularization (contrary to the visual observations) and (ii) the consistently small standard errors (SE) in all reconstructions. Regarding (i), while we observe consistently lower NRMSEs in reconstructions using quadrilateral elements and the weighted prior, the opposite is true in reconstructions using Tikhonov regularization. This, of course, reinforces the significant influence regularization has on inverse solutions, but also that there always exists some relative weighting between the accuracy of the hp-forward model and prior information. In this example, the overall performance of regimes using quadrilateral elements and a less-informative prior (Tikhonov) was not robust. In fact, in some cases, increasing the mesh density resulted in an increase in the NRMSE, which is certainly counterintuitive, since discretization errors should decrease as uniform meshes increase in density.

With respect to point (ii), it is not immediately apparent why the SEs in all reconstructions were quite low, for example in comparison to those observed in the elasticity imaging results (Figs. 7 and 8). It should certainly be expected that the sensitivities of an inverse regime should vary depending on the problem's physics and hp discretization. However, it is unexpected that the hydraulic tomography regimes were, on average, relatively insensitive to differing target $\mathbf{S}$ distributions. Upon closer inspection, we do, however, observe an expected trend: that this sensitivity decreases as degrees of freedom increase. Such a result is anticipated since discretization errors are more prevalent in the coarser meshes and the accuracy of $\mathbf{J}$ is lower compared to finer discretizations. To more closely investigate the effects approximating $\mathbf{J}$ on hydraulic tomography reconstructions, we now consider regimes using Broyden's method. These results, analogous to those previously reported in Fig. 12 are provided in Fig. 13.

From Fig. 13, we immediately observe similar trends – with respect to mesh density and the number of wells – to those observed in regimes using complex differentiation. However, we do note an overall increase in NRMSEs in reconstructions using Broyden's method relative to reconstructions using complex differentiation, which is expected due to the lower accuracy in computing $\mathbf{J}$. This quantitative observation is consistent with the visual comment made earlier regarding the obviously degraded quality of images in Fig. 11. We also observe overall closer performance of regimes using triangular and quadrilateral discretizations than when complex differentiation was used. This indicates that, despite the regularization choice used

in this subsection, the use of Broyden's method had an "averaging" effect on the hydraulic tomography regimes' performance. This observation was true for all levels of well data (25 - 441 points).

To briefly sum up this subsection, we observed that the performance of the hydraulic tomography inverse regimes was dependent on a number of variables. In particular, the mesh density, element type, prior model, and accuracy of the Jacobian were all important contributors to the NRMSE and image quality. Despite the method used for computing $\mathbf{J}$, it was found that the prior model type had a significant role in determining whether regimes using triangular or quadrilateral elements performed better. In fact, when weighted regularization was employed, quadrilateral elements exhibited superior performance to triangular elements and vise verse when Tikhonov regularization was used. With respect to the influence of $\mathbf{J}$: it was found that the lower accuracy method for computing $\mathbf{J}$, Broyden's method, had an averaging effect in regards to the NRMSEs of the regimes. As a whole, it can be concluded that the interplay between all aforementioned variables results in conditions whereby it is impossible to state beforehand (i.e. without having already conducted a large suite of reconstructions) that a hydraulic tomography regime employing a given hp-discretized forward model will out perform another hydraulic tomography regime employing a different hp-discretized forward model.
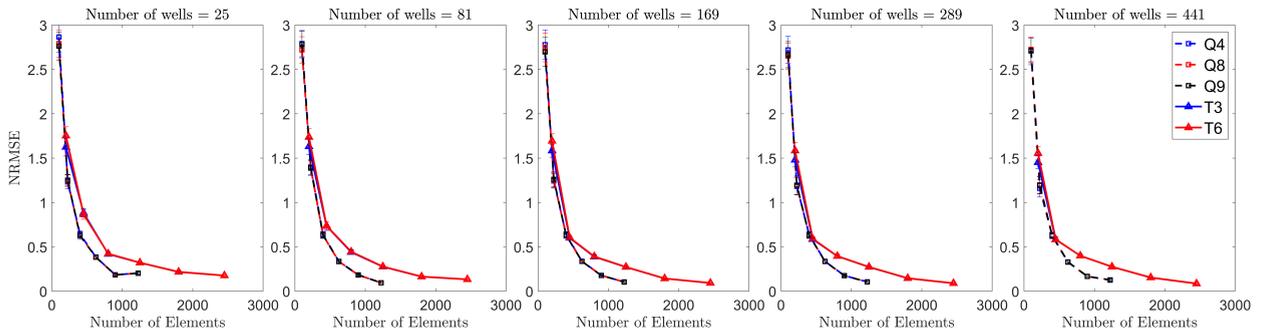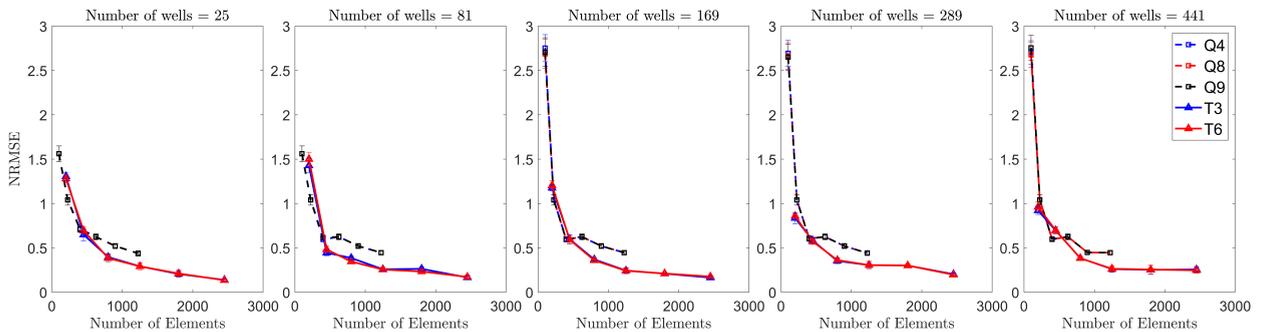
(a) complex differentiation, weighted prior, and 1% noise
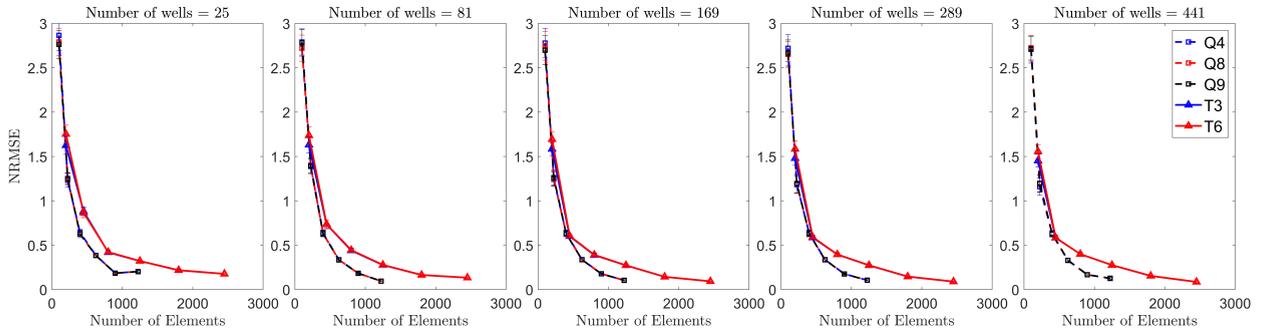
**(b) complex differentiation, weighted prior, and 5% noise**
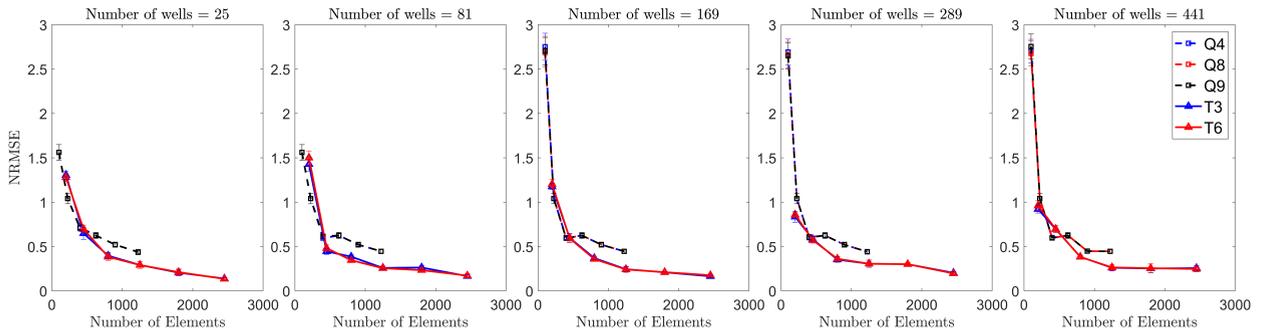
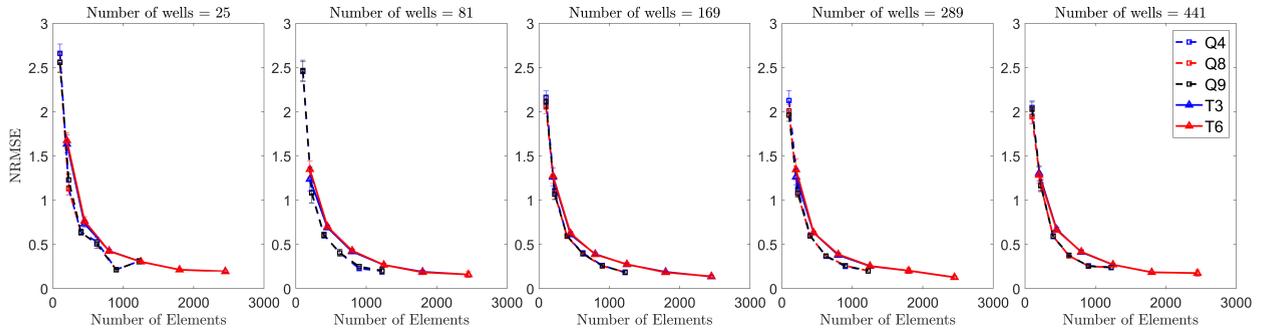**(c) complex differentiation, Tikhonov prior, and 1% noise**

(d) complex differentiation, Tikhonov prior, and 5% noise

Fig. 12. NRMSE plotted against the number of inverse degrees of freedoms for reconstructions of randomized hydraulic conductivity distributions. The columns represent reconstructions using data with 1 and 5% added noise from 25, 81, 169, 289, and 441 wells (from left to right, respectively). Each data point is the mean of 100 reconstructions with the corresponding standard error bars. All reconstructions use complex differentiation for computing J, rows (a)/(b) use a weighted prior model, and rows (c)/(d) use Tikhonov regularization.
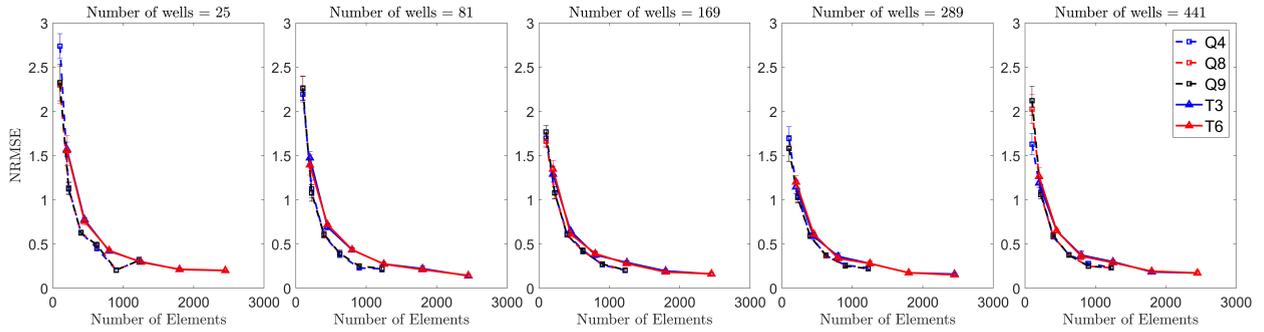
(a) Broyden's method, weighted prior, and 1% noise

(b) Broyden's method, weighted prior, and 5% noise

(c) Broyden's method, Tikhonov prior, and 1% noise
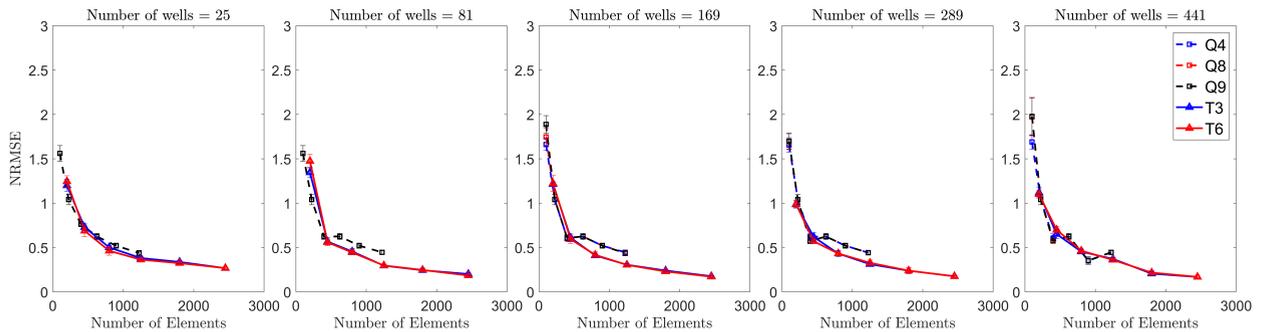

(d) Broyden's method, Tikhonov prior, and 5% noise

Fig. 13. NRMSE plotted against the number of inverse degrees of freedoms for reconstructions of randomized hydraulic conductivity distributions. The columns represent reconstructions using data with 1 and 5% added noise from 25, 81, 169, 289, and 441 wells (from left to right, respectively). Each data point is the mean of 100 reconstructions with the corresponding standard error bars. All reconstructions use Broyden's method for computing J, rows (a)/(b) use a weighted prior model, and rows (c)/(d) use Tikhonov regularization.
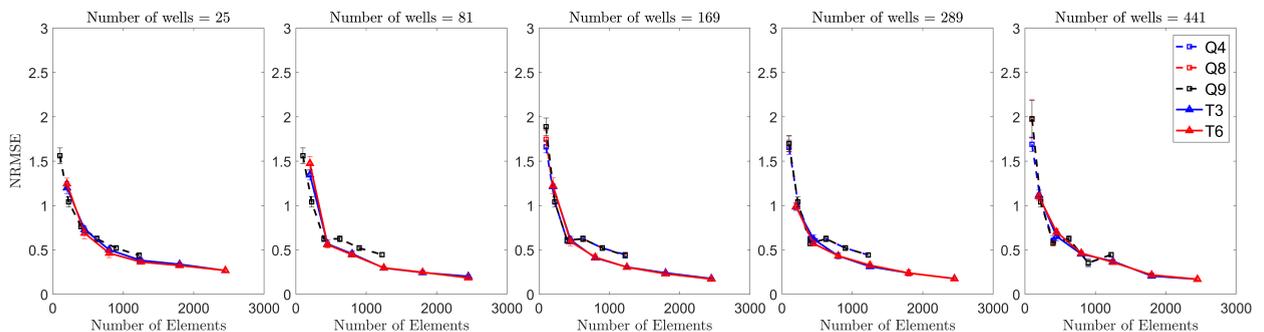
6. Discussion: theoretical considerations

To this point, the article has considered empirical aspects of computational inverse problems solved using hp-forward solvers. As such, some theoretical aspects remain to be discussed. The primary theoretical question arises from the point made in Section 5.2, whereby the importance in matching the modeling errors of the hp-forward operator U and the data errors was noted. In such a discussion, one can begin by investigating the fit between some data d and the forward operator as a function of a estimated parameter θ by writing a functional of the form

$$\Psi_{\text{fit}} = ||d - U(\theta)|| \tag{31}$$

where one would, ideally, want to find a solution approaching the noise amplitude ξ (or, put differently, ξ is the norm of the noise amplitudes) using

$$||d - U(\theta)|| = \xi. \tag{32}$$

Based on the claim in Section 5.2 from [108], we hypothesize that a forward operator with higher accuracy than ξ will not improve reconstruction quality or data fit and, in some cases, degrade the data fit. To investigate this in the context of the inverse problems studied herein, we begin by substituting the appropriate elasticity imaging and hydraulic tomography terms in Eq. 32 resulting in the norms

$$||\mathbf{u}_m - U(\mathbf{E})|| = \xi_{\text{EI}} \text{ and } ||\mathbf{p}_m - U(\mathbf{S})|| = \xi_{\text{S}}, \tag{33}$$

respectively. We remark that, in general, $\xi_{\text{EI}} \neq \xi_{\text{S}}$ since the data sizes and resulting norms of the noise amplitudes are not necessary equivalent. Nonetheless, we may compare the data fidelities (Eq. 33) from reconstructions using various hp discretizations to one another for a fixed noise level to investigate our hypothesis. To do this, we provide data fits for the elasticity imaging ellipsoidal inclusion reconstructions corresponding to 5% noise using Tikhonov regularization and complex differentiation (bottom right corner of Fig. 7. These reconstructions were chosen because they most clearly demonstrated the poor performance of elasticity imaging using high-order hp forward models compared to elasticity imaging using lower-order models. In the case of hydraulic tomography, we selected reconstructions computed with Tikhonov regularization, an intermediate number of well data points (169), complex differentiation, and 5% noise (Fig. 12d, middle column). This selection was made since it clearly showed superior performance of the inversion regimes using lower-order hp models and the reconstructions did not suffer from localized effects resulting from discretization and interpolation errors. Using averaged results at each hp discretization level, the data fits computed using Eq. 33 (and subtracting the Tikhonov term from the regularized cost functionals) are provided in Fig. 14. Each data point represents an average of 100 reconstructions and both plots result from a total of 6,000 reconstructions.
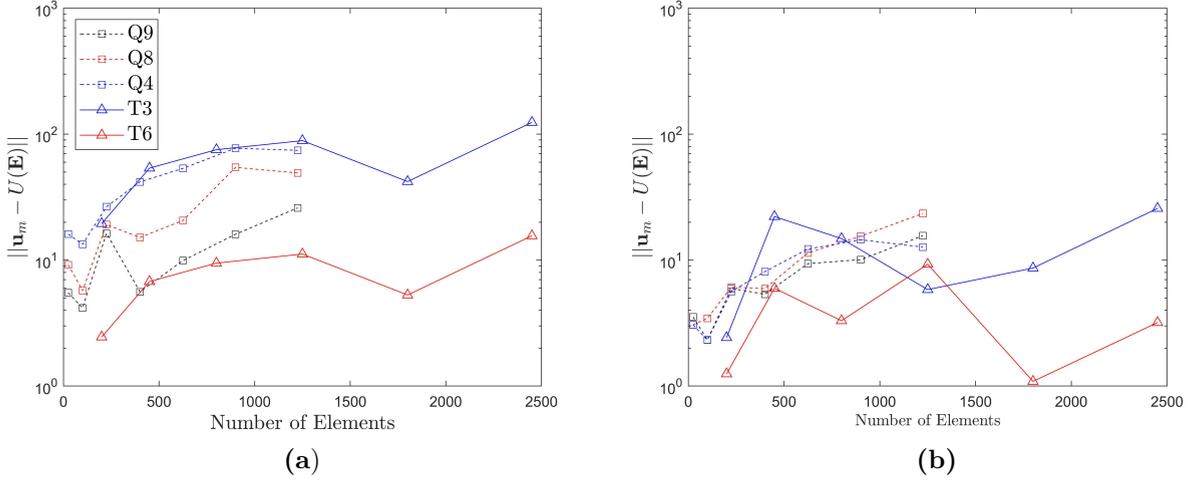
Fig. 14. Data fits as a function of the number of elements (x-axis) and element types for (a) elasticity imaging reconstructions of an ellipsoidal inclusion with 5% noise using Tikhonov regularization and complex differentiation and (b) hydraulic tomography reconstructions of a randomized blob-like distribution computed with Tikhonov regularization, complex differentiation, an intermediate number of wells (169), and 5% noise.

Results from Fig. 14 support the hypothesis and clearly demonstrate that, for a fixed h (x-axis), increasing the order of p does not guarantee an improvement in the data fit. Indeed, in general, reconstructions using T6 elements showed a better fit than all other elements and in some cases, reconstructions using Q4 and Q8 elements resulted in a better fit than the Q9 elements. In addition to this, we would like to note that the increasing trend in the data fit norms is mostly due to the increasing size of the data vectors; therefore the comparison of data fit norms between differing h levels is not particularly meaningful. Inasmuch, while results previously reported in Figs. 7 and Fig. 12d already demonstrated that the test case reconstructions using lower-order hp-discretizations often had lower NRMSEs than the reconstructions using higher-order hp-discretizations, results in Fig. 14 also demonstrate that the data fits may similarly be degraded by using an overly accurate hp-forward model. This realization also provides an logical explanation for the general stagnation and/or fluctuation of NRMSEs in reconstructions using fine meshes at high noise levels.

To summarize this discussion, the results herein demonstrate the importance in (a) selecting the accuracy of the forward solver based on ξ – *in these examples, T6 elements would have been a good choice* and (b) over-discretizing an ill-posed problem is counter-productive. These points are also relevant in the numerical reconstruction regimes used to solve applied inverse problems. For example, in computing a minimizer (cf. Eq. 12), one would want to utilize a solver to the penalty function also reflecting an accuracy of ξ which, for least-squares based approaches, would be reflected in the computation of the Jacobian.

7. Concluding remarks

In this work, we investigated the use of hp-forward models in applied inverse problems. To do this, we chose to study two imaging modalities with different physics and data characteristics. For this, least-squares based elasticity imaging (overdetermined) and hydraulic tomography (underdetermined) inverse

regimes were selected. In a large suite of numerical tests, we quantitatively examined the performance of elasticity imaging and hydraulic tomography regimes in answering the open research question, "Does the use of higher order hp-forward models in applied inverse problems always result in lower error reconstructions than approaches using lower order hp-forward models?"

Through numerous counterexamples, this question can be definitively answered: **no**. In other words, decreasing h (increasing mesh density), increasing p, or simultaneously decreasing h and increasing p does not always guarantee lower error reconstructions in applied inverse problems. Indeed, it was found that there exists a complex interplay between several key variables which prevents a guarantee that an inverse regime employing a higher-order hp-forward model will always out-perform another inverse regime employing a lower-order hp-forward model. These realizations were further explored in a theoretical discussion where it was demonstrated that the selection of a forward model and minimizer should (optimally) be done considering the noise amplitude, ξ. In certain conditions, combinations of these variables may lead to inconsistent or unexpected behavior of the inverse regime employing an hp-forward model. The key variables included (in no particular order):

- The discrepancy between the forward model accuracy and ξ.

- The prior model.

- How underdetermined or overdetermined the problem is.

- The accuracy of the Jacobian.

- The condition number of the Jacobian.

- Finite element mesh density.

- Measurement noise.

- Modeling errors.

- Discretization errors.

- Degrees of freedom in the inverse problem.

A critical takeaway from this work is that employing high-order hp-forward models in applied inverse problems is not always the best choice. More precisely, over-discretizing an ill-posed problem is counter-productive. In general, and consistent with the findings of a recent study [108], the most suitable compromise may be to match the accuracy of the forward model and the data.

Acknowledgments

DS would like to acknowledge the support of the Department of Civil and Structural Engineering at the University of Sheffield. DL was supported by the National Natural Science Foundation of China under Grant No. 61871356 and Anhui Provincial Natural Science foundation under Grant 1708085MA25. Both authors would like to acknowledge the comments from the reviewers, which significantly improved the technical content of this article.

References

[1] J. L. Mueller, S. Siltanen, Linear and nonlinear inverse problems with practical applications, SIAM, 2012.

[2] Y. Chen, J. A. O'Sullivan, D. G. Politte, J. D. Evans, D. Han, B. R. Whiting, J. F. Williamson, Line integral alternating minimization algorithm for dual-energy x-ray ct image reconstruction, IEEE transactions on medical imaging 35 (2016) 685–698.

[3] K. Hämäläinen, L. Harhanen, A. Hauptmann, A. Kallonen, E. Niemi, S. Siltanen, Total variation regularization for large-scale x-ray tomography, International Journal of Tomography & Simulation 25 (2014) 1–25.

[4] E. Maire, P. J. Withers, Quantitative x-ray tomography, International materials reviews 59 (2014) 1–43.

[5] D. Liu, D. Smyl, J. Du, A parametric level set-based approach to difference imaging in electrical impedance tomography, IEEE transactions on medical imaging 38 (2019) 145–155.

[6] D. Liu, D. Gu, D. Smyl, J. Deng, J. Du, B-spline based sharp feature preserving shape reconstruction approach for electrical impedance tomography, IEEE transactions on medical imaging (2019, in press).

[7] C. J. Trepte, C. R. Phillips, J. Solà, A. Adler, S. A. Haas, M. Rapin, S. H. Böhm, D. A. Reuter, Electrical impedance tomography (eit) for quantification of pulmonary edema in acute lung injury, Critical care 20 (2015) 18.

[8] P. Vauhkonen, Image reconstruction in three-dimensional electrical impedance tomography, Ph.D. thesis, University of Kuopio, 2004.

[9] B. Brown, Electrical impedance tomography (EIT): a review, J. Med. Eng. Technol. 27 (2003) 97–108.

[10] J. Kaipio, V. Kolehmainen, E. Somersalo, M. Vauhkonen, Statistical inversion and Monte Carlo sampling methods in EIT, Inverse Probl. 16 (2000) 1487–1522.

[11] T. Saratoon, T. Tarvainen, B. Cox, S. Arridge, A gradient-based method for quantitative photoacoustic tomography using the radiative transfer equation, Inverse Problems 29 (2013) 075006.

[12] K. Maslov, H. F. Zhang, S. Hu, L. V. Wang, Optical-resolution photoacoustic microscopy for in vivo imaging of single capillaries, Optics letters 33 (2008) 929–931.

[13] S. Arridge, J. Kaipio, V. Kolehmainen, M. Schweiger, E. Somersalo, T. Tarvainen, M. Vauhkonen, Approximation errors and model reduction with an application in optical diffusion tomography, Inverse problems 22 (2006) 175.

[14] Y. Zhang, Z. Dong, P. Phillips, S. Wang, G. Ji, J. Yang, Exponential wavelet iterative shrinkage thresholding algorithm for compressed sensing magnetic resonance imaging, Information Sciences 322 (2015) 115–132.

[15] H. Ammari, H. Kwon, Y. Lee, K. Kang, J. K. Seo, Magnetic resonance-based reconstruction method of conductivity and permittivity distributions at the larmor frequency, Inverse Problems 31 (2015) 105001.

[16] S. Goenezen, P. Barbone, A. A. Oberai, Solution of the nonlinear elasticity imaging inverse problem: The incompressible case, Computer Methods in Applied Mechanics and Engineering 200 (2011) 1406–1420.

[17] N. H. Gokhale, P. E. Barbone, A. A. Oberai, Solution of the nonlinear elasticity imaging inverse problem: the compressible case, Inverse Problems 24 (2008) 045010.

[18] A. A. Oberai, N. H. Gokhale, M. M. Doyley, J. C. Bamber, Evaluation of the adjoint equation based algorithm for elasticity imaging, Physics in Medicine and Biology 49 (2004) 2955.

[19] T. Abdelsalhin, A. Maselli, V. Ferrari, Solving the relativistic inverse stellar problem through gravitational waves observation of binary neutron stars, Physical Review D 97 (2018) 084014.

[20] A. Portal, P. Labazuy, J.-F. Lénat, S. Béné, P. Boivin, E. Busato, C. Cârloganu, C. Combaret, P. Dupieux, F. Fehr, et al., Inner structure of the puy de dôme volcano: cross-comparison of geophysical models (ert, gravimetry, muon imaging), Geoscientific Instrumentation, Methods and Data Systems 2 (2013) 47–54.

[21] O. Schueler-Furman, C. Wang, P. Bradley, K. Misura, D. Baker, Progress in modeling of protein structures and interactions, science 310 (2005) 638–642.

[22] D. Smyl, M. Pour-Ghaz, A. Seppänen, Can electrical imaging be used for complex structural cracking patterns?, Computer Methods in Applied Mechanics and Engineering (under review) (2017).

[23] T. Tallman, S. Gungor, G. Koo, C. Bakis, On the inverse determination of displacements, strains, and stresses in a carbon nanofiber/polyurethane nanocomposite from conductivity data obtained via electrical impedance tomography, Journal of Intelligent Material Systems and Structures (2017) 1–13.

[24] J. He, F.-G. Yuan, Lamb wave-based subwavelength damage imaging using the dort-music technique in metallic plates, Structural Health Monitoring 15 (2016) 65–80.

[25] D. Bull, L. Helfen, I. Sinclair, S. Spearing, T. Baumbach, A comparison of multi-scale 3D X-ray tomographic inspection techniques for assessing carbon fibre composite impact damage, Composites Science and Technology 75 (2013) 55–61.

[26] A. K. Saibaba, T. Bakhos, P. K. Kitanidis, A flexible krylov solver for shifted systems with application to oscillatory hydraulic tomography, SIAM Journal on Scientific Computing 35 (2013) A3001–A3023.

[27] W. A. Illman, X. Liu, S. Takeuchi, T.-C. J. Yeh, K. Ando, H. Saegusa, Hydraulic tomography in fractured granite: Mizunami underground research site, japan, Water Resources Research 45 (2009).

[28] J. Zhu, T.-C. J. Yeh, Characterization of aquifer heterogeneity using transient hydraulic tomography, Water Resources Research 41 (2005).

[29] D. Liu, Y. Zhao, A. K. Khambampati, A. Seppanen, J. Du, A parametric level set method for imaging multi-phase conductivity using electrical impedance tomography, IEEE Transactions on Computational Imaging 4 (2018) 552–561.

[30] I. Ismail, J. Gamio, S. A. Bukhari, W. Yang, Tomography for multi-phase flow measurement in the oil industry, Flow Measurement and Instrumentation 16 (2005) 145–155.

[31] D. Smyl, K.-N. Antin, D. Liu, S. Bossuyt, Coupled digital image correlation and quasi-static elasticity imaging of inhomogeneous orthotropic composite structures, Inverse Problems 34 (2018) 124005.

[32] M. Grediac, The use of full-field measurement methods in composite material characterization: interest and limitations, Composites Part A: applied science and manufacturing 35 (2004) 751–761.

[33] D. S. Oliver, A. C. Reynolds, N. Liu, Inverse theory for petroleum reservoir characterization and history matching, Cambridge University Press, 2008.

[34] A. A. Oberai, N. H. Gokhale, G. R. Feijóo, Solution of inverse problems in elasticity imaging using the adjoint method, Inverse Problems 19 (2003) 297.

[35] J. Kaipio, E. Somersalo, Statistical and computational inverse problems, volume 160, Springer Science & Business Media, 2006.

[36] H.-B. An, J. Wen, T. Feng, On finite difference approximation of a matrix-vector product in the jacobian-free newton–krylov method, Journal of Computational and Applied Mathematics 236 (2011) 1399 – 1409.

[37] J. Carrera, A. Alcolea, A. Medina, J. Hidalgo, L. J. Slooten, Inverse problem in hydrogeology, Hydrogeology journal 13 (2005) 206–222.

[38] J. C. de Munck, T. J. Faes, R. M. Heethaar, The boundary element method in the forward and inverse problem of electrical impedance tomography, IEEE transactions on Biomedical Engineering 47 (2000) 792–800.

[39] I. Harari, A survey of finite element methods for time-harmonic acoustics, Computer methods in applied mechanics and engineering 195 (2006) 1594–1607.

[40] H. Jasak, A. Jemcov, Z. Tukovic, et al., Openfoam: A c++ library for complex physics simulations, in: International workshop on coupled methods in numerical dynamics, volume 1000, IUC Dubrovnik, Croatia, 2007, pp. 1–20.

[41] K. S. Surana, J. Reddy, The Finite Element Method for Boundary Value Problems: Mathematics and Computations, CRC Press, 2016.

[42] Q. Lin, J. Lin, Finite element methods: accuracy and improvement, volume 1, Elsevier, 2007.

[43] P. Binev, W. Dahmen, R. DeVore, Adaptive finite element methods with convergence rates, Numerische Mathematik 97 (2004) 219–268.

[44] P. Morin, R. H. Nochetto, K. G. Siebert, Convergence of adaptive finite element methods, SIAM review 44 (2002) 631–658.

[45] C. Geuzaine, J.-F. Remacle, Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities, International journal for numerical methods in engineering 79 (2009) 1309–1331.

[46] T. J. Hughes, The finite element method: linear static and dynamic finite element analysis, Courier Corporation, 2012.

[47] C. Johnson, Numerical solution of partial differential equations by the finite element method, Courier Corporation, 2012.

[48] I. Babuška, U. Banerjee, J. E. Osborn, Survey of meshless and generalized finite element methods: a unified approach, Acta Numerica 12 (2003) 1–125.

[49] S. R. Idelsohn, E. Onate, N. Calvo, F. Del Pin, The meshless finite element method, International Journal for Numerical Methods in Engineering 58 (2003) 893–912.

[50] J. T. Oden, C. Duarte, O. C. Zienkiewicz, A new cloud-based hp finite element method, Computer methods in applied mechanics and engineering 153 (1998) 117–126.

[51] C. A. Duarte, J. T. Oden, An hp adaptive method using clouds, Computer methods in applied mechanics and engineering 139 (1996) 237–262.

[52] G. Liu, K. Dai, T. T. Nguyen, A smoothed finite element method for mechanics problems, Computational Mechanics 39 (2007) 859–877.

[53] M. J. Grote, A. Schneebeli, D. Schötzau, Discontinuous galerkin finite element method for the wave equation, SIAM Journal on Numerical Analysis 44 (2006) 2408–2431.

[54] P. Houston, C. Schwab, E. Süli, Discontinuous hp-finite element methods for advection-diffusion-reaction problems, SIAM Journal on Numerical Analysis 39 (2002) 2133–2163.

[55] M. J. Borden, C. V. Verhoosel, M. A. Scott, T. J. Hughes, C. M. Landis, A phase-field description of dynamic brittle fracture, Computer Methods in Applied Mechanics and Engineering 217 (2012) 77–95.

[56] D. Schrade, R. Mueller, B. Xu, D. Gross, Domain evolution in ferroelectric materials: A continuum phase field model and finite element implementation, Computer methods in applied mechanics and engineering 196 (2007) 4365–4374.

[57] D. Goddeke, S. H. Buijssen, H. Wobker, S. Turek, Gpu acceleration of an unmodified parallel finite element navier-stokes solver, in: High Performance Computing & Simulation, 2009. HPCS'09. International Conference on, IEEE, 2009, pp. 12–21.

[58] M. Paszyński, J. Kurtz, L. Demkowicz, Parallel, fully automatic hp-adaptive 2d finite element package, Computer methods in applied mechanics and engineering 195 (2006) 711–741.

[59] V. Puzyrev, J. Koldan, J. de la Puente, G. Houzeaux, M. Vázquez, J. M. Cela, A parallel finite-element method for three-dimensional controlled-source electromagnetic forward modelling, Geophysical Journal International 193 (2013) 678–693.

[60] I. Babuška, M. Suri, The hp version of the finite element method with quasiuniform meshes, ESAIM: Mathematical Modelling and Numerical Analysis 21 (1987) 199–238.

[61] J. Kaipio, E. Somersalo, Statistical and Computational Inverse Problems, Springer New York, 2005.
[62] P. Ledger, hp-finite element discretisation of the electrical impedance tomography problem, Computer Methods in Applied Mechanics and Engineering 225 (2012) 154–176.
[63] K. Yeo, Y. Hwang, X. Liu, J. Kalagnanam, Development of hp-inverse model by using generalized polynomial chaos, Computer Methods in Applied Mechanics and Engineering 347 (2019) 1–20.
[64] W. Bangerth, A framework for the adaptive finite element solution of large-scale inverse problems, SIAM Journal on Scientific Computing 30 (2008) 2965–2989.
[65] M. Alexe, A. Sandu, Space–time adaptive solution of inverse problems with the discrete adjoint method, Journal of Computational Physics 270 (2014) 21–39.
[66] D. Isaacson, Distinguishability of conductivities by electric current computed tomography, IEEE transactions on medical imaging 5 (1986) 91–95.
[67] D. C. Dobson, Estimates on resolution and stabilization for the linearized inverse conductivity problem, Inverse problems 8 (1992) 71.
[68] M. Cheney, D. Isaacson, J. C. Newell, S. Simske, J. Goble, Noser: An algorithm for solving the inverse conductivity problem, International Journal of Imaging systems and technology 2 (1990) 66–75.
[69] H. R. MacMillan, T. A. Manteuffel, S. F. McCormick, First-order system least squares and electrical impedance tomography, SIAM journal on numerical analysis 42 (2004) 461–483.
[70] R. Winkler, A. Rieder, Resolution-controlled conductivity discretization in electrical impedance tomography, SIAM Journal on Imaging Sciences 7 (2014) 2048–2077.
[71] L. Borcea, V. Druskin, F. G. Vasquez, A. Mamonov, Resistor network approaches to electrical impedance tomography, Inverse Problems and Applications: Inside Out II, Math. Sci. Res. Inst. Publ 60 (2011) 55–118.
[72] H. B. Ameur, B. Kaltenbacher, Regularization of parameter estimation by adaptive discretization using refinement and coarsening indicators, Journal of Inverse and Ill-Posed Problems 10 (2002) 561–583.
[73] P. Mathé, S. V. Pereverzev, Optimal discretization of inverse problems in hilbert scales. regularization and self-regularization of projection methods, SIAM Journal on Numerical Analysis 38 (2001) 1999–2021.
[74] A. Kirsch, An introduction to the mathematical theory of inverse problems, volume 120, Springer Science & Business Media, 2011.
[75] J. Koponen, T. Huttunen, T. Tarvainen, J. P. Kaipio, Bayesian approximation error approach in full-wave ultrasound tomography, IEEE transactions on ultrasonics, ferroelectrics, and frequency control 61 (2014) 1627–1637.
[76] A. Nissinen, L. Heikkinen, J. Kaipio, The bayesian approximation error approach for electrical impedance tomography—experimental results, Measurement Science and Technology 19 (2007) 015501.
[77] T. Tarvainen, A. Pulkkinen, B. T. Cox, J. P. Kaipio, S. R. Arridge, Bayesian image reconstruction in quantitative photoacoustic tomography, IEEE transactions on medical imaging 32 (2013) 2287–2298.
[78] D. Smyl, S. Bossuyt, D. Liu, Stacked elasticity imaging approach for visualizing defects in the presence of background inhomogeneity, Journal of Engineering Mechanics (Submitted, 2018).
[79] M. Bonnet, A. Constantinescu, Inverse problems in elasticity, Inverse Problems 21 (2005) R1.
[80] M. S. Richards, P. E. Barbone, A. A. Oberai, Quantitative three-dimensional elasticity imaging from quasi-static deformation: a phantom study, Physics in Medicine and Biology 54 (2009) 757.
[81] F. Kallel, M. Bertrand, Tissue elasticity reconstruction using linear perturbation method, IEEE Transactions on Medical Imaging 15 (1996) 299–313.
[82] J. Ophir, I. Cespedes, H. Ponnekanti, Y. Yazdi, X. Li, Elastography: a quantitative method for imaging the elasticity of biological tissues, Ultrasonic Imaging 13 (1991) 111–134.
[83] T. Vilhunen, J. Kaipio, P. Vauhkonen, T. Savolainen, M. Vauhkonen, Simultaneous reconstruction of electrode contact impedances and internal electrical properties: I. theory, Measurement Science and Technology 13 (2002) 1848.
[84] J. Martins, P. Sturdza, J. Alonso, The connection between the complex-step derivative approximation and algorithmic differentiation, in: 39th Aerospace Sciences Meeting and Exhibit, 2001, p. 921.
[85] J. R. Martins, P. Sturdza, J. J. Alonso, The complex-step derivative approximation, ACM Transactions on Mathematical Software (TOMS) 29 (2003) 245–262.
[86] C. G. Broyden, A class of methods for solving nonlinear simultaneous equations, Mathematics of computation 19 (1965) 577–593.
[87] M. H. Loke, R. Barker, Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-newton method 1, Geophysical prospecting 44 (1996) 131–152.
[88] E. L. Allgower, K. Georg, Numerical continuation methods: an introduction, volume 13, Springer Science & Business Media, 2012.
[89] R. Battiti, First-and second-order methods for learning: between steepest descent and newton's method, Neural computation 4 (1992) 141–166.
[90] J. Gottlieb, P. Dietrich, Identification of the permeability distribution in soil by hydraulic tomography, Inverse Problems 11 (1995) 353.
[91] M. Cardiff, W. Barrash, 3-d transient hydraulic tomography in unconfined aquifers with fast drainage response, Water Resources Research 47 (2011).
[92] V. Ginting, G. Lin, J. Liu, On application of the weak galerkin finite element method to a two-phase model for subsurface flow, Journal of Scientific Computing 66 (2016) 225–239.
[93] A. Masud, T. J. Hughes, A stabilized mixed finite element method for darcy flow, Computer methods in applied mechanics and engineering 191 (2002) 4341–4370.
[94] J. Liu, F. Sadre-Marandi, Z. Wang, Darcylite: A matlab toolbox for darcy flow computation, Procedia Computer Science 80 (2016) 1301–1312.
[95] P. Vauhkonen, M. Vauhkonen, T. Savolainen, J. Kaipio, Three-dimensional electrical impedance tomography based on

the complete electrode model, IEEE T. Biomedical Eng. 46 (1999) 1150–1160.

[96] M. Benning, M. Burger, Modern regularization methods for inverse problems, Acta Numerica 27 (2018) 1–111.

[97] G. González, V. Kolehmainen, A. Seppänen, Isotropic and anisotropic total variation regularization in electrical impedance tomography, Computers & Mathematics with Applications (2017).

[98] J. P. Kaipio, V. Kolehmainen, M. Vauhkonen, E. Somersalo, Inverse problems with structural prior information, Inverse problems 15 (1999) 713.

[99] M. Lassas, S. Siltanen, Can one use total variation prior for edge-preserving bayesian inversion?, Inverse Problems 20 (2004) 1537.

[100] P. C. Hansen, D. P. O'Leary, The use of the l-curve in the regularization of discrete ill-posed problems, SIAM Journal on Scientific Computing 14 (1993) 1487–1503.

[101] J. Kaipio, E. Somersalo, Statistical inverse problems: discretization, model reduction and inverse crimes, Journal of Computational and Applied Mathematics 198 (2007) 493–504.

[102] P. C. Hansen, The l-curve and its use in the numerical treatment of inverse problems (1999).

[103] T.-C. J. Yeh, S. Liu, Hydraulic tomography: Development of a new aquifer test method, Water Resources Research 36 (2000) 2095–2105.

[104] D. Calvetti, S. Morigi, L. Reichel, F. Sgallari, Tikhonov regularization and the l-curve for large discrete ill-posed problems, Journal of computational and applied mathematics 123 (2000) 423–446.

[105] G. C. Bohling, J. J. Butler Jr, Inherent limitations of hydraulic tomography, Groundwater 48 (2010) 809–824.

[106] W. A. Illman, A. J. Craig, X. Liu, Practical issues in imaging hydraulic conductivity through hydraulic tomography, Groundwater 46 (2008) 120–132.

[107] D. Smyl, S. Bossuyt, W. Ahmad, A. Valilov, D. Liu, An overview of 38 least squares-based frameworks for structural damage tomography, Structural Health Monitoring (under review) (2019).

[108] M. Burger, Y. Korolev, J. Rasch, Convergence rates and structure of solutions of inverse problems with imperfect forward models, Inverse Problems 35 (2019) 024006.

[109] I. Goodfellow, Y. Bengio, A. Courville, Deep learning, MIT press, 2016.

[110] R. Kannan, S. Hendry, N. J. Higham, F. Tisseur, Detecting the causes of ill-conditioning in structural finite element models, Computers & Structures 133 (2014) 79–89.