Highlights

Forward and inverse modelings of fault transmissibility in subsurface flows

Jeonghun J. Lee, Tan Bui-Thanh, Umberto Villa, Omar Ghattas

- A complete investigation of inversion of the fault transmissibility for subsurface flow models under appropriate assumptions on fault structures
- Fault modeling and well-posedness of the forward problems
- Finite element (FEM) discretizations of the forward problem and their rigorous *a priori* convergence analysis
- Well-posedness of the Bayesian inverse problem, FEM discretization of infinite dimensional Bayesian inverse formulation, and its rigorous *a priori* analysis.
- *Maximum a posteriori* computation as deterministic optimization and a Laplace approximation of the Bayesian posterior are also presented.
- Numerical results illustrates the use of the proposed fault model for forward and inverse simulation of subsurface flows in two dimensional domain with multiple faults.

Forward and inverse modelings of fault transmissibility in subsurface flows

Jeonghun J. Lee[a], Tan Bui-Thanh[b,e,1], Umberto Villa[c], Omar Ghattas[d,e]

[a]*Department of Mathematics, Baylor University, Waco, Texas, USA*
[b]*Department of Aerospace Engineering and Engineering Mechanics, UT Austin, Austin, Texas, USA*
[c]*Department of Electrical & Systems Engineering, Washington University in St. Louis, St. Louis, Missouri, USA*
[d]*Department of Geological Sciences, Jackson School of Geosciences, UT Austin, Austin, Texas, USA*
[e]*The Oden Institute for Computational Engineering and Sciences, UT Austin, Austin, Texas, Austin, Texas, USA*

Abstract

Characterizing physical properties of faults, such as their transmissibility, is crucial to perform predictive numerical simulation of subsurface flow problems, such as those encountered in petroleum engineering and remediation of subsurface contamination. This paper provides a complete investigation of inversion of the fault transmissibility for subsurface flow models under appropriate assumptions on fault structures. In particular, the following aspects are considered: 1) fault modeling and well-posedness of the forward problems; 2) finite element (FEM) discretizations of the forward problem and their rigorous *a priori* convergence analysis; 3) Well-posedness of the Bayesian inverse problem, FEM discretization of infinite dimensional Bayesian inverse formulation, and its rigorous *a priori* analysis. *Maximum a posteriori* computation as deterministic optimization and a Laplace approximation of the Bayesian posterior are also presented. Numerical results illustrates the use of the proposed fault model for forward and inverse simulation of subsurface flows in two dimensional domain with multiple faults.

Keywords: subsurface flow, faults transmissibility, inverse problem

Email addresses: jeonghun_lee@baylor.edu (Jeonghun J. Lee), tanbui@ices.utexas.edu (Tan Bui-Thanh), uvilla@wustl.edu (Umberto Villa), omar@ices.utexas.edu (Omar Ghattas)

2000 MSC: 68Q25, 68R10, 68U05

1. Introduction

Accurate modeling and numerical simulation of subsurface flows are important for various applications such as petroleum engineering and contaminant transport. In subsurface flow models the fault (or fracture) structure has a great influence on the fluid flows. Therefore, modeling fault structures with appropriate physical parameters is important for accurate simulations of the subsurface flows when fault structures exist in the subsurface domain (see, e.g., [1, 2, 3, 4] and more references in a recent review article [5]).

Faults in oil reservoirs can be complex geometric structures with relatively thin volumes compared to the whole domain of reservoir models. Therefore, high resolution modeling of faults as volumetric objects requires very fine grid/mesh to describe the full geometric complexity of faults. Such an approach, though providing high fidelity simulations, is at the expense of extremely high computational costs. Moreover, faults' detailed information for high resolution models are usually not available in practical applications. Existing approaches typically model fault structures as a manifold of codimension one in the whole domain. This is meaningful as the thickness of faults is negligibly small compared to the length scale of the whole reservoir model [6, 7, 8]. In this paper, we also adopt this approach for modeling fault structures.

Some previous studies on subsurface flow models with faults consider a geometric multiscale framework in which specifically designed interface conditions are used to couple the flow equations in the whole domain with those in the faults' domain [6, 7]. In this paper, we assume that the permeability parameters on faults are much lower than the surrounding media, so there is no fluid flows along fault structures [9]. As a consequence, we obtain a reduced single-phase flow model that does not need the velocity and the pressure fields on faults as independent unknowns. Moreover, the physical parameters associated to the permeability on faults are reduced to the fault transmissibility fields (see Section 2 and [9]).

In this paper we discuss fault modeling and numerical methods together with their analyses for both forward simulation and inversion of the fault transmissibility for the aforementioned single-phase subsurface flow model. For discretization of the forward model, one can use many different numerical

methods such as the mixed method [6, 9], the multipoint flux approximation [9], the finite volume method [7], and the mimetic finite difference method [10]. In this paper we use the mixed finite element method as it is well-supported by FEniCS [11, 12] and hIPPYlib [13, 14], the two open source software packages for finite element methods and Bayesian inversion problems that we use for numerical results.

The remainder of the paper is organized as follows. In Section 2 we present preliminary background notions on function spaces, derive the single-phase forward partial differential equation (PDE) of subsurface flow id domain with faults, and show its Babška–Brezzi stability. In Section 3 we discretize the model with mixed finite elements and prove *a priori* error estimates. In Section 4 we first present an infinite Bayesian inversion framework and show its well-posedness. This is followed by FEM discretizations of the prior and the likelihood and their error analyses. We then combine prior and likelihood discretizations to construct a FEM discretization of the infinite Bayesian posterior measure and provide its convergence analysis. MAP computation as deterministic optimization and a Laplace approximation of the Bayesian posterior are also presented. Numerical results illustrating our theoretical findings are presented in Section 5. Section 6 concludes the paper with future research directions.

2. Preliminaries

2.1. Notations

Let Ω be a bounded domain in $\mathbb{R}^n$ with $n = 2$ or 3. For a nonnegative integer m, $H^m(\Omega)$, $H^m(\Omega;\mathbb{R}^n)$ denote the standard $\mathbb{R}$ and $\mathbb{R}^n$-valued Sobolev spaces based on L^2 norm, and $\|\cdot\|_m$ with $m \geq 0$ the scalar- or vector-valued H^m-norms. Boldface letters are preserved for vector-valued functions.

For functions $f, g \in L^2(\Omega)$ and $\boldsymbol{f}, \boldsymbol{g} \in L^2(\Omega;\mathbb{R}^n)$ let

$$(f, g) := \int_\Omega f g \, dx, \qquad (\boldsymbol{f}, \boldsymbol{g}) := \int_\Omega \boldsymbol{f} \cdot \boldsymbol{g} \, dx.$$

Similarly, for an $(n-1)$-dimensional submanifold $D \subset \Omega$ and functions $f, g \in L^2(D)$, $\boldsymbol{f}, \boldsymbol{g} \in L^2(D;\mathbb{R}^n)$ we define

$$\langle f, g\rangle_D := \int_D f \cdot g \, ds, \qquad \langle \boldsymbol{f}, \boldsymbol{g}\rangle_D = \int_D \boldsymbol{f} \cdot \boldsymbol{g} \, ds.$$

2.2. Modeling fault transmissibility

In this section we derive the fault transmissibility model under the assumption of an infinitesimal fault thickness. Let us consider the model domain illustrated in Figure 1 comprising two subdomains Ω_+ and Ω_- divided by a fault domain Γ_f with thickness d_f. For simplicity we assume that the two interfaces $\Gamma_+ := \partial\Omega_+ \cap \partial\Gamma_f$, $\Gamma_- := \partial\Omega_- \cap \partial\Gamma_f$ are parallel to each other. The constant permeability parameters on the subdomains are κ_+, κ_-, κ_f, respectively, and p_+, p_- are the pressure values on the interfaces Γ_+, Γ_-. We denote by $\boldsymbol{n}$ the unit normal vector field on Γ_+ from Ω_- to Ω_+. We assume

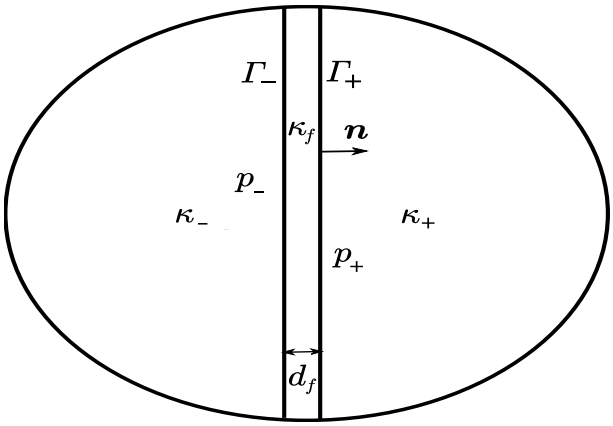

Figure 1: Modeling of fault.

that fluid flows follow the Darcy law, i.e., the velocity of fluid is $-\kappa\nabla p$ for a pressure field p. Here we assume that the pressure field p is continuous, so if we denote the pressure field in the fault of thickness d_f by p_f, then $p_- = p_f|_{\Gamma_-}$ and $p_+ = p_f|_{\Gamma_+}$. We also assume that Γ_f is thin and $\kappa_f \ll \kappa_+, \kappa_-$, so the fault has no absorption, drainage of fluids, and tangential fluid flows. By this assumption and flux conservation

$$-(\kappa_+\nabla p)\cdot\boldsymbol{n}|_{\Gamma_+} = -(\kappa_f\nabla p_f)\cdot\boldsymbol{n}|_{\Gamma_+}, \quad -(\kappa_-\nabla p)\cdot\boldsymbol{n}|_{\Gamma_-} = -(\kappa_f\nabla p_f)\cdot\boldsymbol{n}|_{\Gamma_-}.$$

By the fundamental theorem of calculus along the direction orthogonal to Γ_+, we have

$$-\kappa_f(p_+ - p_-) = -\int_{-d_f/2}^{d_f/2} (\kappa_f\nabla p_f)\cdot\boldsymbol{n}\, dl.$$

In our modeling, d_f is small, so we may assume that κ_f is nearly a constant on the fault. The approximation of the above integral with the trapezoidal rule and the flux continuity give

$$-\kappa_f(p_+ - p_-) \approx -\frac{d_f}{2}((\kappa_+\nabla p)\cdot \boldsymbol{n}|_{\Gamma_+} + (\kappa_-\nabla p)\cdot \boldsymbol{n}|_{\Gamma_-}).$$

and the pressure gradient in Γ_f is high. The difference of p_+ and p_- is thus not negligible. In addition, since we assume that there is no tangential flows along the fault, we have

$$(\kappa_+\nabla p)\cdot \boldsymbol{n}|_{\Gamma_+} = (\kappa_-\nabla p)\cdot \boldsymbol{n}|_{\Gamma_-},$$

which results in the following constitutive equation for the fault

$$[\![p]\!] := p_+ - p_- = t_f(\kappa_+\cdot\nabla p)\cdot \boldsymbol{n}|_{\Gamma_+} = t_f(\kappa_-\cdot\nabla p)\cdot \boldsymbol{n}|_{\Gamma_-}, \qquad t_f = \frac{d_f}{\kappa_f}. \quad (1)$$

Since d_f is much smaller than the characteristic length scale of our subsurface model, Γ_f is considered as a zero thickness fault in our partial differential equation model in Section 2.3.

2.3. Partial differential equation model (forward problem)

A fault Γ is a union of disjoint $(n-1)$-dimensional Lipschitz submanifolds in Ω and we denote $\Omega\setminus\Gamma$ by $\mathring{\Omega}$. We assume that there exist open subdomains $\Omega_+, \Omega_- \subset \mathring{\Omega}$ with Lipschitz boundaries such that $\overline{\Omega} = \overline{\Omega_+}\cup\overline{\Omega_-}$, $\Gamma \subset \partial\Omega_+ \cap \partial\Omega_-$, and only one side of Γ is in contact with Ω_+ or Ω_-. Let $\boldsymbol{n}_+$ and $\boldsymbol{n}_-$ be the two unit normal vector fields on Γ with opposite directions ($\boldsymbol{n}_+ = -\boldsymbol{n}_-$) such that $\boldsymbol{n}_\pm$ correspond to the unit outward normal vector fields from $\Omega_\pm$. Suppose that Γ_D and Γ_N are $(n-1)$-dimensional open submanifolds on $\partial\Omega$ such that $\overline{\Gamma_D}\cup\overline{\Gamma_N} = \partial\Omega$ and $\Gamma_D\cap\Gamma_N = \emptyset$. Here we use the convention that $(\cdot,\cdot)$ is the integration on Ω and $\langle\cdot,\cdot\rangle_\Gamma$ is the integration on Γ.

We also assume that

(A1) for any $q \in L^2(\Omega)$ there exists $\boldsymbol{w} \in H^1(\Omega;\mathbb{R}^n)$ such that $\boldsymbol{w}|_\Gamma = 0$, $\operatorname{div}\boldsymbol{w} = q$ and $\|\boldsymbol{w}\|_1 \le C\|q\|_0$ with a constant $C > 0$ depending only on Ω and Γ.

The assumption (A1) holds, for instance, if both of $\partial\Omega_+ \cap \partial\Omega$ and $\partial\Omega_- \cap \partial\Omega$ have positive $(n-1)$-dimensional Lebesgue measures. A proof can be found in [15, Corollary 2.4].

Let $\gamma_+ : H^1(\Omega_+) \to H^{1/2}(\Gamma)$ and $\gamma_- : H^1(\Omega_-) \to H^{1/2}(\Gamma)$ be the trace operators. For a function $p \in H^1(\mathring{\Omega})$, we define $p_\pm$ on the fault Γ as

$$p_-(x) = \gamma_- p, \qquad p_+(x) = \gamma_+ p, \quad \forall x \in \Gamma.$$

Suppose that $t_f > 0$ is a fault transmissibility function on Γ and $\boldsymbol{u}$ is a vector-valued function on Ω such that $\boldsymbol{u} \cdot \boldsymbol{n}_+ = -\boldsymbol{u} \cdot \boldsymbol{n}_-$ is single-valued on Γ.

Assume that κ is a permeability symmetric positive definite tensor on Ω. The pressure and flux boundary conditions are given as g_D on Γ_D and g_N on Γ_N. Recalling (1), for pressure p and flux $\boldsymbol{u} = \kappa\nabla p$ in $\mathring{\Omega}$, a mixed formulation of the Darcy equation in domain Ω with fault Γ reads:

$$\kappa^{-1}\boldsymbol{u} + \nabla p = 0 \quad \text{in } \mathring{\Omega}, \qquad \operatorname{div} \boldsymbol{u} = f \quad \text{in } \mathring{\Omega}, \tag{2a}$$

$$\boldsymbol{u} \cdot \boldsymbol{n} = g_N \quad \text{on } \Gamma_N, \qquad p = g_D \quad \text{on } \Gamma_D, \tag{2b}$$

$$\boldsymbol{u} \cdot \boldsymbol{n} - t_f^{-1} [\![p]\!] = 0 \quad \text{on } \Gamma. \tag{2c}$$

Hereafter, we assume that $\Gamma_D = \partial\Omega$ and $g_D = 0$ for the simplicity of the exposition. Let $Q = L^2(\Omega)$, and $H(\operatorname{div}, \Omega)$ be the space of $\mathbb{R}^n$-valued L^2 functions on Ω such that its distributional divergence is in $L^2(\Omega)$. We define $\boldsymbol{V}$ as

$$\boldsymbol{V} = \{\boldsymbol{v} \in H(\operatorname{div}, \Omega) \,:\, \boldsymbol{v} \cdot \boldsymbol{n}|_\Gamma \in L^2(\Gamma)\}$$

and the norm on $\boldsymbol{V}$ is defined by

$$\|\boldsymbol{v}\|_{\boldsymbol{V}} = \left(\|\boldsymbol{v}\|_0^2 + \|\operatorname{div} \boldsymbol{v}\|_0^2 + \|\boldsymbol{v} \cdot \boldsymbol{n}\|_{0,\Gamma}^2\right)^{\frac{1}{2}}$$

where $\|q\|_{0,\Gamma} := \langle q, q\rangle_\Gamma^{1/2}$. Here we derive a variational formulation with the interior Robin-type boundary condition (2c) on Γ with an additional regularity assumptions on the exact solution $\boldsymbol{u}$ and p. More specifically, we assume that the exact solution $\boldsymbol{u}$ and p satisfy $p \in H^1(\mathring{\Omega})$ and $\boldsymbol{u} \in \boldsymbol{V}$ with additional regularity satisfying (2c) almost everywhere on Γ. We remark that this approach is inspired by [16] on mixed finite element methods for Robin boundary condition problems.

From the integration by parts of the first equation in (2a), we have

$$\int_\Omega \kappa^{-1}\boldsymbol{u}\cdot\boldsymbol{v}\,dx - \int_\Omega p\ \mathrm{div}\,\boldsymbol{v}\,dx + \langle p_+, \boldsymbol{v}\cdot\boldsymbol{n}_+\rangle_\Gamma + \langle p_-, \boldsymbol{v}\cdot\boldsymbol{n}_-\rangle_\Gamma = 0 \qquad \forall \boldsymbol{v}\in\boldsymbol{V},$$

which, after invoking $\boldsymbol{v}\cdot\boldsymbol{n}_+ = -\boldsymbol{v}\cdot\boldsymbol{n}_-$, $\langle p_+, \boldsymbol{v}\cdot\boldsymbol{n}_+\rangle_\Gamma + \langle p_-, \boldsymbol{v}\cdot\boldsymbol{n}_-\rangle_\Gamma = \langle [\![p]\!], \boldsymbol{v}\cdot\boldsymbol{n}_+\rangle_\Gamma$, and $\boldsymbol{u}\cdot\boldsymbol{n}_+ - t_f^{-1}[\![p]\!] = 0$, becomes

$$\int_\Omega \kappa^{-1}\boldsymbol{u}\cdot\boldsymbol{v}\,dx - \int_\Omega p\ \mathrm{div}\,\boldsymbol{v}\,dx + \langle t_f\boldsymbol{u}\cdot\boldsymbol{n}_+, \boldsymbol{v}\cdot\boldsymbol{n}_+\rangle_\Gamma = 0 \qquad \forall \boldsymbol{v}\in\boldsymbol{V}. \quad (3)$$

In the following, we use $\langle t_f\boldsymbol{u}\cdot\boldsymbol{n}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_\Gamma$ to denote $\langle t_f\boldsymbol{u}\cdot\boldsymbol{n}_+, \boldsymbol{v}\cdot\boldsymbol{n}_+\rangle_\Gamma$ since the bilinear form $\langle t_f\boldsymbol{u}\cdot\boldsymbol{n}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_\Gamma$ is same for $\boldsymbol{n}=\boldsymbol{n}_+$ and $\boldsymbol{n}=\boldsymbol{n}_-$.

In the final variational formulation the exact solution $(\boldsymbol{u}, p)$ satisfies

$$\begin{cases} (\kappa^{-1}\boldsymbol{u}, \boldsymbol{v}) - (p, \mathrm{div}\,\boldsymbol{v}) + \langle t_f\boldsymbol{u}\cdot\boldsymbol{n}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_\Gamma &= 0 & \forall \boldsymbol{v}\in\boldsymbol{V}, \\ (\mathrm{div}\,\boldsymbol{u}, q) &= (f, q) & \forall q\in Q. \end{cases} \quad (4)$$

When $\|t_f\|_{L^\infty(\Gamma)} \to 0$ the above system approaches the mixed formulation of Darcy equation with no faults. On the other hand, the above system becomes the complete sealing fault case as $\|t_f^{-1}\|_{L^\infty(\Gamma)} \to 0$ because $\boldsymbol{u}\cdot\boldsymbol{n}_+ \to 0$ in (2c).

The system (4) can be viewed as a saddle point problem

$$\begin{cases} a(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) &= F(\boldsymbol{v}), \\ b(\boldsymbol{u}, q) &= G(q) \end{cases} \quad (5)$$

with the two bilinear form and two linear forms

$$a(\boldsymbol{u}, \boldsymbol{v}) = \left(\kappa^{-1}\boldsymbol{u}, \boldsymbol{v}\right) + \langle t_f\boldsymbol{u}\cdot\boldsymbol{n}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_\Gamma, \qquad b(\boldsymbol{v}, q) = (\mathrm{div}\,\boldsymbol{v}, q),$$
$$F(\boldsymbol{v}) = 0, \qquad G(q) = -(f, q).$$

Theorem 1. *Suppose that $\|t_f\|_{L^\infty(\Gamma)}, \|t_f^{-1}\|_{L^\infty(\Gamma)} < +\infty$ and F and G are bounded linear functionals on $\boldsymbol{V}$ and $L^2(\Omega)$ in (5). Then, (5) has a unique solution $(\boldsymbol{u}, p)\in\boldsymbol{V}\times L^2(\Omega)$.*

Proof. By the definition of $\|\boldsymbol{v}\|_{\boldsymbol{V}}$ it is easy to check

$$|a(\boldsymbol{v}, \boldsymbol{v}')| \le \|a\|\|\boldsymbol{v}\|_{\boldsymbol{V}}\|\boldsymbol{v}'\|_{\boldsymbol{V}}, \qquad |b(\boldsymbol{v}, q)| \le \|\boldsymbol{v}\|_{\boldsymbol{V}}\|q\|_0$$

with

$$\|a\| := \max\{\|\kappa^{-1}\|_{L^\infty(\Omega)}, \|t_f\|_{L^\infty(\Gamma)}\}. \quad (6)$$

By (A1) we can obtain the inf-sup condition for the Babuška–Brezzi stability theory (cf. [17])

$$\inf_{q\in Q}\sup_{\boldsymbol{v}\in\boldsymbol{V}}\frac{b(\boldsymbol{v},q)}{\|\boldsymbol{v}\|_{\boldsymbol{V}}\|q\|_0}\geq\beta>0 \tag{7}$$

where β depends on Ω. Since we assume $\|t_f^{-1}\|_{L^\infty(\Gamma)}<+\infty$ the coercivity

$$\inf_{\boldsymbol{v}\in\boldsymbol{Z}}\frac{a(\boldsymbol{v},\boldsymbol{v})}{\|\boldsymbol{v}\|_{\boldsymbol{V}}^2}\geq\alpha>0,\qquad \boldsymbol{Z}:=\{\boldsymbol{v}'\in\boldsymbol{V}\,:\,\operatorname{div}\boldsymbol{v}'=0\} \tag{8}$$

with $\alpha=\min\{\|\kappa\|_{L^\infty(\Omega)}^{-1},\|t_f^{-1}\|_{L^\infty(\Gamma)}^{-1}\}$, is obtained by

$$\begin{aligned}a(\boldsymbol{v},\boldsymbol{v})&=\left(\kappa^{-1}\boldsymbol{v},\boldsymbol{v}\right)+\langle t_f\boldsymbol{v}\cdot\boldsymbol{n},\boldsymbol{v}\cdot\boldsymbol{n}\rangle_\Gamma\\&\geq\min\{\|\kappa\|_{L^\infty(\Omega)}^{-1},\|t_f^{-1}\|_{L^\infty(\Gamma)}^{-1}\}\left(\|\boldsymbol{v}\|_0^2+\|\boldsymbol{v}\cdot\boldsymbol{n}\|_{0,\Gamma}^2\right)\\&\geq\min\{\|\kappa\|_{L^\infty(\Omega)}^{-1},\|t_f^{-1}\|_{L^\infty(\Gamma)}^{-1}\}\|\boldsymbol{v}\|_{\boldsymbol{V}}^2\end{aligned}$$

for $\boldsymbol{v}\in\boldsymbol{Z}$. Then, the well-posedness of (4) follows from the Babuška–Brezzi theory [17]. □

For an analysis of inverse problem presented in Section 4, we claim that the solution of (5) depends continuously on t_f under additional assumptions.

Theorem 2. *Suppose that $m,\tilde{m}$ are Hölder continuous functions in $C^{0,t}(\Gamma)$ for some $t>0$. Suppose that $(\boldsymbol{u},p)$ and $(\tilde{\boldsymbol{u}},\tilde{p})$ in $\boldsymbol{V}\times Q$ are the solutions of (5) for $t_f=e^m$ and $t_f=e^{\tilde{m}}$, respectively. Then,*

$$\|\boldsymbol{u}-\tilde{\boldsymbol{u}}\|_{\boldsymbol{V}}+\|p-\tilde{p}\|_0\leq C\|t_f-\tilde{t}_f\|_{L^\infty(\Gamma)}$$

with $C>0$ depending on $\|\kappa^{-1}\|_{L^\infty(\Omega)}$, $\|t_f\|_{L^\infty(\Gamma)}$, and $\|\tilde{t}_f\|_{L^\infty(\Gamma)}$.

Proof. For $t_f=e^m$ and $\tilde{t}_f=e^{\tilde{m}}$ the coercivity constants of $a(\cdot,\cdot)$ are

$$\alpha=\min\{\|\kappa\|_{L^\infty(\Omega)}^{-1},e^{-\|m\|_{L^\infty(\Gamma)}}\},\quad \tilde{\alpha}=\min\{\|\kappa\|_{L^\infty(\Omega)}^{-1},e^{-\|\tilde{m}\|_{L^\infty(\Gamma)}}\}.$$

By [17, Theorem 4.2.3], the solution $(\boldsymbol{u},p)$ of (5) satisfies

$$\|\boldsymbol{u}\|_{\boldsymbol{V}}\leq\frac{1}{\alpha}\|F\|_{\boldsymbol{V}'}+\frac{2\|a\|^{1/2}}{\alpha^{1/2}\beta}\|G\|_0, \tag{9}$$

$$\|p\|_0\leq\frac{2\|a\|^{1/2}}{\alpha^{1/2}\beta}\|F\|_{\boldsymbol{V}'}+\frac{\|a\|}{\beta^2}\|G\|_0 \tag{10}$$

with α, β, $\|a\|$ in (8), (7), (6) where

$$\|F\|_{\boldsymbol{V}'} := \sup_{\boldsymbol{v}\in\boldsymbol{V}} \frac{F(\boldsymbol{v})}{\|\boldsymbol{v}\|_{\boldsymbol{V}}}, \qquad \|G\|_0 := \sup_{q\in L^2(\Omega)} \frac{G(q)}{\|q\|_0}.$$

From (5) it is easy to see that

$$\begin{aligned}\left(\kappa^{-1}(\boldsymbol{u}-\tilde{\boldsymbol{u}}),\boldsymbol{v}\right) + \langle t_f\,(\boldsymbol{u}-\tilde{\boldsymbol{u}})\cdot\boldsymbol{n}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_\Gamma + b(p-\tilde{p},\boldsymbol{v}) + b(\boldsymbol{u}-\tilde{\boldsymbol{u}},q)& \\ = -\left\langle (t_f-\tilde{t}_f)\tilde{\boldsymbol{u}}\cdot\boldsymbol{n}, \boldsymbol{v}\cdot\boldsymbol{n}\right\rangle_\Gamma.&\end{aligned}$$

By (9), (10),

$$\begin{aligned}\|\boldsymbol{u}-\tilde{\boldsymbol{u}}\|_{\boldsymbol{V}} &\le \frac{1}{\alpha}\sup_{\boldsymbol{v}\in\boldsymbol{V}} \frac{-\left\langle (t_f-\tilde{t}_f)\tilde{\boldsymbol{u}}\cdot\boldsymbol{n}, \boldsymbol{v}\cdot\boldsymbol{n}\right\rangle_\Gamma}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} \\ &\le \frac{1}{\alpha} C(m,\tilde{m})\|m-\tilde{m}\|_{L^\infty(\Gamma)}\|\tilde{\boldsymbol{u}}\cdot\boldsymbol{n}\|_{L^2(\Gamma)},\end{aligned}$$

and similarly,

$$\|p-\tilde{p}\|_0 \le \frac{2\|a\|^{1/2}}{\alpha^{1/2}\beta} C(m,\tilde{m})\|m-\tilde{m}\|_{L^\infty(\Gamma)}\|\tilde{\boldsymbol{u}}\cdot\boldsymbol{n}\|_{L^2(\Gamma)}.$$

Finally, $\|\tilde{\boldsymbol{u}}\cdot\boldsymbol{n}\|_{L^2(\Gamma)}$ is bounded by the formulas (9) with $F=0$, $G=f$ where $\|a\|$ and α are replaced by

$$\|\tilde{a}\| := \max\{\|\kappa^{-1}\|_{L^\infty(\Omega)}, \|\tilde{t}_f\|_{L^\infty(\Gamma)}\},$$

and $\tilde{\alpha}$, and it leads to the conclusion. □

3. Discretization with mixed methods and the a priori error analysis

In this section we discuss the finite element discretization and the a priori error analysis of (4). Throughout this section we assume that Ω is a bounded domain with a polygonal/polyhedral boundary.

Let $\mathcal{T}_h$ be a triangulation of Ω with n-dimensional simplices without hanging nodes with discretization parameter $h>0$ which is the maximum diameter of n-dimensional simplices in $\mathcal{T}_h$. $\mathcal{E}_h$ is the corresponding set of $(n-1)$-dimensional simplices generated by $\mathcal{T}_h$. We always assume that a subset of $\mathcal{E}_h$,

denoted by $\mathcal{E}_h^\Gamma$, forms a triangulation of the fault Γ, and $\mathcal{T}_h$ is shape-regular with an upper-bound of shape regularity which is uniform in h (cf. [18]).

For an integer $k \geq 0$ and a set $D \subset \mathbb{R}^n$, $\mathcal{P}_k(D)$ is the space of polynomials defined on D of degree at most k. Similarly, $\mathcal{P}_k(D;\mathbb{R}^n)$ is the space of $\mathbb{R}^n$-valued polynomials of degree at most k. For given $k \geq 1$ let us define

$$\begin{aligned} \boldsymbol{V}_h^{RTN}(T) &= \mathcal{P}_{k-1}(T;\mathbb{R}^n) + \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mathcal{P}_{k-1}(T), \\ \boldsymbol{V}_h^{BDM}(T) &= \mathcal{P}_k(T;\mathbb{R}^n), \qquad T \in \mathcal{T}_h. \end{aligned}$$

Suppose that $\boldsymbol{V}_h \subset \boldsymbol{V}$ is the Raviart–Thomas–Nedelec (RTN) or Brezzi–Douglas–Marini (BDM) element defined by

$$\begin{aligned} \boldsymbol{V}_h &= \{\boldsymbol{v} \in \boldsymbol{V} \,:\, \boldsymbol{v}|_T \in \boldsymbol{V}_h^{RTN}(T), \quad \forall T \in \mathcal{T}_h\}, \\ \text{or } \boldsymbol{V}_h &= \{\boldsymbol{v} \in \boldsymbol{V} \,:\, \boldsymbol{v}|_T \in \boldsymbol{V}_h^{BDM}(T), \quad \forall T \in \mathcal{T}_h\}. \end{aligned}$$

The finite element space Q_h is defined by

$$Q_h = \{q \in Q \,:\, q|_T \in \mathcal{P}_{k-1}(T) \quad \forall T \in \mathcal{T}_h\}.$$

Then $\operatorname{div} \boldsymbol{V}_h = Q_h$ and it is well-known that the pair $(\boldsymbol{V}_h, Q_h)$ satisfies (7) with a discrete inf-sup constant $\bar{\beta}$ independent of h [17, p. 406].

The discretization of (4) is to seek $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$ such that

$$\begin{cases} (\kappa^{-1}\boldsymbol{u}_h, \boldsymbol{v}) - (p_h, \operatorname{div} \boldsymbol{v}) + \langle t_f \boldsymbol{u}_h \cdot \boldsymbol{n}, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_\Gamma &= 0 & \forall \boldsymbol{v} \in \boldsymbol{V}_h, \\ (\operatorname{div} \boldsymbol{u}_h, q) &= (f, q) & \forall q \in Q_h. \end{cases} \tag{11}$$

Since (7) and (8) hold with $(\boldsymbol{V}_h, Q_h)$, well-posedness and the stability of this system follow from the standard Babuška-Brezzi theory.

3.1. The a priori error analysis

For error analysis we consider the difference of (4) and (11) which reads

$$\begin{aligned} &(\kappa^{-1}(\boldsymbol{u} - \boldsymbol{u}_h), \boldsymbol{v}) - (p - p_h, \operatorname{div} \boldsymbol{v}) && && \text{(12a)} \\ &\quad + \langle t_f(\boldsymbol{u} - \boldsymbol{u}_h) \cdot \boldsymbol{n}, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_\Gamma = 0 && \forall \boldsymbol{v} \in \boldsymbol{V}_h, \\ &(\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_h), q) = 0 && \forall q \in Q_h. && \text{(12b)} \end{aligned}$$

Let Π_h be the canonical interpolation operator into $\boldsymbol{V}_h$ by the standard degrees of freedom and P_h be the L^2 projection into Q_h. Then it is known (cf. [17]) that

$$P_h \operatorname{div} \boldsymbol{v} = \operatorname{div} \Pi_h \boldsymbol{v} \tag{13}$$

holds for $\boldsymbol{v} \in \boldsymbol{V} \cap L^r(\Omega; \mathbb{R}^n)$, $r > 2$.

Theorem 3. *Suppose that $(\boldsymbol{u}, p)$ and $(\boldsymbol{u}_h, p_h)$ are the solutions of (4) and (11). Assuming that $\boldsymbol{u}$ and p are sufficiently regular to make the norms below well-defined, then*

$$\begin{aligned} &\left\|\kappa^{-1/2}(\boldsymbol{u} - \boldsymbol{u}_h)\right\|_0 + \left\|t_f^{1/2}(\boldsymbol{u} - \boldsymbol{u}_h) \cdot \boldsymbol{n}\right\|_{0,\Gamma} \qquad (14)\\ &\qquad \le \begin{cases} Ch^{k-1/2} \left\|\boldsymbol{u}\right\|_k & \text{if } \boldsymbol{V}_h \text{ is RTN element} \\ Ch^{k+1/2} \left\|\boldsymbol{u}\right\|_{k+1} & \text{if } \boldsymbol{V}_h \text{ is BDM element} \end{cases}, \\ &\left\|p - p_h\right\|_0 \le \begin{cases} Ch^{k-1/2} \left(\left\|\boldsymbol{u}\right\|_k + \left\|p\right\|_k\right) & \text{if } \boldsymbol{V}_h \text{ is RTN element} \\ Ch^{k} \left(\left\|\boldsymbol{u}\right\|_{k+1/2} + \left\|p\right\|_k\right) & \text{if } \boldsymbol{V}_h \text{ is BDM element} \end{cases}. \qquad (15) \end{aligned}$$

Proof. We decompose the errors as the following

$$\begin{aligned} \boldsymbol{u} - \boldsymbol{u}_h &= (\boldsymbol{u} - \Pi_h \boldsymbol{u}) + \underbrace{(\Pi_h \boldsymbol{u} - \boldsymbol{u}_h)}_{e_{\boldsymbol{u}}} =: (\boldsymbol{u} - \Pi_h \boldsymbol{u}) + e_{\boldsymbol{u}}, \\ p - p_h &= (p - P_h p) + \underbrace{(P_h p - p_h)}_{e_p} =: (p - P_h p) + e_p. \end{aligned}$$

From (12b) $P_h \operatorname{div} \boldsymbol{u} = \operatorname{div} \boldsymbol{u}_h$ holds, therefore $\operatorname{div} e_{\boldsymbol{u}} = 0$ by the definition of $e_{\boldsymbol{u}}$. Taking $\boldsymbol{v} = e_{\boldsymbol{u}}$ in (12a), we have

$$\begin{aligned} &(\kappa^{-1} e_{\boldsymbol{u}}, e_{\boldsymbol{u}}) + \langle t_f e_{\boldsymbol{u}} \cdot \boldsymbol{n}, e_{\boldsymbol{u}} \cdot \boldsymbol{n} \rangle_\Gamma \\ &\qquad = -(\kappa^{-1}(\boldsymbol{u} - \Pi_h \boldsymbol{u}), e_{\boldsymbol{u}}) - \langle t_f (\boldsymbol{u} - \Pi_h \boldsymbol{u}) \cdot \boldsymbol{n}, e_{\boldsymbol{u}} \cdot \boldsymbol{n} \rangle_\Gamma . \qquad (16) \end{aligned}$$

Applying the Cauchy–Schwarz inequality to (16) we obtain

$$\begin{aligned} &\left\|\kappa^{-1/2} e_{\boldsymbol{u}}\right\|_0 + \left\|t_f^{1/2} e_{\boldsymbol{u}} \cdot \boldsymbol{n}\right\|_{0,\Gamma} \\ &\qquad \le 2 \left\|\kappa^{-1/2}(\boldsymbol{u} - \Pi_h \boldsymbol{u})\right\|_0 + 2 \left\|t_f^{1/2}(\boldsymbol{u} - \Pi_h \boldsymbol{u}) \cdot \boldsymbol{n}\right\|_{0,\Gamma} . \end{aligned}$$

The interpolation Π_h gives an optimal order of approximation in $L^2(\Omega)$ [17] but we lose $1/2$ order in the approximation of $\left\|t_f^{1/2} e_{\boldsymbol{u}} \cdot \boldsymbol{n}\right\|_{0,\Gamma}$ because of the

scaling factor in trace inequalities in the theory of finite element methods (cf. [19, Lemma 1.49]).

Assuming that $\boldsymbol{u}$ is sufficiently regular, we obtain error estimates

$$\left\|\kappa^{-1/2} e_{\boldsymbol{u}}\right\|_0 + \left\|t_f^{1/2} e_{\boldsymbol{u}} \cdot \boldsymbol{n}\right\|_{0,\Gamma} \le \begin{cases} Ch^{k-1/2} \|\boldsymbol{u}\|_k & \text{if } \boldsymbol{V}_h \text{ is RTN element} \\ Ch^{k+1/2} \|\boldsymbol{u}\|_{k+1} & \text{if } \boldsymbol{V}_h \text{ is BDM element} \end{cases}$$

where $C > 0$ depends on $\|\kappa^{-1}\|_{L^\infty}$ and $\|t_f\|_{L^\infty}$. Then, (14) is obtained by the triangle inequality.

To estimate $\|p - p_h\|_0$, note that $(p - p_h, \operatorname{div} \boldsymbol{v}) = (e_p, \operatorname{div} \boldsymbol{v})$ for $\boldsymbol{v} \in \boldsymbol{V}_h$ owing to $(p - P_h p, \operatorname{div} \boldsymbol{v}) = 0$ by definition of P_h. By (A1), there exists $\boldsymbol{w} \in H^1(\Omega; \mathbb{R}^n)$ such that $\operatorname{div} \boldsymbol{w} = e_p$, $\|\boldsymbol{w}\|_1 \le C\|e_p\|_0$, $\boldsymbol{w}|_\Gamma = 0$. Then $\boldsymbol{v} = \Pi_h \boldsymbol{w} \in \boldsymbol{V}_h$ satisfies

$$\operatorname{div} \boldsymbol{v} = e_p, \qquad \|\boldsymbol{v}\|_{\operatorname{div}} \lesssim \|e_p\|_0, \qquad \boldsymbol{v} \cdot \boldsymbol{n}|_\Gamma = 0, \tag{17}$$

thanks to (13). Taking this $\boldsymbol{v}$ in (12a), we have

$$\|e_p\|_0^2 = (\kappa^{-1}(\boldsymbol{u} - \boldsymbol{u}_h), \boldsymbol{v}) \le C \|\boldsymbol{u} - \boldsymbol{u}_h\|_0 \|\boldsymbol{v}\|_0 \le C \|\boldsymbol{u} - \boldsymbol{u}_h\|_0 \|e_p\|_0$$

As a consequence,

$$\|e_p\|_0 \le \begin{cases} Ch^{k-1/2} \|\boldsymbol{u}\|_k & \text{if } \boldsymbol{V}_h \text{ is RTN element} \\ Ch^{k+1/2} \|\boldsymbol{u}\|_{k+1} & \text{if } \boldsymbol{V}_h \text{ is BDM element} \end{cases}$$

In this estimate, the convergence order of $\|e_p\|_0$ is 1/2 order superconvergent if the BDM element is used. Finally, (15) is obtained by the triangle inequality. □

3.2. Improved error analysis

In this subsection we show an improved error analysis under additional assumptions on t_f. Recall that a function g on a domain D is in $W^{1,\infty}(D)$ if g is differentiable in the sense of distributions on D and the derivative of g is in $L^\infty(D)$. For more details on $W^{1,\infty}$ we refer to standard references, e.g., [20, 21]. We say that t_f is element-wise $W^{1,\infty}$ on Γ if $t_f|_e \in W^{1,\infty}(e)$ for all $e \in \mathcal{E}_h^\Gamma$, and $\|t_f\|_{W_h^{1,\infty}(\Gamma)} := \max_{e \in \mathcal{E}_h^\Gamma} \|t_f|_e\|_{W^{1,\infty}}$.

Let $\overline{t_f}$ be the piecewise constant function on Γ such that $\overline{t_f}|_e$ is the mean value of t_f on $e \in \mathcal{E}_h^\Gamma$.

Theorem 4. *Assume that t_f is element-wise $W^{1,\infty}$ on Γ. Then,*

$$\left\|\kappa^{-1/2}(\boldsymbol{u}-\boldsymbol{u}_h)\right\|_0 \leq \begin{cases} Ch^k \|\boldsymbol{u}\|_k & \text{if } \boldsymbol{V}_h \text{ is RTN element} \\ Ch^{k+1} \|\boldsymbol{u}\|_{k+1} & \text{if } \boldsymbol{V}_h \text{ is BDM element} \end{cases} \tag{18}$$

and

$$\|p-p_h\|_0 \leq Ch^k \left(\|\boldsymbol{u}\|_k + \|p\|_k\right). \tag{19}$$

Proof. Let P_h^Γ be the L^2 projection to the space

$$\Lambda_h := \{\lambda \in L^2(\Gamma) \,:\, \lambda|_e \in \mathcal{P}_l(e) \quad \forall e \in \mathcal{F}_h\}$$

with $l = k$ if $\boldsymbol{V}_h$ is a BDM element and $l = k-1$ if $\boldsymbol{V}_h$ is an RTN element. From the definition of Π_h one can see $\Pi_h\boldsymbol{u}\cdot\boldsymbol{n}|_e = P_h^\Gamma(\boldsymbol{u}\cdot\boldsymbol{n})|_e$ for all $e \in \mathcal{F}_h$. Then

$$\langle t_f(\boldsymbol{u}-\Pi_h\boldsymbol{u})\cdot\boldsymbol{n}, e_{\boldsymbol{u}}\cdot\boldsymbol{n}\rangle_\Gamma = \left\langle (t_f-\overline{t_f})(\boldsymbol{u}-\Pi_h\boldsymbol{u})\cdot\boldsymbol{n}, e_{\boldsymbol{u}}\cdot\boldsymbol{n}\right\rangle_\Gamma. \tag{20}$$

Applying the Hölder, the Cauchy–Schwarz inequalities, and the Bramble–Hilbert lemma to (20) gives

$$\begin{aligned} &|\langle t_f(\boldsymbol{u}-\Pi_h\boldsymbol{u})\cdot\boldsymbol{n}, e_{\boldsymbol{u}}\cdot\boldsymbol{n}\rangle_\Gamma| \\ &\qquad \leq Ch \|t_f\|_{W_h^{1,\infty}(\Gamma)} \|(\boldsymbol{u}-\Pi_h\boldsymbol{u})\cdot\boldsymbol{n}\|_{0,\Gamma} \|e_{\boldsymbol{u}}\cdot\boldsymbol{n}\|_{0,\Gamma}. \end{aligned} \tag{21}$$

If we use the above inequality to (16), then the Cauchy–Schwarz inequality gives

$$\begin{aligned} &\left\|\kappa^{-1/2}e_{\boldsymbol{u}}\right\|_0 + \left\|t_f^{1/2}e_{\boldsymbol{u}}\cdot\boldsymbol{n}\right\|_{0,\Gamma} \\ &\qquad \leq C\left(\left\|\kappa^{-1/2}(\boldsymbol{u}-\Pi_h\boldsymbol{u})\right\|_0 + h\|t_f\|_{W_h^{1,\infty}} \|(\boldsymbol{u}-\Pi_h\boldsymbol{u})\cdot\boldsymbol{n}\|_{0,\Gamma}\right) \end{aligned} \tag{22}$$

with a constant $C > 0$ depending on $\|t_f^{-1}\|_{L^\infty(\Gamma)}$. From this and the triangle inequality we have an improved estimate

$$\left\|\kappa^{-1/2}(\boldsymbol{u}-\boldsymbol{u}_h)\right\|_0 \leq \begin{cases} Ch^k \|\boldsymbol{u}\|_k & \text{if } \boldsymbol{V}_h \text{ is RTN element} \\ Ch^{k+1} \|\boldsymbol{u}\|_{k+1} & \text{if } \boldsymbol{V}_h \text{ is BDM element} \end{cases} \tag{23}$$

To estimate $\|p - p_h\|_0$ we take $\boldsymbol{v} \in \boldsymbol{V}_h$ satisfying (17) in (12a), then we have

$$\|e_p\|_0^2 = (\kappa^{-1}(\boldsymbol{u} - \boldsymbol{u}_h), \boldsymbol{v}) \leq C \|\boldsymbol{u} - \boldsymbol{u}_h\|_0 \|\boldsymbol{v}\|_0 \leq C \|\boldsymbol{u} - \boldsymbol{u}_h\|_0 \|e_p\|_0 .$$

As a consequence,

$$\|e_p\|_0 \leq \begin{cases} Ch^k \|\boldsymbol{u}\|_k & \text{if } \boldsymbol{V}_h \text{ is RTN element} \\ Ch^{k+1} \|\boldsymbol{u}\|_{k+1} & \text{if } \boldsymbol{V}_h \text{ is BDM element} \end{cases}$$

In this estimate, the convergence order of $\|e_p\|_0$ is one order superconvergent if the BDM element is used. Finally, (19) is obtained by the triangle inequality. □

4. Inversion of Transmissibility

In this section we present a statistical inversion of t_f. We start with a Bayesian formulation, prove its well-posedness, present and analyze FEM discretizations of both the state and parameters, and rigorously establish the convergence of a FEM approximation of the Bayesian posterior measure. We conclude the section with the existence of the maximum a posteriori (MAP) point. A Gaussian approximation of the posterior at the MAP point is also discussed.

4.1. Bayesian inversion on infinite dimensional parameter space

Since t_f is positive on Γ we can define $t_f = e^m$ for $m \in L^2(\Gamma)$. We choose to solve the inverse transmissibility problem using the Bayesian framework as it allows us to take into account of uncertainties. We assume that there is no model inadequacy and thus only uncertainties due to limited noisy data and the prior knowledge are taken into account. The Bayesian framework starts with a prior distribution μ_{pr} of the parameter m and then update it to the posterior distribution μ_{post} using information from observable data by the Bayes' rule

$$\frac{d\mu_{\text{post}}}{d\mu_{\text{pr}}} \propto \pi_{\text{like}}(y_d|m) \tag{24}$$

where $d\mu_{\text{post}}/d\mu_{\text{pr}}$ is the Radon–Nikodym derivative of the posterior measure μ_{post} with respect to the prior measure μ_{pr}, and $\pi_{\text{like}}(y_d|m)$ is the likelihood according to the observation data y_d.

To show that (24) is well-posed, we postulate the prior distribution of m as a Gaussian measure in $L^2(\Gamma)$, i.e., $\mu_{\text{pr}} = \mathcal{N}(m_{\text{pr}}, \mathcal{C}_{\text{pr}})$, where m_{pr} resides in the Cameron-Martin space $\mathcal{M}$ of $\mathcal{C}_{\text{pr}}$. It is sufficient for the prior Gaussian measure to be well-defined if we choose $\mathcal{C}_{\text{pr}} = (\delta I - \gamma\Delta)^{-\sigma}$ for $\sigma > (n-1)/2$ with constants $\delta, \gamma > 0$ [22, 23, 22], where we have assumed that Γ is piecewise flat so that the standard Laplacian operator Δ is meaningful. In this case, the parameter m, distributed under the prior μ_{pr}, is almost surely r-Hölder continuous on Γ for any $0 < r < \min\{1, \sigma - (n-1)/2\}$, i.e., $m \in C^{0,r}(\Gamma)$, and almost surely in $H^r(\Gamma)$ for $0 \leq r < \sigma - (n-1)/2$. By forcing the domain of $\delta I - \gamma\Delta$ to consist of functions having homogeneous Dirichlet and/or Neuman boundary conditions, we can show that the Cameron-Martin space is $\mathcal{M} = H^\sigma(\Gamma)$.

For the likelihood, we adopt the additive noise model

$$y_d = \mathcal{F}(m) + \eta \tag{25}$$

where $\mathcal{F} : L^2(\Gamma) \mapsto \mathbb{R}^{n_{obs}}$ is the parameter to observable map and the noise η follows a Gaussian distribution $\mathcal{N}(0, \Gamma_{\text{noise}})$. For a fixed radius $R > 0$ the local average operator of a function g at a point x is defined by

$$\text{avg}_x(g) := \frac{1}{|b(x,R)|} \int_{b(x,R)} g(y)\, dy, \tag{26}$$

where $b(x, r)$ is the ball of radius R centered at x. The observation operator $\mathcal{B}$ is defined by the linear operator $B : Q \to \mathbb{R}^{n_{obs}}$ which takes the local average of pressure at n_{obs}-points in Ω (with given $0 < R \ll 1$)

$$\mathcal{F}(m) := Bp(m), . \tag{27}$$

where $p(m)$ solves (4) with $t_f = e^m$. In this case the likelihood has a form

$$\pi_{\text{like}}(y_d|m) \propto \exp\left(-\frac{1}{2}\|\mathcal{F}(m) - y_d\|^2_{\Gamma^{-1}_{\text{noise}}}\right).$$

Lemma 1. *The Radon-Nikodym derivative in* (24) *is well-defined in* $L^2(\Gamma)$*, and the parameter* m*, under the posterior* μ_{post}*, almost surely resides in* $X := C^{0,r}(\Gamma)$ *for any* $0 < r < \min\{1, \sigma - (n-1)/2\}$*. Furthermore, the posterior measure is Lipschitz continuous with respect to the data* y_d *in the Hellinger distance.*

Proof. From the definition of observation operator we have, for any $\varepsilon > 0$ and $m \in X$

$$\begin{aligned}\|\mathcal{F}(m)\|_2 \leq \sqrt{n_{obs}} \max_{i=1,\ldots,n_{obs}} \left|\mathrm{avg}_{x_i}(p)\right| &\leq \frac{\sqrt{n_{obs}\,|\Omega|}}{|b(0,R)|}\,\|p\|_0 \\ &\leq \frac{\sqrt{n_{obs}\,|\Omega|}}{|b(0,R)|}\frac{\|a\|}{\beta^2}\,\|f\|_0 \leq \frac{\sqrt{n_{obs}\,|\Omega|}}{|b(0,R)|}\frac{\|f\|_0}{\beta^2}C(\kappa)e^{\|m\|_\infty},\end{aligned}$$

where $C(\kappa)$ is a constant depending only on κ and we have used the bounds on $\|p\|_0$ and $\|a\|$ in the proof of Theorem 2. In addition, from Theorem 2, p, as a function of t_f, and hence m, is Lipschitz continuous, and as a result, the forward map is locally Lipschitz continuous in m, i.e.,

$$\left\|\mathcal{F}\left(m^1\right) - \mathcal{F}\left(m^2\right)\right\|_2 \leq C\left\|m^1 - m^2\right\|_X$$

for some constant C. The well-definedness of the Radon-Nikodym derivative in (24) then follows [22, Theorem 4.1] and the local Lipschitz continuity with respect to the data in the Hellinger distance follows [22, Theorem 4.2]. □

4.2. Analysis of a FEM approximation of the Bayesian posterior

In this section, we first approximate the forward (and hence the likelihood) of the Radon-Nikodym derivative (24) using the mixed FEM presented in section 3. We then approximate the parameter, and hence the prior measure, with a continuous FEM method. We then combine these two approximations to arrive at a FEM approximation of Bayesian posterior. Rigorous analysis of each approximation will be presented.

4.2.1. Likelihood approximation with mixed FEM

For clarity of the exposition, we consider only the mixed FEM with BDM element as the analysis for RTN element follows similarly. We denote by $\mathcal{F}_h(m)$ the observation operator Bp_h where $(\boldsymbol{u}_h, p_h)$ is the solution of (11) with $t_f = e^m$. The following result is a direct consequence of Theorem 3.

Lemma 2. *For any $R > 0$ in (26), and $m \sim \mu_{\mathrm{pr}}$, there exists a constant C independent of the mesh size h such that*

$$\|\mathcal{F}(m) - \mathcal{F}_h(m)\| \leq Ch^k e^{\|m\|_\infty}.$$

4.2.2. Prior approximation with continuous FEM

We denote by $m_{h,1} = \Pi_h m$ the discretization (interpolation) of m using the standard C^0 finite element method of degree $k_m \geq 1$ on the triangulation $\mathcal{E}_h^\Gamma$ of Γ (cf. [18]). Note that due to the low regularity of $m \in C^{0,r}(\Gamma)$, the usual $\mathcal{O}(h^{k_m})$ convergence rate of the FEM interpolation error does not apply here. To obtain such an estimate, we first recall an embedding result of Hölder spaces into Sobolev spaces: for (s,q), $s \geq 0$, $1 \leq q < \infty$ satisfying $(n-1)/q + s = r$ we have

$$\|m\|_{W^{s,q}(\Gamma)} \leq C\|m\|_{C^{0,r}(\Gamma)}, \tag{28}$$

for $C > 0$ depending only on Γ. By choosing a sufficiently large q, we can take $s < r$ as close to r as we desire. For fixed $0 \leq r < \min\{1, \sigma - (n-1)/2)\}$, we can use (28) to conclude that $m \in W^{s,q}(\Gamma)$ for (s,q) satisfying

$$s = r - \frac{\varepsilon}{2}, \quad \frac{n-1}{q} = \frac{\varepsilon}{2},$$

with an arbitrarily small $0 < \varepsilon \ll \min\{1, \sigma - (n-1)/2\}$. For this (ε, s, q), by [18, Theorem 4.4.20], we have

$$\|m - m_{h,1}\|_\infty \leq Ch^{s - \frac{n-1}{q}} \|m\|_{W^{s,q}(\Gamma)} = Ch^{r-\varepsilon} \|m\|_{W^{s,q}(\Gamma)}. \tag{29}$$

Here, the FEM interpolation order k_m is not involved[1] in this estimate due to $r < 1$.

The following is a simple application of Theorem 2 to the discrete system (11) together with the error bound (29), Lemma 2, and (28).

Lemma 3. *Let p_h and $p_{h,1}$ be the solutions of the discrete system (11) corresponding to m and $m_{h,1}$, respectively. Let $m \sim \mu_{\mathrm{pr}}$, and thus $m \in C^{0,r}(\Gamma)$, where $r \in (0, \min\{1, \sigma - (n-1)/2\})$, and $0 < \varepsilon \ll \min\{1, \sigma - (n-1)/2\}$. Then there exists a constant C independent of the meshsize h such that*

$$\|\mathcal{F}_h(m) - \mathcal{F}_h(m_{h,1})\| \leq Ch^{r-\varepsilon} e^{\|m\|_\infty},$$

[1] If $\sigma > (n-1)$, then m resides in $C^{0,r}$ for $0 < r < \min\{1, \sigma - (n-1)/2\}$ and $\|m\|_\infty \leq C\|m\|_{H^r(\Gamma)}$ for $(n-1)/2 < r < \sigma - (n-1)/2$. In this case, a higher order estimate by polynomials of degree k_m can be obtained as

$$\|m - m_{h,1}\|_\infty \leq Ch^{s - \frac{n-1}{2}} \|m\|_{H^s(\Gamma)}, \qquad \frac{n-1}{2} < s \leq \min\{r, k_m\},$$

by [18, Theorem 4.4.20].

and

$$\|\mathcal{F}(m) - \mathcal{F}_h(m_{h,1})\| \leq C\left(h^{r-\varepsilon} + h^k\right) e^{\|m\|_\infty}.$$

Since the interpolation Π_h is a linear operator on $m \in X$ and μ_{pr} is Gaussian, $m_{h,1} = \Pi_h m$ is distributed by the pushforward measure μ_{pr}^h induced by Π_h that is also a Gaussian. Let us denote by $\mu_{\text{pr}}^\perp$ the complement measure of μ_{pr}^h in $X := C^{0,r}(\Gamma)$ such that $\mu_{\text{pr}} = \mu_{\text{pr}}^h \bigotimes \mu_{\text{pr}}^\perp$, as this will be useful for the fully discrete Bayes formula presented below.

4.2.3. FEM approximation of the Bayesian posterior

Let us denote by $\mu_{\text{post}}^{h,1}$ the FEM approximation of the posterior measure μ_{post} using the likelihood and prior approximations in sections 4.2.1–4.2.2.

Theorem 5 (Well-posedness and convergence of FEM posterior $\mu_{\text{post}}^{h,1}$)**.** *The FEM posterior measure $\mu_{\text{post}}^{h,1}$ is absolutely continuous with respect to the prior measure μ_{pr} and the Radon-Nikodym derivative is given by*

$$\frac{d\mu_{\text{post}}^{h,1}}{d\mu_{\text{pr}}} \propto \exp\left(-\frac{1}{2}\|\mathcal{F}_h(m_{h,1}) - y_d\|^2_{\Gamma^{-1}_{\text{noise}}}\right). \tag{30}$$

Furthermore, there hold:

- *the FEM posterior measure $\mu_{\text{post}}^{h,1}$ is locally Lipschitz continuous with respect to the data y_d in the Hellinger distance.*
- *$d_{Hell}\left(\mu_{\text{post}}, \mu_{\text{post}}^{h,1}\right) \leq C\left(h^{r-\varepsilon} + h^k\right)$, where $d_{Hell}(\cdot,\cdot)$ is the Hellinger distance and $0 < \varepsilon \ll \min\{1, \sigma - (n-1)/2\}$. Thus, the FEM posterior $\mu_{\text{post}}^{h,1}$ converges to the true posterior measure μ_{post} as the mesh is refined.*

Proof. Similar to Lemma 1, it is sufficient to show that $\mathcal{F}_h(m_{h,1})$ is exponentially bounded in m and locally Lipschitz continuous with respect to m. For the boundedness, we have

$$\begin{aligned}\|\mathcal{F}_h(m_{h,1})\|_2 &\leq \frac{\sqrt{n_{obs}|\Omega|}}{|b(0,R)|}\|p_{h,1}\|_0 \leq \frac{\sqrt{n_{obs}|\Omega|}}{|b(0,R)|}\frac{\|a\|}{\beta^2}\|f\|_0 \\ &\leq \frac{\sqrt{n_{obs}|\Omega|}}{|b(0,R)|}\frac{\|f\|_0}{\beta^2}C(\kappa)e^{\|m_{h,1}\|_\infty} \leq C\frac{\sqrt{n_{obs}|\Omega|}}{|b(0,R)|}\frac{\|f\|_0}{\beta^2}C(\kappa)e^{\|m\|_\infty},\end{aligned}$$

where, in the second last inequality, we used the linearity and the boundedness of the finite element interpolant Π_h . For the Lipschitz continuity, Theorem 2 is also valid for $p_{h,1}$, and thus

$$\left\|\mathcal{F}_h\left(m_{h,1}^1\right) - \mathcal{F}_h\left(m_{h,1}^2\right)\right\| \leq C\left\|m_{h,1}^1 - m_{h,1}^2\right\|_\infty \leq C\left\|m^1 - m^2\right\|_\infty.$$

□

Now, owing to the fact that $\mu_{pr} = \mu_{pr}^h \bigotimes \mu_{pr}^\perp$, $\mu_{\text{post}}^{h,1}$ is absolutely continuous with respect to μ_{pr}, and the right hand side of (30) is a function of only $m_{h,1}$, we deduce that $\mu_{\text{post}}^{h,1} = \mu_{\text{post}}^h \bigotimes \mu_{\text{pr}}^\perp$, where

$$\frac{d\mu_{\text{post}}^h}{d\mu_{\text{pr}}^h} \propto \exp\left(-\frac{1}{2}\|\mathcal{F}_h(m_{h,1}) - y_d\|_{\Gamma_{\text{noise}}^{-1}}^2\right) \tag{31}$$

is a well-defined finite dimensional approximation of the infinite dimensional Bayes' formula (24). This expression can then be employed in the computations, including the Laplace approximation presented in the next section.

Remark 1. *It should be emphasized that $\mu_{\text{post}}^{h,1}$ is the FEM approximation of the posterior measure μ_{post} while μ_{post}^h is the computable part of $\mu_{\text{post}}^{h,1}$. Thus, unlike the forward error analysis presented in sections 3.1–3.2 for which we can verify the convergence rates numerically, it is not trivial how to verify the theoretical convergence result via Hellinger distance in Theorem 5 as $\mu_{\text{pr}}^\perp$—a measure on infinite dimensional space—is not known.*

4.3. MAP computation and Laplace approximation

The *Maximum a posteriori* (MAP) point of μ_{post} is a solution of

$$\underset{m\in\mathcal{M}}{\operatorname{argmin}} \frac{1}{2}\left(\|\mathcal{F}(m) - y_d\|_{\Gamma_{\text{noise}}^{-1}}^2 + \|m - m_{\text{pr}}\|_{\mathcal{M}}^2\right). \tag{32}$$

The existence of such a MAP point is a direct consequence of the compactness of the Cameron-Martin space $\mathcal{M}$ and the continuity of $\mathcal{F}(m)$.

Therefore the MAP computation can be considered as solving a deterministic inverse problem in which the regularization naturally comes from the prior distribution. When $d\mu_{\text{pr}}$ is Gaussian and the PDE model and $\mathcal{F}$ are linear, then the posterior distribution is Gaussian with mean m_{post} and covariance $\mathcal{C}_{\text{post}}$ of the form

$$m_{\text{post}} = m_{\text{MAP}}, \qquad \mathcal{C}_{\text{post}} = (\mathcal{F}^*\Gamma_{\text{noise}}^{-1}\mathcal{F} + \mathcal{C}_{\text{pr}}^{-1})^{-1} = \mathcal{H}(m_{\text{MAP}})^{-1}.$$

If the linearity assumptions on the PDE model and $\mathcal{F}$ are not true, then the posterior distribution is not Gaussian. Nonetheless it is reasonable to expect in many applications that the Laplace approximation $\mathcal{N}(m_{\text{MAP}}, \mathcal{H}(m_{\text{MAP}})^{-1})$ is still a good approximation of μ_{post}. An approach to compute the MAP point m_{MAP} efficiently using an inexact Newton method is discussed in Appendix Appendix A.

For the Laplace approximation we need to compute $\mathcal{H}(m_{\text{MAP}})^{-1}$ but it is prohibitively costly for infinite dimensional parameter spaces because the Hessian is a full matrix. For efficient approximation of $\mathcal{H}(m_{\text{MAP}})^{-1}$, we, similar to [23], use the relation

$$\mathcal{H}(m_{\text{MAP}})^{-1} = (\mathcal{H}_{\text{misfit}}(m_{\text{MAP}}) + \mathcal{C}_{\text{pr}}^{-1})^{-1},$$

where $\mathcal{H}_{\text{misfit}}(m_{\text{MAP}})$ is the Hessian of the misfit $\frac{1}{2}\|\mathcal{F}(m) - y_d\|^2_{\Gamma_{\text{noise}}^{-1}}$ at m_{MAP}. The dominant information gain from the data in $\mathcal{H}(m_{\text{MAP}})^{-1}$ compared to $\mathcal{C}_{\text{pr}}^{-1}$ can be estimated by finding the dominant eigen-pairs of $\mathcal{C}_{\text{pr}}^{-1/2}\mathcal{H}_{\text{misfit}}(m_{\text{MAP}})\mathcal{C}_{\text{pr}}^{-1/2}$. This allows us to replace $\mathcal{H}_{\text{misfit}}(m_{\text{MAP}})$ by a low-rank approximation with sufficient accuracy. The computation of low-rank approximation can be done independent of the parameter space dimension by randomized singular value decomposition algorithms [24, 25]. Then an approximation of $\mathcal{H}(m_{\text{MAP}})^{-1}$ can be computed efficiently using the Sherman–Morrison–Woodbury formula (see, e.g., [23, 14] for details).

5. Numerical results

In our numerical experiments we use hIPPYlib [13, 14] and FEniCS 2019.1.0 [12].

5.1. Forward problems

In this subsection we numerically assess the convergence rate of the FEM approximation of the subsurface flow model in faulted domains (forward model). In all numerical experiments below we use the lowest order Raviart–Thomas element and piecewise constant element for $\boldsymbol{u}$ and p, so the expected optimal convergence rates are 1 for $\boldsymbol{u}$ and p if they are in $H^1(\Omega; \mathbb{R}^s)$ and $H^1(\Omega)$. In the first numerical experiment we manufacture a solution that has enough regularity to achieve the expected rate of convergence. In the second numerical experiment we consider a fault geometry of practical interest and a manifactured solution $\boldsymbol{u}$ with the Sobolev regularity only H^s for some $s < 1$.

In the first numerical experiment of forward problems we consider a special manufactured solution with smooth functions on $\mathring{\Omega} := \Omega \setminus \Gamma$. Specifically, on the unit square domain $\Omega = [0,1]\times[0,1]$ with fault $\Gamma = \{(x,y) : x = 1/2\}$, let p, $\boldsymbol{u}$ be

$$p(x,y) = \begin{cases} x^3 & \text{if } x < \frac{1}{2} \\ x^3 + 1 & \text{if } x > \frac{1}{2} \end{cases}, \qquad \boldsymbol{u}(x,y) = (3x^2, 0). \tag{33}$$

Note that $\boldsymbol{u}$ can be extended to Γ as a smooth function on Ω, and one can check that (2c) is satisfied with $t_f = \frac{4}{3}$. Let Γ_L, Γ_R, Γ_T, Γ_B be the left, right, top, and bottom boundary parts of Ω. We impose Dirichlet boundary conditions on $\Gamma_L \cup \Gamma_R$ and Neumann boundary conditions on $\Gamma_T \cup \Gamma_B$ with the manufactured solution. For $f = \operatorname{div} \boldsymbol{u}$ convergence rates of numerical solutions are obtained in Table 1. Note that t_f satisfies the assumption in Theorem 4, so the results in Table 1 show the expected convergence rates in our error analysis.

	$\|\boldsymbol{u} - \boldsymbol{u}_h\|$		$\|p - p_h\|$	
$h_{\max}$	error	rate	error	rate
0.3750	2.25e-01	–	7.88e-02	–
0.1875	1.13e-01	0.99	3.93e-02	1.00
0.0938	5.67e-02	1.00	1.96e-02	1.00
0.0469	2.84e-02	1.00	9.80e-03	1.00
0.0234	1.42e-02	1.00	4.90e-03	1.00
0.0117	7.10e-03	1.00	2.45e-03	1.00

Table 1: Convergence rates for the example with the manufactured solution (34).

In the second numerical experiment we consider a domain $\Omega = [0,1]\times[0,1]$ with fault $\Gamma = \{1/2\} \times [1/4, 3/4] \subset \Omega$ (see the left panel of Figure 2). The manufactured solution (see the middle and right panels of Figure 2 for its graph) for this test case is given by

$$p(x,y) = \begin{cases} 0 & \text{if } y < \frac{1}{4} \text{ or } y > \frac{3}{4} \\ \sin\frac{3\pi x}{2}\cos^2\left(2\pi\left(y - \frac{1}{2}\right)\right) & \text{if } x < \frac{1}{2} \text{ and } \frac{1}{4} \geq y \leq \frac{3}{4} \\ -\sin\frac{3\pi(1-x)}{2}\cos^2\left(2\pi\left(y - \frac{1}{2}\right)\right) & \text{if } x > \frac{1}{2} \text{ and } \frac{1}{4} \geq y \leq \frac{3}{4}. \end{cases} \tag{34}$$

Setting $\kappa = 1$ we can compute $\boldsymbol{u} = -\nabla p$ and $f = \operatorname{div} \boldsymbol{u}$ on $\mathring{\Omega}$. It can be verified by direct computation that limits of the normal component of $\boldsymbol{u}$ on Γ is continuous across Γ and the condition $t_f \boldsymbol{u} \cdot \boldsymbol{n} - [\![p]\!] = 0$ holds with $t_f = 4/(3\pi)$. However, by computing $\partial p/\partial y$ of p in (34) and taking limits as $x \to \frac{1}{2}^+$ and $x \to \frac{1}{2}^-$, one can check that the tangential component of $\boldsymbol{u}$ is not continuous on Γ. Therefore, the regularity of $\boldsymbol{u}$ is lower than $H^1(\Omega; \mathbb{R}^2)$. By using Γ_L, Γ_R, Γ_T, Γ_B to denote the boundary parts as before, we impose boundary conditions

$$p = 0 \quad \text{on } \Gamma_L \cup \Gamma_R, \qquad \boldsymbol{u} \cdot \boldsymbol{n} = 0 \quad \text{on } \Gamma_T \cup \Gamma_B.$$

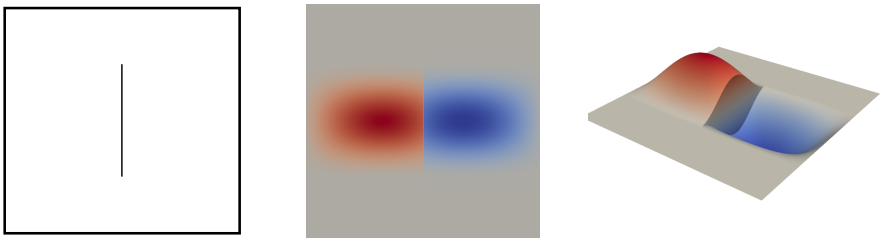

Figure 2: The domain with a vertical fault in numerical experiments (left figure) and the graphs of the pressure field in (34) (middle and right figures).

	$\|\boldsymbol{u} - \boldsymbol{u}_h\|$		$\|p - p_h\|$	
$h_{\max}$	error	rate	error	rate
0.3750	1.79e+00	–	1.53e-01	–
0.1875	1.06e+00	0.77	9.86e-02	0.64
0.0938	5.51e-01	0.94	4.67e-02	1.08
0.0469	2.96e-01	0.90	2.34e-02	1.00
0.0234	1.64e-01	0.86	1.19e-02	0.97
0.0117	9.50e-02	0.78	5.84e-03	1.03
0.0059	5.85e-02	0.70	2.98e-03	0.97

Table 2: Convergence rates for the example with the manufactured solution (34).

For error computation we take element-wise L^2 projection of the manufactured solutions to element-wise polynomial spaces of degree 4, and compute

the L^2 norms of the difference of these projected functions and numerical solutions. The errors and convergence rates due to mesh refinements are presented in Table 2. In the result the asymptotic convergence rates are not clear because we reach the limit of our computational resources due to the mesh refinements. However, it is clear that the convergence rates of $\boldsymbol{u}$ is lower than 1. This low convergence rates are expected because $\boldsymbol{u} \in H^s(\Omega; \mathbb{R}^2)$ for some $s < 1$. We also observe that the pressure errors seem to have higher convergence rates although this is not explained by our error analysis.

5.2. MAP computation as deterministic inverse problem

In our numerical tests for deterministic inverse problems we find the MAP point of the problem (32) using the inexact Newton-CG algorithm in Appendix A. The domain and fault are $\Omega = [0,1]\times[0,1]$, $\Gamma = \{1/2\}\times[1/4, 3/4]$, and boundary conditions are

$$p = 0 \quad \text{on } \Gamma_L, \qquad p = 1 \quad \text{on } \Gamma_R, \qquad \boldsymbol{u}\cdot\boldsymbol{n} = 0 \quad \text{on } \Gamma_T \cup \Gamma_B.$$

We set $\kappa = 1$ on Ω and assume that the true transmissibility field on Γ is $t_f = e^m$ with

$$m(y) = 2\sin\left(8\pi\left(y - \frac{1}{2}\right)\right), \qquad \frac{1}{4} \le y \le \frac{3}{4}. \tag{35}$$

For the prior we use $\mathcal{C}_{\text{pr}} = (\delta - \gamma\Delta)^{-1}$ with $\delta = 0.4$, $\gamma = 0.004$, and $m_{\text{pr}} = 0$. For the likelihood term, pointwise observation of pressure are measured at $k\times k$ ($k = 4, 6, 8$) uniform lattice points on the observation domain $[0.2, 0.8]\times[0.1, 0.9]$ as shown in Figure 3 for the case of 4×4 and 8×8 observation points. We corrupt the observed data vector $\boldsymbol{d} \in \mathbb{R}^N$, $N = 16, 36, 64$, by adding a Gaussian noise vector $\zeta \in \mathbb{R}^N$. Entries of ζ are independent identically distributed and follow a normal distribution $\mathcal{N}(0, \sigma^2)$ with $\sigma = 0.01\|\boldsymbol{d}\|_{l^\infty}$. Then, transmissibility inversion by constrained minimization is done with the noisy data $\boldsymbol{d} + \zeta$.

The approximate MAP points[2] by deterministic inversion for 4×4, 6×6, and 8×8 observation points are shown in Figure 4. Since 4×4 observations do not provide sufficient information to inform the inverse parameter, the

[2]Note that our optimization method is local in nature and thus approximate MAP points can be a local maximum of the posterior.

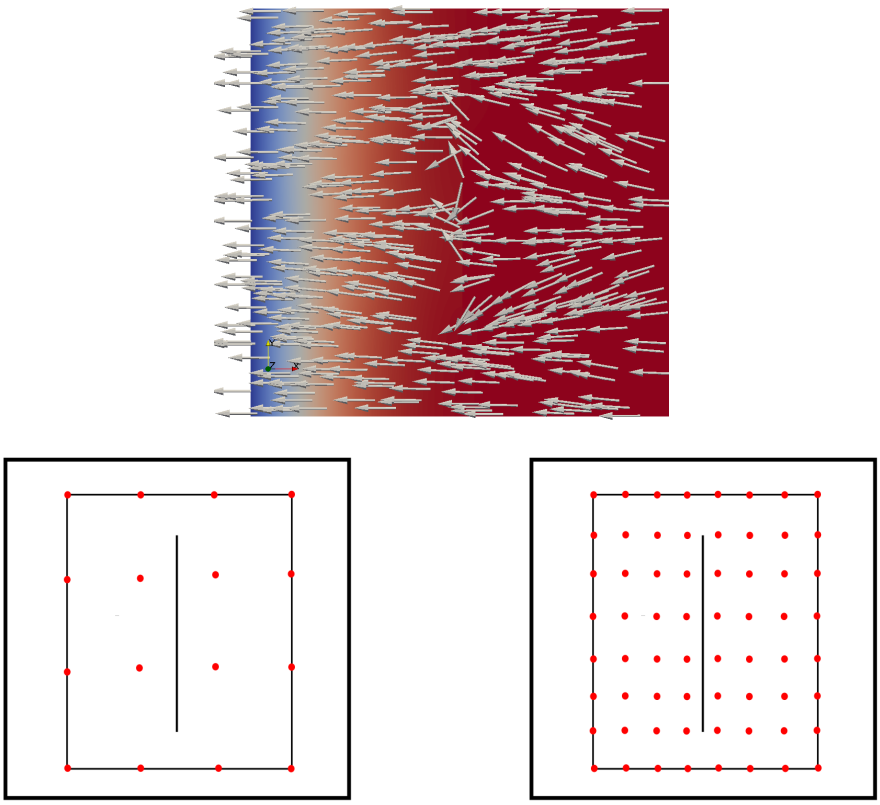

Figure 3: The pressure and velocity fields with fault parameter $t_f = e^m$ for $m = m_{\text{true}}$ in (35), and 4×4 and 8×8 observation points for fault parameter inversion.

approximate MAP point is not much different from the prior mean $m \equiv 0$ (also used as the initial guess in our algorithm). Nevertheless, the approximate MAP points with 6×6 and 8×8 observation points capture major feature of the true parameter field: clearly it is impossible to capture the true parameter field exactly due to limited and noisy data, the ill-posed nature of the inverse problem, and the smoothing effects of the prior.

5.3. Bayesian inverse problem

We limit ourselves to the Laplace approximation $\mathcal{N}(m_{\text{MAP}}, \mathcal{H}(m_{\text{MAP}})^{-1})$ as numerical results for Bayesian inversion. A more complete exploration of the Bayesian posterior can be done by standard sampling methods such as Markov chain Monte Carlo methods [26, 27, 28, 29]. While such sampling

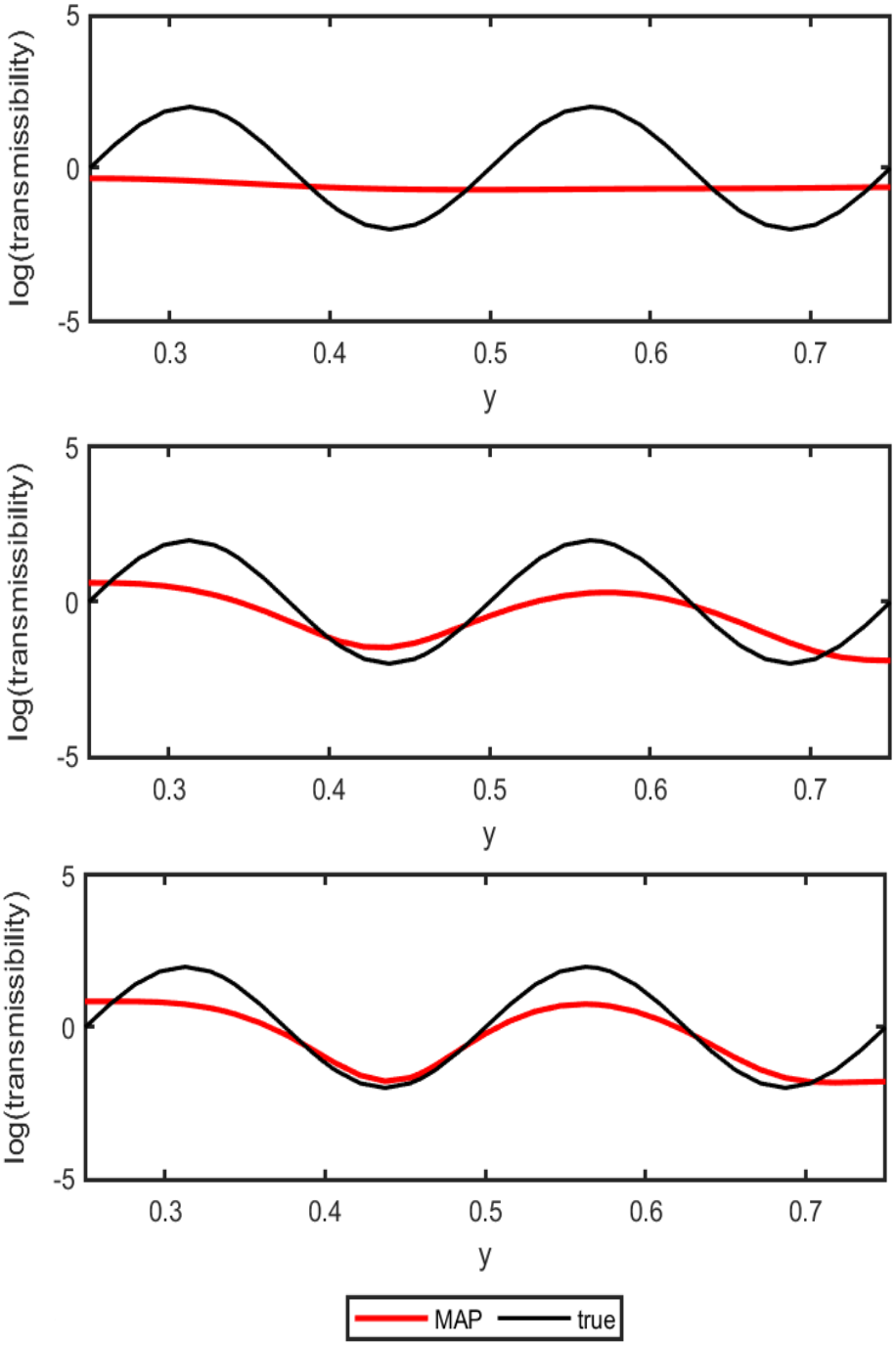

Figure 4: The $\log(t_f)$ fields of true parameter and approximate MAP points for 4×4, 6×6, and 8×8 observation points.

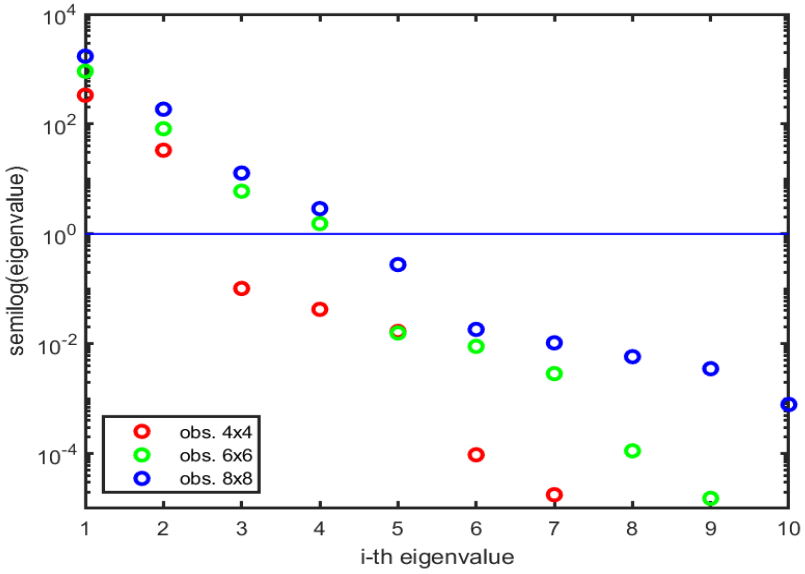

Figure 5: Eigenvalues corresponding to different sets of observation.

methods provide a full characterization of the discrete finite dimensional posterior (31), they do not provide additional insights or novelties to our work, and thus are not considered.

In the first experiment we consider a simple model with one vertical fault shown in Figure 3. We first present the dominant eigenvalues of prior-preconditioned Hessian of data misfit $\mathcal{C}_{\text{pr}}^{-1/2}\mathcal{H}_{\text{misfit}}(m_{\text{MAP}})\mathcal{C}_{\text{pr}}^{-1/2}$ in Figure 5. In all cases, at most four eigenvalues are greater than 1 indicating that only a handful of directions in parameter space can be inferred from the data with high confidence. This allows us to compute the Laplace approximation $\mathcal{N}(m_{\text{MAP}}, \mathcal{H}(m_{\text{MAP}})^{-1})$ efficiently with low rank approximation (see, e.g., [30] for more details).

In Figure 6, we compare the credibility intervals of

$$\mathcal{N}(0, \mathcal{C}_{\text{pr}}) \quad \text{and} \quad \mathcal{N}(m_{\text{MAP}}, \mathcal{H}(m_{\text{MAP}})^{-1})$$

for 4×4 observation points. The shaded (yellow) regions denote the union of pointwise credibility intervals with the length $2\sigma_{\text{std}}$ where σ_{std} are the empirical pointwise variances. We draw 5 random samples from $\mathcal{N}(0, \mathcal{C}_{\text{pr}})$ and $\mathcal{N}(m_{\text{MAP}}, \mathcal{H}(m_{\text{MAP}})^{-1})$ and plot them in the corresponding subfigures.

Similarly, in Figures 7 and 8 are the credibility regions of $\mathcal{N}(0, \mathcal{C}_{\rm pr})$ and $\mathcal{N}(m_{\rm MAP}, \mathcal{H}(m_{\rm MAP})^{-1})$ for the cases with 6×6 and 8×8 observation points, respectively. It is clear that $\mathcal{N}(m_{\rm MAP}, \mathcal{H}(m_{\rm MAP})^{-1})$ gives a much narrower credibility region than the one given by $\mathcal{N}(0, \mathcal{C}_{\rm pr})$. This is expected as there are more observational data. This in turns significantly reduces the uncertainty in the posterior distribution compared to the prior distribution. We also present 5 samples drawn from $\mathcal{N}(0, \mathcal{C}_{\rm pr})$ and $\mathcal{N}(m_{\rm MAP}, \mathcal{H}(m_{\rm MAP})^{-1})$: again those drawn from the Laplace approximation of the posterior are closer to the true parameter field.

In the second numerical experiment we present inverse results for a problem with 3 faults with 6×6 observations in the rectangle $[0.15, 0.85] \times [0.15, 0.85]$ as in Figure 9. The prior $\mathcal{C}_{\rm pr} = (\delta - \gamma\Delta)^{-1}$ with $\delta = 0.4$, $\gamma = 0.004$, and $m_{\rm pr} = 0$ is used for all faults. Observational data are synthetically generated using the same procedure as above, and the Laplace approximations on each fault are computed using low rank approximations. The credibility regions and 5 randomly drawn samples of $\mathcal{N}(0, \mathcal{C}_{\rm pr})$ and $\mathcal{N}(m_{\rm MAP}, \mathcal{H}(m_{\rm MAP})^{-1})$ are shown in Figure 10. This problem is more challenging compared to the one-fault case. As can be seen, the prior mean and samples are very different from the ground truth parameter. With the help of observational data, the posterior results using the Laplace approximation are closer to the ground truth parameter field for each fault. As can be seen, the inversion result for the second fault is better than the others. This is not surprising as the second fault is the longest and is surrounded by the largest number of observation points. In order to obtain better inverse results for faults 1 and 3, more observations are needed. Indeed, Figure 11 shows that the inverse results, both the MAP and uncertainty estimation, are more accurate with 12×12 observation points.

6. Conclusions

This paper concerns the forward and inverse modeling of subsurface flow in domains with faults. In particular, inversion of the fault transmissibility for subsurface flow models under appropriate assumptions on fault structures is considered. Unlike most of the existing works, we address many challenges associated with transmissibility inversion problems simultaneously including fault modeling, well-posedness of the forward PDE, FEM discretizations of the forward PDE and their rigorous analysis, well-posedness of the Bayesian inverse formulation, discretizations of Bayesian inverse formulation and their

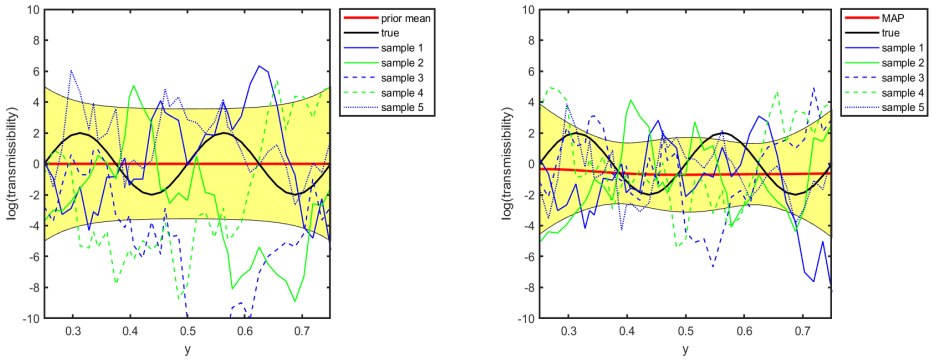

Figure 6: The comparison of pointwise credibility interval, mean, and samples with $\mathcal{N}(0, \mathcal{C}_{\text{pr}})$ and the Laplace approximation $\mathcal{N}(m_{\text{MAP}}, \mathcal{H}(m_{\text{MAP}})^{-1})$ with 4×4 observation points. Left is for the prior and right for the Laplace approximation.

rigorous analysis, and numerical illustrations. We also presented transmissibility inversion problems by a PDE-constrained optimization approach and developed an efficient numerical optimization method utilizing the efficient computation of Hessian action via Lagrangian approach. The results show that our approach can provide reasonable inverse solutions together with their associated uncertainty estimation. Ongoing work is to extend our framework to problem with more realistic networks of faults and three dimensional problems.

Appendix A. MAP computation with an inexact Newton method

Below we present an inexact Newton-Conjugate Gradient (Newton-CG) approach to compute the map point. For concreteness, we consider $n = 2$, and it is sufficient to take $\sigma = 1$ for the Gaussian prior to be well-defined. In this case the Cameron-Martin space is $\mathcal{M} = H^1(\Gamma)$ (in fact an equivalence of $H^1(\Gamma)$ with δ and γ as weights: see the definition $\mathcal{J}$ below). Then the MAP problem becomes the following: solve

$$\operatorname*{argmin}_{m \in \mathcal{M}} \mathcal{J}(m)$$

with

$$\begin{aligned}\mathcal{J}(m) := \frac{1}{2}\left(\|Bp - y_d\|^2_{\Gamma^{-1}_{\text{noise}}} + (\langle \delta(m - \bar{m}), (m - \bar{m})\rangle_\Gamma)\right) \\ + \frac{1}{2}\langle \gamma \nabla(m - \bar{m}), \nabla(m - \bar{m})\rangle_\Gamma\end{aligned}$$

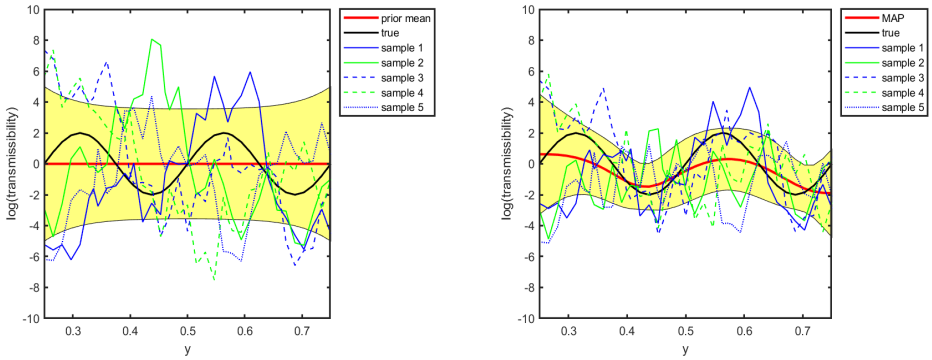

Figure 7: The comparison of pointwise credibility interval, mean, and samples with $\mathcal{N}(0, \mathcal{C}_{\text{pr}})$ and the Laplace approximation $\mathcal{N}(m_{\text{MAP}}, \mathcal{H}(m_{\text{MAP}})^{-1})$ with 6×6 observation points. Left is for the prior and right for the Laplace approximation.

where $(\boldsymbol{u}, p)$ is the solution of forward equation

$$\begin{aligned} \left(\kappa^{-1}\boldsymbol{u}, \boldsymbol{v}\right) + \langle e^m \boldsymbol{u}\cdot\boldsymbol{n}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle - (p, \operatorname{div}\boldsymbol{v}) = 0, \qquad & \boldsymbol{v} \in \boldsymbol{V}, \qquad \text{(A.1a)} \\ (\operatorname{div}\boldsymbol{u}, q) = (f, q) \qquad & q \in Q. \qquad \text{(A.1b)} \end{aligned}$$

We solve this problem using inexact Newton-type methods with the gradient and Hessian-vector project as in [30, 31]. To compute the gradient of $\mathcal{J}(m)$ we use the Lagrangian functional

$$\begin{aligned} &\mathcal{L}^{\mathcal{G}}((\boldsymbol{u}, p), m, (\boldsymbol{v}, q)) \\ &= \mathcal{J}(m) + \left(\kappa^{-1}\boldsymbol{u}, \boldsymbol{v}\right) + \langle e^m \boldsymbol{u}\cdot\boldsymbol{n}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_\Gamma - (p, \operatorname{div}\boldsymbol{v}) + (\operatorname{div}\boldsymbol{u}, q) - (f, q)\,. \end{aligned}$$

with $(\boldsymbol{u}, p), (\boldsymbol{v}, q) \in \boldsymbol{V} \times Q$. The Lagrangian approach for the gradient computation at $m = m_0$ leads to the forward equation (A.1) and the adjoint problem: finding $(\boldsymbol{v}_0, q_0) \in \boldsymbol{V} \times Q$ such that for all $(\tilde{\boldsymbol{u}}, \tilde{p}) \in \boldsymbol{V} \times Q$

$$\begin{aligned} \langle B\tilde{p}, Bp_0 - y_d\rangle_{\Gamma_{\text{noise}}^{-1}} + \left(\kappa^{-1}\tilde{\boldsymbol{u}}, \boldsymbol{v}_0\right) + \langle e^m \tilde{\boldsymbol{u}}\cdot\boldsymbol{n}, \boldsymbol{v}_0\cdot\boldsymbol{n}\rangle_\Gamma - (\tilde{p}, \operatorname{div}\boldsymbol{v}_0) \\ + (\operatorname{div}\tilde{\boldsymbol{u}}, q_0) = 0, \quad \text{(A.2)} \end{aligned}$$

where $(\boldsymbol{u}_0, p_0)$ is the forwar solution obtained from (A.1) for $m = m_0$. Then the gradient of $\mathcal{J}(m)$ at $m = m_0$ in the weak form is: for any $\tilde{m} \in \mathcal{M}$

$$(\mathcal{G}(m_0), \tilde{m}) = \frac{1}{2}\left(\langle \delta(m_0 - \bar{m}), \tilde{m}\rangle_\Gamma + \langle \gamma\nabla m_0, \nabla\tilde{m}\rangle_\Gamma\right) + \langle \tilde{m}e^{m_0}\boldsymbol{u}_0\cdot\boldsymbol{n}, \boldsymbol{v}_0\cdot\boldsymbol{n}\rangle_\Gamma\,. \tag{A.3}$$

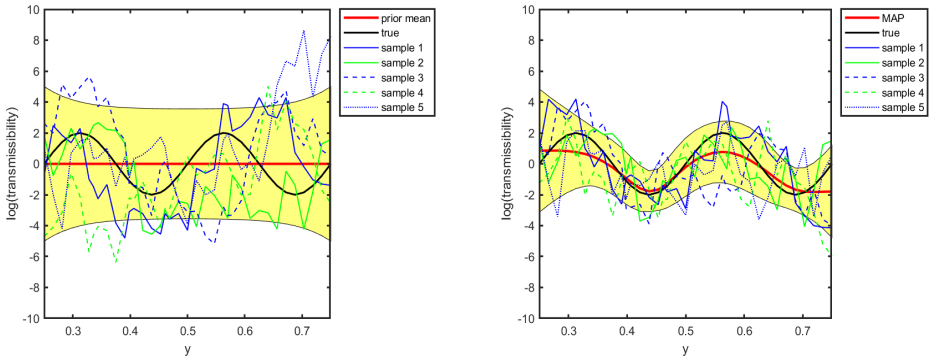

Figure 8: The comparison of pointwise credibility interval, mean, and samples with $\mathcal{N}(0, \mathcal{C}_{\text{pr}})$ and the Laplace approximation $\mathcal{N}(m_{\text{MAP}}, \mathcal{H}(m_{\text{MAP}})^{-1})$ with 8×8 observation points. Left is for the prior and right for the Laplace approximation.

For the Hessian-vector product we define $\mathcal{L}^{\mathcal{H}}$ as

$$\begin{aligned}
&\mathcal{L}^{\mathcal{H}}\left((\boldsymbol{u}, p), m, (\boldsymbol{v}, q); (\hat{\boldsymbol{u}}, \hat{p}), \hat{m}, (\hat{\boldsymbol{v}}, \hat{q})\right) \\
&= (\mathcal{G}(m), \hat{m}) + \left(\kappa^{-1}\boldsymbol{u}, \hat{\boldsymbol{v}}\right) + \langle e^m \boldsymbol{u} \cdot \boldsymbol{n}, \hat{\boldsymbol{v}} \cdot \boldsymbol{n}\rangle - (p, \operatorname{div} \hat{\boldsymbol{v}}) + (\operatorname{div} \boldsymbol{u}, \hat{q}) - (f, \hat{q}) \\
&\quad + \langle B\hat{p}, Bp - y_d\rangle_{\Gamma_{\text{noise}}^{-1}} + \left(\kappa^{-1}\hat{\boldsymbol{u}}, \boldsymbol{v}\right) + \langle e^m \hat{\boldsymbol{u}} \cdot \boldsymbol{n}, \boldsymbol{v} \cdot \boldsymbol{n}\rangle_{\Gamma} - (\hat{p}, \operatorname{div} \boldsymbol{v}) + (\operatorname{div} \hat{\boldsymbol{u}}, q)
\end{aligned}$$

The variational forms of the incremental forward and incremental adjoint equations are

$$(\kappa^{-1}\hat{\boldsymbol{u}}, \tilde{\boldsymbol{v}}) + \langle e^m \hat{\boldsymbol{u}} \cdot \boldsymbol{n}, \tilde{\boldsymbol{v}} \cdot \boldsymbol{n}\rangle_{\Gamma}$$

Thus, for the Hessian of $\mathcal{J}$ evaluated at $m = m_0$, the action of $\mathcal{H}(m_0)$ to $\hat{m} \in \mathcal{M}$ can be written in the weak form as

$$\begin{aligned}
(\tilde{m}, \mathcal{H}(m_0)\hat{m}) = &\langle \delta\hat{m}, \tilde{m}\rangle_{\Gamma} + \langle \gamma\nabla\hat{m}, \nabla\tilde{m}\rangle_{\Gamma} + \langle \tilde{m}\hat{m}e^{m_0}\boldsymbol{u}_0 \cdot \boldsymbol{n}, \boldsymbol{v}_0 \cdot \boldsymbol{n}\rangle_{\Gamma} \\
&+ \left(\kappa^{-1}\boldsymbol{u}_0, \hat{\boldsymbol{v}}\right) + \langle \tilde{m}e^{m_0}\boldsymbol{u}_0 \cdot \boldsymbol{n}, \hat{\boldsymbol{v}} \cdot \boldsymbol{n}\rangle_{\Gamma} \\
&- (p_0, \operatorname{div} \hat{\boldsymbol{v}}) + (\operatorname{div} \boldsymbol{u}_0, \hat{q}) - (f, \hat{q}) \\
&+ \langle \tilde{m}e^{m_0}\hat{\boldsymbol{u}} \cdot \boldsymbol{n}, \boldsymbol{v}_0 \cdot \boldsymbol{n}\rangle_{\Gamma}
\end{aligned}$$

Based on the gradient and the Hessian action computation algorithms, the maximum a posteriori (MAP) point can be found by a standard Newton method. For efficient performance of Newton methods an inexact Newton-CG algorithm can be used. In the inexact Newton-CG algorithm the system $\mathcal{H}(m_k)\delta_k = -\mathcal{G}(m_k)$ for Newton method is solved by a preconditioned CG

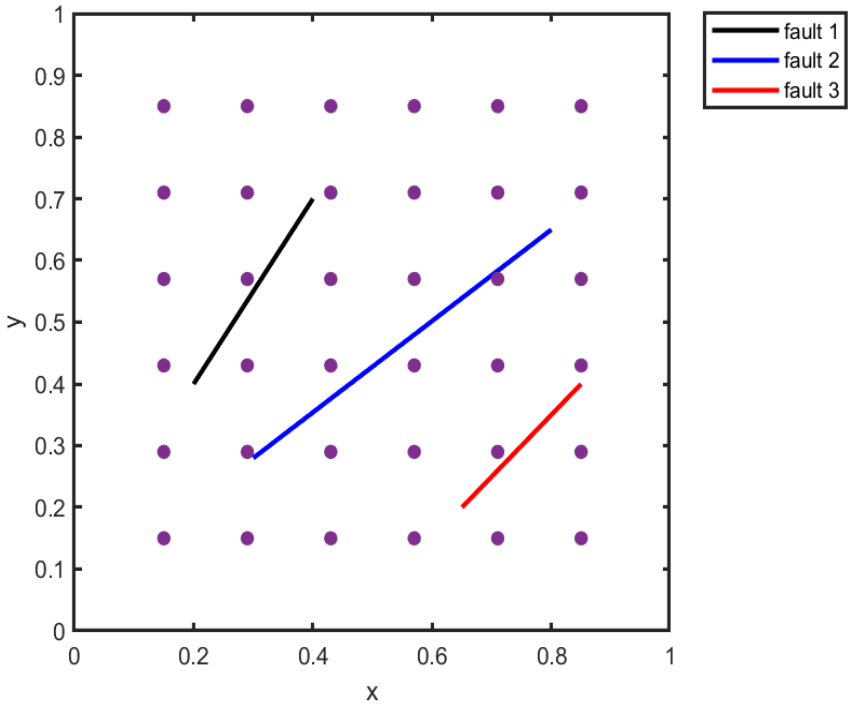

Figure 9: A domain with three faults and observation points.

method with early termination of CG iteration using Eisenstat–Walker criterion (to prevent oversolving) and backtracking algorithm using Steihaug criterion (to avoid negative curvature) (see, e.g.,[32, 14] and the references therein). The full algorithm is presented in Algorithm Appendix A.

$i \leftarrow 0$
while $i < max\_iter$ **do**
  Given m_i solve (A.1) to obtain $(\boldsymbol{u}_i, p_i)$
  Compute $\mathcal{J}_i := \mathcal{J}(m_i)$ using m_0 and $(\boldsymbol{u}_i, p_i)$
  Given m_i and $(\boldsymbol{u}_i, p_i)$ solve (A.2) and obtain $(\boldsymbol{v}_i, q_i)$
  Evaluate $\boldsymbol{g}_i$, the gradient of $\mathcal{J}$ at m_i, using (A.3)
  if $\|\boldsymbol{g}_i\| \leq tol_g$ **then**
    break
  end if

Find a search direction $\hat{m}_i$ such that

$$\|\mathcal{H}_i \hat{m}_i + \boldsymbol{g}_i\| \leq \eta_i \|\boldsymbol{g}_i\|, \qquad \eta_i := (\|\boldsymbol{g}_i\|/\|\boldsymbol{g}_0\|)^{1/2}$$

$j \leftarrow 0$, $\alpha^{(0)} \leftarrow 1$
while $j < max\_backtracking\_iter$ **do**
  Set $m^{(j)} = m_i + \alpha^{(j)} \hat{m}_i$
  Given $m^{(j)}$ solve (A.1) to obtain $(\boldsymbol{u}^{(j)}, p^{(j)})$
  Compute $\mathcal{J}^{(j)}$ using $m^{(j)}$ and $(\boldsymbol{u}^{(j)}, p^{(j)})$
  if $\mathcal{J}^{(j)} < \mathcal{J}^j + \alpha^{(j)} c_{\text{armijo}} \boldsymbol{g}_i^T \hat{m}_i$ **then**
    $m_{i+1} \leftarrow m^{(j)}, \mathcal{J}_{i+1} \leftarrow \mathcal{J}^{(j)}$
    break
  end if
  $\alpha^{(j+1)} \leftarrow \alpha^{(j)}/2, \quad j \leftarrow j + 1$
end while
$i \leftarrow i + 1$
end while

References

[1] P. M. Adler, J.-F. Thovert, V. V. Mourzenko, Fractured porous media, Oxford University Press, Oxford, 2013.

[2] Models of Fractures and Fractured Porous Media, John Wiley & Sons, Ltd, Ch. 8, pp. 213–251. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527636693.ch8, doi:https://doi.org/10.1002/9783527636693.ch8.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9783527636693.ch8

[3] Z. Chen, G. Huan, Y. Ma, Computational methods for multiphase flows in porous media, Vol. 2 of Computational Science & Engineering, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006. doi:10.1137/1.9780898718942.
URL https://doi.org/10.1137/1.9780898718942

[4] K.-A. Lie, An Introduction to Reservoir Simulation Using MATLAB/GNU Octave: User Guide for the MATLAB Reservoir Simulation Toolbox (MRST), Cambridge University Press, 2019. doi:10.1017/9781108591416.

[5] I. Berre, F. Doster, E. Keilegavlen, Flow in fractured porous media: a review of conceptual models and discretization approaches 130 (2019) 215–236. doi:10.1007/s11242-018-1171-6.
URL https://doi.org/10.1007/s11242-018-1171-6

[6] V. Martin, J. Jaffré, J. E. Roberts, Modeling fractures and barriers as interfaces for flow in porous media, SIAM J. Sci. Comput. 26 (5) (2005) 1667–1691. doi:10.1137/S1064827503429363.
URL https://doi.org/10.1137/S1064827503429363

[7] P. Angot, F. Boyer, F. Hubert, Asymptotic and numerical modelling of flows in fractured porous media, M2AN Math. Model. Numer. Anal. 43 (2) (2009) 239–275. doi:10.1051/m2an/2008052.
URL https://doi.org/10.1051/m2an/2008052

[8] E. Ahmed, J. Jaffré, J. E. Roberts, A reduced fracture model for two-phase flow with different rock types, Mathematics and Computers in Simulation 137 (2017) 49–70, mAMERN VI-2015: 6th International Conference on Approximation Methods and Numerical Modeling in Environment and Natural Resources. doi:https://doi.org/10.1016/j.matcom.2016.10.005.
URL https://www.sciencedirect.com/science/article/pii/S0378475416301987

[9] H. M. Nilsen, J. R. Natvig, K.-A. Lie, Accurate modeling of faults by multipoint, mimetic, and mixed methods, SPE Journal 17 (2) (2012) 568–579. doi:10.2118/149690-PA.
URL https://doi.org/10.2118/149690-PA

[10] P. F. Antonietti, L. Formaggia, A. Scotti, M. Verani, N. Verzott, Mimetic finite difference approximation of flows in fractured porous media, ESAIM Math. Model. Numer. Anal. 50 (3) (2016) 809–832. doi:10.1051/m2an/2015087.
URL https://doi.org/10.1051/m2an/2015087

[11] Website, http://fenicsproject.org/.

[12] A. Logg, K.-A. Mardal, G. N. Wells (Eds.), Automated solution of differential equations by the finite element method, Vol. 84 of Lecture Notes in Computational Science and Engineering, Springer, Heidelberg, 2012,

the FEniCS book. doi:10.1007/978-3-642-23099-8.
URL http://dx.doi.org/10.1007/978-3-642-23099-8

[13] U. Villa, N. Petra, O. Ghattas, hIPPYlib: an Extensible Software Framework for Large-scale Deterministic and Bayesian Inverse Problems, Journal of Open Source Software 3 (30) (2018). doi:10.21105/joss.00940.

[14] U. Villa, N. Petra, O. Ghattas, HIPPYlib: An Extensible Software Framework for Large-Scale Inverse Problems Governed by PDEs: Part I: Deterministic Inversion and Linearized Bayesian Inference, ACM Trans. Math. Softw. 47 (2) (Apr. 2021). doi:10.1145/3428447.
URL https://doi.org/10.1145/3428447

[15] V. Girault, P.-A. Raviart, Finite element methods for Navier-Stokes equations, Vol. 5 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1986, theory and algorithms. doi:10.1007/978-3-642-61623-5.
URL http://dx.doi.org/10.1007/978-3-642-61623-5

[16] J. Könnö, D. Schötzau, R. Stenberg, Mixed finite element methods for problems with Robin boundary conditions, SIAM J. Numer. Anal. 49 (1) (2011) 285–308. doi:10.1137/09077970X.
URL https://doi.org/10.1137/09077970X

[17] D. Boffi, F. Brezzi, M. Fortin, Mixed finite element methods and applications, Vol. 44 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2013. doi:10.1007/978-3-642-36519-5.
URL http://dx.doi.org/10.1007/978-3-642-36519-5

[18] S. C. Brenner, L. R. Scott, The mathematical theory of finite element methods, 3rd Edition, Vol. 15 of Texts in Applied Mathematics, Springer, New York, 2008. doi:10.1007/978-0-387-75934-0.
URL https://doi.org/10.1007/978-0-387-75934-0

[19] D. A. Di Pietro, A. Ern, Mathematical aspects of discontinuous Galerkin methods, Vol. 69 of Mathématiques & Applications (Berlin) [Mathematics & Applications], Springer, Heidelberg, 2012. doi:10.1007/978-3-642-22980-0.
URL https://doi.org/10.1007/978-3-642-22980-0

[20] R. A. Adams, Sobolev spaces, Pure and Applied Mathematics, Vol. 65, Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[21] L. C. Evans, Partial differential equations, Vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 1998.

[22] A. M. Stuart, Inverse problems: a Bayesian perspective, Acta Numer. 19 (2010) 451–559. doi:10.1017/S0962492910000061.
URL https://doi.org/10.1017/S0962492910000061

[23] T. Bui-Thanh, O. Ghattas, J. Martin, G. Stadler, A computational framework for infinite-dimensional Bayesian inverse problems Part I: The linearized case, with application to global seismic inversion, SIAM J. Sci. Comput. 35 (6) (2013) A2494–A2523. doi:10.1137/12089586X.
URL https://doi.org/10.1137/12089586X

[24] N. Halko, P. G. Martinsson, J. A. Tropp, Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions, SIAM Rev. 53 (2) (2011) 217–288. doi:10.1137/090771806.
URL https://doi.org/10.1137/090771806

[25] A. K. Saibaba, J. Lee, P. K. Kitanidis, Randomized algorithms for generalized Hermitian eigenvalue problems with application to computing Karhunen-Loève expansion, Numer. Linear Algebra Appl. 23 (2) (2016) 314–339. doi:10.1002/nla.2026.
URL https://doi.org/10.1002/nla.2026

[26] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, Equation of state calculations by fast computing machines, The Journal of Chemical Physics 21 (6) (1953) 1087–1092. doi:10.1063/1.1699114.
URL http://dx.doi.org/10.1063/1.1699114

[27] W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika 57 (1) (1970) 97–109.

[28] J. Martin, L. C. Wilcox, C. Burstedde, O. Ghattas, A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion, SIAM Journal on Scientific Computing 34 (3) (2012) A1460–A1487. doi:10.1137/110845598.

[29] T. Bui-Thanh, M. A. Girolami, Solving large-scale PDE-constrained Bayesian inverse problems with Riemann manifold Hamiltonian Monte Carlo, Inverse Problems Special Issue (2014) 114014.

[30] T. Bui-Thanh, C. Burstedde, O. Ghattas, J. Martin, G. Stadler, L. C. Wilcox, Extreme-scale UQ for Bayesian inverse problems governed by PDEs, in: SC12: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 2012.

[31] T. Bui-Thanh, C. Burstedde, O. Ghattas, J. Martin, G. Stadler, L. C. Wilcox, Extreme-scale UQ for Bayesian inverse problems governed by PDEs, in: SC12: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 2012, Gordon Bell Prize finalist.

[32] J. Nocedal, S. J. Wright, Numerical Optimization, 2nd Edition, Springer Verlag, Berlin, Heidelberg, New York, 2006.

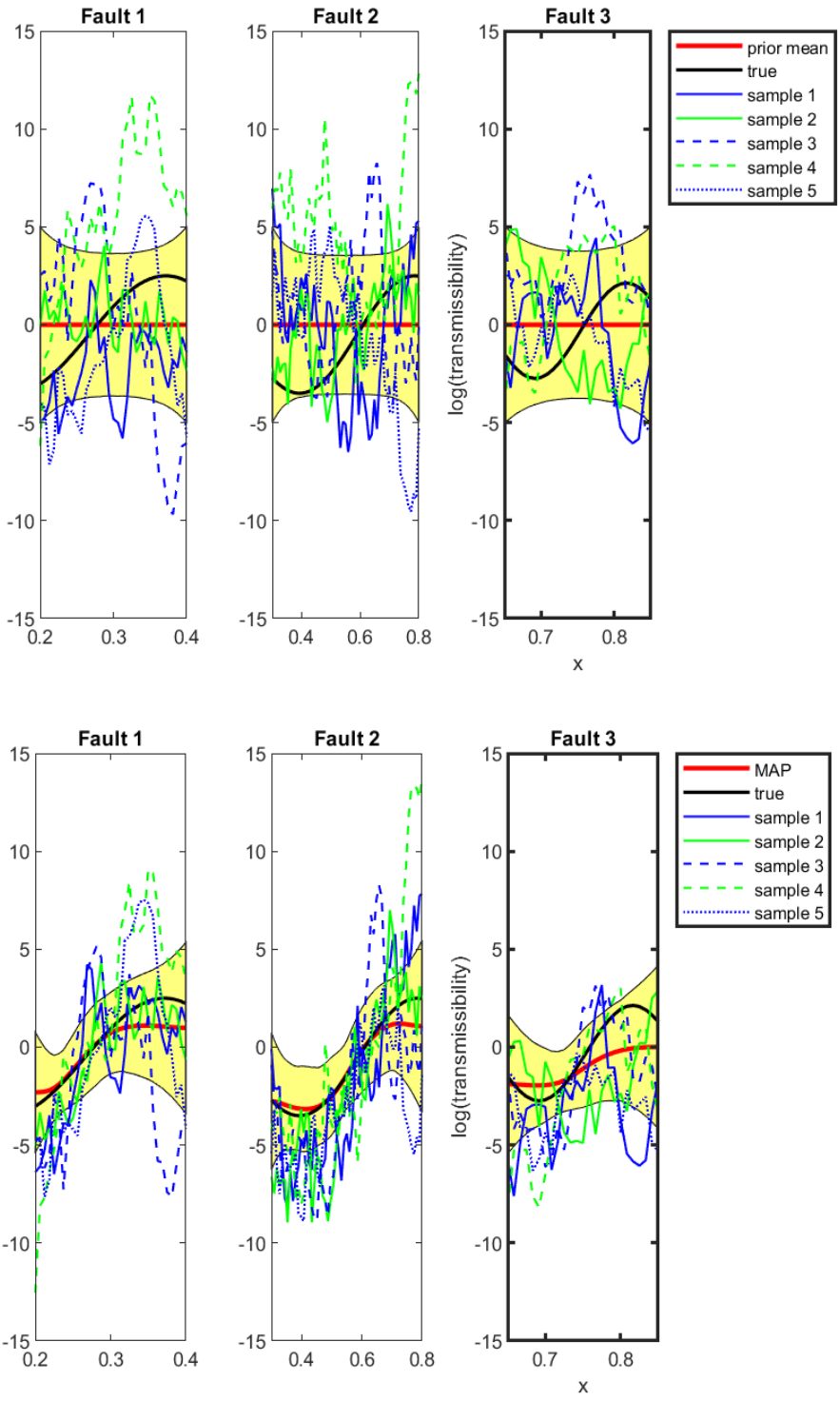

Figure 10: The comparison of pointwise credibility interval, mean, and samples with $\mathcal{N}(0, \mathcal{C}_{\text{pr}})$ and the Laplace approximation $\mathcal{N}(m_{\text{MAP}}, \mathcal{H}(m_{\text{MAP}})^{-1})$ with 6×6 observation points for the problem with three faults. The top row is for the prior and the bottom one is for the posterior using the Laplace approximation.

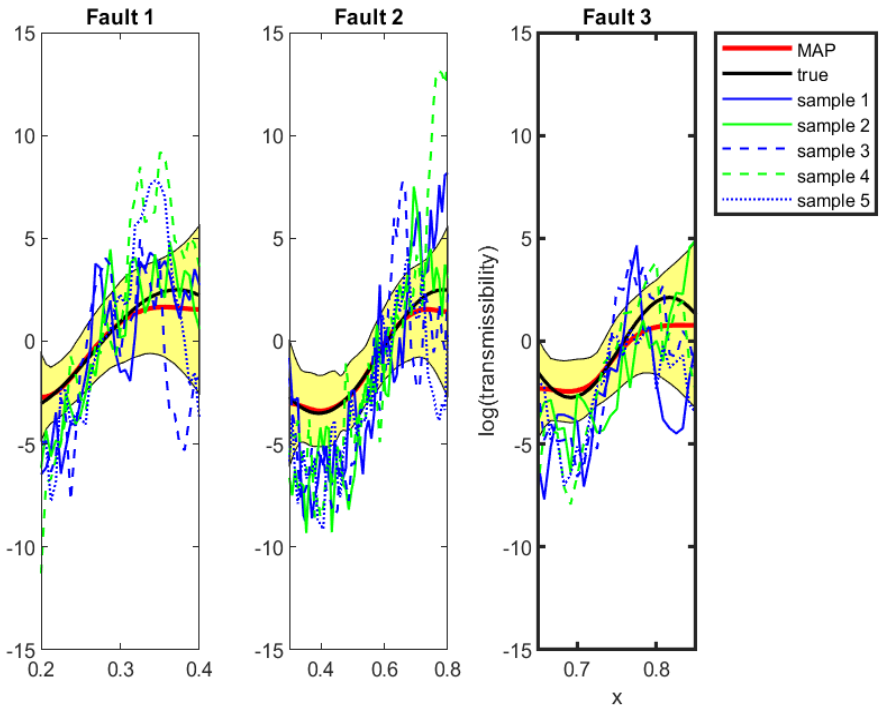

Figure 11: The credibility interval, mean, and samples with the Laplace approximation $\mathcal{N}(m_{\mathrm{MAP}}, \mathcal{H}(m_{\mathrm{MAP}})^{-1})$ obtained with 12×12 observation points for the problem with three faults.