Binary recovery via phase field regularization for first-arrival traveltime tomography

Oliver R. A. Dunbar [1], **Charles M. Elliott**[2]

Mathematics Institute, University of Warwick, UK, CV4 7AL

E-mail: [1] o.dunbar.1@warwick.ac.uk, [2] c.m.elliott@warwick.ac.uk

Abstract. We propose a double obstacle phase field methodology for binary recovery of the slowness function of an Eikonal equation found in first-arrival traveltime tomography. We treat the inverse problem as an optimization problem with quadratic misfit functional added to a phase field relaxation of the perimeter penalization functional. Our approach yields solutions as we account for well posedness of the forward problem by choosing regular priors. We obtain a convergent finite difference and mixed finite element based discretization and a well defined descent scheme by accounting for the non-differentiability of the forward problem. We validate the phase field technique with a Γ – convergence result and numerically by conducting parameter studies for the scheme, and by applying it to a variety of test problems with different geometries, boundary conditions, and source – receiver locations.

1. Introduction

We consider recovery of subsurface structure formed by a disjoint composition of two materials with distinct impedances to ground waves. First-arrival traveltime tomography (FATT) (see [46]) is a subfield of seismic tomography where the observational data for recovery is the first hitting time of a wave at known locations. We formulate the forward problem of FATT as an Eikonal equation ([51, Appendix C] or [3, Chapter 4]), and we use a phase field regularization [10] to model the binary nature of the problem.

The link between the Eikonal equation and first-arrival traveltime problems, has been widely studied [52, 43, 17, 20], where it is shown that if the wave *impedance* or *slowness function* is continuous (or with specific forms of jump discontinuities [21]) then the description admits unique Lipschitz continuous viscosity solutions. The discontinuities arise at boundaries of material regions of different wave impedances [58].

The inverse problem of recovering the slowness function given observed traveltimes is illposed. We work in an optimization framework, and overcome determination issues by restricting optimization of a quadratic data *fidelity* or *misfit* functional over a *prior* or *model subspace* of functions, and we add a regularization functional. For piecewise

constant slowness function we seek to find the regions in which the function is constant; this is equivalent to determining where interfaces between constant values occur. A natural regularization in the optimization framework is to penalize the length the interfaces between constant values [47, 57, 38]. This perimeter regularization is difficult for binary priors because of non-differentiabilty unless one tracks the interfaces. In addition, the solution of the recovery problem is a binary function and unfortunately, the forward problem is not well posed for this class of functions.

In this paper we address these two issues and present a method based on phase field techniques [12] for binary recovery. The slowness is regularized using a phase field function and the perimeter regularization is approximated by a phase field gradient energy. We take an obstacle phase field approach [10, 12] with an H^2 regular functional [29], (for other choices [18, 26]). This method has several attractive features. Primarily it is mathematically grounded, by producing continuous priors for the slowness function so the forward problem is well posed within the theory of viscosity solutions. This leads to an existence theory for solutions of the inverse problem. It is fit for purpose, as the regularization serves as an approximation to the perimeter regularization functional and so is naturally suited to binary recovery, [45, 6, 10]. We are able to construct a convergent discretization scheme for approximating solutions to the inverse problem. This is based on a monotone finite difference method (forward problem) and a mixed finite element method (phase field functional). It is shown in [22] that for a particular finite difference scheme, a derivative of the discrete forward problem exists, and is computed efficiently by a variant of the fast marching method on the discrete adjoint equation. We implement the scheme, using an adjoint equation for the efficient calculation of (a well defined) discrete derivative, and showcase it with a parameter study and apply it to several varied test configurations.

The power of our method comes from a notion of convergence of smooth minimizers to binary functions [45, 6, 10], a well developed numerical analysis [11, 9] and suitability for implementation. It is possible to extend the modelling framework, to one with a constant region linked to a smoothly varying region across an interface jump as in [2]. Although for the forward problem we use the fast marching method (FMM) [49, 50, 48, 1] due to its robustness [30], fast sweeping methods [55, 35, 44, 32] could be used. Also we may use finite difference (rather than finite element) approximations of the phase field functional. Related phase field techniques have been applied to piecewise constant recovery for other inverse problems [24, 13, 23, 7].

With respect to other approaches we note that there are many other representations of the interfaces bounding constant regions. For example, a finite parameter specification [27, 41], or with a finite basis [22] or infinite dimensional subspace [57, 38]. Also, to construct the inverse problem, a probabilistic framework may be used [53, 31, 15], where the prior is a space of probability distributions. Other regularizations to enforce so-called 'blocky' solutions are also used [33, 34, 58]. We also note that, although the solution of the Eikonal equation is not differentiable with respect to the slowness in general, other studies have assumed differentiability to improve efficiency of schemes, as one may then

use an adjoint equation or write optimality conditions [37, 54].

1.0.1. Outline We set out the forward problem in Section 2.1 and inverse problem Section 2.4. We present the phase field regularization in Section 3.2. We discretize the problems in Section 4.1 and Section 4.3, with attention given to the discrete derivative in Section 4.4. We present a numerical scheme and investigate choice of parameters, and the scheme's effectiveness through computations in Section 5.

1.1. Binary recovery

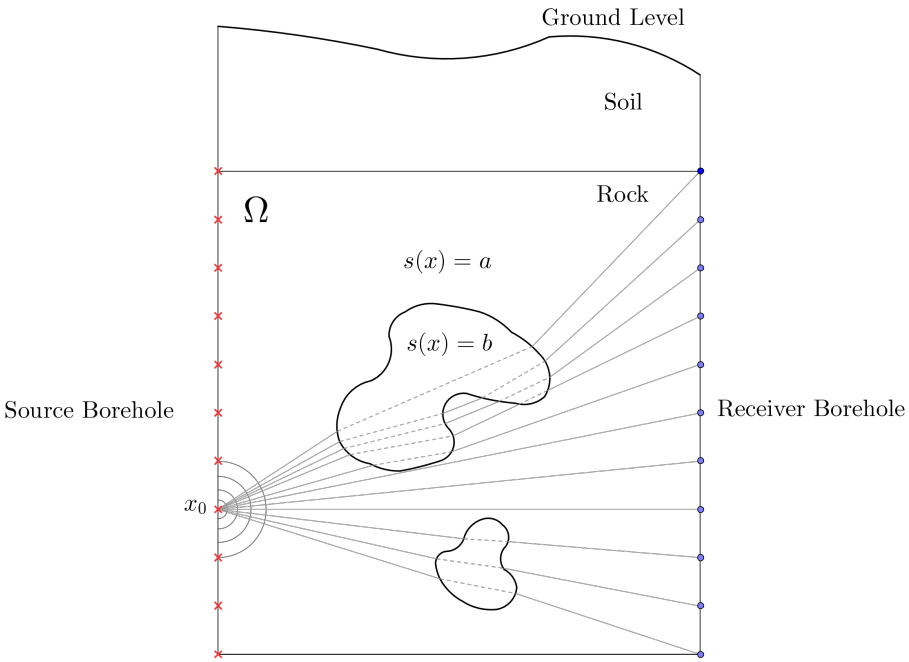

Figure 1. Schematic showing crosswell tomography for binary recovery. The boundaries are drilled boreholes, the left populated with wave sources, the right with receivers. A source at x_0 is producing a wave that moves through a background medium with slowness $s(x) = a$, and through inclusions where $s(x) = b$. Drawing solution ray paths, one sees Snell's law at inclusion boundaries. The Soner boundary condition requires that rays leaving the domain Ω do not re-enter.

We motivate the choices we make in subsequent sections using the following model. Let $\Omega = [0, L_x] \times [0, L_z] \subset \mathbb{R}^2$. The domain is covered by two media, with known distinct slowness $0 < a < b$. we then represent the slowness by a binary function, $s : \Omega \to \{a, b\}$. This representation arises in seismic tomography [5, 58]. The forward problem is the Eikonal equation (2.2) – (2.4) with appropriate boundary conditions. We know this yields a unique solution if jump discontinuities form Lipschitz regular interfaces.

We know location of sources and receivers, and one common configuration is *crosswell* or *cross hole* tomography, where two boreholes are drilled at $x = 0$ (the source borehole), and $x = L_x$ (the receiver borehole). Sources are set of at regular

intervals $[0,(i/M)L_z]$, and recorded at regular intervals $[L_x,(j/N)L_z]$ for $i=1,\dots,M$ and $j=1,\dots,N$. This scenario is depicted in Figure 1. The goal of the inverse problem is: given experimental data of hitting times of seismic waves at receivers, can one reconstruct the slowness $s:\Omega\to\{a,b\}$?

2. The abstract forward and inverse problems

2.1. The forward problem

Let $\Omega\subset\mathbb{R}^d$ ($d=2$ or 3), be an open bounded domain with Lipschitz boundary $\partial\Omega$ and take $x_0\in\Omega$ fixed. We investigate the first arrival time at $x\in\bar{\Omega}$ of a ray originating from a source at x_0. Denote the space of possible ray paths by,

$$\Xi_{x_0}(x) \coloneqq \{\,\xi\in W^{1,\infty}([0,1],\bar{\Omega}) \quad | \quad \xi(0)=x_0,\,\xi(1)=x\,\}.$$

We denote the impedance of the ray in the medium (representing subsurface heterogeneity) by defining a continuous and positive *slowness function* $s:\bar{\Omega}\to\mathbb{R}_+$. The *first-arrival traveltime* $T(x)$ over this set of paths is defined as

$$T(x) \coloneqq \inf_{\xi\in\Xi_{x_0}(x)}\int_0^1 s(\xi(r))\,|\xi'(r)|\,\mathrm{d}r. \tag{2.1}$$

This extremal value is viewed as the shortest arrival time of a ray that obeys Fermat's principle and travels from x_0 to x, with speed $c(x)=s(x)^{-1}$. It has been shown in [42, 52], that $T(x)$ formally satisfies a stationary Hamilton-Jacobi equation, namely the following Eikonal equation:

$$|\nabla T(x)| = s(x), \qquad \forall x\in\Omega\setminus\{x_0\}, \tag{2.2}$$

$$T(x_0)=0, \tag{2.3}$$

$$\nabla T(y)\cdot n(y)\geq 0, \qquad \forall y\in\partial\Omega, \tag{2.4}$$

where n is the outward pointing unit normal. We say T is a viscosity solution of the Eikonal equation (2.2) – (2.4). In the context of (2.1), $\nabla T(x)$ is the direction of the optimal ray and $s(x)^{-1}$ is the speed of the ray at x. The point condition (2.3) ensures the source at x_0 has zero travel time. The Soner boundary condition (2.4) ensures information propagates out of the domain, that is, all ray paths terminate at $\partial\Omega$ [52].

2.2. Well posedness

The well posedness of (2.2) – (2.4) is of importance when considering the inverse problem associated to it. The goal of the theory is to clarify the requirements for this Eikonal equations to produce regular (Lipschitz) solutions. The first proof of well posedness is from [42], where the problem (2.1) was posed with Dirichlet boundary data (replacing (2.4)):

$$s=\varphi \text{ on } \partial\Omega, \text{ where } |\varphi(x)-\varphi(y)|\leq \bar{T}(x,y), \tag{2.5}$$

and $\bar{T}(x,y)$ is the traveltime between a point x and y in $\bar{\Omega}$.

The next development is found in [52], where it is shown that if one constrains paths to lie within $\bar{\Omega}$, a solution of (2.1) is a viscosity solution of (2.2) – (2.4). By viscosity solution here, we mean that $T(x)$ is a subsolution of (2.2) for $x \in \Omega \setminus \{x_0\}$ and a supersolution of (2.2) for $x \in \bar{\Omega} \setminus \{x_0\}$, also used in [21, 22]. In [52], it was shown that if $s \in C^0(\Omega)$ is continuous, positive and bounded, there is a unique optimal value T, that is Lipschitz continuous on $\bar{\Omega}$ with constant bounded by $\|s\|_\infty$. The result holds under regularity conditions satisfied by bounded C^1 domains, or piecewise smooth boundaries containing isolated corners.

The results have been extended in [21] to deal with discontinuous slowness function while still achieving Lipschitz continuous solutions. First note, the regularity requirement on s in [52] can be written: $\forall x, y \in \Omega$,

$$|s(x) - s(y)| \leq w_u(\|x - y\|), \tag{2.6}$$

where w_u is nondecreasing, continuous and $w_u(0) = 0$, differences in s are bounded by a continuous function. In [21], this is extended as follows: $\forall x \in \Omega,\ \exists \varepsilon > 0,\ n \in S^{n-1}$ such that $\forall y \in \Omega,\ r > 0,\ d \in S^{n-1}$, with $|d - n| < \varepsilon$ and $y + rd \in \Omega$, such that

$$s(y + rd) - s(y) \leq w_s(\|x - y\| + \|y + rd - y\|) = w_s(\|x - y\| + r) = w_s^y(r), \tag{2.7}$$

for the $(n-1)$ – dimensional unit sphere S^{n-1}. The property (2.7) requires that at every $x \in \Omega$ one can choose a cone in direction n such that within this cone, the slowness function is bounded by a nondecreasing continuous function $w_u(r)$ (as in (2.6)). This condition holds if s is continuous, and allows for jump discontinuities along an interface where the left and right limits of the jump satisfy strict inequality along the length of the interface. In two dimensions, this can be extended to allow junctions of curves of discontinuity [21].

Remark 1. *The extension is technical. In particular, functions in $W^{1,\alpha}(\Omega)$ or $BV(\Omega, \{a, b\})$ for $0 < a < b$, will not automatically satisfy these cone conditions.*

2.3. Nondifferentiability of the solution operator

We note that the solution of (2.2) is not differentiable in s. Differentiability of the forward problem is often an important feature to solve an inverse problem, as derivative informed methods are often far more efficient for formulating and searching for optimizers. We shall address this through the course of the derivation by making use of a discrete construction found in [22].

2.4. The abstract inverse problem

Let y be observations of the solution of a forward problem, from input slowness s. Assume observations are perturbed by (additive) noise η. The abstract inverse problem is stated as: Given the observed data y of the problem $\mathcal{G}$, find s satisfying

$$y = \mathcal{G}(s) + \eta. \tag{2.8}$$

We choose the forward problem to be (2.2) – (2.4), acting on a slowness function s and followed by an operator $K : T(s) \mapsto \mathcal{G}(s) \in \mathcal{O}$ an observation space. The data $y = T_{obs} \in \mathcal{O}$ is a set of observed first hitting times at fixed known receiver locations in $\bar{\Omega}$ or on densely in a region or curve in $\bar{\Omega}$. Altogether, $\eta = y - \mathcal{G}(s) = T_{obs} - K(T(s))$.

We consider a *mismatch functional* $\mathcal{I}(v)$ to measure the *data misfit* $y - \mathcal{G}(s)$ (chosen in Section 2.5). Our goal is to find a minimizer of $\mathcal{I}$ within a space $\mathcal{A}$, known as the *prior* or *model space*. We write,

$$\text{Find } s \coloneqq \min_{v \in \mathcal{A}} \mathcal{I}(v). \tag{2.9}$$

The choices of $\mathcal{O}$ and $\mathcal{A}$, affect the complexity of the inverse problem. In particular, (2.2) – (2.4) and $\mathcal{I}$, must be well posed for any $s \in \mathcal{A}$, as discussed in Section 2.2.

2.5. Misfit functional

For the misfit we choose a quadratic functional and take the observations $\mathcal{O}$ to be a subspace of a Hilbert space:

$$\mathcal{I}(s) = \|y - \mathcal{G}(s)\|^2_{\mathcal{O}} = \frac{1}{2}\|K(T(s)) - T_{obs}\|^2_{\mathcal{O}}. \tag{2.10}$$

Often the measurements are assumed to be densely defined on union of curves $\Gamma \subset \bar{\Omega}$. In this case, take $T_{obs} : \Gamma \to \mathbb{R}_{>0}$ for $T_{obs} \in \mathcal{O} = L^2(\Gamma)$. The natural candidate for a misfit functional is

$$\mathcal{I}(s) = \frac{1}{2}\int_{\Gamma} |T(s)(x) - T_{obs}(x)|^2 \, \mathrm{d}s_x. \tag{2.11}$$

Commonly $\Gamma \subset \partial\Omega$, as observations are taken at surface seismic stations, or at detectors in a drilled well (see [58, 36, 39] and Figure 1).

We may also consider observations at a finite number of distinct points $x_1, \ldots, x_M \in \bar{\Omega}$ (see [36]). In this case the natural functional (for suitable weights $\{w_i\}$) is,

$$\mathcal{I}(s) = \frac{1}{2}\sum_{i=1}^{M} w_i |T(s)(x_i) - T_{obs}(x_i)|^2. \tag{2.12}$$

In either case, the positions of the observations relative to a source will affect the recovery. We generalize to consider multiple experiments using travel times $T^1(s), \ldots T^S(s)$ produced from sources $x_0^1, \ldots x_0^S$ by summing over the functionals:

$$\mathcal{I}(s) = \frac{1}{2}\sum_{j=1}^{S} \|K(T^j(s)) - T^j_{obs}\|^2_{\mathcal{O}},$$

where T^j_{obs} is a traveltime data set corresponding to the source x_0^j. This assumes independence of experiments $j = 1, \ldots S$.

Hereafter, for the sake of a cleaner analysis in the subsequent sections, we assume a single source and observation space $\mathcal{O} = L^2(\Gamma)$ for a boundary segment $\Gamma \subset \partial\Omega$. The analysis we present can be generalized to multiple source points and applied to a mismatch functional (2.12) by following the same procedures. In fact, with a particular choice of weights, we see that (2.12) can be seen as a discretization of (2.11), see Section 4.2.

2.6. Noisy observations

We often consider the observational data y to be perturbed by a random variable, interpreted as observation imprecision. In the case of M finite observations at discrete points, we introduce this randomness through a normal random variable $\eta \sim \mu = N(0, \Sigma)$ on $\mathbb{R}^M$, where $\Sigma \in \mathbb{R}^{M\times M}$ as a positive-definite covariance.

The misfit (2.12) is naturally weighted with the confidence, for example choose $\Sigma = \nu^2\mathbb{I}$, where $\mathbb{I}$ is the identity on $\mathbb{R}^{M\times M}$ and a constant ν:

$$\|\Sigma^{-\frac{1}{2}}(y - \mathcal{G}(s))\|^2_{\mathcal{O}} = \frac{1}{2}\sum_{i=1}^{M} |\Sigma^{-\frac{1}{2}}(T(s)(x_i) - T_{obs}(x_i))|^2 = \frac{1}{2\nu^2}\sum_{i=1}^{M} |T(s)(x_i) - T_{obs}(x_i)|^2.$$

3. Regularization

3.1. Introduction

We motivate the phase field method by considering a natural regularization for a binary recovery problem. One penalizes length of the interface Γ between constant regions [47], (also [57, 58]), as shown in the following functional for some forward map $\mathcal{F}$,

$$\mathcal{I}^{Per}(s) = \frac{1}{2}\|y - \mathcal{F}(s)\|^2_{\mathcal{O}} + \delta|\Gamma|. \tag{3.1}$$

where $\delta > 0$. If $s \in BV(\Omega, \{a, b\})$, then Γ is the perimeter of $\{s = a\}$ and we can rewrite (3.1) as the total variation of s,

$$\mathcal{I}^{Tik}(s) = \frac{1}{2}\|y - \mathcal{F}(s)\|^2_{\mathcal{O}} + \delta\int_{\Omega} |\nabla s|^{\alpha}\,\mathrm{d}x, \tag{3.2}$$

where $\alpha = 1$. The problem is often relaxed by taking $s \in BV(\Omega)$ and losing the binary nature of the prior. Another common choice is $\alpha = 2$ which acts as a smoothing regularization and so the prior space $\mathcal{A} \subset H^1(\Omega)$. A fundamental problem with either regularization is they lead to prior space $\mathcal{A}$ that is not a subset of the continuous functions, and so we cannot take forward map $\mathcal{F} = \mathcal{G}$ as these spaces are insufficient for the well-posedness of the forward problem (see Remark 1).

3.2. Phase field formulation

Phase field regularization provides regular approximation (via gamma convergence [45, 10]) to (3.2) with $\alpha = 1$ and binary priors. We separate the notation for the slowness function and the *phase field variable* in the regularization. Denote the phase field variable by $u : \Omega \to [-1, 1]$, and we solve (2.2) – (2.4) with the slowness function $s(u(x))$, linear in u, and $s : [-1, 1] \to [s_{\min}, s_{\max}]$

$$s(u) = u(s_{\max} - s_{\min})/2 + (s_{\max} + s_{\min})/2 \tag{3.3}$$

bounded by $0 < s_{\min} < s_{\max}$. As u approaches -1 (or 1), s approaches $s_{\min}$ (or $s_{\max}$). We take u to be our new variable, and we use a functional composed of two parts. Firstly a penalty to second order derivatives $\mathcal{J}^{\gamma}_{1,\varepsilon}(u)$ to ensure $s(u) \in \mathcal{A} \subset H^2(\Omega) \subset C^0(\Omega)$.

Secondly a Ginzburg-Landau type functional $\mathcal{J}^\gamma_{2,\varepsilon}(u)$. Define the phase field functional: for $\sigma, \varepsilon, \gamma > 0$,

$$\mathcal{I}^\gamma_\varepsilon(u) \coloneqq \frac{1}{2}\|y - \mathcal{G}(s(u))\|^2_{\mathcal{O}} + \underbrace{\sigma_1 \int_\Omega \gamma \frac{\varepsilon^3}{2}(\Delta u)^2 \,\mathrm{d}x + \sigma_2 \int_\Omega \frac{\varepsilon}{2}|\nabla u|^2 + \frac{1}{\varepsilon}\Psi(u)\,\mathrm{d}x}_{\coloneqq \sigma_1 \mathcal{J}^\gamma_{1,\varepsilon}(u) + \sigma_2 \mathcal{J}^\gamma_{2,\varepsilon}(u)}, \quad (3.4)$$

where $\Psi(u)$ is given by the double obstacle potential with minima on $\{-1, 1\}$:

$$\Psi(u) \coloneqq \begin{cases} \frac{1}{2}(1-u^2), & \text{if } u \in [-1,1], \\ \infty, & \text{otherwise.} \end{cases} \quad (3.5)$$

The behaviour of a minimizer of $\mathcal{J}^\gamma_{2,\varepsilon}$ (defined in (3.4)) is a function which favours taking the exact values -1 or 1 (and is bounded in $[-1, 1]$) due to the double obstacle potential (3.5), see [12]. The minimizer changes between values in a controlled fashion due to the gradient penalizations, across a thin interface characterized by the width parameter $\varepsilon > 0$. For small ϵ, the function takes these constant values large regions of the domain. A minimizer of $\mathcal{J}^\gamma_{2,\varepsilon}$ undergoes a Laplacian penalization, which ensures the function is continuous. We henceforth assume $\sigma = \sigma_1 = \sigma_2$ and take $\mathcal{J}^\gamma_\varepsilon = \mathcal{J}^\gamma_{1,\varepsilon} + \mathcal{J}^\gamma_{2,\varepsilon}$. The use of similar regularization functionals are found in [24, 16]. Possible boundary conditions that may be imposed are as follows

(i) $u = 1$ or -1 on $\partial\Omega$, $\frac{\partial u}{\partial n} = 0$ on $\partial\Omega$,

(ii) $u = 1$ or -1 or $\partial_D\Omega$, $\frac{\partial u}{\partial n} = 0$ on $\partial_N\Omega$, for $\partial\Omega = \partial_D\Omega \cup \partial_N\Omega$,

(iii) no conditions imposed.

We interpret a Dirichlet condition on the phase field variable, as having knowledge of the value of the slowness there. A zero Neumann condition on the boundary imposes (in the absence of other information) that the interface must touch the boundary orthogonally there. We choose to take (ii) with $\partial_D\Omega \cap \partial_N\Omega = \emptyset$. Our prior space is written:

$$\mathcal{A} \coloneqq \left\{ u \in H^2(\Omega) \;\middle|\; u = 1 \text{ or } -1 \text{ on } \partial_D\Omega,\ \frac{\partial u}{\partial n} = 0 \text{ on } \partial_N\Omega \right\}. \quad (3.6)$$

Remark 2. *Another choice is a finite dimensional prior as in [22]. The problem is reduced to minimizing over a set of coefficients of some given functions: Let $\{\psi_k\}$ satisfy $0 \le \psi_k \in W^{1,\infty}(\Omega)$ $\forall i = 1, \ldots, K$ and $\sum_{i=k}^K \psi_k = 1$ and their support covers $\bar{\Omega}$. Define*

$$\mathcal{A} \coloneqq \{\, s\colon \bar{\Omega} \to \mathbb{R} \mid s(x) = \sum_{k=1}^K s_k \psi_k(x),\ s_k \ge 0 \text{ bounded, and } \phi_k \ge 0 \text{ continuous} \,\}, \quad (3.7)$$

The authors showed that when discretized appropriately, the discrete forward problem (based on (2.2) – (2.4)) can attain a derivative with respect to the state variables s_i. We exploit this setting later to obtain a discrete derivative.

Define $\mathcal{I}^\gamma_\varepsilon$ by (3.4), then our problem is:

$$\text{Find } u \coloneqq \arg\min_{v \in \mathcal{A}} \ \mathcal{I}^\gamma_\varepsilon(v), \quad (3.8)$$

Theorem 3. *Let $\Omega \subset \mathbb{R}^d$ be bounded with $\partial\Omega$ Lipschitz and consider $\mathcal{I}_\varepsilon^\gamma$ defined in (3.4) and $\mathcal{A}$ defined in (3.6). Then there exists a solution to the minimization problem (3.8).*

Proof. We wish to use [19, Theorem 9.3.1] to prove minimizers exist. We require $\mathcal{A}$ to be weakly sequentially closed in $H^2(\Omega)$, and that the functional (3.4):

$$\mathcal{I}_\varepsilon^\gamma(u) = \frac{1}{2}\|y - \mathcal{G}(s(u))\|_{\mathcal{O}}^2 + \sigma\mathcal{J}_\varepsilon^\gamma(u),$$

is coercive and lower semicontinuous over $\mathcal{A}$. Firstly, $\mathcal{A}$ is an unbounded, weakly closed subset of H^2. Due to (3.5), a minimizer will not satisfy $u(\Omega) \not\subset [-1,1]$ so we work over $\tilde{\mathcal{A}} = \mathcal{A} \cap H^2(\Omega\,;[-1,1])$. Over $\tilde{\mathcal{A}}$, we see

$$\sigma\mathcal{J}_\varepsilon^\gamma(u) = \sigma\int_\Omega \Big(\gamma\frac{\varepsilon^3}{2}(\Delta u)^2 + \frac{\varepsilon}{2}|\nabla u|^2 + \frac{1}{2\varepsilon}(1-u^2)\,\mathrm{d}x \geq \sigma\gamma\frac{\varepsilon^3}{2}||\Delta u||_{L^2(\Omega)}^2 + \frac{\varepsilon}{2}||\nabla u||_{L^2(\Omega)}^2$$

and $\|u\|_{L^2(\Omega)}^2 \leq C_u \leq |\Omega|$. From the Poincaré inequality we obtain $\mathcal{J}_\varepsilon^\gamma$ is coercive on $\tilde{\mathcal{A}}$. Weak convergence in $H^2(\Omega)$ implies strong convergence of u, ∇u in $L^2(\Omega)$ and weak convergence of Δu in $L^2(\Omega)$. Quadratic functionals are weak lower semicontinuous, and so $\mathcal{J}_\varepsilon^\gamma$ is too.

The data misfit is nonnegative and so immediately $\mathcal{I}_\varepsilon^\gamma$ is coercive. For lower semicontinuity of the misfit we consider a weakly convergent sequence (u_k) in $H^2(\Omega)$ to $u \in \mathcal{A}$. As $u_k \mapsto \mathcal{G}(s(u_k))$ is well defined and Lipschitz continuous, then the trace onto $\partial\Omega$ is (at least) in $\mathcal{O} = L^2(\partial\Omega)$. Furthermore, u_k converges strongly in $H^1(\Omega)$, and so by the trace theorem for Lipschitz domains [4, Theorem A8.6], u_k converges strongly on $\partial\Omega$ in $L^2(\Omega)$. The continuity of $\mathcal{G}(s(u_k))$ preserves the limit, thus $\mathcal{G}(s(u_k)$ converges to $\mathcal{G}(s(u))$ in $\mathcal{O}$ and the misfit (and so $\mathcal{I}_\varepsilon^\gamma$) is weakly sequentially lower semicontinuous. We then apply [19, Theorem 9.3.1] to complete the proof. □

3.3. Gamma convergence

The strength of the phase field technique lies with the convergence of the phase field functional to the perimeter functional in the sense of Γ – convergence. as the interfacial parameter $\varepsilon \to 0$. The first result [45] was for the Ginzburg-Landau functional (with H^1 minimizers) with a quartic double well potential $\Psi(u) = \frac{1}{2}(1-u^2)^2$. More recently this has been extended to a functional with H^2 minimizers [29] (similar results in [18, 26]). The analysis of [45] was extended to the double obstacle potential (3.5) in [10] by investigation of profile which appear across interfaces of minimizers. With this technique, we extend the analysis of [29] to $\mathcal{J}_\varepsilon^\gamma$ in (3.4) with a double obstacle (3.5). We denote $\int_\Omega |\nabla u|\,\mathrm{d}x$ for $u \in BV(\Omega)$ to be the total variation of u.

Theorem 4. *Let $\Omega \subset \mathbb{R}^d$ with Lipschitz boundary. Define the following functionals:*

$$\mathcal{J}_\varepsilon^\gamma(u) \coloneqq \int_\Omega \gamma\frac{\varepsilon^3}{2}(\Delta u)^2 + \frac{\varepsilon}{2}|\nabla u|^2 + \frac{1}{\varepsilon}\Psi(u)\,\mathrm{d}x, \tag{3.9}$$

and

$$j^\gamma(z) \coloneqq \int_{\mathbb{R}} \frac{1}{2}\gamma(z'')^2 + \frac{1}{2}(z')^2 + \Psi(z)\,\mathrm{d}x, \tag{3.10}$$

where Ψ is the double obstacle potential (3.5). The Γ – limit: $\mathcal{J}_0^\gamma := \Gamma - \lim_{\varepsilon\to 0} \mathcal{J}_\varepsilon^\gamma$, is proportional to the perimeter functional:

$$\mathcal{J}_0^\gamma(u) = \begin{cases} \frac{1}{2}P^\gamma \int_\Omega |\nabla u| \,\mathrm{d}x & \text{if } u \in BV(\Omega, \{-1,1\}), \\ \infty & \text{if } u \in L^1(\Omega) \setminus BV(\Omega, \{-1,1\}). \end{cases}$$

for

$$BV(\Omega, \{a,b\}) := \{\ w \in BV(\Omega) : w(\Omega) \subset \{a,b\}\ \}.$$

The constant $P^\gamma := \inf_{v\in\mathcal{V}} j^\gamma(v)$ is the double obstacle transition energy, where

$$\mathcal{V} = \Big\{ v \in C^2(\mathbb{R}; [-1,1]) \mid \exists \delta > 0, \forall x \in \mathbb{R}\ v(x) = -v(-x), \\ v'(x) \geq 0,\ \textit{and } v(x > \delta) = 1,\ v(x < -\delta) = -1 \Big\}.$$

Proof. We consign the proof of this theorem to Appendix A □

Remark 5. *The relationship between (3.9) and (3.10) is through an ansatz on the phase field variable u:*

$$u(x) = z^\gamma\Big(\frac{d(x)}{\varepsilon}\Big), \qquad z(0) = 0,$$

where $d(x)$ is the signed distance function to the limiting interface at $\{u(x) = 0\}$, and $z^\gamma : \mathbb{R} \to [-1,1]$ is a smooth function found in Appendix A. Inserting the ansatz into (3.9) results in (3.10).

Remark 6. *For $\mathcal{I}_\varepsilon^\gamma$ in (3.4), an equivalent result to Theorem 4 is unknown, due to technicalities in assuring that misfit functional is well defined in the $\varepsilon \to 0$ limit.*

3.4. Mixed formulation

In practice, approximating the high order derivatives as found in the problem (3.8) in the discrete setting is complex when using conforming finite elements. We therefore create a mixed formulation. We begin by introducing a new variable w such that $w = -\Delta u$ weakly on Ω, and setting:

$$\mathcal{I}_\varepsilon^\gamma(u,w) = \frac{1}{2}\|y - \mathcal{G}(s(u))\|_{\mathcal{O}}^2 + \underbrace{\sigma \int_\Omega \gamma\frac{\varepsilon^3}{2}w^2 + \frac{\varepsilon}{2}|\nabla u|^2 + \frac{1}{\varepsilon}\Psi(u)\,\mathrm{d}x}_{\sigma\mathcal{J}_\varepsilon^\gamma(u,w)}. \quad (3.11)$$

The natural optimization problem is then

$$\text{Find } (u,w) := \underset{(\tilde{u},\tilde{w})\in\mathcal{A}_\Delta}{\arg\min}\ \mathcal{I}_\varepsilon^\gamma(\tilde{u},\tilde{w}). \quad (3.12)$$

We find a suitable set $\mathcal{A}_\Delta$ with appropriate boundary conditions and treat the cases $\partial_D\Omega \neq \emptyset$ and $\partial_D\Omega = \emptyset$ separately. We define a bilinear form $B : H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$,

$$B(z,v) = \int_\Omega \nabla z \cdot \nabla v \,\mathrm{d}x,$$

and denote the L^2 inner product by $\langle \cdot, \cdot \rangle$. Define the spaces

$$H^1_D = \{u \in H^1(\Omega) \mid u = g \text{ on } \partial_D\Omega\}, \tag{3.13}$$
$$H^1_{0,D} = \{u \in H^1(\Omega) \mid u = 0 \text{ on } \partial_D\Omega\}. \tag{3.14}$$

then we may define

$$\mathcal{A}_\Delta \coloneqq \begin{cases} \mathcal{A}_{\Delta,D}, & \text{if } \partial_D\Omega \neq \emptyset, \\ \mathcal{A}_{\Delta,N}, & \text{if } \partial_D\Omega = \emptyset, \end{cases} \tag{3.15}$$

where,

$$\mathcal{A}_{\Delta,D} \coloneqq \Big\{(u,w) \in H^1_D(\Omega) \times L^2(\Omega) \;\Big|\; B(u,\zeta) = \langle w, \zeta\rangle,\ \forall \zeta \in H^1_{0,D}(\Omega)\Big\}, \tag{3.16}$$

$$\mathcal{A}_{\Delta,N} \coloneqq \Big\{(u,w) \in H^1(\Omega) \times L^2(\Omega) \;\Big|\; \int_\Omega w \,\mathrm{d}x = 0, \quad B(u,\zeta) = \langle w, \zeta\rangle,\ \forall \zeta \in H^1(\Omega)\Big\}. \tag{3.17}$$

Remark 7. *The compatibility condition $\int_\Omega w \,\mathrm{d}x = 0$ in (3.17) is necessary to make sense of the relation between u and w in the case of the pure Neumann boundary condition, $\partial u/\partial n = 0$ on $\partial\Omega$.*

We now show existence of solutions for the mixed formulation.

Theorem 8. *Let $\Omega \subset \mathbb{R}^d$ be bounded with $\partial\Omega$ either being C^2 or $\partial\Omega$ Lipschitz and Ω convex. Consider $\mathcal{I}^\gamma_\varepsilon$ defined in (3.11) and $\mathcal{A}_\Delta$ as in (3.15). Then there exists a solution to the minimization problem (3.12)*

Proof. The assumptions on Ω imply that if $(u,w) \in \mathcal{A}_\Delta$ then standard results for $-\Delta u = w, w \in L^2(\Omega), u \in H^1(\Omega)$ in elliptic regularity, (see [28, Chapter 2] for convex Lipschitz domain), imply that $u \in H^2(\Omega)$ and is continuous so that the forward problem has a solution and the misfit functional is well defined. Consider $\partial_D\Omega \neq \emptyset$. The set $\mathcal{A}_{\Delta,D}$ is an unbounded, weakly closed subset of $H^1(\Omega) \times L^2(\Omega)$. We seek to apply the theorem of [19, Theorem 9.3.1], and so must show the functional (3.11) is coercive and lower semicontinuous over $\mathcal{A}_{\Delta,D}$. For coercivity, by definition of (3.5), we are done if $u(\Omega) \not\subset [-1,1]$, and so work over $\tilde{\mathcal{A}}_{\Delta,D} = \mathcal{A}_{\Delta,D} \cap (H^1(\Omega,[-1,1]) \times L^2(\Omega))$ with norm $\|(u,w)\|_\Delta = (\|u\|^2_{H^1(\Omega)} + \|w\|^2_{L^2(\Omega)})^{\frac{1}{2}}$.

We see that over $\tilde{\mathcal{A}}_{\Delta,D}$ using (3.5), we have

$$\begin{aligned} \sigma\mathcal{J}^\gamma_\varepsilon(u,w) &= \sigma\int_\Omega \gamma\frac{\varepsilon^3}{2}w^2 + \frac{\varepsilon}{2}|\nabla u|^2 + \frac{1}{2\varepsilon}(1-u^2)\,\mathrm{d}x \\ &\geq \sigma\int_\Omega \gamma\frac{\varepsilon^3}{2}w^2 + \frac{\varepsilon}{2}\Big(|\nabla u|^2 + u^2\Big)\,\mathrm{d}x - (\frac{1}{2\varepsilon} + \frac{\varepsilon}{2})\int_\Omega u^2\,\mathrm{d}x \\ &\geq \min(\sigma\frac{\varepsilon^3}{2}, \frac{\varepsilon}{2})\|(u,w)\|^2_\Delta - (\frac{1}{2\varepsilon} + \frac{\varepsilon}{2})C_u, \end{aligned}$$

where $\|u\|^2_{L^2(\Omega)} \leq C_u \leq |\Omega|$. Therefore $\mathcal{J}^\gamma_\varepsilon$ is coercive on $\tilde{\mathcal{A}}_{\Delta,D}$. Moreover, it is sequentially lower semicontinuous on $\mathcal{A}_{\Delta,D}$ as weak convergence in $H^1(\Omega)\times L^2(\Omega)$ implies

∇u and w weakly, and u strongly converge, and quadratic functionals are weakly lower semicontinuous.

The data misfit is nonnegative and so immediately $\mathcal{I}_\varepsilon^\gamma$ is coercive. For lower semicontinuity of the misfit we consider a weakly convergent sequence (u_k, w_k) in $\mathcal{A}_{\Delta,D}$. Thus $u_k \mapsto \mathcal{G}(s(u_k))$ is well defined and Lipschitz continuous, then the trace onto $\partial\Omega$ is (at least) in $\mathcal{O} = L^2(\partial\Omega)$.

The relationship of (3.16) holds for any $\zeta \in H^1_{0,D}(\Omega)$. One can choose $g \in H^1(\Omega)$ s.t $u_k - g = v_k \in H^1_{0,D}(\Omega)$ and choose $\zeta = v_k$. Then:

$$\int_\Omega |\nabla u_k|^2 \,\mathrm{d}x = \int_\Omega \nabla u_k \cdot \nabla v_k + \nabla u_k \cdot \nabla g \,\mathrm{d}x = \int_\Omega w_k v_k + \nabla u_k \cdot \nabla g \,\mathrm{d}x$$
$$\to \int_\Omega wv + \nabla u \cdot \nabla g \,\mathrm{d}x = \int_\Omega |\nabla u|^2 \,\mathrm{d}x, \qquad \text{as } k \to \infty$$

and so u_k converges strongly in H^1. By the trace theorem for Lipschitz domains [4, Theorem A8.6] we have that u_k converges strongly on $\partial\Omega$ in L^2. The continuity of $\mathcal{G}(s(u_k))$ preserves the limit, thus $\|\mathcal{G}(s(u_k)\|_\mathcal{O} \to \|\mathcal{G}(s(u))\|_\mathcal{O}$ and the misfit (and so $\mathcal{I}_\varepsilon^\gamma$) is weakly sequentially lower semicontinuous.

The proof for the case $\partial_D\Omega = \emptyset$ is similar. □

We now assert that the mixed formulation yields solutions for the original problem:

Theorem 9. *Let $\Omega \in \mathbb{R}^d$, bounded with $\partial\Omega$ either being C^2 or with $\partial\Omega$ Lipschitz and Ω convex. Let $(u, w) \in \mathcal{A}_\Delta$ defined by (3.15), be solutions of the minimization problem (3.12) with $\mathcal{I}_\varepsilon^\gamma(u, w)$ defined by (3.11) . Then, u is also a solution of the minimization problem (2.1) with $\mathcal{I}_\varepsilon^\gamma(u)$ defined by (3.4) and $\mathcal{A}$ defined by (3.6).*

Proof. Let $(u, w) \in \mathcal{A}_\Delta$. Standard regularity theorems for elliptic problems, gives that u is not only in $H^1_D(\Omega)$ (resp. $H^1(\Omega)$), but is actually $u \in H^2_{loc}(\Omega)$. Furthermore, the boundary regularity is sufficient to extend this result to the boundary (see [28, Chapter 2] for convex Lipschitz domain). In this case $w = -\Delta u$ is well defined in $L^2(\Omega)$, so we show the boundary conditions are preserved. If $\partial_D\Omega \neq \emptyset$, we obtain through the relation that for any $\xi \in H^1_{0,D}(\Omega)$:

$$\int_\Omega w\xi \,\mathrm{d}x = \int_\Omega \nabla u \cdot \nabla \xi \,\mathrm{d}x = -\int_\Omega \Delta u \xi \,\mathrm{d}x + \int_{\partial\Omega_N} \xi \nabla u \cdot n \,\mathrm{d}y,$$

and so as $w = -\Delta u$ in Ω, ξ is arbitrary on $\partial\Omega_N$ we obtain that $\partial u/\partial n = 0$ on $\partial\Omega_N$.If $\partial_D\Omega = \emptyset$, and $\int_\Omega w \,\mathrm{d}x = 0$, we know from the elliptic relation that $\forall \xi \in H^1(\Omega)$,

$$\int_\Omega w\xi \,\mathrm{d}x = \int_\Omega \nabla u \cdot \nabla \xi \,\mathrm{d}x = -\int_\Omega \Delta u \xi \,\mathrm{d}x + \int_\Omega \xi \nabla u \cdot n \,\mathrm{d}y,$$

and setting $\xi = 1$, and $w = -\Delta u$ in Ω we obtain $\partial u/\partial n = 0$ on $\partial\Omega$, So $u \in \mathcal{A}$. We have shown that for both sets of boundary conditions, we have $u \in \mathcal{A}$ and so,

$$\inf_{(\bar{u},\bar{w})\in\mathcal{A}_\Delta} \mathcal{I}_\varepsilon^\gamma(\bar{u}, \bar{w}) \leq \inf_{\bar{u}\in\mathcal{A}} \mathcal{I}_\varepsilon^\gamma(\bar{u}, -\Delta\bar{u}) = \inf_{\bar{u}\in\mathcal{A}} \mathcal{I}_\varepsilon^\gamma(\bar{u}).$$

Thus a solution of the mixed problem is also a solution of the original problem. □

Remark 10. *In the continuous setting we are unable to calculate derivatives (which would lead to critical points, optimality conditions of $\mathcal{I}_\varepsilon^\gamma$ etc.), due to the lack of differentiability of the misfit functional with respect to $s(u)$ (briefly mentioned in Remark 2). Hence discussion of derivatives is delayed until Section 4.4.*

4. Discretization

4.1. Discrete forward problem

For the discretization of (2.2) – (2.4) we follow [22]. Here we recall the framework and some key results sufficient for our purpose. Assume that Ω is a convex polygonal domain. We take a regular quadrilateral grid of uniform grid size h, and so $\Omega_h = \Omega \cap \mathbb{Z}_h^2$ where $\mathbb{Z}_h^2 = \{(h\alpha_1, h\alpha_2) \mid (\alpha_1, \alpha_2) \in \mathbb{Z}^2\}$. We assume the point constraint (2.3) in the forward problem lies on a grid point, $x_0 \in \Omega_h$. We index with $\alpha \in \mathbb{Z}^2$, therefore we may say $x_0 = x_{\alpha_0}$ for some $\alpha_0 \in \mathbb{Z}^2$.

Define the discrete boundary $\partial\Omega_h \coloneqq \{y \notin \Omega_h \mid y = x + (-1)^i e_j,\ x \in \Omega_h,\ i, j \in \{1, 2\}\}$ for standard basis vectors e_j and set $\bar{\Omega}_h = \Omega_h \cup \partial\Omega_h$. Finally, we take the set of neighbours $\mathcal{N}_\alpha$ about a point x_α to be the set of neighbours in $\bar{\Omega}_h$. In this way the (2.2) – (2.4) can be discretized using the monotone finite difference scheme,

$$\sum_{x_\beta \in \mathcal{N}_\alpha} \Big[\Big(\frac{\mathcal{T}_\alpha - \mathcal{T}_\beta}{|x_\alpha - x_\beta|}\Big)^+\Big]^2 = s(u_h(x_\alpha))^2, \qquad \text{if } x_\alpha \in \bar{\Omega}_h \setminus \{x_{\alpha_0}\}, \tag{4.1}$$

$$\mathcal{T}_{\alpha_0} = 0, \tag{4.2}$$

where $\mathcal{T}_\delta = \mathcal{T}(x_\delta)$, $u_h : \Omega_h \to \mathbb{R}$ continuous interpolation of u onto Ω_h and $y^+ = \max(y, 0)$, and

$$s : [-1, 1] \to [s_{\min}, s_{\max}], \qquad s(u_h(x_\alpha)) = \frac{s_{\max} - s_{\min}}{2} u_h(x_\alpha) + \frac{s_{\max} + s_{\min}}{2}.$$

It is known [22], for $0 < u \in C^0(\bar{\Omega})$, equations (4.1) – (4.2) gives a unique nonnegative (uniform) Lipschitz continuous solution $\mathcal{T} : \bar{\Omega}_h \to \mathbb{R}_{\geq 0}$ which converges as the mesh size decreases,

$$\max_{x_\alpha \in \Omega_h} |\mathcal{T}_\alpha - T(x_\alpha)| \leq C\sqrt{h}. \tag{4.3}$$

This equation is solved using the fast marching method [49, 50, 48, 22], an efficient and robust technique terminating in a finite number of steps. We label the discrete FMM solution map taking u_h and obtaining the discrete solution $\mathcal{T}$, as $\mathcal{G}_h(s(u_h))$.

4.2. The misfit functional and prior space

We now consider a discrete analogue of problem (3.8). We discretize the misfit functional boundary integral (2.11) to give $\mathcal{I}_h$. More precisely, we take $\mathcal{O}_h = L_h^2(\Gamma_h)$, $\Gamma_h \subset \partial\Omega_h$, a finite difference approximation of $L^2(\Gamma_h)$. Let $\mathcal{T} : \bar{\Omega}_h \to \mathbb{R}_{\geq 0}$, be the solution of (4.1) – (4.2) with slowness function $s(u_h)$. Then,

$$\mathcal{I}_h(u_h) \coloneqq \frac{1}{2} \sum_{i=1}^{M} h|\mathcal{T}(x_{\alpha_i}) - T_{obs}(x_{\alpha_i})|^2, \tag{4.4}$$

where $\{x_{\alpha_i}\}_{i=1}^M = \Gamma_h$. The observations are assumed to be at grid points, and henceforth denoted by $y_h = \{T_{obs}(x_{\alpha_i})\}$. We note that this functional is of the form (2.12), with weights w_i given by the discrete grid spacing.

4.3. Discrete inverse problem

We choose to work with a finite element formulation for approximating u and the regularization functional, though we are not limited to this (one could choose a finite difference method for example). To keep the distinction between the discretization of forward and inverse problems, we use $\hbar$ as the numerical approximation parameter, such that the scheme converges as $\hbar \to 0$. For $\hbar \geq h > 0$, we decompose the polygonal domain Ω into a union of triangles, $\mathscr{T}_\hbar$, whose diameters are bounded below by $\hbar$. Working directly by discretizing (3.8), would require an H^2 conforming finite element space to approximate the function u, (such as Lagrangian $\mathbb{P}^3$ for C^0 elements, or Bell triangular for C^1 elements), unfortunately these are computationally costly, see [14]. Therefore, we present a formulation based on H^1 conforming elements and work with a discretization based on the mixed formulation (3.12). We choose to use P^1 Lagrangian finite elements, given by piecewise linear nodal basis functions $\{\phi_l\}$ for every $x_l \in \Omega$ and satisfying $\phi_l(x_k) = \delta_{kl}$. Define this finite element space,

$$S_\hbar(\Omega) \coloneqq \{v \in C^0(\Omega) \mid v_h|_{\mathcal{T}} \in \mathbb{P}^1, \mathcal{T} \in \mathcal{T}_\hbar\},$$

where the space of polynomials can be written as

$$\mathbb{P}^1 = \{v \mid v(x) = \sum_{l=1}^{L} v_l \phi_l(x) \text{ where } v_l = v(x_l), \quad \forall x_l \in \Omega\}.$$

Remark 11. *We use a finer numerical mesh parameter for the forward problem than the inverse problem i.e $\hbar \leq h$. This is because our problem is to approximate u and the solution of the forward problem is required for resolution. In proofs we often identify the nodes of both finite difference and finite element grids, so that we do not have to include the forward and backward interpolation operators.*

We define the discrete functional,

$$\mathcal{I}^\gamma_{\varepsilon,h,\hbar}(u_\hbar, w_\hbar) \coloneqq \frac{1}{2}\|y_h - \mathcal{G}_h(s(u_\hbar))\|^2_{\mathcal{O}_h} + \sigma \int_\Omega \gamma \frac{\varepsilon^3}{2} w_\hbar^2 + \frac{\varepsilon}{2}|\nabla u_\hbar|^2 + \frac{1}{\varepsilon}\Psi(u_\hbar)\,\mathrm{d}x, \quad (4.5)$$

and the corresponding optimization problem

$$\text{Find } (u_\hbar, w_\hbar) \coloneqq \underset{(\tilde{u}_\hbar, \tilde{w}_\hbar) \in \mathcal{A}_{\Delta,\hbar}}{\arg\min} \quad \mathcal{I}^\gamma_{\varepsilon,h,\hbar}(\tilde{u}_\hbar, \tilde{w}_\hbar). \quad (4.6)$$

We treat $\partial_D\Omega = \emptyset$ or $\partial_D\Omega \neq \emptyset$ separately. Define a bilinear form on $S_\hbar(\Omega) \times S_\hbar(\Omega)$ and discrete L^2 inner product $\langle \cdot, \cdot \rangle_\hbar$:

$$B_\hbar(z_\hbar, v_\hbar) \coloneqq \int_\Omega \nabla z_\hbar \cdot \nabla v_\hbar \,\mathrm{d}x, \qquad \langle z_\hbar, v_\hbar \rangle_\hbar = \int_\Omega z_\hbar v_\hbar \,\mathrm{d}x$$

and denote the space

$$S_{D\hbar} \coloneqq \{u_\hbar \in S_\hbar(\Omega) \mid u_\hbar = 0 \text{ on } \partial_D\Omega\}. \quad (4.7)$$

Then we may define

$$\mathcal{A}_{\Delta,\hbar} \coloneqq \begin{cases} \mathcal{A}_{\Delta,D\hbar}, & \text{if } \partial_D\Omega \neq \emptyset, \\ \mathcal{A}_{\Delta,N\hbar}, & \text{if } \partial_D\Omega = \emptyset, \end{cases} \tag{4.8}$$

given by

$$\mathcal{A}_{\Delta,D\hbar} \coloneqq \Big\{(u_\hbar, w_\hbar) \in S_{D\hbar}(\Omega) \times S_\hbar(\Omega) \;\Big|\; B_\hbar(u_\hbar, \zeta_\hbar) = \langle w_\hbar, \zeta_\hbar\rangle_\hbar,\ \forall \zeta_\hbar \in S_{D\hbar}(\Omega)\Big\}, \tag{4.9}$$

$$\mathcal{A}_{\Delta,N\hbar} \coloneqq \Big\{(u_\hbar, w_\hbar) \in S_\hbar(\Omega) \times S_\hbar(\Omega) \;\Big|\; \int_\Omega w_\hbar \,\mathrm{d}x = 0, \\ B_\hbar(u_\hbar, \zeta_\hbar) = \langle w_\hbar, \zeta_\hbar\rangle_\hbar,\ \forall \zeta_\hbar \in S_\hbar(\Omega)\Big\}. \tag{4.10}$$

Theorem 12. *Let $\Omega \subset \mathbb{R}^d$ be a convex polygonal domain. Define $\mathcal{A}_{\Delta,\hbar}$ as in (4.8), and for any $\gamma, \varepsilon > 0$ define $\mathcal{I}^\gamma_{\varepsilon,h,\hbar}$ as in (4.5), for a convergent discretization $\mathcal{G}_h$ of the forward problem $\mathcal{G}$ (2.2) – (2.4) and data $y \in \mathcal{O}$. Then there exists $(u_\hbar, w_\hbar) \in \mathcal{A}_{\Delta,\hbar}$ such that $\mathcal{I}^\gamma_{\varepsilon,h,\hbar}(u_\hbar, w_\hbar) = \min_{(\tilde{u}_\hbar,\tilde{w}_\hbar)\in\mathcal{A}_{\Delta,\hbar}} \mathcal{I}^\gamma_{\varepsilon,h,\hbar}(\tilde{u}_\hbar, \tilde{w}_\hbar)$. Moreover, every sequence $(u_{\hbar_k}, w_{\hbar_k})_k$ with $\hbar_k \searrow 0$ has a subsequence such that $u_{\hbar_k}$ converges strongly in $H^1(\Omega)$ and $w_{\hbar_k}$ converges weakly in $L^2(\Omega)$ to a minimizer of $I^\gamma_\varepsilon(\cdot,\cdot)$, given by (3.11).*

Proof. As we work over a finite dimensional space, from the proof of Theorem 8 it is straightforward to deduce the coercivity and lower semi continuity of $\mathcal{I}^\gamma_{\varepsilon,h,\hbar}$. Thus there exists a minimum of $\mathcal{I}^\gamma_{\varepsilon,h,\hbar}$.

The cases for $\partial_D\Omega \neq \emptyset$ and $\partial_D\Omega = \emptyset$ follow identically, and so we make no reference to the boundary conditions for u. We also assume for brevity that the nodes for the finite difference and finite element grids are identified (one may rewrite the proof with interpolation operators between them).

Consider a sequence of minimizers $(u_{\hbar_k}, w_{\hbar_k}) \in \mathcal{A}_{\Delta,\hbar_k}$ of $\mathcal{I}^\gamma_{\varepsilon,h_k,\hbar_k}$, for $h_k = \hbar_k$ with $\hbar_k \to 0$ as $k \to \infty$. Then by construction, $(u_{\hbar_k}, w_{\hbar_k})$ is bounded in $H^1(\Omega) \times L^2(\Omega)$. There exists a subsequence relabeled as $(u_{\hbar_k}, w_{\hbar_k})$ and some $(u^*, w^*) \in \mathcal{A}_\Delta$ (as defined in (3.16)) such that, as $k \to \infty$:

$$\begin{aligned} u_{\hbar_k} &\to u^* \text{ weakly in } H^1(\Omega), \\ u_{\hbar_k} &\to u^* \text{ strongly in } L^2(\Omega) \\ w_{\hbar_k} &\to w^* \text{ weakly in } L^2(\Omega) \end{aligned}$$

For convenience we now drop the subscript k in '$\hbar_k$' but note that in the following we mean $(u_\hbar, w_\hbar)$ to denote a subsequence. The elliptic relation of $B_\hbar(u_\hbar, \zeta_\hbar) = \langle w_\hbar, \zeta_\hbar\rangle_\hbar$, holds for any $\zeta_\hbar \in S_\hbar(\Omega)$ (resp $S_{D\hbar}$) so choose $\zeta_\hbar = u_\hbar$. Then as $\hbar \to 0$:

$$\int_\Omega |\nabla u_\hbar|^2 = \int_\Omega w_\hbar u_\hbar \,\mathrm{d}x \to \int_\Omega w^* u^* \,\mathrm{d}x = \int_\Omega |\nabla u^*|^2 \,\mathrm{d}x, \tag{4.11}$$

and therefore $u_\hbar \to u^*$ strongly in $H^1(\Omega)$. Furthermore elliptic regularity implies $u^* \in H^2(\Omega)$. Denote by $\mathcal{K}$ and $\mathcal{K}_\hbar$ the solution operators for the elliptic relation $-\Delta\mathcal{K}\eta = \eta$ and its finite element approximation with either the Dirichlet or Neuman conditions with appropriate data so that

$$u^* = \mathcal{K}w^* \text{ and } u_\hbar = \mathcal{K}_\hbar w_\hbar.$$

By elliptic regularity and standard finite element theory we have

$$||\mathcal{K}\eta||_{H^2(\Omega)} \leq C||\eta||_{L^2(\Omega)} \text{ and } ||\mathcal{K}\eta - \mathcal{K}_\hbar\eta||_{L^\infty(\Omega)} \leq C\hbar^{2-\frac{d}{2}}||\eta||_{L^2(\Omega)}.$$

Decomposing

$$u_\hbar - u^* = (\mathcal{K}_\hbar w_\hbar - \mathcal{K}w_\hbar) + (\mathcal{K}w_\hbar - \mathcal{K}w^*),$$

we see that the first term on the right converges to zero uniformly because of the uniform $L^2(\Omega)$ bound on $w_\hbar$ and the finite element $L^\infty(\Omega)$ error bound. Turning to the second term on the right, we observe that $\mathcal{K}w_h$ converges weakly in $H^2(\Omega)$ because of elliptic regularity and the weak limit is $\mathcal{K}w^*$. Thus by compact embedding we see that $\mathcal{K}w_h - \mathcal{K}w^*$ converges to zero uniformly. Thus $u_\hbar$ converges uniformly to u^*. Stability results for approximations of viscosity solutions imply that $\mathcal{G}_h(s(u_h))$ converges uniformly to $\mathcal{G}(s(u^*))$. It follows that the misfit functional converges

$$\mathcal{I}_h(u_h) \to \mathcal{I}(u^*).$$

We may now prove the claim that that (u^*, w^*) is a minimum of $\mathcal{I}^\gamma_\varepsilon$, we see this by taking $(v, z) \in \mathcal{A}_\Delta$ and a sequence $(v_k, z_k) \to (v, z)$ strongly in $H^1(\Omega) \times L^2(\Omega)$. This may be chosen using suitable interpolations. By definition $\mathcal{I}^\gamma_{\varepsilon,h_k,\hbar_k}(u_{\hbar_k}, w_{\hbar_k}) \leq \mathcal{I}^\gamma_{\varepsilon,h_k,\hbar_k}(v_{\hbar_k}, z_{\hbar_k})$ for all k. Convergence of the misfit functional and the weak lower semi continuity of the functional $\mathcal{I}^\gamma_\varepsilon$ (established in Theorem 8) implies

$$\begin{aligned}\mathcal{I}^\gamma_\varepsilon(u^*, w^*) \leq &\liminf_{k\to\infty} \ \mathcal{I}^\gamma_{\varepsilon,h_k,\hbar_k}(u_{\hbar_k}, w_{\hbar_k}) \leq \limsup_{k\to\infty} \ \mathcal{I}^\gamma_{\varepsilon,h_k,\hbar_k}(u_{\hbar_k}, w_{\hbar_k}) \\ &\leq \lim_{k\to\infty} \mathcal{I}^\gamma_{\varepsilon,h_k,\hbar_k}(v_k, z_k) = \mathcal{I}^\gamma_\varepsilon(v, z). \qquad (4.12)\end{aligned}$$

Thus $\mathcal{I}^\gamma_\varepsilon(u^*, w^*) = \min_{(v,z)\in\mathcal{A}_\Delta} \mathcal{I}^\gamma_\varepsilon(v, z)$.

□

4.4. The discrete derivative

The forward problem (2.2) – (2.4) is not differentiable with respect to the state variable. However using an argument in [22] we recover differentiability for the discrete misfit functional associated to the discretization (2.2) – (2.4). We identify finite difference and finite element nodes (as one can rewrite with interpolation operators) then for $(u_\hbar, w_\hbar) \in \mathcal{A}_{\Delta,\hbar}$, we obtain $u_\hbar = \sum_{i=1}^L u_i\phi_i \in \mathcal{A}_{\Delta,\hbar}$ for $\{\phi_i\}$ basis functions of $S_\hbar$ and get the form discussed in [22] .

We define an adjoint problem associated with the discrete solution $\mathcal{T}$ of (4.1) – (4.2), with discrete slowness function $s(u_\hbar)$. Find $P\colon \bar{\Omega}_h \setminus \{x_{\alpha_0}\}$ so that

$$\sum_{x_\beta\in\mathcal{N}_\alpha} \Big(\frac{\mathcal{T}_\alpha - \mathcal{T}_\beta}{h_{\alpha,\beta}}\Big)^+ \frac{P_\alpha}{h_{\alpha,\beta}} - \Big(\frac{\mathcal{T}_\beta - \mathcal{T}_\alpha}{h_{\alpha,\beta}}\Big)^+ \frac{P_\beta}{h_{\alpha,\beta}} = 0, \qquad x_\alpha \in \Omega_h \setminus \{x_{\alpha_0}\}, \quad (4.13)$$

$$\sum_{x_\beta\in\mathcal{N}_\alpha} \Big(\frac{\mathcal{T}_\alpha - \mathcal{T}_\beta}{h_{\alpha,\beta}}\Big)^+ \frac{P_\alpha}{h_{\alpha,\beta}} - \Big(\frac{\mathcal{T}_\beta - \mathcal{T}_\alpha}{h_{\alpha,\beta}}\Big)^+ \frac{P_\beta}{h_{\alpha,\beta}} = \frac{h_\alpha}{h^2}(T_{obs} - \mathcal{T}_\alpha), \quad x_\alpha \in \Gamma_h, \qquad (4.14)$$

where $h_{\alpha,\beta} = |x_\alpha - x_\beta|$ and the discrete misfit functional is given by (4.4).

Proposition 1. *Let $u_\hbar = \sum_{i=1}^{L} u_i\phi_i \in \mathcal{A}_\hbar$, and $m \in 1,\dots,L$. Let $\mathcal{I}_h$ be defined by (4.4), then,*

$$\frac{\partial \mathcal{I}_h}{\partial u_m}(u_\hbar) = -h^2\frac{s_{\max}-s_{\min}}{2}\sum_{x_\alpha\in\bar{\Omega}_h\setminus\{x_{\alpha_0}\}} P_\alpha s(u_\hbar)(x_\alpha)\phi_m(x_\alpha), \tag{4.15}$$

where P is the solution of (4.13) – (4.14).

Proof. From [22, Theorem 3.6] we trivially extend to linear functions: let $s_h = c_1\sum_{k=1}^{K} s_k\psi_k + c_2$, for $\{\psi_k\}$ basis functions satisfying the properties of Remark 2 and c_1, c_2 constants. If we solve (4.1) – (4.2) with right hand side s_h^2, define

$$\tilde{\mathcal{I}}_h(s_h) \coloneqq \frac{1}{2}\sum_{i=1}^{N} h_{\alpha_i}|\mathcal{T}(x_{\alpha_i}) - T_{obs}(x_{\alpha_i})|^2,$$

and

$$\frac{\partial\tilde{\mathcal{I}}_h}{\partial s_m}(s_h) = -h^2\sum_{x_\alpha\in\bar{\Omega}_h\setminus\{x_{\alpha_0}\}} P_\alpha s_h(x_\alpha)\frac{\partial s_h}{\partial s_m}(x_\alpha) = -c_1h^2\sum_{x_\alpha\in\bar{\Omega}_h\setminus\{x_{\alpha_0}\}} P_\alpha s_h(x_\alpha)\psi_k(x_\alpha).$$

Take $\{\psi_i\}$ to be the the finite element basis functions $\{\phi_i\}$, and as we have assumed the finite difference and finite element nodes are identified, we define

$$s_h \coloneqq s(u_\hbar) = \frac{s_{\max}-s_{\min}}{2}\sum_{i=1}^{L} u_i\phi_i + \frac{s_{\max}+s_{\min}}{2},$$

and the result follows: $\partial\mathcal{I}_h/\partial u_m(u_\hbar) = \partial\tilde{\mathcal{I}}_h/\partial u_m(s_h) = (4.15)$. □

We use the elliptic relationships of (4.9) – (4.10) to describe derivatives of $w_\hbar$. In particular $\partial w_\hbar/\partial u_m$ satisfies

$$\int_\Omega \frac{\partial w_\hbar}{\partial u_m}\zeta_\hbar\,\mathrm{d}x = \int_\Omega \nabla\phi_m\cdot\nabla\zeta_\hbar\,\mathrm{d}x \qquad \forall\zeta_\hbar\in S_{D\hbar}(\Omega)\ \ (\text{resp } S_\hbar(\Omega)).$$

As $w_\hbar = \sum_{i=1}^{K} w_l\phi_l$, where $\{\phi_l\}$ are a basis of $S_\hbar(\Omega)$, set $\zeta_\hbar = \phi_l$. The chain rule gives

$$\int_\Omega \frac{\partial(w_\hbar^2)}{\partial u_m} = \int_\Omega \frac{\partial(w_\hbar^2)}{\partial w_\hbar}\frac{\partial w_\hbar}{\partial u_m} = 2\sum_{l=1}^{L} w_l\int_\Omega \phi_l\frac{\partial w_\hbar}{\partial u_m}\,\mathrm{d}x = 2\sum_{l=1}^{L} w_l\int_\Omega \nabla\phi_m\cdot\nabla\phi_l\,\mathrm{d}x. \tag{4.16}$$

Proposition 2. *Let $(u_\hbar = \sum_{l=1}^{L} u_i\phi_i, w_\hbar = \sum_{l=1}^{L} w_i\phi_i) \in \mathcal{A}_{\Delta,\hbar}$ defined by (4.8) and $m\in\{1,\dots,L\}$. Let $\mathcal{I}_{\varepsilon,h,\hbar}$ be defined by (4.5). For $P:\Omega_h\setminus\{x_{\alpha_0}\}\to\mathbb{R}$, the solution for the adjoint equations (4.13) – (4.14) for the discrete Eikonal equations (4.1) – (4.2) with slowness $s(u_\hbar)$. Then*

$$\begin{aligned}\frac{\partial\mathcal{I}^\gamma_{\varepsilon,h,\hbar}}{\partial u_m}(u_\hbar,w_\hbar) = \ & -h^2\frac{s_{\max}-s_{\min}}{2}\sum_{x_\alpha\in\Omega_h\setminus\{x_{\alpha_0}\}} P_\alpha s(u_\hbar)(x_\alpha)\phi_m(x_\alpha)\\ & + \sigma\sum_{l=1}^{L}\int_\Omega \gamma\varepsilon^3 w_l\nabla\phi_m\cdot\nabla\phi_l + \varepsilon u_l\nabla\phi_m\cdot\nabla\phi_l - \frac{1}{\varepsilon}u_l\phi_m\phi_l\,\mathrm{d}x,\end{aligned} \tag{4.17}$$

where,

$$\sum_{k,l=1}^{L} w_l \int_\Omega \phi_k \phi_l \,\mathrm{d}x = \sum_{k,l=1}^{L} u_l \int_\Omega \nabla\phi_k \cdot \nabla\phi_l \,\mathrm{d}x. \tag{4.18}$$

Proof. Insert the ansatz for $(u_\hbar, w_\hbar)$ into the regularization integral of (4.5) and differentiate with respect to u_m. In view of Proposition 1 and (4.16) we obtain (4.17) – (4.18). □

Remark 13. *Let $\boldsymbol{u} = (u_l)_l$ and $\boldsymbol{w} = (w_l)_l$, we can write the above in a matrix formulation*

$$\begin{aligned}\frac{\partial \mathcal{I}^\gamma_{\varepsilon,h,\hbar}}{\partial u_m}(u_\hbar, w_\hbar) = \; & -h^2 \frac{s_{\max} - s_{\min}}{2} \sum_{x_\alpha \in \Omega_h \setminus \{x_0\}} P_\alpha s(u_\hbar)(x_\alpha)\phi_m(x_\alpha) \\ & + \sigma\Big(\mathbb{S}(\gamma\varepsilon^3 \boldsymbol{w} + \varepsilon \boldsymbol{u}) - \frac{1}{\varepsilon}\mathbb{M}\boldsymbol{u}\Big)_m,\end{aligned} \tag{4.19}$$

and,

$$\mathbb{M}\boldsymbol{w} = \mathbb{S}\boldsymbol{u}. \qquad \Big(\textit{for } \mathbb{M}_{ij} = \int_\Omega \phi_i\phi_j \,\mathrm{d}x, \quad \mathbb{S}_{ij} = \int_\Omega \nabla\phi_i \cdot \nabla\phi_j \,\mathrm{d}x\Big). \tag{4.20}$$

We call $\mathbb{M}$ the mass matrix and $\mathbb{S}$ the stiffness matrix for the discretization.

4.5. An explicit descent scheme

We seek to find an critical point of the discrete functional, given by:

$$\begin{cases} u_m & = \Pi\Big(u_m - \alpha \frac{\partial \mathcal{I}^\gamma_{\varepsilon,h}}{\partial u_m}(u_\hbar, w_\hbar)\Big), \qquad \forall m = 1, \dots, L, \\ \mathbb{M}\boldsymbol{w} & = \mathbb{S}\boldsymbol{u} \end{cases} \tag{4.21}$$

where $\Pi : \mathbb{R} \to [-1, 1]$ is the pointwise projection

$$\Pi(r) := \max\{-1, \min\{r, 1\}\}.$$

We can write down a simple iterative scheme to find this critical point, by updating the phase field variables from the derivative (4.19) – (4.20). Denote coefficient vectors as $\boldsymbol{a} = (a_l)_l$. Set a numerical tolerance $\mathrm{Tol} > 0$, and set $\eta \in (0, 1)$, and $\alpha^{init} \in \mathbb{R}_+$. Define initial coefficient vector $\boldsymbol{u}^{(0)}$ so that $u^{(0)}_\hbar \in S_\hbar$ and set $\mathbb{M}\boldsymbol{w}^{(0)} = \mathbb{S}\boldsymbol{u}^{(0)}$, ensuring that $(\boldsymbol{u}^{(0)} \cdot \boldsymbol{\phi}, \boldsymbol{w}^{(0)} \cdot \boldsymbol{\phi}) \in \mathcal{A}_{\Delta,\hbar}$.

For each $k = 0, 1, 2, \dots$, do the following:

(i) Calculate the discrete functional derivative

$$\begin{aligned}\frac{\partial \mathcal{I}^\gamma_{\varepsilon,h}}{\partial u_m}(u^{(k)}_\hbar, w^{(k)}_\hbar) = \; & -h^2 \frac{s_{\max} - s_{\min}}{2} \sum_{x_\alpha \in \Omega \setminus \{x_0\}} P_\alpha s(u^{(k)}_\hbar)(x_\alpha)\phi_m(x_\alpha) \\ & + \sigma\Big(\mathbb{S}(\gamma\varepsilon^3 \boldsymbol{w}^{(k)} + \varepsilon \boldsymbol{u}^{(k)}) - \frac{1}{\varepsilon}\mathbb{M}\boldsymbol{u}^{(k)}\Big).\end{aligned} \tag{4.22}$$

(ii) Find the largest step $\alpha \in \{\frac{\alpha^{init}}{2^{j-1}} \mid j \in \mathbb{N}\}$ so that, if we define

$$\begin{cases} u_m^{(k+1)} & = \Pi\Big(u_m^{(k)} - \alpha \frac{\partial \mathcal{I}^{\gamma}_{\varepsilon,h}}{\partial u_m}(u_\hbar^{(k)}, w_\hbar^{(k)})\Big), \qquad \forall m = 1, \dots, L, \\ \mathbb{M}\boldsymbol{w}^{(k+1)} & = \mathbb{S}\boldsymbol{u}^{(k+1)}. \end{cases} \tag{4.23}$$

then, the following inequality is satisfied:

$$\mathcal{I}^{\gamma}_{\varepsilon,h}(u_\hbar^{(k+1)}, w_\hbar^{(k+1)}) - \mathcal{I}^{\gamma}_{\varepsilon,h}(u_\hbar^{(k)}, w_\hbar^{(k)}) < -\frac{\eta}{\alpha^2}\|u_\hbar^{(k+1)} - u_\hbar^{(k)}\|^2.$$

(iii) If $\alpha^{-2}\|u_\hbar^{(k+1)} - u_\hbar^{(k)}\|^2 <$ Tol we are done.

(iv) Otherwise go back to step 1. with $k \to k+1$

Remark 14. *One can see from (4.23) that our stopping criteria for the scheme will ensure that both the difference in final iterates $u_\hbar^{(.)}$ are small, and that the final iterate produces a small residual in (4.21).*

5. Numerical results

We validate our model choice and scheme by presenting numerical simulations in two dimensions. We begin by investigating the choice of important model parameters. We then illustrate some different geometries of the true slowness function with different source – receiver configurations designed to show the behaviour of recovery, as well as intuition into the reliability and limitations of solutions.

The solver for the forward problem was constructed in C++, and compiled into a MATLAB mex function. The inverse solver was then computed using MATLAB 2017b.

5.1. Parameter Study

5.1.1. Model Parameters We demonstrate binary recovery of a simple test case to give intuition into sensible model parameter choices. For this we shall use data and source receiver locations as described in Figure 2. We refer to the true field as 'circular disk', defined on $\Omega = [0,1] \times [0,1]$, with $s_{\min} = 2,\ s_{\max} = 4$, and is given by

$$s(x) = \begin{cases} s_{\max}, & (x - \frac{1}{2})^2 + (y - \frac{1}{2})^2 \le (\frac{1}{4})^2, \\ s_{\min}, & \text{otherwise.} \end{cases}$$

The source $x_0 = (1/2, 1/2)$ and data observed on all of $\partial\Omega$. We choose the misfit functional to be a boundary integral of $\partial\Omega$ as in (2.11), and unless specified, in the study we do not have noisy observations. We take our prior space $\mathcal{A}_{\Delta,\hbar}$ as in (4.8), where $\partial_D\Omega = \partial\Omega$. For the regularization we have $\varepsilon, \gamma, \sigma > 0$. The recovery of Figure 2 (right), had parameters $\gamma = 10^{-2}$, ε to produce an interface width of $1/20$, and $\sigma = 10^{-3}$ and with $\nu = 10^{-2}$ (1% noise on observations).

We discretize the inverse problem choosing $\hbar = 1/160$, and take $\Gamma_h = \partial\Omega_h$, for the discrete mismatch functional (4.4). We take $h = \hbar$, and avoid committing an 'inverse crime' by generating data from a forward problem solve on a fine mesh with $h_{\text{dat}} = h/8$. The receivers are densely placed on $\partial\Omega_h$ (spaced h apart). We solve the problem using

the scheme of Section 4.5, with tolerance Tol$= 10^{-12}$ and $\eta = 10^{-5}$, and $\alpha^{init} = 10^4$. We take the initial value of $u_\hbar^0 \equiv -1$ everywhere for all studies.

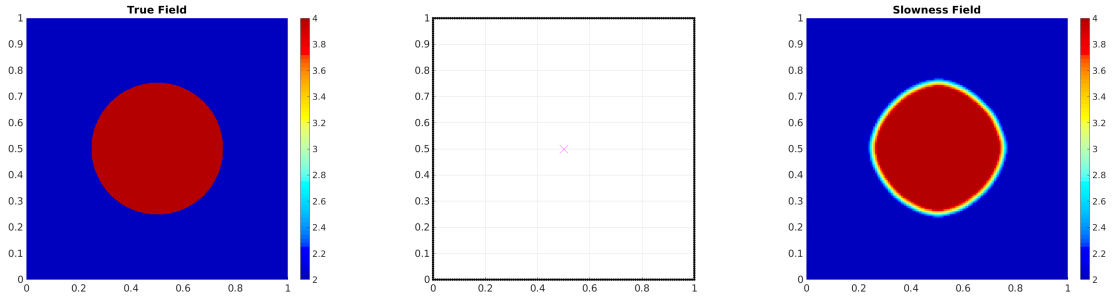

Figure 2. (a) The true slowness function. (b) Position of source (magenta cross) and observation points (black dots). (c) Example recovery from data perturbed with 1% noise at the observations, $\hbar = \frac{1}{160}$.

We first investigate the interfacial resolution. We do so by fixing $\gamma = 10^{-4}$, and construct δ^γ, by numerically solving (A.2). In view of the ansatz in Remark 5, we set $\varepsilon = K\hbar/(2\delta^\gamma)$, which produces an interface width of size $K\hbar$. The simulation run is displayed in Table 1. We observe the $\mathcal{I}_\varepsilon^\gamma$ decreases and appears to converge as interfacial width increases. We choose interfacial width at least $8\hbar$ for good accuracy. Results remained largely constant at different values of γ - so the higher order contributions need little extra resolution.

Table 1. Table listing the interface widths from different ε values, given a fixed value of $\gamma = 10^{-4}$. We display the size of both misfit and regularization terms, and the full functional $\mathcal{I}^\gamma_{\varepsilon_\hbar,h}$.

interface width	Misfit	$\sigma\mathcal{J}_\varepsilon^\gamma(u_h, w_h)$	$\mathcal{I}^\gamma_{\varepsilon_\hbar,h}(u_\hbar, w_\hbar)$
$4\hbar$	2.4442E-08	1.6022E-04	1.6024E-04
$6\hbar$	2.3324E-08	1.5754E-04	1.5756E-04
$8\hbar$	2.2235E-08	1.5656E-04	1.5658E-04
$10\hbar$	2.2193E-08	1.5609E-04	1.5611E-04
$12\hbar$	2.1949E-08	1.5584E-04	1.5586E-04
$14\hbar$	2.1876E-08	1.5568E-04	1.5570E-04

We now assess γ. For each γ, we choose ε to ensure constant interface width of $8\hbar$. The results are listed in Table 2. We see that one obtains consistent values of the functional below a value 10^{-2}, and there is significant difference for greater values. This indicates for consistency we should choose $\gamma \leq 10^{-2}$.

We next investigate σ. We take $\gamma = 10^{-2}$, and interface width $8\hbar$. We validate the choice of σ by using the approximation property of the regularization to the perimeter length of the interface. Theorem 4, shows $\mathcal{J}_\varepsilon^\gamma \xrightarrow{\Gamma} \mathcal{J}_0^\gamma$ and this is proportional to the interfacial length by a factor P^γ (see (A.3)). The truth in Figure 2, has interfacial

Table 2. Table listing varying values of γ, we choose ε appropriately in each simulation to yield interface width $8\hbar$. We display the size of both misfit and regularization terms, and the full functional $\mathcal{I}^{\gamma}_{\varepsilon_\hbar,h}$.

γ	Misfit	$\sigma\mathcal{J}^{\gamma}_{\varepsilon}(u_h, w_h)$	$\mathcal{I}^{\gamma}_{\varepsilon_\hbar,h}(u_h, w_h)$
1.0000E+00	2.3322E-08	1.5839E-04	1.5841E-04
1.0000E-01	2.2446E-08	1.5726E-04	1.5728E-04
1.0000E-02	2.1970E-08	1.5673E-04	1.5675E-04
1.0000E-03	2.2100E-08	1.5659E-04	1.5661E-04
1.000E-04	2.2235E-08	1.5656E-04	1.5658E-04

length $\pi/2$. Table 3 lists the results, along with a difference of the approximate and true interface length $|1/P^{\gamma}\mathcal{J}^{\gamma}_{\varepsilon} - \pi/2|$. We see the best choice of σ is around 10^{-3}. Moreover, accuracy is still good if σ is taken too small, but quickly poor if σ is taken too large.

Table 3. Table listing varying values of σ - the regularization parameter, we fix $\gamma = 10^{-2}$, and choose ε to yield interface width $8\hbar$. We display the size of both misfit and regularization terms, and the difference of predicted and true interfacial length: $\pi/2$.

σ	Misfit	$\sigma\mathcal{J}^{\gamma}_{\varepsilon}(u_h, w_h)$	$\lvert\frac{1}{P^{\gamma}}\mathcal{J}^{\gamma}_{\varepsilon} - \frac{\pi}{2}\rvert$
1.0000E-04	2.2112E-08	1.5748E-04	4.0037E-03
2.0000E-04	6.5955E-08	3.1489E-04	3.6537E-03
4.0000E-04	2.3649E-07	6.2956E-04	3.1037E-03
8.0000E-04	9.0711E-07	1.2582E-03	1.9537E-03
1.6000E-03	3.8804E-06	2.4975E-03	9.8588E-03

Finally we verify the scaling of σ when observations are subject to noise. This uncertainty is incorporated into the discrete misfit functional as in Section 2.6. The results are displayed in Table 4 for three levels of uncertainty with standard deviation $\nu = 1/200,\ 1/100$, and $1/50$ corresponding to 0.5%, 1%, 2% noise. We observe that for larger noise levels the accuracy decreases, as expected. We also observe that choosing the scaling of σ with $1/\nu^2$ is sensible. If one takes $\bar{\sigma}$ too large (4×10^{-3}) we once again encounter a large loss of accuracy.

We note that one must also balance the regularization and data misfit at different data contrasts, one can do this by rescaling σ by $2/(s_{\max} - s_{\min})$.

5.1.2. Discretization parameters We omit the tests of discretization parameters, but observed convergence in the case of fixed ε, and sending $\hbar = Ch \to 0$ for $1 \leq C \in \mathbb{N}$. The test scenario identical to the problem we created for the model parameters.

5.1.3. General parameter recommendations To summarize the previous study, and the intuition from our implementation experience we give advice for making parameter

Table 4. Table listing results produced with detector noise defined by the standard deviation ν of a centred normal random variable, and regularization parameter scaled by $\frac{1}{\nu^2}$. We fix $\gamma = 10^{-2}$, and choose ε to yield interface width $8\hbar$. We display the size of both misfit and regularization terms, and the difference of predicted and true interfacial length: $\pi/2$.

ν, if $\Sigma \sim N(0, \nu^2)$	$\bar{\sigma}$, ($\sigma := \frac{\bar{\sigma}}{\nu^2}$)	Misfit	$\frac{\sigma}{\nu^2}\mathcal{J}_\varepsilon^\gamma(u_h, w_h)$	$\lvert\frac{1}{P^\gamma}\mathcal{J}_\varepsilon^\gamma - \frac{\pi}{2}\rvert$
5.0000E-03	1.0000E-03	4.5291E-01	6.2873E+01	1.0387E-03
1.0000E-02	1.0000E-03	3.7864E-01	1.5734E+01	2.5737E-03
2.0000E-02	1.0000E-03	3.3873E-01	3.9373E+00	4.1237E-03
5.0000E-03	2.0000E-03	5.8286E-01	1.2561E+02	6.8133E-04
1.0000E-02	2.0000E-03	3.9781E-01	3.1382E+01	1.7163E-03
2.0000E-02	2.0000E-03	3.4077E-01	7.8455E+00	1.6963E-03
5.0000E-03	4.0000E-03	1.2109E+00	2.5020E+02	7.0463E-03
1.0000E-02	4.0000E-03	5.8052E-01	6.2531E+01	7.5288E-03
2.0000E-02	4.0000E-03	4.1042E-01	1.5649E+01	5.8963E-03

choices. First we note that ε and σ are the most important parameters for the modelling aspect. The parameter γ serves as a theoretical tool primarily, and the interfacial profile is largely insensitive to it, we choose $\gamma = 10^{-2}$ and have not observed experimentally where changing this value was necessary.

The role of σ balances the relative weight between misfit and regularization. If σ is small, it favours a high fidelity to the data, and if it is large the problem is weighted towards high regularity interfaces (and prevents overfitting to noisy data). The sensitivity to σ is a feature present in inverse problems involving the addition of regularization functionals. Choice of σ is therefore problem specific and we do not provide further intuition for choosing it.

ϵ characterizes the interfacial width, and from the modelling perspective one is using this to approximate zero, with this in mind it should be taken small. In the parameter study we found it important to resolve the interfacial profile, and we see a good rule is to take $\hbar = \varepsilon/8$. Therefore we recommend one takes ε as small as possible so that the computation is feasible with the corresponding numerical grid size.

5.2. Different geometries

We have several different geometries that we wish to recover, which we shall use as "truths" for our inverse problem. We use the discretization parameters as in the previous section, and generate data on a mesh with $h_{\text{dat}} = h/8$. The choices of model parameter configuration are influenced by what we discovered in the parameter study. We take $\nu = 10^{-2}$ for the noise, we also choose (unless otherwise stated) $s_{\min} = 1$, $s_{\max} = 1.1$ $\sigma = 10^{-4}$, $\gamma = 10^{-2}$, and ε so that we have the interface width $8\hbar$. We use two source – receiver configurations as demonstrated in Figure 3 - with receivers densely spaced on the boundary segments.

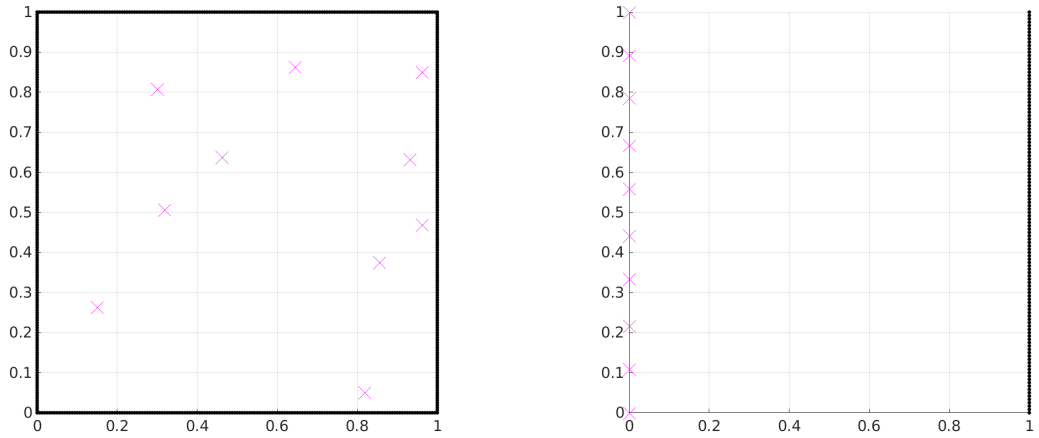

Figure 3. (a) Random: dense boundary observations on $\partial\Omega$; 10 source locations in Ω produced from MATLAB R2017b random number generator (seed=12131415, generator=twister). (b) Wells: dense boundary observations on right hand wall of $\partial\Omega$; 10 sources equidistributed on left hand wall of $\partial\Omega$.

We define four more slowness function "truths":

(i) Banded layers. We take this shape from [38]. The banded regions are sections of two annuli. That is, $\forall(x,y)\in[0,1]$

$$\text{if } \begin{cases} \frac{1}{2}\Big(3.7-\sqrt{2.6^2-(2x-1)^2}\Big)\le y\le \frac{1}{2}\Big(4.1-\sqrt{2.6^2-(2x-1)^2}\Big), \text{ or} \\ \frac{1}{2}\Big(2.8-\sqrt{2.6^2-(2x-1)^2}\Big)\le y\le \frac{1}{2}\Big(3.2-\sqrt{2.6^2-(2x-1)^2}\Big), \end{cases}$$

then $s(x,y)=s_{\max}$, otherwise $s(x,y)=s_{\min}$. The results are found in Figure 5.

(ii) Right angle. This example shape is similar to [23]. $\forall(x,y)\in[0,1]$

$$\text{if } \begin{cases} y\ge \frac{2}{3}x+0.4, \text{ or} \\ y\ge -\frac{3}{2}x+0.9, \end{cases}$$

then $s(x,y)=s_{\max}$, otherwise $s(x,y)=s_{\min}$. The results are found in Figure 6.

(iii) Arbitrary shape. $\forall(x,y)\in[0,1]$

$$\text{if } \begin{cases} (x-\frac{2}{3})^2+(y-\frac{1}{2})^2\le \frac{1}{5^2}, & \text{or} \\ (x-\frac{7}{15})^2+(y-\frac{7}{10})^2\le \frac{1}{6^2}, & \text{or} \\ (x-\frac{7}{15})^2+(y-\frac{3}{10})^2\le \frac{1}{8^2}, \end{cases}$$

then $s(x,y)=s_{\max}$, otherwise $s(x,y)=s_{\min}$. The results are found in Figure 7.

(iv) Shielded disk. $\forall(x,y)\in[0,1]$

$$\text{if } (x-\frac{2}{3})^2+(y-\frac{1}{2})^2\le\frac{1}{6^2}, \text{ or if both } \begin{cases} (x-\frac{2}{3})^2+(y-\frac{1}{2})^2\le\Big(\frac{3}{8}\Big)^2, \\ (x-\frac{4}{9})^2+(y-\frac{1}{2})^2\ge\frac{1}{4^2}, \end{cases}$$

then we set $s(x,y) = s_{\max}$, otherwise $s(x,y) = s_{\min}$. Here we choose values $s_{\min} = 1$, $s_{\max} = 1.4$ to aid recovery in the wells configuration. The results are found in Figure 8, (for $s_{\max} = 1.1$ in wells configuration see Figure 10(a)).

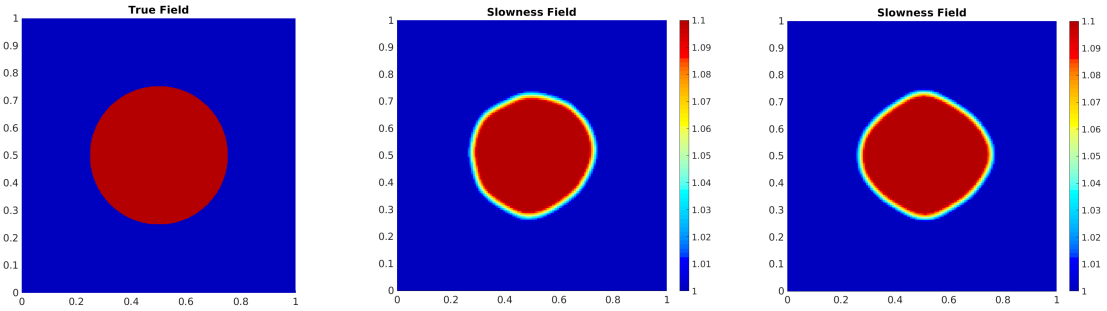

Figure 4. Circular disk recovery. (a) true slowness field, $s_{\min} = 1$, $s_{\max} = 1.1$. (b) slowness field from random configuration. (c) slowness field from wells configuration

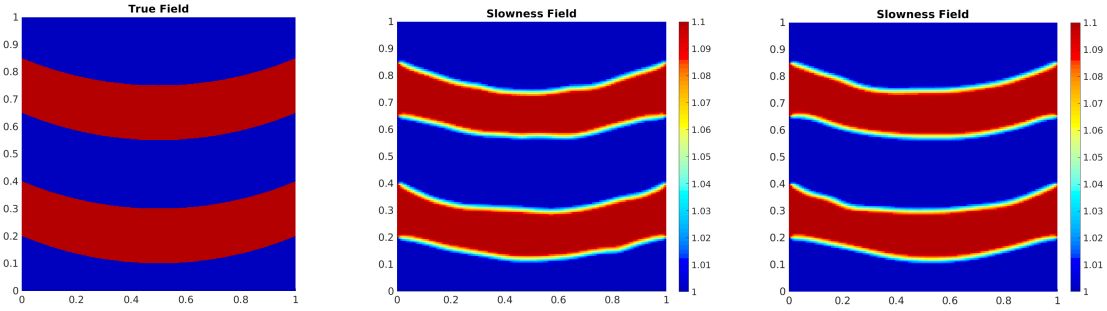

Figure 5. Right angle recovery. (a) true slowness field, $s_{\min} = 1$, $s_{\max} = 1.1$. (b) slowness field from random configuration. (c) slowness field from wells configuration

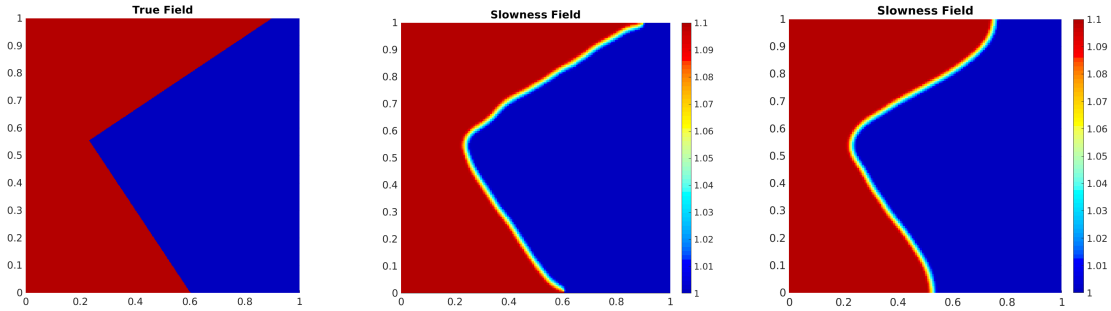

Figure 6. Angular boundary recovery. (a) true slowness field $s_{\min} = 1$, $s_{\max} = 1.1$. (b) slowness field from random configuration. (c) slowness field from wells configuration

The different geometries provide a variety of challenges for the first-arrival traveltime binary recovery problem. The first recovery is of the circular disk, in

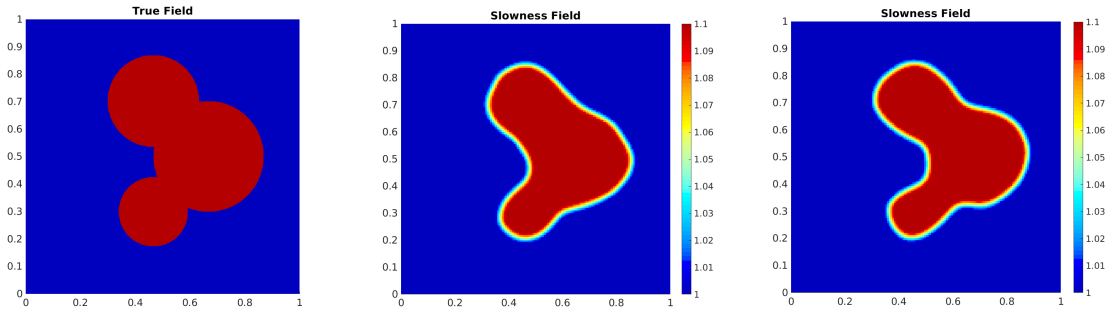

Figure 7. Arbitrary shape recovery. (a) true slowness field $s_{\min} = 1,\ s_{\max} = 1.1$. (b) slowness field from random configuration. (c) slowness field from wells configuration

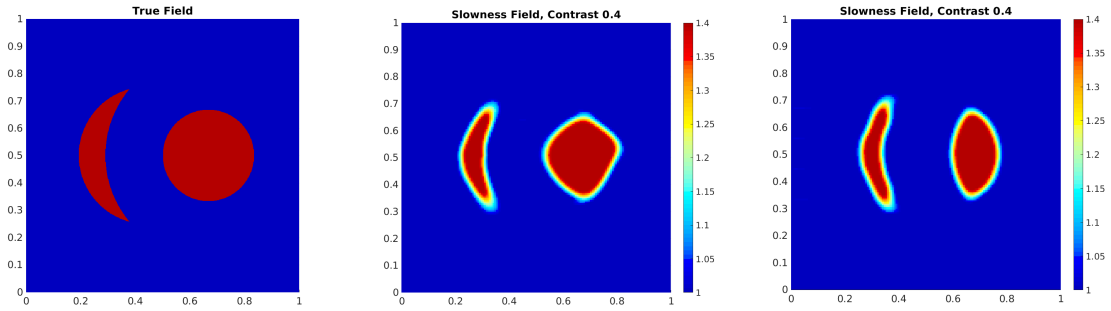

Figure 8. Shielded disk recovery. (a) true slowness field $s_{\min} = 1,\ s_{\max} = 1.4$. (b) slowness field from random configuration. (c) slowness field from wells configuration

Figure 4. We see good recovery in both configurations for this simple geometry. For more complicated inclusions such as Figure 7 we similarly obtain reasonable recovery so long as the interfacial layer is thin relative to the lengthscale of geometric features.

We observe two features of the underlying forward problem in Figure 5. Firstly we see in locations where wave sources are near to interfaces we obtain geometric asymmetry. Secondly, the scheme is efficient if ray paths travel through relatively homogeneous structures and so we found the recovery in the wells configuration is quickly resolved.

The right angle of Figure 6 is well recovered, and we see the effects of the regularization on the boundary conditions. If one takes $\partial_D\Omega_\hbar = \partial\Omega$ then interface positions match the truth (see Figure 6(b)), but when the set $\{y = 0\} \cup \{y = 1\}$ has Neumann conditions (see Figure 6(c)) we see interfaces meet domain boundaries at right angles. In this simulation it is vital for $\partial_D\Omega_\hbar \neq \emptyset$ to obtain a correct local minimizer.

The shielded disk of Figure 8 performs well in the random configuration. In the wells configuration, the disk is shielded by the crescent inclusion and yet the inverse solver still distinguishes two objects and shapes are well recovered. We note this simulation performed better with contrast $s_{\min} = 1$, $s_{\max} = 1.4$, and $\sigma = 5 \times 10^{-4}$.

Remark 15. *We may work under the assumptions that we know only the contrast between slowness phases, and we have slowness(es) S given by the boundary conditions.*

We first remark that, for features of interest contained entirely in the domain interior, we may perform recovery with 'background' material with slowness $S = s_{\min}$ (as in our experiments here) or with $S = s_{\max}$. We further remark that one may quickly determine in which setting ($S = s_{\min}$ or $S = s_{\max}$) we are in, as we see numerically that the incorrect choice quickly obtains a single phase solution. If we further assume that we know an approximate contrast, we observe numerically, stretching and compression of the shape due to under- or overestimation of true slowness, as expected.

Remark 16. *We observe a key phenomena known in FATT: We fix a source – receiver configuration and obtain data from a binary truth with chosen (low) slowness contrast between phases. We then increase the contrast and find that there is a contrast value above which, first-arrival traveltime tomography becomes ineffective. This value depends on the geometry of the interface and the position of source – receivers.*

To illustrate this, we display traveltime fields in Figure 9, from to a single source at $[0, 1/2]$; a single receiver is placed at $[1, 1/2]$. The slowness is given by a continuous approximation of the shielded disk Figure 8(a), with a fixed binary value $s_{\min} = 1$; we now choose values of $s_{\max}$ and define the slowness contrast as $(s_{\max} - s_{\min})/s_{\min}$. For high contrasts (Figure 9(b) and (c)) the obstacles strongly impede the wave, and we see the maximum traveltime occurs within the obstacle. Therefore the wave's first hitting time at the receiver does not have a ray path through the obstacle. This loss of information affects the recovery, see Figure 10 for a wells configuration. Impenetrable obstacles will be investigated in a forthcoming work.

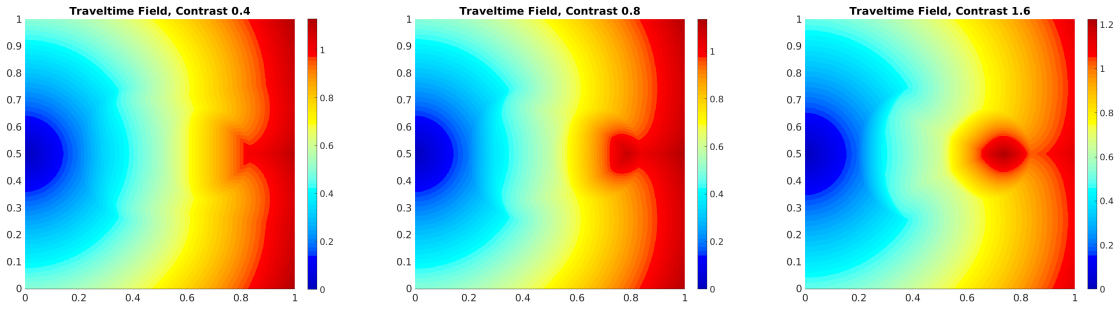

Figure 9. Traveltimes for shielded disk. Source located at $[0, 1/2]$ in the domain. (a) to (c): increasing contrast ($(s_{\max} - s_{\min})/s_{\min}$) values 0.4, 0.8, 1.6. Contrast ≥ 0.8 has maximal traveltime in the domain.

6. Conclusion

6.1. Closing remarks

We have presented a technique for binary recovery based on the phase field methodology with an emphasis of the presence of an underlying mathematical theory.

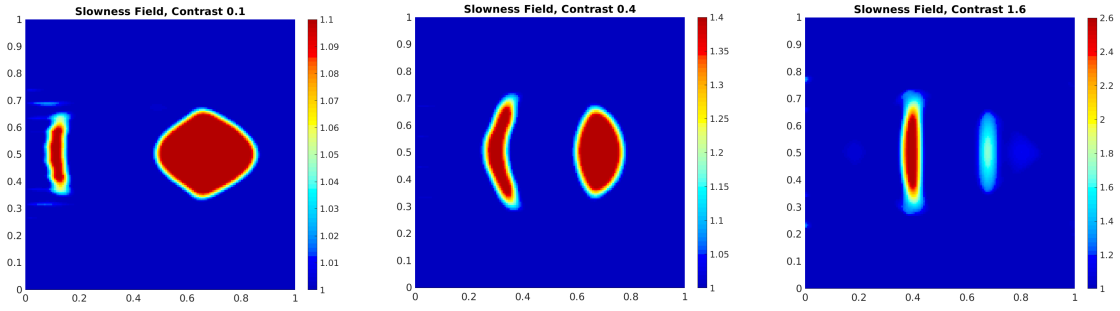

Figure 10. Different contrast ratio for shielded disk, with wells configuration and $s_{\min} = 1$. (a) $s_{\max} = 1.1$, (b) $s_{\max} = 1.4$, (c) $s_{\max} = 2.6$. Situations of too low, acceptable, and too high contrast respectively.

We proved that the method ensures that the solutions exist, and that the forward problem remains well posed for them. We validated the technique with a $\Gamma-$convergence of the phase field regularization to a perimeter penalization technique. We have created a mixed formulation of the problem, and shown that minimizers of this solve the original problem. We have constructed a convergent discrete formulation relying on a monotone finite difference method for the forward problem and mixed finite element method for the inverse problem. Due to our careful treatment of the derivative we were able to use a descent algorithm and we demonstrated its effectiveness of recovery in many geometries and for different configurations of source – receiver pairs, including one found in crosswell tomography.

We would like to emphasize the flexibility of the class of phase field regularizations, which is are not limited to use on the forward problem under investigation. This is demonstrated by the scope of work in the following section.

6.2. Outlook

The framework we have set up suggests several developments worthy of future study. For example, investigation of a hierarchical forward model in which $s_{\max}$ is unknown, as in [38]. One may investigate multiscale forward models, where constant regions contain small perturbations, for example by setting $s_{\max}(x) = \tilde{s}_{\max} + \delta f(x)$ for a continuous function f to be determined with $\delta \ll (\tilde{s}_{\max} - s_{\min})$, similar to the setting of [2].

A natural extension is to investigate a piecewise constant slowness function formulated with a multi-obstacle potential. Modelling impenetrable obstacles ($s_{\max} \to \infty$) - as noted in Remark 16, could also extend the applicability of the tomography. Another direction for model extension is to consider interfacial reflectors within the domain, where one must use reflection data to additionally recover the position of the reflector, as treated by [39, 56]

In practice, gravity acceleration data is also collected in geophysical surveys and one can perform a joint inversion for the slowness and subsurface density. It is a viable extension as the slowness and density are structurally related (with limitations however

[8]), and will share curves of discontinuity. Inversion is performed a weighted sum of the misfit functionals for slowness and gravity acceleration [40].

Acknowledgments

The work of CME was partially supported by the Royal Society via a Wolfson Research Merit Award; the work of ORAD and CME by the EPSRC Programme Grant EQUIP.

References

[1] D. Adalsteinsson and J. A. Sethian, *A fast level set method for propagating interfaces*, Journal of Computational Physics, 118 (1995), pp. 269–277.

[2] H. Aghamiry, A. Gholami, and S. Operto, *Hybrid tikhonov+ total-variation regularization for imaging large-contrast media by full-waveform inversion*, in SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 2018, pp. 1253–1257.

[3] K. Aki and P. G. Richards, *Quantitative seismology*, University Science Books, 2002.

[4] H. W. Alt, *Linear functional analysis: An application oriented introduction*, Universitext, Springer, 2016.

[5] G. Antoine, *Blocky regularization schemes for full-waveform inversion*, Geophysical Prospecting, 60 (2011), pp. 870–884.

[6] G. Bellettini, A. Braides, and G. Riey, *Variational approximation of anisotropic functionals on partitions*, Annali di Matematica Pura ed Applicata, 184 (2005), pp. 75–93.

[7] E. Beretta, L. Ratti, and M. Verani, *A phase-field approach for the interface reconstruction in a nonlinear elliptic problem arising from cardiac electrophysiology*, arXiv:1709.05646, (2017).

[8] F. Billette and S. Brandsberg-Dahl, *The 2004 bp velocity benchmark*, in 67th EAGE Conference & Exhibition, 2005.

[9] L. Blank, H. Garcke, L. Sarbu, and V. Styles, *Primal-dual active set methods for Allen-Cahn variational inequalities with nonlocal constraints*, Numerical Methods for Partial Differential Equations, 29 (2013), pp. 999–1030.

[10] J. F. Blowey and C. M. Elliott, *The Cahn-Hilliard gradient theory for phase separation with non-smooth free energy part i: Mathematical analysis*, European Journal of Applied Mathematics, 2 (1991), pp. 233–280.

[11] ——, *The Cahn-Hilliard gradient theory for phase separation with non-smooth free energy part ii: Numerical analysis*, European Journal of Applied Mathematics, 3 (1992), pp. 147–179.

[12] ——, *Curvature dependent phase boundary motion and parabolic double obstacle problems*, in Degenerate Diffusions, W.-M. Ni, L. A. Peletier, and J. L. Vazquez, eds., New York, NY, 1993, Springer New York, pp. 19–60.

[13] C. Brett, C. M. Elliott, and A. S. Dedner, *Phase field methods for binary recovery*, in Optimization With PDE Constraints, vol. 101, Springer, 2014, pp. 25–63.

[14] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, vol. 15, Springer Science & Business Media, 2012.

[15] T. Bui-Thanh, O. Ghattas, J. Martin, and G. Stadler, *A computational framework for infinite-dimensional Bayesian inverse problems part I: The linearized case, with application to global seismic inversion*, SIAM Journal on Scientific Computing, 35 (2013), pp. A2494–A2523.

[16] M. Burger, T. Esposito, and C. I. Zeppieri, *Second-order edge-penalization in the Ambrosio-Tortorelli functional*, Multiscale Modeling & Simulation, 13 (2015), pp. 1354–1389.

[17] I. Capuzzo-Dolcetta and P.-L. Lions, *Hamilton-Jacobi equations with state constraints*, Transactions of the American Mathematical Society, 318 (1990), pp. 643–683.

[18] M. Chermisi, G. D. Maso, I. Fonseca, and G. Leoni, *Singular perturbation models in*

phase transitions for second-order materials, Indiana University Mathematics Journal, 60 (2011), pp. 367–409.

[19] P. G. Ciarlet, *Linear and nonlinear functional analysis with applications*, vol. 130, Siam, 2013.

[20] M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Transactions of the American Mathematical Society, 277 (1983), pp. 1–42.

[21] K. Deckelnick and C. M. Elliott, *Uniqueness and error analysis for Hamilton-Jacobi equations with discontinuities*, Interfaces and free boundaries, 6 (2004), pp. 329–349.

[22] K. Deckelnick, C. M. Elliott, and V. Styles, *Numerical analysis of an inverse problem for the Eikonal equation*, Numerische Mathematik, 119 (2011), p. 245.

[23] K. Deckelnick, C. M. Elliott, and V. Styles, *Double obstacle phase field approach to an inverse problem for a discontinuous diffusion coefficient*, Inverse Problems, 32 (2016), p. 045008.

[24] M. M. Dunlop, C. M. Elliott, V. H. Hoang, and A. M. Stuart, *Reconciling Bayesian and total variation methods for binary inversion*, arXiv:1706.01960v2, (2018).

[25] H. Federer, *Geometric Measure Theory*, Springer Berlin Heidelberg, 1996.

[26] I. Fonseca and C. Mantegazza, *Second-order singular perturbation models for phase transitions*, SIAM Journal on Mathematical Analysis, 31 (2000), pp. 1121–1143.

[27] G. Gaullier, P. Charbonnier, F. Heitz, and P. Côte, *Introducing shape constraints into object-based traveltime tomography*, Inverse Problems, 32 (2016), p. 095002.

[28] P. Grisvard, *Elliptic problems in nonsmooth domains*, vol. 69, SIAM, 2011.

[29] D. Hilhorst, L. A. Peletier, and R. Schätzle, *Γ-limit for the extended Fisher-Kolmogorov equation*, Proceedings of the Royal Society of Edinburgh: Section A Mathematics, 132 (2002), pp. 141–162.

[30] S.-R. Hysing and S. Turek, *The Eikonal equation: numerical efficiency vs. algorithmic complexity on quadrilateral grids*, in Proceedings of ALGORITMY, vol. 22, 2005.

[31] M. Iglesias, Y. Lu, and A. Stuart, *A Bayesian level set method for geometric inverse problems*, Interfaces and Free Boundaries, 18 (2016), pp. 181–217.

[32] W.-K. Jeong and R. T. Whitaker, *A fast iterative method for Eikonal equations*, SIAM Journal on Scientific Computing, 30 (2008), pp. 2512–2534.

[33] W. Jiang and J. Zhang, *First-arrival traveltime tomography with modified total-variation regularization*, Geophysical Prospecting, 65 (2017), pp. 1138–1154.

[34] W. Jiang and J. Zhang, *3D first-arrival traveltime tomography with modified total-variation regularization*, Journal of Geophysics and Engineering, 15 (2018), p. 207.

[35] H. Lan and Z. Zhang, *Topography-dependent Eikonal equation and its solver for calculating first-arrival traveltimes with an irregular surface*, Geophysical Journal International, 193 (2013), pp. 1010–1026.

[36] P. G. Lelièvre, C. G. Farquharson, and C. A. Hurich, *Joint inversion of seismic traveltimes and gravity data on unstructured grids with application to mineral exploration*, Geophysics, 77 (2012), pp. K1–K15.

[37] S. Leung and J. Qian, *An adjoint state method for three-dimensional transmission traveltime tomography using first-arrivals*, Commun. Math. Sci., 4 (2006), pp. 249–266.

[38] W. Li and S. Leung, *A fast local level set adjoint state method for first arrival transmission traveltime tomography with discontinuous slowness*, Geophysical Journal International, 195 (2013), pp. 582–596.

[39] W. Li, S. Leung, and J. Qian, *A level-set adjoint-state method for crosswell transmission-reflection traveltime tomography*, Geophysical Journal International, 199 (2014), pp. 348–367.

[40] W. Li and J. Qian, *Joint inversion of gravity and traveltime data using a level-set-based structural parameterization*, Geophysics, 81 (2016), pp. G107–G119.

[41] Y. Lin and A. Ortega, *Object-based high contrast traveltime tomography*, IEEE Transactions on Computational Imaging, 3 (2017), pp. 738–748.

[42] P.-L. Lions, *Generalized solutions of Hamilton-Jacobi equations*, vol. 69 of Research Notes in Mathematics, Pitman (Advanced Publishing Program), Boston, Mass.-London, 1982.

[43] P.-L. Lions, *Neumann type boundary conditions for Hamilton-Jacobi equations*, Duke Math. J., 52 (1985), pp. 793–820.
[44] S. Luo, J. Qian, and R. Burridge, *High-order factorization based high-order hybrid fast sweeping methods for point-source Eikonal equations*, SIAM Journal on Numerical Analysis, 52 (2014), pp. 23–44.
[45] L. Modica, *The gradient theory of phase transitions and the minimal interface criterion*, Archive for Rational Mechanics and Analysis, 98 (1987), pp. 123–142.
[46] N. Rawlinson and M. Sambridge, *Seismic traveltime tomography of the crust and lithosphere*, Advances in geophysics, 46 (2003), pp. 81–198.
[47] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D: Nonlinear Phenomena, 60 (1992), pp. 259–268.
[48] J. Sethian and A. Vladimirsky, *Fast methods for the Eikonal and related Hamilton-Jacobi equations on unstructured meshes*, Proceedings of the National Academy of Sciences of the United States of America, 97 (2000), pp. 5699–5703.
[49] J. A. Sethian, *Theory, algorithms, and applications of level set methods for propagating interfaces*, Acta numerica, 5 (1996), pp. 309–395.
[50] J. A. Sethian, *Fast marching methods*, SIAM Review, 41 (1999), pp. 199–235.
[51] P. M. Shearer, *Introduction to seismology*, Cambridge University Press, 2009.
[52] H. M. Soner, *Optimal control with state-space constraint I*, SIAM Journal on Control and Optimization, 24 (1986), pp. 552–561.
[53] A. M. Stuart, *Inverse problems: A Bayesian perspective*, Acta Numerica, 19 (2010), pp. 451–559.
[54] C. Taillandier, M. Noble, H. Chauris, and H. Calandra, *First-arrival traveltime tomography based on the adjoint-state method*, Geophysics, 74 (2009), pp. WCB1–WCB10.
[55] H. Zhao, *A fast sweeping method for Eikonal equations*, Mathematics of Computation, 74 (2005), pp. 603–627.
[56] H. Zhao and Y. Zhong, *A hybrid adaptive phase space method for reflection traveltime tomography*, SIAM Journal on Imaging Sciences, 12 (2019), pp. 28–53.
[57] P. Zheglova, C. G. Farquharson, and C. A. Hurich, *2-D reconstruction of boundaries with level set inversion of traveltimes*, Geophysical Journal International, 192 (2013), pp. 688–698.
[58] T. Zhu and J. M. Harris, *Applications of boundary-preserving seismic tomography for delineating reservoir boundaries and zones of CO_2 saturation*, Geophysics, 80 (2015), pp. M33–M41.

Appendix A. Proof of Γ – convergence

We prove Theorem 4 by first proving a key result about the transition energy.

Lemma 1. *For $\gamma > 0$, there exists a solution $z \in \mathcal{V}$ to the variational inequality*

$$\gamma(z'', \eta'' - z'') + (z', \eta' - z') - (z, \eta - z) \geq 0, \qquad \forall \eta \in H^2(\mathbb{R}, [-1, 1]).$$

where

$$\mathcal{V} = \Big\{ v \in C^2(\mathbb{R}; [-1,1]) \mid \exists \delta > 0, \forall x \in \mathbb{R}\ v(x) = -v(-x), \\ v'(x) \geq 0,\ \textit{and}\ v(x > \delta) = 1,\ v(x < -\delta) = -1 \Big\}.$$

Proof. For $z \in \mathcal{V}$, we can relate z satisfying the critical point, to solutions Z of the following Fisher-Kolmogorov equation. Let $\delta^\gamma > 0$, then find $Z \in C^2(\mathbb{R})$ such that,

$$\gamma Z'''' - Z'' + \psi(Z) = 0 \ \text{ on } (-\delta^\gamma, \delta^\gamma), \qquad \text{where } \psi(s) \coloneqq -s, \tag{A.1}$$

subject to the constraints and boundary conditions imposed by $\mathcal{V}$:

$$Z(x < -\delta^\gamma) = -1, \qquad Z(x > \delta^\gamma) = 1, \qquad Z^{(k)}(-\delta^\gamma) = Z^{(k)}(\delta^\gamma) = 0, \quad \text{for } k = 1, 2.$$

with $|Z| \leq 1$, and $Z'(x) \geq 0$, $\forall x \in (-\delta^\gamma, \delta^\gamma)$, Z odd. To find this, we look at the characteristic equation: $\gamma\lambda^4 - \lambda^2 - 1 = 0$. For $\gamma > 0$ we have two real $\pm\lambda_1^\gamma$and two imaginary roots $\pm\lambda_2^\gamma i$.

$$\lambda_1^\gamma = \sqrt{\frac{1}{2\gamma}(1 + \sqrt{1 + 4\gamma})}, \qquad \lambda_2^\gamma = \sqrt{\frac{1}{2\gamma}(\sqrt{1 + 4\gamma} - 1)}.$$

Thus, for all $\gamma > 0$ we have a general form of solution

$$Z^\gamma(x) = C_1^\gamma e^{\lambda_1^\gamma x} + C_2^\gamma e^{-\lambda_1^\gamma x} + C_3^\gamma \cos(\lambda_2^\gamma x) + C_4^\gamma \sin(\lambda_2^\gamma x).$$

The constants C_k^γ depend on δ^γ. A general odd solution must therefore be of the form $Z_{odd}^\gamma(x) = \tilde{C}_1^\gamma \sinh(\lambda_1^\gamma x) + \tilde{C}_2^\gamma \sin(\lambda_2^\gamma x)$. Using the boundary conditions, one finds coefficients of the form:

$$\tilde{C}_1^\gamma = \frac{(\lambda_2^\gamma)^2}{((\lambda_1^\gamma)^2 + (\lambda_2^\gamma)^2)\sinh(\lambda_1^\gamma \delta^\gamma)}, \qquad \tilde{C}_2^\gamma = \frac{(\lambda_1^\gamma)^2}{((\lambda_1^\gamma)^2 + (\lambda_2^\gamma)^2)\sin(\lambda_2^\gamma \delta^\gamma)}.$$

We require solutions to be $C^2(\mathbb{R})$, so we seek δ^γ so that $(Z^\gamma)'(\delta^\gamma) = (Z^\gamma)''(\delta^\gamma) = 0$. This holds where

$$\lambda_2^\gamma \tan(\lambda_2^\gamma \delta^\gamma) = -\lambda_1^\gamma \tanh(\lambda_1^\gamma \delta^\gamma). \tag{A.2}$$

We take δ^γ to be the first positive solution for (A.2), and note $\delta^\gamma \in (\pi/(2\lambda_2^\gamma), \pi/\lambda_2^\gamma)$ and it is unique over the interval range. Z^γ is in fact strictly monotonic for any $\gamma > 0$ on $(0, \delta^\gamma)$. One may see this from the second derivative, which is negative on the interval, thus the $(Z^\gamma)'(x)$ has a minimum at δ^γ (where it vanishes), and so Z^γ is strictly increasing. To summarize we have a found a unique strictly monotonic solution for the 1D extended obstacle Fisher-Kolmogorov equation. It is bounded on $[-1, 1]$ and for $\gamma > 0$, can be extended to a function $z^\gamma \in C^2(\mathbb{R}; [-1, 1])$ by

$$z^\gamma(t) \coloneqq \begin{cases} -1, & \text{if } t < -\delta^\gamma, \\ \tilde{C}_1^\gamma \sinh(\lambda_1^\gamma t) + \tilde{C}_2^\gamma \sin(\lambda_2^\gamma t), & \text{if } -\delta^\gamma \leq t \leq \delta^\gamma, \\ 1, & \text{if } t > \delta^\gamma. \end{cases}$$

and so $z^\gamma \in \mathcal{V}$. □

Remark 1. *We have found an odd critical point of the variational inequality, and we have that it is unique for all $\gamma > 0$. The structure of the Euler – Lagrange equations for j^γ can be written as the variational inequality of Lemma 1, and so the odd minimizer of j^γ has this form. This follows the idea presented in [29].*

Remark 2. *As $\gamma \to 0$, we have $\delta^\gamma \searrow \frac{\pi}{2}$ and $\tilde{C}_2^\gamma \to 1$ and $\tilde{C}_1^\gamma \to 0$ and therefore $z^\gamma(t)$ converges to $\sin(t)$ on $(-\frac{\pi}{2}, \frac{\pi}{2})$. This is as expected as the fourth order problem reduces to the second order problem seen in [10].*

To prove Theorem 4, we must prove two inequalities. In [29], the authors have a complete proof for the smooth double well potential, so we provide detail of where the double obstacle theory differs.

Lemma 2. *(liminf inequality) Given a sequence* $\{u_\varepsilon\}$ *with* $u_\varepsilon \to u$ *as* $\varepsilon \to 0$ *strongly in* $L^1(\Omega)$*, then*

$$\mathcal{J}_0^\gamma \leq \liminf_{\varepsilon\to 0} \mathcal{J}_\varepsilon^\gamma(u_\varepsilon).$$

Proof. This follows immediately from the proof of [29, Proposition 3.2], which does not explicitly rely upon the double well or obstacle potential. □

Lemma 3. *(limsup inequality) For any* $u \in L^1(\Omega)$*, there exists a sequence* $\{u_\varepsilon\}$ *such that:*

(i) $u_\varepsilon \to u$ *as* $\varepsilon \to 0$ *strongly in* $L^1(\Omega)$*.*
(ii) $\mathcal{J}_0^\gamma(u) \geq \limsup_{\varepsilon\to 0} \mathcal{J}_\varepsilon^\gamma(u_\varepsilon)$*.*

Proof. We follow the proof of [29]: Let $u \in L^1(\Omega)$, we must construct a sequence $\{u_\varepsilon\}$ such that $\lim_{\varepsilon\to 0} u_\varepsilon = u$ in $L^1(\Omega)$ and

$$\limsup_{\varepsilon\to 0} \mathcal{J}_\varepsilon^\gamma(u_\varepsilon) \leq \mathcal{J}_0^\gamma(u).$$

Due to constructions in [45] there exists a set $D \subset \mathbb{R}^d$ open and bounded with $\partial D \in C^\infty$ and $\mathcal{H}^{d-1}(\partial D \cap \partial\Omega) = 0$, such that

$$u = \chi_D - \chi_{\mathbb{R}^d \setminus D}.$$

Now, Let $U \in \mathcal{V}$ be an odd minimizer of the functional j^γ, as discussed in Lemma 1 and Remark 1. By definition $P^\gamma = j^\gamma(U)$. With the specific form for u, (and $|\nabla \cdot|$ in the sense of total variation), we rewrite the limit $\mathcal{J}_0^\gamma(u)$

$$\mathcal{J}_0^\gamma(u) = \frac{1}{2}P^\gamma \int_\Omega |\nabla u| = P^\gamma \int_\Omega |\nabla \chi_D| = j^\gamma(U)\mathcal{H}^{d-1}(\partial D \cap \Omega). \tag{A.3}$$

Let d be the signed distance function to ∂D,

$$d(x) = \begin{cases} \inf\limits_{y\in\partial D} |x-y|, & \text{if } x \in D, \\ -\inf\limits_{y\in\partial D} |x-y|, & \text{if } x \notin D. \end{cases}$$

There exists a neighbourhood N_h of width h to ∂D where d is $C^2(N_h)$, and we define a function $\eta\colon \bar{\Omega} \to \mathbb{R}$

$$\eta(x) = d(x), \quad \text{if } x \in N_h, \qquad |\eta(x)| \geq h \quad \text{if } x \notin N_h.$$

We may extend outside of N_h to ensure $\eta \in C^2(\bar{\Omega})$. As mentioned in Remark 5, a useful rescaling is defined through the following sequence:

$$u_\varepsilon(x) = U\left(\frac{\eta(x)}{\varepsilon}\right), \qquad x \in \Omega.$$

We see that $u_\varepsilon \in C^2(\bar{\Omega})$ and $\lim_{\varepsilon\to 0} u_\varepsilon(x) = u(x)$ in $L^1(\Omega)$ by construction of $U \in \mathcal{V}$. We must now show that this sequence provides the limsup inequality of the Lemma. The chain rule leads to

$$\mathcal{J}_\varepsilon^\gamma(u_\varepsilon) = \frac{1}{\varepsilon}\left(\int_{\Omega\cap N_h} + \int_{\Omega\setminus N_h}\right)\Big(\frac{1}{2}\gamma\Big|U''\left(\frac{\eta}{\varepsilon}\right)|\nabla\eta|^2 + \varepsilon U'\left(\frac{\eta}{\varepsilon}\right)\Delta\eta\Big|^2 + \frac{1}{2}\Big|U'\left(\frac{\eta}{\varepsilon}\right)\Big|^2|\nabla\eta|^2 + \Psi\Big(U\left(\frac{\eta}{\varepsilon}\right)\Big)\Big)\,\mathrm{d}x. \tag{A.4}$$

By construction, $U(z) = 1$ or -1 for $|z| > \delta^\gamma$, and so for ε small enough, the integral in (A.4) over $\Omega \setminus N_h$ is 0 and so this term is done. For the other integral, by construction $\eta = d$ here and so $|\nabla\eta| = 1$, thus $|\Delta\eta| \le C_\eta$.

$$|U''(z) + \varepsilon C_\eta U'(z)|^2 = |U''(z)|^2 + 2\varepsilon C_\eta|U'(z)U''(z)| + C_\eta^2\varepsilon^2|U''(z)|^2,$$

and we apply Young's inequality to the second term with weight $\frac{\nu}{2\gamma}$, where $\nu > 0$:

$$\begin{aligned}|U''(z) + \varepsilon C_\eta U'(z)|^2 &\le |U''(z)|^2 + \frac{\nu}{\gamma}|U'(z)|^2 + \Big(\frac{\gamma(2\varepsilon C_\eta)^2}{\nu} + C_\eta^2\varepsilon^2\Big)|U''(z)|^2\\ &\le (1+\nu)|U''(z)|^2 + \frac{\nu}{\gamma}|U'(z)|^2 + C(\nu)\varepsilon^2,\end{aligned}$$

for constant $C(\nu)$. Notice that after passing to the limit $\varepsilon \to 0$, we could take $\nu > 0$ as arbitrarily small without blowup. We may now bound the integral in (A.4) by

$$\begin{aligned}\mathcal{J}_\varepsilon^\gamma \le\ & \frac{(1+\nu)}{\varepsilon}\int_{\Omega\cap N_h}\Big(\frac{1}{2}\gamma|U''\left(\frac{\eta}{\varepsilon}\right)|^2 + \frac{1}{2}|U'\left(\frac{\eta}{\varepsilon}\right)|^2 + \Psi(U(\omega(x)))\Big)\,\mathrm{d}x + C(\nu)\varepsilon\\ =\ & (1+\nu)\int_{\Omega\cap N_h}\Big(\frac{1}{2}\gamma|U''(\omega(x))|^2 + \frac{1}{2}|U'(\omega(x))|^2 + \Psi(U(\omega(x)))\Big)|\nabla\omega(x)|\,\mathrm{d}x + C(\nu)\varepsilon,\end{aligned}$$

where $\omega(x) = \frac{d(x)}{\varepsilon}$. Note we have used $|\nabla\omega(x)| = \frac{1}{\varepsilon}$, then using the co-area formula [25, Theorem 3.2.12] on $t = \omega(x)$ we obtain

$$\begin{aligned}\mathcal{J}_\varepsilon^\gamma \le\ & (1+\nu)\int_{\mathbb{R}}\int_{\omega^{-1}(t)\cap\Omega\cap N_h}\Big(\frac{1}{2}\gamma|U''(\omega(x))|^2 + \frac{1}{2}|U'(\omega(x))|^2 + \Psi(U(\omega(x)))\Big)|\nabla\omega(x)|\,\mathrm{d}x\\ &+ C(\nu)\varepsilon.\end{aligned}$$

Now we wish to rewrite these integrals. Firstly $\omega^{-1}(t) = \{x \mid \frac{d(x)}{\varepsilon} = t\}$, so

$$\omega^{-1}(t)\cap N_h = \begin{cases}\{x \mid \dfrac{d(x)}{\varepsilon} = t\}, & \text{if } t \le \frac{h}{\varepsilon},\\ \emptyset, & \text{if } t > \frac{h}{\varepsilon}.\end{cases}$$

We may now restate the limits.

$$\begin{aligned}\mathcal{J}_\varepsilon^\gamma \le\ & (1+\nu)\int_{-\frac{h}{\varepsilon}}^{\frac{h}{\varepsilon}}\int_{d(x)=\varepsilon t}\Big(\frac{1}{2}\gamma|U''(t)|^2 + \frac{1}{2}|U'(t)|^2 + \Psi(U(t))\Big)\,\mathrm{d}\mathcal{H}^{d-1}(x)\,\mathrm{d}t\\ &+ C(\nu)\varepsilon\\ =\ & (1+\nu)\int_{-\frac{h}{\varepsilon}}^{\frac{h}{\varepsilon}}\Big(\frac{1}{2}\gamma|U''(t)|^2 + \frac{1}{2}|U'(t)|^2 + \Psi(U(t))\Big)\,\mathrm{d}\mathcal{H}^{d-1}\{x \mid d(x) = \varepsilon t\}\,\mathrm{d}t\end{aligned}$$

$$+ C(\nu)\varepsilon.$$

As $\varepsilon \to 0$, this converges to

$$(1+\nu) j^\gamma(U)\, \mathrm{d}\mathcal{H}^{d-1}\{x \mid d(x) = 0\} = (1+\nu) P^\gamma \mathcal{H}^{d-1}(\partial D \cap \Omega).$$

Therefore we have shown that, in view of (A.3),

$$\limsup_{\varepsilon \to 0} \left(\mathcal{J}_\varepsilon^\gamma(u_\varepsilon)\right) \leq (1+\nu) P^\gamma \mathcal{H}^{d-1}(\partial D \cap \Omega) = (1+\nu)\mathcal{J}_0^\gamma(u),$$

where the choice of ν may be arbitrarily small. Hence the limsup inequality is satisfied □

With both inequalities established, the proof of Theorem 4 is complete.

□