

Schuster, Matthias; Schulz, Volker

Article — Published Version

Second order shape optimization for an interface identification problem constrained by nonlocal models

Computational Optimization and Applications

Suggested Citation: Schuster, Matthias; Schulz, Volker (2025) : Second order shape optimization for an interface identification problem constrained by nonlocal models, Computational Optimization and Applications, ISSN 1573-2894, Springer US, New York, Vol. 91, Iss. 3, pp. 1033-1071, <https://doi.org/10.1007/s10589-025-00678-9>

This Version is available at:

<https://hdl.handle.net/10419/323338>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>



Second order shape optimization for an interface identification problem constrained by nonlocal models

Matthias Schuster¹ · Volker Schulz¹

Received: 22 November 2024 / Accepted: 4 March 2025 / Published online: 1 April 2025
© The Author(s) 2025

Abstract

Since shape optimization methods have been proven useful for identifying interfaces in models governed by partial differential equations, we show how shape optimization techniques can also be applied to an interface identification problem constrained by a nonlocal Dirichlet problem. Here, we focus on deriving the second shape derivative of the corresponding reduced functional and we further investigate a second-order optimization algorithm.

Keywords Shape optimization · Second shape derivative · Nonlocal convection-diffusion

1 Introduction

Nonlocal operators are typically integral operators and therefore allow interactions between two distinct points in space. Moreover, the solution of a nonlocal problem in general needs to satisfy less regularity requirements compared to the solution of a system of partial differential equations. Thus, some physical phenomena like fracture propagation in continuum mechanics [1, 2] or anomalous diffusion effects [3, 4] can be modeled more accurately by nonlocal equations compared to their ‘classic’ description by partial differential equations. Moreover, nonlocal models have been successfully applied to image denoising [5, 6], neural networks [7, 8] or stochastic processes [9, 10], just to name a few.

Shape optimization techniques are mostly developed in the context of PDE-constrained optimization problems and have been investigated in areas like aerodynamics [11, 12], acoustics [13–15], fluid dynamics [16, 17] and many more. In [18], we have already formulated the first shape derivative of an interface identification problem that is governed by a nonlocal Dirichlet problem and we additionally

✉ Volker Schulz
volker.schulz@uni-trier.de

Matthias Schuster
schusterm@uni-trier.de

¹ Department 4, Trier University, Universitätsring 15, 54286 Trier, Rhineland-Palatinate, Germany

investigated a first-order shape optimization algorithm. In this work, we continue by deriving a second shape derivative, which is then used in a Newton-like second-order shape optimization approach.

The first-order shape derivative method of [18] was also employed for interface identification constrained by an energy-based Local-to-Nonlocal coupling in [19]. Moreover, in the context of Peridynamics, a machine-learning approach to detect an interface was tested in [20]. Here, Nader et. al. made use of finite differences to solve a one-dimensional variable horizon Local-to-Nonlocal coupling, and a convolutional neural network was trained to assign domain affiliations to nodes.

In the next section, we start by introducing nonlocal Dirichlet problems and shortly discuss the well-posedness of the corresponding weak formulation. In Sect. 3, we present the necessary basics of shape optimization. Then, we define an interface identification problem that is constrained by a nonlocal Dirichlet problem in Sect. 4 followed by a description of the averaged adjoint method (AAM) in Sect. 5, which we will use to compute the second shape derivative of the reduced functional corresponding to the constrained interface identification problem. After that, we show in Sect. 6, that, under natural assumptions, the prerequisites of AAM are satisfied. Lastly, we introduce in Sect. 7 a second-order optimization approach that is partly a consequence of AAM and which is employed in Sect. 8 in two numerical experiments.

2 Nonlocal terminology and framework

Let $\Omega \subset \mathbb{R}^d$ with $d \in \mathbb{N}$ be a nonempty, open, bounded and simply connected domain. In this work, we consider the *nonlocal convection-diffusion operator*

$$-\mathcal{L}u(\mathbf{x}) := \int_{\mathbb{R}^d} u(\mathbf{x})\gamma(\mathbf{x}, \mathbf{y}) - u(\mathbf{y})\gamma(\mathbf{y}, \mathbf{x}) \, d\mathbf{y}. \quad (1)$$

Here, the function $\gamma : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, \infty)$ is the so-called *kernel*, which plays a vital role in the nonlocal theory. Moreover, problems involving the nonlocal operator (1) have been extensively studied in the literature (see, e.g., [21, 22]). Further, we define a *nonlocal steady-state nonhomogeneous Dirichlet problem* as

$$\begin{aligned} &\text{Find a function } u \text{ s.t.} \\ &-\mathcal{L}u = f \quad \text{on } \Omega \\ &u = g \quad \text{on } \mathcal{I}, \end{aligned} \quad (2)$$

where we set the so-called *nonlocal boundary* (or *nonlocal interaction domain*) \mathcal{I} as

$$\mathcal{I} := \{\mathbf{y} \in \mathbb{R}^d \setminus \Omega : \int_{\Omega} \gamma(\mathbf{x}, \mathbf{y}) + \gamma(\mathbf{y}, \mathbf{x}) \, d\mathbf{x} > 0\},$$

which contains all points outside of the domain Ω that interact with at least one point inside of Ω through the kernel γ . In this section, we will recall the basic theory on problems of type (2). We start by deriving a weak formulation of problem (2). Given

appropriate spaces $V(\Omega \cup \mathcal{I})$ and $V_c(\Omega \cup \mathcal{I})$, which are defined later in this section, we multiply the first equation of (2), where we assume $u \in V(\Omega \cup \mathcal{I})$, with a test vector $v \in V_c(\Omega \cup \mathcal{I})$ and then integrate over Ω , which yields

$$\int_{\Omega} \int_{\Omega \cup \mathcal{I}} v(\mathbf{x}) (u(\mathbf{x})\gamma(\mathbf{x}, \mathbf{y}) - u(\mathbf{y})\gamma(\mathbf{y}, \mathbf{x})) \, d\mathbf{y}d\mathbf{x} = \int_{\Omega} f(\mathbf{x})v(\mathbf{x}) \, d\mathbf{x}.$$

Here, we can rewrite the left-hand side by using $v = 0$ on \mathcal{I} and by applying Fubini's theorem as follows

$$\begin{aligned} & \int_{\Omega} \int_{\Omega \cup \mathcal{I}} v(\mathbf{x}) (u(\mathbf{x})\gamma(\mathbf{x}, \mathbf{y}) - u(\mathbf{y})\gamma(\mathbf{y}, \mathbf{x})) \, d\mathbf{y}d\mathbf{x} \\ &= \frac{1}{2} \int_{\Omega \cup \mathcal{I}} \int_{\Omega \cup \mathcal{I}} (v(\mathbf{x}) - v(\mathbf{y})) (u(\mathbf{x})\gamma(\mathbf{x}, \mathbf{y}) - u(\mathbf{y})\gamma(\mathbf{y}, \mathbf{x})) \, d\mathbf{y}d\mathbf{x}. \end{aligned}$$

This reformulation is numerically more stable for singular symmetric kernels, which will be introduced later in this section. Now, before we present the definition of a weak solution to (2), we define the bilinear form

$$A(u, v) := \frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y})) (u(\mathbf{x})\gamma(\mathbf{x}, \mathbf{y}) - u(\mathbf{y})\gamma(\mathbf{y}, \mathbf{x})) \, d\mathbf{y}d\mathbf{x}$$

and the linear functional

$$F(v) := \int_{\Omega} f(\mathbf{x})v(\mathbf{x}) \, d\mathbf{x}.$$

Then, we set the (semi-)norm $||| \cdot |||$ as

$$|||u||| := \sqrt{A(u, u)}.$$

Additionally, we define the nonlocal energy space $V(\Omega \cup \mathcal{I})$ and the nonlocal volume-constrained energy space $V_c(\Omega \cup \mathcal{I})$ in the following way:

$$\begin{aligned} V(\Omega \cup \mathcal{I}) &:= \{u \in L^2(\Omega \cup \mathcal{I}) : \|u\|_{V(\Omega \cup \mathcal{I})} := |||u||| + \|u\|_{L^2(\Omega \cup \mathcal{I})} < \infty\} \text{ and} \\ V_c(\Omega \cup \mathcal{I}) &:= \{u \in V(\Omega \cup \mathcal{I}) : u = 0 \text{ on } \mathcal{I}\}. \end{aligned}$$

Definition 1 (Weak Solution of a Nonlocal Dirichlet Problem) Given $f \in L^2(\Omega)$ and $g \in V(\Omega \cup \mathcal{I})$, if $\tilde{u} \in V_c(\Omega \cup \mathcal{I})$ satisfies

$$A(\tilde{u}, v) = F(v) - A(g, v) \text{ for all } v \in V_c(\Omega \cup \mathcal{I}),$$

then we call $u := \tilde{u} + g \in V(\Omega \cup \mathcal{I})$ weak solution of the nonlocal Dirichlet problem (2).

For $\delta > 0$ we set $B_\delta(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^d : \|\mathbf{x} - \mathbf{y}\|_2 < \delta\}$. Then, in this work, every kernel γ is assumed to satisfy:

(K1) There exists a constant $\delta > 0$ and a function $\varphi : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, \infty)$ such that

$$\gamma(\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{x}, \mathbf{y}) \chi_{B_\delta(\mathbf{x})}(\mathbf{y}).$$

(K2) There exist constants $\epsilon \in (0, \delta)$ and $\underline{\gamma} > 0$ with

$$\gamma(\mathbf{x}, \mathbf{y}) \geq \underline{\gamma} \text{ for all } \mathbf{x} \in \Omega \text{ and } \mathbf{y} \in B_\epsilon(\mathbf{x}).$$

Further, we consider two types of kernel classes:

- **Integrable Kernels** Set $\gamma^s(\mathbf{x}, \mathbf{y}) := \frac{1}{2}(\gamma(\mathbf{x}, \mathbf{y}) + \gamma(\mathbf{y}, \mathbf{x}))$ and $\gamma^a(\mathbf{x}, \mathbf{y}) := \frac{1}{2}(\gamma(\mathbf{x}, \mathbf{y}) - \gamma(\mathbf{y}, \mathbf{x}))$. Then, $\gamma = \gamma^s + \gamma^a$ holds and we call γ *integrable*, if the following requirements are satisfied:

(1) There exist constants $\underline{\gamma}^s, \bar{\gamma}^s > 0$ such that

$$\underline{\gamma}^s = \inf_{\mathbf{x} \in \Omega} \int_{\Omega \cup \mathcal{I}} \gamma^s(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} < \infty \text{ and } \sup_{\mathbf{x} \in \Omega} \int_{\Omega \cup \mathcal{I}} (\gamma^s(\mathbf{x}, \mathbf{y}))^2 \, d\mathbf{y} = (\bar{\gamma}^s)^2.$$

(2) There exist constants $\underline{\gamma}^a \in \mathbb{R}$ and $\bar{\gamma}^a \geq 0$ with

$$\underline{\gamma}^a = \inf_{\mathbf{x} \in \Omega} \int_{\Omega \cup \mathcal{I}} \gamma^a(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} \text{ and } \sup_{\mathbf{x} \in \Omega} \int_{\Omega \cup \mathcal{I}} |\gamma^a(\mathbf{x}, \mathbf{y})| \, d\mathbf{y} = \bar{\gamma}^a.$$

(3) The lower boundaries $\underline{\gamma}^s$ and $\underline{\gamma}^a$ satisfy

$$\underline{\gamma}^s + \underline{\gamma}^a > 0.$$

- **Singular Symmetric Kernels**

There exist positive constants $0 < \gamma_* \leq \gamma^* < \infty$ and $s \in (0, 1)$ such that

$$\gamma_* \leq \gamma(\mathbf{x}, \mathbf{y}) \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} \leq \gamma^* \text{ for all } \mathbf{x} \in \Omega \text{ and } \mathbf{y} \in B_\delta(\mathbf{x}). \quad (3)$$

Furthermore, the kernel γ is symmetric.

In [21, 22], it is shown that $\|\cdot\| = \sqrt{A(\cdot, \cdot)}$ is indeed a norm on $V_c(\Omega \cup \mathcal{I})$ and that there is a unique weak solution to the Dirichlet problem (2) if the corresponding kernel satisfies (K1) and (K2) and additionally fits into one of those two categories.

Remark 1 In the volume-constrained space $V_c(\Omega \cup \mathcal{I})$ the norms $\|\cdot\|_{L^2(\Omega \cup \mathcal{I})}$ and $\|\cdot\|$ are equivalent if we have chosen an integrable kernel. Thus, it holds that

$$u \in (V_c(\Omega \cup \mathcal{I}), \|\cdot\|) \Leftrightarrow u \in (L_c^2(\Omega \cup \mathcal{I}), \|\cdot\|_{L^2(\Omega \cup \mathcal{I})}),$$

where $L_c^2(\Omega \cup \mathcal{I}) := \{u \in L^2(\Omega \cup \mathcal{I}) : u = 0 \text{ on } \mathcal{I}\}$. On the other hand, if γ meets the requirements of a singular symmetric kernel, we get norm equivalence between $||| \cdot |||$ and $|\cdot|_{H^s(\Omega \cup \mathcal{I})}$ on $V_c(\Omega \cup \mathcal{I})$ and

$$u \in (V_c(\Omega \cup \mathcal{I}), ||| \cdot |||) \Leftrightarrow u \in (H_c^s(\Omega \cup \mathcal{I}), |\cdot|_{H^s(\Omega \cup \mathcal{I})}),$$

where the norm $|\cdot|_{H^s(\Omega \cup \mathcal{I})}$ and the volume-constrained fractional Sobolev space $H_c^s(\Omega \cup \mathcal{I})$ are defined as

$$H_c^s(\Omega \cup \mathcal{I}) := \{u \in L_c^2(\Omega \cup \mathcal{I}) : |u|_{H^s(\Omega \cup \mathcal{I})} < \infty\} \text{ with}$$

$$|u|_{H^s(\Omega \cup \mathcal{I})}^2 := \iint_{(\Omega \cup \mathcal{I})^2} \frac{(u(\mathbf{x}) - u(\mathbf{y}))^2}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} d\mathbf{y} d\mathbf{x}.$$

FOR the proof of these assertions, we refer again to [21, 22].

3 Basics of shape optimization

Here, we present basic notions of shape optimization and we refer to [23–27] for a thorough introduction. In this work, only shapes that are elements of the following set are considered:

Definition 2 (Set of Admissible Shapes) Given a nonempty, Lipschitz open domain $\widehat{\Omega} \subset \Omega$, then we set $\Gamma_0 := \partial \widehat{\Omega}$ and define the corresponding set of admissible shapes $\mathcal{A}(\Gamma_0)$ as

$$\mathcal{A}(\Gamma_0) := \{\mathbf{T}(\Gamma_0) \mid \mathbf{T} : \overline{\Omega} \rightarrow \overline{\Omega} \text{ invertible, } \mathbf{T} \text{ and } \mathbf{T}^{-1} \text{ are Lipschitz continuous}\}.$$

Next, the notion of a shape optimization problem can be introduced:

Definition 3 (Shape Optimization Problem) Let \mathcal{A} be a subset of $\mathcal{A}(\Gamma_0)$ as formulated in Definition 2 and let $J : \mathcal{A} \rightarrow \mathbb{R}$ be a so-called shape functional, then we refer to

$$\min_{\Gamma \in \mathcal{A}} J(\Gamma)$$

as a shape optimization problem.

To define shape derivatives, we make use of a family of mappings $\{\mathbf{F}_t\}_{t \in [0, T]}$, where $\mathbf{F}_t : \overline{\Omega} \rightarrow \mathbb{R}^d$, $\mathbf{F}_0(\mathbf{x}) = \mathbf{x}$ for $\mathbf{x} \in \overline{\Omega}$ and $T > 0$ is supposed to be sufficiently small. In this work, we apply the perturbation of identity:

Definition 4 (Perturbation of Identity) Given a vector field $\mathbf{V} \in C_0^k(\Omega, \mathbb{R}^d)$, $k \in \mathbb{N}$ and $t \geq 0$, the perturbation of identity is defined as

$$\mathbf{F}_t^{\mathbf{V}} : \overline{\Omega} \rightarrow \mathbb{R}^d, \quad \mathbf{F}_t^{\mathbf{V}}(\mathbf{x}) := \mathbf{x} + t\mathbf{V}(\mathbf{x}),$$

where $C_0^k(\Omega, \mathbb{R}^d) := \{\mathbf{V} \in C^k(\Omega, \mathbb{R}^d) : \mathbf{V} \text{ vanishes at the boundary}\}.$

Here, $\mathbf{V} \in C^k(\Omega, \mathbb{R}^d)$ vanishes at the boundary, if the set $\{\mathbf{x} \in \Omega : |\partial^\alpha \mathbf{V}(\mathbf{x})| \geq \epsilon\}$ is compact for all $\epsilon > 0$ and partial derivatives $\partial^\alpha := \partial_{\mathbf{x}_1}^{\alpha_1} \partial_{\mathbf{x}_2}^{\alpha_2} \dots \partial_{\mathbf{x}_d}^{\alpha_d}$ w.r.t. multi-indices $\alpha \in \mathbb{N}_0^d$ with $0 \leq |\alpha| \leq k$. If T is chosen small enough such that $\|t\mathbf{V}\|_{W^{1,\infty}(\Omega, \mathbb{R}^d)} < 1$ for all $t \in [0, T]$, then the function $\mathbf{F}_t^{\mathbf{V}} : \overline{\Omega} \rightarrow \overline{\Omega}$ is a bijective function where $\mathbf{F}_t^{\mathbf{V}}$ as well as the corresponding inverse $(\mathbf{F}_t^{\mathbf{V}})^{-1}$ are both Lipschitz continuous (see [25, 28, 29]). Now, we continue by formulating the definition of the first and second shape derivative:

Definition 5 (First Shape Derivative) The Eulerian derivative of a shape functional J at $\Gamma \in \mathcal{A}(\Gamma_0)$ in the direction $\mathbf{V} \in C_0^k(\Omega, \mathbb{R}^d)$, where $k \in \mathbb{N}$, is defined as the directional derivative

$$DJ(\Gamma)[\mathbf{V}] := \lim_{t \searrow 0} \frac{J(\mathbf{F}_t^{\mathbf{V}}(\Gamma)) - J(\Gamma)}{t}.$$

If $DJ(\Gamma)[\mathbf{V}]$ exists for all vector fields $\mathbf{V} \in C_0^k(\Omega, \mathbb{R}^d)$ and if the function

$$DJ(\Gamma) : C_0^k(\Omega, \mathbb{R}^d) \rightarrow \mathbb{R}, \quad \mathbf{V} \mapsto DJ(\Gamma)[\mathbf{V}]$$

is linear and continuous, then we refer to $DJ(\Gamma)[\mathbf{V}]$ as the shape derivative at Γ in direction \mathbf{V} .

Definition 6 (Second Shape Derivative) Let $k \in \mathbb{N}$. If for all $\mathbf{V}, \mathbf{W} \in C_0^k(\Omega, \mathbb{R}^d)$ the semi-derivative

$$D^2J(\Gamma)[\mathbf{V}, \mathbf{W}] := \lim_{t \searrow 0} \frac{DJ(\mathbf{F}_t^{\mathbf{W}}(\Gamma))[\mathbf{V}] - DJ(\Gamma)[\mathbf{V}]}{t}$$

exists and the function

$$D^2J(\Gamma) : C_0^k(\Omega, \mathbb{R}^d) \times C_0^k(\Omega, \mathbb{R}^d) \rightarrow \mathbb{R}, \quad (\mathbf{V}, \mathbf{W}) \mapsto D^2J(\Gamma)[\mathbf{V}, \mathbf{W}]$$

is bilinear and continuous, then $D^2J(\Gamma)[\mathbf{V}, \mathbf{W}]$ is called the second shape derivative at Γ in the direction (\mathbf{V}, \mathbf{W}) .

The second shape derivative can also be characterized as

$$D^2J(\Gamma)[\mathbf{V}, \mathbf{W}] = \left. \frac{d}{dt} \right|_{t_2 \searrow 0} \left. \frac{d}{dt} \right|_{t_1 \searrow 0} J\left(\mathbf{F}_{t_1}^{\mathbf{V}}(\mathbf{F}_{t_2}^{\mathbf{W}}(\Gamma))\right),$$

which indicates that the second shape derivative is computed by differentiating a functional where the shape Γ is first perturbed along the vector field \mathbf{W} and afterwards along \mathbf{V} as described in Definition 4.

Remark 2 In the literature, the second shape derivative $D^2J(\Gamma)$ is sometimes called the shape Hessian at Γ (see, e.g., [30–32]).

In this work, as mentioned before, we only employ the perturbation of identity, which is invertible if $t > 0$ is small enough. As shown in [31], interpreting shape optimization as optimizing on a subset of $\{\mathbf{T} | \mathbf{T} : \bar{\Omega} \rightarrow \bar{\Omega}\}$ yields the linear second shape derivative

$$J''(\Gamma)[\mathbf{V}, \mathbf{W}] := \left. \frac{d}{dt} \right|_{t_2 \searrow 0} \left. \frac{d}{dt} \right|_{t_1 \searrow 0} J((\mathbf{Id} + t_2 \mathbf{W} + t_1 \mathbf{V})(\Gamma)), \quad (4)$$

which is symmetric, provides the basis of a Taylor series and thus justifies convergence properties of shape Newton methods. In [30, 33, 34] and [23, Section 6.5], it is proven that the classical shape derivative $D^2 J(\Gamma)[\mathbf{V}, \mathbf{W}]$ can be expressed as

$$D^2 J(\Gamma)[\mathbf{V}, \mathbf{W}] = J''(\Gamma)[\mathbf{V}, \mathbf{W}] + DJ(\Gamma)[(D\mathbf{V})\mathbf{W}]. \quad (5)$$

Here, the second term $DJ(\Gamma)[(D\mathbf{V})\mathbf{W}]$ is only symmetric if $(D\mathbf{V})\mathbf{W} = (D\mathbf{W})\mathbf{V}$ holds. In Sect. 6, we first derive the classical shape derivative $D^2 J(\Gamma)[\mathbf{V}, \mathbf{W}]$ by making use of the so-called averaged adjoint method. By omitting the possibly nonsymmetric part $DJ(\Gamma)[(D\mathbf{V})\mathbf{W}]$, we will then obtain the linear second shape derivative, which we utilize in a Newton-like algorithm to solve the constrained interface identification problem (6) introduced in the next section.

Remark 3 In this work, the nonlocal boundary \mathcal{I} will not be deformed since the interface Γ is supposed to be a subset of Ω . Thus, we extend every vector field $\mathbf{V} \in C_0^k(\Omega, \mathbb{R}^d)$ by zero to a function defined on the domain $\Omega \cup \mathcal{I}$, i.e., we only consider vector fields that are elements of the set

$$C_0^k(\Omega \cup \mathcal{I}, \mathbb{R}^d) := \{\mathbf{V} \in C^k(\Omega \cup \mathcal{I}, \mathbb{R}^d) : \mathbf{V} = 0 \text{ on } \mathcal{I} \text{ and } \mathbf{V}|_{\bar{\Omega}} \in C_0^k(\Omega, \mathbb{R}^d)\}.$$

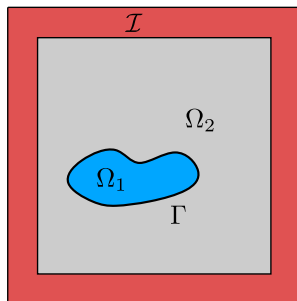
As a consequence, the associated perturbation of identity $\mathbf{F}_t^{\mathbf{V}} : \Omega \cup \mathcal{I} \rightarrow \Omega \cup \mathcal{I}$ is again a diffeomorphism since $\mathbf{F}_t^{\mathbf{V}} = \mathbf{Id}$ on \mathcal{I} and $\mathbf{F}_t^{\mathbf{V}}|_{\Omega} : \Omega \rightarrow \Omega$ is a diffeomorphism. Furthermore, $\mathbf{V} \in C_0^2(\Omega \cup \mathcal{I}, \mathbb{R}^d)$ is sufficient for the calculations in this work.

4 Interface identification constrained by nonlocal models

PDE-constrained interface identification is a well-researched shape optimization problem (see, e.g., [23, 27]) and in the following, we will investigate interface identification governed by nonlocal models.

Here, as illustrated in Fig. 1, we assume Γ to be the boundary of a nonempty, Lipschitz open domain $\Omega_1 \subset \Omega$. Consequently, Ω is decomposed by Γ into two open sets Ω_1 and $\Omega_2 := \Omega \setminus (\Omega_1 \cup \Gamma)$. Further, we suppose that Ω_1 and Ω_2 are connected. Then, we call Γ *interface* and we denote by $\Omega(\Gamma)$ that Ω is decomposed according to Γ , i.e., $\Omega = (\Omega_1 \dot{\cup} \Gamma \dot{\cup} \Omega_2)^\circ$.

Fig. 1 The interface Γ decomposes the domain Ω into the blue subdomain Ω_1 and the gray subdomain Ω_2 . Further, note that the nonlocal boundary \mathcal{I} is depicted in red.



Now, given $\bar{u} \in H^2(\Omega)$, we define *interface identification governed by a nonlocal Dirichlet problem* as follows:

$$\begin{aligned} & \min_{u \in V_c(\Omega(\Gamma) \cup \mathcal{I}), \Gamma \in \mathcal{A}(\Gamma_0)} J(u, \Gamma) \\ \text{s.t. } & -\mathcal{L}_\Gamma u = f_\Gamma \quad \text{on } \Omega, \\ & u = 0 \quad \text{on } \mathcal{I}, \end{aligned} \quad (6)$$

where

$$J(u, \Gamma) := j(u) + \alpha j_{\text{per}}(\Gamma) \quad \text{with } j(u) := \int_\Omega \frac{1}{2} (u - \bar{u})^2 \, d\mathbf{x} \quad \text{and } j_{\text{per}}(\Gamma) := \int_\Gamma 1 \, ds, \\ -\mathcal{L}_\Gamma u(\mathbf{x}) := \int_{\mathbb{R}^d} u(\mathbf{y}) \gamma_\Gamma(\mathbf{x}, \mathbf{y}) - u(\mathbf{y}) \gamma_\Gamma(\mathbf{y}, \mathbf{x}) \, d\mathbf{y} \quad \text{with kernel} \quad (7)$$

$$\gamma_\Gamma := \sum_{i,j=1,2} \gamma_{ij} \chi_{\Omega_i \times \Omega_j} + \sum_{i=1,2} \gamma_{i\mathcal{I}} \chi_{\Omega_i \times \mathcal{I}}. \quad (8)$$

Here, $\gamma_{ij}, \gamma_{i\mathcal{I}} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, \infty)$ are also called subkernels. For ease of presentation, we only consider homogeneous Dirichlet problems. Therefore, we do not need any assumptions on the behavior of γ_Γ on $\mathcal{I} \times \Omega$ and we can choose w.l.o.g. $\gamma_\Gamma = 0$ on $(\mathbb{R}^d \times \mathbb{R}^d) \setminus (\Omega \times (\Omega \cup \mathcal{I}))$ in (8). However, the nonhomogeneous case can be handled analogously. As we can see in (7) and (8), the operator $-\mathcal{L}_\Gamma$ is dependent on the interface since the choice of kernel $\gamma_\Gamma(\mathbf{x}, \mathbf{y})$ varies depending on the location of \mathbf{x} and \mathbf{y} . Furthermore, we suppose that the forcing term can be expressed as $f_\Gamma = f_1 \chi_{\Omega_1} + f_2 \chi_{\Omega_2}$, where $f_1, f_2 \in H^2(\Omega)$. Then, we denote the corresponding variational formulation of the Dirichlet problem (6) as

$$A_\Gamma(u, v) = F_\Gamma(v) \quad \text{for all } v \in V_c(\Omega \cup \mathcal{I}), \quad (9)$$

where the subscripts highlight that the kernel as well as the forcing term are dependent on the decomposition imposed by the interface Γ . The objective functional consists of two terms: The first one, which we refer to as j , is a tracking-type functional since we aim to find the interface Γ such that the solution to (9) approximates the given data \bar{u} as well as possible. The second integral of the objective function, which we denote by j_{per} , is called *perimeter regularization*, which is often used to avoid ill-posedness of

the problem (see [35]). Previously, as mentioned above, first-order shape optimization algorithms for the interface identification problem (6) have already been investigated in [18].

Remark 4 The first and second shape derivative of the perimeter regularization j_{per} can be formulated as

$$Dj_{per}(\Gamma)[\mathbf{V}] = \int_{\Gamma} \operatorname{div}_{\Gamma} \mathbf{V} \, ds = \int_{\Gamma} \operatorname{div} \mathbf{V} - \mathbf{n}^{\top} D\mathbf{V}\mathbf{n} \, ds \quad \text{and} \\ D^2j_{per}(\Gamma)[\mathbf{V}, \mathbf{W}] = \int_{\Gamma} \operatorname{div}_{\Gamma} \mathbf{V} \operatorname{div}_{\Gamma} \mathbf{W} - \operatorname{tr}(D_{\Gamma}\mathbf{V}D_{\Gamma}\mathbf{W}) + \langle (D_{\Gamma}\mathbf{V})^{\top} \mathbf{n}, (D_{\Gamma}\mathbf{W})^{\top} \mathbf{n} \rangle \, ds,$$

where $\operatorname{div}_{\Gamma} \mathbf{V} := \operatorname{div} \mathbf{V} - \mathbf{n}^{\top} D\mathbf{V}\mathbf{n}$ is the tangential divergence and $D_{\Gamma}\mathbf{V} := D\mathbf{V} - D\mathbf{V}\mathbf{n}\mathbf{n}^{\top}$ the tangential Jacobian. The formula for $Dj_{per}(\Gamma)[\mathbf{V}]$ follows directly from [11, Lemma 3.3.4]. The derivation of $D^2j_{per}(\Gamma)[\mathbf{V}, \mathbf{W}]$ can be found in [36, Section 3.1]. For ease of presentation, we exclude $j_{per}(\Gamma)$, $Dj_{per}(\Gamma)[\mathbf{V}]$ and $D^2j_{per}(\Gamma)[\mathbf{V}, \mathbf{W}]$ in the calculations of Sect. 6.

5 Second shape derivative via the averaged adjoint method

Since for every suitable interface Γ the nonlocal Dirichlet problem has a unique solution, the interface identification problem constrained by nonlocal models (6) can be written as a reduced optimization problem

$$\min_{\Gamma \in \mathcal{A}(\Gamma_0)} J^{red}(\Gamma) \quad (10)$$

where $J^{red}(\Gamma) := J(u(\Gamma), \Gamma)$ and $u(\Gamma)$ solves the variational formulation (9). In [18] it was shown how the first shape derivative of (10) can be computed via the averaged adjoint method (AAM) of [26, 37, 38]. Therefore, we refer for the first derivative to [18] and continue by introducing the *averaged adjoint method for a system of state equations* of [26], which we apply to develop a formula for the second shape derivative of (10). The steps to develop the second shape derivative are quite similar to the steps to compute the first shape derivative. Note that there also exist alternative approaches to get a second shape derivative associated with problem (6) like, e.g., a material derivative method [31, 36, 39, 40].

As proven in [18], we can express the first shape derivative of the reduced functional at the shape $\Gamma_t^{\mathbf{W}} := \mathbf{F}_t^{\mathbf{W}}(\Gamma)$ in the direction $\mathbf{V} \in C_0^2(\Omega \cup \mathcal{I}, \mathbb{R}^d)$ as

$$DJ^{red}(\Gamma_t^{\mathbf{W}})[\mathbf{V}] = \mathfrak{J}_{\Gamma^{\mathbf{W}}}(u(\Gamma_t^{\mathbf{W}})) - \mathfrak{F}_{\Gamma^{\mathbf{W}}}(v(\Gamma_t^{\mathbf{W}})) + \mathfrak{A}_{\Gamma^{\mathbf{W}}}(u(\Gamma_t^{\mathbf{W}}), v(\Gamma_t^{\mathbf{W}})),$$

where the three components are the derivatives of the objective functional, the forcing term and the bilinear form. Formulas for these derivatives are derived from (26) and (27) of [18] with respect to $\Gamma_t^{\mathbf{W}}$ and can be found below. Here, the functions $u(\Gamma_t^{\mathbf{W}})$

and $v(\Gamma_t^{\mathbf{W}})$ satisfy the equations

$$\begin{aligned} A_{\Gamma_t^{\mathbf{W}}}(u(\Gamma_t^{\mathbf{W}}), \psi) &= F_{\Gamma_t^{\mathbf{W}}}(\psi) && \text{for all } \psi \in V_c(\mathbf{F}_t^{\mathbf{W}}(\Omega) \cup \mathcal{I}) \text{ and} \\ A_{\Gamma_t^{\mathbf{W}}}(\varphi, v(\Gamma_t^{\mathbf{W}})) &= \tilde{F}_{\Gamma_t^{\mathbf{W}}}(u(\Gamma_t^{\mathbf{W}}), \varphi) && \text{for all } \varphi \in V_c(\mathbf{F}_t^{\mathbf{W}}(\Omega) \cup \mathcal{I}), \end{aligned}$$

where $\tilde{F}_{\Gamma_t^{\mathbf{W}}}(u, \varphi) := - \int_{\mathbf{F}_t^{\mathbf{W}}(\Omega)} (u - \bar{u}) \varphi \, d\mathbf{x}$. Notice that $v(\Gamma_t^{\mathbf{W}})$ is the adjoint of the first shape derivative at $\Gamma_t^{\mathbf{W}}$ (see [18]). Additionally, we set $\mathbf{V}(\mathbf{x}, \mathbf{y}) := (\mathbf{V}(\mathbf{x}), \mathbf{V}(\mathbf{y}))^\top$ as well as

$$\Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}}^1(\mathbf{x}, \mathbf{y}) := \nabla \gamma_{\Gamma_t^{\mathbf{W}}}(\mathbf{x}, \mathbf{y})^\top \mathbf{V}(\mathbf{x}, \mathbf{y}) \text{ and } \Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}}^2(\mathbf{x}, \mathbf{y}) := \gamma_{\Gamma_t^{\mathbf{W}}}(\mathbf{x}, \mathbf{y})(\operatorname{div} \mathbf{V}(\mathbf{x}) + \operatorname{div} \mathbf{V}(\mathbf{y})), \quad (11)$$

where $\nabla \gamma_{\Gamma_t^{\mathbf{W}}} = \left(\nabla_{\mathbf{x}} \gamma_{\Gamma_t^{\mathbf{W}}}, \nabla_{\mathbf{y}} \gamma_{\Gamma_t^{\mathbf{W}}} \right)^\top$. Therefore, the three derivatives of objective, forcing term and bilinear form are formulated in the following way:

$$\begin{aligned} \mathfrak{J}_{\Gamma_t^{\mathbf{W}}}(u) &:= \int_{\mathbf{F}_t^{\mathbf{W}}(\Omega)} -(u(\mathbf{x}) - \bar{u}(\mathbf{x})) \nabla \bar{u}(\mathbf{x})^\top \mathbf{V}(\mathbf{x}) + \frac{1}{2} (u(\mathbf{x}) - \bar{u}(\mathbf{x}))^2 \operatorname{div} \mathbf{V}(\mathbf{x}) \, d\mathbf{x}, \\ \mathfrak{F}_{\Gamma_t^{\mathbf{W}}}(v) &:= \int_{\mathbf{F}_t^{\mathbf{W}}(\Omega)} v(\mathbf{x}) \nabla f(\mathbf{x})^\top \mathbf{V}(\mathbf{x}) + f(\mathbf{x}) v(\mathbf{x}) \operatorname{div} \mathbf{V}(\mathbf{x}) \, d\mathbf{x} \text{ and} \\ \mathfrak{A}_{\Gamma_t^{\mathbf{W}}}(u, v) &:= \frac{1}{2} \iint_{(\mathbf{F}_t^{\mathbf{W}}(\Omega \cup \mathcal{I}))^2} (v(\mathbf{x}) - v(\mathbf{y})) \left(u(\mathbf{x}) \Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}}^1(\mathbf{x}, \mathbf{y}) - u(\mathbf{y}) \Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}}^1(\mathbf{y}, \mathbf{x}) \right) \, dy d\mathbf{x} \\ &\quad + \frac{1}{2} \iint_{(\mathbf{F}_t^{\mathbf{W}}(\Omega \cup \mathcal{I}))^2} (v(\mathbf{x}) - v(\mathbf{y})) \left(u(\mathbf{x}) \Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}}^2(\mathbf{x}, \mathbf{y}) - u(\mathbf{y}) \Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}}^2(\mathbf{y}, \mathbf{x}) \right) \, dy d\mathbf{x}. \end{aligned}$$

Then, we set the corresponding Lagrangian as

$$\begin{aligned} \mathcal{L}(t, u, v, \psi, \varphi) &:= \mathfrak{J}_{\Gamma_t^{\mathbf{W}}}(u) - \mathfrak{F}_{\Gamma_t^{\mathbf{W}}}(v) + \mathfrak{A}_{\Gamma_t^{\mathbf{W}}}(u, v) + A_{\Gamma_t^{\mathbf{W}}}(u, \psi) - F_{\Gamma_t^{\mathbf{W}}}(\psi) \\ &\quad + A_{\Gamma_t^{\mathbf{W}}}(\varphi, v) - \tilde{F}_{\Gamma_t^{\mathbf{W}}}(u, \varphi). \end{aligned}$$

Thus, for the first shape derivative of the reduced functional holds

$$DJ^{red}(\Gamma_t^{\mathbf{W}})[\mathbf{V}] = \mathcal{L}(t, u(\Gamma_t^{\mathbf{W}}), v(\Gamma_t^{\mathbf{W}}), \psi, \varphi) \text{ for any } \psi, \varphi \in V_c(\Omega(\Gamma_t^{\mathbf{W}}) \cup \mathcal{I}).$$

In order to avoid computing derivatives of u , v , ψ or φ to derive the second shape derivative, we use the perturbation of identity, which is a diffeomorphism, as a so-called *pull-back function*. Due to this characteristic of the perturbation of identity, we can find for any function $u \in V_c(\Omega(\Gamma_t^{\mathbf{W}}) \cup \mathcal{I})$ a unique $\tilde{u} \in V_c(\Omega \cup \mathcal{I})$ with $u = \tilde{u} \circ (\mathbf{F}_t^{\mathbf{W}})^{-1}$. Consequently, we can define an alternative Lagrangian as

$$\begin{aligned} \mathfrak{G} &: [0, T] \times V_c(\Omega \cup \mathcal{I})^4 \rightarrow \mathbb{R}, \\ \mathfrak{G}(t, u, v, \psi, \varphi) &:= \mathcal{L}(t, u \circ (\mathbf{F}_t^{\mathbf{W}})^{-1}, v \circ (\mathbf{F}_t^{\mathbf{W}})^{-1}, \psi \circ (\mathbf{F}_t^{\mathbf{W}})^{-1}, \varphi \circ (\mathbf{F}_t^{\mathbf{W}})^{-1}). \end{aligned}$$

Another advantage of this function space parametrization is that the function spaces for u , v , ψ and φ do not depend on t anymore. Further, we set

$$\begin{aligned}\mathfrak{J} &: [0, T] \times V_c(\Omega \cup \mathcal{I}) \rightarrow \mathbb{R}, \quad \mathfrak{J}(t, u) := \mathfrak{J}_{\Gamma^W}(u \circ (\mathbf{F}_t^W)^{-1}), \\ \mathfrak{F} &: [0, T] \times V_c(\Omega \cup \mathcal{I}) \rightarrow \mathbb{R}, \quad \mathfrak{F}(t, u) := \mathfrak{F}_{\Gamma^W}(v \circ (\mathbf{F}_t^W)^{-1}), \\ \mathfrak{A} &: [0, T] \times V_c(\Omega \cup \mathcal{I})^2 \rightarrow \mathbb{R}, \quad \mathfrak{A}(t, u, v) := \mathfrak{A}_{\Gamma^W}(u \circ (\mathbf{F}_t^W)^{-1}, v \circ (\mathbf{F}_t^W)^{-1}), \\ A &: [0, T] \times V_c(\Omega \cup \mathcal{I})^2 \rightarrow \mathbb{R}, \quad A(t, u, v) := A_{\Gamma^W}(u \circ (\mathbf{F}_t^W)^{-1}, v \circ (\mathbf{F}_t^W)^{-1}), \\ F &: [0, T] \times V_c(\Omega \cup \mathcal{I}) \rightarrow \mathbb{R}, \quad F(t, \psi) := F_{\Gamma^W}(\psi \circ (\mathbf{F}_t^W)^{-1}), \\ \tilde{F} &: [0, T] \times V_c(\Omega \cup \mathcal{I})^2 \rightarrow \mathbb{R}, \quad \tilde{F}(t, u, \varphi) := \tilde{F}_{\Gamma^W}(u \circ (\mathbf{F}_t^W)^{-1}, \varphi \circ (\mathbf{F}_t^W)^{-1}).\end{aligned}$$

Additionally, we denote the *directional derivative* of \mathfrak{G} w.r.t. u or v as follows

$$\begin{aligned}D_u \mathfrak{G}(t, u, v, \psi, \varphi)[\tilde{u}] &:= \lim_{s \searrow 0} \frac{\mathfrak{G}(t, u + s\tilde{u}, v, \psi, \varphi) - \mathfrak{G}(t, u, v, \psi, \varphi)}{s} \text{ and} \\ D_v \mathfrak{G}(t, u, v, \psi, \varphi)[\tilde{v}] &:= \lim_{s \searrow 0} \frac{\mathfrak{G}(t, u, v + s\tilde{v}, \psi, \varphi) - \mathfrak{G}(t, u, v, \psi, \varphi)}{s}.\end{aligned}$$

In order to apply AAM, the following assumptions need to be satisfied:

Assumptions:

(D0) For all $t \in [0, T]$ and $u, \tilde{u}, \hat{u}, v, \tilde{v}, \hat{v}, \psi, \varphi \in V_c(\Omega \cup \mathcal{I})$ we assume

- (i) $\mathfrak{G}_1 : [0, 1] \rightarrow \mathbb{R}, s \mapsto \mathfrak{G}(t, u + s\tilde{u}, v, \psi, \varphi)$ and $\mathfrak{G}_2 : [0, 1] \rightarrow \mathbb{R}, s \mapsto \mathfrak{G}(t, u, v + s\tilde{v}, \psi, \varphi)$ are absolutely continuous.
- (ii) $[0, 1] \ni s \mapsto D_u \mathfrak{G}(t, u + s\hat{u}, v, \psi, \varphi)[\tilde{u}]$ and $[0, 1] \ni s \mapsto D_v \mathfrak{G}(t, u, v + s\hat{v}, \psi, \varphi)[\tilde{v}]$ belong to $L^1((0, 1))$.
- (iii) $\psi \mapsto \mathfrak{G}(t, u, v, \psi, \varphi)$ and $\varphi \mapsto \mathfrak{G}(t, u, v, \psi, \varphi)$ are affine-linear.

(D1) For all $u, v, \psi, \varphi \in V_c(\Omega \cup \mathcal{I})$ the function

$$[0, T] \rightarrow \mathbb{R}, t \mapsto \mathfrak{G}(t, u, v, \psi, \varphi)$$

is differentiable.

(D2) For all $t \in [0, T]$ we suppose that

- There exist unique solutions $u^t, v^t \in V_c(\Omega \cup \mathcal{I})$ to the (state) equations

$$A(t, u^t, \psi) = F(t, \psi) \quad \text{for all } \psi \in V_c(\Omega \cup \mathcal{I}), \quad (12)$$

$$A(t, \varphi, v^t) = \tilde{F}(t, u^t, \varphi) \quad \text{for all } \varphi \in V_c(\Omega \cup \mathcal{I}). \quad (13)$$

- There exist unique solutions $\psi^t, \varphi^t \in V_c(\Omega \cup \mathcal{I})$ to the *averaged adjoint equations*

$$\int_0^1 D_u \mathfrak{G}(t, su^t + (1-s)u^0, v^t, \psi^t, \varphi^t)[\tilde{u}] ds = 0 \text{ for all } \tilde{u} \in V_c(\Omega \cup \mathcal{I}), \quad (14)$$

$$\int_0^1 D_v \mathfrak{G}(t, u^0, sv^t + (1-s)v^0, \psi^t, \varphi^t)[\tilde{v}] ds = 0 \text{ for all } \tilde{v} \in V_c(\Omega \cup \mathcal{I}). \quad (15)$$

(D3) For every sequence $(s_n)_{n \in \mathbb{N}} \in [0, T]^{\mathbb{N}}$ with $\lim_{n \rightarrow \infty} s_n = 0$, there exists a subsequence $(s_{n_k})_{k \in \mathbb{N}}$ such that

$$\lim_{k \rightarrow \infty, t \searrow 0} \partial_t \mathfrak{G}(t, u^0, v^0, \psi^{s_{n_k}}, \varphi^{s_{n_k}}) = \partial_t \mathfrak{G}(0, u^0, v^0, \psi^0, \varphi^0).$$

Theorem 1 (Averaged Adjoint Method for a System of State Equations) *If the Assumptions (D0)-(D3) are satisfied, we get for any $\psi, \varphi \in V_c(\Omega \cup \mathcal{I})$ that*

$$\left. \frac{d}{dt} \right|_{t \searrow 0} \mathfrak{G}(t, u^t, v^t, \psi, \varphi) = \partial_t \mathfrak{G}(0, u^0, v^0, \psi^0, \varphi^0).$$

Proof See [26, Theorem 4.5]. □

Corollary 2 *If Assumptions (D0)-(D3) hold, the second shape derivative of the reduced functional can be computed via*

$$D^2 J^{red}(\Gamma)[V, W] = \partial_t \mathfrak{G}(0, u^0, v^0, \psi^0, \varphi^0).$$

Finally, we would like to highlight the fact that the solutions u^0 and v^0 to the state equations (12) and (13) for $t = 0$ are independent of \mathbf{V} . However, since the adjoint equations (14) and (15) involve the shape derivative at Γ in the direction \mathbf{V} , the functions ψ^0 and φ^0 need to be computed for every \mathbf{V} separately. Therefore, we write $\psi(\Gamma, \mathbf{V})$ and $\varphi(\Gamma, \mathbf{V})$ in Sect. 7 to explicitly express which adjoints are calculated during the presented algorithm.

6 Deriving the second shape derivative for the reduced objective function

Before we actually calculate the second shape derivative, we need to prove that the Assumptions (D0)-(D3) hold.

Remark 5 In the following, we use the abbreviations

$$\begin{aligned} \bar{u}^t(\mathbf{x}) &:= \bar{u}(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x})), & \nabla \bar{u}^t(\mathbf{x}) &:= (\nabla \bar{u})(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x})), \\ f^t(\mathbf{x}) &:= f_{\Gamma^t, \mathbf{W}}(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x})), & \nabla f^t(\mathbf{x}) &:= (\nabla f)(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x})), \\ \gamma^t(\mathbf{x}, \mathbf{y}) &:= \gamma_{\Gamma^t, \mathbf{W}}(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x}), \mathbf{F}_t^{\mathbf{W}}(\mathbf{y})), & \nabla \gamma^t(\mathbf{x}, \mathbf{y}) &:= (\nabla \gamma_{\Gamma^t, \mathbf{W}})(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x}), \mathbf{F}_t^{\mathbf{W}}(\mathbf{y})), \\ \mathbf{V}^t(\mathbf{x}) &:= \mathbf{V}(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x})), & \mathbf{V}^t(\mathbf{x}, \mathbf{y}) &:= (\mathbf{V}^t(\mathbf{x}), \mathbf{V}^t(\mathbf{y}))^\top, \\ \mathbf{W}(\mathbf{x}, \mathbf{y}) &:= (\mathbf{W}(\mathbf{x}), \mathbf{W}(\mathbf{y}))^\top, & \xi^t(\mathbf{x}) &:= \det D\mathbf{F}_t^{\mathbf{W}}(\mathbf{x}), \\ \Psi_1^t(\mathbf{x}, \mathbf{y}) &:= \nabla \gamma^t(\mathbf{x}, \mathbf{y})^\top \mathbf{V}^t(\mathbf{x}, \mathbf{y}) \text{ and} \end{aligned}$$

$$\Psi_2^t(\mathbf{x}, \mathbf{y}) := \gamma^t(\mathbf{x}, \mathbf{y})(\operatorname{div} \mathbf{V}^t(\mathbf{x}) + \operatorname{div} \mathbf{V}^t(\mathbf{y})) \text{ for } \mathbf{V}, \mathbf{W} \in C_0^2(\Omega \cup \mathcal{I}, \mathbb{R}^d).$$

As a result, $\Psi_i^t(\mathbf{x}, \mathbf{y}) = \Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}^t}^i(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x}), \mathbf{F}_t^{\mathbf{W}}(\mathbf{y}))$ holds for all $\mathbf{x}, \mathbf{y} \in \Omega \cup \mathcal{I}$ and $i = 1, 2$. Analogously, we set

$$\begin{aligned} \operatorname{Hess}(\bar{u})^t(\mathbf{x}) &:= \operatorname{Hess}(\bar{u})(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x})), \quad \operatorname{Hess}(f)^t(\mathbf{x}) := \operatorname{Hess}(f)(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x})) \text{ and} \\ \operatorname{Hess}(\gamma)^t(\mathbf{x}, \mathbf{y}) &:= \operatorname{Hess}(\gamma)(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x}), \mathbf{F}_t^{\mathbf{W}}(\mathbf{y})). \end{aligned}$$

Lemma 3 *Given any $t \in [0, T]$ and $u, \tilde{u}, \hat{u}, v, \tilde{v}, \hat{v}, \psi, \varphi \in V_c(\Omega \cup \mathcal{I})$, Assumption (D0) is satisfied.*

Proof (i) By using the linearity of A and \mathfrak{A} regarding the second and third argument, we get

$$\begin{aligned} \mathfrak{G}'_1(s) &= D_u \mathfrak{G}(t, u + s\tilde{u}, v, \psi, \varphi)[\tilde{u}] \\ &= D_u \mathfrak{J}(t, u + s\tilde{u})[\tilde{u}] + \mathfrak{A}(t, \tilde{u}, v) + A(t, \tilde{u}, \psi) - D_u \tilde{F}(t, u + s\tilde{u}, \varphi)[\tilde{u}], \\ \mathfrak{G}'_2(s) &= D_v \mathfrak{G}(t, u, v + s\tilde{v}, \psi, \varphi)[\tilde{v}] = -\mathfrak{F}(t, \tilde{v}) + \mathfrak{A}(t, u, \tilde{v}) + A(t, \varphi, \tilde{v}), \end{aligned}$$

where \mathfrak{F} is also linear in the second argument and where we have

$$\begin{aligned} D_u \tilde{F}(t, u, \varphi)[\tilde{u}] &= - \int_{\Omega} \tilde{u} \varphi \xi^t \, d\mathbf{x} \text{ and} \\ D_u \mathfrak{J}(t, u)[\tilde{u}] &= \int_{\Omega} -\tilde{u}(\nabla \bar{u}^t)^\top \mathbf{V}^t \xi^t + (u - \bar{u}^t) \tilde{u} \operatorname{div} \mathbf{V}^t \xi^t \, d\mathbf{x}. \end{aligned}$$

Thus, \mathfrak{G}_1 and \mathfrak{G}_2 are continuously differentiable and therefore absolutely continuous. Lastly, note that $D_u \tilde{F}(t, u, \varphi)[\tilde{u}]$ is independent of the choice of u .

(ii) By examining the computations in 1., we directly see that \mathfrak{G}'_2 is constant on $(0, 1)$. Further, the variable s can be found in only one integral of \mathfrak{G}'_1 and we get as a consequence of Fubini's theorem

$$\begin{aligned} \int_0^1 D_u \mathfrak{G}_1(t, u + s\hat{u}, v, \psi, \varphi)[\tilde{u}] \, ds &= \int_{\Omega} \left(u + \frac{1}{2} \hat{u} - \bar{u}^t \right) \tilde{u} \operatorname{div} \mathbf{V}^t \xi^t - \tilde{u}(\nabla \bar{u}^t)^\top \mathbf{V}^t \xi^t \, d\mathbf{x} \\ &\quad + \mathfrak{A}(t, \tilde{u}, v) + A(t, \tilde{u}, \psi) - D_u \tilde{F}(t, u, v)[\tilde{u}] \\ &= D_u \mathfrak{J}(t, u + \frac{1}{2} \hat{u})[\tilde{u}] + \mathfrak{A}(t, \tilde{u}, v) + A(t, \tilde{u}, \psi) - D_u \tilde{F}(t, u, v)[\tilde{u}]. \end{aligned}$$

(iii) Direct consequence of the linearity of A , F and \tilde{F} in the corresponding argument. \square

Corollary 4 *As seen in the proof of Lemma 3, the averaged adjoint equations (14) and (15) can be formulated as*

Find $\psi^t, \varphi^t \in V_c(\Omega \cup \mathcal{I})$ such that

$$A(t, \tilde{u}, \psi^t) = -D_u \mathfrak{J}(t, \frac{1}{2}(u^t + u^0))[\tilde{u}] - \mathfrak{A}(t, \tilde{u}, v^t) + D_u \tilde{F}(t, u^0, \varphi^t)[\tilde{u}] \quad (16)$$

for all $\tilde{u} \in V_c(\Omega \cup \mathcal{I})$ and

$$A(t, \varphi^t, \tilde{v}) = \mathfrak{F}(t, \tilde{v}) - \mathfrak{A}(t, u^0, \tilde{v}) \text{ for all } \tilde{v} \in V_c(\Omega \cup \mathcal{I}). \quad (17)$$

Typically, (16) and (17) can each be interpreted as a variational formulation of a nonlocal Dirichlet problem. Therefore, the operator $\mathfrak{A}(t, \cdot, \cdot)$ needs to be continuous, which is proven in Lemma 6 below. The remaining requirements of Theorem 1 are satisfied if the following assumptions hold:

Assumption (S0):

- (i) For every $t \in [0, T]$, there exist unique solutions $u^t, v^t \in V_c(\Omega \cup \mathcal{I})$ to the state equations (12) and (13). Moreover, there are unique functions ψ^t, φ^t that solve (16) and (17), i.e., Assumption (D2) is satisfied.
- (ii) • If γ_Γ is an integrable kernel, we assume that there exists a constant $C > 0$ such that

$$A(t, u, u) \geq C \|u\|_{L^2(\Omega \cup \mathcal{I})}^2 \text{ for all } t \in [0, T] \text{ and } u \in V_c(\Omega \cup \mathcal{I}).$$

- If γ_Γ is a singular symmetric kernel, we suppose that there exists a constant $C > 0$ such that

$$A(t, u, u) \geq C \|u\|_{H^s(\Omega \cup \mathcal{I})}^2 \text{ for all } t \in [0, T] \text{ and } u \in V_c(\Omega \cup \mathcal{I}).$$

Assumption (S1):

Each of the two kernel classes has to meet additional conditions:

- For integrable kernels it is assumed to hold

$$\gamma_{ij} \in W^{2,\infty}(\Omega \times \Omega, \mathbb{R}), \quad \gamma_{i\mathcal{I}} \in W^{2,\infty}(\Omega \times \mathcal{I}, \mathbb{R}) \text{ for all } i, j = 1, 2.$$

- For $n \in \mathbb{N}$, define

$$D_n^{\mathbb{R}^d \times \mathbb{R}^d} := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d \times \mathbb{R}^d : \|\mathbf{x} - \mathbf{y}\|_2 \geq \frac{1}{n}\}.$$

Then, for a singular symmetric kernel we assume $\gamma_{ij}, \gamma_{i\mathcal{I}} \in W^{2,\infty}(D_n^{\mathbb{R}^d \times \mathbb{R}^d})$ for $i, j = 1, 2$ and $n \in \mathbb{N}$. Additionally, we suppose that there exists a constant $C > 0$ such that

$$\begin{aligned} |\gamma_{ij}(\mathbf{x}, \mathbf{y})| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} &< C, \quad |\nabla \gamma_{ij}(\mathbf{x}, \mathbf{y})^\top \mathbf{V}(\mathbf{x}, \mathbf{y})| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} < C, \\ |\gamma_{i\mathcal{I}}(\mathbf{x}, \mathbf{y})| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} &< C, \quad |\nabla_{\mathbf{x}} \gamma_{i\mathcal{I}}(\mathbf{x}, \mathbf{y})^\top \mathbf{V}(\mathbf{x})| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} < C, \\ |\mathbf{V}(\mathbf{x}, \mathbf{y})^\top \text{Hess}(\gamma_{ij})(\mathbf{x}, \mathbf{y}) \mathbf{W}(\mathbf{x}, \mathbf{y})| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} &< C \text{ and} \\ |\mathbf{V}(\mathbf{x}) \frac{\partial^2}{\partial \mathbf{x} \partial \mathbf{x}} \gamma_{i\mathcal{I}}(\mathbf{x}, \mathbf{y}) \mathbf{W}(\mathbf{x})| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} &< C \end{aligned}$$

hold for $n \in \mathbb{N}$, $i, j = 1, 2$ and almost every $(\mathbf{x}, \mathbf{y}) \in (\Omega \cup \mathcal{I})^2$.

In the following, we also consider the sets $D_n := D_n^{\mathbb{R}^d \times \mathbb{R}^d} \cap (\Omega \cup \mathcal{I})^2$ for $n \in \mathbb{N}$. As we will see in Lemma 5, Assumption (S1) yields the well-posedness of certain integrals, which are components of the second shape derivative. Moreover, the fact that vector fields \mathbf{V} and \mathbf{W} vanish on \mathcal{I} is already incorporated into the conditions w.r.t. $\gamma_{\mathcal{I}}$ in (S1) in the case of singular symmetric kernels, i.e., we only have to differentiate $\gamma_{\mathcal{I}}$ regarding \mathbf{x} .

Remark 6 Let γ_{Γ} be an integrable kernel and let (S1) be satisfied. Since the indicator function is in general not weakly differentiable, we cannot conclude that

$$\gamma_{\Gamma}^{\mathbf{W}}(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1,2} \gamma_{ij}(\mathbf{x}, \mathbf{y}) \chi_{\mathbf{F}_i^{\mathbf{W}}(\Omega_i) \times \mathbf{F}_j^{\mathbf{W}}(\Omega_j)}(\mathbf{x}, \mathbf{y}) + \sum_{i=1,2} \gamma_{i\mathcal{I}}(\mathbf{x}, \mathbf{y}) \chi_{\mathbf{F}_i^{\mathbf{W}}(\Omega_i) \times \mathcal{I}}(\mathbf{x}, \mathbf{y})$$

is (twice) weakly differentiable with weak derivatives in $L^\infty((\Omega \cup \mathcal{I})^2)$. However, we derive $\gamma_{\Gamma}^{\mathbf{W}} \in W^{2,\infty}(\mathbf{F}_i^{\mathbf{W}}(\Omega_i) \times \mathbf{F}_j^{\mathbf{W}}(\Omega_j))$ and $\gamma_{\Gamma}^{\mathbf{W}} \in W^{2,\infty}(\mathbf{F}_i^{\mathbf{W}}(\Omega_i) \times \mathcal{I})$ for $i, j = 1, 2$. If γ_{Γ} is a singular symmetric kernel and (S1) holds, we get $\gamma_{\Gamma}^{\mathbf{W}} \in W^{2,\infty}(D_n \cap (\mathbf{F}_i^{\mathbf{W}}(\Omega_i) \times \mathbf{F}_j^{\mathbf{W}}(\Omega_j)))$ and $\gamma_{\Gamma}^{\mathbf{W}} \in W^{2,\infty}(D_n \cap (\mathbf{F}_i^{\mathbf{W}}(\Omega_i) \times \mathcal{I}))$ for $i, j = 1, 2$ and $n \in \mathbb{N}$. For ease of presentation, $\nabla \gamma_{\Gamma}^{\mathbf{W}}$ and $\text{Hess}(\gamma_{\Gamma}^{\mathbf{W}})$ are understood as the weak first and second derivatives of the kernel on those subdomains and we do not decompose integrals on $(\Omega \cup \mathcal{I})^2$ or D_n into the sum of integrals on these subdomains in the following. The same convention will be also applied to $f_{\Gamma}^{\mathbf{W}} \in H^2(\mathbf{F}_i^{\mathbf{W}}(\Omega_i))$, where $i = 1, 2$.

We now examine a popular example of a singular symmetric kernel and show that (S1) is satisfied under some additional assumptions:

Example 1 Let a symmetric function $\sigma_{\Gamma} := \sum_{i,j=1,2} \sigma_{ij} \chi_{\Omega_i \times \Omega_j} + \sum_{i=1,2} \sigma_{i\mathcal{I}} \chi_{\Omega_i \times \mathcal{I}}$ be given, where $\sigma_{ij} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, \infty)$ and $\sigma_{i\mathcal{I}} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, \infty)$ for $i, j = 1, 2$. Here, the functions σ_{ii} are supposed to be symmetric for $i = 1, 2$ and we consider $\sigma_{12}(\mathbf{x}, \mathbf{y}) = \sigma_{21}(\mathbf{y}, \mathbf{x})$ to hold for $(\mathbf{x}, \mathbf{y}) \in \Omega \times \Omega$. Additionally, we assume the existence of constants $0 < \sigma_* \leq \sigma^* < \infty$ such that $\sigma_* \leq \sigma_{ij} \leq \sigma^*$ and $\sigma_* \leq \sigma_{i\mathcal{I}} \leq \sigma^*$ on $\mathbb{R}^d \times \mathbb{R}^d$ for all $i, j = 1, 2$. Then, a popular choice of a singular symmetric kernel is the function

$$\gamma_{\Gamma}(\mathbf{x}, \mathbf{y}) := \frac{\sigma_{\Gamma}(\mathbf{x}, \mathbf{y})}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} \chi_{B_s}(\mathbf{y}).$$

Due to the boundedness of σ_{ij} and $\sigma_{i\mathcal{I}}$ from below and above as described before, Assumption (S0)(1) is already satisfied since condition (3) also holds for perturbed domains $\Omega(\Gamma_{\Gamma}^{\mathbf{W}}) \cup \mathcal{I}$. Assumption (S0)(2) can also be easily proven. We now show that Assumption (S1) is also valid if $\sigma_{ij}, \sigma_{i\mathcal{I}} \in W^{2,\infty}(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R})$ and if σ_{ij} and $\sigma_{i\mathcal{I}}$ are translation invariant for $i, j = 1, 2$. Therefore, we set $\tilde{\gamma}(\mathbf{x}, \mathbf{y}) := \frac{1}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}}$, $c_1 := -(d + 2s)$ and $c_2 := 2(d + 2s) + (d + 2s)^2$. For $\tilde{\gamma}$ we derive

$$\frac{\partial}{\partial \mathbf{x}} \tilde{\gamma}(\mathbf{x}, \mathbf{y}) = c_1 \tilde{\gamma}(\mathbf{x}, \mathbf{y}) \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|_2^2} \text{ and } \frac{\partial}{\partial \mathbf{y}} \tilde{\gamma}(\mathbf{x}, \mathbf{y}) = -\frac{\partial}{\partial \mathbf{x}} \tilde{\gamma}(\mathbf{x}, \mathbf{y}).$$

Moreover, the second partial derivatives of $\tilde{\gamma}$ can be expressed as

$$\begin{aligned}\frac{\partial^2}{\partial \mathbf{x} \partial \mathbf{x}} \tilde{\gamma}(\mathbf{x}, \mathbf{y}) &= c_1 \tilde{\gamma}(\mathbf{x}, \mathbf{y}) \frac{\mathbf{I}}{\|\mathbf{x} - \mathbf{y}\|_2^2} + c_2 \tilde{\gamma}(\mathbf{x}, \mathbf{y}) \frac{(\mathbf{x} - \mathbf{y})(\mathbf{x} - \mathbf{y})^\top}{\|\mathbf{x} - \mathbf{y}\|_2^4}, \\ \frac{\partial^2}{\partial \mathbf{x} \partial \mathbf{y}} \tilde{\gamma}(\mathbf{x}, \mathbf{y}) &= \frac{\partial^2}{\partial \mathbf{y} \partial \mathbf{x}} \tilde{\gamma}(\mathbf{x}, \mathbf{y}) = -\frac{\partial^2}{\partial \mathbf{x} \partial \mathbf{x}} \tilde{\gamma}(\mathbf{x}, \mathbf{y}) \text{ and } \frac{\partial^2}{\partial \mathbf{y} \partial \mathbf{y}} \tilde{\gamma}(\mathbf{x}, \mathbf{y}) = \frac{\partial^2}{\partial \mathbf{x} \partial \mathbf{x}} \tilde{\gamma}(\mathbf{x}, \mathbf{y}).\end{aligned}$$

Since σ_{ij} is translation invariant, $\gamma_{ij} = \sigma_{ij} \tilde{\gamma}$ and due to the Lipschitz continuity of \mathbf{V} and \mathbf{W} , we can conclude

$$\begin{aligned}& \left| \nabla \gamma_{ij}(\mathbf{x}, \mathbf{y})^\top \mathbf{V}(\mathbf{x}, \mathbf{y}) \right| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} \\ & \leq \left| c_1 \sigma_{ij}(\mathbf{x}, \mathbf{y}) \frac{(\mathbf{x} - \mathbf{y})^\top (\mathbf{V}(\mathbf{x}) - \mathbf{V}(\mathbf{y}))}{\|\mathbf{x} - \mathbf{y}\|_2^2} \right| + \left| \nabla \sigma_{ij}(\mathbf{x}, \mathbf{y})^\top \mathbf{V}(\mathbf{x}, \mathbf{y}) \right| < C < \infty \text{ and} \\ & \left| \mathbf{V}(\mathbf{x}, \mathbf{y})^\top \text{Hess}(\gamma_{ij})(\mathbf{x}, \mathbf{y}) \mathbf{W}(\mathbf{x}, \mathbf{y}) \right| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} \\ & = \left| (\mathbf{V}(\mathbf{x}) - \mathbf{V}(\mathbf{y}))^\top \left(\sigma_{ij}(\mathbf{x}, \mathbf{y}) \left(c_1 \frac{\mathbf{I}}{\|\mathbf{x} - \mathbf{y}\|_2^2} + c_2 \frac{(\mathbf{x} - \mathbf{y})(\mathbf{x} - \mathbf{y})^\top}{\|\mathbf{x} - \mathbf{y}\|_2^4} \right) \right. \right. \\ & \quad \left. \left. + 2 \frac{\partial}{\partial \mathbf{x}} \sigma_{ij}(\mathbf{x}, \mathbf{y}) \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|_2^2} + \frac{\partial^2}{\partial \mathbf{x} \partial \mathbf{x}} \sigma_{ij}(\mathbf{x}, \mathbf{y}) \right) (\mathbf{W}(\mathbf{x}) - \mathbf{W}(\mathbf{y})) \right| < C < \infty,\end{aligned}$$

for almost every $(\mathbf{x}, \mathbf{y}) \in \Omega \times \Omega$ and for some constant $C > 0$. The remaining conditions on $\gamma_{i\mathcal{I}}$ also hold, which can be proven analogously.

We continue by illustrating that the remaining requirements hold if Assumptions (S0) and (S1) are satisfied. In particular, the existence of $\partial_t \mathfrak{G}(t, u, v, \psi, \varphi)$ is shown in Lemma 5. In the case of a singular symmetric kernel, the following remark is helpful to illustrate that integrals involving γ^t , $\nabla \gamma^t$ or $\text{Hess}(\gamma)^t$ of $\partial_t \mathfrak{G}(t, u, v, \psi, \varphi)$ are finite:

Remark 7 Let γ_Γ be a singular symmetric kernel with $|\gamma_{ij}(\mathbf{x}, \mathbf{y})| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} < \gamma^*$ for $\mathbf{x}, \mathbf{y} \in \Omega$ and $|\gamma_{i\mathcal{I}}(\mathbf{x}, \mathbf{y})| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} < \gamma^*$ for $(\mathbf{x}, \mathbf{y}) \in \Omega \times \mathcal{I}$ for some constant $\gamma^* > 0$.

Since $\mathbf{V} \in C_0^2(\Omega \cup \mathcal{I}, \mathbb{R}^d)$, there exists a Lipschitz constant $L_V > 0$ such that

$$\begin{aligned}\|\mathbf{x} - \mathbf{y}\|_2 &\leq \|\mathbf{x} - \mathbf{y} + t(\mathbf{V}(\mathbf{x}) - \mathbf{V}(\mathbf{y}))\|_2 + t\|\mathbf{V}(\mathbf{x}) - \mathbf{V}(\mathbf{y})\|_2 \\ &\leq \|\mathbf{F}_t^V(\mathbf{x}) - \mathbf{F}_t^V(\mathbf{y})\|_2 + tL_V\|\mathbf{x} - \mathbf{y}\|_2 \text{ for } \mathbf{x}, \mathbf{y} \in \Omega \cup \mathcal{I}.\end{aligned}$$

Let $T > 0$ be sufficiently small such that $1 - TL_V > 0$. Then, we obtain

$$\|\mathbf{F}_t^V(\mathbf{x}) - \mathbf{F}_t^V(\mathbf{y})\|_2 \geq (1 - TL_V)\|\mathbf{x} - \mathbf{y}\|_2 \text{ for } \mathbf{x}, \mathbf{y} \in \Omega \cup \mathcal{I} \text{ and } t \in [0, T].$$

With $L := \frac{1}{1-TL_V}$, we derive for $t \in [0, T]$ and almost every $(\mathbf{x}, \mathbf{y}) \in D_n$ that

$$\gamma^t(\mathbf{x}, \mathbf{y}) = \gamma_{\Gamma_t}(\mathbf{F}_t^V(\mathbf{x}), \mathbf{F}_t^V(\mathbf{y})) \leq \frac{\gamma^*}{\|\mathbf{F}_t^V(\mathbf{x}) - \mathbf{F}_t^V(\mathbf{y})\|_2^{d+2s}} \leq \frac{L\gamma^*}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} < L\gamma^*n^{d+2s}.$$

As a result, we obtain

$$|\gamma^t(\mathbf{x}, \mathbf{y})| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} < L\gamma^* \text{ for every } t \in [0, T] \text{ and almost every } (\mathbf{x}, \mathbf{y}) \in (\Omega \cup \mathcal{I})^2.$$

If Assumption (S1) holds, it can be shown in a similar way that there exists a constant $C > 0$ with

$$\begin{aligned} |\nabla \gamma^t(\mathbf{x}, \mathbf{y})^\top \mathbf{V}(\mathbf{x}, \mathbf{y})| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} &< C \text{ and} \\ |\mathbf{V}(\mathbf{x}, \mathbf{y})^\top \text{Hess}(\gamma)^t(\mathbf{x}, \mathbf{y}) \mathbf{W}(\mathbf{x}, \mathbf{y})| \|\mathbf{x} - \mathbf{y}\|_2^{d+2s} &< C \end{aligned}$$

for almost every $(\mathbf{x}, \mathbf{y}) \in (\Omega \cup \mathcal{I})^2$ and for all $t \in [0, T]$. Note that $\xi^t(\mathbf{x})$ is bounded from above for $t \in [0, T]$ and $\mathbf{x} \in \Omega \cup \mathcal{I}$. Thus, integrals like, e.g.,

$$\begin{aligned} &\iint_{(\Omega \cup \mathcal{I})^2} (u(\mathbf{x}) - u(\mathbf{y}))(v(\mathbf{x}) - v(\mathbf{y})) \gamma^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \, d\mathbf{y} d\mathbf{x} \text{ or} \\ &* \iint_{(\Omega \cup \mathcal{I})^2} (u(\mathbf{x}) - u(\mathbf{y}))(v(\mathbf{x}) - v(\mathbf{y})) \nabla \gamma^t(\mathbf{x}, \mathbf{y})^\top \mathbf{V}(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \, d\mathbf{y} d\mathbf{x} \end{aligned}$$

are well-defined for functions $u, v \in H_c^s(\Omega \cup \mathcal{I})$.

Lemma 5 *If Assumption (S1) holds, then for any $u, v, \psi, \varphi \in V_c(\Omega \cup \mathcal{I})$ and $t \in [0, T]$ the derivative $\partial_t \mathfrak{G}(t, u, v, \psi, \varphi)$ exists, i.e., (D1) is satisfied.*

Proof Since the proof is quite lengthy, it can be found in Appendix A.1. \square

In order to prove the weak convergences $\varphi^t \rightharpoonup \varphi^0$ and $\psi^t \rightharpoonup \psi^0$ in Lemma 7, which are helpful to show that Assumption (D3) is satisfied, we make use of the boundedness of \mathfrak{A} :

Lemma 6 *Let Assumption (S1) hold.*

- If γ_Γ is an integrable kernel, then there exists a constant $C > 0$ such that

$$|\mathfrak{A}(t, u, v)| \leq C \|u\|_{L^2(\Omega \cup \mathcal{I})} \|v\|_{L^2(\Omega \cup \mathcal{I})} \text{ for } t \in [0, T] \text{ and } u, v \in V_c(\Omega \cup \mathcal{I}).$$

- If γ_Γ is a singular symmetric kernel, then there exists a constant $C > 0$ such that

$$|\mathfrak{A}(t, u, v)| \leq C |u|_{H^s(\Omega \cup \mathcal{I})} |v|_{H^s(\Omega \cup \mathcal{I})} \text{ for } t \in [0, T] \text{ and } u, v \in V_c(\Omega \cup \mathcal{I}).$$

Proof By utilizing the triangle inequality, we can directly deduce that

$$\begin{aligned} & |\mathfrak{A}(t, u, v)| \\ & \leq \sum_{i=1,2} \left| \iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x})\Psi_i^t(\mathbf{x}, \mathbf{y}) - u(\mathbf{y})\Psi_i^t(\mathbf{y}, \mathbf{x}))\xi^t(\mathbf{x})\xi^t(\mathbf{y}) \, dyd\mathbf{x} \right|. \end{aligned}$$

Case 1: Integrable Kernel

First, we denote that ξ^t and $\operatorname{div} \mathbf{V}^t$ are continuous and bounded on $\Omega \cup \mathcal{I}$ and therefore

$$\Psi_2^t(\mathbf{x}, \mathbf{y})\xi^t(\mathbf{x})\xi^t(\mathbf{y}) = \gamma^t(\mathbf{x}, \mathbf{y})(\operatorname{div} \mathbf{V}^t(\mathbf{x}) + \operatorname{div} \mathbf{V}^t(\mathbf{y}))\xi^t(\mathbf{x})\xi^t(\mathbf{y})$$

is bounded by a constant $\tilde{C}_1 > 0$ for all $t \in [0, T]$ and for almost every $(\mathbf{x}, \mathbf{y}) \in (\Omega \cup \mathcal{I})^2$ due to Assumption (S1). Then, we can conclude the existence of a constant $C_1 > 0$ such that for the term regarding Ψ_2 holds

$$\begin{aligned} & \left| \iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x})\Psi_2^t(\mathbf{x}, \mathbf{y}) - u(\mathbf{y})\Psi_2^t(\mathbf{y}, \mathbf{x}))\xi^t(\mathbf{x})\xi^t(\mathbf{y}) \, dyd\mathbf{x} \right| \\ & \leq \tilde{C}_1 \iint_{(\Omega \cup \mathcal{I})^2} |(v(\mathbf{x}) - v(\mathbf{y}))u(\mathbf{x})| + |(v(\mathbf{x}) - v(\mathbf{y}))u(\mathbf{y})| \, dyd\mathbf{x} \\ & \leq C_1 \|u\|_{L^2(\Omega \cup \mathcal{I})} \|v\|_{L^2(\Omega \cup \mathcal{I})}. \end{aligned}$$

Since also $\nabla \gamma^t$ is essentially bounded for all $t \in [0, T]$, the term

$$\Psi_1^t(\mathbf{x}, \mathbf{y})\xi^t(\mathbf{x})\xi^t(\mathbf{y}) = \nabla \gamma^t(\mathbf{x}, \mathbf{y})^\top \mathbf{V}^t(\mathbf{x}, \mathbf{y})\xi^t(\mathbf{x})\xi^t(\mathbf{y})$$

is essentially bounded by a constant $\tilde{C}_2 > 0$ for all $t \in [0, T]$ and for almost every $(\mathbf{x}, \mathbf{y}) \in (\Omega \cup \mathcal{I})^2$. Consequently, we get for some $C_2 > 0$ that

$$\begin{aligned} & \left| \iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x})\Psi_1^t(\mathbf{x}, \mathbf{y}) - u(\mathbf{y})\Psi_1^t(\mathbf{y}, \mathbf{x}))\xi^t(\mathbf{x})\xi^t(\mathbf{y}) \, dyd\mathbf{x} \right| \\ & \leq \iint_{(\Omega \cup \mathcal{I})^2} \tilde{C}_2 |(v(\mathbf{x}) - v(\mathbf{y}))u(\mathbf{x})| + \tilde{C}_2 |(v(\mathbf{x}) - v(\mathbf{y}))u(\mathbf{y})| \, dyd\mathbf{x} \\ & \leq C_2 \|u\|_{L^2(\Omega \cup \mathcal{I})} \|v\|_{L^2(\Omega \cup \mathcal{I})}. \end{aligned}$$

Case 2: Singular Symmetric Kernel

Here, we can directly conclude the existence of a constant $C_1 > 0$ with

$$|\Psi_2^t(\mathbf{x}, \mathbf{y})\xi^t(\mathbf{x})\xi^t(\mathbf{y})| \leq \frac{C_1}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} \chi_{B_\delta}(\mathbf{x})(\mathbf{y}) =: \tilde{\gamma}(\mathbf{x}, \mathbf{y})$$

for almost every $(\mathbf{x}, \mathbf{y}) \in (\Omega \cup \mathcal{I})^2$ and for all $t \in [0, T]$ due to Assumption (S1) (see also Remark 7) and the boundedness of \mathbf{V} , $\operatorname{div} \mathbf{V}$ and ξ^t on $\Omega \cup \mathcal{I}$. Therefore, we derive

$$\begin{aligned} & \left| \iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x}) - u(\mathbf{y})) \Psi_2^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \, d\mathbf{y} d\mathbf{x} \right| \\ & \leq \iint_{(\Omega \cup \mathcal{I})^2} |(v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x}) - u(\mathbf{y}))| \tilde{\gamma}(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} d\mathbf{x} \\ & \leq \left(\iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y}))^2 \tilde{\gamma}(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} d\mathbf{x} \right)^{\frac{1}{2}} \left(\iint_{(\Omega \cup \mathcal{I})^2} (u(\mathbf{x}) - u(\mathbf{y}))^2 \tilde{\gamma}(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} d\mathbf{x} \right)^{\frac{1}{2}} \\ & \leq C_1 \|u\|_{H^s(\Omega \cup \mathcal{I})} \|v\|_{H^s(\Omega \cup \mathcal{I})}, \end{aligned}$$

where we obtain the last step by dropping $\chi_{B_\delta(\mathbf{x})}$. From Assumption (S1), as indicated in Remark 7, follows the existence of a constant $C_2 > 0$ with $|\Psi_1^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y})| \leq \frac{C_2}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} \chi_{B_\delta(\mathbf{x})}(\mathbf{y}) =: \bar{\gamma}(\mathbf{x}, \mathbf{y})$ for almost every $(\mathbf{x}, \mathbf{y}) \in (\Omega \cup \mathcal{I})^2$ and for all $t \in [0, T]$. Accordingly, we derive

$$\begin{aligned} & \left| \iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x}) - u(\mathbf{y})) \Psi_1^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \, d\mathbf{y} d\mathbf{x} \right| \\ & \leq \left(\iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y}))^2 \bar{\gamma}(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} d\mathbf{x} \right)^{\frac{1}{2}} \left(\iint_{(\Omega \cup \mathcal{I})^2} (u(\mathbf{x}) - u(\mathbf{y}))^2 \bar{\gamma}(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} d\mathbf{x} \right)^{\frac{1}{2}} \\ & \leq C_2 \|u\|_{H^s(\Omega \cup \mathcal{I})} \|v\|_{H^s(\Omega \cup \mathcal{I})}, \end{aligned}$$

where we again omitted the truncation $\chi_{B_\delta(\mathbf{x})}$ in the last step. \square

Before we prove that the conditions of (D3) are met under some natural assumptions, we need the following weak convergence results.

Lemma 7 *Suppose that Assumptions (S0) and (S1) hold. Then, we derive:*

If γ_Γ is an integrable kernel, we obtain that the families of functions $\{u^t\}_{t \in [0, T]}$, $\{v^t\}_{t \in [0, T]}$, $\{\psi^t\}_{t \in [0, T]}$ and $\{\varphi^t\}_{t \in [0, T]}$ are bounded with respect to $\|\cdot\|_{L^2(\Omega \cup \mathcal{I})}$. Moreover,

$$u^t \rightharpoonup u^0, v^t \rightharpoonup v^0, \psi^t \rightharpoonup \psi^0 \text{ and } \varphi^t \rightharpoonup \varphi^0 \text{ for } t \searrow 0 \text{ in } (L_c^2(\Omega \cup \mathcal{I}), \langle \cdot, \cdot \rangle_{L^2(\Omega \cup \mathcal{I})}).$$

If γ_Γ is a singular symmetric kernel, we get that the families of functions $\{u^t\}_{t \in [0, T]}$, $\{v^t\}_{t \in [0, T]}$, $\{\psi^t\}_{t \in [0, T]}$ and $\{\varphi^t\}_{t \in [0, T]}$ are bounded regarding $\|\cdot\|_{H^s(\Omega \cup \mathcal{I})}$ and $\|\cdot\|_{L^2(\Omega \cup \mathcal{I})}$. Additionally,

$$u^t \rightharpoonup u^0, v^t \rightharpoonup v^0, \psi^t \rightharpoonup \psi^0 \text{ and } \varphi^t \rightharpoonup \varphi^0 \text{ for } t \searrow 0 \text{ in } (H_c^s(\Omega \cup \mathcal{I}), \langle \cdot, \cdot \rangle_{H^s(\Omega \cup \mathcal{I})}).$$

Proof Due to its length, the proof has been moved to Appendix A.2. \square

Lemma 8 *If Assumptions (S0) and (S1) are satisfied, Assumption (D3) of the AAM holds.*

Proof Since the proof is quite lengthy, it can be found in Appendix A.3. \square

Now, we express $D^2 J^{red}(\Gamma)[\mathbf{V}, \mathbf{W}]$ according to the decomposition (5) as the sum of the linear second shape derivative $(J^{red})''(\Gamma)[\mathbf{V}, \mathbf{W}]$ and of the shape derivative $DJ^{red}(\Gamma)[(D\mathbf{V})\mathbf{W}]$. First, note that $(\nabla \operatorname{div} \mathbf{V})^\top \mathbf{W} = \operatorname{div}((D\mathbf{V})\mathbf{W}) - \operatorname{tr}(D\mathbf{V}D\mathbf{W})$ (see, e.g., [31]). Additionally, we can formulate the terms $\nabla(\Psi_{\Gamma, \mathbf{V}}^i(\mathbf{x}, \mathbf{y}))^\top \mathbf{W}(\mathbf{x}, \mathbf{y})$ as a sum of two parts. Therefore, we set

$$\begin{aligned} T_{lin}^1(\mathbf{x}, \mathbf{y}) &= \mathbf{V}(\mathbf{x}, \mathbf{y})^\top \operatorname{Hess}(\gamma_\Gamma)(\mathbf{x}, \mathbf{y}) \mathbf{W}(\mathbf{x}, \mathbf{y}), \\ T_{der}^1(\mathbf{x}, \mathbf{y}) &= \nabla \gamma_\Gamma^\top(\mathbf{x}, \mathbf{y})(D\mathbf{V}(\mathbf{x}, \mathbf{y})) \mathbf{W}(\mathbf{x}, \mathbf{y}), \\ T_{lin}^2(\mathbf{x}, \mathbf{y}) &= (\operatorname{div} \mathbf{V}(\mathbf{x}) + \operatorname{div} \mathbf{V}(\mathbf{y})) \nabla \gamma_\Gamma(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \\ &\quad - \gamma_\Gamma(\mathbf{x}, \mathbf{y})(\operatorname{tr}(D\mathbf{V}(\mathbf{x})D\mathbf{W}(\mathbf{x})) + \operatorname{tr}(D\mathbf{V}(\mathbf{y})D\mathbf{W}(\mathbf{y}))) \text{ and} \\ T_{der}^2(\mathbf{x}, \mathbf{y}) &= \gamma_\Gamma(\mathbf{x}, \mathbf{y})(\operatorname{div}((D\mathbf{V}(\mathbf{x}))\mathbf{W}(\mathbf{x})) + \operatorname{div}((D\mathbf{V}(\mathbf{y}))\mathbf{W}(\mathbf{y}))). \end{aligned}$$

Consequently, we derive

$$\begin{aligned} (\nabla \Psi_{\Gamma, \mathbf{V}}^1(\mathbf{x}, \mathbf{y}))^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) &= T_{lin}^1(\mathbf{x}, \mathbf{y}) + T_{der}^1(\mathbf{x}, \mathbf{y}) \text{ and} \\ (\nabla \Psi_{\Gamma, \mathbf{V}}^2(\mathbf{x}, \mathbf{y}))^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) &= T_{lin}^2(\mathbf{x}, \mathbf{y}) + T_{der}^2(\mathbf{x}, \mathbf{y}). \end{aligned} \quad (18)$$

Corollary 9 (Second Shape Derivative of the Reduced Functional) *Let the functions u^0, v^0 solve (12) and (13). Further, suppose that the functions ψ^0, φ^0 are the solutions to (16) and (17). Additionally, the functions $\Psi_{\Gamma, \mathbf{V}}^1$ and $\Psi_{\Gamma, \mathbf{V}}^2$ are defined as in (11) and their gradient terms can be written as indicated in (18). Then, the second shape derivative of the reduced functional can be expressed as*

$$D^2 J^{red}(\Gamma)[\mathbf{V}, \mathbf{W}] (= \partial_t \mathfrak{G}(0, u^0, v^0, \psi^0, \varphi^0)) = J''(\Gamma)[\mathbf{V}, \mathbf{W}] + DJ(\Gamma)[(D\mathbf{V})\mathbf{W}],$$

where the first component is the linear second shape derivative

$$\begin{aligned} J''(\Gamma)[\mathbf{V}, \mathbf{W}] &= \int_{\Omega} \nabla \bar{u}^\top \mathbf{W} \nabla \bar{u}^\top \mathbf{V} - (u^0 - \bar{u}) \left(\mathbf{V}^\top \operatorname{Hess}(\bar{u}) \mathbf{W} + \nabla \bar{u}^\top \mathbf{V} \operatorname{div} \mathbf{W} \right) d\mathbf{x} \\ &\quad + \int_{\Omega} -(u^0 - \bar{u}) \nabla \bar{u}^\top \mathbf{W} \operatorname{div} \mathbf{V} + \frac{1}{2} (u^0 - \bar{u})^2 (-\operatorname{tr}(D\mathbf{V}D\mathbf{W}) + \operatorname{div} \mathbf{V} \operatorname{div} \mathbf{W}) d\mathbf{x} \\ &\quad + \int_{\Omega} \mathbf{V}^\top \operatorname{Hess}(f_\Gamma) \mathbf{W} v^0 + \nabla f_\Gamma^\top \mathbf{V} v^0 \operatorname{div} \mathbf{W} d\mathbf{x} \\ &\quad + \int_{\Omega} \nabla f_\Gamma^\top \mathbf{W} v^0 \operatorname{div} \mathbf{V} + f_\Gamma v^0 (-\operatorname{tr}(D\mathbf{V}D\mathbf{W}) + \operatorname{div} \mathbf{V} \operatorname{div} \mathbf{W}) d\mathbf{x} \\ &\quad + \sum_{i=1}^2 \frac{1}{2} \left(\iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(u^0(\mathbf{x}) T_{lin}^i(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y}) T_{lin}^i(\mathbf{y}, \mathbf{x})) dy d\mathbf{x} \right. \\ &\quad \left. + \iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(u^0(\mathbf{x}) \Psi_{\Gamma, \mathbf{V}}^i(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y}) \Psi_{\Gamma, \mathbf{V}}^i(\mathbf{y}, \mathbf{x})) (\operatorname{div} \mathbf{W}(\mathbf{x}) + \operatorname{div} \mathbf{W}(\mathbf{y})) dy d\mathbf{x} \right) \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^2 \frac{1}{2} \left(\iint_{(\Omega \cup \mathcal{I})^2} (\psi^0(\mathbf{x}) - \psi^0(\mathbf{y}))(u^0(\mathbf{x})\Psi_{\Gamma, \mathbf{W}}^i(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y})\Psi_{\Gamma, \mathbf{W}}^i(\mathbf{y}, \mathbf{x})) \, dyd\mathbf{x} \right) \\
& + \sum_{i=1}^2 \frac{1}{2} \left(\iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(\varphi^0(\mathbf{x})\Psi_{\Gamma, \mathbf{W}}^i(\mathbf{x}, \mathbf{y}) - \varphi^0(\mathbf{y})\Psi_{\Gamma, \mathbf{W}}^i(\mathbf{y}, \mathbf{x})) \, dyd\mathbf{x} \right) \\
& - \int_{\Omega} \nabla f_{\Gamma}^{\top} \mathbf{W} \psi^0 + f_{\Gamma} \psi^0 \operatorname{div} \mathbf{W} \, d\mathbf{x} + \int_{\Omega} (u^0 - \bar{u}) \varphi^0 \operatorname{div} \mathbf{W} - \nabla \bar{u}^{\top} \mathbf{W} \varphi^0 \, d\mathbf{x}
\end{aligned}$$

and where the second term is the shape derivative at Γ in direction $(DV)\mathbf{W}$, which can be written as

$$\begin{aligned}
DJ(\Gamma)[(DV)\mathbf{W}] &= \int_{\Omega} - \left(u^0 - \bar{u} \right) \nabla \bar{u}^{\top} (DV)\mathbf{W} + \frac{1}{2} (u^0 - \bar{u})^2 \operatorname{div}((DV)\mathbf{W}) \, d\mathbf{x} \\
&+ \int_{\Omega} \nabla f_{\Gamma}^{\top} (DV)\mathbf{W} v^0 + f_{\Gamma} v^0 \operatorname{div}((DV)\mathbf{W}) \, d\mathbf{x} \\
&+ \sum_{i=1}^2 \frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(u^0(\mathbf{x})T_{der}^i(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y})T_{der}^i(\mathbf{y}, \mathbf{x})) \, dyd\mathbf{x}.
\end{aligned}$$

We now continue by putting the derived formula for $(J^{red})''(\Gamma)[\mathbf{V}, \mathbf{W}]$ into numerical practice.

7 Second order shape optimization algorithm

Second-order shape optimization methods have already been introduced and investigated in, e.g., [30–32, 41, 42].

Here, we present a second order shape optimization algorithm where we follow a Newton-like approach that incorporates first and second shape derivatives of the reduced objective functional, which are derived by the averaged adjoint method.

As mentioned before, we follow [31] and optimize on a suitable subset of $\{T|T : \bar{\Omega} \rightarrow \bar{\Omega}\}$. Therefore, we omit the possibly nonsymmetric part of the second shape derivative and only utilize $(J^{red})''(\Gamma)[\mathbf{V}, \mathbf{W}]$. Here, the objective functional also includes the perimeter regularization j_{per} , whose shape derivatives can be found in Remark 4. Moreover, we apply the finite element method and approximate functions by a linear combination of continuous, piecewise linear basis functions. In particular, in every iteration k we will search for a deformation $\mathbf{W}_{\epsilon}^{k,h}$ in the space of such linear combinations of d -dimensional CG ansatz functions, which we denote as CG^d . For a detailed description of how the finite element framework can be implemented in order to solve nonlocal Dirichlet problems, we refer to [43].

However, if $\mathbf{V} = 0$ on Γ , we derive $\mathbf{F}_{\Gamma}^{\mathbf{V}}(\Gamma) = \Gamma$ for $t \in [0, T]$ and, consequently, $DJ^{red}(\Gamma)[\mathbf{V}] = 0$. Therefore, deformations inside of Ω_1 and Ω_2 do not result in a change of the objective functional value. Additionally, $(J^{red})''(\Gamma)[\mathbf{V}, \mathbf{W}] = 0$ if $\mathbf{V} = 0$ on Γ or $\mathbf{W} = 0$ on Γ (see (4)) such that the linear second shape derivative $(J^{red})''(\Gamma)[\mathbf{V}, \mathbf{W}]$ has a huge null space, is not invertible and thus the corresponding stiffness matrix in the finite element method is not positive definite. As a remedy, we

set

$$j_{reg}(\mathbf{V}, \mathbf{W}) := \int_{\Omega} \langle \mathbf{V}, \mathbf{W} \rangle_2 + \langle D\mathbf{V}, D\mathbf{W} \rangle_F d\mathbf{x}$$

and apply a Newton-like approach by solving the following problem in every iteration for a given constant $\epsilon > 0$:

Find $\mathbf{W}_{\epsilon}^{k,h} \in CG^d$ subject to

$$(J^{red})''(\Gamma^k)[\mathbf{V}^h, \mathbf{W}_{\epsilon}^{k,h}] + \epsilon j_{reg}(\mathbf{V}^h, \mathbf{W}_{\epsilon}^{k,h}) = -DJ^{red}(\Gamma^k)[\mathbf{V}^h] \text{ for all } \mathbf{V}^h \in CG^d. \quad (19)$$

Then, there exists a unique solution to problem (19) and the associated finite element stiffness matrix is positive definite. As shown in [31, Theorem 3.5], if $\epsilon \searrow 0$ the corresponding solutions $\mathbf{W}_{\epsilon}^{k,h}$ converge to a solution that is derived by solving the unperturbed problem (19) with $\epsilon = 0$ by a Newton method based on the Moore-Penrose pseudoinverse.

The complete optimization procedure is illustrated in Algorithm 1. Recall that in order to compute $DJ^{red}(\Gamma^k)[\mathbf{V}^h]$ and $(J^{red})''(\Gamma^k)[\mathbf{V}^h, \mathbf{W}^h]$, we need to solve (12) and (13) once in every iteration. However, the solutions $\psi(\Gamma^k, \mathbf{V})$ and $\varphi(\Gamma^k, \mathbf{V})$ to (16) and (17) are calculated for every basis vector field of CG^d .

Algorithm 1 Second Order Shape Optimization Algorithm

while $k \leq \text{maxiter}$ and $\|\mathbf{W}_{\epsilon}^{k,h}\| > \text{tol}$ **do**

Interpolate \bar{u} onto current finite element mesh Ω^k .

$u(\Gamma^k), v(\Gamma^k) \leftarrow$ Solve state and adjoint equation.

Assemble shape derivative $DJ^{red}(\Gamma^k)[\mathbf{V}^h]$ for all basis functions \mathbf{V}^h of CG^d .

Set $DJ^{red}(\Gamma^k)[\mathbf{V}^h] = 0$ if $\text{supp}(\mathbf{V}^h) \cap \Gamma^k = \emptyset$.

$\psi(\Gamma^k, \mathbf{V}^h), \varphi(\Gamma^k, \mathbf{V}^h) \leftarrow$ Solve (16) and (17) for all basis functions \mathbf{V}^h of CG^d .

Assemble linear second shape derivative $(J^{reg})''(\Gamma^k)[\mathbf{V}^h, \mathbf{W}^h]$ for all basis functions $\mathbf{V}^h, \mathbf{W}^h$ of CG^d .

$\mathbf{W}_{\epsilon}^{k,h} \leftarrow$ Get deformation by solving

$$(J^{red})''(\Gamma^k)[\mathbf{V}^h, \mathbf{W}_{\epsilon}^{k,h}] + \epsilon j_{reg}(\mathbf{V}^h, \mathbf{W}_{\epsilon}^{k,h}) = -DJ^{red}(\Gamma^k)[\mathbf{V}^h] \\ \text{for all } \mathbf{V}^h \in CG^d.$$

Update mesh:

$$\Omega^{k+1} = (\mathbf{Id} + \mathbf{W}_{\epsilon}^{k,h})(\Omega^k)$$

$k = k + 1$

end while

8 Numerical experiments

We now describe the implementation of the optimization procedure introduced in the previous Sect. 7 and discuss two numerical examples.

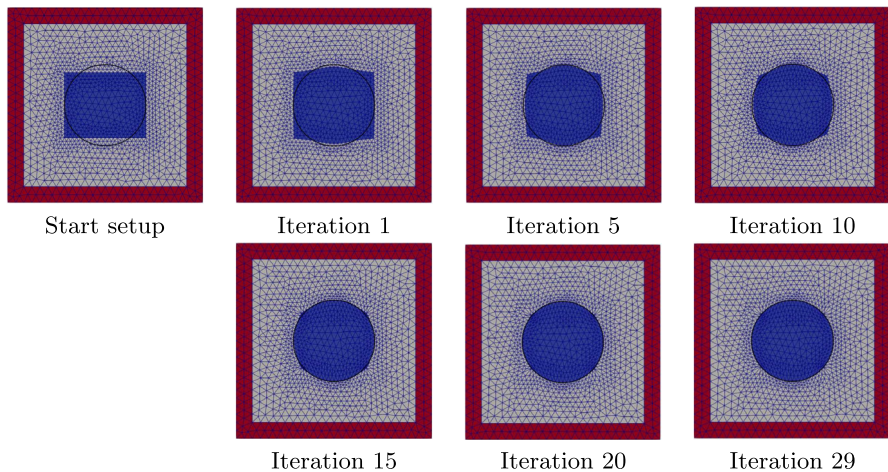


Fig. 2 Example 1: Here we can see snapshots of the domains in different stages of the algorithm, where the domain Ω is decomposed in the blue area Ω_1 and the gray region Ω_2 . The nonlocal boundary is colored in red. Additionally, the black circle can be interpreted as the target shape, i.e., we solve the nonlocal Dirichlet problem given the integrable kernel γ_1 , the parameters chosen as described in the text and the decomposition of Ω in Ω_1 and Ω_2 according to the black circle. The solution is then used as the data \bar{u} in the interface identification problem where we start with the blue square as the initial shape of Ω_1 . After 29 iterations the algorithm ends successfully.

As mentioned before, we apply the finite element method. For our experiments, the underlying mesh is generated by Gmsh [44] and we assemble the stiffness matrix for the bilinear nonlocal form A_Γ with the Python package nlfem [45]. Moreover, the first and second derivatives of the nonlocal bilinear form A_Γ , which are components of the first or the linear second shape derivative, are built by a customized version of nlfem. The regularization function, the forcing term vector and the objective functional along with their derivative are constructed by legacy FEniCS [46, 47].

For the first example, we choose an integrable kernel γ_1 as follows

$$\gamma_1(\mathbf{x}, \mathbf{y}) = \begin{cases} 0.1c_1\chi_{B_\delta(\mathbf{x})}(\mathbf{y}) & \text{if } \mathbf{x}, \mathbf{y} \in \Omega_1, \\ 10c_1\chi_{B_\delta(\mathbf{x})}(\mathbf{y}) & \text{if } \mathbf{x}, \mathbf{y} \in \Omega_2, \\ c_1\chi_{B_\delta(\mathbf{x})}(\mathbf{y}) & \text{else,} \end{cases}$$

where $c_1 := \frac{1}{\delta^4}$. Additionally, we set $\Omega = (0, 1)^2$, $\delta = 0.1$, $\epsilon = 0.3$, $\alpha = 2 \cdot 10^{-3}$, $g = 0$ as boundary data and $f_\Gamma = 10\chi_{\Omega_1} - 10\chi_{\Omega_2}$ in the forcing term. The set-up and the results are shown in Fig. 2. In our experiments, the data \bar{u} , which we would like to approximate as well as possible, is computed by solving the nonlocal Dirichlet problem where Γ is the black circle that can be seen in all the pictures, i.e., Ω_1 is inside and Ω_2 outside of Γ . Then, we start in a set-up where Ω_1 is a square (see Fig. 2) and ideally this blue area should be deformed to a blue disk in the course of the shape optimization algorithm. We stop after 50 iterations or if the $L^2(\Omega, \mathbb{R}^2)$ -norm of the deformation $\mathbf{W}_\epsilon^{k,h}$ is smaller than $5 \cdot 10^{-5}$, which we again compute by using

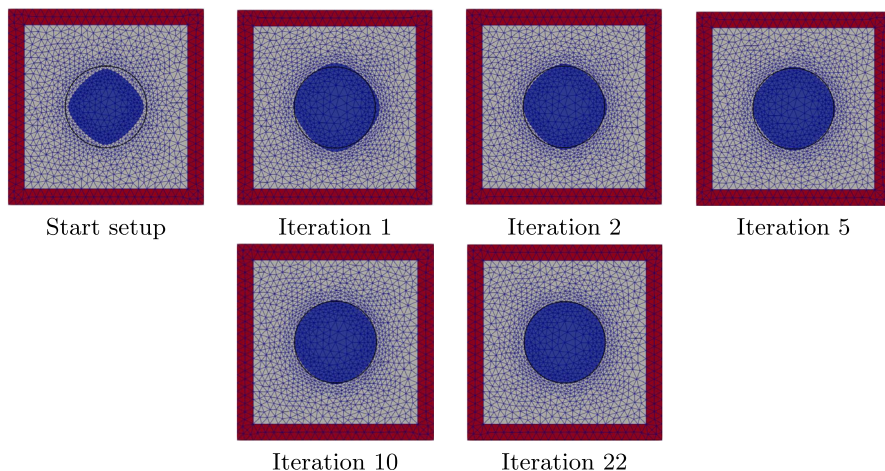


Fig. 3 Example 2: In this picture, we can see snapshots of the domains at different points of the algorithm. Here, Ω_1 is the blue and Ω_2 is the gray area. The nonlocal boundary \mathcal{I} is depicted in red. The black circle is the shape that we try to reach. The given data \bar{u} is the solution of the nonlocal Dirichlet problem where the interface Γ is the black circle and where we apply the singular symmetric kernel γ_2 . Note that the algorithm terminates successfully after 22 iterations.

FEniCS. As we can observe in Fig. 2, we derive roughly a circle after 15 iterations and the algorithm terminates after 29 iterations.

In the second example, we apply the singular symmetric kernel

$$\gamma_2(\mathbf{x}, \mathbf{y}) = \begin{cases} 10.0 \frac{c_2}{\|\mathbf{x} - \mathbf{y}\|_2^{2+2s}} \chi_{B_\delta(\mathbf{x})}(\mathbf{y}) & \text{if } \mathbf{x}, \mathbf{y} \in \Omega_1 \\ 1.0 \frac{c_2}{\|\mathbf{x} - \mathbf{y}\|_2^{2+2s}} \chi_{B_\delta(\mathbf{x})}(\mathbf{y}) & \text{if } \mathbf{x}, \mathbf{y} \in \Omega_2 \\ 5.0 \frac{c_2}{\|\mathbf{x} - \mathbf{y}\|_2^{2+2s}} \chi_{B_\delta(\mathbf{x})}(\mathbf{y}) & \text{else,} \end{cases}$$

where $c_2 := \frac{2-2s}{\pi \delta^{2+2s}}$ and $s = 0.6$. Further, we set $\epsilon = 0.02$, $\alpha = 2 \cdot 10^{-4}$ and $\delta = 0.1$. Again we have $f_\Gamma = 10\chi_{\Omega_1} - 10\chi_{\Omega_2}$ in the forcing term and $g = 0$ on the boundary. The termination criteria as well as the target shape are also the same as in the first experiment. The results are shown in Fig. 3, where we now start from a different shape Γ . In this case, Ω_1 has roughly the shape of the black circle after 5 iterations and the algorithm ends after 22 iterations.

The development of the objective functional values are illustrated in Fig. 4 and the history of the deformation $\mathbf{W}_\epsilon^{k,h}$ in the $L^2(\Omega, \mathbb{R}^2)$ -norm is depicted in Fig. 5. Here, we also tested the integrable kernel γ_1 with the same parameters as in the first example on the second set-up. Analogously, we also used the singular kernel γ_2 with the presented parameters on the starting domain of the first experiment.

We would like to point out that Algorithm 1 is computationally costly since $\psi(\Gamma, \mathbf{V}^h)$ and $\varphi(\Gamma, \mathbf{V}^h)$ are calculated for all basis functions $\mathbf{V}^h \in CG^d$. Thus, we recommend using L-BFGS as in [18]. However, there are certainly some ways

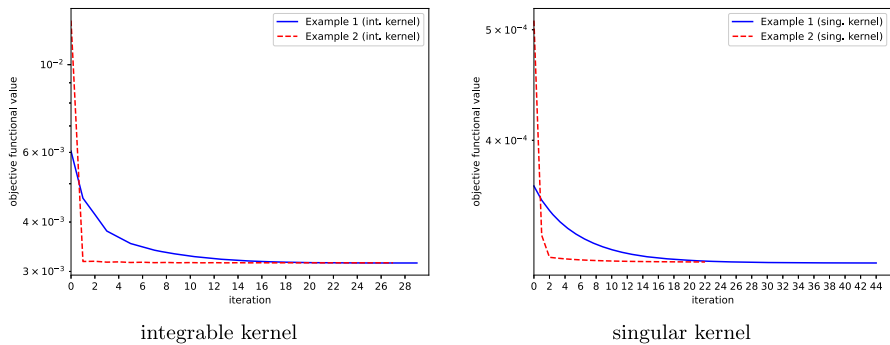
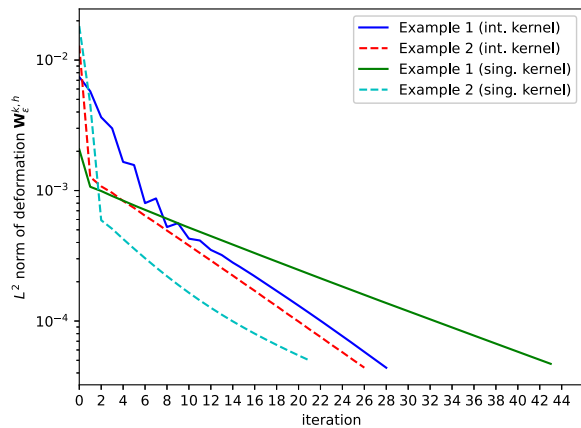


Fig. 4 Development of the objective functional values for both kernels starting on both initial shapes.

Fig. 5 Development of $\|\mathbf{W}_\epsilon^{k,h}\|_2$ during the experiments.



to improve and speed up Algorithm 1. Moreover, note that in our experiments, no remeshing techniques were employed.

9 Conclusion

In this work, we investigated interface identification governed by a nonlocal Dirichlet problem. Further, we developed the (linear) second shape derivative of the reduced functional, which corresponds to this constrained interface identification problem, by applying a version of the averaged adjoint method. The linear second shape derivative was then used in a Newton-like second-order shape optimization algorithm, which was successfully applied to two numerical examples in the previous section. Thus, to the best knowledge of the authors, this is the first study of the numerical usage of linear second shape derivatives for nonlocal problems.

Appendix A: Proofs

Here, we first list some useful lemmata and remarks that are utilized in the proofs of Lemmata 5, 7 and 8.

Lemma 10 (After [24, Proposition 2.32]) *Suppose that $\mathbf{V} \in C_0^1(\Omega, \mathbb{R}^d)$ and $f \in W^{1,k}(\Omega, \mathbb{R})$ with $k \in \mathbb{N}$. Moreover, let $T > 0$ be sufficiently small such that the perturbation of identity $\mathbf{F}_t^{\mathbf{V}} : \overline{\Omega} \rightarrow \overline{\Omega}$ is bijective for all $t \in [0, T]$. Then, $t \mapsto f \circ \mathbf{F}_t^{\mathbf{V}}$ is differentiable in $L^k(\Omega)$ with*

$$\left(\frac{d}{dt} \Big|_{t=r} \left(f \circ \mathbf{F}_t^{\mathbf{V}} \right) \right) (\mathbf{x}) = \nabla f(\mathbf{F}_r^{\mathbf{V}}(\mathbf{x}))^\top \mathbf{V}(\mathbf{x}) \text{ for } r \in [0, T].$$

Proof See proof of [48, Lemma 3.2.11]. □

Moreover, we make use of the following “product rules” for Fréchet derivatives:

Lemma 11 *The following differentiation rules hold:*

- (i) *Let $f_1 : [0, T] \rightarrow L^2(\Omega)$, $t \mapsto f_1(t, \cdot)$ and $f_2 : [0, T] \rightarrow L^2(\Omega)$, $t \mapsto f_2(t, \cdot)$ be Fréchet differentiable in $L^2(\Omega)$ with derivatives $f_1'(t, \cdot)$ and $f_2'(t, \cdot)$. Then, the product $t \mapsto f_1(t, \cdot) f_2(t, \cdot)$ is Fréchet differentiable in $L^1(\Omega)$ with*

$$(f_1 f_2)'(t, \cdot) = f_1(t, \cdot) f_2'(t, \cdot) + f_1'(t, \cdot) f_2(t, \cdot).$$

- (ii) *Given $k \in \mathbb{N}$, let $f_1 : [0, T] \rightarrow L^k(\Omega)$, $t \mapsto f_1(t, \cdot)$ be Fréchet differentiable in $L^k(\Omega)$ with derivative $f_1'(t, \cdot)$ and let $f_2 : [0, T] \rightarrow L^\infty(\Omega)$, $t \mapsto f_2(t, \cdot)$ be Fréchet differentiable in $L^k(\Omega)$. Then, the product $t \mapsto f_1(t, \cdot) f_2(t, \cdot)$ is Fréchet differentiable in $L^k(\Omega)$ with*

$$(f_1 f_2)'(t, \cdot) = f_1(t, \cdot) f_2'(t, \cdot) + f_1'(t, \cdot) f_2(t, \cdot).$$

Proof See proof of [48, Lemma 3.2.9]. □

Remark 8 Let $f \in W^{1,1}(\Omega)$ and $h \in C_0^1(\Omega)$, then fh is weakly differentiable with derivative $\partial^\alpha(fh) = (\partial^\alpha f)h + f(\partial^\alpha h)$ for a multi-index $\alpha \in \mathbb{N}_0^d$ with $|\alpha| = 1$ since there exists a sequence $\{h_n\}_{n \in \mathbb{N}} \in C^\infty(\overline{\Omega})^\mathbb{N}$ with $h_n \rightarrow h$ uniformly (see [49, Appendix C Theorem 7]) such that for every $\varphi \in C_c^\infty(\Omega)$ we derive

$$\int_{\Omega} (\partial^\alpha f) h_n \varphi \, d\mathbf{x} = - \int_{\Omega} f \partial^\alpha (h_n \varphi) \, d\mathbf{x} = - \int_{\Omega} f (\partial^\alpha h_n) \varphi \, d\mathbf{x} - \int_{\Omega} f h_n (\partial^\alpha \varphi) \, d\mathbf{x}.$$

As a consequence, $\nabla f^\top \mathbf{V}$ and $f \operatorname{div} \mathbf{V}$ are weakly differentiable if $f \in H^2(\Omega)$ and $\mathbf{V} \in C_0^2(\Omega \cup \mathcal{I}, \mathbb{R}^d)$.

In order to show Assumption (D3) we further make use of the following lemma and the subsequent remark in the proofs of Lemmas 7 and 8:

Lemma 12 (After [23, Chapter 10.2.4 Lemma 2.1]) Let $\mathbf{V} \in C_0^1(\Omega, \mathbb{R}^d)$ and $n \in \mathbb{N}$. Then, for $\tilde{f} \in L^2(\Omega, \mathbb{R}^n)$, we get

$$\|\tilde{f} \circ \mathbf{F}_t^V - \tilde{f}\|_{L^2(\Omega, \mathbb{R}^n)} \rightarrow 0 \text{ and } \|\tilde{f} \circ (\mathbf{F}_t^V)^{-1} - \tilde{f}\|_{L^2(\Omega, \mathbb{R}^n)} \rightarrow 0 \text{ for } t \searrow 0.$$

Proof See proof of [48, Lemma 3.2.20]. □

Remark 9 Let a sequence $\{t_n\}_{n \in \mathbb{N}} \in [0, T]^{\mathbb{N}}$ with $\lim_{n \rightarrow \infty} t_n = 0$ be given. Moreover, assume two families of functions $\{f^{t_n}\}_{n \in \mathbb{N}}$ and $\{h^{t_n}\}_{n \in \mathbb{N}}$ with $f^{t_n}, h^{t_n} \in L^2(\Omega)$ such that $f^{t_n} \rightarrow f^0$ and $h^{t_n} \rightarrow h^0$ in $L^2(\Omega)$ for $n \rightarrow \infty$. Further, we suppose that the family $\{h^{t_n}\}_{n \in \mathbb{N}}$ is bounded w.r.t. $\|\cdot\|_{L^2(\Omega)}$. Then, we derive

$$\lim_{n \rightarrow \infty} \int_{\Omega} h^{t_n} f^{t_n} d\mathbf{x} = \int_{\Omega} h^0 f^0 d\mathbf{x}. \quad (20)$$

Of course, if $h^{t_n} \rightarrow h^0$ in $L^2(\Omega)$, then (20) also holds. Now, let $p \in \mathbb{N}$ and two families of functions $\{f^t\}_{t \in [0, T]}$ and $\{h^t\}_{t \in [0, T]}$ with $h \in L^\infty([0, T] \times \Omega)$ and $f^t \in L^p(\Omega)$ for $t \in [0, T]$ be given. Further, we assume $f^t \rightarrow f^0$ and $h^t \rightarrow h^0$ in $L^p(\Omega)$. Then, we obtain

$$f^t h^t \rightarrow f^0 h^0 \text{ in } L^p(\Omega) \text{ for } t \searrow 0. \quad (21)$$

In the following, (21) will mainly be used to show that products involving $\xi^t \in L^\infty(\Omega)$ or $\frac{\partial}{\partial r}|_{r=t}(\xi^r) \in L^\infty(\Omega)$ still converge in $L^1(\Omega)$ or $L^2(\Omega)$. Note that ξ^t converges uniformly to $\xi^0 (= 1)$ and $\frac{\partial}{\partial r}|_{r=t}(\xi^r)$ is also uniform convergent to $\operatorname{div} \mathbf{V}$ for $t \searrow 0$, which follows from the Leibniz rule for determinants and $\mathbf{V} \in C_0^2(\Omega \cup \mathcal{I}, \mathbb{R}^d)$.

A.1 Proof of Lemma 5

Proof For the proof of the differentiability of $A(t, u, \psi)$, $A(t, \varphi, v)$ and $F(t, \psi)$ regarding t , we refer to [18]. It remains to show that $\partial_t \tilde{F}(t, u, \varphi)$, $\partial_t \mathfrak{F}(t, u)$, $\partial_t \mathfrak{F}(t, v)$ and $\partial_t \mathfrak{A}(t, u, v)$ exist. Because $\bar{u} \in H^2(\Omega)$ and $\xi^t = \det D\mathbf{F}_t^{\mathbf{W}}$ is continuously differentiable in t on $\overline{\Omega}$ as shown in, e.g., [26, Lemma 2.14], we get by applying Lemma 10 and Lemma 11 that

$$\partial_t \tilde{F}(t, u, \varphi) = \int_{\Omega} (\nabla \bar{u}^t)^\top \mathbf{W} \varphi \xi^t - (u - \bar{u}^t) \varphi \frac{d}{dr} \Big|_{r=t} (\xi^r) d\mathbf{x}.$$

Additionally, since $f^t \in H^2(\Omega_i)$ for $i = 1, 2$ and $\mathbf{V} \in C_0^2(\Omega \cup \mathcal{I}, \mathbb{R}^d)$, we can conclude by making use of Lemma 10 and by applying Lemma 11 that

$$\begin{aligned}
\partial_t \mathfrak{J}(t, u, v) &= \int_{\Omega} (\nabla \bar{u}^t)^\top \mathbf{W} (\nabla \bar{u}^t)^\top \mathbf{V}^t \xi^t \, d\mathbf{x} \\
&- \int_{\Omega} (u - \bar{u}^t) \left((\mathbf{V}^t)^\top \text{Hess}(\bar{u})^t \mathbf{W} \xi^t + (\nabla \bar{u}^t)^\top D \mathbf{V}^t \mathbf{W} \xi^t + (\nabla \bar{u}^t)^\top \mathbf{V}^t \frac{d}{dr} \Big|_{r=t} (\xi^r) \right) d\mathbf{x} \\
&+ \int_{\Omega} -(u - \bar{u}^t) (\nabla \bar{u}^t)^\top \mathbf{W} \operatorname{div} \mathbf{V}^t \xi^t \\
&\quad + \frac{1}{2} (u - \bar{u}^t)^2 \left((\nabla \operatorname{div} \mathbf{V}^t)^\top \mathbf{W} \xi^t + \operatorname{div} \mathbf{V}^t \frac{d}{dr} \Big|_{r=t} (\xi^r) \right) d\mathbf{x} \text{ and} \\
\partial_t \mathfrak{F}(t, v) &= \int_{\Omega} (\mathbf{V}^t)^\top \text{Hess}(f^t) \mathbf{W} v \xi^t + (\nabla f^t)^\top D \mathbf{V}^t \mathbf{W} v \xi^t + (\nabla f^t)^\top \mathbf{V}^t v \frac{d}{dr} \Big|_{r=t} (\xi^r) \, d\mathbf{x} \\
&+ \int_{\Omega} (\nabla f^t)^\top \mathbf{W} v \operatorname{div} \mathbf{V}^t \xi^t + f^t v (\nabla \operatorname{div} \mathbf{V}^t)^\top \mathbf{W} \xi^t + f^t v \operatorname{div} \mathbf{V}^t \frac{d}{dr} \Big|_{r=t} (\xi^r) \, d\mathbf{x}.
\end{aligned}$$

Case 1: Integrable Kernel

Since we assume $\mathbf{V} \in C_0^2(\Omega \cup \mathcal{I}, \mathbb{R}^d)$ and $\gamma_{\Gamma_t^{\mathbf{W}}} \in W^{2,\infty}(\mathbf{F}_t^{\mathbf{W}}(\Omega_i) \times \mathbf{F}_t^{\mathbf{W}}(\Omega_j))$ and $\gamma_{\Gamma_t^{\mathbf{W}}} \in W^{2,\infty}(\mathbf{F}_t^{\mathbf{W}}(\Omega_i) \times \mathcal{I})$ as a consequence of Assumption (S1), we conclude by employing Remark 8 that $\Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}}^1 \in W^{1,\infty}(\mathbf{F}_t^{\mathbf{W}}(\Omega_i) \times \mathbf{F}_t^{\mathbf{W}}(\Omega_j))$ and $\Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}}^1 \in W^{1,\infty}(\mathbf{F}_t^{\mathbf{W}}(\Omega_i) \times \mathcal{I})$ with $\nabla \Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}}^1 = \text{Hess}(\gamma_{\Gamma_t^{\mathbf{W}}}) \mathbf{V} + (\nabla \gamma_{\Gamma_t^{\mathbf{W}}}^\top D \mathbf{V})^\top$. Similarly, we get the weak derivatives

$$\begin{aligned}
\nabla \Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}}^2(\mathbf{x}, \mathbf{y}) &= (\operatorname{div} \mathbf{V}(\mathbf{x}) + \operatorname{div} \mathbf{V}(\mathbf{y})) \nabla \gamma_{\Gamma_t^{\mathbf{W}}}(\mathbf{x}, \mathbf{y}) \\
&\quad + \gamma_{\Gamma_t^{\mathbf{W}}}(\mathbf{x}, \mathbf{y}) (\nabla \operatorname{div} \mathbf{V}(\mathbf{x}) + \nabla \operatorname{div} \mathbf{V}(\mathbf{y})).
\end{aligned}$$

Recall that $\Psi_k'(\mathbf{x}, \mathbf{y}) = \Psi_{\Gamma_t^{\mathbf{W}}, \mathbf{V}^t}^k(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x}), \mathbf{F}_t^{\mathbf{W}}(\mathbf{y}))$ and thus $t \mapsto \Psi_k'(\mathbf{x}, \mathbf{y})$ is Fréchet differentiable in $L^2(\Omega_i \times \Omega_j)$ and in $L^2(\Omega_i \times \mathcal{I})$ due to 10 for $i, j, k = 1, 2$ with essentially bounded derivatives. Consequently, we derive by utilizing Lemma 11 that the partial derivative regarding the nonlocal operator can be expressed as

$$\begin{aligned}
\partial_t \mathfrak{A}(t, u, v) &= \frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y})) \left(u(\mathbf{x}) \nabla \Psi_1^t(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \right. \\
&\quad \left. - u(\mathbf{y}) \nabla \Psi_1^t(\mathbf{y}, \mathbf{x})^\top \mathbf{W}(\mathbf{y}, \mathbf{x}) \right) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \\
&\quad + (v(\mathbf{x}) - v(\mathbf{y})) (u(\mathbf{x}) \Psi_1^t(\mathbf{x}, \mathbf{y}) - u(\mathbf{y}) \Psi_1^t(\mathbf{y}, \mathbf{x})) \frac{d}{dr} \Big|_{r=t} (\xi^r(\mathbf{x}) \xi^r(\mathbf{y})) \, d\mathbf{y} d\mathbf{x} \\
&\quad + \frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y})) \left(u(\mathbf{x}) \nabla \Psi_2^t(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \right. \\
&\quad \left. - u(\mathbf{y}) \nabla \Psi_2^t(\mathbf{y}, \mathbf{x})^\top \mathbf{W}(\mathbf{y}, \mathbf{x}) \right) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \\
&\quad + (v(\mathbf{x}) - v(\mathbf{y})) (u(\mathbf{x}) \Psi_2^t(\mathbf{x}, \mathbf{y}) - u(\mathbf{y}) \Psi_2^t(\mathbf{y}, \mathbf{x})) \frac{d}{dr} \Big|_{r=t} (\xi^r(\mathbf{x}) \xi^r(\mathbf{y})) \, d\mathbf{y} d\mathbf{x}.
\end{aligned}$$

Case 2: Singular Symmetric Kernel

We start by proving Fréchet differentiability of terms involving the subkernels γ_{ij} and $\gamma_{i\mathcal{I}}$. Therefore, for $i, j = 1, 2$ we set

$$\begin{aligned} h_{ij}^1(\mathbf{x}, \mathbf{y}) &:= \nabla \gamma_{ij}(\mathbf{x}, \mathbf{y})^\top \mathbf{V}(\mathbf{x}, \mathbf{y}), \quad h_{i\mathcal{I}}^1(\mathbf{x}, \mathbf{y}) := \nabla \gamma_{i\mathcal{I}}(\mathbf{x}, \mathbf{y})^\top \mathbf{V}(\mathbf{x}, \mathbf{y}), \\ h_{ij}^2(\mathbf{x}, \mathbf{y}) &:= (\operatorname{div} \mathbf{V}(\mathbf{x}) + \operatorname{div} \mathbf{V}(\mathbf{y})) \gamma_{ij}(\mathbf{x}, \mathbf{y}) \text{ and} \\ h_{i\mathcal{I}}^2(\mathbf{x}, \mathbf{y}) &:= (\operatorname{div} \mathbf{V}(\mathbf{x}) + \operatorname{div} \mathbf{V}(\mathbf{y})) \gamma_{i\mathcal{I}}(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Note, that for $k = 1, 2$, we can now express

$$\Psi_k^t(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1,2} h_{ij}^k(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x}), \mathbf{F}_t^{\mathbf{W}}(\mathbf{y})) \chi_{\Omega_i \times \Omega_j}(\mathbf{x}, \mathbf{y}) + \sum_{i=1,2} h_{i\mathcal{I}}^k(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x}), \mathbf{y}) \chi_{\Omega_i \times \mathcal{I}}(\mathbf{x}, \mathbf{y}).$$

Stein's extension theorem [50, Theorem 5.24] combined with Remark 8 yields $h_{ij}^1, h_{ij}^2 \in W^{1,p}(D_n)$ for $n \in \mathbb{N}$ with weak derivatives

$$\begin{aligned} \nabla h_{ij}^1(\mathbf{x}, \mathbf{y}) &= \operatorname{Hess}(\gamma_{ij})(\mathbf{x}, \mathbf{y}) \mathbf{V}(\mathbf{x}, \mathbf{y}) + (\nabla \gamma_{ij}^\top(\mathbf{x}, \mathbf{y}) D\mathbf{V}(\mathbf{x}, \mathbf{y}))^\top \text{ and} \\ \nabla h_{ij}^2(\mathbf{x}, \mathbf{y}) &= (\operatorname{div} \mathbf{V}(\mathbf{x}) + \operatorname{div} \mathbf{V}(\mathbf{y})) \nabla \gamma_{ij}(\mathbf{x}, \mathbf{y}) + \gamma_{ij}(\mathbf{x}, \mathbf{y}) (\nabla \operatorname{div} \mathbf{V}(\mathbf{x}) + \nabla \operatorname{div} \mathbf{V}(\mathbf{y})). \end{aligned}$$

For every $n \in \mathbb{N}$, it can be shown that there exists an integer $n_1 \in \mathbb{N}$ such that $(\mathbf{F}_t^{\mathbf{W}}(\mathbf{x}), \mathbf{F}_t^{\mathbf{W}}(\mathbf{y})) \in D_{n_1}$ for all $(\mathbf{x}, \mathbf{y}) \in D_n$ and $t \in [0, T]$. Further, we set $\tilde{\mathbf{F}}_t^{\mathbf{W}}(\mathbf{x}, \mathbf{y}) := (\mathbf{F}_t^{\mathbf{W}}(\mathbf{x}), \mathbf{F}_t^{\mathbf{W}}(\mathbf{y}))$. Then, by utilizing Stein's extension theorem [50, Theorem 5.24] and Lemma 10 the function $[0, T] \ni t \mapsto h_{ij}^k \circ \tilde{\mathbf{F}}_t^{\mathbf{W}}$ is Fréchet differentiable in $L^p(D_n)$. Analogously, it can be shown that $t \mapsto h_{i\mathcal{I}}^k \circ \tilde{\mathbf{F}}_t^{\mathbf{W}}$ is Fréchet differentiable in $L^p(D_n)$. Thus, we obtain the Fréchet differentiability of $t \mapsto \Psi_k^t(\mathbf{x}, \mathbf{y})$ in $L^2(D_n \cap (\Omega_i \times \Omega_j))$ and $L^2(D_n \cap (\Omega_i \times \mathcal{I}))$ for $i, j, k = 1, 2$. Note that these Fréchet derivatives are again essentially bounded on the subdomains $[0, T] \times (D_n \cap (\Omega_i \times \Omega_j))$ and $[0, T] \times (D_n \cap (\Omega_i \times \mathcal{I}))$ as illustrated in Remark 7. By making use of Lemma 11 multiple times and due to Assumption (S1), which also ensures the well-posedness of the integrals, we get the differentiability of \mathfrak{A} regarding t with

$$\begin{aligned} & \left. \frac{\partial}{\partial r} \right|_{r=t} \mathfrak{A}(t, u, v) \\ &= \lim_{n \rightarrow \infty} \left. \frac{d}{dr} \right|_{r=t} \frac{1}{2} \iint_{D_n} (v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x}) - u(\mathbf{y})) \Psi_1^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \, d\mathbf{y} d\mathbf{x} \\ &+ \lim_{n \rightarrow \infty} \left. \frac{d}{dr} \right|_{r=t} \frac{1}{2} \iint_{D_n} (v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x}) - u(\mathbf{y})) \Psi_2^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \, d\mathbf{y} d\mathbf{x} \\ &= \frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x}) - u(\mathbf{y})) \nabla \Psi_1^t(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \\ &+ (v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x}) - u(\mathbf{y})) \Psi_1^t(\mathbf{x}, \mathbf{y}) \left. \frac{\partial}{\partial r} \right|_{r=t} (\xi^r(\mathbf{x}) \xi^r(\mathbf{y})) \, d\mathbf{y} d\mathbf{x} \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x}) - u(\mathbf{y})) \nabla \Psi_2^t(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \\
& + (v(\mathbf{x}) - v(\mathbf{y}))(u(\mathbf{x}) - u(\mathbf{y})) \Psi_2^t(\mathbf{x}, \mathbf{y}) \frac{\partial}{\partial r} \Big|_{r=t} (\xi^r(\mathbf{x}) \xi^r(\mathbf{y})) \, d\mathbf{y} d\mathbf{x}.
\end{aligned}$$

□

A.2 Proof of Lemma 7

Proof The proof for both assertions regarding $\{u^t\}_{t \in [0, T]}$ as well as $\{v^t\}_{t \in [0, T]}$ can be found in [18]. We follow the same approach in order to show the statements for $\{\varphi^t\}_{t \in [0, T]}$ and $\{\psi^t\}_{t \in [0, T]}$. Thus, for both kernel classes, we first prove the boundedness with respect to the above-mentioned norm and then use these results to show the corresponding weak convergences.

For ease of presentation, we denote every subsequence of a sequence $(t_n)_{n \in \mathbb{N}} \in [0, T]^\mathbb{N}$ by $(t_{n_k})_{k \in \mathbb{N}}$ in the remaining part of this proof. However, technically they could be different subsequences or a subsequence of a subsequence.

Case 1: Integrable Kernel

First, let γ_Γ be an integrable kernel that additionally satisfies Assumptions (S0) and (S1). Then, we derive the existence of constants $C_1, C_2 > 0$ with

$$\begin{aligned}
& \|\varphi^t\|_{L^2(\Omega \cup \mathcal{I})}^2 \leq C_1 |A(t, \varphi^t, \varphi^t)| = C_1 |\mathfrak{F}(t, \varphi^t) - \mathfrak{A}(t, u^0, \varphi^t)| \\
& \leq C_2 \left(\sum_{i=1,2} (\|f^t\|_{L^2(\Omega_i)} + \|\nabla f^t\|_{L^2(\Omega_i)}) + \|u^0\|_{L^2(\Omega)} \right) \|\varphi^t\|_{L^2(\Omega \cup \mathcal{I})} \text{ as follows:}
\end{aligned}$$

We apply in the first step Assumption (S0) (2) and in the second that φ^t solves the averaged adjoint equation (17). Additionally, we used that there exists a $\widehat{C} > 0$ with $|\mathfrak{A}(t, u^0, \varphi^t)| \leq \widehat{C} \|u^0\|_{L^2(\Omega \cup \mathcal{I})} \|\varphi^t\|_{L^2(\Omega \cup \mathcal{I})}$ (see Lemma 6) and $|\mathfrak{F}(t, \varphi^t)| \leq \widehat{C} \sum_{i=1,2} (\|f^t\|_{L^2(\Omega_i)} + \|\nabla f^t\|_{L^2(\Omega_i)}) \|\varphi^t\|_{L^2(\Omega \cup \mathcal{I})}$ due to the boundedness of ξ^t as well as $\mathbf{V}^t \in C_0^2(\Omega \cup \mathcal{I}, \mathbb{R}^d)$ and by utilizing the Cauchy-Schwarz inequality for L^2 -functions. Combined with $f^t \rightarrow f^0$ in $L^2(\Omega_i)$ and $\nabla f^t \rightarrow \nabla f^0$ in $L^2(\Omega_i, \mathbb{R}^d)$ due to Lemma 12, we get $\|f^t\|_{L^2(\Omega_i)}, \|\nabla f^t\|_{L^2(\Omega_i)} < \widetilde{C}$ for $i = 1, 2$ for some constant $\widetilde{C} > 0$. Therefore, we can conclude that $\{\|\varphi^t\|_{L^2(\Omega \cup \mathcal{I})}\}_{t \in [0, T]}$ is bounded.

The boundedness of $\{\|\psi^t\|_{L^2(\Omega \cup \mathcal{I})}\}_{t \in [0, T]}$ can be proven analogously. Consequently, we obtain for every sequence $(t_n)_{n \in \mathbb{N}} \in [0, T]^\mathbb{N}$ with $\lim_{n \rightarrow \infty} t_n = 0$ the existence of a subsequence $\{t_{n_k}\}_{k \in \mathbb{N}}$ and functions $q_1, q_2 \in L_c^2(\Omega \cup \mathcal{I})$ with $\varphi^{t_{n_k}} \rightharpoonup q_1$ and $\psi^{t_{n_k}} \rightharpoonup q_2$ for $k \rightarrow \infty$. We now show that $q_1 = \varphi^0$ and $q_2 = \psi^0$. Therefore, we need several convergences (see (22)). For the proof of $A(t_{n_k}, \varphi^{t_{n_k}}, \widetilde{v}) \rightarrow A(0, q_1, \widetilde{v})$ and $A(t_{n_k}, \widetilde{u}, \psi^{t_{n_k}}) \rightarrow A(0, \widetilde{u}, q_2)$ for $k \rightarrow \infty$ we refer to [18]. Here, we show the remaining convergences by making use of similar arguments.

Since $((\nabla f^t)^\top \mathbf{V}^t + \operatorname{div} \mathbf{V}^t f^t) \xi^t \rightarrow \nabla f_\Gamma^\top \mathbf{V} + \operatorname{div} \mathbf{V} f_\Gamma$ in $L^2(\Omega)$ due to Lemma 12 and observation (21), we obtain by utilizing (20) that

$$\begin{aligned} \mathfrak{F}(t_{n_k}, \tilde{v}) &= \int_{\Omega} \left((\nabla f^{t_{n_k}})^\top \mathbf{V}^{t_{n_k}} \xi^{t_{n_k}} + \operatorname{div} \mathbf{V}^{t_{n_k}} f^{t_{n_k}} \xi^{t_{n_k}} \right) \tilde{v} \, d\mathbf{x} \\ &\rightarrow \int_{\Omega} \left(\nabla f_\Gamma^\top \mathbf{V} + \operatorname{div} \mathbf{V} f_\Gamma \right) \tilde{v} \, d\mathbf{x} = \mathfrak{F}(0, \tilde{v}) \text{ for } k \rightarrow \infty. \end{aligned}$$

We now prove $\lim_{k \rightarrow \infty} \mathfrak{A}(t_{n_k}, u^0, \tilde{v}) = \mathfrak{A}(0, u^0, \tilde{v})$. First, we recall that γ^t and $\nabla \gamma^t$ are essentially bounded by a constant independent of $t \in [0, T]$. Hence, from Lemma 12 and observation (21) follows

$$\begin{aligned} \Psi_1^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) &= \nabla \gamma^t(\mathbf{x}, \mathbf{y})^\top \mathbf{V}^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \\ &\rightarrow \nabla \gamma^0(\mathbf{x}, \mathbf{y})^\top \mathbf{V}^0(\mathbf{x}, \mathbf{y}) = \Psi_1^0(\mathbf{x}, \mathbf{y}) \end{aligned}$$

in $L^2((\Omega \cup \mathcal{I})^2)$ for $t \searrow 0$. Then, for every sequence $(t_n)_{n \in \mathbb{N}}[0, T]^\mathbb{N}$ with $\lim_{n \rightarrow \infty} t_n = 0$ there exists a subsequence $(t_{n_k})_{k \in \mathbb{N}}$ such that $\Psi_1^{t_{n_k}} \rightarrow \Psi_1^0$ almost everywhere on $(\Omega \cup \mathcal{I})^2$. As a consequence of the dominated convergence theorem, we get

$$\begin{aligned} &\frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (\tilde{v}(\mathbf{x}) - \tilde{v}(\mathbf{y})) (u^0(\mathbf{x}) \Psi_1^{t_{n_k}}(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y}) \Psi_1^{t_{n_k}}(\mathbf{y}, \mathbf{x})) \xi^{t_{n_k}}(\mathbf{x}) \xi^{t_{n_k}}(\mathbf{y}) \, d\mathbf{y} d\mathbf{x} \\ &\quad * \rightarrow \frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (\tilde{v}(\mathbf{x}) - \tilde{v}(\mathbf{y})) (u^0(\mathbf{x}) \Psi_1^0(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y}) \Psi_1^0(\mathbf{y}, \mathbf{x})) \, d\mathbf{y} d\mathbf{x}. \end{aligned}$$

Moreover, Lemma 12 and (21) yield

$$\begin{aligned} \Psi_2^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) &= \gamma^t(\mathbf{x}, \mathbf{y}) (\operatorname{div} \mathbf{V}^t(\mathbf{x}) + \operatorname{div} \mathbf{V}^t(\mathbf{y})) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \\ &\rightarrow \gamma^0(\mathbf{x}, \mathbf{y}) (\operatorname{div} \mathbf{V}^0(\mathbf{x}) + \operatorname{div} \mathbf{V}^0(\mathbf{y})) = \Psi_2^0(\mathbf{x}, \mathbf{y}) \text{ in } L^2((\Omega \cup \mathcal{I})^2). \end{aligned}$$

By making use of the same argumentation as before, we conclude

$$\begin{aligned} &\frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (\tilde{v}(\mathbf{x}) - \tilde{v}(\mathbf{y})) (u^0(\mathbf{x}) \Psi_2^{t_{n_k}}(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y}) \Psi_2^{t_{n_k}}(\mathbf{y}, \mathbf{x})) \xi^{t_{n_k}}(\mathbf{x}) \xi^{t_{n_k}}(\mathbf{y}) \, d\mathbf{y} d\mathbf{x} \\ &\quad \rightarrow \frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (\tilde{v}(\mathbf{x}) - \tilde{v}(\mathbf{y})) (u^0(\mathbf{x}) \Psi_2^0(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y}) \Psi_2^0(\mathbf{y}, \mathbf{x})) \, d\mathbf{y} d\mathbf{x}. \end{aligned}$$

Consequently, we get $\lim_{k \rightarrow \infty} \mathfrak{A}(t_{n_k}, u^0, \tilde{v}) = \mathfrak{A}(0, u^0, \tilde{v})$ and therefore $q_1 = \varphi^0$ since the function q_1 is the unique solution to the averaged adjoint equation (15) for

$t = 0$ due to

$$\begin{aligned} A(0, q_1, \tilde{v}) &= \lim_{k \rightarrow \infty} A(t_{n_k}, \varphi^{t_{n_k}}, \tilde{v}) = \lim_{k \rightarrow \infty} \mathfrak{F}(t_{n_k}, \tilde{v}) - \mathfrak{A}(t_{n_k}, u^{t_{n_k}}, \tilde{v}) \\ &= \mathfrak{F}(0, \tilde{v}) - \mathfrak{A}(0, u^0, \tilde{v}). \end{aligned} \quad (22)$$

Thus, we conclude $\varphi^t \rightarrow \varphi^0$ for $t \searrow 0$. **Case 2: Singular Symmetric Kernel**

If γ_Γ is a singular symmetric kernel that satisfies Assumptions (S0) and (S1), we similarly derive for some $C_1, C_2, C_3 > 0$ that

$$\begin{aligned} |\varphi^t|_{H^s(\Omega \cup \mathcal{I})}^2 &\leq C_1 |A(t, \varphi^t, \varphi)| \\ &\leq C_2 \left(\sum_{i=1,2} \left(\|f^t\|_{L^2(\Omega_i)} + \|\nabla f^t\|_{L^2(\Omega_i)} \right) + |u^0|_{H^s(\Omega \cup \mathcal{I})} \right) |\varphi^t|_{H^s(\Omega \cup \mathcal{I})} \leq C_3 |\varphi^t|_{H^s(\Omega \cup \mathcal{I})} \end{aligned}$$

where we again used Assumption (S0) (2) for the first and Lemma 6 in the second step. Thus, $|\varphi^t|_{H^s(\Omega \cup \mathcal{I})} \leq C_3$ for all $t \in [0, T]$. As a result of [21, Lemma 4.2 and Lemma 4.3], we also draw the conclusion $\|\varphi^t\|_{L^2(\Omega \cup \mathcal{I})} \leq \tilde{C} |\varphi^t|_{H^s(\Omega \cup \mathcal{I})} \leq C$ for appropriate constants $\tilde{C}, C > 0$, i.e., the family $\{\varphi^t\}_{t \in [0, T]}$ is bounded with respect to $\|\cdot\|_{H^s(\Omega \cup \mathcal{I})}$.

Note that the boundedness of $\{\|\psi^t\|_{H^s(\Omega \cup \mathcal{I})}\}_{t \in [0, T]}$ can be shown analogously.

Consequently, we derive for every sequence $(t_n)_{n \in \mathbb{N}} \in [0, T]^{\mathbb{N}}$ with $\lim_{n \rightarrow \infty} t_n = 0$ a subsequence $\{t_{n_k}\}_{k \in \mathbb{N}}$ and functions $q_1, q_2 \in H_c^s(\Omega \cup \mathcal{I})$ with $\varphi^{t_{n_k}} \rightarrow q_1$ and $\psi^{t_{n_k}} \rightarrow q_2$ for $k \rightarrow \infty$. As in Case 1, we need to show some convergences.

For the proof of $A(t_{n_k}, \varphi^{t_{n_k}}, \tilde{v}) \rightarrow A(0, q_1, \tilde{v})$ and $A(t_{n_k}, \tilde{u}, \psi^{t_{n_k}}) \rightarrow A(0, \tilde{u}, q_2)$ for $k \rightarrow \infty$ we again refer to [18]. The convergence $\lim_{k \rightarrow \infty} \mathfrak{F}(t_{n_k}, \tilde{v}) = \mathfrak{F}(0, \tilde{v})$ can be found in the proof of Case 1. In the following, we prove $\lim_{k \rightarrow \infty} \mathfrak{A}(t_{n_k}, u^0, \tilde{v}) = \mathfrak{A}(0, u^0, \tilde{v})$.

On every set D_n the kernel γ^t and the corresponding gradient $(\nabla \gamma^t)^\top \mathbf{V}$ are essentially bounded by a constant independent of $t \in [0, T]$ due to Assumption (S1) as explained in Remark 7. Consequently, there exists a constant $\widehat{C} > 0$ such that for $t \in [0, T]$ and almost every $(\mathbf{x}, \mathbf{y}) \in D_n$ we obtain

$$\begin{aligned} \Psi_1^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) &= (\nabla \gamma^t(\mathbf{x}, \mathbf{y}))^\top \mathbf{V}^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) < \widehat{C} \text{ and} \\ \Psi_2^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) &= \gamma^t(\mathbf{x}, \mathbf{y}) (\operatorname{div} \mathbf{V}^t(\mathbf{x}) + \operatorname{div} \mathbf{V}^t(\mathbf{y})) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) < \widehat{C}. \end{aligned}$$

With the same argumentation as in the case of integrable kernels, we derive for every sequence $(t_n)_{n \in \mathbb{N}} \in [0, T]^{\mathbb{N}}$ with $\lim_{n \rightarrow \infty} t_n = 0$ the existence of a subsequence $(t_{n_k})_{k \in \mathbb{N}}$ and

$$\lim_{k \rightarrow \infty} \Psi_i^{t_{n_k}}(\mathbf{x}, \mathbf{y}) \xi^{t_{n_k}}(\mathbf{x}) \xi^{t_{n_k}}(\mathbf{y}) = \Psi_i^0(\mathbf{x}, \mathbf{y})$$

pointwise for almost every $(\mathbf{x}, \mathbf{y}) \in D_n$ and for $i = 1, 2$. Then, by utilizing the dominated convergence theorem in the second and the last step of the following calculation,

we get

$$\begin{aligned}
 & \lim_{k \rightarrow \infty} \sum_{i=1,2} \frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (u^0(\mathbf{x}) - u^0(\mathbf{y}))(\tilde{v}(\mathbf{x}) - \tilde{v}(\mathbf{y})) \Psi_i^{t_{n_k}}(\mathbf{x}, \mathbf{y}) \xi^{t_{n_k}}(\mathbf{x}) \xi^{t_{n_k}}(\mathbf{y}) d\mathbf{y} d\mathbf{x} \\
 &= \lim_{k, n \rightarrow \infty} \sum_{i=1,2} \frac{1}{2} \iint_{D_n} (u^0(\mathbf{x}) - u^0(\mathbf{y}))(\tilde{v}(\mathbf{x}) - \tilde{v}(\mathbf{y})) \Psi_i^{t_{n_k}}(\mathbf{x}, \mathbf{y}) \xi^{t_{n_k}}(\mathbf{x}) \xi^{t_{n_k}}(\mathbf{y}) d\mathbf{y} d\mathbf{x} \\
 &= \lim_{n \rightarrow \infty} \frac{1}{2} \iint_{D_n} (u^0(\mathbf{x}) - u^0(\mathbf{y}))(\tilde{v}(\mathbf{x}) - \tilde{v}(\mathbf{y})) (\Psi_1^0(\mathbf{x}, \mathbf{y}) + \Psi_2^0(\mathbf{x}, \mathbf{y})) d\mathbf{y} d\mathbf{x} \\
 &= \frac{1}{2} \iint_{(\Omega \cup \mathcal{I})^2} (u^0(\mathbf{x}) - u^0(\mathbf{y}))(\tilde{v}(\mathbf{x}) - \tilde{v}(\mathbf{y})) (\Psi_1^0(\mathbf{x}, \mathbf{y}) + \Psi_2^0(\mathbf{x}, \mathbf{y})) d\mathbf{y} d\mathbf{x}.
 \end{aligned}$$

Thus, $\lim_{k \rightarrow \infty} \mathfrak{A}(t_{n_k}, u^0, \tilde{v}) = \mathfrak{A}(0, u^0, \tilde{v})$ and equation (22) also holds for singular symmetric kernels which results in $q_1 = \varphi^0$.

For both kernel cases, it can be shown quite analogously that $q_2 = \psi^0$ and therefore $\psi^t \rightharpoonup \psi^0$ for $t \searrow 0$. However, for the convergence $\lim_{t \searrow 0} \mathfrak{A}(t, \tilde{u}, v^t) = \mathfrak{A}(0, \tilde{u}, v^0)$ we additionally have to make use of $v^t \rightharpoonup v^0$ and observation (20). \square

A.3 Proof of Lemma 8

Proof We show the statement by proving the convergence of each term

$$\begin{aligned}
 & \lim_{k \rightarrow \infty, t \searrow 0} \partial_t \mathfrak{G}(t, u^0, v^0, \psi^{s_{n_k}}, \varphi^{s_{n_k}}) \\
 &= \lim_{k \rightarrow \infty, t \searrow 0} \partial_t \mathfrak{J}(t, u^0) - \partial_t \mathfrak{F}(t, v^0) + \partial_t \mathfrak{A}(t, u^0, v^0) \\
 &\quad + \partial_t A(t, u^0, \psi^{s_{n_k}}) - \partial_t F(t, \psi^{s_{n_k}}) + \partial_t A(t, \varphi^{s_{n_k}}, v^0) - \partial_t \tilde{F}(t, u^0, \varphi^{s_{n_k}}).
 \end{aligned}$$

The convergences

$$\lim_{k \rightarrow \infty, t \searrow 0} \partial_t A(t, u^0, \psi^{s_{n_k}}) = \partial_t A(0, u^0, \psi^0) \text{ and } \lim_{k \rightarrow \infty, t \searrow 0} \partial_t F(t, \psi^{s_{n_k}}) = \partial_t F(0, \psi^0)$$

have already been shown in [18]. Then, $\lim_{k \rightarrow \infty, t \searrow 0} \partial_t A(t, \varphi^{s_{n_k}}, v^0) = \partial_t A(0, \varphi^0, v^0)$ can be illustrated in a similar way. We start by showing the convergence for $\partial_t \tilde{F}(t, u^0, \varphi^{s_{n_k}})$. Here, Lemma 12 and observation (21) yield $\xi^t \nabla \bar{u}^t \rightarrow \nabla \bar{u}$ in $L^2(\Omega, \mathbb{R}^d)$ and $\bar{u}^t \frac{\partial}{\partial r} \Big|_{r=t} \xi^r \rightarrow \bar{u} \operatorname{div} \mathbf{W}$ in $L^2(\Omega)$ for $t \searrow 0$. Furthermore, due to the boundedness of \mathbf{W} , we have $\xi^t (\nabla \bar{u}^t)^\top \mathbf{W} \rightarrow \nabla \bar{u}^\top \mathbf{W}$ in $L^2(\Omega)$. Thus, by (20) we conclude

$$\begin{aligned}
 \lim_{k \rightarrow \infty, t \searrow 0} \partial_t \tilde{F}(t, u^0, \varphi^{s_{n_k}}) &= \lim_{k \rightarrow \infty, t \searrow 0} \int_{\Omega} (\nabla \bar{u}^t)^\top \mathbf{W} \varphi^{s_{n_k}} \xi^t - (u^0 - \bar{u}^t) \varphi^{s_{n_k}} \frac{\partial}{\partial r} \Big|_{r=t} \xi^r d\mathbf{x} \\
 &= \int_{\Omega} \nabla \bar{u}^\top \mathbf{W} \varphi^0 - (u^0 - \bar{u}) \varphi^0 \operatorname{div} \mathbf{W} d\mathbf{x} = \partial_t \tilde{F}(t, u^0, \varphi^0).
 \end{aligned}$$

Now, we continue by proving the convergence with respect to $\partial_t \mathfrak{J}(t, u^0)$. Since $\mathbf{V} \in C_0^2(\Omega \cup \mathcal{I}, \mathbb{R}^d)$ and $\bar{u} \in H^2(\Omega)$, we have $(\nabla \bar{u}^t)^\top \mathbf{V}^t \rightarrow (\nabla \bar{u})^\top \mathbf{V}$ in $L^2(\Omega)$ as well as $(\nabla \bar{u}^t)^\top D\mathbf{V}^t \rightarrow (\nabla \bar{u})^\top D\mathbf{V}$ and $\text{Hess}(\bar{u})^t \mathbf{V}^t \rightarrow \text{Hess}(\bar{u})\mathbf{V}$ in $L^2(\Omega, \mathbb{R}^d)$ as a result of Lemma 12. Combined with the boundedness of $\mathbf{W} \in C_0^2(\Omega \cup \mathcal{I}, \mathbb{R}^d)$ we get $(\mathbf{V}^t)^\top \text{Hess}(\bar{u})^t \mathbf{W} \rightarrow \mathbf{V}^\top \text{Hess}(\bar{u})\mathbf{W}$ and $(\nabla \bar{u}^t)^\top D\mathbf{V}^t \mathbf{W} \rightarrow (\nabla \bar{u})^\top (D\mathbf{V})\mathbf{W}$ in $L^2(\Omega)$. Further, notice that the functions $(t, \mathbf{x}) \mapsto (\nabla \text{div } \mathbf{V}^t(\mathbf{x}))^\top \mathbf{W}(\mathbf{x}) \xi^t(\mathbf{x})$ and $(t, \mathbf{x}) \mapsto \text{div } \mathbf{V}^t(\mathbf{x}) \frac{\partial}{\partial r} \Big|_{r=t} \xi^r(\mathbf{x})$ are bounded on $[0, T] \times \overline{\Omega}$ and converge in $L^2(\Omega)$ for $t \searrow 0$. Consequently, by applying (20) and (21), we have

$$\begin{aligned}
 \lim_{t \searrow 0} \partial_t \mathfrak{J}(t, u^0) &= \lim_{t \searrow 0} \int_{\Omega} (\nabla \bar{u}^t)^\top \mathbf{W} (\nabla \bar{u}^t)^\top \mathbf{V}^t \xi^t d\mathbf{x} \\
 &* - \int_{\Omega} (u^0 - \bar{u}^t) \left((\mathbf{V}^t)^\top \text{Hess}(\bar{u})^t \mathbf{W} \xi^t \right. \\
 &\quad \left. + (\nabla \bar{u}^t)^\top (D\mathbf{V}^t) \mathbf{W} \xi^t + (\nabla \bar{u}^t)^\top \mathbf{V}^t \frac{\partial}{\partial r} \Big|_{r=t} (\xi^r) \right) d\mathbf{x} \\
 &* + \int_{\Omega} -(u^0 - \bar{u}^t) (\nabla \bar{u}^t)^\top \mathbf{W} \text{div } \mathbf{V}^t \xi^t \\
 &* \quad + \frac{1}{2} (u^0 - \bar{u}^t)^2 \left((\nabla \text{div } \mathbf{V}^t)^\top \mathbf{W} \xi^t + \text{div } \mathbf{V}^t \frac{\partial}{\partial r} \Big|_{r=t} \xi^r \right) d\mathbf{x} \\
 &= \int_{\Omega} \nabla \bar{u}^\top \mathbf{W} \nabla \bar{u}^\top \mathbf{V} d\mathbf{x} \\
 &* - \int_{\Omega} (u^0 - \bar{u}) \left(\mathbf{V}^\top \text{Hess}(\bar{u})\mathbf{W} + \nabla \bar{u}^\top (D\mathbf{V})\mathbf{W} + \nabla \bar{u}^\top \mathbf{V} \text{div } \mathbf{W} \right) d\mathbf{x} \\
 &* + \int_{\Omega} -(u^0 - \bar{u}) \nabla \bar{u}^\top \mathbf{W} \text{div } \mathbf{V} + \frac{1}{2} (u^0 - \bar{u})^2 \left((\nabla \text{div } \mathbf{V})^\top \mathbf{W} + \text{div } \mathbf{V} \text{div } \mathbf{W} \right) d\mathbf{x}
 \end{aligned}$$

Further, $f_\Gamma \in H^2(\Omega_i)$ for $i = 1, 2$ and with the same arguments as before, we derive

$$\begin{aligned}
 \lim_{t \searrow 0} \partial_t \mathfrak{F}(t, v^0) \\
 &= \lim_{t \searrow 0} \int_{\Omega} (\mathbf{V}^t)^\top \text{Hess}(f)^t \mathbf{W} v^0 \xi^t \\
 &\quad + (\nabla f^t)^\top (D\mathbf{V}^t) \mathbf{W} v^0 \xi^t + (\nabla f^t)^\top \mathbf{V}^t v^0 \frac{\partial}{\partial r} \Big|_{r=t} (\xi^r) d\mathbf{x} \\
 &\quad + \int_{\Omega} (\nabla f^t)^\top \mathbf{W} v^0 \text{div } \mathbf{V}^t \xi^t + f^t v^0 \left((\nabla \text{div } \mathbf{V}^t)^\top \mathbf{W} \xi^t + \text{div } \mathbf{V}^t \frac{\partial}{\partial r} \Big|_{r=t} (\xi^r) \right) d\mathbf{x} \\
 &= \int_{\Omega} \mathbf{V}^\top \text{Hess}(f_\Gamma) \mathbf{W} v^0 + \nabla f_\Gamma^\top (D\mathbf{V}) \mathbf{W} v^0 + \nabla f_\Gamma^\top \mathbf{V} v^0 \text{div } \mathbf{W} + \nabla f_\Gamma^\top \mathbf{W} v^0 \text{div } \mathbf{V} \\
 &* \quad + f_\Gamma v^0 \left((\nabla \text{div } \mathbf{V})^\top \mathbf{W} + \text{div } \mathbf{V} \text{div } \mathbf{W} \right) d\mathbf{x}
 \end{aligned}$$

it remains to show $\partial_t \mathfrak{A}(t, u^0, v^0) \rightarrow \partial_t \mathfrak{A}(0, u^0, v^0)$ for $t \searrow 0$.

Case 1: Integrable Kernel

First, we denote that the term

$$\begin{aligned} \nabla \Psi_1^t(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) &= \mathbf{V}^t(\mathbf{x}, \mathbf{y})^\top \text{Hess}(\gamma^t)(\mathbf{x}, \mathbf{y}) \mathbf{W}(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \\ &\quad + \nabla \gamma^t(\mathbf{x}, \mathbf{y})^\top (D\mathbf{V}^t(\mathbf{x}, \mathbf{y})) \mathbf{W}(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \end{aligned}$$

is essentially bounded for $t \in [0, T]$ and $(\mathbf{x}, \mathbf{y}) \in (\Omega \cup \mathcal{I})^2$. Moreover, observation (21) and Lemma 12 yield

$$\begin{aligned} \nabla \Psi_1^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) &\rightarrow \nabla \Psi_1^0(\mathbf{x}, \mathbf{y}) \text{ in } L^2((\Omega \cup \mathcal{I})^2, \mathbb{R}^d) \text{ and} \\ \Psi_1^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) &= \nabla \gamma^t(\mathbf{x}, \mathbf{y})^\top \mathbf{V}^t(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \rightarrow \nabla \gamma^0(\mathbf{x}, \mathbf{y})^\top \mathbf{V}^0(\mathbf{x}, \mathbf{y}) = \Psi_1^0(\mathbf{x}, \mathbf{y}) \end{aligned}$$

in $L^2((\Omega \cup \mathcal{I})^2, \mathbb{R})$.

Thus, for every sequence $\{t_n\}_{n \in \mathbb{N}} \in [0, T]^\mathbb{N}$ with $\lim_{n \rightarrow \infty} t_n = 0$ there exists a subsequence $\{t_{n_k}\}_{k \in \mathbb{N}}$ such that $\Psi_1^{t_{n_k}}(\mathbf{x}, \mathbf{y}) \xi^{t_{n_k}}(\mathbf{x}) \xi^{t_{n_k}}(\mathbf{y}) \rightarrow \Psi_1^0(\mathbf{x}, \mathbf{y})$ and $\nabla \Psi_1^{t_{n_k}}(\mathbf{x}, \mathbf{y}) \xi^{t_{n_k}}(\mathbf{x}) \xi^{t_{n_k}}(\mathbf{y}) \rightarrow \nabla \Psi_1^0(\mathbf{x}, \mathbf{y})$ pointwise almost everywhere. By making use of the dominated convergence theorem we derive

$$\begin{aligned} &\iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y})) (u^0(\mathbf{x}) \nabla \Psi_1^{t_{n_k}}(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \\ &\quad * - u^0(\mathbf{y}) \nabla \Psi_1^{t_{n_k}}(\mathbf{y}, \mathbf{x})^\top \mathbf{W}(\mathbf{y}, \mathbf{x}) \xi^{t_{n_k}}(\mathbf{x}) \xi^{t_{n_k}}(\mathbf{y}) \, d\mathbf{y} d\mathbf{x} \\ &\rightarrow \iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y})) (u^0(\mathbf{x}) \nabla \Psi_1^0(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \\ &\quad - u^0(\mathbf{y}) \nabla \Psi_1^0(\mathbf{y}, \mathbf{x})^\top \mathbf{W}(\mathbf{y}, \mathbf{x})) \, d\mathbf{y} d\mathbf{x} \text{ for } k \rightarrow \infty. \end{aligned}$$

Moreover, since ξ^t is continuously differentiable, the derivative $\frac{\partial}{\partial r} \Big|_{r=t} \xi^r$ is also essentially bounded and $\text{div } \mathbf{V}^t \rightarrow \text{div } \mathbf{V}$ pointwise as a composition of continuous functions. Combined with arguments from above, the following integrands are essentially bounded by a constant independent of $t \in [0, T]$ and again for every sequence $\{t_n\}_{n \in \mathbb{N}} \in [0, T]^\mathbb{N}$ with $\lim_{n \rightarrow \infty} t_n = 0$ there exists a subsequence $\{t_{n_k}\}_{k \in \mathbb{N}}$ such that the integrands converge pointwise almost everywhere. Thus, dominated convergence yields

$$\begin{aligned} &\iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y})) \left(u^0(\mathbf{x}) \Psi_1^{t_{n_k}}(\mathbf{x}, \mathbf{y}) \right. \\ &\quad \left. - u^0(\mathbf{y}) \Psi_1^{t_{n_k}}(\mathbf{y}, \mathbf{x}) \right) \frac{\partial}{\partial r} \Big|_{r=t_{n_k}} (\xi^r(\mathbf{x}) \xi^r(\mathbf{y})) \, d\mathbf{y} d\mathbf{x} \\ &\rightarrow \iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y})) (u^0(\mathbf{x}) \Psi_1^0(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y}) \Psi_1^0(\mathbf{y}, \mathbf{x})) (\text{div } \mathbf{W}(\mathbf{x}) + \text{div } \mathbf{W}(\mathbf{y})) \, d\mathbf{y} d\mathbf{x} \end{aligned}$$

and we also conclude

$$\begin{aligned} & \iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(u^0(\mathbf{x}) \nabla \Psi_2^{t_{n_k}}(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \\ & \quad - u^0(\mathbf{y}) \nabla \Psi_2^{t_{n_k}}(\mathbf{y}, \mathbf{x})^\top \mathbf{W}(\mathbf{y}, \mathbf{x})) \xi^{t_{n_k}}(\mathbf{x}) \xi^{t_{n_k}}(\mathbf{y}) \, d\mathbf{y} d\mathbf{x} \\ \rightarrow & \iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(u^0(\mathbf{x}) \nabla \Psi_2^0(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y}) \nabla \Psi_2^0(\mathbf{y}, \mathbf{x})^\top \mathbf{W}(\mathbf{y}, \mathbf{x})) \, d\mathbf{y} d\mathbf{x}. \end{aligned}$$

For the last term, we derive by applying the dominated convergence theorem

$$\begin{aligned} & \iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(u^0(\mathbf{x}) \Psi_2^{t_{n_k}}(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y}) \Psi_2^{t_{n_k}}(\mathbf{y}, \mathbf{x})) \left. \frac{\partial}{\partial r} \right|_{r=t_{n_k}} (\xi^r(\mathbf{x}) \xi^r(\mathbf{y})) \, d\mathbf{y} d\mathbf{x} \\ \rightarrow & \iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(u^0(\mathbf{x}) \Psi_2^0(\mathbf{x}, \mathbf{y}) - u^0(\mathbf{y}) \Psi_2^0(\mathbf{y}, \mathbf{x})) (\operatorname{div} \mathbf{W}(\mathbf{x}) + \operatorname{div} \mathbf{W}(\mathbf{y})) \, d\mathbf{y} d\mathbf{x}. \end{aligned}$$

As a result, we obtain $\partial_t \mathfrak{A}(t_{n_k}, u^0, v^0) \rightarrow \partial_t \mathfrak{A}(0, u^0, v^0)$ for $k \rightarrow \infty$ and

$$\partial_t \mathfrak{A}(t, u^0, v^0) \rightarrow \partial_t \mathfrak{A}(0, u^0, v^0) \text{ for } t \searrow 0.$$

Case 2: Singular Symmetric Kernel

On the sets $[0, T] \times D_n$ the functions

$$\begin{aligned} a_1(t, \mathbf{x}, \mathbf{y}) &:= \nabla \Psi_1^t(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \\ &= \left(\mathbf{V}^t(\mathbf{x}, \mathbf{y})^\top \operatorname{Hess}(\gamma^t)(\mathbf{x}, \mathbf{y}) + \nabla \gamma^t(\mathbf{x}, \mathbf{y})^\top D\mathbf{V}^t(\mathbf{x}, \mathbf{y}) \right) \mathbf{W}(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}), \\ a_2(t, \mathbf{x}, \mathbf{y}) &:= \Psi_1^t(\mathbf{x}, \mathbf{y}) \left. \frac{\partial}{\partial r} \right|_{r=t} (\xi^t(\mathbf{x}) \xi^t(\mathbf{y})) = \nabla \gamma^t(\mathbf{x}, \mathbf{y})^\top \mathbf{V}^t(\mathbf{x}, \mathbf{y}) \left. \frac{\partial}{\partial r} \right|_{r=t} (\xi^t(\mathbf{x}) \xi^t(\mathbf{y})) \\ a_3(t, \mathbf{x}, \mathbf{y}) &:= \nabla \Psi_2^t(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \\ &= (\operatorname{div} \mathbf{V}^t(\mathbf{x}) + \operatorname{div} \mathbf{V}^t(\mathbf{y})) \nabla \gamma^t(\mathbf{x}, \mathbf{y})^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \\ &\quad + \gamma^t(\mathbf{x}, \mathbf{y}) (\nabla \operatorname{div} \mathbf{V}^t(\mathbf{x}) + \nabla \operatorname{div} \mathbf{V}^t(\mathbf{y}))^\top \mathbf{W}(\mathbf{x}, \mathbf{y}) \xi^t(\mathbf{x}) \xi^t(\mathbf{y}) \text{ and} \\ a_4(t, \mathbf{x}, \mathbf{y}) &:= \Psi_2^t(\mathbf{x}, \mathbf{y}) \left. \frac{\partial}{\partial r} \right|_{r=t} (\xi^r(\mathbf{x}) \xi^r(\mathbf{y})) \\ &= \gamma^t(\mathbf{x}, \mathbf{y}) (\operatorname{div} \mathbf{V}^t(\mathbf{x}) + \operatorname{div} \mathbf{V}^t(\mathbf{y})) \left. \frac{\partial}{\partial r} \right|_{r=t} (\xi^r(\mathbf{x}) \xi^r(\mathbf{y})) \end{aligned}$$

are essentially bounded as a result of Assumption (S1) (see also Remark 7). With the same argumentation as in the case of integrable kernels, for every sequence $(t_n)_{n \in \mathbb{N}} \in [0, T]^\mathbb{N}$ with $\lim_{n \rightarrow \infty} t_n = 0$ there exists a subsequence $(t_{n_k})_{k \in \mathbb{N}}$ such that $a_i(t_{n_k}, \mathbf{x}, \mathbf{y}) \rightarrow a_i(0, \mathbf{x}, \mathbf{y})$ pointwise almost everywhere on D_n for $i = 1, \dots, 4$. Thus,

we derive by making use of the dominated convergence theorem

$$\begin{aligned}
 & \sum_{i=1}^4 \lim_{k \rightarrow \infty} \iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(u^0(\mathbf{x}) - u^0(\mathbf{y}))a_i(t_{n_k}, \mathbf{x}, \mathbf{y}) \, d\mathbf{y}d\mathbf{x} \\
 &= \sum_{i=1}^4 \lim_{k, n \rightarrow \infty} \iint_{D_n} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(u^0(\mathbf{x}) - u^0(\mathbf{y}))a_i(t_{n_k}, \mathbf{x}, \mathbf{y}) \, d\mathbf{y}d\mathbf{x} \\
 &= \sum_{i=1}^4 \lim_{n \rightarrow \infty} \iint_{D_n} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(u^0(\mathbf{x}) - u^0(\mathbf{y}))a_i(0, \mathbf{x}, \mathbf{y}) \, d\mathbf{y}d\mathbf{x} \\
 &= \sum_{i=1}^4 \iint_{(\Omega \cup \mathcal{I})^2} (v^0(\mathbf{x}) - v^0(\mathbf{y}))(u^0(\mathbf{x}) - u^0(\mathbf{y}))a_i(0, \mathbf{x}, \mathbf{y}) \, d\mathbf{y}d\mathbf{x}.
 \end{aligned}$$

Consequently, we get $\partial_t \mathfrak{A}(t_{n_k}, u^0, v^0) \rightarrow \partial_t \mathfrak{A}(0, u^0, v^0)$ for $k \rightarrow \infty$ and thus

$$\partial_t \mathfrak{A}(t, u^0, v^0) \rightarrow \partial_t \mathfrak{A}(0, u^0, v^0) \text{ for } t \searrow 0.$$

□

Acknowledgements This work has been partly supported by the German Research Foundation (DFG) within the Research Training Group 2126: 'Algorithmic Optimization' (Grant ID: 264747124).

Author Contributions M.S. conducted the numerical experiments and wrote the manuscript, which was reviewed and annotated by V.S.. Additionally, V.S. advised during the whole process.

Funding Open Access funding enabled and organized by Projekt DEAL. This work has been partly supported by the German Research Foundation (DFG) within the Research Training Group 2126: 'Algorithmic Optimization' (Grant ID: 264747124).

Data Availability Available upon request.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

Ethical approval Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Silling, S.A.: Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**(1), 175–209 (2000)
2. Javili, A., Morasata, R., Oterkus, E., Oterkus, S.: Peridynamics review. *Math. Mech. Solids* **24**(11), 3714–3739 (2019)
3. Suzuki, J.L., Gulian, M., Zayernouri, M., D’Elia, M.: Fractional modeling in action: a survey of nonlocal models for subsurface transport, turbulent flows, and anomalous materials. *J. Peridyn. Nonlocal Model.* **5**, 392–459 (2023)
4. Brockmann, D.: Anomalous diffusion and the structure of human transportation networks. *Eur. Phys. J. Special Topics* **157**, 173–189 (2008)
5. Buades, A., Coll, B., Morel, J.-M.: Image denoising methods. A new nonlocal principle. *SIAM Review* **52**(1), 113–147 (2010)
6. D’Elia, M., Reyes, J.C.D.L., Miniguanco-Trujillo, A.: Bilevel parameter learning for nonlocal image denoising models. *J. Math. Imaging Vis.* **63**, 753–775 (2021)
7. Tao, Y., Sun, Q., Du, Q., Liu, W.: Nonlocal neural networks, nonlocal diffusion and nonlocal modeling. In: Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., Garnett, R., (eds.) *Advances in Neural Information Processing Systems*, vol. 31. Curran Associates, Inc. (2018)
8. You, H., Yu, Y., D’Elia, M., Gao, T., Silling, S.: Nonlocal kernel network (NKN): a stable and resolution-independent deep neural network. *J. Comput. Phys.* **469**, 111536 (2022)
9. Metzler, R., Klafter, J.: The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.* **339**(1), 1–77 (2000)
10. D’Elia, M., Du, Q., Gunzburger, M., Lehoucq, R.: Nonlocal convection-diffusion problems on bounded domains and finite-range jump processes. *Comput. Methods Appl. Math.* **17**(4), 707–722 (2017)
11. Schmidt, S.: Efficient Large Scale Aerodynamic Design Based on Shape Calculus. PhD thesis, Universität Trier (2010)
12. Li, J., Du, X., Martins, J.R.R.A.: Machine learning in aerodynamic shape optimization. *Prog. Aerosp. Sci.* **134**, 100849 (2022)
13. Bängtsson, E., Noreland, D., Berggren, M.: Shape optimization of an acoustic horn. *Comput. Methods Appl. Mech. Eng.* **192**(11–12), 1533–1571 (2003)
14. Bernland, A., Wadbro, E., Berggren, M.: Acoustic shape optimization using cut finite elements. *Int. J. Numer. Meth. Eng.* **113**(3), 432–449 (2018)
15. Schmidt, S., Wadbro, E., Berggren, M.: Large-scale three-dimensional acoustic horn optimization. *SIAM J. Sci. Comput.* **38**(6), 917–940 (2016)
16. Mohammadi, B., Pironneau, O.: Shape optimization in fluid mechanics. *Annu. Rev. Fluid Mech.* **36**, 255–279 (2004)
17. Schlegel, L.: Shape Optimization for the Mitigation of Coastal Erosion. PhD thesis, Universität Trier (2023)
18. Schuster, M., Vollmann, C., Schulz, V.: Shape optimization for interface identification in nonlocal models. *Comput. Optim. Appl.* **88**, 963–997 (2024)
19. Schuster, M., Schulz, V.: Interface Identification Constrained by Local-to-Nonlocal Coupling. *arXiv preprint arXiv:2402.12871* (2024)
20. Nader, N., Diehl, P., D’Elia, M., Glusa, C., Prudhomme, S.: Machine learning-based identification of the interface regions for coupling local and nonlocal models. *J. Mach. Learning Model. Comput.* **5**(3), 1–29 (2024)
21. Du, Q., Gunzburger, M., Lehoucq, R.B., Zhou, K.: Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Rev.* **54**(4), 667–696 (2012)
22. Vollmann, C.: Nonlocal Models with Truncated Interaction Kernels-Analysis, Finite Element Methods and Shape Optimization. PhD thesis, Universität Trier (2019)
23. Delfour, M.C., Zolésio, J.-P.: Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization. SIAM, Philadelphia (2011)
24. Sokolowski, J., Zolésio, J.-P.: Introduction to Shape Optimization. Springer, Berlin (1992)
25. Allaire, G., Dapogny, C., Jouve, F.: Chapter 1—Shape and Topology Optimization. In: *Handbook of Numerical Analysis* vol. 22, pp. 1–132. Elsevier, Amsterdam (2021)
26. Sturm, K.: On Shape Optimization with Non-Linear Partial Differential Equations. PhD thesis, Technische Universität Berlin (2015)

27. Welker, K.: Efficient PDE constrained shape optimization in shape spaces. PhD Dissertation, Trier University (2017)
28. Henrot, A., Pierre, M.: Shape Variation and Optimization : a Geometrical Analysis, vol. 28. European Mathematical Society (2018)
29. Allaire, G.: Conception Optimale de Structures. Springer, Berlin (2007)
30. Sturm, K.: Convergence of Newton's Method in Shape Optimisation via Approximate Normal Functions. arXiv preprint [arXiv:1608.02699](https://arxiv.org/abs/1608.02699) (2016)
31. Schmidt, S., Schulz, V.: A linear view on shape optimization. *SIAM J. Control. Optim.* **61**(4), 2358–2378 (2023)
32. Schulz, V.: A Riemannian view on shape optimization. *Found. Comput. Math.* **14**, 483–501 (2014)
33. Novruzzi, A., Pierre, M.: Structure of shape derivatives. *J. Evol. Equ.* **2**, 365–382 (2002)
34. Delfour, M.C., Zolésio, J.-P.: Anatomy of the shape Hessian. *Annali di Matematica* **159**, 315–339 (1991)
35. Ameer, H.B., Burger, M., Hackl, B.: Level set methods for geometric inverse problems in linear elasticity. *Inverse Prob.* **20**(3), 673–696 (2004)
36. Schmidt, S., Gräßer, M., Schmid, H.-J.: A shape Newton scheme for deforming shells with application to capillary bridges. *SIAM J. Sci. Comput.* **44**(4), 1175–1194 (2022)
37. Sturm, K.: Minimax Lagrangian approach to the differentiability of nonlinear PDE constrained shape functions without saddle point assumption. *SIAM J. Control. Optim.* **53**(4), 2017–2039 (2015)
38. Laurain, A., Sturm, K.: Distributed shape derivative via averaged adjoint method and applications. *ESAIM Math. Modell. Numer. Anal.* **50**(4), 1241–1267 (2016)
39. Gangl, P., Sturm, K., Neunteufel, M., Schöberl, J.: Fully and semi-automated shape differentiation in NGSolve. *Struct. Multidiscip. Optim.* **63**, 1579–1607 (2021)
40. Cesarano, A., Endtmayer, B., Gangl, P.: Homotopy Methods for Higher Order Shape Optimization: A Globalized Shape-Newton Method and Pareto-Front Tracing. arXiv preprint [arXiv:2405.03421](https://arxiv.org/abs/2405.03421) (2024)
41. Hintermüller, M.: Fast level set based algorithms using shape and topological sensitivity information. *Control. Cybern.* **34**(1), 305–324 (2005)
42. Etling, T., Herzog, R., Loayza, E., Wachsmuth, G.: First and second order shape optimization based on restricted mesh deformations. *SIAM J. Sci. Comput.* **42**(2), 1200–1225 (2020)
43. D'Elia, M., Gunzburger, M., Vollmann, C.: A cookbook for approximating Euclidean balls and for quadrature rules in finite element methods for nonlocal problems. *Math. Models Methods Appl. Sci.* **31**(08), 1505–1567 (2021)
44. Geuzaine, C., Remacle, J.-F.: Gmsh: a 3-D finite element mesh generator with built-in pre-and post-processing facilities. *Int. J. Numer. Meth. Eng.* **79**(11), 1309–1331 (2009)
45. Klar, M., Vollmann, C., Schulz, V.: nlfem: a flexible 2D FEM python code for nonlocal convection-diffusion and mechanics. *J. Peridyn. Nonlocal Model.* **6**, 87–117 (2024)
46. Alnæs, M., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M.E., Wells, G.N.: The FEniCS project version 1.5. *Arch. Numer. Softw.* **3**(100), 9–23 (2015)
47. Logg, A., Mardal, K.-A., Wells, G.: Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book. Springer, Berlin Heidelberg (2012)
48. Schuster, M.: Shape Optimization Constrained by Nonlocal Equations. PhD thesis, Universität Trier (2025)
49. Evans, L.C.: Partial Differential Equations, vol. 19. American Mathematical Soc, Providence (2010)
50. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. Elsevier, Amsterdam (2003)