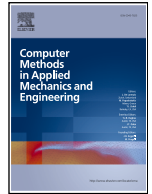




Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

Sparse source identification in transient advection-diffusion problems with a primal-dual-active-point strategy

Marco Mattuschka ^{a,*}, Daniel Walter ^b, Max von Danwitz ^a,
Alexander Popp ^{a,c}

^a Institute for the Protection of Terrestrial Infrastructures, German Aerospace Center (DLR), 53757, Sankt Augustin, Germany

^b Institute for Numerical Mathematics, Johannes Kepler University, 4040, Linz, Austria

^c Institute for Mathematics and Computer-Based Simulations (IMCS), University of the Bundeswehr Munich, 85577, Neubiberg, Germany

ARTICLE INFO

Keywords:

Airborne contaminant transport
Advection-diffusion equation
Source detection
Large-scale inverse problems
Sparse optimization

ABSTRACT

This work presents a mathematical model to enable rapid prediction of airborne contaminant transport based on scarce sensor measurements. The method is designed for applications in critical infrastructure protection (CIP), such as evacuation planning following contaminant release. In such scenarios, timely and reliable decision-making is essential, despite limited observation data. To identify contaminant sources, we formulate an inverse problem governed by an advection-diffusion equation. Given the problem's underdetermined nature, we further employ a variational regularization ansatz and model the unknown contaminant sources as distribution over the spatial domain. To efficiently solve the arising inverse problem, we employ a problem-specific variant of the Primal-Dual-Active-Point (PDAP) algorithm which efficiently approximates sparse minimizers of the inverse problem by alternating between greedy location updates and source intensity optimization. The approach is demonstrated on two- and three-dimensional test cases involving both instantaneous and continuous contaminant sources and outperforms state-of-the-art techniques with L^2 -regularization. Its effectiveness is further illustrated in complex domains with real-world building geometries imported from OpenStreetMap.

1. Introduction

Airborne contaminant transport of hazardous materials poses a significant threat to communities and critical infrastructures. Contaminant release may occur accidentally in industrial leaks or spills, or may be caused intentionally in an act of sabotage or terrorism [1–3]. An example visualization of the dispersion of a contaminant is provided in Fig. 1. In order to stem these threats and predict further spreading of the contaminant, identifying the source of contamination is of utmost importance. In many scenarios, chemical hazards are caused by toxic industrial chemicals (TICs) including colorless gases invisible to the human eye and RGB cameras. Hence, inference about the unknown source is only possible indirectly through scarce measurements of the concentration field, e.g. point observations in the computational domain (green spheres in Fig. 1). Describing the contaminant dispersion by an advection-diffusion equation and assuming models for the measurement process as well as the shape of the contaminant source, leads to an ill-posed inverse problem. In this paper, we revisit a recently popularized variational regularization ansatz for its stable

* Corresponding author.

E-mail addresses: marco.mattuschka@dlr.de (M. Mattuschka), daniel.walter@jku.at (D. Walter), max.vondanwitz@dlr.de (M. von Danwitz), alexander.popp@dlr.de (A. Popp).

<https://doi.org/10.1016/j.cma.2026.118854>

Received 31 October 2025; Received in revised form 17 January 2026; Accepted 13 February 2026

Available online 2 March 2026

0045-7825/© 2026 The Author(s).

Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Published by Elsevier B.V. This is an open access article under the CC BY license

solution and showcase its viability in scenarios with realistic computational domains. Due to its practical relevance for decision-makers in emergency response, numerous approaches to source identification have been explored and documented in the literature. The specific computational method for source identification or source term estimation (STE) follows the underlying dispersion model. The complexity of contaminant dispersion models ranges from simple box models and Gaussian plume models to three-dimensional geometry-aware high fidelity models. An overview of commonly employed dispersion models is given in [4]. For an overview of the model-associated source term estimation methods, the interested reader is referred to [5] and references therein. The more specific task of active source identification additionally involves the routing of a mobile sensor system [6]. Also, the very recent literature documents several practical approaches contributing to gas source localization [7–9].

From the set of available dispersion models, we choose to model the airborne contaminant transport with a linear advection-diffusion equation in which the unknown sources appear as initial condition or right-hand-side term. In general, the passive transport of a substance (pollutant, chemical species, here: contaminant) in an incompressible or compressible fluid can be mathematically modeled by the advection-diffusion problem [10,11]. The wind vector field that drives the transport in the advection-diffusion problem might stem from Computational Fluid Dynamics (CFD) simulations [12,13] or can be inferred from measurements itself [14,15]. In the present work, we consider the wind vector field as a known (fixed) input to the advection-diffusion equation. Throughout the paper, we refer to the process of simulating the latter for a given source and taking measurements of the obtained concentration field as the *forward problem*. Given noisy measurements, the identification of the unknown source then corresponds to the associated ill-posed *inverse problem*.

For the particular application under consideration, various approaches for the stable solution of the inverse problem have been considered in the literature. For example, a Bayesian ansatz for finite-dimensional parametrized sources is presented in which the resulting posterior distribution is sampled via Markov chain Monte Carlo (MCMC) methods [16]. This approach heavily relies on the assumed measurement model which consists of a finite number of point measurements at a fixed time, allowing to compute all required PDE solves in an offline phase. In contrast, the present work considers time series data collected at a finite number of spatially fixed or moving sensor locations, thus preventing similar arguments. Instead, we formulate the source identification problem as a minimization task, incorporate the advection-diffusion equation as a constraint, and compensate the scarcity of measurement data by employing regularization techniques [17–19]. In this regard, Tikhonov regularization with weighted L^2 -type penalty terms leads to quadratic minimization problems which are amenable to Newton methods, Wiedemann et al. [9], Petra and Stadler [20]. This approach can be further developed towards applications with real-time requirements by replacing high-fidelity models with reduced ones that still capture the essential characteristics, yet are much faster to evaluate [21,22]. From a Bayesian perspective, problems of the described form appear in the computation of the maximum a posteriori estimator wherein specific forms of data misfit and penalty term are related to the noise model as well as to the prior distribution of the unknown source, respectively, Stuart [23], Nitzler et al. [24].

However, in realistic scenarios, it is reasonable to assume that contaminant sources are very localized, i.e. around an unknown finite number of locations, and that measurements are scarcely distributed in space, e.g. given by time series concentration data of few fixed sensors. Due to the smoothing behavior of the advection-diffusion process, weighted L^2 -type regularization is unable to capture this structural assumption accurately. As a remedy, in the present paper, we assume that the sought-for contamination source is given by a superposition of finitely many simple atoms of a given shape, e.g. point sources or Gaussians, each characterized by a location parameter as well as a positive amplitude. Since their number is in general also unknown, this modeling assumption leads to a highly nonlinear parameter identification problem.

In order to alleviate these difficulties, we relax it from sums of atoms to parameters represented as integrals of the shape function w.r.t an unknown, positive Radon measure. The identification of the latter is tackled via variational regularization techniques in which the prior assumption on the sparse representation of the unknown source is incorporated by a Radon norm penalty. This approach replaces the combinatorial complexity and strong nonlinearity of the parametrized ansatz with a different set of challenges, namely, nonsmoothness and an infinite-dimensional space of unknowns, while offering the added benefit that the relaxed problem is convex. Moreover, it turns out that this relaxation is exact since optimal reconstructions consisting of finitely many Dirac-Delta functionals exist, see [Theorem 1](#). Variational regularization approaches involving the Radon norm have received increased attention over the last years, see e.g. [25–27] for a nonexhaustive list of references. For the particular application to initial value identification in parabolic problems we refer to [28–31]. In particular, several works have addressed the super-resolution properties of this ansatz, Candès and Fernandez-Granda [32], Duval and Peyré [33], i.e. the capability to recover the correct number of sources together with quantitative convergence guarantees for the associated points and amplitudes despite the presence of measurement noise. In this work, we adopt the framework outlined above for scenarios in contamination transport where measurements are scarcely distributed in space, e.g. given by time series concentration data of few fixed sensors. Our practical realization relies on the Primal-Dual-Active-Point (PDAP) strategy from [34] which efficiently approximates sparse minimizers by alternating between greedy location updates and intensity optimization. Crucially, in comparison to e.g. [35–37], this approach does not rely on nonconvex point-moving steps and is thus compatible with standard finite element discretizations. While, as outlined above, the concept of Radon norm regularization has been studied before, our particular focus lies on its application to practically relevant two and three-dimensional benchmark settings with real building imprints, bridging the gap between mathematical theory and the demand of real-world scenarios. Comparisons to state-of-the-art methods with weighted L^2 -type regularization, reveal excellent recovery results across a variety of examples while the proposed approach also exhibits a favorable computational scaling behavior w.r.t. to the availability of additional information, e.g., by adding new sensors or higher sampling frequency.

The remainder of the present paper is organized as follows. Mathematical modeling of airborne contaminant transport, along with the formulation of a sparsity promoting regularization approach for inverse problems, is presented in [Section 2](#). For the sake

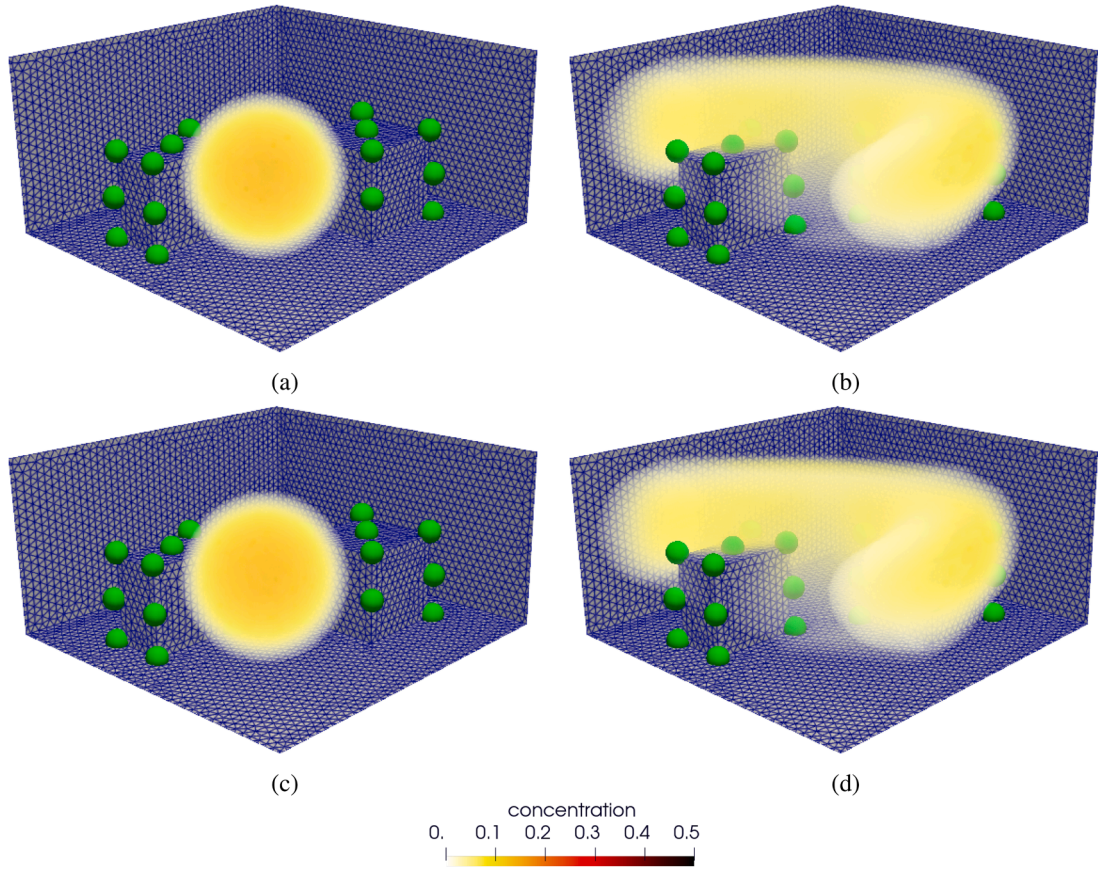


Fig. 1. Three-dimensional gas source identification. Manufactured initial condition (a) and simulation at $t = 3.0$ s (b), reproduced by algorithmic reconstruction of initial contaminant source (c) and simulation at $t = 3.0$ s (d) based on sparse measurements. Positions of noisy point-evaluations (modeled sensors) are marked by green spheres. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of completeness, this section also briefly reviews a classical method with L^2 -regularization. In Section 3, the PDAP algorithm for the efficient solution the source identification problem is presented. Next, Section 4 describes the finite element discretization of the forward and adjoint problem, followed by two- and three-dimensional numerical examples presented in Section 5. Finally, Section 6 offers conclusions and an outlook to future developments.

2. Background

2.1. Forward problem, parameter-to-observable and misfit-to-adjoint map

A mathematical description of the transport of a substance (contaminant) concentration u in a bounded open domain $\Omega \subseteq \mathbb{R}^n$ for $n \in \{2, 3\}$ is given by the following equation:

$$\begin{aligned}
 u_t - \kappa \Delta u + \mathbf{v} \cdot \nabla u &= m_C && \text{in } (0, T) \times \Omega, \\
 \kappa \nabla u \cdot \mathbf{n} &= 0 && \text{in } (0, T) \times (\Gamma_+ \cup \Gamma_0), \\
 u &= 0 && \text{in } (0, T) \times \Gamma_-, \\
 u(0, \cdot) &= m_I && \text{in } \Omega.
 \end{aligned} \tag{P_K}$$

The problem (P_K) depends on two unknown functions $m = (m_I, m_C)$ on Ω , an initial condition m_I and a source term m_C which we assume to be independent of time. The abstract space of admissible parameters will be denoted by D . A commonly used benchmark problem for initial value identification in advective-diffusive transport is illustrated in Fig. 2 (see also [19,20]). Note that the transport depends on a wind vector field, which is assumed to be smooth, bounded and divergence free, i.e., $\nabla \cdot \mathbf{v} = 0$. According to the relation of the wind vector and the outward-pointing boundary normal \mathbf{n} , the domain boundary is subdivided into the outflow boundary $\Gamma_+ \subset \partial\Omega, \mathbf{v} \cdot \mathbf{n} > 0$, the inner boundary $\Gamma_0 \subset \partial\Omega, \mathbf{v} \cdot \mathbf{n} = 0$ and the inflow boundary $\Gamma_- \subset \partial\Omega, \mathbf{v} \cdot \mathbf{n} < 0$, comparable to [38]. Finally, we

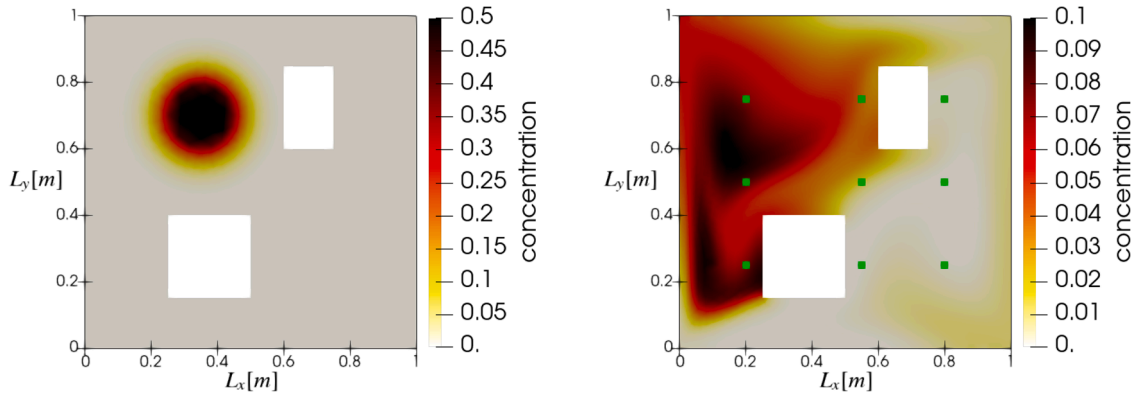


Fig. 2. Illustration of commonly used benchmark problem [19]. Ground truth initial condition m_1 (left), and ground truth concentration field at $t = 5$ s (solution of (P_K)), as well as measurement positions marked by nine green squares (right). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

emphasize that under suitable regularity assumptions on the problem data, i.e. the wind vector field \mathbf{v} as well as the parameter space, (P_K) admits a unique solution u , in an appropriate weak sense, for given m , by standard parabolic regularity theory (see, e.g., [39]). We note that significantly less regularity is required for the parameter m_1 compared to m_C . For this reason, we may also admit Radon measures as initial conditions and, in particular, finite combinations of Dirac-Delta functionals.

Next, we describe the observation setup. For this purpose, let $T_0 > 0$ be a fixed time point and consider a finite number of space-time points $(t_i^{\text{obs}}, x_i^{\text{obs}})$, $i = 1, \dots, N_{\text{obs}}$, with $t_i^{\text{obs}} \in [T_0, T)$ and $x_i^{\text{obs}} \in \Omega$. The discrete space-time observation points are covered by open balls of radius $r_0 > 0$, defining the set $\bar{\Omega}_o := \bigcup_{i=1}^{N_{\text{obs}}} B_{r_0}(x_i^{\text{obs}}) \subset \Omega$. Choosing r_0 sufficiently small ensures that the solution to (P_K) admits a continuous representative on $[T_0, T] \times \bar{\Omega}_o$. This gives rise to a linear and bounded *space-time observation operator*

$$B : C([T_0, T] \times \bar{\Omega}_o) \rightarrow \mathbb{R}^{N_{\text{obs}}}, \quad u \mapsto \sum_{i=1}^{N_{\text{obs}}} \delta_{(t_i^{\text{obs}}, x_i^{\text{obs}})}(u) e_i = \left(u(t_i^{\text{obs}}, x_i^{\text{obs}}) \right)_{i=1}^{N_{\text{obs}}}$$

where $\delta_{(t_i^{\text{obs}}, x_i^{\text{obs}})}$ denotes the Dirac-Delta functional associated to $(t_i^{\text{obs}}, x_i^{\text{obs}})$ and $\{e_i\}_{i=1}^{N_{\text{obs}}}$ is the canonical basis of $\mathbb{R}^{N_{\text{obs}}}$. Consequently, we define the linear and continuous *parameter-to-state* mapping $\mathcal{K} : D \rightarrow C([T_0, T] \times \bar{\Omega}_o)$ as well as the *parameter-to-observable* operator $F : D \rightarrow \mathbb{R}^{N_{\text{obs}}}$ via $\mathcal{K}(m) = u$ and $F(m) = B \circ \mathcal{K}(m)$, respectively. Given a *misfit* vector $\mathbf{y} \in \mathbb{R}^{N_{\text{obs}}}$, e.g., $\mathbf{y} = 1/\sigma^2 (B(u) - \mathbf{d})$ with measurements \mathbf{d} recorded under white noise, the associated *misfit-to-adjoint* map is given by $\mathcal{Q}(\mathbf{y}) = q$, where q is the solution to the final value problem

$$\begin{aligned} -q_t - \kappa \Delta q - \text{div}(q\mathbf{v}) &= \sum_{i=1}^{N_{\text{obs}}} y_i \delta_{(t_i^{\text{obs}}, x_i^{\text{obs}})} && \text{in } (0, T) \times \Omega, \\ (\mathbf{v}q + \kappa \nabla q) \cdot \mathbf{n} &= 0 && \text{on } (0, T) \times (\Gamma_+ \cup \Gamma_0), \\ q &= 0 && \text{on } (0, T) \times \Gamma_-, \\ q(T, \cdot) &= 0 && \text{in } \Omega. \end{aligned} \tag{P_Q}$$

We illustrate it in Fig. 3.

2.2. A sparse inversion framework for contaminant release

For the specific context of airborne contaminant transport, it is reasonable to assume that sources are confined to a finite number of small, localized areas. In the present work, we incorporate this prior knowledge by modeling the sought-for sources, the ground-truth, as

$$m_1^\dagger = \sum_{i=1}^{N_1^\dagger} \lambda_i^{\text{I},\dagger} S_1(x_i^{\text{I},\dagger}, \cdot), \quad m_C^\dagger = \sum_{j=1}^{N_C^\dagger} \lambda_j^{\text{C},\dagger} S_C(x_j^{\text{C},\dagger}, \cdot), \quad \text{where } (\mathbf{x}^{\text{I},\dagger}, \boldsymbol{\lambda}^{\text{I},\dagger}) \in \bar{\Omega}^{N_1^\dagger} \times \mathbb{R}_{>0}^{N_1^\dagger}, (\mathbf{x}^{\text{C},\dagger}, \boldsymbol{\lambda}^{\text{C},\dagger}) \in \bar{\Omega}^{N_C^\dagger} \times \mathbb{R}_{>0}^{N_C^\dagger}$$

denote unknown location-intensity pairs while S_1 as well as S_C are given shape functions, respectively. For a discussion of examples relevant to the present paper, we refer to Section 2.3. Here, and with a slight abuse of notation, we interpret elements in $\bar{\Omega}^N$, $N \in \mathbb{N}$, as matrices whose columns are points in $\bar{\Omega}$.

At this point, we emphasize that we neither assume knowledge of the positions and amplitudes nor of the number of sources. In contrast to the setting of Section 2.4, this leads to a highly nonlinear parameter identification problem. We alleviate these difficulties,

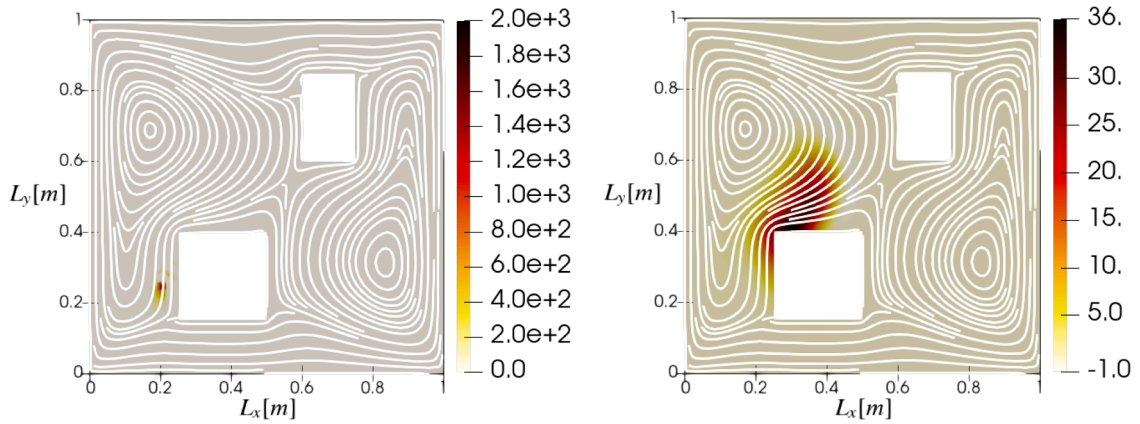


Fig. 3. Adjoint transport problem. Example right-hand side of (P_Q) , i.e., a sensor misfit of “1” applied at $(t_i^{obs}, x_i^{obs}) = (5 \text{ s}, [0.2 \text{ m}, 0.25 \text{ m}])$ (left), solution field of (P_Q) evaluated at $t = 0$.

by first noting that

$$m_I^\dagger = \int_{\Omega} S_I(x, \cdot) d\mu_I^\dagger(x), \quad m_C^\dagger = \int_{\Omega} S_C(x, \cdot) d\mu_C^\dagger(x), \quad \text{where} \quad \mu_I^\dagger = \sum_{i=1}^{N_I^\dagger} \lambda_i^{I,\dagger} \delta_{x_i^{I,\dagger}}, \quad \mu_C^\dagger = \sum_{j=1}^{N_C^\dagger} \lambda_j^{C,\dagger} \delta_{x_j^{C,\dagger}} \tag{1}$$

are conic combinations of Dirac-Delta functionals. Subsequently, we propose to recover the representing measures $(\mu_I^\dagger, \mu_C^\dagger)$ rather than the sources $(m_I^\dagger, m_C^\dagger)$ by solving

$$\min_{\mu_I, \mu_C \in \mathcal{M}^+(\bar{\Omega})} J(\mu_I, \mu_C) := \left[1/(2\sigma^2) \|F(m_I(\mu_I), m_C(\mu_C)) - d\|_{\mathbb{R}^{N_{obs}}}^2 + \alpha (\mu_I(\bar{\Omega}) + \mu_C(\bar{\Omega})) \right], \tag{P_M}$$

where $\alpha > 0$ is a regularization parameter, σ characterizes the measurement noise and

$$m_I(\mu_I) = \int_{\Omega} S_I(x, \cdot) d\mu_I(x), \quad m_C(\mu_C) = \int_{\Omega} S_C(x, \cdot) d\mu_C(x). \tag{2}$$

Here, $\mathcal{M}^+(\bar{\Omega})$ denotes the convex cone of positive Radon measures on $\bar{\Omega}$ and $\mu(\bar{\Omega})$ is the total variation of $\mu \in \mathcal{M}^+(\bar{\Omega})$. On the one hand, we emphasize that the mappings $m_I(\cdot)$ and $m_C(\cdot)$ defined in (2) are linear. Consequently, the optimization problem (P_M) is a convex minimization problem and the existence of minimizers can be proven via the direct method in the calculus of variations. On the other hand, sparse measures, i.e. arbitrary conic combinations of Dirac-Delta functionals,

$$\mathcal{M}_N^+(\bar{\Omega}) := \left\{ \mu[x, \lambda] = \sum_{j=1}^N \lambda_j \delta_{x_j} \mid x \in \bar{\Omega}^N, \lambda \in \mathbb{R}_+^N, N \in \mathbb{N} \right\} \text{ for which } \mu[x, \lambda](\Omega) = |\lambda|_{\rho^1} := \sum_{j=1}^N |\lambda_j|$$

holds, form a proper, but not suitably closed subset of $\mathcal{M}^+(\bar{\Omega})$. Nevertheless, the following theorem establishes the existence of sparse minimizers to (P_M) , thereby bridging the gap between our convex optimization framework and the structural assumption on the ground truth. Moreover, it provides a tool to verify optimality. For this purpose, we introduce the *lifted-parameter-to-observable map* \hat{F} , defined by $\hat{F}(\mu_I, \mu_C) = F(m_I(\mu_I), m_C(\mu_C))$, along with its formal pre-dual $\hat{F}^* : \mathbb{R}^{N_{obs}} \rightarrow C(\bar{\Omega}) \times C(\bar{\Omega})$. Given $y \in \mathbb{R}^{N_{obs}}$, the latter satisfies

$$y^\top \hat{F}(\mu_I, \mu_C) = \int_{\bar{\Omega}} \varphi_I(x) d\mu_I(x) + \int_{\bar{\Omega}} \varphi_C(x) d\mu_C(x) \tag{3}$$

where $(\varphi_I, \varphi_C) := \hat{F}^* y$ satisfies

$$[\hat{F}^* y](x) = \left(\int_{\bar{\Omega}} S_I(x, z) Q(y)(0, z) d\Omega(z), \int_0^T \int_{\bar{\Omega}} S_C(x, z) Q(y)(t, z) d\Omega(z) dt \right) \text{ for all } x \in \bar{\Omega}. \tag{4}$$

Here we recall that $Q(y)$ is given by the solution to an advection-diffusion problem with inverted wind direction v (cf. (P_Q)).

Theorem 1. *Problem (P_M) admits at least one minimizing pair $(\bar{\mu}_I, \bar{\mu}_C) \in \mathcal{M}_N^+(\bar{\Omega})^2$, i.e.,*

$$\bar{\mu}_I = \mu_I[\bar{x}^I, \bar{\lambda}^I], \quad \bar{\mu}_C = \mu_C[\bar{x}^C, \bar{\lambda}^C], \quad \text{where} \quad (\bar{x}^I, \bar{\lambda}^I) \in \bar{\Omega}^{\bar{N}_I} \times \mathbb{R}_{>0}^{\bar{N}_I}, \quad (\bar{x}^C, \bar{\lambda}^C) \in \bar{\Omega}^{\bar{N}_C} \times \mathbb{R}_{>0}^{\bar{N}_C}, \quad \bar{N}_I, \bar{N}_C \geq 0 \tag{5}$$

as well as $\bar{N}_I + \bar{N}_C \leq N_{obs}$. Moreover, the following equivalence holds:

- A pair $(\bar{\mu}_I, \bar{\mu}_C)$ of the form (5) is a minimizer of Problem (P_M) .

- For the dual variables

$$(\bar{\varphi}_I, \bar{\varphi}_C) := -\hat{F}^* \bar{\mathbf{y}} \in C(\bar{\Omega}) \times C(\bar{\Omega}), \quad \text{where } \bar{\mathbf{y}} = 1/\sigma^2 (\hat{F}(\bar{\mu}_I, \bar{\mu}_C) - \mathbf{d}), \tag{6}$$

the following conditions hold:

$$\max_{x \in \bar{\Omega}} \bar{\varphi}_I(x) \leq \alpha, \quad \max_{x \in \bar{\Omega}} \bar{\varphi}_C(x) \leq \alpha, \quad \bar{\varphi}_I(\bar{x}_i^I) = \alpha, \quad \bar{\varphi}_C(\bar{x}_j^C) = \alpha \quad \text{for all } i = 1, \dots, \bar{N}_I, \quad j = 1, \dots, \bar{N}_C. \tag{7}$$

Proof. We start by discussing the claimed equivalence. For this purpose, set

$$f(\mu_I, \mu_C) := 1/(2\sigma^2) \|F(m_I(\mu_I), m_C(\mu_C)) - \mathbf{d}\|_{\mathbb{R}^{N_{\text{obs}}}}^2$$

By means of the chain rule as well as (3), f is Gateaux-differentiable at every pair $(\mu_I, \mu_C) \in \mathcal{M}^+(\Omega)^2$ and there holds

$$f'(\mu_I, \mu_C)(\delta\mu_I, \delta\mu_C) = -\left(\int_{\bar{\Omega}} \varphi_I(x) \, d\delta\mu_I(x) + \int_{\bar{\Omega}} \varphi_C(x) \, d\delta\mu_C(x)\right),$$

where $(\delta\mu_I, \delta\mu_C)$ are two (signed) Radon measures and where we set

$$(\varphi_I, \varphi_C) := -\hat{F}^* \mathbf{y} \in C(\bar{\Omega}) \times C(\bar{\Omega}), \quad \text{with } \mathbf{y} = 1/\sigma^2 (\hat{F}(\mu_I, \mu_C) - \mathbf{d}). \tag{8}$$

Since $(\mathcal{P}_{\mathcal{M}})$ is a convex problem, an arbitrary pair $(\bar{\mu}_I, \bar{\mu}_C) \in \mathcal{M}^+(\Omega)^2$ is a minimizer if and only if the associated dual variables, (6), satisfy

$$\max_{x \in \bar{\Omega}} \bar{\varphi}_I(x) \leq \alpha, \quad \max_{x \in \bar{\Omega}} \bar{\varphi}_C(x) \leq \alpha, \quad \text{supp } \bar{\mu}_I \subset \{x \in \bar{\Omega} \mid \bar{\varphi}_I(x) = \alpha\}, \quad \text{supp } \bar{\mu}_C \subset \{x \in \bar{\Omega} \mid \bar{\varphi}_C(x) = \alpha\}$$

where $\text{supp } \bar{\mu}_I, \text{supp } \bar{\mu}_C$ denote the support of $\bar{\mu}_I$ and $\bar{\mu}_C$, respectively. This can be derived, e.g., analogously to [40, Proposition 3.5, Lemma 3.6]. We conclude the claimed equivalence noting that for measures $(\bar{\mu}_I, \bar{\mu}_C)$ of the form (5), there holds

$$\text{supp } \bar{\mu}_I = \{\bar{x}_i^I \mid i = 1, \dots, \bar{N}_I\}, \quad \text{supp } \bar{\mu}_C = \{\bar{x}_j^C \mid j = 1, \dots, \bar{N}_C\}.$$

It thus remains to show that sparse minimizers with $\bar{N}_I + \bar{N}_C \leq N_{\text{obs}}$ exist. By applying an abstract convex representer theorem, see e.g. [41,42], $(\mathcal{P}_{\mathcal{M}})$ admits at least one solution of the form

$$(\bar{\mu}_I, \bar{\mu}_C) = \sum_{j=1}^{\bar{N}} \bar{\lambda}_j (\mu_I^j, \mu_C^j), \quad \text{with } \bar{\lambda}_j \geq 0, \quad \bar{N} \leq N_{\text{obs}} \quad \text{and} \quad (\mu_I^j, \mu_C^j) \in \text{Ext}(B).$$

Here $\text{Ext}(B)$ denotes the extremal points of the weak*-compact set

$$B := \{(\mu_I, \mu_C) \in \mathcal{M}^+(\bar{\Omega})^2 \mid \alpha(\mu_I(\bar{\Omega}) + \mu_C(\bar{\Omega})) \leq 1\}.$$

Following arguments similar to [43, Lemma 3.2], we obtain

$$\text{Ext}(B) = \{(0, 0)\} \cup \left\{ \frac{1}{\alpha} (\delta_x, 0) \mid x \in \bar{\Omega} \right\} \cup \left\{ \frac{1}{\alpha} (0, \delta_x) \mid x \in \bar{\Omega} \right\},$$

which proves the sparsity of the minimizer, see (Eq. (5)). \square

Remark 1. We emphasize that this is an existence but not a uniqueness result. Quite the contrary, since F has finite rank, Problem $(\mathcal{P}_{\mathcal{M}})$ can admit a multitude of solutions, some of which might even be non-sparse, i.e. elements of $\mathcal{M}^+(\bar{\Omega}) \setminus \mathcal{M}_N^+(\bar{\Omega})$. Additionally, identifying the correct number of sources, $\bar{N}_I = N_I^\dagger$ and $\bar{N}_C = N_C^\dagger$, together with quantitative reconstruction results requires strong nondegeneracy conditions, small measurement noise as well as a noise-adapted regularization parameter α . We refer, e.g., to [33] where these topics are addressed in detail for simpler deconvolution tasks.

We close this section by noting that the support condition in (7) can be used to verify optimality of sparse candidate measures. Moreover, for sparse minimizers, the objective functional decomposes as

$$J(\bar{\mu}_I, \bar{\mu}_C) = J(\mu_I[\bar{x}^I, \bar{\lambda}^I], \mu_C[\bar{x}^C, \bar{\lambda}^C]) = 1/(2\sigma^2) \|\hat{F}(\mu_I[\bar{x}^I, \bar{\lambda}^I], \mu_C[\bar{x}^C, \bar{\lambda}^C]) - \mathbf{d}\|_{\mathbb{R}^{N_{\text{obs}}}}^2 + \alpha \left(|\bar{\lambda}^I|_{\ell_1} + |\bar{\lambda}^C|_{\ell_1} \right),$$

where $\hat{F}(\mu_I[\bar{x}^I, \bar{\lambda}^I], \mu_C[\bar{x}^C, \bar{\lambda}^C]) = \sum_{i=1}^{\bar{N}_I} \bar{\lambda}_i^I \hat{F}(\delta_{\bar{x}_i^I}, 0) + \sum_{j=1}^{\bar{N}_C} \bar{\lambda}_j^C \hat{F}(0, \delta_{\bar{x}_j^C})$ is linear w.r.t the source intensities. Hence, once optimal locations have been identified, the associated intensities are obtained as minimizers of a finite-dimensional, convex problem.

Corollary 1. Consider a minimizing pair $(\bar{\mu}_I, \bar{\mu}_C)$ of the form (5). Then there holds

$$(\bar{\lambda}^I, \bar{\lambda}^C) \in \arg \min_{\substack{\lambda^I \in \mathbb{R}_{\geq 0}^{\bar{N}_I}, \lambda^C \in \mathbb{R}_{\geq 0}^{\bar{N}_C}}} \left[1/(2\sigma^2) \|\hat{F}(\mu_I[\bar{x}^I, \lambda^I], \mu_C[\bar{x}^C, \lambda^C]) - \mathbf{d}\|_{\mathbb{R}^{N_{\text{obs}}}}^2 + \alpha \left(|\lambda^I|_{\ell_1} + |\lambda^C|_{\ell_1} \right) \right]. \tag{9}$$

Vice versa, let two arbitrary sparse measures of the form (5) be given, define $(\bar{\varphi}_I, \bar{\varphi}_C)$ according to (6) and assume that the pairs $(\bar{x}^I, \bar{\lambda}^I) \in \bar{\Omega}^{\bar{N}_I} \times \mathbb{R}_{>0}^{\bar{N}_I}, (\bar{x}^C, \bar{\lambda}^C) \in \bar{\Omega}^{\bar{N}_C} \times \mathbb{R}_{>0}^{\bar{N}_C}$ satisfy (9). Then we have

$$(\bar{\mu}_I, \bar{\mu}_C) \text{ is a minimizer of } (\mathcal{P}_{\mathcal{M}}) \Leftrightarrow \max_{x \in \bar{\Omega}} \bar{\varphi}_I(x) \leq \alpha \text{ and } \max_{x \in \bar{\Omega}} \bar{\varphi}_C(x) \leq \alpha.$$

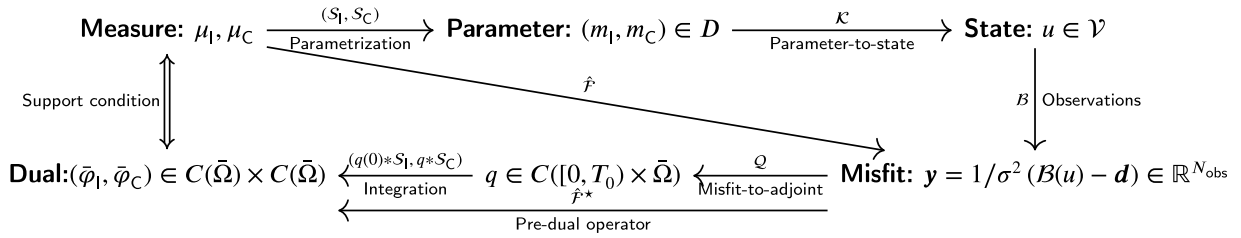


Fig. 4. Summary of the inverse problem with sparse regularization.

Proof. With a slight abuse of notation, define the mapping

$$f(\lambda^I, \lambda^C) := 1/(2\sigma^2) \|\hat{F}(\mu_1[\bar{x}^I, \lambda^I], \mu_C[\bar{x}^C, \lambda^C]) - d\|_{\mathbb{R}^{N_{\text{obs}}}}^2$$

and note that

$$J(\mu_1[\bar{x}^I, \lambda^I], \mu_C[\bar{x}^C, \lambda^C]) = f(\lambda^I, \lambda^C) + \alpha (|\lambda^I|_{\ell_1} + |\lambda^C|_{\ell_1}).$$

For a minimizing pair $(\bar{\mu}_1, \bar{\mu}_C)$ of the form (5), we now estimate

$$J(\mu_1[\bar{x}^I, \bar{\lambda}^I], \mu_C[\bar{x}^C, \bar{\lambda}^C]) = J(\bar{\mu}_1, \bar{\mu}_C) \leq J(\mu_1[\bar{x}^I, \lambda^I], \mu_C[\bar{x}^C, \lambda^C]) \quad \text{for all } \lambda^I \in \mathbb{R}_{\geq 0}^{N_I}, \lambda^C \in \mathbb{R}_{\geq 0}^{N_C}$$

from which we conclude (9). For the second statement, and again by means of the chain rule, note that f is Gateaux-differentiable at every pair $(\lambda^I, \lambda^C) \in \mathbb{R}_{\geq 0}^{N_I} \times \mathbb{R}_{\geq 0}^{N_C}$ and there holds

$$f'(\lambda^I, \lambda^C)(\delta\lambda^I, \delta\lambda^C) = - \left(\sum_{i=1}^{\bar{N}_I} \delta\lambda_i^I \varphi_1(\bar{x}_i^I) + \sum_{j=1}^{\bar{N}_C} \delta\lambda_j^C \varphi_C(\bar{x}_j^C) \right)$$

for all directions $(\delta\lambda^I, \delta\lambda^C) \in \mathbb{R}^{\bar{N}_I} \times \mathbb{R}^{\bar{N}_C}$ and where we define

$$(\varphi_1, \varphi_C) := -\hat{F}^* y \in C(\bar{\Omega}) \times C(\bar{\Omega}), \quad \text{with } y = 1/\sigma^2 (\hat{F}(\mu_1[\bar{x}^I, \lambda^I], \mu_C[\bar{x}^C, \lambda^C]) - d).$$

analogously to Theorem 1. Since (9) is also a convex minimization problem, optimality of $(\bar{\lambda}^I, \bar{\lambda}^C)$ is equivalent to

$$\bar{\varphi}_1(\bar{x}_i^I) = \alpha, \quad \bar{\varphi}_C(\bar{x}_j^C) = \alpha \quad \text{for all } i = 1, \dots, \bar{N}_I, \quad j = 1, \dots, \bar{N}_C.$$

which follows by standard first-order theory, see e.g. [44]. Comparing with the second part of (1), the claimed equivalence follows. \square

The connection between the different components of the sparse regularization approach is illustrated in Fig. 4.

2.3. Modeling

of contaminant sources In this paper, three models of contaminant release are presented and the type of the problem-defining parameter $m = (m_1, m_C)$ is specified accordingly. Numerical examples of the three types are illustrated in Fig. 5. The first model is a radial basis function given in closed form expression as

$$S^1(x_s, r, y) = \min \left\{ 0.5, \exp \left(-\ln(\epsilon) \|y - x_s\|_2^2 / r^2 \right) \right\}. \tag{10}$$

Here $x_s \in \Omega$ is the center, $r > 0$ is the radius of the source and $\epsilon > 0$ is a given threshold. The model S^1 is employed for inverse problems with L^2 -regularization [19,20] and is a highly flexible approach that can be tailored to the specific requirements of the problem. The second model S^2 for a contaminant release is constructed as solution of an elliptic PDE with a Dirac distribution with center x_s on the right-hand side:

$$S^2(x_s, \eta, \gamma, y) = m_{x_s}(y), \quad \text{with } m_{x_s} \text{ given by } \begin{cases} (\eta I - \gamma \Delta) m_{x_s} & = \delta_{x_s} & \text{in } \Omega, \\ \gamma \nabla m_{x_s} \cdot \mathbf{n} + \beta m_{x_s} & = 0 & \text{on } \partial\Omega, \end{cases} \tag{11}$$

where the scalars η, γ and $\beta(\eta, \gamma)$ are used to adjust the initial condition to the specific application. The last formulation

$$S^3(x_s) = \delta_{x_s} \tag{12}$$

is particularly advantageous for large-scale problems in which the source has minimal support.

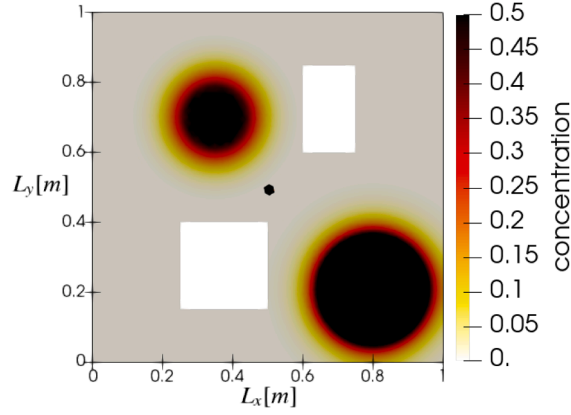


Fig. 5. Mathematical description of contaminant release. Closed form expression $S^1([0.35, 0.7], 0.26, \cdot)$, elliptic PDE solution $S^2([0.8, 0.2], 1.0, 0.001, \cdot)$, and very localized Dirac distribution $S^3([0.5, 0.5], \cdot)$.

2.4. Quadratic regularization in source space

A natural alternative to the parametrization ansatz outlined above is to consider a regularization term directly acting on the source space D leading to problems of the form

$$\min_{(m_1, m_C) \in D} \left[1/(2\sigma^2) \|\mathcal{F}(m_1, m_C) - \mathbf{d}\|_{\mathbb{R}^{N_{\text{obs}}}}^2 + \mathcal{R}(m) \right]. \quad (\mathcal{P}_D)$$

In this regard, a popular approach is to consider $D \subset L^2(\Omega) \times L^2(\Omega)$ and choose quadratic regularization terms of the form $\mathcal{R}(m) := 1/2 \|\mathcal{A}(m - m_{\text{prior}})\|_{L^2(\Omega)^2}^2$ for a reference m_{prior} and, e.g., a differential operator \mathcal{A} . This admits a Bayesian interpretation: Assuming that $m \sim \mathcal{N}(m_{\text{prior}}, \mathcal{A}^{-2})$ with \mathcal{A}^{-2} a trace-class operator and m_{prior} sufficiently regular as well as a linear measurement model $d = \mathcal{F}(m) + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \sigma^2 \text{Id})$, (\mathcal{P}_D) corresponds to the computation of the maximum a posteriori estimator. Since this problem is convex, the associated necessary and sufficient first-order optimality condition implies $\bar{m} = \mathcal{H}^{-1}((1/\sigma^2)\mathcal{F}^* \mathbf{d} + \mathcal{A}^2 m_{\text{prior}})$ where $\mathcal{H} = (1/\sigma^2)\mathcal{F}^* \mathcal{F} + \mathcal{A}^2$. For a given realization of the measurement data, an estimate can thus be obtained by solving this linear equation using an inexact Newton CG method [45] for a suitable discretization [19,20]. Due to the trace-class assumption, the eigenvalues of the positive-definite matrix \mathcal{H} decay rapidly enabling the use of reduced models. This property makes real-time parameter identification feasible through an offline-online decomposition strategy. Restricting ourselves to the case of initial value identification, i.e., assuming $m_C = 0$, we choose a Laplacian-like operator $\mathcal{A} := (\eta I - \gamma \Delta)$, cf. [19,20], together with Robin boundary conditions $\gamma \nabla m_1 \cdot \mathbf{n} + \beta m_1 = 0$ in $(0, T) \times \partial\Omega$, with constants as in [46]. An available implementation of the described method, see [47], is employed as a benchmark for the computations in Section 5.

3. Primal-dual-active-point algorithm for source identification

Finally, we briefly discuss the numerical solution of $(\mathcal{P}_{\mathcal{M}})$ by adapting the Primal-Dual-Active-Point (PDAP) method as described in [34]. In this regard, and in accordance with the original work, we follow an optimize-then-discretize philosophy and formulate the algorithm on the function space level before applying it to suitable discretizations described in Section 4. Conceptually, PDAP is a greedy algorithm which relies on sparse iterates and exploits the existence of a pair of sparse minimizers, see Theorem 1, by alternating between proposing new candidate sources based on the violation of the optimality condition (7) and removing unnecessary points motivated by the observations in Corollary 1. More in detail, it constructs a sequence of parametrized iterates

$$\mu_1^k = \mu_1[\mathbf{x}_k^I, \boldsymbol{\lambda}_k^I], \quad \mu_C^k = \mu_C[\mathbf{x}_k^C, \boldsymbol{\lambda}_k^C] \quad \text{where} \quad (\mathbf{x}_k^I, \boldsymbol{\lambda}_k^I) \in \bar{\Omega}^{N_I^k} \times \mathbb{R}_{>0}^{N_I^k}, \quad (\mathbf{x}_k^C, \boldsymbol{\lambda}_k^C) \in \bar{\Omega}^{N_C^k} \times \mathbb{R}_{>0}^{N_C^k}$$

with the property that $(\boldsymbol{\lambda}_k^I, \boldsymbol{\lambda}_k^C)$ is a minimizing pair of

$$\min_{\boldsymbol{\lambda}^I \in \mathbb{R}_{\geq 0}^{N_I^k}, \boldsymbol{\lambda}^C \in \mathbb{R}_{\geq 0}^{N_C^k}} \left[1/(2\sigma^2) \|\hat{\mathcal{F}}(\mu_1[\mathbf{x}_k^I, \boldsymbol{\lambda}_k^I], \mu_C[\mathbf{x}_k^C, \boldsymbol{\lambda}_k^C]) - \mathbf{d}\|_{\mathbb{R}^{N_{\text{obs}}}}^2 + \alpha \left(|\boldsymbol{\lambda}_k^I|_{\ell_1} + |\boldsymbol{\lambda}_k^C|_{\ell_1} \right) \right] \quad (\mathcal{P}(\mathbf{x}^I, \mathbf{x}^C))$$

for the particular choice of $(\mathbf{x}^I, \mathbf{x}^C) = (\mathbf{x}_k^I, \mathbf{x}_k^C)$. Referring to the terminology of the original paper, cf. Algorithm 1 in [34], \mathbf{x}_k^I and \mathbf{x}_k^C , respectively, collect the currently “active” points comprising the iterates. In each iteration, we start by computing the current misfit and dual variables

$$\mathbf{y}_k = 1/\sigma^2 (\hat{\mathcal{F}}(\mu_1^k, \mu_C^k) - \mathbf{d}) \in \mathbb{R}^{N_{\text{obs}}}, \quad (\varphi_1^k, \varphi_C^k) = -\hat{\mathcal{F}}^* \mathbf{y}_k \in C(\bar{\Omega}) \times C(\bar{\Omega})$$

determine global maximizers $\hat{\mathbf{x}}_k^I$ and $\hat{\mathbf{x}}_k^C$ of the latter and check for optimality according to Corollary 1. If this is not achieved up to a tolerance, we append $\hat{\mathbf{x}}_k^I$ and/or $\hat{\mathbf{x}}_k^C$ to the location vectors yielding $(\mathbf{x}_{k+1/2}^I, \mathbf{x}_{k+1/2}^C)$, respectively. Subsequently, we adjust the iterate

by optimizing the intensities

$$\mu_I^{k+1} = \mu_I \left[\mathbf{x}_{k+1/2}^I, \boldsymbol{\lambda}_{k+1/2}^I \right], \quad \mu_C^{k+1} = \mu_C \left[\mathbf{x}_{k+1/2}^C, \boldsymbol{\lambda}_{k+1/2}^C \right], \quad \left(\boldsymbol{\lambda}_{k+1/2}^I, \boldsymbol{\lambda}_{k+1/2}^C \right) \in \arg \min \left(\mathcal{P} \left(\mathbf{x}_{k+1/2}^I, \mathbf{x}_{k+1/2}^C \right) \right).$$

Finally, location and intensity vectors are pruned by removing entries corresponding to Dirac-Delta functionals with zero weights. We summarize this procedure in **Algorithm 1**. An example for a typical iteration for the simplified case of initial value identification is displayed in **Fig. 6**.

By adapting the results in [34], we point out that (μ_I^k, μ_C^k) weak*-converges, on subsequences, towards minimizers of Problem (\mathcal{P}_M) with a global, sublinear worst-case rate of convergence, i.e. there holds

$$r_J(\mu_I^k, \mu_C^k) := J(\mu_I^k, \mu_C^k) - \min_{\mu_I, \mu_C \in \mathcal{M}^+(\Omega)} J(\mu_I, \mu_C) \leq C/k \quad \text{for all } k \geq 1$$

and some $C > 0$. Additionally, the dual variables satisfy

$$0 \leq \max_{x \in \Omega} \varphi_I^k(x) - \alpha \rightarrow 0, \quad 0 \leq \max_{x \in \Omega} \varphi_C^k(x) - \alpha \rightarrow 0.$$

Hence, for $\text{tol} > 0$, **Algorithm 1** converges in a finite number of steps K and we have

$$r_J(\mu_I^k, \mu_C^k) \leq \frac{J(0, 0)}{\alpha} \text{tol} \tag{13}$$

assuming that the subproblems in Step 3 and 6 of **Algorithm 1** are solved to optimality, see, mutatis mutandis, [48, Proposition 4.12] and the discussion on [48, p. 83]. Moreover, asymptotic linear decrease, i.e. $r_J(\mu_I^{k+1}, \mu_C^{k+1}) \leq \zeta r_J(\mu_I^k, \mu_C^k)$, for all k large enough and some $\zeta \in (0, 1)$, can be ensured under additional structural assumptions related to a suitable no-gap second-order condition, [49]. The theoretical upper bound on the support size from **Theorem 1**, $N_I^k + N_C^k \leq N_{\text{obs}}$ can be ensured throughout all iterations by augmenting the method with additional sparsification strategies, see e.g., [40, Algorithm 2]. However, in practice, this is not required since the method usually terminates before this threshold is even reached. We briefly comment on the practical realization of **Algorithm 1**.

Algorithm 1 Primal-dual-active-point-strategy for source identification.

Require: $S_I, S_C, \mu_I^0 = \mu_C^0 = 0$, empty matrices and vectors $\mathbf{x}_k^I, \boldsymbol{\lambda}_k^I, \mathbf{x}_k^C, \boldsymbol{\lambda}_k^C$

for $k = 0, 1, 2 \dots$ **do**

1. Given $\mu_I^k = \mu_I[\mathbf{x}_k^I, \boldsymbol{\lambda}_k^I]$, $\mu_C^k = \mu_C[\mathbf{x}_k^C, \boldsymbol{\lambda}_k^C]$, compute $\mathbf{y}_k = 1/\sigma^2 (\hat{\mathcal{F}}(\mu_I^k, \mu_C^k) - \mathbf{d})$.

2. Compute convolution

$$(\varphi_I^k, \varphi_C^k) = - \left(\int_{\Omega} S_I(\cdot, z) Q(\mathbf{y}_k)(0, z) \, d\Omega(z), \int_0^T \int_{\Omega} S_C(\cdot, z) Q(\mathbf{y}_k)(t, z) \, d\Omega(z) \, dt \right).$$

3. Determine global maxima

$$\hat{\mathbf{x}}_k^I \in \arg \max_{x \in \Omega} \varphi_I^k(x), \quad \hat{\mathbf{x}}_k^C \in \arg \max_{x \in \Omega} \varphi_C^k(x).$$

4. Set $\mathbf{x}_{k+1/2}^I = \mathbf{x}_k^I$, $\mathbf{x}_{k+1/2}^C = \mathbf{x}_k^C$.

5. Append $\hat{\mathbf{x}}_k^I$ to $\mathbf{x}_{k+1/2}^I$ **if** $\varphi_I^k(\hat{\mathbf{x}}_k^I) > \alpha + \text{tol}$, **append** $\hat{\mathbf{x}}_k^C$ to $\mathbf{x}_{k+1/2}^C$ **if** $\varphi_C^k(\hat{\mathbf{x}}_k^C) > \alpha + \text{tol}$

6. Update source intensities

$$\left(\boldsymbol{\lambda}_{k+1/2}^I, \boldsymbol{\lambda}_{k+1/2}^C \right) \in \arg \min \left(\mathcal{P} \left(\mathbf{x}_{k+1/2}^I, \mathbf{x}_{k+1/2}^C \right) \right).$$

7. Update iterates

$$\mu_I^{k+1} = \mu_I \left[\mathbf{x}_{k+1/2}^I, \boldsymbol{\lambda}_{k+1/2}^I \right], \quad \mu_C^{k+1} = \mu_C \left[\mathbf{x}_{k+1/2}^C, \boldsymbol{\lambda}_{k+1/2}^C \right].$$

8. Obtain $\mathbf{x}_{k+1}^I, \boldsymbol{\lambda}_{k+1}^I$ and $\mathbf{x}_{k+1}^C, \boldsymbol{\lambda}_{k+1}^C$ by pruning.

end for

By saving $\hat{\mathcal{F}}(\delta_{\hat{\mathbf{x}}_k^I}, 0)$ and $\hat{\mathcal{F}}(0, \delta_{\hat{\mathbf{x}}_k^C})$ across iterations and due to the linearity of \mathcal{F} , every iteration requires at most two forward simulations as well as one solution of the adjoint equation. No additional PDE solves are required for the realization of $(\mathcal{P}(\mathbf{x}^I, \mathbf{x}^C))$. Concerning the latter, a semismooth Newton method based on the normal-map approach with multidimensional filter globalization

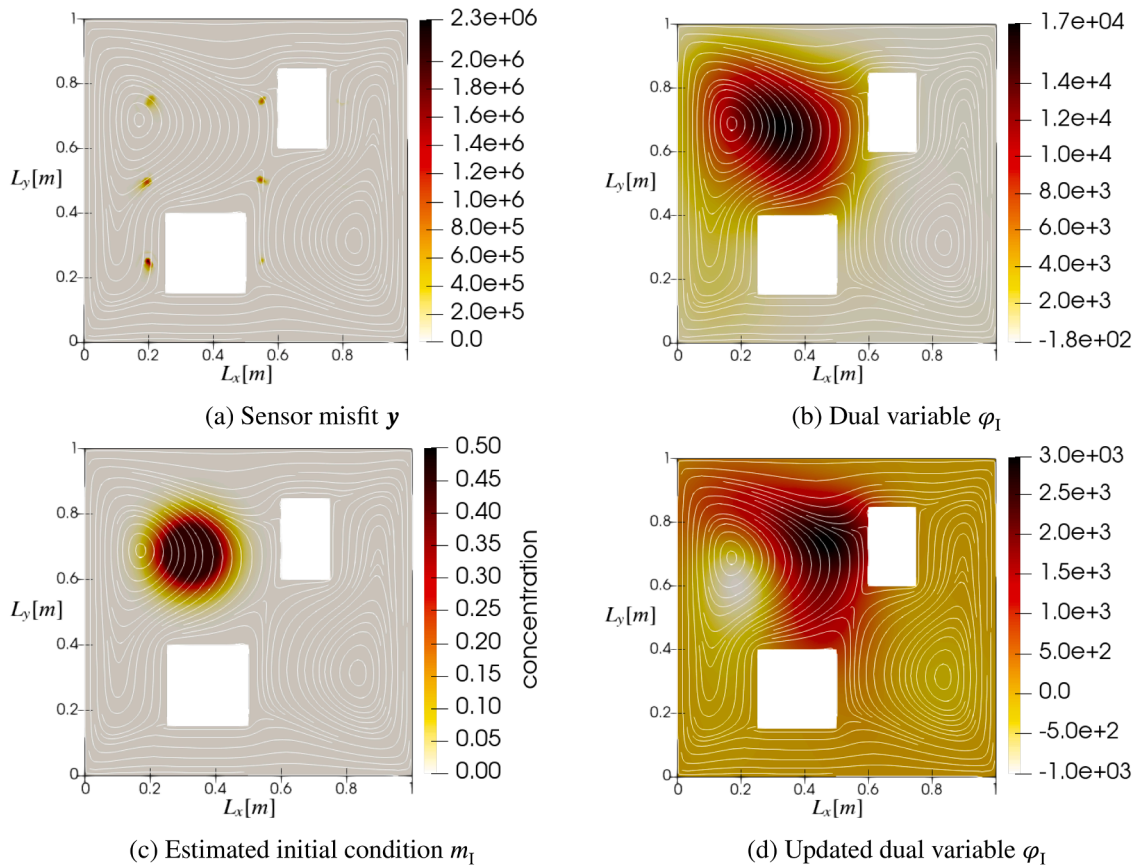


Fig. 6. Iteration of PDAP-algorithm. Starting with sensor misfit y (a) obtained from $(m_1 = 0)$, the dual variable φ_1 is computed (b). Position of $\max(\varphi_1)$ provides a candidate to estimate the initial condition m_1 after the first iteration (c). Reduced value of updated field φ_1 (d) indicates convergence towards stopping criterion $\max(\varphi_1) \leq \alpha$.

is implemented [50], which we warmstart using the current intensity vectors to construct an initial iterate in order to benefit from its potentially fast, local convergence. The Newton solver is stopped once an appropriate residual falls below $\text{tol}_{\text{Newt}} \ll \text{tol}$.

Next, obtaining the dual variables via integration against the shape functions might be expensive if S has large support in Ω . However, for the PDE-based release model S^2 , given in Eq. (11), the integration can be simplified via adjoint calculus leading to a single solve of an elliptic equation, while the Dirac-based formulation of contaminant release S^3 , given in Eq. (12), renders this step trivial. Finally, we acknowledge that the point insertion steps, that is, the computation of global maxima in Step 3 of Algorithm 1, are in general nonconcave maximization problems. Solving solutions that are (approximately) globally optimal is computationally infeasible for general Ω . Approaches to address these challenges can be found in, e.g., [36,51]. In our case, as outlined above, we ultimately apply Algorithm 1 to discretizations of Problem (\mathcal{P}_M) . This also entails replacing $\mathcal{M}^+(\Omega)$ by its discretized counterpart, the cone spanned by Dirac-Deltas in the nodes of the finite elements mesh. This reduces Step 3 of Algorithm 1 to sorting the values of φ_1 and φ_1 , respectively, in the grid nodes, see also Section 4.

4. Finite element discretization of forward and adjoint problem

In this section, we derive the weak formulation of (\mathcal{P}_C) . To this end, we introduce suitable spaces for test functions $H_{\Gamma_-}^{1,2}(\Omega) := \{ \phi_0 \in H^{1,2}(\Omega) \mid \phi_0|_{\Gamma_-} = 0 \}$ in space and time $\mathcal{V} = \{ \phi \mid \phi(t) \in H_{\Gamma_-}^{1,2}(\Omega) \forall t \in (0, T) \}$, cf. [19]. The weak formulation is then given by

$$\int_0^T \int_{\Omega} (u_t + \mathbf{v} \cdot \nabla u) \phi + \kappa \nabla u \cdot \nabla \phi \, d\Omega \, dt + \int_{\Omega} u(0, \cdot) \phi_0 \, d\Omega = \int_0^T \int_{\Omega} m_C \phi \, d\Omega \, dt + \int_{\Omega} m_1 \phi_0 \, d\Omega, \tag{14}$$

$\forall \phi \in \mathcal{V}, \phi_0 \in H_{\Gamma_-}^{1,2}(\Omega)$. Here, the parameter m enters the formulation solely through the right-hand side. Our goal is to discretize this formulation. To this end, we introduce a finite-dimensional subspace $\mathcal{V}_h \subset H_{\Gamma_-}^{1,2}(\Omega)$ of continuous Lagrange nodal basis functions, defined by $\mathcal{V}_h = \text{span}\{\phi_1, \dots, \phi_{n_{\text{dof}}}\}$. To approximate the L^2 -norm, the mass matrix $M \in \mathbb{R}^{n_{\text{dof}} \times n_{\text{dof}}}$ is introduced, with entries $M_{ij} := \int_{\Omega} \phi_i(x) \phi_j(x) \, d\Omega(x)$. An identification between the L^2 -functions, their norm and the discrete elements is required. For this purpose, let

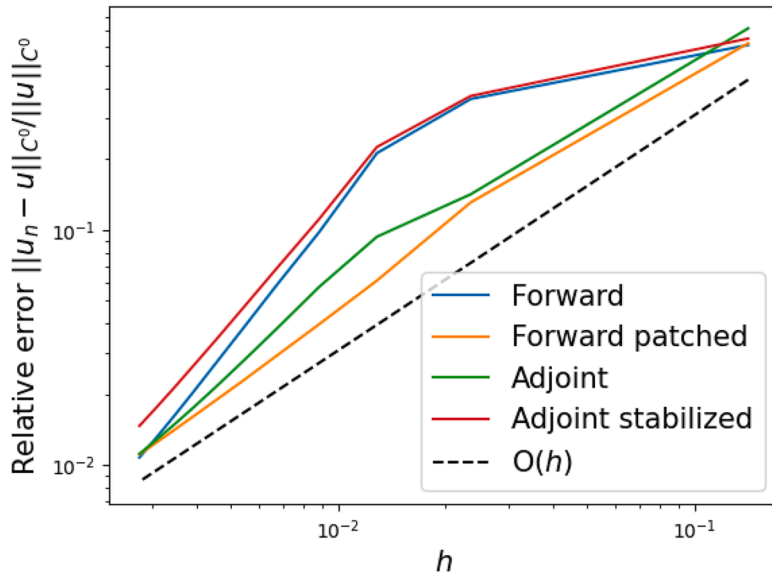


Fig. 7. Asymptotic linear convergence of forward problem, patched forward problem, adjoint problem and stabilized adjoint problem of test case with Dirac initial condition δ_{x_s} on a unit square with the analytical solution $G(x, t) = \exp(-\|x - x_s - vt\|^2 / 4t\kappa) / (4\pi t\kappa)$ in $[0, T] \times \mathbb{R}^2$ for $x_s = [0.5, 0.5]$, $\kappa = 0.001 \text{ m}^2 \text{ s}^{-1}$, $h/\Delta t = \text{const.}$ and $v = [0.1 \text{ m s}^{-1}, 0.1 \text{ m s}^{-1}]$ evaluated at $t = 1 \text{ s}$.

the map $I : \mathbb{R}^{n_{\text{dof}}} \rightarrow \mathcal{V}_h$ be defined by $I(a) = \sum_{i=1}^{n_{\text{dof}}} a_i \phi_i$. Together with the scalar product $\langle a, b \rangle_M := \langle Ma, b \rangle_{\mathbb{R}^{n_{\text{dof}}}}$, the map I becomes an isometry, i.e., $\langle I(a), I(b) \rangle_{L^2(\Omega)} = \langle a, b \rangle_M$. Similarly, the stiffness matrix $K_{ij} = \int_{\Omega} \langle \nabla \phi_i(x), \nabla \phi_j(x) \rangle \text{ d}\Omega(x)$ and the skew-symmetric matrix $V_{ij} = \int_{\Omega} \phi_i(x) \langle \nabla \phi_j(x), v \rangle \text{ d}\Omega(x)$ are defined. Additionally, the well-established SUPG-stabilization technique [52], requires the matrix $S_{ij} = \int_{\Omega} \langle \nabla \phi_i(x), v \rangle \langle \nabla \phi_j(x), v \rangle - \kappa \Delta \phi_j(x) \text{ d}\Omega(x)$. To account for local characteristics of the initial boundary value problem, the stabilization parameter τ is defined as $\tau = \min(h_E^2 / 2\kappa, h_E / \|v\|)$, where $h_E := \sup_{x,y \in E} |x - y|$ denotes the diameter of a finite element E . A more advanced metric-based definition of the stabilization parameter is provided in [53].

The transient problem is addressed using an implicit Euler time-stepping algorithm [54, Equation (10.25)]. We use the approximation $u_t \approx (u^{n+1} - u^n) / \Delta t$ for time instances $(0, \Delta t, \dots, T = \Delta t n_T)$ to find a solution in discrete space-time $\bigoplus_{i=0}^{n_T} \mathcal{V}_h$ of the equation

$$(M + \Delta t V + \Delta t \kappa K + \Delta t \tau S + \tau V^T) u^{n+1} = (M + \tau V^T)(u^n + \Delta t m_C^n), \tag{15}$$

with the initial condition $u_{n=0} = m_I$. The space $\bigoplus_{i=0}^{n_T} \mathcal{V}_h$ represents the temporal product of the spatial ansatz space over all time steps. Function composition gives the discrete linear operator $\mathcal{K}^h : D^h \rightarrow \bigoplus_{i=0}^{n_T} \mathcal{V}_h$ for the chosen time discretization, with the parameters projected into their corresponding finite element representations.

The dual problem for the forward equation has already been derived in (\mathcal{P}_Q) . To solve this terminal-value problem, we apply the change of variables $t = T - \bar{t}$, which transforms the problem (\mathcal{P}_Q) into a well-posed advection-diffusion problem with a prescribed initial condition. The resulting problem can then be discretized using the same methodology as for the forward equation. Under this transformation, only the direction of advection is reversed, i.e., $-v$. See, for example, [55] in the context of the wave equation for a detailed explanation. To obtain the dual or adjoint problem in discrete form, two different approaches can be followed. We pursued the approach of taking the dual of the stabilized primal problem (15), which leads to the following equation

$$(M + \Delta \bar{t} V^T + \Delta \bar{t} \kappa K + \Delta \bar{t} \tau S^T + \tau V) q^n = (M + \tau V) q^{n+1} + M y^{n+1}, \tag{16}$$

with initial $q_{n=0} = 0$. This variant leads to the discrete operator $Q^h : \mathbb{R}^{N_{\text{obs}}} \rightarrow \bigoplus_{i=0}^{n_T} \mathcal{V}_h$.

In the alternative approach, the adjoint problem (\mathcal{P}_Q) , which becomes a transport problem after the transformation $t = T - \bar{t}$, is stabilized and discretized, and the differences in numeric accuracy will be treated in the remainder of this section.

For the discretization, linear Lagrange finite elements of first order are used throughout the following. This choice ensures good numerical accuracy, provided that the discretization is sufficiently fine. Examining the term for S_{ij} in the SUPG-stabilization, it is evident that the diffusion term $-\kappa \Delta \phi_i$ is not included in the residual for first-order elements. To improve accuracy, the second derivatives can be “patched.” Specifically, Eq. (15) is initially solved under the condition $\Delta u^n = 0$. Subsequently, the derivatives of the resulting solution u^n are projected onto a higher-order function space. Subsequently, the patched equation is solved using the approximate second derivatives of Δu^n . This approach enhances the accuracy of the forward problem, but does not improve the observed asymptotic linear convergence in the C^0 -norm, see Fig. 7. In the adjoint problem (Eq. (16)), which is the dual problem of the SUPG-stabilized forward problem, the sensor misfit is not incorporated into the SUPG-stabilization, which would be the case if the adjoint problem (\mathcal{P}_Q) were SUPG-stabilized after performing the change of variables. Despite this omission, the convergence behavior does not deteriorate compared to the “patched” fully SUPG-stabilized adjoint problem, see Fig. 7. For this reason, we have pursued the approach of using the dual problem of the SUPG-stabilized primal problem, which leads to Eq. (16).

To connect primal and dual equation also a discretized form of the observation operator B is required. Therefore, the discrete solution $I(\mathbf{u}^n) = \sum_{i=1}^{n_{\text{dof}}} u_i^n \phi_i$ must be evaluated on the grid as observation points do not necessarily coincide with finite element nodes, i.e., $x_i^{\text{obs}} \notin \{x_1, \dots, x_{n_{\text{dof}}}\}$. This process can be achieved using barycentric interpolation, as described in [54, Problem 1.3]. For the sake of completeness, the method is briefly summarized for triangular and tetrahedral meshes. The interpolation of the function u , given in a discrete form, is discussed. There exists a triangle $\{x_{i_1}, x_{i_2}, x_{i_3}\}$ or a tetrahedron $\{x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4}\}$ for $\Omega \subset \mathbb{R}^{n_{\text{sd}}=2}$ or $\subset \mathbb{R}^{n_{\text{sd}}=3}$ containing x_i^{obs} . The interpolation of the discrete function $I(\mathbf{u}^n)(x_i^{\text{obs}}, x_i^{\text{obs}})$ in a barycentric coordinate system is given by $\sum_j u_j a_j$, with $0 \leq a_j \leq 1$, $\sum_j a_j = 1$ and $x_i^{\text{obs}} = \sum_j a_j x_{i_j}$. In case $t_i^{\text{obs}} \neq k\Delta t$ for $k \in \mathbb{N}_0$ and $k \leq n_T$, interpolation between two time steps is needed additionally to construct the *discrete observation operator* $B^h : \bigoplus_{i=0}^{n_T} \mathcal{V}_h \rightarrow \mathbb{R}^{N_{\text{obs}}}$ and the *discrete parameter-to-observable map* $\mathcal{F}^h : \mathcal{V}_h \rightarrow \mathbb{R}^{N_{\text{obs}}}$, where $\mathcal{F}^h = B^h \circ \mathcal{K}^h$. Moreover, using the barycentric coordinate system, the discrete Dirac distribution is defined as

$$\delta_{x_i^{\text{obs}}} = M^{-1} \sum_j a_j e_{i_j}, \text{ where } a_j \text{ is a barycentric representation of } x_i^{\text{obs}}. \tag{17}$$

Now, the right-hand side of the adjoint problem (\mathcal{P}_Q) can now also be formulated with $\mathbf{y} = y_i \delta_{(t_i^{\text{obs}}, x_i^{\text{obs}})} \in \bigoplus_{i=0}^{n_T} \mathcal{V}_h$, for a given misfit vector $\mathbf{y} \in \mathbb{R}^{N_{\text{obs}}}$. Concerning the discretization of the unknown sources and shape functions, we replace $\mathcal{M}^+(\Omega)$ by the cone of Dirac-Delta functionals supported in the grid nodes,

$$\mathcal{M}_h^+ = \text{cone} \left\{ \delta_{x_j} \mid j = 1, \dots, n_{\text{dof}} \right\} = \left\{ \sum_{j=1}^{n_{\text{dof}}} \lambda_j \delta_{x_j} \mid \boldsymbol{\lambda} \in \mathbb{R}_{\geq 0}^{n_{\text{dof}}} \right\}. \tag{18}$$

Subsequently, the discretized version of the radial basis function S^1 corresponds to the standard L^2 -projection onto the respective elements. The discrete realization of the second source type, S^2 , is obtained by solving $(\eta M + \gamma K + \beta B)\mathbf{m} = \boldsymbol{\delta}_x$ where B with $B_{ij} := \int_{\partial\Omega} \phi_i(x)\phi_j(x) d\partial\Omega(x)$ corresponds to the Robin boundary condition. Finally, the third source type, S^3 , does not require further discretization. Afterwards, Algorithm 1 is applied to the discretized problem where the new candidate locations are computed by

$$\hat{x}_k^I \in \arg \max_{x \in \{x_1, \dots, x_{n_{\text{dof}}}\}} \varphi_1^k(x), \quad \hat{x}_k^C \in \arg \max_{x \in \{x_1, \dots, x_{n_{\text{dof}}}\}} \varphi_C^k(x)$$

which requires $\mathcal{O}(n_{\text{dof}})$ operations.

In summary, this leads to two levels of approximation: On the one hand side, the finite element discretization to simulate the forward and adjoint problem. On the other hand, the discretization of the admissible set, (18). Since, in general, optimal solutions of the limit problem (\mathcal{P}_M) are not supported in grid nodes, the latter leads to clustering effects, i.e. an optimal Dirac-Delta functional in the interior of a triangle/tetrahedron is approximated by various sources in the adjacent grid nodes. In a subsequent heuristic post-processing step, we merge the latter into a single source using barycentric coordinates. This is illustrated in Fig. 9.

Remark 2. For piecewise linear and continuous finite elements as well as $m_1 = S^3$, there holds

$$\max_{x \in \{x_1, \dots, x_{n_{\text{dof}}}\}} \varphi_1(x) = \max_{x \in \Omega} \varphi_1(x)$$

since $\varphi_1 = Q(y)(0, \cdot) \in \mathcal{V}_h$ in this case. Hence, the admissible set is implicitly discretized due to the approximation of the PDE which is also referred to as variational discretization, see, e.g., [56]. Note, however, that this does not alleviate the mentioned clustering phenomena.

5. Numerical results

Numerical results evaluate the performance of the presented algorithm on four test case geometries, namely on a two building benchmark geometry in 2D and 3D, on a campus geometry with wind in inflow-outflow configuration and on a 2D cut of a chemistry plant site. If not stated otherwise, numeric values of physical quantities are reported in standard SI units.

5.1. Benchmark geometry representing two buildings

First, a widely used benchmark geometry is employed to evaluate the algorithm and its implementation. The geometry consists of a unit square with two rectangular obstacles (“buildings”), and the double-lid-driven-cavity setup is utilized to create a dynamic flow scenario [19]:

$$\begin{aligned} -v\nabla^2 \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v} + \nabla \rho &= 0 & \text{in } \Omega, \\ \nabla \cdot \mathbf{v} &= 0 & \text{in } \Omega, \\ \mathbf{v} &= \mathbf{g} & \text{on } \Gamma_D. \end{aligned} \tag{19}$$

In this example of the incompressible Navier-Stokes equations, ρ represents the pressure, and $v = 1/50$ denotes viscosity, resulting in a Reynolds number of 50. The Dirichlet boundary conditions are specified as $\mathbf{g} = \mathbf{e}_2$ on the left wall of the domain and $\mathbf{g} = -\mathbf{e}_2$ on the right wall. No-slip boundary conditions are applied on the remaining boundaries. The wind vector field for this benchmark example

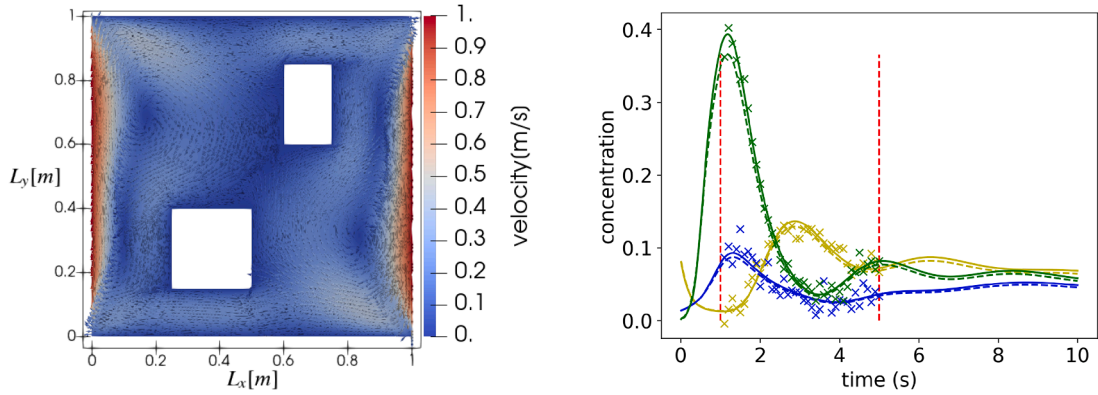


Fig. 8. PDAP-algorithm input. Wind vector field (left) and noisy measurements at 3 sensors from $t = 1$ s to $t = 5$ s (right, marked by crosses), along with exact concentration values (full line), and PDAP-reconstruction (dashed line).

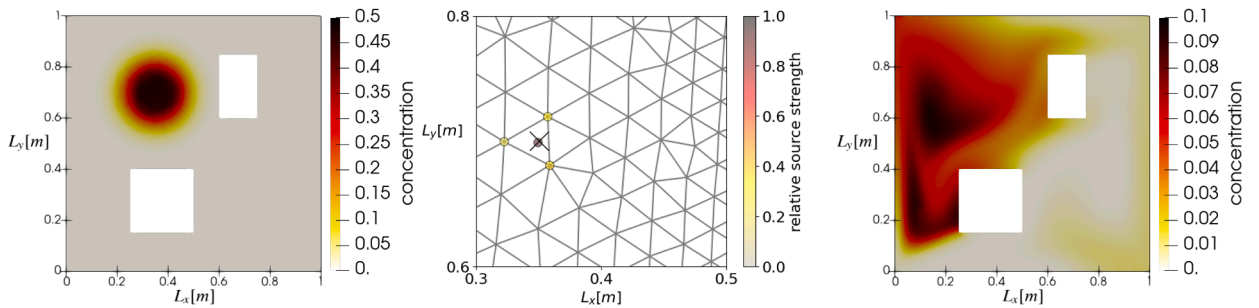


Fig. 9. PDAP-algorithm output for baseline simulation. Reconstructed initial condition (left), identified parameter, namely, reconstructed raw sources $\lambda_i^1 \delta_{x_i^1}$ (yellow dots) and post-processed source $\mu_{\text{post}} = \lambda_{\text{post}} \delta_{x_{\text{post}}}$ (dark dot) next to actual source center (black cross, middle), predicted concentration field at $t = 5$ s (right). Ground truth in Fig. 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1

Robustness of PDAP algorithm with respect to challenging measurements conditions in terms of Euclidean distance from actual source center x_s to algorithmically reconstructed source location x_{post} .

Measurement setting	original	reduced	$\ x_{\text{post}} - x_s\ _{\Omega}$
Baseline simulation (Fig. 9)			0.001 m
Reduced signal-to-noise ratio SNR	≈ 33.3	≈ 6.7	0.003 m
Reduced number of sensors N_d	9	3	0.013 m
Reduced sampling rate f_s	10 Hz	2 Hz	0.019 m
Reduced sampling time T_s	[1.0 s, 5.0 s]	[2.0 s, 3.0 s]	0.017 m
Single moving sensor (Fig. 10)			0.050 m

is depicted in Fig. 8, see also [20]. To reproduce reference results [47,57], the radial basis function $S^1(\cdot, 0.26, \cdot)$ is chosen as the shape function S_i , and the initial condition is defined as $u_0^{\text{bf}} = m_i(\mu_1[x^1, \lambda^1])$, with $x^1 = [(0.35, 0.7)]$ and $\lambda^1 = [1.0]$ (based on Eq. (10)). The diffusion constant $\kappa = 0.001 \text{ m}^2 \text{ s}^{-1}$ has been adjusted accordingly. Synthetic measurement data d is generated with additive white noise following $\mathcal{N}(0, \sigma^2 \text{Id})$. In this case, a relative noise of 3% was multiplied by the maximum measured concentration to determine the variance $\sigma^2 = (0.03 \cdot \max_{i \in \mathbb{R}^{N_{\text{obs}}}} u(i_i^{\text{obs}}, x_i^{\text{obs}}))^2$ for the perturbation. This corresponds to a signal-to-noise ratio of $\text{SNR} \approx 33.3$. The resulting measurement input to the PDAP algorithm is visualized in Fig. 8 (right). With $\alpha = 1000$ as the regularization parameter (in Eq. $(\mathcal{P}(x^1, x^c))$), only eight iterations are required to reconstruct the solution, which corresponds to eight forward solutions of (\mathcal{P}_{κ}) , and eight adjoint solutions of (\mathcal{P}_Q) .

The results displayed in Fig. 9 accurately reproduce the outcomes of established approaches that utilize L^2 -regularization [19, 20,47,57–60]. Note that both presented snapshots of the computed concentration fields are in the extrapolation regime, i.e., at time instances for which no measurement data are available.

Additional experiments are conducted to demonstrate the robustness of Algorithm 1 under challenging measurement conditions, such as reducing the signal-to-noise ratio, the number of sensors, the sampling rate, and the sampling time. Furthermore, we emulate a moving sensor (Fig. 10). The results of these numerical experiments are reported in Table 1. The obtained accuracy of source

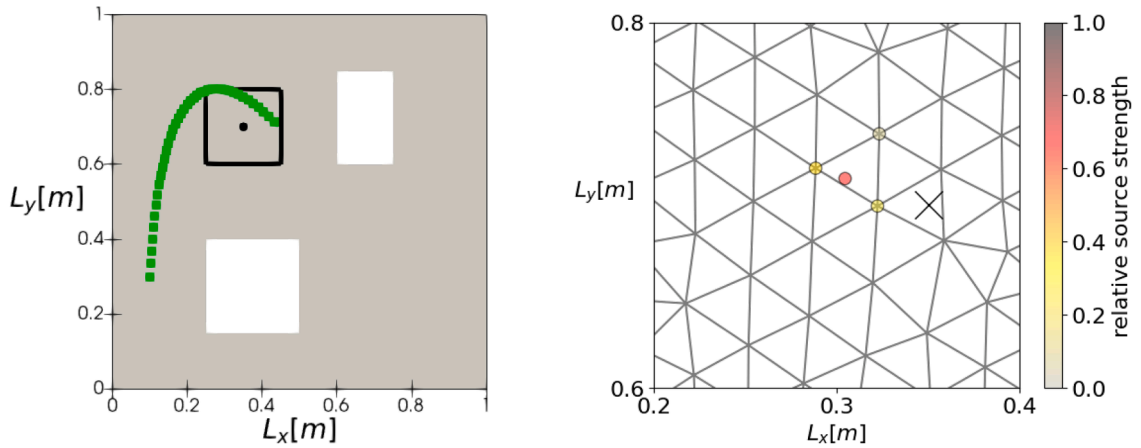


Fig. 10. Single moving sensor experiment. Trajectory of the moving sensor $\gamma(t) = [0.6, 0.5] + t[-11/60, 83/300] + t^2[1/60, -19/300]$ for $t^{\text{obs}} \in \{1.0s, 1.1s, \dots, 5s\}$ (green), source location and excerpt (black, left), predicted raw sources $\lambda_i^1 \delta_{x_i^1}$ (yellow dots) and derived source μ_{post} (red dot) near ground truth source center (black cross, right). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

identification remains acceptable under all five considered perturbations. Thus, we conclude that the algorithm is robust with respect to reduced measurement quality in this test case. Moreover, a single moving sensor with predefined trajectory is sufficient to identify the contaminant source with acceptable accuracy.

To distinguish our method from established L^2 -approaches, we propose a more challenging initial condition with two overlapping sources $u_0^{\text{double}} = m_1(\mu_1[x^1, \lambda^1])$, with $x^1 = [(0.35, 0.7), (0.45, 0.7)]$ and $\lambda^1 = [1.0, 1.0]$, to be identified. Here, a higher signal-to-noise ratio $\text{SNR} = 100$ was used and the regularization parameter $\alpha = 500$ was adapted to the specific problem. The results are shown in Fig. 11. It is observed that the PDAP-algorithm is able to separate the two centers, while the established L^2 -approach recovers a single smooth function and cannot provide the insight that the contaminant stems from two separate sources. For the L^2 -reconstruction, we adhere to the provided numerical example in [19] and employ the same prior distribution, given by $\mathcal{A} = (8I - \Delta)$.

Remark 3. For the numerical examples, we set the tolerance of the Newton algorithm to a very small value, namely $\text{tol}_{\text{Newt}} = 10^{-10}$. The termination of Algorithm 1 is controlled by tol , and an error estimate relative to $J(0, 0) = 1/(2\sigma^2) \|d\|_{\mathbb{R}^{N_{\text{obs}}}}^2$ is given in (13). In our examples, we choose tol/α between 0.01 and 0.1.

5.2. Source identification in three-dimensional flow field

The following scenario extends the two-building benchmark to a 3D geometry (Fig. 1). As initial condition, again the shape function $S^1(\cdot, 0.25, \cdot)$ is chosen and the source is located at $x^1 = [(0.35, 0.7, 0.26)]$, with intensity $\lambda^1 = [1.0]$. The observation pattern corresponds to a sensor array mounted with antennas on the roofs of the two cuboid buildings. The PDAD algorithm requires only 22 iterations (22 forward and 22 adjoint solutions) to accurately reconstruct the initial condition. The center of the reconstructed source is in close vicinity to the true center. A visual comparison of the original simulation and the measurement-based reconstruction is shown in Fig. 1.

5.3. Source identification on complex campus domain with wind in inflow-outflow configuration

To explore a more realistic scenario, a portion of the campus of the University of the Bundeswehr Munich, Germany, is selected as computational domain. The geometry comprises several building imprints directly imported from Open Street Map (OSM). Moreover, a boundary-refined finite element mesh is generated in an automated workflow [61]. At the southern side of the geometry, wind enters the domain resulting in the vector field shown in Fig. 12. For this test case, a diffusion coefficient of $\kappa = 10.0 \text{ m}^2 \text{ s}^{-1}$ is chosen in the transport problem (\mathcal{P}_κ) . The observation time frame $T = [2 \text{ s}, 15 \text{ s}]$ is sampled at a rate of 10Hz, resulting in the input signal displayed in Fig. 12.

5.3.1. Closed form initial condition – comparison with L^2 -approach

This section elaborates how Algorithm 1 compares to an established approach with L^2 -regularization. To achieve this, the previous initial condition is adapted to a more complex scenario in the current environment: $u_0^{\text{bf}} = m_1(\mu_1[x^1, \lambda^1])$, with $x^1 = [(-50, -90), (120, -70), (-40, 0)]$ and $\lambda^1 = [1.0, 0.5, 0.1]$ and radius $r = 10 \text{ m}$. Further algorithmic parameters are reported in Table 2, along with the results for case a and case b for a relatively dense grid of 213 sensors, as well as results for a sparse sensor layout of only 5 sensors (case c and case d). Considering the combined number of PDE solutions in the online and offline phase, the presented PDAP-algorithm clearly outperforms the L^2 -approach. Considering only online solves, an acceptable number of iterations is required

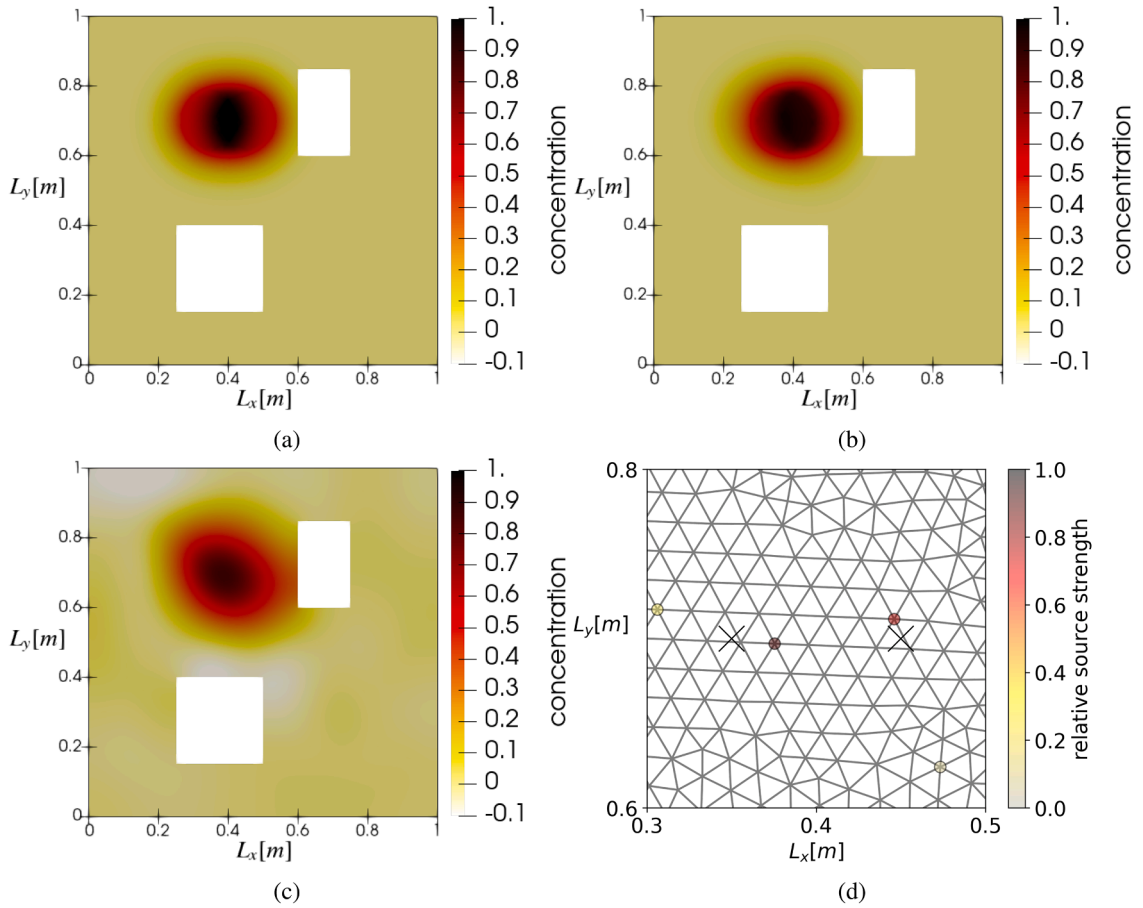


Fig. 11. Two overlapping contaminant sources. Ground truth initial condition u_0^{double} (a), PDAP-reconstructed contaminant sources (b), L^2 -reconstruction (c) and predicted raw sources $\lambda_i^1 \delta_{x_i}$ (d).

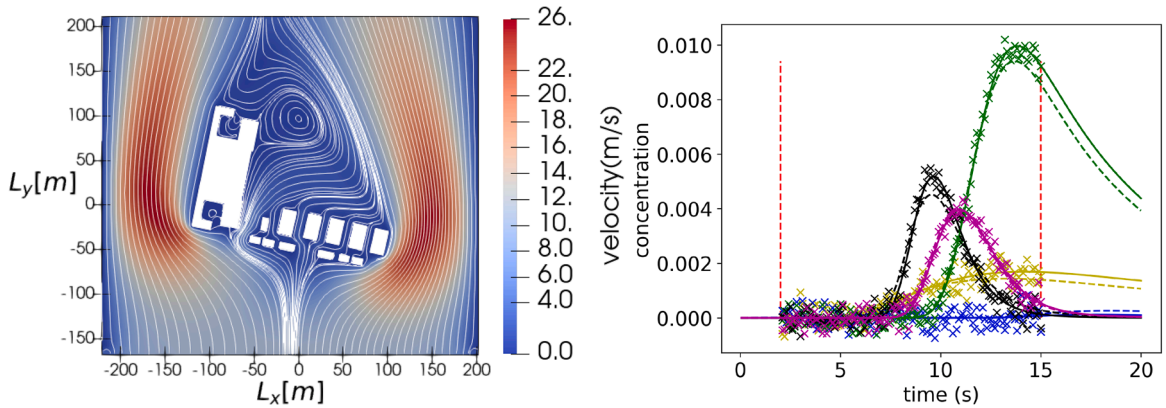


Fig. 12. Campus domain input. Wind vector field (left) and noisy measurements d at 5 sensors (right, marked by crosses, sensor positions shown in Fig. 13), along with exact and PDAP-reconstructed concentration (full, dashed line).

in both algorithms. Notably, the number of PDAP iterations decreases for added information (more sensors) in contrast to the L^2 -approach. Fig. 13 displays the reconstructed initial conditions and demonstrates the better accuracy of the reconstructions obtained with the PDAP-algorithm.

Table 2
Computational cost of established L^2 -approach and PDAP-algorithm reported in terms of required PDE solutions.

case	source type	# sensors	algorithm	parameters	PDE solutions	
					online	offline
a	u_0^{rbf}	213	PDAP	$\alpha = 1000$	17	–
b	u_0^{rbf}	213	L^2	$\eta = 160.0, \gamma = 20.0$	54	1600
c	u_0^{rbf}	5	PDAP	$\alpha = 1500$	20	–
d	u_0^{rbf}	5	L^2	$\eta = 160.0, \gamma = 20.0$	16	120
e	u_0^{ellip}	213	PDAP	$\alpha = 1000$	25	–
f	u_0^{ellip}	5	PDAP	$\alpha = 1500$	22	–
g	u_0^{dirac}	5	PDAP	$\alpha = 2000$	19	–

Table 3
Quantitative results of source localization at chemistry plant site, accompanying Fig. 17.

Source location x_s		Reconstruction x_{post}		$\ x_{\text{post}} - x_s\ _{\Omega}$ [m]	Estimated intensity λ_{post} (truth 1.0)
x	y	x	y		
329.5	-9.1	329.7	-8.9	0.29	0.9
236.1	61.6	234.8	61.7	1.33	1.0
382.7	293.4	386.9	295.1	4.52	0.9
181.9	264.8	192.8	265.3	10.95	0.6
456.6	207.0	456.9	203.4	3.67	0.7
75.1	112.9	72.7	114.2	2.79	0.7
226.0	358.3	226.0	358.6	0.29	1.0
374.2	88.5	374.2	88.5	0.00	0.9

5.3.2. *Implicit source definition with elliptic PDE*

If the initial condition is selected as in the previous case, a convolution of the adjoint solution with the shape functions must be performed at each point in Ω to compute the dual. This process incurs a significant computational cost. To mitigate this, an alternative modeling approach for the initial condition, S^2 , is considered. In this approach, the computation of $\varphi_1(x)$ or $\varphi_C(x)$ requires solving only a single solve of an elliptic problem (Eq. (11)). To motivate this, we utilize the symmetry of the operator \mathcal{A} and observe that

$$\int_{\Omega} S_1^2(x, z) Q(y)(t, z) \, d\Omega(z) = \int_{\Omega} \delta_x(z) \mathcal{A}^{-1}(Q(y)(t, \cdot))(z) \, d\Omega(z) = \mathcal{A}^{-1}(Q(y)(t, \cdot))(x) \tag{20}$$

From this identity, we deduce that

$$\varphi_1(x) = \mathcal{A}^{-1}(Q(y)(0, \cdot))(x), \quad \varphi_C(x) = \int_0^T \mathcal{A}^{-1}(Q(y)(t, \cdot))(x) \, dt \tag{21}$$

and therefore replace the convolution of the shape functions with the elliptic equation solve in Step 2 in Algorithm 1.

In the corresponding experiment (Fig. 14), the initial condition is expressed as $u_0^{\text{ellip}} = m_1(\mu_1[\mathbf{x}^1, \boldsymbol{\lambda}^1])$, with S^2 as shape function, $\mathbf{x}^1 = [(-50, -90), (120, -70), (-40, 0)]$, $\boldsymbol{\lambda}^1 = [1000, 500, 100]$, $\eta = 1$ and $\gamma = 100$. Parameters are selected to closely match the appearance of the initial condition u_0^{rbf} above. Further details of the configuration are provided in case f and very satisfactory results are presented in Fig. 14. In comparison to case a and c, a slight increase of PDAP-iterations is observed. However, this approach is still preferred as Step 2 in Algorithm 1 is greatly simplified, i.e., no convolution is required.

5.3.3. *Instantaneous point sources modeled as Dirac distributions*

As a third example on the campus domain (Fig. 15), we consider a scenario in which the initial condition is highly localized, taking the form of a linear combination of Dirac-Delta functionals. An important advantage of this formulation is that the associated dual, as defined in Theorem 1, coincides with the solution of (P_Q) . Consequently, the convolution with the initial condition is completely omitted, which significantly simplifies the implementation and increases the computational efficiency. The use of a Dirac distribution $S^3 = \delta_{x_s}$ as the initial condition is also physically justified, if the area of the initially affected region is very small relative to the overall domain. In the context of a contaminant release at a chemical plant site, this can be a reasonable assumption. In the numerical experiment, the initial condition is defined as $u_0^{\text{dirac}} = m_1(\mu_1[\mathbf{x}^1, \boldsymbol{\lambda}^1])$, with S^3 as shape function, $\mathbf{x}^1 = [(-50, -90), (120, -70), (-40, 0)]$ and $\boldsymbol{\lambda}^1 = [100, 100, 100]$. The algorithm described in Algorithm 1 successfully identifies the sources with the dense sensor arrangement, and even with the reduced number of five sensors, the method maintains a satisfactory level of accuracy; see Fig. 15 for a visualization of the result.

5.3.4. *Continuous contaminant sources*

The final experiment on the campus domain (Fig. 16) evaluates the capability of Algorithm 1 to identify continuous sources, defined as $u_{\text{cont}} = m_C(\mu_C[\mathbf{x}^C, \boldsymbol{\lambda}^C])$, with S^2 as shape function, $\mathbf{x}^1 = [(-50, -90), (120, -70), (-40, 0)]$, $\boldsymbol{\lambda}^1 = [1000, 500, 100]$, $\eta = 1$ and $\gamma = 100$. As

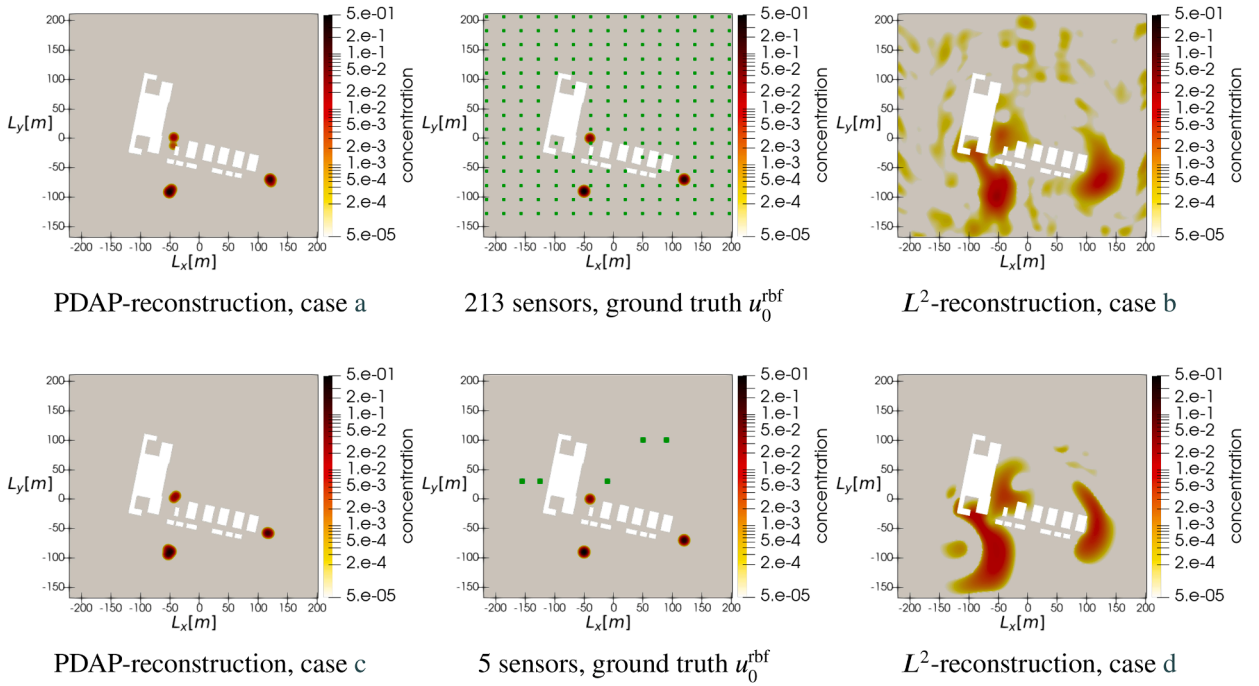


Fig. 13. Comparison of PDAP-reconstruction (left) and L^2 -reconstruction (right) of initial condition u_0^{rbf} (ground truth, middle) based on measurements at sensor positions (green squares, middle) for cases listed in Table 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

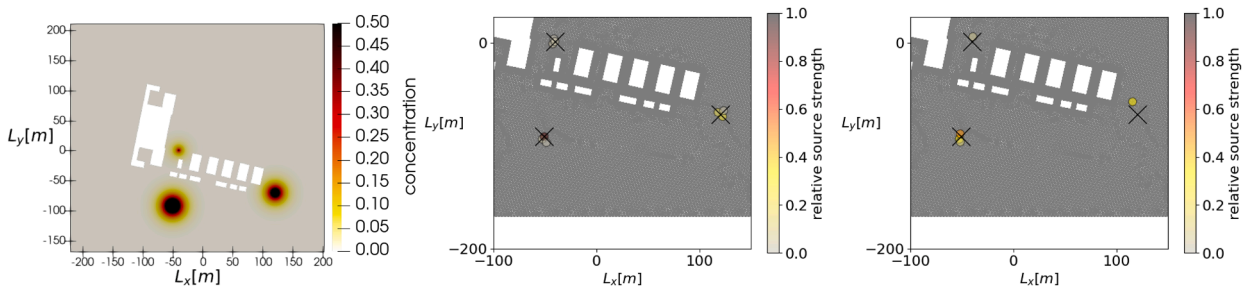


Fig. 14. PDAP reconstruction of elliptic initial condition u_0^{ellip} (ground truth, left) for two sensor configurations. Reconstructed point sources based on all 213 sensors (case e in Table 2, middle), and reconstruction using reduced sensor configuration (case f in Table 2, right).

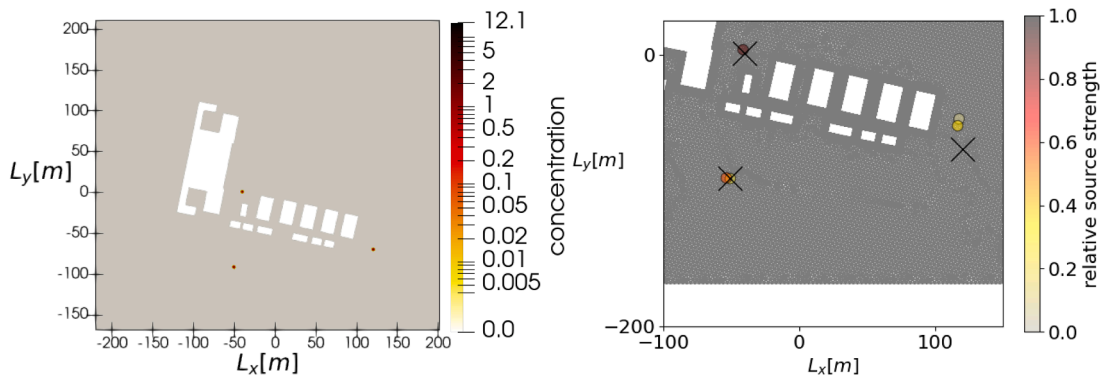


Fig. 15. Campus geometry with very localized initial condition of Dirac type u_0^{dirac} (left), reconstructed point sources for case g listed in Table 2.

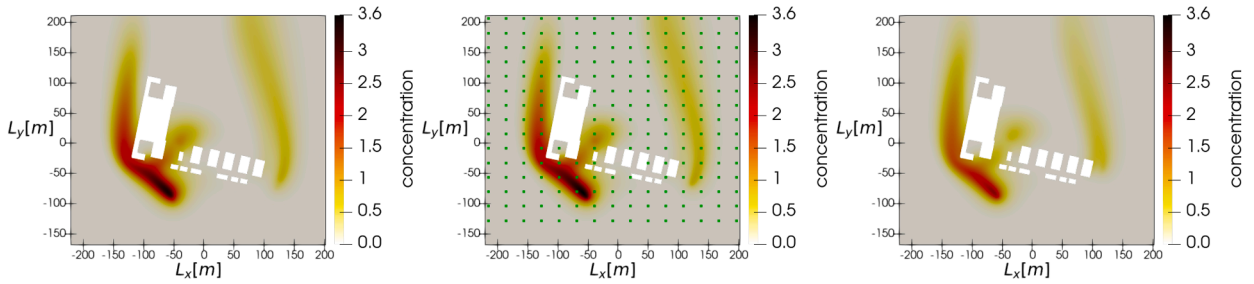


Fig. 16. Identification of continuous contaminant sources. Simulated concentration field with source term u_{cont} evaluated at $t = 20$ s (ground truth, middle), PDAP reconstruction with 213 sensors (left), and with 5 sensors (right).

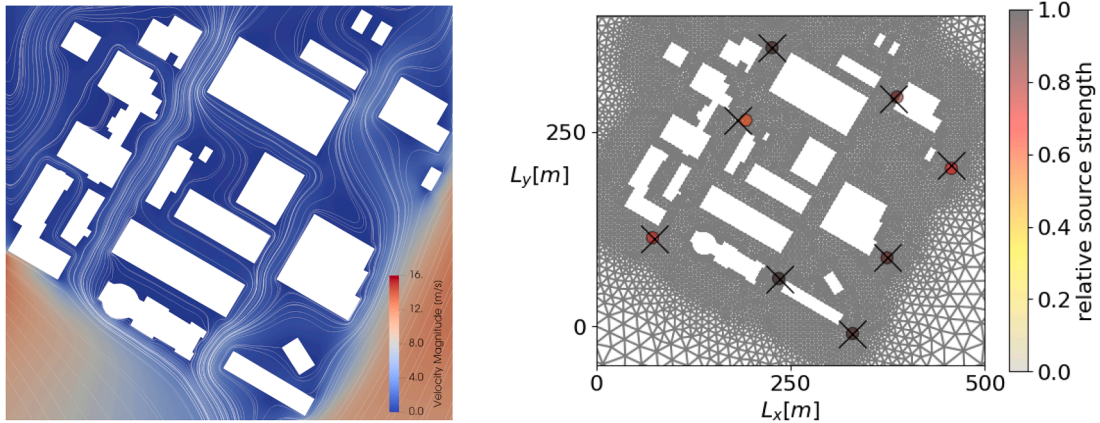


Fig. 17. Estimated wind vector field for chemistry plant site (left), post-processed PDAP-reconstruction of sources (dots) properly indicates locations of actual sources (crosses) on computational domain (right). Geometry adopted from [62].

indicated by the formulation in (\mathcal{P}_Q) , the adjoint solution $q = Q(y)$ may not be smooth at every time instance. However, sufficient regularity of φ_C can still be ensured by selecting an appropriate observation operator B or applying convolution with a suitably regularized kernel. For instance S^2 provides sufficient regularity at φ_C and the problem remains well-posed and solvable. In case of a densely distributed sensor network, the concentration field is again accurately reconstructed; see Fig. 16. However, with a sparse sensor configuration, the problem becomes more difficult. In such cases, careful sensor placement is essential to enable accurate source identification.

5.4. Empirical robustness study of PDAP algorithm for chemistry plant site

The geometry of this test case is taken from the recent work of Schneider et al. [62]. The objective of their study is to investigate a potential application of the Emergency Response Inference Mapping (ERIMap) in collaboration with the fire brigade of a chemistry plant operator. Specifically, an accident of a tank wagon transporting chlorine is defined as test scenario. The dispersion of the released contaminant is modeled in their work using the ALOHA (Areal Locations of Hazardous Atmospheres) software, which employs a Gaussian plume model for dispersion calculations [63]. The accident location is assumed to be known and the primary goal of the ERIMap approach is to identify areas where individuals may have been exposed to hazardous contaminant concentrations, based on discrete sensor measurements and their associated uncertainties. This contribution extends the original approach in two significant ways. First, solving the advection-diffusion equation can provide a more realistic dispersion prediction, accounting for the influence of obstacles, such as buildings, on the contaminant transport. Second, the applicability is extended to cases with unknown source location and intensity, enabling the use of existing sensor networks for source identification.

To estimate the wind vector field (Eq. (19)), a free-slip boundary condition was imposed on the building imprints by means of a Nitsche method as described in [64]. This modeling approach results in slightly increased wind speeds in the inter-building regions, thereby providing a challenging scenario for source localization as visualized in Fig. 17. The scenario is particularly difficult due to the complex interplay between diffusion and advection in the confined spaces between buildings, where flow dynamics are very heterogeneous. Regarding the source identification problem, eight instantaneous concentration sources are distributed across the domain. Concentration measurements were recorded at 154 sensor locations from $t_1 = 2$ s to $t_2 = 120$ s, using a sampling rate of 2 Hz. As in previous experiments, synthetic measurements were generated assuming a signal-to-noise ratio of $SNR \approx 33.3$.

In this scenario, the PDAP-algorithm required 27 iterations to achieve convergence for the results presented in Fig. 17. Based on empirical evaluation, we set the regularization parameter to $\alpha = 1000$, as it provides good adaptation to the specific problem. The

computations were performed on a workstation equipped with an Intel i9-10980XE processor and took less than 2 min. The majority of the computation time was spent in a total of 54 PDE solves (27 forward and 27 adjoint) each with approximately 10,000 degrees of freedom. Despite the scenario's complexity, the proposed [Algorithm 1](#) successfully identifies all eight sources with only a small number of candidate locations. Owing to the sparsity promoting regularization, these candidate locations are highly concentrated around the true source positions. For the employed sensor grid, the maximum absolute error in the prediction of the concentration field with the reconstructed sources, $u(200\text{ s}, \cdot)$ is bounded below 6.4×10^{-3} . Due to the strong clustering, the raw results can be improved by post-processing as in the previous section. Candidates within a circle of 40 m are projected in a post processing step to a single source.

The quantitative results in [Table 3](#) illustrate that the maximum distance from a reconstructed source to an actual source is less than 11 m on the considered computational domain of about $500\text{ m} \times 500\text{ m}$ and on average much smaller. Moreover, it is observed, that the reconstruction is less accurate in areas where diffusion ($\kappa = 2\text{ m}^2\text{ s}^{-1}$) dominates. All in all, [Fig. 17](#) confirms that the final reconstruction provides a reliable and accurate indication for the source locations and intensities.

6. Conclusion and outlook

In this paper, we considered a variational regularization ansatz for the identification of an unknown number of sources from scarce measurements in contaminant dispersion applications. The presented approach combines the theoretical insights of lifting of the problem from an unknown number of location-intensity pairs to a Radon measure on the spatial domain ([Theorem 1](#)) with the practical adaptation of the efficient Primal-Dual-Active-Point method ([Algorithm 1](#)). Numerical results showcased that the presented method provides a reliable reconstruction of contaminant sources in complex environments and that the presented method outperforms state-of-the-art techniques with L^2 -regularization. This is particularly evident in [Fig. 13](#), as well as in [Table 2](#), which demonstrate that accurate source identification is possible with sparse information of few sensors. Furthermore, the presented method has the advantage that adding sensor information does not increase the computational effort in terms of PDE solves, but rather reduces it. As a result, the method scales to scenarios in which a large area is monitored with several sensors (see [Fig. 17](#)).

For future work, the requirement of the known time horizon could be lifted, i.e., the time at which the contaminant was released should be subject to identification as well. Moreover, from a computational perspective, a variety of speed-ups are possible in order to further enhance real-time capabilities of the presented approach. Many of the required PDE solves can be performed in parallel or even be precomputed. For example, by linearity, the solution of the adjoint [Eq. \(P_Q\)](#) in every iteration can be reduced to a single matrix-vector product if we compute and store the fundamental solutions associated to the observation points in an offline library. To further reduce computational (online) cost, suitable reduced-order models or surrogates for the parameter-to-observable map \mathcal{F} can be constructed using, e.g., singular value decomposition [65]. Taking this further, integration with parameter-dependent reduced models for wind fields [61] could enable real-time forecasting in emergency scenarios, offering first responders fast estimates of the release locations and contaminant dispersion.

Finally, the source identification approach proposed here can be combined easily with uncertainty quantification and the optimal design of experiments [27]. In the context of crisis management, an assessment of the overall involved uncertainty is very helpful, e.g., to steer further measurements and support informed decision making in a Bayesian framework [3,62,66]. Given the accurate and robust source identification, an obvious next step is to transfer the method to other applications in critical infrastructure protection (CIP), e.g., to enhance acoustic source identification with targeted applications in structural health monitoring (SHM) [67,68], or to help identifying seismic sources using an elasto-acoustic modeling and simulation approach [69]. A further methodological development might also consider sparsity promoting regularization to support contaminant source identification in water distribution networks [70].

CRedit authorship contribution statement

Marco Mattuschka: Writing – original draft, Visualization, Software, Methodology; **Daniel Walter:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization; **Max von Danwitz:** Writing – review & editing, Validation, Supervision, Project administration; **Alexander Popp:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Max Winkler for discussions leading to the proper problem formulation and Johannes Pfefferer for his helpful comments on optimization strategy and implementation. A.P. acknowledges support by *dtec.bw - Digitalization and Technology Research Center of the Bundeswehr* (project *RISK.twin*). *dtec.bw* is funded by the European Union - NextGenerationEU.

References

- [1] J. Boris, The threat of chemical and biological terrorism: preparing a response, *Comput. Sci. Eng.* 4 (2) (2002) 22–32. <https://doi.org/10.1109/5992.988644>
- [2] G. Patnaik, J.P. Boris, F.F. Grinstein, J.P. Iselin, D. Hertwig, Large scale urban simulations with FCT, in: D. Kuzmin, R. Löhner, S. Turek (Eds.), *Flux-Corrected Transport*, Scientific Computation, Springer Netherlands, Dordrecht, 2012, pp. 91–117. https://doi.org/10.1007/978-94-007-4038-9_4
- [3] M. von Danwitz, J. Bonari, P. Franz, L. Kühn, M. Mattuschka, A. Popp, et al., Contaminant dispersion simulation in a digital twin framework for critical infrastructure protection. <https://doi.org/10.23967/eccomas.2024.301>
- [4] N.S. Holmes, L. Morawska, A review of dispersion modelling and its application to the dispersion of particles: an overview of different dispersion models available, *Atmos. Environ.* 40 (30) (2006) 5902–5928. <https://doi.org/10.1016/j.atmosenv.2006.06.003>
- [5] M. Hutchinson, H. Oh, W.-H. Chen, A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors, *Inf. Fusion* 36 (2017) 130–148. <https://doi.org/10.1016/j.inffus.2016.11.010>
- [6] R. Khodayi-mehr, W. Aquino, M.M. Zavlanos, Model-based active source identification in complex environments, *IEEE Trans. Rob.* 35 (3) (2019) 633–652. <https://doi.org/10.1109/TRO.2019.2894039>
- [7] P. Hinsen, T. Wiedemann, V.S.P. Ruiz, D. Shutin, A.J. Lilienthal, Experimental study of gas propagation: parameter identification and analysis in a wind tunnel, in: 2024 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), IEEE, 2024, pp. 1–3. <https://doi.org/10.1109/ISOEN61239.2024.10556306>
- [8] V.P. Ruiz, P. Hinsen, T. Wiedemann, D. Shutin, C. Christof, Gas source localization using physics-guided neural networks, in: 2024 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), IEEE, 2024, pp. 1–3. <https://doi.org/10.1109/ISOEN61239.2024.10556061>
- [9] T. Wiedemann, P. Hinsen, V.P. Ruiz, D. Shutin, A.J. Lilienthal, Domain knowledge assisted gas tomography, in: 2024 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), IEEE, 2024, pp. 1–3. <https://doi.org/10.1109/ISOEN61239.2024.10556226>
- [10] N.K. Arystanbekova, Application of Gaussian plume models for air pollution simulation at instantaneous emissions, *Math. Comput. Simul.* 67 (4-5) (2004) 451–458. <https://doi.org/10.1016/j.matcom.2004.06.023>
- [11] S. Ulfah, S.A. Awalludin, Wahidin, Advection-diffusion model for the simulation of air pollution distribution from a point source emission, *J. Phys. Conf. Ser.* 948 (2018) 012067. <https://doi.org/10.1088/1742-6596/948/1/012067>
- [12] B. Maronga, S. Banzhaf, C. Burmeister, T. Esch, R. Forkel, D. Fröhlich, V. Fuka, K.F. Gehrke, J. Geletič, S. Giersch, T. Gronemeier, G. Groß, W. Heldens, et al., Overview of the PALM model system 6.0, *Geosci. Model Dev.* 13 (3) (2020) 1335–1372. <https://doi.org/10.5194/gmd-13-1335-2020>
- [13] M. von Danwitz, V. Karyofylli, N. Hosters, M. Behr, Simplex space–time meshes in compressible flow simulations, *Int. J. Numer. Methods Fluids* 91 (1) (2019) 29–48. <https://doi.org/10.1002/flid.4743>
- [14] P. Towers, B.L. Jones, Real-time wind field reconstruction from LiDAR measurements using a dynamic wind model and state estimation, *Wind Energy* 19 (1) (2016) 133–150. <https://doi.org/10.1002/we.1824>
- [15] Y. Chen, D. Wang, D. Feng, G. Tian, V. Gupta, R. Cao, M. Wan, S. Chen, Three-dimensional spatiotemporal wind field reconstruction based on LiDAR and multi-scale PINN, *Appl. Energy* 377 (2025) 124577. <https://doi.org/10.1016/j.apenergy.2024.124577>
- [16] R.A.S. Albani, V.V.L. Albani, H.S. Migon, A.J. Silva Neto, Uncertainty quantification and atmospheric source estimation with a discrepancy-based and a state-dependent adaptive MCMC, *Environ. Pollut.* 290 (2021) 118039. <https://doi.org/10.1016/j.envpol.2021.118039>
- [17] S.M. Gorelick, B. Evans, I. Remson, Identifying sources of groundwater pollution: an optimization approach, *Water Resour. Res.* 19 (3) (1983) 779–790. <https://doi.org/10.1029/WR019i003p00779>
- [18] R. Tsai, S. Osher, Y. Li, Heat source identification based on l_1 constrained minimization, *Inverse Probl. Imaging* 8 (1) (2014) 199–221. <https://doi.org/10.3934/ipi.2014.8.199>
- [19] U. Villa, N. Petra, O. Ghattas, HIPPylib, *ACM Trans. Math. Softw.* 47 (2) (2021) 1–34. <https://doi.org/10.1145/3428447>
- [20] N. Petra, G. Stadler, Model variational inverse problems governed by partial differential equations, (<https://math.nyu.edu/~stadler/papers/PetraStadler11.pdf>).
- [21] F. Key, M. von Danwitz, F. Ballarin, G. Rozza, Model order reduction for deforming domain problems in a time–continuous space–time setting, *Int. J. Numer. Methods Eng.* 124 (23) (2023) 5125–5150. <https://doi.org/10.1002/nme.7342>
- [22] M. Dinkel, C.M. Geitner, G. Robalo Rei, J. Nitzler, W.A. Wall, Solving Bayesian inverse problems with expensive likelihoods using constrained Gaussian processes and active learning, *Inverse Probl.* 40 (9) (2024) 095008. <https://doi.org/10.1088/1361-6420/ad5eb4>
- [23] A.M. Stuart, Inverse problems: a Bayesian perspective, *Acta Numer.* 19 (2010) 451–559. <https://doi.org/10.1017/S09622492910000061>
- [24] J. Nitzler, J. Biehler, N. Fehn, P.-S. Koutsourelakis, W.A. Wall, A generalized probabilistic learning approach for multi-fidelity uncertainty quantification in complex physical simulations, *Comput. Methods Appl. Mech. Eng.* 400 (2022) 115600. <https://doi.org/10.1016/j.cma.2022.115600>
- [25] K. Bredies, H.K. Pikkarainen, Inverse problems in spaces of measures, *ESAIM, Control Optim. Calc. Var.* 19 (1) (2013) 190–218. <https://doi.org/10.1051/cocv/2011205>
- [26] O. Scherzer, B. Walch, Sparsity regularization for radon measures, in: X.-C. Tai, K. Mörken, M. Lysaker, K.-A. Lie (Eds.), *Scale Space and Variational Methods in Computer Vision*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 452–463. https://doi.org/10.1007/978-3-642-02256-2_38
- [27] P.-T. Huynh, K. Pieper, D. Walter, et al., Towards optimal sensor placement for inverse problems in spaces of measures, *Inverse Probl.* 40 (5) (2024) 055007. <https://doi.org/10.1088/1361-6420/ad2cf8>
- [28] E. Casas, K. Kunisch, Using sparse control methods to identify sources in linear diffusion-convection equations, *Inverse Probl.* 35 (11) (2019) 114002. <https://doi.org/10.1088/1361-6420/ab331c>
- [29] D. Leykekhman, B. Vexler, D. Walter, et al., Numerical analysis of sparse initial data identification for parabolic problems, *ESAIM Math. Modell. Numer. Anal.* 54 (4) (2020) 1139–1180. <https://doi.org/10.1051/m2an/2019083>
- [30] U. Biccari, Y. Song, X. Yuan, E. Zuazua, A two-stage numerical approach for the sparse initial source identification of a diffusion–advection equation, *Inverse Probl.* 39 (9) (2023) 095003. <https://doi.org/10.1088/1361-6420/ace548>
- [31] A. Monge, E. Zuazua, Sparse source identification of linear diffusion–advection equations by adjoint methods, *Syst. Control Lett.* 145 (2020) 104801. <https://doi.org/10.1016/j.sysconle.2020.104801>
- [32] E.J. Candès, C. Fernandez-Granda, Towards a mathematical theory of super-resolution, *Comm. Pure Appl. Math.* 67 (6) (2014) 906–956. <https://doi.org/10.1002/cpa.21455>
- [33] V. Duval, G. Peyré, Exact support recovery for sparse spikes deconvolution, *Found. Comput. Math.* 15 (5) (2015) 1315–1355. <https://doi.org/10.1007/s10208-014-9228-6>
- [34] K. Pieper, D. Walter, Linear convergence of accelerated conditional gradient algorithms in spaces of measures, *ESAIM, Control. Optim. Calc. Var.* 27 (2021) 38. <https://doi.org/10.1051/cocv/2021042>
- [35] Q. Denoyelle, V. Duval, G. Peyré, E. Soubies, The sliding Frank-Wolfe algorithm and its application to super-resolution microscopy, *Inverse Probl.* 36 (1) (2020) 42. Id/No 014001. <https://doi.org/10.1088/1361-6420/ab2a29>
- [36] A. Hnatiuk, D. Walter, Lazifying point insertion algorithms in spaces of measures, 2025. <https://doi.org/10.48550/arXiv.2508.03459>
- [37] L. Chizat, Sparse optimization on measures with over-parameterized gradient descent, *Math. Program.* 194 (1-2 (A)) (2022) 487–532. <https://doi.org/10.1007/s10107-021-01636-z>
- [38] H.C. Elman, T. Su, A low-rank solver for the stochastic unsteady Navier–Stokes problem, *Comput. Methods Appl. Mech. Eng.* 364 (2020) 112948. <https://doi.org/10.1016/j.cma.2020.112948>
- [39] A. Friedman, *Partial Differential Equations of Parabolic Type*, Dover Publications, Mineola, New York, dover ed. edition, 2008.
- [40] I. Neitzel, K. Pieper, B. Vexler, D. Walter, A sparse control approach to optimal sensor placement in PDE-constrained parameter estimation problems, *Numer. Math.* 143 (4) (2019) 943–984. <https://doi.org/10.1007/s00211-019-01073-3>
- [41] C. Boyer, A. Chambolle, Y. De Castro, V. Duval, F. de Gournay, P. Weiss, On representer theorems and convex regularization, *SIAM J. Optim.* 29 (2) (2019) 1260–1281. <https://doi.org/10.1137/18M1200750>

- [42] K. Bredies, M. Carioni, Sparsity of solutions for variational inverse problems with finite-dimensional data, *Calc. Var. Partial Differ. Eq.* 59 (1) (2020) Paper No. 14, 26. <https://doi.org/10.1007/s00526-019-1658-1>
- [43] K. Bredies, J.A. Iglesias, D. Walter, On extremal points for some vectorial total variation seminorms, *ESAIM Control Optim. Calc. Var.* 31 (2025) Paper No. 92, 35. <https://doi.org/10.1051/cocv/2025078>
- [44] S.P. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge and New York, NY and Port Melbourne, 2004. <https://doi.org/10.1017/CBO9780511804441>
- [45] T. Steihaug, Local and superlinear convergence for truncated iterated projections methods, *Math. Program.* 27 (2) (1983) 176–190. <https://doi.org/10.1007/BF02591944>
- [46] Y. Daon, G. Stadler, Mitigating the influence of the boundary on PDE-based covariance operators, *Inverse Probl. Imaging* 12 (5) (2018) 1083–1102. <https://doi.org/10.3934/ipi.2018045>
- [47] K. Wu, P. Chen, O. Ghattas, An offline-online decomposition method for efficient linear Bayesian goal-oriented optimal experimental design: application to optimal sensor placement, *SIAM J. Sci. Comput.* 45 (1) (2023) B57–B77. <https://doi.org/10.1137/21M1466542>
- [48] D. Walter, On sparse sensor placement for parameter identification problems with partial differential equations, Ph.D. thesis, Technical University of Munich, 2019. <https://nbn-resolving.org/urn:nbn:de:bvb:91-diss-20190604-1456274-1-8>.
- [49] G. Wachsmuth, D. Walter, No-gap second-order conditions for minimization problems in spaces of measures, 2024. <https://doi.org/10.48550/arXiv.2403.12001>
- [50] A. Milzarek, M. Ulbrich, A semismooth Newton method with multidimensional filter globalization for l_1 -optimization, *SIAM J. Optim.* 24 (1) (2014) 298–333. <https://doi.org/10.1137/120892167>
- [51] A. Flinth, F. de Gournay, P. Weiss, Grid is good. adaptive refinement algorithms for off-the-grid total variation minimization, *OJMO, Open J. Math. Optim.* 6 (2025) 1–27. <https://doi.org/10.5802/ojmo.39>
- [52] A.N. Brooks, T.J.R. Hughes, Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations, *Comput. Methods Appl. Mech. Eng.* 32 (1-3) (1982) 199–259. [https://doi.org/10.1016/0045-7825\(82\)90071-8](https://doi.org/10.1016/0045-7825(82)90071-8)
- [53] M. von Danwitz, I. Voulis, N. Hosters, M. Behr, Time-continuous and time-discontinuous space-time finite elements for advection-diffusion problems, *Int. J. Numer. Methods Eng.* 124 (14) (2023) 3117–3144. <https://doi.org/10.1002/nme.7241>
- [54] H.C. Elman, D.J. Silvester, A.J. Wathen, *Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics*, Numerical mathematics and scientific computation, Oxford University Press, Oxford, second ed. edition, 2014. <https://doi.org/10.1093/acprof:oso/9780199678792.001.0001>
- [55] D. Givoli, Time reversal as a computational tool in acoustics and elastodynamics, *J. Comput. Acoust.* 22 (03) (2014) 1430001. <https://doi.org/10.1142/S0218396X14300011>
- [56] E. Casas, C. Clason, K. Kunisch, Approximation of elliptic control problems in measure spaces with sparse solutions, *SIAM J. Control Optim.* 50 (4) (2012) 1735–1752. <https://doi.org/10.1137/110843216>
- [57] S. Wogrin, A. Singh, D. Allaire, O. Ghattas, K. Willcox, From data to decisions: a real-time measurement-inversion-prediction-steering framework for hazardous events and health monitoring, in: F. Darema, E.P. Blasch, S. Ravela, A.J. Aved (Eds.), *Handbook of Dynamic Data Driven Applications Systems*, Springer International Publishing, Cham, 2023, pp. 195–227. https://doi.org/10.1007/978-3-031-27986-7_8
- [58] A. Alexanderian, N. Petra, G. Stadler, O. Ghattas, A-optimal design of experiments for infinite-dimensional Bayesian linear inverse problems with regularized ℓ_0 -sparsification, *SIAM J. Sci. Comput.* 36 (5) (2014) A2122–A2148. <https://doi.org/10.1137/130933381>
- [59] A. Alexanderian, A.K. Saibaba, Efficient D-optimal design of experiments for infinite-dimensional Bayesian linear inverse problems, *SIAM J. Sci. Comput.* 40 (5) (2018) A2956–A2985. <https://doi.org/10.1137/17M115712X>
- [60] A. Attia, A. Alexanderian, A.K. Saibaba, Goal-oriented optimal design of experiments for large-scale Bayesian linear inverse problems, *Inverse Probl.* 34 (9) (2018) 095009. <https://doi.org/10.1088/1361-6420/aad210>
- [61] J. Bonari, L. Kühn, M. von Danwitz, A. Popp, Towards real-time urban physics simulations with digital twins, in: *2024 28th International Symposium on Distributed Simulation and Real Time Applications (DS-RT)*, IEEE, 2024, pp. 18–25. <https://doi.org/10.1109/DS-RT62209.2024.00013>
- [62] M. Schneider, L. Halekotte, T. Comes, D. Lichte, F. Fiedrich, Emergency response inference mapping (ERIMap): a bayesian network-based method for dynamic observation processing, *Reliab. Eng. Syst. Saf.* 255 (2025) 110640. <https://doi.org/10.1016/j.res.2024.110640>
- [63] S. Brusca, F. Famoso, R. Lanzafame, S. Mauro, A.M.C. Garrano, P. Monforte, Theoretical and experimental study of Gaussian plume model in small scale system, *Energy Procedia* 101 (2016) 58–65. <https://doi.org/10.1016/j.egypro.2016.11.008>
- [64] I. Gjerde, L. Scott, Nitsche's method for Navier–Stokes equations with slip boundary conditions, *Math. Comput.* 91 (334) (2022) 597–622. <https://doi.org/10.1090/mcom/3682>
- [65] N. Halko, P.G. Martinsson, J.A. Tropp, Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions, *SIAM Rev.* 53 (2) (2011) 217–288. <https://doi.org/10.1137/090771806>
- [66] M. Mattuschka, N.A.d. Lan, M. von Danwitz, D. Wolff, A. Popp, Goal-oriented optimal sensor placement for PDE-constrained inverse problems in crisis management. <https://doi.org/10.48550/arXiv.2507.02500>
- [67] S. Sun, T. Wang, F. Chu, J. Tan, Acoustic source identification using an off-grid and sparsity-based method for sound field reconstruction, *Mech. Syst. Signal Process.* 170 (2022) 108869. <https://doi.org/10.1016/j.ymsp.2022.108869>
- [68] M. Scholl, M. Passek, M. Lainer, F. Taddei, F. Schneider, G. Müller, Data-based localisation methods using simulated data with application to small-scale structures, *Arch. Appl. Mech.* 93 (6) (2023) 2411–2433. <https://doi.org/10.1007/s00419-023-02388-2>
- [69] I. Mazzieri, M. Muhr, M. Stupazzini, B. Wohlmuth, Elasto-acoustic modeling and simulation for the seismic response of structures: the case of the Tahtali dam in the 2020 Izmir earthquake, *J. Comput. Phys.* 466 (2022) 111411. <https://doi.org/10.1016/j.jcp.2022.111411>
- [70] D.J. Jerez, H.A. Jensen, M. Beer, M. Broggi, Contaminant source identification in water distribution networks: a Bayesian framework, *Mech. Syst. Signal Process.* 159 (2021) 107834. <https://doi.org/10.1016/j.ymsp.2021.107834>