

LEAK LOCALISATION WITH A MEASURE SOURCE CONVECTION–DIFFUSION MODEL

Thi Tam Dang*

Tuomo Valkonen†

Abstract We study the inverse problem of locating gas leaks from line-of-sight concentration measurements using a convection–diffusion model with the source term a Radon measure. By imposing sparsity-promoting regularisation on this measure, we recover point sources—identifying both their locations and intensities—rather than diffuse approximations. We jointly estimate the underlying physical convection (wind) and diffusion parameters. Our main theoretical contribution is the stability analysis of the convection–diffusion equation with respect to its parameters: the measure, and the convection and diffusion fields. Numerically, we employ a semi-grid-free optimisation approach for reconstructing the source measure. Our experiments demonstrate accurate localisation, highlighting the potential of the method for practical gas emission detection.

1 INTRODUCTION

Modern industrial facilities require reliable gas leak detection to ensure safety and meet regulatory standards. Accurate quantification of greenhouse gas emissions, particularly methane, is essential for effective climate mitigation [3]. Gas emission detection (GED) uses various instruments: remote sensing from aircraft and satellites, point-wise systems such as eddy-covariance towers, and open-path spectrometers. Satellite-based systems are efficient for regional to global emission estimation but often lack accuracy at local scales and in distinguishing different source types (for example, agricultural versus industrial) [19, 25]. Eddy-covariance towers provide high-frequency data but only at fixed locations [4, 26]. In contrast, open-path laser dispersion spectroscopy (LDS) helps to bridge this gap by providing path-averaged concentration measurements along multiple beams, which enables rapid two-dimensional emission mapping that is robust under varying atmospheric conditions [1, 15, 21, 29, 32, 33]. Because all these instruments measure gas concentrations rather than emission rates directly, their data must be combined with an inversion model to localize and quantify the underlying emission sources.

Recovering the full gas concentration field from sparse line-of-sight measurements is an ill-posed inverse problem governed by a parabolic convection-diffusion PDE (2.3). In this work, the source term in the PDE will be a measure, to facilitate modelling discrete sources of leaks, both their locations and release rates:

$$(1.1) \quad \mu = \sum_{i=1}^N \beta_i \delta_{x_i},$$

This research has been supported by the Jane and Aatos Erkko Foundation.

*Department of Mathematics and Statistics, University of Helsinki, Finland; tam.dang@iki.fi, ORCID: 0009-0005-1893-5351

†MODEMAT Research Center in Mathematical Modeling and Optimization, Quito, Ecuador *and* Department of Mathematics and Statistics, University of Helsinki, Finland; tuomo.valkonen@iki.fi, ORCID: 0000-0001-6683-3572

where β_i is the release rate of the leak, and $x_i \in \Omega \subset \mathbb{R}^d$ its location. We denote by δ_x the Dirac measure concentrated at x . To reconstruct such a μ , we consider optimisation problems of the form

$$(1.2) \quad \min_{\mu, k, c} \frac{1}{2} \|Au - b\|^2 + \alpha \|\mu\|_{\mathcal{M}(\Omega)} + \delta_{\geq 0}(\mu) + R_{k,c}(k, c),$$

where the concentration $u : \Omega \times [0, T] \rightarrow \mathbb{R}$ arises as a solution of the convection–diffusion equation for the sources μ , convection field c , and diffusion field k . The observation operator A performs line integration of the concentration from laser sources to mirrors over multiple time instants: our observations are over time, while the length T of the time interval is—for simplicity—assumed to be short enough that μ is time-independent. The Radon-norm regularisation term with parameter $\alpha > 0$ encourages sparsity of solutions [6]: that an optimal μ would have the form (1.1), instead of being diffuse. The term $\delta_{\geq 0}$ (an indicator function, not to be confused with the Dirac δ_x -measure) enforces that there are only sources, no sinks. The convection and diffusion fields’ measurements and regularisation are modelled by $R_{k,c}$.

The dependence of u on $x = (\mu, k, c)$ is nonlinear. Conventionally, if the dependence were linear, and we would not have the auxiliary parameters k and c , this type of measure optimisation problems would be solved by a conditional gradient (Frank–Wolfe) method [5, 7, 8, 17, 18, 28] in a grid-free fashion. Semismooth Newton, particle-based, and semi-infinite approaches have also been developed [10–14, 20]. More recently [30, 31], we introduced more conventional forward-backward and primal-dual methods that more easily extend to nonconvex problems and auxiliary parameters. We will exploit those in Section 5 to solve the problem (1.2). The algorithms are described in more detail in Section 3. To be able to apply these algorithms, we will require the PDE solution operator $S_u : (\mu, k, c) \rightarrow u$ to be Lipschitz-differentiable in appropriate spaces. We will in Section 2 perform the relevant analysis. This forms the principal theoretical contribution of this manuscript.

In Section 4, we introduce a finite element approach that couples the convection-diffusion PDE with a computationally simple quadrature-free observation operator for LDS measurements. Unlike accurate ray-tracing or X-ray transform discretizations prevalent in computed tomography [23] and optical tomography [2], which demand sophisticated mesh intersection algorithms, our approach prioritizes computational simplicity and adjoint consistency over quadrature precision. For strongly regularized inverse problems, this internal consistency between forward and adjoint operators proves more critical than high-fidelity line integration, as discretization artifacts in the adjoint dominate reconstruction errors [22].

Notation For normed spaces X and Y , we write $\mathbb{L}(X; Y)$ for the space of bounded linear operators from X to Y . The dual space of X is denoted by X^* . For a functional $F : X \rightarrow \mathbb{R}$, we denote by $DF(x)$ the Gâteaux derivative of F at x , and by $F'(x) \in X^*$ the Fréchet derivative at x , whenever it exists. If X is a Hilbert space, $\nabla F(x) \in X$ stands for the Riesz representative of $F'(x)$, i.e., the gradient of F at x . For partial derivatives, we use the notation $F^{(u)}(u, x)$ when differentiating with respect to a specific named argument.

The space $\mathcal{M}(\Omega)$ is that of Radon measures on the domain $\Omega \subset \mathbb{R}^d, 1 \leq d \leq 3$.

We write $\langle x^* | x \rangle_{X^*, X} = x^* x$ for the dual pairing, and $\langle x, y \rangle_X$ for the inner product in Hilbert spaces. When the spaces are clear from the context, we do not indicate them.

We generally write $\xi \in \mathbb{R}^d$ for the spatial variable, with x reserved for the overall unknown.

We write δ_A for the $\{0, \infty\}$ -valued indicator function of a set A . We also use δ_x to denote the Dirac mass at the point x .

2 THE CONVECTION–DIFFUSION EQUATION WITH MEASURE SOURCES

This section formulates the PDE-constrained inverse problem under study (Section 2.1). The forward PDE (2.3) is a parabolic convection-diffusion problem with a measure source μ , modelling the locations and intensities of the unknown leaks. Also the convection and diffusion coefficient are unknown. We first establish well-posedness of the PDE (Theorem 2.2, Section 2.2), and Lipschitz continuity of its solution operator (Section 2.3). To differentiate the solution operator, Section 2.4 derives the adjoint equation and corresponding Lipschitz estimates. Finally, Section 2.5 combines the forward and adjoint equations to obtain Lipschitz estimates for the differential of the data term. This is necessary for the application of the algorithms of Section 3.

2.1 PROBLEM SETTING

Let $\Omega \subset \mathbb{R}^d$, ($1 \leq d \leq 3$), be a bounded domain with convex geometry or with boundary Γ of class $C^{1,1}$. For a given end time $T > 0$, write $I := [0, T]$ and $\Omega_T := \Omega \times (0, T)$. Recalling (1.2), we consider the PDE-constrained optimisation problem

$$(2.1) \quad \min_{\mu, k, c} J(S_u(\mu, k, c)) + R(\mu, k, c)$$

for the convection–diffusion PDE solution operator S_u (described below) and the specific choices

$$(2.2) \quad J(u) = \frac{1}{2} \|Au - b\|_2^2 \quad \text{and} \quad R(\mu, k, c) = \alpha \|\mu\|_{\mathcal{M}(\Omega)} + \delta_{\geq 0}(\mu) + R_{k,c}(k, c).$$

The unknowns are the measure $\mu \in \mathcal{M}(\Omega)$, the diffusion coefficient $k \in L^\infty(\Omega_T)$, and the convection coefficient $c \in L^\infty(\Omega_T; \mathbb{R}^d)$. Here $k_{\min} > 0$ ensures uniform ellipticity (non-degenerate diffusion), while $k_{\max}, c_{\max} > 0$ enforce physical bounds for numerical stability. Physically, $k(\xi, t) > 0$ controls diffusive spreading, $c(\xi, t)$ represents background advection (wind currents). The measure μ models the locations and intensities of gas leaks. The Radon-norm regulariser, with regularisation parameter $\alpha > 0$, seeks to impose sparsity of this measure: that it would consist of a finite number of point sources (Dirac masses). The indicator function $\delta_{\geq 0}$ enforces that we only have sources, no sinks. The convection and diffusion fields are regularised by the convex function $R_{k,c} : L^\infty(\Omega_T) \times L^\infty(\Omega_T; \mathbb{R}^d) \rightarrow \bar{\mathbb{R}}$; we will in Section 5 take it as a simple quadratic regulariser combined with constraints that restrict k and c to scalar subspaces

$$\begin{aligned} K &\subset \{k \in L^\infty(\Omega_T) : k_{\min} \leq k(\xi, t) \leq k_{\max} \text{ a.e. } \xi \in \Omega\} \quad \text{and} \\ C &\subset L^\infty(\Omega_T; [-c_{\max}, c_{\max}]^d), \end{aligned}$$

The measurements $b \in \mathbb{R}^m$ are obtained with a system of lasers and mirrors, modelled by the observation operator $A \in \mathbb{L}(L^2(\Omega_T); \mathbb{R}^m)$. For each $x = (\mu, k, c)$, the corresponding gas concentration $u = S_u(x) \in L^p(I; W^{1,p}(\Omega))$ with $p \in [1, \infty)$ is the unique solution of the convection–diffusion equation with initial and boundary values,

$$(2.3) \quad \begin{aligned} \partial_t u - \nabla \cdot (k \nabla u) + c \cdot \nabla u &= \mu && \text{in } \Omega_T, \\ u &= g && \text{on } \Sigma_1 := \Gamma_1 \times (0, T), \\ \frac{\partial u}{\partial n} &= 0 && \text{on } \Sigma_2 := \Gamma_2 \times (0, T), \\ u(\cdot, 0) &= u_0 && \text{in } \Omega, \end{aligned}$$

where $u_0 \in L^2(\Omega)$ denotes the given initial condition, $g \in L^2(\Sigma_1)$ the prescribed boundary flux.

2.2 EXISTENCE OF WEAK SOLUTIONS

Let $I = [0, T]$ and $1 < p < \infty$. We denote by $W^{1,p}(\Omega)$ the Sobolev space of functions in $L^p(\Omega)$ whose distributional derivatives belong to $L^p(\Omega)$, and set $W^{-1,p'}(\Omega)$ to be its dual, where $1/p + 1/p' = 1$. These spaces are reflexive and separable. Consequently, the space

$$L^2(I; W^{1,p}(\Omega)) := \{u : [0, T] \rightarrow W^{1,p}(\Omega) \text{ measurable} \mid \|u\|_{L^2(I; W^{1,p}(\Omega))} < \infty\},$$

equipped with the norm

$$\|u\|_{L^2(I; W^{1,p}(\Omega))} := \left(\int_0^T \|u(t)\|_{W^{1,p}(\Omega)}^2 dt \right)^{1/2},$$

is a separable and reflexive Banach space, whose dual can be identified with $L^2(I; W^{-1,p'}(\Omega))$.

We also work with $L^2(I; C_0(\Omega))$ equipped with the norm

$$\|u\|_{L^2(I; C_0(\Omega))} := \left(\int_0^T \|u(t)\|_{\infty}^2 dt \right)^{1/2} < \infty.$$

The action of μ on time-dependent functions $v \in L^2(I; C_0(\Omega))$ is defined through the canonical duality between $\mathcal{M}(\Omega)$ and $C_0(\Omega)$, integrated over the time interval I . More precisely, we define

$$\langle \mu, v \rangle_{\mathcal{M}(\Omega), L^2(I; C_0(\Omega))} := \int_0^T \langle \mu, v(t) \rangle_{\mathcal{M}(\Omega), C_0(\Omega)} dt,$$

where

$$\langle \mu, v \rangle_{\mathcal{M}(\Omega), C_0(\Omega)} := \int_{\Omega} v d\mu, \quad v \in C_0(\Omega).$$

We may now introduce the concept of a weak solution to (2.3):

Definition 2.1 (Weak Solution). A function $u \in L^2(I; W^{1,p}(\Omega))$ is a weak solution to problem (2.3) if it satisfies the initial condition $u(0) = u_0$ in $L^2(\Omega)$, and, for all $v \in L^2(I; W^{1,p}(\Omega))$,

$$\int_{\Omega_T} (-\partial_t u v + k \nabla u \cdot \nabla v + (c \cdot \nabla u) v) d\xi dt = \int_0^T \langle \mu, v(t) \rangle dt + \int_{\Sigma_1} g v d\sigma dt.$$

Theorem 2.2 (Existence and uniqueness). For every $(\mu, u_0) \in \mathcal{M}(\Omega) \times L^2(\Omega)$, $1 < p < d/(d-1)$, the problem (2.3) admits a unique weak solution $u \in L^2(I; W^{1,p}(\Omega))$. Moreover, there exist constants $C_p > 0$, independent of μ and u_0 , such that

$$\|u\|_{L^2(I; W^{1,p}(\Omega))} \leq C_p (\|\mu\|_{\mathcal{M}(\Omega)} + \|g\|_{L^2(\Sigma_1)} + \|u_0\|_{L^2(\Omega)}).$$

The proof follows the strategy of [11, Theorem 2.2]: approximate $\mu \in \mathcal{M}(\Omega)$ by continuous functions μ_k , obtain uniform $L^2(I; W_0^{1,p}(\Omega))$ bounds via maximal L^p -regularity of the sectorial convection-diffusion operator, pass to weak limit in the variational formulation, and prove uniqueness by energy method using test function solving the dual equation. Key adaptations are replacing the Laplacian Δ by $Ev = -\nabla \cdot (k \nabla v) + c \cdot \nabla v$ (sectorial on $W_0^{1,p}(\Omega)$ for mixed boundary conditions), and adding Dirichlet data $g \in L^2(\Sigma_1)$ via trace estimates.

Remark 2.3 (Sectorial operators). The operator

$$Ev = -\operatorname{div}(k \nabla v) + c \cdot \nabla v$$

is sectorial on $W_0^{1,p}(\Omega)$; that is, its spectrum is contained in the sector

$$\Sigma_{\pi/4} = \{z \in \mathbb{C} : |\arg(z)| \leq \pi/4\} \cup \{0\},$$

and its resolvent is bounded outside larger sectors, i.e., $\|\lambda - E\|^{-1} \leq C/\|\lambda\|$ for all λ outside any larger sector Σ_φ with $\varphi > \pi/4$. This follows from the uniform ellipticity condition

$$k \geq k_{\min} > 0.$$

As a consequence, the operator $\partial_t + E$ satisfies maximal L^p -regularity, that is

$$\partial_t v + E v = f \in L^p(\Omega_T) \implies v, \partial_t v, E v \in L^p(\Omega_T),$$

which is essential for establishing Lipschitz continuity of $S_u(\cdot)$. For more on sectorial operators, see [27, Chap. 2].

2.3 LIPSCHITZ CONTINUITY OF THE SOLUTION MAPPING

For simplicity of notation, we define the space of concentrations (i.e., the state space in the terminology of optimal control), the full space of test functions, and its dual,

$$U := L^2(I; W^{1,p}(\Omega)), \quad W := L^2(I; W_0^{1,p}(\Omega)) \times L^2(\Sigma_1), \quad \text{and} \quad W^* := L^2(I; W^{-1,p'}(\Omega)) \times L^2(\Sigma_1).$$

The full parameter space is the product space

$$X := \mathcal{M}(\Omega) \times L^\infty(\Omega_T) \times L^\infty(\Omega_T; \mathbb{R}^d),$$

equipped with the norm

$$\|x\|_X := \|\mu\|_{\mathcal{M}(\Omega)} + \|k\|_{L^\infty(\Omega_T)} + \|c\|_{L^\infty(\Omega_T; \mathbb{R}^d)} \quad \text{for } x = (\mu, k, c) \in X.$$

Assumption 2.4. We assume throughout that:

- (i) $\Omega \subset \mathbb{R}^d$, $1 \leq d \leq 3$ is a bounded Lipschitz domain, $I = (0, T)$ with $T > 0$, and $\Omega_T = \Omega \times I$.
- (ii) The diffusion coefficient $k : \Omega_T \rightarrow \mathbb{R}$ satisfies

$$0 < k_{\min} \leq k(\xi, t) \leq k_{\max} < \infty \quad \text{for a.e. } (\xi, t) \in \Omega_T,$$

and $k \in L^\infty(\Omega_T)$.

- (iii) The convection field $c : \Omega_T \rightarrow \mathbb{R}^d$ satisfies

$$c \in L^\infty(\Omega_T; \mathbb{R}^d), \quad \|c\|_{L^\infty(\Omega_T; \mathbb{R}^d)} \leq C_c$$

and, when needed for uniqueness and a priori estimates, we assume $\operatorname{div} c \in L^\infty(\Omega_T)$.

- (iv) The boundary data and the initial condition satisfy

$$g \in L^2(\Sigma_1), \quad u_0 \in L^2(\Omega).$$

Under [Assumption 2.4](#), the operator $e : U \times X \rightarrow W^*$ is defined by

$$(2.4a) \quad \langle e(u, \mu, k, c)|v \rangle_{W^*, W} := \int_{\Omega_T} (-\partial_t u v_\Omega + k \nabla u \cdot \nabla v_\Omega + (c \cdot \nabla u) v_\Omega) \, d\xi \, dt \\ - \int_0^T \langle \mu, v_\Omega(t) \rangle \, dt + \int_{\Sigma_1} (\operatorname{trace}_{\Sigma_1} u - g) v_{\Gamma_1} \, d\sigma \, dt$$

for all $v = (v_\Omega, v_{\Gamma_1}) \in W$. Furthermore, we introduce the associated solution mapping

$$(2.4b) \quad S_u : X \rightarrow U, \quad x = (\mu, k, c) \mapsto u.$$

Then $u = S_u(x)$ solve the weak formulation of [Definition 2.1](#) if and only if

$$(2.4c) \quad e(S_u(x), x) = 0 \quad \text{in } W^*.$$

Next we study the Lipschitz continuity of S_u with respect to the parameters.

Lemma 2.5 (Lipschitz continuity of S_u). *Let Assumption 2.4 be fulfilled. Then, for some $L_S \geq 0$,*

$$\|S_u(x_2) - S_u(x_1)\|_U \leq L_S \|x_2 - x_1\|_X \quad \text{for all } x_1, x_2 \in X.$$

Proof. Let $x_1 = (\mu_1, k_1, c_1)$, $x_2 = (\mu_2, k_2, c_2)$. Abbreviate $u_1 = S_u(x_1)$ and $u_2 = S_u(x_2)$. We set $w := u_2 - u_1$. By definition, u_i solves the weak formulation (2.4a). Subtracting the two equations we obtain, for all $v \equiv v_\Omega \in W$ (the boundary term cancels out in the differences),

$$(2.5) \quad \int_{\Omega_T} (\partial_t w v + k_1 \nabla w \cdot \nabla v + (c_1 \cdot \nabla w) v) \, d\xi \, dt = \int_0^T \langle \mu_2 - \mu_1, v(t) \rangle \, dt \\ + \int_{\Omega_T} ((k_2 - k_1) \nabla u_2 \cdot \nabla v + ((c_2 - c_1) \cdot \nabla u_2) v) \, d\xi \, dt.$$

Now, fix x_1 , and consider the dual problem

$$(2.6) \quad -\partial_t z - \nabla \cdot (k_1 \nabla z) + c_1 \cdot \nabla z = w \quad \text{in } \Omega_T, \quad z|_{\Sigma_T} = 0, \quad z(T) = 0.$$

Under the assumptions that Ω is a bounded Lipschitz domain, $k_1 \geq k_{\min} > 0$ uniformly elliptic, $c_1 \in L^\infty(\Omega; \mathbb{R}^d)$, and $w \in L^2(\Omega_T)$, (2.6) admits a unique weak solution

$$z \in L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)) \cap C([0, T]; L^2(\Omega)).$$

Maximal parabolic regularity (Theorem 2.2) and the compactness of the embedding $L^2(0, T; H_0^1(\Omega)) \hookrightarrow L^2(\Omega_T)$ establish a constant $C > 0$ (independent of x_1 in the admissible set) such that

$$\|z\|_{L^2(I; C_0(\Omega))} \leq C \|w\|_{L^2(\Omega_T)}.$$

Choosing $v = z$ in (2.5) and using the dual equation (2.6), we get

$$\|w\|_{L^2(\Omega_T)}^2 = \int_{\Omega_T} w^2 \, d\xi \, dt = \int_{\Omega_T} (\partial_t w z + k_1 \nabla w \cdot \nabla z + (c_1 \cdot \nabla w) z) \, d\xi \, dt \\ = \int_0^T \langle \mu_2 - \mu_1, z(t) \rangle \, dt + \int_{\Omega_T} ((k_2 - k_1) \nabla u_2 \cdot \nabla z + ((c_2 - c_1) \cdot \nabla u_2) z) \, d\xi \, dt.$$

We now estimate the right-hand side term by term.

For the measure difference term, the duality between $\mathcal{M}(\Omega)$ and $C_0(\Omega)$ yields

$$\left| \int_0^T \langle \mu_2 - \mu_1, z(t) \rangle \, dt \right| \leq \|\mu_2 - \mu_1\|_{\mathcal{M}(\Omega)} \|z\|_{L^2(I; C_0(\Omega))} \\ \leq C \|\mu_2 - \mu_1\|_{\mathcal{M}(\Omega)} \|w\|_{L^2(\Omega_T)}.$$

The diffusion coefficient difference term can be estimated by

$$\left| \int_{\Omega_T} (k_2 - k_1) \nabla u_2 \cdot \nabla z \, d\xi \, dt \right| \leq \|k_2 - k_1\|_{L^\infty(\Omega_T)} \|\nabla u_2\|_{L^2(\Omega_T)} \|\nabla z\|_{L^2(\Omega_T)} \\ \leq C \|k_2 - k_1\|_{L^\infty(\Omega_T)} \|u_2\|_U \|w\|_{L^2(\Omega_T)},$$

where we used $\|\nabla z\|_{L^2(\Omega_T)} \leq C \|w\|_{L^2(\Omega_T)}$ from standard energy estimates (see Theorem 2.2) and $\|\nabla u_2\|_{L^2(\Omega_T)} \leq C \|u_2\|_U$. Finally, the convection coefficient difference term yields

$$\left| \int_{\Omega_T} ((c_2 - c_1) \cdot \nabla u_2) z \, dx \, dt \right| \leq \|c_2 - c_1\|_{L^\infty(\Omega_T; \mathbb{R}^d)} \|\nabla u_2\|_{L^2(\Omega_T)} \|z\|_{L^2(\Omega_T)} \\ \leq C \|c_2 - c_1\|_{L^\infty(\Omega_T; \mathbb{R}^d)} \|u_2\|_U \|w\|_{L^2(\Omega_T)}.$$

Collecting these estimates, we obtain

$$\|w\|_{L^2(\Omega_T)}^2 \leq C \left(\|\mu_1 - \mu_1\|_{\mathcal{M}(\Omega)} + \|k_2 - k_1\|_{L^\infty(\Omega_T)} \|u_2\|_U + \|c_2 - c_1\|_{L^\infty(\Omega_T; \mathbb{R}^d)} \|u_2\|_U \right) \|w\|_{L^2(\Omega_T)}.$$

If $w \neq 0$, divide by $\|w\|_{L^2(\Omega_T)}$ to get

$$\|w\|_{L^2(\Omega_T)} \leq C \left(\|\mu_1 - \mu_1\|_{\mathcal{M}(\Omega)} + (\|k_2 - k_1\|_{L^\infty(\Omega_T)} + \|c_2 - c_1\|_{L^\infty(\Omega_T; \mathbb{R}^d)}) \|u_2\|_U \right).$$

By the uniform a priori estimate from [Theorem 2.2](#),

$$\|u_2\|_U \leq C_p (\|\mu_2\|_{\mathcal{M}(\Omega)} + \|g\|_{L^2(\Sigma_1)} + \|u_0\|_{L^2(\Omega)}),$$

and since the parameter set is bounded in X , we have $\|u_2\|_U \leq C$ uniformly in x_2 . Therefore,

$$\|w\|_{L^2(\Omega_T)} \leq C_p \|x_2 - x_1\|_X.$$

Testing the difference equation with $v = w$ yields

$$\frac{d}{dt} \|w(t)\|_{L^2(\Omega)}^2 + k_{\min} \|\nabla w(t)\|_{L^2(\Omega)}^2 \leq C_p \|x_2 - x_1\|_X^2,$$

which implies

$$\|w\|_{L^2(I; H^1(\Omega))} \leq C_p \|x_2 - x_1\|_X.$$

Using the embeddings $H^1(\Omega) \hookrightarrow W_0^{1,p}(\Omega)$ for the range of $p \leq 2d/(d-2)$ in [Theorem 2.2](#), and the definition of $U = L^2(I; W_0^{1,p}(\Omega))$, we finally obtain for some $L_S > 0$ that

$$\|w\|_U \leq L_S \|x_2 - x_1\|_X,$$

where $L_S = C(1 + \sup_{x \in X} \|S_u(x)\|_U)$, $C = C(T, k_{\min}, \|c\|_{L^\infty(\Omega_T; \mathbb{R}^d)})$. □

Lemma 2.6 (Fréchet differentiability of the weak form). *Let [Assumption 2.4](#) hold. The operator e defined in [\(2.4a\)](#) is (partially) Fréchet differentiable at every point $(u, \mu, k, c) \in U \times X$ with respect to each argument. The partial derivatives are bounded linear operators*

$$e^{(u)} \in \mathbb{L}(U; W^*), \quad e^{(\mu)} \in \mathbb{L}(\mathcal{M}(\Omega); W^*), \quad e^{(k)} \in \mathbb{L}(L^\infty(\Omega_T); W^*), \quad \text{and} \quad e^{(c)} \in \mathbb{L}(L^\infty(\Omega_T; \mathbb{R}^d); W^*),$$

defined for $v = (v_\Omega, v_{\Gamma_1}) \in W$ by

$$(2.7a) \quad \langle e^{(u)}(u, \mu, k, c) \tilde{u} | v \rangle_{W^*, W} = \int_{\Omega_T} (-\partial_t \tilde{u} v_\Omega + k \nabla \tilde{u} \cdot \nabla v_\Omega + (c \cdot \nabla \tilde{u}) v_\Omega) \, dx \, dt + \int_{\Sigma_1} g v_{\Gamma_1} \, d\sigma \, dt,$$

$$(2.7b) \quad \langle e^{(\mu)}(u, \mu, k, c) \eta | v \rangle_{W^*, W} = - \int_0^T \langle \eta, v_\Omega(t) \rangle \, dt,$$

$$(2.7c) \quad \langle e^{(k)}(u, \mu, k, c) h | v \rangle_{W^*, W} = \int_{\Omega_T} h \nabla u \cdot \nabla v_\Omega \, d\xi \, dt,$$

$$(2.7d) \quad \langle e^{(c)}(u, \mu, k, c) d | v \rangle_{W^*, W} = \int_{\Omega_T} (d \cdot \nabla u) v_\Omega \, d\xi \, dt,$$

for all $v \in W$, $\tilde{u} \in U$, $\eta \in \mathcal{M}(\Omega)$, $h \in L^\infty(\Omega_T)$, $d \in L^\infty(\Omega_T; \mathbb{R}^d)$.

Proof. We verify Fréchet differentiability with respect to each argument separately, using the definition of $\langle e(u, \mu, k, c)|v \rangle$ from (2.4a).

Proof of (2.7a). Let $u_\varepsilon = u + \varepsilon \tilde{u}$ with $\tilde{u} \in U$. Then for $v = (v_\Omega, v_{\Gamma_1}) \in W$,

$$\begin{aligned} \langle e(u_\varepsilon, \mu, k, c)|v \rangle &= \int_{\Omega_T} (-\partial_t(u + \varepsilon \tilde{u})v_\Omega + k\nabla(u + \varepsilon \tilde{u}) \cdot \nabla v_\Omega + (c \cdot \nabla(u + \varepsilon \tilde{u}))v_\Omega) d\xi dt \\ &\quad - \int_0^T \langle \mu, v_\Omega(t) \rangle dt + \int_{\Sigma_1} ((u + \varepsilon \tilde{u}) - g)v_{\Gamma_1} d\sigma dt \\ &= \langle e(u, \mu, k, c)|v \rangle + \varepsilon \left[\int_{\Omega_T} (-\partial_t \tilde{u} v_\Omega + k\nabla \tilde{u} \cdot \nabla v_\Omega + (c \cdot \nabla \tilde{u})v_\Omega) d\xi dt + \int_{\Sigma_1} \tilde{u}|_{\Sigma_1} v_{\Gamma_1} d\sigma dt \right]. \end{aligned}$$

Dividing by ε and letting $\varepsilon \rightarrow 0$ yields (2.7a). For continuity, we use the standard characterization of the norm in W^* :

$$\|e^{(u)}(u, \mu, k, c)\tilde{u}\|_{W^*} = \sup_{v \in W, \|v\|_W=1} \langle e^{(u)}(u, \mu, k, c)\tilde{u}|v \rangle.$$

For any $\tilde{u} \in U$ and $v \in W$, we estimate

$$\begin{aligned} \left| \int_{\Omega_T} -\partial_t \tilde{u} v_\Omega d\xi dt \right| &\leq \|\partial_t \tilde{u}\|_{L^2(I; W^{-1, p'}(\Omega))} \|v_\Omega\|_{L^2(I; W_0^{1, p'}(\Omega))}, \\ \left| \int_{\Omega_T} k\nabla \tilde{u} \cdot \nabla v_\Omega d\xi dt \right| &\leq \|k\|_{L^\infty(\Omega_T)} \|\nabla \tilde{u}\|_{L^p(\Omega_T)} \|\nabla v_\Omega\|_{L^{p'}(\Omega_T)}, \\ \left| \int_{\Omega_T} (c \cdot \nabla \tilde{u})v_\Omega d\xi dt \right| &\leq \|c\|_{L^\infty(\Omega_T; \mathbb{R}^d)} \|\nabla \tilde{u}\|_{L^p(\Omega_T)} \|v_\Omega\|_{L^{p'}(\Omega_T)}, \quad \text{and} \\ \left| \int_{\Sigma_1} \tilde{u}|_{\Sigma_1} v_{\Gamma_1} d\sigma dt \right| &\leq \|\tilde{u}|_{\Sigma_1}\|_{L^2(\Sigma_1)} \|v_{\Gamma_1}\|_{L^2(\Sigma_1)}. \end{aligned}$$

By trace theorem and Poincaré inequality,

$$\|\tilde{u}|_{\Sigma_1}\|_{L^2(\Sigma_1)} + \|v_\Omega\|_{L^{p'}(\Omega_T)} \leq C_P(\|\tilde{u}\|_U + \|v_\Omega\|_W).$$

Again, by the Poincaré inequality [9, Corollary 9.19] on $W_0^{1, p'}(\Omega)$, we have $\|v\|_{L^{p'}(\Omega_T)} \leq C_P \|\nabla v\|_{L^{p'}(\Omega_T)}$, so all three terms are bounded by

$$C(\|\partial_t \tilde{u}\|_{L^2(I; W^{-1, p'}(\Omega))} + (\|k\|_{L^\infty(\Omega_T)} + \|c\|_{L^\infty(\Omega_T; \mathbb{R}^d)}) \|\nabla \tilde{u}\|_{L^p(\Omega_T)}) \|v\|_W,$$

for some $C > 0$ depending only on Ω, Γ_2, p, T . Taking the supremum over $\|v\|_W = 1$ gives

$$\|e^{(u)}(u, \mu, k, c)\tilde{u}\|_{W^*} \leq C(1 + \|k\|_{L^\infty(\Omega_T)} + \|c\|_{L^\infty(\Omega_T; \mathbb{R}^d)}) \|\tilde{u}\|_U,$$

and, therefore, $e^{(u)} \in \mathbb{L}(U; W^*)$.

Remark 2.7. Since the Fréchet derivatives of e with respect to μ, k, c act only on the volume component v_Ω and are independent of the boundary test function v_{Γ_1} , we may write $v := v_\Omega$ for brevity in subsequent calculations.

Proof of (2.7b). Let μ_ε be a variation of μ in the space $\mathcal{M}(\Omega)$, and denote $\eta \in \mathcal{M}(\Omega)$ so that the directional increment is $\mu_\varepsilon = \mu + \varepsilon\eta$. Since μ is time-independent, we have

$$\langle \mu_\varepsilon, v(t) \rangle = \langle \mu, v(t) \rangle + \varepsilon \langle \eta, v(t) \rangle \quad \text{for all } t \in I.$$

Then

$$\begin{aligned} \langle e(u, \mu_\varepsilon, k, c)|v \rangle &= \int_{\Omega_T} (-\partial_t u v + k \nabla u \cdot \nabla v + (c \cdot \nabla u) v) dx dt - \int_0^T \langle \mu_\varepsilon, v(t) \rangle dt - \int_{\Sigma_1} g v d\sigma dt \\ &= \langle e(u, \mu, k, c)|v \rangle - \varepsilon \int_0^T \langle \eta, v(t) \rangle dt. \end{aligned}$$

The expression is linear in ε , so

$$\langle e^{(\mu)}(u, \mu, k, c) \eta, v \rangle = - \int_0^T \langle \eta, v(t) \rangle dt.$$

Boundedness follows from the fact that

$$|\langle e^{(\mu)}(u, \mu, k, c) \eta|v \rangle| \leq \|\eta\|_{\mathcal{M}(\Omega)} \|v\|_{L^2(I; C_0(\Omega))},$$

using the compact embedding $W \hookrightarrow L^2(I; C_0(\Omega))$, so $e^{(\mu)} \in \mathbb{L}(\mathcal{M}(\Omega); W^*)$.

Proof of (2.7c). Let $k_\varepsilon = k + \varepsilon h$ with $h \in L^\infty(\Omega_T)$. Then

$$\langle e(u, \mu, k_\varepsilon, c)|v \rangle = \langle e(u, \mu, k, c)|v \rangle + \varepsilon \int_{\Omega_T} h \nabla u \cdot \nabla v d\xi dt.$$

Thus

$$\langle e^{(k)}(u, \mu, k, c) h|v \rangle = \int_{\Omega_T} h \nabla u \cdot \nabla v d\xi dt.$$

Boundedness follows from

$$|\langle e^{(k)}(u, \mu, k, c) h|v \rangle| \leq \|h\|_{L^\infty(\Omega_T)} \|\nabla u\|_{L^p(\Omega_T)} \|\nabla v\|_{L^{p'}(\Omega_T)},$$

so $e^{(k)} \in \mathbb{L}(L^\infty(\Omega_T); W^*)$.

Proof of (2.7d). Let $c_\varepsilon = c + \varepsilon d$ with $d \in L^\infty(\Omega_T; \mathbb{R}^d)$. Then

$$\langle e(u, \mu, k, c_\varepsilon)|v \rangle = \langle e(u, \mu, k, c)|v \rangle + \varepsilon \int_{\Omega_T} (d \cdot \nabla u) v dx dt.$$

Thus

$$\langle e^{(c)}(u, \mu, k, c) d|v \rangle = \int_{\Omega_T} (d \cdot \nabla u) v dx dt,$$

and boundedness follows from

$$|\langle e^{(c)}(u, \mu, k, c) d|v \rangle| \leq \|d\|_{L^\infty(\Omega_T; \mathbb{R}^d)} \|\nabla u\|_{L^p(\Omega_T)} \|v\|_{L^{p'}(\Omega_T)},$$

so $e^{(c)} \in \mathbb{L}(L^\infty(\Omega_T; \mathbb{R}^d); W^*)$.

The boundary term $\int_{\Sigma_1} g v_{\Gamma_1} d\sigma dt$ is independent of (μ, k, c) and thus does not contribute to any partial derivative. \square

The following lemma provides Lipschitz continuity estimates for the partial derivatives of the PDE operator e in weak form.

Lemma 2.8. *Let Assumption 2.4 be satisfied and e be defined by (2.4a). Then, the partial derivative $e^{(u)}(u, x) : U \rightarrow W^*$ is uniformly Lipschitz with respect to both $u \in U$ and $x \in X$, i.e., there exist constants $L_{e^{(u)}|_u}, L_{e^{(u)}|_x} \geq 0$ such that*

$$(2.8) \quad \|e^{(u)}(u_2, x_2) - e^{(u)}(u_1, x_1)\|_{\mathbb{L}(U; W^*)} \leq L_{e^{(u)}|_u} \|u_2 - u_1\|_U + L_{e^{(u)}|_x} \|x_2 - x_1\|_X$$

for all $u_1, u_2 \in U$ and $x_1, x_2 \in X$.

Proof. Recall $e^{(u)}(u, x) : U \rightarrow W^*$ is defined by

$$\langle e^{(u)}(u, x) \tilde{u} | v \rangle_{W^*, W} = \int_{\Omega_T} (-\partial_t \tilde{u} v_\Omega + k \nabla \tilde{u} \cdot \nabla v_\Omega + (c \cdot \nabla \tilde{u}) v_\Omega) d\xi dt + \int_{\Sigma_1} \tilde{u} |_{\Sigma_1} v_{\Gamma_1} d\sigma dt,$$

for $\tilde{u} \in U, v = (v_\Omega, v_{\Gamma_1}) \in W$, where $x = (\mu, k, c)$ and X is the parameter space for x . The operator norm in $\mathbb{L}(U; W^*)$ is

$$\|e^{(u)}(u, x)\|_{\mathbb{L}(U; W^*)} = \sup_{\tilde{u} \in U, \|\tilde{u}\|_U=1} \|e^{(u)}(u, x) \tilde{u}\|_{W^*}.$$

Given pairs (u_1, x_1) and (u_2, x_2) in $U \times X$ and letting $v := v_\Omega$ for brevity, the difference of derivatives satisfies

$$\begin{aligned} & \langle (e^{(u)}(u_2, x_2) - e^{(u)}(u_1, x_1)) \tilde{u}, v \rangle \\ &= \int_{\Omega_T} ((k_2 - k_1) \nabla \tilde{u} \cdot \nabla v + (c_2 - c_1) \cdot \nabla \tilde{u} v) d\xi dt, \end{aligned}$$

since the term $-\partial_t v$ is independent of u and x . Estimating each term:

$$\begin{aligned} \left| \int_{\Omega_T} (k_2 - k_1) \nabla \tilde{u} \cdot \nabla v d\xi dt \right| &\leq \|k_2 - k_1\|_{L^\infty(\Omega_T)} \|\nabla \tilde{u}\|_{L^2(\Omega_T)} \|\nabla v\|_{L^2(\Omega_T)}, \\ \left| \int_{\Omega_T} (c_2 - c_1) \cdot \nabla \tilde{u} v d\xi dt \right| &\leq \|c_2 - c_1\|_{L^\infty(\Omega_T; \mathbb{R}^d)} \|\nabla \tilde{u}\|_{L^2(\Omega_T)} \|v\|_{L^2(\Omega_T)}. \end{aligned}$$

By the standard embedding $W \hookrightarrow L^2(\Omega_T)$ and the Poincaré inequality (for $W_0^{1,2}(\Omega)$ in the spatial case), we have

$$\|\nabla \tilde{u}\|_{L^2(\Omega_T)} \leq C \|\tilde{u}\|_U \quad \text{and} \quad \|v\|_{L^2(\Omega_T)} \leq C \|v\|_U.$$

Thus, for $k_1, k_2 \in L^\infty(\Omega_T)$ and $c_1, c_2 \in L^\infty(\Omega_T; \mathbb{R}^d)$, there exists a constant $C_E > 0$ such that

$$\left| \langle (e^{(u)}(u_2, x_2) - e^{(u)}(u_1, x_1)) \tilde{u}, v \rangle \right| \leq C_E (\|k_2 - k_1\|_{L^\infty(\Omega_T)} + \|c_2 - c_1\|_{L^\infty(\Omega_T; \mathbb{R}^d)}) \|\tilde{u}\|_U \|v\|_U.$$

Taking the supremum over $\|\tilde{u}\|_U = \|v\|_U = 1$,

$$\|e^{(u)}(u_2, x_2) - e^{(u)}(u_1, x_1)\|_{\mathbb{L}(U; W^*)} \leq C_E (\|k_2 - k_1\|_{L^\infty(\Omega_T)} + \|c_2 - c_1\|_{L^\infty(\Omega_T; \mathbb{R}^d)}).$$

Since the right-hand side depends only on the difference in $x = (\mu, k, c)$, and $\|x_2 - x_1\|_X$ includes the norms $\|k_2 - k_1\|_{L^\infty}$ and $\|c_2 - c_1\|_{L^\infty}$, and since $e^{(u)}(u, x)$ is actually independent of u , the term involving $\|u_2 - u_1\|_W$ is absent. Hence (2.8) holds with $L_{e^{(u)}|_u} = 0$ and $L_{e^{(u)}|_x} = C_E$. \square

Lemma 2.9 (Lipschitz continuity of $e^{(x)}(S_u(x), x)$). *Suppose Assumption 2.4 holds. Let*

$$e^{(x)}(S_u(x), x) : X \rightarrow W^*, \quad e^{(x)}(S_u(x), x)h := [x \mapsto e^{(x)}(S_u(x), x)]'(x)h,$$

denote the derivative (at x) of the map $x \mapsto e^{(x)}(S_u(x), x)$. This is the total derivative combining the linearisation of the forward map $S'_u(x) : X \rightarrow U$ and the partial derivative of the misfit $e^{(x)}(S_u(x), x) : U \rightarrow W^*$. Then it is Lipschitz continuous, i.e.,

$$\left\| e^{(x)}(S_u(x_2), x_2) - e^{(x)}(S_u(x_1), x_1) \right\|_{\mathbb{L}(X; W^*)} \leq L_e \|x_2 - x_1\|_X.$$

for all $x_1, x_2 \in X$, where $L_e \geq 0$ depends only on bounds from Assumption 2.4. Consequently,

$$M_e := \sup_{x \in X} \left\| e^{(x)}(S_u(x), x) \right\|_{\mathbb{L}(X; W^*)} < \infty,$$

and $x \mapsto e^{(x)}(S_u(x), x)$ is Lipschitz continuous from X to $\mathbb{L}(X; W^*)$.

Proof. By the chain rule for Fréchet derivatives, we have

$$[x \mapsto e(S_u(x), x)]'(x) = e^{(x)}(S_u(x), x) + e^{(u)}(S_u(x), x) \circ S'_u(x).$$

Since $e(S_u(x), x) \equiv 0$, the total derivative vanishes, i.e.,

$$0 = e^{(x)}(S_u(x), x) + e^{(u)}(S_u(x), x)S'_u(x).$$

Rearranging gives

$$e^{(x)}(S_u(x), x) = -e^{(u)}(S_u(x), x)S'_u(x).$$

Let $x_1, x_2 \in X$. Then

$$\begin{aligned} & e^{(x)}(S_u(x_2), x_2) - e^{(x)}(S_u(x_1), x_1) \\ &= -e^{(u)}(S_u(x_2), x_2)S'_u(x_2) + e^{(u)}(S_u(x_1), x_1)S'_u(x_1) \\ &= -[e^{(u)}(S_u(x_2), x_2)S'_u(x_2) - e^{(u)}(S_u(x_2), x_2)S'_u(x_1)] \\ &\quad - [e^{(u)}(S_u(x_2), x_2) - e^{(u)}(S_u(x_1), x_1)]S'_u(x_1). \end{aligned}$$

We have

$$\begin{aligned} \|e^{(x)}(S_u(x_2), x_2) - e^{(x)}(S_u(x_1), x_1)\|_{\mathbb{L}(X;W^*)} &\leq \|e^{(u)}(S_u(x_2), x_2)\|_{\mathbb{L}(U;W^*)} \|S'_u(x_2) - S'_u(x_1)\|_{\mathbb{L}(X;U)} \\ &\quad + \|S'_u(x_1)\|_{\mathbb{L}(X;U)} \|e^{(u)}(S_u(x_2), x_2) - e^{(u)}(S_u(x_1), x_1)\|_{\mathbb{L}(U;W^*)}. \end{aligned}$$

By [Lemma 2.5](#), we have

$$\|S_u(x_2) - S_u(x_1)\|_U \leq L_S \|x_2 - x_1\|_X,$$

and by [Lemma 2.8](#),

$$\|e^{(u)}(S_u(x_2), x_2) - e^{(u)}(S_u(x_1), x_1)\|_{\mathbb{L}(U;W^*)} \leq L_{e^{(u)}|_X} \|x_2 - x_1\|_X.$$

These estimates, together with the uniform bounds

$$\|S'_u(x)\|_{\mathbb{L}(X;U)} \leq L_S, \quad \text{and} \quad \|e^{(u)}(S_u(x), x)\|_{\mathbb{L}(U;W^*)} \leq M_u,$$

imply that $L_e \leq L_S L_{e^{(u)}|_X}$, where the first term vanishes because S'_u is linear (constant operator norm), and the second follows from Lipschitz continuity of $e^{(u)}$. Boundedness of X ensures all constants are uniform. \square

2.4 THE REDUCED ADJOINT EQUATION

In this subsection, we introduce the adjoint equation and its reduced form, which are essential for differentiating the data term ([Section 2.5](#)). We also present the corresponding Lipschitz properties of the adjoint operator, which will be used to derive Lipschitz estimates for the reduced gradient.

We recall that

$$F(x) := J(S_u(x), x),$$

where, following [\(2.4\)](#), $S_u : X \rightarrow U$ arises from the satisfaction of

$$(2.9) \quad e(S_u(x), x) = 0.$$

We will now construct Lipschitz estimates for F' .

By Lemma 2.6, $e(\cdot, x)$ is Fréchet differentiable in u . Since $e^{(u)}(S_u(x), x) : U \rightarrow W^*$ is boundedly invertible by Proposition 2.12, the implicit function theorem (see [34, Thm. 4.3]) yields

$$S'_u(x) = -e^{(u)}(S_u(x), x)^{-1} e^{(x)}(S_u(x), x).$$

Thus, $S_u : X \rightarrow U$ is Fréchet differentiable at x . The total derivative of F with respect to the parameter triple $x = (\mu, k, c)$ is now obtained from the chain rule

$$(2.10) \quad F'(x) = J^{(x)}(S_u(x), x) + J^{(u)}(S_u(x), x)S'_u(x),$$

and

$$\begin{aligned} J^{(x)}(u, x) &= (J^{(\mu)}(u, x), J^{(k)}(u, x), J^{(c)}(u, x)) \quad \text{and} \\ S'_u(x) &= (S_u^{(\mu)}(x), S_u^{(k)}(x), S_u^{(c)}(x)). \end{aligned}$$

Hence,

$$\begin{aligned} F'(x) &= (J^{(u)}(u, x)S_u^{(\mu)}(x) + J^{(\mu)}(u, x)) \\ &\quad + (J^{(u)}(u, x)S_u^{(k)}(x) + J^{(k)}(u, x)) \\ &\quad + (J^{(u)}(u, x)S_u^{(c)}(x) + J^{(c)}(u, x)). \end{aligned}$$

Differentiating (2.9) with respect to x yields

$$(2.11) \quad e^{(u)}(S_u(x), x)S'_u(x) + e^{(x)}(S_u(x), x) = 0.$$

Formally, this gives

$$S'_u(x) = -e^{(u)}(S_u(x), x)^{-1} e^{(x)}(S_u(x), x),$$

where

$$e^{(x)}(S_u(x), x) = (e^{(\mu)}(u, x), e^{(k)}(u, x), e^{(c)}(u, x)) \in \mathbb{L}(X, W^*).$$

Lemma 2.13 will show that there exists a unique adjoint variable $w_x \in U$ such that

$$(2.12) \quad w_x e^{(u)}(S_u(x), x) = J^{(u)}(S_u(x), x) \quad \text{in } W^*,$$

equivalently

$$e^{(u)}(S_u(x), x)^* w_x = J^{(u)}(S_u(x), x) \quad \text{in } U^*.$$

Applying w_x to both sides of (2.11) and using (2.12) yields

$$(2.13) \quad J^{(u)}(S_u(x), x)S'_u(x) + w_x e^{(x)}(S_u(x), x) = 0.$$

Using this expression in (2.10) establishes

$$F'(x) = J^{(x)}(S_u(x), x) - w_x e^{(x)}(S_u(x), x).$$

We call (2.12) the *reduced adjoint equation*, as compared to solving the operator $S'_u(x)$ from the *basic adjoint equation* (2.11), the dimension of the unknown w_x is reduced in discretisations.

Remark 2.10 (Classical adjoint equation). The abstract adjoint $w_x \in U$ from Lemma 2.13 solves the classical backward equation

$$(2.14) \quad \begin{aligned} -\partial_t w_x - \nabla \cdot (k \nabla w_x) - \nabla \cdot (c w_x) &= \nabla_u J(S_u(x), x) && \text{in } \Omega_T, \\ k \frac{\partial w_x}{\partial n} + (c \cdot n) w_x &= 0 && \text{on } \Sigma_1, \\ w_x &= 0 && \text{on } \Sigma_2, \\ w_x(\cdot, T) &= 0 && \text{in } \Omega. \end{aligned}$$

This follows by testing (2.12) against $v \in U$.

Remark 2.11. The reduced gradient (2.12) requires computing $w_x e^{(x)}(S_u(x), x)$ where $w_x \in U$ solves the adjoint equation (2.14). For perturbations $(\eta, h, d) \in X$, this action takes the explicit vector form:

$$w_x e^{(x)}(S_u(x), x)(\eta, h, d) = \begin{pmatrix} -\langle \bar{w}_x, \eta \rangle \\ \int_{\Omega_T} h \nabla u \cdot \nabla w_x \, d\xi \, dt \\ \int_{\Omega_T} (d \cdot \nabla u) w_x \, d\xi \, dt \end{pmatrix},$$

where

$$\bar{w}_x(\xi) := \int_I w_x(\xi, t) \, dt,$$

and the measure term is

$$- \int_{\Omega_T} w_x \, d(\eta \otimes t) = -\langle \bar{w}_x, \eta \rangle.$$

Section 4 shows that these terms arise naturally from the adjoint solve, enabling efficient gradient computation via matrix-free operators. Further details are available in our software implementation [16].

Proposition 2.12 (Uniform adjoint invertibility). *Suppose Assumption 2.4 is satisfied and let e be defined by (2.4a). For $x \in X$, set*

$$e_u(x) := e^{(u)}(S_u(x), x) : U \rightarrow W^*.$$

Then, the adjoint operator

$$e_u(x)^* : W \rightarrow U$$

is uniformly boundedly invertible. More precisely, there exists a constant $C_A > 0$, independent of $x \in X$, such that

$$\|[e_u(x)^*]^{-1}\|_{\mathbb{L}(W;U)} \leq C_A.$$

Proof. Let $E(x) : W_0^{1,p}(\Omega) \rightarrow (W_0^{1,p}(\Omega))^*$ be defined weakly by

$$\langle E(x)v | w \rangle := \langle k \nabla v, \nabla w \rangle + \langle c \cdot \nabla v | w \rangle, \quad \text{for all } v, w \in W_0^{1,p}(\Omega),$$

where $k \geq k_{\min} > 0$ and $c \in L^\infty(\Omega; \mathbb{R}^d)$. Then $e_u(x)[v] = \int_{\Omega_T} (-\partial_t v \varphi + k \nabla v \cdot \nabla \varphi + (c \cdot \nabla v) \varphi) \, d\xi \, dt$ as derived in (2.7a), is the weak form of $\partial_t v + E(x)v = 0$. The operator $E(x)$ is sectorial on $L^p(\Omega)$ under uniform ellipticity and suitable assumptions on the coefficients (see Remark 2.3). Hence, $\partial_t + E(x)$ admits maximal L^p -regularity (see [24, Thm. 4.3.1]). That is,

$$v, \partial_t v, E(x)v \in L^p(\Omega_T) \quad \text{for all } f \in L^p(\Omega_T).$$

and

$$\|v\|_{L^p(\Omega_T)} + \|\partial_t v\|_{L^p(\Omega_T)} + \|E(x)v\|_{L^p(\Omega_T)} \leq C \|f\|_{L^p(\Omega_T)}.$$

The adjoint $e_u(x)^*$ solves the backward equation $-\partial_t w_x + E(x)^* w_x = f$. Since $E(x)$ is sectorial on $L^p(\Omega)$, its adjoint $E(x)^*$ is also sectorial on $L^{p'}(\Omega)$ (see [27, Chap. 1–2]), so by maximal regularity (see [24, Thm. 4.3.1]) and Remark 2.3) gives the uniform estimates

$$\|w_x\|_U + \|\partial_t w_x\|_U + \|E(x)^* w_x\|_U \leq C \|f\|_U,$$

with C independent of $x \in X$ (depends on $k_{\min}, |c|_{L^\infty}, \Omega$). Thus, for $e_u(x)^* w_x = f$,

$$\|w_x\|_U \leq C \|f\|_U \implies \|[e_u(x)^*]^{-1}\|_{\mathbb{L}(W;U)} \leq C = C_A. \quad \square$$

Lemma 2.13 (Solvability of the reduced adjoint equation). *Suppose Assumption 2.4 holds and let $e(u(x), x)$ with $u := S_u(x)$ be defined as in (2.4a). Assume the data fidelity $J : U \times X \rightarrow \mathbb{R}$ is (partially) Fréchet differentiable w.r.t. u with*

$$J^{(u)}(u, x) \in U^* \quad \text{for all } (u, x) \in U \times X.$$

Then for every $x \in X$, there exists a unique adjoint state $w_x \in U$ solving the reduced adjoint equation

$$(2.15) \quad w_x e^{(u)}(S_u(x), x) = J^{(u)}(S_u(x), x) \quad \text{in } W^*.$$

Proof. By Lemma 2.6, $e(\cdot, x)$ is Fréchet differentiable in u at $u := S_u(x)$, so $e^{(u)}(S_u(x), x) : U \rightarrow W^*$ is bounded linear. By Proposition 2.12, its adjoint

$$T(x) := e^{(u)}(S_u(x), x)^* : W \rightarrow U^*$$

is uniformly boundedly invertible, hence $T(x)^{-1} : U^* \rightarrow W$ exists and is bounded.

Let

$$w_x := T(x)^{-1} J^{(u)}(S_u(x), x) \in U.$$

By definition of $T(x)$, we obtain the dual form

$$T(x)w_x = e^{(u)}(S_u(x), x)^* w_x = J^{(u)}(S_u(x), x) \quad \text{in } U^*.$$

The primal form follows immediately by definition of adjoint operator, testing with $v \in U$, we get

$$\langle e^{(u)}(S_u(x), x)^* w_x, v \rangle_{U^*, U} = \langle J^{(u)}(S_u(x), x), v \rangle_{U^*, U} \quad \text{for all } v \in U.$$

By Proposition 2.12, $T(x)$ is uniformly boundedly invertible, i.e., there exists $C_A > 0$, independent of $x \in X$, such that

$$\|T(x)^{-1}v\|_W \leq C_A \|v\|_{U^*} \quad \text{for all } v \in U^*.$$

For uniqueness, suppose $T(x)w_1 = T(x)w_2$. Then $T(x)(w_1 - w_2) = 0$, so

$$w_1 - w_2 = T(x)^{-1}(T(x)(w_1 - w_2)) = 0.$$

Thus $T(x)$ is injective. It follows that w_x is unique in U . □

Example 2.14 (Quadratic data fidelity terms). A standard choice of quadratic data fidelity satisfying Lemma 2.13:

$$J(u, x) = \frac{1}{2} \|Au - z_d\|_Y^2, \quad A \in \mathbb{L}(U; Y), \quad z_d \in Y,$$

where Y is a Hilbert observation space (e.g., $L^2(\Omega_T)$). The Fréchet derivative with respect to u is given by

$$J^{(u)}(u, x) = A^*(Au - z_d) \in U^*.$$

Moreover, it satisfies the bound

$$\|J^{(u)}\|_{U^*} \leq \|A^*\|_{\mathbb{L}(Y; U^*)} (\|A\|_{\mathbb{L}(U; Y)} \|u\|_U + \|z_d\|_Y) \leq M_J < \infty.$$

In addition, the derivative is Lipschitz continuous with respect to u :

$$\|J^{(u)}(u_2, x_2) - J^{(u)}(u_1, x_1)\|_{U^*} \leq \|A^*\|_{\mathbb{L}(Y; U^*)} \|A\|_{\mathbb{L}(U; Y)} \|u_2 - u_1\|_U.$$

This quadratic data fidelity satisfies all assumptions of Lemma 2.13 (Fréchet differentiability $J^{(u)} \in U^*$ and boundedness).

Lemma 2.15 (Lipschitz continuity of the adjoint operator). *Assume that Assumption 2.4 holds and that e is defined by (2.4a). Let*

$$e_u(x) := e^{(u)}(S_u(x), x) : U \rightarrow W^*.$$

Then there exists a constant $L_A \geq 0$ such that

$$\|e_u(x_2)^* - e_u(x_1)^*\|_{\mathbb{L}(W;U)} \leq L_A \|x_2 - x_1\|_X \quad \text{for all } x_1, x_2 \in X.$$

Proof. By definition,

$$e_u(x_2)^* - e_u(x_1)^* = e^{(u)}(S_u(x_2), x_2)^* - e^{(u)}(S_u(x_1), x_1)^*.$$

Since the adjoint map is isometric, we have $\|T^*\|_{\mathbb{L}(W;U)} = \|T\|_{\mathbb{L}(U;W^*)}$. Hence,

$$\|e_u(x_2)^* - e_u(x_1)^*\|_{\mathbb{L}(W;U)} = \|e^{(u)}(S_u(x_2), x_2) - e^{(u)}(S_u(x_1), x_1)\|_{\mathbb{L}(U;W^*)}.$$

By Lemma 2.8,

$$\|e^{(u)}(S_u(x_2), x_2) - e^{(u)}(S_u(x_1), x_1)\|_{\mathbb{L}(U;W^*)} \leq L_{e^{(u)}|_U} \|S_u(x_2) - S_u(x_1)\|_U + L_{e^{(u)}|_X} \|x_2 - x_1\|_X.$$

By Lemma 2.5,

$$\|S_u(x_2) - S_u(x_1)\|_U \leq L_S \|x_2 - x_1\|_X.$$

Thus,

$$\|e_u(x_2)^* - e_u(x_1)^*\|_{\mathbb{L}(W;U)} \leq L_{e^{(u)}|_U} L_S \|x_2 - x_1\|_X + L_{e^{(u)}|_X} \|x_2 - x_1\|_X = (L_{e^{(u)}|_U} L_S + L_{e^{(u)}|_X}) \|x_2 - x_1\|_X.$$

Setting $L_A := L_{e^{(u)}|_U} L_S + L_{e^{(u)}|_X}$ completes the proof. \square

2.5 THE DIFFERENTIAL OF THE DATA TERM

Recalling the expression (2.10), we now proceed with the Lipschitz estimates for F' . We require:

Assumption 2.16 (Data fidelity regularity). The data fidelity $J : U \times X \rightarrow \mathbb{R}$ is Fréchet differentiable with $J^{(u)} : U \times X \rightarrow U^*$ and $J^{(x)} : U \times X \rightarrow X^*$ Lipschitz continuous:

$$\begin{aligned} \|J^{(u)}(u_2, x_2) - J^{(u)}(u_1, x_1)\|_{U^*} &\leq L_{J^{(u)}|_U} \|u_2 - u_1\|_U + L_{J^{(u)}|_X} \|x_2 - x_1\|_X, \quad \text{and} \\ \|J^{(x)}(u_2, x_2) - J^{(x)}(u, x)\|_{X^*} &\leq L_{J^{(x)}|_U} \|u_2 - u\|_U + L_{J^{(x)}|_X} \|x_2 - x\|_X. \end{aligned}$$

for all $u_1, u_2 \in U$ and $x_1, x_2 \in X$ for some $L_{J^{(u)}|_U}, L_{J^{(u)}|_X}, L_{J^{(x)}|_U}, L_{J^{(x)}|_X} \geq 0$.

We first estimate the first factor of the term $J^{(u)}(S_u(x), x) S'_u(x)$ in (2.10).

Lemma 2.17 (Lipschitz continuity of the Jacobian). *Suppose Assumptions 2.4 and 2.16 hold. Then, the mapping $x \mapsto J^{(u)}(S_u(x), x) : X \rightarrow U^*$ is Lipschitz continuous, i.e.,*

$$\|J^{(u)}(S_u(x_2), x_2) - J^{(u)}(S_u(x_1), x_1)\|_{U^*} \leq L_{J^{(u)} \circ S} \|x_2 - x_1\|_X, \quad \text{for all } x_1, x_2 \in X,$$

where $L_{J^{(u)} \circ S} = L_{J^{(u)}|_U} L_S + L_{J^{(u)}|_X} > 0$ with L_S from Lemma 2.5.

Proof. Immediate consequence of Assumption 2.16 and Lemma 2.5. \square

Lemma 2.18 (Boundedness of the differential of the data fidelity). *Suppose Assumptions 2.4 and 2.16 hold. Then there exists a constant $M_J < \infty$ such that*

$$\|J^{(u)}(S_u(x), x)\|_{U^*} \leq M_J \quad \text{for all } x \in X.$$

Proof. By [Assumption 2.16](#), $J^{(u)} : U \times X \rightarrow U^*$ is Lipschitz continuous:

$$\|J^{(u)}(u_2, x_2) - J^{(u)}(u_1, x_1)\|_{U^*} \leq L_{J^{(u)}|_u} \|u_2 - u_1\|_U + L_{J^{(u)}|_x} \|x_2 - x_1\|_X.$$

By maximal L^p -regularity ([\[24, Thm. 4.3.1\]](#)) and [Assumption 2.4](#), the solution operator $S_u : X \rightarrow U$ is uniformly bounded, i.e.,

$$\|S_u(x)\|_U \leq C_S \quad \text{for all } x \in X.$$

Fix a reference point $(\hat{u}, \hat{x}) \in U \times X$. Then

$$\begin{aligned} \|J^{(u)}(S_u(x), x)\|_{U^*} &\leq \|J^{(u)}(S_u(x), x) - J^{(u)}(\hat{u}, \hat{x})\|_{U^*} + \|J^{(u)}(\hat{u}, \hat{x})\|_{U^*} \\ &\leq L_{J^{(u)}|_u} \|S_u(x) - \hat{u}\|_U + L_{J^{(u)}|_x} \|x - \hat{x}\|_X + \|J^{(u)}(\hat{u}, \hat{x})\|_{U^*}. \end{aligned}$$

Since S_u is uniformly bounded, $S_u(x)$ stays in a bounded set in U . For bounded $x \in X$, we obtain

$$\|J^{(u)}(S_u(x), x)\|_{U^*} \leq M_J := L_{J^{(u)}|_u} (C_S + \|\hat{u}\|_U) + L_{J^{(u)}|_x} \text{diam}(X) + \|J^{(u)}(\hat{u}, \hat{x})\|_{U^*} < \infty. \quad \square$$

Finally, we are able to obtain a Lipschitz estimate for the remaining second term of [\(2.10\)](#).

Theorem 2.19 (Partial Lipschitz estimate). *Suppose [Assumptions 2.4](#) and [2.16](#) hold. Then the mapping $x \mapsto J^{(u)}(S_u(x), x)S'(x)$ is Lipschitz continuous from X to $\mathbb{L}(X; U^*)$: there exists $C_{\text{tot}}^{\text{Lip}} > 0$ such that*

$$\|J^{(u)}(S_u(x_1), x_1)S'(x_1) - J^{(u)}(S_u(x_2), x_2)S'(x_2)\|_{\mathbb{L}(X; U^*)} \leq C_{\text{tot}}^{\text{Lip}} \|x_2 - x_1\|_X \quad \text{for all } x_1, x_2 \in X,$$

where

$$(2.16) \quad C_{\text{tot}}^{\text{Lip}} = C_A^2 (L_{e^{(u)}|_u} L_S + L_{e^{(u)}|_x}) M_J + C_A (L_{J^{(u)}|_u} L_S + L_{J^{(u)}|_x}) L_S M_e + C_A M_J L_{e^{(u)}|_x}$$

for constants $L_{e^{(u)}|_u}$, $L_{e^{(u)}|_x}$ ([Lemma 2.8](#)), M_e ([Lemma 2.9](#)), C_A ([Proposition 2.12](#)), $L_{J^{(u)}|_u}$, $L_{J^{(u)}|_x}$ ([Assumption 2.16](#)), L_S ([Lemma 2.17](#)), and M_J ([Lemma 2.18](#)).

Proof. Let w_x solve the reduced adjoint equation [\(2.12\)](#), i.e.,

$$w_x = [e_u(x)^*]^{-1} J^{(u)}(S_u(x)), \quad \text{where } e_u(x) := e^{(u)}(S_u(x), x) : U \rightarrow W^*.$$

Then $w_x e^{(x)}(S_u(x), x) = J^{(u)}(S_u(x), x)S'(x)$ by [\(2.13\)](#), so the claim amounts to showing the Lipschitz continuity of $x \mapsto w_x e^{(x)}(S_u(x), x)$.

Denote $e_x(\tilde{x}) := e^{(x)}(S_u(\tilde{x}), \tilde{x})$. We decompose

$$\Delta := w_{x_2} e_x(x_2) - w_{x_1} e_x(x_1) = (w_{x_2} - w_{x_1}) e_x(x_2) + w_{x_1} (e_x(x_2) - e_x(x_1)).$$

By [Proposition 2.12](#),

$$\|w_{x_i}\|_U \leq C_A \|J^{(u)}(S_u(x_i))\|_{U^*} \leq C_A M_J.$$

[Lemma 2.8](#) gives $\|e_x(x_i)\|_{\mathbb{L}(X; W^*)} \leq M_e$ and [Lemma 2.9](#) gives

$$\|e_x(x_2) - e_x(x_1)\|_{\mathbb{L}(X; W^*)} \leq L_e \|x_2 - x_1\|_X.$$

For the difference, abbreviate $T_i := e_u(x_i)^* : W \rightarrow U^*$. Then

$$(2.17) \quad w_{x_2} - w_{x_1} = T_2^{-1} [J^{(u)}(S_u(x_2)) - J^{(u)}(S_u(x_1))] + (T_2^{-1} - T_1^{-1}) J^{(u)}(S_u(x_1)).$$

By [Lemma 2.17](#), we have

$$(2.18) \quad \|T_2^{-1} [J^{(u)}(S_u(x_2)) - J^{(u)}(S_u(x_1))] \|_U \leq C_A L_{J^{(u)} \circ S} \|x_2 - x_1\|_X.$$

For the resolvent difference, since $u_i = S_u(x_i)$, (2.8) in Lemmas 2.5 and 2.8 give

$$\begin{aligned} \|T_2 - T_1\|_{\mathbb{L}(W;U^*)} &= \|e_u(S_u(x_2), x_2)^* - e_u(S_u(x_1), x_1)^*\| \\ &\leq L_{e^{(u)}|u} \|S_u(x_2) - S_u(x_1)\|_U + L_{e^{(u)}|x} \|x_2 - x_1\|_X \leq L_e \|x_2 - x_1\|_X \end{aligned}$$

for $L_e := L_{e^{(u)}|u} L_S + L_{e^{(u)}|x}$. Since $\|T_i^{-1}\| \leq C_A$, the Neumann series yields

$$\|T_2^{-1} - T_1^{-1}\|_{\mathbb{L}(U^*;U)} = \|T_2^{-1}(T_1 - T_2)T_1^{-1}\|_{\mathbb{L}(U^*;U)} \leq C_A \cdot L_e \cdot C_A \|x_2 - x_1\|_X = C_A^2 L_e \|x_2 - x_1\|_X.$$

Thus,

$$(2.19) \quad \|(T_2^{-1} - T_1^{-1})J^{(u)}(S_u(x_1))\|_U \leq C_A^2 L_e M_J \|x_2 - x_1\|_X.$$

Substituting (2.18) and (2.19) into (2.17), we obtain

$$(2.20) \quad \|w_{x_2} - w_{x_1}\|_U \leq (C_A L_{J^{(u)} \circ S} + C_A^2 L_e M_J) \|x_2 - x_1\|_X.$$

Finally, combining (2.20) with Lemmas 2.9 and 2.17, and using $L_e = L_{e^{(u)}|u} L_S + L_{e^{(u)}|x}$ and $L_{J^{(u)} \circ S} = L_{J^{(u)}|u} L_S + L_{J^{(u)}|x}$ (Lemma 2.17), we obtain for $C_{\text{tot}}^{\text{Lip}}$ as stated that

$$\begin{aligned} \|\Delta\|_{\mathbb{L}(X;U^*)} &\leq \|w_{x_2} - w_{x_1}\|_U \|e_x(x_2)\|_{\mathbb{L}(X;W^*)} + \|w_{x_1}\|_U \|e_x(x_2) - e_x(x_1)\|_{\mathbb{L}(X;W^*)} \\ &\leq (C_A L_{J^{(u)} \circ S} + C_A^2 L_e M_J) M_e \|x_2 - x_1\|_X + C_A M_J L_e \|x_2 - x_1\|_X \\ &= C_{\text{tot}}^{\text{Lip}} \|x_2 - x_1\|_X. \quad \square \end{aligned}$$

Combining Assumption 2.16 regarding $J^{(x)}$ with Lemma 2.5 and Theorem 2.19, we immediately obtain:

Corollary 2.20 (Full Lipschitz estimate). *Suppose Assumptions 2.4 and 2.16 hold. Then $F : x \mapsto J(S_u(x), x)$ is Lipschitz-differentiable:*

$$\|F'(x_2) - F'(x_1)\|_{X^*} \leq (C_{\text{tot}}^{\text{Lip}} + L_{J^{(x)}|u} L_S + L_{J^{(x)}|x}) \|x_2 - x_1\|_X, \quad x_1, x_2 \in X,$$

where $C_{\text{tot}}^{\text{Lip}}$ is defined in (2.16), $L_{J^{(x)}|u}, L_{J^{(x)}|x}$ arise from Assumption 2.16, and L_S from Lemma 2.17.

Proof. Recalling (2.10), we have

$$(2.21) \quad F'(x) = J^{(u)}(S_u(x), x) S'_u(x) + J^{(x)}(S_u(x), x).$$

For the second term, Assumption 2.16 and Lemma 2.5 yield the following

$$\begin{aligned} \|J^{(x)}(S_u(x_2), x_2) - J^{(x)}(S_u(x_1), x_1)\|_{X^*} &\leq L_{J^{(x)}|u} \|S_u(x_2) - S_u(x_1)\|_U + L_{J^{(x)}|x} \|x_2 - x_1\|_X \\ &\leq (L_{J^{(x)}|u} L_S + L_{J^{(x)}|x}) \|x_2 - x_1\|_X. \end{aligned}$$

Thus, using Theorem 2.19 for the first term of (2.21), establishes the claim. \square

3 OPTIMISATION WITH MEASURES

Now that we have developed our data term, and its derivatives, it is time to solve the inverse problem. We recall that we expect there to be a finite number of leaks, modelled by the measure μ . To introduce this a priori information into our problem, we apply Radon-norm regularisation, considering the problem (2.1), which has also the auxiliary unknowns k and c : the diffusion and the convection fields. Considering reasonably short time intervals, we take the latter to be constant in space and time.

We use the specific choices (2.2) of the data fidelity and regularisation term. Since the data term $F = J \circ S_u$ is nonconvex, the existing theory of most conditional gradient methods does not extend to it. In Corollary 2.20 we have, however, proved that F is Lipschitz-differentiable, which is enough to apply forward-backward type methods.

3.1 BASIC ALGORITHM

We apply the forward-backward methods of [30, 31] with a Radon-norm-squared proximal penalty. Recall that we write $x = (\mu, k, c)$, and similarly $x^n = (\mu^n, k^n, c^n)$. For the basic non-sliding variant, an exact update for the measure μ^{n+1} at iteration number $n \in \mathbb{N}$ would be determined by the surrogate problem

$$(3.1a) \quad \mu^{n+1} \in \arg \min_{\mu \in \mathcal{M}(\Omega)} F^{(\mu)}(x^n) + \langle F^{(\mu)}(x^n), \mu - \mu^n \rangle + \alpha \|\mu\|_{\mathcal{M}(\Omega)} + \delta_{\geq 0}(\mu) + \frac{1}{2\tau} \|\mu - \mu^n\|_{\mathcal{M}(\Omega)}^2.$$

The auxiliary variables (convection and diffusion), would be updated by the standard Hilbert-space forward-backward step

$$(3.1b) \quad (k^{n+1}, c^{n+1}) = \text{prox}_{\sigma R_{k,c}}((k^n, c^n) - \sigma \nabla_{k,c} F(x^n)).$$

Here $\tau, \sigma > 0$ are step length parameters that have to satisfy

$$\tau L_\mu < 1 \quad \text{and} \quad \sigma L_{k,c} < 1,$$

where L_μ and $L_{k,c}$ are the Lipschitz factors of F' with respect to the indicated parameters.¹ In [Corollary 2.20](#) we have derived a joint factor $L_\mu = L_{k,c}$.

In practise, following [31], inexact steps are taken for (3.1a) by solving for some iteration-dependent tolerances $\varepsilon^{n+1} > 0$ satisfying $\sum_{n=0}^{\infty} \varepsilon^{n+1} < \infty$, the inexact optimality condition

$$(3.2a) \quad -\varepsilon^{n+1} \leq v^n + w^{n+1} + \|\mu^{n+1} - \mu^n\|_{\mathcal{M}(\Omega)} \omega^{n+1} \leq \varepsilon^{n+1},$$

where $v^n := F^{(\mu)}(x^n)$ and $\omega^{n+1}, w^{n+1} \in C_0(\Omega)$ satisfy

$$(3.2b) \quad -1 \leq \omega^{n+1} \leq 1, \quad \langle \omega^{n+1}, \mu^{n+1} - \mu^n \rangle = \|\mu^{n+1} - \mu^n\|_{\mathcal{M}(\Omega)}, \quad w^{n+1} \leq \alpha, \quad \langle w^{n+1}, \mu^{n+1} \rangle = \|\mu^{n+1}\|_{\mathcal{M}(\Omega)}.$$

To solve these conditions, one finds an (approximate) minimiser $\xi \in \Omega$ of $v^n \in C_0(\Omega)$, and writing $\mu^n = \sum_{i=1}^{m_n} \beta_{n,i} \delta_{\xi_{n,i}}$, forms $\mu^{n+1} = \sum_{i=1}^{m_n} \beta_{n+1,i} \delta_{\xi_{n,i}} + \beta_{n+1,m_n+1} \delta_\xi$ by solving a finite-dimensional optimisation for the escape rates $\{\beta_{n+1,i}\}_{i=1}^{m_n+1}$. We refer to [31] for further details.

3.2 SLIDING ALGORITHM

Our second algorithm applies ideas from optimal transport to first transport the Dirac masses of μ^n to new locations using a transport plan $\gamma^{n+1} = \sum_{i=1}^{m_n} \eta_{n,i} \delta_{(x_{n,i}, y_{n,i})} \in \mathcal{M}_+(\Omega^2)$, where $\eta_{n,i}$ is amount of mass transported from $x_{n,i}$ to the predicted new location $y_{n,i}$. The transport or ‘‘sliding’’ of the locations is based on doing a gradient step on the locations with v^n . Then the method applies the same insertion and weight optimisation steps as above at the transported measure. These steps may have to be repeated several times to satisfy technical curvature and remainder conditions from [30]. As we will see in the numerical experiments, the transport or ‘‘source sliding’’ can significantly improve the performance of the algorithm.

To be precise, define the projections $\pi^i(\xi) = \xi_i$, as well as the push-forward measures $\pi_\#^i \mu$ satisfying $\pi_\#^i \mu(A) = \mu(\{\xi \in \Omega \mid \pi^i(\xi) \in A\})$. A ‘‘sliding’’ version of (3.1) first replaces μ^n by $\check{\mu}^n = \mu^n + (\pi_\#^1 - \pi_\#^0) \gamma^{n+1}$ for a transport plan $\gamma^{n+1} \in \mathcal{M}_+(\Omega^2)$. Then it applies (3.1) with μ^n replaced by $\check{\mu}^n$ and x^n replaced by $\check{x}^n = (\check{\mu}^n, k^n, c^n)$. That is, writing

$$v^n = F^{(\mu)}(x^n), \quad \text{and} \quad \check{v}^n = F^{(\mu)}(\check{x}^n),$$

¹The Lipschitz factor $L_{k,c}$ is denoted L_z in [31].

for some sliding parameter $\theta > 0$, we first construct $\gamma^{n+1} \in \mathcal{M}_+(\Omega^2)$ that satisfies

$$(3.3a) \quad \zeta = \xi - \theta \tau \nabla v^n(\xi) \quad \text{for } \gamma^{n+1}\text{-a.e. } (\xi, \zeta).$$

Then we approximately, as in (3.2), update

$$(3.3b) \quad \mu^{n+1} \in \arg \min_{\mu \in \mathcal{M}(\Omega)} F^{(\mu)}(\check{x}^n) + \langle F^{(\mu)}(\check{x}^n), \mu - \check{\mu}^n \rangle + \alpha \|\mu\|_{\mathcal{M}(\Omega)} + \delta_{\geq 0}(\mu) + \frac{1}{2\tau} \|\mu - \check{\mu}^n\|_{\mathcal{M}(\Omega)}^2,$$

and

$$(3.3c) \quad (k^{n+1}, c^{n+1}) := \text{prox}_{\theta R_{k,c}}((k^n, c^n) - \theta \nabla_{k,c} F(\check{x}^n)).$$

This construction may have to be repeated several times until γ^{n+1} satisfies further restrictions. We typically

1. Start with the source marginal $\pi_{\#}^0 \lambda^{n+1} = \mu^n$, constructing the target marginal $\pi_{\#}^1 \gamma^{n+1}$ using (3.3a). Then $\check{\mu}^n$ is simply μ^n transported along the negative gradient $v^n(\xi)$.
2. We reduce the mass transported by γ^{n+1} , repeating (3.3b) and (3.3c) until further conditions given in [30] are satisfied.

For details, we refer to [30].

4 ASPECTS OF NUMERICAL IMPLEMENTATION

In this section we describe aspects of numerical implementation of the convection–diffusion equation (2.3), as well as its adjoint. We first describe the implementation of the laser observation operator A in Section 4.2. We then apply the finite element method (FEM) and forward Euler steps to discretize (2.3) and obtain the corresponding discrete system (Section 4.1). Finally, we describe some details of the optimisation algorithm in Section 4.3. Further details are available in our software implementation [16].

We introduce a temporal grid $0 = t_0 < t_1 < \dots < t_{N_T} = T$ with time step lengths $\Delta_i := t_{i+1} - t_i$, for $i = 0, \dots, N_T - 1$. For simplicity of presentation, we assume that both the laser observation times and time discretisation agree with this grid.

4.1 DISCRETISATION OF THE CONVECTION–DIFFUSION EQUATION

Discretisation of the concentration (state) We write $I_i = [t_i, t_{i+1}]$, for $i = 0, \dots, N_T - 1$ with $t_0 = 0$, $t_{N_T} = T$. For space discretisation, we choose a finite element basis $\{\varphi_m\}_{m=1}^{N_\varphi}$ that forms $\mathbb{L}^2 \subset L^2(\Omega)$. Then we define the space of discretised concentrations by

$$\mathbb{U} := \left\{ u \in L^2(I \times \Omega) \mid \begin{array}{l} u(\cdot, t_i) \in \mathbb{L}^2 \text{ for all } 0 \leq i \leq N_T, \\ u(x, \cdot)|_{I_i} = \text{constant for all } x \in \Omega, \quad i = 0, \dots, N_T - 1 \end{array} \right\}.$$

Denoting the indicator function of I_i by χ_i , each $u \in \mathbb{U}$ admits for some coefficient vectors $\mathbf{u}_i = (u_1^i, \dots, u_{N_\varphi}^i) \in \mathbb{R}^{N_\varphi}$, ($i = 0, \dots, N_T - 1$) the representation

$$u(\xi, t) = \sum_{i=0}^{N_T-1} \sum_{m=1}^{N_\varphi} u_m^i \chi_i(t) \varphi_m(\xi).$$

We discretize the continuous problem (2.3) in space and time. Let $a(\cdot, \cdot)$ be the bilinear form associated with the operator $-\Delta + c \cdot \nabla$, that is

$$a(u, v; x) = \int_{\Omega} k \nabla u \cdot \nabla v + (c \cdot \nabla u) v \, d\xi.$$

First, the forward Euler scheme on each time interval $I_i = [t_i, t_{i+1}]$ yields

$$(4.1) \quad \left\langle \frac{u(\cdot, t_{i+1}) - u(\cdot, t_i)}{\Delta_i}, v \right\rangle + a(u(\cdot, t_i), v; x) = \langle \mu, v \rangle + \int_{\Gamma_1} g(t_i) v \, d\sigma \, dt, \quad u(\cdot, t_0) = u_0,$$

for all $v \in \mathbb{V}_m \subset \{v_m \in W^{1,p}(\Omega) : v_m|_{\Gamma_1} = 0\}$. Taking $u, v \in \mathbb{U}$, we can write this in terms of coefficients as

$$\sum_{m,r=1}^{N_\varphi} \left(\frac{u_m^{i+1} - u_m^i}{\Delta_i} \langle \varphi_m, \varphi_r \rangle v_r^i + u_m^i a(\varphi_m, \varphi_r; x) v_r^i \right) = \sum_{r=1}^{N_\varphi} \langle \mu, \varphi_r \rangle v_r^i + \sum_{m,r=1}^{N_\varphi} g_m^i \langle \varphi_m, \varphi_r \rangle_{\Gamma_1} v_r^i \quad \text{for all } \mathbf{v}^i \in \mathbb{R}^{N_\varphi}$$

with $u_{m,0} = [u_0]_m$ for all $m = 1, \dots, N_\varphi$. That is,

$$(4.2) \quad \mathbf{M} \mathbf{u}^{i+1} = (\mathbf{M} - \Delta_i \mathbf{K}(x)) \mathbf{u}^i + \Delta_i \mathbf{f}^i \quad \text{for all } i = 0, \dots, N_T - 1,$$

where

$$\mathbf{f}^i = (\langle \mu, \varphi_r \rangle + \mathbf{M}|_{\Gamma_1} g^i)_r, \quad \mathbf{M} = (\langle \varphi_m, \varphi_r \rangle)_{m,r}, \quad \mathbf{K}(x) = (a(\varphi_m, \varphi_r; x))_{m,r} \quad \text{with } m, r \in \{1, \dots, N_\varphi\}.$$

Discretisation of the adjoint variable and equation The discrete adjoint system arises from the same space–time Galerkin discretization as the state equation and is therefore adjoint consistent. The adjoint variable $w \in U_0 := \{v \in U : v|_{\Sigma_1} = 0\}$ is discretised in the same space \mathbb{U} as the concentration. The weak adjoint equation $e_u(x)^* w_x = J^{(u)}(S_u(x), x)$. recalling the adjoint equation (2.14), write $\tilde{a}(\cdot, \cdot; x)$ for the bilinear form associated with the operator $-\Delta - c \cdot \nabla$, that is

$$\tilde{a}(w, v; x) = \int_{\Omega} k \nabla w \cdot \nabla v - (c \cdot \nabla u) v \, d\xi.$$

Similarly to (4.1), we then obtain the system

$$\left\langle \frac{w(\cdot, t_{i+1}) - w(\cdot, t_i)}{\Delta_i}, v \right\rangle + \tilde{a}(w(\cdot, t_i), v; x) = J^{(u(\cdot, t_i))}(S_u(x), x) v, \quad w(\cdot, t_{N_T}) = 0 \quad \text{for all } v \in U_0.$$

for all $i = 1, \dots, N_T$. Taking $w, v \in \mathbb{U}$, similarly to (4.2), this becomes

$$\mathbf{M} \mathbf{w}^i = (\mathbf{M} + \Delta_i \mathbf{L}(x)) \mathbf{w}^{i+1} + \Delta_i \mathbf{M} \mathbf{b}^i \quad \text{for all } i = 0, \dots, N_T - 1,$$

where the matrix and vector

$$(4.3) \quad \mathbf{L}(x) = (\tilde{a}(\varphi_m, \varphi_r; x))_{m,r} \quad \text{and} \quad \mathbf{b}^i(x) = (J^{(u^i)}(S_u(x), x))_r \quad \text{with } m, r \in \{1, \dots, N_\varphi\}.$$

We will refine the expression for $\mathbf{b}^i(x)$ in Section 4.2, after fixing J and the observation operator A within it.

Discretisation of the parameters We now consider the discretisation of the parameters $x = (\mu, k, c)$. Practically, the algorithms of Section 3 ensure that on each iteration n , the measure $\mu^n = \sum_{i=0}^{m_n} \beta_{n,i} \delta_{\xi_{n,i}}$ is a weighted sum of Dirac measures. We do not enforce the source points $\xi_{n,i}$ to lie on any specific grid: aside from the discretisation of u , and of $F^{(\mu)}(x)$, our algorithm is partially grid-free.

In our numerical experiments, the convection and diffusion field k and c are scalars, $c(x) \equiv (c_1, c_2) \in \mathbb{R}^2$ and $k(x) \equiv k_0 \in \mathbb{R}$.

4.2 LASER OBSERVATION OPERATOR

The observation model for gas emission detection is based on path-integrated measurements obtained via Laser Diode Spectroscopy (LDS). Each laser beam provides a path-averaged signal proportional to the line integral of the gas concentration along its trajectory. At each time step t_i , $i = 0, \dots, N_T - 1$, the concentration field $u(\xi, t)$ is observed through N_ℓ laser beams $\{\gamma_\ell\}_{\ell=1}^{N_\ell}$ (see Figure 1) reflected back to their origin from mirrors. This leads to the linear observation model

$$A : L^2(\Omega_T) \rightarrow \mathbb{R}^{N_\ell \times N_T}, \quad [Au]_{\ell,i} = \int_{\gamma_\ell} u(\xi, t_i) ds,$$

where $\ell = 1, \dots, N_\ell$ runs over the laser beams, and $i = 0, \dots, N_T - 1$ over the observation times $t_0, \dots, t_{N_T-1} \in [0, T)$. Thus the path-averaged LDS measurements

$$b_\ell = A\hat{u} + v,$$

for the ground truth concentration $\hat{u} \in L^2(\Omega_T)$, and (Gaussian) noise v .

For $u \in \mathbb{U}$ with the coefficient vectors $\mathbf{u}^1, \dots, \mathbf{u}^{N_T} \in \mathbb{R}^{N_\phi}$ as in Section 4.1, we get

$$[Au]_{\ell,i} = \int_{\gamma_\ell} u(\xi, t_i) ds = \sum_{m=1}^{N_\phi} u_m^i \int_{\gamma_\ell} \varphi_m(\xi) ds = \left(\int_{\gamma_\ell} \varphi_1(\xi) ds, \dots, \int_{\gamma_\ell} \varphi_{N_\phi}(\xi) ds \right) \mathbf{u}^i.$$

For each observation time index $i = 1, \dots, N_T - 1$, the ideal discretisation of $u \mapsto [Au]_{\cdot,i} \in \mathbb{R}^{N_\ell}$ is, therefore,

$$\mathbf{A} \in \mathbb{R}^{N_\ell \times N_\phi} \quad \text{with} \quad \mathbf{A}_{\ell,m} = \int_{\gamma_\ell} \varphi_m(\xi) ds.$$

Taking $J(u, x) = \frac{1}{2} \|Au - z_d\|_2^2$, we now have

$$J(u, x) = \sum_{i=0}^{N_T-1} \frac{1}{2} \|\mathbf{A}\mathbf{u}^i - z_d^i\|_2^2 \quad \text{when} \quad u \in \mathbb{U}.$$

Thus, $\mathbf{b}^i(x)$ defined in (4.3) can be written

$$\mathbf{b}^i(x) = (\mathbf{A}^\top (\mathbf{A}\mathbf{u}^i - z_d^i))_i, \quad i = 0, \dots, N_T - 1.$$

Approximate ray tracing via cell-locator quadrature Due to limitations in the documented API (Application Programming Interface) of the Fenicsx PDE solver and modelling library, we use a simplified and numerically efficient approach, instead of full ray-tracing, to approximate the line integrals. Specifically, \mathbf{A} is approximated via a quadrature-free method as follows:

1. **Beam segmentation:** Partition each beam path γ_ℓ into n_{seg} equal-length segments with midpoints $\bar{\xi}_p \in \Omega$, ($p = 0, \dots, n_{\text{seg}} - 1$).
2. **Cell location:** For each midpoint $\bar{\xi}_p$, identify the containing triangular element T_p .
3. **Uniform quadrature:** Distribute each segment length $l_p = \int_{\gamma_\ell} ds / n_{\text{seg}}$ uniformly over the degrees-of-freedom (DOFs) $h \in \mathcal{D}_p$ the cell T_p , updating $\mathbf{A}_{\ell,h} := \mathbf{A}_{\ell,h} + l_p / |\mathcal{D}_p|$.

This yields a sparse matrix \mathbf{A} whose i -th row approximates the beam integral via midpoint quadrature with uniform cell-DOF averaging. Further details are available in our software implementation [16].

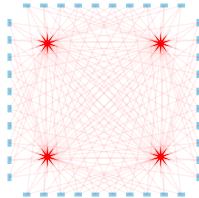


Figure 1: Illustration of the laser and mirror placements. The thick blue lines on the boundary indicate the mirrors, the red stars the laser sources, and the dimmed thin red lines the beams.

4.3 THE OPTIMISATION ALGORITHM

The implementation of the optimisation algorithm is described in [30, 31]. Our present application shares most of the codebase, with the changes related to the implementation and treatment of the forward operator, which in [30, 31] was a simple convolution. In particular, $v = F^{(\mu)}(x)$ now belongs to a P2 (second order polynomial) finite element subspace of $C_0(\Omega)$. Therefore, we do not need to use a branch-and-bound algorithm to find—what were—grid-free minimisers of this function. We can, instead, solve the minima exactly by minimising the function on each element, and taking a best elementwise minimiser as the global minimiser.

5 NUMERICAL RESULTS

We now describe our numerical experiments on leak detection. The details of the numerical implementation of the model (2.1)–(2.3), which is available in [16], are as in Section 4. We describe the specific experimental setup and algorithm parametrisation in Sections 5.1 and 5.2, and then discuss the results in Section 5.3.

5.1 EXPERIMENTAL SETUP

We divide the time interval $I = [0, 1]$ into 50 uniform segments. The domain is $\Omega = [0, 0.5]^2$ with a uniform grid of 32×32 nodes. The laser sources are at $\{(0.1, 0.1), (0.1, 0.4), (0.4, 0.1), (0.4, 0.4)\}$, and each edge of the domain has 10 uniformly spaced mirrors, as illustrated in Figure 1. Further details can be found in our numerical implementation [16]. This gives $4 \cdot 10 \cdot 4 = 160$ data points for each time step, altogether $160 \cdot 50 = 8000$ data points, to be compared with altogether $32 \cdot 32 \cdot 50 = 51200$ nodal values for the diffusion field.

We perform two experiments with different atmospheric conditions and leak sources. In experiment #1, the true convection field is $\hat{c} \equiv 0.5(\cos(30^\circ), \sin(30^\circ))$, and the true diffusion field $\hat{k} \equiv 0.01$ with the true source measure

$$\hat{\mu} = 0.08 \delta_{(0.1,0.3)} + 0.05 \delta_{(0.4,0.25)} + 0.06 \delta_{(0.25,0.13)}$$

In experiment #2, the true convection field is $\hat{c} \equiv 0.1(\cos(120^\circ), \sin(120^\circ))$, and the true diffusion field $\hat{k} \equiv 0.02$ with the true leaks modelled by the measure

$$\hat{\mu} = 0.15 \delta_{(0.2,0.3)} + 0.04 \delta_{(0.4,0.1)}.$$

(The latter leak coincides with one of the laser sources). The measurements \tilde{k} and \tilde{c}_i have, respectively, 2% and 20% Gaussian noise: the PDE is much more sensitive to the diffusion parameter k_0 , and cannot tolerate higher levels of noise. The laser line-of-sight measurements b have 1% Gaussian noise.

The boundary conditions are $u = g := 0$ on $\Gamma_1 = \{0, 1\} \times [0, 1]$ and $\partial u / \partial n = 0$ on $\Gamma_2 = (0, 1) \times \{0, 1\}$. The initial concentration is $u_0 = 0$.

For the regularisation parameter of the measure, by trial and error, we take $\alpha = 1.5 \cdot 10^{-6}$, while we use a combination of box constraints and quadratic fitting with weighting discovered by trial-and-error in

$$R_{k,c}(k, c) = \delta_{[0.001,1]}(k_0) + \frac{3}{2}(k_0 - \tilde{k})^2 + \sum_{i=1,2} \left(\delta_{[-1,1]}(c_i) + \frac{0.0005}{2}(c_i - \tilde{c}_i)^2 \right),$$

where we recall that we take the convection and diffusion fields $c \equiv (c_1, c_2) \in \mathbb{R}^2$ and $k \equiv k_0 \in \mathbb{R}$ to be constant throughout the domain. Their respective measurements are \tilde{k} , \tilde{c}_1 , and \tilde{c}_2 .

5.2 ALGORITHM PARAMETRISATION

We take as the primal step length parameter $\tau = 0.99L$, where $L = 4$ is manually fine-tuned compared to the factor estimate given by [Corollary 2.20](#). The theoretical parameter $L \approx 14.9$ gives an algorithm with comparable optimisation performance, but more superfluous sources (leaks) $x_{k,i}$ of very low release rate $\beta_{k,i} \approx 0$.

The sliding step length parameter is $\theta = 100$, likewise based on manual fine-tuning. We use a merging strategy to reduce the number of sources in the measure iterates $\mu^n = \sum_{i=1}^{m_n} \beta_{n,i} \delta_{\xi_{n,i}}$: every 10 iterations, we replace any two sources $i \neq j$ with $\|\xi_{n,j} - \xi_{n,i}\| \leq 0.1$ with $(\beta_{n,j} + \beta_{n,i}) \delta_{\xi_{n,j}}$ if this does not increase the objective function value, within a decreasing tolerance. This helps to keep the number of sources controlled, and, therefore, avoid the per-iteration computational demands growing.

As initial iterates, we take measured values $c^0 = \tilde{c}$ and $k^0 = \tilde{k}$ of the convection and diffusion parameters. A priori, we assume there to be no leaks, taking as initial iterate of the leak-modelling measure $\mu^0 = 0$.

5.3 NUMERICAL RESULTS

The reconstructed and true concentration u and measure μ for experiments #1 and #2 are in [Figures 2](#) and [3](#), respectively. The performance of both optimisation algorithms (basic and sliding) is graphed in [Figure 4](#) for both experiments. In both cases, the sliding method almost perfectly recovers the true sources, hence the true concentration, while the basic method appears to get stuck at a local stationary point. Although very report the reconstruction after 20000 iterations to ensure good recovery of the sources, in terms of function value, both algorithms are very close to the final value of the objective function already after 5000 iterations.

Overall, we can say that the sliding method performs well at leak detection with simultaneous refinement of the scalar convection (wind) parameter (measured with high 20% noise), and diffusion (affected e.g. by temperature) parameter (measured with low 2% noise). It appears that the sliding helps to avoid local stationary points in this nonconvex optimisation problem.

REFERENCES

- [1] C. Alden, S. Coburn, R. Wright, E. Baumann, K. Cossel, K. Prasad, I. Coddington, and G. Rieker, Single-blind quantification of natural gas leaks from 1 km distance using frequency combs, 2019, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=927129.
- [2] S. R. Arridge, Optical tomography in medical imaging, *Inverse Problems* 15 (1999), R41–R93, doi: [10.1088/0266-5611/15/2/001](https://doi.org/10.1088/0266-5611/15/2/001).
- [3] P. Balcombe, J. F. Speirs, N. P. Brandon, and A. D. Hawkes, Methane emissions: choosing the right climate metric and time horizon, *Environ. Sci.: Processes Impacts* 20 (2018), 1323–1339, doi: [10.1039/c8em00414e](https://doi.org/10.1039/c8em00414e), <http://dx.doi.org/10.1039/C8EM00414E>.

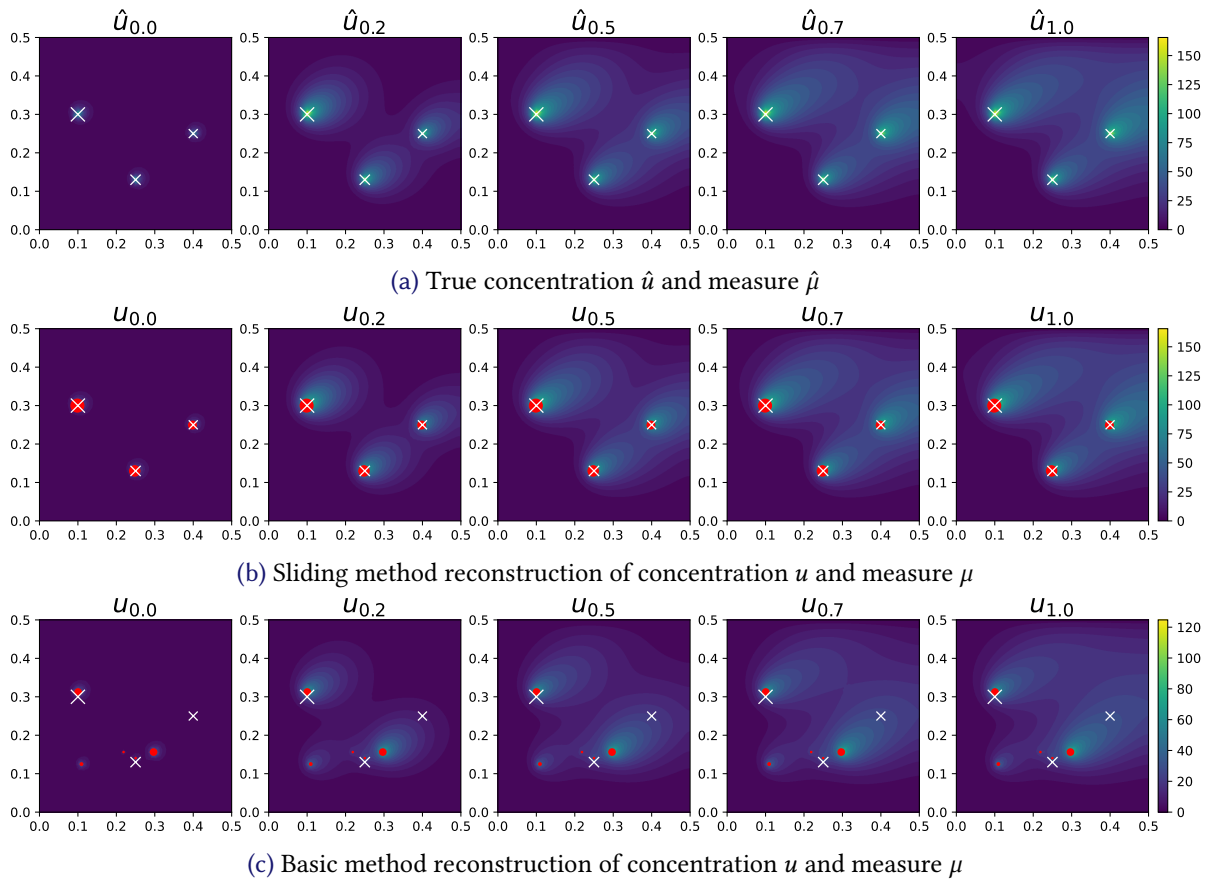


Figure 2: Reconstructions, experiment #1. Time runs from left to right, with the reconstructions shown for $t = 0.0, 0.2, 0.5, 0.7, 1.0$. The shading indicates the concentration field u . The red dots indicate the reconstructed measure, and the white crosses the ground true measure. The size of the dots corresponds to the escape rate.

- [4] D. Baldocchi, Measuring fluxes of trace gases and energy between ecosystems and the atmosphere – the state and future of the eddy covariance method, *Global Change Biol.* 20 (2014), 3600–3609, [doi:10.1111/gcb.12647](https://doi.org/10.1111/gcb.12647).
- [5] L. Blank and C. Rupprecht, An extension of the projected gradient method to a Banach space setting with application in structural topology optimization, *SIAM Journal on Control And Optimization* 55 (2017), 1481–1499, [doi:10.1137/16m1092301](https://doi.org/10.1137/16m1092301).
- [6] K. Bredies and M. Carioni, Sparsity of solutions for variational inverse problems with finite-dimensional data, *Calculus of Variations and Partial Differential Equations* 59 (2019), [doi:10.1007/s00526-019-1658-1](https://doi.org/10.1007/s00526-019-1658-1).
- [7] K. Bredies, M. Carioni, S. Fanzon, and D. Walter, Linear convergence of accelerated generalized conditional gradient methods, 2021, [arXiv:2110.06756](https://arxiv.org/abs/2110.06756).
- [8] K. Bredies and H.K. Pikkarainen, Inverse problems in spaces of measures, *ESAIM: Control, Optimization and Calculus of Variations* 19 (2013), 190–218, [doi:10.1051/cocv/2011205](https://doi.org/10.1051/cocv/2011205).
- [9] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer New York, 2011, [doi:10.1007/978-0-387-70914-7](https://doi.org/10.1007/978-0-387-70914-7).

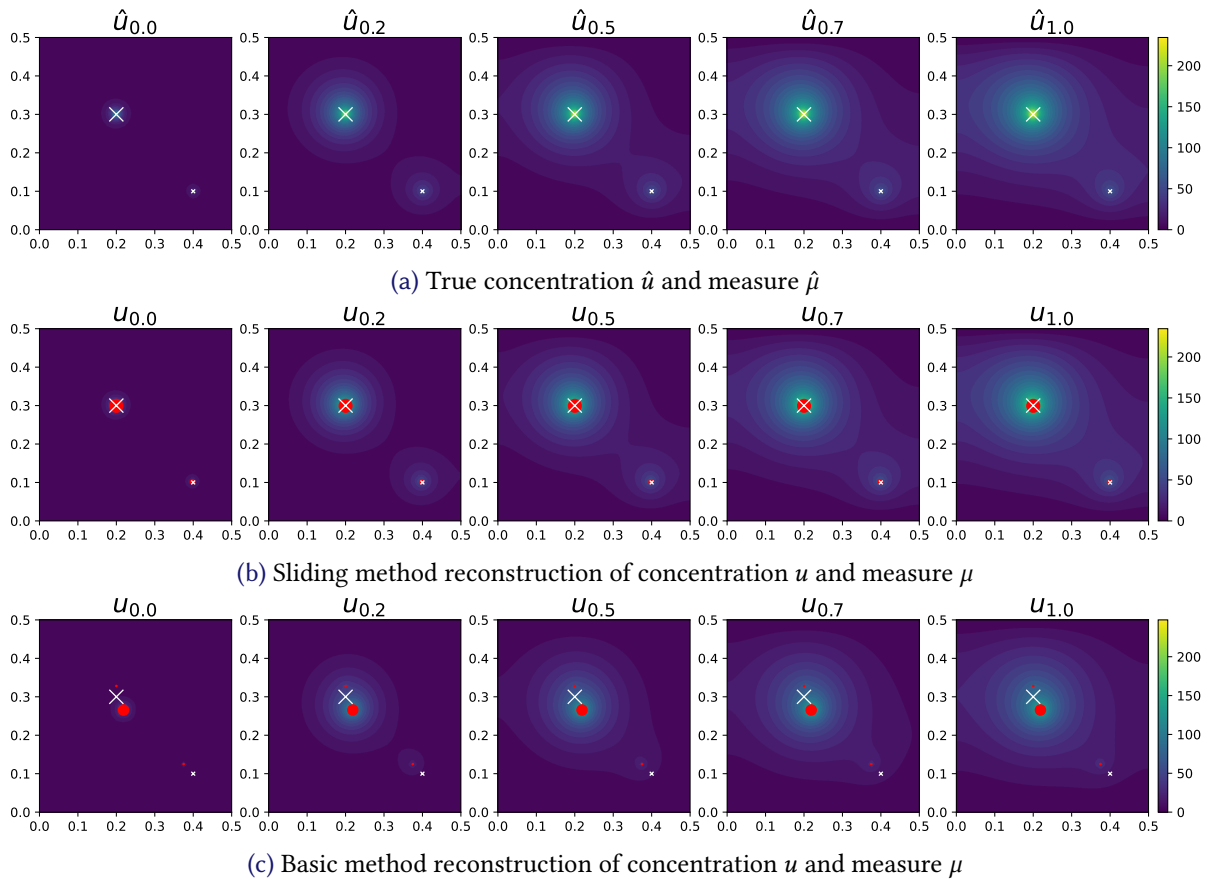


Figure 3: Reconstructions, experiment #2. Time runs from left to right, with the reconstructions shown for $t = 0.0, 0.2, 0.5, 0.7, 1.0$. The shading indicates the concentration field u . The red dots indicate the reconstructed measure, and the white crosses the ground true measure. The size of the dots corresponds to the escape rate.

- [10] E. Casas, C. Clason, and K. Kunisch, Approximation of elliptic control problems in measure spaces with sparse solutions, *SIAM Journal on Control And Optimization* 50 (2012), 1735–1752, doi:10.1137/110843216.
- [11] E. Casas, C. Clason, and K. Kunisch, Parabolic control problems in measure spaces with sparse solutions, *SIAM Journal on Optimization* 51 (2013), 28–63, doi:10.1137/120872395.
- [12] L. Chizat, Sparse optimization on measures with over-parameterized gradient descent, *Mathematical Programming* (2021), doi:10.1007/s10107-021-01636-z.
- [13] L. Chizat, Convergence rates of gradient methods for convex optimization in the space of measures, 2023, arXiv:2105.08368.
- [14] L. Chizat and F. Bach, On the global convergence of gradient descent for over-parameterized models using optimal transport, in *Advances in Neural Information Processing Systems*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (eds.), volume 31, Curran Associates, Inc., 2018, <https://proceedings.neurips.cc/paper/2018/file/a1afc58c6ca9540d057299ec3016d726-Paper.pdf>.
- [15] N. S. Daghestani, R. Brownsword, and D. Weidmann, Analysis and demonstration of atmospheric

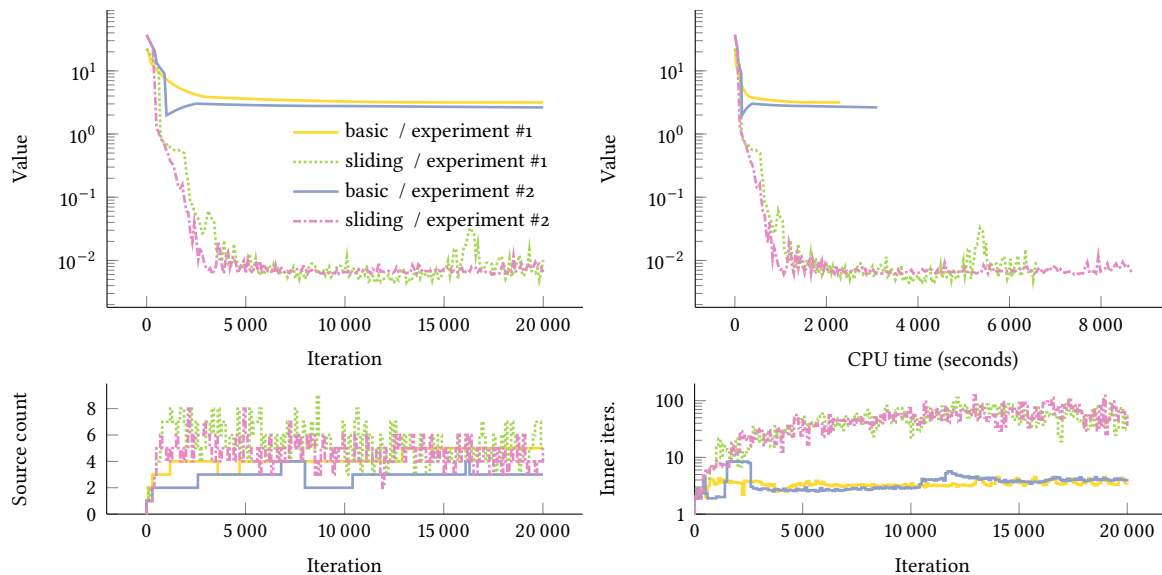


Figure 4: Algorithm performance. Top row: Function value in terms of iteration count (left) and CPU time (right). Bottom row: source count in μ^k (left), and inner weight optimisation algorithm iteration count (right) per outer iteration.

methane monitoring by mid-infrared open-path chirped laser dispersion spectroscopy., *Optics express* 22 Suppl 7 (2014), A1731–43, <https://api.semanticscholar.org/CorpusID:12682072>.

- [16] T. T. Dang and T. Valkonen, Numerical implementation of “Localisation of leaks using a measure source convection–diffusion model”, 2026, [doi:10.5281/zenodo.20090887](https://doi.org/10.5281/zenodo.20090887). Software on Zenodo.
- [17] Q. Denoyelle, V. Duval, G. Peyré, and E. Soubies, The sliding Frank–Wolfe algorithm and its application to super-resolution microscopy, *Inverse Problems* (2019), [doi:10.1088/1361-6420/ab2a29](https://doi.org/10.1088/1361-6420/ab2a29).
- [18] V. Duval and G. Peyré, Sparse regularization on thin grids I: the Lasso, *Inverse Problems* 33 (2017), 055008, [doi:10.1088/1361-6420/aa5e12](https://doi.org/10.1088/1361-6420/aa5e12).
- [19] M. Engram, K.M. Walter Anthony, T. Sachs, K. Kohnert, A. Serafimovich, G. Grosse, and F.J. Meyer, Remote sensing northern lake methane ebullition, *Nature Climate Change* 10 (2020), 511–517, [doi:10.1038/s41558-020-0762-8](https://doi.org/10.1038/s41558-020-0762-8).
- [20] A. Flinth, F. de Gournay, and P. Weiss, On the linear convergence rates of exchange and continuous methods for total variation minimization, *Mathematical Programming* 190 (2020), 221–257, [doi:10.1007/s10107-020-01530-0](https://doi.org/10.1007/s10107-020-01530-0).
- [21] B. Hirst, D. Randell, M. Jones, J. Chu, A. Kannath, N. Macleod, M. Dean, and D. Weidmann, Methane Emissions: Remote Mapping and Source Quantification Using an Open-Path Laser Dispersion Spectrometer, *Geophysical Research Letters* 47 (2020), [doi:10.1029/2019gl086725](https://doi.org/10.1029/2019gl086725).
- [22] J.P. Kaipio, V. Kolehmainen, M. Lassas, and E. Somersalo, Discretization as regularization in statistical inverse problems, *SIAM Journal on Numerical Analysis* 44 (2006), 177–195, [doi:10.1137/040616808](https://doi.org/10.1137/040616808).
- [23] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, 1988.
- [24] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Progress in Non-linear Differential Equations and Their Applications, Birkhäuser Basel, 2011.

- [25] J. McNorton, C. Wilson, M. Gloor, R.J. Parker, H. Boesch, W. Feng, R. Hossaini, and M. P. Chipperfield, Attribution of recent increases in atmospheric methane through 3-D inverse modelling, *Atmospheric Chemistry and Physics* 18 (2018), 18149–18168, doi:10.5194/acp-18-18149-2018, <https://acp.copernicus.org/articles/18/18149/2018/>.
- [26] S.M. Miller, S.C. Wofsy, A.M. Michalak, E.A. Kort, A.E. Andrews, S.C. Biraud, E.J. Dlugokencky, J. Eluszkiewicz, M.L. Fischer, G. Janssens-Maenhout, B.R. Miller, J.B. Miller, S.A. Montzka, T. Nehrkorn, and C. Sweeney, Anthropogenic emissions of methane in the United States, *Proceedings of the National Academy of Sciences* 110 (2013), 20018–20022, doi:10.1073/pnas.1314392110, <http://dx.doi.org/10.1073/pnas.1314392110>.
- [27] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, volume 44 of Applied Mathematical Sciences, Springer-Verlag, New York, 1983.
- [28] K. Pieper and D. Walter, Linear convergence of accelerated conditional gradient algorithms in spaces of measures, *ESAIM: Control, Optimization and Calculus of Variations* 27 (2021), 38, doi:10.1051/cocv/2021042, [arXiv:1904.09218](https://arxiv.org/abs/1904.09218).
- [29] C.W. Rella, T.R. Tsai, C.G. Botkin, et al., Measuring emissions from oil and natural gas well pads using the mobile flux plane technique, *Environ. Sci. Technol.* 49 (2015), 4742–4748, doi:10.1021/es5054082.
- [30] T. Valkonen, Proximal methods for point source localisation, *Journal of Nonsmooth Analysis and Optimization* 4 (2023), 10433, doi:10.46298/jnsao-2023-10433, [arXiv:2212.02991](https://arxiv.org/abs/2212.02991).
- [31] T. Valkonen, Point source localisation with unbalanced optimal transport, 2025, [arXiv:2502.12417](https://arxiv.org/abs/2502.12417), submitted.
- [32] A. Voss, E. Vänskä, D. Weidmann, A. Pulkkinen, and A. Seppänen, Multi-open-path laser dispersion spectroscopy combined with Bayesian state estimation for localizing and quantifying methane emissions, *Atmospheric Environment: X* 22 (2024), 100260, doi:https://doi.org/10.1016/j.aea0a.2024.100260, <https://www.sciencedirect.com/science/article/pii/S2590162124000273>.
- [33] D. Weidmann, B. Hirst, M. Jones, R. Ijzermans, D. Randell, N. Macleod, A. Kannath, J. Chu, and M. Dean, Locating and Quantifying Methane Emissions by Inverse Analysis of Path-Integrated Concentration Data Using a Markov-Chain Monte Carlo Approach, *ACS Earth and Space Chemistry* 6 (2022), 2190–2198, doi:10.1021/acsearthspacechem.2c00093.
- [34] E. Zeidler, *Nonlinear Functional Analysis and Its Applications*, number v. 1 in Nonlinear functional analysis and its applications, Springer, 1986.