

SMAI-JCM
SMAI JOURNAL OF
COMPUTATIONAL MATHEMATICS

Combining reciprocity gap method
and state estimator for source
identification in an advection
diffusion equation

ANTOINE TONNOIR

Volume 11 (2025), p. 233-260.

<https://doi.org/10.5802/smai-jcm.124>

© The authors, 2025.



*The SMAI Journal of Computational Mathematics is a member
of the Centre Mersenne for Open Scientific Publishing*

<http://www.centre-mersenne.org/>

Submissions at <https://smai-jcm.centre-mersenne.org/ojs/submission>

e-ISSN: 2426-8399





Combining reciprocity gap method and state estimator for source identification in an advection diffusion equation

ANTOINE TONNOIR¹

¹ LMI INSA Rouen Normandie, Rouen, France
E-mail address: antoine.tonnoir@insa-rouen.fr.

Abstract. In this paper, we propose to revisit a reciprocity gap method for solving point inverse source problem in an advection diffusion equation. The motivation of this problem is in ecology for pollutant source identification in a river. We propose the construction of original (numerically computed) adjoint functions that allows to consider more realistic geometries and river flows. The method is combined with a state estimator which allows to accelerate the identification process. The proposed method is validated on several examples.

2020 Mathematics Subject Classification. 35K10, 35R30, 65M32.

Keywords. Inverse source problem, advection diffusion equation, reciprocity gap method.

1. Introduction

This work is motivated by the development of numerical methods for source identification in the context of ecology [2, 14, 15, 18, 19]. In particular, we are interested here in pollutant source localization in a river given some downstream data, that is to say data on a part of the boundary of the domain. To describe the pollutant propagation, we consider a simple model from the literature: an advection-diffusion equation [11, 15, 20]. The source term is assumed to be a point to correspond to a localized source. This problem corresponds to a classical inverse source problem, and several studies already have been dedicated to this subject. Let us recall some of them relevant in our context.

A first question that naturally arise in inverse problem is the question of identifiability. In other words, can we recover the unknown information from the available data? For the problem of point source identification in a diffusion equation, it was shown in [3, 7, 12] that we can uniquely recover the source position given measurements on the boundary of the domain. For advection diffusion problem, given upstream and downstream boundary data (that is to say, on a part of the boundary), a similar result was shown in [13, 15, 17] (for different situations, 1D, 2D or 3D) under some assumptions on the flow and the diffusion tensor. The proof of the result relies on the reciprocity gap method. Based on this identifiability result, a non iterative algorithm is proposed to solve the inverse problem. The general idea is to construct analytically appropriate adjoint functions to reformulate the inverse problem as a simple non linear equation (the identifiability result coming from the injectivity of the adjoint functions). A major advantage of this approach comes from the fact that the numerical algorithm of resolution of the inverse problem is fast and no minimization process is required. Yet, a drawback is the analytical construction of the adjoint functions which requires relatively strong assumptions on the geometry of the river, the flow and the diffusion tensor [15, 17].

Other approaches can be found in the literature such as gradient minimization approach, see [10, 25] and references inside, or Kalman filtering approach, see for instance [24, 26], which allow greater flexibility on the hypotheses on the parameters of the model, but usually require larger computational effort to solve the inverse problem. Let us also mention [8, 9] which consider also the problem of source identification using an observer approach but for a different equation (wave equation).

Besides, as initiated in [17], we wish to propose an “online” monitoring procedure that allows to estimate the pollutant source position at each time (making the identification procedure in real time). To be able to do that efficiently, one needs a state estimator for the advection diffusion equation that describes the pollutant propagation. Building this type of estimator is not an easy task, especially with partial boundary data. Let us mention some papers in our context. In [4] a Kalman estimator with an efficient implementation based on matrix compression is proposed for general elliptic equations, and in particular for the advection-diffusion equation. From stabilization process, see [21, 23] and the reference inside, one can also derive a state estimator. Yet, boundary stabilization methods usually rely on the knowledge of the solution in the whole domain, which is not affordable in our case since we consider only measurements on a part of the boundary.

In this work, we propose to revisit the reciprocity gap approach and more precisely the construction of the adjoint functions via a (semi-)numerical computation, the idea being to take advantage as much as possible of the separation of variables method. We wish to propose a method that gives a general way to construct these adjoint functions and allows to fully relax the conditions needed to build them analytically. Thus, it generalizes previous works and keep the advantage of a fast identification procedure. Also, we explain how to implement the identification procedure to get an online monitoring and to identify multiple sources in the case when they are “well-separated” (in a sense that will be clarify hereafter). A second main contribution of this work is the construction of a simple and cheap state estimator based on (partial) boundary data (under some assumptions) for which we give an error bound.

The rest of the paper is organized as follows. In Section 2, we present the model and give the mathematical formulation of the inverse problem. Then, in Section 3 we recall the general approach of the adjoint functions method and give the reformulation of the inverse problem as a simple resolution of two non-linear equations. In Section 4, we present the construction of a cheap state estimator based only on (partial) boundary measurements. In Section 5, we study in details the case of a rectangular river. This particular example allows to illustrate many physical intuitions and to propose a general procedure to construct the adjoint functions. Finally, in Section 6 several generalization and tests are presented to assess the method in more realistic situations. In particular, we consider a synthetic case of two sources identification in the Seine river in Rouen (France).

2. Mathematical model and formulation of the problem

To describe the diffusion and propagation of a quantity u of a pollutant in a river, we consider the following simple linear model:

$$\begin{cases} \partial_t u + \mathbf{V} \cdot \nabla u - \operatorname{div}(D\nabla u) = f & \text{in } \Omega, \\ u = 0 & \text{on } \Gamma_{in}, \\ D\nabla u \cdot \nu = 0 & \text{on } \Gamma_{out} \cup \Gamma, \end{cases} \quad (2.1)$$

where \mathbf{V} describes the velocity field, D is the diffusion tensor¹ (a 2×2 positive definite matrix) and the domain $\Omega \subset \mathbb{R}^2$ represents the geometry of the river, see Figure 2.1. The boundary of the domain $\partial\Omega$ is split into three parts: Γ_{in} corresponding to the in-flow boundary, Γ_{out} corresponding to the out-flow boundary and Γ corresponding to the “lateral” boundary. The homogeneous Dirichlet boundary condition in Γ_{in} corresponds to the fact that we suppose that no pollutant comes from upstream, and the homogeneous Neumann boundary condition in Γ_{out} corresponds to a simple open domain boundary conditions (modeling the fact that the river do not stop at Γ_{out}).

¹Note that \mathbf{V} and D may vary with the position (x, y) .

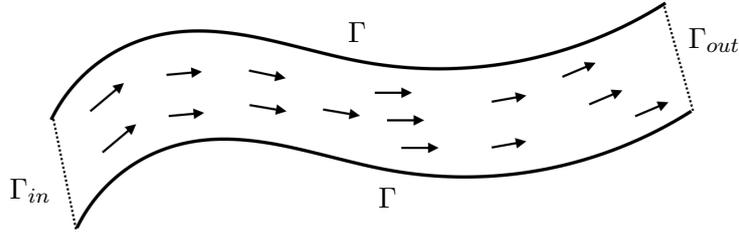


FIGURE 2.1. Geometry of the river Ω . The black arrow corresponds to the velocity field \mathbf{V} .

In the sequel, we will assume the following hypotheses:

- (H1) the initial condition in (2.1) is $u(t = 0, \cdot) = 0$, which corresponds to assume that there is no pollutant in the domain at initial time.
- (H2) the vector field \mathbf{V} is irrotational and divergence free (which amounts to consider that it comes from a potential flow), and $\mathbf{V} \cdot \nu = 0$ on Γ where ν is the outward normal (no flow across the lateral boundary),
- (H3) and the source term is very localized so that we consider

$$f(t, x, y) = \delta_{\mathbf{s}}(x, y)\lambda(t) \quad (2.2)$$

where $\lambda \geq 0$ and $\delta_{\mathbf{s}}$ is the Dirac distribution at position \mathbf{s} corresponding to the pollutant source. The function λ represents the amount of pollutant emitted at each time t and the source is supposed to vanish after time $T_0 > 0$, which means that $\lambda(t) = 0$ for all $t \geq T_0$. This time T_0 is supposed to be unknown, as well as λ .

Remark 2.1. One could also consider a reaction term in (2.1) with a parameter $R \geq 0$. The methodology presented below is the same. In particular, if R is constant then we can always come back to equation (2.1) using the change of unknown $u = \tilde{u} e^{Rt}$ where \tilde{u} is the solution of the problem with the reaction coefficient.

The purpose then is: Given the boundary measurements over time $[0, T]$, $T > T_0$,

$$\mathcal{M} = \{(u, D\nabla u \cdot \nu) \text{ on } [0, T] \times \{\Gamma_{in} \cup \Gamma_{out}\}\} \quad (2.3)$$

find the source position \mathbf{s} in the shortest time. Note that this inverse problem “simply” amounts to find the two coordinates (s_x, s_y) of the source term. An important difference with the previous mentioned works is the fact that we wish here to have an online procedure, that is to say a procedure that gives an estimate of the source position at each time $t \in [0, T]$. This means in particular that we wish to avoid the “direct” least-square minimization approach which would require to know the measurements during all time $[0, T]$.

Also, we can be interested in the reconstruction of the function $\lambda(\cdot)$, but this issue will not be discussed in this work. This can be done a posteriori as in [14, 16]. We will see that even if $\lambda(\cdot)$ is unknown, we can recover the total amount of emitted pollutant and the source position. As already mentioned, the question of identifiability for this inverse problem have been studied in several previous work [7, 15, 17] requiring additional (quite strong) hypotheses on the tensor D , the velocity field \mathbf{V} and the geometry Ω to be ensured. Here, we will focus on a numerical procedure to reconstruct \mathbf{s} or a candidate for \mathbf{s} even if identifiability is not ensured.

Remark 2.2. Using classical regularity results, see [12, 22], one can show that if the source term is not on the boundary, the solution is in $C^0([0, T], H^{1/2}(\Omega))$ and the data $u|_{\Gamma_{out}}$ is in $C^0([0, T], L^2(\Gamma_{out}))$.

Remark 2.3. If we suppose the time dependency of the source $\lambda(\cdot)$ known, then the identifiability of the source with boundary measurements can be obtained using the results of [1]. Indeed, the advection-diffusion equation considered here fits into the general framework proposed in this paper.

3. The general approach

3.1. Reformulation of the inverse problem

The starting point of the method consists in testing equation (2.1) with appropriate adjoint functions. So, multiplying equation (2.1) with a function v and integrating in time and space, we get:

$$\int_0^T \langle u, \mathcal{L}^* v \rangle dt + \mathcal{B}(u, v, T) + \int_{\Omega} u(T, \cdot) v(T, \cdot) = \int_0^T \lambda(t) v(t, \mathbf{s}) dt \quad (3.1)$$

where $\langle \cdot, \cdot \rangle$ is the duality product between $H^{1/2}(\Omega)$ and $H^{-1/2}(\Omega)$, \mathcal{L} is the operator defined by

$$\mathcal{L}(v) = \partial_t v + \mathbf{V} \cdot \nabla v - \operatorname{div}(D\nabla v), \quad \forall v \in C^0([0, T], H^{1/2}(\Omega)), \quad (3.2)$$

and \mathcal{L}^* is the adjoint operator of \mathcal{L}

$$\mathcal{L}^*(v) = -\partial_t v - \mathbf{V} \cdot \nabla v - \operatorname{div}(D\nabla v), \quad \forall v \in C^0([0, T], H^{1/2}(\Omega)). \quad (3.3)$$

The boundary term \mathcal{B} is given by

$$\mathcal{B}(u, v, T) = \int_0^T \int_{\partial\Omega} uv\mathbf{V} \cdot \nu - vD\nabla u \cdot \nu + uD\nabla v \cdot \nu d\Gamma. \quad (3.4)$$

The idea then consists in taking v satisfying $\mathcal{L}^* v = 0$. Also, following the idea of separation of variables, we seek for v of the form: $v(t, x, y) = \mu(t)w(x, y)$. Simple computations then show that $\mu(t) := e^{\alpha t}$ for any real constant α and w must satisfy

$$-\operatorname{div}(D\nabla w) - \mathbf{V} \cdot \nabla w + \alpha w = 0. \quad (3.5)$$

In the sequel, we will consider the case $\alpha = 0$ so that $v(t, x, y) = w(x, y)$ (which means stationary solutions), the motivation of this choice being explained after in Remark 3.5. Then, to fully determine the function w one only need to specify the boundary conditions. They are chosen according to the boundary data knowledge and the goal is to keep in $\mathcal{B}(u, v, T)$ only known terms. Therefore, on Γ we will impose homogeneous Neumann conditions. On Γ_{out} , we will impose a (non homogeneous) Dirichlet boundary conditions and on Γ_{in} either homogeneous Neumann or Dirichlet conditions. To sum up, we will have

Definition 3.1 (Adjoint function). Given $g_{out} \in H^{1/2}(\Gamma_{out})$, the adjoint function w is defined as the unique solution in $H^1(\Omega)$ to

$$\begin{cases} -\operatorname{div}(D\nabla w) - \mathbf{V} \cdot \nabla w = 0 & \text{in } \Omega, \\ D\nabla w \cdot \nu = 0 & \text{on } \Gamma, \\ w = g_{out} & \text{on } \Gamma_{out}, \\ \beta D\nabla w \cdot \nu + (1 - \beta)w = 0 & \text{on } \Gamma_{in}, \end{cases} \quad (3.6)$$

where $\beta = 0$ or $\beta = 1$, depending on the condition one wishes to impose on Γ_{in} .

This problem is well-posed thanks to the Dirichlet condition on Γ_{out} and using Lax–Milgram theorem. Thus, using the boundary conditions satisfied by $v = w$, the term $\mathcal{B}(u, v, T) = \mathcal{B}(u, w, T)$ simplifies and we get

Lemma 3.2. *For any adjoint function w solution to (3.6), we have:*

$$\mathcal{B}(u, w, T) = \int_0^T \int_{\Gamma_{out}} uw \mathbf{V} \cdot \nu + u D \nabla w \cdot \nu \, d\Gamma - \int_0^T \int_{\Gamma_{in}} w D \nabla u \cdot \nu \, d\Gamma. \quad (3.7)$$

The proof of this result is direct noticing that:

$$\begin{aligned} \mathcal{B}(u, w, T) &= \int_0^T \int_{\Gamma_{in} \cup \Gamma_{out}} uw \mathbf{V} \cdot \nu - w D \nabla u \cdot \nu + u D \nabla w \cdot \nu \, d\Gamma, \\ &= \int_0^T \int_{\Gamma_{out}} uw \mathbf{V} \cdot \nu + u D \nabla w \cdot \nu \, d\Gamma - \int_0^T \int_{\Gamma_{in}} w D \nabla u \cdot \nu \, d\Gamma. \end{aligned} \quad (3.8)$$

As mentioned, computing this term requires only the knowledge of u on the out-flow boundary Γ_{out} and $D \nabla u \cdot \nu$ on the in-flow boundary Γ_{in} . Note also that one could avoid the knowledge of the data $D \nabla u \cdot \nu$ on Γ_{in} by taking $w = 0$ on Γ_{in} (i.e. $\beta = 0$ in (3.6)).

Now, with this adjoint function, equation (3.1) can also be simplified and we straightforwardly get

Lemma 3.3. *For any adjoint function w solution to (3.6), we have:*

$$\mathcal{B}(u, w, T) + \int_{\Omega} u(T, \cdot) w(\cdot) = w(\mathbf{s}) \Lambda(T) \quad \text{where} \quad \Lambda(T) = \int_0^T \lambda(t) dt. \quad (3.9)$$

Corollary 3.4. *Taking successively $w = 1$ (which is solution to (3.6) with $\beta = 1$ and $g_{out} = 1$), $w = w_A$ and $w = w_B$ two other solutions to (3.6), we obtain*

$$\begin{cases} \Lambda(T) &= \mathcal{B}(u, 1, T) + (u(T, \cdot), 1)_{L^2(\Omega)}, \\ w_A(s_x, s_y) &= \frac{\mathcal{B}(u, w_A, T) + (u(T, \cdot), w_A)_{L^2(\Omega)}}{\Lambda(T)}, \\ w_B(s_x, s_y) &= \frac{\mathcal{B}(u, w_B, T) + (u(T, \cdot), w_B)_{L^2(\Omega)}}{\Lambda(T)}. \end{cases} \quad (3.10)$$

In the above system, the unknown terms are the final state $u(T, \cdot)$ and the source position $\mathbf{s} = (s_x, s_y)$. $\Lambda(T)$ the total amount of pollutant emitted during time $[0, T_0]$ is given by the known boundary term $\mathcal{B}(u, 1, T)$ and $u(T, \cdot)$. For the final state $u(T, \cdot)$ ², we will approximate it by an estimator \hat{u} that will be explained in the next section. Note that a basic idea of state estimator is to take $\hat{u} = 0$ since the solution u of the direct problem tends to 0 as t tends to $+\infty$. Yet, as we will see, we can do better in many cases.

Then, replacing $u(T, \cdot)$ by $\hat{u}(T, \cdot)$ the quantity $\Lambda(T)$ can be estimated by

$$\Lambda(T) \simeq \mathcal{B}(u, 1, T) + (\hat{u}(T, \cdot), 1)_{L^2(\Omega)}. \quad (3.11)$$

Let us note that this gives an estimate of the total amount of pollutant that has been emitted.

Remark 3.5. Note that if we take $\alpha \neq 0$ in (3.5), one should in fact compute

$$\int_0^T \lambda(t) e^{\alpha t} dt \quad \text{instead of} \quad \Lambda(T) = \int_0^T \lambda(t) dt \quad \text{in (3.9).}$$

Yet, since $v(x, y, t) = e^{\alpha t}$ does not satisfy $\mathcal{L}^* v = 0$ (for $\alpha \neq 0$), we cannot use $\mathcal{B}(u, e^{\alpha t}, T) + (\hat{u}(T, \cdot), e^{\alpha t})_{L^2(\Omega)}$ to estimate the above quantity. This is why we took $\alpha = 0$. Nevertheless, let us note that $\mathcal{B}(u, 1, T) + (\hat{u}(T, \cdot), 1)_{L^2(\Omega)}$ can be a good approximation of the above quantity for small α .

Similarly, we replace $u(T, \cdot)$ by $\hat{u}(T, \cdot)$ in the equations satisfied by w_A and w_B in (3.10) and we get

²Let us remark that in adjoint state method for the computation of the gradient of the functional to minimize, we usually take $v(T, \cdot) = 0$ to eliminate this term.

Definition 3.6 (Reformulation of the inverse problem). We will seek for the estimate source position $\hat{\mathbf{s}} = (\hat{s}_x, \hat{s}_y)$ solution to the system:

$$\begin{cases} w_A(s_x, s_y) = \mathcal{D}_A(T), \\ w_B(s_x, s_y) = \mathcal{D}_B(T), \end{cases} \quad (3.12)$$

where

$$\mathcal{D}_A(T) = \frac{\mathcal{B}(u, w_A, T) + (\hat{u}(T, \cdot), w_A)_{L^2(\Omega)}}{\mathcal{B}(u, 1, T) + (\hat{u}(T, \cdot), 1)_{L^2(\Omega)}} \quad \text{and} \quad \mathcal{D}_B(T) = \frac{\mathcal{B}(u, w_B, T) + (\hat{u}(T, \cdot), w_B)_{L^2(\Omega)}}{\mathcal{B}(u, 1, T) + (\hat{u}(T, \cdot), 1)_{L^2(\Omega)}}. \quad (3.13)$$

This problem is nothing but a system of two non-linear equations. In fact, the coordinates (\hat{s}_x, \hat{s}_y) of the source localization can be interpreted simply as the intersection of the two level sets of w_A and w_B . Let us also remark that the question of identifiability amounts to know if system (3.12) has a unique solution. In particular, if we can find $w_A(x, y) = w_A(x)$ and $w_B(x, y) = w_B(y)$ solution to (3.6), then the two equations are decoupled and identifiability can be easily shown providing monotonicity properties of w_A and w_B . Following this idea, we will propose a way to construct w_A and w_B first on a simple rectangular geometry in Section 5, and then on a general case in Section 6.

3.2. Algorithm for the “online” monitoring

To end this section, let us emphasize that the position estimate can be computed in fact at any time t , not only T . This allows to consider an “online” monitoring of the section of river Ω by solving (3.12) at each time t and replacing therefore $\mathcal{D}_A(T)$ and $\mathcal{D}_B(T)$ by

$$\mathcal{D}_A(t) = \frac{\mathcal{B}(u, w_A, t) + (\hat{u}(t, \cdot), w_A)_{L^2(\Omega)}}{\mathcal{B}(u, 1, t) + (\hat{u}(t, \cdot), 1)_{L^2(\Omega)}} \quad \text{and} \quad \mathcal{D}_B(t) = \frac{\mathcal{B}(u, w_B, t) + (\hat{u}(t, \cdot), w_B)_{L^2(\Omega)}}{\mathcal{B}(u, 1, t) + (\hat{u}(t, \cdot), 1)_{L^2(\Omega)}}. \quad (3.14)$$

As explained, the resolution can be done by computing the level sets associated to each adjoint function. Yet, in the sequel the adjoint functions will be only known numerically on a mesh (we use Lagrange Finite Elements for the discretization and resolution of problem (3.6)), i.e. at some positions $(x_k, y_k)_k$ $k \in \{1, \dots, N_s\}$ with $N_s > 0$. Therefore, instead of computing the level set, we simply consider the set of points for which w_i is in an interval around $\mathcal{D}_i(t)$, $i \in \{A, B\}$ ³:

$$\mathcal{S}_i(t) = \{(x_k, y_k), k \in \{1, N_s\} \text{ s.t. } w_i(x_k, y_k) \in [\mathcal{D}_i(t) - \text{tol}, \mathcal{D}_i(t) + \text{tol}]\}. \quad (3.15)$$

If $\mathcal{S}_A(t) \cap \mathcal{S}_B(t) = \emptyset$, the idea is to gradually increase the value of $\text{tol} > 0$ until the intersection is not empty. Then, the estimate position of the source is defined by

Definition 3.7 (Position estimate). The estimate source position $\hat{\mathbf{s}}(t) = (\hat{s}_x(t), \hat{s}_y(t))$ at time t is given by

$$(\hat{s}_x(t), \hat{s}_y(t)) = \underset{(x,y) \in \mathcal{S}_A(t) \cap \mathcal{S}_B(t)}{\text{argmin}} (w_A(x, y) - \mathcal{D}_A(t))^2 + (w_B(x, y) - \mathcal{D}_B(t))^2. \quad (3.16)$$

Let us underline that this minimization problem is easy to solve since the set $\mathcal{S}_A(t) \cap \mathcal{S}_B(t)$ is very small. Let us also remark that other choices are possible for the functional in (3.16).

Remark 3.8. The parameter tol in (3.15) can also be used to determine a possible area of the source term, in case of uncertainties on the data.

³One could also use an interpolation technique to compute the level sets.

Besides, to get an efficient implementation of $\mathcal{B}(u, w, t)$ we can remark that it can be computed recursively. Indeed, if $\Delta t > 0$ denotes the time stepping of the measurements and $t_n = n\Delta t$, we have:

$$\begin{aligned} \mathcal{B}(u, w, t_{n+1}) &\simeq \mathcal{B}(u, w, t_n) \\ &+ \Delta t \left(\int_{\Gamma_{out}} u(t_{n+1}, \cdot) (w \mathbf{V} \cdot \nu + D \nabla w \cdot \nu) \, d\Gamma - \int_{\Gamma_{in}} w D \nabla u(t_{n+1}, \cdot) \cdot \nu \, d\Gamma \right). \end{aligned} \quad (3.17)$$

The procedure of source position estimation is sum up in Algorithm 1. Note that the source estimate is computed as soon as the estimate of total emitted pollutant $|\mathcal{B}(u, 1, t) + (\hat{u}(t, \cdot), 1)_{L^2(\Omega)}|$ is larger than a given value $\varepsilon > 0$.

Algorithm 1 Source estimation

Require: w_A and w_B appropriate adjoint functions solutions to (3.6)

while $t_n \leq T$ **do**

 Update the term $\mathcal{B}(u, w, t_n)$ using equation (3.17)

 Compute the state estimator \hat{u} at time t_n (See Section 4)

if $|\mathcal{B}(u, 1, t_n) + (\hat{u}(t_n, \cdot), 1)_{L^2(\Omega)}| \geq \varepsilon$ **then**

 Compute $\mathcal{D}_A(t_n)$ and $\mathcal{D}_B(t_n)$ using equation (3.14)

$\text{tol} \leftarrow 10^{-5}$

 Compute the sets of \mathcal{S}_A and \mathcal{S}_B using equation (3.15)

while $\mathcal{S}_A \cap \mathcal{S}_B = \emptyset$ **do**

 Increase tol by a factor > 1

 Update the sets \mathcal{S}_A and \mathcal{S}_B using equation (3.15)

end while

 Solve the minimization problem (3.16) to get the estimate (\hat{s}_x, \hat{s}_y)

end if

 Update time: $t_n \leftarrow t_n + \Delta t$

end while

4. The state estimator

In this section, the purpose is to construct a state estimator \hat{u} of u . Given boundary data, this is a difficult task for an advection diffusion equation, as explained in the introduction. We require the estimator the following criteria:

- It is based only on the measurements of u on Γ_{out} and it converges to u as t tends to $+\infty$,
- The computational cost is low enough to allow “online” monitoring as explained in the previous section.

4.1. Spectral decomposition of the solution

Since the source term is unknown, the idea is to build a state estimator for $t \geq t_0 \geq T_0$ (although we do not know the extinction time T_0 of the source term)⁴ when the source is null. In other words, we wish to estimate for all $t \geq t_0$ u solution to:

$$\begin{cases} \partial_t u + \mathbf{V} \cdot \nabla u - \operatorname{div}(D \nabla u) = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \Gamma_{in}, \\ D \nabla u \cdot \nu = 0 & \text{on } \Gamma_{out} \cup \Gamma, \end{cases} \quad (4.1)$$

⁴The introduction of t_0 will be useful in the sequel

with the unknown initial data $u(t_0, \cdot)$. To construct this estimator, we will rely on a spectral decomposition of the solution. To be able to justify this decomposition, we need to add the following hypothesis on \mathbf{V} and D :

(H4) we suppose that

$$\operatorname{rot}(D^{-1}\mathbf{V}) = 0 \iff \partial_y(D^{-1}\mathbf{V})_1 = \partial_x(D^{-1}\mathbf{V})_2. \quad (4.2)$$

If this hypothesis is not satisfied, then the spectral decomposition we will present hereafter is wrong and so is the associated estimator. Let see now where does this condition comes from. To get the spectral decomposition of u solution to (4.1), one needs to compute the eigenvalues λ and the eigenfunctions φ of the eigenvalue problem:

$$\begin{cases} -\operatorname{div}(D\nabla\varphi) + \mathbf{V} \cdot \nabla\varphi = \lambda\varphi & \text{in } \Omega, \\ \varphi = 0 & \text{on } \Gamma_{in}, \\ D\nabla\varphi \cdot \nu = 0 & \text{on } \Gamma \cup \Gamma_{out}. \end{cases} \quad (4.3)$$

Using the fact that $\operatorname{div}(\mathbf{V}) = 0$ and setting $\varphi(\cdot) := e^{\zeta(\cdot)}\psi(\cdot)$, we get that

$$\begin{aligned} & -\operatorname{div}\left(D\nabla\varphi - \frac{1}{2}\mathbf{V}\varphi\right) + \frac{1}{2}\mathbf{V} \cdot \nabla\varphi = \lambda\varphi, \\ \iff & -\operatorname{div}\left(e^\zeta D\nabla\psi + \left(D\nabla\zeta - \frac{1}{2}\mathbf{V}\right)e^\zeta\psi\right) + \frac{1}{2}\mathbf{V} \cdot \nabla(e^\zeta\psi) = \lambda e^\zeta\psi, \end{aligned} \quad (4.4)$$

Under hypothesis (4.2), one can choose ζ s.t.

$$D\nabla\zeta - \frac{1}{2}\mathbf{V} = 0 \iff \nabla\zeta = \frac{1}{2}D^{-1}\mathbf{V}. \quad (4.5)$$

Thus, equation (4.4) becomes:

$$\begin{aligned} & -\operatorname{div}(D\nabla\psi) - \underbrace{D\nabla\psi \cdot \nabla\zeta + \frac{1}{2}(\mathbf{V} \cdot \nabla\psi)}_{=0} + \frac{1}{2}(\mathbf{V} \cdot \nabla\zeta)\psi = \lambda\psi, \\ \iff & -\operatorname{div}(D\nabla\psi) + \frac{1}{4}(\mathbf{V} \cdot D^{-1}\mathbf{V})\psi = \lambda\psi, \end{aligned}$$

and the eigenvalue problem (4.3) can rewrite as follows

$$\begin{cases} -\operatorname{div}(D\nabla\psi) + \frac{1}{4}(\mathbf{V} \cdot D^{-1}\mathbf{V})\psi = \lambda\psi & \text{in } \Omega, \\ \psi = 0 & \text{on } \Gamma_{in}, \\ D\nabla\psi \cdot \nu + \frac{1}{2}\mathbf{V} \cdot \nu\psi = 0 & \text{on } \Gamma \cup \Gamma_{out}. \end{cases} \quad (4.6)$$

Since the operator $-\operatorname{div}(D\nabla\cdot) + \frac{1}{4}(\mathbf{V} \cdot D^{-1}\mathbf{V})(\cdot)$ (with the above B.C.) is self-adjoint and positive, we are ensured that there is a family $(\psi_k)_{k \in \mathbb{N}}$ of (real) eigenfunctions associated with a non-decreasing sequence of positive eigenvalues $(\lambda_k)_{k \in \mathbb{N}}$ that tends to $+\infty$. Moreover, this family $(\psi_k)_{k \in \mathbb{N}}$ can be normalized to form an orthonormal basis of $L^2(\Omega)$. Thus, we can justify the following decomposition of u :

Proposition 4.1 (Spectral decomposition). *Under the condition (4.2), the solution u to (4.1) is given by*

$$u(t, x, y) = \sum_{k \geq 0} A_k(t_0) e^{-\lambda_k(t-t_0)} \varphi_k(x, y) \quad \text{where} \quad \varphi_k(x, y) = e^{\zeta(x, y)} \psi_k(x, y), \quad \forall k \in \mathbb{N}, \quad (4.7)$$

with $(\psi_k)_{k \in \mathbb{N}}$ the family of orthonormal eigenfunctions solutions to (4.6). Also, the parameters $A_k(t_0)$ are uniquely determined by the value of u at time t_0 and explicitly given by:

$$A_k(t_0) = \left(e^{-\zeta} u(t_0, \cdot), \psi_k \right)_{L^2(\Omega)}, \quad \forall k \in \mathbb{N}.$$

Let us recall that $u(t_0, \cdot)$ is unknown, this is why we need an estimator. One can notice that the family $(\varphi_k)_{k \in \mathbb{N}}$ is not orthonormal according to the scalar product in L^2 but, since e^ζ is uniformly strictly positive, it is an orthonormal basis of $L^2(\Omega)$ according to the weighted scalar product:

$$(\varphi, \tilde{\varphi})_\zeta = \int_\Omega \varphi(x, y) \tilde{\varphi}(x, y) e^{-2\zeta} d\Omega. \quad (4.8)$$

Indeed, using the properties of $(\psi_k)_{k \in \mathbb{N}}$, we clearly have $(\varphi_k, \varphi_l)_\zeta = \delta_{kl}$ and for any function $f \in L^2(\Omega)$:

$$f = \sum_{k \geq 0} (f e^{-\zeta}, \psi_k) \varphi_k = \sum_{k \geq 0} (f, \varphi_k)_\zeta \varphi_k.$$

This remark will simplify some computations in what follows. Let us note also that this scalar product leads to an equivalent norm with the classical L^2 norm. Moreover, we have:

$$f \in L^2(\Omega) \iff \|f\|_\zeta^2 = \sum_{k \geq 0} (f, \varphi_k)_\zeta^2 < +\infty. \quad (4.9)$$

Similarly, we have that:

$$f \in H^1(\Omega) \iff \|f\|_{H^1_\zeta}^2 = \sum_{k \geq 0} (1 + \lambda_k) (f, \varphi_k)_\zeta^2 < +\infty. \quad (4.10)$$

Here again, this defines an equivalent norm with the classical H^1 norm. This last remark will be useful for using the trace theorem in the sequel.

Remark 4.2. Let us note that the hypothesis (4.2) is always satisfied if $D = d\text{Id}$ with $d > 0$ a constant parameter. In that particular case, the function ζ defined in (4.5) can be deduced directly from the potential of \mathbf{V} .

4.2. State estimator

Let us consider the data u on Γ_{out} during a time frame $[t - t_f, t]$, with $t_f > 0$ given. The purpose will be to design an estimate of $u(t, \cdot)$ based on these data. According to Proposition 4.1, we will seek for \hat{u} of the form:

$$\hat{u}(s, x, y) = \sum_{k=0}^N \hat{A}_k(t - t_f) \varphi_k(x, y) e^{-\lambda_k(s-t+t_f)}, \quad \forall s \in [t - t_f, t], \quad (4.11)$$

where N is the given number of eigenfunctions we wish to consider. The coefficients $\hat{A}_k(t - t_f)$ correspond to the decomposition of \hat{u} at time $t_0 = t - t_f$. Once these coefficients are determined, we get the estimate of $u(t, \cdot)$:

Definition 4.3 (State estimator). Under the condition (4.2) and given the coefficients $\hat{A}_k(t - t_f)$, the estimate of $u(t, \cdot)$ is obtained by:

$$\hat{u}(t, x, y) = \sum_{k=0}^N \hat{A}_k(t - t_f) \varphi_k(x, y) e^{-\lambda_k t_f}. \quad (4.12)$$

Since the data are known only on the boundary, we cannot take advantage of the orthonormality of the eigenfunctions in $L^2(\Omega)$ to compute the coefficients $\hat{A}_k(t - t_f)$. Also, the family of traces $(\varphi_k|_{\Gamma_{out}})_k$ is usually not a linearly independent family of functions⁵, so we cannot simply decompose the data $u(t - t_f, \cdot)|_{\Gamma_{out}}$. The idea is then to consider the whole data during the time frame $[t - t_f, t]$ and to solve the minimization problem:

$$\min_{(\hat{A}_k(t-t_f))_{k \in \{0, \dots, N\}}} \int_{t-t_f}^t \|u(s, \cdot)|_{\Gamma_{out}} - \hat{u}(s, \cdot)|_{\Gamma_{out}}\|_{L^2(\Gamma_{out})}^2 ds. \quad (4.13)$$

⁵We can simply show it for instance in the case of a rectangular domain

Proposition 4.4. *The solution to the minimization problem (4.13) is given by the resolution of the linear system:*

$$\mathcal{G}\hat{\mathbf{A}}(t - t_f) = \mathbf{g}, \quad (4.14)$$

where the Gramian matrix \mathcal{G} is given by

$$\begin{aligned} \mathcal{G}_{ij} &= \left(\varphi_j e^{-\lambda_j(\cdot - t + t_f)}, \varphi_i e^{-\lambda_i(\cdot - t + t_f)} \right)_{L^2([t - t_f, t] \times \Gamma_{out})}, \\ &= \left(\varphi_j e^{-\lambda_j(\cdot)}, \varphi_i e^{-\lambda_i(\cdot)} \right)_{L^2([0, t_f] \times \Gamma_{out})}, \quad \forall (i, j) \in \{0, \dots, N\}^2, \end{aligned} \quad (4.15)$$

the vector $\mathbf{g} \in \mathbb{R}^{N+1}$ is defined by

$$\mathbf{g}_k = \left(u|_{\Gamma_{out}}, \varphi_k e^{-\lambda_k(\cdot - t + t_f)} \right)_{L^2([t - t_f, t] \times \Gamma_{out})}, \quad k \in \{0, \dots, N\},$$

and the vector $\hat{\mathbf{A}}(t - t_f) \in \mathbb{R}^{N+1}$ is the vector of components $\hat{A}_k(t - t_f)$ defined in the state estimator (4.12).

Proof. The proof is very classical. Since $\hat{u}(s, \cdot)$ is defined by (4.11), we deduce that

$$\begin{aligned} \|u(s, \cdot)|_{\Gamma_{out}} - \hat{u}(s, \cdot)|_{\Gamma_{out}}\|_{L^2(\Gamma_{out})}^2 &= \left\| \sum_{k=0}^N \hat{A}_k(t - t_f) \varphi_k(\cdot, \cdot)|_{\Gamma_{out}} \right\|_{L^2(\Gamma_{out})}^2 e^{-2\lambda_k(s - t + t_f)} \\ &\quad - 2 \sum_{k=0}^N \hat{A}_k(t - t_f) \left(\varphi_k(x, y)|_{\Gamma_{out}}, \hat{u}(s, \cdot)|_{\Gamma_{out}} \right)_{L^2(\Gamma_{out})} e^{-\lambda_k(s - t + t_f)} \\ &\quad + \|\hat{u}(s, \cdot)|_{\Gamma_{out}}\|_{L^2(\Gamma_{out})}^2 \end{aligned}$$

Integrating in time in $[t - t_f, t]$ and using the definition of \mathcal{G} , \mathbf{g} and $\hat{\mathbf{A}}(t - t_f)$, we get

$$\begin{aligned} \int_{t-t_f}^t \|u(s, \cdot)|_{\Gamma_{out}} - \hat{u}(s, \cdot)|_{\Gamma_{out}}\|_{L^2(\Gamma_{out})}^2 ds &= \left(\mathcal{G}\hat{\mathbf{A}}(t - t_f) \right) \cdot \hat{\mathbf{A}}(t - t_f) - 2\hat{\mathbf{A}}(t - t_f) \cdot \mathbf{g} \\ &\quad + \int_{t-t_f}^t \|\hat{u}(s, \cdot)|_{\Gamma_{out}}\|_{L^2(\Gamma_{out})}^2 ds. \end{aligned}$$

We conclude using the fact that \mathcal{G} is a symmetric matrix so that the minimum of the above quantity is given by the resolution of the linear system (4.14). \blacksquare

Let us make some remarks on linear system (4.14). First, the Gramian matrix is independent of t , so its computations and factorization can be done once and then used for the estimate at any time t . Also, though for N large the matrix \mathcal{G} is (usually) very ill-conditioned, it is still better than the Gramian matrix obtained only with the family $(\varphi_k|_{\Gamma_{out}})_k$ (as we will illustrate in the numerical results hereafter). Intuitively, this can be understood by the fact that the time evolution brings information by the speed of decrease of the solution via the eigenvalues λ_k . Finally, let us mention that at each time t , the main cost for computing the estimator $\hat{u}(t, \cdot)$ is the computation of the coefficients \mathbf{g}_k . To speed up these computations, we can use a recursive formula, as for the coefficient $\mathcal{B}(u, w, t)$. Once

again from a computational point of view (with $\Delta t > 0$ the time stepping and $t_n = n\Delta t$), we have

$$\begin{aligned}
 \mathbf{g}_k(t_{n+1}) &= \int_{t_n+\Delta t-t_f}^{t_n+\Delta t} \int_{\Gamma_{out}} u|_{\Gamma_{out}} \varphi_k|_{\Gamma_{out}} e^{-\lambda_k(s-(t_n+\Delta t-t_f))} d\Gamma ds, \\
 &\simeq e^{\lambda_k \Delta t} \left(\int_{t_n-t_f}^{t_n} \int_{\Gamma_{out}} u|_{\Gamma_{out}} \varphi_k|_{\Gamma_{out}} e^{-\lambda_k(s-(t_n-t_f))} d\Gamma ds \right. \\
 &\quad \left. - \Delta t \int_{\Gamma_{out}} u(t_n-t_f)|_{\Gamma_{out}} \varphi_k|_{\Gamma_{out}} e^{-\lambda_k(t_n-t_f)} d\Gamma \right) \\
 &\quad + \Delta t \int_{\Gamma_{out}} u(t_{n+1})|_{\Gamma_{out}} \varphi_k|_{\Gamma_{out}} e^{-\lambda_k t_{n+1}} d\Gamma, \\
 &\simeq e^{\lambda_k \Delta t} \left(g_k(t_n) - \Delta t \int_{\Gamma_{out}} u(t_n-t_f)|_{\Gamma_{out}} \varphi_k|_{\Gamma_{out}} e^{-\lambda_k t_n-t_f} d\Gamma \right) \\
 &\quad + \Delta t \int_{\Gamma_{out}} u(t_{n+1})|_{\Gamma_{out}} \varphi_k|_{\Gamma_{out}} e^{-\lambda_k t_f} d\Gamma.
 \end{aligned} \tag{4.16}$$

Unfortunately, the recursive formula (4.16) becomes unstable for large n , so in the implementation, every n_s steps we recompute the coefficient $\mathbf{g}_k(t_n)$ using a classical quadrature formula for the integral in time between $[t_n - t_f, t_n]$.

Let us sum up the algorithm for the implementation of the state estimate:

Algorithm 2 State estimator

Require:

Pre-computation of the eigenfunctions φ_k solutions to (4.3) for $k \in \{0, \dots, N\}$

Pre-computation and factorization of the Gramian matrix \mathcal{G} defined in (4.15)

while $t_n \leq T$ **do**

if $n \equiv 0 \pmod{n_s}$ **then**

 Compute each component of $\mathbf{g}(t_n)$ using a quadrature formula

else

 Update each component of $\mathbf{g}(t_n)$ using equation (4.16)

end if

Solve the linear system (4.14)

Compute the state estimate $\hat{u}(t_n, \cdot)$ using equation (4.12)

Update time: $t_n \leftarrow t_n + \Delta t$

end while

4.3. Error estimate

Now, assuming we can solve the linear system (4.14) (i.e. the functions $(\varphi_k|_{\Gamma_{out}} e^{-\lambda_k(\cdot-t_f)})_{k \in \{1, \dots, N\}}$ are linearly independent in $L^2((0, T) \times \Gamma_{out})$), a natural question is: can we have an error bound between \hat{u} and u in Ω at time t ? Indeed, since the family $(\varphi_k|_{\Gamma_{out}} e^{-\lambda_k(\cdot-t_f)})_{k \in \mathbb{N}}$ is not an orthonormal basis, we cannot a priori ensure that the coefficients $\hat{A}_k(t - t_f)$ computed by the resolution of (4.13) are equal (or close) to the coefficients $A_k(t - t_f)$ of the decomposition of u at time $t_0 = t - t_f$.

Let us define by u^N the truncated series:

$$u^N(s, x, y) = \sum_{k=0}^N A_k(t - t_f) \varphi_k(x, y) e^{-\lambda_k(s-t_f)}, \quad \forall s \in [t - t_f, t]. \tag{4.17}$$

We denote by \mathbf{g}^N the associated vector which components are defined by

$$\mathbf{g}_k^N = \left(u^N|_{\Gamma_{out}}, \varphi_k e^{-\lambda_k(\cdot - t + t_f)} \right)_{L^2([t-t_f, t] \times \Gamma_{out})}.$$

The difficulty comes from the fact that we do not have $\mathbf{g}^N = \mathbf{g}$. Yet, the vector $\mathbf{A}(t - t_f)$ which components are given by $A_k(t - t_f)$ for $k \in \{1, \dots, N\}$ is solution to

$$\mathcal{G}\mathbf{A}(t - t_f) = \mathbf{g}^N,$$

so, we can deduce that

$$\|\hat{\mathbf{A}}(t - t_f) - \mathbf{A}(t - t_f)\|_{\mathbb{R}^N} \leq \|\mathcal{G}^{-1}\| \|\mathbf{g} - \mathbf{g}^N\|_{\mathbb{R}^N}, \quad (4.18)$$

where we consider the classical ℓ^1 -norm in \mathbb{R}^N . Now, using the orthogonality of the functions $(\varphi_k)_k$ with respect to the scalar product $(\cdot, \cdot)_\zeta$ we have

$$\begin{aligned} \|\hat{u}(t, \cdot) - u^N(t, \cdot)\|_\zeta &= \left\| \sum_{k=1}^N \left(\hat{A}_k(t - t_f) - A_k(t - t_f) \right) e^{-\lambda_k t_f} \varphi_k \right\|_\zeta, \\ &= \sum_{k=1}^N |\hat{A}_k(t - t_f) - A_k(t - t_f)| e^{-\lambda_k t_f} \leq \|\hat{\mathbf{A}}(t - t_f) - \mathbf{A}(t - t_f)\|_{\mathbb{R}^N}. \end{aligned} \quad (4.19)$$

Using Cauchy–Schwarz inequality, we can bound the term $\|\mathbf{g} - \mathbf{g}^N\|_{\mathbb{R}^N}$ as follows:

$$\begin{aligned} \|\mathbf{g} - \mathbf{g}^N\|_{\mathbb{R}^N} &= \sum_{k=1}^N \left| \left(u(s, \cdot)|_{\Gamma_{out}} - u^N(s, \cdot)|_{\Gamma_{out}}, \varphi_k e^{-\lambda_k(s - t + t_f)} \right)_{L^2((t-t_f, t) \times \Gamma_{out})} \right|, \\ &\leq \|u(s, \cdot)|_{\Gamma_{out}} - u^N(s, \cdot)|_{\Gamma_{out}}\|_{L^2((t-t_f, t) \times \Gamma_{out})} \underbrace{\sum_{k=1}^N \|\varphi_k e^{-\lambda_k(s - t + t_f)}\|_{L^2((t-t_f, t) \times \Gamma_{out})}}_{=C}, \\ &\leq C \|u(t, \cdot)|_{\Gamma_{out}} - u^N(s, \cdot)|_{\Gamma_{out}}\|_{L^2((t-t_f, t) \times \Gamma_{out})}, \end{aligned}$$

where we can note that C is uniformly bounded in t since $C \leq \sqrt{t_f} \sum_{k=1}^N \|\varphi_k\|_{L^2(\Gamma_{out})}$ ⁶. We deduce then that:

$$\|\hat{u}(t, \cdot) - u^N(t, \cdot)\|_\zeta \leq C \|\mathcal{G}^{-1}\| \|u|_{\Gamma_{out}} - u^N|_{\Gamma_{out}}\|_{L^2((t-t_f, t) \times \Gamma_{out})}. \quad (4.20)$$

and thus:

$$\begin{aligned} \|\hat{u}(t, \cdot) - u(t, \cdot)\|_\zeta &\leq \|\hat{u}(t, \cdot) - u^N(t, \cdot)\|_\zeta + \|u^N(t, \cdot) - u(t, \cdot)\|_\zeta, \\ &\leq C \|\mathcal{G}^{-1}\| \|u|_{\Gamma_{out}} - u^N|_{\Gamma_{out}}\|_{L^2((t-t_f, t) \times \Gamma_{out})} + \|u^N(t, \cdot) - u(t, \cdot)\|_\zeta. \end{aligned} \quad (4.21)$$

In the next, the constant C may vary from one line to another. The main point is that it does not depend on t . On the one hand, we have using (4.7) with $t_0 = T_0$ and the properties of the functions $(\varphi_k)_k$:

$$\begin{aligned} \|u^N(t, \cdot) - u(t, \cdot)\|_\zeta &= \left\| \sum_{k \geq N+1} e^{-\lambda_k(t - T_0)} (u(T_0, \cdot), \varphi_k)_\zeta \varphi_k \right\|_\zeta \\ &\leq e^{-\lambda_{N+1}(t - T_0)} \|u(T_0, \cdot)\|_\zeta, \end{aligned} \quad (4.22)$$

⁶Yet, we can already note that C grows with N .

where we recall that T_0 is the extinction time. On other hand, we have using the trace theorem and the equivalence between the H_ζ^1 -norm with the classical H^1 -norm:

$$\begin{aligned} \|u|_{\Gamma_{out}} - u^N|_{\Gamma_{out}}\|_{L^2((t-t_f,t)\times\Gamma_{out})}^2 &= \int_{t-t_f}^t \|u(s,\cdot)|_{\Gamma_{out}} - u^N(s,\cdot)|_{\Gamma_{out}}\|_{L^2(\Gamma_{out})}^2 ds, \\ &\leq \int_{t-t_f}^t \|u(s,\cdot) - u^N(s,\cdot)\|_{H^1(\Omega)}^2 ds, \\ &\leq C \int_{t-t_f}^t \|u(s,\cdot) - u^N(s,\cdot)\|_{H_\zeta^1}^2 ds, \end{aligned} \quad (4.23)$$

where the constant C comes from the equivalence between the two norms. Now, using (4.10) we easily get that:

$$\begin{aligned} \|u|_{\Gamma_{out}} - u^N|_{\Gamma_{out}}\|_{L^2((t-t_f,t)\times\Gamma_{out})}^2 &\leq C \int_{t-t_f}^t \left\| \sum_{k \geq N+1} e^{-\lambda_k(s-T_0)} A_k(T_0) \varphi_k \right\|_{H_\zeta^1}^2 ds, \\ &\leq C \int_{t-t_f}^t e^{-2\lambda_{N+1}(s-T_0)} \sum_{k \geq 0} A_k(T_0)^2 (1 + \lambda_k) ds, \\ &\leq C t_f e^{-2\lambda_{N+1}(t-(t_f+T_0))} \|u(t_0,\cdot)\|_{H_\zeta^1}^2. \end{aligned} \quad (4.24)$$

This leads finally to the following result:

Theorem 4.5 (Error bound). *Under the condition (4.2), the state estimator $\hat{u}(t,\cdot)$ defined in (4.7) satisfies*

$$\|\hat{u}(t,\cdot) - u(t,\cdot)\|_\zeta \leq \|u(t_0,\cdot)\|_\zeta e^{-\lambda_{N+1}(t-T_0)} + C \sqrt{t_f} e^{-\lambda_{N+1}(t-(t_f+T_0))} \|u(T_0,\cdot)\|_{H_\zeta^1}, \quad (4.25)$$

which tends to 0 as $t \rightarrow +\infty$.

According to this result, we can notice that the estimator \hat{u} converge to u for t larger that $t_f + T_0$. This means that t_f should not be too large and if possible, we will take it small. Yet, t_f cannot be too small otherwise the matrix \mathcal{G} is very ill conditioned, as illustrated hereafter. Similarly, N cannot be too large otherwise \mathcal{G} is also very ill conditioned

5. The instructive case of a rectangle river

In this section, let us consider the case where

$$\Omega = [0, L] \times [0, h], \quad \Gamma_{in} = \{0\} \times [0, h] \quad \text{and} \quad \Gamma_{out} = \{L\} \times [0, h], \quad (5.1)$$

with $L > 0$ and $h > 0$. We will consider also a uniform uniaxial flow and a constant diagonal tensor D , that is to say:

$$\mathbf{V} = \begin{bmatrix} V_x \\ 0 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} d_{xx} & 0 \\ 0 & d_{yy} \end{bmatrix} \quad (5.2)$$

where $V_x > 0$, $d_{xx} > 0$ and $d_{yy} > 0$. Let us remark that in that case, condition (4.2) is satisfied so we will be able to use the state estimator we developed.

In this configuration, let us now describe how to compute the adjoint functions w_A and w_B using separation of variables method.

5.1. The longitudinal adjoint function

To recover the x -component of \mathbf{s} , the best situation is to have an adjoint function w_A only depending on x so that its level sets are simply vertical lines. Now, if $w_A(x, y) = w_A(x)$, the boundary conditions on Γ are satisfied and problem (3.6) rewrites as follows:

$$\begin{cases} -d_{xx}\partial_{xx}^2 w_A - V_x \partial_x w_A = 0 & \text{in }]0, L[\\ w_A = 1 & \text{on } \{x = L\} \\ -\beta d_{xx} \partial_x w_A + (1 - \beta) w_A = 0 & \text{on } \{x = 0\} \end{cases} \quad (5.3)$$

where we took $g_{out} = 1$ (if g_{out} is not constant, then w_A depends on y). This equation is a simple linear ODE of order 2 which general solution is given by

$$w_A(x) = K_1 + K_2 e^{-\frac{V_x}{d_{xx}}x} \quad (5.4)$$

where the coefficients K_1 and K_2 are uniquely determined by the boundary conditions. More precisely, if $\beta = 1$ we get $K_1 = 1$ and $K_2 = 0$ which is an invariant solution also in x and therefore cannot be use for solving equation (3.12). If $\beta = 0$ then we get:

$$w_A(x) = \frac{1 - e^{-P_e \frac{x}{L}}}{1 - e^{-P_e}}, \quad (5.5)$$

where P_e is the so-called Péclet number

$$P_e = \frac{LV_x}{d_{xx}} > 0. \quad (5.6)$$

This number gives the ratio between the transport due to advection and the transport due to diffusion. With this expression, we can remark that:

- If P_e is small, then w_A “slowly” increases from 0 to 1 as x goes from 0 to L , see Figure 5.1. In particular, since the function is strictly increasing, this ensure that the first equation of (3.12) has a unique solution. Also, by the mean value theorem we easily show that

$$|x_1 - x_2| \leq L \frac{e^{P_e} - 1}{P_e} |w_A(x_1) - w_A(x_2)|, \quad \forall (x_1, x_2) \in [0, L]^2. \quad (5.7)$$

This stability result shows that for close values of $w_A(x_1)$ and $w_A(x_2)$, the arguments x_1 and x_2 should be close (since P_e is small).

- If P_e is large, then w_A is almost constant equal to 1 and sharply increases for x close to Γ_{in} , see Figure 5.1. Once again, the function is strictly increasing so we are ensure that the first equation of (3.12) has a unique solution. Yet, this time, inequality (5.7) shows that although $w_A(x_1)$ and $w_A(x_2)$ might be close, x_1 and x_2 can be very distant. In other words, recovering the position s_x of the source is not easy and require extremely precise measures, which are usually not affordable.

Remark 5.1. Note that the result for large P_e is coherent with the fact that in absence of diffusion (that is to say $P_e \rightarrow +\infty$ since $d_{xx} \rightarrow 0$), we cannot recover the position s_x . Indeed, in this case with constant parameters \mathbf{V} and D , the solution to (2.1) is given by $u(t, x, y) = u_0(x - tV_x, y)$ where $u_0(x, y) = u(T_0, x, y)$ for $t \geq T_0$, T_0 being the time of extinction of the source. Therefore, we cannot distinguish two source terms shifted in time and space which would lead to the same measurements since $u_0(x - tV_x, y) = u_0((x - t_0V_x) - (t - t_0)V_x, y)$. On the contrary, if $P_e \rightarrow 0$, we can easily show that w_A tends to the linear function $\frac{x}{L}$ which is the best situation for solving (3.12) and recovering s_x .

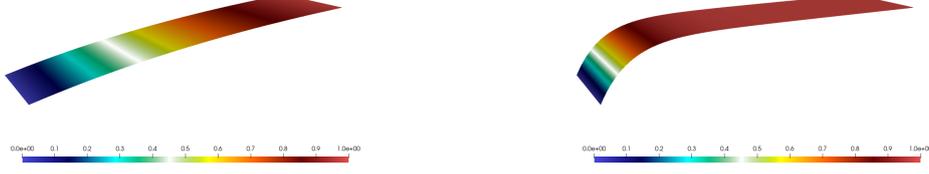


FIGURE 5.1. On the left, the function $w_A(x, y)$ with a low Péclet number $P_e = 1.2$ and on the right, a higher Péclet number $P_e = 12$.

The above remarks can be used to determine the maximum distance between Γ_{in} and Γ_{out} one can consider to still be able to identify s_x with w_A given the velocity \mathbf{V} and the diffusion tensor D . For instance, if we have $V_x = 1$, $d_{xx} = 1$ and we wish to determine the exact position s_x up to a precision of 10^{-1} being able to generate measurements with a precision of 10^{-5} , then L should be small enough to ensure

$$\frac{e^L - 1}{1} 10^{-5} \leq 10^{-1} \iff e^L \leq 1 + 10^4$$

which gives approximatively $L \leq 9.21$. In most interesting cases, as we will see, since the Péclet number will grow with the length L of the river, this adjoint function will in fact not be useful.

5.2. The transverse adjoint function

Now, the purpose is to determine the y -component of \mathbf{s} . The most favorable situation would be to have $w_B(x, y) = w_B(y)$ so that the level sets would be horizontal lines. Yet, we easily show that such solutions are necessarily constant solutions. The idea then is to look for separate variables solutions $w_B(x, y) = \tau(y)\chi(x)$ which leads to

Lemma 5.2. *Separate variables solutions to (3.6) in the rectangular domain Ω defined by (5.1) are given by*

$$w_{k+1} = \cos\left(k\pi\frac{y}{h}\right)\chi_k(x), \quad k \in \mathbb{N}, \quad (5.8)$$

where

$$\chi_k(x) = K_1 e^{r_1(k)\frac{x}{L}} + K_2 e^{r_2(k)\frac{x}{L}}, \quad \begin{cases} r_1(k) = -\frac{P_e}{2} + \frac{\sqrt{P_e^2 + (2\frac{L}{h}k\pi)^2 \frac{d_{yy}}{d_{xx}}}}{2}, \\ r_2(k) = -\frac{P_e}{2} - \frac{\sqrt{P_e^2 + (2\frac{L}{h}k\pi)^2 \frac{d_{yy}}{d_{xx}}}}{2}, \end{cases} \quad (5.9)$$

and the constant K_1 and K_2 are determined by the boundary conditions.

Proof. When looking for solutions of the form $w_B(x, y) = \tau(y)\chi(x)$, we easily get that $\tau(y) = \tau_k(y) = \cos(k\pi\frac{y}{h})$, $k \in \mathbb{N}$, and χ_k must satisfy

$$\begin{cases} -d_{xx}\partial_{xx}\chi_k - V_x\partial_x\chi_k + d_{yy}(\frac{k\pi}{h})^2\chi_k = 0 & \text{in }]0, L[, \\ \chi_k = 1 & \text{on } \{x = L\}, \\ \beta\partial_\nu\chi_k + (1 - \beta)\chi_k = 0 & \text{on } \{x = 0\}. \end{cases} \quad (5.10)$$

Solving this second order ODE leads to the expected result. ■

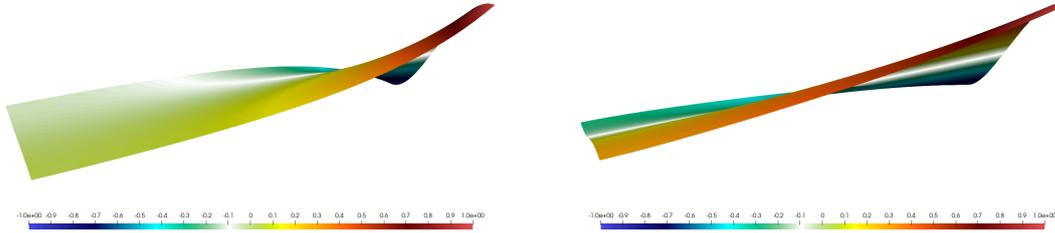


FIGURE 5.2. On the left, the function $w_B = \chi_1(x) \cos(\pi \frac{y}{h})$ with a low Péclet number ($P_e = 1.2$) and on the right the same function with a higher Péclet number ($P_e = 12$).

Remark 5.3. Let us remark that for $k = 0$, we recognize equation (5.3) and the first adjoint function w_A .

In the expression (5.9), we can observe that $r_1(k) \geq 0$ whereas $r_2(k) < 0$ and also $|r_1(k)| \leq |r_2(k)|$. For $k = 1$, the exponential term with $r_1(1)$ will “slowly” increase from 0 to L . The idea then is to choose the parameter β of the boundary condition in order to get K_2 close to 0. A simple way to do so is to take $\beta = 1$ so that we impose an homogeneous Neumann condition on Γ_{in} and we get:

$$\chi_1(x) = \frac{r_2(1)}{e^{r_1(1)r_2(1)} - e^{r_2(1)r_1(1)}} e^{r_1(1)\frac{x}{L}} - \frac{r_1(1)}{e^{r_1(1)r_2(1)} - e^{r_2(1)r_1(1)}} e^{r_2(1)\frac{x}{L}}.$$

In Figure 5.2, we represent the corresponding adjoint function $w_B = \chi_1(x) \cos(\pi \frac{y}{h})$ for two Péclet numbers. As we can see, this time the larger the Péclet number is, the better the situation is for solving (3.12). Indeed, if $P_e \rightarrow +\infty$ and L remains constant, simple computations shows that the function χ_1 tends to 1, so that w_B tends to $\cos(\pi \frac{y}{h})$. This is the opposite situation as for the previous adjoint function w_A ! So, if the advection dominates, w_B is a “good” adjoint function to determine s_y .

Remark 5.4. Note that one could choose β in the Robin condition on Γ_{in} to get exactly $K_2 = 0$. Yet, this is possible only because we know the analytical expression of the solution. In the general case, this will not be possible anymore, this is why we consider here the choice $\beta = 1$ that gives always an approximation of an open domain on Γ_{in} .

For this particular example, we can easily show

Proposition 5.5 (Identifiability). *Taking $w_A = w_1$ and $w_B = w_2$ where w_k is defined in (5.8), we uniquely identify the position \mathbf{s} solution to (3.12).*

Proof. As previously explained, the position s_x of the estimate source is uniquely determined because w_A only depends on x and is monotonous. Then, since $w_B(s_x, \cdot)$ is also monotonous in y , we uniquely determine the position s_y . ■

In more general situations, we may wonder if it is worth considering other adjoint functions for $k \geq 1$

$$w_{k+1} = \left(\frac{r_2(k)}{e^{r_1(k)r_2(k)} - e^{r_2(k)r_1(k)}} e^{r_1(k)\frac{x}{L}} - \frac{r_1(k)}{e^{r_1(k)r_2(k)} - e^{r_2(k)r_1(k)}} e^{r_2(k)\frac{x}{L}} \right) \cos \left(k\pi \frac{y}{h} \right), \quad (5.11)$$

solutions to (3.6) with $g_{out} = \cos(k\pi \frac{y}{h})$ and $\beta = 1$. In fact, this is necessary as soon as the Péclet number is a bit high and the first adjoint function $w_A = w_1$ increases too sharply, so it cannot be used numerically for solving equation (3.12). This will be illustrated in the next section.

Remark 5.6. Let us remark that if we use $(w_A, w_B) = (w_2, w_3)$ for the adjoint functions, we still have identifiability in the case of the rectangular river.

5.3. An example of source localization

To conclude on this particular case of a rectangular river, let us show an example of source identification. We will consider the following parameters:

$$\Omega = [0, 300] \text{ m} \times [0, 100] \text{ m}, \quad \mathbf{V} = \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} \text{ m.s}^{-1} \quad \text{and} \quad D = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \text{ m}^2.\text{s}^{-1}.$$

We consider a source at position $\mathbf{s} = (150, 25) \text{ m}$ and the function λ is defined by

$$\lambda(t) = \begin{cases} 100 \sin(\frac{\pi t}{180}) & \text{if } t \leq 180 \text{ s,} \\ 0 & \text{if not.} \end{cases}$$

The synthetic data on $\Gamma_{in} / \Gamma_{out}$ are recorded during time $[0, T]$ with $T = 3000 \text{ s}$ at interval $\Delta t = 1 \text{ s}$.

In Figure 5.3, we have represented at different time snapshots of the solution u in Ω on the left, the estimator \hat{u} defined by (4.12) (using $N = 5$ and $t_f = 400 \text{ s}$)⁷ and the sets $\mathcal{S}_A(t)$ (in orange), $\mathcal{S}_B(t)$ (in purple) and their intersection $\mathcal{S}_A(t) \cap \mathcal{S}_B(t)$ (in red). The estimated source position $\hat{\mathbf{s}}(t) = (\hat{x}_s(t), \hat{y}_s(t))$ is represented by the red cross. The results are presented for three choices of coupled of adjoint functions. As we can see for small t no estimate is computed. Indeed, this corresponds to the case where no (sufficiently enough) pollutant has reached the boundary Γ_{out} (the data on Γ_{in} are neglected) and therefore $|\mathcal{B}(u, 1, t) + (\hat{u}(t, \cdot), 1)_{L^2(\Omega)}|$ is almost null. As t grows, the estimation is performed and we can see that for the couple of adjoint functions (w_2, w_3) , it converges to the true position (the errors come from the discretization in space and time (numerical quadrature), and from the fact that the data are generated on a different mesh from the one use for the adjoint functions and the estimator). For the two other couples, using w_1 as adjoint function, we see that the algorithm fails to converge. As explained before, this is due to the fact that in that case, this adjoint function is not adapted because the Péclet number is too big.

Now, to evaluate the interest of the state estimator, we have represented in Figure 5.4 the evolution of the error $\|\hat{\mathbf{s}}(t) - \mathbf{s}\|_{\mathbb{R}^2}/L$ (where $L = 300 \text{ m}$ is the length of the domain and $\mathbf{s} = (125, 25) \text{ m}$ is the true source position) versus time when using or not the state estimator (not using it corresponds to take $\hat{u} = 0$). These results correspond to the case where we take for the adjoint functions $(w_A, w_B) = (w_2, w_3)$. As we can see, using the estimator improves the speed of convergence of the method and therefore allows a faster identification procedure (this is particularly true for $N = 1$). At the end, since we almost have $\hat{u} \simeq u (\simeq 0)$, the two curves are superposed. Also, we can observe that the number of modes N has a strong influence on the result. In particular, the best situation is obtained when using only one mode. Note that for this case $N = 1$, we can take t_f small since it has no influence on the condition number of \mathcal{G} . Let us also underline that the state estimator can deteriorate the estimate of the source position when the elapsed time is too short (see around $t = 500 \text{ s}$ for $N = 5$ or $N = 7$). A simple way to check if the state estimator has a chance to give a good estimate (and is worth to be use) consists in computing the differences between $\hat{u}(t, \cdot)|_{\Gamma_{out}}$ and $u(t, \cdot)|_{\Gamma_{out}}$. This leads to the following criterion:

Definition 5.7 (Criterion to use the state estimator). We use the state estimator at each time t for which the following condition is satisfied:

$$\frac{\|\hat{u}(t, \cdot) - u(t, \cdot)\|_{L^2(\Gamma_{out})}}{\|u(t, \cdot)\|_{L^2(\Gamma_{out})}} \leq 1. \quad (5.12)$$

⁷Although in this case we could compute the eigenfunctions of the problem (4.3) analytically, we have computed them numerically.

A. TONNOIR

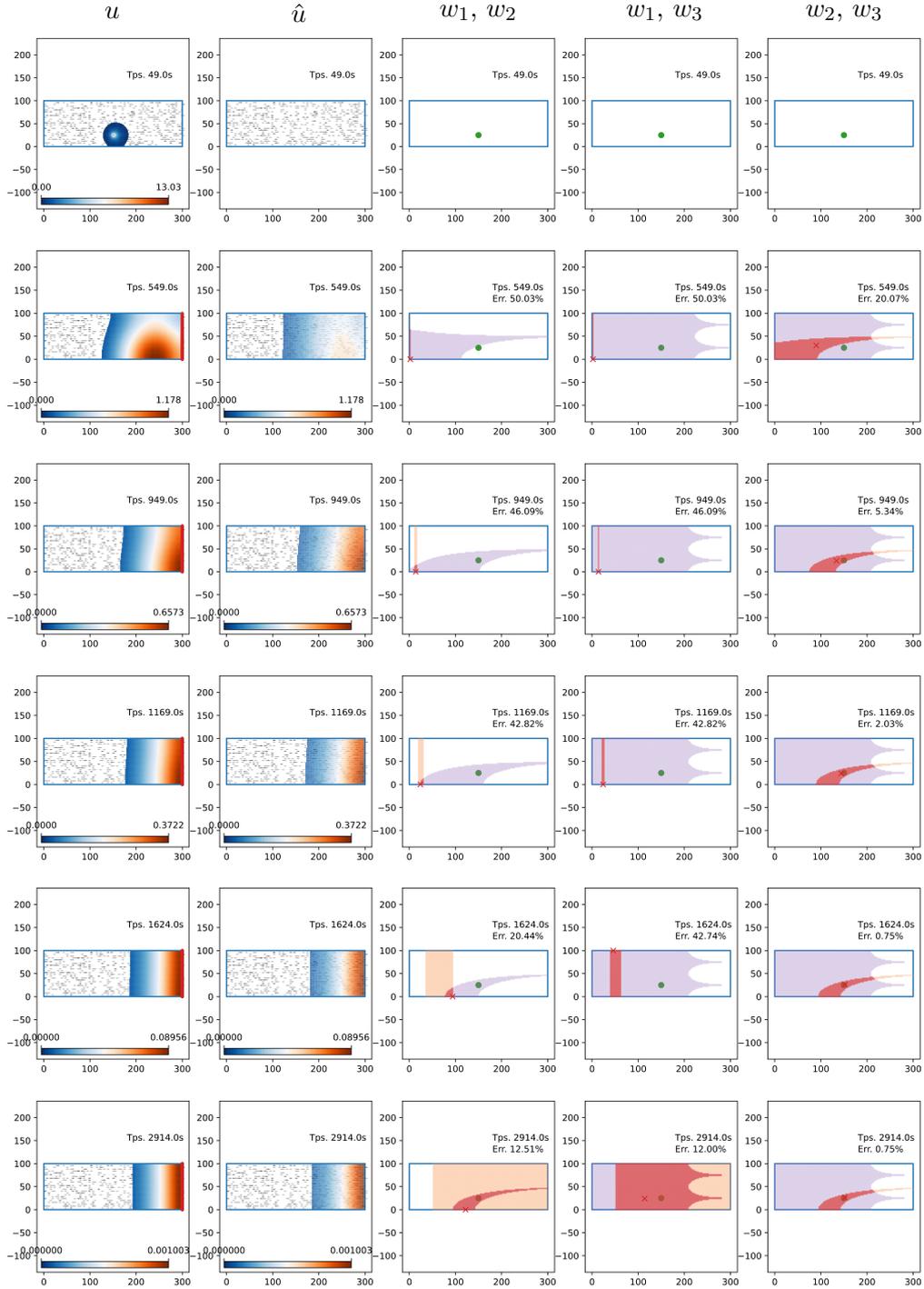


FIGURE 5.3. On the left, the solution of the direct problem at times $t \in \{49, 549, 949, 1169, 1624, 2914\}$ s (from top to bottom). In the second column, the estimator \hat{u} at the same time. In the columns 3 to 5, the set $\mathcal{S}_A(t)$ (in orange), the set $\mathcal{S}_B(t)$ (in purple) and the intersection $\mathcal{S}_A(t) \cap \mathcal{S}_B(t)$ (in red) for different couples of adjoint functions w_A, w_B (from left to right: $(w_A, w_B) = (w_1, w_2)$, $(w_A, w_B) = (w_1, w_3)$ and $(w_A, w_B) = (w_2, w_3)$). The estimate of the source position is the red cross and the green dot is the true position of the source.

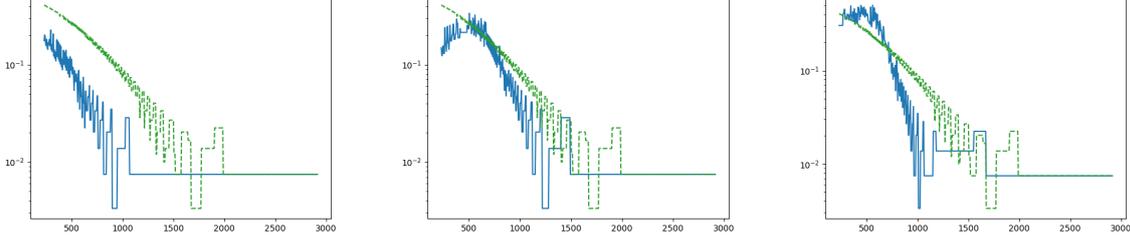


FIGURE 5.4. Error on the source position estimate (log scale) versus time using w_2/w_3 adjoint functions. In green: without the state estimator and in blue: with the state estimator. From left to right: $N = 1$ ($t_f = 10$ s), $N = 5$ ($t_f = 400$ s) and $N = 7$ ($t_f = 400$ s)

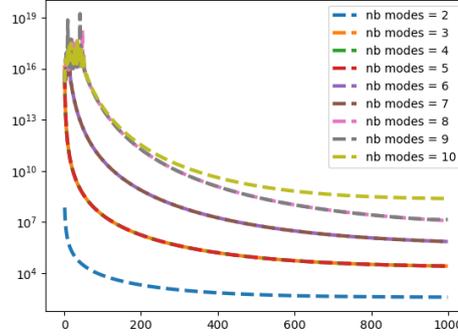


FIGURE 5.5. Condition number of the Gramian matrix \mathcal{G} for $N = \{2, \dots, 10\}$ versus t_f .

This criterion will be used in the numerical examples in the next section. This corresponds to use the state estimator \hat{u} whenever it is closer to u on Γ_{out} than the null estimator. Let us emphasize that, a priori, even if \hat{u} is a better estimator than 0 on the exterior boundary, this does not imply that it is a better estimator in the whole domain Ω (in other words, $\|\hat{u} - u\|_{L^2(\Omega)}$ can be greater than $\|u\|_{L^2(\Omega)}$ even if $\|\hat{u}(t, \cdot) - u(t, \cdot)\|_{L^2(\Gamma_{out})} \leq \|u(t, \cdot)\|_{L^2(\Gamma_{out})}$).

To understand why increasing N can deteriorate the results, it is interesting to have a look to the Gramian matrix \mathcal{G} defined in (4.15). In Figure 5.5, we have represented the condition number of \mathcal{G} considering various number of modes versus t_f . As we can clearly see, the condition number blows up with N the number of modes. Also, increasing t_f reduces the condition number (even though for large N it remains very high) up to a plateau. These results can be explained why the best result in Figure 5.4 is obtained with $N = 1$.

Remark 5.8. In this particular case of a rectangular geometry, we can observe that the condition number is the same for different numbers of modes (for instance for $N = 2, 3, 4, 5$). This is explained by the fact that a part of the modes form an orthonormal basis on the exterior boundary Γ_{out} .

6. Some extensions

In this section, we will explore numerically various extensions of the proposed method.

6.1. A non straight river

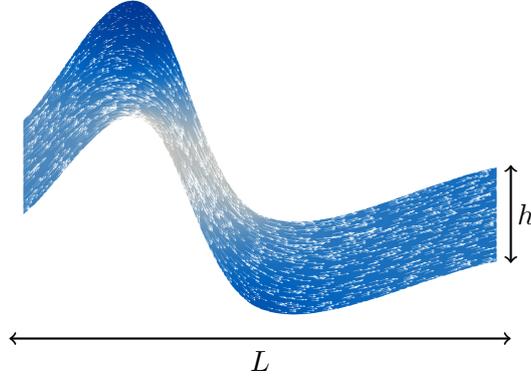


FIGURE 6.1. Illustration of the potential flow (the white arrows correspond to the vector \mathbf{V} and the background color to the norm of \mathbf{V}) in a more realistic geometry.

Let us consider the case of a non rectangular river, as illustrated in Figure 6.1. This time, the adjoint functions w_i cannot be computed analytically and one has to solve numerically the following problems (h corresponds to the river width in Γ_{out}):

$$\left\{ \begin{array}{ll} -\operatorname{div}(D\nabla w_1) - \mathbf{V} \cdot \nabla w_1 = 0 & \text{in } \Omega, \\ D\nabla w_1 \cdot \nu = 0 & \text{on } \Gamma, \\ w_1 = 1 & \text{on } \Gamma_{out}, \\ w_1 = 0 & \text{on } \Gamma_{in}, \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{ll} -\operatorname{div}(D\nabla w_{k+1}) - \mathbf{V} \cdot \nabla w_{k+1} = 0 & \text{in } \Omega, \\ D\nabla w_{k+1} \cdot \nu = 0 & \text{on } \Gamma, \\ w_{k+1} = \cos(k\pi \frac{y}{h}) & \text{on } \Gamma_{out}, \\ D\nabla w_{k+1} \cdot \nu = 0 & \text{on } \Gamma_{in}. \end{array} \right. \quad (6.1)$$

Let us remark that we took $g_{out} = \cos(k\pi \frac{y}{h})$ as in the rectangular case expecting similar properties: w_1 shall vary in the longitudinal direction (that is to say in the flow direction) whereas w_2 shall vary mainly in the transverse direction (that is to say in the direction orthogonal to the flow). Obviously, since the geometry is no more a rectangle, this cannot be perfectly true.

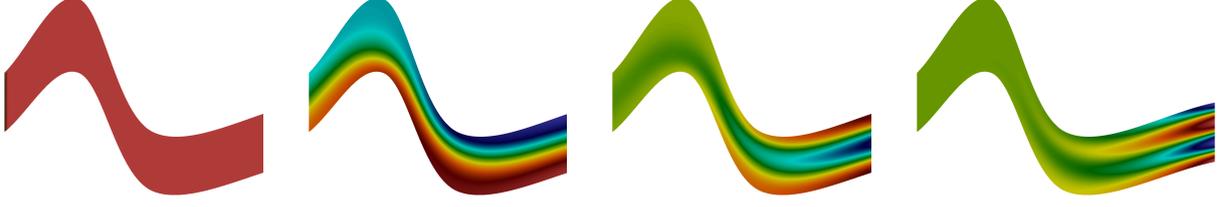
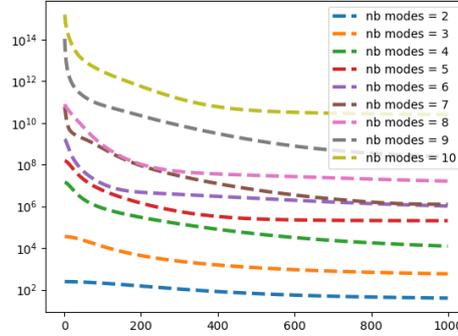
Remark 6.1. In the following examples, we will suppose that the boundary Γ_{out} is straight. For a curved boundary Γ_{out} , the boundary condition in (6.1) is given using a similar idea. For instance, if we consider a boundary Γ_{out} described by the parametric curve $\mathbf{p}(s) = (x(s), y(s))$, $s \in [0, 1]$, we would replace $\frac{y}{h}$ in the boundary condition by the projected coordinate $(\mathbf{p}(s) - \mathbf{p}(0)) \cdot (\mathbf{p}(1) - \mathbf{p}(0)) / (\|\mathbf{p}(s) - \mathbf{p}(0)\|_2 \|\mathbf{p}(1) - \mathbf{p}(0)\|_2)$ along the line $[\mathbf{p}(0), \mathbf{p}(1)]$. Let us emphasize that it has not been tested numerically.

For this new case, we consider the following parameters:

$$(L, h) = (500, 100) \text{ m}, \quad \mathbf{V} = \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} \text{ m.s}^{-1} \text{ on } \Gamma_{in} \cap \Gamma_{out} \quad \text{and} \quad D = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix} \text{ m}^2.\text{s}^{-1}.$$

Inside the domain, the velocity \mathbf{V} is computed solving a potential flow problem (which amounts to a Laplace equation $-\Delta P = 0$ and we have $\mathbf{V} = \nabla P$), see Figure 6.1.

In Figure 6.2, we have represented the adjoint functions w_1 and w_2 solutions to (6.1). As we can see, we cannot expect to recover the longitudinal component with w_1 since this function is almost constant equal to 1 and very sharply increases near Γ_{in} from 0 to 1. To be able to recover the longitudinal component of the source term position, the idea then is to compute w_k for $k \geq 2$. As we can see still in Figure 6.2, the obtained adjoint functions oscillate in the transverse direction and slowly increase (from


 FIGURE 6.2. Adjoint functions w_1 , w_2 , w_3 and w_4 from left to right.

 FIGURE 6.3. Condition number of the Gramian matrix \mathcal{G} for $N = \{2, \dots, 10\}$ versus t_f for the curved river.

left to right) in the longitudinal direction. This slow increase will allow us to recover the longitudinal component of the source term position.

In that case, we consider two sources located respectively at $(50, 125)$ m and $(400, 0)$ m and two associated functions:

$$\lambda_1(t) = \begin{cases} 100 \sin(\frac{\pi t}{180}) & \text{if } t \leq 180s, \\ 0 & \text{if not,} \end{cases} \quad \text{and} \quad \lambda_2(t) = \begin{cases} 100 \sin(\frac{\pi t}{120}) & \text{if } t \leq 120s, \\ 0 & \text{if not.} \end{cases}$$

The data are generated with these source terms solving the direct problem and are recorded with a time step $\Delta t = 1$ s during a period $T = 3600$ s. Here again, we only use the measurements on Γ_{out} and neglect the measurements on Γ_{in} .

For the estimator \hat{u} , the number of modes N is chosen similarly to the straight river case, that is to say $N = 1$ and $t_f = 10$ s. Here again, the condition number of the Gramian matrix \mathcal{G} blows up with N as represented in Figure 6.3. In our numerical tests, we also tried greater values of N but the best results have been obtained with $N = 1$.

To be able to dissociate the two sources in the estimate, we study the evolution in time of $\mathcal{B}(u, 1, t) + (\hat{u}(t, \cdot), 1)_{L^2(\Omega)}$ which gives an estimate of $\Lambda(t)$, see Figure 6.4. If this quantity is not zero and remains (almost) constant during a period $[t_1, t_2]$, we assume that the data between $[0, \frac{t_1+t_2}{2}]$ are explained by a first source and the data for $t \geq \frac{t_1+t_2}{2}$ by a second source (and similarly if more sources are

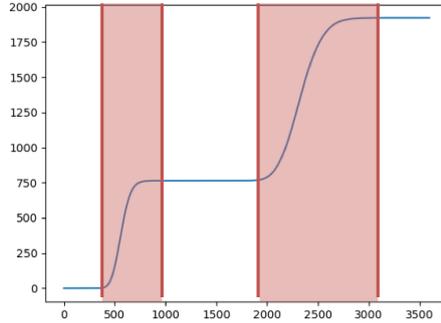


FIGURE 6.4. Evolution in time of the coefficient $\mathcal{B}(u, 1, t)$. The red band corresponds to the identification of one source.

considered). This allows to automatically detect if more than one source is present. In that case, the coefficients \mathcal{D}_A and \mathcal{D}_B (defined in (3.14)) are simply reset to 0 for estimating the position of the next source. Note that this procedure works only if the sources are “well-separated” in space or time. We mean by well separated the fact that during a sufficiently long time $t_2 - t_1 > 0$, the variations of $\mathcal{B}(u, 1, t) + (\hat{u}(t, \cdot), 1)_{L^2(\Omega)}$ induced by the two sources are smaller than a threshold.

In Figure 6.5, we have represented the solution of the direct problem and the reconstructed source term positions at different time. As before, we have represented the sets \mathcal{S}_i (defined in (3.15)), for different couples of adjoint functions. Here again, the data are generated without noise (except the use of different meshes and order of FE for the computations in the direct problem and for the computation of the adjoint functions). As we can see, we recover quite precisely the two source positions.

To end this test, we have represented on Figure 6.6 the error $\frac{\|\hat{\mathbf{s}}(t) - \mathbf{s}\|_{\mathbb{R}^2}}{L}$ (where $\hat{\mathbf{s}}(t)$ is the estimated position) versus time for the two sources localization considering the different couples of adjoint functions and using or not the state estimator. As we can clearly see, for the first source (located close to Γ_{out}), the state estimator works poorly and deteriorates the convergence. This is explained by the fact that the elapsed time between the source extinction and the measurement is too small. Nevertheless, we observe a very sharp convergence around $t = 850$ s since the null estimator replaces \hat{u} because condition (5.12) is not satisfied. On the contrary, for the second source located far from Γ_{out} , we can see that using the state estimator is interesting and the identification of the source is faster (at least for the adjoint functions (w_2, w_3) and (w_3, w_4)).

6.2. Robustness with respect to noisy data

To assess the robustness of the method with respect to noisy data, we have considered on the same example as before perturbed data. The data are perturbed by an additive noise proportional to the measurements. In Figure 6.7, we have represented for a noise of $\{0, 5, 10\}\%$ the evolution of the error $\frac{\|\hat{\mathbf{s}}(t) - \mathbf{s}\|_{\mathbb{R}^2}}{L}$ versus time. As we can see, the method is quite robust with respect to noisy data since the (relative) error stays relatively small. Also, we can observe that recovering a source far from Γ_{out} seems to be more accurate, which can be quite surprising. In fact, the diffusion effect helps to alleviate the effect of the noise.

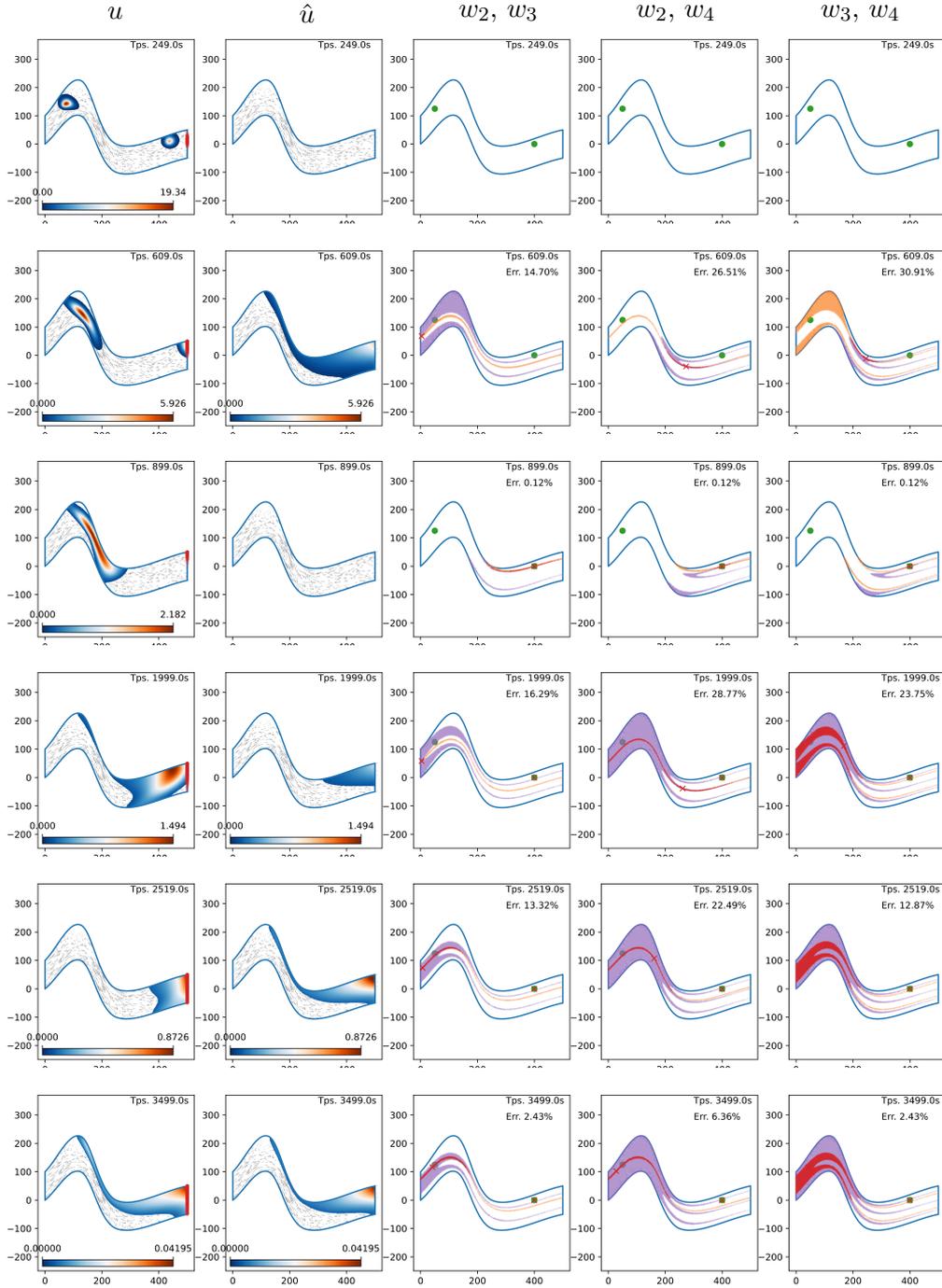


FIGURE 6.5. On the left, the solution of the direct problem at times $t \in \{249, 609, 899, 1999, 2519, 3499\}$ s (from top to bottom). In the second column, the estimator \hat{u} at the same time. In columns 3 to 5, the set $\mathcal{S}_A(t)$ (in orange), the set $\mathcal{S}_B(t)$ (in purple) and the intersection $\mathcal{S}_A(t) \cap \mathcal{S}_B(t)$ (in red) for different couples of adjoint functions (from left to right: (w_2, w_3) , (w_2, w_4) and (w_3, w_4)). The estimate of the source position (red cross). The green dot on the right pictures corresponds to the position of the sources.

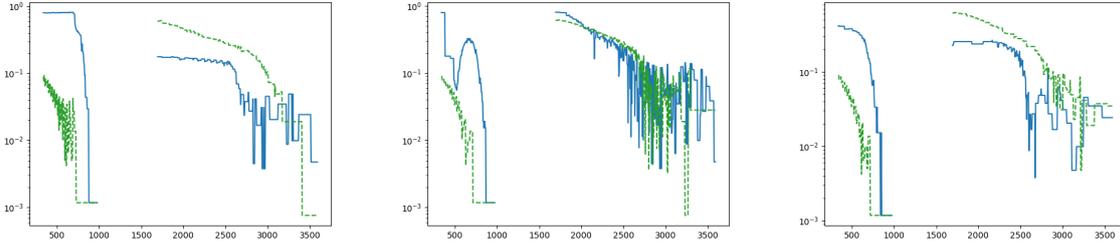


FIGURE 6.6. Error on the source position estimate (log scale) versus time using from left to right the adjoint functions: (w_2, w_3) , (w_2, w_4) and (w_3, w_4) . In green: without the state estimator and in blue: with the state estimator.

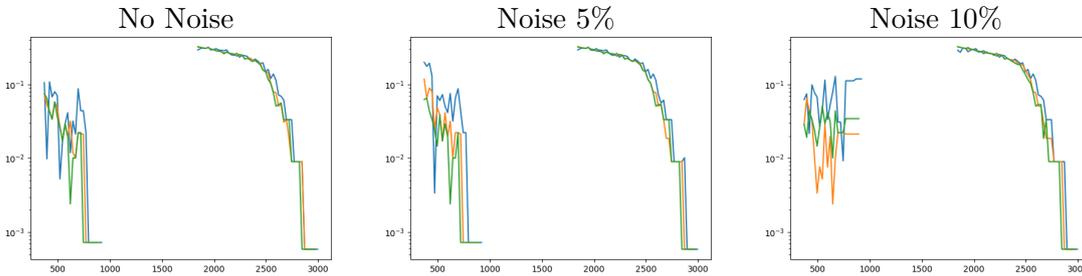


FIGURE 6.7. Evolution in time of the error (log scale) on the estimated source position considering adjoint functions (w_2, w_3) (in blue), (w_2, w_4) (in orange), (w_3, w_4) (in green). For time smaller than 1000 s, we identify the closest source, and for time greater than 1000 s, the second source.

6.3. A more realistic case

To conclude this section on numerical tests, let us consider a more realistic situation. We consider a section of the Seine river in Rouen, France, see Figure 6.8. The flow \mathbf{V} is computed numerically taking $V_{in} = V_{out} = 1 \text{ m.s}^{-1}$. For the diffusion tensor D , we consider the Bear’s hydrodynamic tensor [5, 6]

$$D = (D_m + \alpha_T \|\mathbf{V}\|_2) \text{Id} + \frac{\alpha_L - \alpha_T}{\|\mathbf{V}\|_2} \mathbf{V}\mathbf{V}^t, \quad (6.2)$$

where $D_m := 0.01 \text{ m}^2.\text{s}^{-1}$ is the molecular diffusion, $\alpha_T := 0.1 \text{ m}$ is the transverse diffusion and $\alpha_L := 1 \text{ m}$ is the longitudinal diffusion. In this case, condition (4.2) is not satisfied and we cannot construct the estimator \hat{u} defined in Section 4. As a consequence, we simply take $\hat{u} = 0$.

We consider two sources located as represented in Figure 6.8. In particular, we can see that one of the sources is quite far from the boundary Γ_{out} and in the south of an island (called “Ile Lacroix”). The data are once again synthetic data perturbed by a noise of 5% and we consider only measurements on Γ_{out} recorded with a time step $\Delta t = 10 \text{ s}$ during time $T = 50000 \text{ s}$.

In Figure 6.9, we have represented the reconstruction of the source term positions at different times. Here again, we only use the measurements on Γ_{out} . As we can see, we can recover very well the two source positions after enough time, in particular with the adjoint functions w_2 and w_3 . Let us emphasize that, as we can see in the direct simulation, a part of the pollutant does not cross the boundary Γ_{out} and remains inside the river because it stays in a region where $\mathbf{V} = \mathbf{0}$. This pollutant

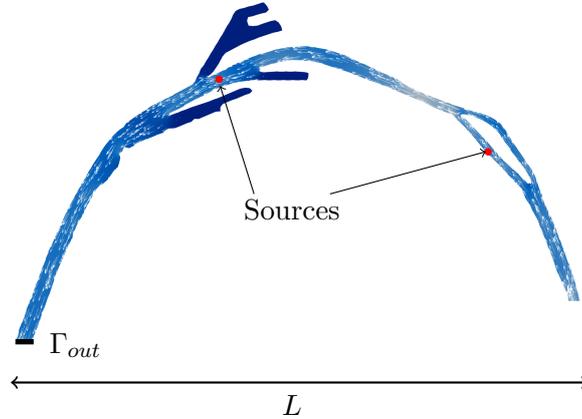


FIGURE 6.8. Potential flow on the Seine river in Rouen, France. The red dots represents the source localization.

is only dissipated by diffusion effect. This explains why the source reconstruction in this case cannot be perfect.

7. Conclusion and discussion

In this work, we have revisited the problem of source reconstruction using adjoint functions. We propose a new method to compute the adjoint functions that allows to consider arbitrary (physical) configurations. The idea is based on the analysis of the straight river case for which we are able to show easily the identifiability result and to have a physical intuition of the effect of the Péclet number on the resolution of the inverse problem. The generalization of this approach turns out to be efficient in the general case and allows to identify sources even far from the boundary Γ_{out} (for instance as in the Seine river case). Also, we have seen that it can be interesting to use several couples of adjoint functions to bring more information on the sources localization. Indeed, some couples of adjoint functions are more accurate for identifying a source located close to Γ_{out} or a source far from Γ_{out} .

The second main contribution of this work is to propose an “online” identification procedure using a state estimator \hat{u} which is cheap to compute and for which we prove the convergence to u (under some assumptions). We have shown that the state estimator can be interesting for accelerating the identification procedure (at least for sources far from the outer boundary). Unfortunately, the estimator suffers from two main limitations. First, it can only be constructed if condition (4.2) is satisfied. It could be interesting to build a more general state estimator valid for any tensor D . The second limitation comes from the parameters N (the number of modes) and t_f the frame time. Indeed, due to the condition number of the Gramian matrix, we cannot take N large. Also, t_f cannot be too large otherwise the state estimator cannot work until at least $t > t_0 + t_f$. This is problematic for sources close to Γ_{out} and it becomes even more problematic for sources that do not vanish. For this last case, it would be interesting to construct another estimator.

To conclude, let us discuss some last questions on the hypothesis made in this work:

- *What if $u|_{\Gamma_{in}} \neq 0$?* This situation corresponds to the idea of a source term located beyond the monitored area. If we suppose that we also have the measurements on Γ_{in} , then there is no problem. Indeed, by a simple lifting we can remove from the data on Γ_{out} the part of

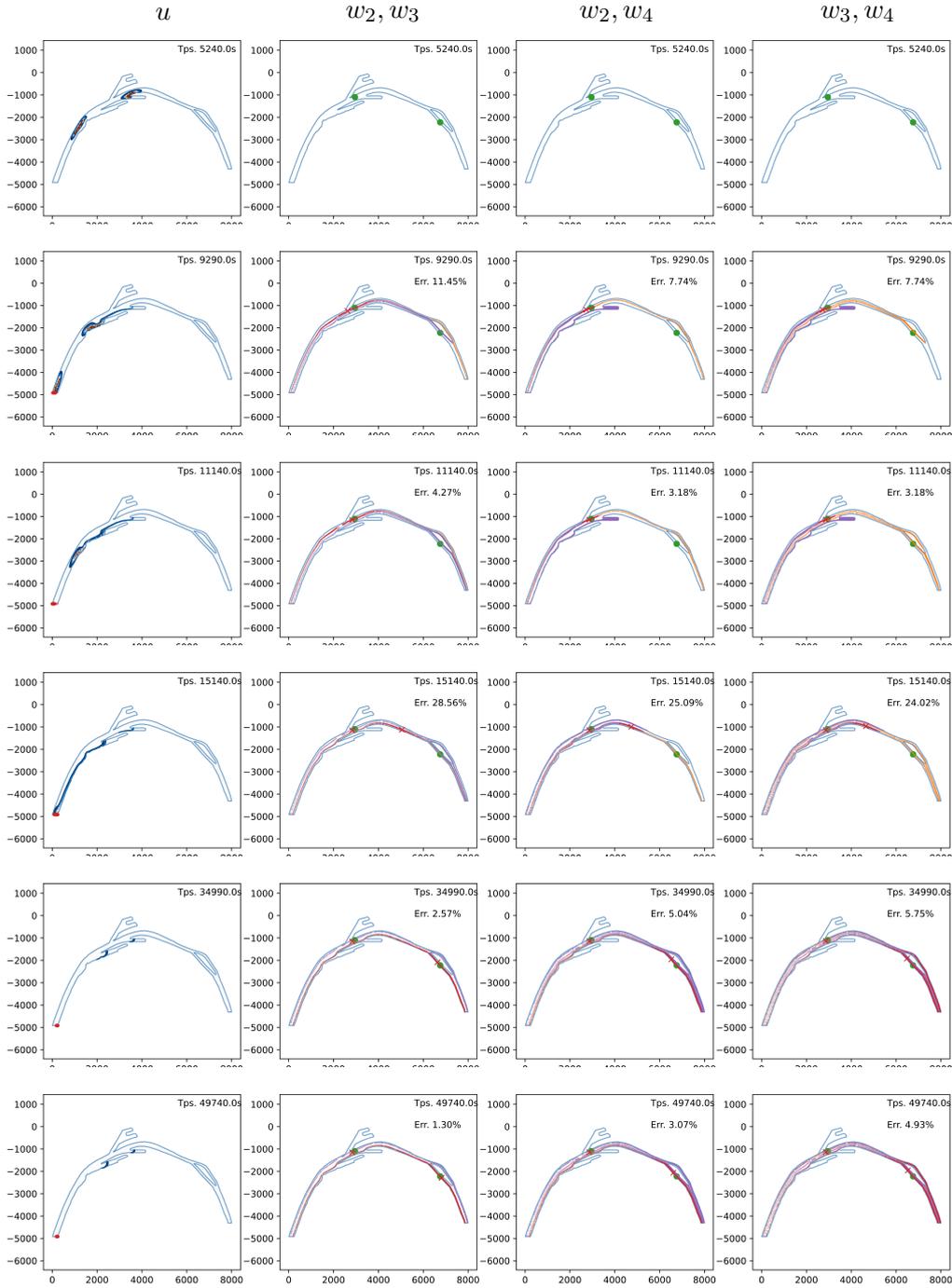


FIGURE 6.9. On the left, the solution of the direct problem at times $t \in \{5240, 9290, 11140, 15140, 34990, 49740\}$ s (from top to bottom). On the right, the set $\mathcal{S}_A(t)$ (in orange), the set $\mathcal{S}_B(t)$ (in purple) and the intersection $\mathcal{S}_A(t) \cap \mathcal{S}_B(t)$ (in red) for different couples of adjoint functions. The estimate of the source position (red cross). The green dot on the right pictures corresponds to the position of the sources.

the pollutant measured coming from the source outside the monitoring domain. If we do not measure u on Γ_{in} , as we wish to consider in this paper, then the reconstructed source position will give a position located on Γ_{in} . In fact, system (3.12) has no solution in that case and the best reconstruction the algorithm gives is a position on Γ_{in} .

- *What if $u(t = 0, \cdot) \neq 0$?* If at the initial time we suppose that there is already an unknown quantity of pollutant in the river, we cannot a priori identify the sources. Yet, if this pollutant comes from a point source that started to emit at time $t_- < 0$ and is such that the measurements of u on Γ_{out} during time $[t_-, 0]$ is negligible (in other words, the main part of the pollutant has not reach at time $t = 0$ the boundary Γ_{out}), then the proposed procedure works fine and we can identify the source position. Indeed, this simply corresponds to a shift in time where we have neglected the (small part of the) measurements of u on Γ_{out} during time $[t_-, 0]$.
- *What if two sources are too close?* In that case, the reconstructed position is between the two source positions. This was already observed in a similar context in [3]. In that case, it is difficult to be able to separate in the data the contribution of the two sources without any additional information.
- *What if the source is not a point source?* If the source has the form $f(t, x, y) = \lambda(t)g(x, y)$, then the method works still well if g has a narrow support as we can expect. If not, then this is similar to the case of close multiple sources.
- *What if the flow is not a potential flow?* In many case, the flow is not constant in time and / or not irrotational. For instance, for Stokes flow, \mathbf{V} is still constant in time but not irrotational. In that case, the methodology presented in this work could still apply but we expect the results to be deteriorated by the presence of vortex (numerical investigations are still in study). For the case of time varying \mathbf{V} , the construction of the adjoint functions presented here does not work anymore.

Acknowledgements

The author would like to thank the two anonymous referees for their useful comments and advices on the document.

References

- [1] Carlos Alves, Ana Leonor Silvestre, Takéo Takahashi, and Marius Tucsnak. Solving inverse source problems using observability. Applications to the Euler–Bernoulli plate equation. *SIAM J. Control Optim.*, 48(3):1632–1659, 2009.
- [2] Michael Andrieu, Fakher Ben Belgacem, and Abdellatif El Badia. Identification of moving pointwise sources in an advection–dispersion–reaction equation. *Inverse Probl.*, 27(2): article no. 025007 (21 pages), 2011.
- [3] Michael Andrieu and Abdellatif El Badia. On an inverse source problem for the heat equation. Application to a pollution detection problem, II. *Inverse Probl. Sci. Eng.*, 23(3):389–412, 2015.
- [4] Matthieu Aussal and Philippe Moireau. Kernel representation of Kalman observer and associated H-matrix based discretization. *ESAIM, Control Optim. Calc. Var.*, 28: article no. 78 (41 pages), 2022.
- [5] Jacob Bear and Yehuda Bachmat. *Introduction to modeling of transport phenomena in porous media*, volume 4 of *Theory and Applications of Transport in Porous Media*. Springer, 2012.
- [6] Jacob Bear and Alexander H.-D. Cheng. *Modeling groundwater flow and contaminant transport*, volume 23 of *Theory and Applications of Transport in Porous Media*. Springer, 2010.

- [7] Faker Ben Belgacem. Uniqueness for an ill-posed reaction-dispersion model. Application to organic pollution in stream-waters. *Inverse Probl. Imaging*, 6(2):163–181, 2012.
- [8] Muriel Boulakia, Maya de Buhan, Tiphaine Delaunay, Sébastien Imperiale, and Philippe Moireau. Solving inverse source wave problem from Carleman estimates to observer design. hal-04788439, version 1, 2024.
- [9] Tiphaine Delaunay, Sébastien Imperiale, and Philippe Moireau. Mathematical analysis of an observer for solving inverse source wave problem. hal-04344193, version 1, 2023.
- [10] Xiaomao Deng, Zi-Ju Liao, and Xiao-Chuan Cai. An efficient two-level overlapping domain decomposition method for recovering unsteady sources of 3D parabolic problems. *Comput. Math. Appl.*, 111:98–108, 2022.
- [11] Bruce A. Egan and James R. Mahoney. Numerical modeling of advection and diffusion of urban area source pollutants. *J. Appl. Meteorol. Climatol.*, 11(2):312–322, 1972.
- [12] Abdellatif El Badia and Tuong Ha-Duong. On an inverse source problem for the heat equation. Application to a pollution detection problem. *J. Inverse Ill-Posed Probl.*, 10(6):585–599, 2002.
- [13] Abdellatif El Badia, Tuong Ha-Duong, and Adel Hamdi. Identification of a point source in a linear advection–dispersion–reaction equation: application to a pollution source problem. *Inverse Probl.*, 21(3):1121–1136, 2005.
- [14] Abdellatif El Badia and Adel Hamdi. Inverse source problem in an advection–dispersion–reaction system: application to water pollution. *Inverse Probl.*, 23(5):2103–2120, 2007.
- [15] Adel Hamdi. Detection-Identification of multiple unknown time-dependent point sources in a 2 D transport equation: application to accidental pollution. *Inverse Probl. Sci. Eng.*, 25(10):1423–1447, 2017.
- [16] Adel Hamdi. A non-iterative method for identifying multiple unknown time-dependent sources compactly supported occurring in a 2d parabolic equation. *Inverse Probl. Sci. Eng.*, 26(5):744–772, 2018.
- [17] Adel Hamdi and Antoine Tonnoir. Dispersion–current adjoint functions for monitoring accidental sources in 3D transport equations. *Inverse Probl. Sci. Eng.*, 29(10):1445–1476, 2021.
- [18] Katrijn M. A. Holvoet, Piet Seuntjens, and Peter A. Vanrolleghem. Monitoring and modeling pesticide fate in surface waters at the catchment scale. *Ecol. Model.*, 209(1):53–64, 2007.
- [19] Cheng-Hung Huang, Jia-Xun Li, and Sin Kim. An inverse problem in estimating the strength of contaminant source for groundwater systems. *Appl. Math. Modelling*, 32(4):417–431, 2008.
- [20] Kartlos Kachiashvili, David Gordeziani, Raytcho Lazarov, and David Melikdzhanian. Modeling and simulation of pollutants transport in rivers. *Appl. Math. Modelling*, 31(7):1371–1396, 2007.
- [21] Hugo Lhachemi, Ionuț Munteanu, and Christophe Prieur. Boundary output feedback stabilisation for 2-D and 3-D parabolic equations. <https://arxiv.org/abs/2302.12460>, 2023.
- [22] Jacques-Louis Lions and Enrico Magenes. Problèmes aux limites non homogènes. II. *Ann. Inst. Fourier*, 11:137–178, 1961.
- [23] Ionuț Munteanu. *Boundary stabilization of parabolic equations*, volume 93 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, 2019.
- [24] James Weimer, Bruno Sinopoli, and Bruce Krogh. Multiple source detection and localization in advection-diffusion processes using wireless sensor networks. In *2009 30th IEEE Real-Time Systems Symposium*, pages 333–342. IEEE, 2009.
- [25] Liu Yang, Zui-Cha Deng, Jian-Ning Yu, and Guan-Wei Luo. Optimization method for the inverse problem of reconstructing the source term in a parabolic equation. *Math. Comput. Simul.*, 80(2):314–326, 2009.
- [26] Sergiy Zhuk, Tigran T. Tchakian, Stephen Moore, Rodrigo Ordonez-Hurtado, and Robert Shorten. On source-term parameter estimation for linear advection-diffusion equations with uncertain coefficients. *SIAM J. Sci. Comput.*, 38(4):A2334–A2356, 2016.