



**HAL**  
open science

# Multi-patient Inverse Estimation of Effective Membrane Diffusion Coefficients in Calcium-Citrate Hemodialysis

Geoffrey Lacour, Nicolae Cîndea, Julien Aniort

## ► To cite this version:

Geoffrey Lacour, Nicolae Cîndea, Julien Aniort. Multi-patient Inverse Estimation of Effective Membrane Diffusion Coefficients in Calcium-Citrate Hemodialysis. 2026. <hal-05532469>

**HAL Id: hal-05532469**

**<https://hal.science/hal-05532469v1>**

Preprint submitted on 2 Mar 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

# Multi-patient Inverse Estimation of Effective Membrane Diffusion Coefficients in Calcium–Citrate Hemodialysis

**Geoffrey Lacour**

Université Paris-Saclay  
INRAE, MaIAGE  
78350 Jouy-en-Josas, France  
geoffrey.lacour@inrae.fr

**Nicolae Cîndea**

Université Clermont Auvergne,  
CNRS, LMBP, F-63000 Clermont-Ferrand,  
France  
nicolae.cindea@uca.fr

**Julien Aniort**

Département de Néphrologie, Dialyse et  
Transplantation,  
CHU Clermont-Ferrand, Hôpital Gabriel  
Montpied,  
58 Rue Montalembert, 63000,  
Clermont-Ferrand, France

&

INRAE UMR 1019,  
Unité de Nutrition Humaine, Université  
Clermont Auvergne,  
Clermont-Ferrand, France  
janiort@chu-clermontferrand.fr

March 2, 2026

## Abstract

We propose a multi-patient inverse modeling framework for identifying effective calcium and citrate diffusion coefficients in hollow-fiber hemodialysis devices. The approach relies on a coupled forward model combining axisymmetric fluid dynamics with multi-species convection-reaction-diffusion, together with a derivative-free optimization strategy to estimate membrane transport parameters from outlet concentration measurements. To account for inter-patient variability, physiological input parameters are first generated from clinical data and complemented by a patient-specific hydraulic calibration step, ensuring physical consistency across the synthetic cohort. The inverse problem is formulated as a global least-squares minimization aggregating residuals over multiple patients. Numerical experiments on synthetic data demonstrate multi-patient identifiability of the diffusion coefficients in the exact-data setting. Robustness with respect to measurement noise is subsequently assessed by perturbing observable outputs at various noise levels, and sensitivity analyses are performed to quantify the influence of membrane transport parameters on model predictions. The methodology is then applied to real clinical data obtained from an AK200 Gambro/Nikkiso DBB07 dialysis system. The results indicate that aggregating information from several patients substantially improves parameter identifiability and stability compared to single-patient inversions. Overall, this work provides a physically consistent and computationally tractable framework for multi-patient parameter estimation in dialysis models,

and opens perspectives for large-scale personalization through physics-informed surrogate modeling.

**Mathematical Subject Classification (2020):** 92C50, 65N21, 76Z05, 76S05, 76V05.

## 1 Introduction

Hemodialysis is a life-sustaining therapy for patients with end-stage renal disease (ESRD). Its primary objective is to restore physiological homeostasis by selectively removing or exchanging solutes and electrolytes between the blood and the dialysate across a semi-permeable membrane as described in J.Himmelfarb, T.Ikizler & T.Alp [19]. Among these solutes, calcium plays a pivotal role in mineral metabolism and cardiovascular function. An inappropriate calcium balance during hemodialysis sessions may contribute to the development of chronic kidney disease–mineral and bone disorder (CKD-MBD), including osteomalacia, osteitis fibrosa, and adynamic bone disease. Moreover, dysregulated calcium homeostasis is strongly associated with cardiovascular complications such as valvular and vascular calcifications, as well as left ventricular hypertrophy (see M.Ketteler et al [22]).

In many clinical protocols, citrate is used as a regional anticoagulant owing to its capacity to chelate ionized calcium, thereby locally inhibiting activation of the coagulation cascade within the extracorporeal circuit. Accurate prediction of calcium transport and local concentration during intermittent hemodialysis therefore represents an important clinical and computational challenge. The coupling between citrate and calcium introduces additional complexity into transport dynamics, as calcium speciation depends on reversible binding reactions, convection, diffusion, and transmembrane exchanges. This coupling motivates detailed modeling of their interactions within the dialysis filter. Most commonly used models rely on one-dimensional approaches assuming equilibrium conditions. Although such models provide global estimates of calcium balance, they do not allow precise investigation of spatial variations in calcium concentration along the dialyzer fibers (see V.Maheshwari et al [24]). However, this spatial resolution is necessary to properly evaluate local anticoagulant efficacy, particularly since effective anticoagulation requires free calcium levels below 0.4 mmol/L within the filter. In previous work, we introduced a time- and space-dependent mathematical model of calcium transport in hollow-fiber dialyzers, based on coupled Navier–Stokes, Darcy, and convection–reaction–diffusion equations [1, 2]. These models yield qualitatively realistic concentration profiles and are consistent with clinical observations. Nevertheless, their quantitative predictive capability strongly depends on membrane transport parameters, especially effective diffusion coefficients, which are difficult to measure directly and remain poorly characterized in practice.

In this work, we address the inverse problem of identifying effective diffusion coefficients in the dialysis membrane from sparse clinical measurements. More precisely, given inlet concentrations and flow rates, together with outlet blood concentrations for several chemical species, we aim at estimating membrane diffusion parameters so that numerical simulations best reproduce observed patient data. This constitutes a nonlinear, low-dimensional inverse problem driven by an expensive forward model, involving coupled fluid dynamics and convection–reaction–diffusion equations. Additional difficulties arise from limited observability, measurement noise, and interpatient variability. Beyond model calibration, identifying effective membrane transport coefficients is a key step toward patient-specific dialysis strategies. Reliable parameter estimates would enable predictive simulations under varying dialysate compositions, opening the possibility of optimizing treatment protocols to achieve targeted calcium balance while minimizing adverse

effects. Such personalized approaches are increasingly sought in clinical practice, but require quantitative models grounded in measurable patient data.

The present study was conducted within the framework of the MARC clinical trial (ClinicalTrials.gov identifier: NCT04530175), carried out at CHU Gabriel Montpied (Clermont-Ferrand, France). The optimization of the mathematical model, was prespecified as a secondary objective of the study protocol. The analysis is based on data from twenty-two patients undergoing intermittent hemodialysis. Building upon the existing forward model, we formulate a multi-patient inverse problem and adopt a derivative-free optimization strategy to estimate two key diffusion coefficients associated with calcium and citrate transport across the membrane. To assess the reliability of the proposed approach, we first validate the method on synthetic data, allowing us to investigate identifiability in the noise-free case and robustness with respect to measurement perturbations. We then apply the methodology to clinical data obtained on AK200 Gambro and Nikkiso DBB07 dialysis machines using hollow-fiber dialyzer cartridges equipped with synthetic polysulfone membranes (Leoced-21 and/or APS series, distributed by Asahi Kasei Medical Co., Ltd.), and discuss the resulting parameter estimates.

The remainder of the paper is organized as follows. Section 2 introduces the forward model, combining fluid dynamics with multi-species transport and chemical reactions in an axisymmetric hollow-fiber geometry, and details the numerical resolution of the stationary problem. Section 3 presents the formulation of the inverse problem together with its numerical implementation. In the latter, we first consider in Section 3.1 a single-patient inverse problem as a baseline validation of the modeling and optimization pipeline. Then, Sections 3.2–3.4 address the multi-patient framework using synthetic data, focusing successively on exact-data identifiability, robustness with respect to measurement noise, and sensitivity to the biomembrane diffusion coefficients. Finally, Section 4 presents an application to clinical data obtained dialysis devices mentioned above. In Section 5, we end the presentation by summarizing the main findings and discuss perspectives, including the integration of physics-informed neural networks to reduce computational cost and improve scalability with respect to patient cohorts.

Although the present contribution builds upon the framework introduced in the works of the present authors & L.Chupin [1] as well as the work of the present authors with F.Richard, F.Thouy, L.Le Guen, C.Philipponnet, C.Garrouste, A.-E.Heng, C.Dupuis, M.Adda, D.Julie, E.Lebredonchel, L.Chupin, D.Bouvier & B.Souweine [2], it should also be viewed within a broader and highly active research landscape. Over the past decades, extensive efforts have been devoted to the modeling and the study of transport phenomena in membrane-based systems, to the mathematical analysis of related mathematical models, and to the identification of effective transport parameters from experimental or clinical data. Significant advances have been achieved in the areas of membrane characterization, inverse modeling, and parameter estimation in complex multiphysics settings, as illustrated for instance in the works of N.Cancilla, L.Gurreri, M.Ciofalo, A.Cipollina, A.Tamburini & M.Giorgio [3], of N.Cancilla, L.Gurreri, G.Marotta, M.Ciofalo, A.Cipollina, A.Tamburini & G.Micale [4], of Y.-L.Chang & C.-J.Lee [5], of D.Donato, A.Boschetti-de-Fierro, C.Zweigart, M.Kolb, S.Eloot, M.Storr, B.Krause, K.Leyboldt & P.Segers [10], of D.Donato, M.Storr & B.Krause [11], of S.Eloot, J.Vierendeels & P.Verdonck [12], of S.Eloot, D.De Wachter, I.Van Tricht & P.Verdonck [13], of A.Fasano & A.Farina [14], of M.Islam & J.Szpunar [20], of M.Jaffrin, B.Gupta & J.Malbrancq [21], of C.Legallais, G.Catapano, B.von Harten & U.Baurmeister [23], of T.Miyasaka & K.Sakai [25], of C.Ronco & W.Clark [26], of D.Schneditz, D.Platzer & J.Daugirdas [27], of A.Werynski & J.Waniewski [29], of T.Yaqoob, M.Ahsan, A.Hussain & I.Ahmad [30], of J.Zhang, X.Chen, J.Ding, K.Fraser, M.Ertan Taskin, B.Griffith & Z.Wu [31] and references therein. We may also

refer to the book of R.Fournier [15] as well as the work of J.Himmelfarb, T.Ikizler & T.Alp [19]. The present work contributes to this ongoing line of research by combining mechanistic partial differential equation modeling, multi-patient inverse identification, and numerical sensitivity analysis in the specific context of hollow-fiber hemodialysis devices.

## 2 Forward model: fluid dynamics and multi-species transport in a hollow-fiber dialyzer

This section presents the forward model that underlies the inverse identification procedure developed later. Our approach follows the modeling framework introduced in the work of the present authors & their collaborators [1, 2]. We recall the main assumptions, describe the governing equations in an axisymmetric setting, and explain the boundary conditions and numerical strategy used to compute the outlet concentrations. For the sake of clarity and reproducibility, we emphasize the physical meaning of each model component, while postponing the full non-dimensionalization details and parameter tables to [1].

A dialyzer module contains a large number of parallel hollow fibers. A classical modeling approach consists in studying a representative fiber and assuming periodicity of the surrounding structure. Each fiber may be divided into three regions: the blood channel (inner lumen), a porous membrane and the dialysate channel (outer region). We may consider as a suitable approximation that the fiber geometry is cylindrical so that flows are approximately radially symmetric, letting us work in cylindrical coordinates and reduce the three-dimensional setting to a two-dimensional axisymmetric domain in variables  $(x, r)$ . Having this in mind, we denote by  $L$  the fiber length and by  $R$  its outer radius. Let  $R_1$  be the blood channel radius and  $R_2$  the outer membrane radius. The computational domain is the rectangle

$$\Omega = \{(x, r) \in \mathbb{R}^2 \mid 0 < x < L, 0 < r < R\},$$

and we can decompose it into

$$\Omega_b = \{0 < r < R_1\}, \quad \Omega_m = \{R_1 < r < R_2\}, \quad \Omega_d = \{R_2 < r < R\},$$

so interfaces are given by  $\Gamma_{bm} = \{r = R_1\}$  and  $\Gamma_{dm} = \{r = R_2\}$ . Also, the boundary is decomposed into inlet/outlet parts for each subdomain and two horizontal boundaries corresponding to the symmetry axis  $\{r = 0\}$  and the outer confinement boundary  $\{r = R\}$  (see Figure 1).

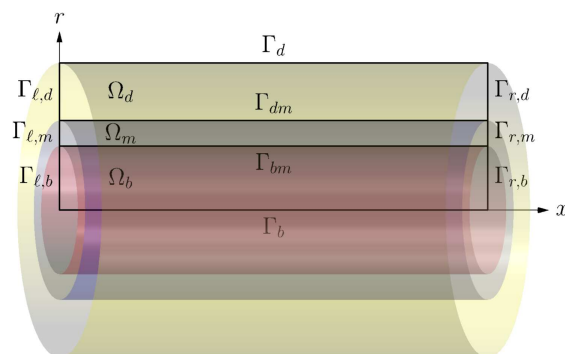


Figure 1: Axisymmetric representation of a representative hollow fiber: blood region  $\Omega_b$ , porous membrane  $\Omega_m$ , and dialysate region  $\Omega_d$ .

Using an axisymmetric formulation in cylindrical coordinates properly accounts for the fiber geometry. In particular, conservation laws naturally lead to operators of the form  $\partial_x + \frac{1}{r}\partial_r(r\cdot)$ , while cross-sectional quantities involve the radial weight  $r$ . This is essential to compute physically meaningful outlet concentrations and to enforce regularity and symmetry conditions at the axis  $r = 0$ .

## 2.1 Flow Model and Velocity Field Computation

Solute transport is driven by the velocity field in the blood and dialysate compartments. Because hollow fibers are well approximated by cylindrical geometries, we formulate the flow model in cylindrical coordinates  $(x, r, \theta)$ , where  $x$  denotes the axial direction,  $r$  the radial variable, and  $\theta$  the azimuth. We denote by  $\mathbf{v} = (v_x, v_r, v_\theta)$  the fluid velocity and by  $p$  the pressure. As a reference continuum model, we consider the incompressible Navier-Stokes equations for a Newtonian fluid, which in cylindrical coordinates are given by

$$\left\{ \begin{array}{l} \rho \left( \partial_t v_x + v_x \partial_x v_x + v_r \partial_r v_x + \frac{v_\theta}{r} \partial_\theta v_x \right) = -\partial_x p + \mu \left( \partial_x^2 v_x + \frac{1}{r} \partial_r (r \partial_r v_x) + \frac{1}{r^2} \partial_\theta^2 v_x \right), \\ \rho \left( \partial_t v_r + v_x \partial_x v_r + v_r \partial_r v_r + \frac{v_\theta}{r} \partial_\theta v_r - \frac{v_\theta^2}{r} \right) = \mu \left( \partial_x^2 v_r + \frac{1}{r} \partial_r (r \partial_r v_r) + \frac{1}{r^2} \partial_\theta^2 v_r - \frac{v_r}{r^2} - \frac{2}{r^2} \partial_\theta v_\theta \right) \\ \quad - \partial_r p, \\ \rho \left( \partial_t v_\theta + v_x \partial_x v_\theta + v_r \partial_r v_\theta + \frac{v_\theta}{r} \partial_\theta v_\theta + \frac{v_r v_\theta}{r} \right) = \mu \left( \partial_x^2 v_\theta + \frac{1}{r} \partial_r (r \partial_r v_\theta) + \frac{1}{r^2} \partial_\theta^2 v_\theta \frac{v_\theta}{r^2} + \frac{2}{r^2} \partial_\theta v_r \right) \\ \quad - \frac{1}{r} \partial_\theta p, \\ \partial_x v_x + \frac{1}{r} \partial_r (r v_r) + \frac{1}{r} \partial_\theta v_\theta = 0, \end{array} \right.$$

where  $\rho$  is the (constant) density and  $\mu$  the dynamic viscosity. We refer to standard textbooks in fluid mechanics for the derivation of these expressions, see e.g. the book of P. Davidson [9]. In hollow-fiber dialyzers, both the geometry and operating conditions are well approximated by an axisymmetric setting, and no significant azimuthal motion is observed. Mathematically, this is represented through  $\partial_\theta(\cdot) = 0$  and  $v_\theta \equiv 0$ . Under such assumptions, the velocity reduces to  $\mathbf{v} = (v_x(x, r), v_r(x, r))$  and the incompressibility condition, which is the ultimate equation in the above Navier-Stokes system, reduces to

$$\partial_x v_x + \frac{1}{r} \partial_r (r v_r) = 0.$$

The momentum equations simplify accordingly, yielding the classical axisymmetric no-swirl Navier-Stokes system

$$\left\{ \begin{array}{l} \rho \left( \partial_t v_x + v_x \partial_x v_x + v_r \partial_r v_x \right) = -\partial_x p + \mu \left( \partial_x^2 v_x + \frac{1}{r} \partial_r (r \partial_r v_x) \right), \\ \rho \left( \partial_t v_r + v_x \partial_x v_r + v_r \partial_r v_r - \frac{v_r}{r} \right) = -\partial_r p + \mu \left( \partial_x^2 v_r + \frac{1}{r} \partial_r (r \partial_r v_r) - \frac{v_r}{r^2} \right) \\ \quad \partial_x v_x + \frac{1}{r} \partial_r (r v_r) = 0. \end{array} \right.$$

At the scale of a single fiber, characteristic radii are of order  $R \sim 10^{-4}$  m, while module-level blood and dialysate flow rates are distributed over approximately  $N \sim 10^4$  parallel fibers. Using

typical clinical values reported in the work of the present authors & L.Chupin [1], this leads to moderate Reynolds numbers at the fiber level, consistent with a laminar flow regime. In the present work, we therefore neglect inertial effects and solve a stationary Stokes-type creeping flow problem in both the blood and dialysate compartments. This may be surprising at first sight, since blood may exhibit non-Newtonian behavior in general. However, in narrow conduits the apparent viscosity decreases due to the Fåhræus-Lindqvist effect, and the flow can therefore be accurately approximated by a Newtonian model. Moreover, numerical comparisons reported in [1] indicate that predicted solute concentrations are only weakly sensitive to the specific rheological law in the regimes considered. This motivates the use of a Newtonian Stokes model for both blood and dialysate in the inverse identification pipeline.

Thus, the above assumptions imply that both the time derivatives and the convective-type terms can be neglected, since we consider an axisymmetric and laminar regime without swirl, so that the governing equations reduce to

$$\begin{cases} \mu \left( \partial_x^2 v_x + \frac{1}{r} \partial_r (r \partial_r v_x) \right) - \partial_x p = 0, \\ \mu \left( \partial_x^2 v_r + \frac{1}{r} \partial_r (r \partial_r v_r) - \frac{v_r}{r^2} \right) - \partial_r p = 0, \\ \partial_x v_x + \frac{1}{r} \partial_r (r v_r) = 0, \end{cases} \quad (\text{F}_c)$$

posed in the blood and dialysate domains with according parameters. The hollow-fiber membrane is modeled as a porous medium. In this region, inertial effects remain negligible and the flow is assumed to obey Darcy law. Denoting by  $p_m$  the pressure in the membrane and by  $\mathbf{u}_m = (u_{m,x}, u_{m,r})$  the fluid velocity across the membrane, the governing equations can be written in axisymmetric cylindrical coordinates

$$\begin{cases} u_{m,x} = -\frac{K}{\mu} \partial_x p_m, \\ u_{m,r} = -\frac{K}{\mu} \partial_r p_m, \\ \partial_x u_{m,x} + \frac{1}{r} \partial_r (r u_{m,r}) = 0, \end{cases} \quad (\text{F}_m)$$

so that the membrane pressure does satisfy the elliptic equation

$$\partial_x^2 p_m + \frac{1}{r} \partial_r (r \partial_r p_m) = 0 \quad \text{in } \Omega_m. \quad (\text{P}_m)$$

Physically, Darcy law expresses a linear relation between pressure gradients and filtration velocity in the porous membrane, with the permeability  $K$  encoding the microstructural properties of the fiber wall. This model is appropriate for slow flows in highly resistive porous media and is usual in dialyzer-scale modeling. The Stokes systems in  $\Omega_b$  and  $\Omega_d$  given by (F<sub>c</sub>) are coupled to the Darcy flow in  $\Omega_m$  given in (F<sub>m</sub>) and (P<sub>m</sub>) through interface conditions enforcing continuity of the normal flux and appropriate stress or pressure transmission. In practice, fiber-level inlet and outlet pressures are not directly measured, whereas blood and dialysate flow rates are available at the machine level. Following [1], we determine consistent pressure boundary values by solving a calibration problem so that the simulated fluxes match the measured ones. The resulting velocity field is then kept fixed and used as input for the solute transport model described in the next subsection. We refer to [1] for further implementation details.

## 2.2 Multi-species Transport and Chemical Reactions

We consider the transport and interaction of five chemical species in blood and dialysate that play a central role in calcium balance under citrate anticoagulation, as described in [1]: free calcium ( $c_1$ ), free albumin binding sites ( $c_2$ ), calcium-albumin complexes ( $c_3$ ), citrate ( $c_4$ ), and calcium-citrate complexes ( $c_5$ ). These species are known to govern ionized calcium availability and buffering mechanisms during dialysis sessions under citrate anticoagulation, see e.g. the textbook of J.Hall & M.Hall [16, Unit V] for a clinical overview.

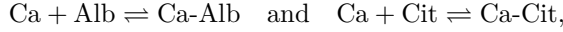
We denote by  $\mathbf{c} = (c_1, c_2, c_3, c_4, c_5)^T$  the vector of concentrations. Then for each species, conservation of mass over a control volume yields a balance equation of the form

$$\partial_t c_i + \operatorname{div}(\mathbf{J}_i) = F_i(\mathbf{c}), \quad (i = 1, \dots, 5)$$

where  $\mathbf{J}_i$  is the total flux of species  $i$  and  $F_i(\mathbf{c})$  represents the net production rate due to local chemical reactions. Namely, it accounts for reversible binding reactions between calcium, albumin and citrate, modeled through mass-action kinetics. Such reaction-diffusion systems naturally lead to coupled convection-reaction-diffusion equations and are classical in the modeling of reactive solute transport in biological fluids (see, e.g. G.Truskey, F.Yuan & D.Katz [28] for a general modeling perspective. Referring to [1] the nonlinear source term which account for interaction between chemical species can is written

$$F(\mathbf{c}) = \begin{pmatrix} F_1(\mathbf{c}) \\ F_2(\mathbf{c}) \\ F_3(\mathbf{c}) \\ F_4(\mathbf{c}) \\ F_5(\mathbf{c}) \end{pmatrix} = \frac{1}{\mathcal{F}d} \begin{pmatrix} c_3 + \delta_2 c_5 - \delta_1 c_1 c_2 - \delta_3 c_1 c_4 \\ c_3 - \delta_1 c_1 c_2 \\ -c_3 + \delta_1 c_1 c_2 \\ \delta_2 c_5 - \delta_3 c_1 c_4 \\ -\delta_2 c_5 + \delta_3 c_1 c_4 \end{pmatrix}, \quad (1)$$

where  $\delta_1, \delta_2, \delta_3$  and  $\mathcal{F}d$  are some constants. More precisely, the reaction network corresponds to reversible bindings



modeled through first-order kinetics with constants that can be associated here (up to a rescaling method) to  $\delta_1, \dots, \delta_3$ . Let us focus on the flux term  $\mathbf{J}_i$ . We decompose it as the sum of an advective flux and a diffusive flux:

$$\mathbf{J}_i = S_i c_i \mathbf{U} - D_i \nabla c_i.$$

Here  $\mathbf{U} = (U_x, U_r)$  is the velocity field solution to (F<sub>c</sub>). The coefficient  $S_i$  is a sieving factor (which is dimensionless), used to represent hindered transport across the membrane:  $S_i = 1$  corresponds to unhindered solute transport, whereas  $S_i = 0$  suppresses convection in the membrane for large molecules such as albumin. The diffusion coefficient  $D_i$  may take different values across  $\Omega_b, \Omega_m$ , and  $\Omega_d$ . In cylindrical coordinates for an axisymmetric setting, such a consideration of  $\mathbf{J}_i$  yields the convection-reaction-diffusion system

$$\partial_t c_i + S_i (U_x \partial_x c_i + U_r \partial_r c_i) - \partial_x (D_i \partial_x c_i) - \frac{1}{r} \partial_r (r D_i \partial_r c_i) = F_i(\mathbf{c}), \quad (i = 1, \dots, 5)$$

posed on the whole domain  $\Omega$ . Although  $D_i$  and  $S_i$  may be discontinuous across subdomains, all quantities are defined on  $\Omega$  and the variational formulation remains meaningful at the global level, as described in [1].

To benefit from a suitable description of the physical underlying phenomenon, we do need to impose boundary conditions that reflect known inlet concentrations and the fact that there is no loss of chemical species in the whole device. To this end, we assume that blood concentrations at the inlet are measured and dialysate concentrations at the inlet are known from the dialysate prescription. This yields Dirichlet boundary conditions given by

$$c_i = \overline{c_{i,\text{input}}} \quad \text{on } \Gamma_{\ell,b} \cup \Gamma_{r,d} \quad (i = 1, \dots, 5).$$

Such conditions represent the measured (or controlled) chemical state into the model. On the remaining boundaries, concentrations are not prescribed. A common modeling choice is to assume that there is no loss through the exterior boundary of the device, which yields homogeneous Neumann conditions

$$\partial_{\mathbf{n}} c_i = 0 \quad \text{on } \Gamma_b \cup \Gamma_{r,b} \cup \Gamma_{\ell,d} \quad (i = 1, 4, 5) \quad \text{and} \quad \partial_{\mathbf{n}} c_j = 0 \quad \text{on } \Gamma_b \cup \Gamma_{r,b} \quad (j = 2, 3).$$

Together, these conditions mix prescribed values (Dirichlet) where the chemical composition is known, and null flux assumptions (Neumann) where it is not. This mixed structure is essential for identifiability since the model is driven by inputs, while outputs remain free and become meaningful observables to compare with clinical measurements. However, it may yield several mathematical issues both from a computational and a theoretical point of views.

Finally, let us remember that we consider ourselves to be in an equilibrium state, which means that we can consider the exchange of chemical species in the dialyzer intrinsically, and thus view the system as being in equilibrium (without time derivative). Namely, this means that the regimes considered, the transient system relaxes rapidly toward a steady state along the fiber. Gathering the above observations, we may describe the considered system to describe chemical interactions into a hollow-fiber of the dialyzer:

- For  $i \in \{1, 4, 5\}$  which we recall respectively corresponds to Calcium, Citrate and Calcium-Citrate, the non-transient system is given as

$$\begin{cases} U_x \partial_x c_i + U_r \partial_r c_i - \frac{1}{P_e} \frac{1}{r} \partial_r (r D_i \partial_r c_i) - \frac{\varepsilon^2}{P_e} \partial_x (D_i \partial_x c_i) = F_i(\mathbf{c}) & \text{in } \Omega \\ c_i = \overline{c_{i,\text{input}}} & \text{on } \Gamma_{\ell,b} \cup \Gamma_{r,d} \\ \partial_{\mathbf{n}} c_i = 0 & \text{on } \Gamma_b \cup \Gamma_{r,b} \cup \Gamma_{\ell,d}; \end{cases} \quad (2)$$

- Also since Albumin and Calcium-Albumin do not cross the membrane, the equations representing species  $i \in \{2, 3\}$  are fully described in  $\Omega_b$  and are given by

$$\begin{cases} U_x \partial_x c_i + U_r \partial_r c_i - \frac{1}{P_e} \frac{1}{r} \partial_r (r D_i \partial_r c_i) - \frac{\varepsilon^2}{P_e} \partial_x (D_i \partial_x c_i) = F_i(\mathbf{c}) & \text{in } \Omega_b \\ c_i = \overline{c_{i,\text{input}}} & \text{on } \Gamma_{\ell,b} \\ \partial_{\mathbf{n}} c_i = 0 & \text{on } \Gamma_b \cup \Gamma_{r,b} \\ D_i \partial_r c_i = c_i U_r & \text{on } \Gamma_{bm}. \end{cases} \quad (3)$$

In systems (2) and (3), the interaction terms are given by (1), and all constants correspond to physical parameters (e.g. a Péclet number and a  $\varepsilon^2$ -weighted longitudinal diffusion term). Since their explicit values are not needed for the present study, we refer to [1] for a detailed description.

### 2.3 Numerical Resolution of the Forward Problem

For each species indexed by  $i$ , the predicted outlet concentration is computed for each species using the formula

$$c_{i, \text{ output}} = \frac{2R^2}{R_1^2} \int_{\Gamma_{r,b}} r c_i dr, \quad (\text{OP})$$

namely the model output is defined as the cross-sectional average concentration in blood at  $x = L$  for a solution  $\mathbf{c}$  of (2)–(3). We discretize  $\Omega$  with a structured triangulation and approximate the concentrations using  $P_1$  finite elements. Note that the stationary coupled system (2)–(3) is nonlinear due to the reaction term  $F(\mathbf{c})$ . We approach the solution of the stationary convection-reaction-diffusion system using a Newton method in a variational form. More exactly we first write the variational formulation associated to (2), so that for  $i \in \{1, 4, 5\}$ , we aim to find  $c_i \in \mathcal{C}_i$  solution of

$$\begin{aligned} \iint_{\Omega} (U_x \partial_x c_i + U_r \partial_r c_i) r \varphi_i dx dr + \frac{1}{\mathcal{P}e} \iint_{\Omega} r D_i \partial_r c_i \partial_r \varphi_i dx dr - \frac{\varepsilon^2}{\mathcal{P}e} \int_{\Gamma_{r,d}} r D_i \partial_x c_i \varphi_i dr \\ + \frac{\varepsilon^2}{\mathcal{P}e} \int_{\Gamma_{\ell,b}} r D_i \partial_x c_i \varphi_i dr + \frac{\varepsilon^2}{\mathcal{P}e} \iint_{\Omega} r D_i \partial_x c_i \partial_x \varphi_i dx dr = \iint_{\Omega} r F_i(\mathbf{c}) \varphi_i dx dr, \end{aligned} \quad (4)$$

for every  $\varphi_i \in \mathcal{C}_i$ , where  $\mathcal{C}_i$  are suitable Sobolev spaces, that in particular take into account the boundary conditions described in (2). Similarly, we can write a variational formulation associated to (3). The main difference with respect to the variational formulation (4) is that the solutions to (3) are supported in  $\Omega_b$  and not in  $\Omega$  and are null in  $\Omega_m$  and  $\Omega_d$ . Another difference is the presence of a boundary term on  $\Gamma_{bm}$  due to the boundary condition. More precisely, for  $i \in \{2, 3\}$  are looking for  $c_i \in \mathcal{C}_i$  which solve

$$\begin{aligned} \iint_{\Omega_b} (U_x \partial_x c_i + U_r \partial_r c_i) r \varphi_i dx dr + \frac{1}{\mathcal{P}e} \iint_{\Omega_b} r D_i \partial_r c_i \partial_r \varphi_i dx dr - \int_{\Gamma_{bm}} \frac{U_r}{\mathcal{P}e} c_i \varphi_i dr \\ + \frac{\varepsilon^2}{\mathcal{P}e} \int_{\Gamma_{\ell,b}} r D_i \partial_x c_i \varphi_i dr + \frac{\varepsilon^2}{\mathcal{P}e} \iint_{\Omega_b} r D_i \partial_x c_i \partial_x \varphi_i dx dr = \iint_{\Omega_b} r F_i(\mathbf{c}) \varphi_i dx dr, \end{aligned} \quad (5)$$

for every  $\varphi_i \in \mathcal{C}_i$ , where once again  $\mathcal{C}_i$  are Sobolev spaces taking into account the boundary conditions present in (3). We highlight that we can gather (4)–(5), and rewrite the variational problem so that we are seeking for  $\mathbf{c} = (c_i)_{1 \leq i \leq 5} \in \mathcal{C}$  solving

$$\mathcal{F}(\mathbf{c}, \varphi) = \mathbf{0}, \quad (6)$$

for every  $\varphi = (\varphi_i)_{1 \leq i \leq 5} \in \mathcal{C} = \prod_{i=1}^5 \mathcal{C}_i$  with  $\mathcal{F} : \mathcal{C} \times \mathcal{C} \rightarrow \mathbb{R}^5$  regrouping the integral terms in the five equalities (4)–(5).

To implement a Newton method, we have to compute the gradient of  $\mathcal{F}$ , denoted as usual  $\nabla \mathcal{F}$ . Moreover, we can write it as follows

$$\nabla \mathcal{F}(\mathbf{c}, \varphi) \bar{\mathbf{c}} = \lim_{h \rightarrow 0} \frac{\mathcal{F}(\mathbf{c} + h \bar{\mathbf{c}}, \varphi) - \mathcal{F}(\mathbf{c}, \varphi)}{h} = \mathcal{L}(\bar{\mathbf{c}}, \varphi) + \mathcal{N}(\mathbf{c}, \bar{\mathbf{c}}, \varphi). \quad (7)$$

In the above, the term  $\mathcal{L}$  brings together the linear terms in the right-hand sides of (2)–(3), while  $\mathcal{N}$  gathers the remaining nonlinear terms, *i.e.*,

$$\mathcal{N}(\mathbf{c}, \bar{\mathbf{c}}, \varphi) = \left( \iint_{\Omega} r (\nabla F(\mathbf{c}) \bar{\mathbf{c}})_i \varphi_i dx dr \right)_{1 \leq i \leq 5}.$$

We give the detail of the variational Newton method employed to approach the solution of (6) in Algorithm 1.

---

**Algorithm 1** Newton method in variational form for equation (6).

---

**Require:**  $\mathcal{F}$ ,  $\mathbf{c}_0$ ,  $tol$ ,  $n_{\max}$   $\triangleright n_{\max}$  is the maximal number of iterations

$n \leftarrow 0$ .

Compute  $\mathbf{c}_{n+1} \in \mathcal{C}$  solution to

$$\nabla \mathcal{F}(\mathbf{c}_n, \varphi) \mathbf{c}_{n+1} = \nabla \mathcal{F}(\mathbf{c}_n, \varphi) \mathbf{c}_n - \mathcal{F}(\mathbf{c}_n, \varphi) \quad \text{for every } \varphi \in \mathcal{C}.$$

**while**  $n \leq n_{\max}$  and  $\|\mathbf{c}_{n+1} - \mathbf{c}_n\| > tol$  **do**  $\triangleright$  The tolerance  $tol$  is taken equal to  $10^{-4}$

$n \leftarrow n + 1$

Compute  $\mathbf{c}_{n+1} \in \mathcal{C}$  solution to

$$\nabla \mathcal{F}(\mathbf{c}_n, \varphi) \mathbf{c}_{n+1} = \nabla \mathcal{F}(\mathbf{c}_n, \varphi) \mathbf{c}_n - \mathcal{F}(\mathbf{c}_n, \varphi) \quad \text{for every } \varphi \in \mathcal{C}.$$

**end while**

**if**  $\|\mathbf{c}_{n+1} - \mathbf{c}_n\| \leq tol$  **then**

The algorithm converged.

The solution of the variational problem (6) is approximated by  $\mathbf{c}_{n+1}$ .

**end if**

---

We are now able to precise a few the underlying finite element structure of the fomulation. We consider structured regular triangulation  $\mathcal{T}_h$  of  $\Omega$  with triangles of diameter  $h$  and approach  $\mathcal{C}_i$  by its finite dimensional subspace  $\mathcal{C}_h$  formed by the continuous function whose restriction to any triangle of  $\mathcal{T}_h$  is a first degree polynomial. All the numerical simulations were done using FreeFEM [18]. In practice, convergence is achieved in a small number of Newton iterations for the parameter regimes considered, and this stationary solver constitutes the forward map used repeatedly within the inverse identification procedure, as described by Figure 2 (please see Table 1).

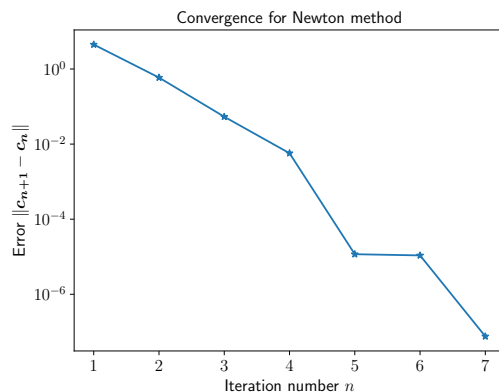


Figure 2: Evolution of the error in Newton algorithm described in Algorithm 1 for boundary data given in Table 1.

### 3 Inverse Problem Formulation and Numerical Implementation

We now turn to the numerical resolution of the inverse problem consisting in the identification of effective membrane diffusion coefficients from partial output measurements. The inversion is formulated as a nonlinear least-squares problem, where each evaluation of the cost functional requires the solution of the forward model described in Section 2.

In order to assess the reliability and stability of the approach, we adopt a progressive validation strategy. We first consider a single-patient configuration, which serves as a baseline validation of the numerical pipeline. We then extend the analysis to a multi-patient setting, allowing us to investigate improvements in parameter identifiability and estimation stability. The robustness of the inversion with respect to noisy observations is subsequently examined, followed by a sensitivity analysis highlighting the respective influence of the diffusion coefficients on the measured outputs. This structured approach enables us to evaluate both the numerical feasibility of the inverse problem and its practical limitations in terms of computational cost and data quality.

#### 3.1 Single-patient Inverse Problem as a Baseline Validation

We first consider a single-patient configuration in order to validate the inverse pipeline on a reference dataset. In this setting, inlet concentrations in blood and dialysate, together with outlet blood concentrations, are available from experimental measurements for one patient. These data are subsequently used to generate synthetic datasets for the multi-patient experiments presented in the following sections.

Diffusion within the membrane is modeled through dimensionless coefficients  $\alpha_i \in (0, 1]$ , defined as fractions of the corresponding diffusion coefficients  $D_i$  in blood (see (5) in Section 2.3). As discussed in Section 2.2, albumin and calcium-albumin complexes do not cross the membrane, and therefore we set  $\alpha_2 = \alpha_3 = 0$ . Moreover, citrate and calcium-citrate are assumed to exhibit similar diffusion properties, leading to the constraint  $\alpha_4 = \alpha_5$ . As a consequence, only two independent parameters remain to be identified, namely  $\alpha_1$  and  $\alpha_4$ . Let  $\overline{\mathbf{c}}_{\text{input}} \in \mathbb{R}^{10}$  denote the measured inlet concentrations in blood and dialysate, and let  $\overline{\mathbf{c}}_{\text{output}} \in \mathbb{R}^5$  be the measured outlet concentrations in blood. For a given vector of diffusion parameters  $\boldsymbol{\alpha} = (\alpha_i)_{1 \leq i \leq 5}$ , the forward model introduced in Section 2.3 provides predicted outlet concentrations

$$\mathcal{B}_{\boldsymbol{\alpha}}(\overline{\mathbf{c}}_{\text{input}}) \in \mathbb{R}^5,$$

obtained by solving the stationary transport problem using FreeFEM. More precisely the output is obtained through the use of the formula (OP) and the inverse problem is formulated as a nonlinear least-squares minimization. To this end, we introduce the admissible set

$$\mathcal{A} = \{ \boldsymbol{\alpha} \in \mathbb{R}^5 \mid \alpha_2 = \alpha_3 = 0, \alpha_4 = \alpha_5, \alpha_i \in (0, 1] \text{ for } i \in \{1, 4, 5\} \},$$

and we seek

$$\min_{\boldsymbol{\alpha} \in \mathcal{A}} J(\boldsymbol{\alpha}),$$

where the above cost functional is defined by

$$J(\boldsymbol{\alpha}) = \sum_{i=1}^5 \frac{|\mathcal{B}_{\boldsymbol{\alpha},i}(\overline{\mathbf{c}}_{\text{input}}) - \overline{c}_{i,\text{output}}|^2}{|\overline{c}_{i,\text{output}}|^2}. \quad (8)$$

Due to the previously mentioned constraints, the optimization reduces to two parameters. Introducing  $\beta = (\beta_1, \beta_2)$  and setting  $\alpha = (\beta_1, 0, 0, \beta_2, \beta_2)$ , the problem becomes the minimization over  $[0, 1]^2$  of the reduced functional

$$\mathcal{J}(\beta) = J(\alpha).$$

In this baseline configuration, the inverse problem is solved directly within the FreeFEM environment using a projected gradient descent method with adaptive step size. For each candidate parameter vector  $\beta$ , the forward problem is solved using the Newton algorithm described in Algorithm 1, and the resulting outlet concentrations are used to evaluate the reduced cost functional  $\mathcal{J}$ . Finite-difference approximations of the gradient are computed by repeated calls to the forward solver, and bound constraints are enforced through projection onto  $[0, 1]^2$ . More precisely, for a small parameter  $h > 0$ , we approximate

$$\nabla \mathcal{J}(\beta) \cdot \mathbf{i} \approx \frac{\mathcal{J}(\beta + h\mathbf{i}) - \mathcal{J}(\beta)}{h}, \quad \nabla \mathcal{J}(\beta) \cdot \mathbf{j} \approx \frac{\mathcal{J}(\beta + h\mathbf{j}) - \mathcal{J}(\beta)}{h},$$

with  $\mathbf{i} = (1, 0)$  and  $\mathbf{j} = (0, 1)$ . The resulting pseudo- gradient vector is then assembled as

$$\nabla \mathcal{J}(\beta) = (\nabla \mathcal{J}(\beta) \cdot \mathbf{i}, \nabla \mathcal{J}(\beta) \cdot \mathbf{j})^T. \quad (9)$$

The complete optimization procedure is summarized in Algorithm 2. Each evaluation of  $\mathcal{J}$  requires the solution of the forward problem, so that several calls to the Newton solver described in Algorithm 1 are performed at every iteration. This single-patient implementation serves as a proof of concept for the inverse methodology and provides a reference solution for subsequent experiments. In the following sections, this approach is extended and embedded into a more general FreeFEM-Python computational framework, allowing for multi-patient inversion and the use of more robust derivative-free optimization strategies.

---

**Algorithm 2** Gradient descent algorithm with adaptive step for the minimization of  $\mathcal{J}$ .

---

**Require:**  $\mathcal{J}, \beta_0, tol, n_{\max}$ . ▷  $n_{\max}$  is the maximal number of iterations  
**Require:** initial\_step ▷ initial\_step is the initial descent step  
**Require:** tol. ▷ tol is a small value

$n \leftarrow 0$ .  
Compute  $\mathcal{J}(\beta_n)$ . ▷ Use Newton algorithm described in Algorithm 1  
Compute  $\nabla \mathcal{J}(\beta_n)$ . ▷ Two more calls of Newton method  
 $s \leftarrow \text{initial\_step}$   
**while**  $\mathcal{J}(\mathcal{P}(\beta_n - s \cdot \nabla \mathcal{J}(\beta_n))) \geq \mathcal{J}(\beta_n)$  and  $\|\mathcal{P}(\beta_n - s \cdot \nabla \mathcal{J}(\beta_n)) - \beta_n\| > tol$  **do**  
▷  $\mathcal{P}(\beta)$  is the projection of  $\beta$  on  $[0, 1]^2$   
 $s \leftarrow s/2$ .  
**end while**  
**while**  $\|\mathcal{P}(\beta_n - s \cdot \nabla \mathcal{J}(\beta_n)) - \beta_n\| > tol$  and  $n < n_{\max}$  **do**  
 $\beta_{n+1} \leftarrow \mathcal{P}(\beta_n - s \cdot \nabla \mathcal{J}(\beta_n))$ .  
 $n \leftarrow n + 1$ .  
Compute  $\nabla \mathcal{J}(\beta_n)$ .  
 $s \leftarrow \text{initial\_step}$ .  
**while**  $\mathcal{J}(\mathcal{P}(\beta_n - s \cdot \nabla \mathcal{J}(\beta_n))) \geq \mathcal{J}(\beta_n)$  and  $s \cdot \|\mathcal{P}(\beta_n - s \cdot \nabla \mathcal{J}(\beta_n)) - \beta_n\| > tol$  **do**  
 $s \leftarrow s/2$ .  
**end while**  
**end while**  
The approximation of the minima of  $\mathcal{J}$  is  $\beta_n$ .

---

As a preliminary numerical illustration, we apply the inverse procedure to the data of one patient from the clinical study. Boundary concentrations at the inlet and outlet of the dialyser are reported in Table 1.

Boundary value	$c_1$	$c_3$	$c_3$	$c_4$	$c_5$
on $\Gamma_{\ell,b}$	0.11	3.71602	0.0577928	5.03048	1.37152
on $\Gamma_{r,d}$	1.25	0	0	0	0
on $\Gamma_{r,b}$	0.96	3.577187	0.48553	0.144108	0.342892

Table 1: Concentrations (in  $\text{mol} \cdot \text{m}^{-3}$ ) of the five chemical species in blood (boundary  $\Gamma_{\ell,b}$ ) and dialysate (boundary  $\Gamma_{r,d}$ ) at the entrance of the dialyser and in blood at the exit of dialyser (boundary  $\Gamma_{r,b}$ ) for the patient number 1.

Starting from an initial guess  $\beta_0 = (0.2, 0.2)$ , the projected gradient descent algorithm converges toward diffusion coefficients that significantly improve the agreement between simulated and measured outlet concentrations.

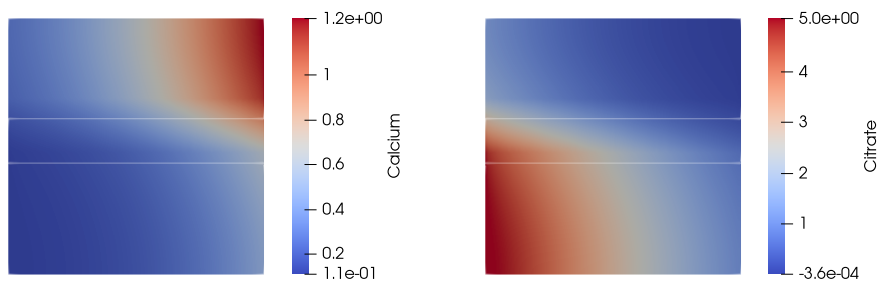


Figure 3: Components  $c_1$  (corresponding to Calcium) and  $c_4$  (corresponding to Citrate) of the numerical approximation of the solution  $c$  computed using Newton's method for the initial data in Table 1 and  $\beta = (0.2, 0.2)$ .

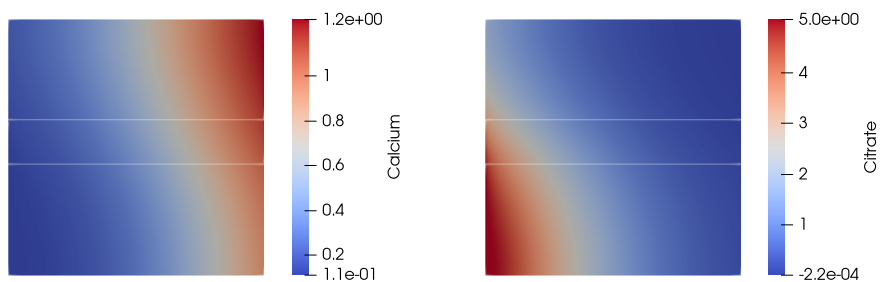


Figure 4: Components  $c_1$  and  $c_4$  of the numerical approximation of the solution  $c$  for  $\beta$  minimizing  $\mathcal{J}$ .

Figures 3 and 4 respectively display the concentrations of Calcium and Citrate for  $\beta_0 = (0.2, 0.2)$  and the computed Calcium and Citrate concentrations for the optimized parameters after the convergence of the procedure described above, showing a close match with experimental values at the dialyser outlet. The corresponding value of the cost functional  $\mathcal{J}$  is reduced by more than two orders of magnitude compared to the initial guess, indicating that the inverse procedure is able to recover physically meaningful membrane diffusion coefficients in this simplified setting.

While preliminary validation can be performed on a single patient, robust identification of diffusion coefficients requires a multi-patient setting. We therefore construct a synthetic cohort whose statistical structure is consistent with the available clinical data. Beyond inter-patient variability, we further investigate the robustness of the inverse reconstruction with respect to parametric uncertainty by considering two complementary sources of perturbations. First, measurement noise is introduced at the level of the synthetic concentration data, reflecting variability and uncertainty in observable quantities. Second, we assess the sensitivity of the inverse problem with respect to the calcium and citrate diffusion coefficients by perturbing their reference values and analyzing the resulting model outputs.

In this latter setting, a single synthetic patient is selected and multiple forward simulations are performed using perturbed diffusion coefficients of the form  $d = d^*(1 + \varepsilon)$ , where  $d^*$  denotes the nominal value and  $\varepsilon$  is a prescribed relative perturbation. The corresponding outputs are then treated as new targets for the inverse procedure. This approach allows us to quantify the stability of the reconstruction with respect to modeling errors in the coefficients, independently of inter-patient variability. Together with the multi-patient experiments based on noisy concentration data, these tests provide a comprehensive numerical assessment of both robustness to measurement noise and sensitivity to parametric perturbations.

### 3.2 Multi-patient Inverse Framework and Exact-data Identifiability

Let  $\mathcal{D}_{\text{real}}$  denote the table of real patient measurements, where each row corresponds to a physiological quantity (inlet/outlet pressures, flows, biochemical concentrations, etc.) and each column corresponds to a patient. Our goal is to generate a fixed number  $N_s$  of synthetic patients while preserving, for each variable, the empirical mean and standard deviation observed in the real cohort.

For each physiological variable  $X$ , we proceed independently as follows. Let  $\{x_j\}_{j=1}^{N_r}$  be the available real measurements for that variable, after removal of missing values. We compute

$$\mu_X = \frac{1}{N_r} \sum_{j=1}^{N_r} x_j \quad \text{and} \quad \sigma_X = \sqrt{\frac{1}{N_r - 1} \sum_{j=1}^{N_r} (x_j - \mu_X)^2},$$

together with the minimum observed value  $m_X = \min_j x_j$ . Synthetic samples are then drawn from a normal distribution

$$X^{(k)} \sim \mathcal{N}(\mu_X, \sigma_X^2), \quad k = 1, \dots, N_s,$$

and truncated from below by a conservative floor equal to half the minimum observed value, namely

$$X^{(k)} \leftarrow \max(X^{(k)}, 0.5 m_X).$$

If  $\sigma_X = 0$  or undefined, all synthetic values are therefore set equal to  $\mu_X$ . This simple truncation prevents negative or unphysical values while preserving the first two empirical moments of the

real data. All variables are generated independently in such a way, yielding a synthetic patient matrix

$$\mathcal{M}_{\text{syn}} \in \mathbb{R}^{M \times N_s},$$

where  $M$  is the number of measured fields. A fixed random seed is used to ensure reproducibility. Importantly, only the physiological input data are synthesized at this stage. We emphasize that this synthetic generation ignores cross-variable correlations and is intended only to provide a controlled numerical testbed, not to reproduce full physiological variability.

The dialysis model outputs (five observable targets per patient) are not generated by noise injection. Instead, these are computed consistently by solving the forward problem (2)–(3) for prescribed diffusion coefficients and subsequently used as reference targets in the inverse problem. The procedure implemented in our pipeline is summarized in Algorithm 3.2.

---

**Algorithm 3** Generation of a synthetic multi-patient cohort from real measurements.

---

**Require:** Real patient data table  $\mathcal{D}_{\text{real}} \in \mathbb{R}^{M \times N_r}$  (rows = fields, columns = patients).

**Require:** Number of synthetic patients  $N_s \in \mathbb{N}$ .

**Require:** Random seed  $s$  (optional, for reproducibility).

**Ensure:** Synthetic patient data table  $\mathcal{M}_{\text{syn}} \in \mathbb{R}^{M \times N_s}$ .

Initialize the random number generator with seed  $s$ .

Initialize  $\mathcal{M}_{\text{syn}}$  as an empty  $M \times N_s$  table.

**for**  $i = 1, \dots, M$  **do** ▷ Loop over physiological fields

Extract the non-missing real values  $\{x_j\}_{j=1}^{N_i}$  from row  $i$ .

**if**  $N_i = 0$  **then**

Set  $\mathcal{M}_{\text{syn}}(i, k) \leftarrow \text{NaN}$  for all  $k = 1, \dots, N_s$  and **continue**.

**end if**

Compute  $\mu_i \leftarrow \text{mean}(\{x_j\}_{j=1}^{N_i})$ .

Compute  $\sigma_i \leftarrow \text{std}(\{x_j\}_{j=1}^{N_i})$  (sample standard deviation).

Compute  $m_i \leftarrow \min(\{x_j\}_{j=1}^{N_i})$ .

Set the lower bound  $\ell_i \leftarrow 0.5 m_i$ . ▷ Conservative floor to avoid unphysical values

**if**  $\sigma_i$  is undefined or  $\sigma_i = 0$  **then**

Set  $y_k \leftarrow \mu_i$  for all  $k = 1, \dots, N_s$ . ▷ Constant row

**else**

**for**  $k = 1, \dots, N_s$  **do**

Draw  $y_k \sim \mathcal{N}(\mu_i, \sigma_i^2)$ .

**end for**

**end if**

**for**  $k = 1, \dots, N_s$  **do**

Enforce the floor:  $y_k \leftarrow \max(y_k, \ell_i)$ .

Set  $\mathcal{M}_{\text{syn}}(i, k) \leftarrow y_k$ .

**end for**

**end for**

**return**  $\mathcal{M}_{\text{syn}}$ .

---

Also, in addition to sampling physiological input variables, a calibration step is performed to construct a consistent set of synthetic reference data. More precisely, for each synthetic patient, auxiliary boundary quantities (in particular pressure profiles or hydraulic parameters) are determined by solving an inverse problem within the forward pipeline. This calibration ensures that the forward model reproduces prescribed reference concentration levels at the dialyzer inlet

and outlet, thereby enforcing physical consistency between inputs and outputs. This preliminary inverse step is not the focus of the present work and is only used to generate coherent synthetic datasets. The calibrated pressure data are subsequently treated as fixed patient-specific inputs, while the diffusion coefficients remain unknowns in the multi-patient inverse problem considered below.

Thus, such a synthetic cohort is subsequently used to generate patient-specific forward solutions of the dialysis model, from which the observable outputs are extracted. These outputs constitute the multi-patient targets employed in the inverse problem for diffusion coefficient identification.

We now turn to the identification of the diffusion coefficients in a multi-patient setting, using the synthetic cohort constructed above. While single-patient experiments provide a first validation of the modeling and inversion pipeline, aggregating information across several patients is needed to improve robustness with respect to noise and to mitigate patient-specific variability.

Let  $\mathcal{P} = \{1, \dots, N_p\}$  denote the set of synthetic patients, namely the columns of  $\mathcal{M}_{\text{syn}}$  described above. For each patient  $p \in \mathcal{P}$  and for a given vector of diffusion coefficients  $\boldsymbol{\beta} = (d_{Ca}, d_{Ci}) \in (0, +\infty)^2$ , we denote by  $F_p(\boldsymbol{\beta}) \in \mathbb{R}^m$  the vector of observable quantities produced by the forward solver, where  $m = 5$  in the present study. The mapping  $F_p$  is obtained by solving the convection-reaction-diffusion model with patient-specific inputs (geometry, calibrated pressure data, and synthetic physiological parameters), and extracting the corresponding outlet quantities at the end of the simulation. From now on, we should denote by  $y_p^{\text{obs}} \in \mathbb{R}^m$  the reference target vector associated to patient  $p$ , generated as described in the previous section.

The diffusion coefficients are estimated by minimizing a global least-squares misfit that aggregates residuals over all patients:

$$\mathcal{J}(\boldsymbol{\beta}) = \sum_{p \in \mathcal{P}} \|W \cdot (F_p(\boldsymbol{\beta}) - y_p^{\text{obs}})\|_2^2 + \lambda \mathcal{R}(\boldsymbol{\beta}). \quad (10)$$

Here  $W \in \mathbb{R}^{m \times m}$  ( $m = 5$ ) is a diagonal scaling operator used to normalize the different components of the output vector, so that no single observable dominates the objective due to its physical units or magnitude. Usually, the diagonal entries of  $W$  are chosen as the inverse of empirical scales (for example mean absolute values or standard deviations computed over the cohort). The optional regularization term  $\mathcal{R}$  enforces weak a priori constraints on the parameter range, and  $\lambda \geq 0$  denotes its weight. In the synthetic experiments reported below,  $\lambda$  is taken small or equals to zero whenever the inverse problem is sufficiently well-conditioned.

The evaluation of  $\mathcal{J}$  requires solving the forward system (2)–(3) for each patient and is therefore computationally expensive. Moreover, the forward solver is treated as a “black box”, so that analytic gradients are unavailable. For these reasons, we adopt a derivative-free direct-search strategy to minimize the functional  $\mathcal{J}$  given in (10). Since there is only two parameters to estimate, classical methods such as Powell algorithm are well suited and typically converge in a moderate number of objective evaluations. Powell method proceeds by iteratively performing one-dimensional line searches along a set of directions in parameter space. Starting from an initial guess, the objective function is minimized successively along each direction, and at the end of each iteration, a new direction is constructed from the net displacement between successive iterates and replaces the least effective previous one. This mechanism allows the method to progressively build approximate curvature information without computing derivatives. In practice, Powell algorithm is expected to converge in a moderate number of objective evaluations and provides stable reconstructions in both noise-free and noisy experiments, we refer the interested reader to the book of A.Conn, K.Scheinberg & L.Vicente [8] for more details concerning Powell-type and more generally derivative-free methods.

Also, to enforce positivity and improve numerical conditioning, the parameters are reparametrized as  $\beta = (e_1^z, e_2^z)$ , and the optimization is performed with respect to  $z \in \mathbb{R}^2$ . Admissible bounds on  $\beta$  are imposed through a penalty term in (10). Forward failures or nonphysical parameter values are handled robustly by returning a large objective value, allowing the optimizer to automatically avoid such regions. In the single-patient setting considered previously, a pseudo-gradient descent with adaptive step size has been efficiently implemented, as finite-difference approximations of the gradient remain affordable. In contrast, in the multi-patient framework each evaluation of  $\mathcal{J}$  already involves  $N_s$  forward solves, and numerical gradients would multiply this cost. Derivative-free optimization therefore provides a more favorable compromise between robustness and computational efficiency in the present low-dimensional parameter setting.

Also, as aforementioned a major computational issue lies in the evaluation of  $\mathcal{J}(\beta)$ , which requires computing the patient-specific forward maps  $(F_p(\beta))_{p \in \mathcal{P}}$ . Since these forward problems are mutually independent, we parallelize their computation on a single workstation by launching multiple patient simulations concurrently using a local multi-process strategy. This reduces the wall-clock time of each objective evaluation almost proportionally to the number of available CPU cores, up to input and output limitations. In addition, optional heavy outputs (such as visualization files) are disabled during the inverse procedure, as they are not required for objective evaluation and significantly impact runtime. The above-mentioned reparametrization  $\beta = (e^{z_1}, e^{z_2})$  is used throughout the optimization, and admissible parameter ranges are enforced through a soft quadratic penalty added to the objective functional. Forward solver failures as well as nonphysical parameter values are handled by returning a large objective value.

While the synthetic cohort provides physiological input parameters, additional boundary quantities required by the forward system (2)–(3), notably the inlet and outlet pressures on both blood and dialysate sides, cannot be prescribed independently, as they must satisfy  $(F_c)$ . For this reason, we recall once again that a preliminary calibration step is carried out for each synthetic patient. In this auxiliary inverse problem, the pressure values at the inlet and outlet boundaries are treated as unknowns and determined so that the forward system reproduces prescribed reference concentration levels, this step ensuring consistency between flow conditions and transport dynamics.

Concretely, for each synthetic patient, the physiological parameters sampled from the real cohort are first fixed. An auxiliary inverse hydraulic problem is then solved to identify inlet and outlet pressures consistent with the model and prescribed reference concentration levels, as aforementioned. These calibrated pressures are subsequently frozen as patient-specific inputs, and the full convection–reaction–diffusion model is finally solved with prescribed diffusion coefficients  $(d_{Ca}, d_{Ci})$  to generate the observable outputs, so that the resulting outlet concentrations constitute the exact synthetic reference targets  $y_p^{\text{obs}}$  used in the inverse study. This two-stage procedure, namely an hydraulic calibration followed by a forward transport simulation, guarantees physical coherence between sampled physiological inputs, boundary conditions, and model outputs. Importantly, the synthetic observables are not only generated by direct sampling or artificial noise injection, but are obtained by solving the complete system (2)–(3) after patient-specific calibration.

Following such a procedure, we generate a synthetic cohort of  $N_s = 40$  patients as described above. For each synthetic patient, patient-specific boundary conditions are first calibrated, and the five observable outputs are then generated by solving the forward system (2)–(3) with a fixed ground-truth diffusion vector

$$\beta^* = (d_{Ca}^*, d_{Ci}^*) = (0.8, 0.4).$$

These model-consistent outputs are stored as exact reference targets  $\{y_p^{\text{obs}}\}_{p \in \mathcal{P}}$ . Before launching

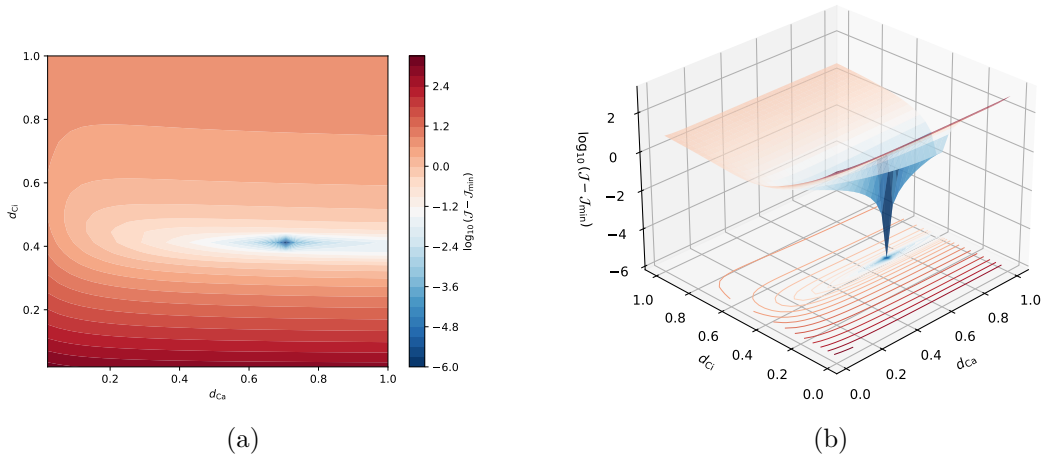


Figure 5: The geometry of the functional  $\mathcal{J}$  as a function of  $\beta = (d_{Ca}, d_{Ci})$  in logarithmic scale.

the derivative-free optimization, we perform a coarse exploratory grid search of the reduced objective functional  $\mathcal{J}(\beta)$  over the domain  $[0.02, 1] \times [0.02, 1]$ , where  $\beta = (d_{Ca}, d_{Ci})$ . The purpose of this preliminary step is not to obtain an accurate minimizer, but rather to visualize the global geometry of  $\mathcal{J}$  and to localize a priori regions of potential local or global minima. Figure 5 displays  $\mathcal{J}$  in logarithmic scale on a  $31 \times 31$  uniform grid. A pronounced valley structure clearly emerges, centered around  $d_{Ci} \approx 0.4$ , while the functional exhibits a much flatter dependence with respect to  $d_{Ca}$ . In particular, the deepest part of the valley is observed for  $0.6 \lesssim d_{Ca} \lesssim 0.9$ , indicating a strong sensitivity of the misfit to citrate diffusion, contrasted with a weaker and more correlated influence of calcium one. This anisotropic geometry suggests that  $d_{Ci}$  is more tightly identifiable than  $d_{Ca}$ , at least in the present illustrating example, the latter lying along a relatively flat direction of the objective landscape. Such a structure is typical of partially correlated parameters and explains why different optimization runs may converge to slightly different values of  $d_{Ca}$  while remaining close in  $d_{Ci}$ . Beyond providing physical insight into parameter sensitivity, this grid search also serves as a diagnostic tool to guide the initialization of Powell method and to mitigate the risk of convergence toward spurious local minima.

To assess numerical identifiability in a multi-patient setting, we then solve the inverse problem (10) using a subset of four patients (here  $p \in \{s12, s13, s14, s15\}$ ). The objective function aggregates the least-squares residuals over these patients, and the resulting two-parameter optimization problem is minimized with the Powell method described above. Convergence is monitored through the sequence of iterates  $(d_{Ca}^{(k)}, d_{Ci}^{(k)})$  and the corresponding objective values  $\mathcal{J}^{(k)}$ .

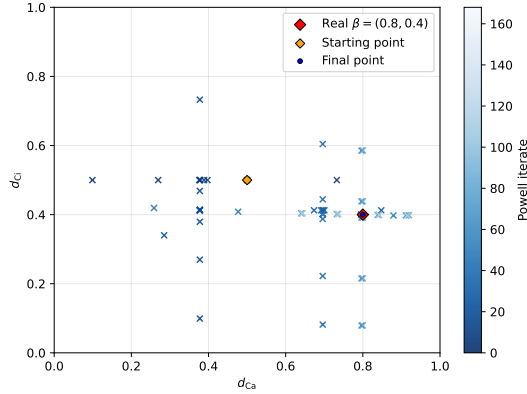


Figure 6: Convergence of the Powell method in the phase plane for the prescribed data  $\beta = (0.8, 0.4)$  without noise

Figure 6 illustrates the behavior of Powell algorithm in parameter space for the exact-data multi-patient experiment. The successive iterates  $(d_{Ca}^{(k)}, d_{Ci}^{(k)})$  are displayed in the  $(d_{Ca}, d_{Ci})$  plane, together with the ground-truth value  $\beta^* = (0.8, 0.4)$ . Starting from an initial guess, the optimization trajectory rapidly enters a neighborhood of the true parameters and subsequently converges toward  $\beta^*$ .

Figure 7 shows the evolution of the Euclidean distance  $\|\beta^{(k)} - \beta^*\|_2$  between the current estimate and the ground-truth diffusion coefficients along the Powell iterations. A rapid decrease of the parameter error is observed during the early stages of the optimization, followed by a slower refinement phase as the iterates approach the exact solution. The error eventually reaches numerical precision, confirming convergence of the reconstructed parameters toward the true values.

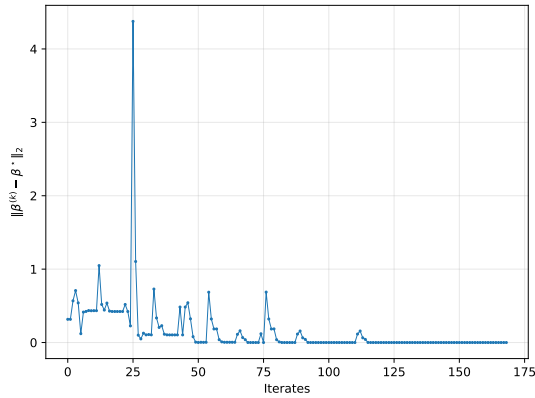


Figure 7: Error of convergence of the Powell method for the prescribed data  $\beta = (0.8, 0.4)$  without noise

The corresponding evolution of the last iterations of the objective functional  $\mathcal{J}$  is shown in Figure 8. A rapid decrease of the misfit is observed during the first iterations, followed by a

slower refinement phase, until numerical convergence is reached, as illustrated in the Figure. Both the parameter trajectory and the objective values indicate stable convergence toward the exact solution.

These results highlight that, in the absence of measurement noise, the multi-patient inverse problem admits a well-resolved minimum at the true diffusion coefficients. In other words, aggregating information from several patients restores numerical identifiability of  $(d_{Ca}, d_{Ci})$  within the proposed framework. This experiment provides a baseline validation of the multi-patient pipeline before investigating robustness with respect to measurement noise and parametric uncertainty.

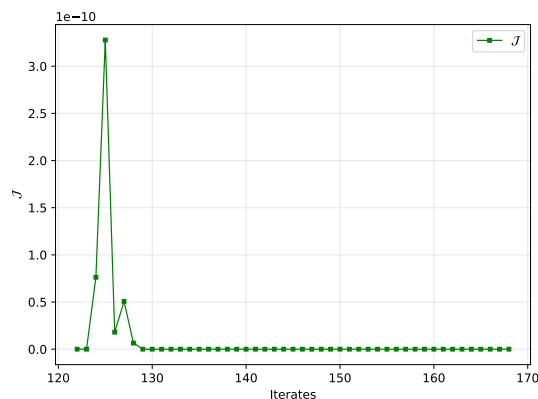


Figure 8: Convergence of the functional for the Powell method with prescribed data  $\beta = (0.8, 0.4)$  without noise

### 3.3 Robustness with respect to Measurement Noise

After establishing numerical identifiability in the exact-data setting, we now investigate the robustness of the multi-patient inverse problem with respect to measurement noise on the observable outputs. Starting from the synthetic cohort constructed in Section 3.2, we generate three additional datasets by perturbing the reference outlet concentrations with artificial noise. Only the observable targets are modified, while all physiological inputs, calibrated boundary conditions, and model parameters are kept unchanged. This allows us to isolate the effect of measurement uncertainty on the inverse reconstruction.

Let  $y_p^{\text{obs}} \in \mathbb{R}^m$  denote the exact target vector associated with patient  $p$ , where  $m = 5$  in the present study. For each component, noisy observations are generated according to a multiplicative perturbation model

$$y_{p,i}^{\text{noisy}} = y_{p,i}^{\text{obs}} (1 + \varepsilon_{p,i}),$$

where the random variables  $\varepsilon_{p,i}$  are drawn independently from a centered normal distribution with prescribed standard deviation  $\sigma$ . To avoid unrealistically large perturbations, the samples are clipped to a bounded interval proportional to  $\sigma$ . Also, we generate the synthetic cohorts by considering three different noise levels, corresponding to relative standard deviations of 1%, 3%, and 5%. Namely, for each such noise level a complete noisy synthetic cohort is generated, yielding three independent inverse test cases. A fixed random seed is used to ensure reproducibility. The inverse problem is then solved using the same multi-patient objective functional and optimization strategy as in the exact-data case.

This procedure mimics measurement uncertainty at the level of clinically observable quantities while preserving full physical consistency of the forward model. It provides a controlled framework to assess the stability of the estimated diffusion coefficients with respect to increasing noise amplitude. For each noise level, we solve the same multi-patient inverse problem as in the exact-data case, using Powell derivative-free method and identical bounds for  $(d_{Ca}, d_{Ci})$ . In order to probe the stability of the reconstruction with respect to both the noise amplitude and the cohort composition, we consider four disjoint sub-cohorts of five patients, namely  $(s1-s5)$ ,  $(s6-s10)$ ,  $(s11-s15)$ , and  $(s16-s20)$ , that we denoted by  $P1$ ,  $P2$ ,  $P3$  and  $P4$  respectively. For each sub-cohort and each noise level, we report the estimated coefficient vector  $\beta^* = (d_{Ca}^*, d_{Ci}^*)$ . Figure 9 summarizes the results in parameter space. Overall, the reconstructions remain close to the ground-truth coefficient  $\beta_{true}$  even when the observable targets are perturbed. As expected from the objective landscape discussed in Section 3.2, the estimates exhibit a stronger dispersion along the  $d_{Ca}$ -direction than along the  $d_{Ci}$ -direction, consistently with a weaker identifiability of  $d_{Ca}$ .

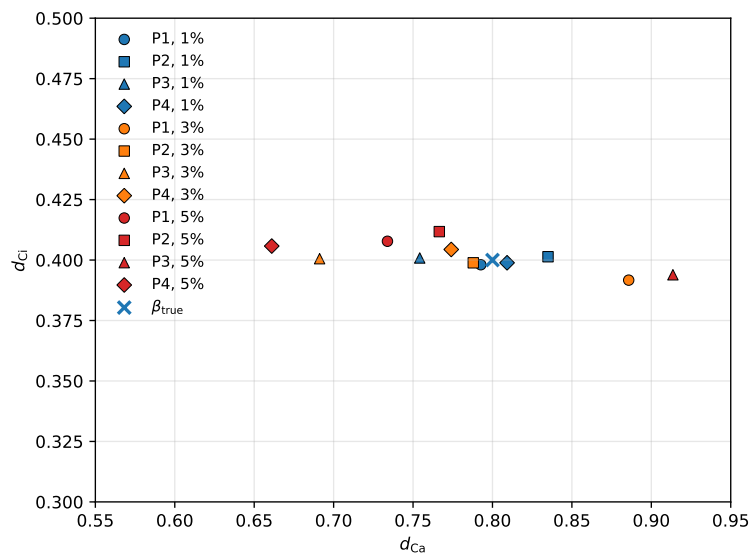


Figure 9: Estimated coefficients  $\beta^*$  for four disjoint sub-cohorts of five synthetic patients, for noise levels 1%, 3%, and 5% on the targets.

To further assess robustness at a larger scale, we also solve the inverse problem for the full cohort  $(s1-s20)$  using the 5% noisy targets. The recovered coefficient  $\beta_{20}^*$  remains close to  $\beta_{true}$ , and Figure 10 reports the resulting relative errors on each component of  $\beta$ , compared to the mean of errors for five-patients synthetic cohorts over the same patients, namely the means of errors of  $P1$ ,  $P2$ ,  $P3$  and  $P4$ . This confirms that, in this controlled setting, measurement noise at the outlet level does not induce an excessive amplification in the reconstructed effective diffusion parameters. Moreover, it highlights that more patients are taken into account the functional; lower is the noise over the diffusion coefficient of Calcium, due to its weaker identifiability. It confirms that such a multi-patient setting is particularly suited for such estimation of the diffusion coefficients.

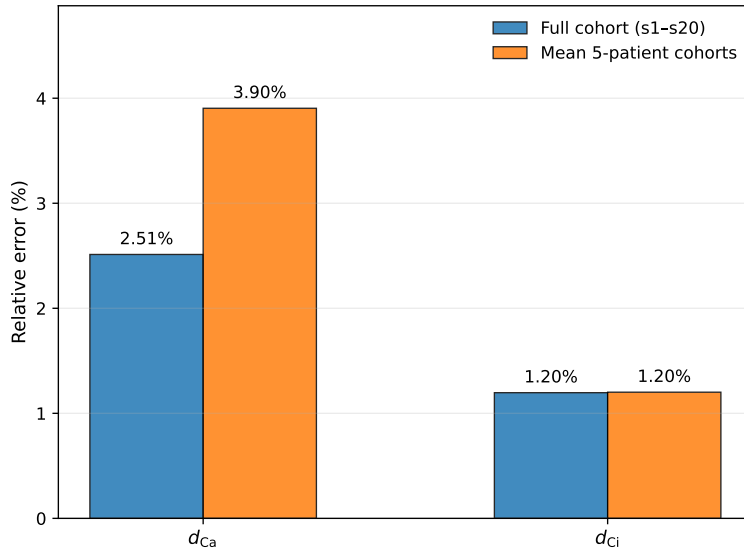


Figure 10: Relative error in percents of the reconstructed coefficients for the full cohort ( $s1-s20$ ) compared to the error of means for sub-cohorts with five synthetic patients in the case of 5% noise on the targets.

While these experiments already provide meaningful evidence of robustness on cohorts of up to  $N = 20$  patients, they remain computationally expensive because each functional evaluation requires solving the forward model for all patients in the cohort. This cost prevents a systematic exploration of larger cohorts, as the full synthetic cohort of  $N_s = 40$  synthetic patients, and more extensive uncertainty quantification through repeated noisy realizations. This motivates the development of computationally efficient surrogate models, which we discuss in Section 5. Finally, the above robustness results concern uncertainty in the observed targets. In the next subsection, we investigate the complementary question of parameter sensitivity, namely how perturbations of  $(d_{Ca}, d_{Ci})$  propagate through the forward model and affect the outlet observables.

### 3.4 Sensitivity with respect to Membrane Diffusion Coefficients

We finally investigate the intrinsic sensitivity of the forward model with respect to perturbations of the membrane diffusion coefficients. This analysis aims at quantifying how uncertainties in  $(d_{Ca}, d_{Ci})$  propagate through the nonlinear convection-reaction-diffusion system and affect the observable outlet concentrations. Starting from the nominal diffusion vector

$$\beta^* = (d_{Ca}^*, d_{Ci}^*) = (0.8, 0.4),$$

as used above, we generate perturbed coefficients for each of the  $N_s = 40$  synthetic patients according to

$$d = d^*(1 + \varepsilon),$$

where  $\varepsilon$  is drawn independently from a centered normal distribution with relative standard deviations of 1%, 3%, and 5, following the same calculations as previously done to generate such

a noise. Then for each noise level, the full forward problem (2)–(3) is solved with these modified coefficients, while all physiological inputs and calibrated hydraulic boundary conditions are kept fixed. The resulting outlet concentrations are written as new synthetic targets. Importantly, noise is not injected directly at the level of observables. Instead, the perturbations are applied solely to  $(d_{Ca}, d_{Ci})$  and propagated through the complete forward solver, ensuring physical consistency of the generated data, that are represented in Figure 11 below.

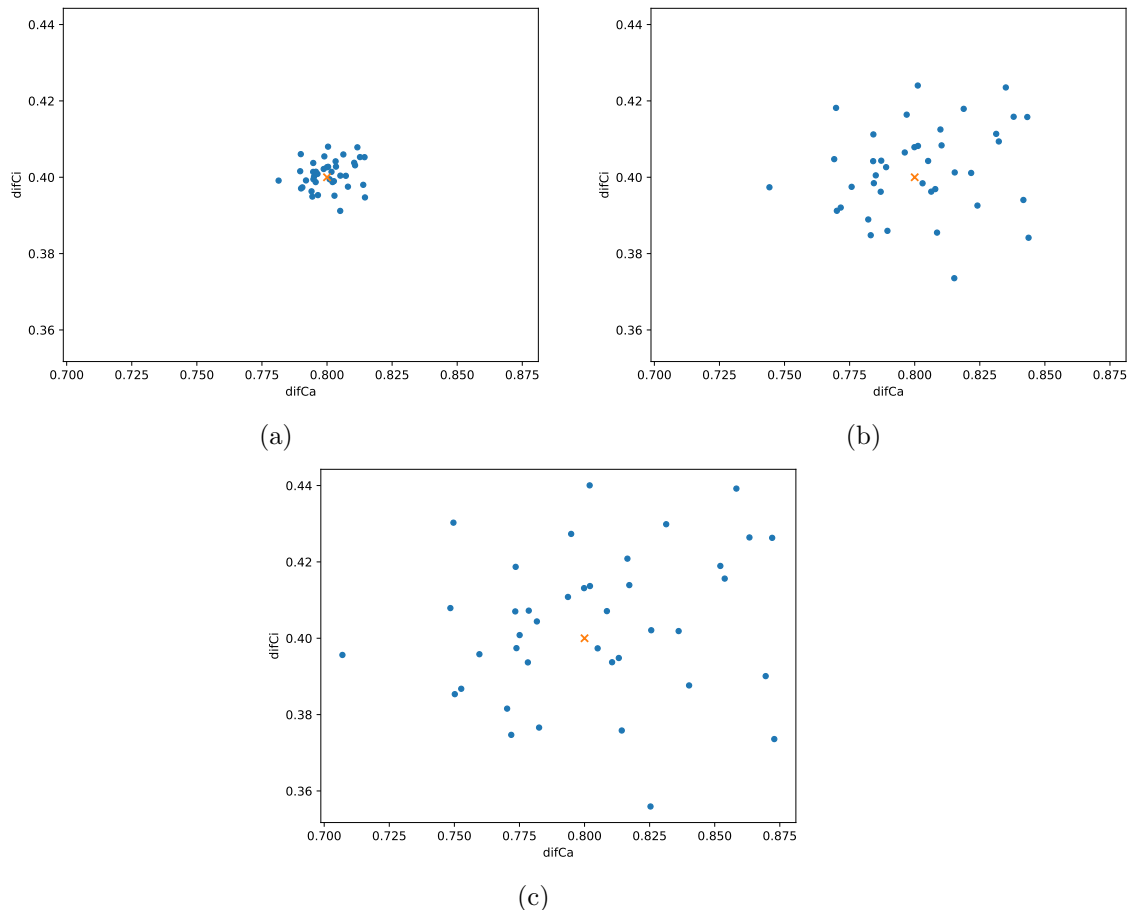


Figure 11: Noisy diffusion coefficients for synthetic cohort with a noise of order (a) 1% (b) 3% and (c) 5%.

Let  $y_p^* \in \mathbb{R}^5$  denote the reference outlet vector for patient  $p$ , and  $y_p^{\text{pert}}$  the corresponding output obtained with perturbed coefficients. For each component we compute relative errors and summarize them by the mean over the five targets. Averaged over the 40 patient cohort, the mean relative error on the outputs is approximately 0.38% for 1% coefficient perturbations, 1.12% for 3%, and 1.87% for 5%. These values remain of the same order as the imposed parameter noise and are slightly smaller on average. Even in the worst cases, the maximal mean error across patients remains bounded (below 1%, 3.1%, and 5.3% respectively). This behavior indicates that, in the considered regime, the forward map does not amplify uncertainties in  $(d_{Ca}, d_{Ci})$  and exhibits a moderate Lipschitz-type sensitivity with respect to these parameters. Figure 12 below

exhibits the mean noise on the targets for each synthetic patient.

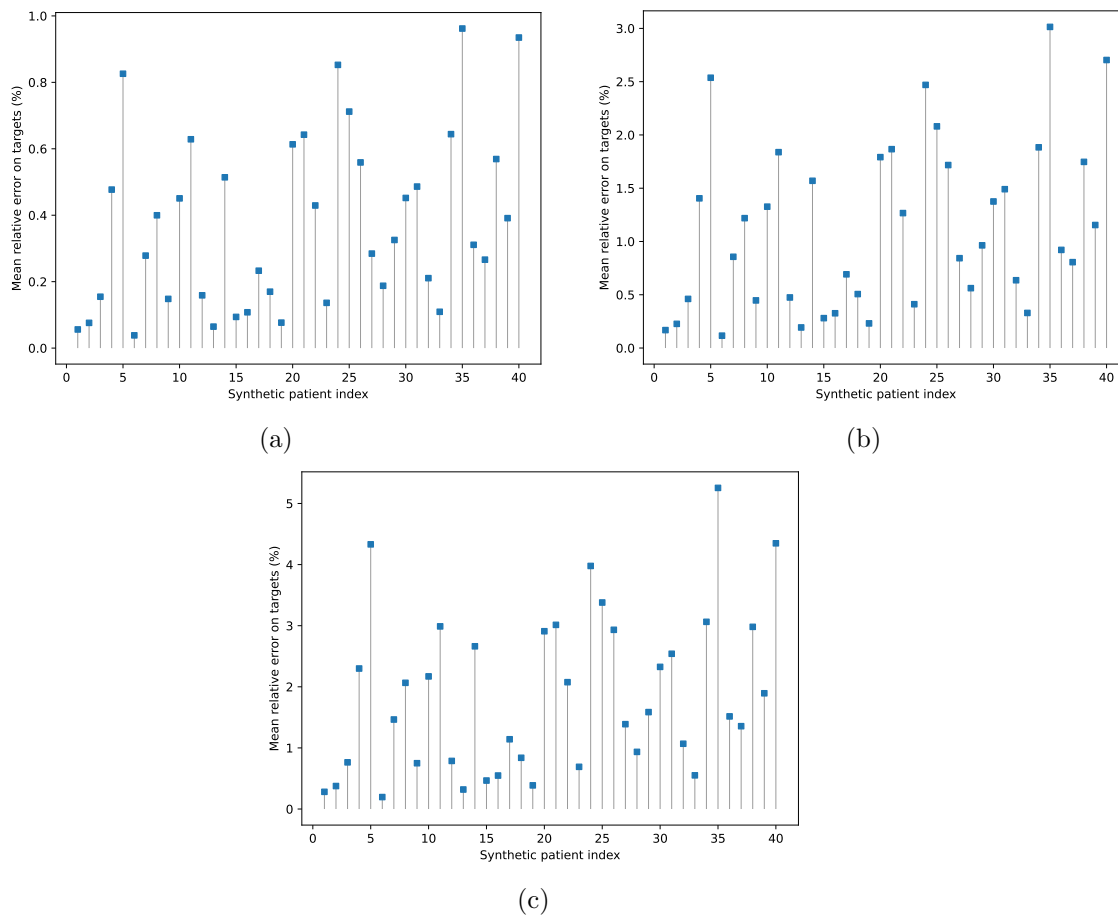
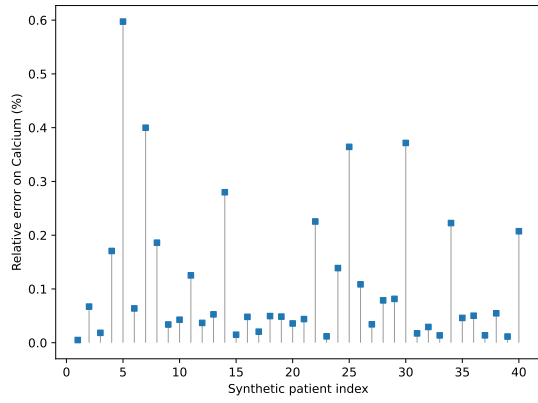
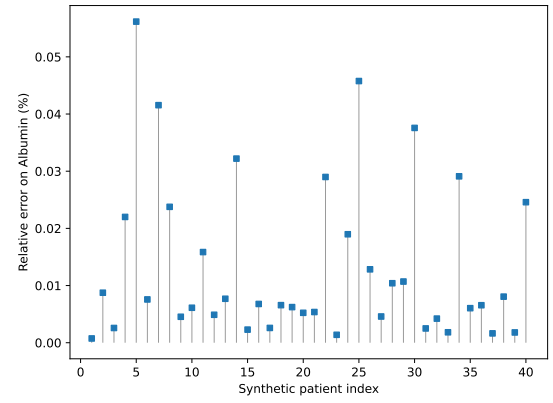


Figure 12: Resulting mean noise on concentrations for synthetic cohort with an initial noise of order (a) 1% (b) 3% and (c) 5%.

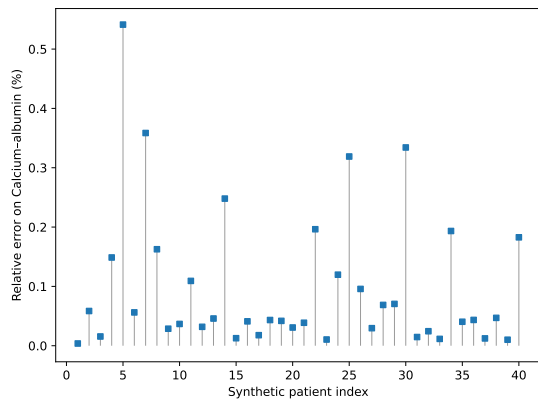
A finer analysis reveals that the five observable quantities do not respond uniformly to parameter perturbations. Averaged over patients, the largest sensitivities are consistently observed on the citrate-related targets, while albumin-related quantities remain weakly affected. For instance, for 1% perturbations, the mean relative errors are approximately 0.88% and 0.78% for citrate and calcium-citrate on the output, compared to about 0.11%, 0.01%, and 0.10% for other concentrations. Similar proportions are observed for 3% and 5% perturbations, with citrate and calcium-citrate reaching mean errors of order 2.6%-4.3%, while the remaining targets stay below 0.6% on average. This distribution is consistent with the structure of the reaction network: citrate and calcium-citrate species depend more directly on membrane diffusion, whereas albumin-related components are only indirectly affected. Figures 13, 14 and 15 below show the noise on the output for each chemical species



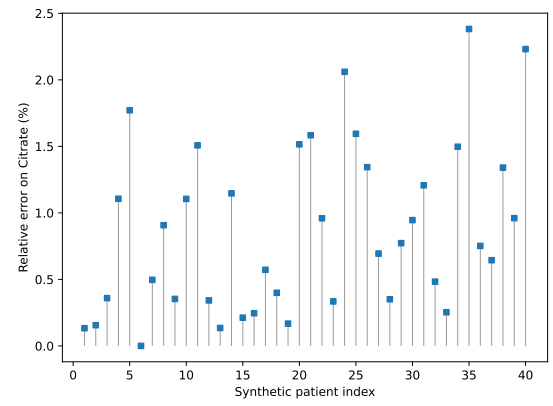
(a)



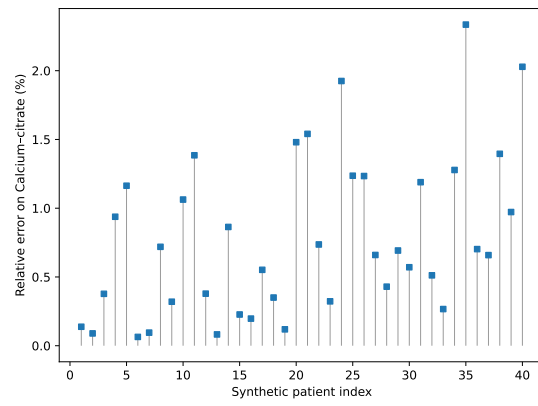
(b)



(c)

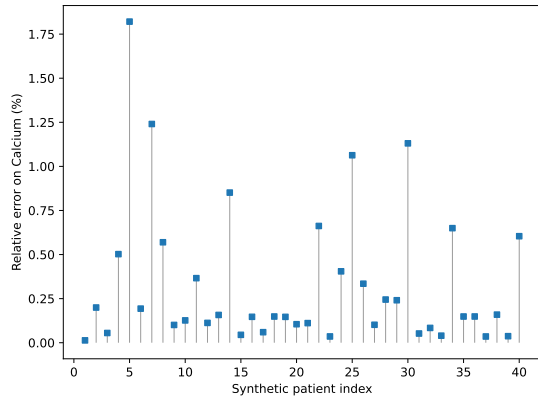


(d)

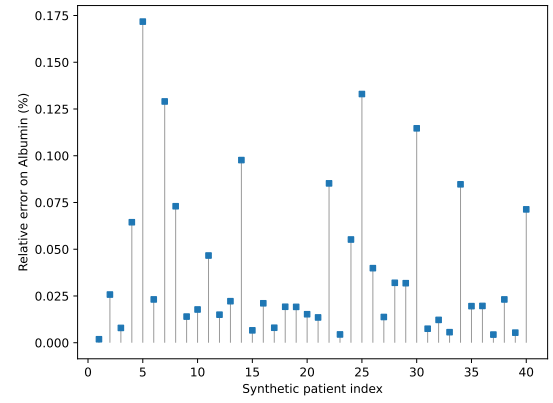


(e)

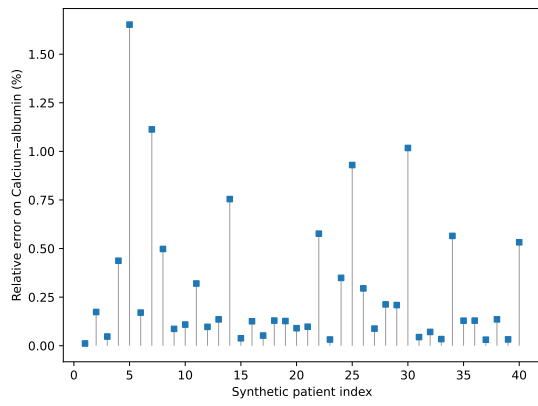
Figure 13: Resulting noise on each species for 1%



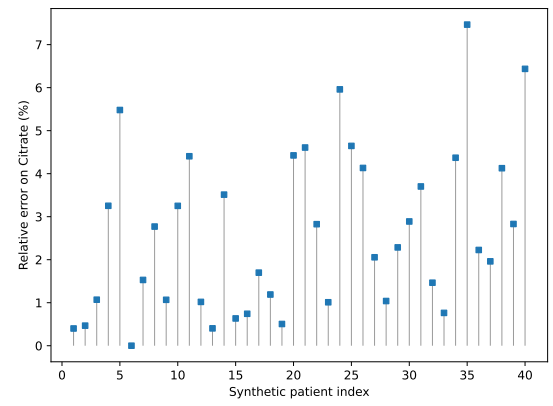
(a)



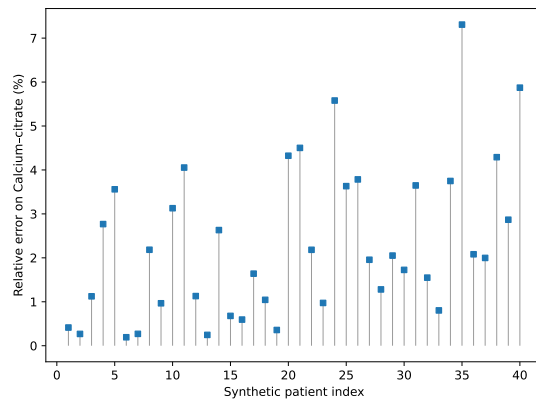
(b)



(c)

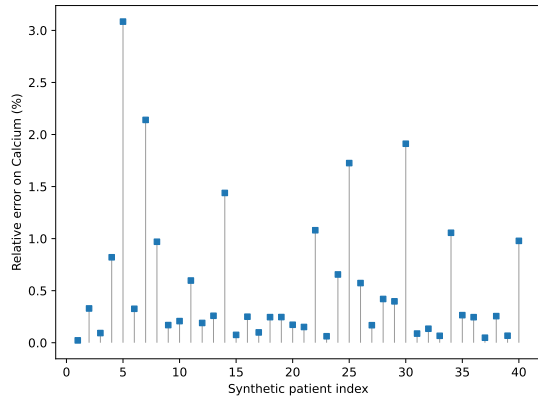


(d)

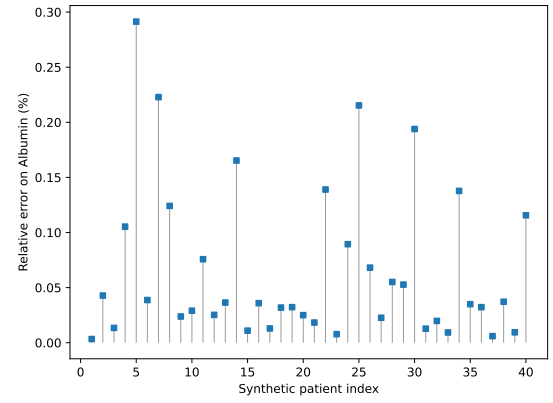


(e)

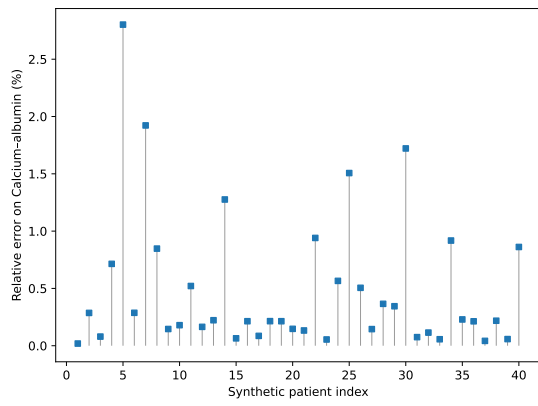
Figure 14: Resulting noise on each species for 3%



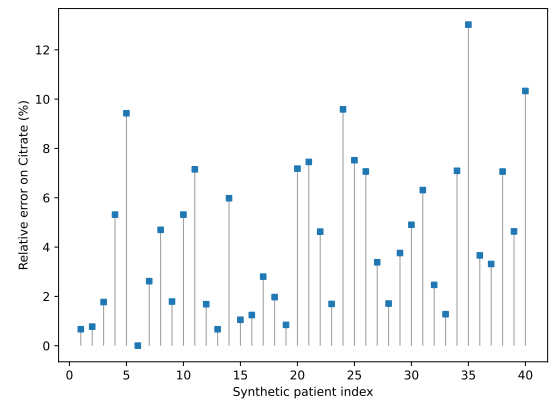
(a)



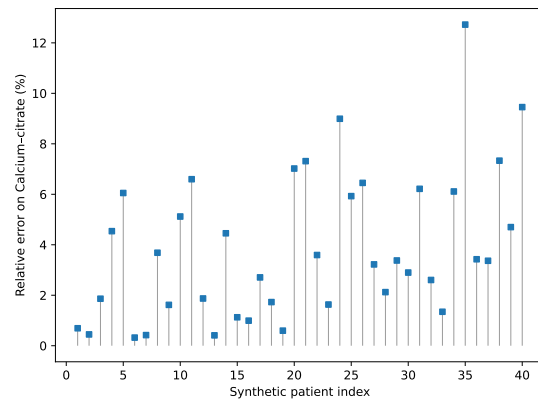
(b)



(c)



(d)



(e)

Figure 15: Resulting noise on each species for 5%

Overall, this sensitivity study shows that moderate perturbations of  $(d_{Ca}, d_{Ci})$  induce comparable or smaller relative variations in the model outputs. The absence of amplification supports the

numerical stability of the forward problem and reinforces the numerically well-posed condition of the inverse identification in the multi-patient setting. Moreover, the target-wise analysis highlights which observables carry the strongest information about membrane transport properties, providing guidance for weighting strategies in the least-squares functional. Also, the heterogeneous sensitivity observed across the five targets provides a direct explanation for the anisotropic geometry of the objective functional reported in Figure 5. Since citrate-related observables exhibit the strongest dependence on membrane diffusion, variations in  $d_{C_i}$  produce comparatively steep changes in  $\mathcal{J}$ , leading to a well-defined minimum in this direction. In contrast, the weaker response of the remaining targets to variations in  $d_{C_a}$  results in a flatter landscape along this axis. This imbalance naturally generates the elongated valley structure observed in the  $(d_{C_a}, d_{C_i})$  plane, with strong identifiability in  $d_{C_i}$  and partial correlation in  $d_{C_a}$ . From an inverse-problem perspective, this confirms that the geometry of  $\mathcal{J}$  is primarily driven by the differential sensitivity of the observable species, and that multi-patient aggregation is essential to stabilize the reconstruction along the weakly informed calcium direction.

## 4 Application to Clinical Data

We now turn to the application of the proposed inverse identification framework to clinical data obtained in the context of the MARC study (ClinicalTrials.gov identifier: NCT04530175), conducted at CHU Gabriel Montpied (Clermont-Ferrand, France). The cohort consists of twenty-two patients undergoing intermittent hemodialysis, for whom inlet concentrations and flow rates are available, together with outlet blood concentrations of calcium and citrate. As described in the introduction, these data provide sparse but clinically relevant observations of membrane transport processes under real operating conditions. Our objective is to identify effective membrane diffusion coefficients for calcium and citrate, denoted by

$$\beta = (d_{C_a}, d_{C_i}),$$

by minimizing the multi-patient least-squares functional (10), which aggregates model-data misfits over all patients and measured chemical species.

As a first exploratory step, we perform a coarse grid search over the rectangular parameter domain  $[0.02, 3] \times [0.02, 3]$  in the  $(d_{C_a}, d_{C_i})$  parameters plane. This global scan reveals a pronounced valley-shaped structure of the objective functional, characterized by a relatively sharp sensitivity with respect to  $d_{C_i}$  and a much flatter dependence on  $d_{C_a}$ . In particular, the lowest values of  $\mathcal{J}$  are consistently observed in a neighborhood of  $d_{C_i} \approx 0.6$  relevant with the previous observations about a strong identifiability of the citrate diffusion coefficient, while calcium diffusion appears less tightly constrained see Figure 16.

To assess the robustness of this geometric structure and to exclude the presence of local oscillations or secondary minima, we subsequently perform a localized grid search restricted to a narrower region around the valley  $[0.25, 2.50] \times [0.52, 0.72]$ . The resulting refined exploration confirms that the minimum of  $\mathcal{J}$  lies within this corridor and that the objective surface is smooth in this region, with no evidence of competing local minima.

These observations provide strong numerical support for the stability of the inverse problem in the vicinity of the identified valley. Building upon these grid-based results, we initialize a derivative-free Powell optimization within the valley region. The algorithm converges toward a minimizer

$$\beta^* = (d_{C_a}, d_{C_i}) \approx (0.5043764449030683, 0.6067775575575455),$$

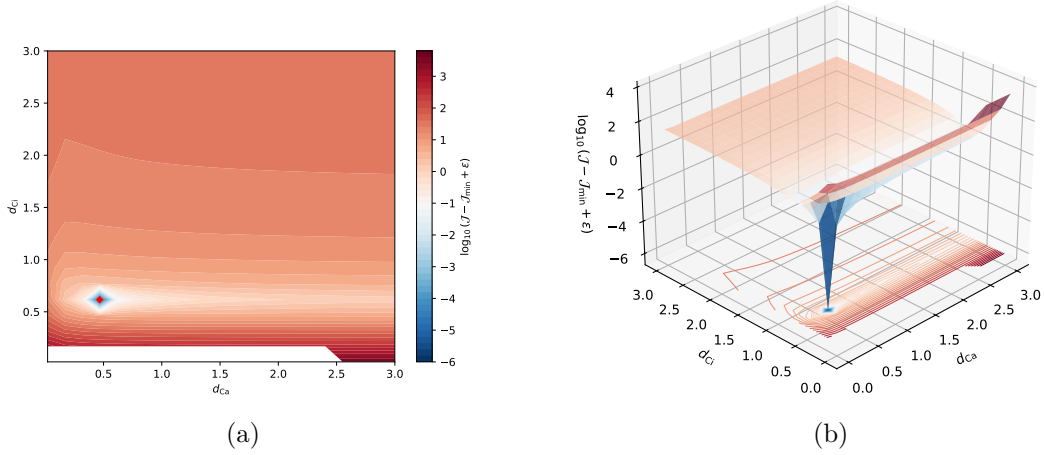


Figure 16: Coarse Grid Search in logarithmic scale (a) two-dimensional representation, the grid minimum is the red dot (b) surface representation in 3D.

which is fully consistent with both the global and localized grid searches. The corresponding value of the objective functional is approximately

$$\mathcal{J}(\beta^*) \approx 12.279445438717667.$$

While this value represents the lowest misfit attained within the explored parameter space, it remains significantly above zero. This residual discrepancy should not be interpreted as a failure of the identification procedure, but rather reflects several intrinsic limitations of the clinical inverse problem. First, outlet concentration measurements are affected by experimental noise and inter-session variability. Second, the forward model relies on effective transport parameters that lump together membrane microstructure, unresolved physicochemical effects, and modeling simplifications. We note that part of the uncertainty in the estimated parameters may also stem from device-related variability, in the sense that although the study was conducted using hollow-fiber dialyzer cartridges equipped with synthetic polysulfone membranes from the same manufacturer, different cartridge series may have been used in clinical practice, and the effective diffusion properties of the membrane can vary across series. Finally, inter-patient physiological variability cannot be fully captured by a single pair of global coefficients. Consequently, the identified parameters must be understood as effective diffusion coefficients, providing the best compromise in a least-squares sense across the entire cohort. In this context, the relatively elevated value of  $\mathcal{J}(\beta^*)$  quantifies a combined effect of measurement uncertainty, model mismatch, and biological heterogeneity, rather than purely numerical suboptimality. In this multi-patient setting, deriving confidence intervals on the estimated parameters directly from  $\mathcal{J}(\beta^*)$  would therefore be misleading, as it would implicitly attribute all discrepancies to observational noise, while significant contributions arise from modeling assumptions and population heterogeneity. As a result, the present framework provides effective population-averaged transport coefficients together with qualitative identifiability information, but does not support a rigorous uncertainty quantification at this stage. Achieving statistically meaningful confidence bounds would require extensive forward evaluations or sampling-based strategies, which are currently computationally prohibitive due to the cost of the underlying partial differential equation model. This limitation

motivates the development of surrogate models or reduced-order representations, as discussed in the concluding section.

Despite these limitations, the inverse pipeline exhibits a coherent and reproducible behavior on real clinical data. More precisely, the geometry of the objective functional is well structured, the identified valley is stable across grid resolutions, and the derivative-free optimization consistently converges toward the same parameter region. These features support the practical identifiability of citrate transport and provide a meaningful estimate of calcium diffusion under clinical operating conditions. Beyond parameter calibration, this result constitutes a first step toward predictive patient-specific simulations. Once effective membrane transport properties are identified, the model may be used to explore the impact of alternative dialysate compositions or operating regimes on calcium balance, thereby opening perspectives for personalized dialysis strategies.

## 5 Conclusion and perspectives

In this work, we developed a multi-patient inverse modeling framework for the identification of effective calcium and citrate diffusion coefficients in hollow-fiber hemodialysis devices. Building upon previous modeling efforts, a first methodological contribution is the direct computation of steady-state solutions of the coupled convection-reaction-diffusion system using a Newton method, rather than time-marching transient simulations. This stationary formulation significantly reduces computational cost while preserving accuracy, making repeated forward evaluations feasible within an inverse identification loop. A second major contribution is the extension from single-patient calibration to a multi-patient inverse setting. Using a synthetic cohort constructed from clinical statistics, we formulated a global least-squares objective aggregating residuals across several patients and solved the resulting low-dimensional nonlinear optimization problem using derivative-free methods. Unsurprisingly, numerical experiments with exact synthetic data demonstrate that aggregating information from multiple patients restores numerical identifiability of the diffusion coefficients and mitigates parameter correlation effects that are observed in single-patient inversions. In particular, the combined analysis of grid-based exploration, sensitivity studies, and Powell optimization highlights a pronounced anisotropy in the objective landscape, with strong identifiability in the citrate diffusion direction and weaker sensitivity with respect to calcium.

The robustness study conducted on synthetic cohorts further strengthens the practical relevance of the proposed inverse framework. By introducing controlled multiplicative perturbations on the outlet concentrations, we demonstrated that moderate measurement noise (up to 5% relative standard deviation) does not induce a significant amplification in the reconstructed effective diffusion coefficients. While the dispersion of the estimates increases with the noise level, the identified parameters remain concentrated in a neighborhood of the ground-truth values, with a variability that reflects the intrinsic anisotropy of the objective landscape. In particular, the weaker practical identifiability of  $d_{Ca}$  compared to  $d_{Ci}$  consistently manifests under noisy conditions. These results indicate that the multi-patient formulation provides a stabilizing effect and that the inverse problem remains well-behaved in clinically realistic noise regimes.

Beyond exploring numerical identifiability and robustness to noise on the targets, we investigated the numerical sensitivity of the forward model with respect to perturbations of the diffusion coefficients. By propagating controlled parameter noise through the full forward solver, we showed that relative variations in the observable outputs remain of the same order as the imposed coeffi-

cient perturbations, and often slightly smaller. This behavior indicates a moderate conditioning of the forward map and supports the numerical stability of the inverse problem in the considered regime. Target-wise analysis further revealed that citrate-related observables dominate the sensitivity, providing a mechanistic explanation for the valley structure observed in the objective functional.

When applied to real clinical data, the inverse procedure yields effective diffusion coefficients that are consistent across patients and compatible with the geometric structure observed in the synthetic setting. The non-zero residual value of the objective functional reflects several unavoidable factors, as measurement uncertainty, inter-patient physiological variability, and modeling approximations inherent to the reduced partial differential equations description. In particular, the use of effective diffusion parameters can be interpreted as a population-averaged representation of heterogeneous transport phenomena occurring at the fiber scale. Despite these simplifications, the identified coefficients lie in a coherent region of parameter space and reproduce the observed outlet concentrations with satisfactory accuracy. This suggests that the proposed framework captures the dominant transport mechanisms while providing clinically interpretable effective parameters.

From a practical perspective, the main computational bottleneck of the present approach remains the repeated evaluation of the high-fidelity forward model, which involves solving coupled fluid dynamics and nonlinear transport equations for each patient. While tractable in the current two-parameter setting, this cost may become prohibitive for larger cohorts, extended parameter spaces, or real-time clinical applications. A natural perspective is therefore the construction of surrogate models to approximate the parameter-to-observable map. Physics-informed neural networks (PINNs) offer a promising route in this direction, as they can embed the governing equations directly into the training process and provide fast evaluations once trained. In the present context, PINNs could be used either to emulate the full forward solver or to approximate the reduced mapping  $(d_{Ca}, d_{Ci}) \mapsto F_p(\beta)$  for each patient, enabling rapid inverse iterations and uncertainty quantification through ensemble evaluations. As an alternative or complementary approach, polynomial or low-rank approximation techniques may also be considered to build reduced-order surrogates of the forward map. Methods based on sparse polynomial expansions or adaptive low-dimensional representations offer strong theoretical guarantees for smooth parametric partial differential equations and can provide interpretable approximations with explicit error control. Such approaches could be particularly attractive in the present low-dimensional parameter setting and would allow systematic exploration of the objective landscape at negligible computational cost once the surrogate is constructed. In particular, Physics-Informed Neural Networks provide a flexible surrogate framework that has already proven effective in several biomedical and microbiological modeling settings, see e.g. the work of P.J.Hossie, B.Laroche, T.Malou, L.Perrin, T.Saigre & L.Sala [17].

Overall, the present framework enables the quantitative identification, from available clinical data, of effective transport properties governing chemical exchanges across the dialysis membrane. These coefficients can be interpreted as intrinsic characteristics of a given dialyzer cartridge, capturing in an aggregated manner membrane permeability, microstructure, and unmodeled effects. Once such parameters are estimated with sufficient accuracy, they provide a calibrated mechanistic model of solute transfer through the biomembrane. From a clinical and mathematical perspective, this opens the door to an optimal control formulation of dialysis treatment. More precisely, given patient-specific inlet blood concentrations, the calibrated forward model may be used to determine dialysate compositions that optimize target concentrations at the outlet of the dialyzer. In this setting, membrane transport coefficients play the role of fixed system

parameters, while dialysate inputs act as control variables. This perspective naturally leads toward personalized dialysis protocols, in which treatment parameters are adapted to individual patients based on predictive simulations, rather than fixed empirical prescriptions.

## References

- [1] J. Aniort, L. Chupin & N. Cîndea: *Mathematical model of calcium exchange during haemodialysis using a citrate containing dialysate*. Math. Med. Biol. (2018).
- [2] J. Aniort, F. Richard, F. Thouy, L. Le Guen, C. Philipponnet, C. Garrouste, A.-E. Heng, C. Dupuis, M. Adda, D. Julie, E. Lebretonchel, L. Chupin, D. Bouvier, B. Souweine & N. Cîndea: *Deciphering simplified regional anticoagulation with citrate in intermittent hemodialysis: a clinical and computational study*. Scientific Reports (2024).
- [3] N. Cancilla, L. Gurreri, M. Ciofalo, A. Cipollina, A. Tamburini & M. Giorgio: *Mathematical modelling of hollow-fiber haemodialysis modules*. Physical Sciences Reviews (2025).
- [4] N. Cancilla, L. Gurreri, G. Marotta, M. Ciofalo, A. Cipollina, A. Tamburini & G. Micale: *A porous media CFD model for the simulation of hemodialysis in hollow fiber membrane modules*. Journal of Membrane Science (2022).
- [5] Y.-L. Chang & C.-J. Lee: *Solute transport characteristics in hemodiafiltration*. Journal of Membrane Science (1988).
- [6] P. Ciarlet: *Introduction to numerical linear algebra and optimisation*. Cambridge University Press, Cambridge Texts in Applied Mathematics (1989).
- [7] CKD-MBD Work Group and others: *KDIGO clinical practice guideline for the diagnosis, evaluation, prevention, and treatment of Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD)*. Kidney international. Supplement (2009).
- [8] A. Conn, K. Scheinberg & L. Vicente: *Introduction to derivative-free optimization*. MPS/SIAM Ser. Optim. (2009).
- [9] P. Davidson: *Dynamics of Rotating Fluids*. Oxford University Press (2024).
- [10] D. Donato, A. Boschetti-de-Fierro, C. Zweigart, M. Kolb, S. Eloot, M. Storr, B. Krause, K. Leyboldt & P. Segers: *Optimization of dialyzer design to maximize solute removal with a two-dimensional transport model*. Journal of Membrane Science (2017).
- [11] D. Donato, M. Storr & B. Krause: *Design optimization of hollow fiber dialyzers to enhance internal filtration based on a mathematical model*. Journal of Membrane Science (2020).
- [12] S. Eloot, J. Vierendeels & P. Verdonck: *Optimisation of solute transport in dialyzers using a three-dimensional finite volume model*. Computer Methods in Biomechanics and Biomedical Engineering (2006).
- [13] S. Eloot, D. De Wachter, I. Van Tricht & P. Verdonck: *Computational Flow Modeling in Hollow-Fiber Dialyzers*. Artificial Organs (2002).
- [14] A. Fasano & A. Farina: *Modeling high flux hollow fibers dialyzers*. Discrete and Continuous Dynamical Systems - B (2012).

- [15] R. Fournier: *Basic Transport Phenomena in Biomedical Engineering*. CRC Press, Third Edition (2011).
- [16] J. Hall & M. Hall: *Guyton and Hall Textbook of Medical Physiology* Elsevier Press, 14th Edition (2020).
- [17] P. J. Hossie, B. Laroche, T. Malou, L. Perrin, T. Saigre & L. Sala: *Surrogate modeling of interactions in microbial communities through Physics-Informed Neural Networks*. ESAIM: ProcS (2025).
- [18] F. Hecht: *New development in FreeFem++*, J. Numer. Math. (2012).
- [19] J. Himmelfarb, T. Ikizler & T. Alp: *Hemodialysis*. New England Journal of Medicine (2010).
- [20] M. Islam & J. Szpunar: *Study of Dialyzer Membrane (Polyflux 210H) and Effects of Different Parameters on Dialysis Performance*. Open Journal of Nephrology (2013).
- [21] M. Jaffrin, B. Gupta & J. Malbrancq: *A One-Dimensional Model of Simultaneous Hemodialysis and Ultrafiltration With Highly Permeable Membranes*. ASME. J Biomech Eng. (1981).
- [22] M. Ketteler, P. Evenepoel, R. Holden, T. Isakova, H. Jørgensen, H. Komaba, T. Nickolas, S. Sinha, M. Vervloet, M. Cheung, J. King, M. Grams, M. Jadoul & R. Moysés: *Conference Participants. Chronic kidney disease-mineral and bone disorder: conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference*. Kidney Int. (2025).
- [23] C. Legallais, G. Catapano, B. von Harten & U. Baurmeister: *A theoretical model to predict the in vitro performance of hemodiafilters*. Journal of Membrane Science (2000).
- [24] V. Maheshwari, A. Cherif, D. Fuertringer, G. Schappacher-Tilp, P. Preciado, S. Thijssen, D. Bushinsky & P. Kotanko: *An in silico method to predict net calcium transfer during hemodialysis*. Annu Int Conf IEEE Eng Med Biol Soc. (2017).
- [25] T. Miyasaka & K. Sakai: *Application of mathematical analysis on dialysis*. J Artif Organs (2023).
- [26] C. Ronco & W. Clark: *Haemodialysis membranes*. Nat. Rev. Nephrol. (2018).
- [27] D. Schneditz, D. Platzer & J. Daugirdas: *A diffusion-adjusted regional blood flow model to predict solute kinetics during haemodialysis*. Nephrology Dialysis Transplantation (2009).
- [28] G. Truskey, F. Yuan & D. Katz: *Transport Phenomena in Biological Systems* Pearson Publishing, 2nd Edition (2008).
- [29] A. Werynski & J. Waniewski: *Theoretical Description of Mass Transport in Medical Membrane Devices*. Artificial Organs (1995).
- [30] T. Yaqoob, M. Ahsan, A. Hussain & I. Ahmad: *Computational Fluid Dynamics (CFD) Modeling and Simulation of Flow Regulatory Mechanism in Artificial Kidney Using Finite Element Method*. Membranes (2020).
- [31] J. Zhang, X. Chen, J. Ding, K. Fraser, M. Ertan Taskin, B. Griffith & Z. Wu: *Computational Study of the Blood Flow in Three Types of 3D Hollow Fiber Membrane Bundles*. ASME. J. Biomech. Eng. (2013).