



MOX-Report No. 39/2024

**Reconstruction of unknown nonlinear operators in semilinear elliptic
models using optimal inputs**

Bartsch, J.; Buchwald, S.; Ciaramella, G.; Volkwein, S.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it

<https://mox.polimi.it>

RECONSTRUCTION OF UNKNOWN NONLINEAR OPERATORS IN SEMILINEAR ELLIPTIC MODELS USING OPTIMAL INPUTS

JAN BARTSCH

Department of Mathematics, University of Konstanz, Universitätsstr. 10, 78464 Konstanz, Germany

SIMON BUCHWALD

Department of Mathematics, University of Konstanz, Universitätsstr. 10, 78464 Konstanz, Germany

GABRIELE CIARAMELLA

*MOX, Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133,
Milano, Italy*

STEFAN VOLKWEIN

Department of Mathematics, University of Konstanz, Universitätsstr. 10, 78464 Konstanz, Germany

ABSTRACT. Physical models often contain unknown functions and relations. The goal of our work is to answer the question of how one should excite or control a system under consideration in an appropriate way to be able to reconstruct an unknown nonlinear relation. To answer this question, we propose a greedy reconstruction algorithm within an offline-online strategy. We apply this strategy to a two-dimensional semilinear elliptic model. Our identification is based on the application of several space-dependent excitations (also called controls). These specific controls are designed by the algorithm in order to obtain a deeper insight into the underlying physical problem and a more precise reconstruction of the unknown relation. We perform numerical simulations that demonstrate the effectiveness of our approach which is not limited to the current type of equation. Since our algorithm provides not only a way to determine unknown operators by existing data but also protocols for new experiments, it is a holistic concept to tackle the problem of improving physical models.

1. INTRODUCTION

The identification of unknown operators in semilinear systems plays a fundamental role in various scientific disciplines, including physics, chemistry, and biology [4, 15, 34]. Accurately characterizing these operators is crucial for understanding and predicting complex system behavior. However, they often remain unknown or only partially known, necessitating efficient and reliable reconstruction methods.

E-mail addresses: jan.bartsch@uni-konstanz.de, simon.buchwald@uni-konstanz.de,
gabriele.ciaramella@polimi.it, stefan.volkwein@uni-konstanz.de.

Date: May 22, 2024.

2020 Mathematics Subject Classification. 35K57, 35R30, 49N45, 49M41, 93B30.

Key words and phrases. Reaction-diffusion equations, greedy reconstruction algorithm, optimal control of partial differential equations, semilinear partial differential equations.

This work was financed by the Deutsche Forschungsgemeinschaft (DFG) within SFB 1432, Project-ID 425217212. The present research is part of the activities of “Dipartimento di Eccellenza 2023-2027”.

Many techniques exist for such identification, drawing from the fields of inverse problems and optimization. Traditional approaches involve inverse problem methods, inferring the operator from measured outputs under controlled inputs. Techniques like gradient descent, least squares minimization, or Bayesian approaches are commonly employed [19, 27]. Optimization methods formulate the identification problem as an optimization task, minimizing the discrepancy between model outputs and data.

With the recent surge of machine learning, neural networks are increasingly applied to identification tasks [25, 31]. While these methods have achieved significant progress, a common challenge lies in data quality. Real-world data often suffers from noise and uncertainties, leading to inaccuracies in the reconstructed operator. Additionally, limited data availability due to cost or time constraints can hinder accurate identification. Furthermore, biased data can lead to biased estimates of the operator, necessitating careful data collection and analysis procedures.

This paper extends the approach introduced in [26] and optimized in [7] to reconstruct unknown nonlinear operators in semilinear elliptic models, addressing some of the challenges mentioned above. We leverage the power of optimal control techniques, formulating the reconstruction problem as an optimization problem where the objective function measures the discrepancy between the model output and observed data. By strategically designing control functions and employing optimization algorithms, we aim to drive the model output toward observed data that provides significant insights into the underlying physical system, thereby revealing the characteristics of the unknown operator. Our method develops optimal protocols to generate data that yields as much insight into the system as possible. In other words, the goal of the algorithm is to convexify the identification problem in order to be able to find a unique solution and speed up the solution process. This optimized algorithm was used to identify different linear and nonlinear unknown operators and distributions [8, 9, 10]. We extend the existing results by considering a system of elliptic partial differential equations (PDEs) and reconstructing infinite-dimensional objects.

The proposed approach offers several advantages. Unlike traditional methods that require large amounts of high-quality data, our framework incorporates an active learning strategy. The algorithm selects new data points to collect, focusing on those that provide the most information about the unknown operator, leading to reduced data dependency. Additionally, the framework provides theoretical guarantees for convergence to the true operator under certain conditions.

Building upon this foundation, we present a computationally efficient numerical scheme to solve the formulated optimization problem, enabling implementation for various semilinear elliptic models. Our semilinear PDE can be motivated as the steady-state solution to PDEs describing models from epidemics, biochemical systems, and nuclear reactor models or general reaction-diffusion models [28, 30]. Similar models are also used to describe membrane-resonators; see, e.g., [37]. We refer to [2, 22, 23, 33] and the references therein for existing investigations on the identification of unknown nonlinearities in semilinear PDEs. Furthermore, we refer the reader to the work [21] on general inverse problems with PDEs.

The paper is organized as follows. In the next section, we introduce our specific setting including the formulation of our reconstruction problem. In Section 3, we analyze the semilinear PDE model and state the existence and uniqueness of solutions in Theorem 3.6. Furthermore, we present continuity results for the control-to-state map, the parameter-to-state map, and the inverse parameter-to-state map in Theorems 3.8, 3.11 and 3.12, respectively. In Section 4, we describe our numerical method to solve the reconstruction problem using a greedy approach (cf. Algorithm 4.1). Afterwards, in Section 5, we report the numerical approximation in terms of spatial discretization and numerical strategy to solve the governing equation. To validate our strategy, we perform numerical experiments and report the results in Section 6. A section of conclusions finishes this work.

Notation. For a natural number $n \in \mathbb{N}$, we define $[n] := \{1, \dots, n\}$. For a vector $\alpha \in \mathbb{R}^n$, we denote by $\|\cdot\|_p$ the standard ℓ^p norms for $1 \leq p \leq \infty$. Furthermore, we denote for vectors $\alpha, \tilde{\alpha} \in \mathbb{R}^n$ the standard scalar product in \mathbb{R}^n as $\alpha^\top \tilde{\alpha} = \alpha \cdot \tilde{\alpha} = \langle \alpha, \tilde{\alpha} \rangle_2$. Whenever two vectors in \mathbb{R}^n are compared with each other, then the comparison has to be understood componentwise.

For a domain $\Omega \in \mathbb{R}^n$ and $1 \leq p \leq \infty$, we set $L^p(\Omega) := L^p(\Omega; \mathbb{R})$ and $\mathbb{L}^p(\Omega) := L^p(\Omega) \times L^p(\Omega)$; analogously, we define $\mathbb{H}^k(\Omega) = H^k(\Omega) \times H^k(\Omega)$ for $k \in \mathbb{N}$. Furthermore, we set $\mathbb{H}_0^1(\Omega) := H_0^1(\Omega) \times$

$H_0^1(\Omega)$ and $\mathbb{C}(\bar{\Omega}) := C(\bar{\Omega}) \times C(\bar{\Omega})$. The space $H_0^1(\Omega)$ is equipped with the inner product

$$\langle \varphi, \psi \rangle_{H_0^1(\Omega)} := \int_{\Omega} \nabla \varphi \cdot \nabla \psi \, dx \quad \text{for } \varphi, \psi \in H_0^1(\Omega).$$

For more details, we refer to the standard notation for Lebesgue and Sobolev spaces; see, e.g., [1, 16].

2. FORMULATION OF THE PROBLEM

We consider a convex two-dimensional spatial domain $\Omega \subsetneq \mathbb{R}^2$ with Lipschitz-continuous boundary $\Gamma = \partial\Omega$. For $x = (x_1, x_2)^\top \in \Omega$ and $\mathbf{y}(x) = (y_1(x), y_2(x))^\top$, we consider the following system of semilinear PDEs with homogeneous Dirichlet boundary conditions

$$(2.1a) \quad -\Delta \mathbf{y}(x) + \mathbf{g}(\mathbf{y}(x)) = \boldsymbol{\varepsilon}(x) \quad \text{in } \Omega,$$

$$(2.1b) \quad \mathbf{y}(x) = 0 \quad \text{on } \Gamma,$$

where $\boldsymbol{\varepsilon} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a control input and $\mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ an unknown nonlinearity. For $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2)^\top$, we define the set of admissible controls as

$$\mathcal{E}_{\text{ad}} := \{ \boldsymbol{\varepsilon} \in \mathbb{L}^2(\Omega) \mid \boldsymbol{\varepsilon}_a \leq \boldsymbol{\varepsilon}(x) \leq \boldsymbol{\varepsilon}_b \text{ for almost all (f.a.a.) } x \in \Omega \},$$

with given bounds $\boldsymbol{\varepsilon}_a, \boldsymbol{\varepsilon}_b \in \mathbb{R}^2$, $\boldsymbol{\varepsilon}_a \leq \boldsymbol{\varepsilon}_b$. Notice that the elements in $\mathcal{E}_{\text{ad}} \subset \mathbb{L}^\infty(\Omega)$ are uniformly bounded in $\mathbb{L}^\infty(\Omega)$ by $\varepsilon_{\max} := \max\{\|\boldsymbol{\varepsilon}_a\|_\infty, \|\boldsymbol{\varepsilon}_b\|_\infty\}$.

The nonlinearity \mathbf{g} is assumed to lie in a K -dimensional space \mathfrak{G} . Thus, given a basis $\mathcal{G} = \{\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_K\}$ of \mathfrak{G} with functions $\boldsymbol{\varphi}_i : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, we can write

$$(2.2) \quad \mathbf{g}(y) = \mathbf{g}_\alpha(y) := \sum_{j=1}^K \alpha_j \boldsymbol{\varphi}_j(y) = \boldsymbol{\Phi}(y) \boldsymbol{\alpha} \in \mathbb{R}^2, \quad y \in \mathbb{R}^2,$$

for some coefficient vector $\boldsymbol{\alpha} = (\alpha_j)_{j=1}^K \in \mathbb{R}^K$ and $\boldsymbol{\Phi}(y) = (\boldsymbol{\varphi}_1(y), \dots, \boldsymbol{\varphi}_K(y)) \in \mathbb{R}^{2 \times K}$. Our true unknown nonlinearity is now assumed to be given by

$$\mathbf{g}_{\boldsymbol{\alpha}^*}(y) = \sum_{j=1}^K \alpha_j^* \boldsymbol{\varphi}_j(y)$$

for an unknown coefficient vector $\boldsymbol{\alpha}^* = (\alpha_j^*) \in \mathbb{R}^K$. Notice that \mathbf{g} is locally Lipschitz continuous if all elements in \mathcal{G} are polynomials or piecewise linear and continuous functions.

For the coefficients $\boldsymbol{\alpha} \in \mathbb{R}^K$, we assume that they are elements of

$$\mathcal{A}_K := \{ \boldsymbol{\alpha} \in \mathbb{R}^K \mid 0 \leq \alpha_i \leq \alpha_{\max} \text{ for } i \in [K] \},$$

with a scalar $\alpha_{\max} \geq 0$. Notice that \mathcal{A}_K is nonempty, compact and convex. Further, we denote by $\mathbf{y}^{\boldsymbol{\alpha}, \boldsymbol{\varepsilon}}$ the solution to (2.1) using the nonlinearity \mathbf{g}_α and applying controls (inputs) $\boldsymbol{\varepsilon} \in \mathcal{E}_{\text{ad}}$ (cf. Section 3). To identify the unknown true nonlinearity $\mathbf{g}_{\boldsymbol{\alpha}^*}$, we generate a K -dimensional set of control functions $\{\boldsymbol{\varepsilon}^m\}_{m=1}^K \subset \mathcal{E}_{\text{ad}}$ to perform K (laboratory) experiments and obtain the data $\mathbf{y}_*^{\boldsymbol{\varepsilon}^m}$ for $m \in [K]$. Then the nonconvex parameter identification problem in order to estimate the nonlinearity $\mathbf{g}_{\boldsymbol{\alpha}^*}$ is given by

$$(2.3) \quad \min_{\boldsymbol{\alpha} \in \mathcal{A}_K} \sum_{m=1}^K \left\| \mathbf{y}^{\boldsymbol{\alpha}, \boldsymbol{\varepsilon}^m} - \mathbf{y}_*^{\boldsymbol{\varepsilon}^m} \right\|_{\mathbb{L}^2(\Omega)}^2.$$

Since $\boldsymbol{\alpha} \in \mathcal{A}_K$ is a vector in \mathbb{R}^K , we did not add an additional regularisation term depending on $\boldsymbol{\alpha}$. The controls $\{\boldsymbol{\varepsilon}^m\}_{m=1}^K \subset \mathcal{E}_{\text{ad}}$ will be chosen properly by the iterative Algorithm 4.1 presented in Section 4.

3. ANALYSIS OF THE GOVERNING EQUATION

From now on, we always assume a nonlinearity

$$\mathbf{g}(y) = \mathbf{g}_\alpha(y) = \sum_{j=1}^K \alpha_j \varphi_j(y), \quad \text{for } \alpha \in \mathcal{A}_K.$$

We omit the subscript α in the first part of this section. Now, we analyze the model (2.1). In particular, we recall results about the existence and uniqueness together with wellposedness results and continuity properties of the control-to-state map. Afterwards, we deal with the Lipschitz continuity (stability) of the equation with respect to the parameters α . Specifically, we formulate the wellposedness of the parameter-to-state map as well as of its inverse, the state-to-parameter map. We refer to [2, 22, 23, 33] for similar investigations. Furthermore, we refer the reader to the work [21] on general inverse problems governed by PDEs.

3.1. Existence of solutions and wellposedness. We apply the results of [11]; see also [36, Chapter 4]. For the basis functions and the nonlinearity \mathbf{g} , we state the following assumptions.

Assumption 3.1. 1) *The basis elements φ_j are continuously differentiable. This implies in particular that they are (locally) Lipschitz continuous, i.e., there exist constants $L_{\varphi_j} > 0$, such that for all $j \in [K]$ and $y, \tilde{y} \in \mathbb{R}^2$ it holds that*

$$\|\varphi_j(y) - \varphi_j(\tilde{y})\|_2 \leq L_{\varphi_j} \|y - \tilde{y}\|_2.$$

We set $L_\varphi := \max_{j \in [K]} L_{\varphi_j}$.

2) *The basis elements are bounded from below and above, i.e., there exist vectors $c_{\varphi_j}, C_{\varphi_j} \in \mathbb{R}^2$ with only positive entries and $c_{\varphi_j} < C_{\varphi_j}$, such that for all $j \in [K]$ and $y \in \mathbb{R}^2$ it holds that*

$$c_{\varphi_j} \leq \varphi_j(y) \leq C_{\varphi_j}.$$

We set $c_\varphi := \min_{j \in [K]} \|c_{\varphi_j}\|_\infty$ and $C_\varphi := \max_{j \in [K]} \|C_{\varphi_j}\|_\infty$.

3) *The basis elements are monotone non-decreasing functions and satisfy*

$$\forall M > 0 \exists \phi_M > 0 : \|\varphi_j(y)\|_2 \leq \phi_M \forall \|y\|_2 \leq M.$$

For the nonlinearity $\mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, we impose the following:

Assumption 3.2. *The Jacobian matrix $\nabla \mathbf{g}(y) \in \mathbb{R}^{2 \times 2}$ of the nonlinearity \mathbf{g} is positive semi-definite.*

These hypotheses on the nonlinearity and the definition of \mathcal{E}_{ad} lead to the following statement.

Corollary 3.3. *Suppose that Assumptions 3.1 and 3.2 hold true. Then it follows that*

- 1) *The nonlinearity \mathbf{g} and its derivative \mathbf{g}' are continuous in \mathbb{R}^2 .*
- 2) *The nonlinearity \mathbf{g} satisfies*

$$\forall M > 0 \exists \varrho_M > 0 : \|\mathbf{g}(y)\|_2 \leq \varrho_M \forall \|y\|_2 \leq M.$$

- 3) *The nonlinearity \mathbf{g} is non-decreasing.*

Proof. Part 1) follows directly from Assumption 3.1-1) and part 2) from Assumption 3.1-2). The part 3) follows from [32, Proposition 12.3]. \square

Remark 3.4. 1) *From Assumption 3.1-2) it follows that for all $j \in [K]$, there exist a constant $I_{\varphi_j} > 0$ such that for any bounded subset $\Upsilon \subset \mathbb{R}^2$ it holds*

$$(3.1) \quad \|\varphi_j(y)\|_{\mathbb{L}^2(\Omega)} \leq I_{\varphi_j}.$$

We define $I_\varphi := \max_{j \in [K]} I_{\varphi_j}$.

- 2) *We need C_{φ_j} for (3.1) and c_{φ_j} for Theorem 3.12.*
- 3) *The monotonicity of Corollary 3.3-3) is needed for the uniqueness of solutions to (2.1) in Theorem 3.6 below.*

Whenever we attempt to solve (2.1), we consider its weak formulation [36, Chapter 4]:

Definition 3.5. A function $\mathbf{y} = (y_1, y_2)^\top \in \mathbb{H}_0^1(\Omega)$ is called a weak solution of (2.1) if the following variational equality is fulfilled

$$(3.2) \quad \int_{\Omega} \langle \nabla \mathbf{y}(x), \nabla \mathbf{v}(x) \rangle_{\mathbb{F}} dx + \int_{\Omega} \mathbf{v}(x) \cdot \mathbf{g}(\mathbf{y}(x)) dx = \int_{\Omega} \mathbf{v}(x) \cdot \boldsymbol{\varepsilon}(x) dx \quad \forall \mathbf{v} \in \mathbb{H}_0^1(\Omega),$$

for fixed $\boldsymbol{\varepsilon} \in \mathcal{E}_{\text{ad}}$ and fixed \mathbf{g} fulfilling Assumptions 3.1 and 3.2. We introduce the bilinearity

$$a[\mathbf{y}, \mathbf{v}] := \int_{\Omega} \langle \nabla \mathbf{y}(x), \nabla \mathbf{v}(x) \rangle_{\mathbb{F}} dx \quad \text{for } \mathbf{y}, \mathbf{v} \in \mathbb{H}_0^1(\Omega).$$

In Definition 3.5, we denote with $\langle \cdot, \cdot \rangle_{\mathbb{F}}$ the Frobenius inner product of matrices and furthermore with $\nabla \mathbf{y}$ the Jacobian matrix of \mathbf{y} . An important property of the bilinearity $a[\cdot, \cdot]$ needed in the sequel is the existence of a constant $C_E > 0$ such that

$$a[\mathbf{v}, \mathbf{v}] \geq C_E \|\mathbf{v}\|_{\mathbb{H}_0^1(\Omega)}^2 \quad \forall \mathbf{v} \in \mathbb{H}_0^1(\Omega).$$

This property is also called the *ellipticity* of the bilinearity. In our specific case, it even holds that

$$(3.3) \quad a[\mathbf{v}, \mathbf{v}] = \|\mathbf{v}\|_{\mathbb{H}_0^1(\Omega)}^2 \quad \forall \mathbf{v} \in \mathbb{H}_0^1(\Omega).$$

Next, we formulate our existence and uniqueness result in the following statement. We introduce the solution space

$$\mathcal{Y} := \mathbb{H}^2(\Omega) \cap \mathbb{H}_0^1(\Omega),$$

equipped with the norm

$$\|\mathbf{v}\|_{\mathcal{Y}} := \|\Delta \mathbf{v}\|_{\mathbb{L}^2(\Omega)}.$$

We remark that the norm $\|\cdot\|_{\mathcal{Y}}$ is equivalent to the standard $\mathbb{H}^2(\Omega)$ norm; see, e.g., [18, Chapter 3]. Furthermore, recall that we use the following norm on $\mathbb{H}_0^1(\Omega)$

$$\|\mathbf{v}\|_{\mathbb{H}_0^1(\Omega)} := \|\nabla \mathbf{v}\|_{\mathbb{L}^2(\Omega)}.$$

Theorem 3.6. Suppose that Assumptions 3.1 and 3.2 hold. Then there exists a constant $C_0 > 0$, independent of $\boldsymbol{\varepsilon}$ and \mathbf{g} , such that for all controls $\boldsymbol{\varepsilon} \in \mathcal{E}_{\text{ad}}$ the equation (2.1) possesses a unique weak solution $\mathbf{y} \in \mathcal{Y}$ in the sense of Definition 3.5 satisfying

$$\|\mathbf{y}\|_{\mathcal{Y}} \leq C_0.$$

Proof. By the virtue of Corollary 3.3, since Ω is convex with Lipschitz boundary Γ , it is possible to apply [11, Theorem 2.10]. Notice in particular that all elements $\boldsymbol{\varepsilon} \in \mathcal{E}_{\text{ad}}$ are uniformly bounded by ε_{max} and that the value of \mathbf{g} at the origin is bounded by our assumption on the basis elements (cf. Assumption 3.1). \square

Remark 3.7. We refer to [5, 6, 35] for related discussions on the analysis of semilinear elliptic PDEs.

We are now able to introduce the *control-to-state* map \mathcal{S} for a fixed nonlinearity $\mathbf{g}_{\boldsymbol{\alpha}}$ for $\boldsymbol{\alpha} \in \mathcal{A}_K$ as

$$\mathcal{S}_{\boldsymbol{\alpha}} : \mathcal{E}_{\text{ad}} \rightarrow \mathcal{Y}, \quad \boldsymbol{\varepsilon} \mapsto \mathbf{y}^{\boldsymbol{\alpha}, \boldsymbol{\varepsilon}} := \mathcal{S}_{\boldsymbol{\alpha}}(\boldsymbol{\varepsilon}),$$

where $\mathcal{S}_{\boldsymbol{\alpha}}(\boldsymbol{\varepsilon})$ is the unique weak solution to (2.1) provided by Theorem 3.6 given the control $\boldsymbol{\varepsilon}$ and the nonlinearity $\mathbf{g}_{\boldsymbol{\alpha}}$. Therefore, we can conclude that $\mathcal{S}_{\boldsymbol{\alpha}}$ is well defined. Furthermore, we have the following result from [11, Lemma 3.5]:

Theorem 3.8. Let Assumptions 3.1 and 3.2 hold and fix a nonlinearity $\mathbf{g}_{\boldsymbol{\alpha}}$ for $\boldsymbol{\alpha} \in \mathcal{A}_K$. Then the control-to-state map $\mathcal{S}_{\boldsymbol{\alpha}}$ is Lipschitz continuous from $\mathbb{L}^2(\Omega)$ to \mathcal{Y} : There exists a constant $L > 0$, such that for all $\boldsymbol{\varepsilon}, \tilde{\boldsymbol{\varepsilon}} \in \mathcal{E}_{\text{ad}}$ it holds

$$\|\mathcal{S}_{\boldsymbol{\alpha}}(\boldsymbol{\varepsilon}) - \mathcal{S}_{\boldsymbol{\alpha}}(\tilde{\boldsymbol{\varepsilon}})\|_{\mathcal{Y}} \leq L \|\boldsymbol{\varepsilon} - \tilde{\boldsymbol{\varepsilon}}\|_{\mathbb{L}^2(\Omega)}.$$

We refer the reader to [36, Chapter 4] as well as [11, 12, 13] for more details.

3.2. Stability with respect to the parameters. In this section, we investigate the wellposedness and Lipschitz continuity (stability) of the parameter-to-state map defined below. In the following analysis, we take account of the iterative structure of the algorithm defined in [Section 4](#). More specifically, we consider a $k \in [K]$ that indicates the k -th iteration of the algorithm in the next results. We introduce for $1 \leq k \leq K-1$ the set

$$(3.4) \quad \mathcal{A}_k := \{\boldsymbol{\alpha} \in \mathcal{A}_K \mid \alpha_j = 0, j = k+1, \dots, K\}.$$

From [Theorem 3.6](#) and [Assumption 3.1](#), we deduce the following proposition.

Proposition 3.9. *Let [Assumptions 3.1](#) and [3.2](#) hold and $\varepsilon \in \mathcal{E}_{\text{ad}}$ be chosen arbitrarily. Introduce the parameter-to-state map \mathcal{P}_ε by*

$$\mathcal{P}_\varepsilon : \mathcal{A}_K \rightarrow \mathcal{Y}, \quad \boldsymbol{\alpha} \mapsto \mathbf{y}^{\boldsymbol{\alpha}, \varepsilon} := \mathcal{P}_\varepsilon(\boldsymbol{\alpha}),$$

where $\mathcal{P}_\varepsilon(\boldsymbol{\alpha})$ is the solution in the sense of [Definition 3.5](#) given the nonlinearity \mathbf{g}_α and the control ε . Then, \mathcal{P}_ε is well-defined.

Since we are interested in the dependence of the solution with respect to the parameters $\boldsymbol{\alpha}$, we assume in the following that the control $\varepsilon \in \mathcal{E}_{\text{ad}}$ is fixed and therefore omit to write ε .

We recall the weak formulation (3.2) for our structure of the nonlinearity given in (2.2) and $\boldsymbol{\alpha} \in \mathcal{A}_K$

$$(3.5) \quad a[\mathbf{y}^\alpha, \mathbf{v}] + \int_\Omega \mathbf{v}(x) \cdot \left(\sum_{j=1}^K \alpha_j \boldsymbol{\varphi}_j(\mathbf{y}^\alpha(x)) \right) dx = \int_\Omega \varepsilon(x) \cdot \mathbf{v}(x) dx \quad \forall \mathbf{v} \in \mathbb{H}_0^1(\Omega).$$

Notice that the difference $\bar{\mathbf{y}} := \mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2}$ of the solutions \mathbf{y}^{α^1} and \mathbf{y}^{α^2} for any $\alpha^1, \alpha^2 \in \mathcal{A}_K$ fulfills the following variational equality

$$(3.6) \quad a[\bar{\mathbf{y}}, \mathbf{v}] + \int_\Omega \mathbf{v}(x) \cdot \left(\sum_{j=1}^K \left(\alpha_j^1 \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^1}(x)) - \alpha_j^2 \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2}(x)) \right) \right) dx = 0 \quad \forall \mathbf{v} \in \mathbb{H}_0^1(\Omega).$$

If we choose $\mathbf{v} = \bar{\mathbf{y}} \in \mathbb{H}_0^1(\Omega)$ we obtain

$$(3.7) \quad a[\bar{\mathbf{y}}, \bar{\mathbf{y}}] + \int_\Omega \bar{\mathbf{y}}(x) \cdot \left(\sum_{j=1}^K \left(\alpha_j^1 \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^1}(x)) - \alpha_j^2 \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2}(x)) \right) \right) dx = 0.$$

We derive the following energy estimate for $\bar{\mathbf{y}}$.

Lemma 3.10. *Let [Assumptions 3.1](#) and [3.2](#) hold and $k \in [K]$. Let $\alpha^1, \alpha^2 \in \mathcal{A}_k$. For the difference $\bar{\mathbf{y}} = \mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2}$ of two corresponding solutions it holds that*

$$\|\mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2}\|_{\mathbb{H}_0^1(\Omega)} \leq C k \|\alpha^1 - \alpha^2\|_\infty,$$

with a constant $C > 0$ independent of $\mathbf{g}_{\alpha^1}, \mathbf{g}_{\alpha^2}, \alpha^1$ and α^2 .

Proof. The claim is true if $\bar{\mathbf{y}} := \mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2} \equiv 0 \in \mathbb{H}_0^1(\Omega)$. Hence we now assume that $\bar{\mathbf{y}} \in \mathbb{H}_0^1(\Omega) \setminus \{0\}$. From (3.7), we obtain

$$(3.8a) \quad \|\bar{\mathbf{y}}\|_{\mathbb{H}_0^1(\Omega)}^2 \stackrel{(3.3)}{=} a[\bar{\mathbf{y}}, \bar{\mathbf{y}}] = \sum_{j=1}^K \int_\Omega \bar{\mathbf{y}}(x) \cdot \left(-\alpha_j^1 \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^1}(x)) + \alpha_j^2 \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2}(x)) \right) dx$$

$$= \sum_{j=1}^k \int_\Omega \bar{\mathbf{y}}(x) \cdot \left[-\alpha_j^1 \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^1}(x)) + \alpha_j^2 \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2}(x)) + \alpha_j^1 \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2}(x)) - \alpha_j^1 \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^1}(x)) \right] dx$$

$$(3.8b) \quad = \underbrace{\int_\Omega -\sum_{j=1}^k \alpha_j^1 \left(\boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^1}(x)) - \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2}(x)) \right) \cdot \bar{\mathbf{y}}(x) dx}_{\leq 0} + \sum_{j=1}^k (\alpha_j^2 - \alpha_j^1) \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2}(x)) \cdot \bar{\mathbf{y}}(x) dx$$

$$(3.8c) \quad \leq \sum_{j=1}^k \int_\Omega (\alpha_j^2 - \alpha_j^1) \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2}(x)) \cdot \bar{\mathbf{y}}(x) dx$$

$$(3.8d) \quad \leq \sum_{j=1}^k \int_{\Omega} \left| (\alpha_j^2 - \alpha_j^1) \bar{\mathbf{y}}(x) \cdot \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2}(x)) \right| dx$$

$$(3.8e) \quad \leq \|\bar{\mathbf{y}}\|_{\mathbb{L}^2(\Omega)} \|\boldsymbol{\alpha}^2 - \boldsymbol{\alpha}^1\|_{\infty} \sum_{j=1}^k \|\boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2})\|_{\mathbb{L}^2(\Omega)}$$

$$(3.8f) \quad \leq k I_{\varphi} \|\boldsymbol{\alpha}^1 - \boldsymbol{\alpha}^2\|_{\infty} \|\bar{\mathbf{y}}\|_{\mathbb{L}^2(\Omega)} \leq k I_{\varphi} C_{\Omega} \|\boldsymbol{\alpha}^1 - \boldsymbol{\alpha}^2\|_{\infty} \|\bar{\mathbf{y}}\|_{\mathbb{H}_0^1(\Omega)}.$$

In (3.8), we use the following: In (3.8c), we apply the monotonicity of the nonlinearity provided by [Assumption 3.1-3](#)), i.e. that we have $(\boldsymbol{\varphi}_j(y) - \boldsymbol{\varphi}_j(\tilde{y})) \cdot (y - \tilde{y}) \geq 0$ for $y, \tilde{y} \in \mathbb{R}^2$ and $j \in [K]$, in (3.8d), we use the triangle inequality, in (3.8e) we use the general version of the Hölder inequality (cf. [3, Lemma 1.18]), and in (3.8f), we use (3.1) and the Poincaré inequality from which we get the constant $C_{\Omega} > 0$. Summarizing, we have

$$\|\bar{\mathbf{y}}\|_{\mathbb{H}_0^1(\Omega)}^2 \leq k I_{\varphi} C_{\Omega} \|\boldsymbol{\alpha}^1 - \boldsymbol{\alpha}^2\|_{\infty} \|\bar{\mathbf{y}}\|_{\mathbb{H}_0^1(\Omega)},$$

which leads to

$$\|\bar{\mathbf{y}}\|_{\mathbb{H}_0^1(\Omega)} \leq C k \|\boldsymbol{\alpha}^1 - \boldsymbol{\alpha}^2\|_{\infty}$$

with $C := I_{\varphi} C_{\Omega}$. □

The result of [Lemma 3.10](#) can be further improved in the sense that we can obtain an estimate in the \mathcal{Y} norm. From this, the continuity of the parameter-to-state map $\mathcal{P}_{\varepsilon}$ follows.

Theorem 3.11. *Let [Assumptions 3.1](#) and [3.2](#) hold and $k \in [K]$. Let $\boldsymbol{\alpha}^1, \boldsymbol{\alpha}^2 \in \mathcal{A}_k$. For the difference $\bar{\mathbf{y}} = \mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2}$ of two corresponding solutions it holds that*

$$\|\mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2}\|_{\mathcal{Y}} \leq C k^2 \|\boldsymbol{\alpha}^1 - \boldsymbol{\alpha}^2\|_{\infty}$$

for a constant $C > 0$ independent of $\mathbf{g}_{\alpha^1}, \mathbf{g}_{\alpha^2}, \boldsymbol{\alpha}^1$ and $\boldsymbol{\alpha}^2$.

Proof. The claim is true if $\bar{\mathbf{y}} \equiv 0 \in \mathcal{Y}$. Hence we now assume that $\bar{\mathbf{y}} \not\equiv 0 \in \mathcal{Y}$. We know from [Theorem 3.6](#) that $\bar{\mathbf{y}} \in \mathcal{Y}$ since it is the sum of two functions in \mathcal{Y} . Hence it fulfils the following equation (cf. (3.8b)) with homogeneous Dirichlet boundary conditions:

$$(3.9) \quad -\Delta \bar{\mathbf{y}} + \sum_{j=1}^k \int_{\Omega} \left(\alpha_j^1 \left(\boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^1}(x)) - \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2}(x)) \right) + (\alpha_j^2 - \alpha_j^1) \boldsymbol{\varphi}_j(\mathbf{y}^{\alpha^2}(x)) \right) \cdot \bar{\mathbf{y}}(x) dx = 0.$$

We now take the \mathbb{L}^2 inner product of (3.9) with $\Delta \bar{\mathbf{y}}$ and obtain

$$\|\Delta \bar{\mathbf{y}}\|_{\mathbb{L}^2(\Omega)}^2 \leq k \left(\|\boldsymbol{\alpha}^1\|_{\infty} L_{\varphi} \|\bar{\mathbf{y}}\|_{\mathbb{L}^2(\Omega)} + C_{\varphi} \|\boldsymbol{\alpha}^1 - \boldsymbol{\alpha}^2\|_{\infty} \right) \|\bar{\mathbf{y}}\|_{\mathbb{L}^2(\Omega)} \|\Delta \bar{\mathbf{y}}\|_{\mathbb{L}^2(\Omega)},$$

where we have used the Lipschitz continuity of the basis elements and the Cauchy-Schwarz inequality. Now we divide by $\|\Delta \bar{\mathbf{y}}\|_{\mathbb{L}^2(\Omega)}$, apply the Poincaré inequality, which gives us the factor $C_{\Omega} > 0$, and use the boundedness of the solutions provided by [Theorem 3.6](#) to obtain

$$\|\Delta \bar{\mathbf{y}}\|_{\mathbb{L}^2(\Omega)} \leq k \alpha_{\max} (L_{\varphi} 2 C_0 + C_{\varphi}) C_{\Omega} \|\bar{\mathbf{y}}\|_{\mathbb{H}_0^1}.$$

Now we can apply [Lemma 3.10](#) to estimate

$$\|\Delta \bar{\mathbf{y}}\|_{\mathbb{L}^2(\Omega)} \leq k^2 \alpha_{\max} (L_{\varphi} 2 C_0 + C_{\varphi}) I_{\varphi} C_{\Omega}^2 \|\boldsymbol{\alpha}^1 - \boldsymbol{\alpha}^2\|_{\infty}.$$

The statement follows by setting $C := \alpha_{\max} (L_{\varphi} 2 C_0 + C_{\varphi}) I_{\varphi} C_{\Omega}^2$. □

3.3. Wellposedness of inverse-map. This section is devoted to the stability analysis of the parameters with respect to small perturbations in the solution. Let us again fix the control $\varepsilon \in \mathcal{E}_{\text{ad}}$. We are interested in the stability estimate of the form

$$\|\boldsymbol{\alpha}^1 - \boldsymbol{\alpha}^2\|_{\infty} \leq C_{\text{inv}} \|\mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2}\|_{\mathcal{Y}}.$$

This inequality implies the Lipschitz continuity of the inverse parameter-to-state map $\mathcal{P}_{\varepsilon}^{-1}$, i.e., the state-to-parameter map given for fixed controls $\varepsilon \in \mathcal{E}_{\text{ad}}$ as

$$\mathcal{P}_{\varepsilon}^{-1} : \mathcal{B} \rightarrow \mathcal{A}_K, \quad \mathbf{y}^{\alpha} \mapsto \boldsymbol{\alpha} = \mathcal{P}_{\varepsilon}^{-1}(\mathbf{y}^{\alpha}),$$

where we introduced the compact set $\mathcal{B} := \mathcal{P}(\mathcal{A}_K) \subset \mathcal{Y}$. Therefore, if the difference of two solutions is small in the \mathcal{Y} norm, then the corresponding parameters will be close with respect to the maximum norm (which is equivalent to all other norms in \mathbb{R}^K). This is the content of [Theorem 3.12](#). Notice that an essential assumption is that the basis elements are bounded away from zero.

Theorem 3.12. *Suppose that [Assumption 3.1](#) and $k \in [K]$ hold. Furthermore, let $\alpha^1, \alpha^2 \in \mathcal{A}_k$. Then there exists a constant $C_{\text{inv}} > 0$ independent of α^1, α^2 such that*

$$\|\alpha^1 - \alpha^2\|_\infty \leq C_{\text{inv}} \|\mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2}\|_{\mathcal{Y}}.$$

Moreover, the constant has the form

$$C_{\text{inv}} := \sqrt{2\tilde{c}}(1 + k L_\varphi \alpha_{\max} C_\Omega) c_\varphi^{-1}$$

for constants $\tilde{c}, C_\Omega > 0$.

Proof. With $\bar{\alpha} := \alpha^1 - \alpha^2$ and $\Phi = (\varphi_1, \dots, \varphi_K)$, we find

$$(3.10a) \quad \|\bar{\alpha}\|_\infty^2 \leq \tilde{c} \|\bar{\alpha}\|_2^2 = \tilde{c} c_\varphi^{-2} \|c_\varphi \bar{\alpha}\|_2^2 \leq \tilde{c} c_\varphi^{-2} \|\Phi(\mathbf{y}^{\alpha^1})\bar{\alpha} + \Phi(\mathbf{y}^{\alpha^2})\alpha_2 - \Phi(\mathbf{y}^{\alpha^2})\alpha_2\|_{\mathbb{L}^2(\Omega)}^2$$

$$(3.10b) \quad \leq 2\tilde{c} c_\varphi^{-2} \left(\|\Phi(\mathbf{y}^{\alpha^1})\alpha_1 - \Phi(\mathbf{y}^{\alpha^2})\alpha_2\|_{\mathbb{L}^2(\Omega)}^2 + \|(\Phi(\mathbf{y}^{\alpha^2}) - \Phi(\mathbf{y}^{\alpha^1}))\alpha_2\|_{\mathbb{L}^2(\Omega)}^2 \right)$$

$$(3.10c) \quad \leq 2\tilde{c} c_\varphi^{-2} \left(\|\Delta(\mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2})\|_{\mathbb{L}^2(\Omega)}^2 + k^2 L_\varphi^2 \alpha_{\max}^2 \|\mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2}\|_{\mathbb{L}^2(\Omega)}^2 \right)$$

$$(3.10d) \quad \leq 2\tilde{c} c_\varphi^{-2} (1 + k L_\varphi \alpha_{\max})^2 \|\mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2}\|_{\mathcal{Y}}^2.$$

In (3.10a) we use the equivalence of norms in \mathbb{R}^K and [Assumption 3.1-2](#); in (3.10b) we apply the triangle inequality and a binomial inequality; in (3.10c) we use that the difference $\mathbf{y}^{\alpha^1} - \mathbf{y}^{\alpha^2}$ fulfills (3.6) and that φ_j are Lipschitz continuous for $j \in [K]$; in (3.10d) we exploit that the Laplace operator is an isometry. \square

4. GREEDY-RECONSTRUCTION ALGORITHM

In this section, we present our strategy to tackle the identification problem (2.3). For this, we recall some notation. We denote by $\mathbf{y}^{\beta, \varepsilon}$ the solution of (2.1) with the control ε and nonlinearity \mathbf{g}_β . We have for $1 \leq k \leq K-1$ the set $\mathcal{A}_k := \{\alpha \in \mathcal{A}_K \mid \alpha_j = 0, j = k+1, \dots, K\}$ (cf. (3.4)). Further, we set $\mathbf{e}^k \in \mathbb{R}^K$ as the k -th canonical vector in \mathbb{R}^K . Hence, $\mathbf{y}^{\mathbf{e}^k, \varepsilon}$ denotes the solution of (2.1) with the control ε and nonlinearity φ_k .

We show now how to construct a set of optimized controls and basis elements, in particular with the goal of improving local convexity properties of (2.3). We follow the idea of [7, 8, 9, 26], where linear and bilinear dynamical systems have been investigated. It is the aim of this work to extend this strategy to the general semilinear PDE case (2.1).

The main idea is to split the reconstruction process of \mathbf{g}_{α^*} into offline and online phases. In the offline phase, a greedy algorithm computes a set of optimized controls $\{\varepsilon^m\}_{m=1}^K$ by exploiting only simulations of the semilinear model (2.1) and without using any laboratory data. In the online phase, the computed controls $\{\varepsilon^m\}_{m=1}^K$ are used experimentally to produce the laboratory data $\mathbf{y}_*^{\varepsilon^m}$ for $m \in [L]$ and to define the nonlinear problem (2.1). While the online phase consists (mathematically) of solving a classical parameter-identification inverse problem, the offline phase requires the greedy algorithm that was first introduced in [26] and analyzed and improved in [8]. The goal of this offline/online framework is to find a good approximation of the unknown operator \mathbf{g}_{α^*} for which the difference between observed experimental data and numerically computed data is the smallest for any control. To do so, the algorithm attempts to distinguish numerical data for any two $\mathbf{g}_{\alpha^1}, \mathbf{g}_{\alpha^2} \in \text{span } \mathcal{G}$ [26]. This is achieved by an iterative procedure that performs a sweep over the basis \mathcal{G} and computes a new control at each iteration. Suppose that at iteration k the control fields $\varepsilon^1, \dots, \varepsilon^k$ are already computed; the new control function ε^{k+1} is obtained by two substeps, the so-called fitting and splitting step. One first solves the identification problem (4.1) which gives the coefficients $\beta^k = (\beta_j^k)_{j=1, \dots, k}$. Then one computes the new control as the solution of the splitting step (4.2).

This general procedure can be optimized using some extensions. More specifically, in each iteration, all elements of \mathcal{G} are considered in parallel and the ‘best’ one is chosen in a greedy way; see [7, Section

6.1] for further explanation of the original and optimized greedy reconstruction algorithm in a linear case.

This offline procedure for our nonlinear case is summarized in [Algorithm 4.1](#). We refer to [\[8\]](#) for a general idea to prove that this algorithm is able to make problem (2.3) (locally) uniquely solvable.

Algorithm 4.1 Optimized Nonlinear Greedy Reconstruction (ONGR) Algorithm

Require: A set of K basis functions $\mathcal{G} = \{\varphi_1, \dots, \varphi_K\}$ and a tolerance $\text{tol}_1 > 0$.

1: Compute ϵ^1 and the index $i_1 \in [K]$ by solving the *initialization* problem

$$\max_{i \in [K]} \max_{\epsilon \in \mathcal{E}_{\text{ad}}} \frac{1}{2} \left\| \mathbf{y}^{0, \epsilon} - \mathbf{y}^{e^i, \epsilon} \right\|_{\mathbb{L}^2(\Omega)}^2 + \frac{\nu}{2} \|\epsilon\|_{\mathbb{L}^2(\Omega)}^2.$$

2: Swap φ_1 and φ_{i_1} in \mathcal{G} , and set $k = 1$ and $f_{\max} = \frac{1}{2} \left\| \mathbf{y}^{0, \epsilon^1} - \mathbf{y}^{e^{i_1}, \epsilon^1} \right\|_{\mathbb{L}^2(\Omega)}^2$.

3: **while** $k \leq K - 1$ and $f_{\max} > \text{tol}_1$ **do**

4: **for** $\ell = k + 1, \dots, K$ **do**

5: *Fitting step:* Find β^ℓ that solves the problem

$$(4.1) \quad \min_{\beta \in \mathcal{A}_k} \frac{1}{2} \sum_{m=1}^k \left\| \mathbf{y}^{\beta, \epsilon^m} - \mathbf{y}^{e^\ell, \epsilon^m} \right\|_{\mathbb{L}^2(\Omega)}^2.$$

6: **end for**

7: *Splitting step:* Find ϵ^{k+1} and i_{k+1} that solve the problem

$$(4.2) \quad \max_{i \in \{k+1, \dots, K\}} \max_{\epsilon \in \mathcal{E}_{\text{ad}}} \frac{1}{2} \left\| \mathbf{y}^{\beta^i, \epsilon} - \mathbf{y}^{e^i, \epsilon} \right\|_{\mathbb{L}^2(\Omega)}^2 + \frac{\nu}{2} \|\epsilon\|_{\mathbb{L}^2(\Omega)}^2.$$

8: Swap φ_{k+1} and $\varphi_{i_{k+1}}$ in \mathcal{G} .

9: Set $f_{\max} = \frac{1}{2} \left\| \mathbf{y}^{\beta^{i_{k+1}}, \epsilon^{k+1}} - \mathbf{y}^{e^{i_{k+1}}, \epsilon^{k+1}} \right\|_{\mathbb{L}^2(\Omega)}^2$.

10: Set $k \leftarrow k + 1$.

11: **end while**

12: **return** \mathcal{G} and the computed $\{\epsilon^m\}_{m=1}^k$.

We now present the optimality system for the optimization problems appearing in [Algorithm 4.1](#). Let us define $\mathbf{g}_k(y) := \varphi_k(y)$ for $y = (y_1, y_2)^\top \in \mathbb{R}^2$ and $k \in [K]$. We start with the discussion of the fitting step at the k -th iteration. For this let us introduce $\mathcal{Y}^k := \mathcal{Y} \times \dots \times \mathcal{Y}$ (k times) and define $\{\mathbf{y}_{k+1}^m\}_{m=1}^k \subset \mathcal{Y}^k$ as the solutions to

$$-\Delta \mathbf{y}_{k+1}^m + \mathbf{g}_{k+1}(\mathbf{y}_{k+1}^m) = c \epsilon^m \quad \text{for } m \in [k]$$

with homogeneous Dirichlet boundary conditions. For the fixed controls $\{\epsilon^m\}_{m=1}^k \subset \mathcal{E}_{\text{ad}}$ and with given $\{\mathbf{y}_{k+1}^m\}_{m=1}^k \subset \mathcal{Y}^k$, the fitting step is given as the following minimization problem (the subscript f stands for “fitting step”)

$$(4.3) \quad \begin{aligned} \min J_f(\beta, \mathbf{y}^1, \dots, \mathbf{y}^k) &:= \sum_{m=1}^k \frac{1}{2} \left\| \mathbf{y}^m - \mathbf{y}_{k+1}^m \right\|_{\mathbb{L}^2(\Omega)}^2 \\ \text{subject to } &-\Delta \mathbf{y}^m + \mathbf{g}_\beta(\mathbf{y}^m) = \epsilon^m, \quad \mathbf{y}_{|\Gamma}^m = 0 \quad \text{for } m \in [k] \\ &\text{and } \beta \in \mathcal{A}_k. \end{aligned}$$

Using the parameter-to-state map \mathcal{P}_ϵ defined in [Proposition 3.9](#), we can introduce the reduced cost

$$\widehat{J}_f(\beta) := J_f(\beta, \mathcal{P}_{\epsilon^1}(\beta), \dots, \mathcal{P}_{\epsilon^k}(\beta))$$

and the reduced problem

$$(4.4) \quad \min \widehat{J}_f(\beta) \quad \text{s.t.} \quad \beta \in \mathcal{A}_k.$$

We can state the following theorem.

Theorem 4.1. *Let [Assumptions 3.1](#) and [3.2](#) hold. Then the optimization problem (4.3) has at least one solution $\widehat{\beta} \in \mathcal{A}_k$.*

Proof. Notice that the reduced cost \widehat{J}_f is non-negative and thus there exists a minimizing sequence $\{\beta^\ell\}_{\ell \in \mathbb{N}} \subset \mathcal{A}_k$ such that

$$0 \leq \bar{J}_f = \inf\{\widehat{J}_f(\beta) \mid \beta \in \mathcal{A}_k\} = \lim_{\ell \rightarrow \infty} \widehat{J}_f(\beta^\ell) < \infty.$$

We set $\mathbf{y}^{\beta, \varepsilon} := \{\mathcal{P}_{\varepsilon^m}(\beta)\}_{m=1}^k \subset \mathcal{Y}^k$. Due to [Theorem 3.6](#), the sequence $\{\mathbf{y}^{\beta^\ell, \varepsilon}\}_{\ell \in \mathbb{N}} \subset \mathcal{Y}^k$ is well defined. Since \mathcal{A}_k is compact in \mathbb{R}^k , there exists a subsequence (which is still labeled by β^ℓ) $\{\beta^\ell\}_{\ell \in \mathbb{N}} \subset \mathcal{A}_k$ of the minimizing sequence and an element $\bar{\beta} \in \mathcal{A}_k$ such that we have the strong convergence $\beta^\ell \rightarrow \bar{\beta}$ in \mathbb{R}^k as $\ell \rightarrow \infty$. Due to [Lemma 3.10](#) we have $\mathbf{y}^{\beta^\ell, \varepsilon} \rightarrow \bar{\mathbf{y}} := \mathbf{y}^{\bar{\beta}, \varepsilon}$ in \mathcal{Y}^k . Using the lower semicontinuity of the norms and the continuity of the parameter-to-state map from \mathcal{A}_K to \mathcal{Y} ([Lemma 3.10](#))

$$\widehat{J}_f(\bar{\beta}) \leq \lim_{\ell \rightarrow \infty} \widehat{J}_f(\beta^\ell) = \bar{J}_f.$$

□

Let us now introduce the adjoint variable $\mathbf{q} = \mathbf{q}(x) = (q_1(x), q_2(x))^\top$. The adjoint equations for the fitting step read as follows

$$-\Delta \mathbf{q}^m + \mathbf{g}'_\beta(\mathbf{y}^m)^\top \mathbf{q}^m = -(\mathbf{y}^m - \mathbf{y}_{k+1}^m), \quad \text{for } m \in [k],$$

completed with homogeneous Dirichlet boundary conditions. Since $\nabla \mathbf{g}_\beta(\mathbf{y})$ is positive semidefinite, we can ensure a constrained qualification for (4.4), so that first-order necessary optimality conditions can be formulated; cf. [\[36, Section 6.1.2\]](#).

Supposing that $\bar{\beta}$ solves (4.4), the optimality condition of the fitting step in iteration k reads

$$\left\langle \nabla \widehat{J}_f(\bar{\beta}), \beta - \bar{\beta} \right\rangle_2 \geq 0 \quad \forall \beta \in \mathcal{A}_k,$$

where $\nabla \widehat{J}_f = (\nabla_{\beta_1} \widehat{J}_f, \dots, \nabla_{\beta_k} \widehat{J}_f)^\top$ is given by

$$\nabla_{\beta_j} \widehat{J}_f = \nu \beta_j + \sum_{m=1}^k \int_{\Omega} \mathbf{q}^m(x) \cdot \varphi_j(\mathbf{y}^m(x)) \, dx, \quad j \in [k].$$

Now we fix β and turn to the splitting step in which the new control ε^{k+1} is computed. The splitting step at iteration $k < K$ is given by (the subscript s stands for “splitting step”)

$$\begin{aligned} \min J_s(\varepsilon, \mathbf{y}^\beta, \mathbf{y}^{k+1}) &:= -\frac{1}{2} \|\mathbf{y}^\beta - \mathbf{y}^{k+1}\|_{\mathbb{L}^2(\Omega)}^2 + \frac{\nu}{2} \|\varepsilon\|_{\mathbb{L}^2(\Omega)}^2 \\ \text{subject to } &\begin{cases} -\Delta \mathbf{y}^\beta + \mathbf{g}_\beta(\mathbf{y}^\beta) = \varepsilon, & \mathbf{y}|_\Gamma = 0, \\ -\Delta \mathbf{y}^{k+1} + \mathbf{g}_{k+1}(\mathbf{y}^{k+1}) = \varepsilon, & \mathbf{y}|_\Gamma = 0, \end{cases} \\ &\text{and } \varepsilon \in \mathcal{E}_{\text{ad}}. \end{aligned}$$

Exploiting the control-to-state map \mathcal{S} defined in [Theorem 3.8](#), we can introduce the reduced cost

$$\widehat{J}_s(\varepsilon) := J_s(\varepsilon, \mathcal{S}_\beta(\varepsilon), \mathcal{S}_{\mathbf{e}^{k+1}}(\varepsilon)),$$

where $\mathbf{e}^k \in \mathbb{R}^K$ denotes again the k -th unit vector in \mathbb{R}^K . We also introduce the reduced problem

$$(4.5) \quad \min \widehat{J}_s(\varepsilon) \quad \text{s.t.} \quad \varepsilon \in \mathcal{E}_{\text{ad}}.$$

We can state the following theorem on the existence of solutions to (4.5).

Theorem 4.2. *Let [Assumptions 3.1](#) and [3.2](#) hold. Then the optimization problem (4.5) has at least one solution $\bar{\varepsilon} \in \mathcal{E}_{\text{ad}}$.*

Proof. Since our \widehat{J}_s is bounded from below by the virtue of [Theorem 3.6](#) and the boundedness of the set of admissible controls \mathcal{E}_{ad} , we can apply [\[11, Theorem 3.1\]](#). □

The adjoint equations for the fitting step read as follows

$$\begin{aligned} -\Delta \mathbf{q}^\beta + \mathbf{g}'_\beta(\mathbf{y}^\beta)^\top \mathbf{q}^\beta &= -\nabla_{\mathbf{y}^\beta} J_s(\varepsilon; \mathbf{y}), \\ -\Delta \mathbf{q}^{k+1} + \mathbf{g}'_{k+1}(\mathbf{y}^{k+1})^\top \mathbf{q}^{k+1} &= -\nabla_{\mathbf{y}^{k+1}} J_s(\varepsilon; \mathbf{y}), \end{aligned}$$

completed with homogeneous Dirichlet boundary conditions. Since $\nabla \mathbf{g}_\beta(y)$ and $\nabla \mathbf{g}_{k+1}(y)$ are positive semidefinite, we can ensure a constrained qualification for (4.4), so that first-order necessary optimality conditions can be formulated; cf. [36, Section 6.1.2].

Suppose that $\bar{\varepsilon}$ solves (4.5), then an optimality condition of the splitting step reads

$$\left\langle \nabla_{\varepsilon} \hat{J}_s(\bar{\varepsilon}), \varepsilon - \bar{\varepsilon} \right\rangle_{\mathbb{L}^2(\Omega)} \geq 0 \quad \forall \varepsilon \in \mathcal{E}_{\text{ad}},$$

where $\nabla_{\varepsilon} \hat{J}_s$ is given $\nabla_{\varepsilon} \hat{J}_s(\varepsilon) = \nu \varepsilon - (\mathbf{q}^\beta + \mathbf{q}^{k+1})$.

5. NUMERICAL APPROXIMATION

In this section, we present the numerical approximation of the governing equation. First, we describe our method to solve the governing model (2.1). We choose now $\Omega = (-x_{\max}, x_{\max})^2$ for a fixed $x_{\max} > 0$. We set a numerical grid that provides a partitioning of Ω in $N \times N$, $N > 1$, equally-spaced non-overlapping square cells of side length $h = 2x_{\max}/N$. We define the nodal points

$$x_1^i = i h - x_{\max}, \quad x_2^j = j h - x_{\max}.$$

Our discrete domain is then given by

$$\Omega_h := \bigcup_{i,j=1}^N \omega_{ij},$$

where an elementary cell is defined as

$$\omega_{ij} := \left\{ (x_1, x_2) \in \Omega \mid (x_1, x_2) \in (x_1^{i-1}, x_1^i) \times (x_2^{j-1}, x_2^j) \right\}, \quad i, j \in [N].$$

In our numerical examples, we use $x_{\max} = 1$.

We define the discrete vectors $\mathbf{y}^N, \varepsilon^N \in \mathbb{R}^{2(N+1)^2}$ as $\mathbf{y}^N = (y_1^0, \dots, y_1^{N^2}, y_2^0, \dots, y_2^{N^2})^\top$ and $\varepsilon^N = (\varepsilon_1^0, \dots, \varepsilon_1^{N^2}, \varepsilon_2^0, \dots, \varepsilon_2^{N^2})^\top$.

After discretization of (2.1), we end up with

$$(5.1) \quad -A \mathbf{y}^N + \mathbf{g}(\mathbf{y}^N) = \varepsilon^N,$$

where A is the discrete Laplace operator using finite differences and zero boundary conditions [24, Section 2.2].

We consider a fixed-point method for solving (5.1) numerically [14]. The procedure is summarized in Algorithm 5.1.

Algorithm 5.1 Fixed-point method

Require: Parameter $0 \leq \lambda_a \leq 1$, tolerance $tol > 0$, maximum iteration depth $\ell_{\max} \gg 1$

- 1: Set $\ell = 0$ and $E \gg tol$
 - 2: Generate initial guess \mathbf{y}_0^N that satisfies $-A \mathbf{y}_0^N = \varepsilon^N$
 - 3: **while** $E < tol$ **and** $\ell < \ell_{\max}$ **do**
 - 4: Calculate $\tilde{\mathbf{y}}^N$ as solution to $-A \tilde{\mathbf{y}}^N = \varepsilon^N - \mathbf{g}(\mathbf{y}_\ell^N)$
 - 5: Set $\mathbf{y}_{\ell+1}^N = \lambda_a \mathbf{y}_\ell^N + (1 - \lambda_a) \tilde{\mathbf{y}}^N$ and $E = h \|\mathbf{y}_{\ell+1}^N - \mathbf{y}_\ell^N\|_2$
 - 6: Set $\ell \leftarrow \ell + 1$
 - 7: **end while**
 - 8: **return** \mathbf{y}_ℓ^N
-

Next, we explain the ingredients for Algorithm 4.1 that we use in our numerical experiment. We choose the following nonlinearity, motivated by the steady state solutions of coupled Lotka-Volterra equations (see, e.g., [17, 20, 29]):

$$\mathbf{g}_\alpha(y) := (\gamma_1 G_\alpha(y), -\gamma_2 G_\alpha(y))^\top,$$

where $\gamma_1 \geq \gamma_2 > 0$ are given constants and

$$(5.2) \quad G_\alpha(y) := \sum_{j=1}^K \alpha_j \varphi_j(y),$$

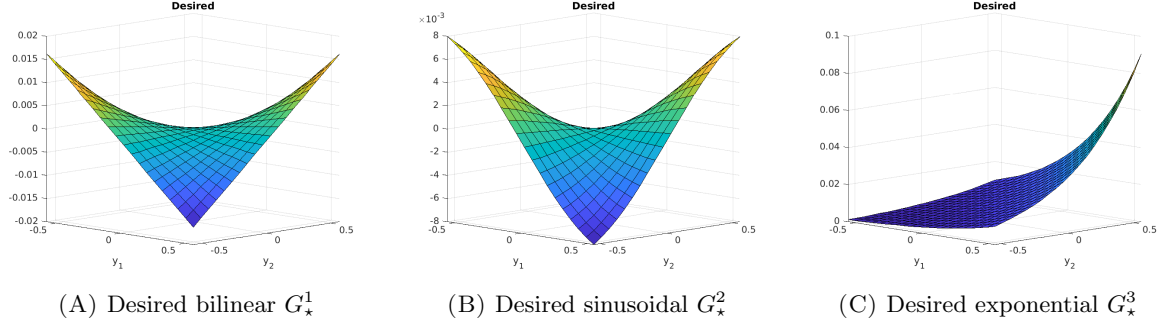


FIGURE 6.1. The three different desired nonlinearities G_\star^i , $i = 1, 2, 3$ defined in (6.1).

with basis functions $\varphi_j : \mathbb{R}^2 \rightarrow \mathbb{R}$, $j \in [K]$.

We use the two-dimensional monomials as basis functions to reconstruct the unknown nonlinearity G_α . More precisely, we approximate the nonlinearity using polynomials up to order d^2 , i.e.,

$$(5.3) \quad G_\alpha(y_1, y_2) = \sum_{i_1=0}^d \sum_{i_2=0}^d \alpha_{i_1, i_2} y_1^{i_1} y_2^{i_2}.$$

With this setting, the support of the solution \mathbf{y}^α of (2.1) does not need to be known in advance. Furthermore, it is possible to make predictions of the nonlinearity outside of the support of the solutions due to the non-locality of the monomials.

In particular, the two-dimensional monomials of order $P \in \mathbb{N}$ are given by

$$\mathcal{G} = \{y_1, y_2, y_1 y_2, y_1^2, y_2^2, y_1^2 y_2, \dots, y_1^P y_2^P\}.$$

The cardinality of \mathcal{G} , which is the maximum number of iterations K in the algorithm, is then given by $\binom{2+P}{P}$. We now discuss Assumption 3.1 for our choice of the basis elements in the following remark.

Remark 5.1. 1) All basis elements are smooth and in particular (locally) Lipschitz continuous since they are monomials.
 2) The basis elements are not bounded away from zero and in general not bounded from above. However, on every bounded set in \mathbb{R}_+^2 they are also bounded and strictly greater than zero.
 3) All monomials are monotonically increasing for nonnegative entries. However, by our special structure (5.2), the nonlinearity \mathbf{g}_α is not monotonically increasing on the full space \mathbb{R}^2 but on a (non-empty) subspace.

Hence our choice of monomials does not fit perfectly with our assumptions. However, they can be fulfilled after restricting ourselves to a non-empty subset of \mathbb{R}^2 . Furthermore, even though the theoretical assumptions are not completely fulfilled, we observe that our strategy works very well in our numerical experiments.

6. NUMERICAL EXPERIMENTS

In the following, we present the results of numerical tests in order to validate the reconstruction ability of our strategy described above.

We consider three different desired nonlinearities G_\star^i , $i = 1, 2, 3$ and aim to reconstruct them. All have different fundamental properties. One of them lies within the span of the basis \mathfrak{G} (cf. Section 2) and the other two can only be approximated by monomials from which one is bounded and enjoys some similarities to the basis elements and the other one is unbounded. Specifically, we choose

$$(6.1) \quad G_\star^1(y) = 0.05 y_1 y_2, \quad G_\star^2(y) = 0.01 \sin(2y_1) \sin(2y_2), \quad G_\star^3(y) = 0.01 \exp(2y_1) \exp(2y_2).$$

These nonlinear functions are plotted in Figure 6.1. Furthermore, we use $\varepsilon_a = (-1, -1)^\top$, $\varepsilon_b = (1, 1)^\top$, $\gamma_1 = \gamma_2 = 0.2$ and run Algorithm 4.1 for three different bases \mathfrak{G} for monomials of order $P = 2, 3, 5$, respectively.

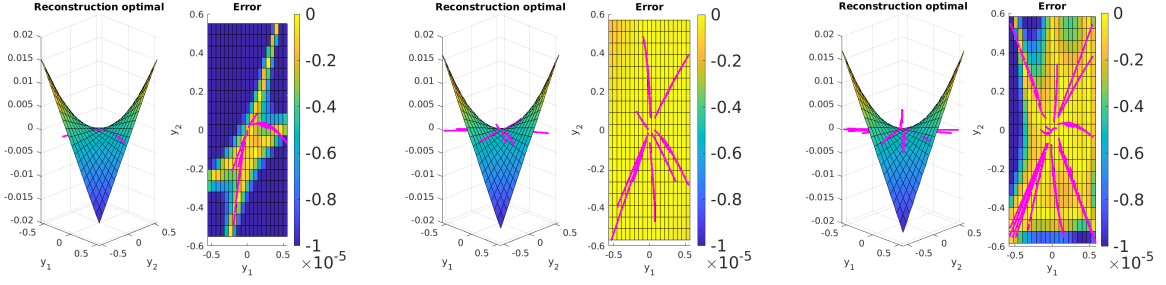


FIGURE 6.2. Reconstruction and error plots (cf. (6.2)) for $P = 2, 3, 5$ and the bilinear nonlinearity with solutions curves in magenta.

In order to evaluate our results, we consider the error

$$(6.2) \quad \mathfrak{e}(y; \alpha) := G_{\star}(y) - g_{\alpha}(y) = G_{\star}(y) - \sum_{j=1}^K \alpha_j \varphi_j(y)$$

between the true and reconstructed nonlinearity, for $y \in \mathbb{R}^2$ and $\alpha \in \mathcal{A}_K$, on a subset Ω_M of \mathbb{R}^2 . This subset Ω_M is the smallest square in \mathbb{R}^2 that contains all points of the corresponding (discrete) solution curves $(\mathbf{y}_{1,m}^N, \mathbf{y}_{2,m}^N)$, $m \in [K]$. Here, we use the notation that $\mathbf{y}_{i,m}^N$ is the i -th component of the discrete solution \mathbf{y}^N while applying the m -th control. Recall, that the true and reconstructed nonlinearities G_{\star} and G_{α} as well as the error \mathfrak{e} are defined on the whole \mathbb{R}^2 since we use nonlocal basis functions.

Remark 6.1. *The final identification only takes the values on the solution curves $(\mathbf{y}_{1,m}^N, \mathbf{y}_{2,m}^N)$, $m \in [K]$ for each control into account because only there we have information from the underlying system of equations. However, since we know the nonlinearities that we try to reconstruct, we are able to define the error on the whole domain Ω_M .*

6.1. Reconstruction of a bilinear function. Let us consider first the bilinear $G_{\star}^1(y) = 0.05y_1y_2$, shown Figure 6.1(A), which clearly lies in the span \mathcal{G} for any order $P \geq 1$. We run Algorithm 4.1, use the selected controls to compute the data, and solve the corresponding online identification problem (2.3). The results are shown in Figure 6.2, where the three different subfigures correspond to the three different bases \mathcal{G} of order $P = 2, 3, 5$.

Each subfigure of Figure 6.2 contains the nonlinearity G_{α} reconstructed by problem (2.3) (left plot) and the error $\mathfrak{e}(y; \alpha)$ between the true and the reconstructed nonlinearity (right plot). In the error plot, the solution curves $(\mathbf{y}_{1,m}^N, \mathbf{y}_{2,m}^N)$, $m \in [K]$, are plotted in magenta.

We observe that the order of magnitude of the error is quite small for all three cases. As discussed above, this is to be expected since the desired nonlinearity G_{\star}^1 is one of the basis elements. However, even though the final identification for $P = 5$ has access to the most data points (as shown by the number of magenta lines), the error is actually smallest for $P = 3$. The reason for this is that for $P = 5$ more basis elements with higher polynomial degrees are taken into account. Hence, in particular, at the boundary of the domain the error \mathfrak{e} might become quite large if the solution curves do not reach this part of the domain; see also Remark 6.1. To avoid this issue, it even seems that Algorithm 4.1 is trying to select the controls such that the corresponding solution curves reach into more parts of Ω_M (cf. beginning of Section 6). This is particularly apparent for $P = 3$ and $P = 5$.

In order to investigate further how Algorithm 4.1 chooses the controls, we consider the solutions

$$y_1(x) = \eta \kappa(x_1, x_2), \quad y_2(x) = -\vartheta \kappa(x_1, x_2) \quad \text{with} \quad \kappa(x_1, x_2) := \sin\left(\frac{x_1+1}{2}\pi\right) \sin\left(\frac{x_2+1}{2}\pi\right).$$

inside the domain $\Omega = (-1, 1)^2$, for some $\eta, \vartheta \in \mathbb{R} \setminus \{0\}$ with $\vartheta > \eta$. Then, the graphs of the solutions (y_1, y_2) in the y_1 - y_2 plane are given by straight lines starting in the origin, having slope $-\vartheta/\eta$ and length $\sqrt{\eta^2 + \vartheta^2}$. Hence, we can reach every point in the y_1 - y_2 plane within the $[\eta, \vartheta]$ rectangle. The

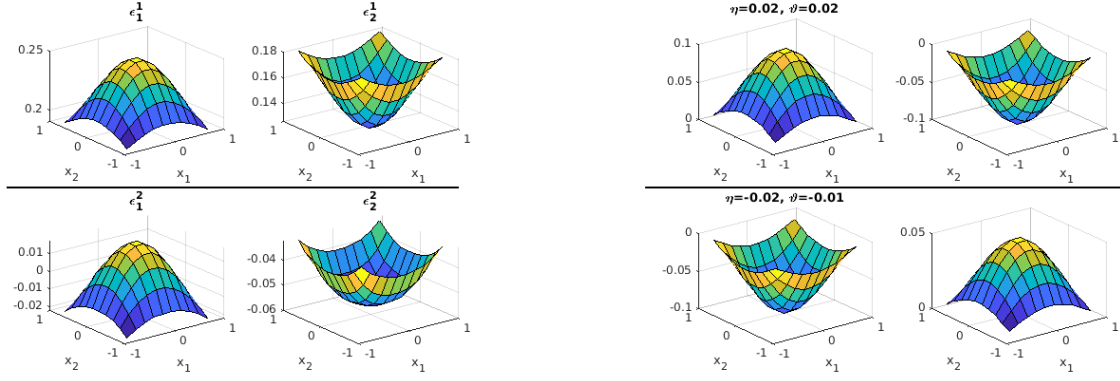


FIGURE 6.3. Controls functions. Left: two pairs of controls obtained by Algorithm 4.1; Right: two pairs of controls with the same structure as the right-hand side of (6.3) with different η, ϑ .

partial derivatives for $i = 1, 2$ are given by

$$\partial_{x_i x_i}^2 y_1(x) = -\frac{\pi^2}{4} y_1(x), \quad \partial_{x_i x_i}^2 y_2(x) = \frac{\pi^2}{4} y_2(x),$$

which implies by the structure of our equation that

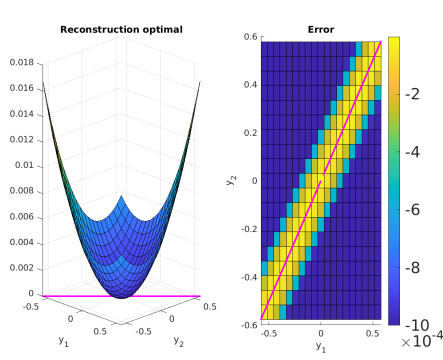
$$(6.3) \quad \begin{aligned} -\Delta y_1 + 0.05 \gamma_1 y_1 y_2 &= \eta \frac{\pi^2}{2} \kappa(x_1, x_2) - 0.05 \gamma_1 \eta \vartheta \kappa(x_1, x_2)^2, \\ -\Delta y_2 - 0.05 \gamma_2 y_1 y_2 &= -\vartheta \frac{\pi^2}{2} \kappa(x_1, x_2) + 0.05 \gamma_2 \eta \vartheta \kappa(x_1, x_2)^2. \end{aligned}$$

Recalling (2.1), we observe that by choosing the controls according to the right-hand side of (6.3), the corresponding solutions are able to reach any point within the $[\eta, \vartheta]$ rectangle.

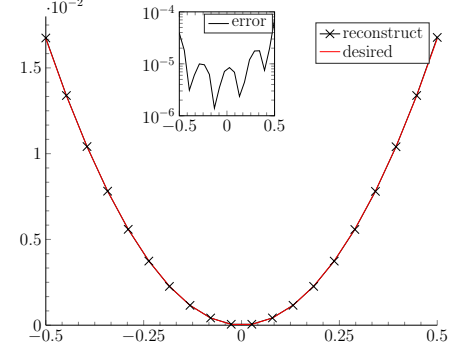
In Figure 6.3, we plot on the left a representative selection of the controls $\{\epsilon^m\}_{m=1}^K$ found by Algorithm 4.1, and on the right some controls chosen according to the right-hand side of (6.3). We observe that Algorithm 4.1 attempts to compute controls that mimic the structure of the right-hand side in (6.3) in order to reach different parts of the domain.

As a way of measuring the effectiveness of our procedure, we compare the behavior of the error when applying random controls for $P = 5$. For this, we consider the bilinear nonlinearity $G_*^1(y_1, y_2) = 0.05 y_1 y_2$. As controls, we (randomly) choose constant functions in \mathcal{E}_{ad} . When we now try to reconstruct the nonlinearity, we obtain the results plotted in Figures 6.4(A) and 6.4(B). In Figure 6.4(A), we plot on the left the reconstructed nonlinearity and on the right the error ϵ together with the solution curves $\{(\mathbf{y}_{1,m}^N, \mathbf{y}_{1,m}^N)\}_{m=1}^K$ applying 19 random controls in \mathcal{E}_{ad} having a constant value. We see clearly that the true nonlinearity is not recovered and that all solution curves lie on one straight line in the y_1 - y_2 -plane. In Figure 6.4(B), we plot the desired and reconstructed nonlinearity on this straight line together with the error between them. We see that there is a very good alignment. Hence the final identification problem (2.3) is solved very precisely, however, this does not directly lead to a good reconstruction of the nonlinearity. The reason for this is, that the controls – and with this the data – were chosen poorly in the sense that they provide barely insights into the behavior of the nonlinearity. On the contrary, the controls that were found by our algorithm are able to reconstruct the nonlinearity very well as we have seen above. Furthermore, some of them remind us of the structure of the controls that we constructed in (6.3) by theoretical considerations. This is visualized in Figure 6.3. Hence, the algorithm seems to provide an effective tool to generate robust controls automatically.

An important property of the algorithm is that it tries to make the problem more convex. This is visualized in Figure 6.5. In both figures, the functional is plotted for different values of the coefficient in front of the basis elements y_1^2 and $y_1 y_2$. Clearly, applying the optimal controls it is evident that the functional is convex and therefore the problem has a unique minimum. For random controls, there



(A) Plot of reconstruction and error.



(B) Cut through the diagonal of the domain.

FIGURE 6.4. Results of robustness test case (using random controls).

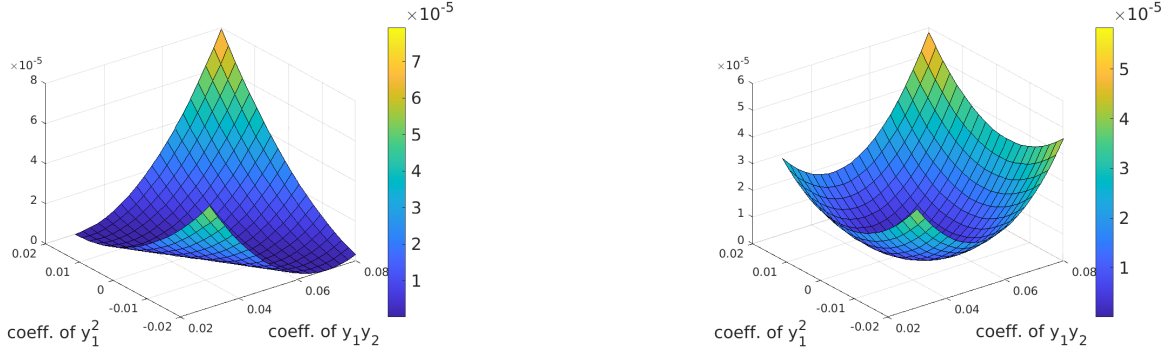
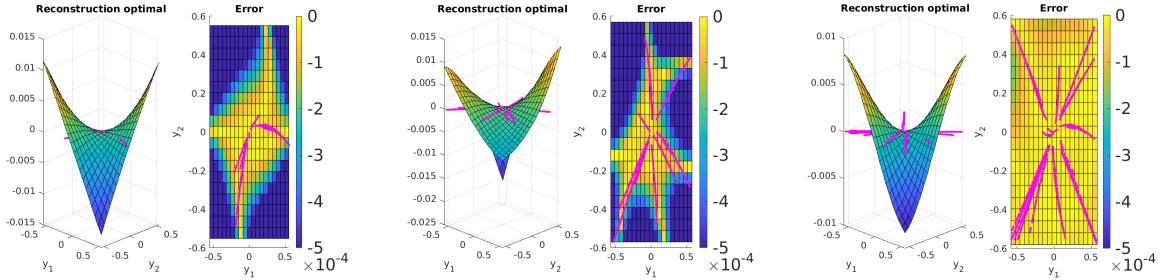


FIGURE 6.5. Making the problem more convex. Left: random controls; Right: optimal controls.

FIGURE 6.6. Reconstruction and error plots (cf. (6.2)) for $P = 2, 3, 5$ and the sinusoidal nonlinearity with the solutions curves in magenta.

seem to be infinitely local minima; at least the functional is quite flat in some directions. See also Tables 6.1-6.3 in [7].

6.2. Sinusoidal and exponential nonlinearity. Let us now consider the two remaining nonlinearities $G_\star^2(y) = 0.01 \sin(2y_1) \sin(2y_2)$ and $G_\star^3(y) = 0.01 \exp(2y_1) \exp(2y_2)$ as shown in Figures 6.1(B) and 6.1(C), respectively. We repeat the experiments from Section 6.1 and obtain the results shown in Figure 6.6 and Figure 6.7, respectively.

From the error plots in Figure 6.7, we observe that we need $P = 5$ to have a sensible reconstruction of the exponential nonlinearity. This is to be expected since the Taylor expansion of the exponential function consists of the sum of all monomials. Furthermore, we see more clearly that the error on

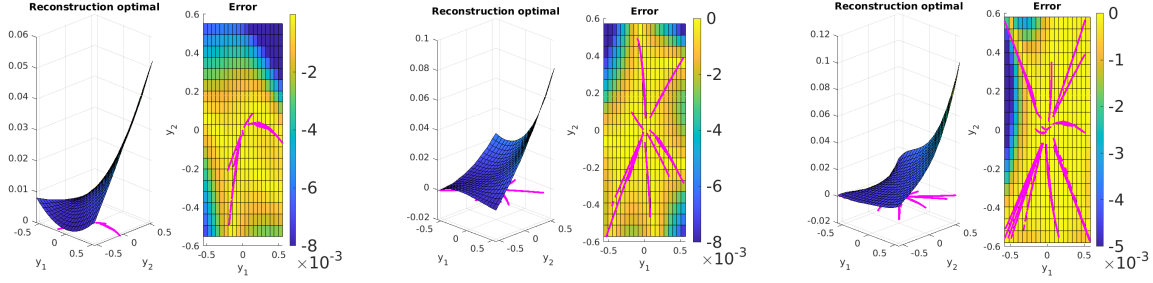


FIGURE 6.7. Reconstruction and error plots (cf. (6.2)) for $P = 2, 3, 5$ and the exponential nonlinearity with the solutions curves in magenta.

each coefficient in the Taylor expansion goes down when increasing the order of reconstruction P (cf. Figure 6.8). The reason is that in every iteration functions that are part of the Taylor function enter the reconstruction.

Since within Ω_M the sinusoidal nonlinearity is very similar to a bilinear one, the error is in principle smaller than for the exponential nonlinearity. We also observe that the error in Figure 6.6 is worse for $P = 3$ than for $P = 2$. This can be explained by the fact that uneven functions enter for $P = 3$ compared to $P = 2$. Notice also that the error in the first subfigure in Figure 6.6 is symmetric with respect to a diagonal line. Hence the error that this function imposes outside of the solution curves might be quite high.

6.3. Error in Taylor coefficients. To round up our numerical experiments, we discuss a different way of measuring the error of the final reconstruction: the absolute value of the difference of the coefficient of each monomial in the Taylor expansion of the true nonlinearity and the reconstructed nonlinearity, respectively.

More precisely, let the Taylor expansion up to order d^2 around the origin given by

$$(6.4) \quad \mathcal{T}_{d^2}(y_1, y_2) = \sum_{i_1=0}^d \sum_{i_2=0}^d t_{i_1, i_2} y_1^{i_1} y_2^{i_2}, \quad t_{i_1, i_2} = \frac{1}{i_1! i_2!} \left(\frac{\partial^{i_1+i_2}}{\partial y_1^{i_1} \partial y_2^{i_2}} g(0, 0) \right).$$

Now, we compare the coefficients $\alpha = (\alpha_{i_1, i_2})_{i_1, i_2=0}^d$ from our reconstruction of the polynomials given in (5.3) with the coefficients t_{i_1, i_2} of the Taylor expansion given in (6.4). The difference in each coefficient is shown in increasing order of the coefficients. In Figure 6.8, the error in the Taylor coefficients (up to the corresponding approximation order) is plotted. In all of these plots, we observe a saturation effect of the error for higher orders of degree. This can be seen most strongly for $P = 3$ and $P = 5$. The reason for this is that for a higher order of polynomial degree the influence of the corresponding monomial grows with the distance to the origin. Since we consider here a quite small domain, it is harder to reconstruct the correct coefficient.

7. CONCLUSION

In this work, we presented a greedy reconstruction algorithm to identify unknown operators in nonlinear elliptic models. We proved the Lipschitz continuity of the parameter-to-state map and its inverse. We performed several numerical experiments that successfully validate our proposed scheme.

This paper represents a significant step towards untangling the mysteries of unknown nonlinear operators within semilinear elliptic models. By harnessing the power of optimal control and active learning, we pave the way for a deeper understanding and more accurate predictions of complex physical phenomena governed by these models.

ACKNOWLEDGMENTS

We would like to express our gratitude to Jacob Körner (University of Würzburg), Jörg Weber (University of Vienna), and Behzad Azmi (University of Konstanz) for their beneficial remarks and fruitful discussions.

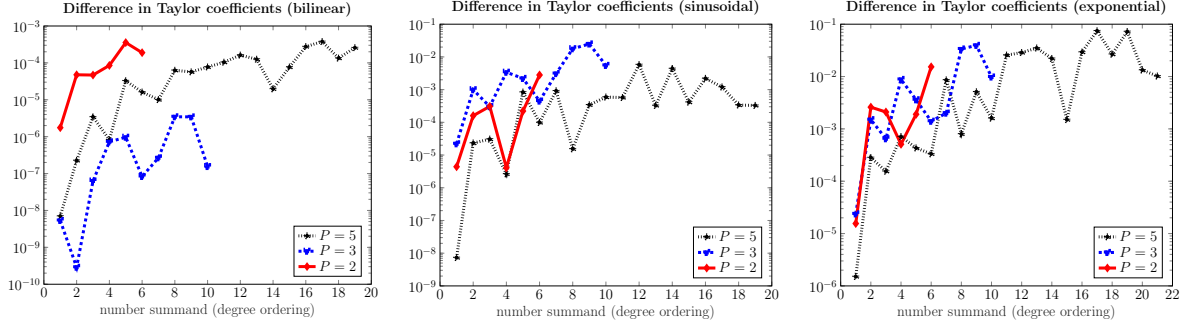


FIGURE 6.8. Difference in coefficients of the (true) Taylor expansion (cf. (6.4)) and the reconstructed nonlinearity for $P = 2, 3, 5$ for bilinear, sinusoidal and exponential nonlinearity.

Competing interests. The authors have no relevant financial or non-financial interests to disclose.
Author contributions. All four authors contributed equally to the present manuscript.

REFERENCES

- [1] ADAMS, R. A., AND FOURNIER, J. J. F. *Sobolev spaces*, second ed., vol. 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, 2003.
- [2] ALESSANDRINI, G. An identification problem for an elliptic equation in two variables. *Ann. Mat. Pura Appl. (4)* 145 (1986), 265–295.
- [3] ALT, H. W. *Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung*. Springer-Verlag, 2012.
- [4] BERETTA, E., AND VOGELIUS, M. An inverse problem originating from magnetohydrodynamics. *Arch. Rational Mech. Anal.* 115, 2 (1991), 137–152.
- [5] BONNANS, J. F., AND CASAS, E. Un principe de Pontryagine pour le contrôle des systèmes semilinéaires elliptiques. *J. Differ. Equ.* 90, 2 (1991), 288–303.
- [6] BRÉZIS, H., AND BROWDER, F. E. Strongly nonlinear elliptic boundary value problems. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 5, 3 (1978), 587–603.
- [7] BUCHWALD, S., CIARAMELLA, G., AND SALOMON, J. Analysis of a greedy reconstruction algorithm. *SIAM J. Control Optim.* 59, 6 (2021), 4511–4537.
- [8] BUCHWALD, S., CIARAMELLA, G., AND SALOMON, J. Gauss–Newton Oriented Greedy Algorithms for the Reconstruction of Operators in Nonlinear Dynamics. *SIAM J. Control Optim.* 62, 3 (2024), 1343–1368.
- [9] BUCHWALD, S., CIARAMELLA, G., SALOMON, J., AND SUGNY, D. Greedy reconstruction algorithm for the identification of spin distribution. *Phys. Rev. A* 104, 6 (2021), Paper No. 063112, 9.
- [10] BUCHWALD, S., CIARAMELLA, G., SALOMON, J., AND SUGNY, D. A SPIRED code for the reconstruction of spin distribution. *Computer Physics Communications* (2024), 109126.
- [11] CASAS, E., MATEOS, M., AND RÖSCH, A. Analysis of control problems of nonmontone semilinear elliptic equations. *ESAIM Control Optim. Calc. Var.* 26 (2020), Paper No. 80, 21.
- [12] CASAS, E., AND TRÖLTZSCH, F. First- and second-order optimality conditions for a class of optimal control problems with quasilinear elliptic equations. *SIAM J. Control Optim.* 48, 2 (2009), 688–718.
- [13] CASAS, E., AND WACHSMUTH, D. A note on existence of solutions to control problems of semilinear partial differential equations. *SIAM J. Control Optim.* 61, 3 (2023), 1095–1112.
- [14] CIARAMELLA, G., AND GANDER, M. J. *Iterative methods and preconditioners for systems of linear equations*, vol. 19 of *Fundamentals of Algorithms*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2022.
- [15] CLERMONT, G., AND ZENKER, S. The inverse problem in mathematical biology. *Math. Biosci.* 260 (2015), 11–15.
- [16] EVANS, L. C. *Partial differential equations*, second ed., vol. 19 of *Graduate Studies in Mathematics*. Amer. Math. Soc., Providence, RI, Providence, RI, 2010.
- [17] GAMBINO, G., LOMBARDI, M. C., AND SAMMARTINO, M. A velocity-diffusion method for a Lotka-Volterra system with nonlinear cross and self-diffusion. *Appl. Numer. Math.* 59, 5 (2009), 1059–1074.
- [18] GRISVARD, P. *Elliptic problems in nonsmooth domains*, vol. 69 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011. Reprint of the 1985 original [MR0775683], With a foreword by Susanne C. Brenner.
- [19] HABER, E., ASCHER, U. M., AND OLDENBURG, D. On optimization techniques for solving nonlinear inverse problems. *Inverse Problems* 16, 5 (2000), 1263–1280. Electromagnetic imaging and inversion of the Earth’s subsurface.
- [20] IBÁÑEZ, A. Optimal control of the lotka–volterra system: turnpike property and numerical simulations. *J. Biol. Dyn.* 11, 1 (2017), 25–41.
- [21] ISAKOV, V. *Inverse problems for partial differential equations*, third ed., vol. 127 of *Applied Mathematical Sciences*. Springer, Cham, 2017.

- [22] ISAKOV, V., AND NACHMAN, A. I. Global uniqueness for a two-dimensional semilinear elliptic inverse problem. *Trans. Amer. Math. Soc.* 347, 9 (1995), 3375–3390.
- [23] JOHANSSON, D., NURMINEN, J., AND SALO, M. Inverse problems for semilinear elliptic pde with a general nonlinearity $a(x, u)$. *arXiv preprint arXiv:2312.12196* (2023).
- [24] JOVANOVIĆ, B. S. S., AND SÜLI, E. *Analysis of finite difference schemes*, vol. 46 of *Springer Series in Computational Mathematics*. Springer, London, 2014. For linear partial differential equations with generalized solutions.
- [25] KUMPATI, S. N., KANNAN, P., ET AL. Identification and control of dynamical systems using neural networks. *IEEE trans. neural netw.* 1, 1 (1990), 4–27.
- [26] MADAY, Y., AND SALOMON, J. A greedy algorithm for the identification of quantum systems. In *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference* (2009), IEEE, pp. 375–379.
- [27] MANDT, S., HOFFMAN, M. D., AND BLEI, D. M. Stochastic gradient descent as approximate Bayesian inference. *J. Mach. Learn. Res.* 18 (2017), Paper No. 134, 35.
- [28] MARTIN, JR., R. H., AND PIERRE, M. Nonlinear reaction-diffusion systems. In *Nonlinear equations in the applied sciences*, vol. 185 of *Math. Sci. Engrg.* Academic Press, Boston, MA, 1992, pp. 363–398.
- [29] MORTUJA, M. G., CHAUBE, M. K., AND KUMAR, S. Dynamic analysis of a predator-prey system with nonlinear prey harvesting and square root functional response. *Chaos Solitons Fractals* 148 (July 2021), 111071.
- [30] PAO, C. V. On nonlinear reaction-diffusion systems. *J. Math. Anal. Appl.* 87, 1 (1982), 165–198.
- [31] QUARANTA, G., LACARBONARA, W., AND MASRI, S. F. A review on computational intelligence for identification of nonlinear dynamical systems. *Nonlinear Dyn.* 99, 2 (2020), 1709–1761.
- [32] ROCKAFELLAR, R. T., AND WETS, R. J.-B. *Variational analysis*, vol. 317 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.
- [33] RÖSCH, A. Stability estimates for the identification of nonlinear heat transfer laws. *Inverse Problems* 12, 5 (1996), 743–756.
- [34] SANTOSA, F., AND WEITZ, B. An inverse problem in reaction kinetics. *J. Math. Chem.* 49, 8 (2011), 1507–1520.
- [35] SEVER, M. An existence theorem for some semilinear elliptic systems. *J. Differential Equations* 226, 2 (2006), 572–593.
- [36] TRÖLTZSCH, F. *Optimal control of partial differential equations: Theory, methods, and applications*, vol. 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, Providence, RI, 2010.
- [37] YANG, F., ROCHAU, F., HUBER, J. S., BRIEUSSEL, A., RASTELLI, G., WEIG, E. M., AND SCHEER, E. Spatial Modulation of Nonlinear Flexural Vibrations of Membrane Resonators. *Phys. Rev. Lett.* 122, 15 (Apr. 2019), 154301.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

- 38/2024** Tonini, A., Regazzoni, F., Salvador, M., Dede', L., Scrofani, R., Fusini, L., Cogliati, C., Pontone, G., Vergara, C., Quarteroni, A.
Two new calibration techniques of lumped-parameter mathematical models for the cardiovascular system
- Fumagalli, A.; Patacchini, F.S.
Numerical validation of an adaptive model for the determination of nonlinear-flow regions in highly heterogeneous porous media
- 37/2024** Begu, B.; Panzeri, S.; Arnone, E.; Carey, M.; Sangalli, L.M.
A nonparametric penalized likelihood approach to density estimation of space-time point patterns
- 36/2024** Torri, V.; Ercolanoni, M.; Bortolan, F.; Leoni, O.; Ieva, F.
A NLP-based semi-automatic identification system for delays in follow-up examinations: an Italian case study on clinical referrals
- 34/2024** Corti, M.
Exploring tau protein and amyloid-beta propagation: a sensitivity analysis of mathematical models based on biological data
- 35/2024** Botti, L.; Botti, M.; Di Pietro, D.A.; Massa, F.C.
Stability, convergence, and pressure-robustness of numerical schemes for incompressible flows with hybrid velocity and pressure
- Corti, M.
Exploring tau protein and amyloid-beta propagation: a sensitivity analysis of mathematical models based on biological data
- 32/2024** Ziarelli, G.; Parolini, N.; Verani, M.
Learning epidemic trajectories through Kernel Operator Learning: from modelling to optimal control
- 31/2024** Gambarini, M.; Agate, G.; Ciaramella, G.; Miglio, E.; Maran, S.
Modeling and optimization for arrays of water turbine OWC devices
- 30/2024** Gregorio, C.; Rea, F.; Ieva, F.; Scagnetto, A.; Indennidade, C.; Cappelletto, C.; Di Lenarda, A.; Barbati, G.
Flexible approaches based on multi-state models and microsimulation to perform real-world

cost-effectiveness analyses: an application to PCSK9-inhibitors