GAUSS-NEWTON ORIENTED GREEDY ALGORITHMS FOR THE RECONSTRUCTION OF OPERATORS IN NONLINEAR DYNAMICS

S. BUCHWALD*, G. CIARAMELLA†, AND J. SALOMON‡

Abstract. This paper is devoted to the development and convergence analysis of greedy reconstruction algorithms based on the strategy presented in [Y. Maday and J. Salomon, Joint Proceedings of the 48th IEEE Conference on Decision and Control and the 28th Chinese Control Conference, 2009, pp. 375–379]. These procedures allow the design of a sequence of control functions that ease the identification of unknown operators in nonlinear dynamical systems. The original strategy of greedy reconstruction algorithms is based on an offline/online decomposition of the reconstruction process and an ansatz for the unknown operator obtained by an a priori chosen set of linearly independent matrices. In the previous work [S. Buchwald, G. Ciaramella and J. Salomon, SIAM J. Control Optim., 59(6), pp. 4511-4537], convergence results were obtained in the case of linear identification problems. We tackle here the more general case of nonlinear systems. More precisely, we introduce a new greedy algorithm based on the linearized system. We show that the controls obtained with this new algorithm lead to the local convergence of the classical Gauss-Newton method applied to the online nonlinear identification problem. We then extend this result to the controls obtained on nonlinear systems where a local convergence result is also proved. The main convergence results are obtained for dynamical systems with linear and bilinear control structures.

Key words. Gauss-Newton method, operator reconstruction, Hamiltonian identification, quantum control problems, inverse problems, greedy reconstruction algorithm, control theory

AMS subject classifications. 65K10, 65K05, 81Q93, 34A55, 49N45, 34H05, 93B05, 93B07

1. Introduction.

This paper is concerned with the development and the analysis of a new class of numerical methods for the operator reconstruction in controlled nonlinear differential systems. The identification of unknown operators and parameters characterizing dynamical systems is a typical problem in several fields of applied science. In general, this is understood as an inverse problem, where the goal is to best fit simulated and experimental data. However, when a system is affected by input forces controlled by an external user, the data used in the fitting process can be manipulated. If the input forces are not properly chosen, the fitting process can result in a very poor quality of the reconstructed operators. Thus, it is natural to look for a set of input forces that allows one to generate good data permitting the best possible reconstruction. This is a typical case in the field Hamiltonian identification [5, 10, 12, 19–23, 32, 34–36], or in engineering in the context of state space realization [18, 26, 31] and optimal design of experiments [1, 4, 7].

In this paper, we focus on the analysis and development of a class of greedy-type reconstruction algorithms (GR) that were introduced in [29] for Hamiltonian identification problems, further developed and analyzed in [13], and later adapted to the identification of probability distributions for parameters in the context of quantum systems in [14]. This approach decomposes the identification process into an offline phase, where the control functions are computed by a GR algorithm, and an online phase, where the controls are used to generate experimental data to be used in an inverse problem for the final reconstruction of the unknown operator. In [13], a

*Mathematics Department, Universität Konstanz, Germany (simon.buchwald@uni-konstanz.de).
†MOX, Dipartimento di Matematica, Politecnico di Milano (gabriele.ciaramella@polimi.it).
‡ANGE team, INRIA Paris, France (julien.salomon@inria.fr).

This is the Author Accepted Manuscript (AAM). The published version is available since May 3, 2024, at https://doi.org/10.1137/23M1552929 under subscription to SIAM: Journal of Control and Optimization, Vol. 62, No. 3, pp. 1343–1368.

first detailed convergence analysis of this strategy was provided for the identification of the control matrix in a linear input/output system. Based on this analysis, the authors developed a new more efficient and robust numerical variant of the standard greedy reconstruction algorithm. It was then shown in [14] that this strategy is also able to reconstruct the probability distribution of control inhomogeneities for a spin ensemble in Nuclear Magnetic Resonance; see, e.g., [11, 24]. Even though the mentioned references focus also on GR algorithms for nonlinear problems, none of them presents a theoretical analysis for nonlinear problems.

When dealing with nonlinear problems it is common to consider linearization techniques. Classical methods for nonlinear problems, like the Gauss-Newton method (GN) [27], are often derived from such linearizations. Following this approach, we introduce a new class of GR algorithms, derived from the linearization of the dynamical system in a neighborhood of the unknown operator. We call these methods LGR (linearized greedy reconstruction algorithms) and present a detailed theoretical analysis. This is the first novelty of this work. Our analysis of LGR reveals that its hidden goal is to construct a set of control functions that attempt to make a specific matrix full rank in a neighborhood of the solution. It turns out that this matrix is actually the GN matrix, namely the Jacobian of the nonlinear residual. In this respect, the second main novelty of this paper is that we can interpret our LGR procedure as a process that computes controls making GN locally well defined and convergent, and thereby the online nonlinear reconstruction problem (locally) solvable. Thus, once the control functions are computed by LGR, GN can be used with guaranteed convergence. We provide a detailed analysis of LGR for two classes of problems: the reconstruction of the drift matrix in linear input/output systems and the reconstruction of an Hamiltonian matrix in skew-symmetric bilinear systems. Both cases represent truly nonlinear problems, since the unknown operators act on the states of the systems. Notice that our analysis for the drift matrix is also valid in the case of the reconstruction of the control matrix in a linear input/output systems, hence includes the case considered in [13, sect. 5]. Thus, this part of the present work is a substantial extension of the results of [13]. The development and analysis of LGR turns out to be also useful to obtain a first analysis of the original GR algorithm applied to nonlinear systems. This is the third main novelty of this paper, and it is achieved by relating the behaviors of GR and LGR: under appropriate controllability and observability assumptions, we show that the controls generated by GR are suitable also for LGR and thus make the GN Jacobian matrix full rank. The hypotheses of this analysis are studied for three classes of problems: the reconstruction of the drift matrix in linear input/output systems, the reconstruction of an Hamiltonian matrix in skew-symmetric bilinear systems, and the reconstruction of the control matrix in nonlinear dynamical systems with linear control structure. The two GR and LGR approaches are compared by direct numerical experiments and by a global convergence analysis in a specific case. These show that GR and LGR are comparable when working locally near the solution. However, GR applied directly to the original nonlinear system is superior when only poor information about the solution is available.

The paper is organized as follows. In section 2, the notation used throughout this work is fixed. Section 3 describes linearized systems and recalls GN for reconstruction problems. LGR is introduced in section 4. In sections 5 and 6, we present analyses of LGR for the reconstruction of drift matrices in linear systems and Hamiltonian matrices in bilinear systems, respectively. Section 7 focuses on GR for nonlinear problems, and a corresponding analysis is provided in section 7.1. In section 7.2, we recall and extend an optimized greedy reconstruction (OGR) algorithm introduced in [13]. In section 7.3, a global analysis is provided for a specific case of a system with bilinear control structure. LGR, GR and OGR are tested numerically in section 8.

2. Notation. Consider a positive natural number N. We denote by $\langle \mathbf{v}, \mathbf{w}\rangle := \mathbf{v}^\top \mathbf{w}$, for any $\mathbf{v}, \mathbf{w} \in \mathbb{R}^N$, the usual real scalar product on $\mathbb{R}^N$, and by $\|\cdot\|_2$ the corresponding norm. For any $A \in \mathbb{R}^{N\times N}$, $[A]_{j,k}$ is the j,k (with $j,k \leq N$) entry of A, and the notation $A_{[1:k,1:j]}$ indicates the upper left submatrix of A of size $k \times j$, namely, $[A_{[1:k,1:j]}]_{\ell,m} := [A]_{\ell,m}$ for $\ell = 1,\dots,k$ and $m = 1,\dots,j$. Similarly, $A_{[1:k,j]}$ denotes the column vector in $\mathbb{R}^k$ corresponding to the first k elements of the column j of A. Additionally, $\mathrm{im}(A)$ is the image of A, and $\ker(A)$ its kernel. We indicate by $\mathfrak{so}(N)$ the space of skew-symmetric matrices in $\mathbb{R}^{N\times N}$. Moreover, when talking about symmetric matrices, PD and PSD stand for positive definite and semidefinite, respectively. By (A, B, C) we denote the input/output dynamical system

$$\boldsymbol{x}(t) = C\boldsymbol{y}(t), \quad \dot{\boldsymbol{y}}(t) = A\boldsymbol{y}(t) + B\boldsymbol{\epsilon}(t), \quad \boldsymbol{y}(0) = \boldsymbol{y}^0. \tag{2.1}$$

The controls $\boldsymbol{\epsilon}$ belong to the admissible set E_{ad}, which is assumed to be a non-empty and weakly compact subset of $L^2(0,T;\mathbb{R}^M)$, $M \in \mathbb{N}$. This guarantees well posedness of the optimal control problems considered in this work. Moreover, we assume that E_{ad} contains $\boldsymbol{\epsilon} \equiv 0$ as an interior point.[1] This hypothesis is used in our analysis in sections 5, 6 and 7. For an interval $X \subset \mathbb{R}$, the notation $\phi : X \rightrightarrows \mathbb{R}^N$ indicates that ϕ is a set-valued correspondence, i.e. $\phi(x) \subset \mathbb{R}^N$ is a set for $x \in X$. Finally, $\mathcal{B}_r^N(x) \subset \mathbb{R}^N$ is the ball with radius $r > 0$ and center $x \in \mathbb{R}^N$.

3. Linearized systems and Gauss-Newton method (GN). Consider the system of ordinary differential equations (ODE)

$$\dot{\boldsymbol{y}}(t) = f(A_\star, \boldsymbol{y}(t), \boldsymbol{\epsilon}(t)),\ t \in (0,T], \quad \boldsymbol{y}(0) = \boldsymbol{y}^0, \tag{3.1}$$

where $\boldsymbol{y}(t) \in \mathbb{R}^N$ with $N \in \mathbb{N}$, $\boldsymbol{y}^0 \in \mathbb{R}^N$ and $\boldsymbol{\epsilon} \in E_{ad}$. The operator $A_\star$ is unknown and assumed to lie in the space spanned by a finite-dimensional set $\mathcal{A} = \{A_1, \dots, A_K\}$, $K \in \mathbb{N}$, and we write $A_\star = \sum_{j=1}^K \boldsymbol{\alpha}_{\star,j} A_j =: A(\boldsymbol{\alpha}_\star)$. We assume that $f : \mathrm{span}(\mathcal{A}) \times \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}^N, (A, \boldsymbol{y}, \boldsymbol{\epsilon}) \mapsto f(A, \boldsymbol{y}, \boldsymbol{\epsilon})$ is differentiable in A and $\boldsymbol{y}$.

To identify the unknown operator $A_\star$ one uses a set of control functions $(\boldsymbol{\epsilon}^m)_{m=1}^K \subset E_{ad}$ to perform K laboratory experiments and obtain the experimental data

$$\boldsymbol{\varphi}^\star_{data}(\boldsymbol{\epsilon}^m) := C\boldsymbol{y}(A_\star, \boldsymbol{\epsilon}^m; T),\ \text{for } m = 1, \dots, K. \tag{3.2}$$

Here, $\boldsymbol{y}(A_\star, \boldsymbol{\epsilon}; T)$ denotes the solution to (3.1) at time $T > 0$, corresponding to the operator $A_\star$ and a control function $\boldsymbol{\epsilon}$. The matrix $C \in \mathbb{R}^{P\times N}$ $(P \leq N)$ is a given observer matrix. The measurements are assumed not to be affected by noise.

Using the set $(\boldsymbol{\epsilon}^m)_{m=1}^K$ and the data $(\boldsymbol{\varphi}^\star_{data}(\boldsymbol{\epsilon}^m))_{m=1}^K \subset \mathbb{R}^P$, the unknown vector $\boldsymbol{\alpha}$ is obtained by solving the least-squares problem

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^K} \frac{1}{2} \sum_{m=1}^K \|\boldsymbol{\varphi}^\star_{data}(\boldsymbol{\epsilon}^m) - C\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; T)\|_2^2\,. \tag{3.3}$$

GN is a typical iterative strategy to solve problems of the form (3.3), and its process is initialized by a vector which we will call $\boldsymbol{\alpha}_\circ \in \mathbb{R}^K$. We denote by $\boldsymbol{\alpha}_c \in \mathbb{R}^K$ the GN iterate, and define $f_m(\boldsymbol{\alpha}) := \frac{1}{2}\sum_{i=1}^P \|(R_m(\boldsymbol{\alpha}))_i\|_2^2 = \frac{1}{2} R_m(\boldsymbol{\alpha})^\top R_m(\boldsymbol{\alpha})$, where

$$R_m(\boldsymbol{\alpha}) := C\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; T) - \boldsymbol{\varphi}^\star_{data}(\boldsymbol{\epsilon}^m), \quad \text{for } m \in \{1, \dots, K\}. \tag{3.4}$$

[1] This is a reasonable assumption, since it is satisfied, e.g., for standard box constraints.

Thus, (3.3) is equivalent to $\min_{\boldsymbol{\alpha}\in\mathbb{R}^K}\sum_{m=1}^K f_m(\boldsymbol{\alpha})$. Given an iterate $\boldsymbol{\alpha}_c\in\mathbb{R}^K$ and the Jacobian $R_m'(\boldsymbol{\alpha}_c)\in\mathbb{R}^{P\times K}$ of R_m at $\boldsymbol{\alpha}_c$, GN finds the new iterate solving

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^K}\sum_{m=1}^K \|R_m'(\boldsymbol{\alpha}_c)(\boldsymbol{\alpha}-\boldsymbol{\alpha}_c)-R_m(\boldsymbol{\alpha}_c)\|_2^2, \tag{3.5}$$

whose first-order optimality condition is $\widehat{W}_c\boldsymbol{\alpha}=\sum_{m=1}^K R_m'(\boldsymbol{\alpha}_c)^\top R_m(\boldsymbol{\alpha}_c)$, where $\widehat{W}_c := \sum_{m=1}^K R_m'(\boldsymbol{\alpha}_c)^\top R_m'(\boldsymbol{\alpha}_c)\in\mathbb{R}^{K\times K}$ is symmetric PSD. Now, we recall the following convergence result from [28, Thm. 2.4.1].

LEMMA 3.1 (local convergence of GN). *Let $\boldsymbol{\alpha}_\star$ solve $\min_{\boldsymbol{\alpha}\in\mathbb{R}^K}\sum_{m=1}^K f_m(\boldsymbol{\alpha})$ such that for all $m\in\{1,\dots,K\}$ the function R_m is Lipschitz continuously differentiable near $\boldsymbol{\alpha}_\star$ and $R_m(\boldsymbol{\alpha}_\star)=0$. If the initialization vector $\boldsymbol{\alpha}_\circ\in\mathbb{R}^K$ is sufficiently close to $\boldsymbol{\alpha}_\star$, and $\widehat{W}_c$ is PD for all iterates $\boldsymbol{\alpha}_c\in\mathbb{R}^K$, then GN converges quadratically to $\boldsymbol{\alpha}_\star$.*

Lemma 3.1 implies that, given an initialization vector $\boldsymbol{\alpha}_\circ$ sufficiently close to the solution $\boldsymbol{\alpha}_\star$, the functions $(\boldsymbol{\epsilon}^m)_{m=1}^K$ should be chosen such that the GN matrix $\widehat{W}_c = \sum_{m=1}^K R_{\boldsymbol{\epsilon}^m}'(\boldsymbol{\alpha}_c)^\top R_{\boldsymbol{\epsilon}^m}'(\boldsymbol{\alpha}_c)$ is PD for all $\boldsymbol{\alpha}_c\in\mathbb{R}^K$ in a neighborhood of $\boldsymbol{\alpha}_\star$. Notice that $\widehat{W}_c$ being PD is equivalent to (3.5) being uniquely solvable. Using (3.4), we can write (3.5) more explicitly. For a direction $\delta\boldsymbol{\alpha}\in\mathbb{R}^K$, we have

$$R_m'(\boldsymbol{\alpha}_c)(\delta\boldsymbol{\alpha}) = C\delta\boldsymbol{y}_c(A(\delta\boldsymbol{\alpha}),\boldsymbol{\epsilon}^m;T), \tag{3.6}$$

where $\delta\boldsymbol{y}_c(A(\delta\boldsymbol{\alpha}),\boldsymbol{\epsilon}^m;T)$ denotes the solution at time T to the linearized equation

$$\begin{cases}\delta\dot{\boldsymbol{y}}_c = \partial_{\boldsymbol{y}} f(A(\boldsymbol{\alpha}_c),\boldsymbol{y}_c,\boldsymbol{\epsilon})\delta\boldsymbol{y}_c + \sum_{j=1}^K \delta\boldsymbol{\alpha}_j\Big(\partial_A f(A(\boldsymbol{\alpha}_c),\boldsymbol{y}_c,\boldsymbol{\epsilon})(A_j)\Big),\ \delta\boldsymbol{y}_c(0)=0,\\ \dot{\boldsymbol{y}}_c = f(A(\boldsymbol{\alpha}_c),\boldsymbol{y}_c,\boldsymbol{\epsilon}),\quad \boldsymbol{y}_c(0)=\boldsymbol{y}^0,\end{cases} \tag{3.7}$$

at $\boldsymbol{\epsilon}=\boldsymbol{\epsilon}^m$. Hence, problem (3.5) can be written as

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^K}\sum_{m=1}^K \|C\delta\boldsymbol{y}_c(A(\boldsymbol{\alpha}-\boldsymbol{\alpha}_c),\boldsymbol{\epsilon}^m;T)-R_m(\boldsymbol{\alpha}_c)\|_2^2. \tag{3.8}$$

The vectors $R_m(\boldsymbol{\alpha}_c)\in\mathbb{R}^P$ are independent of $\boldsymbol{\alpha}$ and can be considered as fixed data when solving (3.8). Now, we recall that the GR algorithm [13, 29] was designed to generate controls $(\boldsymbol{\epsilon}^m)_{m=1}^K$ that make problems of the form (3.8) uniquely solvable.

4. A linearized GR algorithm (LGR).

Let us assume to be provided with a vector $\boldsymbol{\alpha}_\circ$ sufficiently close to $\boldsymbol{\alpha}_\star$ (recall from section 3 that $A_\star = A(\boldsymbol{\alpha}_\star)$). Further, let $\delta\boldsymbol{y}_\circ(A(\boldsymbol{\alpha}-\boldsymbol{\alpha}_\circ),\boldsymbol{\epsilon}^m;T)$ denote the solution at time T to

$$\begin{cases}\delta\dot{\boldsymbol{y}}_\circ = \partial_{\boldsymbol{y}} f(A(\boldsymbol{\alpha}_\circ),\boldsymbol{y}_\circ,\boldsymbol{\epsilon})\delta\boldsymbol{y}_\circ + \sum_{j=1}^K (\boldsymbol{\alpha}_j-\boldsymbol{\alpha}_{\circ,j})\Big(\partial_A f(A(\boldsymbol{\alpha}_\circ),\boldsymbol{y}_\circ,\boldsymbol{\epsilon})(A_j)\Big),\ \delta\boldsymbol{y}_\circ(0)=0,\\ \dot{\boldsymbol{y}}_\circ = f(A(\boldsymbol{\alpha}_\circ),\boldsymbol{y}_\circ,\boldsymbol{\epsilon}),\quad \boldsymbol{y}_\circ(0)=\boldsymbol{y}^0.\end{cases} \tag{4.1}$$

The goal is to generate control functions $(\boldsymbol{\epsilon}^m)_{m=1}^K$ such that (3.8) in $\boldsymbol{\alpha}_\circ$, that is

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^K}\sum_{m=1}^K \|C\delta\boldsymbol{y}_\circ(A(\boldsymbol{\alpha}-\boldsymbol{\alpha}_\circ),\boldsymbol{\epsilon}^m;T)-R_m(\boldsymbol{\alpha}_\circ)\|_2^2, \tag{4.2}$$

is uniquely solvable. Then, in section 5.2, we show that if (4.2) is uniquely solvable, the same holds for (3.8) at all $\boldsymbol{\alpha}$ in a neighborhood of $\boldsymbol{\alpha}_\circ$. Thus, if $\boldsymbol{\alpha}_\circ$ is an initialization vector for GN, then (3.8) is uniquely solvable for all iterates $\boldsymbol{\alpha}_c$ of GN.

Algorithm 4.1 Linearized Greedy Reconstruction Algorithm (LGR)

Require: A set of linearly independent operators $\mathcal{A} = \{A_1, \dots, A_K\}$.

1: Compute the control $\boldsymbol{\epsilon}^1$ by solving

$$\max_{\boldsymbol{\epsilon}\in E_{ad}} \|C\delta\boldsymbol{y}_\circ(A_1,\boldsymbol{\epsilon};T)\|_2^2\,. \tag{4.3}$$

2: **for** $k = 1, \dots, K-1$ **do**

3: Fitting step: Let $A^{(k)}(\boldsymbol{\beta}) := \sum_{j=1}^k \boldsymbol{\beta}_j A_j$, find $\boldsymbol{\beta}^k = (\boldsymbol{\beta}_j^k)_{j=1,\dots,k}$ that solves

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^k} \sum_{m=1}^k \left\| C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}),\boldsymbol{\epsilon}^m;T) - C\delta\boldsymbol{y}_\circ(A_{k+1},\boldsymbol{\epsilon}^m;T)\right\|_2^2. \tag{4.4}$$

4: Splitting step: Find $\boldsymbol{\epsilon}^{k+1}$ that solves

$$\max_{\boldsymbol{\epsilon}\in E_{ad}} \left\| C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k),\boldsymbol{\epsilon};T) - C\delta\boldsymbol{y}_\circ(A_{k+1},\boldsymbol{\epsilon};T)\right\|_2^2. \tag{4.5}$$

5: **end for**

The set $(\boldsymbol{\epsilon}^m)_{m=1}^K$ is computed by Algorithm 4.1. This is the original GR algorithm from [29] applied to (4.2), which involves the linearized equations (4.1). For this reason, we call it Linearized Greedy Reconstruction algorithm (LGR). As for the original GR, the heuristics of LGR is that the set $(\boldsymbol{\epsilon}^m)_{m=1}^K$ must allow distinguishing the states of the system (4.1) corresponding to any two matrices $A(\widehat{\boldsymbol{\alpha}})$ and $A(\widetilde{\boldsymbol{\alpha}})$, for $\widehat{\boldsymbol{\alpha}} \neq \widetilde{\boldsymbol{\alpha}}$. Suppose that the first k controls $(\boldsymbol{\epsilon}^m)_{m=1}^k$ are computed, the new $\boldsymbol{\epsilon}^{k+1}$ is obtained by solving first problem (4.4), identifying two states that cannot be distinguished by the controls $(\boldsymbol{\epsilon}^m)_{m=1}^k$, and then problem (4.5), computing $\boldsymbol{\epsilon}^{k+1}$ with the goal of distinguishing these two states. For more details see [13, 29].

Our goal is to prove that the set $(\boldsymbol{\epsilon}^m)_{m=1}^K$ makes $\widehat{W}_\circ := \sum_{m=1}^K R_m'(\boldsymbol{\alpha}_\circ)^\top R_m'(\boldsymbol{\alpha}_\circ)$ PD, and thus (4.2) uniquely solvable. From (4.1), we have that $\delta\boldsymbol{y}_\circ$ is linear in $\boldsymbol{\alpha}$. Thus, $R_m'(\boldsymbol{\alpha}_\circ)(\delta\boldsymbol{\alpha}) = \delta\boldsymbol{y}_\circ(A(\delta\boldsymbol{\alpha}),\boldsymbol{\epsilon}^m;T) = \sum_{j=1}^K \delta\boldsymbol{\alpha}_j C\delta\boldsymbol{y}_\circ(A_j,\boldsymbol{\epsilon}^m;T)$. Hence, $R_m'(\boldsymbol{\alpha}_\circ)$ is a matrix with columns $R_m'(\boldsymbol{\alpha}_\circ)_j = C\delta\boldsymbol{y}_\circ(A_j,\boldsymbol{\epsilon}^m;T)$ for $j = 1, \dots, K$, and hence

$$[\widehat{W}_\circ]_{i,j} = \sum_{m=1}^K \langle C\delta\boldsymbol{y}_\circ(A_i,\boldsymbol{\epsilon}^m;T), C\delta\boldsymbol{y}_\circ(A_j,\boldsymbol{\epsilon}^m;T)\rangle, \quad i,j\in\{1,\dots,K\}. \tag{4.6}$$

Using (4.6), we can rewrite (4.3), (4.4) and (4.5) in a matrix form in Lemma 4.1, whose proof is similar to the one of [13, Lem. 5.12] and is omitted for brevity.

LEMMA 4.1 (Algorithm 4.1 in matrix form). *Consider Algorithm 4.1. Then:*

- *The initialization problem (4.3) is equivalent to*

$$\max_{\boldsymbol{\epsilon}\in E_{ad}} [W_\circ(\boldsymbol{\epsilon})]_{1,1}, \tag{4.7}$$

where $[W_\circ(\boldsymbol{\epsilon})]_{i,j} := \langle C\delta\boldsymbol{y}_\circ(A_i,\boldsymbol{\epsilon};T), C\delta\boldsymbol{y}_\circ(A_j,\boldsymbol{\epsilon};T)\rangle$ *for* $i,j\in\{1,\dots,K\}$.

- *Let* $\widehat{W}_\circ^{(k)} := \sum_{m=1}^k W_\circ(\boldsymbol{\epsilon}^m)$, *the fitting-step problem (4.4) is equivalent to*

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^k} \langle\boldsymbol{\beta}, [\widehat{W}_\circ^{(k)}]_{[1:k,1:k]}\boldsymbol{\beta}\rangle - 2\langle[\widehat{W}_\circ^{(k)}]_{[1:k,k+1]},\boldsymbol{\beta}\rangle. \tag{4.8}$$

- *Let* $\boldsymbol{v} := [(\boldsymbol{\beta}^k)^\top,\ -1]^\top$, *the splitting-step problem (4.5) is equivalent to*

$$\max_{\boldsymbol{\epsilon}\in E_{ad}} \langle\boldsymbol{v}, [W_\circ(\boldsymbol{\epsilon})]_{[1:k+1,1:k+1]}\boldsymbol{v}\rangle. \tag{4.9}$$

Moreover, problems (4.3)-(4.7), (4.4)-(4.8), and (4.5)-(4.9) are well posed.

The matrix representation given in Lemma 4.1 allows us to nicely describe the mathematical mechanism behind Algorithm 4.1 (see also [13, sect. 5.1]). Assume that at the k-th iteration the set $(\boldsymbol{\epsilon}_m)_{m=1}^k$ has been computed, the submatrix $[\widehat{W}_\circ^{(k)}]_{[1:k,1:k]}$ is PD and $[\widehat{W}_\circ^{(k)}]_{[1:k+1,1:k+1]}$ has a nontrivial (one-dimensional) kernel. Then the fitting step of Algorithm 4.1 identifies this nontrivial kernel. This can be proved by the following technical lemma (for a proof see [13, Lem. 5.3]).

LEMMA 4.2 (kernel of some symmetric PSD matrices). *Consider a symmetric PSD matrix* $\widetilde{G} = \begin{bmatrix} G & \boldsymbol{b} \\ \boldsymbol{b}^\top & c \end{bmatrix} \in \mathbb{R}^{n\times n}$*, where* $G \in \mathbb{R}^{(n-1)\times(n-1)}$ *is symmetric PD, and* $\boldsymbol{b} \in \mathbb{R}^{n-1}$ *and* $c \in \mathbb{R}$ *are such that* $\ker(\widetilde{G})$ *is nontrivial. Then* $\ker(\widetilde{G}) = \operatorname{span}\left\{ \begin{bmatrix} G^{-1}\boldsymbol{b} \\ -1 \end{bmatrix} \right\}$.

In our case, we have $\widetilde{G} = [\widehat{W}_\circ^{(k)}]_{[1:k+1,1:k+1]}$, $G = [\widehat{W}_\circ^{(k)}]_{[1:k,1:k]}$ and $\boldsymbol{b} = [\widehat{W}_\circ^{(k)}]_{[1:k,k+1]}$. In this notation, the solution to (4.8) is given by $\boldsymbol{\beta}^k = G^{-1}\boldsymbol{b}$. Thus, Lemma 4.2 implies that the kernel of $[\widehat{W}_\circ^{(k)}]_{[1:k+1,1:k+1]}$ is spanned by $\boldsymbol{v} := [(\boldsymbol{\beta}^k)^\top,\ -1]^\top$. Now, the splitting step attempts to compute a new control $\boldsymbol{\epsilon}^{k+1}$ such that $[\widehat{W}_\circ(\boldsymbol{\epsilon}^{k+1})]_{[1:k+1,1:k+1]}$ is PD on the span of $\boldsymbol{v}$. If this is successful, then $[\widehat{W}_\circ^{(k+1)}]_{[1:k+1,1:k+1]}$ is PD. The equivalence of (4.5) and (4.9) implies that $[\widehat{W}_\circ(\boldsymbol{\epsilon}^{k+1})]_{[1:k+1,1:k+1]}$ is PD on the span of $\boldsymbol{v}$ if and only if $\boldsymbol{\epsilon}^{k+1}$ satisfies $\|C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k),\boldsymbol{\epsilon}^{k+1};T) - C\delta\boldsymbol{y}_\circ(A_{k+1},\boldsymbol{\epsilon}^{k+1};T)\|_2^2 > 0$. The existence of such a control depends on the controllability and observability properties of system (3.7), as shown in sections 5 and 6.

5. Reconstruction of drift matrix in linear systems.

Consider (3.1) with $f(A,\boldsymbol{y},\boldsymbol{\epsilon}) := A\boldsymbol{y} + B\boldsymbol{\epsilon}$, where A and B are real matrices:

$$\dot{\boldsymbol{y}}(t) = A_\star \boldsymbol{y}(t) + B\boldsymbol{\epsilon}(t),\ t \in (0,T], \quad \boldsymbol{y}(0) = 0. \tag{5.1}$$

This is a linear system, where $B \in \mathbb{R}^{N\times M}$ is a given matrix for $N, M \in \mathbb{N}^+$, and $\boldsymbol{\epsilon} \in E_{ad} \subset L^2(0,T;\mathbb{R}^M)$. The drift matrix $A_\star \in \mathbb{R}^{N\times N}$ is unknown and assumed to lie in the space spanned by a set of linearly independent matrices $\mathcal{A} = \{A_1,\dots,A_K\} \subset \mathbb{R}^{N\times N}$, $1 \le K \le N^2$. We write $A_\star = \sum_{j=1}^K \boldsymbol{\alpha}_{\star,j} A_j =: A(\boldsymbol{\alpha}_\star)$. As stated in section 3, we want to identify the unknown drift matrix $A_\star$ by using a set of control functions $(\boldsymbol{\epsilon}^m)_{m=1}^K \subset E_{ad}$ in order to perform K laboratory experiments and obtain the experimental data $(\boldsymbol{\varphi}^\star_{data}(\boldsymbol{\epsilon}^m))_{m=1}^K \subset \mathbb{R}^P$, as defined in (3.2).

Remark 5.1. The hypothesis $\boldsymbol{y}(0) = 0$ in (5.1) can be made without loss of generality. Indeed, if $\boldsymbol{y}(0) = \boldsymbol{y}^0 \neq 0$, one can use $\boldsymbol{\epsilon} = 0$ (case of uncontrolled system), generate the data $\boldsymbol{\varphi}^\star_{data}(0)$, and then subtract this from all other data $(\boldsymbol{\varphi}^\star_{data}(\boldsymbol{\epsilon}^m))_{m=1}^K$ to get back (by linearity) to the case of system (5.1) with $\boldsymbol{y}(0) = 0$.

Using $(\boldsymbol{\epsilon}^m)_{m=1}^K$ and $(\boldsymbol{\varphi}^\star_{data}(\boldsymbol{\epsilon}^m))_{m=1}^K$, the unknown vector $\boldsymbol{\alpha}_\star$ is obtained by solving (3.3), in which $\boldsymbol{y}(A(\boldsymbol{\alpha}),\boldsymbol{\epsilon}^m;T)$ now solves (5.1), with $A_\star$ replaced by $A(\boldsymbol{\alpha})$. Thus, we use the LGR Algorithm 4.1 to generate $(\boldsymbol{\epsilon}^m)_{m=1}^K$ with the goal of making (4.2) uniquely solvable, that means making PD the GN matrix $\widehat{W}_\circ$, defined in (4.6). In (4.2), $\delta\boldsymbol{y}_\circ(A(\delta\boldsymbol{\alpha}),\boldsymbol{\epsilon};t)$ is now the solution to

$$\begin{cases} \delta\dot{\boldsymbol{y}}_\circ(t) = A(\boldsymbol{\alpha}_\circ)\delta\boldsymbol{y}_\circ(t) + \sum_{j=1}^K \delta\boldsymbol{\alpha}_j A_j \boldsymbol{y}_\circ(t), \quad t \in (0,T], \quad \delta\boldsymbol{y}_\circ(0) = 0, \\ \dot{\boldsymbol{y}}_\circ(t) = A(\boldsymbol{\alpha}_\circ)\boldsymbol{y}_\circ(t) + B\boldsymbol{\epsilon}(t), \quad t \in (0,T], \quad \boldsymbol{y}_\circ(0) = 0. \end{cases} \tag{5.2}$$

In what follows, we show that the LGR Algorithm 4.1 does produce $(\boldsymbol{\epsilon}^m)_{m=1}^K$ that make $\widehat{W}_\circ$ PD under appropriate assumptions. To do so, we recall the observability and controllability properties for an input/output system (A, B, C) of the form (2.1) with $A \in \mathbb{R}^{N\times N}$, $B \in \mathbb{R}^{N\times M}$, $C \in \mathbb{R}^{P\times N}$; see, e.g., [31, Thm. 3, Thm. 23].

Definition & Lemma 5.2 (observable input-output linear systems). *The linear system* (2.1) *is said to be observable if the initial state* $\boldsymbol{y}(0) = \boldsymbol{y}^0$ *can be uniquely determined from input/output measurements. Equivalently,* (2.1) *is observable if and only if the observability matrix* $\mathcal{O}_N(C, A) := \begin{bmatrix} C & CA & \cdots & CA^{N-1}\end{bmatrix}^\top$ *has full rank.*

Definition & Lemma 5.3 (controllable input-output linear systems). *The linear system* (2.1) *is said to be controllable if for any final state* $\boldsymbol{y}^f$ *there exists an input sequence that transfers* $\boldsymbol{y}^0$ *to* $\boldsymbol{y}^f$*. Equivalently,* (2.1) *is controllable if and only if the controllability matrix* $\mathcal{C}_N(A, B) := \begin{bmatrix} B & AB & \cdots & A^{N-1}B\end{bmatrix}$ *has full rank.*

Notice that the analysis that we are going to present is also valid in the case of the reconstruction of a control matrix considered in [13, sect. 5], i.e. $f(A, \boldsymbol{y}, \boldsymbol{\epsilon}) = M\boldsymbol{y} + A\boldsymbol{\epsilon}$, and is therefore an extension of the results obtained in [13].

5.1. Analysis for linear systems. We define $\mathcal{O}_N^\circ := \mathcal{O}_N(C, A(\boldsymbol{\alpha}_\circ))$ and $\mathcal{C}_N^\circ := \mathcal{C}_N(A(\boldsymbol{\alpha}_\circ), B)$ and assume that the system $(A(\boldsymbol{\alpha}_\circ), B, C)$ is observable and controllable, namely $\mathcal{R} := \operatorname{rank}(\mathcal{O}_N^\circ) \cdot \operatorname{rank}(\mathcal{C}_N^\circ) = N^2$. In what follows, we show that this is a sufficient condition for $\widehat{W}_\circ$ to be PD with the controls generated by Algorithm 4.1. First, we need the following result [3, Ch. 3, Thm. 2.11].

Lemma 5.4 (controllability of time-invariant systems). *Consider the system* $\dot{\boldsymbol{x}} = A\boldsymbol{x} + B\boldsymbol{\epsilon}$ *with* $\boldsymbol{x}(0) = 0$ *and its solution* $\boldsymbol{x}(\boldsymbol{\epsilon}, t) := \int_0^t e^{(t-s)A(\boldsymbol{\alpha}_\circ)} B\boldsymbol{\epsilon}(s) ds$*. For any finite time* $t_0 > 0$*, there exists a control* $\boldsymbol{\epsilon}$ *that transfers the state to* $\boldsymbol{w}$ *in time* t_0*, i.e.* $\boldsymbol{x}(\boldsymbol{\epsilon}, t_0) = \boldsymbol{w}$*, if and only if* $\boldsymbol{w} \in \operatorname{im}\big(\mathcal{C}_N(A, B)\big)$*. Furthermore, an appropriate* $\boldsymbol{\epsilon}$ *that will accomplish this transfer in time* t_0 *is given by* $\boldsymbol{\epsilon}(t) = B^\top e^{(t_0-t)A^\top}\boldsymbol{\nu}$*, for* $t \in [0, t_0]$ *and* $\boldsymbol{\nu}$ *such that* $\mathcal{W}_c(0, t_0)\boldsymbol{\nu} = \boldsymbol{w}$*, where* $\mathcal{W}_c(0, T) := \int_0^T e^{\tau A} BB^\top e^{\tau A^\top} d\tau$*.*

Now, we prove the following lemma regarding the initialization problem (4.3) and the splitting step problem (4.5). Notice that the proof of this result is inspired by classical Kalman controllability theory; see, e.g., [17].

Lemma 5.5 (LGR initialization and splitting steps (linear systems)). *Assume that the matrices* $A(\boldsymbol{\alpha}_\circ) \in \mathbb{R}^{N\times N}$*,* $B \in \mathbb{R}^{N\times M}$ *and* $C \in \mathbb{R}^{P\times N}$ *are such that* $\operatorname{rank}(\mathcal{O}_N^\circ) = \operatorname{rank}(\mathcal{C}_N^\circ) = N$*, and let* $\widetilde{A} \in \mathbb{R}^{N\times N} \setminus \{0\}$ *be arbitrary. Then any solution* $\widetilde{\boldsymbol{\epsilon}}$ *of the problem* $\max_{\boldsymbol{\epsilon}\in E_{ad}} \|C\delta\boldsymbol{y}_\circ(\widetilde{A}, \boldsymbol{\epsilon}; T)\|_2^2$ *satisfies* $\|C\delta\boldsymbol{y}_\circ(\widetilde{A}, \widetilde{\boldsymbol{\epsilon}}; T)\|_2^2 > 0$*, where* $\dot{\delta\boldsymbol{y}}_\circ = A(\boldsymbol{\alpha}_\circ)\delta\boldsymbol{y}_\circ + \widetilde{A}\boldsymbol{y}^\circ$*, with* $\delta\boldsymbol{y}_\circ(0) = 0$*, and* $\dot{\boldsymbol{y}}_\circ = A(\boldsymbol{\alpha}_\circ)\boldsymbol{y}_\circ + B\boldsymbol{\epsilon}$ *with* $\boldsymbol{y}_\circ(0) = 0$

Proof. It is sufficient to construct an $\widetilde{\boldsymbol{\epsilon}} \in E_{ad}$ such that $C\delta\boldsymbol{y}_\circ(\widetilde{A}, \widetilde{\boldsymbol{\epsilon}}; T) \neq 0$. Since $\widetilde{A} \neq 0$, there exists $\boldsymbol{w} \in \mathbb{R}^N \setminus \{0\}$ such that $\widetilde{A}\boldsymbol{w} \neq 0$. Since $(A(\boldsymbol{\alpha}_\circ), B, C)$ is observable, there exists $\tilde{t} > 0$ such that $Ce^{\tilde{t}A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\boldsymbol{w} \neq 0$. The map $f : \mathbb{R} \to \mathbb{R}^P, t \mapsto Ce^{tA(\boldsymbol{\alpha}_\circ)}\widetilde{A}\boldsymbol{w}$ is analytic with derivatives $f^{(i)}(t) = CA(\boldsymbol{\alpha}_\circ)^i e^{tA(\boldsymbol{\alpha}_\circ)}\widetilde{A}\boldsymbol{w}$. Since $\mathcal{O}_N^\circ$ has full rank and $e^{\tilde{t}A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\boldsymbol{w} \neq 0$, there exists $i \in \{0, \ldots, N\}$ such that $f^{(i)}(\tilde{t}) = CA(\boldsymbol{\alpha}_\circ)^i e^{\tilde{t}A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\boldsymbol{w} \neq 0$. Hence, f is nonconstant, and there exists $t_0 \in (0, T)$ with $Ce^{t_0 A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\boldsymbol{w} \neq 0$. Now, we use that $\boldsymbol{y}_\circ(\boldsymbol{\epsilon}, s) := \int_0^s e^{(s-\tau)A(\boldsymbol{\alpha}_\circ)} B\boldsymbol{\epsilon}(\tau) d\tau$ is the solution at time s of $\dot{\boldsymbol{y}}_\circ = A(\boldsymbol{\alpha}_\circ)\boldsymbol{y}_\circ + B\boldsymbol{\epsilon}$, with $\boldsymbol{y}_\circ(0) = 0$. Since $\mathcal{C}_N^\circ$ has full rank, we have $\boldsymbol{w} \in \operatorname{im}\big(\mathcal{C}_N^\circ\big)$. Thus, Lemma 5.4 guarantees that $\widehat{\boldsymbol{\epsilon}}(t) = B^\top e^{(t_0-t)A(\boldsymbol{\alpha}_\circ)^\top}\boldsymbol{\nu}$, for $t \in [0, t_0]$ and some $\boldsymbol{\nu} \in \mathbb{R}^N$, satisfies $\boldsymbol{y}_\circ(\widehat{\boldsymbol{\epsilon}}, t_0) = \boldsymbol{w}$. Clearly, $\widehat{\boldsymbol{\epsilon}}$ is analytic in $[0, t_0]$ and thereby the same holds for $\boldsymbol{y}_\circ(\widehat{\boldsymbol{\epsilon}}, s)$. Note that, since $\boldsymbol{\epsilon} \equiv 0$ is an interior point of E_{ad}, there exists $\lambda > 0$ such that $\lambda\widehat{\boldsymbol{\epsilon}} \in E_{ad}$ with $Ce^{t_0 A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\boldsymbol{y}_\circ(\lambda\widehat{\boldsymbol{\epsilon}}, t_0) = \lambda Ce^{t_0 A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\boldsymbol{y}_\circ(\widehat{\boldsymbol{\epsilon}}, t_0) \neq 0$. Hence, we can assume without loss of generality that $\widehat{\boldsymbol{\epsilon}} \in E_{ad}$. In conclusion, we obtain that the map

$\boldsymbol{g}:\mathbb{R}\to\mathbb{R}^p$ defined as $s\mapsto Ce^{(T-s)A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\int_0^s e^{(s-\tau)A(\boldsymbol{\alpha}_\circ)}B\widehat{\boldsymbol{\epsilon}}(\tau)d\tau$, is analytic in $(0,t_0)$ with $\boldsymbol{g}(t_0)\neq 0$. Thus, $\boldsymbol{g}$ is nonzero in an open subinterval of $(0,t_0)$. Hence, there exists $t_1\in(0,t_0)$ such that $\int_0^{t_1}\boldsymbol{g}(s)ds\neq 0$. By choosing $\widetilde{\boldsymbol{\epsilon}}(s) := \begin{cases} 0, & 0\le s< T-t_1,\\ \widehat{\boldsymbol{\epsilon}}(s-t_1), & T-t_1\le s\le T,\end{cases}$ and using that $C\delta\boldsymbol{y}_\circ(\widetilde{A},\widetilde{\boldsymbol{\epsilon}};T)=\int_0^T Ce^{(T-s)A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\int_0^s e^{(s-\tau)A(\boldsymbol{\alpha}_\circ)}B\widetilde{\boldsymbol{\epsilon}}(\tau)d\tau ds$, we obtain $C\delta\boldsymbol{y}_\circ(\widetilde{A},\widetilde{\boldsymbol{\epsilon}};T)=\int_0^{t_1}Ce^{(t_1-s)A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\int_0^s e^{(s-\tau)A(\boldsymbol{\alpha}_\circ)}B\widehat{\boldsymbol{\epsilon}}(\tau)d\tau ds=\int_0^{t_1}\boldsymbol{g}(s)ds\neq 0$. ∎

Lemma 5.5 can be applied to both (4.3) and (4.5), choosing $\widetilde{A}=A_1$ and $\widetilde{A}=\big(A^{(k)}(\boldsymbol{\beta}^k)-A_{k+1}\big)$, respectively. Now, we can prove our first main convergence result.

THEOREM 5.6 (positive definiteness of the GN matrix $\widehat{W}_\circ$ (linear systems)). *Assume that $A(\boldsymbol{\alpha}_\circ)\in\mathbb{R}^{N\times N}$, $B\in\mathbb{R}^{N\times M}$ and $C\in\mathbb{R}^{P\times N}$ are such that $\operatorname{rank}(\mathcal{O}_N^\circ)=\operatorname{rank}(\mathcal{C}_N^\circ)=N$. For $K\le N^2$, let $\mathcal{A}=\{A_1,\dots,A_K\}\subset\mathbb{R}^{N\times N}$ be a set of linearly independent matrices such that $A(\boldsymbol{\alpha}_\circ)\in\operatorname{span}(\mathcal{A})$, and let $\{\boldsymbol{\epsilon}^1,\dots,\boldsymbol{\epsilon}^K\}\subset E_{ad}$ be generated by Algorithm 4.1. Then the GN matrix $\widehat{W}_\circ$, defined in (4.6), is PD.*

Proof. We proceed by induction. Lemma 5.5 guarantees that there exists an $\boldsymbol{\epsilon}^1$ such that $[W_\circ(\boldsymbol{\epsilon}^1)]_{1,1}=\|C\delta\boldsymbol{y}_\circ(A_1,\boldsymbol{\epsilon};T)\|_2^2>0$. Now, we assume that $[\widehat{W}_\circ^{(k)}]_{[1:k,1:k]}=\sum_{m=1}^k[W_\circ(\boldsymbol{\epsilon}^m)]_{[1:k,1:k]}$ is PD. By construction, $[\widehat{W}_\circ^{(k+1)}]_{[1:k+1,1:k+1]}$ is PSD. Thus, if $[\widehat{W}_\circ^{(k)}]_{[1:k+1,1:k+1]}$ is PD, then $[\widehat{W}_\circ^{(k+1)}]_{[1:k+1,1:k+1]}=[\widehat{W}_\circ^{(k)}+W_\circ(\boldsymbol{\epsilon}^{k+1})]_{[1:k+1,1:k+1]}$ is PD as well, since $[W_\circ(\boldsymbol{\epsilon}^k)]_{[1:k+1,1:k+1]}$ is PSD. Assume now that the submatrix $[\widehat{W}_\circ^{(k)}]_{[1:k+1,1:k+1]}$ has a nontrivial kernel. Since $[\widehat{W}_\circ^{(k)}]_{[1:k,1:k]}$ is PD (induction hypothesis), problem (4.4) is uniquely solvable with solution $\boldsymbol{\beta}^k$. Then, by Lemma 4.2 the (one-dimensional) kernel of $[\widehat{W}_\circ^{(k)}]_{[1:k+1,1:k+1]}$ is the span of the vector $\boldsymbol{v}=[(\boldsymbol{\beta}^k)^\top,\,-1]^\top$. Using Lemma 5.5 we obtain that the solution $\boldsymbol{\epsilon}^{k+1}$ to the splitting step problem satisfies $\langle\boldsymbol{v},[W_\circ(\boldsymbol{\epsilon}^{k+1})]_{[1:k+1,1:k+1]}\boldsymbol{v}\rangle=\big\|C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k)-A_{k+1},\boldsymbol{\epsilon};T)\big\|_2^2>0$. Thus, $[W(\boldsymbol{\epsilon}^{k+1})]_{[1:k+1,1:k+1]}$ is PD on the span of $\boldsymbol{v}$, and $[\widehat{W}_\circ^{(k+1)}]_{[1:k+1,1:k+1]}$ is PD. ∎

Having proved that Algorithm 4.1 makes $\widehat{W}_\circ$ PD, the obvious question is whether this is sufficient for the convergence of GN (Lemma 3.1). We answer in section 5.2.

5.2. Positive definiteness of the GN matrix. To guarantee convergence of GN, we need to show that $\widehat{W}(\boldsymbol{\alpha}):=\sum_{m=1}^K R_m'(\boldsymbol{\alpha})^\top R_m'(\boldsymbol{\alpha})$ (defined in section 3) remains PD in a neighborhood of $\boldsymbol{\alpha}_\star$. In section 5.1, we proved that the controls generated by Algorithm 4.1 make $\widehat{W}_\circ=\widehat{W}(\boldsymbol{\alpha}_\circ)$ PD. Thus, it is sufficient to prove that $\widehat{W}(\boldsymbol{\alpha})$ is PD in a neighborhood of $\boldsymbol{\alpha}_\circ$ containing $\boldsymbol{\alpha}_\star$. To do so, we write $\widehat{W}(\boldsymbol{\alpha})$ as

$$[\widehat{W}(\boldsymbol{\alpha})]_{i,j}:=\sum_{m=1}^K\langle\boldsymbol{\gamma}_i(\boldsymbol{\alpha},\boldsymbol{\epsilon}^m),\boldsymbol{\gamma}_j(\boldsymbol{\alpha},\boldsymbol{\epsilon}^m)\rangle,\quad i,j\in\{1,\dots,K\},\tag{5.3}$$

$$\boldsymbol{\gamma}_j(\boldsymbol{\alpha},\boldsymbol{\epsilon}^m):=\int_0^T Ce^{(T-s)A(\boldsymbol{\alpha})}A_j\boldsymbol{y}(A(\boldsymbol{\alpha}),\boldsymbol{\epsilon}^m;s)ds,\quad j\in\{1,\dots,K\},\tag{5.4}$$

and recall the next lemma, which follows from the Bauer-Fike theorem [6].

LEMMA 5.7 (rank stability). *Consider $N_D,M_D\in\mathbb{N}^+$ with $N_D\ge M_D$, and a matrix $D\in\mathbb{R}^{N_D\times M_D}$ with rank $\mathcal{R}_\mathcal{D}$ and (positive) singular values $\sigma_1,\dots,\sigma_{\mathcal{R}_\mathcal{D}}$ in descending order. Then $\min_{\widehat{D}\in\mathbb{R}^{N_D\times M_D}}\{\|\widehat{D}\|_2\mid\operatorname{rank}(D+\widehat{D})<\mathcal{R}_\mathcal{D}\}=\sigma_{\mathcal{R}_\mathcal{D}}$.*

Using this lemma, we can prove the following approximation result.

LEMMA 5.8 (positive definiteness of $\widehat{W}(\boldsymbol{\alpha})$ (linear systems)). *Let $\widehat{W}_\circ$ defined in (4.6) be PD and let $\sigma_K^\circ > 0$ be its smallest singular value. Then, there exists $\delta := \delta(\sigma_K^\circ) > 0$ such that $\widehat{W}(\boldsymbol{\alpha})$ (in (5.3)) is PD for any $\boldsymbol{\alpha} \in \mathbb{R}^K$ with $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}_\circ\|_2 < \delta$.*

Proof. Our first goal is to show that $\widehat{W}(\boldsymbol{\alpha})$ is continuous in $\boldsymbol{\alpha}$. From (5.3) and (5.4) we know that $\widehat{W}(\boldsymbol{\alpha})$ is the sum over products of $\int_0^T Ce^{(T-s)A(\boldsymbol{\alpha})} A_j \boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; s) ds$, where $\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; s) = \int_0^s e^{(s-\tau)A(\boldsymbol{\alpha})} B\boldsymbol{\epsilon}^m(\tau) d\tau$. Now, recall that $A(\boldsymbol{\alpha}) = \sum_{j=1}^K \boldsymbol{\alpha}_j A_j$, meaning that $A(\boldsymbol{\alpha})$ is continuous in $\boldsymbol{\alpha}$. Since the exponential map $\mathbb{R}^N \to \mathbb{R}^{N\times N}, \boldsymbol{\alpha} \mapsto e^{sA(\boldsymbol{\alpha})}$ and the integral map $\mathbb{R}^{N\times N} \to \mathbb{R}^N, X \mapsto \int_0^s XB\boldsymbol{\epsilon}(\tau) d\tau$ are continuous, we obtain that $\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; s)$ is continuous in $\boldsymbol{\alpha}$. Since products of continuous functions are continuous, we obtain that $\widehat{W}(\boldsymbol{\alpha})$ is continuous in $\boldsymbol{\alpha}$.

By assumption, $\widehat{W}_\circ$ is PD, and therefore $\sigma_K^\circ > 0$. Since $\widehat{W}(\boldsymbol{\alpha})$ is continuous in $\boldsymbol{\alpha}$, we obtain that there exists a $\delta := \delta(\sigma_K^\circ) > 0$ such that for any $\boldsymbol{\alpha}$ with $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}_\circ\|_2 < \delta(\sigma_K^\circ)$ it holds that $\left\|\widehat{W}(\boldsymbol{\alpha}) - \widehat{W}(\boldsymbol{\alpha}_\circ)\right\|_2 < \sigma_K^\circ$. Now, let $\widehat{\boldsymbol{\alpha}}$ be such that $\|\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_\circ\|_2 < \delta(\sigma_K^\circ)$ and hence $\left\|\widehat{W}(\widehat{\boldsymbol{\alpha}}) - \widehat{W}(\boldsymbol{\alpha}_\circ)\right\|_2 < \sigma_K^\circ$. Setting $D = \widehat{W}(\boldsymbol{\alpha}_\circ)$ and $\widehat{D} = \widehat{W}(\widehat{\boldsymbol{\alpha}}) - \widehat{W}(\boldsymbol{\alpha}_\circ)$, Lemma 5.7 implies that $K = \operatorname{rank}(\widehat{W}(\boldsymbol{\alpha}_\circ)) \leq \operatorname{rank}(\widehat{W}(\widehat{\boldsymbol{\alpha}}))$. Because of (5.3), $\widehat{W}(\widehat{\boldsymbol{\alpha}}) \in \mathbb{R}^{K\times K}$ meaning that $\operatorname{rank}(\widehat{W}(\widehat{\boldsymbol{\alpha}})) = K$. Since $\widehat{W}(\boldsymbol{\alpha})$ is PSD by construction, $\operatorname{rank}(\widehat{W}(\widehat{\boldsymbol{\alpha}})) = K$ implies that $\widehat{W}(\widehat{\boldsymbol{\alpha}})$ is PD. ∎

Lemma 5.8 implies that the positive definiteness of $\widehat{W}(\boldsymbol{\alpha})$ is locally preserved near $\boldsymbol{\alpha}_\circ$. Now, we can prove our main convergence result.

THEOREM 5.9 (convergence of GN (linear systems)). *Let $\boldsymbol{\alpha}_\circ \in \mathbb{R}^K$ be such that $A(\boldsymbol{\alpha}_\circ) \in \mathbb{R}^{N\times N}$, $B \in \mathbb{R}^{N\times M}$ and $C \in \mathbb{R}^{P\times N}$ satisfy $\operatorname{rank}(\mathcal{O}_N^\circ) \cdot \operatorname{rank}(\mathcal{C}_N^\circ) = N^2$. Let $(\boldsymbol{\epsilon}^m)_{m=1}^K \subset E_{ad}$ be a set of controls generated by Algorithm 4.1. Finally, let $\widehat{\sigma}_K$ be the K-th (smallest) singular value of $\widehat{W}_\circ$ defined in (4.6). Then there exists $\delta = \delta(\widehat{\sigma}_K) > 0$ such that if $\boldsymbol{\alpha}_\star \in \mathbb{R}^K$ satisfies $\|\boldsymbol{\alpha}_\star - \boldsymbol{\alpha}_\circ\| < \delta$, then GN method for the problem*

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^K} \frac{1}{2} \sum_{m=1}^K \|C\boldsymbol{y}(A(\boldsymbol{\alpha}_\star), \boldsymbol{\epsilon}^m; T) - C\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; T)\|_2^2, \tag{5.5}$$

initialized with $\boldsymbol{\alpha}_\circ$, converges to $\boldsymbol{\alpha}_\star$.

Proof. Theorem 5.6 guarantees that $\widehat{W}_\circ$ is PD. Thus, $\widehat{\sigma}_K > 0$ and by Lemma 5.8 there exists $\delta = \delta(\widehat{\sigma}_K) > 0$ such that $\widehat{W}(\boldsymbol{\alpha})$ is PD for $\boldsymbol{\alpha} \in \mathbb{R}^K$ with $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}_\circ\|_2 < \delta$. Moreover, we know from section 3 that $\widehat{W}(\boldsymbol{\alpha}_c)$ is the GN matrix for the iterate $\boldsymbol{\alpha}_c \in \mathbb{R}^K$ of GN for (3.3). As in the proof of Lemma 5.8, one can also show that the functions $R_m(\boldsymbol{\alpha})$, defined in (3.4), are Lipschitz continuously differentiable in $\boldsymbol{\alpha}$ for all $m \in \{1, \dots, K\}$. Hence, if $\|\boldsymbol{\alpha}_\star - \boldsymbol{\alpha}_\circ\| < \delta$, then the result follows by Lemma 3.1. ∎

5.3. Local uniqueness of solutions. Theorem 5.9 says that GN converges to $\boldsymbol{\alpha}_\star$ if an appropriate initialization vector $\boldsymbol{\alpha}_\circ$ is used. However, in the linear case corresponding to (5.1) we can specify the local properties of problem (3.3) around the solution $\boldsymbol{\alpha}_\star$. To this end, we start by rewriting the cost function in a matrix form.

LEMMA 5.10 (online identification problem in matrix form (linear systems)). *Problem (3.3) is equivalent to*

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^K} \frac{1}{2} \langle \boldsymbol{\alpha}_\star - \boldsymbol{\alpha}, \widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha})(\boldsymbol{\alpha}_\star - \boldsymbol{\alpha}) \rangle, \tag{5.6}$$

where $\widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}) \in \mathbb{R}^{K\times K}$ *is defined as*[2] $\widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}) := \sum_{m=1}^K W(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, \boldsymbol{\epsilon}^m)$*, with*

$$[W(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, \boldsymbol{\epsilon}^m)]_{i,j} := \langle \boldsymbol{\gamma}_i(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, \boldsymbol{\epsilon}^m), \boldsymbol{\gamma}_j(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, \boldsymbol{\epsilon}^m)\rangle, \quad i,j \in \{1,\dots,K\}, \tag{5.7}$$

$$\boldsymbol{\gamma}_j(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, \boldsymbol{\epsilon}^m) := \int_0^T Ce^{(T-s)A(\boldsymbol{\alpha}_\star)} A_j \boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; s) ds, \quad j \in \{1,\dots,K\}. \tag{5.8}$$

Proof. Let $J(\boldsymbol{\alpha}) := \frac{1}{2}\sum_{m=1}^K \|C\boldsymbol{y}(A_\star, \boldsymbol{\epsilon}^m; T) - C\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; T)\|_2^2$. For $t \in [0,T]$ and $m \in \{1,\dots,K\}$ define $\Delta \boldsymbol{y}_m(t) := \boldsymbol{y}(A_\star, \boldsymbol{\epsilon}^m; t) - \boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; t)$. Then we have

$$\begin{aligned}\dot{\Delta \boldsymbol{y}}_m(t) &= A(\boldsymbol{\alpha}_\star)\boldsymbol{y}(A_\star, \boldsymbol{\epsilon}^m; t) + B\boldsymbol{\epsilon}^m(t) - A(\boldsymbol{\alpha})\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; t) - B\boldsymbol{\epsilon}^m(t)\\ &= A(\boldsymbol{\alpha}_\star)\Delta \boldsymbol{y}_m(t) + A(\boldsymbol{\alpha}_\star - \boldsymbol{\alpha})\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; t),\end{aligned}$$

whose solution at $t = T$ is $\Delta \boldsymbol{y}_m(T) = \int_0^T e^{(T-s)A(\boldsymbol{\alpha}_\star)}\left[A(\boldsymbol{\alpha}_\star - \boldsymbol{\alpha})\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; s)\right] ds$. Thus, recalling $A(\boldsymbol{\alpha}) = \sum_{j=1}^K \boldsymbol{\alpha}_j A_j$, a direct calculation shows that $J(\boldsymbol{\alpha})$ can be written as $J(\boldsymbol{\alpha}) = \frac{1}{2}\langle \boldsymbol{\alpha}_\star - \boldsymbol{\alpha}, \widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha})(\boldsymbol{\alpha}_\star - \boldsymbol{\alpha})\rangle$. ◻

Now, the set of global solutions to (5.6) is $\mathcal{S}_{global} := \{\boldsymbol{\alpha} \in \mathbb{R}^K : (\boldsymbol{\alpha}_\star - \boldsymbol{\alpha}) \in \ker \widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha})\}$. Since $\widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha})$ is symmetric PSD, (5.6) is locally uniquely solvable if and only if $\widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha})$ is PD for $\boldsymbol{\alpha}$ close to $\boldsymbol{\alpha}_\star$. Now, assume that the system is fully observable and controllable, meaning that $\mathcal{R} = N^2$. Theorem 5.9 guarantees that Algorithm 4.1 computes $(\boldsymbol{\epsilon}_m)_{m=1}^{N^2}$ such that $\widehat{W}(\boldsymbol{\alpha}_\star) = \widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}_\star)$ is PD, if $\boldsymbol{\alpha}_\star$ is close enough to $\boldsymbol{\alpha}_\circ$. Similar to the proof of Lemma 5.8, one can prove that $\widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha})$ is continuous in $\boldsymbol{\alpha}$. Hence, if $\widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}_\star)$ is PD, then the same holds for $\widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha})$, when $\boldsymbol{\alpha}$ is close to $\boldsymbol{\alpha}_\star$, which implies that (5.6) is locally uniquely solvable with $\boldsymbol{\alpha} = \boldsymbol{\alpha}_\star$.

6. Bilinear reconstruction problems.

In this section, we extend the results of section 5 to the case of skew-symmetric bilinear systems. We consider (3.1) with a right-hand side $f(A, \boldsymbol{y}, \epsilon) = (A + \epsilon B)\boldsymbol{y}$, that is

$$\dot{\boldsymbol{y}}(t) = (A_\star + \epsilon(t)B)\boldsymbol{y}(t),\ t \in (0,T], \quad \boldsymbol{y}(0) = \boldsymbol{y}^0, \tag{6.1}$$

where $B \in \mathfrak{so}(N)$ is a given skew-symmetric matrix for $N \in \mathbb{N}^+$, the initial state is $\boldsymbol{y}^0 \in \mathbb{R}^N$, and $\epsilon \in E_{ad} \subset L^2(0,T;\mathbb{R})$ (with $M = 1$; see section 2) is a control function. The matrix $A_\star \in \mathfrak{so}(N)$ is unknown and assumed to lie in the space spanned by a set of linearly independent matrices $\mathcal{A} = \{A_1,\dots,A_K\} \subset \mathbb{R}^{N\times N}$, $1 \le K \le N^2$, and we write $A_\star = \sum_{j=1}^K \boldsymbol{\alpha}_{\star,j} A_j =: A(\boldsymbol{\alpha}_\star)$. Since the matrices $A_\star$ and B are skew-symmetric, system (6.1) is norm preserving, i.e. $\|\boldsymbol{y}(t)\|_2 = \|\boldsymbol{y}^0\|_2$ for all $t \in [0,T]$.[3]

To identify the true matrix $A_\star$, one can consider a set of control functions $(\epsilon^m)_{m=1}^K \subset E_{ad}$ and use it experimentally to obtain the data $(\boldsymbol{\varphi}^\star_{data}(\epsilon^m))_{m=1}^K \subset \mathbb{R}^P$, as defined in (3.2). The unknown vector $\boldsymbol{\alpha}_\star$ is then obtained by solving the problem

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^K} \frac{1}{2}\sum_{m=1}^K \|\boldsymbol{\varphi}^\star_{data}(\epsilon^m) - C\boldsymbol{y}(A(\boldsymbol{\alpha}), \epsilon^m; T)\|_2^2 . \tag{6.2}$$

Assume to have an estimate $\boldsymbol{\alpha}_\circ$ of $\boldsymbol{\alpha}_\star$. We can derive the linearized equation in $\boldsymbol{\alpha}_\circ$:

$$\begin{cases}\delta\dot{\boldsymbol{y}}_\circ(t) = (A_\circ + \epsilon(t)B)\delta\boldsymbol{y}_\circ(t) + \sum_{j=1}^K \delta\boldsymbol{\alpha}_j A_j \boldsymbol{y}_\circ(t), \quad t \in (0,T], \quad \delta\boldsymbol{y}_\circ(0) = 0,\\ \dot{\boldsymbol{y}}_\circ(t) = (A_\circ + \epsilon(t)B)\boldsymbol{y}_\circ(t), \quad t \in (0,T], \quad \boldsymbol{y}_\circ(0) = \boldsymbol{y}^0,\end{cases} \tag{6.3}$$

[2] Notice that this notation is related to (5.3) in the sense that $\widetilde{W}(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = \widehat{W}(\boldsymbol{\alpha})$.

[3] To see this, we observe that $\frac{1}{2}\frac{d}{dt}\|\boldsymbol{y}(t)\|_2^2 = \langle \boldsymbol{y}(t), \dot{\boldsymbol{y}}(t)\rangle = \langle \boldsymbol{y}(t), (A_\star + \epsilon(t)B)\boldsymbol{y}(t)\rangle = 0$.

where $A_\circ := A(\boldsymbol{\alpha}_\circ)$. Denoting by $\delta\boldsymbol{y}_\circ(A(\delta\boldsymbol{\alpha}), \epsilon; t)$ the solution of (6.3) at time $t \in [0, T]$, the GN matrix $\widehat{W}_\circ$ is defined as in (4.6), and LGR is detailed in Algorithm 4.1.

Let us recall the following definition and result from [11, Corollary 4.11].

Definition & Lemma 6.1 (Controllability of skew-symmetric bilinear systems). *Consider a system of the form*

$$\dot{\boldsymbol{y}}(t) = (A_\circ + \epsilon(t)B)\boldsymbol{y}(t), \quad \boldsymbol{y}(0) = \boldsymbol{y}^0, \tag{6.4}$$

where $A_\circ, B \in \mathfrak{so}(N)$. System (6.4) is said to be controllable if for any final state $\boldsymbol{y}^f$ that lies on the sphere of radius $\|\boldsymbol{y}^0\|_2$ there exists a control $\epsilon(t)$ that transfers $\boldsymbol{y}^0$ to $\boldsymbol{y}^f$. Furthermore, if the Lie algebra $L = Lie\{A_\circ, B\} \subset \mathfrak{so}(N)$, generated by the matrices $A_\circ$ and B, has dimension $\frac{N(N-1)}{2}$, then there exists a constant $\tilde{t} \geq 0$ such that for any $T \geq \tilde{t}$ controllability of (6.4) holds.

As in section 5, we also need to make some assumptions on the observability of the linearized equation in (6.3). However, recalling the proof of Lemma 5.5, these assumptions are only required to prove the existence of a control function that guarantees a positive cost function value in the splitting step. If we assume this function to be constant, at least on a subinterval of $[0, T]$, then we get a system of the form

$$\dot{\delta\boldsymbol{y}}_\circ(t) = (A_\circ + cB)\delta\boldsymbol{y}_\circ(t) + A(\delta\boldsymbol{\alpha})\boldsymbol{y}_\circ(t), \tag{6.5}$$

for a scalar $c \in \mathbb{R}$. In this case, system (6.5) is again a linear system, for which observability is defined in Definition 5.2. Hence, the observability matrix is $\mathcal{O}_N(C, A_\circ + cB)$. Let us state our assumptions on controllability and observability of (6.4) and (6.5).

Assumption 6.2. *Let the matrices $A_\circ$, B and C be such that*

1. *The Lie algebra $Lie\{A_\circ, B\} \subset \mathfrak{so}(N)$ has dimension $\frac{N(N-1)}{2}$ and $T > 0$ is large enough, such that the controllability result from Lemma 6.1 holds.*
2. *There exists $c \in \mathbb{R}$ such that $\mathcal{O}_N(C, A_\circ + cB)$ has full rank.*

In addition, let $E_{ad} \subset L^2(0, T; \mathbb{R})$ be chosen such that the controllability result from Lemma 6.1 holds, and such that $\epsilon \equiv c$ is an interior point of E_{ad} (for the above $c \in \mathbb{R}$).

6.1. Analysis for skew-symmetric bilinear systems. We show in this section that Assumption 6.2 is a sufficient condition for the GN matrix $\widehat{W}_\circ$, defined as in (4.6), to be PD if the controls generated by Algorithm 4.1 are used. The idea of the analysis is similar to the one considered in section 5, meaning that we first have to show the existence of a control that makes the cost function of (4.5) strictly positive.

Lemma 6.3 (GR initialization and splitting steps (bilinear systems)). *Let the matrices $A_\circ$, B and C satisfy Assumption 6.2. Let $\widetilde{A} \in \operatorname{span}(\mathcal{A})$ be an arbitrary matrix. If $T > 0$ is sufficiently large, then $\left\|C\delta\boldsymbol{y}_\circ(\widetilde{A}, \widetilde{\epsilon}; T)\right\|_2^2 > 0$ for any solution $\widetilde{\epsilon}$ to*

$$\max_{\epsilon \in E_{ad}} \left\|C\delta\boldsymbol{y}_\circ(\widetilde{A}, \epsilon; T)\right\|_2^2, \text{ s.t. } \dot{\delta\boldsymbol{y}}_\circ(t) = (A_\circ + \epsilon(t)B)\delta\boldsymbol{y}_\circ(t) + \widetilde{A}\boldsymbol{y}_\circ(t), \quad \delta\boldsymbol{y}_\circ(0) = 0,$$
$$\dot{\boldsymbol{y}}_\circ(t) = (A_\circ + \epsilon(t)B)\boldsymbol{y}_\circ(t), \quad \boldsymbol{y}_\circ(0) = \boldsymbol{y}^0.$$

Proof. It is sufficient to construct an $\widehat{\epsilon}_c \in E_{ad}$ such that $C\delta\boldsymbol{y}_\circ(\widetilde{A}, \widehat{\epsilon}_c; T) \neq 0$ for T sufficiently large. Let us define $\widehat{\epsilon}_c(s) := \begin{cases} \widehat{\epsilon}(s), & \text{for } 0 \leq s \leq \widehat{t}, \\ c, & \text{for } \widehat{t} < s \leq T, \end{cases}$ where $c \in \mathbb{R}$, $\widehat{\epsilon} \in E_{ad}$, $T > 0$ and $\widehat{t} \in (0, T)$ are to be chosen. Since $\widetilde{A} \neq 0$, there exists $\boldsymbol{v} \in$

$\mathbb{R}^N, \|\boldsymbol{v}\|_2 = \|\boldsymbol{y}^0\|_2$ such that $\widetilde{A}\boldsymbol{v} \neq 0$. By the first part of Assumption 6.2, we know that (6.4) is controllable on the sphere of radius $\|\boldsymbol{y}^0\|_2$, meaning that there exist $\widehat{t} > 0$ and $\widehat{\epsilon} \in E_{ad}$ such that $\boldsymbol{y}_\circ(\widetilde{A}, \widehat{\epsilon}; \widehat{t}) = \boldsymbol{v}$. Defining $A_c := A_\circ + cB$, we notice that $f_{\boldsymbol{v}}(t) := \widetilde{A}e^{tA_c}\boldsymbol{v}$ is analytic in t, and since $f_{\boldsymbol{v}}(0) = \widetilde{A}\boldsymbol{v} \neq 0$, it is not equal to zero everywhere and therefore has only isolated roots, see, e.g., [30, Thm. 10.18]. Recalling that exponential matrices are always invertible (see, e.g., [25, Thm. 2.6.38]), we obtain that there exists $t_1 > 0$ such that $e^{-t_1(A_c)}\widetilde{A}e^{(t_1-\widehat{t})A_c}\boldsymbol{v} \neq 0$. By defining $\boldsymbol{w} := \delta\boldsymbol{y}_\circ(\widetilde{A}, \widehat{\epsilon}; \widehat{t})$ and $\boldsymbol{g}(t) := \int_{\widehat{t}}^{t} e^{-s(A_c)}\widetilde{A}e^{(s-\widehat{t})A_c}\boldsymbol{v}ds + e^{-\widehat{t}A_c}\boldsymbol{w}$, we observe that $\frac{d\boldsymbol{g}(t_1)}{dt} = e^{-t_1(A_c)}\widetilde{A}e^{(t_1-\widehat{t})A_c}\boldsymbol{v} \neq 0$. Since $\frac{d\boldsymbol{g}(t)}{dt}$ is analytic in t, the same holds for $\boldsymbol{g}(t)$,[4] and since $\frac{d\boldsymbol{g}(t_1)}{dt} \neq 0$ we obtain that $\boldsymbol{g}(t)$ has only isolated roots. Notice that $e^{-tA_c}\delta\boldsymbol{y}(\widetilde{A}, \widehat{\epsilon}_c; t) = e^{-tA_c}\int_{\widehat{t}}^{t} e^{(t-s)(A_c)}\widetilde{A}e^{(s-\widehat{t})A_c}\boldsymbol{v}ds + e^{(t-\widehat{t})A_c}\boldsymbol{w} = \boldsymbol{g}(t)$, for $t > \widehat{t}$. Thus, it remains to show that there exists $T > \widehat{t}$ such that $Ce^{TA_c}\boldsymbol{g}(T) \neq 0$. Assumption 6.2 guarantees that there exists $c \in \mathbb{R}$ such that the observability matrix $\mathcal{O}_N(C, A_\circ + cB)$ has full rank. Hence, for any $\boldsymbol{u} \in \mathbb{R}^N \setminus \{0\}$ there exists a $t_{\boldsymbol{u}} > \widehat{t}$ such that $Ce^{t_{\boldsymbol{u}}A_c}\boldsymbol{u} \neq 0$. Since $t \mapsto Ce^{tA_c}\boldsymbol{u}$ is analytic in t, $Ce^{t_{\boldsymbol{u}}A_c}\boldsymbol{u} \neq 0$ implies that it has only isolated roots. Thus, for $t > \widehat{t}$, $t \mapsto Ce^{tA_c}\boldsymbol{g}(t)$ is the composition of two analytic functions which both have only isolated roots, and is therefore also analytic with isolated roots. Hence, there exists $T > \widehat{t}$ such that $C\delta\boldsymbol{y}(\widetilde{A}, \widehat{\epsilon}_c; T) = Ce^{TA_c}\boldsymbol{g}(T) \neq 0$. ∎

Now, we can prove our main result, whose proof is the same as the one of Theorem 5.6, in which Lemma 6.3 has to be used instead of Lemma 5.5.

THEOREM 6.4 (positive definiteness of the GN matrix $\widehat{W}_\circ$ (bilinear systems)). *Let $\boldsymbol{\alpha}_\circ \in \mathbb{R}^K$ be such that the matrices $A(\boldsymbol{\alpha}_\circ), B \in \mathfrak{so}(N)$ and $C \in \mathbb{R}^{P\times N}$ satisfy Assumption 6.2. For $K \leq N^2$, let $\mathcal{A} = \{A_1, \ldots, A_K\} \subset \mathfrak{so}(N)$ be a set of linearly independent matrices such that $A_\star \in \operatorname{span}\mathcal{A}$, and let $\{\epsilon^1, \ldots, \epsilon^K\} \subset E_{ad}$ be controls generated by Algorithm 4.1. Then the GN matrix $\widehat{W}_\circ$, defined in (4.6), is PD.*

6.2. Positive definiteness of the GN matrix. As in section 5.2, we show that if the GN matrix in $\boldsymbol{\alpha}_\circ$ is PD, then the same is true locally, for all iterates $\boldsymbol{\alpha}_c$ of GN. We start by writing the matrix $\widehat{W}(\boldsymbol{\alpha})$ as a function of $\boldsymbol{\alpha}$:

$$[\widehat{W}(\boldsymbol{\alpha})]_{i,j} := \sum_{m=1}^{K} \langle C\delta\boldsymbol{y}(\boldsymbol{\alpha}, A_i, \epsilon^m; T), C\delta\boldsymbol{y}(\boldsymbol{\alpha}, A_j, \epsilon^m; T)\rangle, \quad i, j \in \{1, \ldots, K\}, \tag{6.6}$$

where $\delta\boldsymbol{y}(\boldsymbol{\alpha}, \widehat{A}, \epsilon; T)$ denotes the solution at time T of

$$\begin{cases} \dot{\delta\boldsymbol{y}}(t) &= (A(\boldsymbol{\alpha}) + \epsilon(t)B)\delta\boldsymbol{y}(t) + \widehat{A}\boldsymbol{y}(t), \quad \delta\boldsymbol{y}(0) = 0, \\ \dot{\boldsymbol{y}}(t) &= (A(\boldsymbol{\alpha}) + \epsilon(t)B)\boldsymbol{y}(t), \quad \boldsymbol{y}(0) = \boldsymbol{y}^0. \end{cases} \tag{6.7}$$

Now, we want to prove the same positive definiteness result as in Lemma 5.8.

LEMMA 6.5 (positive definiteness of $\widehat{W}_\circ$ (bilinear systems)). *Let $\widehat{W}_\circ$, defined in (4.6), be PD and denote by $\sigma_K^\circ > 0$ the smallest singular value of $\widehat{W}_\circ$. Then, there exists $\delta := \delta(\sigma_K^\circ) > 0$ such that for any $\boldsymbol{\alpha} \in \mathbb{R}^K$ with $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}_\circ\|_2 < \delta$, the matrix $\widehat{W}(\boldsymbol{\alpha})$, defined as in (6.6), is also PD.*

Proof. Recalling the proof of Lemma 5.8, it is sufficient to show that the solution $\delta\boldsymbol{y}(\boldsymbol{\alpha}, \widehat{A}, \epsilon; T)$ of (6.7) is continuous in $\boldsymbol{\alpha}$. By [11, Proposition 3.26],[5] we obtain continuity of the map $\boldsymbol{\alpha} \mapsto \boldsymbol{y}(A(\boldsymbol{\alpha}), \epsilon; T)$ and analogously the continuity of $\boldsymbol{\alpha} \mapsto \delta\boldsymbol{y}(\boldsymbol{\alpha}, \widehat{A}, \epsilon; T)$. ∎

[4] This follows directly from the fundamental theorem of calculus.
[5] This result is a special case of the implicit function theorem; see, e.g., [11, Thm. 3.4].

Using the result from Lemma 6.5, we can directly prove our main result.

THEOREM 6.6 (convergence of GN (bilinear systems)). *Let $\boldsymbol{\alpha}_\circ \in \mathbb{R}^K$ be such that the matrices $A(\boldsymbol{\alpha}_\circ)$, B and C satisfy Assumption 6.2, and let $(\epsilon^m)_{m=1}^K \subset E_{ad}$ be generated by Algorithm 4.1. Denote by $\widehat{\sigma}_K$ the smallest singular value of $\widehat{W}_\circ$, defined in (4.6). Then there exists $\delta = \delta(\widehat{\sigma}_K) > 0$ such that, if $\boldsymbol{\alpha}_\star \in \mathbb{R}^K$ satisfies $\|\boldsymbol{\alpha}_\star - \boldsymbol{\alpha}_\circ\| \leq \delta$, then GN for the solution (6.2), initialized with $\boldsymbol{\alpha}_\circ$, converges to $\boldsymbol{\alpha}_\star$.*

Proof. Theorem 6.4 implies that $\widehat{W}_\circ$ is PD, namely $\widehat{\sigma}_K > 0$. As in the proof of Lemma 6.5, one can show that $\boldsymbol{\alpha} \mapsto R_m(\boldsymbol{\alpha})$, defined in (3.4), are Lipschitz continuously differentiable in $\boldsymbol{\alpha}$ for all $m \in \{1, \dots, K\}$. Thus, the result follows by Lemma 6.5. ◻

6.3. Local uniqueness of solutions. Let us study the local properties of problem (6.2) around $\boldsymbol{\alpha}_\star$. We use the same approach as in the linear case, and start by rewriting problem (6.2) in a matrix-vector form.

LEMMA 6.7 (online identification problem in matrix form (bilinear systems)). *Problem (3.3) is equivalent to $\min_{\boldsymbol{\alpha}\in\mathbb{R}^K} \frac{1}{2}\langle \boldsymbol{\alpha}_\star - \boldsymbol{\alpha}, \widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha})(\boldsymbol{\alpha}_\star - \boldsymbol{\alpha})\rangle$, where we define $\widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}) := \sum_{m=1}^K W(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, \epsilon^m) \in \mathbb{R}^{K\times K}$, $[W(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, \epsilon^m)]_{i,j} := \langle \boldsymbol{\gamma}_j, \boldsymbol{\gamma}_i\rangle$, $\boldsymbol{\gamma}_j := C\delta\boldsymbol{y}_m(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, A_j; T)$, for $i, j \in \{1, \dots, K\}$, and $\delta\boldsymbol{y}_m(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, A; t)$ solves*

$$\begin{cases} \dot{\delta\boldsymbol{y}}(t) &= (A(\boldsymbol{\alpha}_\star) + \epsilon^m(t)B)\delta\boldsymbol{y}(t) + A\boldsymbol{y}(t), \quad \delta\boldsymbol{y}(0) = 0, \\ \dot{\boldsymbol{y}}(t) &= (A(\boldsymbol{\alpha}) + \epsilon^m(t)B)\boldsymbol{y}(t), \quad \boldsymbol{y}(0) = \boldsymbol{y}^0. \end{cases}$$

The proof of Lemma 6.7 is analogous to the one of Lemma 5.10 and we omit it here for brevity. Notice that the notations in (6.6) and Lemma 6.7 are related in the sense that $\widehat{W}(\boldsymbol{\alpha}) = \widetilde{W}(\boldsymbol{\alpha}, \boldsymbol{\alpha})$. Now, proceeding as in section 5.3 and defining the set of all global solutions $\mathcal{S}_{global} := \{\boldsymbol{\alpha} \in \mathbb{R}^K : (\boldsymbol{\alpha}_\star - \boldsymbol{\alpha}) \in \ker \widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha})\}$, we obtain the same local uniqueness of the solution $\boldsymbol{\alpha}_\star$ to (6.2), meaning that if $\widehat{W}(\boldsymbol{\alpha}_\star) = \widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}_\star)$ is PD, the same holds for $\widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha})$ when $\boldsymbol{\alpha}$ is close to $\boldsymbol{\alpha}_\star$.

7. Towards general nonlinear GR algorithms. The LGR algorithm introduced in the previous sections only considers the linearized system. Thus it does not have access to the full (nonlinear) dynamics and can only capture the local characteristics of the considered system. Moreover, as we will show in section 8, the standard GR algorithm can outperform LGR when $\boldsymbol{\alpha}_\circ$ is far from the solution. However, the analysis of LGR allows us to better understand the local behavior of GR and prove that locally it is capable to construct control functions that guarantee convergence of GN. This analysis is carried out in section 7.1. This is the first analysis of GR algorithms for nonlinear problems. While section 7.1 focuses on GR, we also briefly discuss its optimized version called optimized GR (OGR), introduced in [13], and propose a slight improvement of the original version.

7.1. A local analysis for nonlinear GR algorithms. This section is concerned with general nonlinear systems of the form $\dot{\boldsymbol{y}}(t) = f(A(\boldsymbol{\alpha}^0) + A(\delta\boldsymbol{\alpha}_\star), \boldsymbol{y}(t), \boldsymbol{\epsilon}(t))$ with the goal of reconstructing $A(\delta\boldsymbol{\alpha}_\star) = A_\star - A(\boldsymbol{\alpha}^0)$. Here, the shift of $A_\star$ is considered to perform a local analysis near $A(\boldsymbol{\alpha}^0)$. The goal is to prove convergence of GN for the controls generated by the GR Algorithm 7.1 using a local analogy to Algorithm 4.1. Notice that there are a few differences between Algorithms 7.1 and 4.1. To derive a local analogy between them, all operators from the set $\mathcal{A}$ are shifted by $A(\boldsymbol{\alpha}_\circ)$. Additionally, the fitting step problem (7.2) only minimizes over a compact set $\mathcal{K}_k \subset \mathbb{R}^k$. However, this is not restrictive since the set $\mathcal{K}_k$ can be chosen arbitrarily

Algorithm 7.1 Nonlinear Greedy Reconstruction Algorithm

Require: A set of linearly independent operators $\mathcal{A} = \{A_1, \dots, A_K\}$, an (initial) operator $A(\boldsymbol{\alpha}_\circ) \in \operatorname{span} \mathcal{A}$ and a family of compact sets $\mathcal{K}_j \subset \mathbb{R}^j$, $j = 1, \dots, K-1$.

Compute the control $\boldsymbol{\epsilon}^1$ by solving

$$\max_{\boldsymbol{\epsilon} \in E_{ad}} \|C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ), \boldsymbol{\epsilon}; T) - C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ) + A_1, \boldsymbol{\epsilon}; T)\|_2^2 . \tag{7.1}$$

for $k = 1, \dots, K-1$ **do**

Fitting step: $A^{(k)}(\boldsymbol{\beta}) := \sum_{j=1}^k \boldsymbol{\beta}_j A_j$, find $\boldsymbol{\beta} = (\boldsymbol{\beta}_j^k)_{j=1,\dots,k}$ that solves

$$\min_{\boldsymbol{\beta} \in \mathcal{K}_k} \sum_{m=1}^k \left\| C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ) + A^{(k)}(\boldsymbol{\beta}), \boldsymbol{\epsilon}^m; T) - C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ) + A_{k+1}, \boldsymbol{\epsilon}^m; T) \right\|_2^2 . \tag{7.2}$$

Splitting step: Find $\boldsymbol{\epsilon}^{k+1}$ that solves

$$\max_{\boldsymbol{\epsilon} \in E_{ad}} \left\| C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ) + A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}; T) - C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ) + A_{k+1}, \boldsymbol{\epsilon}; T) \right\|_2^2 . \tag{7.3}$$

end for

large. Finally, the initialization problem (7.1) is different from the initialization (4.3). This is due to results obtained in [13] which suggest that one should not simply maximize the state corresponding to the first element A_1 in the set, but rather maximize the difference to the state that is observed when no elements from $\mathcal{A}$ are considered.

We recall that, in order to obtain our main results for Algorithm 7.1, it is sufficient to prove two points. First, that the fitting step (7.2) identifies the kernel of the submatrix $[\widehat{W}_\circ^{(k)}]_{[1:k+1,1:k+1]}$. Second, that for the initialization and each splitting step in Algorithm 7.1 there exists at least one control for which the corresponding cost function is strictly positive (making the submatrix $[\widehat{W}_\circ^{(k+1)}]_{[1:k+1,1:k+1]}$ PD).

To prove the fitting step result, we need some continuity properties of the argmin operator. For this purpose, we introduce the following definition of hemi-continuous set-valued correspondences (see, e.g., [8, Chapter VI,§1]).

Definition 7.1 (hemi-continuity). *Let $X \subset \mathbb{R}$ be an open set. A correspondence $c : X \rightrightarrows \mathbb{R}^k$ is called upper hemi-continuous (u.h.c.) if for each $x_0 \in X$ and each open set $G \subset \mathbb{R}^k$ with $c(x_0) \subset G$ there is a neighborhood $U(x_0)$ such that $x \in U(x_0) \Rightarrow c(x) \subset G$ and called lower hemi-continuous (l.h.c.) if for each $x_0 \in X$ and each open set $G \subset \mathbb{R}^k$ with $c(x_0) \cap G \neq \emptyset$ there is a neighborhood $U(x_0)$ such that $x \in U(x_0) \Rightarrow c(x) \cap G \neq \emptyset$. Furthermore, $c : X \rightrightarrows \mathbb{R}^k$ is hemi-continuous if it is u.h.c. and l.h.c.*

Using Definition 7.1, we can recall the Berge maximum theorem [2, Thm. 17.31].

Lemma 7.2 (Berge maximum theorem). *Let $X \subset \mathbb{R}$ be an open interval. Let $J : \mathbb{R}^k \times X \to \mathbb{R}$ be a continuous function and $\phi : X \rightrightarrows \mathbb{R}^k$ be a hemi-continuous, set-valued correspondence such that $\phi(x)$ is nonempty and compact for any $x \in X$. Then the correspondence $c : X \rightrightarrows \mathbb{R}^k$ defined by $c(x) := \operatorname*{arg\,min}_{z \in \phi(x)} J(z; x)$ is u.h.c.*

We will also need the following technical lemma.

Lemma 7.3 (limit of set-valued correspondance). *Let $X \subset \mathbb{R}$ be an open interval with $0 \in X$, and $c : X \rightrightarrows \mathbb{R}^k$ be a u.h.c. correspondence. If $c(0) = \{0\}$, then $\lim_{k\to\infty} c(x_k) = \{0\}$ for any sequence $\{x_k\}_{k=1}^\infty$ such that $\lim_{k\to\infty} x_k = 0$.*

Proof. Consider an arbitrary sequence $\{x_k\}_{k=1}^\infty$ with $\lim_{k\to\infty} x_k = 0$, and let $c(0) = \{0\}$. It is sufficient to show that for any $\epsilon > 0$ there exists $n_\epsilon \in \mathbb{N}$ such that for all $k \geq n_\epsilon$ we have $c(x_k) \subset \mathcal{B}_\epsilon^k(0)$. Let $\epsilon > 0$ and define $G_\epsilon := \mathcal{B}_\epsilon^k(0)$. Since $c(0) = \{0\}$

and c is u.h.c., there exists a neighborhood $U_\epsilon(0) \subset \mathbb{R}$ such that $c(x) \subset G_\epsilon$ for any $x \in U_\epsilon(0)$. Since $U_\epsilon(0)$ is an open neighborhood of 0, there exists $\xi_\epsilon > 0$ such that $(-\xi_\epsilon, \xi_\epsilon) \subset U_\epsilon(0)$. Since $\lim_{k\to\infty} x_k = 0$, there exists n_ϵ such that for all $k \geq n_\epsilon$ we have $x_k \in (-\xi_\epsilon, \xi_\epsilon)$ and hence $c(x_k) \subset \mathcal{B}^k_\epsilon(0)$. ∎

To use Lemmas 7.2 and 7.3, we make the following assumptions.

Assumption 7.4. *Let $k \in \{1, \dots, K-1\}$ and define*

$$J_k(\boldsymbol{\beta}; A_{k+1}) := \sum_{m=1}^{k} \| C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ) + A^{(k)}(\boldsymbol{\beta}), \boldsymbol{\epsilon}^m; T) - C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ) + A_{k+1}, \boldsymbol{\epsilon}^m; T)\|_2^2.$$

- *If $\|A_{k+1}\|$ is small enough, then there exists a $\boldsymbol{\beta}^k = \boldsymbol{\beta}^k(A_{k+1})$ that solves (7.2) with $J_k(\boldsymbol{\beta}^k; A_{k+1}) = 0$.*
- *There exists $\nu > 0$ such that $\mathcal{B}^k_\nu(0) \subset \mathcal{K}_k$ and $\arg\min_{\boldsymbol{\beta} \in \overline{\mathcal{B}^k_\nu(0)}} J_k(\boldsymbol{\beta}; 0) = \{0\}$.*

The first point in Assumption 7.4 guarantees that locally near $A(\boldsymbol{\alpha}_\circ)$, for $\|A_{k+1}\|$ small enough, one can solve (7.2) making the cost function zero, meaning that one can find a linear combination of the first k elements for which the final state cannot be distinguished from the $k+1$-th element by any of the k computed controls. On the other hand, if the minimum function value is strictly positive, then there already exists a control in the set $(\boldsymbol{\epsilon}_m)_{m=1}^k$ that discriminates (splits) these two states.

The second point in Assumption 7.4 ensures that $\{0\} = \arg\min_{\boldsymbol{\beta} \in \overline{\mathcal{B}^k_\nu(0)}} J_k(\boldsymbol{\beta}, 0)$. If this was not true, then for any $\nu > 0$, the ball $\mathcal{B}^k_\nu(0)$ would contain infinitely many $\boldsymbol{\beta} \in \mathbb{R}^k \setminus \{0\}$ satisfying $J_k(\boldsymbol{\beta}, 0) = 0$. Hence, for an infinite number of linear combinations in the set $\{A_1, \dots, A_k\}$, the corresponding states could not be distinguished by any of the previously selected controls. This implies that one of the previous splitting steps was not successful, which contradicts what we assume to reach iteration k.

Now, we can show that the local nonlinear fitting step problem (7.2) is able to identify the kernel of the submatrix $[\widehat{W}^{(k)}_\circ]_{[1:k+1,1:k+1]}$, if it exists.

Theorem 7.5 (nonlinear GR fitting step problems). *Let $k \in \{1, \dots, K\}$ and let $\boldsymbol{\beta}^k$ be a solution to (7.2). If $\|A_{k+1}\|$ is sufficiently small and Assumption 7.4 holds, then $\boldsymbol{\beta}^k$ solves (4.4) with $\sum_{m=1}^k \|C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^m; T) - C\delta\boldsymbol{y}_\circ(A_{k+1}, \boldsymbol{\epsilon}^m; T)\|_2^2 = 0$.*

Proof. Let $A_{k+1} = \delta_k \widetilde{A}_{k+1}$, where $\delta_k = \|A_{k+1}\|$ and $\|\widetilde{A}_{k+1}\| = 1$. Now, define $\widehat{J}_k(\boldsymbol{\beta}, \delta_k) := J_k(\boldsymbol{\beta}, \delta_k \widetilde{A}_{k+1})$. The first point of Assumption 7.4 implies that there is a $\widehat{\delta}_k > 0$ such that $\widehat{J}_k(\boldsymbol{\beta}, \delta_k) = 0$ for all $|\delta_k| < \widehat{\delta}_k$. Thus, Lemma 7.2 guarantees that $c_k : (-\widehat{\delta}_k, \widehat{\delta}_k) \rightrightarrows \mathbb{R}^k, c_k(\delta_k) = \arg\min_{\boldsymbol{\beta} \in \mathcal{K}_k} \widehat{J}_k(\boldsymbol{\beta}; \delta_k)$ is u.h.c.[6] According to the second point of Assumption 7.4, $c_k(0) = 0$ is an isolated solution of (7.2). Hence, the upper hemi-continuity of c_k guarantees that for $\delta_k \to 0$ we have $\boldsymbol{\beta}^k \to 0$ for any corresponding solution $\boldsymbol{\beta}^k = \boldsymbol{\beta}^k(\delta_k)$ of (7.2). Now, let $m \in \{1, \dots, k\}$. If $\widehat{J}_k(\boldsymbol{\beta}^k; \delta_k) = 0$, then

$$C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ) + A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^m; T) - C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ) + \delta_k \widetilde{A}_{k+1}, \boldsymbol{\epsilon}^m; T) = 0. \tag{7.4}$$

We define $g(\boldsymbol{\alpha}) := C\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; T)$. Since $f(A, \boldsymbol{y}, \boldsymbol{\epsilon})$ in (3.1) is assumed to be differentiable with respect to A and $\boldsymbol{y}$, we obtain that the map $A \mapsto \boldsymbol{y}(A, \boldsymbol{\epsilon}; T)$ is differentiable with respect to A by the implicit function theorem (see, e.g., [16, Thm. 17.13-1]). Hence, $C\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}; T)$ is also differentiable with respect to $\boldsymbol{\alpha}$. By Taylor's theorem,

[6]Note that, in this setting, the correspondence $\phi : (-\widehat{\delta}_k, \widehat{\delta}_k) \rightrightarrows \mathbb{R}^k$ mentioned in Lemma 7.2 is defined as $\phi(x) = \mathcal{K}_k$ for any $x \in (-\widehat{\delta}_k, \widehat{\delta}_k)$ with $\mathcal{K}_k$ compact, and is therefore hemi-continuous.

we get $g(\boldsymbol{\alpha}_\circ + \boldsymbol{v}) = g(\boldsymbol{\alpha}_\circ) + g'(\boldsymbol{\alpha}_\circ)(\boldsymbol{v}) + O(\|\boldsymbol{v}\|_2^2)$ for $\boldsymbol{v} \in \mathbb{R}^k$. Defining $\widehat{\boldsymbol{\beta}^k}$ and $\widehat{\boldsymbol{\delta}}_k$ as $\widehat{\boldsymbol{\beta}^k} := [\boldsymbol{\beta}^k, 0, \cdots, 0]^\top \in \mathbb{R}^k$ and $\widehat{\boldsymbol{\delta}}_k := [0, \cdots, 0, \delta_k]^\top \in \mathbb{R}^k$, we can rewrite (7.4) as

$$0 = g(\boldsymbol{\alpha}_\circ + \widehat{\boldsymbol{\beta}^k}) - g(\boldsymbol{\alpha}_\circ + \widehat{\boldsymbol{\delta}}_{k+1}) = g'(\boldsymbol{\alpha}_\circ)(\widehat{\boldsymbol{\beta}^k}) - g'(\boldsymbol{\alpha}_\circ)(\widehat{\boldsymbol{\delta}}_{k+1}) + O(\|\widehat{\boldsymbol{\beta}^k}\|_2^2) + O(|\delta_k|^2).$$

Since $g'(\boldsymbol{\alpha}_\circ)(\widehat{\boldsymbol{\beta}^k}) = C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^m; T)$ and $g'(\boldsymbol{\alpha}_\circ)(\widehat{\boldsymbol{\delta}}_{k+1}) = C\delta\boldsymbol{y}_\circ(\delta_k\widetilde{A}_{k+1}, \boldsymbol{\epsilon}^m; T)$,

$$0 = C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^m; T) - C\delta\boldsymbol{y}_\circ(\delta_k\widetilde{A}_{k+1}, \boldsymbol{\epsilon}^m; T) + O(\|\widehat{\boldsymbol{\beta}^k}\|_2^2) + O(|\delta_k|^2). \tag{7.5}$$

Since $\boldsymbol{\beta}^k = \boldsymbol{\beta}^k(\delta_k) \to 0$ for $\delta_k \to 0$, we know that all four terms vanish for $\delta_k \to 0$. However, $O(|\delta_k|^2)$ converges faster than $C\delta\boldsymbol{y}_\circ(\delta_k\widetilde{A}_{k+1}, \boldsymbol{\epsilon}^m; T)$ and $O(\|\widehat{\boldsymbol{\beta}^k}\|_2^2)$ faster than $C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^m; T)$. Hence, (7.5) can only be true for $\delta_k \to 0$ if $C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^m; T) - C\delta\boldsymbol{y}_\circ(\delta_k\widetilde{A}_{k+1}, \boldsymbol{\epsilon}^m; T) = 0$ for δ_k small enough, which is equivalent to $C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^m; T) - C\delta\boldsymbol{y}_\circ(A_{k+1}, \boldsymbol{\epsilon}^m; T) = 0$ for $\|A_{k+1}\|$ sufficiently small. □

Let us now comment about the assumption on $\|A_{k+1}\|$. The goal of this section is to prove a local result around the approximation $A(\boldsymbol{\alpha}_\circ)$. Since the map $\boldsymbol{\alpha} \mapsto A(\boldsymbol{\alpha})$ is linear, the terms $A(\boldsymbol{\alpha}_\circ) + A^{(k)}(\boldsymbol{\beta}^k)$ and $A(\boldsymbol{\alpha}_\circ) + A_{k+1}$ are perturbations of $A(\boldsymbol{\alpha}_\circ)$. These are clearly used in the proof of Theorem 7.5. In order to remain close to $A(\boldsymbol{\alpha}_\circ)$, we require that the basis elements A_{k+1} are sufficiently small. However, this is not a restrictive assumption since it can be interpreted as a simple rescaling of the operators $A_1, \dots, A_K$, which are only required to be linearly independent. Thus, the sufficiently small assumption of $\|A_{k+1}\|$ has to be understood in the sense that one has to perturb $A(\boldsymbol{\alpha}_\circ)$ by stepping in direction of the (linearly independent) element A_{k+1}. However, the length of the step must be sufficiently small to remain in a neighborhood of $A(\boldsymbol{\alpha}_\circ)$ (related to the Taylor expansion used in the proof of Theorem 7.5).

Regarding the initialization and splitting step result, we make now the assumption that there always exists a control that makes the corresponding cost function value strictly positive, and discuss specific cases where this assumption holds.

Assumption 7.6. *Let $k \in \{1, \dots, K-1\}$ and $\boldsymbol{\beta}^k \in \mathbb{R}^k$ be the solution of (7.2). There exists a solution $\boldsymbol{\epsilon}^{k+1} \in E_{ad}$ to (7.3) that simultaneously satisfies*

$$\|C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ) + A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^{k+1}; T) - C\boldsymbol{y}(A(\boldsymbol{\alpha}_\circ) + A_{k+1}, \boldsymbol{\epsilon}^{k+1}; T)\|_2^2 > 0, \ \textit{and} \tag{7.6}$$

$$\|C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^{k+1}; T) - C\delta\boldsymbol{y}_\circ(A_{k+1}, \boldsymbol{\epsilon}^{k+1}; T)\|_2^2 > 0. \tag{7.7}$$

Let (7.6)-(7.7) also hold for a solution $\boldsymbol{\epsilon}^1 \in E_{ad}$ to (7.1) with $k = 0$ and $\boldsymbol{\beta}^0 = 0$.

In Theorem 7.10, we will investigate Assumption 7.6 for the two settings considered in sections 5 and 6. Now, we state a result relating Algorithms 4.1 and 7.1.

Theorem 7.7 (positive definiteness of the GN matrix $\widehat{W}_\circ$ (general systems)). *Consider the general setting of (3.1) with $\{A_1, \dots, A_K\}$ linearly independent such that $\|A_k\|$ is small enough for all $k \in \{1, \dots, K\}$. Let $(\boldsymbol{\epsilon}^m)_{m=1}^K \subset E_{ad}$ be generated by Algorithm 7.1 such that Assumption 7.4 holds for all $k \in \{1, \dots, K-1\}$ and $\boldsymbol{\epsilon}^m$ satisfies Assumption 7.6 for all $m \in \{1, \dots, K\}$. Then $\widehat{W}_\circ$ is PD.*

The proof of Theorem 7.7 is similar to that of Theorem 5.6 and is omitted for brevity.

In the following subsections 7.1.1 and 7.1.2 we discuss Assumption 7.6 for linear and bilinear systems and for a general class of nonliner systems, respectively.

7.1.1. Linear and bilinear control systems. In this section, we discuss Assumption 7.6 in the settings of sections 5 and 6. First, we require the following results.

LEMMA 7.8 (on analytic functions in Banach spaces [33, p. 1079]). *Let X, Y denote real Banach spaces and $\mathcal{B}_r(x) \subset X$ the open ball with center $x \in X$ and radius $r > 0$. For an open set $D \subset X$, let the functions $f, g : D \to Y$ be analytic. If there exist $x_f, x_g \in D$ such that $f(x_f) \neq 0$ and $g(x_g) \neq 0$, then for any $x \in D$ and any $r > 0$ there exists a $\widetilde{x} \in \mathcal{B}_r(x) \subset D$ such that $f(\widetilde{x}) \neq 0$ and $g(\widetilde{x}) \neq 0$.*

LEMMA 7.9 (analycity of control-to-state maps). *Consider system (3.1) and define the map $c : U \times Y \to Z$ as $c(\boldsymbol{\epsilon}, \boldsymbol{y}) := [\dot{\boldsymbol{y}} - f(A, \boldsymbol{y}, \boldsymbol{\epsilon}), \boldsymbol{y}(0) - \boldsymbol{y}^0]$, where U is the Hilbert space of control functions, Y is the (Banach) space where solutions to (3.1) lie and Z is a Banach space. If c is analytic in $\boldsymbol{\epsilon}$ and $\boldsymbol{y}$, (3.1) has a unique solution $\boldsymbol{y} = \boldsymbol{y}(\boldsymbol{\epsilon}) \in Y$ such that $c(\boldsymbol{y}(\boldsymbol{\epsilon}), \boldsymbol{\epsilon}) = 0$ for each $\boldsymbol{\epsilon} \in E_{ad} \subset U$ and the linearized state equation $\dot{\delta\boldsymbol{y}} = \delta_{\boldsymbol{y}} f(A, \boldsymbol{y}(\boldsymbol{\epsilon}), \boldsymbol{\epsilon})(\delta\boldsymbol{y}) - \varphi$ with $\delta\boldsymbol{y}(0) = \varphi^0$ is uniquely solvable for any $[\varphi, \varphi^0] \in Z$, then the control-to-state map $L : E_{ad} \to Y, \boldsymbol{\epsilon} \mapsto \boldsymbol{y}(\boldsymbol{\epsilon})$ is analytic. If the solution space Y is such that the evaluation map $S_T : Y \to \mathbb{R}^N, \boldsymbol{y} \mapsto \boldsymbol{y}(T)$ is linear and continuous, then also the map $S : E_{ad} \to \mathbb{R}^N, \boldsymbol{\epsilon} \mapsto (\boldsymbol{y}(\boldsymbol{\epsilon}))(T)$ is analytic.*

Proof. First, we prove that the control-to-state map $L : E_{ad} \to Y, \boldsymbol{\epsilon} \mapsto \boldsymbol{y}(\boldsymbol{\epsilon})$ is analytic. This follows directly from the implicit function theorem [33, p. 1081] if we can show that the map $D_{\boldsymbol{y}} c(\boldsymbol{\epsilon}, \boldsymbol{y})$ is an isomorphism of Y on Z for any pair $(\widetilde{\boldsymbol{\epsilon}}, \widetilde{\boldsymbol{y}}) \subset U \times Y$ such that $\widetilde{\boldsymbol{y}}$ is the unique solution to (3.1) for $\widetilde{\boldsymbol{\epsilon}}$, i.e. $c(\widetilde{\boldsymbol{\epsilon}}, \widetilde{\boldsymbol{y}}) = 0$. Since $D_{\boldsymbol{y}} c(\widetilde{\boldsymbol{\epsilon}}, \widetilde{\boldsymbol{y}})(\delta\boldsymbol{y}) = \varphi$, which is equivalent to $\dot{\delta\boldsymbol{y}} = \delta_{\boldsymbol{y}} f(A, \widetilde{\boldsymbol{y}}, \widetilde{\boldsymbol{\epsilon}})(\delta\boldsymbol{y}) - \varphi$ with $\delta\boldsymbol{y}(0) = \varphi^0$, admits a unique solution $\delta\boldsymbol{y} \in Y$ for any $[\varphi, \varphi^0] \in Z$, $D_{\boldsymbol{y}} c(\widetilde{\boldsymbol{\epsilon}}, \widetilde{\boldsymbol{y}})$ is bijective and hence an isomorphism of Y on Z. It remains to show that also the map $S : E_{ad} \to \mathbb{R}^N, \boldsymbol{\epsilon} \mapsto (\boldsymbol{y}(\boldsymbol{\epsilon}))(T)$ is analytic. Consider an arbitrary $\boldsymbol{\epsilon}_0 \in E_{ad}$. Since the control-to-state map L is analytic, there exist (by definition, see, e.g., [33, p. 1078]) ℓ-linear, symmetric and continuous maps $a_\ell : (E_{ad})^\ell \to \mathbb{R}^N, (\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_\ell) \mapsto a_\ell(\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_\ell)$ such that $\boldsymbol{y}(\boldsymbol{\epsilon}) = \sum_{\ell=0}^\infty a_\ell(\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_0)^\ell$. Now, define the maps $b_\ell : (E_{ad})^\ell \to \mathbb{R}^N$ as $b_\ell(\boldsymbol{\epsilon})^\ell := (a_\ell(\boldsymbol{\epsilon})^\ell)(T)$, meaning that $\sum_{\ell=0}^\infty b_\ell(\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_0)^\ell = (\boldsymbol{y}(\boldsymbol{\epsilon}))(T)$. Since $S_T : Y \to \mathbb{R}^N, \boldsymbol{y} \mapsto \boldsymbol{y}(T)$ is linear and continuous, the maps b_ℓ are ℓ-linear, symmetric and continuous. Thus, the map $S : E_{ad} \to \mathbb{R}^N, \boldsymbol{\epsilon} \mapsto (\boldsymbol{y}(\boldsymbol{\epsilon}))(T) = \sum_{\ell=0}^\infty b_\ell(\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_0)^\ell$ is analytic by definition. ∎

In our case, we consider $U = L^2(0, T; \mathbb{R}^M)$ in the linear setting and $U = L^2(0, T; \mathbb{R})$ in the bilinear setting, $Y = H^1(0, T; \mathbb{R}^N)$ and $Z = L^2(0, T; \mathbb{R}^N) \times \mathbb{R}^N$. Then, the assumptions in Lemma 7.9 on the ODE system and its linearization are satisfied for (5.1) and (5.2) in the linear setting, and for (6.1) and (6.3) in the bilinear setting.[7] Notice that all solutions lie in $H^1(0, T; \mathbb{R}^N) \Subset C(0, T; \mathbb{R}^N)$ (see, e.g., [16]), which implies that the evolution map $S_T : H^1(0, T; \mathbb{R}^N) \to \mathbb{R}^N, \boldsymbol{y} \mapsto \boldsymbol{y}(T)$ is also linear and continuous. Now, we can prove our main result.

THEOREM 7.10 (analysis for linear and bilinear systems). *Consider the linear setting (5.1) or the bilinear setting (6.1). Assume that the systems are sufficiently observable and controllable, i.e. fully observable and controllable in the linear case, and satisfying Assumption 6.2 in the bilinear case. If $\|A_{k+1}\|$ is sufficiently small, then the set of controls in E_{ad} that satisfy (7.6)-(7.7) in Assumption 7.6 is nonempty.*

Proof. For brevity, we denote $A_{\boldsymbol{\beta}} := A(\boldsymbol{\alpha}_\circ) + A^{(k)}(\boldsymbol{\beta}^k)$, $A_+ := A(\boldsymbol{\alpha}_\circ) + A_{k+1}$, $\boldsymbol{y}_{\boldsymbol{\beta}}(\boldsymbol{\epsilon}; t) := \boldsymbol{y}(A_{\boldsymbol{\beta}}, \boldsymbol{\epsilon}; t)$ and $\boldsymbol{y}_+(\boldsymbol{\epsilon}; t) := \boldsymbol{y}(A_+, \boldsymbol{\epsilon}; t)$. We start with the linear setting (5.1) from section 5. First, we derive observability and controllability properties for the systems (A_+, B, C) and $(A_{\boldsymbol{\beta}}, B, C)$. Denote by $\sigma_k > 0$ the smallest singular value of $\mathcal{O}_N(C, A(\boldsymbol{\alpha}_\circ))$. Let $k \in \{1, \ldots, K\}$ and $\boldsymbol{\beta}^k \in \mathbb{R}^k$ be the solution of (7.2) for

[7] Existence and uniqueness of all solutions $\boldsymbol{y}, \delta\boldsymbol{y}$ follow by Carathéodory's existence theorem (see, e.g., [31, Thm. 54] and related propositions). For $\boldsymbol{\epsilon} \in L^2(0, T; \mathbb{R}^M)$ in the linear and $\epsilon \in L^2(0, T; \mathbb{R})$ in the bilinear setting, we obtain $\dot{\boldsymbol{y}}, \dot{\delta\boldsymbol{y}} \in L^2(0, T; \mathbb{R}^N)$ and thus $\boldsymbol{y}, \delta\boldsymbol{y} \in H^1(0, T; \mathbb{R}^N)$.

$\|A_{k+1}\| > 0$ sufficiently small such that $\|\mathcal{O}_N(C, A(\boldsymbol{\alpha}_\circ)) - \mathcal{O}_N(C, A_+)\|_2 < \sigma_k$. From the proof of Theorem 7.5, we obtain that also $\boldsymbol{\beta}^k$ can be assumed to be sufficiently small such that $\|\mathcal{O}_N(C, A(\boldsymbol{\alpha}_\circ)) - \mathcal{O}_N(C, A_{\boldsymbol{\beta}})\|_2 < \sigma_k$. Now, Lemma 5.7 guarantees that $\text{rank}(\mathcal{O}_N(C, A_+)) = \text{rank}(\mathcal{O}_N(C, A_{\boldsymbol{\beta}})) = N$. Using the same argument for the rank of the controllability matrices, we obtain that the systems (A_+, B, C) and $(A_{\boldsymbol{\beta}}, B, C)$ are fully observable and controllable.

Next, we consider the state of the difference $\boldsymbol{z}(t) = \boldsymbol{y}(A_+, \boldsymbol{\epsilon}; t) - \boldsymbol{y}(A_{\boldsymbol{\beta}}, \boldsymbol{\epsilon}; t)$ with $\dot{\boldsymbol{z}} = A_+\boldsymbol{z} + (A_+ - A_{\boldsymbol{\beta}})\boldsymbol{y}(A_{\boldsymbol{\beta}}, \boldsymbol{\epsilon}; t)$. Since $A_+ \neq A_{\boldsymbol{\beta}}$, there exists $\boldsymbol{v} \in \mathbb{R}^N$ such that $(A_+ - A_{\boldsymbol{\beta}})\boldsymbol{v} \neq 0$. Recalling that $(A_{\boldsymbol{\beta}}, B)$ is controllable, we can find $\boldsymbol{\epsilon}_{t_1}$ for any $t_1 \in (0, T]$ such that $\boldsymbol{y}_{\boldsymbol{\beta}}(\boldsymbol{\epsilon}_{t_1};) = \boldsymbol{v}$ and therefore $(A_+ - A_{\boldsymbol{\beta}})\boldsymbol{y}_{\boldsymbol{\beta}}(\boldsymbol{\epsilon}_{t_1}; t_1) \neq 0$. We define $\widetilde{\boldsymbol{\epsilon}}(s) := \begin{cases} \boldsymbol{\epsilon}_{t_1}(s), & \text{for } 0 \leq s < t_1, \\ \boldsymbol{c}, & \text{for } t_1 \leq s \leq T, \end{cases}$ where $\boldsymbol{c} \in \mathbb{R}^N$ is chosen later. For $t > t_1$, we define

$$\widetilde{\boldsymbol{z}}(t) := e^{-(t-t_1)A_+}\boldsymbol{z}(t) = \boldsymbol{z}(t_1) + \int_{t_1}^{t} e^{(t_1-s)A_+}(A_+ - A_{\boldsymbol{\beta}})\boldsymbol{y}_{\boldsymbol{\beta}}(\widetilde{\boldsymbol{\epsilon}}; s)ds. \tag{7.8}$$

Notice that for $s > t_1$, the terms $e^{(t_1-s)A_+}$ and $\boldsymbol{y}_{\boldsymbol{\beta}}(\widetilde{\boldsymbol{\epsilon}}; s) = e^{(s-t_1)A_{\boldsymbol{\beta}}}\boldsymbol{v} + \int_0^s e^{(s-\tau)A_{\boldsymbol{\beta}}}B\boldsymbol{c}ds$ are continuous in s. Since exponential matrices are invertible (see, e.g., [25, pag. 369, 5.6.P43]) and $\boldsymbol{z}(t_1)$ is independent of t, there exists a $t > t_1$ such that $\boldsymbol{z}(t_1) + \int_{t_1}^{t} e^{(t_1-s)A_+}(A_+ - A_{\boldsymbol{\beta}})\boldsymbol{y}_{\boldsymbol{\beta}}(\widetilde{\boldsymbol{\epsilon}}; s)ds \neq 0$ and thus $\widetilde{\boldsymbol{z}}(t) \neq 0$. Using (7.8), we obtain

$$C\boldsymbol{z}(t) = Ce^{(t-t_1)A_+}\widetilde{\boldsymbol{z}}(t) = \sum_{j=0}^{\infty} \frac{(t-t_1)^j}{j!} CA_+^j \widetilde{\boldsymbol{z}}(t). \tag{7.9}$$

The observability of (A_+, B, C) guarantees the existence of $i \in \{0, \dots, N-1\}$ such that $CA_+^i \widetilde{\boldsymbol{z}}(t) \neq 0$. We have $\frac{(t-t_1)^i}{i!} > 0$ for $t > t_1$ and all summands in (7.9) converge to zero at different rates. Hence, there exists $t > t_1$ such that $C\boldsymbol{z}(t) \neq 0$. Since $t_1 \in (0, T]$ was chosen arbitrarily, we get $C\boldsymbol{z}(T) \neq 0$ and thus $C\boldsymbol{y}_{\boldsymbol{\beta}}(\widetilde{\boldsymbol{\epsilon}}; T) - C\boldsymbol{y}_+(\widetilde{\boldsymbol{\epsilon}}; T) \neq 0$.

Regarding the linearized system (5.2), we have already shown in Lemma 5.5 that there exists an $\boldsymbol{\epsilon} \in E_{ad}$ such that $C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}; T) - C\delta\boldsymbol{y}_\circ(A_{k+1}, \boldsymbol{\epsilon}; T) \neq 0$.

Finally, the maps $S, S_\delta : L^2(0, T; \mathbb{R}^M) \to \mathbb{R}^N$, $S(\boldsymbol{\epsilon}) := C\boldsymbol{y}_{\boldsymbol{\beta}}(\boldsymbol{\epsilon}; T) - C\boldsymbol{y}_+(\boldsymbol{\epsilon}; T)$, $S_\delta(\boldsymbol{\epsilon}) := C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}; T) - C\delta\boldsymbol{y}_\circ(A_{k+1}, \boldsymbol{\epsilon}; T)$ are analytic by Lemma 7.9. Using Lemma 7.8, we obtain the existence of an $\boldsymbol{\epsilon} \in E_{ad}$ such that $C\boldsymbol{y}(A_{\boldsymbol{\beta}}, \boldsymbol{\epsilon}; T) - C\boldsymbol{y}(A_+, \boldsymbol{\epsilon}; T) \neq 0$ and $C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}; T) - C\delta\boldsymbol{y}_\circ(A_{k+1}, \boldsymbol{\epsilon}; T) \neq 0$.

The proof for the bilinear setting (6.1) from section 6 is analogous to the one above and we omit it here for brevity. ∎

Remark 7.11. Notice that we did not prove exactly Assumption 7.6 in Theorem 7.10, but only the existence of a general control $\boldsymbol{\epsilon} \in E_{ad}$ that satisfies (7.6)-(7.7). However, this implies that any solution $\boldsymbol{\epsilon}^{k+1}$ to (7.3) always satisfies (7.6). Additionally, we recall from the proof of Theorem 7.10 that the maps $S, S_\delta : L^2(0, T; \mathbb{R}^M) \to \mathbb{R}^N$, defined by $S(\boldsymbol{\epsilon}) := \boldsymbol{y}(A, \boldsymbol{\epsilon}; T)$, $S_\delta(\boldsymbol{\epsilon}) := \delta\boldsymbol{y}_\circ(A, \boldsymbol{\epsilon}; T)$ are nonzero. Thus, we obtain by Lemma 7.8 that any neighborhood of $\boldsymbol{\epsilon}^{k+1}$ can contain only isolated controls not satisfying (7.7), and infinitely many $\boldsymbol{\epsilon}$ that do satisfy (7.7). Thus, it is rather unlucky to choose an $\boldsymbol{\epsilon}^{k+1}$ not satisfying (7.7). On the other hand, one can also add inequality (7.7) as a constraint to (7.3) to ensure that both inequalities are met by $\boldsymbol{\epsilon}^{k+1}$.

As a consequence of Theorems 7.7, 7.10 and Remark 7.11, the controls generated by Algorithm 7.1 for (5.1) and (6.1) make the GN matrix $\widehat{W}_\circ$, defined in (4.6), PD under certain assumptions. Thus, the results from sections 5.2 and 6.2 imply that GN for the reconstruction problems (5.5) and (6.2), initialized with $\boldsymbol{\alpha}_\circ$, converges to $\boldsymbol{\alpha}_\star$.

7.1.2. A general class of nonlinear systems. Now, we focus on a class of more general nonlinear systems and prove our second main result, which is analogous to Theorem 7.10. These systems (of the form (3.1)) are characterized by a function f of the form $f(A, \boldsymbol{y}, \boldsymbol{\epsilon}) = g(\boldsymbol{y}) + A\boldsymbol{\epsilon}$, and thus have the form

$$\dot{\boldsymbol{y}}(t) = g(\boldsymbol{y}(t)) + A(\boldsymbol{\alpha}_\star)\boldsymbol{\epsilon}(t), \quad t \in [0, T], \tag{7.10}$$

where $\boldsymbol{y}(t) \in \mathbb{R}^N$, $N \in \mathbb{N}$. We make the following hypotheses.

Assumption 7.12 (on the system (7.10)). *The initial state is $\boldsymbol{y}(0) = 0$. The function g is in C^1 and satisfies $g(0) = 0$.*[8] *Its Lipschitz constant is denoted by L. The approximation of $\boldsymbol{\alpha}_\star$ is given by $\boldsymbol{\alpha}_\circ = 0$. The state $\boldsymbol{y}(T)$ can be fully observed, meaning that the observer matrix is the identity: $C = I \in \mathbb{R}^{N\times N}$.*

Assumption 7.12 and $A(\boldsymbol{\alpha}_\circ = 0) = 0$ imply that the linearized equation in $\boldsymbol{\alpha}_\circ = 0$ is

$$\dot{\delta\boldsymbol{y}}_\circ(t) = g'(\boldsymbol{y}_\circ(t))\delta\boldsymbol{y}_\circ(t) + A(\boldsymbol{\alpha})\boldsymbol{\epsilon}(t), \quad \delta\boldsymbol{y}(0) = 0,$$
$$\dot{\boldsymbol{y}}_\circ(t) = g(\boldsymbol{y}_\circ(t)), \quad \boldsymbol{y}_\circ(0) = 0.$$

Using that $\boldsymbol{y}_\circ(0) = 0$ and $g(0) = 0$ from Assumption 7.12, we get $\boldsymbol{y}_\circ(t) = 0$ for all $t \in [0, T]$, and thus $\dot{\delta\boldsymbol{y}}_\circ(t) = g'(0)\delta\boldsymbol{y}_\circ(t) + A(\boldsymbol{\alpha})\boldsymbol{\epsilon}(t)$. Now, we show that there exists a control satisfying inequality (7.6) in Assumption 7.6. We then introduce the notation $\widetilde{A} := A^{(k)}(\boldsymbol{\beta}^k) - A_{k+1}$, $\boldsymbol{y}_{\boldsymbol{\beta}}(t) := \boldsymbol{y}(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^{k+1}; t)$, $\boldsymbol{y}_{k+1}(t) := \boldsymbol{y}(A_{k+1}, \boldsymbol{\epsilon}^{k+1}; t)$, and $\Delta\boldsymbol{y}(t) := \boldsymbol{y}_{\boldsymbol{\beta}}(t) - \boldsymbol{y}_{k+1}(t)$. Then, $\Delta\boldsymbol{y}(t)$ satisfies

$$\dot{\Delta\boldsymbol{y}}(t) = g(\boldsymbol{y}_{\boldsymbol{\beta}}(t)) - g(\boldsymbol{y}_{k+1}(t)) + \widetilde{A}\boldsymbol{\epsilon}^{k+1}(t), \quad \Delta\boldsymbol{y}(0) = 0. \tag{7.11}$$

By multiplying (7.11) with $\widetilde{A}\boldsymbol{\epsilon}^{k+1}(t)$, and using the Lipschitz continuity of g, we get $\|\widetilde{A}\boldsymbol{\epsilon}^{k+1}(t)\|_2^2 \leq \|\widetilde{A}\boldsymbol{\epsilon}^{k+1}(t)\|_2\Big(\|\dot{\Delta\boldsymbol{y}}(t)\|_2 + L\|\Delta\boldsymbol{y}(t)\|_2\Big)$, where we used Cauchy-Schwarz and triangle inequalities. Thus, we obtain $\|\widetilde{A}\boldsymbol{\epsilon}^{k+1}(t)\|_2 \leq \|\dot{\Delta\boldsymbol{y}}(t)\|_2 + L\|\Delta\boldsymbol{y}(t)\|_2$. Since $A^{(k)}(\boldsymbol{\beta}^k)$ and A_{k+1} are linearly independent, we have that $\widetilde{A} \neq 0$. Thus, there exists $\boldsymbol{v} \in \mathbb{R}^M$ such that $\widetilde{A}\boldsymbol{v} \neq 0$. Now, we define the control $\boldsymbol{\epsilon}_{\boldsymbol{v},\tau}(t) := \begin{cases} 0, & t \in [0, \tau), \\ \boldsymbol{v}, & t \in [\tau, T], \end{cases}$ for an arbitrary $\tau \in (0, T)$. Hence, it holds that $\|\dot{\Delta\boldsymbol{y}}(t)\|_2 + L\|\Delta\boldsymbol{y}(t)\|_2 \geq \|\widetilde{A}\boldsymbol{\epsilon}_{\boldsymbol{v},\tau}(t)\|_2 > 0$ for all $t > \tau$. Recalling that $C = I$ by Assumption 7.12, if $\|\Delta\boldsymbol{y}(T)\|_2 > 0$, the first inequality (7.6) is satisfied and we could move to inequality (7.7). Thus, assume that $\|\Delta\boldsymbol{y}(T)\|_2 = 0$. Then $\|\dot{\Delta\boldsymbol{y}}(T)\|_2 \neq 0$ and therefore $\|\Delta\boldsymbol{y}(T - \eta)\|_2 \neq 0$ for some $\eta > 0$. Notice that η is independent of τ. Since $\Delta\boldsymbol{y}(t) = 0$ for $t \in [0, \tau)$, we can simply shift τ by η, namely considering $\boldsymbol{\epsilon}_{\boldsymbol{v},\tau+\eta}$, to obtain $\|\Delta\boldsymbol{y}(T)\|_2 > 0$. In conclusion, we have found a control, namely $\boldsymbol{\epsilon}^{k+1} = \boldsymbol{\epsilon}_{\boldsymbol{v},\tau+\eta}$, satisfying (7.6).

Now, we show that $\boldsymbol{\epsilon}^{k+1} = \boldsymbol{\epsilon}_{\boldsymbol{v},\tau+\eta}$ satisfies (7.7) in Assumption 7.6 as well. To do so, let $\Delta\boldsymbol{y}_\delta(t) := \delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^{k+1}; t) - \delta\boldsymbol{y}_\circ(A_{k+1}, \boldsymbol{\epsilon}^{k+1}; t)$, which clearly satisfies $\dot{\Delta\boldsymbol{y}}_\delta(t) = g'(0)\Delta\boldsymbol{y}_\delta(t) + \widetilde{A}\boldsymbol{\epsilon}^{k+1}(t)$, with $\Delta\boldsymbol{y}_\delta(0) = 0$. Now, since $\widetilde{A}\boldsymbol{\epsilon}^{k+1}(t) = \widetilde{A}\boldsymbol{\epsilon}_{\boldsymbol{v},\tau+\eta}(t) \neq 0$ for $t \in [\tau + \eta, T]$, thus $\dot{\Delta\boldsymbol{y}}_\delta(\tau + \eta) \neq 0$ Hence, since $t \mapsto \Delta\boldsymbol{y}_\delta(t)$ is continuous, there exists a $\widetilde{\eta} > 0$ such that $\tau + \widetilde{\eta} \in (\tau + \eta, T]$ and $\Delta\boldsymbol{y}_\delta(t) \neq 0$ for all $t \in (\tau + \eta, \tau + \widetilde{\eta}]$. If $\Delta\boldsymbol{y}_\delta(T) \neq 0$ we proved our claim. Thus, we assume that $\tau + \widetilde{\eta} < T$ and $\Delta\boldsymbol{y}_\delta(T) = 0$. In this case, we can use the fact that τ is arbitrary and choose it to shift the interval $(\tau + \eta, \tau + \widetilde{\eta}]$ in order to get that $T \in (\tau + \eta, \tau + \widetilde{\eta}]$. This implies that the control $\boldsymbol{\epsilon}^{k+1} = \boldsymbol{\epsilon}_{\boldsymbol{v},\tau+\eta}$ satisfies both inequalities (7.6) and (7.7) in Assumption 7.6.

We summarize our findings in the next theorem (analogue of Theorem 7.10).

[8]The first two hypotheses imply that the system with $\boldsymbol{\epsilon} = 0$ has an equilibrium point at zero.

THEOREM 7.13 (analysis for systems of the form (7.10)). *Consider the system (7.10) and let Assumption 7.12 hold. Then the set of controls in E_{ad} that satisfy both (7.6) and (7.7) in Assumption 7.6 is nonempty.*

Remark 7.14. Theorem 7.13 is an alternative convergence proof for the linear setting considered in [13, sect. 4], where the unknown is the control matrix. However, it does not cover the setting of section 5, where the unknown is the drift matrix. Moreover, as in Remark 7.11, Theorem 7.13 does not exactly prove Assumption 7.6, but rather that there exist controls that simultaneously satisfy (7.6) and (7.7). On the other hand, the map $\boldsymbol{\epsilon} \mapsto \Delta \boldsymbol{y}_\delta(T) = \delta \boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}; T) - \delta \boldsymbol{y}_\circ(A_{k+1}, \boldsymbol{\epsilon}; T)$ is analytic by Lemma 7.9, and the map $t \mapsto \Delta \boldsymbol{y}_\delta(t)$ is (at least) continuous. In this case, the control functions $\boldsymbol{\epsilon}$ that violate simultaneously (7.6) and (7.7) are isolated points. Therefore, it is theoretically possible but also very unlikely that the solutions $\boldsymbol{\epsilon}^{k+1}$ to the nonlinear splitting step problem (7.3) do not satisfy (7.7) in Assumption 7.6.

7.2. Optimized GR Algorithm (OGR). The analyses discussed in the previous sections are based on certain hypotheses of observability and controllability of the dynamical system. However, as shown already in [13], if they are not satisfied, the choice of the elements of $\mathcal{A}$ becomes very relevant and can strongly affect the online reconstruction process. For this reason, a modified GR called OGR has been introduced in [13] to identify important basis elements by solving in each iteration the fitting and splitting step problems (in parallel) for all remaining basis elements. This also allows us to initialize the algorithm with a larger number of elements $(A_j)_{j=1}^K$, i.e., $K > N^2$. Even though some of the matrices A_j will inevitably be linearly dependent if $K > N^2$, OGR manipulates them to construct a new subset of linearly independent ones. In the spirit of the previous analysis, we add a new feature to the original OGR algorithm. At iteration k, after all fitting step problems have been solved, we check whether there exists $\ell \in \{k+1, \dots, K\}$ for which the optimal cost function value is different from zero (i.e. larger than a tolerance tol_2). If this is the case, then there exists a control $\boldsymbol{\epsilon}^m$, $m \in \{1, \dots, k\}$, that already satisfies $\left\|C\boldsymbol{y}(A^{(k)}(\boldsymbol{\beta}^\ell), \boldsymbol{\epsilon}^m; T) - C\boldsymbol{y}(A_\ell, \boldsymbol{\epsilon}^m; T)\right\|_2^2 > \text{tol}_2$ for at least one index $\ell_{k+1} \in \{k+1, \dots, K\}$ (see Step 8 in Algorithm 7.2). Hence, we can add the element $A_{\ell_{k+1}}$ to the already selected ones without computing a new control. This new improvement can also be motivated with the matrix formulation we used for our analysis. If $\text{rank}(\widehat{W}_\circ^{(k)}) = r > k$, one can appropriately permute rows and columns of $\widehat{W}_\circ^{(k)}$ such that $[\widehat{W}_\circ^{(k)}]_{[1:r,1:r]}$ has rank r and is thus PD. The rank of $\widehat{W}_\circ^{(k)} = \sum_{m=1}^k W_\circ(\boldsymbol{\epsilon}^m)$ is bounded by kP (recall $C \in \mathbb{R}^{P \times N}$). This can be seen by writing $W_\circ(\epsilon^m)$ as $W_\circ(\epsilon^m) = \delta Y_\circ^\top C^\top C \delta Y_\circ$, where $\delta Y_\circ := \left[\delta \boldsymbol{y}_\circ(A_1, \boldsymbol{\epsilon}^m; T), \cdots, \delta \boldsymbol{y}_\circ(A_K, \boldsymbol{\epsilon}^m; T)\right]$. Hence, $\text{rank}(W_\circ(\epsilon^m)) \leq \text{rank}(C) \leq P$, and thus $\text{rank}(\widehat{W}_\circ^{(k)}) \leq kP$. OGR is stated in Algorithm 7.2. The new feature corresponds to the steps 8-9. Algorithm 7.2 can also be formulated for the linearized setting by replacing $\boldsymbol{y}$ with its linearization $\delta \boldsymbol{y}_\circ$. We call OLGR the OGR algorithm for the linearized system.

7.3. Global convergence in a specific case. In the above sections, we discussed the performance of the GR algorithm on the local convergence of GN for the online identification problem. In the context of inverse problems for nonlinear ODEs, global convergence can generally not be expected, independently of the method used to generate control functions or data. A sufficient condition for the global convergence of GN applied to such problems is strict convexity of the online reconstruction problem (3.3). Thus, if GR can produce controls making the cost of (3.3) strictly convex, then GN (or other optimization solvers) exhibits global convergence. One

Algorithm 7.2 Optimized Greedy Reconstruction (OGR) Algorithm

Require: A set of K matrices $\mathcal{A} = \{A_1, \dots, A_K\}$ and two tolerances $\mathrm{tol}_1 > 0$ and $\mathrm{tol}_2 > 0$.

Set $\boldsymbol{\epsilon}^0 = 0$ and compute $\boldsymbol{\epsilon}^1$ and the index ℓ_1 by solving the initialization problem

$$\max_{\ell \in \{1,\dots,K\}} \max_{\boldsymbol{\epsilon} \in E_{ad}} \|C\boldsymbol{y}(0, \boldsymbol{\epsilon}; T) - C\boldsymbol{y}(A_\ell, \boldsymbol{\epsilon}; T)\|_2^2 .$$

Swap A_1 and A_{ℓ_1} in $\mathcal{A}$, and set $k = 1$ and $A^{(0)}(\boldsymbol{\beta}^{\ell_1}) = 0$.

while $k \leq K - 1$ and $\left\|C\boldsymbol{y}(A^{(k-1)}(\boldsymbol{\beta}^{\ell_k}), \boldsymbol{\epsilon}^k; T) - C\boldsymbol{y}(A_k, \boldsymbol{\epsilon}^k; T)\right\|_2^2 > \mathrm{tol}_1$ **do**

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

    Orthogonalize all elements $(A_{k+1}, \dots, A_K)$ with respect to $(A_1, \dots, A_k)$, remove any that are linearly dependent and update K accordingly.

    Fitting step: Find $(\boldsymbol{\beta}_j^\ell)_{j=1,\dots,k}$ that solve the problem

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^k} \sum_{m=1}^{k} \left\|C\boldsymbol{y}(A^{(k)}(\boldsymbol{\beta}), \boldsymbol{\epsilon}^m; T) - C\boldsymbol{y}(A_\ell, \boldsymbol{\epsilon}^m; T)\right\|_2^2,$$

    and set $f_\ell = \sum_{m=1}^k \left\|C\boldsymbol{y}(A^{(k)}(\boldsymbol{\beta}^\ell), \boldsymbol{\epsilon}^m; T) - C\boldsymbol{y}(A_\ell, \boldsymbol{\epsilon}^m; T)\right\|_2^2$.

  end for

  if $\max_{\ell=k+1,\dots,K} f_\ell > \mathrm{tol}_2$ **then** Set $\ell_{k+1} = \arg\max_{\ell=k+1,\dots,K} f_\ell$ **else**

  Extended splitting step: Find $\boldsymbol{\epsilon}^{k+1}$ and ℓ_{k+1} that solve the problem

$$\max_{\ell \in \{k+1,\dots,K\}} \max_{\boldsymbol{\epsilon} \in E_{ad}} \left\|C\boldsymbol{y}(A^{(k)}(\boldsymbol{\beta}^\ell), \boldsymbol{\epsilon}; T) - C\boldsymbol{y}(A_\ell, \boldsymbol{\epsilon}; T)\right\|_2^2 .$$

  end if

  Swap A_{k+1} and $A_{\ell_{k+1}}$ in $\mathcal{A}$, and set $k = k+1$.

end while

example, where strict convexity can be proven under certain assumptions, is given by the linear-quadratic setting discussed in [13, sect. 4]. In what follows, we discuss global convergence (convexity) for a specific case of a bilinear system.

The goal is to reconstruct a matrix $A_\star = \boldsymbol{\alpha}_{\star,1} A_1 + \boldsymbol{\alpha}_{\star,2} A_2 \in \mathbb{R}^{2\times 2}$ of the system

$$\dot{\boldsymbol{y}}(t) = \epsilon(t) A_\star \boldsymbol{y}(t), \quad t \in [0, T], \quad \boldsymbol{y}(0) = \boldsymbol{y}^0, \tag{7.12}$$

where $T = 1$. Assume that $\boldsymbol{y}^0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\boldsymbol{\alpha}_\star = \boldsymbol{\alpha}_\circ = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. We also assume that we can observe the full state $\boldsymbol{y}(T)$, i.e., $C = I \in \mathbb{R}^{2\times 2}$. We consider that E_{ad} is a set of constant controls that are bounded in absolute value by 1, i.e., $\epsilon(t) \equiv \epsilon \in U_{ad} = [-1, 1]$. Thus, the solution to (7.12) can be written explicitly as $\boldsymbol{y}(\boldsymbol{\alpha}, u; t) = e^{t\epsilon A(\boldsymbol{\alpha})} \boldsymbol{y}^0$. Hence, the final reconstruction problem (3.3) reads as

$$\min_{\boldsymbol{\alpha} \in \mathbb{R}^2} J(\boldsymbol{\alpha}) := \sum_{m=1}^{2} \|e^{(\epsilon_m(\boldsymbol{\alpha}_1 A_1 + \boldsymbol{\alpha}_2 A_2))} \boldsymbol{y}^0 - e^{\epsilon_m A_\star} \boldsymbol{y}^0\|_2^2. \tag{7.13}$$

Since $A_\star$ is zero, we have $e^{\epsilon A_\star} \boldsymbol{y}^0 = \boldsymbol{y}^0$. Using the expressions for A_1 and A_2 and the result [9, Cor. 2.3], one gets $e^{\epsilon(\boldsymbol{\alpha}_1 A_1 + \boldsymbol{\alpha}_2 A_2)} \boldsymbol{y}^0 = \frac{1}{\|\boldsymbol{\alpha}\|_2} \begin{bmatrix} \|\boldsymbol{\alpha}\|_2 \cosh(\|\boldsymbol{\alpha}\|_2 \epsilon) + \boldsymbol{\alpha}_1 \sinh(\|\boldsymbol{\alpha}\|_2 \epsilon) \\ \boldsymbol{\alpha}_2 \sinh(\|\boldsymbol{\alpha}\|_2 \epsilon) \end{bmatrix}$. Thus, by a direct calculation, we can write the cost function J in (7.13) as

$$J(\boldsymbol{\alpha}) = 2\cosh(\|\boldsymbol{\alpha}\|_2)^2 + 2\sinh(\|\boldsymbol{\alpha}\|_2)^2 - 4\cosh(\|\boldsymbol{\alpha}\|_2) + 2 =: \widetilde{J}(x = \|\boldsymbol{\alpha}\|_2).$$

Thus, $\widetilde{J}''(x) = 8\cosh(x)^2 + 8\sinh(x)^2 - 4\cosh(x)$. Since $\cosh(x) \geq 1$ and $\sinh(x) \geq 0$ for $x \geq 0$, it holds that $\widetilde{J}''(x) > 0$ for $x \geq 0$. Therefore, J is strictly convex. Thus, a sufficient condition for global convergence of GN is that our proposed methods select the two controls $\epsilon_1 = -1$ and $\epsilon_2 = 1$ (or $\epsilon_1 = 1$ and $\epsilon_2 = -1$). Notice that, if only one control is used among 1 and -1, then J is not convex.

Let us now study the GR algorithm. The initialization problem is given by $\max_{\epsilon\in[-1,1]} J_I(\epsilon) := \|\boldsymbol{y}^0 - e^{\epsilon A_1}\boldsymbol{y}^0\|_2^2$. One can show that $J_I(\epsilon) = (\exp(\epsilon) - 1)^2$, which attains its unique global maximum at $\epsilon = 1$. Thus, assuming that the splitting step solver converges to the unique global maximizer, GR will choose $\epsilon_1 = 1$. Let us now consider the fitting step problem $\min_{\alpha\in\mathbb{R}} J_F(\alpha) := \|e^{\epsilon_1\alpha A_1}\boldsymbol{y}^0 - e^{\epsilon_1 A_2}\boldsymbol{y}^0\|_2^2$. Direct calculations show that $J_F'(\log(\cosh(1))) = 0$ and $J_F'(\alpha)$ is negative for $\alpha < \log(\cosh(1))$ and positive for $\alpha > \log(\cosh(1))$. Since it is continuous in 0, J_F has a global minimum in $\alpha_1 := \log(\cosh(1))$. Now, consider the splitting step: $\max_{\epsilon\in[-1,1]} J_S(\epsilon) := \|e^{\epsilon\alpha_1 A_1}\boldsymbol{y}^0 - e^{\epsilon A_2}\boldsymbol{y}^0\|_2^2$. Proceeding as before, one can get this has a unique global maximizer given by $\epsilon = -1$. Thus, assuming again that the splitting step solver converges to the unique global maximizer, GR will choose $\epsilon_2 = 1$.

In conclusion, GR chooses $\epsilon_1 = 1$ and $\epsilon_2 = -1$, making the final reconstruction problem (7.13) strictly convex and hence leading to global convergence of GN.

The above example is very instructive. On the one hand, as we are going to see, it can be used to reveal some properties (and weaknesses) of GR. On the other hand, it is straightforward to carry out similar calculations to compare GR, LGR and OGR. These calculations (omitted here for brevity) lead to the following remarks.

The above results about GR for the specific case (7.12) depend heavily on the order of the basis elements A_1 and A_2. If one repeats the calculation with the order of A_1 and A_2 reversed, then the initialization problem has two global maximizers at $\epsilon = \pm 1$. Independently of which of these is chosen, the fitting step has two global minimizers. Each of them leads to a splitting step problem having two global maximizers located again at $\epsilon = \pm 1$, independently of the chosen value in the initialization step. Thus, even if one assumes that the optimizations solvers used for intialization, fitting and splitting steps converge to global (maximum/minimum) points, it can happen that the control computed at the splitting step is equal to the one obtained at the initialization step, making the cost of the final identification problem not convex. The importance of the ordering of the elements in $\mathcal{A}$ was already discussed in detail in [13] for the reconstruction of the control matrix B in case of linear systems. This was exactly the reason for designing OGR. In fact, if one repeats the above calculation for OGR, the two controls $\epsilon = 1$ and $\epsilon = -1$ are always obtained, independently of the ordering of A_1 and A_2. Finally, for LGR one can show that, independently of the ordering of A_1 and A_2, it can happen that the controls computed at the initialization and splitting steps are equal, leading to a non-convex final identification problem. The reason is that the linearization process does not allow to fully capture the dynamics of the system. This is also apparent in the numerical experiments in the next section.

8. Numerical experiments.

We study efficiency and robustness of the GR and OGR algorithms by direct numerical experiments. In section 8.1, we consider the reconstruction of a drift matrix. In section 8.2, we focus on a multi-spin reconstruction problem. All optimization problems in the GR algorithms are solved by a BFGS descent-direction method. The online identification problem is solved by GN.

8.1. Reconstruction of drift matrices.

We consider system (5.1) with (full rank) randomly generated matrices $A_\star, B, C \in \mathbb{R}^{3\times 3}$. The final time is $T = 1$ and the initial value is $\boldsymbol{y}^0 = [0, 0, 0]^\top$. First, we study the algorithms for system (5.2). This is obtained by linearizing (5.1) around two different $A_\circ$, which are randomly chosen approximations to $A_\star$, one with 1% and the other with 10% relative error,

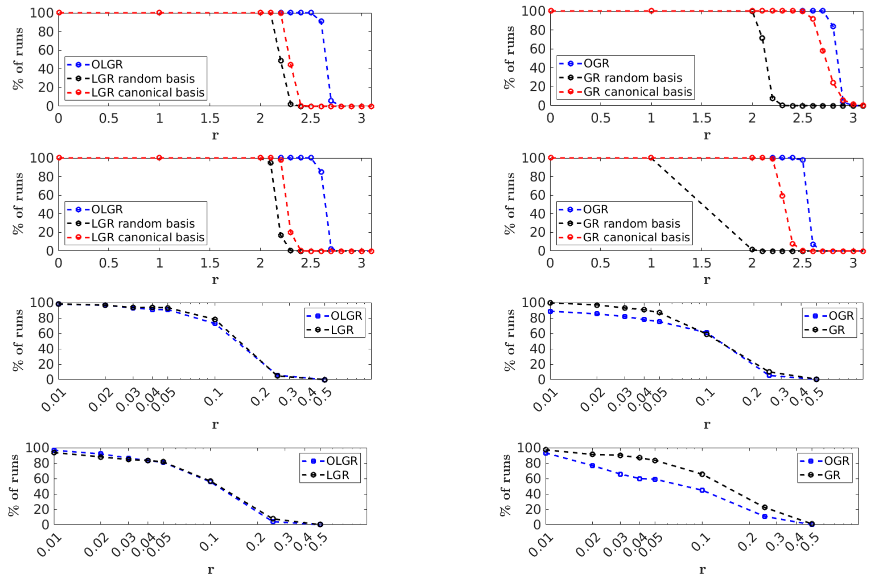

Fig. 8.1: Rows 1 and 2: experiments of section 8.1. Rows 3 and 4: experiments of section 8.2. Percentage of runs that converged to $A_\star$ initialized by randomly chosen vectors on a sphere with radius r, for controls generated by LGR and OLGR for 1% (first and third rows, left) and 10% (second and fourth rows, left) relative error between $A_\star$ and $A_\circ$, and GR and OGR Algorithm 7.1 with (second and fourth rows, right) and without the shift by $A_\circ$ (first and third rows, right).

meaning that, e.g., $\frac{\|A_\star - A_\circ\|_F}{\|A_\star\|_F} = 0.01$ for the one with 1% error, where $\|\cdot\|_F$ is the Frobenius norm. The LGR Algorithm 4.1 is run for two different choices for the basis $\mathcal{A}$: the canonical basis of $\mathbb{R}^{3\times 3}$ and a basis consisting of 9 randomly generated (linearly independent) 3×3 matrices. LGR is also compared with the OLGR Algorithm 7.2, which is run with a set of 18 matrices, namely, the 9 canonical basis elements and the 9 random matrices. The controls generated by the respective algorithms are then used to reconstruct the matrix $A_\star$ by solving the online least-squares problem (3.3) with GN. To test the robustness of the control functions, we consider a nine-dimensional sphere centered in the global minimum $A_\star$ and with given relative radius r, and repeat the minimization for 1000 initialization vectors randomly chosen on this sphere. We then count the percentage of times that GN converges to the global solution $A_\star = A(\boldsymbol{\alpha}_\star)$ up to a tolerance of $Tol = 0.005$ (half of the smallest considered radius), meaning that $\frac{\|A_\star - A(\boldsymbol{\alpha}_{comp})\|_F}{\|A_\star\|_F} \leq Tol$, where $\boldsymbol{\alpha}_{comp}$ denotes the solution computed by GN. Repeating this experiment for different radii of the sphere, we obtain the results reported in the two panels on the left (rows 1 and 2) in Figure 8.1. All control sets make GN capable of reliably reconstructing the global minimum $A_\star$ up to a relative radius $r = 2$, which corresponds to a relative error of 200%. This demonstrates that the choice of the basis is not crucial for fully observable and controllable systems. However, OLGR is able to reduce the number of controls down to 3 and still outperforms any set of 9 controls generated by LGR, while staying reliable up to a relative error of 250%. Thus, OLGR is able to compute a better basis and to omit unnecessary controls.

Next, we repeat the same experiments for the GR Algorithm 7.1. However, we replace the case for the approximation $A_\circ$ with a relative error of 1% by $A_\circ = 0$. This effectively removes the shift and makes the algorithm independent of the choice of

$A_\circ$, which is the version of the algorithm considered in [13,29]. We obtain the results shown in the two panels on the right in Figure 8.1. The performance of the control sets is similar to the ones for the linearized system, with an increase in performance for the GR algorithm with the canonical basis, without the shift by $A_\circ$, and a decrease in performance for GR with the random basis and an $A_\circ$ that has a 10% relative error with respect to $A_\star$. As in the linearized setting, OGR is able to reduce the number of controls to 3 and still outperforms any set of 9 controls generated by LGR.

8.2. Multi-spin (/-qubit) reconstruction problem. Let us consider the case of a 3-qubit system. We use the 64-dimensional real representation of the Liouville master equation (compare, e.g., [11, sect. 2.12.1]):

$$\dot{\boldsymbol{y}}(t) = \Big(A_\star + u(t)\sum_{n=1}^{6} B_n\Big)\boldsymbol{y}(t), \quad t \in [0,T], \quad \boldsymbol{y}(0) = \boldsymbol{y}^0,$$

where $A_\star = 2\pi \sum_{i=1}^{3} \omega_{\star,i}\widehat{A}_i + 2\pi \sum_{j=1}^{2} J_{\star,j}\widetilde{A}_j \in \mathbb{R}^{64\times 64}$, $B_n = 2\pi \widehat{B}_{\frac{n+1}{2}}$ for n odd and $B_n = 2\pi\widetilde{B}_{\frac{n}{2}}$ for n even. Here, the matrices $\widehat{A}_i$ and $\widetilde{A}_j$ correspond to the free evolution and the coupling of the spins, and the matrices $\widehat{B}_n$ and $\widetilde{B}_n$ correspond to the controlled evolution. The goal is to reconstruct $A_\star$, i.e. the coefficients $\omega_{\star,i}$, $J_{\star,j}$, which we assume to be equal to 1. Thus, we define the set $\mathcal{A} = \{A_1, \dots, A_5\}$ as the union of the matrices $\widehat{A}_i$ and $\widetilde{A}_j$. Experimentally, we assume that the final state $\boldsymbol{y}(T)$ is measured against a fixed state $\boldsymbol{y}^1$, meaning that is $C = (\boldsymbol{y}^1)^\top$. The vector $\boldsymbol{y}^1$ consists of zeros except for the second, fifth and 17th entries which are equal to one, while the initial state $\boldsymbol{y}^0$ is equal to one on the fourth, 13th and 49th entries.[9] We perform the same experiments as in section 8.1 with the difference that GR and OGR are run with the same basis $\mathcal{A}$. The results are reported in Figure 8.1 (rows 3 and 4). All algorithms show essentially a similar performance.

9. Conclusion.

We developed and analyzed greedy reconstruction algorithms. We tackled the case of nonlinear problems consisting in the reconstruction of drift operators in linear and bilinear systems. We proved that the controls obtained by GR on linearized systems lead to the local convergence of GN applied to the online nonlinear identification problem. These results were extended to the controls obtained on the fully nonlinear system (without linearization) where a local convergence result was obtained. Future work could focus on the development of theoretical results and numerical strategies for the calculation (or improvement) of the initial value $A(\boldsymbol{\alpha}_\circ)$, for example by maximizing the rank of the observability/controllability matrices, or by refining iteratively $A(\boldsymbol{\alpha}_\circ)$ using new experimental data.

Acknowledgements. G. Ciaramella is member of GNCS Indam. The present research is part of the activities of "Dipartimento di Eccellenza 2023-2027". S. Buchwald is funded by the DFG via the CRC SFB1432, Project-ID 425217212. J. Salomon acknowledges support from ANR-19-CE46-0013/A-HKBU203/19.

REFERENCES

[1] A. Alexanderian and et al., *A fast and scalable method for A-optimal design of experiments for infinite-dimensional Bayesian nonlinear inverse problems*, SIAM J. Sci. Comput., 38 (2016), pp. A243–A272.

[2] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: a Hitchhiker's Guide*, Springer, 2006.

[9] All matrices $\widehat{A}_i$, $\widetilde{A}_j$, $\widehat{B}_n$ and $\widetilde{B}_n$ and the vectors $\boldsymbol{y}^1$ and $\boldsymbol{y}^0$ are generated from the codes of [15].

[3] P. J. Antsaklis and A. N. Michel, *Linear Systems*, Birkhäuser Basel, 2006.
[4] A. C. Atkinson, *Optimum Experimental Design*, Springer Berlin Heidelberg, 2011.
[5] L. Baudouin and A. Mercado, *An inverse problem for Schrödinger equations with discontinuous main coefficient*, Appl. Anal., 87 (2008), pp. 1145–1165.
[6] F. Bauer and C. Fike, *Norms and exclusion theorems*, Numer. Math., 2 (1960), pp. 137–141.
[7] I. Bauer, H. G. Bock, S. Körkel, and J. P. Schlöder, *Numerical methods for optimum experimental design in dae systems*, J. Comput. Appl. Math., 120 (2000), pp. 1 – 25.
[8] C. Berge, *Topological spaces: including a treatment of multi-valued functions, vector spaces and convexity; Translated by E.M. Patterson*, Oliver & Boyd Edinburgh, 1963.
[9] D. Bernstein and W. So, *Some explicit formulas for the matrix exponential*, IEEE Transactions on Automatic Control, 38 (1993), pp. 1228–1232, https://doi.org/10.1109/9.233156.
[10] S. Bonnabel, M. Mirrahimi, and P. Rouchon, *Observer-based Hamiltonian identification for quantum systems*, Automatica, 45 (2009), pp. 1144 – 1155.
[11] A. Borzì, G. Ciaramella, and M. Sprengel, *Formulation and Numerical Solution of Quantum Control Problems*, SIAM, Philadelphia, PA, 2017.
[12] L. Bris and et al., *Hamiltonian identification for quantum systems: Well posedness and numerical approaches*, ESAIM: Contr. Optim. Ca., 13 (2007), pp. 378–395.
[13] S. Buchwald, G. Ciaramella, and J. Salomon, *Analysis of a greedy reconstruction algorithm*, SIAM J. Control Optim., 59 (2021), pp. 4511–4537.
[14] S. Buchwald, G. Ciaramella, J. Salomon, and D. Sugny, *Greedy reconstruction algorithm for the identification of spin distribution*, Phys. Rev. A, 104 (2021), p. 063112.
[15] G. Ciaramella and A. Borzì, *SKRYN: A fast semismooth-Krylov–Newton method for controlling Ising spin systems*, Computer Physics Communications, 190 (2015), pp. 213–223.
[16] P. G. Ciarlet, *Linear and Nonlinear Functional Analysis with Applications*, Applied mathematics, SIAM, Philadelphia, PA, 2013.
[17] J. Coron, *Control and Nonlinearity*, Math. Surveys Monogr., AMS, 2007.
[18] B. De Schutter, *Minimal state-space realization in linear system theory: An overview*, J. Comput. Appl. Math., 121 (2000), pp. 331–354.
[19] A. Donovan and H. Rabitz, *Exploring the Hamiltonian inversion landscape*, Phys. Chem., 16 (2014), pp. 15615–15622.
[20] Y. Fu and G. Turinici, *Quantum Hamiltonian and dipole moment identification in presence of large control perturbations*, ESAIM: Contr. Optim. Ca., 23 (2017), pp. 1129–1143.
[21] J. M. Geremia and H. Rabitz, *Global, nonlinear algorithm for inverting quantum-mechanical observations*, Phys. Rev. A, 64 (2001), p. 022710.
[22] J. M. Geremia and H. Rabitz, *Optimal Hamiltonian identification: The synthesis of quantum optimal control and quantum inversion*, J. Chem. Phys., 118 (2003), pp. 5369–5382.
[23] J. M. Geremia, W. Zhu, and H. Rabitz, *Incorporating physical implementation concerns into closed loop quantum control experiments*, J. Chem. Phys., 113 (2000), pp. 10841–10848.
[24] S. Glaser and et al., *Training Schrödinger's cat: quantum optimal control*, Eur. Phys. J. D, 69 (2015), p. 279.
[25] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Press, 1991.
[26] J. N. Juang and R. S. Pappa, *An eigensystem realization algorithm for modal parameter identification and model reduction*, J. Guid. Control Dynam., 8 (1985), pp. 620–627.
[27] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, De Gruyter, Berlin, New York, 2008.
[28] C. T. Kelley, *Iterative Methods for Optimization*, SIAM, Philadelphia, 1999.
[29] Y. Maday and J. Salomon, *A greedy algorithm for the identification of quantum systems*, in Proceedings of the 48th IEEE Conference on Decision and Control, , 2009, pp. 375–379.
[30] W. Rudin, *Real and Complex Analysis, 3rd Ed.*, McGraw-Hill, Inc., USA, 1987.
[31] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems (2Nd Ed.)*, Springer-Verlag, Berlin, Heidelberg, 1998.
[32] Y. Wang and et al., *A quantum Hamiltonian identification algorithm: Computational complexity and error analysis*, IEEE Trans. Autom. Control, 63 (2018), pp. 1388–1403.
[33] E. F. Whittlesey, *Analytic functions in Banach spaces*, Proceedings of the American Mathematical Society, 16 (1965), pp. 1077–1083.
[34] S. Xue, R. Wu, D. Li, and M. Jiang, *A gradient algorithm for Hamiltonian identification of open quantum systems*, Phys. Rev. A, 103 (2021), p. 022604.
[35] J. Zhang and M. Sarovar, *Quantum Hamiltonian identification from measurement time traces*, Phys. Rev. Lett., 113 (2014), p. 080401.
[36] W. Zhou, S. Schirmer, E. Gong, H. Xie, and M. Zhang, *Identification of Markovian open system dynamics for qubit systems*, Chinese Sci. Bull., 57 (2012), pp. 2242–2246.