MOX-Report No. 09/2023

Gauss-Newton oriented greedy algorithms for the reconstruction of operators in nonlinear dynamics

Buchwald, S.; Ciaramella, G.; Salomon, J.

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

mox-dmat@polimi.it https://mox.polimi.it

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 on 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 show that the controls obtained with the greedy algorithm on the corresponding linearized system 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 obtained. The main convergence results are obtained for the reconstruction of drift operators in linear and bilinear dynamical systems.

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 reconstruction of nonlinear operators in controlled 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 that can be 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 parameters or operators. Thus, it is natural to look for a set of such input forces that allows one to generate good data allowing the best possible reconstruction. This is a typical case in the field Hamiltonian identification in quantum mechanics [5, 9, 17–21, 29, 33, 34, 36–39], or in engineering in the context of state space realization [16, 22, 25, 32] 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 [30] for Hamiltonian identification problems, further developed and analyzed in [11], and later adapted to the identification of probability distributions for parameters in the context of quantum systems in [13]. This approach decomposes the identification process into offline phase, where the control functions are computed by a GR algorithm, and online phase, where the controls are used to generate experimental data to be used in an inverse problem for the final reconstruct of the unknown operator. In [11], a 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

*Universität Konstanz, Germany (simon.buchwald@uni-konstanz.de).

†MOX, Dipartimento di Matematica, Politecnico di Milano (gabriele.ciaramella@polimi.it).

‡INRIA Paris, France (julien.salomon@inria.fr).

more efficient and robust numerical variant of the standard greedy reconstruction algorithm. It was then shown in [13] 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., [10, 23, 28].

The goal of this paper is to further develop the work [11] by considering nonlinear systems, and to relate the greedy-reconstruction procedure to the Gauss-Newton method (GN), which is one of the most famous methods for solving inverse problems [26]. In particular, we assume that the inverse problem in the online phase is solved by GN, and study the effect of the control functions generated by GR algorithms on the convergence of GN. This is achieved in two steps, which represent the main novelties of this manuscript.

First, we introduce a new greedy-type reconstruction approach. In particular, rather than applying GR directly to the nonlinear identification problem, we use it on its linearization. This corresponds to using GR for designing control functions that make the GN matrix, namely the Jacobian of the nonlinear residual, full rank in a neighborhood of the solution, which is a sufficient condition for local convergence of GN. We refer to this strategy as linearized greedy reconstruction algorithm (LGR), and provide a corresponding detailed analysis 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 nonlinear problems, since the unknown operators act on the states of the systems. Notice that the analysis that we are going to presented for the drift matrix is also valid in the case of the reconstruction of the control matrix in a linear input/output systems, as considered in [11, Section 5]. Thus, this part of the present work is a substantial extension of the results of [11].

The second novelty of this work is to provide a first analysis of the original GR algorithm applied to nonlinear systems. This is achieved by relating the behavior of GR (applied to the original nonlinear problem) 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 two GR and LGR approaches are compared by direct numerical experiments. These show that GR and LGR are comparable when working locally near the solution. However, the 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 how GN can be used to solve general reconstruction problems. In order to guarantee convergence of GN, the LGR algorithm is introduced in Section 4. In sections 5 and 6, we present analyses of LGR for the reconstruction of linear drift matrices in linear systems and an Hamiltonian matrix in bilinear systems, respectively. Section 7 focuses on GR for nonlinear problems, and a corresponding analysis is provided in section 7.1. Within section 7.2, we recall and extend an optimized greedy reconstruction (OGR) algorithm introduced in [11]. The LGR, GR and OGR algorithms are then tested numerically in section 8. Finally, our conclusions are drawn in Section 9.

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

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

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, we denote by $\mathcal{B}_r^N(x) \subset \mathbb{R}^N$ the N-dimensional ball with radius $r > 0$ and center $x \in \mathbb{R}^N$.

3. Gauss-Newton method (GN) for reconstruction problems. Consider a state $\boldsymbol{y}(t) \in \mathbb{R}^N$, $N \in \mathbb{N}$, whose time evolution is governed by 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 $\boldsymbol{y}^0 \in \mathbb{R}^N$ is the initial state and $\boldsymbol{\epsilon} \in E_{ad}$ denotes a control function belonging to E_{ad}, a non-empty and weakly compact subset of some Hilbert space of control functions from $[0, T]$ to $\mathbb{R}^M$, $M \in \mathbb{N}$ (e.g., $E_{ad} \subset L^2(0, T; \mathbb{R}^M)$). 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

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

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

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

GN is a typical iterative strategy to solve (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

$$R_m(\boldsymbol{\alpha}) := C\boldsymbol{y}(A(\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; T) - \boldsymbol{\varphi}_{data}^\star(\boldsymbol{\epsilon}^m), \tag{3.4}$$

for $m \in \{1, \dots, K\}$. Thus, the identification problem (3.3) is equivalent to

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^K} \sum_{m=1}^K f_m(\boldsymbol{\alpha}). \tag{3.5}$$

Given an iterate $\boldsymbol{\alpha}_c$, GN computes the new iterate by solving a problem of the form

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^K} \sum_{m=1}^K \|R_m'(\boldsymbol{\alpha}_c)(\boldsymbol{\alpha} - \boldsymbol{\alpha}_c) - R_m(\boldsymbol{\alpha}_c)\|_2^2, \tag{3.6}$$

where $R'_m(\boldsymbol{\alpha}_c) \in \mathbb{R}^{P\times K}$ denotes the Jacobian of R_m at $\boldsymbol{\alpha}_c \in \mathbb{R}^K$. The first-order optimality condition of (3.6) is

$$\sum_{m=1}^{K} \Big(R'_m(\boldsymbol{\alpha}_c)^\top R'_m(\boldsymbol{\alpha}_c)\Big)\boldsymbol{\alpha} = \sum_{m=1}^{K} R'_m(\boldsymbol{\alpha}_c)^\top R_m(\boldsymbol{\alpha}_c), \tag{3.7}$$

where $\sum_{m=1}^{K} R'_m(\boldsymbol{\alpha}_c)^\top R'_m(\boldsymbol{\alpha}_c) =: \widehat{W}_c \in \mathbb{R}^{K\times K}$ is symmetric PSD. Now, we recall the following convergence result from [27, Theorem 2.4.1] (for a proof see also the supplementary material [12]).

LEMMA 3.1 (local convergence of GN). *Consider a problem of the form (3.5). Let $\boldsymbol{\alpha}_\star$ be a minimizer of (3.5) 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 (3.6)-(3.7) being uniquely solvable. Using (3.4), we can write (3.6) more explicitly. For a direction $\delta\boldsymbol{\alpha} \in \mathbb{R}^K$, we have

$$R'_m(\boldsymbol{\alpha}_c)(\delta\boldsymbol{\alpha}) = C\delta\boldsymbol{y}_c(A(\delta\boldsymbol{\alpha}), \boldsymbol{\epsilon}^m; T), \tag{3.8}$$

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

$$\begin{cases} \dot{\delta\boldsymbol{y}}_c = \partial_{\boldsymbol{y}} f(A(\boldsymbol{\alpha}_c), \boldsymbol{y}_c, \boldsymbol{\epsilon})\delta\boldsymbol{y}_c + \sum_{j=1}^{K} \delta\boldsymbol{\alpha}_j \Big(\partial_A f(A(\boldsymbol{\alpha}_c), \boldsymbol{y}_c, \boldsymbol{\epsilon})(A_j)\Big), \ \delta\boldsymbol{y}_c(0) = 0, \\ \dot{\boldsymbol{y}}_c = f(A(\boldsymbol{\alpha}_c), \boldsymbol{y}_c, \boldsymbol{\epsilon}), \quad \boldsymbol{y}_c(0) = \boldsymbol{y}^0. \end{cases} \tag{3.9}$$

Hence, problem (3.6) can be written as

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^K} \sum_{m=1}^{K} \|C\delta\boldsymbol{y}_c(A(\boldsymbol{\alpha}-\boldsymbol{\alpha}_c), \boldsymbol{\epsilon}^m; T) - R_m(\boldsymbol{\alpha}_c)\|_2^2. \tag{3.10}$$

Notice that the vectors $R_m(\boldsymbol{\alpha}_c) \in \mathbb{R}^P$ are independent of $\boldsymbol{\alpha}$ and can therefore be considered as fixed data when solving (3.10). Now, we recall that the GR algorithm, introduced in [30] and further analyzed in [11], was designed specifically to generate control functions $(\boldsymbol{\epsilon}^m)_{m=1}^K$ that make problems of the form (3.10) uniquely solvable.

4. A linearized GR algorithm (LGR).

Let us assume to be provided with an initialization vector $\boldsymbol{\alpha}_\circ$ for GN that is sufficiently close to $\boldsymbol{\alpha}_\star$. Further, let $\delta\boldsymbol{y}_\circ(A(\boldsymbol{\alpha} - \boldsymbol{\alpha}_\circ), \boldsymbol{\epsilon}^m; T)$ denote solution at time T to

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

The goal is to generate control functions $(\boldsymbol{\epsilon}^m)_{m=1}^K$ such that (3.10) in $\boldsymbol{\alpha}_\circ$, that is

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

Algorithm 4.1 Linearized Greedy Reconstruction Algorithm (LGR)

Require: A set of linearly independent operators $\mathcal{A} = \{A_1, \dots, A_K\}$. Recall that $\delta\boldsymbol{y}_\circ(A, \boldsymbol{\epsilon}; T)$ solves (4.1).

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

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

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

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

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

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

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

end for

is uniquely solvable. Then, in Section 5.2, we show that if (4.2) is uniquely solvable, the same holds for (3.10) at all iterates $\boldsymbol{\alpha}_c$ of GN.

The set $(\boldsymbol{\epsilon}^m)_{m=1}^K$ is computed by the LGR Algorithm 4.1, which is the original GR algorithm from [30] applied to (4.2). 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.

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.

Proof. The proof is similar to the ones of [11, Lemma 5.12]. For an arbitrary $k \in \{0, \ldots, K-1\}$ let $\boldsymbol{v} \in \mathbb{R}^{k+1}$ and $A(\boldsymbol{v}) = \sum_{j=1}^{k+1} \boldsymbol{v}_j A_j$. We have $\|C\delta \boldsymbol{y}_\circ(A(\boldsymbol{v}), \boldsymbol{\epsilon}; T)\|_2^2 = \langle \boldsymbol{v}, [W_\circ(\boldsymbol{\epsilon})]_{[1:k+1,1:k+1]}\boldsymbol{v}\rangle$. Recalling that $\delta\boldsymbol{y}_\circ(A(\boldsymbol{v}), \boldsymbol{\epsilon}; T) = \sum_{j=1}^{k+1} \boldsymbol{v}_j \delta\boldsymbol{y}_\circ(A_j, \boldsymbol{\epsilon}; T)$, we obtain the equivalence between (4.7), (4.9), and (4.3), (4.5) for suitable k and $\boldsymbol{v}$. For the equivalence between (4.8) and (4.4), notice that for $\boldsymbol{v} = [\boldsymbol{\beta}^\top, -1]^\top \in \mathbb{R}^{k+1}$ and any $W \in \mathbb{R}^{(k+1)\times(k+1)}$ we have $\langle \boldsymbol{v}, W\boldsymbol{v}\rangle = \langle \boldsymbol{\beta}, [W]_{[1:k,1:k]}\boldsymbol{\beta}\rangle - 2\langle [W]_{[1:k,k+1]}, \boldsymbol{\beta}\rangle + [W]_{k+1,k+1}$. The well-posedness of the three problems follows by standard arguments; see, e.g., [11, Lemma 5.2]. ∎

The matrix representation given in Lemma 4.1 allows us to nicely describe the mathematical mechanism behind Algorithm 4.1 (see also [11, section 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 [11, Lemma 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 (3.9), as shown in sections 5 and 6. We conclude this section with a remark that is useful hereafter.

Remark 4.3. The GN matrix $\widehat{W}_\star := \sum_{m=1}^K R_m'(\boldsymbol{\alpha}_\star)^\top R_m'(\boldsymbol{\alpha}_\star) \in \mathbb{R}^{K\times K}$ can be written as $[\widehat{W}_\star]_{i,j} = \sum_{m=1}^K \langle C\delta\boldsymbol{y}_\star(A_i, \boldsymbol{\epsilon}^m; T), C\delta\boldsymbol{y}_\star(A_j, \boldsymbol{\epsilon}^m; T)\rangle$ for $i,j \in \{1,\ldots,K\}$, where $\delta\boldsymbol{y}_\star(A_i, \boldsymbol{\epsilon}; T)$ denotes the solution at time T of

$$\begin{cases} \delta\dot{\boldsymbol{y}}_\star = \partial_{\boldsymbol{y}} f(A(\boldsymbol{\alpha}_\star), \boldsymbol{y}_\star, \boldsymbol{\epsilon})\delta\boldsymbol{y}_\star + \Big(\partial_A f(A(\boldsymbol{\alpha}_\star), \boldsymbol{y}_\star, \boldsymbol{\epsilon})(A_i)\Big), \quad \delta\boldsymbol{y}_\star(0) = 0, \\ \dot{\boldsymbol{y}}_\star = f(A(\boldsymbol{\alpha}_\star), \boldsymbol{y}_\star, \boldsymbol{\epsilon}), \quad \boldsymbol{y}(0) = \boldsymbol{y}^0. \end{cases}$$

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 $\boldsymbol{\epsilon} \in E_{ad}$ denotes a control function belonging to E_{ad}, a nonempty and weakly compact subset of $L^2(0,T;\mathbb{R}^M)$ that contains $\boldsymbol{\epsilon} \equiv 0$ as an interior point.[1]

[1]This hypothesis is used in our analysis and is a reasonable assumption, since it is, for example, satisfied for standard box constraints, which are quite often used in the applications.

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 on observability and controllability of the considered linear system. Let us recall these 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., [32, Theorem 3, Theorem 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.*

In Section 5.1, we analyze Algorithm 4.1 in the case of fully observable and controllable systems (namely, $\text{rank}(\mathcal{O}_N(C, A(\boldsymbol{\alpha}_\circ))) = \text{rank}(\mathcal{C}_N(A(\boldsymbol{\alpha}_\circ), B)) = N$). However, similar to [11, Section 5.3], one can also formulate the following results for non-fully observable and controllable systems, if appropriate matrices $A_1, \dots, A_K$ are chosen. For further details, we refer the reader to the supplementary material [12].

Notice that the analysis that we are going to presented is also valid in the case of the reconstruction of a linear control matrix considered in [11, Section 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 [11].

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 $\mathcal{R} := \text{rank}(\mathcal{O}_N^\circ) \cdot \text{rank}(\mathcal{C}_N^\circ) = N^2$. 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, Theorem 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 $\boldsymbol{w} \in \mathrm{im}\Big(\mathcal{C}_N(A, B)\Big)$. 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., [15].

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 $\mathrm{rank}(\mathcal{O}_N^\circ) = \mathrm{rank}(\mathcal{C}_N^\circ) = N$, 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 $\delta\dot{\boldsymbol{y}}_\circ = A(\boldsymbol{\alpha}_\circ)\delta\boldsymbol{y}_\circ + \widetilde{A}\boldsymbol{y}^\circ$, 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. To prove the result, 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 $\boldsymbol{w} \in \mathrm{im}\big(\mathcal{C}_N^\circ\big)$. 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

$$\boldsymbol{g} : \mathbb{R} \to \mathbb{R}^p, s \mapsto Ce^{(T-s)A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\int_0^s e^{(s-\tau)A(\boldsymbol{\alpha}_\circ)}B\widehat{\boldsymbol{\epsilon}}(\tau)d\tau$$

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

$$\widetilde{\boldsymbol{\epsilon}}(s) := \begin{cases} 0, & 0 \leq s < T - t_1, \\ \widehat{\boldsymbol{\epsilon}}(s - t_1), & T - t_1 \leq s \leq T, \end{cases}$$

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

$$\begin{aligned} C\delta\boldsymbol{y}_\circ(\widetilde{A}, \widetilde{\boldsymbol{\epsilon}}; T) &= \int_{T-t_1}^T Ce^{(T-s)A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\int_{T-t_1}^s e^{(s-\tau)A(\boldsymbol{\alpha}_\circ)}B\widetilde{\boldsymbol{\epsilon}}(\tau - t_1)d\tau ds \\ &= \int_0^{t_1} Ce^{(t_1-s)A(\boldsymbol{\alpha}_\circ)}\widetilde{A}\int_0^s e^{(s-\tau)A(\boldsymbol{\alpha}_\circ)}B\widehat{\boldsymbol{\epsilon}}(\tau)d\tau ds = \int_0^{t_1} \boldsymbol{g}(s)ds \neq 0. \quad \square \end{aligned}$$

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* $\mathrm{rank}(\mathcal{O}_N^\circ) = \mathrm{rank}(\mathcal{C}_N^\circ) = N$*. For $K \leq N^2$, let $\mathcal{A} = \{A_1, \dots, A_K\} \subset \mathbb{R}^{N\times N}$ be a set of linearly independent matrices such that $A(\boldsymbol{\alpha}_\circ) \in \mathrm{span}(\mathcal{A})$, 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

$$[\widehat{W}_\circ^{(k+1)}]_{[1:k+1,1:k+1]} = [\widehat{W}_\circ^{(k)}]_{[1:k+1,1:k+1]} + [W_\circ(\boldsymbol{\epsilon}^{k+1})]_{[1:k+1,1:k+1]}$$

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

$$\langle \boldsymbol{v}, [W_\circ(\boldsymbol{\epsilon}^{k+1})]_{[1:k+1,1:k+1]}\boldsymbol{v}\rangle = \left\| C\delta\boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k) - A_{k+1}, \boldsymbol{\epsilon}; T)\right\|_2^2 > 0.$$

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.□

Notice that Theorem 5.6 does not require any assumption on the matrices $A_1, \dots, A_K$. These can be arbitrarily chosen with the only constraint to be linearly independent. Also the ordering of these matrices does not affect the result of Theorem 5.6. This is, however, different for non-fully observable and controllable systems, i.e. for $\mathcal{R} < N^2$ (see the supplementary material [12]).

Now that we proved that Algorithm 4.1 makes $\widehat{W}_\circ$ PD, the obvious question is whether this is sufficient for the convergence of GN, as described in Lemma 3.1. We answer this question 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 $\boldsymbol{\alpha}_\star$. Indeed, in Section 5.1, we proved that the control functions generated by Algorithm 4.1 make the GN matrix $\widehat{W}_\circ = \widehat{W}(\boldsymbol{\alpha}_\circ)$ PD. Thus, it is sufficient to prove that $\widehat{W}(\boldsymbol{\alpha})$ remains PD in a neighborhood of $\boldsymbol{\alpha}_\circ$ containing $\boldsymbol{\alpha}_\star$. To do so, let us rewrite $\widehat{W}(\boldsymbol{\alpha})$ as

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

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

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

Lemma 5.7 (rank stability). *Consider two natural numbers N_D and M_D with $N_D \geq M_D$, and an arbitrary matrix $D \in \mathbb{R}^{N_D \times M_D}$ with rank $\mathcal{R}_\mathcal{D}$ and (positive) singular values $\sigma_1, \ldots, \sigma_{\mathcal{R}_\mathcal{D}}$ in descending order. Then it holds that*

$$\min_{\widehat{D} \in \mathbb{R}^{N_D \times M_D}} \{ \|\widehat{D}\|_2 \ | \ \mathrm{rank}(D + \widehat{D}) < \mathcal{R}_\mathcal{D} \} = \sigma_{\mathcal{R}_\mathcal{D}}.$$

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 $K = \mathrm{rank}(\widehat{W}(\boldsymbol{\alpha}_\circ)) \leq \mathrm{rank}(\widehat{W}(\widehat{\boldsymbol{\alpha}}))$. Because of (5.3), $\widehat{W}(\widehat{\boldsymbol{\alpha}}) \in \mathbb{R}^{K \times K}$ meaning that $\mathrm{rank}(\widehat{W}(\widehat{\boldsymbol{\alpha}})) = K$. Since $\widehat{W}(\boldsymbol{\alpha})$ is PSD by construction, $\mathrm{rank}(\widehat{W}(\widehat{\boldsymbol{\alpha}})) = K$ 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 the matrices $A(\boldsymbol{\alpha}_\circ) \in \mathbb{R}^{N \times N}$, $B \in \mathbb{R}^{N \times M}$ and $C \in \mathbb{R}^{P \times N}$ satisfy $\mathrm{rank}(\mathcal{O}_N^\circ) \cdot \mathrm{rank}(\mathcal{C}_N^\circ) = N^2$. 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 $\boldsymbol{\alpha}_j = \boldsymbol{\alpha}_{\star,j}$, $j = 1, \ldots, K$.

Proof. Theorem 5.6 guarantees that $\widehat{W}_\circ$ is PD and hence $\widehat{\sigma}_K > 0$. Thus, by Lemma 5.8 there exists $\delta = \delta(\widehat{\sigma}_K) > 0$ such that, for $\boldsymbol{\alpha} \in \mathbb{R}^K$ with $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}_\circ\|_2 < \delta$, the matrix $\widehat{W}(\boldsymbol{\alpha})$ is also PD. 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). Analogously to 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 $m \in \{1, \ldots, K\}$. 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]

$$\widetilde{W}(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha}) := \sum_{m=1}^{K} W(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha},\boldsymbol{\epsilon}^m), \tag{5.7}$$

with $W(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha},\boldsymbol{\epsilon}^m) \in \mathbb{R}^{K\times K}$ *given by*

$$[W(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha},\boldsymbol{\epsilon}^m)]_{i,j} := \langle \boldsymbol{\gamma}_i(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha},\boldsymbol{\epsilon}^m), \boldsymbol{\gamma}_j(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha},\boldsymbol{\epsilon}^m)\rangle, \quad i,j\in\{1,\dots,K\}, \tag{5.8}$$

$$\boldsymbol{\gamma}_j(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha},\boldsymbol{\epsilon}^m) := \int_0^T Ce^{(T-s)A(\boldsymbol{\alpha}_\star)} A_j \boldsymbol{y}(A(\boldsymbol{\alpha}),\boldsymbol{\epsilon}^m; s)ds, \quad j\in\{1,\dots,K\}. \tag{5.9}$$

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 time T is given by

$$\Delta\boldsymbol{y}_m(T) = \int_0^T e^{(T-s)A(\boldsymbol{\alpha}_\star)}\Big[A(\boldsymbol{\alpha}_\star - \boldsymbol{\alpha})\boldsymbol{y}(A(\boldsymbol{\alpha}),\boldsymbol{\epsilon}^m;s)\Big]ds.$$

Thus, recalling $A(\boldsymbol{\alpha}) = \sum_{j=1}^K \boldsymbol{\alpha}_j A_j$, the function $J(\boldsymbol{\alpha})$ can be written as

$$\begin{aligned}J(\boldsymbol{\alpha}) &= \frac{1}{2}\sum_{m=1}^K \left\| \int_0^T Ce^{(T-s)A(\boldsymbol{\alpha}_\star)}\Big(\sum_{j=1}^K(\boldsymbol{\alpha}_{\star,j} - \boldsymbol{\alpha}_j)A_k\Big)\boldsymbol{y}(A(\boldsymbol{\alpha}),\boldsymbol{\epsilon}^m;s)ds\right\|_2^2\\ &\overset{(5.9)}{=} \frac{1}{2}\sum_{m=1}^K\sum_{i=1}^K\sum_{j=1}^K(\boldsymbol{\alpha}_{\star,i} - \boldsymbol{\alpha}_i)(\boldsymbol{\alpha}_{\star,j} - \boldsymbol{\alpha}_j)\langle\boldsymbol{\gamma}_i(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha},\boldsymbol{\epsilon}^m),\boldsymbol{\gamma}_j(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha},\boldsymbol{\epsilon}^m)\rangle\\ &\overset{(5.8)}{=} \frac{1}{2}\langle\boldsymbol{\alpha}_\star - \boldsymbol{\alpha}, \sum_{m=1}^K W(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha},\boldsymbol{\epsilon}^m)(\boldsymbol{\alpha}_\star - \boldsymbol{\alpha})\rangle = \frac{1}{2}\langle\boldsymbol{\alpha}_\star - \boldsymbol{\alpha}, \widetilde{W}(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha})(\boldsymbol{\alpha}_\star - \boldsymbol{\alpha})\rangle. \quad\square\end{aligned}$$

Now, the set of global solutions to problem (5.6) is given by $\mathcal{S}_{global} := \Big\{\boldsymbol{\alpha}\in\mathbb{R}^K : (\boldsymbol{\alpha}_\star - \boldsymbol{\alpha}) \in \ker \widetilde{W}(\boldsymbol{\alpha}_\star,\boldsymbol{\alpha})\Big\}$. 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

[2] Notice that the notations (5.3) and (5.7) are related in the sense that $\widetilde{W}(\boldsymbol{\alpha},\boldsymbol{\alpha}) = \widehat{W}(\boldsymbol{\alpha})$.

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 the estimate $\boldsymbol{\alpha}_\circ$. 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, we obtain that if the matrix $\widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}_\star)$ is PD, then the same is true 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 problem. 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 $\epsilon \in E_{ad} \subset L^2(0, T; \mathbb{R})$ denotes a control function belonging to E_{ad}, a nonempty, closed, convex and bounded subset of $L^2(0, T; \mathbb{R})$ that contains $\epsilon \equiv 0$ as an interior point. 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 $\mathcal{A} = \{A_1, \dots, A_K\} \subset \mathbb{R}^{N\times N}$, $1 \leq K \leq N^2$, and we write $A_\star = \sum_{j=1}^K \boldsymbol{\alpha}_{\star,j} A_j =: A(\boldsymbol{\alpha}_\star)$. Notice that, 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

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

We assume to be provided with a known estimate $\boldsymbol{\alpha}_\circ$ of $\boldsymbol{\alpha}_\star$. For this estimate, we can derive the linearized state equation

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

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 [10, 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.

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

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 the following conditions are satisfied.*

1. *The Lie algebra $L = Lie\{A_\circ, B\} \subset \mathfrak{so}(N)$, generated by the matrices $A_\circ$ and B, has dimension $\frac{N(N-1)}{2}$.*
2. *The final time $T > 0$ is sufficiently large, such that the controllability result from Lemma 6.1 holds.*
3. *There exists $c \in \mathbb{R}$ such that system (6.5) is observable, i.e. the observability matrix $\mathcal{O}_N(C, A_\circ + cB)$ has full rank.*

In addition, let the set of admissible controls $E_{ad} \subset L^2(0, T; \mathbb{R})$ 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 constant $c \in \mathbb{R}$ mentioned above.

Remark 6.3. The analysis presented in the following sections can be applied to the case where the matrix $A = A_\star$ is assumed to be known and $B = B(\boldsymbol{\alpha}) := \sum_{j=1}^K \boldsymbol{\alpha}_j B_j$ is unknown and to be identified. The main differences in the case of the identification of B is that the state equation is linearized around an initial guess $B_\circ$, leading to

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

Assumption 6.2 would be the same, only with A instead of $A_\circ$ and $B_\circ$ instead of B. Notice that, in this case, we also cover Schrödinger-type systems of the form

$$i\dot{\boldsymbol{\psi}}(t) = (H + \epsilon(t)\mu_\star)\boldsymbol{\psi}(t), \ t \in (0, T], \quad \boldsymbol{\psi}(0) = \boldsymbol{\psi}^0,$$

as considered in [30], for Hermitian matrices $H, \mu_\star \in \mathbb{C}^{N \times N}$. This can be seen by writing $\boldsymbol{\psi} = \boldsymbol{\psi}_R + i\boldsymbol{\psi}_I$, $\boldsymbol{\psi}^0 = \boldsymbol{\psi}_R^0 + i\boldsymbol{\psi}_I^0$, $H = H_R + iH_I$ and $\mu_\star = \mu_{\star,R} + i\mu_{\star,I}$, to get

$$\dot{\boldsymbol{y}}(t) = \Bigg(\underbrace{\begin{bmatrix} H_I & H_R \\ -H_R & H_I \end{bmatrix}}_{=:A} + \epsilon(t) \underbrace{\begin{bmatrix} \mu_{\star,I} & \mu_{\star,R} \\ -\mu_{\star,R} & \mu_{\star,I} \end{bmatrix}}_{:=B_\star} \Bigg) \boldsymbol{y}(t), \tag{6.6}$$

for $\boldsymbol{y}(t) := \begin{bmatrix} \boldsymbol{\psi}_R(t) & \boldsymbol{\psi}_I(t) \end{bmatrix}^\top$ and skew-symmetric matrices $A, B_\star \in \mathbb{R}^{N \times N}$ (compare also [10, Section 2.12.2]).

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 6.4 (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 any solution $\widetilde{\epsilon}$ to the problem*

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

satisfies $\left\| C\delta \boldsymbol{y}_\circ(\widetilde{A}, \widetilde{\epsilon}; T) \right\|_2^2 > 0.$

Proof. It is sufficient to show that there exists a control $\widehat{\epsilon}_c \in E_{ad}$ such that $C\delta\boldsymbol{y}_\circ(\widetilde{A}, \widehat{\epsilon}_c; T) \neq 0$ for T sufficiently large. Let us define $\widehat{\epsilon}_c$ as

$$\widehat{\epsilon}_c(s) := \begin{cases} \widehat{\epsilon}(s), & \text{for } 0 \leq s \leq \widehat{t}, \\ c, & \text{for } \widehat{t} < s \leq T, \end{cases}$$

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 and second 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., [31, Theorem 10.18]. Recalling that exponential matrices are always invertible (see, e.g., [24, Theorem 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 6.4 has to be used instead of Lemma 5.5.

THEOREM 6.5 (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 $\mathcal{A} = \{A_1, \dots, A_K\} \subset \mathfrak{so}(N)$ be a set of linearly independent matrices such that $A_\star \in \operatorname{span} \mathcal{A}$, and let $\{\epsilon^1, \dots, \epsilon^K\} \subset E_{ad}$ be controls generated by Algorithm 4.1. Then the GN matrix $\widehat{W}_\circ$, defined in (4.6), is PD.*

[4]This follows directly from the fundamental theorem of calculus.

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 $\boldsymbol{\alpha}$:

$$[\widehat{W}(\boldsymbol{\alpha})]_{i,j} := \sum_{m=1}^{K} \langle C\delta\boldsymbol{y}(\boldsymbol{\alpha}, A_i, \epsilon^m; T), C\delta\boldsymbol{y}(\boldsymbol{\alpha}, A_j, \epsilon^m; T)\rangle, \quad i,j \in \{1,\dots,K\}, \tag{6.7}$$

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

$$\begin{cases} \dot{\delta\boldsymbol{y}}(t) &= (A(\boldsymbol{\alpha}) + \epsilon(t)B)\delta\boldsymbol{y}(t) + \widehat{A}\boldsymbol{y}(t), \quad \delta\boldsymbol{y}(0) = 0, \\ \dot{\boldsymbol{y}}(t) &= (A(\boldsymbol{\alpha}) + \epsilon(t)B)\boldsymbol{y}(t), \quad \boldsymbol{y}(0) = \boldsymbol{y}^0. \end{cases} \tag{6.8}$$

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

LEMMA 6.6 (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 (6.7), 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 (6.8) is continuous in $\boldsymbol{\alpha}$. By [10, 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 $\boldsymbol{\alpha} \mapsto \delta\boldsymbol{y}(\boldsymbol{\alpha}, \widehat{A}, \epsilon; T)$. ∎

Using the result from Lemma 6.6, we can directly prove our main result.

THEOREM 6.7 (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 6.5 guarantees that $\widehat{W}_\circ$ is PD, meaning that $\widehat{\sigma}_K > 0$. Analogously to the proof of Lemma 6.6, one can also show that the functions $R_m(\boldsymbol{\alpha})$, defined in (3.4), are Lipschitz continuously differentiable in $\boldsymbol{\alpha}$ for all $m \in \{1,\dots,K\}$. Thus, the result follows by Lemma 6.6. ∎

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

$$[W(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, \epsilon^m)]_{i,j} := \langle C\delta\boldsymbol{y}_m(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, A_j; T), C\delta\boldsymbol{y}_m(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, A_j; T)\rangle,$$

for $i,j \in \{1,\dots,K\}$ and where $C\delta\boldsymbol{y}_m(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}, A; T)$ is the solution at time T of

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

[5] This result is a special case of the implicit function theorem; see, e.g., [10, Theorem 3.4].

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 $\mathcal{K}_j \subset \mathbb{R}^j$, $j = 1, \dots, K-1$.

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

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

(I unified the notation here regarding the OGR Algorithm 7.2 and Assumption 7.6. Before, the A_1-state was split against the uncontrolled state)

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

3: 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

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

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

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

5: **end for**

The proof of Lemma 6.8 is analogous to the one of Lemma 5.10 (for details see the supplementary material [12]). Notice that the notations in (6.7) and Lemma 6.8 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 $\mathcal{S}_{global} := \left\{ \boldsymbol{\alpha} \in \mathbb{R}^K : (\boldsymbol{\alpha}_\star - \boldsymbol{\alpha}) \in \ker \widetilde{W}(\boldsymbol{\alpha}_\star, \boldsymbol{\alpha}) \right\}$, 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 [11], 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 $\dot{\boldsymbol{y}}(t) = f(A(\boldsymbol{\alpha}^0) + A(\delta\alpha_\star), \boldsymbol{y}(t), \boldsymbol{\epsilon}(t))$ with the goal of reconstructing $A(\delta\alpha_\star) = A_\star - A(\boldsymbol{\alpha}^0)$. 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 large. Finally, the initialization problem (7.1) is different from the initialization (4.3). This is due to results obtained in [11] 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 4.1, it is sufficient to prove two points. First, that the fitting step 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 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 interval. A set-valued 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 exists 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 $G \subset \mathbb{R}^k$ meeting $c(x_0)$ there exists 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 called hemi-continuous if it is u.h.c. and l.h.c.*

Using Definition 7.1, we can recall the Berge maximum theorem [2, Theorem 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 $c(x) := \underset{z\in\phi(x)}{\arg\min}\, J(z; x)$ 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 $c(x_k) \subset \mathcal{B}_\epsilon^k(0)$. ∎

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 $\mathcal{B}_\nu^k(0) \subset \mathcal{K}_k$ and $\arg\min_{\boldsymbol{\beta}\in\overline{\mathcal{B}_\nu^k(0)}} J_k(\boldsymbol{\beta}; 0) = \{0\}$.*

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 $\{0\} = \arg\min_{\boldsymbol{\beta}\in\overline{\mathcal{B}_\nu^k(0)}} J_k(\boldsymbol{\beta}, 0)$. If this was not true, it would mean that, for any radius $\nu > 0$, the ball $\mathcal{B}_\nu^k(0)$ 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. However, this implies that at least 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 $[\widehat{W}_\circ^{(k)}]_{[1:k+1,1:k+1]}$, 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 $\boldsymbol{\beta}^k$ also 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. Define $\widehat{J}_k(\boldsymbol{\beta}, \delta_k) := J_k(\boldsymbol{\beta}, \delta_k A_{k+1})$ for $\delta_k > 0$. The first point of Assumption 7.4 implies that there exists a $\widehat{\delta}_k > 0$ such that for all $|\delta_k| < \widehat{\delta}_k$ we have $\widehat{J}_k(\boldsymbol{\beta}, \delta_k) = 0$. Thus, Lemma 7.2 guarantees that the correspondence $c_k : (-\widehat{\delta}_k, \widehat{\delta}_k) \rightrightarrows \mathbb{R}^k, c_k(\delta_k) = \arg\min_{\boldsymbol{\beta}\in\mathcal{K}_k} \widehat{J}_k(\boldsymbol{\beta}; \delta_k)$ 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

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

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., [14, Theorem 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, 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 $g'(\boldsymbol{\alpha}_\circ)(\widehat{\boldsymbol{\delta}}_{k+1}) = C\delta\boldsymbol{y}_\circ(\delta_k A_{k+1}, \boldsymbol{\epsilon}^m; T)$, we obtain

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

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 $C\delta\boldsymbol{y}_\circ(\delta_k A_{k+1}, \boldsymbol{\epsilon}^m; T)$ and $O(\|\widehat{\boldsymbol{\beta}^k}\|_2^2)$ faster

[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.

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 $C\delta \boldsymbol{y}_\circ(A^{(k)}(\boldsymbol{\beta}^k), \boldsymbol{\epsilon}^m; T) - C\delta \boldsymbol{y}_\circ(\delta_k A_{k+1}, \boldsymbol{\epsilon}^m; T) = 0$ 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.∎

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). Then there exists a solution $\boldsymbol{\epsilon}^{k+1} \in E_{ad}$ to (7.3) that simultaneously satisfies*

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

and

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

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 the following theorem, relating the two Algorithms 4.1 and 7.1.

THEOREM 7.7. *Consider the general setting of system (3.1) with a set of linearly independent matrices $\{A_1, \dots, A_K\}$ such that $\|A_k\|$ be sufficiently small 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 the GN matrix $\widehat{W}_\circ$, defined in (4.6), is PD.*

The proof of Theorem 7.7 is exactly the one of Theorem 5.6.

It remains to show that Assumption 7.6 holds in the settings considered in sections 5 and 6. First, we require the following results (see, e.g., [35, p. 1079]).

LEMMA 7.8 (on analytic functions in Banach spaces). *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 $g(\widetilde{x}) \neq 0$.*

We also require the following result about the analycity of control-to-state maps, which follows directly from the implicit function theorem (see, e.g., [35, p. 1081]).

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 [35, 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 the equation for the derivative $D_{\boldsymbol{y}} c(\widetilde{\boldsymbol{\epsilon}}, \widetilde{\boldsymbol{y}})(\delta \boldsymbol{y}) = \varphi$, 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 therefore 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., [35, p. 1078]) ℓ-linear, symmetric and continuous maps $a_\ell : (E_{ad})^\ell \to \mathbb{R}^N, (\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_\ell) \mapsto a_\ell(\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_\ell)$ 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 the evaluation map $S_T : Y \to \mathbb{R}^N, \boldsymbol{y} \mapsto \boldsymbol{y}(T)$ 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 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., [14]), 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. *Consider the linear setting* (5.1) *or the bilinear setting* (6.1). *For brevity, we 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 there exists a control $\boldsymbol{\epsilon} \in E_{ad}$ which satisfies* (7.6)-(7.7) *in Assumption 7.6.*

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 <u>linear setting</u> (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 $k \in \{1, \dots, K\}$ and $\boldsymbol{\beta}^k \in \mathbb{R}^k$ be the solution of (7.2) for $\|A_{k+1}\| > 0$ 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 $\operatorname{rank}(\mathcal{O}_N(C, A_+)) = \operatorname{rank}(\mathcal{O}_N(C, A_{\boldsymbol{\beta}})) = N$. 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

$$\widetilde{\boldsymbol{\epsilon}}(s) := \begin{cases} \boldsymbol{\epsilon}_{t_1}(s), & \text{for } 0 \le s < t_1, \\ \boldsymbol{c}, & \text{for } t_1 \le s \le T, \end{cases} \tag{730}$$

[7] Existence and uniqueness of all solutions $\boldsymbol{y}, \delta \boldsymbol{y}$ follow by Carathéodory's existence theorem (see, e.g., [32, Theorem 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 $\boldsymbol{y}, \delta \boldsymbol{y} \in H^1(0, T; \mathbb{R}^N)$.

where $\boldsymbol{c} \in \mathbb{R}^N$ is to be chosen later. For $t > t_1$, we have

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

Multiplying (7.8) with $e^{-(t-t_1)A_+}$ from the left, we get

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

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., [24, 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

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

Now, the observability of (A_+, C) guarantees the existence of some $i \in \{0, \dots, N-1\}$ 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 terms of the sum in (7.9) converge to zero at different rates for different j. Hence, there exists $t > t_1$ such that $C\boldsymbol{z}(t) \neq 0$. Since $t_1 \in (0, T]$ was chosen arbitrarily, we obtain $C\boldsymbol{z}(T) \neq 0$ 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. For a detailed proof, we refer to the supplementary material [12]. ∎

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 analytic and not the zero functional. Thus, we obtain by Lemma 7.8 that any neighborhood of $\boldsymbol{\epsilon}^{k+1}$ contains infinitely many $\boldsymbol{\epsilon}$ that do satisfy (7.7). This implies that it is rather unlucky to choose an $\boldsymbol{\epsilon}^{k+1}$ that does not satisfy (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 the linear (5.1) and bilinear (6.1) setting 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 $\boldsymbol{\alpha}_\star$.

7.2. Optimized GR Algorithm. The analysis discussed in the previous sections are based on certain hypotheses of observability and controllability of the dynamical system. However, as shown already in [11] and also discussed in the supplementary material [12], if these hypotheses are not satisfied, the choice of the elements in the set $\mathcal{A}$ becomes very relevant and can strongly affect the online reconstruction process. For this reason, a modified GR algorithm called Optimized GR (OGR) has been introduced in [11] to identify important basis elements by solving in each iteration the fitting and splitting step problems (in parallel) for all remaining basis elements, and not just the next one. This also allows us to initialize the algorithm with a number of elements $(A_j)_{j=1}^K$ with $K > N^2$. Even though some of the matrices A_j will inevitably be linearly dependent if $K > N^2$, the OGR algorithm 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 $\left\|C\boldsymbol{y}(A^{(k)}(\boldsymbol{\beta}^\ell),\boldsymbol{\epsilon}^m;T) - C\boldsymbol{y}(A_\ell,\boldsymbol{\epsilon}^m;T)\right\|_2^2 > \mathrm{tol}_2$ 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 basis 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 $\mathrm{rank}(\widehat{W}_\circ^{(k)}) = r > k$, 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, where P is the number of rows of the observer matrix C. This can be seen by writing $W_\circ(\epsilon^m)$, as defined in (4.6), 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, $\mathrm{rank}(W_\circ(\epsilon^m)) \le \mathrm{rank}(C) \le P$, and therefore $\mathrm{rank}(\widehat{W}_\circ^{(k)}) \le kP$.

The full OGR algorithm is stated in Algorithm 7.2, where the new feature that we described correspond to the steps 7-8. Algorithm 7.2 can be formulated for the linearized setting considered the previous sections by simply replacing the state $\boldsymbol{y}$ with its linearization $\boldsymbol{y}_\circ$. We call OLGR the OGR algorithm for the linearized system.

8. Numerical experiments.

8. Numerical experiments. In this section, efficiency and robustness of the GR and OGR algorithms are studied by direct numerical experiments. In particular, first we consider the reconstruction of a drift matrix in Section 8.1. Second, we focus on the reconstruction of a bilinear dipole momentum operator as Section 8.2. All optimization problems inside of the GR algorithms are solved by a BFGS descent-direction method, while 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 $A_\star, B, C \in \mathbb{R}^{3\times3}$. 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, 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 $\mathbb{R}^{3\times3}$ 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

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 basis 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

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 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 better basis, thereby optimizing the performance, 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 that was also considered in [11, 30] We obtain the results shown in the two panels on the right in Figure 8.1. The performance of

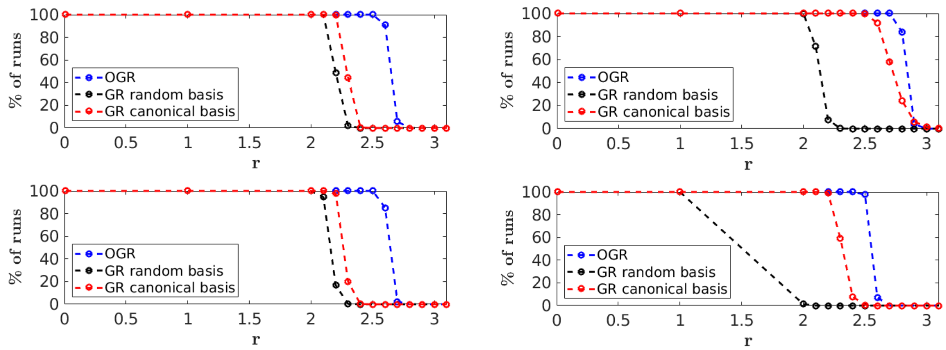

Fig. 8.1: Percentage of runs that converged (up to a tolerance) to the global minimum $A_\star$ starting from randomly chosen vectors on a nine-dimensional sphere with radius r, for controls generated by LGR and OLGR for 1% (top left) and 10% (bottom left) relative error between $A_\star$ and $A_\circ$, and GR and OGR in the version of Algorithm 7.1 (bottom right) and without the shift by $A_\circ$ (top right).

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 the GR algorithm 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 down to 3 and still outperforms any set of 9 controls generated by LGR.

8.2. Bilinear reconstruction problem. Similar to [30] and [11], we consider a Schrödinger-type equation, written as a real system as in (6.6). We also use similar matrices H and $\mu^\star$ as in [11], namely

$$H = H_R = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 16 \end{bmatrix}, \; \mu^\star = \begin{bmatrix} -0.3243 & -3.4790 + 0.7359i & -0.5338 + 1.9254i \\ -3.4790 - 0.7359i & -3.8342 & -1.1697 + 2.0256i \\ -0.5338 - 1.9254i & -1.1697 - 2.0256i & 1.0551 \end{bmatrix}.$$

The final time is $T = 10\pi$ and the initial state is $\boldsymbol{\psi}_0 = [1, 0, 0]^\top$. The observer matrix is $C = [\boldsymbol{\psi}_1, i\boldsymbol{\psi}_1]$, which means that the final state is measured against the fixed state $\boldsymbol{\psi}_1 = \frac{1}{\sqrt{3}}[1, 1, 1]^\top$. Again, we consider two bases, each consisting of 9 elements: the canonical and a random one for the space of Hermitian matrices in $\mathbb{C}^{3\times 3}$. We then perform the same experiments as in Section 8.1. The results are reported in Figure 8.2. We observe that the radii, up to which the control sets make GN capable of reliably reconstructing the global minimum, are much smaller than for the linear setting in Section 8.1. When the initial relative error between $\mu_\circ = \mu(\boldsymbol{\alpha}_\circ)$ and $\mu_\star = \mu(\boldsymbol{\alpha}_\star)$ is very small (1%) then LGR and OLGR have the most stable performance regarding the choice of the basis, making GN capable of reliably reconstructing the global minimum $\mu_\star$ up to a relative error of $4 - 5\%$. However, when the initial relative error is larger (10%) then only the LGR algorithm for the random basis can keep its performance, while even OLGR fails at errors of over 1%. The results for OGR, on the other hand, show the best performance, with and without a shift by $\mu_\circ$. The controls generated by the GR algorithms can not match OGR or LGR and OLGR for small initial errors, but are still more stable with respect to larger initial errors.

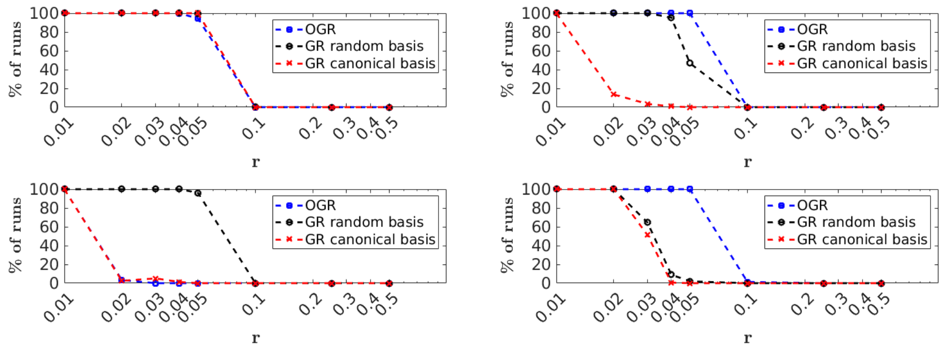

Fig. 8.2: Percentage of runs that converged (up to a tolerance) to the global minimum $\mu_\star$ starting from randomly chosen vectors on a nine-dimensional sphere with radius r, for controls generated by LGR and OLGR for 1% (top left) and 10% (bottom left) relative error between $\mu_\star$ and $\mu_\circ$, and GR and OGR in the version of Algorithm 7.1 (bottom right) and without the shift by $\mu_\circ$ (top right).

9. Conclusion. In this paper, we developed and analyzed greedy reconstruction algorithms based on the strategy presented in [30]. In particular, we tackled the case of nonlinear problems consisting in the reconstruction of drift operators in linear and bilinear dynamical systems. In these cases, we proved that the controls obtained with the greedy algorithm on the corresponding linearized systems lead to the local convergence of the classical Gauss-Newton method 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 also obtained.

Acknowledgements. G. Ciaramella is member of GNCS Indam. Simon Buchwald is funded by the DFG via the collaborative research center SFB1432, Project-ID 425217212. Julien Salomon acknowledges support from the ANR/RGC ALLOWAPP project (ANR-19-CE46-0013/A-HKBU203/19).

REFERENCES

[1] A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas, *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.
[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] S. Bonnabel, M. Mirrahimi, and P. Rouchon, *Observer-based Hamiltonian identification for quantum systems*, Automatica, 45 (2009), pp. 1144 – 1155.
[10] A. Borzì, G. Ciaramella, and M. Sprengel, *Formulation and Numerical Solution of Quan-*

tum Control Problems, SIAM, Philadelphia, PA, 2017.

[11] S. Buchwald, G. Ciaramella, and J. Salomon, *Analysis of a greedy reconstruction algorithm*, SIAM J. Control Optim., 59 (2021), pp. 4511–4537.

[12] S. Buchwald, G. Ciaramella, and J. Salomon, *Gauss-newton oriented greedy-type algorithms for the reconstruction of nonlinear operators: Supplementary material*, (2023).

[13] 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.

[14] P. G. Ciarlet, *Linear and Nonlinear Functional Analysis with Applications*, Applied mathematics, SIAM, Philadelphia, PA, 2013.

[15] J. Coron, *Control and Nonlinearity*, Math. Surveys Monogr., AMS, 2007.

[16] B. De Schutter, *Minimal state-space realization in linear system theory: An overview*, J. Comput. Appl. Math., 121 (2000), pp. 331–354.

[17] A. Donovan and H. Rabitz, *Exploring the Hamiltonian inversion landscape*, Phys. Chem., 16 (2014), pp. 15615–15622.

[18] 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.

[19] J. M. Geremia and H. Rabitz, *Global, nonlinear algorithm for inverting quantum-mechanical observations*, Phys. Rev. A, 64 (2001), p. 022710.

[20] 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.

[21] 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.

[22] E. Gilbert, *Controllability and observability in multivariable control systems*, J. SIAM Control Ser. A, 1 (1963), pp. 128–151.

[23] S. Glaser and et al, *Training schrödinger's cat: quantum optimal control*, Eur. Phys. J. D, 69 (2015), p. 279.

[24] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Press, 1991.

[25] 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.

[26] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, De Gruyter, Berlin, New York, 2008.

[27] C. T. Kelley, *Iterative Methods for Optimization*, SIAM, Philadelphia, 1999.

[28] K. Kobzar, T. E. Skinner, N. Khaneja, S. J. Glaser, and B. Luy, *Exploring the limits of broadband excitation and inversion: Ii. rf-power optimized pulses*, J. Magn. Reson., 194 (2008), pp. 58–66.

[29] C. Le Bris, M. Mirrahimi, H. Rabitz, and G. Turinici, *Hamiltonian identification for quantum systems: Well posedness and numerical approaches*, ESAIM: Contr. Optim. Ca., 13 (2007), pp. 378–395.

[30] 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.

[31] W. Rudin, *Real and Complex Analysis, 3rd Ed.*, McGraw-Hill, Inc., USA, 1987.

[32] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems (2Nd Ed.)*, Springer-Verlag, Berlin, Heidelberg, 1998.

[33] M. Tadi and H. Rabitz, *Explicit method for parameter identification*, J. Guid. Control Dyn., 20 (1997), pp. 486–491.

[34] Y. Wang, D. Dong, B. Qi, J. Zhang, I. R. Petersen, and H. Yonezawa, *A quantum Hamiltonian identification algorithm: Computational complexity and error analysis*, IEEE Trans. Autom. Control, 63 (2018), pp. 1388–1403.

[35] E. F. Whittlesey, *Analytic functions in Banach spaces*, Proceedings of the American Mathematical Society, 16 (1965), pp. 1077–1083.

[36] 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.

[37] J. Zhang and M. Sarovar, *Quantum Hamiltonian identification from measurement time traces*, Phys. Rev. Lett., 113 (2014), p. 080401.

[38] 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.

[39] W. Zhu and H. Rabitz, *Potential surfaces from the inversion of time dependent probability density data*, J. Chem. Phys., 111 (1999), pp. 472–480.

MOX Technical Reports, last issues

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

08/2023 Bonizzoni, F.; Hu, K.; Kanschat, G.; Sap, D.
Discrete tensor product BGG sequences: splines and finite elements

07/2023 Garcia-Contreras, G.; Còrcoles, J.; Ruiz-Cruz, J.A.; Oldoni, M; Gentili, G.G.; Micheletti, S.; Perotto, S.
Advanced Modeling of Rectangular Waveguide Devices with Smooth Profi les by Hierarchical Model Reduction

06/2023 Artoni, A.; Antonietti, P. F.; Mazzieri, I.; Parolini, N.; Rocchi, D.
A segregated finite volume - spectral element method for aeroacoustic problems

05/2023 Fumagalli, I.; Vergara, C.
Novel approaches for the numerical solution of fluid-structure interaction in the aorta

04/2023 Quarteroni, A.; Dede’, L.; Regazzoni, F.; Vergara, C.
A mathematical model of the human heart suitable to address clinical problems

03/2023 Africa, P.C.; Perotto, S.; de Falco, C.
Scalable Recovery-based Adaptation on Quadtree Meshes for Advection-Diffusion-Reaction Problems

01/2023 Zingaro, A.; Bucelli, M.; Piersanti, R.; Regazzoni, F.; Dede', L.; Quarteroni, A.
An electromechanics-driven fluid dynamics model for the simulation of the whole human heart

02/2023 Boon, W. M.; Fumagalli, A.; Scotti, A.
Mixed and multipoint finite element methods for rotation-based poroelasticity

83/2022 Ciaramella, G.; Gander, M.; Mazzieri, I.
Unmapped tent pitching schemes by waveform relaxation

82/2022 Ciaramella, G.; Gander, M.; Van Criekingen, S.; Vanzan, T.
A PETSc Parallel Implementation of Substructured One- and Two-level Schwarz Methods