

Active Estimation of Multiplicative Faults in Dynamical Systems

Gabriel de A. Gleizer^a, Peyman Mohajerin Esfahani^{a,b}, Tamas Keviczky^a

July 8, 2025

Abstract

This paper addresses the problem of estimating multiplicative fault signals in linear time-invariant systems by processing its input and output variables, as well as designing an input signal to maximize the accuracy of such estimates. The proposed real-time fault estimator is based on a residual generator used for fault detection and a multiple-output regressor generator, which feed a moving-horizon linear regression that estimates the parameter changes. Asymptotic performance guarantees are provided in the presence of noise. Motivated by the performance bounds, an optimal input design problem is formulated, for which we provide efficient algorithms and optimality bounds. Numerical examples demonstrate the efficacy of our approach and the importance of the optimal input design for accurate fault estimation.

1 Introduction

In control applications, dynamical systems can occasionally display unexpected behavior due to the presence of faults. Faults can occur due to many reasons, such as unexpected disturbances or malicious attacks, failure of some component, or wear and tear. If not detected timely, faults can lead to economic losses due to out-of-specification performance and system shutdown, or even catastrophic events incurring human, property, and/or environmental damage. For this reason, fault diagnosis methods have been subject of intense research since the early seventies [1], [2] and have since been applied to a myriad of applications, such as robots, transport systems, power systems, manufacturing processes, and chemical processes [3]. Particularly for high-performance control systems, such as high-tech machines, not only the detection of faulty behavior is important, but also the estimation of how much change has occurred, a problem called *fault estimation*, which can be further classified in two types of faults: *additive* and *multiplicative*. This distinction is made depending on the end goal of the technological health monitoring system. If it is sufficient to model the fault in question as an external signal that enters additively in the system equations, one has an *additive fault estimation* problem; in this case, one is satisfied with tracking such an external signal. Instead, if it is more appropriate to model the fault as a change in system parameters, or additional intrusive dynamics, it becomes a problem of *multiplicative fault estimation*. The latter is the topic of the present paper. When faults are additive and the system is linear, the estimation performance generally does not depend on the input to the system, and hence active fault estimation is unnecessary. In contrast, this separation property does not hold for multiplicative faults. For example, changes in friction components of a mechanical system can only be detected when it moves. Therefore, the ability of the input signals to properly excite the system is a critical component of multiplicative fault estimation. When a fault estimation scheme involves input design or control design, it is called *active fault estimation*. Given its impact in the estimation performance, input design is also a central part of this paper.

The idea of designing inputs to increase the diagnosability of faults was probably pioneered by [4] using a statistical framework. Later, a deterministic multi-model approach was proposed by [5], whose objective is to determine which model generates the observed trajectory between two candidates. This has been extended to

This work was supported by the Digital Twin project with project number P18-03 of the research programme TTW-Perspectief, which is partly financed by the Dutch Research Council (NWO). Peyman Mohajerin Esfahani acknowledges the support of the ERC grant TRUST949796. The authors are with (a) Delft University of Technology, The Netherlands; and (b) University of Toronto, Canada.

multiple model candidates in a hybrid stochastic-deterministic framework in [6]. As an alternative to offline design of input signals, [7] considers a closed loop approach where the residual generator also creates a signal to be fed back to the system in order to improve fault detectability and/or tolerance. This problem is still an active area of research [8], [9], [10]. An overview of active fault detection can be found in [11], and of input design for fault diagnosis in [12].

Most of the focus of the literature on active diagnosis has been on classification problems as reported above, as opposed to fault *estimation*. Nonetheless, the related field of experiment design for system identification is much more well established. It emerged in the seventies (see, e.g., [13], [14], [15]) inspired by experiment design in the statistics literature. The idea is to design an input that maximizes a function of the Fischer Information Matrix (FIM), which depends on input data, based on the observation that the variance of a linear estimator is inversely proportional to it. This leads to so-called *T-optimal* and *E-optimal* designs, which maximize, respectively, the trace and the minimum eigenvalue of the FIM. After that, most focus has been given to problems associated to identification for control (e.g., [16]), closed-loop identification (e.g., see [17], which also serves as an overview of this literature), and frequency domain identification, see, e.g., [18], [19]. More recently, [20] addressed the problem of computing the T-optimal input in the MIMO case where the total energy of the input and output are bounded. They provide global optimal results for the case of only two quadratic constraints building on the results of [21]. With the recent surge of data-driven control, experiment design has also been gaining increased attention [22], [23], [24]. In [22], a series of experiments is designed to achieve the most accurate system identification with the least samples, using the classical FIM-optimal approaches as an oracle. In [23], [24], the focus is qualitative rather than quantitative: the goal is to derive sufficient excitation conditions that enable designing a stabilizing controller from an experiment. Since our focus is quantitative, our experiment design problem is closer to that of [13], [14], [15], [16], [17], [20] and related work.

In this work, we propose a fault estimation method for LTI systems described by differential-algebraic equations (DAEs), which include classical ordinary differential and difference equations as a special case [25], [26], [27]. The method is designed to estimate *multiplicative* faults. Additionally, we propose an optimal input design algorithm that maximizes the accuracy of the aforementioned method. The main contributions are summarized in the following list:

1. **Estimation of multiplicative faults affecting latent variables.** We extend the work in [28] to multiple simultaneous faults that possibly multiply latent variables such as states and external disturbances. The fault estimation system has the same structure as [28] (see Fig. 1), including a residual generator, a regressor generator, and a fault estimator. Due to the fault characteristics mentioned above, a novel formulation for the design of residual and regressor generators is required, which we provide in Theorem 3.1. The fault estimation is obtained with a Gauss–Newton-like approach.
2. **Performance characterization: bias and variance.** We characterize the bias and variance bounds of the least-squares fault estimation system in the presence of measurement noise (Theorem 3.2), assuming the regressor signal is persistently exciting. The bias is zero in the special case where the multiplicative faults affect only the measured variables (Corollary 3.1).
3. **Optimal input design.** Motivated by the error characterization formula of Theorem 3.2, we propose an E-optimal periodic input design that maximizes the asymptotic value of the regressor’s singular value in the absence of external disturbances (Definition 4.1).
 - 3a. **Exact formulation and first-order information.** We provide an exact formulation of the optimal periodic input design problem, including exact first-order information via a subgradient (Theorem 4.1). This paves the way to a local optimization method based on convex optimization that yields fast and efficient local optima (Algorithm 1).
 - 3b. **Convex relaxation and sub-optimality gap.** We provide a convex semi-definite relaxation (Proposition 4.2), which gives a conservative certificate of the local solution’s sub-optimality gap.

This paper is organized as follows. Section 2 provides mathematical preliminaries and presents the problem formulation. Section 3 presents the fault estimator design and provides formal results on the estimation performance. Section 4 presents the input design problem and methods to solve it. Section 5 shows numerical

simulations, highlighting how input design plays a major role in the estimation accuracy. Section 6 concludes the paper with discussions and ideas for future work. Finally, Section 7 contains the technical proofs, which we set apart from the main text for improved readability.

Mathematical notation. Throughout the text, we use *Italic typesetting*, e.g., a , for scalars or scalar-valued functions; **bold letters**, e.g., \mathbf{a} , for vectors or vector-valued functions; **bold capital letters**, e.g., \mathbf{A} , for matrices; and **calligraphic letters**, e.g., \mathcal{A} , for sets or operators in Hilbert spaces such as transfer functions. We denote by \mathbb{N}_0 the set of natural numbers including zero, $\mathbb{N} := \mathbb{N}_0 \setminus \{0\}$, $\mathbb{N}_{\leq n} := \{1, 2, \dots, n\}$, and by \mathbb{R}_+ the set of non-negative reals. We denote by $\|\mathbf{x}\|_p$ the p -norm of the vector $\mathbf{x} \in \mathbb{R}^n$, dropping the subscript when $p = 2$. We denote by $\|\mathbf{A}\|_p$ the p -induced norm of the matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$. For a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\lambda(\mathbf{A}) \subset \mathbb{C}^n$ is the set of its eigenvalues, and $\lambda_i(\mathbf{A})$ is the i -th largest-in-magnitude. For any matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, $s(\mathbf{A}) \subset \mathbb{R}_+$ denotes the set of its singular values, with s_i being the i -th largest. The complex conjugate of $z \in \mathbb{C}^n$ is denoted by z^* . The set \mathbb{S}^n denotes the set of symmetric matrices in \mathbb{R}^n . For $\mathbf{P} \in \mathbb{S}^n$, we write $\mathbf{P} \succ \mathbf{0}$ ($\mathbf{P} \succeq \mathbf{0}$) if \mathbf{P} is positive definite (semi-definite); and $\mathbf{P} \prec \mathbf{0}$ ($\mathbf{P} \preceq \mathbf{0}$) if \mathbf{P} is negative definite (semi-definite). The set of positive (semi-)definite matrices in \mathbb{S}^n is denoted by \mathbb{S}_{++}^n (\mathbb{S}_+^n). The partial order induced by \preceq allows us to use $\mathbf{P} \preceq \mathbf{S}$ when $\mathbf{P} - \mathbf{S} \preceq \mathbf{0}$ (and equivalently for $\mathbf{P} \succeq \mathbf{S}$).

For a signal $\mathbf{x} : \mathcal{T} \rightarrow \mathbb{R}^n$, where \mathcal{T} is the time axis (typically \mathbb{N}_0 or \mathbb{R}_+), we denote by $\mathbf{X}_{[a, a+(N-1)h]}^h$ the N by n matrix $[\mathbf{x}(a) \quad \mathbf{x}(a+h) \quad \dots \quad \mathbf{x}(a+(N-1)h)]$. When the sampling interval h and the time range $[a, a+(N-1)h]$ are clear from context, we may drop the subscript and superscript, and the mere capitalization shall indicate the transformation.

2 Preliminaries and problem formulation

2.1 System model

Throughout this work, we consider faulty linear time-invariant dynamical systems described in a differential-algebraic equation (DAE) form,

$$\left(\mathbf{H}(\mathbf{q}) + \sum_{i=1}^m f_i \mathbf{H}'_i(\mathbf{q}) \right) \boldsymbol{\xi} + \left(\mathbf{L}(\mathbf{q}) + \sum_{i=1}^m f_i \mathbf{L}'_i(\mathbf{q}) \right) \mathbf{z} + \mathbf{W}(\mathbf{q}) \mathbf{w} = \mathbf{0}, \quad (1)$$

where $\boldsymbol{\xi}$, \mathbf{z} and \mathbf{w} are (either continuous-time or discrete-time) signals taking values in \mathbb{R}^{n_ξ} , \mathbb{R}^{n_z} and \mathbb{R}^{n_w} , respectively. The matrices $\mathbf{H}(\mathbf{q})$, $\mathbf{L}(\mathbf{q})$, $\mathbf{W}(\mathbf{q})$, $\mathbf{H}'_i(\mathbf{q})$, $\mathbf{L}'_i(\mathbf{q})$, for $i \in \mathbb{N}_{\leq m}$ are polynomial matrices in the operator \mathbf{q} , with n_r rows and compatible number of columns. The operator \mathbf{q} is a linear operator in the signal space where $\boldsymbol{\xi}$, \mathbf{z} and \mathbf{w} are defined, typically the forward shift operator for discrete-time signals or the (right) time derivative for continuous-time (right) differentiable signals. The scalars f_i represent unknown multiplicative faults, which are assumed to be constant for analysis; in Section 3.3 we show how to extend the framework for time-varying faults. The signal $\boldsymbol{\xi}$ comprises the system's latent variables, such as internal states and external disturbances, \mathbf{z} comprises known signals such as control inputs, measured outputs and references, and \mathbf{w} represents unknown i.i.d. noise with zero mean and variance σ^2 . All components of \mathbf{w} can have the same variance without loss of generality, by rescaling the rows of (1). One may notice that we have partitioned unknown variables in two; this is a common approach in the fault diagnosis literature, that allows one to distinguish unknown variables $\boldsymbol{\xi}$ that do not have an a priori statistical description, and are determined either by dynamics or an external agent, from unknown variables \mathbf{w} which may be characterized as random processes. We make the following nominal (disturbance) observability assumption.

Assumption 2.1. (Nominal observability) The polynomial matrix $\mathbf{H}(\mathbf{q})$ has full column rank for all $\mathbf{q} \in \mathbb{C}$. Equivalently, $\mathbf{H}(\mathbf{q})$ has a polynomial left-inverse $\mathbf{H}(\mathbf{q})^\dagger$ such that $\mathbf{H}(\mathbf{q})^\dagger \mathbf{H}(\mathbf{q}) = \mathbf{I}$ (see, e.g., [29]).

We call Assumption 2.1 a nominal observability condition because, when $f_i = 0$ for all i and $\mathbf{w} \equiv \mathbf{0}$, the unmeasured variables $\boldsymbol{\xi}$ satisfy $\boldsymbol{\xi} = -\mathbf{H}(\mathbf{q})^\dagger \mathbf{L}(\mathbf{q}) \mathbf{z}$; that is, they can be determined by combinations of differentials of the measured variables \mathbf{z} . Note that observability is only required in the nominal case, i.e., the faults f_i can be estimated even if $\mathbf{H}(\mathbf{q}) + \sum_{i=1}^m f_i \mathbf{H}'_i(\mathbf{q})$ loses rank. One advantage of the DAE framework used here is that, even though we impose an observability condition, it does not require building (unknown input) observers to obtain a fault estimation system, which has great scalability implications.

Remark 2.1. (Converting state-space ODE to DAE) The DAE framework in (1) encompasses standard differential–algebraic state-space LTI formulations of the form

$$\begin{aligned} \left(\mathbf{G} + \sum_{i=1}^m f_i \mathbf{G}'_i \right) \mathbf{x}(k+1) &= \left(\mathbf{A} + \sum_{i=1}^m f_i \mathbf{A}'_i \right) \mathbf{x}(k) + \left(\mathbf{B}_u + \sum_{i=1}^m f_i \mathbf{B}'_i \right) \mathbf{u}(k) + \left(\mathbf{B}_d + \sum_{i=1}^m f_i \mathbf{B}'_{d,i} \right) \mathbf{d}(k) + \mathbf{B}_w \mathbf{w}(k), \\ \mathbf{y}(k) &= \left(\mathbf{C} + \sum_{i=1}^m f_i \mathbf{C}'_i \right) \mathbf{x}(k) + \left(\mathbf{D}_u + \sum_{i=1}^m f_i \mathbf{D}'_i \right) \mathbf{u}(k) + \left(\mathbf{D}_d + \sum_{i=1}^m f_i \mathbf{D}'_{d,i} \right) \mathbf{d}(k) + \mathbf{D}_w \mathbf{w}(k), \end{aligned} \quad (2)$$

where $\mathbf{x} : \mathbb{N}_0 \rightarrow \mathbb{R}^{n_x}$ is the state trajectory, $\mathbf{u} : \mathbb{N}_0 \rightarrow \mathbb{R}^{n_u}$ is the input signal, $\mathbf{y} : \mathbb{N}_0 \rightarrow \mathbb{R}^{n_y}$ is the output signal, and $\mathbf{d} : \mathbb{N}_0 \rightarrow \mathbb{R}^{n_d}$ is the external disturbance signal. By taking $\mathbf{z} = [\mathbf{y}^\top \ \mathbf{u}^\top]^\top$ and $\boldsymbol{\xi} = [\mathbf{x}^\top \ \mathbf{d}^\top]^\top$, the following conversion from (2) to (1) holds: for all $i = 1, \dots, m$,

$$\begin{aligned} \mathbf{H}(\mathbf{q}) &= \begin{bmatrix} -\mathbf{q}\mathbf{G} + \mathbf{A} & \mathbf{B}_d \\ \mathbf{C} & \mathbf{D}_d \end{bmatrix}, & \mathbf{L}(\mathbf{q}) &= \begin{bmatrix} \mathbf{0} & \mathbf{B}_u \\ -\mathbf{I} & \mathbf{D}_u \end{bmatrix}, & \mathbf{W}(\mathbf{q}) &= \begin{bmatrix} \mathbf{B}_w \\ \mathbf{D}_w \end{bmatrix}, \\ \mathbf{H}'_i(\mathbf{q}) &= \begin{bmatrix} -\mathbf{q}\mathbf{G}'_i + \mathbf{A}'_i & \mathbf{B}'_{d,i} \\ \mathbf{C}'_i & \mathbf{D}'_{d,i} \end{bmatrix}, & \mathbf{L}'_i(\mathbf{q}) &= \begin{bmatrix} \mathbf{0} & \mathbf{B}'_i \\ \mathbf{0} & \mathbf{D}'_i \end{bmatrix}. \end{aligned}$$

2.2 Linear-algebra operations in polynomial matrices

It is possible to translate certain polynomial matrix equations into standard linear algebraic ones, which are suitable for efficient numerical methods from linear algebra and optimization. These techniques have been used in [26] and others. For that, let us introduce some notation: given a polynomial matrix $\mathbf{H}(\mathbf{q}) := \sum_{i=0}^d \mathbf{H}_i \mathbf{q}^i$ of degree d , we denote

$$\begin{aligned} \text{blkrow}(\mathbf{H}(\mathbf{q})) &:= [\mathbf{H}_0 \ \mathbf{H}_1 \ \dots \ \mathbf{H}_d], \\ \tilde{\mathbf{H}} &:= \begin{bmatrix} \mathbf{H}_0 & \dots & \mathbf{H}_d & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_0 & \dots & \mathbf{H}_d & \mathbf{0} & \dots \\ \vdots & & \ddots & & \ddots & \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{H}_0 & \dots & \mathbf{H}_d \end{bmatrix}, \end{aligned}$$

where $\text{blkrow}(\mathbf{H}(\mathbf{q}))$ is the block-row form of $\mathbf{H}(\mathbf{q})$ and $\tilde{\mathbf{H}}$ is its block-Toeplitz form. In the latter, the number of block rows shall be derived from the context of its usage, which we provide next:

Lemma 2.1. (Characterization of polynomial matrix multiplication [27, Lemma 4.2]) Let $\mathbf{H}_1(\mathbf{q})$ and $\mathbf{H}_2(\mathbf{q})$ be polynomial matrices of degrees d_1 and d_2 , respectively, and with dimensions allowing the product $\mathbf{H}_1(\mathbf{q})\mathbf{H}_2(\mathbf{q})$. The following identity holds:

$$\text{blkrow}(\mathbf{H}_1(\mathbf{q})\mathbf{H}_2(\mathbf{q})) = \text{blkrow}(\mathbf{H}_1(\mathbf{q}))\tilde{\mathbf{H}}_2,$$

where the number of block rows of $\tilde{\mathbf{H}}_2$ is $d_1 + 1$.

2.3 Problem statement

In this work we are interested in three problems: (i) to design a fault estimation filter, (ii) to characterize its performance bounds, and (iii) to design an input signal within given constraints that minimizes the fault estimation error. We formalize these problems below.

Problem 2.1 (Estimator design). Consider system (1), and recall that the signal \mathbf{z} is available. We aim to design a (nonlinear) fault estimation filter \mathcal{F} that generates an estimate *signal* $\hat{\mathbf{f}}$ for the multiplicative fault $\mathbf{f} := [f_1 \ f_2 \ \dots \ f_m]^\top$ such that

1. **Causality:** the filter \mathcal{F} is a causal dynamical system taking \mathbf{z} as input;
2. **Estimation error order:** the multiplicative estimation error $|\hat{\mathbf{f}}(t) - \mathbf{f}|$ in the absence of noise is on the order of $|\mathbf{f}|^2$;

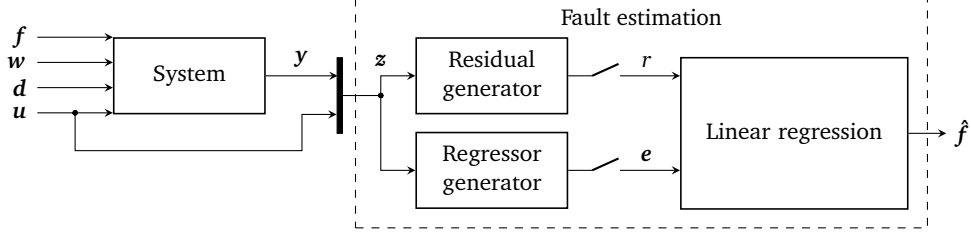


Figure 1: The proposed fault estimation architecture. The switches represent samplers, since the linear regression always operates in discrete-time. The system, residual generator and regressor generator can be either continuous- or discrete-time.

3. **Asymptotic consistency:** for sufficiently high signal-to-noise ratio, the *asymptotic expected squared error* $\lim_{t \rightarrow \infty} \mathbb{E}[|\hat{f}(t) - f|^2]$ is on the order of $\sigma^2(1 + |f|^2)$, with the bias on the order of $|f|^2$.

The estimation error specifications in Problem 2.1 depend on $|f|^2$, making them suitable for small faults f .

The second task is characterizing performance bounds of our designed fault estimation method. Again, focusing on the small-fault scenario, we focus on asymptotic results as f approaches zero. Following the stochastic perturbation theory of [30], we concentrate our attention to the *first-order approximation* \check{f} of the estimate \hat{f} .

Problem 2.2 (Performance bounds). Given system (1) and the designed fault estimator filter \mathcal{F} , determine bounds on the asymptotic bias and variance of the first-order approximation of the error $\check{f}(t) - f(t)$.

Designing a good fault estimation scheme for multiplicative faults may not be sufficient for their accurate estimation. Effective excitation of the system's states is also necessary. Therefore, for active estimation of the faults f , we pose the following experiment design problem:

Problem 2.3 (Input design). Given system (1) and the designed fault estimator filter \mathcal{F} , compute a periodic signal u that minimizes the asymptotic performance bounds obtained by solving Problem 2.2.

Remark 2.2. (Finite-time vs. asymptotic performance) Asymptotic performance criteria measure performance after the contribution of the initial state vanishes. In this case, input periodicity ensures that the performance metric $\mathbb{E}[|\hat{f}(t) - f(t)|^2]$ converges as t grows large. If the initial state is *known*, it can be easily incorporated in the input design problem, and a T -periodic input renders $\mathbb{E}[|\hat{f}(t) - f(t)|^2]$ constant for all $t \geq T$.

3 Fault estimation of multiple multiplicative faults

Inspired by [28], we design a system composed of three blocks as depicted in Fig. 1. One block is a residual generator, which as usual should have zero response if no faults are present, but we shall design to have specific sensitivity properties to the faults f_i that allows them to be distinguishable among each other. The block below it is a regressor generator, similar to what is called pre-filter in [28], whose function is to provide a basis of signals that serves as regressors to the final block, the fault estimator, which in turn performs a linear regression in a moving-horizon fashion.

Both the residual and the regressor generators are LTI systems. They shall be designed, respectively, as the transfer functions $d(q)^{-1}N(q)$ and $d(q)^{-1}M(q)$ with n_z inputs and 1 and m outputs, respectively, where $d(q)$ is a polynomial designed to be stable, provide good noise attenuation, and render the generators proper. As a first requirement, the residual r must satisfy, under small fault assumptions, the approximation

$$r \approx f^T d(q)^{-1} M(q) z. \quad (3)$$

That is, the residual is approximately the linear combination of the regressors. This paves the way for applying linear regression to obtain an estimate $\hat{f}_i(t)$ of each fault f_i , from a sampled trajectory of r and $d(q)^{-1}M(q)z$ from $t - T$ to t .

This framework imposes certain requirements to the residual and regressor generators. In addition to (3), we want that the term $f^T d(q)^{-1} M(q) z$ is not identically zero for nonzero faults f , and that $M(q) z$ spans linearly independent signals for sufficiently exciting z . For the excitation condition, we invoke a recent notion originally proposed for continuous-time systems:

Definition 3.1. (Persistency of excitation [31, Def. 1]) A signal $z : \mathcal{T} \rightarrow \mathbb{R}^m$ is *persistently exciting (PE)* of order d if, for all $v \in \mathbb{R}^{md}$,

$$v^T \begin{bmatrix} z(t) \\ qz(t) \\ \vdots \\ q^{d-1}z(t) \end{bmatrix} = 0 \quad \forall t \implies v = 0. \quad (4)$$

The design requirements are summarized in the definition below:

Definition 3.2. (Independently sensitive residual generator) Consider system (1) with noise $w \equiv 0$. A BIBO stable linear filter $d^{-1}(q)N(q)L(q)$ is called an *independently sensitive residual generator* if there exist a polynomial matrix $M \in \mathbb{R}[q]^{m \times n_z}$ such that the following conditions hold:

1. $\lim_{t \rightarrow \infty} r(t) = 0 \iff f \equiv 0$, i.e., the signal r is a fault detection residual;
2. for any norm $|\cdot|$, $\lim_{|f| \rightarrow 0} r - f^T d(q)^{-1} M(q) z = 0$, i.e., the approximation (3) holds;
3. if z is PE of sufficiently high order, the signals $d(q)^{-1} M(q) z$ are linearly independent.

The corresponding filter $d(q)^{-1} M(q)$ is called a *regressor generator*.

3.1 Residual and regressor generator design

Consider a standard residual generator used for the additive fault case, i.e., one taking the form $d(q)r = N(q)L(q)z$, where $N(q)H(q) \equiv 0$. The first observation is that, when $f_i \neq 0$, $N(q)$ generally ceases to be in the left null space of the system (1). I.e., $N(q)H(q) \equiv 0$ does not imply that $N(q)(H(q) + \sum_i f_i H'_i(q)) \equiv 0$. In this case, the residual output is affected by input-output data of the system z as well as the latent variables ξ . However, this dependency is largely predictable if the faults f_i are sufficiently small, as we show next.

Lemma 3.1 (Residual characterization). Consider system (1) and a residual generator described by $d(q)r = N(q)L(q)z$ satisfying $N(q)H(q) \equiv 0$. If Assumption 2.1 holds, then the residual r satisfies

$$d(q)r = N(q) \sum_i f_i G_i(q) z + \sum_{i,j} f_i f_j J_{i,j}(q) \begin{bmatrix} \xi \\ z \end{bmatrix} + \sum_i f_i F_i(q) w - N(q)W(q)w. \quad (5)$$

where $G_i(q) := H'_i(q)H(q)^\dagger L(q) - L'_i(q)$ for all $i = 1, \dots, m$, $H^\dagger(q)$ is a polynomial matrix satisfying $H^\dagger(q)H(q) = I$ and $J_{i,j}(q)$ and $F_i(q)$ are polynomial matrices. Furthermore, if $H'_i(q) \equiv 0$ for all $i = 1, \dots, m$, then $J_{i,j}(q) \equiv 0$ and $F_i(q) \equiv 0$ for all $i, j = 1, \dots, m$.

Proof: See Section 7.1. ■

Remark 3.1 (Linear-algebraic computation of the left-inverse of a polynomial matrix). From Assumption 2.1, there exists a polynomial matrix $H(q)^\dagger$ satisfying $H(q)^\dagger H(q) = I$. Hence, for a sufficiently large natural number k , there exist $H_i^\dagger, i = 0, \dots, k$ such that

$$H^\dagger(q)H(q) = \begin{bmatrix} H_0^\dagger & \dots & H_k^\dagger \end{bmatrix} \begin{bmatrix} H_0 & \dots & H_d & 0 & \dots & 0 \\ 0 & H_0 & \dots & H_d & 0 & \dots \\ \vdots & & \ddots & & \ddots & \\ 0 & \dots & 0 & H_0 & \dots & H_d \end{bmatrix} \begin{bmatrix} I \\ qI \\ \vdots \\ q^{k+d}I \end{bmatrix} = I.$$

Hence, one can obtain $H^\dagger(q)$ with standard linear algebra by solving

$$\text{blkrow}(H^\dagger(q))\bar{H} = \begin{bmatrix} I & 0 & \dots & 0 \end{bmatrix}$$

where the number of rows of \bar{H} is $k + 1$.

Let us analyze equation (5) further. By dividing it by $d(q)$, we can distinguish r as the sum of the outputs of four LTI systems. The first term performs a linear combination, where the weights are the faults f_i , of the LTI systems $N(q)G_i(q)$ applied to the available signals \mathbf{z} . The second term contains responses of LTI systems to both ξ and \mathbf{z} , but scaled to second-order combinations of the faults; hence, they tend to have a negligible effect on the residual if the faults are small. The third and fourth terms show the effect of noise, which, when the faults are small, is dominated by the fourth term $-N(q)W(q)$. These observations indicate that, for any $N(q)$ satisfying the conditions in Lemma 3.1, the regressor generator $d^{-1}(q)M(q)$ where

$$M(q) := \begin{bmatrix} N(q)G_1(q) \\ N(q)G_2(q) \\ \vdots \\ N(q)G_m(q) \end{bmatrix}, \quad (6)$$

satisfies requirement 2 of Definition 3.2 by taking $T(q) = -N(q)W(q)$. To additionally satisfy requirement (b), the residual must satisfy the following equivalent linear-algebraic conditions:

Theorem 3.1 (Linear-algebraic filter characterization). *Consider system (1), matrices G_i as in Lemma 3.1, and assume Assumption 2.1 holds. Given a polynomial $d(q)$ of sufficiently high order, the system $d^{-1}(q)N(q)L(q)$ is an independently sensitive residual generator with corresponding regressor generator $d^{-1}(q)M(q)$ if $M(q)$ has the form (6) and*

$$\text{blkrow}(N(q))\bar{H} = \mathbf{0}, \quad (7a)$$

$$\text{rank}(\text{blkrow}(M(q))) = m, \quad (7b)$$

where

$$\text{blkrow}(M(q)) = \begin{bmatrix} \text{blkrow}(N(q))\bar{G}_1 \\ \vdots \\ \text{blkrow}(N(q))\bar{G}_m \end{bmatrix}. \quad (8)$$

Proof: See Section 7.1. ■

Unlike residual generator design for additive faults as in [27], conditions (7) do not allow for an efficient linear-programming characterization, by virtue of the rank condition (7b). This rank condition does not appear in [28], since it handles a single multiplicative fault: when $m = 1$, condition (7b) simplifies to $\text{blkrow}(N(q))\bar{G}_1 \neq \mathbf{0}$, which is the same condition as in the design of a filter for additive faults. In addition, in [28], the concern is precisely when the residual generator for a given additive fault *cannot* nullify the effect of the multiplicative fault; that is, the sensitivity to a multiplicative fault is a by-product of the residual design rather than a requirement.

Nonetheless, the matrix $\text{blkrow}(M(q))$ has m rows and $(k_m + 1)n_z$ columns, where k_m is the degree of $M(q)$; as such, if $(k_m + 1)n_z > m$, it is unlikely that $N(q)$ and $M(q)$ violate (7). Therefore, we suggest the following simple procedure to obtain matrices $N(q)$ and $M(q)$, which works satisfactorily for small m :

1. Compute a unitary matrix representation N_H of the left null-space of \bar{H} by, e.g., singular-value decomposition; let b be the number of rows of N_H .
2. Randomly generate K unitary vectors $\mathbf{v}_i \in \mathbb{R}^b$.
3. For each $i = 1, \dots, K$, compute $\text{blkrow}(N^i(q)) = \mathbf{v}_i^T N_H$ and $\text{blkrow}(M^i(q))$ following (8).
4. Compute $i^* = \arg \max_{i \in \mathbb{N}_{\leq K}} (s_{\min}(\text{blkrow}(M^i(q))))$.
5. Select the filter parameters $N(q) = N^{i^*}(q)$ and $M(q) = M^{i^*}(q)$.

The procedure above is a Monte-Carlo optimization where the minimum singular value of $\text{blkrow}(M(q))$ is maximized. This not only ensures full rank of $M(q)$ but additionally controls its condition number. This concludes the residual and regressor design results.

3.2 Fault estimation

With the generators designed according to Theorem 3.1, the main objective of this subsection is to show that our proposed least-squares approach (the rightmost block of Fig. 1) solves Problem 2.1.

Let us call $\mathbf{e} := d(q)^{-1}\mathbf{M}(q)\mathbf{z}$ the regressor signal. Throughout this Section, we fix a sampling interval h and a sample count N to be used in the regression step.¹ The fault estimation is performed by computing, at every time step $t \geq Nh$ multiple of h , a standard least-squares operation, i.e.,

$$\hat{\mathbf{f}}(t) = (\mathbf{E}_{[t-(N-1)h, t]}^h)^\dagger \mathbf{R}_{[t-(N-1)h, t]}^h \quad (9)$$

where $\mathbf{E}^\dagger = (\mathbf{E}^\top \mathbf{E})^{-1} \mathbf{E}^\top$ is the (Moore–Penrose) pseudo-inverse of \mathbf{E} .²

In what follows, we provide performance bounds in the estimation error $\mathbf{f} - \hat{\mathbf{f}}(t)$. We start by representing (5) in discrete time according to the sampling interval h selected for the fault estimation block:

$$\begin{aligned} \mathbf{r} &= \mathbf{f}^\top \mathbf{e} + \sum_{i,j} f_i f_j \mathcal{T}_{i,j}^J(\mathfrak{z}) \begin{bmatrix} \xi \\ \mathbf{z} \end{bmatrix} + \left(\sum_i f_i \mathcal{T}_i^F(\mathfrak{z}) + \mathcal{T}^W(\mathfrak{z}) \right) \mathbf{w}, \\ e_i &= \mathcal{T}_i^M(\mathfrak{z}) \mathbf{z}, \quad i = 1, \dots, m \end{aligned} \quad (10)$$

where the transfer functions $\mathcal{T}_{i,j}^J(\mathfrak{z})$, $\mathcal{T}_i^F(\mathfrak{z})$, $\mathcal{T}^W(\mathfrak{z})$, and $\mathcal{T}_i^M(\mathfrak{z})$ are the exact discrete-time versions (with sampling interval h) of $d^{-1}(q)\mathbf{J}_{i,j}(q)$, $d^{-1}(q)\mathbf{F}_i(q)$, $-d^{-1}(q)\mathbf{N}(q)\mathbf{W}(q)$, and $d^{-1}(q)\mathbf{N}(q)\mathbf{G}_i(q)$, respectively. To prevent clutter, hereafter we abuse the notation by preserving the notation of signals despite the sampling process.

To deal with noise, we follow the stochastic perturbation theory of [30], using a small-noise approach; i.e., we assume the variances of \mathbf{E} and \mathbf{R} are, in an appropriate sense, small in comparison to their respective expected values. This enables approximate computations of the expected error norm of an LS estimator with errors in variables based on a first order approximation of the LS estimate, which is asymptotically exact as variances tend to zero (see [30, Theorem 2.8]). Hence, let us make the contribution of noise explicit by denoting $\tilde{\mathbf{E}} := \mathbf{E} + \mathbf{E}_w$ and $\tilde{\mathbf{R}} := \mathbf{R} + \mathbf{R}_w$, where \mathbf{E}, \mathbf{R} are the noise-free regressor and residual matrices, $\mathbf{E}_w, \mathbf{R}_w$ are the perturbations to these matrices exclusively by noise, and $\tilde{\mathbf{E}}, \tilde{\mathbf{R}}$ are the respective perturbed matrices. From (10), let \mathbf{R}_{NL} be the component of the regressor coming from the second-order terms. It holds that

$$\mathbf{R} = \mathbf{E} \mathbf{f} + \mathbf{R}_{\text{NL}},$$

while the noisy least squares estimation gives

$$\tilde{\mathbf{E}} \hat{\mathbf{f}} = \tilde{\mathbf{R}}.$$

Thus, the following expression relates the fault estimate and the actual fault:

$$\hat{\mathbf{f}} = \tilde{\mathbf{E}}^\dagger (\mathbf{E} \mathbf{f} + \mathbf{R}_w) + \tilde{\mathbf{E}}^\dagger \mathbf{R}_{\text{NL}}, \quad (11)$$

The first-order approximation of the estimate $\hat{\mathbf{f}}$, denoted $\check{\mathbf{f}}$, is given by [30]³

$$\check{\mathbf{f}} = \mathbf{f} + \mathbf{E}^\dagger \mathbf{R}_{\text{NL}} - \mathbf{E}^\dagger (\mathbf{R}_w - \mathbf{E}_w \mathbf{f}) - ((\mathbf{E}^\top \mathbf{E})^{-1} \mathbf{E}_w^\top \mathbf{P}_\perp - \mathbf{E}^\dagger \mathbf{E}_w \mathbf{E}^\dagger) \mathbf{R}_{\text{NL}}, \quad (12)$$

where $\mathbf{P}_\perp := \mathbf{I} - \mathbf{E} \mathbf{E}^\dagger$ is the annihilator (projection) matrix.

Finally, a special type of system norm shall be useful:

¹If the signals are in discrete time, h must be a natural number.

²Because our estimation algorithm is performed every time step, a more computationally efficient online fault estimation implementation should use the recursive least-squares algorithm, see, e.g., [32, Sec. 17.6]. For analysis, we use the standard least-squares solution throughout the text.

³Eq. (12) combines expressions from Sections 3.1 and 3.4 of [30] which give the first-order approximations of, respectively, the perturbed pseudo-inverse and the ordinary-least-squares (OLS) estimate under zero-mean perturbation. The combination is necessary because of the term \mathbf{R}_{NL} , which is nonzero in general.

Definition 3.3 (\mathcal{H}_∞^F norm). Let $\mathcal{T}(q)$ be the transfer function of a BIBO stable system. Its \mathcal{H}_∞^F norm is defined as $\sup_{\omega \in \mathbb{R}} \|\mathcal{T}(j\omega)\|_F$ if it is a continuous-time system, and as $\sup_{\omega \in [0, 2\pi]} \|\mathcal{T}(\exp(j\omega))\|_F$ if it is discrete-time.

It is easy to see that the \mathcal{H}_∞^F metric enjoys the same norm properties as the Frobenius norm, and hence it is itself a norm. It is also submultiplicative thanks to the submultiplicativity of the Frobenius norm. It can be computed by taking the square root of the \mathcal{H}_∞ norm of the system $\sum_{i,j} \mathcal{T}_{ij}^* \mathcal{T}_{ij} = \text{tr}(\mathcal{T}^* \mathcal{T})$.

Let η_F and η_W be the \mathcal{H}_∞^F norms of the transfer functions \mathcal{T}^F and \mathcal{T}^W , respectively, and denote $\eta^2 := \eta_F^2 + \eta_W^2$. In addition, let $\|\mathcal{T}\|_{\infty, \infty}$ be the peak-to-peak gain of \mathcal{T} , and denote by $A := \max_{i,j} \|\mathcal{T}_{i,j}^J\|_{\infty, \infty}$. Our main performance result follows.

Theorem 3.2 (Performance characterization: bias and variance). Consider system (1) satisfying Assumption 2.1 and further assume it is BIBO stable. Let residual r and regressors \mathbf{e} be generated by BIBO stable filters $d^{-1}(q)\mathbf{N}(q)\mathbf{L}(q)\mathbf{z}$ and $d^{-1}(q)\mathbf{M}(q)\mathbf{z}$, respectively. Furthermore, let s_i be the i -th largest singular value of \mathbf{E} . Then, for any N , any fault \mathbf{f} and any time $t \geq Nh$, the first-order approximation of the fault estimate $\check{\mathbf{f}}$ provided by (12) satisfies

$$\text{Bias:} \quad \left| \mathbb{E}(\check{\mathbf{f}}(t) - \mathbf{f}) \right| \leq A\sqrt{Nm} |\mathbf{f}|^2 s_m^{-1} \left\| \begin{bmatrix} \boldsymbol{\xi}^T & \mathbf{z}^T \end{bmatrix} \right\|_\infty =: B, \quad (13a)$$

$$\text{Variance:} \quad \text{tr}(\text{Var}(\check{\mathbf{f}}(t) - \mathbf{f})) \leq \sigma^2 \left(2(|\mathbf{f}|^2 + 1) \eta^2 \sum_{i=1}^m s_i^{-2} + B^2 \eta_F^2 \left(2 \sum_{i=1}^m s_i^{-2} + s_m^{-2} \right) \right). \quad (13b)$$

Proof: See Section 7.2. ■

Since A is the maximal peak-to-peak norm among $\mathcal{T}_{i,j}^J(q) = d(q)^{-1} \mathbf{J}_{i,j}(q)$, Theorem 3.2 with Lemma 3.1 give the following corollary:

Corollary 3.1 (Bias and variance when faults multiply only known signals). Given the premises of Theorem 3.2, suppose additionally that $\mathbf{H}_i'(q) \equiv \mathbf{0}$ for all $i = 1, \dots, m$. Then, more simply, $\left| \mathbb{E}(\check{\mathbf{f}}(t) - \mathbf{f}) \right| = 0$ and $\mathbb{E}(|\check{\mathbf{f}}(t) - \mathbf{f}|^2) \leq (|\mathbf{f}|^2 + 1) \eta^2 \sigma^2 \sum_{i=1}^m s_i^{-2}$.

The involved expressions (13a) and (13b) require some unpacking. We start by the simpler case of Corollary 3.1, where no errors due to nonlinearities occur. Compare it with ordinary least squares (OLS), where the linear model $\mathbf{R} = \mathbf{E}\mathbf{f} + \mathbf{v}$, with i.i.d. v_i with zero mean and variance σ^2 , gives an error variance asymptotically equal to $\sigma^2(\mathbf{E}^T \mathbf{E})^{-1}$. Hence, its trace, which is the expected value of the squared Euclidean norm of the error, converges to $\sigma^2 \sum_{i=1}^m s_i^{-2}$. Our result is the same, but amplified by the factor $(1 + |\mathbf{f}|^2)$ and further scaled by the squared \mathcal{H}_∞^F norm of the systems that filter the white noise \mathbf{w} . The \mathcal{H}_∞^F -norm essentially measures the amplification/attenuation of noise that the systems provide, but also accounts for the added autocorrelation generated by filtering white noise. The amplification factor $(1 + |\mathbf{f}|^2)$ is due to the errors-in-variables nature, and is consistent with the observations of, e.g., [30, Sec. 3.4].

Now, let us inspect the bias term of (13a). As expected from the first-order approach to the estimation problem, it increases with the square of fault magnitude, scaled by the peak-to-peak norm of associated systems $d(q)^{-1} \mathbf{J}_{i,j}(q)$ which map the unmeasured variables into the residual, and the magnitude of the signals involved. It also increases with the number of faults m . However, it *decreases* with s_m/\sqrt{N} , which can be seen as an effective regressor richness measurement, see Remark 3.3.

Returning to the explanation of the total variance bound, (13b) includes a second term, which is the squared bias given in (13a) multiplied by the factor $\eta_F^2 (2 \sum_{i=1}^m s_i^{-2} + s_m^{-2}) \leq \eta_F^2 (2m + 1) s_m^{-2}$. Three observations are in order: (i) this term is the most negligible when faults are small, given the factor $|\mathbf{f}|^4$; (ii) this term is essentially a multiple of s_m^{-4} , while the previously discussed term is a multiple only of s_m^{-2} ; thus, a large value of s_m should also contribute to the low importance of this term; and (iii) this term does increase linearly with N , which is not the case for the previous term.

Remark 3.2 (Relation to Gauss–Newton method for nonlinear least squares). The problem of finding \mathbf{f}_i given \mathbf{z} using (5) can be seen as a nonlinear least-squares (NLS) problem, perturbed by $\boldsymbol{\xi}$. Therefore, using the linear-least-squares approximation in (9) is similar to performing one step of the Gauss–Newton

method for the NLS problem involved. For this reason, a natural extension of the fault estimation approach proposed here is to iteratively (i) compute residual and regressor generators using Theorem 3.1; (ii) run the fault estimation algorithm using Eq. (9) up to a sufficiently long time T ; update the nominal polynomial matrices using $[H(q) \ L(q)] \leftarrow [H(q) \ L(q)] + \sum_i \hat{f}_i(T)[H'_i(q) \ L'_i(q)]$. Under similar convergence conditions as the nominal Gauss-Newton method for NLS (see, e.g., [32, Section 17.4]), the successive incremental fault estimates and residuals converge to zero, and the total fault is the cumulative sum of the incremental estimates. It is important to highlight that, while this process involves switching the dynamics of the generators in real time, it does not hinder their stability, as the poles, dictated by $d(q)$, are unchanged.⁴

Remark 3.3. (Effective singular values) Observation (iii) above may give the impression that, contrary to intuition, more data can actually have a negative impact in the estimation performance, but this is not correct. Generally, $s_{\min}(E)$ grows linearly with \sqrt{N} . E.g., it is easy to verify that the singular values of $[E \ E \ \cdots E]$, where E is repeated k times, are \sqrt{k} times the singular values of E . Hence, when $e(t)$ is periodic, $\lim_{N \rightarrow \infty} \mathbb{E}(|\check{f}(t) - f|^2) = 0$, provided E is full-rank. This asymptotic result does not hold for the bias in (13a): longer experiments cannot cancel the bias introduced by a linear model not properly representing the relation between residual and regressors, which is a known fact for nonlinear regression. Therefore, one can call s_m/\sqrt{N} an effective singular value that measures the richness of data irrespective of its size.

Remark 3.4. (Signal-to-noise ratio) Theorem 3.2 is an asymptotic result using a small-noise assumption. In [30], a metric is suggested to assess whether the first-order approximation for least squares is accurate. In our case, it has the following upper bound (see Section 7.3 for the derivation):

$$c \leq \sqrt{2N} s_m^{-1} \sigma \gamma_F. \quad (14)$$

Due to the asymptotic arguments that justify the theory in [30], this metric must be significantly smaller than 1 for Theorem 3.2 to have practical meaning. Thus, the inverse c^{-1} can be interpreted as a signal-to-noise ratio (SNR), where the signal is measured by the effective richness s_m/\sqrt{N} and the noise by $\sigma \gamma_F$.

Theorem 3.2 indicates the main drivers of estimation performance, among which the smallest singular value of E is the only one depending on real-time deterministic signals. This is the starting point for the experiment design problem we present in Section 4.

3.3 The case of time-varying faults

The approach outlined in this section has only been analyzed for constant faults. However, it can be applied to time-varying faults of the form $f_i = \sum_j p_{ij} \phi_{ij}(t)$, where p_{ij} are unknown parameters and $\phi_{ij}(t)$ are known signals. Its application needs a slight modification, which we explain next.

To accommodate the use of time signals and the q operator, we introduce a notational convention. Throughout this text, the pre-multiplication of a signal with another signal shall be understood as an operator. That is, $a(t)b(t)$ is the result of applying the operator that multiplies signal b by signal a . The main observation is that multiplication is commutative but differentiation does not commute with multiplication. That is, in continuous time, for signals a and b , $ab = ba$ but $aqb \neq qa b$. Hence, we shall always read $abc := a(b(c))$ as usual when applying operators, where a, b, c might be either signals to be multiplied or (differentiation/shift) operators.

With this operator approach, the following identities are important:

- (i) *continuous-time*: for signals a and b , because $q(ab) = (qa)b + aqb$, we write

$$qa = \dot{a} + aq. \quad (15)$$

- (ii) *discrete-time*: for sequences a and b , because $q(ab) = (qa)(qb)$, we write

$$qa = a(k+1)q \quad (16)$$

⁴Stability of switched linear systems can be assessed by finding a common Lyapunov function for all systems. If all modes have the same poles, they admit a state-space realization with the same A matrix; since the Lyapunov function depends only on A , stability is retained with switching zeros.

Modulo the order of operations in the second order terms, one can see that Lemma 3.1 still holds in the time-varying fault case; the only difference is that the second-order term in (5) becomes

$$N(q) \sum_{i,j} f_i(t) \mathbf{H}'_i(q) \mathbf{H}^\dagger(q) f_j(t) \mathbf{H}'_j(q).$$

Therefore, the linear approximation still holds, and, as such, the estimation performance has the same essential characteristics as the constant-fault presented above.

The modification is required on the implementation side. It must remain possible to generate the regressor signals from measured signals and signals that are known in advance. In the time-varying case, the known signals are \mathbf{z} , and, since ϕ_{ij} is known a priori, potentially higher-order derivatives of ϕ_{ij} can be also used. Since the regressors come from the linear term in (5):

$$d^{-1}(q)N(q) \sum_i f_i(t) \mathbf{G}_i(q) = \sum_{ij} p_{ij} d^{-1}(q)N(q) \phi_{ij} \mathbf{G}_i(q),$$

we must be able to implement the regressors

$$d^{-1}(q)N(q) \phi_{ij} \mathbf{G}_i(q).$$

To obtain a realizable filter that avoids derivatives of the measurements, we can simply apply the operations (i) or (ii) above repeatedly so that the regressor ends with the form $d^{-1}(q)N(q)\mathbf{G}'_{ij}(q)\phi'_{ij}$, where ϕ' may contain higher-order derivatives or time advances. When building the regressor, one needs to perform multiplications of \mathbf{z} with ϕ'_{ij} accordingly, and then feed each of these signals to the corresponding LTI filter $d^{-1}N(q)\mathbf{G}'_{ij}(q)$.

This approach is better understood with an example.

Example 3.1. Consider a single fault of the form $f(t) = p_1 + p_2 t$, and let

$$\mathbf{G}(q) = \begin{bmatrix} q^2 + q + 1 & 1 \end{bmatrix}.$$

The only non-trivial element of $\mathbf{G}(q)$ is the first. The two basis functions are $\phi_1(t) = 1$ (constant) and $\phi_2(t) = t$. Applying identity (i), we get $tq = qt - 1$ and $tq^2 = (qt - 1)q = q(qt - 1) - q = q(qt - 1) - q = q^2 t - 2q$. Therefore, $t(q^2 + q + 1) = q^2 t - 2q + qt - 1 + t = (q^2 + q + 1)t - 2q - 1$. Hence, $\phi'_1(t) = 1$, $\phi'_2(t) = t$, and

$$\mathbf{G}'_{11}(q) = \begin{bmatrix} q^2 + q + 1 & 1 \end{bmatrix}, \quad \mathbf{G}'_{12}(q) = \mathbf{0}, \quad \mathbf{G}'_{21}(q) = \begin{bmatrix} -2q - 1 & 0 \end{bmatrix}, \quad \mathbf{G}'_{22}(q) = \begin{bmatrix} q^2 + q + 1 & 1 \end{bmatrix}.$$

Given these matrices, the regressor for p_1 is generated by $d^{-1}(q)N(q)\mathbf{G}'_{11}(q)\mathbf{z}$, while the regressor for p_2 is generated by $d^{-1}(q)N(q)\mathbf{G}'_{21}(q)\mathbf{z} + d^{-1}(q)N(q)\mathbf{G}'_{22}(q)t\mathbf{z}$.

4 Input design for active estimation

In this section, we describe an optimization problem that gives the input design that maximizes the fault estimator performance, followed by our solution approach.

4.1 Optimization problem

Theorem 3.2 tells that the performance of the fault estimator is inversely proportional to $s_{\min}(\mathbf{E})$. Thus, it makes sense to design an input \mathbf{u} that maximizes the minimum singular value of \mathbf{E} , a metric we shall denote as *richness*. Unfortunately, the regressor signal \mathbf{e} is affected by all signals entering the system, including the potential faults \mathbf{f} . Nevertheless, we chose to design the input assuming the nominal case, i.e., $\mathbf{f} = \mathbf{0}$, $\mathbf{w} \equiv \mathbf{0}$, $\mathbf{d} \equiv \mathbf{0}$, where \mathbf{d} is the external disturbance within the partition of ξ . Under this assumption, \mathbf{e} is a function of \mathbf{u} .

Recall that the signals for the estimation are sampled with period h . We shall hereafter assume that an exact discrete-time version $\mathcal{T}_Y(\mathfrak{z})$ of the nominal system of (1) is available, i.e., $\mathbf{y} = \mathcal{T}_Y(\mathfrak{z})\mathbf{u}$. Then, by letting $\mathcal{T}_M(\mathfrak{z})$ be the exact discrete-time representation of $d^{-1}(q)\mathbf{M}(q)$, we have

$$\mathbf{e} = \mathcal{T}(\mathfrak{z})\mathbf{u} := \mathcal{T}_M(\mathfrak{z}) \begin{bmatrix} \mathcal{T}_Y(\mathfrak{z}) \\ \mathbf{I} \end{bmatrix} \mathbf{u}. \quad (17)$$

Before writing down the optimization problem, we must impose a few design choices. First, we remark that a finite-horizon experiment design problem is ill-advised, unless the initial state of the system can be ensured to be zero. This is due to the uncontrollable effect of the unknown initial state in the richness of e . Instead, because the system and filter are stable, it is more convenient to assume that the experiment runs for sufficient long time so that the contribution of the initial state vanishes. Therefore we suggest a moving-horizon approach: we want that $U_{[t-(N-1)h, t]}^h$ maximizes the richness $s_{\min}(E_{[t-(N-1)h, t]}^h)$ for sufficiently large t . This asymptotic richness exists if e is asymptotically periodic, as we show next:

Proposition 4.1 (Asymptotic richness for periodic inputs). *If e is asymptotically N -periodic, then the limit $\lim_{t \rightarrow \infty} s_{\min}(E_{[t, t+N-1]})$ exists.*

Proof: See Section 7.5. ■

Because $\mathcal{T}(z)$ is a linear operator, one could always increase the magnitude of U to increase $s_{\min}(E)$. To make the problem well-posed (and realistic), we assume that U has to be constrained to some compact and bounded set \mathcal{U} . We shall further assume that \mathcal{U} is convex. For simplicity, assume further that $h = 1$. We can define the optimal input design problem as follows:

Definition 4.1 (Optimal input for asymptotic richness). Consider (17) and associated system (1). Assume u is an N -periodic discrete-time signal and denote the *asymptotic richness* by $J_{\infty}(u) := \lim_{t \rightarrow \infty} s_{\min}(E_{[t, t+N-1]})^2$. An N -periodic input signal u^* in a given closed bounded convex subset \mathcal{U} is said to be *optimal for fault estimation* if $u^* = \arg \max_{u \in \mathcal{U}} J_{\infty}(u)$.

Hereafter, let us use the shorthand $U := U_{[0, N-1]}$, and let $\bar{u} \in \mathbb{R}^{Nn_u}$ be the column vector obtained by stacking the columns of U . Furthermore, let $\bar{\mathcal{U}}$ be its associated convex constraint set; since u is N -periodic by definition, there exists a natural bijection between $\bar{\mathcal{U}}$ and \mathcal{U} of Def. 4.1. Consider now a minimal state-space realization of $\mathcal{T}(z)$ and its matrices A, B, C, D , and define the following matrices for all $i = 1, \dots, N$:

$$\begin{aligned} P_i &:= [CA^{i-2}B \quad \dots \quad CB \quad D \quad \mathbf{0}_{m \times n_u(N-i)}], \\ P_N &:= [A^{N-1}B \quad A^{N-2}B \quad \dots \quad AB \quad B], \\ P'_i &:= P_i + CA^{i-1}(I - A^N)^{-1}P_N. \end{aligned} \tag{18}$$

Our main optimization result follows:

Theorem 4.1 (Objective function and first-order information). *Consider (17), (18), and associated system (1), and assume $\mathcal{T}(z)$ is BIBO stable. Let λ_{\min}, v_{\min} be a smallest eigenvalue and associated unitary eigenvector of the p.s.d. matrix $Q(\bar{u}) := \sum_{i=1}^N P'_i \bar{u} \bar{u}^T P'_i$, and define the expressions*

$$\text{Objective function:} \quad J(\bar{u}) := \lambda_{\min}(Q(\bar{u})), \tag{19a}$$

$$\text{Subgradient:} \quad g(\bar{u}) := 2 \sum_{i=1}^N v_{\min}^T P'_i \bar{u} P'_i v_{\min}. \tag{19a}$$

Let u be the signal built by the periodic repetition of the elements in \bar{u} , i.e., $u_i(k) = \bar{u}_{i+n_u(k \bmod N)}$ for all $i = 1, \dots, n_u, k \in \mathbb{N}$; here, mod denotes the remainder operator. Then, $J(\bar{u}) = J_{\infty}(u)$ (Def. 4.1) and $g(\bar{u})$ is a subgradient of $J(\bar{u})$.

Proof: See Section 7.5. ■

Theorem 4.1 enables us to translate the optimal input design of Definition 4.1 into a finite-dimensional optimization problem. Moreover, it provides a subgradient, which is needed as the minimal eigenvalue function is not everywhere differentiable. With the first-order information, convex optimization methods such as projected subgradient ascent can be used.

Remark 4.1 (Convex constraints). Common choices for \mathcal{U} are those described by the following inequalities:

- Component bounds, i.e., $u_i^L \leq u_{i,t} \leq u_i^U, \forall t \in \{1, 2, \dots, N\}, i \in \{1, 2, \dots, m\}$;
- Component energy bounds, i.e., $u_i^T u_i \leq E_i, \forall i \in \{1, 2, \dots, m\}$;

- Total energy bounds, i.e., $\sum_{i=1}^m \mathbf{u}_i^\top \mathbf{u}_i \leq E$.

Nominal output constraints are also possible, but they may need to be robustified against unknown initial state and disturbance.

Remark 4.2. Because the function $f(\mathbf{E}) := \lambda_{\min}(\mathbf{E}^\top \mathbf{E})$ satisfies $f(a\mathbf{E}) = a^2 f(\mathbf{E})$, the solution to Problem (19) is always on the boundary of \mathcal{U} , provided \mathcal{U} is star-shaped at the origin. This is the case, for example, of non-empty convex sets containing the origin. By continuity of f , this implies that it has at least one maximum and one saddle point at the boundary, unless f is constant at the boundary of \mathcal{U} .

4.2 Optimization solution

We propose Algorithm 1, a projected subgradient ascent algorithm for an efficient local-optimal solution to the input design problem. In this method, at each iteration a subgradient of the objective function is computed, a step is taken in such direction, the solution is projected onto the feasible set, and the iteration repeats.

Algorithm 1 Computation of optimal input for (19)

Input: $\mathbf{P}_i, i = 1..N, \mathcal{Q}, \tau, L$
Output: $\bar{\mathbf{u}}$

```

1:  $k \leftarrow 0$ 
2:  $\bar{\mathbf{u}} \leftarrow \text{project}_{\mathcal{Q}}(\text{rand}(Nn_u))$ 
3: while true do
4:    $\mathbf{M} \leftarrow \sum_{i=1}^N \mathbf{P}_i \bar{\mathbf{u}} \bar{\mathbf{u}}^\top \mathbf{P}_i^\top$ 
5:    $(\lambda_{\min}, \mathbf{v}_{\min}) \leftarrow \text{eig}_{\min}(\mathbf{M})$ 
6:    $\mathbf{g} \leftarrow \sum_{i=1}^N \mathbf{v}_{\min}^\top \mathbf{P}_i \bar{\mathbf{u}} \mathbf{P}_i^\top \mathbf{v}_{\min}$ 
7:    $\bar{\mathbf{u}}_{\text{new}} \leftarrow \text{project}_{\mathcal{Q}}(\bar{\mathbf{u}} + \frac{L}{L+k} \tau \mathbf{g}^\top)$ 
8:   if  $|\bar{\mathbf{u}}_{\text{new}} - \bar{\mathbf{u}}|_2 < \varepsilon_u$  or  $|\lambda_{\min}|/k < \varepsilon_\lambda$  then
9:      $\bar{\mathbf{u}} \leftarrow \bar{\mathbf{u}}_{\text{new}}$ 
10:    return
11:  end if
12:   $k \leftarrow k + 1$ 
13:   $\bar{\mathbf{u}} \leftarrow \bar{\mathbf{u}}_{\text{new}}$ 
14: end while
```

Let us explain Algorithm 1 relying on Theorem 4.1. After $\bar{\mathbf{u}}$ is initialized,⁵ the main subgradient ascent loop begins. Lines 4–6 use (19a) to compute the subgradient \mathbf{g} ; next, Line 7 computes the projected subgradient update with step size parameters τ and L , using $\text{project}_{\mathcal{Q}}$ to perform an orthogonal projection onto the convex set \mathcal{U} ; finally, Line 8 presents the two stopping criteria: either the change in the Euclidean norm of the gradient below the tolerance ε_u , or the change in the ergodic mean of the objective function λ_{\min} is smaller than the tolerance ε_λ . Note that, for many practical convex set descriptions, projection is an easy computation; e.g., in the unit ball case, $\mathcal{Q} = \{\mathbf{u} \mid \|\mathbf{u}\|_2 \leq 1\}$, it is computed by normalizing \mathbf{u} to unit Euclidean norm when bigger than 1. In addition, we apply a decaying step size of the form $\frac{L}{L+k} \tau$, which has generally improved convergence properties. For the same reason, the ergodic mean is used as stopping criterion.

Since we have presented a local method, we have no guarantee that the global optimum is attained. Thus, it is sensible to estimate an upper bound to the global optimal value, thereby obtaining an over-approximation of the optimality gap. This can be achieved by semidefinite relaxation. Let us assume for simplicity that the constraint set has the form $\mathcal{Q} = \{\bar{\mathbf{u}} \mid \bar{\mathbf{u}}^\top \mathbf{S}_i \bar{\mathbf{u}} \leq 1, i = 1, \dots, N_c\}$, where N_c is the number of constraints and $\mathbf{S}_i \succeq \mathbf{0}$. This is valid for all constraints laid out in Remark 4.1. We can assert the following.

Proposition 4.2 (Semidefinite relaxation of the optimal experiment design). *Let λ^* be the value of (19),*

⁵We have chosen to initialize $\bar{\mathbf{u}}$ to a pseudorandomly generated vector, but other initialization is possible, as long as it is not the zero vector, which is a global minimum with zero gradient.

where $\mathcal{U} = \{\bar{\mathbf{u}} \mid \bar{\mathbf{u}}^\top \mathbf{S}_i \bar{\mathbf{u}} \leq 1, i = 1, \dots, N_c\}$, where $\mathbf{S}_i \succeq \mathbf{0}$. Then, the value λ^{SDP} of the SDP

$$\begin{aligned} & \max_{\bar{\mathbf{U}}, \lambda} \quad \lambda \\ & \text{subject to} \quad \sum_{i=1}^N \mathbf{P}'_i \bar{\mathbf{U}} \mathbf{P}'_i{}^\top \succeq \lambda \mathbf{I}, \\ & \quad \text{tr}(\mathbf{S}_i \bar{\mathbf{U}}) \leq 1, \quad \forall i = 1, \dots, N_c, \\ & \quad \bar{\mathbf{U}} \succeq \mathbf{0} \end{aligned} \tag{20}$$

is larger than or equal to λ^* . Furthermore, if $\text{rank}(\bar{\mathbf{U}}) = 1$, then $\lambda^{\text{SDP}} = \lambda^*$.

Proof: See Section 7.5 ■

We must remark that the SDP (20) has as its main variable a square Nn_u -dimensional matrix. With large values of N , it may be challenging to solve it reliably.

5 Numerical results

This section presents a series of numerical experiments to illustrate the performance of our method. A Matlab implementation of our methods, including the scripts to reproduce these results, can be found in <https://gitlab.tudelft.nl/ggleizer/aemf>.

5.1 Setup

We consider a continuous-time linearized pendulum-cart system, adapted from [33, Eqs. (3) and (6)], with three parametric faults and one external (unmeasured) disturbance described by the following nominal DAE of the form (1):

$$\begin{bmatrix} (M + M_p)q + b & M_p l q^2 & (M + M_p)g \\ M_p l q & (I + M_p l^2)q^2 + M_p g l & -M_p g l \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \xi + \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -0.1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \mathbf{z} + \begin{bmatrix} 0 \\ \mathbf{I} \end{bmatrix} \mathbf{w} = \mathbf{0},$$

where M is the cart mass, M_p, l, I are the nominal pendulum mass, length, and moment of inertia, respectively, b is the nominal friction coefficient, and g is the acceleration of gravity.⁶ The linearization is performed at downright pendulum position, hence the system is open-loop stable. The measured signals $\mathbf{z}(t) \in \mathbb{R}^4$ can be partitioned as $[\mathbf{y}^\top \quad \mathbf{u}^\top]^\top$ where $\mathbf{y}(t) \in \mathbb{R}^2$ are the car velocity and pendulum angle measurements, and $\mathbf{u}(t) \in \mathbb{R}^2$ are the force applied to move the cart horizontally and a torque applied to the base of the pendulum, in the respective order. Likewise, the unmeasured signals $\xi(t) \in \mathbb{R}^4$ can be partitioned as $[\mathbf{x}^\top \quad \mathbf{d}^\top]^\top$ where $\mathbf{x}(t) \in \mathbb{R}^3$ are the internal states (car velocity, pendulum angle, and pendulum angular velocity, in order), and $\mathbf{d}(t) \in \mathbb{R}^1$ is the unmeasured slope in which the cart is travelling, being zero when the car moves horizontally.

The fault matrices are

$$\begin{aligned} \mathbf{H}'_1(q) &= \begin{bmatrix} b & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{H}'_2(q) = \begin{bmatrix} M_p q & M_p l q^2 & M_p \\ M_p l q & M_p g l + M_p l^2 q^2 & -M_p g \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{H}'_3(q) = \mathbf{0}, \\ \mathbf{L}'_1(q) &= \mathbf{L}'_2(q) = \mathbf{0}, \quad \mathbf{L}'_3(q) = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -0.1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

⁶See [33] for nominal values and units.

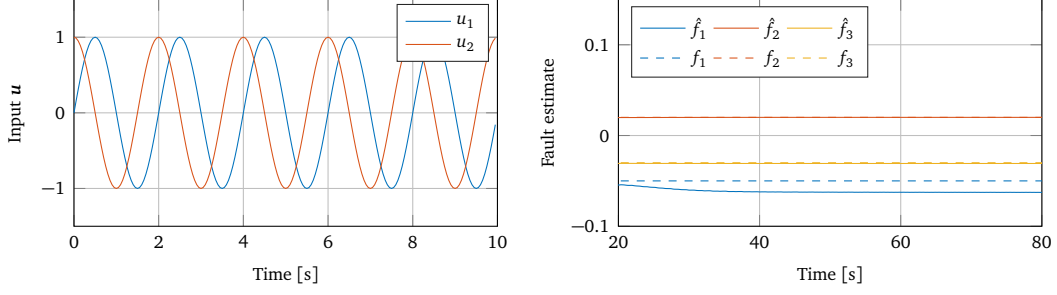


Figure 2: Input (left) and fault estimation results (right) for the first scenario, with simple sinusoidal input and zero noise.

Fault f_1 represents a multiplicative change in the friction of the system; fault f_2 represents a multiplicative change in the mass of the pendulum; and fault f_3 represents a multiplicative factor applied to the battery power of the cart, affecting both force and torque, simultaneously.

5.2 Fault estimator and input design

The matrix $H(q)$ is verified to satisfy Assumption 2.1 with the help of Remark 3.1, where $H^\dagger(q)$ has degree 1. We then solve the equations in Theorem 3.1, which, for degree 2, has the unique solution (up to a scalar)

$$N(q) = [-0.065 \quad 0.026q \quad 0.026 \quad -0.024 - 0.065q \quad 0.994 + 0.026q^2],$$

from which $M(q)$ can be readily derived following (8), having degree 2. To design the denominator $d(q)$, whose degree must be bigger than or equal to 2 for a proper realization, we notice that the nominal system has a peak frequency response at 5.58 rad/s, after which it decays rapidly. Therefore, we choose the poles to be higher than this frequency, and set $d(q) = (q + 10)(q + 20)/200$. For the fault estimator, we have chosen the sampling interval of $h = 0.05$ and estimation window of $N = 400$.

For the input design, we have picked $N = 40$ and chose to bound the 2-norm of each input signal by $\sqrt{N/2}$, so that it has the same 2-norm as a sinusoidal of amplitude 1 and period N .⁷ We have run Alg. 1 with step size parameters $\tau = 10$ and $L = 50$, and tolerances $\varepsilon = 10^{-3}$ and $\varepsilon_\lambda = 10^{-5}$. The algorithm stopped after 3961 iterations, taking only 0.20 second, and yielding an optimal value of 0.212. We used CVX [34] in Matlab to solve the SDP in Prop. 4.2, which took 0.85 second to compute the upper bound to the global optimum of 0.451. This means that the global optimal singular value is between the square roots 0.460 and 0.672.

5.3 Results

Now we present the estimation results for multiple simulations, including small constant faults, large constant faults, and time-varying faults.

5.3.1 Small faults

We have simulated four scenarios. In all of them, the faults are $f_1 = -0.05$, $f_2 = 0.02$, $f_3 = -0.03$, and a disturbance $d(t) = \frac{5\pi}{180} \sin(\pi t)$, corresponding to a terrain oscillating between -5 and 5 degrees of slope. In the first two scenarios we emulate passive fault estimation by setting the input $u(t) = [\sin(\pi t) \quad \cos(\pi t)]^T$, and set noise to zero. In the second scenario, we introduce zero-mean Gaussian white noise with variance 1 through pseudo-randomly generated numbers. In the third and fourth scenarios, we use the optimal input designed using Algorithm 1 for the noiseless case and the noisy output case, respectively.

In the first scenario, without noise, the fault estimation attains a small estimation error, with a more significant bias in the estimation of the friction fault f_1 , see Fig. 2. The non-zero error is in line with Corollary 3.1, since

⁷It is not necessary that the time window for estimation is the same as that of the input signal, provided the former is an integer multiple of the latter to ensure periodicity.

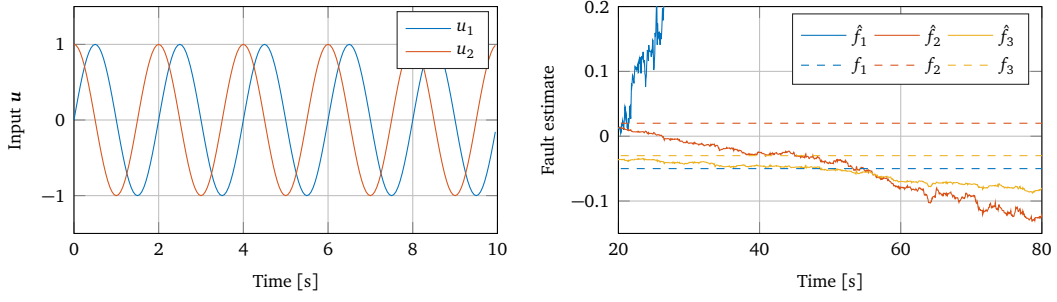


Figure 3: Input (left) and fault estimation results (right) for the second scenario, with simple sinusoidal input and noise.

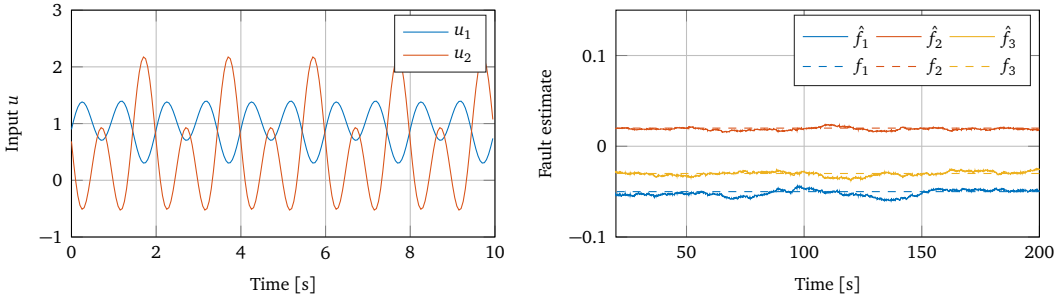


Figure 4: Input (left) and fault estimation results (right) for the third scenario, with optimized input and noise.

$H'_i(q) \neq 0$, but it is still relatively small: $|\hat{f}(t_f) - f|/|f| = 0.2$, where t_f is the final time of the experiment. The performance increases substantially with the optimal input, where the final relative error drops to 0.0134. This is much smaller than the bound of 36.47 obtained by (13a), which indicates that in practice the estimation bias can be much lower than the conservative bound we have devised. In comparison to the noiseless case, it is striking to observe that the accuracy of the estimation is completely unacceptable once noise is introduced (Fig. 3). This is mainly due to the bad excitation properties of the input signal, as illustrated by the estimation performance using the optimized input in Fig. 4, as well as Fig. 5, which shows the minimal singular values of the regressors for both the sinusoidal and optimized inputs. The mean Euclidean norm of the estimation error, using the optimized input, is only about 3.4% of the norm of the fault vector. The mean squared error over the last half of the simulation is $3.3 \cdot 10^{-5}$, orders of magnitude smaller than the bound on the expected squared error given by the first-order-approximation result in Theorem 3.2, which is $5.0 \cdot 10^{-2}$. Finally, to assess the validity of the first-order-approximation, we computed the upper bound for the inverse SNR in (14), obtaining 0.12. Since it is significantly smaller than 1, the first order approximation is deemed accurate.

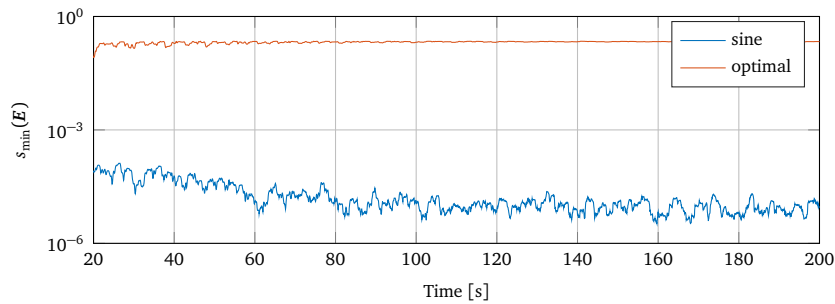


Figure 5: Running minimal singular values of $E^h_{[t-20,t]}$ for the passive sinusoidal input and the optimized input.

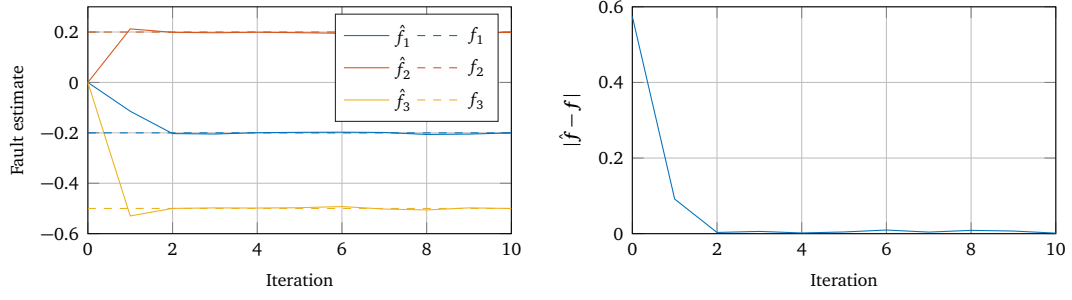


Figure 6: Fault estimation results (left) and estimation error norm (right) for the large fault scenario of Section 5.3.2. Each iteration corresponds to the last estimate after 40 seconds from the last update in the fault estimation filters.

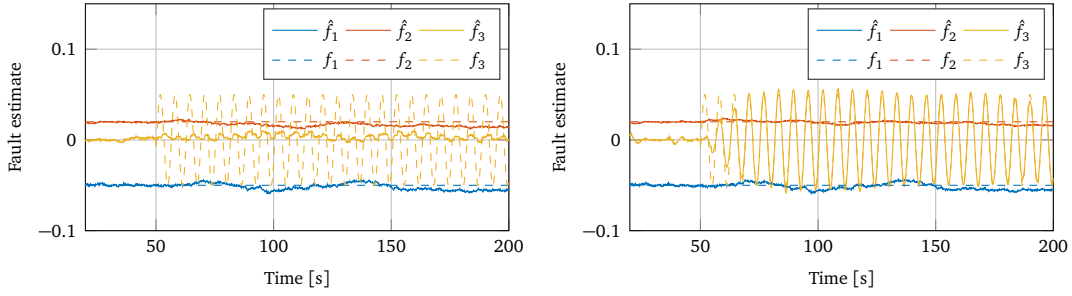


Figure 7: Fault estimation in the time-varying fault case using an estimator that assumes constant faults (left) and the estimation designed for the appropriate time-varying fault signature (right).

5.3.2 Large faults

In the previous subsection, we have seen the performance of our proposed fault estimation scheme for small faults. Now we present the results for significantly larger faults. In this case, the faults are constant with $f_1 = -0.2, f_2 = 0.2, f_3 = -0.5$ and we have followed the procedure suggested in Remark 3.2. The updates on the residual and regressor generators were made every 40 seconds. The same optimal input from Section 5.3.1 was used throughout the simulation. The resulting fault estimation performance is shown in Fig. 6; one may note that convergence is attained after two iterations only, after which the noise effects dominate the estimate variation.

5.3.3 Time-varying fault

In this last set, we simulate one time-varying fault: $f_1(t) \equiv -0.05, f_2(t) \equiv 0.02, f_3(t) = 0.05 \sin(t)$. Following Section 3.3, we compute $G'_{3,j}(q)$. Since in our case study $G_3(q)$ is zero-order, we have only one regressor generator associated with f_3 and $G'_{3,1} = G_3$. Its corresponding input signal is $z \sin(t)$. The optimal input designed in Section 5.3.1 is used in this experiment, and noise variance is kept at 1. Figure 7 shows the difference in fault estimation performance between using the exact same estimator as in Section 5.3.1, i.e., assuming all faults are constant, and using the correct time-varying estimator. Clearly, the performance is far superior with the correct assumption of time-varying fault for estimating f_3 ; in addition, a small improvement in the estimation of f_2 can be observed.

6 Conclusion

We have presented a fault estimation scheme for multiple multiplicative faults using a DAE framework, as well as numerical methods to compute its parameters. The focus is on small parametric faults, which have

small effect in the residual signal, but allow for efficient estimation using a least-squares approach. The fault estimation performance is shown to be highly dependent on the richness of the involved signals, in particular the minimal singular value of a windowed trajectory of the regressor signal, which is highly dependent on the input signal of the plant. This motivated an active estimation approach, by means of an optimal periodic input design, for which we have provided a formulation and an algorithm for its efficient computation. The numerical example, despite being simple, highlights the importance of a properly designed input signal to enable accurate estimation in noisy conditions.

A prototype of the fault estimation method presented here has been successfully implemented in a larger scale realistic use case based on a chip manufacturing machine [35], with a 20-dimensional state space and 8 simultaneous faults. As part of the next steps, we are currently working on the implementation on a more realistic, nonlinear model of this system.

As a potential side contribution of this work, we believe our optimization algorithm can have an impact in other input design applications, such as system identification, active learning, and parameter estimation. There is evidence that the optimization problem (19), despite being non-convex, is well-behaved in the sense that either it does not contain spurious local maxima, or that this is the case with high probability as the problem dimensions increase.

7 Technical proofs

7.1 Proof of Theorem 3.1

Proof of Lemma 3.1: We start by pre-multiplying the system equation (1) with $N(q)(\mathbf{I} - \sum_i f_i \mathbf{H}'_i(q) \mathbf{H}^\dagger(q))$, which is a first-order approximation of the left null space of the system. To simplify notation, we drop the dependency of matrices on the operator q .

$$\begin{aligned} N \left(\mathbf{I} - \sum_i f_i \mathbf{H}'_i \mathbf{H}^\dagger \right) & \left(\mathbf{H} + \sum_i f_i \mathbf{H}'_i \right) \xi \\ & + N \left(\mathbf{I} - \sum_i f_i \mathbf{H}'_i \mathbf{H}^\dagger \right) \left(\mathbf{L} + \sum_i f_i \mathbf{L}'_i \right) \mathbf{z} \\ & + N \left(\mathbf{I} - \sum_i f_i \mathbf{H}'_i \mathbf{H}^\dagger \right) \mathbf{W} \mathbf{w} = \mathbf{0}. \end{aligned}$$

Expanding, we get

$$\begin{aligned} & \left(\mathbf{N} \mathbf{H} + \sum_i f_i \mathbf{N} \mathbf{H}'_i - \sum_i f_i \mathbf{N} \mathbf{H}'_i - \sum_i \sum_j f_i f_j \mathbf{N} \mathbf{H}'_i \mathbf{H}^\dagger \mathbf{H}'_j \right) \xi \\ & + \left(\mathbf{N} \mathbf{L} + \sum_i f_i \mathbf{N} \mathbf{L}'_i - \sum_i f_i \mathbf{N} \mathbf{H}'_i \mathbf{H}^\dagger \mathbf{L} - \sum_i \sum_j f_i f_j \mathbf{N} \mathbf{H}'_i \mathbf{H}^\dagger \mathbf{L}'_j \right) \mathbf{z} \\ & + N \left(\mathbf{I} - \sum_i f_i \mathbf{H}'_i \mathbf{H}^\dagger \right) \mathbf{W} \mathbf{w} = \mathbf{0}. \end{aligned}$$

Since $\mathbf{N} \mathbf{H} = \mathbf{0}$, and $\mathbf{N} \mathbf{L} \mathbf{z} = \mathbf{d}(q) \mathbf{r}$, we obtain

$$\begin{aligned} & - \left(\sum_i \sum_j f_i f_j \mathbf{N} \mathbf{H}'_i \mathbf{H}^\dagger \mathbf{H}'_j \right) \xi + \mathbf{d}(q) \mathbf{r} + \left(\sum_i f_i \mathbf{N} \mathbf{L}'_i - \sum_i f_i \mathbf{N} \mathbf{H}'_i \mathbf{H}^\dagger \mathbf{L} \right) \mathbf{z} \\ & - \left(\sum_i \sum_j f_i f_j \mathbf{N} \mathbf{H}'_i \mathbf{H}^\dagger \mathbf{L}'_j \right) \mathbf{z} - N \left(\sum_i f_i \mathbf{H}'_i \mathbf{H}^\dagger \right) \mathbf{W} \mathbf{w} + \mathbf{N} \mathbf{W} \mathbf{w} = \mathbf{0}. \end{aligned}$$

Finally, let $J_{i,j} := \begin{bmatrix} -NH_i' H_j^\dagger H_j' & -NH_i' H_j^\dagger L_j' \end{bmatrix}$ and $F_i := -NH_i' H^\dagger W$, which are identically zero if $H_i' \equiv \mathbf{0}$ for all i . We arrive at the desired simplified expression:

$$d(q)r = \left(\sum_i f_i NH_i' H^\dagger L - f_i NL_i' \right) z + \sum_{i,j} f_i f_j J_{i,j}(q) \begin{bmatrix} \xi \\ z \end{bmatrix} + \sum_i \mathcal{O}(f_i) F_i(q) w - NWw.$$

■

Now we provide a useful lemma regarding a notion of *symbolic rank* of polynomial matrices.

Lemma 7.1. *Consider a polynomial matrix $H \in \mathbb{R}[q]^{m \times n}$ of degree d . The following statements hold:*

- (i) *There exists a nonzero $v \in \mathbb{R}^n$ such that $H(q)v \equiv \mathbf{0}$ if and only if $\text{blkrow}(H(q)^\top)$ is row-rank deficient;*
- (ii) *There exists a nonzero $v \in \mathbb{R}^m$ such that $v^\top H(q) \equiv \mathbf{0}$ if and only if $\text{blkrow}(H(q))$ is row-rank deficient.*

Proof: We prove item (ii), which immediately gives (i) by transposing the matrices. First, we start by noticing that $\exists v \in \mathbb{R}^n : v^\top H(q) \equiv \mathbf{0}$ iff $\exists v \in \mathbb{R}^n : v^\top H_i = \mathbf{0}$ for all $i = 0, 1, \dots, d$. This, in turn, is equivalent to $\exists v \in \mathbb{R}^n : v^\top \text{blkrow}(H(q)) = \mathbf{0}$, which is the definition of row-rank deficiency. ■

The following Lemma shows that full symbolic rank of a filter gives linearly independent signals provided its input is persistently exciting:

Lemma 7.2. *Consider a polynomial matrix $H \in \mathbb{R}[q]^{m \times n}$ of degree d , where q represents a differential or time-shift operator. Assume $\text{blkrow}(H(q))$ is full-row-rank. Then, if z is PE of order $d + 1$, the output of $H(q)z$ is linearly independent.*

Proof: Suppose $H(q)z$ has a linear dependence. Then there exists a nonzero $v \in \mathbb{R}^m$ such that $v^\top H(q)z(t) = 0$ for all t . By Lemma 7.1, for all nonzero $v \in \mathbb{R}^m$, it holds that $v^\top H(q) \not\equiv \mathbf{0}$. Thus, $a(q)^\top := v^\top H(q)$ is a polynomial vector of degree less than or equal to d . Hence, $a(q)^\top z(t) = 0$ for all t , which can be rewritten as the RHS of (4), contradicting the fact that z is PE of order $d + 1$. ■

Lemmas 3.1, 7.1 and 7.2 allow us to prove Theorem 3.1.

Proof of Theorem 3.1: By Lemma 2.1, $N(q)H(q) \equiv \mathbf{0}$ is equivalent to (7a), and equations (6) and (8) are equivalent. From Lemma 3.1, if Assumption 2.1 holds, then taking $N(q)$ s.t. $N(q)H(q) \equiv \mathbf{0}$ and $M(q)$ according to (6) yields (5). For small f , and taking $T(q) := -N(q)W(q)$, we obtain the approximation (3), fulfilling requirement (b) of Definition 3.2.

Finally, Lemmas 7.1 and 7.2 establish that (7b) implies that requirements (a) and (c) of Definition 3.2 are satisfied. This concludes the proof. ■

7.2 Proof of Theorem 3.2

Let us recall the expression for the first-order approximation of the fault estimation error (12):

$$\check{f} - f = E^\dagger R_{NL} - E^\dagger (R_w - E_w f) - ((E^\top E)^{-1} E_w^\top P_\perp - E^\dagger E_w E^\dagger) R_{NL}.$$

We divide the proof of Theorem 3.2 into two parts: the proof of its bias bound expression (13a), and its total variance bound expression (13b).

Proof of (13a): By assumption, R_w and E_w are zero-mean, and as such, we can state that $\mathbb{E}(\check{f} - f) = E^\dagger R_{NL}$. Thus, its expression can be obtained with $w \equiv \mathbf{0}$. In this case, for any t , (5) gives

$$r(t) = f^\top e(t) + \sum_{i,j} f_i f_j \mathcal{T}_{i,j}^J(z) \begin{bmatrix} \xi(t) \\ z(t) \end{bmatrix}. \quad (21)$$

Hence,

$$|r(t) - f^\top e(t)| \leq \sum_{i,j} |f_i f_j| \left\| \mathcal{T}_{i,j}^J \right\|_{\infty, \infty} \left\| \begin{bmatrix} \xi \\ z \end{bmatrix} \right\|_{\infty}$$

where $\|\mathcal{T}\|_{\infty, \infty}$ is the peak-to-peak gain of \mathcal{T} . Denote by $C := \max_{i,j} \|\mathcal{T}_{i,j}^J\|_{\infty, \infty} \left\| \begin{bmatrix} \xi \\ z \end{bmatrix} \right\|_{\infty}$. We then have

$$|r(t) - f^T e(t)| \leq C \sum_{i,j} |f_i f_j| \leq C m f^T f, \quad (22)$$

where the last inequality follows from Young's inequality. Let $b := C m f^T f$. Then,

$$-b \leq r(t) - f^T e(t) \leq b \implies -b \mathbf{1} \leq R - E f \leq b \mathbf{1},$$

where $\mathbf{1}$ is a vector of ones and the inequality holds element-wise. Hence,

$$|R - E f| \leq b \sqrt{N}. \quad (23)$$

Now, $\hat{f}(t) - f = E^\dagger (R - E f)$. Thus,

$$|\hat{f}(t) - f| = |E^\dagger (R - E f)| \leq \|E^\dagger\|_2 |R - E f| \leq s_{\min}(E)^{-1} b \sqrt{N}, \quad (24)$$

where $\|E^\dagger\|_2 = s_{\min}(E)^{-1}$ comes from the fact that the singular values of A^\dagger are the reciprocals of the singular values of A for any A .

Since E is full rank, $s_{\min}(E)^{-1} < \infty$, and thus the proof is concluded by taking $A = \max_{i,j} \|\mathcal{T}_{i,j}^J\|_{\infty, \infty}$. \blacksquare

To prove Theorem 3.2, we shall use the notion of stochastic norm presented by [30]. For a random matrix A , its stochastic norm is defined by $\|A\|_S^2 = \mathbb{E}(\|A\|_F^2)$. We begin with some useful bounds.

Lemma 7.3. *Let $y = \mathcal{T}(\mathfrak{z})w$ be a discrete-time $m \times n$ LTI w \mathcal{H}_∞^F norm of \mathcal{T} is $\eta < \infty$ and where w_i are independent zero-mean i.i.d. random processes with variance σ^2 . Let $\Sigma \in \mathbb{S}_+^N$ be the auto-covariance matrix defined by $(\Sigma)_{i,j} = \mathbb{E}[y(t)^T y(t - |i - j + 1|)]$. Then, uniformly in N , it holds that $\lambda_{\max}(\Sigma) \leq \eta^2 \sigma^2$.*

Proof: First, we decompose the auto-covariance of y

$$\begin{aligned} \mathbb{E}[y(t)^T y(t - \tau)] &= \mathbb{E} \left[\left(\sum_{i=1}^m \left(\sum_{j=1}^n w_j(t) * h_{i,j}(t) \right) \right)^T \left(\sum_{k=1}^n w_k(t - \tau) * h_{i,k}(t - \tau) \right) \right] \\ &= \mathbb{E} \left[\sum_{i=1}^n \sum_{j=1}^m (w_j(t) * h_{i,j}(t)) (w_j(t - \tau) * h_{i,j}(t - \tau)) \right] \\ &= \sum_{i=1}^n \sum_{j=1}^m \mathbb{E}[(w_j(t) * h_{i,j}(t)) (w_j(t - \tau) * h_{i,j}(t - \tau))], \end{aligned}$$

where $h_{i,j}$ is the impulse response of the entry $\mathcal{T}_{i,j}$, and we have used independence, i.e., $i \neq j \implies \mathbb{E}[w_i(t) w_j(t')] = 0$ for all t, t' , in the third equation. We conclude that $\Sigma = \sum_{i=1}^n \sum_{j=1}^m \Sigma_{i,j}$, where $\Sigma_{i,j}$ is the auto-covariance matrix of $\mathcal{T}_{i,j} w_i$.

Now, consider the auto-covariance function $\rho_{i,j}$ of $\mathcal{T}_{i,j}(\mathfrak{z}) w_i$. According to [36, Chap. 10], the discrete-time Fourier transform $\varsigma_{i,j}$ of $\rho_{i,j}$ satisfies $\varsigma_{i,j}(j\omega) = \sigma^2 |\mathcal{T}_{i,j}(\exp(j\omega))|^2$, where we use the fact that the Fourier transform of the auto-covariance of w_i is constant and equal to σ^2 . By [37, Lemma 4.1], a Toeplitz matrix $\Sigma \in \mathbb{S}_+^T$ satisfies, uniformly in N , $\lambda_{\max}(\Sigma_T) \leq \text{ess sup}_{\omega \in [0, 2\pi]} \varsigma(\exp(j\omega))$. Hence,

$$\lambda_{\max}(\Sigma_T) \leq \sigma^2 \text{ess sup}_{\omega \in [0, 2\pi]} \sum_{i=1}^{n_w} |\mathcal{T}_i(\exp(j\omega))|^2 = \sigma^2 \eta^2,$$

by definition of \mathcal{H}_∞^F norm. \blacksquare

Lemma 7.4. Let γ_F (η_F) and γ_W (η_W) be the \mathcal{H}_2 (\mathcal{H}_∞^F) norms of the transfer functions \mathcal{T}^F and \mathcal{T}^W . Then,

$$\mathbb{E}[\text{tr}(\mathbf{E}_w^T \mathbf{E}_w)] \leq \sigma^2 N \gamma_F^2 \quad (25)$$

$$\lambda_{\max}(\text{Var}(\mathbf{R}_w - \mathbf{E}_w \mathbf{f})) \leq (|\mathbf{f}|^2 + 1) \sigma^2 \eta^2, \quad (26)$$

$$\text{tr Var}(\mathbf{E}_w^T \mathbf{b}) \leq |\mathbf{b}|^2 \sigma^2 \eta_F^2, \quad (27)$$

where $\gamma^2 := \gamma_F^2 + \gamma_W^2$, and $\eta^2 = \eta_F^2 + \eta_W^2$

Proof: Recall that the \mathcal{H}_2 norm γ of an LTI system $\mathbf{y} = \mathcal{T}(z)\mathbf{w}$ satisfies

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^N \mathbb{E}[\mathbf{y}(t)^T \mathbf{y}(t)] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^N \text{tr Var}(\mathbf{y}(t)) = \gamma^2.$$

when \mathbf{w} is an i.i.d. random process with variance \mathbf{I} .

Because the variance $\mathbb{E}[\mathbf{y}(t)^T \mathbf{y}(t)]$ is monotonically increasing with t , this implies that for any t_1 ,

$$\sum_{t=t_1}^{t_1+N} \mathbb{E}[\mathbf{y}(t)^T \mathbf{y}(t)] \leq N \gamma^2. \quad (28)$$

Now, $\mathbb{E}[\text{tr}(\mathbf{E}_w^T \mathbf{E}_w)] = \mathbb{E}[\sum_{i=1}^m (\mathbf{E}_w^i)^T \mathbf{E}_w^i] = \sum_{i=1}^m \mathbb{E}[(\mathbf{E}_w^i)^T \mathbf{E}_w^i] = \sum_{t=N-1}^t \mathbf{e}(t)^T \mathbf{e}(t)$, where \mathbf{E}_w^i is the i -th column of \mathbf{E}_w . Applying (28), while recalling that $\text{Var}(\mathbf{w}) = \sigma^2 \mathbf{I}$, leads to (25).

For $\lambda_{\max}(\text{Var}(\mathbf{R}_w - \mathbf{E}_w \mathbf{f}))$, notice that $\mathbf{r}(t) - \mathbf{e}(t)^T \mathbf{f}$ is the output of the system $\mathcal{T}' := \mathcal{T}^W - \sum_{i=1}^m f_i \mathcal{T}_i^F$. By the triangle inequality, its \mathcal{H}_∞^F norm η' satisfies $\eta' \leq \eta^W + |\mathbf{f}| \eta^F \leq \sqrt{(1 + |\mathbf{f}|^2) \eta^2}$ (by Cauchy-Schwarz). The bound (26) is then obtained by applying Lemma 7.3.

For the last bound,

$$\text{tr}(\text{Var}(\mathbf{E}_w^T \mathbf{b})) = \sum_{i=1}^m \text{Var}(\mathbf{b}^T \mathbf{E}_w^i) = \sum_{i=1}^m \mathbf{b}^T \Sigma_i \mathbf{b} = \mathbf{b}^T \left(\sum_{i=1}^m \Sigma_i \right) \mathbf{b} = \mathbf{b}^T \Sigma \mathbf{b}, \quad (29)$$

where Σ_i is the Toeplitz matrix composed by the auto-covariances of $e_i(t)$, and Σ is the auto-covariance matrix defined by $\mathbb{E}[\mathbf{e}(t)^T \mathbf{e}(t - \tau)]$ as in Lemma 7.3. Hence, this Lemma leads to

$$\text{tr Var}(\mathbf{E}_w^T \mathbf{b}) \leq |\mathbf{b}|^2 \lambda_{\max}(\Sigma) \leq |\mathbf{b}|^2 \sigma^2 \eta_F^2.$$

■

The following fact is taken from [38, Prop. 8.4.13]:

Lemma 7.5. For any two p.s.d. matrices $\mathbf{A}, \mathbf{B} \in \mathbb{S}_{+}^n$, it holds that $\text{tr}(\mathbf{A}\mathbf{B}) \leq \text{tr}(\mathbf{A})\lambda_{\max}(\mathbf{B})$.

Proof of (13b): First, notice that the stochastic norm of the estimation error satisfies $\|\check{\mathbf{f}} - \mathbf{f}\|_S^2 = \mathbb{E}(\check{\mathbf{f}} - \mathbf{f})^2 + \text{tr}(\text{Var}(\check{\mathbf{f}} - \mathbf{f}))$. Hence, we obtain

$$\text{tr}(\text{Var}(\check{\mathbf{f}} - \mathbf{f})) = \left\| \mathbf{E}^\dagger (\mathbf{R}_w - \mathbf{E}_w \mathbf{f}) + ((\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}_w^T \mathbf{P}_\perp - \mathbf{E}^\dagger \mathbf{E}_w \mathbf{E}^\dagger) \mathbf{R}_{\text{NL}} \right\|_S^2.$$

Since $\mathbf{E}^\dagger \mathbf{P}_\perp = \mathbf{0}$, we have, by [30, Theorem 2.4],

$$\begin{aligned} \text{tr}(\text{Var}(\check{\mathbf{f}} - \mathbf{f})) &= \left\| \mathbf{E}^\dagger (\mathbf{R}_w - \mathbf{E}_w \mathbf{f}) - \mathbf{E}^\dagger \mathbf{E}_w \mathbf{E}^\dagger \mathbf{R}_{\text{NL}} \right\|_S^2 + \left\| (\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}_w^T \mathbf{P}_\perp \mathbf{R}_{\text{NL}} \right\|_S^2 \\ &\leq 2 \left\| \mathbf{E}^\dagger (\mathbf{R}_w - \mathbf{E}_w \mathbf{f}) \right\|_S^2 + 2 \left\| \mathbf{E}^\dagger \mathbf{E}_w \mathbf{E}^\dagger \mathbf{R}_{\text{NL}} \right\|_S^2 + \left\| (\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}_w^T \mathbf{P}_\perp \mathbf{R}_{\text{NL}} \right\|_S^2, \end{aligned} \quad (30)$$

where we have used linearity of expectation and the fact that $\left(\sum_{i=1}^N f_i \right)^2 \leq N \sum_{i=1}^N f_i^2$. Now we bound each term, recalling the following facts:

F1. For any matrix \mathbf{A} , the eigenvalues of $\mathbf{A}^\top \mathbf{A}$ are the squared singular values of \mathbf{A} , thus it holds for any $p \in \mathbb{R}$ that $\text{tr}((\mathbf{A}^\top \mathbf{A})^p) = \sum_i s_i^{2p}(\mathbf{A})$.

F2. For a random variable \mathbf{x} with mean \mathbf{m} and covariance matrix Σ , it holds that [39] $\mathbb{E}[\mathbf{x}^\top \mathbf{A} \mathbf{x}] = \text{tr}(\mathbf{A} \Sigma) + \mathbf{m}^\top \mathbf{A} \mathbf{m}$. Hence, for a zero-mean \mathbf{x} , $\mathbb{E}[|\mathbf{A} \mathbf{x}|_2^2] = \text{tr}(\mathbf{A}^\top \mathbf{A} \Sigma)$ and $\text{Var}(\mathbf{f}^\top \mathbf{x}) = \mathbf{f}^\top \Sigma \mathbf{f}$.

Bounding the first term in (30), we obtain

$$\begin{aligned} \|\mathbf{E}^\dagger(\mathbf{R}_w - \mathbf{E}_w \mathbf{f})\|_S^2 &= \mathbb{E} \left[|\mathbf{E}^\dagger(\mathbf{R}_w - \mathbf{E}_w \mathbf{f})|^2 \right] \stackrel{\text{F2}}{=} \text{tr}(\mathbf{E}^{\dagger\top} \mathbf{E}^\dagger \text{Var}(\mathbf{R}_w - \mathbf{E}_w \mathbf{f})) \\ &\stackrel{\text{Lemma 7.5}}{\leq} \text{tr}(\mathbf{E}^{\dagger\top} \mathbf{E}^\dagger) \lambda_{\max}(\text{Var}(\mathbf{R}_w - \mathbf{E}_w \mathbf{f})) = \text{tr}(\mathbf{E}(\mathbf{E}^\top \mathbf{E})^{-1} (\mathbf{E}^\top \mathbf{E})^{-1} \mathbf{E}^\top) \lambda_{\max}(\text{Var}(\mathbf{R}_w - \mathbf{E}_w \mathbf{f})) \\ &= \text{tr}(\mathbf{E}^\top \mathbf{E} (\mathbf{E}^\top \mathbf{E})^{-1} (\mathbf{E}^\top \mathbf{E})^{-1}) \lambda_{\max}(\text{Var}(\mathbf{R}_w - \mathbf{E}_w \mathbf{f})) = \text{tr}((\mathbf{E}^\top \mathbf{E})^{-1}) \lambda_{\max}(\text{Var}(\mathbf{R}_w - \mathbf{E}_w \mathbf{f})) \\ &\stackrel{\text{F1}}{=} \sum_{i=1}^m s_i^{-2} \lambda_{\max}(\text{Var}(\mathbf{R}_w - \mathbf{E}_w \mathbf{f})) \stackrel{\text{Eq. (26)}}{\leq} (|\mathbf{f}|^2 + 1) \sigma^2 \eta^2 \sum_{i=1}^m s_i^{-2}. \end{aligned} \quad (31)$$

For the second and third terms, first note that $\tilde{\mathbf{f}} := \mathbf{E}^\dagger \mathbf{R}_{\text{NL}}$ is precisely the error in estimation for the noiseless case, whose Euclidean-norm upper bound has been provided in (13a). Expanding the second term,

$$\|\mathbf{E}^\dagger \mathbf{E}_w \mathbf{E}^\dagger \mathbf{R}_{\text{NL}}\|_S^2 = \|\mathbf{E}^\dagger \mathbf{E}_w \tilde{\mathbf{f}}\|_S^2 \stackrel{\text{F1}}{=} \mathbb{E} \left[|\mathbf{E}^\dagger \mathbf{E}_w \tilde{\mathbf{f}}|^2 \right] \leq \lambda_{\max}(\text{Var}(\mathbf{E}_w \tilde{\mathbf{f}})) \sum_{i=1}^m s_i^{-2},$$

where the last inequality comes from following the same steps as in (31). Following the steps of the proof of (26) in Lemma 7.4, we get

$$\|\mathbf{E}^\dagger \mathbf{E}_w \mathbf{E}^\dagger \mathbf{R}_{\text{NL}}\|_S^2 \leq |\tilde{\mathbf{f}}|^2 \sigma^2 \eta_F^2 \sum_{i=1}^m s_i^{-2} \stackrel{(13a)}{=} B^2 \sigma^2 \eta_F^2 \sum_{i=1}^m s_i^{-2}. \quad (32)$$

For the third term, let $\mathbf{b} := \mathbf{P}_\perp \mathbf{R}_{\text{NL}}$. We again follow similar steps to (31) to get

$$\begin{aligned} \|(\mathbf{E}^\top \mathbf{E})^{-1} \mathbf{E}_w^\top \mathbf{b}\|_S^2 &\leq \lambda_{\max}((\mathbf{E}^\top \mathbf{E})^{-2}) \text{tr}(\text{Var}(\mathbf{E}_w^\top \mathbf{b})) = \text{tr}(\text{Var}(\mathbf{E}_w^\top \mathbf{b})) s_{\min}(\mathbf{E})^{-4} \\ &\stackrel{(27)}{\leq} |\mathbf{b}|^2 \sigma^2 \eta_F^2 s_{\min}(\mathbf{E})^{-4} \leq |\mathbf{b}|^2 \sigma^2 \eta_F^2 s_{\min}(\mathbf{E})^{-4}. \end{aligned} \quad (33)$$

Since \mathbf{P}_\perp is a projection matrix, $|\mathbf{b}| \leq |\mathbf{R}_{\text{NL}}|$. By the proof of (13a), $|\mathbf{R}_{\text{NL}}| \leq Am|\mathbf{f}|^2 \sqrt{N} \|\begin{bmatrix} \xi^\top & \mathbf{z}^\top \end{bmatrix}\|_\infty = Bs_{\min}(\mathbf{E})$. Therefore,

$$\|(\mathbf{E}^\top \mathbf{E})^{-1} \mathbf{E}_w^\top \mathbf{b}\|_S^2 \leq B^2 \sigma^2 \eta_F^2 s_{\min}(\mathbf{E})^{-2}. \quad (34)$$

Now we can finally replace the bounds (31)–(34) into (30) to get

$$\text{tr}(\text{Var}(\check{\mathbf{f}} - \mathbf{f})) \leq \sigma^2 \left(2(|\mathbf{f}|^2 + 1) \eta^2 \sum_{i=1}^m s_i^{-2} + B^2 \eta_F^2 \left(2 \sum_{i=1}^m s_i^{-2} + s_m^{-2} \right) \right).$$

This concludes the main result. For the special case where $\mathbf{H}'_i \equiv \mathbf{0}$ for all i , we have that $\mathbf{R}_{\text{NL}} = \mathbf{0}$, and then the bounding step in (30) is unnecessary, yielding $\|\check{\mathbf{f}} - \mathbf{f}\|_S^2 = \|\mathbf{E}^\dagger(\mathbf{R}_w - \mathbf{E}_w \mathbf{f})\|_S^2$. The desired expression is then obtained by merely using (31). \blacksquare

7.3 Derivation of the SNR equation (14)

In [30], three metrics are suggested to evaluate the quality of the errors calculated by using the first-order approximation. The first two are related to the approximation of the pseudo-inverse, and the last is specific to

least squares:

$$c_1 := \|\mathbf{E}^\dagger \mathbf{E}_w\|_S, \quad (35)$$

$$c_2 := \sup_{\mathbf{v} \neq \mathbf{0}} \mathbb{E} \left[\frac{\mathbf{v}^\top \mathbf{E}_w^\top \mathbf{P}_\perp \mathbf{E}_w \mathbf{v}}{\mathbf{v}^\top \mathbf{E}^\top \mathbf{E} \mathbf{v}} \right], \quad (36)$$

$$c_3 := \frac{\|\check{\mathbf{E}}^\dagger - \mathbf{E}^\dagger\|_S}{\|\mathbf{E}^\dagger\|_F}. \quad (37)$$

Due to the asymptotic arguments that justify the theory in [30], all these metrics must be significantly smaller than 1 for Theorem 3.2 to have practical meaning; i.e., this ensures that the approximation $\|\hat{\mathbf{f}} - \mathbf{f}\|_S \approx \|\hat{\mathbf{f}} - \mathbf{f}\|_S$ holds. Below we provide some expressions for upper bounds of these metrics, using the same facts and similar algebraic techniques as in the proof of Theorem 3.2. For brevity, we will keep explanations of steps to a minimum.

$$\begin{aligned} c_1^2 &= \mathbb{E}[\text{tr}(\mathbf{E}_w^\top \mathbf{E}^{\dagger\top} \mathbf{E}^\dagger \mathbf{E}_w)] = \mathbb{E}[\text{tr}(\mathbf{E}^{\dagger\top} \mathbf{E}^\dagger \mathbf{E}_w \mathbf{E}_w^\top)] \stackrel{\text{Lemma 7.5}}{\leq} \mathbb{E}[\lambda_{\max}(\mathbf{E}^{\dagger\top} \mathbf{E}^\dagger) \text{tr}(\mathbf{E}_w \mathbf{E}_w^\top)] \\ &= \lambda_{\max}((\mathbf{E}^\top \mathbf{E})^{-1}) \mathbb{E}[\text{tr}(\mathbf{E}_w^\top \mathbf{E}_w)] = s_m^{-2} \mathbb{E}[\text{tr}(\mathbf{E}_w^\top \mathbf{E}_w)] \stackrel{\text{Lemma 7.4}}{\leq} s_m^{-2} \sigma^2 N \gamma_F^2. \end{aligned} \quad (38)$$

For c_2 , we extract from the proof of Lemma 7.4 that, for any constant vector $\mathbf{b} \in \mathbb{R}^m$,

$$\mathbb{E}[\mathbf{b}^\top \mathbf{E}_w^\top \mathbf{E}_w \mathbf{b}] \leq \sigma^2 \eta_F^2 |\mathbf{b}|^2. \quad (39)$$

Now, define $\mathbf{c} := \mathbf{E} \mathbf{v}$ to get

$$\begin{aligned} c_2 &= \sup_{\|\mathbf{c}\|=1} \mathbb{E}[\mathbf{c}^\top \mathbf{E}^{\dagger\top} \mathbf{E}_w^\top \mathbf{P}_\perp \mathbf{E}_w \mathbf{E}^\dagger \mathbf{c}] \stackrel{P_\perp \leq I}{\leq} \sup_{\|\mathbf{c}\|=1} \mathbb{E}[\mathbf{c}^\top \mathbf{E}^{\dagger\top} \mathbf{E}_w^\top \mathbf{E}_w \mathbf{E}^\dagger \mathbf{c}] \\ &\stackrel{(39)}{\leq} \sup_{\mathbf{b} = \mathbf{E}^\dagger \mathbf{c}, \|\mathbf{c}\|=1} \sigma^2 \eta_F^2 |\mathbf{b}|^2 = \sigma^2 \eta_F^2 s_{\max}(\mathbf{E}^\dagger)^2 = \sigma^2 \eta_F^2 s_{\min}(\mathbf{E})^{-2}. \end{aligned} \quad (40)$$

For c_3 , we have from [30] that $\mathbf{E}^\dagger - \check{\mathbf{E}}^\dagger = (\mathbf{E}^\top \mathbf{E})^{-1} \mathbf{E}_w^\top \mathbf{P}_\perp - \mathbf{E}^\dagger \mathbf{E}_w \mathbf{E}^\dagger$. We follow operations similar to those in the proof of Lemma 7.4 to get

$$\begin{aligned} \|\mathbf{E}^\dagger - \check{\mathbf{E}}^\dagger\|_S^2 &= \|(\mathbf{E}^\top \mathbf{E})^{-1} \mathbf{E}_w^\top \mathbf{P}_\perp\|_S^2 + \|\mathbf{E}^\dagger \mathbf{E}_w \mathbf{E}^\dagger\|_S^2 \\ &= \mathbb{E}[\text{tr}(\mathbf{P}_\perp \mathbf{P}_\perp^\top \mathbf{E}_w (\mathbf{E}^\top \mathbf{E})^{-2} \mathbf{E}_w^\top)] + \mathbb{E}[\text{tr}(\mathbf{E}^\dagger \mathbf{E}^{\dagger\top} \mathbf{E}_w^\top \mathbf{E}_w \mathbf{E}^\dagger)] \\ &\leq \mathbb{E}[\text{tr}(\mathbf{E}_w (\mathbf{E}^\top \mathbf{E})^{-2} \mathbf{E}_w^\top)] + s_m^2 \mathbb{E}[\text{tr}(\mathbf{E}^{\dagger\top} \mathbf{E}^\dagger \mathbf{E}_w \mathbf{E}_w^\top)] \\ &\leq 2s_m^{-4} \mathbb{E}[\text{tr}(\mathbf{E}_w^\top \mathbf{E}_w)] \\ &\leq 2s_m^{-4} \sigma^2 N \gamma_F^2, \end{aligned}$$

and, thus,

$$c_3 \leq \frac{\sqrt{2N} s_m^{-2} \sigma \gamma_F}{\sqrt{\sum_{i=1}^m s_m^{-2}}} \leq \frac{\sqrt{2N} s_m^{-2} \sigma \gamma_F}{\sqrt{s_m^{-2}}} = \sqrt{2N} s_m^{-1} \sigma \gamma_F.$$

One may note that bounds satisfy $\sqrt{2} c_1 = c_3$, which are proportional to $\sqrt{N} s_m^{-1}$. The inverse of this quantity, s_m / \sqrt{N} , has been discussed in Section 3 as an “effective” singular value of \mathbf{E} , when considering periodic experiments. Since c_3 is the metric specifically built for least-squares problems, we chose $c := c_3$ in Remark 3.4.

7.4 Proof of Proposition 4.1

Proof: For this proof it suffices to show that for any $i \in \mathbb{N}_{<N}$,

$$\lim_{k \rightarrow \infty} s_{\min}(\mathbf{E}_{[kN, kN+N-1]}) = \lim_{k \rightarrow \infty} s_{\min}(\mathbf{E}_{[kN+i, kN+i+N-1]}).$$

Let \mathbf{e}' be the limit periodic signal of \mathbf{e} , i.e., \mathbf{e}' is N -periodic and, for any $i \in \mathbb{N}_{<N}$, it holds that $\lim_{k \rightarrow \infty} \mathbf{e}(kN+i) = \mathbf{e}'(i)$. We start by noting that

$$\mathbf{E}'_{[i, N+i-1]} = \begin{bmatrix} \mathbf{E}'_{[i, N-1]} \\ \mathbf{E}'_{[0, i-1]} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_i \\ \mathbf{I}_{N-i} & \mathbf{0} \end{bmatrix} \mathbf{E}'_{[0, N-1]},$$

where the subscripts denote the dimension of the corresponding identity matrices. Thus,

$$\mathbf{E}'_{[0, N-1]}^T \mathbf{E}'_{[0, N-1]} = \mathbf{E}'_{[i, N+i-1]}^T \mathbf{E}'_{[i, N+i-1]},$$

which implies that for all i

$$\lim_{k \rightarrow \infty} s_{\min}(\mathbf{E}_{[kN, kN+N-1]}) = s_{\min}(\mathbf{E}'_{[0, N-1]}) = s_{\min}(\mathbf{E}'_{[i, N+i-1]}) = \lim_{k \rightarrow \infty} s_{\min}(\mathbf{E}_{[kN+i, kN+i+N-1]}),$$

proving the existence of the limit. ■

7.5 Proofs of Theorem 4.1 and Proposition 4.2

We start with a Lemma for the subgradient information in Theorem 4.1:

Lemma 7.6 (First-order oracle expression). *Consider the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by*

$$f(\mathbf{x}) := \lambda_{\min} \left(\sum_{i=1}^N \mathbf{P}_i \mathbf{x} \mathbf{x}^T \mathbf{P}_i^T \right).$$

Let \mathbf{v}_{\min} be a unitary eigenvector associated to one of the minimal eigenvalues of $\sum_{i=1}^N \mathbf{P}_i \mathbf{x} \mathbf{x}^T \mathbf{P}_i^T$. Then, the vector

$$\mathbf{g} := 2 \sum_{i=1}^N \mathbf{v}_{\min}^T \mathbf{P}_i \mathbf{x} \mathbf{P}_i^T \mathbf{v}_{\min}$$

is a subgradient of f .

Proof: The first step is to reformulate the minimum eigenvalue function as the solution of a minimization problem. First, notice that the matrix $\mathbf{S}(\mathbf{x}) := \sum_{i=1}^N \mathbf{P}_i \mathbf{x} \mathbf{x}^T \mathbf{P}_i^T$ is symmetric. Therefore, by Rayleigh's Theorem [40, Theorem 4.2.2],

$$f(\mathbf{x}) = \min_{\|\mathbf{v}\|_2=1} \phi(\mathbf{v}, \mathbf{x}) := \min_{\|\mathbf{v}\|_2=1} \mathbf{v}^T \left(\sum_{i=1}^N \mathbf{P}_i \mathbf{x} \mathbf{x}^T \mathbf{P}_i^T \right) \mathbf{v},$$

for which \mathbf{v}_{\min} is in the set of minimizers. Furthermore, the function $\phi(\mathbf{v}, \mathbf{x})$ is convex in both \mathbf{v} and \mathbf{x} . Thus, applying Danskin's Theorem [41], the partial derivative $\partial \phi(\mathbf{x}, \mathbf{v}) / \partial \mathbf{x}$, taken at $\mathbf{v} = \mathbf{v}_{\min}$, is a subgradient of $f(\mathbf{x})$. Applying the dot product rule gives

$$\frac{\partial \phi(\mathbf{v}, \mathbf{x})}{\partial \mathbf{x}} = 2 \sum_{i=1}^N \mathbf{v}^T \mathbf{P}_i \mathbf{x} \mathbf{P}_i^T \mathbf{v}.$$

Replacing \mathbf{v} for \mathbf{v}_{\min} concludes the proof. ■

Proof of Theorem 4.1: Since $\mathcal{T}(z)$ is BIBO stable, by Proposition 4.1 we have that $J_{\infty} = s_{\min}(\mathbf{E}'_{[0, N-1]})^2$, where $\mathbf{E}'_{[0, N-1]}$ is the matrix built by the first N samples of the limit periodic signal \mathbf{e}' .

We now compute \mathbf{e}' as a function of \mathbf{u} . To do this, we must find the initial state \mathbf{x}'_0 such that, for all $k \in \mathbb{N}_0$, $\mathbf{e}'(i + kN) = \mathbf{P}'_i \bar{\mathbf{u}}$. First, recall (18) and notice that, for any \mathbf{x}_0 ,

$$\mathbf{x}(N) = \mathbf{P}_x \bar{\mathbf{u}} + \mathbf{A}^N \mathbf{x}_0.$$

Now, \mathbf{x} is N -periodic iff $\mathbf{x}(N) = \mathbf{x}_0$, hence

$$\mathbf{x}_0 = \mathbf{x}'_0 := (\mathbf{I} - \mathbf{A}^N) \mathbf{P}_x \bar{\mathbf{u}} \quad (41)$$

renders \mathbf{e} N -periodic and equal, by definition, to \mathbf{e}' . From LTI systems theory and by definition of \mathbf{P}_i and $\bar{\mathbf{u}}$, we have that

$$\mathbf{e}'(i) = \mathbf{P}_i \bar{\mathbf{u}} + \mathbf{C} \mathbf{A}^{i-1} \mathbf{x}_0.$$

Substituting \mathbf{x}_0 for \mathbf{x}'_0 in (41), and by definition of \mathbf{P}'_i in (18), we get

$$\mathbf{e}'(i) = \mathbf{P}'_i \bar{\mathbf{u}},$$

which, due to periodicity of \mathbf{x} , implies $\mathbf{e}'(i + kN) = \mathbf{e}'(i)$, $\forall k \in \mathbb{N}_0$.

Notice now that $\mathbf{E}'_{[0, N-1]} = [\mathbf{P}'_1 \bar{\mathbf{u}} \quad \mathbf{P}'_2 \bar{\mathbf{u}} \quad \cdots \quad \mathbf{P}'_N \bar{\mathbf{u}}]^\top$. For simplicity, denote $\mathbf{E}' := \mathbf{E}'_{[0, N-1]}$. Now, recall that $J_\infty(\mathbf{u}) = s_{\min}(\mathbf{E}')^2 = \lambda_{\min}(\mathbf{E}'^\top \mathbf{E}')$. It is easy to see that

$$\mathbf{E}'^\top \mathbf{E}' = \sum_{i=1}^N \mathbf{P}'_i \bar{\mathbf{u}} \bar{\mathbf{u}}^\top \mathbf{P}'_i{}^\top,$$

which establishes that $J_\infty(\mathbf{u}) = J(\bar{\mathbf{u}})$ as desired.

The subgradient result (19a) is an immediate consequence of Lemma 7.6. ■

Proof of Proposition 4.2: First, notice that $\max \left\{ \lambda \mid \sum_{i=1}^N \mathbf{P}'_i \bar{\mathbf{U}} \mathbf{P}'_i{}^\top \succeq \lambda \mathbf{I} \right\} = \lambda_{\min} \left(\sum_{i=1}^N \mathbf{P}'_i \bar{\mathbf{U}} \mathbf{P}'_i{}^\top \right)$, by the properties of eigenvalues of symmetric matrices. Thus, we may rewrite (20) as the equivalent problem

$$\begin{aligned} \max_{\bar{\mathbf{U}}} \quad & \lambda_{\min} \left(\sum_{i=1}^N \mathbf{P}'_i \bar{\mathbf{U}} \mathbf{P}'_i{}^\top \right) \\ \text{subject to} \quad & \text{tr}(\mathbf{S}_i \bar{\mathbf{U}}) \leq 1, \quad \forall i = 1, \dots, N_c, \\ & \bar{\mathbf{U}} \succeq \mathbf{0} \end{aligned} \quad (42)$$

To prove that $\lambda^* \leq \lambda^{\text{SDP}}$, it suffices to show that for any feasible $\bar{\mathbf{u}} \in \tilde{\mathcal{U}}$, $\bar{\mathbf{u}} \bar{\mathbf{u}}^\top$ is a feasible solution of (42). This is easy to see since $\bar{\mathbf{u}}^\top \mathbf{S}_i \bar{\mathbf{u}} = \text{tr}(\mathbf{S}_i \bar{\mathbf{u}} \bar{\mathbf{u}}^\top) \leq 1$, and $\bar{\mathbf{u}} \bar{\mathbf{u}}^\top \succeq \mathbf{0}$.

To prove equality $\lambda^* = \lambda^{\text{SDP}}$ when $\bar{\mathbf{U}}^*$ has rank 1, we can check that in this case $\bar{\mathbf{U}}^* = s_1 \mathbf{v} \mathbf{v}^\top$, where $s_1 > 0$ is the only non-zero singular value of $\bar{\mathbf{U}}^*$. Then let $\bar{\mathbf{u}}^* := \sqrt{s_1} \mathbf{v}$, and it holds that $\bar{\mathbf{U}}^* = \bar{\mathbf{u}}^* \bar{\mathbf{u}}^{*\top}$. Now, $\bar{\mathbf{u}}^* \in \tilde{\mathcal{U}}$, because for all i , $\text{tr}(\mathbf{S}_i \bar{\mathbf{U}}^*) = \text{tr}(\mathbf{S}_i \bar{\mathbf{u}}^* \bar{\mathbf{u}}^{*\top}) = \bar{\mathbf{u}}^{*\top} \mathbf{S}_i \bar{\mathbf{u}}^* \leq 1$. Hence, $\lambda_{\min} \left(\sum_{i=1}^N \mathbf{P}'_i \bar{\mathbf{u}}^* \bar{\mathbf{u}}^{*\top} \mathbf{P}'_i{}^\top \right) \leq \lambda^* \leq \lambda^{\text{SDP}}$. But $\lambda_{\min} \left(\sum_{i=1}^N \mathbf{P}'_i \bar{\mathbf{u}}^* \bar{\mathbf{u}}^{*\top} \mathbf{P}'_i{}^\top \right) = \lambda_{\min} \left(\sum_{i=1}^N \mathbf{P}'_i \bar{\mathbf{U}}^* \mathbf{P}'_i{}^\top \right) = \lambda^{\text{SDP}}$, hence equality is proven. ■

References

- [1] R. V. Beard, *Failure accomodation in linear systems through self-reorganization*. PhD thesis, Massachusetts Inst. Technol., 1971.
- [2] H. L. Jones, *Failure detection in linear systems*. PhD thesis, Massachusetts Inst. Technol., 1973.
- [3] S. X. Ding, *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms and Tools*. Advances in Industrial Control, London: Springer, 2013.
- [4] X. J. Zhang, *Auxiliary signal design in fault detection and diagnosis*. Springer, 1989.

- [5] S. L. Campbell and R. Nikoukhah, *Auxiliary signal design for failure detection*. Princeton Univ. Press, 2015.
- [6] J. K. Scott, G. R. Marseglia, L. Magni, R. D. Braatz, and D. M. Raimondo, "A hybrid stochastic-deterministic input design method for active fault diagnosis," in *52nd IEEE Conf. Decis. Control*, pp. 5656–5661, IEEE, 2013.
- [7] M. imandl and I. Punochá, "Active fault detection and control: Unified formulation and optimal design," *Automatica*, vol. 45, pp. 2052–2059, Sept. 2009.
- [8] J. A. Paulson, T. A. N. Heirung, R. D. Braatz, and A. Mesbah, "Closed-loop active fault diagnosis for stochastic linear systems," in *2018 Annual American Control Conference*, pp. 735–741, IEEE, 2018.
- [9] J. Noom, N. H. Thao, O. Soloviev, and M. Verhaegen, "Closed-loop active model diagnosis using bhat-tacharyya coefficient: Application to automated visual inspection," in *Int. Conf. Intell. Syst. Des. Appl.*, pp. 657–667, Springer, 2020.
- [10] J. Noom, O. Soloviev, C. Smith, and M. Verhaegen, "Online input design for discrimination of linear models using concave minimization," *arXiv preprint arXiv:2401.05782*, 2024.
- [11] Z. Gao, C. Cecati, and S. X. Ding, "A Survey of Fault Diagnosis and Fault-Tolerant TechniquesPart II," *IEEE Trans. Ind. Electron.*, vol. 62, pp. 3768–3774, June 2015.
- [12] T. A. N. Heirung and A. Mesbah, "Input design for active fault diagnosis," *Annual Reviews in Control*, vol. 47, pp. 35–50, 2019.
- [13] M. Aoki and R. Staley, "On input signal synthesis in parameter identification," *Automatica*, vol. 6, no. 3, pp. 431–440, 1970.
- [14] G. Goodwin, "Optimal input signals for nonlinear-system identification," in *Proc. IEE*, vol. 118, pp. 922–926, 1971.
- [15] R. Mehra, "Optimal inputs for linear system identification," *IEEE Trans. Autom. Control*, vol. 19, no. 3, pp. 192–200, 1974.
- [16] H. Hjalmarsson, "From experiment design to closed-loop control," *Automatica*, vol. 41, pp. 393–438, Mar. 2005.
- [17] X. Bombois, M. Gevers, R. Hildebrand, and G. Solari, "Optimal experiment design for open and closed-loop system identification," *Commun. Inf. Syst.*, vol. 11, no. 3, pp. 197–224, 2011.
- [18] N. Dirksen, J. van de Wijdeven, and T. Oomen, "Frequency response function identification for multivariable motion control: Optimal experiment design with element-wise constraints," *Mechatronics*, vol. 71, p. 102440, 2020.
- [19] N. Dirksen, M. Bosselaar, and T. Oomen, "Peak amplitude-constrained experiment design for frf identification of mimo motion systems," in *IEEE 17th Int. Conf. Adv. Motion Control*, pp. 256–261, IEEE, 2022.
- [20] I. M. Mithun, S. Mohan, and B. Bhikkaji, "Optimal finite-dimensional spectral densities for the identification of continuous-time MIMO systems," *Control Theory and Technology*, vol. 17, pp. 276–296, Aug. 2019.
- [21] Y. Ye and S. Zhang, "New results on quadratic mini-mization," *SIAM J. Optim.*, vol. 14, no. 1, pp. 245–267, 2003.
- [22] A. Wagenmaker and K. Jamieson, "Active learning for identification of linear dynamical systems," in *Conference on Learning Theory*, pp. 3487–3582, PMLR, 2020.
- [23] H. J. van Waarde, "Beyond Persistent Excitation: Online Experiment Design for Data-Driven Modeling and Control," *IEEE Control Systems Letters*, vol. 6, pp. 319–324, 2022. Conference Name: IEEE Control Systems Letters.
- [24] C. De Persis and P. Tesi, "Designing experiments for data-driven control of nonlinear systems," *IFAC-PapersOnLine*, vol. 54, no. 9, pp. 285–290, 2021.

- [25] M. Nyberg and E. Frisk, "Residual generation for fault diagnosis of systems described by linear differential-algebraic equations," *IEEE Trans. Autom. Control*, vol. 51, no. 12, pp. 1995–2000, 2006.
- [26] P. Mohajerin Esfahani, M. Vrakopoulou, G. Andersson, and J. Lygeros, "A tractable nonlinear fault detection and isolation technique with application to the cyber-physical security of power systems," in *51st IEEE Conf. Decis. Control*, (Maui, HI, USA), pp. 3433–3438, IEEE, Dec. 2012.
- [27] P. Mohajerin Esfahani and J. Lygeros, "A Tractable Fault Detection and Isolation Approach for Nonlinear Systems With Probabilistic Performance," *IEEE Trans. Autom. Control*, vol. 61, pp. 633–647, Mar. 2016.
- [28] C. Van Der Ploeg, M. Alirezaei, N. Van De Wouw, and P. Mohajerin Esfahani, "Multiple Faults Estimation in Dynamical Systems: Tractable Design and Performance Bounds," *IEEE Trans. Autom. Control*, vol. 67, pp. 4916–4923, Sept. 2022.
- [29] P. Antsaklis, "Some relations satisfied by prime polynomial matrices and their role in linear multivariable system theory," *IEEE Trans. Autom. Control*, vol. 24, no. 4, pp. 611–616, 1979.
- [30] G. W. Stewart, "Stochastic perturbation theory," *SIAM Review*, vol. 32, p. 579610, Dec. 1990.
- [31] P. Rapisarda, M. K. Çamlıbel, and H. Van Waarde, "A persistency of excitation condition for continuous-time systems," *IEEE Control Syst. Lett.*, vol. 7, pp. 589–594, 2022.
- [32] L. Grippo and M. Sciandrone, *Introduction to Methods for Nonlinear Optimization*, vol. 152. Springer Nature, 2023.
- [33] D. Tilbury, B. Messner, L. Oms, J. Pagel, Y. Sun, M. Suri, C. Caruana, D. Kawano, B. Nakai, P. Prempraneerach, J. Luntz, R. Hill, J. Taylor, and S. Das, "Inverted pendulum: System modeling," 1998. <https://ctms.engin.umich.edu/CTMS/index.php?example=InvertedPendulum§ion=SystemModeling> (Accessed on June 17th, 2023).
- [34] CVX Research, Inc., "CVX: Matlab software for disciplined convex programming, version 2.0," Aug. 2012.
- [35] S. de Reij, "Multivariate fault estimation for a high-fidelity model of the AB383 wire bonder: A model-based approach," masters thesis, Delft Univ. Technol., 2023.
- [36] A. V. Oppenheim and G. C. Verghese, *Signals, systems & inference*. Pearson London, 2017.
- [37] R. M. Gray *et al.*, "Toeplitz and circulant matrices: A review," *Found. Trends Commun. Inf. Theory*, vol. 2, no. 3, pp. 155–239, 2006.
- [38] D. S. Bernstein, *Matrix Mathematics: Theory, Facts and Formulas*. Princeton, 2nd ed., 2009.
- [39] M. Brookes, "The matrix reference manual," 2020. <http://www.ee.imperial.ac.uk/hp/staff/dmb/matrix/intro.html>.
- [40] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 2012.
- [41] J. M. Danskin, *The theory of max-min and its application to weapons allocation problems*, vol. 5. Springer Science & Business Media, 1967.