Frequency Response Function Identification for Multivariable Motion Control: Optimal Experiment Design with Element-Wise Constraints

Nic Dirkx\textsuperscript{a,b,}, Jeroen van de Wijdeven\textsuperscript{a}, Tom Oomen\textsuperscript{b}
\textsuperscript{a} ASML Research Mechatronics & Control, Veldhoven, The Netherlands
\textsuperscript{b} Eindhoven University of Technology, Department of Mechanical Engineering, Eindhoven, The Netherlands

Abstract
Frequency Response Functions (FRFs) are essential in mechatronic systems and its application ranges from system design and validation to controller design and diagnostics. The aim of this paper is to optimally design experiments for FRF identification of multivariable motion systems subject to element-wise power constraints. A multivariable excitation design framework is established that explicitly addresses the frequency-wise directionality of the system to be identified. The design problem involves solving a rank-constrained optimization problem, which is non-convex and NP-hard in most cases. Two algorithms to solving this problem approximately are presented that rely on a convex (semi-definite) relaxation of the original problem. Additionally, exact solutions for several special cases are presented. The two algorithms are shown to overcome the limitations of traditional excitation design. This is confirmed by experimental results from a 7 \times 8 wafer stage setup, which show a significant improvement of the FRF quality using the proposed techniques over traditional design approaches.

Key words: Optimal experiment design, system identification, multivariable systems, frequency response function, multisines, rank-constrained optimization

1. Introduction
High-precision positioning systems face ever increasing performance demands in terms of accuracy and productivity (Munnig Schmidt et al., 2014). These mechatronic systems, including wafer scanners, printing systems, and medical scanners, typically exhibit complex multivariable dynamics, e.g., due to flexible mechanics. Meeting the performance demands requires the employment of advanced model-based control strategies (Skogestad and Postlethwaite, 2005). The majority of these control techniques rely on the availability of accurate FRF models (Steinbuch and Norg, 1998; van de Wal et al, 2002; Oomen, 2018). The identification of FRF models from experimental data, as opposed to first principles modeling, is considered fast, inexpensive, and accurate (Pintelon and Schoukens, 2012; Ljung, 1999). The resulting FRF constitutes an intermediate step towards parametric modeling and is also used directly for control design (Karimi and Kammer, 2016) and system diagnostics (Maas et al., 2016).

The quality of the identified FRF depends on the experiment (Hjalmarsson, 2005) and therefore these experiments must be designed carefully. This applies especially to next-generation mechatronic systems, equipped with a large number of actuators (Van Herpen et al., 2014; Oomen, 2018), since the number of experiments increases with the number of inputs.

Optimal Experiment Design (OED) consists in a systematic approach for the design of input signals that maximize the FRF accuracy within limited resources. The topic of OED has received much research interest in the last decades, see the overviews Mehra (1974b); Gevers et al. (2011); Hjalmarsson (2005).

Dealing with system constraints is a necessary aspect in experiment design to achieve practical requirements such as plant-friendly operation (Deshpande and Rivera, 2014). For many mechatronic systems such as motion stages, the constraints are related to physical limitations of specific elements of the system, e.g., a bounded power for a specific actuator (Dirkx et al., 2019a).

OED for Single Input Single Output (SISO) systems is well-developed, which has led to a large number of theoretical results and design frameworks. Design methods for parametric identification of SISO systems based on numerical optimization schemes are developed in, e.g., Ljung (1999), Ch. 13; Goodwin and Payne (1977), Ch. 6; Hjalmarsson (2005); Gevers et al. (2011). Typically, the aim is to formulate the design problem as the minimization of a convex objective function over a convex constraint set (e.g., Bombois et al., 2006; Gerencsér et al., 2016; Lindqvist and Hjalmarsson, 2000; Jansson, 2004). Because the optimal

Preprint submitted to Elsevier June 30, 2020

\textsuperscript{+}This work is part of the research programme VIDI with project number 15698, which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO).

\textsuperscript{*}Corresponding author

Postal address: Technische Universiteit Eindhoven P.O. Box 513 5600 MB Eindhoven, The Netherlands

Email address: nic.dirkx@asml.com (Nic Dirkx)
For Multiple Inputs Multiple Outputs (MIMO) systems, the design problem becomes significantly more complex. For such systems, the problem involves not only the design of excitation magnitudes, but also the excitation directions (Dirkx et al., 2019a). In Pintelon and Schoukens (2012, Ch 2.5), Guillaume et al. (1996); Dobrowiecki et al. (2006) the strategy for dealing with multiple inputs is to choose the excitation directions orthogonal to each other. These orthogonal excitations are (D)-optimal in the case of input power constraints, but not in the general case, e.g., under bounded input and output power. Numerical optimization-based approaches that allow for multivariable design beyond orthogonal excitations are developed in Barenthin et al. (2008); Mehra (1974a). Herein, input design problems under total input or output power constraints are considered, which are typically convex problems and hence can be solved efficiently. For many real-life mechatronic systems, however, constraints on the total power do not adequately represent the actual, often element-wise system limitations.

In OED for MIMO systems, addressing element-wise constraints in a non-conservative manner is crucial to maximize the model quality. In prior work Dirkx et al. (2019a), it is shown that the design problem under element-wise constraints requires solving a non-convex optimization problem, but no general computational tools are provided. In related work Dirkx and Oomen (2019b), preliminary results to solve this problem are reported.

Although many results on OED are present, fast and accurate identification of complex multivariable mechatronic systems is hampered by a lack of design approaches that can deal with element-wise constraints. The aim of this paper is to solve the OED problem for FRF identification of complex multivariable systems under element-wise power constraints.

The main contributions in this paper are:

1. A framework for optimal multivariable excitation design for FRF identification of MIMO systems under element-wise power constraints,
2. a Sequential Semi-Definite Relaxation (SSDR) algorithm to approximately solve the multivariable excitation design problem, including feasibility and convergence results,
3. an excitation design approach based on a Relaxation and Randomization (RR) algorithm, including theoretical and practical performance bounds,
4. a comparison in terms of time-efficiency and performance of both algorithms based on experimental results from a next-generation wafer stage.

In Section 2, the mechatronic system is introduced and the experiment design framework and objective is established. In Section 3, the role of multivariable excitations in identification of MIMO systems is illustrated. Section 4 presents different algorithms to solve the design problem. The experimental results are presented in Section 5, followed by the conclusions in Section 6.

1.1. Notations and definitions

$\mathbb{H}^n$ denotes the set of $n \times n$ Hermitian matrices. Subscripts $(\cdot)^+$ denote positive (semi)-definiteness of a matrix or (semi)-positivity of a vector. $X^*$ represents the complex conjugate of $X \in \mathbb{C}$. Operation $(A, X) = \sum_{i,j} A_{ij} X_{ij} = \text{Tr}(A^H X)$ denotes the Frobenius inner product of equal sized matrices $A$ and $X$. The $\otimes$ operator denotes the Kronecker product, and $X = \text{diag}(x)$ results in a diagonal matrix $X$ with the entries of vector $x$ on the diagonal. $x \in \mathcal{N}(\mathbb{E}\{x\}, \sigma^2\{x\})$ indicates that a sample $x$ belongs to a normal distribution with expected value $\mathbb{E}\{x\}$ and variance $\sigma^2\{x\}$. The Discrete Fourier Transform (DFT) of a discrete time sampled signal $x(t)$ is defined as

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(nT_s)e^{-j2\pi nk/N} \quad (1)$$

with $N$ the number of samples, $T_s$ the sample time, and $k$ the discrete frequency index. The normalization factor of $1/N$ in (1) is employed instead of the conventional factor $1/\sqrt{N}$ to simplify the equations in this paper.

2. System description and OED framework

In Section 2.1, a description of the wafer stage system is provided. The methods in this paper are directly applicable to many other applications, in particular mechatronic systems with multiple actuators and sensors. In Sections 2.2 and 2.3, the identification framework and experiment design objectives are presented, respectively.

2.1. Over-actuated wafer stage system

Complex high-precision positioning systems are considered. Envisaging next-generation control systems, the focus is particularly on systems equipped with a large number of spatially distributed actuators and sensors (Oomen, 2018). One such a system is the next-generation wafer stage system in Fig. 1, which is considered in this paper. To perform fast motion tasks and high accelerations, the stage is designed as a lightweight mechanical structure. Hence, it exhibits complex dynamic behavior, e.g., due to mechanical flexibilities. To achieve nanometer accurate positioning performance, the stage is magnetically levitated using passive gravity compensators. Additionally, the stage is over-actuated and over-sensed, i.e., the number of actuators and sensors is larger than six. This enables control of additional structural deformations, besides the conventional six rigid body motion degrees of freedom (translations and rotations) (Van Herpen et al., 2014). Specifically, the stage is actuated by 8 Lorentz actuators and sensed by 7 encoders, according to the configuration in Fig. 2.
The stage operates in closed-loop position control using a digital PID-like controller that achieves a bandwidth of 50 Hz, with a sampling frequency of 5 kHz. Due to electro-mechanical limitations, the maximum allowable signal power for each of the inputs and outputs is 0.25 W and 0.30 mW, respectively.

Identifying such large dimensional complex MIMO systems in a cost effective manner imposes specific demands on experiment design strategies. In this paper, experiment design strategies are developed that efficiently deal with the large number of inputs by exploiting the design freedom in the excitation signals, yet within the physical constraints of the closed-loop system.

2.2. Identification framework

Consider the identification scheme in Fig. 3, where $G$ represents the $n_y \times n_u$ LTI system to be identified and $K$ is a LTI feedback controller. Signal $w \in \mathbb{R}^{n_u}$ is a (periodic) multisine excitation signal (Pintelon and Schoukens, 2012, Ch. 2). Signals $w$, $u$ and $y$ are measured for system identification. The signal $y$ is perturbed with measurement noise $\nu_y$, characterized as a filtered independent and identically distributed (iid) random sequence. The DFT of the signals $y^{[e]}(t)$ and $u^{[e]}(t)$ measured during an $e$-th experiment, with $e = 1, \ldots, n_u$, are given by

$$
\begin{bmatrix}
Y^{[e]}(k) \\
U^{[e]}(k)
\end{bmatrix} =
\begin{bmatrix}
G(k) \\
I
\end{bmatrix}
S(k)W^{[e]}(k) +
\begin{bmatrix}
V_y^{[e]}(k) \\
V_u^{[e]}(k)
\end{bmatrix}
$$

where $W^{[e]}(k) \in \mathbb{C}^{n_u}$ is the DFT of $w^{[e]}(t)$ and with $[V_y^{[e]}(k)^T \ V_u^{[e]}(k)^T]^T \in \mathbb{C}^{n_y+n_u}$ the DFT of the noise contribution of $\nu_y$ onto $[y^{[e]}(t)^T \ u^{[e]}(t)^T]^T$. Under the assumed conditions on $\nu_y$, terms $V_y^{[e]}(k)$ and $V_u^{[e]}(k)$ converge to a circular complex normal distribution for $N \rightarrow \infty$ (Oomen, 2008; Pintelon and Schoukens, 2012, Ch. 14). System $S(k) = (I + K(k)G(k))^{-1}$ denotes the input Sensitivity function. The excitation matrix $W(k)$ is constructed by performing a number of $n_u$ experiments as

$$
W(k) = \begin{bmatrix} W_1^{[1]}(k) & \ldots & W_1^{[n_u]}(k) \\
\vdots & \ddots & \vdots \\
W_N^{[1]}(k) & \ldots & W_N^{[n_u]}(k)
\end{bmatrix}
$$

Matrices $Y(k)$ and $U(k)$ are composed likewise. The system $G$ is estimated using the Empirical Transfer Function Estimate (ETFE) (Pintelon and Schoukens, 2012, Ch. 2):

$$
\hat{G}(k) = Y(k)U^{-1}(k).
$$

The uncertainty of this estimate is expressed in terms of the covariance matrix

$$
C_G(k) = \left( S(k) \begin{bmatrix} W^{[1]}(k)W^{[1]H}(k) \\
\vdots \\
W^{[n_u]}(k)W^{[n_u]H}(k)
\end{bmatrix} S^H(k) \right)^{-1} \otimes C_Y(k).
$$

Herein,

$$
C_Y(k) = V(k) C_Z(k) V^H(k),
$$

$$
V(k) = \begin{bmatrix} I_{n_y} & -G(k) \end{bmatrix},
$$

and where $C_Z(k)$ is the covariance matrix associated to $[y^{[e]}(k)^T \ u^{[e]}(k)^T]^T$, see Pintelon and Schoukens (2012, Ch. 2).

2.3. Optimal experiment design objective

The OED objective in this paper is to compute the excitations $W = \{W(1), \ldots, W(N)\}$ that minimize a scalar cost function $J(W)$ that expresses, in some sense, the accuracy of estimate $\hat{G}$ in (4), while satisfying pre-specified signal constraints on the signals $w$, $u$, and/or $y$. Therefore, the OED problem is naturally posed as the constrained optimization problem,

$$
\begin{align*}
\text{minimize} & \quad J(W) \\
\text{subject to} & \quad g(W) \leq 0.
\end{align*}
$$

The criterion $J(W)$ and the inequality constraints $g(W)$ are specified in the forthcoming subsections.
Definition 1. The element-wise power of a scalar-valued signal $\xi_i^{[e]}(t)$, $i \in [1,n_\xi]$ out of a vector-valued signal $\xi^{[e]}(t) = [\xi_1^{[e]T}(t), \ldots, \xi_n^{[e]T}(t)]^T$ is given by

$$
P_i^{[e]} = \sum_{k=1}^N |\xi_i^{[e]}(k)|^2,
$$

where $\Xi_i^{[e]}$ is the DFT of $\xi_i^{[e]}$.

Let the signal vector $\xi^{[e]}$ be a function of the excitation signal $w^{[e]}$ through $\xi^{[e]} = G^{[e]}_\xi w^{[e]}$. Then, the power $P^{[e]}_\xi$ is expressed as

$$
P^{[e]}_\xi = \sum_{k=1}^N |\xi^{[e]}(k)|^2 = \sum_{k=1}^N \langle H^{[e]}_\xi(k), W^{[e]}(k) W^{[e]H}(k) \rangle,
$$

where $H^{[e]}_\xi(k) \triangleq G^{[e]^H}_\xi G^{[e]}_\xi$, $\forall k$ and $G^{[e]}_\xi$ denotes the $i$-th row of $G^{[e]}_\xi$. System $G^{[e]}_\xi$ depends on the selected signals in $\xi^{[e]}$. For example, $\xi^{[e]}$ could be composed as $\xi^{[e]} = [w^{[e]}T w^{[e]T} y^{[e]}T]^T$ such that $G^{[e]}_\xi = [I \ ST \ GST]^T$.

The use of constraints on the element-wise powers (9) instead of on the total power (8) complicates the design problem, since it typically results in a non-convex optimization problem, as will be shown in the following. This aspect plays a central role throughout this paper.

2.3.3. OED problem in explicit form

By combining the cost function (7) with the expression for element-wise powers (10), the general design problem (6) is formulated explicitly as

$$
\begin{align*}
\text{minimize} & \sum_{k=1}^N \gamma(k) \text{Tr} \left( S(k) \sum_{e=1}^{n_\xi} W^{[e]}(k) W^{[e]H}(k) S^H(k) \right)^{-1} \\
\text{subject to} & \sum_{k=1}^N \langle H^{[e]}_\xi(k), W^{[e]}(k) W^{[e]H}(k) \rangle \leq c^{[e]}_i, \forall i, e
\end{align*}
$$

(\mathcal{NL}P)

where $\gamma(k) \triangleq \text{Tr}(C^{[e]}_\xi(k))$ and $c^{[e]}_i$ contains the user-defined power limits. The constraints are scaled such that $\max_i c^{[e]}_i = 1$, without loss of generality. Program (\mathcal{NL}P) constitutes the minimization of an inverse quadratic form over an intersection of ellipsoids. This problem is non-convex and NP-hard in general (Nestorov, 1998). Moreover, since the number of scalar decision variables is given by $N_d = N(n_u^2 - n_u)$, the problem dimension tends to become very large, particularly for systems with a large number of inputs.

Solving the non-convex and potentially large-scale problem (\mathcal{NL}P) exactly or with a guaranteed good performance is non-straightforward. This paper presents methods to solve the multivariable OED problem.

2.3.4. Two-step design procedure

Solving the optimal experiment design problem (\mathcal{NL}P) requires prior system knowledge in the form of problem data $\gamma, S, H^{[e]}_\xi$. This data can be acquired from preliminary identification experiments, see e.g., Hildebrand and Gevers (2004); Dirkx et al. (2019a). Likewise, a two-step design approach is considered in this paper:

1. preliminary experiments for prior data acquisition,
2. optimal experiments to produce high quality FRFs.
3. Role of multivariable excitations in OED

In this section, the role of multivariable excitation design in the identification of MIMO systems is investigated. To this end, the limitations of commonly used traditional excitation techniques are pointed out. These limitations underline the necessity of multivariable approaches.

A traditional approach to performing the $n_o$ identification experiments is to excite a single system input at a time. Throughout, this approach will be referred to as the Single Input Multiple Outputs (SIMO) approach. SIMO excitations are of the form,

$$W^{\text{SIMO}}(k) = T_o(k),$$ (11)

where $T_o(k) \in \mathbb{C}^{n_a \times n_u}$ is a diagonal frequency-wise scaling matrix. The primary drawback of using SIMO excitations is that the MIMO system is treated as a multiple of SIMO systems, and thereby the potential in using the plurality of inputs is not exploited.

An alternative traditional approach that does make use of this plurality is the so-called orthogonal excitation design approach (Pintelon and Schoukens, 2012, Sec. 2.7.2). Orthogonal excitations are of the form,

$$W^{\text{orth}}(k) = T_o(k)T_o^T,$$ (12)

where $T_o$ is the (orthogonal) DFT matrix composed as

$$T_{o,m,n} = 1/\sqrt{n_u}e^{-j2\pi(m-1)(n-1)/n_u},$$ (13)

Basically, this DFT matrix performs a rotation of the excitation directions compared to SIMO excitations. Such rotation is beneficial in the case of excitation power constraints. This is illustrated in the following example.

Example 1 (Limitations of SIMO excitations). Consider the identification problem for a system $G = \begin{bmatrix} 1 & 0.7 \\ 0.8 & 1 \end{bmatrix}$ in open-loop, i.e., $S = I$. Let the constraints be imposed upon the excitations, i.e., $|W_i^{[e]}|^2 \leq 1$, $\forall i, e$. Let $C_Y = 1$, by which the cost function in (7) simplifies to $J(W) = \text{Tr} (WW^H)^{-1}$. The best achievable SIMO and orthogonal excitations are visualized in Fig. 5.

As illustrated, orthogonal excitations lead to improved accuracy upon SIMO excitations in the case of excitation power constraints. However, they do not provide a solution for the general case. This is illustrated in the following example, wherein output constraints are considered.

Example 2 (Limitations of orthogonal excitations). Consider the same identification problem as in Example 1, yet for output constraints, i.e., $|Y_i^{[e]}|^2 \leq 1$, $\forall i, e$. The best achievable SIMO and orthogonal excitations are visualized in Fig. 6.

The above examples illustrate that generally neither of the traditional design techniques lead to optimal performance, because the excitation directions are fixed by design. This limitation is addressed in multivariable excitation design, wherein the structure of $W(k)$ is unrestricted, enabling optimization of both the excitation magnitudes and directions.

Example 3 (Multivariable excitations). Consider the same identification problem as in Example 1. The multivariable excitations achieve optimal accuracy and are shown in green in Fig. 6.

The main point of the comparison above is to show that the optimal excitations have a frequency-wise directionality that depends on the directionality of the systems at hand, which is in sharp contrast to the SIMO and orthogonal excitations. This paper presents techniques to optimally design such multivariable excitations.

4. Solutions to the multivariable OED problem

This section presents methods to compute multivariable excitations. Since the OED problem (\text{\textgreek{A}\textgreek{L}\textgreek{P}}) is non-convex and generally NP-hard, the main focus is on solving the problem approximately.

In Section 4.1, (\text{\textgreek{A}\textgreek{L}\textgreek{P}}) is reformulated into a rank-one constrained Semi-Definite Program (SDP), which forms an intermediate yet instrumental step in solving the problem. In Section 4.2, exact solutions to two special cases are given. Section 4.3 presents a Sequential Semi-Definite Relaxation (SSDR) algorithm, wherein the non-convex problem is solved by a sequence of convex programs. In Section 4.4, a Relaxation and Randomization (RR) algorithm is presented that combines a convex relaxation with a subsequent randomization step. The organization of this section is schematically depicted in Figure 4.

All proofs are provided in the related technical report Dirkx et al. (2020).

4.1. Rank-constrained SDP and semidefinite relaxation

In this section, the non-convex problem (\text{\textgreek{A}\textgreek{L}\textgreek{P}}) is reformulated as a rank-constrained SDP, which forms a crucial intermediate step towards actually solving the problem. To this end, the product of vectors $W_i^{[e]}(k) \in \mathbb{C}^{n_u}$ in (\text{\textgreek{A}\textgreek{L}\textgreek{P}})
is ‘lifted’ to the space of rank-one positive semi-definite Hermitian matrices, i.e.,
\[ \Phi_{w}^{[e]}(k) \triangleq W^{[e]}(k)W^{[e]^H}(k), \quad \text{rank} \left( \Phi_{w}^{[e]}(k) \right) = 1, \quad (14) \]
\forall e, k, \text{ see, e.g., Nesterov (1998); Nemirovski et al. (1999).}

Matrix \( \Phi_{w}^{[e]}(k) \in \mathbb{H}^n_k \) represents the excitation spectrum associated with the \( e \)-th experiment. Performing the variable substitution (14) in problem (NLP) leads to a nonlinear program, constituting the minimization of an inverse form over a set of linear constraints and rank constraints. Next, the nonlinear cost function is ‘linearized’ by using the standard result that minimization of \( (S(k)\Phi_{w}(k)S^H(k))^{-1} \) is equivalent to minimizing an auxiliary variable \( Z(k) \in \mathbb{H}^n_k \) under the conic constraint \( Z(k) - S^{-H}(k)\Phi_{w}^{-1}(k)S^{-1}(k) \succeq 0 \). Finally, by applying the Schur-complement, this conic constraint is replaced by a LMI constraint. As a result, (NLP) is equivalent to the rank-one constrained semi-definite program
\[
\begin{align*}
\text{minimize} & \quad \sum_{k=1}^{N} \gamma(k) \text{Tr} (Z(k)) \\
\text{subject to} & \quad \begin{bmatrix} Z(k) & S^{-H}(k) \\
S^{-1}(k) & \sum_{e=1}^{n} \Phi_{w}^{[e]}(k) \end{bmatrix} \succeq 0, \quad \forall k \\
& \quad \sum_{k=1}^{N} (H_{\xi_i}(k), \Phi_{w}^{[e]}(k)) \leq c_{\xi_i}, \quad \forall i, e \\
& \quad \text{rank} \left( \Phi_{w}^{[e]}(k) \right) = 1, \quad \forall k, e.
\end{align*}
\]

Due to the non-convex and discontinuous nature of rank constraints, solving (RSDP) is as difficult as solving (NLP). However, a useful property of (RSDP) is that its fundamental difficulty is captured fully in the rank constraints. This enables formulation of the semi-definite relaxation of (RSDP), which is obtained by simply eliminating the rank
constraints, leading to

\[
\begin{align*}
\text{minimize} & \quad \sum_{k=1}^{N} \gamma(k) \text{Tr}(Z(k)) \\
\text{subject to} & \quad \begin{bmatrix} Z(k) & S^{-H}(k) \\ S^{-1}(k) & \sum_{c=1}^{n_u} \Phi_{uc}(k) \end{bmatrix} \succeq 0, \quad \forall k \\
& \quad \sum_{k=1}^{N} \langle H_{ik}(k), \Phi_{uc}(k) \rangle \leq c_{\xi_i}, \quad \forall i, \forall \epsilon.
\end{align*}
\]

Program \((\mathcal{SDR})\) is convex and can be solved efficiently, e.g., by interior-point algorithms (Nesterov, 1994; Boyd and Vandenberghe, 2004). Even though the solution to \((\mathcal{SDR})\) is generally not a solution to \((\mathcal{RSDP})\), it plays an instrumental role in certifying the quality of an (approximate) solution to \((\mathcal{RSDP})\). Concretely, the optimal cost value \(f_{\mathcal{SDR}}^*\) of \((\mathcal{SDR})\) provides a lower bound to the optimal cost value \(f_{\mathcal{RSDP}}^*\) of the original problem \((\mathcal{RSDP})\), i.e.,

\[f_{\mathcal{SDR}}^* \leq f_{\mathcal{RSDP}}^* . \]

Additionally, besides this lower bound for performance, the solution to \((\mathcal{SDR})\) plays a key role in the algorithms for solving the non-convex problem \((\mathcal{RSDP})\), which is presented in the following.

4.2. Exact solutions to special cases

In special cases, exact solutions to \((\mathcal{RSDP})\) can be obtained by (semi-)analytic approaches. Two such cases and their solutions are investigated in this section.

4.2.1. Rank-1 decomposition for 1 or 2 power constraints

In the special case of two or less power constraints, i.e., \(n_{\xi} \leq 2\), an exact solution to \((\mathcal{RSDP})\) can efficiently be obtained. This is achieved by first solving its relaxation \((\mathcal{SDR})\) and subsequently, at each frequency \(k\), decomposing the solution \(\Phi_{uc}^{(k)} = \sum_{c=1}^{n_u} \Phi_{uc}^{(c)}(k)\) into rank-one matrices \(\Phi_{uc}^{(c)}(k)\) that satisfy both \(\Phi_{uc}^{(k)} = \sum_{c=1}^{n_u} \Phi_{uc}^{(c)}(k)\) and the power constraints in \((\mathcal{RSDP})\). The critical step in this procedure is the rank-one decomposition. Unlike unstructured rank-one decomposition, which can efficiently be computed using the singular value decomposition (Eckart and Young, 1936), general solutions to structured rank-one decompositions, e.g., one that additionally satisfies constraints, do not exist (Markovsky, 2008). Though, for special cases, rank-one decomposition algorithms under linear constraints are reported in Sturm and Zhang (2001); Huang & Zhang (2007). Consider the following lemma.

**Lemma 1 (Theorem 2.1, (Huang & Zhang, 2007)).** Consider \(X \in \mathbb{H}_n^+\) with \(r = \text{rank}(X)\) and let \(A_1, A_2 \in \mathbb{H}_n^+\) be two given matrices. Then, there is a rank-one decomposition of \(X\) of the form \(X = \sum_{i=1}^{r} x_i x_i^H\) such that \((A_j, x_i x_i^H) = \frac{1}{r}(A_j, X), \quad i = 1, \ldots, r; \quad j = 1, 2.\)

Applying this structured rank-one decomposition to the solution of \((\mathcal{SDR})\) leads to the following result.

**Theorem 2.** When \(n_{\xi} \leq 2\), a solution \((\Phi_{uc}^{(c)}(k), Z^*)\) to program \((\mathcal{RSDP})\) exists that achieves \(f_{\mathcal{RSDP}}^* = f_{\mathcal{SDR}}^*\).

The rank-one decomposition in Lemma 1 can be performed efficiently (in polynomial time) through the constructive proof in (Huang & Zhang, 2007). As a result, the solution in Theorem 2 can exactly and efficiently be computed.

4.2.2. Closed-form solution for input power constraints

In the special case wherein power constraints are imposed only on the inputs \(u\), i.e., when \(G_{\xi}(k) = S(k), \forall k\), an analytic solution to \((\mathcal{L}_{\xi})\) exists:

**Theorem 3 (Corollary 1, (Dirkx et al., 2019a)).** If \(G_{\xi}(k) = S(k), \forall k\), a solution \(W^*\) to \((\mathcal{L}_{\xi})\) that achieves \(f_{\mathcal{L}_\xi}^* = f_{\mathcal{SDR}}^*\) exists and is given by

\[
W^*(k) = \sqrt{\eta(k)} S^{-1}(k) T_e T_o, \tag{15}
\]

where \(T_e = \sqrt{\text{diag}(c_{\xi})}, \eta(k) = n_u \sqrt{\gamma(k)} / \sum_{k=1}^{N} \sqrt{\gamma(k)}\) is a frequency-wise scaling factor and \(T_o\) is given in \((19)\).

The solution \((15)\) generalizes the orthogonal excitations reported in Guillaume et al. (1996); Dobrowiecki et al. (2006) to closed-loop systems.

4.3. Approximate solution to general case: a Sequential Semi-Definite Relaxation (SSDR) algorithm

Besides some special cases, no general methods to solve rank-constrained optimization problems exactly do exist (Markovsky, 2008). Therefore, this section presents the SSDR algorithm that aims to solve \((\mathcal{RSDP})\) by approximation. A key property of the SSDR algorithm is that its solution converges to a local optimum of \((\mathcal{RSDP})\).

4.3.1. SSDR algorithm

The main idea of the SSDR algorithm is to substitute the rank constraints by an approximate function that is convex, and then solve the resulting (sequence of) convex yet approximate optimization problem(s). The algorithm builds upon existing alternating projection based methods in Dattorro (2010); Delgado et al. (2014), and the iterative rank minimization or relaxation methods in Sun and Dai (2017); Cao et al. (2017).

The key principle behind the approximation of the rank constraint is based on the property that a \(n \times n\) matrix of rank \(r\) has \(n-r\) zero eigenvalues. Therefore, instead of imposing constraints on the rank of a matrix, constraints are imposed onto its \(n-r\) eigenvalues. Consider the following example.
Matrix $W$ in (16) is referred to as a direction matrix. A closed-form solution to (16) exists, namely, the direction matrix $W \in \Theta$, that minimizes (16) is $W = V_2 V_2^H$ where $V_2 \in \mathbb{C}^{n \times n-r}$ is the matrix of eigenvectors corresponding the $n-r$ smallest eigenvalues of $X$. Now, consider the following result.

**Theorem 5.** When $\varepsilon = 0$ and $X \in \mathbb{H}_n^+$, rank$(X) = 1$ holds if and only if

$$
\langle W, X \rangle \leq \varepsilon \quad \text{and} \quad \langle V_1 V_1^H, X \rangle > 0
$$

with $W = V_2 V_2^H$, and where $V_1 \in \mathbb{C}^{n \times 1}$ and $V_2 \in \mathbb{C}^{n \times n-1}$ contain the eigenvectors corresponding to the largest eigenvalue and to the $n-1$ smallest eigenvalues of $X$, respectively.

By Theorem 5, the non-convex constraint rank$(X) = 1$ can be substituted by the affine constraint (17). However, since the direction matrix $W$ depends directly on the eigenvectors of $X$, it cannot be computed before $X$ is solved. This dependency is addressed by performing an iterative procedure, wherein the direction matrix $W^{(j)}$ at an iteration $j$ is based on the solution $X^{(j-1)}$ obtained in the previous iteration $j-1$. In application to the OED problem ($\mathcal{RSDP}$), the SSDR algorithm solves the following convex program iteratively for $j = 1, \ldots, J_{\text{max}}$:

**Algorithm 1: Sequential Semi-Definite Relaxation**

**Input:** Problem data: $H_\xi, S, \gamma, \xi$. Algorithm parameters: $\alpha_0, e_{f_{\text{min}}}, e_{f_{\text{max}}}, J_{\text{max}}, \varepsilon_{\text{tol}}, \Delta f_{\text{tol}}$. **Output:** Excitation vectors $W^{(j)}, \varepsilon^{(j)}$

1: Initialize Set $j = 1$, $\alpha^{(1)} = 0$, $\varepsilon^{(0)} = \infty$, $W^{(0)} \in \mathbb{H}_n^+$
2: while $j \leq J_{\text{max}}$ & $\varepsilon > \varepsilon_{\text{tol}}$ & $f(j) - f(j-1) \geq \Delta f_{\text{tol}}$
3: Solve ($\mathcal{SSDR}$) and obtain $\Phi^{(j)}_w, Z^{(j)}, \varepsilon^{(j)}$
4: Compute $W^{(j)}(k) = V_2^{(j)}(k) V_2^{(j)H}(k)$
5: Compute $W^{(j)}(k) = Y_2^{(j)}(k) \sqrt{\lambda^{(j)}(k)}$ with $\Phi^{(j)}_w(k)$ and $\lambda^{(j)}(k)$ the principal eigenvector and eigenvalue of $\Phi^{(j)}_w(k)$, respectively
6: Compute true objective value $f(j) = J(W^{(j)})$
7: Update $\alpha^{(j)}$ through (18) and advance $j \leftarrow j + 1$
8: end while

where $f(j) = f(j)$, where $f(j) \triangleq J(W^{(j)})$, see (7), and $f(j) \triangleq \sum_{k=1}^N \gamma(k) \text{Tr} (Z^{(j)}(k))$. Since $e_f^{(j)} = 0$ if rank$(\Phi^{(j)}_w(k)) = 1$, $\forall e, k$, this error can be interpreted as a measure of the rank excess of $\Phi^{(j)}_w$. The weighting parameter $\alpha^{(j)}$ is then selected as:

$$
\alpha^{(j)} = \max \left( \min\{e_f^{(j-1)}, e_f^{\text{max}}\}, e_f^{\text{min}} \right) \alpha_0^{(j)}
$$

where and $e_f^{\text{max}} > e_f^{\text{min}} > 0$ are user-defined upper and lower bounds. Hence, this update law puts more weight on minimizing the rank when the rank excess is large, and vice versa. Additionally, setting $\alpha_0 > 0$ enforces the $n_u - 1$ smallest eigenvalues to zero.

**4.3.2. Convergence**

The convergence behavior of SSDR is characterized by the following lemma and theorem:

**Lemma 6.** Program ($\mathcal{SSDR}$) is feasible in each iteration $j \geq 1$ of Algorithm 1.

**Theorem 7.** When $\lim_{j \rightarrow \infty} \varepsilon^{(j)} = 0$, Algorithm 1 converges to a local minimizer of the original rank-constrained problem ($\mathcal{RSDP}$), i.e., $\lim_{j \rightarrow \infty} (\Phi^{(j)}_w, Z^{(j)}) = (\Phi^{*}_w, Z^*)$, $\forall e$ where $(\Phi^{*}_w, Z^*)$ is a local minimizer of ($\mathcal{RSDP}$).

**4.3.3. Discussion**

The SSDR algorithm provides a systematic approach to approximately solve the nonlinear and generally hard optimization problem ($\mathcal{LSDP}$) using convex optimization tools, e.g., the software package CVX (Grant and Boyd, 2014). Furthermore, convergence to a local optimum of the original problem is guaranteed by Theorem 7.

For large dimensional design problems, however, Algorithm 1 may impose a heavy computational burden, as
it requires solving a large dimensional SDP at each iteration. A second drawback of SSDR is that no quantitative performance guarantees can be given a priori, besides the qualitative guarantee of local optimality. In the next section, an alternative method is presented that addresses these two items.

4.4. Approximate solution to general case: a Relaxation and Randomization (RR) approach

This section presents the RR algorithm to generate multi-variable excitations that approximately solve (NLP). Theoretical approximation accuracy bounds and practical expected performance indicators are provided and the properties of the RR and the SSDR approach are compared.

4.4.1. RR algorithm

The key idea behind the RR method is to extract an approximate solution to (NLP) from the solution of the convex relaxation (SDR) in a randomized fashion. Even though the idea of a randomization-based approach is simple, many theoretically worst-case approximation bounds for semi-definite relaxation techniques are proven through randomization. Well-known results are those of Goemans and Williamson (1995); Nestorov (1998); Zhang (2000) for the max-cut problem, and non-convex quadratic maximization (Nemirovski et al., 1999).

The first step of the algorithm is to solve the convex program (SDR) to obtain the solution \( \Phi_{w}^{\dagger}(k) = \sum_{c=1}^{n} \Phi_{w}^{[c]}(k) \). This solution can be decomposed efficiently as \( \Phi_{w}^{\dagger}(k) = \Delta^{\dagger}(k) \Delta^{H}(k) \) using a (lower triangular) Cholesky decomposition. Different factorizations exist, since for any unitary matrix \( R(k) \), i.e., \( R(k)R^{H}(k) = I \), it holds that

\[
\Phi_{w}^{\dagger}(k) = \Delta^{\dagger}(k)R(k)R^{H}(k)\Delta^{H}(k).
\]

In the second step, the redundancy in (19), constituted by matrix \( R(k) \), is exploited to select the excitations \( W(k) \) in program (NLP). The resulting so-called random Haar unitary is constructed by the following procedure (Réfif, 2005, Ch. 3).

**Procedure 1** (Construction of random Haar unitary).

i) Choose the entries of \( R(k) \) as random mutually independent standard complex normal variables

ii) Perform the Gram-Schmidt orthogonalization procedure onto the column vectors of \( R(k) \).

Notice that since the spectrum \( \Phi_{w}^{\dagger}(k) \) in (19) is invariant to \( R(k) \), the objective value of (SDR) is also invariant to the choice \( R(k) \). This means that any excitation matrix \( W(k) \) constructed as \( W(k) = \Delta^{\dagger}(k)R(k) \) achieves the optimal cost value \( f^{\dagger}_{\text{SDR}} \). Such invariance does not apply to the element-wise power, which is expressed in terms of \( R(k) \) as

\[
P_{\xi,i}^{[c]} = \sum_{k=1}^{N} \left( \Delta^{H}(k)H_{\xi,i}(k)\Delta(k)R^{[c]}(k)R^{[c]H}(k) \right).
\]

**Algorithm 2** Relaxation and Randomization

**Input:** Problem data \( H_{\xi,i}, S, \gamma \). Stopping param.: \( j_{\text{max}} \)

**Output:** Excitations \( W^{*} \)

1. Initialize Set \( f^{\dagger}_{\text{SDR}} = \infty \)
2. Solve (SDR) to obtain \( \Phi_{w}^{\dagger}(k) = \sum_{c=1}^{n} \Phi_{w}^{[c]}(k) \)
3. Decompose \( \Phi_{w}^{\dagger}(k) \) as \( \Phi_{w}^{\dagger}(k) = \Delta^{\dagger}(k)\Delta^{H}(k) \), \( \forall k \)
4. while \( j < j_{\text{max}} \)
5. Construct excitations \( W(k) = \Delta^{\dagger}(k)R(k)\sqrt{T} \), \( \forall k \) with \( R(k) \) a random Haar unitary
6. Compute \( f_{j} = \sum_{k=1}^{N} (\gamma(k)\text{Tr}(S(k)W(k)R^{H}(k)R^{H}(k)))^{-1} \)
7. if \( f_{j} \leq f^{\dagger}_{\text{SDR}} \)
8. Store excitation \( W^{\dagger}(k) \leftarrow W(k) \)
9. Set \( f^{\dagger}_{\text{SDR}} \leftarrow f_{j} \), and advance \( j \leftarrow j + 1 \)
10. end if
11. end while

A feasible set of excitations \( W \), i.e., one that satisfies \( P_{\xi,i}^{[c]} \leq 1 \), \( \forall i, e \), is simply generated by choosing \( W(k) \) as:

\[
W(k) = \Delta^{\dagger}(k)R(k)\sqrt{T},
\]

with \( T = \text{diag}(1/n_{u_{1}}, \ldots, 1/n_{u_{n}}) \).

Based on this principle, an efficient algorithm for approximately solving (NLP) is presented in Algorithm 2.

4.4.2. Performance analysis

In this section, the performance of the RR algorithm is quantified by means of a stochastics-based performance analysis.

By constructing \( R(k) \) according to Procedure 1, the element-wise powers \( P_{\xi,i}^{[c]} \) in (20) become random variables with specific stochastic properties that are determined by the solution of (SDR). This is shown in the following lemma.

**Lemma 8.** Let \( R(k) \), \( \forall k \) be chosen as a random Haar unitary. Then, the element-wise powers \( P_{\xi,i}^{[c]} \) in (20) are random variables with expected value:

\[
E\{P_{\xi,i}^{[c]}\} = \frac{1}{n_{u}} \sum_{k=1}^{n_{u}} P_{\xi,i}^{[c]} \leq 1
\]

and variance:

\[
\sigma^{2}\{P_{\xi,i}^{[c]}\} = \frac{1}{n_{u}} \frac{n_{u} - 1}{n_{u} + 1} \|P_{\xi,i}^{[c]}\|_{2}^{2} \leq \frac{n_{u} - 1}{n_{u} + 1}
\]

where \( P_{\xi,i}^{[c]} = \sum_{e=1}^{n_{u}} |\xi_{i}^{[c]}(1)|^{2}, \ldots, \sum_{e=1}^{n_{u}} |\xi_{i}^{[c]}(N)|^{2} \).

Likewise, by choosing the excitations \( W \) according to (21), the achieved cost \( f_{j} \) of Algorithm 2 becomes a random variable with stochastic properties determined by those of \( P_{\xi,i}^{[c]} \).

This gives rise to the formulation of a theoretically worst-case approximation bound of (SDR), which is a measure of the achievable performance of Algorithm 2.
Consider the system described by mild conditions. Consider first the following lemma:

**Lemma 8.** If $1 - \frac{1}{n_u} \sum_{e=1}^{n_u} |\Xi_e(k)|^2 \leq 1$, $\forall i, k$ and $\sigma^2 \{ |\Xi_e(k)|^2 \} > 0$, $\forall i, e, k$, the powers $\mathcal{P}_e^{[k]}$ converge in distribution to a normal distribution as $N \to \infty$,

$$
\mathcal{P}_e^{[k]} \overset{d}{\to} \mathcal{N} \left( \mathbb{E}\{\mathcal{P}_e^{[k]}\}, \sigma^2 \{\mathcal{P}_e^{[k]}\} \right).
$$

By Lemma 10, the random variables $\mathcal{P}_e^{[k]}$ become normally distributed for $N \to \infty$. This property allows for formulation of a bound on the expected performance of the RR algorithm according to the following theorem:

**Theorem 11.** If the conditions in Lemma 10 hold, then the expected value of the cost $f_{NLP}^*$ of Algorithm 2 is upper bounded by

$$
\mathbb{E}\{f_{NLP}^*\} \leq \left( 1 + \sqrt{\max_{i,e} \sigma^2 \{\mathcal{P}_e^{[k]}\} (n_u n_e - 1)} \right) f_{SDR}^*. \tag{26a}
$$

as $N \to \infty$.

The expected performance bound (26a) is illustrated in the following example.

**Example 5 (Expected performance for large $N$).** Figure 7 illustrates the expected performance bound (26a) the same system as in Example 4. The expected cost is consistently lower than the upper-bound (24). For large $N$, the expected performance of the RR algorithm approaches the fundamentally best achievable performance, i.e., $f_{NLP}^*$ approaches $f_{SDR}^*$.

The bounds in Theorem 11 are based on a Central Limit Theorem (CLT) argument, and hence the result applies already to problems with ‘sufficiently large’ $N$. Commonly, $N > 30$ is considered as sufficiently large for the CLT to hold, which indicates the practical value of Theorem 11.

### 4.4.3. Discussion

The RR algorithm has several advantages over SSDR. First, it requires solving a SDP only once instead of at each iteration, by which the computational complexity is reduced substantially compared to SSDR. Furthermore, the RR approach provides a priori performance indicators, which is not the case for the SSDR algorithm, apart from the local optimality guarantees formulated in Theorem 7. Finally, the achieved performance using RR typically improves for identification problems with a larger number of excitation frequencies, which is particularly beneficial for many real-life identification problems.

### 5. Experimental validation

The developed experiment design algorithms are validated by means of experiments on the wafers stage system described in Section 2.1. This section presents the experimental goals, procedures, and results.
5.1. Experimental goal
The goal of the experimental validation is twofold: 1) to compare the SSDR and the RR algorithm in terms of achievable FRF accuracy and computation time and 2) to illustrate the benefit of multivariable over SIMO excitation design in practice.

5.2. Experiment design and results
Four sets of experiments are performed, resulting from:

1) Preliminary SIMO excitation design
2) Optimized SIMO excitation design
3) Optimized multivariable design, SSDR approach
4) Optimized multivariable design, RR approach

The resulting MIMO FRF is shown in Fig. 11. Herein, the 3×4 upper-left block represents the in-plane stage dynamics, while the 4×4 lower-right block represents the out-of-plane stage dynamics. The wafer stage is designed for high stiffness in the in-plane directions but is rather flexible in the out-of-plane directions, the latter of which is observed from the pronounced flexible dynamics in the 4×4 lower-right block. The results of the out-of-plane dynamics will be analysed in more detail in this section, while the results for the in-plane dynamics are discussed in the appendix. The out-of-plane entry $G_{5,5}^*$ is depicted in close-up in Fig. 9, including the 95% confidence interval. A detailed plot of the dynamics of entry $G_{5,5}^*$ at high frequencies is provided in the appendix. The improvement factors (in terms of the achieved cost value (7)) of the optimized designs with respect to the preliminary design are listed in Table 1. In addition, the computation times are indicated. The designs and results are elaborated below.

1) Preliminary experiments. The preliminary experiments serve solely for acquisition of the prior knowledge required for the subsequent optimal designs. The preliminary excitations are designed as SIMO excitations with a signal power of 0.25W, uniformly distributed over the frequency grid of $f_x = [10 : 5 : 4995]$Hz for each input.

The resulting FRF is shown in black in Fig. 11 and Fig. 9.

2) Optimized SIMO experiments. The SIMO excitations are synthesized using the convex optimization framework reported in Dirkx et al. (2019a). The optimization is performed under the input and output constraints as given in Section 2.1, though only the output constraints are limiting in the design (the input constraints are inactive at the optimal solution). Fig. 9 shows that optimized design results in a quality improvement in entry $G_{5,5}^*$ over the full frequency grid. A more detailed plot of the dynamics of entry $G_{5,5}^*$ at high frequencies is provided in the appendix. Overall, a factor 2.5 improvement is achieved over the preliminary design, see Table 1. The optimization algorithm requires 1 s computation time.

Table 1: Comparative results of the different sets of experiments, where #exc. indicates the number of simultaneously excited inputs per experiment, impr. fac. denotes the achieved improvement factor with respect to set 1), and comp. time is the computation time.

<table>
<thead>
<tr>
<th>Experiment set</th>
<th>#exc.</th>
<th>impr. fac.</th>
<th>comp. time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Prel. SIMO</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2) Opt. SIMO</td>
<td>1</td>
<td>2.5</td>
<td>1 s</td>
</tr>
<tr>
<td>3) SSDR</td>
<td>4</td>
<td>6.3</td>
<td>774 s</td>
</tr>
<tr>
<td>4) RR</td>
<td>8</td>
<td>14.2</td>
<td>125 s</td>
</tr>
</tbody>
</table>

Figure 8: Convergence of SSDR algorithm. Top: True objective value $f^{(j)}$ (x−), approximate objective value $\hat{f}^{(j)}$ (Δ−), and lower bound $f_{\text{ld}}^j$ (−). For visualization purposes, all curves are equally scaled such that $f_{\text{ld}}^j = 1$. Bottom: Monotonic decrease of variable $\varepsilon^{(j)}$.

3) Optimized multivariable experiments, SSDR approach. To attain a computationally tractable problem for SSDR, an excitation design is made wherein only a subset of 4 out of 8 inputs is excited per experiment, resulting in block-diagonal excitations $W(k)$ composed of two 4×4 blocks. The used settings are $j_{\text{max}} = 50$, $\alpha_0 = 1.05$, $e_{\text{fmin}} = 10^{-12}$, $e_{\text{fmax}} = 10^{-4}$, $\epsilon_{\text{tol}} = 10^{-4}$, $\Delta f_{\text{tol}} = 10^{-8}$. The top plot of Fig. 8 shows the convergence of the true and approximate cost value $f^{(j)}$ and $\hat{f}^{(j)}$, as well as the lower bound $f_{\text{ld}}^j$. During the first few iterations, the mismatch between $\hat{f}^{(j)}$ and $f^{(j)}$ is large, which implies a large rank excess. The mismatch reduces over the iterations. In 50 iterations, the cost $f^{(j)}$ has reached a value within a factor 1.5 from the fundamental lowerbound $f_{\text{ld}}^j$. The (slack) variable $\varepsilon^{(j)}$ in (SSDR) used to reduce the rank of the solution $\Phi_k^{(j)}$ is depicted in the bottom plot of Fig. 8.
This variable decreases monotonically, indicating a gradual approach of the rank constraints. Although an exact rank-one solution is not achieved within 50 iterations, i.e., $\hat{f}^{(50)} \neq 0$, the rank excess is rather small, which is supported (implicitly) by the small difference between $f^{(50)}$ and $\hat{f}^{(50)}$ of 0.1 shown in the top plot.

The $\hat{G}_{5,5}$ entry of the resulting FRF in green in Fig. 9 shows an improved quality using optimized multivariable excitations compared to optimized SIMO excitations. This is supported by the improvement factor of 6.3 in Table 1. SSDR requires a substantial computation time of 774 s, see Table 1, which is a main drawback of the algorithm.

4) Optimized multivariable experiments, RR approach. The RR algorithm is applied to synthesize a multivariable excitation design that fully exploits the plurality of the 8 actuators. The number of iterations is $j_{\text{max}} = 50$.

The performance of the algorithm is visualized in Fig. 10. In 50 iterations, a cost value has been achieved that lies within a factor 1.6 of the theoretical lower bound $f^\dagger_{\text{SDR}}$. The mean cost value is 1.8 and is below its expected upper-bound of 2.4 computed by (26a).

The resulting FRF in Fig. 9 shows a further quality improvement compared the SSDR algorithm, which is supported by the improvement factor of 14.2 in Table 1. It should be noted that the achieved improvement over SSDR cannot be ascribed to the algorithm itself, but is due to the use of all 8 actuators simultaneously, rather than a subset of 4, as in the SSDR design.

The computation time of 125 s of the RR algorithm is significantly lower than of the SSDR algorithm, demonstrating that the RR approach is well-suited for application to large dimensional identification problems in practice.

6. Conclusions

The identification of FRFs of mechatronic systems can be made more accurate and/or faster through the multivariable experiment design framework presented in this paper. The developed methods improve upon traditional excitation design techniques by exploiting the design freedom in the multivariable excitation signals, while efficiently dealing with the element-wise constraints of the system. This improvement is supported by experimental results from a 7 × 8 wafer stage, which show a significantly enhanced FRF quality compared to traditional approaches. Furthermore, the results demonstrate that the developed computational techniques are fast and are suitable for application to large dimensional realistic systems. Future work will focus on extending theoretical results and handling different types of constraints.

Appendix A. Additional FRF results

This section discusses the experimental results of additional entries of the FRF in Fig. 11 and zooms in on the high frequency results of entry $\hat{G}_{5,5}$ shown in Fig. 9.

Appendix A1. FRF results of $\hat{G}_{1,1}$ and $\hat{G}_{2,1}$

Since the wafer stage has a high stiffness in the in-plane directions, the flexible dynamics in these directions are not as well-pronounced as in the out-of-plane directions. Consequently, the in-plane stage dynamics are predominantly described by rigid body dynamics (a -2 slope in the Bode magnitude diagram) up to the resonance frequency.
Figure 11: Full 7 x 8 Bode magnitude plot of the FRF obtained from the preliminary SIMO (black), optimized SIMO (blue), optimized multivariable SSDR (green), and optimized multivariable RR (red) measurements.

at approximately 1700Hz. This is observed in entry $\hat{G}_{2,1}$, depicted in Fig. A.12(b). Regarding FRF quality, it is observed that an improvement is achieved using optimized multivariable excitations compared to optimized SIMO excitations, particularly by the RR algorithm. The results are in line with the observations in the out-of-plane entries, as discussed in Section 5.

The results for entry $\hat{G}_{1,1}$ are depicted in Fig. A.12(a). In this entry, no rigid body dynamics are observed. Instead, the FRF is dominated by measurement noise with an increasing magnitude for lower frequencies. These observations are explained as follows. At the location of sensor $y_1$, see Figure 2, no rigid body displacements occur due to inputs $u_1$, since the resulting stage rotation and translation resulting cancel each other out. Furthermore, due to the high in-plane stiffness, the contribution of the remaining flexible dynamics that can be measured in entry $\hat{G}_{1,1}$ is small and falls almost entirely in the uncertainty region.

In terms of absolute uncertainty, represented by the standard deviations in the bottom plots, it is observed that entry $\hat{G}_{1,1}$ is identified with a similar accuracy as entry $\hat{G}_{2,1}$.

Appendix A.2. FRF results of $\hat{G}_{5,5}$, high frequency region

The FRF of entry $\hat{G}_{5,5}$ is depicted in detail for the high frequency range in Fig. A.12(c). The confidence intervals have been omitted for the sake of visualization. The identified FRF using multivariable excitation using the RR approach (red) shows the smallest uncertainty, and provides the clearest insight in the flexible dynamics of the wafer stage, particularly in the frequency region up to 2kHz.

References


Hjalmarsson H. From experiment design to closed-loop control. Automatica, 2005; 41(3):393 - 438. ISSN 0005-1098.


Nemirovski A, Roos C, Terlaky T. On Maximization of Quadratic form over Intersection of Ellipsoids with Common Center. 1999; Mathematical Programming. 86. 10.1007/s101070050100.


