Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Multivariable Experiment Design with Application to a Wafer Stage: a Sequential Relaxation Approach for Dealing with Element-Wise Constraints ⋆

Nic Dirkx *,** **Tom Oomen** **

* *ASML Research Mechatronics & Control, Veldhoven, The Netherlands*
** *Eindhoven University of Technology, Department of Mechanical Engineering, Control Systems Technology, Eindhoven, The Netherlands*

Abstract: Optimal Experiment Design (OED) is an essential aspect in accurate Frequency Response Function (FRF) identification of complex systems. The aim of this paper is to optimally design experiments for FRF identification of multivariable motion systems subject to element-wise power constraints. This design problem involves solving a non-convex and generally NP-hard optimization problem. An algorithm to solving this problem approximately is presented based on sequential semi-definite relaxations. Experimental results on a wafer stage show an improvement of the FRF quality using the proposed techniques over traditional excitation design methods.

Keywords: Optimal experiment design, system identification, multivariable systems, frequency response function, multisines, rank constrained optimization, non-convex optimization

1. INTRODUCTION

Ever increasing performance requirements for high precision positioning systems necessitate the availability of accurate models (Oomen, 2018), both for control design (Karimi and Kammer, 2016) and diagnostics (Van der Maas et al., 2016). These motion systems, including wafer scanners, printing systems, and medical scanners, typically exhibit complex multivariable dynamics, e.g., due to flexible mechanics. For such systems, the identification of FRF models from experimental data, as opposed to first principles modeling, is considered fast, inexpensive, and accurate (Ljung, 1999; Pintelon and Schoukens, 2012).

The choice of input signal is crucial for the quality of the identified FRF, which motivates optimal input design strategies. Optimal experiment design involves the optimization of the inputs to maximize the model accuracy with limited resources, e.g., under bounded inputs or outputs. OED is particularly relevant for Multiple Inputs Multiple Outputs (MIMO) systems, as the number of experiments increases with the number of system inputs.

Optimal experiment design for FRF identification of MIMO systems is addressed in Pintelon and Schoukens (2012), Ch 2,5; Guillaume et al. (1996); Dobrowiecki et al. (2006). Herein, a multiple input signal design approach is proposed using orthogonal excitations. Such orthogonal excitations are optimal (in view of various classical optimality criteria) in the case of input power constraints. However, they do not provide an optimal solution to the general case, e.g., under bounded inputs and outputs. Indeed, due to the inherent directionality in MIMO systems, the optimal excitations, in general, are not orthogonal but have directionality that depends on the system itself (Dirkx et al., 2019).

⋆ 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).

Optimization-based approaches to experiment design 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 problem as the minimization of a convex objective function over a convex set (Jansson, 2004). This concept is also used in Barenthin et al. (2008); Mehra (1974), with a specific focus on OED for MIMO systems. Herein, convexity of the optimization problem is preserved by considering input design problems subject to constraints on total input or total output power, which is affine in the excitation spectrum and hence leads to a convex constraint set.

For multivariable motion systems, however, total power constraints are often not representative for the actual physical limitations of the system. For such systems, limitations related to specific elements of the system are much more important, e.g., bounded power for a specific actuator. A framework for OED design for FRF identification for multivariable systems subject to element-wise constraints is presented in Dirkx et al. (2019). It is shown that the problem consists in solving a rank-constrained optimization problem, and although an exact solution for a special case is reported, no computational tools are available yet to solve the involved non-convex problem.

The aim of this paper is to solve the rank-constrained OED problem for FRF identification of complex multivariable systems under element-wise power constraints. Rank constraints in optimization are considered highly challenging and general methods to solve such problems exactly do not exist (Markovsky, 2008). Therefore, existing methods typically aim to arrive at an approximate solution by approximating the original problem in a certain sense, e.g., through combined linearization and factorization (Hassibi et al., 1999), through a rank approximation heuristic (Recht et al., 2007), or through alternating projection-based schemes, e.g., (Dattorro, 2010; Delgado et al., 2014).

Typically, the choice and design of such methods is largely determined by the specific structure of the problem at hand.
The main contributions in this paper are the following:
C1 A Sequential Semi-Definite Relaxation (SSDR) algorithm to solve the multivariable OED problem in the general case, in Section 3.
C2 An experimental validation of the SSDR algorithm on a prototype wafer stage, in Section 4.

Notations and definitions $\mathbb{H}^n$ denotes the set of $n \times n$ Hermitian matrices. Subscripts (+)+ denote positive (semi)-definiteness of a matrix or (semi)-positivity of a vector. $\overline{X}$ represents the complex conjugate of $X \in \mathbb{C}$. Operation $\langle A, X\rangle = \sum_{i,j} \overline{A_{ij}} X_{ij} = \mathrm{Tr}(A^H X)$ denotes the Frobenius inner product of equal sized matrices A and X. The $\otimes$ operator denotes the Kronecker product. The Discrete Fourier Transform (DFT) of a discrete time sampled signal $x(t)$ is defined as

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

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 chosen over the conventional factor $1/\sqrt{N}$ to simplify the equations in this paper.

2. PROBLEM DEFINITION

2.1 Identification framework

Identification scheme Consider the identification scheme in Fig. 1, where G represents the $n_y \times n_u$ LTI system to be identified. K is a (given) feedback controller. Throughout, $w \in \mathbb{R}^{n_u}$ is a multisine excitation signal (Pintelon and Schoukens (2012), Ch. 2, Ch. 5). Signals w, u and y are measured for system identification. The signal y is corrupted by measurement noise ν_y, characterized as a filtered independent and identically distributed (iid) random sequence. The DFT of the signals $u(t)$ and $y(t)$ measured during an e-th experiment, with $e = 1, \ldots, n_u$, are given by

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

with $W^{[e]}(k) \in \mathbb{C}^{n_u}$ the DFT of w and $V_z^{[e]}(k) \in \mathbb{C}^{n_y+n_u}$ of the noise contribution of ν_y onto $z = [y^T \; u^T]$. Furthermore, $S(k) = (I + K(k)G(k))^{-1}$ denotes the Sensitivity function. The matrix $W(k)$ is formed by performing a number of n_u experiments as

$$W(k) = [\, W^{[1]}(k) \; \ldots \; W^{[n_u]}(k) \,] = \begin{bmatrix} W_1^{[1]}(k) & \ldots & W_1^{[n_u]}(k) \\ \vdots & \ddots & \vdots \\ W_{n_u}^{[1]}(k) & \ldots & W_{n_u}^{[n_u]}(k) \end{bmatrix}. \quad (2)$$

Matrices $Y(k)$, $U(k)$, and $Z(k)$ are constructed similarly. The plant estimate

$$\hat{G}(k) = Y(k) U^{-1}(k). \quad (3)$$

is used throughout.

2.2 Formulation of the multivariable OED problem

The goal of OED is to design the excitations $W = [W(1), \ldots, W(N)]$ such to minimize a scalar cost function

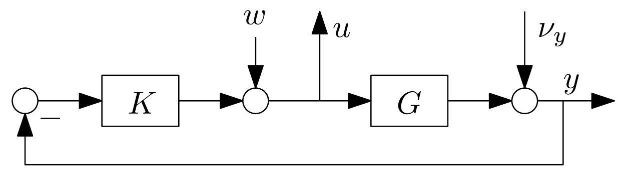

Fig. 1. Closed-loop identification scheme.

$\mathcal{J}(W)$ related to the accuracy of estimate $\hat{G}$ in (3), while satisfying specified signal constraints, e.g., on the signals w, u, or y. The OED problem is therefore naturally posed as,

$$\begin{aligned} \underset{W(k)\in\mathbb{C}^{n_u \times n_u}}{\text{minimize}} \quad & \mathcal{J}(W) \\ \text{subject to} \quad & g(W) \leq 0. \end{aligned} \quad (4)$$

Its components are discussed in more detail in the following, and the explicit form of (4) is presented.

Cost function The cost function in (4) is chosen as the A-optimality criterion,

$$\mathcal{J}(W) = \sum\nolimits_{k=1}^{N} \mathrm{Tr}\left(\mathcal{C}_{\hat{G}}(k)\right), \quad (5)$$

where $\mathcal{C}_{\hat{G}}(k) \in \mathbb{H}_+^{n_y n_u}$ is the covariance matrix associated to the estimation uncertainty in $\hat{G}(k)$,

$$\mathcal{C}_{\hat{G}}(k) = \overline{\left(S(k) \sum_{e=1}^{n_u} W^{[e]}(k) W^{[e]H}(k) S^H(k) \right)^{-1}} \otimes \mathcal{C}_Y(k). \quad (6)$$

Herein,

$$\mathcal{C}_Y(k) = V(k)\, \mathcal{C}_Z(k) V^H(k)$$
$$V(k) = \left[\, I_{n_y} \; -G(k) \,\right],$$

and $\mathcal{C}_Z(k)$ is the covariance matrix associated to $Z(k)$, see Pintelon and Schoukens (2012), Ch 2. Cost function (5) represents the total variance over all frequencies and all entries of the identified FRF of G. Expression (6) requires prior knowledge of $S(k)$, $G(k)$, and $\mathcal{C}_Z(k)$. These quantities can be estimated from a preliminary experiment as illustrated in Dirkx et al. (2019).

Constraints Throughout, the focus is on signal power constraints. Such constraints are directly relevant in typical mechatronic systems, e.g., due to actuator or amplifier limitations. Moreover, a power-constrained design may serve as a preliminary design to a further peak amplitude constrained signal design (Manchester, 2009).
In traditional OED for MIMO systems, constraints on the *total* power are considered, e.g., in Barenthin et al. (2008); Mehra (1974). To understand the notion of total power, consider a set of vector-valued signals, say $\xi^{[e]}(t) \in \mathbb{R}^{n_\xi}$, $e = 1, \ldots, n_u$ with DFT $\Xi(k)$, structured as in (2). The total power is then defined as the scalar quantity $\mathcal{P}_\xi = \sum_{k=1}^{N} \mathrm{Tr}\left(\Xi(k)\Xi^H(k)\right)$ and represents the sum of powers over all n_ξ signals and all n_u experiments.
Although the use of total power constraints eases the design problem, as it typically leads to a convex constraint set, the total power does not necessarily represent a relevant or even meaningful physical quantity in MIMO systems. For motion systems in particular, the truly physically relevant powers are typically those of each of the individual elements $\xi_i^{[e]}(t)$, instead of of their sum. For instance, a signal element $\xi_i^{[e]}(t)$ may directly represent a voltage or current requested from an ith actuator during

the eth experiment. This motivates the use of *element-wise* power constraints, wherein the power in each of the signals and experiments are explicitly distinguished. The element-wise power is formally defined as:

Definition 1. The element-wise power of a scalar-valued signal $\xi_i^{[e]}(t), \quad i \in [1, n_\xi]$ out of a vector-valued signal $\xi(t) = [\xi_1^{[e]T}(t), \ldots, \xi_{n_\xi}^{[e]T}(t)]^T$ is given by

$$\mathcal{P}_{\xi_i}^{[e]} = \sum\nolimits_{k=1}^{N} |\Xi_i^{[e]}(k)|^2$$

with $\Xi_i^{[e]}(k)$ the DFT of $\xi_i^{[e]}(t)$.

Let the signal vector ξ be a function of the excitation signal w through $\xi = G_\xi w$, where system G_ξ depends on the selected signals in ξ. For example, ξ could be composed as $\xi = [w^T \; u^T \; y^T]^T$ such that $G_\xi = [I \quad S^T \; GS^T]^T$. By Definition 1, the associated element-wise power is expressed as

$$\mathcal{P}_{\xi,i}^{[e]} = \sum_{k=1}^{N} |\Xi_i^{[e]}(k)|^2 = \sum_{k=1}^{N} \langle H_i(k), W^{[e]}(k) W^{[e]H}(k) \rangle \quad (7)$$

where $H_{\xi_i}(k) \triangleq G_{\xi_i}^H(k) G_{\xi_i}(k)$, $\forall k$ and G_{ξ_i} denotes the i-th row of G_ξ.

OED problem in explicit form ($\mathcal{NLP}$) Combining the power equations (7) with the cost function (5), the general OED problem (4) to be solved is formulated explicitly as

$$\begin{aligned} \underset{W^{[e]}(k) \in \mathbb{C}^{n_u}}{\text{minimize}} \quad & \sum_{k=1}^{N} \gamma(k) \text{Tr} \left(S(k) \sum_{e=1}^{n_u} W^{[e]}(k) W^{[e]H}(k) S^H(k) \right)^{-1} \\ \text{subject to} \quad & \sum_{k=1}^{N} \langle H_{\xi_i}(k), W^{[e]}(k) W^{[e]H}(k) \rangle \le c_{\xi_i}, \; \forall i, e \end{aligned} \quad (\mathcal{NLP})$$

where $\gamma(k) \triangleq \text{Tr}(\mathcal{C}_Y(k))$ and c_ξ is the user-defined vector containing the power limits. $\text{Tr}(A \otimes B) = \text{Tr}(A)\text{Tr}(B)$ is applied to form the cost function.
Program ($\mathcal{NLP}$) is the minimization of an inverse quadratic form over an intersection of ellipsoids. This problem is non-convex and NP-hard in general, hence solving this problem non-straightforward. An algorithm to solve ($\mathcal{NLP}$) approximately is presented in the next section.

3. SEQUENTIAL SEMI-DEFINITE RELAXATION APPROACH

This section constitutes contribution C1.

The OED problem in ($\mathcal{NLP}$) is non-convex and generally NP-hard. For a special case, an exact analytic solution is reported in Dirkx et al. (2019). For the general case, such exact solutions are not known to exist. Therefore, this section presents a method that solves the OED problem ($\mathcal{NLP}$) approximately. This method consists of 2 steps:

1) A reformulation of problem ($\mathcal{NLP}$) into a rank-constrained Semi-Definite Program (SDP),
2) A Sequential Semi-Definite Relaxation (SSDR) algorithm.

The steps are explained in the following.

1) Rank-constrained SDP and semidefinite relaxation

The non-convex problem ($\mathcal{NLP}$) is reformulated as a rank-constrained SDP. Such reformulation is achieved by first lifting the product of vectors in $W^{[e]}(k) \in \mathbb{C}^{n_u}$ to the space of rank one Hermitian positive semi-definite matrices in $\mathbb{H}_+$,i.e.,

$$\Phi_w^{[e]}(k) \triangleq W^{[e]}(k) W^{[e]H}(k), \quad \text{rank}(\Phi_w^{[e]}(k)) = 1,$$

$\forall e, k$, see, e.g., Nesterov (1998); Nemirovski et al. (1999). Matrix $\Phi_w^{[e]}(k) \in \mathbb{H}_+^{n_u}$ represents the *experiment-wise* excitation spectrum. Expressing problem ($\mathcal{NLP}$) in terms of the new matrix variables $\Phi_w^{[e]}(k)$ leads to a nonlinear minimization problem of a matrix inverse objective over a set of linear constraints and rank constraints. Subsequently, the nonlinear objective is transformed to a linear one by using the standard result that minimization of $\big(S(k)\Phi_w(k)S^H(k)\big)^{-1}$ is equivalent to the minimization of its epigraph (Boyd and Vandenberghe, 2004), which is achieved through minimizing auxiliary variable $Z(k) \in \mathbb{H}^{n_u}$ subject to the conic constraint $Z(k) - S^{-H}(k)\Phi_w^{-1}(k)S^{-1}(k) \succeq 0$. Then, the Schur-complement is applied to replace this conic constraint by a LMI constraint. This allows reformulation of ($\mathcal{NLP}$) as

$$\begin{aligned} \underset{\Phi_w^{[e]}(k), Z(k) \in \mathbb{H}_+^{n_u}}{\text{minimize}} \quad & \sum\nolimits_{k=1}^{N} \gamma(k) \text{Tr}\,(Z(k)) & (\mathcal{RSDP}) \\ \text{subject to} \quad & \begin{bmatrix} Z(k) & S^{-H}(k) \\ S^{-1}(k) & \sum_{e=1}^{n_u} \Phi_w^{[e]}(k) \end{bmatrix} \succeq 0, & \forall k \\ & \sum\nolimits_{k=1}^{N} \langle H_{\xi_i}(k), \Phi_w^{[e]}(k) \rangle \le c_{\xi_i}, & \forall i, e \\ & \text{rank}\left(\Phi_w^{[e]}(k)\right) = 1, & \forall k, e. \end{aligned}$$

Program ($\mathcal{RSDP}$) is a rank-constrained SDP and is equivalent to ($\mathcal{NLP}$). The solution $W^{[e]*}(k)$ to ($\mathcal{NLP}$) can be uniquely reconstructed (up to the sign) from the solution $\Phi_w^{[e]*}(k)$ of ($\mathcal{RSDP}$) through an Eigendecomposition. Due to the discontinuous and non-convex nature of rank constraints, solving ($\mathcal{RSDP}$) is still not straightforward. However, the fundamental difficulty of program ($\mathcal{RSDP}$) is captured fully in the rank constraint. This enables relaxing the problem by eliminating the rank constraints, which leads to:

$$\begin{aligned} \underset{\Phi_w^{[e]}(k), Z(k)}{\text{minimize}} \quad & \sum\nolimits_{k=1}^{N} \gamma(k) \text{Tr}\,(Z(k)) & (\mathcal{SDR}) \\ \text{subject to} \quad & \begin{bmatrix} Z(k) & S^{-H}(k) \\ S^{-1}(k) & \sum_{e=1}^{n_u} \Phi_w^{[e]}(k) \end{bmatrix} \succeq 0, & \forall k \\ & \sum\nolimits_{k=1}^{N} \langle H_{\xi_i}(k), \Phi_w^{[e]}(k) \rangle \le c_{\xi_i}, & \forall i, e \end{aligned}$$

This semidefinite relaxation is convex and can be solved efficiently, e.g., by interior-point algorithms (Nesterov and Nemirovski, 1994; Boyd and Vandenberghe, 2004), such as implemented in package `CVX` (Grant and Boyd, 2014).
The solution to ($\mathcal{SDR}$), say $\Phi_w^{[e]\dagger}(k)$, is generally not of rank one and hence it is not a solution to ($\mathcal{RSDP}$). Nevertheless, the relaxation ($\mathcal{SDR}$) provides valuable information on the solution of the original problem ($\mathcal{RSDP}$). To see this, let the optimal objective values of the two programs be defined as $f_{\mathcal{SDR}}^\dagger$ and $f_{\mathcal{RSDP}}^*$, respectively. Then, $f_{\mathcal{SDR}}^\dagger$ forms a lower bound to $f_{\mathcal{RSDP}}^*$, i.e., $f_{\mathcal{SDR}}^\dagger \le f_{\mathcal{RSDP}}^*$. Although no formal proof will be given in this paper, for many practical OED

Algorithm 1 Sequential Semi-Definite Relaxation

Input: Problem data: $H_{\xi_i}, S, \gamma, c_\xi$. Algorithm parameters: $\alpha_0, e_{f_{\min}}, e_{f_{\max}}, j_{\max}, \varepsilon_{\text{tol}}, \Delta f_{\text{tol}}$. **Output**: Excitation vectors $W^{[e]}$

Initialize Set $j = 1$, $\alpha^{\langle 1 \rangle} = 0$, $\varepsilon^{\langle 0 \rangle} = \infty$, $\mathcal{W}^{[e]\langle 0 \rangle}(k) = I_{n_u}$
while $j \leq j_{\max}$ & $\varepsilon \geq \varepsilon_{\text{tol}}$ || $\frac{f^{\langle j \rangle} - f^{\langle j-1 \rangle}}{f^{\langle j-1 \rangle}} \geq \Delta f_{\text{tol}}$
Solve ($\mathcal{SSDR}$) and obtain $\Phi_w^{[e]\langle j \rangle}, Z^{\langle j \rangle}, \varepsilon^{\langle j \rangle}$
Compute $\mathcal{W}^{[e]\langle j \rangle}(k) = \mathcal{V}_2^{[e]\langle j \rangle}(k)\mathcal{V}_2^{[e]\langle j \rangle H}(k)$
Compute $W^{[e]\langle j \rangle}(k) = \mathcal{V}_1^{[e]\langle j \rangle}(k)\sqrt{\lambda_1^{[e]\langle j \rangle}(k)}$ with $\mathcal{V}_1^{[e]\langle j \rangle}(k)$ and $\lambda_1^{[e]\langle j \rangle}(k)$ the principal eigenvector and eigenvalue of $\Phi_w^{[e]\langle j \rangle}(k)$, respectively
Compute true objective value $f^{\langle j \rangle} = \mathcal{J}(W^{\langle j \rangle})$ by (5)
Update $\alpha^{\langle j \rangle}$ through (10) and update $j \leftarrow j + 1$
end while

problems the gap is quite small. This will also be shown experimentally in Section 4.
Besides providing a lower bound for performance, program ($\mathcal{SDR}$) also plays a key role in the algorithm for solving the non-convex problem ($\mathcal{RSDP}$), which is presented next.

2) Sequential Semi-Definite Relaxation (SSDR) algorithm

The main idea of the SSDR algorithm is to replace the troublesome rank constraint in ($\mathcal{RSDP}$) by an approximative function that is convex, and then solve the resulting (sequence of) convex optimization problem(s). The algorithm is inspired by 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).

Rank approximation The key principle behind SSDR is based on the property that a $n \times n$ matrix of rank r has $n - r$ zero eigenvalues. Therefore, instead of imposing rank constraints, constraints are imposed onto the $n - r$ eigenvalues. The sum of the $n - r$ smallest eigenvalues of a matrix are computed by a result of Fan (1950):

Lemma 1. Consider $X \in \mathbb{H}_+^n$ with eigenvalues in descending order, i.e., $\lambda_1(X) \geq \ldots \geq \lambda_n(X)$, then

$$\sum_{i=r+1}^{n} \lambda_i(X) = \underset{\mathcal{W} \in \Theta_r}{\text{minimize}} \ \langle \mathcal{W}, X \rangle \tag{8}$$

with set $\Theta_r = \{\mathcal{W} \in \mathbb{H}_+^n | \ 0 \preceq \mathcal{W} \preceq I, \ \text{Tr}(\mathcal{W}) = n - r\}$.

Matrix $\mathcal{W}$ in (8) is referred to as a direction matrix. The feasible set Θ_r of direction matrices is the convex hull of outer product of all rank $(n - r)$ orthonormal matrices (Dattorro, 2010). A closed-form solution to (8) exists, stated in the following lemma:

Lemma 2. The direction matrix $\mathcal{W} \in \Theta_r$ that minimizes (8) is $\mathcal{W} = \mathcal{V}_2\mathcal{V}_2^H$ where $\mathcal{V}_2 \in \mathbb{C}^{n \times n-r}$ is the matrix of eigenvectors corresponding the $n - r$ smallest eigenvalues of X.

Now, the rank of a matrix is expressed in terms of the direction matrix according to:

Corollary 1. When $\varepsilon = 0$ and $X \in \mathbb{H}_+^n$, $\text{rank}(X) = 1$ holds if and only if

$$\langle \mathcal{W}, X \rangle \leq \varepsilon \quad \text{and} \quad \langle \mathcal{V}_1\mathcal{V}_1^H, X \rangle > 0 \tag{9}$$

where $\mathcal{W} = \mathcal{V}_2\mathcal{V}_2^H$ and with $\mathcal{V}_1 \in \mathbb{C}^{n \times 1}$ and $\mathcal{V}_2 \in \mathbb{C}^{n \times n-1}$ containing the eigenvectors corresponding to the largest eigenvalue and to the $n - 1$ smallest eigenvalues of X, respectively.

Application of Corollary 1 enables substituting the non-convex constraint $\text{rank}(X) = 1$ by the affine constraint (9). However, note that the direction matrix $\mathcal{W}$ depends directly on the eigenvectors of X and hence cannot be computed before X is solved.

SSDR algorithm The SSDR algorithm addresses this dependency of $\mathcal{W}$ on X by performing a sequential procedure. Herein, the direction matrix $\mathcal{W}^{\langle j \rangle}$ at an iteration j is based on the solution $X^{\langle j-1 \rangle}$ obtained in the previous iteration $j - 1$. Applying the SSDR algorithm to the OED problem ($\mathcal{RSDP}$), the following convex program is solved iteratively for $j = 1, \ldots, j_{\max}$:

$$\begin{aligned}
\underset{\substack{\Phi_w^{[e]\langle j \rangle}(k),\, Z^{\langle j \rangle}(k),\\ \varepsilon^{\langle j \rangle} \in \mathbb{R}_+}}{\text{minimize}} \quad & \sum_{k=1}^{N} \gamma(k)\text{Tr}\left(Z^{\langle j \rangle}(k)\right) + \alpha^{\langle j \rangle}\varepsilon^{\langle j \rangle} \qquad (\mathcal{SSDR})\\
\text{subject to} \quad & \begin{bmatrix} Z^{\langle j \rangle}(k) & S^{-H}(k) \\ S^{-1}(k) & \sum_{e=1}^{n_u} \Phi_w^{[e]\langle j \rangle}(k) \end{bmatrix} \succeq 0, \qquad \forall k\\
& \sum\nolimits_{k=1}^{N} \langle H_{\xi_i}(k), \Phi_w^{[e]\langle j \rangle}(k) \rangle \leq c_{\xi_i}, \qquad \forall i, e\\
& \sum_{k=1}^{N}\sum_{e=1}^{n_u} \langle \mathcal{W}^{[e]\langle j-1 \rangle}(k), \Phi_w^{[e]\langle j \rangle}(k) \rangle \leq \varepsilon^{\langle j \rangle}\\
& \varepsilon^{\langle j \rangle} \leq \varepsilon^{\langle j-1 \rangle}.
\end{aligned}$$

Herein, $\mathcal{W}^{[e]\langle j \rangle}(k) = \mathcal{V}_2^{[e]\langle j \rangle}(k)\mathcal{V}_2^{[e]\langle j \rangle H}(k)$ with $\mathcal{V}_2^{[e]\langle j \rangle}(k) \in \mathbb{C}^{n_u \times n_u - 1}$ the eigenvectors corresponding to the $n_u - 1$ smallest eigenvalues of $\Phi_w^{[e]}(k)$. The slack variable $\varepsilon^{\langle j \rangle} \in \mathbb{R}_+$ is minimized by augmenting it as a penalty term to the cost function, weighted by the externally controlled parameter $\alpha^{\langle j \rangle} \in \mathbb{R}_+$. Additionally, the bottom constraint enforces a monotonic decrease of ε. Doing so, the $n_e N$ rank constraints in ($\mathcal{RSDP}$) are gradually approached, such that when $\varepsilon^{\langle j \rangle} = 0$, a feasible solution to ($\mathcal{RSDP}$) is obtained. The outline of the SSDR algorithm is presented in Algorithm 1. The choice of the weighting parameter $\alpha^{\langle j \rangle}$ affects the solution to ($\mathcal{SSDR}$). Although there are no strict rules for choosing $\alpha^{\langle j \rangle}$, it is proposed to make this parameter dependent on the error $e_f^{\langle j \rangle} = |f^{\langle j \rangle} - \hat{f}^{\langle j \rangle}|$, where $f^{\langle j \rangle} \triangleq \mathcal{J}(W^{\langle j \rangle})$, see (5), and $\hat{f}^{\langle j \rangle} \triangleq \sum_{k=1}^{N} \gamma(k)\text{Tr}\left(Z^{\langle j \rangle}(k)\right)$. This error expresses the mismatch between the true cost based on the excitation vectors $W^{[e]\langle j \rangle}(k)$ and the cost based on the 'approximate' rank one matrices $\Phi_w^{[e]\langle j \rangle}(k)$ used in ($\mathcal{SSDR}$). Notice that $e_f^{\langle j \rangle} = 0$ if $\text{rank}(\Phi_w^{[e]\langle j \rangle}(k)) = 1, \forall e, k$. The weighting parameter $\alpha^{\langle j \rangle}$ is then selected as

$$\alpha^{\langle j \rangle} = \max\left(\min(e_f^{\langle j \rangle}, e_{f_{\max}}), e_{f_{\min}}\right)\alpha_0^j \tag{10}$$

where and $e_{f_{\max}} > e_{f_{\min}} > 0$ are user-defined upper and lower bounds. Hence, this update law puts more weight on minimizing the rank when the cost $\hat{f}^{\langle j \rangle}$ is an inaccurate representation of the true cost $f^{\langle j \rangle}$, and vice versa. Additionally, setting $\alpha_0 > 0$ enforces the $n_u - 1$ smallest eigenvalues to zero.

Discussion The SSDR approach is related to existing iterative approaches to solving rank constrained problems. Though, SSDR is tailored for the potentially large dimensional separable problem ($\mathcal{RSDP}$) and has the following key features that are beneficial in terms of computation load: a) the sequence of programs to be solved consists only of feasible programs, in contrast to Cao et al. (2017), b) the required projection matrices are constructed analytically, instead of optimization-based like in Dattorro (2010); Delgado et al. (2014), and c) the rank constraints are approximated by a single linear constraint, instead of a (set of) conic constraint(s) like in Sun and Dai (2017).

4. EXPERIMENTAL VALIDATION

This section constitutes contribution C2.

4.1 Experimental goal

The goal of the experimental validation is to demonstrate 1) the performance of the SSDR algorithm and 2) the benefit of multivariable excitation design compared to optimal Single Input Multiple Output (SIMO) excitation design (Dirkx et al., 2019), in terms of achievable FRF accuracy. While in multivariable design both the excitation gains and directions are optimized, the SIMO design inherently allows for optimization of the gains only.

4.2 Experimental system and procedure

System The experiments consist in the identification of the $[4 \times 4]$ out-of-plane dynamics of the closed-loop controlled wafer stage depicted in Fig. 2. The output powers are constrained to a maximum value of 0.30mW.

Procedure First, a set of 4 preliminary (non-optimized) SIMO experiments are performed to acquire estimates of H_ξ, S, and γ, that serve as inputs for optimized design. Subsequently, for comparison purposes, 4 experiments are performed using optimized SIMO excitations, applying the SIMO OED framework described in Dirkx et al. (2019). Lastly, a set of 4 experiments is performed using optimized multivariable excitations applying the SSDR approach from Section 3.

4.3 Preliminary design and results

For the preliminary design, SIMO excitations are used with a signal power of 0.25 Watt for each input, uniformly distributed over the frequency grid $f_e = [10:5:4995]$Hz. The resulting identified FRF is shown in black in Fig. 3. Entry $\hat{G}_{1,1}$ is shown in greater detail in black in Fig 4.

4.4 Multivariable OED design and results

The multivariable excitations are generated by the SSDR algorithm. The settings used are $j_{\max} = 50, \alpha_0 = 1.05, e_{f_{\min}} = 10^{-12}, e_{f_{\max}} = 10^{-4}, \varepsilon_{\text{tol}} = 10^{-4}, \Delta f_{\text{tol}} = 10^{-8}$.

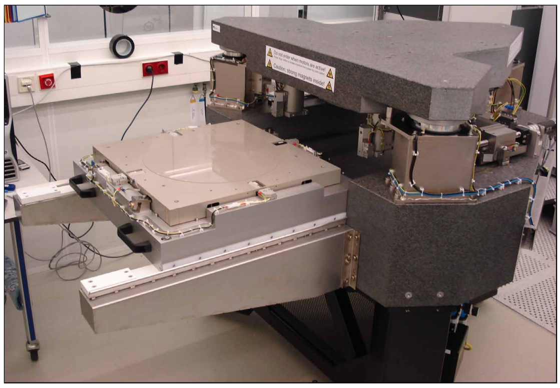

Fig. 2. Wafer stage setup.

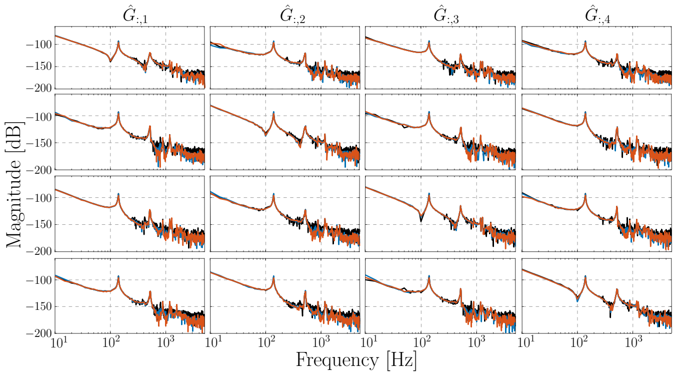

Fig. 3. Full 4×4 Bode magnitude plot of the FRF obtained from the preliminary (black), optimized SIMO (blue), and optimized multivariable (red) measurements.

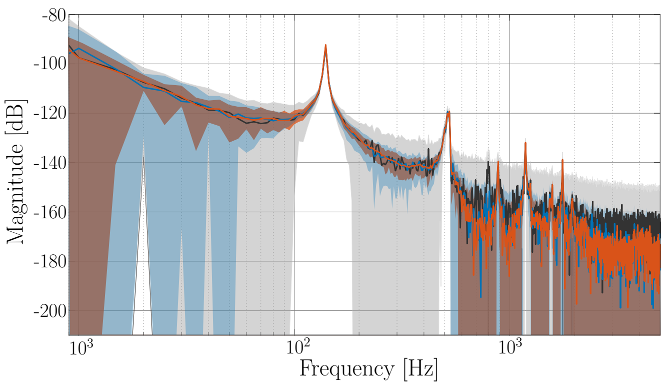

Fig. 4. Bode magnitude plot of FRF entry $\hat{G}_{1,1}$ obtained from the preliminary (black), optimized SIMO (blue), and optimized multivariable (red) measurements, including their respective 95% confidence intervals (shades).

SSDR convergence The progression of the true and approximate objective value $f^{\langle j \rangle}$ and $\hat{f}^{\langle j \rangle}$ are shown in the top plot of Fig. 5, together with the lower bound $f^{\dagger}_{\mathcal{SDR}}$. All curves are equally scaled such that $f^{\dagger}_{\mathcal{SDR}} = 1$. During the first few iterations, $\hat{f}^{\langle j \rangle}$ is an inaccurate approximation of the true objective value $f^{\langle j \rangle}$, but the mismatch reduces over the iterations. In 50 iterations, the objective value $f^{\langle j \rangle}$ has converged towards a solution with a gap of a factor 1.5 from the lower bound $f^{\dagger}_{\mathcal{SDR}}$. The progression of the slack variable $\varepsilon^{\langle j \rangle}$ is depicted in the bottom plot of Fig. 5, which shows a monotonic decrease. Although

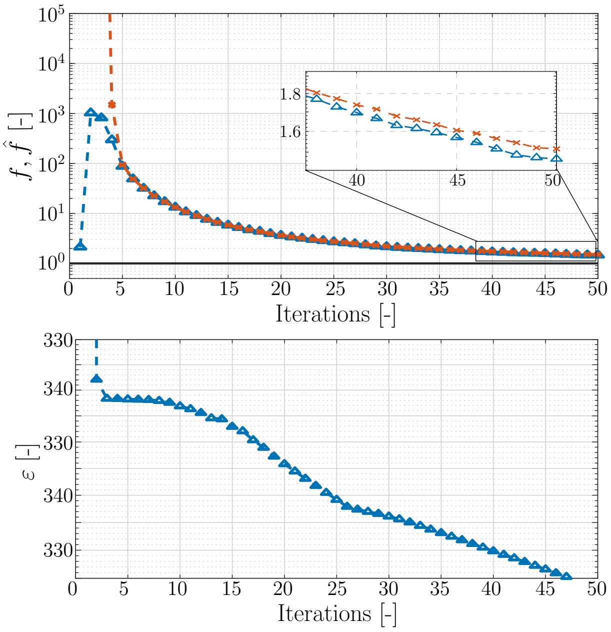

Fig. 5. Convergence of SSDR algorithm. *Top*: True objective value $f^{\langle j \rangle}$ (×–), approximate objective value $\hat{f}^{\langle j \rangle}$ (△–), and lower bound $f^{\dagger}_{SDR}$ (–). *Bottom*: Monotonic decrease of variable $\varepsilon^{\langle j \rangle}$.

a rank one solution is not achieved within 50 iterations, i.e., $\varepsilon^{\langle 50 \rangle} \neq 0$, it is approximated rather accurately, which is supported (implicitly) by the small difference between $f^{\langle 50 \rangle}$ and $\hat{f}^{\langle 50 \rangle}$ of 0.1 shown in the top plot.

FRF identification results The identified FRFs using optimized SIMO and multivariable excitations are shown in blue and red, respectively, in Fig. 3. The $\hat{G}_{1,1}$ entry is shown in more detail in Fig 4. The optimized SIMO excitations improve upon the accuracy of the preliminary FRF by approximately a factor 3. Applying multivariable excitations, yet an additional factor 2 improvement is achieved. This showcases the benefit of optimized excitation directions on top of optimized excitation gains.

5. CONCLUSIONS

An optimization-based multivariable OED framework for the FRF identification of MIMO systems is presented. Herein, optimization of the frequency-wise directions of the system to-be-identified, within given element-wise power constraints, is addressed explicitly. An algorithm that approximately solves this non-convex and generally NP-hard optimization problem is presented. An improvement in the achieved FRF quality over existing excitation design techniques is shown experimentally.

REFERENCES

Barenthin, M., Bombois, X., Hjalmarsson, H., and Scorletti, G. (2008). Identification for control of multivariable systems: Controller validation and experiment design via lmis. *Automatica*, 44(12), 3070 – 3078.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press, New York, NY, USA.

Cao, P., Thompson, J., and Poor, H.V. (2017). A sequential constraint relaxation algorithm for rank-one constrained problems. In *2017 25th EUSIPCO*, 1060–1064.

Dattorro, J. (2010). *Convex optimization & Euclidean distance geometry*. Lulu. com.

Delgado, R.A., Agüero, J.C., and Goodwin, G.C. (2014). A rank-constrained optimization approach: Application to factor analysis. *IFAC Proc. Vol.*, 47(3), 10373 – 10378.

Dirkx, N., Wijdeven, v.d.J., and Oomen, T. (2019). Optimal experiment design for multivariable motion systems: with application to a next-generation wafer stage. In *In Proc. of 8th IFAC Symposium on Mechatronic Systems*.

Dobrowiecki, T.P., Schoukens, J., and Guillaume, P. (2006). Optimized excitation signals for mimo frequency response function measurements. *IEEE Trans. Instrum. Meas.*, 55(6), 2072–2079.

Fan, K. (1950). On a theorem of weyl concerning eigenvalues of linear transformations. *Proc. Natl. Acad. Sci.*, 36(1), 31–35.

Gevers, M., Bombois, X., Hildebrand, R., and Solari, G. (2011). Optimal Experiment Design for Open and Closed-loop System Identification. *CIS*, 11(3), 197–224.

Goodwin, G.C. and Payne, R.L. (1977). *Dynamic system identification: experiment design and data analysis*. Academic Press NY.

Grant, M. and Boyd, S. (2014). Cvx: Matlab software for disciplined convex programming.

Guillaume, P., Pintelon, R., and Schoukens, J. (1996). Accurate estimation of multivariable frequency response functions. *IFAC Proc. Vol.*, 29(1), 4351 – 4356.

Hassibi, A., How, J., and Boyd, S. (1999). A path-following method for solving bmi problems in control. In *Proc. of the 1999 ACC*, volume 2, 1385–1389.

Hjalmarsson, H. (2005). From experiment design to closed-loop control. *Automatica*, 41(3), 393 – 438.

Jansson, H. (2004). *Experiment design with applications in identification for control.* PhD dissertation, Signaler, sensorer och system, Stockholm.

Karimi, A. and Kammer, C. (2016). A data-driven approach to robust control of multivariable systems by convex optimization. *Automatica*, 85.

Ljung, L. (ed.) (1999). *System Identification (2Nd Ed.): Theory for the User.* Prentice Hall PTR, Upper Saddle River, NJ, USA.

Manchester, I.R. (2009). An algorithm for amplitude-constrained input design for system identification. In *Proc. of the 48h IEEE CDC*, 1551–1556.

Markovsky, I. (2008). Structured low-rank approximation and its applications. *Automatica*, 44(4), 891–909.

Mehra, R.K. (1974). Frequency-domain synthesis of optimal inputs for multiinput-multioutput (mimo) systems with process noise. In *1974 IEEE CDC*, 410–418.

Nemirovski, A., Roos, C., and Terlaky, T. (1999). On maximization of quadratic form over intersection of ellipsoids with common center. *Mathematical Programming*, 86(3), 463–473.

Nesterov, Y. and Nemirovski, A. (1994). *Interior-Point Polynomial Algorithms in Convex Programming.* SIAM.

Nesterov, Y. (1998). Semidefinite relaxation and nonconvex quadratic optimization. *Optim. Methods Softw.*, 9(1-3), 141–160.

Oomen, T. (2018). Advanced motion control for precision mechatronics: Control, identification, and learning of complex systems. *IEEJ Journal of Industry Applications*, 7, 127–140.

Pintelon, R. and Schoukens, J. (2012). *System identification: a frequency domain approach (2nd edition).* Hoboken: John Wiley & Sons, Inc.

Recht, B., Fazel, M., and A. Parrilo, P. (2007). Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review*, 52.

Sun, C. and Dai, R. (2017). Rank-constrained optimization and its applications. *Automatica*, 82, 128 – 136.

van der Maas, R., van der Maas, A., Voorhoeve, R., and Oomen, T. (2016). Accurate frf identification of lpv systems: nd-lpm with application to a medical x-ray system. *IEEE Trans.on Contr. Sys. Tech.*, PP, 1–12.