A Möbius Transform Approach for Improved MIMO Frequency Domain Identification

Citation for the original published paper (version of record):
McKelvey, T., Osipov, A., Rylander, T. (2021). A Möbius Transform Approach for Improved MIMO Frequency Domain Identification. IFAC-PapersOnLine, 54(7): 791-796.
http://dx.doi.org/10.1016/j.ifacol.2021.08.458

Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 54-7 (2021) 791–796

A Möbius Transform Approach for Improved MIMO Frequency Domain Identification ⋆

T. McKelvey * A. Ludvig-Osipov * T. Rylander *

* *Chalmers University of Technology, 412 96 Göteborg, Sweden (e-mail: {tomas.mckelvey,osipov,rylander}@chalmers.se).*

Abstract: To improve the numerical conditioning of a frequency domain subspace algorithm (FSID) when identifying continuous time transfer functions a general Möbius transformation can be employed. The Möbius transformation is a more general class of transformations which include the bilinear transformation as a special case. For the state-space model class of MIMO transfer functions we derive the relations describing how the state-space realization of a transfer function is converted under a Möbius transformation. With an extensive numerical example we illustrate that a particular choice of the Möbius transformation leads to transfer function estimates with, in general, significantly improved accuracy.

Keywords: Subspace methods, Frequency domain identification, Continuous time system estimation, System identification, Möbius transformation

1. INTRODUCTION

Frequency domain identification refers to the estimation of models of dynamic systems where the measured data is in the form of Fourier transforms of the the input and output signals Ljung (1999); Pintelon and Schoukens (2001). The state-space model structure is a compact and convenient parametrization of multi input and multi output (MIMO) dynamic linear systems. The class of so called subspace methods are today commonly used for the estimation of MIMO systems both for time-domain Van Overschee and De Moor (1995) as well as for frequency-domain data McKelvey et al. (1996). For discrete time (DT) systems, where the frequency function argument is $e^{j\omega}$ the frequency domain subspace algorithm performs well, see eg. McKelvey et al. (1996). For the continuous time (CT) case where the argument is $j\Omega$ a direct (naive) implementation of the subspace methods run into numerical problems due to ill-conditioned steps in the algorithm McKelvey et al. (1996); Van Overschee and De Moor (1996). Approaches which, to some degree, circumvent this issue, are reported in Van Overschee and De Moor (1996) where a re-parametrization using Forsythe polynomials is used. In McKelvey et al. (1996) the bilinear transformation is employed where the CT problem is transformed to an equivalent DT estimation problem which is solved by the standard DT algorithm and the CT estimate is finally obtained by the inverse bilinear transformation of the DT estimate. In Vakilzadeh et al. (2015) it was illustrated how the numerical conditioning for the method employing the bilinear transformation can be improved by a suitable experiment design. In spite of these methods the numerical conditioning in these CT identification algorithms can still be an issue when the CT data samples have certain structure. Particularly we have noticed this if we have data from a frequency band. In this contribution we propose to employ the more general Möbius transform to move an ill-conditioned rational matrix function approximation problem to a more well conditioned solution domain. For the CT system identification problem we show that a particular Möbius transformation in many cases vastly improves the result as compared to employing the bilinear transformation.

In Section 2 we describe the identification problem and give a short review of the subspace based algorithm considered. In Section 3 we present results relating the Möbius transformation and the state-space models which form basis for the proposed method as well as a proposed selection of the parameters of the transformation. In Section 5 we illustrate the performance of the method using a Monte-Carlo simulation. The paper is summarized in the last section.

2. THE GENERAL ESTIMATION PROBLEM

In this paper we consider the estimation of matrix-valued rational functions given numerical samples of the rational function. The frequency-domain-system-identification problem can be seen as a particular instance to this more general class of problems.

We assume we are faced with data in a set of samples in the form

$$\mathcal{D}_{\mathrm{s}} = \{(s_k, Y_k, U_k)\}_{k=1}^{M} \tag{1}$$

where $s_k \in \mathbb{C}$, $Y_k \in \mathbb{C}^p$, $U_k \in \mathbb{C}^m$ and we seek a rational function matrix $G_{\mathrm{s}}(s) \in \mathbb{C}^{p\times m}$ such that

$$\sum_{k=1}^{M} \|Y_k - G_{\mathrm{s}}(s_k)U_k\|_2^2 \tag{2}$$

is minimized. Here we focus on an algorithm which generates a state-space *realization* $(A_{\mathrm{s}}, B_{\mathrm{s}}, C_{\mathrm{s}}, D_{\mathrm{s}})$ of order n and the estimated rational matrix function is given by

⋆ This research has been carried out in the ChaseOn Centre in a project financed by Vinnova, Chalmers, Food Radar Systems, Keysight Technologies, Medfield Diagnostics, Saab and UniqueSec

10.1016/j.ifacol.2021.08.458

$$G_s(s) = D_s + C_s(sI - A_s)^{-1}B_s. \tag{3}$$

where I is the identity matrix of appropriate size, the matrix A_s is of size $n \times n$ and the size of the other matrices follows from the size of the rational matrix function, i.e. the input and output dimensions. An alternative formulation is obtained if samples of the transfer matrix, i.e. $G_k = G_s(s_k)$ constitute the given data. However such a problem formulation can be recast to the original problem (1)-(2) by associating each column from the transfer function samples G_k as a sample of the output vector Y_k and where the input U_k is selected as the corresponding column of an identity matrix of size m.

If the numerical conditioning when directly solving the problem in (2) is unfavorable, we propose to employ a Möbius transformation method where we introduce a "transformed" problem by forming the alternative data set

$$\mathcal{D}_z = \{(z_k, Y_k, U_k)\}_{k=1}^{M} \tag{4}$$

where $z_k = f^{-1}(s_k)$ and $f(\cdot)$ is a Möbius transformation with parameters selected such that the algorithm employed to minimize

$$\sum_{k=1}^{M} \|Y_k - G_z(z_k)U_k\|_2^2 \tag{5}$$

has an improved numerical conditioning. The rational function solution to the original problem in (2) is then obtained by the Möbius transformation applied to the state-space realization (A_z, B_z, C_z, D_z) obtained when minimizing (5). In Section 3 we give the details of this transformation step.

To use a transformation to move the problem to a more well-conditioned space as outlined above has for CT frequency domain estimation been suggested by McKelvey et al. (1996). They used a bilinear transformation which is a special case of the general Möbius transformation we consider here. The bilinear transformation is often used in control theory to convert theoretical results between DT and CT domains Hitz and Anderson (1969); Doyle et al. (1991) or in signal processing for DT filter design based on CT filter structures Oppenheim and Schafer (1989); Steiglitz (1965).

In this paper we will approach the minimization of (5) by employing the well-known frequency-domain subspace algorithm (FSID) described in McKelvey et al. (1996) which is known to be well-conditioned for DT frequency domain problems where $z_k = e^{j\omega_k}$.

Below in Section 2.1 we summarize the key steps in the (FSID) algorithm, before we proceed with the main contributions of this paper in Section 3.

2.1 The FSID algorithm

The FSID algorithm can be summarized as a multi step procedure:

(i) From the structured data matrices (see below) a low-dimensional range space is approximated which form an estimate of the observability matrix for a state-space realization.
(ii) From the estimated observability matrix the A_z and C_z matrices in the state-space tuple are determined.
(iii) With A_z and C_z fixed, the rational matrix function G_z is linear in D_z and B_z. Expression (5) is minimized w.r.t. (D_z, B_z) with fixed (A_z, C_z). This minimization is a linear least-squares problem.
(iv) With A_z and B_z fixed, the rational matrix function G_z is linear in D_z and C_z. Expression (5) is minimized w.r.t. (D_z, C_z) with fixed (A_z, B_z). This minimization is a linear least-squares problem.

As demonstrated in Gumussoy et al. (2018) iterating between step (iii) and step (iv) 3-4 times can reduce the approximation error significantly for true MIMO systems.

The details of the step (i) are as follows. Let integer n denote the desired model order. From the data set $\mathcal{D}_z$ we define the block matrix with $q > n$ block rows

$$\mathbf{U}_q = \begin{bmatrix} U_1 & U_2 & \cdots & U_M \\ z_1U_1 & z_2U_2 & \cdots & z_MU_M \\ z_1^2U_1 & z_2^2U_2 & \cdots & z_M^2U_M \\ \vdots & \vdots & \vdots & \vdots \\ z_1^{q-1}U_1 & z_2^{q-1}U_2 & \cdots & z_M^{q-1}U_M \end{bmatrix} \tag{6}$$

which we assume has full rank mq. We define the projection matrix

$$\mathbf{P}_q = I_M - \mathbf{U}_q^*(\mathbf{U}_q\mathbf{U}_q^*)^{-1}\mathbf{U}_q \tag{7}$$

which projects onto the nullspace of $\mathbf{U}_q$ so $\mathbf{U}_q\mathbf{P}_q = 0$. Here, I_M is an identity matrix of size M. From the samples Y_k in the data set $\mathcal{D}_z$ we define the block matrix

$$\mathbf{Y}_q = \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_M \\ z_1Y_1 & z_2Y_2 & \cdots & z_MY_M \\ z_1^2Y_1 & z_2^2Y_2 & \cdots & z_M^2Y_M \\ \vdots & \vdots & \vdots & \vdots \\ z_1^{q-1}Y_1 & z_2^{q-1}Y_2 & \cdots & z_M^{q-1}Y_M \end{bmatrix}. \tag{8}$$

If the data in the set $\mathcal{D}_z$ are related as

$$Y_k = (D_z + C_z(z_kI - A_z)^{-1}B_z)U_k \tag{9}$$

and (A_z, B_z, C_z, D_z) is a minimal realization of order n, $q > n$ and $M \geq n + mq$, then

$$\text{range}(\mathbf{Y}_q\mathbf{P}_q) = \text{range}(\mathbf{O}_q(A_z, C_z)). \tag{10}$$

Here,

$$\mathbf{O}_q(A_z, C_z) \triangleq \begin{bmatrix} C_z \\ C_zA_z \\ \vdots \\ C_zA_z^{q-1} \end{bmatrix} \tag{11}$$

is the (extended) observability matrix and range(X) denote the range space of matrix X. See McKelvey et al. (1996) for the details.

A matrix representing the range space is determined by solving the structured matrix approximation problem

$$\hat{\mathbf{Z}}, \hat{\mathbf{R}} = \arg \min_{\mathbf{Z}\in\mathbb{C}^{pq\times n}, \mathbf{R}\in\mathbb{C}^{n\times M}} \|\mathbf{Z}\mathbf{R} - \mathbf{Y}_q\mathbf{P}_q\|_F^2, \tag{12}$$

where $\|\cdot\|_F$ is the Frobenius norm. The solution to this problem can be calculated by the singular value decomposition of the matrix product $\mathbf{Y}_q\mathbf{P}_q$. The matrix $\hat{\mathbf{Z}}\hat{\mathbf{R}}$ is the best rank n Frobenius norm matrix approximation to the matrix $\mathbf{Y}_q\mathbf{P}_q$. If, again, the data is related as in (9) we have range $\hat{\mathbf{Z}}$ = range $\mathbf{O}_q(A_z, C_z)$ which imply that there exists a non-singular matrix L such that $\hat{\mathbf{Z}}L = \mathbf{O}_q(A_z, C_z)$ and hence $\hat{\mathbf{Z}} = \mathbf{O}_q(\hat{A}_z, \hat{C}_z)$ where $C_z = \hat{C}_zL$ and $L_z = L^{-1}\hat{A}_zL$. This yields the estimates

$$\hat{C}_{\rm z} = [I_p \ 0]\, \hat{\mathbf{Z}}$$
$$\hat{A}_{\rm z} = \arg\min_X \| \left[I_{(q-1)p} \ 0\right] \hat{\mathbf{Z}} X - \left[0 \ I_{(q-1)p}\right] \hat{\mathbf{Z}} \|_F^2 \tag{13}$$

where the last problem is a standard linear least-squares problem.

Step (iii) in the algorithm can be formulated as

$$\hat{D}_{\rm z}, \hat{B}_{\rm z} = \arg\min_{D,B} \sum_{k=1}^{M} \|Y_k - (D + \hat{C}_{\rm z}(z_k I - \hat{A}_{\rm z})^{-1} B) U_k\|_2^2 \tag{14}$$

and step (iv) as

$$\hat{D}_{\rm z}, \hat{C}_{\rm z} = \arg\min_{D,C} \sum_{k=1}^{M} \|Y_k - (D + C(z_k I - \hat{A}_{\rm z})^{-1} \hat{B}_{\rm z}) U_k\|_2^2. \tag{15}$$

The FSID algorithm is available as an open source implementation for three software platforms; Python, Julia and Matlab, see McKelvey (2019).

2.2 Numerical conditioning

The FSID algorithm provides a straightforward approach to numerically determine a state-space representation of any arbitrary matrix-valued proper rational function. However, some of the steps involve data processing which can result in numerical ill-conditioning of the performed operations. The core of the algorithm is the construction of the matrices $\mathbf{U}_q$ in (6) and $\mathbf{Y}_q$ in (8). It is clear that if the quotient

$$\frac{\max_k |z_k|^{q-1}}{\min_k |z_k|^{q-1}} \tag{16}$$

is very large it will be numerically challenging to accurately recover the desired subspace in (12). For the CT-system-identification case the argument to the rational matrix function is $j\Omega_k$ and in this case the quotient

$$\frac{\max_k |\Omega_k|^{q-1}}{\min_k |\Omega_k|^{q-1}} \tag{17}$$

is large if the data set contains relatively wide frequency range. The issue becomes more prominent for estimation of models of high orders since q must be selected larger then the model order n. On the other hand if the argument is $z_k = e^{j\omega_k}$, the magnitude is one and hence results in a problem with better conditioned $\mathbf{U}_q$ and $\mathbf{Y}_q$ matrices. Note that $\mathbf{U}_q = (W_q \otimes I_m)\,\mathrm{diag}(U_1, \ldots, U_M)$ where W_q is a Vandermonde matrix defined by the scalars $\{z_k\}_{k=1}^M$. Extensive experimentation has shown that the condition number of the matrix W_q can be used as a proxy to a priori determine if the set of complex numbers $\{z_k\}_{k=1}^M$ will result in a well-conditioned estimation problem.

In the next section we discuss the more general Möbius transformation as a more flexible tool to improve the numerical conditioning of the FSID algorithm when faced with problems which are ill-conditioned in their original formulation.

3. MÖBIUS TRANSFORMATIONS

Consider the scalar rational function $f(z) = \frac{\alpha z + \beta}{\gamma z + \delta}$ where z is a complex variable and $\alpha, \delta, \gamma, \beta$ are complex constants. We will throughout the presentation use the following assumption.

Assumption 1. The constants $\alpha, \beta, \gamma, \delta$ satisfy $\gamma\beta - \alpha\delta \neq 0$.

A function $f(z)$ satisfying Assumption 1 is a Möbius transformation, also called a Linear Fractional Transformation, $s = f(z) = \frac{\alpha z + \beta}{\gamma z + \delta}$ Young (1984); Seppälä and Sorvali (1992). The condition $\gamma\beta - \delta\alpha \neq 0$ ensures that the function is a transformation and hence have an inverse. The inverse function is given by $z = f^{-1}(s) = \frac{\beta - \delta s}{\gamma s - \alpha}$. The following result gives the details when applying the Möbius transformation to the independent variable in a rational matrix function

Theorem 1.

(i) Let $G_{\rm s}(s) = D_{\rm s} + C_{\rm s}(sI - A_{\rm s})^{-1} B_{\rm s}$ be a rational matrix function and assume $P \triangleq (\alpha I - \gamma A_{\rm s})$ is non-singular. Then

$$G_{\rm z}(z) \triangleq G_{\rm s}(\frac{\alpha z + \beta}{\gamma z + \delta}) = D_{\rm z} + C_{\rm z}(zI - A_{\rm z})^{-1} B_{\rm z} \tag{18}$$

where the state-space matrices are given by

$$\begin{bmatrix} A_{\rm z} & B_{\rm z} \\ C_{\rm z} & D_{\rm z} \end{bmatrix} = \begin{bmatrix} (\delta A_{\rm s} - \beta I) P^{-1} & (\alpha\delta - \gamma\beta) P^{-1} B_{\rm s} \\ C_{\rm s} P^{-1} & D_{\rm s} + \gamma C_{\rm s} P^{-1} B_{\rm s} \end{bmatrix} \tag{19}$$

(ii) Let $G_{\rm z}(z) = D_{\rm z} + C_{\rm z}(zI - A_{\rm z})^{-1} B_{\rm z}$ be a rational matrix function and assume $R \triangleq (\delta I + \gamma A_{\rm z})$ is non-singular. Then $G_{\rm s}(s) \triangleq G_{\rm z}(\frac{\beta - \delta s}{\gamma s - \alpha}) = D_{\rm s} + C_{\rm s}(sI - A_{\rm s})^{-1} B_{\rm s}$ where

$$\begin{bmatrix} A_{\rm s} & B_{\rm s} \\ C_{\rm s} & D_{\rm s} \end{bmatrix} = \begin{bmatrix} (\alpha A_{\rm z} + \beta I) R^{-1} & (\gamma\beta - \delta\alpha) R^{-1} B_{\rm z} \\ -C_{\rm z} R^{-1} & D_{\rm z} - \gamma C_{\rm z} R^{-1} B_{\rm z} \end{bmatrix} \tag{20}$$

A proof of the result is given in Appendix A. From the result we note that the order of a realization of a rational transfer function is invariant under a Möbius transformation. The following result clarify that the two non-singular assumptions above are linked.

Theorem 2. Given Assumption 1 then

(i) $(\alpha I - \gamma A_{\rm s})$ is non-singular if and only if $(\delta I + \gamma A_{\rm z})$ is non-singular.
(ii) Let $s_k = \frac{\alpha z_k + \beta}{\gamma z_k + \delta}$ then $\gamma z_k + \delta \neq 0$ if and only if $\gamma s_k - \alpha \neq 0$

Proof: If λ is an eigenvalue to the matrix $A_{\rm s}$ with corresponding eigenvector x, then x is also an eigenvector to $A_{\rm z}$ with corresponding eigenvalue $\lambda_{\rm z} = \frac{\delta\lambda - \beta}{\alpha - \gamma\lambda}$ since

$$\begin{aligned} A_{\rm z}(\alpha I - \gamma A_{\rm s}) x &= (\delta A_{\rm s} - \beta I) x \\ (\alpha - \gamma\lambda) A_{\rm z} x &= (\delta\lambda - \beta) x \\ A_{\rm z} x &= \frac{\delta\lambda - \beta}{\alpha - \gamma\lambda} x = \lambda_{\rm z} x \end{aligned} \tag{21}$$

The matrix $(\alpha - \gamma A_{\rm s})$ is non-singular iff for all eigenvalues λ we have $\alpha - \gamma\lambda \neq 0$. The matrix $\delta I + \gamma A_{\rm z}$ is non singular iff, for all eigenvalues $\lambda_{\rm z}$, we have $\delta + \gamma\lambda_{\rm z} \neq 0$. Thus

$$\begin{aligned} \delta + \gamma\lambda_{\rm z} &= \delta + \frac{\gamma(\delta\lambda - \beta)}{\alpha - \gamma\lambda} \\ &= \frac{\delta\alpha - \delta\gamma\lambda + \delta\gamma\lambda - \gamma\beta}{\alpha - \gamma\lambda} = \frac{\delta\alpha - \gamma\beta}{\alpha - \gamma\lambda} \end{aligned} \tag{22}$$

which is non-zero since, by Assumption 1, $\delta\alpha - \gamma\beta \neq 0$. The converse is proved by similar arguments. ■

If the parameters are selected such that Assumption 1 holds and $(\alpha I - \gamma A_\mathrm{s})$ is non-singular then the mappings in (19) and (20) are hence each others inverses, i.e. they form a transformation.

As outlined in Section 2 the properties above can be used to improve the numerical conditioning in algorithmic procedures involved for example when estimating a realization from numerical data. We note that if for a given value z_k the rational function evaluates to a matrix $G_\mathrm{z}(z_k) = D_\mathrm{z} + C_\mathrm{z}(z_k I - A_\mathrm{z})^{-1}B_\mathrm{z}$ and for a given U_k we have $Y_k \triangleq G(z_k)U_k$. Define $s_k = \frac{\alpha z_k+\beta}{\gamma z_k+\delta}$. Then it follows from above that the rational function $G_\mathrm{s}(s_k) = D_\mathrm{s} + C_\mathrm{s}(s_k I - A_\mathrm{s})^{-1}B_\mathrm{s} = G_\mathrm{z}(z_k)$ and hence $Y_k = G_\mathrm{s}(s_k)U_k$, i.e. $G_\mathrm{s}(s)$ interpolates that data. By employing the Möbius transformation on the complex argument we can associate the same sequence of vectors Y_k, U_k with different rational matrix functions which all preserve the order of the realization. The following corollary forms the foundation on how to use the transformations together with data.

Corollary 1. The following statements hold

(i) Assume the data set $\mathcal{D}_\mathrm{s} = \{(s_k, Y_k, U_k)\}_{k=1}^M$ is interpolated by the realization $(A_\mathrm{s}, B_\mathrm{s}, C_\mathrm{s}, D_\mathrm{s})$, $P = (\alpha I - \gamma A_\mathrm{s})$ is non-singular and $\forall k$, $\gamma s_k - \alpha \neq 0$. Then the realization $(A_\mathrm{z}, B_\mathrm{z}, C_\mathrm{z}, D_\mathrm{z})$ defined by (19) interpolates the data set $\mathcal{D}_\mathrm{z} = \{(z_k, Y_k, U_k)\}_{k=1}^M$ where $z_k \triangleq \frac{\beta-\delta s_k}{\gamma s_k-\alpha}$.

(ii) Assume the data set $\mathcal{D}_\mathrm{z} = \{(z_k, Y_k, U_k)\}_{k=1}^M$ is interpolated by the realization $(A_\mathrm{z}, B_\mathrm{z}, C_\mathrm{z}, D_\mathrm{z})$, $R = (\delta I + \gamma A_\mathrm{z})$ is non-singular and $\forall k$, $\gamma z_k + \delta \neq 0$. Then the realization $(A_\mathrm{s}, B_\mathrm{s}, C_\mathrm{s}, D_\mathrm{s})$ defined by (20) interpolates the data set $\mathcal{D}_\mathrm{s} = \{(s_k, Y_k, U_k)\}_{k=1}^M$ where $s_k \triangleq \frac{\alpha z_k+\beta}{\gamma z_k+\delta}$.

The corollary gives a constructive alternative way to generate a realization which approximates a given data set $\mathcal{D}_\mathrm{s} = \{(s_k, Y_k, U_k)\}_{k=1}^M$.

Algorithm 1.

(i) Generate a new data set $\mathcal{D}_\mathrm{z}$ by transforming the values of the samples of the free variable using a suitable Möbius transformation ($z_k \triangleq \frac{\beta-\delta s_k}{\gamma s_k-\alpha}$) but preserve the samples Y_k, U_k.

(ii) Approximate a realization $(A_\mathrm{z}, B_\mathrm{z}, C_\mathrm{z}, D_\mathrm{z})$ to the modified data set $\mathcal{D}_\mathrm{z} = \{(z_k, Y_k, U_k)\}_{k=1}^M$ using the FSID algorithm.

(iii) Use (20) to generate a realization $(A_\mathrm{s}, B_\mathrm{s}, C_\mathrm{s}, D_\mathrm{s})$.

Since for any k in the data set we have

$$D_\mathrm{z} + C_\mathrm{z}(z_k I - A_\mathrm{z})^{-1}B_\mathrm{z} = D_\mathrm{s} + C_\mathrm{s}(s_k I - A_\mathrm{s})^{-1}B_\mathrm{s} \quad (23)$$

the approximation error in the z-domain, where the approximation was performed, is preserved by the transformation to frequency function in the the s-domain.

4. DOMAIN TRANSFORMATION IN FREQUENCY-DOMAIN SYSTEM IDENTIFICATION

In the following sections, we demonstrate how the domain-transformation technique of Section 3 can be used to improve the performance of the FSID algorithm for the frequency-domain system identification. In the application of the frequency-domain system identification, the samples are taken on an imaginary axis, $\{s_k = j\Omega_k\}_{k=1}^M$, where $\Omega_k \in \mathbb{R}$ are referred as frequencies (in CT). Without loss of generality we assume that the sequence of frequencies $\{\Omega_k\}_{k=1}^M$ is ordered, and hence the frequencies are contained in a frequency interval $\Omega_k \in [\Omega_1, \Omega_M]$. A transformation can be used to move this CT frequency interval, located on an imaginary axis in the s-domain, to a DT frequency interval, located on a unit circle in the z-domain in a way that is numerically favorable for FSID algorithm. The bilinear transformation is conventionally used for this role, as proposed in McKelvey (1996); however, here we show that a usage of a Möbius transformation further improves the system-identification performance.

4.1 Bilinear Transformation

The bilinear transformation is obtained by selecting the Möbius parameters as

$$\alpha = 2, \quad \beta = -2, \quad \gamma = T, \quad \delta = T \quad (24)$$

where T is a positive real valued scaling. For two realizations, $G_\mathrm{s}(s)$ and $G_\mathrm{z}(z)$, related by the bilinear transform $z = f^{-1}(s)$ it holds that

$$G_\mathrm{z}(e^{j\omega}) = G_\mathrm{s}(j\Omega) \quad \text{for} \quad \Omega = 2\tan(\omega/2)/T. \quad (25)$$

The DT frequency function is hence identical to the CT frequency function if the continuous frequency scale is warped according to the right expression in (25). Our numerical investigations (see below) indicate that the system-identification using bilinear transformation significantly degrades in performance when Ω_1 is not close to 0.

4.2 A Möbius transformation method

The inverse Möbius transformation f^{-1} is proposed here as a map between CT and DT domains. We start with excluding from consideration the transformations with $\gamma = 0$, as these cannot map a straight line (subset of the imaginary axis) to an arc (subset of the unit circle). Without loss of generality we can set $\gamma = 1$ and hence the transformation is defined by 3 complex-valued parameters. These parameters can be defined from 3 unique complex-valued equations of the form $e^{j\omega_i} = f^{-1}(j\Omega_i), i \in \{\mathrm{a}, \mathrm{b}, \mathrm{c}\}$, where $\omega_i \in \mathbb{R}$ represent frequency points in DT domain. Note that these three equations can be rewritten as linear when the assumption $\gamma j\Omega - \alpha \neq 0$ of Corollary 1 holds. In choosing the pairs (ω_i, Ω_i) we follow a heuristic approach to minimize the condition number of the Vandermonde matrix W_q formed from the transformed samples $\{z_k = f^{-1}(s_k)\}_{k=1}^M$. We set $\Omega_\mathrm{a} = \Omega_1, \Omega_\mathrm{b} = \frac{\Omega_1+\Omega_M}{2}, \Omega_\mathrm{c} = \Omega_M$ and $\omega_\mathrm{a} = 0, \omega_\mathrm{b} = \frac{\omega_\mathrm{a}+\omega_\mathrm{c}}{2}$, where ω_c is picked such that the condition number associated with the transformed samples is minimized. Note that besides of the parameters of the transformation f^{-1}, the condition number of W_q is dependent on the number of block rows q in $\mathbf{U}_q$, the number of the frequency samples M and their distribution within the interval. It can be computed a priori, without knowledge of $\{Y_k, U_k\}$.

5. NUMERICAL EXAMPLES

In this section, we investigate how the mapping between CT and DT domain influences the numerical stability of the frequency-domain system identification by evaluating the performance of Algorithm 1 with the two mappings: the bilinear transformation and the Möbius transformation as described in Sections 4.1 and 4.2 respectively. We can point out that if the FSID algorithm is used *without* the Möbius or bilinear transformation, significantly larger modeling errors will be obtained.

For each Monte-Carlo iteration, a random state-space realization $(A_\mathrm{s}, B_\mathrm{s}, C_\mathrm{s}, D_\mathrm{s})$ with 10 states, 3 inputs and 4 outputs is generated. For a given set of frequencies $\{\Omega_k\}_{k=1}^M$, the generated realization produces a CT frequency response $\{G_k = G_\mathrm{s}(j\Omega_k)\}_{k=1}^M$ according to (3). A realization $\{n_k\}_{k=1}^M$ of white circular symmetric Complex Gaussian noise with the variance $2\sigma^2$ is added to the frequency response, which yields the noised data points $\{G'_k = G_k + n_k\}_{k=1}^M$. Algorithm 1 with both the Möbius and the bilinear transformations is then used to estimate a rational matrix function of the form (3) from the noised data $\{G'_k\}_{k=1}^M$. The estimates are denoted as G^m and G^b for when the Möbius and the bilinear transformations were used respectively. The normalized approximation errors are evaluated as

$$e_{\mathrm{m/b}} = \frac{\sum_{k=1}^M \|G_k - G^{\mathrm{m/b}}(j\Omega_k)\|_F^2}{\sum_{k=1}^M \|G_k\|_F^2}. \tag{26}$$

In all the examples, $M = 100$ samples are uniformly distributed on normalized angular frequency intervals $[\Omega_1, \Omega_M]$, $\frac{\Omega_M - \Omega_1}{2\pi} = 1$.

The random state-space realization is generated by modified version of `control.matlab.rss()` function from Python Control Systems Library (version 0.8.3 at `https://python-control.org/`), which emulates the `rss` command in MATLAB, used in Input-Output measurements benchmark generation, Chen et al. (2012). Our modifications include setting the true probability of a non-oscillating pole to 0.052 and changing the probability distribution of the imaginary part to normal with unit variance and mean $\Omega_0 = \frac{\Omega_1 + \Omega_M}{2}$. The rest of the state-space-model statistical properties are retained from `control.matlab.rss()`.

Fig. 1 depicts normalized approximation errors sorted in ascending order for individual realizations for the systems observed in the frequency interval $\frac{\Omega}{2\pi} \in [5, 6]$. For each value of noise variance, a pair of curves represents the performance of the FSID algorithm with the Möbius (solid) and the bilinear (dashed) transformations as maps between CT and DT domains. In each curve pair, both approaches are applied to the same set of data, the ascending sorting is however individual for each curve. The Möbius-transformation parameters are calculated according to the scheme in Section 4.2 using the knowledge of the frequency-samples set only. For each value of the noise variance, the bilinear-transformation parameter T giving the lowest median approximation error is picked, that is, the a posteriori (oracle) knowledge is used. The search is performed in the range $T \in [10^{-3}, 10^3]$. The FSID algorithm with $q = 20$ was used, and $\omega_\mathrm{c} = 1.835\pi$, minimizing the condition number for all frequency intervals, was chosen to calculate the Möbius transformation parameters. For each pair of curves, 1001 Monte-Carlo iterations were performed. All the curves behave similarly in logarithmic scale relative to their medians, and hence the median values can be used to describe the performance of the approaches under different standard deviation of noise and frequency intervals.

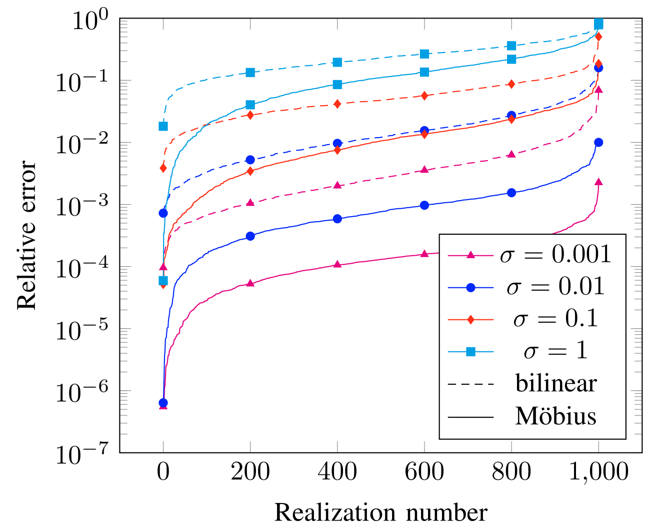

Fig. 1. The normalized approximation errors e_b and e_m sorted in ascending order for individual realizations. The samples are taken uniformly in a range of normalized frequencies $\frac{\Omega}{2\pi} \in [5, 6]$

Fig. 2 shows the median relative error as a function of the left endpoint $\frac{\Omega_1}{2\pi}$ in the frequency interval. The experiments are conducted as in the previous example for Fig. 1 and in fact each curve of Fig. 1 corresponds to a single point on a curve in Fig. 2. While at the low-frequency cases the a-posteriori-informed (oracle) bilinear transformation slightly outperforms the Möbius transformation, the median relative error for the bilinear transformation grows drastically and saturates when we shift to higher frequencies. At the same time the median relative error for Möbius transformation is mostly unaffected by the shift of the frequency interval. To interpret this, note that any shift $\Delta\Omega \in \mathbb{R}$ of the frequency interval can be compensated by the Möbius transformation via simple change of parameter $\beta \to \beta - j\alpha\Delta\Omega$, an operation unavailable for the bilinear transformation which parameters are real-valued. For frequency intervals with $\frac{\Omega_1}{2\pi} \geq 0.25$, the median relative error is more than an order of magnitude lower when the Möbius transformation used instead of the bilinear in the low-noise scenarios. The difference in the performance gets less pronounced with the increase of the noise level, but is still noticeable even at $\sigma = 1$. The smaller difference in the performance for the high-noise scenarios is related to the fact that the error is evaluated by comparison with the noise-free data samples, while the calculations on functions G^m and G^b is performed on the noised data.

To conclude, a Möbius transformation can be used to provide a numerically more favorable mapping between the CT and DT domains than the bilinear transformation.

Appendix A. PROOF OF THEOREM 1

Proof: We start from

$$G_\mathrm{z}(z) = G_\mathrm{s}\left(\frac{\alpha z + \beta}{\gamma z + \delta}\right) = D_\mathrm{s} + C_\mathrm{s}\left(\frac{\alpha z + \beta}{\gamma z + \delta} I - A_\mathrm{s}\right)^{-1} B_\mathrm{s}. \tag{A.1}$$

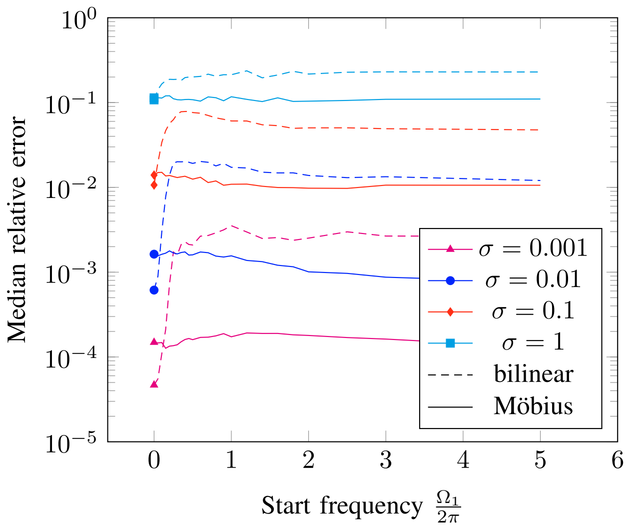

Fig. 2. The median values of the normalized approximation errors e_{b} and e_{m} as a function of the normalized start frequency.

First we focus on the matrix inverse in the expression above

$$
\begin{aligned}
&\left[\frac{\alpha z+\beta}{\gamma z+\delta} I - A_{\mathrm{s}}\right]^{-1} \\
&= \left[(\alpha z I + \beta I - (\gamma z+\delta)A_{\mathrm{s}})\right]^{-1}(\gamma z+\delta) \\
&= P^{-1}\left[(zP - (\delta A_{\mathrm{s}} - \beta I))P^{-1}\right]^{-1}(\gamma z+\delta) \\
&= P^{-1}\left[zI - (\delta A_{\mathrm{s}} - \beta I)P^{-1}\right]^{-1}(\gamma z I+\delta I) \\
&= P^{-1}\left[zI - (\delta A_{\mathrm{s}} - \beta I)P^{-1}\right]^{-1}\gamma\,[zI - \\
&\quad (\delta A_{\mathrm{s}} - \beta I)P^{-1} + (\delta A_{\mathrm{s}} - \beta I)P^{-1} + \delta/\gamma I] \\
&= P^{-1}\gamma + \\
&P^{-1}\left[zI - (\delta A_{\mathrm{s}} - \beta I)P^{-1}\right]^{-1}\gamma\left[(\delta A_{\mathrm{s}} - \beta I)P^{-1} + \delta/\gamma I\right] \\
&= P^{-1}\gamma + \\
&P^{-1}\left[zI - (\delta A_{\mathrm{s}} - \beta I)P^{-1}\right]^{-1}\left[\gamma(\delta A_{\mathrm{s}} - \beta I) + P\delta\right]P^{-1} \\
&= P^{-1}\gamma + P^{-1}\left[zI - (\delta A_{\mathrm{s}} - \beta I)P^{-1}\right]^{-1}(\alpha\delta - \gamma\beta)P^{-1}
\end{aligned}
\tag{A.2}
$$

Now the result follows by inserting (A.2) into (A.1) viz

$$
\begin{aligned}
G_{\mathrm{z}}(z) &= D_{\mathrm{s}} + C_{\mathrm{s}}\left(\frac{\alpha z+\beta}{\gamma z+\delta} I - A_{\mathrm{s}}\right)^{-1} B_{\mathrm{s}} \\
&= D_{\mathrm{s}} + \gamma C_{\mathrm{s}} P^{-1} B_{\mathrm{s}} + \\
&\quad C_{\mathrm{s}} P^{-1}\left[zI - (\delta A_{\mathrm{s}} - \beta I)P^{-1}\right]^{-1}(\alpha\delta - \gamma\beta)P^{-1} B_{\mathrm{s}}
\end{aligned}
\tag{A.3}
$$

The proof of the inverse transformation is analogous by considering

$$
G_{\mathrm{s}}(s) = G_{\mathrm{z}}\left(\frac{\beta-\delta s}{\gamma s-\alpha}\right) = D_{\mathrm{z}} + C_{\mathrm{z}}\left(\frac{\beta-\delta s}{\gamma s-\alpha} I - A_{\mathrm{z}}\right)^{-1} B_{\mathrm{z}}. \tag{A.4}
$$

■

REFERENCES

Chen, T., Ohlsson, H., and Ljung, L. (2012). On the estimation of transfer functions, regularizations and Gaussian processes—Revisited. *Automatica*, 48(8), 1525–1535.

Doyle, J., Packard, A., and Zhou, K. (1991). Review of LFTs, LMIs, and μ. *Proceedings of the IEEE Conference on Decision and Control*, 2, 1227–1232. doi:10.1109/cdc.1991.261572.

Gumussoy, S., Ozdemir, A.A., McKelvey, T., Ljung, L., Gibanica, M., and Singh, R. (2018). Improving Linear State-Space Models with Additional Iterations. *IFAC-PapersOnLine*, 51(15), 341–346.

Hitz, L. and Anderson, B. (1969). Discrete positive-real functions and their application to system stability. *Proceedings of the Institution of Electrical Engineers*, 116(1), 153. doi:10.1049/piee.1969.0031.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, New Jersey, second edition.

McKelvey, T. (1996). Periodic Excitation for Identification of Dynamic Errors-in-Variables systems Operating in Closed Loop. In *Proc. 13th World Congress of International Federation of Automatic Control*, volume J, 155–160. San Francisco, California.

McKelvey, T. (2019). fsid - Frequency Domain Subspace Based Identification - Software. Technical report. URL `https://github.com/tomasmckelvey/fsid`.

McKelvey, T., Akçay, H., and Ljung, L. (1996). Subspace-Based Multivariable System Identification from Frequency Response Data. *IEEE Trans. on Automatic Control*, 41(7), 960–979. doi:10.1109/9.508900.

Oppenheim, A.V. and Schafer, R.W. (1989). *Discrete-Time Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ.

Pintelon, R. and Schoukens, J. (2001). *System Identification - A frequency domain approach*. IEEE Press.

Seppälä, M. and Sorvali, T. (1992). Chapter 1 Geometry of Möbius transformations. In *Geometry of Riemann Surfaces and Teichmüller Spaces*, volume 169 of *North-Holland Mathematics Studies*, 11–57. North-Holland. doi:https://doi.org/10.1016/S0304-0208(08)72779-8.

Steiglitz, K. (1965). The equivalence of digital and analog signal processing. *Information and Control*, 8(5), 455–467. doi:10.1016/S0019-9958(65)90374-8.

Vakilzadeh, M.K., Yaghoubi, V., McKelvey, T., Abrahamsson, T., and Ljung, L. (2015). Experiment design for improved frequency domain subspace system identification of continuous-time systems. *Proc. IFAC Symposium on System Identification, IFAC-PapersOnLine*, 48(28), 886–891.

Van Overschee, P. and De Moor, B. (1995). Choice of State-space Basis in Combined Deterministic-Stochastic Subspace Identification. *Automatica*, 31(12), 1877–1883.

Van Overschee, P. and De Moor, B. (1996). Continuous-time frequency domain subspace system identification. *Signal Processing*, 52(2), 179–194. doi:10.1016/0165-1684(96)00052-7.

Young, N.J. (1984). Linear fractional transformations in rings and modules. *Linear Algebra and its Applications*, 56, 251–290. doi:https://doi.org/10.1016/0024-3795(84)90131-9.