Mahata, Kaushik, Schoukens, Johan, De Cock, Alexander; "Information matrix and D-optimal design with Gaussian inputs for Wiener model identification". Published in Automatica Vol. 69, p. 65-77 (2016)

Available from: http://dx.doi.org/10.1016/j.automatica.2016.02.026

Information matrix and D-optimal design with Gaussian inputs for Wiener model identification*

Kaushik Mahata[†1], Johan Schoukens[‡2], and Alexander De Cock[§3]

[1]Department of Electrical Engineering, University of Newcastle, Callaghan, NSW-2308, Australia.
[2]Department ELEC, Vrije Universiteit, Brussel, Building K, Pleinlaan 2, 1050, Brussels, Belgium.

October 31, 2016

Abstract

We present a closed form expression for the Fischer's information matrix associated with the identification of Wiener models. In the derivation we assume that the input signal is Gaussian. The analysis allows the linear sub-system in the Wiener model to have a generic rational transfer function of arbitrary order. It also allows the static nonlinearity of the Wiener model to be a polynomial of arbitrary degree. In addition, we show how this analysis can be used to design tractable algorithms for D-optimal input design. The idea is further extended to design optimal inputs consisting of a sequence of Gaussian signals with different mean values and variances. By combining Gaussian inputs with different means we can tune the amplitude distribution of the input to achieve the best identification accuracy in D-optimal sense. The analytical results are also illustrated with some numerical simulations.

Keywords Wiener model identification, Fischer's Information matrix, Cramér-Rao bound, D-optimal design, Nevanlinna-Pick Interpolation.

*The research is supported in part by the Australian Research Council under the contract DP130103909, by the Fund for Scientific Research (FWO-Vlaanderen), by the Flemish Government (Methusalem), the Belgian Government through the Inter university Poles of Attraction (IAP VII) Program, and by the ERC advanced grant SNLSID, under contract 320378.

[†]Corresponding author, Email:Kaushik.Mahata@newcastle.edu.au

[‡]Email: Johan.Schoukens@vub.ac.be

[§]Email: Alexander.De.Cock@vub.ac.be

1 Introduction

In this paper we study the design of optimal input signals using a mixture of Gaussian excitations with optimized input spectrum and dc-offsets. We first explain the main ideas in the introduction, next a detailed technical description is given in the rest of this paper.

1.1 Contributions

We analyze how the input signal used to identify a nonlinear Wiener system influences the accuracy of the estimated model. We assume the input signal is a Gaussian stochastic process with some known power spectral density. The analysis presented herein can accommodate a very general model structure. A Wiener model consists of a linear, time-invariant, dynamic sub-system followed by a static nonlinearity. Our analysis allows a linear sub-system with a rational transfer function of arbitrary order, and a static polynomial non-linearity of arbitrary degree. To the best of our knowledge such an analysis with a very general assumption on the model structure is new in literature. The results obtained from this accuracy analysis is also used to choose the input power spectral densities that lead to optimal accuracy of the resulting model. To this end we give a tractable algorithm for solving the underlying optimization problem. We show that the set of all input power spectral densities can be parameterized using a finite number of parameters. In addition, there is a special parameterization which allows us to solve the optimization problem via an one dimensional search over a finite interval. This method shows a new tractable way to handle input design problems for nonlinear systems, which, so far, has been regarded as a problem of substantial difficulty.

We extend the D-optimal design approach to design Gaussian mixture inputs. This novel input design framework utilizes multiple identification experiments using Gaussian inputs with different mean values and variances, and thereby, allows us to tune the effective input amplitude distribution as well as the power spectral distributions at the same time.

1.2 The input design problem

Consider a causal dynamic system $\mathcal{S}$ with an input $u(t)$ and output $y(t)$. We wish to use the input-output data to identify a parametric model $M(\boldsymbol{\vartheta}, \mathbf{u}_t)$ of the system $\mathcal{S}$. Here $\boldsymbol{\vartheta}$ denotes the vector of unknown parameters, and

$$\mathbf{u}_t := [\; u(t) \;\; u(t-1) \;\; u(t-2) \;\; \cdots \;]$$

is potentially an infinite dimensional vector consisting of all the samples of u up to time point t, which determine the output $y(t)$. If we assume the structure of M is rich enough to represent the dynamic behavior of $\mathcal{S}$ with $\overset{\circ}{\boldsymbol{\vartheta}}$ representing the true value of $\boldsymbol{\vartheta}$, then we can write

$$y(t) = M(\overset{\circ}{\boldsymbol{\vartheta}}, \mathbf{u}_t) + e(t).$$

Here $e(t)$ represents the measurement noise, which is assumed to be zero mean, white, with a variance σ_e^2.

The prediction error estimate $\hat{\boldsymbol{\vartheta}}$ obtained from N samples of input-output data is given by [28]

$$\hat{\boldsymbol{\vartheta}} = \arg\min_{\boldsymbol{\vartheta}} \; \frac{1}{N} \sum_{t=1}^{N} \{y(t) - M(\boldsymbol{\vartheta}, \mathbf{u}_t)\}^2.$$

The quality of the identified model depends on the covariance matrix of $\hat{\boldsymbol{\vartheta}}$. For large N the normalized covariance matrix of $\sqrt{N}\hat{\boldsymbol{\vartheta}}$ is given by $\sigma_e^2\mathbf{J}^{-1}$, where $\mathbf{J}$ is the normalized Fischer's information matrix [28]

$$\mathbf{J} := \mathsf{E}\{\mathbf{v}_t\mathbf{v}_t^{\mathsf{T}}\}, \quad \mathbf{v}_t = \frac{\partial M(\mathring{\boldsymbol{\vartheta}}, \mathbf{u}_t)}{\partial \boldsymbol{\vartheta}}.$$

The input design problem, in one way or other, involves optimizing the statistical properties of $\mathbf{u}_t$ in order to maximize some monotonic function of $\mathbf{J}$ [1–3, 21, 22, 24]. In particular, the D-optimal design maximize the determinant of $\mathbf{J}$ [18].

Note that we need to know the true value $\mathring{\boldsymbol{\vartheta}}$ to calculate $\mathbf{J}$, and thus to design the input. In practice, one often uses an estimate in place of $\mathring{\boldsymbol{\vartheta}}$. For this reason, it is common practice in the input design literature to assume that $\mathring{\boldsymbol{\vartheta}}$ is known.

1.3 Input design for linear models

When $\mathcal{S}$ is linear then the situation is somewhat simplified in the sense that there is a linear filter $G_1(z)$ such that

$$\mathbf{v}_t = G_1(z)u(t).$$

Several papers have explored the relation between $\mathbf{J}$ and the statistics of $\mathbf{u}_t$, see e.g., [17, 27, 29–31, 39]. These methods utilize the relation

$$\mathbf{J} = \frac{1}{2\pi}\int_{-\pi}^{\pi} G_1(\mathrm{e}^{\mathrm{i}\omega})\Phi(\mathrm{e}^{\mathrm{i}\omega})G_1^*(\mathrm{e}^{\mathrm{i}\omega})\ \mathrm{d}\omega,$$

which essentially describes a linear map from the input power spectral density $\Phi(\mathrm{e}^{\mathrm{i}\omega})$ to the information matrix $\mathbf{J}$. This map has been analyzed in detail by [29]. It has been shown that the maximization of any concave function of $\mathbf{J}$ can be cast as a convex programming problem in a finite dimensional parameter, see, e.g. [1–3, 21, 22, 24]. Consequently, we can solve the input design problem efficiently via convex optimization tools [4, 19, 20, 37].

1.4 Input design for nonlinear models: the fundamental issues

When $\mathcal{S}$ is a non-linear system, $\mathbf{v}_t$ is no longer linear in u. Hence $\mathbf{J}$ depends on the higher order statistics of $\mathbf{u}_t$. If we compare this scenario with the linear case, we see that $\mathbf{J}$ does not only depend on the power spectral density Φ, but also on the amplitude distribution of $u(t)$. There are two major problems that one must address here:

i) Realization of designed input process

The first problem concerns the realization of the optimal input. Suppose that we have an input design algorithm that allows us to calculate the optimal power spectral density Φ_* and the amplitude distribution of $u(t)$. To the best of our knowledge there is no known way to generate a realization of $u(t)$ with pre-specified power spectral density and amplitude distribution. The standard way to generate a signal $u(t)$ with a pre-specified power spectral density Φ_* is to pass a white noise sequence $\epsilon(t)$ through a filter corresponding to a stable spectral factor of Φ_*. But this process allows no scope for tuning the amplitude distribution of $u(t)$. That is because computing

$u(t)$ involves computing an weighted average of $\epsilon(t), \epsilon(t-1), \epsilon(t-2), \ldots$. The rationale behind the celebrated central limit theorem tells us that the distribution of $u(t)$, so generated, is very close to Gaussian (thereby its amplitude distribution is fixed). This phenomenon is rather prominent even with an FIR filter of an order as low as 3 or 4.

ii) Computation of optimal input attributes

The second problem concerns formulating and solving the optimization problem required to calculate the optimal amplitude and power spectral distributions. This problem is likely to be non-convex and involves a large number of joint statistics of the potentially infinite dimensional vector $\mathbf{u}_t$. To the best of our knowledge, it has not yet been possible formulate the underlying optimization problem in a manner that can be solved via a numerical approach.

Since no satisfactory solutions to the above problems are available, the line of research in the nonlinear input design has undergone a significant paradigm shift. Most of the preliminary studies reported so far [10, 16, 23, 26, 38], have considered a deterministic setting. Among these the multi-level excitation approach [6, 7, 10, 26] appears to be popular lately. These deterministic methods do have their limitations. The multi-level approach is often not tractable when we increase the number of levels. The majority of these methods are unable to handle IIR-type non-linear systems.

1.5 Significance of our results relative to the existing literature

In this paper we address the two key problems described above in the context of Wiener systems. Following [16], we assume the input is a Gaussian stochastic process. The Gaussian assumption is justified by the fact mentioned above - we do not yet know a way to generate a non-white non-Gaussian process. Our solution to the problem of generating inputs with pre-specified amplitude and power spectral distributions is to employ a Gaussian mixture design. We propose to optimally mix the data obtained from a number of system identification experiments while estimating $\boldsymbol{\vartheta}$. Each of these experiments employ a Gaussian input signal with a different mean. By shifting the means of these different inputs and combining them with appropriate weighting factors we can control the effective amplitude distribution of the input. The Gaussian mixture approach decouples the problem of tuning and generating the amplitude distribution from that of the power spectral distribution.

For a single Gaussian, we give a tractable solution to the second problem under very general assumptions. A Wiener system consists of a linear sub-system G followed by a static nonlinearity. Our methods allow a rational model for G. The rational transfer function can be of arbitrary order. We model the static nonlinearity as a polynomial, and can allow an arbitrary polynomial degree. By allowing a rational transfer function for G, we let $\mathbf{u}_t$ to have an infinite length. Thus, it is very challenging to optimize any criterion of $\mathbf{J}$ with respect to the higher order moments of $\mathbf{u}_t$. In fact, it is quite difficult to just compute $\mathbf{J}$. Firstly, the available formulae for calculating higher order moments are quite challenging to program. More importantly, the complexity of the resulting algorithms typically grows exponentially with the length of $\mathbf{u}_t$ [25]. In fact, to the best of our knowledge no previous authors have considered handling this issue when G is not a finite impulse response system. Even when a finite impulse response system is considered in the literature, the order of the system have been restricted to 4 or less in the case studies considered therein. In this paper we show when the input process is Gaussian there is a simple algorithm to compute $\mathbf{J}$, and the complexity of our algorithm can be given as a polynomial in the system order. This analysis also reveals some interesting mathematical structures, that allow us to parameterize the set of all admissible information matrices with a finite number of parameters. Furthermore, $\mathbf{J}$ is linear in all but one of these parameters. In effect, the D-optimal input design problem can be cast as a non-convex optimization problem over a finite interval in one dimension, and such problems can be solved reliably by some suitable line search method. The parameters that we use to parameterize $\mathbf{J}$ are related to the input power spectral density via some interpolation constraints [29]. Therefore, the power spectral density of the optimal input can be computed by solving an analytic interpolation problem.

1.6 Outline of the paper

In Section 2 we present our Gaussian mixture solution to deal with the problem of tuning the amplitude distribution and the power spectral distribution of the input at the same time.

In Section 3.1 we discuss the fundamental constraints that we must consider while parameterizing a Wiener model for the purpose of identification. These constraints must be considered because of underlying non-uniqueness of the Wiener model representation. Suppose $G(z)$ is the transfer function of the linear subsystem, and $\wp(x)$ is the static nonlinear function of a Wiener system. Take any scalar $\alpha \neq 0$, and construct a second Wiener system with a transfer function $\alpha G(z)$, and static nonlinearity $\wp(x/\alpha)$. In terms of the input-output relationship the second model is identical to the first model. Since α can take uncountably many values, one can construct uncountably many models for the same Wiener system. Therefore, to ensure unique identifiability we must impose some additional constraint to remove this unwanted degree of freedom. In Section 3.1 we present a unified framework that allows us to cover several popular ways to impose this constraint. The advantage of using this general framework becomes clear in Remark 6, where we can clearly see the connection between identifiability and the Fischer's information matrix for Wiener model identification problem.

In Section 3.2 we summarize all the main analytical results. It starts with a key observation made in Lemma 1. It is observed that under very general conditions $\mathbf{J}$ is a function of some finite number of moments of a finite dimensional vector valued stationary stochastic process $\mathbf{x}$. In addition, $\mathbf{x}$ is obtained by passing the input u through a single input multiple output linear time-invariant filter. Lemma 1 is valid whenever the input u is a stationary stochastic process, and makes no assumption on the distribution of u.

The fact that $\mathbf{J}$ can be given in terms of finitely many moments of $\mathbf{x}$ is exploited further in Theorem 1, where we obtain a convenient closed form expression of $\mathbf{J}$ for a Gaussian u. This expression is easily computable, and useful, e.g. in input design. Theorem 2 takes the analysis one step further where we show that the determinant of $\mathbf{J}$ admits a simple expression. This expression allows us to derive tractable D-optimal design methods in Section 4.

Section 4 addresses the problem of D-optimal design of Gaussian inputs. The main result is Theorem 3, which gives a tractable algorithm for computing the D-optimal input attributes. The main contribution here is to show that the underlying infinite dimensional problem with infinitely many possible solutions admits a finite dimensional parameterization in terms of a parameter vector $\mathbf{h}$. Moreover, the resulting non-convex problem can be solved via a line-search. This is possible because $\det(\mathbf{J})$ is convex in all but one dimensions in the $\mathbf{h}$-space, and the set of all admissible $\mathbf{h}$ is compact.

The ideas presented in the paper were verified and tested via numerical simulations, and some main findings appear in Section 5. It turns out that the D-optimal design approach with Gaussian inputs is quite comparable with the D-optimal designs with deterministic inputs in terms of estimation accuracy. Furthermore, it has the advantage that it can handle very generic model structures unlike its deterministic counterparts.

2 Gaussian mixture design

In a nonlinear system identification problem we typically identify a model for a pre-specified range of input amplitude. Without of any loss of generality let us assume that the input must satisfy

$$-1 \leq u(t) \leq 1, \qquad \forall t. \tag{1}$$

As mentioned above, we decouple the problem of tuning and generating the input amplitude distribution from that of the power spectral distribution by using a Gaussian mixture design. We propose to optimally mix the data obtained from a number of system identification experiments while computing the estimate $\hat{\boldsymbol{\vartheta}}$. Each of these experiments employ a Gaussian input signal with a different mean. By shifting the means of these different inputs and combining them with appropriate weighting factors we can control the effective amplitude distribution of the input. We re-emphasize that the Gaussian assumption here is not by choice, but is imposed on us from the practical need of generating a non-white input signal.

Suppose we have decided to do p experiments where in the k th experiment we plan to set the mean of the Gaussian input to $\eta_u(k)$. One intuitive way choose $\eta_u(k)$ is to take

$$\eta_u(k) = -1 + \frac{2k}{p+1}, k = 1, 2, \ldots, p. \tag{2}$$

generating an uniform grid. This ensures that $|\eta_u(k)| < 1$ for all $k \in \{1, 2, \ldots, p\}$. We point out that this is not necessarily the best choice. It might be possible to optimally choose the grid based on some prior knowledge, but such an investigation is beyond the current scope. We offer in this section an optimal solution to the global problem.

Although we need a Gaussian $u(t)$ in order to realize a given power spectral density, a Gaussian signal does not satisfy (1) in a strict sense. However, by properly choosing the variance of the Gaussian signal we can ensure that (1) is satisfied with very high probability. For instance if $\eta_u(k) = 0$ then we take the input variance $\varsigma_k = 1/16$ to ensure $\text{Prob}\{|u(t)| > 1\} < 10^{-4}$. This idea can be generalized in mathematical terms as

$$\varsigma_k = \left(\frac{1}{\kappa} \min\{\eta_u(k) + 1, 1 - \eta_u(k)\} \right)^2, \tag{3}$$

where the parameter κ controls the value of $\text{Prob}\{|u(t)| > 1\}$, e.g., if $\kappa = 4$ then $\text{Prob}\{|u(t)| > 1\} = 0.001$. This way of choosing ς_k allows us to maximize the input energy for a target value of the boundary violation probability $\text{Prob}\{|u(t)| > 1\}$.

Now suppose that for each combination $(\eta_u(k), \varsigma_k)$ we can maximize the information matrix in some sense. Let $\mathbf{J}_k$ be the resulting optimal information matrix for the k th optimal experiment obtained in this way. Now we wish to find out the length N_k (in the number of samples) of k-th experiment such that resulting normalized information matrix

$$\bar{\mathbf{J}} = \frac{N_1\mathbf{J}_1 + N_2\mathbf{J}_2 + \cdots + N_p\mathbf{J}_p}{N_1 + N_2 + \cdots + N_p}$$

is optimized in some sense. In practice, the total length of experiment $N_1 + \cdots + N_p$ is known, and we need to determine the fractions

$$\kappa_j = \frac{N_j}{N_1 + N_2 + \cdots + N_p}, \qquad j = 1, 2, \ldots, p.$$

The resulting optimiztion problem takes the general form

$$\begin{aligned}
\underset{\kappa_1,\kappa_2,\ldots,\kappa_p}{\text{maximize}} \quad & \rho(\bar{\mathbf{J}}) && \text{(4a)}\\
\text{subject to} \quad & \bar{\mathbf{J}} = \kappa_1\mathbf{J}_1 + \kappa_2\mathbf{J}_2 + \cdots + \kappa_p\mathbf{J}_p && \text{(4b)}\\
& \kappa_k \geq 0, \quad k = 1, 2, \ldots, p, && \text{(4c)}\\
& \kappa_1 + \kappa_2 + \cdots + \kappa_p = 1, && \text{(4d)}
\end{aligned}$$

which requires to be solved with respect to $\kappa_1, \ldots, \kappa_p$. Here $\rho(\bar{\mathbf{J}})$ is some monotonic function of $\bar{\mathbf{J}}$ of our choice. Some popular choices include $\det(\bar{\mathbf{J}})$, $\lambda_{\min}(\bar{\mathbf{J}})$, etc. For these choices (4) is convex, and can be solved using well-known techniques implemented in popular packages like CVX [20].

The Gaussian mixture design described above allows us to get around the first problem in nonlinear input design. This makes it possible to find an optimum amplitude distribution of the input without worrying about the power spectral density. The decoupling also makes it possible to realize the desired amplitude distribution. Next we address the second problem, where we find the optimum power spectral density for a given pair (η_u, ς).

3 Information matrix and its determinant

In this section we present our main findings about the information matrix $\mathbf{J}$ and its determinant. We start in Section 3.1 with the basic notation and introduce the formal problem setting. In particular, we introduce a generalized framework for setting up the constraint to ensure unique identifiability of the Wiener model. Next in Section 3.2 we list the main results. In particular we use a state space representations of underlying transfer functions. We believe this approach simplifies the analysis, and illuminates the underlying mathematical structures to a significant extent.

3.1 Model parameterization and identifiability

A Wiener system is a cascade of a linear time invariant system followed by a static nonlinearity. We assume that the linear sub-system has a rational transfer function

$$G(z, \boldsymbol{\theta}) = \frac{g_0 + g_1 z^{-1} + \cdots + g_n z^{-n}}{1 + a_1 z^{-1} + \cdots + a_n z^{-n}}, \tag{5}$$

parameterized by the parameter vector $\boldsymbol{\theta}$ defined as

$$\boldsymbol{\theta} = [\ a_1 \quad \cdots \quad a_n \quad g_1 \quad \cdots \quad g_n \quad g_0\ ]^\mathsf{T}. \tag{6}$$

The output of the linear model is denoted by w:

$$w(t, \boldsymbol{\theta}) = G(z, \boldsymbol{\theta})u(t). \tag{7}$$

The static nonlinearity is modeled by a polynomial $\wp$ of order m:

$$\wp(x, \bar{\boldsymbol{\alpha}}) = \alpha_0 + \alpha_1 x + \cdots + \alpha_m x^m,$$

parameterized by the vector of polynomial coefficients

$$\bar{\boldsymbol{\alpha}} = [\ \alpha_0 \quad \alpha_1 \quad \cdots \quad \alpha_m\ ]^\mathsf{T}.$$

Therefore, the Wiener model equation takes the form

$$M(\boldsymbol{\vartheta}, \mathbf{u}_t) = \wp\{G(z, \boldsymbol{\theta})u(t), \bar{\boldsymbol{\alpha}}\}. \tag{8}$$

It is tempting to choose $\boldsymbol{\vartheta} = [\ \bar{\boldsymbol{\alpha}}^\mathsf{T} \quad \boldsymbol{\theta}^\mathsf{T}\ ]^\mathsf{T}$. But this parameterization fails to ensure unique identifiability. We cannot allow all the components of $\boldsymbol{\theta}$ and $\bar{\boldsymbol{\alpha}}$ to vary freely while remaining independent of each other. The reason is straightforward. The transfer operator between u and y does not change by dividing G by a scalar $\varrho \neq 0$, and multiplying α_k by ϱ^k for all $k = 1, 2, \ldots, m$. For this reason we must impose some additional constraint on the parameters. In this paper we allow varying the static gain of G freely, and impose a normalization constraint on $\bar{\boldsymbol{\alpha}}$.

Assumption 1. *There is a known vector*

$$\boldsymbol{\upsilon} = [\ \upsilon_0 \quad \upsilon_1 \quad \cdots \quad \upsilon_m\ ]^\mathsf{T} \tag{9}$$

such that

$$\alpha_0 \upsilon_0 + \alpha_1 \upsilon_1 + \cdots + \alpha_m \upsilon_m = 1, \tag{10}$$

where $\upsilon_\ell \neq 0$ for some known $\ell \in \{1, 2, \ldots, m\}$.

The choice of ℓ is often governed by the prior knowledge on the type of nonlinearity. Typically $\ell \neq 0$, because it is often the case that $\alpha_0 = 0$. For an odd (even) nonlinearity ℓ must be an odd (even) number. In our experience, the choice of ℓ does not influence the asymptotic large sample accuracy of the estimated model.

Example 1. *It is common to take $\boldsymbol{\upsilon} = (0, 1, \ldots, 0)$ or $\boldsymbol{\upsilon} = (0, \ldots, 0, 1)$. Another possibility would be to take $\boldsymbol{\upsilon} = (1, \ldots, 1)$ implying $\wp(1) = 1$. Note that the choice $\boldsymbol{\upsilon} = (1, 0, \cdots, 0)$ is forbidden. It leads to a model that is not identifiable.*

Since $\upsilon_\ell \neq 0$ under Assumption 1, we can rewrite (10) as

$$\alpha_\ell = \frac{1}{\upsilon_\ell} \left\{ 1 - \sum_{\substack{k=0 \\ k \neq \ell}}^{m} \upsilon_k \alpha_k \right\}. \tag{11}$$

Equation (11) can be built into the identification algorithm. We do not identify α_ℓ separately, but express it using (11). We define the parameter vector

$$\boldsymbol{\alpha} := [\ \alpha_{i_1} \quad \cdots \quad \alpha_{i_m}\ ]^\mathsf{T}, \tag{12}$$

where the indices $i_k \in \{0, 1, \ldots, m\}$ are chosen such that $i_k \neq \ell$ for all k, and $i_k \neq i_j$ whenever $k \neq j$. Note that mapping $k \to i_k$ is quite flexible, and we need not impose any further restriction on this mapping. The identification algorithm estimates

$$\boldsymbol{\vartheta} = [\ \boldsymbol{\alpha}^\mathsf{T} \quad \boldsymbol{\theta}^\mathsf{T}\ ]^\mathsf{T}$$

from the data. Defining

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & \cdots & 0 & -\upsilon_{i_1}/\upsilon_\ell \\ 0 & 1 & \cdots & 0 & -\upsilon_{i_2}/\upsilon_\ell \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\upsilon_{i_m}/\upsilon_\ell \end{bmatrix},$$

$$\mathbf{P} = [\ \boldsymbol{e}_{i_1} \quad \cdots \quad \boldsymbol{e}_{i_m} \quad \boldsymbol{e}_\ell\ ]^\mathsf{T}, \tag{13}$$

with $\boldsymbol{e}_k$ denoting the $k+1$ th column of $(m+1) \times (m+1)$ identity matrix, it can be verified from (11) that

$$[\ \boldsymbol{\alpha}^\mathsf{T} \quad \alpha_\ell\ ]^\mathsf{T} = \mathbf{P}\bar{\boldsymbol{\alpha}} = \mathbf{L}^\mathsf{T}\boldsymbol{\alpha}. \tag{14}$$

3.2 Main theoretical results

Let $\mathbf{a} = [\ a_1 \quad \cdots \quad a_n\ ]^\mathsf{T}$, and $\mathbf{g} = [\ g_1 \quad \cdots \quad g_n\ ]^\mathsf{T}$. Then we can write (5) as

$$G(z, \boldsymbol{\theta}) = g_0 + (\mathbf{g} - \mathbf{a}g_0)^\mathsf{T}(z\mathbf{I} - \mathbf{A}_1)^{-1}\mathbf{b}_1, \tag{15}$$

where $(\mathbf{A}_1, \mathbf{b}_1)$ is in controllable canonical form, *i.e.*

$$\mathbf{A}_1 = \begin{bmatrix} -a_1 & \cdots & -a_{n-1} & -a_n \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}, \ \mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{16}$$

Note that we can impose the structure (15) and (16) without any loss of generality. We make the following assumption throughout the paper, where $\mathring{\boldsymbol{\theta}}$ denotes the true value of $\boldsymbol{\theta}$.

Assumption 2. *$G(z, \mathring{\boldsymbol{\theta}})$ is asymptotically stable. Consequently, all the eigenvalues of $\mathring{\mathbf{A}}_1$ (which denotes the true value of $\mathbf{A}_1$) are located inside the unit disc in the complex plane. In addition, the state space realization (15) is minimal.*

Lemma 1. *Define the matrices $\mathbf{A}, \mathbf{b}, \mathbf{c}$ and $\bar{\mathbf{C}}$ as*

$$\bar{\mathbf{C}} = \begin{bmatrix} \mathbf{I} & -\mathring{g}_0\mathbf{I} & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & -\mathring{\mathbf{a}}^\mathsf{T} & 1 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} \mathbf{0}_{n\times 1} \\ \mathbf{0}_{n\times 1} \\ 1 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} \mathbf{0}_{n\times 1} \\ \mathring{\mathbf{g}} \\ \mathring{g}_0 \end{bmatrix},$$

$$\mathbf{A} = \bar{\mathbf{C}} \begin{bmatrix} \mathring{\mathbf{A}}_1 & -\mathbf{b}_1(\mathring{\mathbf{g}} - \mathring{\mathbf{a}}\mathring{g}_0)^\mathsf{T} & \mathbf{0}_{n\times 1} \\ \mathbf{0}_{n\times n} & \mathring{\mathbf{A}}_1 & \mathbf{b}_1 \\ \mathbf{0}_{1\times n} & \mathbf{0}_{1\times n} & 0 \end{bmatrix} \bar{\mathbf{C}}^{-1}, \tag{17}$$

where $\mathring{\mathbf{A}}_1, \mathring{g}_0$, *etc are obtained by setting* $\boldsymbol{\theta} = \mathring{\boldsymbol{\theta}}$ *in* $\mathbf{A}_1, g_0$, *etc. Consider the stochastic process* $\mathbf{x}$ *which is given in state space form as*

$$\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t-1) + \mathbf{b}u(t). \tag{18}$$

Then $w(t, \mathring{\boldsymbol{\theta}}) = \mathbf{c}^\mathsf{T}\mathbf{x}(t)$, *and*

$$\mathbf{v}_t = \begin{bmatrix} \mathbf{L}\mathbf{P}\mathbf{z}(t, \mathring{\boldsymbol{\theta}}) \\ \mathbf{x}(t)\boldsymbol{\alpha}_2^\mathsf{T}\mathbf{z}(t, \mathring{\boldsymbol{\theta}}) \end{bmatrix}, \tag{19}$$

where we define

$$\mathbf{z}(t, \boldsymbol{\theta}) := [\ 1 \quad w(t, \boldsymbol{\theta}) \quad \{w(t, \boldsymbol{\theta})\}^2 \quad \cdots \quad \{w(t, \boldsymbol{\theta})\}^m\ ]^\mathsf{T}, \tag{20}$$

$$\boldsymbol{\alpha}_2 = [\ \mathring{\alpha}_1 \quad 2\mathring{\alpha}_2 \quad \cdots m\mathring{\alpha}_m \quad 0\ ]^\mathsf{T}, \tag{21}$$

with $\mathring{\alpha}_k$ *being the true value of* α_k*.*

Proof: See Appendix A. ■

Remark 1. Lemma 1 does cover the case when G is of finite impulse response type, i.e.,

$$G(z, \boldsymbol{\theta}) = g_0 + g_1 z^{-1} + \cdots + g_n z^{-n}.$$

In this case $\boldsymbol{\theta} = [\ \mathbf{g}^\mathsf{T} \quad g_0]^\mathsf{T}$, and $\mathbf{a} = 0$. The expressions (15) and (16) still hold with $\mathbf{a} = 0$. While finding a realization of G_1 we do not need to consider the derivatives with respect to $\mathbf{a}$. As a result we get

$$\mathbf{A} = \begin{bmatrix} \mathring{\mathbf{A}}_1 & \mathbf{b}_1 \\ \mathbf{0}_{1\times n} & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \mathbf{0}_{n\times 1} \\ 1 \end{bmatrix},$$

$\bar{\mathbf{C}} = \mathbf{I}$ and $\mathbf{c} = \boldsymbol{\theta}$.

The consequence of Lemma 1 is that $\mathbf{J} = \mathsf{E}\{\mathbf{v}_t\mathbf{v}_t^\mathsf{T}\}$ is a function of the moments of the state vector $\mathbf{x}$. For the purpose of setting up an input design problem we can parameterize $\mathbf{J}$ in terms of the moments of the random vector $\mathbf{x}$. In particular, when $u(t)$ is Gaussian, then so is $\mathbf{x}(t)$. Hence for a Gaussian input $\mathbf{J}$ is a function of the mean and the covariance matrix of $\mathbf{x}(t)$. As the next Theorem reveals, we can obtain a closed form expression for $\mathbf{J}$.

Assumption 3. *The input process* $u(t)$ *is stationary Gaussian with mean* η_u*.*

Under Assumption 3, $\mathbf{x}$ is a Gaussian random vector with mean

$$\boldsymbol{\eta} := \mathsf{E}\{\mathbf{x}(t)\} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}\eta_u. \tag{22}$$

Let us define

$$\boldsymbol{\Sigma} = \mathsf{E}\{[\mathbf{x}(t) - \boldsymbol{\eta}][\mathbf{x}(t) - \boldsymbol{\eta}]^\mathsf{T}\}. \tag{23}$$

Consequently, $\mathbf{c}^\mathsf{T}\mathbf{x}(t)$ is a Gaussian random variable such that

$$\gamma := \mathsf{E}\{\mathbf{c}^\mathsf{T}\mathbf{x}(t)\} = \mathbf{c}^\mathsf{T}\boldsymbol{\eta}. \tag{24a}$$

$$\sigma := \mathsf{E}\{\mathbf{c}^\mathsf{T}\mathbf{x}(t) - \gamma\}^2 = \mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{c}. \tag{24b}$$

In the rest of the paper we denote

$$\boldsymbol{\Lambda} := \mathsf{E}\{\mathbf{z}(t, \mathring{\boldsymbol{\theta}})[\mathbf{z}(t, \mathring{\boldsymbol{\theta}})]^\mathsf{T}\}.$$

Remark 2. It is possible to express $\mathbf{\Sigma}$ as well in terms of $\mathbf{A}$, $\mathbf{b}$, and the power spectral density Φ of u. However, we postpone that for a while. We first express $\mathbf{J}$ in terms of $\mathbf{\Sigma}$ and $\boldsymbol{\eta}$, and later connect Φ with $\mathbf{\Sigma}$. This approach suits the purpose of input design, where it is simpler to work with a parameterization of $\mathbf{\Sigma}$ than to work with Φ directly.

Remark 3. The correlation matrix $\mathbf{\Lambda}$ can be given entirely as a function of the mean γ and variance σ of $\mathbf{c}^\intercal\mathbf{x}(t)$. Many different ways are used in the literature to express the moments of the scalar valued normal density. There are some explicit formulae for smaller orders. We find it convenient to use a recursive formula in the implementation. Let us denote $\mu_k(\gamma, \sigma) := \mathsf{E}\{(\mathbf{c}^\intercal\mathbf{x})^k\}$. So μ_k is a function of σ and μ. Then $\mu_k(\gamma, \sigma)$ satisfies the recursion [32, Chapter 5]:

$$\mu_k(\gamma, \sigma) = \gamma^k + \frac{k(k-1)}{2}\int_0^\sigma \mu_{k-2}(\tau, \sigma)\ \mathrm{d}\tau. \tag{25}$$

Note that the recursion (25) needs to be carried out separately for even and odd values of k. For even valued k one can initialize the recursion with $\mu_0(\gamma, \sigma) = 1$, and for the odd values of k we initialize with $\mu_1(\gamma, \sigma) = \gamma$. This allows us to form

$$\mathbf{\Lambda} = \begin{bmatrix} \mu_0(\gamma,\sigma) & \mu_1(\gamma,\sigma) & \cdots & \mu_m(\gamma,\sigma) \\ \mu_1(\gamma,\sigma) & \mu_2(\gamma,\sigma) & \cdots & \mu_{m+1}(\gamma,\sigma) \\ \vdots & \vdots & & \vdots \\ \mu_m(\gamma,\sigma) & \mu_{m+1}(\gamma,\sigma) & \cdots & \mu_{2m+1}(\gamma,\sigma) \end{bmatrix}.$$

Since $\mathbf{x}$ is Gaussian, all the moments of $\mathbf{x}$ can be expressed as functions of $\boldsymbol{\eta}$ and $\mathbf{\Sigma}$. This allows us to derive manageable expressions of $\mathbf{J}$ as a function of $\boldsymbol{\eta}$ and $\mathbf{\Sigma}$. This is shown next.

Theorem 1. *Define*

$$\boldsymbol{\alpha}_1 = [\ 0 \quad \mathring{\alpha}_1, \quad 2\mathring{\alpha}_2 \quad \cdots \quad m\mathring{\alpha}_m\ ]^\intercal, \qquad \beta = \boldsymbol{\alpha}_2^\intercal\mathbf{\Lambda}\boldsymbol{\alpha}_2, \tag{26}$$

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{\sigma} & -\frac{\gamma}{\sigma} \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{F} := [\ \mathbf{\Sigma}\mathbf{c} \quad \boldsymbol{\eta}\ ], \qquad \mathbf{H} = \begin{bmatrix} \beta\sigma & 0 \\ 0 & 0 \end{bmatrix}.$$

Partition $\mathbf{J}$ *as*

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{21}^\intercal \\ \mathbf{J}_{21} & \mathbf{J}_{22} \end{bmatrix},$$

where $\mathbf{J}_{11}$ *is of size* $m \times m$*, while* $\mathbf{J}_{22}$ *is of size* $(2n+1) \times (2n+1)$*. Then*

$$\begin{aligned} \mathbf{J}_{11} &= \mathbf{L}_1\mathbf{\Lambda}\mathbf{L}_1^\intercal, \\ \mathbf{J}_{21} &= \mathbf{F}\mathbf{Q}\mathbf{L}_2\mathbf{\Lambda}\mathbf{L}_1^\intercal \\ \mathbf{J}_{22} &= \mathbf{F}\mathbf{Q}(\mathbf{L}_2\mathbf{\Lambda}\mathbf{L}_2^\intercal - \mathbf{H})\mathbf{Q}^\intercal\mathbf{F}^\intercal + \beta\mathbf{\Sigma}, \end{aligned} \tag{27}$$

where

$$\begin{aligned} \mathbf{L}_1 &:= \mathbf{L}\mathbf{P}, \\ \mathbf{L}_2 &:= \begin{bmatrix} \boldsymbol{\alpha}_1^\intercal \\ \boldsymbol{\alpha}_2^\intercal \end{bmatrix} = \begin{bmatrix} 0 & \mathring{\alpha}_1 & 2\mathring{\alpha}_2 & \cdots & m\mathring{\alpha}_m \\ \mathring{\alpha}_1 & 2\mathring{\alpha}_2 & \cdots & m\mathring{\alpha}_m & 0 \end{bmatrix}. \end{aligned} \tag{28}$$

Proof: See Appendix B. ■

Remark 4. Expressions given by Theorem 1 allow us to compute $\mathbf{J}$ in a simple way. To the best of our knowledge there is no similar expressions in the literature allowing this computational advantage.

The matrices $\mathbf{Q}, \mathbf{H}, \boldsymbol{\Lambda}, \mathbf{L}_1, \mathbf{L}_2$ and β share an interesting property. They depend only on the true parameter vector $\mathring{\boldsymbol{\vartheta}}$ and the second order statistics (consisting of γ and σ) of the stochastic process $w(t, \mathring{\boldsymbol{\theta}}) = \mathbf{c}^\intercal \mathbf{x}(t)$. These quantities remain constant so long γ and σ remain constant, even though the input power spectral density (and thus $\boldsymbol{\Sigma}$) may vary. This is due to the fact that the estimation accuracy of the static nonlinearity depends only on the amplitude distribution of $w(t)$, regardless of $\boldsymbol{\Sigma}$ (or equivalently, Φ). This observation plays a key role in the sequel, and is formalized via the following definition.

Definition 1. *A quantity is called w-dependent if it is a function of $\mathring{\boldsymbol{\vartheta}}$, σ and γ only.*

The expressions given in Theorem 1 may not seem appealing from the point of view of setting up an optimization problem for input design that can be solved in a tractable manner. The next result is more attractive in that regard.

Theorem 2. *The determinant of $\mathbf{J}$ is given by*

$$\det(\mathbf{J}) = \frac{\beta^{2n} r_1^2}{\sigma} \det(\mathbf{J}_{11}) \det(\boldsymbol{\Sigma}). \tag{29}$$

where $r_1 = \boldsymbol{\alpha}_1^\intercal \boldsymbol{\upsilon} (\boldsymbol{\upsilon}^\intercal \boldsymbol{\Lambda}^{-1} \boldsymbol{\upsilon})^{-1/2}$.

Proof: See Appendix C ■

Remark 5. The expression of $\det(\mathbf{J})$ in (29) has some nice structure. The factor

$$f := \beta^{2n} r_1^2 \det(\mathbf{J}_{11})/\sigma \tag{30}$$

is w-dependent, and remains constant when the statistics of $w(t, \boldsymbol{\theta}_0)$ remain invariant. On the other hand it is well-known from the literature on the input design for linear systems that we can parameterize $\det(\boldsymbol{\Sigma})$ in a convex manner using a finite number of parameters. When the mean η_u of the input is kept fixed, then the above facts let us solve the D-optimal design problem for Wiener models via an one dimensional search in σ. To emphasize the w-dependence of f we write it as $f(\gamma, \sigma)$. When we consider a situation where γ is fixed and known, then we write it as $f(\sigma)$.

Remark 6. Note that $\mathbf{J}$ is singular when $r_1 = 0$. This means that the normalization of the form described in Assumption 1 ensures identifiability (and thus a non-singular information matrix) only when

$$0 \neq \boldsymbol{\alpha}_1^\intercal \boldsymbol{\upsilon} = \upsilon_1 \mathring{\alpha}_1 + 2\upsilon_2 \mathring{\alpha}_2 + \cdots + m \upsilon_m \mathring{\alpha}_m, \tag{31}$$

see the definition of r_1 in the statement of Theorem 2. We can easily construct a case where (31) fails to hold. That is $\boldsymbol{\upsilon} = (1, 0, \ldots, 0)$. It is straightforward to see why this choice leads to lack of identifiability: it still allows us to simultaneously vary the gain of the linear subsystem and the factors $\{\alpha_k\}_{k=1}^m$, while the constraint (10) is satisfied.

By imposing the constraint (10) we restrict the search space to the hyperplane

$$\mathcal{H} = \left\{ (\alpha_0, \alpha_1, \cdots \alpha_m) : \sum_{k=0}^{m} \alpha_k \upsilon_k = 1 \right\}$$

By assumption, $(\mathring{\alpha}_0, \mathring{\alpha}_1, \cdots, \mathring{\alpha}_m) \in \mathcal{H}$. The model is identifiable when $\mathcal{H}$ intersects with the manifold

$$\mathcal{M} = \{(\mathring{\alpha}_0, \varrho\mathring{\alpha}_1, \cdots, \varrho^m \mathring{\alpha}_m) : \varrho \neq 0\}$$

only at the point $(\mathring{\alpha}_0, \mathring{\alpha}_1, \cdots, \mathring{\alpha}_m)$, which corresponds to $\varrho = 1$. We have local identifiability at $(\mathring{\alpha}_0, \mathring{\alpha}_1, \cdots, \mathring{\alpha}_m)$ only if $\mathcal{M}$ is not oriented along $\mathcal{H}$ at $(\mathring{\alpha}_0, \mathring{\alpha}_1, \cdots, \mathring{\alpha}_m)$, i.e., $\varrho = 1$. In other words, we do not want the directional derivative $(0, 2\mathring{\alpha}_1, \cdots, m\mathring{\alpha}_m) =: \boldsymbol{\alpha}_1$ of $\mathcal{M}$ at $\varrho = 1$ to be perpendicular to $\boldsymbol{\upsilon}$, which is identical to (31).

4 D-optimal design

Having a Gaussian input enables us to express $\mathbf{J}$ (and thus $\mathbf{J}^{-1}$ and $\det(\mathbf{J})$ solely as a function of $\boldsymbol{\Sigma}$ and $\boldsymbol{\eta}$. Therefore the problem of optimizing the information matrix $\mathbf{J}$ becomes the problem of fine tuning $\boldsymbol{\eta}$ and $\boldsymbol{\Sigma}$ appropriately by choosing the right input power spectral density Φ, and η_u. In Section 2 we have already shown how we use η_u as a tool to fine tune the amplitude distribution. In this section we give a tractable algorithm to compute the D-optimal Φ for a fixed η_u. In Section 4.1 we summarize the main findings in form of Theorem 3, and subsequently, focus on how the underlying algorithm could be implemented to compute an optimal Φ. In Section 4.2 we discuss the underlying theory and the proofs leading to Theorem 3.

4.1 The main result and algorithm implementation

Consider the D-optimal input design problem for a fixed mean η_u and a fixed variance ς of $u(t)$:

$$\text{maximize} \quad \det(\mathbf{J}) \tag{32a}$$

$$\text{subject to} \quad \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi(\mathrm{e}^{\mathrm{i}\omega})\, \mathrm{d}\omega = \varsigma, \tag{32b}$$

which needs to be solved by finding an optimal input power spectral density Φ. Now solving for Φ is equivalent of finding the ‘positive real half spectrum’ $\Phi_\ddagger$ defined as

$$\Phi_\ddagger(z) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{1 + z\mathrm{e}^{-\mathrm{i}\omega}}{1 - z\mathrm{e}^{-\mathrm{i}\omega}} \Phi(\mathrm{e}^{\mathrm{i}\omega})\, \mathrm{d}\omega, \quad |z| < 1. \tag{33}$$

The “equivalence” of Φ and $\Phi_\ddagger$ follows once we expand the integrand in (33) in a convergent power series in z for $|z| < 1$ yielding an expansion

$$\Phi_\ddagger(z) = \frac{\phi_0}{2} + \sum_{k=1}^{\infty} \phi_{-k} z^k,$$

with the expansion coefficients being the Fourier coefficients:

$$\phi_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi(\mathrm{e}^{\mathrm{i}\omega})\ \mathrm{e}^{\mathrm{i}\omega k}\ \mathrm{d}\omega.$$

Now we give the main result of this section as the next theorem. We need a few definitions to state it. First we define the linear map $\mathcal{L}$ from the set of all $2n+1$ dimensional vectors to the set of all $(2n+1) \times (2n+1)$ Hermitian matrices such that $\mathbf{\Pi} = \mathcal{L}(\mathbf{h})$ satisfies the Lyapunov equation $\mathbf{\Pi} = \mathbf{A}\mathbf{\Pi}\mathbf{A}^{\mathsf{T}} + \mathbf{b}\mathbf{h}^{\mathsf{T}} + \mathbf{h}\mathbf{b}^{\mathsf{T}}$. In addition, we define the set

$$\mathbb{K} = \{\mathbf{h} \in \mathbb{R}^{2n+1}\ :\ \mathcal{L}(\mathbf{h}) \succeq 0,\ \mathbf{b}^{\mathsf{T}}\bar{\mathbf{C}}^{-1}\mathbf{h} = \varsigma/2\ \}.$$

Here we write $\mathbf{M} \succeq 0$ to convey that $\mathbf{M}$ is a non-negative definite matrix. That $\mathbb{K}$ is convex is readily verified as it is characterized via a linear equality and a linear matrix inequality in $\mathbf{h}$.

Theorem 3. *Let us define $\sigma_{\min}$ and $\sigma_{\max}$ as the solutions to the convex problems*

$$\sigma_{\min} = \min\ \mathbf{c}^{\mathsf{T}}\mathcal{L}(\mathbf{h})\mathbf{c}, \quad \text{subject to } \mathbf{h} \in \mathbb{K}, \tag{34}$$

$$\sigma_{\max} = \max\ \mathbf{c}^{\mathsf{T}}\mathcal{L}(\mathbf{h})\mathbf{c}, \quad \text{subject to } \mathbf{h} \in \mathbb{K}. \tag{35}$$

Then the optimum value d of the D-optimal design problem (32) is given by the solution to the one dimensional search

$$d = \max_{\sigma_{\min} \leq \sigma \leq \sigma_{\max}} \chi(\sigma) f(\sigma), \tag{36}$$

where f is defined in (30), and the function $\chi(\sigma)$ is defined for any σ satisfying $\sigma_{\min} \leq \sigma \leq \sigma_{\max}$ as the solution to the convex problem

$$\begin{aligned} \chi(\sigma) = \max_{\mathbf{h}} \quad & \det\{\mathcal{L}(\mathbf{h})\} \\ \text{subject to} \quad & \mathbf{h} \in \mathbb{K}, \\ & \mathbf{c}^{\mathsf{T}}\mathcal{L}(\mathbf{h})\mathbf{c} = \sigma. \end{aligned} \tag{37}$$

Let the σ dependent argument minimizer of the problem (37) be denoted by $\mathbf{h}_*(\sigma)$, and suppose that the argument minimizer of (36) is σ_*. Then

$$\mathbf{h}_*(\sigma_*) = \Phi_{\ddagger}(\mathbf{A})\mathbf{b} \tag{38}$$

for any solution Φ of the D-optimal design problem (32).

Theorem 3 will be established in the next section. Here we note some key points on the implementation of the algorithm outlined by Theorem 3.

4.1.1 Computing $\mathbf{h}_*(\sigma_*)$

The optimization problems appearing in (34), (35) are semidefinite programs, while (37) is a max-det problem subject to semidefinite constraints. To express these in the standard semidefinite programming form we characterize $\mathbb{K}$ via a linear matrix inequality and a linear equality. To see the details let us define $\mathbf{\Pi}_i = \mathcal{L}(\mathbf{e}_i)$, where $\mathbf{e}_i$ denotes the i th column of the $(2n+1) \times (2n+1)$ identity matrix. By definition of $\mathbf{\Pi}_i$ and $\mathcal{L}$ we can find $\mathbf{\Pi}_i$ by solving the Lyapunov equation

$$\mathbf{\Pi}_i = \mathbf{A}\mathbf{\Pi}_i\mathbf{A}^\intercal + \mathbf{e}_i\mathbf{b}^\intercal + \mathbf{b}\mathbf{e}_i^\intercal.$$

If h_i denotes the i th component of $\mathbf{h}$, then we can write

$$\mathbf{h} = \sum_{i=1}^{2n+1} \mathbf{e}_i h_i, \ \ \Rightarrow \ \ \mathcal{L}(\mathbf{h}) = \sum_{i=1}^{2n+1} \mathcal{L}(\mathbf{e}_i) h_i = \sum_{i=1}^{2n+1} \mathbf{\Pi}_i h_i.$$

Hence by definition of $\mathbb{K}$, and the definition of $\mathbf{b}$ in (17) we have $\mathbf{h} \in \mathbb{K}$ if and only if

$$\mathbf{\Pi}_1 h_1 + \cdots + \mathbf{\Pi}_{2n+1} h_{2n+1} \succeq 0, \quad \mathbf{b}^\intercal \bar{\mathbf{C}}^{-1}\mathbf{h} = \varsigma/2.$$

The above characterizations of $\mathbb{K}$ and $\mathcal{L}$ can be used in (35), (34) and (37) to reduce them in the familiar forms involving linear matrix inequalities, and linear equalities in $\mathbf{h}$. Each of the convex problems (35), (34) and (37) are solved in an $2n+1$ dimensional variable $\mathbf{h}$, and thus the number of floating point operations per iteration needed to solve these is at most of the order of $(2n+1)^3$.

Theorem 3 outlines a three step procedure to find $\mathbf{h}_*(\sigma_*)$:

1. Solve $\sigma_{\min}$ and $\sigma_{\max}$ by solving the semidefinite programs (34) and (35).

2. Use a convenient line search method[1] to solve (36). The line search algorithm needs to evaluate the functions f and χ for different values of $\sigma \in [\sigma_{\min}, \sigma_{\max}]$. Evaluation of f can be done using the formula (30), while the evaluation of χ can be done by solving the semidefinite program (37).

3. The solution to the line search (36) yields σ_*. We need to calculate $\mathbf{h}_*(\sigma_*)$ by solving (37) once more by setting $\sigma = \sigma_*$ in (37).

For a numerical illustration of the above procedure see Section 5.2.

Remark 7. Note that $f(\sigma)\chi(\sigma)$ is a non-concave function in general, and may have multiple local maximum points in $[\sigma_{\min}, \sigma_{\max}]$. In fact, we have no proof to establish the uniqueness of the global optimum point. However, in the cases examined while producing this manuscript we found that $f(\sigma)\chi(\sigma)$ is often concave on $[\sigma_{\min}, \sigma_{\max}]$, as in the example in Figure 2a.

Remark 8. Whether or not the 3-step approach will reach the global optimum point, depends on the power of the line search method employed. This statement of course assumes that all the convex programs (35), (34) and (37) can be solved globally, which is valid in theory, and is very reasonable in practice. Typically, a line search over a finite interval is a simple problem, and most solvers do find the global optimum. In fact, a plot $f(\sigma)\chi(\sigma)$ with σ over the interval $[\sigma_{\min}, \sigma_{\max}]$ is often informative enough to find σ_*. Therefore, it is reasonable to expect that the 3-step strategy yields the global optimum solution.

[1] In our simulations we have used the `fminbnd` function of Matlab.

4.1.2 From $\mathbf{h}_*(\sigma_*)$ to Φ

Once we know $\mathbf{h}_*(\sigma_*)$, the next task is to find a suitable Φ such that (38) holds. This amounts to solving a Nevanlinna-Pick interpolation problem, see [13], and in general there are infinitely many possible choices of Φ for which $\mathbf{h}_*(\sigma_*) = \Phi_\ddagger(\mathbf{A})\mathbf{b}$. The freedom in choosing Φ can be exploited further to obtain a some spectral shape of our liking [29]. In this section we briefly cover a simple way to construct multi-sine excitations, where Φ is of the form

$$\Phi(\mathrm{e}^{\mathrm{i}\omega}) = 2\pi \sum_{k=1}^{M} q_k\{\delta(\omega - \omega_k) + \delta(\omega + \omega_k)\}, \tag{39}$$

where δ denotes the Dirac's delta function, $q_k \geq 0$, and $0 < \omega_k \leq \pi$ for all k, and the frequency grid $\{\omega_1, \omega_2, \ldots, \omega_M\}$ is given.

We note that there are many methods of finding a Φ with a rational power spectral density, see e.g., [5,8,9,11,12,15,33,34,36]. In particular, [5] gave the celebrated spectral zero assignment approach. When considered in our context, the algorithm in [5] allows us to find a $2n+1$ order rational Φ with spectral zero locations chosen by the user. In [15] the spectral zero assignment approach is generalized. This algorithm allows us to start with a target function Φ_0 and then find an order $2n+1$ rational Φ such that the Kullback-Leibler distance between Φ and Φ_0 is minimized. The contributions in [33,34] gave a more sound numerical algorithm to solve the Kullback-Leibler and the spectral zero assignment problems. In [8] an algorithm to minimize the Hellinger distance between Φ and Φ_0 is proposed. Multi-variable generalizations of the above approach are available in [9,11,36]. Some illustrative examples in the context of input design are available in [29].

To explain our approach to contruct Φ as in (39), let us recall that the matrix $\Phi_\ddagger(\mathbf{A})$ denotes the evaluation of the function $\Phi_\ddagger$ at $\mathbf{A}$. This makes sense because $\Phi_\ddagger$ is analytic whenever $|z| < 1$, and at the same time all eigenvalues of $\mathbf{A}$ are inside the unit disc, (see Assumption 2). Thus, we can write, see (33),

$$\begin{aligned}\Phi_\ddagger(\mathbf{A})\mathbf{b} &= \frac{\phi_0}{2}\mathbf{I} + \sum_{k=1}^{\infty} \phi_{-k}\mathbf{A}^k\mathbf{b}, \\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi} \Phi(\mathrm{e}^{\mathrm{i}\omega})(\mathbf{I} - \mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1}(\mathbf{I} + \mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})\mathbf{b}.\, \mathrm{d}\omega.\end{aligned} \tag{40}$$

When Φ is of the form (39), then (38) holds if and only if

$$\mathbf{h}_* = \sum_{k=1}^{M} \boldsymbol{\psi}_k q_k = \boldsymbol{\Psi}\mathbf{q}, \tag{41}$$

where by (40) we have

$$\begin{aligned}\boldsymbol{\psi}_k = \frac{1}{2}\{&(\mathbf{I} - \mathbf{A}\mathrm{e}^{-\mathrm{i}\omega_k})^{-1}(\mathbf{I} + \mathbf{A}\mathrm{e}^{-\mathrm{i}\omega_k}) \\ &+ (\mathbf{I} - \mathbf{A}\mathrm{e}^{\mathrm{i}\omega_k})^{-1}(\mathbf{I} + \mathbf{A}\mathrm{e}^{\mathrm{i}\omega_k})\}\mathbf{b},\end{aligned} \tag{42}$$

$$\boldsymbol{\Psi} = [\ \boldsymbol{\psi}_1 \ \ \boldsymbol{\psi}_2 \ \ \cdots \ \ \boldsymbol{\psi}_M\ ], \qquad \mathbf{q} = [\ q_1 \ \ q_2 \ \ \cdots \ \ q_M\ ]^\mathsf{T}.$$

Remark 9. We emphasize that the set $\{\omega_1, \omega_2, \ldots, \omega_M\}$ should be sufficiently 'rich' to ensure that there exists a $\mathbf{q}$ with nonnegative components such that (41) holds. Typically one takes $\omega_k = k\pi/M$. In that case we often need $M > 100$. However, in practice one takes M large, typically of the order of 10000 [35]. Hence the feasibility issue is not of much practical importance.

Remark 10. Recall that the analysis presented herein assumes Gaussian input signal. Given a power spectral density function of the form (39), we can synthesize a input signal consistent with (39) as

$$u(t) = \sum_{k=1}^{M} 2\sqrt{q_k}\cos(\omega_k t + \phi_k). \tag{43}$$

By taking $\{\phi_k\}_{k=1}^{M}$ as mutually independent random variables distributed identically and uniformly over $[0, 2\pi)$, this signal will be asymptotically ($M \to \infty$) Gaussian distributed [35]. See Figure 2c in Section 5.2 for an illustration in the present context.

In practice (41) is an over-determined system, with infinitely many solutions even in presence of the constraint $q_k \geq 0, \ \forall k$. We have found that a robust way is to pick a solution which leads to a spectrum as flat as possible. One possible way to get such a solution is to choose $\mathbf{q}$ with the smallest infinity-norm by solving the linear program

$$\begin{aligned} \underset{\ell,\mathbf{q}}{\text{minimize}} \quad & \ell \\ \text{subject to} \quad & 0 \leq q_k \leq \ell, \\ & \mathbf{h}_* = \mathbf{\Psi}\mathbf{q}. \end{aligned} \tag{44}$$

See Figure 2b for an illustration of the nature of solution generated by the optimization (44).

4.2 Proof of Theorem 3 and some relevant discussion

The expression of $\det(\mathbf{J})$ given by Theorem 2 factors $\det(\mathbf{J})$ in two terms. The first term $f(\gamma, \sigma)$ is w-dependent, and depends only on γ and σ. The second term is $\det(\mathbf{\Sigma})$. Since η_u is fixed and given, we know γ, while $\sigma = \mathbf{c}^\intercal \mathbf{\Sigma} \mathbf{c}$ can be calculated if we know $\mathbf{\Sigma}$. Hence the power spectral density Φ controls $\det(\mathbf{J})$ "via" the covariance matrix $\mathbf{\Sigma}$. Now $\mathbf{\Sigma}$ being a $(2n+1) \times (2n+1)$ positive definite matrix, resides in the positive semidefinite cone embedded in the finite dimensional vector space of all $(2n+1) \times (2n+1)$ symmetric matrices. On the other hand Φ, being a function, is an infinite dimensional object. In addition, Φ resides in the cone of all positive functions over $[-\pi, \pi]$. That is because $\Phi(\mathrm{e}^{\mathrm{i}\omega}) \geq 0$ for all frequency ω. Using (18) and the definition of $\mathbf{\Sigma}$ in (23) we know

$$\mathbf{\Sigma} = \frac{1}{2\pi}\int_{-\pi}^{\pi} (\mathbf{I} - \mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1}\mathbf{b}\Phi(\mathrm{e}^{\mathrm{i}\omega})\mathbf{b}^\intercal(\mathbf{I} - \mathbf{A}^\intercal\mathrm{e}^{\mathrm{i}\omega})^{-1}\,\mathrm{d}\omega. \tag{45}$$

Equation (45) reveals that there is a linear map from Φ to $\mathbf{\Sigma}$, and [14] investigated this relation in a comprehensive manner. The result thereof can be summarized as follows.

Theorem 4. *Given a matrix $\mathbf{\Sigma}$ there exists some valid power spectral density function $\Phi(\mathrm{e}^{\mathrm{i}\omega})$ (i.e., $\Phi(\mathrm{e}^{\mathrm{i}\omega}) \geq 0, \ \forall\omega$) satisfying (45) if and only if $\mathbf{\Sigma} = \mathcal{L}(\mathbf{h}) \succeq 0$ for some $2n+1$ dimensional vector* $\mathbf{h}$.

Proof: See [14].

Remark 11. Suppose that we are given a function $\Phi(\mathrm{e}^{\mathrm{i}\omega})$ such that $\Phi(\mathrm{e}^{\mathrm{i}\omega}) > 0$, $\forall\omega$. Then we can calculate the integral in the right hand side of (45) to find $\mathbf{\Sigma}$. Then Theorem 4 says there is a $\mathbf{h}$ such that

$$\mathbf{\Sigma} = \mathbf{A}\mathbf{\Sigma}\mathbf{A}^\intercal + \mathbf{h}\mathbf{b}^\intercal + \mathbf{b}\mathbf{h}^\intercal, \tag{46}$$

which is a linear matrix valued equation in $\mathbf{h}$. Considering the symmetry of the underlying matrices this matrix equality implies $(n+1)(2n+1)$ scalar equations in the $2n+1$ dimensional unknown vector $\mathbf{h}$. However, it can be shown that the system has only $2n+1$ linearly independent equations, and the solution $\mathbf{h}$ is unique.

Remark 12. Conversely, suppose that $\mathbf{h}$ is given, and we have calculated $\mathbf{\Sigma} = \mathcal{L}(\mathbf{h})$ by solving the the Lyapunov equation (46). In addition it turns out that $\mathbf{\Sigma} \succeq 0$. Now we seek a positive function Φ such that (45) holds. In general, we have infinitely many Φ that will solve (45). Only when $\mathbf{\Sigma}$ is positive semidefinite the solution is unique.

Next result shows a direct connection between $\mathbf{h}$ and Φ.

Lemma 2. *Given Φ such that $\Phi(\mathrm{e}^{\mathrm{i}\omega}) \geq 0, \forall\omega$, compute the integral in the right hand side of (45) to obtain $\mathbf{\Sigma}$. Then $\mathbf{\Sigma} = \mathcal{L}\{\Phi_\ddagger(\mathbf{A})\mathbf{b}\}$. Consequently*

$$\mathbf{h} = \Phi_\ddagger(\mathbf{A})\mathbf{b}. \tag{47}$$

Conversely, if $\mathbf{\Sigma}$ is a non-negative definite matrix satisfying (46) for some $\mathbf{h}$, then every solution Φ to (45) satisfies (47).

Proof: The proof is available in [14] for a general setting. For a more direct (and easier) sequence of arguments see Appendix D.1. ■

Remark 13. Suppose that we are given a $\mathbf{h}$ satisfying $\mathcal{L}(\mathbf{h}) \succeq 0$. We are required to find $\Phi_\ddagger$ satisfying (47). This problem is a Nevanlinna-Pick interpolation problem. In particular $\mathbf{h}$ specifies the interpolation data at the eigenvalues of $\mathbf{A}$. This can be checked by transforming the pair $(\mathbf{A}, \mathbf{b})$ into Jordan canonical form representation. This has many interesting consequences in analyzing the role of input in system identification [29]. For instance, it turns out in linear system identification that the variance error is independent of the zeros of the system, and depends solely on the interpolation conditions specified by $\mathbf{h}$.

We need the following result to impose some constraint that the variance of $u(t)$ is ς.

Lemma 3. *If $\mathbf{h} = \Phi_\ddagger(\mathbf{A})\mathbf{b}$, then*

$$\mathbf{b}^\intercal\bar{\mathbf{C}}^{-1}\mathbf{h} = \frac{1}{4\pi}\int_{-\pi}^{\pi} \Phi(\mathrm{e}^{\mathrm{i}\omega})\ \mathrm{d}\omega = \varsigma/2.$$

Proof: See Appendix D.2. ■

Using Theorem 4 and Lemma 3 the following Lemma is immediate.

Lemma 4. *Given a positive definite matrix $\mathbf{\Sigma}$ there exists a power spectral density $\Phi(\mathrm{e}^{\mathrm{i}\omega})$ satisfying (45) and (32b) if and only if $\mathbf{\Sigma} = \mathcal{L}(\mathbf{h})$ for some $\mathbf{h} \in \mathbb{K}$.*

Now combining (24), Theorem 2 and Lemma 4 we can convert the infinite dimensional problem (32) into an equivalent finite dimensional problem in $\mathbf{h}$:

$$\begin{aligned} \underset{\sigma,\mathbf{h}}{\text{maximize}} \quad & f(\sigma)\det\{\mathcal{L}(\mathbf{h})\} && \text{(48a)} \\ \text{subject to} \quad & \sigma - \mathbf{c}^{\mathsf{T}}\mathcal{L}(\mathbf{h})\mathbf{c} = 0, && \text{(48b)} \\ & \mathbf{h} \in \mathbb{K}. && \text{(48c)} \end{aligned}$$

Clearly, due to constraints (48b) and (48c), and by definitions of $\sigma_{\min}$ and $\sigma_{\max}$ in (34) and (35), the feasible set of σ must be $[\sigma_{\min}, \sigma_{\max}]$. Now by the definition of χ in (37), the maximum possible value of the cost function (48a) for any given $\sigma \in [\sigma_{\min}, \sigma_{\max}]$ is $f(\sigma)\chi(\sigma)$. Hence the D-optimal problem (48) can be solved by solving the line search (36).

If σ_* denotes that optimum value of $\sigma \in [\sigma_{\min}, \sigma_{\max}]$, then the above argument also implies that the optimum value of $\mathbf{h}$ in (48) is $\mathbf{h}_*(\sigma_*)$. Consequently, Lemma 2 implies (38). This completes the proof of Theorem 3.

5 Numerical simulation results

5.1 Verification of the accuracy results

In this section we present some simulation results to illustrate the results presented above. First we verify the result in Theorem 1. We consider a Wiener system with

$$G(z, \mathring{\boldsymbol{\theta}}) = \frac{0.1z^{-1} + 0.05z^{-2}}{1 - 1.5z^{-1} + 0.7z^{-2}}, \tag{49}$$

and the polynomial non-linearity is taken as

$$\wp(x, \mathring{\boldsymbol{\alpha}}) = x - 0.5x^3. \tag{50}$$

For this system $m = 3$ and $n = 2$. Hence $\boldsymbol{\vartheta}$ is of dimension $2n + 1 + m = 8$. We take

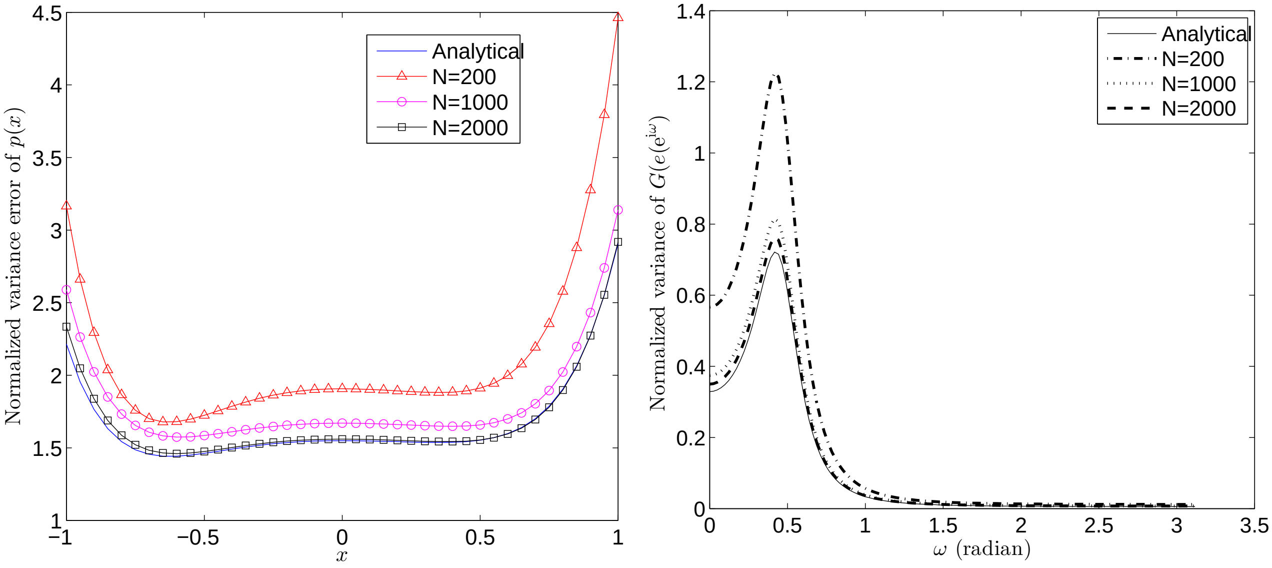
$$\boldsymbol{\upsilon} = [\begin{array}{cccc} 0 & 1 & 0 & 0 \end{array}]^{\mathsf{T}}. \tag{51}$$

Verify that (10) holds. For this choice of $\boldsymbol{\upsilon}$ we must take $\ell = 1$, see Assumption 1. We excite the system with Gaussian input process with mean $\eta_u = -0.5$ and a power spectral density

$$\Phi(\mathrm{e}^{\mathrm{i}\omega}) = \frac{(1 + \mathrm{e}^{\mathrm{i}\omega})(1 + \mathrm{e}^{-\mathrm{i}\omega})}{(1 - 0.7\mathrm{e}^{\mathrm{i}\omega})(1 - 0.7\mathrm{e}^{-\mathrm{i}\omega})}.$$

The variance of the additive, zero mean white measurement noise is set to 1.

In the following $\hat{\boldsymbol{\vartheta}}_N$ denotes the PEM estimate of $\mathring{\boldsymbol{\vartheta}}$ obtained from N samples of input-output data. We plot the results obtained from the Monte-Carlo simulation in Figure 1, where we compare the normalized analytical mean squared errors of $G(\mathrm{e}^{\mathrm{i}\omega}, \hat{\boldsymbol{\vartheta}}_N)$ and $\wp(x, \hat{\boldsymbol{\vartheta}}_N)$ predicted by Theorem 1 with that obtained empirically from Monte-Carlo simulations for three different values of N. For each value of N the empirical variance plots are obtained by averaging the results of 1000 independent simulations. As can be seen in Figure 1 the analytical predictions are very well in agreement with the numerical findings for $N \geq 1000$.



(a) Variance of $\wp(x, \hat{\boldsymbol{\vartheta}}_N)$ as functions of x

(b) Variance of $G(\mathrm{e}^{\mathrm{i}\omega}, \hat{\boldsymbol{\vartheta}}_N)$ as functions of ω

Figure 1: The comparison between the normalized analytical variance obtained from Theorem 1 and the empirical variance obtained from Monte-Carlo simulations.

5.2 Max-det design

In this section we present a numerical example illustrating the max-det design approach outlined in Section 4. For this purpose we consider identification of the Wiener system given in (49) and (50). For this example

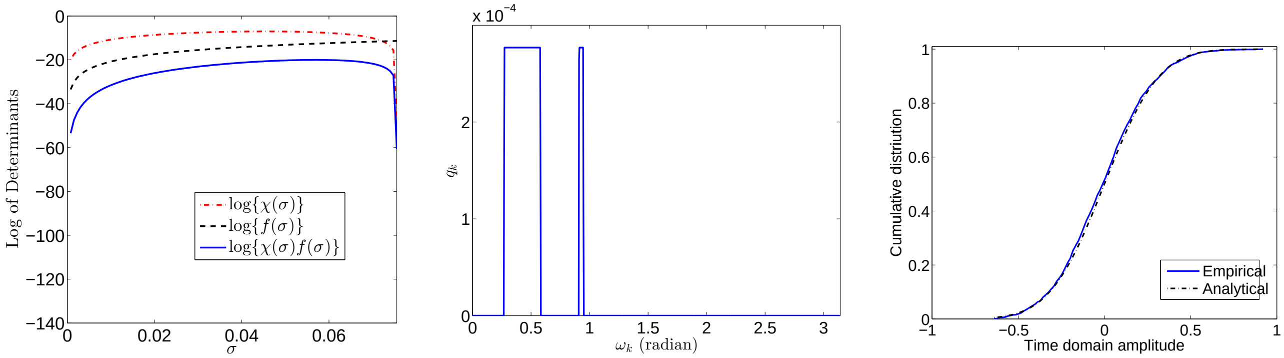
$$\mathbf{A} = \begin{bmatrix} 1.5 & -0.7 & -0.1 & -0.05 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -0.7 & 0 & 1.5 \end{bmatrix}.$$

We take $\eta_u = -0.13$, and $\varsigma = 0.0618$. For this value of ς it is ensured that $|u(t)| \leq 1$ with a probability more that 0.999. By solving (34) and (35) we get $\sigma_{\min} = 1.51 \times 10^{-5}$, and $\sigma_{\max} = 0.0755$. In Figure 2a we plot $\log\{\chi(\sigma)\}, \log\{f(\sigma)\}$ and $\log\{\chi(\sigma)f(\sigma)\}$ as functions of σ. Note that in this case the logarithm of the cost function in (36) is concave, and it has a unique maximum point. The value of σ_* is 0.0569, and

$$\mathbf{h}_*(\sigma_*) = [\ 0.2005 \ \ 0.2201 \ \ 0.1085 \ \ -0.0321 \ \ 0.2161\ ]^\intercal.$$

Corresponding to this value there are infinitely many choices of Φ satisfying (38). We can construct a rational Φ satifying (38), for instance via spectral zero assignment [5]. For instance, if we like all minimum phase zeros of Φ to be located at the origin then we get

$$\Phi(\mathrm{e}^{\mathrm{i}\omega}) = \frac{0.170^2}{|1 - 2.9\mathrm{e}^{\mathrm{i}\omega} + 3.8\mathrm{e}^{2\mathrm{i}\omega} - 2.4\mathrm{e}^{3\mathrm{i}\omega} + 0.7\mathrm{e}^{4\mathrm{i}\omega}|^2}.$$



(a) (b) (c)

Figure 2: Numerical illustration of the 3-step algorithm for the system (49)-(50). (a) The plots of $\log\{\chi(\sigma)\}, \log\{f(\sigma)\}$ and $\log\{\chi(\sigma)f(\sigma)\}$ as functions of σ in the range $[\sigma_{\min}, \sigma_{\max}]$ (b) The plot of q_k as a function of $\omega_k,\ k = 1, 2, \ldots, M$ obtained by solving (44). (c) The empirical cumulative amplitude distribution of a realization of the form (43) of the optimal amplitude spectrum shown in Figure 2b (in solid line) is compared with the corresponding Gaussian cumulative distribution function (in dotted line). The empirical plot is hardly distinguishable from the plot of Gaussian CDF.

For another example, if we like the minimum phase zeros of Φ to be located at $0.9\mathrm{e}^{\pm \mathrm{i}\pi/2}$ and $0.9\mathrm{e}^{\pm \mathrm{i}5\pi/6}$ then we get

$$\Phi(\mathrm{e}^{\mathrm{i}\omega}) = \frac{|2.8 - 4.4\mathrm{e}^{\mathrm{i}\omega} + 4.5\mathrm{e}^{2\mathrm{i}\omega} + 3.5\mathrm{e}^{3\mathrm{i}\omega} + 1.8\mathrm{e}^{4\mathrm{i}\omega}|^2}{|1 - 3.1\mathrm{e}^{\mathrm{i}\omega} + 4.1\mathrm{e}^{2\mathrm{i}\omega} - 2.7\mathrm{e}^{3\mathrm{i}\omega} + 0.7\mathrm{e}^{4\mathrm{i}\omega}|^2} \\ \times 10^{-6}.$$

On the other hand, if we use the algorithm (44) to design a multisine amplitude spectrum, then we get a solution plotted in Figure 2b. Here we have taken $M = 1000$, and $\omega_k = k\pi/(M+1),\ k = 1, 2, \ldots, M$ (note that the domain is discrete, not continuous, and the plot in Figure 2b should not be interpreted as the power spectral density). We synthesize a time domain signal consistent with this solution as in (43), and plot its empirical cumulative amplitude distribution function in Figure 2c. We compare this empirical distriution with the cumulative distribution function of a Gaussian density with zero mean and variance $\varsigma = 0.0618$. Note that the empirical distribution is very well in accordance with Gaussian distribution.

5.3 Gaussian mixture design

In this section we illustrate the Gaussian mixture design approach. We consider the simple example given in [6], where the global optimal input is obtained for identification of a Wiener system with

$$G(z, \mathring{\boldsymbol{\theta}}) = 3 + z^{-1},$$

and the polynomial non-linearity is taken as

$$\wp(x, \mathring{\boldsymbol{\alpha}}) = -0.25x + x^3.$$

For this system $m = 3$ and $n = 1$. Hence $\boldsymbol{\vartheta}$ is of dimension $n + 1 + m = 5$. We take

$$\boldsymbol{\upsilon} = [\, 0 \quad -4 \ \ 0 \ \ 0 \,]^{\intercal}.$$

Verify that (10) holds. For this choice of $\boldsymbol{\upsilon}$ we must take $\ell = 1$, see Assumption 1.

We consider a Gaussian mixture design approach described above. As above, we use p to denote the number of Gaussian considered. We do four different experiments for $p = 1, 5, 11, 19$ respectively. For a given p, we set the means $\eta_u(k)$ of the p Gaussians as in (2), and set the input variance ς_k as in (3), which depends on the parameter κ. Recall that by allowing a smaller κ we increase the input variance, and increase the probability of violating the constraint $|u(t)| < 1$. For each p we do several sub-experiments, while in each sub-experiment we set a new value for κ in the range 0.5 to 5.

Now consider a particular sub-experiment with fixed p and κ. For $k = 1, 2, \ldots,\ p$ we calculate $\eta_u(k)$ and ς_k using (2) and (3). For each k we calculate $\mathbf{J}_k$ by solving the D-optimal design problem in (48) by setting $\eta_u = \eta_u(k)$ and $\varsigma = \varsigma_k$. Subsequently, we find out the numbers $\kappa_1, \ldots, \kappa_p$, see (4b), by solving

$$\begin{aligned} \underset{\kappa_1,\kappa_2,\ldots,\kappa_p}{\text{maximize}} \quad & \det(\kappa_1\mathbf{J}_1 + \kappa_2\mathbf{J}_2 + \cdots + \kappa_p\mathbf{J}_p) \\ \text{subject to} \quad & \kappa_k \geq 0, \quad k = 1, 2, \ldots, p, \\ & \kappa_1 + \kappa_2 + \cdots + \kappa_p = 1. \end{aligned}$$

Because the function $-\log\{\det(\cdot)\}$ is convex, the above optimization problem is convex, and can be solved using the dispersion method, see for example [6, 40].

The above procedure yields an optimal Gaussian mixture design. Next we test the optimal design via Monte-Carlo simulations. Recall that for each k the D-optimal design process gives an optimal $\mathbf{h}$ along with the optimal $\mathbf{J}_k$, see (48). We use this optimal $\mathbf{h}$ to generate an associated Φ by solving the Nevanlinna-Pick interpolation problem, see Section 4.1.2. Due to Gaussian assumption the realizations of input u consistent with Φ may fail to satisfy the constraint $|u(t)| < 1$. Hence we clip the input as

$$u(t) = \max\{\min\{u(t), 1\}, -1\}.$$

Therefore the input signal is always restricted in the interval $[-1, 1]$. However, this approach violates the Gaussian assumption to an extent that depends on the value of κ. The experimental results are plotted in Figure 3, where for each p we compare the analytically predicted value of $\det(\bar{\mathbf{J}})$ with the corresponding empirical value obtained from Monte-Carlo simulations after clipping of the optimal input. As can be seen in Figure 3, the empirical results agree very well with the analytical predictions for $\kappa > 3$. However, the empirical curves diverge from their analytical counterparts for smaller values of κ. This divergence phenomenon for small κ is expected because the analytical design allows some significant probability that $|u(t)| > 1$, while the Monte-Carlo simulations apply clipping to restrict $|u(t)| \leq 1$. It is also interesting to compare the performance obtained by the single Gaussian design with that obtained by multiple Gaussian inputs and mixing them optimally. Note that multiple Gaussian design can achieve significant performance boost for medium and large κ values. The best performance is achieved for $\kappa = 0.5$, for which $\det(\bar{\mathbf{J}})$ obtained from our method was found about 8 times smaller than that obtained via the deterministic design approach presented in [6]. Given that there are 5 parameters being estimated, this performance loss in determinant is rather insignificant.

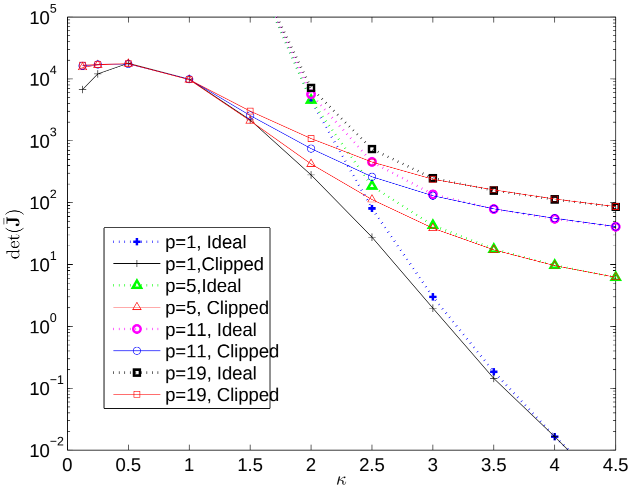


Figure 3: The optimal determinant of $\bar{\mathbf{J}}$ plotted as a function of κ for the FIR model considered in the numerical simulations.

We have also computed the information matrix for the system corresponding to a binary white excitation taking its values in $\{-1, 1\}$ using a Monte-Carlo method. The value of $\det(\mathbf{J})$ in this case turns out to be 1.2392. With a binary process with passband $[-\pi/2, \pi/2]$ the value of $\det(\mathbf{J})$ turns out to be 0.7008.

6 Conclusions

We have presented several new results on the analysis of Wiener model identification using Gaussian input processes. One of the main results in this paper is Theorem 1, which gives a closed form expression of the associated information matrix $\mathbf{J}$. This expression holds under very generic assumptions on the model structure. In addition, unlike other similar formulae available in the literature, our expression for $\mathbf{J}$ is easy to compute. This aspect makes it attractive in input design problems. Theorem 2 gives a simple expression for the determinant of $\mathbf{J}$.

The second major contribution of this paper is to show that the set of all admissible information matrices can be completely parameterized via the finite dimensional parameter vector $\mathbf{h}$. Given an admissible $\mathbf{h}$ there are potentially infinitely many input power spectral densities Φ leading to the same $\mathbf{h}$, and consequently the same $\mathbf{J}$. These input power spectral densities, consistent with a particular $\mathbf{h}$, are characterized by some Nevanlinna-Pick interpolation problem. We reemphasize that the fallibility of having non-unique Φ for a particular $\mathbf{J}$ of our liking can be exploited further. We have the possibility of choosing a Φ of some desirable shape without changing $\mathbf{J}$. This observation might be useful in robust input design.

The accuracy analysis leads to the concept of w-dependence in Definition 1. This concept, which is due to the fact that the estimation accuracy of the static nonlinearity depends only on the amplitude distribution of $w(t)$, plays a vital role in deriving tractable algorithms for D-optimal design. In particular, the expression of $\det \mathbf{J}$ given by Theorem 2 shows that $\det \mathbf{J}$ can be factored in two factors. The first factor is w-dependent. The other term is statistically independent of w. This observation is exploited to derive a tractable algorithm to solve the D-optimal input design problem.

Another main contribution of this paper is to introduce the idea of Gaussian mixture design. This approach allows us to control the effective amplitude distribution and the power spectral density of the excitation at the same time. While the power spectral density controls the accuracy of identifying the linear subsystem, the amplitude distribution determines how well we identify the static nonlinearity.

We have presented some simulation results in support of the claims made in the paper. We found that by using Gaussian input we can indeed achieve an estimation accuracy that is quite close to the global optimum achieved by the deterministic design techniques. While the deterministic design technique can be applied only for FIR Wiener systems with small order, Gaussian design technique can be applied to IIR Wiener systems of any finite order, and finite degree of the static nonlinearity.

References

[1] M. Barenthin, X. Bombois, H. Hjalmarsson, and G. Scorletti. Identification for control of multivariable systems: Controller validation and experiment design via LMIs. *Automatica*, 44:3070–3078, Dec 2008.

[2] M. Barenthin and H. Hjalmarsson. Identification and control: Joint input design and $\mathcal{H}_\infty$ state feedback with ellipsoidal parametric uncertainty via LMIs. *Automatica*, 44(2):543–551, Feb 2008.

[3] X. Bombois, G. Scorletti, M. Gevers, P. M. J. Van den Hof, and R. Hildebrand. Least costly identification experiment for control. *Automatica*, 42:3:1651–1662, 2006.

[4] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge Univ. Press, Cambridge, U.K., 2004.

[5] C. I. Byrnes, T. T. Georgiou, and A. Lindquist. A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint. *IEEE Transactions on Automatic Control*, 46:6:822–839, June 2001.

[6] A. De Cock, M. Gevers, and J. Schoukens. A preliminary study on optimal input design for nonlinear systems. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pages 4931–4936, Dec 2013.

[7] A. De Cock, M. Gevers, and J. Schoukens. D-optimal input design for FIR-type nonlinear systems: A dispersion based approach. 2014. Submitted for publication.

[8] A. Ferrante, M. Pavon, and F. Ramponi. Hellinger vs. Kullback-Leibler multivariable spectrum approximation. *IEEE Transactions on Automatic Control*, 53:5:954–967, May 2008.

[9] A. Ferrante, F. Ramponi, and F. Ticozzi. On the convergence of an efficient algorithm for Kullback-Leibler approximation of spectral densities. *IEEE Transactions on Automatic Control*, 56:3:506–515, March 2011.

[10] M. Forgione, X. Bombois, P. M. J. Van den Hof, and H. Hjalmarsson. Experiment design for parameter estimation in nonlinear systems based on multilevel excitation. In *European Control Conference*, 2014.

[11] T. Georgiou and A. Lindquist. A convex optimization approach to ARMA modeling. *IEEE Transactions Automatic Control*, 53:5:1108–1119, June 2008.

[12] T. T. Georgiou. The interpolation problem with a degree constraint. *IEEE Transactions on Automatic Control*, 44:3:631–635, March 1999.

[13] T. T. Georgiou. Spectral estimation via selective harmonic amplification. *IEEE Transactions on Automatic Control*, 46:1:29–42, January 2001.

[14] T. T. Georgiou. The structure of state covariances and its relation to the power spectrum of the input. *IEEE Transactions on Automatic Control*, 47:7:1056–1066, July 2002.

[15] T. T. Georgiou and A. Lindquist. Kullback-Leibler approximation of spectral density functions. *IEEE Transactions on Information Theory*, 49:11:2910–2917, November 2003.

[16] M. Gevers, M. Caenepeel, and J. Schoukens. Experiment design for the identification of a simple Wiener system. In *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pages 7333–7338, Dec 2012.

[17] M. Gevers, L. Ljung, and P. M. J. Van den Hof. Asymptotic variance expressions for closed loop identification. *Automatica*, 37:781–786, 2001.

[18] G. C. Goodwin and R. L. Payne. *Dynamic System Identification: Experiment design and Data Analysis.* Academic Press, 1977.

[19] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008. `http://stanford.edu/~boyd/graph_dcp.html`.

[20] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, March 2014.

[21] R. Hildebrand and M. Gevers. Identification for control: optimal input design with respect to a worst-case ν-gap cost function. *SIAM journal of control and optimization*, 41:5:1586–1608, 2003.

[22] H. Hjalmarsson and H. Jansson. Closed loop experiment design for linear time invariant dynamical systems via LMIs. *Automatica*, 44(3):623–636, Mar 2008.

[23] H. Hjalmarsson and J. Mårtensson. Optimal input design for identification of non-linear systems: Learning from the linear case. In *American Control Conference, 2007. ACC '07*, pages 1572–1576, July 2007.

[24] H. Jansson and H. Hjalmarsson. Input design via LMIs admitting frequency-wise model specifications in confidence regions. *IEEE Transactions on Automatic Control*, 50(10):1534–1549, Oct 2005.

[25] R. Kan. From moments of sum to moments of product. *Journal of multivariate analysis*, 99:542–554, 2008.

[26] C. A. Larsson, H. Hjalmarsson, and C. R. Rojas. On optimal input design for nonlinear FIR-type systems. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 7220–7225, Dec 2010.

[27] L. Ljung. Asymptotic variance expressions for identified black-box transfer function models. *IEEE Transactions on Automatic Control*, AC-30:834–844, 1985.

[28] L. Ljung. *System Identification - Theory for the User, 2nd edition.* Prentice Hall, Upper Saddle River, NJ, USA, 1999.

[29] K. Mahata. Variance error, interpolation and experiment design. *Automatica*, 49:5:1117–1125, 2013.

[30] B. Ninness and H. Hjalmarsson. Variance error quantifications that are exact for finite model order. *IEEE Transactions on Automatic Control*, 49:1275–1291, 2004.

[31] B. Ninness and H. Hjalmarsson. On the frequency domain accuracy of closed loop estimates. *Automatica*, 41:1109–1122, 2005.

[32] A. Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, New York, 1991.

[33] M. Pavon and A. Ferrante. On the Georgiou-Lindquist approach to constrained Kullback-Leibler approximation of spectral densities. *IEEE Transactions on Automatic Control*, 51:4:639–644, April 2006.

[34] M. Pavon and A. Ferrante. A new algorithm for Kullback-Leibler approximation of spectral densities. In *44th IEEE Conference on Decision and Control*, Sevile, Spain, December, 2005.

[35] R. Pintelon and J. Schoukens. *System Identification: A Frequency Domain Approach.* Wiley-IEEE Press, 2012.

[36] F. Ramponi, A. Ferrante, and M. Pavon. A globally convergent matricial algorithm for multivariate spectral estimation. *IEEE Transactions on Automatic Control*, 54:10:2376–2388, October 2009.

[37] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999. Version 1.05 available from `http://fewcal.kub.nl/sturm`.

[38] P. E. Valenzuela, C. R. Rojas, and H. Hjalmarsson. Optimal input design for non-linear dynamic systems: A graph theory approach. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pages 5740–5745, Dec 2013.

[39] L. L. Xie and L. Ljung. Asymptotic variance expressions for estimated frequency functions. *IEEE Transactions on Automatic Control*, 46:1887–1899, 2001.

[40] Y. Yu. Monotonic convergence of a general algorithm for computing optimal designs. *The Annals of Statistics*, 38:1593–1606, 2010.

A Proof of Lemma 1

By definition of $\mathbf{P}$ in (13) we have $\mathbf{P}\mathbf{P}^\intercal = \mathbf{I}$. Using this in (14) gives

$$\bar{\boldsymbol{\alpha}} = \mathbf{P}^\intercal \mathbf{L}^\intercal \boldsymbol{\alpha}. \tag{52}$$

Using (52) and the definition of $\mathbf{z}(t, \boldsymbol{\theta})$ in (20) in (8) we have

$$M(\boldsymbol{\vartheta}, \mathbf{u}_t) = \bar{\boldsymbol{\alpha}}^\intercal \mathbf{z}(t, \boldsymbol{\theta}) = \boldsymbol{\alpha}^\intercal \mathbf{L}\mathbf{P}\mathbf{z}(t, \boldsymbol{\theta}). \tag{53}$$

Hence

$$\frac{\partial M(\mathring{\boldsymbol{\vartheta}}, \mathbf{u}_t)}{\partial \boldsymbol{\alpha}} = \mathbf{L}\mathbf{P}\mathbf{z}(t, \mathring{\boldsymbol{\theta}}). \tag{54}$$

Also using the definition of $\mathbf{z}(t, \boldsymbol{\theta})$ in (20) and differentiating $M(t, \boldsymbol{\vartheta})$ in (53) with respect to $\boldsymbol{\theta}$ we get

$$\begin{aligned} \frac{\partial M(\mathring{\boldsymbol{\vartheta}}, \mathbf{u}_t)}{\partial \boldsymbol{\theta}} &= \frac{\partial w(t, \mathring{\boldsymbol{\vartheta}})}{\partial \boldsymbol{\theta}} \, \mathring{\boldsymbol{\alpha}}^\intercal \mathbf{L}\mathbf{P} \begin{bmatrix} 0 \\ 1 \\ 2w(t, \mathring{\boldsymbol{\theta}}) \\ \vdots \\ m\{w(t, \mathring{\boldsymbol{\theta}})\}^{m-1} \end{bmatrix} \\ &= \frac{\partial w(t, \mathring{\boldsymbol{\vartheta}})}{\partial \boldsymbol{\theta}} \, \boldsymbol{\alpha}_2^\intercal \mathbf{z}(t, \mathring{\boldsymbol{\theta}}), \end{aligned} \tag{55}$$

where the last equality follows from the definition of $\boldsymbol{\alpha}_2$ in (21) and the definition of $\mathbf{z}(t, \boldsymbol{\theta})$ in (20). The proof for the expression of $\mathbf{v}_t$ in (19) will be complete if we can show

$$\frac{\partial w(t, \mathring{\boldsymbol{\vartheta}})}{\partial \boldsymbol{\theta}} = \frac{\partial G(z, \mathring{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} u(t) = \mathbf{x}(t). \tag{56}$$

This is done next. By differentiating G with respect to $\mathbf{a}$, $\mathbf{g}$ and g_0 we get

$$\begin{aligned} \frac{\partial G(z, \boldsymbol{\theta})}{\partial \mathbf{a}} &= -(z\mathbf{I} - \mathbf{A}_1)^{-1}\mathbf{b}_1 g_0 \\ &\quad - (z\mathbf{I} - \mathbf{A}_1)^{-1}\mathbf{b}_1(\mathbf{g} - \mathbf{a}g_0)^\intercal (z\mathbf{I} - \mathbf{A}_1)^{-1}\mathbf{b}_1, \qquad (57\text{a}) \\ \frac{\partial G(z, \boldsymbol{\theta})}{\partial \mathbf{g}} &= (z\mathbf{I} - \mathbf{A}_1)^{-1}\mathbf{b}_1, \qquad (57\text{b}) \\ \frac{\partial G(z, \boldsymbol{\theta})}{\partial g_0} &= 1 - \mathbf{a}^\intercal (z\mathbf{I} - \mathbf{A}_1)^{-1}\mathbf{b}_1. \qquad (57\text{c}) \end{aligned}$$

Using (57) and (17) it can be verified by direct calculations that

$$\frac{\partial G(z,\mathring{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} = (\mathbf{I} - \mathbf{A}z^{-1})^{-1}\mathbf{b}, \tag{58}$$

implying (56).

To show $w(t,\mathring{\boldsymbol{\theta}}) = \mathbf{c}^{\mathsf{T}}\mathbf{x}(t)$ verify from (15) and (57) that

$$G(z,\mathring{\boldsymbol{\theta}}) = [\ \mathbf{0}^{\mathsf{T}}\ \ \mathring{\mathbf{g}}^{\mathsf{T}}\ \ \mathring{g}_0\ ]\frac{\partial G(z,\mathring{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} = \mathbf{c}^{\mathsf{T}}(\mathbf{I} - \mathbf{A}z^{-1})^{-1}\mathbf{b}.$$

B Proof of Theorem 1

Since $\boldsymbol{\Sigma}$ is positive definite, $\boldsymbol{\Sigma}\mathbf{c} \neq 0$. Hence there exists a full column rank $(2n+1)\times(2n)$ matrix $\mathbf{C}$ such that the column space of $\mathbf{C}$ is the orthogonal complement of $\boldsymbol{\Sigma}\mathbf{c}$, i.e., $\mathbf{C}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{c} = 0$. Hence

$$\begin{bmatrix} \mathbf{c}^{\mathsf{T}} \\ \mathbf{C}^{\mathsf{T}} \end{bmatrix} \boldsymbol{\Sigma}\ [\ \mathbf{c}\ \ \mathbf{C}\ ] = \begin{bmatrix} \sigma & 0 \\ 0 & \boldsymbol{\Sigma}_1 \end{bmatrix}, \tag{59}$$

The block diagonal structure of the matrix in the right hand side of (59) ensures that by premultiplying $\mathbf{x}$ by $[\ \mathbf{c}\ \ \mathbf{C}\ ]^{\mathsf{T}}$ we get two mutually uncorrelated components $\mathbf{c}^{\mathsf{T}}\mathbf{x}$ and

$$\mathbf{x}_1 := \mathbf{C}^{\mathsf{T}}\mathbf{x},$$

with

$$\begin{aligned} \boldsymbol{\gamma} &:= \mathsf{E}\{\mathbf{x}_1\} = \mathbf{C}^{\mathsf{T}}\boldsymbol{\eta}, \\ \boldsymbol{\Sigma}_1 &:= \mathsf{E}\{[\mathbf{x}_1 - \boldsymbol{\gamma}][\mathbf{x}_1 - \boldsymbol{\gamma}]^{\mathsf{T}}\} = \mathbf{C}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{C}. \end{aligned} \tag{60}$$

Because $\mathbf{x}$ is a Gaussian random vector, we conclude that $[\ \mathbf{c}^{\mathsf{T}}\mathbf{x}\ \ \mathbf{x}_1^{\mathsf{T}}\ ]^{\mathsf{T}}$ too is a jointly Gaussian random vector. Since uncorrelated Gaussian variables are independent, $\mathbf{c}^{\mathsf{T}}\mathbf{x}$ and $\mathbf{x}_1$ are mutually independent.

Define the $(m+2n+1)\times(m+2n+1)$ matrix

$$\mathbf{T} = \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{c}^{\mathsf{T}} \\ 0 & \mathbf{C}^{\mathsf{T}} \end{bmatrix}, \tag{61}$$

where the identity matrix appearing in (61) in the north-west corner is of size $m \times m$. Premultiplying $\mathbf{v}_t$ in (20) by $\mathbf{T}$ we note that

$$\mathbf{T}\mathbf{v}_t = \begin{bmatrix} \mathbf{L}_1\mathbf{z}(t,\mathring{\boldsymbol{\theta}}) \\ \mathbf{c}^{\mathsf{T}}\mathbf{x}(t)\boldsymbol{\alpha}_2^{\mathsf{T}}\mathbf{z}(t,\mathring{\boldsymbol{\theta}}) \\ \mathbf{C}^{\mathsf{T}}\mathbf{x}(t)\boldsymbol{\alpha}_2^{\mathsf{T}}\mathbf{z}(t,\mathring{\boldsymbol{\theta}}) \end{bmatrix}. \tag{62}$$

From Lemma 1 recall that $\mathbf{c}^{\mathsf{T}}\mathbf{x}(t) = w(t,\mathring{\boldsymbol{\theta}})$. Then from the definition of $\mathbf{z}(t,\boldsymbol{\theta})$ in (20), the definitions $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ in (26) and (21) we have

$$\mathbf{c}^{\mathsf{T}}\mathbf{x}(t)\boldsymbol{\alpha}_2^{\mathsf{T}}\mathbf{z}(t,\mathring{\boldsymbol{\theta}}) = w(t,\mathring{\boldsymbol{\theta}})\boldsymbol{\alpha}_1^{\mathsf{T}}\mathbf{z}(t,\mathring{\boldsymbol{\theta}}).$$

In addition, $\mathbf{C}^{\mathsf{T}}\mathbf{x}(t) = \mathbf{x}_1(t)$. Hence

$$\mathbf{T}\mathbf{v}_t = \begin{bmatrix} \mathbf{L}_1\mathbf{z}(t,\mathring{\boldsymbol{\theta}}) \\ \boldsymbol{\alpha}_1^{\mathsf{T}}\mathbf{z}(t,\mathring{\boldsymbol{\theta}}) \\ \mathbf{x}_1(t)\boldsymbol{\alpha}_2^{\mathsf{T}}\mathbf{z}(t,\mathring{\boldsymbol{\theta}}) \end{bmatrix}. \tag{63}$$

Since $\mathbf{x}_1(t)$ is independent of $w(t,\mathring{\boldsymbol{\theta}}) = \mathbf{c}^{\mathsf{T}}\mathbf{x}(t)$, it is also independent of $\mathbf{z}(t,\mathring{\boldsymbol{\theta}})$, see (20). Using this and (63) we get

$$\begin{aligned} \mathbf{T}\mathbf{J}\mathbf{T}^{\mathsf{T}} &= \mathsf{E}\left\{[\mathbf{T}\mathbf{v}_t]\,[\mathbf{T}\mathbf{v}_t]^{\mathsf{T}}\right\} \\ &= \begin{bmatrix} \mathbf{L}_1\boldsymbol{\Lambda}\mathbf{L}_1^{\mathsf{T}} & \mathbf{L}_1\boldsymbol{\Lambda}\boldsymbol{\alpha}_1 & \mathbf{L}_1\boldsymbol{\Lambda}\boldsymbol{\alpha}_2\boldsymbol{\gamma}^{\mathsf{T}} \\ \boldsymbol{\alpha}_1^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{L}_1^{\mathsf{T}} & \boldsymbol{\alpha}_1^{\mathsf{T}}\boldsymbol{\Lambda}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_1^{\mathsf{T}}\boldsymbol{\Lambda}\boldsymbol{\alpha}_2\boldsymbol{\gamma}^{\mathsf{T}} \\ \boldsymbol{\gamma}\boldsymbol{\alpha}_2^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{L}_1^{\mathsf{T}} & \boldsymbol{\gamma}\boldsymbol{\alpha}_2^{\mathsf{T}}\boldsymbol{\Lambda}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2^{\mathsf{T}}\boldsymbol{\Lambda}\boldsymbol{\alpha}_2(\boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathsf{T}} + \boldsymbol{\Sigma}_1) \end{bmatrix}. \end{aligned} \tag{64}$$

Define the vector $\mathbf{d}$ and the $(2n+1)\times(2n)$ matrix $\mathbf{D}$ by partitioning the inverse

$$\begin{bmatrix} \mathbf{c}^{\mathsf{T}} \\ \mathbf{C}^{\mathsf{T}} \end{bmatrix}^{-1} = [\ \mathbf{d}\ \ \mathbf{D}\ ]. \tag{65}$$

Then (61) and (64) imply

$$\mathbf{J} = \begin{bmatrix} \mathbf{I} & 0 & 0 \\ 0 & \mathbf{d} & \mathbf{D} \end{bmatrix} (\mathbf{T}\mathbf{J}\mathbf{T}^{\mathsf{T}}) \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{d}^{\mathsf{T}} \\ 0 & \mathbf{D}^{\mathsf{T}} \end{bmatrix}.$$

Using expression of $\mathbf{T}\mathbf{J}\mathbf{T}^{\mathsf{T}}$ in (64) we get

$$\mathbf{J}_{11} = \mathbf{L}_1\boldsymbol{\Lambda}\mathbf{L}_1^{\mathsf{T}}, \tag{66a}$$

$$\mathbf{J}_{21} = [\ \mathbf{d}\ \ \mathbf{D}\boldsymbol{\gamma}\ ]\ \mathbf{L}_2\boldsymbol{\Lambda}\mathbf{L}_1^{\mathsf{T}} \tag{66b}$$

$$\mathbf{J}_{22} = [\ \mathbf{d}\ \ \mathbf{D}\boldsymbol{\gamma}\ ]\ \mathbf{L}_2\boldsymbol{\Lambda}\mathbf{L}_2^{\mathsf{T}}\ [\ \mathbf{d}\ \ \mathbf{D}\boldsymbol{\gamma}\ ]^{\mathsf{T}} + \beta\mathbf{D}\boldsymbol{\Sigma}_1\mathbf{D}^{\mathsf{T}}. \tag{66c}$$

Now from (59) and (65) we obtain

$$\boldsymbol{\Sigma} = [\ \mathbf{d}\ \ \mathbf{D}\ ] \begin{bmatrix} \sigma & 0 \\ 0 & \boldsymbol{\Sigma}_1 \end{bmatrix} \begin{bmatrix} \mathbf{d}^{\mathsf{T}} \\ \mathbf{D}^{\mathsf{T}} \end{bmatrix} = \mathbf{d}\sigma\mathbf{d}^{\mathsf{T}} + \mathbf{D}\boldsymbol{\Sigma}_1\mathbf{D}^{\mathsf{T}}. \tag{67}$$

By definition of $\mathbf{d}$ and $\mathbf{D}$ in (65) we know

$$\begin{bmatrix} \mathbf{c}^{\mathsf{T}} \\ \mathbf{C}^{\mathsf{T}} \end{bmatrix} [\ \mathbf{d}\ \ \mathbf{D}\ ] = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{I} \end{bmatrix},$$

and this implies $\mathbf{C}^{\mathsf{T}}\mathbf{d} = 0, \quad \Rightarrow \mathbf{d} = k\boldsymbol{\Sigma}\mathbf{c}$. In addition, $1 = \mathbf{c}^{\mathsf{T}}\mathbf{d} = k\mathbf{c}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{c} = k\sigma$. Consequently,

$$\mathbf{d} = \frac{1}{\sigma}\boldsymbol{\Sigma}\mathbf{c}. \tag{68}$$

On the other hand

$$\mathbf{I} = [\ \mathbf{d}\ \ \mathbf{D}\ ] \begin{bmatrix} \mathbf{c}^{\mathsf{T}} \\ \mathbf{C}^{\mathsf{T}} \end{bmatrix} = \mathbf{d}\mathbf{c}^{\mathsf{T}} + \mathbf{D}\mathbf{C}^{\mathsf{T}} = \frac{1}{\sigma}\boldsymbol{\Sigma}\mathbf{c}\mathbf{c}^{\mathsf{T}} + \mathbf{D}\mathbf{C}^{\mathsf{T}}, \tag{69}$$

Now multiply both sides of (69) by $\boldsymbol{\eta}$ to get

$$\boldsymbol{\eta} - \frac{\gamma}{\sigma}\mathbf{\Sigma c} = \mathbf{D}\boldsymbol{\gamma} \tag{70}$$

From (68) and (70) it follows that

$$[\ \mathbf{d}\ \ \mathbf{D}\boldsymbol{\gamma}\ ] = \mathbf{FQ}.$$

Now we use (67), (68), and (70) in (66) to eliminate $\mathbf{d}$ and $\mathbf{D}$ from the expressions of $\mathbf{J}_{12}$ and $\mathbf{J}_{22}$. We have

$$\begin{aligned}
\mathbf{J}_{21} &= \mathbf{FQL}_2\mathbf{\Lambda L}_1^\intercal \\
\mathbf{J}_{22} &= \mathbf{FQL}_2\mathbf{\Lambda L}_2^\intercal\mathbf{Q}^\intercal\mathbf{F}^\intercal + \beta(\mathbf{\Sigma} - \mathbf{\Sigma cc}^\intercal\mathbf{\Sigma}/\sigma) \\
&= \beta\mathbf{\Sigma} + \mathbf{FQL}_2\mathbf{\Lambda L}_2^\intercal\mathbf{Q}^\intercal\mathbf{F}^\intercal - \mathbf{FQHQ}^\intercal\mathbf{F}^\intercal \\
&= \mathbf{FQ}(\mathbf{L}_2\mathbf{\Lambda L}_2^\intercal - \mathbf{H})\mathbf{Q}^\intercal\mathbf{F}^\intercal + \beta\mathbf{\Sigma}.
\end{aligned}$$

C Proof of Theorem 2

C.1 Some Schur complement expressions

Lemma 5. *The Schur complement* $\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^\intercal$ *admits an expression*

$$\begin{aligned}
&\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^\intercal \\
&= \beta\mathbf{\Sigma} + \mathbf{FQ}[\mathbf{L}_2\boldsymbol{\upsilon}(\boldsymbol{\upsilon}^\intercal\mathbf{\Lambda}^{-1}\boldsymbol{\upsilon})^{-1}\boldsymbol{\upsilon}^\intercal\mathbf{L}_2^\intercal - \mathbf{H}]\mathbf{Q}^\intercal\mathbf{F}^\intercal.
\end{aligned}$$

Proof: In this proof we let $\mathbf{\Gamma}$ be the Cholesky factor of $\mathbf{\Lambda}$, i.e., $\mathbf{\Lambda} = \mathbf{\Gamma\Gamma}^\intercal$. From the expressions of $\mathbf{J}_{11}$, $\mathbf{J}_{21}$ and $\mathbf{J}_{22}$ in Theorem 1 it follows that

$$\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^\intercal = \beta\mathbf{\Sigma} + \mathbf{FQ}[\mathbf{L}_2\mathbf{\Pi L}_2^\intercal - \mathbf{H}]\mathbf{Q}^\intercal\mathbf{F}^\intercal, \tag{71}$$

where we define

$$\begin{aligned}
\mathbf{\Pi} &= \mathbf{\Lambda} - \mathbf{\Lambda L}_1^\intercal(\mathbf{L}_1\mathbf{\Lambda L}_1^\intercal)^{-1}\mathbf{L}_1\mathbf{\Lambda} \\
&= \mathbf{\Gamma}[\mathbf{I} - \mathbf{\Gamma}^\intercal\mathbf{L}_1^\intercal(\mathbf{L}_1\mathbf{\Gamma\Gamma}^\intercal\mathbf{L}_1^\intercal)^{-1}\mathbf{L}_1\mathbf{\Gamma}]\mathbf{\Gamma}^\intercal. \qquad (72)
\end{aligned}$$

However, the matrix $\bar{\mathbf{\Pi}} := \mathbf{I} - \mathbf{\Gamma}^\intercal\mathbf{L}_1^\intercal(\mathbf{L}_1\mathbf{\Gamma\Gamma}^\intercal\mathbf{L}_1^\intercal)^{-1}\mathbf{L}_1\mathbf{\Gamma}$ is the orthogonal projection operator onto the nullspace of $\mathbf{L}_1\mathbf{\Gamma}$.

From (9), (13) and the definition of $\mathbf{L}_1$ in (28) verify that that $\mathbf{L}_1\boldsymbol{\upsilon} = \mathbf{LP}\boldsymbol{\upsilon} = 0$. This means

$$\mathbf{L}_1\mathbf{\Gamma\Gamma}^{-1}\boldsymbol{\upsilon} = 0,$$

i.e. the vector $\mathbf{\Gamma}^{-1}\boldsymbol{\upsilon}$ spans the one dimensional nullspace of $\mathbf{L}_1\mathbf{\Gamma}$. Hence $\bar{\mathbf{\Pi}}$ is also the orthogonal projection operator onto the span of $\mathbf{\Gamma}^{-1}\boldsymbol{\upsilon}$. Hence

$$\bar{\mathbf{\Pi}} = \mathbf{\Gamma}^{-1}\boldsymbol{\upsilon}(\boldsymbol{\upsilon}^\intercal\mathbf{\Lambda}^{-1}\boldsymbol{\upsilon})^{-1}\boldsymbol{\upsilon}^\intercal\mathbf{\Gamma}^{-\intercal}.$$

Substituting this expression in (72) gives

$$\mathbf{\Pi} = \boldsymbol{v}(\boldsymbol{v}^{\intercal}\mathbf{\Lambda}^{-1}\boldsymbol{v})^{-1}\boldsymbol{v}^{\intercal},$$

which upon substitution in (71) yields the desired result. ■

Define

$$r_i := \boldsymbol{\alpha}_i^{\intercal}\boldsymbol{v}(\boldsymbol{v}^{\intercal}\mathbf{\Lambda}^{-1}\boldsymbol{v})^{-1/2}, \qquad i = 1, 2. \tag{73}$$

Note that

$$\mathbf{L}_2\boldsymbol{v}(\boldsymbol{v}^{\intercal}\mathbf{\Lambda}^{-1}\boldsymbol{v})^{-1}\boldsymbol{v}^{\intercal}\mathbf{L}_2^{\intercal} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}^{\intercal}, \tag{74}$$

see the definition of $\mathbf{L}_2$ in Theorem 1. When $r_2 = 0$ the matrix $\mathbf{L}_2\boldsymbol{v}(\boldsymbol{v}^{\intercal}\mathbf{\Lambda}^{-1}\boldsymbol{v})^{-1}\boldsymbol{v}^{\intercal}\mathbf{L}_2^{\intercal} - \mathbf{H}$ is of rank 1. Then the calculations turn out to be quite different from the case where $r_2 \neq 0$.

Lemma 6. *If* $r_2 = 0$ *then*

$$\det(\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^{\intercal}) = \frac{r_1^2}{\beta\sigma}\det(\beta\mathbf{\Sigma}), \tag{75}$$

Proof: When $r_2 = 0$ then using (73), definition of $\mathbf{Q}$ in Theorem 1 and the expressions given by Lemma 5 we get

$$\begin{aligned}
&\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^{\intercal} \\
&= \beta\mathbf{\Sigma} + \mathbf{F}\mathbf{Q}\begin{bmatrix} r_1^2 - \beta\sigma & 0 \\ 0 & 0 \end{bmatrix}\mathbf{Q}^{\intercal}\mathbf{F}^{\intercal} \\
&= \beta\mathbf{\Sigma} + \mathbf{F}\begin{bmatrix} \frac{1}{\sigma} & -\frac{\gamma}{\sigma} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} r_1^2 - \beta\sigma & 0 \\ 0 & 0 \end{bmatrix}\mathbf{Q}^{\intercal}\mathbf{F}^{\intercal} \\
&= \beta\mathbf{\Sigma} + \mathbf{F}\begin{bmatrix} r_1^2/\sigma - \beta & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{\sigma} & 0 \\ -\frac{\gamma}{\sigma} & 1 \end{bmatrix}\mathbf{F}^{\intercal} \\
&= \beta\mathbf{\Sigma} + \mathbf{F}\begin{bmatrix} r_1^2/\sigma^2 - \beta/\sigma & 0 \\ 0 & 0 \end{bmatrix}\mathbf{F}^{\intercal} \\
&= \beta\mathbf{\Sigma} + (r_1^2/\sigma^2 - \beta/\sigma)\mathbf{\Sigma}\mathbf{c}\mathbf{c}^{\intercal}\mathbf{\Sigma}.
\end{aligned} \tag{76}$$

In this proof we write

$$q = r_1^2/\sigma^2 - \beta/\sigma$$

compactly. From (76) we have

$$\begin{aligned}
&\det(\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^{\intercal}) = \det\left(\beta\mathbf{\Sigma} + q\mathbf{\Sigma}\mathbf{c}\mathbf{c}^{\intercal}\mathbf{\Sigma}\right) \\
&= \det(\beta\mathbf{\Sigma})\det\left(\mathbf{I} + \frac{q}{\beta}\mathbf{c}\mathbf{c}^{\intercal}\mathbf{\Sigma}\right) \\
&= \det(\beta\mathbf{\Sigma})\det\left(1 + \frac{q}{\beta}\mathbf{c}^{\intercal}\mathbf{\Sigma}\mathbf{c}\right) \\
&= \det(\beta\mathbf{\Sigma})\det\left(1 + \frac{q\sigma}{\beta}\right)
\end{aligned}$$

Substituting the value of q we get (75).

■

Lemma 7. *Suppose that* $r_2 \neq 0$. *Then*

$$\det(\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^{\intercal}) = \frac{r_1^2}{\beta\sigma}\det(\beta\boldsymbol{\Sigma}), \tag{77}$$

$$\begin{aligned}(\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^{\intercal})^{-1} &= \left[\frac{1}{r_1^2} - \frac{1}{\beta\sigma}\left(\frac{r_2\gamma}{r_1} - 1\right)^2\right]\mathbf{c}\mathbf{c}^{\intercal} \\ &+ \left(\frac{r_2}{r_1}\mathbf{c}\boldsymbol{\eta}^{\intercal} - \mathbf{I}\right)[\beta\boldsymbol{\Sigma}]^{-1}\left(\frac{r_2}{r_1}\mathbf{c}\boldsymbol{\eta}^{\intercal} - \mathbf{I}\right)^{\intercal}.\end{aligned} \tag{78}$$

Proof: Define

$$\mathbf{B} := \left[\mathbf{Q}\left(\frac{\mathbf{L}_2\boldsymbol{\upsilon}\boldsymbol{\upsilon}^{\intercal}\mathbf{L}_2^{\intercal}}{(\boldsymbol{\upsilon}^{\intercal}\boldsymbol{\Lambda}^{-1}\boldsymbol{\upsilon})^{-1}} - \mathbf{H}\right)\mathbf{Q}^{\intercal}\right]^{-1} \tag{79}$$

Recall that $\zeta = r_1/r_2$. Hence from (74) we get

$$\begin{aligned}&\left(\frac{\mathbf{L}_2\boldsymbol{\upsilon}\boldsymbol{\upsilon}^{\intercal}\mathbf{L}_2^{\intercal}}{(\boldsymbol{\upsilon}^{\intercal}\boldsymbol{\Lambda}^{-1}\boldsymbol{\upsilon})^{-1}} - \mathbf{H}\right)^{-1} \\ &\quad = -\frac{1}{\beta\sigma}\begin{bmatrix} 1 & -r_1/r_2 \\ -r_1/r_2 & r_1^2/r_2^2 - \beta\sigma/r_2^2 \end{bmatrix} \\ &\quad = -\frac{1}{\beta\sigma}\begin{bmatrix} 1 & -\zeta \\ -\zeta & \zeta^2 - \beta\sigma/r_2^2 \end{bmatrix}.\end{aligned}$$

Hence by definition of $\mathbf{Q}$, see Theorem 1, we get

$$\begin{aligned}\beta\mathbf{B} &= -\frac{1}{\sigma}\begin{bmatrix} \sigma & 0 \\ \gamma & 1 \end{bmatrix}\begin{bmatrix} 1 & -\zeta \\ -\zeta & \zeta^2 - \beta\sigma/r_2^2 \end{bmatrix}\mathbf{Q}^{-1} \\ &= -\frac{1}{\beta\sigma}\begin{bmatrix} \sigma & -\sigma\zeta \\ \gamma - \zeta & -\zeta(\gamma - \zeta) - \beta\sigma/r_2^2 \end{bmatrix}\begin{bmatrix} \sigma & \gamma \\ 0 & 1 \end{bmatrix} \\ &= -\frac{1}{\sigma}\begin{bmatrix} \sigma^2 & \sigma(\gamma - \zeta) \\ \sigma(\gamma - \zeta) & (\gamma - \zeta)^2 - \beta\sigma/r_2^2 \end{bmatrix} \\ &= -\begin{bmatrix} \sigma & \gamma - \zeta \\ \gamma - \zeta & \frac{(\gamma - \zeta)^2}{\sigma} - \beta/r_2^2 \end{bmatrix}.\end{aligned} \tag{80}$$

Taking determinant we have

$$\det(\beta\mathbf{B}) = -\frac{\beta\sigma}{r_2^2}. \tag{81}$$

On the other hand, recall that $\mathbf{F} = [\ \boldsymbol{\Sigma}\mathbf{c} \ \ \boldsymbol{\eta}\ ]$. Hence using (24) we get

$$\mathbf{F}^{\intercal}\boldsymbol{\Sigma}^{-1}\mathbf{F} = \begin{bmatrix} \sigma & \gamma \\ \gamma & \boldsymbol{\eta}^{\intercal}\boldsymbol{\Sigma}^{-1}\boldsymbol{\eta} \end{bmatrix}. \tag{82}$$

Combining (80) and (82) we get

$$\beta\mathbf{B} + \mathbf{F}^\intercal\mathbf{\Sigma}^{-1}\mathbf{F} = \begin{bmatrix} 0 & \zeta \\ \zeta & \boldsymbol{\eta}^\intercal\mathbf{\Sigma}^{-1}\boldsymbol{\eta} + \frac{\beta}{r_2^2} - \frac{(\gamma-\zeta)^2}{\sigma} \end{bmatrix} \tag{83}$$

Taking determinant we have

$$\det(\beta\mathbf{B} + \mathbf{F}^\intercal\mathbf{\Sigma}^{-1}\mathbf{F}) = -\zeta^2 \tag{84}$$

Now using Lemma 5 and (79) we know

$$\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^\intercal = \beta\mathbf{\Sigma} + \mathbf{F}\mathbf{B}^{-1}\mathbf{F}^\intercal \tag{85}$$

Hence

$$\begin{aligned} &\det(\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^\intercal) \\ &= \det(\beta\mathbf{\Sigma} + \mathbf{F}\mathbf{B}^{-1}\mathbf{F}^\intercal) \\ &= \det(\beta\mathbf{\Sigma})\det(\mathbf{I} + \mathbf{\Sigma}^{-1}\mathbf{F}(\beta\mathbf{B})^{-1}\mathbf{F}^\intercal) \\ &= \det(\beta\mathbf{\Sigma})\det(\mathbf{I} + \mathbf{F}^\intercal\mathbf{\Sigma}^{-1}\mathbf{F}\{\beta\mathbf{B}\}^{-1}) \\ &= \frac{\det(\beta\mathbf{\Sigma})}{\det(\beta\mathbf{B})}\det\big(\beta\mathbf{B} + \mathbf{F}^\intercal\mathbf{\Sigma}^{-1}\mathbf{F}\big) \\ &= \det(\beta\mathbf{\Sigma})\frac{r_2^2\zeta^2}{\beta\sigma} = \det(\beta\mathbf{\Sigma})\frac{r_1^2}{\beta\sigma}. \end{aligned}$$

C.2 Proof of the formula for $\det(\mathbf{J})$

Using Schur's determinant formula we know

$$\det(\mathbf{J}) = \det(\mathbf{J}_{11})\det(\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^\intercal). \tag{86}$$

The result of Theorem 2 is immediate from (86) once we use the expression for $\det(\mathbf{J}_{22} - \mathbf{J}_{21}\mathbf{J}_{11}^{-1}\mathbf{J}_{21}^\intercal)$ given by (77).

D Proofs Lemma 2 and Lemma 3

D.1 Proof of Lemma 2

First write

$$\begin{aligned} \mathbf{A}(\mathbf{I} - \mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1} &= \mathrm{e}^{\mathrm{i}\omega}(\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})(\mathbf{I} - \mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1} \\ &= \mathrm{e}^{\mathrm{i}\omega}(\mathbf{I} - \mathbf{I} + \mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})(\mathbf{I} - \mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1} \\ &= \mathrm{e}^{\mathrm{i}\omega}\{(\mathbf{I} - \mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1} - \mathbf{I}\} \end{aligned}$$

Hence we have

$$\begin{aligned}
&\mathbf{A}(\mathbf{I}-\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1}\mathbf{b}\mathbf{b}^\intercal(\mathbf{I}-\mathbf{A}^\intercal\mathrm{e}^{\mathrm{i}\omega})\mathbf{A}^\intercal \\
&=\{(\mathbf{I}-\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1}-\mathbf{I}\}\mathbf{b}\mathbf{b}^\intercal\{(\mathbf{I}-\mathbf{A}^\intercal\mathrm{e}^{\mathrm{i}\omega})^{-1}-\mathbf{I}\} \\
&=(\mathbf{I}-\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1}\mathbf{b}\mathbf{b}^\intercal(\mathbf{I}-\mathbf{A}^\intercal\mathrm{e}^{\mathrm{i}\omega})^{-1} \\
&\qquad+\{(\mathbf{I}-\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1}-\mathbf{I}/2\}\mathbf{b}\mathbf{b}^\intercal \\
&\qquad+\mathbf{b}\mathbf{b}^\intercal\{(\mathbf{I}-\mathbf{A}^\intercal\mathrm{e}^{\mathrm{i}\omega})^{-1}-\mathbf{I}/2\} \\
&=(\mathbf{I}-\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1}\mathbf{b}\mathbf{b}^\intercal(\mathbf{I}-\mathbf{A}^\intercal\mathrm{e}^{\mathrm{i}\omega})^{-1} \\
&\qquad+\frac{1}{2}(\mathbf{I}-\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1}(\mathbf{I}+\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})\mathbf{b}\mathbf{b}^\intercal \\
&\qquad+\frac{1}{2}\mathbf{b}\mathbf{b}^\intercal(\mathbf{I}-\mathbf{A}^\intercal\mathrm{e}^{\mathrm{i}\omega})^{-1}(\mathbf{I}+\mathbf{A}^\intercal\mathrm{e}^{\mathrm{i}\omega}) \qquad (87)
\end{aligned}$$

Now use (87) in (45) to get

$$\mathbf{A}\mathbf{\Sigma}\mathbf{A}^\intercal=\mathbf{\Sigma}-\Phi_\ddagger(\mathbf{A})\mathbf{b}\mathbf{b}^\intercal-\mathbf{b}\mathbf{b}^\intercal\Phi_\ddagger(\mathbf{A}^\intercal),$$

and the proof is complete.

D.2 Proof of Lemma 3

This result follows from the structure of **A** and **b**. Let us write

$$\mathbf{F}=\begin{bmatrix}\mathring{\mathbf{A}}_1 & \mathbf{b}_1(\mathring{\mathbf{g}}-\mathring{\mathbf{a}}\mathring{g}_0)^\intercal \\ \mathbf{0}_{n\times n} & \mathring{\mathbf{A}}_1\end{bmatrix}, \qquad \mathbf{g}=\begin{bmatrix}\mathbf{0}_{n\times n} \\ \mathbf{b}_1\end{bmatrix}$$

Then from (17) we note that

$$\bar{\mathbf{C}}^{-1}\mathbf{A}=\begin{bmatrix}\mathbf{F} & \mathbf{g} \\ 0 & 0\end{bmatrix}\bar{\mathbf{C}}^{-1}, \qquad \bar{\mathbf{C}}\mathbf{b}=\mathbf{b}=\begin{bmatrix}0 \\ 1\end{bmatrix}.$$

Hence

$$\begin{aligned}
&\mathbf{b}^\intercal\bar{\mathbf{C}}^{-1}(\mathbf{I}-\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1} \\
&=[\,0\ \ 1\,]\begin{bmatrix}(\mathbf{I}-\mathbf{F}\mathrm{e}^{-\mathrm{i}\omega})^{-1} & (\mathbf{I}-\mathbf{F}\mathrm{e}^{-\mathrm{i}\omega})^{-1}\mathbf{g}\mathrm{e}^{-\mathrm{i}\omega} \\ 0 & 1\end{bmatrix}\bar{\mathbf{C}}^{-1} \\
&=\mathbf{b}^\intercal\bar{\mathbf{C}}^{-1}. \qquad (88)
\end{aligned}$$

On the other hand

$$\begin{aligned}
\bar{\mathbf{C}}^{-1}(\mathbf{I}+\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})\mathbf{b}&=\left\{\mathbf{I}+\begin{bmatrix}\mathbf{F} & \mathbf{g} \\ 0 & 0\end{bmatrix}\mathrm{e}^{-\mathrm{i}\omega}\right\}\bar{\mathbf{C}}^{-1}\mathbf{b} \\
&=\begin{bmatrix}\mathbf{I}+\mathbf{F}\mathrm{e}^{-\mathrm{i}\omega} & \mathbf{g}\mathrm{e}^{-\mathrm{i}\omega} \\ 0 & 1\end{bmatrix}\mathbf{b}=\begin{bmatrix}\mathbf{g}\mathrm{e}^{-\mathrm{i}\omega} \\ 1\end{bmatrix}.
\end{aligned}$$

Hence

$$\begin{aligned}&\mathbf{b}^{\intercal}\bar{\mathbf{C}}^{-1}(\mathbf{I}-\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})^{-1}(\mathbf{I}+\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})\mathbf{b}\\ &=\mathbf{b}^{\intercal}\bar{\mathbf{C}}^{-1}(\mathbf{I}+\mathbf{A}\mathrm{e}^{-\mathrm{i}\omega})\mathbf{b}=\left[\begin{array}{cc}0 & 1\end{array}\right]\left[\begin{array}{c}\mathbf{g}\mathrm{e}^{-\mathrm{i}\omega}\\ 1\end{array}\right]=1.\end{aligned}$$

By multiplying both sides of the above equation by $\Phi(\mathrm{e}^{\mathrm{i}\omega})/(4\pi)$ and integrating with respect to ω we get the desired result using (40).