Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

The Cost of Complexity in Identification of FIR Systems ⋆

Cristian R. Rojas *, **Märta Barenthin** **, **James S. Welsh** *, **Håkan Hjalmarsson** **

* *School of Electrical Engineering and Computer Science, The University of Newcastle, NSW 2308, Australia (e-mail: cristian.rojas@studentmail.newcastle.edu.au).*
** *School of Electrical Engineering, KTH, 100 44 Stockholm, Sweden (e-mail: marta.barenthin@ee.kth.se)*

Abstract: In this paper we investigate the minimum amount of input power required to estimate a given linear system with a prescribed degree of accuracy, as a function of the model complexity. This quantity is defined to be the 'cost of complexity'. The degree of accuracy considered is the maximum variance of the discrete-time transfer function estimator over a frequency range $[-\omega_B, \omega_B]$. It is commonly believed that the cost increases as the model complexity increases. The objective of this paper is to quantify this dependence. In particular, we establish several properties of the cost of complexity. We find, for example, a lower bound for the cost asymptotic in the model order. For simplicity, we consider only systems described by FIR models and assume that there is no undermodelling.

1. INTRODUCTION

The purpose of system identification is to construct mathematical models of dynamical systems from experimental input/output data. To this end, a judicious choice of the input signal is crucial. This has motivated substantial interest in the topic of optimal experiment design. Indeed, many results have appeared on this topic, both in the statistics literature [Cox, 1958, Kempthorne, 1952, Fedorov, 1972] and in the engineering literature [Mehra, 1974, Goodwin and Payne, 1977, Zarrop, 1979, Jansson, 2004].

A key point as to why system identification can work in practice lies in the nature of the input signal: it is noted that experiment design can emphasize system properties of interest, while properties of little or no interest can be 'hidden' [Hjalmarsson, 2005, Hjalmarsson et al., 2006]. As remarked in [Hjalmarsson et al., 2006], some properties can be more easily estimated than others, in the sense that the amount of input power needed to estimate them with a given level of accuracy does not depend on the complexity of the model considered. However, some properties do depend on the model order. For example, it has been shown that the cost of estimating the transfer function at a particular frequency, or one non-minimum phase zero, is independent of the model order [Hjalmarsson et al., 2006].

This paper can be considered as an extension of the study of this phenomenon. Here we investigate the minimum amount of input power needed to estimate a given linear system with a prescribed degree of accuracy, as a function of the model complexity. This quantity is defined to be the 'cost of complexity'. The degree of accuracy considered is the maximum variance of the discrete-time transfer function estimator over a frequency range $[-\omega_B, \omega_B]$. For simplicity, we restrict the model class to systems described by FIR models. Also, we assume that there is no undermodelling, *i.e.* that the true system belongs to the model structure.

⋆ This work was supported by the Swedish Research Council.

The contribution of this paper consists of establishing several properties for the dependence of the cost on the model complexity. We believe that these results can provide a better understanding of the relationship between the amount of information that we ask to be extracted from a system, and the sensitivity of the cost of the identification with respect to the model complexity. This appears to be a key for understanding why system identification works for complex systems.

In order to study the problem posed in this paper, we employ a semidefinite optimization approach [Hildebrand and Gevers, 2003, Jansson and Hjalmarsson, 2005, Bombois et al., 2006]. In particular, the input design problem is formulated in terms of *Linear Matrix Inequalities* (LMIs) and the problem reduces to studying the positivity of a specific Toeplitz matrix.

This paper is organised as follows. The problem is formulated in Section 2. The main results are presented in Section 3 and a numerical example is provided in Section 4. Section 5 concludes the paper.

2. PROBLEM SET-UP

Consider the FIR system with input $u(t)$ and output $y(t)$,

$$y(t) = [\theta_{n_o}^o]^T \Lambda_{n_o}(q) u(t) + e_o(t) = G(q, \theta_{n_o}^o) u(t) + e_o(t),$$

where $\Lambda_{n_o}(q) := [\, 1 \; q^{-1} \; \cdots \; q^{-n_o} \,]^T$ with q^{-1} denoting the backward time shift operator and $\theta_{n_o}^o = [\, b_0^o \; \cdots \; b_{n_o}^o \,]^T$. Furthermore, $e_o(t)$ is zero mean white noise with variance σ_o^2, and the input signal is considered to be wide-sense stationary. The model to be fitted to this system is given by

$$y(t) = \sum_{k=0}^{n} b_k u(t-k) + e(t) = [\theta_n]^T \Lambda_n(q) u(t) + e(t).$$

where $n \geq n_o$. Consider the following autocovariance representation for the power spectrum of $u(t)$:

$$\Phi_u(\omega) := \sum_{k=-\infty}^{\infty} r_k e^{-j\omega k}. \quad (1)$$

Note that r_0 corresponds to the input power, *i.e.* $r_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_u(\omega)d\omega$. The (normalised) associated asymptotic covariance matrix of the estimated parameter vector is

$$\lim_{N\to\infty} N\mathrm{E}[(\hat{\theta}_{N,n} - \theta_n^o)(\hat{\theta}_{N,n} - \theta_n^o)^T] = \sigma_o^2 T_n^{-1},$$

where $\hat{\theta}_{N,n}$ is the Prediction Error (PE) parameter estimator of order n based on N observations of input/output data, $\theta_n^o := [\, b_0^o \cdots b_{n_o}^o \; 0 \cdots 0\,]^T$ and $T_n := T(\{r_k\}_{k=0}^n)$ is a Toeplitz matrix of the vector $[\, r_0 \; r_1 \cdots r_n\,]$ [Ljung, 1999]. In order for Φ_u to define a spectrum, it must satisfy

$$\Phi_u(\omega) \geq 0, \quad |\omega| \leq \pi. \quad (2)$$

We will design the sequence $r_0, r_1, \ldots, r_n$. However, we must ensure that there exists an extension $r_{n+1}, r_{n+2}, \ldots$ such that the nonnegativity constraint (2) holds. A necessary and sufficient condition for the existence of such an extension is that $T_n \geq 0$ [Grenander and Szegö, 1958, Byrnes et al., 2001, Lindquist and Picci, 1996].

In this paper we study the input design problem

$$\begin{aligned} \min_{\Phi_u} \; & \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_u(\omega)d\omega \\ \text{s.t.} \; & \Phi_u(\omega) \geq 0, \quad |\omega| \leq \pi \\ & \lim_{N\to\infty} N \text{ Var}\,\{G(e^{j\omega}, \hat{\theta}_{N,n})\} \leq \frac{1}{\gamma}, \quad |\omega| \leq \omega_B. \end{aligned} \quad (3)$$

where “s.t.” denotes “subject to”.

By the *Gauss' approximation formula* [Ljung, 1999, page 292],

$$\lim_{N\to\infty} N \text{ Var}\{G(e^{j\omega}, \hat{\theta}_{N,n})\} = \sigma_o^2 \Lambda_n^*(e^{j\omega}) T_n^{-1} \Lambda_n(e^{j\omega}). \quad (4)$$

This formula is valid when T_n is non-singular, *i.e.* when $T_n > 0$ (since T_n must be positive semidefinite in order to define a proper spectrum Φ_u). Under this assumption, by applying Schur complements [Boyd et al., 1994, page 7], the second constraint in (3) can be written as

$$T_n - \sigma_o^2\gamma\Lambda_n(e^{j\omega})\Lambda_n^*(e^{j\omega}) \geq 0, \quad |\omega| \leq \pi.$$

Thus, problem (3) can be reformulated, for $\omega_B \in (0, \pi]$, as:

$$\begin{aligned} \min_{r_0,\ldots,r_n} \; & r_0 \\ \text{s.t.} \; & T_n - \sigma_o^2\gamma\Lambda_n(e^{j\omega})\Lambda_n^*(e^{j\omega}) \geq 0, \quad |\omega| \leq \omega_B. \end{aligned} \quad (5)$$

(see *e.g.* [Hjalmarsson et al., 2006]). The constraint $T_n > 0$ has not been included in (5), because it can be shown (see Lemma 10 in Appendix C) that $T_n > 0$ holds for any solution of (5) if $\omega_B > 0$.

The case where $\omega_B = 0$ will be treated separately in Remark 1 of the next section, since, in this case the optimal solution gives a singular matrix T_n.

Let us denote by r_0^{opt} the solution to (5). The focus of this paper is thus to study the dependence of r_0^{opt} on the variables n, ω_B and γ, by analyzing the frequency-wise LMI

$$T_n - \sigma_o^2\gamma\Lambda_n(e^{j\omega})\Lambda_n^*(e^{j\omega}) \geq 0, \quad |\omega| \leq \omega_B. \quad (6)$$

The constraint (6) is infinite dimensional due to the dependence on the continuous variable ω. However, using the Generalised Kalman-Yakubovich-Popov (KYP) Lemma [Iwasaki and Hara, 2005], the dependence on ω is eliminated and thus (6) can be written as a finite dimensional problem. The trade-off is that we add two new matrix variables and that the dimension of the semidefinite program increases.

3. MAIN RESULTS

In this section the main results of this paper are presented. We start by stating some general properties of r_0^{opt} in Propositions 1 and 2. The implication of these propositions is that the more information we require for the model, the larger the cost. In particular, Proposition 1 shows that the cost is a non-decreasing function of n. Proposition 2 shows that the cost is a non-decreasing function of ω_B.

Proposition 1. The optimal cost of (5), r_0^{opt}, is a monotonically non-decreasing function of n.

Proof. Notice that

$$T_{n+1} - \sigma_o^2\gamma\Lambda_{n+1}(e^{j\omega})\Lambda_{n+1}^*(e^{j\omega}) = \begin{bmatrix} A_n & B_{n+1} \\ B_{n+1}^* & r_0 - \sigma_o^2\gamma \end{bmatrix},$$

where

$$A_n := T_n - \sigma_o^2\gamma\Lambda_n(e^{j\omega})\Lambda_n^*(e^{j\omega})$$

$$B_{n+1} := \begin{bmatrix} r_{n+1} - \sigma_o^2\gamma e^{j(n+1)\omega} \\ \vdots \\ r_1 - \sigma_o^2\gamma e^{j\omega} \end{bmatrix}.$$

Thus, if Φ_u satisfies $A_{n+1} \geq 0$, it also satisfies $A_n \geq 0$, for every $\omega \in [-\omega_B, \omega_B]$. This means that r_0^{opt} is monotonically non-decreasing in n. (This result can easily be extended to general model structures.) □

Proposition 2. (Monotonicity of r_0^{opt} with respect to ω_B). Let $r_0^{opt,1}$ and $r_0^{opt,2}$ be the optimal costs of the input design problem (3) for $\omega_B = \omega_{B1}$ and $\omega_B = \omega_{B2}$, respectively, and a fixed model order n. If $0 \leq \omega_{B1} < \omega_{B2} \leq \pi$, then $r_0^{opt,1} \leq r_0^{opt,2}$.

Proof. Follows from the fact that the set of allowable input spectra Φ_u decreases with increasing ω_B. □

The remaining results of this paper are consistent with Propositions 1 and 2. In the next theorem an upper bound for r_0^{opt} is derived by restricting the input spectrum to white noise spectra, *i.e.* $r_k = 0,\; k \neq 0$. This means that the only decision variable in (5) is r_0.

Theorem 1. (White noise input spectrum). For the case of white noise input spectra, we have $r_0^{opt} = r_{whitenoise}^{opt} := (n+1)\sigma_o^2\gamma$.

Proof. White noise corresponds to $r_k = r_0\delta_k$, where δ_k is defined by $\delta_0 = 1$ and $\delta_k = 0$ for $k \neq 0$. From (4) we obtain $\frac{\sigma_o^2}{r_0}\Lambda_n^*(e^{j\omega})\Lambda_n(e^{j\omega}) \leq 1/\gamma$. Since $\Lambda_n^*(e^{j\omega})\Lambda_n(e^{j\omega}) = n+1$, we obtain $r_0^{opt} = (n+1)\sigma_o^2\gamma$. □

From this theorem it is concluded that if we restrict the input to white noise, the cost is proportional to the model order $n+1$ and the precision γ, but independent of the bandwidth ω_B. Note that $r_{whitenoise}^{opt}$ constitutes an upper bound for r_0^{opt} due to the restriction in the structure of the input spectrum. The next theorem appears in [Hjalmarsson et al., 2006]. It considers the case when $\omega_B = 0$, *i.e.* it provides a lower bound for r_0^{opt}. Also, it shows that if we are only interested in estimating the static gain of the system, the optimal input is independent of the model order.

Theorem 2. When $\omega_B = 0$, the optimal cost is given by $r_0^{opt} = \sigma_o^2\gamma$.

Proof. This proof is a particular case of the proof of Theorem 3.1 of [Hjalmarsson et al., 2006]. Here we have that

$$T_n - \sigma_o^2\gamma\Lambda_n(1)\Lambda_n^*(1) \geq 0$$
$$\Leftrightarrow \begin{bmatrix} r_0-\sigma_o^2\gamma & r_1-\sigma_o^2\gamma & \cdots & r_n-\sigma_o^2\gamma \\ r_1-\sigma_o^2\gamma & r_0-\sigma_o^2\gamma & \cdots & r_{n-1}-\sigma_o^2\gamma \\ \vdots & \vdots & \ddots & \vdots \\ r_n-\sigma_o^2\gamma & r_{n-1}-\sigma_o^2\gamma & \cdots & r_0-\sigma_o^2\gamma \end{bmatrix} \geq 0. \quad (7)$$

A necessary condition for (7) to hold is $r_0 \geq \sigma_o^2\gamma$, hence $r_0^{opt} \geq \sigma_o^2\gamma$. On the other hand, if we take $\phi_u(\omega) = (\sigma_o^2\gamma)\delta(\omega)$ (e.g. by taking u to be a constant equal to $\sigma_o^2\gamma$), we have $r_i = \sigma_o^2\gamma$ for $i = 0,\dots,n$, which implies that $r_0^{opt} \leq \sigma_o^2\gamma$. This then implies that $r_0^{opt} = \sigma_o^2\gamma$. □

This theorem presents a loose lower bound for r_0^{opt} due to the fact that ω_B is fixed to zero. In Theorem 3 we derive a more refined lower bound (asymptotic in n) for r_0^{opt}, where ω_B is allowed to vary. However, before stating the theorem, we make a heuristic observation regarding r_0^{opt} by exploiting the asymptotic variance formula in [Ljung, 1985].

Observation 1. Using Ljung's asymptotic variance formula [Ljung, 1985], the condition

$$\lim_{N\to\infty} N\mathrm{Var}G(e^{j\omega}, \hat{\theta}_{N,n}) \leq 1/\gamma$$

can be approximately replaced by $(n+1)\frac{\sigma_o^2}{\Phi_u(\omega)} \leq 1/\gamma$. This implies that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_u(\omega)d\omega \geq \frac{1}{2\pi}\int_{-\omega_B}^{\omega_B} \Phi_u(\omega)d\omega \geq (n+1)\frac{\sigma_o^2\omega_B\gamma}{\pi}, \quad (8)$$

however, if we take

$$\Phi_u(\omega) = \begin{cases} (n+1)\sigma_o^2\gamma, & \text{if } \omega \in [-\omega_B, \omega_B] \\ 0, & \text{otherwise,} \end{cases}$$

(8) turns into an equality. Therefore a heuristic observation is that r_0^{opt} is asymptotically proportional to the model complexity $n+1$, to the accuracy γ and the bandwidth ω_B. This derivation of the asymptotic cost is not entirely rigorous (since $\Phi_u(\omega)$ also depends on n), which calls for some more detailed calculations. □

In the following theorem, we establish the findings made in Observation 1 in a rigorous fashion.

Theorem 3. (Lower bound for the asymptotic cost). Assume that $0 < \omega_B < \pi$. Then, there is an $n_{as} \in \mathbb{N}$, depending on σ_o^2, γ and ω_B, such that, for all $n \geq n_{as}$,

$$r_0^{opt} \geq \left[(n+1)\frac{\omega_B}{\pi} + 1\right]\sigma_o^2\gamma.$$

Proof. See Appendix A. □

Remark 1. The solution of (3) for $\omega_B = 0$ does not give a non-singular matrix T_n. However, if we add a small perturbation, say $\epsilon > 0$, to r_0, we obtain a non-singular T_n. Thus, $r_0^{opt} = \sigma_o^2\gamma$ is the *infimum* value of r_0, but it is not actually attainable, in the sense that the right hand side of (4) is not defined for $\det T_n = 0$, even though the variance of $G(e^{j\omega}, \hat{\theta}_{N,n})$ is meaningful in this case. In fact, in engineering terms, it is possible to generate the solution of this case by using a constant signal, which will give a consistent estimator of the steady state gain of the system. □

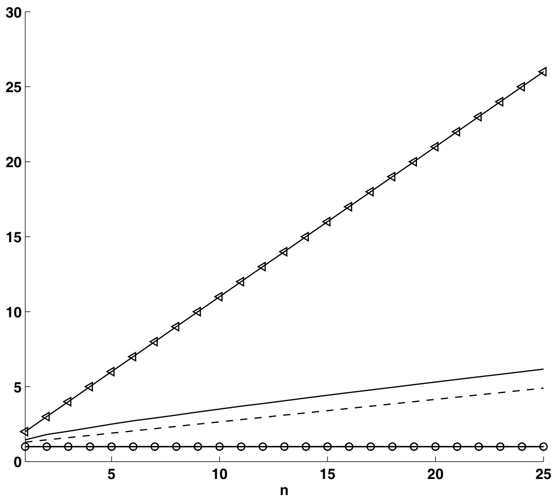

Fig. 1. The optimal cost (r_0^{opt}) from (5) versus model order n (solid); lower bound given by Theorem 2 ($\ominus$); asymptotic lower bound for r_0^{opt}, *c.f.* Theorem 3 (dashed); the white noise solution $r_{whitenoise}^{opt}$ ($\triangleleft$).

To summarize, the results of Propositions 1 and 2 are consistent with the fact that all bounds derived for r_0^{opt} are asymptotically affine in n, ω_B and γ. An upper bound for r_0^{opt} was derived in Theorem 1. The lower bound presented in Theorem 3 can be seen as a refinement of the result in Theorem 2 in the case where ω_B is allowed to vary. Furthermore, Theorem 3 establishes the findings of Observation 1 in a rigorous fashion.

4. NUMERICAL ILLUSTRATION

Let $\sigma_o^2 = 1$, $\omega_B = 0.15\pi$ and $\gamma = 1$. In Figure 1, the optimal solution r_0^{opt} is plotted together with the bounds presented in Section 3. It is seen from the figure that the asymptotic lower bound given in Theorem 3 is a tighter bound (*i.e.* closer to r_0^{opt}) than the simple bound given in Theorem 2.

5. CONCLUSIONS

In this paper we have studied the minimum amount of input power, r_0^{opt}, needed to estimate an FIR model with prescribed precision γ over the frequency range $[-\omega_B, \omega_B]$, as a function of the model order n. It is assumed that n is large enough to capture the true system. Several properties of r_0^{opt} are derived. It is shown that if n is large enough, r_0^{opt} is proportional to n, ω_B and γ. A loose upper bound for r_0^{opt} is given by a white noise input spectrum. The main contribution of this paper is that we provide a tighter asymptotic lower bound for r_0^{opt}. This bound quantifies the cost of extracting more information about the system and overmodelling. In simple terms, it can be concluded that, asymptotically in n,

$$r_0^{opt} \propto n\omega_B\sigma_o^2\gamma.$$

Hence, the results of this paper illustrate that the amount of information we ask to be extracted from the system determines how sensitive the cost of the identification experiment is with respect to the system (and model) complexity. This in turns means that the cost of identification can be kept low for complex systems if features of little or no interest are not excited.

Appendix A. PROOF OF THEOREM 3

By pre- and post-multiplying (6) by $\Lambda_n^*(e^{j\beta})$ and $\Lambda_n(e^{j\beta})$, respectively, where $\beta \in [0, \pi]$, it must hold that

$$\Lambda_n^*(e^{j\beta}) T_n \Lambda_n(e^{j\beta}) \geq \sigma_o^2 \gamma |\Lambda_n^*(e^{j\beta}) \Lambda_n(e^{j\omega})|^2, \\ |\omega| \leq \omega_B, \beta \in [0, \pi]. \tag{A.1}$$

Now,

$$|\Lambda_n^*(e^{j\beta}) \Lambda_n(e^{j\omega})|^2 = \Big| \sum_{k=0}^{n} e^{j(\beta-\omega)k} \Big|^2 = \frac{\sin^2(\frac{n+1}{2}[\beta-\omega])}{\sin^2(\frac{1}{2}[\beta-\omega])}$$

and

$$\Lambda_n^*(e^{j\beta}) T_n \Lambda_n(e^{j\beta}) = \sum_{m=-n}^{n} (n+1-|m|) r_k e^{-j\beta m}.$$

This implies that (A.1) is equivalent to

$$\sum_{m=-n}^{n} \left(1 - \frac{|m|}{n+1}\right) r_k e^{-j\beta m} \geq \sigma_o^2 \gamma \frac{1}{n+1} \frac{\sin^2(\frac{n+1}{2}[\beta-\omega])}{\sin^2(\frac{1}{2}[\beta-\omega])}, \\ |\omega| \leq \omega_B, \beta \in [0, \pi]. \tag{A.2}$$

The right hand side of (A.2) is the Fejér kernel F_n and, by Lemma 6, the left hand side of (A.2) is the convolution of F_n and Φ_u. Thus, (A.2) is equivalent to

$$\frac{1}{2\pi} [\Phi_u * F_n](\beta) \geq \sigma_o^2 \gamma F_n(\omega - \beta), \quad |\omega| \leq \omega_B, \beta \in [0, \pi]. \tag{A.3}$$

This expression can be further simplified by taking the supremum over $\omega \in [-\omega_B, \omega_B]$, and using Lemma 7. This implies that (A.3) is equivalent to

$$\frac{1}{2\pi} [\Phi_u * F_n](\beta) \geq \begin{cases} (n+1)\sigma_o^2 \gamma, & \text{if } \beta \in [0, \omega_B] \\ \sigma_o^2 \gamma \sup\limits_{\beta - \omega_B < x < \beta} F_n(x), & \text{if } \beta \in (\omega_B, \pi]. \end{cases} \tag{A.4}$$

Notice that, by Tonelli's Theorem [Bartle, 1966, page 118], the periodicity of Φ_u, and Lemma 8,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} [\Phi_u * F_n](\beta) d\beta = \int_{-\pi}^{\pi} \Phi_u(\beta) d\beta.$$

Thus, if we integrate both sides of (A.4) using

$$\tilde{F}_n(y) := \sup_{y < x < \beta} F_n(x) - F_n(y) \geq 0, \quad y \in (0, \pi - \omega_B],$$

and divide by 2π, we obtain

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\beta) d\beta \geq (n+1) \frac{\omega_B \sigma_o^2 \gamma}{\pi} + \frac{\sigma_o^2 \gamma}{\pi} \int_0^{\pi - \omega_B} F_n(\beta) d\beta \\ + \frac{\sigma_o^2 \gamma}{\pi} \int_0^{\pi - \omega_B} \tilde{F}_n(\beta) d\beta. \tag{A.5}$$

Let $N \in \mathbb{N}$ be such that, for every $n \geq N$,

$$\epsilon := \frac{\sigma_o^2 \gamma}{\pi} \int_0^{\pi - \omega_B} \tilde{F}_n(\beta) d\beta > 0. \tag{A.6}$$

The existence of such an N comes from the fact that $\sup_{y<x} F_n(x) = F_n(y)$ does not hold for every $y \in (0, \pi]$ (since F_n is not monotonically decreasing), and by Lemma 9,

$$\lim_{n \to \infty} \left[\int_0^{\pi - \omega_B} \tilde{F}_n(\beta) d\beta - \int_0^{\pi} \tilde{F}_n(\beta) d\beta \right] = 0.$$

Moreover, by Lemma 9, there is an $N' \geq N$ such that, for every $n \geq N'$,

$$\int_{\pi - \omega_B}^{\pi} F_n(\beta) d\beta < \frac{\epsilon \pi}{\sigma_o^2 \gamma}.$$

Therefore by Lemmas 8 and 5 we have

$$\frac{\sigma_o^2 \gamma}{\pi} \int_0^{\pi - \omega_B} F_n(\beta) d\beta \tag{A.7}$$
$$= \frac{\sigma_o^2 \gamma}{\pi} \int_0^{\pi} F_n(\beta) d\beta - \frac{\sigma_o^2 \gamma}{\pi} \int_{\pi - \omega_B}^{\pi} F_n(\beta) d\beta > \sigma_o^2 \gamma - \epsilon.$$

Thus, by rewriting (A.5) and (A.7), and using (A.6), we obtain the lower bound

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\beta) d\beta \geq (n+1) \frac{\omega_B \sigma_o^2 \gamma}{\pi} + \sigma_o^2 \gamma, \quad n \geq N.$$

This means that, for n sufficiently large, the optimal cost satisfies the asymptotic lower bound

$$r_0^{opt} \geq \left[(n+1) \frac{\omega_B}{\pi} + 1 \right] \sigma_o^2 \gamma. \tag{A.8}$$

Appendix B. PROPERTIES OF THE FEJÉR KERNEL

The Fejér kernel F_n is defined as

$$F_n(x) := \frac{1}{n+1} \frac{\sin^2(\frac{n+1}{2} x)}{\sin^2(\frac{1}{2} x)}, \quad x \in \mathbb{R}.$$

Properties of the Fejér kernel, necessary to establish the proof of Theorem 3, are given below.

Lemma 4.

$$F_n(x) = \frac{1}{n+1} \sum_{k=0}^{n} \sum_{l=-k}^{k} e^{jlx}.$$

Proof.

$$\begin{aligned} \frac{1}{n+1} \sum_{k=0}^{n} \sum_{l=-k}^{k} e^{jlx} &= \frac{1}{n+1} \sum_{k=0}^{n} \frac{e^{-jkx} - e^{j(k+1)x}}{1 - e^{jx}} \\ &= \frac{1}{n+1} \sum_{k=0}^{n} \frac{e^{j(k+1/2)x} - e^{-j(k+1/2)x}}{e^{jx/2} - e^{-jx/2}} \\ &= \frac{1}{n+1} \sum_{k=0}^{n} \frac{\sin([k+1/2]x)}{\sin(x/2)} \\ &= \frac{1}{n+1} \frac{1}{\sin(x/2)} \mathrm{Im} \left\{ \sum_{k=0}^{n} e^{j(k+1/2)x} \right\} \\ &= \frac{1}{n+1} \frac{1}{\sin(x/2)} \mathrm{Im} \left\{ e^{jx/2} \frac{1 - e^{j(n+1)x}}{1 - e^{jx}} \right\} \\ &= \frac{1}{n+1} \frac{1}{\sin(x/2)} \mathrm{Im} \left\{ \frac{1 - \cos([n+1]x) - j \sin([n+1]x)}{-2j \sin(x/2)} \right\} \\ &= \frac{1}{n+1} \frac{1}{\sin^2(x/2)} \frac{1 - \cos([n+1]x)}{2} \\ &= \frac{1}{n+1} \frac{\sin^2(\frac{n+1}{2} x)}{\sin^2(x/2)} \\ &= F_n(x). \end{aligned}$$

□

Lemma 5. $F_n(x) \geq 0$ and $F_n(-x) = F_n(x)$.

Proof. Follows directly from the definition. □

Lemma 6. If $g : \mathbb{Z} \to \mathbb{R}$ has Fourier transform G, then

$$\sum_{m=-n}^{n} \left[1 - \frac{|m|}{n+1} \right] g_m e^{-j\omega m} = \frac{1}{2\pi} [G * F_n](\omega), \omega \in [-\pi, \pi].$$

Proof.

$$\begin{aligned}
&\sum_{m=-n}^{n}\left[1-\frac{|m|}{n+1}\right]g_m e^{-j\omega m}\\
&=\frac{1}{n+1}\sum_{m=-n}^{n}(n+1-|m|)g_m e^{-j\omega m}\\
&=\frac{1}{n+1}\sum_{k=0}^{n}\sum_{l=0}^{n}g_{k-l}e^{-j\omega(k-l)}\\
&=\frac{1}{n+1}\sum_{k=0}^{n}\sum_{p=-k}^{k}g_p e^{-j\omega p}\\
&=\frac{1}{n+1}\sum_{k=0}^{n}\sum_{p=-k}^{k}\frac{1}{2\pi}\int_{-\pi}^{\pi}G(\beta)e^{j\beta p}d\beta e^{-j\omega p}\\
&=\frac{1}{2\pi}\int_{-\pi}^{\pi}G(\beta)\frac{1}{n+1}\sum_{k=0}^{n}\sum_{p=-k}^{k}e^{j(\beta-\omega)p}d\beta\\
&=\frac{1}{2\pi}\int_{-\pi}^{\pi}G(\beta)F_n(\beta-\omega)d\beta\\
&=\frac{1}{2\pi}[G*F_n],
\end{aligned}$$

where we have used Lemmas 4 and 5. □

Lemma 7. $F_n(x)\le F_n(0)=n+1$ for all $x\in[-\pi,\pi]$.

Proof.

$$F_n(0)=\lim_{n\to\infty}\frac{1}{n+1}\frac{\sin^2(\frac{n+1}{2}x)}{\sin^2(\frac{x}{2})}=n+1.$$

On the other hand, by Lemma 4, for all $x\in[-\pi,\pi]$ we have that

$$\begin{aligned}
|F_n(x)|&=\frac{1}{n+1}\sum_{k=0}^{n}\sum_{l=-k}^{k}e^{jlx}\\
&=\frac{1}{n+1}\sum_{k=0}^{n}\sum_{p=0}^{n}e^{j(k-p)x}\\
&=\frac{1}{n+1}\left|\sum_{p=0}^{n}e^{jkx}\right|^2\\
&\le\frac{1}{n+1}\left(\sum_{p=0}^{n}|e^{jkx}|\right)^2\\
&=1.
\end{aligned}$$

□

Lemma 8.

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}F_n(x)dx=1.$$

Proof. By Lemma 4,

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi}F_n(x)dx&=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{1}{n+1}\sum_{k=0}^{n}\sum_{l=-k}^{k}e^{jlx}\right]dx\\
&=\frac{1}{n+1}\sum_{k=0}^{n}\sum_{l=-k}^{k}\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{jlx}dx\\
&=\frac{1}{n+1}\sum_{k=0}^{n}\sum_{l=-k}^{k}\delta_l\\
&=\frac{1}{n+1}(n+1)\\
&=1.
\end{aligned}$$

□

Lemma 9. For every $\delta>0$,

$$\lim_{n\to\infty}\int_{\delta}^{\pi}F_n(x)dx=0.$$

Proof. For $\delta\le x\le\pi$, we have

$$F_n(x)=\frac{1}{n+1}\frac{\sin^2(\frac{n+1}{2}x)}{\sin^2(\frac{x}{2})}\le\frac{1}{n+1}\frac{1}{\sin^2(\frac{\delta}{2})}.$$

Thus

$$0\le\int_{\delta}^{\pi}F_n(x)dx\le\frac{1}{n+1}\frac{\pi-\delta}{\sin^2(\frac{\delta}{2})}\to 0$$

as $n\to\infty$. □

Appendix C. TECHNICAL LEMMA

Lemma 10. Let $\omega_B\in(0,\pi]$. Then, if $\{r_k\}_{k=0}^n$ is a solution to the input design problem (6), it satisfies $T_n>0$.

Proof. Pick $n+1$ different numbers $\{\omega_k\}_{k=0}^n$ from $[0,\omega_B]$. Then, from (6), the solution $\{r_k\}_{k=0}^n$ satisfies

$$T_n-\sigma_o^2\gamma\Lambda_n(e^{j\omega_1})\Lambda_n^*(e^{j\omega_1})\ge 0,\qquad\text{for } i=0,\dots,n. \tag{C.1}$$

By summing (C.1) over $i=0,\dots,n$, and dividing by $n+1$, we obtain

$$T_n-\frac{\sigma_o^2\gamma}{(n+1)}UU^*\ge 0, \tag{C.2}$$

where

$$U:=[\,\Lambda_n(e^{j\omega_0})\ \cdots\ \Lambda_n(e^{j\omega_n})\,]=\begin{bmatrix}1&\cdots&1\\ e^{-j\omega_0}&\dots&e^{-j\omega_n}\\ \vdots&&\vdots\\ e^{-jn\omega_0}&\dots&e^{-jn\omega_n}\end{bmatrix}.$$

Notice that U^* is a Vandermonde matrix [Horn and Johnson, 1990, page 29], whose determinant is

$$\det(U^*)=\prod_{0\le i<k\le n}(e^{j\omega_k}-e^{j\omega_l})\ne 0.$$

Thus, $UU^*>0$, hence by (C.2) we conclude that $T_n>0$. □

REFERENCES

R. G. Bartle. *The Elements of Integration.* John Wiley & Sons, New York, 1966.

X. Bombois, G. Scorletti, M. Gevers, P. Van den Hof, and R. Hildebrand. Least costly identification experiment for control. *Automatica*, 42(10):1651–1662, October 2006.

S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory.* SIAM Studies in Applied Mathematics, Philadelphia, 1994.

C. I. Byrnes, S. V. Gusev, and A. Lindquist. From finite covariance windows to modelling filters. *SIAM Review*, 43:645–675, 2001.

D. R. Cox. *Planning of Experiments.* Wiley, New York, 1958.

V.V. Fedorov. *Theory of optimal experiments.* Academic Press, New York, 1972.

G. C. Goodwin and R. L. Payne. *Dynamic System Identification: Experiment Design and Data Analysis.* Academic Press, New York, 1977.

U. Grenander and G. Szegö. *Toeplitz forms and their applications.* University of California Press, Berkley, CA., 1958.

R. Hildebrand and M. Gevers. Identification for control: Optimal input design with respect to a worst case ν-gap cost function. *SIAM Journal on Control and Optimization*, 41(5):1586–1608, 2003.

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

H. Hjalmarsson, J. Mårtensson, and B. Wahlberg. On some robustness issues in input design. In *Proceedings of the 14th IFAC Symposium on System Identification*, Newcastle, Australia, 2006.

R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, NY, USA, 1990.

T. Iwasaki and S. Hara. Generalized KYP Lemma: Unified frequency domain inequalities with design applications. *IEEE Transactions on Automatic Control*, 50(1):41–59, 2005.

H. Jansson. Experiment design with applications in identification for control. *Ph.D. thesis, KTH, TRITA-S3-REG-0404*, December 2004.

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, October 2005.

O. Kempthorne. *Design and Analysis of Experiments.* Wiley, New York, 1952.

A. Lindquist and G. Picci. Canonical correlation analysis, approximate covariance extension, and identification of stationary time series. *Automatica*, 32(5):709–733, 1996.

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

L. Ljung. *System Identification - Theory For the User, 2nd ed.* Prentice Hall, Upper Saddle River, New Jersey, 1999.

R. Mehra. Optimal input signals for parameter estimation in dynamic systems - survey and new results. *IEEE Transactions on Automatic Control*, 19(6):753–768, December 1974.

M. Zarrop. *Optimal Experiment Design for Dynamic System Identification.* Lecture Notes in Control and Information Sciences. Sci. 21. Springer Verlag, Berlin, 1979.