

# Noncausal Sampled Signal Reconstruction from Noisy Measurements: A System-Theoretic Approach

Gjerrit Meinsma

Department of Applied Mathematics  
University of Twente  
7500 AE Enschede, The Netherlands  
Email: *g.meinsma@utwente.nl*

Leonid Mirkin

Faculty of Mechanical Engineering  
Technion—IIT  
Haifa 32000, Israel  
Email: *mirkin@technion.ac.il*

**Abstract**—The paper studies the problem of reconstructing a not necessarily bandlimited analog signal from its noisy sampled measurements. Both A/D (analog-to-digital) and D/A (digital-to-analog) devices are treated as design parameters and no causality constraints are imposed on the reconstructor. We derive closed-form formulae for both  $L^2$ - and  $L^\infty$ -optimal reconstructors as well as expressions for the optimal achievable performance in both cases.

## I. INTRODUCTION

Nowadays, most data / signal / image processing operations are carried out by digital hardware. Measurements, pictures, voice, music, video are all stored in digital formats, processed by digital computers, and transmitted through digital communication lines. At the same time, in many cases information is taken from analog sources and utilized in analog forms (just think of music sold through the iTunes store). These facts raise the importance of the interconvertibility between analog and digital forms of representing information.

The problem of recovering analog information from a discretized (sampled) source is called the *sampled signal reconstruction* (SSR) problem [1], [2]. Modern engineering treatments of SSR [3] aim at accounting for practical requirements, such as handling non-bandlimited signals. A promising direction to cope with such requirements is the *system-theoretic* approach put forward in [4]. The idea is to bring concepts and tools of sampled-data control theory [5] (e.g., the lifting technique) to SSR problems. As customary in control, the system-theoretic approach is based on the use of systems to model exogenous signals and the use of system norms to measure the reconstruction performance. These provide a unified framework to treat both deterministic and stochastic signals, facilitate the quantification of the intersample behavior of involved signals, and conceptually simplify imposing causality constraints upon reconstructors. Disadvantages of the approach developed in [5] and its follow-ups (see [6] and the references therein) are that the structure of the original problem does not show up in the final solutions (because of a number of intermediate steps involved) and only rational models can be used (because of the state-space framework adopted).

We recently showed [7], [8] that these disadvantages are not intrinsic to the system-theoretic approach. In particular, in [8], we revised the lifting technique and demonstrated that

its use can lead to transparent solutions that do not rely upon the state-space machinery. With the help of the proposed approach, we provided then a system-theoretic proof of the Whittaker-Kotel'nikov-Shannon (WKS) Sampling Theorem by showing that both the  $L^2$  and the  $L^\infty$  performance criteria produce the sinc-interpolator as the optimal D/A converter and the ideal low-pass filter followed by the ideal sampler as the optimal A/D converter. This result corresponds to the case where the signal to be reconstructed has dominating low-frequency components (up to the Nyquist frequency) and is measured without noise. Remarkably, if these conditions hold true, the optimal reconstructor is independent of properties of the analog signal.

Our purpose in this paper is to study an extension of the Sampling Theorem to the case of noisy measurements. This problem was outlined in [8] as a possible extension, yet no thorough analysis was presented. It is worth emphasizing that whereas the noise-free solutions just recover the known result (the Sampling Theorem [1], [3]), complete solutions in the noisy case are not presently available to the best of our knowledge.

### A. Notation

Throughout,  $h$  denotes the sampling period and  $\omega_N := \frac{\pi}{h}$  is the associated Nyquist frequency. The sinc function with “period”  $h$  is defined as  $\text{sinc}_h(t) := \sin(\omega_N t) / (\omega_N t)$ . Signals are represented by lowercase symbols such as  $y(t) : \mathbb{R} \rightarrow \mathbb{C}$ , overbars indicate discrete-time signals,  $\bar{y}[k] : \mathbb{Z} \rightarrow \mathbb{C}$ . The unit step is denoted by  $\mathbb{1}(t)$  in continuous time and  $\bar{\mathbb{1}}[k]$  in discrete time. Similarly  $\delta(t)$  is the Dirac delta function and  $\bar{\delta}[k]$  is the discrete unit pulse. The number of elements of a vector-valued signal  $v$  is denoted by  $n_v$ .

Uppercase calligraphic symbols, like  $\mathcal{G}$ , denote continuous-time systems, the impulse response/kernel of which is denoted with lowercase symbols, such as  $g$ , and the corresponding transfer function and frequency responses (if exist) are denoted by uppercase symbols,  $G(s)$  and  $G(j\omega)$ . Discrete-time notation is similar modulo the use of overbars, like  $\bar{\mathcal{G}}$ ,  $\bar{g}$ ,  $\bar{G}(z)$ , and  $\bar{G}(e^{j\theta})$ .

## II. PROBLEM FORMULATION

The SSR setup studied in this paper is depicted in Fig. 1. Here  $v$  is an analog scalar signal to be reconstructed from a

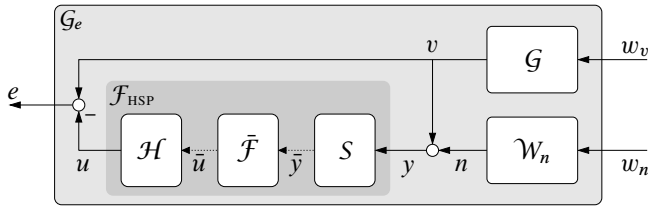


Fig. 1. Sampled signal reconstruction from noisy measurements

sampled version of the measured signal  $y = v + n$ , where  $n$  is measurement noise. We assume that the signals  $v$  and  $n$  are modeled as the outputs of stable LTI (linear time-invariant) systems  $\mathcal{G}$  and  $\mathcal{W}_n$ , respectively, driven by normalized (fictitious) scalar signals  $w_v$  and  $w_n$ . The measurable signal  $y$  is processed by the *hybrid signal processing block* (HSPB)  $\mathcal{F}_{\text{HSP}}$ , which is the cascade of an A/D device (sampler)  $S$ , a digital LTI filter  $\tilde{\mathcal{F}}$ , and a D/A device (interpolator or hold)  $\mathcal{H}$ . By a sampling device  $S$  we understand a linear device transforming a function  $y(t) : \mathbb{R} \rightarrow \mathbb{C}^{n_y}$  into a function  $\bar{y}[k] : \mathbb{Z} \rightarrow \mathbb{C}^{n_{\bar{y}}}$  as follows:

$$\bar{y} = S y : \quad \bar{y}[k] = \int_{-\infty}^{\infty} \psi(kh - s)y(s)ds, \quad k \in \mathbb{Z}, \quad (1)$$

for some  $\psi(t)$ , called the *sampling function*. By a hold device  $\mathcal{H}$  we understand a linear device transforming a function  $\bar{u}[k] : \mathbb{Z} \rightarrow \mathbb{C}^{n_{\bar{u}}}$  into a function  $u(t) : \mathbb{R} \rightarrow \mathbb{C}^{n_u}$  as

$$u = \mathcal{H}\bar{u} : \quad u(t) = \sum_{i \in \mathbb{Z}} \phi(t - ih)\bar{u}[i], \quad t \in \mathbb{R}, \quad (2)$$

for some *hold function*<sup>1</sup>  $\phi(t)$ . The output of  $\mathcal{F}_{\text{HSP}}$ , the analog signal  $u$ , is the estimate of  $v$ . Our goal then is to design the whole HSPB  $\mathcal{F}_{\text{HSP}}$  so that  $u$  is as close to  $v$  as possible under the constraint that the internal discrete signals  $\bar{u}$  and  $\bar{y}$  are scalars.

As the measure of the reconstruction performance we consider a size of the *error system*

$$\mathcal{G}_e := \begin{bmatrix} \mathcal{G} & 0 \end{bmatrix} - \mathcal{F}_{\text{HSP}} \begin{bmatrix} \mathcal{G} & \mathcal{W}_n \end{bmatrix} \quad (3)$$

connecting the exogenous signals  $w_v$  and  $w_n$  with the estimation error  $e = v - u$ . In this paper, we are concerned with the  $L^2$  and  $L^\infty$  norms of  $\mathcal{G}_e$  as the size measures. These are non-causal counterparts of the conventional  $H^2$  and  $H^\infty$  criteria, respectively. Although the error system is in general not time invariant (because  $\mathcal{F}_{\text{HSP}}$  is in general  $h$ -periodic), the definitions of the  $L^2$  and  $L^\infty$  system norms are readily extendible to  $\mathcal{G}_e$  [5], [8]. Roughly, the squared  $L^2$ -norm of  $\mathcal{G}_e$  is the mean squares of the estimation error  $e$  if the spectral densities of  $v$  and  $n$  are  $|G(j\omega)|^2$  and  $|W_n(j\omega)|^2$ , respectively. The  $L^\infty$  norm of  $\mathcal{G}_e$  corresponds to the worst-case energy of  $e$  under all  $v$  and  $n$  satisfying

$$\frac{|v(j\omega)|^2}{|G(j\omega)|^2} + \frac{|n(j\omega)|^2}{|W_n(j\omega)|^2} \leq 1$$

(this, in turn, requires that the spectral densities of  $v$  and  $y$  are bounded by  $|G|^2$  and  $|W_n|^2$ , respectively). Throughout the paper we assume that

<sup>1</sup>Thus,  $\psi$  stands for *sampler* and  $\phi$  for *hold*.

$$\mathcal{A}_1: |G(j\omega)| = |G(-j\omega)| \text{ and } |W_n(j\omega)| = |W_n(-j\omega)|;$$

$$\mathcal{A}_2: |G(j\omega)|^2 + |W_n(j\omega)|^2 > 0 \text{ for all } \omega \geq 0;$$

$$\mathcal{A}_3: |G(j\frac{\theta}{h})| \geq |G(j\frac{\theta+2\pi k}{h})| \text{ for all } \theta \in [0, \pi] \text{ and } k \in \mathbb{Z}.$$

Assumption  $\mathcal{A}_1$  just implies that the signals  $v$  and  $n$  are real valued.  $\mathcal{A}_2$  guarantees that the optimization problems are non-singular. Assumption  $\mathcal{A}_3$  says that the spectrum of  $v$  is dominated by its low-frequency components, namely by the components in the frequency range below the Nyquist frequency  $\omega_N$  (i.e., in the baseband). We refer to this property as the *baseband domination*. This assumption is made to simplify the exposition and can be easily omitted.

### III. ANALOG SOLUTIONS

The requirement that  $\mathcal{F}_{\text{HSP}}$  is the cascade of a sampler, a discrete filter, and a hold, with scalar internal discrete signals, can be viewed as a structural constraint imposed on the reconstructor (estimator). This suggests that the SSR problem can be addressed via the solution of the unconstrained problems, where the  $L^2$  or  $L^\infty$  norms of the error system (3) are minimized by an analog, not necessary time invariant, filter  $\mathcal{F}$ . We thus start with the latter problem following the ideas of [9].

First, remember [9] that the  $L^2$ -norm of  $\mathcal{G}_e$ ,  $\|\mathcal{G}_e\|_2$ , is the square root of the (operator) trace of  $\mathcal{G}_e \mathcal{G}_e^*$  and the  $L^\infty$ -norm of the error system  $\|\mathcal{G}_e\|_\infty \leq \gamma$  iff  $\mathcal{G}_e \mathcal{G}_e^* \leq \gamma^2 I$ . This is to say that the system  $\mathcal{G}_e \mathcal{G}_e^*$  plays a central role in both optimization problems. Now,

$$\begin{aligned} \mathcal{G}_e \mathcal{G}_e^* &= (I - \mathcal{F})\mathcal{G}\mathcal{G}^*(I - \mathcal{F})^* + \mathcal{F}\mathcal{W}_n\mathcal{W}_n^*\mathcal{F}^* \\ &= \mathcal{G}\mathcal{G}^* - \mathcal{F}\mathcal{G}\mathcal{G}^* - \mathcal{G}\mathcal{G}^*\mathcal{F}^* + \mathcal{F}(\mathcal{G}\mathcal{G}^* + \mathcal{W}_n\mathcal{W}_n^*)\mathcal{F}^* \\ &= \mathcal{Q} + (\mathcal{G}\mathcal{G}^*\mathcal{R}^{-1} - \mathcal{F})\mathcal{R}(\mathcal{G}\mathcal{G}^*\mathcal{R}^{-1} - \mathcal{F})^*, \end{aligned} \quad (4)$$

where  $\mathcal{R} := \mathcal{G}\mathcal{G}^* + \mathcal{W}_n\mathcal{W}_n^*$  is invertible by  $\mathcal{A}_2$  and

$$\mathcal{Q} := \mathcal{G}(I - \mathcal{G}^*\mathcal{R}^{-1}\mathcal{G})\mathcal{G}^* = \mathcal{G}\mathcal{G}^*\mathcal{R}^{-1}\mathcal{W}_n\mathcal{W}_n^*.$$

As no causality constraints are imposed, it is readily seen [9] that the optimal solution in both  $L^2$  and  $L^\infty$  cases is

$$\mathcal{F} = \mathcal{F}_a := \mathcal{G}\mathcal{G}^*\mathcal{R}^{-1} = \mathcal{G}\mathcal{G}^*(\mathcal{G}\mathcal{G}^* + \mathcal{W}_n\mathcal{W}_n^*)^{-1}$$

(in the  $L^\infty$  case it might be non-unique). This is an LTI analog system, which is not necessarily of the HSPB form (in fact, it is generally not a HSPB) and as such,  $\mathcal{F}_a$  is not the solution we seek.

Important for us is that (4) can be used to reduce the original SSR problem to a simpler problem, similar to the noise-free problem studied in [8]. This reduction, however, is different in the  $L^2$  and  $L^\infty$  cases.

#### A. $L^2$ optimization

Because of the linearity of the operator trace,

$$\|\mathcal{G}_e\|_2^2 = \|\mathcal{Q}^{1/2}\|_2^2 + \|(\mathcal{F}_a - \mathcal{F})\mathcal{R}^{1/2}\|_2^2. \quad (5)$$

Hence, the  $L^2$  SSR problem is equivalent to the problem of

$$\min_{\mathcal{F}_{\text{HSP}}} \|(\mathcal{F}_a - \mathcal{F}_{\text{HSP}})\mathcal{R}^{1/2}\|_2, \quad (6)$$

which is a one-block problem. In the noise-free setting, the systems  $\mathcal{R}^{1/2}$  and  $\mathcal{F}_a$  should be replaced with  $\mathcal{G}$  and  $I$ , respectively. In the next section, we shall show that this difference does not lead to any conceptual difference though.

### B. $L^\infty$ optimization

The situation here is slightly more complicated than in the  $L^2$  case. Clearly,  $\mathcal{G}_e \mathcal{G}_e^* \leq \gamma^2 I$  iff

$$(\mathcal{F}_a - \mathcal{F})\mathcal{R}(\mathcal{F}_a - \mathcal{F})^* \leq \gamma^2 I - \mathcal{Q}.$$

This obviously requires that  $\gamma \geq \gamma_a$ , where

$$\gamma_a := \sqrt{\|\mathcal{Q}\|_\infty}$$

is the optimal  $L^\infty$  performance achievable with  $\mathcal{F} = \mathcal{F}_a$ .

If  $\gamma > \gamma_a$ , the system  $I - \gamma^{-2}\mathcal{Q}$  is stably invertible and there is a HSPB guaranteeing that  $\|\mathcal{G}_e\|_\infty \leq \gamma$  iff

$$\|(I - \gamma^{-2}\mathcal{Q})^{-1/2}(\mathcal{F}_a - \mathcal{F}_{\text{HSP}})\mathcal{R}^{1/2}\|_\infty \leq \gamma \quad (7)$$

for some  $\mathcal{F}_{\text{HSP}}$ . This is again a one-block problem, although the performance level shows up in both sides of (7). If  $\gamma = \gamma_a$ , the inverse might be unstable or even not well defined. In this case the analysis is a bit more complicated, although the final result remains the same.

*Remark 1:* Curiously, both  $L^2$  and  $L^\infty$  equivalent one-block problems (6) and (7), respectively, can be interpreted as (weighted) approximations of the analog optimal reconstructor  $\mathcal{F}_a$  by  $\mathcal{F}_{\text{HSP}}$ . In other words, the choice of “good” HSPBs can be viewed as an attempt to imitate their analog counterparts. This interpretation merely repeats the main point of [10, Sec. 6] made in the context of the sampled-data feedback control with causal controllers.

## IV. IMPOSING SAMPLED-DATA STRUCTURE

We are now in the position to consider constraints imposed by the sampled-data structure of  $\mathcal{F}_{\text{HSP}}$ . The rationale of the treatment is exactly the same as in the noise-free case studied in [8] (see Section VI for details). The problem is reformulated in the lifted domain, where the error system becomes shift invariant and  $\mathcal{F}_{\text{HSP}}$  is characterized by a finite rank (rank-one in our case) of its lifted frequency response, which is an  $L^2[0, h] \mapsto L^2[0, h]$  operator at each frequency. As no causality constraints are imposed on  $\mathcal{F}_{\text{HSP}}$ , the  $L^2$  and  $L^\infty$  lifted problems are solved for each frequency independently, as rank-one approximation problems in either the Hilbert-Schmidt or the induced norm sense. In both cases, the solution is obtained via the SVD of the lifted analog LTI systems, which can be efficiently carried out in terms of continuous-time frequency responses.

To streamline the exposition, we first formulate the main results postponing the bulky proofs to the next two sections. For the same reason, we do not present here some exotic cases, where the dominant frequency range is affected by properties of the measurement noise. These situations, possible in the  $L^2$  case, are not typical and their consideration would complicate the presentation considerably without changing qualitative conclusions too much.

Introduce the real nonnegative function

$$\rho(\omega) := \frac{|G(j\omega)|^2}{|W_n(j\omega)|^2}, \quad (8)$$

which can be interpreted as the signal-to-noise ratio spectrum. The main result of this paper is then formulated as follows:

*Theorem 1:* Let assumptions  $\mathcal{A}_{1,3}$  hold. Then

- if  $|G(j\omega)|^2 \frac{\rho(\omega)}{1+\rho(\omega)}$  is baseband dominant, the optimal  $L^2$  SSR performance is

$$\gamma_2^2 = \frac{1}{\pi} \int_0^{\omega_N} \frac{|G(j\omega)|^2}{1+\rho(\omega)} d\omega + \frac{1}{\pi} \int_{\omega_N}^{\infty} |G(j\omega)|^2 d\omega;$$

- the optimal  $L^\infty$  SSR performance is

$$\gamma_\infty = \max \left\{ \sup_{\omega \in [0, \omega_N]} \frac{|G(j\omega)|}{\sqrt{1+\rho(\omega)}}, \sup_{\omega \in (\omega_N, \infty)} |G(j\omega)| \right\}.$$

In both cases, the optimal reconstructor  $\mathcal{F}_{\text{HSP}}$  is composed of

- 1) the sinc-sampler, which is a generalized sampler of the form (1) with  $\psi(t) = \frac{1}{h} \text{sinc}_h(t)$ ;
- 2) the digital part  $\bar{\mathcal{F}}$  having the frequency response

$$\bar{F}(e^{j\theta}) = F_a(\theta/h) = \frac{\rho(\theta/h)}{1+\rho(\theta/h)},$$

for all  $\theta \in [-\pi, \pi]$ ;

- 3) the sinc-interpolator, which is a generalized hold of the form (2) with  $\phi(t) = \text{sinc}_h(t)$ ;

and is  $L^2(\mathbb{R})$ -stable.  $\nabla$

The optimal reconstructor of Theorem 1 can be presented in the form shown in Fig. 2. The sinc-sampler  $S_{\text{sinc}}$  can be

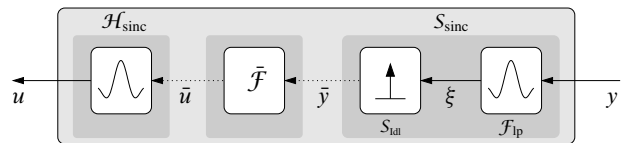


Fig. 2. Optimal  $\mathcal{F}_{\text{HSP}}$

thought of as the (noncausal) ideal low pass filter  $\mathcal{F}_{\text{lp}}$  with frequency support in  $[-\omega_N, \omega_N]$  followed by the ideal sampler  $S_{\text{sam}}$ . Thus, all high frequency components of  $y$  are filtered out before sampled. The output of the  $S_{\text{sinc}}$ ,  $\bar{y}$ , is then processed by the discrete filter  $\bar{\mathcal{F}}$ . As the frequency response of this filter is real valued for all frequencies, it is noncausal too (unless it is static, which happens iff either  $\rho(\omega)$  is constant for all  $\omega$  or  $\mathcal{W}_n = 0$ ). Finally, the output  $\bar{u}$  of  $\bar{\mathcal{F}}$  is interpolated by  $\mathcal{H}_{\text{sinc}}$ , which is exactly the sinc-interpolator from the Sampling Theorem.

The fact that all components of the optimal  $\mathcal{F}_{\text{HSP}}$  are not causal renders it impractical. The truncation of the impulse responses of its components—a common practice in the signal processing literature—might not be efficient as the decay rate if  $\text{sinc}_h(t)$  is quite slow [3]. The significance of Theorem 1, apart from a pure academic interest, is rather in establishing easily calculable lower bounds on the achievable reconstruction performance in the presence of measurement noise.

Some other remarks are in order:

*Remark 2:* The optimal performance indices in Theorem 1,  $\gamma_2$  and  $\gamma_\infty$ , have two components representing two extreme situations. The first of these components reflects the contribution of the baseband,  $[0, \omega_N]$ , and is a size of  $\mathcal{Q}$  in this frequency range. The frequency response of  $\mathcal{Q}$  is actually the spectrum of the estimation error under the optimal analog

reconstruction, see Section III. Thus, the baseband contributes, in a sense, by the optimal analog performance. The second component of  $\gamma_2$  and  $\gamma_\infty$  reflects the contribution of the high-frequency range,  $(\omega_N, \infty)$ , and is a size of  $\mathcal{G}$ . Thus, high frequency components contribute by the estimator-free performance. Thus, in  $[0, \omega_N]$  the sampled-data reconstructor recovers the analog performance, whereas in  $(\omega_N, \infty)$  it does nothing.

*Remark 3:* In the  $L^2$  case, Theorem 1 requires that the function  $|G(j\omega)|^2 \frac{\rho(\omega)}{1+\rho(\omega)}$  is baseband dominant. As  $|G(j\omega)|$  is baseband dominant by  $\mathcal{A}_3$ , this requirement is clearly guaranteed if the signal-to-noise ratio  $\rho(\omega)$  is a non-increasing function of  $\omega$ , which is a reasonable assumption in many applications. The dominance requirement might fail if  $\frac{\rho(\omega)}{1+\rho(\omega)}$  increases faster than  $|G(j\omega)|^2$  decreases. This, in turn, is possible if the signal-to-noise ratio increases considerably faster than the spectrum of  $v$  decays. If  $|G(j\omega)|^2 \frac{\rho(\omega)}{1+\rho(\omega)}$  is not baseband dominant,  $\gamma_2$  still consists of two terms based on the same functions, yet now defined over more complicated, possibly disconnected, frequency ranges. Similarly, in this case the optimal  $\mathcal{F}_{\text{HSP}}$  is more complicated.

*Remark 4:* Remarkably, spectral properties of  $\mathcal{W}_n$  do not affect the baseband dominance in the  $L^\infty$  case.

## V. PRELIMINARY TECHNICALITIES

In this subsection we collect some required preliminary results. The reader is referred to [8] for more details (although some of the results could also be found in the sampled-data control literature, see [5], [11] and the references therein).

### A. Lifting

For any signal  $f : \mathbb{R} \rightarrow \mathbb{C}^{n_r}$ , its *lifting*  $\check{f} : \mathbb{Z} \rightarrow \{[0, h) \rightarrow \mathbb{C}^{n_r}\}$  is defined as the sequence of functions  $\{\check{f}[k]\}$  for which

$$\check{f}[k](\tau) = f(kh + \tau)$$

is satisfied for all  $k \in \mathbb{Z}$  and  $\tau \in [0, h)$ . Fig. 3 explains the

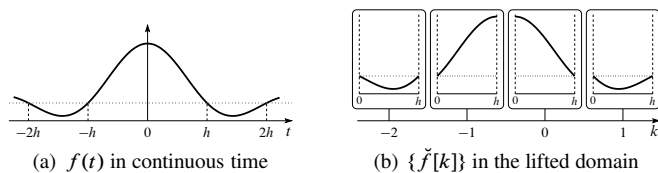


Fig. 3. Lifting analog signal  $f(t) = \text{sinc}_h(t)$

idea. With lifting we consider a function over  $\mathbb{R}$  as a concatenation (sequence) of functions over  $[0, h)$ . Clearly, this incurs no loss of information as it simply is another representation of the signal. Continuous-time systems  $y = \mathcal{G}u$  in the lifted domain are denoted as  $\check{y} = \check{\mathcal{G}}\check{u}$ .

### B. Fourier and $z$ -transform of lifted signals

Naturally the  $z$ -transform  $\mathfrak{Z}_h\{\check{f}\}$  of a lifted signal  $\check{f}$  is defined with respect to the discrete time index: the (lifted)  $z$ -transform  $\mathfrak{Z}_h\{\check{f}\}$  of a lifted signal  $\check{f}$  is defined as

$$\mathfrak{Z}_h\{\check{f}\} = \check{f}(z) := \sum_{k \in \mathbb{Z}} \check{f}[k]z^{-k}, \quad (9)$$

for all  $z \in \mathbb{C}$  for which the series converges. It is worthwhile to emphasize that for each  $z \in \mathbb{C}$ , the  $z$ -transform (if it exists) is still a function of time,

$$\check{f}(z; \tau) = \sum_{k \in \mathbb{Z}} \check{f}[k](\tau)z^{-k}, \quad \tau \in [0, h).$$

Like the  $z$ -transform, also the Fourier transform  $\check{f}(e^{j\theta})$  is still a function of time,

$$\check{f}(e^{j\theta}; \tau) = \sum_{k \in \mathbb{Z}} f(\tau + kh)e^{-jk\theta}, \quad \tau \in [0, h).$$

The lifted Fourier transform is a partial Fourier transform in that only time shifts of multiples  $kh$  are combined in the transformation. With that in mind, it is easy to figure out that it can be completed to a full classic Fourier transform  $F(j\omega)$ :

$$F(j\omega_k) = \int_0^h \check{f}(e^{j\theta}; \tau)e^{-j\omega_k \tau} d\tau, \quad (10)$$

where

$$\omega_k := \frac{\theta + 2\pi k}{h} = \frac{\theta}{h} + 2\omega_N k. \quad (11)$$

This says, in fact, that  $\frac{1}{h}F(j\omega_k)$  for any fixed  $\theta$  is the  $k$ th Fourier series coefficient of the function  $\check{f}(e^{j\theta}; \tau)e^{-j\theta\tau/h}$ . Hence, the classic Fourier theory tells us that for any  $f \in L^2(\mathbb{R})$  the inverse of (10) exists for almost every  $\theta \in [-\pi, \pi]$  with

$$\check{f}(e^{j\theta}; \tau) = \frac{1}{h} \sum_{i \in \mathbb{Z}} F(j\omega_i) e^{j\omega_i \tau}, \quad (12)$$

*Example 1:* To illustrate a use of formula (12), consider the Fourier transform of  $f(t) = \text{sinc}_h(t)$ . Because  $F(j\omega) = \mathbb{1}_{[-\omega_N, \omega_N]}(\omega)$ , (12) yields the following lifted Fourier transform of  $f(t)$ :

$$\check{f}(e^{j\theta}; \tau) = \frac{1}{h} e^{j\theta\tau/h}$$

for  $\theta \in [-\pi, \pi]$  and  $\tau \in [0, h)$ .  $\nabla$

### C. Transfer function and frequency response of lifted systems

Any linear system

$$u = \mathcal{G}y$$

that is  $h$ -shift invariant in continuous time is by construction LTI in the lifted domain with respect to the discrete variable and may hence be written as a convolution

$$\check{u}[k] = \sum_{i \in \mathbb{Z}} \check{\mathcal{G}}[k-i] \check{y}[i].$$

Here  $\check{\mathcal{G}}[k-i] \check{y}[i]$  for each  $i$  is a finite integral, but we need not delve into such matters here, see [8] for details. The transfer function of the lifted system is defined as

$$\check{\mathcal{G}}(z) := \sum_{i \in \mathbb{Z}} \check{\mathcal{G}}[i] z^{-i}. \quad (13)$$

and  $\check{\mathcal{G}}(e^{j\theta})$  is its *frequency response*.

#### D. Norms

By definition, the lifted  $L^2$ -norm of a signal is that of its underlying continuous-time signal,  $\|y\|_2 = \|\check{y}\|_2$ . For an  $h$ -shift invariant system  $\check{u} = \check{G}\check{y}$  the  $L^2$  system norm is defined as

$$\|\check{G}\|_2^2 := \frac{1}{h} \int_0^h \|\check{G}\delta(\cdot - \sigma)\|_{L^2(\mathbb{R})}^2 d\sigma.$$

The frequency-domain version is readily derived from Parseval's theorem, see [12]:

*Lemma 1:* The  $L^2$  system norm of an  $h$ -shift invariant system  $\check{u} = \check{G}\check{y}$  satisfies

$$\|\check{G}\|_2^2 = \frac{1}{2\pi h} \int_{-\pi}^{\pi} \|\check{G}(e^{j\theta})\|_{\text{HS}}^2 d\theta$$

where

$$\begin{aligned} \|\check{G}(e^{j\theta})\|_{\text{HS}}^2 &:= \int_0^h \text{tr}[\check{G}(e^{j\theta})\check{G}(e^{j\theta})^*] d\sigma \\ &:= \int_0^h \|\check{G}(e^{j\theta})\delta(\cdot - \sigma)\|_{L^2[0,h]}^2 d\sigma. \end{aligned}$$

is the Hilbert-Schmidt operator norm.  $\nabla$

For any fixed  $\theta$  the operator  $\check{G}(e^{j\theta})^*$  is the dual operator of  $\check{G}(e^{j\theta})$  with respect to the inner product of  $L^2[0, h]$ . The use of the so defined trace operator  $\text{tr}$  is convenient and resembles that for ordinary matrices.

The  $L^\infty$ -norm  $\|\check{G}\|_\infty$  of a linear system  $\check{u} = \check{G}\check{y}$  is defined as the  $L^2$ -induced norm,

$$\|\check{G}\|_\infty = \sup_{\check{y}} \frac{\|\check{G}\check{y}\|_2}{\|\check{y}\|_2}.$$

This definition makes sense for any linear system, whether LTI or not. If  $\check{G}$  is the lifting of an LTI system  $G$ , it is well known that  $\|\check{G}\|_\infty = \sup_{\omega} \bar{\sigma}(G(j\omega))$ , where  $\bar{\sigma}$  is the largest singular value and  $G(j\omega)$  is the frequency response (the Fourier transform the impulse response, causal or non-causal) of  $G$ . Similarly, as [13] showed, for  $h$ -shift invariant systems we have

$$\|\check{G}\|_\infty = \sup_{\theta \in [-\pi, \pi]} \bar{\sigma}(\check{G}(e^{j\theta})) \quad (14)$$

where now the largest singular value equals

$$\bar{\sigma}(\check{G}(e^{j\theta})) = \sup_{\check{u} \in L^2[0,h]} \frac{\|\check{G}(e^{j\theta})\check{u}\|_{L^2[0,h]}}{\|\check{u}\|_{L^2[0,h]}}.$$

Systems with finite  $L^\infty$  norm are said to be  $L^2$ -stable or simply *stable* because they map  $L^2$ -signals to  $L^2$ -signals and have a finite energy gain.

#### E. Singular value decomposition

If the continuous-time LTI system  $u = Gy$  is stable and strictly proper then its lifted frequency response  $\check{G}(e^{j\theta})$  for each  $\theta$  has an SVD and a remarkable one at that:

$$\check{G}(e^{j\theta})\check{y} = \sum_{i \in \mathbb{N}} \sigma_i \langle \check{y}, \beta_i \rangle \alpha_i, \quad \forall \check{y} \in L^2[0, h], \quad (15)$$

where  $\alpha_i = \beta_i = \frac{1}{\sqrt{h}} e^{j\omega_i \cdot}$  is an orthonormal basis of  $L^2[0, h]$  and  $\sigma_i = G(j\omega_i)$ , where  $\omega_i$  is given by (11). Strictly speaking

this is not yet an SVD as the  $\sigma_i$  here are not nonnegative and real, but this is easily remedied by absorbing phase of  $G(j\omega_i)$  in either  $\alpha_i$  or  $\beta_i$ . The conclusion is that for each  $\theta$  the mapping  $\check{G}(e^{j\theta})$  has countably many singular values  $|G(j\omega_i)|$ ,  $i \in \mathbb{N}$ .

Consider now only the entry with the index  $i = 0$  of (15), mapping from  $\check{y}$  to  $\check{u}$  as follows:

$$\check{u}(e^{j\theta}; \tau) = \frac{1}{h} e^{j\theta/h\tau} G(j\theta/h) \langle \check{y}(e^{j\theta}; \cdot), e^{j\theta \cdot/h} \rangle_{L^2[0,h]}. \quad (16)$$

We reordered the mapping as a concatenation of three mappings. First, with the help of Example 1 it can be shown that the scalar product describes a sampling device of the form (1) with the sampling function  $\psi(t) = \frac{1}{h} \text{sinc}_h(t)$ . This sampler, called the sinc-sampler,  $S_{\text{sinc}}$ , can be viewed as the ideal low-pass filter with support in  $[-\omega_N, \omega_N]$  followed by the ideal sampler, the sampling function of which is  $\psi(t) = \delta(t)$ . Second, the term  $G(j\theta/h)$  is the frequency response of a plain discrete-time system, i.e.,  $\bar{G}(e^{j\theta}) = G(j\theta/h)$ . Given  $G$ , the discrete system  $\bar{G}$  with this property can always be found. Finally, the remaining term,  $\frac{1}{h} e^{j\theta/h\tau}$ , corresponds to the sinc-interpolator (hold), which is a system of the form (2) with the hold function  $\phi(t) = \frac{1}{h} \text{sinc}_h(t)$ . This is exactly the D/A converter from the Sampling Theorem. Fig. 2 explains the idea.

#### F. Rank of hybrid signal processors

The hybrid signal processor of (16) for every  $\theta$  is a mapping that sends  $L^2[0, h]$  functions first to numbers and then back to  $L^2[0, h]$  functions. As a result it is a rank-one mapping. Finite-rank turns out to be a distinguishing feature of hybrid signal processors:

*Lemma 2 ([8]):* Let  $\check{G}$  be the lifting of a continuous-time  $h$ -shift invariant system  $G$ , the impulse response of which  $g \in L^2(\mathbb{R})$ . Then  $\check{G}$  is the lifting of an HSPB iff  $\text{rank } \check{G}(e^{j\theta}) \leq r$  for some  $r \in \mathbb{N}$  and all  $\theta \in [-\pi, \pi]$ .  $\nabla$

In particular if either of the signals  $\bar{y}$  and  $\bar{u}$  is scalar, then the rank is 1 (or zero).

## VI. PROOFS

#### A. $L^2$ optimization

In the lifted domain (6) rewrites as

$$\min_{\text{rank } \check{F}_{\text{HSP}}(e^{j\theta})=1} \frac{1}{2\pi h} \int_{-\pi}^{\pi} \|(\check{F}_a(e^{j\theta}) - \check{F}_{\text{HSP}}(e^{j\theta}))\check{R}^{1/2}(e^{j\theta})\|_{\text{HS}}^2 d\theta,$$

Because no causality constraints are imposed upon  $\mathcal{F}_{\text{HSP}}$ , its lifted frequency response can be treated frequency-wise (i.e., no dependences between the responses at different frequencies are implicitly imposed by analyticity requirements). Therefore, the minimization above can be carried out at each frequency  $\theta$  independently:

$$\min_{\text{rank } \check{F}_{\text{HSP}}(e^{j\theta})=1} \|(\check{F}_a(e^{j\theta}) - \check{F}_{\text{HSP}}(e^{j\theta}))\check{R}^{1/2}(e^{j\theta})\|_{\text{HS}},$$

at each  $\theta \in [-\pi, \pi]$ . In the sequel, we drop the frequency variable to simplify the exposition. Since  $\check{R}$  is nonsingular, the latter minimization can be rewritten as

$$\min_{\text{rank } \check{F}_2=1} \|\check{G}_2 - \check{F}_2\|_{\text{HS}} \quad (17)$$

for  $\check{G}_2 := \check{F}_a \check{R}^{1/2}$  and  $\check{F}_2 := \check{F}_{\text{HSP}} \check{R}^{1/2}$ .

Standard SVD arguments yield then that (17) is minimized by eliminating the largest singular value of  $\check{G}_2$ . As  $\check{G}_2$  is the lifting of

$$\check{G}_2 = \mathcal{F}_a \mathcal{R}^{1/2} = \check{G} \check{G}^* (\check{G} \check{G}^* + \mathcal{W}_n \mathcal{W}_n^*)^{-1/2},$$

which is an LTI system. We thus can use the results of §V-E to see that the singular values of  $\check{G}_2(e^{j\theta})$  are

$$\sigma_i = \frac{|G(j\omega_i)|^2}{\sqrt{|G(j\omega_i)|^2 + |\mathcal{W}_n(j\omega_i)|^2}} = |G(j\omega_i)| \sqrt{\frac{\rho(\omega_i)}{1 + \rho(\omega_i)}},$$

where  $\rho(\omega)$  is defined by (8). The baseband dominance assumption of Theorem 1 guarantees that the largest singular value is always  $\sigma_0$  and the optimal

$$\check{F}_{\text{HSP}}(e^{j\theta}) \check{y} = \alpha_0 F_a(j\theta/h) (\check{y}, \beta_0),$$

which leads to the optimal reconstructor via the discussion in the second part of §V-E.

To calculate the optimal performance, note that the singular values of the optimal  $\check{G}_2 - \check{F}_2$  are those of  $\check{G}_2$  except for  $\sigma_0$ , which is canceled by the optimal  $\check{F}_2$  and thus is zero. Equality (5) yields now that

$$\begin{aligned} \gamma_2^2 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} Q(j\omega) d\omega + \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i \in \mathbb{Z} \setminus \{0\}} \sigma_i^2 d\theta \\ &= \frac{1}{\pi} \int_0^{\infty} \frac{|G(j\omega)|^2}{1 + \rho(\omega)} d\omega + \frac{1}{\pi} \int_{\omega_N}^{\infty} \frac{|G(j\omega)|^2 \rho(\omega)}{1 + \rho(\omega)} d\omega \\ &= \frac{1}{\pi} \int_0^{\omega_N} \frac{|G(j\omega)|^2}{1 + \rho(\omega)} d\omega + \frac{1}{\pi} \int_{\omega_N}^{\infty} |G(j\omega)|^2 d\omega, \end{aligned}$$

where we used  $\mathcal{A}_1$  to restrict our attention to the positive frequency range. This completes the proof.

### B. $L^\infty$ optimization

First, assume that  $\gamma > \gamma_a$ , so that we can start with (7). Using arguments, similar to those in §VI-A, inequality (7) rewrites as

$$\bar{\sigma}(\check{G}_\infty(e^{j\theta}) - \check{F}_\infty(e^{j\theta})) \leq \gamma \quad \text{s.t.} \quad \text{rank } \check{F}_\infty(e^{j\theta}) = 1, \quad (18)$$

which is to be solved at each frequency  $\theta \in [-\pi, \pi]$  independently. Here we denote  $\check{G}_\infty := (I - \gamma^{-2} \check{Q})^{-1/2} \check{F}_a \check{R}^{1/2}$  and  $\check{F}_\infty := (I - \gamma^{-2} \check{Q})^{-1/2} \check{F}_{\text{HSP}} \check{R}^{1/2}$ , where the latter has rank one iff  $\check{F}_{\text{HSP}}$  is a rank-one operator. It is readily seen that  $\check{G}_\infty$  is the lifting of the LTI system

$$\begin{aligned} \check{G}_\infty &= (I - \gamma^{-2} Q)^{-1/2} \mathcal{F}_a \mathcal{R}^{1/2} \\ &= (I - \gamma^{-2} Q)^{-1/2} \check{G} \check{G}^* (\check{G} \check{G}^* + \mathcal{W}_n \mathcal{W}_n^*)^{-1/2} \end{aligned}$$

with the frequency response

$$G_\infty(j\omega) = \frac{|G(j\omega)|^2}{\sqrt{|G(j\omega)|^2 + |\mathcal{W}_n(j\omega)|^2 - \gamma^{-2} |G(j\omega)|^2 |\mathcal{W}_n(j\omega)|^2}}$$

Clearly, (18) is solvable iff the second largest singular value of  $\check{G}_\infty$  is bounded from above by  $\gamma$  and the optimal solution eliminates the largest singular value. By the result of §V-E, the singular values of  $\check{G}_\infty$  are  $G_\infty(j\omega_i)$ . The analysis is greatly simplified by the following result:

*Lemma 3:* Let  $\gamma > \gamma_a$ . Then

$$|G_\infty(j\omega)| \leq \gamma \iff |G(j\omega)| \leq \gamma.$$

*Proof:* Both sides are then equivalent to  $|G_e(j\omega)| \leq \gamma$  for  $F(j\omega) = 0$ .  $\square$

Thus, the second largest singular value of  $\check{G}_\infty$  is bounded from above by  $\gamma$  iff  $|G(j\omega_i)| \leq \gamma$  for all  $i$  but one. Using the baseband dominance assumption  $\mathcal{A}_3$ , it is readily seen that the largest  $|G(j\omega_i)|$  is  $|G(j\omega_0)|$  and we end up with the following solvability condition for (18):

$$|G(j\omega_i)| \leq \gamma \quad \forall \theta \in [-\pi, \pi] \text{ and } i \in \mathbb{Z} \setminus \{0\},$$

which, in turn, is equivalent to

$$\sup_{\omega \in (\omega_N, \infty)} |G(j\omega)| \leq \gamma.$$

Now, this is the solvability condition in the case where

$$\gamma < \gamma_a := \sup_{\omega} \sqrt{Q(j\omega)} = \sup_{\omega} \frac{|G(j\omega)|}{\sqrt{1 + \rho(\omega)}}.$$

The expression for  $\gamma_\infty$  in Theorem 1 then follows by noticing that  $\frac{1}{1 + \rho(\omega)} \leq 1$  for all  $\omega$ .

The expression for the optimal reconstructor follows then by the same arguments as in the  $L^2$  case. The only delicate point is to show that this reconstructor solves the problem if  $\gamma = \gamma_a$ . We omit the details of this because of space limitations.

### ACKNOWLEDGMENTS

G. Meinsma was supported in part at the Technion by a fellowship from the Lady Davis Foundation.

### REFERENCES

- [1] A. J. Jerri, "The Shannon sampling theorem—its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, no. 11, pp. 1565–1596, 1977.
- [2] J. R. Higgins, *Sampling Theory in Fourier and Signal Analysis: Foundations*. Oxford, UK: Oxford University Press, 1996.
- [3] M. Unser, "Sampling—50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, 2000.
- [4] P. P. Khargonekar and Y. Yamamoto, "Delayed signal reconstruction using sampled-data control," in *Proc. 35th IEEE Conf. Decision and Control*, Kobe, Japan, 1996, pp. 1259–1263.
- [5] T. Chen and B. A. Francis, *Optimal Sampled-Data Control Systems*. London: Springer-Verlag, 1995.
- [6] Y. Yamamoto, "A new approach to signal processing via sampled-data control theory," in *Proc. 5th Asian Control Conf.*, vol. 1, Melbourne, Australia, 2004, pp. 17–24.
- [7] G. Meinsma and L. Mirkin, "Sampled signal reconstruction via  $H^2$  optimization," in *Proc. 2006 IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. III, Toulouse, France, 2006, pp. 365–368.
- [8] —, "System theoretic perspectives of the Sampling Theorem," in *Proc. 17th MTNS Symposium*, Kyoto, Japan, 2006, pp. 730–741.
- [9] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite Quadratic Estimation and Control: A Unified Approach to  $H^2$  and  $H^\infty$  Theories*. Philadelphia: SIAM, 1999.
- [10] L. Mirkin, H. Rotstein, and Z. J. Palmor, " $H^2$  and  $H^\infty$  design of sampled-data systems using lifting—Part I: General framework and solutions," *SIAM J. Control Optim.*, vol. 38, no. 1, pp. 175–196, 1999.
- [11] G. E. Dullerud, *Control of Uncertain Sampled-Data Systems*. Boston: Birkhäuser, 1996.
- [12] B. Bamieh and J. B. Pearson, "The  $\mathcal{H}^2$  problem for sampled-data systems," *Syst. Control Lett.*, vol. 19, no. 1, pp. 1–12, 1992.
- [13] —, "A general framework for linear periodic systems with applications to  $H^\infty$  sampled-data control," *IEEE Trans. Automat. Control*, vol. 37, no. 4, pp. 418–435, 1992.