Dynamic Relations in Sampled Processes

Tryphon T. Georgiou, *Fellow, IEEE*, and Anders Lindquist, *Life Fellow, IEEE*

***Abstract*—Linear dynamical relations that may exist in continuous-time, or at some natural sampling rate, are not directly discernable at reduced observational sampling rates. Indeed, at reduced rates, matricial spectral densities of vectorial time series have maximal rank and thereby cannot be used to ascertain potential dynamic relations between their entries. This hitherto undeclared source of inaccuracies appears to plague off-the-shelf identification techniques seeking remedy in hypothetical observational noise. In this letter we explain the exact relation between stochastic models at different sampling rates and show how to construct stochastic models at the finest time scale that data allows. We then point out that the correct number of dynamical dependencies can only be ascertained by considering stochastic models at this finest time scale, which in general is faster than the observational sampling rate.**

***Index Terms*—Identification, stochastic systems.**

I. Introduction

SUPPOSE that we seek to identify linear dynamical relations that may exist between the components of a continuous-time process. What is typically available to us is the discrete-time sampled process (time series) of measurements collected at a given finite sampling rate. In this letter we are not concerned with issues of statistical estimation but instead assume that we can determine sufficiently accurately the spectral density of the sampled process. The theme of this letter is on how, from the model parameters of the sampled process, one may obtain a maximal number of dependencies between the entries of the process at a suitably finer time scale.

Modern-day applications, which aim towards high dimensional data and possibly varying sampling protocols, further underscore the importance of a careful consideration of how sampling affects dependencies. For example, consider a network as in Figure 1 with a continuous stationary

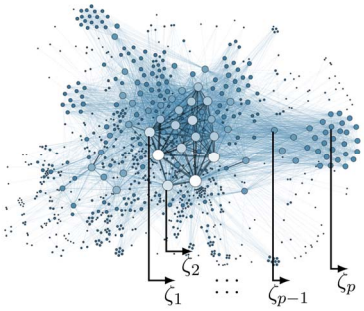

Fig. 1. Processes in a network environment.

stochastic vector process $\zeta = (\zeta_1, \zeta_2, \dots, \zeta_p)'$, whose components correspond to specific nodes. Suppose there is a deterministic dynamical relationship, described by the transfer function $T(s)$, between $u = (\zeta_1, \zeta_2, \dots, \zeta_m)'$ and $y = (\zeta_{m+1}, \zeta_{m+2}, \dots, \zeta_p)'$ as depicted in Figure 2. An important question is whether it is possible to detect this dynamical dependency from an observed sampled process $\zeta_k := \zeta(kh)$, where h is the sampling period. The same issue occurs when the original process is discrete-time and the observed process is subsampled at a slower rate.

The preponderance of techniques in the literature implicitly assume that the observed time series inherits any dynamical dependencies and typically seek to identify relations for the discrete-time process at the observation sampling rate. However this assumption cannot be made in general. Dependencies that may exist in continuous-time, or at some other fine "natural" sampling rate, are obfuscated by the process of sampling or sub-sampling. This is reflected in the fact that while the nullity of the spectral density of the original process coincides with the number of linear relations, the density of the sampled process has generically maximal rank. Evidently, when the observational sampling rate is sufficiently fast, the density of the sampled process is close to a singular one with the correct nullity. In such cases, hypothesizing a "small amount" of observational noise may salvage the situation and suggest a nearly correct model for dynamic dependencies. But in general there are no guarantees that hypothesizing observational noise, especially when there is none, as exemplified in Figure 2, will not exacerbate statistical inaccuracies.

The stochastic nature of the two vector processes u and y in Figure 2 originates in the "environment" of uncertain components and noise sources in feedback, while no observation

Manuscript received May 5, 2018; revised June 28, 2018; accepted July 18, 2018. Date of publication July 25, 2018; date of current version August 10, 2018. This work was supported in part by AFOSR under Grant FA9550-17-1-0435, in part by NSF under Grant ECCS-1509387, in part by ARO under Grant W911NF-17-1-0429, and in part by SSF. Recommended by Senior Editor G. Yin. *(Corresponding author: Tryphon T. Georgiou.)*

T. T. Georgiou is with the Department of Mechanical and Aerospace Engineering, University of California, Irvine, Irvine, CA 92697 USA (e-mail: tryphon@uci.edu).

A. Lindquist is with the Department of Automation and the School of Mathematics, Shanghai Jiao Tong University, Shanghai 200240, China, and also with the Department of Mathematics, KTH Royal Institute of Technology, 10044 Stockholm, Sweden (e-mail: alq@kth.se).

Digital Object Identifier 10.1109/LCSYS.2018.2859481

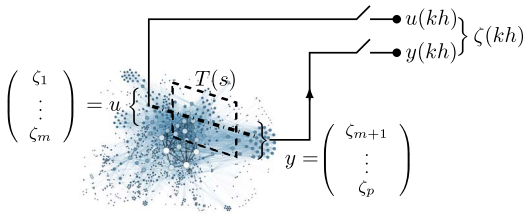

Fig. 2. Dynamical (network) dependencies.

noise is hypothesized. The setting represents a rather typical situation in distributed systems where variables can be accurately measured, albeit at a finite rate. The purpose of this letter is to point out that the statistics/model of the sampled or subsampled process contain enough information to correctly identify a "primitive" model for the underlying finer-time-scale process at its *natural finer time-resolution*. Such a primitive model, by necessity, exposes any dynamical relations between variables that may exist. This observation suggests the following potentially enabling new *dictum for linear system identification*: given time series data that has been collected at a sampling period h, and a derived model at that time resolution, consider the family of models at sampling periods h/q, with q an integer, or in continuous time, for underlying processes that are consistent with the observations at the particular sampling period h. Amongst those models select the one that exposes a maximal number of dynamical relationships. This will be the model with the finest time-resolution, and it will be referred to as the *lifting* of the observation-time-scale model to the natural time-resolution.

The outline of this letter is as follows. In Section II we explain how dynamical dependencies between the entries of a vector process are encoded in the spectral density and how they can be recovered via spectral factorization in the form of a transfer matrix that relates respective components. In Section III we provide the correspondence between continuous-time stochastic systems and their corresponding sampled discrete-time versions. In particular, Theorem 2 provides necessary and sufficient conditions for a discrete-time model to originate via sampling. Then, Section IV details the correspondence between sampled models at different sampling rates. In each of these two sections we present an academic example that illustrates key points. In Section V we investigate general discrete-time linear stochastic model as to when they can be lifted to continuous-time ones. Finally, Section VI concludes with implications of the theory.

II. Dynamic Relations and Spectral Rank

We consider finite-dimensional stationary Gauss-Markov real stochastic processes and are interested in identifying the maximal number of deterministic dynamical relations between the entries of such processes from observational data. Initially, in this section, we explain *how such relations relate to the spectral density of the process at its natural time resolution*, whether this is in continuous-time as depicted in Figure 2 or discrete-time at some basic sampling rate.

To exemplify the task of identifying dynamical dependencies consider the setting of Figure 2 and let

$$\zeta(t) := \begin{pmatrix} u(t) \\ y(t) \end{pmatrix} \tag{1}$$

be a stationary process partitioned into two components. Its stochastic nature originates in the "environment" that may include multiple feedback loops and noise sources. Assume that a deterministic dynamical relation between u and y, represented by the transfer function T, exists. This transfer function does not need to be stable – only the combined (feedback) dynamics need to be. Moreover, causation is not essential in our discussion, although we will see shortly how the entries of ζ can be partitioned into "inputs" and "outputs" so as to assign a consistent *proper* transfer function that models their dependence.

Continuing in the context of Figure 2, the spectral density

$$\Phi_\zeta(i\omega) = \begin{pmatrix} \Phi_u(i\omega) & \Phi_{uy}(i\omega) \\ \Phi_{yu}(i\omega) & \Phi_y(i\omega) \end{pmatrix},$$

of the stochastic process ζ factors as

$$\Phi_\zeta(i\omega) = \begin{pmatrix} I \\ T(i\omega) \end{pmatrix} \Phi_u(i\omega) \begin{pmatrix} I, & T(-i\omega)' \end{pmatrix}. \tag{2}$$

Thus, assuming that Φ_u is a.e. nonsingular, the number of deterministic dynamical relations between the entries of ζ coincides with the nullity of Φ_ζ. Indeed, the transfer function between u and y can be recovered from Φ_ζ since

$$T(s) = \Phi_{yu}(s)\Phi_u(s)^{-1}, \tag{3}$$

and the deterministic dependence between the entries of ζ can be expressed in "kernel" form as $L(i\omega)\Phi_\zeta(i\omega) = 0$ with $L(s) := (-T(s),\ I)$.

We now explain how such dynamical relations can be readily obtained in state-space form via spectral factorization of Φ_ζ. Throughout, we let ζ be a stationary p-vector stochastic process with the $p \times p$ rational spectral density $\Phi_\zeta(i\omega)$ having

$$\operatorname{rank}(\Phi_\zeta(i\omega)) = m, \text{ a.e., for } \omega \in \mathbb{R}.$$

Thus, $m \le p$, and when the inequality is strict, there are $p-m$ dynamical relations between the entries of ζ. For our purposes we assume throughout that $\Phi_\zeta(\infty) = 0$. This ensures that ζ has continuous sample paths, which is needed later on when we define the sampled process $\zeta_k := \zeta(kh)$. Denote by

$$V_\zeta(s) = H(sI - F)^{-1}G$$

a minimal stable $p \times m$ spectral factor of Φ_ζ (see [1, p. 198]). That is,

$$\Phi_\zeta(i\omega) = V_\zeta(i\omega)V_\zeta(-i\omega)' \tag{4}$$

with $H \in \mathbb{R}^{p\times n}$, $F \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{n\times m}$, (F, G) reachable, (F, H) observable, and F stable matrix, i.e., having all eigenvalues in the open left halfplane. In particular, G has full column rank and

$$\operatorname{rank}(G) = m \le n.$$

Moreover, we assume that $\operatorname{rank}(HG) = m$. Below we provide a factorization of V_ζ that displays the transfer function T; for notational convenience, rational proper functions with possibly

a constant term will be displayed in the following standard matrix-notation

$$C(sI-A)^{-1}B+D=:\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

Theorem 1: With V_ζ, F, G, H as above, re-order the rows of H and partition

$$H=\begin{pmatrix} H_0 \\ H_1 \end{pmatrix}$$

so that H_0G is $m\times m$ and invertible. Re-order in the same way the entries of ζ and as in (1) let u represent the first m entries and y the remaining. Then,

$$V_\zeta(s)=\begin{pmatrix} I \\ T(s) \end{pmatrix}M(s)s^{-1}, \text{ where} \tag{5}$$

$$T(s)=\left[\begin{array}{c|c} \Gamma & G(H_0G)^{-1} \\ \hline H_1\Gamma & H_1G(H_0G)^{-1} \end{array}\right], \tag{6a}$$

$$M(s)=\left[\begin{array}{c|c} F & G \\ \hline H_0F & H_0G \end{array}\right], \text{ and} \tag{6b}$$

$$\Gamma=F-G(H_0G)^{-1}H_0F. \tag{6c}$$

Proof: Partition V_ζ conformably with H, to write

$$\begin{aligned} V_\zeta(s) &= \begin{pmatrix} s^{-1}(H_0G+H_0F(sI-F)^{-1}G) \\ s^{-1}(H_1G+H_1F(sI-F)^{-1}G) \end{pmatrix} \\ &=: \begin{pmatrix} M(s) \\ N(s) \end{pmatrix}s^{-1}. \end{aligned}$$

The only thing that needs to be shown is the claimed expression in (6a). One can verify directly that $T(s)M(s)$ coincides with $N(s)=H_1G+H_1F(sI-F)^{-1}G$. To carry out the computations and verify, one needs to write the product in state-space form and cancel the modes corresponding to the system matrix Γ. ∎

A corresponding result can be easily worked out in discrete time where the spectral factor may also have a constant term.

III. Continuous-Time to Discrete-Time and Back

Once again we consider ζ to be a stationary p-vector process with $p\times p$ rational spectral density $\Phi_\zeta(i\omega)$ of constant rank m a.e. on $i\mathbb{R}$, and F, G, H as specified in the previous section. A minimal Markovian representation [1] of ζ is

$$dx(t)=Fx(t)dt+Gdw(t) \tag{7a}$$

$$\zeta(t)=Hx(t) \tag{7b}$$

with $w(t)$ a standard vectorial Wiener process of compatible dimension. The sampled paths of $x(t)$ are also continuous a.s. and the corresponding sampled process $x_k=x(kh)$ satisfies

$$x_{k+1}=e^{Fh}x_k+\int_{kh}^{(k+1)h}e^{F((k+1)h-\tau)}Gdw(\tau).$$

It follows that x_k and $\zeta_k=\zeta(kh)$ satisfy the discrete-time stochastic system of equations

$$x_{k+1}=Ax_k+Bv_k \tag{8a}$$

$$\zeta_k=Cx_k \tag{8b}$$

with,

$$A=e^{Fh} \tag{9a}$$

$$C=H \tag{9b}$$

$$Bv_k=\int_{kh}^{(k+1)h}e^{F((k+1)h-\tau)}Gdw(\tau), \tag{9c}$$

and v_k a sequence of independent random vectors having zero mean and unit variance (i.e., normalized white noise). Then

$$BB'=\textstyle\int_0^h e^{F\tau}GG'e^{F'\tau}d\tau=:Q.$$

Since (F, G) is a controllable pair, $Q=BB'$ is nonsingular and, in particular, we can take

$$B=Q^{1/2}. \tag{9d}$$

At the same time,

$$BB'=\int_0^\infty e^{F\tau}GG'e^{F'\tau}d\tau-\underbrace{e^{Fh}}_{A}\left(\int_0^\infty e^{F\tau}GG'e^{F'\tau}d\tau\right)\underbrace{e^{F'h}}_{A'},$$

and therefore,

$$P:=\int_0^\infty e^{F\tau}GG'e^{F'\tau}d\tau,$$

satisfies simultaneously two Lyapunov equations

$$P=APA'+BB', \text{ and} \tag{10a}$$

$$FP+PF'+GG'=0, \tag{10b}$$

i.e., both a "discrete-time" and a "continuous-time" Lyapunov equation.

We distill the bijection between the continuous-time model and its sampled counterpart in the following theorem.

Theorem 2: Consider the continuous-time stochastic model (7) having parameters (F, G, H) with (F, G) controllable. Sampling with period h gives rise to the discrete-time stochastic model (8) with parameters (A, B, C) satisfying (9). Conversely, if the parameters (A, B, C) of stochastic model (8) (with A a stability matrix) are such that

i) A admits a (principal) matrix logarithm [2] that we denote $\log(A)$, (11a)

ii) $\det(BB')\neq 0$, (11b)

iii) $\log(A)P+P\log(A)'\leq 0$, for P the solution of $P=APA'+BB'$, (11c)

then (8) arises by sampling (7) with $F=\frac{1}{h}\log(A)$, $H=C$, and G is a left factor of $-(FP+PF')$ of full column rank.

Proof: The first claim, to the effect that sampling leads to (9), has already been shown. For the second part, given any sampling period h, let $F=\frac{1}{h}\log(A)$, $H=C$, and G such that

$$GG'=-(FP+PF'), \tag{12}$$

it can be readily verified that conditions (9) hold. That (12) can be solved for a suitable G follows from (11c). ∎

The transfer function corresponding to (8) is

$$W(z)=C(zI-A)^{-1}B,$$

and the spectral density of the sampled process ζ_k is

$$\begin{aligned}\Psi(z) &= W(z)W(z^{-1})' \\ &= C(zI-A)^{-1}Q(z^{-1}I-A')^{-1}C'. \end{aligned} \tag{13}$$

Therefore, since $Q > 0$ and $\operatorname{rank}((zI-A)^{-1}) = n$, we have that $\operatorname{rank}(\Psi(z)) = \operatorname{rank}(C)$ a.e. On the other hand, $m = \operatorname{rank}(\Phi_\zeta(s)) \le \operatorname{rank}(H)$. Hence, since $\operatorname{rank}(C) = \operatorname{rank}(H) \ge m$,

$$p \ge \operatorname{rank}(\Psi(z)) \ge \operatorname{rank}(\Phi_\zeta(s)) = m.$$

Thus the rank of Ψ does not immediately reveal the $p-m$ linear dependencies. However m can be computed as the rank of $\log(A)P + P\log(A)'$ in the second part of Theorem 2.

Finally, since A is invertible (a necessity, since A admits a principal matrix logarithm [2]), $W(0) = -CA^{-1}B$ is finite, and since $W(\infty) = 0$,

$$\Psi(0) = 0. \tag{14}$$

These two conditions, A admitting a matrix logarithm and (14), characterize spectral densities of sampled processes.

Example 1: We now present an illustrative academic example. To this end we consider the continuous-time model (7) with

$$F = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & -5 & -6 & -4 \end{pmatrix}, G = \begin{pmatrix} 1 & -1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{pmatrix},$$

$$H = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 2 \\ \vdots & \ddots & \ddots & \\ 9 & 8 & 7 & 6 \\ 10 & 9 & 8 & 7 \end{pmatrix}. \tag{15}$$

The rank of the corresponding spectral density is 2, in agreement with the rank of G. Next, we consider stochastic models (A, B, C) for the sampled processes with sample periods $h \in \{1, 2^{-1}, 2^{-2}, \ldots, 2^{-7}\}$. Their spectral density has rank 4, and so does BB'. Thus, in Figure 3 we display the four eigenvalues of BB' in logarithmic scale as a function of h.

It is seen that only two eigenvalues remain dominant[1] as $h \to 0$. This is in agreement with the spectral density rank being 2 in the limit. Dynamical relations between the entries of ζ can be obtained from the limiting continuous-time stochastic model with (F, G, H) as in the second part of Theorem 2.

IV. Subsampling in Discrete-Time and Back

In analogy with the correspondence between continuous and discrete-time stochastic models, we now consider a discrete-time process with the following representation

$$x_{k+1} = Fx_k + Gw_k \tag{16a}$$
$$\zeta_k = Hx_k + Jw_k \tag{16b}$$

[1]Their ratio tends to the ratio of the non-zero eigenvalues of GG' since $GG' = \lim_{h\to 0} \frac{1}{h} BB'$.

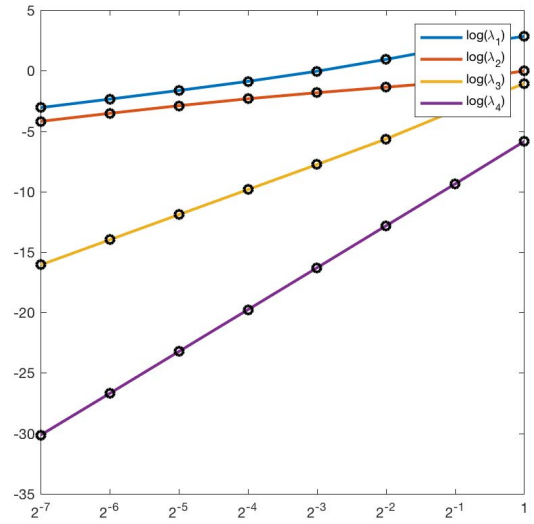

Fig. 3. $\log(\operatorname{eig}(P - e^{Fh}Pe^{F'h}))$ vs. h.

with w_k a discrete-time white noise process and $H \in \mathbb{R}^{p\times n}$, $F \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{n\times m}$, $m \le p$ as before, with (F, G) controllable and (F, H) observable pairs. Once again

$$\Phi_\zeta(e^{i\theta}) = V_\zeta(e^{i\theta})V_\zeta(e^{-i\theta})', \tag{17}$$

is the spectral density, with

$$V_\zeta(z) = H(zI-F)^{-1}G + J$$

a stable spectral factor of Φ_ζ. We use notation (F, G, H), Φ, etc., for this model to reflect the analogy to continuous-time in Section III, as it represents the underlying dynamics at the finest time-resolution. For any given positive integer q, we consider the subsampled process

$$\hat{\zeta}_\ell := \zeta_{\ell q}, \text{ for } \ell \in \{0, 1, 2, \ldots\}.$$

This admits the stochastic model

$$\xi_{\ell+1} = A\xi_\ell + Bv_\ell \tag{18a}$$
$$\hat{\zeta}_\ell = C\xi_\ell + Dv_\ell \tag{18b}$$

where $\xi_\ell := x_{q\ell}$ and

$$A = F^q, \tag{19a}$$
$$B = \begin{bmatrix} G, & FG & \ldots, F^{q-1}G \end{bmatrix}, \tag{19b}$$
$$C = H, \text{ and} \tag{19c}$$
$$D = \begin{bmatrix} 0, & 0 & \ldots, J \end{bmatrix} \tag{19d}$$

with

$$v_\ell = \begin{bmatrix} w'_{\ell q+q-1}, & w'_{\ell q+q-2} & \ldots, & w'_{\ell q} \end{bmatrix}'.$$

We note that $BB' = \sum_{k=0}^{q-1} F^k GG'(F^k)'$ gains in rank as q increases and will attain maximal rank for $q \ge n$, and so the spectral density

$$\Psi(z) = W(z)W(z^{-1})',$$

(where $W(z)$ will be specified in (22)) will have rank that is larger than the rank m of $\Phi_\zeta(z)$. Consequently, as in Section III, nullity of $\Phi_\zeta(z)$ that is due to dynamic relations between the entries of ζ will no longer be reflected in $\Psi(z)$.

The analogous statement to Theorem 2 is the following.

Theorem 3: Consider the discrete-time stochastic model (16) with parameters (F, G, H, J) and (F, G) controllable. Then, with q a positive integer, subsampling ζ_k every q steps gives rise to $\hat{\zeta}_\ell$ in (18) with parameters (A, B, C, D) satisfying (19). Conversely, given a model (18) with parameters (A, B, C, D), it is a model of a q-subsampled discrete-time process ζ_k provided

$$\text{i) } A \text{ admits a } q\text{th root that we denote } A^{1/q}, \text{ and} \tag{20a}$$

$$\text{ii) } M(P) = \begin{bmatrix} P - A^{1/q}P(A^{1/q})' & A^{1/q-1}BD' \\ DB'(A^{1/q-1})' & DD' \end{bmatrix} \geq 0,$$

$$\text{for } P \text{ the solution of } P = APA' + BB'. \tag{20b}$$

Parameters of a model (16) for the corresponding primary process ζ_k are $F = A^{1/q}$, $H = C$, and $\begin{bmatrix} G \\ J \end{bmatrix}$ a left factor of $M(P)$ of full column rank, i.e., so that

$$\begin{bmatrix} G \\ J \end{bmatrix} \begin{bmatrix} G' & J' \end{bmatrix} = M(P). \tag{20c}$$

If $D = 0$, (20b) in condition ii) reduces to

$$\text{ii)}' \quad P - A^{1/q}P(A^{1/q})' \geq 0. \tag{20d}$$

Proof: The first direction in the theorem has already been shown. It only remains to prove the converse direction. By setting $F = A^{1/q}$ and selecting G and J to satisfy (20c), we see that $(A, \hat{B}, C, \hat{D})$, with $\hat{B}$ and $\hat{D}$ given by (19b) and (19d), respectively, is a model of a q-subsampled discrete process. Now, it remains to show that

$$\begin{bmatrix} \hat{B} \\ \hat{D} \end{bmatrix} \begin{bmatrix} \hat{B}' & \hat{D}' \end{bmatrix} = \begin{bmatrix} B \\ D \end{bmatrix} \begin{bmatrix} B' & D' \end{bmatrix}.$$

To this end observe that

$$\hat{B}\hat{B}' = P - F^q P(F^q)' = P - APA',$$

which equals BB' by condition ii). Moreover, $\hat{D}\hat{D}' = JJ'$, which equals DD' by (20c). Finally, by (20c) and (20b),

$$\hat{B}\hat{D}' = F^{q-1}GJ' = A^{1-1/q}GJ' = BD'.$$

Hence, q-subsampling of the output process of (16) gives rise to a process with the same statistics as that in (18). ∎

Example 2: We now present an academic example where the fastest time-resolution is in discrete time. We begin with a discrete-time model having spectral density of rank two and no consistent model running with a faster clock being possible. To this end we take

$$x_{k+1} = \hat{F}x_k + Gu_k$$
$$\zeta_k = Hx_k$$

with zero constant term J, for simplicity, and with $\hat{F} = e^{Fh}$, $h = 0.5$ [time units] and F, G, H as in (15). Assume that data is available at intervals of $h = 2.5$ [time units] and, therefore, provide us with the discrete-time model (18) with parameters $A = \hat{F}^5$, $B = [G, \hat{F}G, \ldots \hat{F}^4G]$ and $C = H$. Selecting $q \in \{2, 3, 4, 5, 6, \ldots\}$ and applying the converse direction in Theorem 3 gives us models with a faster time rate, as long as $P - A^{1/q}P(A^{1/q})' \geq 0$, with P the solution of $P - APA' = BB'$.

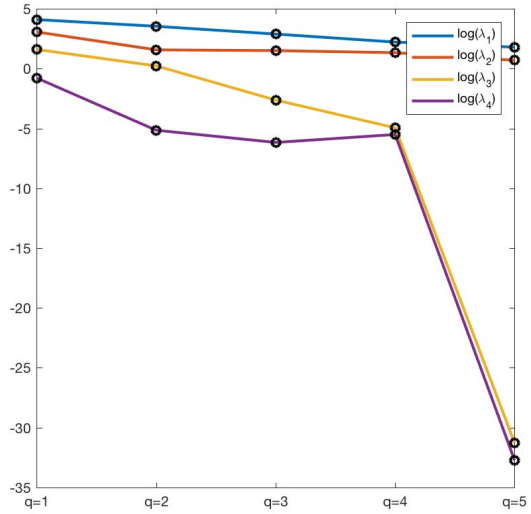

Fig. 4. $\log(\text{eig}(P - A^{1/q}P(A')^{1/q}))$ vs. q.

Naturally, for $q = 5$ we recover the parameters $(\hat{F}, G, H)$. It can be verified that

$$P - A^{1/q}P(A^{1/q})' \not\geq 0 \text{ for } q > 5.$$

Therefore no model with rate faster than 5 times the observational sampling rate is possible. As seen in Figure 4, the lowest two eigenvalues of $P - A^{1/q}P(A^{1/q})'$ separate substantially from the top two and vanish for $q = 5$, leading to $\text{rank}(P - A^{1/q}P(A^{1/q})') = 2$. Once again, the parameters $(\hat{F}, G, H)$ can be obtained from (A, B, C) as in Theorem 3.

V. When Is Lifting to Continuous-Time Possible?

Discrete-time (purely nondeterministic) stationary stochastic models are usually given in the form

$$\xi_{k+1} = A\xi_k + Bv_k \tag{21a}$$
$$\zeta_k = C\xi_k + Dv_k \tag{21b}$$

and, accordingly,

$$W(z) = C(zI - A)^{-1}B + D, \tag{22}$$

with a *nonzero constant term* D. As we already saw in Example 2, lifting to a continuous-time system may not always be possible, even when $D = 0$. Therefore, we now seek conditions on the parameters of (21) under which the output process ζ_k arises via sampling the output of a continuous-time model (7). For later reference we define $\bar{C} := CPA' + DB'$ and $\Lambda_0 := CPC' + DD'$ with P satisfying (10a).

Clearly condition (14) is necessary for ζ_k to be the sampled output process of (7). So we must have

$$\Psi(0) = W(0)W(\infty)' = (D - CA^{-1}B)D' = 0. \tag{23}$$

There are numerous systems (21) with the same output ζ_k, all having the same A and C, but possibly different matrices B and D [1, Ch. 6], each corresponding to a different spectral factor (22). Without loss of generality, we choose (21) to correspond to a minimum phase spectral factor and we let (B_-, D_-) be the corresponding pair of matrices. Then, if the $p \times p$ spectral density $\Psi(e^{i\theta})$ has rank r for all θ,

D_- is a $p \times r$ matrix of full column rank [1, Th. 6.1], and therefore $R := D'_-D_- > 0$. However, (23) implies that $(D_- - CA^{-1}B_-)R = 0$, and therefore

$$D_- - CA^{-1}B_- = 0. \tag{24}$$

Next, letting $x_k := A^{-1}\xi_{k+1}$,

$$x_{k+1} = \xi_{k+1} + A^{-1}B_-v_{k+1} = Ax_k + A^{-1}B_-v_{k+1},$$

and, since $\xi_k = A^{-1}\xi_{k+1} - A^{-1}B_-v_k = x_k - A^{-1}B_-v_k$, we have $\zeta_k = Cx_k + (D_- - CA^{-1}B_-)v_k$. Consequently, by (24),

$$\begin{aligned} x_{k+1} &= Ax_k + A^{-1}B_-v_{k+1} \\ \zeta_k &= Cx_k \end{aligned}$$

which has the form (8) of a sampled system. Now let

$$\hat{P} = A\hat{P}A' + A^{-1}B_-B'_-(A')^{-1}$$

be the corresponding Lyapunov equation. Then $P_- := A\hat{P}A'$ satisfies $P_- = AP_-A' + B_-B'_-$, and hence

$$P_- = \lim_{k\to\infty} \Pi_k \tag{25}$$

[1, Sec. 6.9], where $\Pi_{k+1} = \Lambda(\Pi_k)$, $\Pi_0 = 0$, with

$$\Lambda(\Pi) = \Pi - A\Pi A' + (\bar{C} - C\Pi A')'\Delta(\Pi)^{-1}(\bar{C} - C\Pi A')$$

and $\Delta(\Pi) := \Lambda_0 - C\Pi C'$. We summarize these conclusions in the following theorem.

Theorem 4: The output process ζ_k of the stochastic system (21) is the sampled process of the output of a continuous-time system (7) with parameters (F, G, H) provided $\Psi(0) = 0$ and conditions (11a), (11b) and $\log(A)\hat{P} + \hat{P}\log(A)' \leq 0$ hold, where $\hat{P} = A^{-1}P_-(A')^{-1}$ with P_- given by (25). Then ζ_k arises by sampling (7) with $F = \frac{1}{h}\log(A)$, $H = C$, and G being a left factor of $-(F\hat{P} + \hat{P}F')$ of full column rank.

VI. Concluding Remarks

The most basic underlying principle in modeling is that of parsimony where we seek a minimal number of parameters to explain the data. Historically, statistical reasoning was sought to mathematize the search for exact linear algebraic relations giving rise to methods of principle component analysis, factor analysis and so on, see [3], [4]. On a parallel route a sizable part of the literature has been devoted to statistical rationales for trading off inaccuracy with model complexity, see [5]. At this point, the art of modeling has accumulated a huge arsenal of ideas that includes regression analysis, likelihood methods, information criteria, regularization, and many more. The importance is further underscored by the mere citation-count of modern accounts of these subjects, e.g., [6] which surpassed hundred thousand citations in few short years.

In this letter we dealt with dynamical models. While the topic has a long history [4], issues of how sampling affects dynamical relations have not received sufficient attention. In fact, models are sought at a prespecified time-scale which often coincides with that at which data has been recorded, see the extensive literature on modern subspace identification methods [7]. While sampling rate selection is being discussed in various contexts, in all literature on dynamical relations (e.g., subspace methods, dynamical factor analysis, and so on) the typical assumption is that sampling is "sufficiently fast" with the premise that dynamical relations are not impacted if sought at the time-scale of the observation process.

The point of this letter is to bring attention to the fact that, while the sampling rate may or may not be "fast enough," one may retrieve exact dynamical relations at a finer time-scale from models at a coarse time-resolution. Should the generating mechanism dictate dynamic dependencies that truly originate at time-scale finer than the observational time rate, those dependencies may stay undetected or be poorly identified at the observation rate. Building on the paradigm of linear dynamical models and second-order processes we highlight the consequent dictum to seek models at the finest possible time-resolution that are consistent with the data at the given observation time scale. The implication is that, testing the dimenisionality of "lifted models" is a logical indicator of parsimony in system identification.

In summary, our setting involves dynamic (i.e., difference differential) linear relations between variables and stochastic noise. Guidelines that emerge can be summarized as follows:

a) Given data at observation time-intervals, determine linear dynamical models with parameters (A, B, C, D) using mostly standard techniques.
b) Determine whether the corresponding discrete time model originates from a continuous-time one via sampling the output process, and if not, determine if there is a discrete-time model at a faster time-scale which is consistent.
c) Dynamic relations between variables ought to be sought for the *lifted model at the finest time-scale*.

System identification techniques may be tuned to the idea that parsimony is sought at a scale other than that at which the system (A, B, C, D) is seen to operate. To this end, low-rank regularizers and the postulate of observational noise could be redirected into fitting low complexity models of the *lifted* model (F, G, H) of, e.g., Theorem 2. Thus, the principal contribution of this letter is to introduce the idea of *lifting identified models to a finer time-scale before assessing their complexity*. Sampling or subsampling a random process, and generating data at a rate that is lower than the native time scale, masks the exactness of any pre-existing dynamic relations.

References

[1] A. Lindquist and G. Picci, *Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification*. Heidelberg, Germany: Springer-Verlag, 2015.

[2] W. J. Culver, "On the existence and uniqueness of the real logarithm of a matrix," *Proc. Amer. Math. Soc.*, vol. 17, no. 5, pp. 1146–1151, 1966.

[3] L. Ning, T. T. Georgiou, A. Tannenbaum, and S. P. Boyd, "Linear models based on noisy data and the Frisch scheme," *SIAM Rev.*, vol. 57, no. 2, pp. 167–197, 2015.

[4] G. Picci and S. Pinzoni, "Dynamic factor-analysis models for stationary processes," *IMA J. Math. Control Inf.*, vol. 3, nos. 2–3, pp. 185–210, 1986.

[5] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Selected Papers of Hirotugu Akaike*. New York, NY, USA: Springer, 1998, pp. 199–213.

[6] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken, *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd ed. London, U.K.: Lawrence Erlbaum Assoc., 2003.

[7] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems: Theory-Implementation-Applications*. New York, NY, USA: Springer, 2012.