Autonomous State Space Models for Recursive Signal Estimation Beyond Least Squares

Nour Zalmai, Reto A. Wildhaber*, and Hans-Andrea Loeliger

ETH Zurich, Dept. of Information Technology & Electrical Engineering
* ETH Zurich & Bern University of Applied Sciences, Biel, Switzerland
{zalmai, wildhaber, loeliger}@isi.ee.ethz.ch

***Abstract*—The paper addresses the problem of fitting, at any given time, a parameterized signal generated by an autonomous linear state space model (LSSM) to discrete-time observations. When the cost function is the squared error, the fitting can be accomplished based on efficient recursions. In this paper, the squared error cost is generalized to more advanced cost functions while preserving recursive computations: first, the standard sample-wise squared error is augmented with a sample-dependent polynomial error; second, the sample-wise errors are localized by a window function that is itself described by an autonomous LSSM. It is further demonstrated how such a signal estimation can be extended to handle unknown additive and/or multiplicative interference. All these results rely on two facts: first, the correlation function between a given discrete-time signal and a LSSM signal can be computed by efficient recursions; second, the set of LSSM signals is a ring.**

I. Introduction

Fitting a parameterized signal to discrete-time measurements is a very classical problem. About two hundred years ago, Gauss invented both the least-squares method and its recursive version [1], and successfully applied it to predict the orbit of the newly discovered asteroid Ceres. Recursive least squares (which may be considered as a special case of Kalman filtering [2], [3]), continues to be a key algorithm in digital signal processing. However, the assumptions of linearity and of quadratic costs (or, equivalently, Gaussian noise) are not suitable for some applications, which has motivated nonlinear filters such as the extended Kalman filter (EKF), the unscented Kalman filter (UKF) [4], particle filters, and exact recursive filters [5], [6].

In this paper, we focus on recursive signal estimation rather than filtering. We consider parameterized signals that are generated by an autonomous linear state space model (LSSM) with unknown initial state. Such LSSM signals are highly versatile for modeling, and are of great practical use by virtue of two charming properties (see Sec. II): first, the correlation function between any such signal and any given discrete-time signal can be computed by efficient recursions; second, the element-wise multiplication of two LSSM signals is again a LSSM signal. These two properties are simple but yet extremely valuable. Indeed, in Sec. III, we introduce a general cost function that can still be recursively computed. This cost is obtained by replacing the standard sample-wise squared error by any sample-dependent polynomial cost and by weighting sample-wise errors with a LSSM window. In Sec. IV, we again exploit those two properties to handle signal estimation in the presence of an unknown additive and/or multiplicative interference that can be well modeled with a LSSM. Finally, in Sec. V, we present two illustrative applications of recursive signal estimation beyond least squares.

II. Definition and Properties of LSSM Signals

Definition 1 (LSSM signal): A discrete-time signal $s_j \in \mathbb{R}$, $j \in \mathbb{Z}$, is a LSSM signal (i.e., generated by a two-sided autonomous LSSM) if and only if there exists $C_\ell \in \mathbb{R}^{1\times n_\ell}$, $A_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$, $x_\ell \in \mathbb{R}^{n_\ell}$, and $C_\mathrm{r} \in \mathbb{R}^{1\times n_\mathrm{r}}$, $A_\mathrm{r} \in \mathbb{R}^{n_\mathrm{r}\times n_\mathrm{r}}$, $x_\mathrm{r} \in \mathbb{R}^{n_\mathrm{r}}$, for some, n_ℓ, $n_\mathrm{r} \in \mathbb{N}$ such that

$$s_j = \begin{cases} C_\ell A_\ell^{|j|} x_\ell & \text{for } j \le 0 \\ C_\mathrm{r} A_\mathrm{r}^{j} x_\mathrm{r} & \text{for } j > 0. \end{cases} \tag{1}$$

The changing point of this two-sided model is defined to be at time $j = 0$. However, when performing signal estimation, this is not a restriction since such LSSM signal will be shifted by a time of interest. The signal model can also be made left-sided or right-sided by setting $x_\mathrm{r} = 0$ or $x_\ell = 0$. The set of LSSM signals, which is a vector space, consists of linear combinations of (two-sided) exponentials, polynomials, sinusoids, finite-length signals, and products of those.

The LSSM parameters $\{C, A, x\}$ have to be understood as $\{C_\ell, A_\ell, x_\ell\} \cup \{C_\mathrm{r}, A_\mathrm{r}, x_\mathrm{r}\}$. In the following, the parameters $\{C_\ell, A_\ell, C_\mathrm{r}, A_\mathrm{r}\}$ are assumed to be known while the states $x = \{x_\ell, x_\mathrm{r}\}$ are to be estimated. Thus, a LSSM signal $s_j(x)$, $j \in \mathbb{Z}$, as in (1) is a function of x.

A. Inner Product with LSSM Signals

Given parameters $\{C, A\}$ and any discrete-time signal $y = (y_1, \ldots, y_K) \in \mathbb{R}^K$ of duration $K \in \mathbb{N}$, we define the quantity

$$\xi_k(y, C, A) = \begin{bmatrix} \overrightarrow{\xi}_k(y, C_\ell, A_\ell) \\ \overleftarrow{\xi}_k(y, C_\mathrm{r}, A_\mathrm{r}) \end{bmatrix} \in \mathbb{R}^n, \tag{2}$$

with $n = n_\ell + n_\mathrm{r}$ and

$$\overrightarrow{\xi}_k(y, C_\ell, A_\ell) = \sum_{i=1}^{k} (A_\ell^\mathsf{T})^{k-i} C_\ell^\mathsf{T} y_i \in \mathbb{R}^{n_\ell} \tag{3}$$

$$\overleftarrow{\xi}_k(y, C_\mathrm{r}, A_\mathrm{r}) = \sum_{i=k+1}^{K} (A_\mathrm{r}^\mathsf{T})^{i-k} C_\mathrm{r}^\mathsf{T} y_i \in \mathbb{R}^{n_\mathrm{r}}, \tag{4}$$

for $k \in \{1, \ldots, K\}$ and $\xi_k(y, C, A) = 0$, otherwise. Note that $\xi_k(y, C, A)$ is a linear function of y and can be interpreted as the output of n linear filters.

The quantity (2) is efficiently computed for all $k \in \{1, \ldots, K\}$ using the forward recursion

$$\overrightarrow{\xi}_k(y, C_\ell, A_\ell) = A_\ell^\mathsf{T} \overrightarrow{\xi}_{k-1}(y, C_\ell, A_\ell) + C_\ell^\mathsf{T} y_k, \tag{5}$$

initialized with $\overrightarrow{\xi}_0(y, C_\ell, A_\ell) = 0$ and the backward recursion

$$\overleftarrow{\xi}_k(y, C_\mathrm{r}, A_\mathrm{r}) = A_\mathrm{r}^\mathsf{T} \big(\overleftarrow{\xi}_{k+1}(y, C_\mathrm{r}, A_\mathrm{r}) + C_\mathrm{r}^\mathsf{T} y_{k+1}\big), \tag{6}$$

initialized with $\overleftarrow{\xi}_K(y, C_\mathrm{r}, A_\mathrm{r}) = 0$.

Proposition 1 (Inner Product with a LSSM signal): The inner product between y and a LSSM signal $s(x)$, for any x, as in (1) shifted by a time k and denoted by $s_{\bullet-k}(x)$ is

$$\langle y, s_{\bullet-k}(x)\rangle = x^\mathsf{T} \xi_k(y, C, A), \tag{7}$$

with the convention that $x^\mathsf{T} = [x_\ell^\mathsf{T}, x_\mathrm{r}^\mathsf{T}]$.

Proof of Proposition 1: This proposition follows from

$$\begin{aligned}
\langle y, s_{\bullet-k}(x)\rangle &= \sum_{i=1}^{K} y_i s_{i-k}(x) && (8)\\
&= \sum_{i=1}^{k} C_\ell A_\ell^{k-i} x_\ell y_i + \sum_{i=k+1}^{K} C_\mathrm{r} A_\mathrm{r}^{i-k} x_\mathrm{r} y_i && (9)\\
&= x_\ell^\mathsf{T} \overrightarrow{\xi}_k(y, C_\ell, A_\ell) + x_\mathrm{r}^\mathsf{T} \overleftarrow{\xi}_k(y, C_\mathrm{r}, A_\mathrm{r}) && (10)\\
&= x^\mathsf{T} \xi_k(y, C, A). && (11)
\end{aligned}$$

■

Proposition 1 has several important consequences. First, computing the correlation function between a signal of length K and a LSSM signal $s(x)$ has a complexity of $O(K\underline{n}^2)$, with $\underline{n} = \max(n_\ell, n_\mathrm{r})$. Secondly, the inner product between y and $s_{\bullet-k}(x)$ can be expressed as a standard inner product in $\mathbb{R}^n$ between x and $\xi_k(y, C, A)$. Finally, since $\xi_k(y, C, A)$ is independent of x, the computational effort to obtain the inner product between y and $s_{\bullet-k}(x)$ for any x, is of $O(n)$ only, after having computed $\xi_k(y, C, A)$.

B. Element-wise Product of LSSM signals

Proposition 2 (Product of LSSM signals): Let $s_j^{(1)}$ and $s_j^{(2)}$, $j \in \mathbb{Z}$, be two LSSM signals with respective parameters $\{C_1, A_1, x_1\}$ and $\{C_2, A_2, x_2\}$. Then, $s_j^{(1)} \cdot s_j^{(2)}$, $j \in \mathbb{Z}$, is also a LSSM signal with parameters $\{C_1 \otimes C_2, A_1 \otimes A_2, x_1 \otimes x_2\}$.

To keep the notation concise, $C_1 \otimes C_2$ (and analogously for $A_1 \otimes A_2$ and $x_1 \otimes x_2$) means that the Kronecker product is applied independently for the left-sided part ($C_{\ell,1} \otimes C_{\ell,2}$) and for the right-sided part ($C_{\mathrm{r},1} \otimes C_{\mathrm{r},2}$). We use this convention all along.

Proof of Proposition 2: For $j > 0$, we have

$$\begin{aligned}
s_j^{(1)} \cdot s_j^{(2)} &= (C_{\mathrm{r},1} A_{\mathrm{r},1}^j x_{\mathrm{r},1})(C_{\mathrm{r},2} A_{\mathrm{r},2}^j x_{\mathrm{r},2}) && (12)\\
&= (C_{\mathrm{r},1} A_{\mathrm{r},1}^j x_{\mathrm{r},1}) \otimes (C_{\mathrm{r},2} A_{\mathrm{r},2}^j x_{\mathrm{r},2}) && (13)\\
&= (C_{\mathrm{r},1} \otimes C_{\mathrm{r},2})(A_{\mathrm{r},1} \otimes A_{\mathrm{r},2})^j (x_{\mathrm{r},1} \otimes x_{\mathrm{r},2}). && (14)
\end{aligned}$$

An analog relation holds for $j \leq 0$ with the left-sided parameters, which then concludes the proof. ■

Along with the fact that the constant signal 1 is a LSSM signal (generated with $C_\ell = A_\ell = x_\ell = C_\mathrm{r} = A_\mathrm{r} = x_\mathrm{r} = 1$ and denoted by $C = A = x = \mathbb{1}$), this proposition implies that the set of LSSM signals is a ring. This property will be extremely useful in this paper.

III. Recursive Signal Estimation Beyond Least Squares

Let $y = (y_1, \ldots, y_K) \in \mathbb{R}^K$ be discrete-time observations of duration $K \in \mathbb{N}$. At any given time index $k \in \{1, \ldots, K\}$, we wish to fit a LSSM signal $s(x)$ with parameters $\{C, A, x\}$ and unknown x to the observations. It is well-known that the squared error function

$$J_k(x) = \sum_{i=1}^{K} \big(y_i - s_{i-k}(x)\big)^2 \tag{15}$$

can be computed efficiently with recursions as in Kalman filtering. Indeed, using the function in (2) and Propositions 1 & 2, we have

$$\begin{aligned}
J_k(x) &= \xi_k(y^2, \mathbb{1}, \mathbb{1}) - 2x^\mathsf{T} \xi_k(y, C, A)\\
&\quad + (x \otimes x)^\mathsf{T} \xi_k(y^0, C \otimes C, A \otimes A), && (16)
\end{aligned}$$

where y^p denotes the signal y raised element-wise to the power of $p \in \mathbb{N}$. We now generalize the squared cost in (15) in two different ways while preserving recursive cost computations, and thus, computational efficiency.

A. Time-Dependent Polynomial Cost

Assume that each observation y_i, $i \in \{1, \ldots, K\}$, comes with its own polynomial cost P_i of maximum degree $d \in \mathbb{N}$ and coefficients $p_j^{(i)} \in \mathbb{R}$, $j \in \{0, \ldots, d\}$. Suppose we wish to perform signal estimation at any time k by minimizing the cost function

$$J_k(x) = \sum_{i=1}^{K} P_i\big(y_i - s_{i-k}(x)\big). \tag{17}$$

The polynomials P_i will normally be chosen such that $P_i(u) \geq 0$, for all $u \in \mathbb{R}$, but this is actually not a restriction.

An important special case of (17) consists of polynomials P_i independent of i (i.e., $P_i = P$, for all i), which leads to

$$J_k(x) = \sum_{i=1}^{K} P\big(y_i - s_{i-k}(x)\big). \tag{18}$$

When $P(u) = u^2$, the cost (18) becomes the one of (15).

It turns out that the cost function (17) can be recursively and efficiently computed thanks to the relation

$$J_k(x) = \sum_{q=0}^{d} (-1)^q (\otimes^q x)^\mathsf{T} \xi_k(\tilde{y}^{(q)}, \otimes^q C, \otimes^q A), \tag{19}$$

with $\tilde{y}^{(q)} \in \mathbb{R}^K$ such that

$$\tilde{y}_i^{(q)} = \sum_{j=q}^{d} \binom{j}{q} p_j^{(i)} y_i^{j-q}, \tag{20}$$

for $i \in \{1, \ldots, K\}$, $q \in \{0, \ldots, d\}$, and with

$$\otimes^q C = \underbrace{C \otimes \cdots \otimes C}_{q \text{ times}}, \tag{21}$$

for $q \geq 1$ and $\otimes^0 C = \mathbb{1}$. The proof of (19) follows from the proof of (23), which is given in the next section. In particular, (19) coincides with (16) when $P_i(u) = u^2$, for all i.

B. LSSM-Windowed and Time-Dependent Polynomial Cost

Often, we further want to limit the horizon of the actual signal estimation. For that purpose, we localize the cost function (17) using a LSSM window w_j, $j \in \mathbb{Z}$, centered at time index k. Thus, given the polynomial costs P_i, $i \in \{1, \ldots, K\}$, of maximum degree $d \in \mathbb{N}$, we wish to perform signal estimation at any time k by minimizing the windowed cost function

$$J_k(x) = \sum_{i=1}^{K} w_{i-k} P_i\big(y_i - s_{i-k}(x)\big), \tag{22}$$

where w_j, $j \in \mathbb{Z}$, is a LSSM signal with fixed parameters $\{C_\mathrm{w}, A_\mathrm{w}, x_\mathrm{w}\}$. For instance, when $C_\mathrm{w} = x_\mathrm{w} = \mathbb{1}$ and $A_{\ell,\mathrm{w}} = A_{\mathrm{r},\mathrm{w}} = \gamma$ for some $\gamma \in (0, 1)$, the cost $J_k(x)$ is computed on a symmetric exponentially-decaying window centered at time k. Another example is a finite-length window, which is obtained by choosing $A_{\ell,\mathrm{w}}$ and $A_{\mathrm{r},\mathrm{w}}$ to be nilpotent. Note that (17) is a special case of (22) with $C_\mathrm{w} = A_\mathrm{w} = x_\mathrm{w} = \mathbb{1}$ (i.e., a constant infinite-length window).

The cost (22) is recursively and efficiently computed thanks to the relation

$$J_k(x) = \sum_{q=0}^{d} (-1)^q \big((\otimes^q x) \otimes x_\mathrm{w}\big)^\mathsf{T} \xi_k^{(q)}, \tag{23}$$

with, for $q \in \{0, \ldots, d\}$,

$$\xi_k^{(q)} = \xi_k(\tilde{y}^{(q)}, (\otimes^q C) \otimes C_\mathrm{w}, (\otimes^q A) \otimes A_\mathrm{w}), \tag{24}$$

where $\tilde{y}^{(q)}$ is defined in (20). The quantity $\xi_k^{(q)}$ is a linear function of $\tilde{y}^{(q)}$ but no longer of y. A graphical representation of formula (23) is given in Fig. 1. This formula follows from

$$J_k(x) = \sum_{i=1}^{K} w_{i-k} \sum_{j=0}^{d} p_j^{(i)} \big(y_i - s_{i-k}(x)\big)^j \tag{25}$$

$$= \sum_{j=0}^{d} \sum_{q=0}^{j} \binom{j}{q} (-1)^q \sum_{i=1}^{K} p_j^{(i)} y_i^{j-q} w_{i-k} (s_{i-k}(x))^q \tag{26}$$

$$= \sum_{q=0}^{d} (-1)^q \sum_{i=1}^{K} \sum_{j=q}^{d} \binom{j}{q} p_j^{(i)} y_i^{j-q} w_{i-k} (s_{i-k}(x))^q \tag{27}$$

$$= \sum_{q=0}^{d} (-1)^q \sum_{i=1}^{K} \tilde{y}_i^{(q)} w_{i-k} (s_{i-k}(x))^q, \tag{28}$$

from which we deduce (23) using Proposition 1 since for any $q \in \{0, \ldots, d\}$, $w_{i-k}(s_{i-k}(x))^q$, $i \in \mathbb{Z}$, is a LSSM signal shifted by a time k and with LSSM parameters $\{(\otimes^q C) \otimes C_\mathrm{w}, (\otimes^q A) \otimes A_\mathrm{w}, (\otimes^q x) \otimes x_\mathrm{w}\}$ according to Proposition 2.

The formula (23) also proves that $\{\xi_k^{(q)} : q \in \{0, \ldots, d\}\}$ is a finite-dimensional sufficient statistic for x. The complexity of computing J_k, for all $k \in \{1, \ldots, K\}$, is only of $O(K \underline{n}^{2d} \underline{n}_\mathrm{w}^2)$ (with $\underline{n} = \max(n_\ell, n_\mathrm{r})$ and $\underline{n}_\mathrm{w} = \max(n_{\ell,\mathrm{w}}, n_{\mathrm{r},\mathrm{w}})$), which basically corresponds to the complexity of computing $\xi_k^{(d)}$. In particular, whether the polynomials P_i are time-dependent or not, the computational complexity remains of the same order. Note also the squared dependency of the complexity with respect to the order of the LSSM window.

The cost function (23) is a multivariate polynomial in x. Thus, its minimization can be done using exact algebraic methods such as Gröbner bases or using a relaxation method such as a sum of square formulation solved by semidefinite programming [7].

IV. Signal Estimation in the Presence of Interferences

In several practical applications, signals of interest are altered by some interference signal, which needs to be taken into consideration in the signal estimation problem. For that purpose, in addition to modeling a signal of interest with a LSSM with parameters $\{C, A, x\}$ with unknown x, we also model interferences with a LSSM signal $g_j(x_\mathrm{g})$, $j \in \mathbb{Z}$, with parameters $\{C_\mathrm{g}, A_\mathrm{g}, x_\mathrm{g}\}$ with unknown x_g. In the following, we propose three cost functions $J_k(x, x_\mathrm{g})$ based on (22), which handle interference signals in different ways. Actually, except computational complexity, nothing prevents from combining these three ways of dealing with interferences.

A. Additive Interference

When the interference signal is additive, signal estimation at any time k can be done by minimizing the cost function

$$J_k(x, x_\mathrm{g}) = \sum_{i=1}^{K} w_{i-k} P_i\big(y_i - (g_{i-k}(x_\mathrm{g}) + s_{i-k}(x))\big). \tag{29}$$

Since $g_j(x_\mathrm{g}) + s_j(x)$, $j \in \mathbb{Z}$, is also a LSSM signal obtained by stacking (in a suitable manner) the LSSM parameters $\{C, A, x\}$ with $\{C_\mathrm{g}, A_\mathrm{g}, x_\mathrm{g}\}$, the relation (23) still applies to compute $J_k(x, x_\mathrm{g})$ by replacing the parameters C and A in (24) with the parameters of the stacked model.

B. Multiplicative Interference

In case of multiplicative interference, signal estimation at any time k can be done by minimizing the cost function

$$J_k(x, x_\mathrm{g}) = \sum_{i=1}^{K} w_{i-k} P_i\big(y_i - g_{i-k}(x_\mathrm{g}) s_{i-k}(x)\big). \tag{30}$$

Since $g_j(x_\mathrm{g}) s_j(x)$, $j \in \mathbb{Z}$, is also a LSSM signal (according to Proposition 2) with LSSM parameters $\{C \otimes C_\mathrm{g}, A \otimes A_\mathrm{g}, x \otimes x_\mathrm{g}\}$, the relation (23) still applies to compute $J_k(x, x_\mathrm{g})$ with the substitution $C \leftarrow C \otimes C_\mathrm{g}$, $A \leftarrow A \otimes A_\mathrm{g}$, and $x \leftarrow x \otimes x_\mathrm{g}$.

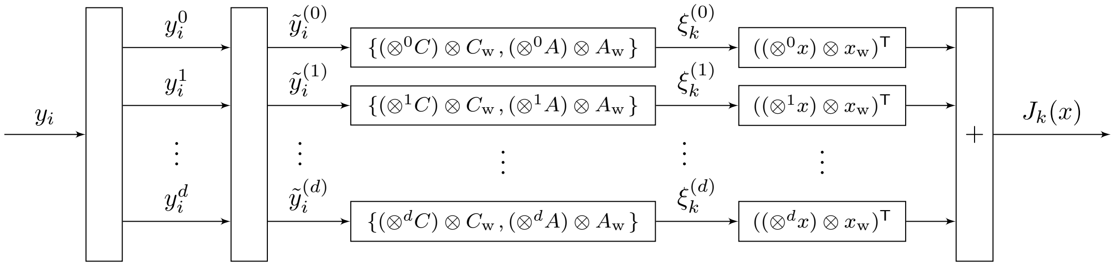

Fig. 1. Graphical representation of the cost computation according to (23) and using (20), (24), and (2).

C. Inverse Multiplicative Interference

In some cases, a multiplicative interference is better modeled as the (element-wise) inverse of a LSSM signal (i.e., $1/g_j(x_\mathrm{g})$, $j \in \mathbb{Z}$) rather than a LSSM signal directly. Unfortunately, the element-wise inverse of a LSSM signal is not in general a LSSM signal. However, instead of multiplying the signal of interest with the interference signal in the cost function, an alternative is to multiply the observations with the element-wise inverse interference. Thus, signal estimation at any time k can be done by minimizing the cost function

$$J_k(x, x_\mathrm{g}) = \sum_{i=1}^{K} w_{i-k} P_i\big(g_{i-k}(x_\mathrm{g}) y_i - s_{i-k}(x)\big). \quad (31)$$

Again, $J_k(x, x_\mathrm{g})$ can still be recursively computed. Indeed, using similar expansions as in (25) and (26), we have

$$\begin{aligned} J_k(x) &= \sum_{j=0}^{d} \sum_{q=0}^{j} (-1)^q \\ &\quad \sum_{i=1}^{K} \binom{j}{q} p_j^{(i)} y_i^{j-q} (g_{i-k}(x_\mathrm{g}))^{j-q} w_{i-k} (s_{i-k}(x))^q \quad (32) \\ &= \sum_{j=0}^{d} \sum_{q=0}^{j} (-1)^q \big(\tilde{x}^{(q,j)}\big)^\mathsf{T} \xi_k\big(\tilde{y}^{(q,j)}, \tilde{C}^{(q,j)}, \tilde{A}^{(q,j)}\big), \quad (33) \end{aligned}$$

with $\tilde{y}^{(q,j)} \in \mathbb{R}^K$ and such that for $(q, j) \in \{0, \ldots, d\}^2$, $q \le j$,

$$\tilde{y}_i^{(q,j)} = \binom{j}{q} p_j^{(i)} y_i^{j-q}, \;\; i \in \{1, \ldots, K\} \quad (34)$$

$$\tilde{C}^{(q,j)} = (\otimes^q C) \otimes (\otimes^{j-q} C_\mathrm{g}) \otimes C_\mathrm{w} \quad (35)$$

$$\tilde{A}^{(q,j)} = (\otimes^q A) \otimes (\otimes^{j-q} A_\mathrm{g}) \otimes A_\mathrm{w} \quad (36)$$

$$\tilde{x}^{(q,j)} = (\otimes^q x) \otimes (\otimes^{j-q} x_\mathrm{g}) \otimes x_\mathrm{w}. \quad (37)$$

Once more, using Proposition 2, those parameters follow from the fact that $(g_{i-k}(x_\mathrm{g}))^{j-q} w_{i-k} (s_{i-k}(x))^q$, $i \in \mathbb{Z}$, (cf. (32)) is a LSSM signal shifted by a time k and with parameters $\{\tilde{C}^{(q,j)}, \tilde{A}^{(q,j)}, \tilde{x}^{(q,j)}\}$.

V. Experimental Results

A. Detection of a Modulated Signal

We want to detect pulses of sinusoidal shape of frequency $\Omega \in \mathbb{R}_+$ in an amplitude-modulated carrier signal of frequency $\Omega_g \in \mathbb{R}_+$ and buried with additive white Gaussian noise. A typical observed signal is displayed in Fig. 2, upper plot.

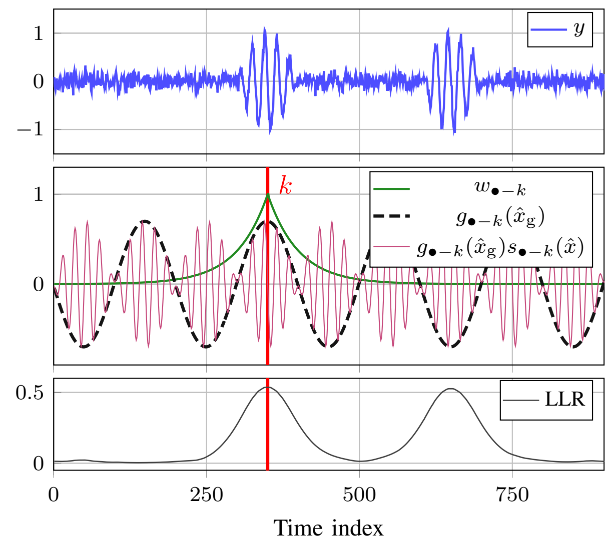

Fig. 2. Synthetic example of amplitude-modulated pulse detection.

For $\omega \in \mathbb{R}$, we denote

$$R(\omega) = \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix} \in \mathbb{R}^{2\times 2}. \quad (38)$$

The carrier signal is seen as an interference signal $g(x_\mathrm{g})$ with LSSM parameters $C_{\mathrm{r},g} = C_{\ell,g} = [1, 0]$, $A_{\mathrm{r},g} = A_{\ell,g}^{-1} = R(\Omega_g)$, and unknown $x_\mathrm{g} \in \mathbb{R}^2 \times \mathbb{R}^2$. The signal of interest $s(x)$, consisting of pulses of sinusoidal shape, is generated with the LSSM parameters $C_\mathrm{r} = C_\ell = [1, 0]$, $A_\mathrm{r} = A_\ell^{-1} = R(\Omega)$, and unknown $x \in \mathbb{R}^2 \times \mathbb{R}^2$. For all time indices k, we recursively compute a cost $J_k(x, x_\mathrm{g})$ as in (30) using $P_i(u) = u^2$, for all i (i.e., standard squared error), but with a two-sided exponential window with parameters $C_\mathrm{w} = x_\mathrm{w} = \mathbb{1}$ and $A_{\ell,\mathrm{w}} = A_{\mathrm{r},\mathrm{w}} = \gamma$ for some $\gamma \in (0, 1)$. In Fig. 2, we illustrate the results of our signal estimation method. In the lower plot, we display the log-likelihood ratio

$$\mathrm{LLR}_k = -\frac{1}{2} \ln\left(\frac{\min_{x, x_\mathrm{g}} J_k(x, x_\mathrm{g})}{J_k(x = 0, x_\mathrm{g} = 0)}\right), \quad (39)$$

which indicates how likely the presence of a signal of interest is. In the middle part of Fig. 2, we plot the estimated signal obtained at index $k = 350$ where LLR_k is maximum. Note the actual separation of the carrier signal from the signal of interest.

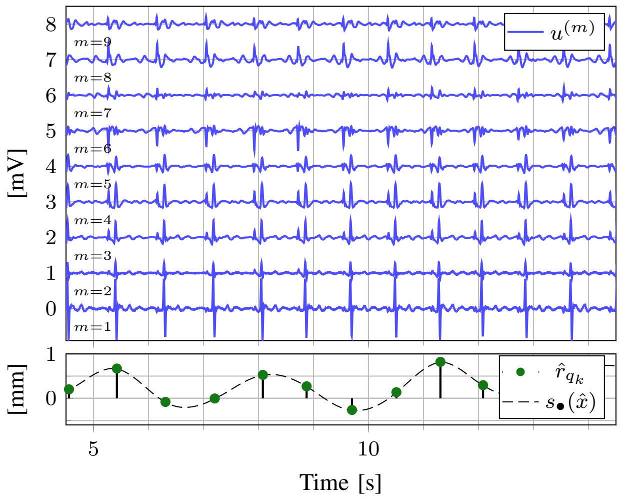

Fig. 3. Multi-channel esophageal ECG signal (blue lines), catheter displacement estimate (green dots) and estimated LSSM signal (black dashed line).

B. Estimation of a Catheter Movement

Unlike surface electrocardiogram (ECG) recordings, esophageal ECG recordings, obtained using electrodes placed in the esophagus, are not commonly used. However, since the esophagus is much closer to the heart than the chest surface, esophageal ECG recordings contain valuable information provided that we can actually extract it. In such recordings the catheter containing the electrodes typically moves due to, among others, the patient's breathing. Given a M-channel esophageal ECG recording, we want to estimate the relative vertical movement of the catheter in the esophagus. The key idea to exploit is that when the catheter slowly moves, the signal shape produced by a heart beat in a given channel is quite similar to the signal shape produced by the previous heart beat in another channel [8].

We consider a catheter that holds $M+1$ ring-shaped electrodes which are located at distances $d_0 < d_1 < \ldots < d_M$ from the catheter tip. For $m \in \{1, \ldots, M\}$, let $u_n^{(m)} \in \mathbb{R}$, $n \in \{1, \ldots, N\}$ be the voltage measured between electrodes m and $m-1$. Within these N samples, we observe $K+1$ heart beats with corresponding R peaks at time indices $q_k \in \mathbb{N}$ and corresponding unknown catheter positions $r_{q_k} \in \mathbb{R}$, $k \in \{0, \ldots, K\}$. The first beat (i.e., for $k=0$) is considered as reference beat with $r_{q_0} = 0$. Each beat k effectively produces signal shapes in the interval $\{q_k + a, \ldots, q_k + b\}$, $(a, b) \in \mathbb{Z}^2$, $a < b$. In order to compare signal shapes produced by the i^{th} heart beat with the ones produced by the reference beat, we define the cost function

$$P_i(r) = \sum_{n=a}^{b} \int_{d_0}^{d_M} \left(\varphi_{q_i+n}(z-r) - \varphi_{q_0+n}(z)\right)^2 \mathrm{d}z, \tag{40}$$

which depends on the displacement r and where $\varphi_n(z)$, $z \in [d_0, d_M]$, is a real polynomial of degree $Q \in \mathbb{N}$, which interpolates the voltage measurements $(u_n^{(1)}, \ldots, u_n^{(M)})$ across channels at time index n. It follows that $P_i(r)$ (plotted in Fig. 4) is also a polynomial in r of degree $2Q+1$.

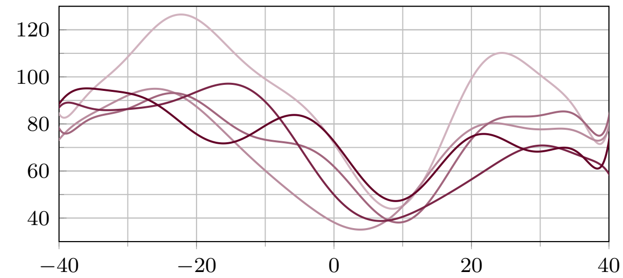

Fig. 4. Selection of polynomials $P_i(r)$ for few indices i.

Then, we model the displacement of the catheter with a LSSM signal $s(x)$ with parameters C and A. Finally, we perform signal estimation by minimizing the cost function

$$J_k(x) = \sum_{i=1}^{K} w_{i-k} P_i(s_{i-k}(x)), \tag{41}$$

where w_{i-k} corresponds to an exponential window, as an example of (22). Fig. 3 shows a snipped of a multi-channel esophageal ECG signal ($M=9$) and the estimated catheter position. For this example, we chose $Q=5$, $C_\mathrm{r} = C_\ell = [1, 0]$, and $A_\mathrm{r} = A_\ell^{-1} = R(\omega)$. Note that the catheter position shows a periodic movement, revealing the subject's breathing activity.

VI. Conclusion

We have introduced a general cost function, which includes sample-dependent polynomial costs along with window weights generated by a LSSM, to fit a parameterized LSSM signal to discrete-time measurements. We have shown how this cost function can be recursively computed and can handle unknown additive and/or multiplicative interferences in a signal estimation problem. The two properties we used are that the correlation function between any given discrete-time signal and a LSSM signal is recursively computed and the set of LSSM signals is a ring. We have also presented two applications, which, however, hardly suffice to illustrate the versatility of the proposed approach.

References

[1] C.-F. Gauss, *Theoria combinationis observationum erroribus minimis obnoxiae*. Henricus Dieterich, Gottingae, 1823.

[2] S. Haykin, *Adaptive filter theory*. Pearson Education India, 2008.

[3] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, "The factor graph approach to model-based signal processing," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1295–1322, 2007.

[4] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[5] F. Daum, "Nonlinear filters: beyond the Kalman filter," *IEEE Aerospace and Electronic Systems Magazine*, vol. 20, no. 8, pp. 57–69, 2005.

[6] V. Beneš, "Exact finite-dimensional filters for certain diffusions with nonlinear drift," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 5, no. 1-2, pp. 65–92, 1981.

[7] P. A. Parrilo and B. Sturmfels, "Minimizing polynomial functions," *Algorithmic and quantitative real algebraic geometry, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, vol. 60, pp. 83–99, 2003.

[8] D. Bruegger, "3D Reconstruction and Simulation of Heart Potentials in the Esophagus," Master's thesis, Bern University of Applied Sciences, Biel, Switzerland, 2016.