Kalman Filtering in Triplet Markov Chains

Boujemaa Ait-El-Fquih and François Desbouvries, *Member, IEEE*

***Abstract*—Let $\mathrm{x} = \{\mathrm{x}_n\}_{n\in\mathbb{N}}$ be a hidden process, $\mathrm{y} = \{\mathrm{y}_n\}_{n\in\mathbb{N}}$ an observed process, and $\mathrm{r} = \{\mathrm{r}_n\}_{n\in\mathbb{N}}$ some additional process. We assume that $\mathrm{t} = (\mathrm{x},\mathrm{r},\mathrm{y})$ is a (so-called "Triplet") vector Markov chain (TMC). We first show that the linear TMC model encompasses and generalizes, among other models, the classical state-space systems with colored process and/or measurement noise(s). We next propose restoration Kalman-like filters for arbitrary linear Gaussian (LG) TMC.**

***Index Terms*—Bayesian signal restoration, hidden Markov chains, Kalman filtering, Markovian models, triplet Markov chains.**

I. Introduction

LET us consider the classical linear dynamical stochastic system

$$\begin{cases} \mathbf{x}_{n+1} = \mathbf{F}_n\mathbf{x}_n + \mathbf{G}_n\mathbf{u}_n \\ \mathbf{y}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{J}_n\mathbf{v}_n \end{cases} \tag{1}$$

in which $\mathbf{x}_n \in \mathbb{R}^{n_\mathbf{x}}$ is the state, $\mathbf{y}_n \in \mathbb{R}^{n_\mathbf{y}}$ is the observation, $\mathbf{u}_n \in \mathbb{R}^{n_\mathbf{u}}$ is the input noise and $\mathbf{v}_n \in \mathbb{R}^{n_\mathbf{v}}$ is the measurement noise. The processes $\mathbf{u} = \{\mathbf{u_n}\}_{n\in\mathbb{N}}$ and $\mathbf{v} = \{\mathbf{v_n}\}_{n\in\mathbb{N}}$ are assumed to be independent,[1] jointly independent and independent of $\mathbf{x}_0$.

Let $\mathbf{x}_{0:n} = \{\mathbf{x}_i\}_{i=0}^n$ and $\mathbf{y}_{0:n} = \{\mathbf{y}_i\}_{i=0}^n$. Let also $p(\mathbf{x}_n)$, $p(\mathbf{x}_{0:n})$ and $p(\mathbf{x}_n|\mathbf{y}_{0:n})$, say, denote the probability density function (pdf) (with regard to Lebesgue measure) of $\mathbf{x}_n$, the pdf of $\mathbf{x}_{0:n}$, and the pdf of $\mathbf{x}_n$, conditional on $\mathbf{y}_{0:n}$, respectively; the other pdf's are defined similarly. A fundamental problem associated with model (1) (the so-called filtering problem) consists in computing the posterior pdf $p(\mathbf{x}_n|\mathbf{y}_{0:n})$. From (1), we get

$$p(\mathbf{x}_{n+1}|\mathbf{x}_{0:n}) = p(\mathbf{x}_{n+1}|\mathbf{x}_n) \tag{2}$$

$$p(\mathbf{y}_{0:n}|\mathbf{x}_{0:n}) = \prod_{i=0}^{n} p(\mathbf{y}_i|\mathbf{x}_{0:n}) \tag{3}$$

$$p(\mathbf{y}_i|\mathbf{x}_{0:n}) = p(\mathbf{y}_i|\mathbf{x}_i) \text{ for all } i,\ 0 \le i \le n. \tag{4}$$

Now from (2) to (4) we get

$$p(\mathbf{x}_{n+1}|\mathbf{x}_{0:n},\mathbf{y}_{0:n}) = p(\mathbf{x}_{n+1}|\mathbf{x}_n), \tag{5}$$

$$p(\mathbf{y}_{n+1}|\mathbf{x}_{0:n+1},\mathbf{y}_{0:n}) = p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1}), \tag{6}$$

and consequently $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ can be computed recursively as

$$p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n+1}) = \frac{p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1})\int p(\mathbf{x}_{n+1}|\mathbf{x}_n)p(\mathbf{x}_n|\mathbf{y}_{0:n})d\mathbf{x}_n}{\int p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1})\left[\int p(\mathbf{x}_{n+1}|\mathbf{x}_n)p(\mathbf{x}_n|\mathbf{y}_{0:n})d\mathbf{x}_n\right]d\mathbf{x}_{n+1}}. \tag{7}$$

If furthermore $\mathbf{x}_0$ and $\mathbf{w}_n = (\mathbf{u}_n,\mathbf{v}_n)$ are Gaussian, then $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ is also Gaussian and is thus described by its mean and covariance matrix. Propagating $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ amounts to propagating its parameters, and (7) reduces to the celebrated Kalman filter (KF) [1] (see also [2]–[4]).

A. Extensions of the KF

Since this pioneering work, the KF has been generalized in many directions. To name just a few examples, robust (i.e., square-root type) or fast (i.e., Chandrasekhar type) algorithms have been proposed; smoothing and prediction algorithms have been developed; the independence assumptions on $\mathbf{u}$ and $\mathbf{v}$ have been dropped; and the extension of (1) to nonlinear and/or non-Gaussian systems has been addressed, leading to approximate solutions such as the extended KF or particle filters. The literature on these extensions is vast, and the interested reader may wish to consult, for instance, [3], [5]–[8], as well as the references therein.

On the other hand, yet another direction in which one can extend the KF consists in releasing some conditional independence assumptions on $\mathbf{x}$ and $\mathbf{y}$. As we have seen, if (1) holds then (2) to (4) hold. In other words, $\mathbf{x}$ is a Markov chain (MC), and since it is known only through the observed process $\mathbf{y}$, (1) is a hidden Markov chain (HMC). Now if (2)–(4) hold, then

$$p(\mathbf{x}_{n+1},\mathbf{y}_n|\mathbf{x}_{0:n},\mathbf{y}_{0:n-1}) = p(\mathbf{x}_{n+1},\mathbf{y}_n|\mathbf{x}_n,\mathbf{y}_{n-1}) \tag{8}$$

i.e. the pair $\{(\mathbf{x}_{n+1},\mathbf{y}_n)\}$ is a (vector) MC, so any HMC is also a so-called "pairwise" Markov chain (PMC). The converse is not true, because if $\{(\mathbf{x}_{n+1},\mathbf{y}_n)\}$ is a vector MC then the marginal process $\{\mathbf{x}_n\}$ is not necessarily an MC, nor does (3) or (4) hold [9]. However, it is still possible to derive restoration algorithms for PMC; from (8), (5) and (6) generalize to

$$p(\mathbf{x}_{n+1}|\mathbf{x}_{0:n},\mathbf{y}_{0:n}) = p(\mathbf{x}_{n+1}|\mathbf{x}_n,\mathbf{y}_n,\mathbf{y}_{n-1}) \tag{9}$$

$$p(\mathbf{y}_{n+1}|\mathbf{x}_{0:n+1},\mathbf{y}_{0:n}) = p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1},\mathbf{y}_n) \tag{10}$$

so (7) becomes (11), shown at the bottom of the next page. In the linear-Gaussian (LG) case, (11) reduces to an algorithm that extends the KF [9]; a particle filtering solution for computing (11) has also been proposed for the general case [10].

Manuscript received January 10, 2005; revised September 6, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jonathon Chambers.

The authors are with the Institut National des Télécommunication Département CITI and CNRS UMR 5157, 91011 Evry, France (e-mail: boujemaa.ait_elfquih@int-evry.fr; francois.desbouvries@int-evry.fr).

Digital Object Identifier 10.1109/TSP.2006.877651

[1] i.e., $\mathbf{u}$ (and similarly $\mathbf{v}$) contains independent variables.

The PMC model can be further generalized to the TMC model [11]–[13], which we now recall. Let $\mathbf{r} = \{\mathbf{r}_n\}_{n\in\mathbb{N}}$ be an additional (possibly artificial) process, and let $\mathbf{t}_n = (\mathbf{x}_n, \mathbf{r}_n, \mathbf{y}_{n-1})$. We say that $\mathbf{t} = \{\mathbf{t}_n\}_{n\in\mathbb{N}}$ is a TMC if $\mathbf{t}$ is a (vector) MC. The interest of TMC is twofold.

1) As far as restoration is concerned, the TMC $(\mathbf{r}, \mathbf{x}, \mathbf{y})$ can be viewed as the PMC $((\mathbf{r}, \mathbf{x}), \mathbf{y})$; so $\mathbf{x}^* = (\mathbf{r}, \mathbf{x})$ can be restored from $\mathbf{y}$ by a PMC algorithm, and finally $\mathbf{x}$ is obtained by marginalization;
2) As far as modelling is concerned, TMC generalize some classical models (including PMC) in the sense that none of the chains $\mathbf{x}$, $\mathbf{r}$, $\mathbf{y}$, $(\mathbf{x}, \mathbf{r})$, $(\mathbf{x}, \mathbf{y})$ or $(\mathbf{r}, \mathbf{y})$ needs to be an MC. On the other hand, in an HMC $p(\mathbf{y}|\mathbf{x})$ is given by (3) and (4), which is too simple in some applications, like e.g., speech recognition [14], [15]; whereas in a TMC, $p(\mathbf{y}|\mathbf{x})$ is the marginal pdf of the MC $p(\mathbf{r}, \mathbf{y}|\mathbf{x})$, and, as such, can be rather complex. In practice, computer experiments have demonstrated the superiority of PMC [16] (respectively, TMC [17]) over HMC in the context of image segmentation.

B. Contributions

Let us turn to the contents of this note. In Section II, we first show that the classical LG models with colored process and/or measurement noise(s) [3], [6], [7], [18]–[22] are, among other models, some important particular cases (mostly with unnoisy measurements) of the linear TMC model; so the triplet model, which initially was designed as an extension of (1), happens also to encompass (and generalize) some of the early classical generalizations (those which deal with the nature of $\mathbf{u}$ and $\mathbf{v}$) of model (1). Now, the restoration algorithm of [12] was designed for regular LG-TMC. In the singular measurements case, it is possible to reduce the dimension of the state vector to be estimated. We thus propose in Section III a restoration algorithm for LG-TMC with singular measurements, which encompasses (and generalizes) some existing algorithms, and we perform some simulations.

II. Linear TMC Model: Definition and Applications

A. Linear TMC Model

Let $\mathbf{x}_n \in \mathbb{R}^{n_\mathbf{x}}$ be the hidden process, $\mathbf{y}_n \in \mathbb{R}^{n_\mathbf{y}}$ the observation, and $\mathbf{r}_n \in \mathbb{R}^{n_\mathbf{r}}$ an additional process. For $n > 0$, let $\mathbf{t}_n = (\mathbf{x}_n, \mathbf{r}_n, \mathbf{y}_{n-1})$, and let $\mathbf{t}_0 = (\mathbf{x}_0, \mathbf{r}_0, \mathbf{0})$. We say that $\mathbf{t} = \{\mathbf{t}_n\}_{n\in\mathbb{N}}$ is a TMC if $\mathbf{t}$ is a (vector) MC, and that $\mathbf{t}$ is a linear TMC if furthermore $\mathbf{t}_n$ satisfies the system, as follows:

$$\underbrace{\begin{bmatrix}\mathbf{x}_{n+1}\\ \mathbf{r}_{n+1}\\ \mathbf{y}_n\end{bmatrix}}_{\mathbf{t}_{n+1}} = \underbrace{\begin{bmatrix}\mathcal{F}_n^{\mathbf{x},\mathbf{x}} & \mathcal{F}_n^{\mathbf{x},\mathbf{r}} & \mathcal{F}_n^{\mathbf{x},\mathbf{y}}\\ \mathcal{F}_n^{\mathbf{r},\mathbf{x}} & \mathcal{F}_n^{\mathbf{r},\mathbf{r}} & \mathcal{F}_n^{\mathbf{r},\mathbf{y}}\\ \mathcal{F}_n^{\mathbf{y},\mathbf{x}} & \mathcal{F}_n^{\mathbf{y},\mathbf{r}} & \mathcal{F}_n^{\mathbf{y},\mathbf{y}}\end{bmatrix}}_{\mathcal{F}_n} \begin{bmatrix}\mathbf{x}_n\\ \mathbf{r}_n\\ \mathbf{y}_{n-1}\end{bmatrix} + \underbrace{\begin{bmatrix}\mathbf{w}_n^{\mathbf{x}}\\ \mathbf{w}_n^{\mathbf{r}}\\ \mathbf{w}_n^{\mathbf{y}}\end{bmatrix}}_{\mathbf{w}_n} \tag{12}$$

where $\mathcal{F}_n$ is deterministic, and $\mathbf{w} = \{\mathbf{w}_n\}_{n\in\mathbb{N}}$ is zero mean, independent, and independent of $\mathbf{t}_0$.

B. Some Particular Linear TMC Models

Before proceeding to restoration algorithms, let us first illustrate the wide applicability of the linear TMC model by noticing that it provides a common framework for some linear stochastic systems (those of Sections II-B-1)to II-B-3) are classical, the others are new). They all differ from one another by the matrices $\mathcal{F}_n$ (some submatrices of which are equal to zero); the physical meaning of $\mathbf{r} = \{\mathbf{r}_n\}_{n\in\mathbb{N}}$; and/or independence assumptions among subvectors of $\mathbf{w}_n$.

1) Autoregressive Process Noise: The case where in (1) $\mathbf{v}$ remains independent but $\mathbf{u}$ becomes an MC has been introduced in [6] (see also [3]). Let $\mathbf{u}_{n+1} = \mathbf{A}_n^{\mathbf{u},\mathbf{u}}\mathbf{u}_n + \boldsymbol{\xi}_n^{\mathbf{u}}$, in which $\boldsymbol{\xi}^{\mathbf{u}} = \{\boldsymbol{\xi}_n^{\mathbf{u}}\}_{n\in\mathbb{N}}$ is zero mean, independent and independent of $\mathbf{u}_0$. Then $\{\mathbf{x}_n\}$ is no longer an MC, but $\{(\mathbf{x}_n, \mathbf{u}_n)\}$ is an MC. The whole $(\mathbf{x}_n, \mathbf{r}_n = \mathbf{u}_n, \mathbf{y}_{n-1})$ model can be rewritten as

$$\begin{bmatrix}\mathbf{x}_{n+1}\\ \mathbf{u}_{n+1}\\ \mathbf{y}_n\end{bmatrix} = \begin{bmatrix}\mathbf{F}_n & \mathbf{G}_n & \mathbf{0}_{n_\mathbf{x}\times n_\mathbf{y}}\\ \mathbf{0}_{n_\mathbf{u}\times n_\mathbf{x}} & \mathbf{A}_n^{\mathbf{u},\mathbf{u}} & \mathbf{0}_{n_\mathbf{u}\times n_\mathbf{y}}\\ \mathbf{H}_n & \mathbf{0}_{n_\mathbf{y}\times n_\mathbf{u}} & \mathbf{0}_{n_\mathbf{y}\times n_\mathbf{y}}\end{bmatrix} \begin{bmatrix}\mathbf{x}_n\\ \mathbf{u}_n\\ \mathbf{y}_{n-1}\end{bmatrix} + \begin{bmatrix}\mathbf{0}_{n_\mathbf{x}\times 1}\\ \boldsymbol{\xi}_n^{\mathbf{u}}\\ \mathbf{J}_n\mathbf{v}_n\end{bmatrix} \tag{13}$$

in which $\{\mathbf{w}_n^{\mathbf{r}} = \boldsymbol{\xi}_n^{\mathbf{u}}\}_{n\in\mathbb{N}}$ and $\{\mathbf{w}_n^{\mathbf{y}} = \mathbf{J}_n\mathbf{v}_n\}_{n\in\mathbb{N}}$ are independent.

2) Autoregressive Measurement Noise: The case where in (1) $\mathbf{u}$ remains independent but $\mathbf{v}$ becomes an MC has been addressed in [20] in case $\mathbf{J}_n$ is the $n_\mathbf{y} \times n_\mathbf{y}$ identity matrix $\mathbf{I}_{n_\mathbf{y}}$ (see also [3, sec. 11.2], [5, pp. 212–215]), then generalized in [6] (see also [3] and [7]). Let $\mathbf{v}_{n+1} = \mathbf{A}_n^{\mathbf{v},\mathbf{v}}\mathbf{v}_n + \boldsymbol{\xi}_n^{\mathbf{v}}$, with $\boldsymbol{\xi}^{\mathbf{v}} = \{\boldsymbol{\xi}_n^{\mathbf{v}}\}_{n\in\mathbb{N}}$ zero mean, independent, and independent of $\mathbf{v}_0$. Then, the $(\mathbf{x}_n, \mathbf{r}_n = \mathbf{v}_n, \mathbf{y}_{n-1})$ model reads

$$\begin{bmatrix}\mathbf{x}_{n+1}\\ \mathbf{v}_{n+1}\\ \mathbf{y}_n\end{bmatrix} = \begin{bmatrix}\mathbf{F}_n & \mathbf{0}_{n_\mathbf{x}\times n_\mathbf{v}} & \mathbf{0}_{n_\mathbf{x}\times n_\mathbf{y}}\\ \mathbf{0}_{n_\mathbf{v}\times n_\mathbf{x}} & \mathbf{A}_n^{\mathbf{v},\mathbf{v}} & \mathbf{0}_{n_\mathbf{v}\times n_\mathbf{y}}\\ \mathbf{H}_n & \mathbf{J}_n & \mathbf{0}_{n_\mathbf{y}\times n_\mathbf{y}}\end{bmatrix} \begin{bmatrix}\mathbf{x}_n\\ \mathbf{v}_n\\ \mathbf{y}_{n-1}\end{bmatrix} + \begin{bmatrix}\mathbf{G}_n\mathbf{u}_n\\ \boldsymbol{\xi}_n^{\mathbf{v}}\\ \mathbf{0}_{n_\mathbf{y}\times 1}\end{bmatrix} \tag{14}$$

in which $\{\mathbf{w}_n^{\mathbf{x}} = \mathbf{G}_n\mathbf{u}_n\}_{n\in\mathbb{N}}$ and $\{\mathbf{w}_n^{\mathbf{r}} = \boldsymbol{\xi}_n^{\mathbf{v}}\}_{n\in\mathbb{N}}$ are independent.

$$p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n+1}) = \frac{p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1}, \mathbf{y}_n) \int p(\mathbf{x}_{n+1}|\mathbf{x}_n, \mathbf{y}_n, \mathbf{y}_{n-1}) p(\mathbf{x}_n|\mathbf{y}_{0:n}) d\mathbf{x}_n}{\int p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1}, \mathbf{y}_n) \left[\int p(\mathbf{x}_{n+1}|\mathbf{x}_n, \mathbf{y}_n, \mathbf{y}_{n-1}) p(\mathbf{x}_n|\mathbf{y}_{0:n}) d\mathbf{x}_n\right] d\mathbf{x}_{n+1}}. \tag{11}$$

3) Autoregressive Process and Measurement Noises: In the two last examples, we have removed the independence assumption either on $\mathbf{u}$ or on $\mathbf{v}$. Sorenson [19] introduced a model in which both $\mathbf{u}$ and $\mathbf{v}$ are MC but remain jointly independent (see also [22, ch. 5]). Let

$$\underbrace{\begin{bmatrix}\mathbf{u}_{n+1}\\ \mathbf{v}_{n+1}\end{bmatrix}}_{\mathbf{n}_{n+1}}=\underbrace{\begin{bmatrix}\mathbf{A}_n^{\mathbf{u},\mathbf{u}} & \mathbf{0}\\ \mathbf{0} & \mathbf{A}_n^{\mathbf{v},\mathbf{v}}\end{bmatrix}}_{\mathbf{A}_n}\begin{bmatrix}\mathbf{u}_n\\ \mathbf{v}_n\end{bmatrix}+\underbrace{\begin{bmatrix}\boldsymbol{\xi}_n^{\mathbf{u}}\\ \boldsymbol{\xi}_n^{\mathbf{v}}\end{bmatrix}}_{\boldsymbol{\xi}_n} \quad (15)$$

in which $\boldsymbol{\xi}^{\mathbf{u}}=\{\boldsymbol{\xi}_n^{\mathbf{u}}\}_{n\in\mathbb{N}}$ (respectively, $\boldsymbol{\xi}^{\mathbf{v}}=\{\boldsymbol{\xi}_n^{\mathbf{v}}\}_{n\in\mathbb{N}}$) is zero mean, independent, and independent of $\mathbf{u}_0$ (respectively, of $\mathbf{v}_0$), and $\boldsymbol{\xi}^{\mathbf{u}}$ and $\boldsymbol{\xi}^{\mathbf{v}}$ are independent. Then, the whole $(\mathbf{x}_n,\mathbf{r}_n=\mathbf{n}_n,\mathbf{y}_{n-1})$ model reads

$$\begin{bmatrix}\mathbf{x}_{n+1}\\ \mathbf{n}_{n+1}\\ \mathbf{y}_n\end{bmatrix}=\begin{bmatrix}\mathbf{F}_n & \overline{\mathbf{G}}_n & \mathbf{0}_{n_\mathbf{x}\times n_\mathbf{y}}\\ \mathbf{0}_{n_\mathbf{n}\times n_\mathbf{x}} & \mathbf{A}_n & \mathbf{0}_{n_\mathbf{n}\times n_\mathbf{y}}\\ \mathbf{H}_n & \overline{\mathbf{J}}_n & \mathbf{0}_{n_\mathbf{y}\times n_\mathbf{y}}\end{bmatrix}\begin{bmatrix}\mathbf{x}_n\\ \mathbf{n}_n\\ \mathbf{y}_{n-1}\end{bmatrix}+\begin{bmatrix}\mathbf{0}_{n_\mathbf{x}\times 1}\\ \boldsymbol{\xi}_n\\ \mathbf{0}_{n_\mathbf{y}\times 1}\end{bmatrix} \quad (16)$$

with $\overline{\mathbf{G}}_n=[\mathbf{G}_n,\mathbf{0}_{n_\mathbf{x}\times n_\mathbf{v}}]$ and $\overline{\mathbf{J}}_n=[\mathbf{0}_{n_\mathbf{y}\times n_\mathbf{u}},\mathbf{J}_n]$.

4) Autoregressive Model Noise: In (15), the independence assumption between $\mathbf{u}$ and $\mathbf{v}$ can be relaxed by only assuming that $\boldsymbol{\xi}=\{\boldsymbol{\xi}_n\}_{n\in\mathbb{N}}$ is independent and independent of $\mathbf{n}_0$. A further step in generalizing (15) consists in assuming that

$$\begin{bmatrix}\mathbf{u}_{n+1}\\ \mathbf{v}_{n+1}\end{bmatrix}=\begin{bmatrix}\mathbf{A}_n^{\mathbf{u},\mathbf{u}} & \mathbf{A}_n^{\mathbf{u},\mathbf{v}}\\ \mathbf{A}_n^{\mathbf{v},\mathbf{u}} & \mathbf{A}_n^{\mathbf{v},\mathbf{v}}\end{bmatrix}\begin{bmatrix}\mathbf{u}_n\\ \mathbf{v}_n\end{bmatrix}+\begin{bmatrix}\boldsymbol{\xi}_n^{\mathbf{u}}\\ \boldsymbol{\xi}_n^{\mathbf{v}}\end{bmatrix} \quad (17)$$

with $\boldsymbol{\xi}=\{\boldsymbol{\xi}_n\}_{n\in\mathbb{N}}$ independent and independent of $\mathbf{n}_0$. Notice that neither $\mathbf{u}$ nor $\mathbf{v}$ needs to be an MC any longer. The model can be rewritten as (16) (but with $\mathbf{A}_n$ given by (17)).

5) Correlation Between Measurement Noise and System State: In (14), $\mathbf{x}$ and $\mathbf{v}$ are both MC but remain independent. Another natural generalization consists in assuming that

$$\begin{bmatrix}\mathbf{x}_{n+1}\\ \mathbf{v}_{n+1}\end{bmatrix}=\begin{bmatrix}\mathbf{F}_n^{\mathbf{x},\mathbf{x}} & \mathbf{F}_n^{\mathbf{x},\mathbf{v}}\\ \mathbf{F}_n^{\mathbf{v},\mathbf{x}} & \mathbf{F}_n^{\mathbf{v},\mathbf{v}}\end{bmatrix}\begin{bmatrix}\mathbf{x}_n\\ \mathbf{v}_n\end{bmatrix}+\begin{bmatrix}\mathbf{u}_n^{\mathbf{x}}\\ \mathbf{u}_n^{\mathbf{v}}\end{bmatrix} \quad (18)$$

with $\tilde{\mathbf{u}}_n=[(\mathbf{u}_n^{\mathbf{x}})^T(\mathbf{u}_n^{\mathbf{v}})^T]^T$ independent and independent of $(\mathbf{x}_0,\mathbf{v}_0)$. Neither $\mathbf{x}$ nor $\mathbf{v}$ needs to be an MC any longer. However $\{(\mathbf{x}_n,\mathbf{r}_n=\mathbf{v}_n,\mathbf{y}_{n-1})\}$ as a whole remains a linear TMC, since

$$\begin{bmatrix}\mathbf{x}_{n+1}\\ \mathbf{v}_{n+1}\\ \mathbf{y}_n\end{bmatrix}=\begin{bmatrix}\mathbf{F}_n^{\mathbf{x},\mathbf{x}} & \mathbf{F}_n^{\mathbf{x},\mathbf{v}} & \mathbf{0}_{n_\mathbf{x}\times n_\mathbf{y}}\\ \mathbf{F}_n^{\mathbf{v},\mathbf{x}} & \mathbf{F}_n^{\mathbf{v},\mathbf{v}} & \mathbf{0}_{n_\mathbf{v}\times n_\mathbf{y}}\\ \mathbf{H}_n & \mathbf{J}_n & \mathbf{0}_{n_\mathbf{y}\times n_\mathbf{y}}\end{bmatrix}\begin{bmatrix}\mathbf{x}_n\\ \mathbf{v}_n\\ \mathbf{y}_{n-1}\end{bmatrix}+\begin{bmatrix}\mathbf{u}_n^{\mathbf{x}}\\ \mathbf{u}_n^{\mathbf{v}}\\ \mathbf{0}_{n_\mathbf{y}\times 1}\end{bmatrix}. \quad (19)$$

6) Linear PMC Models: The linear PMC model introduced in [9] reads

$$\underbrace{\begin{bmatrix}\mathbf{x}_{n+1}\\ \mathbf{y}_n\end{bmatrix}}_{\mathbf{z}_{n+1}}=\begin{bmatrix}\mathbf{F}_n^1 & \mathbf{F}_n^2\\ \mathbf{H}_n^1 & \mathbf{H}_n^2\end{bmatrix}\begin{bmatrix}\mathbf{x}_n\\ \mathbf{y}_{n-1}\end{bmatrix}+\begin{bmatrix}\mathbf{G}_n^{11} & \mathbf{G}_n^{12}\\ \mathbf{G}_n^{21} & \mathbf{G}_n^{22}\end{bmatrix}\underbrace{\begin{bmatrix}\mathbf{u}_n\\ \mathbf{v}_n\end{bmatrix}}_{\mathbf{n}_n} \quad (20)$$

with $\mathbf{n}=\{\mathbf{n}_n\}_{n\in\mathbb{N}}$ zero mean, independent, and independent of $\mathbf{z}_0$. Let us relax this independence hypothesis on $\{\mathbf{n}_n\}$ by assuming that (17) holds. Then, (20) becomes

$$\begin{bmatrix}\mathbf{x}_{n+1}\\ \mathbf{n}_{n+1}\\ \mathbf{y}_n\end{bmatrix}=\begin{bmatrix}\mathbf{F}_n^1 & \overline{\mathbf{G}}_n^1 & \mathbf{F}_n^2\\ \mathbf{0}_{n_\mathbf{n}\times n_\mathbf{x}} & \mathbf{A}_n & \mathbf{0}_{n_\mathbf{n}\times n_\mathbf{y}}\\ \mathbf{H}_n^1 & \overline{\mathbf{G}}_n^2 & \mathbf{H}_n^2\end{bmatrix}\begin{bmatrix}\mathbf{x}_n\\ \mathbf{n}_n\\ \mathbf{y}_{n-1}\end{bmatrix}+\begin{bmatrix}\mathbf{0}_{n_\mathbf{x}\times 1}\\ \boldsymbol{\xi}_n\\ \mathbf{0}_{n_\mathbf{y}\times 1}\end{bmatrix} \quad (21)$$

with $\overline{\mathbf{G}}_n^1=[\mathbf{G}_n^{11},\mathbf{G}_n^{12}]$ and $\overline{\mathbf{G}}_n^2=[\mathbf{G}_n^{21},\mathbf{G}_n^{22}]$.

III. LG-TMC: Restoration Algorithms

The aim of this section is to derive an algorithm for computing recursively, in the Gaussian case, $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ in an arbitrary linear TMC (12). Let us first gather the unobserved variables $\mathbf{x}_n$ and $\mathbf{r}_n$ into a common vector $\mathbf{x}_n^*=(\mathbf{x}_n,\mathbf{r}_n)$. Then, (12) can be rewritten more compactly as

$$\begin{bmatrix}\mathbf{x}_{n+1}^*\\ \mathbf{y}_n\end{bmatrix}=\begin{bmatrix}\mathcal{F}_n^{\mathbf{x}^*,\mathbf{x}^*} & \mathcal{F}_n^{\mathbf{x}^*,\mathbf{y}}\\ \mathcal{F}_n^{\mathbf{y},\mathbf{x}^*} & \mathcal{F}_n^{\mathbf{y},\mathbf{y}}\end{bmatrix}\begin{bmatrix}\mathbf{x}_n^*\\ \mathbf{y}_{n-1}\end{bmatrix}+\begin{bmatrix}\mathbf{w}_n^{\mathbf{x}^*}\\ \mathbf{w}_n^{\mathbf{y}}\end{bmatrix}. \quad (22)$$

Let us moreover assume that

$$\mathbf{x}_0^*\sim\mathcal{N}(\widehat{\mathbf{x}}_0^*,\mathbf{P}_0^*),\ \mathbf{w}_n\sim\mathcal{N}\left(\mathbf{0},\underbrace{\begin{bmatrix}\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{x}^*} & \mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}\\ \mathcal{Q}_n^{\mathbf{y},\mathbf{x}^*} & \mathcal{Q}_n^{\mathbf{y},\mathbf{y}}\end{bmatrix}}_{\mathcal{Q}_n}\right). \quad (23)$$

Since $\mathbf{w}_n$ are independent and independent of $\mathbf{t}_0$, $\mathbf{t}$ is Gaussian, and model (22), (23) actually defines a partially observed Gauss-Markov vector process, in which we observe some components $\{\mathbf{y}_n\}$ and we want to restore (part of) the remaining ones $\{\mathbf{x}_n^*\}$. So our restoration algorithm recursively computes $p(\mathbf{x}_n^*|\mathbf{y}_{0:n})$ (or, in the singular measurements case, $p(\overline{\mathbf{x}}_n|\mathbf{y}_{0:n})$, where $\overline{\mathbf{x}}_n$ is a subvector of $\mathbf{x}_n^*$), and next $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ is obtained by marginalization.

A. Regular LG-TMC

Let us first address the case where $\mathcal{Q}_n^{\mathbf{y},\mathbf{y}}$ is positive definite. In this case a Kalman-like filtering algorithm has been proposed in [9] and [12]; it is recalled here for convenience of the reader.

Proposition 1 (KF for Regular LG-TMC): Let (22) and (23) hold. Let $p(\mathbf{x}_n^*|\mathbf{y}_{0:n})\sim\mathcal{N}(\widehat{\mathbf{x}}_{n|n}^*,\mathbf{P}_{n|n}^*)$ and $p(\mathbf{x}_{n+1}^*|\mathbf{y}_{0:n})\sim\mathcal{N}(\widehat{\mathbf{x}}_{n+1|n}^*,\mathbf{P}_{n+1|n}^*)$. Then, $\widehat{\mathbf{x}}_{n+1|n+1}^*$ and $\mathbf{P}_{n+1|n+1}^*$ can be computed from $\widehat{\mathbf{x}}_{n|n}^*$ and $\mathbf{P}_{n|n}^*$ via the following equations:

$$\begin{aligned}\widehat{\mathbf{x}}_{n+1|n}^*=&\left[\mathcal{F}_n^{\mathbf{x}^*,\mathbf{x}^*}-\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}\left(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}}\right)^{-1}\mathcal{F}_n^{\mathbf{y},\mathbf{x}^*}\right]\widehat{\mathbf{x}}_{n|n}^*\\ &+\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}\left(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}}\right)^{-1}\mathbf{y}_n\\ &+\left[\mathcal{F}_n^{\mathbf{x}^*,\mathbf{y}}-\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}\left(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}}\right)^{-1}\mathcal{F}_n^{\mathbf{y},\mathbf{y}}\right]\mathbf{y}_{n-1}\end{aligned} \quad (24)$$

$$\begin{aligned}\mathbf{P}_{n+1|n}^*=&\left[\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{x}^*}-\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}\left(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}}\right)^{-1}\mathcal{Q}_n^{\mathbf{y},\mathbf{x}^*}\right]\\ &+\left[\mathcal{F}_n^{\mathbf{x}^*,\mathbf{x}^*}-\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}\left(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}}\right)^{-1}\mathcal{F}_n^{\mathbf{y},\mathbf{x}^*}\right]\mathbf{P}_{n|n}^*\\ &\times\left[\mathcal{F}_n^{\mathbf{x}^*,\mathbf{x}^*}-\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}\left(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}}\right)^{-1}\mathcal{F}_n^{\mathbf{y},\mathbf{x}^*}\right]^T\end{aligned} \quad (25)$$

$$\mathbf{K}_{n+1|n+1}^*=\mathbf{P}_{n+1|n}^*\left(\mathcal{F}_{n+1}^{\mathbf{y},\mathbf{x}^*}\right)^T \quad (26)$$

$$\mathbf{L}^*_{n+1|n+1} = \mathcal{Q}^{\mathbf{y},\mathbf{y}}_{n+1} + \mathcal{F}^{\mathbf{y},\mathbf{x}^*}_{n+1}\mathbf{P}^*_{n+1|n}\left(\mathcal{F}^{\mathbf{y},\mathbf{x}^*}_{n+1}\right)^T \tag{27}$$

$$\hat{\mathbf{x}}^*_{n+1|n+1} = \hat{\mathbf{x}}^*_{n+1|n} + \mathbf{K}^*_{n+1|n+1}\left(\mathbf{L}^*_{n+1|n+1}\right)^{-1} \times\left[\mathbf{y}_{n+1} - \mathcal{F}^{\mathbf{y},\mathbf{x}^*}_{n+1}\hat{\mathbf{x}}^*_{n+1|n} - \mathcal{F}^{\mathbf{y},\mathbf{y}}_{n+1}\mathbf{y}_n\right] \tag{28}$$

$$\mathbf{P}^*_{n+1|n+1} = \mathbf{P}^*_{n+1|n} - \mathbf{K}^*_{n+1|n+1}\left(\mathbf{L}^*_{n+1|n+1}\right)^{-1}\left(\mathbf{K}^*_{n+1|n+1}\right)^T \tag{29}$$

and $\hat{\mathbf{x}}^*_{0|0}$ and $\mathbf{P}^*_{0|0}$ are given by (28) and (29), with $\hat{\mathbf{x}}^*_{0|-1} = \hat{\mathbf{x}}^*_0$ and $\mathbf{P}^*_{0|-1} = \mathbf{P}^*_0$.

Remark 1: Let us briefly comment on model (22), and associated algorithm (24)–(29). First, the introduction of PMC in the context of Kalman filtering is not entirely new. A closely related model was introduced independently [23] (see also [24, Corollary 1, p. 72]) in the context of conditionally Gaussian models. In this model, the pair $\mathbf{z}_n = (\mathbf{x}_n, \mathbf{y}_n)$ satisfies a linear equation similar to (20) and thus is an MC. Optimal filtering equations for this model have also been derived. On the other hand, as far as (22)–(29) are concerned, they can be derived from [22, eq. (4.24), p. 64] (by setting $\mathbf{u}_k = \mathbf{v}_{k-1}$). An alternate (prediction) recursive solution can also be obtained from [25, sec. 3.2.4, p. 113].

B. Singular LG-TMC

We now address the restoration problem in the singular measurements case, i.e. the case where $\mathcal{Q}^{\mathbf{y},\mathbf{y}}_n$ is positive semi-definite. Let now $r = \text{rank}(\mathcal{Q}^{\mathbf{y},\mathbf{y}}_n) \in \{0, 1, \ldots, n_{\mathbf{y}} - 1\}$. One can still use (24) to (29) by replacing (if necessary) inverses by generalized inverses. Following [3], [6], and [7], we shall however see that it is possible (under mild sufficient conditions) to exploit the singularity of $\mathcal{Q}^{\mathbf{y},\mathbf{y}}_n$ in order to reduce (via a state-space transform) by $m = n_{\mathbf{y}} - r$ the order of $\mathbf{x}^*_n$; we shall then propose a restoration algorithm for this equivalent, reduced-order system.

1) State-Space Transform: $\mathcal{Q}^{\mathbf{y},\mathbf{y}}_n$ has m zero eigenvalues. So there exists $\mathbf{M}_n$ invertible such that

$$\mathbf{M}_n\mathcal{Q}^{\mathbf{y},\mathbf{y}}_n\mathbf{M}^T_n = \begin{bmatrix}\mathbf{0}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_r\end{bmatrix}$$

in which $\mathbf{0}_m$ denotes the $m \times m$ null matrix. Let $\overline{\mathbf{y}}_n = \mathbf{M}_n\mathbf{y}_n$ and $\mathbf{w}^{\overline{\mathbf{y}}}_n = \mathbf{M}_n\mathbf{w}^{\mathbf{y}}_n$; then from (22) we get

$$\underbrace{\begin{bmatrix}\overline{\mathbf{y}}^p_n\\ \overline{\mathbf{y}}^r_n\end{bmatrix}}_{\overline{\mathbf{y}}_n} = \underbrace{\begin{bmatrix}\mathcal{F}^{\overline{\mathbf{y}}^p,\mathbf{x}^*}_n\\ \mathcal{F}^{\overline{\mathbf{y}}^r,\mathbf{x}^*}_n\end{bmatrix}}_{\mathcal{F}^{\overline{\mathbf{y}},\mathbf{x}^*}_n = \mathbf{M}_n\mathcal{F}^{\mathbf{y},\mathbf{x}^*}_n}\mathbf{x}^*_n + \underbrace{\begin{bmatrix}\mathcal{F}^{\overline{\mathbf{y}}^p,\overline{\mathbf{y}}}_n\\ \mathcal{F}^{\overline{\mathbf{y}}^r,\overline{\mathbf{y}}}_n\end{bmatrix}}_{\mathcal{F}^{\overline{\mathbf{y}},\overline{\mathbf{y}}}_n = \mathbf{M}_n\mathcal{F}^{\mathbf{y},\mathbf{y}}_n\mathbf{M}^{-1}_{n-1}}\overline{\mathbf{y}}_{n-1} + \underbrace{\begin{bmatrix}\mathbf{0}_{m\times 1}\\ \mathbf{w}^{\overline{\mathbf{y}}^r}_n\end{bmatrix}}_{\mathbf{w}^{\overline{\mathbf{y}}}_n} \tag{30}$$

and we see that $\overline{\mathbf{y}}_n$ is divided into a perfect part $(\overline{\mathbf{y}}^p_n)_{m\times 1}$ (the unnoisy part) and a regular one $(\overline{\mathbf{y}}^r_n)_{r\times 1}$. Since m linear functionals of $\mathbf{x}^*_n$ are known once $\overline{\mathbf{y}}_{n-1}$ and $\overline{\mathbf{y}}_n$ are known, there is no need to estimate them, and this is why one can reduce by m the order of the system, as we now see. Let $n_{\mathbf{x}^*} \geq m$, and let us first consider the following alternate partition of $\mathbf{x}^*_n$:

$$\mathbf{x}^*_n = \begin{bmatrix}(\mathbf{x}_n)_{n_{\mathbf{x}}\times 1}\\ (\mathbf{r}_n)_{n_{\mathbf{r}}\times 1}\end{bmatrix} = \begin{bmatrix}(\tilde{\mathbf{x}}_n)_{(n_{\mathbf{x}}+n_{\mathbf{r}}-m)\times 1}\\ (\tilde{\mathbf{r}}_n)_{m\times 1}\end{bmatrix} \tag{31}$$

(for the moment we say nothing about the position of $n_{\mathbf{r}}$ with regard to m). Let us assume that in (30)

$$\text{rank}\left(\mathcal{F}^{\overline{\mathbf{y}}^p,\mathbf{x}^*}_n\right)_{m\times n_{\mathbf{x}^*}} = m. \tag{32}$$

Then, one can chose a $(n_{\mathbf{x}^*} - m) \times n_{\mathbf{x}^*}$ matrix $\mathbf{U}_n$ in such a way that the transform

$$\underbrace{\begin{bmatrix}(\mathbf{U}_n)_{(n_{\mathbf{x}^*}-m)\times n_{\mathbf{x}^*}}\\ \left(\mathcal{F}^{\overline{\mathbf{y}}^p,\mathbf{x}^*}_n\right)_{m\times n_{\mathbf{x}^*}}\end{bmatrix}}_{\mathbf{T}_n}\underbrace{\begin{bmatrix}\tilde{\mathbf{x}}_n\\ \tilde{\mathbf{r}}_n\end{bmatrix}}_{\mathbf{x}^*_n} = \begin{bmatrix}\overline{\mathbf{x}}_n\\ \overline{\mathbf{y}}^p_n - \mathcal{F}^{\overline{\mathbf{y}}^p,\overline{\mathbf{y}}}_n\overline{\mathbf{y}}_{n-1}\end{bmatrix} \tag{33}$$

is reversible, and finally $\mathbf{T}_n$ and $\mathbf{M}_n$ enable to transform (22) into the equivalent system

$$\underbrace{\begin{bmatrix}\mathbf{T}_{n+1} & \mathbf{0}\\ \mathbf{0} & \mathbf{M}_n\end{bmatrix}\begin{bmatrix}\mathbf{x}^*_{n+1}\\ \mathbf{y}_n\end{bmatrix}}_{\mathbf{t}'_{n+1}} = \begin{bmatrix}\mathbf{T}_{n+1} & \mathbf{0}\\ \mathbf{0} & \mathbf{M}_n\end{bmatrix}\mathcal{F}_n\begin{bmatrix}\mathbf{T}^{-1}_n & \mathbf{0}\\ \mathbf{0} & \mathbf{M}^{-1}_{n-1}\end{bmatrix} \times\underbrace{\begin{bmatrix}\mathbf{T}_n & \mathbf{0}\\ \mathbf{0} & \mathbf{M}_{n-1}\end{bmatrix}\begin{bmatrix}\mathbf{x}^*_n\\ \mathbf{y}_{n-1}\end{bmatrix}}_{\mathbf{t}'_n} + \begin{bmatrix}\mathbf{T}_{n+1} & \mathbf{0}\\ \mathbf{0} & \mathbf{M}_n\end{bmatrix}\begin{bmatrix}\mathbf{w}^{\mathbf{x}^*}_n\\ \mathbf{w}^{\mathbf{y}}_n\end{bmatrix}. \tag{34}$$

The first $n_{\mathbf{x}^*}$ equations of (34) can be rewritten as

$$\begin{bmatrix}\overline{\mathbf{x}}_{n+1}\\ \overline{\mathbf{y}}^p_{n+1}\end{bmatrix} = \begin{bmatrix}\overline{\mathcal{F}}^{\overline{\mathbf{x}},\overline{\mathbf{x}}}_n & \overline{\mathcal{F}}^{\overline{\mathbf{x}},\overline{\mathbf{y}}}_n\\ \overline{\mathcal{F}}^{\overline{\mathbf{y}}^p,\overline{\mathbf{x}}}_n & \overline{\mathcal{F}}^{\overline{\mathbf{y}}^p,\overline{\mathbf{y}}}_n\end{bmatrix}\begin{bmatrix}\overline{\mathbf{x}}_n\\ \overline{\mathbf{y}}_n\end{bmatrix} + \begin{bmatrix}\overline{\mathcal{G}}^{\overline{\mathbf{x}}}_n\\ \overline{\mathcal{G}}^{\overline{\mathbf{y}}^p}_n\end{bmatrix}\overline{\mathbf{y}}_{n-1} + \mathbf{T}_{n+1}\mathbf{w}^{\mathbf{x}^*}_n \tag{35}$$

$$\begin{bmatrix}\overline{\mathcal{F}}^{\overline{\mathbf{x}},\overline{\mathbf{x}}}_n & \overline{\mathcal{F}}^{\overline{\mathbf{x}},\overline{\mathbf{y}}}_n\\ \overline{\mathcal{F}}^{\overline{\mathbf{y}}^p,\overline{\mathbf{x}}}_n & \overline{\mathcal{F}}^{\overline{\mathbf{y}}^p,\overline{\mathbf{y}}}_n\end{bmatrix} = \underbrace{\left[\mathbf{T}_{n+1}\mathcal{F}^{\mathbf{x}^*,\mathbf{x}^*}_n\mathbf{T}^{-1}_n, \mathbf{0}\right]}_{n_{\mathbf{x}^*}\times(n_{\mathbf{x}^*}+r)} + \begin{bmatrix}\mathbf{0}_{n_{\overline{\mathbf{x}}}} & \mathbf{0}\\ \mathbf{0} & \mathcal{F}^{\overline{\mathbf{y}}^p,\overline{\mathbf{y}}}_{n+1}\end{bmatrix} \tag{36}$$

$$\begin{bmatrix}\overline{\mathcal{G}}^{\overline{\mathbf{x}}}_n\\ \overline{\mathcal{G}}^{\overline{\mathbf{y}}^p}_n\end{bmatrix} = \mathbf{T}_{n+1}\mathcal{F}^{\mathbf{x}^*,\mathbf{y}}_n\mathbf{M}^{-1}_{n-1} - \mathbf{T}_{n+1}\mathcal{F}^{\mathbf{x}^*,\mathbf{x}^*}_n\mathbf{T}^{-1}_n\begin{bmatrix}\mathbf{0}_{n_{\overline{\mathbf{x}}}\times n_{\mathbf{y}}}\\ \mathcal{F}^{\overline{\mathbf{y}}^p,\overline{\mathbf{y}}}_n\end{bmatrix}. \tag{37}$$

On the other hand, the last $n_{\mathbf{y}}$ equations of (34) are given by (30). The first m equations of (30) coincide with the last m equations of (35) and are thus redundant. Gathering (35) with the last r equations of (30) (written at time $n+1$ thanks to (22)), we get the reduced-order system

$$\underbrace{\begin{bmatrix}\overline{\mathbf{x}}_{n+1}\\ \overline{\mathbf{y}}_{n+1}\end{bmatrix}}_{\overline{\mathbf{t}}_{n+1}} = \underbrace{\begin{bmatrix}\overline{\mathcal{F}}^{\overline{\mathbf{x}},\overline{\mathbf{x}}}_n & \overline{\mathcal{F}}^{\overline{\mathbf{x}},\overline{\mathbf{y}}}_n\\ \overline{\mathcal{F}}^{\overline{\mathbf{y}},\overline{\mathbf{x}}}_n & \overline{\mathcal{F}}^{\overline{\mathbf{y}},\overline{\mathbf{y}}}_n\end{bmatrix}}_{\overline{\mathcal{F}}_n}\begin{bmatrix}\overline{\mathbf{x}}_n\\ \overline{\mathbf{y}}_n\end{bmatrix} + \underbrace{\begin{bmatrix}\overline{\mathcal{G}}^{\overline{\mathbf{x}}}_n\\ \overline{\mathcal{G}}^{\overline{\mathbf{y}}}_n\end{bmatrix}}_{\overline{\mathcal{G}}_n}\overline{\mathbf{y}}_{n-1} + \underbrace{\begin{bmatrix}\mathbf{U}_{n+1} & \mathbf{0}\\ \mathcal{F}^{\overline{\mathbf{y}},\mathbf{x}^*}_{n+1} & \tilde{\mathbf{I}}_r\end{bmatrix}\begin{bmatrix}\mathbf{w}^{\mathbf{x}^*}_n\\ \mathbf{w}^{\overline{\mathbf{y}}^r}_{n+1}\end{bmatrix}}_{\overline{\mathbf{w}}_n} \tag{38}$$

with $\tilde{\mathbf{I}}_r = [\mathbf{0}^T_{m\times r}, \mathbf{I}_r]^T$, and

$$\overline{\mathcal{F}}_n = \underbrace{\begin{bmatrix}\mathbf{T}_{n+1}\\ \mathcal{F}^{\overline{\mathbf{y}}^r,\mathbf{x}^*}_{n+1}\end{bmatrix}}_{(n_{\mathbf{x}^*}+r)\times n_{\mathbf{x}^*}} \underbrace{\left[\mathcal{F}^{\mathbf{x}^*,\mathbf{x}^*}_n \mathbf{T}_n^{-1}, \mathbf{0}\right]}_{n_{\mathbf{x}^*}\times(n_{\mathbf{x}^*}+r)} + \begin{bmatrix}\mathbf{0}_{n_{\overline{\mathbf{x}}}} & \mathbf{0}\\ \mathbf{0} & \mathcal{F}^{\overline{\mathbf{y}},\overline{\mathbf{y}}}_{n+1}\end{bmatrix} \quad (39)$$

$$\overline{\mathcal{G}}_n = \underbrace{\begin{bmatrix}\mathbf{T}_{n+1}\\ \mathcal{F}^{\overline{\mathbf{y}}^r,\mathbf{x}^*}_{n+1}\end{bmatrix}}_{(n_{\mathbf{x}^*}+r)\times n_{\mathbf{x}^*}} \left(\mathcal{F}^{\mathbf{x}^*,\mathbf{y}}_n \mathbf{M}^{-1}_{n-1} - \mathcal{F}^{\mathbf{x}^*,\mathbf{x}^*}_n \mathbf{T}^{-1}_n \begin{bmatrix}\mathbf{0}_{n_{\overline{\mathbf{x}}}\times n_{\mathbf{y}}}\\ \mathcal{F}^{\overline{\mathbf{y}}^p,\overline{\mathbf{y}}}_n\end{bmatrix}\right). \quad (40)$$

2) Restoration Algorithm: Let us next address the restoration of $\overline{\mathbf{x}}_n$ from $\{\overline{\mathbf{y}}_{0:n}\}$ in (38), and, finally, the restoration of $\mathbf{x}_n$ from $\{\mathbf{y}_{0:n}\}$ in (22), which is our ultimate goal. Let $\mathbf{w}^{\mathbf{x}^*}_n$ and $\mathbf{w}^{\mathbf{y}}_n$ be independent, and let

$$\begin{aligned}\overline{\mathcal{Q}}_n &= E\left(\overline{\mathbf{w}}_n\overline{\mathbf{w}}_n^T\right)\\ &= \begin{bmatrix}\mathbf{U}_{n+1} & \mathbf{0}\\ \mathcal{F}^{\overline{\mathbf{y}},\mathbf{x}^*}_{n+1} & \tilde{\mathbf{I}}_r\end{bmatrix}\begin{bmatrix}\mathcal{Q}^{\mathbf{x}^*,\mathbf{x}^*}_n & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_r\end{bmatrix}\begin{bmatrix}\mathbf{U}_{n+1} & \mathbf{0}\\ \mathcal{F}^{\overline{\mathbf{y}},\mathbf{x}^*}_{n+1} & \tilde{\mathbf{I}}_r\end{bmatrix}^T\\ &= \begin{bmatrix}\overline{\mathcal{Q}}^{\overline{\mathbf{x}},\overline{\mathbf{x}}}_n & \overline{\mathcal{Q}}^{\overline{\mathbf{x}},\overline{\mathbf{y}}}_n\\ \overline{\mathcal{Q}}^{\overline{\mathbf{y}},\overline{\mathbf{x}}}_n & \overline{\mathcal{Q}}^{\overline{\mathbf{y}},\overline{\mathbf{y}}}_n\end{bmatrix}. \quad (41)\end{aligned}$$

Then $p(\overline{\mathbf{x}}_n|\overline{\mathbf{y}}_{0:n})$ can be computed recursively via the following algorithm.

Proposition 2 (KF for Singular LG-TMC): Let (22), (23), and (32) hold, let $n_{\mathbf{x}^*} \geq m$, and let moreover $\mathbf{w}^{\mathbf{x}^*}_n$ and $\mathbf{w}^{\mathbf{y}}_n$ be independent. Let $p(\overline{\mathbf{x}}_n|\overline{\mathbf{y}}_{0:n}) \sim \mathcal{N}(\widehat{\overline{\mathbf{x}}}_{n|n}, \overline{\mathbf{P}}_{n|n})$ and $p(\overline{\mathbf{x}}_{n+1}|\overline{\mathbf{y}}_{0:n}) \sim \mathcal{N}\widehat{\overline{\mathbf{x}}}_{n+1|n}, \overline{\mathbf{P}}_{n+1|n})$. Then $\widehat{\overline{\mathbf{x}}}_{n+1|n+1}$ and $\overline{\mathbf{P}}_{n+1|n+1}$ can be computed from $\widehat{\overline{\mathbf{x}}}_{n|n}$ and $\overline{\mathbf{P}}_{n|n}$ via[2]

$$\widehat{\overline{\mathbf{x}}}_{n+1|n} = \overline{\mathcal{F}}^{\overline{\mathbf{x}},\overline{\mathbf{x}}}_n \widehat{\overline{\mathbf{x}}}_{n|n} + \overline{\mathcal{F}}^{\overline{\mathbf{x}},\overline{\mathbf{y}}}_n \overline{\mathbf{y}}_n + \overline{\mathcal{G}}^{\overline{\mathbf{x}}}_n \overline{\mathbf{y}}_{n-1} \quad (42)$$

$$\overline{\mathbf{P}}_{n+1|n} = \overline{\mathcal{Q}}^{\overline{\mathbf{x}},\overline{\mathbf{x}}}_n + \overline{\mathcal{F}}^{\overline{\mathbf{x}},\overline{\mathbf{x}}}_n \overline{\mathbf{P}}_{n|n}\left(\overline{\mathcal{F}}^{\overline{\mathbf{x}},\overline{\mathbf{x}}}_n\right)^T \quad (43)$$

$$\widehat{\overline{\mathbf{y}}}_{n+1|n} = \overline{\mathcal{F}}^{\overline{\mathbf{y}},\overline{\mathbf{x}}}_n \widehat{\overline{\mathbf{x}}}_{n|n} + \overline{\mathcal{F}}^{\overline{\mathbf{y}},\overline{\mathbf{y}}}_n \overline{\mathbf{y}}_n + \overline{\mathcal{G}}^{\overline{\mathbf{y}}}_n \overline{\mathbf{y}}_{n-1} \quad (44)$$

$$\overline{\mathbf{K}}_{n+1|n+1} = \overline{\mathcal{Q}}^{\overline{\mathbf{x}},\overline{\mathbf{y}}}_n + \overline{\mathcal{F}}^{\overline{\mathbf{x}},\overline{\mathbf{x}}}_n \overline{\mathbf{P}}_{n|n}\left(\overline{\mathcal{F}}^{\overline{\mathbf{y}},\overline{\mathbf{x}}}_n\right)^T \quad (45)$$

$$\overline{\mathbf{L}}_{n+1|n+1} = \overline{\mathcal{Q}}^{\overline{\mathbf{y}},\overline{\mathbf{y}}}_n + \overline{\mathcal{F}}^{\overline{\mathbf{y}},\overline{\mathbf{x}}}_n \overline{\mathbf{P}}_{n|n}\left(\overline{\mathcal{F}}^{\overline{\mathbf{y}},\overline{\mathbf{x}}}_n\right)^T \quad (46)$$

$$\widehat{\overline{\mathbf{x}}}_{n+1|n+1} = \widehat{\overline{\mathbf{x}}}_{n+1|n} + \overline{\mathbf{K}}_{n+1|n+1}\overline{\mathbf{L}}^{-1}_{n+1|n+1}(\overline{\mathbf{y}}_{n+1} - \widehat{\overline{\mathbf{y}}}_{n+1|n}) \quad (47)$$

$$\overline{\mathbf{P}}_{n+1|n+1} = \overline{\mathbf{P}}_{n+1|n} - \overline{\mathbf{K}}_{n+1|n+1}\overline{\mathbf{L}}^{-1}_{n+1|n+1}\overline{\mathbf{K}}^T_{n+1|n+1}. \quad (48)$$

As for the initialization, $\widehat{\overline{\mathbf{x}}}_{0|0}$ and $\overline{\mathbf{P}}_{0|0}$ are given by (47) and (48) with $n = -1$, in which $\widehat{\overline{\mathbf{x}}}_{0|-1}$, $\overline{\mathbf{P}}_{0|-1}$, $\widehat{\overline{\mathbf{y}}}_{0|-1}$, $\overline{\mathbf{K}}_{0|0}$ and $\overline{\mathbf{L}}_{0|0}$ are given respectively by $\widehat{\overline{\mathbf{x}}}_{0|-1} = \mathbf{U}_0\widehat{\mathbf{x}}^*_0$, $\overline{\mathbf{P}}_{0|-1} = \mathbf{U}_0\mathbf{P}^*_0\mathbf{U}^T_0$, $\widehat{\overline{\mathbf{y}}}_{0|-1} = \mathcal{F}^{\overline{\mathbf{y}},\mathbf{x}^*}_0\widehat{\mathbf{x}}^*_0$, $\overline{\mathbf{K}}_{0|0} = \mathbf{U}_0\mathbf{P}^*_0(\mathcal{F}^{\overline{\mathbf{y}},\mathbf{x}^*}_0)^T$ and $\overline{\mathbf{L}}_{0|0} = \mathcal{F}^{\overline{\mathbf{y}},\mathbf{x}^*}_0\mathbf{P}^*_0(\mathcal{F}^{\overline{\mathbf{y}},\mathbf{x}^*}_0)^T + \mathrm{diag.}(\mathbf{0}_m, \mathbf{I}_r)$. Finally, conditionally on $\mathbf{y}_{0:n}$, $\mathbf{x}_n \sim \mathcal{N}(\widehat{\mathbf{x}}_{n|n}, \mathbf{P}_{n|n})$, where $\widehat{\mathbf{x}}_{n|n}$ and $\mathbf{P}_{n|n}$ can be computed as

$$\widehat{\mathbf{x}}_{n|n} = [\mathbf{I}_{n_{\mathbf{x}}}, \mathbf{0}_{n_{\mathbf{x}}\times n_{\mathbf{r}}}]\,\mathbf{T}^{-1}_n \begin{bmatrix}\widehat{\overline{\mathbf{x}}}_{n|n}\\ \overline{\mathbf{y}}^p_n - \mathcal{F}^{\overline{\mathbf{y}}^p,\overline{\mathbf{y}}}_n\overline{\mathbf{y}}_{n-1}\end{bmatrix} \quad (49)$$

$$\mathbf{P}_{n|n} = [\mathbf{I}_{n_{\mathbf{x}}}, \mathbf{0}_{n_{\mathbf{x}}\times n_{\mathbf{r}}}]\,\mathbf{T}^{-1}_n \begin{bmatrix}\overline{\mathbf{P}}_{n|n} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}_m\end{bmatrix} \times \left(\mathbf{T}^{-1}_n\right)^T [\mathbf{I}_{n_{\mathbf{x}}}, \mathbf{0}_{n_{\mathbf{x}}\times n_{\mathbf{r}}}]^T. \quad (50)$$

[2]Inverses in (47) and (48) should be replaced by a generalized inverse in case $\overline{\mathbf{L}}_{n+1|n+1}$ is not invertible.

Proof: Model (38) is a particular case of [24, eqs. (13.46)–(13.47)] or [25, model (3.1.1), (3.1.3), and (3.2.20)], so (42) to (48) can be obtained from [24, eqs. (13.56)–(13.57)] or [25, sec. 3.2.4, p. 112]. Let us next consider the initialization. From (30) and (33), we get

$$\begin{bmatrix}\overline{\mathbf{x}}_0\\ \overline{\mathbf{y}}_0\end{bmatrix} = \begin{bmatrix}\mathbf{U}_0\\ \mathcal{F}^{\overline{\mathbf{y}},\mathbf{x}^*}_0\end{bmatrix}\mathbf{x}^*_0 + \begin{bmatrix}\mathbf{0}\\ \mathbf{w}^{\overline{\mathbf{y}}}_0\end{bmatrix}$$

so

$$\begin{bmatrix}\overline{\mathbf{x}}_0\\ \overline{\mathbf{y}}_0\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\widehat{\overline{\mathbf{x}}}_{0|-1}\\ \widehat{\overline{\mathbf{y}}}_{0|-1}\end{bmatrix}, \begin{bmatrix}\overline{\mathbf{P}}_{0|-1} & \overline{\mathbf{K}}_{0|0}\\ \overline{\mathbf{K}}^T_{0|0} & \overline{\mathbf{L}}_{0|0}\end{bmatrix}\right)$$

with $\widehat{\overline{\mathbf{x}}}_{0|-1}$, $\overline{\mathbf{P}}_{0|-1}$, $\widehat{\overline{\mathbf{y}}}_{0|-1}$, $\overline{\mathbf{K}}_{0|0}$ and $\overline{\mathbf{L}}_{0|0}$ computed thanks to (23). On the other hand, $p(\overline{\mathbf{x}}_0|\overline{\mathbf{y}}_0) \sim \mathcal{N}(\widehat{\overline{\mathbf{x}}}_{0|0}, \overline{\mathbf{P}}_{0|0})$ is given by (47) and (48) with $n = -1$. Finally, let us partition $\mathbf{T}^{-1}_n$ as $\mathbf{T}^{-1}_n = [(\mathbf{V}_n)_{n_{\mathbf{x}^*}\times(n_{\mathbf{x}^*}-m)}, (\mathbf{W}_n)_{n_{\mathbf{x}^*}\times m}]$. From (33), $\mathbf{x}^*_n = \mathbf{V}_n\overline{\mathbf{x}}_n + \mathbf{W}_n(\overline{\mathbf{y}}^p_n - \mathcal{F}^{\overline{\mathbf{y}}^p,\overline{\mathbf{y}}}_n\overline{\mathbf{y}}_{n-1})$. So conditionally on $\mathbf{y}_{0:n}$, $\mathbf{x}^*_n \sim \mathcal{N}(\mathbf{V}_n\widehat{\overline{\mathbf{x}}}_{n|n} + \mathbf{W}_n(\overline{\mathbf{y}}^p_n - \mathcal{F}^{\overline{\mathbf{y}}^p,\overline{\mathbf{y}}}_n\overline{\mathbf{y}}_{n-1}), \mathbf{V}_n\overline{\mathbf{P}}_{n|n}\mathbf{V}^T_n)$. Marginalizing with regard to the $n_{\mathbf{x}}$ first components (remember that $\mathbf{x}^*_n = (\mathbf{x}_n, \mathbf{r}_n)$), we eventually get (49) and (50). ■

3) Comments and Remarks:

1) Simplifications occur in some cases. For instance, let us partition $\mathcal{F}^{\overline{\mathbf{y}}^p,\mathbf{x}^*}_n$ defined in (30) as $\mathcal{F}^{\overline{\mathbf{y}}^p,\mathbf{x}^*}_n = [(\mathcal{F}^{\overline{\mathbf{y}}^p,\tilde{\mathbf{x}}}_n)_{m\times n_{\overline{\mathbf{x}}}}, (\mathcal{F}^{\overline{\mathbf{y}}^p,\tilde{\mathbf{r}}}_n)_{m\times m}]$, and let $\mathcal{F}^{\overline{\mathbf{y}}^p,\tilde{\mathbf{r}}}_n$ be invertible. In this case, $\mathbf{U}_n$ can be chosen as $\mathbf{U}_n = [\mathbf{I}_{n_{\tilde{\mathbf{x}}}}, \mathbf{0}]$, and

$$\mathbf{T}^{-1}_n = \begin{bmatrix}\mathbf{I}_{n_{\tilde{\mathbf{x}}}} & \mathbf{0}_{n_{\tilde{\mathbf{x}}}\times m}\\ \mathcal{F}^{\overline{\mathbf{y}}^p,\tilde{\mathbf{x}}}_n & \mathcal{F}^{\overline{\mathbf{y}}^p,\tilde{\mathbf{r}}}_n\end{bmatrix}^{-1} = \begin{bmatrix}\mathbf{I}_{n_{\tilde{\mathbf{x}}}} & \mathbf{0}_{n_{\tilde{\mathbf{x}}}\times m}\\ -\left(\mathcal{F}^{\overline{\mathbf{y}}^p,\tilde{\mathbf{r}}}_n\right)^{-1}\mathcal{F}^{\overline{\mathbf{y}}^p,\tilde{\mathbf{x}}}_n & \left(\mathcal{F}^{\overline{\mathbf{y}}^p,\tilde{\mathbf{r}}}_n\right)^{-1}\end{bmatrix}. \quad (51)$$

Next two cases occur. If $n_{\mathbf{r}} \geq m$, then in (31) $n_{\mathbf{x}} \leq n_{\tilde{\mathbf{x}}}$, and from (49) and (50) $\widehat{\mathbf{x}}_{n|n}$ is simply a subvector of $\widehat{\overline{\mathbf{x}}}_{n|n}$ and $\mathbf{P}_{n|n}$ a submatrix of $\overline{\mathbf{P}}_{n|n}$ (an example where this happens is given in item 3 below). If $m > n_{\mathbf{r}}$, we see from (50) that $\mathbf{P}_{n|n}$ is a $n_{\mathbf{x}} \times n_{\mathbf{x}}$ singular matrix of rank at most $n_{\overline{\mathbf{x}}}$. The reason why is that in (31), $\tilde{\mathbf{x}}_n$ is a subvector of $\mathbf{x}_n$, so, up to an invertible matrix, only $n_{\tilde{\mathbf{x}}}$ components of $\mathbf{x}_n$ need to be estimated from $\mathbf{y}_{0:n}$.

2) In most models of Section II-B $\mathbf{w}^{\mathbf{y}}_n = \mathbf{0}$, so the perfect measurement case is an important particular case of our algorithm. Let $m = n_{\mathbf{y}}$. Then, $\mathcal{Q}^{\mathbf{y},\mathbf{y}}_n = \mathbf{0}_{n_{\mathbf{y}}}$, and one can chose $\mathbf{M}_n = \mathbf{I}_{n_{\mathbf{y}}}$. The last $n_{\mathbf{y}}$ equations of (34) are given by (30), and they coincide with the last $n_{\mathbf{y}}$ equations of (35). So (35) is sufficient, and (38) reduces to (35) (with $\overline{\mathbf{y}}^p_n = \mathbf{y}_n$), (39) to (36) and (40) to (37) (an example where this happens is given in item 3 below).

3) As we have seen in Section II-B, the linear TMC model encompasses some classical models. It happens that the algorithm of Section III-B also generalizes some classical algorithms. Let us e.g. consider the model of Section II-B-2) (which is widely used, in particular in speech enhancement and coding (see, e.g., [21])). If (12) reduces to (14), then $\overline{\mathcal{G}}^{\overline{\mathbf{x}}}_n$ and $\overline{\mathcal{G}}^{\overline{\mathbf{y}}}_n$ vanish, so in (42) to (48) the dependency on $\overline{\mathbf{y}}_{n-1}$ vanishes, and (43), (45), (46) and (47) reduce respectively to [21, eq. 51, p. 1736], [21, eq. 57, p. 1737] and [21, eq. 56, p. 1736]; while (42) (respectively, (48))

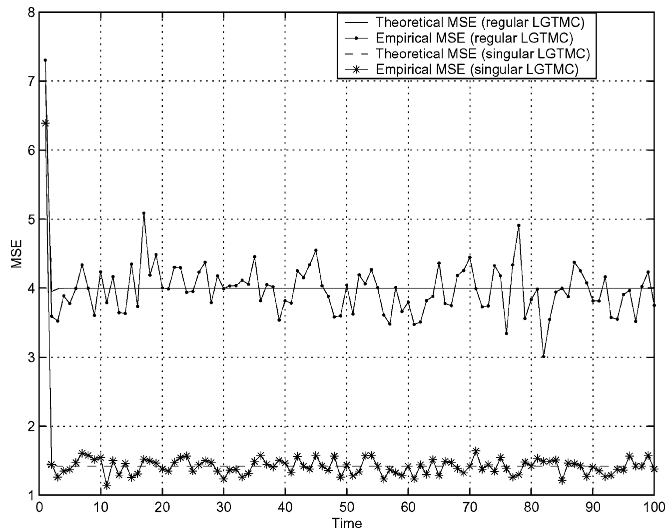

Fig. 1. MSE, regular and singular cases.

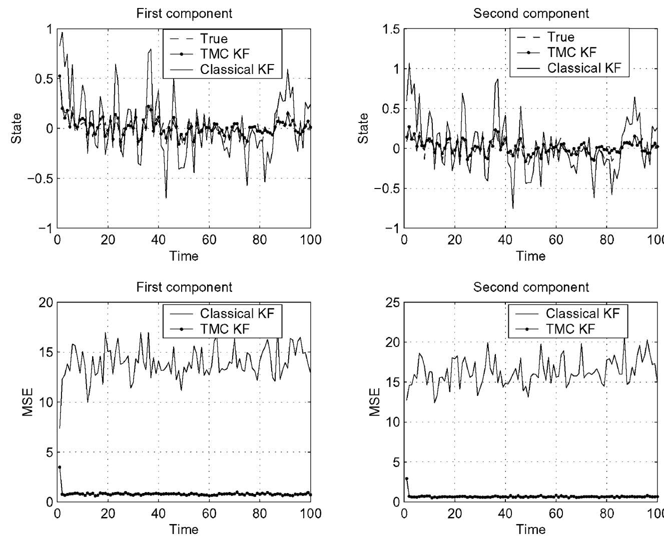

Fig. 2. State and MSE, singular TMC KF, and classical KF.

reduces to an equation that can be obtained as part of [21, eq. 54, p. 1736] (respectively, [21, eq. 52, p. 1736]).

C. Numerical Example

Let us finally perform some simulations. We consider a regular and a singular model, which differ by the covariance matrix $\mathcal{Q}_n^{\mathbf{y},\mathbf{y}}$: $\mathcal{Q}_n^{\mathbf{y},\mathbf{y}} = \mathcal{Q}_{n,r}^{\mathbf{y},\mathbf{y}}$ (respectively, $\mathcal{Q}_n^{\mathbf{y},\mathbf{y}} = \mathcal{Q}_{n,s}^{\mathbf{y},\mathbf{y}}$) in the regular (respectively, singular) model. So let $n_{\mathbf{x}} = n_{\mathbf{r}} = n_{\mathbf{y}} = 2$, $\mathbf{x}_0^* \sim \mathcal{N}([.5, .1, .2, 0]^T, 3.5\mathbf{I}_4)$, $\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}} = \mathbf{0}$, and let

$$\mathcal{F}_n = \begin{bmatrix} .02 & .10 & .01 & .12 & .03 & .11 \\ .10 & .05 & .12 & .04 & .11 & .02 \\ .10 & .10 & .12 & .04 & .02 & .10 \\ .09 & .11 & .04 & .12 & .09 & .10 \\ .04 & .06 & .11 & .12 & .01 & .12 \\ .02 & .09 & .13 & .03 & .05 & .11 \end{bmatrix}$$

$$\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{x}^*} \begin{bmatrix} 2.0 & 1.4 & 1.2 & 1.3 \\ 1.4 & 2.0 & 1.4 & 1.2 \\ 1.2 & 1.4 & 2.0 & 1.4 \\ 1.3 & 1.2 & 1.4 & 2.0 \end{bmatrix} \tag{52}$$

$$\mathcal{Q}_{n,r}^{\mathbf{y},\mathbf{y}} = \begin{bmatrix} 2.0 & 1.4 \\ 1.4 & 2.0 \end{bmatrix} \text{ or } \mathcal{Q}_{n,s}^{\mathbf{y},\mathbf{y}} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}. \tag{53}$$

All simulations results are averaged over 200 independent realizations.

Fig. 1 illustrates the interest of the singular algorithm over the regular one by displaying the empirical and theoretical mean square errors (MSE) of $(\mathbf{x}_n - \widehat{\mathbf{x}}_{n|n})$ for both cases. As expected, the results are consistent with the dimension of the state ($\mathbf{x}_n^*$ or $\overline{\mathbf{x}}_n$) which needs to be estimated: in the regular case, $n_{\mathbf{x}^*} = 4$, while in the singular case $n_{\overline{\mathbf{x}}} = 3$.

Fig. 2 compares the performances of the TMC KF and of a standard KF. The data are generated by the singular TMC model, and $\mathbf{x}$ is restored either by the singular TMC KF or by a classical (singular case) KF. Fig. 2 displays the true value of the state, its two estimates, and the associated empirical MSE; as expected, the TMC KF outperforms the KF.

REFERENCES

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng., Trans. ASME, Series D*, vol. 82, no. 1, pp. 35–45, 1960.

[2] Y. C. Ho and R. C. K. Lee, "A Bayesian approach to problems in stochastic estimation and control," *IEEE Trans. Autom. Control*, vol. 9, no. 4, pp. 333–339, Oct. 1964.

[3] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[4] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, ser. Prentice-Hall Information and System Sciences Series. Upper Saddle River, NJ: Prentice-Hall, 2000.

[5] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. San Diego, CA: Academic, 1970, vol. 64, Mathematics in Science and Engineering.

[6] A. P. Sage and J. L. Melsa, *Estimation Theory With Applications to Communications and Control*. New York: McGraw-Hill, 1971.

[7] P. S. Maybeck, *Stochastic Models, Estimation and Control*. New York: Academic, 1979, vol. 1.

[8] A. Doucet, N. de Freitas, and N. Gordon, Eds., "Sequential Monte Carlo methods in practice," in *Statistics for Engineering and Information Science*. New York: Springer Verlag, 2001.

[9] W. Pieczynski and F. Desbouvries, "Kalman filtering using pairwise Gaussian models," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Hong Kong, Apr. 6–10, 2003, vol. 6, pp. VI-57–VI-60.

[10] F. Desbouvries and W. Pieczynski, "Particle filtering in pairwise and triplet Markov chains," in *Proc. IEEE—EURASIP Workshop Nonlinear Signal Image Processing (NSIP)*, Grado-Gorizia, Italy, Jun. 8–11, 2003.

[11] W. Pieczynski, C. Hulard, and T. Veit, "Triplet Markov chains in hidden signal restoration," in *Proc. SPIE Int. Symp. Remote Sensing*, Crete, Greece, Sep. 22–27, 2002.

[12] F. Desbouvries and W. Pieczynski, "Modèles de Markov triplet et filtrage de Kalman," (in French) *Comptes Rendus de l'Académie des Sciences—Mathématiques*, vol. 336, no. 8, 2003.

[13] W. Pieczynski and F. Desbouvries, "On triplet Markov chains," in *Proc. Int. Symp. Applied Stochastic Models Data Analysis (ASMDA)*, Brest, France, May 17–20, 2005.

[14] C. J. Wellekens, "Explicit time correlation in hidden Markov models for speech recognition," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, 1987, vol. 12, pp. 384–386.

[15] M. Ostendorf, V. V. Digalakis, and O. A. Kimball, "From HMMs to segment models: a unified view of stochastic modeling for speech recognition," *IEEE Trans. Speech Audio Process.*, vol. 4, no. 5, pp. 360–378, Sep. 1996.

[16] S. Derrode and W. Pieczynski, "Signal and image segmentation using pairwise Markov chains," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2477–2489, Sep. 2004.

[17] P. Lanchantin and W. Pieczynski, "Unsupervised non stationary image segmentation using triplet Markov chains," in *Proc. Advanced Concepts for Intelligent Vision Systems (ACVIS)*, Brussels, Belgium, Aug. 31–Sep. 3 2004.

[18] A. E. Bryson, Jr. and D. E. Johansen, "Linear filtering for time-varying systems using measurements containing colored noise," *IEEE Trans. Autom. Control*, vol. AC-10, no. 1, pp. 4–10, Jan. 1965.

[19] H. W. Sorenson, "Kalman filtering techniques," in *Advances in Control Systems Theory and Appl.*, C. T. Leondes, Ed. New York: Academic , 1966, vol. 3, pp. 219–292.

[20] A. E. Bryson, Jr. and L. J. Henrikson, "Estimation using sampled data containing sequentially correlated noise," *J. Spacecr. Rockets*, vol. 5, pp. 662–665, Jun. 1968.

[21] J. D. Gibson, B. Koo, and S. D. Gray, "Filtering of colored noise for speech enhancement and coding," *IEEE Trans. Signal Process.*, vol. 39, no. 8, pp. 1732–1742, Aug. 1991.

[22] C. K. Chui and G. Chen, *Kalman Filtering With Real-Time Applications*. Berlin, Germany: Springer, 1999.

[23] R. S. Lipster and A. N. Shiryaev, "Statistics of conditionally Gaussian random sequences," in *6th Berkeley Symp. Mathematics, Statistics, Probability*, 1972, vol. 2, pp. 389–422.

[24] ——, "Conditionally Gaussian sequences: filtering and related problems," in *Statistics of Random Processes, Vol. 2: Applications*. Berlin, Germany: Springer-Verlag, 2001, ch. 13.

[25] A. C. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge, U.K.: Cambridge Univ. Press, 1989.

Boujemaa Ait-El-Fquih was born in Ouled Teima, Morocco, on April 9, 1977. He received the M.Sc. degree in physics and the Diplome d'Études Approfondies degree in signal and image processing from Cergy-Pontoise University, France, in 2001 and 2002, respectively. He is currently working towards the Ph.D. degree in signal processing at the Institut National des Télécommunications, Évry, France.

François Desbouvries (M'00) received the M.Sc. degree in telecommunications and the Ph.D. degree in signal and image processing from the École Nationale Supérieure des Télécommunications (ENST), Paris, France, in 1987 and 1991, respectively, and the Habilitation à Diriger des Recherches (HDR) degree in signal processing from Marne-la-Vallée University, France, in 2001.

Since 1991, he has been with the Institut National des Télécommunications, where he is currently a Professor. His research interests include Bayesian estimation, sequential Monte Carlo methods, and Markovian models.