BAYESIAN FIXED-INTERVAL SMOOTHING ALGORITHMS IN SINGULAR STATE-SPACE SYSTEMS

B. Ait-El-Fquih

LSS & CNRS UMR 8506
Supélec, 3 rue Joliot-Curie
91192 Gif-sur-Yvette, France

F. Desbouvries

Telecom SudParis/CITI & CNRS UMR 5157
9 rue Charles Fourier
91011 Evry, France

ABSTRACT

Fixed-interval Bayesian smoothing in state-space systems has been addressed for a long time. However, as far as the measurement noise is concerned, only two cases have been addressed so far : the regular case, i.e. with positive definite covariance matrix; and the perfect measurement case, i.e. with zero measurement noise. In this paper we address the smoothing problem in the intermediate case where the measurement noise covariance is positive semi definite (p.s.d.) with arbitrary rank. We exploit the singularity of the model in order to transform the original state-space system into a pairwise Markov model (PMC) with reduced state dimension. Finally, the a posteriori Markovianity of the reduced state enables us to propose a family of fixed-interval smoothing algorithms.

1. INTRODUCTION

Let us consider the state space system

$$\begin{cases} \mathbf{x}_{n+1} & = & \mathbf{F}_n \mathbf{x}_n + \mathbf{G}_n \mathbf{u}_n \\ \mathbf{y}_n & = & \mathbf{H}_n \mathbf{x}_n + \mathbf{v}_n \end{cases}, \qquad (1)$$

in which $\mathbf{x}_n \in \mathbb{R}^{n_\mathbf{x}}$ is the state, $\mathbf{y}_n \in \mathbb{R}^{n_\mathbf{y}}$ the observation, $\mathbf{u}_n \in \mathbb{R}^{n_\mathbf{u}}$ the process noise and $\mathbf{v}_n \in \mathbb{R}^{n_\mathbf{v}}$ the measurement noise. Processes $\mathbf{u} = \{\mathbf{u}_n\}_{n\in\mathbb{N}}$ and $\mathbf{v} = \{\mathbf{v}_n\}_{n\in\mathbb{N}}$ are zero-mean, independent, jointly independent and independent of $\mathbf{x}_0$. So $(\mathbf{x}_n, \mathbf{y}_n)$ is a hidden Markov chain (HMC) with continuous hidden state $\mathbf{x}_n$. Fixed-interval smoothing consists in estimating $\mathbf{x}_n$ from $\mathbf{y}_{0:N}$ for $0 \leq n \leq N$. In the case where the covariance matrix $\mathbf{R}_n$ of $\mathbf{v}_n$ is positive definite, many algorithms have been derived by using such different methods as calculus of variations [1], maximum a posteriori [2] [3], orthogonal projections [4], the innovations approach [5], the two-filter form [6] [7], complementary models [8] or the Bayesian approach [9] [10] (modern surveys can also be found e.g. in [11, ch. 10] [8] or [12]). The most well-known algorithms are now the Bryson-Frazier algorithm [1], the Rauch-Tung-Streibel (RTS) algorithm [3] and the two-filter algorithm [6] [7].

On the other hand, Bayesian restoration in the case where the measurement noise covariance matrix is singular has received little attention so far, even though this case is of interest in a number of applications, including speech enhancement [13] or multiple target tracking [14] (see also [15], in which however only the filtering problem is addressed). In this paper we thus address fixed-interval smoothing in the singular measurement noise case, i.e., the case where $\mathbf{R}_n$ is a singular matrix. More precisely, we exploit the singularity of $\mathbf{R}_n$ in order to transform the original state-space system into an (equivalent) stochastic dynamical system but with reduced state dimension. The transformed system happens to be a PMC model, and as such the hidden process (even though it is not Markovian) is Markovian conditionnally on the observations. This key computational property finally enables us to develop fixed interval smoothing algorithms in the transformed system - and therefore, equivalently, in the original one.

The paper is organised as follows. Section 2 is devoted to the transformation of the HMC model (1) into a reduced state PMC model. In section 3 we propose Bayesian fixed-interval smoothing algorithms in a general (i.e., not necessarily linear and Gaussian) PMC model. Following the Bayesian point of view for deriving minimum mean-square error (MMSE) algorithms for state-space systems (see e.g. [16]), in section 4 we inject the Gaussian assumption into the algorithms of section 3, which eventually yields our fixed interval smoothing algorithms for singular state-space systems. Section 5 concludes the paper.

2. STATE-SPACE TRANSFORM

We address the fixed-interval smoothing problem in the singular measurements case, i.e. in the case where $\mathbf{R}_n$ is positive semi-definite. Some of the classical fixed-interval smoothing algorithms (see e.g. [17] for a recent review), originally designed for regular state-space systems, still hold in the singular case, but some others cannot be used any longer. In this paper we thus develop an alternative technique. Let r = rank$(\mathbf{R}_n) \in \{0, 1, \ldots, n_\mathbf{y}-1\}$. Following [15], we transform the state space model (1) in order to reduce by $m = n_{\mathbf{y}} - r$ the order of state $\mathbf{x}_n$. In sections 3 and 4 it will remain to design smoothing algorithms for this reduced-order linear stochastic system.

2.1. State-space transform

Since $\mathbf{R}_n$ has m zero eigenvalues, there exists a square non-singular matrix $\mathbf{M}_n$ satisfying

$$\mathbf{M}_n \mathbf{R}_n \mathbf{M}_n^T = \begin{bmatrix} \mathbf{0}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_r \end{bmatrix}, \tag{2}$$

in which $\mathbf{0}_m$ denotes the $m \times m$ null matrix and $\mathbf{I}_r$ the $r \times r$ identity matrix. Let $\overline{\mathbf{y}}_n = \mathbf{M}_n \mathbf{y}_n$ and $\overline{\mathbf{v}}_n = \mathbf{M}_n \mathbf{v}_n$; then from (1) we get

$$\underbrace{\begin{bmatrix} \overline{\mathbf{y}}_n^p \\ \overline{\mathbf{y}}_n^r \end{bmatrix}}_{\overline{\mathbf{y}}_n} = \underbrace{\begin{bmatrix} \overline{\mathbf{H}}_n^p \\ \overline{\mathbf{H}}_n^r \end{bmatrix}}_{\overline{\mathbf{H}}_n = \mathbf{M}_n \mathbf{H}_n} \mathbf{x}_n + \underbrace{\begin{bmatrix} \mathbf{0}_{m\times 1} \\ \overline{\mathbf{v}}_n^r \end{bmatrix}}_{\overline{\mathbf{v}}_n}, \tag{3}$$

and we see that $\overline{\mathbf{y}}_n$ is divided into a perfect part $(\overline{\mathbf{y}}_n^p)_{m\times 1}$ (the unnoisy part) and a regular one $(\overline{\mathbf{y}}_n^r)_{r\times 1}$. Since m linear functionals of the state $\mathbf{x}_n$ are known once $\overline{\mathbf{y}}_n$ is known, there is no need to estimate them, and this is why one can reduce by m the order of the state-space system, as we now see. Let us assume that

$$n_{\mathbf{x}} \geq m, \tag{4}$$

and that in (3)

$$\text{rank}(\overline{\mathbf{H}}_n^p)_{m \times n_{\mathbf{x}}} = m. \tag{5}$$

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

$$\underbrace{\begin{bmatrix} (\mathbf{U}_n)_{(n_{\mathbf{x}}-m)\times n_{\mathbf{x}}} \\ (\overline{\mathbf{H}}_n^p)_{m \times n_{\mathbf{x}}} \end{bmatrix}}_{\mathbf{T}_n} \mathbf{x}_n = \begin{bmatrix} \overline{\mathbf{x}}_n \\ \overline{\mathbf{y}}_n^p \end{bmatrix} \tag{6}$$

is reversible, and finally $\mathbf{T}_n$ and $\mathbf{M}_n$ enable us to transform the original linear state-space model (1) into the equivalent state-space system

$$\underbrace{\begin{bmatrix} \overline{\mathbf{x}}_{n+1} \\ \overline{\mathbf{y}}_{n+1}^p \end{bmatrix}}_{\mathbf{T}_{n+1}\mathbf{x}_{n+1}} = \mathbf{T}_{n+1} \mathbf{F}_n \mathbf{T}_n^{-1} \underbrace{\begin{bmatrix} \overline{\mathbf{x}}_n \\ \overline{\mathbf{y}}_n^p \end{bmatrix}}_{\mathbf{T}_n \mathbf{x}_n} + \mathbf{T}_{n+1} \mathbf{G}_n \mathbf{u}_n. \tag{7}$$

On the other hand from (3) and the first equation of (1) we have

$$\begin{aligned} \overline{\mathbf{y}}_{n+1}^r &= \overline{\mathbf{H}}_{n+1}^r \mathbf{x}_{n+1} + \overline{\mathbf{v}}_{n+1}^r \\ &= \overline{\mathbf{H}}_{n+1}^r \mathbf{F}_n \mathbf{T}_n^{-1} \begin{bmatrix} \overline{\mathbf{x}}_n \\ \overline{\mathbf{y}}_n^p \end{bmatrix} + \overline{\mathbf{H}}_{n+1}^r \mathbf{G}_n \mathbf{u}_n + \overline{\mathbf{v}}_{n+1}^r \end{aligned} \tag{8}$$

Gathering (7) and (8), we eventually get the reduced-order linear dynamic stochastic system :

$$\underbrace{\begin{bmatrix} \overline{\mathbf{x}}_{n+1} \\ \overline{\mathbf{y}}_{n+1} \end{bmatrix}}_{\mathbf{z}_{n+1}} = \underbrace{\begin{bmatrix} \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}} & \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\overline{\mathbf{y}}} \\ \overline{\mathcal{F}}_n^{\overline{\mathbf{y}},\overline{\mathbf{x}}} & \overline{\mathcal{F}}_n^{\overline{\mathbf{y}},\overline{\mathbf{y}}} \end{bmatrix}}_{\overline{\mathcal{F}}_n} \begin{bmatrix} \overline{\mathbf{x}}_n \\ \overline{\mathbf{y}}_n \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{U}_{n+1} & \mathbf{0} \\ \overline{\mathbf{H}}_{n+1} & \widetilde{\mathbf{I}}_r \end{bmatrix} \begin{bmatrix} \mathbf{G}_n \mathbf{u}_n \\ \overline{\mathbf{v}}_{n+1}^r \end{bmatrix}}_{\overline{\mathbf{w}}_n}, \tag{9}$$

with

$$\overline{\mathcal{F}}_n = \underbrace{\begin{bmatrix} \mathbf{T}_{n+1} \\ \overline{\mathbf{H}}_{n+1}^r \end{bmatrix}}_{(n_{\mathbf{x}}+r)\times n_{\mathbf{x}}} \underbrace{[\mathbf{F}_n \mathbf{T}_n^{-1} \ , \ \mathbf{0}]}_{n_{\mathbf{x}} \times (n_{\mathbf{x}}+r)}, \tag{10}$$

$$\widetilde{\mathbf{I}}_r = \begin{bmatrix} \mathbf{0}_{m\times r} \\ \mathbf{I}_r \end{bmatrix}. \tag{11}$$

2.2. Markovianity of the reduced state model (9)

$\{\mathbf{z}_n\}_n$ is a Markov chain (MC), so model (9) defines a so-called PMC. PMC models have been first introduced in the discrete state case and have been applied in the context of image segmentation [18]. Kalman Filtering for continuous state PMC has also been addressed, see [19] [20], and parameter estimation via the EM algorithm is available too [21]. The term ”pairwise” here emphasizes the fact that even though $\mathbf{z}_n = (\overline{\mathbf{x}}_n, \overline{\mathbf{y}}_n)$ is an MC, the marginal process $\overline{\mathbf{x}}$ in model (9) is indeed not Markovian. However, in a PMC the conditional distribution $p(\overline{\mathbf{x}}|\overline{\mathbf{y}})$ is Markovian; this property enables the developpement of efficient Bayesian restoration algorithms, and in particular, in the context of this paper, of fixed interval smoothing algorithms.

3. BAYESIAN FIXED-INTERVAL SMOOTHING ALGORITHMS IN PMC

For notational simplicity let us set $\overline{\mathbf{x}}_n = \mathbf{x}_n$, $\overline{\mathbf{y}}_n = \mathbf{y}_n$ and $\mathbf{z}_n = [\mathbf{x}_n^T, \mathbf{y}_n^T]^T$. Let us start by the general case, i.e. the case of the non-linear and/or non Gaussian PMC model :

$$\mathbf{z}_{n+1} = g(\mathbf{z}_n, \overline{\mathbf{w}}_n). \tag{12}$$

The aim of this section is to propose fixed-interval Bayesian smoothing algorithms for model (12), i.e. we want to compute the smoothing probability density function (pdf) $p(\mathbf{x}_n|\mathbf{y}_{0:N})$ for all n, $0 \leq n \leq N$.

Recently most of the existing MMSE smoothing algorithms (as well as some new alternatives) for classical state-space systems (and therefore for HMC with continuous state) have been gathered and classified into a commun unifying framework [17]. That same classification has also been extended to the context of non-symmetrical triplet Markov chains (TMC) [22]. Let us adapt this classification to the context of model (12). All the algorithms described in propositions 2 to 4 combine one or two densities out of the set $\alpha_n \stackrel{\text{def}}{=} p(\mathbf{x}_n|\mathbf{y}_{0:n})$, $\beta_n \stackrel{\text{def}}{=} p(\mathbf{y}_{n+1:N}|\mathbf{z}_n)$, $\gamma_n \stackrel{\text{def}}{=} p(\mathbf{x}_n|\mathbf{y}_{n:N})$

and $\eta_n \stackrel{\text{def}}{=} p(\mathbf{y}_{0:n-1}|\mathbf{z}_n)$. So we begin with the following proposition :

Proposition 1 *$\alpha_n = p(\mathbf{x}_n|\mathbf{y}_{0:n})$ and $\widetilde{\alpha}_n = p(\mathbf{x}_n|\mathbf{y}_{0:n+1})$ can be computed recursively (in the forward direction) as*

$$\left\{ \begin{array}{lcl} \widetilde{\alpha}_n & = & \frac{p(\mathbf{y}_{n+1}|\mathbf{z}_n)\times\alpha_n}{p(\mathbf{y}_{n+1}|\mathbf{y}_{0:n})=\int p(\mathbf{y}_{n+1}|\mathbf{z}_n)\alpha_n d\mathbf{x}_n} \\ \alpha_{n+1} & = & \int p(\mathbf{x}_{n+1}|\mathbf{z}_n,\mathbf{y}_{n+1})\times\widetilde{\alpha}_n d\mathbf{x}_n \end{array} \right. ; \quad (13)$$

$\beta_n = p(\mathbf{y}_{n+1:N}|\mathbf{z}_n)$ and $\widetilde{\beta}_n = p(\mathbf{y}_{n+2:N}|\mathbf{z}_n,\mathbf{y}_{n+1})$ can be computed recursively (in the backward direction) as

$$\left\{ \begin{array}{lcl} \widetilde{\beta}_n & = & \int p(\mathbf{x}_{n+1}|\mathbf{z}_n,\mathbf{y}_{n+1})\times\beta_{n+1}\, d\mathbf{x}_{n+1} \\ \beta_n & = & p(\mathbf{y}_{n+1}|\mathbf{z}_n)\,\widetilde{\beta}_n \end{array} \right. ; \quad (14)$$

$\eta_n = p(\mathbf{y}_{0:n-1}|\mathbf{z}_n)$ and $\widetilde{\eta}_n = p(\mathbf{y}_{0:n-2}|\mathbf{z}_n,\mathbf{y}_{n-1})$ can be computed recursively (in the forward direction) as

$$\left\{ \begin{array}{lcl} \widetilde{\eta}_{n+1} & = & \int p(\mathbf{x}_n|\mathbf{z}_{n+1},\mathbf{y}_n)\times\eta_n\, d\mathbf{x}_n \\ \eta_{n+1} & = & p(\mathbf{y}_n|\mathbf{z}_{n+1})\times\widetilde{\eta}_{n+1} \end{array} \right. ; \quad (15)$$

and $\gamma_n = p(\mathbf{x}_n|\mathbf{y}_{n:N})$ and $\widetilde{\gamma}_n = p(\mathbf{x}_n|\mathbf{y}_{n-1:N})$ can be computed recursively (in the backward direction) as

$$\left\{ \begin{array}{lcl} \widetilde{\gamma}_{n+1} & = & \frac{p(\mathbf{y}_n|\mathbf{z}_{n+1})\times\gamma_{n+1}}{p(\mathbf{y}_n|\mathbf{y}_{n+1:N})=\int p(\mathbf{y}_n|\mathbf{z}_{n+1})\gamma_{n+1}d\mathbf{x}_{n+1}} \\ \gamma_n & = & \int p(\mathbf{x}_n|\mathbf{z}_{n+1},\mathbf{y}_n)\times\widetilde{\gamma}_{n+1}d\mathbf{x}_{n+1} \end{array} \right. . \quad (16)$$

We now turn to the computation of the smoothing density $p(\mathbf{x}_n|\mathbf{y}_{0:N})$ itself. We have the following propositions.

Proposition 2 *$p(\mathbf{x}_n|\mathbf{y}_{0:N})$ can be computed in the backward direction by*

$$p(\mathbf{x}_n|\mathbf{y}_{0:N}) = \int p(\mathbf{x}_{n+1}|\mathbf{y}_{0:N})p(\mathbf{x}_n|\mathbf{x}_{n+1},\mathbf{y}_{0:N})d\mathbf{x}_{n+1}, \quad (17)$$

with

$$\begin{aligned} p(\mathbf{x}_n|\mathbf{x}_{n+1},\mathbf{y}_{0:N}) &\propto p(\mathbf{x}_{n+1}|\mathbf{z}_n,\mathbf{y}_{n+1})\times\widetilde{\alpha}_n && (18)\\ &\propto p(\mathbf{x}_n|\mathbf{z}_{n+1},\mathbf{y}_n)\times\eta_n && (19)\\ &\propto \frac{p(\mathbf{x}_n|\mathbf{z}_{n+1},\mathbf{y}_n)\times\alpha_n}{p(\mathbf{x}_n|\mathbf{y}_n)} && (20)\\ &\propto p(\mathbf{z}_{n+1}|\mathbf{z}_n)\times\eta_n\times p(\mathbf{x}_n|\mathbf{y}_n) && (21) \end{aligned}$$

Proposition 3 *$p(\mathbf{x}_n|\mathbf{y}_{0:N})$ can be computed in the forward direction by*

$$p(\mathbf{x}_{n+1}|\mathbf{y}_{0:N}) = \int p(\mathbf{x}_n|\mathbf{y}_{0:N})p(\mathbf{x}_{n+1}|\mathbf{x}_n,\mathbf{y}_{0:N})d\mathbf{x}_n, \quad (22)$$

with

$$\begin{aligned} p(\mathbf{x}_{n+1}|\mathbf{x}_n,\mathbf{y}_{0:N}) &\propto p(\mathbf{x}_{n+1}|\mathbf{z}_n,\mathbf{y}_{n+1})\times\beta_{n+1} && (23)\\ &\propto p(\mathbf{x}_n|\mathbf{z}_{n+1},\mathbf{y}_n)\times\widetilde{\gamma}_{n+1} && (24)\\ &\propto \frac{p(\mathbf{x}_{n+1}|\mathbf{z}_n,\mathbf{y}_{n+1})\times\gamma_{n+1}}{p(\mathbf{x}_{n+1}|\mathbf{y}_{n+1})} && (25)\\ &\propto p(\mathbf{z}_n|\mathbf{z}_{n+1})\times\beta_{n+1}\times p(\mathbf{x}_{n+1}|\mathbf{y}_{n+1}) && (26) \end{aligned}$$

Proposition 4 *The smoothing pdf $p(\mathbf{x}_n|\mathbf{y}_{0:N})$ can be computed as*

$$\begin{aligned} p(\mathbf{x}_n|\mathbf{y}_{0:N}) &\propto \alpha_n\times\beta_n && (27)\\ &\propto \gamma_n\times\eta_n && (28)\\ &\propto \frac{\alpha_n\times\gamma_n}{p(\mathbf{x}_n|\mathbf{y}_n)} && (29)\\ &\propto \eta_n\times\beta_n\times p(\mathbf{x}_n|\mathbf{y}_n). && (30) \end{aligned}$$

4. MMSE BAYESIAN SMOOTHING ALGORITHMS.

Let us now come back to the linear PMC (9). The general algorithms of Props. 2 to 4 reduce to MMSE fixed-interval smoothing algorithms if we further inject the Gaussian assumption. So from now on we also assume that $\mathbf{z}_0$ and $\overline{\mathbf{w}}_n$ are Gaussian for all n, which in turn holds if the original state-space model (1) is Gaussian, i.e. if $\mathbf{x}_0$, $\mathbf{u}_n$ and $\mathbf{v}_n$ are Gaussian for all n. Let us set

$$\mathbf{x}_0 \sim \mathcal{N}(\widehat{\mathbf{x}}_0,\mathbf{P}_0), \overline{\mathbf{w}}_n \sim \mathcal{N}(\mathbf{0},\begin{bmatrix} \overline{\mathbf{Q}}_n & \overline{\mathbf{S}}_n \\ \overline{\mathbf{S}}_n^T & \overline{\mathbf{R}}_n \end{bmatrix}). \quad (31)$$

In this case, all the pdfs in §3 are Gaussian. Let us set

$$\begin{aligned} p(\mathbf{x}_n|\mathbf{y}_{i:j}) &\sim \mathcal{N}(\widehat{\mathbf{x}}_{n|i:j},\mathbf{P}_{n|i:j}), && (32)\\ \mu_{n|i:j} &= \mathbf{P}_{n|i:j}^{-1}\widehat{\mathbf{x}}_{n|i:j} && (33) \end{aligned}$$

for all n, i, j with $0 \le n \le N$ and $0 \le i \le j \le N$ ($\mu_{n|i:j}$ and $\mathbf{P}_{n|i:j}^{-1}$ are respectively the information vector and matrix associated with $p(\mathbf{x}_n|\mathbf{y}_{i:j})$).

The general algorithms of Props. 2 to 4 compute $p(\mathbf{x}_n|\mathbf{y}_{0:N})$ from α_n (or η_n) and/or γ_n (or β_n). In the Gaussian case, this amounts to computing the parameters of $p(\mathbf{x}_n|\mathbf{y}_{0:N})$ from those of α_n (or η_n) and/or γ_n (or β_n). More precisely, (13) to (30) reduce to equations which compute $\arg\max_{\mathbf{x}_n} p(\mathbf{x}_n|\mathbf{y}_{0:N})$ (i.e., $\widehat{\mathbf{x}}_{n|0:N}$), and the associated covariance matrix, from $\arg\max_{\mathbf{x}_n}\alpha_n = \widehat{\mathbf{x}}_{n|0:n}$ (or $\arg\max_{\mathbf{x}_n}\eta_n$) and/or $\arg\max_{\mathbf{x}_n}\gamma_n = \widehat{\mathbf{x}}_{n|n:N}$ (or $\arg\max_{\mathbf{x}_n}\beta_n$), as well as the associated covariance matrice(s). In practice, these equations can be derived by systematically applying some simple results for Gaussian variables; each one of the twelve general algorithms in Props. 2, 3 and 4 then reduces to a particular Kalman smoothing algorithm; some of them are PMC extensions of such classical Kalman-like smoothing algorithms as, e.g., the RTS algorithm [3], the two-filter algorithm [6] [7], or the general two-filter algorithm [11, Theorem 10.4.1] (see [17] for details). For instance, algorithm (18)-(17) reduces in the Gaussian case to the following PMC RTS algorithm :

Proposition 5 *Let (9) and (31) hold. Let*

$$\mathbf{K}_{n|0:N} = \mathbf{P}_{n|0:n+1}[\overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}} - \overline{\mathbf{S}}_n(\overline{\mathbf{R}}_n)^{-1}\overline{\mathcal{F}}_n^{\overline{\mathbf{y}},\overline{\mathbf{x}}}]^T\mathbf{P}_{n+1|0:n+1}^{-1}. \quad (34)$$

Then

$$\widehat{\mathbf{x}}_{n|0:N} = \widehat{\mathbf{x}}_{n|0:n+1} + \mathbf{K}_{n|0:N}[\widehat{\mathbf{x}}_{n+1|0:N} - \widehat{\mathbf{x}}_{n+1|0:n+1}], \quad (35)$$

$$\mathbf{P}_{n|0:N} = \mathbf{P}_{n|0:n+1} - \mathbf{K}_{n|0:N}[\mathbf{P}_{n+1|0:n+1} - \mathbf{P}_{n+1|0:N}]\mathbf{K}_{n|0:N} \quad (36)$$

So $(\widehat{\mathbf{x}}_{n|0:N}, \mathbf{P}_{n|0:N})$ *can be propagated (in the backward direction) provided* $(\widehat{\mathbf{x}}_{n|0:n+1}, \mathbf{P}_{n|0:n+1})$ *and* $(\widehat{\mathbf{x}}_{n+1|0:n+1}, \mathbf{P}_{n+1|0:n+1})$ *are known; these, in turn, can be computed recursively (in the forward direction) by the PMC Kalman filter algorithm, see [19]* [1].

Let us finally mention the Bryson and Frazier (BF) algorithm [1], which cannot actually be derived from Bayesian considerations, even though it is closely related to the RTS algorithm. The BF algorithm was introduced for the first time (in the continuous time case) as a solution of a deterministic least mean square problem by using the variational approach [1] (see also [11, chap. 10 & 6] and [23, pp. 223-25]). The discrete time version of this algorithm then appeared in various publications like e.g. [24] [3] [25] [11]. The link between the RTS and the BF algorithms has been established for the first time in [3]. As in the classical HMC framework, the PMC RTS algorithm can also be implemented by an algorithm which extends to PMC the BF algorithm :

Proposition 6 *Let (9) and (31) hold. Let*

$$\begin{aligned}
\widehat{\mathbf{y}}_{n+1|0:n} &= \overline{\mathcal{F}}_n^{\overline{\mathbf{y}},\overline{\mathbf{x}}}\widehat{\mathbf{x}}_{n|0:n} + \overline{\mathcal{F}}_n^{\overline{\mathbf{y}},\overline{\mathbf{y}}}\mathbf{y}_n \\
\mathbf{P}^{\mathbf{y}}_{n+1|0:n} &= \overline{\mathcal{F}}_n^{\overline{\mathbf{y}},\overline{\mathbf{x}}}\mathbf{P}_{n|0:n}(\overline{\mathcal{F}}_n^{\overline{\mathbf{y}},\overline{\mathbf{x}}})^T + \overline{\mathbf{R}}_n \\
\mathbf{K}_n^{\lambda} &= \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}}[\mathbf{I} - \mathbf{P}_{n|0:n}(\overline{\mathcal{F}}_n^{\overline{\mathbf{y}},\overline{\mathbf{x}}})^T(\mathbf{P}^{\mathbf{y}}_{n+1|0:n})^{-1}\overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}}].
\end{aligned}$$

Then $(\widehat{\mathbf{x}}_{n|0:N}$ *and* $\mathbf{P}_{n|0:N})$ *can be computed as*

$$\widehat{\mathbf{x}}_{n|0:N} = \widehat{\mathbf{x}}_{n|0:n+1} + \mathbf{P}_{n|0:n}(\mathbf{K}_n^{\lambda})^T\lambda_{n+1}, \quad (37)$$

$$\mathbf{P}_{n|0:N} = \mathbf{P}_{n|0:n+1} - \mathbf{P}_{n|0:n}(\mathbf{K}_n^{\lambda})^T\Lambda_{n+1}\mathbf{K}_n^{\lambda}\mathbf{P}_{n|0:n}, \quad (38)$$

in which the parameters

$$\lambda_n = \mathbf{P}_{n|0:n}^{-1}[\widehat{\mathbf{x}}_{n|0:N} - \widehat{\mathbf{x}}_{n|0:n}] \quad (39)$$

$$\Lambda_n = \mathbf{P}_{n|0:n}^{-1}[\mathbf{P}_{n|0:n} - \mathbf{P}_{n|0:N}]\mathbf{P}_{n|0:n}^{-1}, \quad (40)$$

are computed recursively (with $\lambda_N = \mathbf{0}$ *and* $\Lambda_N = \mathbf{0}$*) as follows :*

$$\lambda_n = (\mathbf{K}_n^{\lambda})^T\lambda_{n+1} + (\overline{\mathcal{F}}_n^{\overline{\mathbf{y}},\overline{\mathbf{x}}})^T(\mathbf{P}^{\mathbf{y}}_{n+1|0:n})^{-1}[\mathbf{y}_{n+1} - \widehat{\mathbf{y}}_{n+1|0:n}], \quad (41)$$

$$\Lambda_n = (\mathbf{K}_n^{\lambda})^T\Lambda_{n+1}\mathbf{K}_n^{\lambda} + (\overline{\mathcal{F}}_n^{\overline{\mathbf{y}},\overline{\mathbf{x}}})^T(\mathbf{P}^{\mathbf{y}}_{n+1|0:n})^{-1}\overline{\mathcal{F}}_n^{\overline{\mathbf{y}},\overline{\mathbf{x}}}. \quad (42)$$

As we can see, the PMC BF algorithm is still a two pass one. In the forward pass, the filtering and one-step smoothing parameters are computed by the PMC KF. One then propagates in the backward pass the new variables λ_n and Λ_n via (41) and (42). The smoothing parameters of interest are finally computed as a combination of the forward and the backward quantities, see (37)-(38).

[1] Due to the symmetry of model (9), $\mathbf{y}_n$ must be replaced by $\mathbf{y}_{n+1}$ in [19].

5. CONCLUSION

In this paper we have addressed the fixed-interval smoothing problem in state-space systems with singular measurement noise. We first transformed the original (singular) HMC model into an equivalent PMC model with regular noise and reduced state dimension. Though the transformed system is no longer an HMC (in particular the hidden state is no longer Markovian), it enables Bayesian restoration because the state is Markovian conditionally on the observations. We finally proposed a set a Bayesian fixed-interval smoothing algorithms.

6. REFERENCES

[1] A.E. Bryson and M. Frazier, “Smoothing for linear and non-linear dynamic systems,” *Tech. Rep., Aero. Syst. Dvi. Wrigth-Patterson Air Force Base*, pp. 353–364, 1963.

[2] H.E. Rauch, “Solutions to smoothing problem,” *IEEE Transactions on Automatic Control*, vol. 8, pp. 371–372, October 1963.

[3] H. E. Rauch, F. Tung, and C. T. Striebel, “Maximum likelihood estimates of linear dynamic systems,” *AIAA J.*, vol. 3, no. 8, pp. 1445–1450, August 1965.

[4] J. S. Meditch, “Orthogonal projection and discrete optimal linear smoothing,” *SIAM J. Contr*, vol. 5, pp. 74–89, January 1967.

[5] T. Kailath and P. Frost, “An innovations approach to least squares estimation, part ii: Linear smoothing in additive white noise,” *IEEE Transactions on Automatic Control*, vol. 13, pp. 655–660, December 1968.

[6] D. Q. Mayne, “A solution to the smoothing problem for linear dynamical systems,” *Automatica*, vol. 4, pp. 73–92, 1966.

[7] D. C. Fraser and J. E. Potter, “The optimum linear smoother as a combination of two optimum linear filters,” *IEEE Tr. Aut. Control*, vol. 7, pp. 387–90, Aug. 1969.

[8] H. L. Weinert, *Fixed interval smoothing for state-space models*, The Kluwer international series in engineering and computer science. Kluwer Academic Publishers, Boston, 2001.

[9] R. C. K. Lee, “Optimal estimation: Identification and control,” 1964, pp. 69–73, Cambridge, MA, M.I.T. Press.

[10] M. Askar and H. Derin, “A recursive algorithm for the Bayes solution of the smoothing problem,” *IEEE*

Transactions on Automatic Control, vol. 26, pp. 558–560, 1981.

[11] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation*, Prentice Hall Information and System Sciences Series. Prentice Hall, Upper Saddle River, NJ, 2000.

[12] O. Cappé, E. Moulines, and T. Rydén, *Inference in Hidden Markov Models*, Springer-Verlag, 2005.

[13] J. D. Gibson, B. Koo, and S. D. Gray, "Filtering of colored noise for speech enhancement and coding," *IEEE Transactions on signal processing*, vol. 39, NO. 8, pp. 1732–42, August 1991.

[14] F. Gustafsson, *Adaptive Filtering and Change Detection*, John Wiley & Sons, Chichester, 2000.

[15] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice Hall, Englewood Cliffs, New Jersey, 1979.

[16] Y. C. Ho and R. C. K. Lee, "A Bayesian approach to problems in stochastic estimation and control," *IEEE Transactions on Automatic Control*, vol. 9, pp. 333–339, October 1964.

[17] B. Ait-El-Fquih and F. Desbouvries, "On Bayesian Fixed-Interval Smoothing Algorithms," *IEEE Transactions on Automatic Control*, vol. 53, no. 10, pp. 2437–42, November 2008.

[18] W. Pieczynski, "Pairwise Markov chains," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 5, pp. 634–39, May 2003.

[19] W. Pieczynski and F. Desbouvries, "Kalman filtering using pairwise Gaussian models," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 03)*, Hong-Kong, 2003.

[20] F. Desbouvries and W. Pieczynski, "Particle filtering in pairwise and triplet Markov chains," in *Proc. IEEE - EURASIP Workshop on Nonlinear Signal and Image Processing*, Grado-Gorizia, Italy, June 8-11 2003.

[21] B. Ait-El-Fquih and F. Desbouvries, "Unsupervised signal restoration in partially observed Markov chains," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 06)*, Toulouse, France, May 15-19 2006.

[22] B. Ait el Fquih and F. Desbouvries, "Exact and approximate Bayesian smoothing algorithms in partially observed Markov chains," in *Proceedings of the IEEE Nonlinear Statistical Signal Processing Workshop (NSSPW'06)*, Cambridge, UK, September 13-15 2006.

[23] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, vol. 64 of *Mathematics in Science and Engineering*, Academic Press, San Diego, 1970.

[24] H. Cox, "On the Estimation of State Variables and Parameters for Noisy Dynamic Systems," *IEEE Transactions on Automatic Conrol*, vol. 9, no. 1, pp. 5–12, January 1964.

[25] A.E. Bryson and Y.C. Ho, *Applied Optimal Control*, Hemisphere, New York, 1975.