RAO-BLACKWELLIZED PARTICLE SMOOTHERS FOR MIXED LINEAR/NONLINEAR STATE-SPACE MODELS

Fredrik Lindsten[⋆], *Pete Bunch*[†], *Simon J. Godsill*[†] *and Thomas B. Schön*[⋆]

[⋆] Division of Automatic Control, Linköping University, Linköping, Sweden
E-mail: lindsten@isy.liu.se,schon@isy.liu.se
[†]Department of Engineering, University of Cambridge, Cambridge, UK
E-mail: pb404@cam.ac.uk,sjg@eng.cam.ac.uk

ABSTRACT

We consider the smoothing problem for a class of conditionally linear Gaussian state-space (CLGSS) models, referred to as mixed linear/nonlinear models. In contrast to the better studied hierarchical CLGSS models, these allow for an intricate cross dependence between the *linear* and the *nonlinear* parts of the state vector. We derive a Rao-Blackwellized particle smoother (RBPS) for this model class by exploiting its tractable substructure. The smoother is of the forward filtering/backward simulation type. A key feature of the proposed method is that, unlike existing RBPS for this model class, the *linear* part of the state vector is marginalized out in both the forward direction and in the backward direction.

***Index Terms*—** Rao-Blackwellization, particle smoothing, backward simulation, sequential Monte Carlo.

1. INTRODUCTION

Particle filters (PF) and particle smoothers (PS) are useful for state inference in nonlinear state-space models (SSM) [1, 2]. It is well recognized that the Rao-Blackwellized PF (RBPF) [3, 4, 5], can be used to address the filtering problem in conditionally linear Gaussian state-space (CLGSS) models. By exploiting the tractable substructure present in these models, the RBPF results in more accurate estimators than a standard PF [6, 7] and it can therefore be used for filtering in even more challenging, e.g. high-dimensional, models.

However, Rao-Blackwellization has not been as well explored for smoothing. Most Rao-Blackwellized particle smoothers (RBPS) have been focused on a type of hierarchical CLGSS models, for which the "nonlinear state" is Markovian [8, 9, 10]. Here, we consider instead a class of mixed linear/nonlinear SSMs, given by

$$u_{t+1} = g(u_t) + B(u_t)z_t + G(u_t)v_t^u, \tag{1a}$$

$$z_{t+1} = f(u_t) + A(u_t)z_t + F(u_t)v_t^z, \tag{1b}$$

$$y_t = h(u_t) + C(u_t)z_t + e_t, \tag{1c}$$

where $v_t^u \sim \mathcal{N}(0, I)$, $v_t^z \sim \mathcal{N}(0, I)$ and $e_t \sim \mathcal{N}(0, R(u_t))$. We assume that $Q(u_t) \triangleq G(u_t)G(u_t)^\mathsf{T}$ and $R(u_t)$ are invertible, but we do not assume invertibility of $F(u_t)F(u_t)^\mathsf{T}$. The state consists of two parts, $x_t = (u_t, z_t)$, where there is a nonlinear dependence on u_t (referred to as the *nonlinear state*) and an affine dependence on z_t (referred to as the *linear state*).

The first and the fourth authors were supported by: the project Calibrating Nonlinear Dynamical Models (Contract number: 621-2010-5876) funded by the Swedish Research Council and CADICS, a Linneaus Center also funded by the Swedish Research Council.

This model is CLGSS, but it is more involved than a hierarchical CLGSS models, since there is a cross-dependence between the linear and the nonlinear states. That is, the nonlinear state process alone is non-Markovian. This class of models arise, for instance, when the observations depend nonlinearly on a subset of the states in a system with linear dynamics. An RBPF for the mixed linear/nonlinear model was derived and used for terrain-aided aircraft navigation in [3]. In this contribution, we derive a novel RBPS, akin to the recent contributions for hierarchical CLGSS models [8, 9], but applicable to the model (1). The proposed method is based on the forward filter/backward simulator (FFBS) by [11]. A key property of the proposed method is that it only samples the nonlinear part of the state, both in the forward and backward directions, as opposed to, for instance [10, 12], who sample the full state in the backward direction.

For a vector μ and a positive semidefinite matrix $\Omega \succeq 0$, we write $\|\mu\|_\Omega^2 \triangleq \mu^\mathsf{T}\Omega\mu$. We write $|A|$ for matrix determinant and $\mathcal{N}(\mu, \Sigma)$ and $\mathcal{N}(x; \mu, \Sigma)$ for the Gaussian distribution and probability density function (PDF), respectively.

2. BACKGROUND

2.1. Particle filtering and smoothing

Consider first a standard, Markovian SSM: $x_{t+1} = f(x_t) + v_t$ and $y_t = h(x_t) + e_t$, where f and h are nonlinear functions, and v_t and e_t have known, tractable densities. A particle filter is a sequential Monte Carlo algorithm used to approximate the intractable filtering density, representing it with a set of weighted particles $\{x_{1:t}^i, w_t^i\}_{i=1}^N$, each of which is a state trajectory $x_{1:t}$,

$$\widehat{p}^N(dx_{1:t} \mid y_{1:t}) \triangleq \sum_{i=1}^N w_t^i \delta_{x_{1:t}^i}(dx_{1:t}). \tag{2}$$

In the simplest particle filter, the t-th set of particles are formed by sampling $x_{1:t-1}$ from the previous distribution and then x_t from an importance distribution. A weight is assigned to each particle to account for the difference between the proposal and the target density. Note that an approximation to $p(x_t \mid y_{1:t})$ is obtained by marginalization of (2), which equates to simply discarding $x_{1:t-1}$.

The term "smoothing" encompasses a number of related inference problems, but here we focus on the estimation of the complete joint smoothing distribution, $p(x_{1:T} \mid y_{1:T})$. This distribution is approximated at the final step of the particle filter [13]. However, this approximation suffer from the problem of path degeneracy, i.e. the number of unique particles decreases rapidly for $t \ll T$. A diverse set of particles may be generated by sampling state trajectories using the forward filtering/backward simulation (FFBS) algorithm [11].

978-1-4799-0356-6/13/$31.00 ©2013 IEEE

FFBS exploits a sequential factorization of the joint smoothing density $p(x_{1:T} \mid y_{1:T}) = p(x_T \mid y_{1:T}) \prod_{t=1}^{T-1} p(x_t \mid x_{t+1:T}, y_{1:T})$. A final state, x'_T is first sampled from the particle filter approximation $\widehat{p}^N(dx_T \mid y_{1:T})$. Then, working sequentially backwards from time T, each subsequent state x'_t is sampled from the backwards kernel, $p(x_t \mid x_{t+1:T}, y_{1:T})$. The resulting trajectory $\{x'_{1:T}\}$ is then a sample from the smoothing distribution.

Using the Markov property, the backward kernel may be expanded as

$$p(x_t \mid x_{t+1:T}, y_{1:T}) \propto p(x_{t+1} \mid x_t) p(x_t \mid y_{1:t}). \tag{3}$$

The second factor is approximated using the particle filter, leading to the representation, $\widehat{p}^N(dx_t \mid x_{t+1:T}, y_{1:T}) \triangleq \sum_{i=1}^N \tilde{w}^i_{t|T} \delta_{x^i_t}(dx_t)$, with $\tilde{w}^i_{t|T} \propto w^i_t \, p(x_{t+1} \mid x^i_t)$.

2.2. Rao-Blackwellized particle filtering

The structure of a CLGSS model (such as (1)) can be exploited by using an RBPF, by exploiting the factorization $p(u_{1:t}, z_t \mid y_{1:t}) = p(z_t \mid u_{1:t}, y_{1:t}) p(u_{1:t} \mid y_{1:t})$. Since the model is CLGSS, it holds that

$$p(z_t \mid u_{1:t}, y_{1:t}) = \mathcal{N}(z_t; \bar{z}_{t|t}(u_{1:t}), P_{t|t}(u_{1:t})). \tag{4}$$

A PF is used to estimate only the nonlinear state marginal density while conditional Kalman filters, one for each particle, are used to compute the moments for the linear state in (4). The resulting RBPF approximation is given by

$$\widehat{p}^N(du_{1:t}, dz_t \mid y_{1:t}) = \sum_{i=1}^N w^i_t \mathcal{N}(dz_t; \bar{z}^i_{t|t}, P^i_{t|t}) \delta_{u^i_{1:t}}(du_{1:t}).$$

The particle weights are given by the ratio of the Gaussian joint density $p(y_t, u_t \mid u_{1:t-1}, y_{1:t-1})$ and the importance density. See [3] for details. The reduced dimensionality of the particle approximation results in a reduction in variance of associated estimators [6, 7].

3. RAO-BLACKWELLIZED PARTICLE SMOOTHING

The new RBPS for the model (1) is an FFBS which uses the RBPF as a forward filter. The novelty lies in the construction of a backward simulator which samples only the nonlinear state in the backward pass. Difficulty arises because the nonlinear state process is non-Markovian. Practically, this means that the backward kernel cannot be expressed in a simple way, as in (3). Furthermore, it implies that the measurement likelihood depends on the complete history $u_{1:t}$; we must therefore sample whole trajectories produced by the RBPF. More precisely, let $u'_{t+1:T}$ be a partial, nonlinear backward trajectory. To extend this trajectory to time t, we draw one of the RBPF particles $\{u^i_{1:t}\}_{i=1}^N$ (with probabilities computed below), set $u'_{t:T} = \{u^i_t, u'_{t+1:T}\}$ and discard $u^i_{1:t-1}$. This procedure is repeated for each time $t = T-1, \ldots, 1$, resulting in a complete backward trajectory.

To compute the backward sampling probabilities, we note that

$$p(u_{1:t} \mid u_{t+1:T}, y_{1:T}) \propto p(y_{t+1:T}, u_{t+1:T} \mid u_{1:t}, y_{1:t}) p(u_{1:t} \mid y_{1:t}). \tag{5}$$

The second factor in this expression can be approximated by the forward RBPF, analogously to a standard FFBS. This results in a point-mass approximation of the backward kernel, given by

$$\widehat{p}^N(du_{1:t} \mid u_{t+1:T}, y_{1:T}) = \sum_{i=1}^N \tilde{w}^i_{t|T} \delta_{u^i_{1:t}}(du_{1:t}), \tag{6}$$

with

$$\tilde{w}^i_{t|T} \propto w^i_t \, p(y_{t+1:T}, u'_{t+1:T} \mid u^i_{1:t}, y_{1:t}). \tag{7}$$

It remains to find an expression for the predictive PDF in this expression (up to proportionality). In fact, this PDF can be computed straightforwardly by running a conditional Kalman filter from time t up to time T. However, using such an approach to calculate the weights at time t would require N separate Kalman filters to run over $T - t$ time steps, resulting in a total computational complexity scaling quadratically with T. To avoid this, we seek an efficient recursion for the weights (7). This is accomplished by propagating a set of statistics backward in time, as the trajectory $u'_{1:T}$ is generated. The idea is similar to that of [14, 9, 8] but our derivation is adapted to the mixed linear/nonlinear model (1). To start with, we express the predictive PDF as

$$\begin{aligned} &p(y_{t+1:T}, u_{t+1:T} \mid u_{1:t}, y_{1:t}) \\ &\quad = \int p(y_{t+1:T}, u_{t+1:T} \mid z_t, u_t) p(z_t \mid u_{1:t}, y_{1:t}) \, dz_t, \end{aligned} \tag{8}$$

where the second factor of the integrand is given by the RBPF (4). Hence, we seek an expression for the first factor of the integrand. The following two propositions, which will be used alternately in the backward simulation, provide the updating equations for a set of sufficient statistics for this PDF. For brevity, we write A_t for $A(u_t)$, etc.

Proposition 1 (Backward prediction). *Given $\widehat{\Omega}_{t+1}$ and $\widehat{\lambda}_{t+1}$ as in Proposition 2, for any $1 \le t \le T-1$,*

$$p(y_{t+1:T}, u_{t+1:T} \mid z_t, u_t) \propto Z_t \exp\left(-\tfrac{1}{2}\left(z_t^\mathsf{T} \Omega_t z_t - 2\lambda_t^\mathsf{T} z_t\right)\right),$$

where Z_t, $\Omega_t \succeq 0$ and λ_t depend on u_t, but are independent of z_t, and the proportionality is w.r.t. (u_t, z_t). The updated statistics are given by,

$$\begin{aligned} Z_t &= |M_t|^{-1/2} |Q_t|^{-1/2} \exp\left(-\tfrac{1}{2}\tau_t\right) \\ \Omega_t &= A_t^\mathsf{T}\left(\widehat{\Omega}_{t+1} - \widehat{\Omega}_{t+1} F_t M_t^{-1} F_t^\mathsf{T} \widehat{\Omega}_{t+1}\right) A_t + B_t^\mathsf{T} Q_t^{-1} B_t, \\ \lambda_t &= A_t^\mathsf{T}\left(I - \widehat{\Omega}_{t+1} F_t M_t^{-1} F_t^\mathsf{T}\right) m_t + B_t^\mathsf{T} Q_t^{-1}(u_{t+1} - g_t), \end{aligned}$$

with $m_t = \widehat{\lambda}_{t+1} - \widehat{\Omega}_{t+1} f_t$ and $M_t = F_t^\mathsf{T} \widehat{\Omega}_{t+1} F_t + I$ and

$$\tau_t = \|(u_{t+1} - g_t)\|^2_{Q_t^{-1}} + \|f_t\|^2_{\widehat{\Omega}_{t+1}} - 2\widehat{\lambda}_{t+1}^\mathsf{T} f_t - \|F_t^\mathsf{T} m_t\|^2_{M_t^{-1}}.$$

Proof. See Section 4. □

Proposition 2 (Update). *Given Ω_t and λ_t as in Proposition 1, for any $1 \le t \le T-1$,*

$$p(y_{t:T}, u_{t+1:T} \mid z_t, u_t) \propto \exp\left(-\tfrac{1}{2}\left(z_t^\mathsf{T} \widehat{\Omega}_t z_t - 2\widehat{\lambda}_t^\mathsf{T} z_t\right)\right),$$

where $\widehat{\Omega}_t \succeq 0$ and $\widehat{\lambda}_t$ depend on u_t, but are independent of z_t, and the proportionality is w.r.t. z_t. The updated statistics are given by,

$$\begin{aligned} \widehat{\Omega}_t &= \Omega_t + C_t^\mathsf{T} R_t^{-1} C_t, \\ \widehat{\lambda}_t &= \lambda_t + C_t^\mathsf{T} R_t^{-1}(y_t - h_t). \end{aligned}$$

Furthermore, at time T, it holds that

$$p(y_T \mid z_T, u_T) \propto \exp\left(-\tfrac{1}{2}\left(z_T^\mathsf{T} \widehat{\Omega}_T z_T - 2\widehat{\lambda}_T^\mathsf{T} z_T\right)\right),$$

with $\widehat{\Omega}_T = C_T^\mathsf{T} R_T^{-1} C_T$ and $\widehat{\lambda}_T = C_T^\mathsf{T} R_T^{-1}(y_T - h_T)$.

Proof. See Section 4. □

We thus have a recursion for updating the statistics Z_t, Ω_t and λ_t. Using these quantities, together with (4), we can solve the integral (8). This is formalized in the next proposition.

Proposition 3. *Let $\bar{z}_{t|t}$ and $P_{t|t} = \Gamma_{t|t}\Gamma_{t|t}^{\mathsf{T}}$ be given as in (4) and let Z_t, Ω_t and λ_t be given as in Proposition 1. Then,*

$$p(y_{t+1:T}, u_{t+1:T} \mid u_{1:t}, y_{1:t}) \propto Z_t |\Lambda_t|^{-1/2} \exp\left(-\tfrac{1}{2}\eta_t\right),$$

where the proportionality is w.r.t. $u_{1:t}$ and where,

$$\begin{aligned}\eta_t &= \|\bar{z}_{t|t}\|^2_{\Omega_t} - 2\lambda_t^{\mathsf{T}}\bar{z}_{t|t} - \|\Gamma_{t|t}^{\mathsf{T}}(\lambda_t - \Omega_t\bar{z}_{t|t})\|^2_{\Lambda_t^{-1}},\\ \Lambda_t &= \Gamma_{t|t}^{\mathsf{T}}\Omega_t\Gamma_{t|t} + I.\end{aligned}$$

Proof. See Section 4. □

By plugging this result into (7), we obtain an expression for the backward sampling weights. The resulting RBPS is given in Algorithm 1.

Algorithm 1 Rao-Blackwellized FFBS

1. **Forward filter:** Run an RBPF for time $t = 1, \ldots, T$. For each t, store $\{u_t^i, w_t^i, \bar{z}_{t|t}^i, \Gamma_{t|t}^i\}_{i=1}^N$.
2. **Initialize:** Draw $u'_T = u_T^i$ with probability w_T^i. Compute $\widehat{\Omega}_T$ and $\widehat{\lambda}_T$ as in Proposition 2.
3. **For** $t = T-1$ **to** 1**:**
   (a) For each forward filter particle, $i = 1, \ldots, N$:
      i. Compute $\{Z_t^i, \Omega_t^i, \lambda_t^i\}$ as in Proposition 1.
      ii. Compute $\{\Lambda_t^i, \eta_t^i\}$ as in Proposition 3.
      iii. Compute $\widetilde{W}_t^i = w_t^i Z_t^i |\Lambda_t^i|^{-1/2}\exp\left(-\frac{1}{2}\eta_t^i\right)$.
   (b) Normalize the weights, $\tilde{w}_{t|T}^i = \widetilde{W}_t^i / \sum_l \widetilde{W}_t^l$.
   (c) Set $J = i$ with probability $\tilde{w}_{t|T}^i$.
   (d) Set $u'_{t:T} = \{u_t^J, u'_{t:T}\}$ and $\{\Omega_t, \lambda_t\} = \{\Omega_t^J, \lambda_t^J\}$.
   (e) Compute $\{\widehat{\Omega}_t, \widehat{\lambda}_t\}$ as in Proposition 2.

As for a standard FFBS, the backward simulation is typically repeated M times, to generate a set of backward trajectories $\{u_{1:T}'^{,j}\}_{j=1}^M$ which can be used to approximate $p(u_{1:T} \mid y_{1:T})$. If we seek smoothed estimates of the linear states, these can be computed, e.g. by running a modified Bryson-Frazier [15] or a Rauch-Tung-Striebel (RTS) [16] smoother for each backward trajectory (alternatively, we can run conditional Kalman filters and fuse the filter estimates with the backward statistics). The total computational complexity of generating M backward trajectories, using N forward filter particles, is $O(NMT)$ (i.e. the same as for a standard FFBS). The computational cost can be reduced by using the rejection-sampling-based FFBS by [17] or the Metropolis-Hastings-based FFBS by [18].

4. PROOFS

In this section, we prove Propositions 1–3. We start with a useful lemma.

Lemma 1. *Let $\xi \sim \mathcal{N}(0, I)$ and let $z = c + Ax + \Gamma\xi$, for some constant vectors c and x and matrices A and Γ of appropriate dimensions. Let $\Omega \succeq 0$ and λ be a constant matrix and vector, respectively. Then $\mathrm{E}\left[\exp\left(-\frac{1}{2}\left(z^{\mathsf{T}}\Omega z - 2\lambda^{\mathsf{T}}z\right)\right)\right] = |M|^{-1/2}\exp\left(-\frac{1}{2}\gamma\right)$ with,*

$$\begin{aligned}\gamma &= \|Ax\|^2_{\Omega - \Omega\Gamma M^{-1}\Gamma^{\mathsf{T}}\Omega} - 2x^{\mathsf{T}}A^{\mathsf{T}}\left(I - \Omega\Gamma M^{-1}\Gamma^{\mathsf{T}}\right)m\\ &\quad + \|c\|^2_{\Omega} - 2\lambda^{\mathsf{T}}c - \|\Gamma^{\mathsf{T}}m\|^2_{M^{-1}},\end{aligned}$$

where $m = \lambda - \Omega c$ and $M = \Gamma^{\mathsf{T}}\Omega\Gamma + I$.

Proof. A detailed proof is omitted due to lack of space. The result follows by plugging in the expression for z and carrying out the integration w.r.t. ξ. □

Propositions 1 and 2 are given by induction. The initialization at time T in Proposition 2 follows directly from (1c),

$$\begin{aligned}p(y_T \mid z_T, u_T) &= \mathcal{N}(y_T; h_T + C_T z_T, R_T)\\ &\propto \exp\left(-\tfrac{1}{2}\left(\|C_T z_T\|^2_{R_T^{-1}} - 2z_T^{\mathsf{T}}C_T^{\mathsf{T}}R_T^{-1}(y_T - h_T)\right)\right). \quad (9)\end{aligned}$$

Hence, assume that Proposition 2 holds at time $t+1$. We have,

$$\begin{aligned}p(y_{t+1:T}, u_{t+1:T} \mid z_t, u_t) &= p(u_{t+1} \mid z_t, u_t)\\ \times \int p(y_{t+1:T}, u_{t+2:T} &\mid z_{t+1}, u_{t+1}) p(z_{t+1} \mid z_t, u_t)\, dz_{t+1}.\end{aligned} \quad (10)$$

The first factor is a Gaussian PDF, given by (1a),

$$\begin{aligned}p(u_{t+1} \mid z_t, u_t) &= \mathcal{N}(u_{t+1}; g_t + B_t z_t, Q_t)\\ &\propto |Q_t|^{-1/2}\exp\left(-\tfrac{1}{2}\left(\|u_{t+1} - g_t\|^2_{Q_t^{-1}}\right)\right)\\ &\times \exp\left(-\tfrac{1}{2}\left(\|B_t z_t\|^2_{Q_t^{-1}} - 2z_t^{\mathsf{T}}B_t^{\mathsf{T}}Q_t^{-1}(u_{t+1} - g_t)\right)\right).\end{aligned}$$

To compute the integral in (10), we use the induction hypothesis and (1b). We then apply Lemma 1 with $c = f_t$, $A = A_t$, $x = z_t$, $\Gamma = F_t$, $\Omega = \widehat{\Omega}_{t+1}$ and $\lambda = \widehat{\lambda}_{t+1}$. Proposition 1 then follows by collecting terms from the two factors.

Next, to prove Proposition 2 for $t < T$, we assume that Proposition 1 holds at time t. We have,

$$p(y_{t:T}, u_{t+1:T} \mid z_t, u_t) = p(y_t \mid z_t, u_t)p(y_{t+1:T}, u_{t+1:T} \mid z_t, u_t).$$

The first factor is given by (1c), analogously to (9), and the second factor is given by Proposition 1. The result follows by collecting terms from the two factors.

Finally, to prove Proposition 3 we note that the sought density is given by (8) where the two factors of the integrand are given by Proposition 1 and by (4), respectively. The result follows by applying Lemma 1 with $c = \bar{z}_{t|t}$, $x = 0$, $\Gamma = \Gamma_{t|t}$, $\Omega = \Omega_t$ and $\lambda = \lambda_t$.

5. NUMERICAL RESULTS

We evaluate the proposed RBPS by comparing its performance with alternative smoothers. The following methods are considered:

- FFBS: A non-Rao-Blackwellizedd FFBS [11].
- RB-F/S: A Rao-Blackwellized Kitagawa filter/smoother [13].
- RB-FF/JBS: Rao-Blackwellized forward filter/joint backward simulator [12].
- RB-FFBS: The proposed method (Algorithm 1).

For all methods, a bootstrap PF [19] or RBPF [3] is used in the forward direction.

The RB-F/S consists of running an RBPF and storing the nonlinear state trajectories. Smoothed linear state estimates are then computed by running constrained RTS smoothers, conditionally on these nonlinear trajectories. The RB-FF/JBS is the "joint backward simulator with constrained RTS smoothing" by [12] (see also [10]). In this method, we run an RBPF in the forward direction, but sample both the nonlinear and the linear states in the backward direction. The method relies on having access to the linear state samples in order to compute the backward sampling probabilities. However, once the backward simulation is complete, the linear parts of the trajectories are discarded. Refined linear state estimates are then computed by, again, running constrained RTS smoothers, one for each nonlinear backward trajectory.

We consider a 5th order mixed linear/nonlinear system. The nonlinear part is given by the time series,

$$u_{t+1} = 0.5u_t + \theta_t \frac{u_t}{1+u_t^2} + 8\cos(1.2t) + 0.071v_t^u, \tag{11a}$$

$$y_t = 0.05u_t^2 + e_t, \tag{11b}$$

for some process $\{\theta_t\}_{t\geq 1}$. The case with a static $\theta_t \equiv 25$ has been studied, among others, by [20, 19]. Here, we assume instead that θ_t is a time varying parameter with known dynamics, given by the output from a 4th order linear system,

$$z_{t+1} = \begin{pmatrix} 3 & -1.691 & 0.849 & -0.3201 \\ 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \end{pmatrix} z_t + 0.1v_t^z \tag{12a}$$

$$\theta_t = 25 + \begin{pmatrix} 0 & 0.04 & 0.044 & 0.008 \end{pmatrix} z_t, \tag{12b}$$

with poles in $0.8\pm0.1i$ and $0.7\pm0.05i$. Combined, (11) and (12) is a mixed linear/nonlinear system. The noises are assumed to be white, Gaussian and mutually independent; $v_t^u \sim \mathcal{N}(0,1)$, $v_t^z \sim \mathcal{N}(0,I)$ and $e_t \sim \mathcal{N}(0,0.1)$.

We generate 1000 batches of data from the system, each with $T = 100$ samples. We run the smoothers two times, first with $N = 300$ and then with $N = 30$ particles. The backward-simulation-based methods use $M = N/3$ backward trajectories, based on the recommendation to set $M \lesssim N$ [21]. Table 1 summarizes the results, in terms of the time averaged RMSE values for the nonlinear state u_t and for the time varying parameter θ_t (note that θ_t is a linear combination of the four linear states z_t).

Table 1. RMSE values averaged over 1 000 runs

	$N = 300$		$N = 30$	
Smoother	u_t	θ_t	u_t	θ_t
FFBS	0.499	0.782	1.203	1.238
RB-F/S	0.424	0.660	0.980	0.909
RB-FF/JBS	0.399	0.579	0.967	0.869
RB-FFBS	0.398	0.564	0.965	0.836

The proposed RB-FFBS gives the most accurate results among the considered smoothers, both for $N = 300$ and $N = 30$. The difference between RB-FFBS and RB-FF/JBS is quite small, but standard statistical hypothesis tests indeed indicate a clear statistical significance. In fact, these two methods are in many respects similar. They use similar forward and backward recursions and

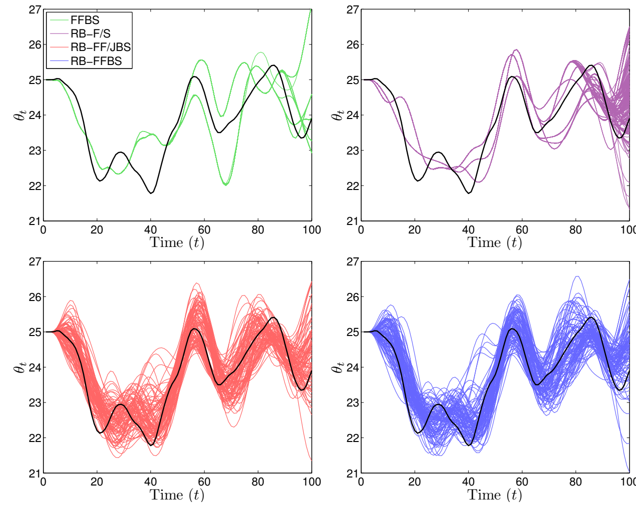

Fig. 1. Estimates of θ_t for $t = 1, \ldots, T$. From top left to bottom right; FFBS, RB-F/S, RB-FF/JBS and RB-FFBS. Each curve corresponds to one particle trajectory ($\theta'_{1:T}$ for FFBS and $\bar{\theta}'_{1:T|T}$ for the remaining smoothers). The true value is shown as a thick black line.

they both use conditional RTS smoothers to compute smoothed estimates of the linear states. Hence, in terms of implementation and computational complexity, they are almost identical. With this in mind, and from the fact that the results in Table 1 are in favor of RB-FFBS, we believe that the RB-FFBS indeed is the preferred method of choice, between these two smoothers. Furthermore, in the authors' opinion, RB-FFBS makes use of a more intuitively correct Rao-Blackwellization, since the marginalization is done both in the forward direction and in the backward direction.

For further comparison, Figure 1 shows the estimates of θ_t for one specific batch of data, using $N = 300$ and $M = 100$. This reveals a clear difference between the methods' abilities of accurately representing the posterior distribution of θ_t. For FFBS and RB-F/S (the top row), there is a clear degeneracy in the trajectories. For RB-F/S, this is expected, as it is a direct effect of the path degeneracy of the RBPF. For the (non-Rao-Blackwellized) FFBS, the degeneracy is caused by the fact that $N = 300$ particles is insufficient to represent the posterior in all five dimensions, resulting in that only a few particles get significantly non-zero weights. This will cause the backward simulator to degenerate, in the sense that many backward trajectories will coincide. The Rao-Blackwellized backward simulators (bottom row) perform much better in this respect, as there is a much larger diversity among the backward trajectories.

6. CONCLUSION

A new smoother for a class of mixed linear/nonlinear state-space models has been presented. The method is a forward filter/backward simulator which uses Rao-Blackwellization to exploit the conditionally linear Gaussian structure of the model. In contrast to previously developed algorithms for this model, the new smoother samples the linear state component in neither the forward nor the backward direction. Instead, a recursion has been derived which allows efficient calculation of backward sampling probabilities. Simulations have been used to demonstrate that the smoother functions well, with improvements in RMSE over previous algorithms.

7. REFERENCES

[1] A. Doucet and A. Johansen, “A tutorial on particle filtering and smoothing: Fifteen years later,” in *The Oxford Handbook of Nonlinear Filtering*, D. Crisan and B. Rozovsky, Eds. Oxford University Press, 2011.

[2] F. Gustafsson, “Particle filter theory and practice with positioning applications,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 25, no. 7, pp. 53–82, 2010.

[3] T. Schön, F. Gustafsson, and P.-J. Nordlund, “Marginalized particle filters for mixed linear/nonlinear state-space models,” *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2279–2289, July 2005.

[4] R. Chen and J. S. Liu, “Mixture Kalman filters,” *Journal of the Royal Statistical Society: Series B*, vol. 62, no. 3, pp. 493–508, 2000.

[5] A. Doucet, S. J. Godsill, and C. Andrieu, “On sequential Monte Carlo sampling methods for Bayesian filtering,” *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[6] F. Lindsten, T. B. Schön, and J. Olsson, “An explicit variance reduction expression for the Rao-Blackwellised particle filter,” in *Proceedings of the 18th IFAC World Congress*, Milan, Italy, Aug. 2011.

[7] N. Chopin, “Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference,” *The Annals of Statistics*, vol. 32, no. 6, pp. 2385–2411, 2004.

[8] S. Särkkä, P. Bunch, and S. Godsill, “A backward-simulation based Rao-Blackwellized particle smoother for conditionally linear Gaussian models,” in *Proceedings of the 16th IFAC Symposium on System Identification*, Brussels, Belgium, July 2012.

[9] N. Whiteley, C. Andrieu, and A. Doucet, “Efficient Bayesian inference for switching state-space models using discrete particle Markov chain Monte Carlo methods,” Tech. Rep., Bristol Statistics Research Report 10:04, 2010.

[10] W. Fong, S. J. Godsill, A. Doucet, and M. West, “Monte Carlo smoothing with application to audio signal enhancement,” *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 438–449, Feb. 2002.

[11] S. J. Godsill, A. Doucet, and M. West, “Monte Carlo smoothing for nonlinear time series,” *Journal of the American Statistical Association*, vol. 99, no. 465, pp. 156–168, Mar. 2004.

[12] F. Lindsten and T. B. Schön, “Rao-Blackwellised particle smoothers for mixed linear/nonlinear state-space models,” Tech. Rep. LiTH-ISY-R-3018, Department of Electrical Engineering, Linköping University, 2011.

[13] G. Kitagawa, “Monte Carlo filter and smoother for non-Gaussian nonlinear state space models,” *Journal of Computational and Graphical Statistics*, vol. 5, no. 1, pp. 1–25, 1996.

[14] R. Gerlach, C. Carter, and R. Kohn, “Efficient Bayesian inference for dynamic mixture models,” *Journal of the American Statistical Association*, vol. 95, no. 451, pp. 819–828, 2000.

[15] G. J. Bierman, “Fixed interval smoothing with discrete measurements,” *International Journal of Control*, vol. 18, no. 1, pp. 65–75, 1973.

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

[17] R. Douc, A. Garivier, E. Moulines, and J. Olsson, “Sequential Monte Carlo smoothing for general state space hidden Markov models,” *Annals of Applied Probability*, vol. 21, no. 6, pp. 2109–2145, 2011.

[18] P. Bunch and S. Godsill, “Improved particle approximations to the joint smoothing distribution using Markov chain Monte Carlo,” *IEEE Transactions on Signal Processing*, vol. 61, no. 4, pp. 956–963, 2013.

[19] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, “Novel approach to nonlinear/non-Gaussian Bayesian state estimation,” *Radar and Signal Processing, IEE Proceedings F*, vol. 140, no. 2, pp. 107 –113, Apr. 1993.

[20] M. L. Andrade Netto, L. Gimeno, and M. J. Mendes, “A new spline algorithm for non-linear filtering of discrete time systems,” in *Proceedings of the 7th Triennial World Congress*, Helsinki, Finland, 1979, pp. 2123–2130.

[21] F. Lindsten and T. B. Schön, “Backward simulation methods for Monte Carlo statistical inference,” *Foundations and Trends in Machine Learning (submitted)*, 2013.