On the nonlinear and nonnormal filter using rejection sampling

Tanizaki, Hisashi

(Citation)
IEEE Transactions on Automatic Control,44(2):314-319

(Issue Date)
1999-02

(Resource Type)
journal article

(Version)
Accepted Manuscript

(URL)
https://hdl.handle.net/20.500.14094/90000577

On the Nonlinear and Nonnormal Filter Using Rejection Sampling

Hisashi Tanizaki

***Abstract*—In this paper, a nonlinear and/or nonnormal filter is proposed using rejection sampling. Generating random draws of the state-vector directly from the filtering density, the filtering estimate is simply obtained as the arithmetic average of the random draws. In the proposed filter, the random draws are recursively generated at each time. The Monte Carlo experiments indicate that the proposed nonlinear and nonnormal filter shows a good performance.**

***Index Terms*—Filtering, nonlinear, nonnormal, proposal density, rejection sampling.**

I. Introduction

Nonlinear filters have been investigated for a long time (e.g., Alspach and Sorenson [1], Sorenson and Alspach [18], and Wishner *et al.* [23]) and we still have numerous density-based nonlinear filtering algorithms. Kitagawa [13] and Kramer and Sorenson [16] proposed the numerical integration procedure. Tanizaki [20] and Tanizaki and Mariano [22] utilized the Monte Carlo integration with importance sampling for nonlinear and nonnormal state-space models. Moreover, Carlin *et al.* [3] and Carter and Kohn [4], [5] proposed a solution to state-space modeling in a Bayesian framework using Gibbs sampling (see also Chib and Greenberg [6]). In addition, Gordon *et al.* [9], Kitagawa [14], and Kitagawa and Gersch [15] proposed the Monte Carlo filter and smoother with the bootstrap method.

An alternative procedure which we propose in this paper also utilizes the Monte Carlo technique. Given random draws of the state-variable at previous time, those at present time are directly generated without evaluating any density function, where the random draws from the filtering densities are recursively obtained. For random number generation, rejection sampling is adopted, which is a method of the random number generation from any distribution function.[1]

II. Overview of Nonlinear and Nonnormal Filter

We consider the following general nonlinear and nonnormal state-space model:

$$\text{Measurement Equation} \quad y_t = h_t(\alpha_t, \epsilon_t) \tag{1}$$

$$\text{Transition Equation} \quad \alpha_t = f_t(\alpha_{t-1}, \eta_t) \tag{2}$$

Manuscript received September 27, 1996; revised August 20, 1997. Recommended by Associate Editor, J. C. Spall.

The author is with the Faculty of Economics, Kobe University, Kobe 657-8501, Japan (e-mail: tanizaki@kobe-u.ac.jp).

Publisher Item Identifier S 0018-9286(99)01278-7.

[1] For rejection sampling, note as follows. When we generate a random draw from $p(x)$, called the target density, we take another distribution function $p_*(x)$, called the proposal density, which is appropriately chosen by a researcher. Define the acceptance probability as $\omega(x) = p(x)/ap_*(x)$, where the assumption $a \equiv \sup_x p(x)/p_*(x) < \infty$ is required. Rejection sampling is implemented as: 1) generate a random draw from $p_*(x)$ and 2) accept it with probability $\omega(x)$. The accepted random draw is taken as a random draw from $p(x)$. In the case where both $p(x)$ and $p_*(x)$ are normally distributed as $N(\mu, \sigma^2)$ and $N(\mu_*, \sigma_*^2)$, it is easily shown that we need $\sigma_*^2 > \sigma^2$ for the condition $a < \infty$, which implies that $p_*(x)$ has to be distributed with larger variance than $p(x)$. Using rejection sampling, we can generate a random draw from any distribution function under the condition that $a < \infty$ is satisfied. However, the disadvantages of rejection sampling are: 1) we need to compute a, which sometimes does not exist and 2) it takes a long time when $\omega(\cdot)$ is close to zero; see, for example, Boswell, *et al.* [2], Geweke [8], and O'Hagan [17] for rejection sampling.

for $t = 1, 2, \cdots, T$, where T denotes the sample size. Let y_t be the observed data and α_t be the state vector. $h_t(\cdot)$ and $f_t(\cdot)$ are vector functions. The error terms ϵ_t and η_t are mutually independently distributed. Define information set up to time s as $Y_s = \{y_s, y_{s-1}, \cdots, y_1\}$. Let $p_y(y_t|\alpha_t)$ and $p_\alpha(\alpha_t|\alpha_{t-1})$ be the density functions obtained from (1) and (2). Under the above setup, the density-based filtering algorithm is known as (see, for example, Kitagawa [13] and Harvey [11])

Prediction Equation

$$p(\alpha_t|Y_{t-1}) = \int p_\alpha(\alpha_t|\alpha_{t-1}) p(\alpha_{t-1}|Y_{t-1})\, \mathrm{d}\,\alpha_{t-1} \tag{3}$$

Updating Equation

$$p(\alpha_t|Y_t) = \frac{p_y(y_t|\alpha_t) p(\alpha_t|Y_{t-1})}{\int p_y(y_t|\alpha_t) p(\alpha_t|Y_{t-1})\, \mathrm{d}\,\alpha_t} \tag{4}$$

for $t = 1, 2, \cdots, T$. The initial condition is given by $p(\alpha_0|Y_0) = p(\alpha_0)$ if α_0 is stochastic and $p(\alpha_1|Y_0) = p_\alpha(\alpha_1|\alpha_0)$, otherwise where $p(\alpha_0)$ denotes the initial density of the state variable. Based on $p_\alpha(\alpha_t|\alpha_{t-1})$ and $p_y(y_t|\alpha_t)$, prediction (3) yields $p(\alpha_t|Y_{t-1})$ given $p(\alpha_{t-1}|Y_{t-1})$ and updating (4) gives us $p(\alpha_t|Y_t)$ from $p(\alpha_t|Y_{t-1})$. Thus, $p(\alpha_t|Y_t)$ can be recursively obtained for $t = 1, 2, \cdots, T$.

When the unknown parameters are in (1) and (2), the likelihood function to be maximized is $p(Y_T) = \prod_{t=1}^{T} \int p_y(y_t|\alpha_t) p(\alpha_t|Y_{t-1})\, \mathrm{d}\,\alpha_t$, which is obtained from the denominator of (4).

Kitagawa [13] and Kramer and Sorenson [16] proposed a nonlinear and nonnormal filter using numerical integration to evaluate each integration in (3) and (4). Tanizaki [20] and Tanizaki and Mariano [22] evaluated the integration using Monte Carlo integration with importance sampling, where a recursive algorithm of the density functions is converted to that of the weight functions, defined as $\omega(\alpha_t|Y_s) = p(\alpha_t|Y_s)/p_*(\alpha_t)$ for $s = t-1, t$. The density $p_*(\alpha_t)$, called the importance density, has to be appropriately specified by a researcher. The random draws $\alpha_{i,t}$, $i = 1, 2, \cdots, n$, are generated from $p_*(\alpha_t)$. A recursive filtering algorithm of $\omega(\alpha_{i,t}|Y_t)$ is derived for $t = 1, 2, \cdots, T$.

Carlin *et al.* [3], Carter and Kohn [4], [5], and Chib and Greenberg [6] introduced the nonlinear and nonnormal state-space models with Gibbs sampling, where we do not have to evaluate each integration in (3) and (4). They investigated the nonlinear state-space models in the Bayesian framework, where the nuisance parameters in the state-space model are assumed to be stochastic. The state-space models that they used are quite restricted because they studied the special state-space models such that it is easy to generate random draws from the underlying assumptions. Their attempt is to generate random draws from the joint density of $\alpha_1, \alpha_2, \cdots, \alpha_T$ given Y_T, which does not yield a recursive algorithm.

In this paper, we consider generating random draws from the filtering density $p(\alpha_t|Y_t)$ without evaluating any integration, where the random draws of α_t given Y_t are recursively obtained given those of α_{t-1} given Y_{t-1}.

III. Rejection Sampling Filter

Let $\alpha_{i,t|s}$ be the ith random draw from $p(\alpha_t|Y_s)$. When the random draws $\alpha_{i,t-1|t-1}$, $i = 1, 2, \cdots, n$, are available, we consider generating $\alpha_{i,t|t}$, $i = 1, 2, \cdots, n$. By substituting (3) into

TABLE I
SIMULATIONS I–III (δ KNOWN)

δ	n	$T=20$ KF	$T=20$ RF (A)	$T=20$ RF (B) $\gamma=4$	$T=20$ RF (B) $\gamma=9$	$T=20$ RF (B) $\gamma=16$	$T=40$ KF	$T=40$ RF (A)	$T=40$ RF (B) $\gamma=4$	$T=40$ RF (B) $\gamma=9$	$T=40$ RF (B) $\gamma=16$
Simulation I (Linear and Normal Model)											
0.5	200	0.7292	0.7317	0.7322	0.7328	0.7323	0.7334	0.7362	0.7365	0.7366	0.7366
	500		0.7309	0.7308	0.7306	0.7315		0.7351	0.7350	0.7352	0.7352
	1000		0.7307	0.7308	0.7308	0.7305		0.7347	0.7349	0.7349	0.7349
0.9	200	0.7760	0.7837	0.7850	0.7835	0.7840	0.7793	0.7880	0.7873	0.7874	0.7872
	500		0.7827	0.7836	0.7831	0.7834		0.7861	0.7862	0.7865	0.7866
	1000		0.7826	0.7825	0.7828	0.7828		0.7862	0.7859	0.7861	0.7861
1.0	200	0.7897	0.7989	0.7989	0.7995	0.7995	0.7928	0.8028	0.8029	0.8028	0.8029
	500		0.7986	0.7973	0.7983	0.7976		0.8011	0.8012	0.8012	0.8013
	1000		0.7977	0.7978	0.7976	0.7975		0.8012	0.8010	0.8008	0.8008
Simulation II (Stochastic Volatility Model)											
0.5	200	1.1487	0.9273	0.9270	0.9271	0.9262	1.1577	0.9391	0.9396	0.9389	0.9392
	500		0.9257	0.9245	0.9246	0.9258		0.9378	0.9378	0.9370	0.9377
	1000		0.9247	0.9249	0.9245	0.9250		0.9368	0.9368	0.9371	0.9369
0.9	200	2.0909	1.2209	1.2219	1.2224	1.2184	2.2135	1.2532	1.2548	1.2552	1.2581
	500		1.2162	1.2160	1.2190	1.2181		1.2499	1.2499	1.2506	1.2502
	1000		1.2152	1.2164	1.2164	1.2161		1.2490	1.2495	1.2493	1.2498
Simulation III (ARCH(1) Model)											
0.5	200	0.7016	0.6909	0.6921	0.6931	0.6926	0.7039	0.6940	0.6963	0.6956	0.6958
	500		0.6903	0.6912	0.6911	0.6915		0.6941	0.6950	0.6950	0.6951
	1000		0.6902	0.6910	0.6911	0.6909		0.6938	0.6946	0.6948	0.6948
0.9	200	0.6748	0.5666	—	—	0.5883	0.6503	0.5585	—	—	0.5830
	500		0.5679	—	—	0.5892		0.5589	—	—	0.5818
	1000		0.5697	—	—	0.5889		0.5601	—	—	0.5806

(4), $p(\alpha_t|Y_t)$ is approximated as

$$\begin{aligned} p(\alpha_t|Y_t) &= \frac{1}{c_t}\int p_y(y_t|\alpha_t)p_\alpha(\alpha_t|\alpha_{t-1})p(\alpha_{t-1}|Y_{t-1})\,\mathrm{d}\,\alpha_{t-1} \\ &\approx \sum_{i=1}^{n}\frac{c_{i,t}}{c_t}\frac{1}{n}\left(\frac{p_y(y_t|\alpha_t)p_\alpha(\alpha_t|\alpha_{i,t-1|t-1})}{c_{i,t}}\right) \\ &\approx \sum_{i=1}^{n}\frac{\hat{c}_{i,t}}{\hat{c}_t}\frac{1}{n}\left(\frac{p_y(y_t|\alpha_t)p_\alpha(\alpha_t|\alpha_{i,t-1|t-1})}{c_{i,t}}\right) \\ &\equiv \sum_{i=1}^{n} q_{i,t}\left(\frac{p_y(y_t|\alpha_t)p_\alpha(\alpha_t|\alpha_{i,t-1|t-1})}{c_{i,t}}\right). \end{aligned} \tag{5}$$

$q_{i,t}$ is defined as $q_{i,t} \equiv \hat{c}_{i,t}/n\hat{c}_t$, and c_t and $\hat{c}_t$ are given by

$$\begin{aligned} c_t &\equiv \iint p_y(y_t|\alpha_t)p_\alpha(\alpha_t|\alpha_{t-1})p(\alpha_{t-1}|Y_{t-1})\,\mathrm{d}\,\alpha_{t-1}\,\mathrm{d}\,\alpha_t \\ &\approx \frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n} p_y(y_t|\alpha_{ji,t|t-1}) \equiv \hat{c}_t \end{aligned}$$

where $\alpha_{ji,t|t-1}$ is obtained from $\alpha_{ji,t|t-1} = f_t(\alpha_{i,t-1|t-1}, \eta_{j,t})$. Moreover, $c_{i,t}$ and $\hat{c}_{i,t}$ are represented as

$$\begin{aligned} c_{i,t} &\equiv \int p_y(y_t|\alpha_t)p_\alpha(\alpha_t|\alpha_{i,t-1|t-1})\,d\alpha_t \\ &\approx \frac{1}{n}\sum_{j=1}^{n} p_y(y_t|\alpha_{ji,t|t-1}) \equiv \hat{c}_{i,t}. \end{aligned}$$

Thus, from (5), $p(\alpha_t|Y_t)$ is approximated as a mixture of n distributions with probability $q_{i,t}$, $i = 1, 2, \cdots, n$. Let $p_*(z)$ be the proposal density. The acceptance probability $\omega(z)$ is defined as

$$\omega(z) = \frac{p_y(y_t|z)p_\alpha(z|\alpha_{i,t-1|t-1})/p_*(z)}{\sup_z\, p_y(y_t|z)p_\alpha(z|\alpha_{i,t-1|t-1})/p_*(z)}.$$

The estimation procedure is as follows: 1) pick up $\alpha_{i,t-1|t-1}$ for i with probability $q_{i,t}$; 2) generate a random draw z from $p_*(\cdot)$ and a uniform random draw u from the interval between zero and one; 3) take z as $\alpha_{j,t|t}$ if $u \le \omega(z)$ and go back to 2) otherwise; 4)

TABLE II
SIMULATION STUDIES I AND II (δ UNKNOWN)

			$T = 20$					$T = 40$				
δ		n	AVE	RMSE	25%	50%	75%	AVE	RMSE	25%	50%	75%
Simulation I (Linear and Normal Model)												
	KF		0.359	0.323	0.17	0.42	0.61	0.428	0.228	0.30	0.47	0.61
0.5	RF (A)	200	0.366	0.326	0.18	0.43	0.62	0.437	0.229	0.31	0.48	0.62
	RF (A)	500	0.367	0.327	0.18	0.44	0.62	0.441	0.229	0.31	0.48	0.62
	RF (A)	1000	0.369	0.330	0.20	0.44	0.63	0.445	0.229	0.33	0.49	0.62
	KF		0.767	0.236	0.67	0.84	0.93	0.845	0.121	0.80	0.88	0.93
0.9	RF (A)	200	0.773	0.236	0.68	0.84	0.94	0.852	0.120	0.81	0.89	0.93
	RF (A)	500	0.775	0.235	0.68	0.85	0.94	0.852	0.120	0.81	0.89	0.93
	RF (A)	1000	0.777	0.234	0.69	0.85	0.94	0.854	0.119	0.81	0.89	0.93
	KF		0.885	0.201	0.83	0.95	1.01	0.951	0.092	0.93	0.98	1.01
1.0	RF (A)	200	0.890	0.201	0.83	0.96	1.02	0.956	0.092	0.93	0.98	1.01
	RF (A)	500	0.892	0.199	0.84	0.96	1.02	0.956	0.091	0.93	0.98	1.01
	RF (A)	1000	0.892	0.198	0.84	0.96	1.02	0.957	0.090	0.93	0.98	1.01
Simulation II (Stochastic Volatility Model)												
	KF		0.509	0.432	0.00	0.63	0.96	0.538	0.434	0.00	0.68	0.98
0.5	RF (A)	200	0.516	0.425	0.02	0.63	0.96	0.543	0.427	0.02	0.68	0.97
	RF (A)	500	0.516	0.425	0.02	0.63	0.95	0.543	0.426	0.02	0.68	0.97
	RF (A)	1000	0.516	0.425	0.02	0.63	0.96	0.543	0.427	0.02	0.68	0.97
	KF		0.785	0.354	0.77	0.99	0.99	0.860	0.293	0.94	0.99	0.99
0.9	RF (A)	200	0.785	0.346	0.79	0.97	0.98	0.857	0.285	0.94	0.98	0.99
	RF (A)	500	0.785	0.346	0.78	0.97	0.98	0.857	0.286	0.94	0.98	0.99
	RF (A)	1000	0.785	0.346	0.79	0.97	0.99	0.857	0.286	0.94	0.98	0.99

repeat 1)–3) n times for $j = 1, 2, \cdots, n$; and 5) repeat 1)–4) T times for $t = 1, 2, \cdots, T$. Note that rejection sampling is utilized in procedures 2) and 3). For a function $g(\cdot)$, the expectation $\mathrm{E}(g(\alpha_t)|Y_s)$ is approximated as $(1/n)\sum_{i=1}^{n} g(\alpha_{i,t|s})$.[2]

The proposed nonlinear and nonnormal filter gives us a general solution to any nonlinear and nonnormal state-space model.

For choice of the proposal density,[3] two candidates are taken, i.e., (A) one is $p_*(\alpha_t) = p_\alpha(\alpha_t|\alpha_{i,t-1|t-1})$ and (B) another is $p_*(\alpha_t) = N(\alpha^*_{t|t}, \gamma\Sigma^*_{t|t})$, where γ is a constant and $\alpha^*_{t|t}$ and $\Sigma^*_{t|t}$ denote the first- and second-moments obtained from the extended Kalman filter (KF).[4] For the proposal density (B), $\gamma = 4, 9, 16$ is taken in Section IV. Note that γ should be greater than one because the proposal density should have larger variance than the target density.

When the unknown parameter is included in (1) and (2), we maximize the likelihood function represented by $p(Y_T) = \prod_{t=1}^{T} c_t \approx \prod_{t=1}^{T} \hat{c}_t$, which comes from the definition of c_t.

IV. NUMERICAL EXAMPLES: MONTE CARLO EXPERIMENTS

We compare the extended KF[5] and the proposed nonlinear filter, which is called the rejection sampling filter (RF) in this paper. The simulation procedure is as follows: 1) generating random numbers of ϵ_t and η_t for $t = 1, 2, \cdots, T$, we obtain a set of data y_t and α_t, $t = 1, 2, \cdots, T$, from (1) and (2), where $T = 20, 40$ is taken;

[2] Consider an example of $g(\alpha_t) = \alpha_t$. Let the estimate of $\alpha_{t|t} \equiv \mathrm{E}(\alpha_t|Y_t)$ be $\hat{\alpha}_{t|t}$, which is the filtering estimate computed by $\hat{\alpha}_{t|t} = (1/n)\sum_{i=1}^{n} \alpha_{i,t|t}$. Recall that $\alpha_{i,t|t}$ is a random draw generated from the filtering density $p(\alpha_t|Y_t)$. Let us assume that variance of $\alpha_{i,t|t}$ is given by $\Sigma_{t|t}$. Then, by the central limit theorem, $\sqrt{n}(\hat{\alpha}_{t|t} - \alpha_{t|t})$ is asymptotically normal distribution with mean zero and variance $\Sigma_{t|t}$. Moreover, denote the estimate of $\Sigma_{t|t}$ by $\hat{\Sigma}_{t|t}$, which is defined as $\hat{\Sigma}_{t|t} = (1/n)\sum_{i=1}^{n}(\alpha_{i,t|t} - \hat{\alpha}_{t|t})'(\alpha_{i,t|t} - \hat{\alpha}_{t|t})$. It is known that $\hat{\Sigma}_{t|t}$ is a consistent estimate of $\Sigma_{t|t}$. Accordingly, $\hat{\alpha}_{t|t}$ is normally distributed with mean $\alpha_{t|t}$ and variance $(1/n)\hat{\Sigma}_{t|t}$ for sufficiently large n. Thus, the confidence bound of the filtering estimate $\hat{\alpha}_{t|t}$ can be easily obtained. There is a great amount of literature on the confidence bounds (see, for example, Spall [19]).

[3] As in Footnote 1, the proposal density has to be more broadly distributed than the target density, but the former is not too different from the latter.

[4] The extended KF is one of the traditional nonlinear filters, where the nonlinear measurement and transition equations are linearized by the first-order Taylor series expansion and the linearized system is directly applied to the standard linear recursive KF algorithm; see, for example, Tanizaki [20].

[5] In the case where the system is linear and normal, the extended KF reduces to the conventional KF. Accordingly, in Simulation I, KF implies the standard Kalman filter.

TABLE III
SIMULATION IV: MARKOV SWITCHING MODEL ($k = 2$)

$p_{11} \backslash p_{22}$		n	$T=20$ 0.2	$T=20$ 0.5	$T=20$ 0.8	$T=40$ 0.2	$T=40$ 0.5	$T=40$ 0.8
	KF		0.4383	0.4354	0.3710	0.4372	0.4359	0.3733
0.2	RF (A)	200	0.4326	0.4351	0.3683	0.4316	0.4352	0.3706
	RF (A)	500	0.4318	0.4337	0.3673	0.4306	0.4344	0.3700
	RF (A)	1000	0.4312	0.4337	0.3672	0.4301	0.4341	0.3700
	KF		0.4349	0.4474	0.4080	0.4343	0.4471	0.4103
0.5	RF (A)	200	0.4341	0.4473	0.4055	0.4338	0.4473	0.4078
	RF (A)	500	0.4330	0.4466	0.4051	0.4327	0.4464	0.4071
	RF (A)	1000	0.4328	0.4465	0.4047	0.4325	0.4462	0.4071
	KF		0.3713	0.4098	0.4323	0.3734	0.4128	0.4365
0.8	RF (A)	200	0.3688	0.4079	0.4276	0.3708	0.4113	0.4298
	RF (A)	500	0.3683	0.4073	0.4259	0.3705	0.4107	0.4297
	RF (A)	1000	0.3680	0.4072	0.4255	0.3702	0.4103	0.4291

2) given Y_T, perform KF and RF, where $n = 200, 500, 1000$ is taken for RF; and 3) repeat 1) and 2) G times and compare the root mean square error (RMSE) for each estimator, which is defined as $\text{RMSE} = (1/T)\sum_{t=1}^{T} \text{MSE}_{t|t}^{1/2}$, where the mean square error (MSE) is defined as $\text{MSE}_{t|t} \equiv (1/G)\sum_{g=1}^{G}(\hat{\alpha}_{t|t}^{(g)} - \alpha_t^{(g)})^2$ and $\hat{\alpha}_{t|t}$ takes the state-variable estimated by KF or RF while α_t denotes the artificially simulated state-variable.[6] Note that the superscript (g) denotes the gth simulation run, where $G = 1000$ is taken.

In Tables I, III, and IV, assuming that δ is known (δ denotes the parameter of the model used in Simulations I–III), we compare the state estimates of KF and RF. In Table II, a comparison between the true parameter and the estimate of δ is shown for each procedure. Note that δ indicates the true value. AVE and RMSE represent the arithmetic average and the RMSE. Using 25, 50, and 75% give us the 0.25, 0.50, and 0.75 quantiles of the G estimates of δ. The maximization of the likelihood function is performed by a simple grid search, in which the function is maximized by changing the parameter value of δ by 0.01. The univariate cases are examined in Simulations I–III (Tables I and II) and the multivariate cases are in Simulation IV (Tables III and IV).

A. Simulation I—Linear and Normal Model

Consider the scalar system $y_t = \alpha_t + \epsilon_t$ and $\alpha_t = \delta\alpha_{t-1} + \eta_t$, where $\delta = 0.5, 0.9, 1.0$ is taken. The initial value α_0 and the error terms ϵ_t and η_t, $t = 1, 2, \cdots, T$, are assumed to be distributed as follows: $\alpha_0 \sim N(0, 1)$ and $(\epsilon_t, \eta_t)' \sim N(0, I_2)$, where I_2 denotes a 2×2 identity matrix. Exactly the same initial conditions are utilized in Simulations II and III.

B. Simulation II—Stochastic Volatility Model

Suppose that the system is represented as $y_t = \exp(0.5\alpha_t)\epsilon_t$ and $\alpha_t = \delta\alpha_{t-1} + \eta_t$, which is called the stochastic volatility model (see Harvey and Streibel [12]). We take $\delta = 0.5, 0.9$.

[6] Note that $\text{MSE}_{t|t}$ goes to $\Sigma_{t|t}$ in Footnote 2, as number of random draws (i.e., n) is large.

C. Simulation III—ARCH Model

The system is given by $y_t = \alpha_t + \epsilon_t$ and $\alpha_t = (1-\delta+\delta\alpha_{t-1}^2)^{1/2}\eta_t$ for $0 \le \delta < 1$, where $\delta = 0.5, 0.9$ is taken.[7] Note that "–" in Table I implies that the denominator of $\omega(z)$ does not exist.

D. Simulation IV—Markov Switching Model

Consider the higher dimensional state-variable, i.e., $\alpha_t = (\alpha_{1t}, \alpha_{2t}, \cdots, \alpha_{kt})'$, where we assume that one of the k elements of α_t is one and the others are zero. The model is specified as $y_t = x_t\alpha_t + \epsilon_t$ and $\alpha_t = P\alpha_{t-1} + \eta_t$, where $x_t = (x_{1t}, x_{2t}, \cdots, x_{kt})$ denotes a vector of exogenous variables and $P = (P_1', P_2', \cdots, P_k')'$ represents the transition probability matrix, where $P_i = (p_{i1}, p_{i2}, \cdots, p_{ik})$. Note that $\sum_{i=1}^{k} p_{ij} = 1$ should be satisfied for all j. ϵ_t is assumed be a standard normal distribution while η_t is distributed as a k-dimensional discrete random variable. The conditional density of α_t given α_{t-1} is represented by $p_\alpha(\alpha_t|\alpha_{t-1}) = \prod_{i=1}^{k}(P_i\alpha_{t-1})^{\alpha_{it}}$, which implies that the probability with which event i occurs at time t is $P_i\alpha_{t-1}$. The initial density is assumed to be $p(\alpha_0) = \prod_{i=1}^{k}(1/k)^{\alpha_{it}}$. This model is called the Markov switching model (see Hamilton [10]). $x_{it} = i$ for $i = 1, 2, \cdots, k$ is assumed in this simulation study.

We consider the two cases, i.e., $k = 2, 3$. The proposal density (A) is examined in both cases. For $k = 2$, p_{11} and p_{22} are the parameters, where $p_{11}, p_{22} = 0.2, 0.5, 0.8$ are taken (see Table III).[8] For $k = 3$, to reduce number of parameters, we assume that P is symmetric, i.e., $p_{11} = 1 - p_{21} - p_{31}$, $p_{12} = p_{21}$, $p_{22} = 1 - p_{21} - p_{32}$, $p_{13} = p_{31}$, $p_{23} = p_{32}$, and $p_{33} = 1 - p_{31} - p_{32}$, where p_{21}, p_{31}, and p_{32} are the parameters (see Table IV).[9]

[7] The transition equation follows the ARCH(1) model (see Engle [7], Harvey [11], and Harvey and Streibel [12]). In this Monte Carlo simulation study the unconditional variance of α_t is assumed to be one.

[8] In the case of $k = 2$, note that $\mathrm{E}(\alpha_{1t}|Y_t) = 1 - \mathrm{E}(\alpha_{2t}|Y_t)$ and $\mathrm{Var}(\alpha_{1t}|Y_t) = \mathrm{Var}(\alpha_{2t}|Y_t)$ because of $\sum_{i=1}^{k} \alpha_{it} = 1$. Therefore, RMSE of α_{1t} is equal to that of α_{2t}. In Table III, RMSE's of α_{1t} are reported.

[9] In this case, note that the conditions $0 \le p_{21}+p_{31} \le 1$, $0 \le p_{21}+p_{32} \le 1$, and $0 \le p_{31} + p_{32} \le 1$ have to be satisfied.

TABLE IV
SIMULATION IV: MARKOV SWITCHING MODEL ($k = 3$)

					$T = 20$			$T = 40$		
p_{21}	p_{31}	p_{32}		n	α_{1t}	α_{2t}	α_{3t}	α_{1t}	α_{2t}	α_{3t}
			KF		0.3845	0.4659	0.3925	0.3850	0.4652	0.3953
0.1	0.2	0.4	RF (A)	200	0.3707	0.4499	0.3864	0.3687	0.4490	0.3897
			RF (A)	500	0.3699	0.4491	0.3860	0.3679	0.4481	0.3892
			RF (A)	1000	0.3697	0.4488	0.3858	0.3677	0.4478	0.3890
			KF		0.3860	0.4523	0.3905	0.3882	0.4524	0.3902
0.1	0.3	0.6	RF (A)	200	0.3777	0.4407	0.3822	0.3787	0.4403	0.3814
			RF (A)	500	0.3769	0.4402	0.3814	0.3786	0.4399	0.3811
			RF (A)	1000	0.3769	0.4395	0.3811	0.3781	0.4394	0.3807
			KF		0.3893	0.4596	0.3917	0.3918	0.4592	0.3906
0.2	0.3	0.6	RF (A)	200	0.3830	0.4464	0.3828	0.3853	0.4470	0.3819
			RF (A)	500	0.3820	0.4458	0.3820	0.3845	0.4459	0.3812
			RF (A)	1000	0.3818	0.4456	0.3819	0.3846	0.4461	0.3811
			KF		0.3934	0.4682	0.3902	0.3961	0.4675	0.3887
0.3	0.4	0.5	RF (A)	200	0.3875	0.4559	0.3834	0.3906	0.4562	0.3816
			RF (A)	500	0.3870	0.4550	0.3830	0.3904	0.4560	0.3814
			RF (A)	1000	0.3865	0.4549	0.3829	0.3901	0.4554	0.3809

E. Results

In Simulation I, it is expected that KF is better than any other estimator, because RF is the simulation-based estimator, which includes the simulation errors.[10] In Table I, for all the simulation studies, $n = 200, 500, 1000$ are very close to each other. We sometimes have the case where RF (B) is not feasible (see the case $\delta = 0.9$ in Simulation III). In such a case, RF (A) is recommended.

We compare KF and RF (A) in Tables II–IV. In Table II, except for the case $\delta = 0.5$, RMSE's of RF are smaller than those of KF. For almost all the cases, AVE goes to the true parameter value as n is large. Similarly, in Tables III and IV, RMSE is small as n is large. In any case, we can conclude that 1) RF performs better than KF in nonlinear cases (RF is very close to KF even in the linear case) and 2) $n = 500$ is not too different from $n = 1000$.[11]

V. SUMMARY

A nonlinear and nonnormal filtering algorithm is proposed in this paper. Given random draws of the state-vector which are directly generated from $p(\alpha_t|Y_t)$, the filtering estimate is recursively obtained. In the proposed filter, we do not evaluate any integration. The proposed filter can be applied to any nonlinear and non-Gaussian state-space model. To implement rejection sampling, the proposal density is utilized to generate random draws from the target density. Two types of candidates are examined for the proposal density, where we obtain the result that RF (A) is recommended in practice, rather than RF (B).

Thus, in this paper, a Monte Carlo procedure of filtering algorithm using the simulation technique is proposed, where we utilize the random draws only. The procedure improves over the other nonlinear filters developed in the past from simplicity of computer programming.

Finally, note as follows. In the proposed nonlinear and nonnormal filter, when the acceptance probability $\omega(z)$ is very small, we need a huge sampling, and accordingly it takes a long time computationally. For example, when we have outlines or structural changes in the system [i.e., when $p(\alpha_t|Y_t)$ is away from $p(\alpha_{t-1}|Y_{t-1})$], the computational burden increases in the rejection sampling procedure.

[10] In Table I, KF does not depend on n.

[11] The result 2) implies that $n = 500, 1000$ is enough large when we want to have the point estimate of α_t. However, note that $n = 500, 1000$ might be too small in the case where we want to obtain a functional form of $p(\alpha_t|Y_t)$ by the random numbers. Since RMSE is taken as a measure of precision of the state estimates, we can conclude in this paper that $n = 500, 1000$ is large enough.

ACKNOWLEDGMENT

This paper is a substantial extension of Tanizaki [21, pp. 94–99]. The author would like to acknowledge the Associate Editor J. C. Spall and three anonymous referees for careful reviews and valuable comments.

REFERENCES

[1] D. L. Alspach and H. W. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 439–448, 1972.

[2] M. T. Boswell, S. D. Gore, G. P. Patil, and C. Taillie, "The art of computer generation of random variables," in *Handbook of Statistics*, vol. 9, C. R. Rao, Ed. Amsterdam, The Netherlands: North-Holland, 1993, pp. 661–721.

[3] B. P. Carlin, N. G. Polson, and D. S. Stoffer, "A Monte Carlo approach to nonnormal and nonlinear state space modeling," *J. American Statistical Assn.*, vol. 87, pp. 493–500, 1992.

[4] C. K. Carter and R. Kohn, "On Gibbs sampling for state space models," *Biometrika*, no. 81, pp. 541–553, 1994.

[5] ——, "Markov chain Monte Carlo in conditionally Gaussian state space models," *Biometrika*, no. 83, pp. 589–601, 1996.

[6] S. Chib and E. Greenberg, "Markov chain Monte Carlo simulation methods in econometrics," *Econometric Theory*, vol. 12, pp. 409–431, 1996.
[7] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of variance of U.K. inflation," *Econometrica*, vol. 50, pp. 987–1008, 1982.
[8] J. Geweke, "Monte Carlo simulation and numerical integration," in *Handbook of Computational Economics*, vol. 1, H. M. Amman, D. A. Kendrick, and J. Rust, Eds. Amsterdam, The Netherlands: North-Holland, 1996, pp. 731–800.
[9] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," in *Proc. Inst. Elec. Eng.*, vol. 140, pp. 107–113, 1993.
[10] J. D. Hamilton, *Time Series Analysis.* Princeton, NJ: Princeton Univ. Press, 1994.
[11] A. C. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge, U.K.: Cambridge Univ. Press, 1989.
[12] A. C. Harvey and M. Streibel, "Testing for a slowly changing level with special preference to stochastic volatility," London School of Economics and Political Science, STICERD-Econometrics Discussion Paper, no. EM/96/306, 1996.
[13] G. Kitagawa, "Non-Gaussian state-space modeling of nonstationary time series," *J. American Statistical Assn.*, vol. 82, pp. 1032–1063, 1987.
[14] ____, "Monte Carlo filter and smoother for non-Gaussian nonlinear state-space models," *J. Computational and Graphical Statistics*, vol. 5, pp. 1–25, 1996.
[15] G. Kitagawa and W. Gersch, *Smoothness Priors Analysis of Time Series*, Lecture Notes in Statistics, no. 116. New York: Springer-Verlag, 1996.
[16] S. C. Kramer and H. W. Sorenson, "Recursive Bayesian estimation using piece-wise constant approximations," *Automatica*, vol. 24, pp. 789–801, 1988.
[17] A. O'Hagan, *Kendall's Advanced Theory of Statistics*, vol. 2B. Edward Arnold, 1994.
[18] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, pp. 465–479, 1971.
[19] J. C. Spall, "The Kantorovich inequality for error analysis of the Kalman filter with unknown noise distributions," *Automatica*, vol. 31, pp. 1513–1517, 1995.
[20] H. Tanizaki, *Nonlinear Filters: Estimation and Applications*, Lecture Notes in Economics and Mathematical Systems, no. 400. New York: Springer-Verlag, 1993.
[21] H. Tanizaki, *Nonlinear Filters: Estimation and Applications*, 2nd ed. New York: Springer-Verlag, 1996.
[22] H. Tanizaki and R. S. Mariano, "Prediction, filtering and smoothing in nonlinear and nonnormal cases using Monte Carlo integration," *J. Appl. Econometrics*, vol. 9, pp. 163–179, 1994.
[23] R. P. Wishner, J. A. Tabaczynski, and M. Athans, "A comparison of three nonlinear filters," *Automatica*, vol. 5, pp. 487–496, 1969.

Steady-State Performance Analysis of Serial Transfer Lines Subject to Machine and Buffer Failure

Robert Lipset, Robert P. Van Til, and Sankar Sengupta

***Abstract*— In modern automated production lines, it is common to connect adjacent machines with buffers. Since these buffers are mechanical devices, they are prone to failure. Previous research concerning the steady-state analytical modeling of serial transfer lines assumed that buffers are completely reliable. This paper considers the unreliable buffer and presents a model of the serial transfer line incorporating this concept. A decomposition technique is developed for the general serial transfer line with unreliable buffers, and an algorithm for computing the solution of the model is presented.**

***Index Terms*—Discrete-event systems, manufacturing systems, performance modeling, unreliable buffer.**

I. Introduction

It is common to connect adjacent machines in serial transfer lines with accumulating conveyors to provide storage and buffering between processing stations. Since conveyors are mechanical devices, they are prone to failure. In the experience of the authors, the failure of accumulating conveyors accounts for a significant portion of lost production in automobile assembly plants.

This paper develops a steady-state analytical model of a serial transfer line that accounts for the failure of both machines and buffers. Previous research concerning the steady-state analytical modeling of serial transfer lines containing more than two machines assumes that buffers are completely reliable [1], [3]–[18], [22]–[26]. An analytical model for a two-machine line where both the machines and the buffer are subject to failure is presented in [20] and [21]. Buffer unreliability has also been discussed in [2].

Zimmern [26] is credited by several authors with first proposing a *continuous* model of the serial transfer line. The results of [26] assumed that machines are unreliable and characterized machine failures as operation-, not time-, dependent. A continuous model in which machine failures are time-dependent was formulated in [22]. *Discrete* models of the serial transfer line characterize the system as a Markovian birth–death process mapping transition probabilities into transition rates [4]. A discrete model that compared various buffering strategies was proposed in [18], but the model assumed that both machines and buffers were reliable. A discrete model that considered machine unreliability was presented in [3].

The simplest serial transfer line is composed of two machines with a single intermediate reliable buffer of finite capacity. Solutions for the continuous model of a two-machine line with a reliable buffer are given in [10] and [15]. In [14], a discrete approach depending on a Markov process model is used to produce closed form expressions for the performance measures.

An exact solution using a discrete model of a serial transfer line with three machines and two intermediate buffers is derived in [16].

Manuscript received February 21, 1997. Recommended by Associate Editor, W.-B. Gong.

R. Lipset is with the Industrial and Manufacturing Systems Engineering Department, Ohio University, Athens, OH 45701 USA (e-mail: lipset@ohio.edu).

R. P. Van Til and S. Sengupta are with the Electrical and Systems Engineering Department, Oakland University, Rochester, MI 48309 USA.

Publisher Item Identifier S 0018-9286(99)01279-9.