Title
A regularized auxiliary particle filtering approach for system state estimation and battery life prediction

Journal
Smart Materials and Structures, 20(7)

ISSN
0964-1726

Authors
Liu, Jie
Wang, Wilson
Ma, Fai

Publication Date
2011-07-01

DOI
10.1088/0964-1726/20/7/075021

Peer reviewed

Smart Mater. Struct. **20** (2011) 075021 (9pp) doi:10.1088/0964-1726/20/7/075021

A regularized auxiliary particle filtering approach for system state estimation and battery life prediction

Jie Liu[1,4], **Wilson Wang**[2] **and Fai Ma**[3]

[1] Department of Mechanical and Aerospace Engineering, Carleton University, Ottawa, ON, K1S 5B6, Canada
[2] Department of Mechanical Engineering, Lakehead University, Thunder Bay, ON, P7B 5E1, Canada
[3] Department of Mechanical Engineering, University of California, Berkeley, CA 94720, USA

E-mail: jliu@mae.carleton.ca

Received 17 January 2011, in final form 27 May 2011
Published 22 June 2011
Online at stacks.iop.org/SMS/20/075021

Abstract
System current state estimation (or condition monitoring) and future state prediction (or failure prognostics) constitute the core elements of condition-based maintenance programs. For complex systems whose internal state variables are either inaccessible to sensors or hard to measure under normal operational conditions, inference has to be made from indirect measurements using approaches such as Bayesian learning. In recent years, the auxiliary particle filter (APF) has gained popularity in Bayesian state estimation; the APF technique, however, has some potential limitations in real-world applications. For example, the diversity of the particles may deteriorate when the process noise is small, and the variance of the importance weights could become extremely large when the likelihood varies dramatically over the prior. To tackle these problems, a regularized auxiliary particle filter (RAPF) is developed in this paper for system state estimation and forecasting. This RAPF aims to improve the performance of the APF through two innovative steps: (1) regularize the approximating empirical density and redraw samples from a continuous distribution so as to diversify the particles; and (2) smooth out the rather diffused proposals by a rejection/resampling approach so as to improve the robustness of particle filtering. The effectiveness of the proposed RAPF technique is evaluated through simulations of a nonlinear/non-Gaussian benchmark model for state estimation. It is also implemented for a real application in the remaining useful life (RUL) prediction of lithium-ion batteries.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Condition-based maintenance is a program that recommends maintenance decisions based on the information collected through system condition monitoring (or system state estimation) and equipment failure prognostics (or system state forecasting) [1–5]. For complex systems whose internal state variables are either inaccessible to sensors or hard to measure under normal operational conditions, inference has to be made from indirect measurements using approaches such as Bayesian learning [6]. The Bayesian learning based system state estimation is to calculate the complete density function of the current system state based on a sequence of noisy measurements, whereas the Bayesian learning based system state forecasting is to compute the complete density function of the future system state based on the prediction model identified in parallel with state estimation [7, 8].

The formal recursion for the density function can be achieved for linear/Gaussian systems by applying the commonly used Kalman filtering [9]; however, there is usually no closed-form solution for nonlinear/non-Gaussian

[4] Author to whom any correspondence should be addressed.

systems [10]. To numerically handle the estimation problem in nonlinear/non-Gaussian systems, particle filters were introduced [11]. The particle filtering is a process to apply the sequential Monte Carlo method in numerical simulations and to represent the posterior distribution of the state variables by a number of weighted particles that are evolved recursively as new observations (or measurements) become available [12, 13]. In the 1990s, there was a surge of research and development (R&D) in particle filtering [11, 14]. Among these developments, two fundamental approaches were primarily used: the standard sampling and importance resampling particle filter (SIR-PF) [15] and the auxiliary particle filtering (APF) [13, 16]. In the standard SIR-PF, the prior distribution is usually chosen as an importance function to simplify computation. Unfortunately, such a selection may not be efficient in many applications in which many of the particles could be generated in the regions with low likelihoods. An improvement over the standard SIR-PF was made by the APF, in which the most recent observation is used to make new proposals and the samples are drawn from the joint distribution of both the prior and the likelihood; such an advantage of the APF, however, can be demonstrated only in some restrictive cases where the process noise is small and the likelihood does not vary dramatically over the prior. Nevertheless, a contradictory scenario may exist: when the prior distribution is tightly peaked and differs significantly from the likelihood, the samples drawn from the joint distribution could suffer a severe reduction in diversity; as a result, many predictive particles could still be generated in the regions with low likelihood. Furthermore, the APF is usually vulnerable to the outlying proposals, and the variance of the resulting importance weights could become extremely large when the process noise rises.

To tackle the aforementioned challenges, a regularized auxiliary particle filtering (RAPF) technique is proposed in this work to improve the effectiveness and robustness of the APF. The goal is to develop a more reliable engineering tool for system state estimation and forecasting. The proposed RAPF technique is new in the following aspects: (1) the empirical density regularization is implemented into the APF and the proposal samples are redrawn from a continuous distribution to diversify the particles; and (2) a rejection/resampling approach is proposed to reject the predictive proposal outliers to improve the robustness of the APF.

The remainder of this paper is organized as follows. The proposed RAPF technique is described in section 2. The effectiveness of the RAPF is demonstrated in section 3 via both simulations of a nonlinear/non-Gaussian benchmark model and an application in battery RUL prognostics. Some important observations and conclusive remarks are summarized in section 4.

2. The regularized auxiliary particle filter

2.1. Problem statement

Consider a general state space model with unobserved states α_t and an observed time series y_t, $t = 1, \ldots, n$. Assume that α_t is a Markov process with an initial density $p(\alpha_0)$ and the probability transition density is represented by $f(\alpha_t|\alpha_{t-1})$. Given $\{\alpha_1, \alpha_2, \ldots, \alpha_t\}$, the observations $Y_t \triangleq \{y_1, y_2, \ldots, y_t\}$ are conditionally independent and have a marginal distribution $f(y_t|\alpha_t)$. The inference of the property of the states α_t relies on the marginal filtering density $f(\alpha_t|Y_t)$. Suppose that the density $f(\alpha_{t-1}|Y_{t-1})$ is available at time instant $t-1$, then the prior density of the state at time instant t can be estimated via the transition density $f(\alpha_t \mid \alpha_{t-1})$ such that

$$f(\alpha_t|Y_{t-1}) = \int f(\alpha_t|\alpha_{t-1}) f(\alpha_{t-1}|Y_{t-1})\, \mathrm{d}\alpha_{t-1}. \tag{1}$$

Correspondingly, the marginal filtering density is computed via the Bayes' theorem,

$$f(\alpha_t|Y_t) = \frac{f(y_t|\alpha_t) f(\alpha_t|Y_{t-1})}{f(y_t|Y_{t-1})}, \tag{2}$$

where the normalizing constant is determined by

$$f(y_t|Y_{t-1}) = \int f(y_t|\alpha_t) f(\alpha_t|Y_{t-1})\, \mathrm{d}\alpha_t. \tag{3}$$

The above equations constitute the formal solution to the Bayesian recursive estimation problem. However, except in very special scenarios, these density functions do not admit a closed-form solution and thus numerical approximations are usually employed.

Particle filtering is a technique for implementing the recursive Bayesian filtering via Monte Carlo simulations, whereby the posterior density function $f(\alpha_t|Y_t)$ is represented by a set of random samples (particles) $\alpha_t^1, \ldots, \alpha_t^M$ and their associated weights $\pi_t^1, \ldots, \pi_t^M$,

$$f(\alpha_t|Y_t) \approx \sum_{i=1}^{M} \pi_t^i \delta(\alpha_t - \alpha_t^i), \qquad \sum_{i=1}^{M} \pi_t^i = 1, \tag{4}$$

where M is the number of particles; the weights π_t^i can be recursively updated using the importance sampling with an importance density $g(\alpha_t|\alpha_{t-1}^i, y_t)$,

$$\pi_t^i \propto \pi_{t-1}^i \frac{f(y_t|\alpha_t^i) f(\alpha_t^i|\alpha_{t-1}^i)}{g(\alpha_t^i|\alpha_{t-1}^i, y_t)}. \tag{5}$$

There are several choices for the importance function $g(\alpha_t|\alpha_{t-1}^i, y_t)$; in terms of minimizing the weight variance [14], the optimal one is given by

$$g(\alpha_t|\alpha_{t-1}^i, y_t) = \frac{f(y_t|\alpha_t, \alpha_{t-1}^i) f(\alpha_t|\alpha_{t-1}^i)}{f(y_t|\alpha_{t-1}^i)}. \tag{6}$$

Substituting (6) into (5) yields

$$\pi_t^i \propto \pi_{t-1}^i f(y_t \mid \alpha_{t-1}^i) = \pi_{t-1}^i \int f(y_t \mid \alpha_t) f(\alpha_t \mid \alpha_{t-1}^i)\, \mathrm{d}\alpha_t. \tag{7}$$

For nonlinear and non-Gaussian systems, the integral in (7) does not admit a closed-form expression unless α_t is a finite set. In the APF, this integration problem was addressed by using a single-point approximation $f(y_t|\alpha_t) = f(y_t|\mu(\alpha_{t-1}))$, where $\mu(\alpha_{t-1})$ is the mean, or the mode, or a random draw associated

with the density $f(\alpha_t|\alpha_{t-1})$. Thus, the samples can be drawn from the joint density,

$$g(\alpha_t, k|Y_t) = g(k|Y_t)g(\alpha_t|k, Y_t) = g(k|Y_t)f(\alpha_t|\alpha_{t-1}^k),$$
$$k = 1, 2, \ldots, M, \propto \pi_{t-1}^k f(y_t|\mu_t^k)f(\alpha_t|\alpha_{t-1}^k), \quad (8)$$

where μ_t^k denotes $\mu(\alpha_{t-1}^k)$; the index k is related to the initial proposal α_{t-1}^k, which was used as an auxiliary variable in the first-stage simulation. The proposal α_{t-1}^k that corresponds to a larger $g(k|y_t)$ is more likely (or to be chosen more times) to survive in the second-stage sampling. It can be seen that the most recent observation y_t has been employed in making new proposals. Accordingly, more particles will be generated in the regions of the state space with larger predictive likelihood $f(y_t|\mu_t^k)$, and thus the statistical efficiency of the sampling process can be improved. Correspondingly, the second-stage weights are updated by

$$\pi_t^i \propto \frac{f(y_t|\alpha_t^i)}{f(y_t|\mu_t^{k^i})}, \qquad i = 1, 2, \ldots, M \quad (9)$$

where the index k^i means that the ith draw α_t^i in the second-stage sampling is related to the initial proposal α_{t-1}^k through the density $f(\alpha_t^i|\alpha_{t-1}^k)$.

Although the APF technique is attractive in new proposal generation, it has two potential limitations in real applications: (1) when the process noise is small (or the transition density $f(\alpha_t|\alpha_{t-1})$ is tightly peaked), and the likelihood $f(y_t|\alpha_t)$ differs significantly from the prior $f(\alpha_t|\alpha_{t-1})$, the index sampling from $g(k|y_t)$ may lead to repeated draws from a specific discrete point, which can result in the problem of loss of diversity among the newly generated particles. (2) The second-stage weight estimation will be very sensitive to the proposal outliers that occur when $f(\alpha_t|\alpha_{t-1})$ is rather diffused and the likelihood $f(y_t|\alpha_t)$ varies dramatically over the prior $f(\alpha_t|\alpha_{t-1})$; in such a scenario, an excessive difference between $\mu_t^{k^i}$ and α_t^i can lead to a large (or even an infinite) importance weight in (9).

2.2. The regularized auxiliary particle filter (RAPF)

The proposed RAPF technique aims to overcome the aforementioned limitations of the APF through two innovative processes.

2.2.1. Regularization of the empirical density $g(k|y_t)$ associated with the first-stage weights. A diversity problem exists in the APF. The first step in the APF is to sample the index k from the density $g(k|y_t)$, where k is the index of the initial proposal α_{t-1}^k. The proposal α_{t-1}^k with a higher probability $\pi_{t-1}^k f(y_t|\mu_t^k)$ will be selected repeatedly in the sampling process, whereas the proposal α_{t-1}^k with a lower probability $\pi_{t-1}^k f(y_t|\mu_t^k)$ is prone to dying out. The surviving particles will then go to the second-stage sampling through the density $f(\alpha_t^i|\alpha_{t-1}^{k^i})$. When the process noise is small, the resulting α_t^i from the repeatedly selected $\alpha_{t-1}^{k^i}$ may drop into a very narrow neighborhood; this diversity problem will become more severe if the likelihood $f(y_t|\alpha_t)$ differs significantly from the prior $f(\alpha_t|\alpha_{t-1})$. In the proposed RAPF technique, this diversity problem will be solved by regularizing the empirical density $g(k|y_t)$ (given α_{t-1}^k, π_{t-1}^k, and y_t) and redrawing the samples (denoted as α_{t-1}^j, $j = 1, 2, \ldots, M$) from a continuous distribution. The rationale behind this approach is to implement the empirical density regularization into the APF to diversify particle generation. Kernel density estimation is a fundamental technique in statistics to estimate the probability density function of a random variable, given a set of samples from a population. The challenging part in the related research actually lies in how to properly implement this technique in particle filtering to effectively diversify the particles. The regularized particle filter in [17] suggested an approach to realize the empirical density regularization in the standard particle filter (SIR-PF). However, the implementation of the empirical density regularization in the auxiliary particle filter (APF) proves to be more indispensable and challenging, because the joint distribution of both the prior and the likelihood usually vary more dramatically than a single distribution. In this work, we will apply the kernel smoothing on the empirical density associated with this joint distribution, which shows it to be an effective approach through both simulations and real-world applications, as demonstrated in section 3. Specifically, the samples α_{t-1}^j are drawn from the following regularized density:

$$g(\alpha_{t-1}|y_t) \approx \sum_{k=1}^{M} \pi_{t-1}^k f(y_t|\mu_t^k) K_h(\alpha_{t-1} - \alpha_{t-1}^k), \quad (10)$$

where μ_t^k can be the mean, the mode, or a random draw associated with the density of $f(\alpha_t|\alpha_{t-1})$. $K_h(\cdot)$ is the rescaled kernel function [17] given by

$$K_h(x) = h^{-n} K(x/h), \quad (11)$$

where $h > 0$ is the scalar kernel bandwidth and n is the dimension of the state vector x. The kernel function is a symmetric probability density:

$$\int x K(x)\,\mathrm{d}x = 0, \qquad \int \| x \|^2 K(x)\,\mathrm{d}x > \infty. \quad (12)$$

The kernel $K_h(\cdot)$ and the bandwidth should be properly chosen to minimize the mean integrated square error (MISE) between the true density $g(\alpha_{t-1}|y_t)$ and its corresponding regularized empirical representation $\sum_{k=1}^{M} \pi_{t-1}^k f(y_t|\mu_t^k) K_h(\alpha_{t-1} - \alpha_{t-1}^k)$. If the weights are equal, the optimal kernel will be

$$K_{\mathrm{opt}}(x) = \begin{cases} \dfrac{n+2}{2c_n}(1 - \| x \|^2) & \text{if } \| x \| > 1 \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

where c_n is the volume of the n-dimensional hypersphere and $K_{\mathrm{opt}}(x)$ is the Epanechnikov kernel. The n-dimensional unit hypersphere is the set of points that are one unit away from the origin and its volume is given by

$$c_n = \begin{cases} 2 & n = 1 \\ \pi & n = 2 \\ 4\pi/3 & n = 3 \\ \vdots & \vdots \\ 2\pi c_{n-2}/n & n = n. \end{cases} \quad (14)$$

When the underlying density is Gaussian with a unit covariance matrix, the optimal bandwidth will be

$$h_{\text{opt}} = [8c_n^{-1}(n+4)(2\pi)^n]^{1/n+4} M^{-1/n+4}. \quad (15)$$

For a general case with an arbitrary underlying density, it is necessary to assume that the density is Gaussian and its covariance matrix S equals the empirical covariance matrix of the samples from this density. Accordingly, the kernel function in (11) can be applied in the form of

$$K_h(x) = (\det A)^{-1} h^{-n} K(A^{-1}x/h), \quad (16)$$

where A is the square root matrix of S such that $AA^T = S$. Interested readers can refer to [19] for more information about the kernel density estimation.

2.2.2. A rejection/resampling approach for the second-stage weight updating. The APF technique may not be used effectively to reduce the variance of the second-stage weights when there exist some outlying proposals from the density $f(\alpha_t|\alpha_{t-1})$; that is, the second-stage weights are actually not upper-bounded [20]. A large (or even an infinite) weight variance could be incurred when $f(\alpha_t|\alpha_{t-1})$ is rather diffused while the likelihood $f(y_t|\alpha_t)$ varies dramatically over the prior $f(\alpha_t|\alpha_{t-1})$. We have observed this problem in our research and we believe it is worth documenting because this pitfall could lead to a serious particle degeneracy problem in applications. Correspondingly in this work we will propose a rejection/resampling approach to address this large-variance issue. The acceptance–rejection method has been used in statistics and signal processing for sampling from probability distributions; however, the criterion/mechanism that drives each acceptance–rejection test usually differs from one application to another, such as those in the Metropolis–Hastings algorithm [21], the prior editing [15], and the resample-move algorithm [22]. The approach proposed in this work aims to artificially boost the new proposals α_t^j in the vicinity of μ_t^j, where $\mu_t^j = \mathrm{E}(\alpha_t|\alpha_{t-1}^j)$. It should be noted that the term μ_t^k has a similar meaning to the term μ_t^j except that the index k is the auxiliary variable originating from the initial proposal α_{t-1}^k whereas the index j is the resampled index. In other words, $k \in \{1, 2, 3, \ldots, M\}$ and can be repeatable, whereas $j = 1, 2, 3, \ldots, M$ and is not repeatable. This approach subjects the proposals drawn from the second-stage sampling to a pragmatic acceptance test by the following steps.

(a) Take the initial proposal α_{t-1}^j at the time instant $t-1$ and pass it through the transition density $f(\alpha_t|\alpha_{t-1})$ to generate a new proposal α_t^j for the time instant t.
(b) Compute the second-stage weight by

$$\pi_t^j \propto \frac{f(y_t|\alpha_t^j)}{f(y_t|\mu_t^j)}. \quad (17)$$

(c) Accept α_t^j if $1/W \leqslant \pi_t^j \leqslant W$; otherwise, reject α_t^j and return to step (a) to redraw a sample α_t^j, and repeat the test.

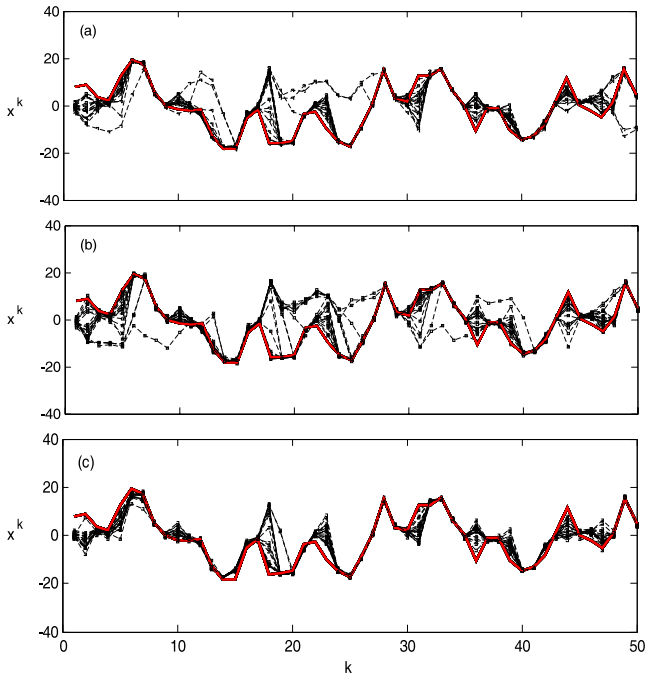

Figure 1. Performance comparison over 20 exemplar runs by (a) SIR-PF; (b) APF; (c) RAPF. The solid (red) line represents the true states whereas the dotted (black) lines represent the estimated states.

(d) Set $j = j + 1$ and repeat steps (a)–(d) until $j > M$.

The weight threshold W is a design parameter. If W is infinitely large, the proposed rejection/resampling approach will accept all of the newly generated proposals and the variance of the weights π_t^j may become extremely large as has been observed in the APF; in this sense, the APF is actually a special case of the developed RAPF technique. As W becomes smaller, the proposed rejection/resampling approach will smooth out the outlying proposals and make the weights π_t^j less variable. By testing, $W = 2$ is employed in the experiments in section 3.

A pseudo-code description of the proposed RAPF technique is summarized in table 1.

3. Performance evaluation

The effectiveness of the proposed RAPF technique will be examined in this section firstly by simulations based on a benchmark model and then it will be implemented for battery state prognosis. Its performance will be compared with the standard SIR-PF and the APF.

3.1. Simulation on a nonlinear/non-Gaussian benchmark model

Consider the following model that has been widely used in the related research [11]:

$$x_k = \frac{1}{2}x_{k-1} + \frac{25x_{k-1}}{1+x_{k-1}^2} + 8\cos[1.2(k-1)] + w_k, \quad (18)$$

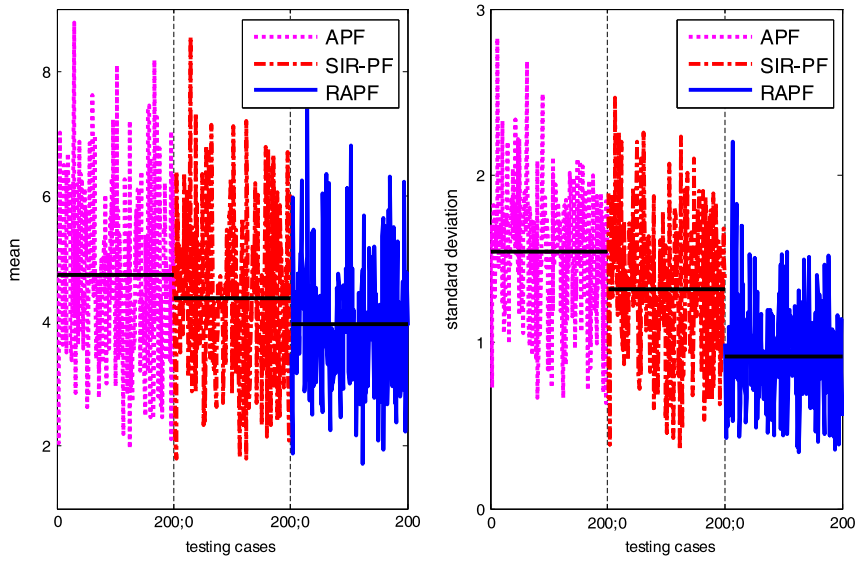

Figure 2. Test results, in terms of (a) mean and (b) standard deviation of RMSE, from three different particle filters when the variance of the process noise is 1. APF: dotted lines (magenta); SIR-PF: dash-dotted lines (red); RAPF: solid lines (blue).

Table 1. The regularized auxiliary particle filter.

$[\{\alpha_t^j, \pi_t^j\}_{j=1}^M] = \text{RAPF}[\{\alpha_{t-1}^i, \pi_{t-1}^i\}_{i=1}^M, y_t]$

- FOR $i = 1 : M$
  - Estimate $\mu_t^i = \text{E}(\alpha_t|\alpha_{t-1}^i)$
  - Calculate $\lambda^i \propto g(\text{i}|Y_t) \propto \pi_{t-1}^i f(y_t|\mu_t^i)$
- END FOR
- Calculate the total weight: $Q = \text{SUM}[\{\lambda^i\}_{i=1}^M]$
- FOR $i = 1{:}M$
  - Normalize $\lambda^i = Q^{-1}\lambda^i$
- END FOR
- Calculate the empirical covariance matrix S_{t-1} of $\{\alpha_{t-1}^i, \lambda^i\}_{i=1}^M$
  - Compute A_{t-1} such that $A_{t-1}A_{t-1}^T = S_{t-1}$ (Cholesky decomposition)
- Perform resampling by using some appropriate resampling algorithm (e.g. multinomial resampling and residual resampling)
  - $[\{\alpha_{t-1}^j, \pi_{t-1}^j = 1/M\}_{j=1}^M] =$ RESAMPLE$[\{\alpha_{t-1}^i, \lambda^i\}_{i=1}^M]$
- FOR $j = 1 : M$
  - Generate $\varepsilon \sim K$, Epanechnikov kernel
  - Calculate h_{opt} using (15)
  - Compute $\alpha_{t-1}^{j*} = \alpha_{t-1}^j + h_{\text{opt}}A_{t-1}\varepsilon$
- END FOR
- FOR $j = 1 : M$
  - Estimate $\mu_t^j = \text{E}(\alpha_t|\alpha_{t-1}^{j*})$
  - Draw $\alpha_t^j \sim f(\alpha_t|\alpha_{t-1}^*)$
  - Compute $\pi_t^j \propto f(y_t|\alpha_t^j)/f(y_t|\mu_t^j)$; if $\pi_t^j < 1/W$ OR $\pi_t^j > W$, go back to the above step to redraw α_t^j and then check if the weight of this new sample satisfies $1/W \leqslant \pi_t^j \leqslant W$.
- END FOR

$$y_k = \tfrac{1}{20}x_k^2 + v_k, \tag{19}$$

where $\{w_k\}$ and $\{v_k\}$ are the Gaussian white noise sequences that have zero means; $k = 50$ in this test. The variance of $\{v_k\}$ is 1, whereas the variance of $\{w_k\}$ is set to 1, 4, and 10, respectively, in three different simulation scenarios. The initial state is taken to be $x_0 = 0.1$. In the tests, all the particle filters have 50 particles (i.e. $M = 50$). The multinomial resampling technique is implemented at each time step. The SIR-PF uses the prior density as the importance function. The testing results from 20 random runs using the same state/observation data set are shown in figure 1. It can be seen that the proposed RAPF outperforms both the standard SIR-PF and the APF. The RAPF can provide a more reliable state estimation thanks to its capability in enhancing the diversity of the particles as well as its robustness to the effects of outlying proposals. The relatively poor performance of the APF is mainly due to its limitations as stated in section 2.

More extensive simulations have been conducted to further verify the effectiveness of the proposed RAPF. The variance of the process noise $\{w_k\}$ is set to be 1, 4, and 10,

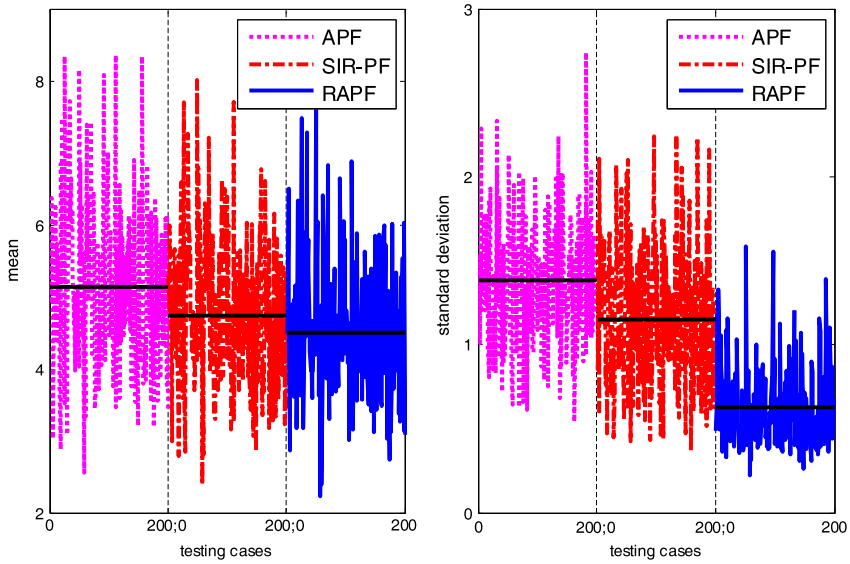

Figure 3. Test results, in terms of (a) mean and (b) standard deviation of RMSE, from three different particle filters when the variance of the process noise is 4. APF: dotted lines (magenta); SIR-PF: dash-dotted lines (red); RAPF: solid lines (blue).

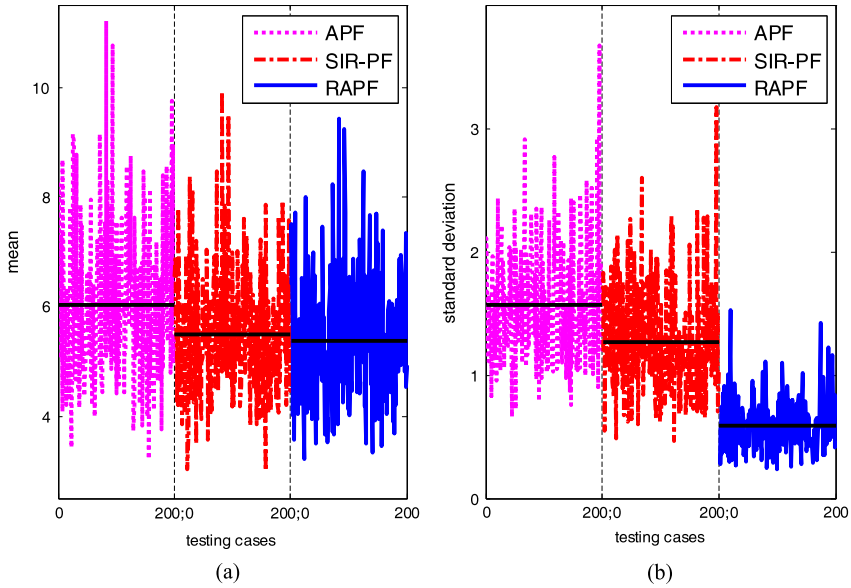

Figure 4. Test results, in terms of (a) mean and (b) standard deviation of RMSE, from three different particle filters when the variance of the process noise is 10. APF: dotted lines (magenta); SIR-PF: dash-dotted lines (red); RAPF: solid lines (blue).

respectively, in three testing scenarios. In each scenario, 200 data sets in total are randomly generated by using the model given by equations (18) and (19). For each state/observation data set, three particle filters are implemented and tested over 50 runs. The root mean square error (RMSE) between the actual states and the estimated states are calculated for each run. The performance of the related particle filters is compared in terms of the mean and standard deviation of these RMSEs over 50 runs, and the testing results are illustrated in figures 2–4, respectively. The averaged mean and standard deviation of RMSEs over 200 data sets are summarized in table 2. It can be seen that the proposed RAPF can provide a more accurate (i.e. with a lower mean) and more reliable (i.e. with a smaller variance) state estimation than the other two classical particle filters. This is because the RAPF technique can effectively diversify the particles and further enhance its robustness against outlying proposals. It is also worth mentioning that in the state estimation the initial density $p(\alpha_0)$ is usually chosen (often arbitrarily) to presumably cover the effective range of the system states. In this work, both uniform distribution and Gaussian distribution have been tested as the initial density; it has been found that there is no clear

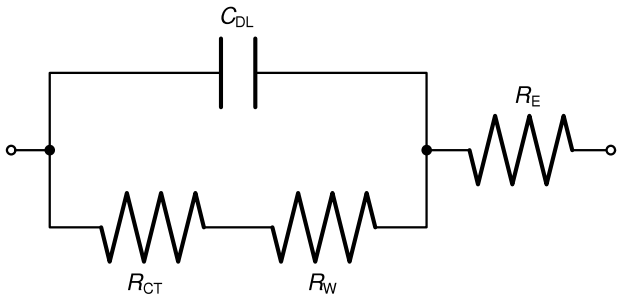

Figure 5. Lumped-parameter model for a lithium-ion battery.

difference in estimation performance when these two different distributions are implemented respectively.

3.2. Application for battery remaining useful life prediction

Batteries are widely used in various engineering and household systems. An effective prognostic tool is critically needed to predict the future condition of a battery in order to estimate its RUL. The reliable RUL information can be used not only for battery fault detection to prevent performance degradation of the related equipment, but also for scheduling battery change which is critical in many applications such as the emerging electric vehicle and aerospace industries. Particle filtering based prognostics has been investigated by the Prognostics Center of Excellence research group at NASA Ames Research Center based on the classical particle filtering techniques [7]. In this work, the developed RAPF is implemented for the RUL prediction of lithium-ion batteries. The lumped-parameter model in [7] is adopted in this investigation, as reproduced in figure 5, where R_E denotes the electrolyte resistance, R_{CT} is the charge transfer resistance, R_W is the Warburg impedance, and C_{DL} is the dual layer capacitance.

In general, a lithium-ion battery is deemed to fail when its capacity $C/1$ fades by 30% of the rated value. The batteries' capacity, however, is usually inaccessible to the measurement; therefore, the lumped parameters R_E and R_{CT} are employed for battery RUL prediction. $R_E + R_{CT}$ is usually inversely proportional to the capacity $C/1$ [18], and can be estimated through the electrochemical impedance spectroscopy test. The state and measurement equations that describe the battery model [7] are given by

$$\begin{cases} Z_0 = C; \qquad \Lambda_0 = \Lambda \qquad Z_k = Z_{k-1} \exp \Lambda_k + w_k \\ \Lambda_k = \Lambda_{k-1} + v_k \qquad X_k = [Z_k; \Lambda_k] \qquad Y_k = Z_k + v_k \end{cases} \tag{20}$$

where the vector Z denotes R_E or R_{CT}, and C and Λ are exponential growth model parameters. The Z and Λ vectors are combined to form the state vector X. The measurement vector Y comprises of the battery parameters inferred from measured data. The value of C takes the initial value of R_E or R_{CT}. The value of Λ will be recursively updated in the estimation process, whose initial value is derived from the training data using a least square estimator. The vectors w, v, and υ are zero mean Gaussian noise. The developed RAPF, the standard SIR-PF, and the APF are implemented for system state estimation and model parameter identification using the first part of R_E or R_{CT} trajectory (measurements); the identified model is then applied to predict the remaining part of the trajectory. In each iteration, 2500 particles are employed for the processing of this two-dimensional problem. The time to trigger the prognosis is arbitrarily chosen. Figure 6 shows both the state tracking and future state prediction for the data R_E and R_{CT}. The prognosis is initiated at week 24. It can be seen that the RAPF provides a more accurate forecasting for the battery parameters. Through a linear transformation as discussed in [7], the derived capacities are plotted in figure 7. The RUL threshold is chosen to be 70% of the rated capacity. The RAPF technique yields an RUL error of 1.54 weeks early, which outperforms the SIR-PF (6.62 weeks early) and the APF (5.71 weeks early), respectively.

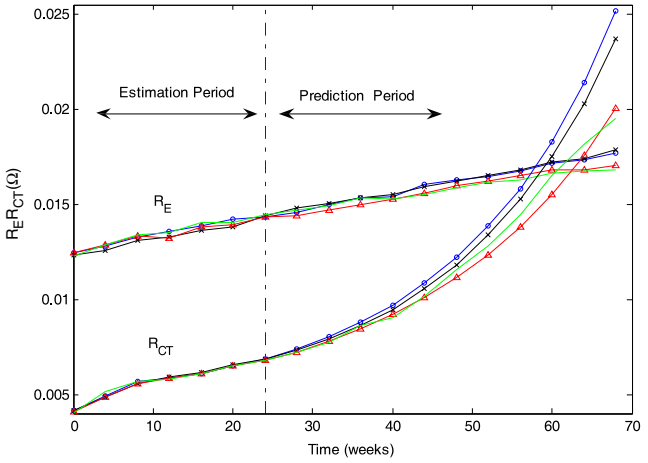

Figure 6. State tracking and future state prediction at week 24 for the battery parameters R_E and R_{CT} by using three different particle filters; solid line (green): ground truth of R_E and R_{CT}; O-mark line (blue): results by SIR-PF; ×-mark line (black): results by APF; Δ-mark line (red): results by the proposed RAPF.

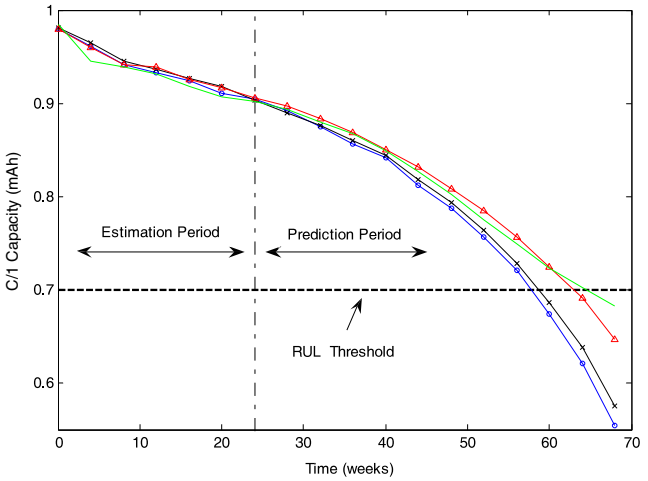

Figure 7. Battery RUL prediction at week 24 using three different particle filters; solid line (green): real measurement data; O-mark line (blue): results by SIR-PF; ×-mark line (black): results by APF; Δ-mark line (red): results by the proposed RAPF.

Table 2. The averaged mean and standard deviation of RMSEs over 200 data sets.

	Averaged mean of RMSE			Averaged standard deviation of RMSE		
Variance of process noise	APF	SIR-PF	RAPF	APF	SIR-PF	RAPF
1	4.731	4.372	3.954	1.546	1.318	0.917
4	5.129	4.731	4.495	1.385	1.151	0.631
10	6.037	5.499	5.379	1.564	1.267	0.584

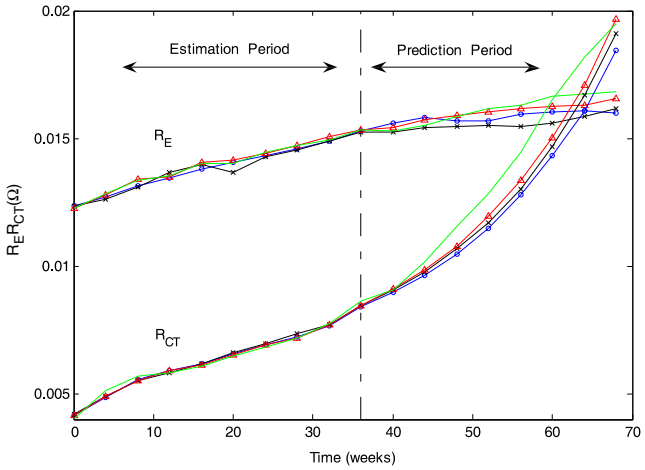

Figure 8. State tracking and future state prediction at week 36 for the battery parameters R_E and R_{CT} by using three different particle filters; solid line (green): ground truth of R_E and R_{CT}; O-mark line (blue): results by SIR-PF; ×-mark line (black): results by APF; Δ-mark line (red): results by the proposed RAPF.

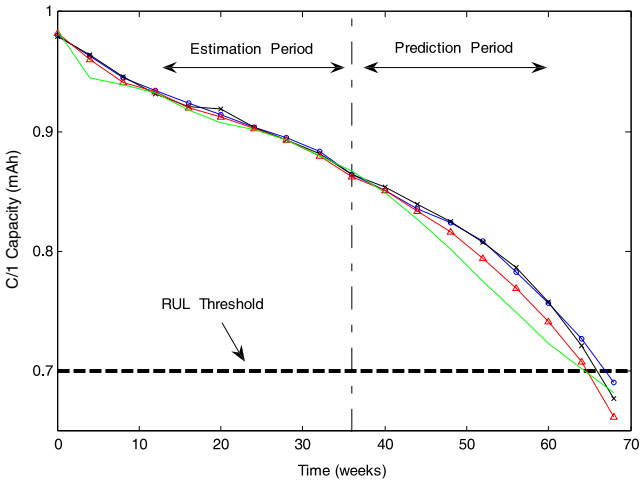

Figure 9. Battery RUL prediction at week 36 by using three different particle filters; solid line (green): real measurement data; O-mark line (blue): results by SIR-PF; ×-mark line (black): results by APF; Δ-mark line (red): results by the proposed RAPF.

Figures 8 and 9 show the testing results when the prognosis is triggered at week 36. It is seen that as more measurements become available over the period of state estimation, all three particle filters can improve learning to update the prediction model and generate more accurate forecasting RUL results. In this test, the RAPF yields an RUL error of 0.37 week late, which is much more accurate than the SIR-PF (2.33 weeks late) and APF (1.56 weeks late), respectively. The superior prognostic performance of the developed RAPF technique over the other two classical particle filters lies in the fact that the RAPF cannot only further diversify the particles but also effectively learns and updates the system states from a limited number of measurements.

4. Conclusions

In this paper, a regularized auxiliary particle filter (RAPF) has been developed for system current state estimation (monitoring) and future state prediction (prognosis). This RAPF technique can enhance the diversity of particles through empirical distribution regularization. A second-stage rejection/resampling technique is used to improve the robustness of the RAPF to outlying proposals. The effectiveness of the developed RAPF has been demonstrated through simulations of a nonlinear and non-Gaussian benchmark model. Testing results have shown that the proposed RAPF technique can provide more accurate and more reliable state estimation than the SIR-PF and the APF. The developed RAPF has also been implemented to predict the remaining useful life (RUL) of lithium-ion batteries. The results of this investigation have demonstrated that the RAPF technique can effectively learn system states from a limited number of measurements to update the nonlinear prediction model. It outperforms the related classical particle filters in battery RUL predictions.

Acknowledgments

The authors would like to thank Drs Kai Goebel, Bhaskar Saha, and Abhinav Saxena from the Prognostics Center of Excellence at NASA Ames Research Center for their assistance in this work. This project was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Carleton University. The authors would also like to thank the reviewers for their valuable suggestions.

References

[1] Kothamasu R, Huang S H and Verduin W H 2006 System health monitoring and prognostics—a review of current paradigms and practices *Int. J. Adv. Manuf. Technol.* **28** 1012–24

[2] Borguet S and Leonard O 2009 Coupling principal component analysis and Kalman filtering algorithms for on-line aircraft engine diagnostics *Control Eng. Practice* **17** 494–502

[3] Omenzetter P and Brownjohn J M W 2006 Application of time series analysis for bridge monitoring *Smart Mater. Struct.* **15** 129–38

[4] Shao Y and Mechefske C K 2009 Gearbox vibration monitoring using extended Kalman filters and hypothesis tests *J. Sound Vibr.* **325** 629–48

[5] Rezaei D and Taheri F 2010 Health monitoring of pipeline girth weld using empirical mode decomposition *Smart Mater. Struct.* **19** 055016
[6] Chinniah Y 2006 The extended Kalman filter as a tool for condition monitoring in hydraulic systems *Int. J. COMADEM* **9** 14–22
[7] Saha B, Goebel K, Poll S and Christophersen J 2009 Prognostics methods for battery health monitoring using a Bayesian framework *IEEE Trans. Instrum. Meas.* **58** 291–6
[8] Caesarendra W, Niu G and Yang B S 2010 Machine condition prognosis based on sequential Monte Carlo method *Expert Syst. Appl.* **37** 2412–20
[9] Kalman R E 1960 A new approach to linear filtering and prediction problems *ASME J. Basic Eng.* **82** 35–45
[10] Van Der Heijden F, Duin R P W, De Ridder D and Tax D M J 2004 *Classification, Parameter Estimation and State Estimation* (New York: Wiley)
[11] Simon D 2006 *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches* (New York: Wiley)
[12] Cadini F, Zio E and Avram D 2009 Monte Carlo-based filtering for fatigue crack growth estimation *Probab. Eng. Mech.* **24** 367–73
[13] Xue S, Tang H and Xie Q 2009 Structural damage detection using auxiliary particle filtering method *Struct. Health Monit.* **8** 101–12
[14] Doucet A, De Freitas N and Gordon N 2001 *Sequential Monte Carlo Methods in Practice* (New York: Springer)
[15] Gordon N, Salmond D J and Smith A F M 1993 Novel approach to nonlinear/non-Gaussian Bayesian state estimation *IEE Proc.* F **140** 107–13
[16] Pitt M K and Shephard N 1999 Filtering via simulation: auxiliary particle filters *J. Am. Stat. Assoc.* **94** 590–9
[17] Musso C, Oudjane N and Le Gland F 2001 Improving regularized particle filters *Sequential Monte Carlo Methods in Practice* (New York: Springer)
[18] Liu J, Saxena A, Goebel K, Saha B and Wang W 2010 An adaptive recurrent neural network for remaining useful life prediction of lithium ion batteries *Proc. Annu. Conf. Prognostics and Health Management Society (Portland, Oregon)*
[19] Silverman B W 1986 Density estimation for statistics and data analysis *Monographs on Statistics and Applied Probability* (London: Chapman and Hall)
[20] Johansen A M and Doucet A 2008 A note on auxiliary particle filters *Stat. Probab. Lett.* **78** 1498–504
[21] Rubinstein R Y and Kroese D P 2007 *Simulations and the Monte Carlo Method* (New York: Wiley)
[22] Gilks W R and Berzuini C 2001 Following a moving target—Monte Carlo inference for dynamic Bayesian models *J. R. Stat. Soc.* B **63** 127–46