13th World Conference on Earthquake Engineering
Vancouver, B.C., Canada
August 1-6, 2004
Paper No. 1092

APPLICATION OF BAYESIAN STATE ESTIMATION IN REAL-TIME LOSS ESTIMATION OF INSTRUMENTED BUIDLINGS

Ching, J. [1], Beck, J.L.[2] and Porter, K.A.[3]

SUMMARY

The focus of this paper is real-time Bayesian state estimation using nonlinear models. A recently developed method, the particle filter, is studied that is based on Monte Carlo simulation. Unlike the well-known extended Kalman filter, it is applicable to highly nonlinear systems with non-Gaussian uncertainties. The particle filter is applied to a real-data case study: a 7-story hotel whose structural system consists of non-ductile reinforced-concrete moment frames, one of which was severely damaged during the 1994 Northridge earthquake. An identification model derived from a nonlinear finite-element model of the building previously developed at Caltech is proposed. The particle filter provides consistent state and parameter estimates, in contrast to the extended Kalman filter. Finally, recorded motions from the 1994 Northridge earthquake are used to illustrate how to do real-time performance evaluation by computing estimates of the repair costs and probability of component damage for the hotel.

INTRODUCTION

State estimation is the process of using dynamic data from a system to estimate quantities that give a complete description of the state of the system according to some representative model of it. State estimation has the potential to be widely applied in civil engineering. Due to the wide applicability, real-time state estimation and identification methods have been studied in civil engineering for various purposes: Beck [1] used an invariant-embedding filter for modal identification; Yun [2] used an extended Kalman filter to study nonlinear fluid-structure interaction; Hoshiya [3] used the extended Kalman filter for structural system identification; Lin [4] developed a real-time identification methodology for better understanding of the degrading behavior of structures subject to dynamic loads; Ghanem [5] presented several different adaptive estimation techniques (e.g. extended Kalman filter, recursive least squares, recursive prediction error methods) and verified them using experimental data (Shinozuka [6]); Glaser [7] used the Kalman filter to identify the time-varying natural frequency and damping of a liquefied soil to get

[1] George W. Housner Research Fellow, Dept. of Civil Engineering, California Institute of Technology, Pasadena, CA 91125, USA. Email: jyching@caltech.edu
[2] Professor of Applied Mechanics and Civil Engineering, Dept. of Civil Engineering, California Institute of Technology, Pasadena, CA 91125, USA. Email: jimbeck@caltech.edu
[3] Senior George W. Housner Research Fellow, Dept. of Civil Engineering, California Institute of Technology, Pasadena, CA 91125, USA. Email: keithp@caltech.edu

insight into the liquefaction phenomenon; Sato [8] derived an adaptive H_∞ filter and applied it to time-varying linear and nonlinear structural systems in which displacements and velocities of the floors are measured; Smyth [9] formulated an on-line adaptive least squares algorithm for identifying multi-degree of freedom nonlinear hysteretic systems for the purpose of on-line control and monitoring.

Among state estimation methodologies, those founded on the Bayesian framework are powerful due to the following facts: (1) they are rigorously based on the probability axioms; therefore, they preserve information; (2) they give the probability density function (PDF) of the system state conditioned on the available information, which may then be used for any probability-based structural health monitoring, system identification, reliability assessment and control techniques. With the PDF available, we can not only estimate the state but also give a description of the associated uncertainties. For the Bayesian state-estimation algorithms, Kalman formulated the well-known Kalman filter (KF) (Kalman [10]) for linear systems with Gaussian uncertainties. Later, KF was modified to give the extended Kalman filter (EKF) (Jazwinski [11]) to accommodate lightly nonlinear systems, and this is basically the dominant Bayesian state-estimation algorithm for nonlinear systems and non-Gaussian uncertainties for the last 30 years. Although EKF has been widely used, it is only reliable for systems that are almost linear on the time scale of the updating intervals. However, civil engineering systems are often highly nonlinear when subject to severe loading events; in this case, the applicability of the Kalman filter and extended Kalman filter is often questionable. These older techniques have been used by civil engineering researchers for decades [1,2,3] although their applicability for nonlinear systems and non-Gaussian uncertainties is seldom verified either empirically or theoretically.

In this paper, we introduce some recent developments (Gordon [12]; Kitagawa [13]; Doucet [14]; van der Merwe [15]) in real-time Bayesian state estimation that use Monte Carlo simulation (MCS). The technique called particle filter (PF) is presented and discussed. These Monte Carlo techniques have the following advantages: (1) they are applicable to highly nonlinear systems with non-Gaussian uncertainties; (2) they are not limited to the first two moments as in KF and EKF; and (3) as the sample size approaches infinity, the resulting state estimates converge to their expected values. The performance of different methods (i.e. EKF and PF) will be compared through a real-data case study.

This paper has the following structure: We first define the general problem of Bayesian state estimation for nonlinear dynamical systems. We then review the Kalman filter and extended Kalman filter algorithms and introduce the particle filter techniques. Finally, we present a case study of a real building to demonstrate the application of the different methods. We also apply the results to illustrate real-time estimation of repair costs and probability of component damage of the building during an earthquake.

STATE ESTIMATION

Consider the following discrete-time state-space model of a dynamical system:

$$x_k = f_{k-1}(x_{k-1}, u_{k-1}, w_k) \qquad y_k = h_k(x_k, u_k, v_k) \qquad k = 1, 2 \ldots T \tag{1}$$

In this equation, $x_k \in R^n$, $u_k \in R^p$ and $y_k \in R^q$ are the system state, input (known excitation) and observed output at time k; $w_k \in R^l$ and $v_k \in R^m$ are introduced to account for unknown disturbances, model errors and measurement noise; f_k is the prescribed state transition function at time k; and h_k is the prescribed observation function at time k. The values of the variables x_k, y_k, w_k and v_k are uncertain and so are modeled as random variables, while u_k is considered to be a known excitation.

For each time k, the dynamical system input u_k and output $\hat{y}_k$ are measured. (In order to avoid confusion, we denote the observed output value by $\hat{y}_k$). We denote $\{\hat{y}_1, \hat{y}_2, ..., \hat{y}_k\}$ and $\{u_1, u_2, ..., u_k\}$ by

$\hat{Y}_k$ and U_k, respectively. The goal of real-time Bayesian state estimation is to evaluate the conditional probability density function (PDF) $p(x_k \mid \hat{Y}_k)$ for the state x_k at every time k in a real-time manner, i.e. to sequentially update this conditional PDF using the observed system input and output up to the current time, based on prescribed probabilistic models for w_k and v_k. Note that the conditioning of every PDF on U_k is not denoted explicitly.

The basic equations for updating $p(x_{k-1} \mid \hat{Y}_{k-1})$ to $p(x_k \mid \hat{Y}_k)$ are the predictor and updater (or corrector) equations that follow from the Theorem of Total Probability and Bayes Theorem, respectively:

$$p(x_k \mid \hat{Y}_{k-1}) = \int p(x_k \mid x_{k-1})\, p(x_{k-1} \mid \hat{Y}_{k-1}) dx_{k-1}$$
$$p(x_k \mid \hat{Y}_k) = \frac{p(\hat{y}_k \mid x_k)\, p(x_k \mid \hat{Y}_{k-1})}{\int p(\hat{y}_k \mid x_k)\, p(x_k \mid \hat{Y}_{k-1}) dx_k} = \frac{p(\hat{y}_k \mid x_k)\, p(x_k \mid \hat{Y}_{k-1})}{p(\hat{y}_k \mid \hat{Y}_{k-1})} \tag{2}$$

where $\hat{Y}_{k-1}$ is dropped in $p(x_k \mid x_{k-1})$ and $p(\hat{y}_k \mid x_k)$ because the models for the state transition and observation PDFs make it irrelevant. The main challenge in Bayesian state estimation for nonlinear systems is that these basic equations cannot be readily evaluated because they involve high-dimensional integrations.

KALMAN FILTER AND EXTENDED KALMAN FILTER

When f_k and h_k in (1) are both linear in u_k, x_k, w_k and v_k, i.e.

$$f_k(x_k, u_k, w_k) = A_k x_k + B_k u_k + G_k w_k \qquad h_k(x_k, u_k, v_k) = C_k x_k + D_k u_k + H_k v_k \tag{3}$$

and w_k and v_k are zero-mean independent Gaussian random variables with identity covariance matrices, the conditional PDF is also Gaussian and can be updated analytically. Furthermore, it is sufficient to update the first two moments because they completely specify this conditional PDF. The updating algorithm is the well-known Kalman filter (KF). It comprises two steps, the predictor (uncertainty propagation) and updater (estimation) steps.

In the uncertainty propagation step, the goal is to compute $p(x_k, y_k \mid \hat{Y}_{k-1})$ from $p(x_{k-1} \mid \hat{Y}_{k-1})$. First, $p(x_k \mid \hat{Y}_{k-1})$ is computed based on $p(x_{k-1} \mid \hat{Y}_{k-1})$ using the following moment equations:

$$x_{k|k-1} \equiv E(x_k \mid \hat{Y}_{k-1}) = A_{k-1} x_{k-1|k-1} + B_{k-1} u_{k-1}$$
$$P_{k|k-1} \equiv Cov(x_k \mid \hat{Y}_{k-1}) = A_{k-1} P_{k-1|k-1} A_{k-1}^T + G_{k-1} G_{k-1}^T \tag{4}$$

Note that the values $x_{0|0}$ and $P_{0|0}$ have to be given prior to the initialization of the algorithm. Second, $p(y_k \mid \hat{Y}_{k-1})$ is computed based on $p(x_k \mid \hat{Y}_{k-1})$ and u_k using the following moment equations:

$$y_{k|k-1} \equiv E(y_k \mid \hat{Y}_{k-1}) = C_k x_{k|k-1} + D_k u_k$$
$$P_{k|k-1}^y \equiv Cov(y_k \mid \hat{Y}_{k-1}) = C_k P_{k|k-1} C_k^T + H_k H_k^T \tag{5}$$

and finally, the conditional covariance between x_k and y_k is the $n \times q$ matrix computed as follows:

$$P_{k|k-1}^{xy} \equiv Cov(x_k, y_k \mid \hat{Y}_{k-1}) = P_{k|k-1} C_k^T \qquad (6)$$

This completes the computation of all the moments needed to specify the Gaussian PDF $p(x_k, y_k \mid \hat{Y}_{k-1})$. In the estimation step, $p(x_k \mid \hat{Y}_k)$ is updated based on the following equations:

$$\begin{aligned} x_{k|k} &= x_{k|k-1} + P_{k|k-1}^{xy} \cdot (P_{k|k-1}^{y})^{-1} \cdot (\hat{y}_k - y_{k|k-1}) \\ P_{k|k} &= P_{k|k-1} - P_{k|k-1}^{xy} \cdot (P_{k|k-1}^{y})^{-1} \cdot P_{k|k-1}^{xy\ T} \end{aligned} \qquad (7)$$

Extended Kalman filter

Many dynamical systems exhibit nonlinear behavior, and the direct use of KF is prohibited. However, if f_k and h_k are only slightly nonlinear, an approximation for KF can be derived by linearizing the uncertainty propagation and estimation steps. The resulting filter is the well-known extended Kalman filter (EKF). For the uncertainty propagation step in EKF, the goal is to find the linear approximations of $x_{k|k-1}$, $P_{k|k-1}$, $y_{k|k-1}$, $P_{k|k-1}^{y}$ and $P_{k|k-1}^{xy}$ based on $x_{k-1|k-1}$ and $P_{k-1|k-1}$. To simplify the notation, we define $z_k = [x_k^T \quad w_k^T \quad v_k^T]^T$ and $z_{k|k} = [x_{k|k}^T \quad 0^T \quad 0^T]^T \in R^{n+l+m}$, so $f_k(x_k, u_k, w_k) = f_k(z_k, u_k)$. When propagating from [$x_{k-1|k-1}$, $P_{k-1|k-1}$] to [$x_{k|k-1}$, $P_{k|k-1}$], we expand $f_{k-1}(z_{k-1}, u_{k-1})$ in the neighborhood of $z_{k-1|k-1}$. With the linear approximation, we get

$$x_{k|k-1} = E[f_{k-1}(z_{k-1}, u_{k-1})] \approx f_{k-1}(z_{k-1|k-1}, u_{k-1}) \equiv x_{k|k-1}^{LN} \qquad (8)$$

and

$$\begin{aligned} P_{k|k-1} &\approx (\nabla_z f_{k-1}) \cdot E\left\{(z_{k-1} - z_{k-1|k-1})(z_{k-1} - z_{k-1|k-1})^T \mid D_{k-1}\right\} \cdot (\nabla_z f_{k-1})^T \\ &= (\nabla_z f_{k-1}) \cdot Cov\left\{z_{k-1} \mid D_{k-1}\right\} \cdot (\nabla_z f_{k-1})^T \equiv P_{k|k-1}^{LN} \end{aligned} \qquad (9)$$

where $\nabla_z f_{k-1} \in R^{n \times (n+l+m)}$ is the Jacobian matrix evaluated at $z_{k-1} = z_{k-1|k-1}$. It can be seen that

$$P_{k|k-1}^{LN} = \left(A_{k-1}^{LN}\right) \cdot P_{k-1|k-1} \cdot \left(A_{k-1}^{LN}\right)^T + \left(G_{k-1}^{LN}\right) \cdot \left(G_{k-1}^{LN}\right)^T \qquad (10)$$

where $A_{k-1}^{LN} \equiv \nabla_x f_{k-1}\big|_{z_{k-1} = z_{k-1|k-1}} \in R^{n \times n}$ and $G_{k-1}^{LN} \equiv \nabla_w f_{k-1}\big|_{z_{k-1} = z_{k-1|k-1}} \in R^{n \times l}$ are the Jacobian matrices. Similarly, h_k is also linearized to get the approximations for $y_{k|k-1}$, $P_{k|k-1}^{y}$ and $P_{k|k-1}^{xy}$:

$$\begin{aligned} y_{k|k-1} &\approx h_{k-1}(x_{k|k-1}^{LN}, u_k, 0) \equiv y_{k|k-1}^{LN} \\ P_{k|k-1}^{y} &\approx \left(C_k^{LN}\right) P_{k|k-1}^{LN} \left(C_k^{LN}\right)^T + \left(H_k^{LN}\right)\left(H_k^{LN}\right)^T \equiv P_{k|k-1}^{y,LN} \\ P_{k|k-1}^{xy} &\approx P_{k|k-1}^{LN} \left(C_k^{LN}\right)^T \equiv P_{k|k-1}^{xy,LN} \end{aligned} \qquad (11)$$

where $C_k^{LN} \equiv \nabla_x h_k\big|_{x_k = x_{k|k-1}, v_k = 0} \in R^{q \times n}$ and $H_k^{LN} \equiv \nabla_v h_k\big|_{x_k = x_{k|k-1}, v_k = 0} \in R^{q \times m}$.

For the estimation step, (7) can still be used as an approximation. If f_k and h_k are indeed linear, EKF is identical to KF. The degree of accuracy of EKF relies on the validity of the linear approximation. EKF is not suitable to track multi-modal or highly non-Gaussian conditional PDFs due to the fact that it only updates the first two moments. When the system parameters are unknown, it is important to also estimate them. Uncertain parameters can be augmented into system states and estimated using EKF.

PARTICLE FILTERS

We have seen that EKF can only propagate and estimate the first two moments of the conditional PDF. For systems with non-Gaussian uncertainties, it is often desirable to propagate and estimate the conditional PDF itself; however, doing so effectively requires an infinite number of parameters to represent the functional form of the conditional PDF. An alternative is to conduct Monte Carlo simulation by drawing samples from the conditional PDF so that the conditional expectation of any function of x_k can be consistently estimated. We focus on the Monte Carlo simulation technique in this section and use the term particle filters (PF) to denote the resulting algorithms (following van der Merwe [15]; Doucet [16]). Similar PF algorithms have been called Monte Carlo filters by Kitagawa [13] and sequential Monte Carlo Bayesian filters by Doucet [14].

Basic equations

We first present some basic equations that are useful throughout this section. Let $X_k = \{x_0, x_1, ..., x_k\}$, then according to Bayes rule,

$$
\begin{aligned}
p(X_k \mid \hat{Y}_k) &= \frac{p(X_k, \hat{Y}_k)}{p(\hat{Y}_k)} = \frac{p(X_{k-1}, x_k, \hat{Y}_{k-1}, \hat{y}_k)}{p(\hat{Y}_k)} \\
&= \frac{p(X_{k-1}, \hat{Y}_{k-1}) \cdot p(x_k, \hat{y}_k \mid X_{k-1}, \hat{Y}_{k-1})}{p(\hat{Y}_k)} = \frac{p(X_{k-1} \mid \hat{Y}_{k-1}) \cdot p(x_k, \hat{y}_k \mid X_{k-1}, \hat{Y}_{k-1})}{p(\hat{y}_k \mid \hat{Y}_{k-1})} \\
&= p(X_{k-1} \mid \hat{Y}_{k-1}) \cdot \frac{p(\hat{y}_k \mid x_k, X_{k-1}, \hat{Y}_{k-1}) \cdot p(x_k \mid X_{k-1}, \hat{Y}_{k-1})}{p(\hat{y}_k \mid \hat{Y}_{k-1})} = p(X_{k-1} \mid \hat{Y}_{k-1}) \cdot \frac{p(\hat{y}_k \mid x_k) \cdot p(x_k \mid x_{k-1})}{p(\hat{y}_k \mid \hat{Y}_{k-1})}
\end{aligned}
\tag{12}
$$

where we have used the fact that $p(\hat{y}_k \mid x_k, X_{k-1}, \hat{Y}_{k-1}) = p(\hat{y}_k \mid x_k)$ and that $p(x_k \mid X_{k-1}, \hat{Y}_{k-1}) = p(x_k \mid x_{k-1})$ based on (1) and the fact that the PDFs for v_k and w_k are prescribed. Evaluating the recursive equation in (12), we get

$$
p(X_k \mid \hat{Y}_k) = p(x_0) \cdot \prod_{m=1}^{k} \frac{p(\hat{y}_m \mid x_m) \cdot p(x_m \mid x_{m-1})}{p(\hat{y}_m \mid \hat{Y}_{m-1})} = \frac{p(x_0)}{p(\hat{Y}_k)} \cdot \prod_{m=1}^{k} p(\hat{y}_m \mid x_m) \cdot p(x_m \mid x_{m-1}) \tag{13}
$$

Monte Carlo simulation for state estimation

Our interest is to develop a MCS algorithm for the conditional PDF $p(X_k \mid \hat{Y}_k)$ that is real-time. In other words, if $\hat{X}_{k-1}$ is a sample from $p(X_{k-1} \mid \hat{Y}_{k-1})$, the sample from $p(X_k \mid \hat{Y}_k)$ must have the form $\hat{X}_k = \{\hat{X}_{k-1}, \hat{x}_k\}$, where $\hat{x}_k$ is the new sample and $\hat{X}_{k-1}$ is the previous sample from $p(X_{k-1} \mid \hat{Y}_{k-1})$. However, such a real-time MCS algorithm cannot be directly implemented. This is because $p(X_{k-1} \mid \hat{Y}_{k-1})$ is different from $p(X_{k-1} \mid \hat{Y}_k) = p(X_{k-1} \mid \hat{Y}_{k-1}, \hat{y}_k)$.

Importance sampling technique

Nevertheless, we can sample from an importance sampling PDF $q(X_k \mid \hat{Y}_k)$ that admits a real-time sampling procedure, i.e. $q(X_{k-1} \mid \hat{Y}_{k-1})$ is identical to $q(X_{k-1} \mid \hat{Y}_k)$. In other words, the structure of $q(X_k \mid Y_k)$ is such that X_{k-1} is independent of y_k conditioned on Y_{k-1}. Drawing N samples $\left\{\hat{X}_k^i : i = 1, ..., N\right\}$ randomly from $q(X_k \mid \hat{Y}_k)$ (at this time, we assume that $q(X_k \mid \hat{Y}_k)$ can be easily sampled), the expectation of any function of X_k, denoted by $r(X_k)$, conditioned on $\hat{Y}_k$ can be estimated using the importance sampling technique as follows:

$$E[r(X_k) \mid \hat{Y}_k] \approx \frac{1}{N}\sum_{i=1}^{N} \hat{\beta}_k^i \cdot r(\hat{X}_k^i) \equiv \hat{r}_{k,N}^1 \tag{14}$$

where $\hat{\beta}_k^i = p(\hat{X}_k^i \mid \hat{Y}_k) \big/ q(\hat{X}_k^i \mid \hat{Y}_k)$ is the non-normalized importance weight of the *i*-th sample. Any quantity of interest can be estimated with different $r(\cdot)$ functions in (14); for instance, if $r(X_k) = X_k$, $E[r(X_k) \mid \hat{Y}_k]$ is simply the conditional expectation $E[X_k \mid \hat{Y}_k]$; if $r(X_k) = X_k X_k^T$, $E[r(X_k) \mid \hat{Y}_k]$ is the conditional second moment $E[X_k X_k^T \mid \hat{Y}_k]$.

Let $\left\{X_k^i : i = 1, ..., N\right\}$ denote the state variables corresponding to N random samples from $q(X_k \mid \hat{Y}_k)$ (before drawing the actual samples). It is readily shown that the estimator $r_{k,N}^1 = \frac{1}{N}\sum_{i=1}^{N} \beta_k^i \cdot r(X_k^i)$ is an unbiased estimator of $E[r(X_k) \mid \hat{Y}_k]$ if the support region for $p(X_k \mid \hat{Y}_k)$ is a subset of that for $q(X_k \mid \hat{Y}_k)$:

$$\begin{aligned} E\left[r_{k,N}^1\right] &= \frac{1}{N}\sum_{i=1}^{N} E_q\left[\beta_k^i \cdot r(X_k^i)\right] = E_q\left[\beta_k \cdot r(X_k)\right] \\ &= \int \left[p(X_k \mid \hat{Y}_k) \big/ q(X_k \mid \hat{Y}_k)\right] r(X_k) \cdot q(X_k \mid \hat{Y}_k) dX_k \\ &= \int r(X_k) \cdot p(X_k \mid \hat{Y}_k) dX_k = E\left[r(X_k) \mid \hat{Y}_k\right] \end{aligned} \tag{15}$$

According to the Central Limit Theorem, $r_{k,N}^1$ converges (as N approaches infinity) to a Gaussian random variable with mean equal to $E[r(X_k) \mid \hat{Y}_k]$ and with variance that decays as $1/N$. Therefore, $r_{k,N}^1$ is a consistent estimator of $E[r(X_k) \mid \hat{Y}_k]$.

Although $r_{k,N}^1$ is unbiased and consistent, it is not a feasible estimator because the non-normalized importance weights $\beta_k^i = p(X_k^i \mid \hat{Y}_k) \big/ q(X_k^i \mid \hat{Y}_k)$ depend on $p(X_k^i \mid \hat{Y}_k)$, which cannot be computed easily since in order to evaluate $p(X_k^i \mid \hat{Y}_k)$, we have to evaluate $p(\hat{Y}_k)$, as shown by (21), which is a difficult task. Nevertheless, it is shown in Ching [17] that the following estimator is computable while it is asymptotically unbiased and consistent:

$$r_{k,N}^{2} \equiv \left(\frac{1}{N} \sum_{i=1}^{N} \beta_k^i \cdot r(X_k^i) \right) \Bigg/ \left(\frac{1}{N} \sum_{j=1}^{N} \beta_k^j \right) = r_{k,N}^{1} \Big/ \bar{\beta}_k^N \tag{16}$$

where

$$\bar{\beta}_k^N = \left(\sum_{j=1}^{N} \beta_k^j \right) \Bigg/ N \tag{17}$$

Note that $\hat{r}_{k,N}^{2}$, unlike $\hat{r}_{k,N}^{1}$, can be computed conveniently from samples $\left\{ \hat{X}_k^i : i = 1, ..., N \right\}$:

$$\hat{r}_{k,N}^{2} = \sum_{i=1}^{N} \left[\beta_k^i \Bigg/ \left(\sum_{j=1}^{N} \beta_k^j \right) \right] \cdot r(\hat{X}_k^i) = \sum_{i=1}^{N} \tilde{\beta}_k^i \cdot r(\hat{X}_k^i) \tag{18}$$

where

$$\begin{aligned} \tilde{\beta}_k^i &= p(\hat{X}_k^i \mid \hat{Y}_k) / q(\hat{X}_k^i \mid \hat{Y}_k) \Bigg/ \left(\sum_{j=1}^{N} p(\hat{X}_k^j \mid \hat{Y}_k) / q(\hat{X}_k^j \mid \hat{Y}_k) \right) \\ &= \frac{p(\hat{x}_0^i)}{q(\hat{X}_k^i \mid \hat{Y}_k)} \cdot \prod_{m=1}^{k} p(\hat{y}_m \mid \hat{x}_m^i) \cdot p(\hat{x}_m^i \mid \hat{x}_{m-1}^i) \Bigg/ \left(\sum_{j=1}^{N} \frac{p(\hat{x}_0^j)}{q(\hat{X}_k^j \mid \hat{Y}_k)} \cdot \prod_{m=1}^{k} p(\hat{y}_m \mid \hat{x}_m^j) \cdot p(\hat{x}_m^j \mid \hat{x}_{m-1}^j) \right) \end{aligned} \tag{19}$$

Therefore, the factor $p(\hat{Y}_k)$ in (13) has been cancelled due to the use of the normalized importance weights $\left\{ \tilde{\beta}_k^i : i = 1, ..., N \right\}$, i.e. $\sum_{i=1}^{N} \tilde{\beta}_k^i = 1$. Also, the likelihood functions $p(\hat{y}_m \mid \hat{x}_m^i)$ and $p(\hat{x}_m^i \mid \hat{x}_{m-1}^i)$ can be readily evaluated using the prescribed PDFs for v_m and w_m if the mappings in (1) uniquely specify v_m and w_m, given y_m, x_m and x_{m-1}. At this time, we assume that $q(\hat{X}_k^i \mid \hat{Y}_k)$ can be easily evaluated too. The selection of an importance sampling PDF that admits a real-time procedure is discussed in Ching [17]. The conclusion is that the following importance sampling PDF performs better:

$$q(X_k \mid \hat{Y}_k) = p(x_0) \cdot \prod_{m=1}^{k} p(x_m \mid x_{m-1}, \hat{y}_m) \tag{20}$$

The corresponding modified non-normalized importance weight is:

$$\beta_k = \prod_{m=1}^{k} \frac{p(\hat{y}_m \mid x_m) \cdot p(x_m \mid x_{m-1})}{p(x_m \mid x_{m-1}, \hat{y}_m)} = \beta_{k-1} \cdot \frac{p(\hat{y}_k \mid x_k) \cdot p(x_k \mid x_{k-1})}{p(x_k \mid x_{k-1}, \hat{y}_k)} \tag{21}$$

Liu [18] discusses the optimality of this importance sampling PDF.
Due to the structure of the algorithm, at any time k, we are only required to store the sampled states and weights in the most recent two time steps, i.e. k and $k-1$, if the quantity of interest is $r(x_k)$ and so depends on the current state (clearly, additional dependence on the previous state x_{k-1} can also be treated). As a result, the following recursive algorithm can be used:

Algorithm 1: Basic PF algorithm

(1) Initialize the N samples: Draw $\hat{x}^i$ from $p(x_0)$ and set $\beta^i = 1/N$, $i = 1, ..., N$.

(2) At time k, store the previous samples and weights

$$\tilde{x}^i = \hat{x}^i \qquad \tilde{\beta}^i = \beta^i \tag{22}$$

For $i = 1,\ldots,N$, draw $\hat{x}^i$ from $p(x_k \mid x_{k-1} = \tilde{x}^i, \hat{y}_k)$ and update the importance weight

$$\beta^i = \tilde{\beta}^i \cdot \frac{p(\hat{y}_k \mid x_k = \hat{x}^i) \cdot p(x_k = \hat{x}^i \mid x_{k-1} = \tilde{x}^i)}{p(x_k = \hat{x}^i \mid x_{k-1} = \tilde{x}^i, \hat{y}_k)} \tag{23}$$

(3) For $i = 1,\ldots,N$, $E[r(x_k) \mid \hat{Y}_k]$ can be approximated based on (16):

$$E[r(x_k) \mid \hat{Y}_k] \approx \sum_{i=1}^{N} \left[\beta^i \Big/ \left(\sum_{j=1}^{N} \beta^j \right) \right] \cdot r(\hat{x}^i) \tag{24}$$

where $r(\cdot)$ is a function that maps from x_k to any quantity of interest.

(4) Do Steps (2) and (3) for $k = 1,\ldots,T$.

Usually, $p(x_k \mid x_{k-1} = \tilde{x}^i, \hat{y}_k)$ in Step 2 is difficult to sample. Note that estimating the first two moments of $p(x_k \mid x_{k-1} = \tilde{x}^i, \hat{y}_k)$ is a problem that can be solved using a single-time-step EKF algorithm. The least-informative PDF (i.e. the maximum entropy PDF; see Jaynes [19]) given the estimated two moments, which is a Gaussian PDF (denoted by $p_{LI}(x_k \mid x_{k-1} = \tilde{x}^i, \hat{y}_k)$; LI subscript means ‘least-informative’), can be used for the importance sampling PDF. The use of $p_{LI}(x_k \mid x_{k-1} = \tilde{x}^i, \hat{y}_k)$ is discussed in van der Merwe [15].

Algorithm 2: Determining $p_{LI}(x_k \mid x_{k-1} = \tilde{x}^i, \hat{y}_k)$

(1) Uncertainty propagation: compute

$$\begin{aligned} E_{LN}(x_k \mid x_{k-1} = \tilde{x}^i) &= f_{k-1}(x_{k-1} = \tilde{x}^i, u_{k-1}, w_{k-1} = 0) \equiv \tilde{x}^i_{k|k-1} \\ Cov_{LN}(x_k \mid x_{k-1} = \tilde{x}^i) &= (G^{LN}_{k-1}) \cdot Cov\{w_{k-1}\} \cdot (G^{LN}_{k-1})^T \equiv \tilde{P}^i_{k|k-1} \end{aligned} \tag{25}$$

where $G^{LN}_{k-1} \equiv \nabla_w f_{k-1}\big|_{x_{k-1} = \tilde{x}^i, w_{k-1} = 0}$ is the Jacobian matrix, and

$$\begin{aligned} E_{LN}(y_k \mid x_{k-1} = \tilde{x}^i) &= h_{k-1}(\tilde{x}^i_{k|k-1}, u_k, 0) \equiv \tilde{y}^i_{k|k-1} \\ Cov_{LN}(y_k \mid x_{k-1} = \tilde{x}^i) &= \left(C^{LN}_k\right) \tilde{P}^i_{k|k-1} \left(C^{LN}_k\right)^T + \left(H^{LN}_k\right)\left(H^{LN}_k\right)^T \equiv \tilde{P}^{y,i}_{k|k-1} \\ Cov_{LN}(x_k, y_k \mid x_{k-1} = \tilde{x}^i) &= \tilde{P}^i_{k|k-1} \left(C^{LN}_k\right)^T \equiv \tilde{P}^{xy,i}_{k|k-1} \end{aligned} \tag{26}$$

where $C^{LN}_k \equiv \nabla_x h_k\big|_{x_k = \tilde{x}^i_{k|k-1}, v_k = 0}$ and $H^{LN}_k \equiv \nabla_v h_k\big|_{x_k = \tilde{x}^i_{k|k-1}, v_k = 0}$.

(2) Estimation: compute

$$\begin{aligned} E_{LI}(x_k \mid x_{k-1} = \tilde{x}^i, \hat{y}_k) &= \tilde{x}^i_{k|k-1} + \tilde{P}^{xy,i}_{k|k-1} \cdot (\tilde{P}^{y,i}_{k|k-1})^{-1} \cdot (\hat{y}_k - \tilde{y}^i_{k|k-1}) \\ Cov_{LI}(x_k \mid x_{k-1} = \tilde{x}^i, \hat{y}_k) &= \tilde{P}^i_{k|k-1} - \tilde{P}^{xy,i}_{k|k-1} \cdot (\tilde{P}^{y,i}_{k|k-1})^{-1} \cdot \tilde{P}^{xy,i\,T}_{k|k-1} \end{aligned} \tag{27}$$

$p_{LI}(x_k \mid x_{k-1} = \tilde{x}^i, \hat{y}_k)$ is then the Gaussian PDF with the two moments in (27).

Reducing degradation – recursive resampling and parallel particle filters

Note that it is desirable to have the importance weights $\{\beta^i : i = 1, 2, ...N\}$ be approximately uniform so that all samples contribute significantly in (24), but they become far from uniform as k grows, which is due to the recursion in (21) and the fact that $q(X_k \mid \hat{Y}_k) \neq p(X_k \mid \hat{Y}_k)$. Ultimately, a few weights become much larger than the rest, so the effective number of samples is small. Nevertheless, the degradation can be reduced, as described in this section and the next.

Instead of letting the N samples evolve through time independently (Algorithm 1), we can resample the samples when the importance weights become highly non-uniform (Kitagawa [13]; Doucet [14]; Liu [18]). After the resampling, the importance weights become uniform, therefore the degradation problem is alleviated. The resampling step tends to terminate small-weight samples and duplicate large-weight samples and, therefore, forces the N samples to concentrate in the high probability region of $p(x_k \mid \hat{Y}_k)$.

Although the resampling step sets the weights back to uniform, the price to pay is that the samples become dependent and therefore collectively carry less information about the state. As a result, the resampling procedure should only be executed when the importance weights become highly non-uniform. This can be done by monitoring the coefficient of variation (c.o.v.) of the importance weights. The resampling procedure is executed only when this c.o.v. exceeds a certain threshold, indicating that the variability in the importance weights is large.

Another way to alleviate the dependency induced by the resampling step is to conduct several independent PF algorithms and combine all of the obtained samples. Although the samples obtained in a single algorithm can be highly dependent, the samples from different algorithms are completely independent. The resulting algorithm is as follows:

Algorithm 3: Parallel PF algorithm with resampling

(1) Perform the following (2)-(6) for $j = 1, ..., L$ independently. Since the processes are completely independent, they can be conducted in parallel.

(2) Initialize N samples: Draw $\tilde{x}^{i,j}$ from $p(x_0)$ and set $\beta^{i,j} = 1/N$ for $i = 1, ..., N$.

(3) At time k, store the previous samples and weights

$$\tilde{x}^{i,j} = \hat{x}^{i,j} \qquad \tilde{\beta}^{i,j} = \beta^{i,j} \tag{28}$$

For $i = 1, ..., N$, draw $\bar{x}^{i,j}$ from $p_{LI}(x_k \mid x_{k-1} = \tilde{x}^{i,j}, \hat{y}_k)$ and update the importance weight

$$\bar{\beta}^{i,j} = \tilde{\beta}^{i,j} \cdot \frac{p(\hat{y}_k \mid x_k = \bar{x}^{i,j}) \cdot p(x_k = \bar{x}^{i,j} \mid x_{k-1} = \tilde{x}^{i,j})}{p_{LI}(x_k = \bar{x}^{i,j} \mid x_{k-1} = \tilde{x}^{i,j}, \hat{y}_k)} \tag{29}$$

(4) Compute the c.o.v. of $\{\bar{\beta}^{i,j} : i = 1, ..., N\}$. If the c.o.v. is larger than the prescribed threshold, then execute the resampling step for $i = 1, ..., N$:

$$\hat{x}^{i,j} = \bar{x}^{i,j} \qquad w.p. \qquad \bar{\beta}^{i,j} \Big/ \sum_{i=1}^{N} \bar{\beta}^{i,j} \tag{30}$$

and set $\beta^{i,j} = 1/N$ for $i = 1, ..., N$. Otherwise, for $i = 1, ..., N$:

$$\hat{x}^{i,j} = \bar{x}^{i,j} \qquad \beta^{i,j} = \bar{\beta}^{i,j} \Big/ \sum_{i=1}^{N} \bar{\beta}^{i,j} \tag{31}$$

(5) $E[r(x_k)\,|\,\hat{Y}_k]$ can be then approximated by

$$E[r(x_k)\,|\,\hat{Y}_k] \approx \left(\sum_{i=1}^{N} \sum_{j=1}^{L} r(\hat{x}^{i,j}) \cdot \beta^{i,j} \right) \Big/ L \tag{32}$$

(6) Do Steps 2 to 5 for $k = 1, \ldots, T$.

Advantages and disadvantages of the PF technique

The advantages of the PF technique include (1) as N (the number of samples per algorithm) approaches infinity, the value of any function of the state x_k estimated by PF converges to its expected value; therefore, the PF technique can be used to validate other methodologies; and (2) parallel computations are possible for PF algorithms. A disadvantage of the PF technique is that it is computationally expensive, especially when the degradation is severe so that we need large N and L to have the algorithm converge. In general, the required N and L grow with the size of the effective support region of $p(x_k\,|\,\hat{Y}_k)$. A simple test for convergence is to add parallel particle filters until the estimated quantities of interest, $r(x_k)$, do not significantly change.

REAL-DATA CASE STUDY

Building description

The selected building for the case study is a 7-story, 66,000 square feet (6,200 m^2) hotel located at 8244 Orion Ave, Van Nuys, CA, at 34.221° north latitude, 118.471° west longitude, in the San Fernando Valley of Los Angeles County. We refer to this building as the Van Nuys hotel. It was built in 1966 according to the 1964 Los Angeles City Building Code. The lateral force-resisting system is a reinforced concrete moment frame in both directions. The building was lightly damaged by the M6.6 1971 San Fernando event, approximately 20 km to the northeast, and severely damaged by the M6.7 1994 Northridge Earthquake, whose epicenter was approximately 4.5 km to the southwest. The building has been studied extensively, e.g., by Jennings [20] and Beck [21]).

The south frame of the building was heavily damaged during the 1994 Northridge earthquake, during which the building was instrumented with sixteen accelerometers. Among the sixteen channels, five channels measured the accelerations in the longitudinal (east-west: E-W) direction at the ground, second, third, sixth floors and the roof. The locations of the five accelerometers were at the south-east corner of the ground floor and near the east wall at the second, third, sixth floors and the roof. Figure **1** shows the measured E-W acceleration time histories during the earthquake.

The purpose of this case study is to examine the use of real-time Bayesian state-estimation techniques in tracking the system state and parameters of the building during the earthquake and to apply the new technique to real-time loss monitoring. We focus exclusively on the dynamics in the E-W direction due to the fact that this involves the south frame that was severely damaged.

Time-varying nonlinear degradation identification model

We employ a time-varying nonlinear stochastic structural model for identification (called the simplified nonlinear degradation model). This model is derived from a more complicated finite-element model previously developed at Caltech and is discussed in detail in Ching [17]. The simplified model is a 7-story shear building model and has five system parameters which control the degradation behavior of the inter-story stiffnesses and dampings. The rationale that we allow the system parameters to drift with time is as follows: it is very unlikely that the simplified nonlinear degradation model has no modeling error. The modeling error can be partially compensated by allowing the system parameters to vary with time, that is, the model is able to adaptively fit the data by slowly changing its parameter values.

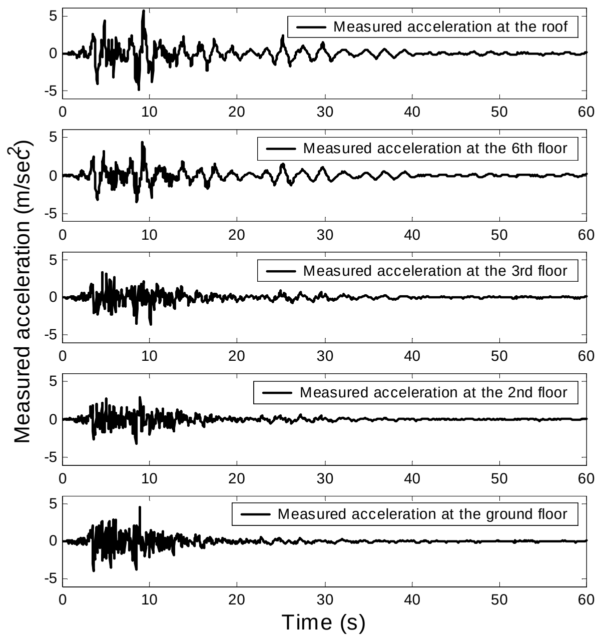

Figure 1. Measured E-W accelerations of the Van Nuys hotel during the Northridge earthquake

Results and discussion

We implement PF with N = 200 and L = 10 and the importance weight c.o.v. threshold = 200% using Algorithm 3. If more samples are used in PF than the $N \cdot L = 2000$ samples, there is little change in the estimated means and variances of the unknown parameters, indicating that the results are nearly converged. Figure **2** shows the evolution of the predicted inter-story drifts. The "measured" inter-story drifts in the first and second stories (obtained from double integrating and high-pass filtering the difference between the accelerations measured at the first story and ground floor base, and second and first stories) are also plotted in the figure. Figure 3 shows the time evolution of the estimated inter-story stiffnesses and dampings from the simplified nonlinear degradation model, where $k_{i,t}$ and $c_{i,t}$ denote the inter-story stiffness and damping in the i-th story at time t. By carefully inspecting Figure **2**, one can see that the predicted maximum inter-story drift is at the fourth story, which is consistent with the actual damage to the south frame in the 1994 Northridge Earthquake (Section 5.3.3 in Beck [21]).

Comparison with EKF results

We also implement the EKF algorithm for the simplified nonlinear degradation model. Although not shown in this paper, the results from EKF are not consistent with those from PF. Since we treat PF as our comparison standard, we conclude that EKF are not consistent. For the detailed results from EKF, please refer to Ching [17].

Real-time performance evaluation

Since the system state x_k of the simplified nonlinear degradation model completely characterizes the model status at time *k*, any desired performance measure of the monitored system at time *k* can be computed with an appropriate response function (the $r(\cdot)$ function appearing in Step 5 in Algorithm 3). With the PF algorithm, at each time *k*, $N \times L$ weighted samples of the state, i.e. $\{\hat{x}^{i,j} : i = 1,...,N;\ j = 1,...,L\}$, are available, so $N \times L$ weighted samples of the performance measure, i.e. $\{r(\hat{x}^{i,j}) : i = 1,...,N;\ j = 1,...,L\}$, can be obtained. The estimated mean of the performance measure at time *k* is given as in Algorithm 3 by

$$E[r(x_k) \mid \hat{Y}_k] \approx \left(\sum_{i=1}^{N} \sum_{j=1}^{L} r(\hat{x}^{i,j}) \cdot \beta^{i,j} \right) \Big/ L \tag{33}$$

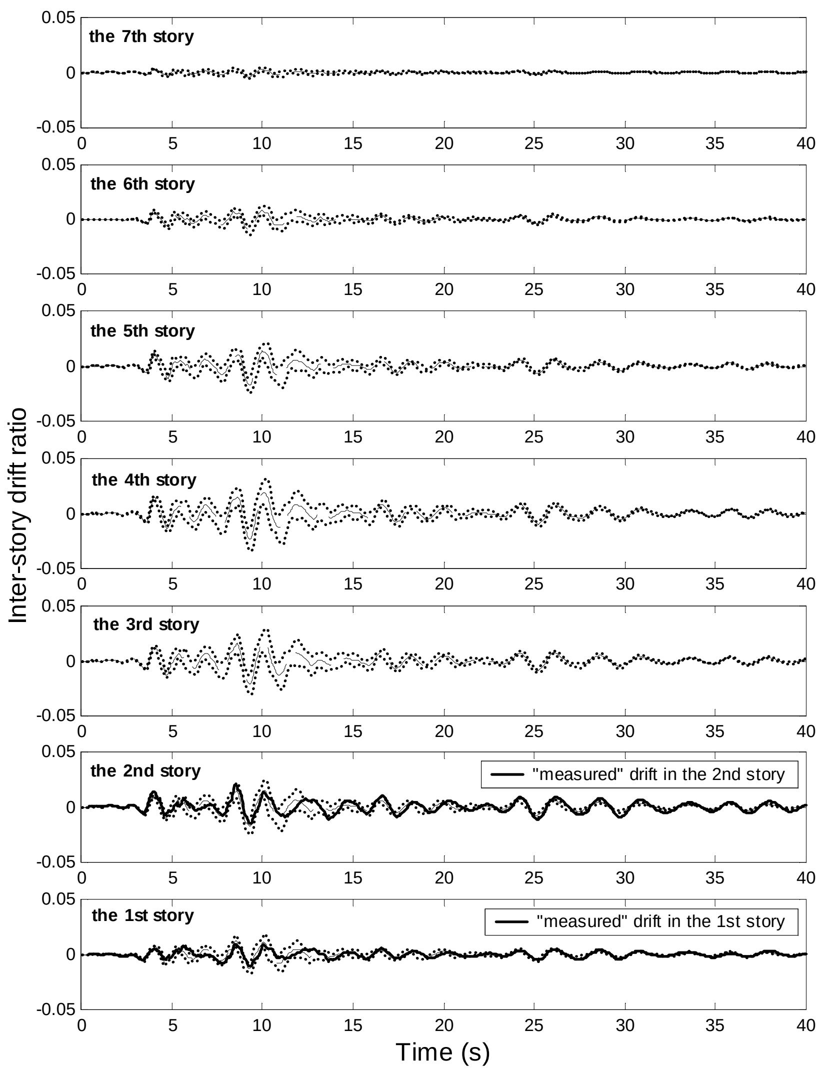

Figure 2. Solid lines are “measured” inter-story drift ratios; dashed lines are predicted inter-story drift ratios; and dotted lines are 95% confidence intervals

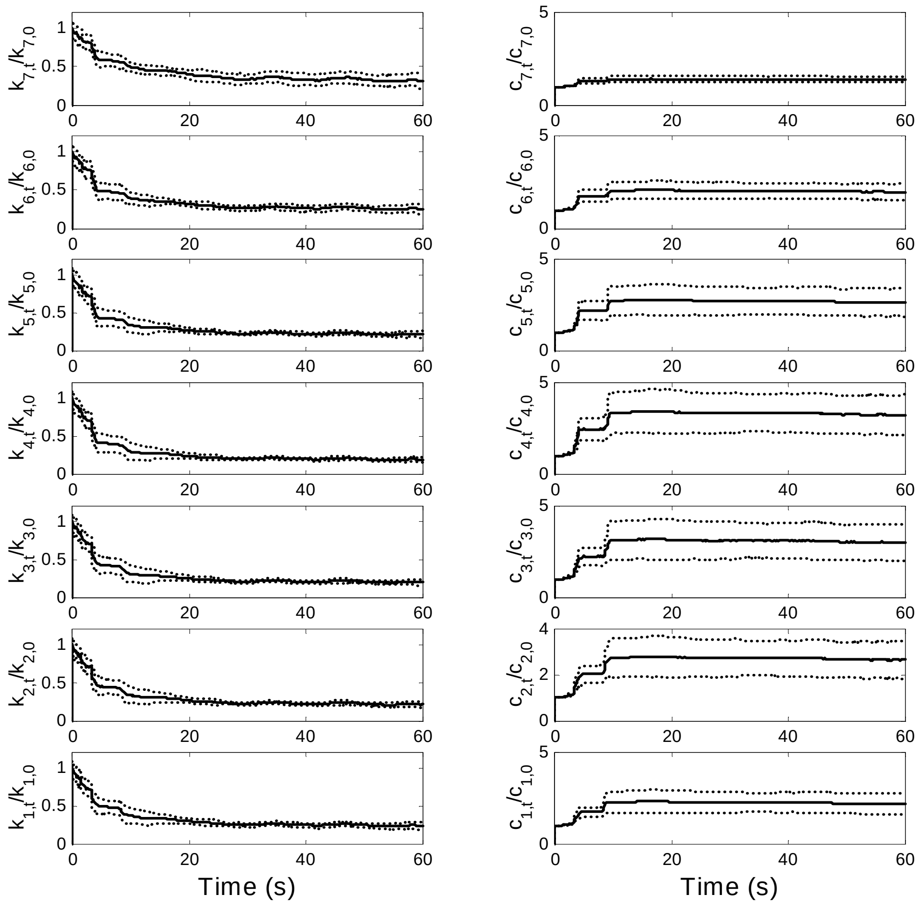

Figure 3. The estimated stiffnesses (left) and dampings (right); the dotted lines are the 95% confidence intervals

There are many possible performance measures of interests. In this paper, we consider the first two moments of the repair costs of the monitored structure (C_R) and the probability of a damage state of a specified damageable assembly. We describe in Ching [17] how to compute the first two moments of the repair costs and probability of damage conditioned on engineering demand parameters (*EDP*), e.g. maximum inter-story drifts, given probability distributions for capacity, unit repair cost and contractor cost uncertainties, by using the assembly-based vulnerability (ABV) approach (Porter [22]). In this paper, we only consider the repair costs and probability of damage of the non-structural components, e.g. windows, drywall finish, drywall partitions, and stucco, of the Van Nuys hotel, and the corresponding *EDP*s are the inter-story drift ratios. There are 41 damageable non-structural assemblies in this hotel, including 6 assemblies of windows (1 assembly per story from 2nd to 7th stories), 14 assemblies of drywall finish (2 assemblies per story from 1st to 7th stories), 14 assemblies of drywall partitions (2 assemblies per story from 1st to 7th stories), and 7 assemblies of stuccos (1 assemblies per story from 1st to 7th stories).

With the results from the PF approach using the simplified nonlinear degradation model, we obtain 2,000 (weighted) samples of system states at each time instant, i.e. 2,000 weighted samples of the first two moments of C_R and probability of damage of the 5th–story partitions (as an example). Figure **4** shows the real-time estimate of the repair costs for the non-structural assemblies. It is clear that there are two time instants, around 3.5 sec and 9 sec, corresponding to a rapid repair cost increase. The c.o.v. for the repair costs is about 15%. Figure **5** shows the real-time estimate of the probability that the drywall partitions in the 5th story are in one of the two damage states (1: visible damage; 2: significant damage).

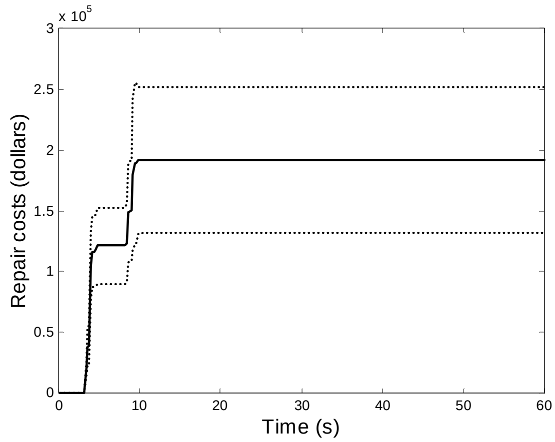

Figure 4. Real-time estimate of the repair costs for the non-structural assemblies in the Van Nuys hotel during the 1994 Northridge earthquake; dotted lines are 95% confidence intervals

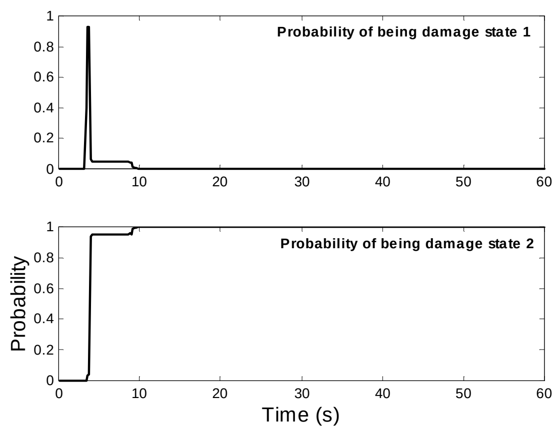

Figure 5. Real-time estimate of the probability for one of the damageable assemblies (drywall partitions in the 5th story) to be in its two damage states (1: visible damage; 2: significant damage) during the 1994 Northridge earthquake

CONCLUSIONS

We have presented two real-time Bayesian state-estimation algorithms in detail, including the extended Kalman filter (EKF) and the particle filter (PF), which is basically a Monte Carlo simulation approach. The Van Nuys hotel is studied using the EKF and PF algorithms, and the results show that EKF can sometimes create misleading results. We further apply the PF samples to compute real-time estimates of the repair costs and the probability of damage of an assembly of the Van Nuys hotel during the 1994 Northridge earthquake.

REFERENCES

1. Beck JL. *Determining Models of Structures from Earthquake Records*, EERL Report 78-01, Earthquake Engineering Research Laboratory, California Institute of Technology, Pasadena, California, 1978. <http://resolver.caltech.edu/CaltechEERL:1978.EERL-78-01>
2. Yun CB, Shinozuka M. "Identification of nonlinear dynamic systems." *Journal of Structural Engineering*, **8**(2), 187-203, 1980.

3. Hoshiya M, Saito E. “Structural identification by extended Kalman filter.” *Journal of Engineering Mechanics*, **110**(12), p. 1757, 1984.
4. Lin CC, Soong TT, Natke HG. “Real-time system identification of degrading structures.” *Journal of Engineering Mechanics*, **116**(10), 2258-2274, 1990.
5. Ghanem R, Shinozuka M. “Structural-system identification. I: Theory.” *Journal of Engineering Mechanics*, **121**(2), 255-264, 1995.
6. Shinozuka M, Ghanem R. “Structural-system identification. II: Experimental verification.” *Journal of Engineering Mechanics*, **121**(2), 265-273, 1995.
7. Glaser SD. “Insight into liquefaction by system identification.” *Geotechnique*, 46(4), 641-655, 1996.
8. Sato T, Qi K. “Adaptive H_∞ filter: Its application to structural identification.” *Journal of Engineering Mechanics*, **124**(11), 1233-1240, 1998.
9. Smyth AW, Masri SF, Chassiakos AG, Caughey TK. “On-line parametric identification of MDOF nonlinear hysteretic systems.” *Journal of Engineering Mechanics*, **125**(2), 133-142, 1999.
10. Kalman RE. “A new approach to linear filtering and prediction problems.” *J. Basic Engr.*, **82**D, 35-45, 1960.
11. Jazwinski AH. *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.
12. Gordon NJ, Salmond DJ, Smith AFM. “Novel approach to nonlinear/non-Gaussian Bayesian state estimation.” *IEEE Proceedings-F*, **140**(2): 107-113, 1993.
13. Kitagawa G. “Monte Carlo filter and smoother for non-Gaussian nonlinear state space models.” *Journal of Computational and Graphical Statistics*, **5**, 1-25, 1996.
14. Doucet A, de Freitas JFG, Gordon N. “Introduction to sequential Monte Carlo methods.” in *Sequential Monte Carlo Methods in Practice*, Doucet, A., de Freitas, J.F.G., and Gordon, N.J. (eds), Springer-Verlag, 2000.
15. van der Merwe R, de Freitas N, Doucet A, Wan EA. *The Unscented Particle Filter*, Report CUED/F-INFENG/TR380, Department of Engineering, Cambridge University, 2000.
16. Doucet A, Andrieu C. *Particle Filtering for Partially Observed Gaussian State Space Models*, Report CUED/F-INFENG /TR393, Department of Engineering, Cambridge University, 2000.
17. Ching J, Beck JL, Porter KA, Shaikhutdinov R. *Real-time Bayesian State Estimation of Uncertain Dynamical Systems*. EERL Report 2004-01, California Institute of Technology, Pasadena, CA. <http://caltecheerl.library.caltech.edu/>
18. Liu JS, Chen R. “Sequential Monte Carlo methods for dynamical systems.” *Journal of the American Statistical Association*, **93**, 1032-1044, 1998.
19. Jaynes ET. Information theory and statistical mechanics, *Physical Review*, **106**, 620-630, 1957. Also in Rosenkrantz RD (Ed) (1982). *E. T. Jaynes, Papers on Probability, Statistics and Statistical Physics*, Dordrecht, Holland: D. Reidel Publishing Co.
20. Jennings PC. *Engineering Features of the San Fernando Earthquake of February 9, 1971*, EERL Report 71-02, California Institute of Technology, Pasadena, CA.
21. Beck JL, Porter KA, Shaikhutdinov R, Au SK, Moroi T, Tsukada Y, Masuda M. *Impact of Seismic Risk on Lifetime Property Values, Final Report, CUREe-Kajima Joint Research Program Phase IV*, Consortium of Universities for Research in Earthquake Engineering, Richmond, CA, 2002. <http://resolver.caltech.edu/CaltechEERL: 2002.EERL-2002-04>
22. Porter KA, Kiremidjian AS, LeGrue JS. “Assembly-based vulnerability of buildings and its use in performance evaluation.” *Earthquake Spectra*, **17** (2), 291-312, 2001.