SPLIT-GAUSSIAN PARTICLE FILTER

Juho Kokkala and Simo Särkkä

Aalto University
Espoo, Finland

ABSTRACT

This paper is concerned with the use of split-Gaussian importance distributions in sequential importance resampling based particle filtering. We present novel particle filtering algorithms using the split-Gaussian importance distributions and compare their performance with several alternatives. Using a univariate nonlinear reference model, we compare the performance off the importance distributions by monitoring the effective number of particles. When using adaptive resampling, the split-Gaussian approximation has the best performance, and the Laplace approximation performs better than importance distributions based on unscented and extended Kalman filters. In addition, we also consider a two-dimensional target-tracking example where the Laplace approximation is not available in closed form and propose fitting the split-Gaussian importance distribution starting from an unscented Kalman filter based approximation.

Index Terms— split-normal distribution, split-Gaussian distribution, particle filter, importance distribution

1. INTRODUCTION

This paper is concerned with particle filtering in discrete-time probabilistic state-space models (see, e.g., [1, 2, 3, 4]) of the form

$$\begin{aligned} \mathbf{x}_k &\sim p(\mathbf{x}_k \mid \mathbf{x}_{k-1}), \\ \mathbf{y}_k &\sim p(\mathbf{y}_k \mid \mathbf{x}_k). \end{aligned} \tag{1}$$

The model describes the dynamics of the state $\mathbf{x}_k \in \mathbb{R}^n$ via the probability distribution for the current state given the state at the previous time step, and the observations $\mathbf{y}_k \in \mathbb{R}^d$ via the probability distribution for the measurement given the current state. In the Bayesian filtering problem (see, e.g., [4]), we desire to estimate the (filtering) posterior distribution of the state $\mathbf{x}_k$ given the current and previous measurements $(\mathbf{y}_1, \ldots, \mathbf{y}_k)$, that is, $p(\mathbf{x}_k \mid \mathbf{y}_{1:k})$. If the dynamic and measurement models are linear with Gaussian noise, the filtering distribution can be exactly computed using the Kalman filter. However, in the general case one must resort to approximations such as the particle filter algorithm considered in this paper.

In sequential importance resampling based particle filtering [1, 2, 3, 4], the posterior distribution of the state is approximated by a discrete set of samples (called particles) and weights associated to them. This particle set is updated sequentially such that after each measurement, a new value for each particle i is sampled from an importance distribution $\pi(\mathbf{x}_k \mid \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_{1:k})$ and the weights are updated to represent the new posterior distribution of the state.

The optimal importance distribution with respect to the weight variance would be the distribution of the current state given the previous state and the new measurement [1]. However, using the optimal importance distribution may not be feasible. Alternatives suggested in literature are Gaussian approximations, for example, the Laplace approximation or extended or unscented Kalman filter based approximations [1, 5, 2].

Geweke [6] has suggested a split-Gaussian importance distribution for (nonsequential) importance sampling. The split-Gaussian distribution is a Gaussian distribution that is scaled with different scaling factors in different directions away from the mode. The scaling factors are selected to match the behavior of the target distribution. In this article, we propose the split-Gaussian importance distribution for particle filtering and compare its performance with alternative importance distributions in terms of the effective number of particles [7] and tracking error.

The use of split-Gaussian distributions in the sequential setting has also been considered before by Guo and Wang [8], who used the split-Gaussian distribution in their sequential quasi-Monte Carlo algorithm. However, their algorithm is different in that they construct a global proposal distribution for all points based on the points from the previous step, whereas in the particle filtering approach we consider here, a separate proposal distri-

We were supported by the Academy of Finland projects 266940 and 273475. We acknowledge the computational resources provided by the Aalto Science-IT project. We thank Aki Vehtari and Enrique Lelo de Larrea Andrade for introducing the split-Gaussian distribution to us and Arno Solin for helpful comments.

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

bution is used for each particle. To our knowledge, the split-Gaussian distribution has not been proposed for the present type particle filters before. Recently, we [9] have also considered split-Gaussian importance distributions for particle filters in a certain class of Poisson regression models. In that work, we motivated the split-Gaussian distribution by convergence analysis and proposed selecting the scaling factor so that certain convergence criteria are fulfilled. The present work is different in that here we use the Geweke-style scaling aiming to match the optimal importance distribution.

In the present paper, we also consider fitting the split-Gaussian importance distribution based on other Gaussian approximations in case the Laplace approximation is not analytically available. For this case, we propose using the grid used to select the scaling factors also to improve the mode of the distribution.

2. PARTICLE FILTERS

In particle filtering, the filtering distribution $p(\mathbf{x}_k \mid \mathbf{y}_{1:k})$ is approximated by a weighted set of particles $\{(w_k^{(i)}, \mathbf{x}_k^{(i)}) : i = 1, \ldots, N\}$. At each measurement $\mathbf{y}_k$, the particle set is updated sequentially as follows:

1. For each $i = 1, \ldots, N$, draw a new point $\mathbf{x}_k^{(i)}$ from the importance distribution $\mathbf{x}_k^{(i)} \sim \pi(\mathbf{x}_k \mid \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_{1:k})$.
2. For each $i = 1, \ldots, N$, update the weights:

$$w_k^{(i)} = w_{k-1}^{(i)} \frac{p(\mathbf{y}_k \mid \mathbf{x}_k^{(i)})\, p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})}{\pi(\mathbf{x}_k \mid \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_{1:k})}.$$

3. Normalize the weights w_k to sum to unity.
4. If the particle set is too degenerate, do resampling.

The quality of the approximation depends on the variance of the weights and therefore on the effective number of particles [7]

$$\mathrm{ESS}(w_k) = 1/\sum_i (w_k^{(i)})^2, \tag{2}$$

which is often used as a measure of the particle degeneracy. A typical resampling condition in step 4 of the above algorithm is a threshold on the effective number of particles.

Thus, in particle filtering, one desires to maximize the effective number of particles, or equivalently, minimize the variance of the weights. The optimal importance distribution in this sense would be $\pi(\mathbf{x}_k \mid \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_{1:k}) = p(\mathbf{x}_k \mid \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k)$ [1]. However, sampling from it is often infeasible and the analytical form of the density might not be known. Therefore, one usually selects the importance distribution based on some approximation of the optimal importance distribution. Alternatives suggested in the literature typically use approximations based on the extended Kalman filter [1] or the unscented Kalman filter [5], more general non-linear Kalman filters [4], or the Laplace approximation [2]. In the so-called bootstrap filter [10], the dynamic model $p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$ is directly used as the importance distribution.

The selection of the importance distribution can also have an effect on the convergence of the particle filter in the limit $N \to \infty$. In particular, many convergence results require the weights to be bounded [11]. In practice, this can usually be guaranteed by selecting the importance distribution so that it has heavier tails than the optimal distribution. Using the Student's t-distribution instead of a Gaussian distribution has been suggested for this purpose [2].

3. SPLIT-GAUSSIAN IMPORTANCE DISTRIBUTION

Geweke [6] proposed using the split-Gaussian as the importance distribution in importance sampling. In this section, we describe how to use it in particle filtering. Note that we use a continuous-density version of the distribution (e.g. [12]), whereas Geweke used a formulation that has a discontinuity at the mode.

In the one-dimensional case, the split-Gaussian distribution is defined in terms of the mean μ and variance σ^2 of an underlying Gaussian random variable as well as scaling factors q and r (e.g. [8]). In the multidimensional case, one starts from mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$ of the underlying Gaussian random variable. The split-Gaussian distribution is constructed by defining scaling factors $\mathbf{q}$ and $\mathbf{r}$ for each principal component of $\boldsymbol{\Sigma}$, that is, each column of a matrix $\mathbf{T}$ such that $\boldsymbol{\Sigma} = \mathbf{T}\mathbf{T}^\mathsf{T}$. Then, a standard multivariate Gaussian $\boldsymbol{\epsilon}$ is drawn and each component of $\boldsymbol{\epsilon}$ is scaled by the scaling factors $\mathbf{q}_i, \mathbf{r}_i$ to form a split-Gaussian variable $\boldsymbol{\eta}$. The variable $\mathbf{x}$ is then obtained by the transformation $\mathbf{x} = \boldsymbol{\mu} + \mathbf{T}\boldsymbol{\eta}$. The density is

$$\mathrm{SN}(\mathbf{x} \mid \boldsymbol{\mu}, \mathbf{T}, \mathbf{q}, \mathbf{r}) = \sqrt{\frac{2^n}{\pi^n}} \frac{1}{|\mathbf{T}| \prod (\mathbf{q}_i + \mathbf{r}_i)} e^{-\frac{1}{2}\boldsymbol{\epsilon}^\mathsf{T}\boldsymbol{\epsilon}}. \tag{3}$$

Note that above, $\mathbf{x}$ does not explicitly appear on the right hand side. Instead, the expression is evaluated by substituting the auxiliary variable $\boldsymbol{\epsilon}$. An expression for the density containing $\mathbf{x}$ explicitly would contain 2^n different cases and therefore is omitted here. Pseudocode for drawing a multivariate split-Gaussian random variable and computing the corresponding density is given in Algorithm 1.

In the particle filter we now aim to approximate the optimal importance distribution as a split-Gaussian distribution. Following Geweke [6], we fit the parameters

function DRAWSPLITGAUSSIAN($\boldsymbol{\mu}, \mathbf{T}, \mathbf{q}, \mathbf{r}$)
  Draw $\boldsymbol{\epsilon} \sim \mathrm{N}(0, \mathbf{I}_n)$
  for $i \in \{1, \ldots, n\}$ **do**
    Draw $u_i \sim \mathrm{Uniform}(0, 1)$
    if $u_i < \mathbf{q}_i/(\mathbf{r}_i + \mathbf{q}_i)$ **then**
      $\boldsymbol{\eta}_i \leftarrow \mathbf{q}_i \, |\boldsymbol{\epsilon}_i|$
    else
      $\boldsymbol{\eta}_i \leftarrow -\mathbf{r}_i \, |\boldsymbol{\epsilon}_i|$
    end if
  end for
  $\mathbf{x} \leftarrow \boldsymbol{\mu} + \mathbf{T}\,\boldsymbol{\eta}$
  density $\leftarrow \sqrt{\frac{2^n}{\pi^n}} \frac{1}{|\mathbf{T}| \prod(\mathbf{q}_i + \mathbf{r}_i)} e^{-\frac{1}{2}\boldsymbol{\epsilon}^\mathsf{T}\boldsymbol{\epsilon}}$
  return $\mathbf{x}$,density
end function

Algorithm 1: Sampling from a split-Gaussian distribution.

function FITSPLITGAUSSIAN($\mathbf{x}_{k-1}, \mathbf{y}_k$)
  $\phi(\mathbf{x}_k) := \log p(\mathbf{y}_k \mid \mathbf{x}_k)\, p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$
  $\boldsymbol{\mu} \leftarrow \arg\max_{\mathbf{x}} \phi(\mathbf{x})$
  $\Sigma \leftarrow -(\frac{\partial^2}{\partial \mathbf{x}^2}\phi(x))^{-1}$ ▷ Inverse Hessian
  $\mathbf{T} \leftarrow \mathbf{T}$ s.t. $\mathbf{T}\,\mathbf{T}^\mathsf{T} = \Sigma$ ▷ E.g., based on singular value decomposition
  for $i \in \{1, \ldots, n\}$ **do**
    for $\delta \in \{\delta_{\min}, \ldots, \delta_{\max}\}$ **do**
      $f_i(\delta) \leftarrow |\delta| \, (2\,(\phi(\boldsymbol{\mu}) - \phi(\boldsymbol{\mu} + \delta\,\mathbf{T}\,\mathbf{e}_i)))^{-1/2}$
    end for
    $\mathbf{q}_i \leftarrow \max_{\delta>0} f_i(\delta)$, $\mathbf{r}_i \leftarrow \max_{\delta<0} f_i(\delta)$
  end for
  return $\boldsymbol{\mu}, \mathbf{T}, \mathbf{q}, \mathbf{r}$
end function

Algorithm 2: Fitting the split-Gaussian approximation to the optimal importance distribution.

function SGPARTICLEFILTER(y,N)
  Draw $x_0^{(1)} \ldots x_0^{(N)}$ from the prior
  for $k \in \{1, \ldots, T\}$ **do**
    for $i \in \{1, \ldots, N\}$ **do**
      $\boldsymbol{\mu}, \mathbf{T}, \mathbf{q}, \mathbf{r} \leftarrow$ FitSplitGaussian($\mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k$)
      $\mathbf{x}_k^{(i)}, \pi \leftarrow$ DrawSplitGaussian($\boldsymbol{\mu}, \mathbf{T}, \mathbf{q}, \mathbf{r}$)
      $w_k^{(i)} \leftarrow w_{k-1}^{(i)} \frac{p(\mathbf{y}_k|\mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})}{\pi}$
    end for
    Normalize the weights to sum to unity
    if ESS(w_k) < threshold **then**
      $w_k, \mathbf{x}_k \leftarrow$Resample($w_k, \mathbf{x}_k$)
    end if
  end for
  return $\{\mathbf{x}_k^{(i)}, w_k^{(i)} \;:\; i = 1, \ldots, N, k = 1, \ldots, T\}$
end function

Algorithm 3: Split-Gaussian particle filter.

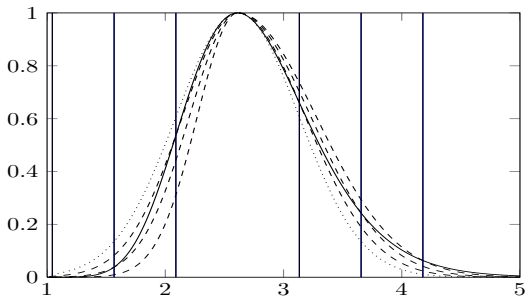

Fig. 1. Illustration of scaling the split-Gaussian approximation. Solid line is the density of Log-N($.2^2$). Dotted line is the Laplace approximation $\mathrm{N}(2.6, 0.51^2)$. Dashed lines are scaled half-Gaussians to match the decay at $\hat{\sigma}, 2\hat{\sigma}, 3\hat{\sigma}$ away from the mode. The final split-Gaussian approximation corresponds to the widest dashed lines in both directions. All densities scaled to have the same maximum.

of the split-Gaussian distribution as follows. We first explain the procedure for the one-dimensional case. First, μ and σ are set as the Laplace approximation of the target distribution. Then, the scaling factors q, r are selected as follows to improve the approximation away from the mode. The target distribution is evaluated in grid points away from the mode. Each grid point defines a candidate scaling factor based on matching the rate of decline of the target distribution to the rate of decline of the normal approximation,

$$f(\delta) = |\delta| \, (2\,(\log p(\mu) - \log p(\mu + \delta\,\sigma)))^{-1/2}, \quad (4)$$

where p is the target distribution, $\mu + \delta\,\sigma$ is the grid point and f is the candidate scaling factor. In each direction, we select the scaling factor that corresponds to the widest distribution,

$$q = \sup_{\delta>0} f(\delta), \qquad r = \sup_{\delta<0} f(\delta). \quad (5)$$

In the multidimensional version, the scaling (selection of q, r) is performed separately along each principal direction of the covariance matrix Σ. In more technical terms, we find $\mathbf{T}$ such that $\mathbf{T}\,\mathbf{T}^\mathsf{T} = \Sigma$, and then for each dimension i, $\delta\,\mathbf{T}\,\mathbf{e}_i$ takes the place of $\delta\,\sigma$ in Equation (5). In Algorithm 2, we give pseudocode for selecting the parameters as a function of the particle value in the previous step and the current measurement. The scaling procedure is illustrated in Figure 1 by showing how a split-Gaussian approximation to a lognormal distribution is constructed.

In the case where the Laplace approximation is not analytically available, one may simply start from some other Gaussian approximation such as the ones used in sigma-point filters (e.g., [4]). We propose a crude numeric algorithm for finding the mode of the optimal importance distribution to be used in conjunction of the

split-Gaussian fitting: if a higher density is found in the grid used for fitting the split-Gaussian distribution, simply restart the split-Gaussian procedure with that mode. We describe the split-Gaussian particle filtering algorithm in Algorithm 3.

4. NUMERICAL RESULTS

4.1. Univariate toy model

We consider the following commonly used toy-model (see, e.g., [1]):

$$\begin{aligned} x_k &\sim \mathrm{N}\left(\frac{1}{2}\,x_{k-1} + 25\frac{x_{k-1}}{1+x_{k-1}^2} + 8\cos(1.2k), Q\right), \\ y_k &\sim \mathrm{N}\left(0.05\,x_k^2, R\right). \end{aligned} \tag{6}$$

For the variances, we used $Q = 1, R = 0.05$, and the initial state was $x_0 = 0$. $1,000$ different datasets were simulated with $T = 25$, and the particle filter was run with $N = 1,000$ particles on all datasets. The following six importance distributions were used: the split-Gaussian approximation, the Laplace approximation, extended Kalman filter (EKF), unscented Kalman filter (UKF), and t-distribution scaled with the EKF/UKF variance.

First, we monitored the effective number of particles (ESS, Eq. 2) without resampling. The results are plotted over time in Figure 2. Second, we ran the particle filters with adaptive resampling, where resampling was performed every time the ESS dropped below $N/4$. The average numbers of resamplings in 25 steps are shown in Table 1.

The results show that the ESS drops quite fast with all tried importance distributions, but the split-Gaussian and Laplace importance distributions have the best performance. The EKF and UKF approximations lead to clearly lower ESS than the split-Gaussian and Laplace approximation, and the t-distributed versions of EKF/UKF perform even worse. Without resampling, the Laplace distribution has higher mean ESS than the split-Gaussian approximation after 4th measurement and later. However, the cumulative number of resamplings with the adaptive resampling threshold is the lowest with split-Gaussian. Furthermore, we also tried resampling after every step, in which case the split-Gaussian distribution had higher mean ESS than the Laplace distribution after most steps.

4.2. Target tracking

Here we consider a tracking experiment where the target location is two-dimensional and follows random walk,

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{q}_k, \tag{7}$$

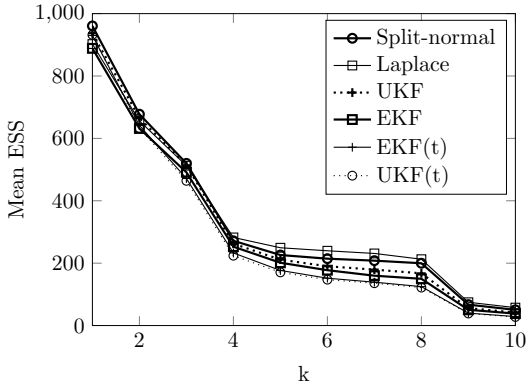

Fig. 2. Effective number of particles during the first 10 time steps using datasets simulated from the reference model. Average over $1,000$ datasets.

Distribution	Mean	95 % CI
Split-Gaussian	4.35	$[4.28, 4.42]$
Laplace	4.51	$[4.43, 4.59]$
UKF	4.96	$[4.88, 5.04]$
EKF	5.30	$[5.21, 5.38]$
UKF (t)	5.47	$[5.39, 5.54]$
EKF (t)	5.48	$[5.40, 5.55]$

Table 1. Number of resamplings required with adaptive resampling in 25 steps. Average over $1,000$ datasets.

$\mathbf{q}_k \sim \mathrm{N}(\mathbf{0}, \mathbf{Q})$, $\mathbf{x}_0 \sim \mathrm{N}(\mathbf{m}_0, \mathbf{P}_0)$. The parameters in this experiment where $\mathbf{Q} = 0.01\,\mathbf{I}_2$, $\mathbf{P}_0 = \mathbf{I}_2$, $\mathbf{m}_0 = \begin{pmatrix} 1 & 0 \end{pmatrix}^\mathsf{T}$. The measurement model consisted of a single sensor tracking both angle and range. The range measurement is assumed to be logarithm of the distance to the target with additive Gaussian noise. Thus, the measurement model is

$$\mathbf{y}_k = \begin{pmatrix} \operatorname{atan2}\left(\mathbf{x}_{k,2} - \mathbf{s}_2, \mathbf{x}_{k,1} - \mathbf{s}_1\right) \\ \log\sqrt{(\mathbf{x}_{k,1} - \mathbf{s}_1)^2 + (\mathbf{x}_{k,2} - \mathbf{s}_2)^2} \end{pmatrix} + \mathbf{r}_k, \tag{8}$$

where $\mathbf{s} = \begin{pmatrix} 0 & 0 \end{pmatrix}^\mathsf{T}$ is the sensor location and the sensor noise is $\mathbf{r}_k \sim \mathrm{N}(\mathbf{0}, \mathbf{R})$ where $\mathbf{R} = \operatorname{diag}(0.1^2,\ 0.01^2)$. In this experiment, we assumed that there is a measurement $\mathbf{y}_0$ associated with the first state.

We simulated $1,000$ datasets consisting of 100 measurements each and run the UKF particle filter and the UKF-based approximation in the split-Gaussian particle filter fitting stage. $1,000$ particles were used with resampling after every step. The performance of the filters was evaluated using the mean squared error of the estimated target location. The results are shown in Table 4.2. Based on median over the $1,000$ replications, the UKF importance distribution performed slightly better, but in some cases it had really poor performance, wherefore the mean performance of the split-Gaussian version is better.

	UKF	Split-Gaussian
Mean SSE	1.57×10^3	2.14
Median SSE	1.18	1.65
95th percentile	17.9	5.58

Table 2. Summary of the distribution of sum of squared tracking error. Mean, median and 95th percentile taken over the $1,000$ replications.

5. CONCLUSION AND DISCUSSION

In this work, we have studied the use of split-Gaussian distribution as an importance distribution for particle filtering. We compared its performance to alternatives in two numeric experiments: a unidimensional toy model as well as a two-dimensional target-tracking example.

In the unidimensional toy model, we compared the split-Gaussian importance distribution to the Laplace approximation and importance distributions based on extended and unscented Kalman filters. We monitored the number of resamplings required when performing adaptive resampling based on an effective sample size threshold. Based on this measure, the split-Gaussian distribution performed better than Laplace approximation, which in turn was better than the extended/unscented Kalman filter approaches. The poor performance of the EKF/UKF based approximations is likely due to the fact that they do not capture the mode of the target distribution correctly, whereas the Laplace approximation is constructed to have the correct mode. However, the Laplace approximation is local approximation, based only on the behavior of the target distribution around its mode. The split-Gaussian approximation improves the Laplace approximation by scaling it according to global behavior of the target distribution. Using the t-distribution instead of Gaussian in the EKF/UKF based importance distribution lead to weaker performance in terms of effective number of particles.

A drawback of the Laplace approximation is that it requires finding the mode and Hessian of the log-density of the target distribution. For cases where they are not analytically available, we considered a modification to the split-Gaussian importance distribution. In this modification, one starts from some Gaussian approximation, and uses the same grid points to both find an approximation to the mode of the density and to compute the split-Gaussian scaling factors. In a two-dimensional tracking simulation, this type of split-Gaussian particle filter outperformed the UKF-based particle filter. However, more work is required to assess how much this explained by the split-Gaussian scaling versus simply finding a better mode for the importance distribution. Furthermore, there is the tradeoff in computational cost between performing the split-Gaussian procedure versus simply adding more particles.

REFERENCES

[1] Arnaud Doucet, Simon Godsill, and Christophe Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[2] Olivier Cappé, Eric Moulines, and Tobias Rydén, *Inference in Hidden Markov Models*, Springer, 2005.

[3] Olivier Cappé, Simon Godsill, and Eric Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proceedings of IEEE*, vol. 95, no. 5, 2007.

[4] Simo Särkkä, *Bayesian Filtering and Smoothing*, Cambridge University Press, 2013.

[5] Rudolph van der Merwe, Arnaud Doucet, Nando de Freitas, and Eric Wan, "The unscented particle filter," in *Proceedings of NIPS*, 2000, pp. 584–590.

[6] John Geweke, "Bayesian inference in econometric models using Monte Carlo integration," *Econometrica: Journal of the Econometric Society*, pp. 1317–1339, 1989.

[7] Jun S Liu and Rong Chen, "Blind deconvolution via sequential imputations," *Journal of the American Statistical Association*, vol. 90, no. 430, pp. 567–576, 1995.

[8] Dong Guo and Xiaodong Wang, "Quasi-Monte Carlo filtering in nonlinear dynamic systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2087–2098, 2006.

[9] Juho Kokkala and Simo Särkkä, "On the (non-)convergence of particle filters with Gaussian importance distributions," in *Proceedings of SYSID*, 2015, To appear.

[10] Neil Gordon, David Salmond, and Adrian Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEEE Proceedings F (Radar and Signal Processing)*, vol. 140, no. 2, pp. 107–113, 1993.

[11] Dan Crisan and Arnaud Doucet, "A survey of convergence results on particle filtering methods for practitioners," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 736–746, 2002.

[12] Mattias Villani and Rolf Larsson, "The multivariate split normal distribution and asymmetric principal components analysis," *Communications in Statistics: Theory & Methods*, vol. 35, no. 6, pp. 1123 – 1140, 2006.