



CHALMERS

Chalmers Publication Library

A Study of MAP Estimation Techniques for Nonlinear Filtering

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

15th international conference on information fusion, July 09-12 2012, Singapore

Citation for the published paper:

Fatemi, M. ; Svensson, L. ; Hammarstrand, L. (2012) "A Study of MAP Estimation Techniques for Nonlinear Filtering". 15th international conference on information fusion, July 09-12 2012, Singapore pp. 1058 - 1065.

Downloaded from: <http://publications.lib.chalmers.se/publication/163062>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

A Study of MAP Estimation Techniques for Nonlinear Filtering

Maryam Fatemi, Lennart Svensson and Lars Hammarstrand

Signals and Systems
Chalmers University of Technology
Gothenburg, Sweden

Email: {maryam.fatemi,lennart.svensson,lars.hammarstrand} @chalmers.se

Mark Morelande

Electrical and Electronic Engineering
University of Melbourne
Victoria, Australia

Email: m.morelande@ee.unimelb.edu.au

Abstract—For solving the nonlinear filtering problem, much attention has been paid to filters based on the Linear Minimum Mean Square Error (LMMSE) estimation. Accordingly, less attention has been paid to MAP estimation techniques in this field. We argue that, given the superior performance of the latter in certain situations, they deserve to be more carefully investigated. In this paper, we look at MAP estimation from optimization perspective. We present a new method that uses this technique for solving the nonlinear filtering problem and we take a look at two existing methods. Furthermore, we derive a new method to reduce the dimensionality of the optimization problem which helps decreasing the computational complexity of the algorithms. The performance of MAP estimation techniques is analyzed and compared to LMMSE filters. The results show that in the case of informative measurements, MAP estimation techniques have much better performance.

Keywords: Nonlinear Filtering, MAP estimation, LMMSE Estimation, Progressive Correction.

I. INTRODUCTION

The nonlinear filtering problem is to estimate the state of a dynamic system using noisy observations (measurements) made on the system. The system has a state-space description comprised of a process equation and a measurement equation. The purpose of filtering is to calculate the posterior density which has all the information about the state [1]. By choosing an appropriate optimality condition, the state can be estimated from this density. The solution includes two steps, prediction and measurement update. If the process equation or/and measurement equation are nonlinear the posterior will be non-Gaussian. We consider Gaussian filters where this density is approximated by a Gaussian density.

When solving nonlinear filtering problems, much effort has been put in finding estimates which are optimal in minimum mean square error (MMSE) sense. Specifically much attention has been paid to the class of Linear MMSE (LMMSE) estimators, such as, Extended Kalman Filter (EKF) [1], Unscented Kalman Filter (UKF) [1] and Cubature Kalman Filter (CKF) [2]. These filters try to improve the moment approximation of the Kalman Filter [3] for the nonlinear situation. As indicated in [4] improved approximation of Kalman filter does not necessarily result in improved approximation of the posterior density. Therefore, in cases where LMMSE estimation methods do not accurately capture the posterior moments, it could

be beneficial to consider the Maximum A Posteriori as the optimality condition. An example of such a situation would be, when the measurement is very informative. It should be noted that, we consider problems where the posterior density is unimodal. In these problems mode of the posterior is usually close to its mean.

This paper focuses on MAP estimation techniques for the nonlinear update step and argues why more attention needs to be paid to these techniques. We describe the condition under which these techniques outperform LMMSE methods. One way of calculating the MAP estimate in the nonlinear filtering problem is to perform an iterative optimization in the measurement update step [5]. Viewing the estimation problem from an optimization perspective can give us new insight and tools to improve estimation performance in certain situations.

Gauss-Newton is a common iterative optimization algorithm which, used to find the MAP estimate, results in the iterative extended Kalman filter (IEKF) [6] [5]. This method is sensitive to the choice of the initial point, i.e., if the initial point is far from the optimum, the convergence is not guaranteed [7]. To alleviate this problem we can adopt one of the following approaches: I)change the optimization algorithm to a one that better fits our problem or II)change our objective function to better fit the Gauss-Newton method. The first approach is developed in [8], where Levenberg-Marquardt replaces the Gauss-Newton method. In our paper we take another look at this method and investigate its cons and pros. Adopting the second approach, we derive a new method which uses the progressive correction idea proposed in [9] along with the Gauss-Newton method to compute the updated state, the result is PC-IEKF filter.

One disadvantage of using iterative optimization methods in the filtering algorithm is the resultant computational complexity. This problem becomes more evident as the dimension of the state vector increases. For solving this issue we propose a new method that reduces the dimensionality of the optimization problem, thus reducing the overall computational complexity of the filtering algorithm. Our method is suitable for the situations where the measurement model is not dependent on all the states.

Bearings-Only Tracking (BOT) and Range-Only Tracking (ROT) are two benchmark scenarios that we use to evaluate

the performance of MAP estimation methods and compare the result to that of LMMSE methods. Furthermore, we investigate how well the MAP methods can approximate the posterior density and compare their approximations to that of LMMSE's, i.e., Monte Carlo Kalman Filter (MCKF) [10]. The results indicate that when the measurement is very informative, MAP techniques perform better than LMMSE in terms of position estimation, moreover, we see that MAP techniques approximate the posterior more accurately.

The paper is structured as follows. Section II describes the nonlinear filtering problem where BOT and ROT are presented as two such problems. In section III we talk about MAP estimation techniques and compare them to LMMSE methods. The Existing algorithms IEKF, LM-IEKF are presented in section IV and our new proposed methods PC-IEKF and Dimension Reduction technique are described in section V. Section VI includes the simulation and results. Finally conclusion will be made in section VII.

II. PROBLEM STATEMENT

We assume that we have a nonlinear discrete-time model which is stated as

$$\begin{aligned}\mathbf{x}_k &= \mathbf{F}\mathbf{x}_{k-1} + \mathbf{v}_{k-1} \\ \mathbf{z}_k &= h(\mathbf{x}_k) + \mathbf{w}_k\end{aligned}\quad (1)$$

where \mathbf{x}_k is the state vector at time k . Process and measurement noise sequences are zero mean Gaussian which are denoted by \mathbf{v}_{k-1} , \mathbf{w}_k with covariances \mathbf{Q} and \mathbf{R} respectively. \mathbf{z}_k is the measurement at time k and \mathbf{Z}_k denotes all the measurements up to and including time k . Since we'd like to concentrate on the update step, we assume the process model to be linear throughout this paper, whereas the measurement model, $h(x)$, is assumed to be nonlinear.

The purpose of filtering is to recursively estimate the posterior density $p(\mathbf{x}_k|\mathbf{Z}_k)$. From this density we extract the unknown state \mathbf{x}_k under an appropriate optimality condition, i.e., MMSE or MAP. To benchmark the filtering solutions considered in this paper, we investigate two problems in this family, i.e., the BOT and the ROT problems.

A. Bearing Only Tracking (BOT)

The Bearings-Only Tracking is to estimate the position and the velocity of a moving target based on noisy observations of the bearing to the target at times t_1, \dots, t_k . The presented model is the one described in [9] and [11]. The target moves with constant velocity along a straight line. A maneuvering sensor (ownship) measures the bearing of the target. The target and the ownship state at time t_k are defined as

$$\mathbf{x}_k^t = [x_k^t, y_k^t, \dot{x}_k^t, \dot{y}_k^t]^T$$

and

$$\mathbf{x}_k^o = [x_k^o, y_k^o, \dot{x}_k^o, \dot{y}_k^o]^T$$

respectively, where the dot notation means differentiation with respect to time. The relative target state is

$$\begin{aligned}\mathbf{x}_k &= \mathbf{x}_k^t - \mathbf{x}_k^o \\ &= [x_k, y_k, \dot{x}_k, \dot{y}_k]^T\end{aligned}$$

and the process model is described as

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \omega_k + \mathbf{v}_k \quad (2)$$

where

$$\begin{aligned}\mathbf{F}_k &= \begin{pmatrix} 1 & t_k - t_{k-1} \\ 0 & 1 \end{pmatrix} \otimes I_2 \\ \omega_k &= \mathbf{F}_k \mathbf{x}_{k-1}^o - \mathbf{x}_k^o\end{aligned}$$

The measurement function is nonlinear

$$h(\mathbf{x}_k) = \arctan\left(\frac{x_k^t - x_k^o}{y_k^t - y_k^o}\right) \quad (3)$$

B. Range Only Tracking (ROT)

The Range-Only Tracking model that was used is the same as presented in [1] and [12]. The purpose is to measure range and range-rate of a target moving in a straight line with constant velocity. The ownship (an airborne observer) moves in a circular trajectory. The state vector is the same as in the BOT problem. The measurement vector at time t_k is

$$\mathbf{z}_k = [r_k \quad \dot{r}_k]^T$$

and the measurement function is described as

$$h(\mathbf{x}_k) = [h_r(\mathbf{x}_k) \quad h_{\dot{r}}(\mathbf{x}_k)]^T$$

Where

$$\begin{aligned}h_r(\mathbf{x}_k) &= \sqrt{x_k^2 + y_k^2} \\ h_{\dot{r}}(\mathbf{x}_k) &= \frac{x_k \dot{x}_k + y_k \dot{y}_k}{\sqrt{x_k^2 + y_k^2}}\end{aligned}$$

The process model is the same as the one given in (2), the only difference is that for the ROT problem we assume the process noise to be zero.

III. MAP ESTIMATION IN NONLINEAR FILTERING

In this section, first we look at using MAP estimation techniques to solve the nonlinear filtering problem. This is done by performing optimization in the update step. Second, we illustrate why these techniques are more accurate and robust than LMMSE methods in certain situations.

The LMMSE family of filters includes EKF, UKF and CKF. These filters have received considerable attention during past years and much research has been dedicated to improve their performance [1] [2] [5] [9]. However, there are some problems inherent in the assumptions upon which these algorithms are constructed, which make them an undesirable choice in certain situations. Considering this fact, we argue that more attention needs to be paid to MAP estimation techniques and their performance in comparison to LMMSE methods.

In the Bayesian framework, the solution to the nonlinear filtering problem comprises of two steps: prediction and update. We study Gaussian filters where the initial prior density

$p(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1})$, the prediction density $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$, and the posterior density $p(\mathbf{x}_k|\mathbf{Z}_k)$ are approximated by Gaussian densities:

$$\begin{aligned} p(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1}) &\approx N(\mathbf{x}_{k-1}; \mathbf{x}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}) \\ p(\mathbf{x}_k|\mathbf{Z}_{k-1}) &\approx N(\mathbf{x}_k; \mathbf{x}_{k|k-1}, \mathbf{P}_{k|k-1}) \\ p(\mathbf{x}_k|\mathbf{Z}_k) &\approx N(\mathbf{x}_k; \mathbf{x}_{k|k}, \mathbf{P}_{k|k}) \end{aligned} \quad (4)$$

Furthermore, having assumed the process model to be linear and the measurement model to be nonlinear, the prediction density is easily calculated. The posterior, however, is more difficult to calculate.

The purpose of the update step is to calculate the posterior density $p(\mathbf{x}_k|\mathbf{Z}_k)$. In a Gaussian filter, it suffices to calculate the posterior mean and covariance. Since true values of the moments are difficult to compute, we need to approximate them. A fair approximation of the posterior mean can be obtained using LMMSE or MAP estimators.

A. MAP estimation

As mentioned earlier, our focus is on the measurement update step. A MAP estimator calculates the posterior mean as

$$\hat{\mathbf{x}}_{k|k}^{MAP} = \arg \max_{\mathbf{x}_k} p(\mathbf{x}_k|\mathbf{Z}_k)$$

According to Bayes' rule, the posterior density can be written as follows

$$p(\mathbf{x}_k|\mathbf{Z}_k) \propto p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Z}_{k-1}) \quad (5)$$

where $p(\mathbf{z}_k|\mathbf{x}_k)$ and $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ are the likelihood and the prior, respectively. Using (4) and (5), we can write the posterior density as:

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{Z}_k) \propto \exp \left(-\frac{1}{2}((\mathbf{z}_k - h(\mathbf{x}_k))^T \mathbf{R}^{-1}(\mathbf{z}_k - h(\mathbf{x}_k)) \right. \\ \left. + (\mathbf{x}_k - \mathbf{x}_{k|k-1})^T \mathbf{P}_{k|k-1}^{-1}(\mathbf{x}_k - \mathbf{x}_{k|k-1})) \right) \end{aligned} \quad (6)$$

Maximizing $p(\mathbf{x}_k|\mathbf{Z}_k)$ is equivalent to minimizing its negative log which means that we have a nonlinear least-squares problem with the following objective function

$$\begin{aligned} L(\mathbf{x}) = \frac{1}{2} \left[(\mathbf{x} - \mathbf{x}_{k|k-1})^T \mathbf{P}_{k|k-1}^{-1}(\mathbf{x} - \mathbf{x}_{k|k-1}) \right. \\ \left. + (\mathbf{z}_k - h(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{z}_k - h(\mathbf{x})) \right] \end{aligned} \quad (7)$$

Iterative minimization of (7) yields an approximate MAP estimate $\hat{\mathbf{x}}_{k|k}^{MAP}$.

Once the iterations converge, we update the covariance $\mathbf{P}_{k|k}$. We use the same covariance update as many existing methods' such as EKF and IEKF which is

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_i \mathbf{H}_i) \mathbf{P}_{k|k-1} \quad (8)$$

where \mathbf{K}_i is the Kalman gain with respect to the last iterate's value $\mathbf{x}_i = \hat{\mathbf{x}}_{k|k}$ and \mathbf{H}_i is the jacobian of the measurement model

$$\begin{aligned} \mathbf{H}_i &= \nabla_{\mathbf{x}_i} h(\mathbf{x}) \\ \nabla_{\mathbf{x}_i} &= \left(\frac{\partial}{\partial(\mathbf{x}_i(1))} \quad \frac{\partial}{\partial(\mathbf{x}_i(2))} \quad \cdots \quad \frac{\partial}{\partial(\mathbf{x}_i(n_x))} \right) \\ \mathbf{x}_i &\in \mathcal{R}^{n_x \times 1} \end{aligned}$$

Bellaire et al. derived the same covariance update in [8] by using the first-order necessary condition for extremism. Furthermore, equation (8) has a close connection to Laplace approximation as explained in the Appendix.

B. MAP vs. LMMSE

In LMMSE methods the measurement prediction $\mathbf{z}_{k|k-1}$ is based on the predicted state $\mathbf{x}_{k|k-1}$. Therefore prediction errors compounded with additional errors due to measurement nonlinearities can cause undesirable errors [5]. LMMSE methods try to improve the moment approximation of the Kalman filter for nonlinear situation. As stated in [4], doing so does not necessarily lead to improvement in the calculation of the posterior. More specifically, when we have a poor prior and an informative measurement, LMMSE methods often fail to approximate the posterior accurately.

To illustrate this, we calculate the 3- σ ellipses for the true posterior and the estimated posterior resulting from both LMMSE and MAP estimation methods for the BOT problem defined in Section II-A. The measurement model for this problem is given in (3). True posterior samples are calculated with the Metropolis-Hastings algorithm [13] using 2×10^4 samples.

We consider two cases, in the first case, the prior mean and covariance are

$$\begin{aligned} (x_{1|0} \ y_{1|0} \ \dot{x}_{1|0} \ \dot{y}_{1|0})^T &= (5 \ 5 \ 0 \ 0)^T \\ \mathbf{P}_{1|0} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

The measurement is assumed to be far from the prior indicating the situation in which we have a poor prior and an informative measurement. The 3- σ ellipses of the prior, true posterior along with posterior approximation by different algorithms are depicted in Figure 1(a). As we can see, the MAP estimation method calculates the posterior more accurately than the LMMSE method. In case II, the measurement is further away from the prior. The prior mean and covariance are

$$\begin{aligned} (x_{1|0} \ y_{1|0} \ \dot{x}_{1|0} \ \dot{y}_{1|0})^T &= (5 \ 2 \ 0 \ 0)^T \\ \mathbf{P}_{1|0} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

This case is depicted in Figure 1(b), as we can see, the MAP estimation of the posterior density is superior to that

of LMMSE's. These results provide a good motivation to investigate the change of the optimality criteria from MMSE to MAP.

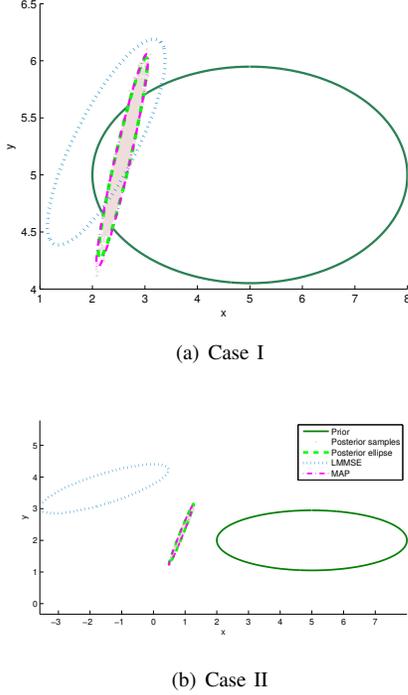


Figure 1. Comparison of posterior density approximation by MAP and LMMSE

IV. EXISTING MAP ESTIMATION TECHNIQUES

The idea of using MAP estimation to perform Gaussian filtering is not new and the use of Gauss-Newton and Levenberg-Marquardt methods has been mentioned in the literature [5] [6] [8] [14]. Properties of each method affect the behavior of the nonlinear filter. For example, while the Gauss-Newton method has possible quadratic convergence near the optimum and is easy to implement in the nonlinear filtering framework, it lacks robustness compared to the Levenberg-Marquardt method. On the other hand, the Levenberg-Marquardt method is more computationally complex than Gauss-Newton. In this section we present an overview of two filters, each of which use one of these two optimization methods in its update step.

A. IEKF

Newton's solution to our optimization problem is

$$\mathbf{x}_{i+1} = \mathbf{x}_i - (\nabla^2 L(\mathbf{x}_i))^{-1} \nabla L(\mathbf{x}_i) \quad (9)$$

where $\nabla^2 L(\mathbf{x}_i)$ is the Hessian of $L(\mathbf{x})$ and $\nabla L(\mathbf{x}_i)$ is its gradient. Following the Gauss-Newton method for finding the Hessian and the gradient and substituting them into (9) will result in the IEKF iterations [5] [6]

$$\mathbf{x}_{i+1} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_i(\mathbf{z}_k - h(\mathbf{x}_i) - \mathbf{H}_i(\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_i)) \quad (10)$$

where \mathbf{H}_i is the Jacobian of the measurement function $h(\mathbf{x})$ at \mathbf{x}_i and \mathbf{K}_i is the Kalman gain which is calculated according to

$$\mathbf{K}_i = \mathbf{P}_{k|k-1} \mathbf{H}_i^T (\mathbf{H}_i \mathbf{P}_{k|k-1} \mathbf{H}_i^T + \mathbf{R})^{-1}$$

Note that with $\mathbf{x}_1 = \hat{\mathbf{x}}_{k|k-1}$, the first iteration of IEKF is equivalent to EKF. After iterating the sequence (10) to convergence, $\hat{\mathbf{x}}_{k|k}$ is set to \mathbf{x}_i and the covariance is updated to $\mathbf{P}_{k|k}$ according to (8).

IEKF, being based on the Gauss-Newton method, approximates the measurement model as linear around \mathbf{x}_i . In cases where this approximation is poor, successive iterates can diverge [15]. Furthermore, when the initial point $\hat{\mathbf{x}}_{k|k-1}$ is far from the actual optimum e.g., when we have an informative measurement, convergence is not guaranteed [7]. In an attempt to overcome these problems, the Levenberg-Marquardt method can be used instead.

B. LM-IEKF

The Levenberg-Marquardt method is an interpolation between the Gauss-Newton method and the steepest descent. The iterations of the Levenberg-Marquardt method are calculated according to

$$\mathbf{x}_{i+1} = \mathbf{x}_i - (\nabla^2 L(\mathbf{x}_i) + \mu_i \mathbf{I})^{-1} \nabla L(\mathbf{x}_i), \quad (11)$$

where μ_i controls the behavior of the algorithm. When the current iterate is far from the optimal point, μ_i should have a large value causing the algorithm to behave similar to steepest descent. When \mathbf{x}_i is close to the optimum, μ_i should have a value close to zero which makes the algorithm behave like Gauss-Newton. Using the Levenberg-Marquardt algorithm for the nonlinear filtering problem, results in the following iterations [8]

$$\begin{aligned} \mathbf{x}_{i+1} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_i(\mathbf{z}_k - h(\mathbf{x}_i) - \mathbf{H}_i(\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_i)) \\ &\quad - \mu_i(\mathbf{I} - \mathbf{K}_i \mathbf{H}_i) \tilde{\mathbf{P}}_{k|k-1}(\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_i) \end{aligned} \quad (12)$$

Where \mathbf{K}_i and $\tilde{\mathbf{P}}_{k|k-1}$ are calculated according to

$$\begin{aligned} \mathbf{K}_i &= \tilde{\mathbf{P}}_{k|k-1} \mathbf{H}_i^T (\mathbf{H}_i \tilde{\mathbf{P}}_{k|k-1} \mathbf{H}_i^T + \mathbf{R})^{-1} \\ \tilde{\mathbf{P}}_{k|k-1} &= \left[\mathbf{I} - \mathbf{P}_{k|k-1} (\mathbf{P}_{k|k-1} + \frac{1}{\mu_i} \mathbf{I})^{-1} \right] \mathbf{P}_{k|k-1} \end{aligned}$$

After the sequence of (12) is iterated to convergence, $\hat{\mathbf{x}}_{k|k}$ is set to the last iteration's value \mathbf{x}_i , and the covariance is updated according to (8). Note that for covariance update $\mathbf{P}_{k|k-1}$ is used.

Although Levenberg-Marquardt is more robust than the Gauss-Newton method, it is more computationally complex and harder to implement.

V. NEW PROPOSED METHODS

As mentioned earlier, Gaussian filters that are based on MAP estimation boil down to solving an optimization problem. There exists a rich optimization literature in which we can find many tools for solving such optimization problems. However, existing methods mentioned in Section IV cover

only two such techniques. In this section we propose a new method (PC-IEKF) that uses a progressive correction technique to modify the objective function. This modification enables us to use the Gauss-Newton method while alleviating some of its disadvantages.

Furthermore, one disadvantage of MAP estimation methods is the added computational complexity due to the iterative optimization process. We derive a new method that reduces the dimensionality of the objective function, thus decreasing the computational complexity. This method is suitable for the situations where the measurement model does not depend on all the states, e.g., the BOT problem given in Section II-A.

A. PC-IEKF

The main idea behind PC-IEKF is to use a homotopy/continuation method to minimize the objective function (7). Homotopy methods are used to deal with difficult problems. The motivation behind such methods is to set up an easy problem and gradually transform this easy problem into the original difficult problem [16].

To interpret our optimization problem in this context we should take a closer look at (7) where the first term is a well behaved (twice differentiable) quadratic function. The second term on the other hand has an element ($h(\mathbf{x})$) which is nonlinear in \mathbf{x} and that is the term that can pose difficulties for the optimization algorithm. To set up an easy problem we use the idea of progressive correction where we split the update step into several steps [9]. We can partition the posterior at time k as:

$$p(\mathbf{x}_k | \mathbf{Z}_k) \propto p(\mathbf{x}_k | \mathbf{Z}_{k-1}) \prod_{i=1}^N p(\mathbf{z}_k | \mathbf{x}_k)^{\omega_i} \quad (13)$$

$$\sum_{i=1}^N \omega_i = 1$$

At the first iteration we have $p(\mathbf{x}_k | \mathbf{Z}_{k-1})p(\mathbf{z}_k | \mathbf{x}_k)^{\omega_1}$ which corresponds to the second term in (7) being multiplied by ω_1 , thus reducing its effect. We continue by finding the MAP estimate for

$$p(\mathbf{x}_k | \mathbf{Z}_{k-1})p(\mathbf{z}_k | \mathbf{x}_k)^{(\sum_{j=1}^i \omega_j)}$$

in each iteration which corresponds to the following objective function

$$L_i(\mathbf{x}) = \frac{1}{2} \left[(\mathbf{x} - \hat{\mathbf{x}}_{k|k-1})^T \mathbf{P}_{k|k-1}^{-1} (\mathbf{x} - \hat{\mathbf{x}}_{k|k-1}) \quad (14)$$

$$+ \left(\sum_{j=1}^i \omega_j \right) (\mathbf{z}_k - h(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{z}_k - h(\mathbf{x})) \right]$$

With this procedure, we start from a simple objective function and make it more difficult one step at a time so that at the last iteration, $L_N(\mathbf{x})$ is equal to the original problem given in (7). The solution to each iteration (subproblem) is used to initiate the next iteration. Hopefully, the initial point and the solution of each iteration are rather close in which case we expect the Gauss-Newton method to work well. It

should be noted that, to solve each subproblem we only run one iteration of the Gauss-Newton algorithm. Following this method, the iterations are calculated according to

$$\begin{aligned} \mathbf{x}_{i+1} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_i (\mathbf{z} - h(\mathbf{x}_i) - \mathbf{H}_i (\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_i)) \\ \mathbf{K}_i &= \mathbf{P}_{k|k-1} \mathbf{H}_i^T (\mathbf{H}_i \mathbf{P}_{k|k-1} \mathbf{H}_i^T + \mathbf{R}_i)^{-1} \\ \mathbf{R}_i &= \frac{\mathbf{R}}{\sum_{j=1}^i \omega_j} \end{aligned} \quad (15)$$

After N iterations, we set $\hat{\mathbf{x}}_{k|k} = \mathbf{x}_N$ and the covariance is updated similar to (8). It should be noted that \mathbf{R}_i given in (15) is different from the one calculated in [9] where $\mathbf{R}_i = \mathbf{R}/\omega_i$.

B. Dimension Reduction

The iterative optimization methods mentioned earlier increase the computational complexity of the filtering algorithms compared to the filters with non-iterative update step. This added complexity becomes more evident as the dimension of the state vector increases. Reducing the dimensionality of the original objective function given in (7) can, hence, decrease the computational complexity.

In many nonlinear filtering problems, measurement model does not depend on all the states. In this section we derive a new method that reduces the dimensionality of the objective function for such problems.

We start by partitioning the state vector as

$$\mathbf{x}_k = [\mathbf{a}^T, \mathbf{b}^T]^T$$

where $\mathbf{a} \in \mathbb{R}^{n_a}$, $\mathbf{b} \in \mathbb{R}^{n_b}$ and $n_a + n_b = n_x$, i.e., the dimension of the state vector. The partitioning is done in a way that we can write the measurement equation of (1) according to

$$\mathbf{z}_k = h(\mathbf{a}) + \mathbf{w}_k \quad (16)$$

Therefore, \mathbf{a} is part of the state vector that the measurement model depends on. Following the same partitioning, we define

$$\hat{\mathbf{x}}_{k|k-1} = [\hat{\mathbf{x}}_a, \hat{\mathbf{x}}_b]$$

Accordingly, we can rewrite the objective function defined in (7) as

$$L(\mathbf{a}, \mathbf{b}) = L_1(\mathbf{a}, \mathbf{b}) + L_2(\mathbf{a}) \quad (17)$$

where

$$\begin{aligned} L_1(\mathbf{a}, \mathbf{b}) &= \left(\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} - \begin{pmatrix} \hat{\mathbf{x}}_a \\ \hat{\mathbf{x}}_b \end{pmatrix} \right)^T \mathbf{P}_{k|k-1}^{-1} \left(\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} - \begin{pmatrix} \hat{\mathbf{x}}_a \\ \hat{\mathbf{x}}_b \end{pmatrix} \right) \\ L_2(\mathbf{a}) &= (\mathbf{z}_k - h(\mathbf{a}))^T \mathbf{R}^{-1} (\mathbf{z}_k - h(\mathbf{a})) \end{aligned} \quad (18)$$

Now we can redefine our minimization problem as

$$\begin{aligned} \min_{\mathbf{x}} L(\mathbf{x}) &= \min_{\mathbf{a}, \mathbf{b}} (L(\mathbf{a}, \mathbf{b})) \\ &= \min_{\mathbf{a}} (\min_{\mathbf{b}} (L_1(\mathbf{a}, \mathbf{b})) + L_2(\mathbf{a})) \end{aligned} \quad (19)$$

First, we find \mathbf{b}_{min} that minimizes $L_1(\mathbf{a}, \mathbf{b})$. This can be done analytically by putting the gradient to zero.

$$\nabla_{\mathbf{b}} L_1(\mathbf{a}, \mathbf{b}_{min}) = 0$$

For presenting the analytical solution of \mathbf{b} , we use the following block matrix definitions for $\mathbf{P}_{k|k-1}$ and its inverse

$$\mathbf{P}_{k|k-1}^{-1} = \begin{pmatrix} \mathbf{\Lambda}_{aa} & \mathbf{\Lambda}_{ab} \\ \mathbf{\Lambda}_{ba} & \mathbf{\Lambda}_{bb} \end{pmatrix}$$

$$\mathbf{P}_{k|k-1} = \begin{pmatrix} \mathbf{P}_{aa} & \mathbf{P}_{ab} \\ \mathbf{P}_{ba} & \mathbf{P}_{bb} \end{pmatrix}$$

Each block of $\mathbf{P}_{k|k-1}^{-1}$ has the following dimensions: $\mathbf{\Lambda}_{aa} \in \mathbb{R}^{n_a \times n_a}$, $\mathbf{\Lambda}_{ab} \in \mathbb{R}^{n_a \times n_b}$, $\mathbf{\Lambda}_{ba} \in \mathbb{R}^{n_b \times n_a}$, $\mathbf{\Lambda}_{bb} \in \mathbb{R}^{n_b \times n_b}$ and so does each corresponding block of $\mathbf{P}_{k|k-1}$. Considering these definitions, \mathbf{b}_{min} is calculated as

$$\mathbf{b}_{min} = \hat{\mathbf{x}}_b - \mathbf{\Lambda}_{bb}^{-1} \mathbf{\Lambda}_{ba} (\mathbf{a} - \hat{\mathbf{x}}_a) \quad (20)$$

By substituting \mathbf{b}_{min} into (18) we have

$$L_1(\mathbf{a}, \mathbf{b}_{min}) = (\mathbf{a} - \hat{\mathbf{x}}_a)^T \mathbf{P}_{aa}^{-1} (\mathbf{a} - \hat{\mathbf{x}}_a) \quad (21)$$

Equation (21) yields a very interesting result, i.e., that the solution to (19) can be found by minimizing $L(\mathbf{a})$

$$L(\mathbf{a}) = \frac{1}{2} \left[(\mathbf{a} - \hat{\mathbf{x}}_a)^T \mathbf{P}_{aa}^{-1} (\mathbf{a} - \hat{\mathbf{x}}_a) + (\mathbf{z}_k - h(\mathbf{a}))^T \mathbf{R}^{-1} (\mathbf{z}_k - h(\mathbf{a})) \right] \quad (22)$$

As we can see $L(\mathbf{a})$ in (22) is the objective function defined in (7) with lower dimensionality. Therefore, instead of performing the iterations for the whole state vector, we only need to iterate (solve) for the measurement dependent part using

$$\mathbf{a}_{min} = \arg \min_{\mathbf{a}} \left[(\mathbf{a} - \hat{\mathbf{x}}_a)^T \mathbf{P}_{aa}^{-1} (\mathbf{a} - \hat{\mathbf{x}}_a) + (\mathbf{z}_k - h(\mathbf{a}))^T \mathbf{R}^{-1} (\mathbf{z}_k - h(\mathbf{a})) \right] \quad (23)$$

After the iterations over \mathbf{a} converge, we substitute \mathbf{a}_{min} into (20) to calculate \mathbf{b}_{min} analytically. Finally the state update is defined as

$$\hat{\mathbf{x}}_{k|k} = [\mathbf{a}_{min}^T, \mathbf{b}_{min}^T]^T$$

This procedure reduces the dimensionality of the objective function which results in lower computational complexity of the whole filtering algorithm.

VI. SIMULATION AND RESULTS

In this section, we compare the performance of UKF, EKF, IEKF, PC-IEKF and LM-IEKF in BOT and ROT problems. We also depict the effect of using Dimension Reduction for LM-IEKF in the BOT problem. As a performance metric we use Root Mean Square Error (RMSE) of the position estimation in both BOT and ROT examples. The comparison is based on 1000 Monte Carlo simulation runs.

Let (x_k^m, y_k^m) and $(\hat{x}_k^m, \hat{y}_k^m)$ denote the true and estimated target position at time k respectively. Then the RMSE is defined as [1]

$$\text{RMSE}_k = \sqrt{\frac{1}{1000} \sum_{m=1}^M (x_k^m - \hat{x}_k^m)^2 + (y_k^m - \hat{y}_k^m)^2}$$

Weights (correction factors) of PC-IEKF filter are calculated according to [9]

$$\omega_i = \frac{10^{3(i-1)/11}}{\sum_{j=1}^{12} 10^{3(j-1)/11}}$$

A. BOT Results

The scenario which we used for testing the performance of the filters is the same as presented in [9] and [11]. It is a single sensor scenario where a target is moving away from the ownship on the course of 45° with the speed of 10 m/s. The ownship initial heading is towards North with the speed of $10/\sqrt{2}$ m/s. It measures the bearing of the target every 20 seconds per scan for total 48 scans. At scans 12 and 36 the ownship makes maneuvers and changes its heading to 90° and 0° respectively. This scenario is depicted in Figure 2. The standard deviation of the bearing is $\sigma_\theta = 0.45^\circ$. We ran this scenario with the initial target range of 1, 2.2 and 10 Km. All filters are initialized with initial range of 50 km. This scenario is used in [9] to show the weakness of EKF when the measurement is informative.

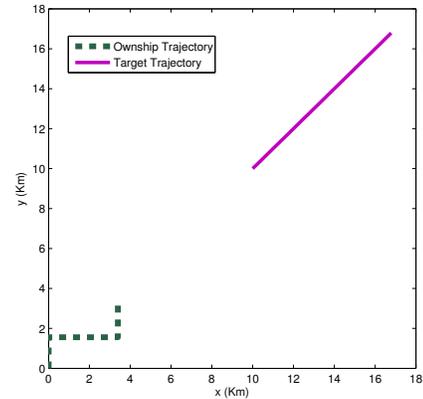
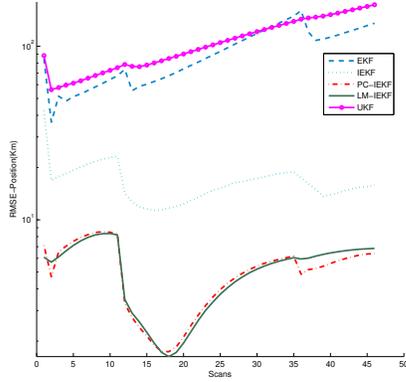


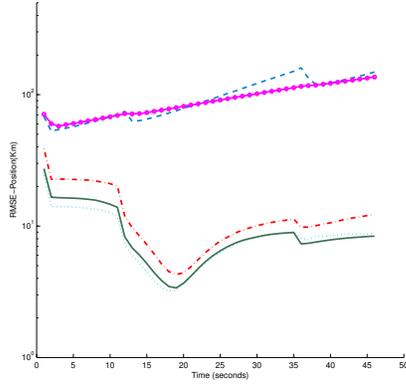
Figure 2. BOT scenario

The result is depicted in Figure 3. As we can see in all three cases MAP estimation techniques, IEKF, LM-IEKF and PC-IEKF perform much better than UKF and EKF. Among the MAP techniques, while IEKF does not perform very well in the first case, PC-IEKF and LM-IEKF perform well in all three cases which indicates their robustness compared to IEKF.

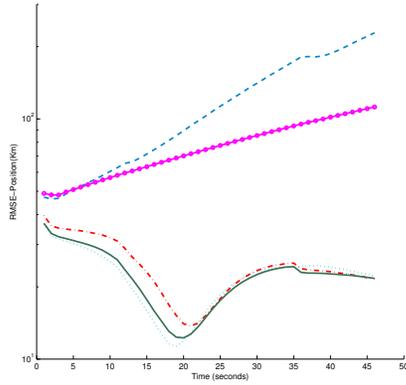
Since LM-IEKF is the most computationally complex among the filtering methods in this paper, we test the Dimension Reduction technique on this algorithm. We compare the computation time for different number of iterations for LM-IEKF with and without Dimension Reduction. The result is depicted in Figure 4. We can see that for the same number of iterations LM-IEKF with Dimension Reduction is 25% faster. It should be noted that the complexity decrease depends on the dimensions of \mathbf{a} and \mathbf{b} . It can also vary for each filtering algorithm.



(a) Target initial range is 1 Km



(b) Target initial range is 2.2 Km



(c) Target initial range is 10 Km

Figure 3. Position RMS error for the BOT problem

B. ROT Results

For the ROT problem we considered the following scenario which is presented in [1] and [12]. The target is moving on a course of 45° with a speed of 8 m/s. The ownship is moving along a circular trajectory with a radius of 15 Km at

a speed of 150 m/s. This scenario is depicted in Figure 5. The initialization of all filters is the same as mentioned in [1] and [12].

$$\begin{aligned} \mathbf{z}_0 &= (r_0 \ \theta_0)^T \\ \mathbf{x}_{0|0} &= (r_0 \sin(\theta_0) \ r_0 \cos(\theta_0) \ -\dot{x}_0^o \ -\dot{y}_0^o)^T \\ \mathbf{P}_{0|0} &= \begin{pmatrix} p_{11} & p_{12} & 0 & 0 \\ p_{21} & p_{22} & 0 & 0 \\ 0 & 0 & p_{33} & 0 \\ 0 & 0 & 0 & p_{44} \end{pmatrix} \end{aligned}$$

where

$$\begin{aligned} p_{11} &= r_0^2 \sigma_\theta^2 \cos^2(\theta_0) + \sigma_r^2 \sin^2(\theta_0) \\ p_{12} &= p_{21} = (\sigma_r^2 - r_0^2 \sigma_\theta^2) \sin(\theta_0) \cos(\theta_0) \\ p_{22} &= r_0^2 \sigma_\theta^2 \sin^2(\theta_0) + \sigma_r^2 \cos^2(\theta_0) \\ p_{33} &= p_{44} = \sigma_v^2 \end{aligned}$$

The RMS error is calculated for the case where $\sigma_v = 10$ m/s which is the practical scenario according to [1] and [12]. The result is depicted in Figure 6, similar to BOT cases, the MAP techniques outperform both UKF and EKF. Among them, PC-IEKF and LM-IEKF perform better than IEKF.

VII. CONCLUSION

In this paper, we discuss using MAP estimation techniques for solving the nonlinear filtering problem. We show why this is an attractive research topic whose potential is not fully explored. We look at two existing MAP estimation methods IEKF and LM-IEKF and we derive a new method PC-IEKF. Furthermore, we derive a new method to reduce the dimensionality of the optimization problem which decreases the computational complexity of the overall filtering algorithm. We compare the performance of the MAP based filters to UKF and EKF as two members of LMMSE family. We also compare how well the MAP estimation technique and the LMMSE method calculate the posterior density. Our results show that MAP estimation techniques have much more accurate calculation of the posterior density in the situation where the measurement is very informative. They also show that, for the same situation, MAP estimation techniques perform better than

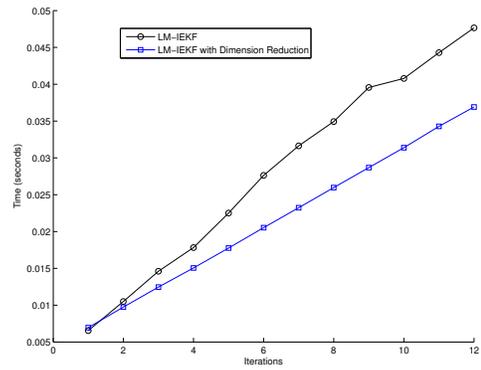


Figure 4. Comparison between the computation time for different number of iterations of LM-IEKF with and without Dimension Reduction

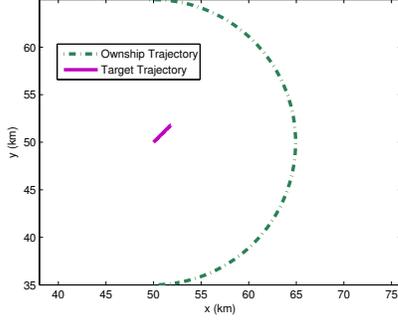


Figure 5. ROT scenario

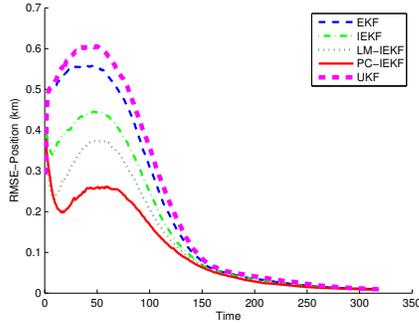


Figure 6. Position RMS error for the ROT problem

LMMSE in position estimation for BOT and ROT problems. Furthermore, the Dimension Reduction technique decreases the computational complexity of LM-IEKF implemented for the BOT problem by 25%.

APPENDIX

According to the laplace approximation [17], the inverse of the posterior covariance can be calculated at the posterior mode as

$$\mathbf{P}_{k|k}^{-1} = -\nabla_{\mathbf{x}_k}^T \nabla_{\mathbf{x}_k} \log(p(\mathbf{x}_k | \mathbf{Z}_k)) \Big|_{\hat{\mathbf{x}}_{k|k}^{MAP}} \quad (24)$$

which is the hessian of our objective function described in (7) at $\hat{\mathbf{x}}_{k|k}^{MAP}$. Therefore $\mathbf{P}_{k|k}^{-1}$ can be written as

$$\begin{aligned} \mathbf{P}_{k|k}^{-1} &= \nabla_{\mathbf{x}_k}^T \nabla_{\mathbf{x}_k} \{L(\mathbf{x}_k)\} \\ &= \mathbf{P}_{k|k-1}^{-1} \\ &\quad - \nabla_{\mathbf{x}_k}^T \{(\mathbf{z}_k - h(\mathbf{x}_k))^T \mathbf{R}^{-1} \nabla_{\mathbf{x}_k} h(\mathbf{x}_k)\} \end{aligned} \quad (25)$$

To derive the exact expression for $\mathbf{P}_{k|k}$, it remains to calculate the second term of equation (25). To continue the derivation first we assume the dimensions of the state and the measurement vector to be $\mathbf{x}_k \in \mathcal{R}^{n_x \times 1}$ and $\mathbf{z}_k \in \mathcal{R}^{n_z \times 1}$. If we write the jacobian of $h(\mathbf{x}_k)$ as

$$\begin{aligned} \mathbf{J}_h &= (\nabla_{\mathbf{x}_k} h(\mathbf{x}_k)) \\ &= (\mathbf{J}_h^1 \mathbf{J}_h^2 \dots \mathbf{J}_h^{n_x})_{n_z \times n_x} \end{aligned}$$

where each column of the jacobian matrix is denoted by $\mathbf{J}_h^i \in \mathcal{R}^{n_z \times 1}$. We can rewrite equation (25) as

$$\begin{aligned} \mathbf{P}_{k|k}^{-1} &= \mathbf{P}_{k|k-1}^{-1} + \mathbf{J}_h^T \mathbf{R}^{-1} \mathbf{J}_h \\ &\quad - (c^1 \ c^2 \ \dots \ c^{n_x}) \end{aligned}$$

where each column of the third term is

$$\begin{aligned} c^i &= \nabla_{\mathbf{x}_k}^T (\mathbf{J}_h^i) \mathbf{R}^{-1} (\mathbf{z}_k - h(\mathbf{x}_k)) \\ c^i &\in \mathcal{R}^{n_x \times 1} \end{aligned}$$

Since we expect $(\mathbf{z}_k - h(\mathbf{x}_k))$ to have a very small value near the posterior mode, we can neglect the third term. If we implement the matrix inversion lemma on the remaining two terms we will end up with the familiar covariance update of the EKF and IEKF stated in equation (8).

REFERENCES

- [1] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter, Particle Filter for Tracking Applications*. Artech House, 2004.
- [2] I. Arasaratnam and S. Haykin, "Cubature kalman filters," *IEEE Trans. Automatic Control*, vol. 54, no. 6, pp. 1254 – 1269, 2009.
- [3] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, pp. 35–45, 1960.
- [4] A. F. G. Fernandez, "Detection and tracking of multiple targets using wireless sensor networks," Ph.D. dissertation, Universidad Politecnica De Madrid, Madrid, Spain, 2011.
- [5] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. WILEY, 1995, pp. 402–437.
- [6] B. M. Bell and F. W. Cathey, "The iterated kalman filter update as a gauss-newton method," *IEEE Trans. Automatic Control*, vol. 38, no. 2, pp. 294–297, 1993.
- [7] D. G. Luenberger, *Linear and nonlinear programming*. Kluwer Academic Publishers, 2003.
- [8] R. L. Bellaire, E. W. Kamen, and S. M. Zabin, "A new iterated filter with applications to target tracking," in *SPIE Signal and Data Processing of Small Targets*, San Diego, CA, USA, 1995.
- [9] X. Wang, M. Morelande, and B. Moran, "Target motion analysis using single sensor bearings-only measurements," in *2nd International Congress on Image and Signal Processing*, Oct 2009, pp. 1–6.
- [10] A. J. Haug, "A tutorial on bayesian estimation and tracking techniques applicable to nonlinear and non-gaussian processes," *MITRE Technical Report, MTR 05W000004*. The MITRE Corporation, 2005.
- [11] N. Peach, "Bearings-only tracking using a set of range-parametrized extended kalman filters," in *Control Theory and Applications, IEEE Proceedings*, 1995.
- [12] B. Ristic, S. Arulampalam, and J. McCarthy, "Target motion analysis using range-only measurements: Algorithms, performance and application to isar data," *Elsevier Signal Processing Journal*, vol. 82, no. 2, pp. 273–296, 2002.
- [13] C. P. Robert, *The Bayesian Choice*. Springer, 2001.
- [14] G. Sibley, G. Sukhatme, and L. Matthies, "The iterated sigma point kalman filter with applications to long range stereo," in *Online Proc. 2nd Robotics: Science and Systems Conf.*, Philadelphia, PA, Aug 2006.
- [15] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.
- [16] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, 2006.
- [17] L. Tierney and J. B. Kadane, "Accurate approximations for posterior moments and marginal densities," *Journal of the American Statistical Association*, vol. 81, no. 393, 1986.