A Weighted Interacting Particle-based Nonlinear Filter

David J. Ballantyne, Surrey Kim, and Michael A. Kouritzin

MITACS-PINTS,
Department of Mathematics at the University of Alberta,
Edmonton, Canada

ABSTRACT

Particle-based nonlinear filters have proven to be effective and versatile methods for computing approximations to difficult filtering problems. We introduce a novel hybrid particle method, thought to possess an excellent compromise between the unadaptive nature of the weighted particle methods and the overly random resampling in classical interactive particle methods, and compare this new method to our previously introduced refining branching particle filter. Our experiments involve various fixed numbers of particles and compare computational efficiency of our new method to the incumbent. The hybrid method is demonstrated to outperform two previous particle filters on our simulated test problems.

To highlight the flexibility of particle filters, we choose to test our methods on a rectangularly-constrained Markov signal that does not satisfy a typical stochastic equation but rather a Skorohod, local time formulation. Whereas normal diffusive behavior occurs in the interior of the rectangular domain, immediate reflections are enforced at the boundary. The test problems involve a fish signal with boundary reflections and is motivated by the fish farming industry.

Keywords: target tracking, hybrid particle system, nonlinear filtering

1. INTRODUCTION

The basic requirements for a particle method are simulation of particles with the same law as the signal and the resampling of these particles to incorporate observation information effectively. Then, the high-particle limit of an empirical measure for the resulting particle system can be anticipated to exist and yield the conditional distribution of the signal at a particular time given the back observations. Beyond these requirements, the precise resampling procedure affects the capacity to construct path-space estimates and also the computational efficiency and filter performance for practical implementations with a fixed number of particles. For example, the "cautious" branching method discussed in Ballantyne, Chan, and Kouritzin[1] has been demonstrated to have path-space estimation capabilities as well superior performance and computational efficiency.

This paper introduces a novel *weighted-interacting* hybrid particle-based method that accepts the extent of resampling as a parameter. The test problem for our simulations involves tracking a fish signal with boundary reflections given a noisy observation sequence. Performance in fidelity and computational complexity of the hybrid filter is compared against the original interacting particle method of Del Moral and Salut[2] and the refining branching particle method.[1]

2. REFINING PARTICLE FILTERS

2.1. Background

The classical weighted particle filter does not adapt its particle locations based upon information supplied from the observations but instead only weights the particles in an observation-dependent manner. Typically, most particle evolutions would differ significantly from the real signal and in normal operation the method assigns significant

Further author information: (Send correspondence to D.J.B.)
D.J.B.: E-mail: dballant@math.ualberta.ca
S.K.: E-mail: surrey@math.ualberta.ca
M.A.K.: E-mail: mkouritz@math.ualberta.ca
PINTS web page: http://www.math.ualberta.ca/pints

Signal Processing, Sensor Fusion, and Target Recognition XI, Ivan Kadar, Editor,
Proceedings of SPIE Vol. 4729 (2002) © 2002 SPIE · 0277-786X/02/$15.00

weights to a very few particles. Consequently, fidelity of performance becomes reliant upon fewer and fewer particles, while computational resources are used by many insignificant particles.

To illustrate this and future considerations, we let $t \to X_t$ be a measurable Markov process estimated through observations $Y_k = h(X_{t_k}) + v_k$ corrupted by an independent noise sequence $\{v_k,\ k = 1, 2, 3, \dots\}$ that is also independent of X. Then, the weight of the j^{th} of N particles ξ^j at time t_k is $W_{1,k}(\xi^j)$, which is calculated for each particle based on the sequence of observations $Y_1, Y_2, \dots, Y_k$. Particles for which $W_{1,k}(\xi^j)$ is relatively small have virtually no effect on the approximate conditional probabilities

$$P\left(X_{t_k} \in A \middle| Y_1, \dots, Y_k\right) \approx \frac{1}{\sum_{j=1}^N W_{1,k}(\xi^j)} \sum_{j=1}^N W_{1,k}(\xi^j) 1_{\xi^j_{t_k} \in A}.$$

In the past decade, the importance of resampling particles and adapting them to the observations has become apparent. We loosely classify a filter as being an *adaptive grid filter* if the particle locations are modified in a reasonable manner to include the information from the observations. For example, in the *adaptive interactive filter* introduced by Del Moral and Salut,[2] the weights need not be stored but rather are used immediately to assign the particles to new sites. Suppose that the particles just prior to the k^{th} observation are denoted $\left\{\xi^1_{t_k-}, \xi^2_{t_k-}, \dots, \xi^N_{t_k-}\right\}$. Then, just after the observation, the i^{th} particle is assigned to the site $\xi^i_{t_k-}$ with probability $\frac{W_{k,k}(\xi^i)}{\sum_{j=1}^N W_{k,k}(\xi^j)}$ independent of the other assignments. The approximate conditional probabilities become

$$P\left(X_{t_k} \in A \middle| Y_1, \dots, Y_k\right) \approx \frac{1}{N} \sum_{j=1}^N 1_{\xi^j_{t_k} \in A}.$$

Whereas the previous method does not resample, this method adds unnecessary noise in the resampling process, thereby degrading performance. We can refer the interested reader to Del Moral, Kouritzin, and Miclo[3] for a detailed explanation of this deficiency. Another consequence of this overly random resampling is that there is no natural association of a particle at time t_k with one at time t_{k-1} and the historical empirical measure satisfies

$$\frac{1}{N} \sum_{j=1}^N 1_{(\xi^j_{t_1}, \xi^j_{t_2}, \dots, \xi^j_{t_k}) \in A_1 \times A_2 \times \cdots \times A_k} \overset{N \to \infty}{\to} P(X_{t_1} \in A_1 | Y_1) P(X_{t_2} \in A_2 | Y_1, Y_2) \cdots P(X_{t_k} \in A_k | Y_1, Y_2, \dots, Y_k),$$

meaning that it does not converge to $P(X_{t_1} \in A_1, X_{t_2} \in A_2, \dots, X_{t_k} \in A_k | Y_1, Y_2, \dots, Y_k)$.

We loosely define a *refining grid filter* as an adaptive grid filter such that the historical empirical measure converges to $P(X_{t_1} \in A_1, X_{t_2} \in A_2, \dots, X_{t_k} \in A_k | Y_1, Y_2, \dots, Y_k)$ as the number of particles increases and suggest that these are the adaptive grid particle filters that do not allow too much randomness in their resampling. Their performance at calculating approximate conditional probabilities $P(X_{t_k} \in A | Y_1, Y_2, \dots, Y_k)$ has been demonstrated to be superior.[1] It is important to note that by this definition a refining grid filter must first be an adaptive grid filter, since, while the historical empirical measure for the classical weighted particle method discussed first does converge to $P(X_{t_1} \in A_1, X_{t_2} \in A_2, \dots, X_{t_k} \in A_k | Y_1, Y_2, \dots, Y_k)$, it is not refining as it is not adaptive. To rephrase, we claim that the refining filters are the filters that undertake the correct amount of resampling.

2.2. Refining Branching Method

The large majority of researchers currently implement adaptive, non-refining filters that provide poorer performance than easily implementable refining filters. The popular weighted, branching, and interactive particle filters all have refining versions. Indeed, Del Moral, Kouritzin, and Miclo[3] introduced the refining version of the popular interacting method and showed that the performance of the refining variant was superior to the original, merely adaptive, one. Practitioners can now replace non-refining filters with refining versions.

To our knowledge, the first and hitherto the best refining filter is the refining branching particle filter introduced by Fleischmann and Kouritzin. (See also Ballantyne, Chan, and Kouritzin[1] for a discussion of it.) It provides improved performance over the previously introduced non-refining branching filter of Crisan and Lyons. The operation of the refining branching particle filter is as follows:

1. Initialize particles $\left\{\xi^1_0, \xi^2_0, \dots, \xi^N_0\right\}$ so that $\frac{1}{N}\sum_{i=1}^N \delta_{\xi^j_{t_k}}$ is a good approximation to the distribution of X_0.

2. Repeat for $k = 1, 2, \dots$

   (a) Evolve the particles forward independently of each other with a simulation technique designed to mimic the law of the signal evolution over this time interval. Call the new particles just prior to the next observation $\left\{\xi^1_{t_k-}, \xi^2_{t_k-}, \dots, \xi^N_{t_k-}\right\}$.

   (b) Repeat for $i = 1, 2, \dots, N$

      i. Collect the k^{th} observation and evaluate $\zeta^i_k = \frac{W_{k,k}(\xi^i)N}{\sum_j W_{k,k}(\xi^j)} - 1$, which is the desired branching functional for particle i.

      ii. Remove particle i with probability $-\zeta^i_k$ if $\zeta^i_k < 0$.

      iii. Otherwise, add $\lfloor \zeta^i_k \rfloor$ particles at the location $\xi^j_{t_k-}$. Then, add one more particle at this site with probability $\zeta^i_k - \lfloor \zeta^i_k \rfloor$.

   (c) Introduce more unbiased particle control to bring the number of particles back to N.

   (d) Relabel the resulting particles $\left\{\xi^1_{t_k}, \xi^2_{t_k}, \dots, \xi^N_{t_k}\right\}$.

The conditional distributions for this algorithm are calculated by

$$P\left(X_{t_k} \in A \middle| Y_1, \dots, Y_k\right) \approx \frac{1}{N} \sum_{j=1}^{N} 1_{\xi^j_{t_k} \in A}.$$

We did not clutter this description of the algorithm with all practical programming efficiencies. The historical estimates come extremely easily, bearing out our refining grid claim: all we have to do is keep the ancestral locations of each current particle. Then, in constructing the historical conditional distribution estimates, we weight earlier particle locations that are ancestors to several current particles by an integer multiple equal to the number of current particles of which it is an ancestor. Notice in this algorithm that most particles will neither be removed nor branched since values of ζ^i_k close to 0 are common with any appreciable amount of noise in the observations.

2.3. New Hybrid Method

Whereas to make the original interacting or branching particle filters refining we had to reduce the amount of resampling, to make the weighted particle system refining we must introduce an effective resampling procedure. We describe this in the current subsection and call the result the *weighted-interacting hybrid* method. However, it could just as well have been called the *refining weighted filter*.

A particle in the hybrid method is updated as a simple weighted particle $\left(\xi^i_{t_k}, W_{1,k}(\xi^i)\right)$ until such time its weight differs significantly from the majority of the particle weights. After each observation extremely weighted particles are resampled so that the range of weights (expressed as a ratio) is less than a specified parameter $\rho > 1$. There are a number of technical implementation details that can be employed to make this method very efficient.

3. FISH TRACKING PROBLEM

Motivated by the fish-farming industry, the test problem is the tracking of a single fish in a tank with boundary reflections.

3.1. Fish Motion Model

For simplicity we choose a 2-dimensional fish defined by the following Skorohod SDE (stochastic differential equation):

$$dX_t = -\alpha \left(X_t - \frac{L}{2} \right) dt + \beta dv_t + \chi_{\partial D}(X_t)\gamma(X_t)d\xi_t \tag{1}$$

where X_0 is a random variable, v is a standard Brownian motion in $\mathbb{R}^2$, $L = (L_1, L_2)^T$ is the size of the tank, $\gamma_i(x) = \sum_{j=1}^2 a_{ij}U_j(x), i = 1, 2$ is the conormal vector field with $a_{11} = a_{22} = \beta^2, a_{12} = a_{21} = 0$, and ξ is the local time for X at the boundary. In our simulations, we take α and β to be parameters and simplify our example by selecting $L_1 = L_2 = 1$.

3.2. Fish Observation Model

The observation process consists of a discrete sequence of images, arriving every $\epsilon = 1.0$ time units at observations times $\{t_k\}_{k=1}^{\infty}$, each observation being a two-dimensional raster $\{Y_{t_k}^{(i,j)}\}_{i,j=1,1}^{R,R}$. $Y_{t_k}^{(i,j)}$ is the (i,j) component of a raster depiction of the observation. We let $h^{(i,j)}(\cdot)$ be the indicator function

$$h^{(i,j)}(x_1, x_2) = 1_{[\frac{x_1 R}{L_1} - \frac{3}{2}, \frac{x_1 R}{L_1} + \frac{3}{2}] \times [\frac{x_2 R}{L_2} - \frac{3}{2}, \frac{x_2 R}{L_2} + \frac{3}{2}]}(i,j) \tag{2}$$

representing a 3×3 pixel square image of the fish and let $V_k^{(i,j)}$ be pixel-by-pixel zero-mean independent Gaussian noise. Then, an observation at time t_k is constructed by superimposing the square image of the signal onto the raster and adding noise according to the formula

$$Y_{t_k}^{(i,j)} = h^{(i,j)}(X_{t_k}) + V_k^{(i,j)}. \tag{3}$$

For our simulations, the size of the length and width of the observation rasters, R, is 128 and the standard deviation of $V_k^{(i,j)}$ is set as a parameter. Observations are not preprocessed; the information from the raster pixels is used directly in the filter algorithms.

3.3. Objective

The problem is to estimate the conditional distribution of the fish position based solely on the noisy observations and the initial conditions X_0, that is,

$$P(X_{t_k} \in dx | \sigma\{Y_{t_l}^{(i,j)}, 1 \le l \le k,\, 1 \le i,j \le 128\}). \tag{4}$$

4. HYBRID FILTER TECHNIQUE

4.1. Update Algorithm

The implementation of the hybrid method has an important computer science component. Considerations such as the data structure used can greatly reflect filter fidelity and efficiency. Simulations were developed and run on a DEC-ALPHA 3000/700 255 MHz system. The adaptive interacting and refining branching algorithms have been described previously. The basic refining hybrid algorithm proceeds as follows:

1. Initialize particles $\{\xi_0^1, \xi_0^2, \ldots, \xi_0^N\}$ uniformly over the domain $[0, L_1] \times [0, L_2]$, thus yielding $\frac{1}{N}\sum_{i=1}^{N} \delta_{\xi_{t_k}^j}$ as a good approximation to the assumed uniform distribution of X_0. Set $\widetilde{W}_0(\xi^j) = 1$, $\forall j = 1, \ldots, N$.
2. Repeat for $k = 1, 2, \ldots$
   (a) Evolve all particles over time interval ϵ using, for example, Euler approximations. Call the new particles just prior to an observation Y_{t_k} $\{\xi_{t_k-}^1, \xi_{t_k-}^2, \ldots, \xi_{t_k-}^N\}$.
   (b) Upon receiving the k^{th} observation, calculate the weight for all $\{\xi_{t_k-}^1, \xi_{t_k-}^2, \ldots, \xi_{t_k-}^N\}$ according to $\widetilde{W}_k(\xi^j) = W_{k,k}(\xi^j)\widetilde{W}_{k-1}(\xi^j)$.
   (c) Resample ξ^i until $\tilde{W}_k(\xi^i) < \rho \tilde{W}_k(\xi^j)$.
   (d) Relabel the resulting particles $\{\xi_{t_k}^1, \xi_{t_k}^2, \ldots, \xi_{t_k}^N\}$.
   (e) The conditional distribution for the fish location is approximated by

$$P\left(X_{t_k} \in A \middle| Y_1, \ldots, Y_k\right) \approx \frac{1}{\sum_{j=1}^{N} \widetilde{W}_k(\xi^j)} \sum_{j=1}^{N} \widetilde{W}_k(\xi^j) 1_{\xi_{t_k}^j \in A}$$

.

   (f) The position estimate for the fish is calculated as the conditional expectation

$$E[X_{t_k} |\, Y_{t_l}, 1 \le l \le k] \approx \frac{1}{\sum_{j=1}^{N} \widetilde{W}_k(\xi^j)} \sum_{j=1}^{N} \widetilde{W}_k(\xi^j) \xi_{t_k}^j.$$

4.2. Parameters N and ρ

The hybrid filter, in concordance with other particle filters, has a parameter N which determines the initial number of particles. Unlike the case for the branching filter algorithm, the number of particles in the hybrid filter is inherently static, thus eliminating excess computation for control of the total number of particles at the end of each update phase. A key feature of the hybrid method is its parameter ρ. At one extreme, choosing $\rho = 1$, the method mimics the interacting method, thus overly resampling. The other extreme, $\rho = \infty$, in which no resampling is done, results in a filtering procedure that does not resample (i.e. the classical weighted method). This parameter allows customization of the extent of resampling to a specific problem. An empirical solution for the optimal selection of ρ is still under investigation. To provide some intuition regarding values for ρ that are reasonable for various problem parameters, we examine five different fish tracking scenarios that differ in the initial number of particles, observation noise, and the chaotic nature of the signal motion.

5. FILTER COMPARISONS

5.1. The Simulation

A graphic representation of the simulations has been constructed and is depicted in Fig. 1. This is one frame of an animation that displays the signal evolution, observation sequence, and filter output. The top left panel indicates the simulated fish position. The upper middle panel is the observation raster as it is presented to the filter algorithms. The bottom three panels each display, on a heat scale, a current filter estimate of the conditional distribution of the signal position, one for each of the three particle-based filter methods applied in these trials. The refining hybrid filter estimate is presented on the left, the refining branching is in the middle, and the adaptive interacting estimate is on the right.

5.2. Problem Scenarios

Five fish tracking scenarios are simulated with the following parameters:

Problem Scenario	α	β	Std. of $V^{(i,j)}$	N	ρ
A	0.0005	0.0125	1.0	10000	1650
B	0.0005	0.0125	1.0	6000	1950
C	0.0005	0.0125	1.25	6000	2400
D	0.00005	0.0200	1.0	50000	2250
E	0.00005	0.0200	1.0	10000	2600

where α and β are signal motion model parameters (Sect. 3.1), Std. of $V^{(i,j)}$ is the standard deviation of the observation noise on a pixel-by-pixel basis (Sect. 3.2), and N and ρ correspond to filter parameters: the initial number of particles and the extent of resampling in the hybrid filter, respectively (Sect. 4.2). Reasonable values of ρ for the hybrid filter were not calculated but instead experimentally determined and are not optimized. Smaller values for α and larger values for β increase the difficulty of tracking the signal by making the fish movement more chaotic. Simulations are repeated for 500 trials with $\epsilon = 1$ and total simulation time $= 100$, where ϵ is the fixed time step between observations. Note that the fish signal has an initial position that is randomly and uniformly selected from the domain $[0, L_1] \times [0, L_2]$, and that the initial location of each particle is uniformly distributed. Because of this, each of the particle filters begins with an estimated position of the fish near the center of the observation area, and thus there is an average mean square error at the start of the simulation that all of the filters share.

5.3. Comparison Data

Graphs of the average mean square error (MSE) in the position estimates over the simulated time for 500 trials are provided in Figures 2 to 6 for each of the five problem scenarios (Sect. 5.2).

Variance in the approximated distribution of the interacting filter is less consistent after a few observations than the variance of the branching and the hybrid filter. This is due to overly extensive resampling during the update phase of the interacting filter that continues after the target has been found. Not only do the two refining filters have a more consistent average distribution variance, but on average these variances are smaller. For example, the average distribution variance for problem scenario D is graphed in Fig. 9. Note that for the simple 2-dimensional

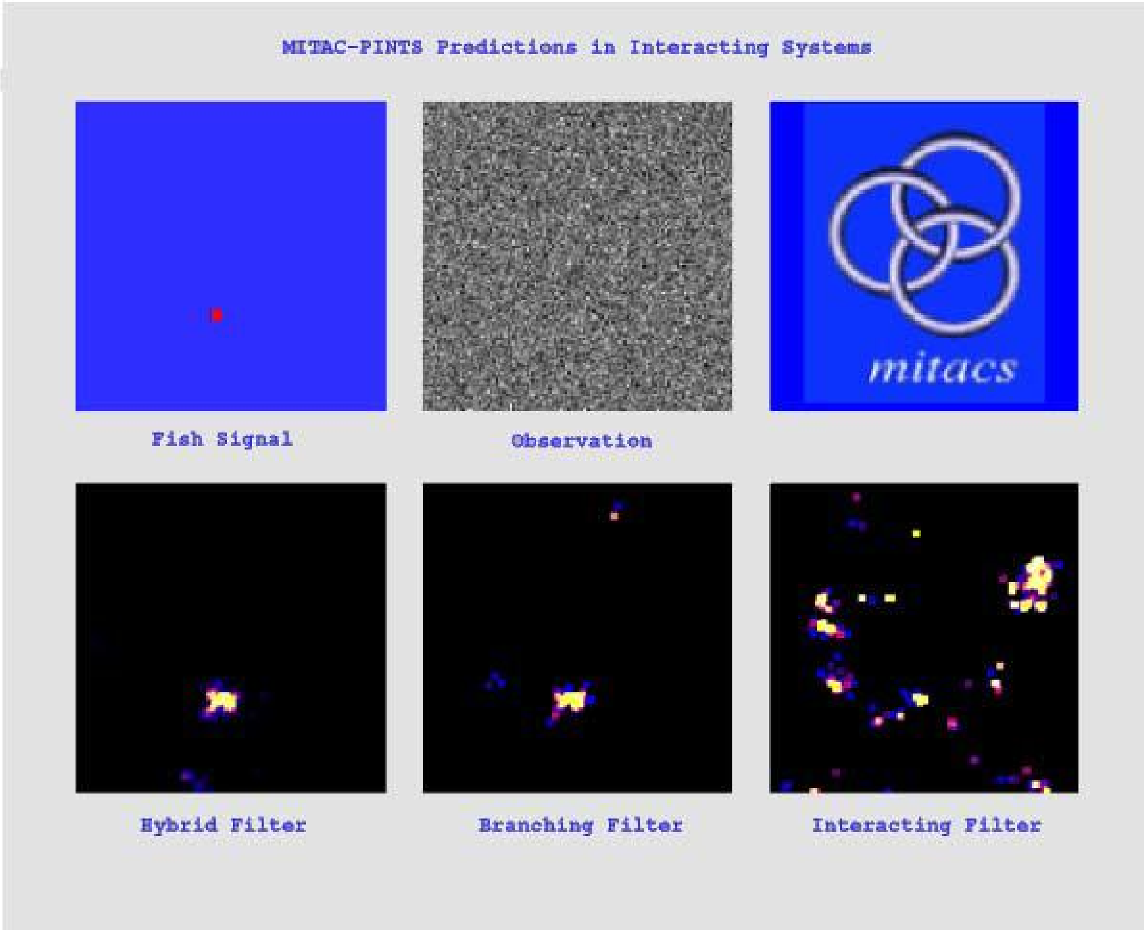

Figure 1. Sample frame from a fish problem animation.

fish signal in moderate observation noise, all three filtering algorithms, being at least adaptive, are readily able to localize the target. Therefore, the comparisons are not as dramatic as they would otherwise be. We expect that if the signal is defined in a higher dimensional domain and exhibits different states (for example, if fish have distinct swimming and eating states), the results will further favor the hybrid and branching methods. The reason for this is that while the refining methods undertake a more "cautious" particle adjustment, the merely adaptive interacting method will often eventually cluster all of its particles to suspected target positions, and thereby, with a practical finite number of particles, will have made erroneous particle adjustments that hamper long-term adaptation.

For scenarios A, B, C, and E the computation time for the hybrid filter is found to be approximately 12% less than that of the branching filter. The interacting filter has the fastest performance, outpacing the hybrid filter by more than 10% (Fig. 7). With $\rho = 1950$, the hybrid method resamples approximately three fifth of the particles at each observation. In problem scenario D the intial number of particles is 50000, and the hybrid filter ρ parameter is set relatively high, thus resampling approximately only two fifth of the particles at each observation. The computation time of the hybrid method is approximately 21% less than that of the branching method and 1% less than that of the interacting method. This gain in performance is due to the increased particle number and increased ρ value (Fig. 8). Following are the results of the MSE comparisons:

- Problem scenario A:
  In this, the easiest problem scenario, we find that the hybrid filter MSE is approximately 35% less than that of the other two filters.

- Problem scenarios B & C:
  For the next two scenarios the number of particles is decreased to 6000 and the observation noise raised (in scenario C), thus making the tracking problem more dificult. All three filters perform poorer in comparison to scenario A. With the hybrid filter ρ value adjusted to 1950 in scenario B, and 2400 in scenario C, the hybrid filter marginally outperforms the branching filter by approximately 7%. Both hybrid and branching filters have approximately 23% lower MSE than the interacting filter.

- Problem scenario D:
  With the signal parameters α modified to one tenth of the original value and β modified from 0.0125 to 0.02, the resulting fish signal evolves much more chaotically over time. The hybrid filter outperforms the other two filters by approximately 20% in the fidelity of its estimates.

- Problem scenario E:
  In the final problem scenario, the same parameters are used as in scenario D but with five times fewer particles. All three filters have trouble tracking the chaotic fish, hence the lack of a downward drift in MSE over time (Fig. 5). The hybrid filter, though, outperforms the other filters by approximately 25%.

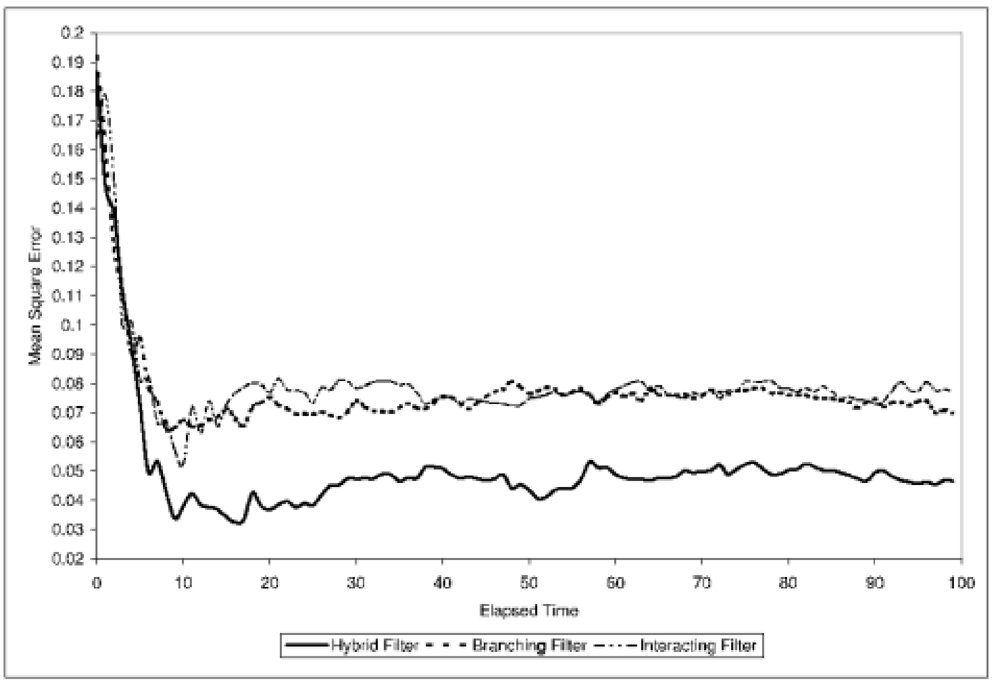

Figure 2. Average mean square error in fish position estimate for problem scenario A.

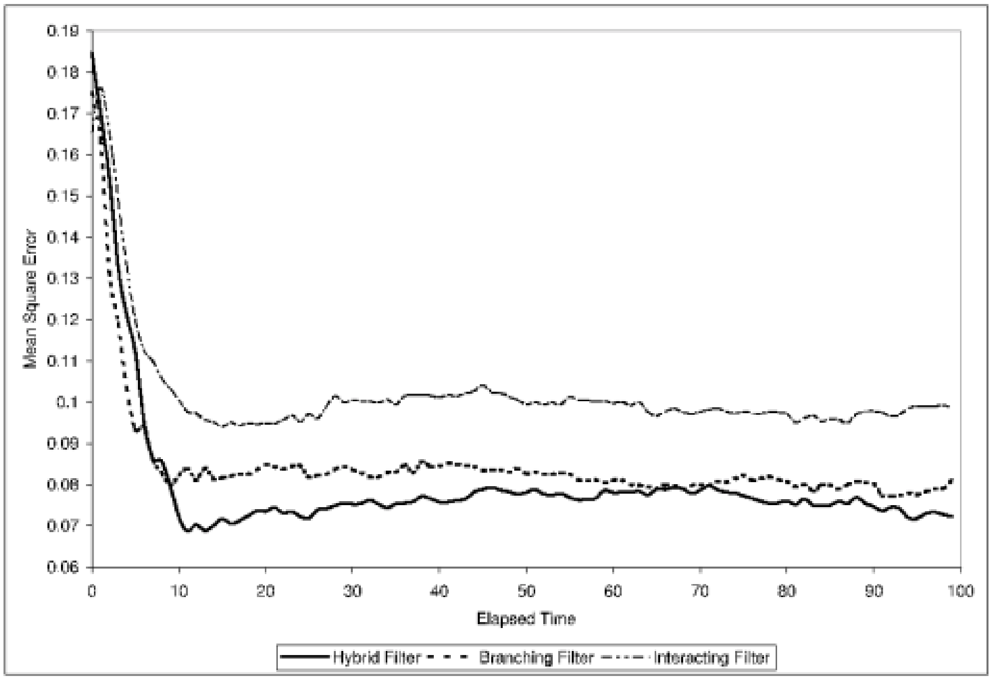

Figure 3. Average mean square error in fish position estimate for problem scenario B.

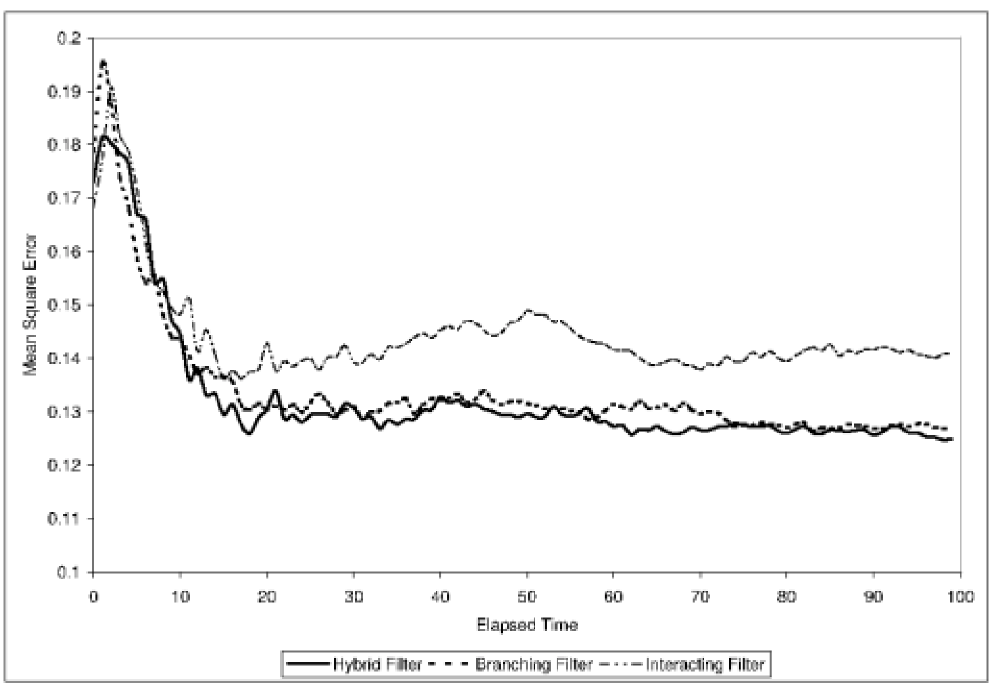

Figure 4. Average mean square error in fish position estimate for problem scenario C.

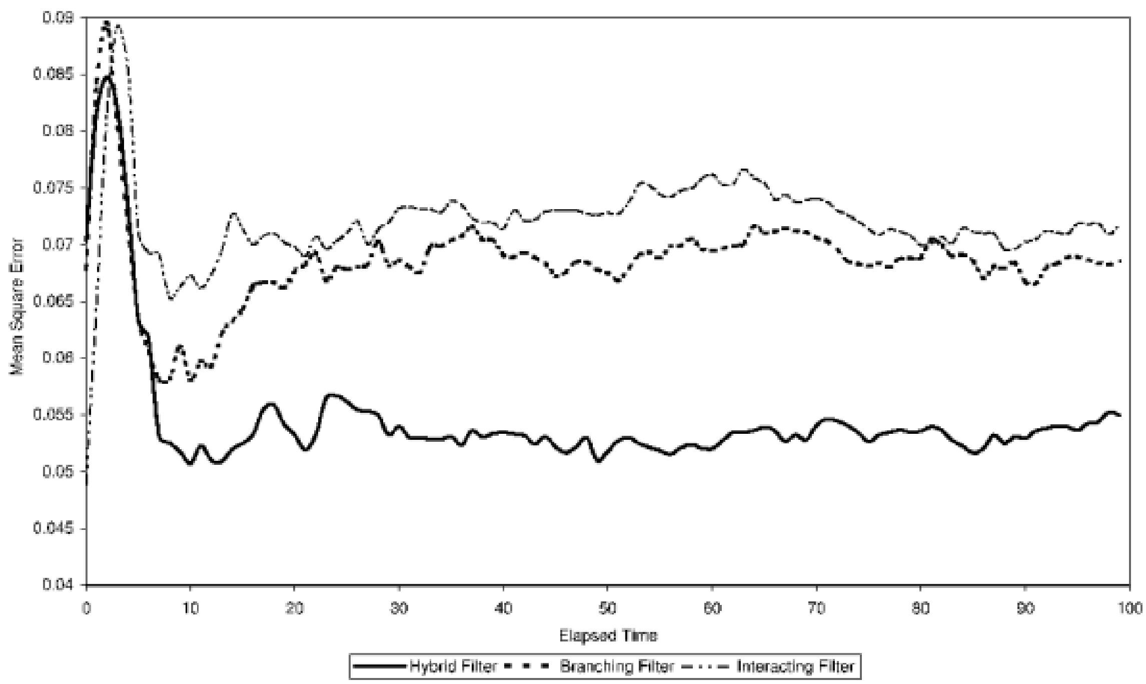

Figure 5. Average mean square error in fish position estimate for problem scenario D.

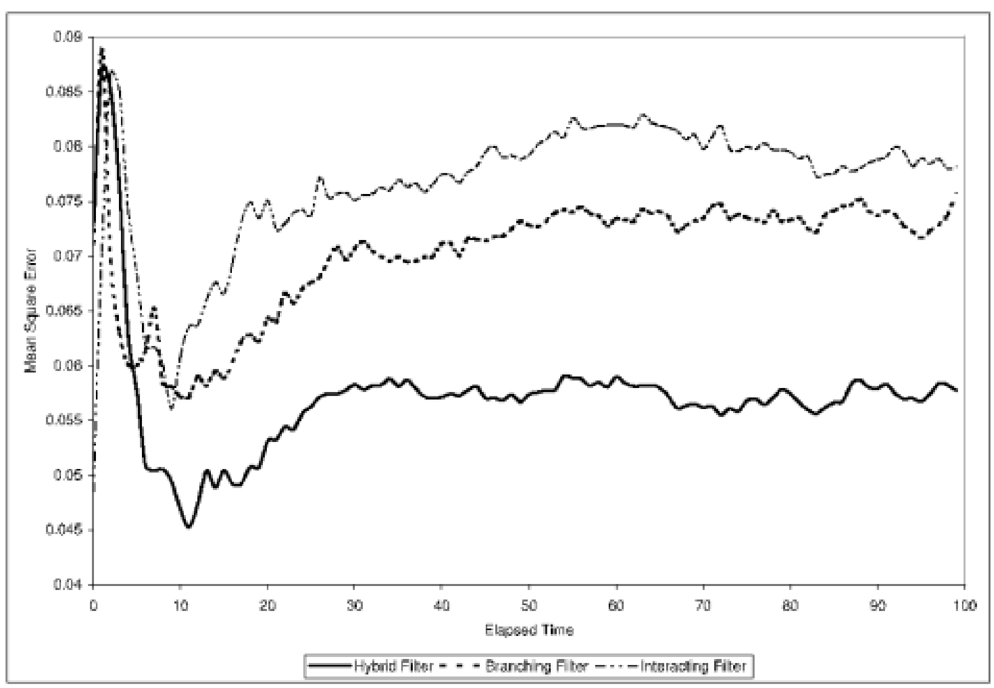

Figure 6. Average mean square error in fish position estimate for problem scenario E.

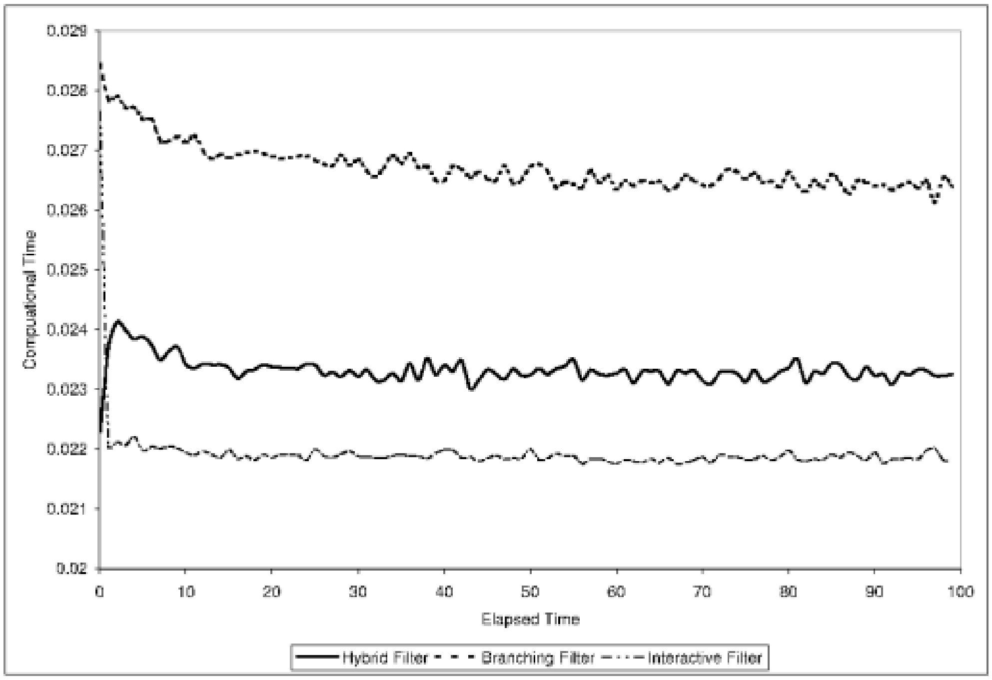

Figure 7. Average computation time for problem scenario B.

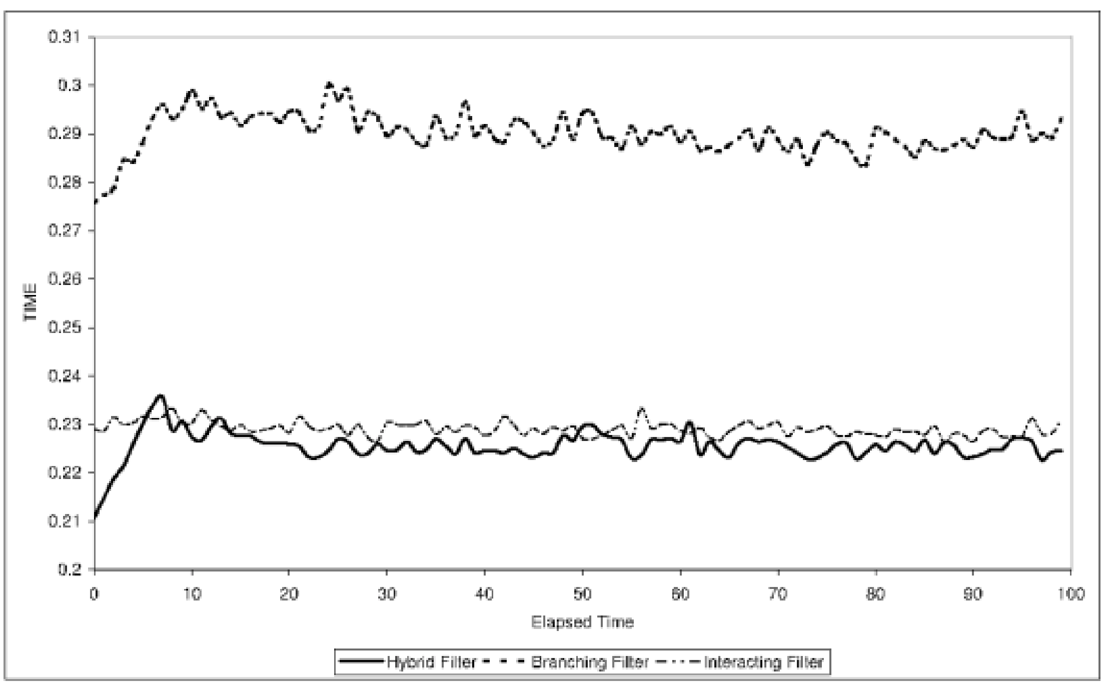

Figure 8. Average computation time for problem scenario D.

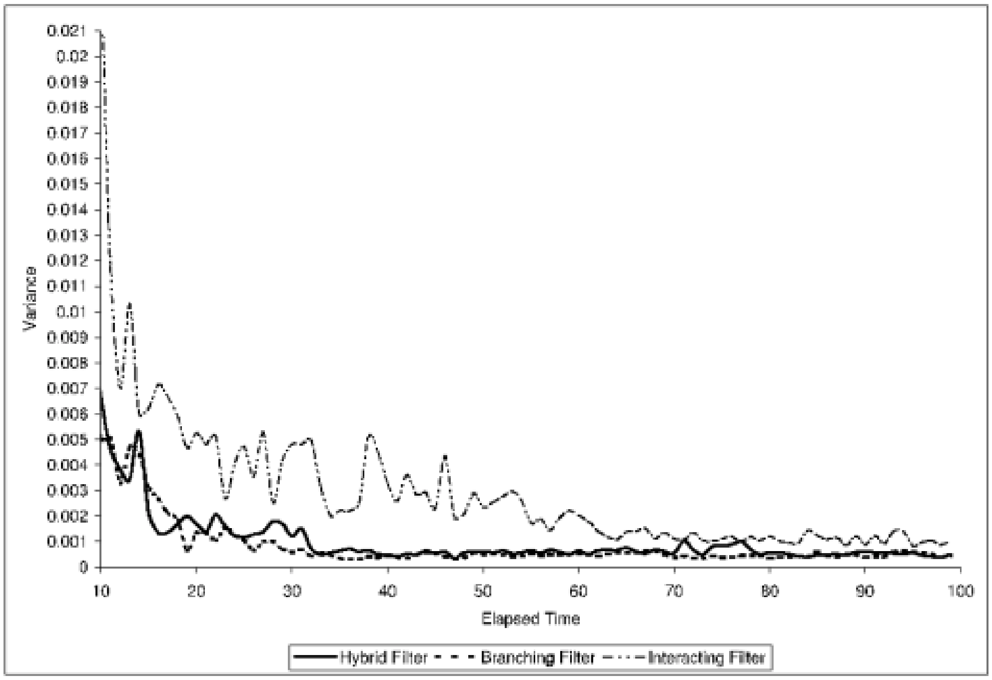

Figure 9. Average variance in the distribution for problem scenario D.

6. CONCLUSIONS

The *weighted-interacting* hybrid method exhibits superior performance in all five problem scenarios of this simulation for the selected ρ values. The parameter ρ is adjusted according to the amount of resampling desired. For problem scenarios with greater observation noise a higher ρ was considered. In our experiments, the diffusiveness of the fish signal motion was varied and corresponding adjustments to ρ led to superior hybrid filter performance in terms of both computing time and track fidelity. We are confident that for signals in higher dimensions and with more complicated structure, the capacity to customize the hybrid filter to the problem will lead to results that are similarly dominating. The desirable amount of resampling may even be dynamic during the filtering process (for example, a different resampling focus may be preferred for pre-localization and post-localization segments).

Whereas the refining branching method relies on deletions and duplications for particle resampling, thus requiring extra computations to control the randomly varying particle count, the hybrid method has an inherently static number of particles and therefore performs observation updates faster.

The advantage of refining grid particle-based nonlinear filters over the larger containing class of adaptive filters is that refining filters undertake a reduced, yet still asymptotically optimal amount of resampling. The introduced *weighted-interacting hybrid* or *refining weighted* filter is currently the best performing refining filter on the set of problems described in this paper and has outperformed other particle filters on all other problems assessed in our research.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support and sponsorship of Lockheed Martin Naval Electronics and Surveillance Systems, Lockheed Martin Canada, the Pacific Institute for the Mathematical Sciences, the Natural Science and Engineering Research Council (NSERC) through the Prediction in Interacting Systems (PINTS) centre of the Mathematics of Information Technology and Complex Systems (MITACS) network of centres of excellence, and the University of Alberta.

REFERENCES

1. D. Ballantyne, H. Chan, and M. Kouritzin, “A novel branching particle method for tracking,” *Proceedings of SPIE, Signal and Data Processing of Small Targets 2000* **4048**, pp. 277–287, 2000.
2. P. Del Moral and G. Salut, “Non-linear filtering using Monte Carlo particle methods,” *C.R. Acad. Sci. Paris* **320, Série I**, pp. 1147–1152, 1995.
3. P. Del Moral, M. Kouritzin, and L. Miclo, “On a class of discrete generation interacting particle systems,” *Electronic Journal of Probability* **paper No. 16**, pp. 1–20, 2001.