

A Bayesian Approach to Low-Thrust Maneuvering Spacecraft Tracking

Enrico M. Zucchelli* and Brandon A. Jones†
University of Texas at Austin, Austin, Texas 78705

<https://doi.org/10.2514/1.G007849>

Bayesian estimation with an explicit transitional prior is required for a tracking algorithm to be embedded in most multitarget tracking frameworks. This paper describes a novel approach capable of tracking maneuvering spacecraft with an explicit transitional prior and in a Bayesian framework, with fewer than two observations passes per day. The algorithm samples thrust profiles according to a multivariate Laplace distribution. It is shown that multivariate Laplace distributions are particularly suited to track maneuvering spacecraft, leading to a log probability function that is almost linear with the thrust. Principles from rare event simulation theory are used to propagate the tails of the distribution. Fast propagation is enabled by multi-fidelity methods. Because of the diffuse transitional prior, a novel k -nearest-neighbor-based ensemble Gaussian mixture filter is developed and used. The method allows Bayesian tracking of maneuvering spacecraft for several scenarios with fewer than two measurement passes per day and with a mismatch between the true and expected thrust magnitude of up to a factor of 200. The validity domain and statistical significance of the method are shown by simulation through several Monte Carlo trials in different scenarios and with different filter settings.

Nomenclature

a	= orbital semimajor axis, km
C_D	= drag coefficient
C_R	= solar radiation pressure coefficient
e	= orbital eccentricity
F	= state transition matrix
$f(\cdot)$	= state transition function
G	= Gramian matrix
$G(\cdot)$	= gating function
H	= measurement Jacobian
$h(\cdot)$	= measurement function
I	= indicator function
\mathbb{I}	= identity matrix
i	= orbital inclination, rad
$K(\cdot)$	= kernel function
N	= number of samples
P	= covariance matrix
$p(\cdot)$	= probability distribution
Q	= process noise matrix
$q(\cdot)$	= importance sampling distribution
R	= measurement noise matrix
R_e	= planet equatorial radius, km
\mathbf{r}	= position, km
r	= range, km
\mathbf{t}	= thrust profile, km/s ²
t	= time, s
\mathbf{v}	= velocity, km/s
w	= particle/component weight
\mathbf{y}	= measurement vector
α	= right ascension, rad
γ	= important sample for the stochastic collocation method
δ	= declination, rad
$\delta(\cdot)$	= Dirac delta function
ϵ	= measurement noise

$\Lambda(\cdot)$	= likelihood function
λ	= Laplace distribution inverse scale parameter
μ	= gravitational parameter, km ³ /s ²
ν	= true anomaly, rad
σ	= standard deviation
τ	= exponential random variable
χ	= augmented state
Ω	= right ascension of ascending node, rad
ω	= argument of periapsis, rad

Subscripts

\cdot^H	= high-fidelity model
\cdot^L	= low-fidelity model

Superscripts

\cdot_p	= pertaining to the measurement pass
\cdot_t	= pertaining to the trajectory

I. Introduction

SATELLITE maneuver capabilities are ever evolving, and not all operators are willing to share data about the paths of their spacecraft. Tracking is a necessary tool for space traffic management and collision avoidance, but unexpected maneuvers may lead to increased uncertainty and even loss of custody. The problem is exacerbated by the low frequency of observations passes and by the variety of modality of maneuvers, ranging from high thrust and impulsive to low thrust and with continuously varying magnitude and direction. A satellite's identity cannot be determined based on a measurement via traditional instruments (e.g., electro-optical and radar sensors), and thus the association between a measurement and a spacecraft has to rely on prior information. It is the propagation of an explicit initial prior into a transitional prior that allows proper data association, required for most multitarget trackers such as the classical multiple hypothesis tracker [1] or the generalized labeled multi-Bernoulli filter [2,3]. While the requirement of an explicit transitional prior can be circumvented [4,5], the absence of it generally leads to non-Bayesian, heuristics-based filters. The goal of this work is to develop a method that can efficiently track low-thrust maneuvering spacecraft when the observation passes are sparse and the maneuver is unknown, while at the same time being compatible with Bayesian multitarget tracking algorithms.

Maneuvering target tracking is a widely researched field [6–9]; however, little literature is available on Bayesian maneuvering target

Received 17 July 2023; accepted for publication 19 February 2024; published online 8 April 2024. Copyright © 2024 by Enrico M. Zucchelli and Brandon A. Jones. Published by the American Institute of Aeronautics and Astronautics, Inc., with permission. All requests for copying and permission to reprint should be submitted to CCC at www.copyright.com; employ the eISSN 1533-3884 to initiate your request. See also AIAA Rights and Permissions www.aiaa.org/randp.

*Ph.D. Candidate, Aerospace Engineering and Engineering Mechanics, 2617 Wichita Street; enricomarco@utexas.edu.

†Associate Professor, Aerospace Engineering and Engineering Mechanics, 2617 Wichita Street, Associate Fellow AIAA.

tracking for space objects. Challenges specific to the low-thrust maneuvering satellite tracking problem are due to 1) infinite dimensionality of the maneuver profile, 2) data sparsity, 3) nonlinearity of the problem, and 4) potentially large mismatches between physics models and actual behavior of the target when a maneuver is performed. A large portion of the literature on maneuvering target tracking relies on the Kalman filter [10] and on the extended Kalman filter (EKF) [11]. The simplest way to track maneuvering targets is by equivalent process noise, in which the process noise of a Kalman filter is inflated depending on how large the difference between actual and expected measurement is [12,13]. While fast and easy to implement, this approach is only successful when the maneuvers are small and when the frequency of observation passes is high because large state uncertainties quickly become non-Gaussian. Moreover, the method requires that measurement and target are already associated because the inflation of the target uncertainty is a function of the measurement residuals. Variable-dimension Kalman filters estimate maneuvers together with the state of the target [14,15]. To allow for observability, a constant acceleration through multiple measurements has to be assumed, which is not generally applicable when working with sparse data and low-thrust propelled space objects. On the other hand, it is possible to track spacecraft after single impulsive maneuvers followed by long-duration coasting arcs [16,17]. Optimal control-based approaches, such as those detailed in Refs. [18] or [19], allow for continuously varying maneuvers between two measurements but also require the data association to be known and assume optimality. Interacting Multiple Model filters [20] are another popular methodology to track noncooperative objects; this approach is, however, only optimal when the number of possible models is finite, which is not the case for a spacecraft. While the technology has been applied to maneuvering spacecraft by merging different maneuvers together [21,22], the finite number of models still introduces limitations to the applicability of the method.

A less common approach to maneuvering target tracking is to use maneuver distributions that are not sub-Gaussian. For the goal of this paper, a sub-Gaussian distribution can informally be described as a distribution with tails that decay as fast as, or faster than, Gaussian tails. In such a framework, the maneuver is seen as process noise, like in equivalent process noise Kalman filters, but its distribution has tails that decay slower than those of a Gaussian distribution. At the same time, the distribution of the process noise is not adapted to the measurement. Filters with non sub-Gaussian, or super-Gaussian, process noise can keep custody of a target better than corresponding Gaussian filters for large deviations in the dynamics. Roth et al. [23] develop a closed-form linearized filter with Student's t process and measurement noises; Huang et al. [24] employ variational Bayesian methods for an optimization-based filter with multivariate Laplace process noise. With rare exceptions, (e.g., Refs. [25,26]), the use of non sub-Gaussian priors for the maneuvers is limited to linear or linearized settings. Representing nonlinear transformations with high accuracy for the super-Gaussian tails can be a challenging task, as they can be orders of magnitude larger than the tails of Gaussian distributions.

In this work, the utility of super-Gaussian process noise distributions, and specifically multivariate Laplace distributions, is extended to nonlinear problems by exploiting concepts from rare event simulation techniques [27,28] and k -nearest-neighbor kernel density estimation. The computational burden caused by the large amount of sample propagation required is mitigated using multi-fidelity methods [29].

The main contribution of this paper is to provide an approach to Bayesian maneuvering target tracking that is capable of tracking low-thrust maneuvering spacecraft when observation passes are very sparse and the thrust profile and magnitude are unknown. To the best of the authors' knowledge, this is the first successful attempt to do so. Bayesian estimation is necessary for multitarget tracking; as satellite tracking is ultimately a multitarget tracking problem, being compatible with multitarget tracking algorithms is a requirement for operational applications of the method. There are six additional contributions in this paper. First, it is shown how to apply multi-fidelity methods to the propagation of a maneuvering satellite, where the dimensionality of the problem increases greatly because of the maneuvers. Second, it is argued and shown that a multivariate Laplace distribution is especially suited to represent the process noise for maneuvering spacecraft, as it is

close to being linearly exponential in the thrust, like the rocket equation. Third, it is proven that, when performing Bayesian estimation with multivariate Laplace priors and Gaussian measurements, rare event simulation and importance sampling greatly reduce the variance and bias of the estimates when large deviations, such as they may be observed in maneuvering target tracking, occur. Fourth, a novel ensemble Gaussian mixture filter algorithm, based on k -nearest neighbors, is proposed and developed. Fifth, a nonlinear filter that is robust to large deviations thanks to multivariate Laplace process noise is derived. The filter has an explicit transitional prior that is independent of the measurement, making it compatible with Bayesian multitarget tracking frameworks, and exploits specific synergy between the rare event simulation technique and the k -nearest-neighbor ensemble Gaussian mixture filter. Sixth, the filter is tested over several challenging scenarios, with large differences between the assumed maneuver magnitude and the actual maneuver magnitude and with very sparse observation passes.

The paper proceeds as follows. Section II describes the multi-fidelity stochastic collocation method and how it is adapted to the maneuvering target tracking problem. Section III describes the distribution of the thrust profile, and Sec. IV introduces a method to efficiently sample from Laplace distributions. In Sec. V, a novel ensemble Gaussian mixture filter is described. Section VI summarizes the filter, and Sec. VII reports the results obtained by simulation. Finally, conclusions are drawn in Sec. VIII.

II. Multi-Fidelity Stochastic Collocation

The most computationally intensive process in orbit determination is the propagation of the state distribution forward in time. To reduce run time, the proposed filter uses multi-fidelity methods [29], which are particularly useful when a large number of samples is needed, like in a particle filter [30] or in a Gaussian Mixture Model (GMM) filter [31]. Even if spacecraft tracking is generally performed on the ground, where there are few constraints on computational cost, a speed-up allows one to use more samples, therefore achieving higher accuracy for same computational costs. Multi-fidelity methods combine models of varying fidelity and produce tradeoffs between existing models. High-fidelity models are expensive and accurate; low-fidelity models offer shorter run time but incur larger errors. This work employs multi-fidelity stochastic collocation [32], a method that opportunely selects a few samples to be propagated with the higher-accuracy model; the outputs from those samples are then used as data to generate a linear mapping that corrects the remaining low-fidelity samples. The corrected samples, or surrogates, provide a local approximation that is almost as accurate as the high-fidelity samples, at a fraction of the computational cost. The stochastic collocation method, which is one approach to uncertainty quantification via surrogate modeling, is chosen over other multi-fidelity methods because it has already been proven and tested in astrodynamical problems [33], where it offered speed-ups by a factor between 50 and 100. The multi-fidelity stochastic collocation method has additionally already been used for orbit determination [34] and for uncertainty quantification in cislunar space [35]; in recent work, computational speed has been further increased using Vinti theory [36,37].

A. High-Level Description of Multi-Fidelity Stochastic Collocation

Let $x_t = f(x_0): x_0 \rightarrow \mathbb{R}^m$ be a function that is evaluated for the random input $x_0 \in \mathbb{R}^d$. Both a high- and a low-fidelity model are available to evaluate the function, referred to by superscripts H and L , respectively. The subscript 0 stands for the initial time, and the subscript t stands for the state evaluated at discrete times along the trajectory. The bifidelity stochastic collocation method consists of three steps:

- 1) Select a finite set $X_0 = [x_0^1, \dots, x_0^n] \in \mathcal{D}(x) \subset \mathbb{R}^d$ that captures the important x variations of $f(x)$, and compute corresponding low-fidelity outputs $X_t^L = [f^L(x_0^1), \dots, f^L(x_0^n)]$.
- 2) Out of the evaluated low-fidelity set X_t^L , select the set of interpolation nodes $\{f^L(x_0^i)\}$ and columns of X_t^L , and compute $\{f^H(x_0^i)\}$,

3) Interpolate the realizations $\{f^H(x_0^i)\}$ to $f^L(x)$ by using the low-fidelity samples to dictate the topology on \mathbb{R}^m .

The low-fidelity model is used at three different moments: first, to inform the choice of the interpolation nodes; then, to shape the interpolating function; and third, to locally evaluate the interpolating function. The key advantage over other interpolation methods is that the approximate topology of the output can be exploited; this fact helps mitigate the curse of dimensionality. Sections II.B and II.C report the methodology developed by Narayan et al. [32]. Details that are not relevant to the problem of this work are neglected. Section II.D describes how the method has been adapted to maneuvering target trajectories and is one of the main contributions of this paper.

B. Interpolation Nodes Selection

Let the low-fidelity distance between a vector $x_0 \in \mathbb{R}^m$ and a subspace $\mathbb{X}_0 \subset \mathbb{R}^m$ be

$$d^L(x_0, \mathbb{X}_0) = \inf_{w \in \mathbb{X}_0} \|f^L(x_0) - f^L(w)\| \quad (1)$$

The selection of the interpolating nodes is iterative, such that the next selected point is the one whose distance from the subspace spanned by the set of the previously selected points is maximum. The nodes $\gamma^1, \dots, \gamma^r$, where r is the number of interpolation nodes, are a subset of the columns of X_t^L and are greedily selected:

$$\gamma^i = \arg \max_{(x_t^j) \in \mathbb{R}^m} d^L(x_t^j, \{\gamma^1, \dots, \gamma^{i-1}\}) \quad (2)$$

Note that for tractability the domain \mathbb{X}_t has now been discretized. This problem can be solved via a single iteration of the pivoted Cholesky decomposition [38,39] of the Gram matrix $G^L = (X_t^L)^T X_t^L$:

$$(X_t^L)^T G^L X_t^L = PLL^T P \quad (3)$$

The pivoted decomposition is operated one pivot at a time; each pivoting solves one iteration of the problem (2). The decomposition is partial because only the first r permutations need to be computed. The iterative pivoted Cholesky decomposition ensures that all elements of $\{\gamma^1, \dots, \gamma^r\}$ are linearly independent, which is a requirement for model synthesis. The computation of the first r rows and columns of the Cholesky factor L of the pivoted Gramian G^L , needed for model synthesis, is a byproduct of the decomposition. For a step-by-step algorithm, see, for example, Ref. [33].

C. Model Synthesis

First, by exploiting the entries of X_t^L , an interpolation operator is constructed. Then, the operator is applied with the data $f^H(\{\gamma^1, \dots, \gamma^r\}) := \Gamma_t^H$. Let $f^L(x_0)$ be a pointwise low-fidelity solution that does not need to be part of the previously included data set; the coefficients c_n are determined by solving

$$\left\langle \sum_{n=1}^N c_n f^L(\gamma^n), \phi \right\rangle = \langle f^L(x_0), \phi \rangle, \quad \forall \phi \in f^L(\{\gamma^1, \dots, \gamma^r\}) \quad (4)$$

which implies

$$c = (LL^T)^{-1} g \quad (5)$$

where c is the stacked vector of the coefficients c_n and g is the stacked vector of $\langle f^L(x_0), \phi \rangle$. Then, the interpolation is applied with the high-fidelity data Γ_t^H :

$$\hat{x}^H = \sum_{n=1}^N c_n f^H(\gamma^n) \quad (6)$$

Because the pivoting matrix is known before performing the last computation, only the columns corresponding to the first r pivots of X_t^H need to be computed. In practice, a linear interpolator is constructed using the low-fidelity topology and data, but then the evaluation is done using the high-fidelity data. The error introduced by the interpolation can then be quantified [34], and it is used to estimate additional process noise in the filter.

D. Multi-Fidelity Stochastic Collocation for Maneuvering Spacecraft

In previous applications of the multi-fidelity stochastic collocation to astrodynamics problems, only the final state was considered by the interpolation. Thus, maneuvering trajectories ending at the same state may have the same correction, even if they started from different initial states and had different thrust profiles. These trajectories would require different corrections from one another because the corresponding propagations are different. Figure 1 illustrates an example of spacecraft trajectories ending at the same state when propagated with the low-fidelity model (blue) and at different states when propagated with the high-fidelity model (red). For simplicity, the thrust (green) is drawn as impulsive instead of continuous. To prevent this issue, the states at the start and end of every thrust segment are interpolated by the multi-fidelity method. The algorithm now effectively interpolates over $6 + 6n_{\Delta V}$ variables, with $n_{\Delta V}$ the number of thrust segments, causing an increase in the required number of high-fidelity samples. In this case, the increase in number of high-fidelity samples is sublinear in the number of variables.

III. Thrust Profile Distribution

This section is one of the main contributions of the paper, as it describes and justifies the selection of the distribution for the thrust profiles. The thrust profile distribution is the model used for process noise in the filter. A real thrust profile has infinite dimensions because the thrust at every instant is a free variable; to keep the problem tractable, the profiles are assumed piecewise constant in the radial, in-track, and cross-track (RIC) frames. Segments of constant thrust are preferred over thrust Fourier coefficients [40] because they provide a direct correlation with total ΔV . This fact enables a sampling method where there is a strict relation between probability density and ΔV . An example of a thrust sample t made of six segments is depicted in Fig. 2.

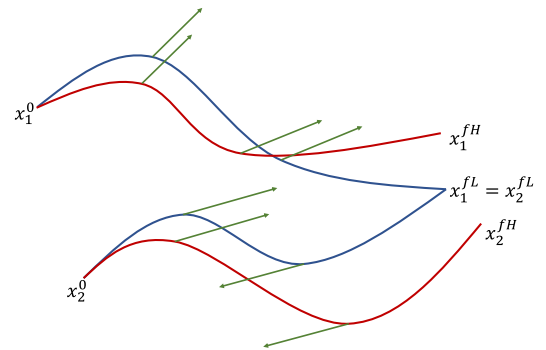


Fig. 1 Example of problematic multi-fidelity maneuvering spacecraft trajectories.

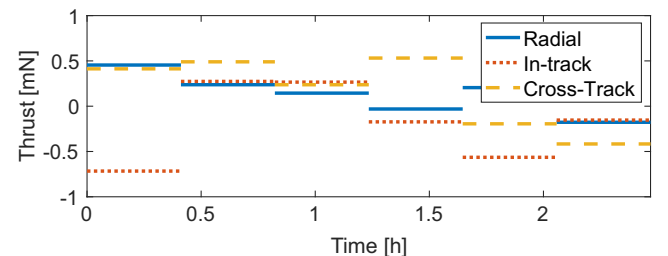


Fig. 2 Example of thrust profile sample between two measurement passes.

The thrust profile is assumed to be distributed according to a multivariate Laplace distribution, which is super-Gaussian. The probability density function (p.d.f.) of a multivariate Laplace distribution with mean $\mathbf{0}$ and covariance Q evaluated at \mathbf{t} is

$$f_{ML}(\mathbf{t}; \mathbf{0}, Q) = \frac{2}{\sqrt{|2\pi Q|}} \left(\frac{\mathbf{t}^T Q^{-1} \mathbf{t}}{2} \right)^{v/2} K_v(\sqrt{2\mathbf{t}^T Q^{-1} \mathbf{t}}) \quad (7)$$

where $v = (2 - d)/2$, the dimension of \mathbf{t} is $d = 3 n_{\Delta V}$, and $K_v(\cdot)$ is the modified Bessel function of the second kind and order v . Let the A -norm of a vector \mathbf{t} be $\|\mathbf{t}\|_A = \sqrt{\mathbf{t}^T A \mathbf{t}}$. By properties of $K_v(\cdot)$, for $\|\mathbf{t}\|_{Q^{-1}} \gg 1$ (which occurs when the thrust is larger than several standard deviations), the multivariate Laplace distribution asymptotically behaves as

$$\lim_{\|\mathbf{t}\|_{Q^{-1}} \rightarrow \infty} f_{ML}(\mathbf{t}; \mathbf{0}, Q) e^{\sqrt{2}\|\mathbf{t}\|_{Q^{-1}}} (\|\mathbf{t}\|_{Q^{-1}})^{1-v} = c \quad (8)$$

where c is a constant. The joint posterior distribution of initial state \mathbf{x}_0 and process noise \mathbf{t} , given a measurement \mathbf{y} and assuming a linearized problem with multivariate Laplace process noise, initial state $\mathbf{x}_0 \sim \mathcal{N}(\bar{\mathbf{x}}, P_0)$, Gaussian measurement noise $\epsilon \sim \mathcal{N}(\mathbf{0}, R)$, and where the innovation distribution is assumed Gaussian, can be approximated by the following log probability distribution when $\|\mathbf{t}\|_{Q^{-1}}$ is large,

$$\begin{aligned} \log p(\mathbf{x}_0, \mathbf{t} | \mathbf{y}) \approx & -\frac{1}{2}(\mathbf{x}_0 - \bar{\mathbf{x}}_0)^T P_0^{-1}(\mathbf{x}_0 - \bar{\mathbf{x}}_0) - \sqrt{2}\|\mathbf{t}\|_{Q^{-1}} + \\ & -\frac{1}{2}(\mathbf{h}(\mathbf{x}_f) - \mathbf{y})^T (H P_f H^T + R)^{-1}(\mathbf{h}(\mathbf{x}_f) - \mathbf{y}) \\ & + (1 - v) \log(\|\mathbf{t}\|_{Q^{-1}}) + C \end{aligned} \quad (9)$$

where C is a normalizing constant, H is the measurement Jacobian, P_f is the propagated state covariance, and $\mathbf{x}_f = \mathbf{f}(\mathbf{x}_0, \mathbf{t}, t_f)$. Let J_t be the part of the probability distribution in Eq. (9) associated with the thrust:

$$\log J_t(\mathbf{t}) = -\sqrt{2}\|\mathbf{t}\|_{Q^{-1}} + (1 - v) \log(\|\mathbf{t}\|_{Q^{-1}}) \quad (10)$$

The derivative of the logarithm is very small when $\|\mathbf{t}\|_{Q^{-1}} \gg 1$, and the term can thus be neglected. Assume now that Q is a multiple of the identity matrix. For a single thrust segment, the log PDF is now affine in the total ΔV . According to Tsolkovsky's equation, the actual cost of maneuvers is exponential in the ΔV , and this formulation of the thrust's PDF reflects that. For multiple thrust segments, one still has that $\log(J_t(\mathbf{t})) \approx \log(a J_t(\mathbf{t}))$. The asymptotically exponential cost in ΔV is indeed the main reason why a multivariate Laplace distribution is preferred over the Gaussian distribution or any other super-Gaussian distributions. The process noise variance Q is still a parameter to be tuned, but thanks to the quasi-linearly exponential dependency in the ΔV , the filter is much less sensitive to the selection of Q than a Gaussian filter would be.

In conclusion, the use of a multivariate Laplace distribution for the thrust profiles of the spacecraft allows having a robust filter, mostly insensitive to the set standard deviation, and a PDF that resembles more closely the actual cost of operating a spacecraft, all while including time-varying thrust within the same measurement gap.

IV. Rare Event Simulation Particle Filter

This section describes the rare event simulation particle filter, one of the main contributions of this paper. When unexpectedly large deviations from the assumed model occur, the approach proposed in this section offers greatly reduced bias and variance with respect to a standard particle filter, thanks to a better representation of the tails of the distributions. As a maneuver can often cause large deviations, such a feature is a key component of the overall algorithm described in this paper. This section provides theoretical and empirical results.

A. Rare Event Simulation for State Estimation

This subsection shows how concepts from rare event simulation, and specifically importance sampling, can be adapted to reduce the variance and the bias of a particle filter estimate when a large deviation, or maneuver, occurs.

In statistical estimation, one generally wishes to obtain the minimum mean square error (MMSE), equal to the following normalized integral:

$$\begin{aligned} x^{MMSE} &= \mathbb{E}_{p(x|y)}[x] = \frac{\mathbb{E}_{p(x)}[x p(y|x)]}{\mathbb{E}_{p(x)}[p(y|x)]} = \frac{\int_{\mathcal{D}(x)} x p(y|x) p(x) dx}{\int_{\mathcal{D}(x)} p(y|x) p(x) dx} \\ &:= \frac{A(y)}{B(y)} \end{aligned} \quad (11)$$

where $p(x)$ is the prior distribution of x , $p(y|x)$ is the measurement distribution conditioned on the state, and $p(x|y)$ is the posterior distribution. The last equality of Eq. (11) is a definition for $A(y)$ and $B(y)$. Note that both A and B are functions of y , a random variable; however, given the data, they are deterministic. For brevity, the dependency on y is dropped. The following analysis is carried out for scalar state and measurement but can be easily generalized to arbitrary dimensions. A particle filter is equivalent to using crude Monte Carlo to approximate both numerator and denominator of Eq. (11):

$$\hat{x}^{MMSE} = \frac{\sum_{i=1}^N \frac{1}{N} x_i p(y|x_i)}{\sum_{i=1}^N \frac{1}{N} p(y|x_i)} = \frac{\hat{A}}{\hat{B}} \quad (12)$$

where $x_i \sim p(x)$, and $\hat{\cdot}$ symbolizes an empirical estimate. The expected value and the covariance of a ratio of two random variables [41] can be approximated by Taylor series:

$$\mathbb{E}_{p(x)} \left[\frac{\hat{A}}{\hat{B}} \right] \approx \frac{A}{B} \left(1 - \frac{\text{Cov}_{p(x)}[\hat{A}, \hat{B}]}{AB} + \frac{\text{Var}_{p(x)}[\hat{B}]}{B^2} \right) \quad (13)$$

$$\text{Var}_{p(x)} \left[\frac{\hat{A}}{\hat{B}} \right] \approx \left(\frac{A}{B} \right)^2 \left(\frac{\text{Var}_{p(x)}[\hat{A}]}{A^2} - 2 \frac{\text{Cov}_{p(x)}[\hat{A}, \hat{B}]}{AB} + \frac{\text{Var}_{p(x)}[\hat{B}]}{B^2} \right) \quad (14)$$

From Eqs. (13) and (14), one can compute the expected root mean square error (RMSE) of the estimate:

$$\mathbb{E}_{p(x)}[\text{RMSE}(\hat{x})] = \sqrt{\left(\frac{A}{B} - \mathbb{E}_{p(x)} \left[\frac{\hat{A}}{\hat{B}} \right] \right)^2 + \text{Var}_{p(x)} \left[\frac{\hat{A}}{\hat{B}} \right]} \quad (15)$$

Note that this RMSE is the expectation over the sampling distribution and refers to the error between the computed MMSE and the exact MMSE; the actual error of the estimate might in fact be larger.

The bootstrap particle filter is also called sequential importance sampling because it exploits a sampling distribution, the state's prior $p(x)$, to compute an expected value for a different distribution $p(x|y)$. Importance sampling for particle filters is necessary because $p(x|y)$ is unknown, and $p(x)$ is just considered a good guess. However, importance sampling in the framework of rare event simulation is different and consists of a variance reduction technique [27,28]. In this context, some instances of importance sampling are also called exponential twisting, or exponential tilting.

Similar to what is normally done to compute the probability of a rare event, it is here proposed to sample from a proposal distribution $q(x)$, strictly positive for every x where the posterior is nonzero. The samples are then given the weights

$$w_i(x_i) = \frac{p(x_i)}{q(x_i)} \quad (16)$$

Because of properties of importance sampling, all the first raw moments of \hat{A} and \hat{B} computed empirically are unaffected by the

choice of $q(x)$, but all second moments are. While rare event simulation generally only reduces the variance of the estimate, in this case, it has an effect on the bias, too, because of Eq. (13). The second moments of \hat{A} and \hat{B} can be computed:

$$\text{Var}_{q(x)}[\hat{A}] = \frac{1}{N} \int_{\mathcal{D}(x)} (xp(y|x) - A)^2 \frac{p^2(x)}{q(x)} dx \quad (17)$$

$$\text{Cov}_{q(x)}[\hat{A}, \hat{B}] = \frac{1}{N} \int_{\mathcal{D}(x)} (xp(y|x) - A)(p(y|x) - B) \frac{p^2(x)}{q(x)} dx \quad (18)$$

$$\text{Var}_{q(x)}[\hat{B}] = \frac{1}{N} \int_{\mathcal{D}(x)} (p(y|x) - B)^2 \frac{p^2(x)}{q(x)} dx \quad (19)$$

From these equations and Eqs. (13–15), opportunely computing expectations over $q(x)$ instead of $p(x)$, one may select $q(x)$ to reduce the expected RMSE of the estimate. It is a known result that the variance of \hat{B} is minimized by sampling from the posterior distribution. Sampling from the posterior distribution is what many inference techniques, such as the auxiliary particle filter [42] and Markov chain Monte Carlo methods [43], attempt to do. Because they require having an associated measurement before propagation, neither of the mentioned techniques are used in this work.

B. Rare Event Simulation Filter with Laplace Prior

Suppose now that one wishes to estimate $x \in \mathbb{R}$, which has a Laplace prior distribution $p(x)$ with zero mean and rate parameter λ_L :

$$p(x) = \frac{\lambda_L}{2} e^{-\lambda_L |x|} \quad (20)$$

The corresponding standard deviation is $\sigma_L = \sqrt{2}/\lambda_L$. The measurement likelihood function $\Lambda(y, x)$ is Gaussian with standard deviation σ_y and mean $x \gg \lambda_L \sigma_y^2$. The numerator of Eq. (11) can be approximated by removing the absolute value from the Laplace distribution:

$$\begin{aligned} B &= \mathbb{E}_{p(x)}[\Lambda(y, x)] \\ &= \int_{-\infty}^{\infty} \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{(x-y)^2}{2\sigma_y^2}} \frac{\lambda_L}{2} e^{-\lambda_L |x|} dx \approx \frac{\lambda_L e^{\frac{1}{2}\lambda_L(\lambda_L \sigma_y^2 - 2y)}}{2} \end{aligned} \quad (21)$$

The assumption is valid for large y because the Gaussian tail decays much faster than the Laplace tail. The same assumption is used in the computation of all the integrals of this section. The MMSE can then be approximated, too:

$$\mathbb{E}_{p(x|y)}[x] \approx y - \lambda_L \sigma_y^2 \quad (22)$$

By drawing N samples from $q(x) = \frac{\tilde{\lambda}_L}{2} e^{-\tilde{\lambda}_L |x|}$, where $\tilde{\lambda}_L$ is a variable to be tuned, the variance of the estimate \hat{B} is

$$\begin{aligned} \text{Var}_{q(x)}[\hat{B}] &= \frac{1}{N} \text{Var}_{q(x)} \left[\Lambda(y, x) \frac{p(x)}{q(x)} \right] \\ &= \frac{1}{N} \int_{-\infty}^{\infty} \Lambda^2(y, x) \frac{1}{2} \lambda_L \frac{e^{-\lambda_L |x|}}{\tilde{\lambda}_L^2 e^{-2\tilde{\lambda}_L |x|}} dx - B^2 \approx \\ &\approx \frac{1}{N} \frac{\lambda_L^2}{\tilde{\lambda}_L} \frac{e^{\frac{1}{4}(\tilde{\lambda}_L - 2\lambda_L)(\sigma_y^2(\tilde{\lambda}_L - 2\lambda_L) + 4y)}}{4\sqrt{\pi}\sigma_y} - B^2 \end{aligned} \quad (23)$$

The variance of \hat{A} and the covariance of \hat{A} and \hat{B} can be approximated in a similar manner, and the computation is not reported for brevity. Finally, the total expected RMSE sampling from $q(x)$ can be computed with Eqs. (13–15), but computing expectations over $q(x)$ instead of $p(x)$,

$$\mathbb{E}_q[\text{RMSE}(\hat{x})] = \sqrt{\left(\frac{A}{B} - \mathbb{E}_q \left[\frac{\hat{A}}{\hat{B}} \right] \right)^2 + \text{Var}_q \left[\frac{\hat{A}}{\hat{B}} \right]} \quad (24)$$

which is therefore a function of $\lambda_L, \tilde{\lambda}_L, y$, and σ_y . After observing y by opportunely selecting $\tilde{\lambda}_L$ one can adjust the second moments of the joint distribution of \hat{A} and \hat{B} such that the RMSE is minimized.

C. Results

Figure 3 shows the theoretical expected log RMSE of the estimate, computed as in Eq. (24), obtained with a particle filter with 2500 samples when $\lambda_L = 1$ and $\sigma_y = 1$. The RMSE is plotted as a function of $\tilde{\sigma}_L = \sqrt{2}/\tilde{\lambda}_L$ and of the measurement y . The plot is capped when log RMSE is larger than 5 because the approximations used to compute the integrals such as Eq. (21) get inaccurate towards the top-left corner. When $q(x) = p(x)$, marked as the vertical black line, the estimate quickly degrades for growing values of y . However, using the optimal $\tilde{\sigma}_L$, the red line in the figure, the decrease in accuracy is mitigated. Note that the optimal value of $\tilde{\sigma}_L$ is a function of y ; therefore, if one has to sample before knowing the value of y , it is best to sample from a mixture of distributions with multiple values of $\tilde{\lambda}_L$. The optimal value of $\tilde{\sigma}_L$ is very close to linear with y , with an empirical regression factor close to $\tilde{\sigma}_{L,\text{opt}} = \sqrt{2}y$.

Figure 4 shows the same quantity of interest as Fig. 3 with same color scale and contour lines, but computed empirically. Every data point provides the RMSE obtained from 1000 particle filters, each with 2500 particles. The left plot is for one-dimensional prior and measurement, and the right plot is for a six-dimensional prior, with one-dimensional measurement. For the latter, only the observed dimension is considered in the RMSE. The red line is the same that minimizes the RMSE of Eq. (24), and the green line minimizes the particle filters' RMSE. Despite the noise, one can qualitatively verify that the red and green lines have similar slopes, and in the one-dimensional case, they tend to overlap. Moreover, the similarity between the left plot in Figs. 3 and 4 extends to the colors and contour lines, implying that, as expected, the values are similar between the two figures. Based on the right plot in Fig. 4, with a higher number of dimensions, the impact of importance sampling is reduced, and the optimal $\tilde{\sigma}$ is smaller, but the benefits of using rare event simulation remain.

In conclusion, this section shows that 1) with Laplace priors, it is necessary to use wider sampling distributions if large deviations occur, as is likely in maneuvering target tracking problems, and 2) the optimal choice of distribution is a function of the measurement. This fact introduces a problem that is solved in Sec. VI.A.1.

V. Nearest-Neighbor Ensemble Gaussian Mixture Filter

This section introduces an ensemble Gaussian mixture filter (EnGMF) aided by a k -nearest-neighbor algorithm. An EGMF [44] has several benefits over particle filters, most importantly the fact that

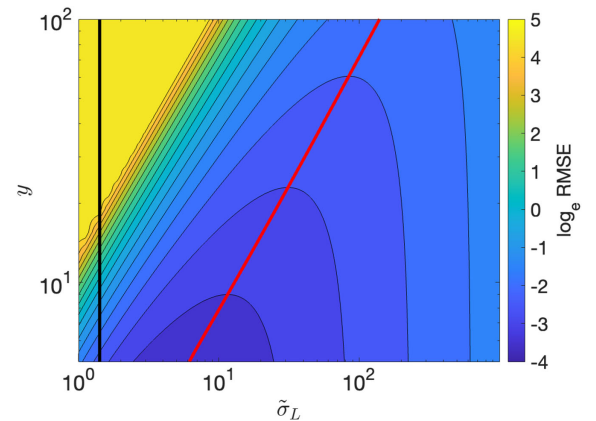


Fig. 3 Expected logarithm of the RMSE according to Eq. (24).

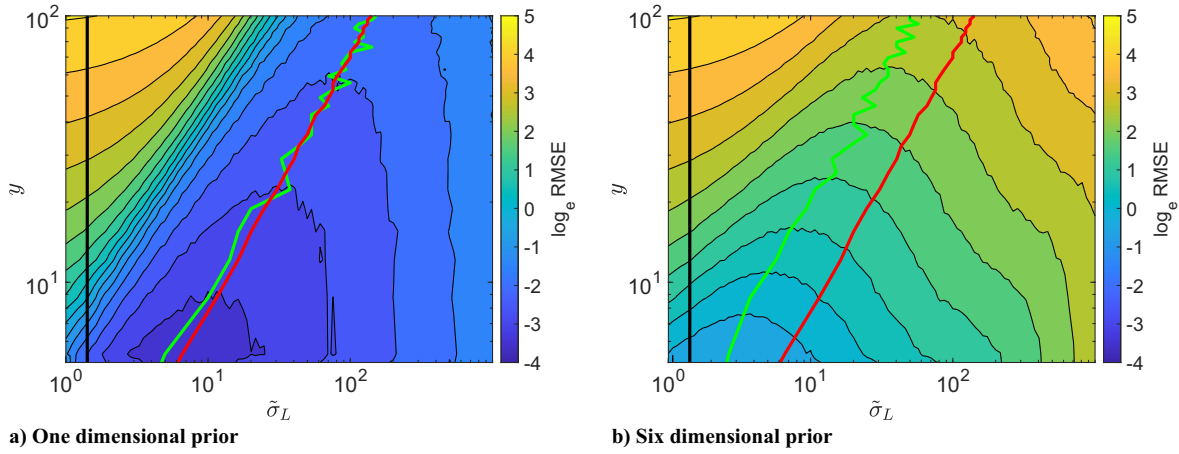


Fig. 4 RMSE for Monte Carlos of particle filters.

it allows for reduced variance in the weights of the updated components, and thus reduces the risks of particle depletion and degeneracy. In an EGMF, one samples particles from an initial distribution; the particles are propagated, and regularized into Gaussian kernels, forming together a GMM. The measurement update then follows that of a Gaussian sum filter. At the following time step, new particles are sampled from the updated GMM. In the context of this work, the development of an EGMF with kernels dependent on their neighboring samples is necessary because of the large diversity in the distance between sampled particles; therefore, using the same kernel for all particles may lead to small isolated components. The EGMF is equivalent to the preregularized particle filter [45,46], except that the measurement update is performed with a bank of weighed Kalman filters. Regularization consists of substituting Dirac delta functions with kernels, making the distribution continuous rather than discrete. In an EGMF, the choice of kernel is limited to Gaussian distributions, with mean equal to the state of the corresponding particle. The only design variable left is the covariance matrix:

$$\sum_{j=1}^N w_j \delta(\mathbf{x}_j, \mathbf{x}) \rightarrow \sum_{j=1}^N w_j \mathcal{N}(\mathbf{x}; \mathbf{x}_j, \Sigma_j)$$

Regularization modifies the original distribution and inflates the overall covariance. This inflation is considered an acceptable drawback that is outweighed by the fact that a discrete distribution is now continuous, and thus closer to the actual distribution under many other metrics, such as the Kullback–Leibler divergence. A commonly used regularization method for the EGMF is Silverman’s rule of thumb [44,47]: the kernel is Gaussian, and the variance is a scaled empirical covariance. Silverman’s regularization is optimal for Gaussian distributions and can be used for non-Gaussian distributions, too. Although it has been used to solve astrodynamics problems before [48], when applied to this problem, it fails because with multivariate Laplace process noise the average distance between particles varies widely when moving from the mode to the tails. Using the same variance for all particles causes the density to either have undesired gaps in the distribution or have kernels so diffuse that too much information is lost. For this reason, a locally adaptive and anisotropic kernel is chosen.

Starting from the approach proposed by Brox et al. [49],

$$P_{k+1|k}^i = \eta \mathbb{I} + \sum_{j=1}^N K_0(\mathbf{x}_i, \mathbf{x}_j) \frac{(\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T}{k-1} \quad (25)$$

$$K_i = \mathcal{N}(\mathbf{x}; \mathbf{x}_i, P_{k+1|k}^i) \quad (26)$$

where K_0 is a particle-independent isotropic weighing function, η is a regularization parameter, and K_i is the kernel applied to particle \mathbf{x}_i . This method alone allows for anisotropic kernels that are dependent

on the local orientation of the density, as opposed to Silverman’s rule of thumb, which provides particle-independent kernels, or the classic version of the k -nearest neighbor regularization such as in [50], which provides isotropic kernels. The method from Eq. (25) still uses a particle-independent, Gaussian function $K(\cdot)$ to weigh the particles contributions to the covariance of the kernel. This fact may lead to very small weights in the tails, where the particles are very far from one another, causing degeneracy of the kernel. Furthermore, the computational cost of computing the covariance for all kernels grows with n^2 because every particle contributes to the covariance of every other kernel. Hence, an improved version of the method is proposed.

A. Proposed Approach

Building on Eq. (25), substitute the particle independent kernel $K_0(\cdot)$ with a particle-dependent gating function $G_i(\mathbf{x}^i, \cdot)$. This substitution solves both of the mentioned problems. The gating function is chosen to be based on k -nearest neighbors, such that

$$G_i(\mathbf{x}_i, \mathbf{x}) = \begin{cases} 1 & \text{if } \|\mathbf{x}_i - \mathbf{x}\| \leq h^i \\ 0 & \text{if } \|\mathbf{x}_i - \mathbf{x}\| > h^i \end{cases} \quad (27)$$

$$P_{k+1|k}^i = \sum_{j=1}^N G_i(\mathbf{x}_i, \mathbf{x}_j) \frac{(\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T}{k-1} \quad (28)$$

$$K_i = \mathcal{N}(\mathbf{x}; \mathbf{x}_i, P_{k+1|k}^i) \quad (29)$$

where h^i is the distance between \mathbf{x}^i and its k th nearest neighbor. For problems with small dimensionality, k -dimensional trees can be built with cost $c_1 n \log n$, and each query costs $\log n$, making the overall regularization cost $c_2 kn + c_1 n \log n$, more manageable than the original approach of Eq. (25). Here h^i is used as a gating distance: This way, instead of spherical kernels, kernels that better follow the data are allowed. The corresponding difference in outcomes will be shown in Sec. VI.A.2. The proposed regularization is equivalent to a sample covariance, centered on the given particle, and where samples are truncated after the k th farthest one. This anisotropic adaptive method has two main benefits over most regularization methods. First, the kernels are smaller where the data are denser and larger where the data are less dense. This fact allows the estimator to more closely approximate the data when it is dense, while avoiding undesired gaps. Second, instead of being spheres, or ellipses oriented along predefined axes, the kernels follow the local orientation of the data. A paper describing an alternative approach for an adaptive EGMF was presented by Popov and Zanetti [51] shortly before this work was submitted for peer review (the method described in this section was developed independently from their paper). In that work, the function $G_i(\mathbf{x}_i, \mathbf{x})$ is no longer a gating function, but it is a

Gaussian kernel with covariance $h^i \mathbb{I}$. Consequently, the computational cost increases with n^2 .

In the spacecraft tracking problem, the state includes position and velocity measures, which cannot be compared with one another. The neighbors are therefore sought in dimensionless canonical units, which are

$$\mathbf{r}_{\text{LU}} = \frac{\mathbf{r}}{R_e} \quad (30)$$

$$\mathbf{v}_{\text{LU}} = \mathbf{v} \sqrt{\frac{R_e}{\mu}} \quad (31)$$

where R_e and μ are the planet's equatorial radius and gravitational parameter, respectively. If additional parameters were to be estimated, such as the drag coefficient, one should find suitable normalization constants, most likely based on how nonlinear are the effects of such parameters.

The total prior PDF can be computed as

$$p(\mathbf{x}) = \sum_{i=1}^N \frac{w_i}{\sqrt{|2\pi P_{k+1|k}^i|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}_i)^T (P_{k+1|k}^i)^{-1} (\mathbf{x} - \mathbf{x}_i)\right) \quad (32)$$

This approach allows for a k -nearest-neighbor method that is anisotropic and that accounts for local correlation between variables, which are characteristics that are often missing in that category of methods.

The combination between prior importance sampling and density-dependent regularization offers the additional advantage that covariance inflation is mitigated: the samples with the highest weights lie in the densest regions, where the k th neighbor is closer than it would be in the less dense regions.

VI. Filter Architecture

This section synthesizes the contributions of the previous sections into a single filter. As with most dynamic parameter estimation filters, this algorithm is divided into two main components: the construction of the transitional prior and the measurement update. In this case, the transitional prior is constructed by sampling particles from a proposal density, weighing them according to importance sampling, propagating the particles, and regularizing. The update is made of two phases: first, the relative weights of all components are updated, and, second, an iterated batch least squares (IBLS) is used to update mean and covariance for each component.

A. Construction of Transitional Prior

This subsection describes how the explicit transitional prior is approximated. Specific attention is given to reproducing the tails of the distribution. By employing a Gaussian sum, as described in Sec. V, the distribution's tails are bound to be sub-Gaussian. The objective of this section is to show how importance sampling described in Sec. IV works in tandem with the nearest-neighbor ensemble Gaussian mixture filter presented in Sec. V; together, they provide a distribution that resembles a super-Gaussian distribution for a large portion of the domain. This result is achieved while representing the shape of the distribution as closely as possible and inflating the total covariance as little as possible. The construction of the transitional prior consists of the following steps. For each particle, 1) sample initial state from a Gaussian sum prior, 2) sample a thrust profile from a proposal density distribution, 3) weigh the particle according to its thrust profile, 4) propagate the particle with its initial state and thrust profile, and 5) perform regularization.

1. Process Noise Importance Sampling

As explained in Sec. IV, the proposed filter samples the process noise from a wider distribution than the one assumed. This way, deviations that are potentially orders of magnitude larger than the assumed standard deviation will still be represented.

Sampling from a multivariate Laplace distribution with mean $\mathbf{0}$ and variance \mathbf{Q} is done in two steps [52]. First, a random sample is drawn from a normal distribution with mean $\mathbf{0}$ and variance \mathbf{Q} . Then, all components of the drawn sample are multiplied by the square root of a sample drawn from an exponential distribution with scale 1:

$$\mathbf{t} = \sqrt{\tau} \mathbf{r} \quad \tau \sim \text{Exp}(1) \quad (33)$$

where $\mathbf{r} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$. As in Sec. IV, let $q(\mathbf{t})$ be the sampling, or proposal, distribution and $p(\mathbf{t})$ be the target distribution. The weight w_i can be computed using Eq. (16). Although there are asymptotic guarantees of convergence whenever the domain of $q(\cdot)$ fully encompasses the domain of $p(\cdot)$, choosing too wide or too small a distribution can lead to very slow convergence when a finite number of samples are used (as was demonstrated in Sec. IV). To be conservative, a mixture of densities is chosen as the proposal distribution. Let $u \sim \mathcal{U}(0, 1)$,

$$\mathbf{t} = \begin{cases} \sqrt{\tau} \mathbf{r} & \text{if } u \leq \alpha_0, \\ a_i \sqrt{\tau} \mathbf{r} & \text{if } \sum_{l=0}^{i-1} \alpha_l < u \leq \sum_{l=0}^i \alpha_l, \quad i \geq 1 \end{cases} \quad (34)$$

where α_i are positive weights such that $\sum_{i=0}^n \alpha_i = 1$ and $a_i > 1$ for every i . In this case, the weight is computed as

$$w_i = \frac{f_{ML}(\mathbf{t}; \mathbf{0}, \hat{\mathbf{Q}})}{\alpha_0 f_{ML}(\mathbf{t}; \mathbf{0}, \hat{\mathbf{Q}}) + \sum_{l=1}^n (1 - \alpha - \sum_{j=1, j \neq l}^n \alpha_j) f_{ML}(\mathbf{t}; \mathbf{0}, a_l^2 \hat{\mathbf{Q}})} \quad (35)$$

Sampling from a mixture of several multivariate Laplace distributions allows the filter to better react to a wider range of thrust magnitudes. Using importance sampling for the thrust still requires the guess of an upper bound for the thrust magnitude. However the practitioner should not worry about exceeding it because it only mildly affects the final estimate. The only drawback caused by using too large a distribution results from wasting some computational power on particles that have very low likelihood.

2. Results

This subsection presents some illustrative results on the generation of the transitional prior. First, different importance sampling distributions are tested. Then, different regularization techniques are tested, using the same proposal density as in Eq. (34). The comparison is both visual and numerical: distributions whose tails decay slower and that follow the target distribution more closely are preferred. At the same time, for similar distributions, the one with the smaller uncertainty inflation is preferred.

To accentuate the effect of the thrust, in every direction the initial uncertainty is 10 m in position and 10 mm/s in velocity. The initial state is

$$\begin{bmatrix} -1,344 \text{ km} & 4,909 \text{ km} & 4,912 \text{ km} \\ -7.375 \text{ km/s} & -1.006 \text{ km/s} & -1.010 \text{ km/s} \end{bmatrix}$$

The standard deviation on the acceleration, assumed constant during the propagation, is $50 \mu\text{m/s}^2$ in each direction. The maximum standard deviation of the acceleration used in the proposal distribution is $1600 \mu\text{m/s}^2$. The propagation lasts 27,120 s, or approximately 7.53 h. Figure 5 shows the explicit transitional prior obtained with and without importance sampling. The importance sampling is here performed with a mixture of distributions as in Eq. (34). Nine distributions are used, logarithmically spaced between the minimum and the maximum. The plots, which can be considered as a reference for the true distributions, are obtained by summing the weights of

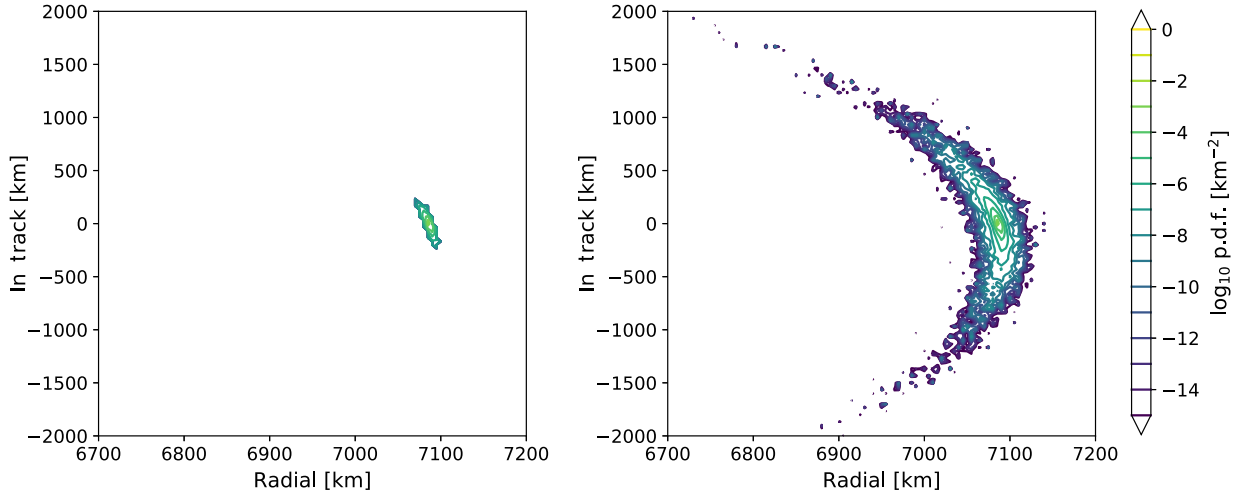


Fig. 5 Transitional prior without (left) and with (right) importance sampling.

Table 1 Standard deviations using different proposal distributions, after time update

Importance sampling	Regularization	Pos. 1σ , km	Vel. 1σ , m/s
No	Yes	2.820	3.117
	No	2.014	2.218
Only largest distribution	Yes	23.137	27.000
	No	8.315	8.709
Mixture of distributions	Yes	5.315	6.069
	No	2.893	3.086

250,000 propagated particles into the corresponding bins. The contour lines are in logarithmic scale and end where the distribution's PDF is less than 10^{-15} km^{-2} . It is clear from the figures that sampling from the original distribution fails to preserve the super-Gaussian tails. Table 1 reports the overall standard deviation obtained by the different importance sampling methods, using 2500 particles and regularizing with 99 neighbors according to the method proposed in Sec. V. The table shows that the proposed regularization moderately increases the overall uncertainty, and so does the importance sampling. Using only a large distribution for importance sampling causes excessive covariance inflation: the variance in the sample weights is much larger in that case, and a single sample receives 25% of the total probability mass. This result is an additional motivation to use a mixture of distributions.

Figure 6 shows the distributions obtained sampling according to Eq. (34), with different regularization techniques. Again, the distributions are obtained by propagating 2500 particles, and, where needed, 99 are used for regularization. Note that, although the proposed regularization approach appears more diffuse than the reference displayed in the right plot of Fig. 5, the total standard deviation does not change much before and after regularization, as evidenced in the last two rows of Table 1. The proposed approach provides tails orders of magnitude larger than both Silverman's regularization and the single Gaussian approach, while at the same time having a comparatively contained overall variance. Moreover, Silverman's approach leaves wide empty spaces; the voids are larger when looking at the actual six dimensions. Larger kernels could be used for Silverman's approach, but that would lead to larger total uncertainty. Further, a standard k -nearest-neighbor regularization approach has been tested. Although in this case the tails are wider, the shape is not appropriately represented. The figure shows how importance sampling alone is not sufficient, and adaptive kernel density estimation is a necessary tool for distributions with super-Gaussian tails. While eventually the tail decay does become sub-Gaussian with the use of any Gaussian kernels, the synergy between adaptive kernel density estimation and importance sampling leads to non sub-Gaussian tails for a large portion of the domain.

B. Filter Update

The proposed filter update combines properties of the Gaussian Sum Filters (GSFs) and of IBLS filters. The GMM obtained as in Sec. VI.A is updated using IBLS. All batches share the same multi-fidelity interpolation, and the weight of each component is updated based on its likelihood over the entire measurement pass. The batch is done over an augmented state that includes a constant acceleration over the pass. As the resulting uncertainty for each batch is Gaussian, marginalization to obtain joint position and velocity distributions of the target is straightforward.

Let the augmented state be

$$\chi^i = [\mathbf{r}^{iT} \quad \mathbf{v}^{iT} \quad \mathbf{t}_p^{iT}]^T$$

where \mathbf{t}_p^i is the thrust. The corresponding covariance is

$$P^i = \begin{bmatrix} P_{rr}^i & P_{rv}^i & 0_{3 \times 3} \\ P_{rv}^i & P_{vv}^i & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & Q_p^i \end{bmatrix}$$

where the thrust is given a prior uncertainty that is a function of the particle:

$$Q_p^i = I_{3 \times 3} \frac{\mathbf{t}_p^T \mathbf{t}_p}{3n_{\Delta V}} \quad (36)$$

The larger the thrust magnitude of the particle during the measurement gap, the larger the thrust uncertainty is through the measurement pass. For a problem where $\chi \sim \mathcal{N}(\bar{\chi}, P)$, $\mathbf{y} = \mathbf{h}(\chi) + \epsilon$, and $\epsilon \sim \mathcal{N}(0, R)$, the MMSE can be approximated by the EKF (e.g., Ref. [11]):

$$K_k^i = P_{k+1|k}^i (H_k^i)^T (H_k^i P_{k+1|k}^i (H_k^i)^T + R_k)^{-1} \quad (37)$$

$$\chi_{k+1|k+1}^i = \chi_{k+1|k}^i + K_k^i (\mathbf{h}(\chi_{k+1|k}^i) - \mathbf{y}_{k+1}) \quad (38)$$

$$P_{k+1|k+1}^i = (I - K_k^i H_k^i) P_{k+1|k}^i \quad (39)$$

where

$$H_{k+1|k}^i = \left. \frac{\partial \mathbf{h}(\chi)}{\partial \chi} \right|_{\chi=\chi_{k+1|k}^i} \quad (40)$$

A GSF update consists of several EKF updates in parallel. Their means and covariances are updated as in Eqs. (37–39), and the relative weight of each filter is also updated:

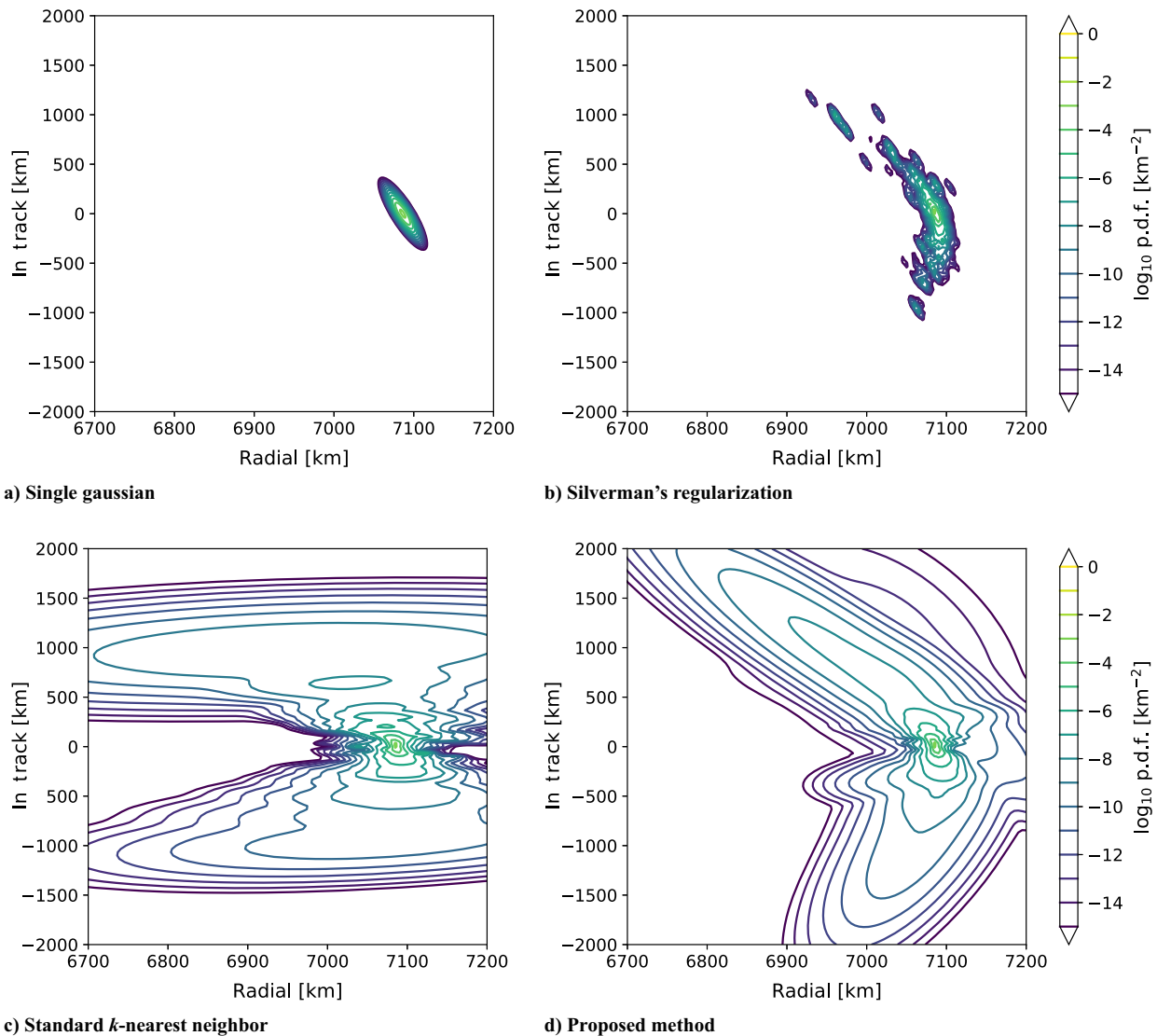


Fig. 6 Transitional prior obtained with different regularization techniques.

$$\hat{w}_{k|k}^i = \frac{w_{k|k-1}^i}{\sqrt{|2\pi(HPH^T + R)|}} e^{-\frac{1}{2}(\mathbf{h}(\mathbf{x}_{k+1|k}^i) - \mathbf{y}_{k+1})^T (HPH^T + R)^{-1} (\mathbf{h}(\mathbf{x}_{k+1|k}^i) - \mathbf{y}_{k+1})} \quad (41)$$

$$w_{k|k}^i = \frac{\hat{w}_{k|k}^i}{\sum_{j=1}^N \hat{w}_{k|k}^j} \quad (42)$$

where $w_{k|k-1}^i$ is the prior weight obtained by importance sampling.

An EGMF usually processes measurements one at a time. Here, because multiple measurements are received within a short time frame, batch processors are used instead, one for each component. To mitigate the effects of the nonlinearities, the batch is iterated. The weights are computed only at the first pass, as if the batch were not iterated, so that components with smaller weights can be pruned before iterating. The weights are obtained by propagating the distribution uncertainty linearly during the pass and computing the total likelihood of the measurement pass. For component i , the posterior weight after processing the batch of observations is

$$\hat{w}_{k|k}^i = w_{k|k-1}^i \prod_{j=0}^{n_{p,k}} e^{-\frac{1}{2}(\bar{\mathbf{y}}_{k_j}^i - \mathbf{y}_{k_j})^T (\mathbf{H}_{k_j}^i \mathbf{P}_{k_j|k-1}^i (\mathbf{H}_{k_j}^i)^T + \mathbf{R}_{k_j})^{-1} (\bar{\mathbf{y}}_{k_j}^i - \mathbf{y}_{k_j})} \sqrt{|2\pi(\mathbf{H}_{k_j}^i \mathbf{P}_{k_j|k-1}^i (\mathbf{H}_{k_j}^i)^T + \mathbf{R}_{k_j})|} \quad (43)$$

where $n_{p,k}$ is the number of measurements in the k th pass, the subscript k_j stands for the j th measurement of the k th pass, $\bar{\mathbf{y}}_{k_j}^i = \mathbf{h}(\mathbf{x}_{k_j|k-1}^i)$ is the expected measurement for component i at measurement k_j , and

$$\mathbf{P}_{k_j|k-1}^i = \mathbf{F}_{k_j,k_0}^i \mathbf{P}_{k_0|k-1}^i (\mathbf{F}_{k_j,k_0}^i)^T \quad (44)$$

where \mathbf{F}_{k_j,k_0}^i is the state transition matrix of component i from the start of the pass to k_j . While multi-fidelity state transition matrix propagation has been developed [53], finite difference methods have been used here. Any component whose posterior weight is less than 10^{-10} times the maximum weight is pruned after the first iteration. After the lower likelihood components are pruned, the update of mean and variance of the mixands can be performed. Also in this phase, multi-fidelity propagation is used.

The process is iterated in order to compute the partials F and H as closely as possible to the actual estimate. Even having the exact partials over the estimate is in theory not sufficient to remove bias from the filter. However, the fact that those partials are exact where the probability density is highest improves the estimate.

Two multi-fidelity interpolations are computed, both along the entirety of the pass. The first one is computed for the full Gaussian mixture and is used for the first iteration only. The second multi-fidelity interpolation is computed for the second iteration, using only the nonpruned components, and is kept for all subsequent iterations.

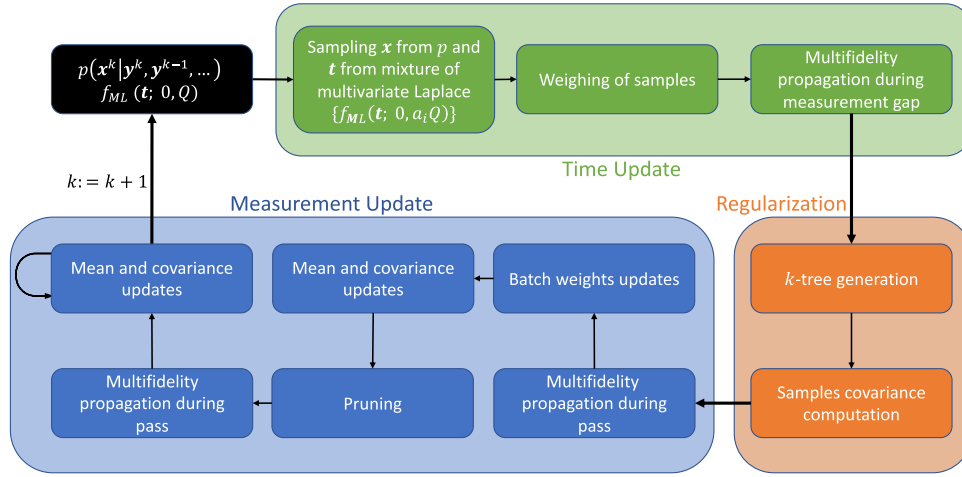


Fig. 7 Overall architecture of the filter.

After the first iteration, two things happen concurrently: 1) many components are pruned, and therefore the interpolation domain is reduced, and 2) the majority of the update for all other components is already applied, driving the interpolation domain toward to the more important region of the state space. These two facts help improve the approximation. Although additional multi-fidelity interpolations may improve the propagation, every different interpolation has different errors, therefore causing the iterative process to oscillate.

C. Filter Summary

This section summarizes the schematics of the filter, depicted in Fig. 7. At the end of a measurement pass, random target states $[\mathbf{x}^i, \mathbf{v}^i]^T$, $i = 1, \dots, N$ are sampled from an initial distribution. Each sample is associated to a thrust profile \mathbf{t}_k^i , sampled from a sampling distribution that is the mixture described in Eq. (34), and are given the prior weight w_i as in Eq. (35). The samples are propagated with the low-fidelity model until the next measurement. Out of those, important samples are selected according to the multi-fidelity stochastic collocation algorithm and are propagated with high-fidelity model. All low-fidelity samples are corrected with the multi-fidelity stochastic collocation method. Then, a 6-dimensional tree is generated for efficient k -nearest-neighbor search. The tree is generated using states transformed into canonical units. Every particle is transformed into a Gaussian distribution, with covariances computed according to Eqs. (27) and (28). Each mean and covariance is propagated with state transition matrices for the duration of the pass, using again the stochastic collocation multi-fidelity approach. Expected measurements and measurement Jacobians are computed for every component at each measurement of the pass. The weights of every component are computed according to Eq. (43) and normalized. Components with relative weights smaller than a given threshold are pruned. All components are updated for one round of batch least squares. After the first update, a new stochastic collocation multi-fidelity interpolation is generated over the pass, and all remaining components are iteratively updated with a batch. The batch is performed over an augmented state that includes a constant acceleration. The prior variance of the acceleration is computed according to Eq. (36). The second multi-fidelity stochastic collocation interpolation is then held constant during the remaining interpolation of the batch least squares. Finally, one obtains the updated augmented state and covariance of each component at the beginning of the measurement pass, together with a constant (in the RIC frame) thrust vector, which need to be propagated to the final epoch of the pass. At this point, one has a posterior distribution at the last measurement of a pass and can proceed forward repeating the algorithm for the next measurement gap.

VII. Results

The proposed algorithm has been tested on several scenarios. All scenarios share the same initial conditions, same true dynamical

models, same filter propagation dynamical models, same measurement models, and same ground station network. The only differences consist of the thrust profiles and the observation passes frequency. The initial conditions are distributed as in Table 2. The model used for the truth is detailed in Table 3 with additional parameters in Table 4. Because in reality one never has a propagation model that is the same as the true physics, the filter's high-fidelity model used in this work neglects drag and solar radiation pressure (SRP). For context, these accelerations in the tested cases are no larger than $0.5 \mu\text{m/s}^2$. There is

Table 2 Orbit state initial conditions and 1σ uncertainty

Variables	a , km	e , -	i , deg	Ω , deg	ω , deg	ν , deg
	7,078	7.7×10^{-4}	45	0	90	0
Variables	x , km	y , km	z , km	\dot{x} , m/s	\dot{y} , m/s	\dot{z} , m/s
	0.0	5,001.048	5,001.048	-7,510.139	0.0	0.0
σ	1.0	1.0	1.0	1.0	1.0	1.0

Table 3 Force models for low- and high-fidelity propagators and for the truth

Model	Low fidelity	High fidelity	Truth
Central body gravity	Two-body, J_2	70×70 [54]	70×70 [54]
Third-body perturbations	None	Sun, moon [55]	Sun, moon [55]
Solar radiation pressure	None	None	Cannonball
Atmospheric drag	None	None	Cannonball
Coordinate System	None	IAU2006 [56]	IAU2006 [56]
Reduction			

Table 4 Force model and satellite parameters for all cases

Parameter	Value
Satellite mass	500 kg
Satellite drag and SRP area	1m^2
C_D	2.0
C_R	1.5
Epoch time	1 January 2000, 12:00
Earth radius R_\oplus	6378.1363 km
Gravitation parameter	$398600.4415 \text{ km}^3/\text{s}^2$

no additional process noise in the filter to account specifically for model mismatch, as the resulting uncertainty is generally smaller than the uncertainty introduced by the maneuvers. The low- and high-fidelity models used for propagation are summarized in Table 3; details can be found in Refs. [54–56].

Three ground stations generate radar measurements. One ground station is located at 44.77° latitude, 83.65° longitude, and at a distance from Earth's center of R_e . The second and third stations have the same longitude and radius, but latitudes of 0° and -44.77° , respectively. Radar measurements are generated every 60 s when two conditions are met: 1) the spacecraft is visible from one of three stations, and 2) when the time from the latest pass is larger than a specified value, dependent on the scenario. The latter condition is included to force the filter to work under conditions of sparse data. During the k th pass, a set of measurements $Y_k = [y_{k_0}, \dots, y_{k_{n_{p,k}}}]$ is obtained, where $n_{p,k}$ is the number of measurements during pass k and $y_{k_i} = [r_{k_i}, \dot{r}_{k_i}, \alpha_{k_i}, \delta_{k_i}]^T$, where r is the range, \dot{r} is the range rate, and α and δ are the topocentric right ascension and declination of the position, respectively. The measurement variance R is

$$R = \text{diag}([(0.1 \text{ km})^2, (0.003 \text{ km/s})^2, (0.15 \text{ deg})^2, (0.15 \text{ deg})^2])$$

and measurement errors are uncorrelated in time.

Except for Sec. VII.C.3, all tested filters use 2500 propagated samples and perform regularization with 99 neighbors. The samples are drawn from a mixture of nine multivariate Laplace distributions with diagonal variance; their standard deviations are equally spaced in logarithmic scale. The minimum standard deviation used is equal to the filter's set 1σ standard deviation, and the maximum is set to $800 \mu\text{m/s}^2$. The distribution with smallest uncertainty has weight 0.3, and all others have equal weight. The iterative batch least squares performs 10 iterations; at each iteration, the estimate is updated by a vector that is the computed updated, multiplied by 0.7.

Four scenarios are tested. Every scenario may include different cases; each case is tested over a Monte Carlo of 50 trajectories. The first scenario, in Sec. VII.A, is a consistency check. The scenario in Sec. VII.B is a constant in-track thrust case, during which the target spirals outward. Tests include different values of the thrust and different data gap durations. The results are compared against another filter from literature. The third scenario, described in Sec. VII.C, is an apogee raising maneuver, which has an intermittent in-track thrust profile, causing the satellite to be harder to track. This scenario is tested over different values of the target's thrust and the filter's assumed thrust, to prove robustness over the filter's parameter. The same scenario is tested in a maneuver initiation case, to show the capability of the approach to react to the beginning of a maneuvering phase, and in cases with even sparser observation passes, as few as one every 16 h, to stretch the filter's capabilities. Finally, in Sec. VII.D, the filter is tested for cases where the thrust is directed in all three directions and changes almost continuously.

A. Consistency Check: Random Thrust

This case is to be considered a toy problem and is used to evaluate the consistency of the proposed filter when its assumptions are met. To this end, the true thrust is generated in a way that is as similar as possible to how the filter samples the thrust profiles. Note that for all other scenarios the filter cannot be consistent. That is because in those scenarios all true trajectories in a same Monte Carlo trial share the same thrust profile.

The actual thrust and the filter's thrust samples are drawn from the same distribution, a three-dimensional multivariate Laplace with standard deviation equal to $50 \mu\text{m/s}^2$ per direction. The true thrust is sampled every 16,800 s, the approximate average time between each measurement pass, and is constant for that duration. Accordingly, only for this scenario and the next, the filter assumes constant thrust between every measurement pass, but magnitude and directions are unknown. Figure 8 displays histograms of the Z-scores of the estimates, the ratios between actual error and the filter's estimated uncertainty. The statistical mean is the black vertical line, and the 1σ statistical standard

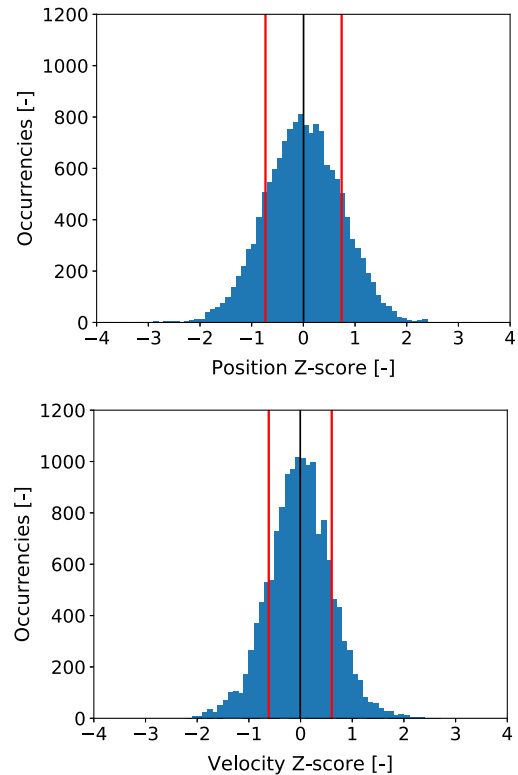


Fig. 8 Z-score histograms for Sec. VII.A and corresponding statistical means and standard deviations.

deviation is marked by the red lines. For a perfectly consistent filter, the latter should be equal to 1. The filter's standard deviation of the error is slightly conservative, as expected, because the regularization inflates the uncertainty of the transitional prior. Most importantly, because there are no large systematic differences between the physics model and the filter's model, the distributions have biases much smaller than their standard deviation. The RMSE of the position is 507 m, and the RMSE of the velocity is 0.978 m/s. Fewer than 0.1% of estimates fall outside of the 3σ bounds.

B. Constant Thrust

In this more realistic scenario, the thrust is constant during the entire simulation and directed in track: the proposed filter does not know the magnitude of the thrust, nor its direction, and assumes constant thrust within every measurement gap, like in Sec. VII.A. The latter assumption is dropped in the next subsection. Only for this scenario, the filter is compared against the variable-state dimension filter (VSDF) proposed by Goff et al. [15]. To reduce complexity, the single model version of the VSDF is used; the maneuver standard deviation for the VSDF is set equal to the magnitude of the true acceleration. The proposed filter here assumes a thrust of $1 \mu\text{m/s}^2$. This gives the VSDF an advantage based solely on tuning parameters. Note that a second advantage for the VSDF comes from the fact that in this scenario the true thrust is constant; while the proposed filter can handle varying thrust, the VSDF cannot.

The filters have been tested for a variety of true thrust magnitudes and for cases with minimum measurement gaps of 15,000 and 20,000 s as shown in Fig. 9. Each data point is obtained from a Monte Carlo trial of 50 trajectories, and every simulation lasts 48 h. Because of the fact that the VSDF linearizes models to approximate the distributions, it has at least one diverging trajectory for any of the cases with 20,000 s measurement gaps, as well as the case with 15,000 s and $600 \mu\text{m/s}^2$ acceleration. Therefore, those results are not reported. On the other hand, the proposed filter can track the target in all cases, albeit some signs of divergence occur when the true thrust is $600 \mu\text{m/s}^2$ and the filter assumes $1 \mu\text{m/s}^2$ acceleration. Further, for the cases in which the VSDF does not diverge, it is always outperformed by the proposed

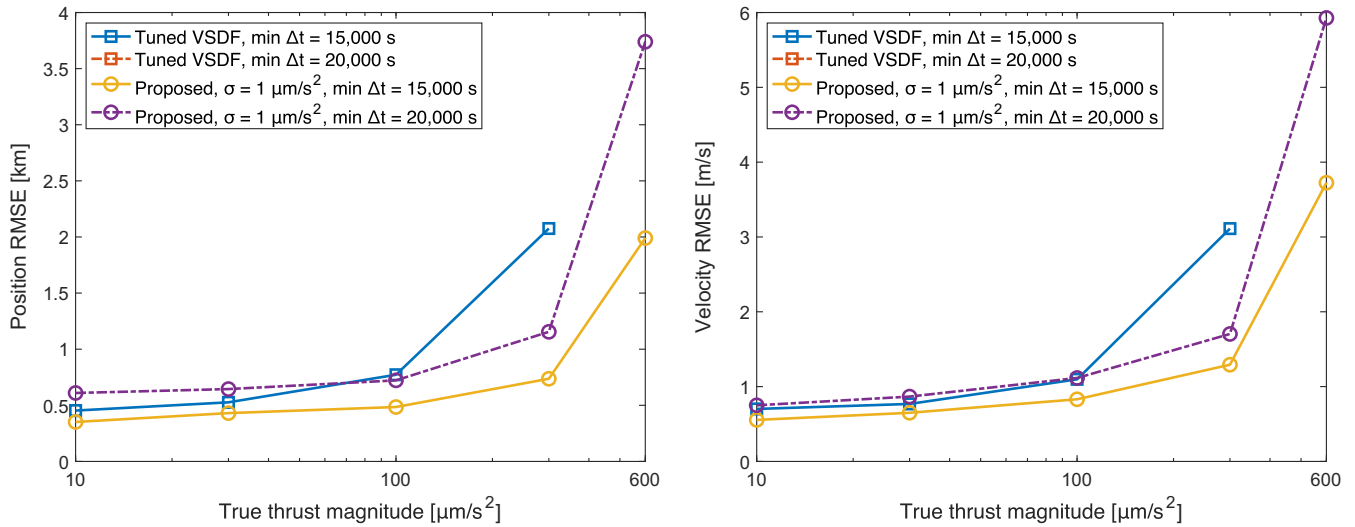


Fig. 9 RMSE for the filters applied to the scenario of Sec. VII.B.

filter. Figure 10 shows the errors estimates of the proposed filter when the true thrust is $300 \mu\text{m/s}^2$. The plot skips over the measurement gaps; measurement passes start and end at, respectively, red and green dashed vertical lines. The blue and cyan dashed lines are, respectively, the maximum and average (over the 50 Monte Carlo runs) 3σ uncertainties expected by the filter at the given time stamp. The plot shows that the filter successfully tracks the maneuvering object, and almost all estimates stay within the 3σ bounds.

C. Apogee Raising Maneuver

In this third scenario, the spacecraft performs an apogee raising maneuver. The intermittent, in-track thrust is on and constant for half of an orbit and off for the remaining half. An example of the orbital radius of the true trajectory is shown in Fig. 11. Because of the intermittency of the thrust, combined with the physics of the problem, filters that assume constant thrust during a measurement gap (like the ones from Sec. VII.B) perform very poorly when applied to this problem. For example, with a true acceleration of $300 \mu\text{m/s}^2$ and an assumed acceleration of $25 \mu\text{m/s}^2$, a proposed filter that assumes constant thrust produces more than one third of the estimates exceeding the 3σ bounds, and the 20σ bounds are reached. For this scenario, the number of thrust segments in a measurement gap is chosen as

$$n_{\Delta V} = \text{floor} \frac{4\Delta t}{2\pi} \sqrt{\frac{\mu}{a^3}} \quad (45)$$

where Δt is the time between the end of the previous measurement pass and the start of the next, and a is the semimajor axis of the

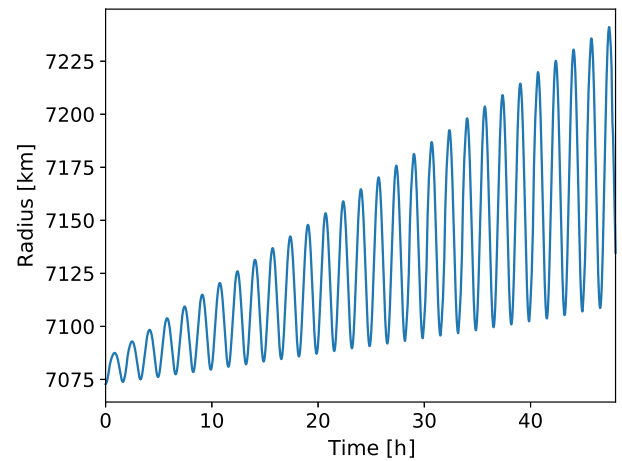


Fig. 11 Example radius of one of the trajectories from Sec. VII.C.

osculating orbit at the start of the measurement gap. Hence, there are approximately four segments per orbit.

1. Base Apogee Raising Scenario

Direction, thrust profile, and thrust magnitude are unknown to the filter. The filter additionally assumes that the thrust is repeated every four thrust segments within the same measurement gap. This assumption is generally acceptable for relatively short measurement gaps, and it is dropped for cases with larger gaps as in Sec. VII.C.3.

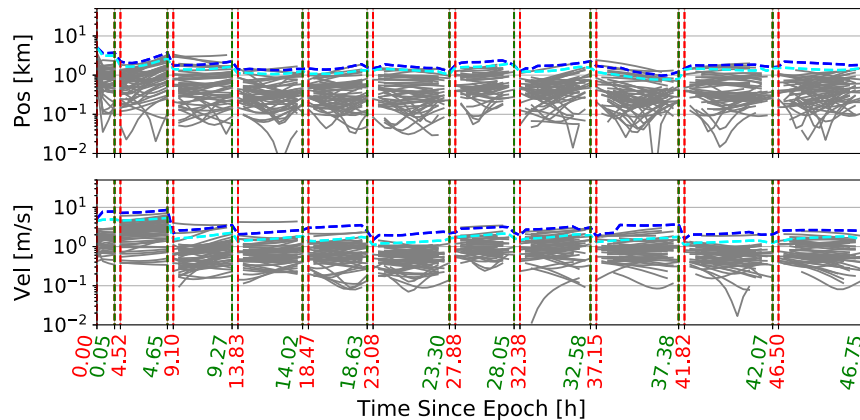


Fig. 10 Errors for the proposed filter, true thrust constant at $300 \mu\text{m/s}^2$.

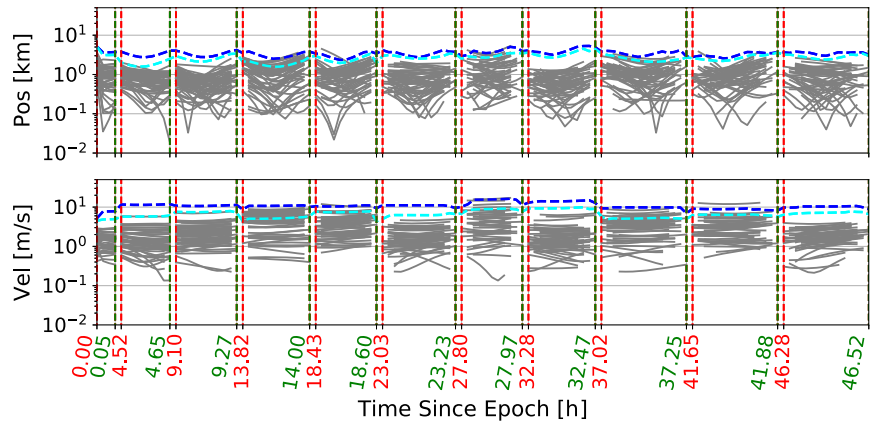


Fig. 12 Estimate errors for case B1 of the scenario from Sec. VII.C.1.

Like in all previous scenarios, there is no assumption of repeating the thrust profiles between different measurement gaps.

Figure 12 shows the error and corresponding maximum and average filter 3σ estimate uncertainty for filter's standard deviation of $1\text{ }\mu\text{m/s}^2$, and with true target acceleration of $300\text{ }\mu\text{m/s}^2$. Comparing the plot with Fig. 10, it is possible to obtain confirmation that this scenario is more challenging than the previous one. In fact, the RMSE is 52% larger in the position and 169% larger in the velocity. Further, the number of outliers is greatly increased, from a negligible number to more than 7%.

Table 5 summarizes the performance of the filters with different settings and different true maneuvers. The ratio between actual thrust and the filter's standard deviations spans three orders of magnitude. The general trend is that the filter is more robust, and more accurate, whenever the actual and expected thrust are most similar in magnitude. For same filter settings, larger thrust leads to larger errors and larger

uncertainties, together with a larger number of outlier estimates. For the highest thrust case, the two filters with standard deviation of 1 and $6.25\text{ }\mu\text{m/s}^2$ have an amount of outliers that exceed the expected amount even for univariate Laplace distributions, which is 1.5%; however, that can be considered acceptable given the systematic error inherent to the scenario. Moreover, as shown in Fig. 12, even the most challenging case does not diverge. Last, given everything else equal, filters whose 1σ uncertainty differs by two orders of magnitude provide estimates whose accuracy differs by less than 30%, showing that the proposed filter is largely insensitive to the process noise tuning.

2. Maneuver Initiation

In this scenario, the target flies ballistically for the first 24 h, and only then the apogee raising maneuver begins. The goal of this test is to show that the filter can also automatically switch between maneuvering and nonmaneuvering situations and does not require any additional or external maneuver detection capabilities to operate. Figure 13 shows position and velocity errors for the extreme case with $1\text{ }\mu\text{m/s}^2$ process noise standard deviation and $300\text{ }\mu\text{m/s}^2$ actual acceleration. The uncertainty initially shrinks because the steady-state covariance for the nonmaneuvering problem is smaller than the initial uncertainty. Once the maneuver begins, the uncertainty increases accordingly.

3. Very Sparse Observations

The filter's capability is further stretched with longer data gaps. The frequency of measurement passes is reduced such that gaps are larger than 12 h in case a and 16 h in case b. The simulation lasts 5 days. As before, the true thrust is intermittent, and the filter assumes continuous segments of thrust of about one-fourth of an orbit. Differently from before, now the thrust is not repeated between orbits, the filter's standard deviation is kept to $1\text{ }\mu\text{m/s}^2$, and the filter uses 10,000 particles, regularized using the 199 nearest neighbors.

Table 5 Filters performance over 50 Monte Carlo trials for the scenario from Sec. VII.C.1

Case	True magnitude, $\mu\text{m/s}^2$	Filter 1σ , $\mu\text{m/s}^2$	Pos. RMSE, m	Vel. RMSE, m/s	Percent outside 3σ
A1	30	1	427	0.773	0.2
A2	30	6.25	420	0.753	<0.1
A3	30	25	414	0.741	<0.1
A4	30	100	524	0.989	<0.1
B1	300	1	1117	3.472	7.7
B2	300	6.25	1047	3.118	4.9
B3	300	25	944	2.678	1.6
B4	300	100	870	2.483	0.2

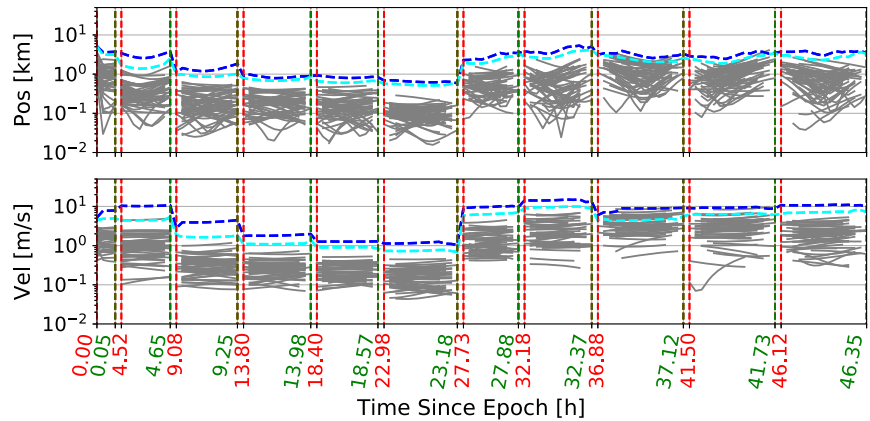


Fig. 13 Estimate errors for the scenario of Sec. VII.C.2.

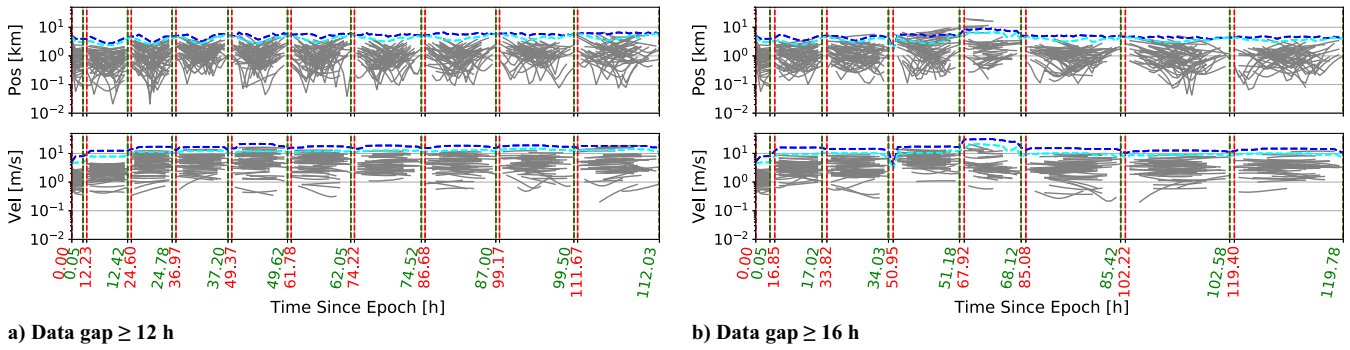


Fig. 14 Estimate errors for the scenario from Sec. VII.C.3.

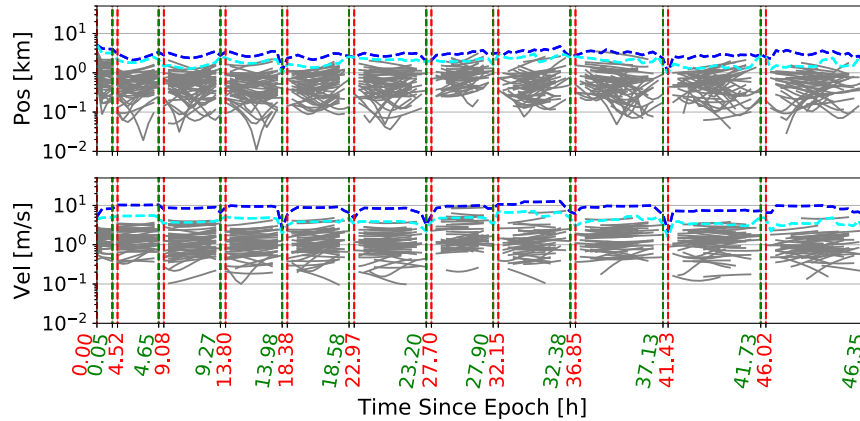


Fig. 15 Estimate errors for the scenario of Sec. VII.D.

Figure 14 shows the errors for a Monte Carlo trial tracking an object accelerating with thrust of $200 \mu\text{m/s}^2$. For case a all trajectories are tracked successfully, with 2.1% of outliers estimates outside the 3σ bounds. For case b, one out of 50 trajectories diverges, not reported in the plot; among the converging trajectories, the number of outliers outside 3σ is 3.6%; the position RMSE is 2.091 km, and the velocity RMSE is 4.522 m/s. With same data gaps as case a but larger true thrust of $300 \mu\text{m/s}^2$, one out of 50 trajectories diverges after 10 days of simulation. The capabilities of the filter can generally be improved by increasing the number of particles, while sublinearly increasing the number of neighbors used for regularization. For example, when tracking case a with 2500 particles, the outliers are 10%, suggesting a trend of improving results with increasing sample number.

D. Quickly Varying Thrust

Because all previous scenarios only involve in-track only true thrust, either constant or intermittent, one scenario is included where the target accelerates with an almost continuously varying thrust in all three directions. The true thrust is drawn from a random process with constant segments of 60 s only. The standard deviation of the true thrust is $100 \mu\text{m/s}^2$ in each direction, with an autocorrelation factor of 0.99. This means that portions of thrust profiles can be much larger than $300 \mu\text{m/s}^2$. The filter used is the same as Sec. VII.C.2. Figure 15 shows the results for a Monte Carlo of 50 trajectories. In every trajectory, the filter can keep custody of the target, with few estimates falling outside of the 3σ bounds. The RMSE is 763 m in position and 1.7 m/s in velocity.

VIII. Conclusions

This paper demonstrates the possibility of using Bayesian estimation to track low-thrust maneuvering spacecraft with sparse observations, while providing an explicit transitional prior. Having an explicit transitional prior is a strict requirement for implementation of any tracking technique into Bayesian multitarget tracking frameworks. The proposed method relies on multivariate Laplace distributions for the

maneuvers and focuses on the nonlinear transformation of the super-Gaussian tails of the distribution. This is achieved using principles from rare event simulation and through a novel k -nearest-neighbor ensemble Gaussian mixture filter. Multi-fidelity methods are exploited to allow fast and efficient propagation. The resulting algorithm is robust to a variety of maneuver profiles and magnitudes and is in large part insensitive to the tuning parameters. Further, the proposed approach does not require any additional maneuver detection techniques. The filter is able to maintain custody of targets that perform complex maneuvers involving intermittent thrust, even when the magnitude is up to 300 times larger than expected and when measurement passes are sparse, one every 2.5 orbital revolutions or more. When observations are even sparser, fewer than two per day, the algorithm still maintains custody of targets with thrust magnitude up to 200 times larger than expected, at $200 \mu\text{m/s}^2$.

Acknowledgments

This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-19-1-0404. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Air Force.

References

- [1] Reid, D. B., "An Algorithm for Tracking Multiple Targets," *IEEE Transactions on Automatic Control*, Vol. 24, No. 6, 1979, pp. 843–854. <https://doi.org/10.1109/TAC.1979.1102177>
- [2] Vo, B.-N., Vo, B.-T., and Hoang, H. G., "An Efficient Implementation of the Generalized Labeled Multi-Bernoulli Filter," *IEEE Transactions on Signal Processing*, Vol. 65, No. 8, 2017, pp. 1975–1987. <https://doi.org/10.1109/TSP.2016.2641392>
- [3] Ravago, N., and Jones, B. A., "Risk-Aware Sensor Scheduling and Tracking of Large Constellations," *Advances in Space Research*, Vol. 68, No. 6, 2021, pp. 2530–2550. <https://doi.org/10.1016/j.asr.2021.04.042>

- [4] Holzinger, M. J., Scheeres, D. J., and Alfriend, K. T., "Object Correlation, Maneuver Detection, and Characterization Using Control Distance Metrics," *Journal of Guidance, Control, and Dynamics*, Vol. 35, No. 4, 2012, pp. 1312–1325.
<https://doi.org/10.2514/1.53245>
- [5] Escribano, G., Sanjurjo-Rivo, M., Siminski, J., Pastor, A., and Escobar, D., "Automatic Maneuver Detection and Tracking of Space Objects in Optical Survey Scenarios Based on Stochastic Hybrid Systems Formulation," *Advances in Space Research*, Vol. 69, No. 9, 2022, pp. 3460–3477.
<https://doi.org/10.1016/j.asr.2022.02.034>
- [6] Li, X. R., and Jilkov, V. P., "Survey of Maneuvering Target Tracking. Part I. Dynamic Models," *IEEE Aerospace and Electronic Systems Magazine*, Vol. 39, No. 4, 2003, pp. 1333–1364.
<https://doi.org/10.1109/TAES.2003.1261132>
- [7] Li, X. R., and Jilkov, V. P., "Survey of Maneuvering Target Tracking. Part II: Motion Models of Ballistic and Space Targets," *IEEE Aerospace and Electronic Systems Magazine*, Vol. 46, No. 1, 2010, pp. 96–119.
<https://doi.org/10.1109/TAES.2010.5417150>
- [8] Li, X. R., and Jilkov, V. P., "Survey of Maneuvering Target Tracking: III. Measurement Models," *Signal and Data Processing of Small Targets 2001*, Vol. 4473, Soc. of Photo-Optical Instrumentation Engineers, Bellingham, WA, 2001, pp. 423–446.
<https://doi.org/10.1117/12.492752>
- [9] Li, X. R., and Jilkov, V. P., "Survey of Maneuvering Target Tracking: Decision-Based Methods," *Signal and Data Processing of Small Targets 2002*, Vol. 4728, Soc. of Photo-Optical Instrumentation Engineers, Bellingham, WA, 2002, pp. 511–534.
<https://doi.org/10.1117/12.478535>
- [10] Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, Vol. 82, No. 1, 1960, pp. 35–45.
<https://doi.org/10.1115/1.3662552>
- [11] Bar-Shalom, Y., Li, X. R., and Kirubarajan, T., *Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software*, Wiley, New York, 2004, pp. 371–420, Chap. 10.
- [12] Gholson, N. H., and Moose, R. L., "Maneuvering Target Tracking Using Adaptive State Estimation," *IEEE Aerospace and Electronic Systems Magazine*, Vol. 13, No. 3, 1977, pp. 310–317.
<https://doi.org/10.1109/TAES.1977.308399>
- [13] Efe, M., and Atherton, D., "Maneuvering Target Tracking with an Adaptive Kalman Filter," *Proceedings of the 37th IEEE Conference on Decision and Control (Cat. No.98CH36171)*, Vol. 1, Inst. of Electrical and Electronics Engineers, New York, 1998, pp. 737–742.
<https://doi.org/10.1109/CDC.1998.760773>
- [14] Bar-Shalom, Y., and Birmiwal, K., "Variable Dimension Filter for Maneuvering Target Tracking," *IEEE Aerospace and Electronic Systems Magazine*, Vol. 18, No. 5, 1982, pp. 621–629.
<https://doi.org/10.1109/TAES.1982.309274>
- [15] Goff, G. M., Black, J. T., and Beck, J. A., "Orbit Estimation of a Continuously Thrusting Spacecraft Using Variable Dimension Filters," *Journal of Guidance, Control, and Dynamics*, Vol. 38, No. 12, 2015, pp. 2407–2420.
<https://doi.org/10.2514/1.G001091>
- [16] Pastor, A., Escribano, G., Sanjurjo-Rivo, M., and Escobar, D., "Satellite Maneuver Detection and Estimation with Optical Survey Observations," *Journal of the Astronautical Sciences*, Vol. 69, No. 3, 2022, pp. 879–917.
<https://doi.org/10.1007/s40295-022-00311-5>
- [17] Porcelli, L., Pastor, A., Cano, A., Escribano, G., Sanjurjo-Rivo, M., Escobar, D., and Di Lizia, P., "Satellite Maneuver Detection and Estimation with Radar Survey Observations," *Acta Astronautica*, Vol. 201, Dec. 2022, pp. 274–287.
<https://doi.org/10.1016/j.actaastro.2022.08.021>
- [18] Lubey, D. P., and Scheeres, D. J., "Identifying and Estimating Mismodeled Dynamics via Optimal Control Policies and Distance Metrics," *Journal of Guidance, Control, and Dynamics*, Vol. 37, No. 5, 2014, pp. 1512–1523.
<https://doi.org/10.2514/1.G000369>
- [19] Pirovano, L., and Armellini, R., "Detection and Estimation of Spacecraft Maneuvers for Catalog Maintenance," *Acta Astronautica*, Vol. 215, Feb. 2024, pp. 387–397.
<https://doi.org/10.1016/j.actaastro.2023.12.016>
- [20] Rong Li, X., and Jilkov, V. P., "Survey of Maneuvering Target Tracking. Part V. Multiple-Model Methods," *IEEE Aerospace and Electronic Systems Magazine*, Vol. 41, No. 4, 2005, pp. 1255–1321.
<https://doi.org/10.1109/TAES.2005.1561886>
- [21] Goff, G. M., Black, J. T., and Beck, J. A., "Tracking Maneuvering Spacecraft with Filter-Through Approaches Using Interacting Multiple Models," *Acta Astronautica*, Vol. 114, Sept. 2015, pp. 152–163.
<https://doi.org/10.1016/j.actaastro.2015.05.009>
- [22] Zucchelli, E. M., McLaughlin, Z. R., and Jones, B. A., "Tracking Maneuvering Targets with Multi-Fidelity Interacting Multiple Model Filters," *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS)*, Curran Assoc., Inc., Red Hook, NY, 2020, pp. 279–292.
- [23] Roth, M., Özkan, E., and Gustafsson, F., "A Student's t Filter for Heavy Tailed Process and Measurement Noise," *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Inst. of Electrical and Electronics Engineers, New York, 2013, pp. 5770–5774.
<https://doi.org/10.1109/ICASSP.2013.6638770>
- [24] Huang, Y., Zhang, Y., Shi, P., Wu, Z., Qian, J., and Chambers, J. A., "Robust Kalman Filters Based on Gaussian Scale Mixture Distributions with Application to Target Tracking," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 49, No. 10, 2019, pp. 2082–2096.
<https://doi.org/10.1109/TSMC.6221021>
- [25] Ikoma, N., Higuchi, T., and Maeda, H., "Tracking of Maneuvering Target by Using Switching Structure and Heavy-Tailed Distribution with Particle Filter Method," *Proceedings of the International Conference on Control Applications*, Vol. 2, Inst. of Electrical and Electronics Engineers, New York, 2002, pp. 1282–1287.
<https://doi.org/10.1109/CCA.2002.1038790>
- [26] Zucchelli, E. M., and Jones, B. A., "A Gaussian Integral Filter with Multivariate Laplace Process Noise," *IEEE International Conference on Information Fusion*, Inst. of Electrical and Electronics Engineers, New York, June 2023.
<https://doi.org/10.23919/FUSION52260.2023.10224147>
- [27] L'Ecuyer, P., Mandjes, M., and Tuffin, B., *Rare Event Simulation Using Monte Carlo Methods*, Vol. 2, Wiley, Hoboken, NJ, 2009, pp. 17–38, Chap. 2.
<https://doi.org/10.1007/s00607-002-1469-6>
- [28] Beck, J. L., and Zuev, K. M., "Rare Event Simulation," *Handbook of Uncertainty Quantification*, edited by R. Ghanem, D. Higdon, and H. Owhadi, Springer International Publ., Cham, 2016, pp. 1–26.
https://doi.org/10.1007/978-3-319-11259-6_24-1
- [29] Peherstorfer, B., Willcox, K., and Gunzburger, M., "Survey of Multi-fidelity Methods in Uncertainty Propagation, Inference, and Optimization," *SIAM Review*, Vol. 60, No. 3, 2018, pp. 550–590.
<https://doi.org/10.48550/arXiv.1806.10761>
- [30] Liu, J. S., and Chen, R., "Sequential Monte Carlo Methods for Dynamic Systems," *Journal of the American Statistical Association*, Vol. 93, No. 443, 1998, pp. 1032–1044.
<https://doi.org/10.1080/01621459.1998.10473765>
- [31] Sorenson, H., and Alspach, D., "Recursive Bayesian Estimation Using Gaussian Sums," *Automatica*, Vol. 7, No. 4, 1971, pp. 465–479.
[https://doi.org/10.1016/0005-1098\(71\)90097-5](https://doi.org/10.1016/0005-1098(71)90097-5)
- [32] Narayan, A., Gittelsohn, C., and Xiu, D., "A Stochastic Collocation Algorithm with Multi-Fidelity Models," *SIAM Journal on Scientific Computing*, Vol. 36, No. 2, 2014, pp. A495–A521.
<https://doi.org/10.1137/130929461>
- [33] Jones, B. A., and Weisman, R., "Multi-Fidelity Orbit Uncertainty Propagation," *Acta Astronautica*, Vol. 155, Feb. 2019, pp. 406–417.
<https://doi.org/10.1016/j.actaastro.2018.10.023>
- [34] Zucchelli, E. M., Delande, E. D., Jones, B. A., and Jah, M. K., "Multi-Fidelity Orbit Determination with Systematic Errors," *Journal of the Astronautical Sciences*, Vol. 68, No. 3, 2021, pp. 695–727.
<https://doi.org/10.1007/s40295-021-00267-y>
- [35] Wolf, T. N., Zucchelli, E. M., and Jones, B. A., "Multi-Fidelity Uncertainty Propagation for Objects in Cislunar Space," *AIAA SciTech Forum*, AIAA Paper 2022-1774, 2022.
<https://doi.org/10.2514/6.2022-1774>
- [36] Vinti, J. P., "New Method of Solution for Unretarded Satellite Orbits," *Journal of Research of the National Bureau of Standards*, Vol. 63, No. 2, 1959, pp. 105–116.
<https://doi.org/10.6028/jres.063B.012>
- [37] Wolf, T. N., and Jones, B. A., "Extending the Utility of Multi-Fidelity Space Object Uncertainty Quantification with Vinti Theory," *AAS/AIAA Space Flight Mechanics Meeting*, American Astronautical Soc. Paper 23-224, Springfield, VA, 2023.
- [38] Harbrecht, H., Peters, M., and Schneider, R., "On the Low-Rank Approximation by the Pivoted Cholesky Decomposition," *Applied Numerical Mathematics*, Vol. 62, No. 4, 2012, pp. 428–440.
<https://doi.org/10.1016/j.apnum.2011.10.001>
- [39] Bebendorf, M., and Rjasanow, S., "Adaptive Low-Rank Approximation of Collocation Matrices," *Computing*, Vol. 70, No. 1, 2003, pp. 1–24.
<https://doi.org/10.1007/s00607-002-1469-6>
- [40] Ko, H. C., and Scheeres, D. J., "Essential Thrust-Fourier-Coefficient Set of Averaged Gauss Equations for Orbital Mechanics," *Journal of Guidance, Control, and Dynamics*, Vol. 37, No. 4, 2014, pp. 1236–1249.
<https://doi.org/10.2514/1.62407>

- [41] Elandt-Johnson, R. C., and Johnson, N. L., *Survival Models and Data Analysis*, Vol. 110, Wiley, New York, 1980, pp. 9–49, Chap. 2. <https://doi.org/10.1002/9781119011040.ch2>
- [42] Pitt, M. K., and Shephard, N., “Filtering via Simulation: Auxiliary Particle Filters,” *Journal of the American Statistical Association*, Vol. 94, No. 446, 1999, pp. 590–599. <https://doi.org/10.1080/01621459.1999.10474153>
- [43] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., “Equation of State Calculations by Fast Computing Machines,” *Journal of Chemical Physics*, Vol. 21, No. 6, 1953, pp. 1087–1092. <https://doi.org/10.1063/1.1699114>
- [44] Liu, B., Ait-El-Fquih, B., and Hoteit, I., “Efficient Kernel-Based Ensemble Gaussian Mixture Filtering,” *Monthly Weather Review*, Vol. 144, No. 2, 2016, pp. 781–800. <https://doi.org/10.1175/MWR-D-14-00292.1>
- [45] Hürzeler, M., and Künsch, H. R., “Monte Carlo Approximations for General State-Space Models,” *Journal of Computational and Graphical Statistics*, Vol. 7, No. 2, 1998, pp. 175–193. <https://doi.org/10.1080/10618600.1998.10474769>
- [46] Oudjane, N., and Musso, C., “Progressive Correction for Regularized Particle Filters,” *Proceedings of the 3rd International Conference on Information Fusion*, Vol. 2, Inst. of Electrical and Electronics Engineers, New York, 2000, pp. THB2/10–THB2/17. <https://doi.org/10.1109/IFIC.2000.859873>
- [47] Silverman, B. W., *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London, U.K., 1986, pp. 75–93, Chap. 4.
- [48] Yun, S., Zanetti, R., and Jones, B. A., “Kernel-Based Ensemble Gaussian Mixture Filtering for Orbit Determination with Sparse Data,” *Advances in Space Research*, Vol. 69, No. 12, 2022, pp. 4179–4197. <https://doi.org/10.1016/j.asr.2022.03.041>
- [49] Brox, T., Rosenhahn, B., Cremers, D., and Seidel, H.-P., “Nonparametric Density Estimation with Adaptive, Anisotropic Kernels for Human Motion Tracking,” *Human Motion—Understanding, Modeling, Capture and Animation*, edited by A. Elgammal, B. Rosenhahn, and R. Klette, Springer-Verlag, Berlin, 2007, pp. 152–165. https://doi.org/10.1007/978-3-540-75703-0_11
- [50] Scott, D. W., *Multivariate Density Estimation: Theory, Practice, and Visualization*, Wiley Series in Probability and Statistics, Wiley, New York, 2015, pp. 137–216, Chap. 6. <https://doi.org/10.1002/9781118575574.ch6>
- [51] Popov, A. A., and Zanetti, R., “Ensemble Gaussian Mixture Filtering with Particle-Localized Covariances,” *2023 26th International Conference on Information Fusion (FUSION)*, Inst. of Electrical and Electronics Engineers, New York, 2023, pp. 1–7. <https://doi.org/10.23919/FUSION52260.2023.10224227>
- [52] Kotz, S., Kozubowski, T., and Podgórski, K., *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*, No. 183, Springer-Verlag, Berlin, 2001, pp. 231–238, Chap. 6. <https://doi.org/10.1007/978-1-4612-0173-1>
- [53] Zucchelli, E. M., and Jones, B. A., “Multi-Fidelity Hamiltonian Monte Carlo for Space Object Tracking with Sparse Data,” *AAS/AIAA Space Flight Mechanics Meeting*, American Astronautical Soc. Paper 23-377, Springfield, VA, 2023.
- [54] Folkner, W. M., Williams, J. G., Boggs, D. H., Park, R. S., and Kuchynka, P., “The Planetary and Lunar Ephemerides DE430 and DE431,” IPN Progress Rept. 42-196, Jet Propulsion Lab., California Inst. of Technology, Feb. 2009, http://ipnpr.jpl.nasa.gov/progress_report/42-196/196C.pdf [retrieved 1 July 2023].
- [55] Pavlis, N. K., Holmes, S. A., Kenyon, S. C., and Factor, J. K., “The Development and Evaluation of the Earth Gravitational Model 2008 (EGM2008),” *Journal of Geophysical Research: Solid Earth*, Vol. 117, No. B4, 2012, Paper B04406. <https://doi.org/10.1029/2011JB008916>
- [56] Petit, G., and Luzum, B., “The 2010 Reference Edition of the IERS Conventions,” *Reference Frames for Applications in Geosciences*, edited by Z. Altamimi, and X. Collilieux, Springer-Verlag, Berlin, 2013, pp. 57–61.