JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 32, No. 1, January–February 2009

Adaptive Particle Filtering for Spacecraft Attitude Estimation from Vector Observations

Avishy Carmi* and Yaakov Oshman†
Technion—Israel Institute of Technology, 32000 Haifa, Israel

DOI: 10.2514/1.35878

An extension is presented to the recently introduced genetic algorithm-embedded quaternion particle filter. Belonging to the class of Monte Carlo sequential methods, the genetic algorithm-embedded quaternion particle filter is an estimator that uses approximate numerical representation techniques for performing the otherwise exact time propagation and measurement update of potentially non-Gaussian probability density functions in the inherently nonlinear attitude estimation problem. The spacecraft attitude is represented via the quaternion of rotation, and a genetic algorithm is used to estimate the gyro biases, allowing one to estimate just the quaternion via the particle filter. An adaptive version of the genetic algorithm-embedded quaternion particle filter is presented herein that extends the applicability of this filter to problems with highly uncertain measurement noise distributions. The adaptive algorithm estimates the measurement noise distribution on the fly, along with the spacecraft attitude and gyro biases. A simulation study is used to demonstrate the performance of the adaptive algorithm using real data obtained from the Technion's TechSAT satellite, whose three-axis magnetometer's data are non-Gaussian. The simulation, which compares the performance of the filter to the nonadaptive genetic algorithm-embedded quaternion particle filter, demonstrates the viability of the new algorithm.

I. Introduction

USING a sequence of vector measurements for attitude determination has been intensively investigated over the last four decades. First proposed 40 years ago by Wahba [1], the problem is to estimate the attitude of a spacecraft based on a sequence of noisy vector observations, resolved in the body-fixed coordinate system and in a reference system. Body-fixed vector observations are typically obtained from onboard sensors, such as star trackers, sun sensors, or magnetometers. Corresponding reference observations are obtained by using an ephemeris routine (for a sun observation), or from orbit data and a magnetic field routine (for a magnetic field observation), or from a star catalog (for star observations).

Inertial reference systems typically use vector measurements in combination with strapdown gyros to estimate both the spacecraft attitude and the gyro drift rate biases. Several approaches have been proposed for the design of such systems, differing mainly in their choice of attitude representation method. The quaternion, a popular rotation specifier, was used in the framework of extended Kalman filtering (EKF) in [2] (reviewing three derivations using the so-called multiplicative approach) and in [3] (using the additive approach). In [4], vector observations were used to estimate both the quaternion and the angular velocity of the spacecraft, in a gyroless attitude determination and control setting. The main advantage of using the quaternion representation is that it is not singular for any rotation. Moreover, its kinematic equation is linear and the computation of the associated attitude matrix involves only algebraic expressions. However, the quaternion representation is not minimal, because it is four-dimensional. This leads to a normalization constraint that has to be addressed in filtering algorithms. Thus, Lefferts et al. [2] assume that the 4×4 quaternion estimation error covariance matrix must be singular and propose reduced-order algorithms to maintain the assumed singularity. On the other hand, Bar-Itzhack and Oshman [3] assume no such singularity, but incorporate a normalization stage within their EKF algorithm, which renders the resulting estimator strictly nonoptimal and increases its workload. In a recent paper [5], an unscented Kalman filter (UKF) has been proposed for the estimation of the rotation quaternion. The UKF does possess a reported advantage over the EKF with regards to dealing with strongly nonlinear systems, because it avoids the linearization associated with the EKF. However, because using the UKF directly with the quaternion attitude parameterization would also yield a non-unit-norm quaternion estimate (after all, the UKF is still a Kalman filter), Crassidis and Markley [5] chose to work with a generalized three-dimensional attitude representation, still using the quaternion for updates to maintain the normalization constraint. It should also be noted that, as a Kalman filter mechanization, the UKF is also sensitive to the statistical distribution of the stochastic processes driving the dynamic model: non-Gaussian distributions guarantee nonoptimality of the estimates.

In a recent work [6], the authors have presented a particle filter (PF) that sequentially and directly estimates the rotation quaternion from vector observations. Also known as sequential Monte Carlo (SMC) methods, particle filters refer to a set of algorithms implementing a recursive Bayesian model using simulation-based methods [7]. Avoiding the underlying assumptions of the Kalman filter, namely, that the state space is linear and Gaussian, these rather general and flexible methods enable solving for the posterior probability distributions of the unknown variables (on which all inference on these variables is based) within a Bayesian framework, exploiting the dramatic recent increase in computer power. Particle filters are not just smart implementations of the Kalman filter or its nonlinear variants/extensions; rather, they are entirely different algorithms that lead to entirely different solutions to the nonlinear, non-Gaussian filtering problem. Contrary to the Kalman filter extensions, the solutions obtained using PF algorithms are approximations to the optimal (in the Bayesian sense) solutions, which can be made arbitrarily close to the exact solutions by increasing the number of particles involved in the computation, thereby also increasing the computation workload.

As a member of the PF class of algorithms, the quaternion PF (QPF) enjoys the aforementioned accuracy-related properties. Furthermore, it naturally maintains the unit-norm property of the

Received 27 November 2007; accepted for publication 15 June 2008.

*Doctoral Student, Department of Aerospace Engineering; avishy@aerodyne.technion.ac.il. Member AIAA.

†Professor, Department of Aerospace Engineering. Holder of the Louis and Helen Rogow Chair in Aeronautical Engineering. Member, Technion's Asher Space Research Institute; yaakov.oshman@technion.ac.il. Fellow AIAA.

rotation quaternion, thus avoiding the incorporation of external ad hoc normalization procedures, such as those used in [2,3]. In an extended version of the filter, called a genetic algorithm-embedded quaternion particle filter (GA-QPF), a genetic algorithm is used to generate a maximum likelihood estimate of the gyro biases, thus alleviating the potential computational burden problem associated with the number of required particles [6]. In fact, running with as few as 150 quaternion particles and a 200-element population for the genetic algorithm bias estimation scheme, the algorithm has been shown in simulations to be amenable to real-time implementations. The GA-QPF algorithm's superiority over the Kalman filter variants is manifested in its faster convergence rate and enhanced robustness to initial conditions uncertainty. Furthermore, a comparison of the estimation error covariance of this filter to the theoretical Cramér-Rao lower bound shows that it is asymptotically efficient in the statistical sense, thus corroborating its optimality and unbiasedness [6].

This paper presents an extension of the GA-QPF algorithm to situations in which the measurement noise probability density function (PDF) is uncertain or even completely unknown. Such situations might occur, in practice, when magnetometers are used for deriving the vector measurements because, then, the actual noise PDF onboard the spacecraft may be different from the one predicted on the ground. This can be the result of various remnant magnetic fields induced by the spacecraft's own electrical instruments or by Earth's magnetic storms. In other cases, various faults may also change the sensor measurement noise PDF. The essential feature of the new algorithm is its ability to estimate, on the fly, the actual measurement noise PDF using a statistical analysis of the filter-generated innovations process.

The remainder of this paper is organized as follows. The next section presents the mathematical model of the quaternion estimation problem. For completeness, a concise summary of SMC methods, and, in particular, of the GA-QPF algorithm, is presented in Sec. III. Constituting the heart of this paper, Sec. IV presents an extended GA-QPF algorithm that adaptively estimates the measurement noise PDF on the fly. Results of an extensive numerical simulation study, which was carried out to assess the performance of the new algorithm and to compare it with its nonadaptive counterpart, the GA-QPF, are then presented. Concluding remarks are offered in the last section.

II. Mathematical Model

In this section, the problem of quaternion estimation from vector observations is mathematically defined.

A. Observation Model

Let $\boldsymbol{r}_k$ and $\boldsymbol{b}_k$ be a pair of corresponding vector measurements acquired at time k in the two Cartesian coordinate systems $\mathcal{R}$ and $\mathcal{B}$, respectively. Let A_k be the rotation matrix (also known as the attitude matrix or the direction cosine matrix) that rotates the axes of $\mathcal{R}$ onto the axes of $\mathcal{B}$ at time k. In general, the reference vector $\boldsymbol{r}_k$ is known exactly, whereas the body vector $\boldsymbol{b}_k$ is measured. This results in the following attitude measurement model:

$$\boldsymbol{b}_k = A_k \boldsymbol{r}_k + \delta \boldsymbol{b}_k \tag{1}$$

where $\{\delta \boldsymbol{b}_k\}_{k=1}^{\infty}$ is the measurement noise process, with known PDF, denoted as $\delta \boldsymbol{b}_k \sim p_{\delta \boldsymbol{b}_k}$.

B. Quaternion Process Model

The discrete-time quaternion stochastic process satisfies the recurrence equation

$$\mathbf{q}_{k+1} = \Phi_k^o \mathbf{q}_k \tag{2}$$

where the process $\{\mathbf{q}_k\}_{k=1}^{\infty}$ denotes the quaternion of rotation from a given reference frame $\mathcal{R}$ onto the body frame $\mathcal{B}$ at times $k = 1, 2, \ldots, \infty$, with some initial PDF $\mathbf{q}_0 \sim p_{\mathbf{q}_0}$. The quaternion process takes its values on the unit three-sphere $\mathbb{S}^3$ and is constructed from vector and scalar parts:

$$\mathbf{q}_k = \begin{bmatrix} \boldsymbol{\varrho}_k^T & \mathrm{q}_{4_k} \end{bmatrix}^T \tag{3}$$

The orthogonal transition matrix Φ_k^o is expressed using $\boldsymbol{\omega}_k^o = [\,\omega_{1k}^o \quad \omega_{2k}^o \quad \omega_{3k}^o\,]^T$, the true angular velocity of $\mathcal{B}$ with respect to $\mathcal{R}$, resolved in $\mathcal{B}$. Assuming that $\boldsymbol{\omega}_k^o$ is constant during the sampling time interval Δt yields

$$\Phi_k^o \triangleq \Phi(\boldsymbol{\omega}_k^o) = \exp\left(\frac{1}{2}\begin{bmatrix} -[\boldsymbol{\omega}_k^o\times] & \boldsymbol{\omega}_k^o \\ -\boldsymbol{\omega}_k^{o^T} & 0 \end{bmatrix}\Delta t\right) \tag{4}$$

where $[\boldsymbol{\omega}_k^o\times]$ denotes the cross-product matrix associated with the vector $\boldsymbol{\omega}_k^o$.

In practice, the true angular velocity vector $\boldsymbol{\omega}_k^o$ is not known; rather, it is measured or estimated. Let $\{\boldsymbol{\omega}_k\}_{k=1}^{\infty}$ be the measured angular velocity stochastic process. Using $\boldsymbol{\omega}_k$ instead of $\boldsymbol{\omega}_k^o$ in Eq. (4) yields the following quaternion process equation:

$$\mathbf{q}_{k+1} = \Phi(\boldsymbol{\omega}_k)\mathbf{q}_k \tag{5}$$

where the process noise is incorporated through the transition matrix.

The relation between the process and observations is established by expressing the attitude matrix as a quadratic function of $\mathbf{q}$, that is,

$$A = A(\mathbf{q}) = [(\mathrm{q}_4)^2 - \boldsymbol{\varrho}^T\boldsymbol{\varrho}]I_{3\times 3} + 2\boldsymbol{\varrho}\boldsymbol{\varrho}^T - 2\mathrm{q}_4[\boldsymbol{\varrho}\times] \tag{6}$$

C. Rate Sensor Measurement Model

When the angular rate is measured, the characterization of the driving process noise depends upon the rate sensor. The most common angular rate sensor onboard spacecraft is the gyro triad. For this sensor, a widely used model is given by [2]

$$\boldsymbol{\omega}_k = \boldsymbol{\omega}_k^o + \boldsymbol{\eta}_k + \boldsymbol{\epsilon}_k \tag{7}$$

where $\boldsymbol{\omega}_k$ denotes the measured angular velocity vector and $\boldsymbol{\epsilon}_k \sim p_{\boldsymbol{\epsilon}_k}$ and $\boldsymbol{\eta}_k \sim p_{\boldsymbol{\eta}_k}$ are the gyros' measurement white noise and bias vectors with their given PDFs, respectively. Commonly, the bias vector is modeled as a random-walk process, that is,

$$\boldsymbol{\eta}_{k+1} = \boldsymbol{\eta}_k + \boldsymbol{\zeta}_k \tag{8}$$

where $\{\boldsymbol{\zeta}_k\}_{k=1}^{\infty}$ is a stationary zero-mean, white noise process with covariance $Q_{\boldsymbol{\zeta}}$. Typically $Q_{\boldsymbol{\zeta}}$ is very small (e.g., with entries on the order of $10^{-7}\ \mu\mathrm{rad}^2/\mathrm{s}^2$).

III. Attitude Estimation via Sequential Monte Carlo Methods

This section briefly overviews some of the key ideas underlying the particle filtering method. A concise summary of the QPF algorithm of [6] is presented thereafter.

A. Sequential Monte Carlo Methods

The optimal solution of the nonlinear estimation problem involves an accurate propagation of the optimal PDF, namely, the conditional PDF of the state given the observation history. Because of the complex nature of nonlinear estimation problems, many estimation algorithms rely on various assumptions to ensure mathematical tractability. It is well known that the Kalman filter is the optimal estimator for linear Gaussian state-space models, but its performance is limited when the aforementioned assumptions do not hold. The optimal PDF admits a Bayesian recursion, which means that it is propagated in accordance with some prior distribution of the state and a likelihood function that relates the states to the incoming observations. In the case of linear Gaussian models, where the PDF can be characterized by its first two moments, the Bayesian approach yields the Kalman filter. In general, for nonlinear, non-Gaussian models, there is no explicit, closed-form solution. Several approximate methods have been proposed. These include the EKF, the Gaussian sum filter, and numerical integration over a state-space grid. Particle filters, or SMC methods [8], refer to a set of algorithms

implementing a recursive Bayesian model by simulation-based methods. This involves representing the required posterior PDF by a set of random samples with associated weights, and deriving the estimates based on these samples. Unlike the other methods, SMC methods are very flexible, easy to implement, and applicable in very general settings.

1. Bayesian Approach to Filtering

Let the unobserved signal (i.e., the process) $\{x_k, k \in \mathbb{N}\}$ be an $\mathbb{R}^n$-valued Markov process with a given initial probability density function p_{x_0} that evolves according to a transition kernel $p_{x_k|x_{k-1}}$. The observation process $\{y_k, k \in \mathbb{N}\}$ is an $\mathbb{R}^p$-valued stochastic process. Given x_k, the observation process is a conditionally independent sequence, possessing the conditional PDF $p_{y_k|x_k}$. Let $\mathcal{X}^k \triangleq \{x_0, \ldots, x_k\}$ and $\mathcal{Y}^k \triangleq \{y_1, \ldots, y_k\}$ be the process and observation time histories up to time k, respectively, and let $X^k \triangleq \{X_0, \ldots, X_k\}$ and $Y^k \triangleq \{Y_1, \ldots, Y_k\}$ be the realizations of $\mathcal{X}^k$ and $\mathcal{Y}^k$, respectively.

In filtering problems, one is commonly interested in estimating the marginal PDF $p_{x_k|\mathcal{Y}^k}$ (filtering density) sequentially in time. Adopting the Bayesian approach to filtering, this density is obtained using a two-step recursion as

$$p_{x_k|\mathcal{Y}^{k-1}}(X_k \mid Y^{k-1}) = \int_{-\infty}^{+\infty} p_{x_k|x_{k-1}}(X_k \mid X_{k-1}) p_{x_{k-1}|\mathcal{Y}^{k-1}}(X_{k-1} \mid Y^{k-1})\, \mathrm{d}X_{k-1} \quad (9a)$$

$$p_{x_k|\mathcal{Y}^k}(X_k \mid Y^k) = \frac{p_{y_k|x_k}(Y_k \mid X_k)}{\int_{-\infty}^{+\infty} p_{y_k|x_k}(Y_k \mid X_k) p_{x_k|\mathcal{Y}^{k-1}}(X_k \mid Y^{k-1})\, \mathrm{d}X_k} p_{x_k|\mathcal{Y}^{k-1}}(X_k \mid Y^{k-1}) \quad (9b)$$

In most cases, one cannot obtain the normalizing density $p_{y_k|\mathcal{Y}^{k-1}}$ and the marginals of the posterior density $p_{\mathcal{X}^k|\mathcal{Y}^k}$. Thus, these expressions can rarely be used in a straightforward implementation. Instead, approximations, based on alternative methods, should be used.

2. Particle Approximation

The PF mechanization approximates Eqs. (9) using a finite number of samples. To understand the rationale behind this method, assume that N independent random samples (called "particles"), denoted by $\{X^k(i)\}_{i=1}^N$, are sampled from the posterior distribution. Then, it follows directly from the strong law of large numbers (SLLN) that, for any function f that is integrable with respect to $p_{\mathcal{X}^k|\mathcal{Y}^k}$ [9],

$$\frac{1}{N}\sum_{i=1}^{N} f(X^k(i)) \to E[f(\mathcal{X}^k) \mid Y^k] \quad (10)$$

where (here and in the sequel) the symbol $\to$ stands for almost sure convergence in N. Equation (10) means that the continuous posterior PDF $p_{\mathcal{X}^k|\mathcal{Y}^k}$ can be effectively approximated by its particles, and that the level of accuracy of this approximation is determined by the number of particles used.

Unfortunately, sampling directly from the posterior distribution is typically infeasible. For this reason, the concept of importance sampling is an integral part of any practical PF [8]. When using importance sampling, the samples are drawn from a so-called importance distribution. The importance distribution can be chosen arbitrarily; the only constraint it should satisfy is that its support must include the support of the posterior distribution. Nevertheless, because the choice of importance distribution greatly affects the behavior of the PF, it constitutes a major consideration during PF design.

Thus, choosing the importance density as $\pi_{\mathcal{X}^k|\mathcal{Y}^k}$, an approximation of the expectation in Eq. (10) is obtained as

$$\frac{1}{N}\sum_{i=1}^{N} w_k(i) f(X^k(i))$$

where

$$w_k(i) \triangleq \frac{p_{\mathcal{X}^k|\mathcal{Y}^k}(X^k(i) \mid Y^k)}{\pi_{\mathcal{X}^k|\mathcal{Y}^k}(X^k(i) \mid Y^k)} \quad (11)$$

is the importance weight of the ith sample. Let

$$\tilde{w}_k(i) \triangleq \frac{w_k(i)}{\sum_{j=1}^{N} w_k(j)} \quad (12)$$

be the normalized importance weight of the ith particle, then it can be easily verified that

$$\sum_{i=1}^{N} \tilde{w}_k(i) f(X^k(i)) \to E[f(\mathcal{X}^k) \mid Y^k] \quad (13)$$

In practice, the PF algorithm exploits the recursive structure of Eqs. (9) to compute the normalized importance weights sequentially in time.

3. Particle Degeneracy and Resampling

Practical implementation of the sequential importance sampling method inevitably results in zero weights for all but, usually, one particle, after just a few iterations. This phenomenon is known as particle degeneracy in the PF literature [8]. Particle degeneracy occurs due to the use of a finite number of particles, which consequently allows only a partial representation of the sample space. A solution to this problem was introduced a decade ago as an ad hoc procedure known as resampling.

Resampling consists of discarding state trajectories whose contributions to the final estimate are small, and multiplying trajectories whose contributions are expected to be significant. This means regeneration of particles with large importance weights and eliminating those with small importance weights. The resampling procedure decreases the particle degeneracy algorithmically, but introduces some practical problems. During the resampling procedure, more likely particles are multiplied, so that the particle cloud is concentrated in regions of interest of the state space. This produces a new particle system in which several particles share the same location. Moreover, if the dynamic noise is small, the particle system ultimately concentrates in a single point in state space. This loss of diversity eventually prevents the filter from correctly representing the posterior. One way of maintaining the particles' diversity is by injecting artificial process noise into the system. This technique is known as regularization or roughening (see [7], p. 247).

B. Quaternion Particle Filter

The QPF algorithm estimates the quaternion from pairs of vector observations. Within this particle filter, each particle is a unit-norm quaternion, so that the norm constraint is inherently preserved. For a detailed presentation of the QPF algorithm, the interested reader is referred to [6].

1. Quaternion Particle Filter Initialization

Large initial attitude errors require a large number of quaternion particles, at least until the zones of high likelihood are populated. A simple initialization procedure that demands a significantly smaller number of particles is used for the QPF. The idea is based on the fact that the first vector observation defines a quaternion of rotation up to 1 degree of freedom. This degree of freedom is used to generate the initial set of particles from the first observation only, $Y_0 = [\boldsymbol{b}_0; \boldsymbol{r}_0]$. In the simulations shown in [6], the QPF has been started with 1500 quaternion particles, reducing their number to 200 after the first two measurement updates.

2. Measurement Update

Denote by $Y^k = \{[\boldsymbol{b}_1; \boldsymbol{r}_1], \ldots, [\boldsymbol{b}_k; \boldsymbol{r}_k]\}$ a set of measurements constructed from pairs of vector observations up to time k. Given a realization $\boldsymbol{q}_k$ of the quaternion $\mathbf{q}_k$ at time k, the measurement $Y_k = [\boldsymbol{b}_k; \boldsymbol{r}_k]$ is statistically independent of past observations. The likelihood of the measurement Y_k associated with a given quaternion is

$$p_{y_k|\mathbf{q}_k}(Y_k \mid \boldsymbol{q}_k) = p_{\delta\mathbf{b}_k}(\boldsymbol{b}_k - A(\boldsymbol{q}_k)\boldsymbol{r}_k) \tag{14}$$

Now let $\{\boldsymbol{q}_{k-1}(i)\}_{i=1}^N$ and $\{\tilde{w}_{k-1}(i)\}_{i=1}^N$ denote N independent unit quaternion samples from the filtering distribution at time $k-1$ and their associated weights, respectively. Setting the importance distribution to be the prior PDF yields the importance weights as

$$w_k(i) = p_{y_k|\mathbf{q}_k}(Y_k \mid \boldsymbol{q}_k(i))\tilde{w}_{k-1}(i) \tag{15}$$

Equation (15) is referred to as the update stage.

3. Filtered Quaternion

At time k, N weighted unit quaternion samples are available. The optimal quaternion estimate can be computed in several ways, depending on the objective. The minimum mean square error unit-norm quaternion estimate is obtained as follows. Letting

$$B_k \triangleq \sum_{i=1}^{N} \tilde{w}_k(i) A_k[\boldsymbol{q}_k(i)] \tag{16}$$

and denoting the singular value decomposition of B_k by

$$B_k = U_k \Sigma_k V_k^T \tag{17}$$

where U_k and V_k are the orthogonal singular vector matrices and Σ_k is the singular value matrix of B_k, the attitude matrix associated with the optimal quaternion is computed as

$$\hat{A}_k = U_k V_k^T \tag{18}$$

The filtered quaternion is then obtained from $\hat{A}_k$.

4. Particle Maintenance

To avoid particle degeneracy, the QPF uses a resampling procedure. The measure of particle degeneracy is the effective sample size, approximated by the following empirical estimate [8]:

$$\hat{N}_{\text{eff}} = \frac{1}{\sum_{i=1}^{N} \tilde{w}_k(i)^2} \tag{19}$$

The resampling procedure is used whenever $\hat{N}_{\text{eff}}$ becomes less than a predetermined threshold N_{th}. The new set of samples is generated by resampling each particle $\boldsymbol{q}_k(i)$ with probability $\tilde{w}_k(i)$. This consists of multiplying each sample according to its associated normalized weight.

5. Particle Evolution

Passing the unit quaternion samples at time $k-1$ through the process equation results in a new set of samples. This is almost equivalent to applying Eq. (9a) to the samples, that is,

$$p_{\mathbf{q}_k|\mathcal{Y}^{k-1}}(\boldsymbol{q}_k \mid Y^{k-1}) = \int_{-\infty}^{+\infty} p_{\mathbf{q}_k|\mathbf{q}_{k-1}}(\boldsymbol{q}_k \mid \boldsymbol{q}_{k-1}) p_{\mathbf{q}_{k-1}|\mathcal{Y}^{k-1}}(\boldsymbol{q}_{k-1} \mid Y^{k-1})\, \mathrm{d}\boldsymbol{q}_{k-1} \tag{20}$$

The slight difference is due to the process noise distribution, which forms the transition kernel $p_{\mathbf{q}_k|\mathbf{q}_{k-1}}$. In the case of a relatively low-intensity process noise, such as the noise characterizing the quaternion evolution model, the new quaternion samples thus obtained represent $p_{\mathbf{q}_k|\mathcal{Y}^{k-1}}$ quite adequately. In other cases, the injection of an additional, artificial noise might be required.

6. Gyro Bias Estimation

The QPF is interlaced with an external maximum likelihood (ML) estimator for the estimation of the gyro bias, thus alleviating the potential computation load resulting from high-dimensional particle filtering. The key idea here is that the particle filter is used for the representation of $p_{\mathbf{q}_k|\boldsymbol{\eta}_k,\mathcal{Y}^k}$ instead of $p_{\mathbf{q}_k,\boldsymbol{\eta}_k|\mathcal{Y}^k}$, thus keeping the dimension of the state low, whereas $\boldsymbol{\eta}_k$ is estimated via an external ML estimator assuming the knowledge of $\hat{\boldsymbol{q}}_k$.

The GA-QPF, which is an extension of the QPF for biased rate gyros, uses a genetic-algorithm-based ML estimator. The proposed genetic algorithm (GA) maximizes the likelihood function of the bias conditioned upon the observations history $\mathcal{L}(\boldsymbol{\eta}_k \mid Y^k)$ sequentially in time, and includes some modifications to cope with the time-varying nature of the biases. The bias parameter population at time k is denoted as $\{\boldsymbol{\Upsilon}_k(i)\}_{i=1}^{N_\eta}$. The bias parameter fitness function is related to the likelihood $\mathcal{L}(\boldsymbol{\Upsilon}_k(i) \mid Y^k)$ and is denoted as $\tilde{\varphi}_k(\boldsymbol{\Upsilon}_k(i))$. A binary coding scheme is used for the representation of the bias vector terms. A single GA-QPF cycle is illustrated in Fig. 1.

IV. Measurement Noise Density Estimation

Forming the core of this paper, this section addresses the problem of attitude estimation under a severe uncertainty in the measurement noise distribution. Thus, it is assumed that the measurement noise PDF $p_{\delta\mathbf{b}_k}$ is either significantly inaccurate or completely unknown. This situation might arise in practice when the actual operating conditions of the sensor differ significantly from the predicted ones. In the case of a three-axis magnetometer, for example, magnetic storms and the spacecraft's own electrical instrumentation may critically affect this sensor's readings. In other cases, faults may also change the measurement noise PDF. The simulation section, in the ensuing, addresses a realistic example where a satellite momentum wheel has induced a magnetic dipole which has resulted in a double-peaked magnetometer noise PDF. The deviation of the assumed measurement noise PDF from the actual one adversely affects the optimality of the filtering scheme. In some extreme cases, this deviation might even result in filter divergence.

To cope with the problem, the GA-QPF algorithm is equipped with a self-learning mechanism that estimates the measurement noise PDF on the fly. The noise distribution estimation scheme is based on an analysis of the filter-generated innovations process.

A. Innovations Process

Because the actual (true) measurement noise PDF $p_{\delta\mathbf{b}_k}$ is unknown, the QPF particles are weighted using some initially assumed or otherwise estimated PDF. After acquiring a measurement, an estimate of the QPF innovations process is computed according to

$$\mathbf{e}_k = \mathbf{b}_k - E^*[A(\mathbf{q}_k)\boldsymbol{r}_k \mid \mathcal{Y}^{k-1}] \tag{21}$$

where the mathematical expectation $E^*[\cdot]$ is computed using the estimated measurement noise PDF (and not the actual one). An approximation of Eq. (21) using the filter samples is

$$\hat{\mathbf{e}}_k = \boldsymbol{b}_k - \left[\sum_{i=1}^{N} \tilde{w}_{k-1}(i) A(\boldsymbol{q}_{k|k-1}(i))\right]\boldsymbol{r}_k \tag{22}$$

where $\boldsymbol{q}_{k|k-1}(i)$ denotes the ith quaternion particle, computed based on $k-1$ measurements and propagated to time k. The relation between the innovations and the measurement noise is obtained by rewriting Eq. (21) as

$$\begin{aligned}\mathbf{e}_k &= \mathbf{b}_k - E^*[A(\mathbf{q}_k)\boldsymbol{r}_k \mid \mathcal{Y}^{k-1}] \\ &= A(\mathbf{q}_k)\boldsymbol{r}_k + \delta\mathbf{b}_k - E^*[A(\mathbf{q}_k)\boldsymbol{r}_k \mid \mathcal{Y}^{k-1}]\end{aligned} \tag{23}$$

Defining

$$\delta\mathbf{h}_k \triangleq A(\mathbf{q}_k)\boldsymbol{r}_k - E^*[A(\mathbf{q}_k)\boldsymbol{r}_k \mid \mathcal{Y}^{k-1}] \tag{24}$$

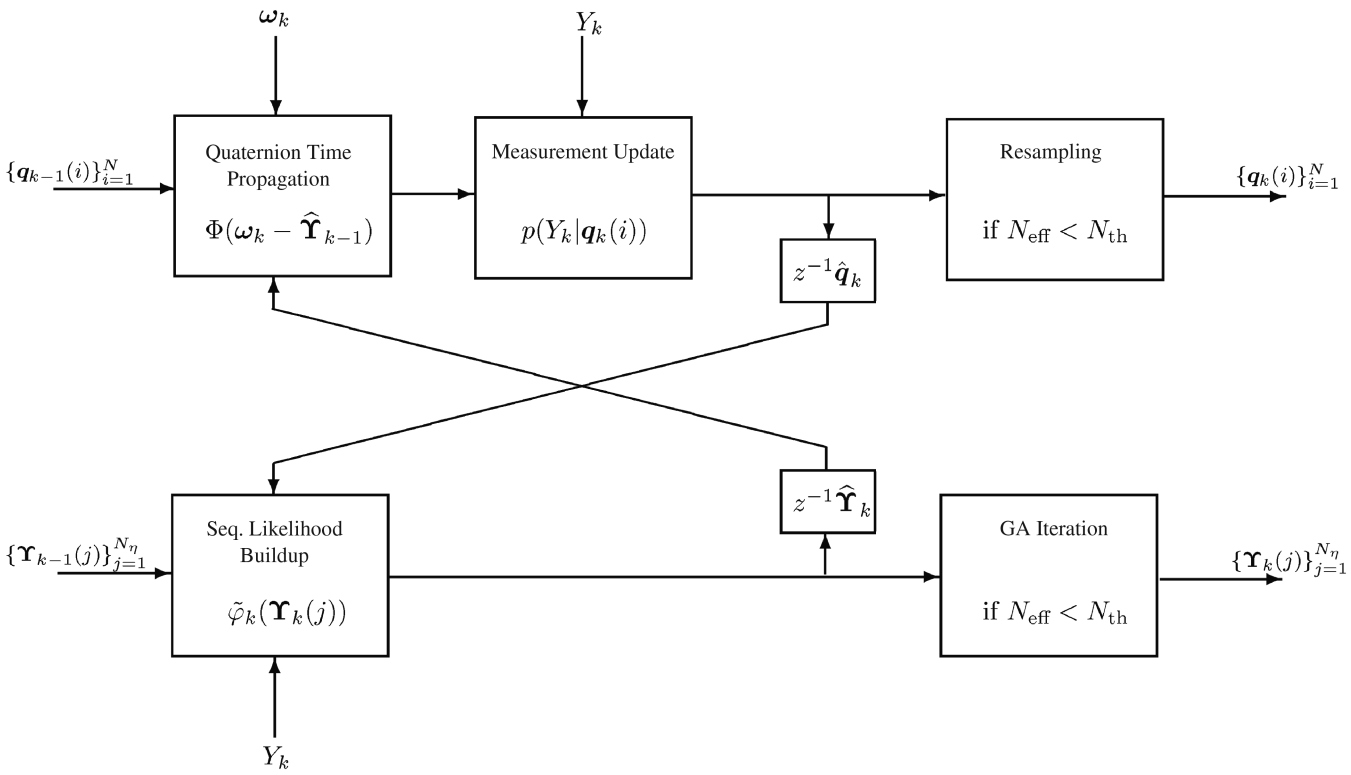


Fig. 1 GA-QPF scheme.

allows rewriting Eq. (23) as

$$\mathbf{e}_k = \delta\mathbf{b}_k + \delta\mathbf{h}_k \tag{25}$$

Since $\delta\mathbf{h}_k$ is independent of $\delta\mathbf{b}_k$, then

$$p_{\mathbf{e}_k} = p_{\delta\mathbf{b}_k} * p_{\delta\mathbf{h}_k} \tag{26}$$

where $*$ denotes the convolution operator. Writing Eq. (26) explicitly yields

$$p_{\mathbf{e}_k}(\boldsymbol{e}_k) = \int_{-\infty}^{+\infty} p_{\delta\mathbf{b}_k}(\boldsymbol{e}_k - \delta\boldsymbol{h}_k) p_{\delta\mathbf{h}_k}(\delta\boldsymbol{h}_k)\,\mathrm{d}\delta\boldsymbol{h}_k = E[p_{\delta\mathbf{b}_k}(\boldsymbol{e}_k - \delta\mathbf{h}_k)] \tag{27}$$

Assuming that $p_{\delta\mathbf{b}_k}$ is a smooth function of $\delta\boldsymbol{b}_k$ and using a first-order Taylor expansion of $p_{\delta\mathbf{b}_k}$ in Eq. (27) yields

$$\begin{aligned} p_{\mathbf{e}_k}(\boldsymbol{e}_k) &\approx E\left[p_{\delta\mathbf{b}_k}(\boldsymbol{e}_k) - \left[\frac{\partial p_{\delta\mathbf{b}_k}(\delta\boldsymbol{b}_k)}{\partial\delta\boldsymbol{b}_k}\right]_{\delta\boldsymbol{b}_k=\boldsymbol{e}_k}^{T} \delta\mathbf{h}_k \right] \\ &= p_{\delta\mathbf{b}_k}(\boldsymbol{e}_k) - \left[\frac{\partial p_{\delta\mathbf{b}_k}(\delta\boldsymbol{b}_k)}{\partial\delta\boldsymbol{b}_k}\right]_{\delta\boldsymbol{b}_k=\boldsymbol{e}_k}^{T} E[\delta\mathbf{h}_k] \end{aligned} \tag{28}$$

Hence, $p_{\mathbf{e}_k} \to p_{\delta\mathbf{b}_k}$ if $\|E[\delta\mathbf{h}_k]\|_2 \to 0$. Obviously, $\delta\mathbf{h}_k$ is related to the quaternion estimation error, so that smaller quaternion estimation errors result in better PDF approximations.

B. Histogram Estimator

In this work it is proposed to estimate the measurement noise PDF using a histogram estimator (see [10] for a brief introduction on density estimation). More precisely, the histogram estimator is applied to the estimation of the PDF of the innovations process $p_{\mathbf{e}}$ that approximates the stationary noise PDF $p_{\delta\mathbf{b}}$, as has been previously shown.

Let $\mathcal{S} \triangleq \text{supp}\{p_{\mathbf{e}}\} \in \mathbb{R}^3$ be the support of $p_{\mathbf{e}}$, and let $\{\mathcal{A}_i\}_{i=1}^{N_p}$, $\mathcal{A}_i \in \mathbb{R}^3$, be a partition of $\mathcal{S}$, with nonzero Lebesgue measures $\{\mu(\mathcal{A}_i)\}_{i=1}^{N_p}$. For practical implementation reasons, it is assumed that the support $\mathcal{S}$ is a bounded set and that some bound of it is available.

The innovations histogram estimator at time k is defined as

$$\hat{p}_{\mathbf{e}_k}(\boldsymbol{e}) \triangleq \frac{1}{k\mu(\mathcal{A}_i)} \sum_{j=1}^{k} \mathbf{1}_{\{\mathbf{e}_j \in \mathcal{A}_i\}}(\boldsymbol{e}_j), \qquad \boldsymbol{e} \in \mathcal{A}_i \tag{29}$$

where $\mathbf{1}_{\{\mathbf{e}_j \in \mathcal{A}_i\}}$ denotes the indicator random variable for the event $\{\mathbf{e}_j \in \mathcal{A}_i\}$.

The histogram estimator draws its power from the SLLN. Under the assumption that the innovations process is an independent and identically distributed sequence and the partition of $\mathcal{S}$ is sufficiently fine, the SLLN implies

$$\begin{aligned} &\lim_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} \mathbf{1}_{\{\mathbf{e}_j \in \mathcal{A}_i\}}(\boldsymbol{e}_j) = E[\mathbf{1}_{\{\mathbf{e}_j \in \mathcal{A}_i\}}] = p_{\mathbf{e}}(\boldsymbol{e})\mu(\mathcal{A}_i) \\ &\qquad \text{with probability 1,} \qquad \boldsymbol{e} \in \mathcal{A}_i \end{aligned} \tag{30}$$

yielding

$$\lim_{k\to\infty} \hat{p}_{\mathbf{e}_k}(\boldsymbol{e}) = p_{\mathbf{e}}(\boldsymbol{e}) \qquad \text{with probability 1,} \qquad \boldsymbol{e} \in \mathcal{A}_i \tag{31}$$

The innovations samples of the QPF at time k are approximated using Eq. (22) after acquiring a new measurement, and the histogram estimator is computed sequentially using $\{\hat{\boldsymbol{e}}_j\}_{j=1}^{k}$. Assuming a uniform partition measure, that is,

$$\mu(\mathcal{A}_i) = \mu, \qquad i = 1, 2, \ldots, N_p \tag{32}$$

yields

$$\hat{p}_{\mathbf{e}_k}(\hat{\boldsymbol{e}}_k) = \frac{k-1}{k} \hat{p}_{\mathbf{e}_{k-1}}(\hat{\boldsymbol{e}}_k) + \frac{1}{k\mu} \mathbf{1}_{\{\hat{\boldsymbol{e}}_k \in \mathcal{A}_i\}}(\hat{\boldsymbol{e}}_k), \qquad \hat{\boldsymbol{e}}_k \in \mathcal{A}_i \tag{33}$$

The GA-QPF is made adaptive by incorporating the estimator $\hat{p}_{\mathbf{e}_k}$ into the basic algorithm described in the previous section. Thus, the likelihood $p_{y_k|\mathbf{q}_k}$ in Eq. (15) is computed as

$$p_{y_k|\mathbf{q}_k}(Y_k \mid \boldsymbol{q}_k(i)) \propto \hat{p}_{\mathbf{e}_k}(\boldsymbol{b}_k - A(\boldsymbol{q}_k(i))\boldsymbol{r}_k) \tag{34}$$

C. Curse of Dimensionality and Probability Density Estimation

The attainable performance of the histogram estimator is highly affected by its resolution (i.e., the number of partition sets). In particular, increasing the number of partition sets may improve the PDF representation. However, as the number of partition sets used grows, the PDF reconstruction procedure demands the acquisition of a larger number of measurements. This also implies a slower convergence of the histogram estimator. This phenomenon constitutes the main problem underlying high-dimensional PDF reconstruction.

One approach for alleviating this problem is by approximating the PDF marginals instead of the joint PDF itself. Thus, each marginal can be approximated using a one-dimensional histogram estimator. Nevertheless, using this approach, it is necessary to model the probabilistic relations between the various marginals for obtaining an adequate joint PDF representation.

V. Simulation Study

A simulation study has been carried out to evaluate the performance of the adaptive GA-QPF algorithm. As part of this evaluation, both the adaptive and the nonadaptive GA-QPF schemes are compared. The comparison is based on a test case involving real data obtained from the Technion's TechSAT satellite.

A. TechSAT Satellite Model

The performance of the adaptive GA-QPF is tested using a realistic three-axis magnetometer (TAM) noise model, taken from the Technion's TechSAT microsatellite. The TechSAT orbit is inclined at 98 deg at an altitude of 820 km and its period is 101 min. In this scenario, the satellite performs an Earth-pointing mission. Analyzing 75 h of TAM data (acquired at a rate of once per 10 s) allows estimating the TAM measurement noise joint PDF. The marginals of this PDF are shown in Fig. 2. Clearly, Fig. 2 implies that the TechSAT TAM's joint PDF is nonzero mean and non-Gaussian. The double-peaked marginal distribution is due to a parasitic magnetic dipole moment along the Y body-frame axis, which was encountered during the momentum wheel slowdown [11]. The TAM noise is contaminated by an additional bias having the magnitude of -1 μT along the Y axis. It is assumed that gyroscopic rate sensors provide the angular rate readings. The rate-integrating gyros' (RIGs') output is contaminated with a measurement noise having two components: a white, zero-mean Gaussian process with intensity 0.1 $(\mu\text{rad})^2/\text{s}$, and a drift bias modeled as an integrated Gaussian white noise with intensity 1×10^{-7} $(\mu\text{rad})^2/\text{s}^3$. The RIGs' initial bias is set to 0.1 deg /h on each axis. The Earth's magnetic field is modeled using the eighth-order international geomagnetic reference field.

In the Monte Carlo simulation study, the initial attitude quaternion is randomly generated according to a uniform distribution on the unit three sphere. The GA-QPFs (both the adaptive and nonadaptive versions) are initialized with $N = 2000$ particles and $N_\eta = 200$ bias parameters, using the initialization scheme previously described. The chromosome length is set to $l = 30$ bit, and the crossover procedure of the GA is applied to only 40% of the chromosomes. After the first two measurement updates, the filters' particle set is reduced to the $N = 200$ unit quaternion particles associated with the largest importance weights. The resampling threshold is set to $N_{\text{th}} = \frac{2}{3}N$ based on tuning runs. (Decreasing N_{th} may reduce the resampling frequency, consequently introducing less Monte Carlo variations into the estimates, however, this might also increase the algorithms' sensitivity to heavy-tailed measurement noise PDFs.) The sampling rate of the TAM is one per 10 s, whereas the gyros are sampled at 1 Hz. The attitude estimation error (in degrees) is computed as

$$\delta\alpha = 2\arccos(\delta q_4) \tag{35}$$

where δq_4 is the scalar component of the error quaternion $\delta\boldsymbol{q}$.

The nonadaptive GA-QPF is fed with an approximation of the correct measurement noise distribution, obtained by fitting a Gaussian mixture to the real data. The corresponding likelihood function $p_{y_k|\mathbf{q}_k}$ used by this filter is given by

$$\begin{aligned} p_{y_k|\mathbf{q}_k}(Y_k|\boldsymbol{q}_k) = \kappa\{&\exp[-\tfrac{1}{2}(\boldsymbol{b}_k - A(\boldsymbol{q}_k)\boldsymbol{r}_k - \bar{\boldsymbol{n}} + \bar{\boldsymbol{m}})^T R^{-1} \\ &\times (\boldsymbol{b}_k - A(\boldsymbol{q}_k)\boldsymbol{r}_k - \bar{\boldsymbol{n}} + \bar{\boldsymbol{m}})] + \exp[-\tfrac{1}{2}(\boldsymbol{b}_k - A(\boldsymbol{q}_k)\boldsymbol{r}_k \\ &+ \bar{\boldsymbol{n}} + \bar{\boldsymbol{m}})^T R^{-1}(\boldsymbol{b}_k - A(\boldsymbol{q}_k)\boldsymbol{r}_k + \bar{\boldsymbol{n}} + \bar{\boldsymbol{m}})]\} \end{aligned} \tag{36}$$

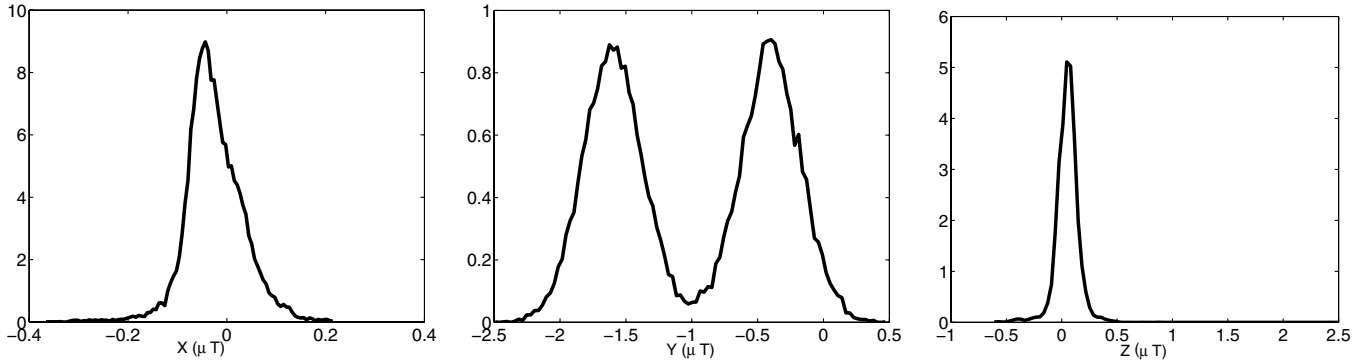
where κ is a normalization constant, $\bar{\boldsymbol{n}} = [0 \quad -0.6 \quad 0]^T$ μT, $\bar{\boldsymbol{m}} = [0 \quad -1 \quad 0]^T$ μT, and $R = \text{diag}\{\sigma_x^2, \sigma_y^2/5.7, \sigma_z^2\}$, where $\sigma_x = 1.06 \times 10^{-1}$ μT, $\sigma_y = 6.42 \times 10^{-1}$ μT, and $\sigma_z = 5.86 \times 10^{-2}$ μT denote the TAM noise data sample standard deviations along the X, Y, and Z axes, respectively.

1. Adaptive Filtering Using Marginal Estimators

The adaptive GA-QPF is not aware of the actual measurement noise distribution. To construct the measurement noise PDF on the fly in a computationally efficient manner, the following approximate procedure is used. Instead of directly estimating the entire three-dimensional joint PDF of the measurement noise, its marginals are first separately estimated via the use of three histogram estimators. This yields the three marginal estimates $\hat{p}_{\mathbf{e}_k^{(i)}}$, $i = 1, 2, 3$, each corresponding to a different TAM channel. Thus, the approximated innovations component $\hat{\boldsymbol{e}}_k^{(i)}$, $i = 1, 2, 3$, which is the ith component of $\hat{\boldsymbol{e}}_k$, is used to construct the ith marginal. Letting $\boldsymbol{g}_k \triangleq [\boldsymbol{b}_k - A(\boldsymbol{q}_k)\boldsymbol{r}_k]$, the likelihood $p_{y_k|\mathbf{q}_k}$ is then approximated as

$$\begin{aligned} p_{y_k|\mathbf{q}_k}(Y_k \mid \boldsymbol{q}_k) \propto &\left\| \left[\hat{p}_{\mathbf{e}_k^{(1)}}\left(\boldsymbol{g}_k^{(1)}\right) \quad \hat{p}_{\mathbf{e}_k^{(2)}}\left(\boldsymbol{g}_k^{(2)}\right) \quad \hat{p}_{\mathbf{e}_k^{(3)}}\left(\boldsymbol{g}_k^{(3)}\right) \right] \right\|_2 \\ &\times \mathbf{1}_{\{\|\boldsymbol{g}_k\|_2 < M\}}(\boldsymbol{g}_k) \end{aligned} \tag{37}$$

where $\boldsymbol{g}_k^{(i)}$, $i = 1, 2, 3$, denotes the ith component of $\boldsymbol{g}_k$. The bound of the PDF support M is arbitrarily set to $6 \times \|[\sigma_x \quad \sigma_y \quad \sigma_z]\|_2$. The initial PDF is set to be zero-mean Gaussian with standard deviation of 100 nT, and each histogram estimator is initialized with 200 fictitious innovation samples from this initial distribution. The number of



a) X axis b) Y axis c) Z axis

Fig. 2 TechSAT TAM noise PDF marginals.

equally measured intervals used for the support partition in each histogram estimator is $N_p = 100$.

To assess the performance of the histogram estimator, the Kullback-Liebler divergence measure for discrepancy between two densities is used. Denoted by J and defined as

$$J(p_1, p_2) \triangleq \int_{-\infty}^{+\infty} \log\left(\frac{p_1(X)}{p_2(X)}\right)[p_1(X) - p_2(X)]\,\mathrm{d}X \tag{38}$$

it is a measure of the difficulty in discriminating between the two densities p_1 and p_2. The measure $J(p_1, p_2) \geq 0$ has all the properties of a metric except for the triangle inequality, and it equals zero only if $p_1 = p_2$. In the present case, where three different histogram estimators are used, a norm Kullback-Liebler divergence is defined as

$$J_k = \|[J(\hat{p}_{\mathbf{e}_k^{(1)}}, p_{\delta\mathbf{b}_k^{(1)}}) \quad J(\hat{p}_{\mathbf{e}_k^{(2)}}, p_{\delta\mathbf{b}_k^{(2)}}) \quad J(\hat{p}_{\mathbf{e}_k^{(3)}}, p_{\delta\mathbf{b}_k^{(3)}})]\|_2 \tag{39}$$

where $p_{\delta\mathbf{b}_k^{(i)}}$, $i = 1, 2, 3$, denotes the ith marginal of the true TAM noise PDF.

2. *Joint PDF Estimation*

The measurement noise joint PDF can be fully approximated using a three-dimensional histogram estimator. Contrary to the one-dimensional case where each partition set is a line interval, in this case, each partition set is a cube. Unfortunately, this type of histogram estimator is highly subjected to the curse of dimensionality. This implies that any increase in the resolution of the histogram (i.e., number of partition sets) results in an exponential growth of the number of innovations samples required for reconstructing the PDF. A reasonable number of 10^3 partition sets is used in this work (i.e., 10 partition sets for each of the three axes).

B. Simulation Results

Figure 3 demonstrates the performance of the adaptive estimation scheme by showing the mean quaternion estimation error of 1000 Monte Carlo runs of both nonadaptive and adaptive versions of the GA-QPF, where the latter uses three marginal histogram estimators. Lacking any knowledge of the true TAM PDF, the adaptive GA-QPF reaches steady-state values of less than 0.1 deg after approximately 3 h. Providing a performance bound for comparison, the steady-state value of the nonadaptive GA-QPF, equipped with the true TAM PDF, is about 0.03 deg.

The advantage of using the adaptive scheme is highlighted by comparing it with the nonadaptive GA-QPF scheme in a case where both algorithms lack complete knowledge regarding the TAM PDF. Figure 4 shows the performance of both schemes, where the nonadaptive GA-QPF is not aware of the true measurement noise distribution but, rather, assumes it to be a Gaussian PDF parameterized by the correct statistical moments. The curve corresponding to the adaptive GA-QPF mean attitude estimation error is identical to that shown in Fig. 3. As can be clearly seen from Fig. 4, the PDF reconstruction procedure enables the adaptive GA-QPF to reach a steady-state estimation error of about 0.1 deg. On the other hand, lacking adaptivity, and assuming an incorrect measurement noise PDF (albeit with the correct statistical moments), the basic GA-QPF algorithm attains a steady-state attitude error on the order of 0.3 deg.

The Monte Carlo mean norm Kullback-Liebler divergence measure is shown in Fig. 5. Starting from values of about 2.5, the divergence measure decreases to about 0.2, which indicates that the estimated marginals are converging to the true TAM marginals.

In another run, the initial distribution of the adaptive GA-QPF is set to a zero-mean Gaussian distribution with standard deviation of 650 nT. The quaternion estimation errors for this case are quite similar to those of the first case (in which the adaptive GA-QPF was

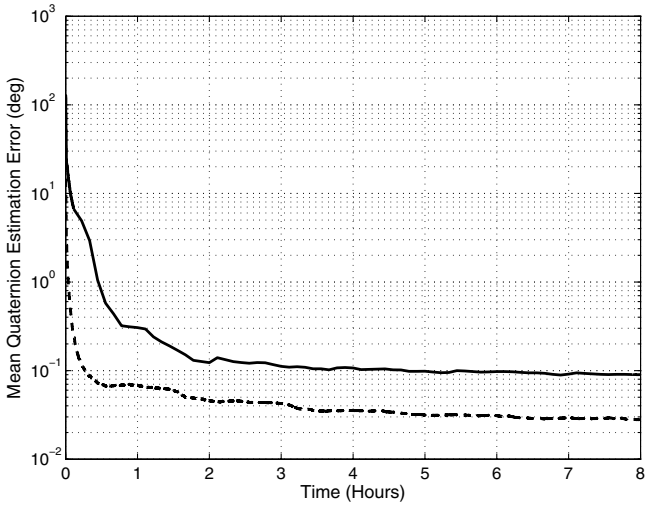


Fig. 3 Mean attitude estimation errors of the adaptive and nonadaptive filters. The nonadaptive GA-QPF (dashed line) is equipped with the correct measurement noise distribution. The adaptive GA-QPF (solid line) is not aware of the correct PDF (and estimates it on the fly); 1000 Monte Carlo runs.

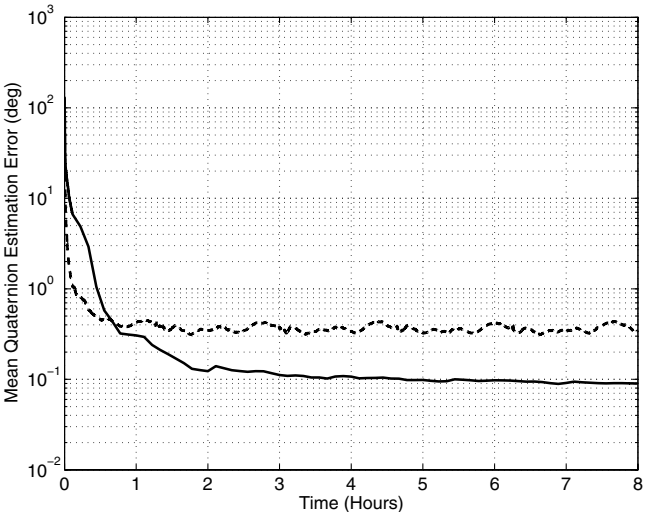


Fig. 4 Mean attitude estimation errors of the adaptive and nonadaptive filters. The adaptive GA-QPF (solid line) is not aware of the correct PDF (and estimates it on the fly). The nonadaptive GA-QPF (dashed line) assumes a Gaussian measurement noise distribution, parameterized by the statistical moments of the correct TAM PDF; 1000 Monte Carlo runs.

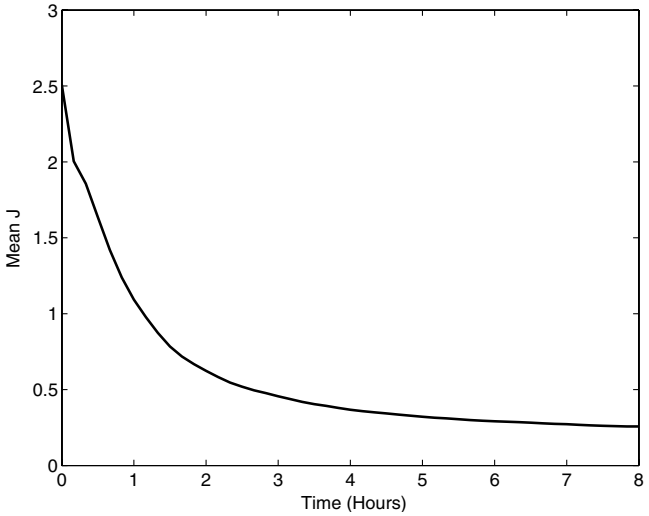


Fig. 5 Mean norm Kullback-Liebler divergence measure; 1000 Monte Carlo runs.

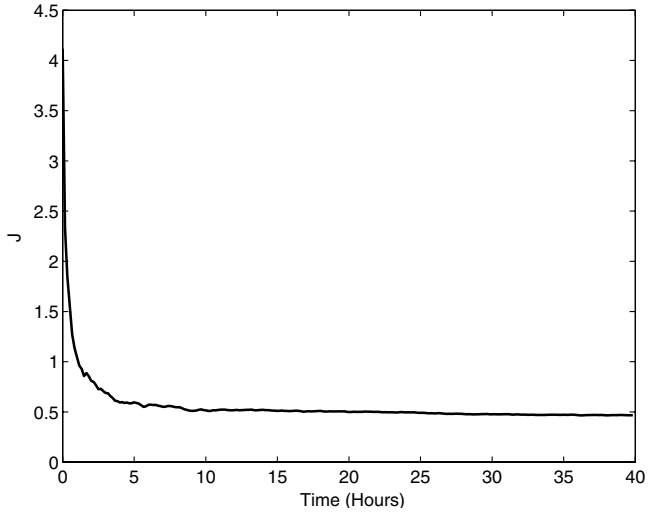
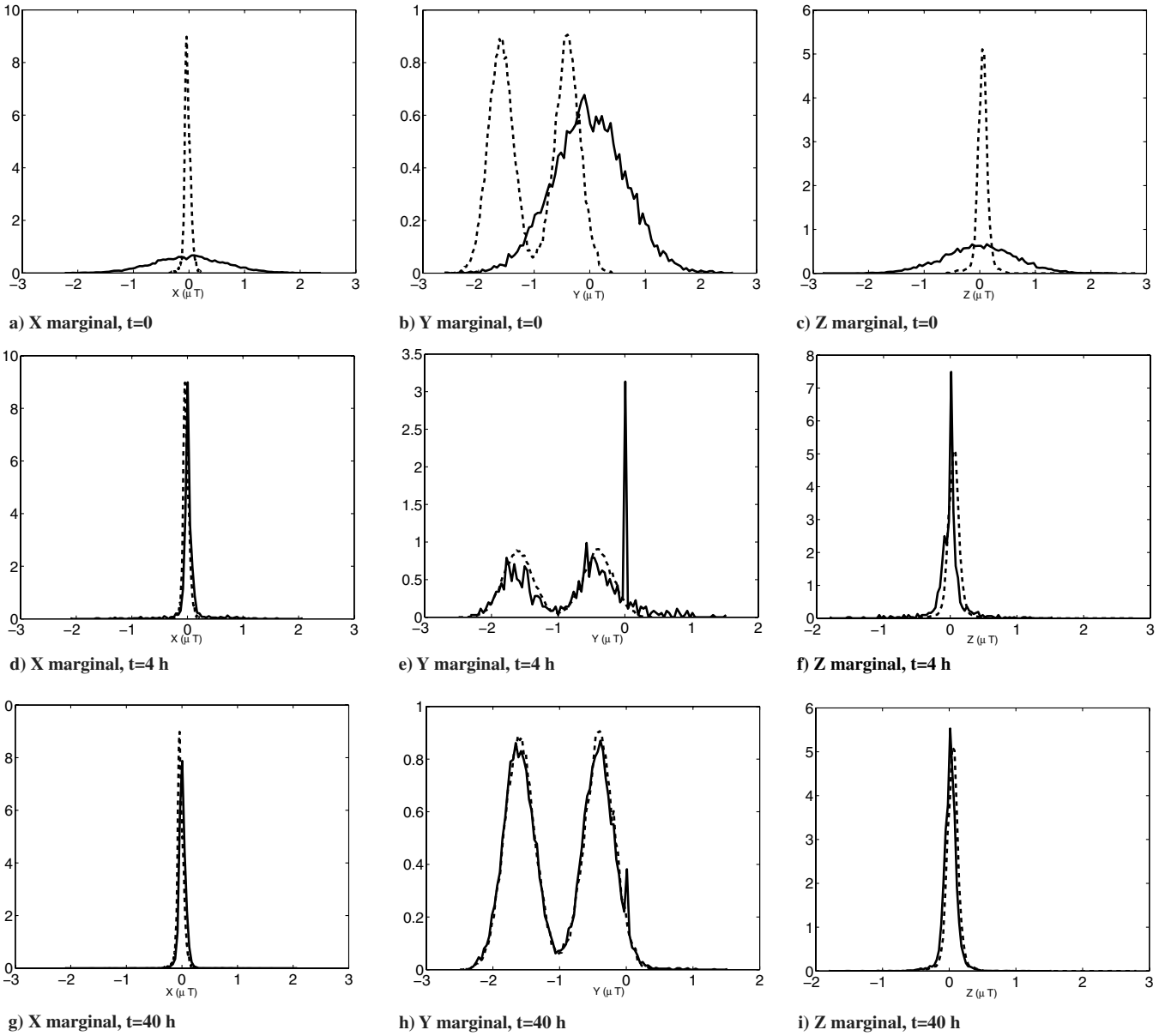


Fig. 6 Norm Kullback-Liebler divergence measure; single run.

initialized with a zero-mean Gaussian distribution with standard deviation of 100 nT). The norm Kullback-Liebler measure, for a single run of 40 h, is shown in Fig. 6. Corresponding to the larger standard deviation of the initial distribution, the initial norm divergence measure for this case is slightly larger than in the previous case. Figure 7 illustrates the TAM PDF marginals reconstruction process for this single run. The upper panel shows the true and the initially assumed marginals (corresponding to a norm divergence measure of about 3.7), whereas the bottom panel shows the true and the estimated marginals after 40 h (corresponding to a norm divergence measure of about 0.5).

Figure 8 shows the mean attitude estimation error based on 1000 Monte Carlo runs of the adaptive GA-QPF using a three-dimensional histogram estimator. As was previously mentioned, this setting is suitable for approximating the measurement noise joint PDF. The histogram estimator resolution is set to 10 partition sets for each axis, which amounts to a total of $N_p = 10^3$ partition sets. When compared with Fig. 3, Fig. 8 clearly depicts the effect of the dimensionality problem on the estimation performance. Whereas the alternative adaptive scheme in Fig. 3, which uses three marginal estimators, attains attitude errors on the order of 0.1 deg, the attainable mean attitude estimation error of this adaptive scheme is approximately 1 deg.



a) X marginal, t=0 **b) Y marginal, t=0** **c) Z marginal, t=0**

d) X marginal, t=4 h **e) Y marginal, t=4 h** **f) Z marginal, t=4 h**

g) X marginal, t=40 h **h) Y marginal, t=40 h** **i) Z marginal, t=40 h**

Fig. 7 True (dashed line) and reconstructed (solid line) TAM PDF marginals at times $t = 0$, 4, 40 h; single run.

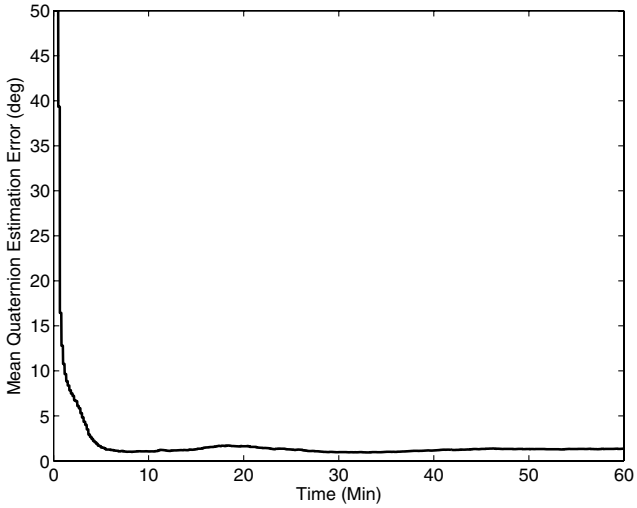


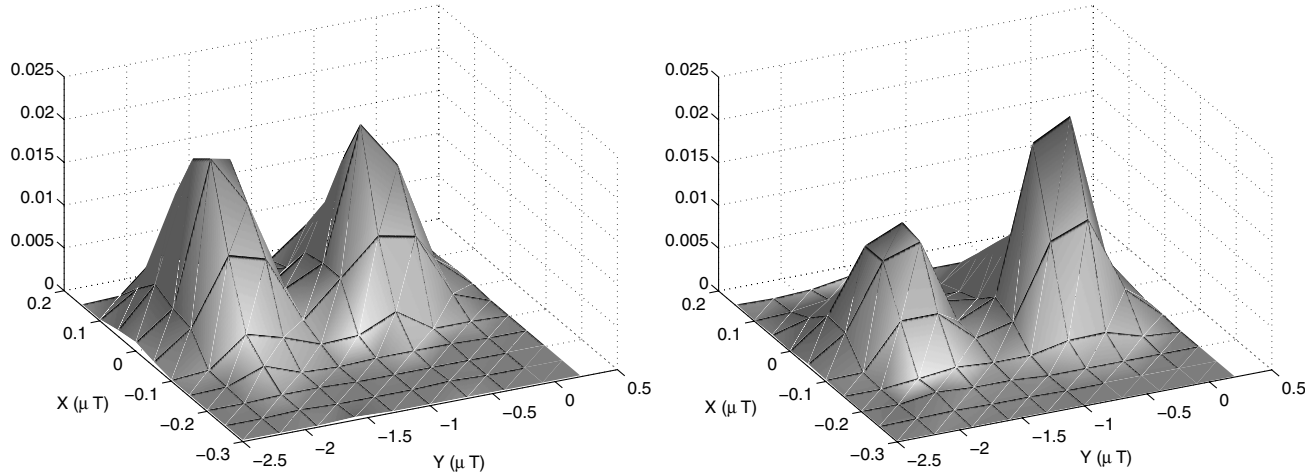
Fig. 8 Mean attitude estimation error of the adaptive GA-QPF using a three-dimensional histogram estimator with 10^3 partition sets; 1000 Monte Carlo runs.

Both Figs. 9 and 10 show the estimated and the true measurement noise joint PDF. Clearly, the histogram estimator manages to capture the essential properties of the joint PDF within a 1 h interval.

VI. Conclusions

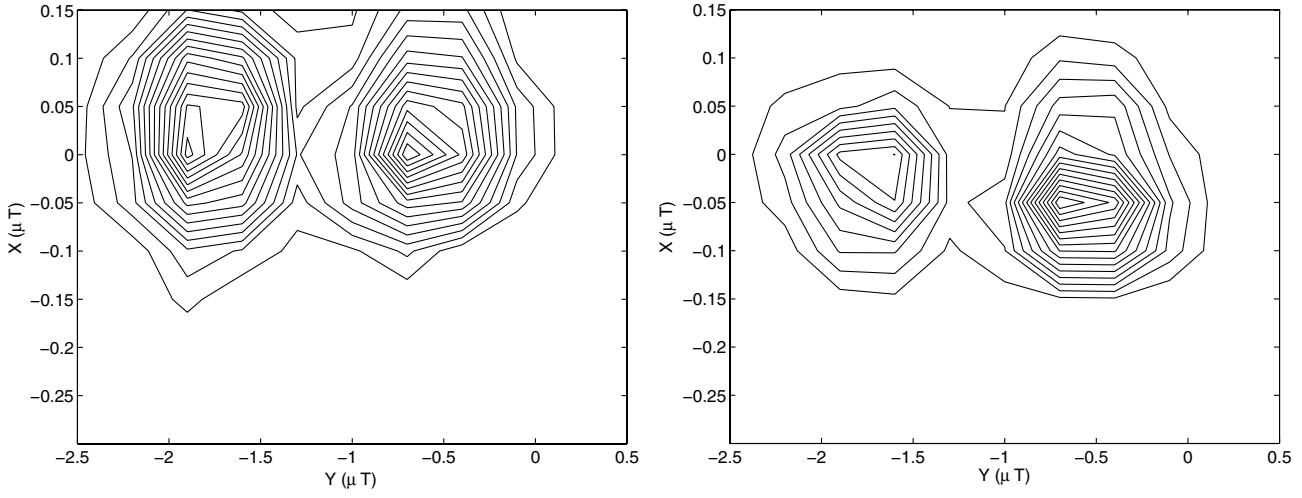
An adaptive algorithm is presented for the estimation of spacecraft attitude from vector observations in scenarios characterized by a high uncertainty of the measurement noise distribution. The filter is based on the recently presented genetic-algorithm-embedded quaternion particle filter, that has been shown to effectively estimate the attitude in nonlinear and non-Gaussian scenarios. In the GA-QPF, the attitude is represented via quaternion particles, thus the quaternion norm constraint is naturally maintained, avoiding the need for ad hoc and external normalization procedures. A genetic algorithm is used to generate a maximum likelihood estimate of the gyro biases, thus alleviating the potential computational burden problem associated with the number of required particles.

In the present work, the GA-QPF is made adaptive by equipping it with a histogram estimator that facilitates the estimation of the measurement noise distribution via a statistical analysis of the filter-generated innovations process. Thus, the filter estimates the measurement noise distribution on the fly, along with the spacecraft attitude quaternion. A simulation study is used to assess the performance of the adaptive algorithm in a scenario involving real,



a) Estimated **b) True**

Fig. 9 Estimated and true TechSAT TAM noise joint PDF (sliced at $Z = 0$) after 1 h.



a) Estimated **b) True**

Fig. 10 Estimated and true TechSAT TAM noise joint PDF (level curves at $Z = 0$) after 1 h.

non-Gaussian data obtained from the Technion's TechSAT satellite. As part of this study, the adaptive filter is compared with the previously introduced nonadaptive GA-QPF. The study demonstrates the viability of the new algorithm and shows that, if an estimate (or an upper bound) of the measurement noise distribution's support is available, an accurate estimate of a non-Gaussian distribution is obtained even when starting from a highly inaccurate initial estimate.

Acknowledgments

This research was supported by the Israel Science Foundation (grant no. 1032/04) and by the Technion's Asher Space Research Fund. The authors thank Moshe Guelman and Alexander Shiryaev of the Technion's Asher Space Research Institute for providing the TechSAT data.

References

[1] Wahba, G., "A Least-Squares Estimate of Satellite Attitude. Problem 65-1," *SIAM Review*, Vol. 7, No. 3, July 1965, p. 409. doi:10.1137/1007077

[2] Lefferts, E. J., Markley, F. L., and Shuster, M. D., "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, Sept.–Oct. 1982, pp. 417–429. doi:10.2514/3.56190

[3] Bar-Itzhack, I. Y., and Oshman, Y., "Attitude Determination from Vector Observations: Quaternion Estimation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-21, No. 1, Jan. 1985, pp. 128–136. doi:10.1109/TAES.1985.310546

[4] Gai, E., Daly, K., Harrison, J., and Lemos, L., "Star-Sensor-Based Satellite Attitude/Attitude Rate Estimator," *Journal of Guidance, Control, and Dynamics*, Vol. 8, No. 5, Sept.–Oct. 1985, pp. 560–565. doi:10.2514/3.56393

[5] Crassidis, J. L., and Markley, F. L., "Unscented Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 4, Aug. 2003, pp. 536–542. doi:10.2514/2.5102

[6] Oshman, Y., and Carmi, A., "Attitude Estimation from Vector Observations Using a Genetic Algorithm-Embedded Quaternion Particle Filter," *Journal of Guidance, Control, and Dynamics*, Vol. 29, No. 4, July–Aug. 2006, pp. 879–891. doi:10.2514/1.17951

[7] Doucet, A., de Freitas, N., and Gordon, N., (eds.), *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and Information Science, Springer, New York, 2001.

[8] Doucet, A., Godsill, S., and Andrieu, C., "On Sequential Monte Carlo Sampling Methods for Bayesian Filtering," *Statistics and Computing*, Vol. 10, No. 3, July 2000, pp. 197–208. doi:10.1023/A:1008935410038

[9] Geweke, J., "Bayesian Inference in Econometric Models Using Monte Carlo Integration," *Econometrica: Journal of the Econometric Society*, Vol. 57, No. 6, Nov. 1989, pp. 1317–1339. doi:10.2307/1913710

[10] Wah, W. K., "A New Adaptive Density Estimator for Particle-Tracing Radiosity," *Proceedings of the 8th Pacific Conference on Computer Graphics and Applications*, IEEE Computer Society, Washington, D.C., 2000, p. 62, ISBN 0-7695-0868-5.

[11] Shiryaev, A., Waller, R., Roznov, M., and Guelman, M., "Design and Experimental Testing of Three-Axis Satellite Attitude Control Systems That Use Only Magnetic Torquers," Norman and Helen Asher Space Research Inst., Technion—Israel Inst. of Technology TR 2002-07, June 2002.