REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
30-08-2011	Final Report	08/01/2008 - 05/31/2011

4. TITLE AND SUBTITLE		5a. CONTRACT NUMBER
Multiple Hypothesis Correlation for Space Situational Awareness		
		5b. GRANT NUMBER
		FA9550-08-1-0419
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		**5d. PROJECT NUMBER**
Aubrey B. Poore, Ph.D.		
Joshua T. Horwood, Ph.D.		5e. TASK NUMBER
		5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Numerica Corporation 4850 Hahns Peak Drive, Suite 200 Loveland, CO 80538	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
USAF, AFRL AF Office of Scientific Research 875 N. Randolph St., Room 312 Arlington, VA 22203	AFOSR/PKR2
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
	AFRL-OSR-VA-TR-2012-0284

12. DISTRIBUTION/AVAILABILITY STATEMENT

DIST. A

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The primary objective of this program is to perform the necessary basic research to support the development of a statistical, multiple hypothesis tracker (MHT) for space surveillance. Such a MHT framework can serve as the next generation space surveillance system to maintain the space catalog, to identify uncorrelated tracks, and to support conjunction analysis and sensor resource management. Key components in such a system include a consistent characterization of uncertainty, physical modeling, multiple model filtering, and the association problem of determining which tracklets/measurements emanate from which object. To achieve a consistent characterization of uncertainty, Numerica has developed an adaptive Gaussian sum filter which correctly represents and propagates uncertainties and adaptively selects the correct the number of Gaussians in the mixture. Realtime online metrics support the coarsening and refining of the filter to maintain consistent uncertainty. Several papers have been published on these results. Numerica also acquired real orbital data and processing of this data has commenced using a new prototype MHT speci

15. SUBJECT TERMS

Multiple Hypothesis Tracking, Anomaly Detection, Multiple Models, Nonlinear Filtering, Uncertainty Management, Data Fusion, Correlation Ambiguity, Uncorrelated Tracks

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Aubrey Poore
UNCLASS	UNCLASS	UNCLASS	SAR	25	19b. TELEPHONE NUMBER *(Include area code)* 970-461-2000 x220

Numerica Corporation
4850 Hahns Peak Drive, Suite 200
Loveland, CO 80538

Office: (970) 461-2000
Fax: (970) 461-2004
E-mail: info@numerica.us

www.numerica.us

Final Technical Report
Multiple Hypothesis Correlation for Space Situational Awareness

In Response to: BAA 2007–08 Topic D – Space Situational Awareness

Contract Number: FA9550-08-1-0419

Government Sponsor: Air Force Office of Scientific Research

Technical Monitor:

Dr. Kent Miller
AFOSR
875 N. Randolph Street, Room 3112
Arlington, VA 22203
Voice: (703) 696-8573
Email: kent.miller@afosr.af.mil

Date: August 29, 2011

Prepared by:

Aubrey B. Poore, Ph.D.
Joshua T. Horwood, Ph.D.
(970) 461-2000, ext. 220

1 Objectives

The space situational awareness (SSA) mission encompasses intelligence, surveillance of all space objects, and the prediction of space events, threats, and activities. Specifically, the mission requires knowledge of the object trajectory (orbit) and type (active satellite, decommissioned satellite, debris, etc.) for all objects in the near-Earth environment. For objects that are currently under active control, one would also like to know their current activities, capabilities, and expected future actions. These types of higher-level knowledge must be supported by a significant amount of raw data collection. Space surveillance is that component of SSA focused on the detection of space objects and the use of multisource data to track and identify space objects. Currently, it can take weeks to establish correct orbits on objects of interest. Our approach to this problem is to use a full multiple hypothesis tracking (MHT) correlation algorithm developed under previous AFOSR grants, but modified to accommodate space surveillance. This new algorithm will be called multiple hypothesis correlation (MHC).

While manual and rudimentary automated correlation schemes may suffice for known orbits or for a small number of uncorrelated tracks (UCTs), an automated system will be required in the future to handle an anticipated substantial increase in the number of space objects arising from space events and higher-resolution sensors. Core challenges in the development of such an automated system include dense target environments, the need to establish robust tracks in a timely fashion, limited data with large coverage gaps, system biases and residual biases and their impact on orbit state propagation, and unresolved closely spaced objects.

From a multiple target tracking perspective, our objectives in this program are to initiate orbit trajectories on newly observed objects as quickly as possible; to improve our ability to maintain continuous tracks in the presence of large coverage gaps; and to piece together UCTs from different sensors. To meet these objectives, research is required in the following three algorithm components.

1. **Uncertainty Modeling and Management.** A prerequisite to a statistically rigorous and information theory-based approach to tracking objects in space, sensor resource management, and conjunction analysis is the *consistent characterization of the uncertainty* exhibited by a stochastic state. In space surveillance, the challenges to achieving this objective are the nonlinear dynamics, long-term propagations, and sparse data environment requiring the development of advanced methods such as the adaptive Gaussian sum filter and the sliding window batch estimation filter. The status of the current research towards this objective is detailed in Subsection 1.1.1.

2. **Multiple Hypothesis Tracking.** The *data association or correlation* problem is that of determining which reports (tracks, tracklets, measurements, features) emanate from which object. For widely spaced objects, reports are assigned to objects if they are within some gate (uncertainty region) of the object. (The method is called nearest neighbor.) On the other hand, multisensor association (correlation) in a dense target environment is a challenging problem and *multiple hypothesis tracking (MHT)* techniques offer the best potential solution. For marginally detectable or dim targets, MHT can be augmented by track-before-detect algorithms that output detections, but not formal tracks, or possibly enhanced resolution techniques. Thus, what is needed is an MHT that can *adapt* to the complexity of the problem thereby covering the range between widely spaced and closely spaced objects as well as luminous and dim targets. Perhaps one of the most overlooked problems in the use of multisource data and modern data fusion and correlation algorithms is that of *sensor biases*. A major new necessity is *joint bias estimation and correlation*, which may be needed for overlapping coverage, but is mandatory in handover situations, especially in large coverage gaps. Just as it is essential to capture the uncertainty in state estimation via nonlinear filtering techniques, it is also essential to capture the uncertainty in the association process. We call this *association ambiguity* and we achieve this objective by producing the "probability of association." Such an assessment is

particularly important for processing UCTs and closely spaced objects. The status of the current research towards this objective is detailed in Subsection 1.1.2.

3. **Unresolved Closely Spaced Objects.** Sensor resolution is a major obstacle in establishing orbits on individual objects with limited data in a timely fashion; thus, it becomes necessary to enhance the resolution of sensors and extract consistent sets of measurements from which individual orbits on individual objects may be established. The status of the current research towards this objective is detailed in Subsection 1.1.3.

While the above objectives remain a focus, Numerica continues the work on network-centric tracking that was initially started with AFOSR funding and has now transitioned to a Phase II SBIR at AFRL/SNAT and a Phase II SBIR with Department of the Army with a transition path to the SIAP Joint Program Office.

1.1 Status of Effort

The status of the current research program is summarized in this section. The individual objectives outlined above are each addressed in the following subsections.

1.1.1 Uncertainty Modeling and Management

In this subsection, we review the general Bayesian framework for nonlinear filtering and describe the most promising methods for achieving correct uncertainty consistency including the Gaussian sum filter and the sliding window batch estimation filter.

1.1.1.1 Nonlinear Estimation and Filtering The problem of nonlinear filtering requires the definition of dynamical and measurement models. It is assumed that the dynamic state $\boldsymbol{x}(t) \in \mathbb{R}^n$ at time t evolves according to the continuous-time stochastic model,

$$\boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}(t), t) + \mathbf{G}(\boldsymbol{x}(t), t)\,\boldsymbol{w}(t), \tag{1}$$

where $\boldsymbol{f} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$, $\mathbf{G} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n \times m}$, and $\boldsymbol{w}(t)$ is an m-dimensional Gaussian white noise process having spectral density matrix $\mathbf{Q}(t)$ with $\mathrm{E}[\boldsymbol{w}(t)\,\boldsymbol{w}(\tau)^T] = \mathbf{Q}(t)\,\delta(t - \tau)$, $t > \tau$. In particular, in (1), the function $\boldsymbol{f}$ encodes the deterministic force components of the dynamics (e.g., gravity, drag, etc.) while the process noise term models the stochastic acceleration. In many tracking applications, it is often convenient to work with a discrete-time formulation of the dynamical model which assumes the form

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}_k(\boldsymbol{x}_k) + \mathbf{G}_k(\boldsymbol{x}_k)\,\boldsymbol{w}_k, \tag{2}$$

where $\boldsymbol{x}_k = \boldsymbol{x}(t_k)$, $\boldsymbol{f}_k : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{G}_k : \mathbb{R}^n \to \mathbb{R}^{n \times m}$, and $\{\boldsymbol{w}_k\}$ is an m-dimensional zero-mean Gaussian white noise sequence possessing a spectral density matrix $\mathbf{Q}_k \in \mathbb{R}^{m \times m}$ such that $\mathrm{E}[\boldsymbol{w}_k\,\boldsymbol{w}_j^T] = \mathbf{Q}_k\,\delta_{kj}$. In space surveillance, the process noise term is often very small and discarded. In such situations, the function $\boldsymbol{f}_k$ is just the solution flow corresponding to the continuous model (1) with $\boldsymbol{w}(t) = 0$.

A sequence of measurements $\boldsymbol{Z}_k \equiv \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k\}$ is related to the corresponding kinematic states $\boldsymbol{x}_k$ via measurement functions $\boldsymbol{h}_k : \mathbb{R}^n \to \mathbb{R}^p$ according to the discrete-time measurement model

$$\boldsymbol{z}_k = \boldsymbol{h}_k(\boldsymbol{x}_k) + \boldsymbol{v}_k. \tag{3}$$

In this equation, $\{\boldsymbol{v}_k\}$ is a p-dimensional zero-mean Gaussian white-noise sequence with $\mathrm{E}[\boldsymbol{v}_k\,\boldsymbol{v}_j^T] = \mathbf{R}_k\,\delta_{kj}$. More general filter models can be formulated from measurement models with non-Gaussian or correlated (e.g., colored) noise terms and sensor biases.

$$p(\boldsymbol{x}_k|\boldsymbol{Z}_{k-1})$$

Correction Step:
$$p(\boldsymbol{x}_k|\boldsymbol{Z}_k) = \frac{p(\boldsymbol{z}_k|\boldsymbol{x}_k)\,p(\boldsymbol{x}_k|\boldsymbol{Z}_{k-1})}{p(\boldsymbol{z}_k|\boldsymbol{Z}_{k-1})}$$
where
$$p(\boldsymbol{z}_k|\boldsymbol{Z}_{k-1}) = \int p(\boldsymbol{z}_k|\boldsymbol{x}_k)\,p(\boldsymbol{x}_k|\boldsymbol{Z}_{k-1})\,\mathrm{d}\boldsymbol{x}_k$$

Prediction Step:
$$p(\boldsymbol{x}_k|\boldsymbol{Z}_{k-1}) = \int p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\,p(\boldsymbol{x}_{k-1}|\boldsymbol{Z}_{k-1})\,\mathrm{d}\boldsymbol{x}_{k-1}$$

Initialization:
$$p(\boldsymbol{x}_0|\boldsymbol{Z}_0) = p(\boldsymbol{x}_0)$$

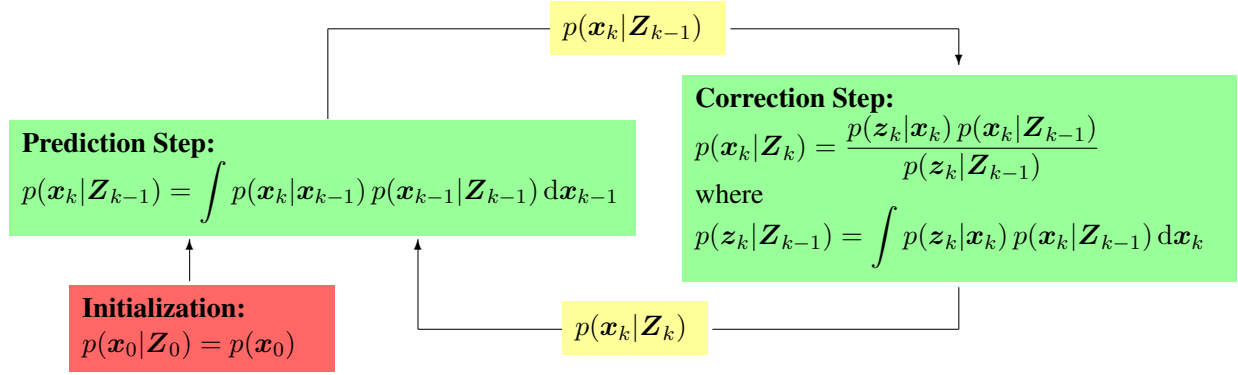$$p(\boldsymbol{x}_k|\boldsymbol{Z}_k)$$

Figure 1: The predictor-corrector step for the recursive Bayesian state estimator.

In the Bayesian approach to dynamic state estimation [1], one constructs the *posterior* probability density function (PDF) of the state based on information of a *prior* state and received measurements. Encapsulating all available statistical information, the posterior PDF $p(\boldsymbol{x}_k|\boldsymbol{Z}_k)$ may be regarded as the complete solution to the estimation problem and various optimal state estimates can be computed from it.

In a recursive filtering approach, measurement data is processed *sequentially*, rather than as a batch. Given the initial density of the state $p(\boldsymbol{x}_0) \equiv p(\boldsymbol{x}_0|\boldsymbol{Z}_0)$, the PDF $p(\boldsymbol{x}_k|\boldsymbol{Z}_k)$ is obtained recursively in two stages, namely prediction and correction, as illustrated in the flowchart of Figure 1. In the case of discrete-time dynamics, the former is obtained from the transitional density $p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})$ in conjunction with the *Chapman-Kolmogorov equation* defined in the prediction step box of Figure 1. In the case of continuous-time dynamics, the time evolution of the predicted PDF $p = p(\boldsymbol{x}, t|\boldsymbol{Z}_{k-1})$ for $t > t_{k-1}$ is governed by the *Fokker-Planck-Kolmogorov equation (FPKE)* [2]

$$\frac{\partial p}{\partial t} = -\nabla_{\boldsymbol{x}}^T(p\boldsymbol{f}) + \frac{1}{2}\,\mathrm{tr}\big[\nabla_{\boldsymbol{x}}\nabla_{\boldsymbol{x}}^T(p\mathbf{G}\mathbf{Q}\mathbf{G}^T)\big], \tag{4}$$

where $\nabla_{\boldsymbol{x}}$ is the gradient with respect to $\boldsymbol{x}$ viewed as a column operator. In the measurement update stage, also called the correction or fusion step, the density $p(\boldsymbol{z}_k|\boldsymbol{x}_k)$ is evaluated from the measurement model (3). The term $p(\boldsymbol{z}_k|\boldsymbol{Z}_{k-1})$ in the denominator of the correction step is called the *prediction error* and appears in the likelihood ratios for scoring an assignment of a report (i.e., $\boldsymbol{z}_k$) to a track (see, for example, Poore [3]). Thus, its accurate evaluation is critical for correct data/track association (correlation).

Analytical solutions to the FPKE and to the prediction and correction equations in Figure 1 are generally intractable and are only known in a few restrictive cases. In practice, models are nonlinear and states can be non-Gaussian; one must be content with an approximate or suboptimal algorithm for the Bayesian state estimator. While the extended Kalman filter (EKF) and unscented Kalman filter (UKF) are used extensively in air and missile tracking, they only represent state uncertainties by a covariance matrix and this need not be adequate in the space surveillance environment. Because of the need to propagate uncertainties over extended time intervals in the absence of measurement updates, higher-order cumulants (e.g., skewness, excess kurtosis) can become non-negligible and must be accounted for in order to achieve uncertainty consistency.

In this AFOSR program and in another STTR funded by AFOSR (topic AF09-BT11), we developed a variety of nonlinear filters specialized from the general Bayesian state estimator and showed that many of them correctly capture statistics beyond a Gaussian state and covariance and retain uncertainty consistency over long propagations. Table 1 summarizes some of the advantages and disadvantages of these filters. Although no special attention was given to the extended Kalman filter (EKF) and particle filter, they are nevertheless included in the table for completeness. We do not propose their use in space surveillance applications. The UKF, while costing about the same to run as the EKF, is more accurate and numerically

Table 1: Comparison of some nonlinear filters.

Filter	Advantages	Disadvantages
Extended Kalman Filter	Ubiquitous	Uncertainty is only represented by a covariance; partial derivatives of the dynamics and measurement models are required; generally less accurate than the UKF
Unscented and Gauss-Hermite Filters	Derivative-free	Uncertainty is only represented by a covariance
Edgeworth Filter	Can represent any uncertainty within a desired accuracy by increasing the number of cumulants in the representation; derivative-free	Computationally feasible only for weakly non-Gaussian densities
Reduced Edgeworth Filter	As in the Edgeworth filter, but optimized to account for possible near Gaussianity in one or more state space coordinates	As in the Edgeworth filter
Adaptive Gaussian Sum Filter	Can represent any uncertainty within a desired accuracy by increasing the number of Gaussians in the mixture; straightforward to implement by running a bank of UKFs in parallel	Representation of weakly non-Gaussian densities is expensive; computationally feasible only in a few dimensions
State Transition Tensor Filter	Can represent any uncertainty within a desired accuracy by increasing the order of the Taylor series expansion; very effective for representing both weakly and highly non-Gaussian densities	Partial derivatives of the dynamics and measurement models are required; integration of the STT ODEs is expensive
Particle Filter	Can represent any uncertainty within a desired accuracy by increasing the number of "particles" in the representation	Very expensive; accuracy scales as $\mathcal{O}(N^{-1/2})$ where N is the number of particles

more stable. The particle filter is usually reserved as a "last resort" and would be too expensive to use in an operational system.

Figure 2 shows how the root-mean-square (RMS) cost error, a metric for quantifying uncertainty consistency, scales according to the dimension of the representation (i.e., the number of degrees of freedom required to store the state PDF) and the number of sigma points (i.e., the number of orbital propagations required in the filter prediction step). The "ideal" filter would achieve a negligible RMS cost error with minimal computational cost either in terms of the dimension of the represented PDF or the number of sigma points. The adaptive Gaussian sum filter (AGSF) and the state transition tensor filter (STTF)* show the most promise for accurately representing a state's uncertainty. *However, beware that the STTF results are based only on unperturbed Keplerian dynamics in which the implementation of the filter is trivial. It remains to see how the STTF performs under a full nonlinear gravity model.*

---

*The state transition tensor filter is not based on the propagation of sigma points, so it does not appear in the right figure.

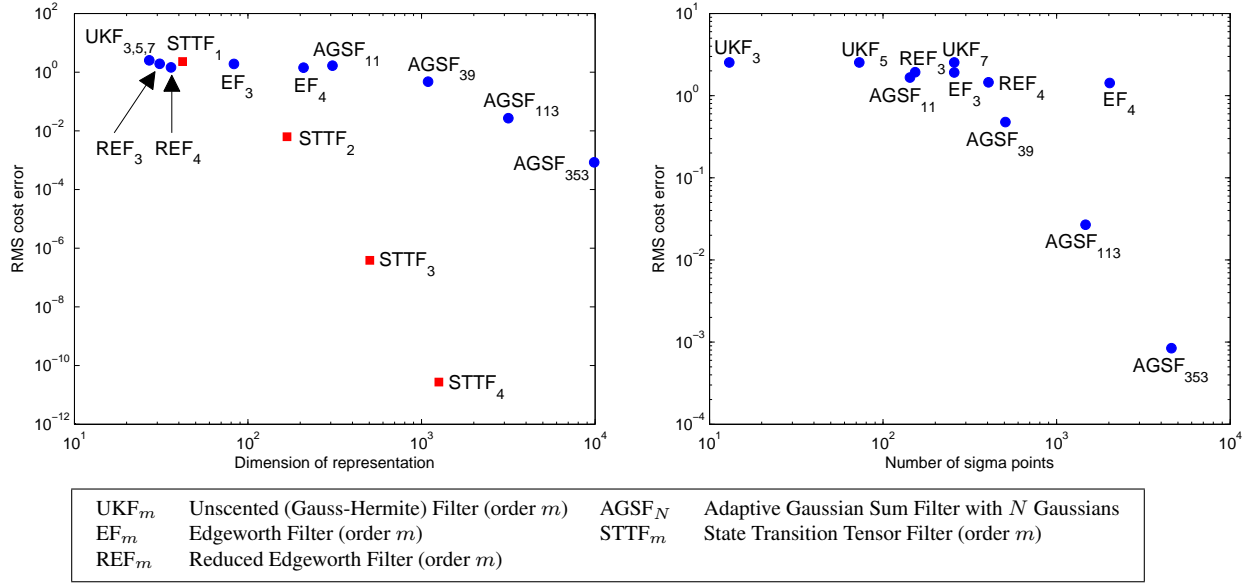| UKF$_m$ | Unscented (Gauss-Hermite) Filter (order m) | AGSF$_N$ | Adaptive Gaussian Sum Filter with N Gaussians |
| EF$_m$ | Edgeworth Filter (order m) | STTF$_m$ | State Transition Tensor Filter (order m) |
| REF$_m$ | Reduced Edgeworth Filter (order m) | | |

Figure 2: Dependence of the root-mean-square (RMS) cost error on the dimension of the representation (left) and the number of sigma points (right). Blue disks (●) imply use of the EGM96 degree and order 70 gravity model while red squares (■) imply use of unperturbed Keplerian dynamics.

1.1.1.2 Gaussian Sum Filters

A Gaussian sum is a mixture density of the form

$$p(\boldsymbol{x}) = \sum_{\alpha=1}^{N} w_\alpha \, \mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}_\alpha, \mathbf{P}_\alpha),$$

where the weights w_α are non-negative scalars which sum to unity and $\mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \mathbf{P})$ denotes the Gaussian PDF with mean $\boldsymbol{\mu}$ and covariance $\mathbf{P}$; i.e.,

$$\mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \mathbf{P}) = \frac{1}{\sqrt{\det(2\pi\mathbf{P})}} \exp\left[-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \mathbf{P}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})\right]. \tag{5}$$

The Gaussian sum filter (GSF) is based on a fundamental result of Alspach and Sorenson [4] which states that any PDF can be approximated arbitrarily close (in the L^1 sense) by a weighted sum (mixture) of Gaussian PDFs henceforth called a Gaussian sum. Thus, Gaussian sums provide a mechanism for modeling non-Gaussian densities and for more accurately approximating the solution of the FPKE. Computationally, the GSF has the added advantage of being *parallelizable* since filters such as the EKF or UKF act *independently* on each component Gaussian in the prediction and correction steps. With regards to updating the weights within the filter, such a scheme is clearly dictated from Bayes' rule following a measurement or track update (fusion). However, there is not complete agreement on how the weights should be updated (if at all) following a prediction.

One key feature of the new GSF developed in this effort and communicated in the paper [5] is the *absence* of a procedure to update the mixture weights following the implementation of the filter prediction step. The accuracy and consistency of the propagated uncertainty is ensured by representing the Gaussian sum in (traditional or alternate) equinoctial orbital elements with component means, covariances, and weights initially selected (by solving an L^2 optimization problem offline) such that the (square-root version of the) UKF [6], when acting in parallel on each component, accurately approximates the solution of the FPKE. The number of Gaussian components required to achieve an accurate approximation is chosen *adaptively* based
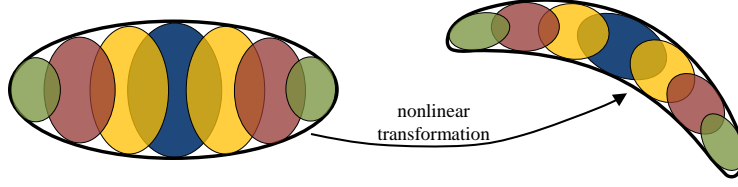
Figure 3: Depiction of a single Gaussian and its Gaussian sum approximation undergoing a nonlinear transformation.

on the length of the propagation time and the initial error (standard deviation) along the radial direction (semi-major axis coordinate). Consequently, by representing the PDF in the equinoctial elements, the new algorithm achieves superior computational efficiency because it only requires a Gaussian sum along *one dimension*.

While others have proposed methods for adapting the weights based on various *online L^2* optimization criteria [7–9], we have found that the application of these methods to our existing GSF does not improve uncertainty consistency, but rather causes the accuracy of the Gaussian sum representation to degrade slightly [5]. We attribute these findings to the fact that our GSF is already very efficient because it solves an L^2 optimization once *offline* which is used thereafter in any scenario through a lookup table. In what follows, we motivate and define the specific optimization problem we solve.

The underlying optimization problem fundamental to our GSF requires the refinement of a single Gaussian PDF into a Gaussian sum. The component means, covariances, and weights of the mixture are chosen such that each Gaussian component remains Gaussian (up to a prescribed accuracy) when propagated by the UKF under the nonlinear dynamics. The refinement methodology is illustrated in Figure 3. Under a nonlinear transformation, a Gaussian (represented by the thick black ellipse) need not be mapped to a Gaussian (e.g., the level surfaces of the transformed distribution could look crescent-shaped). However, in a sufficiently small neighborhood, any (smooth) nonlinear map will be approximately linear. Consequently, Gaussians with smaller covariances (represented by the colored elliptic disks) remain more Gaussian than those with larger covariances under the nonlinear mapping. Therefore, a Gaussian sum refined by approximating each constituent Gaussian by a finer Gaussian sum will exhibit better behavior through nonlinear transformations. It suffices to optimally refine the unit one-dimensional Gaussian $\mathcal{N}(x, 0, 1)$; refinement of a multivariate Gaussian with an arbitrary covariance is obtained by a series of linear transformations detailed in [5].

We now derive a Gaussian sum approximation of the unit one-dimensional Gaussian $\mathcal{N}(x; 0, 1)$. The problem has an obvious trivial solution which perfectly approximates the target, namely the Gaussian sum with a single component. However, as motivated earlier, the idea is to construct a Gaussian sum approximation whose component standard deviations σ_α are smaller than some fixed value $\sigma < 1$ so that the Gaussian sum more accurately represents the true distribution under a nonlinear transformation. This requirement leads to a constrained nonlinear optimization problem. That said, we assume the target Gaussian sum has the form

$$p_{GS}(x) = \sum_{\alpha=1}^{N} w_\alpha \, \mathcal{N}(x; \mu_\alpha, \sigma_\alpha^2),$$

for some predetermined length $N > 1$ and define

$$E = \frac{1}{2} \|\mathcal{N}(x; 0, 1) - p_{GS}(x)\|_{L^2}^2 = \frac{1}{2} \int_{-\infty}^{\infty} \left[\mathcal{N}(x; 0, 1) - \sum_{\alpha=1}^{N} w_\alpha \, \mathcal{N}(x; \mu_\alpha, \sigma_\alpha^2) \right]^2 \, \mathrm{d}x. \qquad (6)$$

The weights w_α, means μ_α, and standard deviations σ_α of each component are determined from the solution
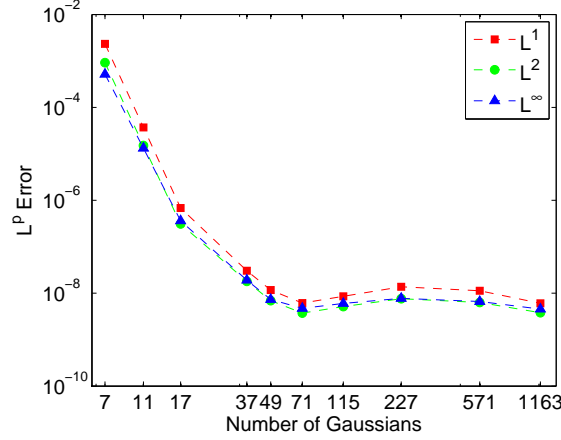
Figure 4: Plots of the L^p error between the unit one-dimensional Gaussian and its Gaussian sum approximation obtained using the suboptimal optimization algorithm.

of the following L^2 optimization problem:

$$\underset{w_1,\ldots,w_N,\,\mu_1,\ldots,\mu_N,\,\sigma_1,\ldots,\sigma_N}{\text{Minimize}} \quad E$$

$$\text{Subject to} \quad \sum_{\alpha=1}^{N} w_\alpha = 1, \quad w_\alpha \geq 0, \quad \alpha = 1,\ldots,N, \tag{7}$$

$$\mu_1 \leq \mu_2 \leq \cdots \leq \mu_N,$$

$$\sigma_\alpha \leq \sigma < 1, \quad \alpha = 1,\ldots,N.$$

Noting the identity

$$\int \mathcal{N}(\boldsymbol{x};\boldsymbol{\mu}_1,\mathbf{P}_1)\,\mathcal{N}(\boldsymbol{x};\boldsymbol{\mu}_2,\mathbf{P}_2)\,\mathrm{d}\boldsymbol{x} = \mathcal{N}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2;\mathbf{0},\mathbf{P}_1 + \mathbf{P}_2),$$

it follows that (6) reduces to

$$E = \frac{1}{2}\boldsymbol{w}^T\mathbf{M}\boldsymbol{w} - \boldsymbol{w}^T\boldsymbol{n} + \frac{1}{4\sqrt{\pi}}, \tag{8}$$

where

$$(\boldsymbol{w})_\alpha = w_\alpha, \quad (\boldsymbol{n})_\alpha = \mathcal{N}(\mu_\alpha;0,\sigma_\alpha^2 + 1), \quad (\mathbf{M})_{\alpha\beta} = \mathcal{N}(\mu_\alpha - \mu_\beta;0,\sigma_\alpha^2 + \sigma_\beta^2).$$

The minimization of (8) over the individual weights, means, and standard deviations subject to the constraints in (7) is a difficult nonlinear optimization problem. A suboptimal yet computationally tractable algorithm which approximates the solution of (7) can be obtained as follows. Specifically, we constrain each of the component Gaussians to have a *common standard deviation* $\sigma_\alpha = \sigma < 1$ and propose *fixing the means* μ_α according to

$$\mu_\alpha = -m + \sigma(\alpha - 1), \tag{9}$$

for $\alpha = 1,\ldots,N$, where $m > 0$ and $N = \lceil 1 + 2m/\sigma \rceil$. Note that the means are evenly distributed with the left-most and right-most components located at $x \approx \pm m$. In order to ensure adequate accuracy around the tails, we propose setting $m = 4$ if $\sigma \geq \frac{1}{2}$, otherwise $m = 6$. With these additional constraints, the optimization problem (7) reduces to a quadratic programming problem which is straightforward to solve using elementary methods.

Figure 4 shows plots of the L^1, L^2, and L^∞ error between the unit one-dimensional Gaussian and
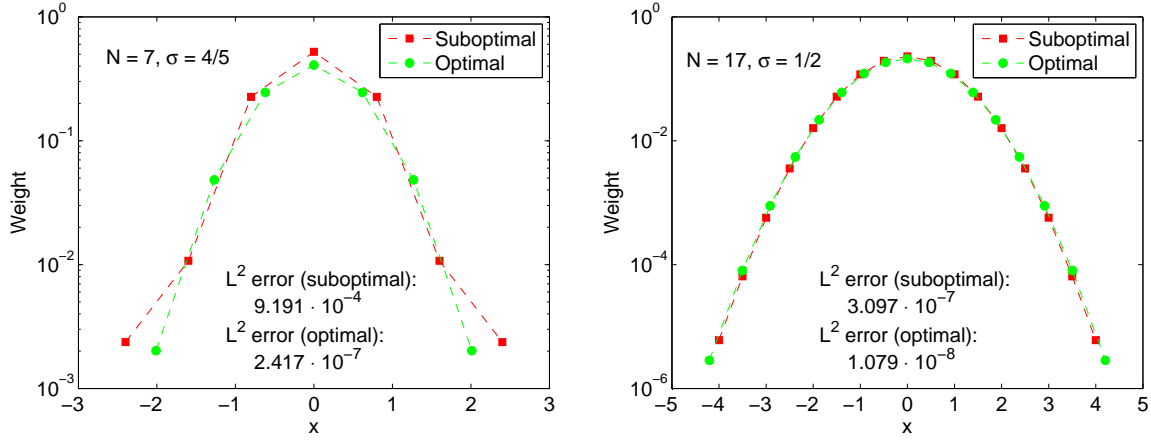
Figure 5: Scatter plots of the component weights and means computed using the suboptimal algorithm (i.e., minimizing only over the weights) and by solving the full (optimal) optimization problem (i.e., minimizing over the weights, means, and standard deviations).

its Gaussian sum approximation computed using the suboptimal approach described above. Up to around $N \approx 100$, the three L^p errors all decrease, but they plateau to around 10^{-8} for $N \gtrsim 100$. We believe this plateau is not because of numerical stability issues (the matrix $\mathbf{M}$ in (8) has a condition number on the order of 10^5) but simply because we are *fixing the standard deviations and means* via (9) and only minimizing over the unknown weights; this need not be optimal. Notwithstanding these comments, the simulation studies considered in the paper [5] use this (suboptimal) refinement scheme with N as high as 1000 without any adverse effects.

We conclude this section by presenting some preliminary solutions of the full optimization problem (7). We have found that the numerical conditioning of this problem worsens as the target length N of the Gaussian sum increases thereby rendering double precision floating-point arithmetic largely inadequate. To facilitate an accurate solution, we have used the MAPLE 'Optimization' package which takes advantage of built-in library routines provided by the Numerical Algorithms Group with the ability to call them within an arbitrary-precision software floating-point environment.

Figure 5 shows various scatter plots of the component weights and means computed using two different methods. The first method uses the suboptimal approach which fixes the means according to (9) and subsequently optimizes over the weights only. The second method solves the full optimization problem (7). Each subplot assumes a different target length N and standard deviation σ. One interesting observation is that the optimal component standard deviations lie on the constraint boundary; i.e., $\sigma_\alpha = \sigma$ for all α. Most importantly, by solving the full optimization problem, the locations of the means are no longer uniformly distributed and the resulting L^2 error is smaller than that obtained by solving the suboptimal problem. *In particular, to get the L^2 error down to 10^{-7} requires about $N = 28$ Gaussians using the suboptimal method but only about $N = 9$ Gaussians when solving the full optimization problem!* We emphasize that the full problem, although very expensive to solve, can nevertheless be done *offline*. One can generate a library of Gaussian sum approximations to $\mathcal{N}(x; 0, 1)$ for various values of N and σ.

1.1.1.3 Batch Estimation and Sliding Window Batch Estimation Given a sequence of m measurements $\boldsymbol{Z}_m \equiv \{\boldsymbol{z}_1, \dots, \boldsymbol{z}_m\}$ at times $t_1, \dots, t_m$ commensurate with the model (3), the objective of initial orbit determination (IOD) and batch estimation is to obtain a representation of the (joint) posterior PDF $p(\boldsymbol{x}_0, \dots, \boldsymbol{x}_m | \boldsymbol{Z}_m)$, where $\boldsymbol{x}_k$ denotes the dynamical state of the system at time t_k, and to extract meaningful statistics (e.g., mean, covariance) from it in a consistent manner. It is assumed that the state evolves

according to the discrete-time model

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}_k(\boldsymbol{x}_k) + \boldsymbol{w}_k, \tag{10}$$

where $\{\boldsymbol{w}_k\}$ is a zero-mean white noise sequence with $\mathrm{E}[\boldsymbol{w}_k\,\boldsymbol{w}_j^T] = \mathbf{Q}_k\,\delta_{kj}$. The following independence assumptions are implied between the prior $\boldsymbol{x}_0$, the measurement noise sequence $\{\boldsymbol{v}_k\}$, and the process noise sequence $\{\boldsymbol{w}_k\}$:

$$\mathrm{E}[\boldsymbol{x}_0\,\boldsymbol{w}_k^T] = 0, \quad \mathrm{E}[\boldsymbol{x}_0\,\boldsymbol{v}_k^T] = 0, \quad \mathrm{E}[\boldsymbol{v}_k\,\boldsymbol{w}_j^T] = 0.$$

Appealing to Bayes' rule and the above assumptions, the joint posterior PDF is derived in Jazwinski [2, §5.3] and is found to be

$$p(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_m|\boldsymbol{Z}_m) = c\,p_0(\boldsymbol{x}_0)\prod_{k=1}^{m} p_{w_k}(\boldsymbol{x}_k - \boldsymbol{f}_{k-1}(\boldsymbol{x}_{k-1}))\prod_{k=1}^{m} p_{v_k}(\boldsymbol{z}_k - \boldsymbol{h}_k(\boldsymbol{x}_k)), \tag{11}$$

where c is a normalizing constant, and p_0 is the prior PDF of the state $\boldsymbol{x}_0$ at time t_0. Further, in (11), the p_{w_k} and p_{v_k}, for $k = 1,\ldots,m$, are the respective PDFs of the process and measurement noise processes. In practice, they are often assumed to be Gaussian with zero mean and covariances of $\mathbf{Q}_k$ and $\mathbf{R}_k$, respectively.

The posterior PDF (11) *is* the complete description of the uncertainty of the state at each of the measurement times. In practice, a finite dimensional representation of the uncertainty is sought. Thus, the emphasis of the batch estimation problem is on how statistical information can be extracted from (11) consistently and accurately. Nonlinear optimization theory provides a framework for computing the *modal trajectory* or *maximum a posteriori (MAP)* estimate of (11). For a Gaussian prior with $\boldsymbol{x}_0 \sim N(\bar{\boldsymbol{x}}_0, \bar{\mathbf{P}}_0)$ and Gaussian noise processes as described above, the modal trajectory is obtained by solving the least squares or batch problem

$$(\hat{\boldsymbol{x}}_0,\ldots,\hat{\boldsymbol{x}}_m)^{MAP} = \arg\max_{\boldsymbol{x}_0,\ldots,\boldsymbol{x}_m} p(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_m|\boldsymbol{Z}_m)$$

$$= \arg\min_{\boldsymbol{x}_0,\ldots,\boldsymbol{x}_m} \frac{1}{2}\|\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0\|_{\bar{\mathbf{P}}_0^{-1}}^2 + \frac{1}{2}\sum_{k=1}^{m}\|\boldsymbol{x}_k - \boldsymbol{f}_{k-1}(\boldsymbol{x}_{k-1})\|_{\mathbf{Q}_{k-1}^{-1}}^2 + \frac{1}{2}\sum_{k=1}^{m}\|\boldsymbol{z}_k - \boldsymbol{h}_k(\boldsymbol{x}_k)\|_{\mathbf{R}_k^{-1}}^2. \tag{12}$$

In initial orbit determination, we do not have a prior; the term $\frac{1}{2}\|\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0\|_{\bar{\mathbf{P}}_0^{-1}}^2$ is removed from the above formulation.

Methods for solving nonlinear least squares problems, such as Gauss-Newton, full Newton, and quasi-Newton updates, along with globalization methods such as line search and trust region methods including Levenberg-Marquardt [10], are efficient and mature and will not be discussed further here. In the astrodynamics community, such solution techniques are called *differential correction* methods. In any nonlinear least squares problem such as (12), one must provide the solver a starting guess in order to initiate the differential correction method. This is the *initial orbit determination (IOD)* problem. In the case of measurement data from a single radar or EO sensor, a first estimate can be obtained using the classical methods of Lambert or Gauss (see, for example, [11]). Additionally, for angle-only observations, a recent algorithm due to Gooding [12] has shown promise for IOD scenarios involving both ground-based and space-based EO sensors [13].

Seen as a hybrid between batch estimation and the sequential prediction-correction filter discussed earlier, the sliding window batch estimation filter (SWBEF) processes n of the m measurements (through a batch process) over a sliding window. The SWBEF is the full batch estimation algorithm when the sliding window encloses the entire measurement sequence ($n = m$) and is the traditional predict-correct sequential filter when $n = 1$.

The use of a SWBEF can be very effective in addressing the problems of anomaly detection and UCT resolution because of the need to propagate states and uncertainties over long time gaps and to precisely evaluate likelihood ratios used in the association problem. In particular, in the problem of orbit determi-

Prediction–Correction Filter: $\cdots \rightarrow \boxed{p(\boldsymbol{x}_{k-1}|\boldsymbol{Z}_{k-1})} \rightarrow \boxed{p(\boldsymbol{x}_k|\boldsymbol{Z}_{k-1})} \rightarrow \boxed{p(\boldsymbol{x}_k|\boldsymbol{Z}_k)} \rightarrow \boxed{p(\boldsymbol{x}_{k+1}|\boldsymbol{Z}_k)} \rightarrow \boxed{p(\boldsymbol{x}_{k+1}|\boldsymbol{Z}_{k+1})} \rightarrow \cdots$

Sliding Window Batch Estimation Filter: $\cdots \rightarrow \boxed{p(\boldsymbol{x}_{k-1}|\boldsymbol{Z}_{k-1})} \longrightarrow \boxed{p(\boldsymbol{x}_k|\boldsymbol{Z}_k)} \longrightarrow \boxed{p(\boldsymbol{x}_{k+1}|\boldsymbol{Z}_{k+1})} \rightarrow \cdots$

Batch Estimation: $\boxed{p(\boldsymbol{x}_0)} \longrightarrow \boxed{p(\boldsymbol{x}_m|\boldsymbol{Z}_m)}$
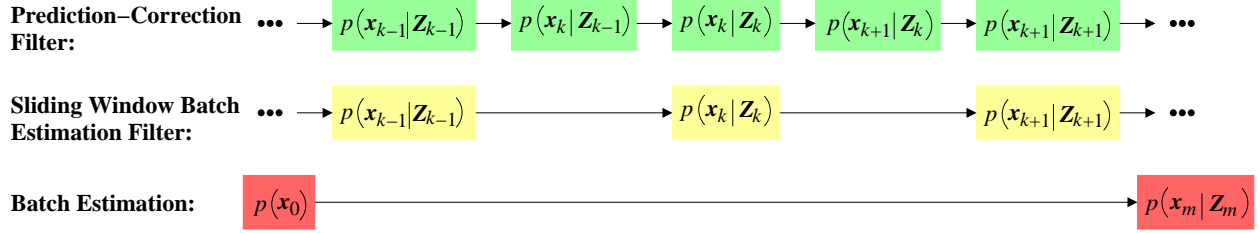
Figure 6: Schematics for the sequential prediction-correction filter, sliding window batch estimation filter (of length $n = 1$), and full batch estimation.

nation in the low Earth orbit (LEO) regime (e.g., a two minute track generated from radar measurements every ten seconds apart), it is observed that the posterior PDF $p(\boldsymbol{x}_k|\boldsymbol{Z}_k)$ is often well approximated by a single Gaussian in equinoctial orbital elements [14] or Cartesian ECI coordinates [15]. However, the predicted PDF $p(\boldsymbol{x}_{k+1}|\boldsymbol{Z}_k)$ can become highly non-Gaussian if the propagation time is long or the uncertainty along the radial direction (semi-major axis) of the posterior is large. One can expend tremendous computational resources in accurately representing the predicted PDF (using, for example, a high fidelity GSF or STTF). Yet, when the predicted PDF $p(\boldsymbol{x}_{k+1}|\boldsymbol{Z}_k)$ is updated with a new report $\boldsymbol{z}_{k+1}$, the resulting density $p(\boldsymbol{x}_{k+1}|\boldsymbol{Z}_{k+1})$ tends to collapse to something which is nearly Gaussian. The SWBEF bypasses the expensive representation of highly non-Gaussian densities arising from the prediction step and instead waits until a new report (measurement, track) becomes available and then performs the prediction and correction step *simultaneously*. Further details on the SWBEF are included in the paper [16].

1.1.1.4 Status A number of papers have been submitted to journals and conference proceedings on uncertainty modeling and management including ones on Gaussian sums [5, 17–19], the Edgeworth filter [19], the sliding window batch estimation filter [16], and the batch estimation problem [15]. Future work on this objective will aim to mature the Gaussian sum filter and the underlying optimization problem. Other alternate nonlinear filtering algorithms will also be pursued including the state transition tensor filter.

1.1.2 Multiple Hypothesis Tracking

Data association or correlation methods for multiple target tracking divide into two broad classes: single frame methods and multiple frame methods. The single frame methods include nearest neighbor, global nearest neighbor based on a two-dimensional assignment solver, and joint probabilistic data association (JPDA) [20]. For many widely spaced objects and a clear background, nearest neighbor may be appropriate, especially for space surveillance. For noisy backgrounds, JPDA can mitigate misassociation by updating the track with a weighted combination of all the measurements within a gate of the projected track state. The most successful of the multiple frame methods are multiple hypothesis tracking (MHT) [21] and multiple frame assignment (MFA) [3, 22]. MFA is an optimization-based formulation of MHT and is now considered to be the industry standard replacement of MHT. MHT/MFA methods mitigate misassociation by the ability to hold difficult correlation decisions in abeyance until additional information is available, as well as provide an opportunity to change past decisions to improve current decisions. In dense tracking environments such as breakups or geoclusters with closely spaced objects, the performance improvements of multiple frame methods over single frame methods are significant, making MHT/MFA the preferred solution.

The challenges of space surveillance, such as large coverage gaps and both sparse and dense target scenarios, require investigation of major extensions to current MHT/MFA architectures, including bias estimation and ambiguity assessment. A major new necessity is joint bias estimation and correlation, which may be needed for overlapping coverage, but is mandatory in handover situations, especially in large coverage gaps. We now discuss these requirements and present some MHT results obtained through this effort.

1.1.2.1   A Brief Review of MHT/MFA   Multiple hypothesis tracking normally uses a moving window over a set of data frames. The concept of a "data frame" is central to most modern data association or correlation algorithms. For initial orbit determination, the data frames correspond to sets of sensor reports in which each target is seen at most once in each frame. For track maintenance or extension, one data frame corresponds to a set of system tracks, while the remaining frames are composed of sensor reports.

Let $\mathcal{Z}_k^j = \{z_k^{i_k^j}\}_{i_k^j=1}^{M_k^j}$ denote a sequence[†] of noise-contaminated measurements of length M_k^j from sensor j at time t_k. Note that as the sensor index j and time index k are varied the number of measurements M_k^j can vary due to field of view issues, false alarms, missed detections, etc.[‡] We now define $\mathcal{A}$ as the list of time-sensor index pairs (k, j) such that $\mathcal{Z}_k^j$ is non-empty. The cardinality of $\mathcal{A}$ is N. Let $\mathcal{Z}^N = \{\mathcal{Z}_k^j\}_{\mathcal{A}}$ denote the data from these N sensor scans, potentially covering multiple times as well as multiple sensors. The central data association problem in multitarget and multisensor data fusion can be generally posed as [3]

$$H^* = \arg \max_{H \in \mathcal{H}} L_H(N), \quad L_H(N) = \frac{\Pr(H|\mathcal{Z}^N)}{\Pr(H_0|\mathcal{Z}^N)}, \tag{13}$$

where H denotes a partition of the data into tracks and false alarms, H_0 denotes a reference partition in which all reports are declared to be false alarms, $\mathcal{H}$ denotes the set of all feasible data partitions of the data $\mathcal{Z}^N$, and H^* denotes the optimal partition. Typically, one replaces the N frame likelihood ratio in (13) with a negative log-likelihood score

$$c_H(N) \equiv -\ln L_H(N), \tag{14}$$

which changes (13) to a minimization problem.

1.1.2.2   The Multidimensional Assignment Problem   A derivation of the problem and the equivalence with MHT using an N-scan back approximation has been previously given by Poore [3] and will not be repeated here. The following is a brief summary of the formulation, but not the derivation. Consider a layered graph $\mathcal{G}$ with N distinct node sets (layers) $A_k = \{0, 1, \ldots, m_k\}$ indexed by $k \in K = \{1, \ldots, N\}$ and arcs $A \subset A_1 \times \cdots \times A_N$. (In tracking applications, a node set A_k is called a frame of data.) A specific arc in A is denoted by the N-dimensional multi-index $a = (a(1), a(2), \ldots, a(N))$ where each $a(k) \in A_k$. Next, define a section $A_{k\ell} = \{a \in A \,|\, a(k) = \ell\}$. For each $k \in K$, associate to each section $A_{k\ell}$ a nonnegative integer $n_{k\ell} \geq 1$ for $\ell = 1, \ldots, m_k$ that will denote the number of times the index (k, ℓ), i.e., can be assigned.

In addition, x_a is assumed to be a zero-one variable for each $a \in A$ with at least two nonzero indices. The N-dimensional problem (multidimensional assignment problem of dimension N) can then be expressed as

$$\text{Minimize } \sum_{a \in A} c_a x_a$$
$$\text{Subject To } \sum_{a \in A_{k\ell}} x_a \leq n_{k\ell} \text{ for } k \in K \text{ and } \ell \in A_k \setminus \{0\}, \tag{15}$$
$$x_a \in \{0, 1\},$$

where each $n_{k\ell}$ is the number of times a report $\ell = 1, \ldots, m_k$ on frame k can be assigned. In most tracking applications, $n_{k\ell} = 1$; however, as suggested by real orbital data provided by JFCC SPACE, an object may be seen more than once in a pass over a sensor and thus there is a need to allow multiassignment. This is a *new feature* of the association problem for space surveillance. The cost c_a is the aforementioned negative

---

[†]No distinction is made between objects being observed as the association (which is being solved for) is not known a priori.
[‡]In space tracking false alarms are rare, but missed detections are quite common.

log of a likelihood ratio represented above as the ratio of two probabilities with the normalizing probability that each tracklet[§] is an uncorrelated track (or false alarm in measurement based tracking).

While the association problem can be formulated as a zero-one linear programming problem, the assignment formulation has been shown to be at least two orders of magnitude faster on larger problem sets. We propose the use of four classes of algorithms: (a) branch and bound based on Lagrangian relaxation to produce a guaranteed optimal, (b) a Lagrangian relaxation algorithm for a heuristic, (c) a partial branch and bound to improve the Lagrangian heuristic, and (d) an anytime algorithm that can control memory usage and runtime. For (a) and (c) we make use of A*-search to compute the k-best solutions for the multidimensional assignment problem. This provides the basis for *ambiguity assessment*.

1.1.2.3 Treatment of Biases

A key challenge in the correlation and fusion of multiple sensors is that of dealing effectively with sensor biases and navigation errors, for which we also use the term "biases." All fusion systems that integrate multiplatform multisensor data require that sensor registration be performed to remove biases before correlation and fusion. Otherwise, the data from each sensor will be misaligned, and performance will degrade. While "biases" are stochastic in nature, they represent systematic errors that do not average out and lead to misassociation and redundant tracks in a multisensor environment. Bias treatment can generally be divided into two cases: (i) association is known or some form of truth is available and (ii) association is unknown and no truth is available.

In the first case, the issue of association has already been separated out and the bias problem can be isolated. This is the case when the association between reports and objects is known, as might be the case with widely spaced targets. This is also the case when truth (or "fuzzy" or "approximate" truth, which might include some small noise) is available. An advantage of the space surveillance tracking environment is the availability of approximate truth data for a small collection of objects with well known trajectories. These can be used to estimate the biases for the sensors in the network, and weekly bias averages are often used to debias the following week's data. This calibration and debiasing of the sensor reports and navigation errors is performed using nonlinear least squares techniques, which can also reveal the observability [23, 24]. The drawback of this approach is that the sensor biases are not constant, and weekly updates can be insufficient for sensors with a larger bias drift. Furthermore, even if this debiasing were perfect, it still does not address the issue of *residual biases*, which is the difference between the true sensor bias and the estimated bias which was used to debias the sensors. Residual biases must be treated by a consider analysis or by the Schmidt-Kalman filter [2, 25]. Even the new multi-billion dollar sensors used in air and ground tracking (e.g., SBX, FBX, TPY-2) are subject to these issues, so the older sensors that make up a large portion of the space surveillance network will necessarily exhibit them as well.

In the second case, the association is unknown (and truth is unavailable) and both the association and bias must be solved for *jointly*. This is a much more difficult problem, and it is generally treated with what is known as "pattern matching" [26–28]. Unlike the first case, which relies on truth or known associations, pattern matching can be performed *online* using only readily available data. Because of this the joint bias and association problem is the key bias issue needed for long term propagation, and the remainder of this subsection will focus on it.

The problem of association and fusion in the presence of biases is illustrated in Figure 7 (with full mathematical details in [26–28]). Both figures show estimates of object positions in two-dimensional space with circles representing the covariance estimates. The large covariances correspond to previously-established tracks that have been predicted into the field of view of the next sensor, which has tracks established with the smaller covariances[¶]. Consider first the situation of Figure 7(a), where bias has not been accounted for. Due to the overlap of the "three-sigma" error circles, four tracks from each sensor will be incorrectly assigned to each other, while three tracks from each sensor remain *uncorrelated*. Compare this with the situation of

---

[§]A tracklet is a short track segment composed from about 10-12 sensor measurements.

[¶]This example is for the track-to-track problem, though the ideas are equally applicable for measurement-to-track.

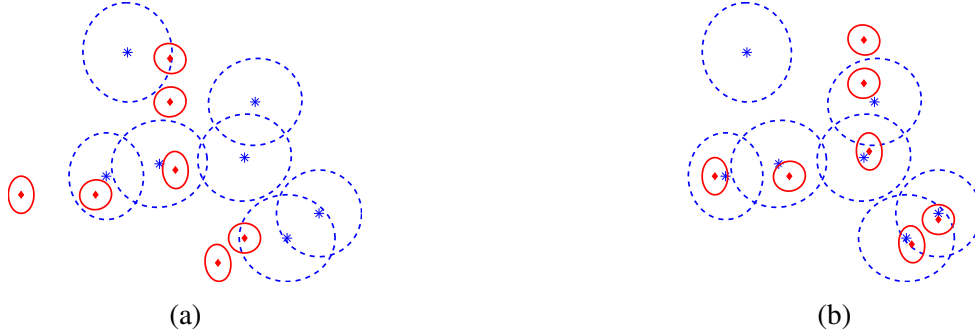(a)   (b)

Figure 7: (a) The effect of bias on data association. (b) Data association after bias correction.

Figure 7(b), where the concurrent bias estimation and association described later have been used to remove residual sensor biases so that the estimates now line up correctly. In this case, most estimates emanating from common objects are correctly correlated, with two estimates that do not correlate.

The *bias problem for handover, track-to-track association (correlation), and fusion* described above has been identified as a major obstacle for missile defense and manifests itself even for the newest MDA sensors; thus, it (along with the analogous measurement-to-track problem) will certainty present difficulties for the UCT problem in space surveillance.

We conclude this subsection with a brief description of the proposals for resolving the problem of joint association and bias estimation. The problem addressed in this work is that of associating two classes of objects in the presence of a bias which is modeled as a displacement between the two classes. Here it is assumed that each object can be assigned at most once (this can be generalized). The first collection of objects is enumerated by $i = 1, \ldots, m$, and the second by $j = 1, \ldots, n$. The formulation of this association problem is that of the two-dimensional inequality-constrained assignment problem, but with an unknown displacement between the objects. Inequality constraints are used since not all objects need be assigned.

$$\text{Minimize}_{(x,d)} \; d^T R^{-1} d + \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}(d) x_{ij}$$

$$\text{Subject To: } \sum_{j=1}^{n} x_{ij} \leq 1 \;\; (i = 1, ..., m), \qquad (16)$$

$$\sum_{i=1}^{m} x_{ij} \leq 1 \;\; (j = 1, ..., n),$$

$$x_{ij} \in \{0, 1\},$$

where $d = (d_x, d_y)$ is the vector of prior estimates of the sensor biases and R is the correlation matrix for these estimates. The correlation costs take the form [21, 28]

$$c_{ij}(d) = \frac{1}{2}(x_i + d_x - (y_j + d_y))^T (P_{ii} + Q_{jj} - S_{ij} - S_{ji})^{-1}(x_i + d_x - (y_j + d_y))$$

$$+ \frac{1}{2} \ln\left(\det(P_{ii} + Q_{jj} - S_{ij} - S_{ji})\right) - \gamma_{ij},$$

where P_{ii} and Q_{jj} are the covariances of the track state estimates x_i and y_j, respectively, while S_{ij} denotes the cross-correlation between x_i and y_j, which may arise from common process noise or common priors in the sensor estimation filters.

As posed, the problem (16) falls within a class of nonconvex mixed-integer nonlinear programming problems (MINLP) [29–32]. The techniques for solving convex MINLP problems include outer approxi-

mation, generalized Bender's decomposition, extended cutting plane, branch and bound, and branch and cut methods. For the nonconvex case, one attempts to extend these methods by using convex envelopes or by developing convex under estimators of the objective function and constraints. In addition, sampling methods such as clustering methods, evolutionary algorithms, simulated annealing, and tabu search may also be utilized.

For branch and bound, the lower bound that one must compute at each node in the branch and bound tree must be a global lower bound. To achieve this, one often relaxes the discrete variables to continuous ones and develops convex underestimators for the objective functions and constraints that facilitate the computation of this global lower bound. An example of this is the α-branch and bound method of Adjiman, Androulakis, and Floudas [33], which can be applied to the current problem.

The branch and bound framework developed herein does not follow these approaches for MINLP problems. The constraints are already convex in that they are affine. The objective function is nonconvex; however, we can develop a convex (affine) lower approximation to the objective function by using

$$b_{ij} = \frac{1}{2}\ln(\det(P_{ii} + Q_{jj} - S_{ij} - S_{ji})) - \gamma_{ij} \quad \text{for } i \neq 0 \text{ and } j \neq 0, \text{ and} \qquad (17)$$
$$b_{ij} = c_{ij} \text{ for } i = 0 \text{ or } j = 0,$$

both of which are independent of the displacement d. Furthermore, $b_{ij} \leq c_{ij}(d)$, thereby providing a basis for the global lower bound needed in the branch and bound procedure. We can make effective use of an assignment solver so there is no need for relaxation of the discrete variables.

For space surveillance, a multidimensional ($N > 2$) version of the joint bias estimation and correlation problem must be formulated and likelihood ratios derived. Such algorithms would be based on Lagrangian relaxation and an anytime branch and bound algorithm similar to that developed for the two-dimensional assignment problem with appropriate upper and lower bounds on optimality.

1.1.2.4 Uncertainty in Correlation

The goal of this algorithm is the development of an approximation to the probability of correct association. This has been achieved for the sequential k-best approximation to MHT based either on a k-best approach [34] or Markov Chain Monte Carlo (MCMC) methods [35], but not yet for the full multidimensional assignment problem.

In the k-best approach, an MHT system directly maintains a set $\mathcal{H}_k$ of k "complete" data association hypotheses $H_i \in \mathcal{H}_k$, $i = 1, \ldots, k$ within a sliding window of sensor data frames to approximate the optimal solution that maximizes the data association problem. Each alternative solution represents a different complete data association hypothesis H_i. Let $\mathcal{H}_k = \{H_1, H_2, \ldots, H_k\}$ denote a set of k *ranked* solutions, such that H_1 corresponds to the "best" solution returned from the assignment solver, H_2 to the "second-best", etc.; that is,

$$c_{H_1}(N) \leq c_{H_2}(N) \leq \ldots \leq c_{H_k}(N),$$

where the costs are defined as negative log-likelihood terms. As long as k is sufficiently large, it is the case that

$$\Pr(H_0|\mathcal{Z}^N) + \sum_{i=1}^{k} \Pr(H_i|\mathcal{Z}^N) \approx 1,$$

and therefore

$$\Pr(H_i|\mathcal{Z}^N) = \frac{\Pr(H_i|\mathcal{Z}^N)}{1} \approx \frac{\Pr(H_i|\mathcal{Z}^N)}{\Pr(H_0|\mathcal{Z}^N) + \sum_{j=1}^{k}\Pr(H_j|\mathcal{Z}^N)}.$$

Using (13) and (14), we obtain [21, 36]

$$\Pr(H_i) \approx \frac{\mathrm{e}^{-c_{H_i}(N)}}{1 + \sum_{j=1}^{k} \mathrm{e}^{-c_{H_j}(N)}},$$

which gives the *probability of a hypothesis in terms of its cost* c_{H_i}.

Now, let χ_{ij} denote the *probabilistic data association event* that measurement x_i associates with y_j. Then, the probability $\Pr(\chi_{ij})$ of this event is given by

$$\Pr(\chi_{ij}) \approx \sum_{H \in \mathcal{H}_{ij}} \Pr(H | \mathcal{Z}^N),$$

where $\mathcal{H}_{ij} \subseteq \mathcal{H}_k$ denotes the subset of those data association hypotheses from the set $\mathcal{H}_k$ that postulate the event χ_{ij}.

1.1.2.5 Results   Though there is still much work to be done in developing an MHT for space surveillance, some initial investigations have been conducted under this effort. Here we will present results from two breakup scenarios, which are described in Table 2. For each scenario we will compare MHT results using different nonlinear filters including (a) a single Gaussian (UKF), (b) a course Gaussian sum, (c) a fine Gaussian sum, and (d) the non-statistical method employed in the AFSPC Astrodynamic Standards ROTAS program. The number of terms in the Gaussian sum is given by N, with $N = 1$ indicating a single Gaussian, and $N = 1^\dagger$ indicating the non-statistical (ROTAS) method.

Table 2: Description of two breakup scenarios.

	Scenario One	Scenario Two
Regime	LEO	LEO
Bias, drag, maneuvers	Not taken into account	
No. of tracklets	27	42
Duration	6.5 hours	24 hours
No. of sensors	4	1

Figure 8 compares MHT results between pairs of methods for the two scenarios. The axes are indices into the tracklets and each pixel tells whether the corresponding pair of tracklets are in the same track in both methods (black), neither method (white), method one only (red), or method two only (green). Note that black and white pixels indicate agreement between the methods, while red and green pixels indicate disagreement. Table 4 supplements the information in the figures with the metrics defined in Table 3.

Table 3: Definitions of the track and summary metrics.

Track Metric T_1	Number of tracks in the first method
Track Metric T_2	Number of tracks in the second method
Track Metric T_b	Number of tracks that appear in both methods with the same associated tracklets
Summary Metric S	Percentage of tracklet pairs for which both methods agree on whether or not they emanate from a common object

The left three plots in Figure 8 and the left half of Table 4 show the results for Scenario One. For this scenario, the single Gaussian (UKF), coarse Gaussian sum ($N = 17$), and fine Gaussian sum ($N = 56$)
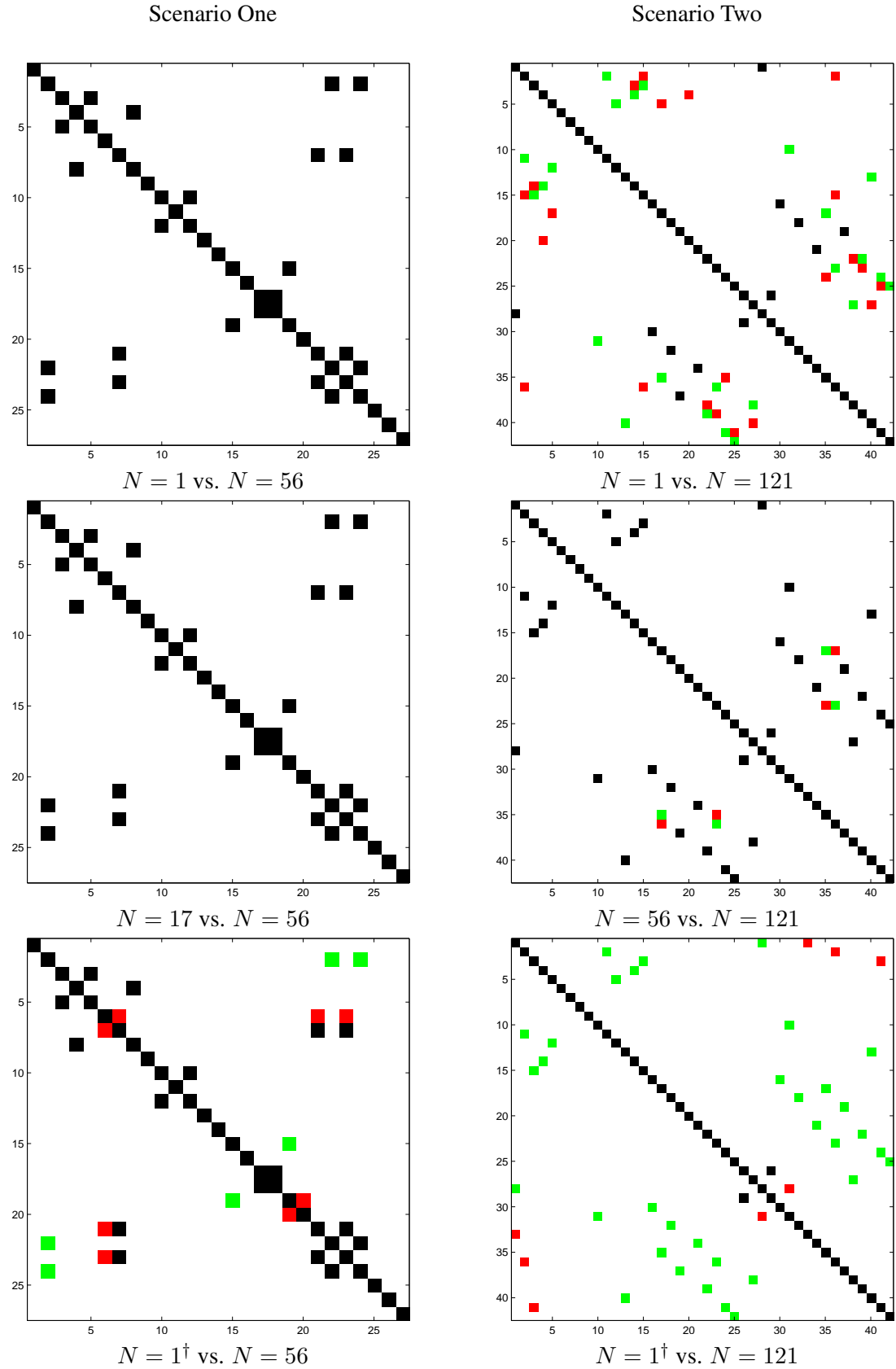
Scenario One

Scenario Two

Figure 8: Plots comparing MHT results using different nonlinear filters.

Table 4: Track and summary metrics (defined in Table 3) for Scenario One (left) and Scenario Two (right)

N_1	N_2	(T_1, T_2, T_b)	S		N_1	N_2	(T_1, T_2, T_b)	S
1	56	(18,18,18)	1.000000		1	121	(26,24,11)	0.973923
17	56	(18,18,18)	1.000000		56	121	(24,24,22)	0.995465
$1^\dagger$	56	(18,18,13)	0.980796		$1^\dagger$	121	(37,24,6)	0.976190

all give the same answers. Specifically, all three methods produce the exact same 18 tracks. From this we can conclude that in breakup scenarios with short time intervals between updates, propagating a state and covariance using the UKF is robust and sufficient – the higher order representations have validated this conclusion. In contrast, the non-statistical approach also produces 18 tracks, but only 13 of those tracks are the same as the tracks produced with the statistical methods. This indicates that even over a short scenario the non-statistical approach is insufficient and need not produce the correct solution.

The right three plots in Figure 8 and the right half of Table 4 show the results for Scenario Two. Here we see a reasonable amount of disagreement between the single Gaussian (UKF) and fine Gaussian sum ($N = 121$), but almost exact agreement between the coarse Gaussian sum ($N = 56$) and the fine Gaussian sum. This "convergence" as the fidelity of the filter increases provides evidence that advanced filtering techniques such as adaptive Gaussian sum filters can be used to produce the correct solutions. It is interesting to look a bit closer at the form of the discrepancy between the two Gaussian sum filters. The coarse Gaussian sum associates tracklets (17, 36) and (23, 35), while the fine Gaussian sum switches these pairs to (17, 35) and (23, 36) – the only difference in results between these two methods is a single association decision. Moving to the non-statistical approach, we see a stark contrast. The non-statistical approach produced 37 tracks, indicating that it had trouble associating objects together (we can also see this by the large number of green pixels in the figure). In particular, it made very few associations between tracklets that were separated by long time intervals – the statistical methods did not have this problem.

Finally, we mention that the results presented here only focus on the *best* hypothesis from each method. In cases where the best hypothesis has less than 100% of the total likelihood, the ambiguity must also be considered in order to get "the full picture." Even in these relatively small scenarios there was significant ambiguity, though for simplicity the ambiguity is not included in the presented results.

1.1.2.6 Status The small breakup scenarios in LEO discussed above were presented at the AFOSR review meeting on Sept. 11, 2010 and at the DARPA meeting on UCTs at Washington DC on Nov. 17, 2010. Extensions to GEO and the joint association and bias estimation problem are ongoing.

1.1.3 Unresolved Closely Spaced Objects

Much of this section discusses algorithm components that have been developed under a Phase II SBIR, "Hierarchical Image Processing for Closely Spaced Object (CSO) Resolution". The core research problem addressed in this part of the project is that of augmenting measurement generation algorithms to supply high-quality measurement covariances and estimates of the uncertainty in object count in a small region of space, called *resolution uncertainty*, for use in generating accurate orbit state covariances and data association ambiguity information.

The innovative part of our approach is the use of multiframe ideas in two manners. First, we propose to use multiple images from the sensor to generate high-quality individual images using image superresolution techniques. Second, given a set of images, either directly from the sensor or preprocessed using the above techniques, we will use multiple-frame, multiple-hypothesis expectation maximization (EM)-based measurement generation to solve the problem of targets appearing in some frames while not appearing in

others, the so-called *stochastic resolution problem.* A combination of such methods promises to substantially improve the resolution and tracking of closely spaced objects.

1.1.3.1 Background on Image Processing The field of image processing for astronomy and space surveillance has involved a significant amount of research focused on the problem of improving image quality prior to extracting target measurement data. Single-frame image restoration methods involve various techniques intended to reduce the noise in the image, reduce the blurring that results from optical diffraction and atmospheric turbulence, or both. In general, noise reduction and blur compensation are at odds; most methods that reduce blur will tend to amplify noise, and vice versa. These techniques include basic spatial filtering [37], deconvolution approaches [38–40], variations on the CLEAN algorithm [41], and newer techniques based on wavelet transform filtering [42].

All of the above methods for image restoration are intended to function using a single input frame of data and, usually, some form of prior knowledge (such as the image noise statistics and the point spread function) to generate a restored scene. One of the core innovations in this work is to broaden the view of the problem to include multiframe methods, in which data from more than one collected image may be used to generate a result. This additional data permits both resolution enhancement and consideration of measurement stability in the presence of unresolved closely spaced objects. Multiframe image restoration methods in the literature range from shift-and-add formulations with anti-aliasing through hybrid approaches such as the Drizzle algorithm [43] all the way up through to image superresolution techniques.

Most of the image restoration algorithms discussed in the literature are pure image transforms; that is, they take as input one or more images and produce a higher-quality output image. This alone is insufficient for the purposes of tracking; we must convert the raw image data into a sequence of measurements that can be correlated and fused to produce orbit state estimates. This process can be challenging during events such as on-orbit collisions, where potentially large numbers of objects may reside within the field of view of a sensor. Prior work by Numerica on tracking from infrared sensor data [44] has shown that multiple-frame, multiple-hypothesis expectation maximization (EM)-based measurement generation has the ability to solve the stochastic resolution problem by considering information from several consecutive frames of data to determine the correct measurement count in the scene. Such algorithms stabilize the observed object scene in the presence of potentially significant numbers of closely spaced objects and are discussed below.

1.1.3.2 Proposed Implementation Consider the case of an imaging sensor viewing one or more objects as point sources (that is, the sensor resolution is insufficient to observe the objects in more than one pixel). The observed image $\mathbf{Y}$ is a function of object position and magnitude, plus noise:

$$\mathbf{Y} = \sum_{i=1}^{K} \mathbf{G}_i(\theta_i) + \mathbf{E},$$

where θ_i is a vector $[\mu_{ui}, \mu_{vi}, I_i]'$ containing the centroid focal plane coordinates u, v and intensity magnitude I for object i. Further, $\mathbf{E}$ is the additive noise present in the image, and $\mathbf{G}_i$ is the (nonlinear) imaging function. We would like to find an estimate $\hat{\theta}_i$ for each object i that maximizes the likelihood of the observed image. It may be observed that the structure of this problem is similar to that of the multiframe image restoration problem presented by Elad and Feuer [45]; this analogy is intentional. We will discuss the combined image restoration and measurement extraction problem below.

In the special case where the imaging function can be approximated by a Gaussian blur function resulting from sensor limitations and atmospheric turbulence, we can be more precise in the formulation of the above equation. We collapse the 2D image data into a column vector for ease of manipulation. For a single

measured pixel y_j, the relationship between the point sources and the pixel intensity is given by

$$y_j = \sum_{i=1}^{K} I_i \int_{vmin_j}^{vmax_j} \int_{umin_j}^{umax_j} \frac{1}{2\pi\sigma_u\sigma_v\sqrt{1-\rho^2}} e^{-\frac{1}{2}\frac{1}{1-\rho^2}\left(\frac{(u-\mu_{ui})^2}{\sigma_u^2}+\frac{(v-\mu_{vi})^2}{\sigma_v^2}-\frac{2\rho(u-\mu_{ui})(v-\mu_{vi})}{\sigma_u\sigma_v}\right)} dudv + E_j$$

for an arbitrary Gaussian blur function $\mathbf{B}$. If, as is typical, the blur function is spherical (that is, $\mathbf{B} = \sigma^2\mathbf{I}$), then $\rho = 0$ in the above equation. This form is separable, and therefore y_j may be expressed as a closed-form solution in terms of the error function for computational efficiency.

We now have a model for the intensity of each pixel in the image that we can use to formulate a nonlinear estimation problem for the point source locations and intensities. This nonlinear estimation problem, however, tends to be very poorly conditioned and have local minima. Good starting guesses for the point source locations are required if the estimator is to converge to the correct solution. Also, it is clear that this model requires some prior estimate of the number K of point sources to be sought in the image. In some cases this estimate will be available from the sensor tasking (if the objects in the field of view are well-characterized); however, this is not typically the case for objects resulting from new space events. The key to solving this problem lies in the *multiple hypothesis* approach. We have constructed a range of alternate approaches to generating high-probability prior object counts and locations. Then, given these priors, we can construct several different measurement hypotheses for the image.

We have observed that the most effective method for generating a measurement prior, if the input datarate supports use of the technique, is that of first constructing a synthetic high-resolution image from a sequence of low-resolution images (multiple frame image restoration). In this case, the formulation of the problem is now

$$\mathbf{Y} = \mathbf{H}(\mathbf{X}) + \mathbf{E},$$

$$\mathbf{X} = \sum_{i=1}^{K} \mathbf{G}_i \theta_i$$

with a linear transform $\mathbf{H}$ representing blur, warp, and decimation functions between a theoretical high-resolution image $\mathbf{X}$ and the measured low-resolution images. We accumulate sufficient measured images to solve the sparse linear estimation problem for the high-resolution image, and then use standard peak extraction techniques to obtain a prior for the nonlinear parameter estimation problem described above.

While the use of multiframe image restoration for astronomy applications is not novel, our unique approach combines the image restoration problem with source position estimation for tracking in the presence of unresolved closely spaced objects. Figure 9 shows the multiple frame image restoration used in conjunction with model-based measurement generation to achieve maximum point source resolving power. Panel (a) shows a raw measured image (single-band data in false color) of a set of point sources at locations marked with white '+' symbols, along with measurements and covariances (blue '*' and ellipses) generated from a simple cluster and centroid approach. Panel (b) focuses on a small subset of the image containing several unresolved point sources. Panel (c) is a synthetic high-resolution image of the same scene generated from a sequence of images such as (a), with measurements constructed from the nonlinear optimization problem described above. The advanced measurement generation algorithms have the ability to resolve point sources with much greater efficacy than would be possible from a single-frame approach.

Using multiple consecutive frames of data, we consider whether the evolution of measurements over time produces a consistent track picture. In order to solve this problem within an orbit determination context, we require two additional items of information: the covariance associated with each of the estimated object locations, and the uncertainty of the *object resolution hypothesis*. The measurement covariance is a function of the signal-to-noise ratio for the object being observed. If there are multiple objects within a confusable distance (overlapping blur regions) of one another, then the error associated with the centroid estimate is
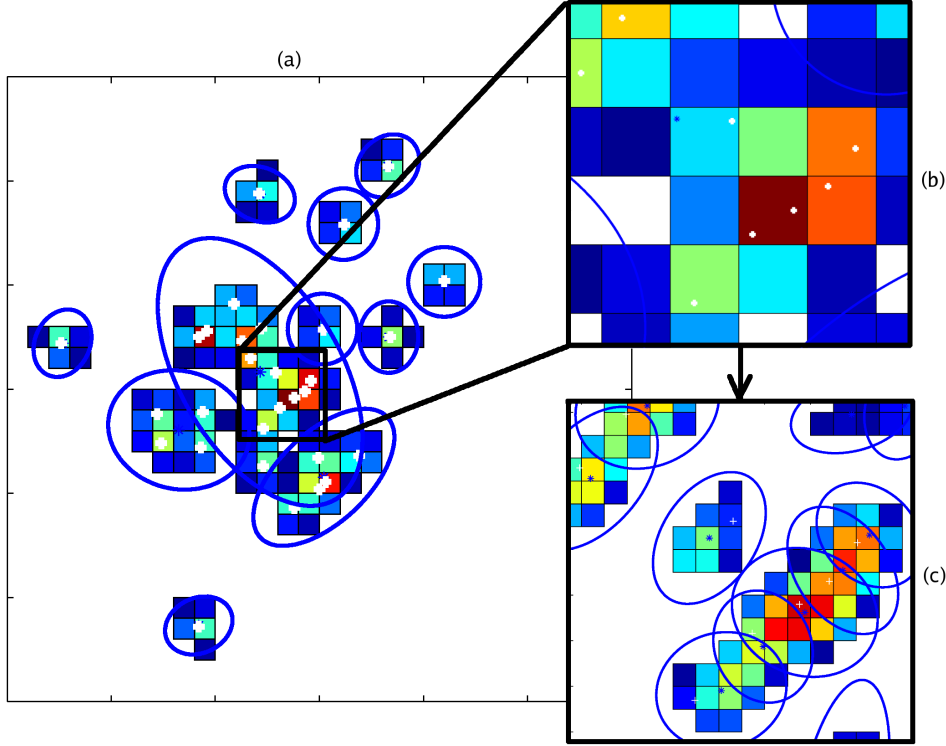
Figure 9: Measurement generation from sequence of linear and nonlinear inverse problems.

also a function of the distance to, and brightness of, the nearby objects. Failure to account for these effects can result in an incorrect measurement covariance estimate. We present here the development for object resolution uncertainty given two objects within a confusion distance from one another; development of the more general form is anticipated as part of this research program. Suppose we have:

$$\mathbf{Y} = \mathbf{G}(\theta) + \mathbf{E} \sim N(\mathbf{G}(\theta), \sigma_n^2\mathbf{I});$$

that is, the intensity noise in the image is Gaussian and characterized by noise power σ_n^2.

Let us consider the hypotheses that $\theta_0 = \mathbf{0}$ (the data consists of noise only), $\theta_1 = [\theta_{obj1}; 0]$ (the data consists of object 1 only), $\theta_2 = [0; \theta_{obj2}]$ (the data consists of object 2 only), and $\theta_3 = [\theta_{obj1}; \theta_{obj2}]$ (both objects must be present to explain the data). The generalized likelihood of each of these hypotheses can then be expressed as

$$\hat{f}_i(y) = (2\pi)^{-n/2}\sigma_n^{-n}\exp\left[-\frac{1}{2\sigma_n^2}(\mathbf{Y} - \mathbf{G}(\hat{\theta}_i))^T(\mathbf{Y} - \mathbf{G}(\hat{\theta}_i))\right].$$

These likelihoods allow us to resolve the hypotheses described above. Critical for the resolution question is the ratio of the likelihood that the data is explained by one object versus two. If the algorithm has hypothesized the presence of two objects, but it is more likely that the data can be explained by only one object, then we either (a) really do only have one object, or (b) have two objects that cannot be distinctly resolved. These likelihood ratios then form the basis for incorporating measurement hypothesis probabilities into the tracking problem.

1.1.3.3 Conclusion We have discussed two innovative approaches that combine to address tracking problems for closely spaced objects, both using multiframe ideas, but in two subtly different ways. First, we discussed using multiple images from the sensor to generate high-quality individual images using image

superresolution techniques. Second, given a set of images, either directly from the sensor or preprocessed using the above techniques, we showed how to use multiple-frame, multiple-hypothesis expectation maximization (EM)-based measurement generation to solve the problem of targets appearing in some frames while not appearing in others, giving rise to a possible solution to both stochastic resolution and resolution uncertainty problems. We have constructed a prototype of the combined multiframe approach; future work will focus on use of this formulation to construct measurement covariances and resolution uncertainties for tracking.

1.1.3.4 Status This work is ongoing. The processing of data in GEO has not yet started.

1.2 Accomplishments / New Findings

The primary objective of this program is to perform the necessary basic research to support the development of a statistical, multiple hypothesis tracker (MHT) for space surveillance. Such a MHT framework can serve as the next generation space surveillance system to maintain the space catalog, to identify uncorrelated tracks, and to support conjunction analysis and sensor resource management. Key components in such a system include a consistent characterization of uncertainty, physical modeling, multiple model filtering, and the association problem of determining which tracklets/measurements emanate from which object. To achieve a consistent characterization of uncertainty, Numerica has developed an adaptive Gaussian sum filter which correctly represents and propagates uncertainties and adaptively selects the correct the number of Gaussians in the mixture. Realtime online metrics support the coarsening and refining of the filter to maintain consistent uncertainty. A sliding window batch estimation filter has also been developed and shown to provide an accurate evaluation of the prediction error critical for correct anomaly detection and resolution of UCTs. Several papers have been published on these uncertainty management and propagation. Numerica has also acquired orbital data from JFCC SPACE and processing of this data has commenced using a new prototype MHT specifically suited to space surveillance in conjunction with the aforementioned techniques on uncertainty management.

2 Personnel Supported

a. PI: Aubrey B. Poore

b. Colleagues at Numerica: Joshua T. Horwood, Nathan D. Aragon, and Scott Danford

3 Publications

Journal papers submitted for publication under this effort are cited as References [5, 17] in the bibliography. Conference proceedings papers produced under this effort are cited as References [15, 16, 18, 19] in the bibliography.

4 Participation / Presentations at Meetings, Conferences, Seminars, etc.

1. SPIE Defense, Security, and Sensing 2010: Signal and Data Processing of Small Targets. Orlando, FL. Apr. 5–9, 2010. Presented paper [15].

2. George H. Born Symposium. Boulder, CO. May 13-14, 2010.

3. Kyle T. Alfriend Astrodynamics Symposium. Monterey, CA. May 17–19, 2010. Presented paper [18].

4. Advanced Maui Optical and Space Surveillance Technologies Conference 2010. Wailea, HI. Sept. 15–17, 2010. Presented paper [19].

5. DARPA UCT meeting. Washington, DC. Nov. 17, 2010.

6. 21st AAS/AIAA Space Flight Mechanics Meeting. New Orleans, LA. Feb. 14–17, 2011. Presented paper [16].

5 New Discoveries, Inventions, or Patents

No patents resulted from this effort.

References

[1] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Boston: Artech House, 2004.

[2] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Dover, 1970.

[3] A. B. Poore, "Multidimensional assignment formulation of data association problems arising from multitarget tracking and multisensor data fusion," *Computational Optimization and Applications*, vol. 3, pp. 27–57, 1994.

[4] D. Alspach and H. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Transactions on Automatic Control*, vol. 17, no. 4, pp. 439–448, 1972.

[5] J. T. Horwood, N. D. Aragon, and A. B. Poore, "Gaussian sum filters for space surveillance: theory and simulations," *Journal of Guidance, Control, and Dynamics*, 2011. (To Appear).

[6] R. van der Merwe and E. A. Wan, "The square-root unscented Kalman filter for state and parameter estimation," in *IEEE Int. Conf. Acoust., Speech, Signal Process (ICASSP)*, vol. 6, pp. 3461–3464, 2001.

[7] G. Terejanu, P. Singla, T. Singh, and P. D. Scott, "Uncertainty propagation for nonlinear dynamic systems using Gaussian mixture models," *Journal of Guidance, Control and Dynamics*, vol. 31, no. 6, pp. 1623–1633, 2008.

[8] K. DeMars, M. Jah, D. Giza, and T. Kelecy, "Orbit determination performance improvements for high area-to-mass ratio space object tracking using an adaptive Gaussian mixtures estimation algorithm," in *21st International Symposium on Space Flight Dynamics*, (Toulouse, France), 2009.

[9] D. R. Giza, P. Singla, J. L. Crassidis, R. Linares, P. J. Cefola, and K. Hill, "Entropy-based space object data association using an adaptive Gaussian sum filter," in *AIAA Guidance, Navigation, and Control Conference*, (Toronto, Canada), August 2010.

[10] R. Fletcher, *Practical methods of optimization*. Chichester: John Wiley & Sons, 1991.

[11] D. A. Vallado, *Fundamentals of Astrodynamics and Applications*. New York: Springer, 2007.

[12] R. H. Gooding, "A new procedure for the solution for the classical problem of minimal orbit determination from three lines of sight," *Celestial Mechanics and Dynamical Astronomy*, vol. 66, no. 1, pp. 387–423, 1997.

[13] D. A. Vallado, "Evaluating Gooding angles-only orbit determination of space based space surveillance measurements," in *Proceedings of the AAS George Born Astrodynamics Symposium*, (Boulder, CO), May 2010. (In Press).

[14] C. Sabol, T. Sukut, K. Hill, K. T. Alfriend, B. Wright, Y. Li, and P. Schumacher, "Linearized orbit covariance generation and propagation analysis via simple Monte Carlo simulations," in *Proceedings of the 2010 Space Flight Mechanics Conference*, (San Diego, CA), February 2010. AAS 10-134.

[15] J. T. Horwood, N. D. Aragon, and A. B. Poore, "Covariance consistency for track initiation using Gauss-Hermite quadrature," in *SPIE Proceedings: Signal and Data Processing of Small Targets 2010*, vol. 7698, 2010.

[16] J. T. Horwood, N. D. Aragon, and A. B. Poore, "Sliding window batch estimation filtering for enhanced anomaly detection and uncorrelated track resolution," in *Proceedings of the 21st AAS/AIAA Space Flight Mechanics Meeting*, (New Orleans, LA), February 2011. Paper AAS-11-152.

[17] J. T. Horwood and A. B. Poore, "Adaptive Gaussian sum filters for space surveillance," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1777–1790, 2011.

[18] J. T. Horwood, N. D. Aragon, and A. B. Poore, "Adaptive Gaussian sum filters for space surveillance tracking," in *Proceedings of the AAS Kyle T. Alfriend Astrodynamics Symposium*, (Monterey, CA), May 2010. Paper AAS-10-321.

[19] J. T. Horwood, N. D. Aragon, and A. B. Poore, "Edgeworth filters for space surveillance tracking," in *Proceedings of the 2010 Advanced Maui Optical and Space Surveillance Technologies Conference*, (Wailea, HI), September 2010.

[20] Y. Bar-Shalom and X.-R. Li, "Multitarget-multisensor tracking: Principles and techniques," Lecture Notes, Department of Electrical and Systems Engineering, University of Connecticut, 1995.

[21] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Boston: Artech House, 1999.

[22] A. B. Poore and A. J. Robertson, III, "A new class of Lagrangian relaxation based algorithms for a class of multidimensional assignment problems," *Computational Optimization and Applications*, vol. 8, p. 129, 1997.

[23] S. M. Herman and A. B. Poore, "Nonlinear least-squares estimation for sensor and navigation biases," in *SPIE Proceedings: Signal and Data Processing of Small Targets 2006*, vol. 6236, 2006.

[24] B. D. Kragel, S. Danford, S. M. Herman, and A. B. Poore, "Bias estimation using targets of opportunity," in *SPIE Proceedings: Signal and Data Processing of Small Targets 2007*, vol. 6699, 2007.

[25] R. Paffenroth, R. Novoselov, S. Danford, M. Teixeira, S. Chan, and A. Poore, "Mitigation of biases using the Schmidt-Kalman filter," in *SPIE Proceedings: Signal and Data Processing of Small Targets 2007*, vol. 6699, 2007.

[26] S. Danford, B. Kragel, and A. Poore, "Joint MAP bias estimation and data association: algorithms," in *SPIE Proceedings: Signal and Data Processing of Small Targets 2007*, vol. 6699, 2007.

[27] S. Danford, B. Kragel, and A. Poore, "Joint MAP bias estimation and data association: simulations," in *SPIE Proceedings: Signal and Data Processing of Small Targets 2007*, vol. 6699, 2007.

[28] M. Levedahl, "An explicit pattern matching assignment algorithm," in *SPIE Proceedings: Signal and Data Processing of Small Targets 2002*, vol. 4728, 2002.

[29] B. Borchers and J. E. Mitchell, "A computational comparison of branch and bound and outer approximation algorithms for 0-1 mixed integer nonlinear programs," *Computers and Operations Research*, vol. 24, no. 8, pp. 699–701, 1997.

[30] C. A. Floudas, *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*. New York: Oxford University Press, 1995.

[31] I. E. Grossman, "Review of nonlinear mixed-integer and disjunctive programming techniques," *Optimization and Engineering*, vol. 3, pp. 227–252, 2002.

[32] M. Tawarmalani and N. V. Sahinidis, *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming*, vol. 65 of *Nonconvex Optimization and Its Applications*. Kluwer Academic Publishers, 2002.

[33] C. S. Adjiman, I. P. Androulakis, and C. A. Floudas, "Global optimization of mixed-integer nonlinear problems," *AIChE Journal*, vol. 46, no. 9, pp. 1769–1797, 2000.

[34] S. Gadaleta, S. Herman, S. Miller, F. Obermeyer, B. J. Slocumb, and A. B. Poore, "Short-term ambiguity assessment to augment tracking data association information," in *Eighth International Conference on Information Fusion*, (Philadelphia, PA), 2005.

[35] S. Oh, L. Schenato, and S. Sastry, "A hierarchical multiple-target tracking algorithm for sensor networks," in *International Conference on Robotics and Automation*, (Barcelona, Spain), April 2005.

[36] R. Popp, K. Pattipati, and Y. Bar-Shalom, "m-best S-D assignment algorithm with application to multitarget tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37, pp. 22–39, 2001.

[37] J. Russ, *The Image Processing Handbook*. CRC Press, 2002.

[38] J. L. Starck, E. Pantin, and F. Murtagh, "Deconvolution in astronomy: A review," *Publications of the Astronomical Society of the Pacific*, vol. 114, pp. 1051–1069, 2002.

[39] W. Waniak, "Image restoration by simple adaptive deconvolution," *Astronomy and Astrophysics Supplement*, vol. 124, pp. 197–203, 1997.

[40] S. Gull and J. Skilling, "Maximum entropy method in image processing," *IEEE Proc.*, vol. 131-F, pp. 646–659, 1984.

[41] U. J. Schwarz, "Mathematical-statistical description of the iterative beam removing technique (method CLEAN).," *Astronomy and Astrophysics*, vol. 65, pp. 345–356, 1978.

[42] D. L. Donoho, "Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition," 1995.

[43] R. N. Hook and A. S. Fruchter, "Variable-pixel linear combination," *Astronomical Data Analysis Software and Systems VI*, vol. 125, pp. 147–150, 1997.

[44] S. Gadaleta, A. Poore, and B. Slocumb, "Pixel-cluster decomposition tracking for multiple IR-sensor surveillance," in *SPIE Vol. 5204, Signal and Data Processing of Small Targets*, 2003.

[45] M. Elad and A. Feuer, "Super-resolution reconstruction of image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 9, pp. 817–834, 1999.