Analysis of Costs for the GNP Problem

MARK LEVEDAHL
JOHN D. GLASS

Track-to-track data association in a multisensor framework involves score functions to determine a solution. When sensor errors include both random noise and unknown bias terms, several options are available. Of these, two options are the global nearest pattern match (GNPM) and marginal track-to-track association (MTTA) scores. The former involves a joint likelihood of bias and association hypothesis and the latter is the result of integrating the total probability space over the unknown bias to remove the bias likelihood. Analytically, we show that the difference between these scores is the determinant of the a-posteriori bias covariance, and that the same bias estimation is inherent in both. Using a simple numerical example, we compare the weight each score formulation apportions to track assignment hypotheses based on the quality of the bias estimate, and show that GNPM tends to favor hypotheses with low a-posteriori bias covariance. Additionally, through evaluation of the incremental cost structure, we argue that the non-assignment cost used in both scores is nearly optimal, in the sense of correct associations, for GNPM. However, the same non-assignment cost is not optimal for the MTTA score, and the significance depends upon the uncertainty of bias and the number of associations made.

Manuscript received November 23, 2020; revised May 19, 2021; released for publication June 30, 2021.

Dr. Ramona Georgescu handled refereeing of this contribution.

Mark Levedahl and John D. Glass are with Raytheon Technologies, Waltham, MA 02451, USA (e-mail: {mark.d.levedahl, john.d.glass}@rtx.com).

I. INTRODUCTION

ASSOCIATING sets of observations from sensor systems is fundamental in multi-sensor tracking. With reliable multi-sensor track assignment, track fusion can achieve improved accuracy and allow handover of data from one sensor to another [4]. Furthermore, distributed sensor systems allow coverage of larger areas with different viewing angles and facilitate the formation of a complete track picture [6]. Basic complications that prohibit perfect track-to-track association are unknown residual bias errors, random errors contained in the observations of a sensor, unknown true target motion, and heterogeneous sensor coverage. Residual bias may arise from imperfections in sensor registration, transformation errors, and other sources, whereas random errors arise from stochastic effects of sensor systems such as thermal noise. Missed detections are often the result of sensor sensitivity/phenomenology and other aspects such as sensor resolution, thus causing heterogeneous sensor coverage. Unknown target motion may also induce error in the estimated track state, regardless of other errors, yielding cross-correlated error across sensors. Mathematical models of these sensor errors form the foundation of modern track-to-track association algorithms.

Track association in a multisensor framework involves score functions to assess alternate association hypotheses. Any hypothesized association of tracks implies a set of observed targets and locations, with the score function providing the probability the given tracks arise from common targets specified by the association hypothesis. These score functions in general have unknown, possibly random, parameters (e.g., location of targets) implicitly set such that the score is maximized at the observed values [12], [13], and thus evaluate how well each hypothesis fits the data. The classic formulation is termed the global nearest neighbor (GNN) problem and addresses random errors with heterogeneous sensor coverage, but assumes independent errors per track, ignoring bias errors [5], [6], [16]. Since the assignment score of a track tuple is independent of others, GNN is an N-D assignment problem with costs in the form of negative log-likelihoods based on assumed statistical models. Solving the two-sensor case is very efficient with solvers such as the auction or Jonker–Volgenant–Castanon algorithms. For a survey on solution methods to GNN see [22]. To handle heterogeneous sensor coverage, GNN algorithms include the cost of particular tracks not assigning, based on a-priori assumptions by which targets may appear. Uniform spatial distribution of targets in the surveillance volume with the total number of targets as Poisson distributed are standard assumptions in the literature.

When multiple sensors track the same target, errors of those tracks can become cross-correlated, assuming these errors arise due to common process noise. The basic ideas are in [2], including discussion of handling more than two sensors. As shown in [2], the scores of track tuples remain independent with inclusion of cross-correlated errors of this form, maintaining the ability to use GNN solution methods. A more difficult problem arises when there are cross-correlated errors across a set of tracks from a sensor, herein called "bias" errors, though these errors need not be time-invariant nor 100% cross-correlated. This problem is very important when the magnitude of the residual bias errors is significant compared to the sum of target spacing and independent errors in tracks. A simple mitigation with a GNN algorithm is to inflate covariances to cover both the random errors and the residual bias. However, as shown in [16], the method of covariance inflation gives poor association performance as the magnitude of bias grows. Approaches that are more sophisticated attempt to recognize the bias and provide specific treatment. Early techniques involved sequential methods that first attempt to estimate and remove the bias, then use GNN as if the tracks are unbiased [24]. More recent techniques jointly solve for residual bias and assignments within the mathematical formulation of the problem.

A full treatment of bias errors requires different scoring functions than used in GNN. The global nearest pattern match (GNPM) function, presented in [16] for the case of two sensors, includes the most probable bias per hypothesis in the score. A variant of this approach based upon marginalizing the bias estimate, termed marginalized track-to-track association (MTTA), is presented in [20], again for the case of two sensors. These scoring functions facilitate solutions to what we call the global nearest pattern (GNP) problem. Compared to early solutions to the GNP problem that focused on independent bias estimation and assignment steps, the novelty of GNPM and MTTA is in the explicit treatment of sensor bias in the scoring functions, leading to joint assignment and bias estimation [16], [20]. As shown in [16], this joint approach can provide significantly improved data association performance compared to GNN even when bias errors are a small fraction of the independent random errors. The work of [7] and [14] extends the GNPM function for the N-sensor case, including cross-correlation due to process noise.

The GNP assignment problem is much more difficult to solve than the related GNN problem as the costs are not separable into independent costs per track pair. Instead, GNP gives coupled costs based upon the hypothesis dependent bias estimate, breaking the assumption underlying use of standard assignment solvers for this problem. For problems with only a handful of hypotheses to choose from, a feasible solution is to enumerate and score all. However, in many real-world cases, this approach is infeasible. The 6×8 association problem we investigate in Section III gives 93,289 possible hypotheses, illustrating how even a handful of tracks give a high number of total hypotheses. Addressing this problem, Levedahl in [17] provides a Dijkstra shortest path technique for providing the K-best solutions to the GNP problem, applicable to both the GNPM and MTTA cost functions, and discuses performance (both runtime and accuracy) compared to GNN in [16]. Papageorgiou in [21] provides additional specialized mathematical programs for solving these problems, again including discussion of accuracy and runtime issues. The techniques discussed above have proved practical and useful in real time for problems much larger than the 6×8 problem included here.

It is worth noting that, in general, these techniques assume the targets within a single GNP problem have a common bias offset represented in the same dimension of the state space. Strictly speaking, the assumption of a common relative bias offset to the sets of data is seldom true in practice. For example, a registration bias is often modeled as additive constants in the measurement space of range and angle as in [19], which affects Cartesian tracks in a non-linear fashion. Thus, a small azimuth bias $\delta\theta$ affects position as range multiplied by $\delta\theta$, but also the velocity as the latter vector estimate has been rotated by $\delta\theta$. So long as the targets are not widely dispersed, an assumption of common bias is reasonable. We prefer to think of the common bias assumption as a linear approximation of a non-linear bias model about the centroid of the targets of interest. Conversely, a non-linear bias model of specific range and azimuth offsets for each sensor needs wide dispersion among targets to yield favorable observability [19], and requires non-linear estimation techniques. In addition, widely dispersed targets tend to unambiguous association problems where GNN covariance inflation approaches may suffice. GNP methods are appropriate where bias errors are significant compared to the noise error and inter-target spacing, and targets in the problem are not widely dispersed such that the common bias model is unreasonable. Regardless, the common bias representation is an approximation made by numerous authors, including [7], [8], [9], [14], [16], [17], [20], and [24], and is the focus of this work. We leave any extensions to non-linear bias models as future work, in part because such extensions preclude the closed-form solutions essential to the comparisons made in this paper.

The key objective of this paper is to understand the full mathematical foundation and relationship of GNPM and MTTA, along with sound mathematical rationale for selection between them. Therefore, we ignore any extension to greater than two sensors, and ignore any extension to general cross-correlated errors beyond sensor bias. Generally speaking, GNPM assumes that the most probable bias is a key variable in the association problem, while MTTA treats bias as a nuisance parameter and marginalizes bias in the score. Ferry in [8] and [9] also makes arguments based in Bayesian methodology in agreement with Papageorgiou's treatment of bias, but Ferry incorporates fundamentally different a-priori target assumptions, most importantly that targets appear spatially according to a Gaussian distribution rather than uniform as GNPM/MTTA assumes. A benefit of the Gaussian assumption is closed-form integrals rather than the approximations needed for the uniform case, but results in equations much more complex than GNPM/MTTA and are hard to decompose in a fashion that allow efficient solution. We note that the perceived value of bias marginalization by the authors of [20] is in ambiguity management, claiming that bias likelihood can be a corrupting presence in correctly determining the probabilities of various association hypotheses. Other authors have attributed the difficulty of reliable probability determination to the integral approximations inherent in the posterior [15], precisely the integral targeted in the work of Ferry in [8] and [9]. Although the work of [15] demonstrates that this integral approximation becomes less accurate in dense target scenarios where ambiguity management is critical, the role this integral plays was not discussed in [20]. We further note that the key metric used in this work, association accuracy, directly scores whether the highest probability hypothesis is most correct, and is a necessary but insufficient criterion to achieving correct hypothesis probabilities. For the problems investigated here, our findings show MTTA is sometimes worse, and never better, than GNPM in association accuracy.

We investigate GNPM/MTTA against the key criterion of maximizing the number of correct assignments, as that is the fundamental objective of data association. In Section II of this contribution, we start from the basic assumptions of the track association problem in the presence of bias to derive the exact difference between the GNPM and MTTA score functions. A part of this derivation includes expressing the GNPM and MTTA scores as a new, yet equivalent, stacked Gaussian density equation. We show that although MTTA has the bias term removed through integration, the same relative bias estimation of GNPM is implicit in the MTTA assignment score. Intuitively, we show that the difference between the two scores is the determinant of the covariance of the a-posteriori bias estimate, very similar to the conclusions made in [12] for marginalization of target locations. Leveraging this result, in Section III we elaborate on the practical differences between the GNPM and MTTA scores through analytic and numerical examples. Critically, we evaluate association performance for various non-assignment costs and show that the non-assignment cost often cited in the GNN/GNP literature is nearly optimal for GNPM in the sense of maximizing the probability of correct association. However, as the uncertainty of residual bias grows, this non-assignment cost can be far from optimal for MTTA. The covariance of the a-posteriori bias estimate in the MTTA cost is precisely the source of sub-optimal assignment performance, therefore any adjustments made to MTTA that maximize correct assignments give GNPM. Our findings show that GNPM can be much more accurate than MTTA for few assignments or large bias errors, and MTTA is never more accurate than GNPM. Therefore, we recommend the use of GNPM for the problems described here. We provide concluding remarks in Section IV.

II. GNP SCORES AND COSTS

In this section, we start from a basic mathematical description of the track association problem and derive the necessary total and conditional distributions required to reveal the relationship between the GNPM and MTTA assignment scores. For the nomenclature used in this paper, $\mathcal{N}(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$ denotes a multivariate normal distribution of mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. All vectors are assumed column vectors denoted in lowercase bold, and matrices are uppercase bold. To reduce nomenclature complexity, we use $\mathbf{0}$ to denote either the zero matrix or the zero vector, which is obvious in the context of usage.

A. Observation Model

Assume n_t targets denoted as $\mathbf{x}_t,\ t = 1, .., n_t$, observed by sensors $\mathcal{A}$ and $\mathcal{B}$, each observing a potentially different subset of targets. Assume sensor $\mathcal{A}$ develops m distinct observations and $\mathcal{B}$ develops n distinct observations with no false or redundant observations from either sensor. Without loss of generality, assume the observations satisfy $m \le n \le n_t$. Sensor $\mathcal{A}$ observations are corrupted by zero-mean random noise with covariance $\mathbf{S}_{\mathcal{A},i}$, uncorrelated for each observation. Therefore the observations from $\mathcal{A}$ take the form

$$\mathbf{x}_i^{\mathcal{A}} = \mathbf{x}_{\alpha_i} + \mathbf{n}_i^{\mathcal{A}}, \tag{1}$$

$$p\left(\mathbf{n}_i^{\mathcal{A}}\right) = \mathcal{N}\left(\mathbf{0},\ \mathbf{S}_{\mathcal{A},i}\right), \tag{2}$$

where α_i is an unknown index to the target tracked. Observations from sensor $\mathcal{B}$ follow a similar model but with errors specific to that sensor including an unknown relative bias term $\mathbf{b}$. Therefore

$$\mathbf{x}_j^{\mathcal{B}} = \mathbf{x}_{\eta_j} - \mathbf{b} + \mathbf{n}_j^{\mathcal{B}}, \tag{3}$$

$$p\left(\mathbf{n}_j^{\mathcal{B}}\right) = \mathcal{N}\left(\mathbf{0},\ \mathbf{S}_{\mathcal{B},j}\right), \tag{4}$$

where η_j is an unknown index to the target tracked for sensor $\mathcal{B}$. The single relative bias term (relative to the coordinate frame of sensor $\mathcal{A}$) is assumed common to all observations from sensor $\mathcal{B}$ and has the probability distribution

$$p\left(\mathbf{b}\right) = \mathcal{N}\left(\mathbf{0},\ \mathbf{R}\right). \tag{5}$$

The covariances $\mathbf{S}_{\mathcal{A},i}$, $\mathbf{S}_{\mathcal{B},j}$, and $\mathbf{R}$ are all assumed to be symmetric positive-definite matrices, and the dimension of all sensor observations is assumed to be of dimension d.

The goal of track-to-track assignment is to determine the underlying truth commonality in the observations. Truth commonality is represented as the i and j indexes such that $\alpha_i = \eta_j$. Since the actual ordering of the targets is arbitrary and unknown, we equivalently seek the assignment of tracks from sensor $\mathcal{A}$ to sensor $\mathcal{B}$. Define the assignment vector as $\mathbf{h} = [h_1 \ldots h_m]^T$ of length m

where the i^{th} element indicates the index in $\mathcal{B}$ that is assigned to the i^{th} observation in $\mathcal{A}$. Unassigned observations in $\mathcal{A}$ are indicated with an h_i of zero. Therefore, let $\mathcal{J} = \{i : h_i \neq 0\}$ be the set of assigned track indexes and $n_a = |\mathcal{J}|$ be the number of assignments in $\mathbf{h}$. Assuming a uniform prior on each $\mathbf{x}_i$ location and that the number of targets in the surveillance volume is Poisson distributed, following the derivation in [14], the posterior probability of an assignment hypothesis and bias can be written as[1]

$$\begin{aligned} &\Pr\left(\mathbf{h},\ \mathbf{b}|\mathbf{x}_1^{\mathcal{A}}, .., \mathbf{x}_m^{\mathcal{A}},\ \mathbf{x}_1^{\mathcal{B}}, .., \mathbf{x}_n^{\mathcal{B}}\right) \\ &= C\frac{1}{\sqrt{|2\pi\mathbf{R}|}}e^{-\frac{1}{2}\mathbf{b}^T\mathbf{R}^{-1}\mathbf{b}} \\ &\times \left(\beta P_{\mathcal{A}\bar{\mathcal{B}}}\right)^{n-n_a}\left(\beta P_{\bar{\mathcal{A}}\mathcal{B}}\right)^{m-n_a}\left(\beta P_{\mathcal{A}\mathcal{B}}\right)^{n_a} \\ &\times \prod_{i\in\mathcal{J}}\frac{1}{\sqrt{|2\pi\mathbf{S}_i|}}e^{-\frac{1}{2}\left(\mathbf{x}_i^{\Delta}-\mathbf{b}\right)^T\mathbf{S}_i^{-1}\left(\mathbf{x}_i^{\Delta}-\mathbf{b}\right)}, \end{aligned} \tag{6}$$

with the difference terms and associated covariances expressed as

$$\begin{aligned} \mathbf{x}_i^{\Delta} &= \mathbf{x}_i^{\mathcal{A}} - \mathbf{x}_{h_i}^{\mathcal{B}}, \\ \mathbf{S}_i &= \mathbf{S}_{\mathcal{A},i} + \mathbf{S}_{\mathcal{B},h_i}, \end{aligned} \tag{7}$$

for all $i \in \mathcal{J}$. The β term is the spatial density of the targets, $P_{\mathcal{A}\mathcal{B}}$ is the probability that both sensor $\mathcal{A}$ and $\mathcal{B}$ observe a target, $P_{\mathcal{A}\bar{\mathcal{B}}}$ is the probability that sensor $\mathcal{A}$ but not $\mathcal{B}$ observe a target, $P_{\bar{\mathcal{A}}\mathcal{B}}$ is the probability that sensor $\mathcal{B}$ but not $\mathcal{A}$ observe a target, and C is a normalizing constant.[2] Of significance in (6) is the sufficient statistic of an assignment hypothesis as the absolute difference between the track states, $\mathbf{x}_i^{\Delta}$. As noted in Corollary 1 of [14], incorporation of cross-correlated errors between $\mathbf{x}_i^{\mathcal{A}}$ and $\mathbf{x}_{h_i}^{\mathcal{B}}$ due to common process noise involves a simple subtraction term to $\mathbf{S}_i$, which can be easily inserted into (7). We choose to leave that term omitted since we have not studied the effects of common process noise in our numerical simulations, but anticipate no impact upon the conclusions reached. As will be discussed in upcoming sections, GNPM is the joint posterior of (6), while MTTA requires the additional step of marginalizing $\mathbf{b}$.

B. Probability Distributions of Bias and Errors

Any joint probability density has an equivalence with marginal and conditional densities. Block forms of the random vectors described in the observation model allow the use of the fundamental equations of linear estimation [3] to give marginal and conditional densities. Defining $\boldsymbol{\gamma} = [\gamma_1, \ \ldots, \gamma_{n_a}]$ to be a length n_a vector that contains an ordering of the indices in $\mathcal{J}$, the stacked vector of absolute differences of assigned tracks from (7) as

$$\mathbf{x}_{\Delta} = \begin{bmatrix} \mathbf{x}_{\gamma_1}^{\Delta} \\ \vdots \\ \mathbf{x}_{\gamma_{n_a}}^{\Delta} \end{bmatrix}, \tag{8}$$

and the block identity matrix as

$$\mathbf{H} = \begin{bmatrix} \mathbf{I} \\ \vdots \\ \mathbf{I} \end{bmatrix}, \tag{9}$$

with n_a blocks of $d \times d$ identity matrices, the following marginal and conditional distributions are derived in Appendix A:

$$p\left(\mathbf{x}_{\Delta}\right) = \mathcal{N}\left(\mathbf{0},\ \mathbf{Q}_{\mathbf{x}_{\Delta}}\right), \tag{10}$$

$$p(\mathbf{x}_{\Delta}|\mathbf{b}) = \mathcal{N}\left(\boldsymbol{\mu}_{\mathbf{x}_{\Delta}|\mathbf{b}},\ \mathbf{Q}_{\mathbf{x}_{\Delta}|\mathbf{b}}\right), \tag{11}$$

$$p\left(\mathbf{b}|\mathbf{x}_{\Delta}\right) = \mathcal{N}\left(\boldsymbol{\mu}_{\mathbf{b}|\mathbf{x}_{\Delta}},\ \mathbf{Q}_{\mathbf{b}|\mathbf{x}_{\Delta}}\right), \tag{12}$$

with the corresponding elements as

$$\mathbf{Q}_{\mathbf{x}_{\Delta}} = \mathbf{Q}_{\mathbf{x}_{\Delta}|\mathbf{b}} + \mathbf{H}\mathbf{R}\mathbf{H}^T, \tag{13}$$

$$\boldsymbol{\mu}_{\mathbf{x}_{\Delta}|\mathbf{b}} = \mathbf{H}\mathbf{b}, \tag{14}$$

$$\mathbf{Q}_{\mathbf{x}_{\Delta}|\mathbf{b}} = \begin{bmatrix} \mathbf{S}_{\gamma_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S}_{\gamma_{n_a}} \end{bmatrix}, \tag{15}$$

$$\boldsymbol{\mu}_{\mathbf{b}|\mathbf{x}_{\Delta}} = \mathbf{P}_{\mathbf{b}\mathbf{x}_{\Delta}}^T\mathbf{Q}_{\mathbf{x}_{\Delta}}^{-1}\mathbf{x}_{\Delta} \tag{16}$$

$$\mathbf{Q}_{\mathbf{b}|\mathbf{x}_{\Delta}} = \mathbf{R} - \mathbf{P}_{\mathbf{b}\mathbf{x}_{\Delta}}^T\mathbf{Q}_{\mathbf{x}_{\Delta}}^{-1}\mathbf{P}_{\mathbf{b}\mathbf{x}_{\Delta}}, \tag{17}$$

$$\mathbf{P}_{\mathbf{b}\mathbf{x}_{\Delta}} = \mathbf{H}\mathbf{R}. \tag{18}$$

Each of these probability densities relate to the likelihood of a track assignment hypothesis and bias. Upon conversion into the block structure, (11) is the final term in (6), therefore

$$\begin{aligned} &\Pr\left(\mathbf{h},\ \mathbf{b}|\mathbf{x}_1^{\mathcal{A}}, ..\mathbf{x}_m^{\mathcal{A}},\ \mathbf{x}_1^{\mathcal{B}}, ..\mathbf{x}_n^{\mathcal{B}}\right) \\ &= C\left(\beta P_{\mathcal{A}\bar{\mathcal{B}}}\right)^{n-n_a}\left(\beta P_{\bar{\mathcal{A}}\mathcal{B}}\right)^{m-n_a}\left(\beta P_{\mathcal{A}\mathcal{B}}\right)^{n_a}p\left(\mathbf{x}_{\Delta}|\mathbf{b}\right)p\left(\mathbf{b}\right)^{.} \end{aligned} \tag{19}$$

It is worth noting that $p(\mathbf{x}_{\Delta}|\mathbf{b})$ in (19) is the likelihood function of the bias and hypothesis given the data, although the conditioning term only mentions $\mathbf{b}$. By inspection of (7), the $\mathbf{x}_{\Delta}$ notation depends on the hypothesis, and therefore we do not add $\mathbf{h}$ as a conditioning term. We use the $p(\mathbf{x}_{\Delta}|\mathbf{b})$ notation to identify

[1] The authors in [14] generalized to more than two sensors, with a separate bias term per sensor instead of a single relative bias.

[2] A slight distinction with the derivation in [14] is the detection probabilities as hypothesized in $\mathbf{h}$, which are conditioned on the event that at least one sensor detected the target (i.e., undetected targets do not enter the assignment problem). Some authors have also made this distinction explicit as in [11] or [18]. We also note that C scales all hypothesis scores equally so is not needed for finding the best hypothesis, and in general is not determined as doing so may require enumerating all possible assignment hypotheses.

that the likelihood is a function of the differences of assigned tracks in a particular hypothesis, in addition to simplicity compared to the more formal, yet equivalent, $p(\mathbf{x}_1^{\mathcal{A}}, ..\mathbf{x}_m^{\mathcal{A}},\ \mathbf{x}_1^{\mathcal{B}}, ..\mathbf{x}_n^{\mathcal{B}}|\mathbf{h},\ \mathbf{b})$. Careful readers may also notice that (19) is not written explicitly as a function of $p(\mathbf{h})$. However, following the derivation in [14] $p(\mathbf{h})$ is part of the product of scalar terms in (19). More specifically, the derivation of (6) in [14] involves conditioning $\mathbf{h}$ to the abstract α_i and η_j indices from (1) and (3), respectively, along with the unknown number of targets, which upon simplification gives $p(\mathbf{h})$ as being a contributor to the term $C(\beta P_{\mathcal{A}\bar{\mathcal{B}}})^{n-n_a}(\beta P_{\bar{\mathcal{A}}\mathcal{B}})^{m-n_a}(\beta P_{\mathcal{A}\mathcal{B}})^{n_a}$. Rigorous technical details of the posterior density derivation exist in previous literature, including [9] and [14]. We also note that many authors use the word "likelihood" liberally when referring to posteriors and related terms, sometimes by admission as the authors of MTTA in [20]. In this work, we prefer to maintain more strict terminology usage, particularly with the use of the word likelihood as a specific contribution to the posterior. Furthermore, we define the product of likelihood and bias prior, $p(\mathbf{x}_\Delta|\mathbf{b})p(\mathbf{b})$, as the kinematic score.

The GNPM and MTTA scores differ only in kinematic terms, which are those depending upon $\mathbf{x}_\Delta$ or $\mathbf{b}$. These terms reveal the relationship of the GNPM/MTTA scores using $p(\mathbf{x}_\Delta)$ decomposed through Bayes law:

$$p\left(\mathbf{x}_\Delta\right) = \frac{p\left(\mathbf{x}_\Delta|\mathbf{b}\right)\ p\left(\mathbf{b}\right)}{p\left(\mathbf{b}|\mathbf{x}_\Delta\right)}, \tag{20}$$

which is valid for any realization of $\mathbf{b}$.

C. GNPM and $p(\mathbf{x}_\Delta)$ Equivalence

In this section, we provide the relationship between the GNPM score of [16] and the distribution of the total errors. First, with algebraic manipulations (16) can be expressed as[3]

$$\begin{aligned}
\mu_{\mathbf{b}|\mathbf{x}_\Delta} &= \mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}^T \mathbf{Q}_{\mathbf{x}_\Delta}^{-1}\mathbf{x}_\Delta \\
&= \mathbf{R}\mathbf{H}^T\left(\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}} + \mathbf{H}\mathbf{R}\mathbf{H}^T\right)^{-1}\mathbf{x}_\Delta \\
&= \mathbf{R}\mathbf{H}^T\left(\mathbf{I} + \mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{H}\mathbf{R}\mathbf{H}^T\right)^{-1}\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{x}_\Delta \\
&= \mathbf{R}\left(\mathbf{I} + \mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}\mathbf{H}\mathbf{R}\right)^{-1}\mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{x}_\Delta \\
&= \left(\mathbf{R}^{-1} + \mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{x}_\Delta.
\end{aligned} \tag{21}$$

Recognizing that $\mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{H} = \Sigma_{\mathrm{i}=1}^{\mathrm{m}}\ \mathbf{S}_i^{-1}$, removal of the block form in (21) reveals an equivalence to the $\bar{\mathbf{x}}$ from [16]

$$\begin{aligned}
\mu_{\mathbf{b}|\mathbf{x}_\Delta} &= \left(\mathbf{R}^{-1} + \mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{x}_\Delta \\
&= \left(\mathbf{R}^{-1} + \sum_{i\in\mathcal{J}}\mathbf{S}_i^{-1}\right)^{-1}\sum_{i\in\mathcal{J}}(\mathbf{S}_i^{-1}\mathbf{x}_i^\Delta) = \bar{\mathbf{x}},
\end{aligned} \tag{22}$$

which is the bias estimate that maximizes the kinematic score for a given assignment hypothesis. We subsequently refer to $\mu_{\mathbf{b}|\mathbf{x}_\Delta}$ as $\bar{\mathbf{x}}$, avoiding excessive use of subscripts and to clarify connections to previous literature. By inspection of (6) and (8) in [16], nomenclature translation allows the GNPM kinematic score to be written as

$$\begin{aligned}
K_{GNPM} &= \frac{1}{\sqrt{|2\pi\mathbf{R}|}}e^{-\frac{1}{2}\bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}}} \\
&\quad\times\prod_{i\in\mathcal{J}}\frac{1}{\sqrt{|2\pi\mathbf{S}_i|}}e^{-\frac{1}{2}\left(\mathbf{x}_i^\Delta-\bar{\mathbf{x}}\right)^T\mathbf{S}_i^{-1}\left(\mathbf{x}_i^\Delta-\bar{\mathbf{x}}\right)} \\
&= \frac{1}{\sqrt{|2\pi\mathbf{R}|}}e^{-\frac{1}{2}\bar{\mathbf{x}}^T\ \mathbf{R}^{-1}\bar{\mathbf{x}}} \\
&\quad\times\frac{1}{\sqrt{\left|2\pi\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}\right|}}e^{-\frac{1}{2}\left(\mathbf{x}_\Delta-\mathbf{H}\bar{\mathbf{x}}\right)^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\left(\mathbf{x}_\Delta-\mathbf{H}\bar{\mathbf{x}}\right)}.
\end{aligned} \tag{23}$$

Notice that the first term of (23) is (5) evaluated at $\mathbf{b} = \bar{\mathbf{x}}$ and the second term is (11), also evaluated at $\mathbf{b} = \bar{\mathbf{x}}$ by (14). Further observing from (22) and (12) that $p(\mathbf{b}|\mathbf{x}_\Delta)$ evaluated at $\mathbf{b} = \bar{\mathbf{x}}$ gives $1/\sqrt{|2\pi\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}|}$, the relationship between K_{GNPM} and $p(\mathbf{x}_\Delta)$ is

$$\begin{aligned}
p\left(\mathbf{x}_\Delta\right) &= \frac{p\left(\mathbf{x}_\Delta|\mathbf{b}\right)p\left(\mathbf{b}\right)}{p\left(\mathbf{b}|\mathbf{x}_\Delta\right)}\Big|_{\mathbf{b}=\bar{\mathbf{x}}} \\
&= \frac{K_{GNPM}}{p(\mathbf{b}|\mathbf{x}_\Delta)|_{\mathbf{b}=\bar{\mathbf{x}}}} = K_{GNPM}\sqrt{\left|2\pi\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}\right|}.
\end{aligned} \tag{24}$$

D. MTTA and $p(\mathbf{x}_\Delta)$ Equivalence

The derivation of MTTA in [20] began with GNPM, shown in the previous section to be $p(\mathbf{x}_\Delta|\mathbf{b})p(\mathbf{b})$, followed by integration of bias out of the score. Therefore, due to $\int p(\mathbf{x}_\Delta|\mathbf{b})p(\mathbf{b})d\mathbf{b} = p(\mathbf{x}_\Delta)$, we expect the MTTA likelihood to be equivalent to the distribution of the total errors. Here, we show the equivalence using the expansion of (20) about the point $\mathbf{b} = \mathbf{0}$. As a preliminary step, we rewrite (17) into an equivalent expression using the matrix inversion

[3]An algebraic step here uses the relationship $(\mathbf{I}+\mathbf{P}\mathbf{Q})^{-1}\mathbf{P} = \mathbf{P}(\mathbf{I}+\mathbf{Q}\mathbf{P})^{-1}$ from traditional literature on the matrix inversion lemma [10].

lemma and removal of the block form,

$$\begin{aligned}\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta} &= \mathbf{R} - \mathbf{P}^T_{\mathbf{b}\mathbf{x}_\Delta}\mathbf{Q}^{-1}_{\mathbf{x}_\Delta}\mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta} \\ &= \mathbf{R} - \mathbf{R}\mathbf{H}^T\left(\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}} + \mathbf{H}\mathbf{R}\mathbf{H}^T\right)^{-1}\mathbf{H}\mathbf{R} \\ &= \left(\mathbf{R}^{-1} + \mathbf{H}^T\mathbf{Q}^{-1}_{\mathbf{x}_\Delta|\mathbf{b}}\mathbf{H}\right)^{-1} \\ &= \left(\mathbf{R}^{-1} + \sum_{i\in\mathcal{J}}\mathbf{S}_i^{-1}\right)^{-1}. \end{aligned} \tag{25}$$

Upon nomenclature translation, the MTTA kinematic score as given for (8) in [20] is

$$K_{MTTA} = \frac{\sqrt{|2\pi\mathbf{V}|}}{\sqrt{\prod_{i\in\mathcal{J}^*}|2\pi\mathbf{S}_i|}}e^{-\frac{1}{2}\zeta}, \tag{26}$$

with

$$\mathbf{V} = \left(\sum_{i\in\mathcal{J}^*}\mathbf{S}_i^{-1}\right)^{-1}, \tag{27}$$

$$\zeta = \left(\sum_{i\in\mathcal{J}^*}\left(\mathbf{x}_i^\Delta\right)^T\mathbf{S}_i^{-1}\mathbf{x}_i^\Delta\right) - \mathbf{u}^T\mathbf{V}\mathbf{u}, \tag{28}$$

$$\mathbf{u} = \sum_{i\in\mathcal{J}^*}\mathbf{S}_i^{-1}\mathbf{x}_i^\Delta, \tag{29}$$

and the definitions $\mathbf{S}_0 = \mathbf{R}$, $\mathbf{x}_0^\Delta = \mathbf{0}$, and $\mathcal{J}^* = \{\mathcal{J}, 0\}$. From (22) and (25), we observe that $\mathbf{V} = \mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}$ and $\mathbf{u} = \mathbf{V}^{-1}\bar{\mathbf{x}}$, therefore

$$\begin{aligned}\mathbf{u}^T\mathbf{V}\mathbf{u} &= \bar{\mathbf{x}}^T\mathbf{V}^{-1}\mathbf{V}\mathbf{V}^{-1}\bar{\mathbf{x}} \\ &= \bar{\mathbf{x}}^T\mathbf{V}^{-1}\bar{\mathbf{x}} \\ &= \bar{\mathbf{x}}^T\mathbf{Q}^{-1}_{\mathbf{b}|\mathbf{x}_\Delta}\bar{\mathbf{x}}, \end{aligned} \tag{30}$$

and the full expansion of ζ can be rewritten as

$$\zeta = \left(\sum_{i\in\mathcal{J}}\left(\mathbf{x}_i^\Delta\right)^T\mathbf{S}_i^{-1}\mathbf{x}_i^\Delta\right) - \bar{\mathbf{x}}^T\mathbf{Q}^{-1}_{\mathbf{b}|\mathbf{x}_\Delta}\bar{\mathbf{x}}. \tag{31}$$

Substituting the expansions of $\mathbf{V}$, ζ, and rearranging terms in (26) to expose the specific Gaussian densities, we demonstrate the desired equivalency of MTTA and $p(\mathbf{x}_\mathbf{\Delta})$ following similar steps as in (23) and (24):

$$\begin{aligned}K_{MTTA} &= \frac{\sqrt{\left|2\pi\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}\right|}}{e^{-\frac{1}{2}\left(\bar{\mathbf{x}}^T\mathbf{Q}^{-1}_{\mathbf{b}|\mathbf{x}_\Delta}\bar{\mathbf{x}}\right)}} \\ &\quad\times\prod_{i\in\mathcal{J}}\frac{1}{\sqrt{|2\pi\mathbf{S}_i|}}e^{-\frac{1}{2}\left(\mathbf{x}_i^\Delta\right)^T\mathbf{S}_i^{-1}\mathbf{x}_i^\Delta} \\ &\quad\times\frac{1}{\sqrt{|2\pi\mathbf{R}|}} \\ &= \left(\frac{1}{p\left(\mathbf{b}|\mathbf{x}_\mathbf{\Delta}\right)}\times p\left(\mathbf{x}_\mathbf{\Delta}|\mathbf{b}\right)\times p\left(\mathbf{b}\right)\right)\Bigg|_{\mathbf{b}=\mathbf{0}} \\ &= p\left(\mathbf{x}_\mathbf{\Delta}\right). \end{aligned} \tag{32}$$

E. Remarks on GNP Assignment Scores

Combining (24) and (32) gives the key result relating the kinematic scores and the distribution of the absolute errors

$$K_{MTTA} = K_{GNPM}\sqrt{\left|2\pi\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}\right|} = p\left(\mathbf{x}_\mathbf{\Delta}\right). \tag{33}$$

Consequently, although the MTTA formulation integrated the bias from the GNPM kinematic score, it inherently uses the same bias estimate that maximizes the association hypothesis as in GNPM. This result for bias mirrors the conclusions found for marginalizing the unknown target locations by Kaplan in [12] and [13]. In both cases, the difference between using the maximum likelihood value versus marginalizing reduces to a factor of the a-posteriori covariance. Furthermore, (33) implies the additional insight that bias estimation does not need to be a separate step in the calculation of a GNP score due to the equivalence with (10). In other words, combining the terms raised to the exponent in (23) gives an expression equivalent to $\mathbf{x}_\Delta^T\mathbf{Q}^{-1}_{\mathbf{x}_\Delta}\mathbf{x}_\Delta$. To solidify this result, we show the following equivalence algebraically in Appendix B:

$$\mathbf{x}_\Delta^T\mathbf{Q}^{-1}_{\mathbf{x}_\Delta}\mathbf{x}_\Delta = \left(\mathbf{x}_\Delta - \mathbf{H}\bar{\mathbf{x}}\right)^T\mathbf{Q}^{-1}_{\mathbf{x}_\Delta|\mathbf{b}}\left(\mathbf{x}_\Delta - \mathbf{H}\bar{\mathbf{x}}\right) + \bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}} \tag{34}$$

which follows from the matrix inversion lemma along with several algebraic manipulations. A corollary of (34) is that $(\mathbf{x}_\Delta - \mathbf{H}\bar{\mathbf{x}})^T\mathbf{Q}^{-1}_{\mathbf{x}_\Delta|\mathbf{b}}(\mathbf{x}_\Delta - \mathbf{H}\bar{\mathbf{x}}) + \bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}}$ is a chi-square random variable with dn_a degrees of freedom, since $\mathbf{x}_\Delta^T\mathbf{Q}^{-1}_{\mathbf{x}_\Delta}\mathbf{x}_\Delta$ is a chi square random variable of dimension dn_a. This may not be immediately obvious at first glance, since with removal of the block form of the right hand side, (34) is the sum of $(n_a + 1)$ terms. In other words, degrees of freedom are lost through the estimation of $\bar{\mathbf{x}}$ with the data. An additional observation of (34) is that the left hand side is a function of $\mathbf{R}$, but the right hand side is a function of $\mathbf{R}^{-1}$. This allows various interpretations and simplifications if $\mathbf{R}$ is assumed arbitrarily large or arbitrarily small.

As an additional remark, equating the normalization terms inherent in (33) gives

$$\frac{\sqrt{\left|2\pi\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}\right|}}{\sqrt{|2\pi\mathbf{R}|}\prod_{i\in\mathcal{J}}\sqrt{|2\pi\mathbf{S}_i|}} = \frac{1}{\sqrt{\left|2\pi\mathbf{Q}_{\mathbf{x}_\mathbf{\Delta}}\right|}}, \tag{35}$$

which after removal of the square roots and factoring out the constants, gives a simpler expression that relates the determinant terms of the structured matrices in the GNP problem:

$$\left|\mathbf{Q}_{\mathbf{x}_\mathbf{\Delta}}\right| = \frac{|\mathbf{R}|\prod_{i\in\mathcal{J}}|\mathbf{S}_i|}{\left|\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}\right|}. \tag{36}$$

F. GNP Costs

In this section, we provide the GNP costs in a form that includes the non-assignment gate with the same structure as found in [16] and [20]. Multiplying the joint posterior of (6) by the hypothesis-invariant term $\frac{1}{C}P_{\mathcal{AB}}^{-m}\beta^{-n}P_{\mathcal{A}\bar{\mathcal{B}}}^{(m-n)}(2\pi)^{d(m+1)/2}\sqrt{|\mathbf{R}|}$ gives[4]

$$\Pr\left(\mathbf{h},\ \mathbf{b}|\mathbf{x}_1^{\mathcal{A}},..,\mathbf{x}_m^{\mathcal{A}},\ \mathbf{x}_1^{\mathcal{B}},..,\mathbf{x}_n^{\mathcal{B}}\right)$$
$$\propto e^{-\frac{1}{2}\mathbf{b}^T\mathbf{R}^{-1}\mathbf{b}}\left(\frac{P_{\mathcal{AB}}}{(2\pi)^{\frac{d}{2}}\beta P_{\bar{\mathcal{A}}\mathcal{B}}P_{\mathcal{A}\bar{\mathcal{B}}}}\right)^{-(m-n_a)}$$
$$\times\prod_{i\in\mathcal{J}}\frac{1}{\sqrt{|\mathbf{S}_i|}}e^{-\frac{1}{2}\left(\mathbf{x}_i^{\Delta}-\mathbf{b}\right)^T\mathbf{S}_i^{-1}\left(\mathbf{x}_i^{\Delta}-\mathbf{b}\right)}. \quad (37)$$

For a given hypothesis, the GNPM approach selects $\mathbf{b}$ that maximizes the score, thus GNPM is (37) evaluated at $\mathbf{b} = \bar{\mathbf{x}}$. As a critical note, the results of [1] stress the use of unitless likelihood ratios. Since the first term in (37) is unitless and the units on the remaining terms are V^{n_a-m} and V^{-n_a}, respectively, where V is a unit hypersphere of the surveillance volume, the units of (37) are hypothesis invariant as V^{-m}. Therefore, hypotheses with varying numbers of assignments have the same units and one may safely use (37) within a specific GNP problem. However, if the GNPM cost is used in a higher context application, for example, in a suboptimal solution for the association of more than two sensor data, care must be taken with the units of (37). We prefer to keep units in the score to be consistent with [16] and [20].

Taking the negative logarithm of (37) evaluated at $\mathbf{b} = \bar{\mathbf{x}}$ and multiplying by 2 gives the GNPM cost as

$$C_{GNPM}(\mathbf{h}) = \bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}}$$
$$+2(m-n_a)\log G_0$$
$$+\sum_{i\in\mathcal{J}}\left[\log(|\mathbf{S}_i|) + (\mathbf{x}_i^{\Delta}-\bar{\mathbf{x}})^T\mathbf{S}_i^{-1}\left(\mathbf{x}_i^{\Delta}-\bar{\mathbf{x}}\right)\right], \quad (38)$$

where

$$G_0 = \frac{P_{\mathcal{AB}}}{(2\pi)^{d/2}\beta P_{\bar{\mathcal{A}}\mathcal{B}}P_{\mathcal{A}\bar{\mathcal{B}}}} \quad (39)$$

is the non-assignment gate value used in track-to-track assignment problems [5].[5]

Applying the equivalence from (33) and multiplying by the hypothesis-invariant term $\frac{1}{C}P_{\mathcal{AB}}^{-m}\beta^{-n}P_{\mathcal{A}\bar{\mathcal{B}}}^{(m-n)}(2\pi)^{dm/2}\sqrt{|\mathbf{R}|}$ gives

$$\Pr\left(\mathbf{h}|\mathbf{x}_1^{\mathcal{A}},..,\mathbf{x}_m^{\mathcal{A}},\ \mathbf{x}_1^{\mathcal{B}},..,\mathbf{x}_n^{\mathcal{B}}\right)$$
$$\propto\sqrt{\left|\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}\right|}e^{-\frac{1}{2}\bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}}}\left(\frac{P_{\mathcal{AB}}}{(2\pi)^{\frac{d}{2}}\beta P_{\bar{\mathcal{A}}\mathcal{B}}P_{\mathcal{A}\bar{\mathcal{B}}}}\right)^{-(m-n_a)}$$
$$\times\prod_{i\in\mathcal{J}}\frac{1}{\sqrt{|\mathbf{S}_i|}}e^{-\frac{1}{2}\left(\mathbf{x}_i^{\Delta}-\bar{\mathbf{x}}\right)^T\mathbf{S}_i^{-1}\left(\mathbf{x}_i^{\Delta}-\bar{\mathbf{x}}\right)}, \quad (40)$$

which has different units than GNPM through the determinant of the a-posteriori bias covariance. Converting to cost format the MTTA cost is

$$C_{MTTA}(\mathbf{h}) = \bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}} - \log\left(\left|\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}\right|\right)$$
$$+2(m-n_a)\log G_0$$
$$+\sum_{i\in\mathcal{J}}\left[\log(|\mathbf{S}_i|) + (\mathbf{x}_i^{\Delta}-\bar{\mathbf{x}})^T\mathbf{S}_i^{-1}\left(\mathbf{x}_i^{\Delta}-\bar{\mathbf{x}}\right)\right]. \quad (41)$$

Note that (41) is not written exactly as was provided [20], but is equivalent through the result of (33) with hypothesis invariant terms removed.

G. Equivalence with GNN

Intuitively, the GNP problem in both the GNPM and MTTA form is expected to reduce to the classic GNN problem as $\mathbf{R} \to \mathbf{0}$. However, this does not directly follow from (38) due to the indeterminate 0/0 that arises in $\bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}}$. As shown in Appendix C, application of the key results of Section II.E avoids this issue and both the GNPM and MTTA costs reduce to GNN as $\mathbf{R} \to \mathbf{0}$, thus

$$C_{GNPM}(\mathbf{h})\big|_{\mathbf{R}\to\mathbf{0}} = C_{MTTA}(\mathbf{h})\big|_{\mathbf{R}\to\mathbf{0}} = C_{GNN}(\mathbf{h}), \quad (42)$$

where

$$C_{GNN}(\mathbf{h}) = 2(m-n_a)\log G_0$$
$$+\sum_{i\in\mathcal{J}}\left[\log(|\mathbf{S}_i|) + (\mathbf{x}_i^{\Delta})^T\mathbf{S}_i^{-1}\left(\mathbf{x}_i^{\Delta}\right)\right]. \quad (43)$$

Therefore, when $\mathbf{R}$ is sufficiently small, a GNN algorithm is suitable since GNPM and MTTA effectively give the same answer as GNN, as demonstrated in [16].

III. PRACTICAL CONSIDERATIONS

In this section, we provide further insight into the cost differences and elaborate on the practical relevance. We begin with a discussion on behavior of non-assignment costs, and then conclude with a discussion on bias estimation within the costs.

A. Optimal Non-Assignment Costs

Motivated by solution algorithms, we prefer to think of the track assignment problem in an incremental cost

[4]The determinant identity $\sqrt{|2\pi\mathbf{\Sigma}|} = (2\pi)^{d/2}\sqrt{|\mathbf{\Sigma}|}$ is used in (6) to allow the 2π term to be factored.

[5]A minor difference in the gate compared to previous literature is the density of false tracks, which we have taken as zero. For applications that need false target densities, we recommend using the gate value in [20], which is a trivial adjustment of (39). With false target densities as zero, the gate value in [20] is exactly (39).

structure, which starts from no assignments and incrementally seeks additional assignments that lower the cost as in the algorithm of [17]. In light of this concept, by inspection of (43), the incremental cost of adding track tuple $(i,\ h_i)$ to the assignment set in GNN is $(\mathbf{x}_i^\Delta)^T\mathbf{S}_i^{-1}(\mathbf{x}_i^\Delta) + \log(|\mathbf{S}_i|)$. Therefore, the optimal decision in GNN is to accept the assignment for consideration if the statistical distance (of d degrees of freedom) does not exceed the covariance-dependent threshold

$$(\mathbf{x}_i^\Delta)^T\mathbf{S}_i^{-1}\left(\mathbf{x}_i^\Delta\right) < 2\log G_0 - \log\left(|\mathbf{S}_i|\right). \tag{44}$$

We interpret the physical meaning of the GNN assignment threshold as evaluating the probability that the tracks in the pair are on different targets that randomly appeared in the containment volume of the covariance, based on the a-priori spatial density that tracked targets may appear. Therefore, as the statistical distance of the pair increases, corresponding to a larger containment volume, it is more likely that one of the tracks is on a different target. Note that satisfying this inequality does not necessarily guarantee any particular assignment, as there may be other assignment pairs with lower cost. Once the best available assignment fails this inequality, no additional assignments may be added and all unassigned tracks remain as singletons.

Seeking an analogous threshold for the GNPM cost of (38) is challenging since the acceptance of a new track assignment adjusts the bias estimation within the hypothesis. Recalling the equivalence found in (22), algebraic manipulations give the expected value of the bias term $\bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}}$ in (38) as

$$\begin{aligned}
&E[\bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}}]\\
&= E\left[tr\left(\bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}}\right)\right]\\
&= tr\left(\mathbf{R}^{-1}E\left[\bar{\mathbf{x}}\bar{\mathbf{x}}^T\right]\right)\\
&= tr\left(\mathbf{R}^{-1}E\left[\mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}^T\mathbf{Q}_{\mathbf{x}_\Delta}^{-1}\mathbf{x}_\Delta\left(\mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}^T\mathbf{Q}_{\mathbf{x}_\Delta}^{-1}\mathbf{x}_\Delta\right)^T\right]\right)\\
&= tr\left(\mathbf{R}^{-1}\mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}^T\mathbf{Q}_{\mathbf{x}_\Delta}^{-1}\mathbf{Q}_{\mathbf{x}_\Delta}\mathbf{Q}_{\mathbf{x}_\Delta}^{-1}\mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}\right)\\
&= tr\left(\mathbf{R}^{-1}\mathbf{R}\mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta}^{-1}\mathbf{H}\mathbf{R}\right)\\
&= tr\left(\mathbf{H}^T\left(\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}} + \mathbf{H}\mathbf{R}\mathbf{H}^T\right)^{-1}\mathbf{H}\mathbf{R}\right)\\
&= tr\left(\mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\left(\mathbf{I} + \mathbf{H}\mathbf{R}\mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\right)^{-1}\mathbf{H}\mathbf{R}\right)\\
&= tr\left(\mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{H}\mathbf{R}\left(\mathbf{I} + \mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{H}\mathbf{R}\right)^{-1}\right) \qquad (45)\\
&= tr\left(\mathbf{I} - \left(\mathbf{I} + \mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{H}\mathbf{R}\right)^{-1}\right)\\
&= d - tr\left(\left(\mathbf{I} + \mathbf{H}^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\mathbf{H}\mathbf{R}\right)^{-1}\right)\\
&= d - tr\left(\left(\mathbf{I} + \sum_{i\in\mathcal{J}}\mathbf{S}_i^{-1}\mathbf{R}\right)^{-1}\right),
\end{aligned}$$

which is limited to $[0, d]$ since each $\mathbf{S}_i$ and $\mathbf{R}$ are symmetric and positive definite matrices.[6] Therefore, when the final term in (45) vanishes, the incremental cost of the i^{th} assignment, in an expected value sense, is completely contained in the $\log(|\mathbf{S}_i|) + (\mathbf{x}_i^\Delta - \bar{\mathbf{x}})^T\mathbf{S}_i^{-1}(\mathbf{x}_i^\Delta - \bar{\mathbf{x}})$ term. Thus, under this assumption and by inspection of (38), the analogous threshold from GNPM follows the same structure as GNN

$$(\mathbf{x}_i^\Delta - \bar{\mathbf{x}})^T\mathbf{S}_i^{-1}\left(\mathbf{x}_i^\Delta - \bar{\mathbf{x}}\right) < 2\log G_0 - \log\left(|\mathbf{S}_i|\right), \tag{46}$$

which is a statistical distance of d degrees of freedom compared to a threshold that is dependent upon the covariance used in that statistical distance.

As discussed and demonstrated in [18], since GNPM follows the same threshold decision structure as GNN, G_0 is a nearly optimal gate for GNPM. Critically, the gate is optimal when the final term in (45) vanishes, which occurs after several assignments are made or after the first assignment when $\mathbf{R} \gg \mathbf{S}_i$. In [18], an optimal gate was provided for the case where only one assignment is made, but we do not recommend this in practice since intuitively the notion of a pattern match is only meaningful with multiple assignments.

The determination of incremental cost for MTTA is further complicated by the $\log(|\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}|)$ term in (41), which introduces dependence upon the specific assignments made, including the incremental addition of tuple $(i,\ h_i)$. To allow an approximate analysis, we make the simplifying assumption that each $\mathbf{S}_i = \mathbf{S}$ (this condition is not required by GNPM or MTTA) and that enough assignments are made such that the final term in (45) vanishes. With these assumptions after n_a assignments are made, $\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta} = (\mathbf{R}^{-1} + \sum_{i\in\mathcal{J}}\mathbf{S}^{-1})^{-1} \approx (n_a\mathbf{S}^{-1})^{-1} = \mathbf{S}/n_a$. Given $n_a - 1$ assignments made before incrementally adding the tuple, an approximation of the incremental cost of the $-\log(|\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}|)$ term is

$$-\log\left(|\mathbf{S}/n_a|\right) + \log\left(\left|\mathbf{S}/\left(n_a - 1\right)\right|\right) = \log\left(\frac{n_a}{n_a - 1}\right). \tag{47}$$

Therefore, for MTTA, the approximate incremental cost is $\log(|\mathbf{S}|) + (\mathbf{x}_i^\Delta - \bar{\mathbf{x}})^T\mathbf{S}^{-1}(\mathbf{x}_i^\Delta - \bar{\mathbf{x}}) + \log(n_a/(n_a - 1))$ and the analogous decision threshold is

$$\begin{aligned}(\mathbf{x}_i^\Delta - \bar{\mathbf{x}})^T\mathbf{S}^{-1}\left(\mathbf{x}_i^\Delta - \bar{\mathbf{x}}\right) &< 2\log G_0 - \log\left(|\mathbf{S}|\right)\\ &\quad - \log\left(\frac{n_a}{n_a - 1}\right),\end{aligned} \tag{48}$$

[6] We use the relationships $(\mathbf{I} + \mathbf{P}\mathbf{Q})^{-1}\mathbf{P} = \mathbf{P}(\mathbf{I} + \mathbf{Q}\mathbf{P})^{-1}$ and $\mathbf{A}(\mathbf{I} + \mathbf{A})^{-1} = \mathbf{I} - (\mathbf{I} + \mathbf{A})^{-1}$ in (45).

which does not take the same form as the GNN threshold and G_0 is therefore not an optimal gate for MTTA. However, with many assignments, the $\log(n_a/(n_a-1))$ term eventually vanishes and we expect G_0 to be nearly optimal for MTTA in problems with a high number of common targets.

The incremental cost of the i^{th} assignment is not the only mechanism of non-assignment behavior, the cost of the null hypothesis (i.e., the hypothesis of no assignments) also plays a significant role. Since the bias estimate in the null assignment is $\bar{\mathbf{x}}_0 = \mathbf{0}$ and recognizing from (25) that $\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta} = \mathbf{R}$ if there are no assignments, the GNPM and MTTA costs for the null hypothesis $\mathbf{h}_0$ are

$$C_{GNPM}(\mathbf{h}_0) = 2m \log G_0, \tag{49}$$

$$C_{MTTA}(\mathbf{h}_0) = 2m \log G_0 - \log(|\mathbf{R}|). \tag{50}$$

Therefore, with large $\mathbf{R}$, the null hypothesis in MTTA can dominate over other hypotheses. This trait is not present with GNPM, which can generally be expected to provide assignments using an arbitrarily large $\mathbf{R}$.

We demonstrate the analytical results for non-assignment behavior with numerical simulations. Consider a scenario where sensor $\mathcal{A}$ observes six targets and sensor $\mathcal{B}$ observes eight, with three targets in common. By the formula given in [16], a 6×8 track association problem has a total of 93,289 possible hypotheses. Assume that each of the 11 total targets are randomly generated in a hypersphere of dimension $d = 3$ with a uniform distribution, giving a target density of $\beta = 11$. These numbers are sufficient to evaluate the parameters in G_0 as $P_{\mathcal{A}\mathcal{B}} = 3/11$, $P_{\bar{\mathcal{A}}\mathcal{B}} = 5/11$, and $P_{\mathcal{A}\bar{\mathcal{B}}} = 3/11$, and thus $G_0 = 0.0127$. Assume that the track covariances in each hypothesis satisfy $\mathbf{S}_i = \mathbf{S} = \sigma^2\mathbf{I}$ and that the bias covariance satisfies $\mathbf{R} = \sigma_b^2\mathbf{I}$. In Monte Carlo experiments, we evaluate the probability of correct association, which is the total number of correct entries in the most likely assignment vector $\mathbf{h}$ as evaluated for the GNPM and MTTA costs. In the Monte Carlo experiments, a test gate, G_{test}, offset from the optimal gate of (39) is used in the cost functions and 10^4 Monte Carlo trials are performed for each G_{test}. The structure of these experiments is very similar to the numerical results of [23], which evaluated the fraction of correct assignments using various non-assignment thresholds.

In the first experiment, the Monte Carlo simulation varies σ while maintaining $\sigma_b = 5\sigma$, and these results are provided in Fig. 1. As shown, G_0 gives very close to optimal performance for GNPM, but a G_{test} slightly larger than G_0 gives maximal probability of correct association for MTTA. This illustrates the analytic result of (48) which, with several assumptions, predicts that G_0 is generally not an optimal gate for MTTA, particularly if there are few assignments made. The performance loss using G_0 for MTTA in this case is likely negligible as it causes less than a percentage point from maximal performance if that maximal performance is above 90%.

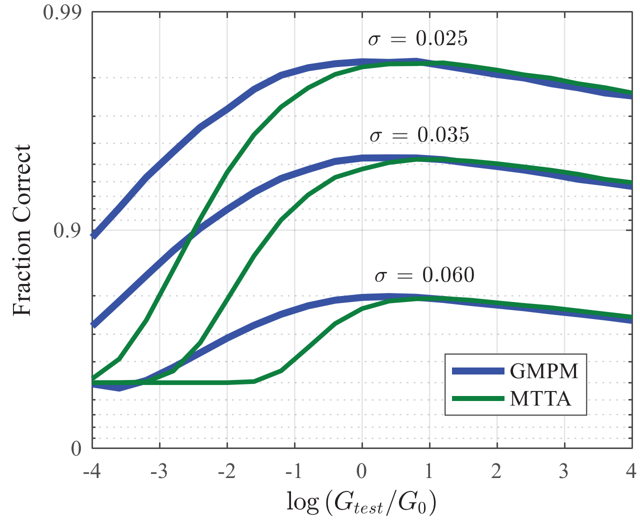

Fig. 1. Probability of correct track-to-track association for various covariance sizes. In this case, G_0 is a nearly optimal gate for GNPM.

In the next experiment, we maintain σ as the single value of 0.025, but set σ_b to values of 0.5σ, 5σ, and 60σ. The results are provided in Fig. 2, which illustrates that G_0 is not an optimal gate for MTTA when σ_b is large, while GNPM maintains G_0 as a nearly optimal gate. This illustrates the analytic result of (50), which states that the null hypothesis can dominate over other hypotheses if $\mathbf{R}$ is large. As discussed in the derivation of (48) and (50), the effects of the $\log(|\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}|)$ term in the MTTA cost cause performance loss with G_0. However, by the key result of (33), any removal of the effects $\log(|\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}|)$ cause in the incremental cost structure for MTTA effectively gives the GNPM cost.

Additionally, the result in Fig. 2 corresponding to the lowest σ_b illustrates (42), which states that GNPM and MTTA are equivalent as $\mathbf{R} \rightarrow \mathbf{0}$. As a final observation, maximal performance of both GNPM and MTTA reduces as σ_b grows. This is the intuitive result that

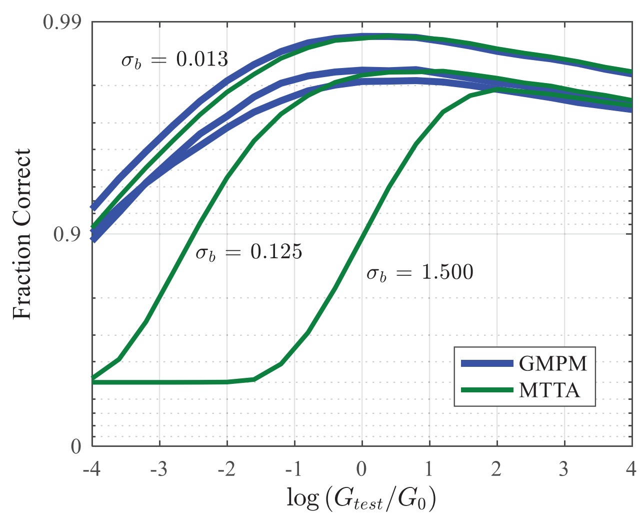

Fig. 2. Probability of correct track-to-track association for various σ_b values. With large σ_b, G_0 is not an optimal non-assignment gate for MTTA.

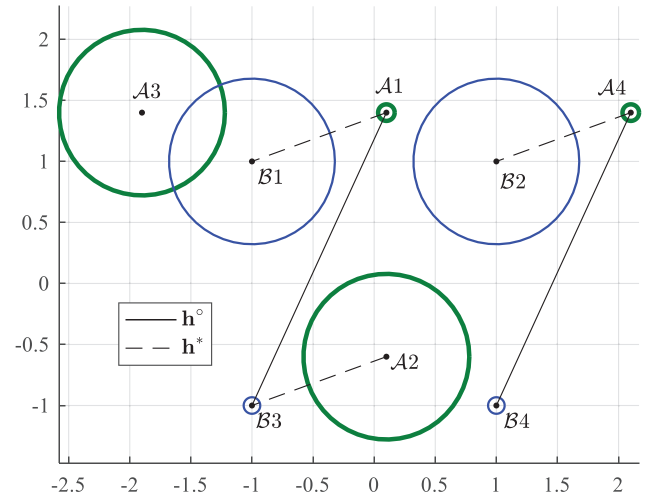

Fig. 3. 4×4 track assignment example. Targets 1 and 4 from Sensor $\mathcal{A}$ and 3 and 4 from Sensor $\mathcal{B}$ have a variable covariance $\boldsymbol{\sigma}^2\mathbf{I}$. All other tracks have covariance $0.1\mathbf{I}$. Circles represent 90% containment areas with $\boldsymbol{\sigma}^2 = 10^{-3}$.

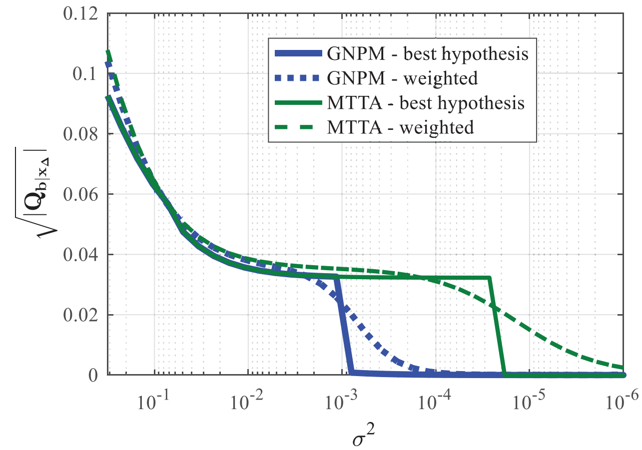

Fig. 4. A-posteriori bias covariance sizes of GNPM/MTTA hypotheses from the track sets in Fig. 3. GNPM determines $\mathbf{h}^\circ$ as the best hypothesis near $\sigma^2 = 10^{-3}$ and MTTA near $\sigma^2 = 10^{-5}$.

some track association performance is lost when bias is added to the GNN problem, and this was also reported in [16].

B. A-Posteriori Bias Covariance

As evidenced by (33), the difference between the GNPM and MTTA costs is in the a-posteriori bias covariance. Inspired by the example in [8], we use the 4×4, $d = 2$ numerical example in Fig. 3 to illustrate the practical difference between the cost functions. A 4×4 track assignment problem gives 209 possible hypotheses. In this scenario, we let the covariance values of tracks 1 and 4 from Sensor $\mathcal{A}$ and tracks 3 and 4 from Sensor $\mathcal{B}$ vary from very high to very low values, but let the others maintain the value of $0.1\mathbf{I}$. If the variable covariances are large, the hypothesis of three assignments, $\mathbf{h}^* = [1\ 3\ 0\ 2]$ (i.e., $\mathcal{A}1 \rightarrow \mathcal{B}1$, $\mathcal{A}2 \rightarrow \mathcal{B}3$, and $\mathcal{A}4 \rightarrow \mathcal{B}2$ as illustrated in Fig. 3)**,** is preferable since the track states align and only a small shift is needed for the alignment. However, as the variable covariances shrink to very small size, the hypothesis of $\mathbf{h}^\circ = [3\ 0\ 0\ 4]$ becomes more probable. In other words, given the a-priori assumptions that targets appear at random locations in the surveillance volume, the probability that the pattern difference $[(\mathcal{A}1 - \mathcal{A}4) - (\mathcal{B}3 - \mathcal{B}4)]^2 < \sigma^2$ occurs by random chance is essentially zero as $\sigma^2 \rightarrow 0$.

To illustrate the practical difference in the cost formulations, we find the track covariance size for GNPM and MTTA that gives $\mathbf{h}^\circ$ as the definitive hypothesis. Provided in Fig. 4 is the a-posteriori bias covariance of the top hypothesis from the GMPM and MTTA costs. For this numerical experiment, we let $\mathbf{R} = \mathbf{I}$ and $G_0 = 19.2$. As shown, GNPM determines the definitive hypothesis with a larger σ than MTTA. This example illustrates that GNPM generally tends to prefer (and score favorably) hypotheses that give smaller $|\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}|$. Further illustrating this concept, we also provide the posterior-weighted $\sqrt{|\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}|}$ in Fig. 4, using all 209 posteriors (from (37) and (40)) normalized to sum to unity. As shown, hypotheses with large a-posteriori bias covariance scored by GNPM have nearly zero weight as $\sigma^2 < 10^{-4}$, while MTTA maintains significant on those hypotheses.

IV. CONCLUDING REMARKS

GNP costs have their typical use in track-to-track association problems. Compared to traditional literature for track-to-track association, the GNP problem includes unknown sensor bias into the observation model. The two types of GNP costs discussed in this work are the GNPM and MTTA costs. GNPM involves the joint likelihood of both a hypothesis and a-posteriori bias estimate, while MTTA marginalizes bias from the problem. Here, we showed the intuitive result that the analytic difference between GNPM and MTTA kinematic scores is the determinant of the a-posteriori bias covariance. Several key insights arise through that result, including equivalences with the distribution of total errors and the role of bias estimation as a separate step in cost calculations. Leveraging this result, through an inspection of the GNPM incremental assignment cost, we argue that the non-assignment cost G_0 is nearly optimal for GNPM and demonstrate with numerical examples. However, through similar inspection of the MTTA incremental cost, G_0 is not optimal for MTTA and the significance diminishes for problems with many assignments but grows with large $\mathbf{R}$. Removal of the covariance of the a-posteriori bias from the MTTA non-assignment cost to give maximal probability of correct association effectively yields the GNPM cost. Therefore, if the goal of a GNP algorithm is to maximize the probability of correct association, we recommend GNPM. As a final experiment, through a simple two-dimensional example, we show that GNPM tends to favor hypotheses with smaller a-posteriori bias covariance compared to MTTA. In conclusion, the results contained here expand upon previous literature to reveal important design considerations for specific track-to-track association problems.

APPENDIX A

Provided in this appendix is a derivation of the marginal and conditional densities for the random vectors within the track assignment problem. Explicitly writing $\mathbf{x}_i^{\Delta}$ from (7) to expose the noise terms gives

$$\begin{aligned}\mathbf{x}_i^{\Delta} &= \mathbf{x}_i^{\mathcal{A}} - \mathbf{x}_{h_i}^{\mathcal{B}} \\ &= \mathbf{x}_i + \mathbf{n}_i^{\mathcal{A}} - \left(\mathbf{x}_i - \mathbf{b} + \mathbf{n}_{h_i}^{\mathcal{B}}\right) \\ &= (\mathbf{n}_i^{\mathcal{A}} - \mathbf{n}_{h_i}^{\mathcal{B}}) + \mathbf{b},\end{aligned} \quad (51)$$

since, given $\mathbf{h}$ is the correct hypothesis, each track is an observation of $\mathbf{x}_i$. Defining the combined noise term as $\mathbf{n}_i = (\mathbf{n}_i^{\mathcal{A}} - \mathbf{n}_{h_i}^{\mathcal{B}})$ which is zero mean with covariance $\mathbf{S}_i = \mathbf{S}_{\mathcal{A},i} + \mathbf{S}_{\mathcal{B},h_i}$, we have

$$\mathbf{x}_i^{\Delta} = \mathbf{n}_i + \mathbf{b}. \quad (52)$$

Let $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_{n_a}]$ to be a length n_a vector that contains an ordering of the indices in $\mathcal{J}$, which, in other words, is simply a list of the track indices from sensor $\mathcal{A}$ that are assigned to a track from sensor $\mathcal{B}$. Assuming all error terms are uncorrelated, the stacked vector of error terms is a normally distributed random vector with a block diagonal covariance, expressed as

$$\begin{bmatrix}\mathbf{n}_{\gamma_1} \\ \vdots \\ \mathbf{n}_{\gamma_{na}} \\ \mathbf{b}\end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix}\mathbf{S}_{\gamma_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \mathbf{S}_{\gamma_{na}} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{R}\end{bmatrix}\right). \quad (53)$$

Defining the stacked vector of error terms as

$$\mathbf{x}_{\Delta} = \begin{bmatrix}\mathbf{x}_{\gamma_1}^{\Delta} \\ \vdots \\ \mathbf{x}_{\gamma_{na}}^{\Delta}\end{bmatrix}, \quad (54)$$

left multiplication of (53) by the transform matrix

$$\mathbf{V}_{\Delta} = \begin{bmatrix}\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I} & \mathbf{I}\end{bmatrix}, \quad (55)$$

gives the distribution of absolute error between the observations as a normally distributed random vector

$$p(\mathbf{x}_{\Delta}) = \mathcal{N}(\mathbf{0}, \mathbf{Q}_{\mathbf{x}_{\Delta}}), \quad (56)$$

with covariance

$$\begin{aligned}\mathbf{Q}_{\mathbf{x}_{\Delta}} &= \mathbf{V}_{\Delta}\begin{bmatrix}\mathbf{S}_{\gamma_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \mathbf{S}_{\gamma_{na}} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{R}\end{bmatrix}\mathbf{V}_{\Delta}^T \\ &= \begin{bmatrix}\mathbf{S}_{\gamma_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S}_{\gamma_{na}}\end{bmatrix} + \begin{bmatrix}\mathbf{R} & \cdots & \mathbf{R} \\ \vdots & \ddots & \vdots \\ \mathbf{R} & \cdots & \mathbf{R}\end{bmatrix}.\end{aligned} \quad (57)$$

Next, we separate the distribution of absolute errors into conditional distributions. Defining the joint vector of absolute errors and bias as

$$\mathbf{x}_{\mathbf{b}} = \begin{bmatrix}\mathbf{x}_{\Delta} \\ \mathbf{b}\end{bmatrix}, \quad (58)$$

left multiplication of (53) by a similar transformation matrix

$$\mathbf{V} = \begin{bmatrix}\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I}\end{bmatrix}, \quad (59)$$

gives the joint distribution of absolute error and bias as a zero mean normally distributed random vector

$$p(\mathbf{x}_{\mathbf{b}}) = \mathcal{N}(\mathbf{0}, \mathbf{Q}), \quad (60)$$

with covariance written in block partition form

$$\begin{aligned}\mathbf{Q} &= \mathbf{V}\begin{bmatrix}\mathbf{S}_{\gamma_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \mathbf{S}_{\gamma_{na}} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{R}\end{bmatrix}\mathbf{V}^T \\ &= \begin{bmatrix}\mathbf{Q}_{\mathbf{x}_{\Delta}} & \mathbf{P}_{\mathbf{b}\mathbf{x}_{\Delta}} \\ \mathbf{P}_{\mathbf{b}\mathbf{x}_{\Delta}}^T & \mathbf{R}\end{bmatrix}.\end{aligned} \quad (61)$$

The cross-correlation matrix is the block matrix

$$\mathbf{P}_{\mathbf{b}\mathbf{x}_{\Delta}} = \begin{bmatrix}\mathbf{R} \\ \vdots \\ \mathbf{R}\end{bmatrix}. \quad (62)$$

Applying the fundamental equations of linear estimation from [3] gives the conditional distributions according to both $\mathbf{x}_\Delta$ and $\mathbf{b}$. Defining a stacked matrix of identity matrices as $\mathbf{H} = [\mathbf{I} \ldots \mathbf{I}]^T$, the conditional distribution of the absolute errors given the relative bias is a normally distributed random vector

$$p\left(\mathbf{x}_\Delta|\mathbf{b}\right) = \mathcal{N}\left(\mu_{\mathbf{x}_\Delta|\mathbf{b}},\ \mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}\right), \tag{63}$$

with mean

$$\begin{aligned}\mu_{\mathbf{x}_\Delta|\mathbf{b}} &= \mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}\mathbf{R}^{-1}\mathbf{b} \\ &= \mathbf{H}\mathbf{b},\end{aligned} \tag{64}$$

and corresponding covariance

$$\begin{aligned}\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}} &= \mathbf{Q}_{\mathbf{x}_\Delta} - \mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}\mathbf{R}^{-1}\mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}^T \\ &= \mathbf{Q}_{\mathbf{x}_\Delta} - \mathbf{H}\mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}^T \\ &= \mathbf{Q}_{\mathbf{x}_\Delta} - \begin{bmatrix}\mathbf{R} & \cdots & \mathbf{R} \\ \vdots & \ddots & \vdots \\ \mathbf{R} & \cdots & \mathbf{R}\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{S}_{\gamma_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S}_{\gamma_{n_a}}\end{bmatrix}.\end{aligned} \tag{65}$$

Next we apply conditional density relationships to (60) to write the converse distribution. The conditional distribution of the bias given the absolute errors is a normally distributed random vector

$$p\left(\mathbf{b}|\mathbf{x}_\Delta\right) = \mathcal{N}\left(\mu_{\mathbf{b}|\mathbf{x}_\Delta},\ \mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta}\right), \tag{66}$$

with mean and covariance

$$\mu_{\mathbf{b}|\mathbf{x}_\Delta} = \mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}^T\mathbf{Q}_{\mathbf{x}_\Delta}^{-1}\mathbf{x}_\Delta, \tag{67}$$

$$\mathbf{Q}_{\mathbf{b}|\mathbf{x}_\Delta} = \mathbf{R} - \mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}^T\mathbf{Q}_{\mathbf{x}_\Delta}^{-1}\mathbf{P}_{\mathbf{b}\mathbf{x}_\Delta}, \tag{68}$$

thus completing the derivation of the desired probability distributions.

APPENDIX B

In this appendix, we algebraically show the equivalence of (34). To reduce cumbersome nomenclature, we drop the subscripts used in (34). Specifically, we establish the following equivalence:

$$\mathbf{x}^T\mathbf{Q}^{-1}\mathbf{x} = \left(\mathbf{x} - \mathbf{H}\mathbf{b}\right)^T\mathbf{S}^{-1}\left(\mathbf{x} - \mathbf{H}\mathbf{b}\right) + \mathbf{b}^T\mathbf{R}^{-1}\mathbf{b}, \tag{69}$$

given $\mathbf{b} = (\mathbf{R}^{-1} + \mathbf{H}^T\mathbf{S}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{S}^{-1}\mathbf{x}$, $\mathbf{Q} = \mathbf{S} + \mathbf{H}\mathbf{R}\mathbf{H}^T$, and $\mathbf{H} = [\mathbf{I} \ldots \mathbf{I}]^T$. Assume that all necessary matrix inverses exist.

Beginning with expression for $\mathbf{b}$, multiplication of both sides by $(\mathbf{R}^{-1} + \mathbf{H}^T\mathbf{S}^{-1}\mathbf{H})$ gives the useful preliminary relationship,

$$\begin{aligned}\left(\mathbf{R}^{-1} + \mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\right)\mathbf{b} &= \mathbf{H}^T\mathbf{S}^{-1}\mathbf{x} \\ \mathbf{R}^{-1}\mathbf{b} + \mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\mathbf{b} &= \mathbf{H}^T\mathbf{S}^{-1}\mathbf{x} \\ \mathbf{R}^{-1}\mathbf{b} &= \mathbf{H}^T\mathbf{S}^{-1}\mathbf{x} - \mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\mathbf{b}.\end{aligned} \tag{70}$$

Application of the matrix inversion lemma to $\mathbf{Q}^{-1}$ gives

$$\begin{aligned}\mathbf{Q}^{-1} &= \left(\mathbf{S} + \mathbf{H}\mathbf{R}\mathbf{H}^T\right)^{-1} \\ &= \mathbf{S}^{-1} - \mathbf{S}^{-1}\mathbf{H}\Big(\mathbf{R}^{-1} + \mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\Big)^{-1}\mathbf{H}^T\mathbf{S}^{-1},\end{aligned} \tag{71}$$

therefore, the full chi-square term can be written as

$$\begin{aligned}&\mathbf{x}^T\mathbf{Q}^{-1}\mathbf{x} \\ &= \mathbf{x}^T\left[\mathbf{S}^{-1} - \mathbf{S}^{-1}\mathbf{H}\Big(\mathbf{R}^{-1} + \mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\Big)^{-1}\mathbf{H}^T\mathbf{S}^{-1}\right]\mathbf{x} \\ &= \mathbf{x}^T\mathbf{S}^{-1}\mathbf{x} - \mathbf{x}^T\mathbf{S}^{-1}\mathbf{H}\Big(\mathbf{R}^{-1} + \mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\Big)^{-1}\mathbf{H}^T\mathbf{S}^{-1}\mathbf{x}.\end{aligned} \tag{72}$$

Since the expression for $\mathbf{b}$ appears in (72), we have

$$\mathbf{x}^T\mathbf{Q}^{-1}\mathbf{x} = \mathbf{x}^T\mathbf{S}^{-1}\mathbf{x} - \mathbf{x}^T\mathbf{S}^{-1}\mathbf{H}\mathbf{b}. \tag{73}$$

Recognizing that (73) is a portion of the quadratic expansion of $(\mathbf{x} - \mathbf{H}\mathbf{b})^T\mathbf{S}^{-1}(\mathbf{x} - \mathbf{H}\mathbf{b})$, rewriting to include the addition of terms that complete the quadratic expansion gives

$$\begin{aligned}\mathbf{x}^T\mathbf{Q}^{-1}\mathbf{x} &= \left(\mathbf{x} - \mathbf{H}\mathbf{b}\right)^T\mathbf{S}^{-1}\left(\mathbf{x} - \mathbf{H}\mathbf{b}\right) \\ &\quad + \mathbf{b}^T\left(\mathbf{H}^T\mathbf{S}^{-1}\mathbf{x} - \mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\mathbf{b}\right).\end{aligned} \tag{74}$$

Substituting (70) into (74) gives the desired equivalency of (69).

APPENDIX C

In this appendix, we establish the equivalence between the GNPM, MTTA, and GNN costs as $\mathbf{R} \rightarrow \mathbf{0}$. Unfortunately, direct substitution of $\mathbf{R} = \mathbf{0}$ into the GNPM and MTTA costs of (38) and (41) gives indeterminate terms. Applying the key results of Section II.E avoids this issue, and allows simplification to the GNN cost. Converting the chi-square terms of GNPM into a block structure followed by application of (36) gives the following equivalence:

$$\begin{aligned}&\bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}} + \sum_{i\in\mathcal{J}}\left(\mathbf{x}_i^\Delta - \bar{\mathbf{x}}\right)^T\mathbf{S}_i^{-1}\left(\mathbf{x}_i^\Delta - \bar{\mathbf{x}}\right) \\ &= \bar{\mathbf{x}}^T\mathbf{R}^{-1}\bar{\mathbf{x}} + \left(\mathbf{x}_\Delta - \mathbf{H}\bar{\mathbf{x}}\right)^T\mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}^{-1}\left(\mathbf{x}_\Delta - \mathbf{H}\bar{\mathbf{x}}\right) \\ &= \mathbf{x}_\Delta^T\mathbf{Q}_{\mathbf{x}_\Delta}^{-1}\mathbf{x}_\Delta\end{aligned} \tag{75}$$

By inspection of (13), $\mathbf{Q}_{\mathbf{x}_\Delta} = \mathbf{Q}_{\mathbf{x}_\Delta|\mathbf{b}}$ if $\mathbf{R} = \mathbf{0}$, therefore, the limiting form of GNPM can be written as

$$
\begin{aligned}
C_{GNPM}(\mathbf{h})\big|_{\mathbf{R}\to\mathbf{0}} &= \mathbf{x}_{\Delta}^{T}\mathbf{Q}_{\mathbf{x}_{\Delta}|\mathbf{b}}^{-1}\mathbf{x}_{\Delta} + 2(m-n_a)\log G_0 \\
&\quad + \sum_{i\in\mathcal{J}} \log\left(|\mathbf{S}_i|\right) \\
&= 2(m-n_a)\log G_0 \\
&\quad + \sum_{i\in\mathcal{J}} \left[\log\left(|\mathbf{S}_i|\right) + \left(\mathbf{x}_i^{\Delta}\right)^T \mathbf{S}_i^{-1}\left(\mathbf{x}_i^{\Delta}\right)\right] \\
&\equiv C_{GNN}(\mathbf{h}), \qquad (76)
\end{aligned}
$$

which establishes the desired equivalence of GNPM with GNN. Evaluating MTTA as $\mathbf{R} \to \mathbf{0}$ involves the additional complication of $\mathbf{Q}_{\mathbf{b}|\mathbf{x}_{\Delta}}$, which includes $\mathbf{R}^{-1}$ by inspection of (25). Rearranging (36) gives

$$
\frac{\left|\mathbf{Q}_{\mathbf{b}|\mathbf{x}_{\Delta}}\right|}{|\mathbf{R}|} = \frac{\prod_{i\in\mathcal{J}}|\mathbf{S}_i|}{\left|\mathbf{Q}_{\mathbf{x}_{\Delta}}\right|}, \qquad (77)
$$

which for $\mathbf{R} \to \mathbf{0}$ can be reduced to

$$
\frac{\prod_{i\in\mathcal{J}}|\mathbf{S}_i|}{\left|\mathbf{Q}_{\mathbf{x}_{\Delta}|\mathbf{b}}\right|} = \frac{\prod_{i\in\mathcal{J}}|\mathbf{S}_i|}{\prod_{i\in\mathcal{J}}|\mathbf{S}_i|} = 1. \qquad (78)
$$

Therefore, by reintroducing the hypothesis-invariant term $|\mathbf{R}|$ into (40) and applying (75) and (78) gives the MTTA cost as the same functional form of (76) since $\log(1) = 0$, which establishes $C_{MTTA}(\mathbf{h})|_{\mathbf{R}\to\mathbf{0}} = C_{GNN}(\mathbf{h})$.

REFERENCES

[1] Y. Bar-Shalom, S. Blackman, and R. Fitzgerald
"Dimensionless score function for multiple hypothesis tracking,"
IEEE Trans. Aerosp. Electron. Syst., vol. 43, no. 1, pp. 392–400, 2007.

[2] Y. Bar-Shalom and H. Chen
"Multisensor track-to-track association for tracks with dependent errors,"
J. Adv. Inf. Fusion, vol. 1, no 1, pp. 3–14, Jul. 2006.

[3] Y. Bar-Shalom, X. R. Li, and T. Kirbuarajan
Estimation with Applications to Tracking and Navigation. New York, NY, USA: John Wiley and Sons, 2001.

[4] Y. Bar-Shalom, P. K. Willett, and X. Tian
Tracking and Data Fusion: A Handbook of Algorithms. Storrs, CT, USA: YBS Publishing, 2011.

[5] S. S. Blackman and R. Popoli
Design and Analysis of Modern Tracking Systems. Norwood, MA, USA: Artech House, 1999.

[6] C. Y. Chong, S. Mori, and K. C. Chang
"Distrubuted multitarget multisensor tracking,"
in *Multitarget Multisensor Tracking: Advanced Applications*. Boston, MA, USA: Artech House, 1990, pp. 247–295.

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

[8] J. Ferry
"Exact bias removal for the track-to-track association problem,"
in *Proc. 12th Int. Conf. Inf. Fusion*, Seattle, WA, USA, 2009.

[9] J. Ferry
"Exact association probability for data with bias and features,"
J. Adv. Inf. Fusion, vol. 5, no. 1, pp. 41–67, 2010.

[10] H. V. Henderson and S. R. Searle
"On deriving the inverse of a sum of matrices,"
SIAM Review, vol. 23, no. 1, pp. 53–60, 1981.

[11] M. B. Hurley
Track association with Bayesian probability theory, MIT Lincoln Laboratory,
Tech. Rep., Lexington, MA, USA, 2003.

[12] L. M. Kaplan, Y. Bar-Shalom, and W. D. Blair
"Assignment costs for multiple sensor track-to-track association,"
IEEE Trans. Aerosp. Electron. Syst., vol. 44, no. 2, pp. 655–677, 2008.

[13] L. M. Kaplan and W. D. Blair
"Assignment costs for muliple sensor track-to-track association,"
in *Proc. 7th Int. Conf. Inf. Fusion*, Stockholm, Sweden, 2004.

[14] B. Kragel, S. Danford, and A. Poore
"Concurrent MAP data association and absolute bias estimation with an arbitrary number of sensors,"
in *Proc. SPIE, Signal and Data Process. Small Targets,* vol. 6969, 2008.

[15] B. D. Kragel, S. M. Herman, and N. J. Roseveare
"Efficiency and sensitivity of methods for assessing ambiguity in data association decisions,"
in *Proc. SPIE, Signal and Data Process. Small Targets*, Orlando, FL, USA, 2008.

[16] M. Levedahl
"An explicit pattern matching assignment algorithm,"
in *Proc. SPIE, Signal and Data Process. Small Targets,* vol. 4728, pp. 461–469, 2002.

[17] M. Levedahl
Method and system for assigning observations.
U.S. Patent 7092924, 28 February 2002.

[18] M. Levedahl and J. D. Glass
"Optimal non-assignment costs for the GNP problem,"
in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, USA, 2020.

[19] X. Lin, T. Kirubarajan, and Y. Bar-Shalom
"Multisensor bias estimation with local tracks without a priori association,"
in *Proc. SPIE, Signal and Data Process. Small Targets*, vol. 5204, Bellingham, 2003.

[20] D. J. Papageorgiou and M. Holender
"Track-to-track association and ambiguity management in the presence of sensor bias,"
J. Adv. Inf. Fusion, vol. 6, no. 2, pp. 77–100, 2011.

[21] D. J. Papageorgiou and J. D. Sergi
"Simultaneous track-to-track association and bias removal using multistart local search,"
in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, USA, 2008.

[22] K. R. Pattipati, T. Kirubarajan, and R. L. Popp
"Survey of assignment techniques for multitarget tracking,"
in *Multitarget-Multisensor Tracking: Applications and Advances, vol.* 3, Norwood, MA, USA: Artech House, 2000, pp. 77–159.

[23] L. Stone, T. M. Tran, and M. Williams
"Improvement in track-to-track association from using an adaptive threshold,"
in *Proc. 12th Int. Conf. Inf. Fusion*, Seattle, WA, USA, 2009.

[24] L. D. Stone, M. L. Williams, and T. M. Tran
"Track to track association and bias removal,"
in *Proc. SPIE, Signal and Data Process. Small Targets,* vol. 4728, pp. 315–329, 2002.

Mark Levedahl is a Raytheon Engineering Fellow. He has a bachelor's degree in mechanical engineering from McGill University and has been employed by Raytheon since 1991. Mr. Levedahl has more than 40 years of experience working in systems engineering, guidance, navigation, and control-system development, battle management, multi-target multi-sensor tracking, and simulation primarily related to missile defense.

John D. Glass is a Systems Engineer with Raytheon Technologies. In 2009, he graduated from the University of Tennessee with a bachelor of science in electrical engineering, and in 2010 at Georgia Tech with a Master of science in electrical and computer engineering. In May 2015, he completed the Ph.D. degree in electrical and computer engineering at Georgia Tech, focusing on monopulse processing and target tracking. From 2013 to 2017, Dr. Glass was a Research Engineer at Georgia Tech Research Institute working on radar simulation and multi-sensor fusion analysis. From 2011 to 2017, Dr. Glass was a Member of the Editorial Board for the Aerospace and Electronic Systems Magazine as Associate Editor for Student Research, recruiting and handling student highlight articles. Currently, Dr. Glass is a Session Organizer for the IEEE Aerospace Conference. His research interests include target tracking, multisensor fusion, and detection/estimation applied to radar.