arXiv:1012.4795v1 [math.OC] 21 Dec 2010

On the Equivalence of the General Covariance Union (GCU) and Minimum Enclosing Ellipsoid (MEE) Problems

Ottmar Bochardt and **Jeffrey Uhlmann**
Department of Computer Science, 201 EBW,
University of Missouri, Columbia, MO 65211
Phone: 573 884-2129, Fax: 573 882-8318

Abstract

In this paper we describe General Covariance Union (GCU) and show that solutions to GCU and the Minimum Enclosing Ellipsoid (MEE) problems are equivalent. This is a surprising result because GCU is defined over positive semidefinite (PSD) matrices with statistical interpretations while MEE involves PSD matrices with geometric interpretations. Their equivalence establishes an intersection between the seemingly disparate methodologies of covariance-based (e.g., Kalman) filtering and bounded region approaches to data fusion.

1 Introduction

Positive semidefinite matrices are often chosen to represent uncertainty in state estimation and control algorithms because of their special linear-algebraic properties. The Kalman filter, for example, uses PSD matrices to reprepresent covariance upper bounds on the second central moments of unknown probability distributions relating to the state of a system of interest. Bounded region filters, by contrast, use PSD matrices to define ellipsoidal regions which bound the state of the system of interest.

The choice between representing uncertainty with covariance upper bounds versus ellipsoidal bounded regions leads to very different data fusion algorithms with very different filtering and control properties. The statistical interpretation associated with covariance matrices leads to use of the Kalman fusion equations in the case of independent estimates or Covariance Intersection (CI) when the statistical relationships among estimates to be fused cannot be established[9]. Under their respective assumptions, Kalman and CI yield a fused estimate with the smallest possible covariance that is guaranteed to be an upper bound on the second central moment of

the unknown probability distribution defining the state of the system. A bounded region filter, on the other hand, determines the minimum-size ellipsoid that bounds the intersection of the given ellipsoids which are assumed to bound the true state of the system. In summary, one framework is statistical while the other is geometric.

Despite the use of PSD matrices and the exploitation of linearity properties, covariance and ellipsoidal frameworks employ very different mathematical techniques and have almost completely disjoint literatures. (This relationship also holds true more generally between statistics and computational geometry.) A consequence of this fact is that the sophisticated tools developed separately within each framework seem to have no applicability within the other framework. In this paper we make progress toward bridging this gap by establishing a surprising equivalence result.

The structure of this paper is as follows: The next section provides background for the Kalman filter, Covariance Intersection (CI), and Covariance Union (CU). We then present General Covariance Union (GCU) and show that its solution is equivalent to determining the Minimum Enclosing Ellipsoid (MEE) of the ellipsoids defined by its covariance matrix arguments. In other words, GCU and MEE are equivalent. We end with a discussion of the implications of this unexpected result.

2 The Covariance Upper Bound Framework for Data Fusion

An estimate of the squared error, or *covariance*, associated with measurements from a particular sensor *can* be modeled empirically by examining measurements taken of reference objects whose true states are known. This permits an error covariance matrix to be associated with subsequent measurements of objects whose true states are not known. For example, the measured position of an object in two dimensions can be represented as a vector $\mathbf{a}$ consisting of the object's estimated mean position, e.g., $\mathbf{a} = [\mathbf{x}, \mathbf{y}]^T$, and an error covariance matrix $\mathbf{A}$ that expresses the uncertainty associated with the estimated mean. If the error in the estimated mean vector is denoted as $\tilde{\mathbf{a}}$, then the error covariance matrix is an estimate of the expected squared error, $\mathrm{E}[\tilde{\mathbf{a}}\tilde{\mathbf{a}}^T]$.

Ideally, $\mathbf{A}$ would equal $\mathrm{E}[\tilde{\mathbf{a}}\tilde{\mathbf{a}}^T]$, but this is not generally possible to achieve in practice because measurement and process models are never perfect. To accommodate the effect of model errors, prediction and measurement covariances are typically overestimated so as to avoid underestimating the actual squared errors. In other words, a more conservative overestimate of errors is deemed preferable to underestimating the errors. One of many reasons for this preference is that it avoids the consequences of spuriously small errors causing a covariance matrix to become singular or numerically unstable.

Formally, a mean and covariance estimate is said to be consistent (or conservative) if and only if $\mathbf{A}$ - $\mathrm{E}[\tilde{\mathbf{a}}\tilde{\mathbf{a}}^T]$ is positive semidefinite (i.e., has no negative eigenvalues):

$$\mathbf{A} \succeq \mathrm{E}[\tilde{\mathbf{a}}\tilde{\mathbf{a}}^T] \tag{1}$$

The statistical properties associated with a mean vector and its associated covariance upper bound are completely defined by the definion of consistency: all that can be said about a mean and covariance pair, $(\mathbf{a}, \mathbf{A})$, is that the covariance matrix, $\mathbf{A}$), is greater than the expected squared error in its associated mean, $\mathbf{a}$).

Having established a rigorous definition of what constitutes a *consistent* or *conservative* estimate, it is possible to certify the performance of the Kalman filter and Covariance Intersection.

2.1 The Kalman Filter

Given two mean and covariance estimates $(\mathbf{a}, \mathbf{A})$ and $(\mathbf{b}, \mathbf{B})$, the data fusion problem of interest in this paper consists of determining a fused estimate $(\mathbf{c}, \mathbf{C})$ that is guaranteed to be consistent and summarizes the information in the two estimates with error (in terms of the size of $\mathbf{C}$) that is less than or equal to that of either estimate. If the two estimates are consistent and presumed to be statistically independent, then a joint estimate can be constructed as:

$$\left(\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \right). \tag{2}$$

Letting $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$ denote the errors in the respective mean estimates, the key property of the joint covariance estimate is that it satisfies:

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \succeq \begin{bmatrix} \mathrm{E}[\tilde{\mathbf{a}}\tilde{\mathbf{a}}^T] & \mathbf{0} \\ \mathbf{0} & \mathrm{E}[\tilde{\mathbf{b}}\tilde{\mathbf{b}}^T] \end{bmatrix}, \tag{3}$$

where the RHS matrix represents the true but unknown joint error covariance, which has zero cross covariance, $\tilde{\mathbf{a}}\tilde{\mathbf{b}}^T = \mathbf{0}$, due to the assumption of statistical independence. The estimated joint covariance is a conservative estimate of the true joint covariance because in practice $\mathbf{A} \succeq \mathrm{E}[\tilde{\mathbf{a}}\tilde{\mathbf{a}}^T]$ and $\mathbf{B} \succeq \mathrm{E}[\tilde{\mathbf{b}}\tilde{\mathbf{b}}^T]$. The latter inequalities hold by design in that intentional efforts are made to ensure that estimate error covariances do not underestimate the actual squared errors associated with sensor and kinematic models.

Given a consistent joint covariance for two given n-dimensional estimates, the Kalman filter defines the optimal linear projection of the $2n$-dimensional joint estimate back to the n-dimensional state space of interest. The result of the Kalman projection is a mean and covariance estimate $(\mathbf{c}, \mathbf{C})$ that represents the optimal fusion of the two given mean and covariance estimates. In fact, if there is no additional information

available (e.g., distribution information), then the Kalman fusion estimate is optimal according to virtually any error criteria [7].

In the case of statistically independent estimates $(\mathbf{a}_1, \mathbf{A}_1), (\mathbf{a}_2, \mathbf{A}_2), ..., (\mathbf{a}_m, \mathbf{A}_m)$, the Kalman fusion equations have a particularly simple form [7]:

$$\begin{aligned} \mathbf{C} &= \left(\mathbf{A}_1^{-1} + \mathbf{A}_2^{-1} + ... + \mathbf{A}_m^{-1}\right)^{-1} \\ \mathbf{c} &= \mathbf{C}\left(\mathbf{A}_1^{-1}\mathbf{a}_1 + \mathbf{A}_2^{-1}\mathbf{a}_2 + ... + \mathbf{A}_m^{-1}\mathbf{a}_m\right). \end{aligned} \tag{4}$$

If its underlying assumptions hold (i.e., consistency and independence), then the above Kalman equations ensure that the fused estimate $(\mathbf{c}, \mathbf{C})$ is consistent, and $\mathbf{C} \preceq \mathbf{A}_i \ \forall i, 1 \leq i \leq m$. However, any presumption of statistical independence in practical data fusion contexts should be carefully considered. Specifically, virtually any sensor is subject to time-correlated errors induced by the particular conditions of its use (e.g., changes in temperature, platform vibrations, relative humidity), and errors associated with the nonlinear transformation of its measurements (e.g., from local spherical coordinates to a global coordinate frame) are deterministic and therefore non-independent.

If estimates $(\mathbf{a}_1, \mathbf{A}_1), (\mathbf{a}_2, \mathbf{A}_2), ..., (\mathbf{a}_m, \mathbf{A}_m)$ are each consistent, but not completely independent, then it is possible for the Kalman fused estimate $(\mathbf{c}, \mathbf{C})$ to be *inconsistent*. In fact, if $\mathbf{A}_i = \tilde{\mathbf{a}}_i\tilde{\mathbf{a}}_i^T \ \forall i, 1 \leq i \leq m$, then *any* degree of correlation guarantees inconsistency, i.e., $\mathbf{C} \not\succeq \mathrm{E}[\tilde{\mathbf{c}}\tilde{\mathbf{c}}^T]$. The key point is that a Kalman fused estimate is not guaranteed to be consistent even if each of its given estimates are consistent. The reason why the Kalman filter fails is because although two given estimates $(\mathbf{a}, \mathbf{A})$ and $(\mathbf{b}, \mathbf{B})$ may be individually consistent, the implicit joint covariance may not be if independence is assumed when the cross covariance between the estimates is $\mathbf{X} = \mathrm{E}[\tilde{\mathbf{a}}\tilde{\mathbf{b}}^T] \neq \mathbf{0}$. Specifically:

$$\begin{bmatrix} \mathrm{E}[\tilde{\mathbf{a}}\tilde{\mathbf{a}}^T] & \mathbf{0} \\ \mathbf{0} & \mathrm{E}[\tilde{\mathbf{b}}\tilde{\mathbf{b}}^T] \end{bmatrix} \not\succeq \begin{bmatrix} \mathrm{E}[\tilde{\mathbf{a}}\tilde{\mathbf{a}}^T] & \mathbf{X} \\ \mathbf{X}^T & \mathrm{E}[\tilde{\mathbf{b}}\tilde{\mathbf{b}}^T] \end{bmatrix}. \tag{5}$$

In other words, the Kalman filter only fails to produce a consistent fused estimate when the implicit *joint* estimate is inconsistent. Although the equations are very simple and elegant for independent estimates, the Kalman filter is also defined generally for any consistent joint covariance with $\mathbf{X} \neq \mathbf{0}$ [7]. Therefore, the Kalman filter is guaranteed to produce consistent estimates as long as the given estimates are consistent and their cross covariance is known [3]. Unfortunately, this poses significant challenges. The first challenge is that the cross covariance information must in principle be determined *exactly*. This can be seen by examining the difference between two joint covariance matrices with different cross terms:

$$\begin{bmatrix} \mathbf{A} & \mathbf{X} \\ \mathbf{X}^T & \mathbf{B} \end{bmatrix} - \begin{bmatrix} \mathbf{A} & \mathbf{Y} \\ \mathbf{Y}^T & \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{X}-\mathbf{Y} \\ (\mathbf{X}-\mathbf{Y})^T & \mathbf{0} \end{bmatrix}. \tag{6}$$

The difference matrix is not PSD for any case in which $\mathbf{X} \neq \mathbf{Y}$. The need for absolutely perfect cross covariance information presents difficulties when estimates are

the products of nonlinear operations (e.g., coordinate transformations, kinematic time projections, human-derived estimates) because the error processes are not perfectly modeled. For example, the same approximate nonlinear transformation equations may be applied to convert different radar observations of an object to a common coordinate frame, so the errors committed are clearly not independent, but it may not be possible to determine exact cross covariances.

Kalman's original derivation of his eponymous filter was based on orthogonal projection theory, and the fact that there exists a simple Bayesian interpretation of the result when error distributions are Gaussian was presented only as an interesting special case [5]. However, many subsequent references to the Kalman filter incorrectly suggest that the Kalman filter *requires* assumptions of Gaussianity[1]. It turns out that the assumption of estimate independence is actually the only problematic assumption because it typically cannot be guaranteed in practice. Relaxing the independence assumption leads to the more general fusion equations of the Covariance Intersection (CI) method which has proven invaluable for a wide variety of practical applications[2, 4].

3 Covariance Intersection (CI)

The general mean and covariance data fusion problem can be formulated in terms of the joint covariance structure that implicitly exists between a given pair of estimates $(\mathbf{a}, \mathbf{A})$ and $(\mathbf{b}, \mathbf{B})$:

$$\begin{bmatrix} \mathbf{A} & \mathbf{X} \\ \mathbf{X}^T & \mathbf{B} \end{bmatrix} \tag{7}$$

where $\mathbf{X}$ represents the actual, but unknown, cross covariance between the two estimates. If $\mathbf{X}$ were known, then it would be possible to apply more general formulations of the Kalman filter equations to produce an optimal fused estimate. Unfortunately, these generalizations only guarantee consistency if the cross covariance is known *exactly*, i.e., it cannot be conservatively approximated in any way analogous to the way conservative covariance estimates are used.

Without knowledge of $\mathbf{X}$, the only way to ensure consistency in the application of the Kalman filter is to identify a joint covariance that is guaranteed to be consistent based on the information available. In the present context, therefore, a joint covariance

[1] In fact, one commonly-cited motivation for investigating the application of neural networks and fuzzy logic is the claim that the Kalman filter imposes restrictive Gaussianity assumptions that cannot be satisfied in many applications. The fact is that the use of covariance upper bounds was recognized as necessary when the first Kalman filters were implemented in the late 1960s, and such bounds are incompatible with PDF interpretations. It was shown by Jazwinski in his classic 1970 book that the standard practice of using covariance upper estimates does not undermine the integrity of the Kalman filter[3].

matrix $\mathbf{M}$ must be determined such that:

$$\mathbf{M} \succeq \begin{bmatrix} \mathbf{A} & \mathbf{X} \\ \mathbf{X}^T & \mathbf{B} \end{bmatrix} \tag{8}$$

for every possible cross covariance $\mathbf{X}$. It can be inferred from the symmetry of the unknown cross covariance information (i.e., $\mathbf{M}$ must be consistent for any instantiation $\mathbf{X} = \mathcal{X}$ and for $\mathbf{X} = -\mathcal{X}$) that the off-diagonal blocks of $\mathbf{M}$ should be zero, and its diagonal blocks must be sufficiently larger than $\mathbf{A}$ and $\mathbf{B}$ to account for the effects of all possible degrees of correlation among the error components of the mean estimates $\mathbf{a}$ and $\mathbf{b}$.

It has been shown (appendix 14 of [9]) that a consistent and tight joint covariance $\mathbf{M}$ can be generated by selecting a scalar value ω, $0 \leq \omega \leq 1$ as:

$$\begin{bmatrix} \frac{1}{\omega}\mathbf{A} & \mathbf{0} \\ \mathbf{0} & \frac{1}{(1-\omega)}\mathbf{B} \end{bmatrix} \succeq \begin{bmatrix} \mathbf{A} & \mathbf{X} \\ \mathbf{X}^T & \mathbf{B} \end{bmatrix} \tag{9}$$

where ω is chosen to minimize the size (e.g., determinant) of the covariance produced by the Kalman filter update equations for the estimates $(\mathbf{a}, \frac{1}{\omega}\mathbf{A})$ and $(\mathbf{b}, \frac{1}{1-\omega}\mathbf{B})$. This can be generalized for an arbitrary number of estimates:

$$\begin{bmatrix} \frac{1}{\omega_1}\mathbf{A}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \frac{1}{\omega_2}\mathbf{A}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \frac{1}{\omega_m}\mathbf{A}_m \end{bmatrix} \succeq \begin{bmatrix} \mathbf{A}_1 & \mathbf{X}_{1,2} & \dots & \mathbf{X}_{1,m} \\ \mathbf{X}_{2,2} & \mathbf{A}_2 & \dots & \mathbf{X}_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{X}_{m,1} & \mathbf{X}_{m,2} & \dots & \mathbf{A}_m \end{bmatrix} \tag{10}$$

where $\sum_{i=1}^{m} \omega_i = 1$, and the parameters are determined to minimize the covariance resulting from the fusion of the n estimates. The above inequality provides a general (and optimal) mechanism for obtaining a consistent joint covariance when given only the block diagonals of an unknown joint covariance matrix. The use of appropriate ω-paramterized covariances in the Kalman update equations yields the CI fusion equations.

The general approach of determining consistent joint covariances can also be applied to solve related problems. For example, given estimates $(\mathbf{a}, \mathbf{A})$ and $(\mathbf{b}, \mathbf{B})$ that are correlated to an unknown extent, the covariance of $\mathbf{a} + \mathbf{b}$ can be computed as $\frac{1}{\omega}\mathbf{A} + \frac{1}{1-\omega}\mathbf{B}$. This is referred to as Covariance Addition (CA).

3.1 Covariance Union (CU)

Covariance Intersection addresses the general form of the data fusion problem for mean and covariance estimates, but in practice a different problem can arise before data fusion can even be performed. Specifically, what is to be done if two estimates $(\mathbf{a}, \mathbf{A})$ and $(\mathbf{b}, \mathbf{B})$, purportedly relating to the state of the same real-world object, are

determined to be mutually inconsistent with each other, i.e., the differences between their means is much larger than what can be expected based on their respective error covariance estimates? For example, if two mean position estimates differ by more than a kilometer, but their respective covariances suggest that each mean is accurate to within a meter, then clearly something is wrong.

One mechanism for detecting statistically significant deviations between estimates is to compute Mahalanobis distances [6]. The Mahalanobis distance between estimates $(\mathbf{a}, \mathbf{A})$ and $(\mathbf{b}, \mathbf{B})$ is defined as:

$$(\mathbf{a} - \mathbf{b})^T (\mathbf{A} + \mathbf{B})^{-1} (\mathbf{a} - \mathbf{b}), \tag{11}$$

which is essentially just the squared distance between the means as normalized by the sum of their respective covariances. Intuitively, if the covariances are large, then a large difference between the mean vectors $\mathbf{a}$ and $\mathbf{b}$ is not surprising, so the Mahalanobis distance is small. However, if the covariances are very small, then even small differences between the means may yield a large Mahalanobis distance. A large Mahalanobis distance may tend to indicate that the estimates are not consistent with each other, but a user-defined threshold is required to define what constitutes an acceptable deviation[2]. When the threshold is exceeded, the estimates are regarded as being contradictory and some kind of action must be taken. Resolving such inconsistencies among estimates is sometimes referred to as *deconfliction* [8].

The Covariance Intersection method guarantees consistency as long as the estimates to be fused are each consistent. In the deconfliction problem it is only known that one of the estimates, either $(\mathbf{a}, \mathbf{A})$ or $(\mathbf{b}, \mathbf{B})$, is a consistent estimate of the state of the object of interest. Because it is not generally possible to know which estimate is spurious, the only way to rigorously combine the estimates is to form a unioned estimate, $(\mathbf{u}, \mathbf{U})$, that is guaranteed to be consistent with respect to *both* of the two estimates. Such a unioned estimate can be constructed[10] by computing a mean vector $\mathbf{u}$ and covariance matrix $\mathbf{U}$ such that:

$$\begin{aligned} \mathbf{U} &\succeq \mathbf{A} + (\mathbf{u} - \mathbf{a})(\mathbf{u} - \mathbf{a})^T \\ \mathbf{U} &\succeq \mathbf{B} + (\mathbf{u} - \mathbf{b})(\mathbf{u} - \mathbf{b})^T \end{aligned} \tag{12}$$

where some measure of the size of $\mathbf{U}$, e.g., determinant, is minimized. This Covariance Union (CU) of the two estimates can be subsequently fused with other consistent estimates using CI. The CU equations generalize directly for the case of $m > 2$ two

[2]It must be emphasized that the use of a threshold on Mahalanobis distance is not the only possible mechanism for identifying potentially spurious estimates, but *some* user-defined mechanism is required. Otherwise there is no way to distinguish fault conditions from low probability events. In other words, models for fault conditions are inherently application-specific.

estimates:

$$\begin{aligned} \mathbf{U} &\succeq \mathbf{A}_1 + (\mathbf{u}-\mathbf{a}_1)(\mathbf{u}-\mathbf{a}_1)^T \\ \mathbf{U} &\succeq \mathbf{A}_2 + (\mathbf{u}-\mathbf{a}_2)(\mathbf{u}-\mathbf{a}_2)^T \\ &\vdots \\ \mathbf{U} &\succeq \mathbf{A}_m + (\mathbf{u}-\mathbf{a}_m)(\mathbf{u}-\mathbf{a}_m)^T \end{aligned} \tag{13}$$

Intuitively, the CU equations simply say that if the estimate $(\mathbf{a}, \mathbf{A})$ is consistent, then the translation of the vector $\mathbf{a}$ to $\mathbf{u}$ will require its covariance to be enlarged by the addition of a matrix at least as large as the outer product of $(\mathbf{u}-\mathbf{a})$ in order to be consistent. The same reasoning applies if the estimate $(\mathbf{b}, \mathbf{B})$ is consistent. Covariance Union therefore determines the mean vector $\mathbf{u}$ having the smallest covariance $\mathbf{U}$ that is large enough to guarantee consistency regardless of which of the two given estimates is consistent. The resulting covariance may be significantly larger than either of the given covariances, but this is an accurate reflection of the actual uncertainty that exists due to the conflict between the two estimates. The key fact is that the CU estimate satisfies the definition of consistency.

As a simple example of a CU construction, consider two estimates $(\mathbf{a}, \mathbf{A})$ and $(\mathbf{b}, \mathbf{B})$ of the location of an object observed from two nodes in a network. The estimate from the first node places the mean position at $\mathbf{a} = [0, 0]^T$, and the second node places it at $\mathbf{b} = [4, 4]^T$, and each has an error covariance equal to the identity matrix $\mathbf{I}$. If it is determined that the two estimates are statistically inconsistent with each other, thus implying that one of the estimates is not a consistent estimate of the object's location, then deconfliction must be performed. The optimal CU deconflicted estimate is:

$$\mathbf{u} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \qquad \mathbf{U} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \tag{14}$$

It is straightforward to verify that this estimate $(\mathbf{u}, \mathbf{U})$ is in fact consistent with respect to either/both of the estimates $(\mathbf{a}, \mathbf{A})$ and $(\mathbf{b}, \mathbf{B})$. If $(\mathbf{a}, \mathbf{A})$ is a consistent estimate of the target's state, then the covariance $\mathbf{U}$ for mean $\mathbf{u}$ must be greater than or equal to $\mathbf{A} + (\mathbf{u}-\mathbf{a})(\mathbf{u}-\mathbf{a})^T$, which it is. It can be verified that the estimate $(\mathbf{u}, \mathbf{U})$ is similarly consistent with respect to the estimate $(\mathbf{b}, \mathbf{B})$. Therefore, if either of the two estimates represents a consistent estimate of the state of the object, then the CU estimate is also consistent.

4 General Covariance Union (GCU)

Consistency of the CU equations rests on an implicit assumption that the estimates to be combined are not correlated. Specifically, the CU inequalities for the estimate

$(\mathbf{u}, \mathbf{U})$:

$$\begin{aligned} \mathbf{U} &\succeq \mathbf{A}_1 + (\mathbf{u} - \mathbf{a}_1)(\mathbf{u} - \mathbf{a}_1)^T \\ \mathbf{U} &\succeq \mathbf{A}_2 + (\mathbf{u} - \mathbf{a}_2)(\mathbf{u} - \mathbf{a}_2)^T \\ &\vdots \\ \mathbf{U} &\succeq \mathbf{A}_m + (\mathbf{u} - \mathbf{a}_m)(\mathbf{u} - \mathbf{a}_m)^T \end{aligned} \tag{15}$$

technically hold only if the errors associated with $\mathbf{A}_i$ are uncorrelated with $\mathbf{a}_i$. This is true for most types of common process models, but not for certain kinds of recursive control and colored noise models.

For example, suppose that we are given a $1D$ estimate $(0, 0)$, i.e., zero mean and zero variance. The underpinning assumptions of CU imply that if the mean is translated to 1, the new consistent estimate must have variance $0+1^2 = 1$. If instead the mean is translated to 2, the variance must be $0+2^2 = 4$. However, suppose a 2-step translation is performed: First the estimate is translated 1 unit to produce $(0+1, 0+1^2) = (1, 1)$, then that estimate is translated 1 more unit. According to CU the result should be $(1 + 1, 1 + 1^2) = (2, 2)$, but this is clearly incorrect because translating the estimate $(0, 0)$ by 2 units must produce $(2, 0+2^2) = (2, 4)$ to ensure consistency. The problem, of course, is that the translations in the sequence of steps are not independent.

More generally, given a mean and covariance estimate $(\mathbf{a}, \mathbf{A})$, the covariance of $\mathbf{a}+\mathbf{x}$ is equal to $\mathbf{A}+\mathbf{x}\mathbf{x}^T$ if and only if the estimate error, $\tilde{\mathbf{a}}$, is independent of the translation $\mathbf{x}$. If they are correlated to an unknown extent, then the covariance of $\mathbf{a}+\mathbf{x}$ must be formulated using Covariance Addition (CA) as $\frac{1}{\omega}\mathbf{A} + \frac{1}{1-\omega}\mathbf{x}\mathbf{x}^T$. Applying CA to the CU inequalities (15) gives the generalized GCU formulation:

$$\begin{aligned} \mathbf{U} &\succeq \frac{1}{\omega_1}\mathbf{A}_1 + \frac{1}{1-\omega_1}(\mathbf{u} - \mathbf{a}_1)(\mathbf{u} - \mathbf{a}_1)^T \\ \mathbf{U} &\succeq \frac{1}{\omega_2}\mathbf{A}_2 + \frac{1}{1-\omega_2}(\mathbf{u} - \mathbf{a}_2)(\mathbf{u} - \mathbf{a}_2)^T \\ &\vdots \\ \mathbf{U} &\succeq \frac{1}{\omega_m}\mathbf{A}_m + \frac{1}{1-\omega_m}(\mathbf{u} - \mathbf{a}_m)(\mathbf{u} - \mathbf{a}_m)^T \end{aligned} \tag{16}$$

where the optimization problem now involves n free parameters ($0 \leq \omega_i \leq 1$), like CI, but with the semidefinite inequalities of standard CU. This appears to represent a daunting challenge to solve efficiently. However, plots of solutions for low-dimensional problems provide a critical insight. Specifically, Figure 1 depicts the 1-sigma solution contour that minimally encloses the 1-sigma contours of the estimates that are unioned. This suggests – but of course does not prove – that the GCU solution is equivalent to the corresponding MEE solution obtained by interpreting the covariances as ellipsoids defined by their 1-sigma contours. That they are in fact equivalent is established in the following section.

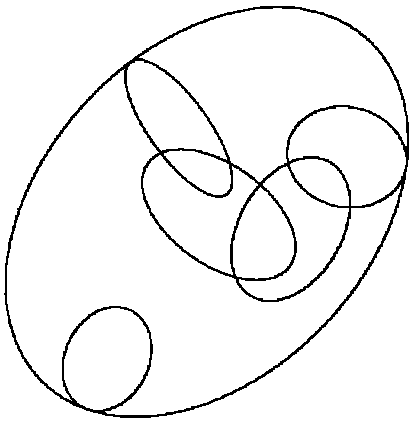

Figure 1: A typical General Covariance Union (GCU) optimization problem. The five small covariance ellipses represent the five estimates to be unioned. Surprisingly, they seem to be circumscribed by the minimal-determinant GCU solution.

5 The Equivalence of GCU and MEE

This section provides a formal proof of the equivalence of MEE and GCU[3] Our treatment of ellipsoid enclosure follows the discussion in §3.7.1 of [1], so several of its equations are repeated here for easy reference.

In [1], the enclosing and enclosed ellipsoids are represented via parameter triples $\{A_i, b_i, c_i\}$ where index 0 refers to the enclosing ellipsoid:

$$\{x \in R^n \mid x^T A_i x + 2x^T b_i + c_i \leq 0\}, \quad i = 0 \ldots m, \tag{17}$$

$$A_i = A_i^T \succ 0 \quad \text{and} \quad b_i^T A_i^{-1} b_i - c_i > 0$$

The triples $\{A_i, b_i, c_i\}$ are homogenous so the authors normalize the enclosing ellipsoid's parameters via $c_0 = b_0^T A_0^{-1} b_0 - 1$ and then use the S-lemma to represent the ellipsoid-enclosure constraint as a matrix inequality:

$$\begin{bmatrix} A_0 & b_0 \\ b_0^T & b_0^T A_0^{-1} b_0 - 1 \end{bmatrix} - \tau_i \begin{bmatrix} A_i & b_i \\ b_i^T & c_i \end{bmatrix} \preceq 0, \quad \tau_i \geq 0, \quad i = 1 \ldots m. \tag{18}$$

That inequality is nonlinear so they use a Schur complement argument to expand it:

$$\begin{bmatrix} A_0 & b_0 & 0 \\ b_0^T & -1 & b_0^T \\ 0 & b_0 & -A_0 \end{bmatrix} - \tau_i \begin{bmatrix} A_i & b_i & 0 \\ b_i^T & c_i & 0^T \\ 0 & 0 & 0 \end{bmatrix} \preceq 0, \quad \tau_i \geq 0, \quad i = 1 \ldots m. \tag{19}$$

[3]To make the exposition concise our proof assumes that the covariance matrices are nonsingular, but the equivalence holds generally.

The minimum-determinant ellipsoid enclosure can then be posed as a Maxdet [11] optimization:

$$\begin{array}{ll} \text{minimize} & \log\det A_0^{-1} \\ \text{subject to} & A_0 \succ 0, \quad \tau_i \geq 0, \quad i = 1 \ldots m. \\ & \left[\begin{array}{ccc} A_0 & b_0 & 0 \\ b_0^T & -1 & b_0^T \\ 0 & b_0 & -A_0 \end{array}\right] - \tau_i \left[\begin{array}{ccc} A_i & b_i & 0 \\ b_i^T & c_i & 0 \\ 0 & 0 & 0 \end{array}\right] \preceq 0, \quad i = 1 \ldots m. \end{array} \tag{20}$$

5.1 Covariance ellipsoid enclosure as a Maxdet

In this section, we translate the ellipse-enclosure formulation (20) from the original $\{A, b, c\}$ triples to covariance/mean estimation parameters (u, U).

The covariance ellipsoid $\mathcal{E}(u, U)$ for an estimate at mean u with covariance matrix U is defined as:

$$\begin{aligned} \mathcal{E}(u,U) &= \{\, x \in R^n \mid (x-u)^T U^{-1}(x-u) \leq 1\} \\ &= \{\, x \in R^n \mid x^T U^{-1} x + 2x^T(-U^{-1}u) + (u^T U^{-1} u - 1) \leq 0\} \end{aligned} \tag{21}$$

Comparison of (17) and (21) reveals the following correspondences:

$$A \Leftrightarrow U^{-1}, \quad b \Leftrightarrow -U^{-1}u, \quad c \Leftrightarrow u^T U^{-1} u - 1 \tag{22}$$

Apply (22) to the nonlinear matrix inequality (18) - we will use the result in §5.2. The result is a GCU equation so A_i now denotes a constraint covariance, and a_i its mean:

$$\left[\begin{array}{cc} U^{-1} & -U^{-1}u \\ (-U^{-1}u)^T & u^T U^{-1} u - 1 \end{array}\right] - \tau_i \left[\begin{array}{cc} A_i^{-1} & -A_i^{-1}a_i \\ (-A_i^{-1}a_i)^T & a_i^T A_i^{-1} a_i - 1 \end{array}\right] \preceq 0, \quad \tau_i \geq 0, \quad i = 1 \ldots m. \tag{23}$$

Returning to the covariance ellipsoid Maxdet, apply (22) to (20):

$$\begin{array}{ll} \text{minimize} & \log\det U \\ \text{subject to} & U \succ 0, \quad \tau_i \geq 0, \quad i = 1 \ldots m, \end{array} \tag{24}$$

$$\left[\begin{array}{ccc} U^{-1} & -U^{-1}u & 0 \\ (-U^{-1}u)^T & -1 & (-U^{-1}u)^T \\ 0 & -U^{-1}u & -U^{-1} \end{array}\right] - \tau_i \left[\begin{array}{ccc} A_i^{-1} & -A_i^{-1}a_i & 0 \\ (-A_i^{-1}a_i)^T & a_i^T A_i^{-1} a_i - 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \preceq 0, \quad i = 1 \ldots m.$$

The matrix inequalities in (24) are nonlinear, due to the (quadratic) $U^{-1}u$ terms. Linearize them with change of variables $v = -U^{-1}u$. Also, let $W = U^{-1}$ to obtain a more standardized Maxdet formulation:

$$\begin{array}{ll} \text{minimize} & \log\det W^{-1} \\ \text{subject to} & W \succ 0, \quad \tau_i \geq 0, \quad i = 1 \ldots m, \\ & \left[\begin{array}{ccc} W & v & 0 \\ v^T & -1 & v^T \\ 0 & v & -W \end{array}\right] - \tau_i \left[\begin{array}{ccc} A_i^{-1} & -A_i^{-1}a_i & 0 \\ (-A_i^{-1}a_i)^T & a_i^T A_i^{-1} a_i - 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \preceq 0, \quad i = 1 \ldots m. \end{array} \tag{25}$$

After solving (25) for v and W, the optimal (u, U) are recovered as $U = W^{-1}$ and $u = -Uv$.

5.2 Enclosure of a single ellipsoid

Consider a GCU constraint from (16) for a single estimate (a, A) where A is full rank:

$$U \succeq \frac{1}{\omega}A + \frac{1}{1-\omega}(u-a)(u-a)^T, \quad \omega \in [0,1] \tag{26}$$

We shall prove that (26) defines the set of all estimates (u, U) whose covariance ellipsoid $\mathcal{E}(u, U)$ encloses $\mathcal{E}(a, A)$. It then follows that the combined GCU constraints (16) define the set of all covariance ellipsoids which enclose all the $\mathcal{E}(a_k, A_k)$ so the solution to the (25) Maxdet formulation is a minimum-determinant solution to (16).

Note that (26) only depends on the difference between u and a rather than their absolute values, and must hold for any arbitrary value of u. Therefore we may, without loss of generality, assume that $a = 0$ (simple coordinate shift) which simplifies (26) to:

$$U \succeq \frac{1}{\omega}A + \frac{1}{1-\omega}uu^T, \quad \omega \in [0,1] \tag{27}$$

A similar argument can be applied to a single instance from the ellipsoid-enclosure inequalities (23): it must hold for arbitrary values of u, and ellipsoid enclosure is a geometric property unaffected by coordinate shifts. So we may again assume that $a = 0$, which simplifies (23) to:

$$\begin{bmatrix} U^{-1} & -U^{-1}u \\ (-U^{-1}u)^T & u^TU^{-1}u - 1 \end{bmatrix} - \tau \begin{bmatrix} A^{-1} & 0 \\ 0^T & -1 \end{bmatrix} \preceq 0, \quad \tau \geq 0 \tag{28}$$

Equation (28) implies the following scalar inequality for the lower right-hand main diagonal entry: $u^TU^{-1}u - 1 + \tau \leq 0$. Since $u^TU^{-1}u \geq 0$ we must have $\tau \leq 1$. Make that restriction explicit, to bring τ in line with the equation (27) ω parameter:

$$\begin{bmatrix} U^{-1} & -U^{-1}u \\ (-U^{-1}u)^T & u^TU^{-1}u - 1 \end{bmatrix} - \tau \begin{bmatrix} A^{-1} & 0 \\ 0^T & -1 \end{bmatrix} \preceq 0, \quad \tau \in [0,1] \tag{29}$$

We will prove equivalence by demonstrating that (27) and (29) are mutual inverses. But first we must consider the singular points. For (27) they are at $\omega = 0$ and $\omega = 1$.

As ω approaches 0, U becomes unbounded and therefore $\mathcal{E}(u, U)$ encloses any (bounded) covariance ellipsoid $\mathcal{E}(0, A)$. The behavior for $\omega = 1$ depends on the value of u: if $u = 0$ then the ellipsoids are concentric and (27) reduces to the expected enclosure requirement $U \succeq A$. If $u \neq 0$ then the right-hand side approaches the sum of A and an unbounded rank-1 adjustment in the direction of u. The result is an elliptic cylinder - with radial axis through 0 and u - which tightly encloses $\mathcal{E}(0, A)$. So both singular points lead to enclosure of $\mathcal{E}(0, A)$.

Now that the singular points have been considered, we can simply invert (27): move the rank-1 term to the left-hand side and apply the Sherman-Morrison formula. The result simplifies to:

$$U^{-1} + \frac{(U^{-1}u)(U^{-1}u)^T}{1 - \omega - u^T U^{-1} u} \preceq \omega A^{-1} \tag{30}$$

The rank-1 term in (30) is bounded and positive semidefinite, since it is related to the bounded and positive semidefinite rank-1 term $\frac{1}{1-\omega}\, uu^T$ from (27). Therefore, $1 - \omega - u^T U^{-1} u > 0$ so we can apply a Schur complement:

$$\begin{bmatrix} \omega A^{-1} - U^{-1} & U^{-1}u \\ (U^{-1}u)^T & 1 - \omega - u^T U^{-1} u \end{bmatrix} \succeq 0 \tag{31}$$

Finally, rearrange terms:

$$\begin{bmatrix} U^{-1} & -U^{-1}u \\ (-U^{-1}u)^T & u^T U^{-1} u - 1 \end{bmatrix} - \omega \begin{bmatrix} A^{-1} & 0 \\ 0^T & -1 \end{bmatrix} \preceq 0, \quad \omega \in (0, 1) \tag{32}$$

Equation (32) has the same form as (29) (except for the singular points $\omega = 0$ and $\omega = 1$ which were considered earlier). This completes the proof.

6 Discussion

In this paper we have established a connection between covariance-based and bounded-region models of uncertainty. In particular, we have shown that GCU and MEE have equivalent solutions. Thus, techniques for solving one problem can be directly applied to solve instances of the other. This equivalence is also suggestive of a potentially more general framework that subsumes the covariance and bounded-region approaches.

Acknowledgements: This work was supported by grants from the Office of Naval Research and the Naval Research Laboratory.

References

[1] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. *Linear matrix inequalities in system and control theory*, volume 15 of *SIAM studies in applied mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1994.

[2] J. Ferreira and J. Waldmann. Covariance intersection-based sensor fusion for sounding rocket tracking and impact area prediction. *Control Engineering Practice*, 15(4):389–409, April 2007.

[3] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

[4] S. J. Julier and J. K. Uhlmann. Using covariance intersection for SLAM. *Robotics and Autonomous Systems*, 55(1):3–20, January 2007.

[5] R.E. Kalman. A new approach to linear filtering and prediction problems. *ASME J. Basic Eng.*, 82:34–45, 1960.

[6] P. Mahalanobis. On the generalized distance in statistics. *Proc. Natl. Inst. Science*, 12:49–55, 1936.

[7] P. S. Maybeck. *Stochastic Models, Estimation, and Control*, volume 1. Academic Press, 1979.

[8] Ranjeev Mittu and Frank Segaria. Common operational picture (COP) and common tactical picture (CTP) management via a consistent networked information stream (CNIS). *Proceedings of the Command and Control Research and Technology Symposium*, 2000.

[9] J. K. Uhlmann. *Dynamic Map Building and Localization for Autonomous Vehicles*. PhD thesis, University of Oxford, 1995.

[10] Jeffrey Uhlmann. Covariance consistency methods for fault-tolerant distributed data fusion. *Information Fusion*, 4(3):201–215, Sept 2003.

[11] S.-P. Wu, L. Vandenberghe, and S. Boyd. Maxdet: Software for determinant maximization problems, alpha version. *Stanford University*, April 1996.