AN ELLIPSOIDAL APPROXIMATION TO THE HADAMARD PRODUCT OF ELLIPSOIDS

R. Lorenz, S. Boyd

Stanford University
Department of Electrical Engineering
Stanford, CA, 94305

ABSTRACT

This paper introduces a computationally efficent outer approximation to the Hadamard, *i.e.*, element-wise, product of two ellipsoids. This element-wise product corresponds to multiplicative uncertainties, which arrive commonly in practice. We consider the case where both ellipsoids describe real numbers and the case in which the ellipsoids correspond to the direct-sum representation of complex numbers.

1. INTRODUCTION

Uncertainties that are multiplicative in nature arise often in practice; consider, for example, beamforming using the amplified output of an antenna array in which the gains and phases of the electronics paths that are not precisely known. This is depicted schematically in Figure 1. The gains may be known to have some a-priori uncertainty; in other applications, these quantities are estimated in terms of a mean vector and covariance matrix. In both cases, this uncertainty is well described by an ellipsoid.

Assume that the range of possible values of the array manifold is described by an ellipsoid $\mathcal{E}_1 = \{Au + b \mid \|u\| \leq 1\}$. Similarly assume the multiplicative uncertainties lie within a second ellipsoid $\mathcal{E}_2 = \{Cv + d \mid \|v\| \leq 1\}$. The set of possible values of the array manifold in the presence of multiplicative uncertainties is described by the numerical range of the Hadamard *i.e.* element-wise product of $\mathcal{E}_1$ and $\mathcal{E}_2$.

1.1. Ellipsoid descriptions

A n-dimensional ellipsoid can be defined as the image of a n-dimensional Euclidean ball under an affine mapping from $\mathbf{R}^n$ to $\mathbf{R}^n$, *i.e.*,

$$\mathcal{E} = \{Au + c \mid \|u\| \leq 1\}, \tag{1}$$

where $A \in \mathbf{R}^{n \times n}$ and $c \in \mathbf{R}^n$. The set $\mathcal{E}$ describes an ellipsoid whose center is c and whose *principal semiaxes* are the unit-norm left singular vectors of A scaled by the corresponding singular values. We say that an ellipsoid is *flat* if this mapping is not injective, *i.e.*, one-to-one. Flat ellipsoids can be described by (1) in the proper affine subspaces of $\mathbf{R}^n$. In this case, $A \in \mathbf{R}^{n \times l}$ and $u \in \mathbf{R}^l$. An interpretation of a flat uncertainty ellipsoid is that some linear combinations of the data are known exactly; see [1].

Unless otherwise specified, an ellipsoid in $\mathbf{R}^n$ will be parameterized in terms of its center $c \in \mathbf{R}^n$ and a symmetric non-negative

Research supported in part DARPA (under F33615-99-C-3014), and AFOSR URI award F49620-01-1-0365.

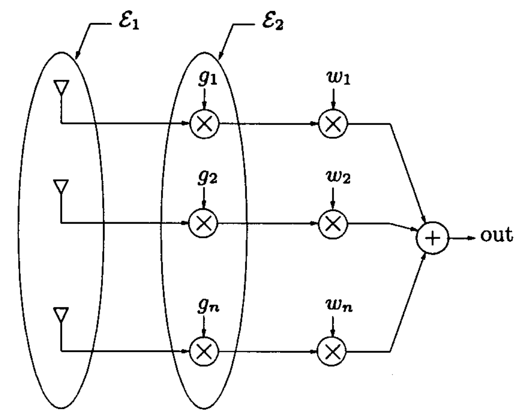

Fig. 1. The possible values of array manifold are contained in ellipsoid $\mathcal{E}_1$; the values of gains are described by ellipsoid $\mathcal{E}_2$. The design variable w need to consider the multiplicative effect of these uncertainties.

definite configuration matrix $Q \in \mathbf{R}^{n \times n}$ as

$$\mathcal{E}(c, Q) = \{Q^{1/2}u + c \mid \|u\| \leq 1\} \tag{2}$$

where $Q^{1/2}$ is any matrix square root satisfying $Q^{1/2}(Q^{1/2})^T = Q$. When Q is full rank, the non-degenerate ellipsoid $\mathcal{E}(c, Q)$ may also be expressed as

$$\mathcal{E}(c, Q) = \{x \mid (x - c)^T Q^{-1} (x - c) \leq 1\}. \tag{3}$$

The first representation (2) is more natural when $\mathcal{E}$ is degenerate or poorly conditioned. Using the second description (3), one may quickly determine whether a point is within the ellipsoid.

There are two common techniques for fitting an ellipsoid to a collection of points: finding an ellipsoid based on the first and second order statistics of the points and finding the minimum volume ellipsoid containing these points.

1.2. Ellipsoid computation using mean and covariance of data

The mean and covariance are calculated in the usual fashion. If the underlying distribution is multivariate normal, the $k - \sigma$ ellipsoid would be expected to contain a fraction of points equal to $1 - \chi^2(k^2, n)$, where n is the dimension of the random variable. It is prudent to examine the relationship between the scaling factor of the ellipsoid and the fraction of points contained therein.

1.3. Minimum volume ellipsoid (MVE)

Let $\mathcal{S} = \{s_1, \cdots, s_m\} \in \mathbf{R}^{2n}$ be a set of samples of possible values of the array manifold $a(\cdot)$. Assume that $\mathcal{S}$ is bounded. In the case of a full rank ellipsoid, the problem of finding the minimum volume ellipsoid containing the convex hull of $\mathcal{S}$ can be expressed as the following semidefinite program (SDP):

$$\begin{array}{ll} \text{minimize} & \log\det F^{-1} \\ \text{subject to} & F = F^T \succ 0 \\ & \|Fs_i - g\| \leq 1, \quad i = 1, \cdots, m. \end{array} \tag{4}$$

See Vandenberghe and Boyd [2] and Wu and Boyd [3]. Equation (4) is a convex problem in variables F and g. For A full rank,

$$\{x \mid \|Fx - g\| \leq 1\} \equiv \{Au + c \mid \|u\| \leq 1\} \tag{5}$$

with $A = F^{-1}$ and $c = F^{-1}g$.

Compared to an ellipsoid based on the first and second order statistics of the data, a minimum volume ellipsoid is robust in the sense that it is guaranteed to cover all the data points used in the description; the MVE is *not robust* to data outliers. The computation of the covering ellipsoid is relatively complex; see Vandenberghe et al. [4].

1.4. The sum of two ellipsoids

Recall that we can parameterize an ellipsoid in $\mathbf{R}^n$ in terms of its center $c \in \mathbf{R}^n$ and a symmetric non-negative definite configuration matrix $Q \in \mathbf{R}^{n\times n}$ as

$$\mathcal{E}(c, Q) = \{Q^{1/2}u + c \mid \|u\| \leq 1\}$$

where $Q^{1/2}$ is any matrix square root satisfying $Q^{1/2}(Q^{1/2})^T = Q$. Let $x \in \mathcal{E}_1 = \mathcal{E}(c_1, Q_1)$ and $y \in \mathcal{E}_2 = \mathcal{E}(c_2, Q_2)$. The range of values of the geometrical (or Minkowski) sum $z = x + y$ is contained in the ellipsoid

$$\mathcal{E} = \mathcal{E}(c_1 + c_2, Q(p)) \tag{6}$$

for all $p > 0$ where

$$Q(p) = (1 + p^{-1})Q_1 + (1 + p)Q_2; \tag{7}$$

see Kurzhanski and Vályi [5]. The value of p is commonly chosen to minimize either the determinant of $Q(p)$ or the trace of $Q(p)$. Minimizing the trace of Q in equation (7) affords two computational advantages over minimizing the determinant. First, computing the optimal value of p can be done with $\mathcal{O}(n)$ operations; minimizing the determinant requires $\mathcal{O}(n^3)$. Second, the minimum trace calculation may be used without worry with degenerate ellipsoids.

An example of the geometrical sum of two ellipses for various values of p is shown in Figure 2.

1.5. Minimum trace

There exists an ellipsoid of minimum trace, *i.e.*, sum of squares of the semiaxes, that contains the sum $\mathcal{E}_1(c_1, Q_q) + \mathcal{E}_2(c_2, Q_2)$; it is described by $\mathcal{E}(c_1 + c_2, Q(p^\star))$, where $Q(p)$ is as in (7),

$$p^\star = \sqrt{\frac{\mathbf{Tr}\, Q_1}{\mathbf{Tr}\, Q_2}}, \tag{8}$$

and $\mathbf{Tr}$ denotes trace. This fact, noted by Kurzhanski and Vályia [5, §2.5], may be verified by direct calculation.

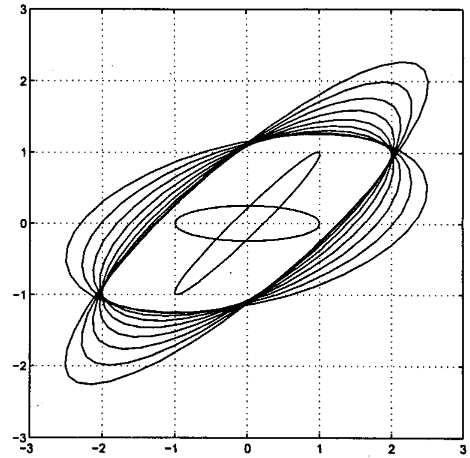

Fig. 2. Outer approximations of the sum of two ellipses (center) for different configuration matrices $Q(p)$

2. AN OUTER APPROXIMATION TO THE HADAMARD PRODUCT OF TWO ELLIPSOIDS

In this section we will develop outer approximations to the Hadamard product of two ellipsoids. In §2.2, we consider the case where both ellipsoids describe real numbers; the case of complex values is considered in §2.3. Prior to this, we will review some basic facts about Hadamard products.

2.1. Preliminaries

The Hadamard product of matrices has considerable structure; the interested reader is referred to Horn and Johnson [6]. The Hadamard product of vectors is the element-wise product of the entries. We denote the Hadamard product of vectors x and y as

$$x \circ y = \begin{bmatrix} x_1y_1 \\ x_2y_2 \\ \vdots \\ x_ny_n \end{bmatrix}.$$

The Hadamard product of two matrices is similarly denoted and also corresponds to the element-wise product. As with other operators, we shall consider the Hadamard product operator $\circ$ to have lower precedence than ordinary matrix multiplication.

Lemma 1 *For any* $x, y \in \mathbf{R}^n$

$$(x \circ y)(x \circ y)^T = (xx^T) \circ (yy^T).$$

Lemma 2 *Let* $x \in \mathcal{E}_x = \{Au \mid \|u\| \leq 1\}$ *and* $y \in \mathcal{E}_y = \{Cv \mid \|v\| \leq 1\}$. *Then the field of values of the Hadamard product* $x \circ y$ *is contained in the ellipsoid*

$$\mathcal{E}_{xy} = \{(AA^T \circ CC^T)^{1/2}w \mid \|w\| \leq 1\}.$$

Lemma 3 *Let* $\mathcal{E}_1 = \{Au \mid \|u\| \leq 1\}$ *and let d be any vector in* $\mathbf{R}^n$. *The Hadamard product of* $\mathcal{E}_1 \circ d$ *is contained in the ellipsoid*

$$\mathcal{E} = \{(AA^T \circ dd^T)^{1/2}w \mid \|w\| \leq 1\}$$

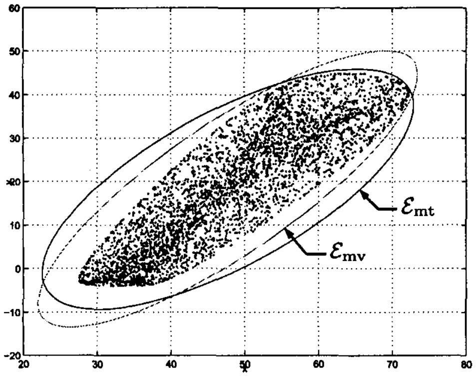

Fig. 3. Samples of the Hadamard product of two ellipsoids. The outer approximations based on the minimum volume and minimum trace metrics are labeled $\mathcal{E}_{\rm mv}$ and $\mathcal{E}_{\rm mt}$.

2.2. Outer approximation

Let $\mathcal{E}_1 = \{Au + b \mid \|u\| \le 1\}$ and $\mathcal{E}_2 = \{Cv + d \mid \|v\| \le 1\}$ be ellipsoids in $\mathbf{R}^n$. Let x and y be n dimensional vectors taken from ellipsoids $\mathcal{E}_1$ and $\mathcal{E}_2$ respectively. Expanding the Hadamard product $x \circ y$, we have

$$\begin{aligned} x \circ y &= b \circ d + Auu^TA^T \circ Cvv^TC^T + \\ & \quad Auu^TA^T \circ dd^T + bb^T \circ Cvv^TC^T. \end{aligned} \tag{9}$$

By Lemmas 2 and 3, the field of values of the Hadamard product

$$x \circ y \in \{(Au + b) \circ (Cv + d) \mid \|u\| \le 1, \|v\| \le 1\}$$

is contained in the geometrical sum of three ellipsoids

$$\begin{aligned} \mathcal{S} &= \mathcal{E}(b \circ d, AA^T \circ CC^T) + \\ & \quad \mathcal{E}(0, AA^T \circ dd^T) + \mathcal{E}(0, bb^T \circ CC^T). \end{aligned} \tag{10}$$

Ignoring the correlations between terms in the above expansion, we find that $\mathcal{S} \subset \mathcal{E}(b \circ d, Q)$, where

$$\begin{aligned} Q &= \left(1 + \tfrac{1}{p_1}\right)\left(1 + \tfrac{1}{p_2}\right) AA^T \circ CC^T + \\ & \quad (1 + p_1)\left(1 + \tfrac{1}{p_2}\right) AA^T \circ dd^T + \\ & \quad (1 + p_1)(1 + p_2) CC^T \circ bb^T \end{aligned} \tag{11}$$

for all $p_1 > 0$ and $p_2 > 0$. The values of p_1 and p_2 may be chosen to minimize the trace or the determinant of Q.

As a numerical example, we consider the Hadamard product of two ellipsoids in $\mathbf{R}^2$. The ellipsoid $\mathcal{E}_1$ is described by

$$A = \begin{bmatrix} -0.6452 & -1.5221 \\ 0.2628 & 2.2284 \end{bmatrix}, \; b = \begin{bmatrix} -5.0115 \\ 1.8832 \end{bmatrix};$$

the parameters of $\mathcal{E}_2$ are

$$C = \begin{bmatrix} -1.0710 & 0.7919 \\ 0.8744 & 0.7776 \end{bmatrix}, \; d = \begin{bmatrix} -9.5254 \\ 9.7264 \end{bmatrix}.$$

Samples of the Hadamard product of $\mathcal{E}_1 \circ \mathcal{E}_2$ are shown in Figure 3 along with the outer approximations based on the minimum volume and minimum trace metrics, $\mathcal{E}_{\rm mv}$ and $\mathcal{E}_{\rm mt}$ respectively.

2.3. The complex case

We now extend the results of §2.2 to the case of complex values. In this section, for reasons of numerical efficiency, we will compute the approximating ellipsoid using the minimum trace metric. As before, we will consider complex numbers to be represented by the direct sum of their real and imaginary components. While it is possible to cover the field of values with a complex ellipsoid in $\mathbf{C}^n$, doing so generally results in a larger ellipsoid than if the direct sum of the real and imaginary components are covered in $\mathbf{R}^{2n}$. Let $x \in \mathbf{R}^{2n}$ and $y \in \mathbf{R}^{2n}$ be the direct sum representations of $\alpha \in \mathbf{C}^n$ and $\beta \in \mathbf{C}^n$ respectively; *i.e.*,

$$x = \begin{bmatrix} \mathbf{Re}\,\alpha \\ \mathbf{Im}\,\alpha \end{bmatrix}, \quad y = \begin{bmatrix} \mathbf{Re}\,\beta \\ \mathbf{Im}\,\beta \end{bmatrix}.$$

We can represent the real and imaginary components of $\gamma = \alpha \circ \beta$ as

$$\begin{aligned} z &= \begin{bmatrix} \mathbf{Re}\,\gamma \\ \mathbf{Im}\,\gamma \end{bmatrix} = \begin{bmatrix} \mathbf{Re}\,\alpha \circ \mathbf{Re}\,\beta + \mathbf{Im}\,\alpha \circ \mathbf{Im}\,\beta \\ \mathbf{Im}\,\alpha \circ \mathbf{Re}\,\beta + \mathbf{Re}\,\alpha \circ \mathbf{Im}\,\beta \end{bmatrix} \\ &= F_1 x \circ F_2 y + F_3 x \circ F_4 y, \end{aligned} \tag{12}$$

where

$$F_1 = \begin{bmatrix} I_n & 0 \\ 0 & I_n \end{bmatrix}, \quad F_2 = \begin{bmatrix} I_n & 0 \\ I_n & 0 \end{bmatrix},$$

and

$$F_3 = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}, \quad F_4 = \begin{bmatrix} 0 & I_n \\ 0 & I_n \end{bmatrix}.$$

Note that multiplications associated with matrices $F_1, \ldots, F_4$ correspond to reordering of the calculations, not general matrix multiplies. Applying (12) to $x \in \mathcal{E}_1 = \{Au + b \mid \|u\| \le 1\}$ and $y \in \mathcal{E}_2 = \{Cv + d \mid \|v\| \le 1\}$ yields:

$$\begin{aligned} z &= F_1 b \circ F_2 d + \\ & \quad F_3 b \circ F_4 d + \\ & \quad F_1 Auu^TA^TF_1^T \circ F_2 Cvv^TC^TF_2^T + \\ & \quad F_1 Auu^TA^TF_1^T \circ F_2 dd^TF_2^T + \\ & \quad F_1 bb^TF_1^T \circ F_2 Cvv^TC^TF_2^T + \\ & \quad F_3 Auu^TA^TF_3^T \circ F_4 Cvv^TC^TF_4^T + \\ & \quad F_3 Auu^TA^TF_3^T \circ F_4 dd^TF_4^T + \\ & \quad F_3 bb^TF_3^T \circ F_4 Cvv^TC^TF_4^T. \end{aligned} \tag{13}$$

The direct sum representation of the field of values of the complex Hadamard product $\alpha \circ \beta$ is contained in the geometrical sum of ellipsoids

$$\begin{aligned} \mathcal{S} &= \mathcal{E}(F_1 b \circ F_2 d, F_1 AA^TF_1^T \circ F_2 CC^TF_2^T) + \\ & \quad \mathcal{E}(F_3 b \circ F_4 d, F_1 AA^TF_1^T \circ F_2 dd^TF_2^T) + \\ & \quad \mathcal{E}(0, F_1 bb^TF_1^T \circ F_2 CC^TF_2^T) + \\ & \quad \mathcal{E}(0, F_3 AA^TF_3^T \circ F_4 CC^TF_4^T) + \\ & \quad \mathcal{E}(0, F_3 AA^TF_3^T \circ F_4 dd^TF_4^T) + \\ & \quad \mathcal{E}(0, F_3 bb^TF_3^T \circ F_4 CC^TF_4^T). \end{aligned} \tag{14}$$

As before, we compute $\mathcal{E}(c, Q) \supseteq \mathcal{S}$ where the center of the covering ellipsoid is given by the sum of the first two terms of (13); the configuration matrix Q is calculated by repeatedly applying (6) and (7) to the remaining terms of (13), where p is chosen according to (8).

2.4. An improved approximation

We now make use of two facts which generally lead to tighter approximations. First, the ellipsoidal outer approximation ignores any correlation between the terms in expansion (13); hence, it is productive to reduce the *number* of these terms. If, for example, $b^T(AA^T)^{-1}b \ll d^T(CC^T)^{-1}d$, *i.e.*, the relative size of the first uncertainty region is much larger than that of the second, by changing coordinates, we may eliminate a significant term from this expansion.

Consider a Given's rotation matrix of the form:

$$T = \begin{bmatrix} \cos\theta_1 & & & \sin\theta_1 & & \\ & \ddots & & & \ddots & \\ & & \cos\theta_n & & & \sin\theta_n \\ -\sin\theta_1 & & & \cos\theta_1 & & \\ & \ddots & & & \ddots & \\ & & -\sin\theta_n & & & \cos\theta_n \end{bmatrix}. \quad (15)$$

The effect of premultiplying a direct sum representation of a complex vector by T is to shift the phase of each of component by the corresponding angle θ_i. It is not surprising, then, that for all T_x and T_y of the form (15) we have

$$\begin{aligned} &T_x^{-1}T_y^{-1}(F_1T_xx \circ F_2T_yy + F_3T_xx \circ F_4T_yy) \\ &= F_1x \circ F_2y + F_3x \circ F_4y, \end{aligned} \quad (16)$$

which does not hold for unitary matrices in general.

We now compute rotation matrices T_b and T_d such that the entries associated with the imaginary components of products T_xb and T_yd respectively are set to zero. In computing T_b, we choose the values of θ in (15) according to $\theta_i = \angle(b(i)+\sqrt{-1}\times b(n+i))$. T_y is similarly computed using the values of d; *i.e.*, $\theta_i = \angle(d(i) + \sqrt{-1} \times d(n+i))$. We change coordinates according to

$$\begin{aligned} A &\leftarrow T_bA \\ b &\leftarrow T_bb \\ C &\leftarrow T_bC \\ d &\leftarrow T_bd \end{aligned}$$

The rotated components associated with the ellipsoid centers have the form

$$T_bb = \begin{bmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad T_dd = \begin{bmatrix} \tilde{d}_1 \\ \vdots \\ \tilde{d}_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (17)$$

zeroing the term $F_3T_bAA^TT_b^TF_3^T \circ (F_4T_ddd^TT_d^TF_4^T)$ in (13). The desired outer approximation is computed as the geometrical sum of outer approximations to the remaining five terms *i.e.*,

$$\begin{aligned} \mathcal{E}(c,Q) \supseteq\ &\mathcal{E}(F_1b \circ F_2d, F_1AA^TF_1^T \circ F_2CC^TF_2^T) + \\ &\mathcal{E}(F_3b \circ F_4d, F_1AA^TF_1^T \circ F_2dd^TF_2^T) + \\ &\mathcal{E}(0, F_1bb^TF_1^T \circ F_2CC^TF_2^T) + \\ &\mathcal{E}(0, F_3AA^TF_3^T \circ F_4CC^TF_4^T) + \\ &\mathcal{E}(0, F_3bb^TF_3^T \circ F_4CC^TF_4^T). \end{aligned} \quad (18)$$

Second, while the Hadamard product is commutative, the outer approximation based on covering the individual terms in the expansion (13) *is* sensitive to ordering; simply interchanging the dyads $\{A, b\}$ and $\{C, d\}$ results in different qualities of approximations. The ellipsoidal approximation associated with this interchanged ordering is given by:

$$\begin{aligned} \mathcal{E}(c,Q) \supseteq\ &\mathcal{E}(F_1d \circ F_2b, F_1CC^TF_1^T \circ F_2AA^TF_2^T) + \\ &\mathcal{E}(F_3d \circ F_4b, F_1CC^TF_1^T \circ F_2bb^TF_2^T) + \\ &\mathcal{E}(0, F_1dd^TF_1^T \circ F_2AA^TF_2^T) + \\ &\mathcal{E}(0, F_3CC^TF_3^T \circ F_4AA^TF_4^T) + \\ &\mathcal{E}(0, F_3dd^TF_3^T \circ F_4AA^TF_4^T). \end{aligned} \quad (19)$$

Since our goal is to find the smallest ellipsoid covering the numerical range of z we compute the trace associated with both orderings and choose the smaller of the two. This determination can be made without computing the minimum trace ellipsoids explicitly, making use of the following fact. Let $\mathcal{E}_0$ be the minimum trace ellipsoid covering $\mathcal{E}_1 + \ldots + \mathcal{E}_p$. The trace of $\mathcal{E}_0$ is given by:

$$\mathbf{Tr}\,\mathcal{E}_0 = \left(\sqrt{\mathbf{Tr}\,\mathcal{E}_1} + \sqrt{\mathbf{Tr}\,\mathcal{E}_2} + \ldots + \sqrt{\mathbf{Tr}\,\mathcal{E}_p}\right)^2,$$

which may be verified by direct calculation. Hence, determining which of (18) and (19) yields the smaller trace can be performed in $\mathcal{O}(n)$ calculations. After making this determination, we perform the remainder of the calculations to compute the desired configuration matrix Q. We then transform Q back to the original coordinates according to:

$$Q \leftarrow (T_b^{-1}T_d^{-1})Q(T_b^{-1}T_d^{-1})^T.$$

3. CONCLUSIONS

Ellipsoidal calculus techniques may be used to efficiently propagate uncertainty ellipsoids in the presence of multiplicative uncertainties. The use of the minimum trace metric allows these computations to be performed efficiently. The ideas extend readily to the case where the ellipsoids describe complex numbers.

4. REFERENCES

[1] A. Ben-Tal and A. Nemirovski, *Convex Optimization In Engineering: Modeling, Analysis, Algorithms*, Technion - Israel Institute of Technology, 1999.

[2] L. Vandenberghe and S. Boyd, "Semidefinite programming," in *Siam Review*, 1995.

[3] S.-P. Wu and S. Boyd, "SDPSOL: a parser/solver for semidefinite programs with matrix structure," in *Advances in Linear Matrix Inequality Methods in Control*, L. El Ghaoui and S.-I. Niculescu, Eds., chapter 4, pp. 79–91. SIAM, 2000.

[4] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM J. on Matrix Analysis and Applications*, vol. 19, no. 2, pp. 499–533, Apr. 1998.

[5] A. Kurzhanski and I. Vályi, *Ellipsoidal Calculus for Estimation and Control*, Systems & Control: Foundations & Applications. Birkhauser, Boston, 1997.

[6] R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[7] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Transactions on Signal Processing*, submitted October, 2001, available at: `http://www.stanford.edu/ boyd/rmvb.html`.