A Tight Bound for the Joint Covariance of Two Random Vectors with Unknown but Constrained Cross–Correlation

Uwe D. Hanebeck*, Kai Briechle*, Joachim Horn†

*Institute of Automatic Control Engineering
Technische Universität München
80290 München, Germany
Uwe.Hanebeck@ieee.org
Kai.Briechle@ei.tum.de

†Siemens AG, Corporate Technology
Information and Communications
81730 München, Germany
Joachim.Horn@ieee.org

Abstract

This paper derives a fundamental result for processing two correlated random vectors with unknown cross–correlation, where constraints on the maximum absolute correlation coefficient are given. A tight upper bound for the joint covariance matrix is derived on the basis of the individual covariances and the correlation constraint. For symmetric constraints, the bounding covariance matrix naturally possesses zero cross covariances, which further increases their usefulness in applications. Performance is demonstrated by recursively propagating a state through a linear dynamical system suffering from stochastic noise correlated with the system state.

1 Introduction

In many applications correlated random vectors have to be processed, which requires their joint statistics to be available [1]. However, the cross covariances may either be too expensive to maintain or are simply not available. Unfortunately, simply neglecting the cross covariances by setting them to zero gives wrong results [2, 5]. Hence, an upper bound for the joint covariance matrix is desired, which is compatible with all possible cross covariances.

For the case that the correlation between the considered random vectors is unconstrained, i.e., the maximum absolute correlation coefficient is less than or equal to one ($|r| \leq 1$), a covariance bound exists [3, 4].

However, the existing covariance bound is too conservative, i.e., is not tight enough, when a constraint of the form $|r| \leq r_{max} < 1$ is available. Hence, the purpose of this paper is to derive a tight bound for the case of constrained correlation.

The problem of bounding two correlated random vectors with a given cross–correlation constraint is formulated in Sec. 2. An appropriate bound is derived in Sec. 3 and then discussed in detail in Sec. 4. The advantage of using the new bound in applications is demonstrated in Sec. 5, where a state is recursively propagated through a linear dynamic system corrupted by correlated noise.

2 Problem Formulation

We are given two random vectors $\underline{x} \in \mathbb{R}^N$, $\underline{y} \in \mathbb{R}^M$ with expected values

$$E\{\underline{x}\} = \hat{\underline{x}} \ , \ E\{\underline{y}\} = \hat{\underline{y}}$$

and individual covariances

$$\text{Cov}\{\underline{x}\} = \mathbf{E}_{xx} \ , \ \text{Cov}\{\underline{y}\} = \mathbf{E}_{yy} \ ,$$

where $\underline{x}$ and $\underline{y}$ are assumed to be correlated. Their cross covariances $\text{Cov}\{\underline{x}, \underline{y}\} = \mathbf{E}_{xy}$ and $\text{Cov}\{\underline{y}, \underline{x}\} = \mathbf{E}_{yx}$, however, are not explicitly known. It is only known that the correlation coefficient r is limited according to

$$|r| \leq r_{max} \ . \tag{1}$$

Hence, a constraint for the cross covariances is given by

$$\mathbf{E}_{yx}\mathbf{E}_{xx}^{-1}\mathbf{E}_{xy} \leq r_{max}^2 \mathbf{E}_{yy} \ , \tag{2}$$

where, in general, for two positive definite matrices $\mathbf{A}$ and $\mathbf{B}$, an expression of the form $\mathbf{A} > \mathbf{B}$ ($\mathbf{A} \geq \mathbf{B}$) is interpreted as $\mathbf{A} - \mathbf{B}$ positive definite (positive semi–definite). By defining the matrix

$$\mathbf{C} = r_{max}^2 \mathbf{E}_{yy} - \mathbf{E}_{yx}\mathbf{E}_{xx}^{-1}\mathbf{E}_{xy} \ ,$$

verification of (2) can be performed by Sylvester's criterion according to

$$\det\left(\mathbf{C}(1:i, 1:i)\right) \geq 0$$

for $i = 1, \ldots, M$.

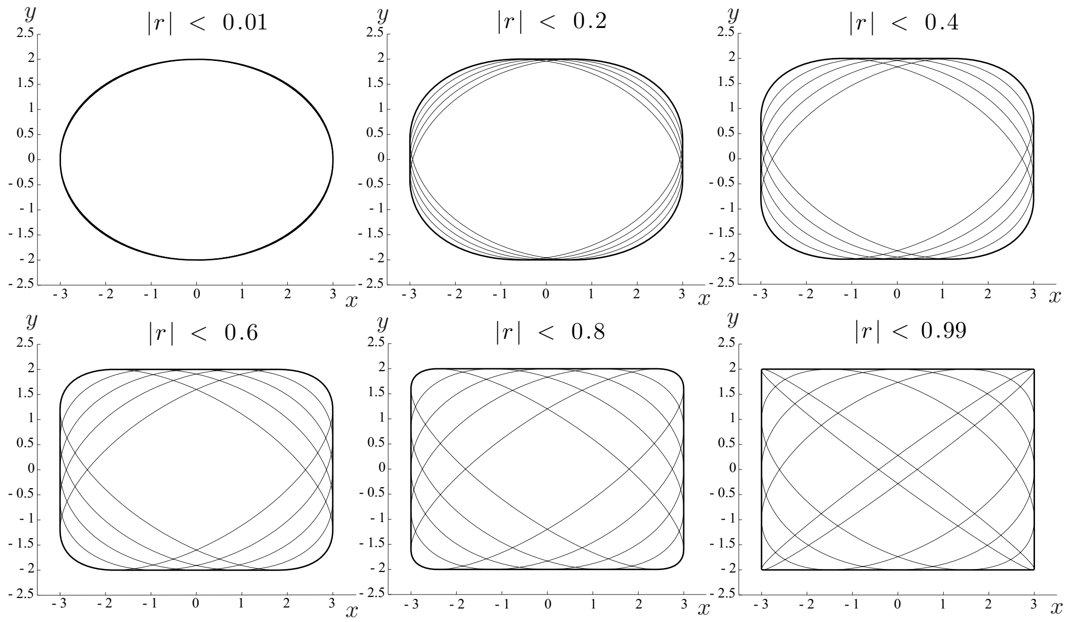

Figure 1: Some members of the family of possible covariances (1–sigma–bounds) for different constraints on the maximum absolute correlation coefficient. The unconstrained case corresponds to $|r| \leq 1$.

Example 2.1 For two scalar random variables x and y with individual variances $E_{xx} = 9$ and $E_{yy} = 4$, some members of the family of possible joint covariance matrices for different constraints on the maximum absolute correlation coefficient are visualized in Fig. 1 by plotting the respective 1–sigma–bounds [1]. The unconstrained case would correspond to $|r| \leq 1$.

The goal is now to find a family of bounding covariances $\mathbf{B}$ with

$$\mathbf{B} \geq \mathbf{E}(r) \tag{3}$$

for all possible joint covariances $\mathbf{E}(r)$ defined by

$$\mathbf{E}(r) = \begin{bmatrix} \mathbf{E}_{xx} & \mathbf{E}_{xy} \\ \mathbf{E}_{yx} & \mathbf{E}_{yy} \end{bmatrix}$$

with r according to (1) and $\mathbf{E}_{xy}$, $\mathbf{E}_{yx}$ such that (2) holds.

3 Derivation of Covariance Bound

For deriving the desired covariance bound we use the fact that the union of the 1-sigma–bounds of all possible joint covariances forms a convex set aligned with the coordinate axes. Hence, the cross covariances of the bounding covariance matrix have to be zero matrices. For the simplest case of two scalar random variables x and y this is visualized in Fig. 1.

[1] 1–sigma–bounds will be used throughout the paper without loss of generality.

In addition, for achieving an upper bound, the covariance matrices $\mathbf{E}_{xx}$ and $\mathbf{E}_{yy}$ have to be individually scaled. Combining both conditions yields

$$\mathbf{B} = \begin{bmatrix} k_x \mathbf{E}_{xx} & \mathbf{0} \\ \mathbf{0} & k_y \mathbf{E}_{yy} \end{bmatrix} . \tag{4}$$

k_x, k_y have to be selected in such a way that (3) holds.

Theorem 3.1 *The scale factors k_x, k_y in (4) are given by*

$$k_x = \frac{1}{\eta - \kappa} , \; k_y = \frac{1}{\eta + \kappa} \tag{5}$$

with

$$\kappa^2 \leq \frac{1 - 2\eta}{1 - r_{max}^2} + \eta^2 \tag{6}$$

and

$$0.5 \leq \eta \leq \frac{1}{1 + r_{max}} . \tag{7}$$

PROOF. For proving (3), the difference matrix

$$\mathbf{D} = \mathbf{B}(\eta,\kappa) - \mathbf{E}(r)$$
$$= \begin{bmatrix} \frac{1}{\eta-\kappa}\mathbf{E}_{xx} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\eta+\kappa}\mathbf{E}_{yy} \end{bmatrix} - \begin{bmatrix} \mathbf{E}_{xx} & \mathbf{E}_{xy} \\ \mathbf{E}_{yx} & \mathbf{E}_{yy} \end{bmatrix}$$
$$= \begin{bmatrix} \left(\frac{1}{\eta-\kappa}-1\right)\mathbf{E}_{xx} & -\mathbf{E}_{xy} \\ -\mathbf{E}_{yx} & \left(\frac{1}{\eta+\kappa}-1\right)\mathbf{E}_{yy} \end{bmatrix}$$

is considered. According to Sylvester's criterion, the matrix $\mathbf{D}$ is positive semi–definite, if the determinants of all submatrices $\mathbf{D}(1:N+i, 1:N+i)$ for $i = 1,\dots,M$ are larger than or equal to zero. $\left(\frac{1}{\eta-\kappa}-1\right)\mathbf{E}_{xx}$ is positive definite and does not need to be tested. The determinants are given by

$$|\mathbf{D}(1:N+i, 1:N+i)|$$
$$= \left|\left(\frac{1}{\eta-\kappa}-1\right)\mathbf{E}_{xx}\right| \left|\left(\frac{1}{\eta+\kappa}-1\right)\mathbf{E}_{yy}(1:i,1:i) - \left(\frac{1}{\eta-\kappa}-1\right)^{-1}\mathbf{E}_{yx}(1:i,1:N)\mathbf{E}_{xx}^{-1}\mathbf{E}_{xy}(1:N,1:i)\right| \ .$$

With (2) we obtain

$$\left|\left[\left(\frac{1}{\eta-\kappa}-1\right)\left(\frac{1}{\eta+\kappa}-1\right) - r_{max}^2\right]\mathbf{E}_{yy}(1:i,1:i)\right| \overset{!}{\geq} 0$$

for $i = 1,\dots,M$, which is equivalent to

$$\left(\frac{1}{\eta-\kappa}-1\right)\left(\frac{1}{\eta+\kappa}-1\right) - r_{max}^2 \geq 0$$

and yields (6). The constraint on η in (7) then follows by claiming a non–negative right–hand–side in (6). □

The parameter set for η and κ from the Theorem is redundant in the sense that it specifies scaled variants of a bounding covariance with the same form and orientation. Hence, it is sufficient to restrict attention to the smallest of these scaled variants. The appropriate parameter values are specified in the following Lemma.

LEMMA 3.1 *A family of bounding covariances* $\mathbf{E}(\kappa)$ *depending on a parameter* κ *is given by (4) with* k_x, k_y *in (5). The parameter* κ *may vary according to*

$$|\kappa| \leq 0.5 \ . \tag{8}$$

η *is a function of* κ *given by*

$$\eta(\kappa) = \frac{1-\sqrt{r_{max}^2 + \kappa^2\left(1-r_{max}^2\right)^2}}{1-r_{max}^2} \ . \tag{9}$$

The admissible values for η and κ resulting from Lemma 3.1 are visualized for different values of r_{max} in Fig. 2.

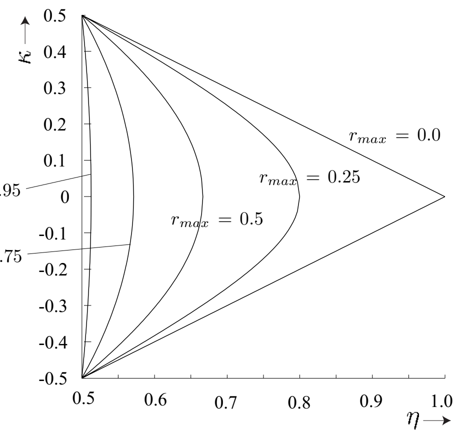

Figure 2: The admissible values for $\eta(\kappa)$ and κ resulting from Lemma 3.1.

4 Discussion of Result

The resulting family of bounding covariance matrices $\mathbf{B}(\kappa)$ given in Lemma 3.1 is now discussed regarding optimality, selection of one member, and more complicated correlation constraints.

Optimality: An important feature of the new approach is that every member of the family of bounding covariance matrices $\mathbf{B}(\kappa)$ bounds every possible covariance $\mathbf{E}(r)$ for r according to (1). Even more, consider the union $U(r_{max})$ of the 1–sigma–bounds of all possible covariance matrices fulfilling the given correlation constraint and the intersection $I(\kappa)$ of the 1–sigma–bounds of the proposed family of bounding covariances.

The previously known bound from [3] ensures that the set $U(r_{max})$ is always a subset of the set $I(\kappa)$ according to

$$U(r_{max}) = \bigcup_{|r|\leq r_{max}} \mathbf{E}(r) \subset \bigcap_{\kappa} \mathbf{B}(\kappa) = I(\kappa)$$

However, it is apparent from the example given in Fig. 3 that the approximation is not tight for $r_{max} < 1$. In contrast, for the new bound the set $U(r_{max})$ is **equivalent** for every r_{max} to the set $I(\kappa)$, i.e.,

$$U(r_{max}) \equiv I(\kappa) \ ,$$

which is visualized in Fig. 4.

Selection of κ: Up to now, the complete family of bounding covariance matrices $\mathbf{B}(\kappa)$ has been considered. Of course, when applying the new bound, a specific value κ^* has to be selected, which results in a single joint covariance matrix $\mathbf{B}(\kappa^*)$.

More complex constraints: The case of more general constraints on the correlation coefficient of the form

$$-1 \leq r_{min} \leq r \leq r_{max} \leq 1$$

is not considered here, since it is regarded to be of minor practical importance. Furthermore, it leads to a more complicated family of bounding covariance matrices $\mathbf{B}(\kappa)$ with nonzero cross covariances.

5 Application Example

For demonstrating the performance of the new bound, a typical application problem is solved: Propagation of a given state recursively through a linear system. The system equation is given by

$$\underline{x}_{k+1} = \underline{x}_k + \underline{w}_{k+1} \ ,$$

where the noise vector $\underline{w}_k$ is correlated with the state $\underline{x}_k$. The level of correlation between $\underline{w}_k$ and $\underline{x}_k$ is unknown but constrained by (1) with a given r_{max}. The individual covariances of the initial state and of the (time–invariant) noise are selected as

$$\mathbf{E}_0^{xx} = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{E}_k^{ww} = \begin{bmatrix} 1 & -0.9 \\ -0.9 & 1 \end{bmatrix} \ ,$$

respectively. The goal is to calculate a bounding covariance matrix for all the possible covariance matrices $\mathbf{E}_k^{xx}(r_k)$ of the state $\underline{x}_k$ for several time steps recursively.

For reference purposes, the covariance matrices $\mathbf{E}_k^{xx}(r_k)$ are calculated according to

$$\mathbf{E}_{k+1}^{xx}(r_{k+1}) = \mathbf{T} \begin{bmatrix} \mathbf{E}_k^{xx}(r_k) & \mathbf{E}_{k+1}^{xw}(r_{k+1}) \\ \mathbf{E}_{k+1}^{wx}(r_{k+1}) & \mathbf{E}_{k+1}^{ww} \end{bmatrix} \mathbf{T}^T$$

for $k = 0, 1, 2, \ldots$ and all possible cross covariances $\mathbf{E}_{k+1}^{xw}(r_{k+1})$, $\mathbf{E}_{k+1}^{wx}(r_{k+1})$ compatible with the given bound (1), where $\mathbf{T}$ is given by

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \ .$$

Applying the new bound gives

$$\mathbf{E}_{k+1}^{xx}(\kappa_{k+1}) = \frac{\mathbf{E}_k^{xx}(\kappa_k)}{\eta_{k+1} - \kappa_{k+1}} + \frac{\mathbf{E}_{k+1}^{ww}}{\eta_{k+1} + \kappa_{k+1}}$$

for $k = 0, 1, 2, \ldots$ and η_k, κ_k from Lemma 3.1. For comparison purposes, the existing bound [3] has been applied to the propagation problem.

Results are shown in Fig. 5 for $r_{max} = 0.2$ and in Fig. 6 for $r_{max} = 0.6$. The shaded regions correspond to the convex hull of the 1–sigma–bounds of all possible joint covariance matrices. In addition, the resulting 1–sigma–bounds of applying the new and the existing bound are plotted. It is obvious that the new bound gives much less conservative results, since the existing bound does not exploit the given correlation constraints.

6 Conclusions

The problem of calculating an upper bound for the joint covariance matrix of two correlated random vectors has been considered for the case that the level of correlation is limited, i.e., the maximum absolute correlation coefficient is less than a prespecified value. This problem has been solved by scaling the individual covariance matrices in such a way that the joint covariance matrix provides a tight upper bound for the set of all possible true joint covariances fulfilling the correlation constraint.

The new bound generalizes and enhances a known result for unconstrained correlation between two random vectors, which gives rather conservative result when a correlation constraint is available. Simulations impressively demonstrate the advantage of the new bound.

As a byproduct, the new covariance bound yields an uncorrelated representation of the joint statistics of the two random vectors under consideration. This provides the basis for the derivation of a state estimation algorithm in the presence of correlated noise with a prespecified maximum correlation level, which generalizes the results in [3].

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice–Hall, 1979.

[2] J. A. Castellanos, J. D. Tardós, and G. Schmidt, "Building a Global Map of the Environment of a Mobile Robot: The Importance of Correlations", *Proc. of the 1997 IEEE International Conference on Robotics and Automation (ICRA'97)*, Albuquerque, New Mexico, USA, 1997, pp. 1053–1059.

[3] U. D. Hanebeck, K. Briechle, "New Results for Stochastic Prediction and Filtering with Unknown Correlations", this conference.

[4] S. J. Julier and J. K. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations", *Proc. of the 1997 American Control Conference (ACC'97)*, Albuquerque, New Mexico, 1997, pp. 2369-2373.

[5] J. A. Castellanos, J. M. Martínez, J. Neira, J. D. Tardós, "Simultaneous Map Building for Mobile Robots: A Multisensor Fusion Approach", *Proc. of the IEEE International Conference on Robotics and Automation (ICRA'98)*, Leuven, Belgium, 1998, pp. 1244-1249.

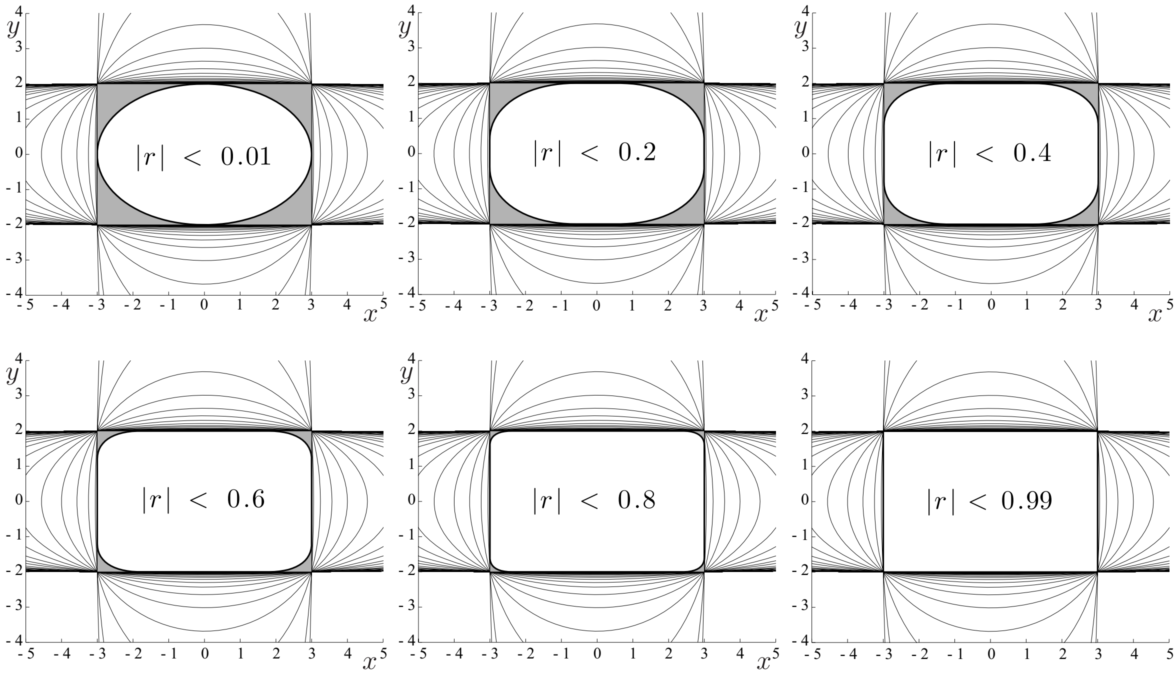

Figure 3: Results of applying the existing bound to the joint covariance of two scalar random variables x and y with individual covariances in accordance with Example 2.1. The result is not tight when the correlation is constrained (the approximation error is shown by the shaded area).

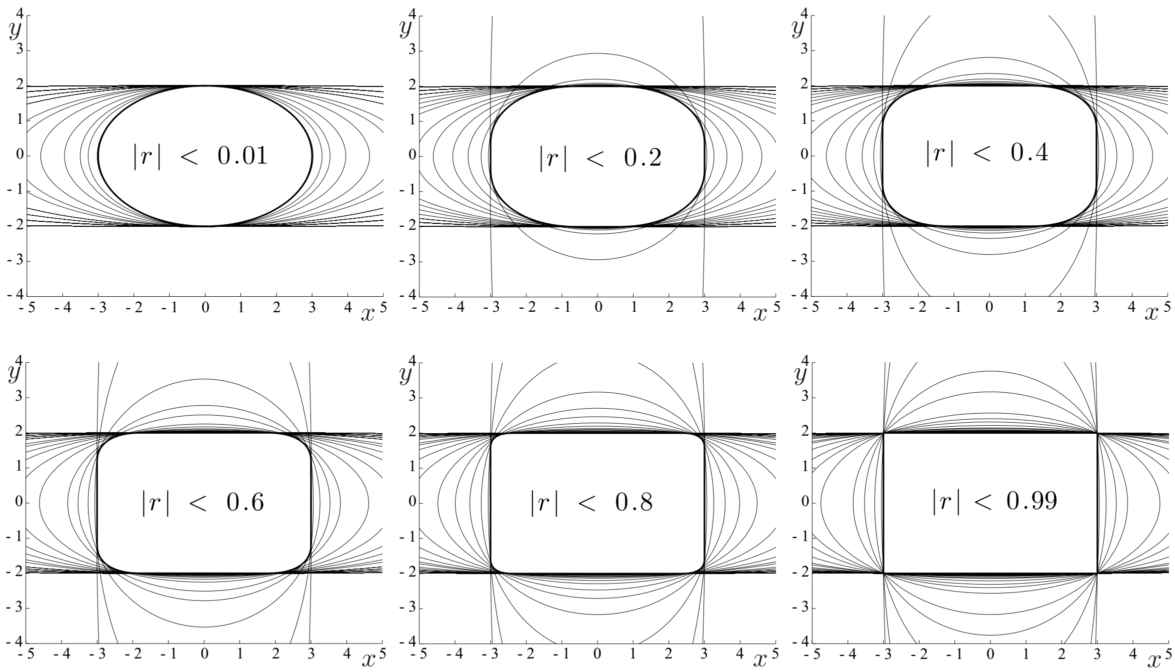

Figure 4: Results of applying the proposed new bound to the joint covariance of two scalar random variables x and y with individual covariances in accordance with Example 2.1. Here the result is tight for all constraints.

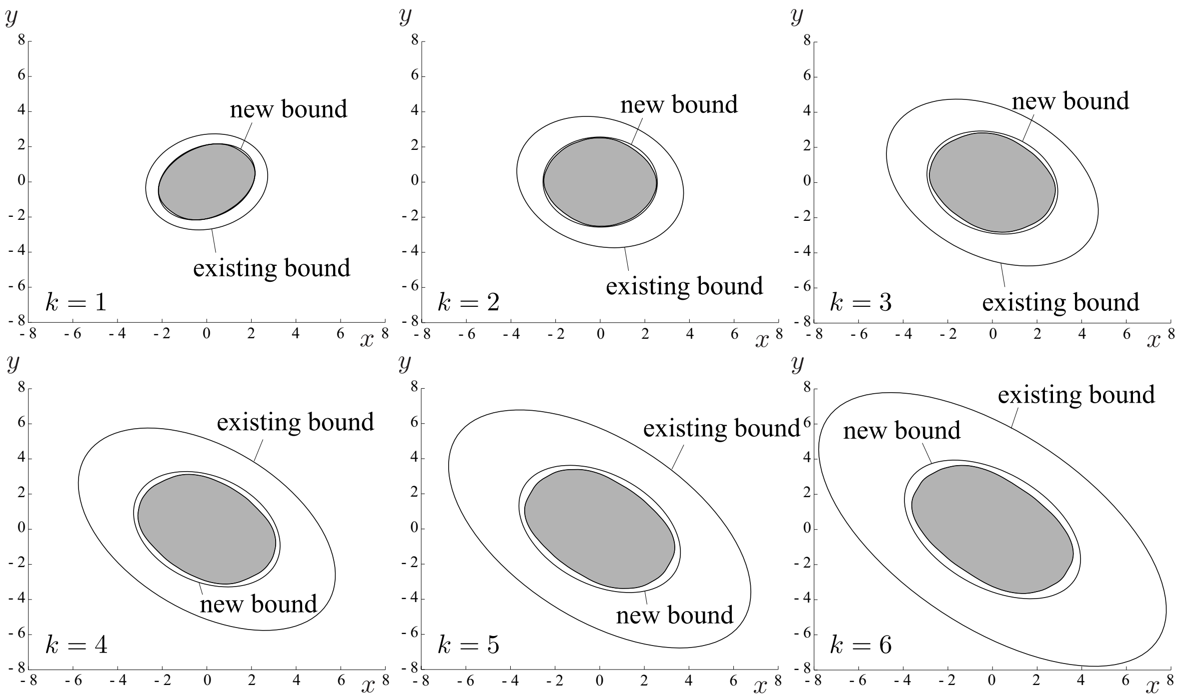

Figure 5: Results of the recursive propagation of a given state through a linear system model suffering from additive noise correlated with the system state ($r_{max} = 0.2$). Shaded: The convex hull of all possible true covariances.

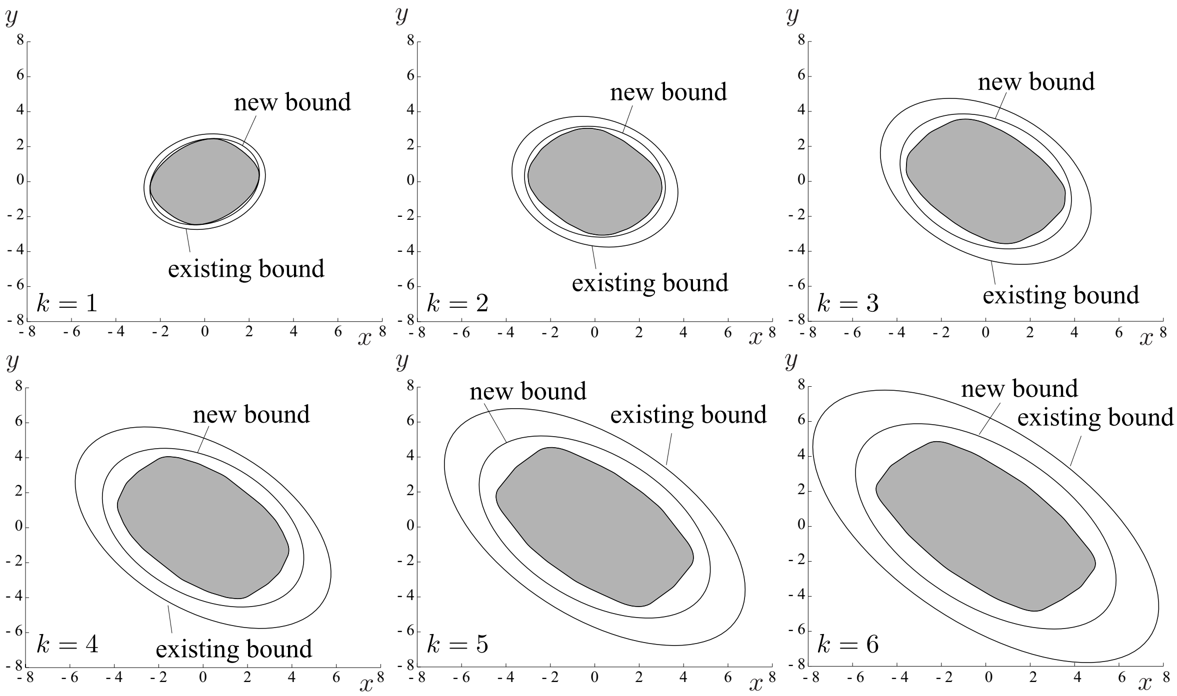

Figure 6: Results of the recursive propagation of a given state through a linear system model suffering from additive noise correlated with the system state ($r_{max} = 0.6$). Shaded: The convex hull of all possible true covariances.