Partitioned Covariance Intersection

Arne Petersen
Christian Albrechts University
Kiel, Germany
E-Mail: apetersen@mip.informatik.uni-kiel.de
Phone: +49 (0)431 880 1418

Marc-André Beyer
Raytheon Anschütz GmbH
Kiel, Germany
E-Mail: Marc-Andre_Beyer@raykiel.com
Phone: +49 (0)431 3019 540

Abstract

In this paper, a new approach for data fusion of two random variables is presented for the case of partially unknown cross-correlations. The new algorithm is called *Partitioned Covariance Intersection* (PCI). PCI yields informative and consistent estimates by partitioning the error estimates into an unknown correlated part and a known correlated part. The application to Kalman Filtering problems is described, where time-correlated sensor noise results in cross-correlations between the state vector and the sensor noise.

The paper is organized as follows: Section 1 introduces the reader to the data fusion problem adressed in this contribution. In section 2 the new algorithm is developed along with required theorems and proofs. Section 3 shows the benefit of the algorithm by means of two examples:

- a simple two dimensional data fusion problem;
- a Kalman filtering example where time-correlation in the sensor noise leads to unknown cross-correlation between sensor noise and the Kalman Filter state.

1 Problem Statement

We want to discuss the problem that two pieces of information, labeled X and Y, are to be fused to yield an output Z. This is a general type of data fusion problem, e.g. in the case where X and Y represent measurements of different input sensors, or X could be a prediction from a system model and Y could be sensor information. The input sources are erroneous, thus they are defined as random variables $\vec{x} \sim \mathcal{N}\left(\bar{x}, \boldsymbol{P}^{\tilde{x}\tilde{x}}\right)$ and $\vec{y} \sim \mathcal{N}\left(\bar{y}, \boldsymbol{P}^{\tilde{y}\tilde{y}}\right)$ respectively[1]. Here $\bar{x} \in \mathbb{R}^n$, $\bar{y} \in \mathbb{R}^m$ give the mean vectors and $\boldsymbol{P}^{\tilde{x}\tilde{x}}$, $\boldsymbol{P}^{\tilde{y}\tilde{y}}$ the covariance matrices for the according random variables.

In the following it is assumed that the true statistics of these variables are not known exactly. The only information which is available are consistent (see A, page 18) estimates $\{\hat{x}, \boldsymbol{P}^{xx}\}, \{\hat{y}, \boldsymbol{P}^{yy}\}$ of the means and covariances of $\vec{x}$ and $\vec{y}$. In addition partial information $\boldsymbol{P}^{xy}$ about the correlations $\boldsymbol{P}^{\tilde{x}\tilde{y}}$ might be available. The problem is to fuse the information from $\vec{x}$ and $\vec{y}$ so that a new estimate $\{\hat{z}, \boldsymbol{P}^{zz}\}$ is determined which minimizes some form of cost function but guarantees consistency to the true fused random variable $\vec{z} \sim \mathcal{N}\left(\bar{z}, \boldsymbol{P}^{\tilde{z}\tilde{z}}\right)$.

The most common linear data fusion algorithms (best linear unbiased estimators - BLUE) compute a weighted mean of input variables. The covariance of the resulting estimation is determined using linear error propagation. Although this is an optimal approach when the statistics are exactly known, i.e. $\boldsymbol{P}^{xy} = \boldsymbol{P}^{\tilde{x}\tilde{y}}$, inconsistencies result in the presence of unknown cross correlations. I.e. the Gauss Markov Model (GMM, see [1]) uses a linear update rule of the form

$$\boldsymbol{P}^{\tilde{z}\tilde{z}} = \left[\boldsymbol{H}^T \begin{bmatrix} \boldsymbol{P}^{\tilde{x}\tilde{x}} & \boldsymbol{P}^{\tilde{x}\tilde{y}} \\ \boldsymbol{P}^{\tilde{y}\tilde{x}} & \boldsymbol{P}^{\tilde{y}\tilde{y}} \end{bmatrix}^{-1} \boldsymbol{H}\right]^{-1} \tag{1}$$

$$\vec{z} = \boldsymbol{P}^{\tilde{z}\tilde{z}} \boldsymbol{H}^T \begin{bmatrix} \boldsymbol{P}^{\tilde{x}\tilde{x}} & \boldsymbol{P}^{\tilde{x}\tilde{y}} \\ \boldsymbol{P}^{\tilde{y}\tilde{x}} & \boldsymbol{P}^{\tilde{y}\tilde{y}} \end{bmatrix}^{-1} \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \tag{2}$$

Where:

$$\begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} = \boldsymbol{H}\vec{z} \tag{3}$$

[1] $\boldsymbol{P}^{\tilde{a}\tilde{a}} = \mathbb{E}[\tilde{a}\tilde{a}^T] = \mathbb{E}[(\vec{a} - \bar{a})(\vec{a} - \bar{a})^T]$

However, this estimation requires that all correlations and covariances are known. That is, applying this algorithm to the problem defined above

$$\boldsymbol{P}^{zz} = \left[\boldsymbol{H}^T \begin{bmatrix} \boldsymbol{P}^{xx} & \boldsymbol{P}^{xy} \\ \boldsymbol{P}^{yx} & \boldsymbol{P}^{yy} \end{bmatrix}^{-1} \boldsymbol{H} \right]^{-1} \quad (4)$$

$$\hat{z} = \boldsymbol{P}^{zz} \boldsymbol{H}^T \begin{bmatrix} \boldsymbol{P}^{xx} & \boldsymbol{P}^{xy} \\ \boldsymbol{P}^{yx} & \boldsymbol{P}^{yy} \end{bmatrix}^{-1} \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} \quad (5)$$

might result in an inconsistent estimation. This problem is addressed by the Covariance Intersection (CI) algorithms, which are well known (e.g. see [2], [3]).

Consider the problem that two instances of X and Y need to be fused where only parts of the correlation are known in form of partitions of the vectors. Specifically it is assumed that the random variables X and Y can be decomposed into two terms

$$\vec{x} = [\vec{x}_{cu} \mid \vec{x}_{ck}]^T \in \mathbb{R}^{n \times n} \quad (6)$$

$$\vec{y} = [\vec{y}_{cu} \mid \vec{y}_{ck}]^T \in \mathbb{R}^{m \times m} \quad (7)$$

Where $\vec{\star}_{cu}$ is the portion of the respective variable in which components are potentially correlated with one another. $\vec{\star}_{ck}$ denotes the fraction of X and Y respectively whose correlations with all other components are known. Since the correlations within $\vec{\star}_{cu}$ are unknown standard BLUEs cannot be applied.

This paper introduces an algorithm that transforms the given statistics similar to the CI, so that standard BLU-estimators can be applied. Moreover it takes advantage of the partialy known correlations between the input informations. This way the resulting estimation is consistent but more informative (not as conservative) compared to the CI.

2 Partitioned Covariance Intersection

In the simple case when the assumed and actual variables are uncorrelated or the correlation is exactly known ($\boldsymbol{P}^{xy} = \boldsymbol{P}^{\tilde{x}\tilde{y}} \ \vee \boldsymbol{P}^{xy} = \boldsymbol{P}^{\tilde{x}\tilde{y}} = \boldsymbol{0}$) it is obvious that

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{P}^{xx} & \boldsymbol{P}^{xy} \\ \boldsymbol{P}^{yx} & \boldsymbol{P}^{yy} \end{bmatrix} \geq \begin{bmatrix} \boldsymbol{P}^{\tilde{x}\tilde{x}} & \boldsymbol{P}^{\tilde{x}\tilde{y}} \\ \boldsymbol{P}^{\tilde{y}\tilde{x}} & \boldsymbol{P}^{\tilde{y}\tilde{y}} \end{bmatrix} = \tilde{\boldsymbol{P}} \tag{8}$$

and thus, equations (4) and (5) yield a consistent estimate $\{\bar{z}, \boldsymbol{P}^{zz}\}$ (e.g. see [2]). However, when $\boldsymbol{P}^{xy} \neq \boldsymbol{P}^{\tilde{x}\tilde{y}}$, it is difficult to guarantee a consistent update.

If the structure of the correlation is partially known, it is possible to achieve better estimation accuracy with a modification of the CI algorithm by employing this additional information. The Partitioned Covariance Intersection (PCI) algorithm extends CI by allowing some information about the cross correlations to be exploited.

By introducing the partition mentioned above the true joint system covariance matrix can be written as follows:

$$\begin{aligned} \tilde{\boldsymbol{P}} &= \begin{bmatrix} \boldsymbol{P}^{\tilde{x}\tilde{x}} & \boldsymbol{P}^{\tilde{x}\tilde{y}} \\ \boldsymbol{P}^{\tilde{y}\tilde{x}} & \boldsymbol{P}^{\tilde{y}\tilde{y}} \end{bmatrix} \\ &= \left[\begin{array}{cc|cc} \boldsymbol{P}^{\tilde{x}_{cu}\tilde{x}_{cu}} & \boldsymbol{P}^{\tilde{x}_{cu}\tilde{x}_{ck}} & \boldsymbol{P}^{\tilde{x}_{cu}\tilde{y}_{cu}} & \boldsymbol{P}^{\tilde{x}_{cu}\tilde{y}_{ck}} \\ \boldsymbol{P}^{\tilde{x}_{ck}\tilde{x}_{cu}} & \boldsymbol{P}^{\tilde{x}_{ck}\tilde{x}_{ck}} & \boldsymbol{P}^{\tilde{x}_{ck}\tilde{y}_{cu}} & \boldsymbol{P}^{\tilde{x}_{ck}\tilde{y}_{ck}} \\ \hline \boldsymbol{P}^{\tilde{y}_{cu}\tilde{x}_{cu}} & \boldsymbol{P}^{\tilde{y}_{cu}\tilde{x}_{ck}} & \boldsymbol{P}^{\tilde{y}_{cu}\tilde{y}_{cu}} & \boldsymbol{P}^{\tilde{y}_{cu}\tilde{y}_{ck}} \\ \boldsymbol{P}^{\tilde{y}_{ck}\tilde{x}_{cu}} & \boldsymbol{P}^{\tilde{y}_{ck}\tilde{x}_{ck}} & \boldsymbol{P}^{\tilde{y}_{ck}\tilde{y}_{cu}} & \boldsymbol{P}^{\tilde{y}_{ck}\tilde{y}_{ck}} \end{array}\right], \end{aligned} \tag{9}$$

To allow a consistent fusion operation with unknown correlations $\boldsymbol{P}^{\tilde{x}_{cu}\tilde{y}_{cu}} = (\boldsymbol{P}^{\tilde{y}_{cu}\tilde{x}_{cu}})^T$, a larger matrix $\boldsymbol{P}_\alpha \geq \tilde{\boldsymbol{P}}$ needs to be found.

Theorem 1. *Given a positive definite symmetric matrix $\boldsymbol{P} \in \mathbb{R}^{(n+m)\times(n+m)}$ with*

$$\begin{aligned} \boldsymbol{P} &= \begin{bmatrix} \boldsymbol{P}^{xx} & \boldsymbol{P}^{xy} \\ \boldsymbol{P}^{yx} & \boldsymbol{P}^{yy} \end{bmatrix} \\ &= \left[\begin{array}{cc|cc} \boldsymbol{P}^{x_{cu}x_{cu}} & \boldsymbol{P}^{x_{cu}x_{ck}} & \boldsymbol{P}^{x_{cu}y_{cu}} & \boldsymbol{P}^{x_{cu}y_{ck}} \\ \boldsymbol{P}^{x_{ck}x_{cu}} & \boldsymbol{P}^{x_{ck}x_{ck}} & \boldsymbol{P}^{x_{ck}y_{cu}} & \boldsymbol{P}^{x_{ck}y_{ck}} \\ \hline \boldsymbol{P}^{y_{cu}x_{cu}} & \boldsymbol{P}^{y_{cu}x_{ck}} & \boldsymbol{P}^{y_{cu}y_{cu}} & \boldsymbol{P}^{y_{cu}y_{ck}} \\ \boldsymbol{P}^{y_{ck}x_{cu}} & \boldsymbol{P}^{y_{ck}x_{ck}} & \boldsymbol{P}^{y_{ck}y_{cu}} & \boldsymbol{P}^{y_{ck}y_{ck}} \end{array}\right], \end{aligned} \tag{10}$$

a larger matrix $\boldsymbol{P}_\alpha \geq \boldsymbol{P}$ is given by

$$\begin{aligned}\boldsymbol{P}_\alpha &= \left[\begin{array}{cc|cc} \frac{1}{\alpha}\boldsymbol{P}^{x_{cu}x_{cu}} & \boldsymbol{P}^{x_{cu}x_{ck}} & \mathbf{0} & \boldsymbol{P}^{x_{cu}y_{ck}} \\ \boldsymbol{P}^{x_{ck}x_{cu}} & \boldsymbol{P}^{x_{ck}x_{ck}} & \boldsymbol{P}^{x_{ck}y_{cu}} & \boldsymbol{P}^{x_{ck}y_{ck}} \\ \hline \mathbf{0} & \boldsymbol{P}^{y_{cu}x_{ck}} & \frac{1}{1-\alpha}\boldsymbol{P}^{y_{cu}y_{cu}} & \boldsymbol{P}^{y_{cu}y_{ck}} \\ \boldsymbol{P}^{y_{ck}x_{cu}} & \boldsymbol{P}^{y_{ck}x_{ck}} & \boldsymbol{P}^{y_{ck}y_{cu}} & \boldsymbol{P}^{y_{ck}y_{ck}} \end{array}\right] \\ &= \left[\begin{array}{c|c} \boldsymbol{P}_\alpha^{xx} & \boldsymbol{P}_\alpha^{xy} \\ \hline \boldsymbol{P}_\alpha^{yx} & \boldsymbol{P}_\alpha^{yy} \end{array}\right] \end{aligned} \tag{11}$$

with $0 < \alpha < 1$.

Proof 1. *For $\boldsymbol{P}_\alpha - \boldsymbol{P}$ we obtain*

$$\boldsymbol{P}_\alpha - \boldsymbol{P} = \left[\begin{array}{cc|cc} (\frac{1}{\alpha}-1)\boldsymbol{P}^{x_{cu}x_{cu}} & \mathbf{0} & -\boldsymbol{P}^{x_{cu}y_{cu}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline -\boldsymbol{P}^{y_{cu}x_{cu}} & \mathbf{0} & (\frac{1}{1-\alpha}-1)\boldsymbol{P}^{y_{cu}y_{cu}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right]. \tag{12}$$

This matrix can be rotated without changing its eigenvalues by using a similarity transformation induced by a rotation matrix $\boldsymbol{R}$ and we get

$$\begin{aligned} &\boldsymbol{R}(\boldsymbol{P}_\alpha - \boldsymbol{P})\boldsymbol{R}^T \\ &= \left[\begin{array}{cc|cc} (\frac{1}{\alpha}-1)\boldsymbol{P}^{x_{cu}x_{cu}} & -\boldsymbol{P}^{x_{cu}y_{cu}} & \mathbf{0} & \mathbf{0} \\ -\boldsymbol{P}^{y_{cu}x_{cu}} & (\frac{1}{1-\alpha}-1)\boldsymbol{P}^{y_{cu}y_{cu}} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right] \\ &= \left[\begin{array}{cc|cc} \lambda\boldsymbol{P}^{x_{cu}x_{cu}} & -\boldsymbol{P}^{x_{cu}y_{cu}} & \mathbf{0} & \mathbf{0} \\ -\boldsymbol{P}^{y_{cu}x_{cu}} & \frac{1}{\lambda}\boldsymbol{P}^{y_{cu}y_{cu}} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right] \end{aligned} \tag{13}$$

The lower right matrix is positive semidefinite. The quadratic form of the upper left matrix is given by

$$\begin{aligned} &\boldsymbol{Q} \\ &= \begin{bmatrix} \vec{a}^T & \vec{b}^T \end{bmatrix} \begin{bmatrix} \lambda\boldsymbol{P}^{x_{cu}x_{cu}} & -\boldsymbol{P}^{x_{cu}y_{cu}} \\ -\boldsymbol{P}^{y_{cu}x_{cu}} & \frac{1}{\lambda}\boldsymbol{P}^{y_{cu}y_{cu}} \end{bmatrix} \begin{bmatrix} \vec{a} \\ \vec{b} \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{\lambda}\vec{a}^T & -\frac{\vec{b}^T}{\sqrt{\lambda}} \end{bmatrix} \underbrace{\begin{bmatrix} \boldsymbol{P}^{x_{cu}x_{cu}} & \boldsymbol{P}^{x_{cu}y_{cu}} \\ \boldsymbol{P}^{y_{cu}x_{cu}} & \boldsymbol{P}^{y_{cu}y_{cu}} \end{bmatrix}}_{\boldsymbol{Z}} \begin{bmatrix} \sqrt{\lambda}\vec{a} \\ -\frac{\vec{b}}{\sqrt{\lambda}} \end{bmatrix} \end{aligned} \tag{14}$$

with arbitrary vectors $\vec{a} \in \mathbb{R}^n$ *and* $\vec{b} \in \mathbb{R}^m$. *Since* $\boldsymbol{Z}$ *is positive definite, we have* $\boldsymbol{Q} \geq \boldsymbol{0}$ *what implies that* $\boldsymbol{P}_\alpha - \boldsymbol{P}$ *is positive semidefinite.*

■

Corollary 1. *From* $\boldsymbol{P}_\alpha \geq \boldsymbol{P}$ *it follows directly, that* $\boldsymbol{P}_\alpha > \boldsymbol{0}$, *i.e.* $\boldsymbol{P}_\alpha$ *is positive definite, since* $\boldsymbol{P} > \boldsymbol{0}$.

From the derived theorem 1 and corollary 1 and assuming that the input estimates of X and Y obey consistency conditions (63) and (64), the PCI algorithm can be summarized by

$$\boldsymbol{P}^{zz} = \left[\boldsymbol{H}^T \begin{bmatrix} \boldsymbol{P}_\alpha^{xx} & \boldsymbol{P}_\alpha^{xy} \\ \boldsymbol{P}_\alpha^{yx} & \boldsymbol{P}_\alpha^{yy} \end{bmatrix}^{-1} \boldsymbol{H}\right]^{-1} \tag{15}$$

$$\bar{z} = \boldsymbol{P}^{zz}\boldsymbol{H}^T \begin{bmatrix} \boldsymbol{P}_\alpha^{xx} & \boldsymbol{P}_\alpha^{xy} \\ \boldsymbol{P}_\alpha^{yx} & \boldsymbol{P}_\alpha^{yy} \end{bmatrix}^{-1} \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} \tag{16}$$

Here $\boldsymbol{H}$ represents the transformation matrix, that transforms the input estimates of X and Y into the space of the output estimate of Z. The modified input estimate covariance matrices $\boldsymbol{P}_\alpha^{xx}$ and $\boldsymbol{P}_\alpha^{yy}$ are given by

$$\boldsymbol{P}_\alpha^{xx} = \begin{bmatrix} \frac{1}{\alpha}\boldsymbol{P}^{x_{cu}x_{cu}} & \boldsymbol{P}^{x_{cu}x_{ck}} \\ \boldsymbol{P}^{x_{ck}x_{cu}} & \boldsymbol{P}^{x_{ck}x_{ck}} \end{bmatrix} \tag{17}$$

$$\boldsymbol{P}_\alpha^{yy} = \begin{bmatrix} \frac{1}{1-\alpha}\boldsymbol{P}^{y_{cu}y_{cu}} & \boldsymbol{P}^{y_{cu}y_{ck}} \\ \boldsymbol{P}^{y_{ck}y_{cu}} & \boldsymbol{P}^{y_{ck}y_{ck}} \end{bmatrix} \tag{18}$$

$$\boldsymbol{P}_\alpha^{xy} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{P}^{x_{cu}y_{ck}} \\ \boldsymbol{P}^{x_{ck}y_{cu}} & \boldsymbol{P}^{x_{ck}y_{ck}} \end{bmatrix} \tag{19}$$

$$\boldsymbol{P}_\alpha^{yx} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{P}^{y_{cu}x_{ck}} \\ \boldsymbol{P}^{y_{ck}x_{cu}} & \boldsymbol{P}^{y_{ck}x_{ck}} \end{bmatrix} \tag{20}$$

The parameter $\alpha \in (0, 1)$ can be estimated using a non-linear optimization. It has to be chosen to minimize some sort of norm (i.e. trace, determinant etc.) of the resulting covariance matrix $\boldsymbol{P}^{zz}$. Solutions to this problem can be found in [4].

3 Examples

3.1 Fusing two pieces of information

We want to fuse the two pieces of information X and Y, where consistent estimates of X and Y are given by

$$\begin{aligned} \boldsymbol{P}^{xx} &= \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} && (21) \\ \boldsymbol{P}^{yy} &= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} && (22) \end{aligned}$$

Both true errors can be partitioned into

$$\begin{aligned} \tilde{x} &= [\tilde{x}_{cu} \mid \tilde{x}_{ck}]^T = [\tilde{x}_1 | \ \tilde{x}_2]^T && (23) \\ \tilde{y} &= [\tilde{y}_{cu} \mid \tilde{y}_{ck}]^T = [\tilde{y}_1 | \ \tilde{y}_2]^T, && (24) \end{aligned}$$

so we assume that only the correlation between $\tilde{x}_1$ and $\tilde{y}_1$ is unknown. Moreover, the true error cross covariance matrix is given by

$$\boldsymbol{P}^{\tilde{x}\tilde{y}} = (\boldsymbol{P}^{\tilde{y}\tilde{x}})^T = \begin{bmatrix} \boldsymbol{P}^{\tilde{x}_1\tilde{y}_1} & 0 \\ 0 & 0 \end{bmatrix}. \qquad (25)$$

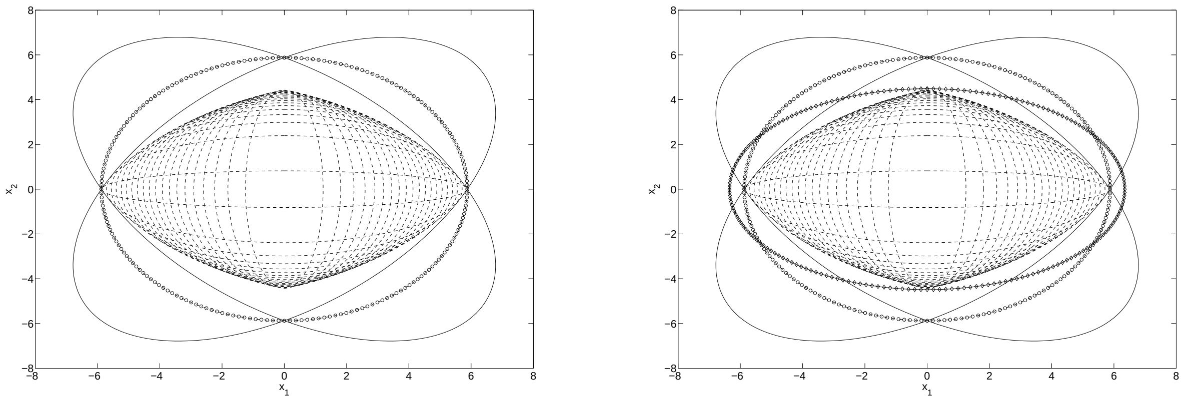
$\boldsymbol{P}^{\tilde{x}_1\tilde{y}_1} = (\boldsymbol{P}^{\tilde{y}_1\tilde{x}_1})^T$ is completely unknown but both estimates can be fused in a consistent and informative fashion using the PCI algorithm.

The results displayed in Figure 1 show that PCI allows more progressive estimates to be computed and is not as conservative as CI while maintaining consistency given the partial information about cross correlation. In this particular example the determinant of the PCI fused estimate reduces by $\approx 30\,\%$ compared to CI.

3.2 PCI in Kalman Filtering

In signal processing Kalman Filters are frequently used for the estimation in dynamic systems as described in [6], [7]. The system under investigation is a ship traveling a constant course $\psi = 30\,°$. The state vector is defined as

$$\vec{x} = \begin{bmatrix} p^n & p^e & u_r & v_r & u_c & v_c \end{bmatrix}^T \qquad (26)$$



(a) Covariance Intersection

(b) Partitioned Covariance Intersection

Figure 1: The shape of the updated covariance ellipses for the CI and PCI algorithms. The covariance ellipses for input estimates X and Y are the solid outer ellipses. The dash-dotted lines show the ellipses of $\boldsymbol{P}^{zz}$ for various values of $\boldsymbol{P}^{\tilde{x}_1\tilde{y}_1}$. The determinant minimizing solution of the CI fused estimates ($\alpha \approx 0.51$) is shown with the solid line with circle markers. In the right figure, the PCI fused covariance ellipse of $\boldsymbol{P}^{zz}$ ($\alpha = 0.5$) is shown with the solid line with diamond markers.

where p^n, p^e denote north and east position, u_r and v_r water relative velocity (surge and sway) of the vessel along the vessels body coordinates. u_c and v_c denote north and east water current velocity. The dynamic stochastic system can be described by

$$\begin{aligned}
\vec{x}_{k+1} &= \boldsymbol{\Phi}\vec{x}_k + \boldsymbol{G}\vec{\kappa}_k && (27)\\
\boldsymbol{\Phi} &= e^{\boldsymbol{F}\Delta t} && (28)\\
\boldsymbol{F} &= \begin{bmatrix} 0 & 0 & +\cos(\psi) & -\sin(\psi) & 1 & 0 \\ 0 & 0 & +\sin(\psi) & +\cos(\psi) & 0 & 1 \\ \boldsymbol{0}^{4\times 6} & & & & & \end{bmatrix} && (29)\\
\boldsymbol{G} &= \begin{bmatrix} \boldsymbol{0}^{2\times 4} \\ \boldsymbol{I}^{4\times 4} \end{bmatrix}. && (30)
\end{aligned}$$

$\vec{\kappa}_k \in \mathbb{R}^{4\times 1}$ is a zero-mean random variable with covariance matrix

$$\boldsymbol{P}_k^{\kappa\kappa} = \mathrm{diag}([\sigma^2_{\kappa_{u_r}} \ \sigma^2_{\kappa_{v_r}} \ \sigma^2_{\kappa_{u_c}} \ \sigma^2_{\kappa_{v_c}}]). \quad (31)$$

Assume that the vessel is equipped with two sensors, one measuring positions and one measuring the water relative velocities.

The position sensor is assumed to be an Inertial Navigation System (INS) that operates in a free inertial mode, i.e. no guidance by e.g. GPS. INS compute a position by integrating noise corrupted measurements of accelerometers and gyroscopes. Thus, the error of such INS is not Gaussian distributed but has typically a 2nd-order random walk characteristic, so the measurement equation[2] is given by

$$\vec{y}_{1,k} = \boldsymbol{H}_1 \vec{x}_k + \vec{\gamma}_{1,k} \tag{32}$$

$$\boldsymbol{H}_1 = [\boldsymbol{I}^{2\times 2}\ \mathbf{0}^{2\times 4}] \tag{33}$$

$$\vec{\gamma}_1 = \vec{\Gamma}_{1,3} \tag{34}$$

$$\vec{\Gamma}_{k+1} = \boldsymbol{\Phi}_\Gamma \vec{\Gamma}_k + \boldsymbol{G}_\Gamma \vec{\mu}_k \tag{35}$$

$$\boldsymbol{\Phi}_\Gamma = \text{blkdiag}(\begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}) \tag{36}$$

$$\boldsymbol{G}_\Gamma = \begin{bmatrix} 0 & 0 \\ \Delta t & 0 \\ 0 & 0 \\ 0 & \Delta t \end{bmatrix}, \tag{37}$$

where $\Gamma = [e_{p^n}\ e_{v^n}\ e_{p^e}\ e_{v^e}]^T$ is the error vector of velocities and positions in north and east direction respectively. $\mu \in \mathbb{R}^{2\times 2}$ is zero-mean white accelerometer noise driving the random-walk error of the INS with covariance matrix $\boldsymbol{P}^{\mu\mu}$. The covariance matrices of Γ and $\vec{\gamma}_1$ are given by

$$\boldsymbol{P}_{k+1}^{\Gamma\Gamma} = \boldsymbol{\Phi}_\Gamma \boldsymbol{P}_k^{\Gamma\Gamma} \boldsymbol{\Phi}_\Gamma^T + \boldsymbol{G}_\Gamma \boldsymbol{P}_k^{\mu\mu} \boldsymbol{G}_\Gamma^T \tag{38}$$

$$\boldsymbol{P}_k^{\gamma_1\gamma_1} = \begin{bmatrix} \boldsymbol{P}_k^{\Gamma_1\Gamma_1} & \boldsymbol{P}_k^{\Gamma_1\Gamma_3} \\ \boldsymbol{P}_k^{\Gamma_3\Gamma_1} & \boldsymbol{P}_k^{\Gamma_3\Gamma_3} \end{bmatrix} \tag{39}$$

An INS typically outputs this covariance matrix $\boldsymbol{P}^{\gamma_1\gamma_1}$, often even only the diagonal elements.

The water relative velocity sensor is assumed to be an Electromagnetic-Log with two sensitive axes. The measurement equation is given by

$$\vec{y}_{2,k} = \boldsymbol{H}_2 \vec{x}_k + \vec{\gamma}_{2,k} \tag{40}$$

$$\boldsymbol{H}_2 = [\mathbf{0}^{2\times 2}\ \boldsymbol{I}^{2\times 2}\ \mathbf{0}^{2\times 2}] \tag{41}$$

γ_2 is zero-mean white noise with covariance matrix $\boldsymbol{P}^{\gamma_2\gamma_2}$.

[2]Note, that the error propagation equations used here are strongly simplified compared to real strapdown INS equations. Anyhow, the objective is not to show an application to true strapdown INS problems, but give a qualitatively comparable example for this class of fusion problems. For this purpose, the used equations are sufficient.

The severe problem that arises from time-correlated sensor noise is that it leads to cross-correlation between the state vector and the measurement noise, see Figure 2. In the sequel, we compare three

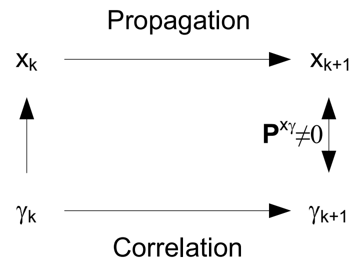


Figure 2: Cross-correlation between sensor noise and state vector arising from time-correlated sensor noise.

different estimators for this data fusion problem.

3.2.1 Kalman Filter

The Kalman Filter uses the supplied covariance matrix from the INS, where the KF equations[3]

$$\boldsymbol{P}^{xx}_{k|k} = \left[\boldsymbol{H}^T \begin{bmatrix} \boldsymbol{P}^{xx}_{k|k-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}^{\gamma\gamma}_k \end{bmatrix}^{-1} \boldsymbol{H}\right]^{-1} \quad (42)$$

$$\boldsymbol{P}^{xx}_{k+1|k} = \boldsymbol{\Phi}\boldsymbol{P}^{xx}_{k|k}\boldsymbol{\Phi}^T + \boldsymbol{G}\boldsymbol{P}^{\kappa\kappa}_k\boldsymbol{G}^T \quad (43)$$

are used. Here the equivalent GMM form is given as introduced in [8]. $\boldsymbol{H}$ and $\boldsymbol{P}^{\gamma\gamma}$ are given by:

$$\boldsymbol{H} = \left[(\boldsymbol{I}^{6\times6})\ (\boldsymbol{H}_1)^T\ (\boldsymbol{H}_2)^T\right]^T \quad (44)$$

$$\boldsymbol{P}^{\gamma\gamma} = \mathrm{blkdiag}(\boldsymbol{P}^{\gamma_1\gamma_1}, \boldsymbol{P}^{\gamma_2\gamma_2}) \quad (45)$$

3.2.2 Kalman Filter with Covariance Intersection

Using standard Covariance Intersection, we assume that the complete state vector may be correlated with the measurement noise, so $\boldsymbol{P}^{x\gamma_1} \neq \boldsymbol{0}$. The modified matrices for the Kalman update are given by

$$\boldsymbol{P}^{xx,\alpha}_{k|k-1} = \frac{1}{1-\alpha}\boldsymbol{P}^{xx}_{k|k-1} \quad (46)$$

$$\boldsymbol{P}^{\gamma_1\gamma_1,\alpha}_k = \frac{1}{\alpha}\boldsymbol{P}^{\gamma_1\gamma_1}_k \quad (47)$$

leading to KF covariance equations

$$\boldsymbol{P}^{xx}_{k|k} = \left[\boldsymbol{H}^T \begin{bmatrix} \boldsymbol{P}^{xx,\alpha}_{k|k-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}^{\gamma\gamma,\alpha}_k \end{bmatrix}^{-1} \boldsymbol{H}\right]^{-1} \quad (48)$$

$$\boldsymbol{P}^{xx}_{k+1|k} = \boldsymbol{\Phi}\boldsymbol{P}^{xx}_{k|k}\boldsymbol{\Phi}^T + \boldsymbol{G}\boldsymbol{P}^{\kappa\kappa}_k\boldsymbol{G}^T, \quad (49)$$

[3] For brevity only the covariance equations are shown

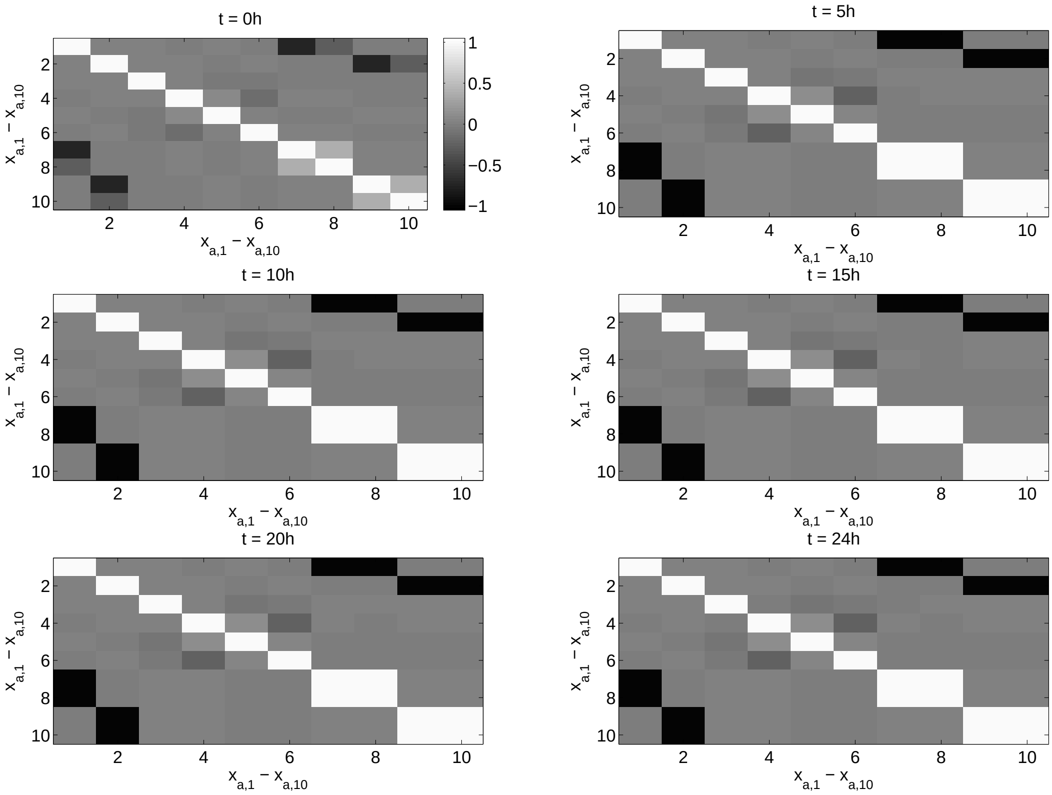


Figure 3: Correlation matrix of a KF with augmented state $\vec{x}_a = [(\vec{x})^T \ (\vec{\Gamma})^T]^T$.

where $\boldsymbol{P}^{\gamma\gamma,\alpha} = \text{blkdiag}(\boldsymbol{P}^{\gamma_1\gamma_1,\alpha}, \boldsymbol{P}^{\gamma_2\gamma_2})$.

3.2.3 Kalman Filter with Partitioned Covariance Intersection

Using equations (42) and (43) the algorithm presented in chapter 2 can be applied directly to the Kalman Filter. This is due to the fact, that both models are BLU-estimators and by this equivalent. That is both systems - Kalman Filter update and GMM - can be transformed directly into each other.

Figure 3 shows the development of the correlation matrix over time of a KF estimator which exactly knows the random walk process of the sensor noise and uses state augmentation to deal with the time correlation (for state augmentation see [9]). Idealistically, this would be the KF estimator with best results, i.e. highest informativity and consistency. Anyhow, in many applications, the random-walk process is not known and alternatives to deal with the time correlation need to be applied (e.g. CI and PCI).

We can see, that the correlation of the position measurement error ($\vec{\gamma}_1 = \vec{x}_{a_{7,9}}$) is only very strong with the position states ($\vec{x}_{a_{1,2}}$). So it is wise to partition the state vector into

$$\vec{x}_{cu} = \vec{x}_{1,2} \tag{50}$$

$$\vec{x}_{ck} = \vec{x}_{3,4,5,6} \tag{51}$$

$$\vec{\gamma_1}_{cu} = \vec{\gamma_1} \tag{52}$$

$$\vec{\gamma_1}_{ck} = [] \tag{53}$$

where [] denotes an empty vector. This yields the compound covariance matrix (in partitioned form)

$$\begin{aligned}
\boldsymbol{P} &= \begin{bmatrix} \boldsymbol{P}^{xx} & \boldsymbol{P}^{x\gamma_1} \\ \boldsymbol{P}^{\gamma_1 x} & \boldsymbol{P}^{\gamma_1\gamma_1} \end{bmatrix} \\
&= \left[\begin{array}{cc|cc} \boldsymbol{P}^{x_{cu}x_{cu}} & \boldsymbol{P}^{x_{cu}x_{ck}} & \boldsymbol{P}^{x_{cu}\gamma_{1cu}} & \boldsymbol{P}^{x_{cu}\gamma_{1ck}} \\ \boldsymbol{P}^{x_{ck}x_{cu}} & \boldsymbol{P}^{x_{ck}x_{ck}} & \boldsymbol{P}^{x_{ck}\gamma_{1cu}} & \boldsymbol{P}^{x_{ck}\gamma_{1ck}} \\ \hline \boldsymbol{P}^{\gamma_{1cu}x_{cu}} & \boldsymbol{P}^{\gamma_{1cu}x_{ck}} & \boldsymbol{P}^{\gamma_{1cu}\gamma_{1cu}} & \boldsymbol{P}^{\gamma_{1cu}\gamma_{1ck}} \\ \boldsymbol{P}^{\gamma_{1ck}x_{cu}} & \boldsymbol{P}^{\gamma_{1ck}x_{ck}} & \boldsymbol{P}^{\gamma_{1ck}\gamma_{1cu}} & \boldsymbol{P}^{\gamma_{1ck}\gamma_{1ck}} \end{array}\right] \\
&= \left[\begin{array}{cc|c} \boldsymbol{P}^{x_{cu}x_{cu}} & \boldsymbol{P}^{x_{cu}x_{ck}} & \boldsymbol{P}^{x_{cu}\gamma_{1cu}} \\ \boldsymbol{P}^{x_{ck}x_{cu}} & \boldsymbol{P}^{x_{ck}x_{ck}} & \boldsymbol{P}^{x_{ck}\gamma_{1cu}} \\ \hline \boldsymbol{P}^{\gamma_{1cu}x_{cu}} & \boldsymbol{P}^{\gamma_{1cu}x_{ck}} & \boldsymbol{P}^{\gamma_{1cu}\gamma_{1cu}} \end{array}\right],
\end{aligned} \tag{54}$$

where $\boldsymbol{P}^{\gamma_{1cu}x_{ck}} = \boldsymbol{P}^{x_{ck}\gamma_{1cu}} \approx \boldsymbol{0}$, which is deduced from the correlation matrix plot. This leads to the modified covariance matrices

$$\boldsymbol{P}^{xx,\alpha}_{k|k-1} = \begin{bmatrix} \frac{1}{1-\alpha}\boldsymbol{P}^{x_{1:2}x_{1:2}} & \boldsymbol{P}^{x_{1:2}x_{3:6}} \\ \boldsymbol{P}^{x_{3:6}x_{1:2}} & \boldsymbol{P}^{x_{3:6}x_{3:6}} \end{bmatrix}_{k|k-1} \tag{55}$$

$$\boldsymbol{P}^{\gamma_1\gamma_1,\alpha}_{k} = \frac{1}{\alpha}\boldsymbol{P}^{\gamma_1\gamma_1}_{k} \tag{56}$$

for the KF covariance equations

$$\boldsymbol{P}^{xx}_{k|k} = \left[\boldsymbol{H}^T \begin{bmatrix} \boldsymbol{P}^{xx,\alpha}_{k|k-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}^{\gamma\gamma,\alpha}_{k} \end{bmatrix}^{-1} \boldsymbol{H}\right]^{-1} \tag{57}$$

$$\boldsymbol{P}^{xx}_{k+1|k} = \boldsymbol{\Phi}\boldsymbol{P}^{xx}_{k|k}\boldsymbol{\Phi}^T + \boldsymbol{G}\boldsymbol{P}^{\kappa\kappa}_{k}\boldsymbol{G}^T, \tag{58}$$

where again $\boldsymbol{P}^{\gamma\gamma,\alpha} = \mathrm{blkdiag}(\boldsymbol{P}^{\gamma_1\gamma_1,\alpha}, \boldsymbol{P}^{\gamma_2\gamma_2})$.

Discussion of Results

Figure 4 depicts the results of the position estimation accuracy of the three estimators as well as the determinant and trace of the state covariance matrix $\boldsymbol{P}^{xx}_{k|k}$ for the CI and the PCI filter. To allow

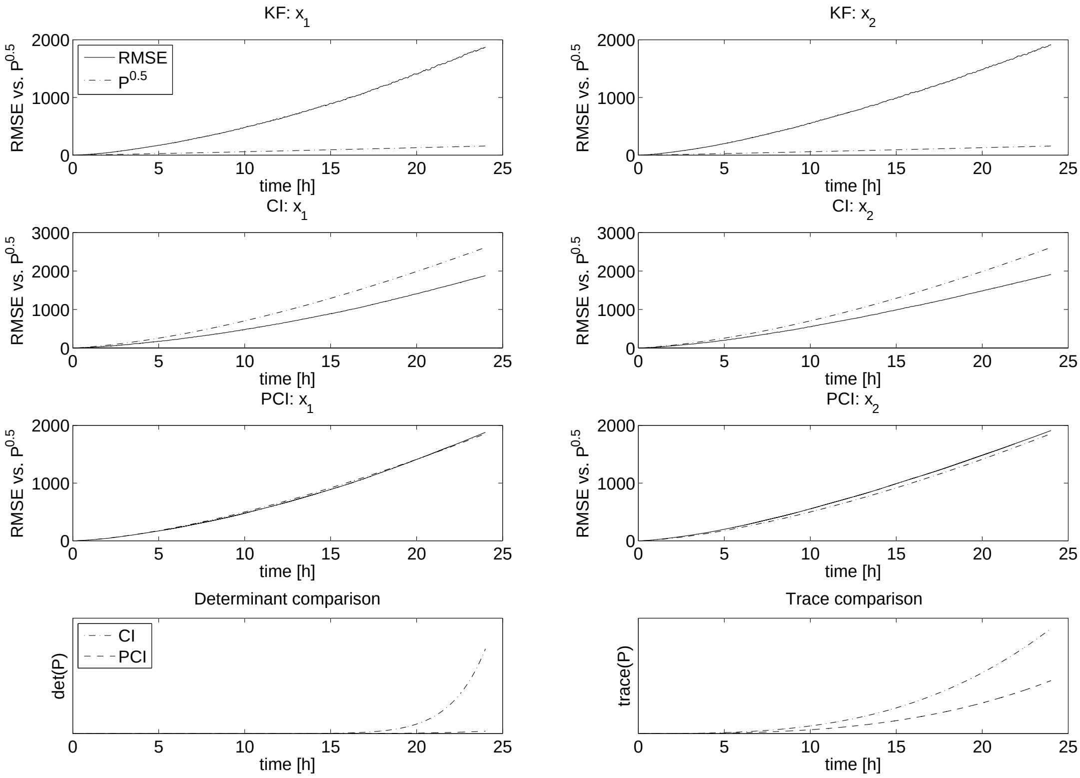


Figure 4: Filter results KF vs. CI vs. PCI for $MC = 100$ simulations:
$RMSE$ vs. $\boldsymbol{P}^{0.5} = \sqrt{\mathrm{diag}(\boldsymbol{P}^{xx}_{1:K,1:K})}$.

a fair comparison, the results were averaged over 100 independent Monte Carlo runs. Parameters used can be found in appendix B.

The root mean squared error

$$RMSE = \sqrt{\frac{1}{MC}\sum_{i=1}^{MC}(\vec{x}_{1:K} - \hat{\vec{x}}_{1:K})^2} \tag{59}$$

is compared to the Filter standard deviation. The estimation accuracy of all three estimators is comparable, anyhow the results of the Filter covariance estimates differ substantially.

The severe disadvantage of the KF approach is, that the resulting covariance matrix is inconsistent, i.e. the assumed error variance is smaller than the true variance. It is clear, that for navigation purposes this behavior is completely unacceptable, since over-confidence might lead to decisions of the vessel crew than can cause severe damage to material and human.

The results of the CI as well as the PCI approach are consistent, i.e. the true error variance is smaller than the assumed variance. Thus, both estimators give acceptable results. Anyhow, accuracy is desired in filtering and from the (P)CI error plots and the trace and determinant plots we can see, that the PCI approach outperforms the CI approach.

4 Conclusion

The problem of fusing two pieces of information, where parts of the correlation in the error estimates are unknown, has been considered in this paper. This problem has been solved by partitioning the error vectors into a unknown correlated and known correlated part. The potentially correlated parts of the error estimate covariance have been scaled in such way that the joint covariance matrix provides a tight upper bound for the set of all possible true cross-covariances.

The new bound generalizes known results for unknown cross-correlation between the complete error estimates, which gives too conservative results when parts of the correlation are known.

The benefit of PCI in Kalman Filtering problems where cross-correlations between the state vector and the measurement noise occur (arising from time-correlation in the noise) has been demonstrated. As result PCI allows more informative, yet consistent estimates to be computed than with standard CI.

References

[1] J.C. McGlone. *Manual of Photogrammetry*. ASPRS, 2004.

[2] S.J Julier and J.K. Uhlmann. *A Non-divergent Estimation Algorithm in the Presence of Unknown Correlations*. Proceedings of the American Control Conference, vol. 4, p. 2369 - 2373, 1997.

[3] U.D. Hanebeck and K. Briechle. *New Results for Stochastic Prediction and Filtering with Unknown Correlations*. Proceedings of the IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems, p. 147-152, 2001.

[4] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2008.

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

[6] G. Welch and G. Bishop. *An Introduction to the Kalman Filter*. University of North Carolina at Chapel Hill, 2006.

[7] R.E. Kalman. *A New Approach to Linear Filtering and Prediction Problems*. Journal of Basic Engineering, vol. 82, 1960.

[8] A. Petersen and R. Koch. *Statistical Analysis of Kalman Filters by Conversion to Gauss-Helmert Models with Applications to Process Noise Estimation*. Proceedings of ICPR2010, Istanbul, Turkey, 2010.

[9] M.S. Grewal and A.P. Andrews. *Kalman Filtering: Theory and Practice Using MATLAB, Second Edition*. John Wiley & Sons, Inc, 2001.

A Consistency

Consider that two estimates for random variables $\vec{x} \in \mathbb{R}^n$ and $\vec{y} \in \mathbb{R}^m$ are given by their assumed means and covariances $\bar{x}$, $\bar{y}$, $\boldsymbol{P}^{xx}$, and $\boldsymbol{P}^{yy}$ respectively. The deviations of the assumed means from the true values are $\tilde{x} \triangleq \vec{x} - \bar{x}$ and $\tilde{y} \triangleq \vec{y} - \bar{y}$. In general these deviations are not zero mean, and the mean squared error and cross correlations are

$$\boldsymbol{P}^{\tilde{x}\tilde{x}} = \mathbb{E}[\tilde{x}\tilde{x}^T] \tag{60}$$

$$\boldsymbol{P}^{\tilde{y}\tilde{y}} = \mathbb{E}[\tilde{y}\tilde{y}^T] \tag{61}$$

$$\boldsymbol{P}^{\tilde{x}\tilde{y}} = \mathbb{E}[\tilde{x}\tilde{y}^T]. \tag{62}$$

The actual values are not known, but rather are approximated by the values $\boldsymbol{P}^{xx}$ and $\boldsymbol{P}^{yy}$. These approximations are consistent if

$$\boldsymbol{P}^{xx} - \boldsymbol{P}^{\tilde{x}\tilde{x}} \geq \mathbf{0} \tag{63}$$

$$\boldsymbol{P}^{yy} - \boldsymbol{P}^{\tilde{y}\tilde{y}} \geq \mathbf{0}, \tag{64}$$

where, in general, for two positive definite matrices $\boldsymbol{V}$ and $\boldsymbol{Z}$, an expression of the form $\boldsymbol{V} > \boldsymbol{Z}$ ($\boldsymbol{V} \geq \boldsymbol{Z}$) is interpreted as $\boldsymbol{V} - \boldsymbol{Z}$ positive definite (positive semidefinite). Note that the cross-correlation matrix $\boldsymbol{P}^{\tilde{x}\tilde{y}}$ is unknown and will not - in general - be $\mathbf{0}$. This definition conforms to the standard definition of consistency used in [5].

B Simulation Parameters

For the second example, we chose the following simulation parameters:

$$\Delta t = 1\,\mathrm{s} \tag{65}$$

$$\boldsymbol{P}^{\kappa\kappa} = \mathrm{diag}([(0.2\,{}^{m}\!/\mathrm{s})^2\ (0.1\,{}^{m}\!/\mathrm{s})^2\ (0.05\,{}^{m}\!/\mathrm{s})^2\ (0.05\,{}^{m}\!/\mathrm{s})^2]) \tag{66}$$

$$\boldsymbol{P}^{\mu\mu} = \boldsymbol{I}^{2\times 2}(0.0001265\,{}^{m}\!/\mathrm{s})^2 \tag{67}$$

$$\boldsymbol{P}^{\gamma_2\gamma_2} = \boldsymbol{I}^{2\times 2}(0.1\,{}^{m}\!/\mathrm{s})^2 \tag{68}$$

The variance of the INS error has been chosen such that the position error has a 1σ-value of $1\,\mathrm{nmi} \approx 1852\,\mathrm{m}$ after 24 hours, which is typical for accurate INS systems.