

Pose Fusion with Biased and Dependent Data for Automated Driving

Christian Merfels* Tobias Riemenschneider[†]
Cyrill Stachniss[‡]

June 15, 2016

Automated driving requires at all times robust pose estimates. Multiple redundant and complementary localization systems are therefore installed on most automated vehicles. This work proposes a novel pose fusion approach for combining multiple pose estimates into a single estimate. Our technique is based on a nonlinear least squares optimization of a pose graph that consists of odometry and global pose measurements. We provide effective methods to estimate biases of different pose sources and to fuse correlated pose sources guaranteeing conservative estimates. The proposed approaches are evaluated on simulated data as well as on data gathered with a prototype vehicle. The results suggest that we can substantially reduce the impact of a time-varying bias on a GPS receiver and precisely estimate the vehicle poses.

1 Introduction

Advanced driver assistance systems and automated driving functionality rely on robust and highly available pose estimates. In the past, different localization systems have been proposed that are usually tailored to a specific sensor set, which typically includes Global Positioning System (GPS), vision, or lidar sensors. Different types of localization systems and sensors have individual failure modes such as satellite-denied regions for GPS or darkness for visual systems, in which the performance degrades substantially or the system fails to provide a reasonable estimate of the vehicle's pose. We approach this problem by proposing a generic pose

*Christian Merfels is with Volkswagen Group Research, Wolfsburg, and Institute of Geodesy and Geoinformation, University of Bonn, Germany. christian.merfels@volkswagen.de.

[†]Tobias Riemenschneider is with c4c Engineering, Braunschweig.

[‡]Cyrill Stachniss is with Institute of Geodesy and Geoinformation, University of Bonn.

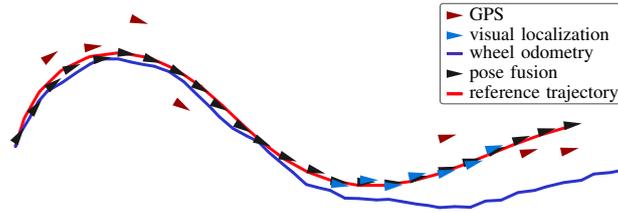


Figure 1: A coarse localization (GPS, red triangles), a precise but only temporary available localization (visual localization, blue triangles), and odometry as dead reckoning trajectory (wheel odometry, blue) are used to estimate the true trajectory (reference trajectory, red) of a vehicle. The estimated poses are shown as black triangles (pose fusion): the goal is to approximate the unknown red line as closely as possible with the black triangles.

fusion approach, which is able to effectively merge several relative and global pose sources. This enables the combination of orthogonal or redundant pose sources and has the potential to increase the availability, reliability, and accuracy of the resulting localization. Fig. 1 illustrates the key idea behind the pose fusion: pose estimates from multiple sources are combined into a single estimate by formulating the fusion as a joint optimization problem. The optimized trajectory then provides the fused estimate of the current state.

Besides improving the availability, reliability, and accuracy of the localization, our pose fusion also enables new use cases. Some localization systems only work in specific areas, such as GPS in outdoor scenarios or marker-based visual systems in particularly prepared parking garages, and our proposed pose fusion enables the seamless transition from one area to the other.

The main contribution of this paper is a generic pose fusion approach, in which we consider pose fusion as multi-sensor data fusion for pose estimation. We view pose estimation as the optimization problem of computing the set of recent poses that define the most likely trajectory given odometry and absolute pose measurements. To this end, we propose a nonlinear least squares estimation implemented in a sliding window fashion for efficient and effective pose estimation. In this system, the measurements are incorporated in a generic way so as to allow us to plug in third-party localization modules for which source code is unavailable. As a consequence, we treat common effects such as biased measurements or correlated noise between pose sources without knowing their cause. We employ a Covariance Intersection (CI) framework to handle dependent pose measurements and to prevent overconfidence. For pose measurements with systematic error components, we develop an online bias estimation scheme and demonstrate the effective error reduction.

In brief, the key contributions of this paper are:

- the presentation of a pose fusion concept implemented as a sliding window pose graph;
- the conservative treatment of dependent information such as correlated errors between pose sources by employing a CI framework;
- an online bias estimation technique that is fast, effective, and simple to implement.

2 Related Work

Multi-sensor data fusion applied to localization enables the integration of observations from multiple sources to estimate the pose of the system. Algorithms of this field are commonly designed to fuse a specific set of pose sources, such as GPS and inertial measurement units (IMU), by conventionally employing a Kalman or particle filter [1–3]. This work differs from these approaches in that we consider a generic set of pose sources instead of specializing the fusion for a specific set. Extended Kalman Filter (EKF)-based approaches have also been proposed for generic pose fusion [4,5]. Filtering approaches are frequently used for online estimation problems but we argue in Section 3.2 that we generally expect smoothing approaches to outperform them due to the possibility of relinearization.

Ranganathan *et al.* [6] propose a fixed-lag smoothing algorithm for pose estimation. The technique is based on square root smoothing and allows incorporating delayed and out-of-sequence measurements in their pose graph representation. The authors model GPS errors as unbiased Gaussian distributions and discard unsuitable GPS measurements. Sibley *et al.* [7] introduce the concept of sliding window filters to estimate surface structure with a stereo camera during a planetary entry, descent, and landing scenario. We use a similar methodology but apply it to the problem of generic pose fusion, which leads to different challenges such as bias estimation and treatment of dependent information.

Bias estimation in the context of localization is treated in the literature primarily for GPS systems. A common approach consists in augmenting the state vector of a filter to allow for more sophisticated error models and correcting the bias by a second pose source. Jo *et al.* [8] correct systematic GPS errors that change slowly over time by comparing visual observations of the road structure to a given road map database. Laneurit *et al.* [9] empirically model GPS error as an additive Gaussian distribution plus a time-dependent bias and white noise. They estimate the bias by computing the difference of the sensor fusion result to the GPS measurement. Significant bias changes are determined by testing whether the prediction based on the last GPS position lies within the one sigma error ellipse of the current measurement. Tao *et al.* [10] construct a first order autoregressive model for GPS error estimation. While this model captures the autocorrelation of the bias, strong bias variations in the form of position jumps are only treated by rejecting the corresponding GPS fixes. The authors further compare visual observations of lane markings to a map in order to correct GPS errors. These approaches specifically define a second, unbiased pose estimation to subsequently eliminate the GPS bias. Our bias estimation scheme is inspired by the same idea but we generalize it to work with any unbiased pose source.

Another challenge besides accounting for biased data in generic pose fusion algorithms is the treatment of dependent information. Ignoring the effects of correlated noise between pose sources can be detrimental for any multi-sensor fusion approach because of overconfident estimates, corrupt uncertainty estimates, and estimator divergence. Generic pose fusion systems do not know about the specific source of correlation and have to perform *fusion under unknown correlation*. One approach to this is ellipsoidal intersection [11], which maximizes the common information given that the sources share a common prior. Another method is CI [12], which combines estimates with unknown error correlations. It provides a conservative estimate of the actual mean square error matrix. CI has been employed for a range of applications, including graph-based [13] and filtering-based [14] Simultaneous Localization and Mapping

(SLAM). Reinhardt *et al.* [15] propose closed-form solutions for two- and three-dimensional matrices, on top of which we design our treatment of dependent information as detailed in Section 3.4.

3 Pose Fusion

Our multi-sensor data fusion approach combines pose measurements from multiple pose sources to compute the current best pose estimate. Odometry and global pose sources have orthogonal strengths and weaknesses. On the one hand, odometry measurements are usually available at high frequencies and do not require a priori knowledge about the environment. However, they accumulate drift with growing distance and they are not globally referenced. On the other hand, global pose measurements are globally referenced and their error is independent of the covered distance. However, they are usually only available at low frequency and require a priori knowledge about or preparation of the environment (such as maps or satellite placement). Combining both types of measurements in our sliding window pose fusion allows us to estimate a trajectory which is both globally referenced and locally smooth.

This requires merging poses from two different coordinate systems, which are detailed in Section 3.1. We present the estimation procedure in Section 3.2 and subsequently approach two specific problems: the online estimation of systematic biases in Section 3.3 and the treatment of correlated noise between input sources in Section 3.4.

3.1 Coordinate Systems

The pose fusion combines measurements in a global coordinate system with odometry measurements in a vehicle reference frame. The global coordinate system of the pose fusion is a two-dimensional Cartesian coordinate system for which we choose the Universal Transverse Mercator (UTM) coordinate system. Poses in this coordinate system are dubbed *global poses*. We convert different but prevalent global coordinates (such as pairs of latitude and longitude in the World Geodetic System (WGS)) to UTM to support a wide range of input sources. The vehicle reference frame is defined according to ISO 8855 [16]. It is also commonly referred to as body frame or body coordinate system.

3.2 Sliding Window Graph-based Pose Fusion

We formulate the pose fusion as the problem of estimating the most likely trajectory given a set of odometry and global pose measurements. We assume the noise of these measurements to be additive, white, and normally distributed with zero mean. In Section 3.3 we detail how to minimize systematic biases if the mean is not zero. The output of the pose fusion is the most recent state of the estimated trajectory. The underlying estimation procedure relies on nonlinear least squares optimization. Our approach exploits the state-of-the-art graph optimization framework g2o [17] and for the most part, we adopt the notation used by Kümmerle *et al.*

Let $\mathbf{x} = (\mathbf{x}_1^\top, \dots, \mathbf{x}_m^\top)^\top$ be the state vector. The key idea in nonlinear least squares estimation is that given a set of measurements, where $z_{i,j}$ is the mean and $\mathbf{\Omega}_{i,j}$ the inverse measurement covariance matrix of a single measurement relating \mathbf{x}_i to \mathbf{x}_j (with \mathcal{C} being all

pairs of indices for which a measurement is available), least squares estimation seeks the state

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \sum_{(i,j) \in \mathcal{C}} \mathbf{e}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{z}_{i,j})^\top \boldsymbol{\Omega}_{i,j} \mathbf{e}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{z}_{i,j}) \quad (1)$$

that best explains all measurements given the ℓ_2 norm. The vector error function $\mathbf{e}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{z}_{i,j})$ measures how well the constraint from the measurement $\mathbf{z}_{i,j}$ is satisfied. Solving (1) requires approximating the solution of the current step and refining it by successive iterations. Only the states that are currently contained in the state vector can be relinearized and thus refined in the Gauss-Newton or Levenberg-Marquardt iterations. For more details, we refer the reader to Kümmerle *et al.* [17].

Optimizing the entire trajectory is commonly referred to as batch estimation. It leads to a maximum a posteriori (MAP) estimate over the joint probability of vehicle poses and produces a statistically optimal result. A drawback of this approach is that the state vector grows unboundedly over time, thus limiting its online applicability. The EKF approaches this issue by restricting the state vector to the most recent state, hence collapsing the trajectory estimation into a single pose estimation problem. This, however, prevents relinearization of previous states as they are already marginalized out. Also, the current state is not relinearized and the Jacobians are evaluated only once. The Iterated EKF solves this drawback by iterating and relinearizing the solution of the current state, thus converging to the optimal state estimate. However, older states are not explicitly available due to the involved marginalization, which implies that they cannot be relinearized and refined anymore. The Sliding Window Filter [7] weakens this drawback by extending the state vector to the set of the most recent states. It can be seen as an Iterated EKF with an augmented state vector because the Iterated Kalman Filter update is for many problem instances algebraically identical to the Gauss-Newton method [18] when both the prediction and update steps are iterated [19].

Fig. 2 illustrates the relationship between the (Iterated) EKF, the Sliding Window Filter, and batch estimation. Motivated by the need for a powerful estimator and constrained by the requirement of an online solution, we design our pose estimation with the same key concepts as a Sliding Window Filter.

We represent our nonlinear least square problem using a pose graph. The reduction to a sliding window problem fixes \mathbf{x} to the M most recent states such that $\mathbf{x} = (\mathbf{x}_{t-M+1}^\top, \dots, \mathbf{x}_t^\top)^\top$. This limitation of the number of states results in bounded runtime requirements. Old states are marginalized out and new ones are appended to the front of the graph. State variables are represented by *hidden nodes*. Global pose measurements are encoded in *observed nodes* and linked to hidden nodes via edges. Odometry measurements result in edges between hidden nodes. The joint optimization seeks the set of hidden nodes that best explain all pose and odometry measurements.

3.3 Estimation of Systematic Biases

As mentioned in Section 3.2, we assume the noise of the input pose measurements for the pose fusion to be ideally additive, white, and normally distributed with zero mean. However,

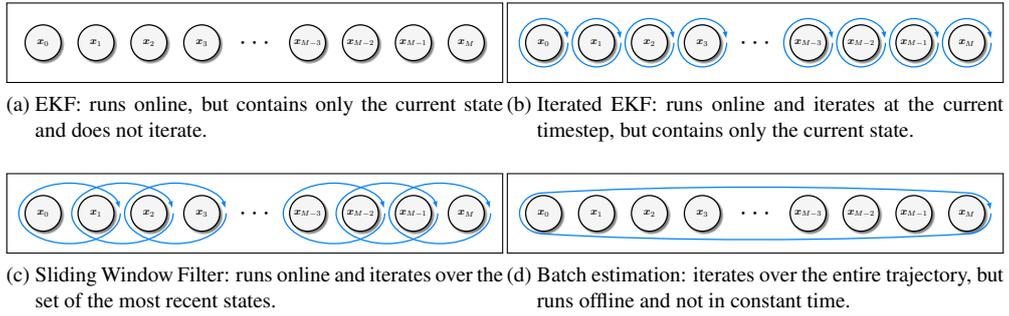


Figure 2: Comparison of iterative state estimation techniques. The figure is inspired by Barfoot [19, p. 145].

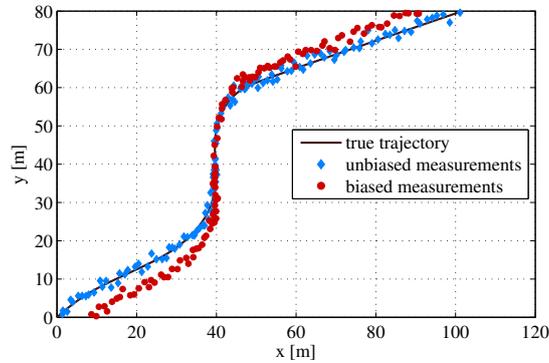


Figure 3: Illustration of the effects of a time-varying bias. The true trajectory (black line) is measured by an unbiased (blue diamonds) and a biased (red dots) pose source.

a systematic bias in the measurements of a global pose source is common [8–10] and results in a mean error unequal to zero when compared to the true trajectory. Fig. 3 illustrates the effects of unbiased and biased pose measurements. Systematic biases of this kind result in detrimental effects for the pose fusion approach: the mean error of the fused estimate will generally be unequal to zero and the estimator potentially diverges. We propose an effective technique to estimate systematic biases online. This estimation enables pose sources to better satisfy the noise assumptions by reducing the mean error, thus leading to more consistent and precise estimates.

Systematic biases are often varying over time. The duration, in which a systematic bias stays roughly constant, depends on the specific sensor, its field of application, and environmental conditions. A prominent example of a class of global pose sources, which commonly suffers from time-varying biases, is GPS. Clock errors and multipath effects are common sources for their biases [8]. The bias might for example stay roughly constant as long as the GPS receiver sees the same amount of satellites or suffers from the same multipath effects. Other, rather

sensor-independent, sources of systematic error include calibration and time synchronization issues.

Our approach does not assume any knowledge about the implementation details of the input sources, which makes it challenging to eliminate the unknown source of biases. We therefore propose a generic bias estimation method. To this end, we analyze the global pose measurements z^b with inverse measurement covariance matrices Ω^b from a biased source and compare them to a source which is noisy, but produces unbiased measurements z^k with inverse measurement covariance matrices Ω^k . More formally, we model the noise as

$$z_i^k = p_i + v_i^k \quad (2)$$

$$z_i^b = p_i + v_i^b + c_i^b, \quad (3)$$

where p_i is the true pose for the i -th measurement, v^k and v^b are white noise such that $v^k \sim \mathcal{N}(\mathbf{0}, (\Omega^k)^{-1})$ and $v^b \sim \mathcal{N}(\mathbf{0}, (\Omega^b)^{-1})$, and c_i^b denotes the bias. Equation (2) is a common noise model in probabilistic robotics and we will show how to gain knowledge about c_i^b by observing the corresponding unbiased pose source.

The key idea of the bias estimation is that the difference of the mean estimates of a biased source compared to an unbiased source is roughly equal to the respective bias. Estimating the bias in a sliding window of length s , such that the estimated bias $\hat{c}_i^b(s)$ for the i -th measurement is a function of s , allows it to adapt to time-varying biases. The estimate of the bias $\hat{c}_i^b(s)$ also depends on the unbiased pose source k , but we omit the additional index for the sake of notational clarity. We refine this method by computing the mean errors as weighted averages such that pose measurements with higher variances account for less impact. In total, we define the estimated bias $\hat{c}_i^b(s)$ accordingly as

$$\hat{c}_i^b(s) = \frac{1}{s} \frac{1}{\sum_{j=i-s}^i \Omega_j^k} \sum_{j=i-s}^i \Omega_j^k z_j^k - \frac{1}{s} \frac{1}{\sum_{j=i-s}^i \Omega_j^k} \sum_{j=i-s}^i \Omega_j^k z_j^b \quad (4)$$

$$= \frac{1}{s} \frac{1}{\sum_{j=i-s}^i \Omega_j^k} \sum_{j=i-s}^i \Omega_j^k (z_j^k - z_j^b). \quad (5)$$

Fig. 4 illustrates the application of this method. This technique is straight forward and effective.

A key issue is the determination of the size s of the sliding window. On the one hand, the estimated bias adapts too slowly to the true time-varying bias if the window size is too large. On the other hand, we violate the assumption, that the unbiased pose source has a zero-mean error over the sliding window, if the sliding window size is too small. A sliding window size that is too small additionally leads to a significant correlation of the error terms of the biased and unbiased source. In Section 3.4.1 we provide an insight on the order of the error of the uncertainty estimate given a certain cross-correlation and thus allows us to gauge what cross-correlation is tolerable.

We formulate the search for the optimal sliding window size s^* as an offline optimization over a dataset with L data points per pose source. This dataset contains unbiased measurements z^k , biased measurements z^b , and ground truth measurements p . The optimal sliding

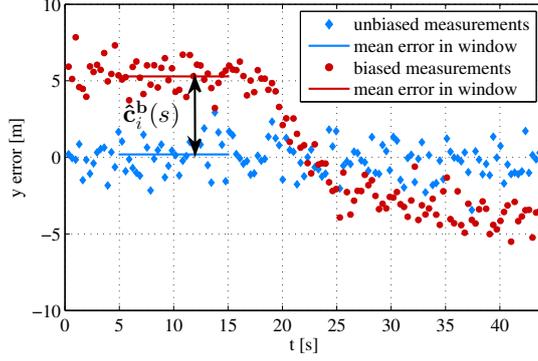


Figure 4: Measurement error over time of the unbiased (blue diamonds) and biased (red dots) pose sources. The difference of their mean errors in a sliding window of 10 s leads to the computation of $\hat{c}_i^b(s) \approx 5$ m.

window size s^* is given by

$$s^* = \arg \min_s \frac{1}{L} \sum_{i=1}^L \|(\mathbf{z}_i^b - \hat{\mathbf{c}}_i^b(s)) - \mathbf{p}_i\|, \quad (6)$$

$$s.t. \quad \rho(\mathbf{z}^b - \mathbf{p}, \mathbf{z}^k - \mathbf{p}) \leq \rho_{\text{tol}}, \quad (7)$$

$$\tau(\mathbf{z}^k, \mathbf{p}, s) \leq \tau_{\text{tol}}. \quad (8)$$

The first condition in (7) is true if the errors of the bias-corrected source and the unbiased source are correlated less than a threshold ρ_{tol} . For this, $\rho(\cdot, \cdot)$ is an auxiliary function which computes the correlation between two input vectors. The second condition in (8) is true if the sliding window is large enough to justify the assumption of a zero-mean error of the unbiased source. For this, $\tau(\cdot, \cdot, \cdot)$ is an auxiliary function such that 95% of all terms $\frac{1}{s} \sum_{j=i-s}^i \|\Omega_j^k \mathbf{z}_j^k - \mathbf{p}_j\|$ for $i = 1, \dots, L$ are less or equal than its function value. The usage of the 95% percentile robustifies the function against outliers. As this is an offline optimization without timing requirements, we content ourselves to solve (6) with an exhaustive search over all sliding window sizes of interest.

Fig. 5 shows the output of the optimization over a given dataset. The optimal sliding window size in terms of mean GPS error should be equal to one because the unbiased source should be much more precise than GPS. However, this makes the GPS signal equal to the unbiased signal and yields a maximum correlation $\rho(\mathbf{z}^b - \mathbf{p}, \mathbf{z}^k - \mathbf{p})$ of their error terms. Also, that sliding window is too small to justify the assumption of a near-zero mean of the error of the unbiased source as indicated by the value of $\tau(\mathbf{z}^k, \mathbf{p}, s)$. In total, this approach allows for choosing the optimal tradeoff for the sliding window size s^* .

This approach can easily be extended to incorporate multiple unbiased pose sources by estimating $\hat{\mathbf{c}}^b(s)$ for each of these sources. In general, we assume that the majority of sources are unbiased. We can easily determine whether a source is biased by observing all pose sources over a certain period of time and performing a simple version of random sample consensus.

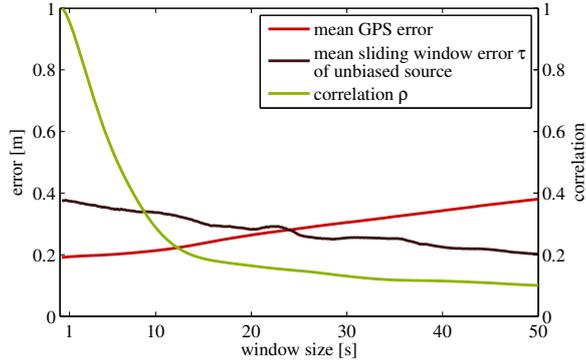


Figure 5: We find the optimal sliding window size by minimizing the mean GPS error (see (6)) while limiting the correlation of the errors of GPS and the unbiased source (see (7)). These two objectives conflict because the mean GPS error increases (as the adaption of the sliding window is too slow) and the correlation decreases to random correlations with larger window sizes. Additionally, we ensure that the sliding window is large enough by inspecting the maximum mean error over all sliding windows (see (8)).

The search for the optimal sliding window size s^* takes into account the correlation between the error of the unbiased and bias-corrected measurements. We highlight the importance of analyzing this kind of measurement correlation in the next section.

3.4 Correlated Noises between Pose Sources

The noise of different pose sources can in general be correlated. Not accounting for correlated noise leads to divergent and overconfident pose estimates. Noise is correlated between pose sources when for example the same sensor or map data is being used in different algorithms or when the same algorithm runs on two physically different sensors (e.g., two GPS receivers). Two of the major sources are *common process noise* and *common prior information* [15]. Both lead to common noise that potentially influences all affected sources. Ideally, we would be able to eliminate the source of the common noise, but as we perform generic pose fusion we lack the insight what this common source is. Therefore, we try to minimize its impact in the fusion process by using CI to produce conservative estimates.

In the following, we first elaborate on the optimal fusion which requires knowledge about the true correlation, subsequently detail the naive fusion which ignores the correlation, and finally investigate two ways of implementing CI. Fig. 6 motivates why naive fusion is sub-optimal as its result is overconfident. The true correlation between C_1 and C_2 is unknown and randomly drawn positive correlations (displayed in gray) show that the two CI methods (trace minimization with $\omega = 0.69$ and determinant minimization with $\omega = 0.29$) provide conservative bounds.

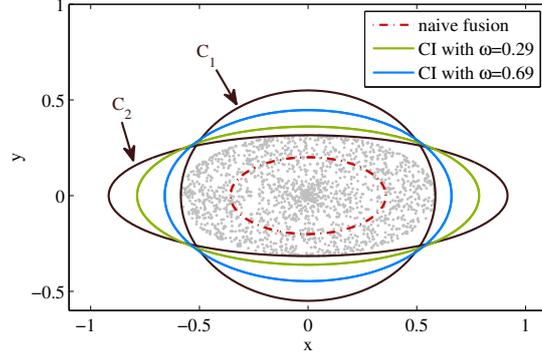


Figure 6: C_1 and C_2 are two example covariance matrices with unknown correlation that we want to fuse. CI covariance ellipses for trace ($\omega = 0.69$) and determinant ($\omega = 0.29$) minimization. True covariance values for randomly chosen positive correlations are displayed as gray dots. The naive fusion (dashed, red) is overconfident.

3.4.1 Optimal and Naive Fusion of Correlated Estimates

Two normally distributed states x_1 and x_2 with covariance matrices C_1 and C_2 can generally be combined by

$$x = A_1 x_1 + A_2 x_2 \quad (9)$$

where A_1 and A_2 denote quadratic matrices. To ensure that the mean error is equal to zero, $A_1 + A_2 = I$ must hold [20]. The variance law of error propagation implies for the resulting covariance matrix

$$C = (A_1 \ A_2) \begin{pmatrix} C_1 & C_{12} \\ C_{12}^\top & C_2 \end{pmatrix} \begin{pmatrix} A_1^\top \\ A_2^\top \end{pmatrix} \quad (10)$$

$$= A_1 C_1 A_1^\top + A_1 C_{12} A_2^\top + A_2 C_{12}^\top A_1^\top + A_2 C_2 A_2^\top. \quad (11)$$

To obtain a fusion with a small covariance matrix, A_1 and A_2 are set such that C is optimal in some sense, e.g., minimal trace or determinant. The optimal solution with respect to trace minimization yields

$$C = \left[(I \ I) \begin{pmatrix} C_1 & C_{12} \\ C_{12}^\top & C_2 \end{pmatrix}^{-1} \begin{pmatrix} I \\ I \end{pmatrix} \right]^{-1}, \quad (12)$$

see [20]. This is the covariance of the *optimal fusion*. However, it requires knowledge about the cross-correlation matrix C_{12} . If it is unknown, the *naive approach* is to assume $C_{12} = 0$. In that case (12) can be simplified to

$$C = (C_1^{-1} + C_2^{-1})^{-1} \quad (13)$$

which corresponds to the Kalman gain [20]. This, however, leads to overconfident estimates if $C_{12} > 0$, as illustrated in Fig. 6. The error of the naive fusion can be estimated from

(11) by $\|\mathbf{A}_1\mathbf{C}_{12}\mathbf{A}_2^\top + \mathbf{A}_2\mathbf{C}_{12}^\top\mathbf{A}_1^\top\|$. If the signals correlate with a correlation coefficient ρ , i.e., $\mathbf{C}_{12} = \rho\mathbf{I}$, the error is linear in ρ . If we can find an upper bound for the cross-correlation, we can decide whether the error is small and tolerable or if we need to take the cross-correlation into account. In general however, we need a method to produce conservative uncertainty estimates without knowledge of their correlation. For this, we detail our CI framework in Section 3.4.2.

3.4.2 Covariance Intersection

CI is the optimal algorithm to fuse two estimates $\mathbf{x}_1, \mathbf{x}_2$ with associated covariance matrices $\mathbf{C}_1, \mathbf{C}_2$ if the cross-correlations \mathbf{C}_{12} between their errors are unknown [21]. The resulting covariance matrix \mathbf{C}^ω and the fused state \mathbf{x}^ω are computed according to

$$\mathbf{C}^\omega = (\omega\mathbf{C}_1^{-1} + (1-\omega)\mathbf{C}_2^{-1})^{-1} \quad (14)$$

$$\mathbf{x}^\omega = \mathbf{C}^\omega (\omega\mathbf{C}_1^{-1}\mathbf{x}_1 + (1-\omega)\mathbf{C}_2^{-1}\mathbf{x}_2) \quad (15)$$

where $\omega \in [0, 1]$. The parameter ω is chosen in such a way that the covariance matrix \mathbf{C}^ω is minimized regarding a given objective function J such that

$$\omega^* = \arg \min_{\omega \in [0,1]} J(\mathbf{C}^\omega) \quad (16)$$

in order to minimize the upper bound of the corresponding mean square error matrix. Common choices for J are the trace and determinant of \mathbf{C}^ω . We evaluate them on simulated data in Section 4.1.

Reinhardt *et al.* [15] derive explicit formulas for the CI optimization using joint diagonalization. We briefly introduce the main concepts here, adopt their notation, and extend their algorithm for an important yet untreated corner case. The underlying optimization problems for these closed-form formulas are for determinant minimization given by

$$\omega^* = \arg \max_{\omega \in [0,1]} \prod_{i=1}^n \omega + (1-\omega)\bar{d}_i \quad (17)$$

and for trace minimization given by

$$\omega^* = \arg \min_{\omega \in [0,1]} \sum_{i=1}^n \frac{a_i}{\omega + (1-\omega)\bar{d}_i} \quad (18)$$

where the terms \bar{d}_i arise in the development of the closed-form solutions [15]. They play a crucial role as they are used to compute the auxiliary variables $\tilde{d}_i = \frac{\bar{d}_i}{1-\bar{d}_i}$, from which we can finally derive the candidate solutions to (16). We refer the reader for further details to Julier and Uhlmann [12, 14]. It is this definition of \tilde{d}_i for which we provide an extension for the corner case of $\bar{d}_i = 1$, where we derive how to compute ω^* .

3.4.3 Extension of CI to a Corner Case

The closed-form expressions [15] for trace and determinant minimization are derived independent of the properties of the covariance matrices to which they are applied. We give a simple example in which they cannot be applied, briefly highlight the underlying problem, and show how we can reduce the dimensionality of the optimization problem for both trace and determinant minimization so that we can again apply the analytic solutions.

The issue is that we can only set $\bar{d}_i = \frac{\bar{d}_i}{1-\bar{d}_i}$ if $\bar{d}_i \neq 1$. This, however, is not always the case as the following example shows. Without loss of generality, let

$$\mathbf{C}_1 = \begin{pmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \sigma^2 \end{pmatrix}, \quad \mathbf{C}_2 = \begin{pmatrix} \boldsymbol{\Sigma}_2 & \mathbf{0} \\ \mathbf{0} & \sigma^2 \end{pmatrix} \quad (19)$$

with $\sigma > 0$. We then verify that $\bar{d}_n = 1$. This is the case whenever at least one element of \mathbf{C}_1 and \mathbf{C}_2 is identical. For notational simplicity, we assume in the following that this is the last element in \mathbf{C}_1 and \mathbf{C}_2 . This important class of matrices constitutes the corner case of interest. We will now investigate how to treat the case $\bar{d}_i = 1$ for both trace and determinant minimization.

First, we are interested in determinant minimization and directly develop (17) to

$$\omega^* = \arg \max_{\omega} \prod_{i=1}^n \omega + (1 - \omega)\bar{d}_i \quad (20)$$

$$= \arg \max_{\omega} \prod_{i=1}^{n-1} \omega + (1 - \omega)\bar{d}_i. \quad (21)$$

Hence, the dimension of the CI problem can be reduced by one by solving the CI problem for $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$ instead of $(\mathbf{C}_1, \mathbf{C}_2)$.

Second, we turn to trace minimization and develop (18) to

$$\omega^* = \arg \max_{\omega} \sum_{i=1}^n \frac{a_i}{\omega + (1 - \omega)\bar{d}_i} \quad (22)$$

$$= \arg \max_{\omega} \sum_{i=1}^{n-1} \frac{a_i}{\omega + (1 - \omega)\bar{d}_i} + \frac{a_n}{1} \quad (23)$$

$$= \arg \max_{\omega} \sum_{i=1}^{n-1} \frac{a_i}{\omega + (1 - \omega)\bar{d}_i}. \quad (24)$$

Again we observe that this corner case leads to the reduction of the dimensionality of the problem. The key insight is that this problem is now well-defined and that we can apply our common set of CI tools to solve it.

This derivation is easily extended to the case that $\bar{d}_i = 1$ for multiple i . This reduces the set of viable candidates for ω^* by the amount of i for which $\bar{d}_i = 1$. In the extreme case that $\bar{d}_i = 1$ for all $i = 1, \dots, n$ it can easily be shown that $\mathbf{C}_1 = \mathbf{C}_2$, thus $\omega \in [0, 1]$ can be chosen arbitrarily.

4 Evaluation

The experimental section is designed to support our claims that our treatment of unknown noise correlations between different input sources leads to conservative estimates, that the bias estimation effectively reduces systematic biases, and that the resulting pose fusion is a precise pose estimation algorithm. We provide experiments on data gathered on a real prototype vehicle and on simulated data.

4.1 Correlated Noise between Pose Sources

The first experiment is designed to show that naive fusion produces overconfident estimates whereas CI produces conservative estimates. To this end, we generate a time series of random covariance matrices for a fixed cross-correlation $\rho = 0.6$ by sampling from a normal distribution, subsequently correlating the noise signals, and finally modifying them to have the desired mean and variance. With this data, we perform naive fusion, optimal fusion, and both (trace and determinant minimizing) CI fusions. Fig. 7 shows that the naive fusion is overly confident whereas CI produces uncertainty estimates equal to or greater than the optimal uncertainty. In this use case, CI with trace minimization produces similar results as CI with determinant minimization. The mean covariance of the optimal fusion is equal to $\sigma_{\text{optimal}} = 0.90$. Both CI with trace and determinant minimization estimate on average a covariance of $\sigma_{\text{CI}} = 0.98$, whereas the naive fusion estimates a mean covariance of $\sigma_{\text{naive}} = 0.70$. The proposed CI implementation therefore provides conservative estimates.

Furthermore, we evaluate how often the corner case of $\bar{d}_i = 1$ arises, for which we proved in Section 3.4.3 the reduction of the problem dimension. The frequency of this corner case depends on the type of data and the threshold for the difference of two floating point numbers below which we consider both numerically identical. For numerical stability, we choose this threshold to be 0.01. With typical GPS errors (as simulated this experiment) the corner case occurs in approximately 2–5% of all measurements.

4.2 Pose Fusion with Bias Removal

The following experiment is designed to show that the bias estimation effectively reduces the time-varying, systematic bias of GPS data gathered on a real prototype vehicle and that the pose fusion produces precise estimates. The vehicle is an Audi A6 Avant equipped with two GPS receivers of different quality, and is able to additionally measure its wheel odometry. The data was recorded on a route of about 16 km in rural and urban areas in Germany.

The errors of the two GPS receivers are significantly correlated with a factor of $\rho = 0.28$. We therefore apply our CI method to combine sources with correlated noises leading to a single virtual GPS source. However, as the covariance ellipse of the cheaper receiver is always magnitudes larger than that of the second receiver, the intersection of the covariance ellipses is equal to the covariance ellipse of the more expensive receiver. Therefore, the solution of CI degenerates in this special case simply to the estimate of the second receiver.

Subsequently, we fuse the GPS measurements with odometry data. In this process, we estimate the bias with the help of the presented sliding window technique by comparing against a third pose measurement source, which is afterwards ignored for the remainder of the pose

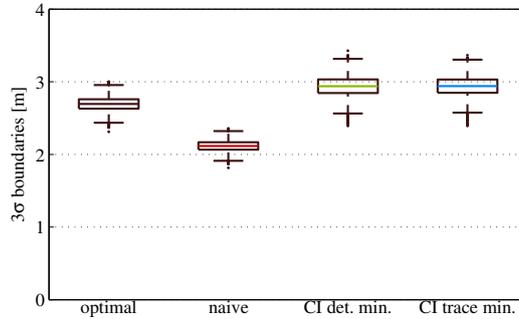


Figure 7: Fusion of correlated covariance matrices with different methods for simulated data. The 3σ uncertainty boundaries of the estimate after fusion are shown. The naive fusion ignores the correlation and thus produces overconfident covariance estimates. The optimal fusion requires knowledge about the unknown correlation ρ . CI generates conservative estimates.

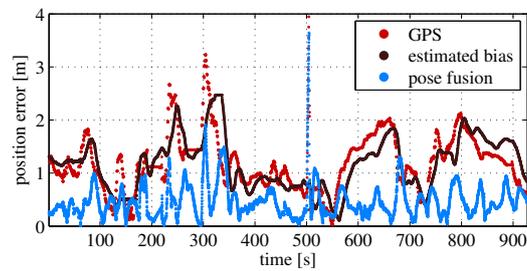


Figure 8: Position error of a GPS receiver in a prototype vehicle over time. The estimated bias is shown in black. It serves to reduce the systematic, time-varying bias of the GPS. The biased GPS (red) exhibits a higher position error than the bias-reduced pose fusion output (blue).

fusion for the sake of clarity. Fig. 8 shows the position errors over time of the GPS source, the estimated bias, and the error of the resulting output of the pose fusion.

The GPS signal exhibits a root mean square (RMS) position error of 1.31 m with a standard deviation of 0.52 m. The pose fusion drastically reduces the RMS position error to 0.57 m with a standard deviation of 0.33 m, thus showing the effectiveness of the bias estimation.

5 Conclusion

We presented a multi-sensor data fusion approach for pose estimation of an automated vehicle based on smoothing the most recent part of the trajectory. The underlying optimization problem is formulated as a sliding window nonlinear least squares minimization of the pose graph of odometry and global pose measurements. We incorporated these measurements in a generic way and account for correlated noise between input sources, for which we apply a Covariance Intersection framework. To minimize the effect of systematic biases, we estimate them online. We have evaluated these approaches with experiments on a real vehicle and demonstrated their effectiveness. Our experiments suggest that we can substantially reduce the impact of the time-varying bias of a GPS receiver by comparing it to an unbiased pose source. The key ideas of this paper can also be applied to other domains than automated driving.

References

- [1] C. Eling, L. Klingbeil, and H. Kuhlmann, “Real-time single-frequency GPS/MEMS-IMU attitude determination of lightweight UAVs,” *Sensors*, vol. 15, no. 10, pp. 26 212–26 235, 2015.
- [2] N. Steinhardt and S. Leinen, “Data Fusion for Precise Localization,” in *Handbook of Driver Assistance Systems*. Springer International Publishing, 2015, pp. 1–34.
- [3] A. Giremus, A. Doucet, A.-C. Escher, and J.-Y. Tournet, “Nonlinear filtering approaches for INS/GPS integration,” in *Proc. European Signal Process. Conf.* IEEE, 2005, pp. 873–876.
- [4] T. Moore and D. Stouch, “A Generalized Extended Kalman Filter Implementation for the Robot Operating System,” in *Proc. Int. Conf. Intelligent Autonomous Systems (IAS)*. Springer, 2016, pp. 335–348.
- [5] D. Ratasich and R. Grosu, “Generic Sensor Fusion Package for ROS,” in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Syst. (IROS)*, 2015, pp. 286–291.
- [6] A. Ranganathan, M. Kaess, and F. Dellaert, “Fast 3D Pose Estimation With Out-of-Sequence Measurements,” in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Syst. (IROS)*, 2007, pp. 2486–2493.
- [7] G. Sibley, L. Matthies, and G. Sukhatme, “Sliding Window Filter with Application to Planetary Landing,” *J. Field Robot.*, vol. 27, no. 5, pp. 587–608, 2010.

- [8] K. Jo, K. Chu, and M. Sunwoo, "GPS-bias correction for precise localization of autonomous vehicles," in *IEEE Intell. Vehicles Symp.*, 2013, pp. 636–641.
- [9] J. Laneurit, R. Chapuis, and F. Chausse, "Accurate Vehicle Positioning on a Numerical Map," *Int. J. Control, Automation and Syst.*, vol. 3, no. 1, pp. 15–31, 2005.
- [10] Z. Tao, P. Bonnifait, V. Frémont, and J. Ibañez-Guzman, "Mapping and Localization Using GPS, Lane Markings and Proprioceptive Sensors," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Syst. (IROS)*, 2013, pp. 406–412.
- [11] J. Sijs and M. Lazar, "State fusion with unknown correlation: Ellipsoidal intersection," *Automatica*, vol. 48, no. 8, pp. 1874–1878, 2012.
- [12] S. J. Julier and J. K. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations," in *Proc. Amer. Control Conf. (ACC)*, 1997, pp. 2369–2373.
- [13] B. Noack, S. J. Julier, and U. D. Hanebeck, "Treatment of Biased and Dependent Sensor Data in Graph-based SLAM," in *Proc. Int. Conf. Inform. Fusion (FUSION)*. IEEE, 2015, pp. 1862–1867.
- [14] S. J. Julier and J. K. Uhlmann, "Using Covariance Intersection for SLAM," *J. Robot. Auton. Syst.*, vol. 55, no. 1, pp. 3–20, 2007.
- [15] M. Reinhardt, B. Noack, and U. D. Hanebeck, "Closed-form Optimization of Covariance Intersection for Low-Dimensional Matrices," in *Proc. Int. Conf. Inform. Fusion (FUSION)*. IEEE, 2012, pp. 1891–1896.
- [16] ISO, "Road Vehicles – Vehicle Dynamics and Road-holding Ability – Vocabulary," International Organization for Standardization, Geneva, Switzerland, ISO 8855:2011, 2011.
- [17] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "g2o: A General Framework for Graph Optimization," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2011, pp. 3607–3613.
- [18] B. Bell and F. Cathey, "The Iterated Kalman filter Update as a Gauss-Newton Method," *IEEE Trans. Autom. Control*, vol. 38, no. 2, pp. 294–297, 1993.
- [19] T. D. Barfoot, *State Estimation For Robotics: A Matrix-Lie-Group Approach*. Draft in preparation for publication by Cambridge University Press, 2016.
- [20] L. Chen, P. O. Arambel, and R. K. Mehra, "Estimation under unknown correlation: Covariance intersection revisited," *IEEE Trans. Autom. Control*, vol. 47, no. 11, pp. 1879–1882, 2002.
- [21] J. K. Uhlmann, "Dynamic map building and localization for autonomous vehicles," Ph.D. dissertation, University of Oxford, 1995.