Distributed Kalman Filtering using a Common Prior

Felix Gustafsson, Roland Hostettler, Subhrakanti Dey
Department of Electrical Engineering; Signals and Systems, Uppsala University, Uppsala, Sweden
Email: {felix.gustafsson, roland.hostettler, subhrakanti.dey}@angstrom.uu.se

***Abstract*—This paper presents an optimal Kalman filtering (KF) method for distributed systems where agents have unequal state vectors. In contrary to traditional distributed KF methods, the prediction is done centrally at the server. The proposed method allows for an analytically optimal linear estimator that facilitates data transfer efficiency, privacy, and scalability, particularly in scenarios where subsystems have large numbers of measurements. The approach avoids the dimensionality challenges in centralized systems by maintaining local estimates at the distributed agents and minimizing the data transmitted to a central server.**

***Index Terms*—Distributed Kalman Filter, Unequal state vectors, Distributed estimation, Kalman Filtering**

I. Background

In dynamic systems, the choice of estimation method is driven by the system's characteristics. For systems that are linear with Gaussian noise, the centralized Kalman Filter (KF) provides an optimal Bayesian estimate [1]. The centralized KF assumes a centralized system that has access to all measured data. If the data is collected in a distributed manner, this may not be possible in a practical setting. Accessing all data can be problematic due to privacy concerns, highlighted in Federated Learning (FL) [2], and communication limitations [3].

An alternative to centralized estimation is Distributed Data Fusion (DDF). In DDF, each agent estimates the system states individually, which are then fused to obtain a more accurate estimate than any individual estimate. The Bar-Shalom/Campo (BC) fuser [4] is an optimal way of fusing two estimates of the same state. A limitation of the BC fuser is that it requires knowledge of the cross-covariance between estimates, which is rarely available in practical applications. If the cross covariance is (wrongly) assumed to be zero, the resulting fusion is called naive. Covariance Intersection (CI) [5] offers an alternative to naive fusion, it provides conservative fusion irrespective of cross covariances, meaning that it does not underestimate the covariance of the fused estimate. However, CI cannot be directly applied when distributed agents estimate different subspaces of a full global state, a scenario known as distributed estimation of unequal states [6].

In this paper, we propose a novel strategy for distributed estimation of unequal state vectors by using a global prior for the distributed agents. This differs from the typical setup seen in literature where priors are local to each agent. To the best of the authors knowledge, this has not previously been considered. This maintains the privacy of the measurements and if the distributed agents have more measurements than states, it also reduces the complexity of centralized computations.

This work was supported by the Swedish Research Council under grant 2022-04505.

II. Problem formulation

This paper considers the linear estimation of a global state using local distributed estimates of its subspaces. The global system is characterized by the discrete time process

$$x_n^* = \mathbf{A}_n x_{n-1}^* + w_n^* \tag{1}$$

where $\mathbf{A}_n \in \mathbb{R}^{d_{x^*} \times d_{x^*}}$, $x_n^* \in \mathbb{R}^{d_{x^*}}$ and $w_n^* \in \mathbb{R}^{d_{x^*}}$ are the system matrix, the state vector, and zero mean white Gaussian process noise with covariance matrix $\mathbf{Q}_n^*$ at time n, respectively. The superscript x^* denotes the global state vector. Each distributed subsystem state is given by

$$x_n^l = \mathbf{M}_n^l x_n^* \qquad l = 1, \ldots, L \tag{2}$$

where $x_n^l \in \mathbb{R}^{d_{x^l}}$, denotes the lth local state and $\mathbf{M}_n^l \in \mathbb{R}^{d_{x^l} \times d_{x^*}}$ is a mapping from the global state to the subsystem.

Each agent measures an arbitrary subset of the local states

$$y_n^l = \mathbf{H}_n^l x_n^l + v_n^l \tag{3}$$

where $y_n^l \in \mathbb{R}^{d_{y^l}}$ is the measurement, $\mathbf{H}_n^l \in \mathbb{R}^{d_{y^l} \times d_{x^l}}$ is the observation matrix, and $v_n^l \in \mathbb{R}^{d_{y^l}}$ is zero mean white Gaussian measurement noise with covariance matrix $\mathbf{R}_n^l \in \mathbb{R}^{d_{y^l} \times d_{y^l}}$. It is assumed that there is no relation between measurements in different subsystems, which means that local subsystems only measure their own states and $\mathrm{Cov}(v_n^i, v_n^j) = 0, \ \forall \ i \neq j$.

Assuming that all measurement data y_n^*, with corresponding covariance matrix $\mathbf{R}_n^*$, and global observation matrix $\mathbf{H}_n^* \in \mathbb{R}^{d_{y^*} \times d_{x^*}}$, are available at a centralized server, the optimal estimate of the global system, is given by the centralized KF.

The global KF will provide an optimal estimate, but there are practical and privacy concerning challenges regarding the measurements of the subsystems. Transferring the raw data may require using a lot of bandwidth, and the distributed agents may not want to share their data with other agents or the centralized server. Hence, we consider the case where each agent performs a local measurement update based on common prior information obtained from the global system, which is maintained by a central server. The central server performs the global prediction and transfers the (potentially) reduced prior state and covariance estimates to each respective agent according to

$$\hat{x}_{n|n-1}^l = \mathbf{M}_n^l \, \hat{x}_{n|n-1}^* \tag{4a}$$

$$\mathbf{P}_{n|n-1}^l = \mathbf{M}_n^l \, \mathbf{P}_{n|n-1}^* \, (\mathbf{M}_n^l)^\intercal. \tag{4b}$$

TABLE I
TOY EXAMPLE OF THE PROPOSED METHOD

$\mathbf{H}^*$	$\bar{\mathbf{H}}^*$	y	$\bar{y}$	$\mathbf{R}$	$\bar{\mathbf{R}}$	**RMSRE**
$\begin{bmatrix} 1.0 & 0.3 & 0 & 0 \\ 2.0 & 2.0 & 0 & 0 \\ 0.3 & 1.0 & 0 & 0 \\ 0 & 0 & 2.0 & 1.0 \\ 0 & 0 & 1.0 & 3.0 \\ 0 & 0 & 0.3 & 0.5 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -0.10 \\ 0.06 \\ 0.76 \\ -0.72 \\ -0.83 \\ -0.90 \end{bmatrix}$	$\begin{bmatrix} -0.56 \\ 0.67 \\ -0.29 \\ -0.21 \end{bmatrix}$	$0.5\mathbf{I}_{6\times 6}$	$\begin{bmatrix} 0.54 & -0.48 & 0 & 0 \\ -0.48 & 0.54 & 0 & 0 \\ 0 & 0 & 0.20 & -0.10 \\ 0 & 0 & -0.10 & 0.10 \end{bmatrix}$	4.37×10^{-16}

The updated reduced-state estimates and covariances are then transferred back to the centralized server to perform a global update without needing the local measurements, measurement covariances or observation matrices.

III. EQUIVALENT MEASUREMENTS AND COVARIANCES

By utilizing the fact that the global state has a common prior distribution and that the locally calculated estimates are available, it is possible to calculate equivalent, but transformed, measurements, measurement covariances and observation matrices without compromising the integrity of the distributed agents. If the calculated observation matrices $\bar{\mathbf{H}}_n^l$ are augmented with zero-columns for the non-common state of the lth distributed agent and the global state, we can form

$$\bar{y}_n^* = \begin{bmatrix} \bar{y}_n^1 \\ \vdots \\ \bar{y}_n^L \end{bmatrix}, \bar{\mathbf{R}}_n^* = \begin{bmatrix} \bar{\mathbf{R}}_n^1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \bar{\mathbf{R}}_n^L \end{bmatrix}, \bar{\mathbf{H}}_n^* = \begin{bmatrix} \bar{\mathbf{H}}_n^{1,\ aug} \\ \vdots \\ \bar{\mathbf{H}}_n^{L,\ aug} \end{bmatrix},$$

which are the equivalent measurement vector, noise covariance, and the joint estimated observation matrix, respectively. These can be used for an equivalent global KF update.

IV. RESULTS

In order to demonstrate the feasibility of the proposed method, we show a toy example for one time step. Consider the global system

$$x_n = \frac{1}{100}\begin{bmatrix} 82 & 8 & 0 & 0 \\ 8 & 75 & 7 & 0 \\ 0 & 7 & 68 & 6 \\ 0 & 0 & 6 & 61 \end{bmatrix} x_{n-1} + w_n, w_n \sim \mathcal{N}\left(\mathbf{0}, \frac{1}{100}\begin{bmatrix} 8 & 1 & 0 & 0 \\ 1 & 8 & 1 & 0 \\ 0 & 1 & 7 & 1 \\ 0 & 0 & 1 & 6 \end{bmatrix}\right),$$

that contains 4 states, and is divided up into 2 subsystems, $\mathbf{A}$ and $\mathbf{B}$, each containing 2 substates of the global states.

A centralized server runs a Kalman prediction of the global state, and transfers marginalizations of the global state (4) to $\mathbf{A}$ and $\mathbf{B}$. Subsystems $\mathbf{A}$ and $\mathbf{B}$ each measure their respective states using a total of 3 observations each, with their respective observations matrices $\mathbf{H}^{\mathbf{A}}$ and $\mathbf{H}^{\mathbf{B}}$. For comparison, a centralized KF is used to estimate the system, using the global observation matrix

$$\mathbf{H}^* = \begin{bmatrix} \mathbf{H}^{\mathbf{A}} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}^{\mathbf{B}} \end{bmatrix} \in \mathbb{R}^{6\times 4}. \quad (5)$$

Table I presents the RMSRE [7] between the centralized KF and the method. The results highlight the advantages of the proposed method in terms of privacy. Neither the true measurement, nor the measurement covariance, nor the number of measurements are revealed at the server. The benefits in terms of communication and server side computation are also highlighted. If the number of measurements increase, the calculated observation matrix $\bar{\mathbf{H}}^*$ will still be a $\mathbf{I}_{4\times 4}$.

V. CONCLUSIONS

The proposed method demonstrates near identical results to the centralized KF in simulations, with the main differences being in the order of magnitude of machine precision. The proposed method allows for an analytically optimal linear estimator for distributed systems of unequal states, which has advantages over the centralized KF in terms of data transfer, privacy and computational time, particularly when the subsystems have a large number of measurements. Even if the number of measurements far exceeds the number of states at the local agent, the calculations done at the central server assume that the number of measurements equals the number of states. This can significantly reduce the dimensionality of the matrix inversions at the centralized server. In addition to this, since only the state estimates and covariance matrices are sent back to the centralized server, the integrity of the distributed agents is not compromised.

REFERENCES

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[2] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," 2023, arXiv:1602.05629.

[3] H. Qiu, F. Ahmad, F. Bai, M. Gruteser, and R. Govindan, "AVR: Augmented vehicular reality," in *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '18, 2018, p. 81–95.

[4] Y. Bar-Shalom and L. Campo, "The effect of the common process noise on the two-sensor fused-track covariance," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-22, no. 6, pp. 803–805, 1986.

[5] S. Julier and J. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations," in *Proceedings of American Control Conference*, vol. 4, 1997, pp. 2369–2373 vol.4.

[6] B. Noack, J. Sijs, and U. D. Hanebeck, "Fusion strategies for unequal state vectors in distributed Kalman filtering," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 3262–3267, 2014, 19th IFAC World Congress.

[7] X. Li and Z. Zhao, "Relative error measures for evaluation of estimation algorithms," in *2005 7th International Conference on Information Fusion*, vol. 1, 2005, p. 8.