Conservative Compression of Information Matrices using Event-Triggering and Robust Optimization

Christopher Funk and Benjamin Noack
Autonomous Multisensor Systems Group
Institute for Intelligent Cooperating Systems
Otto-von-Guericke University, Magdeburg, Germany
christopher.funk@ovgu.de, benjamin.noack@ieee.org

***Abstract*—Distributed sensor fusion requires the transmission of intermediate fusion results, consisting of point estimates and associated error covariance or information matrices. Bandwidth constraints necessitate data compression techniques for error covariance and information matrices, which typically dominate data volume. To ensure the safe use of the fusion results for decision-making, these techniques must be conservative, i.e., not lead to the compressed error covariance or information matrices underestimating the true estimate error. This work introduces a novel approach for the conservative compressed transmission of information matrices, that builds on a previous event-based method for covariance matrices. The proposed method allows the entire sensor fusion pipeline to operate in 'information space', facilitating efficient fusion operations without the need to compute corresponding covariance matrices. Contributions include an event-trigger for information matrices and a robust-optimization-based bounding mechanism ensuring conservativeness. The proposed approach is evaluated in the context of transmitting error information matrices generated by extended information filter SLAM to a receiver for further processing.**

***Index Terms*—Distributed Data Fusion, Data Reduction, Conservative Data Fusion**

I. Introduction

Distributed sensor fusion aims to combine and refine information extracted from sensor data without requiring all of the data acquisition and processing pipeline to be colocated. This has several advantages, such as distributing the computational load, greater robustness to failure and better coverage, but requires communication between the devices involved. In particular, intermediate fusion results, typically consisting of a point estimate and an associated error covariance or information matrix, need to be transmitted. To ensure the safe and reliable operation of the physical systems utilizing the fusion results, the uncertainty encoded in these matrices must not underestimate the actual uncertainty of the point estimate. Since the transmission bandwidth is often limited, data compression techniques need to be applied, particularly to the error covariance and information matrices, as they cause the majority of the data volume to be transmitted. Critically, data compression must not lead to underestimation of the actual point estimate uncertainty by the compressed matrices, motivating the notion of *conservative* data compression methods for covariance and information matrices. The use of information matrices instead of covariance matrices is advantageous in certain applications. For instance, efficient simultaneous localization and mapping (SLAM) algorithms exploit the approximately [1] or exactly [2] sparse nature of the information matrix compared to the covariance matrix. In addition, many fusion algorithms such as least squares (LS), the Kalman filter (KF), and covariance intersection (CI) [3] can be naturally formulated in terms of information matrices instead of covariance matrices, thereby avoiding potentially costly matrix inversions. In fact, it is possible to perform most sensor fusion operations without converting to covariance matrices until the final result is to be extracted. Consequently, it is clearly useful to consider the compressed transmission of information matrices for cases where the application at hand dictates their use. As will be seen, this poses unexpected theoretical and practical challenges, compared to the transmission of covariance matrices.

Only recently has the need for conservative covariance matrix compression methods for sensor fusion been recognized and investigated in the literature. Existing approaches for covariance matrices include conservative diagonal [4] or block-diagonal approximations [5] which use induced sparsity as the main mechanism for data reduction. In contrast, projection-based methods [6]–[9] utilize low-dimensional linear projections of estimates and covariance matrices in conjunction with conservative fusion algorithms such as CI to achieve data reduction. Quantization-based approaches map individual matrix elements to lower resolution numerical representations and prevent the potential loss of conservativeness through appropriate diagonal shifts that are obtained through solving an optimization problem [10] or computing a modified Cholesky decomposition [11]. In the event-based approach to covariance matrix compression [12], [13] individual covariance matrix elements are only transmitted if they satisfy a so-called trigger condition that depends on past values, e.g., of the same element. The received elements are stored in a buffer and, with the implicit information contained in not receiving an element, used by a bounding mechanism inside the receiver to obtain a conservative covariance matrix. In contrast to the above methods for covariance matrices, to the best of the authors' knowledge, no *application-independent* methods for the compression of information matrices have been proposed. However, approaches specific to certain applications such as SLAM exist [1], [14] and can in fact be used together with the approach presented in the following.

The extended event-based method from [13] forms the basis for this work. While the overall structure of the method, consisting of a trigger condition, an element buffer and a bounding mechanism remains, now, the transmission of information matrices instead of covariance matrices is considered. Surprisingly, this necessitates a theoretically substantially different bounding mechanism and as a consequence a more complex trigger condition. The resulting approach enables the compressed transmission of sequences of information matrices without having to compute the corresponding covariance matrix, which allows for the entire sensor fusion pipeline to operate in 'information space'. The contributions of this work are 1) a bounding mechanism for the event-based transmission of information matrices, 2) a trigger condition that ensures the feasibility of the bounding mechanism, and 3) an evaluation of the new approach in a scenario, where extended information filter SLAM results are to be sent to a receiver.

II. Background

In this section, the mathematical notation, the central concepts of positive semidefiniteness, diagonal dominance, and conservativeness for information matrices, as well as their interrelationships are introduced. In addition, the original event-based method for covariance matrix compression, as proposed in [12], [13], is outlined and recapitulated.

A. Mathematical Notation

In this work, scalars are denoted by lowercase letters (e.g., a), vectors by lowercase underlined letters (e.g., $\underline{a}$), and matrices by uppercase bold letters (e.g., $\mathbf{A}$). Elements of a vector $\underline{a}$ or a matrix $\mathbf{A}$ are indicated by $[\underline{a}]_i$ or $[\mathbf{A}]_{ij}$, where i and j are the element's row- and column index, respectively. The identity matrix is denoted by $\mathbf{I}$ and the all one vector and matrix by $\underline{1}$ and $\mathbf{1}$, respectively. Elementwise multiplication (the Hadamard product) of two matrices $\mathbf{A}$ and $\mathbf{B}$ is indicated by $\mathbf{A} \odot \mathbf{B}$. Furthermore, $|\mathbf{A}|$ denotes taking the elementwise absolute value of the matrix $\mathbf{A}$. For symmetric matrices $\mathbf{A}$ and $\mathbf{B}$ the notation $\mathbf{A} \succeq \mathbf{B}$ is synonymous with $\mathbf{A} - \mathbf{B}$ being positive semidefinite. The minimum and maximum eigenvalue of a matrix $\mathbf{A}$ are denoted as $\lambda_{min}(\mathbf{A})$ and $\lambda_{max}(\mathbf{A})$, respectively. The notation $\mathbf{A} \geq \mathbf{B}$ expresses the elementwise (standard) inequality of the matrices $\mathbf{A}$ and $\mathbf{B}$. Sets are denoted by standard uppercase letters, e.g., A. The set of real symmetric $n \times n$ matrices is denoted as S_n, the set of real positive semidefinite $n \times n$ matrices as S_n^+, the set of real positive definite $n \times n$ matrices as S_n^{++}, and the set of real diagonally dominant $n \times n$ matrices as DD_n.

B. Positive Semidefiniteness and Diagonal Dominance

Positive semidefiniteness (PSD) and properties related to it play an important role in definitions and derivations in the remainder of this work. The following lemmas provide useful tools to work with and bound PSD matrices (above / below) in terms of the generalized inequality $\mathbf{A} \succeq \mathbf{B}$ iff $\mathbf{A} - \mathbf{B} \in S_n^+$.

Lemma 1. *If $\mathbf{A} \in S_n$, then $\lambda_{min}(\mathbf{A})\mathbf{I} \preceq \mathbf{A} \preceq \lambda_{max}(\mathbf{A})\mathbf{I}$.*

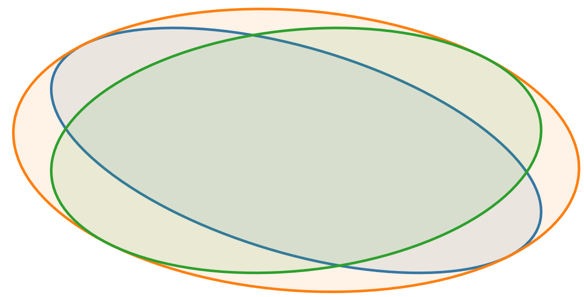

Fig. 1. Credible set of covariance matrix $\mathbf{B}$ (blue), credible set of covariance matrix $\mathbf{A}_c$ conservative with respect to $\mathbf{B}$ (orange), and credible set of covariance matrix $\mathbf{A}_{nc}$ not conservative with respect to $\mathbf{B}$ (green). The credible set of $\mathbf{B}$ is contained in that of $\mathbf{A}_c$, but not in that of $\mathbf{A}_{nc}$.

Proof. Let $\mathbf{A}$ have the eigenvalues $\lambda_1 \leq \ldots \leq \lambda_n$. Then $\mathbf{A} + \alpha\mathbf{I}$ has the eigenvalues $\lambda_i' = \lambda_i + \alpha$, $i = 1, \ldots, n$, since $(\mathbf{A} + \alpha\mathbf{I})\underline{x}_i = (\lambda_i + \alpha)\underline{x}_i$, where $\underline{x}_i$ is the eigenvector associated with λ_i. Hence, $\lambda_{max}(\mathbf{A})\mathbf{I} - \mathbf{A}$ has eigenvalues $\lambda_i' = \lambda_n - \lambda_i \geq 0$, $i = 1, \ldots, n$ and consequently is PSD. The same argument shows that $\mathbf{A} - \lambda_{min}(\mathbf{A})\mathbf{I}$ is PSD, as claimed. $\square$

Lemma 2. *If $\mathbf{A}, \mathbf{B} \in S_n^{++}$, then $\mathbf{A} \succeq \mathbf{B}$ if and only if $\mathbf{B}^{-1} \succeq \mathbf{A}^{-1}$ [15, Cor. 7.7.4].*

Diagonal dominance is essential as a tractable approximation of positive semidefiniteness and is thus introduced next.

Definition 1. *A matrix $\mathbf{A} \in S_n$ is diagonally dominant (DD) if the inequality $[\mathbf{A}]_{ii} \geq \sum_{\substack{j=1 \\ j \neq i}}^{n} |[\mathbf{A}]_{ij}|$ holds for all $i = 1, \ldots, n$.*

The following lemma demonstrates that diagonal dominance implies positive semidefiniteness and therefore $DD_n \subseteq S_n^+$.

Lemma 3. *If $\mathbf{A} \in DD_n$, then $\mathbf{A} \in S_n^+$ [12, Lem. 1].*

In the context of this work, the set of DD matrices provides an approximation of the set of positive semidefinite matrices that allows approximately solving optimization problems with an infinite number of PSD constraints.

C. Conservativeness of Covariance and Information Matrices

The concepts of conservativeness of covariance / information matrices and diagonal dominance are used extensively in this work and thus are introduced here for future reference.

Definition 2. *Let $\mathbf{A}, \mathbf{B} \in S_n^{++}$ be two covariance matrices. $\mathbf{A}$ is said to be conservative with respect to $\mathbf{B}$ if $\mathbf{A} \succeq \mathbf{B}$ holds.*

If and only if $\mathbf{A}$ is conservative with respect to the error covariance matrix $\mathbf{B} = \mathbb{E}_{\underline{x},\underline{y}}[(\underline{x} - \hat{\underline{x}}(\underline{y}))(\underline{x} - \hat{\underline{x}}(\underline{y}))^\top]$ of an estimate $\hat{\underline{x}}$ derived from observations $\underline{y}$, the ellipsoidal credible sets $A = \{\underline{x} | (\underline{x} - \hat{\underline{x}})^\top \mathbf{A}^{-1} (\underline{x} - \hat{\underline{x}}) \leq \alpha\}$ and $B = \{\underline{x} | (\underline{x} - \hat{\underline{x}})^\top \mathbf{B}^{-1} (\underline{x} - \hat{\underline{x}}) \leq \alpha\}$, where α is a parameter tied to the credibility level, satisfy $A \supseteq B$, as can be seen from Lemma 2 and as is illustrated in Fig. 1. Consequently, conservativeness of $\mathbf{A}$ with respect to $\mathbf{B}$ guarantees that the estimate uncertainty encoded in $\mathbf{A}$ does not underestimate that encoded in $\mathbf{B}$. Information matrices are related to their

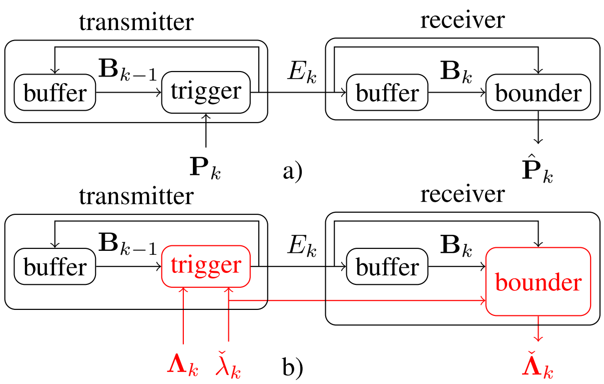

Fig. 2. Overview of conservative event-triggered a) covariance matrix and b) information matrix transmission approaches. a) Covariance matrix to be transmitted $\mathbf{P}_k$, transmitter buffer matrix $\mathbf{B}_{k-1}$, event-set E_k, receiver buffer matrix $\mathbf{B}_k$ and upper bound $\hat{\mathbf{P}}_k$ on original information matrix $\mathbf{P}_k$. b) Information matrix to be transmitted $\mathbf{\Lambda}_k$ with lower bound $\check{\lambda}_k$ on its minimum eigenvalue, transmitter buffer matrix $\mathbf{B}_{k-1}$, event-set E_k, receiver buffer matrix $\mathbf{B}_k$ and lower bound $\check{\mathbf{\Lambda}}_k$ on original information matrix $\mathbf{\Lambda}_k$. Conceptually different components of b) compared to a) are highlighted.

corresponding covariance matrices by inversion: Let $\mathbf{A} \in S_n^{++}$ be a covariance matrix, then $\mathbf{A}^{-1}$ is the corresponding information matrix. Due to Lemma 2, conservativeness of covariance matrices can equivalently be expressed in terms of information matrices.

Definition 3. *Let $\mathbf{A}, \mathbf{B} \in S_n^{++}$ be information matrices. $\mathbf{A}$ is said to be conservative with respect to $\mathbf{B}$ if $\mathbf{A} \preceq \mathbf{B}$ holds.*

This definition has the same interpretation in terms of credible sets as the previous one using covariance matrices. Since this work considers the transmission of information matrices, this is the definition that will be used in the following.

D. Event-Based Covariance Matrix Compression

The event-based approach for covariance matrix compression from [12], [13] aims to transmit a sequence of covariance matrices $\mathbf{P}_k \in S_n^{++}$, $k \in \mathbb{N}$, and consists of three components that are interrelated as shown in Fig. 2 a). At each time step k a covariance matrix $\mathbf{P}_k$ is input at the transmitter and a covariance matrix $\hat{\mathbf{P}}_k \in S_n^{++}$ that is conservative with respect to $\mathbf{P}_k$ is output at the receiver. The buffers at transmitter and receiver contain the same approximation $\mathbf{B}_k \in S_n$ of $\mathbf{P}_k$ that is continually updated, as described in more detail below. From the buffer value $\mathbf{B}_{k-1}$ and the covariance matrix $\mathbf{P}_k$, the trigger generates the event set of index-value pairs

$$E_k = \{((i,j), [\mathbf{P}_k]_{ij}) | |[\mathbf{P}_k - \mathbf{B}_{k-1}]_{ij}| > [\mathbf{T}]_{ij}, i \leq j\} \quad (1)$$

of elements of $\mathbf{P}_k$ that are to be transmitted. Here, elements $[\mathbf{P}_k]_{ij}$ that have changed more than a threshold $[\mathbf{T}]_{ij}$, where $\mathbf{T} \in S_n$, $\mathbf{T} \geq \mathbf{0}$, are transmitted. This trigger rule is called *absolute change trigger* and represents the simplest of many possible trigger rules, see [13]. Both at the transmitter and the receiver, the buffer value $\mathbf{B}_{k-1}$ is updated according to $[\mathbf{B}_k]_{ij} = [\mathbf{B}_k]_{ji} = [\mathbf{P}_k]_{ij}$ for all $((i,j), [\mathbf{P}_k]_{ij}) \in E_k$. While the buffer $\mathbf{B}_k$ is an approximation of $\mathbf{P}_k$, it is not necessarily conservative with respect to $\mathbf{P}_k$. This necessitates the bounder, which computes the conservative covariance matrix $\hat{\mathbf{P}}_k$ based on $\mathbf{B}_k$ and E_k. It does so by solving a robust optimization problem with solution $\hat{\mathbf{P}}_k = \mathbf{B}_k + \mathrm{diag}((\mathbf{T} \odot \mathbf{E}_k)\underline{1})$, where $[\mathbf{E}_k]_{ij} = [\mathbf{E}_k]_{ji} = 0$ if $((i,j), \cdot) \in E_k$ and $= 1$ otherwise.

III. Information Matrix Compression using Approximate Robust Optimization

In this section, a sequence of information matrices $\mathbf{\Lambda}_k \in S_n^{++}$, $k \in \mathbb{N}$, is considered for transmission instead of a sequence of covariance matrices $\mathbf{P}_k \in S_n^{++}$, $k \in \mathbb{N}$ as in Sec. II-D. Note that information and covariance matrices are related by $\mathbf{\Lambda}_k = \mathbf{P}_k^{-1}$. The overall approach resembles the one recapitulated in Sec. II-D. Its components and their interplay are illustrated in Fig. 2 b). Similar to the approach for covariance matrices, both buffers contain an approximation $\mathbf{B}_k$ of $\mathbf{\Lambda}_k$. As before the trigger creates an event-set E_k consisting of index-value tuples $((i,j), [\mathbf{\Lambda}_k]_{ij})$ based on the buffer value $\mathbf{B}_{k-1}$ and the current information matrix $\mathbf{\Lambda}_k$. The main difference is that the approach now also requires a lower bound $\check{\lambda}_k$ on $\mathbf{\Lambda}_k$'s minimum eigenvalue for its operation. The event-set is transmitted to the receiver and both the transmitter and receiver buffer are updated according to $[\mathbf{B}_k]_{ij} = [\mathbf{B}_k]_{ji} = [\mathbf{\Lambda}_k]_{ij}$ for all $((i,j), [\mathbf{\Lambda}_k]_{ij}) \in E_k$. In contrast to the approach for covariance matrices, an additional value, $\check{\lambda}_k$, needs to be transmitted. The updated buffer $\mathbf{B}_k$, the event-set E_k, and the eigenvalue lower bound $\check{\lambda}_k$ are then used by the bounder to generate an information matrix $\check{\mathbf{\Lambda}}_k$ conservative with respect to $\mathbf{\Lambda}_k$. As will be seen, the bounder for the information matrix approach proposed here has different properties and requirements to that used in the covariance matrix approach, which necessitates using a modified trigger rule. In the remainder of this section, the bounder, its requirements and properties, and a trigger rule that is compatible with the bounder are presented.

A. Information Matrix Bounder

The goal of the bounder is to determine $\check{\mathbf{\Lambda}}_k$ such that it is conservative with respect to $\mathbf{\Lambda}_k$ and maximizes $\mathrm{tr}(\check{\mathbf{\Lambda}}_k)$, where the latter ensures that $\check{\mathbf{\Lambda}}_k$ is not overly conservative. By Definition 3, $\mathbf{\Lambda}_k \succeq \check{\mathbf{\Lambda}}_k$ must hold for $\check{\mathbf{\Lambda}}_k$ to be conservative with respect to $\mathbf{\Lambda}_k$. In addition, $\check{\mathbf{\Lambda}}_k \succeq \mathbf{0}$ needs to be ensured to make $\check{\mathbf{\Lambda}}_k$ a valid information matrix. The latter condition is a major difference to the covariance matrix approach, where conservativeness is ensured via Definition 2, i.e., $\hat{\mathbf{P}}_k \succeq \mathbf{P}_k$, which implies $\hat{\mathbf{P}}_k \succeq \mathbf{0}$ automatically, due to $\mathbf{P}_k \succeq \mathbf{0}$. The above goals are expressed by the semidefinite program (SDP)

$$\begin{aligned} \underset{\check{\mathbf{\Lambda}}_k \in S_n}{\text{maximize}} \quad & \mathrm{tr}(\check{\mathbf{\Lambda}}_k) \\ \text{subject to} \quad & \mathbf{\Lambda}_k \succeq \check{\mathbf{\Lambda}}_k \succeq \mathbf{0} \end{aligned}$$

which unfortunately cannot be solved at the receiver, as $\mathbf{\Lambda}_k$ is not available there. However, similarly to [12], [13], the buffer $\mathbf{B}_k$, the event-set E_k, and the eigenvalue bound $\check{\lambda}_k$, which are available at the receiver, can be used to construct a sufficient conservativeness condition that does not require knowledge of the information matrix $\mathbf{\Lambda}_k$. For that, define the buffer error as $\mathbf{\Delta}_k = \mathbf{\Lambda}_k - \mathbf{B}_k$ and the candidate conservative information matrix as $\check{\mathbf{\Lambda}}_k = \mathbf{B}_k - \mathbf{S}_k$, where $\mathbf{S}_k \in S_n$ is to be determined, such that $\check{\mathbf{\Lambda}}_k$ satisfies the conservativeness condition $\mathbf{\Lambda}_k \succeq \check{\mathbf{\Lambda}}_k$, is PSD, and has maximum trace. The conservativeness condition can be rewritten as $\mathbf{S}_k + \mathbf{\Delta}_k \succeq \mathbf{0}$ by combining it with the definition of the buffer error and that of the candidate conservative information matrix. Identically to [12], [13], it is assumed that, although $\mathbf{\Delta}_k$ is not known at the receiver, it is known to lie within a set of the form

$$F_k = \left\{ \mathbf{\Delta} \in S_n \middle| |\mathbf{\Delta}| \leq \hat{\mathbf{\Delta}}_k \right\} \tag{2}$$

where $\hat{\mathbf{\Delta}}_k \in S_n$, $\hat{\mathbf{\Delta}}_k \geq \mathbf{0}$, is determined by the choice of the trigger rule, the event-set E_k, and the eigenvalue bound $\check{\lambda}_k$. In other words, the employed trigger rule together with E_k and $\check{\lambda}_k$ must allow to derive a set F_k of the above form that contains the actual buffer error $\mathbf{\Delta}_k$. With the assumption that $\mathbf{\Delta}_k \in F_k$, it is evident that if $\mathbf{S}_k + \mathbf{\Delta} \succeq \mathbf{0}$ for all $\mathbf{\Delta} \in F_k$ then $\mathbf{S}_k + \mathbf{\Delta}_k \succeq \mathbf{0}$ and consequently $\mathbf{\Lambda}_k \succeq \check{\mathbf{\Lambda}}_k$ follow. Therefore, $\mathbf{S}_k + \mathbf{\Delta} \succeq \mathbf{0}$ for all $\mathbf{\Delta} \in F_k$ is a sufficient condition for $\check{\mathbf{\Lambda}}_k$ to be conservative with respect to $\mathbf{\Lambda}_k$. This sufficient condition leads to the robust SDP for $\check{\mathbf{\Lambda}}_k$ in terms of $\mathbf{S}_k$

$$\begin{aligned} \underset{\mathbf{S}_k \in S_n}{\text{maximize}} \quad & \operatorname{tr}(\check{\mathbf{\Lambda}}_k) = \operatorname{tr}(\mathbf{B}_k - \mathbf{S}_k) \\ \text{subject to} \quad & \mathbf{\Delta} + \mathbf{S}_k \succeq \mathbf{0} \quad \forall \mathbf{\Delta} \in F_k \\ & \check{\mathbf{\Lambda}}_k = \mathbf{B}_k - \mathbf{S}_k \succeq \mathbf{0} \end{aligned} \tag{RBP}$$

which can be solved at the receiver and will next be simplified using diagonal dominance to avoid the high computational effort required for its solution. To simplify (RBP), the conservativeness constraint $\mathbf{\Delta} + \mathbf{S}_k \succeq \mathbf{0}\ \forall \mathbf{\Delta} \in F_k$ is approximated by $\mathbf{\Delta} + \mathbf{S}_k \in DD_n\ \forall \mathbf{\Delta} \in F_k$, or in terms of Definition 1, by

$$[\mathbf{\Delta}]_{ii} + [\mathbf{S}_k]_{ii} \geq \sum_{\substack{j=1 \\ j \neq i}}^{n} |[\mathbf{\Delta}]_{ij} + [\mathbf{S}_k]_{ij}|, \quad \begin{matrix} \forall \mathbf{\Delta} \in F_k \\ i = 1, \ldots, n \end{matrix}.$$

Due to Lemma 3, the above robust DD constraint still ensures that the original robust PSD constraint holds. In addition, the following lemma from [13] implies that the above robust DD constraint is equivalent to a finite number of constraints.

Lemma 4. *Let F_k be of the form* (2)*, then the sets*

$$A = \left\{ \mathbf{S}_k \in S_n \middle| \begin{matrix} [\mathbf{\Delta}]_{ii} + [\mathbf{S}_k]_{ii} \geq \sum_{\substack{j=1 \\ j \neq i}}^{n} |[\mathbf{\Delta}]_{ij} + [\mathbf{S}_k]_{ij}|, \\ \forall \mathbf{\Delta} \in F_k, i = 1, \ldots, n \end{matrix} \right\}$$

and

$$B = \left\{ \mathbf{S}_k \in S_n \middle| \begin{matrix} [\mathbf{S}_k]_{ii} \geq \sum_{j=1}^{n} [\hat{\mathbf{\Delta}}_k]_{ij} + \sum_{\substack{j=1 \\ j \neq i}}^{n} |[\mathbf{S}_k]_{ij}|, \\ i = 1, \ldots, n \end{matrix} \right\}$$

are identical [13, Lem. 2].

In summary, replacing the robust PSD constraint with a robust DD constraint is an approximation, but still ensures the conservativeness of the optimization result. The newly introduced robust DD constraint, which comprises an infinite number of inequality constraints, can in turn be expressed equivalently by a finite number of inequality constraints. These considerations lead to the standard semidefinite program

$$\begin{aligned} \underset{\mathbf{S}_k \in S_n}{\text{maximize}} \quad & \operatorname{tr}(\check{\mathbf{\Lambda}}_k) = \operatorname{tr}(\mathbf{B}_k - \mathbf{S}_k) \\ \text{subject to} \quad & \begin{matrix} [\mathbf{S}_k]_{ii} \geq \sum_{j=1}^{n} [\hat{\mathbf{\Delta}}_k]_{ij} + \sum_{\substack{j=1 \\ j \neq i}}^{n} |[\mathbf{S}_k]_{ij}|, \\ i = 1, \ldots, n \end{matrix} \\ & \check{\mathbf{\Lambda}}_k = \mathbf{B}_k - \mathbf{S}_k \succeq \mathbf{0} \end{aligned} \tag{ARBP}$$

to be solved by the bounder. Its optimal solution $\mathbf{S}_k^*$ gives the optimal conservative information matrix $\check{\mathbf{\Lambda}}_k^* = \mathbf{B}_k - \mathbf{S}_k^*$. As will be shown next, (ARBP) has a simple solution.

Theorem 1. (ARBP) *is feasible iff* $\mathbf{B}_k \succeq \operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1})$ *and* $\mathbf{S}_k^* = \operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1})$ *is an optimal solution.*

Proof. First assume that $\mathbf{B}_k \succeq \operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1})$ holds. Then $\mathbf{S}_k = \operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1}) \in S_n$ satisfies the PSD and DD constraints with equality and (ARBP) is feasible. Now assume that (ARBP) is feasible but that $\mathbf{B}_k \not\succeq \operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1})$. Choose $\mathbf{S}_k$ feasible, then $\mathbf{B}_k \succeq \mathbf{S}_k$ holds. The diagonal of $\mathbf{S}_k$ can be decreased arbitrarily without violating the PSD constraint. Therefore, the smallest $[\mathbf{S}_k]_{ii}$ for which $\mathbf{S}_k$ remains feasible is attained when $[\mathbf{S}_k]_{ii}$ is equal to the right hand side of the respective DD constraint. This choice of diagonal elements $[\mathbf{S}_k]_{ii}$ results in

$$\mathbf{S}'_k = \operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1}) + \begin{bmatrix} \sum_{\substack{j=1 \\ j \neq 1}}^{n} |[\mathbf{S}_k]_{1j}| & \cdots & [\mathbf{S}_k]_{1n} \\ \vdots & \ddots & \vdots \\ [\mathbf{S}_k]_{n1} & \cdots & \sum_{\substack{j=1 \\ j \neq n}}^{n} |[\mathbf{S}_k]_{nj}| \end{bmatrix},$$

where the second summand can be seen to be DD and thus PSD. Since this $\mathbf{S}'_k$ is feasible and the second summand is PSD, $\mathbf{B}_k \succeq \mathbf{S}'_k \succeq \operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1})$ follows, which contradicts $\mathbf{B}_k \not\succeq \operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1})$ and simultaneously shows that $\mathbf{S}''_k = \operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1})$ is feasible (setting the off-diagonal elements of $\mathbf{S}'_k$ to zero neither violates the first nor the second constraint). The objective function is bounded above by $\operatorname{tr}(\mathbf{B}_k) - \underline{1}^\top \hat{\mathbf{\Delta}} \underline{1}$ on the intersection of S_n and the set defined by the DD constraints. This follows, because the right hand sides of the DD constraints are bounded below by $\sum_{j=1}^{n} [\hat{\mathbf{\Delta}}_k]_{ij}$. The feasible set is a subset of the aforementioned intersection and the upper bound is attained by the feasible $\mathbf{S}''_k = \operatorname{diag}(\hat{\mathbf{\Delta}} \underline{1})$. Therefore, we have found an optimal solution of (ARBP). □

While the above theorem provides a simple closed-form solution to (ARBP), it also indicates that feasibility issues can arise if the potential buffer errors $\hat{\mathbf{\Delta}}_k$ are too large compared to $\mathbf{\Lambda}_k$, since $\mathbf{B}_k \succeq \operatorname{diag}(\hat{\mathbf{\Delta}} \underline{1})$ must hold for feasibility and $\mathbf{\Lambda}_k \approx \mathbf{B}_k$. The issue arises because, in contrast to the approach for covariance matrices, which does not suffer from this, the constraint $\check{\mathbf{\Lambda}}_k \succeq \mathbf{0}$ had to be introduced. To avoid the above feasibility issues, next, a trigger rule that leads to the feasibility condition being satisfied by construction, is developed.

B. Feasibility-Ensuring Trigger

In this section, the absolute change trigger for covariance matrices, recapitulated in Sec. II-D, is modified to function with information matrices and to guarantee the feasibility of the information matrix bounder from Sec. III-A. The trigger for information matrices is obtained by replacing all occurences of the covariance matrix $\mathbf{P}_k$ in the event-set (1) with the information matrix $\mathbf{\Lambda}_k$ and scaling the trigger thresholds $\mathbf{T}$ by a factor $\alpha_k \in [0,1]$ chosen to ensure feasibility. Thus, the new information matrix trigger yields event-sets of the form

$$E_k = \{((i,j),[\mathbf{\Lambda}_k]_{ij}) | |[\mathbf{\Lambda}_k - \mathbf{B}_{k-1}]_{ij}| > \alpha_k[\mathbf{T}]_{ij}, i \leq j\},$$

that contain the elements $[\mathbf{\Lambda}_k]_{ij}$ of the upper triangle of the information matrix that deviate more than the corresponding scaled threshold $\alpha_k[\mathbf{T}]_{ij}$ (where $\mathbf{T} \in S_n$, $\mathbf{T} \geq \mathbf{0}$) from the corresponding buffer value $[\mathbf{B}_{k-1}]_{ij}$. The event-set is transmitted together with either the scaling factor itself or the data needed to compute it. The possible buffer errors that can occur with the information matrix trigger are given by the set

$$F_k = \{\mathbf{\Delta} \in S_n | |\mathbf{\Delta}| \leq \alpha_k \mathbf{T} \odot \mathbf{E}_k =: \hat{\mathbf{\Delta}}_k\}, \tag{3}$$

where $\mathbf{E}_k \in S_n$ with $[\mathbf{E}_k]_{ij} = 0$ if $((i,j),[\mathbf{\Lambda}_k]_{ij}) \in E_k$ and $[\mathbf{E}_k]_{ij} = 1$ otherwise. The above can be verified by noting that $[\mathbf{\Delta}_k]_{ij} = 0$ if $((i,j),[\mathbf{\Lambda}_k]_{ij}) \in E_k$, since $[\mathbf{B}_k]_{ij} = [\mathbf{\Lambda}_k]_{ij}$ due to the buffer getting updated. Similarly, $((i,j),[\mathbf{\Lambda}_k]_{ij}) \notin E_k$ implies $|[\mathbf{\Lambda}_k - \mathbf{B}_{k-1}]_{ij}| \leq \alpha_k[\mathbf{T}]_{ij}$ and $[\mathbf{B}_k]_{ij} = [\mathbf{B}_{k-1}]_{ij}$ as the buffer is not updated. Thus $|[\mathbf{\Delta}_k]_{ij}| \leq \alpha_k[\mathbf{T}]_{ij}$ holds.

It remains to determine a suitable value for α_k such that the feasibility condition is satisfied. The next theorem gives a *sufficient* condition for the feasibility condition to be satisfied.

Theorem 2. *If $\lambda_{min}(\mathbf{\Lambda}_k) \geq \|\operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1}) + \hat{\mathbf{\Delta}}_k\|_F$ holds then the feasibility condition $\mathbf{B}_k \succeq \operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1})$ is satisfied.*

Proof. First note that $\mathbf{B}_k = \mathbf{\Lambda}_k - \mathbf{\Delta}_k$. Therefore, the feasibility condition can be written as $\mathbf{\Lambda}_k \succeq \mathbf{R}_k := \operatorname{diag}(\hat{\mathbf{\Delta}}_k \underline{1}) + \mathbf{\Delta}_k$. By Lemma 1, $\mathbf{\Lambda}_k \succeq \lambda_{min}(\mathbf{\Lambda}_k)\mathbf{I}$ and $\lambda_{max}(\mathbf{R}_k)\mathbf{I} \succeq \mathbf{R}_k$. Since $\mathbf{R}_k$ is symmetric, $|\lambda_{max}(\mathbf{R}_k)| = \|\mathbf{R}_k\|_2$. The spectral norm is bounded above by the Frobenius norm so that $|\lambda_{max}(\mathbf{R}_k)| \leq \|\mathbf{R}_k\|_F$. Finally, $\|\mathbf{R}_k\|_F \leq \|\operatorname{diag}(\hat{\mathbf{\Delta}}\underline{1}) + \hat{\mathbf{\Delta}}_k\|_F$ because

$$\begin{aligned}\|\mathbf{R}_k\|_F^2 &= \sum_{i=1}^{n}\left([\mathbf{\Delta}_k]_{ii} + \sum_{j=1}^{n}[\hat{\mathbf{\Delta}}_k]_{ij}\right)^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{n}[\mathbf{\Delta}_k]_{ij}^2 \\ &\leq \sum_{i=1}^{n}\left([\hat{\mathbf{\Delta}}]_{ii} + \sum_{j=1}^{n}[\hat{\mathbf{\Delta}}_{ij}]\right)^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{n}[\hat{\mathbf{\Delta}}_k]_{ij}^2\end{aligned}$$

follows from $|[\mathbf{\Delta}_k]_{ij}| \leq [\hat{\mathbf{\Delta}}_k]_{ij}$, $[\mathbf{\Delta}_k]_{ii} + \sum_{j=1}^{n}[\hat{\mathbf{\Delta}}_k]_{ij} \geq 0$, and $(\cdot)^2$ being increasing for non-negative arguments. Consequently, $\lambda_{min}(\mathbf{\Lambda}_k) \geq \|\operatorname{diag}(\hat{\mathbf{\Delta}}_k\underline{1}) + \hat{\mathbf{\Delta}}_k\|_F$ implies $\mathbf{\Lambda}_k \succeq \mathbf{R}_k$, i.e., the feasibility condition $\mathbf{B}_k \succeq \operatorname{diag}(\hat{\mathbf{\Delta}}_k\underline{1})$. ☐

With the information matrix trigger considered here, the sufficient condition stated in the above Theorem 2 becomes

$$\lambda_{min}(\mathbf{\Lambda}_k) \geq \alpha_k \|\operatorname{diag}((\mathbf{T} \odot \mathbf{E}_k)\underline{1}) + \mathbf{T} \odot \mathbf{E}_k\|_F, \tag{4}$$

due to $\hat{\mathbf{\Delta}}_k = \alpha_k \mathbf{T} \odot \mathbf{E}_k$, as defined in (3). The right hand side of (4) depends on $\mathbf{E}_k$ and thus on the trigger decisions, which in turn depend on α_k. This complicates choosing an α_k that satisfies the sufficient condition. Fortunately, the right hand side of (4) attains its trigger decision independent maximum value $\|\operatorname{diag}(\mathbf{T}\underline{1}) + \mathbf{T}\|_F$ for $\mathbf{E}_k = \mathbf{1}$, i.e., when E_k is empty. Therefore, $\lambda_{min}(\mathbf{\Lambda}_k) \geq \alpha_k\|\operatorname{diag}(\mathbf{T}\underline{1}) + \mathbf{T}\|_F$ implies the above sufficient condition for feasibility, meaning that

$$\alpha_k = \min\left\{\frac{\check{\lambda}_k}{\|\operatorname{diag}(\mathbf{T}\underline{1}) + \mathbf{T}\|_F}, 1\right\},$$

where $0 \leq \check{\lambda}_k \leq \lambda_{min}(\mathbf{\Lambda}_k)$, guarantees that the information matrix bounder can operate. Since the bounder needs to know $\hat{\mathbf{\Delta}}_k = \alpha_k \mathbf{T} \odot \mathbf{E}_k$, either α_k or $\check{\lambda}_k$ need to be sent.

In addition to ensuring the feasibility of (ARBP) and thus the functioning of the bounder, it is advantageous to be able to specify elementwise worst-case deviations of $\check{\mathbf{\Lambda}}_k$ from the information matrix $\mathbf{\Lambda}_k$ by judicious selection of $\mathbf{T}$. The following theorem, essentially equivalent to [13, Thm. 2] but for information matrices, provides a way to relate the thresholds $\mathbf{T}$ to the worst-case elementwise bound errors.

Theorem 3. *Let $\check{\mathbf{\Lambda}}_k^* = \mathbf{B}_k - \mathbf{S}_k^*$, where $\mathbf{S}_k^*$ is an optimal solution of* (ARBP)*, then the following bound holds*

$$\|\mathbf{W} \odot (\check{\mathbf{\Lambda}}_k^* - \mathbf{\Lambda}_k)\|_F \leq \|\mathbf{W} \odot (\operatorname{diag}(\hat{\mathbf{\Delta}}_k\underline{1}) + \hat{\mathbf{\Delta}}_k)\|_F,$$

where $\mathbf{W} \in \{0,1\}^{n\times n}$ allows selecting submatrices.

Proof. The proof of [13, Thm. 2] applies to the information matrix case considered here with straightforward changes. ☐

Particularizing the above result for the information matrix bounder considered here with $\hat{\mathbf{\Delta}}_k = \alpha_k \mathbf{T} \odot \mathbf{E}_k$, one obtains

$$\|\mathbf{W}\odot(\check{\mathbf{\Lambda}}_k^* - \mathbf{\Lambda}_k)\|_F \leq \alpha_k\|\mathbf{W}\odot(\operatorname{diag}((\mathbf{T}\odot\mathbf{E}_k)\underline{1}) + \mathbf{T}\odot\mathbf{E}_k)\|_F, \tag{5}$$

which depends on α_k and $\mathbf{E}_k$ and thus the trigger decisions. An error bound that is independent from $\mathbf{E}_k$ and α_k can be derived from the observation that the right hand side of (5) attains its maximum for $\mathbf{E}_k = \mathbf{1}$ and $\alpha_k = 1$. Hence, the deviation of $\check{\mathbf{\Lambda}}_k$ from $\mathbf{\Lambda}_k$ is bounded above, according to

$$\|\mathbf{W} \odot (\check{\mathbf{\Lambda}}_k^* - \mathbf{\Lambda}_k)\|_F \leq \|\mathbf{W} \odot (\operatorname{diag}(\mathbf{T}\underline{1}) + \mathbf{T})\|_F.$$

Choosing $\mathbf{W}$ to extract bounds on individual elements gives

$$|[\check{\mathbf{\Lambda}}_k^* - \mathbf{\Lambda}_k]_{ij}| \leq \begin{cases}[\mathbf{T}]_{ii} + \sum_{j=1}^{n}[\mathbf{T}]_{ij} & , i = j\\ [\mathbf{T}]_{ij} & , \text{otherwise}\end{cases},$$

which allows to first select the off-diagonal thresholds, followed by choosing the diagonal thresholds, both according to some per-element worst-case error specification.

IV. Results and Discussion

In this section, the information matrix compression approach is evaluated in terms of conservativeness and data reduction in a scenario, where extended information filter SLAM results are sent to a receiver, e.g., to fuse them with other results or to make them available for decision-making. The case where the minimum eigenvalue of the information matrix is computed to high precision, is considered first. Next,

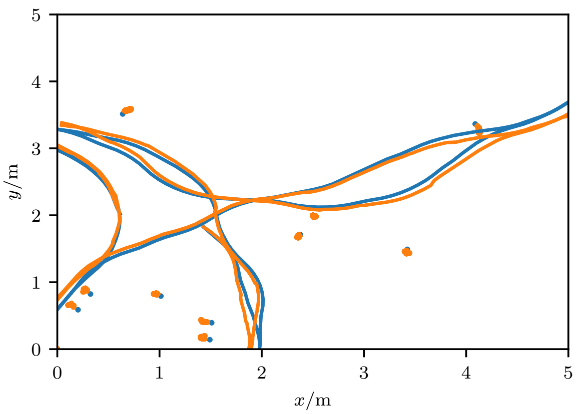

Fig. 3. One of 100 SLAM scenarios used for the evaluation. True robot trajectory (—) and landmark positions (·) shown in blue. Estimated robot trajectory (—) and landmark positions (·) for each time step shown in orange.

the case where a lower bound on said eigenvalue is obtained using a variety of methods, is investigated. Advantages and challenges of the abovementioned cases and different methods are discussed in the context of the proposed approach.

A. Experimental Setup

For the following evaluation 100 2D-SLAM scenarios of 500 time steps each were generated using a sampling rate of 0.1 Hz. In detail, a $5\,\mathrm{m}\times 5\,\mathrm{m}$ area with 10 uniformly distributed landmarks was considered. The robot's initial position was uniformly distributed in the central $2.5\,\mathrm{m}\times 2.5\,\mathrm{m}$ area and its heading initialized according to a uniform distribution on the unit circle. The robot was simulated to move longitudinally at a velocity of $v = 0.5\,\mathrm{m/s}$ and to rotate with an angular velocity of $\omega = 0\,\mathrm{rad/s}$, both superimposed with Gaussian zero-mean noises of variances $\sigma_v^2 = 1\cdot 10^{-2}\,\mathrm{m}^2/\mathrm{s}^2$ and $\sigma_\omega^2 = 1\cdot 10^{-2}\,\mathrm{rad}^2/\mathrm{s}^2$, respectively. Upon hitting the boundary of the considered area, the robot was simulated to stop, turn 180° on the spot, and continue at the previous longitudinal velocity. Range-bearing measurements $\underline{z}_k = (r_k, \theta_k)$ were generated from the resulting robot trajectory and the landmarks every eight time steps and Gaussian zero-mean noises with variances $\sigma_r^2 = 9\cdot 10^{-4}\,\mathrm{m}^2$ and $\sigma_\theta^2 = 9\cdot 10^{-4}\,\mathrm{rad}^2$ were added. The simulated measurements were then used as input to a SLAM algorithm based on an extended information filter (EIF-SLAM). The SLAM algorithm was initialized with the information matrix $\mathbf{\Lambda}_0 = 10^{12}\cdot\mathbf{I}\in S_3^{++}$ and employed the motion and measurement model used to simulate the robot with the same parameters for system and measurement noise. Gating using the Mahalanobis distance $d(\hat{\underline{z}}_k, \underline{z}_k)$ between the predicted range-bearing measurement $\hat{\underline{z}}_k$ and the actual measurement $\underline{z}_k$ combined with nearest neighbor search was used to perform data association (gate $d(\hat{\underline{z}}_k, \underline{z}_k) < 8$) and to detect new landmarks (gate $d(\hat{\underline{z}}_k, \underline{z}_k) > 100$). The ground truth, robot trajectory and landmark estimates from a single simulated EIF-SLAM scenario are shown in Fig. 3.

The information matrices resulting from the estimation process were processed using the proposed event-based approach for information matrix compression with thresholds $\mathbf{T} = T\mathbf{1}$,

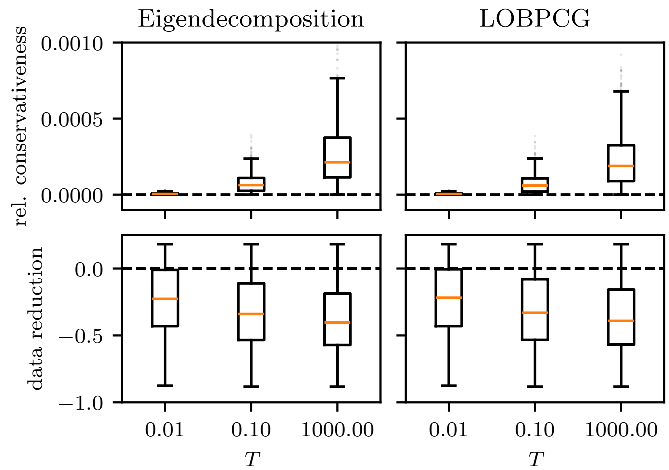

Fig. 4. Summary statistics of relative conservativeness c_k and data reduction d_k over all time steps and EIF-SLAM scenarios, using an eigendecomposition / LOBPCG to compute the minimum eigenvalue of $\mathbf{\Lambda}_k$.

where $T \in \{0.01, 0.1, 1000\}$[1]. Various methods to obtain $\check{\lambda}_k \leq \lambda_{min}(\mathbf{\Lambda}_k)$ were investigated. For an information matrix $\mathbf{\Lambda}_k \in S_n^{++}$ the data reduction was measured using

$$d_k = \frac{64(|E_k|+1) + n(n+1)/2 - 64n(n+1)/2}{64n(n+1)/2},$$

i.e., the number of bits to transmit E_k and α_k relative to that to transmit all elements of the upper triangle of the information matrix, assuming that 64 bit floating-point numbers are used and a bitmask encodes which elements are included in E_k.

The amount of over-conservativeness of the $\check{\mathbf{\Lambda}}_k$ was measured in terms of 'relative conservativeness'

$$c_k = \frac{\mathrm{tr}(\mathbf{\Lambda}_k) - \mathrm{tr}(\check{\mathbf{\Lambda}}_k)}{\mathrm{tr}(\mathbf{\Lambda}_k)}.$$

Relative conservativeness is not just a measure of conservativeness but also of accuracy, as $c_k = 0$ implies $\mathbf{\Lambda}_k = \check{\mathbf{\Lambda}}_k$ due to positive (semi)definiteness of the matrices. The computer used for the evaluation contained an AMD Ryzen 7 PRO 4750U processor clocked at 1.7 GHz and 16 GB RAM.

B. Results with Exact Minimum Eigenvalue

The minimum eigenvalue $\lambda_{min}(\mathbf{\Lambda}_k)$ or a lower bound on it must be known to be able to apply the trigger proposed in Sec. III-B. In this section, the minimum eigenvalue is determined numerically to great precision and $\check{\lambda}_k = \lambda_{min}(\mathbf{\Lambda}_k)$ is used. Specifically, two methods to compute the minimum eigenvalue are considered. The first and more computationally expensive method to obtain $\lambda_{min}(\mathbf{\Lambda}_k)$, is to perform an eigendecomposition of $\mathbf{\Lambda}_k$ and to extract the minimum computed eigenvalue. Alternatively, locally optimal block preconditioned conjugate gradient (LOBPCG) [16], an iterative method that supports warm-starts and can compute subsets of eigenvalues, is considered to determine the minimum eigenvalue. Assuming

[1]When the dimension of the information matrix $\mathbf{\Lambda}_k$ increased, the buffer $\mathbf{B}_k$ was zero-padded and the thresholds $\mathbf{T} = T\mathbf{1}$ were T-padded to match the size of $\mathbf{\Lambda}_k$. This does not affect the proper functioning of the approach.

that $\mathbf{\Lambda}_k$ does not change rapidly, neither do its eigenvalues [15, Thm. 2.4.9.2], so that ideally a few iterations of warm-started LOBPCG suffice to update the minimum eigenvalue with little computational cost compared to an eigendecomposition.

The data reduction and relative conservativeness obtained using the minimum eigenvalue determined through both, eigendecomposition and LOBPCG (for the three smallest eigenvalues using isotropic Gaussian initialization, warm-starts, and block-diagonal preconditioning) are shown in Fig. 4. The displayed summary statistics were computed over all time steps and all scenarios. As can be seen, using eigendecomposition / LOBPCG, data reduction increases for increasing T, with minimum median data reduction of $-2.27 \cdot 10^{-1}$ / $-2.18 \cdot 10^{-1}$ achieved for $T = 0.1$ and maximum median data reduction of $-4.03 \cdot 10^{-1}$ / $-3.91 \cdot 10^{-1}$ achieved for $T = 1000$. For $T = 1000$ using eigendecomposition / LOBPCG, it was observed that $\alpha_k < 1$ for essentially all information matrix instances, implying that data reduction is limited by how small $\check{\lambda}_k = \lambda_{min}(\mathbf{\Lambda}_k)$ is. Therefore, the abovementioned maximum median data reduction is the best attainable using the proposed approach and sufficient feasibility condition. Note that data reduction for $T = 1000$ using LOBPCG is more skewed towards zero, than using eigendecomposition. This is caused by LOBPCG occasionally not reaching the desired tolerance, in which case the inaccurately computed minimum eigenvalue is discarded, and $\check{\lambda}_k = 0$ is used instead, leading to the transmission of all information matrix elements that have changed, even infinitesimally, compared to the buffer. Furthermore, for every T there are instances where data reduction is positive, implying an increase in the amount of transmitted data. This is due to the scaling factor α_k and the bitmask (see Sec. IV-A) that are transmitted in addition to the matrix elements themselves, as well as the small dimension of the information matrix at the start of each EIF-SLAM information matrix sequence, when no landmarks have been added to the state estimate yet.

The relative conservativeness, both using eigendecomposition and LOBPCG, can be seen to be non-negative, due to the conservativeness of the bounder output. While the relative conservativeness increases as T increases, it is generally small, indicating that only mild over-conservativeness occurs. Using eigendecomposition / LOBPCG the lowest median relative conservativeness of $4.34 \cdot 10^{-6}$ / $4.11 \cdot 10^{-6}$ is achieved for $T = 0.01$ and the worst median relative conservativeness of $2.13 \cdot 10^{-4}$ / $1.89 \cdot 10^{-4}$ for $T = 1000$, where in the latter case the trigger uses the largest scaled thresholds that still guarantee feasibility of the bounder. Finally, it can be observed that the relative conservativeness using LOBPCG and $T = 1000$ is slightly lower than when using eigendecomposition. This is simply due to more data being transmitted on average when using LOBPCG, as discussed in the previous paragraph.

The eigendecomposition has the same asymptotic computational complexity of $\mathcal{O}(n^3)$ as a matrix inversion but is usually significantly more computationally expensive. Consequently, computing the minimum eigenvalue through eigendecomposition is only advisable, when maximum possible data

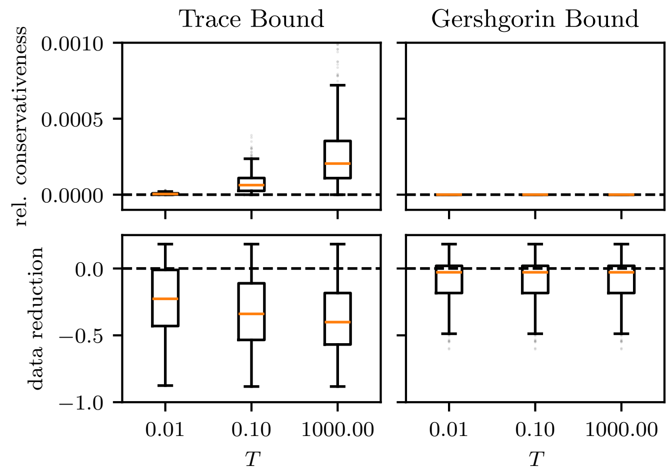

Fig. 5. Summary statistics of relative conservativeness c_k and data reduction d_k over all time steps and EIF-SLAM scenarios, using Gershgorin discs / the trace bound (6) to bound $\lambda_{min}(\mathbf{\Lambda}_k)$ from below.

reduction (i.e., $\check{\lambda}_k = \lambda_{min}(\mathbf{\Lambda}_k)$) is required and inverting the information matrix at the transmitter, transmitting the resulting covariance matrix using [12], [13], and inverting the conservative covariance matrix produced at the receiver is not feasible, e.g., because the receiver does not have the required computational resources. Especially for large ($n \gg 1000$) and sparse information matrices using LOBPCG will likely be faster than computing an eigendecomposition. In the considered EIF-SLAM example however, LOBPCG exhibited a median execution time of $9.4\,\mathrm{ms}$ as compared to the eigendecomposition's median execution time of $0.96\,\mathrm{ms}$, which can likely be attributed to the fact that the considered information matrices were not sparse and of low dimensionality, as well as LOBPCG being implemented in an interpreted language.

C. Results with Eigenvalue Lower Bounds

While the methods in IV-B attempt to compute the minimum eigenvalue exactly, true lower bounds that are easy to compute can also be used: One possibility are eigenvalue inclusion sets such as Gershgorin discs [15, Thm. 6.1.1] or Brauer ellipsoids [15, Thm. 6.4.7] from which eigenvalue lower bounds $\check{\lambda}_k \leq \lambda_{min}(\mathbf{\Lambda}_k)$ can be derived. An issue with the above inclusion sets is that for information matrices with large off-diagonal elements, the resulting lower bounds are not guaranteed to be positive. This can lead to the best obtainable non-negative eigenvalue bound being $\check{\lambda}_k = 0$, in which case data reduction is severly limited, as it only occurs when there are elements of $\mathbf{\Lambda}_k$ that do not change over time. However, guaranteed positive lower bounds exist, a simple yet effective example being

$$0 \leq \check{\lambda}_k = \frac{1}{\sqrt{\mathrm{tr}(\mathbf{\Lambda}_k^{-2})}} = \frac{1}{\sqrt{\|\mathbf{\Lambda}_k^{-1}\|_F^2}} \leq \lambda_{min}(\mathbf{\Lambda}_k). \quad (6)$$

The above bound follows from the fact that the trace of a matrix is the sum of its eigenvalues and that the eigenvalues of the inverse matrix are the inverses of the original eigenvalues.

Fig. 5 shows the data reduction and relative conservativeness obtained using the trace bound (6) and Gershgorin discs.

As in Sec. IV-B the summary statistics were obtained from all time steps in all scenarios. The results using the trace bound show essentially the same qualitative behavior as those obtained in Sec. IV-B. Quantitatively, the lowest median data reduction of $-2.27 \cdot 10^{-1}$ is achieved for $T = 0.01$ and the highest of $-4.01 \cdot 10^{-1}$ for $T = 1000$. The lowest median relative conservativeness of $4.35 \cdot 10^{-5}$ is achieved for $T = 0.01$ and the highest of $2.05 \cdot 10^{-4}$ for $T = 1000$. These results are only slightly worse than those obtained from using an eigendecomposition. This indicates that the eigenvalue lower bound (6) is very accurate for the information matrix sequences occuring in the considered EIF-SLAM scenarios. In contrast, using the Gershgorin disc bound yields a negligible median data reduction of $-2.83 \cdot 10^{-2}$ and consequently virtually zero relative conservativeness for all T. This is due to the Gershgorin-based bound being negative for essentially all information matrix instances, such that $\check{\lambda}_k = 0$ must be used, causing all elements that are not unchanged compared to the buffer to be transmitted. The fact that any data reduction occurs at all is a consequence of the landmark variances staying constant during the prediction step of EIF-SLAM. The above result clearly demonstrates that eigenvalue lower bounds based on Gershgorin discs cannot be used effectively in the considered EIF-SLAM scenarios. Nevertheless, it is conceivable that in other scenarios where the matrices under consideration have smaller off-diagonal elements, Gershgorin-based eigenvalue bounds could provide meaningful results.

Computing a lower bound $\check{\lambda}_k$ using Gershgorin discs has asymptotic computational complexity $\mathcal{O}(n^2)$ and is very fast in practice. In contrast, the lower bound (6) requires the inversion of $\mathbf{\Lambda}_k$, e.g., using a Cholesky decomposition, which has asymptotic computational complexity $\mathcal{O}(n^3)$, but is still significantly cheaper than an eigendecomposition in practice. Since only a single inversion needs to be carried out, using the lower bound (6) is preferable over inverting the information matrix, transmitting the resulting covariance matrix using [12], [13], and inverting the conservative covariance matrix produced at the receiver. In the EIF-SLAM scenario, computing the Gershgorin-based bound had a median runtime of $37.5\,\mu$s whereas computing the lower bound (6) had one of $76.0\,\mu$s.

V. Conclusion

This work proposed an event-based method for the compressed transmission of information matrix sequences. The method only transmits information matrix elements when they have changed sufficiently, compared to the last transmitted value. Additionally, a scalar value, which depends on a lower bound on the minimum eigenvalue of the information matrix, must be transmitted. Closed-form robust optimization is then used to extract a conservative information matrix from the entirety of received values. The experiments demonstrated that, using various approaches to compute the minimum eigenvalue lower bound, the method can achieve a median data reduction of up to 40.3% with a median decrease in information matrix trace (as a measure of conservativeness) of $2.13 \cdot 10^{-2}\%$ in an extended information filter SLAM scenario. However, data reduction cannot be traded for conservativeness at will, since a (sufficient) feasibility condition must be satisfied for the robust optimization problem to be solved at the receiver. The sufficient condition derived in this work is practical but conservative, which negatively impacts the achievable data reduction. Furthermore, there exists a tradeoff between the accuracy of the minimum eigenvalue lower bounds, which directly affects the achievable data reduction, and computational expense. Therefore, it is suggested that future work investigate less conservative sufficient conditions for feasibility, as well as fast and accurate methods to bound the minimum eigenvalue of the information matrices to be transmitted from below.

References

[1] M. R. Walter, R. M. Eustice, and J. J. Leonard, “Exactly Sparse Extended Information Filters for Feature-based SLAM,” *The International Journal of Robotics Research*, vol. 26, no. 4, pp. 335–359, Apr. 2007.

[2] G. Grisetti, R. Kümmerle, C. Stachniss, and W. Burgard, “A Tutorial on Graph-Based SLAM,” *IEEE Intelligent Transportation Systems Magazine*, vol. 2, no. 4, pp. 31–43, 2010.

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

[4] R. Forsling, Z. Sjanic, F. Gustafsson, and G. Hendeby, “Consistent Distributed Track Fusion Under Communication Constraints,” in *2019 22th International Conference on Information Fusion (FUSION)*, Jul. 2019, pp. 1–8.

[5] J. K. Uhlmann, “Covariance consistency methods for fault-tolerant distributed data fusion,” *Information Fusion*, vol. 4, no. 3, pp. 201–215, Sep. 2003.

[6] R. Forsling, Z. Sjanic, F. Gustafsson, and G. Hendeby, “Communication Efficient Decentralized Track Fusion Using Selective Information Extraction,” in *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, Jul. 2020, pp. 1–8.

[7] ——, “Optimal Linear Fusion of Dimension-Reduced Estimates Using Eigenvalue Optimization,” in *2022 25th International Conference on Information Fusion (FUSION)*, Jul. 2022, pp. 1–8.

[8] R. Forsling, F. Gustafsson, Z. Sjanic, and G. Hendeby, “Decentralized Data Fusion of Dimension-Reduced Estimates Using Local Information Only,” in *2023 IEEE Aerospace Conference*, Mar. 2023, pp. 1–10.

[9] R. Forsling, “The Dark Side of Decentralized Target Tracking : Unknown Correlations and Communication Constraints,” Ph.D. dissertation, Linköping University Electronic Press, Linköping, 2023.

[10] C. Funk, B. Noack, and U. D. Hanebeck, “Conservative Quantization of Fast Covariance Intersection,” in *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, Sep. 2020, pp. 68–74.

[11] ——, “Conservative Quantization of Covariance Matrices with Applications to Decentralized Information Fusion,” *Sensors*, vol. 21, no. 9, p. 3059, Jan. 2021.

[12] C. Funk and B. Noack, “Conservative Data Reduction for Covariance Matrices Using Elementwise Event Triggers,” in *2023 26th International Conference on Information Fusion (FUSION)*. Charleston, SC, USA: IEEE, Jun. 2023, pp. 1–6.

[13] ——, “An event-based approach for the conservative compression of covariance matrices,” *arXiv:2403.05977 [cs.RO]*, Mar. 2024, submitted to IEEE Transactions on Automatic Control.

[14] S. Thrun, D. Koller, Z. Ghahramani, H. Durrant-Whyte, and A. Y. Ng, “Simultaneous Mapping and Localization with Sparse Extended Information Filters: Theory and Initial Results,” in *Algorithmic Foundations of Robotics V*, ser. Springer Tracts in Advanced Robotics, J.-D. Boissonnat, J. Burdick, K. Goldberg, and S. Hutchinson, Eds. Berlin, Heidelberg: Springer, 2004, pp. 363–380.

[15] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge University Press, 2012.

[16] A. V. Knyazev, “Toward the Optimal Preconditioned Eigensolver: Locally Optimal Block Preconditioned Conjugate Gradient Method,” *SIAM Journal on Scientific Computing*, vol. 23, no. 2, pp. 517–541, Jan. 2001.