Sparse Extended Information Filters: Insights into Sparsification

Ryan Eustice
Department of Applied Ocean Physics and Engineering
Woods Hole Oceanographic Institution
Woods Hole, MA, USA
ryan@whoi.edu

Matthew Walter and John Leonard
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, MA, USA
{mwalter,jleonard}@mit.edu

***Abstract*— Recently, there have been a number of variant Simultaneous Localization and Mapping (SLAM) algorithms which have made substantial progress towards large-area scalability by parameterizing the SLAM posterior within the information (canonical/inverse covariance) form. Of these, probably the most well-known and popular approach is the Sparse Extended Information Filter (SEIF) by Thrun et al. While SEIFs have been successfully implemented with a variety of challenging real-world data sets and have lead to new insights into scalable SLAM, open research questions remain regarding the approximate sparsification procedure and its effect on map error and consistency.**

In this paper, we examine the constant-time SEIF sparsification procedure in depth and offer new insight into issues of consistency. In particular, we show that exaggerated map inconsistency occurs within the *global* reference frame where estimation is performed, but that empirical testing shows that relative *local* map relationships are preserved. We then present a slightly modified version of their sparsification procedure which is shown to preserve sparsity while also generating both local *and* global map estimates comparable to those obtained by the non-sparsified SLAM filter; this modified approximation, however, is no longer constant-time. We demonstrate our findings by benchmark comparison of the modified and original SEIF sparsification rule using simulation in the linear Gaussian SLAM case and real-world experiments for a nonlinear dataset.

I. Introduction

Since its inception with the fundamental work of Smith et al. [1] and Moutarlier and Chatila [2], roboticists have been trying to address scalability issues associated with an Extended Kalman Filter (EKF) based approach to SLAM. While this approach is often considered the "standard" [3] and is attractive in its simplicity (because it only requires tracking first and second moments of the joint landmark-robot distribution), a well known fact is that EKF SLAM inference requires quadratic complexity in the number of landmarks *per update* to maintain the joint-posterior correlations. As a consequence, the direct application of the EKF to SLAM is limited to relatively small environments (e.g., less than than 100 landmarks).

Recently, a new class of scalable SLAM algorithms have been proposed by Thrun et al. [4], Paskin [5], and Frese [6], [7] — all based upon the canonical-form which has the nice interpretation as a Gaussian graphical model [5], [8]. As Thrun et al. [4] empirically first showed, and Frese later analytically proved [9], the inverse covariance matrix (i.e., information matrix) of feature-based SLAM exhibits a "natural" sparseness where many of the off-diagonal elements (i.e., graphical constraints) are relatively "weak". This insight has spawned the development of scalable SLAM algorithms founded upon pruning these weak constraints and exploiting the resulting sparse representation [4], [5], [7].

For example, Paskin (Thin-Junction-Tree Filters) [5] and Frese (Treemap Filters) [7] both employ tree-based approximations to sparsify the canonical-form and have developed very efficient inference algorithms for this representation. One drawback to their techniques, though, is that their tree-representations cannot explicitly model cyclic environments nor has data association been addressed. Alternatively, the Sparse Extended Information Filter (SEIF) proposed by Thrun et al. [4], probably the most well known SLAM information formulation, is based upon representing the SLAM posterior through the dual of the EKF (i.e., an Extended Information Filter). SEIFs maintain a sparse information matrix representation which has been demonstrated to be efficient, scalable, allows for explicit representation of cyclic environments, and addresses data association [10]. The delicate issue, however, which is at the core of the SEIFs paper, is *how* to perform the necessary sparsification step required to keep the information matrix representation sparse by eliminating weak robot-landmark constraints.

In the following, we explore in depth the approximation employed by SEIFs to enforce sparseness. We show that a particular assumption in SEIF's sparsification derivation leads to inconsistency of the global map error covariance estimates, however, empirical testing indicates that local map relations and relative uncertainties are preserved. In addition, we present a slightly modified derivation which yields a sparsification rule that is shown to produce both global and local map estimates comparable to the full-covariance EKF while maintaining the same sparse representation as SEIFs — however, sparsification is no longer constant-time. We demonstrate our insights by concluding with a benchmark comparison for a linear Gaussian SLAM simulation and in addition present results for a nonlinear experimental dataset.

II. Background

Expanding the quadratic in the exponential of the Gaussian random variable $\boldsymbol{\xi}_t \sim \mathcal{N}(\boldsymbol{\mu}_t, \Sigma_t)$ yields the canonical parameterization $\boldsymbol{\xi}_t \sim \mathcal{N}^{-1}(\boldsymbol{\eta}_t, \Lambda_t)$ where $\boldsymbol{\eta}_t$ and Λ_t are the information vector and information matrix, respectively [11]. Equation (1) shows how the two forms are mathematically related while Table I expresses the dual relationship they have with respect to marginalization and conditioning. For a general discussion of the mechanics and properties of the ensuing information filter the reader is referred to [4], [11].

$$\Lambda_t = \Sigma_t^{-1} \qquad \boldsymbol{\eta}_t = \Lambda_t \boldsymbol{\mu}_t \tag{1}$$

TABLE I

Summary of Marginalization and Conditioning Operations on a Gaussian Distribution Expressed in Covariance and Information Form

$$p(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}_\alpha \\ \boldsymbol{\mu}_\beta \end{bmatrix}, \begin{bmatrix} \Sigma_{\alpha\alpha} & \Sigma_{\alpha\beta} \\ \Sigma_{\beta\alpha} & \Sigma_{\beta\beta} \end{bmatrix}\right) = \mathcal{N}^{-1}\left(\begin{bmatrix} \boldsymbol{\eta}_\alpha \\ \boldsymbol{\eta}_\beta \end{bmatrix}, \begin{bmatrix} \Lambda_{\alpha\alpha} & \Lambda_{\alpha\beta} \\ \Lambda_{\beta\alpha} & \Lambda_{\beta\beta} \end{bmatrix}\right)$$

	Marginalization $p(\boldsymbol{\alpha}) = \int p(\boldsymbol{\alpha}, \boldsymbol{\beta}) d\boldsymbol{\beta}$	Conditioning $p(\boldsymbol{\alpha} \mid \boldsymbol{\beta}) = p(\boldsymbol{\alpha}, \boldsymbol{\beta}) / p(\boldsymbol{\beta})$
Cov. Form	$\boldsymbol{\mu} = \boldsymbol{\mu}_\alpha$ $\Sigma = \Sigma_{\alpha\alpha}$	$\boldsymbol{\mu}' = \boldsymbol{\mu}_\alpha + \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)$ $\Sigma' = \Sigma_{\alpha\alpha} - \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta\alpha}$
Info. Form	$\boldsymbol{\eta} = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\boldsymbol{\eta}_\beta$ $\Lambda = \Lambda_{\alpha\alpha} - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\Lambda_{\beta\alpha}$	$\boldsymbol{\eta}' = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\boldsymbol{\beta}$ $\Lambda' = \Lambda_{\alpha\alpha}$

A. Controlling Feature-Based SLAM Sparsity

Most SLAM approaches are feature-based which assumes that the robot can extract an abstract representation of features in the environment from its sensor data and then use re-observation of these features for localization [1]. In this approach a landmark map is explicitly built and maintained. The process of concurrently performing localization *and* feature map building are inherently coupled, therefore, the robot must then represent a joint-distribution over landmarks and current pose, i.e.,

$$p(\boldsymbol{\xi}_t \mid \mathbf{z}^t, \mathbf{u}^t) = \mathcal{N}(\boldsymbol{\mu}_t, \Sigma_t) = \mathcal{N}^{-1}(\boldsymbol{\eta}_t, \Lambda_t) \tag{2}$$

where $\boldsymbol{\xi}_t = [\mathbf{x}_t^\top, \mathbf{M}^\top]^\top$ represents the current robot state and landmark map respectively, $\mathbf{z}^t$ are the measurements, and $\mathbf{u}^t$ the control inputs. In (2) we have explicitly modeled this distribution as being jointly-Gaussian based upon additive white noise models and first-order linearizations of our process and observation models as described in [1], [4]. The key behind scalable SLAM algorithms in the canonical-form is based upon the insight that the information matrix Λ_t *naturally* tends to exhibit strong and weak constraints as shown in Fig. 1.

What Thrun et al. [4] insightfully observed was that the time-projection step is the cause for creating these weak constraints, and furthermore that by bounding the number of nonzero off-diagonal elements linking the robot to landmarks, that they could eliminate the generation of many of these weak links and enforce exact sparsity in the information matrix. Their concept was to partition the landmark map variable $\mathbf{M}$ (i.e., $\mathbf{M} = \{\mathbf{m}^+, \mathbf{m}^-\}$) into a set of active features $\mathbf{m}^+$ (i.e., those with a nonzero off-diagonal element linking them to the robot $\mathbf{x}_t$) and a set of passive features $\mathbf{m}^-$ (i.e., those with no link to $\mathbf{x}_t$). They showed that by enforcing an upper bound on the number of active features $\mathbf{m}^+$, that they could control the resulting fill-in of the information matrix.

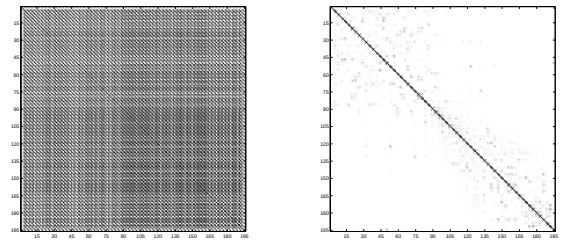


Fig. 1. A comparison of the structure of the covariance and information matrices as is typically seen in feature-based SLAM implementations; darker shades represent larger magnitudes. (left) The correlation matrix is dense and requires quadratic storage. (right) The normalized information matrix exhibits "natural" sparsity with a majority of the elements being orders of magnitude smaller than the few dominant entries.

A simple explanation for the effectiveness of their strategy comes from viewing motion prediction as a two-step process of state augmentation of $\mathbf{x}_{t+1}$ followed by marginalization over $\mathbf{x}_t$. Referring to Fig. 2(a) we see that filtering naturally tends to fill-in the information matrix by creating new links between all active features through elimination of $\mathbf{x}_t$ (see Table I for details of marginalization in the information form) while passive features remain unaffected; for a more in depth discussion of this phenomenon the reader is referred to [5], [12]. Insightfully, as Fig. 2(b) shows, we can *control* the active feature fill-in of the information matrix by *bounding* the number of links connected to $\mathbf{x}_t$ *before* marginalization occurs. This key insight motivates the concept behind sparsification which is the process of how we remove these links to satisfy our active feature bound.

Before moving on to discuss *how* SEIFs actually enforce the upper bound on the number of active features, it will prove useful to first elucidate the conditional independence relationship implied by active and passive features.

B. Implied Conditional Independence

A very useful property of the canonical-form is that the information matrix has the direct interpretation as a non-directed Bayes Net [13] where: random variables are nodes, nonzero off-diagonal elements are edges/constraints, and zero-valued off-diagonal elements are *missing* edges implying available conditional independencies [8]. Applying this property to SEIF's partitioning of landmarks into active and passive features, we see that Fig. 3 correctly illustrates the resulting block information matrix and non-directed Bayes net for the SEIF SLAM posterior over robot and features. In particular, Fig. 3 clearly shows that there is a missing edge between the robot $\mathbf{x}_t$ and the passive features $\mathbf{m}^-$ implying that the two are conditionally independent *given* the active features $\mathbf{m}^+$.

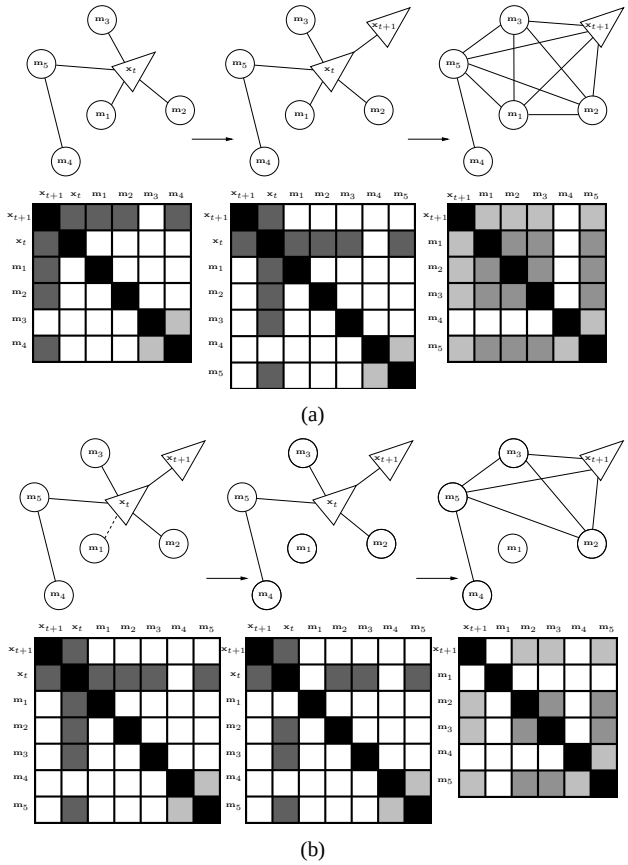


Fig. 2. A graphical explanation of SEIF's methodology for controlling sparsity in the information matrix. (a) A sequence of illustrations depicting the evolution of the Bayes Net and corresponding information matrix resulting from time projection when viewed as a two-step process of state augmentation followed by marginalization. Gray shades imply magnitude with white being exactly zero. From left to right we have: (1) the robot $\mathbf{x}_t$ connected to four active features, $\mathbf{m}_{1:3}$ and $\mathbf{m}_5$; (2) state augmentation of the time-propagated robot pose $\mathbf{x}_{t+1}$; (3) marginalized distribution where the old pose, $\mathbf{x}_t$, has been eliminated. (b) A sequence of illustrations highlighting the concept behind sparsification. If feature $\mathbf{m}_1$ can first be made passive by eliminating its link to the old pose, $\mathbf{x}_t$, then marginalization over $\mathbf{x}_t$ will not link it to any of the other active features. This implies that we can control fill-in of the information matrix by bounding the number of currently active features.

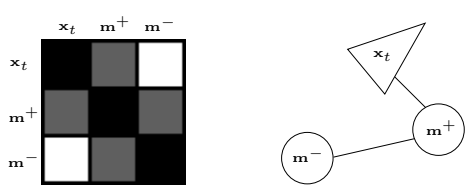


Fig. 3. An illustration of SEIF's concept of active and passive features and their relation to the robot. (left) A schematic of the block 3×3 SEIF information matrix. Dark squares correspond to nonzero block-elements while white squares corresponds to exactly zero block elements. (right) The SEIFs information matrix expressed as a non-directed Bayes-Net. The missing edge between $\mathbf{x}_t$ and $\mathbf{m}^-$ implies available conditional independence.

Mathematically, we can also easily prove this relationship by noting that conditional independence for a Gaussian distribution implies that the conditional posterior $p(\mathbf{x}_t, \mathbf{m}^- \mid \mathbf{m}^+, \mathbf{z}^t, \mathbf{u}^t)$ must have a block-diagonal covariance matrix. In the information form, conditioning on the active features $\mathbf{m}^+$ corresponds to simply extracting the $\{\mathbf{x}_t, \mathbf{m}^-\}$ sub-block from the information matrix Λ_t (see Table I). Referring again to Fig. 3, we note that extracting this sub-block results in a block-diagonal conditional information matrix over $\mathbf{x}_t$ and $\mathbf{m}^-$ whose inverse is a block-diagonal covariance matrix, thus, conditional independence is proved.

As we show next, we can exploit this conditional independence relationship to derive a sparsification rule which allows us to bound the number of active features.

III. Sparsification

In feature-based SLAM, landmarks become active through observation causing them to become linked to the robot through a shared off-diagonal constraint — this constraint decays over time if the landmark is not re-observed, but will never become exactly zero (i.e., passive) unless it is "sparsified". Sparsification refers to the operation where these weak robot-landmark constraints are pruned and features are made passive. It is a useful *approximation* which allows sparsity to be enforced in the information matrix by bounding the number of active features as described in §II-A.

A. SEIF Sparsification Rule

Sparsification is required whenever the active feature threshold is exceeded through landmark observation. SEIF's strategy for sparsification is based upon partitioning the landmark map $\mathbf{M}$ into a union of three disjoint sets $\mathbf{M} = \{\mathbf{m}^0 \cup \mathbf{m}^+ \cup \mathbf{m}^-\}$ where in a slight abuse of our previous notation: $\mathbf{m}^-$ are the currently passive features which will *remain* passive after sparsifying, $\mathbf{m}^+$ are the currently active features which will *remain* active after sparsifying, and $\mathbf{m}^0$ are the currently active features which will *become* passive after sparsifying.

We begin our derivation of the SEIF sparsification approximation by factorizing the SLAM posterior over the robot and map as:

$$\begin{aligned} p\big(\mathbf{x}_t, &\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) \\ &= p\big(\mathbf{x}_t \mid \mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) p\big(\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) && \text{(3a)} \\ &= p\big(\mathbf{x}_t \mid \mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^- = \boldsymbol{\alpha}\big) p\big(\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) && \text{(3b)} \end{aligned}$$

where for notational convenience we have omitted explicitly writing out the conditioning on $\mathbf{z}^t$ and $\mathbf{u}^t$. The above factorization uses the available conditional independence discussed in §II-B between the robot and passive features to arbitrarily assign a value to the passive features in the conditional (3b) (i.e., $\mathbf{m}^- = \boldsymbol{\alpha}$) without influencing the conditional robot posterior (i.e. $p\big(\mathbf{x}_t \mid \mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) \equiv p\big(\mathbf{x}_t \mid \mathbf{m}^0, \mathbf{m}^+\big)$). Note that in the derivation presented in [4] $\boldsymbol{\alpha}$ is simply set to zero while we leave it a free parameter for the purposes of exposition.

The SEIF sparsification approximation is derived from (3b) by imposing that $\mathbf{m}^0$ be passive via dropping it from the robot posterior as

$$\begin{aligned} \tilde{p}_{\text{SEIFs}}\big(\mathbf{x}_t, &\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) \\ &= p\big(\mathbf{x}_t \mid \mathbf{m}^+, \mathbf{m}^- = \boldsymbol{\alpha}\big) p\big(\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) && \text{(4a)} \\ &= \frac{p_B\big(\mathbf{x}_t, \mathbf{m}^+ \mid \mathbf{m}^- = \boldsymbol{\alpha}\big)}{p_C\big(\mathbf{m}^+ \mid \mathbf{m}^- = \boldsymbol{\alpha}\big)} p_D\big(\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) && \text{(4b)} \end{aligned}$$

where (4b) merely expresses the conditional of (4a) as a ratio and the subscripts p_B, p_C, p_D are used for notational convenience to reference the different pdfs involved in its calculation. While the factorization expressed in (3b) is theoretically *exact* due to the conditional independence between $\mathbf{x}_t$ and $\mathbf{m}^-$ given the active features, equation (4) is in *error* because $\mathbf{x}_t$ is no longer conditionally independent of $\mathbf{m}^-$ given only a partial set of the active features (i.e., the set of all active features is $\mathbf{m}^0 \cup \mathbf{m}^+$). This implies that the particular value of $\boldsymbol{\alpha}$ we choose modifies the posterior approximation.

Equations (5)–(8) summarize the SEIF sparsified posterior (4) as expressed in both covariance and information form — due to space limitations we omit their derivation and only present the resulting expressions. For ease of comparison we use the same notation as [4] where S denotes a projection matrix over the state space $\boldsymbol{\xi}_t$ (e.g., $\mathbf{x}_t = S_{x_t}^\top \boldsymbol{\xi}_t$ extracts the robot pose). Note that the mean update in equation (6) clearly shows that the original mean vector $\boldsymbol{\mu}_t$ is modified during the sparsification step for values of $\boldsymbol{\alpha} \neq S_{m^-}^\top \boldsymbol{\mu}_t$ indicating $\boldsymbol{\alpha}$'s influence on the term $p\big(\mathbf{x}_t \mid \mathbf{m}^+, \mathbf{m}^- = \boldsymbol{\alpha}\big)$ used in the approximation (4)[1].

Covariance Form

$$\tilde{\Sigma}_t = \big(S_{x_t,m^+}\Sigma_B^{-1}S_{x_t,m^+}^\top - S_{m^+}\Sigma_C^{-1}S_{m^+}^\top + S_{m^0,m^+,m^-}\Sigma_D^{-1}S_{m^0,m^+,m^-}^\top\big)^{-1} \quad (5)$$

$$\tilde{\boldsymbol{\mu}}_t = \boldsymbol{\mu}_t + \tilde{\Sigma}_t\big(S_{x_t,m^+}\Sigma_B^{-1}S_{x_t,m^+}^\top - S_{m^+}\Sigma_C^{-1}S_{m^+}^\top\big)\times \Sigma_t S_{m^-}\big(S_{m^-}^\top\Sigma_t S_{m^-}\big)^{-1}\big(\boldsymbol{\alpha} - S_{m^-}^\top\boldsymbol{\mu}_t\big) \quad (6)$$

where

$$\Sigma_B = S_{x,m^+}^\top\Big(\mathrm{I} - \Sigma_t S_{m^-}\big(S_{m^-}^\top\Sigma_t S_{m^-}\big)^{-1}S_{m^-}^\top\Big)\Sigma_t S_{x,m^+}$$
$$\Sigma_C = S_{m^+}^\top\Big(\mathrm{I} - \Sigma_t S_{m^-}\big(S_{m^-}^\top\Sigma_t S_{m^-}\big)^{-1}S_{m^-}^\top\Big)\Sigma_t S_{m^+}$$
$$\Sigma_D = S_{m^0,m^+,m^-}^\top\Sigma_t S_{m^0,m^+,m^-}$$

Information Form

$$\tilde{\Lambda}_t = S_{x_t,m^+}\Lambda_B S_{x_t,m^+}^\top - S_{m^+}\Lambda_C S_{m^+}^\top + S_{m^0,m^+,m^-}\Lambda_D S_{m^0,m^+,m^-}^\top \quad (7)$$

$$\tilde{\boldsymbol{\eta}}_t = S_{x_t,m^+}\boldsymbol{\eta}_B - S_{m^+}\boldsymbol{\eta}_C + S_{m^0,m^+,m^-}\boldsymbol{\eta}_D \quad (8)$$

where

$$\boldsymbol{\eta}_\alpha = \Sigma_t S_{m^-}\boldsymbol{\alpha}$$
$$\Lambda_B = S_{x_t,m^+}^\top\Big(\mathrm{I} - \Lambda_t S_{m^0}\big(S_{m^0}^\top\Lambda_t S_{m^0}\big)^{-1}S_{m^0}^\top\Big)\Lambda_t S_{x_t,m^+}$$
$$\boldsymbol{\eta}_B = S_{x_t,m^+}^\top\Big(\mathrm{I} - \Lambda_t S_{m^0}\big(S_{m^0}^\top\Lambda_t S_{m^0}\big)^{-1}S_{m^0}^\top\Big)(\boldsymbol{\eta}_t - \boldsymbol{\eta}_\alpha)$$
$$\Lambda_C = S_{m^+}^\top\Big(\mathrm{I} - \Lambda_t S_{x_t,m^0}\big(S_{x_t,m^0}^\top\Lambda_t S_{x_t,m^0}\big)^{-1}S_{x_t,m^0}^\top\Big)\Lambda_t S_{m^+}$$
$$\boldsymbol{\eta}_C = S_{m^+}^\top\Big(\mathrm{I} - \Lambda_t S_{x_t,m^0}\big(S_{x_t,m^0}^\top\Lambda_t S_{x_t,m^0}\big)^{-1}S_{x_t,m^0}^\top\Big)(\boldsymbol{\eta}_t - \boldsymbol{\eta}_\alpha)$$
$$\Lambda_D = S_{m^0,m^+,m^-}^\top\Big(\mathrm{I} - \Lambda_t S_{x_t}\big(S_{x_t}^\top\Lambda_t S_{x_t}\big)^{-1}S_{x_t}^\top\Big)\Lambda_t S_{m^0,m^+,m^-}$$
$$\boldsymbol{\eta}_D = S_{m^0,m^+,m^-}^\top\Big(\mathrm{I} - \Lambda_t S_{x_t}\big(S_{x_t}^\top\Lambda_t S_{x_t}\big)^{-1}S_{x_t}^\top\Big)\boldsymbol{\eta}_t$$

[1]The expression for the sparsified information vector as presented in [4] corresponds to setting $\boldsymbol{\alpha} = S_{m^-}^\top\boldsymbol{\mu}_t$, (i.e., the mean of the passive features) and not $\boldsymbol{\alpha} = \mathbf{0}$ as stated in their paper.

B. Modified Sparsification Rule

In the previous section we showed that the derivation of the SEIF sparsification rule introduces a conditioning on a specific realization of the passive features — i.e., their mean estimate. This conditioning *influences* the outcome of the sparsification approximation and in particular can modify the resulting mean estimate as evident by the functional dependence on $\boldsymbol{\alpha}$ in (6). In the following we show that we can easily modify the original SEIFs approximation to derive a more correct version of the sparsification rule by explicitly using $\mathbf{x}_t$'s conditional independence of the passive features $\mathbf{m}^-$ to drop its dependence. This modified version of the SEIFs sparsification rule will be shown to preserve the state mean and, as demonstrated in §IV, provide a high fidelity approximation yielding results comparable to the full-covariance EKF.

We begin by factorizing the posterior $p\big(\mathbf{x}_t, \mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big)$ using Bayes rule like in equation (3a) of the SEIF derivation, but this time we explicitly employ the available conditional independence between the robot and passive features given the active features which allows us to drop $\mathbf{m}^-$ from the posterior over $\mathbf{x}_t$ as

$$p\big(\mathbf{x}_t, \mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) = p\big(\mathbf{x}_t \mid \mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big)p\big(\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) \quad (9a)$$
$$\overset{\text{C.I.}}{=} p\big(\mathbf{x}_t \mid \mathbf{m}^0, \mathbf{m}^+\big)p\big(\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) \quad (9b)$$
$$= \frac{p\big(\mathbf{x}_t, \mathbf{m}^0 \mid \mathbf{m}^+\big)}{p\big(\mathbf{m}^0 \mid \mathbf{m}^+\big)}p\big(\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) \quad (9c)$$

The posterior factorization shown above is exact where for convenience equation (9c) merely re-expresses the conditional over $\mathbf{x}_t$ in (9b) as a ratio. To obtain the sparsified posterior approximation, we now impose conditional independence between $\mathbf{x}_t$ and $\mathbf{m}^0$ as

$$\breve{p}_{\textsc{ModRule}}\big(\mathbf{x}_t, \mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) = \frac{p\big(\mathbf{x}_t \mid \mathbf{m}^+\big)p\big(\mathbf{m}^0 \mid \mathbf{m}^+\big)}{p\big(\mathbf{m}^0 \mid \mathbf{m}^+\big)}p\big(\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) \quad (10a)$$
$$= p\big(\mathbf{x}_t \mid \mathbf{m}^+\big)p\big(\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) \quad (10b)$$
$$= \frac{p_U\big(\mathbf{x}_t, \mathbf{m}^+\big)}{p_V\big(\mathbf{m}^+\big)}p_D\big(\mathbf{m}^0, \mathbf{m}^+, \mathbf{m}^-\big) \quad (10c)$$

where again for convenience equation (10c) simplifies the sparsified posterior to a ratio of marginals and the subscripts p_U, p_V, p_D are used to notationally reference the different pdfs involved. As equations (10a)–(10b) show, sparsification is equivalent to imposing conditional independence, which in turn is equivalent to dropping dependence on the set of features we wish to deactivate (i.e., $\mathbf{m}^0$). The resulting modified sparsification rule is summarized by equations (11)–(14) which express it in both covariance and information form.

Covariance Form

$$\breve{\Sigma}_t = \big(S_{x_t,m^+}\Sigma_U^{-1}S_{x_t,m^+}^\top - S_{m^+}\Sigma_V^{-1}S_{m^+}^\top + S_{m^0,m^+,m^-}\Sigma_D^{-1}S_{m^0,m^+,m^-}^\top\big)^{-1} \quad (11)$$

$$\breve{\boldsymbol{\mu}}_t = \boldsymbol{\mu}_t \quad (12)$$

where

$$\Sigma_U = S^\top_{x_t,m^+}\Sigma_t S_{x_t,m^+} \qquad \Sigma_V = S^\top_{m^+}\Sigma_t S_{m^+}$$
$$\Sigma_D = S^\top_{m^0,m^+,m^-}\Sigma_t S_{m^0,m^+,m^-}$$

Information Form

$$\breve{\Lambda}_t = S_{x_t,m^+}\Lambda_U S^\top_{x_t,m^+} - S_{m^+}\Lambda_V S^\top_{m^+} + S_{m^0,m^+,m^-}\Lambda_D S^\top_{m^0,m^+,m^-} \quad (13)$$

$$\breve{\boldsymbol{\eta}}_t = S_{x_t,m^+}\boldsymbol{\eta}_U - S_{m^+}\boldsymbol{\eta}_V + S_{m^0,m^+,m^-}\boldsymbol{\eta}_D \quad (14)$$

where

$$\Lambda_U = S^\top_{x_t,m^+}\Big(\mathrm{I} - \Lambda_t S_{m^0,m^-} \times \big(S^\top_{m^0,m^-}\Lambda_t S_{m^0,m^-}\big)^{-1} S^\top_{m^0,m^-}\Big)\Lambda_t S_{x_t,m^+}$$

$$\boldsymbol{\eta}_U = S^\top_{x_t,m^+}\Big(\mathrm{I} - \Lambda_t S_{m^0,m^-}\big(S^\top_{m^0,m^-}\Lambda_t S_{m^0,m^-}\big)^{-1} S^\top_{m^0,m^-}\Big)\boldsymbol{\eta}_t$$

$$\Lambda_V = S^\top_{m^+}\Big(\mathrm{I} - \Lambda_t S_{x_t,m^0,m^-} \times \big(S^\top_{x_t,m^0,m^-}\Lambda_t S_{x_t,m^0,m^-}\big)^{-1} S^\top_{x_t,m^0,m^-}\Big)\Lambda_t S_{m^+}$$

$$\boldsymbol{\eta}_V = S^\top_{m^+}\Big(\mathrm{I} - \Lambda_t S_{x_t,m^0,m^-} \times \big(S^\top_{x_t,m^0,m^-}\Lambda_t S_{x_t,m^0,m^-}\big)^{-1} S^\top_{x_t,m^0,m^-}\Big)\boldsymbol{\eta}_t$$

$$\Lambda_D = S^\top_{m^0,m^+,m^-}\Big(\mathrm{I} - \Lambda_t S_{x_t}\big(S^\top_{x_t}\Lambda_t S_{x_t}\big)^{-1} S^\top_{x_t}\Big)\Lambda_t S_{m^0,m^+,m^-}$$

$$\boldsymbol{\eta}_D = S^\top_{m^0,m^+,m^-}\Big(\mathrm{I} - \Lambda_t S_{x_t}\big(S^\top_{x_t}\Lambda_t S_{x_t}\big)^{-1} S^\top_{x_t}\Big)\boldsymbol{\eta}_t$$

In particular, note that equation (12) shows that the modified sparsification rule clearly maintains the mean estimate. Furthermore, as seen by careful inspection of the projection matrices involved in equation (13), it simultaneously deactivates the map features $\mathbf{m}^0$ (i.e., S_{x_t,m^+} only populates the robot/active feature sub-block of the resulting information matrix $\breve{\Lambda}_t$). However, a major drawback to the modified rule's correctness is that sparsification is no longer a scalable operation as evident by the expressions for Λ_U and Λ_V which require large matrix inversions over the passive features $\mathbf{m}^-$.

IV. Results

In this section we investigate the implications associated with the different sparsification rules by considering two different scenarios in which we compare the sparsified information filters to that of the standard Kalman Filter (KF) formulation. In the first scenario we consider a linear Gaussian (LG) SLAM simulation in which the KF is the optimal Bayes estimator and provides a benchmark measure against which to compare the different sparsification routines. Subsequently, in the second scenario we test the algorithms on a real-world indoor nonlinear dataset to understand their performance in practice.

A. Linear Simulation

In an effort to compare the effects of the two sparsification strategies in a controlled manner, we start by applying the three estimators to a synthetic dataset. The vehicle motion is purely translational, generated by a linear, constant-velocity model corrupted by additive white Gaussian noise. As the robot moves around in the environment, it measures the relative position of local point features, again perturbed by white noise. The desired sparsity is expressed by limiting the number of active features to ten percent of the total number of landmarks in the environment. As the simulation has been restricted to LG SLAM, we can compare the effects of the two sparsification routines relative to the optimal Kalman Filter.

To test the consistency between the different filter uncertainties and the true state estimation errors, we use the normalized estimation error squared (NEES) [11] computed based upon a series of Monte Carlo simulations and two different error metrics. The first metric relates the ground truth to the direct output of the filters and provides a measure of *global* error. The second metric computes the state estimate as expressed relative to the first feature that was observed $\mathbf{x}_m$ via the standard compounding operation: $\mathbf{x}_{mi} = \ominus\mathbf{x}_m \oplus \mathbf{x}_i$ and provides a measure of *local/relative* error. Fig. 4(a) and Fig. 4(c) compare the two error metrics for the vehicle position NEES score for the KF and information filters. Similarly, Fig. 4(b) and Fig. 4(d) show the normalized errors for a single map feature and are representative of the performance for other map elements. The horizontal threshold signifies the $97.5^{\%}$ upper bound for the chi-square test. Looking at the estimate of vehicle and map positions in the world frame, the modified-rule yields errors nearly identical to those of the KF not only in regards to magnitude, but also behavior over time. In comparison, the SEIF global errors are seen to be noticeably larger, though in contrast, the normalized relative errors are roughly equivalent to those of the KF and modified filter. This apparent discrepancy is indicative that the relative map estimates for all three filters have converged while the global SEIF estimate has an absolute state estimate which is inconsistent.

We can gain further insight into the consequences of sparsification by looking at the covariances associated with each filter. Fig. 5 provides a histogram comparing the ratio of determinants of the absolute and relative feature covariance matrices for the KF to those of the information filters. To aid in interpreting the ratio as a metric: values of one represent ideal, while those larger than one indicate the amount of over-confidence. For both map representations, we see that the two sparsified uncertainty measures are over-confident with respect to those of the standard KF and, in turn, are inconsistent with the true estimation error. However, the difference in magnitude between the confidence regions associated with the modified-rule and the standard KF are nearly negligible in both a global and local sense while the SEIF rule has absolute uncertainty which is significantly more over-confident[2]. Upon referencing the state estimates relative to the first observed feature, though, the SEIF covariance matrix reflects nearly the same estimate of uncertainty as the KF and the modified-rule. Note that in the process of root-shifting the map to the first feature, the original world origin is now included as a state element. While the

[2]The exception is with the first feature added to the map which is the source of the outliers shown in the plots in Fig. 5(a).

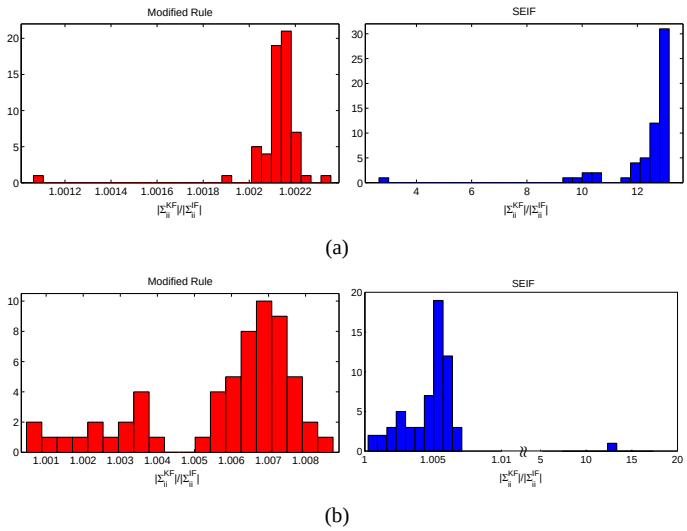


Fig. 5. The KF estimates of feature uncertainty as a fraction of the uncertainties obtained using the SEIF and modified-rule. In particular, we show histograms over the ratio of determinants for the map element covariance matrix sub-blocks. The uncertainties depicted in (a) are computed directly from the absolute covariances maintained by the filters. The histograms in (b) describe the *relative* map determinants for the covariance matrices corresponding to the state as expressed relative to the first observed feature. Ratios greater than one indicate that both sparsification methods result in estimates of feature position which are over-confident. However, the modified-rule produces uncertainty estimates which are comparable to those of the KF for both the absolute and relative maps. On the other hand, SEIF maintains absolute estimates which are significantly over-confident while the relative uncertainty estimates are approximately equivalent to those of the KF and modified-rule. In both plots in (a), the outlier nearest to one is that of the first mapped feature. The outlier shown in the SEIF histogram in (b) corresponds to the representation of the original world origin in the root-shifted reference frame and is a consequence of the over-confident absolute map.

world origin uncertainty estimate for the modified-rule agrees with those of the rest of the relative map estimates, the same is not true for the SEIF's uncertainty measure of the world-origin as indicated by the outlier in Fig. 5(b). This indicates that while the relative SEIF map estimate has converged, its estimate of the global world origin remains inconsistent.

The effect of sparsification on the covariance estimates is consistent with what is observed with the normalized errors. Though there is little difference between the three sets of feature position estimates, the errors for the absolute SEIF map are larger due to the higher confidence attributed to the estimates. In the case of root-shifting the state, the negligible difference that we see between a feature's auto-covariance sub-block for the relative map leads to nearly identical normalized errors.

B. Experimental Validation

Simulations are helpful in investigating our findings without having to take into consideration the effects of linearization. More often than not, though, real-world SLAM problems involve nonlinear vehicle motion and perception models and include noise which is not truly Gaussian. For that reason, we tested the estimation algorithms on a typical, nonlinear dataset.

For our experimental setup, depicted in Fig. 6, a wheeled robot was manually driven in an environment consisting of a set of 64 track hurdles positioned on four adjacent tennis courts which provide ground truth for the experiment. The vehicle made observations of the environment using a SICK laser range finder and was equipped with wheel encoders for determining the motion control inputs. The problem of correctly pairing measurement data with the corresponding hurdle was addressed offline and, thus, data association is identical for each SLAM filter. Under our feature-based representation, each hurdle serves as an individual coordinate frame, parameterized by a base leg position and its orientation.

Fig. 6. Photograph of experimental setup on MIT tennis court with hurdles. Ground truth is determined from the court baselines.

An Extended Kalman Filter (EKF) is applied concurrently with the two information filters, relying upon the sparsification routines to maintain a limit of ten active features. The resulting state estimates exhibit much the same behavior as we see in the LG SLAM case with the contrasting absolute and relative performance of the SEIF. This is perhaps best revealed by the SLAM maps generated by the three filters. In Fig. 7(a) we provide plots of the SLAM maps in terms of the global state representation. Enforcing sparsity with the modified-rule leads to estimates of both feature pose and uncertainty which exhibit negligible difference from the results of the EKF. As in the LG simulation, the SEIF yields global map estimates which are inconsistent as a majority of the true hurdle positions fall well outside the three-sigma uncertainty regions. Alternatively, root-shifting the map relative to the first feature instantiated into the map, as depicted in Fig. 7(b), reveals that both sparsified filters as well as the EKF maintain relative posteriors which are very similar.

V. Discussion

Both the simulated and nonlinear experimental datasets reveal that the modified sparsification rule yields error estimates which are nearly identical to those of the standard EKF in both a global and local sense, while the SEIF estimate is only comparable locally. Despite its advantages, close inspection reveals that the modified-rule estimates are still slightly over-confident with respect to the KF. This section seeks to explain the cause of this inconsistency.

Instructively, it can be shown that this inconsistency is a natural result of the fact that the sparsification routine *imposes*

conditional independence between the robot and the deactivated features $\mathbf{m}^0$. To illustrate this, consider a general three state distribution, $p(a,b,c) = p(a|b,c)p(b,c)$ and its sparsified approximation where any possible dependence between a and b is ignored

$$\tilde{p}(a,b,c) = p(a|c)p(b,c) \tag{15}$$

To understand the effect of this approximation on LG SLAM suppose the true distribution is given by

$$p(a,b,c) = \mathcal{N}\left(\begin{bmatrix}\mu_a\\ \mu_b\\ \mu_c\end{bmatrix}, \begin{bmatrix}\sigma_a^2 & \rho_{ab}\sigma_a\sigma_b & \rho_{ac}\sigma_a\sigma_c\\ \rho_{ab}\sigma_a\sigma_b & \sigma_b^2 & \rho_{bc}\sigma_b\sigma_c\\ \rho_{ac}\sigma_a\sigma_c & \rho_{bc}\sigma_b\sigma_c & \sigma_c^2\end{bmatrix}\right) \tag{16}$$

Applying the sparsification approximation of (15) we have

$$\tilde{p}(a,b,c) = \mathcal{N}\left(\begin{bmatrix}\mu_a\\ \mu_b\\ \mu_c\end{bmatrix}, \begin{bmatrix}\sigma_a^2 & \rho_{ac}\rho_{bc}\sigma_a\sigma_b & \rho_{ac}\sigma_a\sigma_c\\ \rho_{ac}\rho_{bc}\sigma_a\sigma_b & \sigma_b^2 & \rho_{bc}\sigma_b\sigma_c\\ \rho_{ac}\sigma_a\sigma_c & \rho_{bc}\sigma_b\sigma_c & \sigma_c^2\end{bmatrix}\right) \tag{17}$$

A necessary and sufficient condition for the approximation to be consistent is that the covariance matrices obey the inequality, $\tilde{\Sigma} - \Sigma \geq 0$ [14]. A sufficient condition test for positive semi-definiteness is that the determinant of all upper left submatrices be positive [15]. Applying this test we see that (17), does not, in general, satisfy the positive semi-definiteness condition because the determinant of the upper left 2×2 of $\tilde{\Sigma} - \Sigma$ (i.e., $\det\left[\begin{smallmatrix}0 & (\rho_{ac}\rho_{bc}-\rho_{ab})\sigma_a\sigma_b\\ (\rho_{ac}\rho_{bc}-\rho_{ab})\sigma_a\sigma_b & 0\end{smallmatrix}\right]$) is less than zero for $\rho_{ac}\rho_{bc} \neq \rho_{ab}$. Hence, extending this insight we see that imposing conditional independence between the robot and the deactivated features, $\mathbf{m}^0$, results in an approximation to the joint posterior which is inconsistent. Therefore, though the modified-rule estimates are comparable to the KF, this explains the cause of their slight over-confidence.

VI. Conclusion

In conclusion, recent novel insights into the canonical formulation of SLAM have revealed sparseness as a "natural" characteristic of the information parameterization and have lead to promising new research into scalable algorithms. Many of these new approaches are founded upon pruning relatively weak constraints in the information form to achieve exact sparsity. The delicate issue these methods must deal with then is "how to approximate the posterior with an exactly sparse representation in a consistent manner?"

In this paper, we have demonstrated that the method of enforcing sparsity employed by SEIFs leads to an inconsistent *absolute* map, while empirical testing indicates that the *relative* map relationships are preserved. We then showed that by exploiting the conditional independence between the robot and the passive features given the active map, that a new modified version of the SEIF sparsification rule can be derived. It was shown that this modified-rule yields a sparsified posterior comparable to that of the EKF in regards to the mean and uncertainty estimates for both the *absolute* and *relative* maps. Unfortunately, this accuracy comes at a cost as the modified rule requires the inversion of a matrix of the size of the number of passive features and, thus, is no longer constant-time.

Furthermore, despite the modified-rule's accuracy with respect to the KF, LG SLAM simulation results indicate that *both* sparsification routines lead to over-confident state estimates. We investigated the cause of this inconsistency in §V and concluded that this over-confidence is a direct result of the approximation by which the two filters achieve an exactly sparse representation of the posterior. It appears then that a computationally efficient *and* theoretically correct approximation for maintaining a sparse information matrix representation for SLAM remains an open research task.

References

[1] R. Smith, M. Self, and P. Cheeseman, *Estimating Uncertain Spatial Relationships in Robotics*, ser. Autonomous Robot Vehicles. Springer-Verlag, 1990.

[2] P. Moutarlier and R. Chatila, "An Experimental System for Incremental Environment Modeling by an Autonomous Mobile Robot," in *Proceedings of the 1st International Symposium on Experimental Robotics*, Montreal, Canada, June 1989.

[3] M. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, "A Solution to the Simultaneous Localization and Map Building (SLAM) Problem," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 3, pp. 229–241, June 2001.

[4] S. Thrun, Y. Liu, D. Koller, A. Ng, Z. Ghahramani, and H. Durrant-Whyte, "Simultaneous Localization and Mapping with Sparse Extended Information Filters," *International Journal of Robotics Research*, Accepted, To Appear.

[5] M. Paskin, "Thin Junction Tree Filters for Simultaneous Localization and Mapping," University of California, Berkeley, Technical Report CSD-02-1198, September 2002.

[6] U. Frese and G. Hirzinger, "Simultaneous Localization and Mapping - a Discussion," in *Proceedings of the IJCAI Workshop Reasoning with Uncertainty in Robotics*, Seattle, WA, 2001, pp. 17–26.

[7] U. Frese, "Treemap: An O(Log N) Algorithm for Simultaneous Localization and Mapping," in *Spatial Cognition IV*, C. Freksa, Ed. Springer Verlag, 2004.

[8] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Prentice Hall, 2003.

[9] U. Frese, "A Proof for the Approximate Sparsity of SLAM Information Matrices," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Barcelona, Spain, Accepted, To Appear.

[10] Y. Liu and S. Thrun, "Results for Outdoor-SLAM Using Sparse Extended Information Filters," in *IEEE International Conference on Robotics and Automation*, vol. 1, September 2003, pp. 1227–1233.

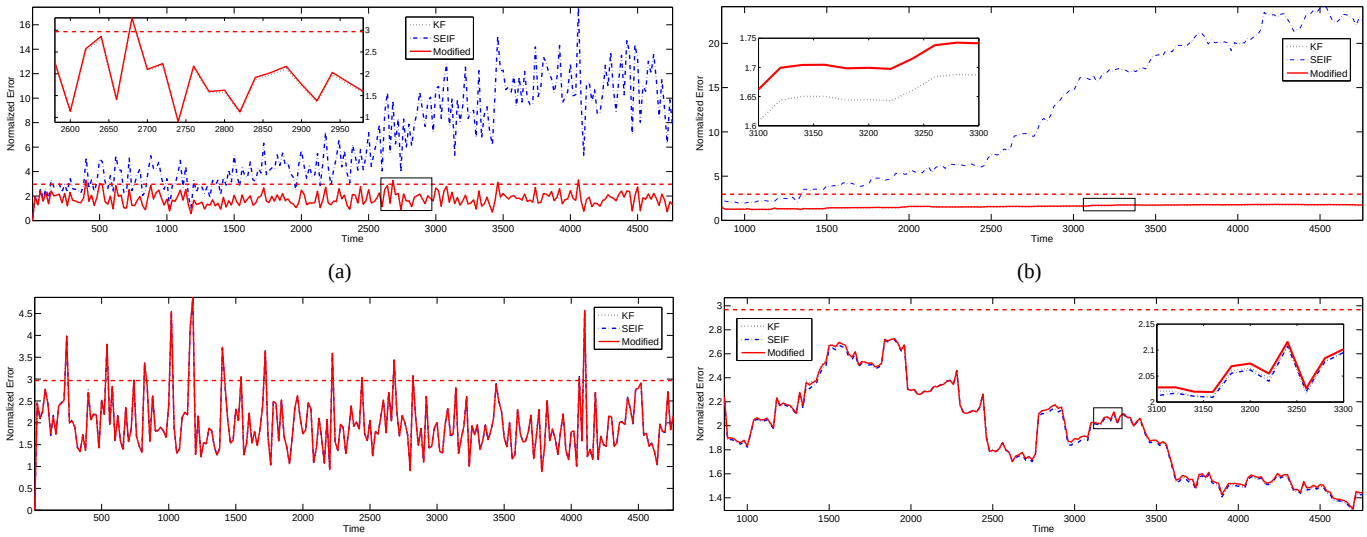
[11] Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York: John Wiley & Sons, Inc., 2001.

[12] R. Eustice, H. Singh, and J. Leonard, "Exactly Sparse Delayed State Filters," in *Proceedings of the International Conference on Robotics and Automation*, Barcelona, Spain, Aceepted, To Appear.

[13] M. Paskin, "Thin Junction Tree Filters for Simultaneous Localization and Mapping," in *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, San Francisco, CA, 2003, pp. 1157–1164.

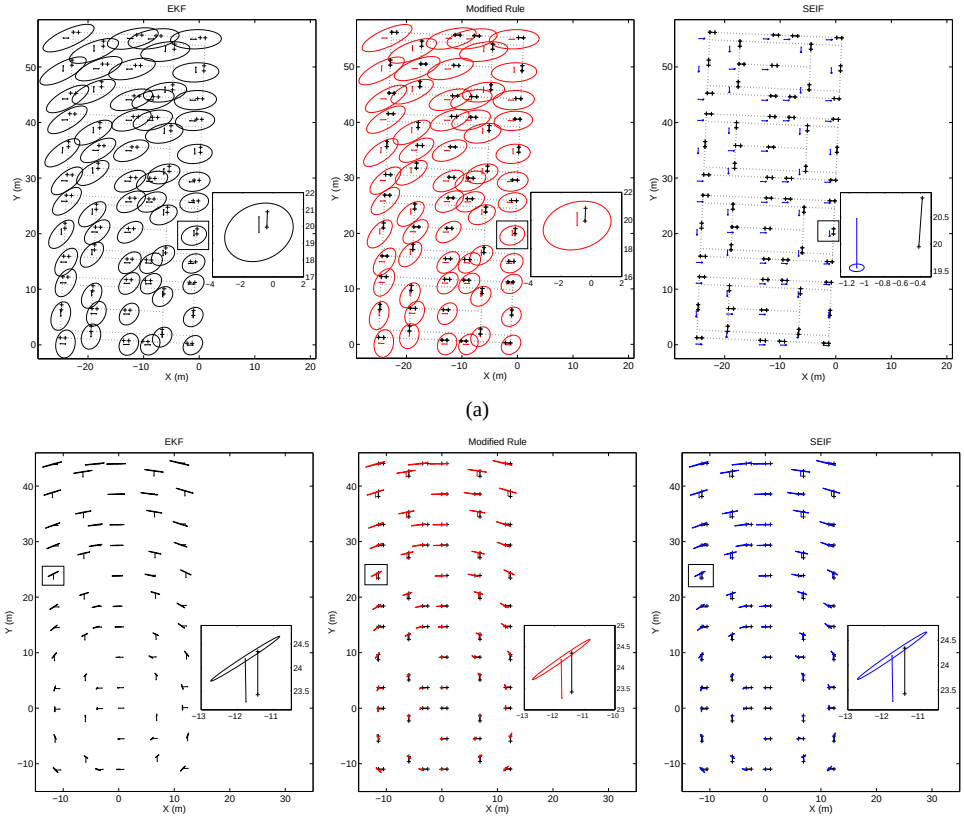
[14] Covariance Intersection Working Group, "A Culminating Advance in the Theory and Practice of Data Fusion, Filtering, and Decentralized Estimation, Technical Report," 1997.

[15] G. Strang, *Linear Algebra and Its Applications*, 2nd ed. New York: Academic Press, 1980.



(a) (b) (c) (d)

Fig. 4. The normalized estimation error squared for the vehicle, (a) and (c), and one of the features, (b) and (d), as estimated based upon 20 linear Gaussian Monte Carlo simulations. The horizontal line signifies the the 97.5% chi-square upper bound. The error shown in the top two plots, (a) and (b), corresponds to a direct comparison of the filter estimates to the ground truth, and represents a measure of *global* consistency. In the bottom two plots, (c) and (d), we plot the *local* normalized error computed relative to the first feature instantiated in the map: $\mathbf{x}_{mi} = \ominus\mathbf{x}_m \oplus \mathbf{x}_i$. The larger global normalized error associated with the SEIF is a result of an absolute state which is significantly over-confident. The relative map error, on the other hand, is nearly identical to that of the modified-rule and KF, empirically indicating that the SEIF yields locally consistent estimates.



(a)

(b)

Fig. 7. Comparison of the EKF and information filter SLAM maps with ground truth (cross-hairs) for the hurdles experimental dataset. The plots in (a) correspond to the absolute feature poses as directly estimated by the three SLAM algorithms together with the three-sigma confidence bounds. Shown in (b) are the relative maps and corresponding three-sigma uncertainty ellipses transformed relative to the first hurdle added to the map. As indicated by the right-most plot of (a), the SEIF maintains global feature estimates which are significantly over-confident. The modified-rule, meanwhile, yields estimates for absolute feature pose and uncertainty which are nearly identical to those of the EKF. Considering the relative map structure, the two sparsified filters perform similarly to the EKF.