Rotational Outlier Identification in Pose Graphs Using Dual Decomposition

Arman Karimian, Ziqi Yang, and Roberto Tron

Boston University, Boston MA 02215, USA
{armandok,zy259,tron}@bu.edu

Abstract In the last few years, there has been an increasing trend to consider Structure from Motion (SfM, in computer vision) and Simultaneous Localization and Mapping (SLAM, in robotics) problems from the point of view of *pose averaging* (also known as *global SfM*, in computer vision) or *Pose Graph Optimization* (PGO, in robotics), where the motion of the camera is reconstructed by considering only relative rigid body transformations instead of including also 3-D points (as done in a full Bundle Adjustment). At a high level, the advantage of this approach is that modern solvers can effectively avoid most of the problems of local minima, and that it is easier to reason about outlier poses (caused by feature mismatches and repetitive structures in the images). In this paper, we contribute to the state of the art of the latter, by proposing a method to detect incorrect orientation measurements prior to pose graph optimization by checking the geometric consistency of rotation measurements. The novel aspects of our method are the use of Expectation-Maximization to fine-tune the covariance of the noise in inlier measurements, and a new approximate graph inference procedure, of independent interest, that is specifically designed to take advantage of evidence on cycles with better performance than standard approaches (Belief Propagation). The paper includes simulation and experimental results that evaluate the performance of our outlier detection and cycle-based inference algorithms on synthetic and real-world data.

Keywords: Pose averaging; outliers; inference in graphical models.

1 Introduction

Reconstructing a 3-D scene from a collection of ordered or unordered images or videos is one of the most prominent classical problems in computer vision and robotics. In computer vision, this task is known as *Structure from Motion* (SfM), and is traditionally performed using images alone. The typical solution pipeline for this problem [26, 35, 39] includes three steps: *1)* estimate relative poses between pairs of images using matched features [7, 17, 31] and robust fitting techniques [21,24]; *2)* combine the pairwise estimates, either in sequential stages [3, 4, 22, 43, 44], or by combining poses alone (without considering a 3-D structure) in a *pose-averaging* [1, 5, 14, 25, 33, 50, 51] or *pose-graph* [11] approach; *3)* use

Bundle Adjustment (BA) [18, 26, 49], which minimizes the reprojection error by considering jointly the motion and the structure.

In robotics, a very similar task is known as *Simultaneous Localization and Mapping* (SLAM, [10]), and it usually includes the use of additional information such as wheel odometry, inertial measurements, or laser scans. Visual SLAM is a variant of the SLAM problem where only visual information obtained from a camera is used for the task [48]. Similarly to the case of SfM, the state of the art approach for SLAM is based on a pose graph formulation where nodes represent robot poses at different times, and edges represent relative pose measurements between pairs of nodes. One typical difference between typical SfM and SLAM applications is that, in the latter, the images are mostly ordered; hence, edges in the graph can be divided into two categories: *ego motion* edges which correspond to temporally close measurements; e.g., visual odometry measurements (for which temporal correlations can easily predict the presence of outliers), and *loop closure* edges which correspond to temporally distant measurements, e.g., when the same physical location is revisited at different times.

In both pose averaging for SfM, and PGO for SLAM, the absolute poses (nodes) in the graph are estimated from all the measurements (edges) via a Maximum Likelihood (ML) formulation [16, 39], which typically involves solving a nonlinear least squares error minimization problem, and is highly sensitive to initialized values and the unavoidable presence of outlier measurements. For the problem of initialization, the most effective solutions use relaxations based on eigenvector computations or semi-definite programming [5, 11, 33, 51]. More recent techniques can certify the global optimality of their ML estimates [9, 20, 27, 41].

For the problem of outliers, traditional approaches rely on local optimization from an initial guess, and either discount outliers using robust cost functions [2,30,38]), or attempt to directly identify them [12,23,29,46,47]. In the latter group, there exist techniques based on inference on graphical models [53]. Empirically, these methods work well when the outliers are only few, and embedded in a dense graph of otherwise valid measurements; their performance decreases in more challenging regimes, such as in the alignment of multiple maps in SLAM, where many of the associations (loop closures) can be erroneous, and, for instance, finding a good initial guess for the alignment is more challenging. Existing solutions for this problem are limited to either a single map (as the optimization based approach in [28]) or two maps (as the set maximization approach of [32]).

Paper contributions. We propose a probabilistic approach for outlier detection between any number of maps. Our algorithm checks for the geometric consistency of the rotation measurements in loops within the graph of poses, and decides if each edge is an inlier or outlier without relying on a trajectory estimate. We use a Gaussian additive noise model for rotation measurements and use the rotational error over cycles as evidence to infer the inlier/outlier probabilities. We use the Expectation-Maximization algorithm to fine-tune the parameters of the distribution of measurement errors and present simulation results. For the inference step required by our algorithm, we first apply Belief Propagation (BP), and highlight its shortcomings in this setting, next we present a novel inference

algorithm based on a novel cycle-based dual decomposition and the Alternating Direction Method of Multipliers (ADMM) which has local convergence guarantees. We evaluate the performance of our proposed solution using simulations and in the alignment of four real-world maps produced by a standard SLAM algorithm.
Paper outline. In the remainder of the paper, we first review a probabilistic graphical model for error propagation on the space of rotations, and errors on cycles of poses (Section 2). We then review Belief Propagation for performing inference on the graphical model, describe our ADMM-based alternative (Section 3); and show how this inference can be embedded in an Expectation-Maximization procedure to estimate the variances of the inliers and the outliers (Section 4). Finally, we present our simulations and experiments (Section 5).

2 Probabilistic Model

In this section, we describe the approximate additive Gaussian noise model on rotations used for modeling the errors on single edges and along graph cycles, as well as the graphical model used to relate the inlier versus outlier probabilities for each edge with the evidence provided by the geometric consistency of cycles.

2.1 Gaussian Noise Model and Uncertainty Propagation

We denote the graph of poses as $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{T})$ with vertices $\mathcal{V} = \{1, \dots, n\}$ representing absolute poses that need to be estimated, and edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ representing the existence of measured relative transformations $\tilde{\mathbf{T}}_{ij} \in \mathcal{T}$ between them, i.e., $\tilde{\mathbf{T}}_{ij} \approx \mathbf{T}_j \mathbf{T}_i^{-1}$. Each pose $\mathbf{T}_i$ is represented as a member of a matrix Lie group, i.e., a group whose elements and group operation are representable by square matrices, and that is also a smooth differentiable manifold.

In this paper, we limit our attention to SO(3), leaving the applications of our methods to other Lie groups (e.g., SE(d) or Sim(d)) as future work. As we will show, this choice already provides significant benefits in the detection of outliers.

We model errors over rotations through a Gaussian distribution in local exponential coordinates, i.e., the distribution is defined in the tangent space at the mean, and mapped to the Lie group via the exponential map. Formally:

$$\begin{aligned} \boldsymbol{\epsilon} &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \\ \tilde{\mathbf{R}} &= \exp(\hat{\boldsymbol{\epsilon}})\, \mathbf{R} \end{aligned} \tag{1}$$

where $\boldsymbol{\epsilon} \in \mathbb{R}^3$ is a zero-mean Gaussian random variable with covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{3\times 3}$, and $\hat{\boldsymbol{\epsilon}} \in \mathfrak{so}(3)$ is a skew-symmetric matrix given by the *hat operator*

$$\hat{\boldsymbol{\epsilon}} = \begin{bmatrix} 0 & -\epsilon_3 & \epsilon_2 \\ \epsilon_3 & 0 & -\epsilon_1 \\ -\epsilon_2 & \epsilon_1 & 0 \end{bmatrix}. \tag{2}$$

We assume that, for inlier measurements, the magnitude of the vector $\boldsymbol{\epsilon}$, which represents the magnitude of the noise, is small, thus justifying the following.

Lemma 1 ([6], 7.3). *The first order approximation of the uncertainty in the composition of two rotations* $\tilde{\mathbf{R}}_1 \sim \mathcal{N}_{\mathrm{SO}(3)}(\mathbf{R}_1, \boldsymbol{\Sigma}_1)$ *and* $\tilde{\mathbf{R}}_2 \sim \mathcal{N}_{\mathrm{SO}(3)}(\mathbf{R}_2, \boldsymbol{\Sigma}_2)$ *is:*

$$\tilde{\mathbf{R}}_2\tilde{\mathbf{R}}_1 \sim \mathcal{N}_{\mathrm{SO}(3)}(\mathbf{R}_2\mathbf{R}_1, \boldsymbol{\Sigma}_2 + \mathbf{R}_2\boldsymbol{\Sigma}_1\mathbf{R}_2^{\mathsf{T}}) \tag{3}$$

This approximation, which comes from the truncation of the BCH formula, is justified by our assumption that the inlier errors are relatively small. In addition, we make the following assumption about the noise covariance $\boldsymbol{\Sigma}$:

Assumption 1 *The rotation distributions are isotropic, i.e.* $\boldsymbol{\Sigma}_i = \sigma_i^2\mathbf{I}_3$*, where* $\mathbf{I}_3$ *is the identity matrix.*

Combining Assumption 1 with (3), the distribution of the composition of a subset $\mathcal{S} \subset \mathcal{V}$ of noisy rotations is given by:

$$\prod_{i\in\mathcal{S}} \tilde{\mathbf{R}}_i \sim \mathcal{N}_{\mathrm{SO}(3)}\Big(\textstyle\prod_{i\in\mathcal{S}} \mathbf{R}_i \,,\, (\sum_{i\in\mathcal{S}} \sigma_i^2)\mathbf{I}_3\Big). \tag{4}$$

If all the variances σ_i are equal, the resultant covariance matrix is given by $m\sigma^2\mathbf{I}_3$, where $m = |\mathcal{S}|$. Since the expected length of a zero-mean spherical Gaussian random variable $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \varsigma^2\mathbf{I}_d)$ is tightly bounded as $\frac{d}{\sqrt{d+1}}\varsigma \leq \mathbb{E}(\|\boldsymbol{\varepsilon}\|) \leq \sqrt{d}\varsigma$ [13, Definition 3.1], for small enough m and σ the expected value of noise is proportional to $\sqrt{m}$; this was experimentally validated in [19, Figure 3].

2.2 Inlier and Outlier Gaussian Mixture Model

We model the distribution for each measurement R_e along an edge $e \in \mathcal{E}$ with a Gaussian mixture model with two modes, one for inliers and the other for outliers. We use the Bernoulli indicator variable $x_e \in \{0, 1\}$ to denote e as an inlier ($x_e = 0$) or an outlier ($x_e = 1$), with respective (user-defined) prior probabilities $p(x_e = 0) = \pi_e$ and $p(x_e = 1) = 1 - \pi_e \doteq \bar{\pi}_e$. Building upon Assumption 1, we assume that every inlier edge has uncertainty $\sigma^2\mathbf{I}_3$ and every outlier edge has uncertainty $\bar{\sigma}^2\mathbf{I}_3$, where $\bar{\sigma} \gg \sigma$; note that a sufficiently large value of $\bar{\sigma}$ in practice leads to an approximation of the uniform distribution.

2.3 Graphical Model for Evidence over Cycles

A simple cycle is a closed chain of edges where each edge appears only once. Every simple cycle c in our pose graph corresponds to an ordered set of rotation measurements along the edges of the cycle, and the composition $\tilde{\mathbf{R}}_c$ of these rotations $\tilde{\mathbf{R}}_c = \prod_{e\in c} \tilde{\mathbf{R}}_e$ should, ideally, be close to the identity (i.e., by transforming a reference frame along a cycle return it to its initial pose). Defining z_c to be the geodesic distance of $\tilde{\mathbf{R}}_c$ from the identity, i.e.,

$$z_c = \frac{1}{\sqrt{2}}\|\log(\tilde{\mathbf{R}}_c)\|_F = \arccos\Big(\frac{\mathrm{tr}(\tilde{\mathbf{R}}_c) - 1}{2}\Big), \tag{5}$$

where $\|\cdot\|_F$ is the Frobenius norm, we can use (4) to obtain a probabilistic mode (distribution) of $\tilde{\mathbf{R}}_c$. Note that the variance of $\tilde{\mathbf{R}}_c$ mainly depends on the length and the number of outliers of the cycle.

Similarly to previous work that aims to use geometric relations in cycles in Structure from Motion [19,53], we model the relation between errors on edges and cycles by using a Bayesian network in which every edge $e \in \mathcal{E}$, and every cycle $c \in \mathcal{C}$ of the original pose graph is modeled by a node in the Bayesian network, and each edge e is connected to the cycles c to which it belongs (Figure 1b); the cycles serve as evidence for inferring the hidden inlier/outlier state random variables x_e for each edge $e \in \mathcal{E}$. The joint probability distribution given by this model for hidden states $\boldsymbol{x} \in \{0,1\}^{|\mathcal{E}|}$ and cycle-consistency errors $\boldsymbol{z} \in \mathbb{R}^{|\mathcal{C}|}$ is

$$p(\boldsymbol{x}, \boldsymbol{z}) = \prod_{e \in \mathcal{E}} p(x_e) \prod_{c \in \mathcal{C}} p(z_c \,|\, \boldsymbol{x}_c), \tag{6}$$

where $\mathcal{C}$ is a set of cycles in $\mathcal{G}$, and $p(x_e)$ is the prior probability of edge e, and $\boldsymbol{x}_c$ is the vector containing x_e values for every $e \in c$. Equation (6) can be graphically represented using a factor graph (Figure 1c).

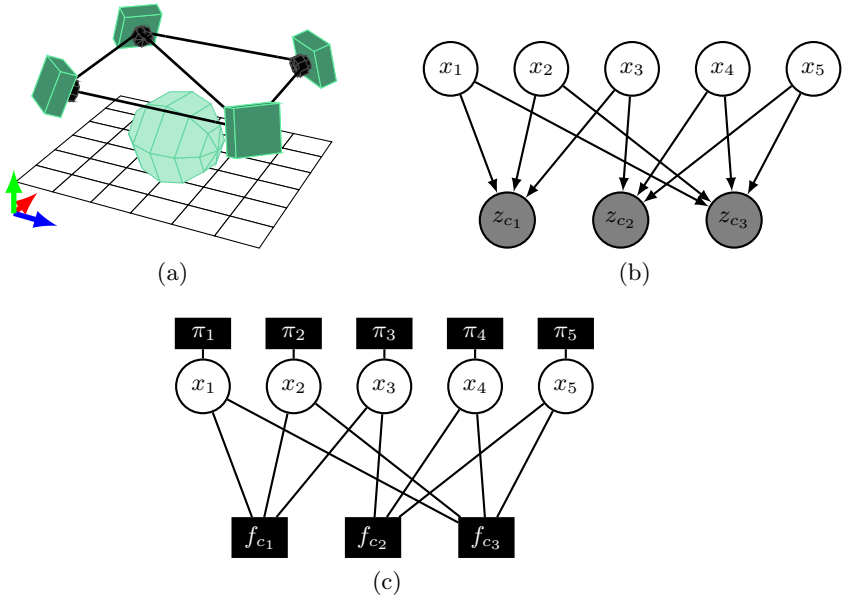


Figure 1: (a) Example of a small problem with four poses, five measurements, and three cycles; (b) A Bayesian network representation where the upper nodes correspond to the edges and the bottom nodes correspond to cycles (shaded in gray because they are observed variables); (c) The factor graph representation, with $\{\pi_e\}$ representing the edge prior probabilities, and $\{f_c\}$ the cycle evidence.

Letting $s_c = \mathbf{1}^\top \boldsymbol{x}_c$ be the number of outliers in c for the configuration $\boldsymbol{x}$, the distribution $p(z_c \,|\, \boldsymbol{x}_c)$ is obtained from (4), where the covariance matrix is given by $\varsigma_c^2(\boldsymbol{x}_c)\mathbf{I}_3 = \big(s_c\bar{\sigma}^2 + (|c| - s_c)\sigma^2\big)\mathbf{I}_3$ and $|c|$ is the length of the cycle.

Note that in robotics applications (see also our experiments), we can limit the inference of x_e to loop closure edges $\mathcal{E}_{lc}$; in modern systems, ego motion edges are unlikely to contain outlier measurements, hence we set the priors $\pi_e = 1$ for any ego motion edge $e \in \mathcal{E}$. Moreover, using all possible cycles is neither necessary nor practical for this task. The total number of cycles in a graph, in general, grows combinatorially with the size of the graph, leading to a proportional increase in the computational cost. To deal with this issue, we restrict ourselves to cycles from a *Minimum Cycle Basis* $\mathcal{C}_{min} \in 2^{\bar{\mathcal{C}}}$ of the pose graph, obtained using the de Pina's method [34]. This reduces the number of cycles to $\mathcal{O}(|\mathcal{E}_{lc}|)$, covers all the edges in bi-connected components of the pose-graph, and every other cycle can be obtained as a combination of cycles in the basis. Moreover, the MCB, which is minimal in the sense of the number of times each edge appears in cycles in $\mathcal{C}_{min}$, has the benefits of *1)* reducing the number of connections in the Bayesian graphical model (Figure 1b); and *2)* short cycles reduce the uncertainty in the observations z_c along cycles with only inliers (see the discussion in Section 2.1). In future work, we will explore the option of finding a basis that is minimal in the sense of the sum of the errors z_c.

3 Inference for Graphical Models

In this section, we assume that the set of parameters $\Theta = \{\sigma, \bar{\sigma}, \Pi\}$ where $\Pi = \{\pi_e\}_{e \in \mathcal{E}_{lc}}$ is given, and that we aim to find the marginal probabilities from (6) for each edge, i.e., $\gamma_e \triangleq p(x_e|\boldsymbol{z})$, $e \in \mathcal{E}$ (in Section 4, we will extend the procedure to concurrently estimate σ,$\bar{\sigma}$ from the same data). An exact solution to this probabilistic inference problem can easily become intractable, as the complexity increases exponentially with the number of edges. Resorting to approximation methods, we consider two options *1)* Loopy Belief Propagation (BP), which represents the standard traditional choice for approximate inference in graphs, although it is not guaranteed to converge for general graphs; and *2)* our novel inference algorithm based on dual decomposition along cycles with the Alternating Direction Method of Multipliers (ADMM), which instead has local convergence guarantees. Section 5 shows that, in our setting, the latter is superior in terms of outlier detection.

3.1 Belief Propagation

Belief Propagation is one of the most well known inference algorithms for finding marginal and conditional probabilities; it is a Variational Inference approach based on the minimization of the Bethe free energy [52]. For graphical models with loops, vanilla BP may fail to converge, and, even if it converges, the solution is generally not exact.

We review here the factor graph version of BP via the example of Figure 1. In BP, messages are sent between neighboring variables and factors according to the following equations [52]:

$$n_{e\to f_c}(x_e) = \prod_{f\in N(e)\setminus f_c} m_{f\to e}(x_e), \qquad m_{f_c\to e}(x_e) = \sum_{\boldsymbol{x}_c\setminus x_e} f_c(\boldsymbol{x}_c) \prod_{i\in N(f_c)\setminus e} n_{i\to f_c}(x_i), \tag{7}$$

where $n_{e\to f_c}$ is the message from variable e to factor f_c, and $m_{f_c\to e}$ is the message from factor f_c to variable e; $N(e)$ and $N(f_c)$ denote the factors that are connected to the random variable x_e and the factor f_c, respectively (the former includes the prior π_e, which is constant). These messages are passed in an asynchronous order until convergence of beliefs (approximate marginals), which are computed as:

$$b_e(x_e) \propto \prod_{f\in N(e)} m_{f\to e}(x_e), \qquad b_c(\boldsymbol{x}_c) \propto f_c(\boldsymbol{x}_c) \prod_{e\in N(f_c)} n_{e\to f_c}(x_e), \tag{8}$$

where b_e is the belief for the indicator variable of an edge and an approximation of $\gamma_e = p(x_e|\boldsymbol{z})$, and b_c is the belief of all random variables connected to factor f_c, and an approximation of $\gamma_c \triangleq p(\boldsymbol{x}_c|\boldsymbol{z})$ (note that γ_c is used in the Expectation-Maximization procedure in Section 5). To force convergence of the BP iterations, we introduce a damping factor as suggested in [40, Chapter 22] (we use 0.5 in our experiments).

3.2 Alternating Direction Method of Multipliers

The Alternating Direction Method of Multipliers (ADMM) provides a robust and decomposable algorithm for optimization problems by breaking them into smaller and easier to handle sub-problems [8]. For convex problems, ADMM guarantees global linear convergence rate [37]. It can also be used in non-convex problems, although in that case it will convergence to a local minimum.

In order to estimate γ_e and γ_c, instead of marginalizing $p(\boldsymbol{x}, \boldsymbol{z})$ over each edge directly, we propose to marginalize over the each cycle, i.e.,

$$p_c(\boldsymbol{x}_c, z_c) = p(z_c|\boldsymbol{x}_c) \prod_{e\in c} p(x_e), \tag{9}$$

and then force the marginals of each edge e from different overlapping cycles to agree on a common value. Intuitively, our approximation strategy aims to preserve the statistical correlation (joint distribution) between edges in the same cycle, while ignoring the correlations across cycles.

More in detail, we can implement this strategy by solving a *consensus* problem with ADMM [8, Chapter 7]. We denote as $\hat{\boldsymbol{v}}_c \in \mathbb{R}^{2^{|c|}}$ the vector containing all probabilities $p_c(\boldsymbol{x}_c|z_c)$ obtained from (9) evaluated over all possible values of $\boldsymbol{x}_c \in \{0,1\}^{|c|}$. For each cycle c, we try to estimate a vector $\boldsymbol{v}_c$ such that *1)* $\boldsymbol{v}_c$ is close to $\hat{\boldsymbol{v}}_c$, and *2)* when two distributions $\boldsymbol{v}_c$, $\boldsymbol{v}_{c'}$ for two overlapping cycles

c, c' are marginalized with respect to a common edge $e \in (c \cap c')$, the two results agree. We will parameterize the marginal distribution γ_e by keeping track of the inlier probability alone, denoted as $w_e = p(x_e = 0|\boldsymbol{z})$. We can then formulate the following minimization problem:

$$\begin{aligned} \min_{\boldsymbol{w},\{\boldsymbol{v}_c\}} \ & \sum_{c\in\mathcal{C}} h_c(\boldsymbol{v}_c), \\ \text{subject to } & \mathbf{p}_{e,c}^{\mathsf{T}}\boldsymbol{v}_c = w_e, \forall c \in \mathcal{C}, e \in c, \\ & 0 \le \boldsymbol{w} \le 1, \end{aligned} \tag{10}$$

where $\boldsymbol{w} \in \mathbb{R}^{|\mathcal{E}_{\iota_c}|}$ is the vector of all $\{w_e\}$, and the indicator vectors $\mathbf{p}_{e,c} \in \{0,1\}^{2^{|c|}}$ are a vectorial representation for obtaining the marginal inlier probability w_e given the cycle distribution $\boldsymbol{v}_c$. In (10), each h_c (i.e., each cycle), is considered a subproblem with its own local constraints that can be solved in a distributed fashion. As stated earlier, we want $\boldsymbol{v}_c$ to be close to $\hat{\boldsymbol{v}}_c$ with respect to some metric. If we choose the 2-Wasserstein metric, h_c will be formulated as follows:

$$h_c(\boldsymbol{v}_c) = \begin{cases} \|\boldsymbol{v}_c - \hat{\boldsymbol{v}}_c\|^2 \text{ if } \mathbf{1}^{\mathsf{T}}\boldsymbol{v}_c = 1, 0 \le \boldsymbol{v}_c \le 1, \\ +\infty \qquad\qquad \text{otherwise.} \end{cases} \tag{11}$$

(In future work, we plan to evaluate other measures of similarity between c and $\hat{c}$, such as the Kullback–Leibler divergence.) Subproblems c, c' that share an edge are forced to agree through the constraints $\mathbf{p}_{e,c}^{\mathsf{T}}\boldsymbol{v}_c = \mathbf{p}_{e,c'}^{\mathsf{T}}\boldsymbol{v}_{c'} = w_e$. This problem formulation is very similar to a consensus optimization problem, with the difference being that a linear combination of the variables $\boldsymbol{v}_c$ should reach consensus instead of the full vector, plus the global constraint $0 \le \boldsymbol{w} \le 1$. To apply ADMM, we write (10) using the indicator function $g(\boldsymbol{w})$ which returns $+\infty$ if the global constraint $0 \le \boldsymbol{w} \le 1$ is violated,

$$\begin{aligned} \min_{w_e} & \sum_{c\in\mathcal{C}} h_c(\boldsymbol{v}_c) + g(\boldsymbol{w}), \\ \text{subject to } & \mathbf{P}_c\boldsymbol{v}_c = \boldsymbol{w}_c, \ \forall c \in \mathcal{C}, \end{aligned} \tag{12}$$

where the vector $\boldsymbol{w}_c \in \mathbb{R}^{|c|}$ contains the elements w_e of $\boldsymbol{w}$ for every $e \in c$ and $\mathbf{P} \in \mathbb{R}^{|c|\times 2^{|c|}}$ is obtained by horizontally stacking the vectors $\mathbf{p}_{e,c}^{\mathsf{T}}$ column-wise. The augmented Lagrangian for (12) is:

$$L_\rho = \sum_{c\in\mathcal{C}} \Big(h_c(\boldsymbol{v}_c) + \boldsymbol{y}_c^{\mathsf{T}}(\mathbf{P}_c\boldsymbol{v}_c - \boldsymbol{w}_c) + \frac{\rho}{2}\|\mathbf{P}_c\boldsymbol{v}_c - \boldsymbol{w}_c\|^2 \Big) + g(\boldsymbol{w}), \tag{13}$$

with dual variables $\boldsymbol{y}_c \in \mathbb{R}^{|c|}$, and penalty parameter ρ. The ADMM iterations for this problem are given by [8]:

$$\begin{aligned} \boldsymbol{v}_c^{k+1} &= \operatorname*{argmin}_{\boldsymbol{v}_c} \Big(h_c(\boldsymbol{v}_c) + \boldsymbol{y}_c^{k\mathsf{T}}\mathbf{P}_c\boldsymbol{v}_c + \frac{\rho}{2}\|\mathbf{P}_c\boldsymbol{v}_c - \boldsymbol{w}_c^k\|^2 \Big) \\ \boldsymbol{w}^{k+1} &= \operatorname*{argmin}_{\boldsymbol{w}} \Big(g(\boldsymbol{w}) + \sum_{c\in\mathcal{C}} \big(-\boldsymbol{y}_c^{k\mathsf{T}}\boldsymbol{w}_c + \frac{\rho}{2}\|\mathbf{P}_c\boldsymbol{v}_c - \boldsymbol{w}_c\|^2 \big) \Big) \\ \boldsymbol{y}_c^{k+1} &= \boldsymbol{y}_c^k + \rho(\mathbf{P}_c\boldsymbol{v}_c - \boldsymbol{w}_c^k). \end{aligned} \tag{14}$$

Note that the variables $\boldsymbol{v}_c$ and $\boldsymbol{y}_c$ can be updated in parallel for each cycle. The solution for $\boldsymbol{v}_c$ is obtained by solving a quadratic programming problem (which can be done efficiently), while the solution for the global consensus variable $\boldsymbol{w}$ is given by:

$$w_e^{k+1} = \max(0, \min(1, \omega_e^{k+1})), \tag{15}$$

$$\omega_e^{k+1} = \frac{\sum_{c;e\in c}\left(\mathbf{p}_{e,c}^{\mathsf{T}}\boldsymbol{v}_c^{k+1} + \frac{1}{\rho}[\boldsymbol{y}_c^k]_e\right)}{\sum_{c;e\in c} 1}. \tag{16}$$

The denominator in (16) is the number of times edge e appears in different cycles, and therefore ω_e^{k+1} is the average of marginalized values for edge e plus the component of $\boldsymbol{y}_c^k$ that corresponds to e over cycles that contain that edge. In (15), the values of ω^{k+1} are projected to be between zero and one.

This problem will reach (local) optimality when the primal residual r^k and dual residuals t^k converge to zero, where:

$$\begin{aligned} r^k &= \sum_{c\in\mathcal{C}} \left\|\mathbf{P}_c\boldsymbol{v}_c^k - \boldsymbol{w}_c^k\right\|^2 \\ t^k &= \rho^2 \sum_{e\in\mathcal{E}_{lc}} \sum_{c;e\in c} (w_e^k - w_e^{k-1})^2 \end{aligned} \tag{17}$$

The penalty parameter ρ plays a very important role in the convergence speed of this method. Intuitively, small ρ allows intermediate solutions to have a much lower cost while somewhat ignoring the primal feasibility, and makes the solution less impacted by initial value and easier to escape from the local minima, whereas a large ρ will place a large penalty on violating the consensus constraints, but tends to produce small primal residuals. As suggested in [8, Chapter 3], we start with a small ρ, and gradually change the value of ρ based on primal and dual residual, using the following dynamic update rule:

$$\rho^{k+1} = \begin{cases} \tau^{incr}\rho^k & \text{if } r^k \le \mu t^k \\ \rho^k/\tau^{decr} & \text{if } t^k \le \mu r^k \\ \rho^k & \text{otherwise,} \end{cases} \tag{18}$$

where $\mu > 1$, $\tau^{decr} > 1$, and $\tau^{decr} > 1$ are constant parameters.

A disadvantage of our method is that the local variables $\{\boldsymbol{v}_c\}$ have dimensions that grow exponentially with the length of the cycles; however, in our experiments we noted that cycles of length up to $|c| = 15$ remain tractable, and longer cycles could be discarded, since they are likely to provide only very weak evidence.

4 Expectation-Maximization

In the previous sections, we assumed that the parameters $\Theta = \{\sigma, \bar{\sigma}, \Pi\}$ (the inlier and outlier standard deviations, and the edge prior probabilities, respectively)

were given. However, this assumption is not true and these parameters need to be estimated. By including parameters in the distribution, we rewrite (6) as:

$$p(\boldsymbol{x}, \boldsymbol{z}|\Theta) = \prod_{e\in\mathcal{E}} p(x_e|\pi_e) \prod_{c\in\mathcal{C}} p(z_c \,|\, \varsigma_c(\boldsymbol{x}_c)) \tag{19}$$

where the first term is a given by Bernoulli distribution. With some abuse of notation, we assume π_e is $p(x_e = 0)$ and $\bar{\pi}_e = 1 - \pi_e$ which yields $p(x_e|\pi_e) = \pi_e^{1-x_e}\bar{\pi}_e^{x_e}$. The second term is a wrapped Gaussian mixture distribution:

$$p(z_c \,|\, \varsigma_c(\boldsymbol{x}_c)) = \frac{1}{\psi_c} \frac{\varsigma_c^{-3}}{\phi(\varsigma_c)} \exp(\frac{-z_c^2}{2\varsigma_c^2}) \tag{20}$$

with $\varsigma_c(\boldsymbol{x}) = \sqrt{(\mathbf{1}^\mathsf{T}\boldsymbol{x}_c)\bar{\sigma}^2 + (|c| - \mathbf{1}^\mathsf{T}\boldsymbol{x}_c)\sigma^2}$, $\phi(\varsigma_c)$ is a normalizing constant for the wrapped Gaussian, ψ_c normalizes over all the possible values for $\boldsymbol{x}$,

$$\psi_c = \sum_{s=0}^{|c|} \binom{|c|}{s} \frac{\varsigma_c^{-3}(s)}{\phi(\varsigma_c(s))} \exp(\frac{-z_c^2}{2\varsigma_c^2(s)}), \tag{21}$$

and the term ς_c^{-3} comes from the denominator of the Gaussian probability density function, $\sqrt{\det(\varsigma_c^2\mathbf{I}_3)}$. The value of $\varsigma_c(\boldsymbol{x}_c)$ only depends on the number of outliers $s_c = \mathbf{1}^\mathsf{T}\boldsymbol{x}_c$, hence we can denote it is $\varsigma_c(s_c)$. The log-likelihood function is:

$$\begin{aligned}
&\mathcal{L}(\Theta; \boldsymbol{x}, \boldsymbol{z}) = \log(p(\boldsymbol{x}, \boldsymbol{z}|\Theta)) \\
&= \sum_{e\in\mathcal{E}} (1 - x_e)\log(\pi_e) + x_e \log(\bar{\pi}_e) + \sum_{c\in\mathcal{C}} -3\log\varsigma_c - \frac{z_c^2}{2\varsigma_c^2} - \log\left(\psi_c\phi(\varsigma_c)\right)
\end{aligned} \tag{22}$$

In the Expectation step, we find the expectation of the log likelihood of Θ^i with respect to the current distribution of $\boldsymbol{x}$ given $\boldsymbol{z}$ and previous estimate of parameters Θ^{i-1}:

$$Q(\Theta^{(i)}|\Theta^{(i-1)}) = \mathbb{E}_{\boldsymbol{x}|\boldsymbol{z},\Theta^{(i-1)}}[\mathcal{L}] = \sum_{\boldsymbol{x}\in\mathbb{Z}_2^{|\mathcal{E}|}} \mathcal{L}(\Theta^{(i)}; \boldsymbol{x}, \boldsymbol{z}) p(\boldsymbol{x}|\boldsymbol{z}, \Theta^{(i-1)}) \tag{23}$$

We use $\gamma_e^i = p(x_e|\boldsymbol{z}, \Theta^i)$ for the marginal of edge e, and $\gamma_c^i = p(\boldsymbol{x}_c|\boldsymbol{z}, \Theta^i)$ for the marginal of cycle c, given the parameters Θ^i (these are approximated via either BP or ADMM). Now, by expanding (23) we get:

$$Q(\Theta^{(i)}|\Theta^{(i-1)}) = \sum_{e\in\mathcal{E}} \sum_{x_e\in\mathbb{Z}_2} p(x_e|\boldsymbol{z}, \Theta^{(i-1)}) \log p(x_e|\pi_e^{(i)}) \tag{24a}$$

$$+ \sum_{c\in\mathcal{C}} \sum_{\boldsymbol{x}_c\in\mathbb{Z}_2^{|c|}} p(\boldsymbol{x}_c|\boldsymbol{z}, \Theta^{(i-1)}) \log p(z_c|\boldsymbol{x}_c, \sigma^{(i)}, \bar{\sigma}^{(i)}) \tag{24b}$$

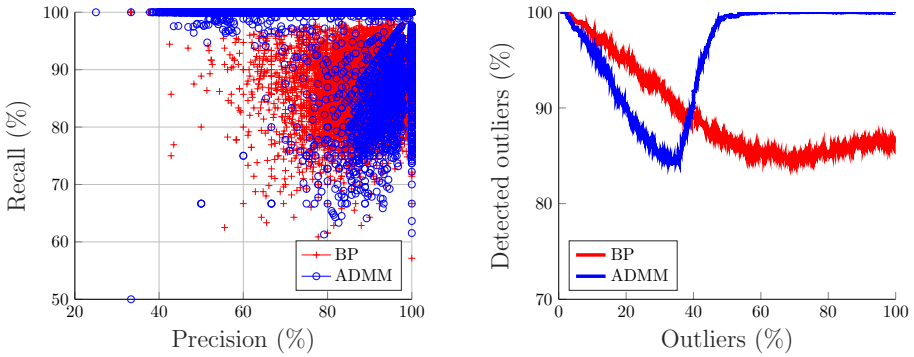
In the Maximization step, we find $\Theta^{(i)} = \text{argmax}_\Theta\, Q(\Theta|\Theta^{(i-1)})$. For $\Pi^{(i)}$, we have $\pi_e^{(i)} = \gamma_e^{(i-1)}$, but for $\sigma^{(i)}$ and $\bar{\sigma}^{(i)}$ it is not as straightforward. Each term in the summation in (24b) is a quasiconcave function, but their sum need not be quasiconcave. Therefore, we use a grid-search to find σ and $\bar{\sigma}$ at each iteration.

5 Simulations and Experiments on Map Merging

In this section, we provide performance results of our outlier detection algorithm over synthetic and real data. For the synthetic data, we repeatedly generate a pose graph with two maps of 15 nodes each and random poses. At every iteration, m edges are added between the two maps, where m varies from 10 to 200 with increments of 5. For every given m, from 1 to $m-1$ edges are selected to be outliers (with unitary increments). Inlier and outlier edges are given a random noise rotation with a random direction, and the magnitude of noise uniformly selected within $2.4^\circ \leq \|\boldsymbol{\epsilon}\| \leq 3.6^\circ$ for inliers and $72^\circ \leq \|\bar{\boldsymbol{\epsilon}}\| \leq 108^\circ$ for outliers (although we obtained similar results with different outlier distributions). The total number of generated graphs is $8,112$ and both BP and ADMM inference algorithms were used on the same graphs.

In Fig. 2a, we plot each simulation as a point on the precision-recall plane. Fig. 2b, the ratio of detected outliers is plotted versus the ratio of the outlier edges to total loop closure edges. It is clear that our ADMM inference performs better than BP, as it has overall higher precision and recall. In addition, as the ratio of outliers to loop closure edges increase, the performance of BP continuously deteriorates, while ADMM presents a V-shaped curve; we hypothesize that this is due to the fact that situations with a majority of inliners or outliers represent easier cases (there is little discrepancy between the results of the local inferences over the different cycles), while mixed situations are more difficult to reconcile.

In Fig. 3 we present an SfM experiment on the Castle-P30 dataset [45]. We obtain 83 pairwise relative rotation measurements by estimating the essential matrix (shown in red), and further add 30 random pairwise relative measurements



(a) (b)

Figure 2: (a) Precision ($\frac{TP}{TP+FP}$)-Recall ($\frac{TP}{TP+FN}$) for each simulated case using a threshold for γ_e of 0.5; points toward the top and right boundaries are better. (b) The ratio of detected outliers to the total number of outliers (Recall) versus the ratio of outlier loop closure edges to total loop closure edges; higher is better.

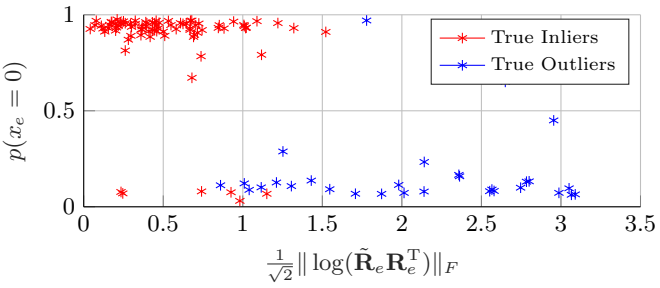


Figure 3: Error in relative pairwise orientation measurement $\tilde{\mathbf{R}}_e$ w.r.t ground-truth $\mathbf{R}_e$ ($x-$axis) versus the probability of detection as inlier ($y-$axis) by our EM-ADMM algorithm.

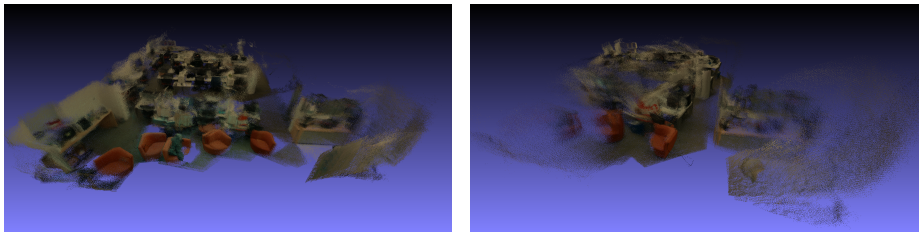
by sampling a random vector in SO(3) with a magnitude uniformly sampled between 0 and π (shown in blue). The given results indicate that most of the edges with small noise in their measurements are classified as inliers and those with high noise are classified as outliers.

In Fig. 4 we present the result of implementing our classifier on actual data obtained from an office environment and compare its performance with the method in [53]. Four independent sequences of RGB-D images were obtained using an Intel RealSense D435 camera, and were processed with ORB-SLAM2 [36]. The result of merging maps is shown with and without removing outliers, after initial alignment and pose graph optimization using GTSAM [15].

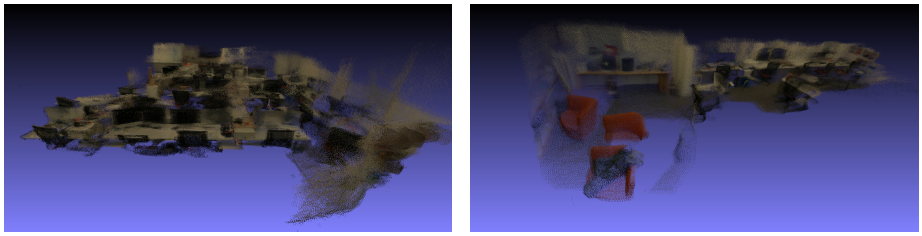
Data associations between the maps are obtained using the ORB-SLAM2's place recognition module in addition to an object detector (MobileNet-SSD [42]). In Table 1a the number of image pairs (RGB and Depth) for each map and the total number of images is given. A subset of these images is picked as Keyframes and constitute nodes in the joint pose graph. In Table 1b, the number of loop closure edges between the maps is given. There are no loop closure edges within each map. In Table 1c, the length of the cycles and frequency of cycles of those lengths are given from $\mathcal{C}_{min}$ which is used by our algorithm and also by the algorithm from [53]. We removed cycles from $\mathcal{C}_{min}$ with length greater than 15 due to increased complexity in the subproblem 11 and reduced the quality of evidence gathered from long cycles. As stated before, the length of the cycle is the number of loop-closure edges (total of 247) that appear in a bigger cycle which includes ego motion edges of robots' pose graphs. The outlier detection algorithm from [53] detects 13 outliers, whereas our algorithm finds 21 outliers.

6 Conclusion

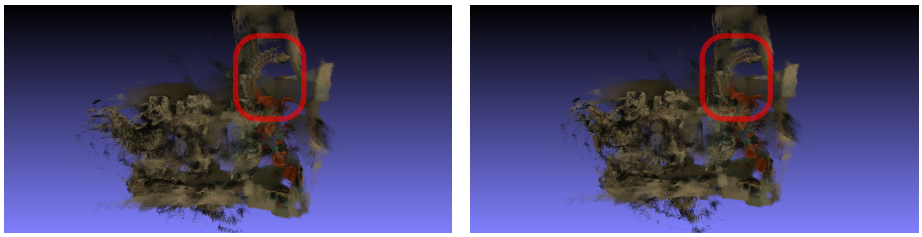
In this paper, we presented a probabilistic outlier detection algorithm which detects outliers based on the geometric consistency of rotation measurements over the cycles of a pose graph. We introduced a novel discrete inference algorithm with convergence guarantees that performed better than Belief Propagation.



(a) Map 1 (b) Map 2



(c) Map 3 (d) Map 4



(e) Outliers present (f) Outliers removed by [53]

(g) Outliers removed by our method

Figure 4: In (a)-(d), point clouds from four different sequences of images are depicted. The final point cloud made from joining all the four pointclouds without outlier detection is shown in (e) and with outlier detection is shown in (f)(from [53]) and (g) (our method). In (e) and (f), phantoms can be observed (e.g., see marked area) and the overall shape of the environment is not correct due to misalignment caused by outliers, whereas with out method in (g) these issues are considerably less severe.

	Map 1	Map 2	Map 3	Map 4	Total
Nodes	637	471	447	220	1,775
Images	3,233	2,289	2,641	926	9,089

(a) Images and nodes in each map and their total numbers from our dataset.

Map-pair Indices	1 - 2	1 - 3	2 - 3	1 - 4	3 - 4
Freq. of edges	50	8	70	70	49

(b) Frequency of loop-closure edges between map pairs.

Cycle length	3	4	5	6	7	12	15
Cycle freq.	80	102	45	8	2	1	1

(c) Length of the cycles in our dataset versus their frequency.

Table 1: Statistics on the dataset used for the experiments.

References

1. Aftab, K., Hartley, R., Trumpf, J.: Generalized weiszfeld algorithms for lq optimization. IEEE Transactions on Pattern Analysis and Machine Intelligence **37**(4), 728–745 (2015)
2. Agarwal, P., Tipaldi, G.D., Spinello, L., Stachniss, C., Burgard, W.: Robust map optimization using dynamic covariance scaling. In: 2013 IEEE International Conference on Robotics and Automation. pp. 62–69. Citeseer (2013)
3. Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S.M., Szeliski, R.: Building Rome in a day. Communications of the ACM **54**(10), 105–112 (2011)
4. Agarwal, S., Snavely, N., Seitz, S.M., Szeliski, R.: Bundle adjustment in the large. In: IEEE European Conference on Computer Vision, pp. 29–42. Springer (2010)
5. Arie-Nachimson, M., Kovalsky, S., Kemelmacher-Shlizerman, I., Singer, A., Basri, R.: Global motion estimation from point matches. In: International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission. pp. 81–88 (2012)
6. Barfoot, T.D.: State Estimation for Robotics. Cambridge University Press, USA, 1st edn. (2017)
7. Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-up robust features (SURF). Computer Vision and Image Understanding **110**(3), 346–359 (2008)
8. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends® in Machine learning **3**(1), 1–122 (2011)
9. Briales, J., Gonzalez-Jimenez, J.: Cartan-sync: Fast and global se (d)-synchronization. IEEE Robotics and Automation Letters **2**(4), 2127–2134 (2017)
10. Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., Leonard, J.J.: Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. IEEE Transactions on robotics **32**(6), 1309–1332 (2016)
11. Carlone, L., Tron, R., Daniilidis, K., Dellaert, F.: Initialization techniques for 3D SLAM: a survey on rotation estimation and its use in pose graph optimization. In: IEEE International Conference on Robotics and Automation (2015)

12. Carlone, L., Censi, A., Dellaert, F.: Selecting good measurements via ? 1 relaxation: A convex approach for robust estimation over graphs. In: Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on. pp. 2667–2674. IEEE (2014)
13. Chandrasekaran, V., Recht, B., Parrilo, P.A., Willsky, A.S.: The convex geometry of linear inverse problems. Foundations of Computational mathematics **12**(6), 805–849 (2012)
14. Chatterjee, A., Govindu, V.M.: Efficient and robust large-scale rotation averaging. In: IEEE International Conference on Computer Vision. pp. 521–528 (2013)
15. Dellaert, F.: Factor graphs and gtsam: A hands-on introduction. Tech. rep., Georgia Institute of Technology (2012)
16. Dellaert, F., Kaess, M.: Square root sam: Simultaneous localization and mapping via square root information smoothing. The International Journal of Robotics Research **25**(12), 1181–1203 (2006)
17. Dong, J., Soatto, S.: Domain-size pooling in local descriptors: DSP-SIFT. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 5097–5106 (2015)
18. Engels, E.C., Stewénius, H., Nistér, D.: Bundle adjustment rules. In: Photogrammetric Computer Vision. vol. 2, pp. 124–131 (2006)
19. Enqvist, O., Kahl, F., Olsson, C.: Non-sequential structure from motion. In: Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on. pp. 264–271. IEEE (2011)
20. Eriksson, A., Olsson, C., Kahl, F., Chin, T.J.: Rotation averaging and strong duality. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 127–135 (2018)
21. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM **24**(6), 381–395 (1981)
22. Frahm, J.M., Fite-Georgel, P., Gallup, D., Johnson, T., Raguram, R., Wu, C., Jen, Y.H., Dunn, E., Clipp, B., Lazebnik, S., Pollefeys, M.: Building Rome on a cloudless day. In: IEEE European Conference on Computer Vision, pp. 368–381. Springer (2010)
23. Graham, M.C., How, J.P., Gustafson, D.E.: Robust incremental slam with consistency-checking. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 117–124. IEEE (2015)
24. Hartley, R., Li, H.: An efficient hidden variable approach to minimal-case camera motion estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence (2012)
25. Hartley, R., Trumpf, J., Dai, Y., Li, H.: Rotation averaging. International Journal of Computer Vision **103**(3), 267–305 (2013)
26. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, second edn. (2004)
27. Kasten, Y., Geifman, A., Galun, M., Basri, R.: Algebraic characterization of essential matrices and their averaging in multiview settings. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 5895–5903 (2019)
28. Lajoie, P.Y., Hu, S., Beltrame, G., Carlone, L.: Modeling perceptual aliasing in slam via discrete–continuous graphical models. IEEE Robotics and Automation Letters **4**(2), 1232–1239 (2019)
29. Latif, Y., Cadena, C., Neira, J.: Robust loop closing over time for pose graph slam. The International Journal of Robotics Research **32**(14), 1611–1626 (2013)

30. Lee, G.H., Fraundorfer, F., Pollefeys, M.: Robust pose-graph loop-closures with expectation-maximization. In: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 556–563. IEEE (2013)
31. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60**(2), 91–110 (2004)
32. Mangelson, J.G., Dominic, D., Eustice, R.M., Vasudevan, R.: Pairwise consistent measurement set maximization for robust multi-robot map merging. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). pp. 2916–2923. IEEE (2018)
33. Martinec, D., Pajdla, T.: Robust rotation and translation estimation in multiview reconstruction. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–8 (2007)
34. Mehlhorn, K., Michail, D.: Implementing minimum cycle basis algorithms. Journal of Experimental Algorithmics (JEA) **11**, 2–5 (2007)
35. Moulon, P., Monasse, P., Marlet, R.: Global fusion of relative motions for robust, accurate and scalable structure from motion. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 3248–3255 (2013)
36. Mur-Artal, R., Tardós, J.D.: Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. IEEE Transactions on Robotics **33**(5), 1255–1262 (2017)
37. Nishihara, R., Lessard, L., Recht, B., Packard, A., Jordan, M.I.: A general analysis of the convergence of admm. arXiv preprint arXiv:1502.02009 (2015)
38. Olson, E., Agarwal, P.: Inference on networks of mixtures for robust robot mapping. The International Journal of Robotics Research **32**(7), 826–840 (2013)
39. Olson, E., Leonard, J., Teller, S.: Fast iterative alignment of pose graphs with poor initial estimates. In: Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on. pp. 2262–2269. IEEE (2006)
40. Robert, C.: Machine learning, a probabilistic perspective (2014)
41. Rosen, D.M., Carlone, L., Bandeira, A.S., Leonard, J.J.: Se-sync: A certifiably correct algorithm for synchronization over the special euclidean group. The International Journal of Robotics Research **38**(2-3), 95–125 (2019)
42. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4510–4520 (2018)
43. Snavely, N., Seitz, S.M., Szeliski, R.: Photo tourism: exploring photo collections in 3D. In: ACM Transactions on Graphics. vol. 25, pp. 835–846 (2006)
44. Snavely, N., Seitz, S.M., Szeliski, R.: Skeletal graphs for efficient structure from motion. In: IEEE Conference on Computer Vision and Pattern Recognition. vol. 1, p. 2 (2008)
45. Strecha, C., Von Hansen, W., Van Gool, L., Fua, P., Thoennessen, U.: On benchmarking camera calibration and multi-view stereo for high resolution imagery. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–8. Ieee (2008)
46. Sünderhauf, N., Protzel, P.: Switchable constraints for robust pose graph slam. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 1879–1884. IEEE (2012)
47. Sünderhauf, N., Protzel, P.: Towards a robust back-end for pose graph slam. In: 2012 IEEE International Conference on Robotics and Automation. pp. 1254–1261. IEEE (2012)
48. Taketomi, T., Uchiyama, H., Ikeda, S.: Visual slam algorithms: A survey from 2010 to 2016. IPSJ Transactions on Computer Vision and Applications **9**(1), 16 (2017)

49. Triggs, B., McLauchlan, P.F., Hartley, R.I., Fitzgibbon, A.W.: Bundle adjustment – a modern synthesis. In: Vision algorithms: theory and practice, pp. 298–372. Springer (2000)
50. Tron, R., Vidal, R.: Distributed 3-d localization of camera sensor networks from 2-d image measurements. IEEE Transactions on Automatic Control **59**(12), 3325–3340 (2014)
51. Wang, L., Singer, A.: Exact and stable recovery of rotations for robust synchronization. Information and Inference **2**(2), 145–193 (2013)
52. Yedidia, J.S., Freeman, W.T., Weiss, Y.: Constructing free-energy approximations and generalized belief propagation algorithms. IEEE Transactions on information theory **51**(7), 2282–2312 (2005)
53. Zach, C., Klopschitz, M., Pollefeys, M.: Disambiguating visual relations using loop constraints. In: Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. pp. 1426–1433. IEEE (2010)