**IEEE Robotics and Automation Letters (RAL) paper presented at the
2019 International Conference on Robotics and Automation (ICRA)
Palais des congres de Montreal, Montreal, Canada, May 20-24, 2019**

Modeling Perceptual Aliasing in SLAM
via Discrete-Continuous Graphical Models

Pierre-Yves Lajoie[1], Siyi Hu[2], Giovanni Beltrame[1], Luca Carlone[2]

Abstract— Perceptual aliasing is one of the main causes of
failure for Simultaneous Localization and Mapping (SLAM)
systems operating in the wild. Perceptual aliasing is the
phenomenon where different places generate a similar visual
(or, in general, perceptual) footprint. This causes spurious
measurements to be fed to the SLAM estimator, which typically
results in incorrect localization and mapping results. The
problem is exacerbated by the fact that those outliers are *highly
correlated*, in the sense that perceptual aliasing creates a large
number of mutually-consistent outliers. Another issue stems
from the fact that most state-of-the-art techniques rely on a
given trajectory guess (e.g., from odometry) to discern between
inliers and outliers and this makes the resulting pipeline brittle,
since the accumulation of error may result in incorrect choices
and recovery from failures is far from trivial. This work
provides a unified framework to *model* perceptual aliasing in
SLAM and provides practical algorithms that can cope with
outliers without relying on any initial guess. We present two
main contributions. The first is a *Discrete-Continuous Graphical
Model* (DC-GM) for SLAM: the continuous portion of the DC-
GM captures the standard SLAM problem, while the discrete
portion describes the selection of the outliers and models their
correlation. The second contribution is a semidefinite relaxation
to perform inference in the DC-GM that returns estimates
with provable sub-optimality guarantees. Experimental results
on standard benchmarking datasets show that the proposed
technique compares favorably with state-of-the-art methods
while not relying on an initial guess for optimization.

I. INTRODUCTION

Simultaneous Localization and Mapping (SLAM) is the
backbone of several robotics applications. SLAM is already
widely adopted in consumer applications (e.g., robot vacuum
cleaning, warehouse maintenance, virtual/augmented reality),
and is a key enabler for truly autonomous systems operating
in the wild, ranging from unmanned aerial vehicles operating
in GPS-denied scenarios, to self-driving cars.

Despite the remarkable advances in SLAM, both re-
searchers and practitioners are well aware of the brittleness of
current SLAM systems. While SLAM failures are a tolerable
price to pay in some consumer applications, they may put
human life at risk in several safety-critical applications. For
this reason, SLAM is often avoided in those applications
(e.g., self-driving cars) in favor of alternative solutions where
the map is built beforehand in an offline (and typically
human-supervised) manner, even though this implies extra
setup costs.

[1]P. Lajoie and G. Beltrame are with the Department of Computer and
Software Engineering, École Polytechnique de Montréal, Montreal, Canada
{pierre-yves.lajoie,giovanni.beltrame}@polymtl.ca
[2]S. Hu and L. Carlone are with the Laboratory for Information & Decision
Systems (LIDS), Massachusetts Institute of Technology, Cambridge, USA,
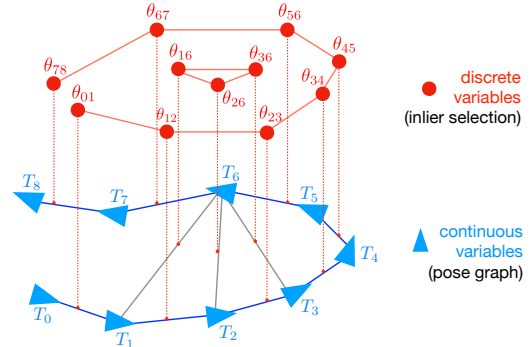{siyi,lcarlone}@mit.edu

Fig. 1. We introduce a *Discrete-Continuous Graphical Model* (DC-GM) to
model perceptual aliasing in SLAM. The model describes the interactions
between continuous variables (e.g., robot poses) and discrete variables (e.g.,
the binary selection of inliers and outliers), and captures the correlation
between the discrete variables (e.g., due to perceptual aliasing).

Arguably, the main cause of SLAM failure is the presence
of incorrect data association and outliers. Incorrect data
association is caused by *perceptual aliasing*, the phenomenon
where different places generate a similar visual (or, in
general, perceptual) footprint. Perceptual aliasing leads to
incorrectly associating the measurements taken by the robot
to the wrong portion of the map, which may lead to map
deformations and potentially to catastrophic failure of the
mapping process. The problem is exacerbated by the fact
that those outliers are *highly correlated*: due to the temporal
nature of the data collection, perceptual aliasing creates a
large number of mutually-consistent outliers. This correlation
makes it even harder to judge if a measurement is an out-
lier, contributing to the brittleness of the resulting pipeline.
Surprisingly, while the SLAM literature extensively focused
on mitigating the effects of perceptual aliasing, none of
the existing approaches attempt to explicitly model positive
correlation between outliers.

Contribution. This work provides a unified framework to
model perceptual aliasing and outlier correlation in SLAM.
We propose a novel approach to obtain provably-robust
SLAM algorithms: rather than developing techniques to *mit-
igate* the impact of perceptual aliasing, we *explicitly model*
perceptual aliasing using a *discrete-continuous graphical
model* (DC-GM). A simple illustration is given in Fig. 1.
The figure shows a DC-GM where the continuous variables,
shown in blue, describe a standard SLAM formulation, i.e., a
pose graph, where the triangles represent the trajectory of a
moving robot while the edges represent measurements. The
figure shows that we associate a discrete variable (large red
circles) to each edge/measurement in the pose graph. The
discrete variables decide between accepting or rejecting a
measurement. The red edges in the top portion of the figure
model the correlation between discrete variables. The expert
reader can recognize the top of the figure (graph in red), to be
a discrete *Markov Random Field* (MRF) [1]. The proposed

model can naturally capture positive correlation between outliers: for instance, we can model the correlation between three nearby edges, (T_1, T_6), (T_2, T_6), (T_3, T_6) in Fig. 1, as a clique involving the corresponding discrete variables $(\theta_{16}, \theta_{26}, \theta_{36})$ in the MRF (red triangle in the figure). Similarly, we can capture the temporal correlation of wheel slippage episodes by connecting variables corresponding to consecutive edges.

Our second contribution is the design of a semidefinite (SDP) relaxation that computes a near-optimal estimate of the variables in the DC-GM. Inference in DC-GM is intractable in general, due to the nonconvexity of the corresponding estimation problem and to the presence of discrete variables. We show how to obtain an SDP relaxation with per-instance sub-optimality guarantees, generalizing previous work on provably-correct SLAM without outliers [2], [3], [4], [5], [6]. The SDP relaxation can be solved in polynomial time by off-the-shelf convex solvers without relying on an initial guess.

Our last contribution is an experimental evaluation on standard SLAM benchmarking datasets. The experimental results show that the proposed DC-GM model compares favorably with state-of-the-art methods, including Vertigo [7], RRR [8] and DCS [9]. Moreover, they confirm that modeling outlier correlation further increases the resilience of the proposed model, which is able to compute correct SLAM estimates even when 50% of the loop closures are highly-correlated outliers. Our current (Matlab) implementation is slow, compared to state-of-the-art methods, but the proposed approach can be sped-up by designing a specialized solver along the lines of [6]. We leave these numerical aspects (which are both interesting and non-trivial on their own) for future work.

II. PRELIMINARIES AND RELATED WORK

This section reviews basic concepts about Markov Random Fields and Pose Graph Optimization.

A. Markov Random Fields (MRFs)

Markov Random Fields (MRFs) are a popular graphical model for reconstruction and recognition problems in computer vision and robotics [10], [1], [11]. A *pairwise MRF* is defined by a set of ℓ *nodes* we want to label, and a set of *edges* or *potentials*, representing probabilistic constraints involving the labels of a single or a pair of nodes. Here we consider *binary* MRFs, where we associate a binary label $\theta_i \in \{-1, +1\}$ to each node $i = 1, \ldots, \ell$.

The *maximum a posteriori* (MAP) estimate of the variables in the MRF is the assignment of the node labels that attains the maximum of the posterior distribution of an MRF, or, equivalently, the minimum of the negative log-posterior [10]:

$$\min_{\substack{\theta_i \in \{-1, +1\} \\ i=1,\ldots,\ell}} -\sum_{i \in \mathcal{U}} \bar{c}_i \theta_i - \sum_{(i,j) \in \mathcal{B}} \bar{c}_{ij} \theta_i \theta_j \quad (1)$$

where $\mathcal{U} \subseteq \{1, \ldots, \ell\}$ is the set of *unary potentials* (terms involving a single node), $\mathcal{B} \subseteq \{1, \ldots, \ell\} \times \{1, \ldots, \ell\}$ is the set of *binary potentials* (involving a pair of nodes). Intuitively, if $\bar{c}_i > 0$ (resp. $\bar{c}_i < 0$), then the unary terms encourage $+1$ (resp. -1) labels for node i. Similarly, if $\bar{c}_{ij} > 0$, then the binary term (i,j) encourages nodes i

and j to have the same label (positive correlation) since that decreases the cost (1) by $\bar{c}_{ij}$. While several choices of unary and binary potentials are possible, the expression in eq. (1) is a very popular model, and is referred to as the *Ising model* [1, Section 1.4.1].

Related works consider extensions of (1) to continuous [12], discrete-continuous [13], or discretized [14] labels, while to the best of our knowledge, our paper is the first to propose a semidefinite solver for discrete-continuous models and use these models to capture perceptual aliasing in SLAM.

B. Pose Graph Optimization (PGO)

Pose Graph Optimization (PGO) is one of the most popular models for SLAM. PGO consists in the estimation of a set of poses (i.e., rotations and translations) from pairwise relative pose measurements. In computer vision a similar problem (typically involving only rotation) is used as a preprocessing step for bundle adjustment in Structure from Motion (SfM) [15].

PGO estimates n poses from m relative pose measurements. Each to-be-estimated pose $T_i \doteq [R_i \ t_i]$, $i = 1, \ldots, n$, comprises a *translation* vector $t_i \in \mathbb{R}^d$ and a *rotation* matrix $R_i \in \mathrm{SO}(d)$, where $d = 2$ in planar problems or $d = 3$ in three-dimensional problems. For a pair of poses (i,j), a relative pose measurement $[\bar{R}_{ij} \ \bar{t}_{ij}]$, with $\bar{t}_{ij} \in \mathbb{R}^d$ and $\bar{R}_{ij} \in \mathrm{SO}(d)$, describes a noisy measurement of the relative pose between T_i and T_j. Each measurement is assumed to be sampled from the following generative model:

$$\bar{t}_{ij} = R_i^\mathsf{T}(t_j - t_i) + t_{ij}^\epsilon, \qquad \bar{R}_{ij} = R_i^\mathsf{T} R_j R_{ij}^\epsilon \quad (2)$$

where $t_{ij}^\epsilon \in \mathbb{R}^d$ and $R_{ij}^\epsilon \in \mathrm{SO}(d)$ represent translation and rotation measurement noise, respectively. PGO can be thought as an MRF[1] with variables living on manifold: we need to assign a pose to each node in a graph, given relative measurements associated to edges $\mathcal{E}$ of the graph. The resulting graph is usually referred to as a *pose graph*.

Assuming the translation noise is Normally distributed with zero mean and information matrix $\omega_t \mathbf{I}_d$ and the rotation noise follows a Langevin distribution [4], [3] with concentration parameter ω_r, the MAP estimate for the unknown poses can be computed by solving the following optimization problem:

$$\min_{\substack{t_i \in \mathbb{R}^d \\ R_i \in \mathrm{SO}(d)}} \sum_{(i,j) \in \mathcal{E}} \omega_t \|t_j - t_i - R_i \bar{t}_{ij}\|_2^2 + \frac{\omega_r}{2} \|R_j - R_i \bar{R}_{ij}\|_F^2 \quad (3)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. The derivation of (3) is given in [3, Proposition 1]. The estimator (3) involves solving a nonconvex optimization, due to the nonconvexity of the set $\mathrm{SO}(d)$. Recent results [3], [6] show that one can still compute a globally-optimal solution to (3), when the measurement noise is reasonable, using convex relaxations.

Unfortunately, the minimization (3) follows from the assumption that the measurement noise is light-tailed (e.g., Normally distributed translation noise) and it is known to produce completely wrong pose estimates when this assumption is violated, i.e., in presence of outlying measurements.

---

[1] An interpretation of landmark-based SLAM as a pairwise MRF with continuous variables is given by Dellaert in [16].
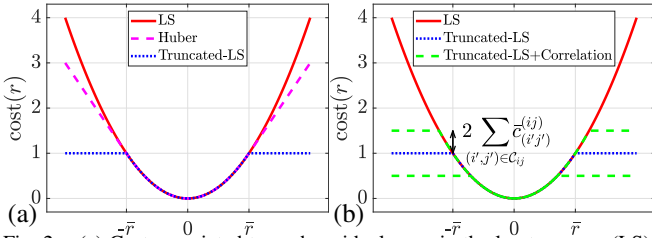
Fig. 2. (a) Cost associated to each residual error in the least squares (LS), Huber, and truncated LS estimators. (b) The correlation terms $\bar{c}_{(i'j')}^{(ij)}$ in eq. (9) have the effect of altering the error threshold in the truncated LS estimator.

C. Robust PGO

The sensitivity to outliers of the formulation (3) is due to the fact that we minimize the squares of the *residual errors* (quantities appearing in the squared terms): this implies that large residuals corresponding to spurious measurements dominate the cost. Robust estimators reduce the impact of outliers by adopting cost functions that grow slowly (i.e., less than quadratically) when the residual exceeds a given upper bound $\bar{r}$. This is the idea behind robust M-estimators, see [17]. For instance, the Huber loss in Fig. 2 grows linearly outside the quadratic region $[-\bar{r}, +\bar{r}]$. Ideally, one would like to adopt a truncated least squares (LS) formulation (Fig. 2) where the impact of arbitrarily large outliers remains bounded. Such a formulation, however, is non-convex and non-differentiable, typically making the resulting optimization hard.

Traditionally, outlier mitigation in SLAM and SfM relied on the use of robust M-estimators, see [18], [15]. Agarwal *et al.* [9] propose *Dynamic Covariance Scaling* (DCS), which dynamically adjusts the measurement covariances to reduce the influence of outliers. Olson and Agarwal [19] use a max-mixture distribution to accommodate multiple hypotheses on the noise distribution of a measurement. Casafranca *et al.* [20] minimize the ℓ_1-norm of the residual errors. Lee *et al.* [21] use expectation maximization. An alternative set of approaches attempts to explicitly *identify* and reject outliers. Early techniques include RANSAC [22] and branch & bound [23]. Sünderhauf and Protzel [7], [24] propose Vertigo, which augments the PGO problem with latent binary variables (then relaxed to continuous variables) that are responsible for deactivating outliers. Latif *et al.* [8], Carlone *et al.* [25], Graham *et al.* [26], Mangelson *et al.* [27] look for large sets of "mutually consistent" measurements. Pfingsthorn and Birk [28], [29] model ambiguous measurements using hyperedges and mixture of Gaussians, and provide a measurement selection approach that also constructs an initial guess for PGO. Both [19] and [29] implicitly model negative correlation (or, more precisely, mutual exclusivity) between multiple edge hypotheses. The introduction of discrete variables has also been used to reconcile data association and semantic SLAM [30], and to deal with unknown data association in SfM [31].

III. DISCRETE-CONTINUOUS GRAPHICAL MODELS FOR ROBUST POSE GRAPH OPTIMIZATION

We propose a novel approach for robust PGO that addresses the three main limitations of the state of the art. First, rather than mitigating outlier correlation, we explicitly model it. Second, our PGO method (Section IV) does not rely on any initial guess. Third, we go beyond recently proposed convex relaxations for robust rotation and pose estimation [32], [5], [33], and use a *nonconvex* loss, namely, the *truncated LS* cost in Fig. 2. This circumvents issues with convex robust loss functions which are known to have low breakdown point (e.g., the Huber loss [5] or ℓ_1 norm [32], [5], [33] can be compromised by the presence of a single "bad" outlier).

A. A unified view of robust PGO

Let us partition the edges of the pose graph into odometric edges $\mathcal{E}_{od}$ and loop-closure edges $\mathcal{E}_{lc}$. Perceptual aliasing affects exteroceptive sensors, hence —while we can typically trust odometric edges— loop closures may include outliers.

According to the discussion in Section II-C, an ideal formulation for robust PGO would use a truncated LS cost for the loop-closure edges in $\mathcal{E}_{lc}$:

$$
\begin{aligned}
\min_{\substack{\boldsymbol{t}_i \in \mathbb{R}^d \\ \boldsymbol{R}_i \in \mathrm{SO}(d)}} & \sum_{(i,j)\in\mathcal{E}_{od}} \omega_t \|\boldsymbol{t}_j - \boldsymbol{t}_i - \boldsymbol{R}_i \bar{\boldsymbol{t}}_{ij}\|_2^2 + \frac{\omega_r}{2} \|\boldsymbol{R}_j - \boldsymbol{R}_i \bar{\boldsymbol{R}}_{ij}\|_F^2 \\
& + \sum_{(i,j)\in\mathcal{E}_{lc}} \omega_t f_{\bar{c}_t}(\|\boldsymbol{t}_j - \boldsymbol{t}_i - \boldsymbol{R}_i \bar{\boldsymbol{t}}_{ij}\|_2) + \frac{\omega_r}{2} f_{\bar{c}_R}(\|\boldsymbol{R}_j - \boldsymbol{R}_i \bar{\boldsymbol{R}}_{ij}\|_F)
\end{aligned}
\tag{4}
$$

where, for a positive scalar c, the function $f_c(\cdot)$ is:

$$
f_c(x) = \left\{ \begin{array}{ll} x^2 & \text{if } |x| \le c \\ c^2 & \text{otherwise} \end{array} \right.
\tag{5}
$$

While the formulation (4) would be able to tolerate arbitrarily "bad" outliers, it has two main drawbacks. First, $f_c(\cdot)$ is non-convex, adding to the non-convexity already induced by the rotations ($\mathrm{SO}(d)$ is a non-convex set). Second, the cost is non-differentiable, as shown in Fig. 2, hence also preventing the use of fast (but local) smooth optimization techniques.

The first insight behind the proposed approach is simple but powerful: we can rewrite the truncated LS cost (5) as a minimization over a binary variable:

$$
f_c(x) = \min_{\theta \in \{-1; +1\}} \frac{(1+\theta)}{2} x^2 + \frac{(1-\theta)}{2} c^2
\tag{6}
$$

To show the equivalence between (6) and (5), we observe that for any $\bar{x}$ such that $\bar{x}^2 < c^2$ (or $|\bar{x}| < c$), the minimum in (6) is attained for $\theta = +1$ and $f_c(\bar{x}) = \bar{x}^2$; on the other hand, for any $\hat{x}$ such that $\hat{x}^2 > c^2$ (or $|\hat{x}| > c$), the minimum in (6) is attained for $\theta = -1$ and $f_c(\hat{x}) = c^2$.

We can now use the expression (6) to rewrite the cost function (4) by introducing a binary variable for each rotation and translation measurement:

$$
\begin{aligned}
\min_{\substack{\boldsymbol{t}_i \in \mathbb{R}^d \\ \boldsymbol{R}_i \in \mathrm{SO}(d) \\ \theta_{ij}^t \in \{-1;+1\} \\ \theta_{ij}^R \in \{-1;+1\}}} & \sum_{(i,j)\in\mathcal{E}_{od}} \omega_t \|\boldsymbol{t}_j - \boldsymbol{t}_i - \boldsymbol{R}_i \bar{\boldsymbol{t}}_{ij}\|_2^2 + \frac{\omega_r}{2} \|\boldsymbol{R}_j - \boldsymbol{R}_i \bar{\boldsymbol{R}}_{ij}\|_F^2 + \\
& + \sum_{(i,j)\in\mathcal{E}_{lc}} \omega_t \frac{(1+\theta_{ij}^t)}{2} \|\boldsymbol{t}_j - \boldsymbol{t}_i - \boldsymbol{R}_i \bar{\boldsymbol{t}}_{ij}\|_2^2 + \omega_t \frac{(1-\theta_{ij}^t)}{2} \bar{c}_t^2 \\
& + \frac{\omega_r}{2} \frac{(1+\theta_{ij}^R)}{2} \|\boldsymbol{R}_j - \boldsymbol{R}_i \bar{\boldsymbol{R}}_{ij}\|_F^2 + \frac{\omega_r}{2} \frac{(1-\theta_{ij}^R)}{2} \bar{c}_R^2
\end{aligned}
\tag{7}
$$

where $\bar{c}_t$ and $\bar{c}_R$ are simply the largest admissible residual errors for a translation and rotation measurement to

be considered an inlier. Intuitively, θ_{ij}^t decides whether a translation measurement is an inlier ($\theta_{ij}^t = +1$) or an outlier ($\theta_{ij}^t = -1$); θ_{ij}^R has the same role for rotation measurements. While eq. (7) resembles formulations in the literature, e.g., Sünderhauf's switchable constraints [7], establishing connections with the truncated LS cost provides a physically meaningful interpretation of the parameters $\bar{c}_t$ and $\bar{c}_R$ (maximum admissible residuals). Moreover, we will push the boundary of the state of the art by modeling the outlier correlation (next sub-section) and proposing global semidefinite solvers (Section IV).

B. Modeling outlier correlation and perceptual aliasing

The goal of this section is to introduce extra terms in the cost (7) to model the correlation between subsets of binary variables, hence capturing outlier correlation. For the sake of simplicity, we assume that a unique binary variable is used to decide if both the translation and the rotation components of measurement (i, j) are accepted, i.e., we set $\theta_{ij}^t = \theta_{ij}^R \doteq \theta_{ij}$. This assumption is not necessary for the following derivation, but it allows using a more compact notation. In particular, we rewrite (7) more succinctly as:

$$\min_{\substack{T_i \in \mathrm{SO}(d) \times \mathbb{R}^d \\ \theta_{ij} \in \{-1; +1\}}} \sum_{(i,j) \in \mathcal{E}_{od}} \|T_j - T_i \bar{T}_{ij}\|_{\Omega}^2$$
$$+ \sum_{(i,j) \in \mathcal{E}_{lc}} \frac{(1 + \theta_{ij})}{2} \|T_j - T_i \bar{T}_{ij}\|_{\Omega}^2 + \frac{(1 - \theta_{ij})}{2} \bar{c} \tag{8}$$

where for two matrices M and Ω of compatible dimensions $\|M\|_{\Omega}^2 \doteq \mathrm{tr}\left(M \Omega M^{\mathsf{T}}\right)$, and –following [34]– we defined:

$$T_i \doteq [R_i\ t_i], \quad \bar{T}_{ij} \doteq \begin{bmatrix} \bar{R}_{ij} & \bar{t}_{ij} \\ \mathbf{0}_d^{\mathsf{T}} & 1 \end{bmatrix}, \quad \Omega \doteq \begin{bmatrix} \frac{\omega_r}{2} \mathbf{I}_d & \mathbf{0}_d \\ \mathbf{0}_d^{\mathsf{T}} & \omega_t \end{bmatrix}$$

and for simplicity we called $\bar{c} \doteq \omega_r \bar{c}_t^2 + \frac{\omega_t}{2} \bar{c}_R^2$.

We already observed in Section II-A that to model the correlation between two discrete variables θ_{ij} and $\theta_{i'j'}$ we can add terms $-\bar{c}_{(i'j')}^{(ij)} \theta_{ij} \theta_{i'j'}$ to the cost function, which penalize a mismatch between θ_{ij} and $\theta_{i'j'}$ whenever the scalar $\bar{c}_{(i'j')}^{(ij)}$ is positive. This leads to generalizing problem (8) as follows:

$$\min_{\substack{T_i \in \mathrm{SO}(d) \times \mathbb{R}^d \\ \theta_{ij} \in \{-1; +1\}}} \sum_{(i,j) \in \mathcal{E}_{od}} \|T_j - T_i \bar{T}_{ij}\|_{\Omega}^2$$
$$+ \sum_{(i,j) \in \mathcal{E}_{lc}} \frac{(1 + \theta_{ij})}{2} \|T_j - T_i \bar{T}_{ij}\|_{\Omega}^2 + \frac{(1 - \theta_{ij})}{2} \bar{c} - \sum_{(i,j),(i'j') \in \mathcal{C}} \bar{c}_{(i'j')}^{(ij)} \theta_{ij} \theta_{i'j'} \tag{9}$$

where the set $\mathcal{C}$ contains pairs of edges that are correlated, i.e., pairs of edges $(i, j), (i', j')$ for which if (i, j) is found to be an outlier, it is likely for (i', j') to be an outlier as well.

In the supplemental material [35], we show that the correlation terms have the effect of altering the threshold $\bar{c}$. For instance, if all neighbors (i', j') of an edge (i, j) are inliers ($\theta_{i'j'} = 1$), the correlation terms become $\bar{c}_{(i'j')}^{(ij)} \theta_{ij}$ and they have the effect of increasing $\bar{c}$. We also show that perturbations of $\bar{c}$ are bounded in the interval $[\bar{c} - 2 \sum_{(i',j') \in \mathcal{C}_{ij}} \bar{c}_{(i'j')}^{(ij)}, \bar{c} + 2 \sum_{(i',j') \in \mathcal{C}_{ij}} \bar{c}_{(i'j')}^{(ij)}]$, where $\mathcal{C}_{ij}$ is the

set of edges correlated to the edge (i, j), see Fig.2(b) for an illustration.

Problem (9) describes a *discrete-continuous graphical model* (DC-GM) as the one pictured in Fig. 1: the optimization problems returns the most likely assignment of variables in the graphical model, which contains both continuous variables (T_i) and discrete variables (θ_{ij}). The reader can notice that if the assignment of discrete variables is given, (9) reduces to PGO, while if the continuous variables are given, then (9) becomes an MRF, where the second sum in (9) defines the unary potentials for each discrete variable in the MRF.

IV. INFERENCE IN DC-GM VIA CONVEX RELAXATION

The DC-GM presented in Section III captures two very desirable aspects: (i) it uses a robust truncated LS loss function and (ii) it can easily model outlier correlation. On the downside, the optimization (9) is intractable in general, due to the presence of discrete variables and the non-convex nature of the rotation set SO(d).

Here we derive a convex relaxation that is able to compute near-optimal solutions for (9) in polynomial time. While we do not expect to compute exact solutions for (9) in all cases in polynomial time (the problem is NP-hard in general), our goal is to obtain a relaxation that works well when the noise on the inliers is reasonable (i.e., similar to the one found in practical applications) and whose quality is largely insensitive to the presence of a large number of (arbitrarily "bad") outliers.

In order to derive our convex relaxation, it is convenient to reformulate (9) using a more compact matrix notation. Let us first "move" the binary variables inside the norm and drop constant terms from the objective in (9):

$$\min_{\substack{T_i \in \mathrm{SO}(d) \times \mathbb{R}^d \\ \theta_{ij} \in \{-1; +1\}}} \sum_{(i,j) \in \mathcal{E}_{od}} \|T_j - T_i \bar{T}_{ij}\|_{\Omega}^2$$
$$+ \sum_{(i,j) \in \mathcal{E}_{lc}} \|\frac{(1 + \theta_{ij})}{2} (T_j - T_i \bar{T}_{ij})\|_{\Omega}^2$$
$$- \sum_{(i,j) \in \mathcal{E}_{lc}} \frac{\theta_{ij}}{2} \bar{c} - \sum_{(i,j),(i'j') \in \mathcal{C}} \bar{c}_{(i'j')}^{(ij)} \theta_{ij} \theta_{i'j'} \tag{10}$$

where we noted that $\frac{(1+\theta_{ij})}{2}$ is either zero or one, hence it can be safely moved inside the norm, and we dropped $\frac{1}{2} \bar{c}$.

We can now stack pose variables into a single $d \times (d+1)n$ matrix $T \doteq [T_1 \ldots T_n]$. We also use a *matrix representation* for the binary variables $\Theta \doteq [\Theta_1 \ldots \Theta_\ell] \in \{-I_d; +I_d\}^\ell$ where $\ell = |\mathcal{E}_{lc}|$ denotes the number of loop closures and I_d denotes the identity matrix of size d. Finally, we define:

$$X = [T \quad \Theta \quad I_d] \in (\mathrm{SO}(d) \times \mathbb{R}^d)^n \times \{-I_d; +I_d\}^\ell \times I_d$$
$$\left(\text{note: } X^{\mathsf{T}} X = \begin{bmatrix} T^{\mathsf{T}} T & T^{\mathsf{T}} \Theta & T^{\mathsf{T}} \\ \Theta^{\mathsf{T}} T & \Theta^{\mathsf{T}} \Theta & \Theta^{\mathsf{T}} \\ T & \Theta & I_d \end{bmatrix} \right) \tag{11}$$

The following proposition provides a compact reformulation of problem (10) using the matrix X in (11):

Proposition 1 (Inference in DC-GM). *Problem* (10) *can be equivalently written in compact form using the matrix*

variable $\boldsymbol{X}$ in (11) *as follows:*

$$\min_{\boldsymbol{X}} \; \operatorname{tr}\left(\boldsymbol{Q}\boldsymbol{X}^\mathsf{T}\boldsymbol{X}\right) + \sum_{e=(i,j)\in\mathcal{E}_{lc}} \operatorname{tr}\left(\boldsymbol{U}_e\boldsymbol{X}^\mathsf{T}\boldsymbol{X}\boldsymbol{W}_e\boldsymbol{X}^\mathsf{T}\boldsymbol{X}\right)$$

$$\text{subject to} \quad \boldsymbol{X} \in (SO(d)\times\mathbb{R}^d)^n \times \{-\boldsymbol{I}_d; +\boldsymbol{I}_d\}^\ell \times \boldsymbol{I}_d \tag{12}$$

where $\boldsymbol{Q}, \boldsymbol{U}_e, \boldsymbol{W}_e \in \mathbb{R}^{(n(d+1)+d\ell+d)\times(n(d+1)+d\ell+d)}$ *are sparse matrices (for all loop closures $e \in \mathcal{E}_{lc}$). The expressions for these (known) matrices are given in Appendix.*

Intuitively, $\boldsymbol{Q}$ in (12) captures the terms in the first, third, and fourth sum in (10), while the sum including $\boldsymbol{U}_e, \boldsymbol{W}_e$ (one term for each loop closure e) captures the terms in the second sum in (10) which couples discrete and continuous variables.

The final step before obtaining a convex relaxation is to write the "geometric" constraints in (12) in terms of linear algebra. Towards this goal, we relax the set SO(d) (rotation matrices) to O(d) (orthogonal matrices), i.e., we drop the constraint that rotation matrices need to have determinant $+1$. In related work, we found the determinant constraint to be redundant [36]. Moreover, this is done for the sake of simplicity, while the determinant constraints can be still modeled as shown in [36]. Then, we obtain an SDP relaxation of Problem (12) by (i) introducing a matrix variable $\boldsymbol{Z} = \boldsymbol{X}^\mathsf{T}\boldsymbol{X}$ and rewriting (12) as a function of $\boldsymbol{Z}$, (ii) noting that any matrix $\boldsymbol{Z} = \boldsymbol{X}^\mathsf{T}\boldsymbol{X}$ is a positive-semidefinite ($\boldsymbol{Z} \succeq 0$) rank-d matrix, and (iii) relaxing the non-convex rank-d constraint.

Proposition 2 (Semidefinite Relaxation of DC-GM). *The following SDP is a convex relaxation of Problem* (12):

$$\min_{\boldsymbol{Z}} \; \operatorname{tr}\left(\boldsymbol{Q}\boldsymbol{Z}\right) + \sum_{e=(i,j)\in\mathcal{E}_{lc}} \operatorname{tr}\left(\boldsymbol{U}_e\boldsymbol{Z}\boldsymbol{W}_e\boldsymbol{Z}\right)$$

$$\text{subject to} \quad [\boldsymbol{Z}]_{ii} = \begin{bmatrix} \boldsymbol{I}_d & * \\ * & * \end{bmatrix} \quad i=1,\dots,n$$
$$[\boldsymbol{Z}]_{ii} = \boldsymbol{I}_d \qquad\qquad i=n+1,\dots,n+\ell+1$$
$$[\boldsymbol{Z}]_{ij} = \operatorname{idiag}([\boldsymbol{Z}]_{ij}) \quad i,j=n+1,\dots,n+\ell+1$$
$$\boldsymbol{Z} \succeq 0$$

$$\tag{13}$$

where $[\boldsymbol{Z}]_{ij}$ denotes the block of $[\boldsymbol{Z}]$ in block row i and block column j, the symbol "$*$" denotes entries that are unconstrained (we follow the notation of [34]), and where $[\boldsymbol{Z}]_{ij} = \operatorname{idiag}([\boldsymbol{Z}]_{ij})$ enforces the block $[\boldsymbol{Z}]_{ij}$ to be an isotropic diagonal matrix, i.e., a scalar multiple of $\boldsymbol{I}_d$.

Let us explain the constraints in (13), by using the block structure of $\boldsymbol{Z}$ described in (11). For $i = 1,\dots,n$, the diagonal blocks $[\boldsymbol{Z}]_{ii}$ are in the form of $\boldsymbol{T}_i^\mathsf{T}\boldsymbol{T}_i$, hence the first constraint in (13) captures the orthogonality of the rotation matrix included in each pose $\boldsymbol{T}_i$. For $i = n+1,\dots,n+\ell$, the diagonal blocks $[\boldsymbol{Z}]_{ii}$ are in the form of $\Theta_i^\mathsf{T}\Theta_i$ and since $\Theta_i \in \{-\boldsymbol{I}_d, +\boldsymbol{I}_d\}$, $\Theta_i^\mathsf{T}\Theta_i = \boldsymbol{I}_d$, which is captured in the second constraint in (13); similar considerations hold for $i = \ell+1$. Finally, the products $\Theta_i^\mathsf{T}\Theta_j$ (captured by the blocks $[\boldsymbol{Z}]_{ij}$ when $i,j = n+1,\dots,n+\ell+1$) must be diagonal matrices, producing the third constraint in (13).

The SDP relaxation can be solved using off-the-shelf convex solvers. In particular, we note that the constraint $[\boldsymbol{Z}]_{ij} = \operatorname{idiag}([\boldsymbol{Z}]_{ij})$ can be implemented as a set of linear equality constraints. Indeed, this constraint can be rewritten as $[\boldsymbol{Z}]_{ij} = [\boldsymbol{Z}]_{ij,11} \cdot \boldsymbol{I}_d$ where $[\boldsymbol{Z}]_{ij,11}$ is the top left entry of $[\boldsymbol{Z}]_{ij}$. Therefore, the constraint enforces that the matrix

has offdiagonal elements equal to zero and diagonal elements equal to a single scalar $[\boldsymbol{Z}]_{ij,11}$. The SDP relaxation (13) enjoys the typical per-instance optimality guarantees described in related work [3], [4], [5], [6]. In particular, if the solution $\boldsymbol{Z}^\star$ of (13) has rank d, then the relaxation solves (12) *exactly*. Moreover, the optimal objective of (13) is a lower bound for the optimal objective (12), a property that can be used to evaluate how sub-optimal a given estimate is, see [3], [4].

V. EXPERIMENTS

This section presents two sets of experiments. Section V-A reports the results of Monte Carlo runs on a synthetic dataset and shows that the proposed technique compares favorably with the state of the art, and that modeling outlier correlation leads to performance improvements. Section V-B evaluates the proposed techniques in three real benchmarking datasets and shows that our approach outperforms related techniques while not requiring any initial guess.

A. Experiments On Synthetic Dataset

Methodology. For this set of experiments, we built a synthetic dataset composed of a simple trajectory on a grid of 20 by 10 nodes. Then we added random groups of loop closures between the rows as described in [7]. Typically, in presence of perceptual aliasing, the outliers are in mutually-consistent groups, e.g., the SLAM front-end generates multiple false loop closures in sequence. To simulate this phenomenon, we set the loop closures in each group to be either all inliers or all outliers. We set the standard deviation of the translation and rotation noise for the inlier measurements (odometry and correct loop closures) to 0.1m and 0.01rad. The maximum admissible errors for the truncated LS (5) is set to 1σ of the measurement noise. We tested the performance of our techniques for increasing levels of outliers, up to the case where 50% of the loop closure are outliers. Fig. 4 shows the overlay of multiple trajectories (5 runs) estimated by our techniques versus the ground truth trajectory (green), when 50% of the loop closures are outliers.

Compared Techniques. We evaluate the performance of the proposed technique, DC-GM, which solves the minimization problem (9). In order to show that capturing outlier correlation leads to performance improvements, we also test a variation of the proposed approach, called DC-GMd, which implements the minimization problem (8), where outliers are assumed uncorrelated (the "d" stands for decoupled). In both DC-GM and DC-GMd, we solve the SDP using CVX [37] in Matlab. If the resulting matrix does not have rank $d = 2$ (in which case we are not guaranteed to get an exact solution to the non-relaxed problem), we round the result to detect the set of outliers, and re-run the optimization without the outliers.

We benchmarked our approach against three other robust PGO techniques, i.e., Vertigo [7], RRR [8] and DCS [9]. For Vertigo we use the default parameters, while for RRR and DCS we report results for multiple choices of parameters, since these parameters have a significant impact on performance. In particular, for RRR we consider three cluster sizes ($t_g = \{1, 5, 10\}$) and for DCS we considered three values of the parameter $\Phi = \{1, 10, 100\}$ [9]. For all these techniques, we used the odometric estimate as initial guess.
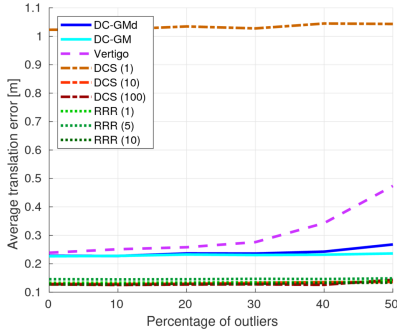
Fig. 3. Average translation error of the 9 approaches tested in this paper with an increasing percentage of outliers.
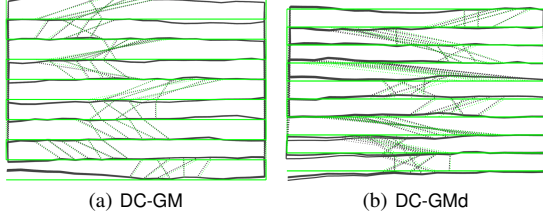

(a) DC-GM   (b) DC-GMd

Fig. 4. Trajectory estimates computed by the proposed techniques (black, overlay of 5 runs) versus ground truth (green) for the simulated grid dataset.

Fig. 5. Percentage of rejected outliers for the proposed techniques.

(a) CSAIL   (b) FR079   (c) FRH

Fig. 6. Trajectory estimates computed by DC-GM (black) versus ground truth (green) for the real datasets CSAIL, FR079, and FRH.

Results and Interpretation. Fig. 3 reports the average translation error for all the compared approaches and for increasing percentage of outliers. Vertigo's error grows quickly beyond 30% of outliers. For DCS, the performance heavily relies on correct parameter tuning: for some choice of parameters ($\Phi = \{10, 100\}$) it has excellent performance while the approach fails for $\Phi = 1$. Unfortunately, these parameters are difficult to tune in general (we will observe in Section V-B that the choice of parameters mentioned above may not produce the best results in the real tests). The proposed techniques, DC-GMd and DC-GM, compare favorably against the state of the art while they are slightly less accurate than RRR, which produced the best results in simulation.

In order to shed light on the performance of DC-GM and DC-GMd, Fig. 5 reports the average percentage of outliers rejected by these two techniques. While from the scale of the y-axis we note that both techniques are able to reject most outliers, DC-GM is able to reject *all* outliers in all tests even when up to 50% of the loop closures are spurious. As expected, modeling outlier correlation as in DC-GM improves outlier rejection performance. We also recorded the number of *incorrectly rejected inliers*: both approaches do not reject any inlier and for this reason we omit the corresponding figure.

In our tests, the SDP relaxation (13) typically produces low-rank solutions with 2 relatively large eigenvalues, followed by 2 smaller ones (the remaining eigenvalues are numerically zero). The interested reader can find statistics on the average rank, results for different choices of the thresholds $\bar{c}$ and $\bar{c}_{(i'j')}^{(ij)}$, and additional tests in a simulated Manhattan World in the supplemental material [35].

B. Experiments On Real Datasets

Methodology. In this section, we consider three real-world standard benchmarking datasets, the CSAIL dataset (1045 poses and 1172 edges), the FR079 dataset (989 poses and 1217 edges), and the FRH dataset (1316 poses and 2820 edge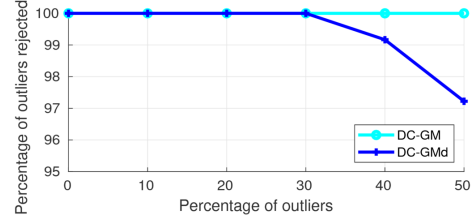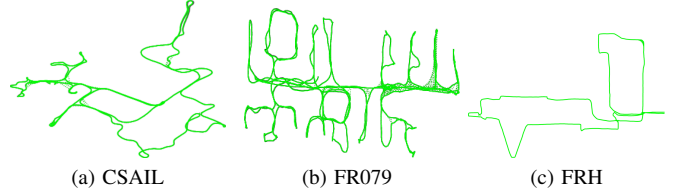s). We spoiled those datasets with 20 randomly grouped outliers. We add correlation terms with $\bar{c}_{(i'j')}^{(ij)} = 1$ for each pair of edges connecting consecutive nodes, e.g., (i, j) and $(i \pm 1, j \pm 1)$. We benchmarked our approach against Vertigo, RRR, and DCS.

Results and Interpretation. Table I presents the average translation error (computed with respect to the optimized trajectory without outliers) for all datasets and techniques. We also report the average translation error of the odometric estimate. All compared techniques achieve very good results on the FRH dataset. This is probably due to the fact that this dataset provides a very good initial guess, hence the techniques that rely on iterative optimization are favored. This intuition is confirmed by the high accuracy of the odometry. The results on the FR079 dataset are more interesting. In this case, DC-GM and RRR achieve the best results with a slight advantage towards DC-GM. However, Vertigo performs poorly and DCS performance remains worse than the proposed technique even with its best parameter choice. DC-GM has also the best performance on the CSAIL dataset. Again, RRR achieves very good results while Vertigo and DCS have poor performance except for some parameter choice (e.g., DCS performs well for $\Phi = \{100\}$). We attribute this performance boost to the fact that the proposed approach provides a more direct control on the maximum admissible error of each measurement, while the parameters in DCS and Vertigo have a less clear physical interpretation. This translates to the fact that it is more difficult for DCS and Vertigo to strike a balance between outlier rejection and inlier selection. Therefore, even when these approaches are able to discard most outliers, they may lose accuracy since they also tend to discard good measurements. The difficulty in performing parameter tuning for DCS is confirmed by the fact that the value $\Phi = 1$ (recommended by Agarwal *et al.* [9]) leads to good results on FRH, but fails on FR079 and CSAIL.

Fig. 6 shows the trajectory estimates produced by DC-GM for the three real datasets, CSAIL, FR079, and FRH.

VI. CONCLUSION

We introduced a discrete-continuous graphical model (DC-GM) to capture perceptual aliasing and outlier correlation in

	DC-GM	Vertigo	RRR (t_g=1)	RRR (t_g=5)	RRR (t_g=10)	DCS ($\Phi = 1$)	DCS ($\Phi = 10$)	DCS ($\Phi = 100$)	Odometry
FRH	0.0008	0.0005	0.0004	0.0003	0.0003	0.0004	0.0004	0.0004	0.0021
FR079	0.0438	0.2751	0.0546	0.0520	0.0521	0.2721	0.1804	0.1250	0.2836
CSAIL	0.0430	1.4625	0.0495	0.0613	0.0506	1.4576	1.4240	0.0521	1.4480

SLAM. Then we developed a semidefinite (SDP) relaxation to perform near-optimal inference in the DC-GM and obtain robust SLAM estimates. Our experiments show that the proposed approach compares favorably with the state of the art while not relying on an initial guess for optimization. Our approach also enables a more intuitive tuning of the parameters (e.g., the maximum admissible residual $\bar{c}$). The supplemental material [35] contains extra results to provide more insights on the performance and limitations of the proposed approach. This paper opens several avenues for future work. First, our Matlab implementation is currently slow: we plan to develop specialized solvers to optimize the SDP relaxations presented in this paper efficiently, leveraging previous work [6]. Second, we plan to extend our testing to 3D SLAM problems: the mathematical formulation in this paper is general, while for numerical reasons we had to limit our tests to relatively small 2D problems. Third, it would be useful to develop incremental solvers that can re-use computation when the measurements are presented to the robot in online (rather than batch) fashion. Finally, it would be interesting to provide a theoretical bound on the number of outliers the proposed technique can tolerate.

APPENDIX

This appendix proves Proposition 1 by showing how to reformulate problem (10) using the matrix $\boldsymbol{X}$ in (11). Let us start by rewriting problem (10) and replacing the (scalar) discrete variables $\theta_{ij} \in \{-1, +1\}$ with "binary" selection matrices $\Theta_{ij} \in \{-\boldsymbol{I}_d; +\boldsymbol{I}_d\}$:

$$\min_{\substack{\boldsymbol{T}_i \in \mathrm{SO}(d) \times \mathbb{R}^d \\ \Theta_{ij} \in \{-\boldsymbol{I}_d; +\boldsymbol{I}_d\}}} \sum_{(i,j) \in \mathcal{E}_{od}} \|\boldsymbol{T}_j - \boldsymbol{T}_i \bar{\boldsymbol{T}}_{ij}\|^2_{\boldsymbol{\Omega}}$$
$$- \sum_{(i,j) \in \mathcal{E}_{lc}} \frac{\bar{c}}{2d} \mathrm{tr}\left(\Theta_{ij}\right) - \sum_{(i,j),(i',j') \in \mathcal{P}} \frac{\bar{c}^{(ij)}_{(i'j')}}{d} \mathrm{tr}\left(\Theta_{ij}^\mathsf{T} \Theta_{i'j'}\right) \quad (14)$$
$$+ \sum_{(i,j) \in \mathcal{E}_{lc}} \frac{1}{4} \|\boldsymbol{T}_j - \boldsymbol{T}_i \bar{\boldsymbol{T}}_{ij} + \Theta_{ij}^\mathsf{T}(\boldsymbol{T}_j - \boldsymbol{T}_i \bar{\boldsymbol{T}}_{ij})\|^2_{\boldsymbol{\Omega}}$$

where we also rearranged the summands. Note the division by d in the second and third sum in (14), needed to compensate for the fact that we are now working with $d \times d$ matrices Θ_{ij}.

The first summation in (14) can be written as

$$\sum_{(i,j) \in \mathcal{E}_{od}} \|\boldsymbol{T}_j - \boldsymbol{T}_i \bar{\boldsymbol{T}}_{ij}\|^2_{\boldsymbol{\Omega}}$$
$$= \sum_{(i,j) \in \mathcal{E}_{od}} \mathrm{tr}\left((\boldsymbol{T}_j - \boldsymbol{T}_i \bar{\boldsymbol{T}}_{ij})\boldsymbol{\Omega}(\boldsymbol{T}_j - \boldsymbol{T}_i \bar{\boldsymbol{T}}_{ij})^\mathsf{T}\right) \quad (15)$$
$$= \mathrm{tr}\left(\boldsymbol{L}(\mathcal{G}_{od})\boldsymbol{T}^\mathsf{T}\boldsymbol{T}\right)$$

where $\boldsymbol{L}(\mathcal{G}_{od}) \in \mathbb{R}^{(d+1)n \times (d+1)n}$ is the *Connection Laplacian* [6] of the graph $\mathcal{G}_{od} = (\mathcal{V}, \mathcal{E}_{od})$, which has the same set of nodes $\mathcal{V}$ as the original pose graph, but only includes odometric edges $\mathcal{E}_{od}$. We can use a derivation similar to [34] to show that the *Connection Laplacian* of a generic graph $\mathcal{G} = (V, E)$ can be written as

$$\boldsymbol{L}(\mathcal{G}) = \boldsymbol{A}(\mathcal{G})\boldsymbol{\Omega}(\mathcal{G})\boldsymbol{A}(\mathcal{G})^\mathsf{T} \quad (16)$$

where, the matrices $\boldsymbol{A}(\mathcal{G}) \in \mathbb{R}^{(d+1)|V| \times (d+1)|E|}$ and $\boldsymbol{\Omega}(\mathcal{G}) \in \mathbb{R}^{(d+1)|E| \times (d+1)|E|}$ are given as follows:

$$[\boldsymbol{A}(\mathcal{G})]_{r,e} \doteq \begin{cases} -\bar{\boldsymbol{T}}_{i_e, j_e} & \text{if } r = i_e, \\ +\boldsymbol{I}_{d+1} & \text{if } r = j_e, \\ \boldsymbol{0}_{d+1} & \text{otherwise.} \end{cases} \quad \text{for } e = 1, \ldots, |E| \quad (17)$$

$$\boldsymbol{\Omega}(\mathcal{G}) \doteq \mathrm{blkdiag}(\boldsymbol{\Omega}_1, ..., \boldsymbol{\Omega}_{|E|}) \quad (18)$$

The notation $[\boldsymbol{A}(\mathcal{G})]_{r,e}$ denotes the $(d+1) \times (d+1)$ block of $\boldsymbol{A}(\mathcal{G})$ at block row r and block column e, while the e-th edge in E is denoted as (i_e, j_e).

The second summation in (14) can be developed as follows:

$$\sum_{(i,j) \in \mathcal{E}_{lc}} -\frac{\bar{c}}{2d} \mathrm{tr}\left(\Theta_{ij}\right) = -\frac{\bar{c}}{2d} \mathrm{tr}\left(\boldsymbol{I}_{d,\ell}^\mathsf{T}\Theta\right) \quad (19)$$

where $\boldsymbol{I}_{d,\ell} \doteq [\boldsymbol{I}_d \ldots \boldsymbol{I}_d]$ is a row of ℓ identity matrices.

Similarly, the third summation in (14) can be written as:

$$-\sum_{(i,j),(i',j') \in \mathcal{P}} \frac{\bar{c}^{(ij)}_{(i'j')}}{d} \mathrm{tr}\left(\Theta_{ij}^\mathsf{T}\Theta_{i'j'}\right) = -\frac{1}{2d}\mathrm{tr}\left(\boldsymbol{N}(\mathcal{C})\Theta^\mathsf{T}\Theta\right) \quad (20)$$

where $\boldsymbol{N}(\mathcal{C}) \in \mathbb{R}^{d\ell \times d\ell}$ has $d \times d$ blocks in the form:

$$[\boldsymbol{N}(\mathcal{C})]_{e,e'} \doteq \begin{cases} \bar{c}^{(ij)}_{(i'j')}\boldsymbol{I}_d & \text{if } e = (i,j), e' = (i', j') \in \mathcal{C} \\ \bar{c}^{(i'j')}_{(ij)}\boldsymbol{I}_d & \text{if } e = (i',j'), e' = (i, j) \in \mathcal{C} \\ \boldsymbol{0}_d & \text{otherwise.} \end{cases} \quad (21)$$

The first three terms in (14) are linear with respect to parts of the matrix $\boldsymbol{X}^\mathsf{T}\boldsymbol{X}$ in (11), so we write the sum of (15), (19), (20) compactly as $\mathrm{tr}\left(\boldsymbol{Q}\boldsymbol{X}^\mathsf{T}\boldsymbol{X}\right)$ where

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{L}(\mathcal{G}_{od}) & \boldsymbol{0}_{(d+1)n,d\ell} & \boldsymbol{0}_{(d+1)n,d} \\ \boldsymbol{0}_{d\ell,(d+1)n} & -\frac{1}{2d}\boldsymbol{N}(\mathcal{C}) & -\frac{\bar{c}}{4d}\boldsymbol{I}_{d,\ell}^\mathsf{T} \\ \boldsymbol{0}_{d,(d+1)n} & -\frac{\bar{c}}{4d}\boldsymbol{I}_{d,\ell} & \boldsymbol{0}_{d,d} \end{bmatrix} \quad (22)$$

which is the first term in eq. (12). Here $\boldsymbol{0}_{p,q}$ denotes a zero matrix of size $p \times q$.

In order to complete the proof, we only need to show that the last sum in (14) can be written as $\sum_{e=(i,j) \in \mathcal{E}_{lc}} \mathrm{tr}\left(\boldsymbol{U}_e \boldsymbol{X}^\mathsf{T}\boldsymbol{X}\boldsymbol{W}_e \boldsymbol{X}^\mathsf{T}\boldsymbol{X}\right)$, *cf.* (12). Towards this goal, we develop each squared norm in the last sum using a derivation similar to (15) and get:

$$\|\boldsymbol{T}_j - \boldsymbol{T}_i \bar{\boldsymbol{T}}_{ij} + \Theta_{ij}^\mathsf{T}(\boldsymbol{T}_j - \boldsymbol{T}_i \bar{\boldsymbol{T}}_{ij})\|^2_{\boldsymbol{\Omega}}$$
$$= \mathrm{tr}\left(\boldsymbol{L}(\mathcal{G}_e)\boldsymbol{T}^\mathsf{T}\boldsymbol{T}\right) + \mathrm{tr}\left(\boldsymbol{L}(\mathcal{G}_e)(\Theta_{ij}^\mathsf{T}\boldsymbol{T})^\mathsf{T}(\Theta_{ij}^\mathsf{T}\boldsymbol{T})\right) \quad (23)$$
$$+ \mathrm{tr}\left(\boldsymbol{L}(\mathcal{G}_e)(\Theta_{ij}^\mathsf{T}\boldsymbol{T})^\mathsf{T}\boldsymbol{T}\right) + \mathrm{tr}\left(\boldsymbol{L}(\mathcal{G}_e)\boldsymbol{T}^\mathsf{T}(\Theta_{ij}^\mathsf{T}\boldsymbol{T})\right)$$

where $\mathcal{G}_e = (\mathcal{V}, \mathcal{E}_{ij})$ denotes a graph with a single edge $e = (i, j)$. We can write $\boldsymbol{T}$ and $\Theta^\mathsf{T}\boldsymbol{T}$ as matrix blocks in $\boldsymbol{X}^\mathsf{T}\boldsymbol{X}$:

$$
\begin{aligned}
\boldsymbol{T} =& [\boldsymbol{0}_{d,(d+1)n+d\ell} \quad \boldsymbol{I}_d]\boldsymbol{X}^\mathsf{T}\boldsymbol{X} \\
& [\boldsymbol{I}_{(d+1)n} \quad \boldsymbol{0}_{((d+1)n,d(\ell+1))}]^\mathsf{T} \\
\Theta_{ij}^\mathsf{T}\boldsymbol{T} =& [\boldsymbol{0}_{d,(d+1)n+d(e-1)} \quad \boldsymbol{I}_d \quad \boldsymbol{0}_{d,d(\ell-e+1)}]\boldsymbol{X}^\mathsf{T}\boldsymbol{X} \\
& [\boldsymbol{I}_{(d+1)n} \quad \boldsymbol{0}_{((d+1)n,d(\ell+1))}]^\mathsf{T}
\end{aligned}
\tag{24}
$$

which enables to write each squared norm in terms of $\boldsymbol{X}^\mathsf{T}\boldsymbol{X}$ as follows:

$$
\|\boldsymbol{T}_j - \boldsymbol{T}_i\bar{\boldsymbol{T}}_{ij} + \Theta_{ij}^\mathsf{T}(\boldsymbol{T}_j - \boldsymbol{T}_i\bar{\boldsymbol{T}}_{ij})\|_{\boldsymbol{\Omega}}^2 = \operatorname{tr}\left(\boldsymbol{U}_e\boldsymbol{X}^\mathsf{T}\boldsymbol{X}\boldsymbol{W}_e\boldsymbol{X}^\mathsf{T}\boldsymbol{X}\right)
\tag{25}
$$

where

$$
\begin{aligned}
\boldsymbol{U}_e =& \begin{bmatrix} \boldsymbol{L}(\mathcal{G}_e) & \boldsymbol{0}_{(d+1)n,d\ell} & \boldsymbol{0}_{(d+1)n,d} \\ \boldsymbol{0}_{d\ell,(d+1)n} & \boldsymbol{0}_{d\ell,d\ell} & \boldsymbol{0}_{d\ell,d} \\ \boldsymbol{0}_{d,(d+1)n} & \boldsymbol{0}_{d,d\ell} & \boldsymbol{0}_{d,d} \end{bmatrix} \\
\boldsymbol{W}_e =& [\boldsymbol{0}_{d,(d+1)n+d(e-1)} \quad \boldsymbol{I}_d \quad \boldsymbol{0}_{d,d(m-e)} \quad \boldsymbol{I}_d]^\mathsf{T} \\
& [\boldsymbol{0}_{d,(d+1)n+d(e-1)} \quad \boldsymbol{I}_d \quad \boldsymbol{0}_{d,d(m-e)} \quad \boldsymbol{I}_d]
\end{aligned}
\tag{26}
$$

Summing over all loop-closure edges results in the second term in eq. (12), concluding the proof.

REFERENCES

[1] A. Blake, P. Kohli, and C. Rother, *Markov Random Fields for Vision and Image Processing*. The MIT Press, 2011.

[2] L. Carlone and F. Dellaert, "Duality-based verification techniques for 2D SLAM," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2015, pp. 4589–4596, (pdf) (code).

[3] L. Carlone, G. Calafiore, C. Tommolillo, and F. Dellaert, "Planar pose graph optimization: Duality, optimal solutions, and verification," *IEEE Trans. Robotics*, vol. 32, no. 3, pp. 545–565, 2016, (pdf) (code).

[4] L. Carlone, D. Rosen, G. Calafiore, J. Leonard, and F. Dellaert, "Lagrangian duality in 3D SLAM: Verification techniques and optimal solutions," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2015, pp. 125–132, (pdf) (code) (datasets: (web)).

[5] L. Carlone and G. Calafiore, "Convex relaxations for pose graph optimization with outliers," *IEEE Robotics and Automation Letters (RA-L)*, vol. 3, no. 2, pp. 1160–1167, 2018, arxiv preprint: 1801.02112, (pdf).

[6] D. Rosen, L. Carlone, A. Bandeira, and J. Leonard, "SE-Sync: A certifiably correct algorithm for synchronization over the Special Euclidean group," in *Intl. Workshop on the Algorithmic Foundations of Robotics (WAFR)*, San Francisco, CA, December 2016, extended arxiv preprint: 1611.00128, (pdf) (pdf) (code).

[7] N. Sünderhauf and P. Protzel, "Switchable constraints for robust pose graph SLAM," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2012.

[8] Y. Latif, C. D. C. Lerma, and J. Neira, "Robust loop closing over time." in *Robotics: Science and Systems (RSS)*, 2012.

[9] P. Agarwal, G. Tipaldi, L. Spinello, C. Stachniss, and W. Burgard, "Robust map optimization using dynamic covariance scaling," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2013.

[10] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A Comparative Study of Energy Minimization Methods for Markov Random Fields with Smoothness-Based Priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 6, pp. 1068–1080, 2008.

[11] J. H. Kappes, B. Andres, F. A. Hamprecht, C. Schnörr, S. Nowozin, D. Batra, S. Kim, B. X. Kausler, T. Kröger, J. Lellmann, N. Komodakis, B. Savchynskyy, and C. Rother, "A Comparative Study of Modern Inference Techniques for Structured Discrete Energy Minimization Problems," *Intl. J. of Computer Vision*, vol. 115, no. 2, pp. 155–184, 2015.

[12] A. Fix and S. Agarwal, "Duality and the continuous graphical model," in *European Conf. on Computer Vision (ECCV)*, 2014, pp. 266–281.

[13] C. Zach and P. Kohli, "A convex discrete-continuous approach for Markov random fields," in *European Conf. on Computer Vision (ECCV)*, 2012, pp. 386–399.

[14] D. Crandall, A. Owens, N. Snavely, and D. Huttenlocher, "SfM with MRFs: Discrete-continuous optimization for large-scale structure from motion," *IEEE Trans. Pattern Anal. Machine Intell.*, 2012.

[15] R. Hartley, J. Trumpf, Y. Dai, and H. Li, "Rotation averaging," *IJCV*, vol. 103, no. 3, pp. 267–305, 2013.

[16] F. Dellaert, "Square Root SAM: Simultaneous location and mapping via square root information smoothing," in *Robotics: Science and Systems (RSS)*, 2005.

[17] P. Huber, *Robust Statistics*. John Wiley & Sons, New York, NY, 1981.

[18] M. Bosse, G. Agamennoni, and I. Gilitschenski, "Robust estimation and applications in robotics," *Foundations and Trends in Robotics*, vol. 4, no. 4, pp. 225–269, 2016.

[19] E. Olson and P. Agarwal, "Inference on networks of mixtures for robust robot mapping," in *Robotics: Science and Systems (RSS)*, July 2012.

[20] J. Casafranca, L. Paz, and P. Piniés, "A back-end ℓ_1 norm based solution for factor graph SLAM," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2013, pp. 17–23.

[21] G. H. Lee, F. Fraundorfer, and M. Pollefeys, "Robust pose-graph loop-closures with expectation-maximization," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2013.

[22] M. Fischler and R. Bolles, "Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography," *Commun. ACM*, vol. 24, pp. 381–395, 1981.

[23] J. Neira and J. Tardós, "Data association in stochastic mapping using the joint compatibility test," *IEEE Trans. Robot. Automat.*, vol. 17, no. 6, pp. 890–897, December 2001.

[24] N. Sunderhauf and P. Protzel, "Towards a robust back-end for pose graph SLAM," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2012, pp. 1254–1261.

[25] L. Carlone, A. Censi, and F. Dellaert, "Selecting good measurements via ℓ_1 relaxation: a convex approach for robust estimation over graphs," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2014.

[26] M. Graham, J. How, and D. Gustafson, "Robust incremental SLAM with consistency-checking," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, Sept 2015, pp. 117–124.

[27] J. Mangelson, D. Dominic, R. Eustice, and R. Vasudevan, "Pairwise consistent measurement set maximization for robust multi-robot map merging," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2018.

[28] M. Pfingsthorn and A. Birk, "Simultaneous localization and mapping with multimodal probability distributions," *Intl. J. of Robotics Research*, vol. 32, no. 2, pp. 143–171, 2013.

[29] ——, "Generalized graph SLAM: Solving local and global ambiguities through multimodal and hyperedge constraints," *Intl. J. of Robotics Research*, vol. 35, no. 6, pp. 601–630, 2016.

[30] S. Bowman, N. Atanasov, K. Daniilidis, and G. Pappas, "Probabilistic data association for semantic slam," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2017, pp. 1722–1729.

[31] F. Dellaert, S. Seitz, C. Thorpe, and S. Thrun, "Structure from motion without correspondence," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, June 2000.

[32] L. Wang and A. Singer, "Exact and stable recovery of rotations for robust synchronization," *Information and Inference: A Journal of the IMA*, vol. 30, 2013.

[33] F. Arrigoni, B. Rossi, P. Fragneto, and A. Fusiello, "Robust synchronization in SO(3) and SE(3) via low-rank and sparse matrix decomposition," *Comput. Vis. Image Underst.*, 2018.

[34] J. Briales and J. Gonzalez-Jimenez, "Cartan-sync: Fast and global SE(d)-synchronization," *IEEE Robot. Autom. Lett*, vol. 2, no. 4, pp. 2127–2134, 2017.

[35] P. Lajoie, S. Hu, G. Beltrame, and L. Carlone, "Modeling perceptual aliasing in SLAM via discrete-continuous graphical models," Tech. Rep., 2018, arXiv preprint: 1810.11692, (pdf).

[36] R. Tron, D. Rosen, and L. Carlone, "On the inclusion of determinant constraints in lagrangian duality for 3D SLAM," in *Robotics: Science and Systems (RSS), Workshop "The problem of mobile sensors: Setting future goals and indicators of progress for SLAM"*, 2015, (pdf).

[37] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming." [Online]. Available: http://cvxr.com/cvx