## Authors' Response to Reviews of

# Decentralized Pose Graph Optimization on Manifolds for Multi-Robot Object-based Collaborative Mapping

Yixian Zhao, Jinming Xu, Yang Xu, Liang Li, Jiming Chen
*IEEE Robotics and Automation Letters,* No.RA-L-24-3622

---

**RC: Reviewer Comment**, AR: Authors' Response, ☐ Manuscript text

Dear Prof. Lucia Pallottino,

We would like to express our sincere gratitude for your efforts devoted to carefully handling our manuscript "*Decentralized Pose Graph Optimization on Manifolds for Multi-Robot Object-based Collaborative Mapping*". In response to the reviewers' constructive comments, we have carefully revised the manuscript and hereby resubmit the revised version for your consideration. We also thank all the reviewers for their precious time and valuable comments which have helped us improve this manuscript.

In what follows, we provide a detailed point-by-point response to the reviewers' comments. For the convenience of the Associate Editor and reviewers, we use "**RC**" to represent the Associate Editor and reviewers' comments highlighted in bold black, and "AR" to represent our response. Besides, for easier reference, we have included the relevant sections of the manuscript, with the revised parts highlighted in blue.

Sincerely yours,

The Authors

**Submission Info:**

1. *Name of the Editor handling the Original Submission*: **Prof. Lucia Pallottino**
2. *Title of Original Submission*: **Decentralized Pose Graph Optimization on Manifolds for Multi-Robot Object-based Collaborative Mapping**
3. *Paper Number of Original Submission*: **RAL-2024-3622**
4. *Original Submission Date*: **2024-10-30**
5. *Original Submission Date that Reviews were returned*: **2024-12-27**
6. *First Resubmission Date*: **2025-1-26**
7. *Last Decision Date*: **2025-2-28**

## 1. Response to Associate Editor

**RC: The paper has improved in the current iteration. However, technical questions persist regarding the alignment of the update directions for the public variables. The transition between global optimization problem and its augmented Lagrangian and the local counterparts is missing. Either show the decomposition that enables the distributed computation or cite a paper where this derivation is performed.**

**Further clarification on object-free vs cluttered settings is required. Lastly, the phrasing of text "in an object-contained space" added in V.B needs improvement. Perhaps, "in a space containing objects" instead. Do to RA-L policy that allows only one "Revise and Resubmit" decision, we invite the authors to revise the paper and possibly submit the revised manuscript as a new submission.**

AR: We thank the Associate Editor as well as the reviewers for their meticulous work. The valuable comments have helped us improve the manuscript. We have endeavored to address all the comments and issues raised by the reviewers carefully. In particular, we have revised the methodology sections to explain more clearly how our method leverages second-order information to track the global Newton direction. In addition, we have also made several minor revisions to improve clarity and overall presentation. A detailed, point-by-point response to the reviewers' comments is provided below.

## 2. Response to Reviewer #7 (No. 389889)

**RC: Follow up on the discussion of Newton tracking and the connection between local and global quadratic approximations: the reviewer acknowledges that the penalty terms in the augmented Lagrangian are intended to pull the $y_i$ values toward consensus. However, even if the $y_i$ values converge to a common point, it is not immediately clear whether the update directions $v_i$ will also reach consensus. For $v_i$ to align, we would need the condition $H_i^{-1}g_i = H_j^{-1}g_j$ to hold for all robot pairs (i, j), but the paper does not explicitly address this aspect. To strengthen the theoretical foundation of the work, it is recommended to include a formal proof to clarify this condition. Such an addition would enhance the reader's understanding and provide greater confidence in the proposed approach.**

AR: We thank the reviewer again for the insightful and valuable comment. We acknowledge that the consensus of the $y_i$ values does not immediately imply that the update directions $v_i$ will also converge to a common point. As a result, following the reviewer's suggestion, we have now provided a relaxed but rigorous convergence analysis *under standard convexity assumptions*, demonstrating that once the $y_i$ values converge, all the update directions $v_i$ will indeed converge to zero, implying the consensus of $v_i$. Moreover, to illustrate this in a more intutive way, we have conducted additional experimental validation (c.f., Fig 2.1) to show that the consistency error among Newton direction will be vanishing. While we believe these above results are sufficient to establish the algorithm's validity for real-world applications, we remark that a more stringent theoretical proof in general non-convex settings remains one of our important ongoing works.

The specific theoretical analysis and experimental results are provided below:

**[Theoretical analysis under a simplified setting]**

To show that the update directions $v_i$ will eventually align and thus mimic Newton tracking, we conduct a rigorous analysis which constitute two stages as follows. First, by examining the primal-dual dynamics under standard convexity assumptions, we prove asymptotic convergence to the limit points. Second, we perform a thorough limit point analysis, which shows that both quantities converge to a consistent equilibrium point that satisfies the global optimality conditions, i.e., both $y_i$ and the update directions $v_i$ converging to certain consensus value eventually (note that $\boldsymbol{v}$ vanishes as well). The detailed derivation are given as below.

Fist of all, for notational convenience, let $z_i := [x_i, y_i]$ denote the concatenation of private and public variables of robot $i$. The consistency constraint for public variables between any two robots $i \neq j$ can be expressed as $\varphi(A_i z_i, A_j z_j)$, where matrices $\{A_i\}_{i\in[N]}$ extract the public components (Without loss of generality, we omit these matrices in subsequent analysis for simplicity). Let $\boldsymbol{z} := [z_1, \dots, z_N]$ denote the stacked vector of all robot variables, and $\boldsymbol{\lambda} := [\lambda_{ij}]_{i \neq j \in [N]}$ represent all dual variables. As a result, with $f(\boldsymbol{z}) = \sum_{i=1}^{N} f_i(\boldsymbol{z}_i)$, the global Lagrange function becomes

$$L(\boldsymbol{z}, \lambda) = f(\boldsymbol{z}) + \sum_{i=1}^{N} \left( \sum_{j \in \mathcal{N}_i} \langle \lambda_{ij}, w_{ij} \varphi(z_i, z_j) \rangle + \frac{\beta}{2} \| w_{ij} \varphi(z_i, z_j) \|^2 \right). \tag{2.1}$$

**STEP 1: Convergence to certain cluster point**

For analytical simplicity, let the consistency constraint $\varphi(\cdot)$ be defined with $\varphi(z_i, z_j) := z_i - z_j$ . Then, the primal-dual update process becomes

$$\boldsymbol{z}^{k+1} = \underset{\boldsymbol{z}_i, x_i}{\arg\min}\, L\left(\boldsymbol{z}^k, \boldsymbol{\lambda}^k\right), \tag{2.2}$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \beta D_W B \boldsymbol{z}^{k+1}. \tag{2.3}$$

with

$$L\left(\boldsymbol{z}^k, \boldsymbol{\lambda}^k\right) = f\left(\boldsymbol{z}^k\right) + \left\langle \boldsymbol{\lambda}^k, D_W B \boldsymbol{z}^k \right\rangle + \frac{\beta}{2} \boldsymbol{z}^k (D_W B)^\top D_W B \boldsymbol{z}^k. \tag{2.4}$$

where $B$ is the incidence matrix and $D_W$ is a diagonal matrix with weights $w_{ij}$ as the diagonal entries.

According to first-order optimality condition, we have

$$\begin{aligned} &\nabla f\left(\boldsymbol{z}^{k+1}\right) + (D_W B)^\top \boldsymbol{\lambda}^k + \beta (D_W B)^\top D_W B \boldsymbol{z}^{k+1} = 0 \\ &\Leftrightarrow \nabla f\left(\boldsymbol{z}^{k+1}\right) + (D_W B)^\top \left(\boldsymbol{\lambda}^k + \beta D_W B \boldsymbol{z}^{k+1}\right) = 0 \\ &\Leftrightarrow \nabla f\left(\boldsymbol{z}^{k+1}\right) + (D_W B)^\top \boldsymbol{\lambda}^{k+1} = 0. \end{aligned} \tag{2.5}$$

where we have used the dual update in Eq. (2.3).

Let $(\boldsymbol{z}^*, \lambda^*)$ with $\boldsymbol{z}^* = 1 \cdot z^*$ be a saddle point. Then, we have

$$\nabla f(\boldsymbol{z}^*) + (D_W B)^\top \boldsymbol{\lambda}^* + \beta (D_W B)^\top D_W B \boldsymbol{z}^* = 0. \tag{2.6}$$

Since $B\boldsymbol{z}^* = 0$, we further have

$$\nabla f(\boldsymbol{z}^*) + (D_W B)^\top \boldsymbol{\lambda}^* = 0. \tag{2.7}$$

Besides, due to the convexity of $f(\boldsymbol{z})$, we obtain

$$\left\langle \nabla f\left(\boldsymbol{z}^{k+1}\right) - \nabla f(\boldsymbol{z}^*), \boldsymbol{z}^{k+1} - \boldsymbol{z}^* \right\rangle \geq 0. \tag{2.8}$$

Then, substituting (2.5) and (2.7) into (2.8), we have

$$\begin{aligned} &\left\langle (D_W B)^\top \left(\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\right), \boldsymbol{z}^{k+1} - \boldsymbol{z}^* \right\rangle \\ &= \left\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, D_W B \left(\boldsymbol{z}^{k+1} - \boldsymbol{z}^*\right) \right\rangle \\ &= \left\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, D_W B \boldsymbol{z}^{k+1} \right\rangle \\ &= \left\langle \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*, \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^k \right\rangle \leq 0. \end{aligned} \tag{2.9}$$

where we have used the dual update in Eq. (2.3).

Knowing the fact that $2\langle a,b\rangle = \|a\|^2 + \|b\|^2 - \|a-b\|^2$, letting $a := \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*$, $b := \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^k$ we have

$$\left\|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\right\|^2 - \left\|\boldsymbol{\lambda}^{k} - \boldsymbol{\lambda}^*\right\|^2 \leq - \left\|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^k\right\|^2. \tag{2.10}$$

According to the standard limit point analysis [A1], the sequence of variables$(\boldsymbol{z}^k, \boldsymbol{\lambda}^k)$ will gradually converge to a certain limit point $(\boldsymbol{z}^*, \boldsymbol{\lambda}^*)$ . It follows that the limit $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\lambda}^*)$ is a fixed point of the iteration.

[A1] *Chambolle, Antonin, and Thomas Pock. "A first-order primal-dual algorithm for convex problems with applications to imaging." Journal of mathematical imaging and vision 40 (2011): 120-145.*

**STEP 2: Limit point analysis**

As $k \to \infty$, we analyze the algorithm update procedure of robot $i$ as follows,

$$y_i^\infty = Retr_{y_i^\infty}(-\alpha v_i^\infty), \tag{2.11}$$

$$x_i^\infty = Retr_{x_i^\infty}(-\gamma u_i^\infty), \tag{2.12}$$

$$\lambda_{ij}^\infty = \lambda_{ij}^\infty + \beta w_{ij} \varphi^2\left(y_i^\infty, y_j^\infty\right). \tag{2.13}$$

From Eq. (2.13), we have

$$\varphi^2\left(y_i^\infty, y_j^\infty\right) = 0. \tag{2.14}$$

It suggests that the public variables of all robots can converge on the consensus.

Similarly, for Eq. (2.11) when $k \to \infty$, we have

$$v_i^\infty = -\frac{1}{\alpha} Retr_{y_i^\infty}^{-1}(y_i^\infty) = \mathbf{0}, \tag{2.15}$$

which implies that the update directinos $v_i$ eventually converge to zero. Thus, we complete the proof.

In Euclidean settings, Zhang et al.[25] demonstrated that distributed Newton tracking achieves linear convergence by approximating the global Newton direction through local updates despite heterogeneity in $H_i$ and $g_i$. However, formal proof for the Riemannian manifold case with general residual function $\varphi$ is more challenging due to non-linear geometry (e.g., curvature, retractions), which is left for our ongoing work.

**[Experimental Validation]**

To validate that those $v_i$ directions achieve practical consensus, we define the metrics of Newton direction consistency error (NDCE) as

$$NDCE = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i} \parallel v_i - v_j \parallel^2 \tag{2.16}$$

where we normalize the $v_i \leftarrow \frac{v_i}{\|v_i\|}$ to prevent the effects of the scale. We experimented with five trials of arbitrarily generated communication topologies with connectivity of $0.5$ in the Dingo dataset. The normalized discrepancy norms (NDCE) were computed for each trial by scaling the initial NDCE value to unity, enabling direct comparison of convergence rates despite differing initial magnitudes. As shown in Fig. 2.1, all normalized NDCE trajectories exhibit decay toward zero, aligning with our theoretical convergence prediction. This behavior persists across diverse network topologies, demonstrating the algorithm's robustness to initial disparities. Low $NDCE$ in Fig. 2.1 confirm that $v_i$ updates to a consistent and accurate solution.

However, due to the limited space of RA-L, we have made the following revisions in Section IV. A. to strengthen the theoretical intuition:

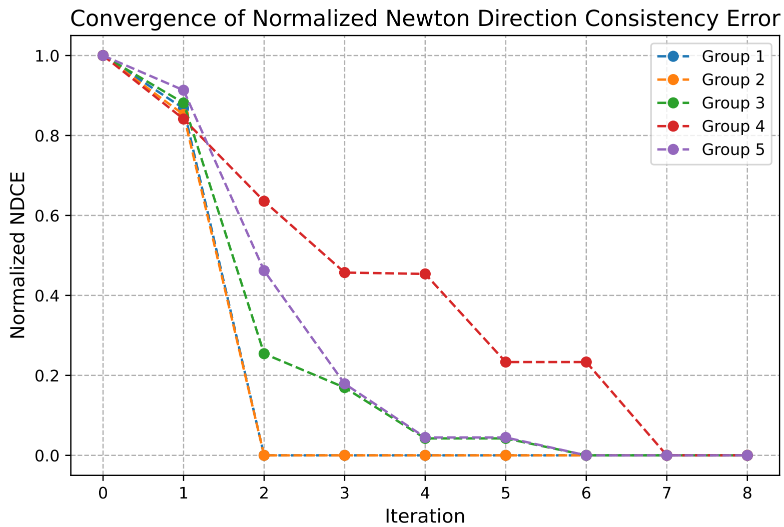


Figure 2.1: The consensus error of Newton's direction

> **(Section IV:) Remark 3.** *The decentralized Riemannian approximate Newton method achieves faster convergence for distributed optimization problems with heterogeneous local objective functions by utilizing local second-order information. As consistency constraints ensure that public variables $y$ reach consensus, the local updates directions of the public variables (c.f., Line 11 in Alg.1) asymptotically converge to zero and will approximate the global Newton gradient direction when they reach consensus enough, mitigating the effects of heterogeneity. The theoretical analysis is provided in [].…*

**RC: As a related comment, in the current method section, the paper directly starts with defining the augmented Lagrangian for each local cost fi(xi, yi). The reviewer finds this to be unconventional, because typically the Lagrangian is first defined for the global optimization problem (Problem 2), which is then decomposed in some way so that the resulting updates can be performed distributedly. See for example Chapter 7 of Boyd, Stephen, et al. "Distributed optimization and statistical learning via the alternating direction method of multipliers." Foundations and Trends in Machine learning 3.1 (2011): 1-122.**

AR: We sincerely appreciate the reviewer's insightful observation regarding the Lagrangian formulation. Our original presentation intentionally emphasized the decentralized structure to better reflect real-world robotic constraints, where each robot operates independently using only local data and neighboring communications. The local augmented Lagrangian for robot $i$ inherently enforces global consensus through interactions with neighboring robots. The dual variables $\lambda_{ij}$ and quadratic penalty terms collectively approximate the global Lagrangian when aggregated across all robots, resembling the edge-based consensus framework in Boyd et al. [1]. We recognize the value of explicitly connecting our approach to classical ADMM frameworks. To avoid any potential ambiguity, we have restructured Section III.B as follows:

(**Section IV:**) Firstly, we construct an augmented Lagrangian for global cost function $f(\boldsymbol{x}, \boldsymbol{y})$ as follows:

$$
\begin{aligned}
L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda}) = f(\boldsymbol{x}, \boldsymbol{y}) &+ \sum_{i=1}^{n} \sum_{j \in \mathcal{N}_i} \langle \lambda_{ij}, w_{ij}\varphi(y_i, y_j) \rangle \\
&+ \sum_{i=1}^{n} \sum_{j \in \mathcal{N}_i} \frac{\beta}{2} \| w_{ij}\varphi(y_i, y_j) \|^2 .
\end{aligned} \tag{xxx}
$$

where $\lambda_i = [\lambda_{ij}]_{j\in\mathcal{N}_i}$ denotes the dual variable associated with robot $i$, $\boldsymbol{\lambda} := [\lambda_i]_{i\in[N]}$ is the stack vector of all dual variables and $\beta$ is the penalty coefficient. The weights $w_{ij}$ are introduced in the augmented Lagrangian function to modulate the influence of the consistency constraint between neighboring robots $i$ and $j$. The process of solving the augmented Lagrangian in Eq.(4) can be expressed as:

$$
y_i^{k+1}, x_i^{k+1} = \arg\min_{y_i, x_i} L\left(\boldsymbol{x}^k, \boldsymbol{y}^k, \boldsymbol{\lambda}^k\right), \tag{2.17a}
$$

$$
\lambda_{ij}^{k+1} = \lambda_{ij}^{k} + \beta w_{ij}\varphi\left(y_i^{k+1}, y_j^{k+1}\right), \tag{2.17b}
$$

where $k$ is the iteration index in the optimization process.

**RC:** **Following up on the experiments with datasets that do not contain objects: based on the authors' response, the reviewer understands that the proposed method directly communicates separators rather than performing consensus on them. However, it is unclear how the method operates in this case, as the consensus variables yi would presumably be empty. Clarifying this point would be important to improve the reader's understanding. Additionally, it is recommended that the paper provide a detailed explanation early on regarding how cases with and without objects are handled, to avoid potential confusion.**

AR: Thank you for the valuable comment. We remark that our method is primarily designed for decentralized multi-robot collaborative optimization involving public objects, where all robots jointly estimate object poses, and it seamlessly extends to object-free cases under our framework. In the following, we provide a comprehensive explanation of how the method works with decentralized optimization.

When no objects exist in the environment, the consensus variables $\{y_i\}$ are indeed absent. In this case, the coupling between robots arises solely from **inter-robot loop closure constraints**, which are explicitly modeled through separators in the robots' trajectories, as we mentioned in Section III. Specifically:

$$
f_i(x_i) = f_i^{\text{intra}}(x_i) + f_i^{\text{inter}}(x_i), \tag{2.18}
$$

During the communication phase, robots exchange their separator poses $x_i$ (not full trajectories) with neighbors. For the iteration $k$ of the robot $i$, the quadratic approximation to this objective function is made to solve for the second order direction $u_i^k$ of $x_i$.

$$
m_i(u_i) = f_i(x_i) + \langle g_{ix}, u_i \rangle + \frac{1}{2}\langle u_i, M_i u_i \rangle \tag{2.19}
$$

Further, update $x_i$ by $x_i^{k+1} = Retr_{x_i^k}(-\gamma u_i^{k+1})$ (c.f., Line 11 in Alg.1).

Empirically, experiments on G2O benchmark datasets (Table I) demonstrate robust performance in object-free environments. To clarify this, we have now added the following paragraph in Section IV.

**Section IV**: In object-free cases, inter-robot loop closures enforce relative positional relationships between robots in the global coordinate system via separators, preserving the same optimization/communication framework as object-based scenarios, thereby ensuring methodological parity.