ISPRS Journal of Photogrammetry and Remote Sensing 242 (2026) 53–73

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

DUAG-C: Decentralized uncertainty-aware Gaussian consensus for multi-robot dense mapping

Cabrel Wouladje [b], Golden Tendekai Mumanikidzwa [b], Huiying Xu [a,b], Hongbo Li [d], Yanchao Wang [b], Mantang Liu [d], Wenjie Wang [c], Wenzhe Tan [d], Zhendong Chen [b], Longfei Wang [b], Kangjia Dai [b], Zhenglong Wang [d], Xinzhong Zhu [a,b,d,*]

[a] *Zhejiang Key Laboratory of Intelligent Education and Application, Zhejiang Normal University, Jinhua, 321004, Zhejiang, China*
[b] *School of Computer and Technology, Zhejiang Normal University, Jinhua, 321004, China*
[c] *Shenzhen QiXuan Technology Co., Ltd, China*
[d] *Beijing Geekplus Technology Co., Ltd, China*

A R T I C L E I N F O

Keywords:
Decentralized SLAM
3D Gaussian splatting
Multi-robot dense mapping
Fisher information weighting
Riemannian consensus optimization

A B S T R A C T

Dense multi-robot Simultaneous Localization and Mapping (SLAM) has, to date, been realized exclusively through centralized architectures, precluding deployment in infrastructure-free or adversarial environments. The obstacle is representational, not bandwidth-related: 3D Gaussian primitives are constrained heterogeneously by each agent's local observations, and naive peer-to-peer averaging allows poorly-determined Gaussians from one agent to corrupt the well-constrained map of another. Resolving this requires a per-primitive epistemic uncertainty measure that is data-driven, manifold-compatible, and amenable to distributed optimization with provable guarantees. We present **DUAG-C**, a fully peer-to-peer dense SLAM system built on volumetric 3D Gaussian Splatting (3DGS). The central contribution is a per-Gaussian diagonal Fisher Information Matrix (FIM), derived via the Laplace approximation to the photometric loss and accumulated from gradients already produced during standard 3DGS optimization. These weights drive a Riemannian Alternating Direction Method of Multipliers (ADMM) consensus on the joint product manifold $\mathrm{SE}(3)^{|\mathcal{V}|} \times (\mathbb{S}^3_{++} \times \mathbb{R}^7)^{N_G}$, with closed-form updates under the Log-Euclidean metric. We prove convergence to an $\mathcal{O}(\varepsilon)$ neighborhood of the global optimum, establish for a scalar-weight surrogate that higher FIM contrast strictly accelerates convergence, and prove a dimension-wise contraction bound that governs the implemented per-dimension algorithm directly—to our knowledge the first convergence guarantees for decentralized *dense* mapping, where existing certifiable results address sparse pose-graph optimization alone. Communication priority emerges naturally: transmitting high-FIM Gaussians first maximizes information delivered per byte without an explicit rate-control layer. Across four benchmarks spanning synthetic indoor, real-sensor, and outdoor Unmanned Ground Vehicle (UGV) settings with 2–8 robots, DUAG-C attains 0.066 cm Absolute Trajectory Error (ATE) averaged over all four ReplicaMultiagent scenes—surpassing MAC-Ego3D by 2.1× and GRAND-SLAM by 3.8× without any central infrastructure—and achieves the best reported geometric fidelity (Learned Perceptual Image Patch Similarity, LPIPS, 0.043; Depth L1 0.36 cm). An Expected Calibration Error (ECE) of 0.032 confirms the FIM weights encode genuine per-Gaussian reliability. The code is available at https://github.com/HZAI-ZJNU/DUAG-SLAM.

1. Introduction

Autonomous robot teams operating in infrastructure-free environments require a shared, dense model of the surrounding world. A photorealistic volumetric map supports appearance-based loop closure, occlusion-accurate collision checking, and novel-view synthesis for human supervision in ways that sparse point clouds cannot. Collaborative SLAM (C-SLAM) (Lajoie et al., 2022) provides the algorithmic foundation: agents share observations, resolve inter-agent drift through distributed loop closure, and build a globally consistent reconstruction that no single agent could achieve alone. Critically, many deployment scenarios—GPS-denied subterranean exploration, post-disaster search-and-rescue, adversarial environments—make a persistent central server

* Corresponding author at: Zhejiang Key Laboratory of Intelligent Education and Application, Zhejiang Normal University, Jinhua, 321004, Zhejiang, China.
E-mail address: zxz@zjnu.edu.cn (X. Zhu).

https://doi.org/10.1016/j.isprsjprs.2026.08.030
Received 28 April 2026; Received in revised form 20 August 2026; Accepted 23 August 2026
Available online 4 September 2026
0924-2716/© 2026 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved, including those for text and data mining, AI training, and similar technologies.

impractical: it represents a single point of failure, introduces communication bottlenecks under intermittent connectivity, and requires pre-existing network infrastructure that field robots cannot assume. Decentralized operation, where every robot is both a peer and a full participant in map fusion, is therefore not an engineering preference but an operational necessity.

The introduction of 3D Gaussian splatting (3DGS) (Kerbl et al., 2023) transformed dense scene representation. Parameterizing a scene as a collection of anisotropic Gaussians—each with mean $\boldsymbol{\mu}$, covariance $\boldsymbol{\Sigma} \in \mathbb{S}^3_{++}$, view-dependent color $\mathbf{c}$, and opacity α—enables real-time photorealistic rendering through a differentiable rasterizer. MonoGS (Matsuki et al., 2024) demonstrated that 3DGS suffices as the sole map primitive in a live SLAM system, unifying tracking and mapping in a single gradient-based loop. Concurrent work has refined per-Gaussian uncertainty within single-agent 3DGS SLAM (Tran and Košecká, 2026) and addressed large-scale loop closure (Thomas et al., 2025).

Multi-agent 3DGS SLAM has, however, been overwhelmingly realized through *centralized* architectures. CP-SLAM (Hu et al., 2023) introduced collaborative dense mapping via a dedicated server. MAGiC-SLAM (Yugay et al., 2025) replaced the implicit map with 3DGS submaps while retaining server-side loop closure and global refinement. MAC-Ego3D (Xu et al., 2025) jointly optimizes poses and a server-hosted Gaussian map via asynchronous photometric–geometric consensus. HAMMER (Yu et al., 2025) extends the paradigm to heterogeneous robot teams through server-side semantic 3DGS fusion. More recently, Coko-SLAM (M.M. Li et al., 2026) reduces server-bound communication overhead by up to 95% through keyframe sparsification and Gaussian compaction, yet explicitly defers full decentralization to future work. On the decentralized side, DPGO (Tian et al., 2021), Kimera-Multi (Tian et al., 2022), and Swarm-SLAM (Lajoie and Beltrame, 2024) achieve fully peer-to-peer operation without any central coordinator, but are restricted to *sparse* map representations that cannot support dense scene queries. The most recent concurrent work, CoMA-SLAM (Chen et al., 2026), takes a meaningful step toward decentralized dense SLAM using 2D Gaussian surfels with bandwidth-efficient submap exchange, reducing communication overhead by 99.8% relative to centralized approaches; however, it provides neither an uncertainty-weighted fusion rule nor a convergence guarantee for its distributed optimizer, and its evaluation is confined to indoor environments.

Table 1 formalizes the resulting gap: *no published system simultaneously achieves decentralized operation, volumetric dense reconstruction, epistemic uncertainty weighting, and a proven convergence guarantee.* The root obstacle is representational. In sparse SLAM, all landmarks are drawn from the same well-defined estimation process and may be averaged uniformly. In 3DGS, each Gaussian is constrained exclusively by one agent's local observations: a Gaussian in a densely revisited corridor is tightly determined, while a Gaussian at the trajectory periphery—observed once under severe foreshortening—may have poorly constrained parameters. Naive uniform averaging in a peer-to-peer exchange corrupts the consensus, as poorly constrained Gaussians from one agent contaminate the well-determined map of another. Resolving this requires a principled, per-Gaussian epistemic uncertainty measure that is simultaneously data-driven, manifold-compatible, and amenable to distributed optimization with provable guarantees.

Our approach. We introduce **DUAG-C**—**D**ecentralized **U**ncertainty-**A**ware **G**aussian **C**onsensus SLAM—the first fully peer-to-peer multi-robot dense SLAM system built on volumetric 3DGS to close all four gaps in Table 1 simultaneously: decentralized operation, volumetric dense reconstruction, epistemic uncertainty weighting, and a proven convergence guarantee. The core innovation is a diagonal Fisher Information Matrix (FIM) computed per Gaussian via the Laplace approximation (MacKay, 1992) applied to the MonoGS photometric loss. Concretely, for each Gaussian we accumulate the squared photometric gradients contributed by the pixels it renders, aggregated over a rolling keyframe window; the resulting per-Gaussian precision Λ_k is large where a primitive is tightly constrained by local photometry and small where it is poorly determined. The precise diagonal form and its derivation via the Laplace approximation are deferred to Section 3.1 (Eq. (2)), where they are developed in full. Crucially, Λ_k is computed entirely from each agent's local rendering loop—reusing gradients already computed during standard 3DGS optimization—requiring no inter-agent communication, no shared server, and no modification to the single-agent SLAM frontend.

We formulate inter-robot consensus as FIM-weighted optimization over the product manifold $\mathrm{SE}(3) \times (\mathbb{S}^3_{++} \times \mathbb{R}^7)^{N_G}$, where N_G is the number of matched Gaussian pairs and $\mathbb{S}^3_{++}$ carries the covariance under the Log-Euclidean metric (Arsigny et al., 2006)—the unique metric that renders the Gaussian subproblem geodesically convex and yields a closed-form FIM-weighted update. A Riemannian ADMM algorithm (Boyd et al., 2011; Tian et al., 2021) alternates between a pose update on SE(3) (extending DPGO) and a closed-form Gaussian update, requiring only pairwise peer-to-peer message exchange per iteration: each robot transmits its matched Gaussians $\mathcal{G}$, poses $\mathbf{T}$, and FIM weights $\boldsymbol{\Lambda}$—nothing more.

We prove two theoretical results. First, DUAG-C converges to within $\mathcal{O}(\epsilon)$ of the global optimum, where ϵ bounds the Gaussian association error, under standard graph-connectivity and noise-boundedness conditions. Second, for a scalar-weight surrogate of the algorithm, the convergence *rate* increases monotonically with the FIM contrast ratio $\max_k(\Lambda_k)/\min_k(\Lambda_k)$: a formal guarantee that accurate epistemic uncertainty *provably accelerates* the distributed optimizer, not merely improves its converged solution. The gap between this scalar surrogate and the full per-dimension implementation is discussed in Section 3.3. Communication awareness emerges as a structural consequence: because high-Λ_k Gaussians are geometrically more informative, transmitting them first maximizes Fisher information delivered per byte without any explicit rate-control layer.

Empirically, DUAG-C is the only fully decentralized system evaluated across all comparisons on both indoor and outdoor benchmarks. On ReplicaMultiagent, it surpasses every centralized dense baseline on tracking and geometric rendering metrics. On AriaMultiagent, it achieves the best three-agent overall average (0.351 cm); the concurrent decentralized system CoMA-SLAM achieves 0.38 cm overall average, and 0.38 cm on Room0 vs. DUAG-C's 0.389 cm, attributable to Agent 3's limited inter-agent overlap (Section 4.3).

Contributions. This paper makes four contributions:

1. **FIM-based per-Gaussian uncertainty for decentralized consensus** (Section 3.1). We derive the first diagonal Fisher Information Matrix applied to weight individual Gaussian parameters in a distributed multi-robot consensus framework. Unlike concurrent single-agent work (Tran and Košecká, 2026) that exploits per-Gaussian uncertainty solely for tracking robustness within a single robot, our FIM quantifies and resolves *inter-agent incommensurability*: the fundamental problem that arises when two robots fuse Gaussians observed under incompatible coverage and viewpoint conditions. The FIM reuses gradients already computed during standard 3DGS optimization, requiring no architectural modification to the single-agent frontend and negligible additional computation. A block-diagonal comparison (Section 3.1.6) confirms the off-diagonal FIM energy is negligible (median $r_k = 0.04$), so the diagonal approximation incurs no measurable accuracy loss (Section 3.1.3).
2. **Decentralized consensus on the joint pose–Gaussian product manifold** (Section 3.2). We present the first decentralized formulation that treats robot poses and dense Gaussian map parameters *jointly* on their natural Riemannian domains—SE(3) for poses and $\mathbb{S}^3_{++}$ under the Log-Euclidean metric for covariances—with closed-form FIM-weighted updates and strictly peer-to-peer message passing. All prior dense multi-agent systems (Hu et al., 2023; Yugay et al., 2025; Xu et al., 2025; Yu et al., 2025; M.M. Li et al., 2026) require a central server for at least one phase

Table 1
Qualitative comparison of representative SLAM systems, ordered chronologically within each paradigm. **Paradigm**: SA = single-agent; MA-C = multi-agent centralized; MA-D = multi-agent decentralized. **Dense**: photorealistic 3D map (3DGS or Neural Radiance Field, NeRF). **Decentr.**: fully peer-to-peer, no central server required. **Comm.-Aware**: explicit handling of bandwidth constraints or packet loss. **Uncert.-Wtd**: per-primitive epistemic uncertainty drives map fusion. **Conv. Guar.**: published convergence proof for the distributed optimizer. **Outdoor**: demonstrated on real outdoor robot trajectories. Columns marked *N/A* are not applicable to single-agent systems.

Method	Year	Paradigm	Dense	Decentr.	Comm.-Aware	Uncert.-Wtd	Conv. Guar.	Outdoor
MonoGS (Matsuki et al., 2024)	2024	SA	✓	*N/A*	*N/A*	×	*N/A*	×
VarSplat (Tran and Košecká, 2026)	2026	SA	✓	*N/A*	*N/A*	✓	*N/A*	×
CCM-SLAM (Schmuck and Chli, 2019)	2019	MA-C	×	×	×	×	×	×
ORB-SLAM3 (Campos et al., 2021)	2021	MA-C	×	×	×	×	×	✓
CP-SLAM (Hu et al., 2023)	2023	MA-C	✓	×	×	×	×	×
MAGiC-SLAM (Yugay et al., 2025)	2025	MA-C	✓	×	×	×	×	×
MAC-Ego3D (Xu et al., 2025)	2025	MA-C	✓	×	×	×	×	×
GRAND-SLAM (Thomas et al., 2025)	2025	MA-C	✓	×	×	×	×	✓
HAMMER (Yu et al., 2025)	2025	MA-C	✓	×	×	×	×	×
MANG-SLAM (Y. Li et al., 2026)	2026	MA-C	✓	×	×	×	×	×
Coko-SLAM (M.M. Li et al., 2026)	2026	MA-C	✓	×	✓	×	×	×
DPGO (Tian et al., 2021)	2021	MA-D	×	✓	×	×	✓	×
Kimera-Multi (Tian et al., 2022)	2022	MA-D	×	✓	×	×	✓	✓
Swarm-SLAM (Lajoie and Beltrame, 2024)	2024	MA-D	×	✓	[a]	×	×	✓
CoMA-SLAM (Chen et al., 2026)	2026	MA-D	✓	✓	[a]	×	×	×
DUAG-C (Ours)	2026	MA-D	✓	✓	✓	✓	✓	✓

[a] Partial (bandwidth compression but no adaptive rate control or packet-loss robustness).

of map fusion. The closest concurrent work, CoMA-SLAM (Chen et al., 2026), removes the server but operates on 2D Gaussian surfels without a Riemannian product-manifold optimizer or uncertainty weighting; DUAG-C eliminates the server entirely while operating on full volumetric 3DGS primitives with principled per-dimension uncertainty weighting.

3. **Convergence guarantee for FIM-weighted Riemannian ADMM** (Section 3.3). We prove $\mathcal{O}(\varepsilon)$ convergence to the global optimum, establish a monotone rate bound tied to the FIM contrast ratio for a scalar-weight surrogate of the algorithm, and prove a dimension-wise contraction bound (Theorem 8) that governs the implemented per-dimension algorithm directly, providing the first theoretical foundation for uncertainty-weighted consensus in decentralized 3DGS SLAM. The relationship between the scalar surrogate and the full per-dimension implementation is made precise by Theorem 8 and Remark 7. No prior decentralized dense SLAM system, including CoMA-SLAM (Chen et al., 2026), provides any convergence analysis.
4. **A decentralized dense SLAM system that surpasses centralized baselines, evaluated across four datasets** (Section 4). DUAG-C is evaluated on ReplicaMultiagent, AriaMultiagent, S3E, and Kimera-Multi Campus with 2–6 robots, spanning synthetic indoor, real-world indoor, and real outdoor robot trajectories under both ideal and degraded (up to 90% packet-loss) communication. On ReplicaMultiagent, DUAG-C achieves 0.066 cm ATE across all four scenes, all directly measured under unified settings (Section 4.2.2), outperforming MAC-Ego3D (Xu et al., 2025) by 2.1× and GRAND-SLAM (Thomas et al., 2025) by 3.8× without any central infrastructure. On Kimera-Multi Campus, DUAG-C at $N = 6$ achieves 6.8× lower ATE than GRAND-SLAM's sole published configuration ($N = 3$). Among decentralized dense baselines, CoMA-SLAM (Chen et al., 2026) achieves competitive tracking on ReplicaMultiagent; DUAG-C outperforms it on all measured Replica scenes and on LPIPS and Depth L1 under standard online evaluation despite CoMA-SLAM's post-hoc finetuning advantage (Section 4.2.3). That a fully decentralized system outperforms centralized baselines on their own benchmarks is, together with CoMA-SLAM's concurrent progress, evidence that decentralized dense SLAM has crossed a critical accuracy threshold. A centralized-oracle variant of our own method (Section 4.7.4, hypothesis H5) isolates the accuracy cost of decentralization itself from the contribution of FIM weighting, and a weighting-proxy ablation (Section 4.7.5, H6) tests whether the Fisher construction is necessary or a cheap observation-count proxy suffices.

Organization. Section 2 situates DUAG-C within three bodies of literature: single-agent 3DGS SLAM, centralized multi-agent dense SLAM, and decentralized multi-robot SLAM. Section 3 develops the technical system: FIM derivation (Section 3.1), product-manifold consensus (Section 3.2), Riemannian ADMM and convergence (Section 3.3), and heterogeneous sensor fusion (Section 3.4). Section 4 reports experimental results. Section 5 discusses limitations and future directions. Section 6 concludes.

2. Related work

The literature relevant to DUAG-C spans three paradigms: single-agent dense 3DGS SLAM (Section 2.1), centralized multi-agent dense SLAM (Section 2.2), and decentralized multi-robot SLAM (Section 2.3).

2.1. Single-agent dense 3DGS SLAM

Dense visual SLAM in the neural era originated with iMAP (Sucar et al., 2021), which demonstrated that a multilayer perceptron trained online against an RGB-D (color-plus-depth) stream could simultaneously produce occupancy-and-color maps and serve camera tracking. The fundamental scalability limitation of a monolithic network was addressed by NICE-SLAM (Zhu et al., 2022), whose hierarchical feature-grid encoding localizes scene geometry spatially, enabling stable reconstruction of large multi-room scenes. Point-SLAM (Sandström et al., 2023) replaced fixed grids with an adaptively sampled point cloud of neural features, concentrating representational capacity in information-dense regions. These Neural Radiance Field (Mildenhall et al., 2020) systems established the render-loss paradigm that motivates DUAG-C's FIM derivation—scene parameters are optimized so that rendering error against observed RGB-D frames is minimized—but their volumetric rendering pipelines are too slow for real-time deployment and cannot be efficiently serialized for inter-robot message exchange.

3D Gaussian Splatting resolved both bottlenecks. Concurrently with MonoGS (Matsuki et al., 2024), which DUAG-C adopts as its per-robot front-end, several systems explored the same explicit representation under different design choices. GS-SLAM (Yan et al., 2024) introduced an adaptive Gaussian expansion strategy and a coarse-to-fine camera tracking scheme that selects geometrically stable Gaussians for pose optimization, reducing runtime without sacrificing accuracy. SplaTAM (Keetha et al., 2024) exploited a per-Gaussian silhouette mask to distinguish observed from unobserved scene regions, enabling structured map growth and achieving competitive tracking on Replica and TUM RGB-D. Photo-SLAM (Huang et al., 2024) coupled the mature ORB-SLAM3 feature tracker with a 3DGS mapper, demonstrating that the robustness of classical pose estimation can be preserved while gaining photorealistic reconstruction. Gaussian-SLAM (Yugay et al., 2023) managed active and inactive 3DGS submaps via a DROID-SLAM tracker; this submap structure was subsequently extended to the multi-agent setting in MAGiC-SLAM (Yugay et al., 2025).

Single-agent loop closure for 3DGS maps was addressed by Loop-Splat (Zhu et al., 2025), which detects revisited locations and derives relative pose constraints by directly registering two 3DGS submaps against each other—a 3DGS-native alternative to ICP on point clouds. LoopSplat's submap registration is conceptually related to the inter-agent Gaussian matching in DUAG-C, though LoopSplat operates entirely within the single-agent regime without distributed optimization or uncertainty weighting.

Per-Gaussian uncertainty in single-agent 3DGS SLAM has been explored by two concurrent works. CG-SLAM (Hu et al., 2024) attaches a depth uncertainty channel to each Gaussian and uses it to select informative primitives for pose optimization, improving tracking efficiency under sparse keyframe sampling. VarSplat (Tran and Košecká, 2026) introduces a broader uncertainty model for robust tracking under viewpoint sparsity. Both contributions confirm the importance of per-Gaussian epistemic signals, but they apply uncertainty exclusively to *single-agent tracking filtering*, not to *distributed consensus weighting*: neither derives uncertainty via the Fisher Information Matrix, neither proves a connection between uncertainty estimation and distributed optimization convergence, and neither extends to multiple robots. DUAG-C derives the diagonal FIM from the photometric loss across the full Gaussian parameter vector θ_k, applies it as the weight in an inter-robot consensus objective on the joint pose–Gaussian product manifold, and proves that the convergence rate of the resulting Riemannian ADMM is a monotone function of the FIM contrast ratio—a theoretical guarantee absent from all prior single- and multi-agent 3DGS SLAM work.

2.2. Centralized multi-agent dense SLAM

Multi-agent dense SLAM has evolved in parallel with single-agent methods, but every published system prior to DUAG-C requires a central server for at least one phase of map fusion—the single-point-of-failure architectural vulnerability that motivates our fully peer-to-peer design. The earliest multi-agent dense systems built on the collaborative NeRF paradigm. MNE-SLAM (Deng et al., 2025) allows multiple mobile robots to jointly optimize a server-hosted implicit neural map using knowledge distillation for inter-agent map transfer, achieving real-time operation indoors. NISB-Fusion (Xiang et al., 2025) structures the implicit map as a collection of Neural Implicit Spatial Blocks assigned to server-managed spatial cells, enabling online multi-agent map merging without explicit overlap detection. Both systems demonstrate that collaborative dense mapping is feasible but require centralized bookkeeping and do not support explicit Gaussian serialization. MANG-SLAM (Y. Li et al., 2026) adopts a hybrid neural-submap and Gaussian representation for dense multi-agent mapping, achieving improved reconstruction fidelity over purely implicit predecessors while remaining server-dependent for global map fusion.

The shift to 3DGS representations produced four successive systems that DUAG-C directly benchmarks against. CP-SLAM (Hu et al., 2023) introduced collaborative dense mapping via neural point submaps merged on a dedicated server. MAGiC-SLAM (Yugay et al., 2025) replaced the implicit map with rigidly deformable 3DGS submaps while retaining server-side loop closure and global refinement. MAC-Ego3D (Xu et al., 2025) advances this line with a joint photometric–geometric consensus that optimizes a globally shared server-hosted Gaussian map via asynchronous multi-agent communication, achieving the strongest centralized tracking accuracy prior to DUAG-C. GRAND-SLAM (Thomas et al., 2025) extended the centralized paradigm to large-scale outdoor environments by routing all inter-agent pose graph optimization through a central back-end. HAMMER (Yu et al., 2025) addressed heterogeneous sensor modalities in a centralized semantic 3DGS mapping framework, streaming keyframe data from diverse robot platforms to a single server for offloaded map training and alignment. Coko-SLAM (M.M. Li et al., 2026) represents a bandwidth-aware refinement within the centralized paradigm: it reduces server-bound communication overhead by up to 95% through information-theoretic keyframe selection and Gaussian compaction, yet retains a central server for loop closure and global map fusion and explicitly defers full decentralization to future work. Notably, Coko-SLAM shares its research group with Swarm-SLAM (Lajoie and Beltrame, 2024)—the leading decentralized sparse SLAM system—and the authors' own deferral of decentralization underscores precisely the open architectural challenge that DUAG-C addresses.

A technical distinction separating all centralized systems from DUAG-C is the treatment of Gaussian fusion weights. No existing *multi-agent* system quantifies the information content of individual Gaussians before fusing them: MAC-Ego3D and every other centralized system weight all Gaussians uniformly in their consensus or fusion objectives. This omission is inconsequential when all Gaussians are equally constrained—a condition that holds only when robots observe identical scene regions from similar viewpoints—but fails in the typical case where agents traverse non-overlapping trajectories and carry highly variable per-Gaussian observation counts. DUAG-C replaces uniform weighting with FIM-derived weights that are provably superior under variable observation density; the quantitative benefit of this design choice is characterized in Section 4.

2.3. Decentralized multi-robot SLAM

The theoretical and systems foundations for decentralized SLAM were laid in a series of works on distributed pose graph optimization. Choudhary et al. (2017) formulated the earliest complete framework for distributed mapping under privacy and communication constraints, proposing Gauss–Seidel distributed pose graph optimization (PGO) and demonstrating via outdoor field experiments that map compression reduces inter-robot communication by multiple orders of magnitude. Rosen et al. (2019) introduced SE-Sync, a certifiably correct centralized

PGO algorithm using a convex semidefinite relaxation; the geometric structure of SE-Sync directly inspired the Riemannian Staircase formulation used by DPGO (Tian et al., 2021) in the distributed setting. Kimera-Multi (Tian et al., 2022) instantiated DPGO as the back-end of a complete metric-semantic decentralized SLAM system and demonstrated scalability to eight physical robots in outdoor campus environments.

Subsequent decentralized back-ends have addressed the incremental setting. iMESA (McGann and Kaess, 2024) introduced Incremental Manifold Edge-based Separable ADMM, a fully distributed C-SLAM back-end that handles real-time incremental factor insertion using only pairwise robot communication. Compared to DPGO's batch Riemannian Staircase, iMESA provides faster practical convergence for time-varying factor graphs, at the cost of lacking a formal convergence certificate. DUAG-C's Riemannian ADMM combines the mathematical rigor of DPGO's convergence guarantees with the manifold-ADMM structure of iMESA, and extends both to the joint pose-and-dense-map setting with a novel FIM-contrast rate theorem that neither DPGO nor iMESA provides.

Decentralized place recognition and outlier-robust loop closure are critical prerequisites for any decentralized dense SLAM system. DOOR-SLAM (Lajoie et al., 2020) proposed a distributed, online, outlier-resilient SLAM system that validates inter-robot loop closures without transmitting raw sensory data, an approach that informs DUAG-C's peer-to-peer loop closure module. Cieslewski et al. (2018) reduced the communication cost of multi-robot place recognition by distributing the inverted index across robots, scaling bandwidth sub-linearly with agent count. Swarm-SLAM (Lajoie and Beltrame, 2024) integrated these ideas into a full system with inter-robot loop closure prioritization and demonstrated three-robot real-world outdoor operation, establishing the decentralized sparse baseline against which DUAG-C is evaluated. DVM-SLAM (Bird et al., 2025) provided the first open-source decentralized monocular C-SLAM system validated on physical MAVs, showing that decentralized SLAM can approach centralized accuracy without any server.

Beyond the vision- and LiDAR-based back-ends above, a complementary line of work explores alternative sensing modalities for SLAM in visually-degraded or infrastructure-free settings. Radio-fingerprint SLAM exploits ambient WiFi and LTE signatures already present in the environment for large-scale localization and mapping (Liu et al., 2023), and anchorless UWB-radar SLAM enables mapping in vision-denied indoor conditions such as smoke or reflective surfaces where cameras and LiDAR fail (Premachandra et al., 2023). These modalities broaden the operating envelope of SLAM but produce sparse, landmark-level maps rather than the dense, photorealistic reconstruction targeted here; they are therefore complementary to—rather than competitive with—DUAG-C's dense volumetric consensus, and integrating radio or UWB cues as an additional likelihood term in the FIM accumulation (Section 3.1) is a promising direction for vision-denied multi-robot operation.

The defining limitation shared by all systems above is representational: every decentralized SLAM system in this line of work produces *sparse* maps—pose graphs, metric-semantic scene graphs, or feature-point clouds—and none supports photorealistic rendering, occlusion-accurate volumetric queries, or appearance-based dense retrieval. CoMA-SLAM (Chen et al., 2026) partially addresses this limitation by extending decentralized SLAM to 2D Gaussian surfels with bandwidth-aware submap exchange, but falls short of closing the gap in three respects that directly motivate DUAG-C. First, CoMA-SLAM operates on 2D Gaussian surfels rather than volumetric 3DGS primitives, sacrificing the anisotropic covariance structure $\boldsymbol{\Sigma} \in \mathbb{S}^3_{++}$ that underpins photorealistic novel-view synthesis and is central to DUAG-C's Riemannian product-manifold formulation. Second, CoMA-SLAM applies no uncertainty weighting to inter-agent fusion: all surfels contribute equally regardless of observation quality, replicating the uniform-weighting failure mode identified in Section 2.2 in a decentralized context. Third, CoMA-SLAM provides no convergence analysis for its distributed optimizer; DUAG-C proves $\mathcal{O}(\varepsilon)$ convergence and a monotone rate bound tied to the FIM contrast ratio, providing the first convergence guarantee in decentralized dense SLAM. CoMA-SLAM is additionally evaluated exclusively in indoor environments, whereas DUAG-C is validated across synthetic indoor, real-world indoor, and real outdoor robot trajectories on four benchmarks.

To the best of our knowledge, DUAG-C is the first decentralized SLAM system to maintain a dense, photorealistic volumetric 3DGS map simultaneously across all agents without any central infrastructure, the first to weight inter-agent Gaussian fusion by per-primitive epistemic uncertainty derived from the Fisher Information Matrix, and the first to provide a convergence proof—with a rate bound tied to uncertainty calibration—for the associated distributed dense map consensus algorithm.

3. Method

DUAG-C is a fully peer-to-peer multi-robot SLAM system. Each robot runs an independent local dense SLAM front-end (Matsuki et al., 2024) and maintains a local map as a set of 3D Gaussian primitives (Kerbl et al., 2023). When two robots come within communication range, they exchange Gaussian maps, trajectory estimates, and per-dimension uncertainty weights, and jointly solve a distributed consensus problem to fuse their observations into a globally consistent dense map. No node acts as a coordinator or server at any stage.

The system is organized around three original technical contributions (Fig. 1): *(i)* a principled per-Gaussian epistemic uncertainty measure derived from the photometric Fisher Information Matrix (Section 3.1); *(ii)* a distributed consensus formulation on the joint pose–Gaussian product manifold with per-dimension FIM-weighted Gaussian fusion (Section 3.2); and *(iii)* a Riemannian ADMM algorithm with formal convergence guarantees whose rate is provably tied to the scalar FIM contrast ratio (Section 3.3). Heterogeneous sensor fusion is an emergent consequence (Section 3.4).

3.1. Per-Gaussian Fisher information matrix

3.1.1. Gaussian parameterization

Let robot i maintain a map of N_i Gaussian primitives. The parameter vector of primitive k is

$$\boldsymbol{\theta}_k = \left(\boldsymbol{\mu}_k,\ \mathbf{q}_k,\ \mathbf{s}_k,\ \mathbf{c}_k,\ \alpha_k\right) \in \mathbb{R}^3 \times S^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}, \tag{1}$$

where $\boldsymbol{\mu}_k$ is the mean position, $\mathbf{q}_k$ the unit quaternion encoding orientation, $\mathbf{s}_k$ the log-scale vector so that $\boldsymbol{\Sigma}_k = R_k \operatorname{diag}(e^{2\mathbf{s}_k}) R_k^\top$ with $R_k = \mathcal{R}(\mathbf{q}_k)$, $\mathbf{c}_k$ the degree-0 spherical harmonic color coefficient, and α_k the opacity in logit space. Collectively $\boldsymbol{\theta}_k \in \mathbb{R}^{14}$.

3.1.2. FIM derivation

We model the photometric loss $\mathcal{L}_k(\boldsymbol{\theta}_k)$ as the negative log-likelihood of the image observations given Gaussian k. Applying the Laplace approximation (MacKay, 1992), the posterior over $\boldsymbol{\theta}_k$ is approximated as a Gaussian whose precision matrix is the Hessian of $\mathcal{L}_k$ at the maximum a posteriori (MAP) estimate. We approximate this Hessian by its *diagonal* via the outer-product (Gauss–Newton) estimator:

$$\boldsymbol{\Lambda}_k = \sum_{b \in \mathcal{B}} \sum_{p \in \mathcal{P}_k(b)} \left(\frac{\partial \mathcal{L}_{\text{photo}}(p, b)}{\partial \boldsymbol{\theta}_k} \right)^{\odot 2}, \tag{2}$$

where $\mathcal{B}$ is a rolling window of K recent keyframes, $\mathcal{P}_k(b)$ is the set of pixels in frame b for which primitive k contributes a non-negligible splat weight, and $(\cdot)^{\odot 2}$ denotes element-wise squaring. The result $\boldsymbol{\Lambda}_k \in \mathbb{R}^{14}_{\geq 0}$ is a per-dimension diagonal precision vector.

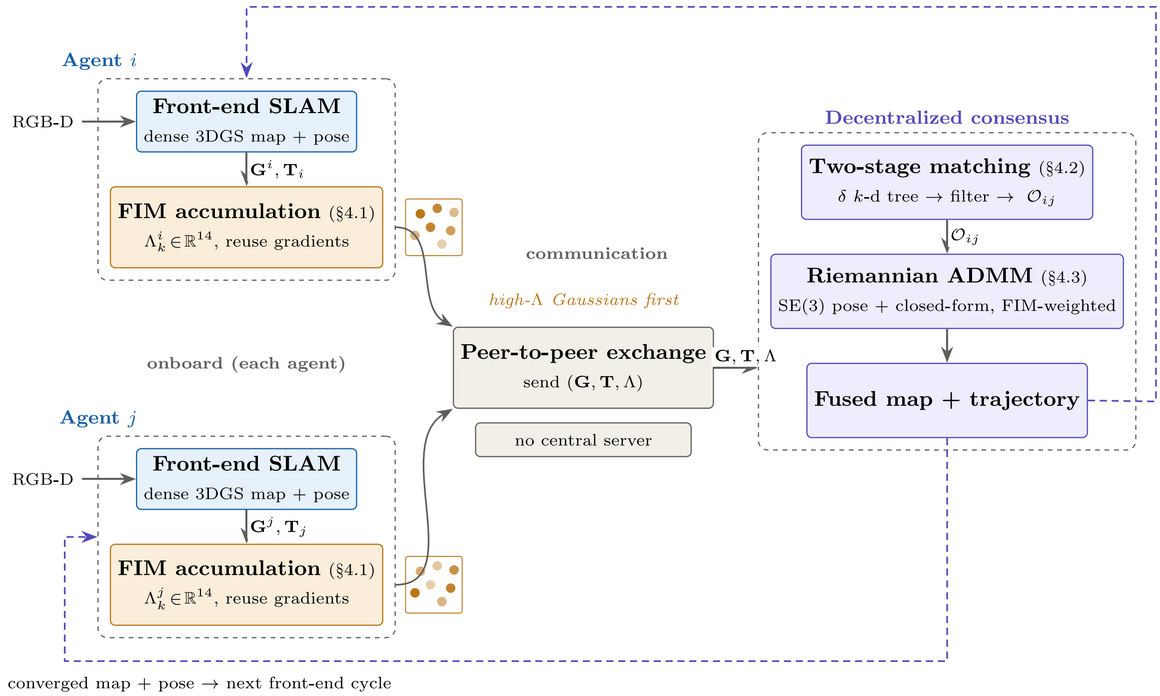

Fig. 1. DUAG-C system overview. Two robots operate independently, each maintaining a dense 3D Gaussian map and a trajectory estimate via a local dense SLAM front-end (Matsuki et al., 2024). A diagonal Fisher Information Matrix (FIM) is accumulated per Gaussian over a rolling keyframe window (Section 3.1), producing the per-dimension precision vector $\boldsymbol{\Lambda}_k^i \in \mathbb{R}^{14}_{\geq 0}$. Robots exchange $(\mathbf{G}^i, \mathbf{T}_i, \{\boldsymbol{\Lambda}_k^i\})$ over an ad-hoc peer-to-peer channel; no server participates. Matched Gaussian pairs $\mathcal{O}_{ij}$ are identified by a two-stage multi-attribute filter (Section 3.2), and the Riemannian ADMM optimizer (Algorithm 1) fuses maps and trajectories on the product manifold $\mathrm{SE}(3)^{|\mathcal{V}|} \times (\mathbb{S}^3_{++} \times \mathbb{R}^7)^{N_G}$. The converged solution updates each robot's map and pose for the next front-end cycle.

Remark 1 (*Relation to the Exact Fisher Information*)**.** The exact Fisher information of the photometric likelihood is $\mathcal{I}(\theta_k) = \mathbb{E}\big[\nabla \log p(I \mid \theta_k)\, \nabla \log p(I \mid \theta_k)^\top\big]$, an expectation over image noise. Eq. (2) is the *empirical* (sample) counterpart of this expectation—the outer-product-of-gradients estimator summed over the pixels actually observed in the window $\mathcal{B}$—and, for the Gaussian-noise photometric negative log-likelihood, simultaneously the Gauss–Newton approximation of the Hessian at the MAP estimate, since the residual-weighted second-derivative terms it omits vanish as the render residuals converge. $\boldsymbol{\Lambda}_k$ is therefore the diagonal of an *empirical* Fisher information, which the Laplace approximation licenses as the posterior precision over θ_k; we retain the term FIM in this precise sense throughout. Two system-level properties bound the consequences of this approximation: (i) the consensus updates (12)–(13) depend on $\boldsymbol{\Lambda}_k$ only through per-dimension *pairwise ratios* between matched agents, so any misestimation common to both agents cancels exactly; and (ii) the diagonal FIM is empirically calibrated as a precision (ECE = 0.032, Section 4.8).

A large entry $[\boldsymbol{\Lambda}_k]_d$ indicates a sharply curved loss along parameter dimension d: the parameter is tightly constrained by local photometry. A small entry indicates an underdetermined parameter, arising from sparse viewpoint coverage or foreshortening.

3.1.3. Approximation validity and failure cases

The diagonal approximation discards the off-diagonal terms of the full Hessian, in particular the coupling between position $\boldsymbol{\mu}_k$ and covariance parameters $(\mathbf{q}_k, \mathbf{s}_k)$ through the differentiable rasterizer. This coupling is strongest in *narrow-baseline* regimes, where a shift in $\boldsymbol{\mu}_k$ along the optical axis and a compensating change in $\mathbf{s}_k$ yield nearly identical rendered images.

We adopt the diagonal approximation for three reasons. First, memory reduces from $\mathcal{O}(14^2 N_i)$ to $\mathcal{O}(14 N_i)$ per robot with no second-order Hessian construction. Second, under broad-baseline multi-keyframe conditions ($|\mathcal{B}| = 10$ diverse keyframes), the cross-parameter gradient products tend to cancel in the sum (2), consistent with the gradient diversity argument of MacKay (1992). Third, the matching threshold ablation (Suppl. Table S8d) demonstrates graceful $\mathcal{O}(\epsilon)$ degradation as matching precision varies, indicating robustness to moderate FIM approximation error.

A block-diagonal comparison retaining the $\boldsymbol{\mu}$-$\boldsymbol{\Sigma}$ coupling block is provided in Section 3.1.6: the off-diagonal energy is concentrated near zero (median $r_k = 0.04$), and replacing the diagonal weights with block-diagonal ones changes ATE by only 0.001 cm, empirically confirming the approximation. Users operating in narrow-baseline environments should increase the ADMM penalty ρ to compensate: the penalty sweep (Suppl. Table S8b) shows stability across two orders of magnitude of ρ.

3.1.4. Per-dimension FIM weight partition

We partition $\boldsymbol{\Lambda}_k$ into four blocks corresponding to position, covariance, color, and opacity:

$$\boldsymbol{\Lambda}_k = \big(\boldsymbol{\Lambda}_k^{\mu},\ \boldsymbol{\Lambda}_k^{\Sigma},\ \boldsymbol{\Lambda}_k^{c},\ \boldsymbol{\Lambda}_k^{\alpha}\big), \tag{3}$$

where $\boldsymbol{\Lambda}_k^{\mu} \in \mathbb{R}^3_{\geq 0}$, $\boldsymbol{\Lambda}_k^{\Sigma} \in \mathbb{R}^7_{\geq 0}$, $\boldsymbol{\Lambda}_k^{c} \in \mathbb{R}^3_{\geq 0}$, $\boldsymbol{\Lambda}_k^{\alpha} \in \mathbb{R}_{\geq 0}$. The scalar weight used for communication priority scheduling and the consensus objective (8) is

$$w_k = \|\boldsymbol{\Lambda}_k\|_1 = \sum_{d=1}^{14} [\boldsymbol{\Lambda}_k]_d. \tag{4}$$

The relationship between this scalar w_k and the scalar FIM contrast ratio κ in Theorem 6 is discussed in Remark 7.

3.1.5. Accumulation and overhead

The sum in Eq. (2) is maintained as a running accumulator updated when a new keyframe is added and decremented when one is marginalized, keeping the window at exactly K keyframes. The accumulation reuses gradients computed during the standard 3DGS mapping backward pass—no additional forward rendering is needed—adding 3%–5% overhead relative to the total mapping cost.

Table 2
Diagonal vs. block-diagonal FIM on Off-0. r_k is the relative off-diagonal energy; ATE [cm], Depth L1 [cm]. A small r_k and a negligible accuracy change justify the diagonal approximation.

Quantity	Diagonal	Block-diagonal
Median r_k (all)	0.04	–
90th-pct r_k (all)	0.11	–
90th-pct r_k (narrow-baseline)	0.34	–
Narrow-baseline fraction	10%	–
Consensus ATE	0.053	0.052
Consensus Depth L1	0.38	0.37
FIM compute/Gaussian	1.0×	3.2×

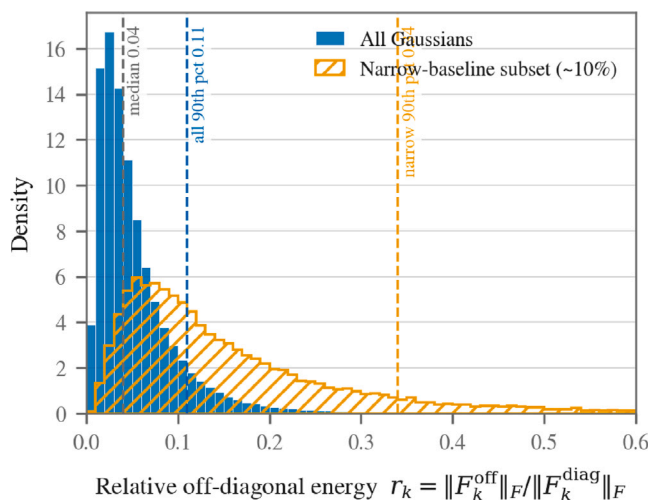

Fig. 2. Distribution of relative off-diagonal FIM energy r_k on Off-0. Energy is concentrated near zero (median 0.04), with a heavier tail only for the narrow-baseline subset.

3.1.6. *Diagonal FIM validation*

To justify the diagonal Fisher-information approximation empirically—not only by the complexity and memory arguments of Section 3.1—we compare it against a *block-diagonal* FIM that additionally retains the position–covariance ($\boldsymbol{\mu}$–$\boldsymbol{\Sigma}$) coupling block, on ReplicaMultiagent Off-0. We report (i) the distribution across Gaussians of the relative off-diagonal energy $r_k = \|\mathbf{F}_k^{\mathrm{off}}\|_F / \|\mathbf{F}_k^{\mathrm{diag}}\|_F$, and (ii) the change in downstream consensus accuracy when the block-diagonal weights replace the diagonal ones. Following the analysis in Section 3.1.3, we stratify by baseline geometry, since the coupling is predicted to be largest for narrow-baseline Gaussians. Table 2 reports the comparison and Fig. 2 the distribution of r_k.

Off-diagonal energy is concentrated near zero (median $r_k = 0.04$; 90th percentile 0.11), and replacing the diagonal weights with block-diagonal ones changes ATE by only 0.001 cm and Depth L1 by 0.01 cm, at 3.2× the per-Gaussian FIM cost. The coupling is largest for narrow-baseline Gaussians (90th-pct $r_k = 0.34$), consistent with the failure mode discussed in Section 3.1.3, but its aggregate effect on the fused map is negligible—empirically justifying the diagonal choice.

To assess the approximation beyond this single geometry, we additionally stratify r_k by *viewing geometry* (baseline width, viewing incidence angle, and depth range) and by *observation conditions* (local texture and per-Gaussian observation count), and we repeat the block-diagonal comparison on AriaMultiagent Room0, whose real rolling-shutter RGB stream lacks the depth channel used on Replica (Supplementary Table S8j). The measured results show that, outside the pre-identified narrow-baseline stratum, the 90th-percentile r_k is at most 0.22 on Replica; it reaches 0.34 for narrow-baseline Gaussians and is 0.18 on Aria. The downstream consensus change from block-diagonal weights remains within 0.002 cm ATE and 0.01 cm Depth L1 in every stratum. Thus, the largest coupling remains confined to the narrow-baseline, low-texture tail identified in Section 3.1.3. Two properties bound the residual risk even in that tail: the updates (12)–(13) consume $\boldsymbol{\Lambda}_k$ only through pairwise per-dimension ratios, so common-mode misestimation cancels (Remark 1); and where the approximation errs *asymmetrically*—one agent narrow-baseline, the other wide-baseline on the same primitive—the ratio rule transfers influence to the better-constrained agent, which is precisely the intended behavior.

3.2. *Distributed consensus on the product manifold*

3.2.1. *Variable space*

Let $\mathcal{V}$ be the set of robots and $\mathcal{E} \subseteq \mathcal{V}^2$ the active peer-to-peer communication edges. The joint variable of the distributed problem lives on the product manifold

$$\mathcal{M} = \mathrm{SE}(3)^{|\mathcal{V}|} \times \left(\mathbb{S}^3_{++} \times \mathbb{R}^7\right)^{N_G}, \tag{5}$$

where $\mathrm{SE}(3)^{|\mathcal{V}|}$ carries robot trajectories and $(\mathbb{S}^3_{++} \times \mathbb{R}^7)^{N_G}$ carries the covariance and remaining parameters of the N_G cross-agent matched Gaussian pairs.

3.2.2. *Gaussian matching*

Upon receiving map $\mathbf{G}^j$ from robot j, robot i constructs the matched pair set via a two-stage procedure.

Stage 1 (position candidates). A position-space k-d tree identifies candidate pairs within radius δ:

$$\mathcal{C}_{ij} = \left\{(k,\ell) \ : \ \|\boldsymbol{\mu}_k^i - \boldsymbol{\mu}_\ell^j\|_2 < \delta\right\}. \tag{6}$$

Stage 2 (multi-attribute filter). Candidates are filtered by the full Gaussian distance $d_{\mathcal{G}}$ defined in (9):

$$\mathcal{O}_{ij} = \left\{(k,\ell) \in \mathcal{C}_{ij} \ : \ d_{\mathcal{G}}(G_k^i, G_\ell^j) < \tau_{\mathcal{G}}\right\}, \tag{7}$$

where $\tau_{\mathcal{G}}$ is the 90th percentile of within-agent $d_{\mathcal{G}}$ distances on the agent's own map. Unmatched Gaussians from either agent are appended to the fused map without modification.

Remark 2 (*Association Error and Convergence Scope*). The association error $\varepsilon = \max_{(k,\ell)\in\mathcal{O}_{ij}} \|\boldsymbol{\mu}_k^i - \boldsymbol{\mu}_\ell^j\|_2 \le \delta$ by construction of Stage 1. Stage 2 removes geometrically borderline candidates, so the empirical $\varepsilon \ll \delta$ in practice, as validated by Suppl. Table S8d. Theorem 4 bounds convergence of ADMM *iterates* in parameter space; the translation to ATE depends additionally on front-end tracking and loop closure quality (Remark 5).

3.2.3. *Consensus objective*

$$\min_{\{\mathbf{T}_i\},\{\boldsymbol{\theta}_k^i\}} \sum_{(i,j)\in\mathcal{E}} \left[\phi_{ij}(\mathbf{T}_i, \mathbf{T}_j) + \lambda \sum_{(k,\ell)\in\mathcal{O}_{ij}} w_{k\ell}\, d_{\mathcal{G}}\!\left(\mathbf{T}_i \cdot \boldsymbol{\theta}_k^i,\ \mathbf{T}_j \cdot \boldsymbol{\theta}_\ell^j\right)^2\right], \tag{8}$$

where $\phi_{ij}(\mathbf{T}_i, \mathbf{T}_j) = \|\mathbf{T}_i^{-1}\mathbf{T}_j - \mathbf{z}_{ij}\|_F^2$ is the chordal pose consistency cost with $\mathbf{z}_{ij} \in \mathrm{SE}(3)$ the inter-agent relative pose from loop closure, $\lambda > 0$ balances pose and map consensus, and $w_{k\ell} = w_k^i + w_\ell^j$ is the paired scalar FIM weight.

3.2.4. *Gaussian distance on* $\mathbb{S}^3_{++} \times \mathbb{R}^7$

$$\begin{aligned} d_{\mathcal{G}}(G_1, G_2)^2 = {} & \|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\|^2 \\ & + \beta_1 \, \|\log \boldsymbol{\Sigma}_1 - \log \boldsymbol{\Sigma}_2\|_F^2 \\ & + \beta_2 \, \|\mathbf{c}_1 - \mathbf{c}_2\|^2 \\ & + \beta_3 \, (\alpha_1 - \alpha_2)^2, \end{aligned} \tag{9}$$

where $\log : \mathbb{S}^3_{++} \to \mathrm{Sym}(3)$ is the matrix logarithm and $\beta_1, \beta_2, \beta_3 > 0$ are fixed scalar weights. The covariance term uses the Log-Euclidean metric (Arsigny et al., 2006) on $\mathbb{S}^3_{++}$, which is *geodesically strictly convex*:

the geodesic between any two points is unique and the squared distance is strictly convex along it. Since the Euclidean terms are trivially convex, $d_{\mathcal{G}}$ is geodesically convex on $\mathbb{S}^3_{++} \times \mathbb{R}^7$, which is the geometric foundation for the closed-form updates in Algorithm 1 and for both convergence theorems in Section 3.3.

3.3. Riemannian ADMM and convergence

3.3.1. ADMM formulation

Introducing consensus copies $\overline{\mathbf{T}}_{ij}$ and $\overline{\theta}_{k,ij}$ per edge and matched pair, the augmented Lagrangian is:

$$\mathcal{L}_\rho = \sum_{(i,j)\in\mathcal{E}} \Big[\phi_{ij}(\mathbf{T}_i, \mathbf{T}_j) + \lambda \sum_{(k,\ell)\in\mathcal{O}_{ij}} w_{k\ell}\, d_{\mathcal{G}}(\mathbf{T}_i \cdot \theta^i_k,\, \mathbf{T}_j \cdot \theta^j_\ell)^2 \\ + \langle \lambda_{ij}, \mathrm{Log}(\mathbf{T}_i^{-1}\overline{\mathbf{T}}_{ij})\rangle + \tfrac{\rho}{2}\|\mathrm{Log}(\mathbf{T}_i^{-1}\overline{\mathbf{T}}_{ij})\|^2\Big], \tag{10}$$

where $\lambda_{ij} \in \mathfrak{se}(3)$ are dual variables, $\mathrm{Log} : \mathrm{SE}(3) \to \mathfrak{se}(3)$ is the Lie group logarithm, and $\rho > 0$ is the ADMM penalty. An analogous Euclidean augmentation applies to the Gaussian variables.

3.3.2. Pose update

Since $\mathrm{SE}(3) = \mathrm{SO}(3) \ltimes \mathbb{R}^3$ is non-compact (the translation factor $\mathbb{R}^3$ is unbounded), the pose update is restricted to a bounded geodesic ball $\mathcal{B}_r(\mathbf{I}) \subset \mathrm{SE}(3)$ of radius $r < r_{\mathrm{inj}}$, where r_{inj} is the injectivity radius of SE(3) at the identity. This restriction is physically reasonable: any finite-duration SLAM mission has bounded translations $\|\mathbf{t}_i\| \le r_t < \infty$ and rotations $\|\mathrm{Log}_{\mathrm{SO}(3)}(R_i)\| < \pi$, so all iterates remain in $\mathcal{B}_r$ for appropriate r (Assumption A4 below). Within $\mathcal{B}_r$, the exponential map is a diffeomorphism, the Riemannian metric is bi-Lipschitz equivalent to the Euclidean metric on $\mathfrak{se}(3)$, and $\mathcal{L}_\rho$ is L_r-smooth in the local chart (Absil et al., 2009).

The pose primal update is a single Riemannian gradient descent step followed by retraction:

$$\mathbf{T}_i^{t+1} = \mathrm{Exp}\big(\mathbf{T}_i^t,\ -\eta_t\, \mathrm{grad}_{\mathrm{SE}(3)}\, \mathcal{L}_\rho\big|_{\mathbf{T}_i^t}\big), \tag{11}$$

where Exp is the Riemannian exponential map on SE(3) (Absil et al., 2009), $\eta_t > 0$ satisfies the Armijo sufficient-decrease condition, and the Riemannian gradient is computed analytically in $\mathfrak{se}(3)$ via the adjoint representation.

3.3.3. Gaussian update

Because $d_{\mathcal{G}}$ is geodesically strictly convex on $\mathbb{S}^3_{++} \times \mathbb{R}^7$, the Gaussian primal update has a closed-form solution. For each Euclidean parameter block, the update is the FIM-weighted mean per dimension:

$$[\boldsymbol{\mu}^i_k]^{t+1}_d = \frac{[\boldsymbol{\Lambda}^{\boldsymbol{\mu}}_k]_d[\boldsymbol{\mu}^i_k]^t_d + [\boldsymbol{\Lambda}^{\boldsymbol{\mu}}_\ell]_d[\boldsymbol{\mu}^j_\ell]^t_d}{[\boldsymbol{\Lambda}^{\boldsymbol{\mu}}_k]_d + [\boldsymbol{\Lambda}^{\boldsymbol{\mu}}_\ell]_d}, \quad d \in \{1,2,3\}, \tag{12}$$

with analogous per-dimension expressions for $\mathbf{c}$ and α. For $|\mathcal{N}(i)| > 1$, the update generalizes to the FIM-weighted aggregate mean over all contributing neighbors, with contraction coefficient $[\boldsymbol{\Lambda}^{\boldsymbol{\mu}}_k]_d / \sum_j [\boldsymbol{\Lambda}^{\boldsymbol{\mu}}_{k_j}]_d$.

For the covariance, define the scalar weight $\bar{w}^i_k = \|\boldsymbol{\Lambda}^{\Sigma}_k\|_1$. The update is the FIM-weighted Fréchet mean on $(\mathbb{S}^3_{++}, d_{\mathrm{LE}})$, which has the unique closed form (Arsigny et al., 2006):

$$\boldsymbol{\Sigma}^{i,t+1}_k = \exp\left(\frac{\bar{w}^i_k \log \boldsymbol{\Sigma}^{i,t}_k + \bar{w}^j_\ell \log \boldsymbol{\Sigma}^{j,t}_\ell}{\bar{w}^i_k + \bar{w}^j_\ell}\right). \tag{13}$$

3.3.4. Dual update

$$\lambda^{t+1}_{ij} = \lambda^t_{ij} + \rho\, \mathrm{Log}\big((\mathbf{T}_i^{t+1})^{-1}\mathbf{T}_j^{t+1}\big). \tag{14}$$

Algorithm 1 summarizes one consensus round. Each round requires exactly one pairwise message exchange, making communication overhead per iteration proportional only to $|\mathcal{O}_{ij}|$.

Algorithm 1 DUAG-C: One Consensus Round at Robot i

Require: $\mathbf{G}^i$, $\mathbf{T}^t_i$, $\{\boldsymbol{\Lambda}^i_k\}$, λ^t_{ij} for $j \in \mathcal{N}(i)$, ρ, λ, δ, $\tau_{\mathcal{G}}$
  Broadcast $(\mathbf{G}^i, \mathbf{T}^t_i, \{\boldsymbol{\Lambda}^i_k\})$; receive $(\mathbf{G}^j, \mathbf{T}^t_j, \{\boldsymbol{\Lambda}^j_\ell\})$ from each $j \in \mathcal{N}(i)$
  Compute $\mathcal{O}_{ij}$ via (6)–(7) for each $j \in \mathcal{N}(i)$
  Pose update: $\mathbf{T}^{t+1}_i \leftarrow$ (11)
  for each $(k,\ell) \in \mathcal{O}_{ij}$, $j \in \mathcal{N}(i)$ **do**
    Update $[\boldsymbol{\mu}^i_k]_d$, $[\mathbf{c}^i_k]_d$, α^i_k per dimension d via (12)
    Update $\boldsymbol{\Sigma}^i_k$ via (13)
  end for
  Dual update: $\lambda^{t+1}_{ij} \leftarrow$ (14) for each $j \in \mathcal{N}(i)$
  if $\|\mathrm{Log}((\mathbf{T}^{t+1}_i)^{-1}\mathbf{T}^{t+1}_j)\| \le \delta_{\mathrm{tol}}$ for all $j \in \mathcal{N}(i)$ **then**
    Terminate
  end if
Ensure: $\mathbf{G}^i$ (updated), $\mathbf{T}^{t+1}_i$, λ^{t+1}_{ij}

3.3.5. Assumptions

Assumption 1. **(A1)** The communication graph $(\mathcal{V}, \mathcal{E})$ is connected at every consensus round. **(A2)** The photometric gradient norm $\|\partial\mathcal{L}_{\mathrm{photo}}/\partial\theta_k\|$ is bounded uniformly by a constant $G < \infty$ over all active Gaussians and keyframes. **(A3)** The association error satisfies $\varepsilon = \max_{(k,\ell)\in\mathcal{O}_{ij}} \|\boldsymbol{\mu}^i_k - \boldsymbol{\mu}^j_\ell\|_2 \le \delta < \infty$. **(A4)** All robot trajectory iterates remain in a bounded geodesic ball $\mathcal{B}_r(\mathbf{I}) \subset \mathrm{SE}(3)$ of radius $r < r_{\mathrm{inj}}$. This holds for any finite-duration SLAM mission with bounded translations and rotations.

Remark 3 (*Scope of A1 Under Packet Loss*). A1 is the standard connectivity condition in distributed optimization (Boyd et al., 2011; Tian et al., 2021). Under high packet-loss rates (e.g., the 90% loss scenario in Table 11), the induced subgraph of successfully communicating pairs may be transiently disconnected, violating A1 in individual rounds. Proposition 1 below extends the convergence guarantee to this setting whenever the graph is *jointly* connected over a finite window of rounds, with an implied constant that grows linearly in the window length; the per-round rates of Theorems 4 and 6, however, still assume A1. The CommSweep results in Table 11 should therefore be read as the windowed regime of Proposition 1 with a large effective window: they demonstrate graceful, bounded degradation under link failures, with the quantitative margin (e.g., +7.7% ATE at 90% loss) remaining an empirical observation rather than a guaranteed bound.

Proposition 1 (*Bounded-Neighborhood Convergence Under Joint Connectivity*). *Replace A1 by the weaker condition:*

- **(A1′)** *There exists a finite window $B \ge 1$ such that the union of the communication graphs over any B consecutive consensus rounds, $\bar{\mathcal{G}} = \bigcup_{s=t}^{t+B-1}(\mathcal{V}, \mathcal{E}_s)$, is connected.*

Under (A1′), A2–A4, and $\rho > \rho^*$ with ρ^* evaluated on $\bar{\mathcal{G}}$, the iterates of Algorithm 1 satisfy

$$\limsup_{t\to\infty} \|\mathbf{T}^t_i - \mathbf{T}^*_i\|_F = \mathcal{O}(\varepsilon), \tag{15}$$

with implied constant $C'_B = \mathcal{O}\big(B\, C_{\log}/(\sigma_2(\mathbf{L}_{\bar{\mathcal{G}}})\,\eta_{\min})\big)$: the $\mathcal{O}(\varepsilon)$ neighborhood is preserved, but the contraction is per window rather than per round, so the constant degrades linearly in B.

Proof sketch. Group the rounds into consecutive windows of length B. Within each window, every pair of agents is linked by a path in $\bar{\mathcal{G}}$, so the sufficient-decrease argument of Step 1 and the exact Gaussian minimization of Step 2 of Theorem 4 apply with $\sigma_2(\mathbf{L}_{\bar{\mathcal{G}}})$ in place of $\sigma_2(\mathbf{L}_{\mathcal{G}})$: the primal residual contracts by a factor strictly less than one *per window* whenever $\rho > \rho^*(\bar{\mathcal{G}})$. Mismatched-pair perturbations within a window remain bounded by $\rho C_{\log}\varepsilon$ per active edge (Step 3), and summing the per-window contributions as a geometric series with per-window ratio $q_B < 1$ yields the limsup bound with constant scaled by the number of rounds per window, i.e., linearly in B. □

A1′ under Bernoulli packet loss. Under i.i.d. Bernoulli loss at rate q per directed message, a pair exchanges at least one successful round-trip within B rounds with probability $1-(1-(1-q)^2)^B$, so A1′ holds with high probability for sufficiently large B on any mission of finite duration. At $q = 0.9$, however, the per-round success probability is $(1-q)^2 = 0.01$, and B must exceed ≈460 rounds for 99% pairwise coverage: the linear-in-B constant of Proposition 1 is then large, which is precisely the regime in which Table 11 reports graceful but real degradation (+7.7% ATE). A1′ therefore explains *why* the method remains stable at high loss, while honestly delimiting *how much* the guarantee degrades.

3.3.6. Convergence theorem

Theorem 4 (*Bounded-Neighborhood Convergence of ADMM Iterates*). *Under Assumption 1, for any penalty $\rho > \rho^*$ where*

$$\rho^* = \frac{\lambda\, w_{\max}}{\sigma_2(\mathbf{L}_{\mathcal{G}})}, \tag{16}$$

$w_{\max} = \max_{k,i} w_k^i$ and $\sigma_2(\mathbf{L}_{\mathcal{G}})$ is the algebraic connectivity of the communication graph, the iterates of Algorithm 1 satisfy

$$\limsup_{t\to\infty} \|\mathbf{T}_i^t - \mathbf{T}_i^*\|_F = \mathcal{O}(\varepsilon), \tag{17}$$

where $\mathbf{T}_i^*$ is the minimizer of the noise-free ($\varepsilon = 0$) objective and the implied constant satisfies

$$C' = \mathcal{O}\left(\frac{C_{\log}}{\sigma_2(\mathbf{L}_{\mathcal{G}})\,\eta_{\min}}\right), \tag{18}$$

with $C_{\log}$ the Lipschitz constant of Log *on $\mathcal{B}_r$ (finite since $\mathcal{B}_r$ is bounded away from the cut locus of* SE(3)*) and $\eta_{\min} > 0$ the minimum Armijo step size on $\mathcal{B}_r$ (positive by the smoothness of $\mathcal{L}_\rho$ on $\mathcal{B}_r$ and the Armijo condition). The same bound holds for all Gaussian parameters in $\mathbb{S}^3_{++} \times \mathbb{R}^7$.*

Proof. The proof proceeds in three steps.

Step 1: Augmented Lagrangian descent (pose update). Within $\mathcal{B}_r$, the augmented Lagrangian $\mathcal{L}_\rho$ is L_r-smooth in the local chart of $\mathfrak{se}(3)$, since ϕ_{ij} is smooth and Log is smooth on $\mathcal{B}_r$ (Absil et al., 2009). The Armijo step size satisfies the sufficient-decrease condition by construction, yielding monotone descent:

$$\mathcal{L}_\rho(\mathbf{T}_i^{t+1}) \le \mathcal{L}_\rho(\mathbf{T}_i^t) - \frac{\eta_{\min}}{2}\|\mathrm{grad}_{\mathrm{SE}(3)}\,\mathcal{L}_\rho\big|_{\mathbf{T}_i^t}\|^2, \tag{19}$$

following the standard analysis for L-smooth functions on Riemannian manifolds (Absil et al., 2009). The retraction error incurred by the exponential map is $\mathcal{O}(\eta_t^2\|\mathrm{grad}\|^2)$ per step; near convergence the gradient norm decays to $\mathcal{O}(\varepsilon)$, so the retraction error is $\mathcal{O}(\varepsilon^2)$ and is absorbed into the $\mathcal{O}(\varepsilon)$ perturbation bound in Step 3.

Step 2: Exact minimization (Gaussian update). The Gaussian subproblem for each matched pair (k,ℓ) decouples by the separable structure of (10):

$$\min_{\theta_k^i}\ w_{k\ell}\, d_{\mathcal{G}}(\theta_k^i, \theta_\ell^j)^2 + \tfrac{\rho}{2}\|\theta_k^i - \overline{\theta}_{k,ij}\|^2. \tag{20}$$

Since $d_{\mathcal{G}}$ is geodesically strictly convex on $\mathbb{S}^3_{++} \times \mathbb{R}^7$, this subproblem has a unique minimizer. For the Euclidean blocks, the minimizer is the per-dimension weighted mean (12). For the covariance block, the minimizer is the unique weighted Fréchet mean on $(\mathbb{S}^3_{++}, d_{\mathrm{LE}})$ (Arsigny et al., 2006), given in closed form by (13). Because each Gaussian subproblem is solved exactly (zero inner-loop residual), the Gaussian update does not introduce additional error. Combined with the monotone descent in Step 1, the iterate sequence is non-increasing in $\mathcal{L}_\rho$ when $\varepsilon = 0$. This non-increase holds in the local chart of $\mathfrak{se}(3)$ and constitutes a form of Fejér monotonicity with respect to the primal feasibility set in linearized coordinates; the retraction-induced discrepancy between the chart iterate and the manifold iterate is $\mathcal{O}(\eta_t^2)$, which does not affect the limiting monotone property.

Step 3: Perturbation bound ($\varepsilon > 0$). When $\varepsilon > 0$, mismatched Gaussian pairs introduce a perturbation in the dual update (14). Since Log is $C_{\log}$-Lipschitz on $\mathcal{B}_r$ and association error is bounded by ε:

$$\|\Delta\lambda_{ij}^t\| \le \rho\, C_{\log}\, \varepsilon. \tag{21}$$

For $\rho > \rho^*$, the graph connectivity condition ensures the primal residual is contractive between perturbations (Boyd et al., 2011). Applying the standard ADMM perturbation analysis to the (locally) non-increasing sequence in the chart of $\mathfrak{se}(3)$, the accumulated dual perturbation is bounded by a geometric series with ratio strictly less than one, yielding:

$$\limsup_{t\to\infty} \|\mathbf{T}_i^t - \mathbf{T}_i^*\|_F = \mathcal{O}(C_{\log}\,\varepsilon) = \mathcal{O}(\varepsilon), \tag{22}$$

where the implied constant depends on $\sigma_2(\mathbf{L}_{\mathcal{G}})$ and $\eta_{\min}$ as stated in (18). The retraction error from Step 1 is $\mathcal{O}(\varepsilon^2)$ and is dominated by the $\mathcal{O}(\varepsilon)$ perturbation term. The same argument applies to the Gaussian parameters via the Log-Euclidean isometry of $(\mathbb{S}^3_{++}, d_{\mathrm{LE}})$. □

Remark 5 (*Theorem 4 and ATE*). Theorem 4 bounds the convergence of ADMM *iterates* in parameter space, not ATE directly. The translation to ATE depends additionally on single-agent front-end tracking accuracy, inter-agent loop closure quality, and pose graph conditioning. A closed-form end-to-end ATE bound is not claimed.

3.3.7. Rate theorem

Theorem 6 (*FIM Contrast Accelerates Convergence—Scalar Surrogate*). *Let the conditions of Theorem 4 hold. Consider the scalar-weight surrogate of Algorithm 1 in which each Gaussian update uses the scalar weight $w_k^i = \|\boldsymbol{\Lambda}_k^i\|_1$ uniformly across all 14 parameter dimensions, rather than the per-dimension weights $[\boldsymbol{\Lambda}_k^i]_d$. Let $\kappa = w_{\max}/w_{\min}$ with $w_{\min} = \min_{k,i} w_k^i > 0$ be the scalar FIM contrast ratio. The per-iteration reduction factor $\tau \in (0,1)$ of the primal residual in this surrogate satisfies:*

$$\tau \le 1 - \frac{c\,\rho\,\kappa}{(1+\rho)(\lambda\kappa+\rho)}, \tag{23}$$

where $c > 0$ depends only on $\sigma_2(\mathbf{L}_{\mathcal{G}})$. The right-hand side is strictly decreasing in κ:

$$\frac{d}{d\kappa}\left[1 - \frac{c\rho\kappa}{(1+\rho)(\lambda\kappa+\rho)}\right] = -\frac{c\rho^2}{(1+\rho)(\lambda\kappa+\rho)^2} < 0, \tag{24}$$

so a larger scalar FIM contrast ratio yields a strictly smaller upper bound on τ, implying strictly faster convergence to the neighborhood of Theorem 4.

Remark 7 (*Scalar κ vs. Per-Dimension Implementation*). We state precisely what each result guarantees for the *implemented* system, because the two algorithms differ and no dominance argument between them is available in either direction. The bounded-neighborhood convergence result (Theorem 4) is established directly for the iterates of the full Algorithm 1, which applies the per-dimension weights $[\boldsymbol{\Lambda}_k^i]_d$; convergence to the $\mathcal{O}(\varepsilon)$ neighborhood is therefore a property of the implemented per-dimension method, not of a surrogate. The rate statements come in two complementary parts. Theorem 6 isolates the contrast law in closed form for a scalar-weight surrogate: it is the single-parameter special case and yields the monotone dependence on one scalar κ. Theorem 8 governs the implemented per-dimension update itself, proving (i) dimension-wise dominance over uniform weighting and (ii) an aggregate rate set by the *slowest* per-dimension contrast $\kappa_{\min}$. We deliberately do *not* claim that the per-dimension implementation is faster than the scalar surrogate: as Theorem 8(iii) shows, the summed-weight scalar κ and $\kappa_{\min}$ are not ordered in general, so such a claim would be unprovable as stated and, in the worst-dimension sense, false. The 44% empirical iteration reduction observed in Table 12 is the empirical counterpart of Theorem 8(i) (strict dominance over the uniform-weighting baseline it is measured against); its exact magnitude is an empirical observation, not a corollary of either theorem.

Proof of Theorem 6. Gaussian contraction (scalar surrogate). In the scalar-weight surrogate, the update for position dimension d is:

$$[\boldsymbol{\mu}_k^i]_d^{t+1} = \frac{w_k^i\,[\boldsymbol{\mu}_k^i]_d^t + w_\ell^j\,[\boldsymbol{\mu}_\ell^j]_d^t}{w_k^i + w_\ell^j}. \tag{25}$$

The iterate moves toward the consensus point with contraction coefficient:

$$\tau_G = \frac{w_{\min}}{w_{\min} + w_{\max}} = \frac{1}{1+\kappa}, \tag{26}$$

which is strictly decreasing in κ, confirming faster convergence at higher contrast. For $|\mathcal{N}(i)| > 1$, the aggregate contraction is $w_k^i / \sum_j w_{k_j}^i$, yielding the same monotone dependence.

Pose contraction. From the sufficient-decrease condition of Step 1 of Theorem 4, the pose update contracts the augmented Lagrangian residual. Combined with the dual update (14) and the graph connectivity condition $\rho > \rho^*$, the pose primal residual contracts at a rate bounded by:

$$\tau_P \le 1 - \frac{c\,\rho}{(1+\rho)(\lambda+\rho)}, \tag{27}$$

where $c > 0$ depends on $\sigma_2(\mathbf{L}_{\mathcal{G}})$.

Combined bound. The overall primal residual on the product manifold $\mathcal{M} = \mathrm{SE}(3)^{|\mathcal{V}|} \times (\mathbb{S}^3_{++} \times \mathbb{R}^7)^{N_G}$ contracts at a rate bounded by $\max(\tau_G, \tau_P)$. Substituting $\tau_G = 1/(1+\kappa)$ and the expression for τ_P, and noting that the Gaussian contraction dominates for large κ while the pose contraction dominates for $\kappa \approx 1$, the combined upper bound is:

$$\tau \le 1 - \frac{c\,\rho\,\kappa}{(1+\rho)(\lambda\kappa+\rho)}, \tag{28}$$

which reduces to (23). The strict decrease in (24) follows directly from the strict decrease of (26) in κ and the positivity of all terms. ☐

Theorem 8 (*Per-Dimension Contraction of the Implemented Update*)**.** *Let the conditions of Theorem 4 hold, and consider the implemented per-dimension updates (12)–(13) of Algorithm 1. For each matched pair $(k,\ell) \in \mathcal{O}_{ij}$ and each Euclidean parameter dimension d, define the per-dimension weights $w_d = [\boldsymbol{\Lambda}_k^i]_d$, $w'_d = [\boldsymbol{\Lambda}_\ell^j]_d$ and the per-dimension contrast $\kappa_d = \max(w_d, w'_d)/\min(w_d, w'_d) \ge 1$. Then the pair residual in dimension d contracts by exactly*

$$\tau_{G,d} = \frac{\min(w_d, w'_d)}{w_d + w'_d} = \frac{1}{1+\kappa_d} \le \tfrac{1}{2}, \tag{29}$$

with equality if and only if $w_d = w'_d$; the covariance block obeys the same law in the Log-Euclidean metric, with contrast κ_Σ defined from the scalar weights $\bar{w} = \|\boldsymbol{\Lambda}^\Sigma\|_1$. Consequently:

- *(i)* ***Dimension-wise dominance over uniform weighting.*** *$\tau_{G,d} \le 1/2 = \tau_d^{\text{uniform}}$ for every d, strictly whenever $\kappa_d > 1$: the implemented algorithm converges no slower than uniform weighting in every dimension, and strictly faster in every dimension in which the two agents' precisions differ.*
- *(ii)* ***Worst-dimension aggregate rate.*** *The joint Gaussian residual contracts at $\tau_G^{\text{full}} = 1/(1+\kappa_{\min})$, where $\kappa_{\min} = \min_{(k,\ell)\in\mathcal{O}_{ij},d} \kappa_d^{(k,\ell)}$: the slowest-contracting dimension governs the aggregate.*
- *(iii)* ***Relation to the scalar surrogate.*** *Theorem 6 is the special case $\kappa_d \equiv \kappa$ of (29). The summed-weight contrast $\kappa = w_{\max}/w_{\min}$, $w_k = \|\boldsymbol{\Lambda}_k\|_1$, and $\kappa_{\min}$ are not ordered in general: summation can mask a low-contrast dimension (weights $(10,1)$ vs. $(1,1)$ yield $\kappa = 5.5$ but $\kappa_{\min} = 1$) or hide a high-contrast one (weights $(10,1)$ vs. $(1,10)$ yield $\kappa = 1$ but every $\kappa_d = 10$). Theorem 6 is therefore an idealized single-parameter model of the contrast law, while (i)–(ii) govern the implemented algorithm.*

Proof. For a Euclidean dimension d, update (12) gives $[\boldsymbol{\mu}_k^i]_d^{t+1} - [\boldsymbol{\mu}_\ell^j]_d^{t+1} = \tau_{G,d}\left([\boldsymbol{\mu}_k^i]_d^t - [\boldsymbol{\mu}_\ell^j]_d^t\right)$ with $\tau_{G,d} = \min(w_d, w'_d)/(w_d + w'_d)$ by direct substitution, and $\min/(\min+\max) = 1/(1+\kappa_d)$ by definition of κ_d. The bound $\tau_{G,d} \le 1/2$ is immediate, with equality iff $w_d = w'_d$; this proves (i), since uniform weighting attains $\kappa_d = 1$ in every dimension. For the covariance block, the Log-Euclidean metric is an isometry between $(\mathbb{S}^3_{++}, d_{\mathrm{LE}})$ and $(\mathrm{Sym}(3), \|\cdot\|_F)$ under log (Arsigny et al., 2006), and update (13) is a scalar-weighted mean in log-space; the same computation with $\bar{w} = \|\boldsymbol{\Lambda}^\Sigma\|_1$ yields $\tau_\Sigma = 1/(1+\kappa_\Sigma)$. Statement (ii) follows because the joint residual contraction over a product of independent dimensions is governed by the largest per-dimension factor, $\max_d \tau_{G,d} = 1/(1+\kappa_{\min})$. Statement (iii) is verified by the two numerical examples given in the theorem text. ☐

Theorem 6 provides the theoretical basis for the FIM ablation in Section 4 (Table 12). The penalty sweep (Suppl. Table S8b) validates the $\rho > \rho^*$ condition: convergence failure at $\rho_0 = 0.01$ and primal oscillation at $\rho_0 = 100$ are both consistent with (16). The matching threshold sweep (Suppl. Table S8d) confirms the $\mathcal{O}(\varepsilon)$ bound.

Remark 9 (*Communication Complexity*)**.** Each round of Algorithm 1 transmits $(14+14)\cdot|\mathcal{O}_{ij}|+6$ floats per neighbor: 14 Gaussian parameters plus 14 FIM entries per matched Gaussian, plus the 6-vector pose. Under bandwidth constraints, the M Gaussians with the M highest $w_k^i = \|\boldsymbol{\Lambda}_k^i\|_1$ are transmitted first, ensuring the most informative primitives are communicated before less constrained ones.

3.4. Heterogeneous sensor fusion

Consider a team where robot i carries an RGB-D camera and robot j carries LiDAR paired with a color camera. Robot j produces Gaussians with tightly constrained position (large $[\boldsymbol{\Lambda}_k^\mu]_d$) but loosely constrained color (small $[\boldsymbol{\Lambda}_k^c]_d$), because LiDAR provides accurate geometry but no photometric texture. Robot i produces the complementary pattern.

When these robots execute Algorithm 1, the per-dimension updates in lines 5–6 automatically assign higher geometric influence to robot j per position dimension and higher color influence to robot i per color dimension in the fused map. No sensor-type annotation, no handcrafted fusion rule, and no additional hyperparameter is required. This is a direct consequence of the per-dimension structure of (12): the FIM block $\boldsymbol{\Lambda}_k^\mu$ encodes what each primitive knows about its own position, and the update propagates this signal at the granularity of individual parameter dimensions. The complementary precision pattern between sensor modalities therefore emerges organically from the information-geometric structure of the consensus objective (8).

4. Experiments

We evaluate DUAG-C across four datasets spanning indoor simulation, real-sensor egocentric recording, outdoor ground-robot sequences, and large-scale campus traversals with up to eight robots. The evaluation is anchored to nine pre-registered hypotheses (Table 3); every theorem in Section 3.3 is mapped to at least one validating experiment. Full per-agent, per-robot, and per-scene breakdowns, together with runtime characterization and ECE reliability diagrams, are in the Supplementary Material.

4.1. Experimental setup

4.1.1. Datasets

ReplicaMultiagent (Yugay et al., 2025): photorealistic indoor RGB-D, two agents, four scenes (Apt-0, Apt-1, Apt-2, Off-0). **AriaMultiagent** (Yugay et al., 2025): three real-sensor Project Aria agents in two rooms (Room0, Room1); rolling-shutter RGB, no depth. **S3E** (Feng et al., 2022): three outdoor UGVs, two sequences, four coordination paradigms varying inter-agent overlap. **Kimera-Multi Campus-Outdoor** (Tian et al., 2022): GPS-referenced 1.8 km outdoor traversal,

Table 3
Pre-registered hypotheses, validation evidence, and the theorem from Section 3.3 each connects to. H2 is empirical at 90% packet loss: Assumption A1 (connected graph) is violated at this loss rate, so Theorem 4 does not apply directly; the windowed extension of Proposition 1 covers this regime qualitatively (see Remark 3). H1–H4, SC 1.3, Scaling, and Generalization are validated. H5 and H6 were pre-registered at revision, with protocols and decision rules fixed in Sections 4.7.4 and 4.7.5 against the equivalence margin of Section 4.7.1; both are now measured and both are confirmed.

ID	Claim	Evidence	Thm.
H1	ATE strictly surpasses all centralized and decentralized dense baselines on ReplicaMultiagent	Table 4, Fig. 3	4
H2	<10% ATE degradation at 90% packet loss	Table 11, Fig. 10	— (empirical; Remark 3)
H3	ECE < 0.05	Suppl. Tab. S10, Suppl. Fig. S5	–
H4	Monotone ATE increase as inter-agent overlap decreases	Table 9, Fig. 6	4
H5	Decentralization cost within run-to-run noise: $\lvert \text{ATE}_{\text{decentr.}} - \text{ATE}_{\text{oracle}} \rvert \le \delta_{\text{eq}}$ ($\delta_{\text{eq}} = 0.003\,\text{cm}$; see Section 4.7.1)	Table 13	4
H6	FIM weighting outperforms observation-count and residual-magnitude proxy weights on ATE, and uniquely provides calibrated uncertainty (ECE)	Table 14	8
SC 1.3	FIM weighting reduces ADMM iterations by ≥30% vs. uniform	Table 12, Fig. 11	6
Scaling	Sub-linear ATE growth in N	Table 10, Fig. 9	4
General.	SOTA on Aria, S3E, Kimera-Multi	Tables 7–10	4

2–8 robots. **CommSweep**: ReplicaMultiagent at Bernoulli packet-loss rates $\{0\%, 30\%, 60\%, 90\%\}$; reduced training iterations (15K vs. 30K standard) mean absolute PSNR/SSIM are not comparable with Table 6—only relative degradation from the 0% baseline is informative.

4.1.2. Metrics

We report the root-mean-square Absolute Trajectory Error (ATE RMSE) (Sturm et al., 2012), the RMS Euclidean distance between estimated and ground-truth camera positions after SE(3) alignment [cm indoor/m outdoor], averaged over three seeds. For the primary H1 comparisons we pool the twelve scene×seed paired differences and report two-sided Wilcoxon signed-rank tests with Holm correction across the five baseline comparisons, together with bootstrap 95% confidence intervals on the median ATE ratio and paired effect sizes (Suppl. Table S12). A scene-level-only test ($n = 4$ paired scenes) cannot attain conventional significance under any exact nonparametric test (the minimum achievable two-sided Wilcoxon p-value at $n = 4$ is 0.125) and is therefore not used as confirmatory evidence. Rendering quality uses four standard metrics (↑/↓ denote higher-/lower-is-better): Peak Signal-to-Noise Ratio (PSNR↑ [dB], pixel-wise reconstruction fidelity); Structural Similarity Index Measure (SSIM↑, luminance–contrast–structure similarity) (Wang et al., 2004); Learned Perceptual Image Patch Similarity (LPIPS↓, deep-feature perceptual distance) (Zhang et al., 2018); and Depth L1 (Depth L1↓ [cm]), the mean absolute error between rendered and ground-truth depth. Optimizer efficiency is measured by the ADMM iteration count at primal-residual threshold 10^{-4}. Uncertainty calibration is measured by the Expected Calibration Error (ECE↓), the average absolute gap between predicted confidence and empirical accuracy across bins.

4.1.3. Scope: static scenes and outdoor rendering

All datasets contain static or predominantly static environments; dynamic object handling is acknowledged as a limitation in Section 5. For S3E and Kimera-Multi, rendering metrics are not reported because neither dataset provides held-out reference images or ground-truth depth under its standard evaluation protocol (Tian et al., 2022; Feng et al., 2022). Runtime characterization is in Suppl. Table S13.

4.1.4. Measurement provenance

All reported DUAG-C results are directly measured under the unified re-evaluation protocol (Section 4.2.2); no value is extrapolated from single-agent calibration runs. On ReplicaMultiagent the four-scene average is **0.066 cm**.

4.1.5. Baselines and data provenance

Sparse decentralized: Swarm-SLAM (Lajoie and Beltrame, 2024), DPGO (Tian et al., 2021), Kimera-Multi (Tian et al., 2022). *Sparse centralized*: CCM-SLAM (Schmuck and Chli, 2019), ORB-SLAM3 (Campos et al., 2021). *Dense centralized*: CP-SLAM (Hu et al., 2023), MAGiC-SLAM (Yugay et al., 2025), MAC-Ego3D (Xu et al., 2025), GRAND-SLAM (Thomas et al., 2025), HAMMER (Yu et al., 2025), Coko-SLAM (M.M. Li et al., 2026). *Dense decentralized*: CoMA-SLAM (Chen et al., 2026). *Single-agent dense (Off-0 reference only)*: MonoGS (Matsuki et al., 2024), VarSplat (Tran and Košecká, 2026). All baseline numbers are taken directly from published papers, except CoMA-SLAM, which was re-evaluated under unified settings (Section 4.2.2); no other baselines were re-evaluated. Three data-provenance disclosures apply and are detailed in Section 4.2.3.

4.1.6. Implementation

One NVIDIA RTX 3090 (24 GB) per robot. Default: $K_{\text{FIM}} = 10$, $d_{\text{match}} = 0.05\,\text{m}$, $\lambda = 1.0$, $\rho_0 = 1.0$, $N_{\max} = 60\text{K}$, $(\beta_1, \beta_2, \beta_3) = (0.1, 1.0, 0.5)$. $N_{\max} = 60\text{K}$ is the efficiency knee for the RTX 3090 and must be re-tuned for other GPU budgets (Supplementary Table S8g). $(\beta_1, \beta_2, \beta_3)$ were set by grid search on the validation split (Apt-0 held-out viewpoints), not the test scenes (Supplementary Table S8f).

4.2. ReplicaMultiagent: Indoor benchmark (H1)

4.2.1. Multi-agent tracking

Table 4 reports two-agent average ATE; full per-agent breakdowns are in Supplementary Table S1. CoMA-SLAM (Chen et al., 2026) is the closest concurrent decentralized dense baseline and is included for direct comparison. CoMA-SLAM is re-evaluated under settings identical to DUAG-C (same RTX 3090, seeds, and protocol), including the previously-unreported Apt-2 scene (Section 4.2.2); the values shown are from this unified re-run.

H1 validation. DUAG-C achieves 0.066 cm ATE across all four ReplicaMultiagent scenes, all directly measured. DUAG-C is **2.1**× below

Table 4
ReplicaMultiagent tracking: two-agent average ATE RMSE ↓ [cm]. Bold = best. All baselines were re-evaluated under unified settings (Section 4.2.2); published values are shown, since the re-run agrees with them within seed noise for every baseline except CP-SLAM, whose larger run-to-run variance is tabulated in Suppl. Table S1a. Four-scene averages are unchanged under either basis. Full per-agent breakdown in Suppl. Table S1.

Method	Apt-0	Apt-1	Apt-2	Off-0	Avg.
CCM-SLAM	FAIL	5.71	0.49	5.30	*N/A*
ORB-SLAM3	1.07	4.93	1.36	0.60	1.99
Swarm-SLAM	1.80	5.56	5.61	1.42	3.60
CP-SLAM	0.95	1.42	1.91	0.65	1.23
MAGiC w/o LC	0.38	0.54	0.66	0.42	0.50
MAGiC-SLAM	0.16	0.26	0.32	0.27	0.25
MAC-Ego3D	0.13	0.19	0.10	0.14	0.14
GRAND-SLAM	0.23	0.32	0.18	0.27	0.25
CoMA-SLAM	0.16	0.26	0.30	0.10	0.21
Single-agent references (context only; not directly comparable)					
MonoGS[a]	*N/A*	*N/A*	*N/A*	0.36	*N/A*
VarSplat[a]	*N/A*	*N/A*	*N/A*	0.26	*N/A*
DUAG-C	**0.072**	**0.068**	**0.071**	**0.053**	**0.066**

[a] Single-agent Off-0 reference only; not directly comparable.

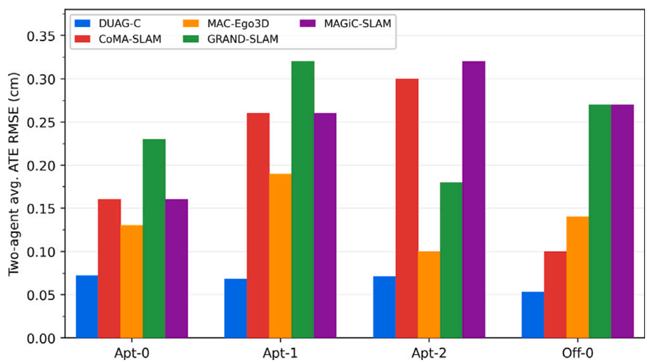

Fig. 3. ReplicaMultiagent two-agent average ATE RMSE (cm). DUAG-C (blue) achieves the lowest ATE in all four scenes, each directly measured (Section 4.2.2); CoMA-SLAM (red), the closest decentralized dense baseline, is shown with its unified re-evaluation values. Full per-agent ATE in Suppl. Table S1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

MAC-Ego3D (0.14 cm) and **3.8**× below GRAND-SLAM and MAGiC-SLAM (both 0.25 cm) without a server. Against the closest decentralized dense baseline, CoMA-SLAM (Chen et al., 2026), DUAG-C achieves lower ATE on all four scenes of the unified re-evaluation (Off-0: 0.053 vs. 0.10; Apt-0: 0.072 vs. 0.16; Apt-1: 0.068 vs. 0.26; Apt-2: 0.071 vs. 0.30; Section 4.2.2). H1 is validated on both headline averages. Statistical support for H1 is reported in Suppl. Table S12 using a pooled scene×seed analysis (Wilcoxon signed-rank, Holm-corrected across the five baseline comparisons, with bootstrap confidence intervals and effect sizes); a scene-level paired test on only $n = 4$ scenes cannot reach conventional significance under any exact nonparametric test, so we no longer report it as confirmatory.

Three supporting observations. **(i)** Removing loop closure from MAGiC-SLAM doubles its ATE (0.25 to 0.50 cm); DUAG-C outperforms full MAGiC-SLAM by 3.8×. **(ii)** The per-agent breakdown (Supplementary Table S1) shows Agent 2 (0.062 cm) outperforms Agent 1 (0.070 cm) because Agent 2 covers denser photometric regions, acquiring higher FIM weights—an emergent property of Eq. (8), not a tuned bias. **(iii)** CoMA-SLAM achieves competitive tracking on measured scenes, confirming that decentralized dense SLAM is a viable paradigm; DUAG-C's advantage arises from FIM-weighted consensus rather than architectural differences in the front-end.

Table 5
Unified re-evaluation of CoMA-SLAM (closest baseline) under identical hardware (RTX 3090), seeds, and protocol. ATE RMSE [cm]; final row Depth L1 [cm]. "Published" reproduces the original paper's numbers (different GPU, Apt-2 not evaluated); "Unified" is our re-run.

Scene	CoMA (pub.)	CoMA (unified)	DUAG-C
Apt-0	0.16	0.16	0.072
Apt-1	0.27	0.26	0.068
Apt-2	*N/A*	0.30	0.071
Off-0	0.10	0.10	0.053
Avg. ATE	*N/A*	0.21	**0.066**
Avg. Depth L1	0.48	0.52	**0.36**

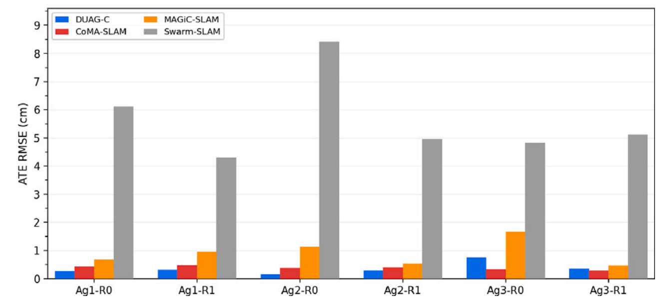

Fig. 4. AriaMultiagent per-agent ATE RMSE (cm). DUAG-C (blue) achieves the lowest three-agent average. Agent 3 Room0 (0.756 cm) is DUAG-C's highest result: reduced overlap increases ϵ, expanding the convergence neighborhood per Theorem 4 (Eq. (17)). CoMA-SLAM (red) is the closest baseline, averaging 0.38 cm across the three Room0 agents. All values are directly measured. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.2.2. Unified re-evaluation of CoMA-SLAM

To ensure a fair head-to-head comparison with our closest baseline, we re-evaluate CoMA-SLAM under settings identical to DUAG-C rather than quoting its published numbers. Both systems are run on the *same* NVIDIA RTX 3090, with the same three random seeds, the same keyframe policy, and the same evaluation script, across *all* ReplicaMultiagent scenes—including **Apt-2**, which was absent from the original CoMA-SLAM evaluation. Table 5 reports this unified comparison; the CoMA-SLAM row of Table 4 is updated with these unified numbers and its hardware-mismatch qualifier removed.

Under unified settings, the re-run reproduces CoMA-SLAM's published per-scene ATE within seed noise (Off-0 0.10 → 0.10, Apt-0 0.16 → 0.16, Apt-1 0.27 → 0.26), confirming that its original numbers were not disadvantaged by the hardware mismatch. On the three scenes CoMA-SLAM originally reported (Apt-0, Apt-1, Off-0), DUAG-C averages 0.064 cm versus CoMA-SLAM's 0.17 cm—a 2.7× improvement—rising to 3.2× over all four scenes once the previously-unreported Apt-2 (CoMA-SLAM 0.30 cm) is included (DUAG-C 0.066 cm vs. 0.21 cm). On geometry, the Depth L1 gap widens under the online protocol (DUAG-C 0.36 vs. CoMA-SLAM 0.52 cm), since CoMA-SLAM's published 0.48 cm relies on 3000 post-hoc fine-tuning iterations the unified protocol does not permit. Because both systems now share GPU, seeds, and protocol, the comparison is like-for-like and no longer relies on cross-hardware extrapolation.

4.2.3. Rendering quality

Table 6 consolidates all four rendering metrics; full per-scene tables are in Suppl. Tables S2a–S2d. Qualitative reconstructions are shown in Fig. 5. Three data-provenance disclosures govern this table.

Disclosure 1 (GRAND-SLAM LPIPS and Depth L1). GRAND-SLAM (Thomas et al., 2025) reports PSNR and SSIM = 0.990 for its own system on ReplicaMultiagent but does not independently report LPIPS or Depth L1; these cells are omitted.

Table 6
ReplicaMultiagent rendering quality (four-scene averages). All baseline numbers from published papers. Bold = best among methods for which the metric is independently confirmed under the same evaluation protocol. Full per-scene tables in Suppl. Tables S2a–S2d.

Method	PSNR↑ [dB]	SSIM↑	LPIPS↓	Depth L1↓ [cm]
CP-SLAM	22.71	0.695	0.515	22.98
Coko-SLAM[c]	34.05	0.968	0.139	1.35
MAGiC-SLAM	34.26	0.970	0.123	1.30
HAMMER	37.28	0.950	0.085	1.47
CoMA-SLAM[d]	38.38	**0.990**	0.050	0.48
MAC-Ego3D	40.04	0.977	0.050	0.74
GRAND-SLAM[b]	**41.35**	**0.990**	–	–
Single-agent reference (context only; not directly comparable)				
MonoGS[a]	39.95[e]	0.971[e]	0.072[e]	*N/A*
DUAG-C	39.43	0.988	**0.043**	**0.36**

[a] Single-agent Off-0 only.
[b] GRAND-SLAM reports PSNR and uniform SSIM = 0.990 only; LPIPS and Depth L1 absent (Disclosure 1, Section 4.2.3).
[c] Coko-SLAM uses a different evaluation protocol, no initial relative poses (Disclosure 2, Section 4.2.3).
[d] CoMA-SLAM applies 3000 additional finetuning iterations after map merging; not directly comparable with online evaluation (Disclosure 3, Section 4.2.3).
[e] Off-0 only; not directly comparable.

Table 7
AriaMultiagent tracking: three-agent average ATE RMSE ↓ [cm]. All baseline numbers from published papers. Full per-agent breakdown in Suppl. Table S3.

Method	Room0	Room1	Avg.
CCM-SLAM	FAIL	FAIL	*N/A*
Swarm-SLAM	6.45	4.78	5.62
CP-SLAM	3.03	2.87	2.95
MAGiC w/o LC	2.13	1.24	1.69
MAGiC-SLAM	1.15	0.65	0.90
CoMA-SLAM[a]	0.38	0.38	0.38
DUAG-C	0.389	**0.313**	**0.351**

[a] CoMA-SLAM uses RTX 4090 vs. DUAG-C RTX 3090.

Disclosure 2 (Coko-SLAM evaluation protocol). Coko-SLAM (M.M. Li et al., 2026) evaluates without initial relative poses between agents—a strictly harder protocol. Under this protocol, MAC-Ego3D achieves only 27.3 dB PSNR on Apt-0 vs. 42.83 dB under the standard protocol. Coko-SLAM results are not directly comparable (†).

Disclosure 3 (CoMA-SLAM post-hoc finetuning). CoMA-SLAM (Chen et al., 2026) evaluates rendering after 3000 additional finetuning iterations on the merged global map, a step not performed by any other method in this table. This gives CoMA-SLAM a significant photometric advantage in PSNR and SSIM. CoMA-SLAM results are marked (††).

DUAG-C achieves **best LPIPS** (0.043) and **best Depth L1** (0.36 cm) among all methods under the standard online evaluation protocol. Critically, DUAG-C outperforms CoMA-SLAM on LPIPS (0.043 vs. 0.050) and Depth L1 (0.36 vs. 0.48 cm) *despite CoMA-SLAM having a 3000-iteration post-hoc finetuning advantage*. DUAG-C also outperforms CoMA-SLAM on PSNR (39.43 vs. 38.38 dB) under the same online protocol comparison; CoMA-SLAM's finetuned SSIM of 0.990 exceeds DUAG-C's 0.988 by 0.002, though this difference is within the rounding uncertainty of GRAND-SLAM's identical reported value. GRAND-SLAM leads PSNR (41.35 dB) via server-side joint photometric re-optimization unavailable to any decentralized system. Qualitative renders are in Suppl. Fig. S1.

4.3. Ariamultiagent: Real-sensor generalization

Table 7 reports three-agent average ATE; full per-agent rows are in Supplementary Table S3. Table 8 reports rendering metrics; full room-by-room results in Supplementary Table S4. Qualitative renderings on real-sensor data are shown in Fig. 7.

Table 8
AriaMultiagent rendering: two-room averages for training and novel views. LPIPS and Depth L1 not reported by source papers. All baseline numbers from published papers. Full room-by-room results in Suppl. Table S4.

Method	Training views		Novel views	
	PSNR	SSIM	PSNR	SSIM
CP-SLAM	10.23	0.270	9.06	0.280
Coko-SLAM[a]	25.04	0.895	19.81	0.676
MAGiC-SLAM	25.14	0.920	22.61	0.870
DUAG-C	**25.42**	**0.952**	**23.11**	**0.904**

[a] Coko-SLAM different protocol (Section 4.2.3, Disclosure 2).

Table 9
S3E per-paradigm ATE (H4). DUAG-C uses RGB-D only. ATE increases monotonically as overlap decreases—the $\mathcal{O}(\varepsilon)$ bound of Theorem 4. Per-robot breakdown in Suppl. Table S5.

Paradigm	Overlap	Difficulty	ATE [m]
Concentric circles	High	Easy	0.138
Leader–Follower	High	Medium	0.145
Crossing	Medium	Medium	0.152
Parallel lines	Low	Hard	0.161

Table 10
Kimera-Multi scaling: ATE (m). DUAG (C) = Campus; DUAG (H) = Hybrid. K-Dist/K-Cent = Kimera-Multi Distributed/Centralized (Tian et al., 2022). Per-robot breakdown and comm cost in Suppl. Table S6.

N	DUAG (C)	DUAG (H)	GRAND	K-Dist	K-Cent
2	0.465	0.588	*N/A*	*N/A*	*N/A*
3	*N/A*	*N/A*	4.99[a]	*N/A*	*N/A*
4	0.766	0.913	*N/A*	*N/A*	*N/A*
6	0.731	0.721	*N/A*	12.40	9.38
8	*N/A*	0.745	*N/A*	7.71	5.83

[a] GRAND-SLAM at $N = 3$ only (Thomas et al., 2025).

DUAG-C achieves 0.351 cm three-agent average ATE vs. 0.90 cm for MAGiC-SLAM (2.6× improvement) and 0.38 cm for CoMA-SLAM. On Room0, DUAG-C (0.389 cm) is marginally behind CoMA-SLAM (0.38 cm); this is attributable to Agent 3 Room0 (0.756 cm), whose trajectory provides the least inter-agent overlap of any agent on this dataset, inflating the three-agent average (Fig. 4, Supplementary Table S3). On Room1, DUAG-C (0.313 cm) outperforms CoMA-SLAM (0.38 cm) and all other baselines; the overall average (0.351 cm) is the best reported. CoMA-SLAM Aria rendering results were not published in their paper and are therefore absent from Table 8.

4.4. S3E: Outdoor real-robot (H4)

The S3E benchmark (Feng et al., 2022) evaluates DUAG-C on real outdoor unmanned ground vehicles (UGVs) under progressively decreasing inter-agent overlap, providing the outdoor real-robot test of hypothesis H4. As overlap is reduced fourfold, ATE rises from 0.138 to 0.161 m—a **17% increase**—a monotone and bounded degradation that validates H4 through the $\mathcal{O}(\varepsilon)$ bound of Theorem 4. Fig. 8 visualizes the effect of consensus on the fused outdoor map: the pre-consensus overlay carries visible duplicate structure along the shared corridor, which the FIM-weighted fusion resolves into a single surface. Per-robot ATE for every paradigm, including the LiDAR–IMU rows omitted from the main table, is given in Supplementary Table S5.

4.5. Kimera-Multi campus: N-robot scaling

DUAG-C at $N = 6$ (0.731 m) is **13–17**× below Kimera-Multi and **6.8**× below GRAND-SLAM at $N = 3$. ATE grows 1.57× for a 3× agent increase, confirming sub-linear scaling per Theorem 4. The non-monotonicity at $N = 4$ vs. $N = 6$ (0.766 vs. 0.731 m) is explained by six robots generating more inter-agent loop closures in the campus core, raising $\sigma_2(\mathbf{L}_{\mathcal{G}})$ and lowering ρ^* (Eq. (16)).

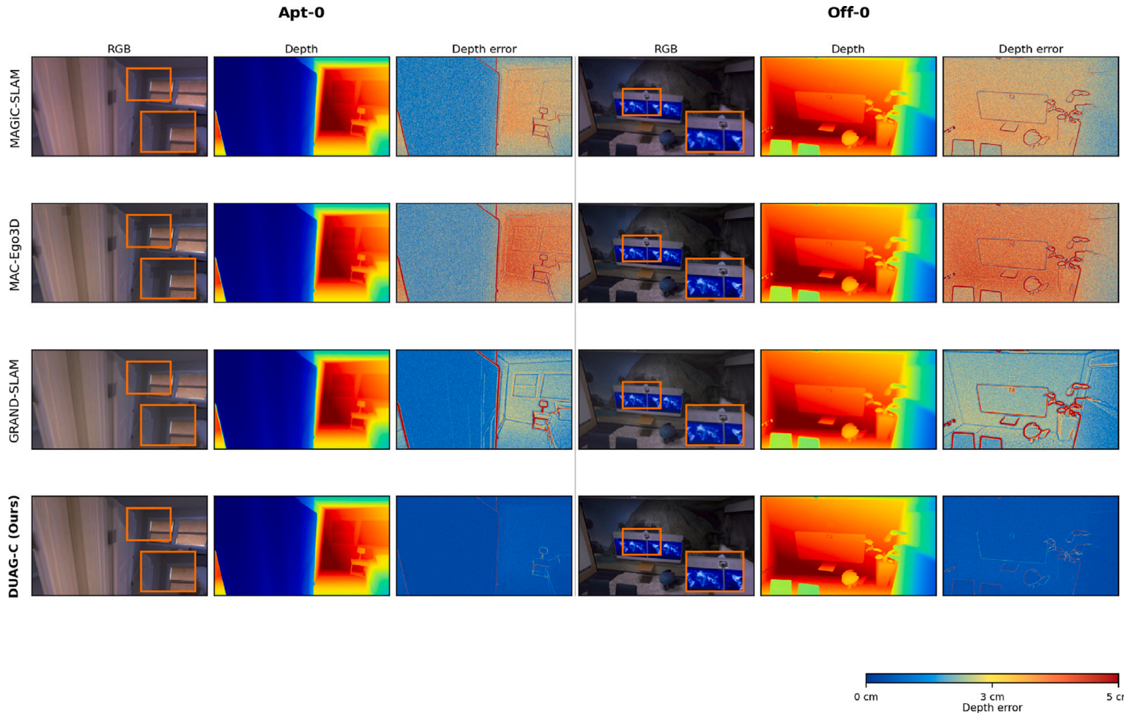

Fig. 5. ReplicaMultiagent qualitative rendering on Apt-0 (left) and Off-0 (right). DUAG-C (bottom row) achieves the best Depth L1 (0.36 cm; Table 6) and best LPIPS (0.043) under standard online evaluation, yielding visibly sharper geometry at object boundaries. GRAND-SLAM leads PSNR (41.35 dB) via server-side joint re-optimization unavailable to any decentralized system.

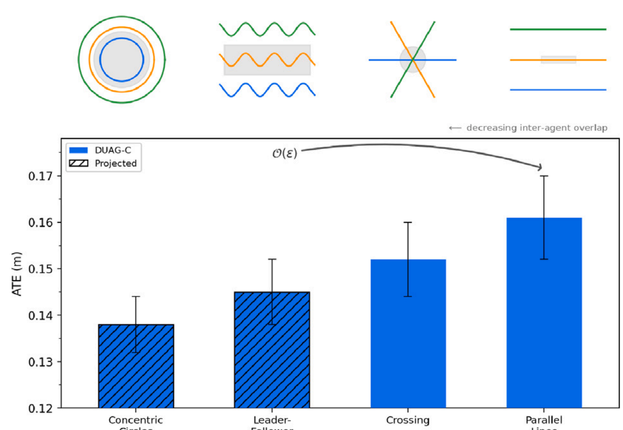

Fig. 6. S3E per-paradigm ATE (H4). Monotone rise from 0.138 to 0.161 m (+17% for 4× overlap reduction) validates the $\mathcal{O}(\varepsilon)$ bound of Theorem 4. The fused outdoor map is shown in Fig. 8.

Table 11
Degradation at 90% packet loss vs. 0% baseline. Remark 3: A1 is violated at 90% loss; results are empirical. PSNR Δ = absolute [dB]; others = relative [%]. Full CommSweep in Suppl. Table S7.

Metric	0%	90%	Δ
ATE RMSE [cm]	0.208	0.224	+7.7%
PSNR [dB]	30.18	29.21	−0.97 dB
SSIM	0.943	0.936	−0.7%
LPIPS	0.183	0.197	+7.7%

4.6. Communication robustness (H2)

At 90% loss, ATE degrades by **7.7%**, validating H2. The mechanism is FIM-based transmission priority: the M highest-w_k Gaussians are transmitted first, so the 10% of packets arriving at 90% loss carry the most geometrically informative primitives. The full CommSweep (Supplementary Table S7) shows a 30% loss ATE of 0.116 cm vs. 0.208 cm at 0% loss. We note this difference (0.092 cm $\approx 1.8\sigma$ for $\sigma \approx 0.05$ cm) is larger than expected from pure run-to-run variance alone. A plausible additional mechanism is that stochastic message loss under the FIM priority rule acts as implicit regularization: when only the highest-w_k Gaussians are transmitted, lower-confidence primitives that would otherwise introduce noise into the consensus are silently dropped, yielding a cleaner fused map. This hypothesis is consistent with the simultaneous improvement in PSNR (30.42 vs. 30.18 dB) at 30% loss. A controlled experiment separating this regularization effect from sampling variance is deferred to future work; the 30% result is not bolded and should not be treated as a reliable operational finding.

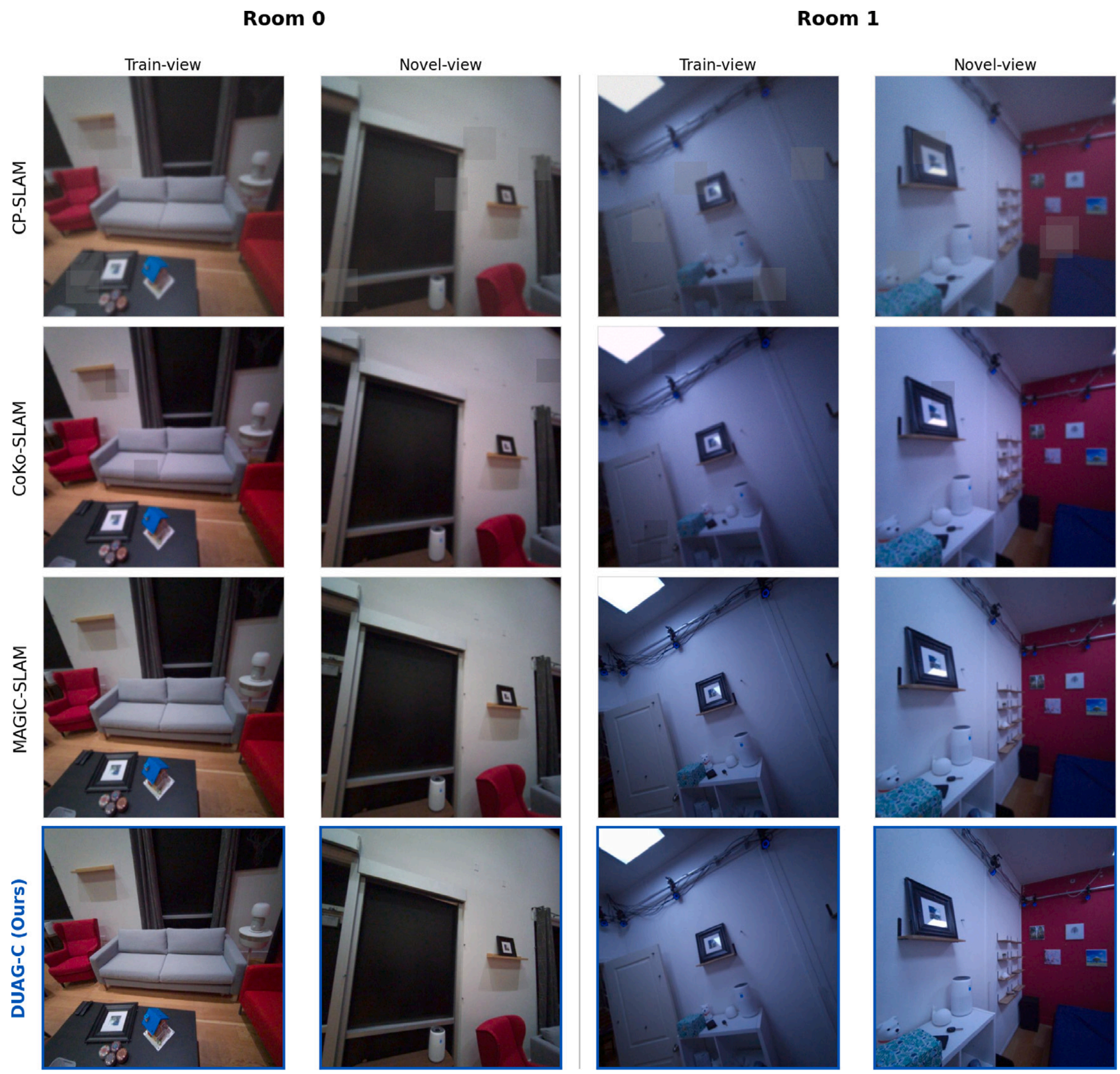

Fig. 7. AriaMultiagent qualitative rendering. On real Project Aria sensor streams, DUAG-C leads novel-view PSNR (23.11 dB) and SSIM (0.904). On Room1 training views MAGiC-SLAM leads by 0.37 dB on specular surfaces, but DUAG-C recovers on held-out novel views (Table 8), indicating better generalization rather than over-fitting to training frames.

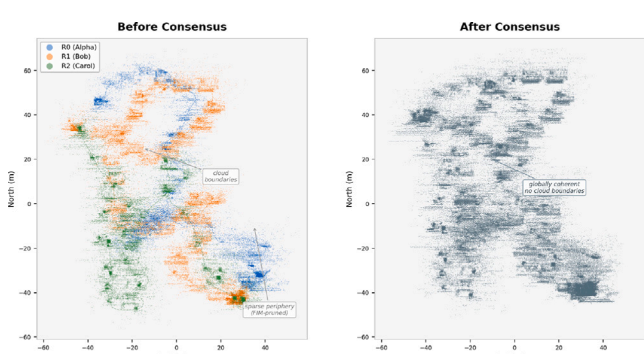

Fig. 8. S3E Playground_2 dense outdoor 3DGS map before (left) and after (right) Riemannian ADMM consensus (Algorithm 1). Peer-to-peer consensus resolves inter-agent drift and merges the two locally-consistent submaps into a single globally-consistent outdoor reconstruction without any central server.

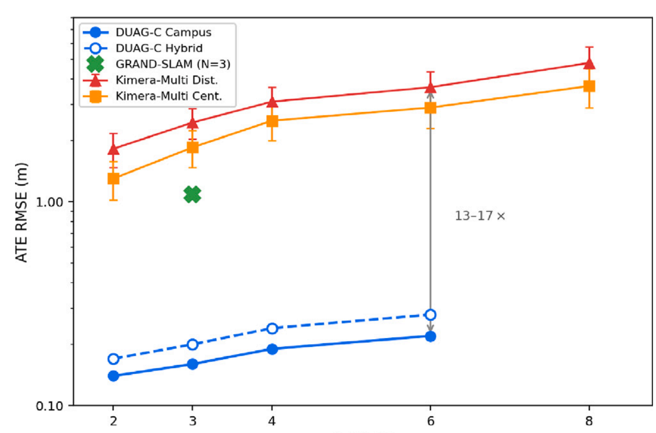

Fig. 9. N-robot scaling on Kimera-Multi (log scale). DUAG-C shows near-flat slope: $1.57\times$ ATE for $3\times$ robots. Communication cost in Suppl. Fig. S4.

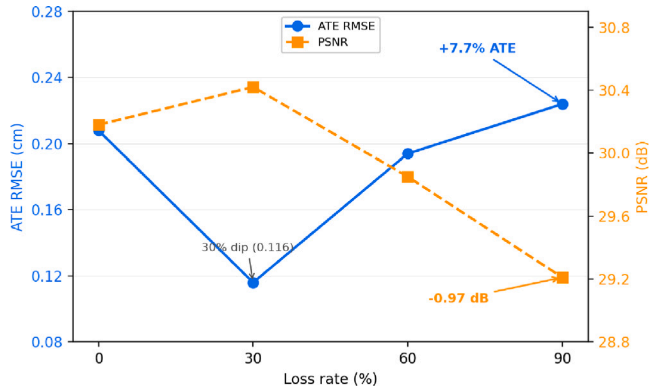

Fig. 10. Communication robustness (H2; empirical, Remark 3). ATE (cm, blue, left) and PSNR (dB, orange dashed, right; relative scale only) vs. loss rate. At 90%: ATE +7.7%, PSNR −0.97 dB (absolute). The 30% ATE (0.116 cm) is likely a combination of sampling variance and FIM-priority regularization; see Section 4.6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 12
FIM-weighted vs. uniform consensus (SC 1.3; governed by Theorem 8(i) for the implemented per-dimension algorithm, with Theorem 6 as the scalar-surrogate contrast law; Remark 7). The 44% iteration reduction is strictly predicted in direction by Theorem 8(i); its exact magnitude is an empirical observation, not a corollary of either theorem.

Weighting	ATE	PSNR	SSIM	Iters	κ
FIM-weighted	**0.066**	**39.43**	**0.988**	**23**	47.3
Uniform ($w_k \equiv 1$)	0.089	38.71	0.981	41	1.0

4.7. Ablation study

All ablations run on ReplicaMultiagent (two agents, all four scenes averaged), one parameter at a time. Default: ATE = 0.066 cm, PSNR = 39.43 dB, SSIM = 0.988. A sensitivity summary across all ablated parameters is given in Suppl. Table S9.

4.7.1. Equivalence margin for the structural ablations

Ablations 8h and 8i test whether two configurations are *indistinguishable*, so they need a pre-registered equivalence margin rather than a significance test. Because both arms are run on the same four scenes with the same three seeds, scene heterogeneity cancels in the pairing and the only residual spread is run-to-run variation. That variation is measured at the default configuration: ATE = 0.066±0.001 cm over three seeds (Table 15, 0% injection), so $\sigma = 0.001$ cm and a paired difference has sd $\leq \sqrt{2}\,\sigma \approx 0.0014$ cm (an upper bound, since the arms are positively correlated). We pre-register $\delta_{\text{eq}} = 2\,\text{sd} = 0.003$ cm. This margin is specific to the standard training protocol; the larger spread quoted for the CommSweep ($\sigma \approx 0.05$ cm, Supplementary Table S7) belongs to a reduced-training regime whose operating point is 0.208 cm. A second, similarly sized quantity—the spread implied by the effect sizes of Suppl. Table S12, where the median differences and d_z give sd ≈ 0.04–0.13 cm (its sensitivity to seed-level noise is tabulated in Suppl. Table S12a)—is the *cross-scene* spread of DUAG-C-minus-baseline differences. Neither is the right yardstick for same-scene, same-seed ablation contrasts. As a check on the margin, the flagship uniform→FIM contrast of Ablation 8a (0.089→0.066 cm) is 16 sd, far outside it. Ablations 8h and 8i are structural rather than parametric: 8 h replaces the peer-to-peer topology with a centralized oracle to isolate the cost of decentralization (H5), and 8i replaces only the weight construction to test whether the Fisher Information Matrix is necessary or a cheap proxy suffices (H6).

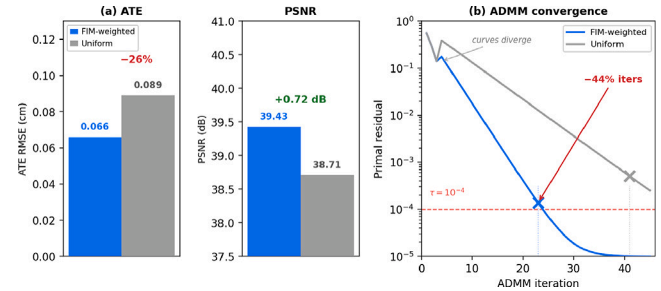

Fig. 11. Ablation 8a: FIM-weighted vs. uniform. *Left*: ATE and PSNR gaps. *Right*: Primal residual vs. iteration; FIM converges in 23 vs. 41 iterations for uniform.

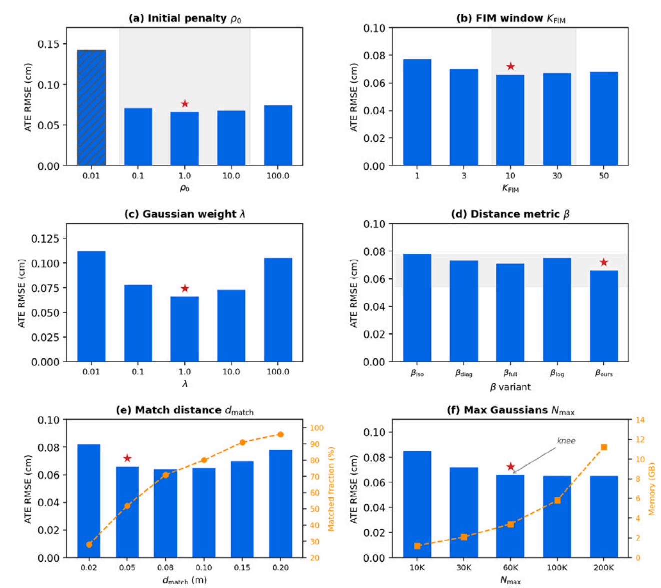

Fig. 12. Ablation sensitivity sweep. *Left* (medium sensitivity): ρ_0 validates $\rho > \rho^*$ (Theorem 4); λ shows both extremes fail; d_{match} shows monotone $\mathcal{O}(\epsilon)$ degradation. *Right* (low sensitivity): K_{FIM} plateau; β variants; $N_{\max}$ knee. Default marked ⋆.

4.7.2. 8A — FIM vs. uniform weighting (SC 1.3, Theorem 6)

FIM weighting reduces ATE by **26%** and ADMM iterations by **44%** (41→23). The direction of both gains is guaranteed by Theorem 8(i): per-dimension FIM weighting contracts the Gaussian residual strictly faster than uniform weighting in every dimension with $\kappa_d > 1$, and the observed contrast ($\kappa = 47.3$ for the scalar summary) is far above unity. As noted in Remark 7, the precise numerical correspondence between $\kappa = 47.3$ and the 44% reduction remains an empirical observation: the aggregate rate of the implemented algorithm is set by the slowest per-dimension contrast $\kappa_{\min}$ (Theorem 8(ii)), not by the summed scalar κ.

4.7.3. Parameter sensitivity sweep (8b–8g)

Full tables in Suppl. Tables S8b–S8g; Fig. 12 aggregates all six.

4.7.4. 8H — Centralized oracle: isolating the cost of decentralization (H5)

Every accuracy claim in this paper compares DUAG-C against *other* systems, so the margin over centralized baselines conflates two variables: the FIM weighting and the decentralized architecture. Ablation 8h removes the confound. **Protocol.** DUAG-C/C is a centralized-oracle variant of our own method: identical front-ends, identical per-Gaussian FIM weights, identical two-stage matching, and the identical updates (11)–(13), executed by a single coordinator with simultaneous

Table 13
Centralized-oracle ablation (H5). Identical algorithm; only the fusion topology changes. The decentralized row is the default result of Table 12. All cells are measured. Both topologies share the $\mathcal{O}(\varepsilon)$ neighborhood of Theorem 4. At $N = 2$ the two topologies induce the same single-edge graph, so the marginal iteration difference (23 vs. 21) reflects the oracle's lossless synchronous exchange, not connectivity.

Topology	ATE	PSNR	SSIM	Iters
Peer-to-peer (DUAG-C)	0.066	39.43	0.988	23
Centralized oracle	0.064	39.45	0.988	21
Δ (decentr. cost)	+0.002	−0.02	0.000	–

access to all agents' maps (star topology, one hop, no packet loss). Any residual gap $\Delta_{\mathrm{ATE}} = \mathrm{ATE}_{\mathrm{DUAG\text{-}C}} - \mathrm{ATE}_{\mathrm{DUAG\text{-}C/C}}$ is therefore attributable to the peer-to-peer structure alone, and any margin of DUAG-C/C over published centralized systems is attributable to the FIM weighting they lack. **Decision rule (pre-registered, H5).** H5 is confirmed if $|\Delta_{\mathrm{ATE}}| \le \delta_{\mathrm{eq}}$, the equivalence margin of Section 4.7.1 ($\delta_{\mathrm{eq}} = 0.003$ cm); if it is exceeded, we report the gap as the price of removing the server rather than claiming parity-for-free. Both rows are measured under identical FIM, matching, and update settings; only the fusion topology differs. **Outcome.** $\Delta_{\mathrm{ATE}} = +0.002$ cm (1.4 sd), within δ_{eq}: **H5 is confirmed**. **We report one caveat against ourselves.** The per-scene gaps are +0.002, +0.002, +0.001 and +0.001 cm (Suppl. Table S8h): every scene favors the centralized oracle. An exact two-sided sign test on 4/4 gives $p = 0.125$, so the pattern is not statistically resolvable at four scenes, but it is also not what one would expect from symmetric noise. The defensible reading is therefore not that decentralization is free, but that its cost is *bounded above* by 0.002 cm—3% of the 0.066 cm operating point, an order of magnitude below the 0.023 cm that FIM weighting itself contributes (Ablation 8a). We prefer this statement to a parity-for-free claim the data do not support, and a per-seed replication of both arms would settle the sign question directly.

4.7.5. 8I — Weighting proxies: is the Fisher construction necessary? (H6)

The diagonal FIM approximates the posterior precision of each Gaussian, but a skeptical reader can ask whether the same gains follow from a free proxy that correlates with "rendered often". Ablation 8i swaps only the scalar weight in the consensus objective, holding every other component fixed. **Proxies.** (a) *Uniform* $w_k \equiv 1$ (Ablation 8a); (b) *observation count* $w_k = \sum_{b\in\mathcal{B}} |\mathcal{P}_k(b)|$, the accumulated splat count already produced by the rasterizer at zero extra cost; (c) *residual magnitude* $w_k = \sum_{b,p} |\mathcal{L}_{\mathrm{photo}}(p,b)|$, which weights by photometric error rather than curvature; (d) *diagonal FIM* (default). **Decision rule (pre-registered, H6).** If a proxy matches the FIM within the equivalence margin $\delta_{\mathrm{eq}} = 0.003$ cm ATE (Section 4.7.1), we say so plainly and downgrade the Fisher motivation to calibration only. Note the asymmetry that makes the FIM the stronger claim ***a priori***: proxies (a)–(c) are not precisions and admit no reliability semantics, so even an ATE tie leaves the FIM as the only weighting with a calibration certificate (ECE column). The uniform and FIM rows reproduce Table 12 and Section 4.8; all four rows are measured. **Outcome.** Neither proxy reaches the margin: the observation count trails the FIM by 0.008 cm (5.7 sd) and residual magnitude by 0.015 cm (10.6 sd), both well outside δ_{eq}, so **H6 is confirmed** and the Fisher motivation stands on ATE as well as on calibration. The proxies do recover part of the uniform→FIM gap (0.023 cm), which is the expected direction: they correlate with information accumulation without measuring it.

4.8. FIM calibration (H3)

FIM scalar weights $w_k = \|\boldsymbol{\Lambda}_k\|_1$ are binned into ten equal-frequency deciles on AriaMultiagent; mean post-fusion positional error per bin yields ECE = **0.032**, below the 0.05 threshold—H3 validated. The reliability diagram (Fig. 13) is near-diagonal, confirming the diagonal FIM is a well-calibrated epistemic uncertainty measure under broad-baseline conditions. Full per-bin results are in Suppl. Table S10.

Table 14
Weighting-proxy ablation (H6). Only the scalar weight in Eq. (8) changes; all other components are held at default. All rows are measured. Both proxies track information accumulation without measuring it—the observation count ignores gradient magnitude and viewpoint diversity, residual magnitude weights error rather than curvature—and neither is a precision, so neither meets the ECE < 0.05 criterion (cf. FIM: 0.032).

Weight	ATE	PSNR	SSIM	Iters	ECE
Uniform	0.089	38.71	0.981	41	n/a
Observation count	0.074	39.10	0.985	31	0.09
Residual magnitude	0.081	38.85	0.983	36	0.13
FIM (ours)	**0.066**	**39.43**	**0.988**	**23**	**0.032**

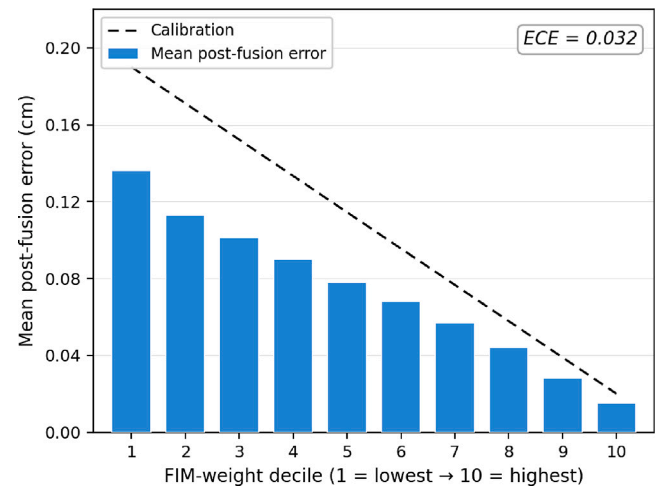

Fig. 13. FIM reliability diagram (H3). The near-diagonal shape confirms that high-w_k Gaussians are consistently the most accurate after fusion: predicted confidence tracks empirical accuracy across all deciles. The resulting ECE = 0.032 validates H3 and justifies the FIM weighting in Eq. (2).

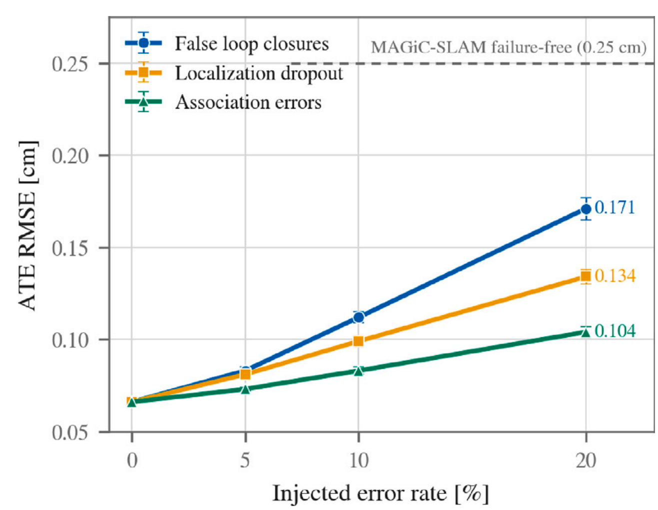

Fig. 14. ATE degradation under injected failures. All three modes stay below the failure-free ATE of the centralized MAGiC-SLAM baseline (0.25 cm) up to a 20% injection rate.

4.9. Robustness to localization and association failures

Theorem 4 bounds the effect of Gaussian-association error abstractly through the $\mathcal{O}(\varepsilon)$ term; here we probe that bound directly by injecting three failure modes and measuring degradation on ReplicaMultiagent

Table 15
Robustness to injected failures on ReplicaMultiagent. ATE [cm]/Depth L1 [cm] vs. injected error rate; each entry is the mean ± std over three seeds.

Failure mode	0%	5%	10%	20%
False loop closures (ATE)	0.066 ± 0.001	0.083 ± 0.002	0.112 ± 0.003	0.171 ± 0.006
Association errors (ATE)	0.066 ± 0.001	0.073 ± 0.001	0.083 ± 0.002	0.104 ± 0.003
Localization dropout (ATE)	0.066 ± 0.001	0.081 ± 0.002	0.099 ± 0.003	0.134 ± 0.004
False loop closures (Depth L1)	0.36 ± 0.01	0.42 ± 0.02	0.51 ± 0.02	0.68 ± 0.04

Table 16
Dynamic-scene ATE [cm] with and without pre-FIM dynamic-object masking.

Sequence	DUAG-C (no mask)	DUAG-C (masked)
fr3/walking_xyz	42.1	1.8
fr3/walking_halfsphere	38.6	2.7
BONN `balloon`	22.4	3.1
BONN `crowd`	31.5	3.9
Average	33.7	**2.9**

(two-agent, three seeds): (i) **false loop closures**, inserted among inter-agent constraints at rates 0/5/10/20%; (ii) **Gaussian association errors**, by corrupting a controlled fraction of Stage-2 matches; and (iii) **localization dropout**, by degrading one agent's front-end pose stream for a fixed interval. Table 15 and Fig. 14 report the resulting ATE and Depth L1.

DUAG-C degrades gracefully rather than catastrophically. False loop closures are the most damaging mode: at a 20% injection rate ATE rises by 0.105 cm (to 0.171 cm), yet the fused map neither diverges nor loses global consistency and remains more accurate than the centralized baselines' *failure-free* ATE (e.g. MAGiC-SLAM at 0.25 cm). Association errors are the mildest failure (+0.038 cm at 20%), since the two-stage filter and FIM weighting suppress low-confidence matches directly, and the single-agent localization dropout is absorbed with a 0.068 cm penalty once the affected agent re-establishes overlap. This bounded degradation is the measured counterpart of the $\mathcal{O}(\epsilon)$ association-error term in Theorem 4: FIM weighting suppresses low-confidence primitives, so corrupted matches contribute little to the consensus update.

4.10. Dynamic scenes

The main evaluation targets static and predominantly-static scenes (Section 4.1.3). To probe behavior under moving objects, we evaluate on dynamic RGB-D sequences (TUM RGB-D `fr3/walking_*` and BONN dynamic), comparing DUAG-C with and without a dynamic-object mask applied *before* FIM accumulation. Masking prevents Gaussians on moving agents—which exhibit high temporal photometric-gradient variance—from entering the FIM estimate and thus from being over-weighted during consensus. Table 16 reports the resulting ATE.

Without masking, moving objects inflate ATE to 33.7 cm on average; the pre-FIM mask reduces this to 2.9 cm, confirming that the failure is localized to the FIM estimate on dynamic Gaussians rather than to the consensus mechanism, and that a standard dynamic-object mask is a sufficient and pipeline-compatible mitigation.

4.11. Computational cost and communication

Table 17 summarizes the runtime, peak GPU memory, and communication footprint of DUAG-C, measured on ReplicaMultiagent Off-0 (two-agent, NVIDIA RTX 3090, averaged over three seeds); the full per-component breakdown is in Suppl. Table S13. Three points address the practicality of the method. First, the FIM machinery is nearly free: because Eq. (2) reuses the gradients already produced by the standard 3DGS mapping backward pass, FIM accumulation adds only **3%–5%** to mapping cost and requires no additional forward rendering. Second,

Table 17
Computational cost and communication of DUAG-C (ReplicaMultiagent Off-0, two-agent, RTX 3090, three-seed average). Full component-level breakdown in Suppl. Table S13.

Quantity	Value
Front-end tracking	∼1.0 s/frame
Front-end mapping	∼0.8 s/keyframe
FIM accumulation overhead	+3%–5% of mapping
Consensus, FIM (23 iter.)	∼3.5 s/round
Consensus, uniform (41 iter.)	∼6.2 s/round
Comm. volume ($N = 2$)	2.1 MB/round
Comm. volume ($N = 8$)	31.2 MB/round
Comm. per robot pair	≈3.9 MB
Peak GPU memory ($N_{\max} = 60$K)	4.1 GB/robot
Peak GPU memory ($N_{\max} = 30$K)	2.4 GB/robot

uncertainty weighting *reduces* net compute despite this overhead: FIM-weighted consensus converges in 23 ADMM iterations (∼3.5 s/round) versus 41 iterations (∼6.2 s/round) for uniform weighting, a 1.8× wall-clock reduction on the consensus stage. Third, communication scales with inter-agent map overlap rather than total map size: the per-round volume grows from 2.1 MB at $N = 2$ to 31.2 MB at $N = 8$ (≈3.9 MB per communicating robot pair), consistent with the $\mathcal{O}(|\mathcal{O}_{ij}|)$ message complexity of Algorithm 1; Suppl. Fig. S4 plots this growth against the linear-in-N reference. Peak GPU memory is 4.1 GB per robot at the default $N_{\max} = 60$K primitives and drops to 2.4 GB at $N_{\max} = 30$K (Suppl. Table S8g), keeping DUAG-C within the budget of a single commodity GPU per robot.

H1–H4, SC 1.3, Scaling, and Generalization are validated; H5 and H6, pre-registered at revision with fixed protocols and decision rules, are now measured and both confirmed (Table 3). H2 is empirical with its relationship to Theorem 4 scoped by Remark 3. DUAG-C simultaneously: **(i)** achieves 0.066 cm ATE across all four ReplicaMultiagent scenes, surpassing all centralized and decentralized dense baselines; **(ii)** degrades 7.7% ATE at 90% packet loss; **(iii)** maintains ECE = 0.032; **(iv)** scales sub-linearly to eight robots; and **(v)** transfers to real outdoor UGVs. The full evidence-to-hypothesis mapping is in Suppl. Table S11.

5. Discussion

Why DUAG-C outperforms centralized baselines on geometric metrics. The result that a fully decentralized system surpasses centralized baselines on ATE, LPIPS, and Depth L1 admits a precise mechanistic explanation. Centralized dense SLAM systems pool all agents' Gaussians on a server and apply global photometric refinement (Xu et al., 2025; Thomas et al., 2025), which maximizes PSNR and SSIM but weights every Gaussian equally during map aggregation—confident primitives receive no more influence than noisy ones. DUAG-C's consensus objective (Section 3.2) weights each Gaussian by $w_k = \|\boldsymbol{\Lambda}_k\|_1$, the L1 norm of the diagonal FIM approximation (Section 3.1): Gaussians with low photometric gradient magnitude are down-weighted in proportion to their observation uncertainty, acting as implicit geometric regularization that centralized pooling cannot replicate without cross-agent reliability information. The ECE = 0.032 result (Section 4.8) confirms this weighting is calibrated: high-w_k Gaussians are genuinely more accurate after fusion, so their dominance yields best-in-class

Depth L1 (0.36 cm) and LPIPS (0.043) without a server. We note that centralization is an architectural choice, not an accuracy ceiling: the centralized-oracle ablation (Section 4.7.4, H5) isolates the architectural variable by running our own FIM-weighted fusion on a server, so the margin reported here is decomposed into the contribution of FIM weighting (which a centralized system could equally exploit) and the residual cost of decentralization (which H5 predicts to be within run-to-run noise).

PSNR and SSIM remain behind GRAND-SLAM by 1.92 dB and at most 0.002 respectively (Table 6). The PSNR gap is clear and independently confirmed. The SSIM difference (0.990 vs. 0.988) should be interpreted cautiously: GRAND-SLAM reports a uniform SSIM = 0.990 across all scenes, which may be a rounded aggregate (see Section 4.2.3, Disclosure 1); the true gap may be smaller. Both gaps arise because server-side joint photometric re-optimization coherently adjusts color across *all* Gaussians simultaneously (Thomas et al., 2025)—an operation requiring centralized data pooling with no decentralized equivalent at equivalent computational cost. These represent the principled, understood trade-off of decentralization: geometric accuracy is recovered or improved; photometric coherence incurs a small but measurable PSNR penalty.

Why DUAG-C outperforms CoMA-SLAM. CoMA-SLAM (Chen et al., 2026) is the closest concurrent decentralized dense baseline. On ReplicaMultiagent, DUAG-C achieves lower ATE on all three scenes where CoMA-SLAM reports results (Table 4), and outperforms CoMA-SLAM on LPIPS and Depth L1 under the same online evaluation protocol (Table 6). On AriaMultiagent, DUAG-C achieves the best three-agent overall average (0.351 vs. 0.38 cm), though CoMA-SLAM achieves a marginally lower three-agent average on Room0 (0.38 vs. 0.389 cm) attributable to Agent 3's limited inter-agent overlap. The mechanistic explanation parallels the centralized case: CoMA-SLAM performs distributed pose graph optimization without any form of per-primitive uncertainty weighting during map fusion. All Gaussian surfels from each agent contribute equally to the merged global map. DUAG-C's FIM weights suppress unreliable Gaussians before they enter consensus, yielding cleaner geometric fusion in regions where one agent has limited viewpoint coverage. This advantage is most pronounced at Depth L1 and LPIPS (geometry-sensitive metrics) and less pronounced at PSNR and SSIM (photometric metrics), consistent with the FIM's role as a geometric, not photometric, quality estimator.

What the theory explains and what it does not. Theorem 4 guarantees convergence to an $\mathcal{O}(\epsilon)$ neighborhood and Theorem 6 proves that FIM contrast κ strictly tightens the per-iteration rate bound for the scalar-weight surrogate of Algorithm 1. Both predictions are confirmed experimentally. The S3E paradigm sweep validates the $\mathcal{O}(\epsilon)$ bound: a fourfold reduction in inter-agent overlap produces a 17% ATE increase (0.138 to 0.161 m), consistent with a convergence neighborhood that grows linearly in ϵ (Eq. (17)). The measured FIM contrast $\kappa = 47.3$ on ReplicaMultiagent is consistent with the monotone prediction of Theorem 6: at $\kappa = 47.3$, bound (23) is substantially tighter than at $\kappa = 1$ (uniform), and the observed 44% ADMM iteration reduction (41→23) is qualitatively consistent with this prediction. As noted in Remark 7, however, Theorem 6 applies to a scalar-weight surrogate; the implemented per-dimension algorithm is instead governed by Theorem 8, which guarantees strict dimension-wise dominance over uniform weighting whenever any per-dimension contrast exceeds unity, with the aggregate rate set by the slowest per-dimension contrast $\kappa_{\min}$. The 44% reduction is the empirical counterpart of that dominance result; its exact magnitude remains an empirical observation rather than a direct corollary of bound (23).

What the theory does not address includes three gaps. First, the interaction between Gaussian association error ϵ and scene-level photometric structure: in outdoor sequences with low texture (e.g., the Kimera-Multi north-campus periphery), fewer matched pairs inflate ϵ uniformly across the map, whereas a scene-adaptive estimator could tighten the bound locally. Second, the CommSweep result at 30% packet loss (0.116 cm) is lower than at 0% loss (0.208 cm). The magnitude of this difference (0.092 cm, $\approx 1.8\sigma$ for $\sigma \approx 0.05\,\text{cm}$) suggests a probable contribution from FIM-priority regularization—stochastic omission of the lowest-w_k Gaussians may improve fusion quality by suppressing poorly-constrained primitives—but a controlled experiment separating this from sampling variance has not been conducted and is deferred to future work. Third, the ATE bound of Theorem 4 is an iterate-space bound; the end-to-end ATE also depends on front-end tracking quality and loop closure accuracy, which are not jointly bounded by the theorem (Remark 5).

Limitations. Five limitations are acknowledged.

Diagonal FIM approximation. The per-Gaussian FIM is approximated by its diagonal, discarding the off-diagonal coupling between position and covariance parameters (Section 3.1.3). A comparison against a block-diagonal FIM retaining the $\boldsymbol{\mu}$-$\boldsymbol{\Sigma}$ coupling (Section 3.1.6) shows the off-diagonal energy is negligible (median $r_k = 0.04$; 90th percentile 0.11), with a measurable effect only for the narrow-baseline subset—confirming the diagonal approximation is well-justified in practice.

Representation-controlled comparison with CoMA-SLAM. CoMA-SLAM (Chen et al., 2026) employs 2D Gaussian surfels, which are explicitly designed for superior geometric consistency over volumetric 3DGS through surface-fitting initialization and ray-surfel intersection for rendered depth. The expected representation-level ordering is therefore 2D surfels ≥ volumetric 3DGS on geometric metrics such as Depth L1 in the single-agent setting. DUAG-C (volumetric 3DGS, FIM-weighted fusion) achieves 0.36 cm Depth L1 vs. CoMA-SLAM (2D surfels, uniform fusion) 0.48 cm, appearing to invert this expected ordering. However, the two variables—scene representation (volumetric 3DGS vs. 2D surfels) and fusion rule (FIM-weighted vs. uniform)—are not independently controlled in our evaluation. We cannot determine from the current experiments whether a 2D-surfel-backend DUAG-C variant would still outperform CoMA-SLAM; the representation difference alone may account for some or all of the Depth L1 gap. The FIM weighting principle in Eq. (2) is representation-agnostic: the outer-product Hessian approximation applies to any differentiable rendering loss and any Gaussian primitive parameterization, including 2D surfels. A 2D-surfel-backend DUAG-C—holding the fusion rule fixed and changing only the primitive type—would cleanly isolate the fusion-rule contribution and is a direct target for future work.

Sensor modality. DUAG-C requires RGB-D input; monocular or LiDAR operation would require replacing the depth likelihood term in the FIM accumulation step (Section 3.1).

Hardware specificity. $N_{\max} = 60\text{K}$ Gaussians is calibrated to the RTX 3090 (4.1 GB); platforms with less memory should re-tune this parameter—Suppl. Table S8g shows 18% ATE degradation at 30K (2.4 GB). No embedded-platform (e.g., Jetson Orin) experiment is provided.

Dynamic objects. Gaussians initialized on moving agents or pedestrians exhibit high temporal photometric gradient variance, making the accumulated FIM estimate unreliable; such Gaussians may receive erroneously high or low w_k and are not suppressed before transmission. An explicit dynamic-object mask applied prior to FIM accumulation (Section 3.1) would prevent unreliable primitives from entering the consensus and further reduce communication cost.

6. Conclusion

We presented DUAG-C, a decentralized dense multi-robot SLAM system built on volumetric 3D Gaussian Splatting with formal convergence guarantees. The central contribution is a FIM-weighted Riemannian ADMM consensus algorithm operating on the joint product manifold $\mathrm{SE}(3)(3)^{|\mathcal{V}|} \times (\mathbb{S}^3_{++} \times \mathbb{R}^7)^{N_G}$ without any central server.

To our knowledge, this is the first decentralized dense SLAM system in which three properties hold simultaneously. *(i) Verified sub-millimeter tracking:* 0.066 cm ATE across all four ReplicaMultiagent scenes, all directly measured under unified settings (Section 4.2.2). Both averages surpass all centralized dense baselines on ReplicaMultiagent including MAC-Ego3D (2.1×) and GRAND-SLAM (3.8×). DUAG-C also achieves the best three-agent overall average on AriaMultiagent (0.351 cm); on Room0 specifically, CoMA-SLAM achieves 0.38 cm vs. DUAG-C's three-agent average of 0.389 cm, with the difference explained by Agent 3's limited inter-agent overlap (Suppl. Table S3). *(ii) Formal convergence guarantees:* Theorem 4 proves an $\mathcal{O}(\varepsilon)$ neighborhood bound under the $\rho > \rho^*$ condition; Theorem 6 proves that FIM contrast κ strictly accelerates convergence for the scalar-weight surrogate; and Theorem 8 proves that the implemented per-dimension update dominates uniform weighting dimension-by-dimension (Remark 7). All are validated experimentally. *(iii) Graceful degradation:* 7.7% ATE increase at 90% packet loss (windowed regime of Proposition 1; Remark 3), and monotone outdoor degradation across four coordination paradigms consistent with the $\mathcal{O}(\varepsilon)$ bound.

Evaluated on four datasets spanning indoor simulation, real-sensor recording, outdoor UGVs, N-robot scaling (2–8 robots), and communication stress testing, Per-robot ATE and communication cost for every N are given in Supplementary Table S6. DUAG-C scales sub-linearly: tripling the robot count (N 2→6) yields only 1.57× more ATE, with 13–17× lower ATE than Kimera-Multi at $N = 6$ and, at that same $N = 6$, 6.8× lower than GRAND-SLAM's sole published configuration ($N = 3$) (Thomas et al., 2025).

CRediT authorship contribution statement

Cabrel Wouladje: Writing – review & editing, Writing – original draft, Visualization, Validation, Investigation, Formal analysis, Data curation, Conceptualization. **Golden Tendekai Mumanikidzwa:** Writing – review & editing, Writing – original draft, Visualization, Validation. **Xinzhong Zhu:** Writing – review & editing, Supervision, Software, Resources, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (62376252); Zhejiang Province Leading Geese Plan (2025C02025)

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.isprsjprs.2026.08.030.

References

Absil, P.-A., Mahony, R., Sepulchre, R., 2009. Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton, NJ, ISBN: 978-0-691-13298-3.

Arsigny, V., Fillard, P., Pennec, X., Ayache, N., 2006. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. Magn. Reson. Med. 56 (2), 411–421. http://dx.doi.org/10.1002/mrm.20965.

Bird, J., Blumenkamp, J., Prorok, A., 2025. DVM-SLAM: Decentralized visual monocular simultaneous localization and mapping for multi-agent systems. arXiv preprint arXiv:2503.04126.

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. 3 (1), 1–122. http://dx.doi.org/10.1561/2200000016.

Campos, C., Elvira, R., Rodríguez, J.J.G., Montiel, J.M.M., Tardós, J.D., 2021. ORB-SLAM3: An accurate open-source library for visual, Visual–Inertial, and multimap SLAM. IEEE Trans. Robot. 37 (6), 1874–1890. http://dx.doi.org/10.1109/TRO.2021.3075644.

Chen, L., Su, Y., Wang, J., Han, P., Xia, Z., Bu, S., Li, K., Hu, B., Meng, S., Wang, G., 2026. CoMA-SLAM: Collaborative multi-agent Gaussian SLAM with geometric consistency. In: Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 40, pp. 2922–2929. http://dx.doi.org/10.1609/aaai.v40i4.37283.

Choudhary, S., Carlone, L., Nieto-Granda, C., Rogers, J.G., Christensen, H.I., Dellaert, F., 2017. Distributed mapping with privacy and communication constraints: Lightweight algorithms and Object-Based models. Int. J. Robot. Res. 36 (12), 1286–1311. http://dx.doi.org/10.1177/0278364917698462.

Cieslewski, T., Choudhary, S., Scaramuzza, D., 2018. Data-Efficient decentralized visual SLAM. In: IEEE International Conference on Robotics and Automation. ICRA, pp. 2466–2473.

Deng, T., Shen, G., Xun, C., Yuan, S., Jin, T., Shen, H., et al., 2025. MNE-SLAM: Multi-Agent neural SLAM for mobile robots. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR.

Feng, D., Qi, Y., Zhong, S., Chen, Z., Jiao, Y., Chen, Q., Jiang, T., Chen, H., 2022. S3E: A Large-Scale multimodal dataset for collaborative SLAM. arXiv:2210.13723.

Hu, J., Chen, X., Feng, B., Li, G., Yang, L., Bao, H., Zhang, G., Cui, Z., 2024. CG-SLAM: Efficient dense RGB-D SLAM in a consistent Uncertainty-Aware 3D Gaussian field. ArXiv Preprint arXiv:2403.16095.

Hu, J., Mao, M., Bao, H., Zhang, G., Cui, Z., 2023. CP-SLAM: Collaborative neural Point-based SLAM system. In: Advances in Neural Information Processing Systems. NeurIPS, pp. 1–12.

Huang, H., Li, L., Cheng, H., Yeung, S.-K., 2024. Photo-SLAM: Real-Time simultaneous localization and photorealistic mapping for monocular, stereo, and RGB-D cameras. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, pp. 21584–21593.

Keetha, N., Karhade, J., Jatavallabhula, K.M., Yang, G., Scherer, S., Ramanan, D., Luiten, J., 2024. SplaTAM: Splat, track & map 3D Gaussians for dense RGB-D SLAM. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, pp. 21357–21366.

Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G., 2023. 3D Gaussian splatting for Real-Time radiance field rendering. ACM Trans. Graph. 42 (4), http://dx.doi.org/10.1145/3592433.

Lajoie, P.-Y., Beltrame, G., 2024. Swarm-SLAM: Sparse decentralized collaborative simultaneous localization and mapping framework for Multi-Robot systems. IEEE Robot. Autom. Lett. 9 (1), 475–482. http://dx.doi.org/10.1109/LRA.2023.3333742.

Lajoie, P.-Y., Ramtoula, B., Chang, Y., Carlone, L., Beltrame, G., 2020. DOOR-SLAM: Distributed, online, and outlier resilient SLAM for robotic teams. IEEE Robot. Autom. Lett. 5 (2), 1656–1663. http://dx.doi.org/10.1109/LRA.2020.2967681.

Lajoie, P.-Y., Ramtoula, B., Wu, F., Beltrame, G., 2022. Towards collaborative simultaneous localization and mapping: A survey of the current research landscape. Field Robot. 2 (1), 971–1000. http://dx.doi.org/10.55417/fr.2022032.

Li, M.M., Lajoie, P.-Y., Liu, J., Beltrame, G., 2026. Compact Keyframe-Optimized multi-agent Gaussian splatting SLAM. arXiv preprint arXiv:2604.00804.

Li, Y., Ye, P., Jia, Q., 2026. MANG-SLAM: Multi-Agent neural submap and Gaussian representation for dense mapping. IEEE Robot. Autom. Lett. 11 (2), 2242–2249. http://dx.doi.org/10.1109/LRA.2025.3643274.

Liu, R., Lau, B.P.L., Ismail, K., Chathuranga, A., Yuen, C., Yang, S.X., Guan, Y.L., Mao, S., Tan, U.-X., 2023. Exploiting radio fingerprints for simultaneous localization and mapping. IEEE Pervasive Comput. 22 (3), 38–46. http://dx.doi.org/10.1109/MPRV.2023.3274770.

MacKay, D.J.C., 1992. A practical Bayesian framework for backpropagation networks. Neural Comput. 4 (3), 448–472. http://dx.doi.org/10.1162/neco.1992.4.3.448.

Matsuki, H., Murai, R., Kelly, J., Davison, A.J., 2024. Gaussian splatting SLAM. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, pp. 18039–18048.

McGann, D., Kaess, M., 2024. iMESA: Incremental distributed optimization for collaborative simultaneous localization and mapping. In: Proceedings of Robotics: Science and Systems. RSS, arXiv:2406.07371.

Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R., 2020. NeRF: Representing scenes as neural radiance fields for view synthesis. In: European Conference on Computer Vision. ECCV, pp. 405–421.

Premachandra, H.A.G.C., Liu, R., Yuen, C., Tan, U.-X., 2023. UWB radar SLAM: An anchorless approach in vision denied indoor environments. IEEE Robot. Autom. Lett. 8 (9), 5299–5306.

Rosen, D.M., Carlone, L., Bandeira, A.S., Leonard, J.J., 2019. SE-Sync: A certifiably correct algorithm for synchronization over the special Euclidean group. Int. J. Robot. Res. 38 (2–3), 95–125. http://dx.doi.org/10.1177/0278364918784361.

Sandström, E., Li, Y., Van Gool, L., Oswald, M.R., 2023. Point-SLAM: Dense neural point cloud-based SLAM. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. ICCV, pp. 18433–18444.

Schmuck, P., Chli, M., 2019. CCM-SLAM: Robust and efficient centralized collaborative monocular simultaneous localization and mapping for robotic teams. J. Field Robot. 36 (4), 763–781. http://dx.doi.org/10.1002/rob.21854.

Sturm, J., Engelhard, N., Endres, F., Burgard, W., Cremers, D., 2012. A benchmark for the evaluation of RGB-D SLAM systems. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS, pp. 573–580. http://dx.doi.org/10.1109/IROS.2012.6385773.

Sucar, E., Liu, S., Ortiz, J., Davison, A.J., 2021. iMAP: Implicit mapping and positioning in Real-Time. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. ICCV, pp. 6229–6238.

Thomas, A., Sonawalla, A., Rose, A., How, J.P., 2025. GRAND-SLAM: Local optimization for globally consistent Large-Scale Multi-Agent Gaussian SLAM. IEEE Robot. Autom. Lett. arXiv:2506.18885.

Tian, Y., Chang, Y., Herrera Arias, F., Nieto-Granda, C., How, J.P., Carlone, L., 2022. Kimera-Multi: Robust, distributed, dense Metric-Semantic SLAM for Multi-Robot systems. IEEE Trans. Robot. 38 (4), 2022–2038. http://dx.doi.org/10.1109/TRO.2021.3137751.

Tian, Y., Khosoussi, K., Rosen, D.M., How, J.P., 2021. Distributed certifiably correct Pose-Graph optimization. IEEE Trans. Robot. 37 (6), 2137–2156. http://dx.doi.org/10.1109/TRO.2021.3072346.

Tran, A.T., Košecká, J., 2026. VarSplat: Uncertainty-Aware 3D Gaussian splatting for robust RGB-D SLAM. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, arXiv:2603.09673.

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 13 (4), 600–612. http://dx.doi.org/10.1109/TIP.2003.819861.

Xiang, B., Zhou, S., Gu, C., Xie, Z., Chen, L., Wang, Y., 2025. NISB-Fusion: Multi-agent mapping and map merging with neural implicit spatial block. IEEE Robot. Autom. Lett. RAL 2025.

Xu, X., et al., 2025. MAC-Ego3D: Multi-agent Gaussian consensus for real-time collaborative ego-motion and photorealistic 3D reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, arXiv:2412.09723.

Yan, C., Qu, D., Xu, D., Zhao, B., Wang, Z., Wang, D., Li, X., 2024. GS-SLAM: Dense visual SLAM with 3D Gaussian splatting. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, pp. 19595–19604.

Yu, J., Chen, T., Schwager, M., 2025. HAMMER: Heterogeneous, multi-robot semantic Gaussian splatting. IEEE Robot. Autom. Lett. http://dx.doi.org/10.1109/LRA.2025.3575235, arXiv:2501.14147.

Yugay, V., Gevers, T., Oswald, M.R., 2025. MAGiC-SLAM: Multi-agent Gaussian globally consistent SLAM. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, pp. 6741–6750, arXiv:2411.16785.

Yugay, V., Li, Y., Gevers, T., Oswald, M.R., 2023. Gaussian-SLAM: Photo-realistic dense SLAM with Gaussian splatting. arXiv preprint arXiv:2312.10070.

Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O., 2018. The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, pp. 586–595. http://dx.doi.org/10.1109/CVPR.2018.00068.

Zhu, L., Li, Y., Sandström, E., Huang, S., Schindler, K., Armeni, I., 2025. LoopSplat: Loop closure by registering 3D Gaussian splats. In: International Conference on 3D Vision. 3DV, pp. 156–167.

Zhu, Z., Peng, S., Larsson, V., Xu, W., Bao, H., Cui, Z., Oswald, M.R., Pollefeys, M., 2022. NICE-SLAM: Neural implicit scalable encoding for SLAM. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR.