Parallel Rigidity Matters for Bundle Adjustment Supplementary Material

Lalit Manam
Mitsubishi Electric Research Labs (MERL)
Cambridge, MA 02139, USA
manam@merl.com

Venu Madhav Govindu
Indian Institute of Science
Bengaluru, KA 560012, INDIA
venug@iisc.ac.in

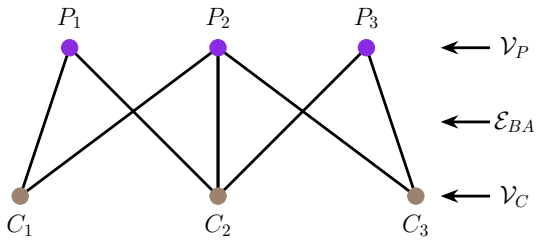
A. Graph Notations

In this section, we provide the graph notations used in Sec. 4 of the main paper along with a visual representation in Fig. S1 for clarity. The graphs involved in the paper:

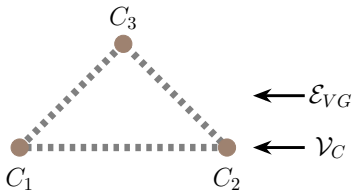
- $\mathcal{G}_{BA} = (\{\mathcal{V}_C \cup \mathcal{V}_P\}, \mathcal{E}_{BA})$: bipartite BA graph (Fig. S1a),
- $\mathcal{G}_{VG} = (\mathcal{V}_C, \mathcal{E}_{VG})$: viewgraph (Fig. S1b),

where

- $\mathcal{V}_C$: set of cameras (brown nodes in Figs. S1a and S1b),
- $\mathcal{V}_P$: set of 3D points (violet nodes in Fig. S1b),
- $\mathcal{E}_{BA}$: cam-pt edges in $\mathcal{G}_{BA}$ representing 2D observations such that $\mathcal{E}_{BA} \subseteq \mathcal{V}_C \times \mathcal{V}_P$ (black lines in Fig. S1a),
- $\mathcal{E}_{VG}$: cam-cam edges in $\mathcal{G}_{VG}$ representing matched 2D observations such that $\mathcal{E}_{VG} \subseteq \mathcal{V}_C \times \mathcal{V}_C$ (thick dashed gray lines in Fig. S1b).



(a) Bipartite BA graph ($\mathcal{G}_{BA}$).



(b) Viewgraph ($\mathcal{G}_{VG}$).

Figure S1. Notations of the two graphs used in the paper. Bipartite BA graph: $\mathcal{G}_{BA} = (\{\mathcal{V}_C \cup \mathcal{V}_P\}, \mathcal{E}_{BA})$. Viewgraph: $\mathcal{G}_{VG} = (\mathcal{V}_C, \mathcal{E}_{VG})$. See text for details.

B. Additional Details on the Proposed Method

In Sec. 4 of the main paper, we presented our proposed method. Here, we provide details on maintaining subgraphs with union-find data structure and on the computation complexity of GPRBA.

Subgraphs with union-find data structure: In Step 2 of our method, we noted that matched observations in each cam-cam edge in the viewgraph $\mathcal{G}_{VG}$ form subgraphs of the bipartite BA graph $\mathcal{G}_{BA}$ that are generically parallel rigid (GPR), through Type-2 interactions of 4-length loops in $\mathcal{G}_{BA}$. Then, we considered Type-1 interactions of 4-length loops in $\mathcal{G}_{BA}$ via common tracks between two cam-cam edges in $\mathcal{G}_{VG}$. To efficiently record which cam-cam edges in $\mathcal{G}_{VG}$ together form a GPR subgraph in $\mathcal{G}_{BA}$, we utilize the union-find data structure on cam-cam edges.

The union-find data structure stores a collection of disjoint sets to efficiently query if two elements belong to the same set. Every disjoint set is identified by a representative element. In our method, each disjoint set represents a set of cam-cam edges in $\mathcal{G}_{VG}$ that together form a GPR subgraph of $\mathcal{G}_{BA}$. The data structure can be represented as a set of tuples $\mathcal{UF}$. In each tuple $(e, r) \in \mathcal{UF}$, e is an element denoting a cam-cam edge in $\mathcal{E}_{VG}$ and r denotes the representative element of the disjoint set in which e belongs to.

Before starting Step 2 of our method, we initialize $\mathcal{UF}$ by considering every element in $\mathcal{UF}$ as a disjoint set, i.e. for every $(e, r) \in \mathcal{UF}$, $r \leftarrow e$. Let $(e_1, r_1), (e_2, r_2) \in \mathcal{UF}$. While executing Step 2 of our method, when two cam-cam edges in $\mathcal{G}_{VG}$, denoted by elements e_1 and e_2, are identified to belong to the same set, then $r_2 \leftarrow r_1$. This means the representative elements of e_1 and e_2 are assigned the same, and hence, they belong to the same disjoint set in $\mathcal{UF}$. After Step 2 of our method is completed, we extract the cam-pt edges in $\mathcal{G}_{BA}$ that correspond to the cam-cam edges

in $\mathcal{G}_{VG}$ having the same representative element in $\mathcal{UF}$. Each of these disjoint sets of cam-pt edges forms a GPR subgraph in $\mathcal{G}_{BA}$, which are used in Step 3 of our method.

Computational complexity of GPRBA: Let $\mathcal{N}$ be the subgraphs obtained at the end of Step 2. Our method has an average case computational complexity of $O(z|\mathcal{E}_{VG}| + z|\mathcal{V}_P| + |\mathcal{N}|^2)$, with $O(z|\mathcal{E}_{VG}| + z|\mathcal{V}_P|)$ for Step 1, $O(|\mathcal{E}_{VG}|)$ for Step 2 and $O(|\mathcal{N}|^2)$ for Step 3. The value of z depends on the number of cam-cam edges and tracks having less than two observations. In practice, for a dataset, $z \ll |\mathcal{E}_{VG}|$ and $|\mathcal{N}|^2 \ll |\mathcal{E}_{VG}| \ll |\mathcal{V}_P|$.

C. Additional Results

In Sec. 5 of the main paper, we provided camera translation errors with retriangulation on IMC 2022 [34] and visual reconstruction results with our HOF and GPRBA using GLOMAP [64]. Here, we present the translation errors both without and with retriangulation on IMC 2022 [34] and visual results on other datasets. We use the following notations in this supplementary material, same as that of the main paper:

- $\#N_C$: Number of cameras reconstructed.
- $\#N_P$: Number of 3D points reconstructed (in 10^3).

In Tab. S1, we provide reconstruction statistics both without and with retriangulation in GLOMAP [64]. We also provide camera translation errors after applying HOF and GPRBA individually along with HOF+GPRBA for completeness. However, we note that applying GPRBA individually without HOF does not guarantee a GPR graph, and is strongly discouraged. This is because hanging observations are not supported by any cam-cam edge in $\mathcal{G}_{VG}$, as discussed in practical considerations in Sec. 4. Thus, they are not accounted in GPRBA, which relies on cam-cam edges.

We observe that using GPRBA individually results in similar number of cams and more 3D points without retriangulation but similar number of 3D points with retriangulation on most datasets, compared to only HOF and HOF+GPRBA (Tab. 1 of the main paper). Overall, the median translation errors with retriangulation are less with HOF+GPRBA.

The retriangulation in GLOMAP is done in an incremental manner, which initializes with the output of without retriangulation phase. It also uses the matched observations in cam-cam edges that were not used in without retriangulation phase, due to which the impact of a GPR bipartite BA graph is difficult to assess exclusively from the results with retriangulation. So, we also provide the translation errors without retriangulation in Tab. S1. It can be seen that, without retriangulation, each of HOF and GPRBA improves over GLOMAP, while HOF+GPRBA performs overall best. This clearly shows that a GPR graph improves camera translations.

In Figs. S2 and S3, we show reconstructions with retriangulation in GLOMAP [64]. It can be seen that misplaced parts in the GLOMAP reconstructions are not removed by applying our HOF. Applying our HOF + GPRBA results in significantly clean reconstructions.

We also provide reconstruction results without retriangulation in Figs. S4 and S5. It can be seen that applying our HOF results in more reconstructed 3D points but has misplaced reconstruction parts. Similar to results with retriangulation, applying our HOF + GPRBA gives significantly clean reconstructions. We note that due to a small number of reconstructed 3D points by GLOMAP without retriangulation, misplaced parts are not clearly visible. As noted in the main paper, GLOMAP only retains observations with a low reprojection error, which removes many reconstructed 3D points. However, results with retriangulation clearly indicate that cameras belonging to misplaced parts are present in the reconstructions obtained without retriangulation, which can only be removed by applying our HOF + GPRBA.

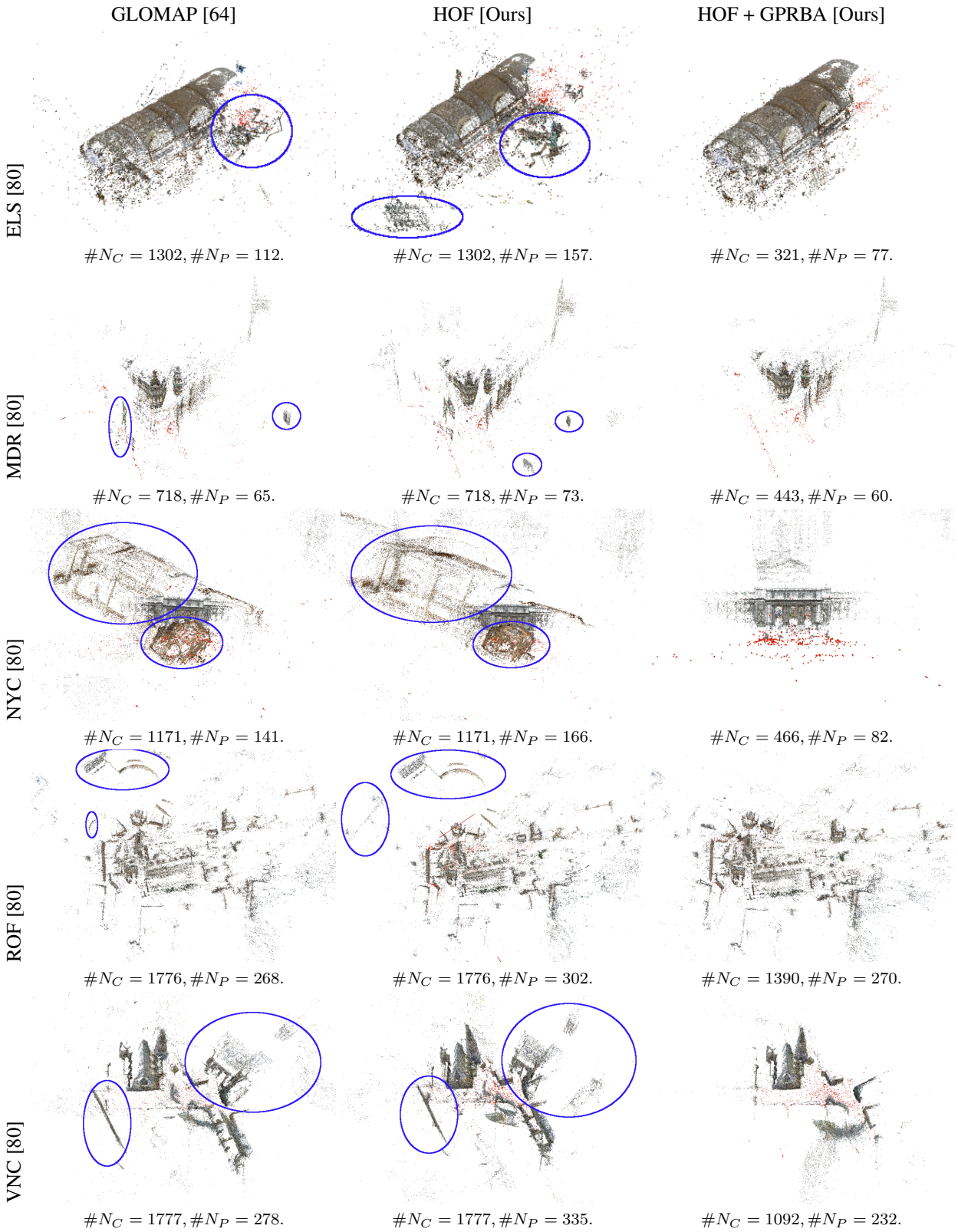


Figure S2. Reconstruction results with retriangulation in GLOMAP [64]. Applying our HOF does not remove misplaced parts (in blue), which are present in GLOMAP reconstructions. Our approach, HOF+GPRBA, designed to extract GPR subgraphs in BA, results in significantly clean reconstructions by removing misplaced cameras and 3D points.

Dataset	GPRBA (individual)			Median Translation Errors							
	# Cameras	# 3D Points (in 10^3)		w/o retriangulation (w/o retri.)				w/ retriangulation (w/ retri.)			
		w/o retri.	w/ retri.	GLO.	HOF	GPRBA	HOF+GPRBA	GLO.	HOF	GPRBA	HOF+GPRBA
BDG	347	18	30	1.04	0.61	**0.42**	0.52	**0.21**	0.26	0.28	0.26
BRM	176	16	26	0.72	0.35	0.38	**0.34**	0.16	0.15	0.15	**0.14**
BKP	445	36	81	6.18	0.28	2.39	**0.25**	0.69	0.21	0.33	**0.17**
COE	498	61	98	3.44	0.86	0.85	**0.81**	0.60	0.41	0.42	**0.40**
LMS	214	9	17	0.59	**0.05**	0.18	0.16	0.04	**0.03**	0.19	0.05
NDF	901	101	158	0.61	**0.14**	**0.14**	0.18	**0.04**	0.11	0.09	**0.04**
PNE	320	41	61	0.44	**0.13**	0.19	**0.13**	0.08	0.06	0.07	**0.05**
PSM	68	14	23	0.85	0.90	0.83	**0.30**	0.33	**0.20**	0.22	0.22
SCR	281	29	44	5.09	**0.12**	1.25	**0.12**	0.23	0.12	**0.11**	0.16
SAF	90	33	41	0.16	**0.06**	0.07	**0.06**	**0.07**	**0.07**	**0.07**	**0.07**
SPC	142	19	33	0.40	**0.24**	0.30	0.28	0.30	**0.16**	0.29	0.27
TNJ	217	20	28	2.26	0.83	**0.82**	0.84	0.74	0.29	0.29	**0.28**
TRF	700	117	165	0.06	0.06	**0.05**	**0.05**	0.05	**0.04**	0.09	0.08

Table S1. Reconstruction statistics on IMC 2022 [34] before and after applying our HOF and GPRBA methods. **Bold** and underlined camera translation errors are the best and second best camera translation errors in their respective category. A GPR bipartite BA graph with HOF+GPRBA results in overall best camera translations.

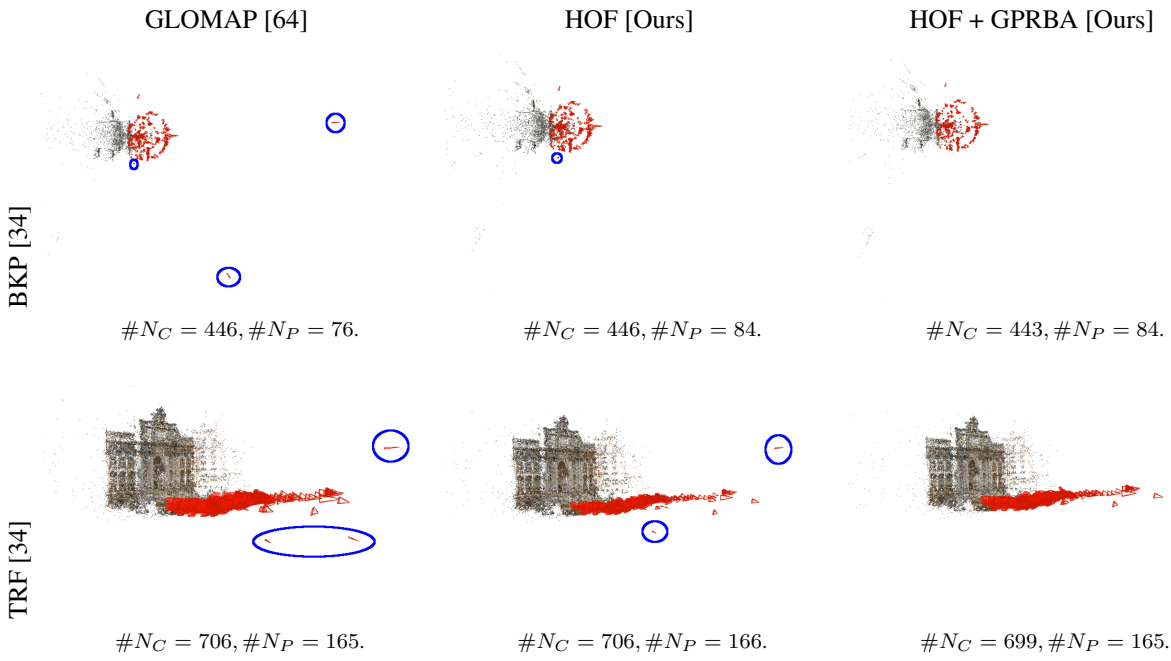


Figure S3. Reconstruction results with retriangulation in GLOMAP [64]. Similar observations can be made as in Fig. S2.

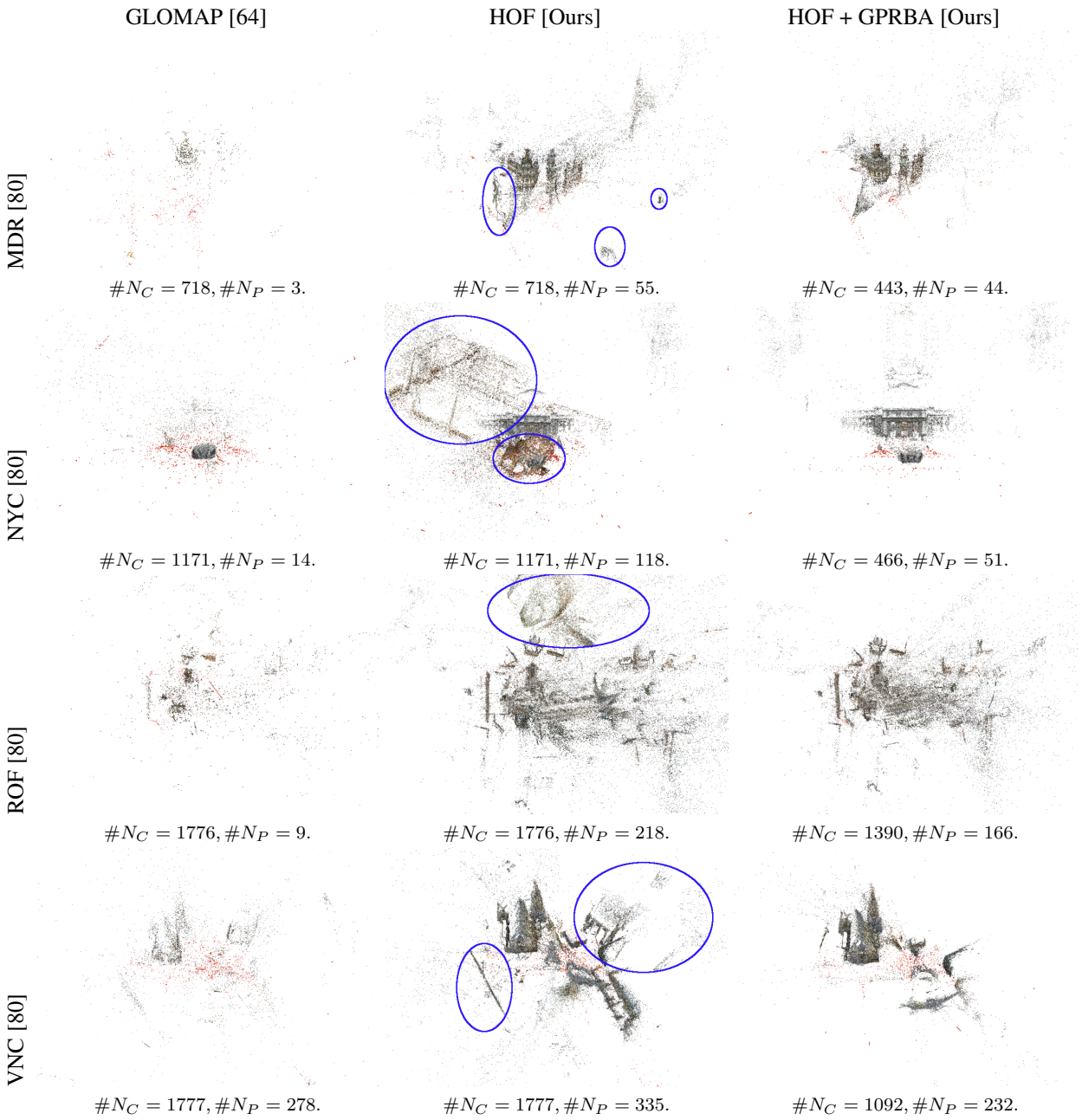


Figure S4. Reconstruction results without retriangulation in GLOMAP [64]. Applying our HOF increases reconstructed 3D points significantly but does not remove misplaced parts (blue ellipses) present in the reconstructions. Applying our HOF+GPRBA leads to significantly cleaner reconstructions by removing misplaced cameras and 3D points.

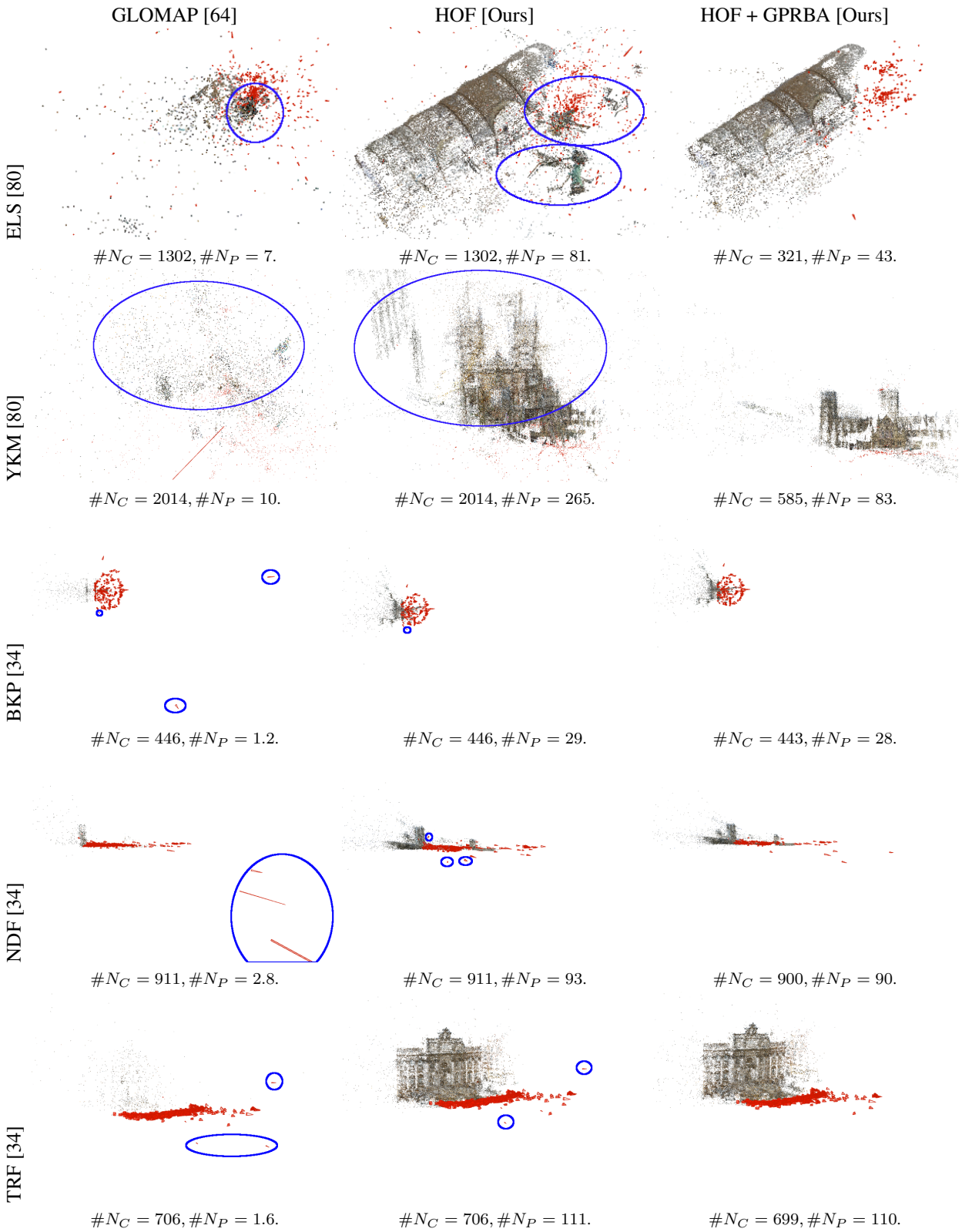


Figure S5. Additional reconstruction results without retriangulation in GLOMAP [64]. Similar observations can be made as in Fig. S4.