Sensitivity Analysis in Translation Averaging Supplementary Material

Lalit Manam and Venu Madhav Govindu

A Proofs for Theorems

We first state a result from matrix theory, which will be helpful in proving Thm. 4 of the main paper.

Result 1. *Let $\mathbf{X} \in \mathbb{R}^{n\times n}$ be a square block diagonal matrix given as*

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \cdots & \mathbf{0} \\ & \vdots & \vdots & \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_M \end{bmatrix}, \tag{S1}$$

where $\mathbf{X}_k \in \mathbb{R}^{n_k\times n_k}$, $k \in \{1, 2, \cdots M\}$ are the (square) diagonalizable matrices at the block diagonal of $\mathbf{X}$. Then, using the definition of the determinant (Sec. 6.1 in [42]), we get

$$\det(\mathbf{X}) = \prod_{k=1}^{M} \det(\mathbf{X}_k). \tag{S2}$$

Then, the following equation holds:

$$\det(\mathbf{X} - \lambda\mathbf{I}_n) = \prod_{k=1}^{M} \det(\mathbf{X}_k - \lambda\mathbf{I}_{n_k}), \tag{S3}$$

where $\lambda \in \mathbb{C}$, and $\mathbf{I}_p$ is the $p \times p$ identity matrix. Equating Eqn. S3 to zero gives us the characteristic equation of $\mathbf{X}$, which is solved to obtain the eigen values of $\mathbf{X}$. Hence, the eigen values of a square block diagonal matrix are a collection of the eigen values of each of the (square) diagonalizable matrices at the block diagonals.

Now, we prove the theorems stated in Secs. 3 and 4 of the main paper.

Theorem 1. *Given two directions $\mathbf{v}_1$ and $\mathbf{v}_2$, and the matrix $\mathbf{V} = [\ \mathbf{v}_1 \mid \text{-}\mathbf{v}_2\ ]$, the condition number of the matrix $\mathbf{V}^T\mathbf{V}$ with matrix-2 norm is given as*

$$\kappa(\mathbf{V}^T\mathbf{V}) = \begin{cases} \cot^2(\phi_{12}/2) & \text{if } 0^\circ < \phi_{12} \leq 90^\circ; \\ \tan^2(\phi_{12}/2) & \text{if } 90^\circ < \phi_{12} < 180^\circ; \\ \infty & \text{if } \phi_{12} \in \{0^\circ, 180^\circ\}, \end{cases} \tag{S4}$$

where ϕ_{12} is the angle between the two directions $\mathbf{v}_1$ and $\mathbf{v}_2$.

Proof.

$$\kappa(\mathbf{X}) = \frac{\sigma_{max}(\mathbf{X})}{\sigma_{min}(\mathbf{X})} = \sqrt{\frac{\lambda_{max}(\mathbf{X}^T\mathbf{X})}{\lambda_{min}(\mathbf{X}^T\mathbf{X})}}, \tag{S5}$$

- *Lalit Manam was with, and Venu Madhav Govindu is with the Department of Electrical Engineering, Indian Institute of Science, Bengaluru, Karnataka - 560012, INDIA. E-mail: lalitmanam@gmail.com, venug@iisc.ac.in*

where $\sigma(\mathbf{X})$ and $\lambda(\mathbf{X})$ denote a singular value and an eigen value of the matrix $\mathbf{X}$, respectively, and their subscript denotes maximum or minimum value among their possible values for $\mathbf{X}$. From Eqn. S5, we get

$$\begin{aligned}\kappa(\mathbf{V}^T\mathbf{V}) =& \sqrt{\frac{\lambda_{max}\left(\left(\mathbf{V}^T\mathbf{V}\right)^T\left(\mathbf{V}^T\mathbf{V}\right)\right)}{\lambda_{min}\left(\left(\mathbf{V}^T\mathbf{V}\right)^T\left(\mathbf{V}^T\mathbf{V}\right)\right)}} = \sqrt{\frac{\lambda_{max}\left(\left(\mathbf{V}^T\mathbf{V}\right)^2\right)}{\lambda_{min}\left(\left(\mathbf{V}^T\mathbf{V}\right)^2\right)}} \\ =& \sqrt{\frac{\lambda^2_{max}(\mathbf{V}^T\mathbf{V})}{\lambda^2_{min}(\mathbf{V}^T\mathbf{V})}} = \frac{\lambda_{max}(\mathbf{V}^T\mathbf{V})}{\lambda_{min}(\mathbf{V}^T\mathbf{V})},\end{aligned} \tag{S6}$$

where we use that fact that $\mathbf{V}^T\mathbf{V}$ is positive semi-definite, and hence, $\lambda\left(\left(\mathbf{V}^T\mathbf{V}\right)^2\right) = \lambda^2\left(\mathbf{V}^T\mathbf{V}\right)$ and $\sqrt{\lambda^2(\mathbf{V}^T\mathbf{V})} = \lambda(\mathbf{V}^T\mathbf{V})$. The matrix $\mathbf{V}^T\mathbf{V}$ is given as

$$\mathbf{V}^T\mathbf{V} = \begin{bmatrix}\mathbf{v}_1^T\mathbf{v}_1 & -\mathbf{v}_1^T\mathbf{v}_2 \\ -\mathbf{v}_1^T\mathbf{v}_2 & \mathbf{v}_2^T\mathbf{v}_2\end{bmatrix} = \begin{bmatrix}1 & -\mathbf{v}_1^T\mathbf{v}_2 \\ -\mathbf{v}_1^T\mathbf{v}_2 & 1\end{bmatrix}, \tag{S7}$$

and its eigen values are given as $\lambda(\mathbf{V}^T\mathbf{V}) = 1 \pm \mathbf{v}_1^T\mathbf{v}_2 = 1 \pm \cos(\phi_{12})$, where ϕ_{12} is the angle between $\mathbf{v}_1$ and $\mathbf{v}_2$. Appropriately placing maximum and minimum eigen values of $\mathbf{V}^T\mathbf{V}$ in Eqn. S6, we get

$$\begin{aligned}\kappa(\mathbf{V}^T\mathbf{V}) &= \begin{cases}\frac{1+cos(\phi_{12})}{1-cos(\phi_{12})} & \text{if } 0^\circ < \phi_{12} \leq 90^\circ; \\ \frac{1-cos(\phi_{12})}{1+cos(\phi_{12})} & \text{if } 90^\circ < \phi_{12} < 180^\circ; \\ \infty & \text{if } \phi_{12} \in \{0^\circ, 180^\circ\},\end{cases} \\ \implies \kappa(\mathbf{V}^T\mathbf{V}) &= \begin{cases}\cot^2(\phi_{12}/2) & \text{if } 0^\circ < \phi_{12} \leq 90^\circ; \\ \tan^2(\phi_{12}/2) & \text{if } 90^\circ < \phi_{12} < 180^\circ; \\ \infty & \text{if } \phi_{12} \in \{0^\circ, 180^\circ\}.\end{cases}\end{aligned} \tag{S8}$$

■

Theorem 2. *Given two directions $\mathbf{v}_1$ and $\mathbf{v}_2$, the matrix $\mathbf{V} = [\ \mathbf{v}_1 \mid \text{-}\mathbf{v}_2\ ]$, and the restricted dot product matrix for two directions $\mathbf{RDP}_{12}$, the condition number of both $\mathbf{RDP}_{12}$ and $\mathbf{V}^T\mathbf{V}$ with matrix-2 norm are the same.*

Proof. We rewrite the definition of $\mathbf{RDP}_{12}$ for reference.

$$\mathbf{RDP}_{12} = \begin{bmatrix}1 & |\mathbf{v}_1^T\mathbf{v}_2| \\ |\mathbf{v}_1^T\mathbf{v}_2| & 1\end{bmatrix}. \tag{S9}$$

At first, we check the eigen values of $\mathbf{RDP}_{12}$, which are given as $\lambda(\mathbf{RDP}_{12}) = 1 \pm |\mathbf{v}_1^T\mathbf{v}_2| = 1 \pm \mathbf{v}_1^T\mathbf{v}_2$. Since $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = 1$, $|\mathbf{v}_1^T\mathbf{v}_2| \leq \|\mathbf{v}_1\|\|\mathbf{v}_2\| = 1$ (from Cauchy-Schwarz inequality). Hence $\lambda(\mathbf{RDP}_{12}) \geq 0$ implying $\mathbf{RDP}_{12}$ is a positive semi-definite matrix. The condition number of $\mathbf{RDP}_{12}$ with matrix-2 norm is given as

$$\begin{aligned}\kappa(\mathbf{RDP}_{12}) =& \sqrt{\frac{\lambda_{max}\left(\mathbf{RDP}_{12}^T\mathbf{RDP}_{12}\right)}{\lambda_{min}\left(\mathbf{RDP}_{12}^T\mathbf{RDP}_{12}\right)}} = \sqrt{\frac{\lambda_{max}\left(\left(\mathbf{RDP}_{12}\right)^2\right)}{\lambda_{min}\left(\left(\mathbf{RDP}_{12}\right)^2\right)}} \\ =& \sqrt{\frac{\lambda^2_{max}(\mathbf{RDP}_{12})}{\lambda^2_{min}(\mathbf{RDP}_{12})}} = \frac{\lambda_{max}(\mathbf{RDP}_{12})}{\lambda_{min}(\mathbf{RDP}_{12})},\end{aligned}$$

where $\lambda\left(\left(\mathbf{RDP}_{12}\right)^2\right) = \lambda^2\left(\mathbf{RDP}_{12}\right)$ and $\sqrt{\lambda^2(\mathbf{RDP}_{12})} = \lambda(\mathbf{RDP}_{12})$ since $\mathbf{RDP}_{12}$ is a positive semi-definite matrix.

From Eqn. S7 in the proof of Thm. 1, the eigen values of $\mathbf{V}^T\mathbf{V}$ are given as $\lambda(\mathbf{V}^T\mathbf{V}) = 1 \pm \mathbf{v}_1^T\mathbf{v}_2$. Since the eigen values of $\mathbf{RDP}_{12}$ and $\mathbf{V}^T\mathbf{V}$ are the same, the condition number of both $\mathbf{RDP}_{12}$ and $\mathbf{V}^T\mathbf{V}$, with matrix-2, norm are the same and is given by

$$\kappa(\mathbf{RDP}_{12}) = \begin{cases}\cot^2(\phi_{12}/2) & \text{if } 0^\circ < \phi_{12} \leq 90^\circ; \\ \tan^2(\phi_{12}/2) & \text{if } 90^\circ \leq \phi_{12} < 180^\circ; \\ \infty & \text{if } \phi_{12} \in \{0^\circ, 180^\circ\},\end{cases} \tag{S10}$$

with ϕ_{12} being the angle between $\mathbf{v}_1$ and $\mathbf{v}_2$. ■

Theorem 3. *Consider p directions, $\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_p$, and the restricted dot product matrix using the p directions, $\mathbf{RDP}$. The matrix $\mathbf{RDP}$ is well-conditioned if the angle between all pairs of directions is not close to 0° or 180°.*

Proof. The entries of the restricted dot product matrix $\mathbf{RDP}$ for p directions are given as

$$rdp^{ij} = |\mathbf{v}_i^T \mathbf{v}_j|, \tag{S11}$$

where rdp^{ij} denotes the entry in i^{th} row and j^{th} column of $\mathbf{RDP}$.

Let us assume that angles between all pairs of directions are not close to 0° or 180°. Since this condition is satisfied, $|\mathbf{v}_i^T \mathbf{v}_j| \neq 1$. Thus, no two columns in $\mathbf{RDP}$ can have exactly same entries because the diagonal elements are 1 ($\mathbf{v}_i^T \mathbf{v}_i = 1$ since $\|\mathbf{v}_i\| = 1$). This makes $\mathbf{RDP}$ non-singular when angles between all pairs of directions are not close to 0° or 180°.

Next, we look at the conditioning of $\mathbf{RDP}$ matrix. We check cosine of the angle between two columns in $\mathbf{RDP}$ corresponding to i^{th} and j^{th} directions, which is given as

$$\begin{aligned} \cos\angle(\mathbf{RDP}^{*i}, \mathbf{RDP}^{*j}) =& \frac{\sum_{k=1}^{p} |\mathbf{v}_k^T \mathbf{v}_i| \cdot |\mathbf{v}_k^T \mathbf{v}_j|}{\sqrt{\sum_{k=1}^{p} |\mathbf{v}_k^T \mathbf{v}_i|^2} \cdot \sqrt{\sum_{k=1}^{p} |\mathbf{v}_k^T \mathbf{v}_j|^2}} \\ \implies \cos\angle(\mathbf{RDP}^{*i}, \mathbf{RDP}^{*j}) =& \frac{2|\mathbf{v}_i^T \mathbf{v}_j| + \sum_{k \neq \{i,j\}} |\mathbf{v}_k^T \mathbf{v}_i| \cdot |\mathbf{v}_k^T \mathbf{v}_j|}{\sqrt{1 + \sum_{k \neq i} |\mathbf{v}_k^T \mathbf{v}_i|^2} \cdot \sqrt{1 + \sum_{k \neq j} |\mathbf{v}_k^T \mathbf{v}_j|^2}}. \end{aligned} \tag{S12}$$

When angles between all pairs of directions are not close to 0° or 180°, then $|\mathbf{v}_k^T \mathbf{v}_l| < 1$. Thus, $\cos\angle(\mathbf{RDP}^{*i}, \mathbf{RDP}^{*j})$ cannot be close to 1, and any two columns in the $\mathbf{RDP}$ matrix are not similar.

Hence, if angles between all pairs of directions are not close to 0° or 180°, then $\mathbf{RDP}$ is non-singular, and its columns are not similar, implying that $\mathbf{RDP}$ is well-conditioned. ■

Theorem 4. *Consider a bearing-based network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the restricted dot product matrix $\mathbf{RDP}[i]$ for every node $i \in \mathcal{V}$, and the block diagonal matrix $\mathbf{RDP}_\mathcal{G}$. Then, the condition number, with matrix-2 norm, of the matrix $\mathbf{RDP}_\mathcal{G}$ is lower bounded by the condition number of $\mathbf{RDP}[i]$, for any $i \in \mathcal{V}$, i.e.*

$$\kappa\left(\mathbf{RDP}_\mathcal{G}\right) \geq \kappa\left(\mathbf{RDP}[i]\right), \ \forall i \in \mathcal{V}. \tag{S13}$$

Proof. We restate the definition of $\mathbf{RDP}_\mathcal{G}$ for reference. $\mathbf{RDP}_\mathcal{G}$ is a block diagonal matrix, defined as

$$[\mathbf{RDP}_\mathcal{G}]_i = \mathbf{RDP}[i]. \tag{S14}$$

Using Eqn. S5, the condition numbers of the matrices $\mathbf{RDP}_\mathcal{G}$ and $\mathbf{RDP}[i]$, with matrix-2 norm, are given as

$$\kappa\left(\mathbf{RDP}_\mathcal{G}\right) = \sqrt{\frac{\lambda_{max}\left(\mathbf{RDP}_\mathcal{G}^T \mathbf{RDP}_\mathcal{G}\right)}{\lambda_{min}\left(\mathbf{RDP}_\mathcal{G}^T \mathbf{RDP}_\mathcal{G}\right)}}, \tag{S15}$$

$$\kappa\left(\mathbf{RDP}[i]\right) = \sqrt{\frac{\lambda_{max}\left(\mathbf{RDP}[i]^T \mathbf{RDP}[i]\right)}{\lambda_{min}\left(\mathbf{RDP}[i]^T \mathbf{RDP}[i]\right)}}. \tag{S16}$$

Since $\mathbf{RDP}_\mathcal{G}$ is a block diagonal matrix, $\mathbf{RDP}_\mathcal{G}^T \mathbf{RDP}_\mathcal{G}$ is also a block diagonal matrix which is symmetric. Moreover, as $\mathbf{RDP}_\mathcal{G}^T \mathbf{RDP}_\mathcal{G}$ is block diagonal, the eigen values of $\mathbf{RDP}_\mathcal{G}^T \mathbf{RDP}_\mathcal{G}$ are the collection of eigen values from $\mathbf{RDP}[i]^T \mathbf{RDP}[i]$ $\forall i \in \mathcal{V}$ (using Result 1). Hence, the following relations hold:

$$\lambda_{max}\left(\mathbf{RDP}_\mathcal{G}^T \mathbf{RDP}_\mathcal{G}\right) = \max_{i \in \mathcal{V}} \left\{\lambda_{max}\left(\mathbf{RDP}[i]^T \mathbf{RDP}[i]\right)\right\}, \tag{S17}$$

$$\lambda_{min}\left(\mathbf{RDP}_\mathcal{G}^T \mathbf{RDP}_\mathcal{G}\right) = \min_{i \in \mathcal{V}} \left\{\lambda_{min}\left(\mathbf{RDP}[i]^T \mathbf{RDP}[i]\right)\right\}. \tag{S18}$$

Thus, using Eqns. S17 and S18 for any $i \in \mathcal{V}$, and the fact that $\mathbf{RDP}[i]^T\mathbf{RDP}[i], \forall i \in \mathcal{V}$ and $\mathbf{RDP}_{\mathcal{G}}^T\mathbf{RDP}_{\mathcal{G}}$ are positive semi-definite matrices, the following inequalities hold $\forall i \in \mathcal{V}$:

$$\lambda_{max}\left(\mathbf{RDP}_{\mathcal{G}}^T\mathbf{RDP}_{\mathcal{G}}\right) \geq \lambda_{max}\left(\mathbf{RDP}[i]^T\mathbf{RDP}[i]\right) \geq 0, \tag{S19}$$

$$\lambda_{min}\left(\mathbf{RDP}[i]^T\mathbf{RDP}[i]\right) \geq \lambda_{min}\left(\mathbf{RDP}_{\mathcal{G}}^T\mathbf{RDP}_{\mathcal{G}}\right) \geq 0. \tag{S20}$$

Using Eqns. S19 and S20 in Eqn. S15, we get

$$\kappa\left(\mathbf{RDP}_{\mathcal{G}}\right) \geq \kappa\left(\mathbf{RDP}[i]\right), \ \forall i \in \mathcal{V}. \tag{S21}$$

Thus, the matrix-2 norm condition number of $\mathbf{RDP}_{\mathcal{G}}$ is lower bounded by the matrix-2 norm condition number of the restricted dot product matrix $\mathbf{RDP}[i]$ for any node $i \in \mathcal{V}$ in a bearing-based network $\mathcal{G}$. ■

B Formulations for Solving Translation Averaging

In Sec. 6 of the main paper, we used two translation averaging formulations, Revised LUD [47],[68] and BATA [68], for the experiments. Here, we provide the corresponding optimization problems for reference.

Revised LUD [47],[68] (compares relative displacements):

$$\begin{aligned} &\min_{\{\mathbf{T}_i\}_{i\in\mathcal{V}},\{\lambda_{ij}\}_{(i,j)\in\mathcal{E}}} \sum_{(i,j)\in\mathcal{E}} \|\mathbf{T}_j - \mathbf{T}_i - \lambda_{ij}\mathbf{v}_{ij}\|_2 \\ &\text{s.t.} \sum_{i\in\mathcal{V}} \mathbf{T}_i = \mathbf{0}, \sum_{(i,j)\in\mathcal{E}} \langle \mathbf{T}_j - \mathbf{T}_i, \mathbf{v}_{ij}\rangle = 1, \lambda_{ij} \geq 0 \ \forall (i,j) \in \mathcal{E}. \end{aligned} \tag{S22}$$

BATA [68] (compares relative directions):

$$\begin{aligned} &\min_{\{\mathbf{T}_i\}_{i\in\mathcal{V}},\{\gamma_{ij}\}_{(i,j)\in\mathcal{E}}} \sum_{(i,j)\in\mathcal{E}} \rho\left(\|\left(\mathbf{T}_j - \mathbf{T}_i\right)\gamma_{ij} - \mathbf{v}_{ij}\|_2\right) \\ &\text{s.t.} \sum_{i\in\mathcal{V}} \mathbf{T}_i = \mathbf{0}, \sum_{(i,j)\in\mathcal{E}} \langle \mathbf{T}_j - \mathbf{T}_i, \mathbf{v}_{ij}\rangle = 1, \gamma_{ij} \geq 0 \ \forall (i,j) \in \mathcal{E}. \end{aligned} \tag{S23}$$

The zero centroid and dot product constraints in Eqns. S22 and S23 fix the origin and the global scale ambiguities, respectively. ρ denotes the Cauchy loss function with its scale value $\beta = 0.1$ used in the experiments, as suggested in [68]. λ_{ij} and γ_{ij} are non-negative variables that are ideally equal to edge scale and its inverse for the edge (i,j), respectively.

C Additional Results

In this section, we provide additional experimental results before and after applying our filter.

Outlier-Free Data: In Sec. 6.2 of the main paper, we presented results on outlier-free data with BATA [68] for unfiltered and filtered bearing-based networks using Algo. 1. Here, we provide results using RLUD [47],[68] in Table S1. It can be seen that the overall performance of RLUD improves after improving the conditioning of bearing-based networks. Moreover, as observed with BATA translation estimates, even with RLUD, absolute translation errors of the nodes removed due to applying our filter are high, suggesting that the removed nodes were not estimated properly.

Aggressive vs Non-Aggressive Pruning: In Sec. 6 of the main paper, we presented results with non-aggressive pruning of edges in our method. Here, we compare the networks on real data with non-aggressive and aggressive pruning in Table S2. We observe that aggressive pruning of edges results in the loss of a significant fraction of nodes and edges in the network, which is much higher than non-aggressive pruning. This is expected because many directions (edges) have minimum restricted angles close to 0° or 180°, as seen in Fig. 5 of the main paper. These directions will be marked in Stage-1 of our method, which are then pruned either aggressively or non-aggressively in Stage-2.

We note that N_{rem} and M_{rem} in Table S2 denote the number of nodes and edges removed only due to ill-conditioned nodes (same as in the main paper). Some nodes are removed as all edges incident on them are removed during the pruning process. Aggressive pruning results in a significant loss of input data, which is not desirable as it leads to a far smaller number of cameras reconstructed, heavily diminishing the benefits of improving the conditioning of translation averaging. Hence, we recommend non-aggressive pruning of edges in practice.

Impact of varying τ: In Sec. 6 of the main paper, we used $\tau = 0.5^\circ$ for all the experiments. Here, we provide the impact of varying τ on real world data. In Table S3, we provide the network details on 1DSfM [63] datasets for $\tau = \{0.25^\circ, 0.5^\circ, 1^\circ\}$ with non-aggressive pruning. It can be seen that with an increase in τ, the number of nodes and edges reduces. Between

TABLE S1: Absolute translation errors (in meters) on 1DSfM [63] datasets without outliers. The results are obtained using RLUD [47],[68]. t_{RLUD}: Time taken by RLUD. '-': No nodes were removed after applying our filter.

Dataset	Mean Errors		RMS Errors		Removed Node Errors		t_{RLUD}
	w/o filter	w/ filter	w/o filter	w/ filter	Mean	RMS	(sec)
ALM	**2.9**	**2.9**	**4.8**	**4.8**	6.1	7.8	5
ELS	1.5	**1.4**	**1.8**	**1.8**	-	-	1
GMM	38.7	**37.7**	58.8	**56.5**	24.6	34.9	5
MDR	**9.0**	10.8	**18.9**	21.3	13.2	15.4	2
MND	2.5	**2.3**	3.5	**3.3**	2.2	2.8	3
NYC	**2.4**	2.5	**5.0**	5.2	1.3	1.3	2
ND	**3.5**	4.0	**5.6**	6.4	3.3	4.9	7
PDP	7.7	**4.5**	8.7	**5.7**	6.0	6.3	2
PIC	**3.0**	**3.0**	**5.9**	6.0	3.7	5.2	25
ROF	14.1	**12.9**	26.6	**24.9**	-	-	8
TOL	13.9	**13.3**	27.5	**22.0**	10.8	13.1	3
TFG	**8.1**	**8.1**	13.0	**12.9**	15.2	20.9	63
USQ	**5.7**	5.8	**8.0**	8.1	-	-	3
VNC	6.3	**5.3**	9.8	**7.8**	10.8	12.3	4
YKM	9.4	**7.9**	20.9	**15.0**	71.8	85.6	3

TABLE S2: Network details on 1DSfM [63] datasets with non-aggressive (Non-Agg fil.) and aggressive (Agg fil.) pruning of edges in Step 7 of Algo. 1. N_{rem}, M_{rem}: Number of nodes and edges removed only due to ill-conditioned configuration of directions. 'w/o filter' indicates without applying our filter i.e. the input network.

Dataset	w/o filter		$\#N_{rem}$		$\#M_{rem}$	
	#Nodes	#Edges	Non-Agg fil.	Agg fil.	Non-Agg fil.	Agg fil.
ALM	822	16192	5	440	9571	14962
ELS	325	7430	0	144	2611	6085
GMM	962	13930	4	452	7937	12243
MDR	435	4521	8	204	2288	3793
MND	581	18356	16	362	14809	17407
NYC	704	7782	8	267	3183	6234
ND	1421	70771	213	1261	65449	70478
PDP	980	15007	34	514	10469	13876
PIC	2826	45944	15	1178	24603	40199
ROF	1404	19049	1	501	6724	15130
TOL	633	8960	40	390	6715	8485
TFG	6957	113832	42	2732	66895	101107
USQ	1002	13133	2	301	5396	10155
VNC	1284	26023	24	571	18776	23101
YKM	990	12115	10	375	5221	9513

$\tau = 0.25°$ and $\tau = 0.5°$, the number of nodes reduces slightly, but the number of edges reduces significantly. From $\tau = 0.5°$ to $\tau = 1°$, both the number of nodes and edges drastically reduce, which is not desirable. This phenomenon is observed because of the following reasons.

1) *At a given node:* There are many directions (edges) whose minimum of restricted angles are close to $0°$, as shown in Fig. 5 of the main paper. Increasing τ will mark more edges at every node in Stage-1 of Algo. 1 to be considered for removal.
2) *Across all nodes:* Increasing τ leads to more edges being marked for removal from both the nodes they are incident to. This is evident from Fig. 4 of the main paper that shows angles between all pairs of directions at all nodes. Thus, more edges are removed even with non-aggressive pruning in our method.

Hence, we choose $\tau = 0.5°$ for our experiments to maintain a trade-off between improving the conditioning and retaining sufficient data.

We also provide a comparison in terms of absolute translation errors with varying τ in Table S4. It can be seen that filtering with any value of τ leads to overall improvement in translation estimates compared to the unfiltered network. We also observe that, for some datasets, the improvement is not consistent with increasing τ. This is because translation estimates not only depend on the conditioning of the problem but also on the noise and outliers in the directions. In Table S5, we provide the errors of nodes in the unfiltered network which were removed in the filtered network (RNE errors, same as Table 5 of the main paper). We observe that with an increase in τ, the mean and RMS errors reduce. This is expected because with increasing τ, we increasingly remove directions that less affect the conditioning of the problem.

TABLE S3: Network details on 1DSfM [63] datasets with varying τ.

	#Nodes				#Edges			
Dataset	w/o filter	$\tau = 0.25$	$\tau = 0.5$	$\tau = 1$	w/o filter	$\tau = 0.25$	$\tau = 0.5$	$\tau = 1$
ALM	822	814	801	538	16192	10389	6602	2067
ELS	325	325	325	293	7430	6703	4819	1516
GMM	962	962	928	785	13930	11201	5942	3216
MDR	435	403	381	237	4521	3378	2131	976
MND	581	575	537	349	18356	7424	3513	1493
NYC	704	677	664	612	7782	6257	4493	2630
PDP	980	947	850	375	15007	8202	4350	1188
PIC	2826	2812	2769	2405	45944	34708	21273	10588
ROF	1404	1400	1391	1275	19049	16662	12306	5897
TOL	633	617	451	82	8960	4467	1993	196
TFG	6957	6851	6559	5568	113832	74707	46142	23732
USQ	1002	989	987	896	13133	11413	7685	4007
VNC	1284	1250	1084	623	26023	10830	6566	2645
YKM	990	980	937	821	12115	10449	6824	3983

TABLE S4: Absolute translation errors on 1DSfM [63] datasets with varying τ using BATA [68]. **Bold** indicates that the translation error is better after filtering (for a given τ) compared to before filtering (w/o filter).

	Mean				RMS			
Dataset	w/o filter	$\tau = 0.25$	$\tau = 0.5$	$\tau = 1$	w/o filter	$\tau = 0.25$	$\tau = 0.5$	$\tau = 1$
ALM	5.2	**5.1**	5.3	**4.0**	9.6	**9.6**	9.8	**7.7**
ELS	22.8	**22.3**	**22.4**	**16.0**	51.9	**48.4**	**48.5**	**31.9**
GMM	40.7	**40.5**	**40.6**	**38.5**	61.0	**60.8**	**60.4**	**57.5**
MDR	13.7	**12.6**	**12.5**	**10.2**	30.4	**28.7**	**27.5**	**23.4**
MND	4.4	**3.8**	**3.4**	4.5	10.2	**7.4**	**6.2**	**6.5**
NYC	6.5	**5.0**	**4.7**	**5.3**	13.6	**8.4**	**8.5**	**11.4**
PDP	7.7	**7.7**	**7.1**	17.9	12.9	13.0	**10.6**	31.1
PIC	5.5	**5.5**	**5.2**	**5.0**	11.9	**11.7**	**10.8**	**9.7**
ROF	12.5	**11.5**	**11.2**	**10.8**	25.3	**23.8**	**23.7**	**23.2**
TOL	15.7	**14.9**	**17.3**	**11.8**	29.0	**27.3**	31.4	**17.8**
TFG	19.7	**18.1**	**14.4**	**14.0**	69.5	**55.4**	**27.7**	**24.3**
USQ	15.9	**15.7**	**13.9**	**14.5**	25.1	**24.9**	**20.7**	**22.3**
VNC	13.9	**13.2**	**8.9**	**5.8**	27.7	**27.2**	**13.2**	**8.4**
YKM	27.9	**17.9**	**14.2**	**22.8**	39.6	**25.6**	**22.5**	**33.9**

TABLE S5: Removed node errors on 1DSfM [63] datasets with varying τ using BATA [68]. RNE denotes errors of nodes in the unfiltered network which were removed in the filtered network. '-': No nodes were removed after applying our filter.

	RNE Mean			RNE RMS		
Dataset	$\tau = 0.25$	$\tau = 0.5$	$\tau = 1$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 1$
ALM	17.3	8.4	5.6	17.4	11.2	8.8
ELS	-	-	77.8	-	-	110.4
GMM	-	87.9	51.0	-	104.3	69.5
MDR	4.2	6.7	10.2	4.2	11.5	17.5
MND	78.0	14.8	6.5	78.0	34.5	16.2
NYC	-	73.7	42.5	-	73.7	54.0
PDP	16.9	7.1	7.0	16.9	11.3	16.1
PIC	10.8	12.8	8.5	14.4	20.5	17.1
ROF	-	36.0	32.8	-	36.0	49.7
TOL	30.2	14.0	12.6	40.3	28.8	26.7
TFG	99.8	60.7	24.9	138.2	96.3	51.2
USQ	-	43.2	21.9	-	43.3	30.9
VNC	18.0	32.7	18.0	28.3	55.0	34.4
YKM	72.2	45.1	35.2	72.2	54.2	47.1