

# A Subspace-Constrained Tyler’s Estimator and its Applications to Structure from Motion \*

Feng Yu <sup>†</sup>  
University of Minnesota  
fyu@umn.edu

Teng Zhang  
University of Central Florida  
teng.zhang@ucf.edu

Gilad Lerman <sup>‡</sup>  
University of Minnesota  
lerman@umn.edu

## Abstract

We present the subspace-constrained Tyler’s estimator (STE) designed for recovering a low-dimensional subspace within a dataset that may be highly corrupted with outliers. STE is a fusion of the Tyler’s M-estimator (TME) and a variant of the fast median subspace. Our theoretical analysis suggests that, under a common inlier-outlier model, STE can effectively recover the underlying subspace, even when it contains a smaller fraction of inliers relative to other methods in the field of robust subspace recovery. We apply STE in the context of Structure from Motion (SfM) in two ways: for robust estimation of the fundamental matrix and for the removal of outlying cameras, enhancing the robustness of the SfM pipeline. Numerical experiments confirm the state-of-the-art performance of our method in these applications. This research makes significant contributions to the field of robust subspace recovery, particularly in the context of computer vision and 3D reconstruction.

## 1. Introduction

In many applications, data has been collected in large quantities and dimensions. It is a common practice to represent such data within a low-dimensional subspace that preserves its essential information. Principal Component Analysis (PCA) is frequently employed to identify this subspace. However, PCA faces challenges when dealing with data contaminated by outliers. Consequently, the field of Robust Subspace Recovery (RSR) aims to develop a framework for outlier-robust PCA. RSR is particularly relevant to problems in computer vision, such as fundamental matrix estimation, which involves recovering a hidden subspace associated with “good correspondence pairs” among highly corrupted measurements.

Various algorithms have been proposed to address RSR, employing methods such as projection pursuit [1, 5, 14, 23, 29, 35, 38], subspace energy minimization (in particular least absolute

deviations and its relaxations) [9, 24, 27, 33, 34, 42, 43, 51], robust covariance estimation [50], filtering-based methods [4, 8, 49] and exhaustive subspace search methods [10, 16]. An in-depth exploration and comprehensive overview of robust subspace recovery and its diverse algorithms can be found in [25].

Methods based on robust covariance estimators, such as the Tyler’s M-estimator (TME), offer additional useful information on the shape of the data within the subspace, similarly to PCA in the non-robust setting. They also offer maximum-likelihood interpretation, which is missing in many other methods. Application of the TME [47] to RSR has been shown to be successful on basic benchmarks [25, 50]. Moreover, under a model of inliers in a general position on a subspace and outliers in general position in the complement of the subspace, TME was shown to recover the subspace within a desirable fraction of inliers [50]. Below this fraction it was proved to be Small Set Expansion (SSE) hard to solve the RSR problem [16].

One may still succeed with solving the RSR problem with a computationally efficient algorithm when the fraction of inliers is lower than the one required by [16], considering a more restricted data model or violating other assumptions made in [16]. For example, some special results in this direction are discussed in [32]. Also, [33] proposes the generalized haystack model of inliers and outliers to demonstrate the possibility of handling lower fractions of inliers by an RSR algorithm. This model extends the limited standard haystack model [27], where basic methods (such as PCA filtering) can easily work with low fractions of outliers. Nevertheless, it is unclear how practical the above theoretical ideas are for applied settings.

One practical setting that requires a fraction of inliers significantly lower than the one stated in [16] arises in the problem of robust fundamental (or essential) matrix estimation. The fundamental matrix encompasses the epipolar geometry of two views in stereo vision systems. It is typically computed using point correspondences between the two projected images. This computation requires finding an 8-dimensional subspace within a 9-dimensional ambient space. In this setting, the theoretical framework of [16] requires that the fraction of inliers be at least  $8/9 \approx 88.9\%$ , which is clearly unreasonable to require.

To date, the RANdom Sample Consensus (RANSAC)

\*This work was supported by NSF DMS awards 2124913 and 2318926.

<sup>†</sup>Supplementary code: <https://github.com/alexfgg/STE>

<sup>‡</sup>Corresponding author. All authors equally contributed.

method [10] is the only RSR method that has been highly successful in addressing this nontrivial scenario, gaining widespread popularity in computer vision. RANSAC is an iterative method that randomly selects minimal subsets of the data and fits models, in particular subspaces, to identify the best consensus set, that is, the set in most agreement with the hypothesized model. There are numerous approaches proposed to improve RANSAC, especially for this particular application, including locally optimized RANSAC (LO-RANSAC, [6]), maximum likelihood estimator RANSAC (MLE-SAC) [45]), degeneracy-check enabled RANSAC (DEGENSAC) [7]) and M-estimator guided RANSAC (MAGSAC) [2]). A near recovery theory for a variant of RANSAC under some assumptions on the outliers was suggested in [32]. Nevertheless, in general, RANSAC is rather slow and its application to higher-dimensional problems is intractable.

This work introduces a novel RSR algorithm that is guaranteed to robustly handle a lower fraction of outliers than the theoretical threshold proposed by [16], under special settings. Our basic idea is to adapt Tyler’s M-Estimator to utilize the information of the underlying  $d$ -dimensional subspace, while avoiding estimation of the full covariance. By using less degrees of freedom we obtain a more accurate subspace estimator than the one obtained by TME with improved computational complexity. We show that STE is a fusion of the Tyler’s M-estimator (TME) and a variant of the fast median subspace (FMS) [24] that aims to minimize a subspace-based  $\ell_0$  energy.

Our theory shows that our proposed subspace-constrained Tyler’s estimator (STE) algorithm can effectively recover the underlying subspace, even when it contains a smaller fraction of inliers relative to other methods. We obtain this nontrivial achievement first in a generic setting, where we establish when an initial estimator for STE is sufficiently well-conditioned to guarantee the desired robustness of STE. We then assume the asymptotic generalized haystack model and show that under this model, TME itself is a well-conditioned initial estimator for STE, and that unlike TME, STE with this initialization can deal with a lower fraction of inliers than the theoretical threshold specified in [16].

We demonstrate competitive performance in robust fundamental matrix estimation, relying solely on subspace information without additional methods for handling degenerate scenarios, in contrast to [7, 12, 37]. We also propose a potential application of RSR for removing bad cameras in order to enhance the SfM pipeline and show competitive performance of STE. This is a completely new idea and it may require additional exploration to make it practical. Nevertheless, it offers a very different testbed where  $N = D$  is very large and RANSAC is generally intractable.

The rest of the paper is organized as follows: §2 introduces the STE framework, §3 establishes theoretical guarantees of STE, §4 applies STE to two different problems in SfM, demonstrating its competitive performance relative to existing algorithms, and §5 provides conclusions and future directions.

## 2. The STE Algorithm

We present our proposed STE. We review basic notation in §2.1 and Tyler’s original estimator in §2.2. We describe our method in §2.3, its computational complexity in §2.4, its algorithmic choices in §2.5 and an interpretation for it as a fusion of TME and FMS with  $p=0$  in §2.6.

### 2.1. Notation

We use bold upper and lower case letters for matrices and column vectors, respectively. Let  $\mathbf{I}_k$  denote the identity matrix in  $\mathbb{R}^{k \times k}$ , where if  $k$  is obvious we just write  $\mathbf{I}$ . For a matrix  $\mathbf{A}$ , we denote by  $\text{tr}(\mathbf{A})$  and  $\text{Im}(\mathbf{A})$  the trace and image (i.e., column space) of  $\mathbf{A}$ . We denote by  $S_+(D)$  and  $S_{++}(D)$  the sets of positive semidefinite and definite matrices in  $\mathbb{R}^{D \times D}$ , respectively. We denote by  $O(D, d)$  the set of semi-orthogonal  $D \times d$  matrices, i.e.,  $\mathbf{U} \in O(D, d)$  if and only if  $\mathbf{U} \in \mathbb{R}^{D \times d}$  and  $\mathbf{U}^\top \mathbf{U} = \mathbf{I}_d$ . We refer to linear  $d$ -dimensional subspaces as  $d$ -subspaces. For a  $d$ -subspace  $L$ , we denote by  $\mathbf{P}_L$  the  $D \times D$  matrix representing the orthogonal projector onto  $L$ . We also arbitrarily fix  $\mathbf{U}_L$  in  $O(D, d)$  such that  $\mathbf{U}_L \mathbf{U}_L^\top = \mathbf{P}_L$  (such  $\mathbf{U}_L$  is determined up to right multiplication by an orthogonal matrix in  $O(d, d)$ ). Throughout the paper,  $\mathcal{X} = \{\mathbf{x}_i\}_{i=1}^N \subset \mathbb{R}^D$  is assumed to be a given centered dataset, that is,  $\sum_{i=1}^N \mathbf{x}_i = \mathbf{0}$ .

### 2.2. Tyler’s Estimator and its Application to RSR

Tyler’s M-estimator (TME) [47] robustly estimates the covariance  $\Sigma^*$  of the dataset  $\mathcal{X} = \{\mathbf{x}_i\}_{i=1}^N \subset \mathbb{R}^D$  by minimizing

$$\frac{D}{N} \sum_{i=1}^N \log(\mathbf{x}_i^\top \Sigma^{-1} \mathbf{x}_i) + \log \det(\Sigma) \quad (1)$$

over  $\Sigma \in S_{++}(D)$  such that  $\text{tr}(\Sigma) = 1$ . The cost function in (1) can be motivated by writing the maximum likelihood of the multivariate  $t$ -distribution and letting its degrees of freedom parameter,  $\nu$ , approach zero [31]. This cost function is invariant to dilations of  $\Sigma$ , and the constraint on  $\text{tr}(\Sigma)$ , whose value can be arbitrarily chosen, fixes a scale. TME also applies to scenarios where the covariance matrix does not exist. In such cases, TME estimates the shape (or scatter matrix) of the distribution, which is defined up to an arbitrary scale. More direct interpretations of TME as a maximum likelihood estimator can be found in [11, 46]. When  $D$  is fixed and  $N$  approaches infinity, TME is the “most robust” estimator of the shape matrix for data i.i.d. sampled from a continuous elliptical distribution [47] in a minimax sense, that is, as a minimizer of the maximal variance.

Tyler [47] proposed the following iterative formula for computing TME:

$$\Sigma^{(k)} = \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^\top}{\mathbf{x}_i^\top (\Sigma^{(k-1)})^{-1} \mathbf{x}_i} / \text{tr} \left( \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^\top}{\mathbf{x}_i^\top (\Sigma^{(k-1)})^{-1} \mathbf{x}_i} \right).$$

Kent and Tyler [22] proved that if any  $d$ -subspace of  $\mathbb{R}^D$ , where  $1 \leq d \leq D - 1$ , contains fewer than  $Nd/D$  data points, then

the above iterative procedure converges to TME. Linear rate of convergence was proved for the regularized TME in [15] and for TME in [13].

One can apply the TME estimator to solve the RSR problem with a given dimension  $d$  by forming the subspace spanned by the top  $d$  eigenvectors of TME. Zhang [50] proved that as long as there are more than  $Nd/D$  inliers lying on a subspace, and the projected coordinates of these inliers on the  $d$ -subspace and the projected coordinates of the outliers on the  $(D-d)$ -dimensional orthogonal complement of the subspace are in general position, then TME recovers this subspace. Zhang [50] also showed that in this setting the above iterative formula converges (note that the condition of convergence in [22] does not apply in this case). The above lower bound of  $Nd/D$  on the number of inliers coincides with the general bound for the noiseless RSR problem, beyond which the problem becomes Small Set-Expansion (SSE) hard [16].

Numerical experiments in [50] and [25] indicated state-of-the-art accuracy of TME compared to other RSR algorithms in various settings. The computational complexity of TME is of order  $O(K(ND^2 + D^3))$ , where  $K$  is the number of iterations. On the other hand, the cost of faster RSR algorithms is of order  $O(KNDd)$  [24, 25, 33].

### 2.3. Motivation and Formulation of STE

We aim to use more cleverly the  $d$ -subspace information within the TME framework to form an RSR algorithm, instead of first estimating the full covariance. By using less degrees of freedom we can obtain a more accurate subspace estimator, especially when the fraction of outliers can be large. Furthermore, our idea allows us to improve the computational cost to become state-of-the-art for high-dimensional settings.

Many RSR algorithms can be formulated as minimizing a best orthogonal projector onto a  $d$ -subspace [24, 25, 27, 33, 51]. We are going to do something similar, but unlike using an orthogonal projector, we will still use information from TME to get the shape of the data on the projected subspace. We will make the rest of the eigenvalues (i.e., bottom  $D-d$  ones) equal and shrink them by a parameter  $0 < \gamma < 1$ . We thus use a regularized version of a reduced-dimension covariance matrix. This parameter  $\gamma$  plays a role in our theoretical estimates. Making  $\gamma$  smaller helps with better subspace recovery, whereas making  $\gamma$  bigger enhances the well-conditioning of the estimator.

Following these basic ideas, we formulate our method, STE. For simplicity, we utilize covariance matrices and their inverses. Since these covariance matrices are essentially  $d$ -dimensional and include an additional simple regularizing component, our overall computations can be expressed in terms of the computation of the top  $d$  singular values of an  $N \times D$  matrix (see §2.4).

At iteration  $k$  we follow a similar step to that of TME:

$$\mathbf{Z}^{(k)} := \sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^\top / (\mathbf{x}_i^\top (\boldsymbol{\Sigma}^{(k-1)})^{-1} \mathbf{x}_i).$$

We compute the eigenvalues  $\{\sigma_i\}_{i=1}^D$  of  $\mathbf{Z}^{(k)}$  and replace each of the bottom  $(D-d)$  of them with  $\gamma \cdot \sigma_{d+1,D}$ , where

$$\sigma_{d+1,D} := \frac{1}{D-d} \sum_{i=d+1}^D \sigma_i. \quad (2)$$

We also compute the eigenvectors of  $\mathbf{Z}^{(k)}$  and form the matrix  $\boldsymbol{\Sigma}^{(k)}$  with the same eigenvectors as those of  $\mathbf{Z}^{(k)}$  and the modified eigenvalues, scaled to have trace 1. We iteratively repeat this procedure until the two estimators are sufficiently close. Algorithm 1 summarizes this procedure. Note that it is invariant to scaling of the data, similarly to TME.

---

#### Algorithm 1 STE: Subspace-Constrained Tyler's Estimator

---

- 1: **Input:**  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathbb{R}^{D \times N}$ : centered data matrix,  $d$ : subspace dimension,  $K$ : maximum number of iterations,  $\tau, \gamma$ : parameters.
  - 2: **Output:**  $L$ :  $d$ -subspace in  $\mathbb{R}^D$
  - 3:  $\boldsymbol{\Sigma}^{(0)} = \mathbf{I}_D / D$
  - 4: **for**  $k = 1, 2, \dots$  **do**
  - 5:    $\mathbf{Z}^{(k)} \leftarrow \sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^\top / (\mathbf{x}_i^\top (\boldsymbol{\Sigma}^{(k-1)})^{-1} \mathbf{x}_i)$
  - 6:    $[\mathbf{U}^{(k)}, \mathbf{S}^{(k)}, \mathbf{U}^{(k)}] \leftarrow \text{EVD}(\mathbf{Z}^{(k)})$
  - 7:    $\sigma_i \leftarrow [\mathbf{S}^{(k)}]_{ii}$  and  $\sigma_{d+1,D} \leftarrow \sum_{i=d+1}^D \sigma_i / (D-d)$
  - 8:    $\tilde{\mathbf{S}}^{(k)} \leftarrow \text{diag}(\sigma_1, \dots, \sigma_d, \gamma \cdot \sigma_{d+1,D}, \dots, \gamma \cdot \sigma_{d+1,D})$ ,
  - 9:    $\boldsymbol{\Sigma}^{(k)} \leftarrow \mathbf{U}^{(k)} \tilde{\mathbf{S}}^{(k)} (\mathbf{U}^{(k)})^\top / \text{tr}(\mathbf{U}^{(k)} \tilde{\mathbf{S}}^{(k)} (\mathbf{U}^{(k)})^\top)$
  - 10:   Stop if  $k \geq K$  or  $\|\boldsymbol{\Sigma}^{(k)} - \boldsymbol{\Sigma}^{(k-1)}\|_F < \tau$
  - 11: **end for**
  - 12:  $L = \text{Span}$  of the first  $d$  columns of  $\mathbf{U}^{(k)}$
- 

### 2.4. Computational Complexity

Setting  $w_i^{(k)} = (\mathbf{x}_i^\top (\boldsymbol{\Sigma}^{(k-1)})^{-1} \mathbf{x}_i)^{-1}$ , we can express  $\mathbf{Z}^{(k)}$  as  $\mathbf{Z}^{(k)} = \tilde{\mathbf{X}} \tilde{\mathbf{X}}^\top$ , where  $\tilde{\mathbf{X}} = [(w_1^{(k)})^{1/2} \mathbf{x}_1, \dots, (w_N^{(k)})^{1/2} \mathbf{x}_N]$ . With some abuse of notation we denote by  $\sigma_1, \dots, \sigma_D$  the eigenvalues of  $\boldsymbol{\Sigma}^{(k-1)}$  (and not  $\boldsymbol{\Sigma}^{(k)}$ ). Since they are scaled to have trace 1,  $\sigma_{d+1,D} = (1 - \sum_{j=1}^d \sigma_j) / (D-d)$ . We thus only need the top  $d$  eigenvectors and top  $d$  eigenvalues of  $\boldsymbol{\Sigma}^{(k-1)}$  to update  $\tilde{w}_i^{(k)}$ . Therefore, the complexity of STE can be of order  $O(KNDd)$  if a special fast algorithm is utilized for computing only the top  $d$  eigenvectors.

### 2.5. Implementation Details

STE depends on the parameters  $K$ ,  $\tau$  and  $\gamma$  and the initialization of  $\boldsymbol{\Sigma}^{(0)}$ . The first two parameters are rather standard in iterative procedures and do not raise any concern.

Our theory sheds some light on possible choices of  $\gamma$  and in particular it indicates that the algorithm can be more sensitive to choices of  $\gamma$  when the quantity defined later in (3) is relatively small. In this case, it may be beneficial to try several values of  $\gamma$ . We propose here a constructive way of doing it. We first form a sequence of  $0 < \gamma \leq 1$ , e.g.,  $\gamma_k = 1/k$ ,  $k = 1, \dots, m$ . In order to

determine the best choice of  $\gamma$ , we compute the distance of each data point  $\mathbf{x}$  to each subspace  $L_k$ , corresponding to the choice of  $\gamma_k$ , where  $\text{dist}(\mathbf{x}, L_k) = \|\mathbf{x} - \mathbf{P}_{L_k} \mathbf{x}\|$ . We set a threshold  $\zeta$ , obtained by the median among all points and all subspaces and for each subspace,  $L_k$ , we count the number of the inliers with distance below this threshold. The best  $\gamma_k$  is determined according to the subspace yielding the largest number of inliers. We describe this procedure in Algorithm 2.

For simplicity, we initialize with  $\Sigma^{(0)} = \mathbf{I}_D / D$  and note that with this choice  $\Sigma^{(1)}$  reflects the trimmed covariance matrix and thus reflects the PCA subspace. One can also initialize with TME or other subspaces (see §3 where the theory of STE is discussed). One can further try several initialization (with possible random components) and use a strategy similar to Algorithm 2 to choose the best one.

At last, we remark that when computing  $\mathbf{Z}^{(k)}$  we want to ensure that  $\mathbf{x}_i^\top (\Sigma^{(k-1)})^{-1} \mathbf{x}_i$  cannot be zero and we thus add the arbitrarily small number  $10^{-15}$  to this value.

---

**Algorithm 2** Estimating best  $\gamma$  for STE

---

- 1: **Input:**  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathbb{R}^{D \times N}$ : centered data matrix,  $d$ : subspace dimension,  $\{\gamma_1, \dots, \gamma_m\}$ : a set of pre-selected  $\gamma$ 's.
  - 2: **Output:**  $\gamma^*$ : optimal  $\gamma$  among  $\{\gamma_1, \dots, \gamma_m\}$
  - 3: **for**  $j = 1, 2, \dots, m$  **do**
  - 4:    $L^{(j)} \leftarrow \text{STE}(\mathbf{X}, d, \gamma_j)$
  - 5:    $\mathcal{D}^{(j)} \leftarrow \{\text{dist}(\mathbf{x}_i, L^{(j)}) \mid \mathbf{x}_i \in \mathcal{X}\}$ .
  - 6: **end for**
  - 7: Set  $\zeta = \text{median}(\{\mathcal{D}^{(1)}, \dots, \mathcal{D}^{(m)}\})$
  - 8:  $j^* = \text{argmax}_{1 \leq j \leq m} |\mathcal{D}^{(j)}| < \zeta$
  - 9:  $\gamma^* = \gamma_{j^*}$
- 

## 2.6. STE fuses TME and a Variant of FMS

STE is formally similar to both TME and FMS. Indeed, at each iteration these algorithms essentially compute  $\Sigma^{(k+1)} = \sum_{i=1}^N w_i \mathbf{x}_i \mathbf{x}_i^\top$ , where  $w_i \equiv w_i(\Sigma^{(k)})$ . We summarize the formal weights for FMS (with any choice of  $p$  for minimizing an  $\ell_p$  energy in [24]), TME and STE. We ignore an additional scaling constant for TME and STE, obtained by dividing  $w_i \mathbf{x}_i \mathbf{x}_i^\top$  above by its trace, and a regularization parameter  $\delta$  for FMS. We express these formulas using the eigenvalues  $\sigma_1, \dots, \sigma_D$  and eigenvectors  $\mathbf{u}_1, \dots, \mathbf{u}_D$  of the weighted sample covariance,  $\sum_{i=1}^N w_i \mathbf{x}_i \mathbf{x}_i^\top$  for each method and  $\beta := \gamma \cdot \sigma_{d+1, D}$  (see (2)) as follows:

$$w_i^{\text{FMS}} = \frac{1}{\left(\sum_{j=d+1}^D (\mathbf{x}_i^\top \mathbf{u}_j)^2\right)^{(2-p)/2}},$$

$$w_i^{\text{TME}} = \frac{1}{\sum_{j=1}^D \sigma_j^{-1} (\mathbf{x}_i^\top \mathbf{u}_j)^2},$$

$$w_i^{\text{STE}} = \frac{1}{\sum_{j=1}^d \sigma_j^{-1} (\mathbf{x}_i^\top \mathbf{u}_j)^2 + \beta^{-1} \sum_{j=d+1}^D (\mathbf{x}_i^\top \mathbf{u}_j)^2}.$$

These weights aim to mitigate the impact of outliers in different ways. For FMS,  $\sum_{j=d+1}^D (\mathbf{x}_i^\top \mathbf{u}_j)^2$  is the squared distance of a data point  $\mathbf{x}_i$  to the subspace  $L$ . Thus for  $p < 2$ ,  $w_i^{\text{FMS}}$  is smaller for “subspace-outliers”, where the robustness to such outliers increases when  $p \geq 0$  decreases.

The weights of TME are inversely proportional to the squared Mahalanobis distance of  $\mathbf{x}_i$  to the empirical distribution. They mitigate the effect of “covariance-outliers”. If the dataset is concentrated on a  $k$ -subspace where  $k < d$ , then TME can provide smaller weights to points lying away from this subspace, unlike FMS that does not distinguish between points within the larger  $d$ -subspace.

We note that the weights of STE fuse the above two weights. Within a  $d$ -subspace, they use the shape of the data. They can thus avoid outliers within this  $d$ -subspace. Within the orthogonal component of this subspace, they use a term proportional to that of FMS with  $p=0$ . We remark that such  $\ell_0$  minimization has a clear interpretation for RSR, though is generally hard to guarantee. Indeed, [24] has no guarantees for FMS with  $p=0$ . It can also yield unwanted spurious stationary points [26].

## 3. Theory

We review a theoretical guarantee for STE, whose proof is given in [28]. It requires some conditions and we verify they hold with high probability under the asymptotic generalized haystack model. We assume a noiseless inliers-outliers RSR model. Let  $L_*$  denote the underlying  $d$ -subspace in  $\mathbb{R}^D$ ,  $\mathcal{X}_{in} = \mathcal{X} \cap L_*$  and  $\mathcal{X}_{out} = \mathcal{X} \setminus \mathcal{X}_{in}$  be the set of inliers and outliers, respectively, and  $n_1 = |\mathcal{X}_{in}|$  and  $n_0 = |\mathcal{X}_{out}|$  be the number of inliers and outliers. Our first assumption is a mild one on how well-conditioned the inliers are in  $L_*$  (compare e.g., other assumptions in [25, 32]).

**Assumption 1:** Any  $k$ -subspace of  $L_*$ ,  $1 \leq k \leq d$ , contains at most  $n_1 k / d$  points.

**Motivation for Assumption 2:** The ratio of inliers per outliers,  $n_1 / n_0$ , in RSR is often referred to as the SNR (signal-to-noise ratio) [25, 32, 33]. The smaller it is, the best the subspace recovery is. We define the dimension-scaled SNR (DS-SNR) as the SNR obtained when scaling  $n_1$  and  $n_0$  by their respective dimensions (of  $L_*$  and  $L_*^\perp$ ):

$$\text{DS-SNR} := \frac{n_1 / d}{n_0 / (D - d)}. \quad (3)$$

Zhang [50] showed that exact recovery by TME is guaranteed whenever  $\text{DS-SNR} > 1$  (assuming general position assumptions on the inliers and outliers) and Hardt and Moitra [16] showed that when considering general datasets with general position assumptions on the inliers and outliers, the RSR problem is SSE hard if the DS-SNR is lower than 1. We aim to show that under the following weaker generic condition, STE can obtain exact recovery with DS-SNR, strictly lower than 1.

**Assumption 2:**  $\text{DS-SNR} > \gamma$ , where  $\gamma < 1$  is the STE parameter.

Our last assumption requires a sufficiently good initialization for STE, but also implicitly involves additional hidden

assumptions on the inliers and outliers. This is expected, since Assumption 1 does not require anything from the outliers and also has a very weak requirement from the inliers. To formulate the new assumption we define below some basic condition numbers for good initialization (which are more complicated than the one for initialization by PCA suggested by [33] and [32]) and also quantities similar to the ones used to guarantee landscape stability in the theory of RSR [25, 27, 33, 51].

**Definitions required for Assumption 3:** Recall that  $\Sigma^{(0)}$  denotes the initial value in Algorithm 1, and denote

$$\Sigma_{L_1, L_2}^{(0)} = \mathbf{U}_{L_1}^\top \Sigma^{(0)} \mathbf{U}_{L_2}.$$

We define the following condition number

$$\kappa_1 = \frac{\sigma_d \left( \Sigma_{L_*, L_*}^{(0)} - \Sigma_{L_*, L_*^\perp}^{(0)} \Sigma_{L_*^\perp, L_*^\perp}^{(0)-1} \Sigma_{L_*^\perp, L_*}^{(0)} \right)}{\sigma_1 \left( \Sigma_{L_*^\perp, L_*^\perp}^{(0)} \right)}.$$

To get a better intuition to this primary quantity of Assumption 3, we first express the initial estimator  $\Sigma^{(0)}$ , using basis vectors for  $L_*$  and  $L_*^\perp$ , as a  $2 \times 2$  block matrix

$$\begin{pmatrix} \Sigma_{L_*, L_*}^{(0)} & \Sigma_{L_*, L_*^\perp}^{(0)} \\ \Sigma_{L_*^\perp, L_*}^{(0)} & \Sigma_{L_*^\perp, L_*^\perp}^{(0)} \end{pmatrix}.$$

Defining  $\Sigma' = \Sigma_{L_*, L_*}^{(0)} - \Sigma_{L_*, L_*^\perp}^{(0)} \Sigma_{L_*^\perp, L_*^\perp}^{(0)-1} \Sigma_{L_*^\perp, L_*}^{(0)}$ , we decompose this block matrix as

$$\begin{pmatrix} \Sigma' & \Sigma_{L_*, L_*^\perp}^{(0)} \\ \Sigma_{L_*^\perp, L_*}^{(0)} & \Sigma_{L_*^\perp, L_*^\perp}^{(0)} \end{pmatrix} + \begin{pmatrix} \Sigma_{L_*, L_*}^{(0)} - \Sigma' & 0 \\ 0 & 0 \end{pmatrix}.$$

We note that the numerator of  $\kappa_1$  is the  $d$ -th eigenvalue of the second matrix in the above sum. We show in [28] that this eigenvalue is positive if  $\Sigma^{(0)}$  is positive definite, which can be easily enforced. The condition number is thus the ratio between the smallest positive eigenvalue of the second matrix of the sum and the largest eigenvalue of the component of the first matrix associated with  $L_*^\perp$ . Therefore,  $\kappa_1$  expresses a ratio between a quantifier of a  $d$ -dimensional component of  $\Sigma^{(0)}$ , associated with  $L_*$ , and a quantifier of the projection onto  $L_*^\perp$  of a full rank component of  $\Sigma^{(0)}$ .

We also define  $\Sigma_{in,*}$  as the TME solution to the set of the projected inliers  $\{\mathbf{U}_{L_*} \mathbf{x} \mid \mathbf{x} \in \mathcal{X}_{in}\} \subset \mathbb{R}^d$  and the following two condition numbers

$$\kappa_2 = \frac{\sigma_1 \left( \Sigma_{L_*^\perp, L_*^\perp}^{(0)} \right)}{\sigma_D \left( \Sigma^{(0)} \right)} \quad \text{and} \quad \kappa_{in} = \frac{\sigma_1 \left( \Sigma_{in,*} \right)}{\sigma_d \left( \Sigma_{in,*} \right)}.$$

We note that  $\kappa_{in}$  is analogous to the condition number in (25) of [32], where we replace the sample covariance by the TME estimator. An analog to the alignment of outliers statistic [27, 33] for STE is

$$\mathcal{A} = \left\| \sum_{\mathbf{x} \in \mathcal{X}_{out}} \frac{\mathbf{x} \mathbf{x}^\top}{\|\mathbf{U}_{L_*^\perp} \mathbf{x}\|^2} \right\|.$$

An analog to the stability statistic [27, 33] for STE is

$$\mathcal{S} = \sigma_{d+1, D} \left( \sum_{\mathbf{x} \in \mathcal{X}} \frac{\mathbf{x} \mathbf{x}^\top}{\|\mathbf{x}\|^2} \right),$$

where  $\sigma_{d+1, D}(\mathbf{X})$  was defined in (2).

**Assumption 3:** There exists  $C = C(\gamma, \text{DS-SNR}) > 0$  such that

$$\kappa_1 \geq C \frac{d \kappa_{in} \mathcal{A}}{n_1} \left( \kappa_{in} + \frac{\mathcal{A}}{\frac{n_1}{d} - \gamma \frac{n_0}{D-d}} + \frac{\kappa_2 \mathcal{A}}{\gamma \mathcal{S}} (1 + \kappa_{in}) \right). \quad (4)$$

The exact technical requirement on  $C$  is specified in [28]. In general, the larger the RHS of (4), the more restricted the choice of  $\Sigma^{(0)}$  is. In particular, when  $\kappa_1 = \infty$ , the definition of  $\kappa_1$  implies that  $\text{Im}(\Sigma^{(0)}) = L_*$ , so the subspace is already recovered by the initial estimate. Therefore, reducing the lower bound of  $\kappa_1$  may allow some flexibility, so a marginally suboptimal initialization could still work out. In [28], we show that under the asymptotic generalized haystack model, Assumption 3 can be interpreted as an upper bound on the largest principal angle between the initial and ground truth subspaces.

**Generic Theory:** The next theorem suggests that under assumptions 1-3, STE nicely converges to an estimator that recovers  $L_*$ . The main significance of this theory is that its assumptions can allow DS-SNR lower than 1 for special instances of datasets (for which the assumptions hold), unlike the general recovery theories of [16] and [50].

**Theorem 1.** *Under assumptions 1-3, the sequence  $\Sigma^{(k)}$  generated by STE converges to  $\mathbf{U}_{L_*} \Sigma_{in,*} \mathbf{U}_{L_*}^\top$ , the TME solution for the set of inliers  $\mathcal{X}_{in}$ . In addition, let  $L^{(k)}$  be the subspace spanned by the top  $d$  eigenvectors of  $\Sigma^{(k)}$ , then the angle between  $L^{(k)}$  and  $L_*$ ,  $\angle(L^{(k)}, L_*) = \cos^{-1}(\|\mathbf{U}_{L^{(k)}}^\top \mathbf{U}_{L_*}\|)$ , converges  $r$ -linearly to zero.*

We discuss insights of this theory on choices of the algorithms and further verify the above stated advantage of STE over TME assuming a common probabilistic model.

**Choice of  $\gamma$  for subspace recovery:** In order to avoid too large lower bound for  $\kappa_1$  in (4), which we motivated above, it is good to find  $\epsilon_1$  and  $\epsilon_2 > 0$ , such that  $\gamma$  lies in  $(\epsilon_1, \text{DS-SNR} - \epsilon_2)$  (to notice this, observe the terms involving  $\gamma$  in the denominators of the last two additive terms in (4)). We thus note that if the DS-SNR is expected to be sufficiently larger than 1, we can use, e.g.,  $\gamma = 0.5$ , but when the DS-SNR can be close to 1 or lower (e.g., in fundamental matrix estimation), it is advisable to choose small values of  $\gamma$  according to Algorithm 2 and their sizes may depend on the expected value of the DS-SNR.

**Possible ways of Initialization:** If one expects an initial estimated subspace  $\hat{L}$  to have a sufficiently small angle  $\theta$  with  $L_*$ , where  $\theta = \angle(\hat{L}, L_*)$ , then for  $\Sigma^{(0)} := \Pi_{\hat{L}} + \epsilon \mathbf{I}$  it can be shown that  $\kappa_1 > O(1/(\epsilon + \theta))$  and  $\kappa_2 < O(1 + \frac{\theta}{\epsilon})$ . Thus one may use a trusted RSR method, e.g., FMS. As discussed in §2.5, the choice  $\Sigma^{(0)} = \mathbf{I}$  (or a scaled version of it) corresponds to  $\hat{L}$  being the

PCA subspace (obtained at iteration 1). Also, using the TME solution for  $\Sigma^{(0)}$  corresponds to using the TME subspace as  $\hat{L}$ .

**Theory under a probabilistic model:** We show that under a common probabilistic model, the assumptions of Theorem 1, where  $\Sigma^{(0)}$  is obtained by TME, hold. Moreover, we show that STE (initialized by TME) can recover the correct subspace in situations with DS-SNR < 1, whereas TME cannot recover the underlying subspace in such cases. We follow [33] and study the Generalized Haystack Model, though for simplicity, we assume Gaussian instead of sub-Gaussian distributions and an asymptotic setting. We assume  $n_1$  inliers i.i.d. sampled from a Gaussian distribution  $N(0, \Sigma^{(in)}/d)$ , where  $\Sigma^{(in)} \in S_+(D)$  and  $L_* = \text{Im}(\Sigma^{(in)})$  (so  $\Sigma^{(in)}$  has  $d$  nonzero eigenvalues), and  $n_0$  outliers are i.i.d. sampled from a Gaussian distribution  $N(0, \Sigma^{(out)}/D)$ , where  $\Sigma^{(out)}/D \in S_{++}(D)$ . We define the following condition numbers of inliers (in  $L_*$ ) and outliers:

$$\kappa_{in} = \frac{\sigma_1(\Sigma^{(in)})}{\sigma_d(\Sigma^{(in)})} \quad \text{and} \quad \kappa_{out} = \frac{\sigma_1(\Sigma^{(out)})}{\sigma_D(\Sigma^{(out)})}.$$

Clearly, Assumption 1 holds under this model, and Assumption 2 constrains some of its parameters. Our next theorem shows that Assumption 3 holds under this model when the initial estimate  $\Sigma^{(out)}$  for STE is obtained by TME. It also shows that in this case STE can solve the RSR problem even when DS-SNR < 1, unlike TME. For simplicity, we formulate the theory for the asymptotic case, where  $N \rightarrow \infty$  and the theorem holds almost surely. It is possible to formulate it for a very large  $N$  with high probability, but it requires stating complicated constants depending on various parameters.

**Theorem 2.** *Assume data generated from the above generalized haystack model. Assume further that for  $0 < \mu < 1$ , which can be arbitrarily small,  $d < (1 - \mu)D - 2$ . Then, for any chosen  $0 < c_0 < 1$ , which is a lower bound for  $\gamma$ , there exists  $\eta := \eta(\kappa_{in}, \kappa_{out}, c_0, \mu) < 1$  such that if DS-SNR  $\geq \eta$  and  $\Sigma^{(0)}$  is obtained by TME, then Assumption 3 for  $\Sigma^{(0)}$  is satisfied with  $c_0 < \gamma < \eta - c_0$  almost surely as  $N \rightarrow \infty$ . Consequently, the output of the STE algorithm, initialized by TME and with the choice of  $c_0 < \gamma < \eta - c_0$ , recovers  $L_*$ . On the other hand, if  $\Sigma_{L_*, L_*^\perp}^{(out)} \neq 0$  and DS-SNR < 1, then the top  $d$  eigenvectors of TME do not recover  $L_*$ .*

There are three different regimes that the theorem covers. When DS-SNR  $\geq 1$ , both TME+STE (i.e., STE initialized by TME) and TME solve the RSR problem. When  $\eta \leq \text{DS-SNR} < 1$ , TME+STE solves the RSR problem and TME generally fails. When  $\gamma \leq \text{DS-SNR} < \eta$ , TME+STE might also fail, but STE with extremely good initialization (that satisfies Assumption 3) can still solve the problem.

To get a basic idea of the dependence of  $\eta$  on its parameters, we remark that  $\eta \rightarrow 1$  if either  $c_0 \rightarrow 0$ ,  $\kappa_{in} \rightarrow \infty$ ,  $\kappa_{out} \rightarrow \infty$  or  $\mu \rightarrow 0$ , where the parameter  $\mu$  is somewhat artificial and might be removed with a tighter proof. Therefore, successful

performance of TME+STE requires a DS-SNR that is close to 1 when  $\gamma$  is close to either 0 or  $\eta$  (so that  $c_0$  is very small) or when either the inlier or outlier distribution is highly non-symmetric, that is, either  $\kappa_{in}$  or  $\kappa_{out}$  is large.

## 4. Applications to Structure from Motion

We apply STE to problems relevant to SfM: robust estimation of fundamental matrices (see §4.1), and initial screening of undesirable cameras (see §4.2).

### 4.1. Robust Fundamental Matrix Estimation

Fundamental matrix estimation from noisy and inexact keypoint matches is a core computer vision problem. It provides a challenging setting for applying RSR methods.

We review this setting as follows. Let  $(\mathbf{x}, \mathbf{x}') \in \mathbb{R}^3 \times \mathbb{R}^3$  be a correspondence pair of two points in different images that are projections of the same 3D point in the scene, where  $\mathbf{x}$  and  $\mathbf{x}'$  are expressed by homogeneous coordinates of planar points. The fundamental matrix  $\mathbf{F} \in \mathbb{R}^{3 \times 3}$  relates these corresponding points and the epipolar lines they lie on as follows:  $\mathbf{x}'^\top \mathbf{F} \mathbf{x} = 0$  [17], or equivalently,

$$\text{vec}(\mathbf{F}) \cdot \text{vec}(\mathbf{x}\mathbf{x}'^\top) = 0. \quad (5)$$

where  $\text{vec}(\cdot)$  denotes the vectorized form of a matrix. Therefore, ideally, the set of all vectors in  $\mathbb{R}^9$  of the form  $\text{vec}(\mathbf{x}\mathbf{x}'^\top)$ , where  $(\mathbf{x}, \mathbf{x}') \in \mathbb{R}^3 \times \mathbb{R}^3$  is a correspondence pair, lies on an 8-subspace in  $\mathbb{R}^9$  and its orthogonal complement yields the fundamental matrix. In practice, the measurements of correspondence pairs can be highly corrupted due to poor matching. Moreover, some choices of correspondence pairs and the corruption mechanism may lead to concentration on low-dimensional subspaces within the desired 8-subspace. Furthermore, the corruption mechanism can lead to nontrivial settings of outliers. Lastly, since  $d=8$  and  $D=9$ , the theoretical threshold of [16] translates to having the fraction of inliers among all data points be at least  $8/9 \approx 88.9\%$ .

Therefore, this application is often a very challenging setting for direct RSR methods. The best performing RSR methods to date for fundamental matrix estimation are variants of RANSAC [10]. RANSAC avoids any subspace-modeling assumptions, but estimates the subspace based on testing myriads of samples, each having 7 or 8 point correspondences [17].

We test the performance of STE in estimating the fundamental matrix on the Photo Tourism database [41], where the image correspondences are obtained by SIFT feature similarities [30]. We compare STE with the following 3 top RSR performers according to [25]: FMS [24], spherical FMS (SFMS) [24] and TME [47, 50]. We also compared with vanilla RANSAC [10] and two of its specialized extensions, which are state-of-the-art performers for estimating fundamental matrices: locally optimized RANSAC (LO-RANSAC) [6] and degeneracy-check enabled RANSAC (DEGENSAC) [7]. For the RSR methods we used codes from the supplementary material of [25]

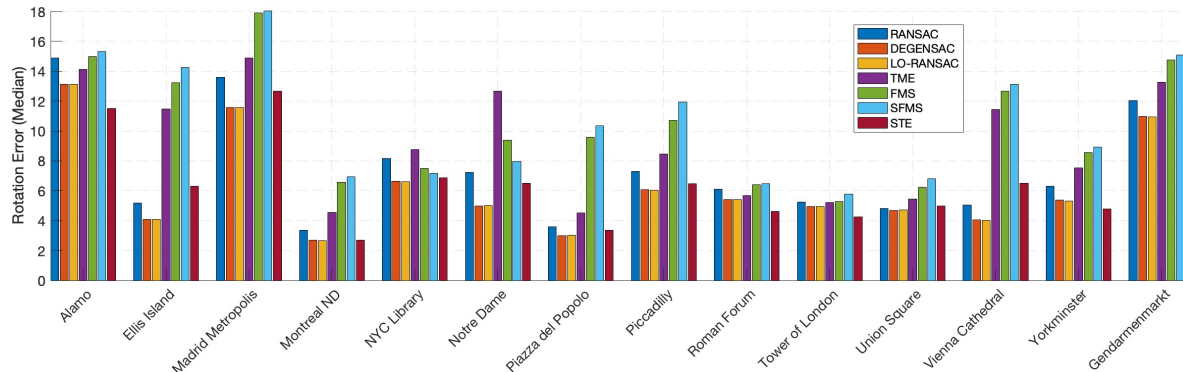


Figure 1. Median (relative) rotation errors obtained by seven algorithms for the 14 datasets of Photo Tourism.

with their default options. We further used the Python package `pydegensac` for implementing LO-RANSAC and DEGENSAC with the inlier threshold  $\eta=0.75$ . For STE, we used Algorithm 2 to estimate the best  $\gamma$  from  $\{(2i)^{-1}\}_{i=1}^5$ .

We measure the accuracy of the results according to the median and mean errors of relative rotation and direction vectors directly obtained by the fundamental matrices for each method. For computing these errors, we compared with ground-truth values provided by [41, 48]. Figure 1 describes the result of the mean errors for relative rotation per dataset of Photo Tourism, where the other three errors and  $\text{mAA}(10^\circ)$  are in the supplemental material. STE is significantly better than top RSR performers (TME, FMS and SFMS). Overall, it appears that STE performs better than vanilla RANSAC, except for the Ellis Island and Vienna Cathedral datasets, where RANSAC outperforms STE. STE is still competitive when compared with LO-RANSAC and DEGENSAC, except for Notre Dame and the latter two datasets.

## 4.2. Initial Camera Removal for SfM

We propose a novel application of RSR for SfM and test STE for this application. Even though our framework is not sufficiently practical at this point, it allows testing STE in a different setting where  $N=D$  is very large and  $d=6$ . Our idea is to use RSR within the SfM pipeline right after estimating the fundamental matrices, in order to remove some cameras that result in inaccurate estimated fundamental matrices. The hope is that eventually such methods may reduce corruption and speed up the costly later computationally intensive stages of the global SfM pipeline.

There are two main reasons to question such a process. One may first question the gain in improving accuracy. Indeed, since the rest of the pipeline already identifies corrupted pairwise measurements, this process may not improve accuracy and may even harm it as it removes whole cameras and not pairs of cameras. That is, it is possible that a camera, which results in bad pairwise measurement, also contributes to some other accurate pairwise estimates that can improve the overall accuracy. The second concern is in terms of speed. In general,

the removal of cameras may result in higher or comparable speed. Indeed, the LUD global pipeline [36], which we follow, examines the parallel rigidity of the viewing graph and extracts the maximal parallel rigid subgraph. Thus earlier removal of cameras may worsen the parallel rigidity of the graph and increase the computation due to the need of finding a maximal parallel rigid subgraph. For example, [40] removes cameras in an earlier stage of the LUD pipeline, but results in higher computational cost than the LUD pipeline. Therefore, improvement of speed for the LUD pipeline by removing cameras is generally non-trivial. Moreover, currently we use scale factors obtained by first running LUD, so we do not get a real speed improvement. Nevertheless, the proposed method is insightful whenever it may indicate clear improvement in accuracy for a dataset, since one can then infer that the current pipeline is not effective enough in handling corrupted measurements, which can be easily recognized by a simple method. Furthermore, improvement in “speed” can be indicative of maintaining parallel rigidity.

Our RSR formulation is based on a fundamental observation by Sengupta et al. [39] on the low-rank of the  $n$ -view essential (or fundamental) matrix. The  $n$ -view essential matrix  $\mathbf{E}$  of size  $3n \times 3n$  is formed by stacking all  $\binom{n}{2}$  essential matrices, while being appropriately scaled. That is, the  $ij$ -th block of  $\mathbf{E}$  is the essential matrix for the  $i$ -th and  $j$ -th cameras, where each  $\mathbf{E}_{ij}$  is scaled by a factor  $\lambda_{ij}$  in accordance with the global coordinate system (see [20, 21, 39]). It was noticed in [39] that  $\mathbf{E}$  has rank 6. Moreover, [39] characterized the set of  $n$ -view essential matrices whose camera centers are not all collinear by the satisfaction of a few algebraic conditions, where the major one is  $\text{rank}(\mathbf{E})=6$ . Further explanation appears in [20].

We propose a straightforward application of RSR, utilizing these ideas to initially eliminate cameras that introduce significant corruption to the essential matrices. For this purpose, we compute the essential matrices (by computing first the fundamental matrices and then using the known camera calibration) and scale each matrix according to the factor obtained by the LUD pipeline [36] (note that this is the initial

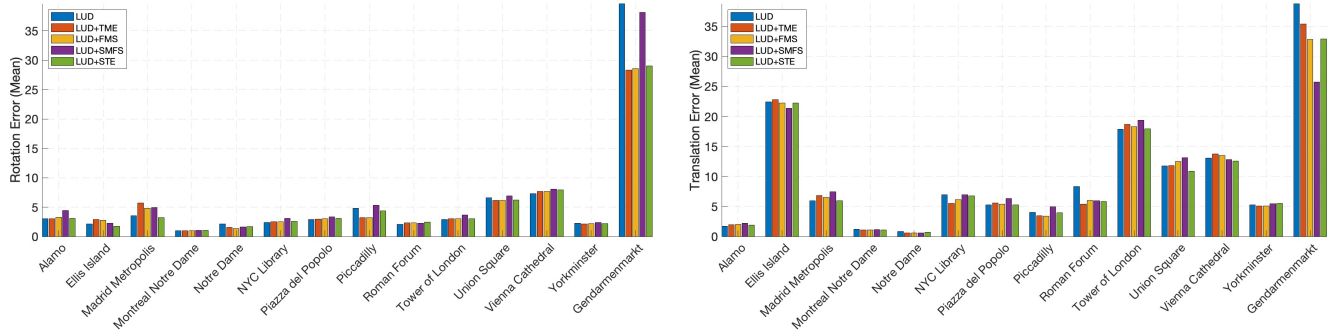


Figure 2. Mean (absolute) rotation errors (in degrees, left) and mean translation errors (in degrees, right) of LUD and four RSR methods used to initially screen bad cameras within LUD applied to the 14 datasets of Photo Tourism.

scaling applied in [20, 21, 39] before applying a non-convex and nontrivial optimization procedure that refines such scales). Using these appropriately scaled essential matrices, we form the  $n$ -view essential matrix  $\mathbf{E}$  of size  $3n \times 3n$ . We denote the  $3n \times 3$  column blocks of  $\mathbf{E}$  by  $\mathbf{E}_{:,1}, \dots, \mathbf{E}_{:,n}$  (since  $\mathbf{E}$  is symmetric they are the same as the row blocks transposed). We treat  $\mathbf{E}$  as a data matrix with  $D = N = 3n$ , where the columns of  $\mathbf{E}$  are the data points. We apply RSR with  $d=6$ , recover a  $d$ -dimensional robust subspace and identify the outlying columns whose distance is largest from this subspace. To avoid heuristic methods for the cutoff of outliers we assume a fixed percentage of 20% outlying columns. If a column block,  $\mathbf{E}_{:,i}$  contains an outlying column, we remove its corresponding camera  $i$ . Consequently, a smaller percentage of cameras (about 10 – 15%) will be eliminated.

We use the Photo Tourism database [41] with precomputed pairwise image correspondences provided by [39] (they were obtained by thresholding SIFT feature similarities). To compute scale factors for the essential matrices we use the output of the LUD pipeline [36] as follows (following an idea proposed in [39] for initializing these values): Given the essential matrix for cameras  $i$  and  $j$  computed at an early stage of our pipeline,  $\mathbf{E}_{ij}$ , and the one obtained by the full LUD pipeline,  $\mathbf{E}_{ij}^{\text{LUD}}$ , the scaling factor is  $\lambda_{ij} = \langle \mathbf{E}_{ij}, [\mathbf{E}_{ij}^{\text{LUD}}] \rangle / \|\mathbf{E}_{ij}^{\text{LUD}}\|_F^2$ . Since many values of  $\mathbf{E}_{ij}$  are missing, we also apply matrix completion.

We compare the LUD pipeline with the LUD pipeline combined with the filtering processes achieved by STE, FMS, SFMS, and TME. For STE we fix  $\gamma=1/3$ , though any other value we tried yielded the same result. We report both mean and median errors of rotations and translations and runtime of the standard LUD and the RSR+LUD methods with initial screening of cameras. Figure 2 shows the mean rotation and translation errors, where the rest of the figures and a summarizing table are in the supplementary material. In general, STE demonstrates slightly higher accuracy compared to other RSR methods. Improved accuracy is particularly notable when matrix completion is not utilized, as demonstrated in the supplementary material. We observe that LUD+STE generally improves the estimation of camera parameters (both rotations and translations) over LUD.

The improvement of LUD+STE is noticeable in Roman Forum and Gendarmenmarkt. In the supplementary material we show further improvement for Gendarmenmarkt with the removal of 45% outlying columns. While the resulting errors are still large, their improvement shows some potential in dealing with difficult SfM structure by initially removing cameras in a way that may help eliminate some scene ambiguities, which are prevalent in Gendarmenmarkt. In terms of runtime, both LUD+STE and LUD+SFMS demonstrate significant improvements, where LUD+SFMS is even faster than LUD+STE. While this does not yet imply faster handling of the datasets (as we use initial scaling factors obtained by LUD), it indicates the efficiency of the removal of outliers in maintaining parallel rigidity.

## 5. Conclusions

We introduce STE, a meticulously crafted adaptation of TME designed to address challenges within RSR. Theoretical guarantees demonstrate its ability to recover the true underlying subspace reliably, even with a smaller fraction of inliers compared to the well-known theoretical threshold. Under the generalized haystack model, we show that this initialization can be chosen as TME itself, leading to improved handling of a smaller fraction of inliers compared to TME. Our exploration extends to practical applications, where STE proves effective in two 3D vision tasks: robust fundamental matrix estimation and screening of bad cameras for improved SfM.

Several avenues for future research include: • Exploring adaptations of other robust covariance estimation methods to RSR. • Studying effective initialization for STE both in theory and in practice. • In-depth theoretical exploration of the optimal choice of the parameter  $\gamma$ . • Study of alternative ways of adapting TME to RSR problems. • Improving STE for fundamental matrix estimation following ideas similar to those in [7, 12, 37] for addressing challenging degeneracies. • Enhancing our initial idea of initial removal of bad cameras, specifically attempting to use it to rectify challenging scene ambiguities. • Testing our methods for SfM using more recent feature matching algorithms.



## References

- [1] Larry P. Ammann. Robust singular value decompositions: A new approach to projection pursuit. *Journal of the American Statistical Association*, 88(422):pp. 505–514, 1993. [1](#)
- [2] Daniel Barath, Jiri Matas, and Jana Noskova. MAGSAC: marginalizing sample consensus. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10197–10205, 2019. [2](#)
- [3] Jian-Feng Cai, Emmanuel J. Candès, and Zuowei Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on optimization*, 20(4):1956–1982, 2010. [17](#)
- [4] Yeshwanth Cherapanamjeri, Prateek Jain, and Praneeth Netrapalli. Thresholding based outlier robust PCA. In *Conference on Learning Theory*, pages 593–628. PMLR, 2017. [1](#)
- [5] Vartan Choulakian.  $L_1$ -norm projection pursuit principal component analysis. *Computational Statistics & Data Analysis*, 50(6):1441–1451, 2006. [1](#)
- [6] Ondřej Chum, Jiri Matas, and Josef Kittler. Locally optimized RANSAC. In *Pattern Recognition: 25th DAGM Symposium, Magdeburg, Germany, September 10-12, 2003. Proceedings 25*, pages 236–243. Springer, 2003. [2](#), [6](#), [13](#)
- [7] Ondrej Chum, Tomas Werner, and Jiri Matas. Two-view geometry estimation unaffected by a dominant plane. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pages 772–779. IEEE, 2005. [2](#), [6](#), [8](#), [13](#)
- [8] Ilias Diakonikolas, Gautam Kamath, Daniel M. Kane, Jerry Li, Ankur Moitra, and Alistair Stewart. Robustly learning a gaussian: Getting optimal error, efficiently. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, page 2683–2702, USA, 2018. Society for Industrial and Applied Mathematics. [1](#)
- [9] Chris Ding, Ding Zhou, Xiaofeng He, and Hongyuan Zha. R1-PCA: rotational invariant  $L_1$ -norm principal component analysis for robust subspace factorization. In *ICML '06: Proceedings of the 23rd international conference on Machine learning*, pages 281–288, New York, NY, USA, 2006. ACM. [1](#)
- [10] Martin A. Fischler and Robert C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. [1](#), [2](#), [6](#), [13](#)
- [11] Gabriel Frahm and Uwe Jaekel. A generalization of Tyler’s M-estimators to the case of incomplete data. *Computational Statistics & Data Analysis*, 54(2):374–393, 2010. [2](#)
- [12] J-M Frahm and Marc Pollefeys. RANSAC for (quasi-) degenerate data (QDEGSAC). In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, pages 453–460. IEEE, 2006. [2](#), [8](#)
- [13] William Cole Franks and Ankur Moitra. Rigorous guarantees for Tyler’s M-estimator via quantum expansion. In *Conference on Learning Theory*, pages 1601–1632. PMLR, 2020. [3](#)
- [14] Jerome H. Friedman and John W. Tukey. A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers*, C-23(9):881–890, 1974. [1](#)
- [15] John Goes, Gilad Lerman, and Boaz Nadler. Robust sparse covariance estimation by thresholding Tyler’s M-estimator. *The Annals of Statistics*, 48(1):86 – 110, 2020. [3](#)
- [16] Moritz Hardt and Ankur Moitra. Algorithms and hardness for robust subspace recovery. In *Conference on Learning Theory*, pages 354–375. PMLR, 2013. [1](#), [2](#), [3](#), [4](#), [5](#), [6](#)
- [17] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003. [6](#), [11](#), [12](#), [13](#)
- [18] Yuhe Jin, Dmytro Mishkin, Anastasiia Mishchuk, Jiri Matas, Pascal Fua, Kwang Moo Yi, and Eduard Trulls. Image matching across wide baselines: From paper to practice. *International Journal of Computer Vision*, 129(2):517–547, 2021. [13](#)
- [19] Arman Karimian and Roberto Tron. Essential matrix estimation using convex relaxations in orthogonal space. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 17142–17152, 2023. [15](#)
- [20] Yoni Kasten, Amnon Geifman, Meirav Galun, and Ronen Basri. Algebraic characterization of essential matrices and their averaging in multiview settings. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5895–5903, 2019. [7](#), [8](#)
- [21] Yoni Kasten, Amnon Geifman, Meirav Galun, and Ronen Basri. Gpsfm: Global projective sfm using algebraic constraints on multi-view fundamental matrices. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3264–3272, 2019. [7](#), [8](#)
- [22] John T. Kent and David E. Tyler. Maximum likelihood estimation for the wrapped Cauchy distribution. *Journal of Applied Statistics*, 15(2):247–254, 1988. [2](#), [3](#)
- [23] Nojun Kwak. Principal component analysis based on  $L_1$ -norm maximization. *IEEE transactions on pattern analysis and machine intelligence*, 30(9):1672–1680, 2008. [1](#)
- [24] Gilad Lerman and Tyler Maunu. Fast, robust and non-convex subspace recovery. *Information and Inference: A Journal of the IMA*, 7(2):277–336, 2018. [1](#), [2](#), [3](#), [4](#), [6](#), [13](#)
- [25] Gilad Lerman and Tyler Maunu. An overview of robust subspace recovery. *Proceedings of the IEEE*, 106(8):1380–1410, 2018. [1](#), [3](#), [4](#), [5](#), [6](#)
- [26] Gilad Lerman and Teng Zhang.  $l_p$ -recovery of the most significant subspace among multiple subspaces with outliers. *Constr. Approx.*, 40(3):329–385, 2014. [4](#)
- [27] Gilad Lerman, Michael B. McCoy, Joel A. Tropp, and Teng Zhang. Robust computation of linear models by convex relaxation. *Found. Comput. Math.*, 15(2):363–410, 2015. [1](#), [3](#), [5](#)
- [28] Gilad Lerman, Feng Yu, and Teng Zhang. Theoretical guarantees for the subspace-constrained Tyler’s estimator, 2024. [4](#), [5](#), [11](#)
- [29] Guoying Li and Zhonglian Chen. Projection-pursuit approach to robust dispersion matrices and principal components: primary theory and monte carlo. *Journal of the American Statistical Association*, 80(391):759–766, 1985. [1](#)
- [30] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60:91–110, 2004. [6](#)
- [31] Ricardo A. Maronna and Víctor J. Yohai. Robust estimation of multivariate location and scatter. *Wiley StatsRef: Statistics Reference Online*, pages 1–12, 2014. [2](#)
- [32] Tyler Maunu and Gilad Lerman. Robust subspace recovery with adversarial outliers, 2019. [1](#), [2](#), [4](#), [5](#)

- [33] Tyler Maunu, Teng Zhang, and Gilad Lerman. A well-tempered landscape for non-convex robust subspace recovery. *J. Mach. Learn. Res.*, 20(1):1348–1406, 2019. [1](#), [3](#), [4](#), [5](#), [6](#)
- [34] Tyler Maunu, Chenyu Yu, and Gilad Lerman. Stochastic and private nonconvex outlier-robust PCAs. In *Proceedings of Mathematical and Scientific Machine Learning*, pages 173–188. PMLR, 2022. [1](#)
- [35] Michael McCoy and Joel A. Tropp. Two proposals for robust PCA using semidefinite programming. *Electronic Journal of Statistics*, 5(none):1123 – 1160, 2011. [1](#)
- [36] Onur Ozyesil and Amit Singer. Robust camera location estimation by convex programming. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2674–2683, 2015. [7](#), [8](#)
- [37] Rahul Raguram, Ondrej Chum, Marc Pollefeys, Jiri Matas, and Jan-Michael Frahm. USAC: A universal framework for random sample consensus. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):2022–2038, 2012. [2](#), [8](#)
- [38] Elvezio M. Ronchetti and Peter J. Huber. *Robust statistics*. John Wiley & Sons Hoboken, NJ, USA, 2009. [1](#)
- [39] Soumyadip Sengupta, Tal Amir, Meirav Galun, Tom Goldstein, David W Jacobs, Amit Singer, and Ronen Basri. A new rank constraint on multi-view fundamental matrices, and its application to camera location recovery. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4798–4806, 2017. [7](#), [8](#), [16](#)
- [40] Yunpeng Shi, Shaohan Li, Tyler Maunu, and Gilad Lerman. Scalable cluster-consistency statistics for robust multi-object matching. In *International Conference on 3D Vision, 3DV 2021, London, United Kingdom, December 1-3, 2021*, pages 352–360. IEEE, 2021. [7](#)
- [41] Noah Snavely, Steven M. Seitz, and Richard Szeliski. Photo tourism: exploring photo collections in 3d. In *ACM siggraph 2006 papers*, pages 835–846. Association for Computing Machinery, 2006. [6](#), [7](#), [8](#)
- [42] Nathan Srebro and Tommi Jaakkola. Weighted low-rank approximations. In *Proceedings of the 20th international conference on machine learning (ICML-03)*, pages 720–727, 2003. [1](#)
- [43] Jacob Steinhardt, Moses Charikar, and Gregory Valiant. Resilience: A criterion for learning in the presence of arbitrary outliers. In *9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11-14, 2018, Cambridge, MA, USA*, pages 45:1–45:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. [1](#)
- [44] Ben Tordoff and David W. Murray. Guided sampling and consensus for motion estimation. In *Computer Vision—ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part I 7*, pages 82–96. Springer, 2002. [15](#)
- [45] Philip H.S. Torr and Andrew Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *Computer vision and image understanding*, 78(1):138–156, 2000. [2](#)
- [46] David E. Tyler. Statistical analysis for the angular central gaussian distribution on the sphere. *Biometrika*, 74(3):579–589, 1987. [2](#)
- [47] David E. Tyler. A distribution-free m-estimator of multivariate scatter. *The Annals of Statistics*, pages 234–251, 1987. [1](#), [2](#), [6](#), [13](#)
- [48] Kyle Wilson and Noah Snavely. Robust global translations with ldsfm. In *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part III 13*, pages 61–75. Springer, 2014. [7](#)
- [49] Huan Xu, Constantine Caramanis, and Sujay Sanghavi. Robust PCA via outlier pursuit. *Advances in neural information processing systems*, 23, 2010. [1](#)
- [50] Teng Zhang. Robust subspace recovery by Tyler’s M-estimator. *Information and Inference: A Journal of the IMA*, 5(1):1–21, 2016. [1](#), [3](#), [4](#), [5](#), [6](#), [13](#)
- [51] Teng Zhang and Gilad Lerman. A novel M-estimator for robust PCA. *The Journal of Machine Learning Research*, 15(1): 749–808, 2014. [1](#), [3](#), [5](#)