Adaptive Annealing for Robust Geometric Estimation

Chitturi Sidhartha, Lalit Manam and Venu Madhav Govindu
Indian Institute of Science, Bengaluru, India-560012
{chitturis,lalitmanam,venug}@iisc.ac.in

Abstract

Geometric estimation problems in vision are often solved via minimization of statistical loss functions which account for the presence of outliers in the observations. The corresponding energy landscape often has many local minima. Many approaches attempt to avoid local minima by annealing the scale parameter of loss functions using methods such as graduated non-convexity (GNC). However, little attention has been paid to the annealing schedule, which is often carried out in a fixed manner, resulting in a poor speed-accuracy trade-off and unreliable convergence to the global minimum. In this paper, we propose a principled approach for adaptively annealing the scale for GNC by tracking the positive-definiteness (i.e. local convexity) of the Hessian of the cost function. We illustrate our approach using the classic problem of registering 3D correspondences in the presence of noise and outliers. We also develop approximations to the Hessian that significantly speeds up our method. The effectiveness of our approach is validated by comparing its performance with state-of-the-art 3D registration approaches on a number of synthetic and real datasets. Our approach is accurate and efficient and converges to the global solution more reliably than the state-of-the-art methods.

1. Introduction

Geometric estimation problems in computer vision use point locations or correspondences extracted from images or scans, e.g. epipolar geometry, bundle adjustment, 3D registration etc. Such point observations are noisy and also have a number of corrupting outliers. While RANSAC and related approaches [11,22,40,53] work by removing observations classified as outliers, an important class of methods use M-estimation [2,52,56,57], which reduces the influence of outliers. Denoting the geometric model parameters as $\mathbf{x}$ and the residual of the model fitting of the i-th observation as $\mathbf{r}_i(\mathbf{x})$, the robust estimation problem is

$$\min_{\mathbf{x}} \sum_{i=1}^{N} \rho_\sigma(\|\mathbf{r}_i(\mathbf{x})\|) \tag{1}$$

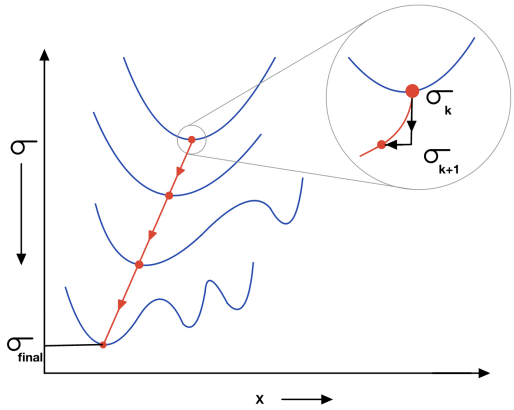


Figure 1. Annealing of scale σ in GNC: We start with a large value of σ that makes the problem easy to solve. The red curve indicates the path of the minimum solution for robust estimation as σ is gradually reduced. We desire an annealing scheme that will ensure that we reach the final global minimum for σ_{final}.

where $\rho(.)$ is a statistical loss function and σ is a scale parameter. σ signifies the scale of the noise in the observations and defines the distinction between what we classify as inliers and outliers.

Eqn. 1 is a non-linear optimization problem that is usually solved using a series of local optimization steps. A common choice is the Iteratively Reweighted Least Squares (IRLS) approach [28,31]. In IRLS, we take the gradient of the cost function and equate it to zero, resulting in

$$\boldsymbol{\nabla}\rho_\sigma(\mathbf{x}) = \sum_{i=1}^{N} \rho'(\|\mathbf{r}_i\|)\boldsymbol{\nabla}(\|\mathbf{r}_i\|) = 0 \tag{2}$$

$$\Rightarrow \sum_{i=1}^{N} \frac{\rho'(\|\mathbf{r}_i\|)}{\|\mathbf{r}_i\|}\boldsymbol{\nabla}\left(\frac{\|\mathbf{r}_i\|^2}{2}\right) = 0 \tag{3}$$

$$\Rightarrow \sum_{i=1}^{N} \phi(\|\mathbf{r}_i\|)\boldsymbol{\nabla}\left(\frac{\|\mathbf{r}_i\|^2}{2}\right) = 0 \tag{4}$$

In a given iteration, we treat the weights $\phi_i = \phi(\|\mathbf{r}_i\|)$ as constants, which gives us a weighted least squares problem. Given a new estimate of $\mathbf{x}$, these weights are updated, and the process is repeated until convergence.

Graduated Non-Convexity: Since the cost function in Eqn. 1 can have a large number of local minima, a careful initialization of IRLS is necessary to reach the global minimum that we desire. To this end, a popular approach developed in the computer vision literature is Graduated Non-Convexity (GNC) [10, 47, 66, 71, 72]. In GNC, we start with a large value, σ_0, that smoothens the underlying cost function making it easier to optimize. When $\sigma_0 \to \infty$, $\rho_{\sigma_0}(\|\mathbf{r}\|) \to \|\mathbf{r}\|^2/2$ leading to a least squares problem. The initial cost is, thus, very smooth and is either convex or has a large basin of convexity around the global minimum, thereby enabling practical optimization algorithms to reliably find the global minimum. In many cases, there exist certifiable global solvers [12, 74], and in the special case of 3D registration, there exists a closed-form solution for the global minimum [55] of the initial cost. After obtaining the global minimum for the initial smoothened cost, we proceed by progressively *annealing* or reducing σ and re-solving Eqn. 1. Given an estimate $\mathbf{x}_k$ at scale σ_k at the k-th stage, we update scale to σ_{k+1} and solve Eqn. 1 using $\mathbf{x}_k$ as the initialization, yielding the updated estimate $\mathbf{x}_{k+1}$. This is repeated till we terminate our procedure at a desired final scale σ_{final}.

If we incrementally update σ from σ_k to σ_{k+1}, we can ensure that the resulting local minimum of the new cost function, obtained using the previous solution as the initialization, is a very good local minimum if not the global minimum of the new cost. When the data is corrupted with more than 50% outliers in an adversarial manner, any robust method would give a model that is biased towards the majority outlier observations. This is because the global minimum of the robust cost is not close to the inlier model parameters. Thus, it is reasonable for any GNC strategy to focus on obtaining the global minimum of the robust cost on data corrupted with $< 50\%$ outliers.

Fig. 1 illustrates a family of cost functions parametrized by σ used in GNC. Starting from a very large σ which gives us a smooth, easily optimizable cost function in Eqn. 1, we can move $\mathbf{x}$ along the smooth red curve which is the locus of the minimum of the cost function as we drive σ down to its final value σ_{final}. The GNC procedure seeks to ensure that at every stage $(k+1)$, the initialization $\mathbf{x}_k$ falls within the basin of convergence for the global minimum of the current cost function for σ_{k+1}. This, in turn, increases the likelihood that the final estimate of $\mathbf{x}$ at σ_{final} is the global minimum, i.e. we try to avoid falling into the basin of a local minimum. We note that as we anneal σ, the cost function starts becoming non-convex with multiple local minima in the cost function landscape. However, if throughout we remain in the same basin of convergence as we modify σ, we can reach the global minimum more

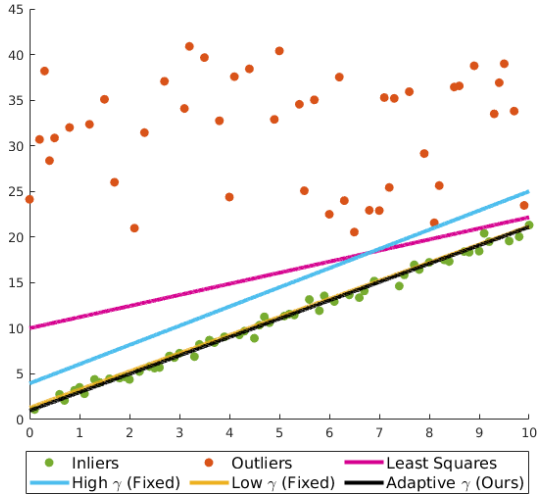


Figure 2. Adaptive annealing for GNC for line fitting in the presence of outliers. See text for details.

reliably for the desired cost function specified by σ_{final}, in most of the cases with outliers $< 50\%$. Even in cases where the outlier corruption is $\geq 50\%$, the real-world data is not always adversarial, thereby enabling GNC to perform well. It will be noted that this is akin to the approach of *homotopy continuation* [24, 37, 38, 41] in optimization.

Annealing Scheme: The inset in Fig. 1 further illustrates the GNC procedure. At the k-th stage, we are at the minimum of Eqn. 1 for σ_k. We keep the current $\mathbf{x}_k$ and update σ to σ_{k+1}, as indicated by the downward vertical arrow. Given this scale σ_{k+1}, we solve Eqn. 1, which moves the current estimate of the model parameters to $\mathbf{x}_{k+1}$ as indicated by the leftward horizontal arrow. This procedure is repeated till we reach the desired scale σ_{final}. The performance of this scheme critically depends on the annealing steps used. While GNC is used in many geometric estimation approaches [66, 68, 75], the majority of them use a simple annealing scheme, wherein, at each stage, the scale σ is decremented by a *fixed* factor, i.e. $\sigma_{k+1} = \frac{\sigma_k}{\gamma}$ where $\gamma > 1$. It will be immediately noted that the choice of γ has a significant impact on the performance of GNC. If we take γ close to 1, say $\gamma = 1 + \epsilon$ for very small $\epsilon > 0$, we move slowly in the cost function landscape. This will ensure that we stay on the red curve leading to the global minimum for σ_{final}. However, this comes at the cost of using a large number of annealing stages, thereby making the computational cost prohibitive. In contrast, if we take a large γ, we move rapidly in the scale space σ but can possibly end up in a poor quality local minimum.

In this paper, we propose a principled approach to *adaptively* anneal the scale σ. At each stage, we seek to use

GNC strategy	# of stages in GNC	Accuracy
Small γ (fixed)	$\uparrow$	$\uparrow$
Large γ (fixed)	$\downarrow$	$\downarrow$
Adaptive γ [Ours]	$\downarrow$	$\uparrow$
Desirable	$\downarrow$	$\uparrow$

Table 1. Impact of different annealing strategies. Our adaptive approach achieves high accuracy with fewer annealing stages.

as large a γ as possible while ensuring that we always stay within the desired basin of convergence of the global minimum. While our approach is general, for simplicity, throughout this paper, we use the Geman-McClure loss function:

$$\rho_\sigma(e) = \frac{e^2}{2(1+\frac{e^2}{\sigma^2})} \tag{5}$$

In Fig. 2, we illustrate the impact of our adaptive annealing scheme for a simple problem of 2D line fitting in the presence of outliers. When we set the annealing factor to a small value ($\gamma = 2$), GNC converges to the global minimum but needs 14 stages. For a large annealing factor ($\gamma = 10$), GNC terminates in 4 stages but converges to a poor local minimum. In contrast, our adaptive approach uses 8 stages and converges to the correct solution at the global minimum. We can summarize our observations as shown in Table. 1.

1.1. Adaptive Annealing

In this subsection, we describe our adaptive annealing scheme. From the above discussion, we have noted that at each stage of GNC, we seek to anneal σ while ensuring that we remain in the basin of convergence of the global minimum. We propose to achieve this objective by examining the Hessian of the cost function in Eqn. 1 with respect to the parameters $\mathbf{x}$. For an N_p parameter model, the $N_p \times N_p$ Hessian matrix is

$$\begin{aligned} [\mathbf{H}_i]_{(r,s)} &= \frac{\partial^2 \rho_\sigma(\|\mathbf{r}_i(\mathbf{x})\|)}{\partial x_r \partial x_s}\Big|_{\mathbf{x}_k} \\ \Rightarrow \mathbf{H} &= \sum_{i=1}^{N} \mathbf{H}_i \end{aligned} \tag{6}$$

where i is the index for the individual observations in Eqn. 1. Thus, $\mathbf{H}$ is the Hessian of the robust cost function evaluated at the k-th stage estimate $\mathbf{x}_k$. Since $\mathbf{x}_k$ is a minimum of the cost function in Eqn. 1 evaluated for σ_k, we note that $\mathbf{H}$ is *locally convex*, i.e. positive-definite. This is true since we have converged to $\mathbf{x}_k$ through an optimization of Eqn. 1.[1] Additionally, in the $(k+1)$-th iteration, if we update the scale to σ_{k+1} ensuring that the Hessian $\mathbf{H}$ evaluated at σ_{k+1} remains positive-definite, then we have likely ensured that the new estimate, $\mathbf{x}_{k+1}$, is in the same basin of convergence as that of the previous iteration. This is true since we initialize our optimization for the $(k+1)$-th iteration at $\mathbf{x}_k$ and converge to the local minimum using IRLS. We finally note that we begin the adaptive GNC procedure for a very large σ, thereby ensuring that the cost function is very smooth and convex over a larger domain. This makes it easier to obtain the global minimum of the smoothened cost function, which serves as the first estimate $\mathbf{x}_0$. To summarize, if we iteratively choose σ_k and solve for $\mathbf{x}_k$, such that at each stage, the corresponding Hessian $\mathbf{H}$, in Eqn. 6, remains positive-definite, then the resulting solution, $\mathbf{x}_{final}$, is very likely to be the global minimum of the cost at σ_{final}.

In the $(k+1)$-th iteration, in the interest of a faster estimate, we seek to reduce σ_{k+1} as much as possible *while* maintaining local convexity of $\mathbf{H}$. Since a positive-definite $\mathbf{H}$ has all positive eigen values, we achieve our objective by tracking the sign of the smallest eigen value λ_{min} of $\mathbf{H}$. Therefore, our criteria for maintaining local convexity translates to one of seeking the smallest σ_{k+1} (i.e. the largest annealing) while maintaining $\lambda_{min}(\mathbf{H}) > 0$. While a $\lambda_{min}(\mathbf{H})$ close to zero can make $\mathbf{H}$ highly ill-conditioned, we note that we never use $\mathbf{H}$ in our estimation process and only use the criteria of $\lambda_{min}(\mathbf{H}) > 0$ to determine σ_{k+1}. We can also choose to terminate our search when $\lambda_{min}(\mathbf{H})$ is close to a threshold $\lambda_T > 0$, which will ensure that $\mathbf{H}$ is never ill-conditioned and the optimization of Eqn. 1 using σ_{k+1} is a well-defined problem.

Binary Search: In general, we do not have closed-form expressions for $\lambda_{min}(\mathbf{H})$. Consequently, we estimate σ_{k+1} using a divide-and-conquer approach using the criteria that $\lambda_{min}(\mathbf{H}) > 0$. We define a search interval below σ_k and perform a binary search. We note that the cost of evaluating $\lambda_{min}(\mathbf{H})$ is cheap since $\mathbf{H}$ is a small $N_p \times N_p$ matrix. In the problem of interest in this paper, *i.e.* 3D registration, $N_p = 6$. We note here that while the cost in Eqn. 1 is non-linear, it is smooth and differentiable. Implicit to our binary search strategy is the assumption that λ_{min} is monotonically decreasing as we reduce σ in the binary search interval. We observe that, in practice, this assumption is true in most cases. We can also ensure the reliability of this assumption by making the search interval smaller, although that would increase the overall computational time since now we will need more stages of GNC to reach the desired scale σ_{final}.

Our approach of adaptive annealing of scale σ to solve Eqn. 1 is general in nature. In the remainder of this paper,

[1] We emphasize that the positive-definiteness of $\mathbf{H}$ indicates local convexity of our cost function and does not imply that our cost function is globally convex.

we illustrate our approach with a concrete example of solving for robust 3D registration given point correspondences that are noisy and include a large number of outliers. Before we develop our method for 3D registration, we briefly discuss some of the relevant literature.

2. Related Work on 3D Registration

There are primarily two types of methods for pairwise registration of 3D point clouds. One direction of work considers the problem of finding correspondences between the point clouds and the registration of the correspondences as a joint problem and solves it using alternating, iterative [7, 14, 17, 32, 35, 39, 44, 48, 69, 70] methods. Another direction that is of relevance to our approach deal with the registration of point clouds given putative correspondences.

Robust Methods: Traditionally, Random Sampling Consensus (RANSAC) [11, 22, 26, 40, 53] is used to embed robustness into the registration problem. A few methods [13, 45] perform preprocessing before RANSAC to remove outliers using either deterministic geometric methods [13] or stochastic game theoretical methods [1]. A class of methods which are based on consensus maximization [3, 18, 19, 25, 33, 43, 54, 63] and Branch-and-Bound (BNB) techniques [6, 15, 29] perform robust *rotation search* (also known as the *Wahba problem*). However, all the aforementioned methods become intractable and inaccurate in the presence of high outlier rates and/or large number of correspondences.

M-estimation: Another class of methods pose the robust registration problem as minimization of a robust cost (*e.g.* [36]). [8] minimizes a $\ell_{\frac{1}{2}}$ robust loss using Lie-algebraic optimization on the $\mathbb{SE}(3)$ manifold. [75] uses Graduated Non-Convexity (GNC) while exploiting the equivalence of robust cost and line processes [9]. [68] proposes a certifiably global optimal algorithm for solving the registration problem along with the scale by filtering outliers using a graph theoretical framework [49] and using certifiably robust rotation estimation [67].

Deep Learning Methods: Many recent deep learning methods [4, 34, 60, 61] learn the end-to-end problem of finding learnable features [46, 62], matching to get putative correspondences, and registering them using differentiable registration modules. [20] obtains putative correspondences from FCGF [21] features and uses different modules for correspondence weighting, differentiable registration and refinement. [30] uses an overlap-attention block to register point clouds with low overlap. Some methods [5, 16, 34] use spatial consistency measure to eliminate outliers.

3. Proposed Method for 3D Registration

The optimization in Eqn. 1 can be specifically written for the problem of robust 3D registration as

$$\min_{(\mathbf{R},\mathbf{t})} \sum_{i=1}^{N} \rho_\sigma(\|\mathbf{a}_i - \mathbf{R}\mathbf{b}_i - \mathbf{t}\|) \tag{7}$$

where $\{\mathbf{a}_i\}$'s and $\{\mathbf{b}_i\}$'s are sets of 3D correspondences (with noise and outliers) that need to be registered. Further, the parameters $\mathbf{x}$, in this case, are $\mathbf{R} \in \mathbb{SO}(3)$ and $\mathbf{t} \in \mathbb{R}^3$, which are the rotation and translation, respectively, that align the two point clouds $\mathbf{A}$ and $\mathbf{B}$. In the context of our approach of using GNC for optimization, we note that for $\sigma \to \infty$, we have the well-known least squares 3D registration problem *i.e.*

$$\min_{(\mathbf{R},\mathbf{t})} \frac{1}{2} \sum_{i=1}^{N} \|\mathbf{a}_i - \mathbf{R}\mathbf{b}_i - \mathbf{t}\|^2 \tag{8}$$

for which the global minimum can be found in a closed form (Umeyama's method [55]). For other finite values of σ, the resulting cost function can be solved using IRLS. For each IRLS iteration, fixing the weights ϕ_i results in a weighted least squares problem as follows:

$$\min_{(\mathbf{R},\mathbf{t})} \quad \frac{1}{2} \sum_{i=1}^{N} \phi_i \|\mathbf{r}_i\|^2 = \frac{1}{2} \sum_{i=1}^{N} \phi_i \|\mathbf{a}_i - \mathbf{R}\mathbf{b}_i - \mathbf{t}\|^2 \tag{9}$$

which can also be solved in a closed-form, that we refer to as the 'weighted Umeyama' solution. The weighted Umeyama algorithm is a variant of Umeyama's method [55], and the pseudo-code is provided in the supplementary material. We summarize our adaptive annealing GNC procedure for robust 3D registration in Algorithm 1.

Hessian of the Umeyama Costs: Unlike the general case, the Hessian for the 3D registration cost functions in Eqn. 7 and Eqn. 8 have a closed form solution that can be stated as follows:

Theorem 3.1 *The gradient $\mathbf{g}_{LSQ}$ and Hessian $\mathbf{H}_{LSQ}$ of the least squares cost (Eqn. 8), at the point $(\mathbf{R}, \mathbf{t})$, are given by:*

$$\mathbf{g}_{LSQ} = \sum_{i=1}^{N} \mathbf{g}_{LSQ,i}; \quad \mathbf{g}_{LSQ,i} = \begin{bmatrix} -[\mathbf{b}_i]_\times \mathbf{R}^\top \mathbf{r}_i \\ -\mathbf{r}_i \end{bmatrix} \tag{10}$$

$$\mathbf{H}_{LSQ} = \sum_{i=1}^{N} \mathbf{H}_{LSQ,i} \tag{11}$$

$$\mathbf{H}_{LSQ,i} = \begin{bmatrix} (\mathbf{p}_i^\top \mathbf{R}\mathbf{b}_i)\mathbf{I} - \frac{\mathbf{b}_i \mathbf{p}_i^\top \mathbf{R}}{2} - \frac{\mathbf{R}^\top \mathbf{p}_i \mathbf{b}_i^\top}{2} & [\mathbf{b}_i]_\times \mathbf{R}^\top \\ -\mathbf{R}[\mathbf{b}_i]_\times & \mathbf{I} \end{bmatrix} \tag{12}$$

where $\mathbf{p}_i = \mathbf{a}_i - \mathbf{t}$ and $\mathbf{I}$ is the 3×3 identity matrix.

Algorithm 1: Robust 3D Registration using Adaptively Annealed GNC (GNCp)

Input: N corresponding points in the two point clouds $\{\mathbf{a}_i\}$'s, $\{\mathbf{b}_i\}$'s; σ_{final}
Output: Rotation $\mathbf{R}$, Translation $\mathbf{t}$
Initialization: $k = 0$, $\sigma = \sigma_0$
(No initialization required for $\mathbf{R}$,$\mathbf{t}$).
while $\sigma_k \geq \sigma_{final}$ **do**
  $\mathbf{r}_i = \mathbf{a}_i - \mathbf{R}\mathbf{b}_i - \mathbf{t}$
  $\phi_i = \frac{1}{\left(1+\frac{\|\mathbf{r}_i\|^2}{\sigma_k^2}\right)^2}$
  `/* Solve for R and t using weighted Umeyama method */`
  $\mathbf{R}, \mathbf{t}$ = *weightedUmeyama*($\{\mathbf{a}_i\}$,$\{\mathbf{b}_i\}$, $\{\phi_i\}$)
  Obtain $\mathbf{H}(\sigma)$ using Eqn. 6
  Perform *binary search* on λ_{min} $(\mathbf{H}_{approx})$ to obtain σ_{k+1}
  $k = k + 1$
end

Theorem 3.2 *The Hessian* $\mathbf{H}$ *of the robust Umeyama cost (Eqn. 7), at the point* $(\mathbf{R}, \mathbf{t})$*, is given by:*

$$\mathbf{H} = \sum_{i=1}^{N} \left(-l_i \frac{\mathbf{g}_{LSQ,i}\mathbf{g}_{LSQ,i}^\top}{\|\mathbf{r}_i\|^2} + m_i \mathbf{H}_{LSQ,i} \right), \quad (13)$$

$$\textit{where } l_i = \frac{\rho'(\|\mathbf{r}_i\|)}{\|\mathbf{r}_i\|} - \rho''(\|\mathbf{r}_i\|), m_i = \frac{\rho'(\|\mathbf{r}_i\|)}{\|\mathbf{r}_i\|} \quad (14)$$

For the Geman-McClure loss, we have,

$$l_i = \frac{4\|\mathbf{r}_i\|^2}{\sigma^2\left(1+\frac{\|\mathbf{r}_i\|^2}{\sigma^2}\right)^3}, \quad m_i = \frac{1}{\left(1+\frac{\|\mathbf{r}_i\|^2}{\sigma^2}\right)^2} \quad (15)$$

and the Hessian is

$$\mathbf{H} = \sum_{i=1}^{N} \frac{-4\mathbf{g}_{LSQ,i}\mathbf{g}_{LSQ,i}^\top}{\sigma^2\left(1+\frac{\|\mathbf{r}_i\|^2}{\sigma^2}\right)^3} + \frac{1}{\left(1+\frac{\|\mathbf{r}_i\|^2}{\sigma^2}\right)^2}\mathbf{H}_{LSQ,i} \quad (16)$$

The proofs for Theorems 3.1 and 3.2 are provided in the supplementary material.

Piecewise Polynomial Approximation of l_i and m_i: Given the large number of points (N) summed over in Eqn. 16, computing $\mathbf{H}$ in Eqn. 16 for different values of σ can be expensive during binary search. However, we note that the expressions for l_i and m_i, in Eqn. 15, can be approximated by:

$$l_i = \frac{\alpha_i}{\hat{\sigma}_i^2} + \kappa_i, \quad m_i = \frac{\zeta_i}{\hat{\sigma}_i^2} + \chi_i \quad (17)$$

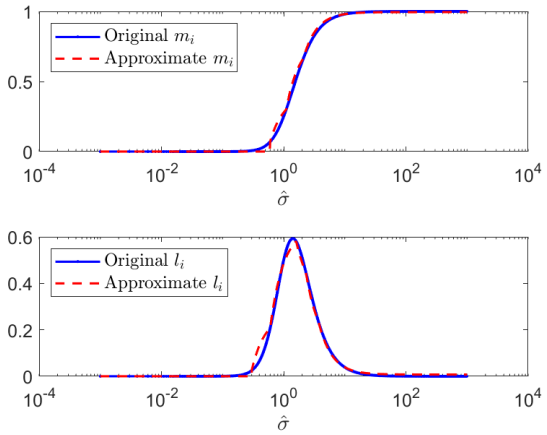


Figure 3. Original and piecewise approximations of m_i and l_i.

where $\hat{\sigma}_i = \sigma/\|\mathbf{r}_i\|$ and the terms $\{\alpha_i, \kappa_i, \zeta_i, \chi_i\}$ are piecewise constant functions over the search interval for σ. In our implementation, we choose 6 and 4 piecewise constant intervals for l_i and m_i, respectively, and search for the piecewise constant terms that best approximate l_i and m_i over σ. We illustrate our approximations for the expressions l_i and m_i in Fig. 3. Using this approximation, we can write the approximated Hessian as

$$\mathbf{H}_{approx} = \frac{\mathbf{C}}{\sigma^2} + \mathbf{D} \quad (18)$$

where $\mathbf{C}$ and $\mathbf{D}$ are matrices that are *constant* in the piecewise constant intervals of l_i and m_i. As a result, during the binary search, we need to evaluate the matrices $\mathbf{C}$ and $\mathbf{D}$ only a few times, which drastically reduces the computational cost of our overall scheme.

To further increase the speed, after the first 5 iterations, we perform a stochastic approximation to the Hessian ($\mathbf{H}$) at every annealing stage. We sample a subset of observations with probability proportional to the weights of the observations. This is done because the contribution of the lesser weighted observations to the Hessian is small. Also, we reiterate the fact that the Hessian is used only for finding the next σ and is not used in the optimization step. We observe that this improves efficiency without any negative impact on the accuracy of the solution.

4. Results

In this section, we evaluate the performance of our principled GNC-based method (GNCp) with the state-of-the-art methods on synthetic and real-world datasets. We compare our method with Fast Global Registration (FGR) [75][2],

[2] https://github.com/isl-org/FastGlobalRegistration

which uses fixed annealing scheme, GORE [13][3], which is a RANSAC based method, SE3Reg [8], which uses $\ell_{\frac{1}{2}}$ loss for robustness, and TEASER++ [68][4], which is a GNC based method that uses maximum clique detection to prune outliers. The source code made available by the authors are mentioned in their respective footnotes. The implementation of SE3Reg was provided by the authors. Our method is implemented in C++. For our method, we fix $\sigma_{final} = 0.1$ unless stated otherwise. All experiments are performed on a PC with Intel Xeon Silver 4210 processor with 256 GB RAM. Finally, MRE and MTE denote Mean Rotation Errors and Mean Translation Errors, respectively.

4.1. Synthetic Data

We first carry out experiments on synthetic data to validate our proposed method. We use datasets from the Stanford 3-D Scanning repository [23] and create the synthetic datasets as suggested in TEASER++ [68]. The first point cloud $\mathbf{B}$ (containing $\{\mathbf{b}_i\}$'s) is obtained by downsampling the original point cloud to N points and resized to fit inside the $[0, 1]^3$ cube. Then, we apply a random transformation $(\mathbf{R}, \mathbf{t})$ on the point cloud $\mathbf{B}$ and then add a random bounded noise $\boldsymbol{\epsilon_i} \sim \mathcal{N}(\mathbf{0}, \eta^2\mathbf{I})$ ($\|\boldsymbol{\epsilon}_i\|^2 \leq \beta_i$ with $\beta_i = 5.54\eta$, as chosen in [68]) to obtain the second point cloud $\mathbf{A}$ (containing $\{\mathbf{a}_i\}$'s), *i.e.* $\mathbf{a}_i = \mathbf{R}\mathbf{b}_i + \mathbf{t} + \boldsymbol{\epsilon}_i$, where $\mathbf{R} \in \mathbb{SO}(3)$ and $0 \leq \|\mathbf{t}\| \leq 1$. The outlier correspondences are obtained by randomly sampling a percentage p_{outl} of the points $\mathbf{a}_i$'s and replacing them with vectors uniformly sampled in the sphere of radius 5 units. We create two types of datasets by choosing (N, η) as: (i) $(100, 0.01)$ *i.e.* small size with low noise level, called as *Type-1*, and (ii) $(10000, 0.1)$ *i.e.* large size with high noise level, called as *Type-2*. *Type-2* datasets are harder to solve than *Type-1* datasets. We execute 40 randomized trials for each dataset type.

From Tables 2 and 3, it can be seen that the rotation and translation errors are the least when using our method for both *Type-1* and *Type-2* datasets. Specifically, there is a significant improvement in error metrics for *Type-2* datasets compared to other methods. This reveals that our method is able to handle outliers effectively in both small and large input sizes. Computation time for all methods is reported in the supplementary material, where we observe that our method scales very well with large input size (*Type-2*). We note that GORE has a very high computation time on *Type-2* datasets due to a large number of correspondences and thus not reported in Table 3.

We also perform experiments on the synthetic CAD model dataset - ModelNet [64]. We use the features obtained from Predator [30] and match them using nearest neighbour matching, which gives 1266 point cloud pairs. The obtained correspondences have a lesser noise level. Hence, we choose $\sigma_{final} = 0.01$ for our method. We use the criteria of successful registration, defined in [68], as the rotation error being less than 10° and the translation error being less than 30 cm with respect to the ground truth. From Table 4, it can be seen that our method has the least rotation and translation errors compared to other methods. Also, our method has a similar success percentage compared to the highest one (TEASER++). GORE takes the least amount of time due to the small input size ($N \approx 100$ to 300) of the ModelNet dataset. SE3Reg follows in terms of speed, while other methods including ours take similar amounts of time.

[3] https://cs.adelaide.edu.au/ aparra/project/gore/
[4] https://github.com/MIT-SPARK/TEASER-plusplus

Dataset	FGR [75]	GORE [13]	SE3Reg [8]	TEASER++ [68]	GNCp (Ours)
	Mean Rotation Errors (deg) ↓				
armadillo	1.04	3.36	0.88	0.77	**0.70**
bunny	0.87	3.42	0.72	0.66	**0.59**
buddha	1.34	4.21	1.05	0.95	**0.88**
dragon	0.87	3.50	0.71	0.65	**0.58**
	Mean Translation Errors ($\times 10^{-3}$) ↓				
armadillo	11.31	26.52	7.58	6.38	**6.23**
bunny	10.06	32.54	7.18	6.20	**5.90**
buddha	10.81	24.88	6.79	**5.74**	5.80
dragon	10.18	30.44	6.73	5.84	**5.56**

Table 2. Rotation and translation mean errors (over all instances) on *Type-1* synthetic datasets for $p_{outl} = 50\%$. This dataset is small with a low noise level and thus easy to solve.

Dataset	FGR [75]	SE3Reg [8]	TEASER++ [68]	GNCp (Ours)
	Mean Rotation Errors (deg) ↓			
armadillo	0.89	1.12	13.87	**0.79**
bunny	0.93	1.27	13.63	**0.81**
buddha	1.22	1.32	22.32	**1.02**
dragon	0.88	1.05	13.05	**0.76**
	Mean Translation Errors ($\times 10^{-3}$) ↓			
armadillo	16.68	48.06	106.59	**9.27**
bunny	14.53	46.76	152.33	**9.18**
buddha	18.97	45.57	148.82	**11.09**
dragon	15.89	47.21	119.91	**9.08**

Table 3. Rotation and translation mean errors (over all instances) on *Type-2* synthetic datasets for $p_{outl} = 50\%$. This dataset is large with a high noise level and thus hard to solve.

Dataset	Success % ↑	MRE ↓	MTE ($\times 10^{-2}$) ↓	Time (ms)
FGR [75]	93.1	2.92	2.96	10
GORE [13]	93.3	2.87	2.86	**2**
SE3Reg [8]	93.0	2.60	2.59	3
TEASER++ [68]	**94.2**	2.40	2.28	7
GNCp (Ours)	94.1	**2.26**	**2.26**	8

Table 4. Evaluation on ModelNet dataset [64]. Our method has the least mean errors.

Dataset	Success % ↑				Mean Rotation Errors (deg) ↓				Mean Translation Errors (m) ↓			
	FGR [75]	SE3Reg [8]	TEASER++ [68]	GNCp (Ours)	FGR [75]	SE3Reg [8]	TEASER++ [68]	GNCp (Ours)	FGR [75]	SE3Reg [8]	TEASER++ [68]	GNCp (Ours)
MIT lab	72.7	75.3	71.4	**77.9**	13.46	12.48	14.64	**9.18**	0.42	0.44	0.63	**0.33**
home1	93.6	92.9	92.9	**96.1**	5.91	6.21	8.74	**4.21**	0.19	0.19	0.29	**0.15**
home2	79.3	78.8	78.8	**81.7**	20.48	19.56	**16.24**	20.46	0.38	**0.32**	0.38	0.38
hotel1	93.8	93.8	94.7	**95.1**	6.95	7.13	7.14	**6.63**	**0.18**	**0.18**	**0.18**	**0.18**
hotel2	88.5	89.4	86.5	**91.3**	14.93	**14.48**	14.64	15.34	0.33	**0.32**	0.45	0.36
hotel3	85.2	87.0	85.2	**88.9**	23.38	20.98	**13.43**	20.81	0.46	0.40	**0.38**	0.40
kitchen	95.3	92.7	96.0	**96.6**	4.91	5.48	4.60	**4.29**	0.12	0.14	0.14	**0.11**
study	79.8	82.5	**86.0**	84.6	16.05	12.97	15.17	**10.88**	0.51	0.43	0.57	**0.35**

Table 5. Results on 3D Match dataset [73]. TEASER++ fails to converge in some instances even after 2 hours and the code fails to execute in some instances. Such instances are considered as failure cases for TEASER++, and the mean is computed for the remaining instances. GORE also fails to converge for data instances with a large set ($>= 1000$) of correspondences and hence is not reported.

4.2. Real Data

In this subsection, we present the results on real datasets. First, we consider the 3DMatch dataset [73], which is constructed based on the 7-Scenes dataset [50] and the SUN3D dataset [65]. The dataset comprises 62 real-world scenes, which is split into 8 scenes for testing and 54 scenes for training. We compute the FCGF features [21] and generate correspondences for the testing split using nearest-neighbor matching, as done in [75]. This leads to 1623 pairs of matched point clouds. The putative correspondences are directly given as inputs to the geometric registration algorithms. The outlier percentage in the resulting correspondences ranges from ≈ 0 to 99%. Therefore, there are bound to be failures in a few data instances.

As evident from the results in Table 5, our method achieves a higher percentage of successful registrations (criteria for successful registration chosen to be the same as in ModelNet) compared to the other methods. We observe that our method achieves superior or similar performance in terms of the rotation and translation errors for most of the scenes. In Table 6, we can see that our method is the fastest among all and also has the highest percentage of successful registrations. We note that TEASER++ fails to converge in some instances even after 2 hours of compute time, and the code fails to execute in a few other instances. We treat such instances as unsuccessful instances. The mean errors are computed over the remaining instances, thereby favouring TEASER++, due to which, the mean errors for TEASER++ are lower in a few sequences. In Fig. 4, we visually compare the alignment for an instance of the MIT Lab sequence in the 3DMatch dataset where the overlap is very low. It can be seen that our method performs the best and is close to the ground truth alignment.

Next, we evaluate the performance on KITTI odometry [27] dataset. This dataset contains 11 sequences of LiDAR-scanned outdoor driving scenarios. We extract

Dataset	FGR [75]	SE3Reg [8]	TEASER++ [68]	GNCp (Ours)
MIT lab	71.4	21.6	8216	**7.4**
home1	54.6	14.6	3964	**5.6**
home2	47.1	12.2	5555	**6.0**
hotel1	54.1	16	4416	**6.4**
hotel2	52.4	15.1	3969	**7.6**
hotel3	56.6	15.1	5849	**6.3**
kitchen	45.1	15.4	1978	**6.0**
study	55.2	16.2	3195	**8.9**

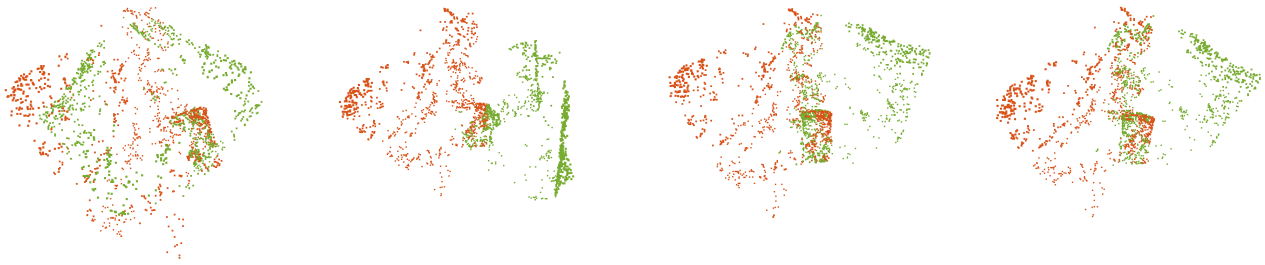
Table 6. Mean time taken (in ms) on 3DMatch dataset.

	Success % ↑	MRE ↓	MTE (m) ↓	Time (ms)
FGR [75]	73.5	0.94	0.38	195
GORE [13]	>= 30 minutes (per instance)			-
SE3Reg [8]	84.7	0.88	0.32	40
TEASER++ [68]	Code did not execute			-
GNCp (Ours)	**85.6**	**0.74**	**0.29**	**13**

Table 7. Evaluation on KITTI dataset [27]. Our method performs the best in all metrics.

FCGF features [21] on the test set (sequences 8-10), which has 555 pairs of matched point clouds. In Table 7, we compare performance on this dataset. It can be seen that our method has the highest success percentage ($< 5^\circ$ rotation errors and < 40 cm translation error) and has the least MRE and MTE while being the fastest method. The criteria for successful registration in this dataset is chosen based on the fact that motions have a very small rotation component and a large translation component. We note that TEASER++ fails to execute in most instances, and GORE takes longer than 30 mins for each instance, hence they are not reported here. We also perform experiments on the challenging low-overlap dataset - 3DLoMatch [30] in which the global minimum of the robust cost is close to the ground truth only in $\approx 60\%$ of the instances. Therefore, our method, which can only guarantee convergence to the global minimum of the robust cost, achieves 55.1% success percentage. More details can be found in the supplementary material.

Convergence to Global Minimum: In Table 8, we



(a) SE3Reg (b) TEASER++ (c) GNCp (Ours) (d) Ground Truth

Figure 4. Two point clouds (red and green) with a low overlap in the MIT Lab sequence of 3D Match dataset. Our method registers the point clouds correctly compared to the ground truth reference. FGR fails and is not included owing to space constraints. Refer to supplementary material for comparison of all methods.

analyze adaptive and fixed annealing schemes in terms of their relative success in reaching the global minimum of the cost function. We use different γ values for fixed annealing schemes on 3DMatch dataset ($\gamma = 1.4, 2.0$ for small and large, respectively). The global minimum is computed using 60 runs of a differential evolution algorithm with basin hoping [42, 51, 59], which is implemented using SciPy [58] in Python. We only consider the instances with $< 50\%$ outliers. We declare that the global minimum is attained if the difference between the solution and the global minimum is within 1° rotation and 1 cm translation. It can be seen that our adaptive annealing scheme reaches the global minimum most frequently with the least number of stages. We note that each stage in adaptive annealing involves computing the Hessian of the cost, which has some computational overheads.

Critical Failure Cases and their Remedy: Finally, we observe that tracking the local minimum using our strategy sometimes leads to a solution far away ($> 10^\circ$ rotation error) from the ground truth, especially in high outlier scenarios. These are observed in $\approx 12\%$, 0.36% and 0.08% of the instances in 3DMatch, KITTI and ModelNet datasets, respectively. To overcome these failure cases, at every stage of GNC, we perturb the current estimate by a large random rotation in order to jump out of a local minimum. Then, we apply local optimization at the perturbed point to get a new estimate. If the cost function at the new estimate reduces compared to the current estimate, we accept the new estimate and continue with our adaptive annealing strategy. We note that this modification does not influence the working of our strategy in any of the successful instances. Acceptance of the new estimate occurs at most once in the failure instances during the entire sequence of σ's, eventually reducing the failures to $\sim 9\%$ in 3DMatch and 0% in other datasets. We note that perturbation applied at the initial stages of GNC is beneficial due to the lesser number of local minima compared to the final stages of GNC, where the cost is highly non-convex. Therefore, in practice, we observe that the reduction in failure instances is consistent across multiple runs despite having stochasticity.

Dataset	Small γ (Fixed)		Large γ (Fixed)		GNCp (Ours)	
	$S\%$	#Stages	$S\%$	#Stages	$S\%$	#Stages
MIT lab	89.3	14	80.4	7	91.1	6.5
home1	94.8	14	84.3	7	100	6.1
home2	99.3	14	95.8	7	99.3	5.4
hotel1	97.9	14	94.7	7	100	5.8
hotel2	100	14	98.8	7	100	5.9
hotel3	95.7	14	87	7	100	6.7
kitchen	97.7	14	94	7	99.7	5.4
study	90.9	14	75.5	7	96.2	7.9

Table 8. Comparison of different annealing schemes for instances with $< 50\%$ outliers. $S\%$ refers to percentage of instances reaching global minimum.

5. Conclusion

We propose a principled approach for adaptively annealing the scale in Graduated Non-Convexity (GNC) by tracking the positive definiteness of the Hessian of the cost function. We apply the scheme to the problem of robust pairwise registration of 3D point clouds. This approach is shown to be efficient and accurate and achieves better results compared to state-of-the-art methods.

Acknowledgements: Chitturi Sidhartha and Lalit Manam are supported by the Prime Minister's Research Fellowship, Government of India. This research was supported in part by a Core Research Grant from Science and Engineering Research Board, Department of Science and Technology, Government of India.

References

[1] Andrea Albarelli, Emanuele Rodola, and Andrea Torsello. A game-theoretic approach to fine surface registration without initial motion estimation. In *2010 IEEE computer society conference on computer vision and pattern recognition*, pages 430–437. IEEE, 2010. 4

[2] Robert Andersen. *Modern methods for robust regression*. Number 152. Sage, 2008. 1

[3] Erik Ask, Olof Enqvist, and Fredrik Kahl. Optimal geometric fitting under the truncated l2-norm. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1722–1729, 2013. 4

[4] Armen Avetisyan, Manuel Dahnert, Angela Dai, Manolis Savva, Angel X Chang, and Matthias Nießner. Scan2cad: Learning cad model alignment in rgb-d scans. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, pages 2614–2623, 2019. 4

[5] Xuyang Bai, Zixin Luo, Lei Zhou, Hongkai Chen, Lei Li, Zeyu Hu, Hongbo Fu, and Chiew-Lan Tai. Pointdsc: Robust point cloud registration using deep spatial consistency. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 15859–15869, 2021. 4

[6] Jean-Charles Bazin, Yongduek Seo, and Marc Pollefeys. Globally optimal consensus set maximization through rotation search. In *Asian Conference on Computer Vision*, pages 539–551. Springer, 2012. 4

[7] Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In *Sensor fusion IV: control paradigms and data structures*, volume 1611, pages 586–606. Spie, 1992. 4

[8] Uttaran Bhattacharya and Venu Madhav Govindu. Efficient and robust registration on the 3d special euclidean group. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5885–5894, 2019. 4, 6, 7

[9] Michael J Black and Anand Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International journal of computer vision*, 19(1):57–91, 1996. 4

[10] Andrew Blake and Andrew Zisserman. *Visual reconstruction*. MIT press, 1987. 2

[11] Robert C Bolles and Martin A Fischler. A ransac-based approach to model fitting and its application to finding cylinders in range data. In *IJCAI*, volume 1981, pages 637–643, 1981. 1, 4

[12] Jesus Briales, Laurent Kneip, and Javier Gonzalez-Jimenez. A certifiably globally optimal solution to the non-minimal relative pose problem. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 145–154, 2018. 2

[13] Alvaro Parra Bustos and Tat-Jun Chin. Guaranteed outlier removal for point cloud registration with correspondences. *IEEE transactions on pattern analysis and machine intelligence*, 40(12):2868–2882, 2017. 4, 6, 7

[14] Dylan Campbell and Lars Petersson. Gogma: Globally-optimal gaussian mixture alignment. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5685–5694, 2016. 4

[15] Dylan Campbell, Lars Petersson, Laurent Kneip, and Hongdong Li. Globally-optimal inlier set maximisation for simultaneous camera pose and feature correspondence. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1–10, 2017. 4

[16] Zhi Chen, Kun Sun, Fan Yang, and Wenbing Tao. Sc2-pcr: A second order spatial compatibility for efficient and robust point cloud registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13221–13231, 2022. 4

[17] Dmitry Chetverikov, Dmitry Svirko, Dmitry Stepanov, and Pavel Krsek. The trimmed iterative closest point algorithm. In *2002 International Conference on Pattern Recognition*, volume 3, pages 545–548. IEEE, 2002. 4

[18] Tat-Jun Chin, Zhipeng Cai, and Frank Neumann. Robust fitting in computer vision: Easy or hard? *International Journal of Computer Vision*, 128, 03 2020. 4

[19] Tat-Jun Chin and David Suter. The maximum consensus problem: recent algorithmic advances. *Synthesis Lectures on Computer Vision*, 7(2):1–194, 2017. 4

[20] Christopher Choy, Wei Dong, and Vladlen Koltun. Deep global registration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2514–2523, 2020. 4

[21] Christopher Choy, Jaesik Park, and Vladlen Koltun. Fully convolutional geometric features. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8958–8966, 2019. 4, 7

[22] Ondřej Chum, Jiří Matas, and Josef Kittler. Locally optimized ransac. In *Joint Pattern Recognition Symposium*, pages 236–243. Springer, 2003. 1, 4

[23] Brian Curless and Marc Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 303–312, 1996. 6

[24] Daniel M Dunlavy and Dianne P O'Leary. Homotopy optimization methods for global optimization. *Report SAND2005-7495, Sandia National Laboratories*, 2005. 2

[25] Olof Enqvist, Erik Ask, Fredrik Kahl, and Kalle Åström. Robust fitting for multiple view geometry. In *European conference on computer vision*, pages 738–751. Springer, 2012. 4

[26] 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. 4

[27] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE conference on computer vision and pattern recognition*, pages 3354–3361. IEEE, 2012. 7

[28] Peter J Green. Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives. *Journal of the Royal Statistical Society: Series B (Methodological)*, 46(2):149–170, 1984. 1

[29] Richard I Hartley and Fredrik Kahl. Global optimization through rotation space search. *International Journal of Computer Vision*, 82(1):64–79, 2009. 4

[30] Shengyu Huang, Zan Gojcic, Mikhail Usvyatsov, Andreas Wieser, and Konrad Schindler. Predator: Registration of 3d point clouds with low overlap. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, pages 4267–4276, 2021. 4, 6, 7

[31] Peter J Huber. *Robust statistics*, volume 523. John Wiley & Sons, 2004. 1

[32] Bing Jian and Baba C Vemuri. Robust point set registration using gaussian mixture models. *IEEE transactions on pattern analysis and machine intelligence*, 33(8):1633–1645, 2010. 4

[33] Huu Le, Tat-Jun Chin, Anders Eriksson, Thanh-Toan Do, and David Suter. Deterministic approximate methods for maximum consensus robust fitting. *IEEE transactions on pattern analysis and machine intelligence*, 43(3):842–857, 2019. 4

[34] Junha Lee, Seungwook Kim, Minsu Cho, and Jaesik Park. Deep hough voting for robust global registration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 15994–16003, 2021. 4

[35] Yinlong Liu, Chen Wang, Zhijian Song, and Manning Wang. Efficient global point cloud registration by matching rotation invariant features through translation search. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 448–463, 2018. 4

[36] Kirk MacTavish and Timothy D Barfoot. At all costs: A comparison of robust cost functions for camera correspondence outliers. In *2015 12th conference on computer and robot vision*, pages 62–69. IEEE, 2015. 4

[37] Hossein Mobahi and John W Fisher. Coarse-to-fine minimization of some common nonconvexities. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 71–84. Springer, 2015. 2

[38] Hossein Mobahi and John Fisher III. A theoretical analysis of optimization by gaussian continuation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29, 2015. 2

[39] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence*, 32(12):2262–2275, 2010. 4

[40] David Nistér. Preemptive ransac for live structure and motion estimation. *Machine Vision and Applications*, 16(5):321–329, 2005. 1, 4

[41] Jorge Nocedal and Stephen J Wright. *Numerical optimization*. Springer, 1999. 2

[42] Brian Olson, Irina Hashmi, Kevin Molloy, and Amarda Shehu. Basin hopping as a general and versatile optimization framework for the characterization of biological macromolecules. *Advances in Artificial Intelligence (16877470)*, 2012. 8

[43] Carl Olsson, Olof Enqvist, and Fredrik Kahl. A polynomial-time bound for matching and registration with outliers. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2008. 4

[44] Alvaro Parra Bustos, Tat-Jun Chin, and David Suter. Fast rotation search with stereographic projections for 3d registration. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3930–3937, 2014. 4

[45] Alioscia Petrelli and Luigi Di Stefano. Pairwise registration by local orientation cues. In *Computer Graphics Forum*, volume 35, pages 59–72. Wiley Online Library, 2016. 4

[46] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 652–660, 2017. 4

[47] Anand Rangarajan and Rama Chellappa. Generalized graduated nonconvexity algorithm for maximum a posteriori image estimation. In *[1990] Proceedings. 10th International Conference on Pattern Recognition*, volume 2, pages 127–133. IEEE, 1990. 2

[48] Szymon Rusinkiewicz and Marc Levoy. Efficient variants of the icp algorithm. In *Proceedings third international conference on 3-D digital imaging and modeling*, pages 145–152. IEEE, 2001. 4

[49] Jingnan Shi, Heng Yang, and Luca Carlone. Robin: a graph-theoretic approach to reject outliers in robust estimation using invariants. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 13820–13827. IEEE, 2021. 4

[50] Jamie Shotton, Ben Glocker, Christopher Zach, Shahram Izadi, Antonio Criminisi, and Andrew Fitzgibbon. Scene coordinate regression forests for camera relocalization in rgb-d images. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2930–2937, 2013. 7

[51] Rainer Storn and Kenneth Price. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4):341–359, 1997. 8

[52] Yuliana Susanti, Hasih Pratiwi, Sri Sulistijowati, Twenty Liana, et al. M estimation, s estimation, and mm estimation in robust regression. *International Journal of Pure and Applied Mathematics*, 91(3):349–360, 2014. 1

[53] Philip HS 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. 1, 4

[54] Vasileios Tzoumas, Pasquale Antonante, and Luca Carlone. Outlier-robust spatial perception: Hardness, general-purpose algorithms, and guarantees. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5383–5390. IEEE, 2019. 4

[55] Shinji Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 13(04):376–380, 1991. 2, 4

[56] Sara A Van de Geer and Sara van de Geer. *Empirical Processes in M-estimation*, volume 6. Cambridge university press, 2000. 1

[57] Aad W Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000. 1

[58] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni

Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020. 8

[59] David J Wales and Harold A Scheraga. Global optimization of clusters, crystals, and biomolecules. *Science*, 285(5432):1368–1372, 1999. 8

[60] Yue Wang and Justin M Solomon. Deep closest point: Learning representations for point cloud registration. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 3523–3532, 2019. 4

[61] Yue Wang and Justin M Solomon. Prnet: Self-supervised learning for partial-to-partial registration. *Advances in neural information processing systems*, 32, 2019. 4

[62] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *Acm Transactions On Graphics (tog)*, 38(5):1–12, 2019. 4

[63] Fei Wen, Rendong Ying, Zheng Gong, and Peilin Liu. Efficient algorithms for maximum consensus robust fitting. *IEEE Transactions on Robotics*, 36(1):92–106, 2019. 4

[64] Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1912–1920, 2015. 6

[65] Jianxiong Xiao, Andrew Owens, and Antonio Torralba. Sun3d: A database of big spaces reconstructed using sfm and object labels. In *Proceedings of the IEEE international conference on computer vision*, pages 1625–1632, 2013. 7

[66] Heng Yang, Pasquale Antonante, Vasileios Tzoumas, and Luca Carlone. Graduated non-convexity for robust spatial perception: From non-minimal solvers to global outlier rejection. *IEEE Robotics and Automation Letters*, 5(2):1127–1134, 2020. 2

[67] Heng Yang and Luca Carlone. A quaternion-based certifiably optimal solution to the wahba problem with outliers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1665–1674, 2019. 4

[68] Heng Yang, Jingnan Shi, and Luca Carlone. Teaser: Fast and certifiable point cloud registration. *IEEE Transactions on Robotics*, 37(2):314–333, 2020. 2, 4, 6, 7

[69] Jiaolong Yang, Hongdong Li, Dylan Campbell, and Yunde Jia. Go-icp: A globally optimal solution to 3d icp point-set registration. *IEEE transactions on pattern analysis and machine intelligence*, 38(11):2241–2254, 2015. 4

[70] Jiaolong Yang, Hongdong Li, and Yunde Jia. Go-icp: Solving 3d registration efficiently and globally optimally. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1457–1464, 2013. 4

[71] Ming Ye, Robert M Haralick, and Linda G Shapiro. Estimating piecewise-smooth optical flow with global matching and graduated optimization. *IEEE transactions on pattern analysis and machine intelligence*, 25(12):1625–1630, 2003. 2

[72] Christopher Zach and Guillaume Bourmaud. Descending, lifting or smoothing: Secrets of robust cost optimization. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 547–562, 2018. 2

[73] Andy Zeng, Shuran Song, Matthias Nießner, Matthew Fisher, Jianxiong Xiao, and T Funkhouser. 3dmatch: Learning the matching of local 3d geometry in range scans. In *CVPR*, volume 1, page 4, 2017. 7

[74] Ji Zhao, Wanting Xu, and Laurent Kneip. A certifiably globally optimal solution to generalized essential matrix estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12034–12043, 2020. 2

[75] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Fast global registration. In *European conference on computer vision*, pages 766–782. Springer, 2016. 2, 4, 5, 6, 7