A Closed-Form Uncertainty Propagation in Non-Rigid Structure from Motion

Jingwei Song, Mitesh Patel, Ashkan Jasour, and Maani Ghaffari

Abstract—Semi-Definite Programming (SDP) with low-rank prior has been widely applied in Non-Rigid Structure from Motion (NRSfM). A low-rank constraint avoids the inherent ambiguity of the basis number selection in conventional base-shape or base-trajectory methods. Despite SDP-based NRSfM’s efficiency, it remains unclear how to propagate the noisy tracked feature points’ uncertainty to the 3D recovered shape in SDP-based NRSfM formulation. This paper presents a closed-form statistical inference for the element-wise uncertainty propagation of the estimated deforming 3D shape points in the exact low-rank SDP-based NRSfM. Then, we extend the exact low-rank uncertainty propagation to the approximate low-rank scenario with an optimal numerical rank selection method. The proposed method provides an independent module to the SDP-based method and only requires the statistical information of the input 2D trackings. Extensive experiments show that the major uncertainty in the recovered 3D points follows normal distribution, the proposed method quantifies the uncertainty accurately, and it has desirable effects on the routinely SDP low-rank based NRSfM solver. The code is open-sourced at https://github.com/JingweiSong/NRSfM_uncertainty

Index Terms—NRSfM, uncertainty quantification, SDP

I. INTRODUCTION

NON-RIGID Structure from Motion (NRSfM) is the topic of recovering the camera motion and the 3D time-varying shape simultaneously from sequential 2D trajectories in the monocular video. It contributes to 3D shape perception in various applications, including 3D reconstruction and scene understanding, using consumer-level digital cameras. Uncertainty propagation has been heavily analyzed in robotics [1] and, Structure from Motion (SfM) [2], [3] as the noise of the observation is not negligible. Uncertainty generally includes aleatoric uncertainty (relating to measurement noise), epistemic uncertainty (relating to model parameters), and structure uncertainty (relating to model structure) [4], but the uncertainty in SfM and Simultaneous Localization and Mapping (SLAM) is normally referred as aleatoric. The uncertainty propagation of the estimated state provides important statistical descriptions for further applications such as quality assessment [5], localization [6], mapping [7], path planning [8], and multi-source information fusion [9], [10]. In NRSfM community, [11], [12] propagate uncertainty sequentially with their modified Extended Kalman Filter (EKF), which is coupled with map deformation modeling. We do not find any uncertainty estimation algorithm for the widely used Semi-Definite Programming (SDP) based NRSfM. The lack of uncertainty analysis in SDP-based NRSfM may be attributed to the difficulty in linearizing the objective function or the discontinuity of the objective function. Unlike in robotics and SfM, where the sensor-to-object functions can be linearized easily, NRSfM involves complicated nonlinear formulation such as factorization or nuclear norm minimization, which hinders the process of uncertainty propagation.

This work proposes a closed-form solution to the element-wise uncertainty propagation for the group of SDP-based NRSfM [13] formulations. Extensive experiments show that the proposed uncertainty estimation algorithm describes the estimated shape well and can be applied in measuring confidences of the recovered 3D deformable shape. Additionally, we adapt the exact rank algorithm with an empirical approximate rank estimation method for the prior-free SDP-based approaches. Results show that the rank has a remarkably small impact on the element-wise variance estimation.

It should be noticed that we only aim at uncertainty esti-
mation of the recovered deforming shape. Uncertainty of the camera rotation matrix cannot be obtained due to ambiguity. [14] shows that the camera rotation matrix and estimated shape cannot be retrieved simultaneously due to their ambiguity. Normally the camera pose is estimated with other off-the-shelf techniques or rigid background [13], [15]. Thus, our work only focuses on the 3D shape uncertainty in SDP-based NRSfM. In particular, this work has the following contributions.

1) We propose a closed-form element-wise uncertainty propagation algorithm for SDP-based NRSfM.
2) A numerical rank estimation method is proposed to define the best rank for the confidence estimation algorithm.
3) The code is publicly available for future research and benchmark comparison. 1

II. RELATED WORK

This section briefly reviews the state-of-the-art works in NRSfM and rigid uncertainty propagation methods.

A. NRSfM

Following the rigid SfM community, early NRSfM works [16] directly adopted factorization to recover the time-varying 3D shapes. Later, [17] showed ambiguous solvability in NRSfM that factorization alone is insufficient in solving the ill-posed NRSfM problem. Since then, priors have been introduced to constrain the problem into a low-rank subspace to ensure solvability. The low-rank priors include base shape [18], [19], base trajectory [18], [20], [21], base shape-trajectory [22], [24] and force model [25]. There is a duality in the formulation of all these proposed base prior, which is, the estimated state (shape or trajectory) is a linear combination of all bases. Innovations include ‘coarse-to-fine’ [18], probabilistic principal component analysis [19], kernel trick [22], and Procrustean analysis in consecutive shapes [26] were proposed. As [14] pointed out, the internal constraint orthonormality of camera orientation and external constraint of fixed base shapes enable the unique shape structure.

One milestone was achieved by [27] and [13] who enforced low-rank constraints for spatial-temporal smoothness of the shape. Different from conventional basis formulation, which is in a hard low-rank constraint, the SDP-based formulation is in the soft low-rank loss (nuclear norm for the objective function’s convexity) minimization. As [13] pointed out, the number of basis is not essential in their SDP formulation and thus can be classified as prior-free. The only constraint needed is the low nuclear norm constraint in their SDP-based formulation. The automatic prior-free SDP achieves high accuracy with fewer parameter configurations. SDP and their derived methods are the most widely used algorithms in the state-of-the-art NRSfM community. Following researchers pushed the low-rank formulation toward spanning the model with a union of low dimensional shape subspace [15], [28]–[37] for multiple body reconstruction. [38] successfully filled the gap between the theory and application. Other prior-based learning-based methods include [39]–[43]. These ‘training and testing’ methods fall out of the scope of this article. This study concentrates on the prior-free and training-free SDP-based methods.

B. Rigid Uncertainty Propagation

The uncertainty propagation has been extensively studied in SfM, and SLAM communities. The confidences are of equal importance as the predictions [44]. In SfM or SLAM, uncertainty is defined as the impact of input perturbation and measured as covariance matrices. SfM process normally linearizes the objective function and solves with Gauss-Newton algorithm [45], which defines the covariance as the inverse of the second-order derivations known as ‘Hessian matrix’. Numerically, the Hessian matrix is approximated by $H = J^T J$ where $J$ is the first-order derivative. In sequential state estimation, SLAM [46] propagates the uncertainty through the linearized state propagation function in the EKF. With known initial pose and map uncertainties, SfM system balances the weights to yield the optimal estimation and the associated uncertainty. Built on the EKF theory and deformation modeling, the sequential NRSfM [11], [12] algorithms, or termed deformable SLAM, achieved camera orientation and map uncertainty propagation iteratively in the dynamic process.

The closed-form uncertainty estimation in SDP-based NRSfM, however, is far from the straightforward solutions. Although numerous researches are conducted in the rigid scenario or sequential NRSfM, no work analyzes its related uncertainty. Due to the nonlinear nuclear norm constraint, SDP fails to pin the formulation down to a closed-form version as SLAM or SfM. The nuclear norm term is an obstruct for closed-form uncertainty propagation. Recently, a breakthrough has been achieved in optimal uncertainty propagation and inference for noisy matrix completion [47]. The relaxed convex optimization in matrix completion shares the similarity in SDP in NRSfM [13]. It measures the confidence interval by developing simple de-biased estimators that admit tractable and accurate distributional characterizations. This article is inspired by [47] and provides a solution to quantify the uncertainty of the estimated time-varying shape. Moreover, [47] only presented a solution to the matrix with an exact low-rank structure. We provide a numerical solution in NRSfM to allow uncertainty propagation for the approximate low-rank matrix form.

III. METHODOLOGY

A. Problem definition and the SDP solver

The classic SDP-based NRSfM [13] formulates the time-varying shape recovery as minimizing
\[
\min_{S, R} \mu \|S\|_* + \frac{1}{2} \|W - RS\|_F^2, \quad \text{such that}
\]
\[
S^2 = g(S) = [P_X P_Y P_Z] \odot (I_3 \otimes S)
\]
\[
RR^T = I_{2F}
\]
where $\|\cdot\|_*$ is the nuclear norm, $\|\cdot\|_F$ is the Frobenius norm and $\otimes$ is the Kronecker product. $I_{2F}$ is the identity matrix with the
size $2F$. $\mu$ is the hyper-parameter balancing two constraints. The time-varying shape $S$ is a $3F \times N$ matrix with $F$ frames and $N$ 3D feature points $(p_{ij}) = [x_{ij}, y_{ij}, z_{ij}]$ point $j$ in frame $i$ which are sequentially permuted. $S^*$ is the repermuted matrix of $S$ and in the size $3N \times F$. It is the rearranged form of the matrix $S$ for conciseness. Let $S^d = g(S)$ maps $S$ to $S^d$ while $S = g^{-1}(S^d)$ does the opposite. $P_X, P_Y, P_Z \in \mathbb{R}^{F \times 3F}$ are some properly defined 0-1 valued ‘row-selection’ matrices (similar to the ‘permutation matrix’). $R$ is the camera rotation matrix composed of the diagonal block matrix $R_1$. $R_1$ is the first two rows of its full rotation form $R_1 \in SO(3)$. $R$ defines the 3D to 2D projection. As $[15]$ pointed out, there are several off-the-shelf methods to estimate the camera rotation matrix $R$ prior to [1]. Thus, $R$ is estimated firstly and assumed noiseless. $W \in \mathbb{R}^{2F \times N}$ is the noisy observation matrix of the 2D tracking as

$$W_{ij} = W^*_d + E_{ij}, E_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$$

(2)

where $\Omega \subseteq \{1, \ldots, 2F\} \times \{1, \ldots, N\}$ is the subset of indexes, $W^*_d$ is the noise-free observation at point $j$ in frame $i$. The uniform Gaussian noise $E_{ij}$ denotes the spatial disturbance of element $(ij)$ in measurement. $\sigma^2$ is the variance of the Gaussian noise. The permuted version $E^d = g(E)$. 

Equation (1) was proposed to overcome the difficulty in selecting the optimal rank in searching for $S$ [13]. The low-rank constraint makes the objective function nonconvex. Therefore, nuclear norm minimization is adopted to relax the problem to convex form [13, 27].

B. Uncertainty propagation in the low-rank structure

Fig. 2 presents the relationship of the proposed method and previous research [13]. In this paper, we propose a novel module (in orange) to quantify the uncertainty of the SDP-based deformable 3D shape estimation (in blue). This section presents a method to retrieve the closed-form uncertainty in the scenario that the time-varying shape is in the exact low-rank. Section III-C extends it to the approximate low-rank version. Note that in the SDP formulation (1), the ground truth of the time-varying deforming shape is in approximate low-rank structure, meaning it is strictly full-rank but has a small nuclear norm, which can be viewed as the approximated low-rank structure. It should be noticed that the impact of noise on the rotation matrix $R$ is not considered because it is obtained separately.

Denote $S^d \in \mathbb{R}^{3N \times F}$ as the rank-$r$ matrix estimated from the noise-free observation $W^*$ (the exact low-rank structure after solving (1)). Please note that $S^d$ is not the ground truth but the noise-free estimation. Denote $S^d_\mu \in \mathbb{R}^{3N \times F}$ as the rank-$r$ matrix estimated from the noisy observation $W$. We enforce Singular Value Decomposition (SVD) and we have $S = U\Sigma V^T$ and $S^d = U^d\Sigma^d V^dT$. We define two auxiliary matrices $X := U\Sigma^{1/2} \in \mathbb{R}^{3N \times r}$ and $Y := V\Sigma^{1/2} \in \mathbb{R}^{F \times r}$, where $X$ and $Y$ are the globally rectified matrices. Similarly, define $X^d \in \mathbb{R}^{3N \times r}$, $Y^d \in \mathbb{R}^{F \times r}$ as unrectified estimations. The following rules apply.

$$\hat{X}^T \hat{X} = \hat{Y}^T \hat{Y} = \Sigma, \ S^d = \hat{X} \hat{Y}^T$$

$$X^d \hat{X} = Y^d \hat{Y} = \Sigma^d, \ S^d_\mu = X^d \hat{Y}^dT$$

(3)

In conventional NRSfM in shape basis formulation, $\hat{Y}$ is termed as the basis matrix and $X$ is the linear coefficient matrix [20]. Following the ambiguity issue in the coefficient matrix and basis matrix raised by [14], an optimal global rectification matrix $H_d \in \mathbb{R}^{r \times r}$ is needed to align the noisy estimation $(X^d, Y^d)$ to the noise-free and rectified estimation $(\hat{X}, \hat{Y})$. $H_d$ can be obtained as

$$H_d := \arg \min_{H \in \mathcal{A}^{r \times r}} \|X^d H - \hat{X}\|_F^2 + \|Y^d H - \hat{Y}\|_F^2$$

(4)

where $\mathcal{A}^{r \times r}$ denotes orthogonal matrix manifold.

Theorem 1. On condition that the observation noise matrix $E$ is i.i.d., follows Gaussian and is relatively small, the errors between the estimated operator $(\hat{X}, \hat{Y})$ from the noisy observation $W$ and the estimated operator $(\hat{X}, \hat{Y})$ from the noise-free observation $W^*$ are

$$X^d \hat{X} - \hat{X} = Z_X + \Psi_X$$

$$Y^d \hat{Y} - \hat{Y} = Z_Y + \Psi_Y$$

(5)

where $\Psi_X, \Psi_Y \in \mathbb{R}^{3N \times r}$ and $\Psi_Y \in \mathbb{R}^{F \times r}$ are the non-Gaussian residual matrices. $\Psi_X$ and $\Psi_Y$ are always smaller than Gaussian error matrices $Z_X$ and $Z_Y$. The rows of the error matrix $Z_X \in \mathbb{R}^{3N \times r}$ (resp. $Z_Y \in \mathbb{R}^{F \times r}$) are independent and obey

$$Z^T_X e_j \overset{i.i.d.}{\sim} \mathcal{N}(0, \frac{3}{2} \sigma^2 (\Sigma^d)^{-1}), \ \text{for } 1 \leq j \leq r$$

$$Z^T_Y e_j \overset{i.i.d.}{\sim} \mathcal{N}(0, \frac{3}{2} \Sigma (\Sigma^d)^{-1}), \ \text{for } 1 \leq j \leq r$$

(6)

where $e_j$ is the base vector with one element 1 and rest 0. (4) strictly holds on condition that the frame size $F$ (observed from different orientations) is infinite. Appendix proves Theorem 2.

Based on Theorem 1, the element-wise difference between the estimation $S^d_\mu$ from the noisy observation and the estimation $S^d$ from the noise-free observation is

$$S^d_\mu - S^d = [X^d H^d (Y^d H^d)^T - \hat{X} \hat{Y}^T]_{ij}$$

$$= [e_i^T X^d H^d (Y^d H^d)^T e_j - e_i^T (\hat{X} \hat{Y}^T) e_j]$$

$$= [e_i^T X^d H^d (Y^d H^d)^T e_j - e_i^T (X^d H^d Z_Y)^T e_j]$$

(7)

$$= [e_i^T (Z_X (Y^d H^d)^T e_j + e_i^T (X^d H^d Z_Y^T e_j)]$$

where the bases $e_i, e_j$ localize the elements involved in calculating the elements in location $(i, j)$. The higher order error
we have the element-wise variance of the error as
\[ \Psi \approx \text{Var}(\varepsilon_i (X^d H^d) (Y^d H^d)^T e_j) + \text{Var}(\varepsilon_i ((X^d H^d) Z Y^T e_j)) \]
\[ = \frac{1}{2} \sigma_0^2 e_i^T (Y^d H^d) (\Sigma^d)^{-1} (Y^d H^d)^T e_i + e_i^T (X^d H^d) (\Sigma^d)^{-1} (X^d H^d)^T e_i \]
\[ \approx \frac{1}{2} \sigma_0^2 (||U_{i,}||_2^2 + ||V_{j,}||_2^2) \]
\[ \text{(8)} \]

where (i) is from Theorem [1] since \( Z_X \) and \( Z_Y \) are independent. (ii) is from [5]. (iii) \( U_{i,} \) and \( V_{j,} \) are the \( i \) and \( j \)th row of \( U^d \) and \( V^d \). For conciseness, we define \( v_{ij}^d = (||U_{i,}||_2^2 + ||V_{j,}||_2^2) \). Similarly, the covariance between the element \((i, j)\) and \((m, n)\) is
\[ \text{Cov}(S^d_{ij} - \hat{S}^d_{ij}, S^d_{mn} - \hat{S}^d_{mn}) = \frac{3}{2} \sigma_0^2 e_i^T (Y^d H^d) (\Sigma^d)^{-1} (Y^d H^d)^T e_m + e_i^T (X^d H^d) (\Sigma^d)^{-1} (X^d H^d)^T e_n \]
\[ \text{(9)} \]

C. Uncertainty for approximate low-rank structure

In the presence of noisy observations, an exact subspace projection can be used to mitigate the impact of noises. The original approximate SDP-based formulation [13] is mainly built for modeling noise-free observations. In solving the SDP contaminated with Gaussian noise, the original method [13] falls into the trap of over-fitting the noises. To efficiently handle this issue, the converged shape from the approximate SDP solver can be projected onto an exact low-rank subspace addressing the relation of shape and residual. With an optimal exact low-rank structure, the reprojection error should be consistent with the Gaussian noise.

The optimal rank selection is straightforward (Algorithm 1) and can be coupled with SDP-based NRSfM [13] as the post-processing module. The noise of the observation follows Gaussian distribution [2]. We enforce this constraint by iteratively searching the optimal rank of \( S^d \) whose reprojection residual \( ||W - RS^d|| \) is closest to the prior distribution \( W \). Specifically, the testing rank \( r \) is traversed from 1 to 10 in maximum in searching for the optimal rank.

D. Potential applications

Uncertainty is indispensable in fusing multiple data sources. In contrast to batch optimization for dense or long trajectory cases, we demonstrate a preliminary test to segment the entire trajectory into several overlapping sub-trajectories. Each sub-trajectory is fed into SDP solver individually (and in parallel with multi-core CPU). Then, all sub-trajectories are fused based on the estimated uncertainty. In this segmented process, the heavy time consumption on iterative SDP decomposition can be significantly reduced because the computation is on a much smaller matrix. Take SDV solver from [49] as an example, the time complexity is \( O(\min (mn^2, m^2 n)) \) (\( m \) and \( n \) are matrix size). Therefore, the uncertainty estimation in this research enables the uncertainty-aware fusion of the overlapping sub-trajectories.

Algorithm 1: Rank searching for noisy observations.

Input: \( W, \sigma_0, S^d \) from SDP
Output: Optimal rank \( r \) and exact version of \( S^d \)

1. while \( r < \max(3N, F) \) do
2. \( S^d = g(\pi_r(g^{-1}(S^d))) \)
3. \( r = r + 1 \)
4. /* Check Convergence */
5. 95% elements in \( ||W - Rg^{-1}(S^d)|| \) within 1.96\( \sigma_0 \)
end

Notation: \( \pi_r(\cdot) \) projects to exact rank with hinge loss.

IV. RESULTS AND DISCUSSION

Similar to previous works [13], [15], [28], the proposed method has been validated quantitatively on the classic MoCap data set [50] which uses 41 markers distributed on the surface of the human body. Five sequences were chosen and their 3D points were projected onto the 2D space with a virtual orthographic camera following a relative circular motion around the object at a stable angular speed. The five data sets were: ‘Drink’, ‘Pickup’, ‘Yoga’, ‘Stretch’ and ‘Dance’. Two real-world dense deforming ‘paper’ and ‘T-shirt’ datasets obtained by Kinect with 3D ground truth were tested [51]. Validity masks were enforced on them. Moreover, two dense face data sets from [42] with ground truth were also adopted to test the performance. The ground truth rotations were directly adopted for the dense data sets. All shapes were normalized to the range [0, 1]. All computations were conducted on a commercial desktop with CPU i5-9400 and Matlab 2020a.

Table I reveals that the proposed exact rank searching makes the SDP-based formulation more robust to noisy observations as claimed in Section III-C. It indicates that the proposed noise-aware exact rank searching (Algorithm 1) is necessary for noisy observations (at least more than 1% standard deviation). We project the approximate \( S^d \) to the subspace correctly. The proposed noise-aware module helps the SDP-based NRSfM achieves better accuracy in the presence of the Gaussian noises. In experiment, we also notice that projecting the time-varying shape onto subspace with a smaller rank achieves better accuracy, especially with heavy noises.

A. Aleatoric uncertainty propagation

Monte Carlo tests were conducted on the 2D observations with different levels of noise. For each test, Gaussian noise
was imposed on the 2D tracked points with different standard deviations $\sigma_0$. All parameter settings STRICTLY followed the work [13]. We also tested the augmented Lagrange multiplier solver used in [15] (our implementation). All the numerical results from Lagrange multiplier and augmented Lagrange multiplier solvers are very close since the cost functions are the same in (1). Thus, only the results from the Lagrange multiplier [13] are presented. Besides, we need to clarify that most state-of-the-art SDP-based algorithms are not open-sourced and it is difficult to test more. Monte Carlo test of $T = 100$ times were performed on each data set with the given $\sigma_0$. Since the noise-free estimation $\hat{\Psi}$ from Algorithm 1 suffers from local minima, the Monte Carlo’s average $\hat{\Psi}_{\text{avg}}$ is used instead. The element-wise error in trial $k$ is defined as

$$e_{i,j}^{(k)} = \hat{\Psi}_{i,j}^{(k)} - \hat{\Psi}_{i,j}^{\text{avg}}$$

We intend to validate that the estimated variance $\frac{1}{2} \sigma_0^2 v_{i,j}^{(k)}$ is correct, that is, 95% of $e_{i,j}^{(k)}$ fall within the range of

$$-1.96 \sigma_0 \sqrt{\frac{1}{2} v_{i,j}^{(k)}}, 1.96 \sigma_0 \sqrt{\frac{1}{2} v_{i,j}^{(k)}}$$

$(v_{i,j}^{(k)} \text{ is the } k\text{th trial of } v_{i,j}^{(k)}).$ Thus, define the coverage rate as the ratio of errors fall in the 1.96 bound. Table II shows the general statistical results over all elements of the 5 data sets, and it indicates that the closed-form uncertainty quantification coverage rate is close to 95%. The mean and standard deviation count the element-wise coverage rate since the element-wise coverage rates cannot be presented individually. Table II shows that the coverage rate is close to 95% and is more accurate in large $\sigma_0$. Thus, the impact of the non-Gaussian residuals $\Psi_X$ and $\Psi_Y$ is small regarding the Gaussian residuals. Fig. 1 and the attached video are provided for better visualization of the error ellipsoid with $1.96 \sigma_0$ bound ($\sigma_0 = 0.05$). Finally, it should be addressed that the uncertainty is aleatoric and should be validated with Monte Carlo test (routine in SLAM and SfM). Comparing directly with ground truth involves epistemic and structure uncertainty.

In addition to the general coverage rate of the proposed aleatoric uncertainty propagation approach, we also validate that the error of the estimated shape $\hat{\Psi}_{\text{avg}}$ follows the Gaussian distribution. The Quantile-Quantile (Q-Q) plots (Fig. 3) of $\hat{\Psi}_{i,j}^{(k)}$ (element (1,1) of ‘drink’), $\hat{\Psi}_{i,j}^{(k)}$ (element (3,4) of ‘pickup’), $\hat{\Psi}_{i,j}^{(k)}$ (element (5,8) of ‘stretch’), $\hat{\Psi}_{i,j}^{(k)}$ (element (30,30) of ‘yoga’) and $\hat{\Psi}_{i,j}^{(k)}$ (element (25,60) ‘yoga’) are selected randomly and presented in the case of $\sigma_0 = 0.05$ against the standard normal distribution. Table III is presented to show that the errors of the recovered 3D points follow the normal distribution. The $p$-value reveals that 15 out of 16 samples are significant in

1Readers are encouraged to test the provided code.
numerical estimation process. Among the 16 tests, 9 achieve
error, and 12 achieve 3% errors regarding Monte Carlo average’s 95% coverage rate. Moreover, the standard deviations do not increase significantly with larger rank disturbance in Table IV. Therefore, the presented uncertainty propagation approach is robust to the disturbances on the rank numerically estimated in Algorithm 1.

### TABLE V: The comparisons between the original work [13] and the overlapping sub-trajectory fusion. The optimized sub-trajectories are fused based on the closed-form uncertainty. Time is measured in 1000s. Time* is obtained by parallel computing on 6 cores CPU.

|       | Original | Uncertainty-aware |
|-------|----------|------------------|
|       | Accuracy | Time | Accuracy | Time | Time* |
| Face3 | 0.052    | 7.363 | 0.053    | 5.978 | 1.533 |
| Face4 | 0.044    | 6.972 | 0.044    | 5.193 | 1.266 |

The normal distribution (> 0.05). Since all elements in the recovered shape should be consistent (either obey or reject the normal distribution), we can conclude that the shape noise is predominated by Gaussian noise and the non-Gaussian residuals $Ψ_\mathbf{X}$ and $Ψ_\mathbf{Y}$ are trivial.

### B. Robustness of the approximate structure

Apart from the coverage rate tests, we further validate that the proposed approach is robust to the numerical rank selection process. Numerically, $\mathbf{V}^{\dagger}_{ij} \in [0, 2]$ is dependent on the rank selection, as it is 0 when the matrix is null and 2 when it is full-rank. Different ranks in Algorithm 1 were tested and presented in Table IV.

Table IV illustrates the coverage rate with different disturbances on the rank numerically estimated in Algorithm 1. Among the 16 tests, 9 achieve 2% error, and 12 achieve 3% errors regarding Monte Carlo average’s 95% coverage rate. Moreover, the standard deviations do not increase significantly with larger rank disturbance in Table IV. Therefore, the presented uncertainty propagation approach is robust to the optimal rank selection in the numerical estimation process. The standard deviations reveal the robustness to subspace rank selection.

### C. Potential applications

One potential application of the confidence quantification is data fusion. To reduce the time consumption, the entire trajectory can be divided into several submaps and joined after processing [52]. Two dense face data sets from [42] with ground truth were adopted. Each sequential data set was divided into 6 sub-trajectories with 20% overlapping and processed in parallel with 6 CPU cores. Table V shows that parallel processing is much faster than batch optimization. The accuracy of the sub-trajectories fusion is similar to SDP in a batch. Further tests indicate that the accuracy of the uncertainty-aware fused shape on the overlapping regions is from 3% to 5% higher than the average fusion.

V. CONCLUSION AND FUTURE WORKS

This is the first research to address the uncertainty propagation in SDP-based NRSfM problems. We propose a closed-form element-wise uncertainty propagation algorithm for the state-of-the-art SDP-based approach. The proposed method only requires the statistical distribution of the errors in 2D observation. Monte Carlo tests validate that the coverage rate of the element-wise variance precisely describes the estimated time-varying shape distribution. Robustness tests demonstrate that the uncertainty is not sensitive to the rank estimated with our modified SDP workflow. Our closed-form uncertainty propagation approach may benefit applications like parallel processing of SDP-based methods.

Although the uncertainty cannot directly improve the accuracy, it allows a more statistically sound way to utilize the estimation results, such as parallel processing, fusion with other sensors or risk evaluation of the system in a safety-critical situation (e.g., autonomous vehicle). Future work may focus on exploiting solutions to estimate uncertainty quantification for 2D trackings with missing or occluded elements.

ACKNOWLEDGMENT

Toyota Research Institute provided funds to support this work. Funding for M. Ghaffari was in part provided by NSF Award No. 2118818.

APPENDIX.

The appendix is the proof of Theorem 1. This section first explicitly derives the relation between $\mathbf{S}$ and $\mathbf{R}$ in (3). Then, the uncertainties of $\mathbf{X}^d$ and $\mathbf{Y}^d$ are inferred taking advantage of the zero derivatives of (1) at its local minima ($\mathbf{X}^d$ and $\mathbf{Y}^d$).

**Remark 1.** Given two i.i.d. variables $x \sim \mathcal{N}(\gamma, \delta_x^2)$ and $y \sim \mathcal{N}(\gamma, \delta_y^2)$, the product $xy$’s variance $\delta_{xy}^2 = \delta_x^2 + \delta_y^2 - 2\gamma^2$ if $\gamma = 0$, the variance degrades to $\delta_{xy}^2 = \gamma^2 \delta_x^2$.

**Remark 2.** Denote the non-Gaussian noise of $\mathbf{G}_1^d = \mathbf{X}^d \mathbf{HY}^T - \mathbf{XY}^T$ and $\mathbf{G}_2^d = \hat{\mathbf{X}}^H \mathbf{Y}^d \mathbf{T} - \hat{\mathbf{X}} \hat{\mathbf{Y}}^T$ resulted from the local optimization. Since shape noise $\mathbf{E}^d = \mathbf{X}^d \mathbf{Y}^d \mathbf{T} - \mathbf{XY}^T$ and the global rectification $\mathbf{R}$, the non-Gaussian noises $\mathbf{G}_1^d$ and $\mathbf{G}_2^d$ are always smaller than $\mathbf{E}^d$.

**Remark 3.** Given a random orthogonal matrix $\mathbf{U}_i \in \mathbb{A}^{3N \times 3N}$, the element-wise expected value $\mathcal{E}(\mathbf{U}_i) = \frac{1}{\sqrt{3N}} \mathbf{1}$, where $\mathbf{1}$ is the all 1 matrix with size $3N \times 3N$. For the matrix $\hat{\mathbf{U}}_i$, which is composed of the first $2N$ rows of $\mathbf{U}_i$, $\mathcal{E}(\hat{\mathbf{U}}_i^T \hat{\mathbf{U}}_i) = p \mathbf{I}_N$ and $p = \frac{2}{3}$.
where $\tilde{U} = \Pi(\tilde{R}_0) \tilde{X} \tilde{Y}^T_{\tilde{F}}, \cdots, \Pi(\tilde{R}_F) \tilde{X} \tilde{Y}^T_{\tilde{F}} \approx U_{\Delta} \tilde{X} \tilde{Y}^T$, (11)

$$
\Pi(\tilde{R}_i) = \begin{bmatrix} I_N \tilde{R}_{i(11)} & I_N \tilde{R}_{i(12)} & I_N \tilde{R}_{i(13)} \\
I_N \tilde{R}_{i(21)} & I_N \tilde{R}_{i(22)} & I_N \tilde{R}_{i(23)} \\
I_N \tilde{R}_{i(31)} & I_N \tilde{R}_{i(32)} & I_N \tilde{R}_{i(33)} \end{bmatrix},
$$

(12)

where $U_{\Delta} \in \mathbb{R}^{3N \times 3N}$ is the orthogonal matrix satisfying all $\Pi(\tilde{R}_i), i \in [1, F]$. The orthogonal matrix $\Pi(\tilde{R}_i) \in \mathbb{R}^{3N \times 3N}$ is defined in [21]. $\tilde{R}_{i(uv)}$ is the $(u, v)$ element of $R_i, I_N$ is the identity matrix with size $N$. Next, we discuss the existence of $U_{\Delta}$.

**Remark 4.** On the manifold of all factorization estimators of $S^d$, there are infinite exact $U_{\Delta}$ on condition that $\frac{3N(3N-1)}{2} > F$, one exact $U_{\Delta}$ if $\frac{3N(3N-1)}{2} = F$ and an optimal approximate $U_{\Delta}$ if $\frac{3N(3N-1)}{2} < F$. For all three situations, with the sample $U_{\Delta}$ large enough (or frame numbers $F$), the expectant $E(U_{\Delta}^T U_{\Delta}) = p I_{3N}$ ($p = \frac{2}{3}$) and $E(U_{\Delta}^T U_{\Delta}) = \frac{1}{\sqrt{3N}} I_{3N}$.

This follows remark [3] $\tilde{U}_{\Delta}$ is the first $2N$ rows of $U_{\Delta}$. The accuracy is dependent on the number of frames $F$.

We next take advantage of the zero derivatives and propagate the noise $E$ to $X^d$ and $Y^d$. Denote $Q$ as the 0-1 valued ‘row selection matrix’ to select components of $W$. $\|W\|_F^2$ and $\|Q \circ W\|_F^2$ are equivalent ($\circ$ is the Hadamard product). Following remark [4] and by reaching the optimal solution, the nuclear norm of the objective function [4] is known and there are exact $U_{\Delta}$ satisfying

$$
\begin{align*}
\frac{\partial f(X^d, Y^d)}{\partial X^d} & \approx \frac{1}{|2|} \frac{1}{\sqrt{3N}} E^d Y^d (Y^d Y^d)^{-1} H \sum_{i=1}^{F} \frac{1}{\sqrt{1+\|Y_{i,1,1}\|^2}} Y_{i,1,1} E^d_{i} H^{-1} Y_{i,1,1} E^d_{i} \sum_{i=1}^{F} \frac{1}{\sqrt{1+\|Y_{i,1,1}\|^2}} Y_{i,1,1} E^d_{i},
\end{align*}
$$

(13)

where $Q^d = g(Q)$ and $W \in \mathbb{R}^{3F \times N}$ is the 3D version of $W$ with zero values on Z direction. Note that first $2N$ rows of $Q^d$ ($3N \times F$) are all 1 and last $N$ rows are 0.

With regard to the objective function [13], the optimized shrinkage optimizer $X^d$ and $Y^d$ should satisfy the local minima of (13) denoted as $f(X^d, Y^d)$. $f(X^d, Y^d)$ is close to noise-free estimators $f(X, Y)$ meaning the first-order expansion is close to the zero matrix defined as $O$ (arbitrary size). For simplicity, we denote $\tilde{U}_{\Delta} \tilde{X}$ as $\tilde{X}$ and $\tilde{U}_{\Delta} X^d$ as $X^d$. 1 is the first $2N$ rows of 1.

$$
\begin{align*}
\frac{\partial f(X^d, Y^d)}{\partial X^d} & \approx U_{\Delta} (U_{\Delta} X^d Y^d - g(W)) Y^d \\
& = \left[ U_{\Delta}^T (U_{\Delta} X^d Y^d - \tilde{U}_{\Delta} \tilde{X} \tilde{Y}^T - E^d) \right] Y^d \\
& = (U_{\Delta}^T \tilde{U}_{\Delta} X^d Y^d - U_{\Delta}^T \tilde{U}_{\Delta} \tilde{X} \tilde{Y}^T - U_{\Delta}^T E^d + U_{\Delta}^T \tilde{U}_{\Delta} G_i^2) Y^d = O,
\end{align*}
$$

(14)
