AN APPROXIMATE EXPECTATION-MAXIMIZATION FOR TWO-DIMENSIONAL MULTI-TARGET DETECTION

Shay Kreymer[⋆]*, Amit Singer*[†]*, and Tamir Bendory*[⋆]

[⋆]School of Electrical Engineering, Tel Aviv University, Tel Aviv, Israel
[†]Department of Mathematics and PACM, Princeton University, Princeton, NJ, USA

ABSTRACT

We consider the two-dimensional multi-target detection (MTD) problem of recovering a target image from a noisy measurement that contains multiple copies of the image, each randomly rotated and translated. The MTD model serves as a mathematical abstraction of the structure reconstruction problem in single-particle cryo-electron microscopy, the chief motivation of this study. We focus on high noise regimes, where accurate detection of image occurrences within a measurement is impossible. We suggest an expectation-maximization (EM) framework to estimate the image directly from a measuremnt. Since EM is intractable for the MTD problem, we develop an approximate EM, and demonstrate image recovery in highly noisy environments. We conduct numerical experiments and suggest improvement in estimation accuracy in comparison to previously studied autocorrelation analysis. The code to reproduce all numerical experiments is publicly available at `https://github.com/krshay/MTD-2D-EM`.

***Index Terms*—** Expectation-maximization, multi-target detection, cryo-electron microscopy.

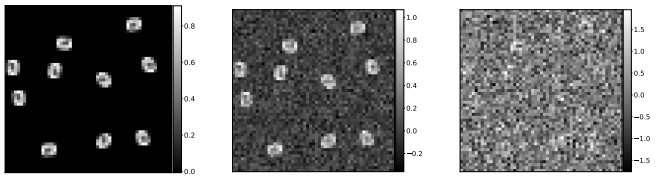
1. INTRODUCTION

We study the multi-target detection (MTD) problem of estimating a target image $f : \mathbb{R}^2 \to \mathbb{R}$ from a noisy measurement that contains multiple copies of the image, each randomly rotated and translated [1, 2, 3, 4, 5, 6, 7]. We consider a measurement $M \in \mathbb{R}^{N\times N}$ of the form

$$M[\vec{\ell}] = \sum_{i=1}^{p} F_{\phi_i}[\vec{\ell} - \vec{\ell}_i] + \varepsilon[\vec{\ell}], \tag{1}$$

where $F_{\phi_i}[\vec{\ell}] := f_{\phi_i}(\vec{\ell}/n)$ is a discrete copy of f, rotated by angle ϕ_i about the origin; n is the radius of the image in pixels; $\{\phi_i\}_{i=1}^{p} \sim \mathrm{Unif}[0, 2\pi)$ are uniformly distributed rotations; $\{\vec{\ell}_i\}_{i=1}^{p} \in \{n+1, \dots, N-n\}^2$ are arbitrary translations; and $\varepsilon[\vec{\ell}]$ is i.i.d. Gaussian noise with zero mean and variance σ^2. The rotations, translations and the number of occurrences of f in M, denoted by p, are unknown. Importantly, since the rotations are unknown, it is possible to reconstruct the target image only up to a rotation. We consider an image f, which is supported on the unit disk $D = \{\vec{x} \in \mathbb{R}^2 : |\vec{x}| \le 1\}$. We assume that f has a finite expansion in the basis of Dirichlet Laplacian eigenfunctions (see [3, 4, 5] for details). We further consider a discrete

S.K. is supported by the Yitzhak and Chaya Weinstein Research Institute for Signal Processing. A.S. is partly supported by AFOSR Award FA9550-20-1-0266, the Simons Foundation Math+X Investigator Award, NSF BIGDATA Award IIS1837992, NSF Award DMS-2009753, and NIH/NIGMS Award R01GM136780- 01. T.B. is supported in part by NSF-BSF grant no. 2019752.



(a) No noise (b) SNR $= 50$ (c) SNR $= 2$

Fig. 1: Three measurements at different SNRs: (a) no noise; (b) SNR $= 50$; (c) SNR $= 2$. Each measurement contains multiple rotated versions of the target image. We focus on the low SNR regime (e.g., panel (c)) in which the locations and rotations of image occurrences cannot be detected reliably, and aim to estimate the target image directly from the measurement.

version of the rotated f_ϕ, $F_\phi : \mathbb{Z}^2 \to \mathbb{R}$, which is given by

$$F_\phi[\vec{\ell}] = f_\phi(\vec{\ell}/n) = \sum_{(\nu,q):\lambda_{\nu,q}\le\lambda} \alpha_{\nu,q}\Psi_{\nu,q}[\vec{\ell}]e^{i\nu\phi}, \tag{2}$$

where $\Psi_{\nu,q}[\vec{\ell}] = \psi_{\nu,q}(\vec{\ell}/n)$, $\psi_{\nu,q}(r,\theta) = J_\nu(\lambda_{\nu,q}r)\,e^{i\nu\theta}$, for $r \le 1$, $\nu \in \mathbb{Z}_{\ge 0}$, J_ν is the ν-th order Bessel function of the first kind, $\lambda_{\nu,q} > 0$ is the q-th positive root of J_ν, and α is the vector of expansion coefficients. In this work, estimating the image means eastimating the vector of expansion coefficients.

We focus on the well-separated case of the 2-D MTD problem, which was introduced in [3, 4]. In this case, each transaltion in the measurement M is separated by at least a full image diameter from its neighbors. Specifically, we assume that

$$|\vec{\ell}_{i_1} - \vec{\ell}_{i_2}| > 4n, \quad \text{for all } i_1 \ne i_2. \tag{3}$$

Figure 1 presents an example of a measurement M at different signal-to-noise ratios (SNRs). We define SNR $:= \frac{\|F_0\|_{\mathrm{F}}^2}{A\sigma^2}$, where A is the area in pixels of F_0 (the unrotated image).

The MTD model serves as a mathematical abstraction of the cryo-electron microscopy (cryo-EM) technology for macromolecular structure determination [8, 9, 10]. In a cryo-EM experiment [11], individual copies of the target biomolecule are dispersed at unknown 2-D locations and 3-D orientations in a thin layer of vitreous ice, from which 2-D tomographic projection images are produced by an electron microscope. It is necessary to keep the electron dose low in order to minimize ~~the irreversible~~ structural damage. Consequently, the projection images are considerably noisy. In the current data processing pipeline for cryo-EM [12, 13, 14], the 2-D projections are first detected and extracted from the micrograph, and later rotationally and translationally aligned to reconstruct the 3-D molecular structure. This approach fails for small molecules, which are difficult to detect and align [7, 8, 12, 15].

The MTD model was devised in [7] in order to study the recovery of small molecules directly from the micrograph, below the current detection limit of cryo-EM [8, 16]. In [3, 4, 5], an autocorrelation analysis technique was implemented for the two-dimensional MTD problem. Autocorrelation analysis is a special case of the method of moments, and it consists of finding an image that best matches the empirical autocorrelations of the measurement. Computing the autocorrelations is straightforward and requires only one pass over the data, which is advantageous for massively large datasets, such as cryo-EM datasets [12].

In this work, we propose an approximate version of the expectation maximization (EM) algorithm [17] that estimates the target image F directly, circumventing estimating the image locations in the measurement ~~by estimating the target image F directly.~~ In the EM algorithm, we aim to iteratively maximize the data likelihood, which marginalizes over the rotations and translations. Importantly, EM is guaranteed to increase the likelihood at each iteration [17]. However, the number of admissible configurations for the locations grows exponentially with the problem size. Therefore, the direct application of the EM algorithm to the MTD problem is computationally intractable, even for very small measurements. Thus, following [2], we suggest mitigating the computational burden by approximating the likelihood.

Previous works suggest that EM outperforms autocorrelation analysis ~~(or the equivalent method of moments)~~ in terms of estimation accuracy, for the one-dimensional MTD problem [2] and for the closely related multireference alignment [18]. In the context of cryo-EM, EM was successfully applied to single-particle reconstruction [19, 20], and enjoys well-established implementations, such as RELION [13] and cryoSPARC [14]. Moreover, a recent paper [21] shows that likelihood optimization in the low SNR regime reduces to a sequence of least squares optimization problems that match the moments of the estimate to the ground truth moments one by one, and by that suggests that EM has the potential to surpass the estimation accuracy achieved by autocorrelation analysis ~~(which uses autocorrelations only up to third~~ order). On the other hand, in contrast to autocorrelation analysis, the approximate EM algorithm scans through the whole dataset in each iteration, and hence requires much longer computational time.

The main contribution of this paper is in developing an approximate EM framework for the two-dimensional MTD problem. We demonstrate a successful reconstruction in noisy regimes (see Section 3). ~~Moreover,~~ we conduct extensive comparison experiments between the approximate EM framework and the ~~previously developed~~ autocorrelation analysis scheme, and suggest significant improvement in estimation accuracy. It is thus a significant step towards efficiently estimating a molecular structure directly from a noisy cryo-EM datasets [7].

2. APPROXIMATE EXPECTATION-MAXIMIZATION

Given the measurement M that follows the MTD model (1) ~~in the well-separated~~ case, the maximum marginal likelihood estimator (MMLE) for the vector of coefficients α, that represents the target image F, is the maximizer of the likelihood function $p(M|\alpha)$. ~~Within the EM framework [17],~~ the locations and rotations of the target images within the measurement are treated as nuisance variables ~~drawn from some distribution~~. The EM algorithm estimates the MMLE by iteratively applying the expectation (E) and maximization (M) steps. Specifically, given the current estimate α_k, the E-step computes the expected log-likelihood function, where the expectation is taken over over all admissible configurations of locations and rotations. The estimate is then updated in the M-step by maximizing the function with respect to α. The major drawback of this approach is that the number of admissible configurations for the locations grows exponentially with the problem size, rendering the direct application of ~~the~~ EM ~~algorithm~~ computationally intractables. Hence, based on [2], we suggest an approximate EM framework.

In our framework of the approximate EM algorithm, we first partition the measurement M into $N_d = (N/L)^2$ non-overlapping patches, each of size $L \times L$, where $L = 2n+1$ is the diameter of the target image F. The image is estimated by the maximizer of the approximate likelihood function

$$p(M_0, M_1, \ldots, M_{N_d-1}|\alpha) \approx \prod_{m=0}^{N_d-1} p(M_m|\alpha), \tag{4}$$

where M_m is the m^{th} patch, and we ignore the dependencies between patches. Our approximate EM algorithm works by applying the EM algorithm to estimate the MMLE of (4).

Overall, a rotated image may appear in $4L^2 - 1$ different ways when it is present in a patch. The separation condition (3) implies that each patch contains parts of at most one rotated image. Depending on the position of image occurrences, the patch M_m can be modeled by

$$M_m = CT_{\vec{l}_m} ZF_{\phi_m} + \varepsilon_m, \quad \varepsilon_m \sim \mathcal{N}(0, \sigma^2 I_{L\times L}). \tag{5}$$

The operator Z zero-pads L entries to the right and to the bottom of the rotated copy of F, and $T_{\vec{l}_m}$ circularly shifts the zero-padded image by $\vec{l}_m$ positions, that is

$$T_{\vec{l}_m} ZF_{\phi_m}[i,j] = ZF_{\phi_m}\left[(i+\ell_{mx}) \bmod 2L, (j+\ell_{my}) \bmod 2L\right], \tag{6}$$

where $\vec{\ell}_m \in \mathbb{L} := \{0, 1, \ldots, 2L-1\}^2$, and is treated as a random variable. The operator C then crops the first L entries in the vertical and horizontal axes, and the result is further corrupted by additive white Gaussian noise. Since the EM algorithm assigns probabilities to different rotations (in the expectation step), we need to discretize the rotations of target images within the measurement. The discretization is represented by the parameter K:

$$\phi_m \in \Phi := \left\{ k\frac{2\pi}{K} \right\}_{k=0}^{K-1}. \tag{7}$$

In the E-step, our algorithm calculates the expected log-likelihood function

$$Q(\alpha|\alpha_k) = \sum_{m=0}^{N_d-1} \sum_{\vec{\ell}\in\mathbb{L}} \sum_{\phi\in\Phi} p(\vec{\ell}, \phi|M_m, \alpha_k) \log p(M_m, \vec{\ell}, \phi|\alpha) \tag{8}$$

given the current coefficients estimate α_k, where

$$p(M_m|\vec{\ell}, \phi, \alpha) \propto \exp\left(-\frac{\|M_m - CT_{\vec{l}} ZF_\phi\|_{\text{F}}^2}{2\sigma^2} \right), \tag{9}$$

with the normalization $\sum_{\vec{\ell}\in\mathbb{L}} \sum_{\phi\in\Phi} p(M_m|\vec{\ell}, \phi, \alpha) = 1$. From Bayes' rule, we have

$$p(\vec{\ell}, \phi|M_m, \alpha_k) = \frac{p(M_m|\vec{\ell}, \phi, \alpha_k) p(\vec{\ell}, \phi|\alpha_k)}{\sum_{\vec{\ell}'\in\mathbb{L}} \sum_{\phi'\in\Phi} p(M_m|\vec{\ell}', \phi', \alpha_k) p(\vec{\ell}', \phi'|\alpha_k)}, \tag{10}$$

which is the normalized likelihood function $p(M_m|\vec{\ell}, \phi, \alpha_k)$, weighted by the prior distribution $p(\vec{\ell}, \phi)$. We assume that $p(\vec{\ell}, \phi)$ is independent of the model α_k ~~and can be estimated simultaneously with the image~~. ~~Since the rotations and translations are independent~~, we have

that

$$p(\vec{\ell},\phi|\alpha_k) = p(\vec{\ell},\phi) = p(\vec{\ell})p(\phi) := \rho[\vec{\ell}]\frac{1}{K}, \tag{11}$$

where we assume the rotations are uniformly distributed in the set Φ from (7), and $\rho[\vec{\ell}]$ is the distribution of image locations within the patch. ~~This distribution depends on the density of the target images within the measurment~~. We can rewrite (8) as (up to an irrelevant constant)

$$Q(\alpha,\rho|\alpha_k,\rho_k) = \sum_{m=0}^{N_d-1}\sum_{\vec{\ell}\in\mathbb{L}}\sum_{\phi\in\Phi} p(M_m|\vec{\ell},\phi,\alpha_k)\rho_k[\vec{\ell}] \times \left(\log p(M_m|\vec{\ell},\phi,\alpha) + \log\rho[\vec{\ell}]\right), \tag{12}$$

where ~~$p(M_m|\vec{\ell},\phi,\alpha_k)$ and~~ $p(M_m|\vec{\ell},\phi,\alpha)$ are given by (9).

The M-step updates the image estimate and ρ by maximizing $Q(\alpha,\rho|\alpha_k,\rho_k)$ under the constraint that ρ lies on the simplex $\Delta_{2L\times 2L}$ for ρ:

$$\alpha_{k+1},\rho_{k+1} = \arg\max_{\alpha,\rho} Q(\alpha,\rho|\alpha_k,\rho_k) \text{ s.t. } \rho\in\Delta_{2L\times 2L}. \tag{13}$$

The constrained maximization of (13) can be achieved with the unconstrained maximization of the Lagrangian

$$\mathcal{L}(\alpha,\rho,\eta) = Q(\alpha,\rho|\alpha_k,\rho_k) + \eta\left(1-\sum_{\vec{\ell}\in\mathbb{L}}\rho[\vec{\ell}]\right), \tag{14}$$

where η denotes the Lagrange multiplier. We note that the constraints in (13) involve the inequalities that the priors are non-negative. Such constrained maximization in general cannot be achieved by maximizing the Lagrangian, for the inequalities might be violated. As we will see later, however, these inequalities are automatically satisfied at the computed maximum (or local maximum) of the Lagrangian, which justifies this approach.

Since $Q(\alpha,\rho|\alpha_k,\rho_k)$ is additively separable for α and ρ, we maximize $\mathcal{L}(\alpha,\rho,\eta)$ with respect to α and ρ separately. At the maximum of $\mathcal{L}(\alpha,\rho,\eta)$, we have

$$0 = \frac{\partial\mathcal{L}}{\partial(\alpha)_{\nu,q}} = \sum_{m=0}^{N_d-1}\sum_{\vec{\ell}\in\mathbb{L}}\sum_{\phi\in\Phi} p(M_m|\vec{\ell},\phi,\alpha_k)\rho_k[\vec{\ell}] \times \frac{\partial\log p(M_m|\vec{\ell},\phi,\alpha)}{\partial(\alpha)_{\nu,q}}, \tag{15}$$

resulting in a linear set of equations which is solved to achieve the updated α.

In order to update ρ, we maximize $\mathcal{L}(\alpha,\rho,\eta)$ with respect to ρ:

$$0 = \frac{\partial\mathcal{L}}{\rho[\vec{\ell}]} = \sum_{m=0}^{N_d-1}\sum_{\phi\in\Phi} p(M_m|\vec{\ell},\phi,\alpha_k)\rho_k[\vec{\ell}]\frac{1}{\rho[\vec{\ell}]} - \eta, \tag{16}$$

for $\vec{\ell}\in\mathbb{L}$. We thus obtain the update rule for ρ as

$$\rho[\vec{\ell}] = \frac{1}{\eta}\sum_{m=0}^{N_d-1}\sum_{\phi\in\Phi} p(M_m|\vec{\ell},\phi,\alpha_k)\rho_k[\vec{\ell}], \tag{17}$$

and we can immediately solve $\eta = N_d$ from the normalization $\sum_{\vec{\ell}\in\mathbb{L}}\rho[\vec{\ell}] = 1$.

3. NUMERICAL EXPERIMENTS

In this section, we present numerical results for the approximate EM described in Section 2. As a baseline, we compare the results to the recovery achieved using autocorrelation analysis, based on the framework (and code) of [4, 5]. To take the in-plane rotation symmetry into account, we measure the estimation error by

$$\text{relative error}_\alpha := \min_{\phi\in[0,2\pi)}\frac{\|\alpha^*-\alpha_\phi\|_2}{\|\alpha^*\|_2}, \tag{18}$$

where α^* is the true vector of expansion coefficients, and α_ϕ is the vector of coefficients of the estimated image, rotated by angle ϕ. In all experiments, the well-separated measurements were generated according to (1) with density $\gamma = 0.04$, where $p = \gamma\frac{N^2}{L^2}$. The target images are of diameter $L = 5$ pixels. Each entry of the target images was drawn i.i.d. from a uniform distribution on $[0,1]$. Then, each image was normalized such that $\|F\|_{\text{F}} = 10$, and expanded using its first 10 coefficients. We try to estimate the image from initial guesses that were drawn from the same distribution as the ground truth image ~~(the number of initial guesses changes varies experiments),~~ and $\gamma_{\text{init}} = 0.03$. We then calculate the estimation error of the image estimate whose final objective function is minimal. Figures 2 and 3 present the mean error over 40 trials. The code to reproduce all experiments is publicly available at `https://github.com/krshay/MTD-2D-EM`.

3.1. Recovery from noisy measurements

In [2, 5], a comparison of autocorrelation analysis to a naive method that detects and extracts the image occurrences, and then averages, was conducted. While this method works well in a high SNR environment, it fails in low SNR regimes—the focal point of this paper—and thus we do not conduct comparison experiments with the approximate EM framework. Nevertheless, we use the results of [5, Section III-E] as a threshold between high and low SNR environments—the naive method fails there for SNR < 10. Therefore, we report results for the recovery errors in estimating a target image from a measurement with SNR $= 50$ (a high SNR environment) and SNR $= 2$ (a low SNR environment), using approximate EM and autocorrelation analysis. We consider a measurement of size $N = 5000$ pixels, and use $K = 16$. The SNRs are visualized in Figure 1. For SNR $= 50$, the relative error is TBD using approximate EM, and TBD using autocorrelation analysis. For SNR $= 2$, the relative error is TBD using approximate EM, and 0.073 using autocorrelation analysis. Evidently, approximate EM outperforms autocorrelation analysis in both SNR regimes.

3.2. Recovery error as a function of the measurement size

Figure 2 presents recovery error and running time as a function of the measurement size, with SNR $= 5$, $K = 16$, and 5 random initial guesses for α. Using approximate EM, the error decays as $1/\sqrt{N^2}$. The same trend is visible also for autocorrelation analysis for sufficiently large measurements. ~~This is the same estimation rate as if the translations and rotations were known (that is, the estimation rate of averaging over i.i.d. Gaussian variables)~~. We achieve a significant improvement in recovery accuracy using approximate EM, although the search space in the angular direction is coarsely sampled. However, the computation time using the method is significantly larger, and grows linearly with the measurement size.

3.3. Recovery error as a function of discretization of rotations

Figure 3 presents recovery error and running time as a function of K, the size of the search space in the angular direction ~~in the approximate EM~~, for measurements with SNR $= 5$, $N = 1500$ pixels, and one random initial guess for α. The results are compared to the results using autocorrelation analysis. Remarkably, even when the EM

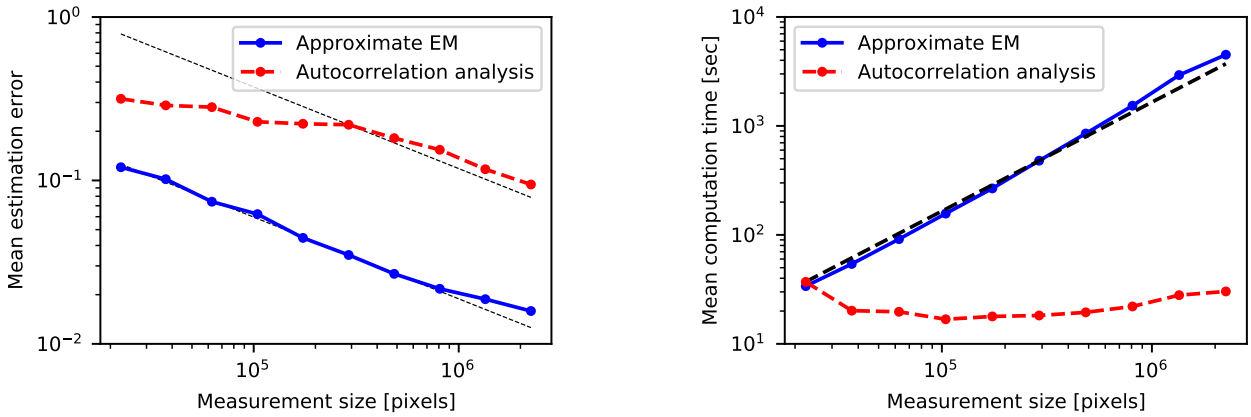


Fig. 2: The mean estimation error (left) and running time (right) ~~recovering the target image~~ F, as a function of measurement size, by: (a) approximate EM; (b) autocorrelation analysis. For the estimation error, the black dashed lines illustrate a slope of $-1/2$, as predicted by the law of large numbers. For the computation time, the black dashed line illustrates a slope of 1, which implies a linear increase in computation time, as the number of patches $N_d = N^2/L^2$ grows linearly in N^2.

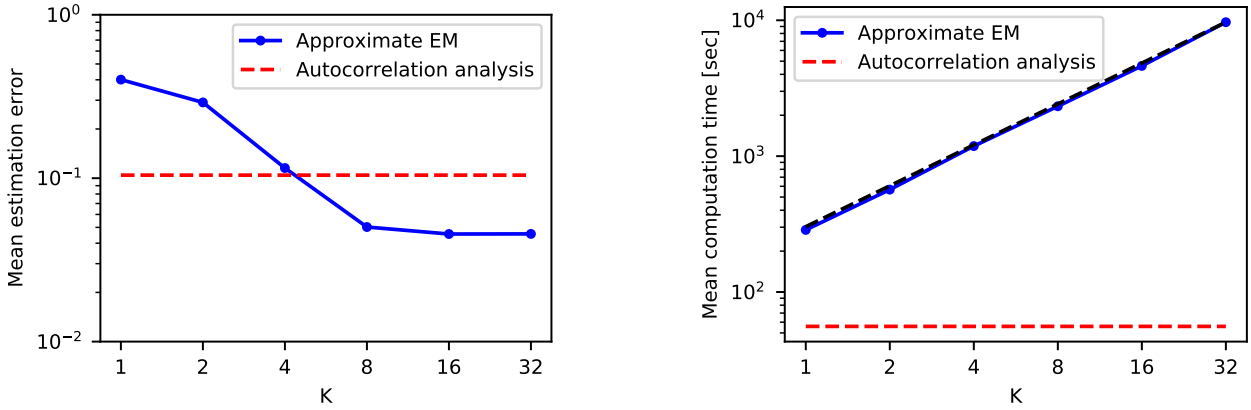


Fig. 3: The mean estimation error (left) and time (right) of recovering the target image F, as a function of K, the size of the search space in the angular direction, by approximate EM. The recovery error and running time using autocorrelation analysis are marked by ~~the~~ dashed horizontal red lines. For the running time, the black dashed line illustrates a slope of 1, which implies a linear increase in computation time, as the number of computations per patch depends linearly on K.

searches over only 4 rotations, we achieve estimation errors similar to the estimation errors of autocorrelation analysis, which takes all possible rotations into account. As expected, the computation time grows linearly with the parameter K.

4. CONCLUSION

This paper is motivated by the effort of reconstructing small 3-D molecular structures using cryo-EM, below the current detection limit. The main contribution of this paper is in introducing an approximate EM scheme for the two-dimensional ~~well separated~~ MTD problem, and comparing it numerically to ~~the previously studied~~ autocorrelation analysis. The numerical experiments suggest an improvement in estimation accuracy, at the cost of ~~an increase in required~~ computational time. As Figure 3 shows, the parameter K provides an accuracy-running time trade-off. A possible improvement is increasing the resolution of the search space as the iterations progress [2]; this is a standard procedure in current cryo-EM software packages [13, 14].

Our ultimate goal is developing an approximate EM scheme for recovering small molecular structures using cryo-EM [7]. Although it significantly outperforms autocorrelation analysis in terms of accuracy, a major drawback of the approximate EM method is the computation time. Therefore, in order to achieve a computationally tractable algorithm for the 3-D case of the MTD problem, parallel processing and randomized algorithms, such as stochastic or online EM [22, 23, 24, 25, 26], must be utilized. In addition, working on a lower dimensional space, see for instance SubspaceEM [27], might reduce the computational burden. Furthermore, in the future, EM may be replaced by more intricate techniques, such as variational inference [28] or variational auto-encoders ~~(VAEs)~~ [29]. Finally, adding a prior on the target image should improve the robustness and accuracy of the scheme, at the cost of introducing model bias ~~[30, 31, 32].~~

5. REFERENCES

[1] Tamir Bendory, Nicolas Boumal, William Leeb, Eitan Levin, and Amit Singer, "Multi-target detection with application to cryo-electron microscopy," *Inverse Problems*, vol. 35, no. 10, pp. 104003, 2019.

[2] Ti-Yen Lan, Tamir Bendory, Nicolas Boumal, and Amit Singer, "Multi-target detection with an arbitrary spacing distribution," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1589–1601, 2020.

[3] Nicholas F Marshall, Ti-Yen Lan, Tamir Bendory, and Amit Singer, "Image recovery from rotational and translational invariants," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 5780–5784.

[4] Tamir Bendory, Ti-Yen Lan, Nicholas F Marshall, Iris Rukshin, and Amit Singer, "Multi-target detection with rotations," *arXiv preprint arXiv:2101.07709*, 2021.

[5] Shay Kreymer and Tamir Bendory, "Two-dimensional multi-target detection: an autocorrelation analysis approach," *arXiv preprint arXiv:2105.06765*, 2021.

[6] Ye'Ela Shalit, Ran Weber, Asaf Abas, Shay Kreymer, and Tamir Bendory, "Generalized autocorrelation analysis for multi-target detection," *arXiv preprint arXiv:2109.11813*, 2021.

[7] Tamir Bendory, Nicolas Boumal, William Leeb, Eitan Levin, and Amit Singer, "Toward single particle reconstruction without particle picking: breaking the detection limit," *arXiv preprint arXiv:1810.00226*, 2018.

[8] Richard Henderson, "The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules," *Quarterly Reviews of Biophysics*, vol. 28, no. 2, pp. 171–193, 1995.

[9] Eva Nogales, "The development of cryo-EM into a mainstream structural biology technique," *Nature methods*, vol. 13, no. 1, pp. 24–27, 2016.

[10] Xiao-Chen Bai, Greg McMullan, and Sjors HW Scheres, "How cryo-EM is revolutionizing structural biology," *Trends in Biochemical Sciences*, vol. 40, no. 1, pp. 49–57, 2015.

[11] Joachim Frank, *Three-dimensional electron microscopy of macromolecular assemblies: visualization of biological molecules in their native state*, Oxford University Press, 2006.

[12] Tamir Bendory, Alberto Bartesaghi, and Amit Singer, "Single-particle cryo-electron microscopy: Mathematical theory, computational challenges, and opportunities," *IEEE Signal Processing Magazine*, vol. 37, no. 2, pp. 58–76, 2020.

[13] Sjors HW Scheres, "RELION: implementation of a Bayesian approach to cryo-EM structure determination," *Journal of Structural Biology*, vol. 180, no. 3, pp. 519–530, 2012.

[14] Ali Punjani, John L Rubinstein, David J Fleet, and Marcus A Brubaker, "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination," *Nature methods*, vol. 14, no. 3, pp. 290–296, 2017.

[15] Cecilia Aguerrebere, Mauricio Delbracio, Alberto Bartesaghi, and Guillermo Sapiro, "Fundamental limits in multi-image alignment," *IEEE Transactions on Signal Processing*, vol. 64, no. 21, pp. 5707–5722, 2016.

[16] Edoardo D'Imprima and Werner Kühlbrandt, "Current limitations to high-resolution structure determination by single-particle cryoEM," *Quarterly Reviews of Biophysics*, vol. 54, 2021.

[17] Arthur P Dempster, Nan M Laird, and Donald B Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[18] Tamir Bendory, Nicolas Boumal, Chao Ma, Zhizhen Zhao, and Amit Singer, "Bispectrum inversion with application to multireference alignment," *IEEE Transactions on signal processing*, vol. 66, no. 4, pp. 1037–1050, 2017.

[19] Fred J Sigworth, "A maximum-likelihood approach to single-particle image refinement," *Journal of structural biology*, vol. 122, no. 3, pp. 328–339, 1998.

[20] Chao Ma, Tamir Bendory, Nicolas Boumal, Fred Sigworth, and Amit Singer, "Heterogeneous multireference alignment for images with application to 2D classification in single particle reconstruction," *IEEE Transactions on Image Processing*, vol. 29, pp. 1699–1710, 2019.

[21] Anya Katsevich and Afonso Bandeira, "Likelihood maximization and moment matching in low SNR gaussian mixture models," *arXiv preprint arXiv:2006.15202*, 2020.

[22] Søren Feodor Nielsen, "The stochastic EM algorithm: estimation and asymptotic results," *Bernoulli*, pp. 457–489, 2000.

[23] Jianfei Chen, Jun Zhu, Yee Whye Teh, and Tong Zhang, "Stochastic expectation maximization with variance reduction," in *NeurIPS*, 2018, pp. 7978–7988.

[24] Percy Liang and Dan Klein, "Online EM for unsupervised models," in *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 2009, pp. 611–619.

[25] Olivier Cappé and Eric Moulines, "On-line expectation–maximization algorithm for latent data models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 71, no. 3, pp. 593–613, 2009.

[26] Olivier Cappé, "Online EM algorithm for hidden Markov models," *Journal of Computational and Graphical Statistics*, vol. 20, no. 3, pp. 728–749, 2011.

[27] Nicha C Dvornek, Fred J Sigworth, and Hemant D Tagare, "SubspaceEM: A fast maximum-a-posteriori algorithm for cryo-EM single particle reconstruction," *Journal of Structural Biology*, vol. 190, no. 2, pp. 200–214, 2015.

[28] David M Blei, Alp Kucukelbir, and Jon D McAuliffe, "Variational inference: A review for statisticians," *Journal of the American Statistical Association*, vol. 112, no. 518, pp. 859–877, 2017.

[29] Dan Rosenbaum, Marta Garnelo, Michal Zielinski, Charlie Beattie, Ellen Clancy, Andrea Huber, Pushmeet Kohli, Andrew W Senior, John Jumper, Carl Doersch, et al., "Inferring a continuous distribution of atom coordinates from cryo-EM images using VAEs," *arXiv preprint arXiv:2106.14108*, 2021.

[30] Richard Henderson, "Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise," *Proceedings of the National Academy of Sciences*, vol. 110, no. 45, pp. 18037–18041, 2013.

[31] Peter B Rosenthal and John L Rubinstein, "Validating maps from single particle electron cryomicroscopy," *Current Opinion in Structural Biology*, vol. 34, pp. 135–144, 2015.

[32] Sjors HW Scheres and Shaoxia Chen, "Prevention of overfitting in cryo-EM structure determination," *Nature Methods*, vol. 9, no. 9, pp. 853–854, 2012.