Automated target segmentation and real space fast alignment methods for high-throughput classification and averaging of crowded cryo-electron subtomograms — Supplementary Document

Min Xu and Frank Alber*

Molecular and Computational Biology, University of Southern California, Los Angeles, CA 90089, USA.

In all the tables in this supplementary document, each row corresponds to a specific Signal-to-noise ratio (SNR) level, and each column corresponds to a tilt angle range.

1 Model based clustering

To test the effectiveness of our model based clustering, we carry out the following simulation study. We first define two 2D Gaussian distribution with mean (25,50) and (75,50), and $\sigma^2 = 10$. We then sample 100 points from each distribution, representing a point cluster. We then randomly sample 1000 uniformly distributed points as noise, which will disturb the clustering process. Then we compose an image (of size 100 by 100 voxels) whose intensity is proportional to the density of the mixture of the two Gaussian distributions. We then predict clusters of these 1200 points using the standard model based clustering method and our modified model based clustering method, with weight proportional to the image intensity of corresponding point locations. We test the cluster number from 1 to 5 and use Bayesian Information Criterion to select the optimal set of clusters. We measure the clustering performance by the consistency between predicted clusters and true clusters:

$$\frac{|S_{\text{Pred}} \cap S_{\text{True}}|}{|S_{\text{Pred}} \cup S_{\text{True}}|}$$

, where S_{True} is the true set of points of a cluster, and S_{Pred} is the predicted cluster that maximally matches S_{True}.

We repeat the above randomized test for 1000 times, on average, the cluster consistency obtained from our extended model based clustering is 19.6% higher than the consistency obtained by the standard model based clustering, showing that the inclusion of image intensity information into the cluster model significantly increases the clustering accuracy.

*to whom correspondence should be addressed

2 Tests on the phantom model and ribosome complex

2.1 Un-crowded case

2.1.1 Center of mass and alignment performance

Table 1: Phantom model: median of center of mass invariance $\{d_{\text{ccm\_inverse},ij}\}$ across 100 tests

SNR	±70°	±60°	±50°
0.05	1.3345	1.3105	1.3421
0.04	1.3621	1.2261	1.5499
0.03	1.1736	1.4115	1.5609
0.02	1.0426	1.6353	1.5535
0.01	0.9750	1.1361	1.3694
0.008	0.8867	1.0874	1.2698
0.006	0.8843	1.4320	1.3819
0.005	1.0109	1.5544	1.4633
0.004	1.0584	1.2901	1.3744
0.003	1.2036	1.3767	1.8786
0.002	1.2881	1.7993	2.5086
0.001	5.0891	5.6215	6.1297

Table 2: Phantom model: median of rotational alignment error $d_{\text{rot\_err}}$ across 100 tests. Left of comma, our current method. Right of comma: our previous alignment method [2]. The green regions corresponds to the distortion levels at which both method can successfully perform alignments (at a cutoff of 0.5). The yellow regions correspond to the distortion levels at which our new method outperforms our previous method. In other words, it aligns correctly the subtomograms while our previous method [2] fails to do so. The gray region corresponds to the distortion levels at which both methods fail to align correctly.

SNR	±70°	±60°	±50°
0.05	0.1185, 0.0591	0.1485, 0.0593	0.1568, 0.0623
0.04	0.1374, 0.0623	0.1552, 0.0646	0.1655, 0.0692
0.03	0.1365, 0.0677	0.1557, 0.0729	0.1533, 0.0803
0.02	0.1181, 0.0870	0.1715, 0.0936	0.1819, 0.1163
0.01	0.1207, 0.1556	0.1400, 0.1634	0.1563, 0.1961
0.008	0.1258, 0.2161	0.1328, 0.2196	0.1552, 0.2669
0.006	0.1238, 0.3536	0.1544, 0.3050	0.1904, 2.1153
0.005	0.1377, 0.5027	0.1957, 0.7283	0.1932, 2.4784
0.004	0.1661, 1.9331	0.1813, 2.3974	0.2598, 2.5179
0.003	0.1706, 2.4709	0.1809, 2.5352	1.3786, 2.4758
0.002	0.2272, 2.5547	0.5248, 2.5127	2.1547, 2.4719
0.001	2.6797, 2.5559	2.5915, 2.5390	2.5725, 2.4170

Table 3: Phantom model: correlation between center of mass invariance $\{d_{\text{ccm\_inverse},ij}\}$ and rotation alignment errors $d_{\text{rot\_err}}$

SNR	$\pm 70°$	$\pm 60°$	$\pm 50°$
0.05	0.3696	0.6111	0.5716
0.04	0.3682	0.5766	0.5305
0.03	0.4682	0.4502	0.5157
0.02	0.4065	0.4719	0.5169
0.01	0.1287	0.2814	0.5268
0.008	0.1758	0.2526	0.3870
0.006	0.3150	0.4333	0.3574
0.005	0.1901	0.3956	0.3526
0.004	0.3654	0.4317	0.3762
0.003	0.4779	0.4177	0.4560
0.002	0.2982	0.3164	0.3542
0.001	0.4015	0.2451	0.2700

Table 4: Ribosome: median of center of mass invariance $\{d_{\text{ccm\_inverse},ij}\}$ across 100 tests

SNR	$\pm 70°$	$\pm 60°$	$\pm 50°$
0.05	1.2871	1.7094	1.7426
0.04	1.4306	1.7124	1.9367
0.03	1.4384	1.8253	2.1951
0.02	1.3641	1.6818	2.2620
0.01	1.0681	1.6490	2.7085
0.008	1.1000	1.3114	2.4627
0.006	1.4384	2.1349	2.6077
0.005	1.3628	3.3202	2.3993
0.004	1.2351	1.7153	2.0158
0.003	1.5100	2.1521	2.3435
0.002	1.8740	3.2342	3.5714
0.001	8.5367	8.6766	9.1065

Table 5: Ribosome: median of rotational alignment error $d_{\text{rot\_err}}$ across 100 tests. Left of comma, our current method. Right of comma: our previous alignment method [2]. The green regions corresponds to the distortion levels at which both method can successfully perform alignments (at a cutoff of 0.5). The yellow regions correspond to the distortion levels at which our new method outperforms our previous method. In other words, it aligns correctly the subtomograms while our previous method [2] fails to do so. The gray region corresponds to the distortion levels at which both methods fail to align correctly.

SNR	±70°	±60°	±50°
0.05	0.2216, 0.0811	0.2411, 0.0864	0.2206, 0.0908
0.04	0.2504, 0.0936	0.2161, 0.0969	0.2759, 0.1005
0.03	0.2201, 0.1026	0.2326, 0.1207	0.2691, 0.1289
0.02	0.2416, 0.1259	0.2631, 0.1738	0.3224, 0.1899
0.01	0.2299, 0.2549	0.3090, 0.2833	1.6313, 1.6327
0.008	0.2590, 0.3349	0.2802, 0.3588	1.8159, 2.2497
0.006	0.3116, 0.4867	0.4146, 0.5001	2.2657, 2.4514
0.005	0.3551, 0.5449	1.0790, 2.0474	2.2544, 2.4254
0.004	0.3625, 2.1660	0.4467, 2.3535	2.0954, 2.5018
0.003	0.4564, 2.5217	1.7761, 2.5103	2.1758, 2.5235
0.002	0.8436, 2.6238	2.0817, 2.5658	2.4681, 2.5732
0.001	2.5453, 2.6724	2.4244, 2.5869	2.5777, 2.5377

Table 6: Ribosome: correlation between center of mass invariance $\{d_{\text{ccm\_inverse},ij}\}$ and rotation alignment errors $d_{\text{rot\_err}}$

SNR	±70°	±60°	±50°
0.05	0.3351	0.4131	0.3639
0.04	0.4342	0.3837	0.3230
0.03	0.5538	0.5042	0.3548
0.02	0.4323	0.4709	0.5952
0.01	0.7703	0.7557	0.6526
0.008	0.7695	0.7453	0.5887
0.006	0.7667	0.7272	0.5552
0.005	0.7521	0.7076	0.4797
0.004	0.6721	0.6137	0.4509
0.003	0.6464	0.6690	0.4290
0.002	0.5405	0.4879	0.1725
0.001	0.1601	0.2000	-0.1245

2.2 Crowded case

2.2.1 Detection of target complex region

Table 7: Phantom model: median true-positive rate $o_{\text{true\_positive\_rate}}$ across 100 tests

SNR	±70°	±60°	±50°
0.05	1.0000	1.0000	1.0000
0.04	1.0000	1.0000	1.0000
0.03	1.0000	1.0000	1.0000
0.02	1.0000	1.0000	1.0000
0.01	1.0000	1.0000	1.0000
0.008	1.0000	1.0000	1.0000
0.006	1.0000	1.0000	1.0000
0.005	1.0000	1.0000	1.0000
0.004	1.0000	1.0000	1.0000
0.003	1.0000	1.0000	0.9990
0.002	1.0000	0.9996	0.9969
0.001	0.9957	0.9936	0.9936

Table 8: Phantom model: median false-positive rate $o_{\text{false\_positive\_rate}}$ across 100 tests

SNR	±70°	±60°	±50°
0.05	0.0000	0.0000	0.0000
0.04	0.0000	0.0000	0.0000
0.03	0.0000	0.0000	0.0000
0.02	0.0000	0.0000	0.0000
0.01	0.0000	0.0000	0.0000
0.008	0.0000	0.0000	0.0000
0.006	0.0000	0.0000	0.0000
0.005	0.0000	0.0000	0.0000
0.004	0.0000	0.0000	0.0000
0.003	0.0000	0.0000	0.0000
0.002	0.0000	0.0000	0.0000
0.001	0.0000	0.0001	0.0001

Table 9: Phantom model: median $o_{\text{surrounding structure}}$ across 100 tests

SNR	±70°	±60°	±50°
0.05	0.0000	0.0001	0.0018
0.04	0.0000	0.0001	0.0013
0.03	0.0000	0.0000	0.0005
0.02	0.0000	0.0000	0.0005
0.01	0.0000	0.0000	0.0002
0.008	0.0000	0.0000	0.0000
0.006	0.0000	0.0000	0.0000
0.005	0.0000	0.0000	0.0000
0.004	0.0000	0.0000	0.0000
0.003	0.0000	0.0000	0.0000
0.002	0.0000	0.0000	0.0055
0.001	0.1592	0.2138	0.3409

Table 10: Ribosome: median true-positive rate $o_{\text{true\_positive\_rate}}$ across 100 tests

SNR	±70°	±60°	±50°
0.05	0.9916	0.9930	0.9953
0.04	0.9915	0.9933	0.9954
0.03	0.9915	0.9931	0.9939
0.02	0.9894	0.9896	0.9927
0.01	0.9817	0.9804	0.9658
0.008	0.9768	0.9768	0.9579
0.006	0.9654	0.9555	0.9433
0.005	0.9594	0.9440	0.9254
0.004	0.9476	0.9192	0.9161
0.003	0.9368	0.9225	0.9147
0.002	0.9257	0.9322	0.9272
0.001	0.9386	0.9361	0.9089

Table 11: Ribosome: median false-positive rate $o_{\text{false\_positive\_rate}}$ across 100 tests

SNR	±70°	±60°	±50°
0.05	0.0003	0.0003	0.0002
0.04	0.0003	0.0003	0.0002
0.03	0.0003	0.0003	0.0002
0.02	0.0004	0.0004	0.0003
0.01	0.0007	0.0007	0.0013
0.008	0.0009	0.0009	0.0016
0.006	0.0013	0.0017	0.0022
0.005	0.0015	0.0021	0.0028
0.004	0.0020	0.0031	0.0032
0.003	0.0024	0.0030	0.0033
0.002	0.0028	0.0026	0.0028
0.001	0.0023	0.0024	0.0035

Table 12: Ribosome: median $o_{\text{surrounding structure}}$ across 100 tests

SNR	±70°	±60°	±50°
0.05	0.0000	0.0000	0.0007
0.04	0.0000	0.0000	0.0032
0.03	0.0000	0.0000	0.0012
0.02	0.0000	0.0000	0.0011
0.01	0.0000	0.0000	0.0000
0.008	0.0000	0.0000	0.0000
0.006	0.0000	0.0000	0.0000
0.005	0.0000	0.0000	0.0000
0.004	0.0000	0.0000	0.0000
0.003	0.0000	0.0000	0.0000
0.002	0.0000	0.0049	0.0083
0.001	0.1981	0.2488	0.3419

2.2.2 Center of mass and alignment performance

Table 13: Phantom model: median of center of mass invariance $\{d_{\text{ccm\_inverse},ij}\}$ across 100 tests

SNR	±70°	±60°	±50°
0.05	0.9084	1.1582	1.2600
0.04	1.2680	1.2279	1.4276
0.03	0.7440	0.9805	1.6729
0.02	1.0028	1.1295	1.6236
0.01	1.0158	1.3856	1.6098
0.008	1.2107	1.4643	1.4827
0.006	1.3162	1.9139	1.7148
0.005	1.3607	1.7308	1.9660
0.004	1.7045	2.0921	2.3869
0.003	2.1125	2.5776	3.4607
0.002	3.8180	7.0615	8.1721
0.001	11.2168	10.9190	11.9960

Table 14: Phantom model: median of rotational alignment error $d_{\text{rot\_err}}$ across 100 tests. Left of comma, our current method. Right of comma: our previous alignment method [2]. The green regions corresponds to the distortion levels at which both method can successfully perform alignments (at a cutoff of 0.5). The yellow regions correspond to the distortion levels at which our new method outperforms our previous method. In other words, it aligns correctly the subtomograms while our previous method [2] fails to do so. The gray region corresponds to the distortion levels at which both methods fail to align correctly.

SNR	±70°	±60°	±50°
0.05	0.1075, 0.2247	0.1298, 0.3288	0.1238, 0.4652
0.04	0.1297, 0.2586	0.1357, 0.4781	0.1365, 0.8521
0.03	0.1154, 0.4141	0.1163, 0.8702	0.1462, 1.8636
0.02	0.1543, 1.2313	0.1271, 1.6796	0.1861, 2.2714
0.01	0.1739, 2.4951	0.2044, 2.4778	0.2117, 2.4633
0.008	0.1716, 2.5874	0.2949, 2.5529	0.5589, 2.4090
0.006	0.2245, 2.5159	0.3462, 2.5591	2.0082, 2.4407
0.005	0.2626, 2.6235	0.4178, 2.5958	2.1376, 2.4407
0.004	0.4316, 2.5734	2.3065, 2.5391	2.4781, 2.4684
0.003	2.0515, 2.5390	2.5177, 2.5391	2.6469, 2.4415
0.002	2.6056, 2.4849	2.5025, 2.5102	2.5470, 2.4282
0.001	2.6234, 2.5434	2.5458, 2.5763	2.4472, 2.5192

Table 15: Phantom model: correlation between center of mass invariance $\{d_{\text{ccm\_inverse},ij}\}$ and rotation alignment errors $d_{\text{rot\_err}}$

SNR	±70°	±60°	±50°
0.05	0.5430	0.6484	0.3516
0.04	0.4459	0.6178	0.4814
0.03	0.5154	0.4317	0.4909
0.02	0.6248	0.3974	0.4655
0.01	0.3880	0.3776	0.3714
0.008	0.1942	0.3227	0.3441
0.006	0.3473	0.4423	0.2830
0.005	0.3016	0.3083	0.2155
0.004	0.2064	0.3116	0.1997
0.003	0.3005	0.2702	0.2058
0.002	0.2500	0.1231	0.0314
0.001	0.0855	0.0874	0.2706

Table 16: Ribosome: median of center of mass invariance $\{d_{\text{ccm\_inverse},ij}\}$ across 100 tests

SNR	±70°	±60°	±50°
0.05	0.6167	0.9498	1.3683
0.04	0.7073	1.0744	1.4591
0.03	0.7793	1.0934	1.9163
0.02	0.7147	1.1100	1.6862
0.01	1.1672	1.2653	1.5640
0.008	1.0736	1.5328	1.6158
0.006	1.5372	2.0610	2.0727
0.005	1.6844	2.2711	2.6776
0.004	1.7001	2.3601	3.7305
0.003	2.5555	4.5776	5.0098
0.002	6.4040	6.9970	6.8650
0.001	9.4265	10.1695	11.2316

Table 17: Ribosome: median of rotational alignment error $d_{\text{rot\_err}}$ across 100 tests. Left of comma, our current method. Right of comma: our previous alignment method [2]. The green regions corresponds to the distortion levels at which both method can successfully perform alignments (at a cutoff of 0.5). The yellow regions correspond to the distortion levels at which our new method outperforms our previous method. In other words, it aligns correctly the subtomograms while our previous method [2] fails to do so. The gray region corresponds to the distortion levels at which both methods fail to align correctly.

SNR	±70°	±60°	±50°
0.05	0.1801, 0.4468	0.1881, 0.6214	0.2288, 2.4884
0.04	0.2135, 0.5048	0.2105, 1.0579	0.2284, 2.3461
0.03	0.2047, 0.9570	0.2380, 1.7144	0.3168, 2.4549
0.02	0.2136, 1.8366	0.2616, 2.3060	0.3020, 2.5626
0.01	0.2861, 2.5569	0.3830, 2.5142	1.9574, 2.4949
0.008	0.3150, 2.6737	0.5629, 2.5241	2.2035, 2.4920
0.006	0.4625, 2.6163	2.1803, 2.6171	2.4087, 2.4211
0.005	0.6754, 2.6638	2.1371, 2.5906	2.4455, 2.4809
0.004	1.8862, 2.6056	2.5807, 2.5923	2.4701, 2.5070
0.003	2.4155, 2.5772	2.5253, 2.5532	2.5695, 2.4809
0.002	2.5481, 2.6178	2.6412, 2.5260	2.5039, 2.4836
0.001	2.6113, 2.6394	2.5025, 2.5199	2.5412, 2.4244

Table 18: Ribosome: correlation between center of mass invariance $\{d_{\text{ccm\_inverse},ij}\}$ and rotation alignment errors $d_{\text{rot\_err}}$

SNR	±70°	±60°	±50°
0.05	0.8174	0.1090	0.3866
0.04	0.5255	0.5975	0.6436
0.03	0.7768	0.4453	0.5100
0.02	0.1697	0.7585	0.4023
0.01	0.7488	0.5553	0.3341
0.008	0.6960	0.4502	0.3833
0.006	0.6331	0.4064	0.3433
0.005	0.6135	0.4358	0.1935
0.004	0.3897	0.4131	0.1785
0.003	0.2976	0.0449	0.0933
0.002	0.1680	0.1612	-0.0555
0.001	0.1264	0.0624	-0.0762

2.2.3 Target complex region detection and alignment performance respect to level of crowdedness

In this section, we report the target region detection and alignment performance respect to crowdedness levels at certain distortion levels. We use

$$\frac{\text{the volume of the extra structures}}{\text{the volume not occupied by the target structure to segment}}$$

in the ground truth density map as a measure of level of crowdedness.

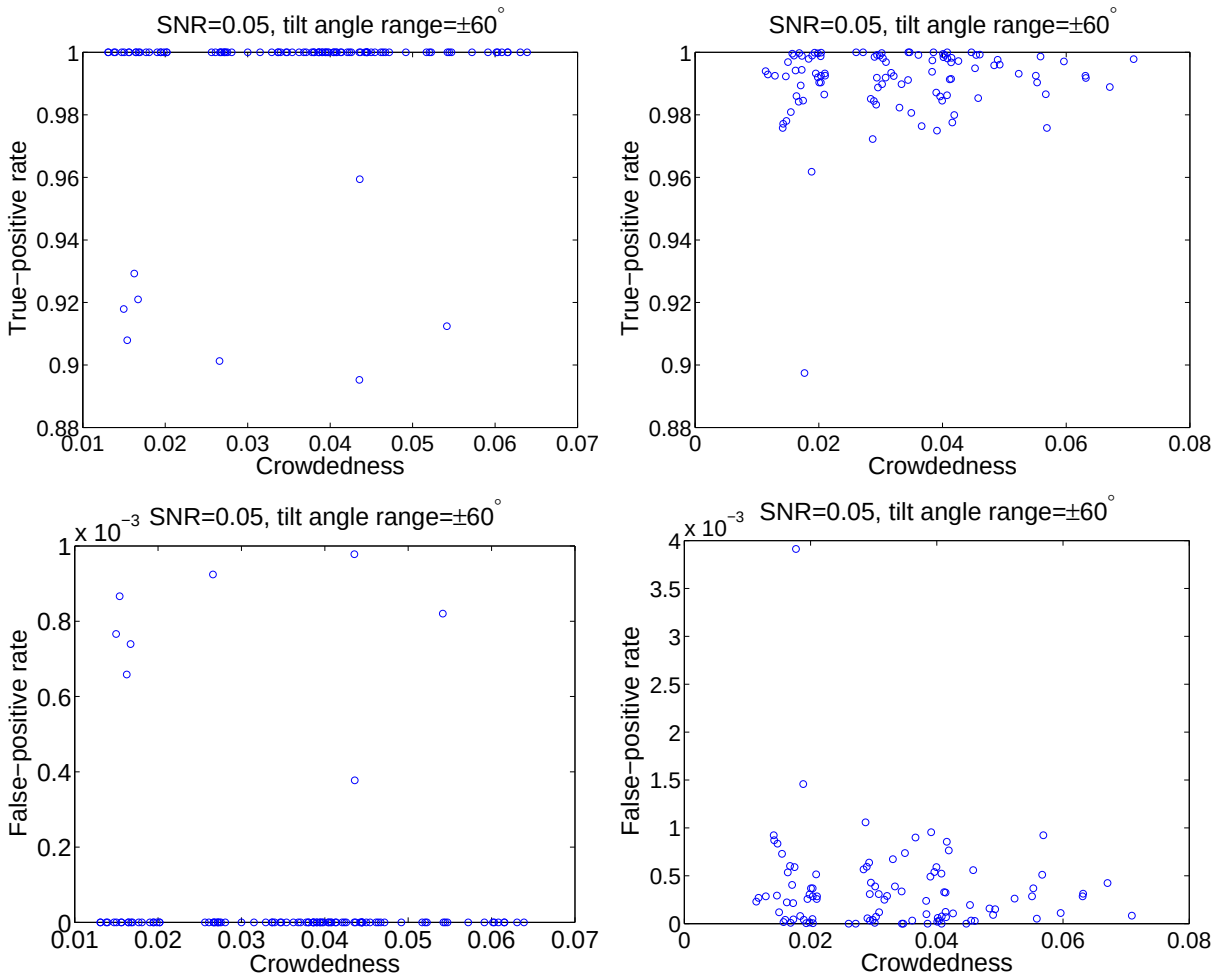

Figure 1: True-positive rate $o_{\text{true\_positive\_rate}}$ and false-positive rate $o_{\text{false\_positive\_rate}}$ of target complex detection respect to crowdedness. Subtomograms are simulated at SNR=0.05, tilt angle range ±60°. Left colomn: phantom model. Right column, Ribosome. Upper row: true-positive rate $o_{\text{true\_positive\_rate}}$. Lower row: false-positive rate $o_{\text{false\_positive\_rate}}$.

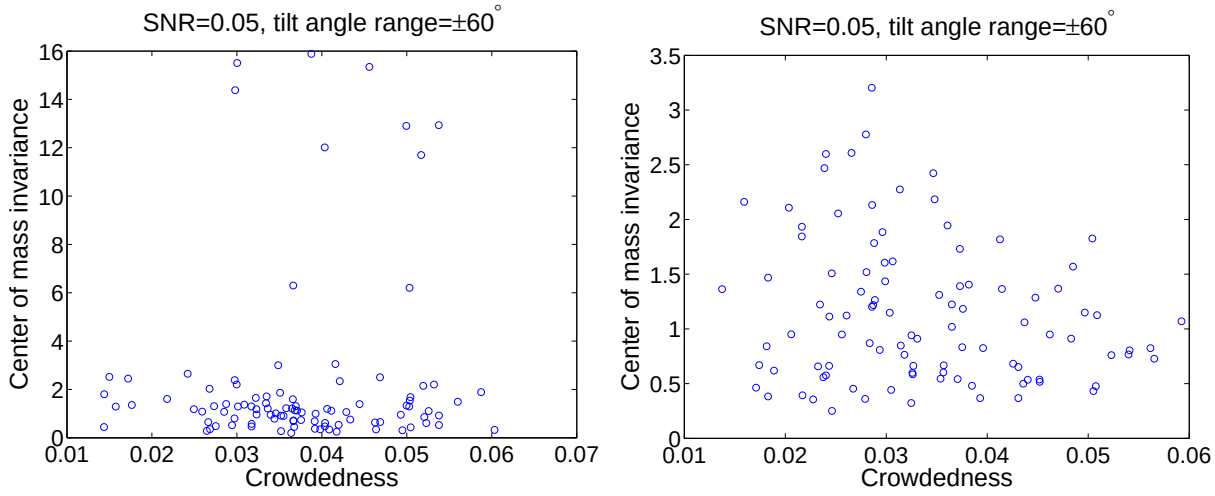

Figure 2: Center of mass invariance $\{d_{\text{ccm\_inverse},ij}\}$ respect to average crowdedness of two subtomograms to be aligned. Subtomograms are simulated at SNR=0.05, tilt angle range ±60°. Left colomn: phantom model. Right column, Ribosome.

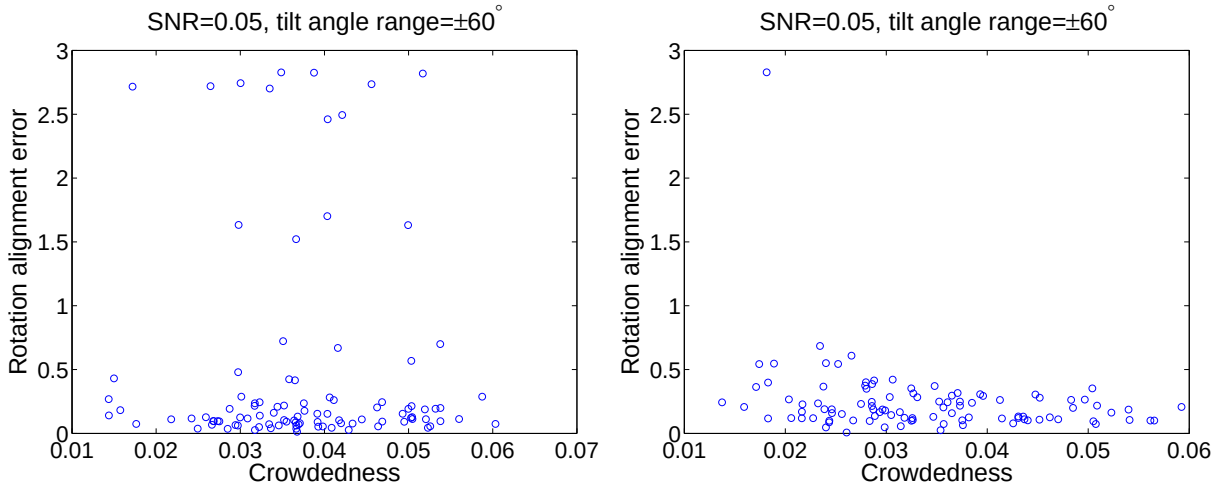

Figure 3: Rotational alignment errors $d_{\text{rot\_err}}$ respect to average crowdedness of two subtomograms to be aligned. Subtomograms are simulated at SNR=0.05, tilt angle range ±60°. Left colomn: phantom model. Right column, Ribosome.

2.3 Tests on the crowded subtomograms of GroEL, Carboamoyl phosphate synthase, Octameric enolase and ClpP

We further tested our method for extraction and aligning target complexes in simulated crowed subtomograms. These complexes were used for benchmarking of our previously proposed alignment method [2]. Because of the multiple solutions brought by rotation symmetry, the measurement of rotation angle error is not straightforward. Instead, we compare the

alignment score resulted from rotation and consequent translation alignments. The score is defined in Equation 2, which is equivalent to the scores in Equation 8 of [1] and Equation 6 of [2].

Table 19: GroEL (PDBID: 1KP8): median of rotational alignment score across 100 tests. Left of comma, our current method. Right of comma: our previous alignment method [2].

SNR	±70°	±60°	±50°
0.05	0.7537, 0.2390	0.7224, 0.2592	0.7034, 0.2865
0.04	0.7284, 0.2106	0.7107, 0.2294	0.6689, 0.2514
0.03	0.6990, 0.1711	0.6709, 0.1868	0.6550, 0.2080
0.02	0.6446, 0.1269	0.6204, 0.1346	0.5831, 0.1579
0.01	0.5295, 0.0714	0.4977, 0.0759	0.5295, 0.0940
0.008	0.4813, 0.0601	0.4604, 0.0663	0.4942, 0.0899
0.006	0.4257, 0.0509	0.4109, 0.0581	0.4481, 0.0874
0.005	0.3971, 0.0468	0.3756, 0.0558	0.4219, 0.0871
0.004	0.3631, 0.0446	0.3563, 0.0552	0.3842, 0.0869
0.003	0.3028, 0.0434	0.3028, 0.0546	0.3212, 0.0865
0.002	0.2033, 0.0432	0.1697, 0.0549	0.1882, 0.0874
0.001	0.0605, 0.0433	0.0642, 0.0553	0.0932, 0.0881

Table 20: Carboamoyl phosphate synthase (PDBID: 1BXR): median of rotational alignment score across 100 tests. Left of comma, our current method. Right of comma: our previous alignment method [2].

SNR	±70°	±60°	±50°
0.05	0.6370, 0.1672	0.5916, 0.1708	0.5459, 0.1877
0.04	0.5372, 0.1431	0.5499, 0.1444	0.5180, 0.1628
0.03	0.4482, 0.1166	0.4677, 0.1170	0.4578, 0.1341
0.02	0.3813, 0.0845	0.3879, 0.0869	0.4104, 0.1076
0.01	0.3117, 0.0510	0.3297, 0.0583	0.3601, 0.0890
0.008	0.2993, 0.0467	0.3145, 0.0557	0.3369, 0.0879
0.006	0.2681, 0.0441	0.2992, 0.0545	0.3053, 0.0888
0.005	0.2553, 0.0436	0.2775, 0.0545	0.3027, 0.0880
0.004	0.2384, 0.0432	0.2638, 0.0545	0.2832, 0.0881
0.003	0.2063, 0.0432	0.2239, 0.0545	0.2702, 0.0878
0.002	0.1754, 0.0430	0.1607, 0.0544	0.1733, 0.0888
0.001	0.0570, 0.0431	0.0683, 0.0543	0.0953, 0.0891

Table 21: Octameric enolase (PDBID: 1W6T): median of rotational alignment score across 100 tests. Left of comma, our current method. Right of comma: our previous alignment method [2].

SNR	±70°	±60°	±50°
0.05	0.6204, 0.1582	0.5699, 0.1687	0.5330, 0.1900
0.04	0.6268, 0.1395	0.5232, 0.1459	0.5029, 0.1674
0.03	0.5957, 0.1096	0.5165, 0.1189	0.4447, 0.1411
0.02	0.5301, 0.0828	0.4772, 0.0877	0.4806, 0.1102
0.01	0.4259, 0.0537	0.4249, 0.0614	0.4301, 0.0891
0.008	0.3898, 0.0484	0.3868, 0.0581	0.4021, 0.0874
0.006	0.3164, 0.0455	0.3435, 0.0560	0.3875, 0.0868
0.005	0.2855, 0.0442	0.3127, 0.0554	0.3575, 0.0870
0.004	0.2723, 0.0434	0.2983, 0.0551	0.3464, 0.0874
0.003	0.2520, 0.0432	0.2680, 0.0549	0.3295, 0.0874
0.002	0.2323, 0.0432	0.2361, 0.0546	0.2647, 0.0880
0.001	0.0878, 0.0429	0.0801, 0.0551	0.1223, 0.0883

Table 22: ClpP (PDBID: 1YG6): median of rotational alignment score across 100 tests. Left of comma, our current method. Right of comma: our previous alignment method [2].

SNR	±70°	±60°	±50°
0.05	0.6959, 0.1952	0.6442, 0.2001	0.6114, 0.2204
0.04	0.7014, 0.1656	0.6229, 0.1729	0.5806, 0.1880
0.03	0.6718, 0.1352	0.5976, 0.1389	0.5233, 0.1509
0.02	0.6144, 0.1006	0.5594, 0.1050	0.5434, 0.1150
0.01	0.5203, 0.0612	0.4998, 0.0674	0.4920, 0.0914
0.008	0.4780, 0.0546	0.4716, 0.0600	0.4684, 0.0877
0.006	0.4493, 0.0475	0.4395, 0.0563	0.4377, 0.0851
0.005	0.4201, 0.0445	0.4045, 0.0557	0.4288, 0.0867
0.004	0.3826, 0.0436	0.3676, 0.0553	0.3912, 0.0852
0.003	0.3243, 0.0428	0.3232, 0.0555	0.3481, 0.0852
0.002	0.2367, 0.0428	0.2260, 0.0545	0.1706, 0.0869
0.001	0.0602, 0.0433	0.0701, 0.0548	0.0953, 0.0857

3 Extraction of Ribosome complexes from crowded subtomograms of a human pathogen Leptospira interrogans cell

We use Ribosome template search result to define the target complex ground truth region in a subtomogram. From such ground truth region, we calculate true-positive and false-positive rates at voxel level. When measuring crowdedness, because the ground truth of neighbor structures is unknown, we use the set of detected neighbor structural segments (i.e. those segments in $S_{\text{max\_structural}}$ that are not selected as target complex region) to define neighbor

structure region. The true-positive rate and false-positive rate respect to crowdedness level of the 20 subtomograms with top template matching scores is plotted in Figure 4.

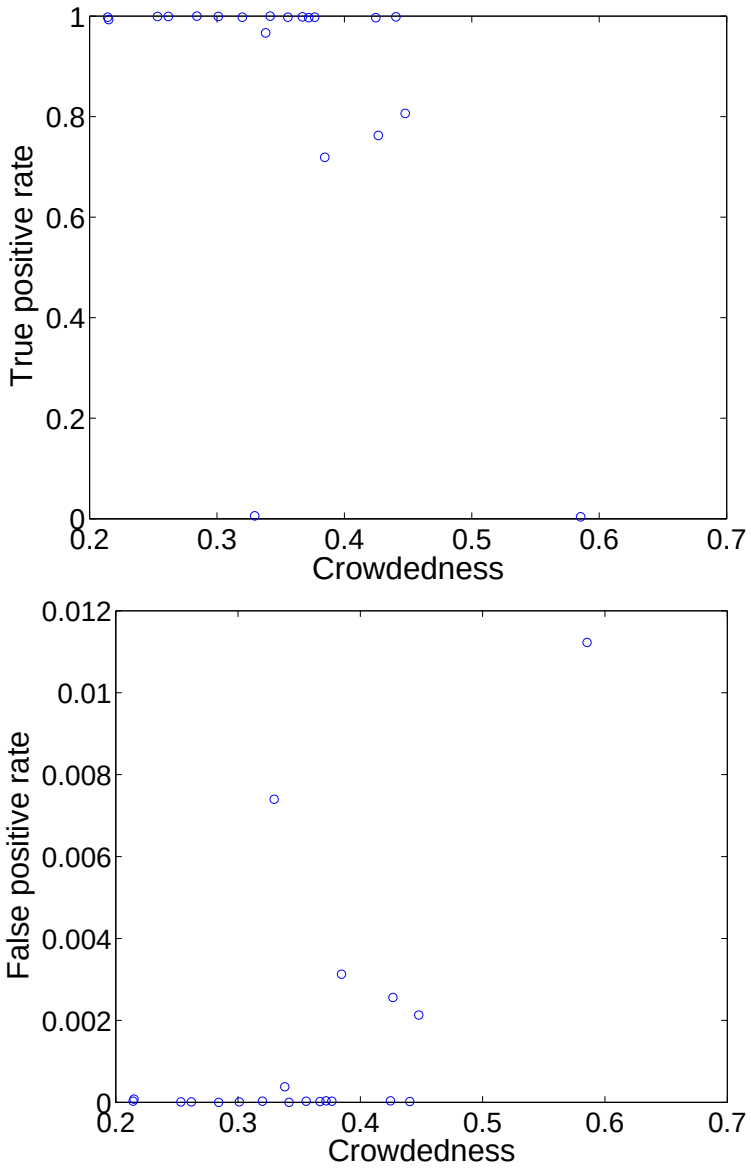

Figure 4: True-positive rate $o_{\text{true\_positive\_rate}}$ and false-positive rate $o_{\text{false\_positive\_rate}}$ of target complex region detection in crowded subtomograms of a human pathogen Leptospira interrogans cell.

References

[1] F. Förster, S. Pruggnaller, A. Seybert, and A.S. Frangakis. Classification of cryo-electron sub-tomograms using constrained correlation. *Journal of structural biology*, 161(3):276–286, 2008.

[2] M. Xu, M. Beck, and F. Alber. High-throughput subtomogram alignment and classification by Fourier space constrained fast volumetric matching. *Journal of Structural Biology*, 178(2):152–164, 2012.