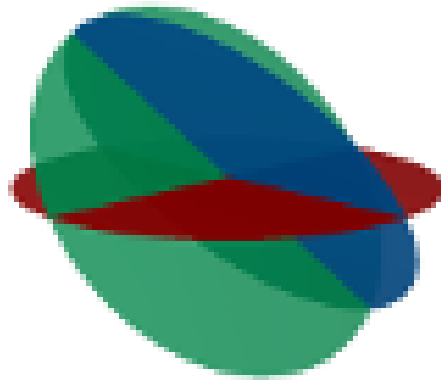# NYC Computational Cryo-EM Summer Workshop

Report of Contributions

Contribution ID: 4　　　　　　　　　　　　　　　　　Type: **not specified**

Welcome Remarks

Thursday, August 8, 2019 9:00 AM (10 minutes)

Presenter:　Dr GREENGARD , Leslie (CCM Director)

Contribution ID: **5**   Type: **not specified**

Session 1

Presenter: CHAIR: AMIT SINGER

Contribution ID: **6**   Type: **not specified**

Joachim Frank; Mapping of conformational states of macromolecules

Thursday, August 8, 2019 9:10 AM (40 minutes)

The ability of single-particle cryo-EM to furnish hundreds of thousand images of molecules fast-frozen in solution means that even rarely sighted conformations, corresponding to high-energy states, are represented in the data set. As a consequence, a continuum of states is experimentally accessible (Frank, Biochemistry 2018). Current practice of cryo-EM structure research does not take advantage of this opportunity. In collaboration with Abbas Ourmazd and Peter Schwander at the University of Wisconsin in Milwaukee we have developed an approach to map this continuum into a low-dimensional space by employing manifold embedding (Dashti et al., PNAS 2014; bioarxiv 2018; biorxiv 2019). From the observed occupancies the free energies of the molecule can be computed. The resulting free-energy landscape reveals the trajectories of states readily accessible by the molecule in the thermal bath, and the way the molecule's function is encoded in these trajectories. A Python implementation of this software, ManifoldEM, is being prepared for distribution.

Presenter: Dr FRANK , Joachim

Session Classification: Chair: Amit Singer

Contribution ID: **7**   Type: **not specified**

Joe Kileel; Autocorrelation analysis for cryo-EM and XFEL

Thursday, August 8, 2019 11:40 AM (40 minutes)

In 1980, Zvi Kam introduced an autocorrelation-based approach to cryo-electron microscopy (cryo-EM) single particle reconstruction, in which moments of the 2D images are computed and the 3D molecule is recovered by solving a polynomial system of equations. Recently, the method has also been used in X-ray free electron laser (XFEL) imaging.

This talk addresses important challenges in applying Kam's method. Drawing on invariant theory and tensor decomposition, we find the first provable solver for recovering the 3D molecule from the moments of the 2D images, for the setting of uniformly-distributed viewing angles. In XFEL (where the uniformity assumption holds), experiments show numerical stability and a speed-up over existing solvers by orders of magnitude. Additionally, we extend autocorrelation analysis to include an unknown, non-uniform distribution of viewing angles, which is the relevant case for cryo-EM. Rather unexpectedly, for referred orientations in cryo-EM, significantly fewer images are required than for the uniform distribution case in XFEL. A covariance-based non-convex optimization scheme for cryo-EM reconstruction is presented.

Presenter: Dr KILEEL, Joe

Session Classification: Chair: Carlos Oscar Sorzano

Contribution ID: **8**   Type: **not specified**

Jose Maria Carazo; Local Directional Resolution: Its impact in map validation, map sharpening and beyond

Thursday, August 8, 2019 2:40 PM (40 minutes)

I will introduce the notion of Directional Local Resolution, extending the current practice in Cry-oEM around local resolution. In this way, local resolution becomes a tensor. I will then present the wide range of new opportunities that this extension opens, while focusing during the presentation mostly in map validation and sharpening

Presenter: Dr MARIA CARAZO, Jose

Session Classification: Chair: Marin van Heel

Contribution ID: **9**   Type: **not specified**

Pilar Cossio; Unbiased cross-validation of cryo-EM maps

Thursday, August 8, 2019 2:00 PM (40 minutes)

Cryo-electron microscopy is a revolutionary technique that can Provide 3D density maps at near atomic resolution. However, map validation is still an open issue in the field. Despite several efforts from the community, it is possible to overfit the reconstructions to noisy data. Here, inspired by modern statistics, we develop a novel methodology that uses a small independent particle set to validate the 3D maps. The main idea is to monitor how the BioEM [1] probability evolves over the control set during the refinement. High-quality maps should increase in probability for higher refinement iterations. We then show that the similarity between the probability distributions of the two reconstructions from the gold-standard procedure is an additional quality indicator. We tested the method over several systems, some which were overfitted. Our method is able to clearly discriminate the overfitted sets from the non-overfitted ones. We conclude that having a control particle set, not used for the refinement, is essential for cross-validating cryo-EM maps.

References:
1. P. Cossio and G. Hummer. Journal of Structural Biology 184, 427-437 (2013).

Presenter: Dr COSSIO , Pilar

Session Classification: Chair: Marin van Heel

Contribution ID: **10**   Type: **not specified**

Bridget Carragher; CryoEM and CryoET: A brief intro and some thoughts about future computational needs

Thursday, August 8, 2019 9:50 AM (40 minutes)

Presenter: Dr CARRAGHER, Bridget

Session Classification: Chair: Amit Singer

Contribution ID: **11**   Type: **not specified**

Ayelet Heimowitz; Automatic particle picking and centering

Thursday, August 8, 2019 11:00 AM (40 minutes)

"Particle picking is a crucial first step in the computational pipeline of single-particle cryo-electron microscopy. Selecting particles from the micrographs is difficult especially for small particles with low contrast. As high-resolution reconstruction typically requires hundreds of thousands of particles, including rare views, manual picking is often too time-consuming. While template-based particle picking is currently a popular approach, it may introduce bias into the selection process. In addition, these methods may disregard the rare views of the particle. In an effort to avoid this, automatic particle pickers typically select any region of the micrograph that may possibly contain a particle projection. They therefore do not impose alignment restrictions, resulting in the presence of outliers as well as off-centered picked particles.

In this talk I will introduce the APPLE picker (Automatic Particle Picking with Low user Effort), a simple and novel approach for fast, accurate, and template-free particle picking. I will also discuss a novel method for alignment, wherein each particle is aligned independently of all other picked particles. In this way, the alignment is done directly by the APPLE-picker."

Presenter: Dr HEIMOWITZ , Ayelet

Session Classification: Chair: Carlos Oscar Sorzano

Contribution ID: **12**   Type: **not specified**

Alberto Bartesaghi; Data-driven Strategies for Unsupervised Single-particle Reconstruction

Thursday, August 8, 2019 3:50 PM (40 minutes)

Automated protocols for data acquisition used in modern cryo-EM facilities combined with the newer generation of larger and faster cameras have dramatically increased the throughput of data collection, exacerbating the need to develop effective strategies for structure determination that can operate without a user in the loop in order to keep up with the rate of data production. I'll present data-driven computational strategies we are using to enable the unsupervised execution of the single-particle cryo-EM reconstruction pipeline and show several examples where the results produced by these methods match the resolution of maps obtained by experienced users.

Presenter: Dr BARTESAGHI , Alberto

Session Classification: Chair: Fred Sigworth

Contribution ID: **13**   Type: **not specified**

Richard Henderson; Getting the most out of cryo-EM images

Thursday, August 8, 2019 4:30 PM (45 minutes)

Single particle cryo-EM images are still not perfect. I will discuss some work that Chris Russo and I have done recently to identify the remaining sources of imperfection and what might be done either by computation or by experimentation to overcome these problems.

Presenter: Dr HENDERSON, Richard

Session Classification: Chair: Fred Sigworth

Contribution ID: **14**   Type: **not specified**

Marin Van Heel; The statistical boundary between Science Fiction and Science Fact

Friday, August 9, 2019 9:00 AM (40 minutes)

In recent years we have seen an avalanche of cryo-EM (cryogenic Electron Microscopy) publications presenting beautiful biological structures at resolution levels of even better than ~3Å! This true "resolution revolution"has culminated in the 2017 Nobel prize for Chemistry being awarded for single-particle cryo-EM. Impressive as these results may be (and continue to be), various fundamental - mainly statistical - errors have been introduced in the early days of biological electron microscopy that are now interfering with the progress of the field. For example, a generic a-priori assumption made in the derivation of virtually all current resolution criteria in cryo-EM, is that signal and noise in the data are independent of each other, and that thus the cross-terms between signal and noise can be left out of the equations. Leaving them out of the equations, however, is equivalent to stating that the signal vectors and the corresponding noise vectors are orthogonal - which is incorrect - and not equivalent to stating that the signal vectors and the noise vectors are independent - which would have been correct. There are serious consequences to using flawed metrics in comparing the results of independently conducted experiments or in using such metrics for optimisation in automatic refinement procedures. The persistence of the field to move away from these flawed metrics has now led to an accumulation of errors by building new methodologies upon shaky foundations. Ignoring the Whittaker-Shannon sampling rules, for example, and justifying the results obtained from under-sampled data by pointing at "how beautiful they look" , brings us into swampy scientific territory.

Presenter: Dr VAN HEEL, Marin

Session Classification: Chair: Yoel Shkolnisky

Contribution ID: **15**
Type: **not specified**

Gregory Chirikjian; Cross-Validating and Fusing Information from SAXS and EM

Friday, August 9, 2019 9:40 AM (40 minutes)

This talk discusses information fusion algorithms for biomolecular structure determination using data obtained from both Small-Angle X-Ray Scattering (SAXS) and Single-Particle Electron Microscopy (EM). As a precursor to information fusion, methods for cross-validation to ensure data consistency are first addressed. As such, the theories behind SAXS and EM data acquisition are both reviewed. Whereas the information content in a set of class-averaged EM images is much higher than that of SAXS, it is shown that SAXS nevertheless can be used to cover "blind spots" resulting from cones of missing projection directions in EM.

Acknowledgments: This work was funded under grant NIH R01GM113240. The collaborations with the co-authors on the papers listed below are greatly appreciated.

References

Lyu, S., Wuelker, C., Jayaraman, A.S., Cai, Y., Zheng, J., Chirikjian, G.S., "Information Fusion Between SAX and EM," (in preparation), 2019

Kim, J.S., Afsari, B., Chirikjian, G.S., "Cross-Validation of Data Compatibility Between SAXS and Cryo-EM," Journal of Computational Biology, 24(1):13–30, 2017.

Dong, H., Kim, J.S., Chirikjian, G.S., "Computational Analysis of SAXS Data Acquisition," Journal of Computational Biology, 22(9): 787-805, 2015.

Presenter: Dr CHIRIKJIAN, Gregory

Session Classification: Chair: Yoel Shkolnisky

Contribution ID: **16**   Type: **not specified**

Joakim Andén; 3D covariance and graph Laplacians for continuous heterogeneity

Friday, August 9, 2019 11:00 AM (40 minutes)

A single biological molecule imaged in cryo-EM typically exhibits multiple structural configurations, each with a different function. These configurations may exist along a continuum of states, or conformations, giving rise to a manifold of continuous variability known as the conformational manifold. We propose to estimate this manifold by combining low-resolution reconstruction methods and graph Laplacian techniques. A covariance-based method is used to first obtain low-resolution estimates, which are used to construct a graph over the projection images. Computing the graph Laplacian and extracting its eigenvectors then allows us to characterize the underlying conformational manifold. Among other things, these Laplacian eigenvectors allow us to visualize the topology of the manifold, but they also provide a means for constructing higher-resolution molecular reconstructions through the method of "spectral volumes."Both applications are evaluated on synthetic and experimental datasets.

Presenter: ANDÉN, Joakim

Session Classification: Chair: Alex Barnett

Contribution ID: **17**   Type: **not specified**

Roy Lederman; Hyper-Molecules: High Dimensional Maps of Molecular Conformations

Friday, August 9, 2019 11:40 AM (40 minutes)

One of the open problems in cryo-EM is mapping complex heterogeneity, such as continuous heterogeneity. We begin our discussion with the question what does it mean to recover a heterogeneous structure, compared to a homogeneous structure or several distinct structures? We introduce "hyper-molecules,"a mathematical formulation which captures the continuum of states and the relationships between them, and a Bayesian framework for recovering these "hyper-molecules." We present preliminary implementations and results, which demonstrates how the heterogeneous structures can be recovered from synthetic and real data, and discuss some of the practical challenges and solutions.

We discuss next steps in this work on a scalable framework which would map complex heterogeneous structures, and would optionally allow researchers to explicitly encode prior knowledge, such as general physical properties, and existing knowledge about the specific structure, when such knowledge is available, in order to resolve complex structures which could otherwise require unrealistic amounts of data.

This is joint work with Joakim Andén and Amit Singer.

Presenter: Dr LEDERMAN, Roy

Session Classification: Chair: Alex Barnett

Contribution ID: **18**   Type: **not specified**

Erik Lindahl; Characterization of Molecular Motion and Refinement in single-particle Cryo-EM

Friday, August 9, 2019 2:00 PM (40 minutes)

Presenter: Dr LINDAHL, Erik

Session Classification: Chair: David Fleet

Contribution ID: **19**
Type: **not specified**

Jasenko Zivanov; Estimation of Higher-Order Aberrations and Anisotropic Magnification in Single-Particle Cryo-EM

Friday, August 9, 2019 2:40 PM (40 minutes)

Three new computational methods will be presented that aim to improve the maximal attainable resolution of single-particle cryo-EM reconstructions by measuring and correcting for optical effects in the microscope. All methods work on a given data set, so they can be applied after the images have been collected. The first two methods allow us to measure the symmetrical and antisymmetrical higher-order optical aberrations (which deform the CTF and induce a phase shift, respectively), while the third method measures the magnification anisotropy. The methods are computationally efficient, and implementations will be provided with the upcoming release of Relion.

Presenter: Dr ZIVANOV, Jasenko

Session Classification: Chair: David Fleet

Contribution ID: **20**   Type: **not specified**

Ali Punjani; Cryo-EM Resolution Estimation

Friday, August 9, 2019 3:50 PM (40 minutes)

Resolution estimation is a critical and practical aspect of cryo-EM reconstruction. In this project, we explore a new way to think about resolution estimation in cryo-EM. We apply the principle of cross-validation as a general idea to develop a framework, from which tasks like resolution estimation can be formulated, and solutions can be derived from first principles. The resulting solutions yield interesting properties that can be useful in practice.

Presenter: Dr PUNJANI, Ali

Session Classification: Chair: Marina Spivak

Contribution ID: **21**　　　　　　　　　　　　　　　　　　　Type: **not specified**

Gabor Herman; The Efficacy of Iterative Algorithms for Reconstruction from Cryo-EM Data

Friday, August 9, 2019 4:30 PM (40 minutes)

• Tomographic data collection methods (of which cryo-EM is an example) provide us with approximations of line integrals of some parameter.

• Reconstruction is the attempted recovery of the distribution of the parameter values from such data.

• The two major categories of reconstruction techniques are:

–transform methods such as weighted backprojection (WBP) and FBP; and

–series expansion methods such as the algebraic reconstruction techniques (ART) and SIRT.

• The former methods are in wide use because of their fast speed and simple implementation.

• The latter methods have been claimed to provide greater detail with incomplete or noisy data.

–Series expansion methods specify the distribution as a linear combination of some basis functions.

–They estimate the coefficients in such an expansion by an iterative reconstruction (IR) algorithm.

• In many fields of application of tomography (e.g., in medical CT), IR has become the norm.

• This has not happened in cryo-EM, where there seems to be a greater inertia against its use.

• The reluctance in the EM community to use IR goes back fifty years and is still strong.

• The purpose of this presentation is to discuss this phenomenon.

–Examples are given for the efficacy of IR in EM and other applications.

–Objections to the use of IR in cryo-EM are listed and critically examined.

Presenter:　Dr HERMAN, Gabor

Session Classification:　Chair: Marina Spivak

Contribution ID: **22**   Type: **not specified**

New Methods for Denoising and CTF Correction

Thursday, August 8, 2019 5:15 PM (2 hours)

We will present recent work on the problem of simultaneously denoising cryo-EM images and correcting for the effects of the contrast transfer function (CTF). The methods used are based on new results from high-dimensional principal component analysis and matrix recovery in the spiked covariance model. We use new spectral shrinkers that account for the effects of both the CTF and the colored noise distribution. This is joint work with Joakim Andén and Amit Singer.

Primary author: LEEB, William (University of Minnesota)

Co-authors: CHAIR: AMIT SINGER;   ANDÉN, Joakim

Presenter: LEEB, William (University of Minnesota)

Session Classification: Poster Session and Reception

Contribution ID: **24**   Type: **not specified**

Computing and Understanding Statistical Models for Heterogeneous Biological Nano-machines

Thursday, August 8, 2019 5:15 PM (2 hours)

The goal of this research is to understand the dynamical motion of nanoscale biological machines such as viruses directly from large sets of data. The ideal data would be 4-D measurements (3 spatial and 1 temporal) on each instance of the machine. The most informative available data are 2-D cryo-electron microscopy projection images of flash-frozen instances, one image for each instance. Because the instances are flash frozen, each image represents a sample from the statistical mechanical ensemble of the machine and the fact that many instances are available is the key to learning the 4-D behavior of the machine. Our proven learning methods based on maximum likelihood estimators can provide the 6-D spatial covariance function of the electron scattering intensity of the machine and the covariance function includes dynamical motion information. But 6-D information is challenging to interpret. Our goal is to understand the 4-D behavior of the machine by learning a generative mathematical mechanical model of the machine based on the connection between the model and the experiment that is provided by statistical mechanics. From such a mathematical model we can bring the machine back to life in 4-D and compute all of the machine's dynamical behaviors.

Primary authors: GONG, Yunye (Cornell University); Prof. DOERSCHUK, Peter C. (Cornell University)

Presenters: GONG, Yunye (Cornell University); Prof. DOERSCHUK, Peter C. (Cornell University)

Session Classification: Poster Session and Reception

Contribution ID: **25**

Type: **not specified**

Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs

Thursday, August 8, 2019 5:15 PM (2 hours)

Cryo-electron microscopy is a popular method for protein structure determination. Identifying a sufficient number of particles for analysis can take months of manual effort. Current computational approaches find many false positives and require significant ad hoc post-processing, especially for unusually-shaped particles. To address these shortcomings, we develop Topaz, an efficient and accurate particle picking pipeline using neural networks trained with a novel positive-unlabeled (PU) learning method. This framework enables state-of-the-art particle detection models to be trained with few, sparsely labeled particles and no labeled negatives. Topaz retrieves many more real particles than conventional picking methods while maintaining low false positive rates, is uniquely capable of picking challenging unusually-shaped proteins (e.g. small, non-globular, and asymmetric), produces more representative particle sets, and does not require post hoc curation; we demonstrate these results on two currently difficult datasets and three conventional datasets. Our PU learning method is general-purpose and outperforms existing PU learning approaches. Topaz is modular, standalone, free, and open source.

Primary author: BEPLER, Tristan

Co-authors: MORIN, Andrew (MIT); RAPP, Micah (Columbia University); BRASCH, Julia (Columbia University); LAWRENCE, Shapiro (Columbia University); NOBLE, Alex (New York Structural Biology Center); BONNIE, Berger (MIT)

Presenter: BEPLER, Tristan

Session Classification: Poster Session and Reception

Contribution ID: **26**   Type: **not specified**

Z-contrast enhancement for small protein cryo-EM structure determination

Thursday, August 8, 2019 5:15 PM (2 hours)

Single-particle EM 3D reconstruction comprises two aspects: data acquisition and data analysis. While "data analysis" aims to extract as much information as possible from noisy raw images, "data acquisition" strives to maximize the amount of information in data recording.

Electron scattering cross-sections of heavy atoms quantitatively differ from light atoms (namely C, N, O) abundant in protein molecules, and this physical property endorses enhanced amplitude contrast in cryo-EM. Here we demonstrate that a 45 kDa copper storage protein (CSP) can be visualized by single-particle cryo-EM *without phase-plate* imaging and its density map has been reconstructed to 3.5 Å resolution upon multiset-CTF correction. The Z-contrast enhancement technique has broader applications in single-particle EM structural biology.

Primary authors: Mr OKEN, Adam (Chen Lab, OHSU); CHEN, James

Presenter: CHEN, James

Session Classification: Poster Session and Reception

Contribution ID: **28** Type: **not specified**

3D Alignment based on Regionalized Deep Learning for Cryo-EM Reconstruction

Thursday, August 8, 2019 5:15 PM (2 hours)

Cryo Electron Microscopy (Cryo-EM) is currently one of the main tools to reveal the structural information of biological specimens. However, in a common Cryo-EM processing workflow, the 3D alignment step, due to the very low signal-to-noise ratio of Cryo-EM images, is a prone error process. Thus, the reconstructed 3D maps can show areas with low resolution.

In this work, a novel method to align sets of projection images in the 3D space is presented. Our proposal is based on deep learning networks. Specifically, we propose to design several deep networks on a regionalized basis, creating a bank of networks to classify the projection images in sub-regions of the 3D space and, then, making a refinement of the final 3D alignment parameters. We show that the method applied to experimental data results in accurately aligned images.

Primary author: Dr JIMENEZ-MORENO, Amaya (National Center of Biotechnology (CSIC))

Co-authors: Dr S. SORZANO, Carlos Oscar (National Center of Biotechnology (CSIC)); Dr CARAZO, Jose Maria (National Center of Biotechnology (CSIC))

Presenter: Dr JIMENEZ-MORENO, Amaya (National Center of Biotechnology (CSIC))

Session Classification: Poster Session and Reception

Contribution ID: **29**   Type: **not specified**

Non-Uniformity of Projection Distributions Attenuates Resolution in Cryo-EM

Thursday, August 8, 2019 5:15 PM (2 hours)

Virtually every single-particle cryo-EM experiment currently suffers from specimen adherence to the air-water interface, leading to a non-uniform distribution in the set of projection views. Non-uniform (anisotropic) distributions can negatively affect map quality, elongate structural features, and in some cases, prohibit interpretation altogether. Although some consequences of non-uniform sampling have been described qualitatively, we know little about how sampling quantitatively affects resolution in cryo-EM. Here, we show how inhomogeneity in any projection distribution scheme attenuates the global Fourier Shell Correlation (FSC) in relation to the number of particles and a single geometrical parameter, which we term the sampling compensation factor (SCF). The reciprocal of the SCF is defined as the average over Fourier shells of the reciprocal of the per-particle sampling and normalized to unity for uniform distributions. The SCF ranges from one to zero, with values close to the latter implying large regions of poorly sampled or completely missing data in Fourier space. Using two synthetic test cases, influenza hemagglutinin and human apoferritin, we demonstrate how any amount of sampling inhomogeneity always attenuates the FSC compared to a uniform distribution.

Primary author: Dr BALDWIN, Philip (The Salk Institute)

Co-author: Prof. LYUMKIS, Dmitry (The Salk Institute)

Presenter: Dr BALDWIN, Philip (The Salk Institute)

Session Classification: Poster Session and Reception

Contribution ID: **30**   Type: **not specified**

DeepRes: A New Deep Learning-based Local Resolution Method

Thursday, August 8, 2019 5:15 PM (2 hours)

We present a method to estimate a new local quality measure for 3D cryoEM maps that adopts the form of a local resolution-type of information. The algorithm (DeepRes) is based on deep learning 3D features detection. DeepRes is fully automatic and parameter free and avoids issues of most current methods, such as their insensitivity to enhancements due to B-factor sharpening (unless the 3D mask is changed), among others, which is an issue virtually neglected in the cryoEM field until now. In this way, DeepRes can be applied to any map, detecting subtle changes of local quality after applying enhancement processes, like isotropic filters or substantially more complex procedures, such as local sharpening or denoising, that may be very difficult to follow by current methods. The comparison with traditional local resolution indicators is also addressed. The comparison between current methods and DeepRes also allows detecting over-sharpening in CryoEM maps.

Primary authors: Dr RAMÍREZ-APORTELA, Erney (CNB-CSIC (Madrid, Spain)); Prof. CARAZO, Jose Maria (CNB-CSIC (Madrid, Spain)); Dr SORZANO, Carlos Oscar S. (CNB-CSIC (Madrid, Spain))

Presenter: Dr RAMÍREZ-APORTELA, Erney (CNB-CSIC (Madrid, Spain))

Session Classification: Poster Session and Reception

Contribution ID: **32** Type: **not specified**

ManifoldEM –Mapping Continuous Conformations, Energy Landscapes & Functional Pathways

Thursday, August 8, 2019 5:15 PM (2 hours)

We have developed a manifold-based machine-learning approach for analyzing cryoEM single-particle data. This approach is capable of mapping continuous conformational changes of biological molecules along any user-selected trajectory on the energy landscape, without timing information, supervision, or templates. Our unbiased approach (1) reveals the number of degrees of freedom exercised during the observations, (2) retrieves energy landscapes explored by the biological molecule, (3) determines least-action functional pathways, and (4) compiles 3D movies of the continuous conformational changes associated with functional pathways. These capabilities constitute a powerful platform for quantitative study of a wide range of key biological processes.

Primary author: Prof. SCHWANDER, Peter (University of Wisconsin Milwaukeee)

Co-authors: Dr DASHTI, Ali (University of Wisconsin-Milwaukee); Dr MASHAYEKHI, Ghoncheh (University of Wisconsin-Milwaukee); Dr LIAO, Hstau (Columbia University, New York); Dr MAJI, Suvrajit (Columbia University, New York); Mr SEITZ, Evan (Columbia University, New York); Dr FRANK , Joachim (Columbia University, New York); Prof. OURMAZD, Abbas (University of Wisconsin-Milwaukee)

Presenter: Prof. SCHWANDER, Peter (University of Wisconsin Milwaukeee)

Session Classification: Poster Session and Reception

Contribution ID: **33**   Type: **not specified**

Deep learning and image analysis tools for automated analysis and assessment of cryo-EM datasets

Thursday, August 8, 2019 5:15 PM (2 hours)

The growth of cryo-EM into a mainstream structural biology tool has led to its widespread adoption for users across a range of expertise, where experts represents a small fraction of cryo-EM users. Considering the manual and subjective decisions involved in solving a structure, such as the programs, parameters and determination of good micrographs and good 2D class averages, cryo-EM frustrates many users. To make cryo-EM data processing more user-friendly, we have developed an automated pipeline for cryo-EM data preprocessing and assessment using a combination of deep learning and image analysis tools to help streamline the process of cryo-EM structure determination. Specifically, we have built a deep-learning based generic micrograph classifier that can assess the quality of a micrograph with an accuracy of 96% allowing bad micrographs to be removed without user decision. We have also built a 2D class average classifier that can identify the good 2D class averages from RELION and help to find the optimal parameters in 2D classification. We have verified the performance of our pipeline on multiple datasets including both EMPIAR and real-world datasets. We propose that our automatic pipeline will make cryo-EM preprocessing more convenient for cryo-EM users from a range of backgrounds.

Primary authors: LI, Yilai (University of Michigan); CIANFROCCO, Michael (University of Michigan)

Presenter: LI, Yilai (University of Michigan)

Session Classification: Poster Session and Reception

Contribution ID: **34**   Type: **not specified**

Multi-tapered CTF estimation

Thursday, August 8, 2019 5:15 PM (2 hours)

We present a novel method for contrast transfer function (CTF) estimation. Our method is based on the multi-taper method for power spectral density estimation, which aims to reduce the bias and variance of the estimator. Furthermore, we use known properties of the CTF and of the background of the power spectrum to increase the accuracy of our estimation. We will show that the resulting estimates capture the zero-crossings of the CTF in low-mid frequencies. Our estimation can be incorporated into existing cross-correlation based CTF estimation techniques, and, in addition, paves the way for new, zero-crossing based estimation methods.

Primary authors: HEIMOWITZ, Ayelet (Princeton University); ANDÉN, Joakim; CHAIR: AMIT SINGER

Presenter: HEIMOWITZ, Ayelet (Princeton University)

Session Classification: Poster Session and Reception

Contribution ID: **35**   Type: **not specified**

Deep Learning tools for particle picking

Thursday, August 8, 2019 5:15 PM (2 hours)

Cryo-EM workflows require from tens of thousands of high-quality particle projections to unveil the three-dimensional structure of macromolecules. Current methods for automatic particle-picking tend to suffer from high false-positive rates, hurdling the reconstruction process. Usually, the failures of one particle-picking algorithm are typically not the failures of another. Therefore, a smart consensus over the output of different particle-picking algorithms, named *DeepConsensus* is presented in this work. *DeepConsensus* is based on a deep-convolutional neural network that is trained on a semi-automatically generated data set, resulting in a set of particles with a lower false-positive ratio than the initial sets[1]. However, one common false-positives source for most of the particle-picking algorithms is the presence of carbon and different types of high-contrast contaminations. In order to avoid those areas affected with this kind of contaminants, we have developed a deep-learning approach named *MicrographCleaner*, designed to discriminate the regions of micrographs suitable for particle picking from those labeled as contaminated[2]. *Micrograph-Cleaner* implements a U-net-like model trained on a manually curated dataset, compiled from over five hundred micrographs.

[1] Sanchez-Garcia, Ruben et al. IUCrJ vol. 5,Pt 6 854-865. 30 Oct. 2018, doi:10.1107/S2052252518014392

[2] Sanchez-Garcia, Ruben et al. Submitted at BioInformatics, available in BioRxiv doi: https://doi.org/10.1101/677542

Primary authors: Dr MALUENDA, David (National Center of Biotechnology (CSIC)); Mr SANCHEZ--GARCIA, Ruben (National Center of Biotechnology (CSIC)); Dr SEGURA, Joan (San Diego Supercomputer Center); Dr SANCHEZ-SORZANO, Carlos Oscar (National Center of Biotechnology (CSIC)); Dr CARAZO, Jose Maria (National Center of Biotechnology (CSIC))

Presenter: Dr MALUENDA, David (National Center of Biotechnology (CSIC))

Session Classification: Poster Session and Reception

Contribution ID: **36**   Type: **not specified**

Manifold embedding analysis of TRPV1

Thursday, August 8, 2019 5:15 PM (2 hours)

The TRPV1 ion channel is a heat sensor that plays a key role in pain sensing pathways. Recent advances in cryo-electron microscopy (cryo-EM) have facilitated a recent explosion in the availability of TRP channel structures. Despite these structures, the temperature-sensing mechanism of any TRP channel remains poorly understood. The manifold embedding method in cryo-EM allows the identification of free energy landscapes and related continuous molecular trajectories. Here, we apply this method to the original apo TRPV1 dataset at low temperature to find the conformations accessible via thermal motions, including the identification of motions of specific transmembrane helices. This result sets the stage for a better understanding of the TRPV1 temperature sensing mechanism via cryo-EM.

Primary authors: Dr HANSON, Sonya M. (Max Planck Institute of Biophysics);   Prof. FRANK, Joachim (Columbia University)

Presenter: Dr HANSON, Sonya M. (Max Planck Institute of Biophysics)

Session Classification: Poster Session and Reception

Contribution ID: **37**   Type: **not specified**

Reconstructing continuous distributions of 3D protein structure from cryo-EM images

Thursday, August 8, 2019 5:15 PM (2 hours)

In single particle cryo-EM, the central problem is to reconstruct the three-dimensional structure of a protein from $10^4 - 10^7$ noisy and randomly oriented two-dimensional projections. However, the imaged protein molecules may exhibit structural variability, which complicates reconstruction and is typically addressed using discrete clustering approaches that fail to capture the full range of protein dynamics. Here, we present a novel framework using deep neural networks for cryo-EM reconstruction that extends naturally to modeling continuous generative factors of protein structural heterogeneity. We demonstrate that our framework, termed CryoNN, can perform *ab initio* reconstruction of 3D protein structures from simulated and real cryo-EM image data. To our knowledge, CryoNN is the first neural network-based approach for cryo-EM reconstruction and the first end-to-end method for directly reconstructing continuous ensembles of protein structures from cryo-EM images.

Primary author: ZHONG, Ellen (Massachusetts Institute of Technology)

Co-authors: BERGER, Bonnie (Massachusetts Institute of Technology); Prof. DAVIS, Joseph (Massachusetts Institute of Technology); BEPLER, Tristan

Presenter: ZHONG, Ellen (Massachusetts Institute of Technology)

Session Classification: Poster Session and Reception

Contribution ID: **38**   Type: **not specified**

Spectral volumes for cryo-EM reconstruction with continuous heterogeneity

Thursday, August 8, 2019 5:15 PM (2 hours)

In this work, we address the continuous heterogeneity problem. We parametrize the 3D density maps of the particles being imaged using a low-dimensional manifold of conformations. This parametrization is based on low-resolution reconstructions and Laplacian eigenmaps. We use this parametrization to form a generalized tomographic reconstruction problem which reconstructs a density map at each point on the manifold. The resulting set of equations is high-dimensional, but by exploiting certain properties we recast the problem as a deconvolution. It can be solved efficiently using the conjugate gradient method, where the forward operator kernel is calculated non-uniform FFT. The solution of this problem is given by a set of spectral volumes. These volumes are used to calculate high-resolution reconstructions of the projected particles as well as providing insight into the nature of the conformational changes. We present results for this method on both simulated models and experimental data.

Primary authors: HALEVI, Amit (Princeton University); Dr MOSCOVICH, Amit (Princeton University); ANDÉN, Joakim; CHAIR: AMIT SINGER

Presenters: HALEVI, Amit (Princeton University); Dr MOSCOVICH, Amit (Princeton University)

Session Classification: Poster Session and Reception

Contribution ID: **42**

Type: **not specified**

Fast rigid image alignment via factorization of the translation kernel

Thursday, August 8, 2019 5:15 PM (2 hours)

An important subproblem in image analysis is the comparison of different images, which usually involves the calculation of inner products between pairs of images. In this poster we provide a new method for computing such inner products for arbitrary rotations and a user-specified range of translations. Our method takes advantage of the Fourier–Bessel basis to efficiently handle rotations, while at the same time using an optimal factorization of the translation kernel. For standard applications in cryomicroscopy our method is roughly an order of magnitude faster than the state of the art.

Primary authors: Prof. RANGAN, Aaditya (New York University); Dr SPIVAK, Marina (Flatiron Institute); ANDÉN, Joakim; Dr BARNETT, Alex (Flatiron Institute)

Presenter: Prof. RANGAN, Aaditya (New York University)

Session Classification: Poster Session and Reception

Contribution ID: **43** Type: **not specified**

Introduction to Deep Learning in cryo-EM

Saturday, August 10, 2019 9:00 AM (1h 30m)

Session Classification: Breakout Session

Contribution ID: 44 Type: **not specified**

Collaborative Research with Code Ocean

Saturday, August 10, 2019 9:00 AM (1h 30m)

Presenter: FEI, Xu (Code Ocean)

Session Classification: Breakout Session

Contribution ID: **45** Type: **not specified**

Show and Tell

Saturday, August 10, 2019 10:30 AM (2 hours)

Session Classification: Breakout Session

Contribution ID: **46**   Type: **not specified**

Radial function based ab-initio tomographic reconstruction for cryo electron microscopy

Thursday, August 8, 2019 5:15 PM (2 hours)

The scope of this paper is the tomographic reconstruction of the observed object in the ab-initio case where the volume has to be estimated only from a raw projection dataset. A new fast approach based on a parametric model of the volume is presented. The description of the model and the search of the parameters are detailed. The accuracy and robustness of the proposed reconstruction method is shown on synthetic and real databases.

Primary author: Dr BAUDRIER, Étienne

Presenter: Dr BAUDRIER, Étienne

Session Classification: Poster Session and Reception

Contribution ID: 47   Type: **not specified**

Cryo-Electron Microscopy Image Analysis Using Multi-Frequency Vector Diffusion Maps

Thursday, August 8, 2019 5:15 PM (2 hours)

Cryo-electron microscopy (EM) single particle reconstruction is an entirely general technique for 3D structure determination of macromolecular complexes. However, because the images are taken at low electron dose, it is extremely hard to visualize the individual particle with low contrast and high noise level. In this paper, we propose a novel approach called multi-frequency vector diffusion maps (MFVDM) to improve the efficiency and accuracy of cryo-EM 2D image classification and denoising. This framework incorporates different irreducible representations of the estimated alignment between similar images. In addition, we propose a graph filtering scheme to denoise the images using the eigenvalues and eigenvectors of the MFVDM matrices. Through both simulated and publicly available real data, we demonstrate that our proposed method is efficient and robust to noise compared with the state-of-the-art cryo-EM 2D class averaging and image restoration algorithms.

Primary authors: Mrs FAN, Yifeng; Dr ZHAO, Zhizhen

Presenter: Mrs FAN, Yifeng

Session Classification: Poster Session and Reception