Erik Henning Thiede

Research Fellow, Flatiron Institute CCM

Education

2013–2019 **Doctorate in Chemistry**, *University of Chicago*.

2009–2013 **Bachelor of Science in Chemistry**, *UNC Chapel Hill*, Graduated with honors and highest distinction.

Professional Appointments

2020– **Research Fellow**, *Center for Computational Mathematics, Flatiron Institute*, Working in collaboration with Prof. Risi Kondor, Dr. Pilar Cossio, and Dr. Sonya Hanson.

2019–2020 **Postdoctoral Scholar**, *University of Chicago*, Advised by Prof. Risi Kondor.

2014–2019 **Graduate Research Associate**, *University of Chicago*, Jointly advised by Prof. Aaron Dinner and Prof. Jonathan Weare.

2013 **Research Associate**, *University of Chicago*, Advised by Prof. David Mazziotti.

Honors and Fellowships

2021 **Blavatnik Regional Award Finalist in Chemistry**, *Blavatnik Awards for Young Scientists*.
Award recognizing outstanding postdoctoral scientists from institutions in New York, New Jersey, and Connecticut. Each year, one winner and two finalists are chosen in each discipline.

2017 **MolSSI Software Fellow**, *Molecular Sciences Software Institute*.
Fellowship for the development of algorithms and software for computational molecular sciences. Initially awarded for Fall 2017, successfully renewed for Spring 2018 to Summer 2019.

2013 **Freud Fellow**, *University of Chicago*.
Initial grant of $10,000 doubled to extend to 2014.

2013 **Hypercube Scholar**, *University of North Carolina at Chapel Hill*.

2012 **Award for Undergraduate Excellence in Physical Chemistry**, *University of North Carolina at Chapel Hill*.

2009 **Carolina Scholar**, *University of North Carolina at Chapel Hill*.
Highest merit-based scholarship offered by UNC Chapel Hill.

Publications

Asterisks denote equal authorship.

2022 Li X.S., Van Koten B., Dinner A.R., **Thiede E.H.** "Understanding the Sources of Error in MBAR through Asymptotic Analysis" arXiv:2203.01227 (submitted to Journal of Chemical Physics) https://arxiv.org/abs/2203.01227

2021 **Thiede E.H.**, Zhou W., Kondor R. "Autobahn: Automorphism-based Graph Neural Nets" NeurIPS 34 https://proceedings.neurips.cc/paper/2021/file/faf02b2358de8933f480a146f4d2d98e-Paper.pdf

2021 Bratholm L.A., Gerrard W., Anderson B., Bai S., Choi S., Dang L., Hanchar P., Howard A., Kim S., Kolter Z., Kondor R., Kornbluth M., Lee Y., Lee Y., Mailoa J.P., Nguyen T.T., Popovic M., Rakocevic G., Reade W., Song W., Stojanovic L., **Thiede E.H.**, Tijanic N., Torrubia A., Willmott D., Butts C.P., Glowacki D.R. "A community-powered search of machine learning strategy space to find NMR property prediction models" PLOS One 16, e0253612 https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0253612

2021 Giraldo-Barreto J., Ortiz S., **Thiede E.H.**, Palacio-Rodriguez K., Carpenter B., Barnett A.H., Cossio P. "A Bayesian approach for extracting free energy profiles from cryo-electron microscopy experiments using a path collective variable" Scientific Reports 11, 13657 https://www.nature.com/articles/s41598-021-92621-1

2021 Webber R.J., **Thiede E.H.**, Dow D., Dinner A.R., Weare J. "Error bounds for dynamical spectral estimation" SIMODS 3, 225 https://epubs.siam.org/doi/abs/10.1137/20M1335984

2020 Lorpaiboon C.*, **Thiede E.H.*,** Webber R.J.*, Weare J., Dinner A.R. "Integrated Variational Approach to Conformational Dynamics: A Robust Strategy for Identifying Eigenfunctions of Dynamical Operators" The Journal of Physical Chemistry B 124, 9354-9364 https://pubs.acs.org/doi/abs/10.1021/acs.jpcb.0c06477

2020 Antoszewski A., Feng C.J., Vani B.P., **Thiede E.H.**, Hong L., Weare J. Tokmakoff A., Dinner A.R. "Insulin dissociates by diverse mechanisms of coupled unfolding and unbinding" The Journal of Physical Chemistry B 124, 5571-5587 (2020), https://aip.scitation.org/doi/abs/10.1063/1.5063730

2020 **Thiede E.H.**, Son H.T., Kondor R. "The general theory of permutation equivariant neural networks and higher order graph variational encoders" arXiv:2004.03990 https://arxiv.org/abs/2004.03990

2019 **Thiede E.H.**, Giannakis D., Dinner A.R., Weare J. "Galerkin Approximation of Dynamical Quantities using Trajectory Data" Journal of Chemical Physics 150, 24411 https://aip.scitation.org/doi/abs/10.1063/1.5063730

2018 Hong L., Vani B.P., **Thiede E.H.**, Rust M.J., Dinner A.R. "Molecular dynamics simulations of nucleotide release from the circadian clock protein KaiC reveal atomic-resolution functional insights" PNAS 115, E11475 https://www.pnas.org/content/115/49/E11475.short

2017 Dinner A.R.*, **Thiede E.H.*,** Van Koten B.*, Weare J.* "Stratification as a general variance reduction method for Markov chain Monte Carlo" SIAM/ASA Journal on Uncertainty Quantification 8, 1139–1188 https://epubs.siam.org/doi/abs/10.1137/18M122964X

2016 **Thiede E.H.**, Van Koten B., Weare J., Dinner A.R. "Eigenvector method for umbrella sampling enables error analysis." J. Chem. Phys. 145, 084115 https://aip.scitation.org/doi/abs/10.1063/1.4960649

2015 **Thiede E.H.*,** Van Koten B.*, Weare J.* "Sharp entrywise perturbation bounds for Markov chains" SIAM J. Matrix Anal. Appl. 36, 917 https://epubs.siam.org/doi/abs/10.1137/140987900

Presentations

2022 **Seminar at John Hopkins University Dept. of CS**, *Invited Talk*, How Group Theory and Graph Neural Networks Come Together.

2022 **Multiscale Materials Modeling**, *Oral Presentation*, Autobahn: Constructing Neural Networks from Molecular Substructures.

2022 **Physical Chemistry Seminar at UNC Chapel Hill**, *Invited Talk*, Deconstructing Neural Networks for Molecules.

2022 **APS March Meeting**, *Oral Presentation*, Graph Neural Networks that incorporate Physical Structure.

2021 **GAMM Young Researchers Minisymposium on Molecular Dynamics**, *Invited Talk*, Dynamical Galerkin Approximation: long-time rates from short-time data.

2019 **Minisymposium in SIAM Conference on Applications of Dynamical Systems**, *Invited Talk*, Long-time Dynamical Estimates from Short-time Data using Dynamical Galerkin Approximation.

2018 **Berkeley Statistical Mechanics Meeting**, *Poster*, Quantitative Information from Trajectory Data through Dynamical Galerkin Approximation.

2017 **IPAM Program: Complex High-Dimensional Energy Landscapes: Workshop III**, *Poster*, Quantitative Information from Trajectory Data through Dynamical Galerkin Approximation.

2017 **Berkeley Statistical Mechanics Meeting**, *Poster*, Estimating Dynamical Information using Diffusion Maps.

2015 **Berkeley Statistical Mechanics Meeting**, *Poster*, Umbrella Sampling: Insights from Numerical Analysis.

Scientific Leadership

Teaching Experience

Mar.–Jun. 2014, 2015 **Teaching Assistant for Computational Methods in Chemistry and Biology**, *University of Chicago*.
- Graded student work, helped develop curricula and homework assignments.

Oct. 2013–Mar. 2014 **Teaching Assistant for General Chemistry**, *University of Chicago*.
- Graded student work, led discussion sections and supervised laboratory work.

Outreach

2021 **Career in the Sciences: An Interactive Panel**, *CUNY*.
- Recruiting session focused on encouraging underrepresented students to believe that they have a place in academia and apply to graduate school.

2020 **Flatiron-wide Algorithms and Mathematics**, *Flatiron Institute*.
- Gave an introduction to advanced methods in Markov chain Monte Carlo aimed at non-experts.

2016, 2018 **SESAME Program**, *University of Chicago*.
- Taught at the SESAME program, which offers professional development courses for Chicago Public School teachers.
- In 2016, gave a lecture on the statistics of biomolecules.
- In 2018, assisted a lecture on enzymes.

2017 **Chicago Area Undergraduate Research Symposium**, *University of Chicago*.
- Volunteered to judge posters for the CAURS conference.

2013, 2015, 2017 **Physics with a Bang**, *University of Chicago*.
- Volunteer at Physics with a Bang, the annual open-house for the James Franck Institute.
- Performed scientific demonstrations for elementary and middle school children on particle jamming (2013) and molecular simulation (2015).
- In 2017, performed overtone singing to demonstrate Fourier decomposition of sound.

Participation in Scientific Competitions

2019 **Predicting Molecular Properties**, *Kaggle*, 12th place (Gold).
- Along with other contestants, attempted to regress NMR couplings on a collection of small molecules.

Software Development

EMUS Package Open source implementation of the **E**igenvector **M**ethod for **U**mbrella **S**ampling available at `https://github.com/ehthiede/EMUS`. The package provides:
- Rapid convergence to the maximum likelihood estimate of averages calculated by umbrella sampling, leading to accuracy equivalent to or surpassing WHAM.
- Support for large, high-dimensional umbrella sampling datasets. Provides support for calculating averages over the state space, as well as free energy surfaces in collective variable spaces with dimension larger than 2.
- Error analysis of free energy calculations, detailed in Thiede et al. JCP 2016.

PyDiffMap Codeveloper of an open source implementation of the Diffusion Map algorithm, available at https://github.com/DiffusionMapsAcademics/pyDiffMap. Among other features, the implementation provides
- Automatic bandwidth parameter selection.
- Variable bandwidth diffusion kernels.
- Out-of-sample extension.
- Intelligent neighborlist construction using scikit-learn's neighbor class.
- Support of target-measure diffusion maps.

PyEDGAR Open source implementation of the dynamical Galerkin approximation algorithm available at `https://github.com/ehthiede/PyEDGAR`. The package implements dynamical Galerkin approximation, a novel method for estimating long-time statistics from collections of short-time trajectories detailed in Thiede et al., arXiv:1810.01841. This algorithm generalizes Markov state modeling, and is closely related to Koopman operator methods such as the extended dynamic mode decomposition and the Variational approach.

Cormorant Contributing author and maintainer of an implementation of the Cormorant machine learning architecture available at `https://github.com/risilab/cormorant`.

Enhanced Sampling Toolkit Contributing author to the Enhanced Sampling Toolkit at `https://github.com/jtempkin/enhanced_sampling_toolkit`, a software package designed for rapid prototyping and implementation of enhanced sampling algorithms. Designed and developed the walker API, which provides an easy interface for controlling molecular dynamics simulations.

PyGElib Author of PyGELib `https://github.com/ehthiede/python-pygelib`, a pytorch interface to routines in the GELib package for group equivariant networks. Currently in development.

Autobahn Coauthor of the Autobahn repository `https://github.com/risilab/Autobahn`, an implementation of a specific Autobahn network using the pytorch, pytorch-lightning, pytorch-geometric, and hydra frameworks.