Matey N. Neykov, Ph.D.

2006 Sheridan Road, Evanston, IL 60208

mneykov@northwestern.edu

Current Position

Assistant Professor, Department of Statistics and Data Science — Evanston IL
Northwestern University — **2023 –**

Education

Harvard University — Cambridge MA

Ph.D., Biostatistics, May 2015
Dissertation: *Three Aspects of Biostatistical Learning Theory*
Advisors: Jun S. Liu, Tianxi Cai
A.M., Biostatistics, May 2012

Sofia University — Sofia Bulgaria

B.S., Applied Mathematics, September 2009

Past Positions

Associate Professor (without tenure), Department of Statistics & Data Science — Pittsburgh PA
Carnegie Mellon University — **2022 – 2023**

Assistant Professor, Department of Statistics & Data Science — Pittsburgh PA
Carnegie Mellon University — **2017 – 2022**

Postdoctoral Research Associate, ORFE, **Princeton University** — Princeton NJ
Mentor: Han Liu — **2015 – 2017**

Research Interests

High Dimensional Inference, Graphical Models, Statistical Machine Learning, Convex Analysis and Optimization, Empirical Processes and Random Matrix Theory, Statistical Applications in Biomedical Settings

Honors and Awards

Princeton University — Princeton NJ
IMS Travel Award, 2016
NIPS Travel Award, 2016
CMACS Travel Award, 2016

Harvard University — Cambridge MA
Certificate of Distinction in Teaching, 2015
IMS Travel Award, 2015
PQG Travel Award, 2014
Robert Balentine Reed Prize, 2012
HUSEC graduate student fellowship, 2011 & 2012

Sofia University — Sofia Bulgaria
Graduated with honors, first in the class, 2009
National Programming Contest for College Students, 2009 (3rd place)

ACM ICPC South-Eastern Europe Regional competition, 2007 (top 15)
ACM ICPC South-Eastern Europe Regional competition, 2006 (HM)
National Programming Contest for College Students, 2006 (3rd place)

Sofia High School of Mathematics Sofia Bulgaria
All Russian Mathematics Olympiad, 2005 (HM)
Balkan Mathematics Olympiad, 2005 (Bronze Medal)
All Russian Mathematics Olympiad, 2004 (HM)
International Mathematics Olympiad Tyimaada, 2004 (Silver Medal)
Balkan Mathematics Olympiad, 2004 (Silver Medal)
International Mathematics Olympiad Tyimaada, 2003 (Silver Medal)
Junior Balkan Mathematics Olympiad, 2001 (Silver Medal)

Papers & Publications

Y. Yi and **M. Neykov** "Non-Asymptotic Bounds for the ℓ_∞ Estimator in Linear Regression with Uniform Noise", *Bernoulli, to appear 2023+*

M. Neykov "On the minimax rate of the Gaussian sequence model under bounded convex constraints", *IEEE Transactions on Information Theory, to appear 2022+*

I. Kim, **M. Neykov**, S. Balakrishnan and L. Wasserman "Local permutation tests for conditional independence", *The Annals of Statistics, to appear*, 2022+

M. Li, **M. Neykov** and S. Balakrishnan "Minimax Optimal Conditional Density Estimation under Total Variation Smoothness", *Electronic Journal of Statistics*, 2022

Y. Zhang, Molei Liu, **M. Neykov** and T. Cai, "Prior Adaptive Semi-supervised Learning with Application to EHR Phenotyping", *Journal of Machine Learning Research (JMLR)*, 2022

M. Neykov, S. Balakrishnan, L. Wasserman, "Minimax Optimal Conditional Independence Testing", *The Annals of Statistics*, 2021, *https://arxiv.org/pdf/2001.03039.pdf*

Y. Cao, **M. Neykov**, H. Liu, "High-Temperature Structure Detection in Ferromagnets", *Information and Inference: A Journal of the IMA, 2020+*, *arXiv preprint, arXiv:1809.08204*

M. Neykov and H. Liu, "Property testing in high dimensional Ising models", *The Annals of Statistics*, 2019, *arXiv preprint, arXiv:1709.06688*

M. Neykov "Isotonic Regression Meets LASSO", *Electronic Journal of Statistics*, 2019

Z. Yang, L. F. Yang, E. X. Fang, T. Zhao, Z. Wang, **M. Neykov**, "Misspecified Nonconvex Statistical Optimization for Sparse Phase Retrieval", *arXiv preprint, arXiv:1510.08986*, *Mathematical Programming Series B*, 2019

M. Neykov, "Tossing Coins Under Monotonicity", *Proceedings of Machine Learning Research Volume 89, AISTATS*, 2019

M. Neykov, "Gaussian Regression with Convex Constraints", *Proceedings of Machine Learning Research Volume 89, AISTATS*, 2019

M. Neykov, J. Lu and H. Liu, "Combinatorial inference for graphical models", *The Annals of Statistics*, 2019, *arXiv preprint, arXiv:1608.03045*

M. Neykov, Y. Ning, J. S. Liu and H. Liu, "A unified theory of confidence regions and testing for high dimensional estimating equations", *Statistical Science*, 2018, *arXiv preprint, arXiv:1510.08986*

M. Neykov, Z. Wang and H. Liu, "Agnostic estimation for misspecified phase retrieval models", *Journal of Machine Learning Research (JMLR)*, 2020, *Advances in Neural Information Processing Systems (NIPS)*, 2016 (short version)

M. Neykov, Q. Lin and J. S. Liu, "Signed support recovery for single index models in high-dimensions", *Annals of Mathematical Sciences and Applications*, 2016

M. Neykov, J. S. Liu and T. Cai, "On the characterization of a class of Fisher-consistent loss functions and its application to boosting", *Journal of Machine Learning Research (JMLR)*, 2016

M. Neykov, J. S. Liu and T. Cai, "L_1–regularized least squares for support recovery of high dimensional single index models with Gaussian designs", *Journal of Machine Learning Research (JMLR)*, 2016

M. Neykov, B. Hejblum and J. Sinnott, "Kernel machine score test for pathway analysis in the presence of semi-competing risks", *Statistical Methods in Medical Research*, 2016

R. Payne*, **M. Neykov***, M. K. Jensen and T. Cai, "Kernel machine testing for risk prediction with stratified case cohort studies", *Biometrics*, 2015; **equal contribution*

S. Yu, K.K. Kumamaru, E. George, R.M. Dunne, A. Bedayat, **M. Neykov**, A.R. Hunsaker, K.E. Dill, T. Cai, and F.J. Rybicki "Classification of CT pulmonary angiography reports by presence, chronicity, and location of pulmonary embolism with natural language processing", *Journal of Biomedical Informatics*, 2014

Under Review

M. Neykov, L. Wasserman, I. Kim, S. Balakrishnan "Nearly Minimax Optimal Wasserstein Conditional Independence Testing", 2023

I. Kim, **M. Neykov**, S. Balakrishnan , L. Wasserman "Conditional Independence Testing for Discrete Distributions: Beyond χ^2- and G-tests", 2023

Y. Yi and **M. Neykov** "A New Perspective on Debiasing Linear Regressions", 2021

Y. Yi and **M. Neykov** "Non-Sparse PCA in High Dimensions via Cone Projected Power Iteration", 2020

S. Shrotriya and **M. Neykov** "Adversarial Sign-Corrupted Isotonic Regression", 2022

S. Shrotriya and **M. Neykov** "Revisiting Le Cam's Equation: Exact Minimax Rates over Convex Density Classes", 2022

Unpublished Manuscripts

J. Lu, **M. Neykov** and H. Liu, "Adaptive inferential method for monotone graph invariants", *arXiv preprint, arXiv:1707.09114*, 2017

A. Chakrabortty, **M. Neykov**, R. Carroll and T. Cai, "Surrogate aided unsupervised recovery of sparse signals in single index models for binary outcomes", *arXiv preprint, arXiv:1701.05230*, 2016

Software

R package `kernscr`; Kernel Machine Score Test for Pathway Analysis in the Presence of Semi-Competing Risks; Available on [CRAN]

Academic Experience

Princeton University — Princeton NJ
Postdoctoral Research Associate, Mentor: Han Liu — **2015 – 2017**

Harvard University — Cambridge MA
Dissertation Research, Advisors: Jun S. Liu, Tianxi Cai — **2010 – 2015**

Teaching Experience

Carnegie Mellon University — Pittsburgh PA

Instructor: Advanced Statistical Theory I, 36-709, Spring 2023

Instructor: Advanced Statistical Inference I (MINI), 36-747, Fall 2022

Instructor: Advanced Statistical Inference II (MINI), 36-748, Fall 2022

Instructor: Multivariate Analysis I (MINI), 36-755, Spring 2022

Instructor: Multivariate Analysis II (MINI), 36-756, Spring 2022

Instructor: Advanced Statistical Theory I, 36-709, Spring 2021

Instructor: Introduction to Statistical Inference, 36-226, Spring 2020

Co-Instructor: Advanced Statistical Theory II, 36-710, Fall 2019

Instructor: Statistical Graphics and Visualization, 36-315, Fall 2017, 2018, Spring 2018, Fall 2020

Harvard University — Cambridge MA
Instructor:

Linear Algebra and Real Analysis, *Math camp for incoming PhD Students* — **August, 2014**
Key Concepts: Linear Operators, Hilbert Spaces, Spectral Theorem, Fundamental Theorems of Calculus, Convergent Sequences & Series

Probability Theory, *Math camp for incoming PhD Students* — **August, 2013**
Key Concepts: Elementary Set & Measure Theory, Combinatorics, Discrete & Continuous Random Variables, CLT, LLN, Concentration Inequalities

Problem Solving in Advanced Statistics, *PhD level class* — **Fall, 2012; Spring, 2013**
Key Concepts: Coach for the theory part of the Qualifying Exam

Teaching Assistant:

Statistical Inference II, *PhD level class* — **Spring, 2015**
Professors: Andrea Rotnitzky & Giovanni Parmigiani
Key Concepts: Semi-parametric & Decision Theory, Influence Functions, Minimax & Bayesian Estimation and Hypothesis Testing

Statistical Inference II, *PhD level class* — **Spring, 2014**

Professor: Tianxi Cai
Key Concepts: Asymptotics of M-Estimation, Kernel Smoothing, U-Statistics, Empirical Processes

Analysis of Multivariate and Longitudinal Data, *PhD level class* **Spring, 2013**
Professor: Xihong Lin
Key Concepts: Multivariate Analysis, Generalized Linear Models, Linear & Generalized Linear Mixed Models, Generalized Estimating Equations

Analysis of Rates and Proportions, *Master of Public Health level class* **Spring, 2011**
Professor: Robert Glynn
Key Concepts: One and two sample T-tests, ANOVA, MANOVA

Sofia University Sofia Bulgaria
Teaching Assistant: **Fall, Spring, 2009; 2010**
Professors: Leda Minkova, Marusia Bozhkova
Led the exercise section of the Theory of Probability (for Math/Stat majors), Probability and Statistics (for CS majors) courses. Taught in class, proposed problems for examinations and graded the students.

Conference Presentations

Invited Talks

"Isotonic Regression Meets LASSO"
- CMStatistics, Virtual Conference, Dec 2020

"High Temperature Structure Detection in Ferromagnets"
- DSSV, Virtual Conference, Jul 2020
- SDSS, Virtual Conference, Jun 2020

"Minimax optimal conditional independence testing"
- University of Pittsburgh, Feb 2020
- University of Florida, Nov 2019

"Property testing in high dimensional Ising models"
- Workshop on Graphical Models, Columbia University, Oct 2019
- University of Geneva, Nov 2018

"High Dimensions, Inference and Combinatorics. A Journey Through the Data Jungle"
- Rutgers University, Feb 2017
- Cornell University, Feb 2017
- University of Wisconsin, Madison, Feb 2017
- Carnegie Mellon University, Feb 2017
- University of Maryland, College Park, Jan 2017
- Rice University, Jan 2017
- EPFL, Jan 2017, Lausanne Switzerland
- McGill University, Jan 2017, Montreal QC
- UIUC, Jan 2017
- Florida State University, Jan 2017

"Agnostic estimation for misspecified phase retrieval models"

- IMS 18th Meeting of New Researchers in Statistics and Probability, University of Wisconsin-Madison, Jul 2016

"On the characterization of a class of Fisher-consistent loss functions and its application to boosting for hierarchical outcomes",
- Harvard University, Oct 2013

Contributed Talks & Posters

"Agnostic estimation for misspecified phase retrieval models"
- NeurIPS, Barcelona, Spain, Dec 2016
- Cornell Day of Statistics, Cornell University, Sep 2016

"Structure testing for sparse high dimensional graphical models: lower bounds and algorithms",
- JSM, Chicago Il, Aug 2016

"A unified theory of confidence regions and testing for high dimensional estimating equations",
- ICORS, Geneva Switzerland, Jul 2016
- ENAR, Austin TX, Mar 2016

"L_1–regularized least squares for support recovery of high-dimensional single index models with Gaussian designs",
- IOS, Princeton NJ, Mar 2016

"SVM with bootstrap for soft clustering of populations",
- ENAR, Baltimore MD, Mar 2014

"On the characterization of a class of Fisher-consistent loss functions and its application to boosting for hierarchical outcomes",
- JSM, Montreal Canada, Aug 2013
- NESS, Storrs CT, April 2013

"Kernel machine based testing procedure for assessing the overall effect of multiple markers on the risk of developing a clinical disease",
- Harvard University, Aug 2011

Grants Experience

- NSF DMS-2113684, $250,000 (co-PI with S.Balakrishnan and L. Wasserman)

Student Supervision

Graduate Students
- Yufei Yi (primary thesis advisor)
- Shamindra Shrotriya (primary thesis advisor)
- Raghav Bansal (ADA project)

Undergraduate Students
- Michael Li (worked on conditional density estimation)

Professional Affiliations

Memberships

Institute of Mathematical Statistics, American Statistical Association

Referee Service

The Annals of Statistics; Journal of the American Statistical Association; Journal of the Royal

Statistical Society Series B; Biometrika; Electronic Journal of Statistics; Statistical Science; Mathematical Reviews; Journal of Machine Learning Research; Machine Learning Journal; IEEE Journal on Selected Areas of Information Theory; Annales de l'Institut Henri Poincaré (B) Probabilités et Statistique; Bernoulli; NIPS 2016, 2017, 2018; COLT 2018, 2019;

Session Chair
High-dimensional statistics, IMS Contributed Papers, JSM 2016

Professional Experience

Multimedia Solutions Sofia Bulgaria

Researcher and Software Developer (part-time, internship) **2007 – 2009**

Improved significantly the performance of OpenCV's implementation of Viola Jones' algorithm for face detection with on arm processors; Worked on projects for Bayesian networks, continuous autofocus, image stitching and stabilizing images for CMOS sensors using homotopy and ransac algorithm.

Languages

Bulgarian – native; English – fluent; Russian – limited working proficiency.

Computer Skills

- Statistical Packages: R, SPSS, experience with SAS
- Languages: C/C++, Java, MATLAB, Python, Ruby
- Applications: Ruby on Rails, LaTeX
- Operating Systems: Linux, Mac OS, Windows