

JAKE M. HOFMAN

175 W 79th St #15B
New York, NY 10024
646.831.7272

jmh2045@columbia.edu
<http://jakehofman.com>

ACADEMIC DEGREES

Columbia University	Physics	Ph.D. 2008
Columbia University	Physics	M. Phil. 2007
Columbia University	Applied Physics	M.S. 2004
Boston University	Electrical Engineering	B.S. 2002

ACADEMIC EXPERIENCE

Columbia University, New York, NY

Adjunct Assistant Professor

2012

- Designed and taught *Data-driven modeling*, an applied data mining and machine learning course in the Department of Applied Mathematics

Columbia University, New York, NY

Research Assistant

2003-2008

- Studied machine learning and statistical inference for analysis of biological network and image data, with emphasis on variational techniques and complexity control for generative models
- Developed open source software platform to automate characterization of cell spreading and migration

Instructor

2006

- Created and taught *Concepts in Biological Physics*, Science Honors Program course for advanced high school students

Graduate Teaching Fellow

2003-2005

- Instructor for introductory physics laboratory for pre-medical students

Boston University, Boston, MA

Research Assistant

2001-2002

- Designed and implemented plug & play auto-alignment system for maximal coupling of entangled photons into dual single-mode fiber channels
- Developed correlated data acquisition system for real-time simultaneous detection of spatially distinct photo-events

Tutor

2001-2002

- One of four engineering undergraduates selected to assist peers with assignments for required physics, mathematics, and engineering courses

Teaching Assistant

2001

- Lectured for *Introduction to Robotics* recitation

PROFESSIONAL EXPERIENCE

Yahoo!, Inc., New York, NY

Research scientist

2008-Present

- Applying methods in machine learning and statistical inference for analysis of social data

Advantage Testing, New York, NY

Tutor

2003-2008

- Tutored individual high school students for academic courses and standardized tests in mathematics and physics

Intel Corporation, Folsom, CA

Applications Engineer (8 month co-op)

2000

- Developed automated test environment for platform component validation
- Performed schematic capture for memory subsystems of customer reference motherboards

Boston University Engineering IT, Boston, MA

Computer Consultant/Analyst

2000-2001

- Installed and maintained computer and network systems for engineering faculty and administration

PUBLICATIONS

2012

WHO DOES WHAT ON THE WEB: A LARGE-SCALE STUDY OF BROWSING BEHAVIOR. *4th International Conference on Web Search and Data Mining*. 2012. M. I. Sirer, J. M. Hofman, and S. Goel.

2011

WHO SAYS WHAT TO WHOM ON TWITTER. *20th International World Wide Web Conference*. 2011. S. Wu, J. M. Hofman, W. A. Mason, and D. J. Watts.

EVERYONE'S AN INFLUENCER: QUANTIFYING INFLUENCE ON TWITTER. *4th International Conference on Web Search and Data Mining*. 2011. E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts.

2010

PREDICTING CONSUMER ACTIVITY WITH WEB SEARCH. *Proceedings of the National Academy of Sciences*. 2010 (to appear). S. Goel, J. M. Hofman, S. Lahaie, D. M. Pennock, and D. J. Watts.

INFERRING RELEVANT SOCIAL NETWORKS FROM INTERPERSONAL COMMUNICATION. *19th International World Wide Web Conference*. 2010. M. De Choudry, W. A. Mason, J. M. Hofman, and D. J. Watts.

FILTERING: A TECHNIQUE FOR COUNTING TRIANGLES, COMPUTING MSTs AND CLUSTERING IN MAPREDUCE. *Second Workshop on Massive Data Algorithmics*. 2010. J. M. Hofman, S. Suri, and S. Vassilvitskii

2009

CHARACTERIZING INDIVIDUAL COMMUNICATION PATTERNS. *15th ACM SIGKDD Conference*. 2009. R. Dean Malmgren, J. M. Hofman, L. A. N. Amaral, and D. J. Watts.

CENTMAIL: RATE LIMITING VIA CERTIFIED MICRO-DONATIONS. *18th World Wide Web Conference Developers Track*. 2009. S. Goel, J. M. Hofman, J. Langford, D. Pennock and D. Reeves.

LEARNING RATES AND STATES FROM BIOPHYSICAL TIME SERIES: A BAYESIAN APPROACH TO MODEL SELECTION AND SINGLE-MOLECULE FRET DATA. *Biophysical Journal*. 2009. J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, Jr, and C. H. Wiggins.

ALLOSTERIC COLLABORATION BETWEEN ELONGATION FACTOR G AND THE RIBOSOMAL L1 STALK DIRECTS TRNA MOVEMENTS DURING TRANSLATION. *Proceedings of the National Academy of Sciences*. 2009. J. Fei, J. E. Bronson, J. M. Hofman, R. L. Srinivas, C. H. Wiggins and R. L. Gonzalez, Jr.

2008

A BAYESIAN APPROACH TO NETWORK MODULARITY. *Physical Review Letters*, 100(25):258701, 2008. J. M. Hofman and C. H. Wiggins.

QUANTIFICATION OF EDGE VELOCITIES AND TRACTION FORCES REVEALS DISTINCT MOTILITY MODULES DURING CELL SPREADING. *PLoS ONE*, 2008. Accepted for publication. B. J. Dubin-Thaler, J. M. Hofman, H. S. Xenias, I. Spielman, A. V. Shneidman, L. A. David, H. G. Döbereiner, and C. H. Wiggins, and M. P. Sheetz.

2007

OPPOSING EFFECTS OF PKC θ AND WASP ON SYMMETRY BREAKING AND RELOCATION OF THE IMMUNOLOGICAL SYNAPSE. *Cell*, 129(4):773–85, 2007. T. N. Sims, T. J. Soos, H. S. Xenias, B. Dubin-Thaler, J. M. Hofman, J. C. Waite, T. O. Cameron, V. K. Thomas, R. Varma, C. H. Wiggins, M. P. Sheetz, and D. R. Littman, and M. L. Dustin.

2006

LATERAL MEMBRANE WAVES CONSTITUTE A UNIVERSAL DYNAMIC PATTERN OF MOTILE CELLS. *Physical Review Letters*, 97(3):038102, 2006. H.G. Döbereiner, B. J. Dubin-Thaler, J. M. Hofman, H. S. Xenias, T. N. Sims, G. Giannone, M. L. Dustin, and C. H. Wiggins, and M. P. Sheetz.

THE SMALL GTPASE R-RAS REGULATES ORGANIZATION OF ACTIN AND DRIVES MEMBRANE PROTRUSIONS THROUGH THE ACTIVITY OF PLC- ϵ . *Journal of Cell Science*, page 4364, 2006. A. S. Ada-Nguema, H. Xenias, J. M. Hofman, C. H. Wiggins, and M. P. Sheetz, and P. J. Keely.

NONMUSCLE MYOSIN IIA-DEPENDENT FORCE INHIBITS CELL SPREADING AND DRIVES F-ACTIN FLOW. *Biophysical Journal*, 91(10):3907–3920., 2006. Y. Cai, N. Biais, G. Giannone, M. Tanase, G. Jiang, J. M. Hofman, C. H. Wiggins, P. Silberzan, A. Buguin, and B. Ladoux, and M. P. Sheetz.

PRESENTATIONS AND CONFERENCES

2011

INFERRING THE STRUCTURE AND SCALE OF MODULAR NETWORKS. *Invited presentation*, Department of Computer Science, Rutgers University. New Brunswick, NJ.

LEARNING FROM ONLINE ACTIVITY. *Invited presentation*, Department of Statistics, UCLA, Los Angeles, CA.

DATA BOOTCAMP. *Invited tutorial*, Strata 2011. Santa Clara, CA.

2010

LARGE-SCALE SOCIAL MEDIA ANALYSIS WITH HADOOP. *Invited tutorial*, ICWSM 2010. Washington, D.C.

INFERRING THE STRUCTURE AND SCALE OF MODULAR NETWORKS. *Invited presentation*, Information Theory and Applications 2010. San Diego, CA.

AN INTRODUCTION TO BAYESIAN INFERENCE. *Invited presentation*, NYC Machine Learning Meetup. New York, NY.

2009

A BAYESIAN APPROACH TO COMMUNITY DETECTION. *Invited Presentation*, Department of Statistics Student Seminar, Columbia University, New York, NY.

SOCIAL NETWORK ANALYSIS WITH HADOOP. *Presentation*, Hadoop World 2009. New York, NY.

BAYESIAN INFERENCE: PRINCIPLES AND PRACTICE. *Invited presentation*, NYC R Meetup. New York, NY.

2008

COMMUNITY DETECTION: MODEL FITTING, COMPARISON, AND UTILITY. *Invited presentation*, Workshop on Statistical Inference for Complex Networks. Santa Fe, NM.

BAYESIAN STATISTICS AND MASSIVE DATA STREAMS. *Invited session chair*, 11th Annual Japanese-American Kavli Frontiers of Science Symposium. Irvine, CA.

INFERRING THE STRUCTURE AND SCALE OF MODULAR NETWORKS. *Best student presentation*, 6th International Workshop on Mining and Learning with Graphs (MLG). Helsinki, Finland.

A BAYESIAN APPROACH TO NETWORK MODULARITY. *Presentation*, American Physical Society, March Meeting, New Orleans, LA.

A QUANTITATIVE ANALYSIS OF CELL MOTILITY. *Poster*, Workshop on Bio-Image Informatics, Santa Barbara, CA.

2007

A BAYESIAN APPROACH TO NETWORK MODULARITY. *Poster spotlight presentation*, Neural Information Processing Systems conference workshop on statistical models of networks. Whistler, British Columbia.

A BAYESIAN APPROACH TO NETWORK MODULARITY. *Invited presentation*, Princeton University. Princeton, NJ.

SIMPLE MATH FOR A COMPLEX WORLD: RANDOM WALKS IN BIOLOGY AND FINANCE. *Invited presentation*, Dalton High School, New York, NY.

A BAYESIAN APPROACH TO NETWORK MODULARITY . *Poster*, New York Academy of Sciences 2nd annual Machine Learning Symposium. New York, NY.

A BAYESIAN APPROACH TO NETWORK MODULARITY . *Research seminar*, Department of Applied Mathematics and Applied Physics, Columbia University, New York, NY.

MACHINE LEARNING TUTORIAL. *Invited presentation*, Boulder Biophysics Summer School, Boulder, CO.

BIOIMAGE INFORMATICS: QUANTITATIVE ANALYSIS OF CELL MOTILITY. *Invited presentation*, Instituto Gulbenkian de Ciência, Oeiras, Portugal.

2006

ARCLENGTH PARAMETERIZATION AND OPTICAL FLOW CALCULATION OF MEMBRANE VELOCITY REVEALS ROLE OF ACTIN IN THE PHASES OF CELL SPREADING. *Poster*, Frontiers in Live Cell Imaging, National Institutes of Health, Bethesda, MD.

ACTIVITIES

2011

Program committee member for World Wide Web (WWW) 2012, Web Search and Datamining (WSDM) 2012, Knowledge Discovery and Datamining (KDD) 2011, International Conference on Weblogs and Social Media (ICWSM) 2011,

Reviewer for Transactions on Knowledge and Data Engineering (TKDE), Transactions on Knowledge Discovery from Data (TKDD), Physical Review Letters (PRL), Very Large Databases (VLDB)

Organizer for the National Academy of Sciences Japanese-American Frontiers of Science (JAFOS) Meeting 2012

2010

Program committee member for World Wide Web (WWW) 2011

Reviewer for Transactions on Knowledge and Data Engineering (TKDE), Journal of Statistical Computation and Simulation (JSCS), Statistical Modelling

Organizer for the National Academy of Sciences Japanese-American Frontiers of Science (JAFOS) Meeting 2011

2009

Program committee member for Web Search and Data Mining (WSDM) 2010, World Wide Web (WWW) 2010, Conference on Information and Knowledge Management (CNIKM) 2010 Workshop on complex networks

Reviewer for Neural Information Processing (NIPS) 2010, Proceedings of the National Academy of Sciences (PNAS), Physics Review Letters & Physics Review E, Annals of Applied Statistics (AOAS), Europhysics Letters, Bioinformatics

Organizer for the National Academy of Sciences Japanese-American Frontiers of Science (JAFOS) Meeting 2010

2008

Co-organized Neural Information Processing (NIPS) 2008 workshop *Analyzing Graphs: Theory and applications*

Released *vbmod*, an open source project for identifying modules (i.e. communities) in networks
(vbmod.sourceforge.net)

2007

Founded seminar on computing and research tools in Columbia University's Applied Mathematics and Applied Physics Department

2006

Released *noVel*, an open source project for quantitative analysis of cell motility
(cellmap.sourceforge.net)

Built, administered, and maintained 32-core Linux computing cluster, bling.apam.columbia.edu, for Wiggins Lab computation

2005

Designed and taught preparatory course for incoming graduate students taking Columbia University's Physics Department Ph.D. qualifying exam

Selected to participate in admissions committee for Columbia University's Physics Ph.D program

COURSEWORK

Columbia University, New York, NY

GPA 3.87

Graduate courses

2002-2008

- Advanced Machine Learning, Scientific Computing, Introduction to Biophysical Modeling, Statistical Mechanics, Thermal and Statistical Physics, Quantum Mechanics I & II, Electromagnetics II, Geometric Concepts in Physics, Solid State Physics I, Seminar in Quantum Electronics, Applied Partial Differential Equations, General Relativity, Classical Mechanics

Boston University, Boston, MA

GPA 3.74

Graduate courses

2001-2002

- Quantum Mechanics I & II, Electromagnetics II

Undergraduate courses

1998-2002

- Introductory C Programming, Introductory HPVEE Programming, Introduction to Ethics, Electric Circuit Theory, Hume, Kant & 18th Century Philosophy, Introduction to Electronics, Introduction to Digital Logic, Linear Signals and Systems, Advanced C++ Data Structures, Modern Physics, Introduction to Electromagnetics, Analog Electronics, Philosophy of Science, DSP, Chaos Theory, Introduction to Photonics, History of Mathematics

TECHNICAL SKILLS

Programming Languages

MapReduce/Hadoop, Python, MATLAB, R, C, C++, Bash, Perl, PHP, MySQL, \LaTeX , HTML, XML, Javascript, Java, Pig, Applescript, x86 Assembly, LabView, LabWindows, HPVEE

Operating Systems

Mac OS X, Linux/Unix (Debian, Gentoo, RedHat, Rocks cluster), Windows (all versions), DOS

Software tools

Sun Grid Engine, Apache HTTP Server, AutoCAD, ConceptHDL, ViewDraw, Pspice, Open Office, Microsoft Office

DISTINCTIONS

Recipient of Best Student Presentation Award, MLG 2008

Recipient of Selvaggi Award in physics, Columbia University 2007

Graduated Summa Cum Laude, Boston University 2002

Tau Beta Pi Engineering Honor Society Nominee, Boston University 2002

Golden Key National Honors Society, Boston University 2002

Deans list, all semesters, Boston University 1998-2002