

Aritra Bose

Computational Biology Center
IBM T.J. Watson Research Center
1101 Kitchawan Road
Yorktown Heights, NY 10598

Email: a.bose@ibm.com, bosearitra08@gmail.com
Web: <https://aritra90.github.io/>
Phone: +1(518) 522-7975

Research Interests Statistical Genetics, Machine Learning, Computational Biology, Bioinformatics, Data Mining, Randomized Numerical Linear Algebra.

Education **Purdue University** West Lafayette, IN, USA
Ph.D. in *Computer Science* Aug 2016 - Aug 2019
Advisor: Prof. Petros Drineas

Rensselaer Polytechnic Institute Troy, NY, USA
M.S. in *Computer Science* Aug 2014 - Jul 2016

West Bengal University of Technology Kolkata, WB, India
B.Tech in *Information Technology* Aug 2009 - Jun 2013

Experience **Postdoctoral Researcher**, IBM T.J. Watson Research Center Yorktown, NY, USA
Sep 2019 - Present

Research Assistant, Purdue University West Lafayette, IN, USA
Aug 2016 - Aug 2019

Research Intern, IBM T.J. Watson Research Center Yorktown Heights, NY, USA
May 2018 - Jul 2018

Research Intern, IBM T.J. Watson Research Center Yorktown Heights, NY, USA
May 2017 - Aug 2017

Research Intern, IBM T.J. Watson Research Center Yorktown Heights, NY, USA
May 2016 - Aug 2016

Research Assistant, Rensselaer Polytechnic Institute Troy, NY, USA
Aug 2015 - May 2016

Teaching Assistant, Rensselaer Polytechnic Institute Troy, NY, USA
Aug 2014 - May 2015

Analyst, Teradata Corporation Hyderabad, India
Oct 2013 - April 2014

Research Trainee, CoE in Bioinformatics, Bose Institute Kolkata, India
Sep 2012 - Oct 2013

Summer Intern, Indian Institute of Technology Guwahati, India
May 2012 - Jul 2012

Winter Intern, Indian Statistical Institute Kolkata, India
Dec 2011 - Mar 2012

Publications

Journals

1. **A. Bose**, M.C. Burch, A. Chowdhury, P. Paschou, P. Drineas, *Structure informed clustering adjusts for population stratification in association studies*, BioRxiv, 2020.
2. **A. Bose**, F. Utro, D.E. Platt, L. Parida, *Multiple Loci Selection with Multi-way Epistasis in Coalescence with Recombinations*, Under review.
3. **A. Bose**, D.E. Platt, L. Parida, P. Paschou, P. Drineas, *Integrating linguistics, social structure, and geography to model genetic diversity within India*, Mol. Biol. Evol., to appear.
4. **A. Bose**, V. Kalantzis, E. Kontopoulou, M. Elkady, P. Paschou, P. Drineas, *TeraPCA: a fast and scalable software package to study genetic variation in tera-scale genotypes*, Bioinformatics, Volume 35, Issue 19, 1 October 2019, Pages 36793683.
5. **A. Bose**, D.E. Platt, L. Parida, P. Paschou, P. Drineas, *Dissecting Population Substructure in India via Correlation Optimization of Genetics and Geodemographics*, BioRxiv, 2017.
6. G. Stamatoyannopoulos, **A. Bose**, A. Teodasiadis, F. Tsetsos, A. Plantiga, N. Psatha, N. Zogas, E. Yannaki, P. Zalloua, K.K. Kidd, B.L. Browning, J. Stamatoyannopoulos, P. Paschou, P. Drineas, *Genetics of the Peloponnesean populations and the theory of the extinction of the medieval Peloponnesean Greeks*, Eur. J. Hum. Genet., 25(5), pp. 637-645, 2017.
7. S.Hassan, P. Pal Choudhury and **A. Bose**,(2011), *A Quantitative model for Human Olfactory Receptors*, Nature Precedings, npre20126967-2, 2012.

Conferences

8. **A. Bose**, M.C. Burch, A. Chowdhury, P. Paschou, P. Drineas, *CluStrat: a structure informed clustering strategy for population stratification*, In International Conference on Research in Computational Molecular Biology (pp. 234-236). Springer, Cham.

Abstracts (peer reviewed only)

9. **A. Bose**, M.C. Burch, A. Chowdhury, P. Paschou, P. Drineas, *Structure informed clustering for population stratification and genetic risk prediction*, Annual Meeting of the American Society of Human Genetics, 2019.
10. **A. Bose**, F. Utro, D.E. Platt, L. Parida, *Algorithms to modulate ARG by Selection*, RECOMB-Genetics, 2018. **Selected for Platform presentation.**
11. **A. Bose**, V. Kalantzis, E. Kontopoulou, M. Elkady, P. Paschou, P. Drineas, *TeraPCA: a fast and scalable software package to study genetic variation in tera-scale genotypes*, Annual Meeting of the American Society of Human Genetics, 2017.
12. **A. Bose**, D.E. Platt, L. Parida, P. Paschou, P. Drineas, *Correlation Optimization of Genetics and Geodemographics*, Annual Meeting of the American Society of Human Genetics, 2016. **Selected for Platform presentation.**

Dissertation

13. **A.Bose**, *Computational Methods for Population Genetics*, <https://doi.org/10.25394/PGS.9752924.v1>, Purdue University, 2019.

News Articles

1. *Genetic testing has a data problem. New software can help.* Available at: <https://www.purdue.edu/newsroom/releases/2019/Q2/genetic-testing-has-a-data-problem.-new-software-can-help..html>.

Talks and Posters

- **Platform** presentation on *CluStrat: a structure informed clustering strategy for population stratification* at **RECOMB** meeting held virtually on June 2020.
- Poster presentation on *Structure informed clustering for population stratification and genetic risk prediction* at **ASHG 2019** meeting held at Houston, TX.
- **Platform** presentation on *Algorithms to modulate ARG by Selection* at the **RECOMB-Genetics** meeting held in Paris on April, 2018. (This talk is given by Dr. Laxmi Parida)
- Poster presentation on *TeraPCA: A fast and scalable method to study genetic variation in tera-scale genotypes* at **Conference on Scientific Computing and Approximation**, held at Purdue University, IN.
- Poster presentation on *TeraPCA: A fast and scalable method to study genetic variation in tera-scale genotypes* at **ASHG 2017** meeting held at Orlando, FL.
- Poster presentation on *COGG* at Summer Intern Showcase held in IBM T.J. Watson Research Center, NY.
- Poster presentation on *COGG* at **Biology of Genomes (BOG) 2017** meeting held at Cold Spring Harbor Labs, NY.
- **Platform** talk on *COGG (Correlation Optimization of Genetics and Geodemographics)*, a novel method to model sociolinguistic stratification in India at **ASHG 2016**, Annual Meeting in Vancouver, BC, Canada. Abstract selected in top 8% out of over 6000 submissions.
- Presented a poster at Student Research Showcase at Computer Science Department, Purdue University, IN.
- Presented a poster at **BOG 2016** conference held in Cold Spring Harbor Labs in Cold Spring, NY.
- Presented posters in Student Research Symposium at Computer Science Department, Rensselaer Polytechnic Institute.
- **Summer school** on "Mathematics of Data", organized by **Park City Mathematics Institute (PCMI)** and the Institute for Advanced Study (IAS), held in, Midway, Utah, USA.
- **ASHG 2015**, Annual Meeting in Baltimore, MD, USA as a trainee researcher.
- **Gene Golub SIAM Summer School 2015**, held in, Delphi, Greece.

Awards of Merit

- ISCB (International Society for Computational Biology) Travel Fellowship for RECOMB 2020 in Padua, Italy.
- NSF Travel Grants to the following conferences:
 - * Biology of Genomes: 2016 and 2017.
 - * American Society of Human Genetics (ASHG), 2015 - 2019.

* International Conference for Distributed Computing and Internet Technologies (ICDCIT) 2017 meeting held at Bhubaneswar, Odisha, India.

- Received a 4 year fellowship from Ministry of Human Resource Development (M.H.R.D), Government of India for significant achievement in Higher Secondary Examination
- Ranked 1 out of 65 students in Department of Information Technology in Meghnad Saha Institute of Technology, Kolkata.
- Won several school, college and corporate level Quiz competitions across India.

Professional Activities

- Reviewer for the following:
 - Journals
 - American Journal of Medical Genetics; BMC Bioinformatics
 - Conferences
 - KDD; NeurIPs; RECOMB; WABI
- Student member: American Society of Human Genetics, International Society for Computational Biology.
- Peer Adviser to incoming graduate students in the Computer Science Department in Rensselaer Polytechnic Institute as well as in Purdue University.
- Co-Founder of the Robotics club of Meghnad Saha Institute of Technology which has over 400 students now.
- Presented a paper, “Automated Hybrid Bot” in Student’s Convention, 2011 at Indian Statistical Institute, Kolkata, India organized by, Computer Society of India.

Computer Skills

- **Languages:** C, C++, Python, MATLAB, Java, R, AWK, PL/SQL, Scripting(bash,etc), Perl, HTML, LaTeX
- **Operating Systems:** GNU/Linux, Unix, Windows
- **Computational Biology:** GATK, PLINK, GSEA, MeV, samtools, bcftools, PeakSeq, Cytoscape and other computational biology and population genetic tools.
- **Databases:** MySQL, TERADATA, Oracle, DB2

Graduate Coursework (selected)

Machine Learning, Computational Linear Algebra, Parallel Computing, Foundations of Data Science, Algorithms Design, Frontiers of Network Science, Distributed Systems, Randomized Algorithms, Theory of Computation

Independent Coursework

Coursera.org: Algorithms: Design and Analysis; Bioinformatics I and II; Python for Genomic Data Science

References

Prof. Petros Drineas
Professor
Department of Computer Science
Purdue University
West Lafayette, IN, USA
drineas@gmail.com

Prof. Peristera Paschou
Associate Professor
Department of Biological Sciences
Purdue University
West Lafayette, IN, USA
ppaschou@gmail.com

Dr. Laxmi Parida
IBM Fellow & Manager,
Computational Genomics,
IBM T.J. Watson Research Center
Yorktown Heights, NY, USA
parida@us.ibm.com

**Additional
Information**

- *Date of Birth:* August 8, 1990
- *Marital Status:* Single
- *Citizenship:* Indian