Seong-Hwan Jun, Ph.D.
Email: seonghwan_jun@urmc.rochester.edu
https://junseonghwan.github.io/

EDUCATION

The University of British Columbia
Ph.D. in Statistics Sep 2013 - Feb 2018
Thesis: Scalable sequential Monte Carlo methods and probabilistic approach to combinatorial problems
Advisors: Alexandre Bouchard-Côté, Ph.D. and James V. Zidek, Ph.D., FRSC

MSc. in Statistics Sep 2011 - Aug 2013
Thesis: Entangled Monte Carlo
Advisor: Alexandre Bouchard-Côté, Ph.D.

University of Waterloo
Bachelor of Mathematics, Honours, Co-op Jan 2004 - May 2009
Major: Computer Science
Minor: Combinatorics and Optimization

EMPLOYMENT

University of Rochester Medical Center
Assistant Professor Jan 2024 - Current
Department of Biostatistics and Computational Biology

University of Rochester Medical Center
Research Assistant Professor Oct 2022 - Dec 2023
Department of Biostatistics and Computational Biology

Fred Hutchinson Cancer Research Center
Postdoctoral Research Fellow Jan 2020 - Sep 2022
Computational Biology Program, Public Health Sciences Division
Mentors: Frederick Matsen, Ph.D. and Raphael Gottardo Ph.D.

Science for Life Laboratory, KTH Royal Institute of Technology
NSERC Postdoctoral Research Fellow Nov 2017 - Nov 2019
Department of EECS
Mentor: Jens Lagergren, Ph.D.

REFEREED PUBLICATIONS

11. H. Koptagel, **S-H. Jun**, J. Hård, J. Lagergren. Scuphr: A probabilistic framework for cell lineage tree reconstruction. PLoS Comput. Biol. 20, e1012094 (2024).
10. **S-H. Jun**, H. Nasif, C. Jennings-Shaffer, D. H. Rich, A. Kooperberg, M. Fourment, C. Zhang, M. A. Suchard, F. A. Matsen IV. A topology-marginal composite likelihood via a generalized phylogenetic pruning algorithm. Algorithms Mol. Biol. 18, 10 (2023).
9. D. A. Oyong, F. J. Duffy, M. L. Neal, Y. Du, J. Carnes, K. V. Schwedhelm, N. Hertoghs, **S-H. Jun**, H. Miller, J. D. Aitchison, S. C. De Rosa, E. W. Newell, M. J. McElrath, S. M. McDermott, and K. D. Stuart. Distinct immune responses associated with vaccination status and protection outcomes after malaria challenge. PLoS Pathog. 19, e1011051 (2023).

8. **S-H. Jun**, H. Toosi, J. Mold, C. Engblom, X. Chen, C. O'Flanagan, M. Hagemann-Jensen, R. Sandberg, S. Aparicio, J. Hartman, A. Roth, J. Lagergren, Reconstructing clonal tree for phylo-phenotypic characterization of cancer using single-cell transcriptomics. Nature Communications. 14, 982 (2023).
7. X. Chen, E.G. Sifakis, S. Robertson, S.Y. Neo, **S-H. Jun**, J. Lövrot, V. Jovic, J. Bergh, T. Foukakis, J. Lagergren, A. Lundqvist, R. Ma, and J. Hartman. Breast cancer patient-derived whole-tumor cell culture model for efficient drug profiling and treatment response prediction. Proceedings of the National Academy of Sciences of the United States of America. 120, e2209856120 (2023).
6. M. M. Neyshabouri, **S-H. Jun**, and J. Lagergren. Inferring tumor progression in large datasets. PLOS Computational Biology. 16(10), p.e1008183 (2020).
5. **S-H. Jun**, S. Wong, J. Zidek, and A. Bouchard-Côté. Sequential decision model for inference and prediction on non-uniform hypergraphs with application to knot matching from computational forestry. The Annals of Applied Statistics. 13(3), pp. 1678-1707 (2019).
4. E. Haber, L. Ruthotto, E. Holtham, **S-H. Jun**. Learning across scales - A multiscale method for convolution neural networks. Association for the Advancement of Artificial Intelligence (AAAI). (2018).
   Acceptance rate: 933/3800.
3. **S-H. Jun**, A. Bouchard-Côté, S. Wong, and J. Zidek. Sequential graph matching with sequential Monte Carlo. International Conference on Artificial Intelligence and Statistics (AISTATS). pp. 1075–1084 (2017).
   Acceptance rate: 168/530.
2. **S-H. Jun** and A. Bouchard-Côté. Memory (and time) efficient sequential Monte Carlo. International Conference on Machine Learning (ICML). pp. 514-–522 (2014).
   Acceptance rate: 310/1238.
1. **S-H. Jun**, L. Wang, and A. Bouchard-Côté. Entangled Monte Carlo. Advances in Neural Information Processing Systems 25 (NIPS). pp. 2735–2743 (2012).
   Acceptance rate: 370/1467. Spotlight talk: 72/1467.

TEACHING EXPERIENCE

Spring 2024 – BST 434: Genomics Data Analysis
Instructor
Lectures on bulk and single-cell RNA expression analysis, batch effect correction and normalization, microbiome and phylogenetic analysis.

Spring 2023 – IND 419: Introduction to Quantitative Biology
Lecturer
Lectures on R graphcs using `ggplot2` and differential gene expression analysis.

Fall 2018 – KTH DD2447: Statistical methods in applied computer science
Guest lecturer
Delivered lectures on importance sampling, sequential Monte Carlo, and particle MCMC methods.

2016-2017 –UBC Master of Data Science program
Academic assistant
Developed assignments and lab materials for newly launched master program. Example of topics covered:

- Analyzing Google N-grams using Map-Reduce on Amazon Web Services (AWS).

- Training deep neural network with Tensorflow using GPU instances on AWS.
- Designing an A/B testing for web interface using R Shiny.

Fall 2015 – UBC STAT 300: Intermediate statistics for applications
Head teaching assistant
Topics: Non-parametric tests including Kruskal-Wallis, permutation test, and fisher's exact test.

Winter 2014 – UBC STAT 547: Statistical modelling with stochastic processes
Teaching assistant
Topics: Selected topics in non-parameteric Bayesian methods, continuous time Markov processes, point processes.

Summer and Fall 2012 – UBC STAT 447B/547B: Methods for statistical learning
Course developer and teaching assistant
Developed course materials on Boosting, generalized additive models, splines, regression trees and random forest, LASSO, K-NN classifier.

Winter 2012 – UBC STAT 441 Multivariate statistical methods
Teaching assistant
Topics: Multivariate hypothesis testing and ANOVA, PCA, latent variable analysis, and discriminant analysis.

Fall 2011 – UBC STAT 203: Statistical methods
Head teaching assistant
Introduction to statistics including central limit theorem and hypothesis testing.

PRESENTATIONS

Research in Computational Molecular Biology (RECOMB) Apr, 2024
Boston, MA, USA
Poster presentation: Statistical modeling of microRNA-sequencing data.

The Classification Society Annual Meeting June, 2023
Rochester, NY, USA
Oral presentation: Reconstructing cancer evolution as a Bayesian co-clustering problem.

PhyloMania Nov, 2020
Virtual conference.
Oral presentation: Generalized phylogenetic pruning algorithm.

Probabilistic modelling in genomics Nov, 2018
Cold Spring Harbor, NY, USA
Poster presentation: Reconstruction of tumor phylogeny from single-cells via joint probabilistic analysis of bulk DNA and scRNA-seq.

Conference on machine and other intelligence Sep, 2018
Norrköping, Sweden
Poster and oral presentation: Large scale machine learning for the single cell revolution.

International Conference on AI and Statistics (AISTATS) Oct, 2017
Fort Lauderdale, FL, USA

Poster presentation: Sequential graph matching with sequential Monte Carlo.

Joint Statistical Meeting (JSM) Aug, 2017
Baltimore, MD, USA
Poster presentation: Sequential graph matching and streaming sequential Monte Carlo.

International Conference on Machine Learning July, 2014
Beijing, China
Oral presentation: Memory (and time) efficient sequential Monte Carlo.

Randomized Algorithm Workshop at NeurIPS Dec, 2013
Lake Tahoe, NV, USA
Poster presentation: Using a stochastic map view of sequential Monte Carlo for memory and network efficiency.

Annual Meeting of the Statistical Society of Canada May, 2013
Edmonton, AB, Canada
Poster and oral presentation: Exploring spatial and temporal heterogeneity of environmental noise in Toronto.
Winner of the case study competition.

NeurIPS Dec, 2012
Lake Tahoe, NV, USA
Spotlight talk and poster presentation: Entangled Monte Carlo.

UBC-SFU joint seminar Sep, 2012
Vancouver, BC, Canada
Oral presentation: Importance sampling, sequential importance sampling, and bootstrap particle filter.

PROFESSIONAL SERVICES

Reviewer

- British Journal of Cancer
- BMC Bioinformatics
- International Conference on Research In Computational Molecular Biology (RECOMB).
- Bayesian Analysis
- International Conference on Artificial Intelligence and Statistics (AISTATS)
- Neural Information Processing Systems (NeurIPS)
- International Conference on Machine Learning (ICML)
- International Conference on Learning Representations (ICLR)

Manager of Statistical consulting services 2016 - 2017
Department of Statistics, UBC

- Served in the steering committee.
- Developed operating guidelines for the consulting services.

Senior consultant 2015 - 2017
Department of Statistics, UBC

- Provided statistical advice to graduate students and postdoctoral researchers.
- Recruited and mentored junior consultants.

Graduate student seminar organizer 2014 - 2016
Department of Statistics, UBC

- Invited speakers for weekly seminar.
- Organized lecture series on parallel computing in R, statistical analysis of network data, deep neural networks, and sports analytics.

Academic guide 2013 - 2014, 2015 - 2016
International Graduate Student Preparation Program, UBC

- Cultivate research interests and develop research statements with prospective graduate students.

AWARDS

2024-25	$35,000 USD	URMC CTSI NBEM Pilot Grant
2018-20	$90,000 CAD	NSERC Postdoctoral Fellowship
2017	$850 CAD	CRM Industrial Problem Workshop Travel Award
2017	$1,000 USD	AISTATS Travel Award
2013-17	$18,000 CAD	Faculty of Science Graduate Award (Ph.D)
2014	$500 USD	ICML Travel Award
2013	$500 CAD	SSC Case Study Competition Winner
2011-13	$1,000 CAD	Faculty of Science Graduate Award (MSc.)
2012	$400 USD	NIPS Travel Award
2011	$5,500 CAD	NSERC Undergraduate Student Research Award