Seong-Hwan Jun, Ph.D.
Email: seonghwan_jun@urmc.rochester.edu
http://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
Research Assistant Professor Oct 2022 - Current
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

1. **Jun, S-H**., Toosi, H., Mold, J., Engblom, C., Chen, X., O'Flanaga C., Hagemann-Jensen, M., Sandberg, R., Aparicio, S., Hartman, J., Roth, A., and Lagergren, J., 2023. Reconstructing clonal tree for phylo-phenotypic characterization of cancer using single-cell transcriptomics. *Nature Communications*, 14 (1):982.
2. Chen, X., Sifakis, E.G., Robertson, S., Neo, S.Y., **Jun S-H**., Lövrot, J., Jovic, V., Bergh, J., Foukakis, T., Lagergren, J., Lundqvist, A., Ma, R., and Hartman, J, 2022. Breast cancer patient-derived whole-tumor cell culture model for efficient drug profiling and treatment response prediction. *Proc Natl Acad Sci USA*, 120(1):e2209856120
3. Mohaghegh Neyshabouri, M., **Jun, S-H**., and Lagergren, J., 2020. Inferring tumor progression in large datasets. *PLoS Computational Biology*, 16(10):e1008183.
4. **Jun, S-H**., Wong, S.W., Zidek, J.V., and Bouchard-Côté, A., 2019. 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):1678-1707.

5. Haber, E., Ruthotto, L., Holtham, E., and **Jun S-H**., 2018. Learning across scales – A multiscale method for convolution neural networks. *AAAI Conference on Artificial Intelligence*. arXiv:1703.02009
   Acceptance rate: 933/3800.
6. **Jun, S-H**., Wong, S.W., Zidek, J.V., and Bouchard-Côté, A., 2017. Sequential Graph Matching with sequential Monte Carlo. *International Conference on Artificial Intelligence and Statistics*. 20:1075–1084.
   Acceptance rate: 168/530.
7. **Jun, S-H**. and Bouchard-Côté, A., 2014. Memory (and time) efficient sequential Monte Carlo. *International Conference in Machine Learning*. 31:514–522.
   Acceptance rate: 310/1238.
8. **Jun, S-H**., Wang, L., and Bouchard-Côté, A., 2012. Entangled Monte Carlo. *Advances in Neural Information Processing Systems*. 25:2735–2743.
   Acceptance rate: 370/1467. Spotlight talk: 72/1467.

MANUSCRIPTS UNDER REVIEW

1. **Jun, S-H**., Nasif, H., Jennings-Shaffer, C., Rich, D., Kooperberg, A., Fourment, M., Zhang, C., Suchard, M., and Matsen F.A. A Generalized phylogenetic pruning algorithm. Under review at *Algorithms for Molecular Biology*.
2. Koptagel, H., **Jun, S-H**., Hård, J., and Lagergren, J. SCuPhr: A probabilistic framework for cell lineage tree reconstruction. doi: `https://doi.org/10.1101/357442`. Under review at *PLOS Computational Biology*.

MANUSCRIPTS IN PREPARATION

1. **Jun, S-H**. Phylo-SMC: An R package for large scale phylogenetics inference using particle Markov chain Monte Carlo methods. In preparation for submission to Journal of Statistical Software.
2. **Jun, S-H**., Zheng, Y. and Gottardo, R. CITE-seq tree: immunophenotyping for CITE-seq data. In preparation.
3. Zheng, Y., Jun, S-H., Tian, Y., Mair, F., and Gottardo, R. 2022. Robust Normalization and Integration of Single-Cell Protein Expression across CITE-Seq Datasets. bioRxiv. `https://doi.org/10.1101/2022.04.29.489989`.

TEACHING EXPERIENCE

Winder 2022/2023 – 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

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.

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 mem-

ory 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

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