Sohan Seth

3.12 Bayes Center, 47 Potterrow, Edinburgh EH8 9BT, UK
sohan.seth@ed.ac.uk ⋆ +44 (0) 131 651 3126 ⋆ https://sohanseth.github.io

Education

Doctor of Philosophy, Electrical and Computer Engineering, Jul. 2011
University of Florida, Gainesville, FL, USA

Master of Science, Electrical and Computer Engineering, Dec. 2008
University of Florida, Gainesville, FL, USA

Bachelor of Engineering, Instrumentation and Electronics Eng. Jul. 2005
Jadavpur University, Kolkata, India

Employment

Senior Data Scientist, Apr. 2018 – Present
University of Edinburgh, Edinburgh, UK

Postdoctoral Researcher, Nov. 2014 – Apr. 2018
University of Edinburgh, Edinburgh, UK

Postdoctoral Researcher, Sep. 2011 – Nov. 2014
Helsinki Institute for Information Technology, Espoo, Finland

Research Assistant, Aug. 2006 – Aug. 2011
University of Florida, Gainesville, FL, USA

Junior Research Fellow, Nov. 2005 – Jul. 2006
Jadavpur University, Kolkata, India

Research

Data Science for Health, Interpretable Models, Unsupervised Learning, AI for Social Good,

Funding

National Institute for Health Research, £4.9 M 2021 – 2024
Multimorbidity Clustering in Individuals, Space and Clinical Context, co-I

Scottish Funding Council, £148 K 2020 – 2021
Bottom-up Population Density Estimate from Survey Data and Satellite Imaging, PI

Legal and General, £6.9 M 2020 – 2025
Advanced Care Research Center, co-I

Natural Environment Research Council, £809 K 2020 – 2022
4D Quantification of Micro-scale Feedbacks in Dehydrating, Deforming Rocks, co-I

Scottish Funding Council, £109 K 2020 – 2021
Child Poverty and Access to Services, co-I

Engineering and Physical Sciences Research Council, £4.5 K 2021
Uniting real and reciprocal space with machine learning, co-I

Engineering and Physical Sciences Research Council, £50 K 2020 – 2021
Interpretable Crystal Descriptions Across Length Scales for Materials Discovery, co-I

Advising

Alexandra Adams, Principal Supervisor 2020

Samuel Fielding, Co-supervisor 2021

Syed Muhammad Abbas Rizvi, Co-supervisor 2021

Teaching

Special course on multi view and multitask learning, Fall 2012
Evaluating presentations by students, preparing and grading assignments

Adaptive signal processing, Spring 2010, 2011
Preparation and presentation of lectures when instructor is absent

Neural networks for signal processing, Fall 2009, 2010
Preparation and presentation of lectures when instructor is absent

Invited Talks

Model Criticism in Latent Space, Sep 2019
Universitä Degli Studi Di Napoli Federico II

Probabilistic Archetypal Analysis, Sep 2019
12-th Scientific Meeting Classification and Data Analysis Group

Estimating Bacterial and Cellular Load in FCFM Imaging, Sep 2018
University of Edinburgh Medical Informatics Seminar Series

Visits

University College London Jul. 2013 – Jun. 2014
John Shawe-Taylor

European Bioinformatics Institute Nov. 2013 – Jun. 2014
Alvis Brazma and Ugis Sarkans

Activity

Journal reviewing
Bioinformatics, Entropy,
IEEE Transactions on Cybernetics,
IEEE Transactions on Neural Networks and Learning Systems,
IEEE Transactions on Signal Processing,
International Journal of Environmental Research and Public Health,
International Journal of Pattern Recognition and Artificial Intelligence,
Journal of Neural Processing Letters,
Machine Learning,
Neurocomputing,
Pattern Recognition Letters,
PLoS Computational Biology,
Signal Processing,
Technometrics,
The American Statistician

Conference reviewing

AAAI Conference on Artificial Intelligence (AAAI)	2018,2019
International Conference on Artificial Intelligence and Statistics (AISTATS)	2017, 2018, 2019
European Conference on Machine Learning (ECML)	2016, 2017
European Signal Processing Conference (EUSIPCO)	2014
International Conference on Artificial Neural Networks (ICANN)	2010, 2011
International Conference on Learning Representations (ICLR)	2018, 2019
International Conference on Machine Learning (ICML)	2016, 2017, 2018, 2019
International Joint Conference on Neural Networks (IJCNN)	2010, 2013, 2014
IEEE International Workshop on Machine Learning for Signal Processing (MLSP)	2009
Neural Information Processing Systems (NeurIPS)	2016, 2017, 2018, 2019,2020

Awards

Principal's Medal as part of StopCovid Team, 2021

Cover Story in Journal of Imaging, 2018

University of Florida, Office of Research, travel grant, 2010

Publications

Under review

Jonathan E Millar, Lucile Neyton, **Sohan Seth**, Jake Dunning, Laura Merson, Srinivas Murthy, Clark D Russell, Sean Keating, Maaike Swets, Carole H Sudre, Timothy D Spector, Sebastien Ourselin, Claire J Steves, Jonathan Wolf, ISARIC-4C Investigators, Annemarie B Docherty, Ewen M Harrison, Peter JM Openshaw, Malcolm G Semple, and J. Kenneth Baillie. Robust, reproducible clinical patterns in hospitalised patients with COVID-19. preprint, Infectious Diseases (except HIV/AIDS), August 2020.

Book chapters

Weifeng Liu, Puskal Pokharel, Jianwu Xu, and **Sohan Seth**. Correntropy for Random Variables: Properties and Applications in Statistical Inference. In *Information Theoretic Learning*, pages 385–413. Springer New York, New York, NY, 2010. Series Title: Information Science and Statistics.

Journal papers

Olivia V Swann, Karl A Holden, Lance Turtle, Louisa Pollock, Cameron J Fairfield, Thomas M Drake, **Sohan Seth**, Conor Egan, Hayley E Hardwick, Sophie Halpin, Michelle Girvan, Chloe Donohue, Mark Pritchard, Latifa B Patel, Shamez Ladhani, Louise Sigfrid, Ian P Sinha, Piero L Olliaro, Jonathan S Nguyen-Van-Tam, Peter W Horby, Laura Merson, Gail Carson, Jake Dunning, Peter J M Openshaw, J Kenneth Baillie, Ewen M Harrison, Annemarie B Docherty, and Malcolm G Semple. Clinical characteristics of children and young people admitted to hospital with covid-19 in United Kingdom: prospective multicentre observational cohort study. *BMJ*, page m3249, August 2020.

Katerina Boufea, **Sohan Seth**, and Nizar N. Batada. scID Uses Discriminant Analysis to Identify Transcriptionally Equivalent Cell Types across Single-Cell RNA-Seq Data with Batch Effect. *iScience*, 23(3):100914, March 2020.

Sohan Seth, Iain Murray, and Christopher K. I. Williams. Model Criticism in Latent Space. *Bayesian Analysis*, 14(3):703–725, September 2019.

Sohan Seth, Ahsan R. Akram, Kevin Dhaliwal, and Christopher K. I. Williams. Estimating Bacterial and Cellular Load in FCFM Imaging. *Journal of Imaging*, 4(1):11, January 2018.

Debaditya Choudhury, Michael. G. Tanner, Sarah McAughtrie, Fei Yu, Bethany Mills, Tushar. R. Choudhary, **Sohan Seth**, Tom H. Craven, Jim M. Stone, Ioilia. K. Mati, Colin J. Campbell, Mark Bradley, Christopher K. I. Williams, Kevin Dhaliwal, Timothy A. Birks, and Robert R. Thomson. Endoscopic sensing of alveolar pH. *Biomedical Optics Express*, 8(1):243–259, December 2016.

Sohan Seth, Ahsan R. Akram, Paul McCool, Jody Westerfeld, David Wilson, Stephen McLaughlin, Kevin Dhaliwal, and Christopher K. I. Williams. Assessing the utility of autofluorescence-based pulmonary optical endomicroscopy to predict the malignant potential of solitary pulmonary nodules in humans. *Scientific Reports*, 6:31372, August 2016.

Paul Blomstedt, Ritabrata Dutta, **Sohan Seth**, Alvis Brazma, and Samuel Kaski. Modelling-based experiment retrieval: a case study with gene expression clustering. *Bioinformatics*, 32(9):1388–1394, May 2016.

Sohan Seth and Manuel J. A. Eugster. Probabilistic archetypal analysis. *Machine Learning*, 102(1):85–113, January 2016.

Sohan Seth and Manuel J. A. Eugster. Archetypal Analysis for Nominal Observations. *IEEE transactions on pattern analysis and machine intelligence*, 38(5):849–861, May 2016.

Sohan Seth, Niko Välimäki, Samuel Kaski, and Antti Honkela. Exploration and retrieval of whole-metagenome sequencing samples. *Bioinformatics*, page btu340, May 2014.

Sohan Seth and José C. Príncipe. Learning dependence from samples. *International Journal of Bioinformatics Research and Applications*, 10(1):43, 2014.

Il Memming Park, **Sohan Seth**, Antonio R.C. Paiva, Lin Li, and Jose C. Principe. Kernel Methods on Spike Train Space for Neuroscience: A Tutorial. *IEEE Signal Processing Magazine*, 30(4):149–160, July 2013.

Sohan Seth and José C. Príncipe. Learning dependence from samples. *International Journal of Bioinformatics Research and Applications*, 2014.

Bilal Fadlallah, **Sohan Seth**, Andreas Keil, and José C. Príncipe. Quantifying cognitive state from eeg using dependence measures. *IEEE Transactions on Biomedical Engineering*, 59:2773–81, 2012.

Lin Li, Il Park, Austin Brockmeier, Badong Chen, **Sohan Seth**, Joe Francis, Justin Sanchez, and José C. Príncipe. Adaptive inverse control of neural spatiotemporal spike patterns with a reproducing kernel Hilbert space (RKHS) framework. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2012.

Lin Li, Il Park, **Sohan Seth**, Justin Sanchez, and José C. Príncipe. Functional connectivity dynamics among cortical neurons: A dependence-graph anaysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, pages 18–30, 2012.

Il Park, **Sohan Seth**, Murali Rao, and José C. Príncipe. Strictly positive definite kernels for point process divergence. *Neural Computation*, pages 2223–2250, 2012.

Sohan Seth and José C. Príncipe. Conditional association. *Neural Computation*, pages 1882–1905, 2012.

Sohan Seth and José C. Príncipe. Assessing Granger non-causality using nonparametric measure of conditional independence. *IEEE Transactions on Neural Networks*, 2011.

Sohan Seth, Murali Rao, Il Park, and José C. Príncipe. A unified framework for quadratic measures of independence. *IEEE Transactions on Signal Processing*, 2011.

Murali Rao, **Sohan Seth**, Jianwu Xu, Yunmei Chen, Hemant Tagare, and José C. Príncipe. A test of independence based on a generalized correlation function. *Signal Processing*, 91, January 2011.

Nabarun Bhattacharyya, **Sohan Seth**, Bipan Tudu, Pradip Tamuly, Arun Jana, Devdulal Ghosh, Rajib Bandyopadhyay, Manabendra Bhuyan, and Santanu Sabhapandit. Detection of optimum fermentation time for black tea manufacturing using electronic nose. *Sensors and Actuators B: Chemical*, 122:627–634, March 2007.

Conference papers

Debaditya Choudhury, Michael G Tanner, Sarah McAughtrie, Fei Yu, Bethany Mills, Tushar R Choudhary, **Sohan Seth**, Thomas H Craven, James M Stone, Ioulia K Mati, Colin J Campbell, Mark Bradley, Christopher K I Williams, Kevin Dhaliwal, Timothy A Birks, and Robert R Thomson. Endoscopic sensing of distal lung physiology. *Journal of Physics: Conference Series*, 1151:012009, January 2019.

Sohan Seth, Ahsan R. Akram, Kevin Dhaliwal, and Christopher K. I. Williams. Estimating Bacterial Load in FCFM Imaging. In María Valdés Hernández and Víctor González-Castro, editors, *Medical Image Understanding and Analysis*, pages 909–921, Cham, 2017. Springer International Publishing.

Debaditya Choudhury, Michael G. Tanner, Sarah McAughtrie, Fei Yu, Bethany Mills, Tushar R. Choudhary, **Sohan Seth**, Thomas Craven, James M. Stone, Loulia K. Mati, Colin J. Campbell, Mark Bradley, Christopher K. I. Williams, Kevin Dhaliwal, Timothy A. Birks, and Robert R. Thomson. Endoscopic sensing of pH in the distal lung (Conference Presentation). In Melissa J. Suter, Stephen Lam, and Matthew Brenner, editors, *Optical Techniques in Pulmonary Medicine II*, page 10, San Francisco, United States, April 2017. SPIE.

Sohan Seth, John Shawe-Taylor, and Samuel Kaski. Retrieval of Experiments by Efficient Comparison of Marginal Likelihoods. In Chu Kiong Loo, Keem Siah Yap, Kok Wai Wong, Andrew Teoh, and Kaizhu Huang, editors, *Neural Information Processing*, volume 8835, pages 135–142. Springer International Publishing, Cham, 2014.

Rosha Pokharel, **Sohan Seth**, and Jose C. Principe. Quantized mixture kernel least mean square. In *2014 International Joint Conference on Neural Networks (IJCNN)*, pages 4168–4174, Beijing, China, July 2014. IEEE.

Rosha Pokharel, **Sohan Seth**, and Jose C. Principe. Mixture kernel least mean square. In *The 2013 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7, Dallas, TX, USA, August 2013. IEEE.

Il Park, **Sohan Seth**, and Steven Vaerenbergh. Probabilistic kernel least mean squares algorithms. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014.

Bilal Fadlallah, Austin Brockmeier, **Sohan Seth**, Andreas Keil, and José C. Príncipe. An association framework to analyze dependence structure in time series. In *International Conference on Engineering in Medicine and Biology Society (EMBC)*, 2012.

Bilal Fadlallah, **Sohan Seth**, Andreas Keil, and José C. Príncipe. Robust EEG preprocessing for dependence-based condition discrimination. In *International Conference on Engineering in Medicine and Biology Society (EMBC)*, 2011.

Sohan Seth, Austin Brockmeier, John Choi, Mulugeta Semework, Joseph Francis, and José C. Príncipe. Estimating dependence in spike train metric spaces. *IEEE International Joint Conference on Neural Networks (IJCNN)*, 2011.

Verónica Bolón-Canedo, **Sohan Seth**, Noelia Sánchez-Maroño, Amparo Alonso-Betanzos, and José C. Príncipe. Statistical dependence measure for feature selection in microarray datasets. In *Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*, 2011.

Il Park, **Sohan Seth**, Murali Rao, and José C. Príncipe. Estimating symmetric chi-square divergence for point processes. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011.

Sohan Seth, Austin Brockmeier, and José C. Príncipe. A metric based approach toward point process divergence. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011.

Sohan Seth, Il Park, Austin Brockmeier, Mulugeta Semework, John Choi, Joseph Francis, and José C. Príncipe. A novel family of non-parametric cumulative based divergences for point processes. In *Advances in Neural Information Processing Systems 23 (NIPS)*. 2010.

Sohan Seth and José C. Príncipe. Variable selection: A statistical dependence perspective. In *Proceedings of the International Conference on Machine Learning and Applications (ICMLA)*, 2010.

Lin Li, Il Park, **Sohan Seth**, Justin Sanchez, and José C. Príncipe. Neuronal functional connectivity dynamics in cortext: An MSC-based approach. In *Proceedings of the IEEE International Conference on Engineering in Medicine and Biology Society (EMBC)*, 2010.

Sohan Seth and José C. Príncipe. A test of granger non-causality based on nonparametric conditional independence. In *Proceedings of the International Conference on Pattern Recognition (ICPR)*, 2010.

Sohan Seth and José C. Príncipe. A conditional distribution function based approach to design nonparametric tests of independence and conditional independence. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2010.

Jung Phil Kwon, **Sohan Seth**, Andreas Keil, and José C. Príncipe. Estimation of instantaneous power in the EEG to assess brain connectivity with high temporal resolution. In *Proceedings*

of the IEEE International Conference on Engineering in Medicine and Biology Society (EMBC), 2009.

Lin Li, Il Park, **Sohan Seth**, Justin Sanchez, and José C. Príncipe. Estimation and visualization of neuronal functional connectivity in motor task. In *Proceedings of the IEEE International Conference on Engineering in Medicine and Biology Society (EMBC)*, 2009.

Sohan Seth and José C. Príncipe. Estimation of density ratio and its application to design a measure of dependence. In *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2009.

Sohan Seth and José C. Príncipe. On speeding up computation in information theoretic learning. In *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN)*, 2012.

Sohan Seth, Il Park, and José C. Príncipe. A new nonparametric measure of conditional independence. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2009.

Sohan Seth and José C. Príncipe. Compressed signal reconstruction using the correntropy induced metric. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2008.

Sohan Seth, Mustafa C. Ozturk, and José C. Príncipe. Signal processing with echo state networks in the complex domain. In *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2007.

Workshop papers

Isaac Neal, **Sohan Seth**, Gary Watmough, and Mamadou Saliou Diallo. Towards Sustainable Census-Independent Population Estimation in Mozambique. In *Artificial Intelligence for Public Health (AI4PH)*, 2021.

Lin Li, Il Park, **Sohan Seth**, John Choi, Joseph Francis, Justin Sanchez, and José C. Príncipe. An adaptive decoder from spike trains to micro-stimulation using kernel least-mean-square KLMS algorithm. *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2011.

Sohan Seth and José C. Príncipe. A conditional independence perspective of variable selection. In *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2010.

Sohan Seth and José C. Príncipe. Estimation of density ratio and its application to design a measure of dependence. In *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2009.

Sohan Seth, Mustafa C. Ozturk, and José C. Príncipe. Signal processing with echo state networks in the complex domain. In *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2007.

Extended Abstract

Alexander Przybylski, Edwin van Beek, Kevin Dhaliwal, Aziz Sheikh, Azin Salimian, **Sohan Seth**, Giorgos Papanastasiou, Jeremy Walker, and Nik Hirani. Stratification of Fibrotic Lung Disease: Integration of Molecular Endotyping and Quantitative CT via Machine Learning. In *Futuristic Medicine symposium, Royal College of Physicians Edinburgh*, 2019.

Sohan Seth, Niko Valimaki, Antti Honkela, and Samuel Kaski. Differential analysis of whole-genome shotgun sequences. In *Machine Learning in Systems Biology*, 2014.