# Information Geometry, Privacy and Monte Carlo

Title & Abstract

July 4–7, 2026

## Programme

### Day 1 — Saturday, July 4

*Afternoon only*

| | |
|---|---|
| 12:30–13:00 | Registration |
| 13:00–13:15 | Opening remarks |
| 13:15–14:25 | Oral 1–2 |
| | Chair: Alexandros Beskos |
| | Oral 1: Ajay Jasra |
| | Oral 2: Christophe Andrieu |
| 14:25–14:40 | Coffee break |
| 14:40–16:25 | Oral 3–5 |
| | Chair: Alexandros Beskos |
| | Oral 3: Maria De Iorio |
| | Oral 4: Takuya Koriyama |
| | Oral 5: Alexandros Beskos |
| 16:45–18:30 | Poster Session (1st) |

### Day 2 — Sunday, July 5

*Full day*

| | |
|---|---|
| 09:00–10:10 | Oral 6–7 |
| | Chair: Shuhei Mano |
| | Oral 6: Federica Milinanni |
| | Oral 7: Kenji Fukumizu |
| 10:10–10:25 | Coffee break |
| 10:25–12:10 | Oral 8–10 |
| | Chair: Christian P. Robert |
| | Oral 8: Hongsheng Dai |
| | Oral 9: Stefano Favaro |
| | Oral 10: Joshua Bon |
| 12:10–13:30 | Lunch |
| 13:30–13:50 | Group photo |
| 13:50–15:20 | Poster Session (2nd) |

### Day 3 — Monday, July 6

*Full day*

| | |
|---|---|
| 09:00–10:10 | Oral 11–12 |
| | Chair: Dootika Vats |
| | Oral 11: Radu V. Craiu |

| | |
|---|---|
| | Oral 12: Galin Jones |
| 10:10–10:25 | Coffee break |
| 10:25–12:10 | Oral 13–15 |
| | Chair: Dootika Vats |
| | Oral 13: James Flegal |
| | Oral 14: Michiko Okudo |
| | Oral 15: Marta Catalano |
| 12:10–13:30 | Lunch |
| 13:30–14:40 | Oral 16–17 |
| | Chair: Keisuke Yano |
| | Oral 16: Takemasa Miyoshi |
| | Oral 17: Manon Michel |
| 14:40–14:55 | Coffee break |
| 14:55–16:40 | Oral 18–20 |
| | Chair: Michael Choi |
| | Oral 18: Geoffrey Wolfer |
| | Oral 19: Daniel Paulin |
| | Oral 20: Shahab Asoodeh |

## Day 4 — Tuesday, July 7

*Morning only*

| | |
|---|---|
| 09:00–10:10 | Oral 21–22 |
| | Chair: Michael Choi |
| | Oral 21: Weiming Feng |
| | Oral 22: Michael Choi |
| 10:10–10:25 | Coffee break |
| 10:25–12:10 | Oral 23–25 |
| | Chair: Keisuke Yano |
| | Oral 23: Yuga Iguchi |
| | Oral 24: Edric Tam |
| | Oral 25: Victor Elvira |
| 12:10–12:30 | Closing and general discussion |
| 12:30 | Workshop ends |

## Titles and Abstracts

### Oral 1: Particle Filtering for a Class of State-Space Models with Low and Degenerate Observational Noise

**Speaker:** Ajay Jasra
**Affiliation:** CUHK-Shenzhen

**Abstract.** We consider the discrete-time filtering problem in scenarios where the observation noise is degenerate or low. We focus on the case where the observation equation is a linear function of the state and that additive noise is low or degenerate, however, we place minimal assumptions on the hidden state process. In this scenario we derive new particle filtering (PF) algorithms and, under assumptions, in such a way that as the noise becomes more degenerate a PF which approximates the low noise filtering problem provably inherits the properties of the PF used in the degenerate case. We extend our framework to the case where the hidden states are drawn from a diffusion process. In this scenario we develop new PFs which are robust to both low noise and fine levels of time discretization. We illustrate our algorithms numerically on several examples.

### Oral 2: The Oracle Metropolis-Hastings

**Speaker:** Christophe Andrieu
**Affiliation:** University of Bristol, UK

**Abstract.** We present novel and less novel MCMC updates.

### Oral 3: Title TBA

**Speaker:** Maria De Iorio
**Affiliation:** TBA

**Abstract.** Abstract TBA.

### Oral 4: Alignment from Pairwise Preferences: A Markov Chain Approach

**Speaker:** Takuya Koriyama
**Affiliation:** University of Chicago

**Abstract.** TBA.

### Oral 5: Particle Based Inference for Continuous-Discrete State Space Models

**Speaker:** Alexandros Beskos
**Affiliation:** University College London

**Abstract.** We develop a methodology allowing application of the complete machinery of particle-based inference upon the class of continuous-discrete State Space Models (CD-SSMs). This class corresponds to a latent continuous-time diffusion observed with noise at discrete-time instances. Due to the continuous-time nature of the signal, standard Feynman-Kac formulations and their

particle-based approximations have to overcome several challenges, including: (i) finite-time transition densities of the signal are typically intractable; (ii) ancestors of sampled signals are determined w.p. 1, thus cannot be resampled; (iii) diffusivity parameters given a sampled signal yield Dirac distributions. We overcome such issues by building a framework of carefully designed path proposals and reparameterisations thereof, and obtain new expressions for the Feynman-Kac model that accommodate the effects of continuous-time signals. Our formulations enable use of the full range of particle-based algorithms for CD-SSMs: for filtering/smoothing and parameter inference, online or offline. Our framework is compatible with guided proposals in the filtering steps which are essential for efficient performance under informative observations or in higher dimensions, and is applicable for a general class of CD-SSMs, including the case when the signal is a hypo-elliptic diffusion.

### Oral 6: Rapid mixing of stereographic MCMC for heavy-tailed sampling

**Speaker:** Federica Milinanni
**Affiliation:** Northwestern University

**Abstract.** Sampling from high-dimensional, heavy-tailed distributions is a fundamental challenge in computational statistics, as many standard Markov chain Monte Carlo (MCMC) methods mix poorly in such settings. Recently, Stereographic MCMC [Yang et al., 2024] was proposed as a promising family of samplers for such task. However, its non-asymptotic convergence properties remain unexplored. In this work, we fill this gap by establishing non-asymptotic upper bounds on the mixing time of the Stereographic Random Walk Metropolis. Our results demonstrate that, under certain conditions on the initial distribution, the mixing time is polynomial in dimension for a broad class of distributions, including light- and heavy-tailed cases. Our proof is based on a conductance argument and leverages weighted isoperimetric inequalities for heavy-tailed distributions. These proof techniques provide new insights into the geometric properties of heavy tailed distributions that govern rapid mixing in high dimensions.

### Oral 7: Title TBA

**Speaker:** Kenji Fukumizu
**Affiliation:** The Institute of Statistical Mathematics

**Abstract.** Abstract TBA.

### Oral 8: Online federated learning framework for classification

**Speaker:** Hongsheng Dai
**Affiliation:** Newcastle University

**Abstract.** This presentation introduces a new online federated learning framework for classification that is designed for streaming data from multiple clients while preserving data privacy and computational efficiency. The approach builds on the generalized distance-weighted discriminant method and a key contribution is an efficient optimization strategy based on the Majorization–Minimization principle, combined with a renewable estimation procedure. This enables the model to be updated incrementally as new data arrive, avoiding the need for costly full retraining. To further address privacy concerns, differential privacy mechanisms are incorporated to protect sensitive client information while maintaining predictive performance.

## Oral 9: Bayesian nonparametric privacy-preserving synthetic data generation

**Speaker:** Stefano Favaro
**Affiliation:** University of Torino and Collegio Carlo Alberto

**Abstract.** Synthetic data generation is a well-established approach for releasing data while protecting individual privacy. From a statistical perspective, a valid release mechanism should simultaneously guarantee differential privacy (DP) and preserve the statistical utility of the data. We propose a new release mechanism based on Bayesian nonparametric modeling. Specifically, we model the private dataset as a sample from a Pitman–Yor process (PYP) prior and generate synthetic observations from its posterior predictive distribution. We provide a detailed analysis of both the privacy and utility properties of the proposed mechanism. Concerning privacy, we focus on the recently introduced notion of instance-level DP. Our results reveal three qualitatively distinct regimes depending on the value of the discount parameter $\sigma$ of the PYP. If $\sigma < 0$, instance-level $(\varepsilon, \delta)$-DP can be achieved, with guarantees depending on the size of the released synthetic sample. If $\sigma = 0$, corresponding to the Dirichlet process, instance-level $(\varepsilon, \delta)$-DP holds provided that the original dataset contains no singletons and for suitable choices of the total mass parameter $\theta$. In contrast, if $\sigma > 0$, instance-level DP can only be obtained under extreme and practically restrictive conditions. To assess statistical utility, we concentrate on the case $\sigma = 0$ and derive convergence rates for the expected 1-Wasserstein distance between the empirical measure of the released synthetic data and the true data-generating distribution. The analysis relies on a novel concentration result for the posterior distribution of a Dirichlet process, which may be of independent interest.

## Oral 10: Persuasive Privacy

**Speaker:** Joshua Bon
**Affiliation:** Adelaide University

**Abstract.** We propose a novel framework for measuring privacy from a Bayesian game-theoretic perspective. This framework enables the creation of new, purpose-driven privacy definitions that are rigorously justified, while also allowing for the assessment of existing privacy guarantees through game theory. We show that pure and probabilistic differential privacy are special cases of our framework, and provide new interpretations of the post-processing inequality in this setting. Further, we demonstrate that privacy guarantees can be established for deterministic algorithms, which are overlooked by current privacy standards.

## Oral 11: MCMC for Directed Acyclic Graphs via Birth-Death Processes

**Speaker:** Radu V. Craiu
**Affiliation:** University of Toronto

**Abstract.** Inferring a directed acyclic graph (DAG) given data is computationally challenging. Current state-of-the-art MCMC methods for graph inference efficiently scan the space by first considering a restricted search space and iteratively expand the space until a stopping criterion is met. Here, we estimate the error introduced from current methods that use restricted spaces compared to the full space, and develop a novel MCMC method that reduces this error. Our method is an adaptive algorithm which allows for either expansion or contraction of the search space at any iteration. Both the expansion and contraction are determined by a birth-death process. Extensive simulations demonstrate the efficiency of the new algorithm, compare its

performance with existing methods, and consider applications in the field of imaging proteomics. This is joint work with Kieran Campbell and Morris Greenberg.

### Oral 12: Title TBA

**Speaker:** Galin Jones
**Affiliation:** University of Minnesota

**Abstract.** Abstract TBA.

### Oral 13: Simultaneous confidence bands for (Markov chain) Monte Carlo simulations

**Speaker:** James Flegal
**Affiliation:** University of California, Riverside

**Abstract.** We consider the construction of simultaneous confidence bands (CBs) for functional parameters over uncountable sets in Monte Carlo simulations. This setting arises naturally in estimating one or more marginal posterior densities, which may be done nonparametrically or via Rao-Blackwellisation. Although our motivation is primarily in estimating and visualizing densities, our approach can be applied in more general functional estimation, such as in estimation of likelihood functions for likelihood-based inference and when functional parameters are of interest in Bayesian inference. In short, we provide principled uncertainty quantification in the form of a simultaneous CB and hence increase the reliability of the resulting functional inference.

### Oral 14: Applications of information geometry to Bayesian prediction and estimation in curved exponential families

**Speaker:** Michiko Okudo
**Affiliation:** Chiba University

**Abstract.** I will review several of our recent results on Bayesian prediction and estimation from an information-geometric viewpoint. We propose a construction of predictive distributions for curved exponential families by extending estimators beyond the parameter space and considering a geometric projection of Bayesian predictive distributions. Using this construction, we show that for multivariate normal distributions with unknown mean vectors and known covariance matrices, we can construct a predictive distribution that improves on the Bayesian predictive distribution based on the Jeffreys prior by additionally introducing an estimator for covariance matrices.

### Oral 15: Distances on random measures for Bayesian nonparametrics

**Speaker:** Marta Catalano
**Affiliation:** Luiss University

**Abstract.** Random measures are a key component of many nonparametric models in Bayesian Statistics. Their infinite-dimensionality guarantees remarkable flexibility and generality but makes the investigation of theoretical and inferential properties more demanding. In this talk we underline how several of these properties can be investigated through suitable distances between the laws of the random measures. Some crucial desiderata for such distances are metrization of weak convergence, numerical estimation through samples, and tractability of analytical bounds: we achieve them by relying on optimal transport and integral probability metrics. Applications of our findings include the measurement of dependence in Bayesian nonparametric models, the definition of merging rates of opinions, two-sample tests for measure-valued data, and the quantification of the error in approximate posterior inference.

### Oral 16: Title TBA

**Speaker:** Takemasa Miyoshi
**Affiliation:** RIKEN

**Abstract.** Abstract TBA.

### Oral 17: Latent Monte Carlo

**Speaker:** Manon Michel
**Affiliation:** CNRS, Université Clermont-Auvergne

**Abstract.** TBA.

### Oral 18: Characterization of Exponential Families of Lumpable Stochastic Matrices

**Speaker:** Geoffrey Wolfer
**Affiliation:** Tokyo University of Agriculture and Technology

**Abstract.** A Markov chain is said to be lumpable when its state space can be partitioned into groups so that the aggregated process over those groups remains Markovian. In this talk, we examine the geometric structure of lumpable families of Markov chains within the information geometric framework introduced by Nagaoka. Although the set of lumpable Markov chains on a finite state space with respect to a fixed lumping map does not, in general, form an exponential family (e-family) of stochastic matrices, it can nevertheless be endowed with the structure of an e-foliated manifold. A natural question is which properties of the connection graph and the lumping map ensure that a lumpable family is exponential. We provide polynomially verifiable necessary and sufficient conditions, including combinatorial criteria on the connection graph and a more broadly applicable dimension-based method for determining whether a lumpable family forms an e-family. The talk is based on joint work with Shun Watanabe.

### Oral 19: Stochastic gradient Langevin dynamics: convergence and bias

**Speaker:** Daniel Paulin
**Affiliation:** Nanyang Technological University

**Abstract.** Using stochastic gradients in Langevin dynamics can dramatically reduce the computational cost of such algorithms. In this talk, we will present some new non-asymptotic results about convergence and bias for such stochastic gradient-based Langevin methods, improving upon earlier results in the literature.

## Oral 20: Recent Advances in Metropolis–Hastings Algorithms

**Speaker:** Shahab Asoodeh
**Affiliation:** McMaster University & Vector Institute

**Abstract.** I will present several new information-theoretic coupling methods and demonstrate their applications to obtaining sharper bounds on the meeting time of Metropolis–Hastings algorithms. Time permitting, I will also extend the analysis to the Multiple-Try Metropolis framework.

## Oral 21: Faster mixing of the Jerrum-Sinclair chain

**Speaker:** Weiming Feng
**Affiliation:** The University of Hong Kong

**Abstract.** We show that the Jerrum-Sinclair Markov chain on matchings mixes in time $O(\Delta^2 m \operatorname{polylog} n)$ on any graph with $n$ vertices, $m$ edges, and maximum degree $\Delta$, for any constant edge weight $\lambda > 0$. For general graphs with arbitrary, potentially unbounded $\Delta$, this provides the first improvement over the classic $O(n^2 m \operatorname{polylog} n)$ mixing time bound of Jerrum and Sinclair (1989) and Sinclair (1992). To achieve this, we develop a general framework for analyzing mixing times, combining ideas from the classic canonical path method with the local-to-global approaches recently developed in high-dimensional expanders, introducing key innovations to both techniques.

## Oral 22: Optimising two-block averaging kernels to speed up Markov chains

**Speaker:** Michael Choi
**Affiliation:** National University of Singapore

**Abstract.** We study the problem of selecting optimal two-block partitions to accelerate the mixing of finite Markov chains under group-averaging transformations. The main objectives considered are the Kullback-Leibler (KL) divergence and the Frobenius distance to stationarity. We establish explicit connections between these objectives and the induced projection chain. In the case of the KL divergence, this reduction yields explicit decay rates in terms of the log-Sobolev constant. For the Frobenius distance, we identify a Cheeger-type functional that characterises optimal cuts. This formulation recasts two-block selection as a structured combinatorial optimisation problem admitting difference-of-submodular decompositions. We further propose several algorithmic approximations, including majorisation–minimisation and coordinate descent schemes, as computationally feasible alternatives to exhaustive combinatorial search. Our numerical experiments reveal that optimal cuts under the two objectives can substantially reduce total variation distance to stationarity and demonstrate the practical effectiveness of the proposed approximation algorithms. This is based on joint work with Ryan Lim (NUS) and Youjia Wang (NUS).

## Oral 23: Dynamical Regimes of Denoising Diffusion Models: Geometric Perspectives

**Speaker:** Yuga Iguchi
**Affiliation:** School of Mathematical Sciences, Lancaster University

**Abstract.** TBA.

### Oral 24: Some Fundamental Limits to Neural Monte Carlo

**Speaker:** Edric Tam
**Affiliation:** Stanford University

**Abstract.** Deep generative models (DGMs) are increasingly used in modern Monte Carlo methods to learn transport maps that push a simple source distribution, such as a Gaussian, to a complex target distribution. Despite strong empirical performance, these approaches often struggle with heavy-tailed targets. In this work, we analyze this limitation through functional inequalities. Using Poincare inequalities, we derive quantitative lower bounds on the gradient and Jacobian norms of any transport map pushing a Gaussian source to a heavy-tailed target, showing that as the target's moments grow, the transport map must necessarily exhibit increasingly large Jacobian norms. We extend these results to general log-concave source distributions. Notably, these constraints arise from the intrinsic geometry of the transport problem and therefore hold independently of neural architectures or training procedures. Our results highlight fundamental limitations of neural transport approaches for Monte Carlo, particularly in the presence of challenging targets that exhibit heavy-tails, and suggest the need for more robust hybrid sampling strategies.

### Oral 25: Title TBA

**Speaker:** Victor Elvira
**Affiliation:** TBA

**Abstract.** Abstract TBA.