Poster Titles and Abstracts

Daniel Adams
Gradient flows, entropy regularized optimal transport, and large deviations for particle dynamics
The theory of Wasserstein gradient flows in the space of probability measures constitutes a unified and powerful framework in the study of dissipative partial differential equations (PDE). The recently developed entropic regularisation technique paves the way for fast and efficient numerical methods for solving these gradient flows via the celebrated 'JKO scheme'. Here I discuss variants of the JKO scheme, designed to solve PDE which, a priori, do not have a gradient flow structure, however they do fit into the Lyapunov theory. A micro-macro passage from large deviations to gradient flows was established in the work of Stefan Adams, Nicolas Dirr, Mark Peletier, and Johannes Zimmer. In this vein, we will also use microscopic dynamics to justify the required scaling between the entropic regularization and the JKO time-stepping - first imposed in the work 'Convergence of Entropic Schemes for Optimal Transport and Gradient Flows', but until now, lacked theoretical motivation.

Francesca Basini
Improving the linear noise approximation for applications in developmental biology

Xingyuan Chen
Simulation of MV-SDEs with super-linear growth in measure and space
We present a particle system-based scheme for the simulation of McKean–Vlasov Stochastic Differential Equations (MV-SDEs) with drifts of super-linear growth in the measure and spatial component, and non-constant Lipschitz diffusion coefficient. The super-linear growth in the measure component stems from convolution operations with super-linear growth functions allowing in particular application to the granular media equation with possibly multi-well confining potentials. From a methodological point of view, we avoid altogether functional inequality arguments (as we allow for non-linear diffusion maps).

Fabian Germ
Learning the conditional law: signatures and conditional GANs in filtering and prediction

Matei Hanu
Subsampling in Ensemble Kalman Inversion
Inverse problems seek to find parameters that generate a set of observed data in a mathematical model. Several methods have been proposed to solve inverse problems such are optimisation and Bayesian approaches. We consider the Ensemble Kalman Inversion (EKI) which has been recently introduced as an efficient, gradient-free method. Based on the continuous-time Ensemble Kalman Filter, it uses an ensemble of particles and a linearisation technique to essentially estimate the posterior distribution in an underlying Bayesian inverse problem. Unfortunately, the algorithm becomes inefficient or even computationally infeasible if the considered data set is too large. A similar problem appears in large-scale optimisation with gradient descent algorithms in, e.g., machine learning. Here, randomised algorithms like stochastic gradient descent (SGD) have become popular: those use only a random subset of the data in each iteration. These are so-called subsampling techniques.

In the past, subsampling has been employed in the EKI, but not analysed. Based on a recent analysis of a continuous-time representation of SGD, we propose, analyse, and apply a subsampling-technique within EKI. Indeed, we propose two different subsampling techniques: either every particle observes the same data subset (single subsampling) or every particle observes a different data subset (batch subsampling). We present convergence results of the method (and some variants) in the setting of linear inverse problems. Then we illustrate our results in PDE-based inverse problems and image reconstruction.

Thomas Hodgson

Non-mean field Vicsek-type models for collective behaviour

We consider Interacting Particle dynamics with Vicsek-type interactions, and their macroscopic PDE limit, in the non-mean-field regime; that is, we consider the case in which each particle/agent in the system interacts only with a prescribed subset of the particles in the system (for example, those within a certain distance). In this non-mean-field regime the influence between agents (i.e. the interaction term) can be scaled either by the total number of agents in the system (global scaling) or by the number of agents with which the particle is effectively interacting (local scaling). We compare the behaviour of the globally scaled and the locally scaled system in many respects; in particular we observe that, while both models exhibit multiple stationary states, such equilibria are unstable (for certain parameter regimes) for the globally scaled model, with the instability leading to travelling wave solutions, while they are always stable for the locally scaled one. This observation is based on a careful numerical study of the model, supported by further analysis.

Martin Kolodziejczyk

Invariant measures for McKean-Vlasov SPDEs

In this poster we consider a non-linear parabolic Fokker-Planck PDE (the non-linearity is intended in the sense of McKean) and we are going to characterize its stationary states (i.e., solutions to the stationary PDE). Once this is done, we perturb the Fokker-Planck PDE with a stochastic term. Namely, a space time-white noise and we are going to study the well-posedness of the corresponding SPDE. After that, once the well-posedness is established we study whether the uniqueness of the stationary state (i.e., invariant measure) for the SPDE is recovered.

Jonas Latz

Stochastic gradient descent in continuous time: discrete and continuous data

Optimisation problems with discrete and continuous data appear in statistical estimation, machine learning, functional data science, robust optimal control, and variational inference. The 'full' target function in such an optimisation problem is given by the integral over a family of parameterised target functions with respect to a discrete or continuous probability measure. Such problems can often be solved by stochastic optimisation methods: performing optimisation steps with respect to the parameterised target function with randomly switched parameter values. In this talk, we discuss a continuous-time variant of the stochastic gradient descent algorithm. This so-called stochastic gradient process couples a gradient flow minimising a parameterised target function and a continuous-time `index' process which determines the parameter. While stochastic gradient descent is not an interacting particles system, it is worth noting that many ideas can be carried over to other continuous-time data science algorithms/methods.

We first briefly introduce the stochastic gradient processes for finite, discrete data which uses pure jump index processes. Then, we move on to continuous data. Here, we allow for very general index processes: reflected diffusions, pure jump processes, as well as other Lévy processes on compact spaces. Thus, we study multiple sampling patterns for the continuous data space. We show that the stochastic gradient process can approximate the gradient flow minimising the full target function at any accuracy. Moreover, we give convexity assumptions under which the stochastic gradient process with constant learning rate is geometrically ergodic. In the same setting, we also obtain ergodicity and convergence to the minimiser of the full target function when the learning rate decreases over time sufficiently slowly.

We illustrate the applicability of the stochastic gradient process in a simple polynomial regression problem with noisy functional data, as well as in physics-informed neural networks approximating the solution to certain partial differential equations.

Maria Lefter

Subsampling particle systems for McKean-Vlasov processes

The value of a function of the solution to an SDE with non-linear drift in the sense of McKean is usually approximated using the empirical mean of the associated *N* particle system. Recently it has been proposed a more computationally efficient alternative for sampling: one can use another associated, randomly subsampled approximating system instead of the usual *N*--particle one. Although there is plenty of evidence of a performance improvement with this approach, quantitative results are missing in the literature. We obtain uniform in time bounds for the order of weak error in the Euler scheme approximation. These computations are based on decaying bounds in time for derivatives of the solution to the Backwards Kolmogorov Equation associated to an associated stochastic process governed by a time dependent dynamics. Extra detail is introduced by additional dependence of an equally subsampled external measure into the coefficient, which is of outstanding relevance for applications to 1--layer neuronal networks.

Andrés Miniguano-Trujillo

A nonlocal PDE-constrained optimisation model for containment of infectious diseases

Non-pharmaceutical interventions have proven crucial in the containment and prevention of Covid-19 outbreaks. In particular public health policy makers have to assess the effects of strategies such as social distancing and isolation to avoid exceeding social and economical costs. In this work, we study an optimal control approach for parameter selection applied to a dynamical density functional theory model. This is applied in particular to a spatially-dependent SIRD model where social distancing and isolation of infected persons are explicitly taken into account. Special attention is paid when the strength of these measures is considered as a function of time and their effect on the overall infected compartment. A first order optimality system is presented, and numerical simulations are presented using a spectral-Newton method. This work could potentially provide some mathematical insights into the management of disease outbreaks.

This is an ongoing joint work with John Pearson and Ben Goddard

Conor Osborne

Towards convergence analysis for deep Gaussian process regression

Gaussian processes have proved to be a powerful and flexible tool in the reconstruction of functions given a set of known data points, with applications in machine learning, optimisation and data assimilation. However, they can be limited when the functions being reconstructed are of a non-stationary or anisotropic nature. Deep Gaussian processes, constructed using a hierarchical process where the inputs to a Gaussian process are themselves Gaussian processes, aim to give a more flexible approach to function reconstruction.

We look at convergence rates of these deep Gaussian processes in terms of the number of known data points. We also show that deep Gaussian process regression achieves considerably better results than standard Gaussian process regression when reconstructing non-stationary and anisotropic functions.

Jonna Roden

PDE-constrained optimization for multiscale particle dynamics

Louis Sharrock

Online Parameter Estimation for the McKean Stochastic Differential Equation

We consider the problem of online parameter estimation for a stochastic McKean-Vlasov equation and the associated system of weakly interacting particles. We propose an online estimator for the parameters of the McKean SDE, which evolves according to a continuous-time stochastic gradient descent algorithm on the asymptotic log-likelihood of the interacting particle system. We prove that this estimator converges in L1 to the stationary points of the asymptotic log-likelihood of the McKean SDE in the joint limit as $N \to \infty$ and $t \to \infty$, under suitable conditions which guarantee ergodicity and uniform-in-time propagation of chaos. We then demonstrate, under the additional condition of global strong concavity, that our estimator converges in L2 to the unique maximiser of this asymptotic log-likelihood function, and establish an L2 convergence rate. We also obtain analogous results under the assumption that, rather than observing multiple trajectories of the interacting particle system, we instead observe multiple independent replicates of the McKean SDE itself or, less realistically, a single sample path of the McKean SDE and its law. Our theoretical results are demonstrated via several numerical examples, including a stochastic opinion dynamics model.

Iain Souttar

Asymptotic properties of averaging procedures on particle systems.

Multiscale methods are widely used in the applied sciences in order to approximate two-scale particle systems with simplified equations. It is important, when implementing these methods, that one is sure they will preserve the long-time behaviour of the original system, and I will discuss how one can guarantee such a property.

Tiffany Vlaar

Multirate training of neural networks

In this work we illustrate the presence of latent multiple time scales in deep learning applications. We introduce the use of multirate techniques for neural network training and show that different features present in the data can be learned by training the network on different time scales simultaneously. Further, we propose multirate training within neural networks and propose linear drift of the slow parameters during the fast parameter update. By choosing appropriate partitioning of the parameters into fast and slow parts we show that deep neural networks for transfer learning applications in vision and natural language processing can be trained in half the time, without losing much (if any) generalization performance. Our work unlocks the potential of using multirate techniques for neural network training and we discuss potential applications in the physical sciences.