Self-interacting processes, supersymmetry, and Bayesian statistics

organized by
Persi Diaconis, Margherita Disertori, Christophe Sabot, and Pierre Tarres

Workshop Summary

Overview. This workshop was devoted to self-interacting processes and their interaction with supersymmetry, Bayesian statistics and de Finetti-type theorems.

The focus was on the following areas:

(1.) self-interacting processes,
(2.) supersymmetry and Anderson localisation,
(3.) de Finetti type theorems and excheangeability.

The participants in the workshop had expertise in different areas - some primarily had background in stochastic processes and excheangeability, while others were more familiar with statistical mechanics and supersymmetry. Thus the first goal of the workshop was to the familiarize the participants with the basics of the above areas. The second goal was to identify promising future directions in these areas.

The workshop followed the AIM format. The mornings were dedicated to talks. These talks were always of an introductory nature, focused on providing an up-do-date overview of some particular subject. Monday afternoon was dedicated to two tutorials, on supersymmetry and excheangeability respectively. Tuesday afternoon consisted in a productive open problem session. In the afternoons for the rest of the week, the participants worked in groups on problems. The outcomes of these working groups are summarized below.

Manhattan lattice: classical and random. In the classical Manhattan lattice model in $\mathbb{Z}^2$, all horizontal and vertical lines are oriented in alternating directions, whereas in the randomly-oriented Manhattan one, every line is assigned a uniform random direction. In both cases the walker may either make a turn with probability p or continue straight on.

The classical Manhattan lattice is associated to a network model of quantum localisation problem belonging to the symmetry class C (see Beamond, Owczarek and Cardy, 2003). By placing a mirror at 45° at each of the vertices where the path turns through ±90°, the possible trajectories may be thought of as forming clusters of bonds on a square lattice at 45° to the original one and therefore, by percolation argument, the trajectories are a.s. finite if $p > 1/2$ (Chalker, 2003). Aside from that nice observation, no rigorous mathematical study was so far completed for that model.

On the other hand, on the randomly-oriented Manhattan lattice, a new approach was proposed by Ledger, Tóth and Valkó (2019) in the case where $p = 1/2$, based on the observation that in that case the configuration of randomly oriented lines seen from the position of the random walker has a simple invariant measure given by the product Bernoulli measure on all possible orientations. Using resolvent techniques they can show that the variance is between $t^{5/4}$ and $t^{4/3}$ at time t, in a time averaged version. The conjectured behavior is $t^{4/3}$.

We tried to understand the techniques from that new paper and deduce from them the transience of the process in the randomly-oriented case, which looks doable. We also made some interesting guesses on possible scaling limits for both processes.

New models for reinforcement from Bayesian statistics? From $H^{2|2}$ to Bayesian statistics? During the last decade it emerged that the edge reinforced random walk (ERRW) has several complementary faces: it is a self-interacting process, it can be interpreted in terms of Bayesian statistics as it gives an interesting conjugate prior for reversible Markov chains, and it is explicitly related to a supersymmetric hyperbolic nonlinear sigma model ($H^{2|2}$).

The question considered here emerged from the natural aim to find other models that would share this phenomenology. These models could emerge either from the algebraic side from other supersymmetric sigma models, in particular coming from the large class identified by Zirnbauer. Or it could emerge from the Bayesian statistics side: some other interesting models have already been exhibited by Baccalado.

This is a difficult and long term question and we did not really make decisive progress, but it gave the opportunity to experts in Bayesian statistics and mathematical physics to learn from each other and to share their views on the different aspects of this question.

A signaling problem. The so-called reinforcement learning model in signaling games was introduced by Skyrms and models how two interacting players learn to signal each other and thus create a common language.

The first rigorous analysis was done by Argiento, Pemantle, Skyrms and Volkov (2009) with 2 states and 2 signals, who showed that the system converges to perfect information in that case. Hu and Tarrès (2011) studied the general case with N states M signals, possibly different. They proved that the expected payoff increases in average (and thus converges a.s.), and that a limit bipartite graph emerges, such that no signal-state correspondence is associated to both a synonym and an informational bottleneck. Also any graph correspondence with the above property is a limit configuration with positive probability, which means that the system can converge to asymptotically imperfect information, even in the case where $M = N$.

The aim of the group was to understand the case where invention of new signals or states were allowed, corresponding to mutator balls in the reinforced urn model, whose draw would imply the addition of a new color, as in Hoppe urn. The question was to determine whether that model would "learn better", that is, whether that continuous addition of new signals would result in an asymptotically perfect information transfer. We did not solve the problem but we made some useful observations that reduced to information known for the version with no new signals.

$H^{2|2}$ model on a hierarchical lattice. A group of participants worked on the vertex-reinforced jump process (VRJP) and the hyperbolic supersymmetric nonlinear sigma model $H^{2|2}$ on the hierarchical lattice. This lattice is indeed more tractable and gives interesting ensign on several models of statistical physics. The group working on that question showed that the VRJP exhibits a perfect renormalization by scaling.

More precisely, define the hierarchical lattice $\mathbb{L} = \mathbb{L}(d, L)$ of dimension d which is invariant under scaling by L^{-1}. The vertex reinforced jump process X on $\mathbb{L}$ is such that

$$\mathbb{P}_\lambda\big(X_{t+\delta t} = j \,\big|\, \{X_s\}_{s\le t}\big) = \beta_{X_t,j}(1 + \lambda\tau_{t,j})\delta t + o(\delta t), \quad j \neq X_t, \tag{1}$$

where β_{ij} are conductances, $\frac{1}{\lambda}$ is the initial local time for all vertices and $\tau_{t,j}$ is the additional local time spent by X at vertex j up to time t. We choose β_{ij} such that the free ($\lambda = 0$) process has Green's function $|i-j|^{-(d-2)}$.

We proved that the scaled process $Z_t = L^{-1}X_{L^2 t}$ satisfies the same law as X and the susceptibility $\chi_{\nu,\lambda}$ with mass $\nu > 0$ satisfies $\chi_{\nu,\lambda} = \sigma\chi_{L^2\nu, L^{-(d-2)}\lambda}$ with $\sigma = L^2$. By iteration this implies $\chi_{\lambda,\nu} = \frac{1}{\nu}$ for $d \geq 3$.

Rate of convergence of edge ratios to final ratios. We investigated the rate of convergence for ERRW and for VRJP, that is, how long it takes for the random environment to become close to its eventual limit.

We carried out a few calculations, for Polya's urn with 2 and then n colors, and found out that after time t the distribution of colours is likely within $O(1/\sqrt{t})$ of its limiting distribution in both cases - not depending on n, which we thought was surprising. We spent some time discussing what sort of techniques might allow us to say things about more complicated examples, but didn't get any results there.

Deriving properties of ERRW from VRJP, in particular find cover, relaxation and hitting time. As it is known from the work of Ding and Peres, and subsequent works, there is a very precise relation between the cover time of a reversible Markov chain and the maxima of the Gaussian Free Field (GFF) on the corresponding conductance network.

The cover time of the ERRW on a finite conductance network should exhibit a rather different behavior compared to the Markov chain due to the possible strong localization. On the other hand, it is known that ERRW can be written as a mixture of reversible Markov chains. The corresponding mixing measure is directly related to the $H^{2|2}$ nonlinear sigma model. As a consequence, the question about the cover time of the ERRW can be translated in terms of the maxima of mixture of GFF. The discussion could only be started on this question, but hopefully it will be continued and solved after the workshop.