Pierre-Loïc Méliot
Assistant Professor (*maître de conférences*)
Université Paris-Sud – Faculté des Sciences d'Orsay
Laboratoire de mathématiques – Bâtiment 307
F-91405 Orsay – France
pierre-loic.meliot@math.u-psud.fr

Research statement

This document presents my research activities and the corresponding publications; it mostly focuses on my works since my nomination as an assistant professor at University Paris-Sud Orsay (September 2013). It also includes a detailed research program. Further details can be found on my webpage: `http://www.imo.universite-paris-saclay.fr/~meliot/`.

My researches consist in studying random objects stemming from various areas of mathematics (random matrices, random graphs, arithmetic functions of random integers, combinatorial objects, *etc.*) by using tools from harmonic analysis. Here, harmonic analysis can be understood in two ways: the classical Fourier analysis on the real line (Fourier transform of distributions of numerical random variables), and the representation theory of finite, compact or Lie groups (non-commutative Fourier transform of functions on a group). This document is split in two sections which explain these two aspects of my works. Each section details the corresponding publications and future research projects.

1. Asymptotic behavior of random variables and mod-ϕ convergence

1.1. Ratios of Fourier and Laplace transforms.

A first half of my research work consist in developing general techniques which enable the study of the asymptotic behavior of random variables. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real-valued random variables. If there exists a renormalisation $(X_n/s_n)_{n\in\mathbb{N}}$ which satisfies a central limit theorem, then most of the time one can make this result more precise by giving a sequence of parameters $(t_n)_{n\in\mathbb{N}}$ growing to $+\infty$ and a continuous function $\psi : \mathbb{C} \to \mathbb{C}$ such that

$$\mathbb{E}[\mathrm{e}^{zX_n}]\,\mathrm{e}^{-\frac{t_n z^2}{2}} \to_{n\to\infty} \psi(z) \quad \text{locally uniformly on the complex plane.}$$

In other words, X_n equals a Gaussian random variable with large variance t_n, plus a residue which is asymptotically encoded in the Laplace or Fourier sense by the function ψ. We then say that $(X_n)_{n\in\mathbb{N}}$ converges in the mod-Gaussian sense with parameters $(t_n)_{n\in\mathbb{N}}$ and limiting function ψ. The theory of mod-Gaussian convergence has been developed during the last decade, with the following double objective:

- To identify the probabilistic consequences of a mod-Gaussian convergence: precisions on the central limit theorem for $X_n/\sqrt{t_n}$, large deviation principles, speed of convergence estimates, local limit theorems, concentration inequalities.
- To find large classes of random variables which converge in the mod-Gaussian sense (in general this requires much more work than the proof of the central limit theorem).

Thus, the convergence of ratios of Laplace transforms provides a unified approach to all the possible kinds of asymptotic results for the distributions of a sequence of random variables (§1.2). Besides, one can give quite general sufficient conditions which imply this convergence of the ratios (§1.3). Let us notice that a generalisation of this method is possible with, instead of the normal law, a reference law ϕ which is an arbitrary infinitely divisible distribution: this is the theory of mod-ϕ convergence (in particular, with a Lévy–Khintchine exponent $\mathrm{e}^z - 1$ instead of z^2, one obtains the mod-Poisson convergence).

1.2. **Probabilistic consequences and main examples.** In (A6), it is shown that the mod-ϕ convergence implies sharp large or moderate deviation principles (by sharp we mean with an asymptotic equivalent of the probability of deviation, instead of its logarithm). In (A5) and (A4), these results are completed by estimates of the speed of convergence in the central limit theorem, and by local limit theorems for mod-Gaussian or mod-stable sequences of random variables. The case of discrete random variables which converge in the mod-Poisson case is detailed in (A3): in this setting, one obtains estimates of the total variation distance, and techniques in order to improve the speed of convergence of these approximation schemes. Finally, in a restricted case of mod-Gaussian convergence (the case of the method of cumulants, see §1.3), the previous results have been completed by concentration inequalities in (A2).

Let us list the main applications of these theoretical results (for each example, the new estimates which have been found with our methods are indicated by the corresponding article, according to the paragraph above):

- number of cycles or rises in a random permutation (A3)-(A6); more generally, statistics of random combinatorial objects whose generating series have logarithmic-algebraic singularities (A3);
- number of prime divisors of a random integer; more generally, arithmetic function whose L-function satisfies the Selberg–Delange hypotheses (A3)-(A6);
- characteristic polynomials of large random matrices chosen in compact Lie groups (A4)-(A6);
- magnetisation of the Ising model: (A7) for the dimension 1 and the related critical Curie–Weiss model, and (A5) for the dimension $d \geq 2$;
- number of points of a determinantal point process with locally square-integrable kernel (A4)-(A5)-(A6);
- motives in a random graph or permutation (A2)-(A5)-(A6);
- polynomial observable of a random integer partition chosen according to a central spectral measure, or of a model of random metric measure space (A1)-(A2)-(A5)-(A6).

1.3. **Techniques of analysis and the method of cumulants.** For each object previously listed, we had to perfect some already existing techniques that give the asymptotics of Fourier or Laplace transform; or, we developed new techniques. Let us give a (non-complete) list of these techniques, and their respective scopes.

- *Selberg–Delange method (A3)-(A6).* It enables to transfer the analytic properties of L-functions of arithmetic functions to Tauberian theorems satisfied by the Laplace transforms of these arithmetic functions evaluated on a random integer $n \in [\![1, N]\!]$.
- *Singularity analysis (A3)-(A6).* this is the analogue of the previous method when the random integer is replaced by a random combinatorial object, and the L-function is replaced by the generating series of the statistics under study.
- *Selberg integrals and Toeplitz determinants (A4)-(A6).* An exact formula for a matrix integral, possibly involving a Toeplitz or Fredholm determinant, can be analysed in order to understand the asymptotic behavior of some observables of random matrices (characteristic polynomial, number of eigenvalues in a domain of the complex plane).

Another important method in order to establish a mod-Gaussian convergence and its probabilistic consequences is the study of the cumulants

$$\kappa^{(r)}(X_n) = [z^r]\ \left(\log \mathbb{E}[\mathrm{e}^{zX_n}]\right)$$

of the sequence of random variables. These quantities have many interesting combinatorial properties, and in some cases they fit much better than the moments for the asymptotic study. The reason is that, given an asymptotically normal sequence $(Y_n)_{n\in\mathbb{N}}$, in many cases one has not only

$$\kappa^{(r)}(Y_n) \to 0 \qquad \forall r \geq 3$$

but in fact $\kappa^{(r)}(Y_n) = O((\varepsilon_n)^{r-2})$ with $\varepsilon_n \to 0$. In other words, the higher-order cumulants vanish faster than the first cumulants. This property leads to speed of convergence estimates, large deviation principles or concentration inequalities. In (A5)-(A6), some fundamental inequalities for the control of cumulants have been established when the random model admits a dependency graph or a weighted dependency graph. These techniques led in (A2) to the construction of large families of random models whose observables are jointly and generically mod-Gaussian convergent. These families of random models have been called mod-Gaussian moduli spaces, and an important question in this framework is the identification of the singular models in these families. In (A1), it has been proved that the approximation of a complete metric measure space by a random discrete model is singular (with non-Gaussian fluctuations which are much smaller than generically) if and only if the metric measure space approximated is a compact homogeneous space. There is a similar conjecture for the singular graphon models, see the next paragraph.

1.4. **Research projects.** We now detail some research projects connected to the theory of mod-Gaussian convergence; they are listed by decreasing advancement.

- *Singular graphon models.* The article (A2) proves in particular the mod-Gaussian convergence of the subgraph counts in a random graph associated to a graphon ω; the fluctuations of the densities of subgraphs are generically of order $n^{-1/2}$. In (A6), the same result has been established for Erdős–Rényi random graphs, but with fluctuations of order n^{-1}. We conjecture that the only graphons which give smaller fluctuations of order n^{-1} are the constant graphons corresponding to the Erdős–Rényi random graphs. Recently, I have shown that every singular graphon satisfy a central limit theorem analogous to the one for constant graphons; the proof relies on hidden equations satisfied by the cumulants of any singular model in a mod-Gaussian moduli space (see the preprint (A0)). This advance reduces the conjecture on singular graphs to a problem of random perturbation of the spectrum of a linear operator.
- *Large deviations of matrix models.* In (A6), we used the mod-Gaussian convergence of the logarithms $\log(\det(I_N - M_N))$ of the characteristic polynomials of random matrices chosen in classical compact groups in order to obtain moderate deviation estimates for their fluctuations of order $O(\log N)$. Very recently, we have discovered that the same method can be perfected to obtain the sharp large deviations, up to order $O(N)$. Besides, a comparison theorem enables one to extend the results to the case of the circular β ensembles, the case $\beta = 2$ being the only one known so far for the large deviations. As the random variables under study are here complex-valued, we also expect to observe some non-trivial multi-dimensional phenomena when computing the deviation probabilities.
- *Speed of convergence of the circular β ensembles.* In the same framework of the CβE ensembles, a very famous result due to Johansson ensures that if $\beta = 2$, then the traces of powers $\mathrm{Tr}((M_N)^k)$ converge for $k \geq 1$, $N \to \infty$ towards independent complex Gaussian variables, with a speed of convergence for the Kolmogorov distance which is super-exponentia and of order $O(\exp(-cN \log N))$. If $\beta \neq 2$, it is conjectured that the speed of convergence for the analogue central limit theorem (recently established by Jiang–Matsumoto) is much smaller. Some exact computations on the cumulants of these random variables, by using either the Jack polynomials or the underlying penta-diagonal matrix models, lead one to believe that the order of magnitude ot the speed of convergence for $\beta \neq 2$ is $O(N^{-1})$. The adaptation of the method of cumulants in this setting is an important problem: it would enable a better understanding of the central limit theorems without renormalisation, which occur frequently in random matrix theory.
- *Method of cumulants for the mixing dynamical systems.* In (A5), it has been shown thanks to the method of cumulants that the linear functionals of Markov chains satisfy central

limit theorems with Berry–Esseen type estimates. The proof relies on a combinatorial argument and the use of weighted dependency graphs; and on the Perron–Frobenius theorem for the location of the largest eigenvalues of the transition matrix of a Markov chain. One can adapt the proof to the case of a quasi-compact Markov operator in infinite dimension (in particular, one can study with these tools the Brownian motions on compact Lie groups). More generally, the argument should generalise to mixing dynamical systems whose linear functionals have Fourier transforms with a Nagaev–Guivarc'h representation involving a quasi-compact operator. An important example is provided by products of random matrices and by the Furstenberg–Kesten central limit theorem. In this setting, the method of cumulants would allow one to prove results without using the perturbation theory of operators (this is often the delicate point of the proofs). We should also obtain new concentration inequalities for these dynamical systems.

(A0) *A central limit theorem for singular graphons*. Submitted, 2021. https://www.math.u-psud.fr/~meliot/files/singular_final.pdf

(A1) *Fluctuations of the Gromov–Hausdorff sample model*, with Jacques De Catelan. To appear in Electronic Journal of Probability, 2021. https://www.math.u-psud.fr/~meliot/files/samplemodel.pdf.

(A2) *Graphons, permutons and the Thoma simplex: three mod-Gaussian moduli spaces*, with Valentin Féray and Ashkan Nikeghbali. Proc. London Math. Soc., 121(4):876-926, 2020.

(A3) *Mod-ϕ convergence: Approximation of discrete measures and harmonic analysis on the torus*, with Reda Chhaibi, Freddy Delbaen and Ashkan Nikeghbali. Ann. Inst. Fourier, 70(3):1115-1197, 2020.

(A4) *Local limit theorems and mod-ϕ convergence*, with Martina dal Borgo and Ashkan Nikeghbali. Latin American Journal of Probability and Mathematical Statistics, 16(1):817-853, 2019.

(A5) *Mod-ϕ convergence, II: Estimates on the speed of convergence*, with Valentin Féray and Ashkan Nikeghbali. Séminaire de Probabilités L, 405-478, LNM 2252, Springer-Verlag, 2019.

(A6) *Mod-ϕ convergence: Normality Zones and Precise Deviations*, with Valentin Féray and Ashkan Nikeghbali. Springer Briefs in Probability and Mathematical Statistics, 152+xii p., Springer-Verlag, 2016.

(A7) *Mod-Gaussian convergence and its applications for models of statistical mechanics*, with Ashkan Nikeghbali. In Memoriam Marc Yor – Séminaire de Probabilités XLVII, 369-425, LNM 2137, Springer-Verlag, 2015.

---

2. Random objects chosen in a group, an homogeneous space or a dual of a group

2.1. Probability and duality in groups. The second half of my research work consists in using the linear representations of groups in order to study random objects connected to these algebraic structures, in particular when the size of the group or of the object goes to infinity. The general method is as follows. Let G be a (finite, or compact, or reductive Lie) group, and denote G^* the set of its irreducible representations $\lambda = (V^\lambda, \rho^\lambda : G \to \mathrm{GL}(V^\lambda))$.

- Suppose given a G-valued random variable X, or a random object drawn on G and constructed from such random variables: for instance, a random walk $(X_t)_{t\geq 0}$ drawn on G, or a graph which connects random vertices $X_1, \ldots, X_n$ chosen in G. Then, the random matrices $\rho^\lambda(X)$ and the random character values $\mathrm{ch}^\lambda(X) = \mathrm{tr}\, \rho^\lambda(X)$ with λ running over G^* provide a family of matrix- or numerical observables, and these observables enable computations on the distribution of the variable X or of the random object constructed from X. The method can be adapted to the case where X takes its values in a quotient G/H (homogeneous space): we then need to replace G^* by the set of spherical representations of the pair (G, H), and the characters ch^λ by the zonal spherical functions zon^λ.

- In a dual way, let us consider a random variable Λ with values in G^*. The irreducible representations of the classical groups are labeled by sequences of integers, and numerous classes of random objects correspond to natural distributions on these sequences, and therefore on the irreducible representations of a group (we then speak of a spectral measure on the representations). In order to study Λ, we can as previously consider the random matrices $\rho^\Lambda(g)$ or the random character values $\mathrm{ch}^\Lambda(g)$, but this time with g running over G. These observables determine the law of the variable Λ, and they enable computations on this distribution.

The second part is classically involved in the solution of Ulam's problem of the longest increasing subsequence in a random permutation. Certain central limit theorems related to these problems, to the so-called Plancherel measures and to other spectral measures on integer partitions were the subject of my Ph.D. thesis, see (B4)-(B5)-(B6)-(B7)-(B8). Let us mention two recent works on this topic. The monography (B2) explains the combinatorics of the representations of the symmetric groups, and their use in the study of random partition models; it contains notably some new central limit theorems for the so-called central measures on integer partitions (see also (B5)). The proof of these algebraic central limit theorems uses the cumulant method explained in the first section of this document, and in fact, these exotic examples were one of the starting point of the theory of mod-Gaussian convergence. In particular, the Thoma simplex which labels the extremal characters of the group $\mathfrak{S}(\infty)$ and the central measures on integer partitions is a mod-Gaussian moduli space (A2), with a subset of singular models in bijection with $\mathbb{Z}$ and with the so-called Schur–Weyl measures. Some extensions of these results are one of our projects detailed in §2.4.

2.2. **Speed of convergence of random processes drawn on graphs.** A classical application of the first half of the general method described above is the estimation of the speed of convergence of the law of a random walk $(X_t)_{t\geq 0}$ on a finite or compact G; this has been made popular by the works of Diaconis in the 1980's. Indeed, the Parseval formula for compact groups yields an upper bound on the total variation distance between the law μ_t of X_t, and the Haar measure μ_∞ of the group. This upper bound is a series whose terms are labeled by the irreducibles representations of G (or the spherical representations of the pair (G, H) in the case of a G-invariant random walk on a quotient G/H). Often, the main term of this series also provides a lower bound in short time. In particular, a cut-off phenomenon frequently occurs, with a total variation distance which stays close to 1 for a long time, and then decreases abruptly to 0. The article (B3) proves this cut-off phenomenon for every Brownian motion on a compact symmetric space, with a cut-off time $t = c \log N$ if the space has rank N; this solves a conjecture of Saloff–Coste.

2.3. **Large random geometric graphs.** More recently, similar techniques have been used in order to study the spectrum of large random geometric graphs drawn on compact symmetric spaces. Let S be such a space (for instance, $S = \mathrm{SU}(3)$ or $S = \mathbb{S}^3$), and L be a fixed level. The random geometric graph with N points and level L is the graph $\Gamma_{\mathrm{geom}}(N, L)$ whose vertices are N independent random points $X_1, \ldots, X_N$ chosen uniformly on S, and whose edges connect the vertices X_i and X_j if $d(X_i, X_j) \leq L$, d being the geodesic distance on S. When N goes to infinity, the spectrum $e_1(N, L) \geq \cdots \geq e_N(N, L)$ of the adjacency matrix of $\Gamma_{\mathrm{geom}}(N, L)$ has remarkable asymptotic properties, that are explained in (B1). To simplify, let us assume that $S = G$ is a compact simply connected Lie group.

1. If $N \to \infty$ and L is fixed, the largest eigenvalues converge almost surely after scaling towards the eigenvalues of an integral Hilbert–Schmidt operator. These limiting eigenvalues can be made entirely explicit by using Bessel functions on the space of weights of the Lie algebra of the group G.
2. If $N \to \infty$ and $L = L_N = O(N^{-1/\dim G})$, the random graph $\Gamma_{\mathrm{geom}}(N, L_N)$ converges in the local sense of Benjamini–Schramm, and this implies that the spectral measure $\mu_N = N^{-1} \sum_{i=1}^{N} \delta_{e_i(N,L_N)}$ converges in probability towards a deterministic measure μ, which is determined by its moments.

The results for the second regime (Poissonian regime) are related to a general conjecture on certain functionals of the irreducible representations of the compact group G. Besides, the limiting measure μ does not really depend on the group G: it only depends on its dimension and on the parameter $\ell = \lim_{N\to\infty} L_N\, N^{1/\dim G}$. This limiting measure is the spectral measure of an infinite Poisson geometric graph in $\mathbb{R}^{\dim G}$; it seems to be connected to important problems in continuous percolation.

2.4. **Research projects.** As in the first section, we now explain four research projects connected to the techniques of harmonic analysis explained above; again, they are listed by decreasing order of advancement.

- *Spectral measure of a Poisson geometric graph.* In (B1), we constructed the spectral measures $\mu(d, \ell)$ of the large Poisson geometric graphs, but we did not give much information on these probability distributions, beyond the fact that they only depend on the dimension d of the group and on the scaled connection level ℓ. A in-depth study of these laws $\mu(d, \ell)$ show that: they are never compactly supported; they always have an atomic part; they seem to also have an absolutely continuous part with respect to the Lebesgue measure when $\ell > \ell_c$, ℓ_c being some critical parameter. This phase transition is related to the phase transition of the continuous percolation on $\mathbb{R}^d$: an absolutely continuous part in $\mu(d, \ell)$ appears when the parameter ℓ exceeds the critical threshold of the Poisson–Boolean percolation model in $\mathbb{R}^d$. We conjecture the following:
  (1) For every parameter ℓ, the measure $\mu(d, \ell)$ has exponential decay.
  (2) The spectral measures $\mu(d, \ell)$ vary continuously with the parameter ℓ (in particular, in the neighborhood of ℓ_c).

  The second item is very important, because it would be a result of continuity in percolation theory. For $d \geq 3$, it is unknown whether the percolation probability $p_\infty(d, \ell)$ is continuous at the critical parameter ℓ_c. Thus, the study of the limiting measures of the random geometric graphs on compact spaces opens the way to spectral methods in the theory of continuous percolation. The fact that such methods have not yet been developed is related to the fact that we are considering the spectra of infinite random graphs; this theory is very recent.
- *Stability of the cut-off phenomenon with respect to the generator of a Brownian motion.* In (B3), the cut-off phenomenon for the convergence to stationarity is proved for standard G-invariant Brownian motions on compact Lie groups G and compact symmetric spaces G/K. However, the convergence towards the Haar measure holds for much more general random processes: for instance, for every G-invariant Lévy process on G whose continuous part admits an infinitesimal generator which is hypoelliptic (a partial Laplace–Beltrami operator $\sum_{i=1}^{e \leq \dim G} \partial_i \otimes \partial_i$ with G-invariant vector fields $(\partial_i)_{1\leq i\leq e}$ which span the Lie algebra of G). One can then wonder under which conditions the convergence towards stationarity still exhibit a cut-off, and at which time. For instance, if $G = \mathrm{SU}(N)$, $d = N^2 - 1$ and the standard Laplace–Beltrami operator $\sum_{i=1}^{d} X_i \otimes X_i$ is replaced by a hypoelliptic operator $\sum_{i=1}^{e<d} X_i \otimes X_i$, does the corresponding diffusion still exhibit a cut-off? The representation theory of G should provide us with an upper

bound on the distance to stationarity, but the manipulation of this series seems much more complicated than in the elliptic case.

- *Speed of convergence for non-Euclidean central limit theorems.* The papers (A3)-(A5) of the previous section give general conditions on the Fourier transforms of real-valued random variables in order to measure the distance of their laws to a reference distribution. Similarly, in a framework of non-commutative harmonic analysis, the Diaconis–Shashahani upper bound (used in (B3)) controls the total variation distance to the Haar measure of a compact homogeneous space. More generally, one can try to use tools from harmonic analysis in order to measure distances between probability distributions in other geometric settings, for instance for random variables with values in a non-compact symmetric space. A starting point for this general question is the central limit theorem for convolutions of measures on a non-compact space G/K, for instance the Poincaré half-plane $\mathbb{H} = \mathrm{SL}(2)/\mathrm{SO}(2)$. The computation of the speed of convergence of this central limit theorem is an interesting problem, whose solution should involve the well-known harmonic analysis of $\mathrm{SL}(2)$. In this theory, there is a non-discrete part in the Parseval and Fourier inversion formula; this is one of the main differences with the compact case considered in the previous paragraph.
- *Mod-Gaussian moduli spaces stemmming from asymptotic representation theory.* As explained in the first paragraph of this section, the Thoma simplex is an important example of compact topological space which labels the extremal characters of an infinite group (the symmetric group $\mathfrak{S}(\infty)$), and such that the random finite approximations of these characters exhibit mod-Gaussian fluctuations. There are other examples of such classifying spaces in the representation theory of infinite groups; in particular, Borodin and Bufetov have proved a central limit theorem which is the analogue of the previous one, but for the extremal characters of the infinite unitary group $\mathrm{U}(\infty)$; this is related to the Gaussian free field in dimension 2. It would be interesting to understand these results by using the method of cumulants from §1.3; and to extend them to other classifying spaces, in particular the quantum analogue of the previous example corresponding to the characters of $\mathrm{U}_q(\infty)$, and which has in particular been studied by Gorin. Another important question which links the representation theory and the theory of mod-Gaussian convergence is the study of the fluctuations of a singular point of the Thoma simplex, namely, the one corresponding to the Plancherel measures. The fluctuations of the Plancherel model of random partitions are again asymptotically normal, but with covariances which are much smaller than for a generic point of the Thoma simplex, and no large deviation estimate known to this day. The mod-Gaussian convergence of these models is conjectured, in connection with combinatorial questions of enumeration of surfaces and of factorisations in symmetric groups. Let us notice that some recent works of Moll replace these geometric objects by combinatorial objects whose enumeration is simpler (ribbon paths); this might open the way to a solution of this conjecture.

(B1) *Asymptotic representation theory and the spectrum of a random geometric graph on a compact Lie group*. Electronic Journal of Probability, 24(43):1-85, 2019.

(B2) *Representation Theory of Symmetric Groups*. Discrete Mathematics and Applications, 666 +xvi p., CRC Press, 2017.

(B3) *The cut-off phenomenon for Brownian motions on compact symmetric spaces*, Potential Analysis, 40(4):427-509, 2014.

(B4) *Partial isomorphisms over finite fields*, Journal of Algebraic Combinatorics, 40(1):83-136, 2014.

(B5) *Fluctuations of central measures on partitions*, Proceedings of the 24th International Conference on Formal Power Series and Algebraic Combinatorics (Nagoya, Japan), p. 387-398, 2012.

(B6) *Asymptotics of q-Plancherel measures*, with Valentin Féray. Probability Theory and Related Fields, 152(3-4):589-624, 2012.

(B7) *Kerov's central limit theorem for Schur–Weyl and Gelfand measures,* Proceedings of the 23rd International Conference on Formal Power Series and Algebraic Combinatorics (Reykjavík, Iceland), p. 669-680, 2011.

(B8) *Products of Geck–Rouquier conjugacy classes and the algebra of composed permutations*, Proceedings of the 22nd International Conference on Formal Power Series and Algebraic Combinatorics (San Francisco, USA), p. 789-800, 2010.
