Research Statement

Subhajit Goswami
Institut des Hautes Études Scientifiques

During my graduate studies at the University of Chicago and subsequently my tenure as a postdoctoral fellow at the IHÉS, a unifying theme of my research has been to understand two fundamental phenomena, namely *percolation* and *diffusion*, in *correlated* environments. Studies of these phenomena find their origins in physics and all my research problems are motivated by questions arising in statistical mechanics.

On the percolation side, my contributions can be roughly classified into three subareas. In [31], we answered a fundamental open question about the phase transition of Ising model by studying a corresponding problem for the FK-Ising model — a member of a class of dependent percolation models known as the *random-cluster models* first introduced in [38] (see also [42]). These models are very important in statistical mechanics, especially because of their relation to other models. The second subarea of my research concerns percolation models driven by some strongly correlated process in the background. A canonical representative of such a model is the level-set of the Gaussian free field (see [46, 12, 64]) and in a recent work [33] we obtained a detailed characterization of the nature of its phase transition. Interestingly, both of these models turned out to be key ingredients in a different work [32] where we proved the existence of phase transition for *independent* (or *Bernoulli*) percolation on general graphs. Another type of models I have worked on is first-passage percolation defined using the *exponential* of some planar *log-correlated* Gaussian field. In the works [21, 22], we studied the typical order of first-passage percolation distance between two generic points in such models which, apart from being interesting on its own, is relevant for understanding the distance of Liouville quantum gravity (LQG) [57, 37, 63] — a major challenge in contemporary probability theory.

In addition to first-passage percolation, I have also worked on a diffusion process driven by the exponential of planar log-correlated field (see [9]). This was originally studied by physicists [13, 14] as a model for a particle (or a defect) diffusing in a disordered media with log-correlations. This process exhibits several interesting features, e.g. anomalous diffusive behavior, which are significantly different from that of a simple random walk. In [9] we were able to rigorously establish some of the predictions made in [13, 14] in this regard.

I am also interested in problems in probability theory that arise from applications in statistics and machine learning. Recently we revisited the problem of estimating piecewise constant functions from noisy data in [16, 15] where we analyzed the statistical risk associated with several methodologies including some new ones proposed by us.

In the following sections, I describe some of the aforementioned works in more detail and also discuss my long term research goals inspired by these projects. The titles of all my works can be found in my list of publications.

1 Truncated two-point function of the Ising model and the FK-Ising percolation

Originally conceived to explain the phenomenon of phase transition in ferromagnets, the Ising model is an archetypal example of a model undergoing an order-disorder phase transition. As one of the most well-studied models in mathematical physics, it has inspired a great amount of research activity among theoretical physicists and mathematicians alike. Apart from providing a deep understanding of the model itself and many other related systems, the study of Ising model has led to the development of various mathematical tools which has found applications in other areas of mathematics.

Formally an Ising model on $\mathbb{Z}^d$ (the regular d-dimensional lattice) is a probability measure $\langle\cdot\rangle$ on the space of *spin configurations* $\{\pm1\}^{\mathbb{Z}^d}$ characterized by the following property. Conditionally on the spins $\sigma_{\mathbb{Z}^d\setminus\Lambda} \coloneqq (\sigma_x : x \in \mathbb{Z}^d \setminus \Lambda) \in \{\pm1\}^{\mathbb{Z}^d\setminus\Lambda}$ outside a finite set $\Lambda \subset \mathbb{Z}^d$, $\langle\cdot\rangle$ penalizes the configuration $\sigma_\Lambda \coloneqq (\sigma_x : x \in \Lambda)$ by $\mathrm{e}^{-\beta H(\sigma_\Lambda)}$, where $\beta \in (0,\infty)$ is called the *inverse temperature* and $H(\sigma_\Lambda)$ is the Hamiltonian of the system given by $H(\sigma_\Lambda) \coloneqq -\sum_{x\in\Lambda,\, y\in\mathbb{Z}^d,\, x\sim y} \sigma_x\sigma_y$. Any such measure is called a *Gibbs measure* on $\mathbb{Z}^d$ for the (nearest-neighbor ferromagnetic) Ising model and the fundamental quantities to look at are the *magnetization* $\langle\sigma_x\rangle$ and the *two-point (correlation) function* $\langle\sigma_x\sigma_y\rangle$. Of particular interest is the Ising measure with + *boundary condition*, denoted as $\langle\cdot\rangle^+_\beta$, for which one can establish the existence of a *critical* inverse temperature $\beta_c = \beta_c(\mathbb{Z}^d) \in (0,\infty)$ such that the magnetization vanishes when $\beta < \beta_c$, and is strictly positive for $\beta > \beta_c$.

For a given Gibbs measure $\langle\cdot\rangle$, the *truncated two-point (correlation) function* is defined as:

$$\langle\sigma_0;\sigma_x\rangle \coloneqq \langle\sigma_0\sigma_x\rangle - \langle\sigma_0\rangle\langle\sigma_x\rangle\,.$$

It is well-known that sufficiently far from the critical phase, most systems of statistical physics exhibit exponential relaxation of truncated correlations [26], in both the equilibrium and the dynamical sense. It is more challenging to narrow the range of exceptions to a set of points, or lines, in the model's phase space. In [31], we *completed* that task for the d-dimensional nearest-neighbor ferromagnetic Ising model by showing:

Theorem 1.1 (Theorem 1.1, [31])**.** *For the nearest-neighbor Ising model on $\mathbb{Z}^d$ in dimension $d \geq 3$, for any $\beta > \beta_c$ there exists $c = c(\beta, d) > 0$ such that for every $x, y \in \mathbb{Z}^d$,*

$$0 \leq \langle\sigma_x;\sigma_y\rangle^+_\beta \leq \exp[-c\|x-y\|]\,.$$

Jointly with previously known results [45, 1, 2], this implies that for the nearest-neighbor Ising model in *any* dimension it is only at the critical point $\beta = \beta_c$ that the pure state's truncated two-point function fails to decay exponentially fast.

We proved Theorem 1.1 by first converting it into a corresponding statement for a *dependent* percolation model, called the random-cluster model, or the *Fortuin-Kasteleyn (FK) percolation*. A random-cluster measure is a probability measure on the space of (bond) *percolation configurations* $\{\text{open}, \text{closed}\}^{\text{Edges}(\mathbb{Z}^d)}$ satisfying the following property. Conditionally on any configuration ξ outside a finite *subgraph* Λ of $\mathbb{Z}^d$, the probability of $\omega \in \{\text{open}, \text{closed}\}^{\text{Edges}(\Lambda)}$ is proportional to $p^{\#\text{open}(\omega)}(1-p)^{\#\text{closed}(\omega)}q^{k_\xi(\omega)}$, where $p \in [0, 1]$ and $q \geq 1$ are parameters of the model and $k_\xi(\omega)$ is the number of clusters intersecting Λ of the percolation configuration $\overline{\omega}$ on $\mathbb{Z}^d$ that agrees with ω in $\text{Edges}(\Lambda)$ and ξ outside. ξ is also called a *boundary condition* on Λ.

In any percolation model one is primarily concerned with the connectivity properties of the underlying configuration, i.e., whether there exists a nearest-neighbor path between two given subsets S and T of $\mathbb{Z}^d$ consisting only of open edges (we henceforth denote this event by $S \leftrightarrow T$). In fact, the *analogous* quantities for percolation corresponding to the magnetization and two-point function are the probabilities of the events $x \leftrightarrow \infty$ and $x \leftrightarrow y$ respectively. For the random-cluster model $\phi_p[\cdot]$ with cluster-weight $q = 2$ (there exists a unique such measure, see [11, 60]), this analogy is even more direct in the form of *Edwards-Sokal coupling* which states that

$$\phi_p[x \leftrightarrow y] = \langle \sigma_x \sigma_y \rangle_\beta^+, \ \phi_p[x \leftrightarrow \infty] = \langle \sigma_x \rangle_\beta^+$$

for $p = 1 - \mathrm{e}^{-2\beta}$. Because of this, $\phi_p[\cdot]$ is also called the *FK-Ising model*. Like the critical inverse temperature β_c for Ising, there exists a constant $p_c = p_c(d) \in (0, 1)$ such that $\phi_p[x \leftrightarrow \infty]$ is equal to 0 for every $p < p_c$, and is strictly positive for every $p > p_c$. As a consequence of the Edwards-Sokal coupling these two critical values are related by $p_c = 1 - \mathrm{e}^{-2\beta_c}$. In particular this implies

$$\langle \sigma_x; \sigma_y \rangle_\beta^+ = \phi_p[x \leftrightarrow y] - \phi_p[x \leftrightarrow \infty]\phi_p[y \leftrightarrow \infty]$$

for some $p > p_c$ as soon as $\beta > \beta_c$. The exponential decay of the quantity on the right for Bernoulli percolation follows readily from the phenomenon of *percolation in slabs* for supercritical p (see [43]). While the percolation in slabs for the FK-Ising model is known (see Bodineau [10]), a major challenge to derive Theorem 1.1 from this stems from the *correlations* between events supported on distant sets, which could *a priori* be large (notice that such correlations vanish for Bernoulli percolation). In view of this we prove in [31], the following *exponential mixing property*:

Theorem 1.2 (Theorem 1.3, [31]). *For every $d \geq 3$ and $p > p_c$, there exists a constant $c > 0$ such that for every $n \geq 1$,*

$$|\phi_p[A \cap B] - \phi_p[A]\phi_p[B]| \leq \exp(-cn)\,,$$

where A and B are any two events depending on edges in $[-n, n]^d \cap \mathbb{Z}^d$ and outside $[-2n, 2n]^d \cap \mathbb{Z}^d$ respectively.

Future directions:

Using our techniques in [31], it is actually possible to improve the bound in Theorem 1.2 to $\exp(-cn)\phi_p[A]\phi_p[B]$. In literature this is known as the *ratio-weak mixing property*. This falls short of the *ratio-strong mixing property* related to the phenomenon of *boundary phase transition* for Ising models (see [51]). Although this stronger property is absent for Ising models in dimensions larger than 2 at low temperature, it is expected to hold in the entire subcritical phase.

Question 1. For $d \geq 3$ and $p < p_c$, prove that

$$|\phi_p[A \cap B] - \phi_p[A]\phi_p[B]| \leq \exp(-cd_{A,B})\phi_p[A]\phi_p[B]$$

where $d_{A,B}$ is the distance between the *supports* of the events A and B.

Another important improvement would be to understand the case of the *Potts models* with $q \geq 3$ colors. The Potts model is a generalization of the Ising model to more than two possible spins and has become an object of great interest in the last four decades (see [42, 30] and the references therein). Like Ising, the q-Potts model is also related to the random-cluster model with cluster weight q via the Edwards-Sokal coupling. In a recent paper by Duminil-Copin, Raoufi and Tassion [34], the exponential decay of the two-point function was proved in every dimension for the subcritical phase. However the study of the supercritical phase is still limited except in dimension 2 (see [36] and also [35]) which leads us to the following natural question.

Question 2. Prove that the truncated two-point function for the (nearest-neighbor) q-Potts model with monochromatic boundary conditions decays exponentially fast for all $q \geq 2$ in any dimension.

There seems to be, at this moment, one major hurdle to overcome before we can reduce Question 2 to a statement similar to Theorem 1.2. This obstacle arises from the fact that we do not have an analogue of Bodineau's result [10] for the random-cluster models with $q > 2$ (a partial result was obtained in [35]). Bodineau showed the equality between the so-called *slab percolation threshold* $\hat{p}_c$ — the minimum value of p above which there is an infinite cluster in sufficiently thick slabs for *any* boundary conditions — and the percolation threshold p_c for $q = 2$ and all $d \geq 3$. Therefore an important intermediate question to answer is

Question 3. Show that $\hat{p}_c = p_c$ for all integer values of $q \geq 3$ and $d \geq 3$.

An important ingredient of our proof of Theorem 1.2 is the *uniqueness* of infinite-volume measure for the FK-Ising model. While this is true for any random-cluster model for *all but countably many* values of p, the question of uniqueness at all (non-critical) p is still open except in dimension 2.

Question 4. Let $d \geq 3$ and $q \geq 3$ be an integer. Prove that there exists a unique random-cluster measure for all $p > p_c$.

Another important machinery used in the proof of Theorem 1.2 is the *random-current representation* of the Ising model (see [29] and the references therein) which itself is not available for the Potts model.

Question 5. Let $d \geq 3$ and $p > p_c$ be such that there exists a unique random-cluster measure with cluster weight $q \geq 3$ (an integer). Prove an analogue of Theorem 1.2 for this measure.

2 Percolation of level-sets of the Gaussian free field

A Gaussian free field (GFF) on a graph (or network) with boundary is a Gaussian process with covariances given by the Green's function of the associated simple random walk. It is one of the most natural Gaussian Markov fields defined on a graph and has deep connections with simple random walk, potential theory and harmonic analysis. In the context of the d-dimensional lattice $\mathbb{Z}^d$ it is also called the *Euclidean bosonic massless free field* in physics and as such plays an important role for many constructions in quantum field theory (see the exposition by Sheffield [68] and the references therein). From a mathematical perspective it is a rich source of interesting problems, partly stemming from its Markovian and Gaussian nature. Furthermore, on lattice graphs it is a canonical representative of strongly-correlated fields. It is therefore natural to study the percolative properties of the landscape of GFF on $\mathbb{Z}^d$ (in the transient regime, i.e., $d \geq 3$) as a model for percolation with long-range dependence.

The simplest way to define a percolation model using GFF (or for that matter, any random field) is to look at its level-sets, i.e., the set of vertices where its value lies above a certain threshold $h \in \mathbb{R}$. This gives a non-increasing one-parameter family of (site) percolation models indexed by $\mathbb{R}$. Consequently there is a critical point $h_* = h_*(d)$ such that the probability of the origin lying in an infinite cluster of the level-set above h is strictly positive for $h < h_*$ and 0 for $h > h_*$. It is a difficult question whether h_* is finite in which case the phase transition is called non-trivial. The non-triviality of phase transition was proved in [12] for $d = 3$ along with the non-negativity of h_* for all $d \geq 3$. It took almost three decades before a proof of non-triviality was found by Rodriguez and Sznitman [64] for all $d \geq 3$.

Like in the case of the FK-Ising percolation or any other models in statistical mechanics, we might expect the (appropriately) truncated two-point function of the GFF level-set percolation to decay (stretched) exponentially fast for h far away from h_*. In the subcritical phase this was proved for all dimensions in [64] (exponential decay for $d \geq 4$, with logarithmic corrections when $d = 3$, was obtained in [58]). An analogous result for the supercritical phase was proved by Sznitman in [71]. It is an important question whether this stretched-exponential decay of the truncated two-point function holds for all non-critical h. In [33] we answered this question in the affirmative.

Theorem 2.1 ([33])**.** *For all $d \geq 3$ and $h \neq h_*$, there exists $\rho = \rho(d, h) \in (0, 1]$ such that for every $x, y \in \mathbb{Z}^d$ sufficiently far,*

$$\mathbb{P}\left[x \text{ and } y \text{ lie in a } \textbf{finite} \text{ component of the level-set above } h\right] \leq \exp(-\|x - y\|^{\rho})\,.$$

In the common parlance of statistical mechanics, Theorem 2.1 implies that the phase transition of the level-set percolation of GFF is *sharp*. In fact, this theorem is one of the main corollaries of a more general result where we proved the identity between three — *a priori* different — notions of critical thresholds for GFF level-sets that were introduced in earlier works. One of these critical thresholds corresponds to a *strongly percolating* phase where several geometric and dynamical properties of the infinite cluster have already been obtained, e.g. the comparability of the chemical (intrinsic) distance on the infinite cluster with the Euclidean one, scaling limits of balls in the chemical distance [27], quenched invariance principle for the random walk [59] among others. In view of our new result, all the above results now are known to hold in the entire supercritical phase.

Our approach in [33] consists in obtaining a comparison result with a *finite-range* model in a *fictitious* non-perturbative regime, i.e., an interval outside which sharpness follows from (sophisticated) renormalization arguments. This finite-range model has a lot in common with Bernoulli (site) percolation and as such its sharpness — in the subcritical as well as the supercritical phase — follows from arguments similar to those for the Bernoulli percolation (see [43] and the recent paper [34]). However this contradicts our comparison result implying that the fictitious regime is indeed empty and the phase transition of the original model is sharp.

2.1 Relationship to Bernoulli percolation and other models

The general strategy of comparing the GFF level-sets to finite-range models can provide valuable insights about the Bernoulli percolation as well. In fact this is one of the main ideas used in [32], where we proved:

Theorem 2.2 (Theorem 1.1, [32])**.** *Consider any bounded degree graph G. Suppose that the probability of a simple random walk on G returning to its starting vertex at any given time n (the diagonal heat kernel) is uniformly bounded $c/n^{d/2}$ for some $d > 4$ and $c > 0$. Then the phase transition of Bernoulli percolation on G is non-trivial.*

As a corollary, we obtained that the phase transition of Bernoulli percolation on infinite quasi-transitive graphs (in particular, Cayley graphs) with super-linear growth is non-trivial, thus answering a conjecture of Benjamini and Schramm [7] dating back to 1996.

Whether a model undergoes a non-trivial phase transition or not is one of the most fundamental questions in statistical mechanics. In [55], Peierls introduced a combinatorial technique, known as *Peierls argument*, to prove that the critical temperature of the Ising model is non-zero on $\mathbb{Z}^d$ for $d \geq 2$. This argument found many applications to other models, including Potts models as well as Bernoulli percolation and the random-cluster model. Peierls argument has two drawbacks. First, it often does not apply to continuous spin models, for instance the spin $O(n)$ models. In this case, the technique may sometimes be replaced by two other techniques: Reflection Positivity and the Renormalization Group. More precisely, Frohlich, Simon and Spencer [39] proved that the spin $O(n)$ model undergoes a nontrivial order/disorder phase transition on $\mathbb{Z}^d$ with $d \geq 3$ [39] using *Reflection Positivity*, and Balaban and coauthors (see [6] and references therein) proved delicate properties of the low-temperature regime using the Renormalization Group. Another problem with Peierls argument is that it requires a precise understanding of so-called cut sets, i.e., sets of edges which disconnect certain sets of vertices from infinity. On planar graphs, this boils down to the understanding of circuits in the dual graph. On non-planar graphs, the question is a much more complex combinatorial problem and it is not completely understood in general. In this context we believe that the new technique introduced by us in [32] can be useful to prove the existence of a phase transition for various models. In particular, our results can be immediately extended to finite-range percolation and random-cluster models via classical comparisons. Furthermore, the Edwards-Sokal coupling between the random-cluster models and the Ising/Potts model mentioned previously implies that the results translate into results on the latter. Our hope is to be able to use this method to prove the existence of a phase transition for the spin $O(n)$ models that does not rely on reflection positivity.

Another important feature of our work is that we were able to connect two very fundamental but different processes that can be defined on a graph, namely the random walk — from which we get the heat kernel as well as the Green's function governing the law of the GFF — and percolation. The discovery of this link between a *dynamical* and an *equilibrium* property of a graph is very interesting on its own.

Future directions:

Ideally the ρ in Theorem 2.1 should be 1 but it is a bit more subtle in the case of GFF level-set percolation. For example, in the subcritical phase the probability that the GFF is above h along the line $\mathrm{L_n}$ connecting 0 and $(n, 0, 0, \dots)$ is bounded below by $\mathrm{e}^{-c\,\mathrm{Cap}(\mathrm{L_n})}$ where $\mathrm{Cap}(\mathrm{L_n})$ is the random walk capacity of $\mathrm{L_n}$. The latter is linear in all dimensions strictly larger than 3 whereas in dimension 3 it is of order $n/\log n$. Therefore in dimension 3 the best decay one can hope for has indeed a log-corrected linear exponent as reflected in the bounds obtained in [58]. The supercritical phase is even more unclear at this moment. Thus, a natural follow-up to Theorem 2.1 would be:

Question 1. Can we replace $\|x-y\|^\rho$ with $\|x-y\|/(\log\|x-y\|)^c$ in Theorem 2.1 for some $c > 0$? Is it possible to identify the precise *large deviation speed*?

A different direction is to see if the strategy developed in [33] can be applied to prove sharpness for models coming from fields with similar large-scale behavior as the GFF. One such instance is the model of random interlacements which was introduced by Sznitman in [69], motivated by the broad question about the disconnection of discrete cylinders and tori by the trace of simple random walk. The relevant family of random subsets in this case are the so-called *vacant sets* which are, roughly speaking, the sets of vertices *not* covered by a Poissonian ensemble of simple random walk trajectories. In the series of works by Lupu and Sznitman [70, 48, 72] a coupling was obtained between the vacant set of random interlacement and the level-set of GFF for a general class of graphs. However, the implication of this coupling is not clear as to the question of sharpness.

Question 2. Prove an analogue of Theorem 2.1 for the percolation of vacant sets of random interlacements.

Another interesting model with similar decay of correlation as the GFF level-sets where we expect our strategies to work is the family of stationary distributions of the so-called *voter model* (see [47], and also [61] for the percolation problem).

An exciting area to explore is the critical behavior of the level-sets of GFF. In the setting of *metric graphs* introduced by Lupu [48] the critical point was shown to be 0 by explicit computation of connection probabilities. In the same set-up, Ding and Wirth [24] recently obtained bounds on the critical one-arm probability with matching exponents for dimension 3. One can expect that the critical exponents of the two models should be same. The predictions in physics literature (see, e.g., section V in [74]) seem to suggest that the *correlation-length exponent* for long-range percolation models with the same rate of decay as the GFF level-set percolation should be $2/(d-2)$ for $d \leq 5$ and $1/2$ (the mean-field exponent) for higher dimensions.

Question 3. Show the existence of relevant critical exponents for GFF level-set percolation on $\mathbb{Z}^d$ and obtain sharp bounds on it.

As an intermediate step one may first work with the model on the metric graph which possesses additional structures and try to improve the bounds in [24]. It is quite possible that one can obtain precise estimates in this case.

Question 4. Study Question 3 for the GFF level-set percolation on the metric graph of $\mathbb{Z}^d$. Obtain precise bounds on the critical one-arm exponents.

3 First-passage percolation on the exponential of Gaussian free field

Consider the two-dimensional box V_N ($\subseteq \mathbb{Z}^2$) of side length N and a GFF η on V_N with zero boundary condition. For any fixed $\gamma > 0$ and $v, w \in V_N$, we define

$$D_{\gamma,N}(v,w) = \min_{\pi} \sum_{u \in \pi} \mathrm{e}^{\gamma \eta_u},$$

to be the *first-passage percolation* distance between v and w where each vertex u is assigned a weight $\mathrm{e}^{\gamma\eta_u}$ and π ranges over all paths in V_N connecting v and w. Our primary motivation behind this model comes from the random distance associated with Liouville quantum gravity (LQG) [57, 37, 63]. Informally, LQG is a random surface whose "Riemannian metric tensor" can be described as $\mathrm{e}^{\gamma X(x)} dx^2$, where X is a Gaussian free field on some planar domain D. Therefore the distance $D_{\gamma,N}(\cdot,\cdot)$ appears as a natural discrete approximation for the LQG distance. To emphasize this connection we refer to $D_{\gamma,N}$ as the *Liouville first-passage percolation (LFPP) distance.*

As in the classical case, a fundamental problem in the study of LFPP is to understand the behavior of the typical LFPP distance between two macroscopically separated vertices in the graph V_N. In [22] we were able to show that for small enough γ, the expected value of the LFPP distance between any two vertices $v, w \in V_N$ can be at most

$O(N^{1-c\gamma^{4/3}/\log\gamma^{-1}})$. This came as a surprise since it appears to be in disagreement with a non-rigorous prediction made by Watabiki [73] which, according to some "reasonable" interpretations (see [5, Equation (17), (18)]), implies that the LFPP distance between two generic vertices should behave like $N^{1-\Theta(\gamma^2)}$ for small γ and large N.

In [22] we obtained *similar exponents* for two other (related) notions of random distance which also contradict relevant interpretations of Watabiki's prediction when γ is small. One of this is an integrated version of $D_{\gamma,N}(\cdot,\cdot)$ defined with respect to a mollified version of the continuum GFF on a planar domain (this is probably a better approximation of the conjectured LQG distance than $D_{\gamma,N}$). The other notion comes from the *Liouville quantum gravity (LQG) measure* (see, e.g., [44, 37, 62, 63, 67]) which is, roughly speaking, the volume measure corresponding to the LQG metric. An intuitively natural way to recover the distance between two points u and v from a volume measure is to look at, for any given (small) δ, the minimum number of balls with volume at most δ whose union contains a path between u and v. When the underlying measure is the LQG measure, we call this the *Liouville graph distance* [22].

The main idea behind the proofs is a multi-scale analysis approach to construct an "economical path" connecting any two given points. A similar approach was used earlier in a different work [21] where we obtained a weaker bound on the exponent in a different (and simpler) set-up.

Future directions:

A natural question is whether the $\gamma^{4/3}$ term in all these distance exponents is optimal. In a recent work [25] by Ding, Zeitouni and Zhang, the existence of the distance exponent for the Liouville graph distance was established. In the same paper the authors also established the existence of an exponent involving the off-diagonal heat kernel of Liouville Brownian motion (LBM) — the natural diffusion process on the Liouville surface (see [40, 8]) — and related these two exponents. In a follow-up work [23], Ding and Gwynne extended this relationship to a larger class of regularized LQG distance exponents as well as relevant exponents from random planar map models in the LQG universality class. In view of these recent developments, it seems now even more important to obtain more precise estimates of any (and hence all) of these exponents.

Question 1. Find upper and lower bounds on the exponent of LFPP (or, equivalently, the Liouville graph distance) that match in their dependence on γ up to at most a log factor.

One can also ask if the bound for the LFPP distance is universal over some "suitable" class of log-correlated Gaussian fields.

Question 2. If φ is an arbitrary mean-zero Gaussian field satisfying $|\mathbb{E}\varphi_v\varphi_u - \log\frac{N}{1+\|u-v\|}| \leq K$, is it true that the corresponding expected LFPP distance is bounded above by $C_{K,\gamma}N^{1-c_K\gamma^{4/3}/\log\gamma^{-1}}$ for some $C_{K,\gamma}, c_K > 0$?

4 Random walk driven by the two-dimensional Gaussian free field

Let η be an instance of the GFF on $\mathbb{Z}^2$ pinned to 0 at the origin. Now for $\gamma > 0$ and conditional on the sample η, consider the random walk $\{X_t\}_{t\geq 0}$ on $\mathbb{Z}^2$ among random networks where the conductance of edge (v, w) is given by $\mathrm{e}^{\gamma(\eta(v)+\eta(w))}$. Random walks of this type were considered in the physics literature [13, 14] in a more general setup where the underlying Gaussian field is only required to satisfy $\mathrm{Var}(\eta(v) - \eta(w)) \asymp \log|v - w|$. The two-dimensional GFF is arguably the most canonical instance of such a field.

We studied this model in [9] where we showed ([9, Theorem 1.1]) that, for almost every η, $\{X_t\}_{t\geq 0}$ is recurrent, and with probability tending to 1 as $T \to \infty$ the return probability of $\{X_t\}_{t\geq 0}$ at time $2T$ decays like $T^{-1+o(1)}$. We also proved a version of *subdiffusive* behavior by showing ([9, Theorem 1.2]) that the expected exit time from a ball of radius N scales as $N^{\psi(\gamma)+o(1)}$ with $\psi(\gamma) > 2$ (explicitly defined) for all $\gamma > 0$. One can then expect $|X_T|$ to scale like $T^{1/\psi(\gamma)+o(1)}$ for large T which would make $1/\psi(\gamma)$ the *diffusive exponent* of the walk. In fact, we were able to prove a corresponding lower bound in [9, Theorem 1.3]. Our bounds are consistent with the predictions made in [13, 14].

The main difficulty of working with this model is that the network law is not shift invariant which makes most of the existing theory in *random conductance model* inapplicable. The proofs of our results exploit the connection between random walks and effective resistances for the underlying network (see, e.g., [49]) and involve delicate control on the latter. In particular, we showed ([9, Theorem 1.4]) that the effective resistance between two vertices at Euclidean distance N behaves as $N^{o(1)}$. Considering that effective resistance is a fundamental metric for a graph, our work in [9] sheds light on a metric property of planar GFF which is different from the LFPP. The proof of [9, Theorem 1.4] relies heavily on planarity and uses a novel combination of duality, Gaussian concentration inequality and the Russo-Seymour-Welsh theory. A point worth noting from our discussions in the previous and current section is that putting random weights/conductances as exponential of the GFF substantially distorts the graph distance of $\mathbb{Z}^2$ but does not significantly distort the resistance metric of $\mathbb{Z}^2$.

Future directions:

Our method of estimating effective resistances can be easily adapted to some other two-dimensional log-correlated Gaussian fields (e.g., those considered in [50]). An interesting question therefore would be:

Question 1. Characterize the right universality class of log-correlated Gaussian fields for the $N^{o(1)}$ growth of the effective resistances.

One important ingredient in our estimate of effective resistance is the Gaussian concentration inequality. This poses the question:

Question 2. Can we derive an analogue of our estimate on the effective resistance when the underlying random media is not a Gaussian process? A natural class of processes for which one can try this is that of random gradient fields.

There are many open problems related to the random walk $\{X_t\}_{t\geq 0}$. For example, a long term goal is:

Question 3. Derive an appropriate scaling limit of the whole problem. This includes the walk as well as the resistance metric and associated current and voltage configurations.

However, our immediate goals are more modest.

Question 4. Compute the *spectral dimension* of $\{X_t\}_{t\geq 0}$ which amounts to an almost sure version of a result we already prove in [9]. Also show that the lower bound $1/\psi(\gamma)$ we obtain on the diffusive exponent is in fact sharp.

5 Other works

Percolation of averages in the stochastic mean field model. Equip the edges of a complete graph on n vertices with i.i.d. exponential weights of mean n. This is known as the *stochastic mean-field* (SMF_n) model. Now given any $\lambda > 0$, consider the length $L(n, \lambda)$ of the longest path whose average edge weight (i.e., the total edge weight divided by the number of edges) is at most λ. The study of this object was initiated by Aldous [3] who proved that with high probability $L(n, \lambda) = O(\log n)$ for $\lambda < 1/\mathrm{e}$ and $\Theta(n)$ for $\lambda > 1/\mathrm{e}$. The critical behavior was established in [19], where it was shown that with high probability $L(n, \lambda)$ is $\Theta((\log n)^3)$ when λ is around $1/\mathrm{e}$ within a window of order $(\log n)^{-2}$. A natural question is the asymptotic behavior of $L(n, \lambda)/n$ in the supercritical regime $\lambda > 1/\mathrm{e}$ where a power law behavior was conjectured by Aldous [4] for λ only slightly bigger than $1/\mathrm{e}$ (the *near-supercritical* regime). More specifically Aldous predicted by a non-rigorous analysis that $L(n, \lambda) \asymp n(\lambda - 1/\mathrm{e})^3$ when $\lambda - 1/\mathrm{e} > 0$ is fixed but sufficiently small.

We revisited this problem in [20] and found $L(n, \lambda)$ to be much smaller to be described by any power law. In particular, we showed ([20, Theorem 1.1]) that in the near-supercritical regime, $L(n, \lambda)/n$ behaves like $\mathrm{e}^{-\Theta(1)/\sqrt{\lambda - 1/\mathrm{e}}}$. This is quite interesting given that power-law behavior often seems "natural" in similar situations (for instance, by analogy with the case of percolation as in [4]). Our proof of [20, Theorem 1.1] is based on first and second moment methods and a delicate use of sprinkling argument.

Finite size scaling of random XORSAT. Consider a random instance of k-XORSAT which is a system of n equations over m variables in $\mathbb{F}_2$. Each equation is of the form $y_1 + y_2 + \cdots + y_k = b$ where k is fixed, $y_1, y_2, \cdots$ are variables (not necessarily distinct) and $b \in \mathbb{F}_2$. The equations are chosen independently and uniformly at random with replacement. Let $P_k(m, n)$ denote the probability that this system is solvable in $\mathbb{F}_2$. It is known that $P_k(m,n)$ exhibits a sharp phase transition around a critical value ρ_k of the ratio m/n when $k \geq 3$. More precisely, $\lim_{m,n\to\infty; m/n\to\rho} P_k(m,n) = 1$ or 0 accordingly as $\rho > \rho_k$ or $< \rho_k$ respectively. This was first shown by Dubois and Mandler [28] for $k = 3$ and independently by Dietzfelbinger et al. [18] and Pittel and Sorkin [56] for all $k \geq 3$.

The critical behavior of $P_k(m,n)$ around the satisfiability threshold is a natural question which we investigated in [41]. Our main result ([41, Theorem 1.1]) showed that for any $r \in \mathbb{R}$, $m = \lfloor n\rho_k + rn^{1/2} \rfloor$ and all large n, $|P_k(m,n) - \Phi(rs_k)| \leq n^{-c^*}$ where $\Phi(.)$ is the standard Gaussian distribution function, $s_k > 0$ depends only on k and c^* is some positive constant. Our proof draws upon the tools developed in [17] for studying the finite size scaling behavior of a different but related problem as well as a result from [56] on the transition window of a random XORSAT model with more "constraints".

Estimation of functions under regularity constraints. Originally introduced in [65], two-dimensional Total Variation Denoising (TVD) is a widely used technique for image denoising. It is also an important nonparametric regression method for estimating functions with heterogeneous smoothness. Recent results have shown the TVD estimator to be nearly minimax rate optimal for the class of functions with bounded variation. In our work in [16], we complement these worst case guarantees by investigating the adaptivity of the TVD estimator to functions which are piecewise constant on axis aligned rectangles. We rigorously show that, when the truth is piecewise constant, the ideally tuned TVD estimator performs better than in the worst case. We also study the issue of choosing the tuning parameter. In particular, we propose a fully data driven version of the TVD estimator which enjoys similar worst case risk guarantees as the ideally tuned TVD estimator. In a forthcoming work [15] we propose a computationally efficient methodology which we prove to have nearly-optimal statistical risk for similar class of estimation problems in more general set-up.

6 Research outlook

My broader goal through all the research questions mentioned in this statement and beyond is to investigate physically relevant problems in statistical mechanics by rigorous methods in probability theory. The liaison between mathematical physics and probability theory is old and beneficial for both fields. In the context of two-dimensional conformal field theory, the probabilistic methods have already been proven to be successful in the case of Ising model and Bernoulli percolation (see [30] and the references therein). Very recently, the conformal structure of the limit of random triangulations on sphere and its connection to Liouville quantum gravity has been described in a series of works by Miller and Sheffield [52, 53, 54].

Although many important questions still remain unanswered for planar models, the understanding in intermediate dimensions, i.e., between the planar and the mean-field regime, is comparatively limited. Partly this can be attributed to the lack of integrability in standard three (or higher) dimensional models in statistical mechanics and partly to the fact that the conformal groups in these dimensions are more restricted. Through my research endeavors, I would like to bring new insights into this area in two different ways. Firstly, I want to *modify* the models so as to allow more integrability in certain sense. It may be even possible to ensure, at the heuristic level, that these modified models belong to the same universality class as their original counterparts. Owing to the presence of additional structure, analyzing the behavior of such models can be hoped to be more tractable and would thus provide valuable information. Secondly, I would like to explore the recently developed technology of *conformal bootstrap* (see, e.g., [66]) in conformal field theory which has been very effective in estimating the universal exponents for the classical models in statistical mechanics in dimension larger than 2. My long term goal in this regard is to understand the key ideas in this domain and develop their parallel in probability theory. A significant part of my effort would be to communicate and collaborate with physicists working in related fields.

Apart from statistical physics, rich sources of new research problems in probability theory are machine learning and modern statistics. New methodologies are being developed in neural networks, pattern recognition, nonparametric regression among others. While the practical utility of these methods are being tested in different contexts, there is an enormous scope for their theoretical understanding which may lead to the development of new ideas in probability theory. Presence of a theoretical framework to analyze these methodologies is also helpful for researchers working in more applied domains who can use the insights obtained from it to design new methods. I look forward to continuing my collaboration with statisticians and scientists working in machine learning to explore this exciting territory of research.

References

[1] M. Aizenman, D. J. Barsky, and R. Fernández. The phase transition in a general class of Ising-type models is sharp. *J. Statist. Phys.*, 47(3-4):343–374, 1987.

[2] M. Aizenman, H. Duminil-Copin, and V. Sidoravicius. Random Currents and Continuity of Ising Model's Spontaneous Magnetization. *Communications in Mathematical Physics*, 334:719–742, 2015.

[3] D. Aldous. On the critical value for "percolation" of minimum-weight trees in the mean-field distance model. *Combin. Probab. Comput.*, 7(1):1–10, 1998.

[4] D. J. Aldous. Percolation-like scaling exponents for minimal paths and trees in the stochastic mean field model. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 461(2055):825–838, 2005.

[5] J. Ambjørn and T. G. Budd. Geodesic distances in Liouville quantum gravity. *Nuclear Phys. B*, 889:676–691, 2014.

[6] T. Balaban and M. O'Carroll. Low temperature properties for correlation functions in classical N-vector spin models. *Comm. Math. Phys.*, 199(3):493–520, 1999.

[7] I. Benjamini and O. Schramm. Percolation beyond $\mathbf{Z}^\mathbf{d}$, many questions and a few answers. *Electron. Comm. Probab.*, 1:no. 8, 71–82 (electronic), 1996.

[8] N. Berestycki. Diffusion in planar Liouville quantum gravity. to appear in *Ann. Inst. Henri Poincaré Probab. Stat.*

[9] M. Biskup, J. Ding, and S. Goswami. Return probability and recurrence for the random walk driven by two-dimensional Gaussian free field. To appear in *Communications in Mathematical Physics*. Preprint available at https://arxiv.org/abs/1611.03901.

[10] T. Bodineau. Slab percolation for the Ising model. *Probab. Theory Related Fields*, 132(1):83–118, 2005.

[11] T. Bodineau. Translation invariant Gibbs states for the Ising model. *Probab. Theory Related Fields*, 135(2):153–168, 2006.

[12] J. Bricmont, J. L. Lebowitz, and C. Maes. Percolation in strongly correlated systems: the massless Gaussian field. *J. Statist. Phys.*, 48(5-6):1249–1268, 1987.

[13] D. Carpentier and P. Le Doussal. Glass transition of a particle in a random potential, front selection in nonlinear rg and entropic phenomena in Liouville and sinhgordon models. *Phys. rev. E.*, 63:026110, 2001.

[14] H. E. Castillo and P. Le Doussal. Freezing of dynamical exponents in low dimensional random media. *Phys. Rev. Lett.*, 86:4859, 2001.

[15] S. Chatterjee and S. Goswami. Adaptive estimation of multivariate piecewise polynomials and bounded variation functions by optimal decision trees. 2019. Preprint, available at http://arxiv.org/abs/1911.11562.

[16] S. Chatterjee and S. Goswami. New risk bounds for 2d total variation denoising. 2019. Preprint, available at http://arxiv.org/abs/1902.01215.

[17] A. Dembo and A. Montanari. Finite size scaling for the core of large random hypergraphs. *Ann. Appl. Probab.*, 18(5):1993–2040, 2008.

[18] M. Dietzfelbinger, A. Goerdt, M. Mitzenmacher, A. Montanari, R. Pagh, and M. Rink. *Tight Thresholds for Cuckoo Hashing via XORSAT*, pages 213–225. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[19] J. Ding. Scaling window for mean-field percolation of averages. *Ann. Probab.*, 41(6):4407–4427, 2013.

[20] J. Ding and S. Goswami. Percolation of averages in the stochastic mean field model: the near-supercritical regime. *Electron. J. Probab.*, 20:no. 124, 21, 2015.

[21] J. Ding and S. Goswami. First passage percolation on the exponential of two-dimensional branching random walk. *Electronic Communications in Probability*, 22(69), 2017.

[22] J. Ding and S. Goswami. Upper bounds on Liouville first passage percolation and Watabiki's prediction. *Commun. Pure Appl. Math.*, 72, 2019. Preprint available at http://arxiv.org/abs/1610.09998.

[23] J. Ding and E. Gwynne. The fractal dimension of liouville quantum gravity: universality, monotonicity, and bounds. 2018. Preprint, available at http://arxiv.org/abs/1807.01072.

[24] J. Ding and M. Wirth. Percolation for level-sets of gaussian free fields on metric graphs. 2018. Preprint, available at http://arxiv.org/abs/1807.11117.

[25] J. Ding, O. Zeitouni, and F. Zhang. Heat kernel for liouville brownian motion and liouville graph distance. 2018. Preprint, available at http://arxiv.org/abs/1807.00422.

[26] R. L. Dobrushin and S. B. Shlosman. Completely analytical interactions: constructive description. *J. Statist. Phys.*, 46(5-6):983–1014, 1987.

[27] A. Drewitz, B. Ráth, and A. Sapozhnikov. On chemical distances and shape theorems in percolation models with long-range correlations. *J. Math. Phys.*, 55(8):083307, 30, 2014.

[28] O. Dubois and J. Mandler. The 3-XORSAT threshold. *C. R. Math. Acad. Sci. Paris*, 335(11):963–966, 2002.

[29] H. Duminil-Copin. Random currents expansion of the Ising model. arXiv:1607:06933, 2016.

[30] H. Duminil-Copin. Lectures on the ising and potts models on the hypercubic lattice. 2017. Preprint, available at http://arxiv.org/abs/1707.00520.

[31] H. Duminil-Copin, S. Goswami, and A. Raoufi. Exponential decay of truncated correlations for the ising model in any dimension for all but the critical temperature. *Commun. Math. Phys.*, 2019. Preprint available at http://arxiv.org/abs/1808.00439.

[32] H. Duminil-Copin, S. Goswami, A. Raoufi, F. Severo, and A. Yadin. Existence of phase transition for percolation using the Gaussian free field. Preprint, available at http://arxiv.org/abs/1806.07733, 2018.

[33] H. Duminil-Copin, S. Goswami, P.-F. Rodriguez, and F. Severo. Equality of critical parameters for percolation of Gaussian free field level-sets. 2020. Preprint available at http://arxiv.org/abs/2002.07735.

[34] H. Duminil-Copin, A. Raoufi, and V. Tassion. Sharp phase transition for the random-cluster and Potts models via decision trees. *Ann. of Math. (2)*, 189(1):75–99, 2019.

[35] H. Duminil-Copin and V. Tassion. A note on schramm's locality conjecture for random-cluster models. 2017. Preprint, available at http://arxiv.org/abs/1707.07626.

[36] H. Duminil-Copin and V. Tassion. Renormalization of crossing probabilities in the planar random-cluster model. 2019. Preprint, available at http://arxiv.org/abs/1901.08294.

[37] B. Duplantier and S. Sheffield. Liouville quantum gravity and KPZ. *Invent. Math.*, 185(2):333–393, 2011.

[38] C. M. Fortuin and P. W. Kasteleyn. On the random-cluster model. I. Introduction and relation to other models. *Physica*, 57:536–564, 1972.

[39] J. Fröhlich, B. Simon, and T. Spencer. Infrared bounds, phase transitions and continuous symmetry breaking. *Comm. Math. Phys.*, 50(1):79–95, 1976.

[40] C. Garban, R. Rhodes, and V. Vargas. On the heat kernel and the Dirichlet form of Liouville Brownian motion. *Electron. J. Probab.*, 19:no. 96, 25, 2014.

[41] S. Goswami. Finite size scaling of random XORSAT. Preprint, available at https://arxiv.org/abs/1610.07431.

[42] G. Grimmett. *The random-cluster model*, volume 333 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.

[43] G. R. Grimmett and J. M. Marstrand. The supercritical phase of percolation is well behaved. *Proc. Roy. Soc. London Ser. A*, 430(1879):439–457, 1990.

[44] J.-P. Kahane. Sur le chaos multiplicatif. *Ann. Sci. Math. Québec*, 9(2):105–150, 1985.

[45] J. L. Lebowitz and O. Penrose. Analytic and clustering properties of thermodynamic functions and distribution functions for classical lattice and continuum systems. *Comm. Math. Phys.*, 11:99–124, 1968/1969.

[46] J. L. Lebowitz and H. Saleur. Percolation in strongly correlated systems. *Physica A: Statistical Mechanics and its Applications*, 138(1-2):194–205, 1986.

[47] T. M. Liggett. *Interacting particle systems*. Classics in Mathematics. Springer-Verlag, Berlin, 2005. Reprint of the 1985 original.

[48] T. Lupu. From loop clusters and random interlacements to the free field. *Ann. Probab.*, 44(3):2117–2146, 05 2016.

[49] R. Lyons and Y. Peres. *Probability on Trees and Networks*. Cambridge University Press, 2016. Available at http://pages.iu.edu/~rdlyons/.

[50] T. Madaule. Maximum of a log-correlated Gaussian field. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(4):1369–1431, 2015.

[51] F. Martinelli, E. Olivieri, and R. H. Schonmann. For 2-D lattice spin systems weak mixing implies strong mixing. *Comm. Math. Phys.*, 165(1):33–47, 1994.

[52] J. Miller and S. Sheffield. Liouville quantum gravity and the brownian map I: The QLE(8/3,0) metric. Preprint, available at http://arxiv.org/abs/1507.00719.

[53] J. Miller and S. Sheffield. Liouville quantum gravity and the brownian map ii: geodesics and continuity of the embedding. 2016. Preprint, available at http://arxiv.org/abs/1605.03563.

[54] J. Miller and S. Sheffield. Liouville quantum gravity and the brownian map III: the conformal structure is determined. 2016. Preprint, available at https://arxiv.org/abs/1608.05391.

[55] R. Peierls. On Ising's model of ferromagnetism. *Math. Proc. Camb. Phil. Soc.*, 32:477–481, 1936.

[56] B. Pittel and G. B. Sorkin. The satisfiability threshold for k-XORSAT. *Combinatorics, Probability and Computing*, 25(02):236–268, 2016.

[57] A. M. Polyakov. Quantum geometry of bosonic strings. *Phys. Lett. B*, 103(3):207–210, 1981.

[58] S. Popov and B. Ráth. On decoupling inequalities and percolation of excursion sets of the Gaussian free field. *J. Stat. Phys.*, 159(2):312–320, 2015.

[59] E. B. Procaccia, R. Rosenthal, and A. Sapozhnikov. Quenched invariance principle for simple random walk on clusters in correlated percolation models. *Probab. Theory Related Fields*, 166(3-4):619–657, 2016.

[60] A. Raoufi. Translation-invariant gibbs states of ising model: general setting. 2017. Preprint, available at http://arxiv.org/abs/1710.07608.

[61] B. Ráth and D. Valesin. Percolation on the stationary distributions of the voter model. *Ann. Probab.*, 45(3):1899–1951, 2017.

[62] R. Rhodes and V. Vargas. KPZ formula for log-infinitely divisible multifractal random measures. *ESAIM Probab. Stat.*, 15:358–371, 2011.

[63] R. Rhodes and V. Vargas. Gaussian multiplicative chaos and applications: a review. *Probab. Surv.*, 11:315–392, 2014.

[64] P.-F. Rodriguez and A.-S. Sznitman. Phase transition and level-set percolation for the Gaussian free field. *Comm. Math. Phys.*, 320(2):571–601, 2013.

[65] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268, 1992.

[66] S. Rychkov. *EPFL lectures on conformal field theory in D 3 dimensions.* Springer, 2017.

[67] A. Shamov. On Gaussian multiplicative chaos. *J. Funct. Anal.*, 270(9):3224–3261, 2016.

[68] S. Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Related Fields*, 139(3-4):521–541, 2007.

[69] A.-S. Sznitman. Vacant set of random interlacements and percolation. *Ann. of Math. (2)*, 171(3):2039–2087, 2010.

[70] A.-S. Sznitman. An isomorphism theorem for random interlacements. *Electron. Commun. Probab.*, 17:no. 9, 9, 2012.

[71] A.-S. Sznitman. Disconnection and level-set percolation for the gaussian free field. *J. Math. Soc. Japan*, 67(4):1801–1843, 10 2015.

[72] A.-S. Sznitman. Coupling and an application to level-set percolation of the Gaussian free field. *Electron. J. Probab.*, 21:Paper No. 35, 26, 2016.

[73] Y. Watabiki. Analytic study of fractal structure of quantized surface in two-dimensional quantum gravity. *Progress of Theoretical Physics Supplement*, 1993(114):1–17, 1993.

[74] A. Weinrib and B. Halperin. Critical phenomena in systems with long-range-correlated quenched disorder. *Physical Review B*, 27(1):413, 1983.