Yang-Mills Learning Seminar: Shen-Zhu-Zhu '22, Part II

Ansh S.

October 17, 2025

0 Summary

We continue working through [1]. In these notes we basically see the following (strict subset of section 4.1 in the original paper):

$$
\begin{aligned}
\text{Estimate on Hessian (Lemma 2)} &\implies \text{Bakry-Émery condition (9) is satisfied} \\
&\iff \text{Temporal mixing under Langevin dynamics (Theorem 1)} \\
&\implies \text{Log-Sobolev inequality (Theorem 2)} \\
&\implies \text{Decay of correlations and other niceties.}
\end{aligned}
$$

The last implication will be covered in Ron's talk on October 24, 2025.

1 Review

Let $\Lambda_L = \mathbb{Z}^d \cap L\mathbb{T}^d$; $d > 1$ be a finite d-dimensional lattice with side length L and unit lattice spacing.

$$
\begin{aligned}
E_L^+ &:= \{\text{+vely oriented edges}\} \\
G &= SO(N) \text{ or } SU(N) \\
\mathcal{P}_L^+ &:= \{\text{+vely oriented plaquettes}\}.
\end{aligned}
$$

We assign a group element $Q_e \in G$ to each edge $e \in E_L^+$, and extend to all edges by requiring that $Q_{e^{-1}} = Q_e^{-1} = Q_e^*$ (for the groups under consideration) where $e \in E_L^+$ and e^{-1} is the same edge but in negative orientation. The lattice Yang-Mills measure now is a probability measure on the space of such configurations $\mu_{L,N,\beta} \in \mathcal{M}^1(G^{E_L^+})$, given by

$$
d\mu_{L,N,\beta}(Q) = \frac{1}{Z_{L,N,\beta}} \exp(\mathcal{S}(Q))\, d\sigma_N(Q) \tag{1}
$$

where σ_N is the Haar measure on the product group $G^{E_L^+}$ and

$$\mathcal{S}(Q) := N\beta\Re \sum_{p\in\mathcal{P}_L^+} \mathrm{Tr}(Q_p) \tag{2}$$

where $Q_p = Q_{e_1}Q_{e_2}Q_{e_3}^*Q_{e_4}^*$ if $e_1, e_2, e_3, e_4 \in E_L^+$ appear in $p \in \mathcal{P}_L^+$ in that order. Usually there is no N in the action, it being there is telling us we can take the inverse temperature to scale like a constant times N. This is the t'Hooft scaling – more on this later.

Here's what Kunal talked about in his talk:

1. We have global well-posedness of the following SDE (Langevin dynamics) on the configuration space:

$$dQ = \nabla\mathcal{S}(Q)\, dt + \sqrt{2}\, d\mathfrak{B} \tag{3}$$

   where $\mathfrak{B} = (\mathfrak{B}_e)_{e\in E_L^+}$ are independent brownian motions on G. $d\mathfrak{B}$ can be seen as the white noise w.r.t. the inner product on $T_Q G^{E_L^+}$ which will be recalled in the next section.

2. By global well-posedness above, the solutions form a Markov process in $G^{E_L^+}$. Let the associated semigroup be $(P_t^L)_{t\geq 0}$, i.e. $\forall f \in C^\infty(G^{E_L^+} \to \mathbb{R})$, $P_t^L f(x) = \mathbb{E}[f(Q(t,x))]$ where $Q(t,x)$ denotes the solution at time t to (3) starting from initial data x.

3. (1) is invariant under (3). That is, for any $f \in C^\infty(G^{E_L^+} \to \mathbb{R})$ we have that

$$\int P_t^L f(x)\, d\mu_{L,N,\beta}(x) = \int f(x)\, d\mu_{L,N,\beta}(x). \tag{4}$$

2 On to the results

Let the lie algebra associated with G be $\mathfrak{g}$, which we can view as a subset of $M_{N\times N}(\mathbb{C})$ and endow with the Hilbert-Schmidt inner product given by $\Re\,\mathrm{Tr}(XY^*)$. This inner product on $\mathfrak{g} = T_{I_N}G$ can be used to induce an inner product on every tangent space $T_Q G$ using the group structure of G. This Riemannian metric is bi-invariant. Now we can decompose the tangent space $T_Q G^{E_L^+} = \oplus_{e\in G^{E_L^+}} T_{Q_e}G$, and get a Riemannian metric on $G^{E_L^+}$.

Accordingly, we use the distance from the Riemannian metric on G to induce a distance on the space of configurations $G^{E_L^+}$ by

$$\rho_L(Q,Q')^2 := \sum_{e\in E_L^+} \rho(Q_e,Q'_e)^2;\ Q,Q' \in G^{E_L^+}$$

where ρ is the distance on G.

We also define the Wasserstein distance on $\mathcal{M}^1(G^{E_L^+})$ by

$$W_{L,p}(\mu,\nu) := \inf_{\pi\in\mathcal{C}(\mu,\nu)} \left(\int_{G^{E_L^+}\times G^{E_L^+}} |\rho_L(x,y)|^p \, d\pi(x,y)\right)^{1/p};\ \mu,\nu\in\mathcal{M}^1(G^{E_L^+})$$

where $\mathcal{C}(\mu,\nu)$ is the set of couplings between μ and ν – that is, $\pi \in \mathcal{M}^1(G^{E_L^+}\times G^{E_L^+})$ such that it has μ and ν as its marginals. This means that π restricted to the first entry is μ, and restricted to the second entry it is ν.

A final piece of notation: for $\nu \in \mathcal{M}^1(G^{E_L^+})$, we define $\nu P_t^L \in M^1(G^{E_L^+})$ to be the measure such that

$$\forall f \in C^\infty(G^{E_L^+}), \int f(x)\, d\nu P_t(x) = \int P_t f(x)\, d\nu.$$

The way to interpret this is that the Markov semigroup $(P_t)_{t\geq 0}$ acts dually on $\mathcal{M}^1(G^{E_L^+})$.

We now state the main results we will discuss today.

Theorem 1. *$K_\mathcal{S} = K_\mathcal{S}(N,\beta)$ is a constant described below. For any N,β, the following is true:*

1. *The dynamic defined by (3) is exponentially ergodic in the sense that*

$$W_{L,2}(\delta_x P_t^L, \delta_y P_t^L) \leq e^{-K_\mathcal{S} t}\rho_L(Q,Q');\ t\geq 0, x,y\in G^{E_L^+}. \tag{5}$$

2. *For $1 < p < 2$,*

$$W_{L,p}(\mu P_t^L, \nu P_t^L) \leq e^{-K_\mathcal{S} t} W_{L,p}(\mu,\nu);\ t\geq 0, \mu,\nu\in\mathcal{M}^1(G^{E_L^+}), \tag{6}$$

which in particular implies uniqueness of the invariant measure when $K_\mathcal{S} > 0$.

3. *$\forall f \in C^\infty(G^{E_L^+}\to\mathbb{R})$,*

$$P_t^L(f^2\log f^2) - (P_t^L f^2)\log(P_t^L f^2) \leq \frac{2(1-e^{-2K_\mathcal{S} t})}{K_\mathcal{S}} P_t^L|\nabla f|^2, \tag{7}$$

which we will use to prove the log-Sobolev inequality below.

Above,

$$K_\mathcal{S} = \begin{cases} \frac{N+2}{4} - 1 - 8N|\beta|(d-1);\ G = SO(N) \\ \frac{N+2}{2} - 1 - 8N|\beta|(d-1);\ G = SU(N) \end{cases}. \tag{8}$$

In order to have decay in time, we will make the following assumption:

Assumption 1. *Suppose that $K_{\mathcal{S}} > 0$, which is equivalent to the following strong coupling assumption:*

$$|\beta| < \begin{cases} \frac{1}{32(d-1)} - \frac{16}{N(d-1)} & ; G = SO(N) \\ \frac{1}{16(d-1)} & ; G = SU(N) \end{cases}.$$

Under this assumption, in particular, we obtain that the invariant measure of $(P_t^L)_{t\geq 0}$ is unique. Indeed, if μ is an invariant measure then μP_t^L satisfies $\int f\, d\mu P_t^L = \int P_t^L f\, d\mu = \int f\, d\mu$ and hence $\mu P_t^L = \mu$. If we had two invariant measures μ and ν, then we'd get $W_{L,p}(\mu,\nu) = W_{L,p}(\mu P_t^L, \nu P_t^L) \leq e^{-K_{\mathcal{S}} t} W_{L,p}(\mu,\nu) \to 0$ as $t \to \infty$ using (6). In Kunal's talk we saw that $\mu_{L,N,\beta}$ is unique under the dynamics $(P_t^L)_{t\geq 0}$. The above shows that it is the unique invariant measure as long as the inverse temperature is less than N times some small constant.

The following proposition reduces Theorem 1 to checking a condition.

Proposition 1. *Each assertion of Theorem 1 and the following Bakry-Émery condition are all equivalent: for every $v \in T_Q G^{E_L^+}$ at any Q,*

$$\mathrm{Ric}(v,v) - \mathrm{Hess}_{\mathcal{S}}(v,v) \geq K_{\mathcal{S}} |v|^2. \tag{9}$$

Proof. For the full proof see [2, Theorem 5.6.1 (1), (11), (12), (6)]. Here, we will only prove that (9) $\Longrightarrow$ (5) in a flat space.

Let the configuration space be $\mathbb{R}^n$. Suppose $S : \mathbb{R}^n \to \mathbb{R}$ satisfies (9), and consider the following SDE on $\mathbb{R}^n$:

$$dq = \nabla S(q) + \sqrt{2}\, dB$$

where B is just the standard brownian motion on $\mathbb{R}^n$. The Bakry-Émery condition then tells us that

$$-\nabla^2 S \geq K$$

which consequently gives us an estimate on $(\nabla S(x) - \nabla S(y)) \cdot (x-y)$ as follows:

$$\nabla S(x) - \nabla S(y) = \int_0^1 \nabla^2 S(\gamma(t))(x-y)\, dt$$

and hence

$$\begin{aligned}(\nabla S(x) - \nabla S(y)) \cdot (x-y) &\leq \int_0^1 (x-y)^T \nabla^2 S(\gamma(t))(x-y)\, dt \\ &\leq -K|x-y|^2.\end{aligned}$$

Next, couple $q(t,x)$ and $q(t,y)$ with the same Brownian motion. That is, suppose that they are solutions to the above SDE at time t starting with data x and y respectively, driven by the same Brownian motion. Then the difference $q(t,x)-q(t,y)$ actually satisfies an ODE. We get,

$$\begin{aligned}\frac{d}{dt}|q(t,x)-q(t,y)|^2 &= 2(\nabla S(q(t,x)) - \nabla S(q(t,y))\cdot(q(t,x)-q(t,y)) \\ &\leq -2K|q(t,x)-q(t,y)|^2.\end{aligned}$$

Gronwall's inequality now implies

$$|q(t,x)-q(t,y)|^2 \leq e^{-2Kt}|x-y|^2$$

which implies the result for the particular coupling used above. Minimizing over all couplings readily gives (5). ☐

I again want to emphasize that the assertions in Theorem 1 are equivalent to the Bakry-Émery condition (9) regardless of the value of $K_{\mathcal{S}}$. Of course, these assertions are only useful for us under Assumption 1. Under this assumption we obtain the Log-Sobolev inequality, which uses the temporal mixing seen in Theorem 1 to get spatial mixing.

Theorem 2. *(Log-Sobolev inequality) For any $F: G^{E_L^+} \to \mathbb{R}$, let $\mu_{L,N,\beta}(F) := \int f(x)\, d\mu_{L,N,\beta}(x)$. Then for any $f \in C^\infty(G^{E_L^+} \to \mathbb{R})$, under Assumption 1 we have*

$$\mu_{L,N,\beta}(f^2 \log f^2) \leq \frac{2}{K_{\mathcal{S}}}\mu_{L,N,\beta}(|\nabla f|^2) + \mu(f^2)\log\mu(f^2). \tag{10}$$

Proof. From (7) we have that

$$P_t^L(f^2\log f^2) \leq \frac{2(1-e^{-2K_{\mathcal{S}}t})}{K_{\mathcal{S}}}P_t^L|\nabla f|^2 + (P_t^L f^2)\log(P_t^L f^2). \tag{11}$$

taking integral w.r.t. $\mu_{L,N,\beta}$ on both sides and using invariance of measure (4) gives

$$\mu_{L,N,\beta}(f^2\log f^2) \leq \frac{2(1-e^{-2K_{\mathcal{S}}t})}{K_{\mathcal{S}}}\mu_{L,N,\beta}(|\nabla f|^2) + \mu_{L,N,\beta}((P_t^L f^2)\log(P_t^L f^2)). \tag{12}$$

Intuitively, the Bakry-Émery condition is saying that there is a spectral gap in the spectrum of the generator of our semigroup. Therefore any deviation from the equilibrium of the sort $P_t^L F - \mu(F)$ should decay exponentially in time in the L^2 sense. That is, $P_t^L F \to \mu(F)$ in $L^2(\mu_{L,N,\beta})$, using which we get

$$\to \frac{2}{K_{\mathcal{S}}}\mu_{L,N,\beta}(|\nabla f|^2) + \mu_{N,L,\beta}(f^2)\log\mu(f^2). \tag{13}$$

☐

The reason we care about the Log-Sobolev inequality is that by choosing f judiciously we can obtain things like decay of correlations. Note that we trivially have decay of correlations when $\beta = 0$ in the original action without the t'Hooft scaling. It gets harder to show this the larger we take the inverse temperature to be, and that's what's impressive about this result: we can show decay of correlations even when the inverse temperature is increasing linearly with N (times some small constant).

Remember that the Log-Sobolev inequality is implied by satisfying the Bakry-Émery condition as long as the inverse temperature is less than some function of N. It is then interesting to ask if we can get Log-Sobolev without satisfying the condition, to perhaps get things like decay of correlation for all values of the inverse temperature.

Anyhow, we now show that the Bakry-Émery condition is satisfied. This follows from the following two lemmas and the definition of $K_{\mathcal{S}}$ in (8).

Lemma 1. *For any tangent vector $u \in G$ we have*

$$\mathrm{Ric}(u, u) = \left(\frac{\alpha(N+2)}{4} - 1 \right) |u|^2.$$

Proof. Instead of a proof, we give an intuitive argument on why the Ricci curvature scales linearly with N. Our Riemannian metric is invariant under the adjoint action. For such metrics, the Ricci curvature is given by

$$\mathrm{Ric}[X, X] = \frac{1}{4} \sum_i |[X, E_i]|^2$$

where the sum is over the orthonormal basis of $\mathfrak{g}$. When $\mathfrak{g} = \mathfrak{so}(N)$, X gives the infinitesimal rotation in some plane. The question then, is how many other planes is this rotation witnessed in? It'll be precisely these planes which will fail to commute with X. If $X = E_{jk}$, then the planes will be those containing either j or k – which will be $N - 2$ many. Hence the linear scaling! □

Lemma 2. *For $v \in T_Q G^{E_L^+}$ at any point Q we have*

$$|\operatorname{Hess}_{\mathcal{S}}(v, v)| \le 8(d-1)N|\beta||v|^2.$$

Before proceeding with the proof of this lemma, we note that $K_{\mathcal{S}}$ is simply the difference between Ric and the above estimate on $\operatorname{Hess}_{\mathcal{S}}$. The computation of the Ricci curvature is exact: it grows linearly. If we had a worse than linear estimate on the Hessian, then $K_{\mathcal{S}}$ would no longer be positive for every N regardless of how small we chose β to be. Naïvely, we'd expect the Hessian to be of the order N^2: suppose the trace is $O(N)$, and then there is an N on the outside from the definition of the Hessian – but this is too big! The novel contribution in [1] is precisely Lemma 2. The better estimates we can get on the Hessian, the larger we can take our inverse temperature to be and, for example, still get decay of correlations.

Proof. Let $v, w \in T_Q G^{E_L^+}$ and X and Y be vector fields $G^{E_L^+} \to TG^{E_L^+}$ such that $X(Q) = v$ and $Y(Q) = w$. Then the hessian is given by

$$\mathrm{Hess}_{\mathcal{S}}(v, w) = X(Y(S)) - (\nabla_X Y)(S).$$

The LHS is a tensor field and only depends on the vectors at Q, but the individual objects on the RHS by themselves aren't and can depend on the vector field locally around Q. This means that we can choose vector fields conveniently to make the computation of the RHS simpler, and our choice won't affect the LHS that we want to estimate! To this end, we will consider a right-invariant vector field. For $v \in T_Q G^{E_L^+}$ recall that we can write $v = XQ$ where $X \in \mathfrak{g}$. Then, we construct our right-invariant vector field $\tilde{X}$ by assigning the vector $\tilde{X}(Q') = XQ'$ at every point $Q' \in G^{E_L^+}$. For such vector fields, the Levi-Civita connection is quite simple:

$$\nabla_{\tilde{X}} \tilde{Y} = \frac{1}{2}[\tilde{X}, \tilde{Y}].$$

And thus, the second term in the hessian that we want to compute vanishes:

$$\nabla_{\tilde{X}} \tilde{X} = \frac{1}{2}[\tilde{X}, \tilde{X}] = 0.$$

Also note that due to the decomposition of tangent space we have

$$v = XQ = \sum_{e \in E_L^+} X_e Q_e \tag{14}$$

and thus

$$\begin{aligned} \mathrm{Hess}_{\mathcal{S}}(v, v) &= \tilde{X}(\tilde{X}(S)) \\ &= \sum_{e, \bar{e}} (X_{\bar{e}} Q_{\bar{e}})(X_e Q_e) S. \end{aligned} \tag{15}$$

Here we first try to compute $X_e Q_e(S)$. This should be the derivative of S in the direction of $X_e Q_e$. This is computed by considering a curve $\gamma : [0, 1] \to G$ such that $\gamma(0) = Q_e$ and the derivative of the curve at $t = 0$ is the vector XQ. This curve is $e^{tX_e} Q_e$. Thus, supposing $Q_p = Q_{e_1} Q_{e_2} Q_{e_3}^* Q_{e_4}^*$ and that $e = e_3$ we get

$$\begin{aligned} X_{e_3} Q_{e_3}(Q_p) &= Q_{e_1} Q_{e_2} (X_{e_3} Q_{e_3}) Q_{e_3}^* Q_{e_4}^* \\ &= Q_{e_1} Q_{e_2} \frac{d}{dt} (\exp(t X_{e_3}) Q_{e_3})^* Q_{e_4}^* \\ &= Q_{e_1} Q_{e_2} Q_{e_3}^* X_{e_3}^* Q_{e_4}^*. \end{aligned}$$

Next, we want to compute $(X_{\bar{e}} Q_{\bar{e}})(X_e Q_e) Q_p$. This leads to two cases:

<u>Case 1:</u> $e = \bar{e}$. Let $\bar{e} = e = e_3$ in the above computation. Then,

$$\begin{aligned} (X_{e_3} Q_{e_3})(X_{e_3} Q_{e_3}) Q_p &= (X_{e_3} Q_{e_3}) Q_{e_1} Q_{e_2} Q_{e_3}^* X_{e_3}^* Q_{e_4}^* \\ &= Q_{e_1} Q_{e_2} Q_{e_3}^* (X_{e_3}^*)^2 Q_{e_4}. \end{aligned}$$

We now estimate the trace:

$$|(X_{e_3}Q_{e_3})(X_{e_3}Q_{e_3})\Re\operatorname{Tr}(Q_p)| \le |\operatorname{Tr}(Q_{e_1}Q_{e_2}Q_{e_3}^*(X_{e_3}^*)^2Q_{e_4})|,$$

using the cyclic invariance of the trace

$$\le |\operatorname{Tr}(X_{e_3}^*(Q_{e_3}Q_{e_2}^*Q_{e_1}^*Q_{e_4}X_{e_3})^*)|$$

writing $Q' = I_N$, $Q'' = Q_{e_3}Q_{e_2}^*Q_{e_1}^*Q_{e_4}$, and using Cauchy-Schwarz for the Hilbert-Schmidt inner product gives

$$\le (\operatorname{Tr}(Q'X_{e_3}(Q'X_{e_3})^*))^{1/2}(\operatorname{Tr}(Q''X_{e_3}(Q''X_{e_3})^*))^{1/2}$$

which using the bi-invariance of the metric is just

$$= |X_{e_3}|^2.$$

The contribution from these diagonal terms to the Hessian will thus be bounded by

$$N|\beta| \sum_{e=\bar{e}\in E_L^+} \sum_{p\in\mathcal{P}_L^+} |X_e|^2 \mathbb{1}_{e\in p} = N|\beta| \sum_{e\in E_L^+} |X_e|^2 \sum_{p\in P_L^+} \mathbb{1}_{e\in p},$$

there will be $2(d-1)$ such plaquettes containing some fixed edge e, and so

$$= 2N|\beta|(d-1) \sum_{e\in E_L^+} |X_e|^2$$

which using (14) is just

$$= 2N|\beta|(d-1)|v|^2. \tag{16}$$

Case 2: $e \neq \bar{e}$. Let's say $\bar{e} = e_1$ in the original computation. Then

$$\begin{aligned}(X_{e_1}Q_{e_1})(X_{e_3}Q_{e_3})Q_p &= (X_{e_1}Q_{e_1})Q_{e_1}Q_{e_2}Q_{e_3}^*X_{e_3}^*Q_{e_4}^* \\ &= X_{e_1}Q_{e_1}Q_{e_2}Q_{e_3}^*X_{e_3}^*Q_{e_4}\end{aligned}$$

and the trace as before can be bounded by

$$|(X_{e_1}Q_{e_1})(X_{e_3}Q_{e_3})\Re\operatorname{Tr}(Q_p)| \le |\operatorname{Tr}(X_{e_1}Q_{e_1}Q_{e_2}Q_{e_3}^*(Q_{e_4}X_{e_3})^*)|,$$

setting $Q' = Q_{e_1}Q_{e_2}Q_{e_3}^*$ and $Q'' = Q_{e_4}$ gives

$$\begin{aligned}&\le (\operatorname{Tr}(X_{e_1}Q'(X_{e_1}Q')^*))^{1/2}(\operatorname{Tr}(Q''X_{e_3}(Q''X_{e_3})^*))^{1/2} \\ &= |X_{e_1}||X_{e_3}|\end{aligned}$$

which using AM-GM inequality is

$$\le \frac{|X_{e_1}|^2 + |X_{e_3}|^2}{2}.$$

Other non-diagonal terms can be estimated in a similar way. The contribution from these terms to the Hessian can now be bounded by

$$
\begin{aligned}
N|\beta| \sum_{e \neq \bar{e} \in E_L^+} \sum_{p \in \mathcal{P}_L^+} \frac{|X_e|^2 + |X_{\bar{e}}|^2}{2} \mathbb{1}_{e, \bar{e} \in p} &= N|\beta| \sum_{e \neq \bar{e} \in E_L^+} \sum_{p \in P_L^+} |X_e|^2 \mathbb{1}_{e \in p} \mathbb{1}_{\bar{e} \in p} \\
&= N|\beta| \sum_{e \in E_L^+} |X_e|^2 \sum_{p \in \mathcal{P}_L^+} \mathbb{1}_{e \in p} \sum_{\substack{\bar{e} \in E_L^+ \\ \bar{e} \neq e}} \mathbb{1}_{\bar{e} \in p}
\end{aligned}
$$

where for a fixed edge e and plaquette $p \ni e$, there are only three choices for $\bar{e} \in p$ such that $\bar{e} \neq e$. Thus

$$
\begin{aligned}
&= 3N|\beta| \sum_{e \in E_L^+} |X_e|^2 \sum_{p \in \mathcal{P}_L^+} \mathbb{1}_{\bar{e} \in p} \\
&= 6N|\beta|(d-1) \sum_{e \in E_L^+} |X_e|^2 \\
&= 6N|\beta|(d-1)|v|^2. \qquad (17)
\end{aligned}
$$

Finally, using (16) and (17) in (15) gives

$$
\begin{aligned}
|\operatorname{Hess}_{\mathcal{S}}(v,v)| &\leq \sum_{e, \bar{e}} |(X_{\bar{e}} Q_{\bar{e}})(X_e Q_e) S| \\
&\leq \sum_{e = \bar{e} \in E_L^+} |(X_{\bar{e}} Q_{\bar{e}})(X_e Q_e) S| + \sum_{e \neq \bar{e} \in E_L^+} |(X_{\bar{e}} Q_{\bar{e}})(X_e Q_e) S| \\
&\leq 8N|\beta|(d-1)|v|^2.
\end{aligned}
$$

□

Getting better estimates on the Hessian will help obtain the Log-Sobolev inequality for a larger range of β. Note that we expect the corollaries of the Log-Sobolev ineqality (Poincare and consequently decay of correlations and so on) to be true $\forall \beta > 0$ given any N (in particular we care about small N, for which the range of admissible β is even smaller), so we are still a long way off!

References

[1] Hao Shen, Rongchan Zhu, and Xiangchan Zhu. "A Stochastic Analysis Approach to Lattice Yang–Mills at Strong Coupling". In: *Communications in Mathematical Physics* 400.2 (Dec. 2022), pp. 805–851. ISSN: 1432-0916. DOI: 10.1007/s00220-022-04609-1. URL: http://dx.doi.org/10.1007/s00220-022-04609-1.

[2] F. Wang. *Functional inequalities, Markov semigroups and spectral theory.* Elsevier, 2006.