A STOCHASTIC ANALYSIS APPROACH TO LATTICE YANG–MILLS AT STRONG COUPLING

HAO SHEN, RONGCHAN ZHU, AND XIANGCHAN ZHU

Abstract. We develop a new stochastic analysis approach to the lattice Yang–Mills model at strong coupling in any dimension $d > 1$, with t' Hooft scaling βN for the inverse coupling strength. We study their Langevin dynamics, ergodicity, functional inequalities, large N limits, and mass gap.

Assuming $|\beta| < \frac{N-2}{32(d-1)N}$ for the structure group $SO(N)$, or $|\beta| < \frac{1}{16(d-1)}$ for $SU(N)$, we prove the following results. The invariant measure for the corresponding Langevin dynamic is unique on the entire lattice, and the dynamic is exponentially ergodic under a Wasserstein distance. The finite volume Yang–Mills measures converge to this unique invariant measure in the infinite volume limit, for which Log-Sobolev and Poincaré inequalities hold. These functional inequalities imply that the suitably rescaled Wilson loops for the infinite volume measure has factorized correlations and converges in probability to deterministic limits in the large N limit, and correlations of a large class of observables decay exponentially, namely the infinite volume measure has a strictly positive mass gap. Our method improves earlier results or simplifies the proofs, and provides some new perspectives to the study of lattice Yang–Mills model.

Contents

Date: December 5, 2022.
2010 *Mathematics Subject Classification.* 37A25; 39B62; 60H10.

1. Introduction

The purpose of this paper is to apply stochastic analysis and ergodic theory for Markov processes to study the lattice Yang–Mills model with structure group $G \in \{SO(N), SU(N)\}$. In particular, we will consider the Langevin dynamics of these models, and under explicit strong coupling assumptions, we will prove uniqueness of invariant measures in infinite volume, log-Sobolev and Poincaré inequalities, with some application in large N limits of Wilson loops and exponential decay of correlations.

Lattice discretizations of the Yang–Mills theories were first proposed in the physics literature by Wilson [Wil74] which lead to well-defined Gibbs measures on collections of matrices. We refer to [Cha19b] for a nice review on the Yang–Mills model and its gauge invariant discretization as well as the fundamental questions for the model. Among the literature we only mention that approximate computations of the Wilson loop expectations as the size N of the structure group becomes large was first suggested by 't Hooft [tH74], where the Yang–Mills Hamiltonian is multiplied by βN (known as the 't Hooft scaling), which is closely related to our present article.

The problems we discuss in this paper have been of interest and studied for decades in mathematical physics. A closely related earlier paper is by Osterwalder–Seiler [OS78], which showed that for the lattice Yang–Mills theory, when the coupling is sufficiently strong, the cluster expansion (or high-temperature expansion in statistical mechanics language) for the expectation values of local observables (i.e. bounded functions of finitely many edge variables) is convergent, uniformly in volume. The proof of this convergent cluster expansion was sketched in [OS78] since it follows similarly as [GJS73] for $P(\phi)_2$ model (and also [Spe75]); in fact it is simpler than the $P(\phi)_2$ model in [GJS73] since the fields are bounded in lattice Yang–Mills theory. Moreover, as explained in [OS78], the existence of a mass gap (exponential clustering) follows from convergence of the cluster expansion, so do existence of the infinite volume limit and analyticity of Schwinger functions in the inverse coupling. Uniqueness of infinite volume limit should also follow from cluster expansion, see e.g. [AHKZ89] for the case of the $P(\phi)_2$ model. We also refer to the book [Sei82] for these expansion techniques and results. As for the large N limits, in the recent papers, factorization property of the Wilson loop expectations was proved in [Cha19a, Corollary 3.2] and [Jaf16] under the assumption that β is sufficiently small.

Given the earlier work, we revisit these problems in this article for a number of reasons. First of all, the earlier work [OS78] didn't consider 't Hooft scaling, but if we translate their results into 't Hooft scaling where the Hamiltonian is multiplied by βN then their condition amounts to requiring βN to be small. However, to our best knowledge, under the 't Hooft scaling βN uniqueness was not known for β in a fixed small neighborhood of the origin when N is arbitrarily large (see for instance the discussion after [Cha19a, Theorem 3.1]); this is the reason that [Cha19a] and [Jaf16] formulated their large N results on a sequence of N-dependent finite volumes. One aim of this paper is to establish uniqueness of infinite volume measures for β in a *fixed and explicit* small neighborhood of the origin which is *uniform* in N, which allows us to prove the existence of a mass gap and large N limits of Wilson loops directly in infinite volume for this range of β.

Secondly, as another motivation of this paper, we develop new methods based on stochastic analysis and give new proofs to these results. In these methods, the curvature properties of the Lie groups are better exploited via the verification of the Bakry–Émery condition. In particular, this allows us to perform more delicate calculations and obtain more explicit smallness condition on inverse coupling. As another novelty we study the Langevin dynamics (or stochastic quantization) and we prove uniqueness of the infinite volume measures by showing that the dynamic on the entire $\mathbb{Z}^d$ has a unique invariant measure. To this end we employed coupling methods for our stochastic dynamics, which is a variant of Kendall–Cranston's coupling. Such stochastic coupling arguments were used earlier in the stochastic analysis on manifolds, but to our best knowledge this appears to be the first time that such coupling arguments are used in the setting of statistical physics or lattice quantum field theory models with manifold target spaces. For our coupling arguments we will also need to introduce suitable weighted distances on the product manifolds, and in our calculations a subtle comparison between the weight parameter and the curvature plays a key role in order to obtain ergodicity.

As the third motivation, it appears to us that some of the proofs in this paper are simpler. For instance, the large N results on Wilson loops follow quickly from the Poincaré inequality, which simply comes from the Bakry–Émery condition. Our proof of exponential decay relies on some earlier ideas of Guionnet–Zegarlinski [GZ03] together with our explicit bounds on commutators between derivatives and Markov generators on Lie groups. This seems to be simpler than cluster expansion, or at least provides some new perspectives.

1.1. **Lattice Yang–Mills.** We first recall the basic setup and definitions of the model.

Let $\Lambda_L = \mathbb{Z}^d \cap L\mathbb{T}^d$ be a finite d dimensional lattice with side length L and unit lattice spacing, and we will consider various functions on it with periodic boundary conditions. We will sometimes write $\Lambda = \Lambda_L$ for short. We say that a lattice edge of $\mathbb{Z}^d$ is positively oriented if the beginning point is smaller in lexographic order than the ending point. Let E^+ (resp. E^-) be the set of positively (resp. negatively) oriented edges, and denote by $E^+_{\Lambda_L}$, $E^-_{\Lambda_L}$ the corresponding subsets of edges with both beginning and ending points in Λ_L. Define $E \stackrel{\text{def}}{=} E^+ \cup E^-$ and let $u(e)$ and $v(e)$ denote the starting point and ending point of an edge $e \in E$, respectively.

We write G for the Lie group $SO(N)$ or $SU(N)$ and $\mathfrak{g}$ for the associated Lie algebra $\mathfrak{so}(N)$ or $\mathfrak{su}(N)$. Note that we always view G as a real manifold (even for $SU(N)$), and $\mathfrak{g}$ as a real vector space, and we will write $d(\mathfrak{g}) = \dim_{\mathbb{R}} \mathfrak{g}$.

To define the lattice Yang–Mills theory we need more notation, for which we closely follow [Cha19a] and [SSZ22].

A *path* is defined to be a sequence of edges $e_1 e_2 \cdots e_n$ with $e_i \in E$ and $v(e_i) = u(e_{i+1})$ for $i = 1, 2, \cdots, n-1$. The path is called closed if $v(e_n) = u(e_1)$. A *plaquette* is a closed path of length four which traces out the boundary of a square. Also, let $\mathcal{P}_{\Lambda_L}$ be the set of plaquettes whose vertices are all in Λ_L, and $\mathcal{P}^+_{\Lambda_L}$ be the subset of plaquettes $p = e_1 e_2 e_3 e_4$ such that the beginning point of e_1 is lexicographically the smallest among all the vertices in p and the ending point of e_1 is the second smallest.

The lattice Yang-Mills theory (or lattice gauge theory) on Λ_L for the structure group G, with $\beta \in \mathbb{R}$ the inverse coupling constant, is the probability measure $\mu_{\Lambda_L,N,\beta}$ on the set of all collections $Q = (Q_e)_{e \in E^+_{\Lambda_L}}$ of G-matrices, defined as

$$\mathrm{d}\mu_{\Lambda_L,N,\beta}(Q) \stackrel{\text{def}}{=} Z^{-1}_{\Lambda_L,N,\beta} \exp\Big(\mathcal{S}(Q)\Big) \prod_{e \in E^+_{\Lambda_L}} \mathrm{d}\sigma_N(Q_e)\,, \tag{1.1}$$

with

$$\mathcal{S}(Q) \stackrel{\text{def}}{=} N\beta\text{Re} \sum_{p\in\mathcal{P}^+_{\Lambda_L}} \text{Tr}(Q_p), \tag{1.2}$$

where $Z_{\Lambda_L,N,\beta}$ is the normalizing constant, $Q_p \stackrel{\text{def}}{=} Q_{e_1}Q_{e_2}Q_{e_3}Q_{e_4}$ for a plaquette $p = e_1e_2e_3e_4$, and σ_N is the Haar measure on G. Note that for $p \in \mathcal{P}^+_{\Lambda_L}$ the edges e_3 and e_4 are negatively oriented, so throughout the paper we define $Q_e \stackrel{\text{def}}{=} Q^{-1}_{e^{-1}}$ for $e \in E^-$, where e^{-1} denotes the edge with orientation reversed. Also, Re is the real part, which can be omitted when $G = SO(N)$.

1.2. **Main Results.** We will assume the following in our main results on lattice Yang–Mills.

Assumption 1.1. *Suppose that*

$$K_{\mathcal{S}} \stackrel{\text{def}}{=} \begin{cases} \dfrac{N+2}{4} - 1 - 8N|\beta|(d-1) > 0, & G = SO(N)\ , \\ \dfrac{N+2}{2} - 1 - 8N|\beta|(d-1) > 0\ , & G = SU(N)\ . \end{cases}$$

Assumption 1.1 is equivalent to the following *strong coupling* assumption:

$$|\beta| < \begin{cases} \dfrac{1}{32(d-1)} - \dfrac{1}{16N(d-1)}, & G = SO(N)\ , \\ \dfrac{1}{16(d-1)}, & G = SU(N)\ . \end{cases} \tag{1.3}$$

Define the (product) topological space $\mathcal{Q} \stackrel{\text{def}}{=} G^{E^+}$, which will serve as our infinite volume configuration space. By Tychonoff's theorem $\mathcal{Q}$ is compact. For each $a > 1$ we define the distance $\rho_{\infty,a}$ on $\mathcal{Q}$ by

$$\rho^2_{\infty,a}(Q,Q') \stackrel{\text{def}}{=} \sum_{e\in E^+} \frac{1}{a^{|e|}}\rho^2(Q_e,Q'_e), \tag{1.4}$$

with $|e|$ being the distance from 0 to e in $\mathbb{Z}^d$. Here $\rho(\cdot,\cdot)$ is the Riemannian distance on G. The distances for different choices of a give equivalent topologies, and we just write ρ_∞ when there is no confusion. $\mathcal{Q}$ is then a Polish space w.r.t. ρ_∞. By standard results in topology, the topology induced by ρ_∞ is equivalent with the product topology on $\mathcal{Q}$.

We can easily extend the measure $\mu_{\Lambda_L,N,\beta}$ to the infinite volume configuration space $\mathcal{Q}$ by periodic extension, which is still denoted as $\mu_{\Lambda_L,N,\beta}$. Namely, we can construct a random variable with law given by $\mu_{\Lambda_L,N,\beta}$ and extend the random variable periodically, and the law of the periodic extension gives the desired extension of measure. Since G and $\mathcal{Q}$ are compact, $\{\mu_{\Lambda_L,N,\beta}\}_{L\geqslant 1}$ form a tight set.

We will consider the Langevin dynamic on $\mathcal{Q}$, formally given by

$$\mathrm{d}Q = \nabla\mathcal{S}(Q)\mathrm{d}t + \sqrt{2}\mathrm{d}\mathfrak{B}\ , \tag{1.5}$$

with $\mathfrak{B} = (\mathfrak{B}_e)_{e\in E^+}$ being independent Brownian motions on G. This is formal since we will need to "extend" $\nabla\mathcal{S}$ to infinite volume in a suitable sense. More precisely, the Langevin dynamic we consider is the following SDE system parametrized by $e \in E^+$:

$$\mathrm{d}Q_e = -\frac{1}{2}N\beta \sum_{p\in\mathcal{P}, p\succ e} (Q_p - Q_p^*)Q_e\mathrm{d}t - \frac{1}{2}(N-1)Q_e\mathrm{d}t + \sqrt{2}\mathrm{d}B_eQ_e, \qquad \text{if} \quad G = SO(N),$$

$$\mathrm{d}Q_e = -\frac{1}{2}N\beta \sum_{p\in\mathcal{P}, p\succ e} \Big((Q_p - Q_p^*) - \frac{1}{N}\mathrm{Tr}(Q_p - Q_p^*)I_N\Big)Q_e\mathrm{d}t \tag{1.6}$$

$$- \frac{N^2-1}{N}Q_e\mathrm{d}t + \sqrt{2}\mathrm{d}B_eQ_e, \qquad \text{if} \quad G = SU(N).$$

Here $B = (B_e)_{e\in E}$ is a collection of independent Brownian motions on the Lie algebra $\mathfrak{g}$ of G, and the terms linear in Q_e arise from Casimir elements of the Lie algebras; we will review these in Section 2.

Remark 1.1. *We note that the above SDE (in finite volume) was used earlier in [SSZ22] to derive the loop equations (i.e. Dyson–Schwinger or Makeenko–Migdal equations) for Wilson loops of the model (1.1). These loop equations also hold for any infinite volume tight limit of the measures, and in particular for the unique invariant measure for β satisfying (1.3) as given in Theorem 1.2.*

The study of a quantum field theory of the form (1.1) via a dynamic (1.5) is also called stochastic quantization as first proposed by [Nel66, PW81].

We will prove that there exists a unique probabilistically strong solutions to SDE (1.6) starting from any initial data in $\mathcal{Q}$ in Proposition 3.4. Hence the solutions form a Markov process in $\mathcal{Q}$ and the related semigroup is denoted by $(P_t)_{t\geqslant 0}$.

Our first main result is as follows.

Theorem 1.2 (Uniqueness and ergodicity)**.** *Under Assumption 1.1, the following statements hold.*

(1) The invariant measure of the Markov semigroup $(P_t)_{t\geqslant 0}$ for the Langevin dynamic (1.6) is unique. We denote this invariant measure by $\mu^{\mathrm{YM}}_{N,\beta}$.

(2) Furthermore, every tight limit of $\{\mu_{\Lambda_L,N,\beta}\}_L$ is the same, and the whole sequence $\{\mu_{\Lambda_L,N,\beta}\}_L$ converges to $\mu^{\mathrm{YM}}_{N,\beta}$ as $L\to\infty$.

(3) Finally, the Markov semigroup $(P_t)_{t\geqslant 0}$ is exponentially ergodic in the following sense: there exists a constant $a > 1$ such that for any $\nu \in \mathscr{P}(\mathcal{Q})$

$$W_2^{\rho_{\infty,a}}(\nu P_t, \mu^{\mathrm{YM}}_{N,\beta}) \leqslant C(a)e^{-\widetilde{K}_{\mathcal{S}}t}, \quad t\geqslant 0, \tag{1.7}$$

for some $\widetilde{K}_{\mathcal{S}} > 0$ which only depends on the constant a, d, β and G (in particular N).

Here $W_2^{\rho_{\infty,a}}$ is the Wasserstein distance w.r.t. $\rho_{\infty,a}$ given for any $\mu,\nu \in \mathscr{P}(\mathcal{Q})$

$$W_2^{\rho_{\infty,a}}(\mu,\nu) \stackrel{\mathrm{def}}{=} \inf_{\pi\in\mathscr{C}(\mu,\nu)} \pi(\rho^2_{\infty,a})^{1/2},$$

with $\mathscr{C}(\mu,\nu)$ being the set of couplings between μ and ν. Remark that $\widetilde{K}_{\mathcal{S}}$ can be explicitly given by (5.15) below and gives a lower bound of spectral gap for $(P_t)_{t\geqslant 0}$ in Wasserstein distance. In Theorem 1.4 we will see that $K_{\mathcal{S}}$ gives a lower bound of spectral gap in $L^2(\mu^{\mathrm{YM}}_{N,\beta})$.

Remark 1.3. *The periodic boundary condition in the definition of $\{\mu_{\Lambda_L,N,\beta}\}_L$ is not essential. By the same argument as in Theorem 3.5 the tight limit of $\{\mu_{\Lambda_L,N,\beta}\}_L$ when changing the periodic boundary condition to Dirichlet or other boundary conditions is also the invariant measure of the SDE (1.6), hence, is the same as $\mu^{\mathrm{YM}}_{N,\beta}$.*

We remark that uniqueness for small β could possibly also be proven using the method of Dobrushin, see e.g. [Dob70]. To this end one would also need to consider the related Wasserstein metric with respect to the Riemannian distance similarly as we do in this paper. However as we understand such an argument has not been carried out in detail for lattice Yang Mills in the literature. Here we give a proof based on a new idea which is a variant of Kendall–Cranston's coupling used earlier in the stochastic analysis on manifold.

The idea for the proof of Theorem 1.2 is to use finite dimensional approximation, for which we construct a suitable coupling and find a suitable distance such that the associated Wasserstein distance between the two finite dimensional approximations starting from different initial distributions decays exponentially fast in time with *uniform* speed.

We define the cylinder functions $C^\infty_{cyl}(\mathcal{Q})$ by

$$C^\infty_{cyl}(\mathcal{Q})=\Big\{F:F=f(Q_{e_1},\dots,Q_{e_n}),n\in\mathbb{N},e_i\in E^+,f\in C^\infty(G^n)\Big\}. \tag{1.8}$$

We then obtain the following log-Sobolev inequality for $\mu^{\text{YM}}_{N,\beta}$ based on Bakry–Émery's criterion.

Theorem 1.4 (Log-Sobolev inequality)**.** *Under Assumption 1.1, the log-Sobolev inequality holds for the measure $\mu^{\text{YM}}_{N,\beta}$ in Theorem 1.2. Namely, for all cylinder functions $F\in C^\infty_{cyl}(\mathcal{Q})$ with $\mu^{\text{YM}}_{N,\beta}(F^2)=1$,*

$$\mu^{\text{YM}}_{N,\beta}(F^2\log F^2)\leqslant\frac{2}{K_{\mathcal{S}}}\sum_{e\in E^+}\mu^{\text{YM}}_{N,\beta}(|\nabla_e F|^2). \tag{1.9}$$

This implies the Poincaré inequality, i.e. for all cylinder functions $F\in C^\infty_{cyl}(\mathcal{Q})$,

$$\mu^{\text{YM}}_{N,\beta}(F^2)\leqslant\frac{1}{K_{\mathcal{S}}}\sum_{e\in E^+}\mu^{\text{YM}}_{N,\beta}(|\nabla_e F|^2)+\mu^{\text{YM}}_{N,\beta}(F)^2, \tag{1.10}$$

with ∇_e the gradient w.r.t. the variable Q_e.

Theorem 1.4 follows from Theorem 1.2 and Corollary 4.5, which states the log-Sobolev inequality for every tight limit of $(\mu_{\Lambda_L,N,\beta})_{L\geqslant1}$. The RHS of (1.9) is the Dirichlet form associated with the Langevin dynamic (see Proposition 3.7). Hence, (1.9) holds for the functions in the domain of Dirichlet form by lower-semicontinuity.

As some simple applications of the Poincaré inequality, we show certain "susceptibility" bounds on the field Q_e and $\mathrm{Tr}(Q_p)$, see Corollaries 4.7 and 4.8. These examples demonstrate how to choose suitable functions in these functional inequalities to yield interesting bounds for the model.

Log-Sobolev and Poincaré inequalities in Theorem 1.4 follow by checking the Bakry–Émery criteria [BE85, BGL14] directly for the finite dimensional approximation on the product manifolds. As the Ricci curvatures of the target manifolds G are given by positive constants, so are the Ricci curvatures of the configuration space (i.e. the product manifolds). The Hessian of the Hamiltonian could also be bounded by the Ricci curvatures in the strong coupling regimes.

As a corollary of Theorem 1.4 we obtain the following large N properties of the Wilson loops. For the rest of this paper, by a *loop* we mean an equivalent class of closed paths (as defined in Section 1.1), where the equivalence relation $\sim$ is given by cyclic permutations $e_1e_2\cdots e_n\sim e_ie_{i+1}\cdots e_ne_1e_2\cdots e_{i-1}$ for any $i\in\{1,...,n\}$, and it has no two successive edges of the form $e^{-1}e$. We will always assume that a loop is non-empty, i.e. has positive number of edges. Given a loop $\ell = e_1 e_2 \cdots e_n$, recall that the Wilson loop variable W_ℓ is defined as

$$W_\ell \stackrel{\text{def}}{=} \mathrm{Tr}(Q_{e_1} Q_{e_2} \cdots Q_{e_n}) \,. \tag{1.11}$$

Corollary 1.5 (Large N limit of Wilson loops)**.** *Under Assumption 1.1, for every Wilson loop (1.11), writing* Var *and* $\mathbf{E}$ *for the variance and expectation under the measure* $\mu^{\text{YM}}_{N,\beta}$ *in Theorem 1.2, one has*

$$\mathrm{Var}\Big(\frac{1}{N}W_\ell\Big) \leqslant \frac{n(n-3)}{K_{\mathcal{S}}N}, \quad G = SO(N); \quad \mathrm{Var}\Big(\frac{1}{N}W_\ell\Big) \leqslant \frac{4n(n-3)}{K_{\mathcal{S}}N}, \quad G = SU(N)\,. \tag{1.12}$$

In particular, we obtain the convergence

$$\Big|\frac{W_\ell}{N} - \mathbf{E}\frac{W_\ell}{N}\Big| \to 0 \qquad \text{as } N \to \infty \tag{1.13}$$

in probability, and the factorization property of Wilson loops, i.e. for any loops $\ell_1, \dots, \ell_m$

$$\lim_{N\to\infty} \Big|\mathbf{E}\frac{W_{\ell_1}\dots W_{\ell_m}}{N^m} - \prod_{i=1}^m \mathbf{E}\frac{W_{\ell_i}}{N}\Big| = 0.$$

Corollary 1.5 is proven in Section 4. Our proof is novel which is based on the Poincaré inequality (1.10). Note that our formulation of the result is different from [Cha19a] and [Jaf16] in which the factorization property of Wilson loops was obtained by taking a sequence of increasing finite lattices $\mathbb{Z}^d = \cup_{N=1}^\infty \Lambda_N$, considering the correlations of Wilson loops over Λ_N, and taking infinite volume limit simultaneously as the large N limit when sending $N \to \infty$. In our approach, we work directly in infinite volume, which seems more natural. The subtlety here, as mentioned above and also explained in [Cha19a], is that the 't Hooft coupling places $N\beta$ instead of β in front of the Hamiltonian so one would require $N\beta$ to be sufficiently small to obtain the infinite volume limit, which would appear to be problematic when taking the large N limit afterwards. However, thanks to our precise smallness condition on β in (1.3), we can take β small uniformly in N. This also allows us to derive bounds on the variances of Wilson loops which are explicit in terms of N. Our proof based on the Poincaré inequality which follows from the Bakry–Émery condition also appears to be simpler than the arguments in aforementioned previous work.

Furthermore, we obtain the following exponential decay property of the covariance. Consider $f \in C^\infty_{cyl}(\mathcal{Q})$ and we write Λ_f for the set of the edges f depends on. Let $|\Lambda_f|$ denote the cardinality of Λ_f. We define

$$|\!|\!| f |\!|\!|_\infty \stackrel{\text{def}}{=} \sum_{e\in\Lambda_f} \|\nabla_e f\|_{L^\infty},$$

where $C^\infty_{cyl}(\mathcal{Q})$ is introduced in (1.8) and ∇_e is introduced in Section 3. We also write $d(A,B)$ for the distance between $A, B \subset E^+$, which is given by the nearest distance between the vertices in A and B.

Corollary 1.6 (Mass gap)**.** *Suppose that Assumption 1.1 holds. Writing* Cov *for the covariance under the measure* $\mu^{\text{YM}}_{N,\beta}$ *in Theorem 1.2. For* $f, g \in C^\infty_{cyl}(\mathcal{Q})$*, suppose that* $\Lambda_f \cap \Lambda_g = \varnothing$*. It holds that*

$$|\mathrm{Cov}(f,g)| \leqslant c_1 d(\mathfrak{g}) e^{-c_N d(\Lambda_f,\Lambda_g)} \Big(|\!|\!| f |\!|\!|_\infty |\!|\!| g |\!|\!|_\infty + \|f\|_{L^2(\mu^{\text{YM}}_{N,\beta})} \|g\|_{L^2(\mu^{\text{YM}}_{N,\beta})} \Big)\,,$$

where c_1 *depends on* $|\Lambda_f|$, $|\Lambda_g|$ *and* c_N *depends on* $K_{\mathcal{S}}$, N *and* d.

Note that f and g in the above corollary can be chosen to be Wilson loops, or functions of an arbitrary number of Wilson loops, which are of particular interest in physics.

We also remark that exponential decay of correlations is also related to Wilson's area law for Wilson loops – see [Cha21, Theorem 2.4] in which it is proved that exponential decay of correlations is a sufficient condition for "unbroken center symmetry", which implies confinement (slightly weaker than Wilson's area law).

The proof of Corollary 1.6 is given in Section 4.

We conclude this subsection by some brief comments on the challenges or subtleties in the proofs of the above results. One of the important ingredients in the proofs is to estimate the Hessian or the general second order derivatives for the interaction $\mathcal{S}$ defined in (1.2). Note that a term of the form $N\mathrm{Tr}(Q_p)$ in the interaction $\mathcal{S}$ would "appear" to be of order N^2, which would be too large for us to obtain the desired results. In our proofs we will properly arrange terms and apply certain properties of the Lie groups and we will show that the relevant second order derivatives are actually at most of order $|\beta|N$. See the explanations before Lemma 4.1 and the proof of Theorem 1.2 in Section 5 for more details. This is one of the crucial reasons which allow us to compare the Ricci curvatures of the Lie groups and Hessians to verify the Bakry–Émery condition by choosing β small, and also prove ergodicity using a suitable weighted distance.

1.3. **Relevant literature and possible directions.** The study of properties of lattice gauge theories recently attracts much interest. Besides the aforementioned work by [Cha19a] and [Jaf16], [Cha16] computed the leading terms of free energies, [BG18] provided an elaborate description of loop expectations in the planar setting, and [CJ16] derived $1/N$ expansions in the $SO(N)$ case at strong coupling. Wilson loops (and also Wilson lines when coupled with Higgs) for gauge theories with finite structure groups were studied in [Cha20, FLV22, FLV21, Cao20, For21, Adh21]; see also [GS21] for the $U(1)$ case. Moreover, exponential correlation decay for lattice gauge theories with finite abelian structure groups was obtained by [For22] using coupling argument, and for finite non-abelian structure groups this was proved by [AC22] at weak coupling.

Our present article provides a new approach to study these models via stochastic analysis and dynamical perspective; see also [SSZ22] for a new derivation of loop i.e. Makeenko–Migdal equations for Wilson loops by such methods.

Remark 1.7. *Here by "stochastic analysis approach", we do not mean the stochastic analysis approach for 2D Yang–Mills in continuum developed earlier by* [GKS89] *and* [Dri89] *(See Def. 3.3 therein) in which parallel translations (which are related with Wilson loops) are formulated as stochastic differential equations. See* [Dri19] *and references therein for more recent literature in this direction. Our Yang–Mills SDE on the other hand is the stochastic dynamic for the connection fields on a lattice with fixed spacing, which is along the line of stochastic quantization.*

We remark that the choice of constant positive curvature Lie groups $SO(N)$ and $SU(N)$ in this article is a technical simplification for demonstrating our method, and it should apply as well for other compact target spaces with constant or non-constant positive curvatures. For instance it should apply to a lattice $SO(N)$ Yang–Mills model coupled with a Higgs field Φ which takes values in a sphere in $\mathbb{R}^N$ (i.e. rotator model) via a gauge-covariant derivative term, whose action takes the form $\mathrm{Re}\sum_p \mathrm{Tr}(Q_p) + \sum_e |Q_e\Phi_{v(e)} - \Phi_{u(e)}|^2$.

It would certainly be interesting to show if log-Sobolev inequalities still hold when the lattice spacing vanishes, in the situations where the continuum limits of these models are shown or expected to exist. In this direction, on the two dimensional torus, the continuum limit of lattice approximations of the Yang–Mills measures on 1-forms was recently obtained by Chevyrev [Che19], who also showed that a certain class of Wilson loop observables of this random 1-form coincide in law with the corresponding observables under the Yang–Mills measure in the sense of [Lév03]. Note that the Langevin dynamics for Yang–Mills models on the two and three dimensional continuous torus were recently constructed in [CCHS22a, CCHS22b] (see [Che22] for a review of these results), and as mentioned in [CCHS22a] it would be interesting to show that the lattice dynamics of the type (1.5) converge to the processes constructed in the above papers in two and three dimensions. For some of the recent progress along this direction, see the proof of log-Sobolev inequalities for the $\Phi^4_{2,3}$ and sine-Gordon models [BD22, BB21], and the 1D nonlinear σ-model (see [AD99, Hai16, BGHZ21]) for which the log-Sobolev inequality, ergodicity and non-ergodicity (depending on the curvature of the target manifolds) were obtained in [RWZZ20, CWZZ21]. It would also be interesting to see if the methods developed in this paper can be applied to weak coupling i.e. large β regime (when the structure group G is finite, see the recent progress [AC22] on mass gap in weak coupling regime).

Notation. Given a Polish space E, we write $C([0,T];E)$ for the space of continuous functions from $[0,T]$ to E. We use $\mathscr{P}(E)$ to denote all the probability measures on E with Borel σ-algebra.

Acknowledgments. We would like to thank Scott Smith, Feng-Yu Wang, Xicheng Zhang and Xin Chen for very helpful discussions. H.S. gratefully acknowledges financial support from NSF grants DMS-1954091 and CAREER DMS-2044415. R.Z. and X.Z. are grateful to the financial supports by National Key R&D Program of China (No. 2022YFA1006300). R.Z. gratefully acknowledges financial support from the NSFC (No. 11922103, 12271030). X.Z. is grateful to the financial supports by National Key R&D Program of China (No. 2020YFA0712700) and the NSFC (No. 12090014, 12288201) and the support by key Lab of Random Complex Structures and Data Science, Youth Innovation Promotion Association (2020003), Chinese Academy of Science. This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID 317210226–SFB 1283. Rongchan Zhu is the corresponding author.

2. Notation and preliminaries

In this section we collect some notation and standard facts about Riemannian geometry, Lie groups and Brownian motions.

Riemannian manifolds. Let M be a Riemannian manifold of dimension d. We denote by $C^\infty(M)$ the space of real-valued smooth functions on M. For $x \in M$ we denote by T_xM the tangent space at x with inner product $\langle \cdot, \cdot \rangle_{T_xM}$. For $X \in T_xM$, we write Xf or $X(f)$ for the differentiation of f along X at x. For a smooth curve $\gamma : [\alpha, \beta] \to M$ the tangent vector along γ is defined by

$$\dot\gamma_t f = \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma_t), \quad f \in C^\infty(M).$$

Let ∇ be the Levi-Civita connection, which is a bilinear operation associating to vector fields X and Y a vector field $\nabla_Y X$. Recall that $(\nabla_Y X)(x)$ depends on Y only via $Y(x)$ for $x \in M$ (e.g. [dC92, Remark 2.3]).

For $f \in C^\infty(M)$, we denote by ∇f the gradient vector field of f. We also write $\mathrm{Hess}(f)$ for the Hessian. It can be calculated in the following ways

$$\mathrm{Hess}_f(X,Y) \stackrel{\text{def}}{=} \mathrm{Hess}(f)(X,Y) = \langle \nabla_X \nabla f, Y\rangle = X(Yf) - (\nabla_X Y)f\ . \tag{2.1}$$

It is a two-tensor: $\mathrm{Hess}_f(\varphi X,Y) = \mathrm{Hess}_f(X,\varphi Y) = \varphi \mathrm{Hess}_f(X,Y)$ for any $\varphi \in C^\infty(M)$ so $\mathrm{Hess}_f(X,Y)(x)$ depends only on $X(x)$ and $Y(x)$. Since Levi-Civita connection is torsion-free, $\mathrm{Hess}(f)$ is symmetric in X,Y.

The Riemann curvature tensor $\mathscr{R}(\cdot,\cdot)$ associated to vector fields X,Y is an operator defined by

$$\mathscr{R}(X,Y)Z = \nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) - \nabla_{[X,Y]}Z.$$

Let $\{W_i\}_{i=1}^d$ be an orthonormal basis of T_xM. The Ricci curvature tensor is defined by

$$\mathrm{Ric}(X,Y) = \sum_{i=1}^d \langle \mathscr{R}(X,W_i)W_i, Y\rangle_{T_xM} \tag{2.2}$$

and is independent of the choice of $\{W_i\}$. Note that $\mathrm{Ric}(X,Y)(x)$ depends on X,Y only via $X(x),Y(x)$ for $x \in M$.

Let γ be a geodesic. A smooth vector field J is called a Jacobi field along $\gamma : [0,t] \to M$ if $\nabla_{\dot\gamma}\nabla_{\dot\gamma}J + \mathscr{R}(J,\dot\gamma)\dot\gamma = 0$. For any $X \in T_{\gamma_0}M$ and $Y \in T_{\gamma_t}M$, there exists a Jacobi field J along γ satisfying $J_0 = X$ and $J_t = Y$ (c.f. [CE75, Section 1.5], [Wan06, Section 0.4]).

Lie groups and algebras. For any matrix M we write M^* for the conjugate transpose of M. Let $M_N(\mathbb{R})$ and $M_N(\mathbb{C})$ be the space of real and complex $N \times N$ matrices.

For Lie groups $SO(N)$, $SU(N)$, we write the corresponding Lie algebras as $\mathfrak{so}(N)$, $\mathfrak{su}(N)$ respectively. Every matrix Q in one of these Lie groups satisfies $QQ^* = I_N$, and every matrix X in one of these Lie algebras satisfies $X + X^* = 0$. Here I_N denotes the identity matrix.

We endow $M_N(\mathbb{C})$ with the Hilbert-Schmidt inner product

$$\langle X,Y\rangle = \mathrm{Re}\mathrm{Tr}(XY^*) \qquad \forall X,Y \in M_N(\mathbb{C}). \tag{2.3}$$

We restrict this inner product to our Lie algebra $\mathfrak{g}$, which is then invariant under the adjoint action. In particular for $X,Y \in \mathfrak{so}(N)$ or $\mathfrak{su}(N)$ we have $\langle X,Y\rangle = -\mathrm{Tr}(XY)$. Note that $\mathrm{Tr}(XY) \in \mathbb{R}$ since we have $\mathrm{Tr}((XY)^*) = \mathrm{Tr}(Y^*X^*) = \mathrm{Tr}(XY)$, and $\mathrm{Tr}(A^*) = \overline{\mathrm{Tr}(A)}$ for any $A \in M_N(\mathbb{C})$.

Below G is always understood as $SO(N)$ or $SU(N)$. Every $X \in \mathfrak{g}$ induces a right-invariant vector field $\widetilde{X}$ on G, and for each $Q \in G$, $\widetilde{X}(Q)$ is just given by XQ since G is a matrix Lie group. Indeed, given any $X \in \mathfrak{g}$, the curve $t \mapsto e^{tX}Q$ is well approximated near $t = 0$ by $Q + tXQ$ up to an error of order t^2.

The inner product on $\mathfrak{g}$ induces an inner product on the tangent space at every $Q \in G$ via the right multiplication on G. Hence, for $X,Y \in \mathfrak{g}$, we have $XQ, YQ \in T_QG$, and their inner product is given by $\mathrm{Tr}((XQ)(YQ)^*) = \mathrm{Tr}(XY^*)$. This yields a bi-invariant Riemannian metric on G.

For any function $f \in C^\infty(G)$ and $X \in \mathfrak{g}$, the right-invariant vector field $\widetilde{X}$ induced by X acts on f at $Q \in G$ by the right-invariant derivative

$$\widetilde{X}f(Q) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(e^{tX}Q). \tag{2.4}$$

We have

$$\widetilde{[X,Y]} = [\widetilde{X},\widetilde{Y}], \qquad \text{namely,} \qquad ([X,Y]Q)f(Q) = [XQ,YQ]f(Q),$$

where the $[\cdot,\cdot]$ is the Lie bracket on $\mathfrak{g}$ on the LHS and the vector fields commutator on the RHS. Also, for the Levi-Civita connection ∇ we have

$$\nabla_{\widetilde{X}}(\widetilde{Y}) = \frac{1}{2}\widetilde{[X,Y]}. \tag{2.5}$$

We refer the above facts to [AGZ10, Appendix F], e.g. Lemma F.27 therein.

Brownian motions. Denote by $\mathfrak{B}$ and B the Brownian motions on a Lie group G and its Lie algebra $\mathfrak{g}$ respectively. The Brownian motion B is characterized by

$$\mathbf{E}\Big[\langle B(s),X\rangle\langle B(t),Y\rangle\Big] = \min(s,t)\langle X,Y\rangle \qquad \forall X,Y\in\mathfrak{g}. \tag{2.6}$$

By [LÍ7, Sec. 1.4], the Brownian motions $\mathfrak{B}$ and B are related through the following SDE:

$$\mathrm{d}\mathfrak{B} = \mathrm{d}B\circ\mathfrak{B} = \mathrm{d}B\,\mathfrak{B} + \frac{c_{\mathfrak{g}}}{2}\mathfrak{B}\mathrm{d}t, \tag{2.7}$$

where $\circ$ is the Stratonovich product, and $\mathrm{d}B\,\mathfrak{B}$ is in the Itô sense. Here the constant $c_{\mathfrak{g}}$ is determined by $\sum_\alpha v_\alpha^2 = c_{\mathfrak{g}} I_N$ where $(v_\alpha)_{\alpha=1}^{d(\mathfrak{g})}$ is an orthonormal basis of $\mathfrak{g}$. Moreover, by [LÍ7, Lem. 1.2], [1]

$$c_{\mathfrak{so}(N)} = -\frac{1}{2}(N-1), \quad c_{\mathfrak{su}(N)} = -\frac{N^2-1}{N}. \tag{2.8}$$

Denote by δ the Kronecker function, i.e. $\delta_{ij}=1$ if $i=j$ and 0 otherwise. For any matrix M, we write M^{ij} for its (i,j)th entry. The following holds by straightforward calculations (see e.g. [SSZ22, (2.5)]):

$$\mathrm{d}\langle B^{ij},B^{k\ell}\rangle = \frac{1}{2}(\delta_{ik}\delta_{j\ell}-\delta_{i\ell}\delta_{jk})\mathrm{d}t, \qquad \mathfrak{g}=\mathfrak{so}(N); \tag{2.9a}$$

$$\mathrm{d}\langle B^{ij},B^{k\ell}\rangle = \big(-\delta_{i\ell}\delta_{jk}+\frac{1}{N}\delta_{ij}\delta_{k\ell}\big)\,\mathrm{d}t\ , \qquad \mathfrak{g}=\mathfrak{su}(N). \tag{2.9b}$$

2.1. **Product manifolds and Lie groups.** For Riemannian manifolds M_1, M_2, the tangent space of the product manifolds $T_{(x_1,x_2)}(M_1\times M_2)$ is isomorphic with $T_{x_1}M_1\oplus T_{x_2}M_2$ which is endowed with the inner product

$$\langle u_1+u_2, v_1+v_2\rangle_{T_{(x_1,x_2)}(M_1\times M_2)} = \langle u_1,v_1\rangle_{T_{x_1}M_1} + \langle u_2,v_2\rangle_{T_{x_2}M_2}.$$

For a finite collection of Riemannian manifolds $(M_e)_{e\in A}$ where A is some finite set, the product is defined analogously.

If all M_e are the same manifold M, the product is written as M^A. In this case, given a point $x=(x_e)_{e\in A}\in M^A$, if $u_e\in T_{x_e}M_e$ for some $x_e\in M_e$, we will sometimes view u_e as a tangent vector in T_xM^A which has zero components for all $\bar e\neq e$. Continuing with this notation, if $(v_e^i)^{i=1,\dots,d}$ is a basis (resp. orthonormal basis) of $T_{x_e}M_e$, then $(v_e^i)_{e\in A}^{i=1,\dots,d}$ is a basis (resp. orthonormal basis) of T_xM^A.

For Lie groups G_1, G_2, the group multiplication is defined on $G_1\times G_2$ componentwise. The Lie algebra $\mathfrak{g}$ of $G_1\times G_2$ is isomorphic to $\mathfrak{g}_1\oplus\mathfrak{g}_2$ where $\mathfrak{g}_i$ is the Lie algebra of G_i. The Lie bracket on $\mathfrak{g}_1\oplus\mathfrak{g}_2$ is defined componentwise. If $X=(X_1,X_2)\in\mathfrak{g}$, then induced

[1]Note that in [LÍ7, Lem. 1.2], the scalar product differs from (2.3) by a scalar multiple depending on N and $\mathfrak{g}$, so we accounted for this in the expression for $c_{\mathfrak{g}}$ above.

the right-invariant vector field $\widetilde{X}(x)$ for every $x \in G_1 \times G_2$ is equal to $(\widetilde{X}_1(x), \widetilde{X}_2(x))$. In particular, (2.5) still holds for any two right-invariant vector fields on the Lie group product.

With similar notation as above we can define product G^A and its Lie algebra $\mathfrak{g}^A$ for a finite set A. Given $X \in \mathfrak{g}^A$, the exponential map $t \mapsto \exp(tX)$ is also defined pointwise as $\exp(tX)_e \stackrel{\text{def}}{=} e^{tX_e}$ for each $e \in A$.

In the following we choose G to be one of the matrix Lie groups as before. Define the configuration space as the Lie group product $\mathcal{Q}_L = G^{E^+_{\Lambda_L}}$, consisting of all maps $Q : e \in E^+_{\Lambda_L} \mapsto Q_e \in G$. Let $\mathfrak{q}_L = \mathfrak{g}^{E^+_{\Lambda_L}}$ be the corresponding direct sum of $\mathfrak{g}$. Note that $\mathfrak{q}_L$ is the Lie algebra of the Lie group $\mathcal{Q}_L$. For any matrix-valued functions A, B on $E^+_{\Lambda_L}$, we denote by AB the pointwise product $(A_e B_e)_{e \in E^+_{\Lambda_L}}$.

As above, the tangent space at $Q \in \mathcal{Q}_L$ consists of the products $XQ = (X_e Q_e)_{e \in E^+_{\Lambda_L}}$ with $X \in \mathfrak{q}_L$, and given two such elements XQ and YQ, their inner product is defined by

$$\langle XQ, YQ \rangle_{T_Q \mathcal{Q}_L} = \sum_{e \in E^+_{\Lambda_L}} \mathrm{Tr}(X_e Y_e^*).$$

The basis of the tangent space $T_Q \mathcal{Q}_L$ is given by $\{X^i_e Q : e \in E^+_{\Lambda_L}, 1 \leqslant i \leqslant d(\mathfrak{g})\}$ where for each e, $\{X^i_e\}_i$ is a basis for $\mathfrak{g}$.

Given any function $f \in C^\infty(\mathcal{Q}_L)$, the right-invariant derivative is given by $\frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} f(\exp(tX)Q)$. For each $Q \in \mathcal{Q}_L$, the gradient $\nabla f(Q)$ is an element of the tangent space at Q which satisfies for each $X \in \mathfrak{q}_L$

$$\langle \nabla f(Q), XQ \rangle_{T_Q \mathcal{Q}_L} = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(\exp(tX)Q) = (XQ)f. \tag{2.10}$$

We can write $\nabla f = \sum_{i=1}^{d(\mathfrak{g})} \sum_{e \in E^+_{\Lambda_L}} (v^i_e f) v^i_e$ with $\{v^i_e : e \in E^+_{\Lambda_L}, i = 1, \cdots, d(\mathfrak{g})\}$ being an orthonormal basis of $T_Q \mathcal{Q}_L$. We then define

$$\nabla_e f \stackrel{\text{def}}{=} \sum_{i=1}^{d(\mathfrak{g})} (v^i_e f) v^i_e, \qquad \Delta_e f \stackrel{\text{def}}{=} \mathrm{div} \nabla_e f = \sum_{i=1}^{d(\mathfrak{g})} \langle \nabla_{v^i_e} \nabla_e f, v^i_e \rangle.$$

Here ∇_e and Δ_e can be viewed as the gradient and the Laplace–Beltrami operator (w.r.t. the variable Q_e) on G endowed with the metric given above.

3. Yang–Mills SDE

In this section we first recall the Langevin dynamics (1.6) associated to the lattice Yang–Mills model in finite volume from [SSZ22]. We then extend the dynamics from finite volume to the whole $\mathbb{Z}^d$ and prove the global well-posdness of the SDE (1.6). Furthermore, we prove that every tight limit of $(\mu_{\Lambda_L,N,\beta})_L$ is an invariant measure for the SDE (1.6).

We consider the Langevin dynamic for the measure (1.1), which is the following SDE on $\mathcal{Q}_L$

$$\mathrm{d}Q = \nabla \mathcal{S}(Q)\mathrm{d}t + \sqrt{2}\mathrm{d}\mathfrak{B}\ , \tag{3.1}$$

with $\mathfrak{B} = (\mathfrak{B}_e)$ being independent Brownian motions on G. Here $\mathrm{d}\mathfrak{B}$ can be viewed as the white noise w.r.t. the inner product on $T_Q \mathcal{Q}_L$.

We now recall the explicit expression for $\nabla \mathcal{S}$. To this end, we introduce the following notation. For a plaquette $p = e_1 e_2 e_3 e_4 \in \mathcal{P}$, we write $p \succ e_1$ to indicate that p is a plaquette that starts from edge e_1. Note that for each edge e, there are $2(d-1)$ plaquettes in $\mathcal{P}$ such that $p \succ e$. For *any* Lie algebra $\mathfrak{g}$ embedded into $M_N(\mathbb{C})$ (denoted as M_N below for short), it forms a closed subspace of M_N, and therefore M_N has an orthogonal decomposition $M_N = \mathfrak{g} \oplus \mathfrak{g}^\perp$. Given $M \in M_N$, we denote by $\mathbf{p}M \in \mathfrak{g}$ the orthogonal projection onto $\mathfrak{g}$.

Lemma 3.1. *Writing $\cdot$ for matrix multiplication, for each $e \in E^+_{\Lambda_L}$ we have*

$$\nabla \mathcal{S}(Q)_e = N\beta \sum_{p \in \mathcal{P}_{\Lambda_L}, p \succ e} \mathbf{p} Q_p^* \cdot (Q_e^*)^{-1} \,. \tag{3.2}$$

Proof. See [SSZ22, Lemma 3.1]. We remark that in this calculation of the gradient of $\mathcal{S}$, for each fixed e, we replace the Q_p in (1.2) where p contains e or e^{-1} by a product of the form $Q_e Q.Q.Q.$, which does not change the trace. This motivates our introduction of the notation $p \succ e$. □

The above result holds for general matrix Lie groups, and for our specific choices of Lie groups, we have the SDE system (3.1) on the finite lattice Λ_L more explicitly as

$$\mathrm{d}Q_e = \nabla \mathcal{S}(Q)_e \mathrm{d}t + c_{\mathfrak{g}} Q_e \mathrm{d}t + \sqrt{2}\mathrm{d}B_e Q_e \,, \qquad (e \in E^+_\Lambda) \tag{3.3}$$

$$\nabla \mathcal{S}(Q)_e = \begin{cases} -\frac{1}{2} N\beta \sum\limits_{p \in \mathcal{P}_{\Lambda_L}, p \succ e} (Q_p - Q_p^*) Q_e \,, & G = SO(N)\,, \\ -\frac{1}{2} N\beta \sum\limits_{p \in \mathcal{P}_{\Lambda_L}, p \succ e} \Big((Q_p - Q_p^*) - \frac{1}{N} \mathrm{Tr}(Q_p - Q_p^*) I_N \Big) Q_e \,, & G = SU(N)\,. \end{cases} \tag{3.4}$$

We recall the following two results from [SSZ22, Lemmas 3.2-3.3].

Lemma 3.2. *For fixed $N \in \mathbb{N}$ and any initial data $Q(0) = (Q_e(0))_{e \in E^+_{\Lambda_L}} \in \mathcal{Q}_L$, there exists a unique solution $Q = (Q_e)_{e \in E^+_{\Lambda_L}} \in C([0,\infty); \mathcal{Q}_L)$ to (3.3).*

Lemma 3.3. *(1.1) is invariant under the SDE system (3.3).*

By global well-posedness of the SDE (3.3), the solutions form a Markov process in $\mathcal{Q}_L$. We use $(P_t^L)_{t \geqslant 0}$ to denote the associated semigroup, i.e. for $f \in C^\infty(\mathcal{Q}_L)$, $(P_t^L f)(x) = \mathbf{E} f(Q(t,x))$ for $x \in \mathcal{Q}_L$, where $Q(t,x)$ denotes the solution at time t to (3.3) starting from $x \in \mathcal{Q}_L$. We can also write down the Dirichlet form associated with $(P_t^L)_{t \geqslant 0}$. More precisely, for $F \in C^\infty(\mathcal{Q}_L)$ we consider the following symmetric quadratic form

$$\begin{aligned} \mathcal{E}^L(F,F) &\stackrel{\text{def}}{=} \int \langle \nabla F, \nabla F \rangle_{T_Q \mathcal{Q}_L} \mathrm{d}\mu_{\Lambda_L, N, \beta} \\ &= \sum_{e \in E^+_{\Lambda_L}} \int \langle \nabla_e F, \nabla_e F \rangle \mathrm{d}\mu_{\Lambda_L, N, \beta} \\ &= \sum_{e \in E^+_{\Lambda_L}} \int \mathrm{Tr}(\nabla_e F (\nabla_e F)^*) \mathrm{d}\mu_{\Lambda_L, N, \beta}. \end{aligned}$$

Using integration by parts formula for the Haar measure, we have that $(\mathcal{E}^L, C^\infty(\mathcal{Q}_L))$ is closable, and its closure $(\mathcal{E}^L, D(\mathcal{E}^L))$ is a regular Dirichlet form on $L^2(\mathcal{Q}_L, \mu_{\Lambda_L, N, \beta})$. (c.f. [FOT94].)

Recall $\mathcal{S}$ in (1.2). We write the generator associated to the above Dirichlet form for $F \in C^\infty(\mathcal{Q}_L)$ as

$$\mathcal{L}_L F = \sum_{e \in E^+_{\Lambda_L}} \Delta_e F + \sum_{e \in E^+_{\Lambda_L}} \langle \nabla \mathcal{S}(Q)_e, \nabla_e F \rangle. \tag{3.5}$$

We use $D(\mathcal{L}_L)$ to denote the domain of the generator. Moreover, $\mathcal{E}^L(F,G) = -\int \mathcal{L}_L F G \mathrm{d}\mu_{\Lambda_L,N,\beta}$ for $F, G \in C^\infty(\mathcal{Q}_L)$. It is easy to see that $(P^L_t)_{t \geqslant 0}$ is a $\mu_{\Lambda_L,N,\beta}$-version of the $L^2(\mu_{\Lambda_L,N,\beta})$-semigroup associated with the Dirichlet form $(\mathcal{E}^L, D(\mathcal{E}^L))$. (c.f. [MR92] or [FOT94].)

Recall that $\mathcal{Q} = G^{E^+}$. Now we extend Λ_L to $\mathbb{Z}^d$ and consider the SDE (1.6) on the entire space. To this end, we write $M_N^{E^+} = \prod_{e \in E^+} M_N$ for the direct product of infinitely many vector spaces M_N (i.e. an element of $M_N^{E^+}$ is allowed to have infinitely many non-zero components). We define a norm on $M_N^{E^+}$ by

$$\|Q\|^2 \stackrel{\text{def}}{=} \sum_{e \in E^+} \frac{1}{2^{|e|}} |Q_e|^2, \tag{3.6}$$

with $|Q_e|^2 = \langle Q_e, Q_e \rangle$ for $\langle \cdot, \cdot \rangle$ as in (2.3) and $|e|$ given by the distance from 0 to e in $\mathbb{Z}^d$. (More precisely, $|e|$ is the minimum of the distances from the two vertices of e to 0.) Now we give existence and uniqueness of solutions to (1.6).

Proposition 3.4. *Fix $N \in \mathbb{N}, \beta \in \mathbb{R}$. For any $Q^0 \in \mathcal{Q}$, there exists a unique probabilistically strong solution Q to (1.6) in $C([0,\infty); \mathcal{Q})$ starting from any Q^0. Namely, for a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and Brownian motion $(B_e)_{e \in E^+}$ on it, there exists an $(\mathcal{F}_t)_{t \geqslant 0}$-adapted process $Q \in C([0,\infty); \mathcal{Q})$ and Q satisfy (1.6) $\mathbb{P}$-a.s. with $(\mathcal{F}_t)_{t \geqslant 0}$ given by normal filtration generated by the Brownian motion $(B_e)_{e \in E^+}$.*

Proof. For every initial data $Q^0 \in \mathcal{Q}$, we can easily find $Q^L(0) \in \mathcal{Q}_L$ with the periodic extension still denoted by $Q^L(0)$ such that $\|Q^L(0) - Q^0\| \to 0$ as $L \to \infty$. Indeed, we can set $Q^L(0)$ as follows (the specification on $E^+_{\Lambda_L} \backslash E^+_{\Lambda_{L-1}}$ is just to ensure periodic boundary condition):

$$Q^L_e(0) = \begin{cases} Q_e(0) & e \in E^+_{\Lambda_{L-1}} \\ I_N & e \in E^+_{\Lambda_L} \backslash E^+_{\Lambda_{L-1}} \end{cases}.$$

By Lemma 3.2 we obtain a unique solution $Q^L \in C([0,\infty); \mathcal{Q}_L)$ to (3.3) from $Q^L(0)$. We could also extend Q^L to $\mathcal{Q}$ by periodic extension. Since $\mathcal{Q}$ is compact, the marginal laws of $\{Q^L\}$ at each $t \geqslant 0$ form a tight set in $\mathcal{Q}$.

Furthermore, using the SDE (3.3), for $t \geqslant s \geqslant 0$

$$Q^L_e(t) - Q^L_e(s) = \int_s^t \nabla \mathcal{S}(Q^L)_e \mathrm{d}r + \int_s^t c_{\mathfrak{g}} Q^L_e \mathrm{d}r + \sqrt{2} \int_s^t \mathrm{d}B_e Q^L_e.$$

By Itô's formula and the fact that $Q^L_e \in G$ which is *compact*, we have the following bound for $p \geqslant 1$ and $0 \leqslant s, t \leqslant T$

$$\mathbf{E}|Q^L_e(t) - Q^L_e(s)|^{2p} \leqslant C_{N,\beta,p,T}(|t-s|^{2p} + |t-s|^p),$$

where $C_{N,\beta,p,T}$ is a positive constant and may change from line to line. Since the above constant is independent of e, we have

$$\mathbf{E}\|Q^L(t) - Q^L(s)\|^{2p} \leqslant C_{N,\beta,p,T}(|t-s|^{2p} + |t-s|^p).$$

Hence, by Kolmogorov criterion we have for $\alpha < 1/2$

$$\sup_L \mathbf{E}\Big(\sup_{s\neq t\in[0,T]} \frac{\|Q^L(t)-Q^L(s)\|}{|t-s|^\alpha}\Big) < \infty.$$

Hence, the laws of $\{Q^L\}$, which are denoted by $\{\mathbb{P}^L\}$ form a tight set in $C([0,\infty);\mathcal{Q})$ equipped with the distance

$$\widetilde{\rho}(Q,Q') \stackrel{\text{def}}{=} \sum_{n=0}^\infty 2^{-n}\Big(1\wedge \sup_{t\in[n,n+1]}\|Q(t)-Q'(t)\|\Big), \quad Q,Q'\in C([0,\infty);\mathcal{Q}).$$

We write $\mathbb{P}^Q$ for one tight limit. For simplicity we still write $\{\mathbb{P}^L\}$ for the converging subsequence. Since $(C([0,\infty);\mathcal{Q}),\widetilde{\rho})$ is a Polish space, existence follows from the usual Skorohod Theorem and taking limit on the both sides of the equation. More precisely, there exists a stochastic basis $(\tilde{\Omega},\tilde{\mathscr{F}},\tilde{\mathbb{P}})$ and $C([0,\infty);\mathcal{Q})$-valued random variables $\{\tilde{Q}^L\}_L,\tilde{Q}$ on it such that $\tilde{Q}^L \stackrel{d}{=} \mathbb{P}^L$, $\tilde{Q}\stackrel{d}{=}\mathbb{P}^Q$ and $\tilde{Q}^L\to\tilde{Q}$ in $C([0,\infty);\mathcal{Q})$ $\tilde{\mathbb{P}}$-a.s., $L\to\infty$. As a result, for every $F\in C^\infty_{cyl}(\mathcal{Q})$, which can be viewed as function on $\mathcal{Q}_L$ for L large enough, we know that

$$F(\tilde{Q}^L(t)) - F(Q^L(0)) - \int_0^t \mathcal{L}_L F(\tilde{Q}^L(s))\mathrm{d}s$$

is a $\tilde{\mathbb{P}}$-martingale, where $\mathcal{L}_L$ is as in (3.5). Letting $L\to\infty$

$$F(\tilde{Q}(t)) - F(Q^0) - \sum_{e\in E^+}\int_0^t \Big(\Delta_e F(\tilde{Q}(s)) + \langle Z_e,\nabla_e F\rangle(\tilde{Q}(s))\Big)\mathrm{d}s$$

is a $\tilde{\mathbb{P}}$-martingale with

$$Z_e(Q) \stackrel{\text{def}}{=} \begin{cases} -\frac{1}{2}N\beta \sum\limits_{p\in\mathcal{P},p\succ e}(Q_p-Q_p^*)Q_e\ , & G=SO(N)\ ,\\ -\frac{1}{2}N\beta\sum\limits_{p\in\mathcal{P},p\succ e}\Big((Q_p-Q_p^*)-\frac{1}{N}\mathrm{Tr}(Q_p-Q_p^*)I_N\Big)Q_e\ , & G=SU(N)\ . \end{cases} \tag{3.7}$$

We then obtain a martingale solution to (1.6). By martingale representation theorem we could construct a stochastic basis and on it Brownian motions $(\overline{B}_e)_{e\in E^+}$ and $\overline{Q}\in C([0,\infty),\mathcal{Q})$ with law given by $\mathbb{P}^Q$ such that $\overline{Q}$ and $(\overline{B}_e)_{e\in E^+}$ satisfy SDE (1.6), which gives the existence of probabilistically weak solutions to SDE (1.6).

Now we prove pathwise uniqueness: Consider two solutions $Q,Q'\in C([0,T];\mathcal{Q})$ starting from the same initial data $Q(0)\in\mathcal{Q}$ and we apply Itô's formula to calculate $\mathrm{d}\|Q-Q'\|^2$. Since $Q_e,Q'_e\in G$ for every $e\in E^+$, by the Burkholder–Davis–Gundy inequality and (2.9) for the stochastic integral we obtain

$$\begin{aligned}&\mathbf{E}\sup_{t\in[0,T]}|Q_e-Q'_e|^2\\ &\leqslant C_{N,\beta,T}\int_0^T \mathbf{E}|Q_e-Q'_e|^2\mathrm{d}s + C_{N,\beta,T}\sum_{p\in\mathcal{P},p\succ e}\sum_{\bar{e}\in p}\int_0^T \mathbf{E}|Q_e-Q'_e||Q_{\bar{e}}-Q'_{\bar{e}}|\mathrm{d}s,\end{aligned}$$

where $C_{N,\beta,T}$ may change from line to line. We then use

$$2|Q_{\bar{e}}-Q'_{\bar{e}}||Q_e-Q'_e| \leqslant |Q_{\bar{e}}-Q'_{\bar{e}}|^2 + |Q_e-Q'_e|^2,$$

to obtain

$$\frac{1}{2^{|e|}}\mathbf{E}\sup_{t\in[0,T]}|Q_e-Q'_e|^2\leqslant C_{N,\beta,T}\frac{1}{2^{|e|}}\int_0^T\mathbf{E}|Q_e-Q'_e|^2\mathrm{d}s$$
$$+C_{N,\beta,T}\sum_{p\in\mathcal{P},p\succ e}\sum_{e\neq\bar{e}\in p}\frac{1}{2^{|\bar{e}|}}\int_0^T\mathbf{E}|Q_{\bar{e}}-Q'_{\bar{e}}|^2\mathrm{d}s.$$

Summing over e we get

$$\mathbf{E}\sup_{t\in[0,T]}\|Q-Q'\|^2\leqslant C_{N,\beta,T}\int_0^T\mathbf{E}\|Q-Q'\|^2\mathrm{d}s.$$

Hence, pathwise uniqueness follows by Gronwall's lemma. By Yamada–Watanabe Theorem [Kur07], weak existence and pathwise uniqueness gives us existence and uniqueness of strong solution. In particular one can consider other boundary conditions for finite L and the infinite volume limit solution is the same which is the unique solution to SDE (1.6). □

By Proposition 3.4, the solutions to (1.6) form a Markov process in $\mathcal{Q}$. We denote by $(P_t)_{t\geqslant 0}$ the associated semigroup. As we are in the compact setting, it is easy to obtain the tightness of the field $(\mu_{\Lambda_L,N,\beta})_L$ in $\mathcal{Q}$ as $L\to\infty$. Since by Lemma 3.3 $\mu_{\Lambda_L,N,\beta}$ is an invariant measure for (3.3), we then obtain the following result.

Theorem 3.5. *Every tight limit $\mu_{N,\beta}$ of $\{\mu_{\Lambda_L,N,\beta}\}$ is an invariant measure for* (1.6).

Proof. Suppose that a subsequence – still denoted by $\mu_{\Lambda_L,N,\beta}$ for simplicity – converges to $\mu_{N,\beta}$ weakly in $\mathcal{Q}$. We start from the unique solutions Q_L to equation (3.3) with initial distribution $\mu_{\Lambda_L,N,\beta}$ and the unique solutions Q to (1.6) with initial distribution $\mu_{N,\beta}$. By exactly the same arguments as in the proof of Proposition 3.4, we know that the laws of $\{Q_L\}$ are also tight in $C([0,T];\mathcal{Q})$ and every tight limit satisfies equations (1.6) with initial distribution $\mu_{N,\beta}$. By uniqueness of solution to (1.6) from Proposition 3.4, we have that the whole sequence of the laws of $\{Q_L\}$ converge to the law of Q in $C([0,T];\mathcal{Q})$. Since by Lemma 3.3 $\mu_{\Lambda_L,N,\beta}$ is an invariant measure for (3.3), the result follows. □

The dynamic (1.6) is gauge covariant in the following sense. For every G-valued function g on $\mathbb{Z}^d$, one can define the gauge transformation $g\circ Q$ by $(g\circ Q)_e=g_xQ_eg_y^{-1}$ where $e=\{xy\}$. If Q is a solution to (1.6), it is easy to check that $g\circ Q$ also satisfies (1.6) with B_e replaced by $g_xB_eg_x^{-1}$, which is still a Brownian motion in $\mathfrak{g}$. By Proposition 3.4, the uniqueness in law to SDE (1.6) holds. Hence, we obtain the following result.

Proposition 3.6. *Fix $N\in\mathbb{N},\beta\in\mathbb{R}$. Let Q and $\bar{Q}$ be the unique solutions to* (1.6) *with initial conditions Q^0 and $\bar{Q}^0$ in $\mathcal{Q}$ respectively. If $\bar{Q}^0=g\circ Q^0$ for some G valued function g on $\mathbb{Z}^d$, then, $\bar{Q}(t)$ and $g\circ Q(t)$ are equal in law for all $t\geqslant 0$.*

We can also write the Dirichlet form and generator associated with (1.6). Recall $C^\infty_{cyl}(\mathcal{Q})$ defined in (1.8). For every tight limit $\mu_{N,\beta}$ and $F\in C^\infty_{cyl}(\mathcal{Q})$ we define the following symmetric quadratic form

$$\begin{aligned}\mathcal{E}^{\mu_{N,\beta}}(F,F) &\stackrel{\text{def}}{=} \sum_{e\in E^+}\int \langle \nabla_e F, \nabla_e F\rangle \mathrm{d}\mu_{N,\beta}\\ &= \sum_{e\in E^+}\int \mathrm{Tr}(\nabla_e F(\nabla_e F)^*)\mathrm{d}\mu_{N,\beta}.\end{aligned} \tag{3.8}$$

By (3.5) and letting $L\to\infty$ we have that for $F,G\in C^\infty_{cyl}(\mathcal{Q})$

$$\mathcal{E}^{\mu_{N,\beta}}(F,G) = -\int \mathcal{L}F\,G\,\mathrm{d}\mu_{N,\beta}, \tag{3.9}$$

with

$$\mathcal{L}F \stackrel{\text{def}}{=} \sum_{e\in E^+}\Delta_e F + \sum_{e\in E^+}\langle Z_e, \nabla_e F\rangle,$$

for Z_e in (3.7).

Proposition 3.7. *$(\mathcal{E}^{\mu_{N,\beta}}, C^\infty_{cyl}(\mathcal{Q}))$ is closable and its closure $(\mathcal{E}^{\mu_{N,\beta}}, D(\mathcal{E}^{\mu_{N,\beta}}))$ is a regular Dirichlet form on $L^2(\mathcal{Q},\mu_{N,\beta})$.*

Proof. It is sufficient to prove the closability of $(\mathcal{E}^{\mu_{N,\beta}}, C^\infty_{cyl}(\mathcal{Q}))$. Let $\{F_n\}_{n\in\mathbb{N}}\subset C^\infty_{cyl}(\mathcal{Q})$ be such that

$$\lim_{n\to\infty}\mu_{N,\beta}(F_n^2)=0, \qquad \lim_{n,m\to\infty}\mathcal{E}^{\mu_{N,\beta}}(F_n-F_m,F_n-F_m)=0.$$

Using (3.9), for $G\in C^\infty_{cyl}(\mathcal{Q})$ we have

$$\mathcal{E}^{\mu_{N,\beta}}(G,F_n) = -\int \mathcal{L}G\,F_n\,\mathrm{d}\mu_{N,\beta}\to 0.$$

Hence, the result follows from [MR92, Chapter I. Lemma 3.4]. $\square$

4. Log-Sobolev and Poincaré inequalities and applications

In this section we prove log-Sobolev inequality under the usual Bakry–Émery condition (see (4.7) below). As applications we obtain large N limit, factorization property of rescaled Wilson loops and the exponential decay of a large class of observables.

4.1. **Log-Sobolev and Poincaré inequalities.** In this section we first prove Log-Sobolev and Poincaré inequalities for the probability measure $\mu_{\Lambda_L,N,\beta}$ on the finite dimensional compact manifold $\mathcal{Q}_L$. We then let $L\to\infty$ to derive the log-Sobolev inequality for every tight limit of $\mu_{\Lambda_L,N,\beta}$. As simple application we give a proof of correlation bounds (or susceptibility bounds) for the field Q, and that for the "microscopic Wilson loops" $\mathrm{Tr}(Q_p)$ for plaquettes p.

Below to verify the Bakry–Émery's condition we need to calculate $\mathrm{Hess}_{\mathcal{S}}(v,v)(Q)$ and $\mathrm{Ric}(v,v)(Q)$ for $v\in T_Q\mathcal{Q}_L$ and $Q\in\mathcal{Q}_L$, and we recall (2.1)(2.2) for their definitions. Following the convention in Section 2.1, we write

$$v=(v_e)_{e\in E^+_\Lambda}=\sum_{e\in E^+_\Lambda}X_eQ_e \tag{4.1}$$

with $X_e\in\mathfrak{q}_L$ being zero for all components except for the component e. We also write $|v|^2=\langle v,v\rangle_{T_Q\mathcal{Q}_L}$.

We first compute $\mathrm{Hess}_{\mathcal{S}}(v,v)$ for $v \in T_Q\mathcal{Q}_L$. Note that as a "naive" guess, $\mathcal{S}$ defined in (1.2) would appear to be of order N^2, since the trace of the orthogonal or unitary matrix Q_p would be generally bounded by N and there is another factor N outside the summation. If the Hessian of $\mathcal{S}$ was indeed of order N^2, or N^p for any $p>1$, then in Assumption 1.1 we would never be able to fix β small uniformly in N and ensure that $K_{\mathcal{S}}$ is strictly positive when N gets large. Fortunately in the next lemma by properly arranging terms and using Hölder inequalities we prove that the Hessian is actually at most of order N.

Lemma 4.1. *For $v = XQ \in T_Q\mathcal{Q}_L$ we have*

$$|\mathrm{Hess}_{\mathcal{S}}(v,v)| \leqslant 8(d-1)N|\beta||v|^2. \tag{4.2}$$

Proof. In the proof we omit the subscript L for simplicity. We view $v = XQ \in T_Q\mathcal{Q}_L$ as a right-invariant vector field on $\mathcal{Q}_L$ generated by $X \in \mathfrak{q}_L$. By (2.5) (and the discussion in Sec. 2.1) we have $\nabla_v v = 0$. We apply the second identity in (2.1) and using (4.1) we have

$$\mathrm{Hess}_{\mathcal{S}}(v,v) = v(v(\mathcal{S})) = \sum_{e,\bar{e}\in E^+_\Lambda} (X_{\bar{e}}Q_{\bar{e}})(X_eQ_e)\mathcal{S}. \tag{4.3}$$

Recall $\mathcal{S}$ from (1.2), which is a sum over plaquettes $p \in \mathcal{P}^+_{\Lambda_L}$, and is also a linear (affine) function in each variable Q_e. Since $p \in \mathcal{P}^+$, we can write $Q_p = Q_{e_1}Q_{e_2}Q^*_{e_3}Q^*_{e_4}$ where $e_1,e_2,e_3,e_4 \in E^+$. Then it is easy to calculate the derivatives using the definition (2.4), for instance

$$(X_{e_3}Q_{e_3})Q_p = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} Q_{e_1}Q_{e_2}(e^{tX_{e_3}}Q_{e_3})^*Q^*_{e_4} = Q_{e_1}Q_{e_2}Q^*_{e_3}X^*_{e_3}Q^*_{e_4} \tag{4.4}$$

which is effectively just inserting the matrix $X^*_{e_3}$.

Note that:

(1) For the terms with $e = \bar{e}$, the term involving plaquette p in $\mathcal{S}$ is non-zero if and only if the plaquette p contains e or e^{-1}. In this case, we write $p \in \mathcal{P}_e$, and there will be $2(d-1)$ such plaquettes p. Direct calculation as in (4.4) yields a result of the following form

$$(X_eQ_e)(X_eQ_e)\mathrm{Tr}(Q_p) = \mathrm{Tr}(Y_1Q_{e_1}Y_2Q_{e_2}Q^*_{e_3}Y^*_3Q^*_{e_4}Y^*_4)$$

for some matrices Y_1,Y_2,Y_3,Y_4, in which three of them are I_N and one of them is X^2_e.

(2) For the terms with $e \neq \bar{e}$, the summand on the RHS of (4.3) is non-zero if and only if there exists a plaquette p which contains both e or e^{-1} and $\bar{e}$ or $\bar{e}^{-1}$. In this case, we write $p \in \mathcal{P}_{e,\bar{e}}$ and there will be only one such plaquette, and

$$(X_{\bar{e}}Q_{\bar{e}})(X_eQ_e)\mathrm{Tr}(Q_p) = \mathrm{Tr}(Y_1Q_{e_1}Y_2Q_{e_2}Q^*_{e_3}Y^*_3Q^*_{e_4}Y^*_4)$$

where two of Y_1,Y_2,Y_3,Y_4 are I_N, and the other two are X_e and $X_{\bar{e}}$.

In either case, there are two occurrences of X, so by cyclic invariance of trace we can write the result into one of the following forms:

$$\mathrm{Tr}(QX_e\cdot\tilde{Q}X_{\bar{e}}), \qquad \mathrm{Tr}(QX_e\cdot(\tilde{Q}X_{\bar{e}})^*), \qquad \mathrm{Tr}((QX_e)^*\cdot\tilde{Q}X_{\bar{e}}), \qquad \mathrm{Tr}((QX_e)^*\cdot(\tilde{Q}X_{\bar{e}})^*)$$

for some $Q,\tilde{Q} \in G$. By the Cauchy–Schwarz inequality for the Hilbert–Schmidt inner product, each of these terms is bounded by

$$\Big(\mathrm{Tr}(QX_e\cdot(QX_e)^*)\Big)^{\frac12}\Big(\mathrm{Tr}(\tilde{Q}X_{\bar{e}}\cdot(\tilde{Q}X_{\bar{e}})^*)\Big)^{\frac12} = |X_e||X_{\bar{e}}|$$

and this is bounded by $\frac12(|X_e|^2+|X_{\bar{e}}|^2)$.

Therefore we have

$$\frac{1}{N}\sum_{e=\bar e\in E_\Lambda^+}|(X_{\bar e}Q_{\bar e})(X_eQ_e)\mathcal{S}|\leqslant\sum_{e\in E_\Lambda^+}\sum_{p\in\mathcal{P}_e}|\beta||X_e|^2=\sum_{e\in E_\Lambda^+}2|\beta|(d-1)|X_e|^2=2|\beta|(d-1)|v|^2$$

and

$$\begin{aligned}\frac{1}{N}\sum_{e\neq\bar e\in E_\Lambda^+}|(X_{\bar e}Q_{\bar e})(X_eQ_e)\mathcal{S}|&\leqslant\sum_{e\neq\bar e\in E_\Lambda^+}\sum_{p\in\mathcal{P}_{e,\bar e}}\frac{|\beta|}{2}(|X_e|^2+|X_{\bar e}|^2)\\&=\sum_{p\in\mathcal{P}_\Lambda^+}\sum_{e\neq\bar e\in E_\Lambda^+}1_{p\in\mathcal{P}_{e,\bar e}}\cdot\frac{|\beta|}{2}(|X_e|^2+|X_{\bar e}|^2)=\sum_{p\in\mathcal{P}_\Lambda^+}\sum_{e\in E_\Lambda^+}1_{p\in\mathcal{P}_e}\cdot 3|\beta||X_e|^2\\&=\sum_{e\in E_\Lambda^+}\sum_{p\in\mathcal{P}_e}3|\beta||X_e|^2=\sum_{e\in E_\Lambda^+}6(d-1)|\beta||X_e|^2=6(d-1)|\beta||v|^2\end{aligned}$$

which implies (4.2). □

We denote the Riemannian distance on G by ρ. We write ρ_L for the induced Riemannian distance on $\mathcal{Q}_L$ given by

$$\rho_L(Q,Q')^2\stackrel{\text{def}}{=}\sum_{e\in E_{\Lambda_L}^+}\rho(Q_e,Q'_e)^2,\quad Q,Q'\in\mathcal{Q}_L.$$

For any $\mu,\nu\in\mathscr{P}(\mathcal{Q}_L)$, we introduce the Wasserstein distance as

$$W_p^{\rho_L}(\mu,\nu)\stackrel{\text{def}}{=}\inf_{\pi\in\mathscr{C}(\mu,\nu)}\pi(\rho_L^p)^{1/p},$$

with $\mathscr{C}(\mu,\nu)$ being the set of couplings between μ and ν.

We then have the following result using the Bakry–Émery condition (4.7) and [Wan06, Theorem 5.6.1], which was first proved by [Bak97] and [vRS05].

Theorem 4.2. *Under Assumption 1.1, the following hold.*

(1) The dynamic defined by the SDE (3.3) is exponentially ergodic in the sense that

$$W_2^{\rho_L}(\delta_QP_t^L,\delta_{\bar Q}P_t^L)\leqslant e^{-K_{\mathcal{S}}t}\rho_L(Q,\bar Q),\quad t\geqslant0,\quad Q,\bar Q\in\mathcal{Q}_L.\tag{4.5}$$

(2) For $1<p<2$

$$W_p^{\rho_L}(\mu P_t^L,\nu P_t^L)\leqslant e^{-K_{\mathcal{S}}t}W_p^{\rho_L}(\mu,\nu),\quad t\geqslant0,\quad \mu,\nu\in\mathscr{P}(\mathcal{Q}_L),\tag{4.6}$$

In particular, invariant measure of $(P_t^L)_{t\geqslant0}$ is unique.

Proof. Using [Wan06, Theorem 5.6.1(1)(11)(12)] we know that (4.5), (4.6) and the following condition are all equivalent: for every $v=XQ\in T_Q\mathcal{Q}_L$,

$$\mathrm{Ric}(v,v)-\langle\nabla_v\nabla\mathcal{S},v\rangle\geqslant K_{\mathcal{S}}|X|^2.\tag{4.7}$$

Here we recall that $|v|^2=|X|^2$ and $\langle\nabla_v\nabla\mathcal{S},v\rangle=\mathrm{Hess}_{\mathcal{S}}(v,v)$. By [AGZ10, (F.6)], for any tangent vector u of G,

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

with $\alpha = 1, 2$ for $SO(N)$ and $SU(N)$ respectively. Since $\mathrm{Ric}(v, v) = \sum_e \mathrm{Ric}(v_e, v_e)$ and $|X|^2 = \sum_e |X_e|^2$, we have

$$\mathrm{Ric}(v, v) = \Big(\frac{\alpha(N + 2)}{4} - 1\Big)|X|^2. \tag{4.8}$$

By Lemma 4.1 and definition of $K_{\mathcal{S}}$ in Assumption 1.1, we obtain (4.7), and therefore (4.5), (4.6) follow.

Uniqueness of invariant measure follows from (4.6) by letting $t \to \infty$. □

Remark 4.3. *In general, if we do not require $K_{\mathcal{S}}$ to be strictly positive as in Assumption 1.1, (4.5)-(4.6) still hold, and (4.7) is also equivalent with the following statements: for any $t \geqslant 0, f \in C^1(\mathcal{Q}_L)$*

$$|\nabla P_t^L f| \leqslant e^{-K_{\mathcal{S}} t} P_t^L |\nabla f|, \tag{4.9}$$

$$P_t^L(f^2 \log f^2) - (P_t^L f^2) \log(P_t^L f^2) \leqslant \frac{2(1 - e^{-2K_{\mathcal{S}} t})}{K_{\mathcal{S}}} P_t^L |\nabla f|^2. \tag{4.10}$$

We refer to [Wan06, Theorem 5.6.1] *for these results and more equivalent statements.*

As (4.7) is the Bakry–Émery's condition, we have the following log-Sobolev inequality (c.f. [Wan06, Theorem 5.6.2]). In fact, it follows from taking integral w.r.t. $\mu_{\Lambda_L,N,\beta}$ on the both sides of (4.10) and letting $t \to \infty$.

Corollary 4.4. *Under Assumption 1.1, the log-Sobolev inequality holds for each $L > 1$, i.e. for $F \in C^\infty(\mathcal{Q}_L)$ with $\mu_{\Lambda_L,N,\beta}(F^2) = 1$,*

$$\mu_{\Lambda_L,N,\beta}(F^2 \log F^2) \leqslant \frac{2}{K_{\mathcal{S}}} \mathcal{E}^L(F, F).$$

This implies the Poincaré inequality: for $F \in C^\infty(\mathcal{Q}_L)$,

$$\mu_{\Lambda_L,N,\beta}(F^2) \leqslant \frac{1}{K_{\mathcal{S}}} \mathcal{E}^L(F, F) + \mu_{\Lambda_L,N,\beta}(F)^2. \tag{4.11}$$

We could view any probability measure ν in $\mathscr{P}(\mathcal{Q}_L)$ as a probability measure in $\mathscr{P}(\mathcal{Q})$ by periodic extension. Namely, we can construct a random variable with law given by $\nu \in \mathscr{P}(\mathcal{Q}_L)$ and extend the random variable periodically. The law of the periodic extension gives the desired extension of ν. Since G is compact, $\{\mu_{\Lambda_L,N,\beta}\}_L$ form a tight set and passing to a subsequence we obtain a tight limit, which is denoted by $\mu_{N,\beta}$. Hence, by approximation we have the following results.

Corollary 4.5. *Under Assumption 1.1, the log-Sobolev inequality holds, i.e. for cylinder functions $F \in C^\infty_{cyl}(\mathcal{Q})$ with $\mu_{N,\beta}(F^2) = 1$,*

$$\mu_{N,\beta}(F^2 \log F^2) \leqslant \frac{2}{K_{\mathcal{S}}} \mathcal{E}^{\mu_{N,\beta}}(F, F). \tag{4.12}$$

This implies the Poincaré inequality: for cylinder functions $F \in C^\infty_{cyl}(\mathcal{Q})$

$$\mu_{N,\beta}(F^2) \leqslant \frac{1}{K_{\mathcal{S}}} \mathcal{E}^{\mu_{N,\beta}}(F, F) + \mu_{N,\beta}(F)^2. \tag{4.13}$$

In Section 5 we will prove Theorem 1.2 which will then identify the tight limit $\mu_{N,\beta}$ in (4.12) and (4.13) as the measure $\mu^{\mathrm{YM}}_{N,\beta}$ in Theorem 1.2; this then proves Theorem 1.4.

Remark 4.6. *By the Poincaré inequality* (4.11) *and* (4.13)*, the semigroup* $(P_t^L)_{t\geqslant 0}$ *and* $(P_t)_{t\geqslant 0}$ *satisfy*

$$\|P_t^L f - \mu_{\Lambda_L,N,\beta}(f)\|_{L^2(\mu_{\Lambda_L,N,\beta})} \leqslant e^{-tK_{\mathcal{S}}}\|f\|_{L^2(\mu_{\Lambda_L,N,\beta})},$$

and

$$\|P_t f - \mu_{N,\beta}(f)\|_{L^2(\mu_{N,\beta})} \leqslant e^{-tK_{\mathcal{S}}}\|f\|_{L^2(\mu_{N,\beta})}.$$

(c.f. [Wan06, Theorem 1.1.1]*). However, this does not imply the uniqueness of the invariant measure for* $(P_t)_{t\geqslant 0}$*.*

The following two results are simple applications of the Poincaré inequality.

Corollary 4.7. *Under Assumption 1.1, for every* $e_0 \in E^+$ *and every unit vector* E *in* M_N *we have*

$$\sum_{e\in E^+_{\Lambda_L}} \mathrm{Cov}_{N,\beta,L}\Big(\langle Q_{e_0}, E\rangle, \langle Q_e, E\rangle\Big) \leqslant \begin{cases} 1/K_{\mathcal{S}}, & G = SO(N), \\ 2/K_{\mathcal{S}}, & G = SU(N). \end{cases}$$

Here $\mathrm{Cov}_{N,\beta,L}$ *means covariance w.r.t. the measure* $\mu_{\Lambda_L,N,\beta}$*. In particular,*

$$\Big|\sum_{e\in E^+_{\Lambda_L}\setminus\{e_0\}} \mathrm{Cov}_{N,\beta,L}\Big(\langle Q_{e_0}, E\rangle, \langle Q_e, E\rangle\Big)\Big| \leqslant \begin{cases} 2/K_{\mathcal{S}}, & G = SO(N), \\ 4/K_{\mathcal{S}}, & G = SU(N). \end{cases}$$

Proof. Let $f = |E^+_{\Lambda_L}|^{-\frac{1}{2}}\sum_{e\in E^+_{\Lambda_L}}\langle Q_e, E\rangle$. By direct calculation, one has $\nabla\langle Q_e, E\rangle = \mathbf{p}(EQ_e^*)Q_e$, which implies that $|\nabla f|^2 \leqslant \gamma$ with $\gamma = 1$ for $G = SO(N)$ and $\gamma = 2$ for $G = SU(N)$, where for $G = SU(N)$ we used that for any matrices $Q, Q' \in M_N$

$$\mathrm{Tr}\Big((Q - \frac{1}{N}\mathrm{Tr}(Q)I_N)(Q' - \frac{1}{N}\mathrm{Tr}(Q')I_N)\Big) = \mathrm{Tr}(QQ') - \frac{1}{N}\mathrm{Tr}(Q)\mathrm{Tr}(Q'). \tag{4.14}$$

Hence, by the Poincaré inequality (4.11) we get

$$\frac{1}{|E^+_{\Lambda_L}|}\sum_{e,e'\in E^+_{\Lambda_L}} \mathrm{Cov}_{N,\beta,L}\Big(\langle Q_e, E\rangle, \langle Q_{e'}, E\rangle\Big) \leqslant \frac{\gamma}{K_{\mathcal{S}}}.$$

With periodic boundary condition we have translation invariance, so for fixed edge e_0

$$\sum_{e\in E^+_{\Lambda_L}} \mathrm{Cov}_{N,\beta,L}\Big(\langle Q_{e_0}, E\rangle, \langle Q_e, E\rangle\Big) \leqslant \frac{\gamma}{K_{\mathcal{S}}},$$

which implies the first result and

$$\Big|\sum_{e\in E^+_{\Lambda_L}\setminus\{e_0\}} \mathrm{Cov}_{N,\beta,L}\Big(\langle Q_{e_0}, E\rangle, \langle Q_e, E\rangle\Big)\Big| \leqslant \frac{\gamma}{K_{\mathcal{S}}} + \mathrm{Var}_{N,\beta,L}\Big(\langle Q_{e_0}, E\rangle\Big),$$

where we used triangle inequality and $\mathrm{Var}_{N,\beta,L}$ means variance under $\mu_{\Lambda_L,N,\beta}$. Now we take $g = \langle Q_{e_0}, E\rangle$ and have $|\nabla g|^2 \leqslant \gamma$. Then by the Poincaré inequality (4.11)

$$\mathrm{Var}_{\Lambda_L,N,\beta}\Big(\langle Q_{e_0}, E\rangle\Big) \leqslant \frac{\gamma}{K_{\mathcal{S}}}.$$

Thus the second result follows. $\square$

Corollary 4.8. *Under Assumption 1.1, it holds that for every plaquette p in $\mathcal{P}^+_{\Lambda_L}$*

$$\sum_{\bar p\in\mathcal{P}_{\Lambda_L}}\mathrm{Cov}_{N,\beta,L}\Big(\mathrm{ReTr}Q_p,\mathrm{ReTr}Q_{\bar p}\Big)\leqslant\begin{cases}8N(d-1)/K_{\mathcal{S}}, & G=SO(N),\\ 16N(d-1)/K_{\mathcal{S}}, & G=SU(N).\end{cases}$$

Here $\mathrm{Cov}_{N,\beta,L}$ means covariance w.r.t. the measure $\mu_{\Lambda_L,N,\beta}$. In particular, for every plaquette p in $\mathcal{P}^+_{\Lambda_L}$

$$\Big|\sum_{\bar p\in\mathcal{P}_{\Lambda_L},\bar p\neq p}\mathrm{Cov}_{N,\beta,L}\Big(\mathrm{ReTr}Q_p,\mathrm{ReTr}Q_{\bar p}\Big)\Big|\leqslant\begin{cases}\big(8N(d-1)+4N\big)/K_{\mathcal{S}}, & G=SO(N),\\ \big(16N(d-1)+8N\big)/K_{\mathcal{S}}, & G=SU(N).\end{cases}$$

Proof. Let $f=|\mathcal{P}^+_{\Lambda_L}|^{-\frac12}\sum_{\bar p\in\mathcal{P}^+_{\Lambda_L}}\mathrm{ReTr}Q_{\bar p}$. By the same calculation as (3.2) and (3.4) for the action $\mathcal{S}$ we have

$$\nabla f_e=\begin{cases}-\frac{1}{2|\mathcal{P}^+_{\Lambda_L}|^{1/2}}\sum_{p\in\mathcal{P}_\Lambda,p\succ e}(Q_p-Q_p^*)Q_e\;, & G=SO(N)\;,\\ -\frac{1}{2|\mathcal{P}^+_{\Lambda_L}|^{1/2}}\sum_{p\in\mathcal{P}_\Lambda,p\succ e}\Big((Q_p-Q_p^*)-\frac1N\mathrm{Tr}(Q_p-Q_p^*)I_N\Big)Q_e\;, & G=SU(N)\;.\end{cases}$$

Thus in $SO(N)$ case we have

$$|\nabla f_e|^2=\frac{1}{4|\mathcal{P}^+_{\Lambda_L}|}\sum_{p,\bar p\in\mathcal{P}_\Lambda,p,\bar p\succ e}\mathrm{Tr}\Big((Q_p-Q_p^*)(Q_{\bar p}-Q_{\bar p}^*)^*\Big)\leqslant\frac{4N(d-1)^2}{|\mathcal{P}^+_{\Lambda_L}|},$$

which implies that

$$|\nabla f|^2\leqslant 8N(d-1).$$

Here we used $|\mathcal{P}^+_{\Lambda_L}|=\frac{2(d-1)}{4}|E^+_{\Lambda_L}|$, since each plaquette has 4 edges and each edge is adjacent to $2(d-1)$ plaquettes. Hence, applying the Poincaré inequality to f we get

$$\frac{1}{|\mathcal{P}^+_{\Lambda_L}|}\sum_{p,\bar p\in\mathcal{P}^+_{\Lambda_L}}\mathrm{Cov}_{N,\beta,L}\Big(\mathrm{ReTr}Q_p,\mathrm{ReTr}Q_{\bar p}\Big)\leqslant\frac{8N(d-1)}{K_{\mathcal{S}}}.$$

We choose periodic boundary condition to have translation invariance and we get for fixed p

$$\sum_{\bar p\in\mathcal{P}^+_{\Lambda_L}}\mathrm{Cov}_{N,\beta,L}\Big(\mathrm{ReTr}Q_p,\mathrm{ReTr}Q_{\bar p}\Big)\leqslant\frac{8N(d-1)}{K_{\mathcal{S}}}.$$

Thus the first result follows and

$$\Big|\sum_{\bar p\in\mathcal{P}^+_{\Lambda_L},\bar p\neq p}\mathrm{Cov}_{N,\beta,L}\Big(\mathrm{ReTr}Q_p,\mathrm{ReTr}Q_{\bar p}\Big)\Big|\leqslant\frac{8N(d-1)}{K_{\mathcal{S}}}+\mathrm{Var}\Big(\mathrm{ReTr}Q_p\Big).$$

Moreover, take $f=\mathrm{ReTr}Q_{\bar p}$ and $|\nabla f|^2\leqslant 4N$ and we obtain

$$\mathrm{Var}\Big(\mathrm{ReTr}Q_p\Big)\leqslant\frac{4N}{K_{\mathcal{S}}}.$$

Thus the second result follows for $SO(N)$ case. The result for the $SU(N)$ case follows by similar arguments and using (4.14). ☐

4.2. **Application I: large N limit of Wilson loops.** In the following we give the proof of Corollary 1.5 by applying the Poincaré inequality.

Proof of Corollary 1.5. Since Theorem 1.2 identifies any tight limit $\mu_{N,\beta}$ as the measure $\mu^{\text{YM}}_{N,\beta}$, it suffices to prove the result for any tight limit $\mu_{N,\beta}$. We apply the Poincaré inequality (4.13) to Wilson loops defined in (1.11). Consider the $SO(N)$ case. Let

$$f(Q)=\frac{1}{N}W_\ell=\frac{1}{N}\mathrm{Tr}(Q_{e_1}Q_{e_2}\dots Q_{e_n}).$$

We get

$$\mu_{N,\beta}(f^2)-\mu_{N,\beta}(f)^2=\mathrm{Var}\Big(\frac{1}{N}W_\ell\Big).$$

We then need to calculate ∇f which appears on the RHS of the Poincaré inequality. For an edge which appears in the location x of the loop ℓ, we write

$$Q_\ell=\prod_{i=1}^n Q_{e_i},\qquad Q_{a_x}=\prod_{i=1}^{x-1}Q_{e_i},\qquad Q_{b_x}=\prod_{i=x+1}^{n}Q_{e_i}.$$

We then have $W_\ell=\mathrm{Tr}(Q_\ell)$. For each $e\in E^+$, we may have an edge e_x in ℓ which is e or e^{-1}, so by straightforward calculation we have

$$\begin{aligned}(\nabla W_\ell)_e=&-\frac12\sum_{x=1}^n \mathbf{1}_{e_x=e}(Q_{e_x}Q_{b_x}Q_{a_x}-Q^*_{a_x}Q^*_{b_x}Q^*_{e_x})Q_e\\&+\frac12\sum_{x=1}^n \mathbf{1}_{e_x=e^{-1}}(Q_{b_x}Q_{a_x}Q^*_e-Q_eQ^*_{a_x}Q^*_{b_x})Q_e.\end{aligned}\tag{4.15}$$

Here, the calculation is similar as in Lemma 3.1 (see [SSZ22]). Namely, when $e_x=e$, by cyclic invariance of trace, we write $W_\ell=\mathrm{Tr}(Q_{e_x}Q_{b_x}Q_{a_x})$, and for $X\in\mathfrak{g}$ we compute

$$\begin{aligned}\partial_t|_{t=0}\mathrm{Tr}(e^{tX}Q_{e_x}Q_{b_x}Q_{a_x})=\mathrm{Tr}(XQ_{e_x}Q_{b_x}Q_{a_x})&=\langle X,\mathbf{p}(Q_{e_x}Q_{b_x}Q_{a_x})^*\rangle\\&=\langle XQ_{e_x},\mathbf{p}(Q_{e_x}Q_{b_x}Q_{a_x})^*Q_{e_x}\rangle\end{aligned}\tag{4.16}$$

where $\mathbf{p}$ is defined in Lemma 3.1 and is the orthogonal projection of a matrix to the Lie algebra $\mathfrak{g}$ of skew-symmetric matrices, and XQ_{e_x} is a tangent vector at Q_{e_x}. On the other hand if $e_x=e^{-1}$, we write $W_\ell=\mathrm{Tr}(Q_{b_x}Q_{a_x}Q^*_e)$, so for $X\in\mathfrak{g}$ similar calculation as above yields

$$\mathrm{Tr}(Q_{b_x}Q_{a_x}Q^*_eX^*)=\langle X,\mathbf{p}(Q_{b_x}Q_{a_x}Q^*_e)\rangle=\langle XQ_e,\mathbf{p}(Q_{b_x}Q_{a_x}Q^*_e)Q_e\rangle\tag{4.17}$$

which gives the second term on the RHS of (4.15).

Using (4.15) we have

$$\begin{aligned}|(\nabla W_\ell)_e|^2=&\frac14\sum_{x,y=1}^n\mathbf{1}_{e_x=e_y=e}\mathrm{Tr}\Big((Q_{e_x}Q_{b_x}Q_{a_x}Q_{e_x}-Q^*_{a_x}Q^*_{b_x})(Q_{e_y}Q_{b_y}Q_{a_y}Q_{e_y}-Q^*_{a_y}Q^*_{b_y})^*\Big)\\&-\frac14\sum_{x,y=1}^n\mathbf{1}_{e_x^{-1}=e_y=e}\mathrm{Tr}\Big((Q_{b_x}Q_{a_x}-Q_eQ^*_{a_x}Q^*_{b_x}Q_e)(Q_{e_y}Q_{b_y}Q_{a_y}Q_{e_y}-Q^*_{a_y}Q^*_{b_y})^*\Big)\\&-\frac14\sum_{x,y=1}^n\mathbf{1}_{e_x=e_y^{-1}=e}\mathrm{Tr}\Big((Q_{e_x}Q_{b_x}Q_{a_x}Q_{e_x}-Q^*_{a_x}Q^*_{b_x})(Q_{b_y}Q_{a_y}-Q_eQ^*_{a_y}Q^*_{b_y}Q_e)^*\Big)\end{aligned}$$

$$+\frac{1}{4}\sum_{x,y=1}^{n}\mathbf{1}_{e_x=e_y=e^{-1}}\mathrm{Tr}\Big((Q_{b_x}Q_{a_x}-Q_eQ_{a_x}^*Q_{b_x}^*Q_e)(Q_{b_y}Q_{a_y}-Q_eQ_{a_y}^*Q_{b_y}^*Q_e)^*\Big).$$

Note that the trace of any $SO(N)$ matrix is bounded by N, and therefore each of the four traces above is bounded by $4N$. Summing over $e\in E^+$, we see that the Dirichlet form term in the Poincaré inequality is bounded as follows:

$$\begin{aligned}\mathcal{E}^{\mu_{N,\beta}}(f,f)&=\sum_{e\in E^+}\mu_{N,\beta}(|\nabla_e f|^2)\\&\leqslant\frac{1}{N}\sum_{e\in E^+}\sum_{x,y=1}^{n}\Big(\mathbf{1}_{e_x=e_y=e}+\mathbf{1}_{e_x^{-1}=e_y=e}+\mathbf{1}_{e_x=e_y^{-1}=e}+\mathbf{1}_{e_x=e_y=e^{-1}}\Big).\end{aligned}\tag{4.18}$$

For any edge $e\in E^+$ we let $A(e)$ be the number of locations in ℓ where e occurs and $B(e)$ be the number of locations in ℓ where e^{-1} occurs. (4.18) is then bounded by

$$\frac{1}{N}\sum_{e\in E^+}(A(e)+B(e))^2\leqslant\frac{n(n-3)}{N},$$

where we used $\sum_{e\in E^+}(A(e)+B(e))=n$ and $A(e)+B(e)\leqslant n-3$. The Poincaré inequality then yields

$$\mathrm{Var}\Big(\frac{1}{N}W_\ell\Big)\leqslant\frac{1}{K_{\mathcal{S}}}\frac{n(n-3)}{N}.$$

Letting $N\to\infty$, (1.13) follows for the $SO(N)$ case.

For $G=SU(N)$ we choose, with $\iota=\sqrt{-1}$,

$$\begin{aligned}f_R(Q)&=\frac{1}{N}\mathrm{Re}W_\ell=\frac{1}{N}\mathrm{ReTr}\Big(Q_{e_1}Q_{e_2}\dots Q_{e_n}\Big),\\f_I(Q)&=\frac{1}{N}\mathrm{Im}W_\ell=-\frac{1}{N}\mathrm{ReTr}\Big(\iota Q_{e_1}Q_{e_2}\dots Q_{e_n}\Big)\end{aligned}$$

to obtain the result for the real and imaginary parts. It is sufficient to calculate $\nabla\mathrm{Re}W_\ell$. Besides the terms in (4.15) we also have the following additional terms

$$\begin{aligned}&\frac{1}{2}\sum_{x=1}^{n}\mathbf{1}_{e_x=e}\frac{1}{N}\mathrm{Tr}(Q_{e_x}Q_{b_x}Q_{a_x}-Q_{a_x}^*Q_{b_x}^*Q_{e_x}^*)Q_e\\&-\frac{1}{2}\sum_{x=1}^{n}\mathbf{1}_{e_x=e^{-1}}\frac{1}{N}\mathrm{Tr}(Q_{b_x}Q_{a_x}Q_{e_x}-Q_eQ_{a_x}^*Q_{b_x}^*)Q_e.\end{aligned}$$

(This is similar with how the second case in (3.4) was derived, namely, the projection $\mathbf{p}$ appearing in (4.16)(4.17) should also make the matrices traceless in the $SU(N)$ case.) Noting that (4.14) we have

$$|(\nabla\mathrm{Re}W_\ell)_e|^2\leqslant 2\sum_{x,y=1}^{n}\Big(\mathbf{1}_{e_x=e_y=e}+\mathbf{1}_{e_x^{-1}=e_y=e}+\mathbf{1}_{e_x=e_y^{-1}=e}+\mathbf{1}_{e_x=e_y=e^{-1}}\Big)N.$$

Summing over $e\in E^+$ we get

$$\mathcal{E}(f_R,f_R)\leqslant\frac{2}{N}\sum_{e\in E^+}\sum_{x,y=1}^{n}\Big(\mathbf{1}_{e_x=e_y=e}+\mathbf{1}_{e_x^{-1}=e_y=e}+\mathbf{1}_{e_x=e_y^{-1}=e}+\mathbf{1}_{e_x=e_y=e^{-1}}\Big)$$

$$\leqslant \frac{2n(n-3)}{N}.$$

Similarly, we get

$$\mathcal{E}(f_I, f_I) \leqslant \frac{2n(n-3)}{N}.$$

Hence, (1.13) holds for $SU(N)$.

To prove the factorization property, by the Cauchy–Schwarz inequality we have

$$\begin{aligned}
&N^{-n}\Big|\mathbf{E}(W_{\ell_1}\dots W_{\ell_n}) - \mathbf{E}(W_{\ell_1}\dots W_{\ell_{n-1}})\mathbf{E}W_{\ell_n}\Big| \\
&\leqslant N^{-n}\Big|\mathbf{E}(W_{\ell_1}\dots W_{\ell_{n-1}}(W_{\ell_n} - \mathbf{E}W_{\ell_n})\Big| \\
&\leqslant \mathrm{Var}\Big(\frac{W_{\ell_n}}{N}\Big)^{1/2} \to 0.
\end{aligned}$$

Hence, the result follows by induction. □

4.3. **Application II: mass gap.** In this section we use the Poincaré inequality to prove the existence of mass gap for lattice Yang–Mills. To this end, for $f \in C^\infty_{cyl}(\mathcal{Q})$, recall that Λ_f is the set of edges f depends on. We define

$$|||f|||_\infty \stackrel{\text{def}}{=} \sum_{e\in\Lambda_f} \|\nabla_e f\|_{L^\infty}.$$

In this section it will be convenient for the calculations to consider an explicit choice of an orthonormal basis of $\mathfrak{g}$. This choice is standard, see e.g. [AGZ10, Proposition E.15]. Let $e_{kn} \in M_N$ for $k, n = 1, \dots, N$ be the elementary matrices, namely its (k, n)-th entry is 1 and all the other entries are 0. For $1 \leqslant k < N$ and $\iota = \sqrt{-1}$, let

$$D_k = \frac{\iota}{\sqrt{k+k^2}}\Big(-ke_{k+1,k+1} + \sum_{i=1}^k e_{ii}\Big).$$

For $1 \leqslant k, n \leqslant N$, let

$$E_{kn} = \frac{e_{kn} - e_{nk}}{\sqrt{2}}, \quad F_{kn} = \frac{\iota e_{kn} + \iota e_{nk}}{\sqrt{2}}. \tag{4.19}$$

Then:

- $\{E_{kn} : 1 \leqslant k < n \leqslant N\}$ is an orthonormal basis of $\mathfrak{so}(N)$, and,
- $\{D_k : 1 \leqslant k < N\} \cup \{E_{kn}, F_{kn} : 1 \leqslant k < n \leqslant N\}$ is an orthonormal basis of $\mathfrak{su}(N)$.

This then determines an orthonormal basis $\{v_e^i\}$ of $\mathfrak{g}^{E^+_{\Lambda_L}}$, which consists of right-invariant vector fields on $\mathcal{Q}_L$.

We first prove the following lemma for Lie brackets.

Lemma 4.9. *It holds that for every* v_e^i

$$\sum_j |[v_e^i, v_e^j]f|^2 \leqslant \frac{1}{2}|\nabla_e f|^2 \quad \textit{for} \quad G = SO(N),$$

$$\sum_j |[v_e^i, v_e^j]f|^2 \leqslant \frac{9}{2}|\nabla_e f|^2 \quad \textit{for} \quad G = SU(N).$$

Proof. By direct calculation we have

$$e_{ij}e_{mn} = \delta_{jm}e_{in}.$$

Using this and (4.19), we deduce

$$\begin{aligned}
2[E_{kn}, E_{lm}] &= [e_{kn} - e_{nk}, e_{lm} - e_{ml}] \\
&= \delta_{nl}e_{km} - \delta_{km}e_{ln} - \delta_{nm}e_{kl} + \delta_{lk}e_{mn} \\
&\quad - \delta_{kl}e_{nm} + \delta_{mn}e_{lk} + \delta_{km}e_{nl} - \delta_{nl}e_{mk} \\
&= \delta_{nl}(e_{km} - e_{mk}) + \delta_{km}(e_{nl} - e_{ln}) + \delta_{nm}(e_{lk} - e_{kl}) + \delta_{lk}(e_{mn} - e_{nm}).
\end{aligned} \tag{4.20}$$

With this calculation, observe that if we fix (k, n) and vary (l, m), we either get 0 or one of the orthonormal basis vectors of $\mathfrak{so}(N)$ up to a factor $\pm\frac{1}{\sqrt{2}}$, and in the latter case different values of (l, m) yield different basis vectors. This implies that for $G = SO(N)$, $\{[v_e^i, v_e^j], j = 1, \dots, d(\mathfrak{g})\}$ is a subset of orthonormal basis of $T_{Q_e}G$ up to a factor $\pm\frac{1}{\sqrt{2}}$. Hence, the result holds for $SO(N)$ by definition of $|\nabla_e f|^2$.

The proof for the $SU(N)$ case is similar but requires a bit more calculations. We have

$$\begin{aligned}
\frac{2}{\iota}[E_{kn}, F_{lm}] &= [e_{kn} - e_{nk}, e_{lm} + e_{ml}] \\
&= \delta_{nl}e_{km} - \delta_{km}e_{ln} + \delta_{nm}e_{kl} - \delta_{lk}e_{mn} \\
&\quad - \delta_{kl}e_{nm} + \delta_{mn}e_{lk} - \delta_{km}e_{nl} + \delta_{nl}e_{mk} \\
&= \delta_{nl}(e_{km} + e_{mk}) + \delta_{km}(-e_{nl} - e_{ln}) + \delta_{nm}(e_{lk} + e_{kl}) + \delta_{lk}(-e_{mn} - e_{nm}).
\end{aligned} \tag{4.21}$$

When $k = l \neq n = m$

$$\frac{2}{\iota}[E_{kn}, F_{lm}] = \delta_{nm}(e_{lk} + e_{kl}) + \delta_{lk}(-e_{mn} - e_{nm}) = 2e_{kk} - 2e_{nn}. \tag{4.22}$$

Furthermore

$$\begin{aligned}
-2[F_{kn}, F_{lm}] &= [e_{kn} + e_{nk}, e_{lm} + e_{ml}] \\
&= \delta_{nl}e_{km} - \delta_{km}e_{ln} + \delta_{nm}e_{kl} - \delta_{lk}e_{mn} \\
&\quad + \delta_{kl}e_{nm} - \delta_{mn}e_{lk} + \delta_{km}e_{nl} - \delta_{nl}e_{mk} \\
&= \delta_{nl}(e_{km} - e_{mk}) + \delta_{km}(e_{nl} - e_{ln}) + \delta_{nm}(e_{kl} - e_{lk}) + \delta_{lk}(e_{nm} - e_{mn}).
\end{aligned} \tag{4.23}$$

For Lie brackets involving D we have

$$\begin{aligned}
[E_{kn}, e_{mm}] &= (\delta_{mn}F_{mk} - \delta_{mk}F_{mn})/\iota. \\
[F_{kn}, e_{mm}] &= \iota(\delta_{mn}E_{km} + \delta_{mk}E_{nm}).
\end{aligned}$$

By this we obtain, for $k < n$,

$$[E_{kn}, D_m] = \begin{cases} 0 & m+1 < k \\ \frac{m}{\sqrt{m+m^2}}F_{kn} & m+1 = k \\ -\frac{1}{\sqrt{m+m^2}}F_{kn} & k \leqslant m < m+1 < n \\ -\frac{m+1}{\sqrt{m+m^2}}F_{kn} & k \leqslant m < m+1 = n \\ 0 & k, n \leqslant m, \end{cases} \tag{4.24}$$

and

$$[F_{kn}, D_m] = \begin{cases} 0 & m+1<k \\ -\frac{m}{\sqrt{m+m^2}}E_{kn} & m+1=k \\ \frac{1}{\sqrt{m+m^2}}E_{kn} & k\leqslant m<m+1<n \\ \frac{m+1}{\sqrt{m+m^2}}E_{kn} & k\leqslant m<m+1=n \\ 0 & k,n\leqslant m. \end{cases} \tag{4.25}$$

For $v_e^i = E_{kn}Q_e$ or $F_{kn}Q_e$, we decompose $\{v_e^p, p=1,\dots,d(\mathfrak{g})\}$ into the following three sets

$$I_1 = \{E_{lm}Q_e, F_{lm}Q_e, 1\leqslant l<m\leqslant N, l\neq k \text{ or } n\neq m\}, \quad I_2=\{E_{kn}Q_e, F_{kn}Q_e\},$$

and

$$I_3 = \{D_kQ_e, 1\leqslant k\leqslant N-1\}.$$

By (4.20)-(4.23) we view $\{[v_e^i, v_e^j], v_e^j\in I_1\}$ as a subset of orthonormal basis of $T_{Q_e}G$ up to a factor $\pm\frac{1}{\sqrt{2}}$. Hence,

$$\sum_{v_e^j\in I_1} |[v_e^i, v_e^j]f|^2 \leqslant \frac{1}{2}|\nabla_e f|^2.$$

We further use (4.22) to have

$$\sum_{v_e^j\in I_2} |[v_e^i, v_e^j]f|^2 \leqslant |\nabla_e f|^2 |e_{kk}-e_{nn}|^2 = 2|\nabla_e f|^2.$$

We also use (4.24)-(4.25) to have

$$\begin{aligned}\sum_{v_e^j\in I_3} |[v_e^i, v_e^j]f|^2 &\leqslant \Big(\frac{(k-1)^2}{k-1+(k-1)^2}\mathbf{1}_{k\geqslant 2} + \frac{n^2}{n-1+(n-1)^2} + \sum_{m=k}^{n-2}\frac{1}{m+m^2}\Big)|\nabla_e f|^2 \\ &= \Big(2-\frac{1}{k}+\frac{1}{n-1}+\sum_{m=k}^{n-2}(\frac{1}{m}-\frac{1}{m+1})\Big)|\nabla_e f|^2 = 2|\nabla_e f|^2.\end{aligned}$$

As a consequence, the result holds for $v_e^i = E_{kn}Q_e$ or $F_{kn}Q_e$.

For $v_e^i = D_mQ_e$ as $[D_l, D_m]=0$ we also use (4.24)-(4.25) to view $\{[v_e^i, v_e^j], j=1,\dots,d(\mathfrak{g})\}$ as a subset of orthonormal basis up to a factor with absolute value smaller than $\sqrt{2}$. We then have

$$\sum_j |[v_e^i, v_e^j]f|^2 \leqslant 2|\nabla_e f|^2.$$

Hence, the result follows. □

We first prove the following lemma. We write $\bar{e}\sim e$ if $\bar{e}$ and e appear in the same plaquette; more precisely, if there exists $p\in\mathcal{P}$ such that $\{e, e^{-1}\}\cap p\neq\varnothing$ and $\{\bar{e},\bar{e}^{-1}\}\cap p\neq\varnothing$.

Lemma 4.10. *Let $\{v_e^i\}$ be the orthonormal basis given above. For every $f\in C^\infty(\mathcal{Q}_L)$ and every $e\in E^+_{\Lambda_L}$, one has*

$$|[v_e^i, \mathcal{L}_L]f(Q)| \leqslant \sum_{E^+_{\Lambda_L}\ni\bar{e}\sim e} a_{e,\bar{e}}|\nabla_{\bar{e}}f(Q)|, \qquad \forall Q\in\mathcal{Q}_L,$$

with $a_{e,\bar{e}} = N|\beta|\sqrt{d(\mathfrak{g})}$ for $e\neq\bar{e}$ and

$$a_{e,e} = 2(d-1)N|\beta|(\sqrt{d(\mathfrak{g})}+\sqrt{2}N^{1/2}\gamma)\ ,$$

where $\gamma = 1$ *when* $G = SO(N)$ *and* $\gamma = 3\sqrt{2}$ *for* $G = SU(N)$.

Proof. In this proof all the sums over $\bar{e}$ are restricted to $E^+_{\Lambda_L}$. Since the metric on G is bi-invariant and each v^i_e is right-invariant which generates a one-parameter family of isometries, v^i_e commutes with the Beltrami-Laplacian Δ_e. So we have

$$\begin{aligned}[v^i_e, \mathcal{L}_L]f &= v^i_e \mathcal{L}_L f - \mathcal{L}_L v^i_e f \\ &= \sum_{\bar{e}\sim e} \left\langle \nabla_{v^i_e} \nabla_{\bar{e}} \mathcal{S}, \nabla_{\bar{e}} f \right\rangle + \left\langle \nabla_e \mathcal{S} \, , \, \nabla_{v^i_e} \nabla_e f - \nabla_e v^i_e f \right\rangle .\end{aligned}$$

Writing $\nabla_{\bar{e}}\mathcal{S} = \sum_j (v^j_{\bar{e}} \mathcal{S}) v^j_{\bar{e}}$, and using (2.5), the first term on the RHS is equal to

$$\sum_{\bar{e}\sim e} \Big\langle \sum_j (v^i_e v^j_{\bar{e}} \mathcal{S}) v^j_{\bar{e}}, \nabla_{\bar{e}} f \Big\rangle + \frac{1}{2} \sum_j \Big\langle (v^j_e \mathcal{S}) [v^i_e, v^j_e], \nabla_e f \Big\rangle .$$

For the second term we use $\nabla_e f = \sum_j (v^j_e f) v^j_e$ and (2.5) to write it as

$$\begin{aligned}&\sum_j \Big\langle \nabla_e \mathcal{S}, (v^i_e v^j_e f) v^j_e + v^j_e f \nabla_{v^i_e} v^j_e - (v^j_e v^i_e f) v^j_e \Big\rangle \\ &= \sum_j (v^j_e \mathcal{S}) \Big\langle [v^i_e, v^j_e], \nabla_e f \Big\rangle + \frac{1}{2} \sum_j v^j_e f \Big\langle \nabla_e \mathcal{S}, [v^i_e, v^j_e] \Big\rangle .\end{aligned}$$

Therefore,

$$\begin{aligned}[v^i_e, \mathcal{L}_L]f = &\sum_{\bar{e}\sim e} \Big\langle \sum_j (v^i_e v^j_{\bar{e}} \mathcal{S}) v^j_{\bar{e}}, \nabla_{\bar{e}} f \Big\rangle + \frac{3}{2} \sum_j (v^j_e \mathcal{S}) \Big\langle [v^i_e, v^j_e], \nabla_e f \Big\rangle \\ &+ \frac{1}{2} \sum_j v^j_e f \Big\langle \nabla_e \mathcal{S}, [v^i_e, v^j_e] \Big\rangle \stackrel{\text{def}}{=} \sum_{k=1}^3 I_k.\end{aligned}$$

For I_1, by similar calculation as in the proof of Lemma 4.1, we have $|v^i_e v^j_{\bar{e}} \mathcal{S}| \leqslant N|\beta|$ for $e \neq \bar{e}$; also, $|v^i_e v^j_e \mathcal{S}| \leqslant 2(d-1)N|\beta|$ since for each edge e there are $2(d-1)$ plaquettes containing e or e^{-1}. Combining with Hölder's inequality we have

$$\begin{aligned}|I_1| = \Big| \sum_{\bar{e}\sim e} \sum_j (v^i_e v^j_{\bar{e}} \mathcal{S}) v^j_{\bar{e}} f \Big| &\leqslant \sum_{\bar{e}\sim e} \Big(\sum_j |v^i_e v^j_{\bar{e}} \mathcal{S}|^2 \Big)^{1/2} \Big(\sum_j |v^j_{\bar{e}} f|^2 \Big)^{1/2} \\ &\leqslant N|\beta| \sqrt{d(\mathfrak{g})} \sum_{\bar{e}\sim e, \bar{e}\neq e} |\nabla_{\bar{e}} f| + 2(d-1)N|\beta| \sqrt{d(\mathfrak{g})} |\nabla_e f|.\end{aligned}$$

For I_2 and I_3, fixing the edge e we recall our choice of the orthonormal basis $\{v^i_e\}_{1\leqslant i\leqslant d(\mathfrak{g})}$ above. Using Lemma 4.9 we then have

$$\begin{aligned}|I_2 + I_3| &\leqslant \frac{3}{2} \Big(\sum_j |v^j_e \mathcal{S}|^2 \Big)^{1/2} \Big(\sum_j |[v^i_e, v^j_e] f|^2 \Big)^{1/2} + \frac{1}{2} \Big(\sum_j |v^j_e f|^2 \Big)^{1/2} \Big(\sum_j |[v^i_e, v^j_e] \mathcal{S}|^2 \Big)^{1/2} \\ &\leqslant \sqrt{2}\gamma_1 \Big(\sum_j |v^j_e \mathcal{S}|^2 \Big)^{1/2} \Big(\sum_j |v^j_e f|^2 \Big)^{1/2} = \sqrt{2}\gamma_1 |\nabla_e \mathcal{S}| |\nabla_e f| \\ &\leqslant 2\sqrt{2}(d-1)N^{3/2}\gamma|\beta||\nabla_e f|,\end{aligned}$$

where $\gamma_1 = 1$ for $G = SO(N)$ and $\gamma_1 = 3$ for $G = SU(N)$ and we use (3.4) to bound $|\nabla_e \mathcal{S}|$ by $2(d-1)N^{3/2}|\beta|\gamma/\gamma_1$ in the last inequality. Hence, the result follows. $\square$

The next corollary together with uniqueness in Section 5 proves Corollary 1.6.

Corollary 4.11. *Suppose that Assumption 1.1 holds. For $f, g \in C^\infty_{cyl}(\mathcal{Q})$, suppose that $\Lambda_f \cap \Lambda_g = \varnothing$. Then one has*

$$|\mathrm{Cov}(f,g)| \leqslant c_1 d(\mathfrak{g}) e^{-c_N d(\Lambda_f,\Lambda_g)}(|\!|\!|f|\!|\!|_\infty |\!|\!|g|\!|\!|_\infty + \|f\|_{L^2}\|g\|_{L^2}),$$

where c_1 depends on $|\Lambda_f|$, $|\Lambda_g|$, and c_N depends on $K_{\mathcal{S}}$, N and d. Here the covariance and L^2 are with respect to every tight limit of $\{\mu_{\Lambda_L,N,\beta}\}_L$.

Proof. With the calculations and bounds obtained in the previous lemmas, together with our Poincaré inequality, to prove exponential decay we can then apply an argument essentially from [GZ03, Section 8.3]. We write $\mu_L = \mu_{\Lambda_L,N,\beta}$ for simplicity and consider

$$\begin{aligned}|\mathrm{Cov}_{\mu_L}(f,g)| &= |\mu_L(fg) - \mu_L(f)\mu_L(g)| = |\mu_L(P^L_t(fg)) - \mu_L(P^L_t f)\mu_L(P^L_t g)| \\ &= |\mu_L(P^L_t(fg) - P^L_t f P^L_t g) + \mathrm{Cov}_{\mu_L}(P^L_t f, P^L_t g)| \\ &\leqslant |\mu_L(P^L_t(fg) - P^L_t f P^L_t g)| + \mathrm{Var}_{\mu_L}(P^L_t f)^{1/2}\mathrm{Var}_{\mu_L}(P^L_t g)^{1/2}. \end{aligned} \tag{4.26}$$

Recall that the Poincaré inequality is equivalent to the following: $\mathrm{Var}(P^L_t f) \leqslant e^{-2tK_{\mathcal{S}}}\|f\|^2_{L^2(\mu_L)}$ (see Remark 4.6). Therefore by the Poincaré inequality, the last term in (4.26) is bounded by

$$\mathrm{Var}_{\mu_L}(P^L_t f)^{1/2}\mathrm{Var}_{\mu_L}(P^L_t g)^{1/2} \leqslant e^{-2tK_{\mathcal{S}}}\|f\|_{L^2(\mu_L)}\|g\|_{L^2(\mu_L)}. \tag{4.27}$$

As $\mathcal{L}_L$ is uniform elliptic operator with smooth coefficient, by Hörmander's Theorem (c.f. [Nua06, Theorem 2.3.3]) $P^L_t f \in C^\infty(\mathcal{Q}_L)$. Now we consider $P^L_t(fg) - P^L_t f P^L_t g$ in (4.26) and we omit L for notation simplicity. Recall that P_t and $\mathcal{L}$ commute on the domain $D(\mathcal{L})$ (see e.g. [MR92, Chap. I Exercise 1.9]). We have

$$\begin{aligned}P_t(fg) - P_t f P_t g &= \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}[P_s(P_{t-s}fP_{t-s}g)]\mathrm{d}s \\ &= \int_0^t \Big[P_s\mathcal{L}(P_{t-s}fP_{t-s}g) - P_s(\mathcal{L}P_{t-s}fP_{t-s}g + P_{t-s}f\mathcal{L}P_{t-s}g)\Big]\mathrm{d}s \\ &= 2\sum_e \int_0^t P_s\langle\nabla_e P_{t-s}f, \nabla_e P_{t-s}g\rangle\mathrm{d}s = 2\sum_{e,i}\int_0^t P_s[(v^i_e P_{t-s}f)\cdot(v^i_e P_{t-s}g)]\mathrm{d}s.\end{aligned}$$

Here, to obtain the third line from the second line, recalling the definition of $\mathcal{L}$, by $\nabla(fg) = g\nabla f + f\nabla g$ the first order terms cancel, and it then follows from $\Delta(fg) = g\Delta f + f\Delta g + 2\langle\nabla f, \nabla g\rangle$. Note that for every e, i, we have $(P_s v^i_e f)\cdot(P_s v^i_e g) = 0$ since $\Lambda_f \cap \Lambda_g = \varnothing$. From this we then have

$$\begin{aligned}\sum_{e,i}(v^i_e P_{t-s}f)(v^i_e P_{t-s}g) &= \sum_{e,i}(v^i_e P_{t-s}f - P_{t-s}v^i_e f)\cdot(v^i_e P_{t-s}g - P_{t-s}v^i_e g) \\ &\quad + \sum_{e,i}(v^i_e P_{t-s}f - P_{t-s}v^i_e f)\cdot(P_{t-s}v^i_e g) \\ &\quad + \sum_{e,i}(v^i_e P_{t-s}g - P_{t-s}v^i_e g)\cdot(P_{t-s}v^i_e f) \stackrel{\text{def}}{=} \sum_e(I^1_e + I^2_e + I^3_e).\end{aligned}$$

Suppose for the moment that we can prove the following: for any $c > 0$ and $f \in C^\infty_{cyl}(\mathcal{Q}_L)$, there exists $B > 0$ such that for $d(e, \Lambda_f) \geqslant Bt$ one has

$$\sum_i \|v^i_e P_t f - P_t v^i_e f\|_{L^\infty} \leqslant d(\mathfrak{g}) e^{-2cd(e,\Lambda_f)} |\!|\!| f |\!|\!|_\infty. \tag{4.28}$$

We choose $t \sim d(\Lambda_f, \Lambda_g)/B$ below. Applying (4.28) to the function g with $e \in \Lambda_f$ (in which case $I^2_e = 0$ since $v^i_e g = 0$) and using (4.9)

$$\begin{aligned}\|I^1_e + I^3_e\|_{L^\infty} &\leqslant \sum_i \|v^i_e P_{t-s} f\|_{L^\infty} \|v^i_e P_{t-s} g - P_{t-s} v^i_e g\|_{L^\infty} \\ &\leqslant d(\mathfrak{g}) e^{-2cd(\Lambda_f,\Lambda_g)} |\!|\!| f |\!|\!|_\infty |\!|\!| g |\!|\!|_\infty.\end{aligned}$$

Similarly for $e \in \Lambda_g$, $I^3_e = 0$ and

$$\|I^1_e + I^2_e\|_{L^\infty} \leqslant d(\mathfrak{g}) e^{-2cd(\Lambda_g,\Lambda_f)} |\!|\!| g |\!|\!|_\infty |\!|\!| f |\!|\!|_\infty.$$

For $e \notin \Lambda_f \cup \Lambda_g$ we have $I^2_e = I^3_e = 0$ and $d(e, \Lambda_f) \geqslant d(\Lambda_g, \Lambda_f)/2$ or $d(e, \Lambda_g) \geqslant d(\Lambda_g, \Lambda_f)/2$. For both cases we have

$$\|I^1_e\|_{L^\infty} \leqslant d(\mathfrak{g}) e^{-cd(\Lambda_f,\Lambda_g) - c(d(e,\Lambda_f)\wedge d(e,\Lambda_g))} |\!|\!| f |\!|\!|_\infty |\!|\!| g |\!|\!|_\infty.$$

With these bounds on I^1_e, I^2_e, I^3_e, we sum over e and obtain that for $d(\Lambda_f, \Lambda_g) \geqslant Bt$

$$\|P_t(fg) - P_t f P_t g\|_{L^\infty} \leqslant c_1 d(\mathfrak{g}) e^{-cd(\Lambda_f,\Lambda_g)} |\!|\!| f |\!|\!|_\infty |\!|\!| g |\!|\!|_\infty. \tag{4.29}$$

Substituting (4.27) and (4.29) into (4.26) we get

$$|\mathrm{Cov}_{\mu_L}(f, g)| \leqslant c_1 d(\mathfrak{g}) e^{-cd(\Lambda_f,\Lambda_g)} |\!|\!| f |\!|\!|_\infty |\!|\!| g |\!|\!|_\infty + e^{-2tK_S} \|f\|_{L^2} \|g\|_{L^2},$$

where c_1 depends on $|\Lambda_f|$ and $|\Lambda_g|$ and is independent of L. Since $t \sim d(\Lambda_f, \Lambda_g)/B$, letting $L \to \infty$ the result follows.

It remains to check the claimed bound (4.28). We use a similar argument as in [GZ03, Theorem 8.2] which we adapt into our setting. We have

$$v^i_e P_t f - P_t v^i_e f = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} \left(P_{t-s} v^i_e P_s f\right) \mathrm{d}s = \int_0^t P_{t-s}[v^i_e, \mathcal{L}] P_s f \mathrm{d}s. \tag{4.30}$$

By Lemma 4.10, we have

$$\|[v^i_e, \mathcal{L}] P_s f\|_{L^\infty} \leqslant \sum_{\bar{e} \sim e} a_{e,\bar{e}} \|\nabla_{\bar{e}} P_s f\|_{L^\infty},$$

for constants $a_{e,\bar{e}}$ which are uniformly bounded in $e, \bar{e}$. Hence, by (4.30)

$$\sum_i \|v^i_e P_t f\|_{L^\infty} \leqslant \sum_i \|v^i_e f\|_{L^\infty} + \int_0^t \sum_{\bar{e}} D_{e,\bar{e}} \|\nabla_{\bar{e}} P_s f\|_{L^\infty} \mathrm{d}s,$$

with a matrix D such that $D_{e,\bar{e}} = d(\mathfrak{g}) a_{e,\bar{e}}$ if $e \sim \bar{e}$ and $D_{e,\bar{e}} = 0$ otherwise. Since $e \notin \Lambda_f$ we get $v^i_e f = 0$ and by iteration

$$\sum_i \|v^i_e P_t f\|_{L^\infty} \leqslant \sum_{n=N_e}^\infty \frac{t^n}{n!} \sum_{\bar{e}} D^{(n)}_{e,\bar{e}} \sum_i \|v^i_{\bar{e}} f\|_{L^\infty},$$

with $N_e = d(e, \Lambda_f)$ and $D^{(n)}_{e,\bar{e}} \leqslant C_0^n$ with $C_0 = d(\mathfrak{g})(a_{e,e} + 6(d-1)a_{e,\bar{e}})$. As a result, using $n! \geqslant e^{n \log n - 2n}$, for $2 - \log B + \log C_0 + \frac{C_0}{B} \leqslant -2c$ and $d(e, \Lambda_f) \geqslant Bt$ we have

$$\sum_i \|v_e^i P_t f\|_{L^\infty} \leqslant \sum_{n=N_e}^{\infty} \frac{t^n}{n!} C_0^n d(\mathfrak{g}) \|\!|f|\!\|_\infty \leqslant \frac{(C_0 t)^{N_e}}{N_e!} e^{tC_0} d(\mathfrak{g}) \|\!|f|\!\|_\infty \leqslant d(\mathfrak{g}) e^{-2cd(e,\Lambda_f)} \|\!|f|\!\|_\infty.$$

Hence, (4.28) follows. □

Remark 4.12. *From the above proof one can see*

$$c_N \sim \frac{K_{\mathcal{S}}}{d(\mathfrak{g})(a_{e,e} + 6(d-1)a_{e,\bar{e}})},$$

but this is not necessarily optimal.

5. Uniqueness of invariant measure

In this section we prove Theorem 1.2. As the results (4.5) and (4.6) in Theorem 4.2 depend on ρ_L, we cannot simply send $L \to \infty$ to conclude the result for $(P_t)_{t\geqslant 0}$ on $\mathcal{Q}$. The idea of our proof is to construct a suitable coupling and find a suitable distance $\rho_{\infty,a}$ such that for any $\mu, \nu \in \mathscr{P}(\mathcal{Q}_L)$, the Wasserstein distance w.r.t. $\rho_{\infty,a}$ between μP_t^L and νP_t^L decays exponentially fast in time. Recall that $\rho_{\infty,a}$ is given in (1.4) and we will choose a suitable parameter $a > 1$ below.

We denote $C_{\mathrm{Ric},N} = \frac{\alpha(N+2)}{4} - 1$ which is a constant arising from Ricci curvature in (4.8), where $\alpha = 1, 2$ for $SO(N)$ and $SU(N)$ respectively. For any $\mu, \nu \in \mathscr{P}(\mathcal{Q})$, we introduce the Wasserstein distance

$$W_p^{\rho_{\infty,a}}(\mu, \nu) \stackrel{\text{def}}{=} \inf_{\pi \in \mathscr{C}(\mu,\nu)} \pi(\rho_{\infty,a}^p)^{1/p}.$$

Recall that the generator $\mathcal{L}_L$ is given by

$$\mathcal{L}_L F = \sum_{e \in E^+_{\Lambda_L}} \Delta_e F + \sum_{e \in E^+_{\Lambda_L}} \langle \nabla \mathcal{S}(Q)_e, \nabla_e F \rangle. \tag{5.1}$$

For fixed $Q \in \mathcal{Q}_L$ define

$$C \stackrel{\text{def}}{=} \{(Q, Q') : Q' \in \mathrm{cut}(Q)\}, \quad D \stackrel{\text{def}}{=} \{(Q, Q) : Q \in \mathcal{Q}_L\}, \tag{5.2}$$

where $\mathrm{cut}(Q)$ consists of conjugate points of Q and points having more than one minimal geodesics to Q.

In the following we prove the result for any $a > 1$.

Lemma 5.1. *Suppose that* $\widetilde{K}_{\mathcal{S}} \stackrel{def}{=} C_{\mathrm{Ric},N} - (4 + 4\sqrt{a})N|\beta|(d-1) > 0$. *Then for every* $L \in \mathbb{Z}$,

$$W_2^{\rho_{\infty,a}}(\mu P_t^L, \nu P_t^L) \leqslant e^{-\widetilde{K}_{\mathcal{S}} t} W_2^{\rho_{\infty,a}}(\mu, \nu), \qquad t \geqslant 0, \quad \mu, \nu \in \mathscr{P}(\mathcal{Q}_L).$$

Here we use periodic extension to view every measure as a probability on $\mathcal{Q}$.

Proof. To prove the statement we will construct a suitable coupling $(Q(t), Q'(t))_{t\geqslant 0}$ between the two Markov processes associated to the generator $\mathcal{L}_L$ starting from two different points (Q, Q'). We will then use Itô's fomula to calculate $\mathrm{d}\rho^2_{\infty,a}(Q(t), Q'(t))$ and obtain

$$\rho^2_{\infty,a}(Q(t), Q'(t)) \leqslant e^{-2\tilde{K}_{\mathcal{S}} t} \rho^2_{\infty,a}(Q(0), Q'(0)), \quad t \geqslant 0. \tag{5.3}$$

Suppose that (5.3) holds and we use $\mathbb{P}_t^{Q,Q'}$ to denote the distribution of the coupling $(Q(t), Q'(t))$. Then for any $\mu, \nu \in \mathscr{P}(\mathcal{Q}_L)$ and $\pi \in \mathscr{C}(\mu,\nu)$ we set

$$\pi_t \stackrel{\text{def}}{=} \int \mathbb{P}_t^{Q,Q'} \pi(\mathrm{d}Q, \mathrm{d}Q') \in \mathscr{C}(\mu P_t^L, \nu P_t^L).$$

Hence, for $t \geqslant 0$

$$W_2^{\rho_{\infty,a}}(\mu P_t, \nu P_t)^2 \leqslant \int \rho_{\infty,a}^2 \mathrm{d}\pi_t \leqslant e^{-2\widetilde{K}_{\mathcal{S}} t} \pi(\rho_{\infty,a}^2),$$

and the result follows. In the following we prove (5.3) in three steps.

STEP 1. Construction of coupling $(Q(t), Q'(t))_{t\geqslant 0}$ and calculation of $\mathrm{d}\rho^2(Q_e(t), Q'_e(t))$.

The usual coupling for Brownian motions and diffusions on Riemannian manifolds is the Kendall–Cranston's coupling (c.f. [Ken86]). In our case we adapt a construction in [Wan06, Proposition 2.5.1] to cancel the noise part, with one of the key modifications due to our new weighted distance on our product manifold.

More precisely, let $(Q(t), Q'(t))$ be the coupling on $\mathcal{Q}_L \times \mathcal{Q}_L$ starting from (Q, Q') given by the following generator

$$\mathcal{L}^c = \sum_{e \in E^+_{\Lambda_L}} \Delta_{Q_e} + \sum_{e \in E^+_{\Lambda_L}} \Delta_{Q'_e} + 2 \sum_{i,j=1}^{\dim \mathfrak{q}_L} \Big\langle P_{Q,Q'} v_i, v'_j \Big\rangle_{T_{Q'}\mathcal{Q}_L} v_i v'_j + \nabla \mathcal{S}(Q) + \nabla \mathcal{S}(Q'), \qquad (5.4)$$

where $\Delta_{Q_e} f(Q, Q') = (\Delta_e f(\cdot, Q'))(Q)$, $\Delta_{Q'_e} f(Q, Q') = (\Delta_e f(Q, \cdot))(Q')$ and $\{v_i\}, \{v'_j\}$ are orthonormal bases of tangent spaces at Q and Q', and $P_{Q,Q'} : T_Q\mathcal{Q}_L \to T_{Q'}\mathcal{Q}_L$ is the parallel translation along the geodesic from Q to Q'. It is easy to see that $\mathcal{L}^c$ is independent of the choices of the basis $\{v_i\}, \{v'_j\}$. In fact, to construct such coupling we need to avoid the cut locus C and the diagonal set D by suitable cut-off approximation and we refer to Appendix A and [Wan06, Section 2.1] for more details on the construction.

We intend to apply Itô's formula to ρ_e^2 with $\rho_e \stackrel{\text{def}}{=} \rho(Q_e, Q'_e)$. To this end, we consider the projection map $\pi_e : \mathcal{Q}_L \to G$ defined by $\pi_e Q \stackrel{\text{def}}{=} Q_e$. We then write $\hat{\rho}_e(\,\cdot\,; Q'_e)$ for the pull-back of the function $\rho(\cdot, Q'_e)$ via the map π_e. Namely, fixing any $Q'_e \in G$, the function $\hat{\rho}_e(\,\cdot\,; Q'_e)$ is a function on $\mathcal{Q}_L$ defined by

$$\hat{\rho}_e(Q; Q'_e) \stackrel{\text{def}}{=} \rho(\pi_e Q, Q'_e) = \rho(Q_e, Q'_e) \qquad \text{for } Q \in \mathcal{Q}_L.$$

Similarly we define function $\hat{\rho}_e(Q_e;\,\cdot\,)$ on $\mathcal{Q}_L$ as

$$\hat{\rho}_e(Q_e; Q') \stackrel{\text{def}}{=} \rho(Q_e, \pi_e Q') = \rho(Q_e, Q'_e) \qquad \text{for } Q' \in \mathcal{Q}_L.$$

We can also write $\rho_e = \rho(\pi_e Q, \pi_e Q')$ and view ρ_e as a function on $\mathcal{Q}_L \times \mathcal{Q}_L$.

For $R \in \mathbb{N}$, we choose a smooth cut-off function $\chi_R : [0, \infty) \to [0, \infty)$ satisfying $\chi_R(x) = x$ for $x \geqslant 1/R$ and $\chi_R|_{[0, \frac{1}{2R}]} = 0$ and $\chi'_R \geqslant 0$.

Since ρ_e^2 is smooth near the diagonal, we claim that by Itô's formula (see [Wan06, Section 2.1], [Hsu02, Section 6.5]), and writing $\rho_e(t) = \rho(Q_e(t), Q'_e(t))$, we have

$$\mathrm{d}\chi_R(\rho_e^2(t)) \leqslant 2\chi'_R(\rho_e^2(t))\rho_e(t) J(Q(t), Q'(t))\mathrm{d}t \qquad (5.5)$$

for $t < T \stackrel{\text{def}}{=} \inf\{t \geqslant 0, Q(t) = Q'(t)\}$ where J is a continuous function on $\mathcal{Q}_L \times \mathcal{Q}_L$ such that $J \geqslant I_{\mathcal{S}}$ on $(D \cup C)^c$. Here

$$I_{\mathcal{S}}(Q, Q') \stackrel{\text{def}}{=} I(Q_e, Q'_e) + \Big((\nabla \mathcal{S})\hat{\rho}_e(\,\cdot\,; Q'_e)\Big)(Q) + \Big((\nabla \mathcal{S})\hat{\rho}_e(Q_e;\,\cdot\,)\Big)(Q')\,, \qquad (5.6)$$

and $I(x,y)$ is the index along $\gamma : [0,\rho(x,y)] \to G$ which is the minimal geodesic from x to y in G:

$$I(x,y) \stackrel{\text{def}}{=} \sum_{i=1}^{\dim\mathfrak{g}-1} \int_0^{\rho(x,y)} \Big(|\nabla_{\dot\gamma} J_i|^2 - \langle \mathscr{R}(J_i,\dot\gamma)\dot\gamma, J_i\rangle \Big)_s \mathrm{d}s,$$

where $\{J_i\}_{i=1}^{\dim\mathfrak{g}-1}$ are Jacobi fields along γ such that at x and y, they, together with $\dot\gamma$, form an orthonormal basis. Note that the reason to derive a bound in terms of J instead of $I_{\mathcal{S}}$ in (5.5) is that J is defined everywhere on $\mathcal{Q}_L \times \mathcal{Q}_L$ whereas $I_{\mathcal{S}}$ is not well-defined on $C \cup D$. In Step 2 below we control $I_{\mathcal{S}}$ by a continuous function on $\mathcal{Q}_L \times \mathcal{Q}_L$, which can also control J.

The rigorous derivation of (5.5) follows by cut-off approximation to avoid the cut locus C and the diagonal set D (c.f. [Wan06, Theorem 2.1.1], [Hsu02, Theorem 6.6.2]). In the following we give the idea on how the terms in (5.6) arise and we put more details of the construction and derivation of (5.5) in Appendix A.

For $t < T$ and $(Q(t),Q'(t)) \notin C \cup D$, on the support of $\chi_R(\rho_e^2)$, $I_{\mathcal{S}}(Q,Q')$ is given by $\mathcal{L}^c\rho_e$. To prove this, since $\rho_e = \rho(\pi_e Q, \pi_e Q') = \rho(Q_e, Q'_e)$ only depends on Q_e, Q'_e (i.e. independent of the values of Q, Q' on the other edges), we can write the first three terms in $\mathcal{L}^c\rho_e$ as (see the RHS of (5.4))

$$\Delta_{Q_e}\rho(Q_e,Q'_e) + \Delta_{Q'_e}\rho(Q_e,Q'_e) + 2\sum_{i,j=1}^{d(\mathfrak{g})} \Big\langle P_{Q_e,Q'_e} v_{e,i}\ ,\ v'_{e,j} \Big\rangle_{T_{Q'_e}G} v_{e,i} v'_{e,j}\, \rho(Q_e,Q'_e), \tag{5.7}$$

with $\{v_{e,i}\}$ and $\{v'_{e,j}\}$ being an orthonormal basis of the tangent space at Q_e, Q'_e and $P_{Q_e,Q'_e} : T_{Q_e}G \to T_{Q'_e}G$ being the parallel translation along the geodesic from Q_e to Q'_e. Here we used the fact that

$$P_{Q_e,Q'_e}(v_e) = (P_{Q,Q'}v)_e \qquad \forall v \in T_Q\mathcal{Q}_L$$

(in particular the e component of the geodesic from Q to Q' is the geodesic from Q_e to Q'_e.) By the second variational formula (c.f. [CE75, p21-22], [Ken86, Theorem 2], [Hsu02, Lemma 6.6.1]) we know that (5.7) is equal to $I(Q_e, Q'_e)$.

Moreover, the last two terms involving $\nabla\mathcal{S}$ in $\mathcal{L}^c$ give rise to the last two terms in (5.6).

The quadratic variation of the martingale part from applying Itô's formula to ρ_e is

$$|\langle\dot\gamma,\dot\gamma\rangle(Q'_e) - \langle\dot\gamma,\dot\gamma\rangle(Q_e)|^2 + \sum_{i=1}^{d(\mathfrak{g})-1} |\langle J_i,\dot\gamma\rangle(Q'_e) - \langle J_i,\dot\gamma\rangle(Q_e)|^2 \tag{5.8}$$

by the first variation formula (c.f. [CE75, p5], [Hsu02, Section 6.6]). Since $\{J_i\}_{i=1}^{\dim\mathfrak{g}-1}$ together with $\dot\gamma$ form an orthonormal basis, each term in (5.8) is zero, which implies that the martingale part is zero. We also refer to the derivation of (A.5) in Appendix A for more details on the calculation of the quadratic variation.

STEP 2. Estimate the RHS of (5.6).

In this step we estimate the RHS of (5.6) and prove that for $t < T$

$$\partial_t\rho_e^2 \leqslant -2C_{\mathrm{Ric},N}\rho_e^2 + 2N|\beta| \sum_{p,p\succ e} \rho_e\Big(\rho_e + \sum_{e\neq\bar e\in p} \rho_{\bar e}\Big). \tag{5.9}$$

By the index lemma (see [Wan06, Theorem 2.1.4] or [Hsu02, Lemma 6.7.1]), for $x = Q_e$, $y = Q'_e$ with $\gamma : [0,\rho_e] \to G$ the minimal geodesic from Q_e to Q'_e, where we recall that

$\rho_e = \rho(Q_e, Q'_e)$, we have

$$I(Q_e, Q'_e) \leqslant -\int_0^{\rho_e} \mathrm{Ric}(\dot\gamma, \dot\gamma)\mathrm{d}s = -C_{\mathrm{Ric},N}\rho_e. \tag{5.10}$$

In the following we consider the last two terms in (5.6). Given $Q, Q' \in \mathcal{Q}_L$ as above, we define a path $\Gamma : [0, \rho_e] \to \mathcal{Q}_L$ which goes from Q to Q' as follows. For any $\bar e \in E^+_{\Lambda_L}$, we can find a geodesic $\gamma^{\bar e} : [0, \rho_{\bar e}] \to G_{\bar e}$ from $Q_{\bar e}$ to $Q'_{\bar e}$. Here $\rho_{\bar e}$ is the length of the geodesic. We then set

$$\Gamma(s) = \Big(\tilde\gamma^{\bar e}(s)\Big)_{\bar e \in E^+_{\Lambda_L}} \in \mathcal{Q}_L \qquad (s \in [0, \rho_e]) \qquad \text{where} \qquad \tilde\gamma^{\bar e}(s) = \gamma^{\bar e}(\rho_{\bar e} s/\rho_e).$$

We can check that we indeed have $\Gamma(0) = Q$ and $\Gamma(\rho) = Q'$ and Γ is the geodesic from Q to Q'. Also, we have

$$\pi_e(\Gamma(s)) = \gamma(s) \qquad (\forall s \in [0, \rho_e]).$$

With the above notation at hand, we write the last two terms in (5.6) as

$$\Big((\nabla\mathcal{S})\hat\rho_e(\,\cdot\,; Q'_e)\Big)(Q) + \Big((\nabla\mathcal{S})\hat\rho_e(Q_e;\,\cdot\,)\Big)(Q') = \langle \nabla\mathcal{S}, \dot\gamma\rangle(Q') - \langle \nabla\mathcal{S}, \dot\gamma\rangle(Q), \tag{5.11}$$

with $\dot\gamma$ extended as a tangent vector field on $\mathcal{Q}_L$ along the curve Γ, which is still denoted by $\dot\gamma$, by setting all the other components as zero.

We then write (5.11) as

$$\int_0^{\rho_e} \Big(\frac{\mathrm{d}}{\mathrm{d}s}\Big\langle \nabla\mathcal{S}(\Gamma(s)), \dot\gamma(s)\Big\rangle\Big)\mathrm{d}s = \int_0^{\rho_e} \Big(\dot\Gamma\langle\nabla\mathcal{S}, \dot\gamma\rangle\Big)(\Gamma(s))\mathrm{d}s.$$

Hence, we get

$$(5.11) = \int_0^{\rho_e} \Big(\dot\Gamma\langle\nabla\mathcal{S}, \dot\gamma\rangle\Big)(\Gamma(s))\mathrm{d}s.$$

Below we estimate the above integral. With a slight abuse of notation, for an edge $e \in E^+_{\Lambda_L}$ we write $e \in p$ if $\{e, e^{-1}\} \cap p \neq \varnothing$, namely we view edges as undirected in the calculation below. We also extend $\dot{\tilde\gamma}^{\bar e}, \bar e \in E^+_{\Lambda_L}$ as tangent vector field on $\mathcal{Q}_L$ along Γ, which is still denoted by $\dot{\tilde\gamma}^{\bar e}$, by setting all the components other than $\bar e$ to be zero. Then recalling our formula for $\mathcal{S}$ we have

$$\begin{aligned}
\int_0^{\rho_e} \Big(\dot\Gamma\langle\nabla\mathcal{S}, \dot\gamma\rangle\Big)(\Gamma(s))\mathrm{d}s &= N\beta \sum_{p\succ e}\sum_{\bar e\in p}\int_0^{\rho_e} \dot{\tilde\gamma}^{\bar e}(\dot\gamma \mathrm{ReTr}(Q_p))\mathrm{d}s \\
&\leqslant N|\beta| \sum_{p\succ e}\sum_{\bar e\in p}\int_0^{\rho_e} |\dot{\tilde\gamma}^{\bar e}||\dot\gamma|\,\mathrm{d}s \\
&\leqslant N|\beta| \sum_{p\succ e}\Big(\rho_e + \sum_{e\neq\bar e\in p}\rho_{\bar e}\Big),
\end{aligned} \tag{5.12}$$

where we used $|\dot\gamma| = 1$ and $|\dot{\tilde\gamma}^{\bar e}| = \rho_{\bar e}/\rho_e$. Here we calculate $\dot{\tilde\gamma}^{\bar e}(\dot\gamma \mathrm{ReTr}(Q_p))$ as follows: for $Q_p = Q_e Q_{\bar e} Q_1 Q_2$ with $Q_1, Q_2 \in G$ we get

$$\begin{aligned}
\dot{\tilde\gamma}^{\bar e}(\dot\gamma \mathrm{ReTr}(Q_p)) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\mathrm{ReTr}\Big(\gamma(s)\tilde\gamma^{\bar e}(t)Q_1Q_2\Big) \\
&= \mathrm{ReTr}\Big(\dot\gamma\dot{\tilde\gamma}^{\bar e}Q_1Q_2\Big),
\end{aligned}$$

the absolute value of which by Hölder's inequality for trace is bounded by $|\dot{\tilde\gamma}^{\bar e}|\,|\dot\gamma|$. Similar calculation holds for $Q_p = Q_e Q_1 Q_{\bar e} Q_2$ and $Q_p = Q_e Q_1 Q_2 Q_{\bar e}$ and we use similar argument as in the proof of Lemma 4.1 to control $\dot{\gamma}\dot{\gamma}\mathrm{ReTr}(Q_p)$ by $|\dot{\gamma}|^2$. Hence, by (5.5), (5.6), (5.10), (5.12), we get

$$\partial_t \chi_R(\rho_e^2) \leqslant -2C_{\mathrm{Ric},N}\,\chi_R'(\rho_e^2)\,\rho_e^2 + 2\chi_R'(\rho_e^2)N|\beta|\sum_{p,p\succ e}\Big(\rho_e^2 + \sum_{e\neq\bar{e}\in p}\rho_{\bar{e}}\rho_e\Big).$$

Letting $R\to\infty$ and by dominated convergence theorem and the fact that χ_R' is uniformly bounded in R, (5.9) holds.

Step 3. Derivation of (5.3).

We extend $(Q(t),Q'(t))_{t\geqslant 0}$ periodically as a process on $\mathcal{Q}\times\mathcal{Q}$, which is still denoted as $(Q(t),Q'(t))_{t\geqslant 0}$. (5.9) also holds for the extension. By (5.9) we have

$$\partial_t\rho_e^2 \leqslant -2C_{\mathrm{Ric},N}\,\rho_e^2 + 2N|\beta|\Big(2(d-1)\rho_e^2 + \sum_{p,p\succ e}\sum_{e\neq\bar{e}\in p}\rho_{\bar{e}}\rho_e\Big).$$

In the following we bound $\rho_{\bar{e}}\rho_e$. To obtain the desired rate given by $\widetilde{K}_{\mathcal{S}}t$, we need to control $\rho_{\bar{e}}\rho_e$ in different ways depending on the relations between $|e|$ and $|\bar{e}|$. We first fix a plaquette p and consider two edges $\bar{e}\neq e$.

For the edges satisfying $|e|=|\bar{e}|$ [2] we have

$$2\rho_{\bar{e}}\rho_e \leqslant \rho_{\bar{e}}^2+\rho_e^2.$$

For the edges satisfying $|e|\neq|\bar{e}|$ we have

$$\frac{2}{\sqrt{a}}\rho_{\bar{e}}\rho_e \leqslant \frac{1}{a}\rho_{\bar{e}}^2+\rho_e^2.$$

The reason for the choice of the above weight is as follows: there is one plaquette p such that only one edge $\bar{e}\neq e$ in p with the same distance as $|e|$ and other edges with the distance larger than $|e|$. Thus, since for each edge e there are $2(d-1)$ plaquettes in $\mathcal{P}$ such that $p\succ e$, we get

$$\begin{aligned}
&2N|\beta|\sum_{p,p\succ e}\sum_{e\neq\bar{e}\in p}\rho_{\bar{e}}\rho_e\\
&\leqslant \sqrt{a}N|\beta|\sum_{p,p\succ e}\sum_{|e|\neq|\bar{e}|\in p}\Big(\frac{1}{a}\rho_{\bar{e}}^2+\rho_e^2\Big) + N|\beta|\sum_{p,p\succ e}\sum_{|e|=|\bar{e}|\in p,e\neq\bar{e}}\Big(\rho_{\bar{e}}^2+\rho_e^2\Big)\\
&= \sqrt{a}N|\beta|\sum_{p,p\succ e}\sum_{|e|\neq|\bar{e}|\in p}\frac{1}{a}\rho_{\bar{e}}^2 + (4\sqrt{a}+2)(d-1)N|\beta|\rho_e^2 + N|\beta|\sum_{p,p\succ e}\sum_{|e|=|\bar{e}|\in p,e\neq\bar{e}}\rho_{\bar{e}}^2
\end{aligned}$$

where the first sum for $\bar{e}$ with $|e|\neq|\bar{e}|$ includes two edges and the second sum for $\bar{e}$ with $|e|=|\bar{e}|$ contains only one edge. Note that we also get an extra $\frac{1}{a}$ before $\rho_{\bar{e}}^2$ with $|\bar{e}|=|e|+1$, which can be put into the weight $\frac{1}{a^{|\bar{e}|}}$. Substituting the above calculation into (5.9) and using again the fact that for each edge e there are $2(d-1)$ plaquettes in $\mathcal{P}$ such that $p\succ e$ we get

$$\begin{aligned}
\frac{1}{a^{|e|}}\partial_t\rho_e^2 \leqslant &-2C_{\mathrm{Ric},N}\frac{1}{a^{|e|}}\rho_e^2 + (4\sqrt{a}+6)N|\beta|(d-1)\frac{1}{a^{|e|}}\rho_e^2\\
&+N|\beta|\sum_{p,p\succ e}\Big(\sqrt{a}\sum_{|e|\neq|\bar{e}|\in p}\frac{1}{a^{|\bar{e}|}}\rho_{\bar{e}}^2 + \sum_{|e|=|\bar{e}|\in p,e\neq\bar{e}}\rho_{\bar{e}}^2\Big).
\end{aligned}$$

[2] If two edges $e\neq\bar{e}$ share the same vertex and this vertex is closer to origin, we may have $|e|=|\bar{e}|$.

Taking sum over e we notice that ρ_e^2 also appears when calculating $\frac{1}{a^{|\bar e|}}\partial_t\rho_{\bar e}^2$ with $\bar e$ and e in the same plaquette, which at most gives $2\sqrt{a}N|\beta|\frac{1}{a^{|e|}}\rho_e^2$ and $N|\beta|\frac{1}{a^{|e|}}\rho_e^2$ from $\frac{1}{a^{|\bar e|}}\partial_t\rho_{\bar e}^2$ with $|\bar e|\neq|e|$ and $|\bar e|=|e|$, respectively. Since for each edge e there are $2(d-1)$ plaquettes in $\mathcal{P}$ such that $p\succ e$, we get

$$\sum_{e\in E^+}\frac{1}{a^{|e|}}\partial_t\rho_e^2\leqslant -2C_{\mathrm{Ric},N}\sum_{e\in E^+}\frac{1}{a^{|e|}}\rho_e^2+(8+8\sqrt{a})N|\beta|(d-1)\sum_{e\in E^+}\frac{1}{a^{|e|}}\rho_e^2. \tag{5.13}$$

Hence, (5.3) follows from Gronwall's lemma. □

Remark 5.2. *In general, even if $\tilde K_S\leqslant 0$, Lemma 5.1 still holds, but in that case the bound would not be useful for us.*

Now we prove Theorem 1.2. One of the important ingredients in the proof is that under Assumption 1.1, the condition of Lemma 5.1 can *indeed* be satisfied by tuning the weight parameter $a>1$ to be sufficiently close to 1, see Eq. (5.15) below. The crucial reason for this proof to work is that the last term in our bound (5.13) is of order N, rather than N^p for some $p>1$. This is a nontrivial point: indeed, that term comes from bounding the $\nabla\mathcal{S}$ terms on the right-hand side of (5.6), but note that $\mathcal{S}$ defined in (1.2) would appear to be of order N^2 if one naively bound $\mathrm{Tr}(Q_p)\leqslant N$, in which case the proof would break down. In fact in the previous proof we instead apply the property of the Lie group G and Hölder inequality to separate different vector fields appearing in the second order derivative of $\mathcal{S}$, which could finally be bounded by sum of Riemannian distances up to a factor $N|\beta|$.

Proof of Theorem 1.2. For any two invariant measures μ,ν of (1.6), we can find two sequences $\{\mu_L\},\{\nu_L\}\subset\mathscr{P}(\mathcal{Q}_L)$ such that their periodic extensions over the entire $\mathcal{Q}$, which are still denoted by μ_L,ν_L, converge to μ,ν weakly in $\mathcal{Q}$, with the distance induced by $\|\cdot\|$ defined in (3.6). Indeed, let $Q(0):\Omega\to\mathcal{Q}$ be a random variable such that $\mathrm{Law}(Q(0))=\mu$, and then define $\mu_L\stackrel{\mathrm{def}}{=}\mathrm{Law}(Q^L(0))$ where $Q^L(0):\Omega\to\mathcal{Q}_L$ is given by

$$Q_e^L(0)=\begin{cases}Q_e(0) & e\in E^+_{\Lambda_{L-1}}\\ I_N\,, & e\in E^+_{\Lambda_L}\backslash E^+_{\Lambda_{L-1}}\,,\end{cases}$$

then $\{\mu_L\}$ satisfy the desired property. The sequence $\{\nu_L\}$ can be constructed in the same way.

By Lemma 3.2 we obtain the unique solution $Q^L\in C([0,\infty);\mathcal{Q}_L)$ to (3.3) starting from the initial distribution $\mu_L\in\mathscr{P}(\mathcal{Q}_L)$. By periodic extension we view $Q^L\in C([0,\infty);\mathcal{Q})$. Recall that $(P_t^L)_{t\geqslant0}$ is the Markov semigroup associated with the solution to (3.3). By global well-posedness of (3.3), we obtain for $F\in C^\infty_{cyl}(\mathcal{Q})$ and $t\geqslant0$

$$\int P_t^L F\mathrm{d}\mu_L=\mathbf{E}F(Q^L(t)).$$

Similarly, using Proposition 3.4 we obtain unique solutions $Q\in C([0,\infty);\mathcal{Q})$ to (1.6) starting from the initial distribution μ. Recall that $(P_t)_{t\geqslant0}$ is the Markov semigroup for the Markov process associated to (1.6). By uniqueness in law of the solution to (1.6) we have

$$\int P_tF\mathrm{d}\mu=\mathbf{E}F(Q(t)).$$

As μ_L converges to μ weakly in $\mathcal{Q}$, by the same argument as in the proof Proposition 3.4, the law of $\{Q^L\}$ is tight in $C([0,\infty);\mathcal{Q})$ and the tight limit satisfies the limit equation (1.6) with the initial distribution μ. By uniqueness in law of equation (1.6), which follows from pathwise uniqueness in Proposition 3.4 and Yamada–Watanabe Theorem, the law of Q^L converges weakly to the law of Q in $C([0,\infty);\mathcal{Q})$, as $L\to\infty$. As a result, for $F\in C^\infty_{cyl}(\mathcal{Q})$ we have

$$\int P^L_t F\mathrm{d}\mu_L=\mathbf{E}F(Q^L(t))\to\mathbf{E}F(Q(t))=\int P_tF\mathrm{d}\mu,\quad L\to\infty. \tag{5.14}$$

Similarly, we obtain

$$\int P^L_tF\mathrm{d}\nu_L\to\int P_tF\mathrm{d}\nu,\quad L\to\infty.$$

Moreover, by the condition $K_{\mathcal{S}}>0$, there exists $a>1$ such that

$$\widetilde{K}_{\mathcal{S}}=C_{\mathrm{Ric},N}-(4+4\sqrt{a})N|\beta|(d-1)>0. \tag{5.15}$$

We then invoke Lemma 5.1 to have

$$\begin{aligned}&\Big|\int F\mathrm{d}\mu-\int F\mathrm{d}\nu\Big|=\Big|\int P_tF\mathrm{d}\mu-\int P_tF\mathrm{d}\nu\Big|=\lim_{L\to\infty}\Big|\int P^L_tF\mathrm{d}\mu_L-\int P^L_tF\mathrm{d}\nu_L\Big|\\&=\lim_{L\to\infty}\inf_{\pi\in\mathscr{C}(\mu_LP^L_t,\nu_LP^L_t)}\Big|\int(F(x)-F(y))\mathrm{d}\pi(x,y)\Big|\leqslant C_F\lim_{L\to\infty}W_2^{\rho_{\infty,a}}(\mu_LP^L_t,\nu_LP^L_t)\\&\leqslant C_Fe^{-\widetilde{K}_{\mathcal{S}}t}\lim_{L\to\infty}W_2^{\rho_{\infty,a}}(\mu_L,\nu_L)\leqslant C(a)e^{-\widetilde{K}_{\mathcal{S}}t}\end{aligned}$$

where C_F only depends on F, and the constant $C(a)$ is independent of L by boundedness of $\rho_{\infty,a}$, i.e. $\sup_{Q,Q'\in\mathcal{Q}}\rho_{\infty,a}(Q,Q')<\infty$. Letting $t\to\infty$ we have

$$\Big|\int F\mathrm{d}\mu-\int F\mathrm{d}\nu\Big|=0.$$

Hence, $\mu=\nu$. This gives the uniqueness of invariant measure, as denoted by $\mu^{\mathrm{YM}}_{N,\beta}$ in the theorem.

By Theorem 3.5, every tight limit is the invariant measure of (1.6). Hence, it is also unique and the second result of the theorem follows.

To prove the last statement (1.7), taking now an *arbitrary* probability measure ν on $\mathcal{Q}$ we also have $\{\nu_L\}$ constructed similarly as above. We denote by Q^{ν_L} and Q^ν the processes starting from ν_L and ν, respectively. We also have, as in (5.14),

$$\int P^L_tF\mathrm{d}\nu_L=\mathbf{E}F(Q^{\nu_L}(t))\to\mathbf{E}F(Q^\nu(t))=\int P_tF\mathrm{d}\nu,\quad L\to\infty. \tag{5.16}$$

Recall $\{\mu_L\}$ and $\mu=\mu^{\mathrm{YM}}_{N,\beta}$ as the unique invariant measure given above. By triangle inequality and Lemma 5.1 we have for $t\geqslant 0$

$$\begin{aligned}W_2^{\rho_{\infty,a}}(\nu P_t,\mu)&\leqslant W_2^{\rho_{\infty,a}}(\nu_LP^L_t,\nu P_t)+W_2^{\rho_{\infty,a}}(\nu_LP^L_t,\mu_LP^L_t)+W_2^{\rho_{\infty,a}}(\mu,\mu_L)\\&\leqslant W_2^{\rho_{\infty,a}}(\nu_LP^L_t,\nu P_t)+e^{-\widetilde{K}_{\mathcal{S}}t}W_2^{\rho_{\infty,a}}(\mu_L,\nu_L)+W_2^{\rho_{\infty,a}}(\mu,\mu_L)\\&\leqslant\mathbf{E}\rho^2_{\infty,a}(Q^{\nu_L}(t),Q^\nu(t))+C(a)e^{-\widetilde{K}_{\mathcal{S}}t}+W_2^{\rho_{\infty,a}}(\mu,\mu_L).\end{aligned} \tag{5.17}$$

As $\mathcal{Q}$ is compact w.r.t. the distance ρ^a_∞, $Q^{\nu_L}(t)$ is tight in $(\mathcal{Q},\rho^a_\infty)$. Using (5.16) we then have for $t\geqslant 0$

$$\mathbf{E}\rho^2_{\infty,a}(Q^{\nu_L}(t),Q^\nu(t))\to 0,\quad L\to\infty.$$

Letting $L\to\infty$ in (5.17), we have

$$W_2^{\rho_{\infty,a}}(\nu P_t,\mu)\leqslant C(a)e^{-\widetilde{K}_{\mathcal{S}}t},$$

which is (1.7). It is clear from (5.15) that $\widetilde{K}_{\mathcal{S}}$ only depends on the constant a, d, β and dimension of G. $\square$

Appendix A. Construction of coupling

In this appendix, we follow [Ken86], [Wan06, Section 2.1] to construct the coupling

$$(Q(t), Q'(t))_{t\geqslant 0}$$

starting from (Q, Q') by approximation, and prove (5.5). The coupling argument presented here is similar with [Wan06, Chapter 2] but a main difference is that the above reference applies Itô's formula to a distance on a given manifold – which would be ρ_L (not ρ_e) in our case, but we will apply Itô's formula to the quantity $\chi_R(\rho_e^2)$.

Before proceeding we recall the basic definitions and notations (c.f. [Hsu02, Chapter 2] for more detailed explanations). Recall that $\mathcal{Q}_L$ is a Riemannian manifold with dimension $d = |E^+_{\Lambda_L}| d(\mathfrak{g})$. Let $\mathscr{O}(\mathcal{Q}_L)$ be the orthonormal frame bundle over $\mathcal{Q}_L$, which is a $d(d+1)/2$-dimensional Riemannian manifold. Given $l \in \mathbb{R}^d$, let H_l be the corresponding horizontal vector field on $\mathscr{O}(\mathcal{Q}_L)$. Denote by $\pi : \mathscr{O}(\mathcal{Q}_L) \to \mathcal{Q}_L$ be the canonical projection. For any $\Phi \in \mathscr{O}(\mathcal{Q}_L)$ we have $\Phi l \in T_{\pi\Phi}\mathcal{Q}_L$ and $H_l(\Phi) \in T_\Phi \mathscr{O}(\mathcal{Q}_L)$ is the horizontal lift of $\Phi l \in T_{\pi\Phi}\mathcal{Q}_L$ to Φ. In particular, let $\{l_i\}_{i=1}^d$ be an orthonormal basis of $\mathbb{R}^d$, define the horizontal Laplace operator

$$\Delta_{\mathscr{O}(\mathcal{Q}_L)} \stackrel{\text{def}}{=} \sum_{i=1}^d H_{l_i}^2,$$

which is independent of the choice of the basis $\{l_i\}$. Moreover, for any vector field Z on $\mathcal{Q}_L$ we define its horizontal lift by $H_\Phi Z \stackrel{\text{def}}{=} H_{\Phi^{-1}Z}(\Phi)$ for $\Phi \in \mathscr{O}(\mathcal{Q}_L)$, where $\Phi^{-1}Z$ is the unique vector $l \in \mathbb{R}^d$ such that $Z_{\pi\Phi} = \Phi l$.

As in [Hsu02, Chapter 6], for a function f defined on $\mathscr{O}(\mathcal{Q}_L) \times \mathscr{O}(\mathcal{Q}_L)$, we denote by $H_{l_i,1}f$ and $H_{l_i,2}f$ the derivatives of f with respect to the horizontal vector field H_{l_i} on the first and the second variable respectively. The horizontal Laplacian on the first and the second variable are

$$\Delta_{\mathscr{O}(\mathcal{Q}_L),1} = \sum_{i=1}^d H_{l_i,1}^2, \quad \Delta_{\mathscr{O}(\mathcal{Q}_L),2} = \sum_{i=1}^d H_{l_i,2}^2.$$

Construction of coupling. Consider the following Stratonovich SDE with $Q(t) \stackrel{\text{def}}{=} \pi(\Phi_t)$

$$\mathrm{d}\Phi_t = \sum_{i=1}^d H_{\Phi_t}(\Phi_t) \circ \mathrm{d}N_t, \qquad \mathrm{d}N_t = \sqrt{2}\mathrm{d}B_t + \Phi_t^{-1}\nabla\mathcal{S}(Q(t))\mathrm{d}t, \tag{A.1}$$

where $(B_t)_{t\geqslant 0}$ is a standard d-dimensional Brownian motion and $\pi\Phi_0 = Q$. Then $Q(t)$ is an $\mathcal{L}_L$-diffusion process ($\mathcal{L}_L$ as in (3.5)) starting from Q and Φ_t is called its horizontal lift.

As the Riemannian distance is not smooth on C and D defined in (5.2), we introduce cut-off approximation as follows: For any $n \geqslant 1$ and $\varepsilon \in (0,1)$, let $h_{n,\varepsilon} \in C^\infty(\mathcal{Q}_L \times \mathcal{Q}_L)$ such that $0 \leqslant h_{n,\varepsilon} \leqslant 1-\varepsilon$, $h_{n,\varepsilon}|_{C_n^c} = 1-\varepsilon$ and $h_{n,\varepsilon}|_{C_{2n}} = 0$, where

$$C_n \stackrel{\text{def}}{=} \Big\{(Q, Q') : \rho_{\mathcal{Q}_L\times\mathcal{Q}_L}((Q,Q'), C) \leqslant \frac{1}{n}\Big\}, \quad n \geqslant 1,$$

with $\rho_{\mathcal{Q}_L\times\mathcal{Q}_L}$ the Riemannian distance on $\mathcal{Q}_L\times\mathcal{Q}_L$. Let $g_n\in C^\infty(\mathcal{Q}_L\times\mathcal{Q}_L)$ such that $0\leqslant g_n\leqslant 1$, $g_n(Q,Q')=0$ if $\rho_L(Q,Q')\leqslant\frac{1}{2n}$ and $g_n(Q,Q')=1$ if $\rho_L(Q,Q')\geqslant\frac{1}{n}$. Let $\Psi_t^{n,\varepsilon}$ and $N_t^{n,\varepsilon}$ solve the following SDE with $\widetilde{Q}^{n,\varepsilon}(t)\stackrel{\text{def}}{=}\pi\Psi_t^{n,\varepsilon}$

$$\begin{aligned}
\mathrm{d}\Psi_t^{n,\varepsilon}&=\sum_{i=1}^d H_{\Psi_t^{n,\varepsilon}}(\Psi_t^{n,\varepsilon})\circ\mathrm{d}N_t^{n,\varepsilon},\\
\mathrm{d}N_t^{n,\varepsilon}&=\sqrt{2}(h_{n,\varepsilon}g_n)(Q(t),\widetilde{Q}^{n,\varepsilon}(t))(\Psi_t^{n,\varepsilon})^{-1}P_{Q(t),\widetilde{Q}^{n,\varepsilon}(t)}\Phi_t\,\mathrm{d}B_t\\
&\quad+\sqrt{2(1-(h_{n,\varepsilon}g_n)^2(Q(t),\widetilde{Q}^{n,\varepsilon}(t)))}\,\mathrm{d}B_t'+(\Psi_t^{n,\varepsilon})^{-1}\nabla\mathcal{S}(\widetilde{Q}^{n,\varepsilon}(t))\,\mathrm{d}t,
\end{aligned}\tag{A.2}$$

where B_t' is a Brownian motion in $\mathbb{R}^d$ independent of B_t, $\pi\Psi_0=Q'$ and $P_{Q,Q'}:T_Q\mathcal{Q}_L\to T_{Q'}\mathcal{Q}_L$ is parallel translation along the geodesic from Q to Q'. As the coefficients are smooth on the compact manifold, we have unique solutions $(\Phi_t,\Psi_t^{n,\varepsilon})$ to (A.1) and (A.2).

The generator for $(\Phi_t,\Psi_t^{n,\varepsilon})$ is then given by

$$\mathcal{L}^{n,\varepsilon}_{\mathscr{O}(\mathcal{Q}_L)}=\Delta_{\mathscr{O}(\mathcal{Q}_L),1}+\Delta_{\mathscr{O}(\mathcal{Q}_L),2}+2\sum_{i=1}^d(h_{n,\varepsilon}g_n)H_{l_i^*,2}H_{l_i,1}+H_\Phi\nabla\mathcal{S}+H_\Psi\nabla\mathcal{S},$$

with $l_i^*(\Phi,\Psi)=\Psi^{-1}P_{\pi\Phi,\pi\Psi}\Phi l_i\in\mathbb{R}^d$.

We then consider the following approximation to the generator $\mathcal{L}^c$ defined in (5.4).

$$\mathcal{L}^{n,\varepsilon}=\sum_{e\in E^+_{\Lambda_L}}\Delta_{Q_e}+\sum_{e\in E^+_{\Lambda_L}}\Delta_{Q_e'}+2g_nh_{n,\varepsilon}\sum_{i,j=1}^{\dim\mathfrak{q}_L}\langle P_{Q,Q'}v_i,v_j'\rangle_{T_{Q'}\mathcal{Q}_L}v_iv_j'+\nabla\mathcal{S}(Q)+\nabla\mathcal{S}(Q'),$$

with $\{v_i\},\{v_j'\}$ as in the definition of $\mathcal{L}^c$ in (5.4).

It is easy to see that $\mathcal{L}^{n,\varepsilon}_{\mathscr{O}(\mathcal{Q}_L)}$ is a lift of $\mathcal{L}^{n,\varepsilon}$. Namely, for $f\in C^2(\mathcal{Q}_L\times\mathcal{Q}_L)$ and $F(\Phi,\Psi)=f(\pi\Phi,\pi\Psi)$, one has $\mathcal{L}^{n,\varepsilon}_{\mathscr{O}(\mathcal{Q}_L)}F(\Phi,\Psi)=\mathcal{L}^{n,\varepsilon}f(\pi\Phi,\pi\Psi)$. We then know $(Q(t),\widetilde{Q}_t^{n,\varepsilon})=(\pi\Phi_t,\pi\Psi_t)$ starting from (Q,Q') is generated by $\mathcal{L}^{n,\varepsilon}$ (c.f. [Wan06, Section 2.1]). Since the marginal operators of $\mathcal{L}^{n,\varepsilon}$ coincide with $\mathcal{L}_L$, $(Q_t,\widetilde{Q}_t^{n,\varepsilon})$ gives a coupling of $\mathcal{L}_L$-diffusions starting from different initial data.

Let $\mathbb{P}^Q$ denote the law of $\mathcal{L}_L$-diffusion $(Q(t))_{t\geqslant0}$ starting from $Q\in\mathcal{Q}_L$ in $C([0,\infty);\mathcal{Q}_L)$ endowed with the distance

$$\widetilde{\rho}_L(Q,Q')\stackrel{\text{def}}{=}\sum_{n=0}^\infty 2^{-n}\Big(1\wedge\sup_{t\in[n,n+1]}\rho_L(Q(t),Q'(t))\Big),\quad Q,Q'\in C([0,\infty);\mathcal{Q}_L).$$

As the marginal law of $(Q,\widetilde{Q}^{n,\varepsilon})_{n,\varepsilon}$ is tight in $C([0,\infty);\mathcal{Q}_L)$, the joint law $\mathbb{P}^{Q,Q'}_{n,\varepsilon}$ of $(Q,\widetilde{Q}^{n,\varepsilon})_{n,\varepsilon}$ is also tight. Therefore, for every $\varepsilon>0$ there exists a probability measure $\mathbb{P}^{Q,Q'}_\varepsilon$ and a subsequence, which is still denoted by $\mathbb{P}^{Q,Q'}_{n,\varepsilon}$ such that $\mathbb{P}^{Q,Q'}_{n,\varepsilon}\to\mathbb{P}^{Q,Q'}_\varepsilon$ weakly in $C([0,\infty);\mathcal{Q}_L)$. Moreover, we could find $\mathbb{P}^{Q,Q'}_{\varepsilon_k}$ and $\mathbb{P}^{Q,Q'}$ such that $\mathbb{P}^{Q,Q'}_{\varepsilon_k}\to\mathbb{P}^{Q,Q'}$ weakly in $C([0,\infty);\mathcal{Q}_L)$. $\mathbb{P}^{Q,Q'}$ is then the desired coupling of $\mathbb{P}^Q$ and $\mathbb{P}^{Q'}$.

Proof of inequality (5.5)**.** In the following we prove (5.5).

Following [Ken86], [Wan06, Section 2.1] we apply Itô's formula to $\chi_R(\rho^2)(Q_e(t),\widetilde{Q}_e^{n,\varepsilon}(t))$ and use $\chi_R'\geqslant 0$ to obtain

$$\begin{aligned}
&\mathrm{d}\chi_R(\rho^2)(Q_e(t),\widetilde{Q}_e^{n,\varepsilon}(t))\\
&\quad=\mathrm{d}M_t^{n,\varepsilon}+2(4\rho^2\chi_R''(\rho^2)+2\chi_R'(\rho^2))(1-g_nh_{n,\varepsilon})(Q(t),\widetilde{Q}^{n,\varepsilon}(t))\mathrm{d}t-\mathrm{d}L_t^{n,\varepsilon}\\
&\qquad+\mathbf{1}_{C^c\cap D^c}2\chi_R'(\rho^2)\rho\Big(g_nh_{n,\varepsilon}I_{\mathcal{S}}+(1-g_nh_{n,\varepsilon})Z\Big)(Q(t),\widetilde{Q}^{n,\varepsilon}(t))\mathrm{d}t.
\end{aligned}\tag{A.3}$$

Here $\rho=\rho(Q_e(t),\widetilde{Q}_e^{n,\varepsilon}(t))$ and $f(\rho^2)=f(\rho^2)(Q_e(t),\widetilde{Q}_e^{n,\varepsilon}(t))$ for $f\in\{\chi_R',\chi_R''\}$. The term $M_t^{n,\varepsilon}$ is a martingale with quadratic variation process given by

$$\int_0^t 4(2\chi_R'(\rho^2)\rho)^2(1-g_nh_{n,\varepsilon})(Q(s),\widetilde{Q}^{n,\varepsilon}(s))\mathrm{d}s,$$

and $L_t^{n,\varepsilon}$ is a non-decreasing process which increases only when $(Q(t),\widetilde{Q}^{n,\varepsilon}(t))\in C$. The term $I_{\mathcal{S}}$ is given in (5.6) and finally

$$Z(Q,Q')=\Delta\rho(\cdot,Q_e')(Q_e)+\Delta\rho(Q_e,\cdot)(Q_e')+\Big((\nabla\mathcal{S})\hat{\rho}_e(\cdot,Q_e')\Big)(Q)+\Big((\nabla\mathcal{S})\hat{\rho}_e(Q_e,\cdot)\Big)(Q').\tag{A.4}$$

with Δ being the Laplace–Beltrami operator on G.

In fact, to derive (A.3), we may first apply Itô's formula to $\mathrm{d}\rho(Q_e(t),\widetilde{Q}_e^{n,\varepsilon}(t))$ and then apply Itô's formula again to $\chi_R(\rho^2)$. Suppose for now that the quadratic variation process of the martingale part M^ρ for ρ is given by

$$\langle M_t^\rho\rangle=4\int_0^t(1-g_nh_{n,\varepsilon})(Q(s),\widetilde{Q}^{n,\varepsilon}(s))\mathrm{d}s.\tag{A.5}$$

Then the second term on the r.h.s. of (A.3) comes from $\langle M_t^\rho\rangle$. As explained in (5.6)–(5.7) the last line in (A.3) comes from $\mathcal{L}^{n,\varepsilon}\rho$.

We now verify (A.5). Using (A.1) and (A.2) we have, for $\widetilde{\rho}(\Phi,\Psi)=\rho(\pi_e\pi\Phi,\pi_e\pi\Psi)$,

$$\mathrm{d}\langle M_t^\rho\rangle=2\sum_{i=1}^d\Big[\Big|H_{l_i,1}\widetilde{\rho}+g_nh_{n,\varepsilon}H_{l_i^*,2}\widetilde{\rho}\Big|^2+\big(1-(g_nh_{n,\varepsilon})^2\big)\big|H_{l_i,2}\widetilde{\rho}\big|^2\Big]\mathrm{d}t.$$

As $\widetilde{\rho}$ only depends on the component at e, this now boils down to standard arguments. Namely, since the above quantity is independent of the choice of basis, we can write the above terms as vector fields on G, and choose basis in such a way that one basis vector is tangent to the geodesic and the others are perpendicular to the geodesic, and use the first variation formula to derive (A.5) (c.f. [Hsu02, Section 6.6]).

In the following we send $n\to\infty,\varepsilon\to0$ to derive (5.5).

Let $(x(t),y(t))$ be the canonical process on $(C([0,\infty);\mathcal{Q}_L)\times C([0,\infty);\mathcal{Q}_L),\mathscr{F}\times\mathscr{F})$ and let $\{\mathscr{F}_t\}_{t\geqslant0}$ be the natural filtration. On the support of χ_R using Laplacian comparison theorem (c.f. [Hsu02, Corollary 3.4.4]) we know that Z defined in (A.4) satisfies $Z\leqslant C_R$ for some constant $C_R>0$. Since the support of $\chi_R(\rho^2)\subset\{\rho^2\geqslant\frac{1}{2R}\}$ and for n large enough $\rho_L(x,y)\geqslant\rho(x_e,y_e)\geqslant1/\sqrt{2R}\geqslant1/n$, $g_n(x,y)=1$ on the support of $\chi_R(\rho^2)$, we obtain that

$$\begin{aligned}
\chi_R(\rho^2)(Q_e(t),\widetilde{Q}_e^{n,\varepsilon}(t))-\int_0^t&2(4\rho^2\chi_R''(\rho^2)+2\chi_R'(\rho^2))(1-h_{n,\varepsilon})(Q(s),\widetilde{Q}^{n,\varepsilon}(s))\\
&+2\chi_R'(\rho^2)\rho\Big(h_{n,\varepsilon}J+(1-h_{n,\varepsilon})C_R\Big)(Q(s),\widetilde{Q}^{n,\varepsilon}(s))\,\mathrm{d}s
\end{aligned}$$

is a supermartingale, where J is as in (5.5). Therefore,

$$S_t^{n,\varepsilon} \stackrel{\text{def}}{=} \chi_R(\rho^2)(x_e(t),y_e(t)) - \int_0^t 2(4\rho^2\chi_R''(\rho^2)+2\chi_R'(\rho))(x_e(s),y_e(s))(1-h_{n,\varepsilon})(x(s),y(s)) + 2(\chi_R'(\rho^2)\rho)(x_e(s),y_e(s))\Big(h_{n,\varepsilon}J+(1-h_{n,\varepsilon})C_R\Big)(x(s),y(s))\,\mathrm{d}s$$

is a $\mathbb{P}^{Q,Q'}_{n,\varepsilon}$-supermartingale.

Furthermore, since by [Wan06, Lemma 2.1.2] $\mathbb{P}^{Q,Q'}_{\varepsilon}((x(s),y(s))\in C)=0$ for $s>0$, using the same argument as in [Wan06, Proof part (b) of Theorem 2.1.1] we let $n\to\infty$ and obtain that

$$S_t^{\varepsilon} \stackrel{\text{def}}{=} \chi_R(\rho^2)(x_e(t),y_e(t)) - \int_0^t 2\varepsilon(4\rho^2\chi_R''(\rho^2)+2\chi_R'(\rho))(x_e(s),y_e(s)) + 2(\chi_R'(\rho^2)\rho)(x_e(s),y_e(s))\Big((1-\varepsilon)J+\varepsilon C_R\Big)(x(s),y(s))\,\mathrm{d}s$$

ia a $\mathbb{P}^{Q,Q'}_{\varepsilon}$-supermartingale.

Letting $\varepsilon\to 0$ we obtain that

$$S_t \stackrel{\text{def}}{=} \chi_R(\rho^2)(x_e(t),y_e(t)) - 2\int_0^t (\chi_R'(\rho^2)\rho)(x_e(s),y_e(s))J(x(s),y(s))\,\mathrm{d}s$$

is a $\mathbb{P}^{Q,Q'}$-supermartingale. Hence, by Doob–Meyer's decomposition

$$\mathrm{d}\chi_R\Big(\rho^2(x_e(t),y_e(t))\Big) = \mathrm{d}M_t + 2(\chi_R'(\rho^2)\rho)\Big(x_e(t),y_e(t)\Big)J(x(t),y(t))\mathrm{d}t - \mathrm{d}L_t, \tag{A.6}$$

with M a martingale and L a predictable increasing process.

In the following we prove $M=0$. Similarly we use the above argument for $f(\chi_R(\rho^2(x_e(t),y_e(t))))$ with $0\leqslant f\in C^2(\mathbb{R}^+)$, $f'\geqslant 0$ and we have that

$$f\Big(\chi_R(\rho^2)(x_e(t),y_e(t))\Big) - 2\int_0^t \Big(f'(\chi_R(\rho^2))\chi_R'(\rho^2)\rho\Big)\Big(x_e(s),y_e(s)\Big)J(x(s),y(s))\,\mathrm{d}s$$

is a $\mathbb{P}^{Q,Q'}$-supermartingale. Choosing $f(r)=\exp(mr), m\in\mathbb{N}$ and setting

$$\Xi_t \stackrel{\text{def}}{=} \exp\Big(m\chi_R\Big(\rho^2(x_e(t),y_e(t))\Big)\Big),$$

we have that

$$\Xi_t - 2m\int_0^t \Xi_s^1 J(x(s),y(s))\,\mathrm{d}s$$

is a $\mathbb{P}^{Q,Q'}$-supermartingale with

$$\Xi_s^1 = \Big(\exp(m\chi_R(\rho^2))\chi_R'(\rho^2)\rho\Big)(x_e(s),y_e(s)).$$

By Doob–Meyer's decomposition,

$$\mathrm{d}\Xi_t = \mathrm{d}\widetilde{M}_t + 2m\Xi_t^1 J(x(t),y(t))\mathrm{d}t - \mathrm{d}\widetilde{L}_t, \tag{A.7}$$

where $\widetilde{M}_t$ is a martingale and $\widetilde{L}_t$ is predictable increasing.

On the other hand, applying Itô's formula to (A.6) we obtain

$$\mathrm{d}\Xi_t = m\,\Xi_t\,\mathrm{d}M_t + \frac{1}{2}m^2\,\Xi_t\,\mathrm{d}\langle M_t,M_t\rangle + 2m\,\Xi_t^1\,J(x(t),y(t))\,\mathrm{d}t - m\,\Xi_t\,\mathrm{d}L_t. \tag{A.8}$$

Comparing (A.7) and (A.8) we obtain $\mathrm{d}\langle M,M\rangle_t \leqslant \frac{2}{m}\mathrm{d}L_t$. Letting $m\to\infty$, we get $\mathrm{d}\langle M,M\rangle_t = 0$. Hence, (5.5) follows.

References

[AC22] A. Adhikari and S. Cao. Correlation decay for finite lattice gauge theories at weak coupling. *arXiv preprint arXiv:2202.10375*, 2022.

[AD99] L. Andersson and B. K. Driver. Finite dimensional approximations to Wiener measure and path integral formulas on manifolds. *Journal of functional analysis*, 165(2):430–498, 1999.

[Adh21] A. Adhikari. Wilson loop expectations for non-abelian gauge fields coupled to a Higgs boson at low and high disorder. *arXiv preprint arXiv:2111.07540*, 2021.

[AGZ10] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[AHKZ89] S. Albeverio, R. Hoegh-Krohn, and B. Zegarliński. Uniqueness of Gibbs states for general $P(\varphi)_2$-weak coupling models by cluster expansion. *Comm. Math. Phys.*, 121(4):683–697, 1989.

[Bak97] D. Bakry. On Sobolev and logarithmic Sobolev inequalities for Markov semigroups. In *New trends in stochastic analysis (Charingworth, 1994)*, pages 43–75. World Sci. Publ., River Edge, NJ, 1997.

[BB21] R. Bauerschmidt and T. Bodineau. Log-Sobolev inequality for the continuum Sine-Gordon model. *Communications on Pure and Applied Mathematics*, 74(10):2064–2113, 2021.

[BD22] R. Bauerschmidt and B. Dagallier. Log-Sobolev inequality for the φ^4_2 and φ^4_3 measures. *arXiv preprint arXiv:2202.02295*, 2022.

[BE85] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[BG18] R. Basu and S. Ganguly. SO(N) lattice gauge theory, planar and beyond. *Comm. Pure Appl. Math.*, 71(10):2016–2064, 2018.

[BGHZ21] Y. Bruned, F. Gabriel, M. Hairer, and L. Zambotti. Geometric stochastic heat equations. *J. Amer. Math. Soc.*, 35(1):1–80, 2021.

[BGL14] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 103. Springer, 2014.

[Cao20] S. Cao. Wilson loop expectations in lattice gauge theories with finite gauge groups. *Comm. Math. Phys.*, 380(3):1439–1505, 2020.

[CCHS22a] A. Chandra, I. Chevyrev, M. Hairer, and H. Shen. Langevin dynamic for the 2D Yang–Mills measure. *Publ. Math. IHS*, 2022.

[CCHS22b] A. Chandra, I. Chevyrev, M. Hairer, and H. Shen. Stochastic quantisation of Yang-Mills-Higgs in 3D. *arXiv preprint arXiv:2201.03487*, 2022.

[CE75] J. Cheeger and D. G. Ebin. *Comparison theorems in Riemannian geometry.* North-Holland Mathematical Library, Vol. 9. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1975.

[Cha16] S. Chatterjee. The leading term of the Yang-Mills free energy. *J. Funct. Anal.*, 271(10):2944–3005, 2016.

[Cha19a] S. Chatterjee. Rigorous solution of strongly coupled $SO(N)$ lattice gauge theory in the large N limit. *Comm. Math. Phys.*, 366(1):203–268, 2019.

[Cha19b] S. Chatterjee. Yang-Mills for probabilists. In *Probability and analysis in interacting physical systems*, volume 283 of *Springer Proc. Math. Stat.*, pages 1–16. Springer, Cham, 2019.

[Cha20] S. Chatterjee. Wilson loops in Ising lattice gauge theory. *Comm. Math. Phys.*, 377(1):307–340, 2020.

[Cha21] S. Chatterjee. A probabilistic mechanism for quark confinement. *Comm. Math. Phys.*, 385(2):1007–1039, 2021.

[Che19] I. Chevyrev. Yang-Mills measure on the two-dimensional torus as a random distribution. *Comm. Math. Phys.*, 372(3):1027–1058, 2019.

[Che22] I. Chevyrev. Stochastic quantisation of Yang-Mills. *arXiv preprint arXiv:2202.13359*, 2022.

[CJ16] S. Chatterjee and J. Jafarov. The $1/N$ expansion for $SO(N)$ lattice gauge theory at strong coupling. *arXiv preprint arXiv:1604.04777*, 2016.

[CWZZ21] X. Chen, B. Wu, R. Zhu, and X. Zhu. Stochastic heat equations for infinite strings with values in a manifold. *Trans. Amer. Math. Soc.*, 374(1):407–452, 2021.

[dC92] M. P. do Carmo. *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

[Dob70] R. L. Dobrushin. Prescribing a system of random variables by conditional distributions. *Theory of Probability & Its Applications*, 15(3):458–486, 1970.

[Dri89] B. K. Driver. YM_2: continuum expectations, lattice convergence, and lassos. *Comm. Math. Phys.*, 123(4):575–616, 1989.

[Dri19] B. K. Driver. A functional integral approaches to the Makeenko-Migdal equations. *Comm. Math. Phys.*, 370(1):49–116, 2019.

[FLV21] M. P. Forsström, J. Lenells, and F. Viklund. Wilson loops in the abelian lattice Higgs model. *arXiv e-prints*, July 2021.

[FLV22] M. P. Forsström, J. Lenells, and F. Viklund. Wilson loops in finite Abelian lattice gauge theories. *Ann. Inst. Henri Poincaré Probab. Stat.*, 58(4):2129–2164, 2022.

[For21] M. P. Forsström. Wilson lines in the abelian lattice Higgs model. *arXiv preprint arXiv:2111.06620*, 2021.

[For22] M. P. Forsström. Decay of correlations in finite Abelian lattice gauge theories. *Comm. Math. Phys.*, 393(3):1311–1346, 2022.

[FOT94] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1994.

[GJS73] J. Glimm, A. Jaffe, and T. Spencer. The particle structure of the weakly coupled $P(\phi)_2$ model and other applications of high temperature expansions. In *Constructive quantum field theory*, pages 132–198. Springer, 1973.

[GKS89] L. Gross, C. King, and A. Sengupta. Two-dimensional Yang-Mills theory via stochastic differential equations. *Ann. Physics*, 194(1):65–112, 1989.

[GS21] C. Garban and A. Sepúlveda. Improved spin-wave estimate for Wilson loops in $U(1)$ lattice gauge theory. *arXiv e-prints*, July 2021.

[GZ03] A. Guionnet and B. Zegarlinski. Lectures on logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXVI*, volume 1801 of *Lecture Notes in Math.*, pages 1–134. Springer, Berlin, 2003.

[Hai16] M. Hairer. The motion of a random string. *ArXiv e-prints*, May 2016.

[Hsu02] E. P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[Jaf16] J. Jafarov. Wilson loop expectations in $SU(N)$ lattice gauge theory. *arXiv preprint arXiv:1610.03821*, 2016.

[Ken86] W. S. Kendall. Nonnegative Ricci curvature and the Brownian coupling property. *Stochastics*, 19(1-2):111–129, 1986.

[Kur07] T. G. Kurtz. The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities. *Electron. J. Probab.*, 12:951–965, 2007.

[LÍ7] T. Lévy. The master field on the plane. *Astérisque*, (388):ix+201, 2017.

[Lév03] T. Lévy. Yang-Mills measure on compact surfaces. *Mem. Amer. Math. Soc.*, 166(790):xiv+122, 2003.

[MR92] Z. M. Ma and M. Röckner. *Introduction to the theory of (nonsymmetric) Dirichlet forms*. Universitext. Springer-Verlag, Berlin, 1992.

[Nel66] E. Nelson. Derivation of the Schrödinger Equation from Newtonian Mechanics. *Phys. Rev.*, 150:1079–1085, Oct 1966.

[Nua06] D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[OS78] K. Osterwalder and E. Seiler. Gauge field theories on a lattice. *Annals of Physics*, 110(2):440–471, 1978.

[PW81] G. Parisi and Y. S. Wu. Perturbation theory without gauge fixing. *Sci. Sinica*, 24(4):483–496, 1981.

[RWZZ20] M. Röckner, B. Wu, R. Zhu, and X. Zhu. Stochastic Heat Equations with Values in a Manifold via Dirichlet Forms. *SIAM J. Math. Anal.*, 52(3):2237–2274, 2020.

[Sei82] E. Seiler. *Gauge theories as a problem of constructive quantum field theory and statistical mechanics*, volume 159 of *Lecture Notes in Physics*. Springer-Verlag, Berlin, 1982.

[Spe75] T. Spencer. The decay of the Bethe-Salpeter kernel in $P(\phi)_2$ quantum field models. *Comm. Math. Phys.*, 44(2):143–164, 1975.

[SSZ22] H. Shen, S. A. Smith, and R. Zhu. A new derivation of the finite n master loop equation for lattice Yang–Mills. *arXiv preprint arXiv:2202.00880*, 2022.

[tH74] G. 't Hooft. A planar diagram theory for strong interactions. *Nuclear Phys. B*, 72:461–473, 1974.

[vRS05] M.-K. von Renesse and K.-T. Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.

[Wan06] F. Wang. *Functional inequalities, Markov semigroups and spectral theory*. Elsevier, 2006.

[Wil74] K. G. Wilson. Confinement of quarks. *Physical review D*, 10(8):2445, 1974.

(H. Shen) Department of Mathematics, University of Wisconsin - Madison, USA

E-mail address: pkushenhao@gmail.com

(R. Zhu) Department of Mathematics, Beijing Institute of Technology, Beijing 100081, China; Beijing Key Laboratory on MCAACI, Beijing, China

E-mail address: zhurongchan@126.com

(X. Zhu) Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

E-mail address: zhuxiangchan@126.com