2022-04-12

Statistical Properties of 2D Navier-Stokes Equations Driven by Quasi-Periodic Force and Degenerate Noise

Rongchang Liu
Brigham Young University

Statistical Properties of 2D Navier-Stokes Equations Driven by Quasi-Periodic Force and Degenerate Noise

Rongchang Liu

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Kening Lu, Chair
Lennard Frank Bakker
Mark A. Allen
Christopher P. Grant
Benjamin Zachary Webb

Department of Mathematics

Brigham Young University

ABSTRACT

Statistical Properties of 2D Navier-Stokes Equations Driven by Quasi-Periodic Force and Degenerate Noise

Rongchang Liu
Department of Mathematics, BYU
Doctor of Philosophy

We consider the incompressible 2D Navier-Stokes equations on the torus driven by a deterministic time quasi-periodic force and a noise that is white in time and extremely degenerate in Fourier space. We show that the asymptotic statistical behavior is characterized by a uniquely ergodic and exponentially mixing quasi-periodic invariant measure. The result is true for any value of the viscosity $\nu > 0$. By utilizing this quasi-periodic invariant measure, we show the strong law of large numbers and central limit theorem for the continuous time inhomogeneous solution processes. Estimates of the corresponding rate of convergence are also obtained, which is the same as in the time homogeneous case for the strong law of large numbers, while the convergence rate in the central limit theorem depends on the Diophantine approximation property on the quasi-periodic frequency and the mixing rate of the quasi-periodic invariant measure. We also prove the existence of a stable quasi-periodic solution in the laminar case (when the viscosity is large). The scheme of analyzing the statistical behavior of the time inhomogeneous solution process by the quasi-periodic invariant measure could be extended to other inhomogeneous Markov processes.

Keywords: Navier-Stokes equations, quasi-periodic invariant measure, unique ergodicity, mixing, limit theorems, rate of convergence, Diophantine condition

Acknowledgements

I would like to thank Professor Kening Lu from the depth of my heart, for introducing me to this amazing interdisciplinary research field, and providing lots of valuable suggestions and encouragement as well as so much time for discussion to improve my writing and research, so that I can complete this study.

I would also like to express my sincere gratitude to the committee members, and all the professors and staff in the Department of Mathematics for their kind help and endless support.

I am also thankful for my families, especially my parents who encourage me all the time. Without their support I cannot arrive at BYU and complete this dissertation.

Contents

Chapter 1. Introduction

We study the asymptotic statistical properties of the time inhomogeneous solution process of the incompressible 2D Navier-Stokes equation driven by a deterministic time-dependent quasi-periodic force and a highly degenerate stochastic force. There are three main ingredients of our work for this system:

(1) We extend the Harris-like theorem in the infinite dimensional hypoelliptic setting developed by Hairer and Mattingly [41] to the time quasi-periodic inhomogeneous case. This allows us to view the two-parameter Markov transition operator as a contractive deterministic non-autonomous dynamical system acting on the space of probability measures endowed with an appropriate Wasserstein metric. We then show that the non-autonomous system has a globally exponentially attracting quasi-periodic trajectory, serving as the uniquely ergodic and exponentially mixing quasi-periodic invariant measure, which describes the quasi-periodically statistical steady states of the system.

(2) The quasi-periodic invariant measure in (1) enables us to develop a martingale approximation scheme, to show the strong law of large numbers and central limit theorem for the time inhomogeneous solution process starting at every deterministic point and for Hölder observables weighted by a Lyapunov function. The martingale approximation is applied to the corresponding homogenized process obtained by taking the skew product of the solution process with the irrational rotation flow induced by the quasi-periodic force. However, the homogenized process is not mixing so the martingale approximation cannot be applied directly as in the usual way. This is resolved by centering the observables using the quasi-periodic invariant measure.

(3) We also obtain an estimate on the rate of convergence in the limit theorems. For the strong law of large numbers, the convergence rate is the same as in the time homogeneous case. While in the case of the central limit theorem, the convergence rate is related to the mixing rate of the quasi-periodic invariant measure and the convergence rate of the Birkhoff average of the irrational rotation flow with particular observable functions involving the quasi-periodic invariant measure. The latter has a close connection with the Diophantine approximation property of the frequency.

We confine ourselves to the 2D Navier-Stokes system to simplify the presentation of the main ideas. However, the results can be applied to a class of time inhomogeneous Markov process generated by stochastic semilinear equations driven by an additional time dependent deterministic force for which the inhomogeneity can be modeled by a dynamical system (the Bebutov shift for example). In what follows we will briefly describe our result and the background. A more detailed and technical description of the main results is given in Chapter 3.

1.1 Description of Results

Consider the incompressible 2D Navier-Stokes on the two dimensional torus $\mathbb{T}^2 := \mathbb{R}^2/(2\pi)\mathbb{Z}^2$ in the vorticity form

$$dw(t,x) + B(\mathcal{K}w, w)(t,x)dt = \nu\Delta w(t,x)dt + f(t,x)dt + \sum_{i=1}^{d} g_i dW_i(t), \quad t > s, \quad w(s) = w_0, \tag{1.1}$$

where $w(t,x)$ is the vorticity field, and $\mathcal{K}w$ is the divergence free velocity field. The phase space is chosen as $H := \left\{w \in \mathrm{L}^2\left(\mathbb{T}^2, \mathbb{R}\right) : \int_{\mathbb{T}^2} w dx = 0\right\}$ whose norm is denoted by $\|\cdot\|$. The deterministic force f is quasi-periodic in t in the sense that $f(t,x) = \Psi(\alpha t, x)$ for some $\Psi \in C(\mathbb{T}^n, H)$. Here the frequency $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ and $\{\alpha_k\}_{k=1}^n$ are rationally independent. $W = (W_1, W_2, \cdots, W_d)$ is a two-sided $\mathbb{R}^d$-valued standard Wiener process over the sample space $(\Omega, \mathcal{F}, \mathbf{P})$ where $\mathbf{P}$ is the Wiener measure, and $\{g_i\}$ are elements of H. Under appropriate spatial regularity conditions on the external forces (see the next section), the equation is well posed with a time inhomogeneous Markov solution process $w_{s,t}(w_0)$. It generates a two-parameter Markov transition operator $\mathcal{P}_{s,t}$ acting on the space of bounded measurable functions $B_b(H)$ as

$$\mathcal{P}_{s,t}\phi(w_0) = \mathbf{E}\phi(w_{s,t}(w_0)), \quad \forall\phi \in B_b(H), w_0 \in H.$$

It acts on the space of probability measures $\mathcal{P}(H)$ by duality

$$\mathcal{P}_{s,t}^*\mu(A) = \int_H \mathcal{P}_{s,t}\mathbb{I}_A(w)\mu(dw), \text{ for } \mu \in \mathcal{P}(H), A \in \mathfrak{B},$$

where $\mathfrak{B}$ is the Borel σ-algebra of H and $\mathbb{I}_A$ is the indicator function of A. For $\eta > 0$ small, recall the geodesic metric ρ weighted by a Lyapunov function introduced in [41]

$$\rho(w_1, w_2) = \inf_{\gamma} \int_0^1 e^{\eta\|\gamma(t)\|^2}\|\dot{\gamma}(t)\|dt, \quad \forall w_1, w_2 \in H, \tag{1.2}$$

where the infimum is taken over all differentiable paths γ connecting $w_1, w_2 \in H$. We endow $\mathcal{P}(H)$ with the topology of weak convergence and denote by $\mathcal{P}_1(H)$ the set of probability measures that have finite first moment with respect to the 1-Wasserstein metric induced from the metric ρ in H. A quasi-periodic invariant measure is a continuous quasi-periodic map with values in $\mathcal{P}(H)$

satisfying the following invariance condition

$$\mathcal{P}_{s,t}^{*}\mu_s = \mu_t, \quad s \leq t. \tag{1.3}$$

To state the main result, we recall a condition on the structure of the degenerate noise from [42]. Define the set A_∞ by setting $A_1 = \{g_l : 1 \leq l \leq d\}$, $A_{k+1} = A_k \cup \{\widetilde{B}(h, g_l) : h \in A_k, g_l \in A_1\}$, and $A_\infty = \overline{\operatorname{span}(\cup_{k\geq 1} A_k)}$, where $\widetilde{B}(u, w) = -B(\mathcal{K}u, w) - B(\mathcal{K}w, u)$ is the symmetrized nonlinear term. These sets reflect the mechanism of the propagation of the extremely degenerate noise to the phase space that yields a smoothing effect of the dynamics.

Besides spatial regularity conditions on the external forces, the only remaining assumption for our main results is $A_\infty = H$. In particular, the noise is allowed to be extremely degenerate to have $A_\infty = H$, for example it can be excited only through four directions [40].

Remark. Note that our result does not rely on any condition on the viscosity $\nu > 0$, nor conditions on the strength of the external forces. In particular, we do not need the range condition as in [41] and our result in the case when f is time independent verifies a conjecture made by Hairer and Mattingly [41] (see Remark 1.3) that the spectral gap (as well as unique ergodicity and exponentially mixing) holds without any restriction on f other than it be sufficiently smooth.

1.1.1 Unique Ergodicity and Mixing of the Quasi-periodic Invariant Measure.

One of the main ingredients of this dissertation is to provide a perspective to investigate the asymptotic statistical behavior of the time inhomogeneous Markov solution process, by analyzing the dynamics of the non-autonomous system on $\mathcal{P}(H)$ induced by the action of the two parameter Markov transition operators $\mathcal{P}_{s,t}^{*}$. It turns out that the asymptotic behavior of $\mathcal{P}_{s,t}^{*}$ is characterized by a complete quasi-periodic trajectory in $\mathcal{P}(H)$. More precisely, we prove the existence of a unique quasi-periodic path $\{\mu_t\}_{t\in\mathbb{R}}$ in $\mathcal{P}(H)$ satisfying the invariance condition (1.3), such that

$$\rho(\mathcal{P}_{s,t}^{*}\mu, \mu_t) \leq Ce^{-\varpi(t-s)}\rho(\mu, \mu_s), \quad \forall s \leq t, \mu \in \mathcal{P}(H), \tag{1.4}$$

where C, ϖ are positive constants. Such a unique quasi-periodic path is called a uniquely ergodic and exponentially mixing quasi-periodic invariant measure. One of the classical methods to show the unique ergodicity and mixing in the homogeneous setting, which dates back to the early works of Doeblin [18] and Harris [37], is to show that the action of the Markov semigroup on $\mathcal{P}(H)$ is a contraction under an appropriate metric, whose unique fixed point gives the uniquely ergodic and mixing invariant measure. The work of Harris was generalized to the hypoelliptic case for infinite dimensional systems by Hairer and Mattingly [41]. Our result can be regarded as an extension of their work to the time (quasi-periodically) inhomogeneous setting.

The proof of (1.4) will be given in chapter 5 by first proving $\mathcal{P}_{s,t}^*$ is a contraction on $\mathcal{P}(H)$ in Chapter 4 and then using a fixed point argument to the induced action (through a pull-back procedure) on the space of quasi-periodic graphs $C(\mathbb{T}^n, \mathcal{P}_1(H))$. Besides the Lyapunov structure and a particular type of irreducibility (in a form that is more uniform than the usual topological irreducibility), the proof of the contraction for $\mathcal{P}_{s,t}^*$ requires a deep analysis of its gradient in the time inhomogeneous hypoelliptic setting, which was first developed in the time homogeneous case by Hairer and Mattingly in their celebrated works [40, 41, 42]. In addtion, we prove the particular type of irreducibility here by combining the parabolic regularizing effect of the equation and the usual topological irreducibility, where the latter is a consequence of the well known controllability results of Agrachev and Sarychev [1, 2].

1.1.2 Limit Theorems in Terms of the Quasi-periodic Invariant Measure. In the time homogeneous setting, the strong law of large numbers and central limit theorem show that the asymptotic behavior of an observation along a Markov process can be characterized by the unique invariant measure. Suppose that we are given a homogeneous Markov process X_t with a unique invariant measure μ_* with a certain independence condition (mixing for example) and $\phi \in C(H, \mathbb{R})$ is an observable function with some regularity, say Hölder continuous. Then as $T \to \infty$, one has (see [48] for example) the strong law of large numbers

$$\frac{1}{T}\int_0^T \phi(X_t)dt \xrightarrow{\text{a.s.}} \langle \mu_*, \phi \rangle, \tag{1.5}$$

and the central limt theorem

$$\frac{1}{\sqrt{T}}\int_0^T \Big(\phi(X_t) - \langle \mu_*, \phi \rangle\Big)dt \xrightarrow{\mathcal{D}} N(0, \sigma^2), \tag{1.6}$$

where $\langle \mu_*, \phi \rangle$ is the integral of ϕ with respect to μ_*, and a.s. denotes the almost sure convergence w.r.t. the Wiener measure $\mathbf{P}$, while $\mathcal{D}$ represents the convergence in distribution and $N(0, \sigma^2)$ is the centered normal distribution with variance $\sigma^2 \geq 0$. The strong law of large numbers in the form (1.5) also bears the name of the ergodic theorem, which states that the time average of the observations converges to the ensemble average for almost every sample, regardless of the initial condition. And the central limit theorem measures the size of fluctuations around the ensemble average. We notice that the following form of the strong law of large numbers

$$\frac{1}{T}\int_0^T \Big(\phi(X_t) - \langle \mu_*, \phi\rangle\Big) dt \xrightarrow{\text{a.s.}} 0 \tag{1.7}$$

captures another role played by the invariant measure in the sense that, the average of observations centered by the invariant measure converges to the corresponding asymptotic mean. We will see that the quasi-periodic invariant measure has the same feature, which is the key to extend the limit theorems to the time inhomogeneous setting.

The celebrated Dobrusion's theorem [17] is of particular importance in the time inhomogeneous case. It shows that for any discrete time inhomogeneous Markov chain X_k with a certain compatibility condition between the minimal ergodic coefficients, the observable functions ϕ_k and the variance, one has the convergence as $n \to \infty$,

$$\frac{S_n - \mathbf{E}[S_n]}{\sqrt{\mathrm{Var}(S_n)}} \xrightarrow{\mathcal{D}} N(0,1), \tag{1.8}$$

where $S_n = \sum_{k=1}^{n} \phi_k(X_k)$ and $\mathrm{Var}(S_n)$ is the variance. Although this theorem is quite general in its own right, it is not applicable to the Navier-Stokes equation in our context. Indeed, as mentioned in [41], the transition probabilities in infinite dimensional systems are likely to be mutually singular, especially in the case when the strong Feller property does not hold. Hence the non-degeneracy condition (characterized by the total variational metric) on the minimal ergodic coefficients in Dobrusion's theorem may not be satisfied. Besides, to apply the Dobrusion's theorem, one needs to compute the expectation and variance along each observation as indicated in (1.8).

Our second result shows how the uniquely ergodic and mixing quasi-periodic invariant measure μ_t enables us to give the limit theorems for the time inhomogeneous solution process. Indeed, in view of the role played by the invariant measure in the homogeneous case, and the fact that the ergodicity and mixing of the quasi-periodic invariant measure shows that the distribution of any solution is exponentially attracted by the quasi-periodic path in $\mathcal{P}(H)$, one may expect that the quasi-periodic invariant measure carries the information that one needs to center the observations appropriately to derive the associated limit theorems. In fact, we obtain the following limit theorems. For any Hölder continuous observable function ϕ, and any initial data $w_0 \in H$, the solution $w_{s,s+t}(w_0)$ of the Navier-Stokes system satisfies as $T \to \infty$ the strong law of large numbers

$$\frac{1}{T}\int_0^T \Big(\phi(w_{s,s+t}(w_0)) - \langle \mu_{s+t}, \phi\rangle\Big)dt \xrightarrow{\text{a.s.}} 0, \tag{1.9}$$

and the central limit theorem

$$\frac{1}{\sqrt{T}}\int_0^T \Big(\phi\left(w_{s,s+t}(w_0)\right) - \langle \mu_{s+t}, \phi\rangle\Big)dt \xrightarrow{\mathcal{D}} N(0,\sigma^2). \tag{1.10}$$

These two limit theorems are proved in Chapter 6 through a martingale approximation procedure and by applying martingale limit theorems. The idea of the martingale approximation was originally due to Gordin [36] and further developed by Knips and Varadhan [44]. In [62], a new proof of Dobrusion' s theorem was given by using similar ideas. Our approach here is different from that in [62], since we apply the martingale approximation to the associated homogenized Markov process on the extended phase space $H \times \mathbb{T}^n$, where $\mathbb{T}^n$ is equipped with the irrational rotation flow with rotation frequency α from the quasi-periodic force. It is worth mentioning that this homogenized process is not mixing since the irrational rotation flow is never mixing. Hence the method of martingale approximation for uniformly mixing time homogeneous Markov processes [48] cannot be applied in a straightforward way. Fortunately, the exponentially mixing quasi-periodic invariant measure centers the observations in the limit theorems in an appropriate way that enables us to derive a martingale approximation.

1.1.3 Rate of convergence in the limit theorems. In the time homogeneous and essentially elliptic setting, the estimates of the rate of convergence of the limit theorems were obtained for the 2D stochastic Navier-Stokes equation, which are close to being optimal [61]. Our third result shows that similar estimates hold in the inhomogeneous hypoelliptic context. Namely, we will show the following convergence rate (which is the same as in the time homogeneous case) for the strong law of large numbers: for any $s \in \mathbb{R}$ and $\varepsilon > 0$, there is an almost surely finite random time $T_{s,\varepsilon}$ such that for all $T \geq T_{s,\varepsilon}$,

$$\left|\frac{1}{T}\int_0^T \Big(\phi(w_{s,s+t}(w_0)) - \langle \mu_{s+t}, \phi\rangle\Big)dt\right| \leq CT^{-\frac{1}{2}+\varepsilon}, \tag{1.11}$$

and a Berry-Esseen type rate of convergence in the central limit theorem measured by the Kolmogorov uniform distance:

$$\sup_{z\in\mathbb{R}}\left(\xi_\sigma(z)\left|\mathbf{P}\left\{\frac{1}{\sqrt{T}}\int_0^T \Big(\phi\left(w_{s,s+t}(w_0)\right) - \langle \mu_{s+t}, \phi\rangle\Big)dt \leq z\right\} - \Phi_\sigma(z)\right|\right) \leq C_{\varepsilon_0}T^{-\frac{1}{4}+\varepsilon_0}, \tag{1.12}$$

where $\xi_\sigma \equiv 1$ for $\sigma > 0, \xi_0(z) = 1 \wedge |z|$, and $\Phi_\sigma(z)$ is the distribution function of the centered Gaussian distribution. Here $\varepsilon_0 \in (0, \frac{1}{4})$ is a constant depending on the mixing rate of the quasi-periodic invariant measure and the convergence rate for the Birkhoff ergodic sums of the irrational rotation for a particular observable function involving the quasi-periodic invariant measure.

These estimates are proved in Chapter 6. The general idea is to derive first the estimates for the approximating martingale and then pass to inequalities (1.11) and (1.12) by invoking the martingale approximation. In particular, estimate (1.12) is derived from a combination of several ideas from [61] with an important Berry-Esseen type result for martingales from [39]. We stress here that estimate (1.12) requires a detailed analysis on the Hölder regularity of a particular induced observable function (see Proposition 6.13) on the torus $\mathbb{T}^n$ involving the quasi-periodic invariant measure. This is a main feature in our context that is different from the time homogeneous case. We will also see that the result from [39] enables us to show (1.12) for observable functions that are allowed to have exponential growth at infinity, while in the time homogeneous case [61] the estimate is valid for observable functions growing at infinity not faster than a polynomial. However, due to the interaction between the mixing of the solution process and the irrational rotation inherited from the quasi-periodic force, the convergence rate in our context cannot be arbitrarily close to the likely optimal rate usually obtained in the time homogeneous case. Indeed, from our proof, one can see that estimate (1.12) is a mixture of the convergence rate to the variance of the time inhomogeneous solution process, and the convergence rate of the Birkhoff sum for the irrational rotation with a particular observable function involving the quasi-periodic invariant measure that is related to the Diophantine approximation property of the frequency. The combination of the two rates prohibits the possibility of ε_0 being arbitrarily small. Besides, estimate (1.11) is obtained by combining the martingale approximation and the Borel-Cantelli lemma with an estimation on the convergence rate for the moments of the time average of the observations centered by the quasi-periodic invariant measure.

1.2 Additional Historical Backgrounds

In this section, we give a brief description about the history of the statistical theory of stochastic Navier-Stokes equations. Since the literature on this topic is too vast a subject to review here, we limit ourselves to topics that are closely related to the works in the present dissertation.

1.2.1 Two Dimensional Turbulence.

The following system of two dimensional incompressible Navier-Stokes equations describes the evolution of the velocity field $u(t,x)$ of a given fluid subject to a deterministic external force $\overline{F}(x,t)$,

$$\begin{aligned} \partial_t u(x,t) - \nu\Delta u(x,t) + u(x,t)\cdot\nabla u(x,t) + \nabla P(x,t) &= \overline{F}(x,t), \\ \nabla\cdot u(x,t) &= 0, \\ u(x,s) &= u_0(x), \end{aligned} \tag{1.13}$$

where ν is the viscosity, P is the pressure and x takes values on some bounded two dimensional domain U with appropriate boundary conditions. When $U = \mathbb{T}^2$, the equation is equivalent to the corresponding vorticity equation (1.1) without noise, see the next chapter on this point.

The standard theory of two dimensional turbulence is the study of the dynamical behavior of the Navier-Stokes equation when the viscosity is small $0 < \nu \ll 1$ for many external force $\overline{F}$ (for some forces the turbulence can be absent for any value of the viscosity [51]), which corresponds to the case of large Reynolds number. Various conjectures and experimental discoveries have not been rigorously treated due to the chaotic nature of the system in the turbulent regime [32]. It is widely believed that the statistical behavior of the turbulence should be described by a particular invariant measure supported on the attractor of the system, the so called Sinai-Ruelle-Bowen (SRB) measure [30]. However, the existence of such a canonical measure is still a challenging open problem.

The approach to the above scenario is accessible when the external force is random. For example, we may take

$$\overline{F}(x,t) = \overline{f}(x,t) + \sum_{i=1}^{\infty} b_i e_i(x)\dot{W}_i(t), \tag{1.14}$$

where b_i are constants, $\{e_i\}$ is an orthonormal basis of the phase space, $W_i(t)$ are independent standard Wiener processes and $\dot{W}_i(t)$ are white noise processes. The Navier-Stokes equation (1.13) is then a stochastic differential equation with such a random external force. Under additional mild conditions, it can be proved that the Markov solution process of (1.13) has a unique ergodic invariant measure (which may further be shown to be mixing, see the next section). The advantage of working with such a random external force is that we have a canonical invariant measure to analyze the statistical behavior of the turbulence, though in a mean value sense (by taking average over the samples). The idea that the turbulence should be described by the Navier-Stokes equations driven by a random force dates back to Kolmogorov [64].

The concept of determining modes is useful when analyzing the Navier-Stokes dynamics [30]. It is roughly the first N modes of the equation (when rewriting the equation in terms of the eigenvectors of the linear part) that determines the dynamics of the system, where N depends on the viscosity ν and the strength of the external force, and N goes to infinity if ν tends to 0. Since we are considering the Navier-Stokes equations on the torus, one can use Fourier expansions in terms of the eigenvectors of the Laplacian to rewrite the equation as well as the force (1.14) in the Fourier space. We say that the stochastic equation (1.13) is **elliptic** if all $b_i \neq 0$ in (1.14), and it is called **essentially elliptic** if $b_i \neq 0$ for all $1 \leq i \leq N$, where the first N modes are precisely the determining modes. It is called **hypoellptic** if the noise is extremely degenerate such that it does not act on all determining modes, i.e., $N_0 := \#\{i \geq 1 : b_i \neq 0\} < N$, and N_0 is independent of ν and the strength of the external force. These terms correspond to ellipticity (or essential ellipticity, hypoellipticity) of the Fokker-Planck-Kolmogorov equation associated with (1.13) that governs the evolution of the distribution of the solution process, which is a deterministic parabolic equation in an infinite dimensional Hilbert space [5].

1.2.2 The Time Homogeneous Setting.

When the deterministic force $\overline{f}(x,t)$ in (1.14) (or $f(x,t)$ in (1.1)) is independent of time, the solution process is a time homogeneous Markov process with corresponding Markov transition semigroup $\mathcal{P}_t$. A probability measure μ is unique if $\mathcal{P}_t^*\mu = \mu$ for $t \geq 0$. The existence of such a measure is usually guaranteed by the Krylov-Bogolyubov theorem and a compactness argument from the dissipative nature of the Navier-Stokes system. However the uniqueness and mixing need more effort.

There are mainly two approaches to prove unique ergodicity. One way is the Doob-Khasminskii type argument that combines the (asymptotic) strong Feller property that shows ergodic invariant measures have disjoint supports, with topological irreducibility which shows that any invariant measure is supported on the whole phase space. Another method is the coupling approach that shows the contraction property of the transition operators by choosing an appropriate copy of the solution process and analyzing the coupled process. One usually proves (exponential) mixing first by coupling methods and then shows the unique ergodicity as a consequence. While the Doob-Khasminskii type argument is powerful when proving unique ergodicity, it gives less information on (exponential) mixing. It is worth mentioning that in the hypoellptic case when the driving noise is white in time, there is no existing work that proves the unique ergodicity and mixing through coupling methods.

The Essentially Elliptic Case. When the noise is not degenerate or acting on all determining modes, the statistical properties of the stochastic Navier-Stokes system have been extensively studied over the decades, see for example [3, 4, 24, 25, 27, 28, 38, 45, 46, 47, 55, 56, 58, 61, 34, 8] and references therein. See also the monograph [48] for a summary of the existing results on unique ergodicity and mixing, as well as limit theorems with convergence rates. The results in the cited works require the random forcing to act on all determining modes. In particular, the dimension of the random forcing goes to infinity as viscosity approaches to zero. The methods involved are either the coupling argument or the Doob-Khasminskii type argument with strong Feller property and topological irreducibility.

The Hypoelliptic Case. When the noise is extremely degenerate and not all determining modes are activated, a major breakthrough in this case was made by Hairer and Mattingly in their seminal work [40], where they introduced the asymptotic strong Feller property to show unique ergodicity of the Navier-Stokes equation driven by an extremely degenerate noise. To obtain this asymptotic smoothing property, they developed a theory of infinite dimensional Malliavin calculus and established an infinite dimensional Hörmander type theorem. This asymptotic smoothing effect also allowed them to develop an infinite dimensional Harris-like theorem in [41], to prove the exponential convergence to the unique invariant measure under the Wasserstein metric ρ induced by (1.2). The result in [41] in turn led to a proof of the weak law of large numbers and the central limit theorem in [50]. The results in [40, 41, 50] are independent of the strength of the external force and the viscosity but require a range condition when the time independent force is nonzero, i.e., they require the range of the deterministic force $\overline{f}(x)$ to be contained in the span of the noise. Later in [35] the authors proved the unique ergodicity without the range condition but the exponential mixing without the range condition remains unproved. We prove this exponential mixing without the range condition in the present work and extend the results to the time inhomogeneous case.

1.2.3 The Time Inhomogeneous Setting.

When the deterministic force $\overline{f}(x,t)$ depends on time, the only existing result is the work of Da Prato and Debussche [16] where they considered a time periodic force $\overline{f}(x,t)$. They proved the unique ergodicity and exponential mixing in the essentially elliptic case by a coupling argument. Unique ergodicity in the hypoelliptic case was claimed but without proof in [16] with a range condition on $\overline{f}(x,t)$, i.e., the range of the deterministic force should be contained in the span of the noise for all $t \in \mathbb{R}$. We take an approach completely different from that in [16] and prove the exponential mixing, unique ergodicity, as well as limit theorems and convergence rates in the hypoellptic case, without any condition on the deterministic force other than some spatial regularity.

We also note that the existence of a continuous (in time) periodic invariant measure was obtained in [16] by first disintegrating an invariant measure of the associated homogenized Markov process whose existence relies on the Krylov-Bogolyubov theorem, and then proving the existence of a continuous version. The extension of this method to the quasi-periodic case becomes challenging when proving the continuity. In this dissertation we extend the Harris-like theorem to the time inhomogeneous setting, which allows us to prove the existence of a unique quasi-periodic invariant measure as a fixed point in the space of continuous quasi-periodic measures. This fixed point naturally has the continuity and can be shown to have a Hölder continuity if the deterministic force does, which plays an important role in the convergence rate of the central limit theorem.

The concept of quasi-periodic invariant measure with its ergodicity and mixing for stochastic 2D Navier-Stokes equation, is introduced in the present work. However, the quasi-periodic invariant measure for ordinary stochastic differential equations has been introduced in [31]. In fact, the probability measure valued path μ_t satisfying (1.3) has been studied by Dynkin in [21, 22] where it is called an entrance law. In this context, a quasi-periodic invariant measure is an entrance law that possesses additional dynamical structure inherited from the quasi-periodic force.

The martingale limit theory has been widely studied as a generalization of the limit theory for the sum of independent random variables, see the monograph [39]. It becomes a powerful tool when studying the limit theorems of Markov processes, in the case that the Gordin's martingale approximation is available. This approach has been summarized in [48, 61] in the case of uniformly mixing time homogeneous Markov processes with applications to randomly forced PDE's including the 2D Navier-Stokes equation. As we mentioned above, the result in [48, 61] cannot be applied directly here due to the time inhomogeneity and non-mixing feature of the homogenized process.

For discrete time inhomogeneous Markov chains, progress on limit theorems has been made over the decades, which is best summarized in [20]. However, the problem and method we consider here, are different from that in [20] where they mainly focus on the local limit theorems in the stationary regime for discrete time Markov chains and take a different perspective. Furthermore, despite the uniform boundedness condition on the observables, the uniform elliptic condition in [20] does not hold in our hypoelliptic setting so that the results cannot be applied. In fact, we are not aware of any existing results on limit theorems for continuous time inhomogeneous Markov processes, as well as applications to randomly forced PDE's both in elliptic and hypoelliptic settings.

Chapter 2. Settings and Preliminaries

In this section, we give definitions, basic settings on the equation and technical preliminaries that will be used throughout the work.

2.1 Basic Settings on The Equation

This section consists of a brief description of the two dimensional Navier-Stokes equations driven by a time dependent deterministic quasi-periodic force and a random force that is spatially regular and white in time. We first give two definitions related to quasi-periodic functions. Let (M, d) be a metric space with metric d and $C_b(\mathbb{R}, M)$ the space of bounded continuous functions endowed with the uniform convergence topology generated by the following metric

$$\underline{d}(q_1, q_2) = \sup_{t\in\mathbb{R}} d(q_1(t), q_2(t)).$$

Definition 2.1 (Quasi-periodic functions)**.** A function $q \in C_b(\mathbb{R}, M)$ is quasi-periodic with frequency $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n) \in \mathbb{R}^n$ if there is $Q \in C(\mathbb{T}^n, M)$ such that

$$q(t) = Q(\alpha t) = Q(\alpha_1 t, \alpha_2 t, \cdots, \alpha_n t), \tag{2.1}$$

where $\alpha_1, \alpha_2, \cdots, \alpha_n$ are rationally independent real numbers and $\mathbb{T}^n = \mathbb{R}^n/(2\pi)\mathbb{Z}^n$ is the n-dimensional torus.

The following Diophantine condition is closely related to the convergence rate in limit theorems for solutions of the stochastic Navier-Stokes system.

Definition 2.2 (Diophantine condition)**.** A frequency $\alpha \in \mathbb{T}^n$ is said to satisfy a Diophantine condition if there exist $K > 0$ and $A > n$ such that

$$\operatorname{dist}(k \cdot \alpha, \mathbb{Z}) \geq \frac{K}{\|k\|^A}, \tag{2.2}$$

for all $k \in \mathbb{Z}^n$ with $\|k\| \neq 0$, where $\|k\| := \max_{1 \leq i \leq n} |k_i|$, and $k \cdot \alpha = k_i\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$.

2.1.1 The Equation.

We consider the incompressible Navier-Stokes equations on the two dimensional torus $\mathbb{T}^2$, which describes the evolution of an incompressible fluid. They are usually written as follows in terms of the velocity field

$$\begin{aligned} \partial_t u(x,t) - \nu\Delta u(x,t) + u(x,t)\cdot\nabla u(x,t) + \nabla P(x,t) = \overline{F}(x,t), \\ \nabla\cdot u(x,t) = 0, \\ u(x,s) = u_0(x). \end{aligned}$$

Here $u(x,t) \in \mathbb{R}^2$ for $x \in \mathbb{T}^2, t \in \mathbb{R}$ is the velocity field of the fluid. $s \in \mathbb{R}$ is the initial time and $u_0(x)$ is the initial condition. $\nu > 0$ is the kinematic viscosity, P is the pressure and $\overline{F}$ is the external force. We assume that u_0 and $\overline{F}$ have zero mean when averaged on the torus, so that the solution will have zero mean for all time.

In the two dimensional case, it is convenient to consider the following equivalent equations for the scalar vorticity field $w = \nabla \wedge u = \partial_1 u_2 - \partial_2 u_1$, which is obtained by taking curl on the equations for the velocity,

$$\partial_t w(x,t) - \nu\Delta w(x,t) + B(\mathcal{K}w, w)(x,t) = F(x,t), \quad w(x,s) = w_0(x). \tag{2.3}$$

Here $B(\mathcal{K}w, w) = \mathcal{K}w \cdot \nabla w$ is the nonlinear term, where $u = \mathcal{K}w$ and $\mathcal{K}$ is the Biot-Savart integral operator that recovers the velocity from the vorticity through the conditions

$$w = \nabla \wedge u, \quad \nabla \cdot u = 0, \quad \int_{\mathbb{T}^2} u(x)dx = 0.$$

We will consider equation (2.3) on the space H of square integrable functions over $\mathbb{T}^2$ that have zero mean, i.e., $H := \left\{w \in \mathrm{L}^2\left(\mathbb{T}^2, \mathbb{R}\right) : \int_{\mathbb{T}^2} w(x)dx = 0\right\}$, where the norm is denoted by $\|\cdot\|$ and the inner product is $\langle\cdot,\cdot\rangle$. We also define the interpolation spaces $H_s = \left\{w \in H^s\left(\mathbb{T}^2, \mathbb{R}\right) : \int_{\mathbb{T}^2} w(x)dx = 0\right\}$ and the corresponding norms $\|\cdot\|_s$ by $\|w\|_s = \left\|(-\Delta)^{s/2} w\right\|$.

2.1.2 The Forcing.

The external force $F(x,t)$ in equation (2.3) consists of a time dependent deterministic part and a random part that is white in time and regular in space. More specifically, we assume $F(x,t) = f(x,t) + G\frac{dW(t)}{dt}$.

The notation $\frac{dW(t)}{dt}$ denotes symbolically the time derivative of the two-sided Brownian motion $W(t)$ and equation (2.3) is understood in the integral form as a stochastic differential equation in the Ito sense. Here the standard d dimensional two-sided Brownian motion $W(t)$ is obtained as follows. Let $W^{\pm}(t)$ be two independent standard d dimensional Brownian motion, then define

$$W(t) := \begin{cases} W^{+}(t) \ , \ t \geq 0, \\ W^{-}(-t) \ , \ t < 0. \end{cases}$$

The sample space is denoted by $(\Omega, \mathcal{F}, \mathbf{P})$, where $\Omega = \{\omega \in C(\mathbb{R}, \mathbb{R}^d) : \omega(0) = 0\}$ endowed with the compact open topology, $\mathcal{F}$ is the Borel σ-algebra and $\mathbf{P}$ is the Wiener measure associated with the Brownian motion W. Denote by $\mathcal{F}_t$ the filtration of σ-algebras generated by $W(t)$. The coefficient of the noise is a bounded linear operator $G : \mathbb{R}^d \to H_\infty := \bigcap_{s>0} H_s$, such that $Ge_i = g_i$, where $\{e_i\}$ is the standard basis of $\mathbb{R}^d$ and $g_i \in H_\infty$ for $i = 1, 2, \cdots, d$. Then the noise can be expressed as

$$GW(t) = \sum_{i=1}^{d} g_i W_i(t). \tag{2.4}$$

Also for integer $k \geq 0$, let $\mathcal{B}_k := \sum_{i=1}^{d} \|g_i\|_k^2$ be the various norms of the energy input from the noise.

We assume that the deterministic force $f \in C_b(\mathbb{R}, H_2)$ is quasi-periodic with frequency $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$. Then by Definition 2.1 there is a function $\Psi \in C(\mathbb{T}^n; H_2)$ such that $f(t,x) = \Psi(\alpha t, x)$. The regularity condition imposed on f is to ensure the spatial regularity of the solution that is needed to show the asymptotic smoothing effect of the dynamics.

2.1.3 Well-Posedness of 2D Stochastic Navier-Stokes Equations. Having introduced the external force $F(x,t)$ in the previous subsection, we can now rewrite equation (2.3) as an Ito's stochastic differential form

$$dw(x,t) - \nu\Delta w(x,t)dt + B(\mathcal{K}w, w)(x,t)dt = f(x,t)dt + GdW(t), \quad w(x,s) = w_0(x). \tag{2.5}$$

Under the conditions imposed on the forcing term as in the previous subsection, the existence and uniqueness of the solution to equation (2.5) is well known, see for example [7, 48] and references therein. To be specific, we have the following

Proposition 2.3. *Assume $f \in C_b(\mathbb{R}, H_2)$ and $g_i \in H_\infty$ (see (2.4)) for all $1 \le i \le d$. Then for any initial time $s \in \mathbb{R}$, every $T > s$ and $w_0 \in H$, equation (2.5) under the above conditions has a unique strong solution $w(t, \omega; s, w_0)$, $t \in [s, T]$, i.e., it solves the integral equation*

$$
\begin{aligned}
w(x,t) - w_0(x) - \int_s^t \nu \Delta w(x,t)dt + \int_s^t B(\mathcal{K}w, w)(x,t)dt \\
= \int_s^t f(x,t)dt + G(W(t) - W(s)), \ \mathbf{P} - \text{a.s.}
\end{aligned}
$$

The solution is adapted to the filtration $\mathcal{F}_t$, and generates a stochastic flow $\Phi(t, \omega; s, \cdot) : H \to H$ such that $\Phi(t, \omega; s, w_0) = w(t, \omega; s, w_0)$ for $s \le t, w_0 \in H$ and

$$
w \in C([s,T]; H) \cap C((s,T]; H_3), \ \mathbf{P} - \text{a.s.}
$$

Here by a stochastic flow $\Phi(t, \omega; s, w_0)$, we mean that it is a modification of the solution of equation (2.5) *satisfying the following conditions:*

(i) It is adapted to $\mathcal{F}_t$ and for almost all ω, $\Phi(t, \omega; s, w_0)$ is continuous in (t, s, w_0) and $\Phi(s, \omega; s, w_0) = w_0$.

(ii) For almost all ω,

$$
\Phi(t+\tau, \omega; s, w_0) = \Phi(t+\tau, \omega; t, \Phi(t, \omega; s, w_0)),
$$

for $s \le t$, $\tau > 0$ and $w_0 \in H$.

2.2 Inhomogeneous Markovian Formulations

The solution to (2.5) generates a two parameter Markov transition operator $\mathcal{P}_{s,t} : B_b(H) \to B_b(H)$ defined by

$$
\mathcal{P}_{s,t}\varphi(w_0) := \mathbf{E}[\varphi(w(t; s, w_0))], \quad \varphi \in B_b(H), \tag{2.6}
$$

where $B_b(H)$ is the space of bounded Borel measurable functions on H with the supremum norm. We denote the transition probabilities as $\mathcal{P}_{s,t}(w, A) := \mathcal{P}_{s,t}\mathbb{I}_A(w)$ for $A \in \mathfrak{B}$, the Borel σ-algebra of H, where $\mathbb{I}_A$ is the characteristic function of A. By duality, the transition operator $\mathcal{P}^*_{s,t}$ acts on the space $\mathcal{P}(H)$ of probability measures on H by

$$
\mathcal{P}^*_{s,t}\mu(A) = \int_H \mathcal{P}_{s,t}(w, A)\mu(dw), \ \text{for } \mu \in \mathcal{P}(H), A \in \mathfrak{B}.
$$

For $\eta > 0$ small, recall the metric ρ on H weighted by the Lyapunov function $e^{\eta\|w\|^2}$ as in [41],

$$\rho(w_1, w_2) = \inf_{\gamma} \int_0^1 e^{\eta\|\gamma(t)\|^2} \|\dot{\gamma}(t)\| dt, \quad \forall w_1, w_2 \in H, \tag{2.7}$$

where the infimum is taken over all differentiable paths γ connecting w_1 and w_2. This metric naturally induces a Wasserstein metric (allowed to take values in $[0, \infty]$) on $\mathcal{P}(H)$ by

$$\rho(\mu_1, \mu_2) = \inf_{\mu \in \mathcal{C}(\mu_1, \mu_2)} \int_{H \times H} \rho(u, v) \mu(du dv), \tag{2.8}$$

where $\mathcal{C}(\mu_1, \mu_2)$ is the set of couplings of $\mu_1, \mu_2 \in \mathcal{P}(H)$. The subset

$$\mathcal{P}_1(H) := \{\mu \in \mathcal{P}(H) : \rho(\mu, \delta_0) < \infty\} \tag{2.9}$$

is complete under the metric ρ [12, 65], where δ_0 is the Dirac measure at 0. For this Wasserstein metric, the following Monge-Kantorovich duality is well-known [12, 65],

$$\rho\left(\mu_1, \mu_2\right) = \sup_{\mathrm{Lip}_\rho(\phi) \leq 1} \left| \int \phi(x) \mu_1(dx) - \int \phi(x) \mu_2(dx) \right|, \quad \forall \mu_1, \mu_2 \in \mathcal{P}_1(H), \tag{2.10}$$

where $\mathrm{Lip}_\rho(\phi)$ is the Lipschitz constant of the function ϕ on H endowed with the metric ρ.

Remark. We use the metric (2.8) on $\mathcal{P}(H)$ to measure the convergence to the quasi-periodic invariant measure. The reason for working with the Wasserstein metric is that the transition probabilities in infinite dimensional systems are likely to be mutually singular, especially when the strong Feller property does not hold [41] (which is the case when the driving noise is extremely degenerate). Hence the convergence to the invariant measure often fails under the total variation metric and one would like to replace it by a weaker Wasserstein metric.

We now give the definition of a quasi-periodic invariant measure.

Definition 2.4 (Quasi-periodic invariant measures)**.** A quasi-periodic invariant measure of system (2.5) is a quasi-periodic function $\mu \in C(\mathbb{R}, \mathcal{P}(H))$ that is invariant under the Markov transition operators

$$\int_H \mathcal{P}_{s,t} \varphi(w) \mu_s(dw) = \int_H \varphi(w) \mu_t(dw), \quad s \leq t, \quad \varphi \in B_b(H),$$

or equivalently

$$\mathcal{P}^*_{s,t} \mu_s = \mu_t, \quad s \leq t.$$

It is called uniquely ergodic if such measure is unique. It is exponentially mixing under the Wasserstein metric (2.8) if there is a constant $\varpi > 0$ such that

$$\rho(\mathcal{P}_{s,t}^{*}\mu, \mu_t) \leq Ce^{-\varpi(t-s)}\rho(\mu, \mu_s), \quad \forall s \leq t, \mu \in \mathcal{P}(H).$$

2.3 Homogenization through Skew-Product

To deal with the time inhomogeneity, we adopt a classical method that has been widely used in the study of non-autonomous problems arising from deterministic differential equations and dynamical systems. Let $\mathcal{H}(f)$ be the closure in $C_b(\mathbb{R}, H_2)$ of $\{f(\cdot + s)|s \in \mathbb{R}\}$, the set of time translations of f. The set $\mathcal{H}(f)$ is called the hull of f, which is compact since f is quasi-periodic. It is the minimal invariant set of the translation group $\{T(t)\}$ acting on $C_b(\mathbb{R}; H_2)$ by $(T(t)g)(s) = g(t+s)$, which is called the Bebutov shift flow. The Bebutov shift dynamics has the advantage of capturing the nonautonomy caused by the quasi-periodicity of f and the compactness of the hull $\mathcal{H}(f)$ allows us to apply tools from dynamical systems to analyze related problems. For each $g \in H(f)$, there is a unique solution $w_{s,t,g}(\omega, w_0)$ of (2.5) by replacing f with g, which is called the process corresponding to problem (2.5) with time symbol g [14]. The homogenization process associated with $\Phi_{s,t}(\omega, w_0)$ is then the homogeneous Markov process in the extended phase space $H \times \mathcal{H}(f)$ defined by $S(t, \omega, w_0, g) := (w_{0,t,h}(\omega, w_0), T(t)g)$.

In the quasi-periodic case, it turns out that $\mathcal{H}(f) = \{\Psi(\alpha t + h_0, x)|h_0 \in \mathbb{T}^n\}$ [14], where $\Psi \in C(\mathbb{T}^n, H_2)$ is the function corresponding to f. Hence it is more convenient to work with $\mathbb{T}^n$ instead of $\mathcal{H}(f)$. The irrational rotation flow $\beta_t h := h + \alpha t$ on torus corresponds to the Bebutov shift $T(t)$ on $\mathcal{H}(f)$ through the continuous map Ψ. Instead of working on $H \times \mathcal{H}(f)$, we will study the associated homogeneous Markov process $X_{s,t}(w_0, h_0)$ on $H \times \mathbb{T}^n$ given by the solution of the following equation

$$\begin{cases} dw(t,x) + B(\mathcal{K}w, w)(t,x)dt = \nu\Delta w(t,x)dt + \Psi(\beta_t, x)dt + GdW(t), \\ d\beta_t = \alpha dt, \\ w(s) = w_0, \beta_s = h_0. \end{cases} \tag{2.11}$$

Note that the solution

$$X_{s,t}(w_0, h_0) = (\Phi_{s,t,\beta_{-s}h_0}(w_0), \beta_{t-s}h_0)$$

for $(w_0, h_0) \in H \times \mathbb{T}^n$, where $\Phi_{s,t,\beta_{-s}h_0}(w_0)$ is the solution to (2.5) by replacing $f(t,x) = \Psi(\alpha t, x)$ with $\Psi(\alpha t - \alpha s + h_0, x)$. We will use $w_{s,t,h}(w_0)$ to denote the solution of (2.5) with $f(t,x)$ replaced by $\Psi(\beta_t, x)$ with $\beta_t h = h + \alpha t$.

Let $\mathcal{P}_{s,t,h}\phi(w_0) = \mathbf{E}\phi\left(\Phi_{s,t,h}(w_0)\right)$ be the transition operator corresponding to the time symbol h. In particular, $\mathcal{P}_{s,t,0} = \mathcal{P}_{s,t}$ as defined in (2.6). It follows from the uniqueness of solution that the following translation identity holds:

$$\mathcal{P}_{s+\tau,t+\tau,h} = \mathcal{P}_{s,t,\beta_\tau h},\ \tau \in \mathbb{R},\ h \in \mathbb{T}^n. \tag{2.12}$$

For $\varphi \in B_b(H \times \mathbb{T}^n)$, the Markov transition operator associated with the homogenized process is given by

$$P_{s,t}\varphi(w_0, h_0) = \mathbf{E}\varphi\left(X_{s,t}\right) = \mathbf{E}\varphi\left(\Phi_{s,t,\beta_{-s}h_0}(w_0), \beta_{t-s}h_0\right) = \mathcal{P}_{s,t,\beta_{-s}h_0}\varphi(\cdot, \beta_{t-s}h_0)(w_0). \tag{2.13}$$

In view of the translation identity (2.12) and time homogeneity of $X_{s,t}$, we can assume the initial time of the homogenized process $s = 0$ for simplicity.

2.4 Path-wise Random Quasi-periodic Solutions

The existence of a quasi-periodic invariant measure implies the existence of a solution process of (2.5), whose distribution changes in time quasi-periodically. However, information about path-wise dynamics is not easy to obtain from the quasi-periodic invariant measure. We now give the following definition of the random quasi-periodic solution in the path-wise sense. And we will show that when the viscosity is large, then the dissipation dominates and the system (2.5) has a trivial dynamics, in the sense that it has a globally stable random quasi-periodic solution which supports the unique quasi-periodic invariant measure.

Definition 2.5 (Random quasi-periodic solutions)**.** A random quasi-periodic solution of system (2.5) is a progressively measurable stochastic process $w^*(t,\omega)$ defined on $\mathbb{R} \times \Omega$ that satisfies the following property:

(i) (Invariance property) $\Phi(t,\omega; s, w^*(s,\omega)) = w^*(t,\omega)$, for almost all ω;

(ii) (Quasi-periodicity) $w^*(t,\theta_{-t}\omega)$ is a quasi-periodic function for almost every sample ω. That is, there is a function $Q : \mathbb{T}^n \times \Omega \to H$, where $Q(h,\omega)$ is continuous in $h \in \mathbb{T}^n$ for almost every $\omega \in \Omega$ and measurable in ω for each fixed h, such that for every $t \in \mathbb{R}$,

$$w^*(t,\theta_{-t}\omega) = Q(\alpha t, \omega), \quad \mathbf{P} - \text{a.s.}$$

Here $\theta : \mathbb{R} \times \Omega \to \Omega$ given by

$$\theta_t \omega(\cdot) := \omega(t + \cdot) - \omega(t) \tag{2.14}$$

is the Wiener shift.

Chapter 3. Main Results

In this subsection, we formulate the main results of the present paper in details. We recall a condition on the structure of the degenerate noise from [42]. To do that, define the set A_∞ by setting

$$A_1 = \{g_l : 1 \leq l \leq d\}, A_{k+1} = A_k \cup \{\widetilde{B}(h, g_l) : h \in A_k, g_l \in A_1\}, \text{ and } A_\infty = \overline{\text{span}(\cup_{k \geq 1} A_k)}, \tag{3.1}$$

where $\widetilde{B}(u, w) = -B(\mathcal{K}u, w) - B(\mathcal{K}w, u)$ is the symmetrized nonlinear term. These sets reflect the mechanism of the propagation of the extremely degenerate noise to the phase space that yields a smoothing effect of the dynamics. With these notations, the condition we need on the structure of the noise is an infinite dimensional version of the Hörmander's Lie bracket condition $A_\infty = H$. In finite dimensional settings, the Hörmander's Lie bracket condition ensures the invertibility of the Malliavin matrix and shows the smoothness of the density of the transition probabilities of a degenerate stochastic differential equation. This hypoelliptic theory was first extended to the stochastic Navier-Stokes equation driven by an extremely degenerate noise in [57]. Then in [40] the authors proved the unique ergodicity of the stochastic Navier-Stokes equation in the hypoelliptic setting by introducing the asymptotic strong Feller property. As was shown in [40, 42], it is notable that the noise is allowed to be extremely degenerate to have $A_\infty = H$, for example it can be excited only through four directions.

The following Theorem 3.1-3.4 are our main results under the standing assumption:

$$f \in C_b(R, H_2) \text{ is quasi-periodic}; g_i \in H_\infty, \forall 1 \leq i \leq d; \text{ and } A_\infty = H. \tag{3.2}$$

The first result is the following unique ergodicity and exponentially mixing of the quasi-periodic invariant measure for (2.5) under the Wasserstein metric (2.8). Note that the metric ρ (2.8) weighted by the Lyapunov function $V(w) = e^{\eta \|w\|^2}$ depends on the parameter $\eta > 0$. And the following estimate (A.3) is from Appendix A.

Theorem 3.1 (Ergodicity and mixing). *There is a unique quasi-periodic invariant measure* μ_t *for* (2.5) *given by a unique map* $\Gamma \in C(\mathbb{T}^n, \mathcal{P}(H))$, *i.e.,* $\mu_t = \Gamma_{\beta_t 0}$, *and* $\mathcal{P}^*_{s,t}\mu_s = \mu_t$. *Moreover, there exists* $\eta_0 > 0$, *such that for every* $\eta \in (0, \eta_0]$, *there are constants* $C, \varpi > 0$, *such that* $\Gamma \in C(\mathbb{T}^n, \mathcal{P}_1(H))$ *and*

$$\rho(\mathcal{P}^*_{s,s+t}\mu, \mu_{s+t}) \leq Ce^{-\varpi t}\rho(\mu, \mu_s), \quad \forall s \in \mathbb{R}, t \geq 0, \mu \in \mathcal{P}(H), \tag{3.3}$$

where C, ϖ *do not depend on* s. *Furthermore,* $\Gamma \in C^{\zeta}(\mathbb{T}^n, (\mathcal{P}_1(H), \rho))$ *if* $\Psi \in C^{\gamma}(\mathbb{T}^n, H)$, *where* $\zeta = \frac{\varpi\gamma}{r+\varpi}$ *with* $r = 64c_0^6\eta^{-3}\nu^{-5} + \eta C(f, \mathcal{B}_0)$ *from estimate* (A.3).

In an equivalent form that involves the transition operator acting on observables, we have for every $\phi \in C^1_\eta$,

$$\left\| \mathcal{P}_{s,s+t,h}\phi - \int_H \phi(w)\mu_{s+t}(dw) \right\|_\eta \leq Ce^{-\varpi t}\|\phi\|_\eta. \tag{3.4}$$

Here

$$C^1_\eta := \left\{ \phi \in C^1(H) : \|\phi\|_\eta := \sup_{w \in H} e^{-\eta\|w\|^2}\big(|\phi(w)| + \|\nabla\phi(w)\|\big) < \infty \right\}.$$

Theorem 3.1 will be proved in Chapter 5, by combining a fixed point argument with the following uniform contraction property proved in Chapter 4.

Theorem 3.2. (*Contraction on* $\mathcal{P}(H)$) *There exists* $\eta_0 > 0$ *such that for* $\eta \in (0, \eta_0]$, *there are positive constants* C *and* ϖ *such that*

$$\rho(\mathcal{P}^*_{s,s+t,h}\mu_1, \mathcal{P}^*_{s,s+t,h}\mu_2) \leq Ce^{-\varpi t}\rho(\mu_1, \mu_2), \tag{3.5}$$

for every $s \in \mathbb{R}$, $t \geq 0$, $h \in \mathbb{T}^n$ *and any* $\mu_1, \mu_2 \in \mathcal{P}(H)$.

The second result is on the strong law of large numbers (SLLN) and the central limit theorem (CLT) for the solution process. The proof will be given in Chapter 6. To state the results, we first define the space of observable functions. For $\gamma \in (0, 1]$, let $C^\gamma_\eta(H)$ be the space of Hölder continuous functions with finite norms weighted by the Lyapunov function $e^{\eta\|w\|^2}$,

$$C^\gamma_\eta(H) := \{\phi : H \to \mathbb{R} : \|\phi\|_{\gamma,\eta} < \infty\}, \tag{3.6}$$

where

$$\|\phi\|_{\gamma,\eta} := \sup_{w \in H} \frac{|\phi(w)|}{e^{\eta\|w\|^2}} + \sup_{0 < \|w_1 - w_2\| \leq 1} \frac{|\phi(w_1) - \phi(w_2)|}{\|w_1 - w_2\|^\gamma \left(e^{\eta\|w_1\|^2} + e^{\eta\|w_2\|^2}\right)}.$$

Recall that $w_{s,s+t}(w_0)$ is the solution to (2.5) starting from $w_0 \in H$ at time $s \in \mathbb{R}$.

Theorem 3.3. *There is a constant $\eta_0 > 0$, such that for every $\eta \in (0, \eta_0]$, $\phi \in C^\gamma_\eta(H)$ and $w_0 \in H$, we have*

1. *SLLN for the time inhomogeneous solution process:*
$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \Big(\phi(w_{s,s+t}(w_0)) - \langle \mu_{s+t}, \phi\rangle\Big) dt \xrightarrow{\text{a.s.}} 0. \tag{3.7}$$

2. *CLT for the time inhomogeneous solution process:*
$$\lim_{T\to\infty} \frac{1}{\sqrt{T}} \int_0^T \Big(\phi(w_{s,s+t}(w_0)) - \langle \mu_{s+t}, \phi\rangle\Big) dt \overset{\mathcal{D}}{=} N(0, \sigma^2), \tag{3.8}$$

where $N(0, \sigma^2)$ a centered normal variable with variance σ^2 and $\mathcal{D}$ represents the convergence in distribution. The variance
$$\sigma^2 = \sigma^2_\phi = \lim_{T\to\infty} \frac{1}{T} \mathbf{E}\left[\int_0^T \phi(w_{s,s+t}(w_0)) - \langle \mu_{s+t}, \phi\rangle dt\right]^2,$$
where $\sigma_\phi \geq 0$ is independent of s.

The third result is an estimate of the rate of convergence in the limit theorems. The proof will also be given in Chapter 6.

Theorem 3.4. *There is a constant $\eta_0 > 0$ such that the following estimates hold.*

1. *(Rate of convergence in SLLN) Let $\varepsilon > 0$, for every integer $p \geq 3$ satisfying $2^p > 1/\varepsilon$, every $\eta \in (0, 2^{-p-1}\eta_0]$, and every $\phi \in C^\gamma_{\eta,H}(H \times \mathbb{T}^n)$, $w_0 \in H$, $s \in \mathbb{R}$, there is an almost surely finite random time $T_0(\omega) \geq 1$, depending on $p, \varepsilon, \|\phi\|_{\gamma,\eta,H}, s, \|w_0\|$ such that for all $T > T_0$, we have*
$$\left|\frac{1}{T} \int_0^T \Big(\phi(w_{s,s+t}(w_0)) - \langle \mu_{s+t}, \phi\rangle\Big) dt\right| \leq C T^{-\frac{1}{2}+\varepsilon},$$
where $C > 0$ is a constant that does not depend on the above parameters. Moreover, for every $0 < \ell < \min\{2^p\varepsilon - 1, 2^{p-2} - 1\}$, there is a constant $C_p = C_p(\|\phi\|_{\gamma,\eta,H}, \ell, \varepsilon)$ such that
$$\mathbf{E}T_0^\ell \leq C_p e^{2^{p+1}\eta\|w_0\|^2}.$$

2. *(Rate of convergence in CLT) Assume $\Psi \in C^\gamma(\mathbb{T}^n, H)$ and the frequency α satisfies the Diophantine condition* (2.2) *with constant A and dimension n (of the torus). Let Φ_σ be the distribution function of $N(0, \sigma^2)$. Also let $\Lambda = \frac{\varpi}{5(2-\gamma)}$, $\zeta = \frac{\varpi\gamma}{r+\varpi}$ and $\overline{\gamma}_0 = \frac{\Lambda\zeta}{5(\Lambda+r)(2-\gamma)}$, where ϖ is the mixing rate from Theorem 3.1 and $r = 64c_0^6\eta^{-3}\nu^{-5} + \eta C(f, \mathcal{B}_0)$ is the constant from* (A.3)*.*

(1). For any integer $p \geq 2$, $\eta \in (0, 2^{-p-1}\eta_0]$, and $\phi \in C^\gamma_\eta(H)$ with $\sigma^2_\phi > 0$, and $w_0 \in H$, there

are constants $C_p = C_p(\|\phi\|_{\gamma,\eta}, \|w_0\|) > 0$ *and* $T_0 > 0$ *such that for all* $T \geq T_0$,

$$\sup_{z\in\mathbb{R}} \left| \mathbf{P}\left\{ \frac{1}{\sqrt{T}} \int_0^T \Big(\phi(w_{s,s+t}(w_0)) - \langle \mu_{s+t}, \phi \rangle \Big) dt \leq z \right\} - \Phi_{\sigma_\phi}(z) \right| \leq C_p \left(T^{-\frac{1}{4}} + T^{-\frac{2^{p-2}}{2^p+1}} + T^{-\frac{2^{p-1}\overline{\gamma}_0}{(2^p+1)(A+n)}} \right),$$

(2). For $\eta \in (0, 2^{-7}\eta_0]$ *and* $\phi \in C^\gamma_\eta(H)$ *such that* $\sigma^2_\phi = 0$, *and* $w_0 \in H$, *there is a constant* $C = C(\|\phi\|_{\gamma,\eta}, \|w_0\|) > 0$ *such that for all* $T \geq 1$,

$$\sup_{z\in\mathbb{R}} (|z| \wedge 1) \left| \mathbf{P}\left\{ \frac{1}{\sqrt{T}} \int_0^T \Big(\phi(w_{s,s+t}(w_0)) - \langle \mu_{s+t}, \phi \rangle \Big) dt \leq z \right\} - \Phi_0(z) \right| \leq C \left(T^{-\frac{1}{4}} + T^{-\frac{\overline{\gamma}_0}{2(A+n)}} \right).$$

The following result shows that if the viscosity is large, then the dynamics of system (2.5) is actually trivial. Let c_0 be the constant from Ladyzhenskaya's inequality

$$||w||^2_{\mathrm{L}^4(\mathbb{T}^2)} \leq c_0 \|w\|_1 \|w\|. \tag{3.9}$$

And let $G = \sqrt{\|f\|^2_\infty/\nu^4 + \mathcal{B}_0/\nu^3}$ be the Grashof number for the whole system, where $\|f\|_\infty = \sup_{t\in\mathbb{R}} \|f(t)\|$. The following theorem will be proved in Chapter 7. Note that we do not need the assumption $A_\infty = H$ here.

Theorem 3.5. *Let* $\overline{\eta} > 0$. *Assume* $\Psi \in C^\gamma(\mathbb{T}^n, H)$. *If*

$$Gc_0 \leq \sqrt{1/2}, \quad \textit{and} \;\; \nu^3 > 8(n+\overline{\eta})c_0^2\mathcal{B}_0\gamma^{-1}, \tag{3.10}$$

then there exists a random quasi-periodic solution $w^*(t,\omega)$ *of equation* (2.5) *in the sense of Definition 2.5, where the function* $Q(h,\omega)$ *associated with* $w^*(t,\theta_{-t}\omega)$ *has a continuous (with respect to* h*) modification that is* η*-Hölder continuous for all* $0 < \eta < \frac{\overline{\eta}\gamma}{2(n+\overline{\eta})}$. *Moreover,* $w^*(t,\omega)$ *exponentially attracts all other solutions both in forward and pullback times. The law of* $w^*(t,\omega)$ *gives the unique quasi-periodic invariant measure.*

Chapter 4. Contraction on The Space of Probability Measures

In this chapter, we prove Theorem 3.2, the "uniform fiber-wise" contraction property (3.5) of the transition operators $\{\mathcal{P}^*_{s,t,h}\}_{h\in\mathbb{T}^n}$ when acting on $\mathcal{P}(H)$. The result can be regarded as an extension of the Harris-like theorem for infinite dimensional systems first established by Hairer and Mattingly [41] in the time homogeneous setting, where they proved the existence of a unique invariant measure that is exponentially mixing under the Wasserstein metric ρ. The idea behind their result dates back to the early work of Dobelin [18] and Harris [37] for finite dimensional systems. The proof here will be accomplished through a combination of a particular type of irreducibility, a Lyapunov structure and a gradient estimate on the transition operator that requires an infinite dimensional Malliavin calculus and a Hörmander-type condition. We first prove these three ingredients through Sections 4.1-4.3, and then prove the contraction as in Theorem 3.2 in Section 4.4. Note that the proofs of the Lyapunov structure and gradient estimate are essentially adaptations of the scheme developed in [40, 41] to our time inhomogeneous setting, while the irreducibility is apparently new and does not require any condition on the deterministic force other than it being time quasi-periodic and having certain spatial regularity as in the standing assumption (3.2). In particular, we do not need the range condition as in [41] and our result in the case when f is time independent verifies a conjecture made by Hairer and Mattingly [41] (see Remark 1.3) that the spectral gap (as well as unique ergodicity and exponentially mixing) holds without restrictions on f other than it be sufficiently smooth.

Recall from (2.11) that $w_{s,t,h}(w_0)$ and $\Phi_{s,t,h}(w_0)$ denote the solution of (2.5) with $f(t,x)$ replaced by $\Psi(\beta_t h, x)$ with $\beta_t h = h + \alpha t$ and

$$\sup_{t\in\mathbb{R}} \|\Psi(\beta_t h)\| = \sup_{h\in\mathbb{T}^n} \|\Psi(h)\| = \sup_{t\in\mathbb{R}} \|f(t)\| := \|f\|_\infty .$$

For $\eta > 0, 0 < r \le 1$, define the metric ρ_r on H weighted by the Lyapunov function $V(w) = e^{\eta\|w\|^2}$ as in [41],

$$\rho_r(w_1, w_2) = \inf_\gamma \int_0^1 V^r(\gamma(t))\|\dot\gamma(t)\|dt, \quad \forall w_1, w_2 \in H, \tag{4.1}$$

where the infimum is taken over all differentiable paths γ in H connecting w_1 and w_2. When $r = 1$ we have $\rho_r = \rho$ as given in (2.7).

4.1 The Lyapunov Structure

In this section we prove the following estimates, which show a Lyapunov structure that reveals the dissipation property of the Navier-Stokes system (2.5).

Proposition 4.1 (The Lyapunov Structure)**.** *Let* $V(w) = \exp(\eta\|w\|^2)$, $\eta_0 > 0$ *be the constant from* (A.1) *and* $\alpha(t) = e^{-\nu t}$*. For any* $\kappa \ge 1$*, and any* $\eta \in (0, \frac{\eta_0}{2\kappa}]$*, there is a constant* $C > 0$ *such that for any* $s \in \mathbb{R}$*, and* $(w, h) \in H \times \mathbb{T}^n$*, we have*

$$\mathbf{E}V^\kappa(\Phi_{s,s+t,h}(w)) \le CV^{\kappa\alpha(t)}(w), \tag{4.2}$$

$$\mathbf{E}\rho(0, w_{s,s+t,h}(w))^{\kappa} \le CV^{2\kappa}(w), \tag{4.3}$$

for all $t \ge 0$ *and*

$$\mathbf{E}V^{\kappa}(\Phi_{s,s+t,h}(w))(1 + \|\nabla\Phi_{s,s+t,h}(w))\xi\|) \le CV^{\kappa\alpha(t)}(w) \tag{4.4}$$

for every $t \in [0,1], h \in \mathbb{T}^n$ *and* $\xi \in H$ *with* $\|\xi\| = 1$*.*

Proof. Inequality (4.2) is a reformulation of (A.1). The estimate (4.3) follows from (A.1). Indeed, by the definition of ρ,

$$\begin{aligned}\mathbf{E}\rho(0, w_{s,s+t,h}(w))^{\kappa} &\le \mathbf{E}\|w_{s,s+t,h}(w)\|^{\kappa}V^{\kappa}(w_{s,s+t,h}(w)) \\ &\le C\mathbf{E}V^{2\kappa}(w_{s,s+t,h}(w)) \le C\exp(2\kappa\eta\|w\|^2),\end{aligned}$$

for $\eta \le \frac{\eta_0}{2\kappa}$, where η_0 is from (A.1). Estimate (A.6) shows that for any $\eta > 0$, there is a constant $C > 0$ such that

$$\|\nabla\Phi_{s,s+t,h}(w)\xi\| \le \exp\left(\nu\eta\int_s^{s+t}\|\Phi_{s,r,h}(w)\|_1^2 dr + Ct\right).$$

The estimate (4.4) then follows from (A.1) and (A.2). □

With the help of this proposition, one can show the following lemma on a contraction property under the metric (4.1) weighted by the Lyapunov function.

Lemma 4.2. *Fix any* $0 < r_0 < 1$*. For any* $r \in [r_0, 1]$*, there are constants* $\sigma \in (0,1)$ *and* $C, K > 0$ *such that*

$$\mathbf{E}\rho_r(\Phi_{s,s+t,h}(u), \Phi_{s,s+t,h}(v)) \le C\rho_r(u,v), \tag{4.5}$$

$$\mathbf{E}\rho_r(\Phi_{s,s+n,h}(u), \Phi_{s,s+n,h}(v)) \le \sigma^n\rho_r(u,v) + K, \tag{4.6}$$

for any $s \in \mathbb{R}$*,* $n \in \mathbb{N}$*,* $t \in [0,1]$*,* $h \in \mathbb{T}^n$ *and* $u, v \in H$*.*

Proof. For any $\varepsilon > 0$, there is a path γ connecting u and v such that

$$\rho_r(u,v) \le \int_0^1 V^r(\gamma(t))\|\dot\gamma(t)\|dt \le \rho_r(u,v) + \varepsilon.$$

Let $\tilde{\gamma}(\tau) = \Phi_{s,s+t,h}(\gamma(\tau))$ for $t \in [0,1]$. Then by (4.4),

$$\begin{aligned}
\mathbf{E}\rho_r(\Phi_{s,s+t,h}(u), \Phi_{s,s+t,h}(v)) &\le \mathbf{E}\int_0^1 V^r(\tilde{\gamma}(\tau))\|\dot{\tilde{\gamma}}(\tau)\|d\tau \\
&\le \mathbf{E}\int_0^1 V^r(\tilde{\gamma}(\tau))\left\|\nabla\Phi_{s,s+t,h}(\gamma(\tau))\frac{\dot{\gamma}(\tau)}{\|\dot{\gamma}(\tau)\|}\right\|\|\dot{\gamma}(\tau)\|d\tau \\
&\le C\int_0^1 V^{r\alpha(t)}(\gamma(\tau))\|\dot{\gamma}(\tau)\|d\tau \le C\rho_r(u,v) + C\varepsilon.
\end{aligned}$$

The inequality (4.5) then follows since ε is arbitrary. Now let R be large enough so that $CV^{r\alpha(1}(u) \le \sigma V^r(u)$ for some $\sigma \in (0,1)$ and for all u with $\|u\| \ge R$. Then

$$\begin{aligned}
\mathbf{E}\rho_r(\Phi_{s,s+1,h}(u), \Phi_{s,s+1,h}(v)) &\le C\int_0^1 V^{r\alpha(1)}(\gamma(\tau))\|\dot{\gamma}(\tau)\|d\tau \\
&\le \sigma\rho_r(u,v) + C\int_0^1 \mathbb{I}_{B_R(0)}(\gamma(\tau))V^{r\alpha(1)}(\gamma(\tau))\|\dot{\gamma}(\tau)\|d\tau + \sigma\varepsilon \\
&\le \sigma\rho_r(u,v) + CV(R)\int_0^1 \mathbb{I}_{B_R(0)}(\gamma(\tau))\|\dot{\gamma}(\tau)\|d\tau + \varepsilon. \qquad (4.7)
\end{aligned}$$

Note that if we set u_0, v_0 to be the points where γ first enters and last exits the ball $B_R(0)$, then we have

$$\int_0^1 \mathbb{I}_{B_R(0)}(\gamma(\tau))\|\dot{\gamma}(\tau)\|d\tau \le \int_0^1 \mathbb{I}_{B_R(0)}(\gamma(\tau))V^r(\gamma(\tau))\|\dot{\gamma}(\tau)\|d\tau < \rho_r(u_0,v_0) + \varepsilon. \qquad (4.8)$$

The second inequality of (4.8) is true since otherwise, one has

$$\int_0^1 \mathbb{I}_{B_R(0)}(\gamma(\tau))V^r(\gamma(\tau))\|\dot{\gamma}(\tau)\|d\tau \ge \rho_r(u_0,v_0) + \varepsilon,$$

so that $\int_0^1 V^r(\gamma(\tau))\|\dot{\gamma}(\tau)\|d\tau \ge \rho_r(u,v) + \varepsilon$ by observing that

$$\int_0^1 \mathbb{I}_{B_R^c(0)}(\gamma(\tau))V^r(\gamma(\tau))\|\dot{\gamma}(\tau)\|d\tau \ge \rho_r(u,u_0) + \rho_r(v,v_0),$$

which contradicts the choice of the path $\gamma(\tau)$. Note that by considering the straight line that connects u_0 and v_0, we have $\rho_r(u_0,v_0) \le 2RV^r(R)$. Hence

$$\int_0^1 \mathbb{I}_{B_R(0)}(\gamma(\tau))\|\dot{\gamma}(\tau)\|d\tau \le 2RV^r(R) + \varepsilon. \qquad (4.9)$$

By arbitrariness of ε, we conclude from (4.7) and (4.9) that for $K = 2CRV^{r+1}(R)$,

$$\mathbf{E}\rho_r(\Phi_{s,s+1,h}(u), \Phi_{s,s+1,h}(v)) \le \sigma\rho_r(u,v) + K.$$

For $n > 1$, we apply the Markov property. Observe that

$$\begin{aligned}
&\mathbf{E}\left[\rho_r(\Phi_{s,s+n,h}(u), \Phi_{s,s+n,h}(v))|\mathcal{F}_{s+(n-1)}\right] \\
&= \mathbf{E}\left[\rho_r(\Phi_{s+(n-1),s+n,h}(\Phi_{s,s+(n-1),h}(u)), \Phi_{s+(n-1),s+n,h}(\Phi_{s,s+(n-1),h}(v)))|\mathcal{F}_{s+(n-1)}\right] \\
&\leq \sigma\rho_r(\Phi_{s,s+(n-1),h}(u), \Phi_{s,s+(n-1),h}(v)) + K,
\end{aligned}$$

and

$$\begin{aligned}
&\mathbf{E}\left[\rho_r(\Phi_{s,s+n,h}(u), \Phi_{s,s+n,h}(v))|\mathcal{F}_{s+(n-2)}\right] \\
&\leq \sigma\left(\mathbf{E}\left[\rho_r(\Phi_{s+(n-2),s+(n-1),h}(\Phi_{s,s+(n-2),h}(u)), \Phi_{s+(n-2),s+(n-1),h}(\Phi_{s,s+(n-2),h}(v)))|\mathcal{F}_{s+(n-2)}\right] + K\right) \\
&\leq \sigma^2\rho_r(\Phi_{s,s+(n-2),h}(u), \Phi_{s,s+(n-2),h}(v)) + (\sigma+1)K.
\end{aligned}$$

By iterating the procedure, we obtain that

$$\mathbf{E}\rho_r(\Phi_{s,s+n,h}(u), \Phi_{s,s+n,h}(v)) \leq \sigma^n\rho_r(u,v) + K\sum_{j=0}^{n-1}\sigma^j,$$

which completes the proof. □

4.2 The Irreducibility

In this section we prove a particular form of irreducibility (slightly stronger than the usual topological irreducibility) that concerns the small positive probability of being contraction of the dynamics on the phase space. It deals with the contraction (3.5) at intermediate scales.

Proposition 4.3 (The Irreducibility)**.** *For any $R, \varepsilon, T > 0$, and $r \in (0,1]$, there exists $a > 0$ so that*

$$\inf_{\|w_1\|,\|w_2\|\leq R}\ \sup_{\Gamma\in\mathcal{C}\left(\mathcal{P}^*_{s,s+T,h}\delta_{w_1}, \mathcal{P}^*_{s,s+T,h}\delta_{w_2}\right)} \Gamma\left\{(w_1', w_2') \in H\times H : \rho_r(w_1', w_2') < \varepsilon\right\} \geq a, \tag{4.10}$$

for all $s \in \mathbb{R}, h \in \mathbb{T}^n$. Here the metric ρ_r is given by (4.1).

The proof is based on the topological irreducibility (Lemma 4.4) which is a consequence of the celebrated controllability results of Agrachev and Sarychev [1, 2], and a parabolic regularizing property (Lemma 4.5) of the Navier-Stokes equations. We first state and prove the two lemmas and then give the proof of Proposition 4.3 at the end.

The first lemma states that uniformly for the initial positions in any given compact set, the system has a positive probability to enter into any neighborhood of a given position.

Lemma 4.4. *For any compact set $K \subset H$, any $\delta, T > 0$, and $v \in H$, there is $a_0 := a_0(K, \delta, T, v) > 0$ such that*

$$\inf_{w\in K, h\in\mathbb{T}^n} \mathcal{P}_{0,T,h}(w, B_\delta(v)) \geq a_0.$$

Proof. By the well known controllability argument of Agrachev and Sarychev [1, 2] (see also [35] and Section B.3 in the appendix) and the condition $A_\infty = H$, it follows that for each $h \in \mathbb{T}^n$ the Navier-Stokes equation (2.5) with the deterministic force f replaced by $\Psi(\beta_t h)$ is approximately controllable in H. Hence it is topologically irreducible, i.e., for all $w, v \in H$, $\delta > 0, T > 0$ there is $\varepsilon = \varepsilon(w, v, \delta, T, h) > 0$ such that

$$\mathcal{P}_{0,T,h}\mathbb{I}_{B_{\delta/2}(v)}(w) = \mathcal{P}_{0,T,h}(w, B_{\delta/2}(v)) \geq \varepsilon.$$

Now let $0 \leq L_\delta \leq 1$ be a Lipschitz continuous function on H such that $L_\delta(w) = 1$ on $B_{\delta/2}(v)$ and $L_\delta(w) = 0$ outside of $B_\delta(v)$. Then

$$\mathcal{P}_{0,T,h}\mathbb{I}_{B_\delta(v)}(w) \geq \mathcal{P}_{0,T,h}L_\delta(w) \geq \mathcal{P}_{0,T,h}\mathbb{I}_{B_{\delta/2}(v)}(w) \geq \varepsilon.$$

The Lipschitz continuity of L_δ and estimates (A.3) and (A.4) imply that the function

$$(w, h) \to \mathcal{P}_{0,T,h}L_\delta(w)$$

is continuous. Hence on the compact set $K \times \mathbb{T}^n$, it attains the infimum at some (w_0, h_0), which implies

$$\inf_{(w,h)\in K\times\mathbb{T}^n} \mathcal{P}_{0,T,h}\mathbb{I}_{B_\delta(v)}(w) \geq \inf_{(w,h)\in K\times\mathbb{T}^n} \mathcal{P}_{0,T,h}L_\delta(w) = \mathcal{P}_{0,T,h_0}L_\delta(w_0) \geq \varepsilon(w_0, v, \delta, T, h_0).$$

The proof is then completed by taking $a_0 = \varepsilon(w_0, v, \delta, T, h_0)$. □

The second lemma is a result of the parabolic regularizing effect, which states that any given bounded set has a positive probability to be mapped into a compact set by the solution map.

Lemma 4.5. *For any $R > 0, T > 0$, there are $R_1 := R_1(R, T) > 0$ and $a_1 := a_1(R, R_1, T) > 0$ such that*

$$\inf_{w\in B_R(0), h\in\mathbb{T}^n} \mathcal{P}_{0,T,h}\left(w, \overline{B}^{H_1}_{R_1}(0)\right) \geq a_1,$$

where $B_R(0)$ is the ball in H centered at 0 with radius R, and $\overline{B}^{H_1}_{R_1}(0)$ is the closed ball in H_1 centered at 0 with radius R_1.

Proof. Let $w_{0,t,h}(w_0)$ be the solution of (2.5) with the deterministic force f replaced by $\Psi(\beta_t h)$. We eliminate the additive noise by subtracting from (2.5) the following auxiliary Ornstein-Uhlenbeck process

$$dV_t = \nu\Delta V_t dt + GdW_t, \quad V(0) = 0. \tag{4.11}$$

It is readily seen that the process $u_{t,h} := w_{0,t,h}(w_0) - V_t$ solves the equation

$$\partial_t u = \nu\Delta u - B\left(\mathcal{K}\left(u + V_t\right), u + V_t\right) + \Psi(\beta_t h), \quad u_{0,h} = w_0. \tag{4.12}$$

For $\delta > 0, T > 0$, consider the event that the d-dimensional Wiener process has small amplitude on a finite interval

$$\Omega_{\delta,T} = \{\omega \in \Omega : |W(\omega,t)| \leq \delta, \quad \text{for all} \quad t \in [0,T]\}, \tag{4.13}$$

where $|\cdot|$ denote the usual norm in $\mathbb{R}^d$. Then there exists $a_1 = a_1(\delta,T)$ such that

$$\mathbf{P}(\Omega_{\delta,T}) \geq a_1 > 0. \tag{4.14}$$

By the properties of stochastic convolutions, one can show that the Ornstein-Uhlenbeck process V_t has the following property on the event $\Omega_{\delta,T}$.

Lemma 4.6. *[23] For any $\delta > 0, T > 0$, there exists a positive deterministic constant $\varepsilon_{\delta,T}$ such that $\varepsilon_{\delta,T} \to 0$ as $\delta \to 0$ for T fixed, and*

$$\sup_{t\in[0,T]} \|V(t,\omega)\|_1 \leq \varepsilon_{\delta,T}, \ \textit{for all} \quad \omega \in \Omega_{\delta,T}. \tag{4.15}$$

Proof. Note that

$$V_{s+t} = \int_s^{s+t} e^{\nu\Delta(t+s-r)} GdW(r) = \sum_{i=1}^d \int_s^{s+t} e^{\nu\Delta(t+s-r)} g_i dW_i(r)$$

is the stochastic convolution that solves equation (4.11) with initial condition $V(s) = 0$. From the integration by parts formula for the Wiener integral (see for example [63]), we have

$$\begin{aligned} V_{s+t} &= \sum_{i=1}^d \left(g_i W_i(s+t) - e^{\nu\Delta t} g_i W_i(s) + \int_s^{s+t} \nu\Delta e^{\nu\Delta(t+s-r)} g_i W_i(r) dr \right) \\ &= \sum_{i=1}^d \left(g_i \left(W_i(s+t) - W_i(s)\right) + \int_0^t \nu\Delta e^{\nu\Delta(t-r)} g_i \left(W_i(r+s) - W_i(s)\right) dr \right). \end{aligned}$$

Note that $g_i \in H_\infty$ for each $1 \le i \le d$. Hence on $\Omega_{\delta,T}$, we have

$$\sup_{t\in[s,s+T]} \|V(t,\omega)\|_1 \le C(1+T) \sup_{t\in[s,s+T]} |W(s+t)-W(s)| \le C(1+T)\delta,$$

where the constant C is independent of δ, T and $s \in \mathbb{R}$. The lemma then follows by taking $\varepsilon_{\delta,T} = C(1+T)\delta$ and $s = 0$. □

Return to the proof of Lemma 4.5. Taking H inner product with $2u$ on both sides of (4.12) and applying standard estimates on the nonlinear term (see for example [40]), we have (we write u or u_t for $u_{t,h}$ for notational simplicity and denote $\|\Psi\|_\infty = \sup_{h\in\mathbb{T}^n} \|\Psi(h)\|$)

$$\begin{aligned}
\partial_t \|u\|^2 &= \langle \nu\Delta u, 2u\rangle - \langle B\left(\mathcal{K}\left(u+V_t\right), u+V_t\right), 2u\rangle + \langle \Psi(\beta_t h), 2u\rangle \\
&\le -2\nu\|u\|_1^2 - 2\langle B\left(\mathcal{K}V_t, V_t\right), u\rangle - 2\langle B\left(\mathcal{K}u, V_t\right), u\rangle + \frac{4}{\nu}\|\Psi\|_\infty^2 + \nu\|u\|^2 \\
&\le -\nu\|u\|_1^2 + C\|V_t\|_1^2\|u\| + C\|u\|^2\|V_t\|_1 + \frac{4}{\nu}\|\Psi\|_\infty^2 \qquad (4.16) \\
&\le C\|V_t\|_1^2\|u_t\|^2 + C(\|V_t\|_1^4 + \|\Psi\|_\infty^2).
\end{aligned}$$

Hence the Gronwall's inequality and Lemma 4.6 yield

$$\begin{aligned}
\|u_t\|^2 &\le \|w_0\|^2 \exp\left(C\int_0^t \|V_\tau\|_1^2 d\tau\right) + C\int_0^t \left(\|V_\tau\|_1^2 + \|\Psi\|_\infty^2\right) \exp\left(C\int_\tau^t \|V_r\|_1^2 dr\right) d\tau \\
&\le C(\|w_0\|^2 + 1) \qquad (4.17)
\end{aligned}$$

for all $t \in [0,T]$ and $\omega \in \Omega_{\delta,T}$, where the last constant $C = C(\varepsilon_{\delta,T}, T, \|\Psi\|_\infty, \nu)$. Using this estimate and integrating (4.16), we also have on $\Omega_{\delta,T}$,

$$\nu\int_0^T \|u_t\|_1^2 dt \le \|w_0\|^2 + \int_0^T \left(C\|V_t\|_1^2\|u_t\| + C\|u_t\|^2\|V_t\|_1 + \frac{4}{\nu}\|\Psi\|_\infty^2\right) dt \le C(\|w_0\|^2 + 1) \quad (4.18)$$

with constant C only depending on $\varepsilon_{\delta,T}, T, \|\Psi\|_\infty, \nu$.

Differentiating $t\|u_t\|_1^2$ with respect to t and using (4.12) yields

$$\begin{aligned}
\partial_t(t\|u_t\|_1^2) &= \|u\|_1^2 + 2t\langle \partial_t u, -\Delta u\rangle \\
&= \|u\|_1^2 - 2t\nu\|u\|_2^2 + 2t\left\langle B\left(\mathcal{K}u, u\right) + B\left(\mathcal{K}V_t, V_t\right) + B\left(\mathcal{K}u, V_t\right) + B\left(\mathcal{K}V_t, u\right) - \Psi(\beta_t h), \Delta u\right\rangle.
\end{aligned}$$

It follows from standard estimates on the nonlinear term and interpolation inequalities that

$$\begin{aligned}
|\langle B\left(\mathcal{K}u, u\right), \Delta u\rangle| &\le C\|u\|\|u\|_{\frac{3}{2}}\|u\|_2 \le C\|u\|\|u\|^{\frac{1}{4}}\|u\|_2^{\frac{3}{4}}\|u\|_2 \\
&\le C(\nu)\|u\|^{10} + \frac{\nu}{5}\|u\|_2^2,
\end{aligned}$$

and

$$|\langle B\left(\mathcal{K}V_t,V_t\right),\Delta u\rangle| \le C\|V_t\|_1^2\|u\|_2 \le C(\nu)\|V_t\|_1^4+\frac{\nu}{5}\|u\|_2^2,$$
$$|\langle B\left(\mathcal{K}u,V_t\right),\Delta u\rangle| \le C\|V_t\|_1\|u\|_1\|u\|_2 \le C(\nu)\|u\|_1^2\|V_t\|_1^2+\frac{\nu}{5}\|u\|_2^2,$$
$$|\langle B\left(\mathcal{K}V_t,u\right),\Delta u\rangle| \le C\|V_t\|_1\|u\|_1\|u\|_2 \le C(\nu)\|u\|_1^2\|V_t\|_1^2+\frac{\nu}{5}\|u\|_2^2,$$
$$|\langle \Psi(\beta_t h),\Delta u\rangle| \le \|\Psi\|_\infty\|u\|_2 \le C(\nu)\|\Psi\|_\infty^2+\frac{\nu}{5}\|u\|_2^2.$$

Therefore

$$\partial_t(t\|u_t\|_1^2) \le \|u\|_1^2 + C(\nu)t\left(\|u\|^{10}+\|V_t\|_1^4+\|u\|_1^2\|V_t\|_1^2+\|\Psi\|_\infty^2\right).$$

Integrating this inequality from 0 to T, using Lemma 4.6 and estimates (4.17)-(4.18), one has

$$\|u_{T,h}\|_1^2 \le C(1+\|w_0\|^2)^5,$$

on $\Omega_{\delta,T}$, where C depends on $\varepsilon_{\delta,T}, T, \|\Psi\|_\infty, \nu$, but is independent of $h \in \mathbb{T}^n$. Hence on $\Omega_{\delta,T}$ we have

$$\|w_{0,T,h}(w_0)\|_1^2 \le 2(\|u_{T,h}\|_1^2+\|V_T\|_1^2) \le C(1+\|w_0\|^{10}),$$

which completes the proof by taking $R_1=\sqrt{C(1+R^{10})}$. □

We are now in a position to give a proof of Proposition 4.3.

Proof of Proposition 4.3. Fix any $v \in H$. For any $\delta, T>0$, $h\in\mathbb{T}^n$, and $w \in B_R(0)$, we have by the Chapman-Kolmogorov relation, the translation identity (2.12) that

$$\mathcal{P}_{0,T,h}(w,B_\delta(v)) = \int_H \mathcal{P}_{\frac{T}{2},T,h}(y,B_\delta(v))\mathcal{P}_{0,\frac{T}{2},h}(w,dy) = \int_H \mathcal{P}_{0,\frac{T}{2},\beta_{\frac{T}{2}}h}(y,B_\delta(v))\mathcal{P}_{0,\frac{T}{2},h}(w,dy).$$

It follows from Lemma 4.5 that there are $R_1 := R_1(R,T/2)>0$, and $a_1 := a_1(R,R_1,T/2)>0$ such that

$$\inf_{w\in B_R(0),h\in\mathbb{T}^n} \mathcal{P}_{0,T/2,h}\left(w,\overline{B}_{R_1}^{H_1}(0)\right) \ge a_1.$$

Combining this with Lemma 4.4 with $K=\overline{B}_{R_1}^{H_1}(0)$, we have the existence of $a_0 := a_0(K,\delta,T,v)>0$ such that

$$\begin{aligned}\inf_{w\in B_R(0),h\in\mathbb{T}^n}\mathcal{P}_{0,T,h}(w,B_\delta(v)) &\ge \inf_{w\in B_R(0),h\in\mathbb{T}^n}\int_K \mathcal{P}_{0,\frac{T}{2},\beta_{\frac{T}{2}}h}(y,B_\delta(v))\mathcal{P}_{0,\frac{T}{2},h}(w,dy)\\ &\ge a_1 \inf_{y\in K,h\in\mathbb{T}^n}\mathcal{P}_{0,\frac{T}{2},h}(y,B_\delta(v)) \ge a_0a_1.\end{aligned}$$

Therefore by the translation identity (2.12), one has

$$\mathcal{P}_{s,s+T,h}(w, B_\delta(v)) = \mathcal{P}_{0,T,\beta_s h}(w, B_\delta(v)) \geq a_0 a_1, \quad \forall w \in B_R(0), s \in \mathbb{R}, h \in \mathbb{T}^n.$$

Now for any $\varepsilon > 0$, we choose $\delta = \frac{\varepsilon}{2V(\|v\|+\varepsilon)}$, here $V(w) = e^{\eta\|w\|^2}$ for $w \in H$ is the Lyapunov function from Proposition 4.1. Then $\widetilde{B}_\delta(v) := B_\delta(v) \times B_\delta(v) \subset \{(w_1', w_2') \in H \times H | \rho_r(w_1', w_2') < \varepsilon\}$. Hence

$$\begin{aligned}
&\inf_{w_1,w_2 \in B_R(0)} \sup_{\Gamma \in \mathcal{C}\left(\mathcal{P}^*_{s,s+T,h}\delta_{w_1}, \mathcal{P}^*_{s,s+T,h}\delta_{w_2}\right)} \Gamma\left\{(w_1', w_2') \in H \times H : \rho_r(w_1', w_2') < \varepsilon\right\} \\
&\geq \inf_{w_1,w_2 \in B_R(0)} \mathcal{P}^*_{s,s+T,h}\delta_{w_1} \otimes \mathcal{P}^*_{s,s+T,h}\delta_{w_2}(\widetilde{B}_\delta(v)) \\
&= \inf_{w_1,w_2 \in B_R(0)} \mathcal{P}_{s,s+T,h}(w_1, B_\delta(v))\mathcal{P}_{s,s+T,h}(w_2, B_\delta(v)) \geq (a_0 a_1)^2.
\end{aligned}$$

The proof is complete with $a = (a_0 a_1)^2$. ☐

4.3 The Gradient Estimate

In this subsection, we will show that the transition operator $\mathcal{P}_{s,t,h}$ has the following gradient estimate.

Proposition 4.7. *Assume $A_\infty = H$. Then for every $\eta > 0$ and $a > 0$ there exists constants $C = C(\eta, a) > 0$ such that*

$$\|\nabla \mathcal{P}_{s,s+t,h}\phi(w)\| \leq C \exp(p\eta\|w\|^2) \left(\sqrt{\left(\mathcal{P}_{s,s+t,h}|\phi|^2\right)(w)} + e^{-at}\sqrt{\left(\mathcal{P}_{s,s+t,h}\|\nabla\phi\|^2\right)(w)}\right) \tag{4.19}$$

for some $p \in (0,1)$, for every Frechet differentiable function ϕ, every $w \in H, h \in \mathbb{T}^n, s \in \mathbb{R}$ and $t \geq 0$. Here $C(\eta, a)$ does not depend on initial condition (s, w) and ϕ.

Proof. The proof is a combination of inequality (4.27), Proposition 4.16 and Proposition 4.17 below. ☐

The general scheme for the proof of the above gradient inequality is quite standard in the time homogeneous case after the groundbreaking works [40, 41, 42, 57]. However, there is no known proof for the time inhomogeneous case in the literature, hence we supply a proof in this section, using the same arguments as that in the time homogeneous case. We first apply the integration by parts formula from the theory of Malliavin calculus, to transfer the variation on the initial condition in a solution to a variation v on the Wiener path. The problem of obtaining estimate (4.19) is then reduced to finding an appropriate v with bounded cost to approximately compensate the variation on the initial condition, so that the error of the two variations in the solution goes to 0 as time goes to infinity.

The invertibility of the Malliavin matrix is crucial when constructing such a desired control v. However, it is not easy to verify the invertibility in the infinite dimensional case, hence the inverse of its Tikhonov regularization is taken into consideration. Moreover, since the noise here is extremely degenerate, the unstable directions of the system are not directly forced. Hence one needs an infinite dimensional version of Lie bracket condition $A_\infty = H$ as in Hörmander's theorem to ensure the propagation of the noise to those unstable directions and obtain a spectral estimate of the Malliavin matrix to control the dynamics on the determining modes. Since there is no existing proof of the spectral property of the Malliavin matrix in the time inhomogeneous setting, we give a proof here.

4.3.1 The Malliavin matrix. In this subsection, we recall several facts about the Malliavin matrix and give a specific description about the construction of the control v. To introduce the Malliavin derivative of the solution process, we first consider its linearized equations.

As in [40, 57], the linearized flow $J_{\tau,r,h}\xi \in C([\tau,t];H) \cap \mathrm{L}^2((\tau,t];H_1)$ is the solution to the equation

$$\partial_r J_{\tau,r,h}\xi = \nu\Delta J_{\tau,r,h}\xi + \widetilde{B}(w_{s,r,h}, J_{\tau,r,h}\xi), \quad r > \tau, \quad J_{\tau,\tau,h}\xi = \xi, \tag{4.20}$$

for any $r > \tau \geq s$ and $\xi \in H$, where $\widetilde{B}(u,w) = -B(\mathcal{K}u, w) - B(\mathcal{K}w, u)$, where $w_{s,r,h} = w_{s,r,h}(w_0)$ is the solution of (2.5) by replacing the force $f(t)$ with $\Psi(\beta_t h)$, see (2.11).

It is also helpful to consider the time-reversed, H-adjoint $U_h^{t,(\cdot)}(r)$ of $J_{r,t,h}(\cdot)$ to analyze the Malliavin matrix. Here we use the notation $U_h^{t,\varphi}(r)$ to emphasize that the time t is the initial time and $\varphi \in H$ is the initial data, and the process $U_h^{t,\varphi}(r)$ runs backward in time for $s \leq r \leq t$. It follows that $U_h^{t,\varphi}(r) \in C([s,t];H) \cap \mathrm{L}^2([s,t);H_1)$ is the unique solution to the backward random PDE

$$\begin{cases} \partial_r U_h^{t,\varphi}(r) = -\nu\Delta U_h^{t,\varphi}(r) + B\left(\mathcal{K}w_{s,r,h}, U_h^{t,\varphi}(r)\right) + C\left(\mathcal{K}U_h^{t,\varphi}(r), w_{s,r,h}\right), \quad s \leq r < t, \\ U_h^{t,\varphi}(t) = \varphi. \end{cases} \tag{4.21}$$

Here $C(\mathcal{K}(\cdot), w_{s,r,h})$ is the adjoint of $B(\mathcal{K}(\cdot), w_{s,r,h})$ determined by the relation

$$\langle B(\mathcal{K}u, w_{s,r,h}), v\rangle = \langle C(\mathcal{K}v, w_{s,r,h}), u\rangle.$$

The second derivative $K_{\tau,t,h}$ of $w_{s,r,h}$ with respect to its initial condition is the solution of the following equation

$$\begin{cases} \partial_t K_{\tau,t,h}(\xi,\xi') = \nu\Delta K_{\tau,t,h}(\xi,\xi') + \widetilde{B}(w_{s,t,h}, K_{\tau,t,h}(\xi,\xi')) + \widetilde{B}(J_{\tau,t,h}\xi', J_{\tau,t,h}\xi), \\ K_{\tau,\tau,h}(\xi,\xi') = 0. \end{cases} \tag{4.22}$$

By the variation of constants formula $K_{\tau,t,h}(\xi,\xi')$ is given by

$$K_{\tau,t,h}\left(\xi,\xi'\right) = \int_\tau^t J_{r,t,h}\widetilde{B}\left(J_{\tau,r,h}\xi', J_{\tau,r,h}\xi\right) dr.$$

Note that the solution $w_{s,t,h}(\omega, w_0)$ is a functional of the two sided Wiener process via the Ito map $\Phi^{w_0}_{s,t,h} : C([s,t],\mathbb{R}^d) \to H$ with $w_{s,t,h}(\omega, w_0) = \Phi^{w_0}_{s,t,h}\left(W_{[s,t]}\right)$, where $W_{[s,t]}$ is the restriction of the Wiener process on $[s,t]$. The Cameron-Martin space associated with the Wiener space $(\Omega, \mathcal{F}, \mathbf{P})$ is

$$\mathcal{CM} = \left\{V \in \mathrm{L}^2\left(\mathbb{R},\mathbb{R}^d\right) : \partial_t V \in \mathrm{L}^2\left(\mathbb{R},\mathbb{R}^d\right),\ V(0) = 0\right\},$$

endowed with the norm $\|V\|^2_{\mathcal{CM}} := \int_{\mathbb{R}} |\partial_t V(t)|^2_{\mathbb{R}^d} dt$, which is a Hilbert space isometric to $\mathcal{CM}' = \mathrm{L}^2(\mathbb{R},\mathbb{R}^d)$. As in Section 4.1 in [42], for any $V \in \mathcal{CM}$, denote the directional derivative of the H valued random variable $w_{s,t,h}$ along the direction V as

$$\mathrm{D}\Phi^{w_0}_{s,t,h}V := \lim_{\varepsilon\to 0} \frac{\Phi^{w_0}_{s,t,h}(W+\varepsilon V) - \Phi^{w_0}_{s,t,h}(W)}{\varepsilon},$$

which exists and satisfies

$$\mathrm{D}\Phi^{w_0}_{s,t,h}V = \int_s^t J_{r,t,h}GV'(r)dr. \tag{4.23}$$

Now for any $v \in \mathcal{CM}'$, define $V(t) = \int_0^t v(t)dt$. Then $D^v w_{s,t,h} = D^v\Phi^{w_0}_{s,t,h} := \mathrm{D}\Phi^{w_0}_{s,t,h}V$ is called the Malliavin derivative of the random variable $w_{s,t,h}$ in the direction v. Since $D^v w_{s,t,h}$ is a (random) bounded linear operator from $\mathcal{CM}'$ to H, by Riesz's representation theorem, there exists a random element, $Dw_{s,t,h} \in \mathcal{CM}' \otimes H$ such that for every $v \in \mathcal{CM}'$,

$$D^v w_{s,t,h} = \langle Dw_{s,t,h}, v\rangle_{\mathcal{CM}'} = \int_{\mathbb{R}} (D_r w_{s,t,h})\, v(r)dr.$$

The random element $Dw_{s,t,h}$ is the Malliavin derivative of $w_{s,t,h}$, which can be regarded as a stochastic process $(D_r w_{s,t,h})_{r\in\mathbb{R}}$ with values in $\mathbb{R}^d \otimes H$. From equation (4.23), we see that $D_r w_{s,t,h} = J_{r,t,h}G$ for $r \in [s,t]$ and $D_r w_{s,t,h} = 0$ for $r \in \mathbb{R}\setminus[s,t]$. The operator $D : \mathrm{L}^2(\Omega,\mathbb{R}) \otimes H \to$

$L^2_{\text{ad}}(\Omega, \mathcal{F}_t, \mathcal{CM}') \otimes H$ is actually a closed unbounded linear operator, which is called the Malliavin derivative. Here L^2_{ad} is the space of L^2 functions adapted to the filtration $\mathcal{F}_t$ [42].

For any $t \geq \tau \geq s$, define the random operator $A_{\tau,t,h} : \mathcal{CM}' \to H$ by

$$A_{\tau,t,h}v := \langle Dw_{s,t,h}, v\mathbb{I}_{[\tau,t]}\rangle_{\mathcal{CM}'} = \int_\tau^t J_{r,t,h}Gv(r)dr,$$

and its adjoint $A^*_{\tau,t,h}$ by the relation $\langle A_{\tau,t,h}v, u\rangle_H = \langle v, A^*_{\tau,t,h}u\rangle_{\mathcal{CM}'}$. Then the Malliavin matrix is defined as $M_{\tau,t,h} := A_{\tau,t,h}A^*_{\tau,t,h}$. We have for $\xi \in H$,

$$\langle M_{\tau,t,h}\xi, \xi\rangle = \sum_{i=1}^d \int_\tau^t \langle J_{r,t,h}g_i, \xi\rangle^2 dr = \sum_{i=1}^d \int_\tau^t \left\langle g_i, U_h^{t,\xi}(r)\right\rangle^2 dr, \tag{4.24}$$

where the second identity is due to the fact that $U_h^{t,(\cdot)}(\tau)$ is the adjoint of $J_{\tau,t,h}$ in H, which has been proved in [42, 57].

For any $v \in \mathcal{CM}'$, denote by $v_{\tau,t} = v\mathbb{I}_{[\tau,t]}$ the restriction of v on the interval $[\tau, t]$. Set the error

$$\mathfrak{R}_{t,h} = J_{s,t,h}\xi - A_{s,t,h}v_{s,t} \tag{4.25}$$

caused by the infinitesimal variation on the Wiener path W by v that is used to compensate the variation on the initial condition of the solution process. Applying the integration by parts formula [59] for the Malliavin derivative, we have for any Fréchet differentiable $\varphi : H \to \mathbb{R}$, any initial condition $w_0 \in H$ and $\alpha := t - s \geq 0$

$$\begin{aligned}
\langle \nabla \mathcal{P}_{s,t,h}\varphi(w_0), \xi\rangle &= \mathbf{E}\left\langle \nabla\Big(\varphi\big(w_{s,t,h}(w_0)\big)\Big), \xi\right\rangle = \mathbf{E}\Big(\big(\nabla\varphi\big)\big(w_{s,t,h}(w_0)\big)J_{s,t,h}\xi\Big) \\
&= \mathbf{E}\Big(\big(\nabla\varphi\big)\big(w_{s,t,h}(w_0)\big)A_{s,t,h}v_{s,t}\Big) + \mathbf{E}\Big(\big(\nabla\varphi\big)\big(w_{s,t,h}(w_0)\big)\mathfrak{R}_{t,h}\Big) \\
&= \mathbf{E}\Big(D^{v_{s,t}}\varphi\big(w_{s,t,h}(w_0)\big)\Big) + \mathbf{E}\Big(\big(\nabla\varphi\big)\big(w_{s,t,h}(w_0)\big)\mathfrak{R}_{t,h}\Big) \\
&= \mathbf{E}\left(\varphi\big(w_{s,t,h}(w_0)\big)\int_s^t v(r)dW(r)\right) + \mathbf{E}\Big(\big(\nabla\varphi\big)\big(w_{s,t,h}(w_0)\big)\mathfrak{R}_{t,h}\Big) \qquad (4.26) \\
&\leq \left(\mathbf{E}\left|\int_s^{s+\alpha} v(r)dW(r)\right|^2\right)^{1/2} \sqrt{\mathcal{P}_{s,s+\alpha,h}|\varphi|^2(w_0)} \\
&\quad + \sqrt{\mathcal{P}_{s,s+\alpha,h}\|\nabla\varphi\|^2(w_0)}\left(\mathbf{E}\|\mathfrak{R}_{s+\alpha,h}\|^2\right)^{1/2}, \qquad (4.27)
\end{aligned}$$

where we used Hölder's inequality at the last step. Fix $\|\xi\| = 1$, where $\xi \in H$ represents the direction of the variation of the solution on the initial condition. To show the gradient inequality, we will choose an appropriate random process v with sample paths in $\mathcal{CM}'$ to make sure the existence of constant $C > 0$ and $p \in (0,1)$, such that

$$\begin{aligned} &\mathbf{E}\left|\int_s^{s+\tau} v(r)dW(r)\right| < C\exp(p\eta\|w_0\|), \\ &\mathbf{E}\,\|\mathfrak{R}_{s+\tau,h}\| \le Ce^{-a\tau}\exp(p\eta\|w_0\|). \end{aligned} \tag{4.28}$$

for some $a > 0$ and all $\tau \ge 0, w_0 \in H$. Note once we fixed the initial time s, the values of v before s do not affect the gradient estimate (4.27). Hence we will set $v(r) = 0$ for $r < s$ and mainly focus on the construction of v after the initial time s.

The proof of Proposition 4.7 is then reduced to finding such an appropriate control v, which involves the inverse of the Malliavin matrix. However, it is unclear if the Malliavin matrix is invertible or not in the present infinite dimensional setting, therefore we consider its Tikhonov regularization $\widetilde{M}_{\tau,t,h} := M_{\tau,t,h} + \beta$ for small constant $\beta > 0$, which is invertible. For integer values $n \ge 0$, define $J_n = J_{s+n,s+n+1,h}$, $A_n = A_{s+n,s+n+1,h}$, $M_n = A_n A_n^*$, $\widetilde{M}_n = \beta + M_n$. Note that we omit the dependence on h for notational simplicity. The process v is then recursively defined as

$$v(r) = \begin{cases} A_{2n}^*\widetilde{M}_{2n}^{-1}J_{2n}\mathfrak{R}_{s+2n} & \text{for } r \in [s+2n, s+2n+1), n \ge 0, \\ 0 & \text{for } r \in [s+2n+1, s+2n+2), n \ge 0, \end{cases} \tag{4.29}$$

where $\mathfrak{R}_s = \xi$, and $\mathfrak{R}_t = J_{s,t,h}\xi - A_{s,t,h}v_{s,t}$. The definition is not circular since the construction of $v(r)$ for $r \in [s+2n, s+2n+2]$ only requires the knowledge of $\mathfrak{R}_{s+2n}$, which depends only on $v(r)$ for $r \in [s, s+2n]$. For instance, for known $\mathfrak{R}_s = \xi$, we obtain the definition of $v(r)$ for $r \in [s, s+2]$ from formula (4.29), and $\mathfrak{R}_t = J_{s,t,h}\xi - A_{s,t,h}v_{s,t,h}$ for $t \in [s, s+2]$. Then we use $\mathfrak{R}_{s+2}$ to construct $v(r)$ for $r \in [s+2, s+4]$ and iterate this procedure.

In what follows we first prove a spectral property on the Malliavin matrix in Section 4.3.2 and then give the desired estimates as in (4.28) in Section 4.3.3.

4.3.2 A Spectral Property of The Malliavin Matrix. We need the following important result about the spectral property of the Malliavin matrix over the unstable modes to have the desired controls on the dynamics. The same result in the time homogeneous setting has been obtained in [7, 40, 42, 57]. Since there is no known proof for a time inhomogeneous system such as (2.5), we supply a proof here.

Theorem 4.8. *Assume $A_\infty = H$. For any $p \ge 1$, positive α, η, n, and any orthogonal projection $\Pi : H \to H$ on a finite dimensional subspace of H, there exist $C = C(p, n, \eta, \nu, \Pi, f, \mathcal{B})$ and*

$\varepsilon_0 = \varepsilon_0(n, \alpha, \Pi, f, \mathcal{B})$ *such that*

$$\mathbf{P}\left(\langle M_{s,s+n,h}\varphi, \varphi\rangle < \varepsilon\|\varphi\|^2\right) \leq C\varepsilon^p \exp\left(\eta\,\|w_0\|^2\right)$$

holds for every (random) vector $\varphi \in H$ *satisfying* $\|\Pi\varphi\| \geq \alpha\|\varphi\|$ *almost surely, for every* $\varepsilon \in (0, \varepsilon_0), s \in \mathbb{R}$, *and for every* $w_0 \in H$, $h \in \mathbb{T}^n$.

The proof is based on the approach in Section 6 of [42] along with some estimates on the solution of the Navier-Stokes equation (2.5) and its linearization. We first prove several lemmas and then give the proof of the theorem at the end.

Fixing $T > 0$, we will consider the problem in the interval $[s, s+T]$. To avoid the singularities at the initial and terminal times, we introduce $I_\delta := [s + \frac{T}{2}, s + T - \delta]$, where $\delta = \frac{T}{4}\varepsilon^r$ for $0 < \varepsilon < 1$ and some $r > 0$ that will be determined later. Also for $\alpha \in (0,1)$, and for a given orthogonal projection $\Pi : H \to H$, we define $S_\alpha \subset H$ by

$$S_\alpha = \{\varphi \in H\backslash\{0\} : \|\Pi\varphi\| \geq \alpha\|\varphi\|\}.$$

The following estimates about the process $U_h^{t,\varphi}$ in the time homogeneous case have been proved in [42]. Since the setting is a bit different here due to time inhomogeneity, we give the proof below for the reader's convenience.

Lemma 4.9. *For any* $\delta \in (0, T/2]$, $p > 0, \eta > 0$, *one has the bound*

$$\mathbf{E}\sup_{\|\varphi\|\leq 1}\left\|U_h^{s+T,\varphi}(s+T-\delta) - e^{\delta\nu\Delta}\varphi\right\|^{2p} \leq C\delta^p \exp\left(p\eta\|w_0\|^2\right),$$

$$\mathbf{E}\sup_{\|\varphi\|\leq 1}\left\|U_h^{s+T,\varphi}(s+T-\delta) - \varphi\right\|_{-1}^{2p} \leq C\delta^p \exp\left(p\eta\|w_0\|^2\right).$$

Proof. We first reverse the time of the process by setting $\bar{w}_{s,r} = w_{s,T+2s-r,h}(w_0)$, and $\bar{U}_r = U_h^{s+T,\varphi}(T + 2s - r)$. Then $\bar{U}_r$ solves the parabolic equation

$$\begin{cases} \partial_r \bar{U}_r = \nu\Delta\bar{U}_r + B\left(\mathcal{K}\bar{w}_{s,r}, \bar{U}_r\right) - C\left(\mathcal{K}\bar{U}_r, \bar{w}_{s,r}\right), & s < r \leq s+T, \\ \bar{U}_s = \varphi. \end{cases}$$

It then follows from the variation of constant formula that,

$$\bar{U}_{s+\delta} = e^{\delta\nu\Delta}\varphi + \int_s^{s+\delta} e^{\nu\Delta(s+\delta-r)}\left[B\left(\mathcal{K}\bar{w}_{s,r}, \bar{U}_r\right) - C\left(\mathcal{K}\bar{U}_r, \bar{w}_{s,r}\right)\right] dr.$$

Since both $\|B\left(\mathcal{K}\bar{w}_{s,r}, \bar{U}_r\right)\|$ and $\|C\left(\mathcal{K}\bar{U}_r, \bar{w}_{s,r}\right)\|$ are bounded by $C\|\bar{w}_{s,r}\|_1\|\bar{U}_r\|_1$, one has

$$\left\|\bar{U}_{s+\delta} - e^{\delta\nu\Delta}\varphi\right\| \leq C \sup_{T+s-\delta\leq r\leq s+T}\|w_{s,r,h}\|_1 \int_s^{s+\delta}\|\bar{U}_r\|_1 dr.$$

To estimate $\|\bar{U}_r\|_1$, set $\zeta_r = \|\bar{U}_r\|^2 + \nu(r-s)\|\bar{U}_r\|_1^2$. As in the proof of inequality (A.16), one obtains from the equation for $\bar{U}_r$ that

$$\|\bar{U}_r\|_1 \leq \frac{C}{\sqrt{r-s}}\|\varphi\| \exp\left(\int_s^{s+T} \eta\|w_{s,r}\|_1^2 dr\right), \tag{4.30}$$

where C is a constant depending on ν, η, T. Now it follows from the estimate (A.2) that

$$\mathbf{E} \sup_{\|\varphi\|\leq 1} \left\|\bar{U}_{s+\delta} - e^{\delta\nu\Delta}\varphi\right\|^{2p} \leq C\delta^p \exp\left(p\eta\|w_0\|^2\right),$$

which is the first inequality of Lemma (4.9). The second inequality follows from the first one and the following fact

$$\left\|\bar{U}_{s+\delta} - \varphi\right\|_{-1} \leq \left\|\bar{U}_{s+\delta} - e^{\delta\nu\Delta}\varphi\right\|_{-1} + \left\|e^{\delta\nu\Delta}\varphi - \varphi\right\|_{-1} \leq \left\|\bar{U}_{s+\delta} - e^{\delta\nu\Delta}\varphi\right\| + C\delta. \qquad \square$$

The next lemma allows one to transfer the properties of φ back from the terminal time.

Lemma 4.10. *Fix any orthogonal projection* Π *of* H *onto a finite dimensional subspace of* H *spanned by elements of* H_1*. There exists a constant* $c \in (0,1)$ *such that for every* $r > 0$ *and every* $\alpha > 0$*, the event*

$$\Omega_{\delta,\Pi} := \left\{\omega \in \Omega : \varphi \in S_\alpha \Longrightarrow U_h^{s+T,\varphi}(T+s-\delta) \in S_{c\alpha} \text{ and } \|\Pi U_h^{s+T,\varphi}(T+s-\delta)\| \geq \frac{\alpha}{2}\|\varphi\|\right\}$$

satisfies $\mathbf{P}\left(\Omega_{\delta,\Pi}^c\right) \leq C\exp\left(\eta\|w_0\|^2\right)\varepsilon^p$ *for every* $p \geq 1$*. Note that here* C *depends on* $1/r$*.*

Proof. Since we have proved Lemma 4.9, this is a reformulation of Lemma 6.15 in [42]. $\square$

Since the randomness spreads over the state space through the nonlinear term, we define recursively the following sets $\{A_k\}_{k=1}^\infty$ formed by the symmetrized nonlinear term $\widetilde{B}(u,w) = -B(\mathcal{K}u, w) - B(\mathcal{K}w, u)$. Set $A_1 = \{g_k : 1 \leq k \leq d\}$, and $A_{k+1} = A_k \cup \{\widetilde{B}(h, g_l) : h \in A_k, g_l \in A_1\}$. Also define $A_\infty = \overline{\operatorname{span}(\cup_{k\geq 1} A_k)}$. Note that each A_k here, consisting of constant vector fields in H, is a subset of the k-th Hörmander bracket defined in Section 6 of [42]. To each A_n we associate a quadratic form $\mathcal{Q}_n$ by $\langle\varphi, \mathcal{Q}_n\varphi\rangle = \sum_{h\in \mathrm{A}_n}\langle\varphi, h\rangle^2$. Just as in [7, 42, 57], it is typical to apply arguments that use local time regularity to replace the analysis of non-adapted processes. Hence for $\theta \in (0,1]$ we define the following (semi-)norm for functions $g : I_\delta \to H$ by

$$\|g\|_{\theta,s} := \sup_{r,t\in I_\delta} \frac{\|g(t) - g(r)\|_s}{|t-r|^\theta}, \quad \text{and} \quad \|g\|_{\infty,s} := \sup_{t\in I_\delta}\|g\|_s.$$

Also for $g : I_\delta \to \mathbb{R}$, we use the following notation for the corresponding norms

$$\|g\|_\infty = \sup_{t\in I_\delta}|g(t)|, \quad \|g\|_\theta := \sup_{r,t\in I_\delta}\frac{|g(t) - g(r)|}{|t-r|^\theta}.$$

The following lemma is the key to prove theorem 4.8. The proof requires a technical result which roughly states that two distinct monomials in a Wiener polynomial cannot cancel each other out [7, 42, 57].

Lemma 4.11. *For every* $p \geq 1$ *and integer* $N > 0$*, there exist* $0 < \varepsilon_N < 1, r_N > 0, p_N > 0$ *and* $q_N = q_N(p) > 0$ *such that, provided that* $r \leq r_N$*, the event*

$$\Omega_{\varepsilon,N} := \left\{ \omega \in \Omega : \langle \varphi, M_{s,s+T,h}\varphi \rangle \leq \varepsilon \|\varphi\|^2 \Longrightarrow \sup_{h\in \mathrm{A}_N} \sup_{t\in I_\delta} \left| \left\langle U_h^{s+T,\varphi}(t), h \right\rangle \right| \leq \varepsilon^{p_N} \|\varphi\| \right\}$$

satisfies

$$\mathbf{P}\left(\Omega^c_{\varepsilon,N}\right) \leq C_{q_N} \exp\left(\eta \|w_0\|^2\right) \varepsilon^p,$$

for $\varepsilon \in (0, \varepsilon_N]$ *and* $\eta \in (0, \eta_0]$.

Proof. The proof proceeds by induction on N. It suffices to show the result for φ with unit norm $\|\varphi\| = 1$. We first prove that the result is true for A_1. Assume that $\langle \varphi, M_{s,s+T,h}\varphi \rangle \leq \varepsilon$, then by representation (4.24), one has

$$\sup_{1\leq k\leq d} \int_{I_\delta} \left\langle g_k, U_h^{s+T,\varphi}(\tau) \right\rangle^2 d\tau \leq \varepsilon.$$

Setting $R(t) = \int_{s+\frac{T}{2}}^t \left\langle g_k, U_h^{s+T,\varphi}(\tau) \right\rangle d\tau$, Lemma 6.14 in [42] implies that

$$\begin{aligned} \sup_{t\in I_\delta} \left\langle g_k, U_h^{s+T,\varphi}(t) \right\rangle = \|\partial_t R\|_\infty &\leq 4\|R\|_\infty \max\left\{ \frac{1}{|I_\delta|}, \|R\|_\infty^{-\frac{1}{2}} \|\partial_t R\|_1^{\frac{1}{2}} \right\} \\ &\leq C_T \max\left\{ \varepsilon^{\frac{1}{2}}, \varepsilon^{\frac{1}{4}} \|\partial_t R\|_1^{\frac{1}{2}} \right\}, \end{aligned}$$

where $C_T = 4\max\left\{1, \frac{2}{\sqrt{T}}, \left(\frac{T}{2}\right)^{1/4}\right\}$. It follows from Lemma E.1 of [57] that

$$\|U_h^{s+T,\varphi}\|_{1,0} \leq C\left(1 + \|w_{s,\cdot,h}\|^2_{\infty,1} + \|U_h^{s+T,\varphi}\|^2_{\infty,1}\right).$$

From estimate (4.30), Lemma A.1, we have that for any $p' \geq 1$,

$$\mathbf{E}\|\partial_t R\|_1^{p'} \leq \mathbf{E}\|g_k\|^{p'} \|U_h^{s+T,\phi}\|_{1,0}^{p'} \leq C\mathbf{E}\|U_h^{s+T,\phi}\|_{1,0}^{p'} \leq C\exp\left(\eta\|w_0\|^2\right)\delta^{-p'}.$$

Therefore by the Markov inequality and recalling that $\delta = \frac{T}{4}\varepsilon^r$, one has for $\tilde{p} > 0, \alpha > 0$,

$$\begin{aligned} \mathbf{P}\left(\|\partial_t R\|_1 > \alpha\varepsilon^{-\tilde{p}}\right) &= \mathbf{P}\left(\|\partial_t R\|_1^{p'} > \alpha^{p'}\varepsilon^{-\tilde{p}p'}\right) \\ &\leq \alpha^{-p'}\varepsilon^{\tilde{p}p'}\mathbf{E}\|\partial_t R\|_1^{p'} \leq C(\alpha, p')\exp\left(\eta\|w_0\|^2\right)\varepsilon^{\tilde{p}p' - rp'}. \end{aligned}$$

Then on a set $\widetilde{\Omega}_{\varepsilon,1} \subset \Omega$, such that $\mathbf{P}(\widetilde{\Omega}^c_{\varepsilon,1}) \leq C(\alpha, p')\exp\left(\eta\|w_0\|^2\right)\varepsilon^{\tilde{p}p'-rp'}$, we have

$$\|\partial_t R\|_1 \leq \alpha\varepsilon^{-\tilde{p}}.$$

Now choose $\alpha = C_T^{-2}$, $\tilde{p} = \frac{1}{4}$, and $r_1 = \frac{1}{2}\tilde{p}$. Then provided $r < r_1$, for any $p \geq 1$, by choosing $p' = \frac{p}{\tilde{p}-r}$, one has on $\widetilde{\Omega}_{\varepsilon,1}$,

$$\sup_{t \in I_\delta} \left\langle g_k, U_h^{s+T,\varphi}(t) \right\rangle \leq C_T \max\left(\varepsilon^{\frac{1}{2}}, \alpha^{\frac{1}{2}} \varepsilon^{\frac{1}{8}} \right) \leq \varepsilon^{\frac{1}{8}}$$

and $\mathbf{P}(\widetilde{\Omega}_{\varepsilon,1}^c) \leq C \exp\left(\eta \|w_0\|^2\right) \varepsilon^p$ for $\varepsilon \in (0, \varepsilon_1]$, where $\varepsilon_1 = C_T^{-8/3}$. Observe that the event set $\widetilde{\Omega}_{\varepsilon,1}$ does not depend on the choice of $g_k \in A_1$ and it is contained in $\Omega_{\varepsilon,1}$. So

$$\mathbf{P}(\Omega_{\varepsilon,1}^c) \leq \mathbf{P}(\widetilde{\Omega}_{\varepsilon,1}^c) \leq C \exp\left(\eta \|w_0\|^2\right) \varepsilon^p.$$

Hence the proof for the base case is complete with $p_1 = r_1 = 1/8, q_1 = p/r_1$ and $\varepsilon_1 = C_T^{-8/3}$.

The inductive step is accomplished through the following lemma.

Lemma 4.12. *For $N \geq 2$, fix $\mathfrak{g} \in A_{N-1}$, and suppose that $q := p_{N-1}$ has been given. Then for $p \geq 1$, provided $r < r_N$, the event*

$$\widetilde{\Omega}_{N,\varepsilon} := \left\{ \sup_{t \in I_\delta} \left| \langle U_h^{s+T,\varphi}(t), \mathfrak{g} \rangle \right| \leq \varepsilon^q \Longrightarrow \sup_{1 \leq k \leq d} \sup_{t \in I_\delta} \left| \langle U_h^{s+T,\varphi}(t), \widetilde{B}(\mathfrak{g}, g_k) \rangle \right| \leq \varepsilon^{p_N} \right\}$$

satisfies $\mathbf{P}(\widetilde{\Omega}_{\varepsilon,N}^c) \leq C_{q_N} \exp\left(\eta \|w_0\|^2\right) \varepsilon^p$ for $\varepsilon \in (0, \varepsilon_N)$. Here $p_N = q/24$, $r_N = q/12$, $q_N = 12p/q$ and $\varepsilon_N = C_T^{-8/(7q)}$ with $C_T = 4 \max\{4/T, 1\}$.

Proof. Let $R(t) = \partial_t \langle U_h^{s+T,\varphi}(t), \mathfrak{g} \rangle$. Then Lemma 6.14 in [42] (with $\alpha = 1/3$) implies that

$$\|R\|_\infty \leq 4 \max\left\{ \frac{1}{|I_\delta|} \varepsilon^q, \varepsilon^{\frac{q}{4}} \|R\|_{1/3}^{3/4} \right\} \leq C_T \max\left\{ \varepsilon^q, \varepsilon^{\frac{q}{4}} \|R\|_{1/3}^{3/4} \right\}.$$

Next we show that $\|R\|_{1/3}$ has a bounded expectation. As in [42], we consider the process

$$v_{s,t,h} := w_{s,t,h} - \sum_{k=1}^d g_k \left(W_k(t) - W_k(s) \right), \quad t \geq s,$$

since it has more time regularity. Note

$$R(t) = \left\langle -\Delta \mathfrak{g} + \widetilde{B}(v_{s,t,h}, \mathfrak{g}) + \sum_{k=1}^d \widetilde{B}(g_k, \mathfrak{g}) \left(W_k(t) - W_k(s) \right), U_h^{s+T,\varphi}(t) \right\rangle.$$

And recall that we assumed elements of $\{g_k\}_{k=1}^d$ are smooth, so each $\mathfrak{g}$ has bounded H_θ norm for any θ. Observe that

$$\begin{aligned}
&\left\| \left\langle \widetilde{B}(v_{s,t,h}, \mathfrak{g}), U_h^{s+T,\varphi}(t) \right\rangle \right\|_{1/3} \leq C \|U_h^{s+T,\varphi}\|_{1,0} \left(\|v_{s,t,h}\|_{\infty,1} + \|v_{s,t,h}\|_{1/3,1} \right) \\
&\leq C \|U_h^{s+T,\varphi}\|_{1,0} \left(\|w_{s,t,h}\|_{\infty,1} + \sum_{k=1}^d \|g_k\|_1 \|W_k\|_\infty + \|\partial_t v_{s,t,h}\|_{\infty,1} \right) \\
&\leq C \|U_h^{s+T,\varphi}\|_{1,0} \left(1 + \|w_{s,t,h}\|_{\infty,3}^2 + \sum_{k=1}^d \|W_k\|_\infty^2 \right),
\end{aligned} \tag{4.31}$$

for constant C depending on $\|\mathfrak{g}\|_2, \mathcal{B}_1$ and $\|f\|_{\infty,1} = \sup_{t\in\mathbb{R}} \|\Psi(\beta_t h)\|_1$, where we used the fact that

$$\begin{aligned}\|\partial_t v_{s,t,h}\|_{\infty,1} &= \|\Delta w_{s,t,h} - B(\mathcal{K}w_{s,t,h}, w_{s,t,h}) + \Psi(\beta_t h)\|_{\infty,1} \\ &\le \|w_{s,t,h}\|_{\infty,3} + C\|w_{s,t,h}\|_{\infty,3}^2 + \|f\|_{\infty,1} \le C\left(1 + \|w_{s,t,h}\|_{\infty,3}^2\right).\end{aligned}$$

For other terms in the expression of $R(t)$, one has

$$\|\langle -\Delta\mathfrak{g}, U_h^{s+T,\varphi}\rangle\|_{1/3} \le C\|U_h^{s+T,\varphi}\|_{1,0}, \tag{4.32}$$

and

$$\left\|\left\langle \sum_{k=1}^d \widetilde{B}(g_k,\mathfrak{g})\left(W_k(t) - W_k(s)\right), U_h^{s+T,\varphi}(t)\right\rangle\right\|_{1/3} \le C\sum_{k=1}^d \|W_k\|_{1/3}\|U_h^{s+T,\varphi}\|_{1,0}. \tag{4.33}$$

Therefore one obtains

$$\|R\|_{1/3} \le C\|U_h^{s+T,\varphi}\|_{1,0}\left(1 + \|w_{s,t,h}\|_{\infty,3}^2 + \sum_{k=1}^d \|W_k\|_\infty^2 + \sum_{k=1}^d \|W_k\|_{1/3}^2\right).$$

From the proof of Lemma 7.12 in [57], we know that $\mathbf{E}\|W_k\|_\alpha^\gamma < C(T,\gamma)$ for all $\gamma \ge 1$ and $\alpha \in [0, \frac{1}{2})$. This fact, together with the estimates for the solution from Lemma A.1, implies that for any $p' \ge 1$, $\alpha > 0$ and $\tilde{p} > 0$,

$$\mathbf{P}\left(\|R\|_{1/3} > \alpha\varepsilon^{-\tilde{p}}\right) \le C(p',\alpha)\exp\left(\eta\|w_0\|^2\right)\delta^{-p'}\varepsilon^{\tilde{p}p'} \le C(p',\alpha)\exp\left(\eta\|w_0\|^2\right)\varepsilon^{\tilde{p}p'-p'r}.$$

Therefore on a set $\widetilde{\Omega}_{\varepsilon,N,1} \subset \Omega$, such that $\mathbf{P}(\widetilde{\Omega}_{\varepsilon,N,1}^c) \le C\exp\left(\eta\|w_0\|^2\right)\varepsilon^{\tilde{p}p'-p'r}$, we have

$$\|R\|_{1/3} \le \alpha\varepsilon^{-\tilde{p}}.$$

Now choose p_N, r_N and ε_N as stated in the Lemma and let $\tilde{p} = q/6$, $\alpha = C_T^{-4/3}$. Then on the set $\widetilde{\Omega}_{\varepsilon,N,1}$, we have

$$\|R\|_\infty \le C_T \max\left\{\varepsilon^q, \alpha^{\frac{3}{4}}\varepsilon^{\frac{q}{8}}\right\} \le \varepsilon^{\frac{q}{8}}.$$

for all $\varepsilon \in (0, \varepsilon_N]$. Note that α and ε_N are determined when taking the maximum. Observing that for any $p \ge 1$, provided $r < r_N$, we can take $p' = \frac{p}{\tilde{p}-r}$, so that

$$\mathbf{P}(\widetilde{\Omega}_{\varepsilon,N,1}^c) \le C_{q_N}\exp\left(\eta\|w_0\|^2\right)\varepsilon^p.$$

Denote $\mathcal{R}_0 = \left\langle -\Delta\mathfrak{g} + \widetilde{B}(v_{s,t,h},\mathfrak{g}), U_h^{s+T,\varphi}(t)\right\rangle$ and $\mathcal{R}_k = \left\langle \widetilde{B}(g_k,\mathfrak{g}), U_h^{s+T,\varphi}(t)\right\rangle$ for $k = 1, 2, \cdots, d$, which are the coefficients of the Wiener polynomial $R(t)$, where the Wiener process $\overline{W}$ is the shift of the original two sided Wiener process as in (2.4), i.e.,

$$\overline{W}(r) = W(r+s) - W(s) = (\theta_s\omega)(r), \quad r \ge 0,$$

where θ_s is the Wiener shift as defined in (2.14). Then the technical Theorem 7.1 from [42] implies that on a set $\widetilde{\Omega}_{\varepsilon,N,2} \subset \Omega$, one has

$$\|R\|_\infty \leq \varepsilon^{q/8} \Rightarrow \begin{cases} \text{either } \sup_{0\leq k\leq d} \|\mathcal{R}_k\|_\infty \leq \varepsilon^{q/24}, \\ \text{or } \sup_{0\leq k\leq d} \|\mathcal{R}_k\|_1 \geq \varepsilon^{-q/72}, \end{cases}$$

and $\mathbf{P}(\widetilde{\Omega}^c_{\varepsilon,N,2}) \leq C\varepsilon^p$, where C depends only on p and the events $\widetilde{\Omega}_{\varepsilon,N,2}$ depends on the processes R_k only through the value of the highest degree of the Wiener polynomial, which is 1 here. The Markov inequality and the estimates (4.31)-(4.33) imply that there is an event $\widetilde{\Omega}_{\varepsilon,N,3}$, on which $\|R_k\|_1 < \varepsilon^{-q/72}$ for each k, and

$$\mathbf{P}(\widetilde{\Omega}^c_{\varepsilon,N,3}) \leq C\mathbf{P}\left(\|\mathcal{R}_k\|_1^{72p/q} \geq \varepsilon^{-p}\right) \leq C\exp\left(\eta\|w_0\|^2\right)\varepsilon^p.$$

Now observe that $\bigcap_{i=1}^3 \widetilde{\Omega}_{\varepsilon,N,i} \subset \widetilde{\Omega}_{\varepsilon,N}$, hence

$$\mathbf{P}(\widetilde{\Omega}^c_{\varepsilon,N}) \leq \sum_{i=1}^3 \mathbf{P}(\widetilde{\Omega}^c_{\varepsilon,N,i}) \leq C\exp\left(\eta\|w_0\|^2\right)\varepsilon^p.$$

This completes the proof of the induction step. □

The proof of Lemma 4.11 is then complete. □

Proof of Theorem 4.8. Now we give a proof of Theorem 4.8 by combining the above lemmas. Since $A_\infty = H$, by Lemma 8.3 in [42], for any fixed finite dimensional projection Π, there exists $N > 0$ (N depends on the projection Π, so that p_N, r_N, ε_N depends on Π) such that for each $\alpha > 0$, there exists a constant $\Lambda_\alpha > 0$, such that for every $n \geq N$,

$$\inf_{\varphi\in S_\alpha} \frac{|\langle\varphi, \mathcal{Q}_n\varphi\rangle|}{\|\Pi\varphi\|^2} \geq \Lambda_\alpha.$$

On the other hand, it follows from Lemma 4.10 and Lemma 4.11 that there exist constants $p_N, r_N, \varepsilon_N, c > 0$ such that for every $\alpha > 0$, on the set $\Omega_{\varepsilon,N} \bigcap \Omega_{\delta,\Pi}$, the condition

$$\varphi \in S_\alpha \quad \text{and} \quad \langle\varphi, M_{s,s+T,h}\varphi\rangle \leq \varepsilon\|\varphi\|^2$$

implies that

$$U_h^{s+T,\varphi}(T+s-\delta) \in S_{c\alpha}, \quad \|\Pi U_h^{s+T,\varphi}(T+s-\delta)\| \geq \frac{\alpha}{2}\|\varphi\|,$$

$$\text{and} \quad \sup_{h\in \mathrm{A}_N}\sup_{t\in I_\delta}\left|\left\langle U_h^{s+T,\varphi}(t), h\right\rangle\right| \leq \varepsilon^{p_N}\|\varphi\|,$$

and $\mathbf{P}\left(\Omega^c_{\varepsilon,N} \bigcup \Omega^c_{\delta,\Pi}\right) \leq C\exp\left(\eta\|w_0\|^2\right)\varepsilon^p$, for any $\varepsilon \in (0, \varepsilon_N)$ and $p \geq 1$. Then it follows that on the set $\Omega_{\varepsilon,N}\bigcap\Omega_{\delta,\Pi}$, one has

$$\frac{\alpha}{2}\|\varphi\|\leq\|\Pi U_h^{s+T,\varphi}(T+s-\delta)\|\leq C\Lambda_{c\alpha}^{-1/2}\sup_{h\in A_N}\langle U_h^{s+T,\varphi}(T+s-\delta),h\rangle\leq C\Lambda_{c\alpha}^{-1/2}\varepsilon^{p_N}\|\varphi\|.$$

This in turn shows that on $\Omega_{\varepsilon,N}\bigcap\Omega_{\delta,\Pi}$, $\langle\varphi,M_{s,s+T,h}\varphi\rangle\leq\varepsilon\|\varphi\|^2$ and $\varphi\in S_\alpha$ implies that

$$\frac{\alpha}{2}<C\varepsilon^{p_N},$$

which is not true for $\varepsilon\leq\varepsilon_0:=\min\left\{\varepsilon_N,\left(\frac{\alpha}{2C}\right)^{1/p_N}\right\}$.

Hence $\mathbf{P}\left(\langle M_{s,s+T,h}\varphi,\varphi\rangle<\varepsilon\|\varphi\|^2\right)\leq C\varepsilon^p\exp\left(\eta\left\|w_0\right\|^2\right)$ for $\varphi\in S_\alpha$. □

4.3.3 Estimate of The Error $\mathfrak{R}$ and The Control v. Now it remains to check (4.28). This will be accomplished through Proposition 4.16 and Proposition 4.17. We first establish several lemmas. Recall that the error $\mathfrak{R}$ and control v, as well as related quantities have been given when defining (4.29).

The following lemma is a version of the well known Foias–Prodi estimate. It shows that the linearized system of equation (2.5) has only a finite number of unstable directions along the low modes. The proof of the asymptotic regularizing inequality (4.19) relies on an estimate of the spectrum of the Malliavin matrix on such determining modes.

Lemma 4.13. *For any constants $p\geq1,T,\gamma,\eta>0$, there exists an orthogonal projection $\pi_\ell:=\pi_\ell(p,T,\gamma,\eta)$ onto a finite dimensional subspace of H such that*

$$\begin{aligned}\mathbf{E}\left\|(1-\pi_\ell)\,J_{s,s+T,h}\right\|^p&\leq\gamma\exp\left(\eta\left\|w_0\right\|^2\right),\\ \mathbf{E}\left\|J_{s,s+T,h}\,(1-\pi_\ell)\right\|^p&\leq\gamma\exp\left(\eta\left\|w_0\right\|^2\right),\end{aligned}\tag{4.34}$$

for every $w_0\in H$, $s\in\mathbb{R}$ and $h\in\mathbb{T}^n$.

Proof. Let $\{\lambda_n\}$ be the eigenvalues of $-\Delta$ associated with (2.5), and Π_N the projection onto the subspace of H spanned by the first N eigenfunctions. Let $\Pi_N^\perp=1-\Pi_N$. Since $\left\|\Pi_N^\perp J_{s,s+T,h}\xi\right\|\leq\frac{1}{N}\left\|J_{s,s+T,h}\xi\right\|_1$, from bound (A.16) for the linearization flow and (A.2), we obtain

$$\mathbf{E}\left\|\Pi_N^\perp J_{s,s+T,h}\right\|^p\leq\frac{1}{N^p}\mathbf{E}\left\|J_{s,s+T,h}\right\|_1^p\leq\gamma\exp\left(\eta\left\|w_0\right\|^2\right),$$

for any $\gamma>0$, $\eta\in(0,\eta_0]$ by choosing N sufficiently large, where η_0 is from Lemma A.1. It is readily seen that the inequality still holds for $\eta\geq\eta_0$ hence is true for any $\eta>0$.

It follows from Proposition 6.1 in [13] that

$$\|\widetilde{B}(w_{s,r,h},J_{s,r,h}\xi)\|_{-1/4}\leq\|w_{s,r,h}\|\|J_{s,r,h}\xi\|_1+\|w_{s,r,h}\|_1\|J_{s,r,h}\xi\|.\tag{4.35}$$

Denote $\bar{\xi}_N = \Pi_N^{\perp}\xi$. From inequality (4.35), equation (4.20) and using the variation of constant formula, and the analyticity of $e^{t\Delta}$, we have that

$$\begin{aligned}
\|J_{s,s+T,h}\bar{\xi}_N\| &= \left\| e^{\nu t\Delta}\bar{\xi}_N + \int_s^{s+t} e^{\nu(s+t-r)\Delta}\widetilde{B}(w_{s,r,h}, J_{s,r,h}\bar{\xi}_N)dr \right\| \\
&\le \|e^{\nu t\Delta}\bar{\xi}_N\| + C\int_s^{s+t}(s+t-r)^{-1/4}\|\widetilde{B}(w_{s,r,h}, J_{s,r,h}\bar{\xi}_N)\|_{-1/4}dr \\
&\le e^{-\nu t\lambda_{N+1}}\|\xi\| + C\sup_{r\in[s,s+T]}C(r)\int_s^{s+t}(s+t-r)^{-1/4}(r-s)^{-1/2}dr \\
&\le e^{-\nu t\lambda_{N+1}}\|\xi\| + t^{1/4}C\sup_{r\in[s,s+T]}C(r),
\end{aligned} \tag{4.36}$$

where $C(r) = (\|w_{s,r,h}\|\|J_{s,r,h}\xi\|_1 + \|J_{s,r,h}\xi\|\|w_{s,r,h}\|_1)(r-s)^{1/2}$. It then follows from the estimate (A.16), (A.2), (A.6) and (A.8) from Lemma A.1 that there exists a constant $C > 0$ independent of s, such that

$$\sup_{r\in[s,s+T]}C(r) \le C\exp\left(\eta\|w_0\|^2\right). \tag{4.37}$$

Hence for every $p \ge 1, \gamma > 0, \eta > 0$, by first choosing sufficiently small $t > 0$ and then choosing sufficiently large N, we have by (4.36) and (A.6), (A.2) that

$$\mathbf{E}\|J_{s,s+T,h}\Pi_N^{\perp}\|^p \le \left(\mathbf{E}\|J_{s+t,s+T,h}\|^{2p}\mathbf{E}\|J_{s,s+t,h}\Pi_N^{\perp}\|^{2p}\right)^{1/2} \le \gamma\exp\left(\eta\|w_0\|^2\right).$$

The proof is complete by setting $\pi_\ell = \Pi_N$ for a large enough N. □

The following lemma gives a quantitative control of the error between the Malliavin matrix and its regularization.

Lemma 4.14. *Fix $\xi \in H$ and set*

$$\zeta = \beta(\beta + M_0)^{-1}J_0\xi.$$

Then for any constants $p \ge 1$, $\gamma, \eta > 0$ and every finite dimensional orthogonal projector π_ℓ, there exists a small $\beta_0 := \beta_0(p,\gamma,\eta) > 0$ such that for every $\beta \in (0, \beta_0]$,

$$\mathbf{E}\|\pi_\ell\zeta\|^p \le \gamma\exp\left(\eta\|w_0\|^2\right)\|\xi\|^p.$$

Proof. For $\alpha > 0$, define $A_\alpha := \{\omega \in \Omega : \|\pi_\ell\zeta\|(\omega) > \alpha\|\zeta\|(\omega)\}$. Let $\zeta_\alpha(\omega) = \zeta(\omega)\mathbb{I}_{A_\alpha}(\omega)$ and $\bar{\zeta}_\alpha(\omega) = \zeta(\omega) - \zeta_\alpha(\omega) = \zeta(\omega)\mathbb{I}_{A_\alpha^c}(\omega)$, where $\mathbb{I}_A$ is the characteristic function of the set A. Since

$\|\beta(\beta+M_0)^{-1}\| \le 1$, it follows from estimates (A.2) and (A.6) in Lemma A.1 that

$$\mathbf{E}\left\|\pi_\ell \bar{\zeta}_\alpha\right\|^p \le \alpha^p \mathbf{E}\|\zeta\|^p \le \alpha^p \mathbf{E}\left\|J_0\xi\right\|^p \le \frac{\gamma}{2}\exp\left(\eta\left\|w_0\right\|^2\right)\|\xi\|^p \tag{4.38}$$

by choosing α sufficiently small. Fix such an α. We also have

$$\langle\zeta_\alpha, M_0\zeta_\alpha\rangle \le \langle\zeta, M_0\zeta\rangle \le \langle\zeta,(M_0+\beta)\,\zeta\rangle = \left\langle\beta\left(M_0+\beta\right)^{-1}J_0\xi, \beta J_0\xi\right\rangle \le \beta\left\|J_0\xi\right\|^2.$$

By Theorem 4.8, we know that for every $p \ge 1$ and $\alpha > 0$, there exists a constant C and ε_0 such that

$$\mathbf{P}\left(\langle M_0\zeta_\alpha, \zeta_\alpha\rangle < \varepsilon\left\|\zeta_\alpha\right\|^2\right) \le C\varepsilon^p\exp\left(\eta\left\|w_0\right\|^2\right)$$

holds for every $w_0 \in H$ and every $\varepsilon \in (0,\varepsilon_0)$. Therefore

$$\mathbf{P}\left(\frac{\|\zeta_\alpha\|^2}{\|J_0\xi\|^2} > \frac{\beta}{\varepsilon}\right) \le \mathbf{P}\left(\langle M_0\zeta_\alpha,\zeta_\alpha\rangle < \varepsilon\left\|\zeta_\alpha\right\|^2\right) \le C\varepsilon^p\exp\left(\eta\left\|w_0\right\|^2\right).$$

Choosing $\beta = \varepsilon^2$, and noting $\frac{\|\zeta_\alpha\|}{\|J_0\xi\|} \le 1$, we find that

$$\mathbf{E}\left(\frac{\|\zeta_\alpha\|^{2p}}{\|J_0\xi\|^{2p}}\right) \le \mathbf{P}\left(\frac{\|\zeta_\alpha\|^{2p}}{\|J_0\xi\|^{2p}} > \frac{\beta^p}{\varepsilon^p}\right) + \frac{\beta^p}{\varepsilon^p} \le C\varepsilon^p\exp\left(\eta\left\|w_0\right\|^2\right). \tag{4.39}$$

Note that

$$\mathbf{E}\left\|\pi_\ell\zeta_\alpha\right\|^p \le \mathbf{E}\left\|\zeta_\alpha\right\|^p \le \sqrt{\mathbf{E}\left(\left\|\zeta_\alpha\right\|^{2p}\left\|J_0\xi\right\|^{-2p}\right)\mathbf{E}\left\|J_0\xi\right\|^{2p}}.$$

Then combining this with (A.6) from Lemma A.1 and (4.39), it follows that

$$\mathbf{E}\left\|\pi_\ell\zeta_\alpha\right\|^p \le \frac{\gamma}{2}e^{\eta\|w_0\|^2}\|\xi\|^p \tag{4.40}$$

by choosing ε sufficiently small, which in turn gives the desired β_0. The lemma then follows from (4.38) and (4.40) by observing that $\mathbf{E}\left\|\pi_\ell\zeta\right\|^p = \mathbf{E}\left\|\pi_\ell\zeta_\alpha\right\|^p + \mathbf{E}\left\|\pi_\ell\bar{\zeta}_\alpha\right\|^p$. □

Remark. By the Markov property in its generalized form (see for example Theorem 9.18 in [23]), it follows from Lemma 4.14 that for each positive integer n and $\zeta = \beta\left(\beta+M_n\right)^{-1}J_n\xi$, one has

$$\mathbf{E}\left(\left\|\pi_\ell\zeta\right\|^p \,|\, \mathcal{F}_{s+n}\right) \le \gamma e^{\eta\left\|w_{s,s+n,h}\right\|^2}\|\xi\|^p.$$

Lemma 4.15. *For any constants $\gamma, \eta > 0$ and $p \ge 1$, there exists a constant $\beta_0 := \beta_0(p,\gamma,\eta) > 0$ such that whenever $0 < \beta \le \beta_0$, we have*

$$\mathbf{E}\left(\left\|\mathfrak{R}_{s+2(n+1)}\right\|^p \,\middle|\, \mathcal{F}_{s+2n}\right) \le \gamma e^{\eta\left\|w_{s,s+2n,h}\right\|^2}\left\|\mathfrak{R}_{s+2n}\right\|^p,\ \mathbf{P}\text{-a.s.}$$

Proof. The proof is mainly based on Lemma 4.14 and Lemma 4.13. Let $\zeta = \beta\widetilde{M}_{2n}^{-1}J_{2n}\mathfrak{R}_{s+2n}$.

Observe that

$$\mathfrak{R}_{s+2(n+1)} = J_{2n+1}\mathfrak{R}_{s+2n+1} = J_{2n+1}\zeta. \tag{4.41}$$

Also note that $\|\beta\widetilde{M}_{2n}^{-1}\| \leq 1$ and $\mathfrak{R}_{s+2n}$ is $\mathcal{F}_{s+2n}$ measurable. Hence by the estimate (A.6) in Lemma A.1 and the Markov property, one has

$$\mathbf{E}\left(\|\zeta\|^p \,|\, \mathcal{F}_{s+2n}\right) \leq \|\mathfrak{R}_{s+2n}\|^p\, \mathbf{E}\left(\|J_{2n}\|^p \,|\, \mathcal{F}_{s+2n}\right) \leq Ce^{\frac{\eta}{2}\left\|w_{s,s+2n,h}\right\|^2} \|\mathfrak{R}_{s+2n}\|^p. \tag{4.42}$$

Applying Lemma 4.13, Hölder's inequality and estimate (4.42), it follows that there exists a projection π_l on a finite dimensional subspace of H such that

$$\begin{aligned}\mathbf{E}\left(\|J_{2n+1}\left(1-\pi_\ell\right)\zeta\|^p \,|\, \mathcal{F}_{s+2n}\right) &\leq \sqrt{\mathbf{E}\left(\|J_{2n+1}\left(1-\pi_\ell\right)\|^{2p} \,\middle|\, \mathcal{F}_{s+2n}\right)\mathbf{E}\left(\|\zeta\|^{2p} \,\middle|\, \mathcal{F}_{s+2n}\right)} \\ &\leq \tilde{\gamma} e^{\frac{\eta}{2}\left\|w_{s,s+2n,h}\right\|^2}\left(Ce^{\frac{\eta}{2}\left\|w_{s,s+2n,h}\right\|^2}\|\mathfrak{R}_{s+2n}\|^p\right) \qquad (4.43) \\ &\leq \gamma e^{\eta\left\|w_{s,s+2n,h}\right\|^2}\|\mathfrak{R}_{s+2n}\|^p. \qquad (4.44)\end{aligned}$$

From Lemma 4.14 and the Markov property, it follows that for an arbitrarily small $\tilde{\gamma}$, one can choose β sufficiently small such that

$$\mathbf{E}\left(\|\pi_\ell\zeta\|^p \,|\, \mathcal{F}_{s+2n}\right) \leq \tilde{\gamma} e^{\frac{\eta}{2}\left\|w_{s,s+2n,h}\right\|^2}\|\mathfrak{R}_{s+2n}\|^p.$$

Again applying Hölder's inequality and the estimate (A.6) on the Jacobian J_{2n+1}, one can deduce that for any $\gamma > 0$,

$$\mathbf{E}\left(\|J_{2n+1}\pi_\ell\zeta\|^p \,|\, \mathcal{F}_{s+2n}\right) \leq \gamma e^{\eta\|w_{s+2n}\|^2}\|\mathfrak{R}_{s+2n,h}\|^p. \tag{4.45}$$

by choosing β sufficiently small. The proof is then complete by combining (4.41), (4.43) and (4.45). □

The following result gives a desired estimate on the error between the variations on the initial condition and that on the Wiener path.

Proposition 4.16. *There is $p \in (0,1)$ such that for any $\eta > 0$ and $a > 0$ there are constants $C = C(\eta, a, p)$, so that*

$$\left(\mathbf{E}\|\mathfrak{R}_{s+t}\|^2\right)^{1/2} \leq C\exp(p\eta\|w_0\|^2)e^{-at},$$

for all $s \in \mathbb{R}$ and $t \geq 0$.

Proof. The proof is based on Lemma 4.15, Lemma A.1 and an iteration procedure. Let $C_n = \frac{\|\mathfrak{R}_{s+2n+2}\|^{10}}{\|\mathfrak{R}_{s+2n}\|^{10}}$, where we set $C_n = 0$ if $\mathfrak{R}_{s+2n} = 0$. Note that $\|\mathfrak{R}_{s+2N}\|^{10} = \prod_{n=0}^{N-1} C_n$ since $\|\mathfrak{R}_s\| =$

$\|\xi\| = 1$. Also one observes that $\left\|\beta\widetilde{M}_{2n}^{-1}\right\| \leq 1$, and $\|\mathfrak{R}_{s+2n+2}\| \leq \left\|J_{2n+1}\beta\widetilde{M}_{2n}^{-1}J_{2n}\right\| \|\mathfrak{R}_{s+2n}\|$ in view of (4.41). So by the estimate (A.6) on the Jacobian from Lemma A.1, it follows that for every $\eta > 0$, there exists a constant $C := C(\eta, \nu)$ such that

$$C_n \leq \left\|J_{2n+1}\beta\widetilde{M}_{2n}^{-1}J_{2n}\right\|^{10} \leq \|J_{2n+1}\|^{10}\|J_{2n}\|^{10} \leq \exp\left(\eta\int_{s+2n}^{s+2n+2}\|w_{s,r,h}\|_1^2\,dr + C\right), \ \mathbf{P}-\text{a.s.} \tag{4.46}$$

Now define for $\eta, R > 0$,

$$C_{n,R} = \begin{cases} e^{-\eta R} & \text{if } \|w_{s,s+2n,h}\|^2 \geq 2R, \\ e^{\eta R}C_n & \text{otherwise.} \end{cases}$$

Note that both C_n and $C_{n,R}$ are $\mathcal{F}_{s+2n+2}$ measurable. We denote

$$\Omega_R := \left\{\omega \in \Omega : \|w_{s,s+2n,h}\|^2 \geq 2R\right\}$$

and $\overline{\Omega}_R$ its complement. The probabilities of theses events could depend on the fixed initial time s and parameter $h \in \mathbb{T}^n$, but this dependency will be eliminated when we take expectation later.

It follows from Lemma 4.15 that for every $R > \eta^{-1}$, there is $\beta > 0$ making γ sufficiently small such that

$$\begin{aligned}\mathbf{E}\left(C_{n,R}^2 \,\middle|\, \mathcal{F}_{s+2n}\right) &= \mathbf{E}\left(\mathbb{I}_{\Omega_R}e^{-2\eta R} + \mathbb{I}_{\bar{\Omega}_R}e^{2\eta R}C_n^2 \,\middle|\, \mathcal{F}_{s+2n}\right) = \mathbb{I}_{\Omega_R}e^{-2\eta R} + \mathbb{I}_{\bar{\Omega}_R}e^{2\eta R}\mathbf{E}\left(C_n^2 \,\middle|\, \mathcal{F}_{s+2n}\right) \\ &\leq \mathbb{I}_{\Omega_R}e^{-2\eta R} + \mathbb{I}_{\bar{\Omega}_R}\gamma e^{4\eta R} \leq \frac{1}{2}, \ \mathbf{P}-\text{a.s.}\end{aligned} \tag{4.47}$$

It now follows from the definition of C_n and inequality (4.46) that

$$C_n \leq C_{n,R}\exp\left(\eta\int_{s+2n}^{s+2n+2}\|w_{s,r,h}\|_1^2\,dr + \eta\|w_{s,s+2n,h}\|^2 + C - \eta R\right), \ \mathbf{P}-\text{a.s.}$$

Therefore by the Cauchy-Schwarz inequality,

$$\begin{aligned}\prod_{n=0}^{N-1} C_n &\leq \prod_{n=0}^{N-1} C_{n,R}^2 + \prod_{n=0}^{N-1}\exp\left(2\eta\int_{s+2n}^{s+2n+2}\|w_{s,r,h}\|_1^2\,dr + 2\eta\|w_{s,s+2n,h}\|^2 + 2C - 2\eta R\right) \\ &\leq \prod_{n=0}^{N-1} C_{n,R}^2 + \exp\left(4\eta\sum_{n=0}^{N-1}\|w_{s,s+2n,h}\|^2 + 2N(C - \eta R)\right) \\ &\quad + \exp\left(4\eta\int_s^{s+2N}\|w_{s,r,h}\|_1^2\,dr + 2N(C-\eta R)\right).\end{aligned} \tag{4.48}$$

Now from inequality (4.47) one has

$$\mathbf{E}\left(\prod_{n=0}^{N-1} C_{n,R}^2 \,\middle|\, \mathcal{F}_{s+2(N-1)}\right) \leq \frac{1}{2}\prod_{n=0}^{N-2} C_{n,R}^2, \ \mathbf{P}-\text{a.s.}$$

Taking conditional expectation repeatedly, one obtains

$$\mathbf{E}\left(\prod_{n=0}^{N-1} C_{n,R}^2\right) \le \frac{1}{2^N}. \tag{4.49}$$

Fix $\eta > 0$ such that $\eta \le \min\{\frac{1}{4}\eta_0\nu, \frac{1}{4}\eta_1\}$. Then the bounds from inequalities (A.2) and (A.5) imply that

$$\begin{aligned} \mathbf{E}\exp\left(4\eta \int_s^{s+2N} \|w_{s,r,h}\|_1^2\,dr + 2N\,(C-\eta R)\right) &\le C\exp\left(4\eta\nu^{-1}\|w_0\|^2 + 2N(C-\eta R)\right), \\ \mathbf{E}\exp\left(4\eta \sum_{n=0}^{N-1} \|w_{s,s+2n,h}\|^2 + 2N\,(C-\eta R)\right) &\le \exp\left(4a\eta\|w_0\|^2 + N\,(\gamma + 2C - 2\eta R)\right). \end{aligned} \tag{4.50}$$

Choose R sufficiently large such that these two terms satisfy the desired bounds. Then choose β sufficiently small so that the estimate (4.49) holds and hence by (4.48)-(4.50) we have

$$\mathbf{E}\,\|\mathfrak{R}_{s+2N}\|^{10} \le \frac{C\exp\left(\eta\|w_0\|^2\right)}{2^N} \tag{4.51}$$

for every $N \in \mathbb{N}$.

Note that for $t \in [2n, 2n+1)$, one has (we omit the dependence on $h \in \mathbb{T}^n$ for notational simplicity)

$$\begin{aligned} \mathfrak{R}_{s+t} &= J_{s+2n,s+t}J_{s,s+2n}\xi - A_{s,s+t}v_{s,s+t} \\ &= J_{s+2n,s+t}\mathfrak{R}_{s+2n} + J_{s+2n,s+t}A_{s,s+2n}v_{s,s+2n} - A_{s,s+t}v_{s,s+t} \\ &= J_{s+2n,s+t}\mathfrak{R}_{s+2n} + A_{s,s+t}v_{s,s+t} - A_{s+2n,s+t}v_{s+2n,s+t} - A_{s,s+t}v_{s,s+t} \\ &= J_{s+2n,s+t}\mathfrak{R}_{s+2n} - A_{s+2n,s+t}v_{s+2n,s+t}. \end{aligned}$$

Hence by the definition of v as in (4.29), and the fact that $\|A_{2n}^*\widetilde{M}_{2n}^{-1}\| \le \beta^{-1/2}$, we have that

$$\begin{aligned} \|\mathfrak{R}_{s+t}\| &\le \|J_{s+2n,s+t}\mathfrak{R}_{s+2n}\| + \|A_{s+2n,s+t}v_{s+2n,s+t}\| \\ &\le C\beta^{-1/2}\left(1 + \sup_{\tau\in[s+2n,s+t]}\|J_{\tau,s+t}\|^2\right)\|\mathfrak{R}_{s+2n}\|. \end{aligned} \tag{4.52}$$

And for $t \in [2n+1, 2n+2)$, we have

$$\begin{aligned} \mathfrak{R}_{s+t} &= J_{s+2n,s+t}J_{s,s+2n}\xi - A_{s,s+t}v_{s,s+t} \\ &= J_{s+2n+1,s+2n+2}J_{s,s+2n+1}\xi - A_{s,s+t}v_{s,s+t} = J_{s+2n+1,s+t}\mathfrak{R}_{s+2n+1}. \end{aligned}$$

Note that $\|\mathfrak{R}_{s+2n+1}\| = \|\beta\widetilde{M}_{2n}^{-1}J_{2n}\mathfrak{R}_{s+2n}\| \le \|J_{2n}\|\|\mathfrak{R}_{s+2n}\|$. Hence

$$\|\mathfrak{R}_{s+t}\| \le \sup_{\tau\in[s+2n+1,s+t]}\|J_{\tau,s+t}\|\|\mathfrak{R}_{s+2n+1}\| \le \sup_{\tau\in[s+2n,s+t]}\|J_{\tau,s+t}\|^2\|\mathfrak{R}_{s+2n+1}\|. \tag{4.53}$$

Combining the above inequalities (4.52) and (4.53) with estimates (A.2), (A.6) and inequality

(4.51), one has

$$\left(\mathbf{E}\,\|\mathfrak{R}_{s+t}\|^2\right)^{1/2} \le C\exp\left(p\eta\|w_0\|^2\right)e^{-at}$$

for some $p \in (0,1)$ and all $t \ge 0$. □

The following result shows that the cost of the variation v on the Wiener path can be bounded. Since the proof is the same as that in [40] once we obtain the estimate (4.51), we omit it here.

Proposition 4.17. *There is $p \in (0,1)$ such that for any $\eta > 0$, there exists a constant $C = C(f, \mathcal{B}_0, \eta, \nu, p)$ so that for all $t \ge 0$,*

$$\mathbf{E}\left|\int_s^{s+t} v(r)dW(r)\right|^2 \le \frac{C}{\beta^2}e^{p\eta\|w_0\|^2}\sum_{n=0}^{\infty}\left(\mathbf{E}\,\|\mathfrak{R}_{s+2n}\|^{10}\right)^{\frac{1}{5}}.$$

As a byproduct, we have the following asymptotic strong Feller property. Note that the constant C is independent of the initial time s compared with the aymptotic strong Feller property proposed in [16].

Corollary 4.18. *Under the same condition as in Proposition 4.7, with $t_n = 2n$ and $\delta_n = 2^{-n}$, we have some $\eta_0 > 0$, such that for $\eta \in (0, \eta_0]$, there is a constant $C = C(\eta) > 0$ such that*

$$\|\nabla\mathcal{P}_{s,s+t_n,h}\varphi(w)\| \le C\exp(\eta\|w\|)\left(\|\varphi\|_\infty + \delta_n\|\nabla\varphi\|_\infty\right)$$

for all $\varphi \in C_b^1(H)$, $s \in \mathbb{R}$, $n \in \mathbf{N}$ and $w \in H$, $h \in \mathbb{T}^n$.

Proof. The inequality follows by (4.26), estimate (4.51) and Proposition 4.17. □

4.4 Proof of Theorem 3.2

As in [41], to prove Theorem 3.2, we use a metric d on H that is equivalent to ρ but easier to handle with the estimates from previous subsections. Fix any $r_0 > 0$ as in the Lyapunov structure in Proposition 4.1, and ρ_r is the metric defined as in (4.1). For constants $r \in [r_0, 1)$, $\delta > 0$ and $\beta \in (0,1)$, the metric d is defined as

$$d(w_1, w_2) = \left(1 \wedge \frac{\rho_r(w_1, w_2)}{\delta}\right) + \beta\rho(w_1, w_2), \tag{4.54}$$

which is equivalent to ρ since $\beta\rho(w_1, w_2) \le d(w_1, w_2) \le (\delta^{-1} + \beta)\rho(w_1, w_2)$.

We first give a lemma that can reduce the contraction (3.5) to a relatively simpler case. The first part of the lemma allows us to extend the contraction of the transition operator on H (embedded in

$\mathcal{P}(H)$) to a contraction on $\mathcal{P}(H)$. And the second part asserts that one can obtain the contraction for all times once the transition operator is a contraction at a particular time.

Lemma 4.19. *We have*

(i) For $s \in \mathbb{R}, t \geq 0$, and any given distance d on H, if $d(\mathcal{P}^*_{s,s+t,h}\delta_{w_1}, \mathcal{P}^*_{s,s+t,h}\delta_{w_2}) \leq \alpha d(w_1, w_2)$, for any $w_1, w_2 \in H$, then $d(\mathcal{P}^*_{s,s+t,h}\mu_1, \mathcal{P}^*_{s,s+t,h}\mu_2) \leq \alpha d(\mu_1, \mu_2)$ for any $\mu_1, \mu_2 \in \mathcal{P}(H)$.

(ii) If there are $N \in \mathbb{N}$, $\alpha \in (0,1)$, such that for any $w_1, w_2 \in H$, $r \in [r_0, 1]$ and $s \in \mathbb{R}$, $\rho_r(\mathcal{P}^*_{s,s+N,h}\delta_{w_1}, \mathcal{P}^*_{s,s+N,h}\delta_{w_1}) \leq \alpha\rho_r(w_1, w_2)$. Then there are $C > 0, \gamma > 0$ such that

$$\rho_r(\mathcal{P}^*_{s,s+t,h}\mu_1, \mathcal{P}^*_{s,s+t,h}\mu_2) \leq Ce^{-\gamma t}\rho_r(\mu_1, \mu_2),$$

for any $t \geq 0$.

Proof. Fix $s \in \mathbb{R}$, $h \in \mathbb{T}^n$ and $t \geq 0$. By Theorem 4.4.3 from [49] on the existence of optimal couplings for probability kernels, it follows that for the transition probability kernel

$$\mathcal{P}_{s,s+t,h}(\cdot,\cdot) : H \times \mathcal{B}(H) \to \mathbb{R},$$

there is an optimal coupling kernel Q in the sense that

$$Q : (H \times H) \times (\mathcal{B}(H) \otimes \mathcal{B}(H)) \to \mathbb{R},$$

is a probability kernel on $(H \times H, \mathcal{B}(H) \otimes \mathcal{B}(H))$ such that for every $(w_1, w_2) \in H \times H$, $Q((w_1, w_2), \cdot)$ is an optimal coupling of the transition probabilities $\mathcal{P}^*_{s,s+t,h}\delta_{w_1}$ and $\mathcal{P}^*_{s,s+t,h}\delta_{w_2}$:

$$d(\mathcal{P}^*_{s,s+t,h}\delta_{w_1}, \mathcal{P}^*_{s,s+t,h}\delta_{w_2}) = \inf_{\mu\in\mathcal{C}} \int_{H\times H} d(u,v)\mu(dudv) = \int_{H\times H} d(u,v)Q((w_1,w_2), dudv), \quad (4.55)$$

where $\mathcal{C} = \mathcal{C}(\mathcal{P}^*_{s,s+t,h}\delta_{w_1}, \mathcal{P}^*_{s,s+t,h}\delta_{w_2})$ is the set of all couplings of the transition probabilities $\mathcal{P}^*_{s,s+t,h}\delta_{w_1}$ and $\mathcal{P}^*_{s,s+t,h}\delta_{w_2}$.

Define the operator P_Q acting on $B_b(H \times H)$ by

$$P_Q\phi(w_1, w_2) = \int_{H\times H} \phi(u,v)Q((w_1,w_2), dudv),$$

which induces an operator P^*_Q on $\mathcal{P}(H \times H)$ by duality, $P^*_Q\mu(A \times B) = \int_{H\times H} Q((w_1,w_2), A \times B)\mu(dw_1dw_2)$. One can verify that if μ is a coupling of μ_1, μ_2, then $P^*_Q\mu$ is a coupling of $\mathcal{P}^*_{s,s+t,h}\mu_1$

and $\mathcal{P}^*_{s,s+t,h}\mu_2$. Suppose μ_* is an optimal coupling of μ_1, μ_2. Hence by (4.55),

$$
\begin{aligned}
d(\mathcal{P}^*_{s,s+t,h}\mu_1, \mathcal{P}^*_{s,s+t,h}\mu_2) &= \inf_{\mu \in \mathcal{C}(\mathcal{P}^*_{s,s+t,h}\mu_1, \mathcal{P}^*_{s,s+t,h}\mu_2)} \int_{H\times H} d(u,v)\mu(dudv) \le \int_{H\times H} d(u,v) P^*_Q \mu_*(dudv) \\
&= \int_{H\times H}\int_{H\times H} d(u,v) Q((w_1,w_2), dudv)\mu_*(dw_1 dw_2) \\
&= \int_{H\times H} d(\mathcal{P}^*_{s,s+t,h}\delta_{w_1}, \mathcal{P}^*_{s,s+t,h}\delta_{w_2})\mu_*(dw_1 dw_2) \\
&\le \alpha \int_{H\times H} d(w_1,w_2)\mu_*(dw_1 dw_2) = \alpha d(\mu_1,\mu_2).
\end{aligned}
$$

This completes the proof for the first part of the lemma. From Lemma 4.2, one has

$$
\rho_r(\mathcal{P}^*_{s,s+t,h}\delta_{w_1}, \mathcal{P}^*_{s,s+t,h}\delta_{w_2}) \le \mathbf{E}\rho_r(\Phi_{s,s+t,h}(w_1), \Phi_{s,s+t,h}(w_2)) \le C\rho_r(w_1,w_2),
$$

for any $s \in \mathbb{R}$ and $t \in [0,1]$. Observe that

$$
\begin{aligned}
\rho_r(\mathcal{P}^*_{s,s+2,h}\delta_{w_1}, \mathcal{P}^*_{s,s+2,h}\delta_{w_2}) &= \rho_r(\mathcal{P}^*_{s+1,s+2,h}\mathcal{P}^*_{s,s+1,h}\delta_{w_1}, \mathcal{P}^*_{s+1,s+2,h}\mathcal{P}^*_{s,s+1,h}\delta_{w_2}) \\
&\le C\rho_r(\mathcal{P}^*_{s,s+1,h}\delta_{w_1}, \mathcal{P}^*_{s,s+1,h}\delta_{w_2}) \le C^2\rho_r(w_1,w_2).
\end{aligned}
$$

So by iteration we have for any $n \in \mathbb{N}$,

$$
\rho_r(\mathcal{P}^*_{s,s+n,h}\delta_{w_1}, \mathcal{P}^*_{s,s+n,h}\delta_{w_2}) \le C^n\rho_r(w_1,w_2).
$$

Now for any $0 \le t \le N$, we can write $t = k + \beta$ for a unique integer $k \ge 0$ such that $k \le N$ and $\beta \in [0,1)$. Therefore

$$
\rho_r(\mathcal{P}^*_{s,s+t,h}\delta_{w_1}, \mathcal{P}^*_{s,s+t,h}\delta_{w_2}) = \rho_r\left(\mathcal{P}^*_{s+\beta,s+\beta+k,h}\mathcal{P}^*_{s,s+\beta,h}\delta_{w_1}, \mathcal{P}^*_{s+\beta,s+\beta+k,h}\mathcal{P}^*_{s,s+\beta,h}\delta_{w_2}\right) \le C^{N+1}\rho_r(w_1,w_2).
$$

Hence for any $\gamma > 0$, choosing $\widetilde{C} = C^{N+1}e^{\gamma N}$, we have

$$
\rho_r(\mathcal{P}^*_{s,s+t,h}\delta_{w_1}, \mathcal{P}^*_{s,s+t,h}\delta_{w_2}) \le \widetilde{C}e^{-\gamma t}\rho_r(w_1,w_2), \tag{4.56}
$$

while for $t > N$, one has $t = kN + \beta$, where $k \in \mathbb{N}$, and $0 \le \beta < N$. By assumption, $\rho_r(\mathcal{P}^*_{s,s+N,h}\delta_{w_1}, \mathcal{P}^*_{s,s+n,h}\delta_{w_1}) \le \alpha\rho_r(w_1,w_2)$ for all $s \in \mathbb{R}$. So for any $k \in \mathbb{N}$, by iteration,

$$
\rho_r(\mathcal{P}^*_{s,s+kN,h}\delta_{w_1}, \mathcal{P}^*_{s,s+kN,h}\delta_{w_1}) \le \alpha^k\rho_r(w_1,w_2). \tag{4.57}
$$

It then follows from inequality (4.56), (4.57) and the first part of Lemma 4.19 that

$$
\begin{aligned}
\rho_r(\mathcal{P}^*_{s,s+t,h}\delta_{w_1}, \mathcal{P}^*_{s,s+t,h}\delta_{w_2}) &= \rho_r(\mathcal{P}^*_{s+\beta,s+\beta+kN,h}\mathcal{P}^*_{s,s+\beta,h}\delta_{w_1}, \mathcal{P}^*_{s+\beta,s+\beta+kN,h}\mathcal{P}^*_{s,s+\beta,h}\delta_{w_2}) \\
&\le \alpha^k\rho_r(\mathcal{P}^*_{s,s+\beta,h}\delta_{w_1}, \mathcal{P}^*_{s,s+\beta,h}\delta_{w_2}) \le \alpha^{\frac{t-\beta}{N}}C^{N+1}\rho_r(w_1,w_2)] \\
&\le Ce^{-\gamma t}\rho_r(w_1,w_2),
\end{aligned}
$$

for appropriate constants $C, \gamma > 0$. The proof is then complete by invoking again the first part of

Lemma 4.19 . ☐

The irreducibility, Lyapunov structure and gradient inequality give us the contractions on the state space at different scales. This fact is summarized in the following lemmas. Lemma 4.20 deals with those points that are far apart, where the contraction is guaranteed by the Lyapunov structure. The gradient inequality and Lyapunov structure give the contraction at small scales in Lemma 4.21. And the contraction of the intermediate scale in Lemma 4.22 is given by the irreducibility and Lyapunov structure.

The proof for the lemmas is almost the same as that in [41] since we obtain the irreducibility, Lyapunov structure and the gradient inequality that are uniform in the initial time and $h \in \mathbb{T}^n$ in the previous subsections. We still give the proof here for completeness. Recall that the metric d is defined in (4.54), which depends on δ, β and r.

Lemma 4.20. *There is a constant $L > 0$ such that for any $\delta > 0$, $\beta \in (0,1)$ and $r \in [r_0, 1)$, there is $\alpha_1 \in (0,1)$ such that*

$$\left.\begin{aligned} \rho(w_1, w_2) &\geq L \\ \rho_r(w_1, w_2) &\geq \delta \end{aligned}\right\} \implies d\left(\mathcal{P}^*_{s,s+n,h}\delta_{w_1}, \mathcal{P}^*_{s,s+n,h}\delta_{w_2}\right) \leq \alpha_1 d(w_1, w_2),$$

for all $n \in \mathbb{N}$.

Proof. By Lemma 4.2, we know that there are constants $\alpha \in (0,1)$ and $K > 0$ such that for any w_1, w_2 with $\rho(w_1, w_2) \geq L$, we have

$$\mathbf{E}\rho(\Phi_{s,s+n,h}(w_1), \Phi_{s,s+n,h}(w_2)) \leq \alpha^n \rho(w_1, w_2) + K \leq (\alpha + K/L)\rho(w_1, w_2).$$

Choose L large such that $\alpha_0 := \alpha + K/L \in (0,1)$. Then

$$\begin{aligned} d\left(\mathcal{P}^*_{s,s+n,h}\delta_{w_1}, \mathcal{P}^*_{s,s+n,h}\delta_{w_2}\right) &\leq \mathbf{E}d(\Phi_{s,s+n,h}(w_1), \Phi_{s,s+n,h}(w_2)) \\ &\leq 1 + \beta\mathbf{E}\rho(\Phi_{s,s+n,h}(w_1), \Phi_{s,s+n,h}(w_2)) \leq 1 + \alpha_0\beta\rho(w_1, w_2). \end{aligned}$$

Since $\rho_r(w_1, w_2) > \delta$, by definition of the metric d, one has $d(w_1, w_2) = 1 + \beta\rho(w_1, w_2) \geq 1 + \beta L$. Therefore

$$1 - \alpha_0 \leq (1 - \alpha_0)\frac{d(w_1, w_2)}{1 + \beta L} = \frac{1 + \alpha_0\beta L}{1 + \beta L} d(w_1, w_2) - \alpha_0 d(w_1, w_2).$$

As a result

$$d\left(\mathcal{P}^*_{s,s+n,h}\delta_{w_1}, \mathcal{P}^*_{s,s+n,h}\delta_{w_2}\right) \leq 1 + \alpha_0\beta\rho(w_1, w_2) = 1 - \alpha_0 + \alpha_0 d(w_1, w_2) \leq \frac{1 + \alpha_0\beta L}{1 + \beta L} d(w_1, w_2),$$

where $\alpha_1 := \frac{1+\alpha_0\beta L}{1+\beta L} \in (0,1)$. □

Lemma 4.21. *For any $\alpha_2 \in (0,1)$ there exist $n_0 > 0$, and $r \in [r_0, 1)$, $\delta > 0$ such that*

$$\rho_r(w_1, w_2) < \delta \Longrightarrow d\left(\mathcal{P}^*_{s,s+n,h}\delta_{w_1}, \mathcal{P}^*_{s,s+n,h}\delta_{w_2}\right) \le \alpha_2 d(w_1, w_2),$$

for all $n > n_0$ and $\beta \in (0,1)$.

Proof. By the Monge-Kantorovich duality (2.10), one has

$$d(\mu_1, \mu_2) = \sup_{\mathrm{Lip}_d(\phi)\le 1} \left| \int_H \phi(w)\mu_1(dw) - \int_H \phi(w)\mu_2(dw) \right|.$$

Without loss of generality, in the above formula we could assume the test function $\phi \in C_b^1(H)$ and $\phi(0) = 0$. Then $\mathrm{Lip}_d(\phi) \le 1$ implies that $\|\nabla\phi(w)\| \le (\delta^{-1} + \beta)V(w)$. Also by Proposition 4.1 we have for any $\kappa > 1$,

$$|\phi(w)| \le 1 + \beta\|w\|V(w) \le 1 + \beta C V^\kappa(w) \le 1 + \beta C V^\kappa(w).$$

Now combining Proposition 4.1 and Proposition 4.7, one has

$$\begin{aligned}
&\|\nabla \mathcal{P}_{s,s+t,h}\phi(w)\| \\
&\le C(\eta, a)V^p(w)\left(\sqrt{(\mathcal{P}_{s,s+t,h}|\phi|^2)(w)} + e^{-at}\sqrt{(\mathcal{P}_{s,s+t,h}\|\nabla\phi\|^2)(w)}\right) \\
&\le C(\eta, a)V^p(w)\left[\Big(1 + \beta^2 C^2 \mathbf{E}V^{2\kappa}\big(\Phi_{s,s+t,h}(w)\big)\Big)^{\frac{1}{2}} + e^{-at}(\delta^{-1} + \beta)\Big(\mathbf{E}V^2\big(\Phi_{s,s+t,h}(w)\big)\Big)^{\frac{1}{2}}\right] \\
&\le C(\eta, a)V^{\kappa\alpha(t)+p}(w)(1 + e^{-at}\delta^{-1}) = \delta^{-1}V^{\kappa\alpha(t)+p}(w)(\delta C(\eta, a) + C(\eta, a)e^{-at}).
\end{aligned}$$

For any $\alpha_2 \in (0,1)$, choose large $T_0 > 0$ so that $C(\eta, a)e^{-at} < \frac{\alpha_2}{2}$ for all $t \ge T_0$. From the formula for $\alpha(t)$ in Proposition 4.1, we see that there is a large time $T > T_0$ such that for all $t \ge T$, one has $\kappa\alpha(t) + p < 1$. Choosing $r = \max\{r_0, \kappa\alpha(t) + p\} < 1$ and letting δ be small such that $\delta C(\eta, a) < \frac{\alpha_2}{2}$, then we have

$$\|\nabla \mathcal{P}_{s,s+t,h}\phi(w)\| \le \delta^{-1}V^r(w)\alpha_2.$$

Note that for any $w_1, w_2 \in H$ and any $\varepsilon > 0$, there is a differentiable path $\gamma : [0,1] \to H$ with $\gamma(0) = w_1$ and $\gamma(1) = w_2$ such that

$$\rho_r(w_1, w_2) \le \int_0^1 V^r(\gamma(\tau))\|\dot\gamma(\tau)\|d\tau \le \rho_r(w_1, w_2) + \varepsilon.$$

Then

$$|\mathcal{P}_{s,s+t,h}\phi(w_1) - \mathcal{P}_{s,s+t,h}\phi(w_2)| = \left|\int_0^1 \langle \nabla \mathcal{P}_{s,s+t,h}\phi(\gamma(\tau)), \dot\gamma(\tau)\rangle \, d\tau\right|$$

$$\leq \delta^{-1}\alpha_2 \int_0^1 V^r(\gamma(\tau))\|\dot\gamma(\tau)\|d\tau \leq \delta^{-1}\alpha_2\rho_r(w_1,w_2) + \delta^{-1}\alpha_2\varepsilon \leq \alpha_2 d(w_1,w_2) + \delta^{-1}\alpha_2\varepsilon,$$

where in the last step we use the fact that $\rho_r(w_1,w_2) < \delta$ implies $d(x,y) = \delta^{-1}\rho_r(w_1,w_2) + \beta\rho(w_1,w_2)$. Since $\varepsilon > 0$ is arbitrary, we have for any $w_1, w_2 \in H$

$$\sup_{\mathrm{Lip}_d(\phi)\leq 1} |\mathcal{P}_{s,s+t,h}\phi(w_1) - \mathcal{P}_{s,s+t,h}\phi(w_2)| \leq \alpha_2 d(w_1,w_2).$$

Hence by the Monge-Kantorovich duality,

$$d\left(\mathcal{P}^*_{s,s+n,h}\delta_{w_1}, \mathcal{P}^*_{s,s+n,h}\delta_{w_2}\right) \leq \alpha_2 d(w_1,w_2).$$

The proof is complete. □

Lemma 4.22. *For any $L, \delta > 0$, $r \in (0,1]$, there is some $n_1 > 0$ such that for any $n > n_1$, there are $\beta, \alpha_3 \in (0,1)$ such that*

$$\left.\begin{aligned} &\rho(w_1,w_2) < L \\ &\rho_r(w_1,w_2) \geq \delta \end{aligned}\right\} \implies d\left(\mathcal{P}^*_{s,s+n,h}\delta_{w_1}, \mathcal{P}^*_{s,s+n,h}\delta_{w_2}\right) \leq \alpha_3 d(w_1,w_2).$$

Proof. For $L, \delta > 0$, and $r \in [r_0, 1)$, Lemma 3.11 of [41] shows that the set $S = \{(w_1,w_2) : \rho_r(w_1,w_2) \geq \delta, \rho(w_1,w_2) < L\}$ is a bounded set in $H \times H$. So there exists $R = R(L,\delta,r) > 0$ such that $S \subset \{(w_1,w_2) : \|w_1\|, \|w_2\| \leq R\}$. By Proposition 4.3, we know that for every $n > 0$, there is positive constant $a > 0$ so that for any $(w_1,w_2) \in S$, there is a coupling $(X_{s,s+n,h}, Y_{s,s+n,h})$ of the transition probabilities $\mathcal{P}_{s,s+n,h}(w_1,\cdot)$ and $\mathcal{P}_{s,s+n,h}(w_2,\cdot)$, such that $\mathbf{P}\left(\rho_r(X_{s,s+n,h}, Y_{s,s+n,h}) < \frac{\delta}{2}\right) > a > 0$. Note that there is a constant $C > 0$ such that for any $w \in H$

$$\rho(w,0) \leq \int_0^1 V(\tau w)\|w\|d\tau \leq \|w\|V(\|w\|) \leq CV^\kappa(w).$$

Therefore

$$\begin{aligned} \mathbf{E}\rho(X_{s,s+n,h}, Y_{s,s+n,h}) &\leq \mathbf{E}\rho(X_{s,s+n,h}, 0) + \mathbf{E}\rho(0, Y_{s,s+n,h}) \\ &\leq C\left(\mathbf{E}V^\kappa\left(X_{s,s+n,h}\right) + \mathbf{E}V^\kappa\left(Y_{s,s+n,h}\right)\right) \\ &= C\left(\mathbf{E}V^\kappa\left(\Phi_{s,s+n,h}(w_1)\right) + \mathbf{E}V^\kappa\left(\Phi_{s,s+n,h}(w_2)\right)\right) \\ &\leq C(V^{\kappa\alpha(n)}(w_1) + V^{\kappa\alpha(n)}(w_2)) \leq R_n, \end{aligned}$$

where $R_n = CV^{\kappa\alpha(n)}(R)$. For given random variable X and a measurable set A, recall the notation

$\mathbf{E}(X;A) = \mathbf{E}X\mathbb{I}_A$. Then

$$\begin{aligned}
&\mathbf{E}d(X_{s,s+n,h}, Y_{s,s+n,h}) = \mathbf{E}\Big(1 \wedge \frac{\rho_r(X_{s,s+n,h}, Y_{s,s+n,h})}{\delta}\Big) + \beta\mathbf{E}\rho(X_{s,s+n,h}, Y_{s,s+n,h}) \\
&= \mathbf{E}\Big(1 \wedge \frac{\rho_r(X_{s,s+n,h}, Y_{s,s+n,h})}{\delta}; \rho_r(X_{s,s+n,h}, Y_{s,s+n,h}) < \frac{\delta}{2}\Big) \\
&+ \mathbf{E}\Big(1 \wedge \frac{\rho_r(X_{s,s+n,h}, Y_{s,s+n,h})}{\delta}; \rho_r(X_{s,s+n,h}, Y_{s,s+n,h}) \geq \frac{\delta}{2}\Big) + \beta\mathbf{E}\rho(X_{s,s+n,h}, Y_{s,s+n,h}). \\
&\leq \frac{1}{2} + \frac{1}{2}\mathbf{P}\left(\rho_r(X_{s,s+n,h}, Y_{s,s+n,h}) \geq \frac{\delta}{2}\right) + \beta R_n \leq \frac{1}{2} + \frac{1}{2}(1-a) + \beta R_n = 1 - \frac{a}{2} + \beta R_n.
\end{aligned}$$

Letting β be small enough so that $\alpha_3 := 1 - \frac{a}{2} + \beta R_n < 1$, then since $\rho_r(w_1, w_2) \geq \delta$ implies $d(w_1, w_2) \geq 1$, we have

$$d\left(\mathcal{P}^*_{s,s+n,h}\delta_{w_1}, \mathcal{P}^*_{s,s+n,h}\delta_{w_2}\right) \leq \mathbf{E}d(X_{s,s+n,h}, Y_{s,s+n,h}) \leq \alpha_3 d(w_1, w_2),$$

which completes the proof. □

Now we prove Theorem 3.2 with the help of the above lemmas.

Proof of Theorem 3.2. By Lemma 4.19 and the equivalence of the two metrics ρ and d, it suffices to show that

$$d(\mathcal{P}^*_{s,s+n,h}\delta_{w_1}, \mathcal{P}^*_{s,s+n,h}\delta_{w_2}) \leq \alpha d(w_1, w_2),$$

for some $N \in \mathbb{N}$ and $0 < \alpha < 1$ and for every $(w_1, w_2) \in H \times H$. By Lemma 4.21, fixing an $\alpha_2 \in (0,1)$, then there are n_0, r, δ such that for those (w_1, w_2) with $\rho_r(w_1, w_2) < \delta$, one has

$$d\left(\mathcal{P}^*_{s,s+n,h}\delta_{w_1}, \mathcal{P}^*_{s,s+n,h}\delta_{w_2}\right) \leq \alpha_2 d(w_1, w_2)$$

for all $n > n_0$ and $\beta \in (0,1)$. Now fixing L as in Lemma 4.20, then by Lemma 4.22, for the fixed L, δ, r, there is some n_1 such that for $n > n_1$, there exist β, α_3 such that the implication in Lemma 4.22 holds true. Now for fixed δ, β, r, L, there is α_1 such that the implication of Lemma 4.20 holds true. So the conclusion follows by taking $N > \max\{n_0, n_1\}$ and $\alpha = \max\{\alpha_1, \alpha_2, \alpha_3\} < 1$. □

Chapter 5. Unique Ergodicity and Exponential Mixing

In this chapter, we prove Theorem 3.1 by applying a fixed point argument and the uniform contraction (3.5) proved in the previous chapter. The idea is to regard the two parameter family of Markov transition operators $\mathcal{P}^*_{s,t}$ as a non-autonomous dynamical system over the space of probability measures, whose associated pull back mapping is a contraction semigroup acting on an appropriate family of closed subsets of the space of "quasi-periodic graphs" $C(\mathbb{T}^n, \mathcal{P}_1(H))$. And the fixed point of the semigroup gives the desired quasi-periodic invariant measure. We further show that this fixed point has a Hölder regularity if the function Ψ that generates the quasi-periodic force does, and prove the weak convergence of time averages of the transition probabilities to the unique invariant measure for the homogenized process. These results play important roles in the study of the limit theorems in the next chapter. In the last section, we prove the exponential mixing (3.4) in terms of particular observable functions.

5.1 A Fixed Point Argument

In this section we will prove the exponential mixing (5.1) in the following Theorem 5.1 by applying a fixed point argument. Note that Theorem 3.1 follows from it with $\mu_s = \Gamma_{\beta_s 0}$ by taking $h = \beta_s 0$ in (5.1) and (5.7), and applying the translation identity (2.12).

Theorem 5.1. *There is a unique map* $\Gamma \in C(\mathbb{T}^n, \mathcal{P}(H))$, *such that* $\mathcal{P}^*_{0,t,h}\Gamma_h = \Gamma_{\beta_t h}$ *for any* $h \in \mathbb{T}^n$. *Furthermore, there is a constant* $\eta_0 > 0$, *such that for every* $\eta \in (0, \eta_0]$, *there are constants* $C, \varpi > 0$, *such that* $\Gamma \in C(\mathbb{T}^n, \mathcal{P}_1(H))$ *and*

$$\rho(\mathcal{P}^*_{0,t,h}\mu, \Gamma_{\beta_t h}) \leq Ce^{-\varpi t}\rho(\mu, \Gamma_h), \quad t \geq 0, \mu \in \mathcal{P}(H), h \in \mathbb{T}^n, \tag{5.1}$$

where C, ϖ *dose not depend on* h. *Also* $\int_H \exp\left(2\kappa\eta\|w\|^2\right)\Gamma_h(dw) \leq C$ *for all* $h \in \mathbb{T}^n$, *where* $\kappa \geq 2$ *is the constant from Proposition 4.1.*

Moreover, $\Gamma \in C^\zeta(\mathbb{T}^n, (\mathcal{P}_1(H), \rho))$ *if* $\Psi \in C^\gamma(\mathbb{T}^n, H)$, *where* $\zeta = \frac{\varpi\gamma}{r+\varpi}$ *with* $r = 64c_0^6\eta^{-3}\nu^{-5} + \eta C(f, \mathcal{B}_0)$ *from estimate* (A.3).

Proof. Recall that $\mathcal{P}_1(H)$ is defined by (2.9). By Theorem 3.2, for any $t \geq 0$, $\mathcal{P}^*_{0,t,h}$ maps $\mathcal{P}_1(H)$ to itself. Denote for convenience

$$\varphi : \mathbb{R}_+ \times \mathcal{P}_1(H) \times \mathbb{T}^n \to \mathcal{P}_1(H), \text{ by } \varphi(t, \mu, h) = \mathcal{P}^*_{0,t,h}\mu.$$

It follows from the translation identity (2.12) that φ has the cocycle property over the base dynamical system $(\mathbb{T}^n, \mathbb{R}, \beta)$ since for all $\tau, t \geq 0$ and $h \in H(f)$, $\mu \in \mathcal{P}_1(H)$,

$$\mathcal{P}^*_{0,t+\tau,h}\mu = \mathcal{P}^*_{t,t+\tau,h}\mathcal{P}^*_{0,t,h}\mu = \mathcal{P}^*_{0,\tau,\beta_t h}\mathcal{P}^*_{0,t,h}\mu.$$

Hence the pull-back map S^t induced from φ, is defined on the space of quasi-periodic graphs $C(\mathbb{T}^n, \mathcal{P}_1(H))$, i.e.,

$$S^t(\gamma)(h) := \varphi(t, \gamma(\beta_{-t}h), \beta_{-t}h)),\ \gamma \in C(\mathbb{T}^n, \mathcal{P}_1(H)),$$

which satisfies the semigroup property $S^{t_1}S^{t_2}\gamma(h) = S^{t_1+t_2}\gamma(h)$. We would like to apply the fixed point theorem for S^t on $C(\mathbb{T}^n, \mathcal{P}_1(H))$ endowed with the metric (which is complete since $(\mathcal{P}_1(H), \rho)$ is complete)

$$p(\gamma_1, \gamma_2) := \max_{h \in \mathbb{T}^n} \rho(\gamma_1(h), \gamma_2(h)),\ \gamma_1, \gamma_2 \in C(\mathbb{T}^n, \mathcal{P}_1(H)).$$

However, the continuity of $\varphi(t, \mu, h)$ with respect to (μ, h) is unclear due to the Lyapunov structure of the solution of (2.5). Hence $C(\mathbb{T}^n, \mathcal{P}_1(H))$ may not be invariant under the map S^t.

Indeed, from the definition of ρ as in (2.7), one has

$$\rho(w_1, w_2) \leq \|w_1 - w_2\| \left(e^{\eta\|w_1\|^2} + e^{\eta\|w_2\|^2}\right), \quad \forall w_1, w_2 \in H. \tag{5.2}$$

It is known [12, 65] that for any $\mu_1, \mu_2 \in \mathcal{P}(H)$,

$$\rho(\mu_1, \mu_2) = \inf \mathbf{E}\rho(X_1, X_2), \tag{5.3}$$

where the infimum is taken over all couplings (X_1, X_2) for (μ_1, μ_2). Combining (5.2)-(5.3) with estimates (A.1) and (A.3), it follows that

$$\begin{aligned}
&\rho(\mathcal{P}^*_{0,t,h_1}\delta_w, \mathcal{P}^*_{0,t,h_2}\delta_w) \leq \mathbf{E}\rho(w_{0,t,h_1}(w), w_{0,t,h_2}(w)) \\
&\leq \left(\mathbf{E}\|w_{0,t,h_1}(w) - w_{0,t,h_2}(w)\|^2\right)^{\frac{1}{2}} \left(2\mathbf{E}\left[\exp(2\eta\|w_{0,t,h_1}(w)\|^2) + \exp(2\eta\|w_{0,t,h_2}(w)\|^2)\right]\right)^{\frac{1}{2}} \\
&\leq Ce^{\gamma t} g(w) \sup_{t \in \mathbb{R}} \|\Psi(\beta_t h_1) - \Psi(\beta_t h_2)\|,
\end{aligned}$$

where $r = 64c_0^6\eta^{-3}\nu^{-5} + \eta C(f, \mathcal{B}_0)$ is from (A.3), $g(w) = V^{2\kappa}(w) = \exp\left(2\kappa\eta\|w\|^2\right)$, and the Lyapunov function $e^{\eta\|w\|^2}$ along with κ, η are from Proposition 4.1. Therefore by the Markov property,

$$\rho(\mathcal{P}^*_{0,t,h_1}\mu, \mathcal{P}^*_{0,t,h_2}\mu) \leq Ce^{rt} \int_H g(w)\mu(dw) \sup_{t \in \mathbb{R}} \|\Psi(\beta_t h_1) - \Psi(\beta_t h_2)\|, \tag{5.4}$$

It is unclear if each $\mu \in \mathcal{P}_1(H)$ yields $\int_H g(w)\mu(dw)$ finite, therefore we confine ourselves to those measures that make the integral finite to ensure the continuity. To be specific, consider the family of closed subsets of $\mathcal{P}_1(H)$,

$$\mathcal{P}_R := \{\mu \in \mathcal{P}(H) : \int_H g(w)\mu(dw) \leq R\}, \quad R > 0.$$

For each fixed R, $\mathcal{P}_R$ is indeed a closed subset of the complete space $\mathcal{P}_1(H)$ defined as in (2.9). For any $\mu \in \mathcal{P}_R$, one has

$$\rho(\mu, \delta_0) = \int_H \rho(w, 0)\mu(dw) \leq \int_H \|w\| e^{\eta\|w\|^2}\mu(dw) \leq C \int_H g(w)\mu(dw) < \infty,$$

so that $\mu \in \mathcal{P}_1(H)$. Let μ_n be a Cauchy sequence in $\mathcal{P}_R$ under the metric ρ. Then there is a unique $\mu \in \mathcal{P}_1(H)$ such that $\rho(\mu_n, \mu) \to 0$ as $n \to \infty$. So μ_n converges to μ weakly. For $N > 0$, let $g_N(w) = \min\{g(w), N\}$, then $g_N \in C_b(H)$, and

$$\int_H g_N(w)\mu(dw) = \lim_{n\to\infty} \int_H g_N(w)\mu_n(dw) \leq \lim_{n\to\infty} \int_H g(w)\mu_n(dw) \leq R.$$

Hence by the monotone convergence theorem, one has $\int_H g(w)\mu(dw) \leq R$. Therefore $\mu \in \mathcal{P}_R$, which shows that $\mathcal{P}_R$ is closed.

By the contraction property in Theorem 3.2, we have that for any $R > 0$ and $\mu \in \mathcal{P}_R$, φ is continuous in μ, uniformly with respect to h. And by inequality (5.4), it is continuous in h uniformly for μ. Hence φ is jointly continuous in $(\mu, h) \in \mathcal{P}_R \times \mathbb{T}^n$. Then the fixed point argument will be applied on the complete subset $C(\mathbb{T}^n, \mathcal{P}_R)$. However the trade off for the joint continuity is the loss of the invariance of $C(\mathbb{T}^n, \mathcal{P}_R)$ under S^t uniformly for any $t \geq 0$. Indeed, it follows from Proposition 4.1 that for $\mu \in \mathcal{P}_R$, and any $t \geq 0$,

$$\begin{aligned}\int_H \mathcal{P}_{0,t,h} g(w)\mu(dw) &= \int_H \mathbf{E} g(\Phi_{0,t,h}(w))\mu(dw) \\ &\leq C \int_H g^{\alpha(t)}(w)\mu(dw) \leq C \left(\int_H g(w)\mu(dw) \right)^{\alpha(t)} \leq CR^{\alpha(t)},\end{aligned}$$

where we used Jensen's inequality in the penultimate step. One can check that it is impossible to choose a common $R > 0$ such that $CR^{\alpha(t)} \leq R$ for any $t \geq 0$ since $\alpha(t) \to 0$ as $t \to 0$. However, note that for each fixed $t_0 > 0$, if we choose $R = R_{t_0} := C^{\frac{1}{1-\alpha(t_0)}}$ then $CR^{\alpha(t)} \leq CR^{\alpha(t_0)} = R_{t_0}$, which gives the invariance under S^t uniformly for $t \geq t_0$.

Now for any fixed $t_0 \in (0, 1)$, the above analysis shows that the map $S^t : C(\mathbb{T}^n, \mathcal{P}_{R_{t_0}}) \to C(\mathbb{T}^n, \mathcal{P}_{R_{t_0}})$ is well defined for $t \geq t_0$. It remains to show that it is a contraction. Indeed, by Theorem 3.2, one has

$$\begin{aligned}p(S^t\gamma_1, S^t\gamma_2) &= \max_{h\in\mathbb{T}^n} \rho(\varphi(t, \gamma_1(\beta_{-t}h), \beta_{-t}h), \varphi(t, \gamma_2(\beta_{-t}h), \beta_{-t}h)) \\ &= \max_{h\in\mathbb{T}^n} \rho(\mathcal{P}^*_{0,t,\beta_{-t}h}\gamma_1(\beta_{-t}h), \mathcal{P}^*_{0,t,\beta_{-t}h}\gamma_2(\beta_{-t}h)) \\ &\leq Ce^{-\varpi t} \max_{h\in\mathbb{T}^n} \rho(\gamma_1(\beta_{-t}h), \gamma_2(\beta_{-t}h)) = Ce^{-\varpi t} p(\gamma_1, \gamma_2).\end{aligned}$$

Therefore for large $T > t_0$, $p(S^T\gamma_1, S^T\gamma_2) \le cp(\gamma_1, \gamma_2)$ for some $c \in (0,1)$. Fix such a T, then S^T is a contraction over the complete metric space $C(\mathbb{T}^n, \mathcal{P}_{R_{t_0}})$, so there is a unique fixed point $\Gamma_{t_0} \in C(\mathbb{T}^n, \mathcal{P}_{R_{t_0}})$ of S^T. Noting for any $t \ge t_0$, S^t maps $C(\mathbb{T}^n, \mathcal{P}_{R_{t_0}})$ to itself, hence

$$S^T(S^t\Gamma_{t_0}) = S^t(S^T\Gamma_{t_0}) = S^t(\Gamma_{t_0})$$

implies that $S^t(\Gamma_{t_0}) = \Gamma_{t_0}$ by the uniqueness of the fixed point, which shows that Γ_{t_0} is a fixed point of S^t for $t \ge t_0$. For $0 < t_1 \le t_0$, one has $R_{t_1} \ge R_{t_0}$, so $\mathcal{P}_{R_{t_0}} \subset \mathcal{P}_{R_{t_1}}$. And for the same $T > 0$, S^T is a contraction on $C(\mathbb{T}^n, \mathcal{P}_{R_{t_1}})$, which has a unique fixed point Γ_{t_1}. By the uniqueness, $\Gamma_{t_1} = \Gamma_{t_0}$, hence Γ_{t_0} is also a fixed point of S^t for $t \ge t_1$. Since t_1 is arbitrary, we see that $\Gamma := \Gamma_{t_0}$ is a fixed point of S^t for $t \ge 0$, that is, $\varphi(t, \Gamma(\beta_{-t}h), \beta_{-t}h) = \Gamma(h)$ for all $h \in \mathbb{T}^n$. Replacing h with $\beta_t h$ we have $\varphi(t, \Gamma(h), h) = \Gamma(\beta_t h)$ which by definition is

$$\mathcal{P}^*_{0,t,h}\Gamma(h) = \Gamma(\beta_t h).$$

Hence the invariance follows, and the exponential mixing (5.1) then follows from the invariance and Theorem 3.2. Note that by replacing h with $\beta_s h$ in the invariance identity and using the translation identity (2.12), we have

$$\mathcal{P}^*_{s,s+t,h}\Gamma(\beta_s h) = \Gamma(\beta_{s+t}h), \quad \forall s \in \mathbb{R}, t \ge 0, h \in \mathbb{T}^n. \tag{5.5}$$

To show the uniqueness of Γ, suppose that there is another $\widetilde{\Gamma} \in C(\mathbb{T}^n, \mathcal{P}(H))$ that is invariant. Then $\widetilde{\Gamma} \in C(\mathbb{T}^n, \mathcal{P}_1(H))$ by the Lyapunov structure (4.2). Indeed, for $R > 0$, let

$$g_R(w) = \begin{cases} e^{2k\eta\|w\|^2}, & \text{if } \|w\| \le R, \\ e^{2k\eta R^2}, & \text{if } \|w\| \ge R. \end{cases}$$

Then by the invariance of $\widetilde{\Gamma}$ and estimate (4.2), we have for any $M, N > 0$,

$$\begin{aligned} \int_H g_R(w)\widetilde{\Gamma}_h(dw) &= \int_H \mathcal{P}_{-N,0,h}g_R(w)\widetilde{\Gamma}_{\beta_{-N}h}(dw) \\ &\le \int_{\{\|w\|\le M\}} \mathcal{P}_{-N,0,h}g_R(w)\widetilde{\Gamma}_{\beta_{-N}h}(dw) + \int_{\{\|w\|\ge M\}} \mathcal{P}_{-N,0,h}g_R(w)\widetilde{\Gamma}_{\beta_{-N}h}(dw) \\ &\le \int_{\{\|w\|\le M\}} \mathbf{E}g(w_{-N,0,h}(w))\widetilde{\Gamma}_{\beta_{-N}h}(dw) + e^{2k\eta R^2}\widetilde{\Gamma}_{\beta_{-N}h}(\{\|w\| \ge M\}) \\ &\le Ce^{2\kappa\eta\alpha(N)M^2} + e^{2\kappa\eta R^2}\widetilde{\Gamma}_{\beta_{-N}h}(\{\|w\| \ge M\}). \end{aligned}$$

Since $\mathbb{T}^n$ is compact, and $\widetilde{\Gamma} \in C(\mathbb{T}^n, \mathcal{P}(H))$, where $\mathcal{P}(H)$ is endowed with the topology of weak convergence, therefore $\{\widetilde{\Gamma}_h\}_{h\in\mathbb{T}^n}$ is compact and hence tight by Prokhorov's theorem: for any $\varepsilon > 0$, there is a compact subset K_ε of H such that

$$\widetilde{\Gamma}_h(H\backslash K_\varepsilon) < \varepsilon, \quad \forall h \in \mathbb{T}^n.$$

Hence for any $R > 0$ and $\varepsilon = e^{-2\kappa\eta R^2}$, there is a compact subset K_ε of H such that

$$\widetilde{\Gamma}_{\beta_{-N}h}(H\backslash K_\varepsilon) < \varepsilon, \quad \forall N > 0.$$

Now we can choose M large enough such that $K_\varepsilon \subset \{\|w\| \leq M\}$ so that

$$e^{2\kappa\eta R^2}\widetilde{\Gamma}_{\beta_{-N}h}(\{\|w\| \geq M\}) \leq 1.$$

Since $\alpha(N) \to 0$ as $N \to \infty$, we can choose N large such that $e^{2\kappa\eta\alpha(N)M^2} \leq 1$ as well. Therefore we have

$$\int_H g_R(w)\widetilde{\Gamma}_h(dw) \leq C, \quad \forall R > 0,$$

which, by the monotone convergence theorem, in turn implies that

$$\int_H g(w)\widetilde{\Gamma}_h(dw) \leq C, \quad \forall h \in \mathbb{T}^n,$$

and hence $\widetilde{\Gamma} \in C(\mathbb{T}^n, \mathcal{P}_1(H))$. This ensures that

$$\sup_{h\in\mathbb{T}^n} \rho(\Gamma(h), \widetilde{\Gamma}(h)) < \infty.$$

Now by the translation identity (2.12) and Theorem 3.2, we have for $h \in \mathbb{T}^n, t \geq s$,

$$\begin{aligned}\rho(\Gamma(\beta_t h), \widetilde{\Gamma}(\beta_t h)) &= \rho(\mathcal{P}^*_{s,t,h}\Gamma(\beta_s h), \mathcal{P}^*_{s,t,h}\widetilde{\Gamma}(\beta_s h)) \\ &\leq Ce^{-\varpi(t-s)}\rho(\Gamma(\beta_s h), \widetilde{\Gamma}(\beta_s h)) \leq C \sup_{h\in\mathbb{T}^n} \rho(\Gamma(h), \widetilde{\Gamma}(h))e^{-\varpi(t-s)}.\end{aligned}$$

By letting $s \to -\infty$, it follows that $\Gamma(\beta_t h) = \widetilde{\Gamma}(\beta_t h)$ for $t \in \mathbb{R}$. In particular this is true for $t = 0$ and any $h \in \mathbb{T}^n$, hence $\Gamma = \widetilde{\Gamma}$.

To show that $\Gamma \in C^\zeta(\mathbb{T}^n, (\mathcal{P}_1(H), \rho))$ if $\Psi \in C^\gamma(\mathbb{T}^n, H)$, where $\zeta = \frac{\varpi\gamma}{r+\varpi}$ with $r = 64c_0^6\eta^{-3}\nu^{-5} + \eta C(f, \mathcal{B}_0)$ from estimate (A.3), observing that for any $t \geq 0$ and $h_1, h_2 \in \mathbb{T}^n$, by the invariance of

Γ and estimate (A.3), one has

$$\rho\Big(\Gamma(h_1),\Gamma(h_2)\Big) = \rho\Big(\varphi\left(t,\Gamma(\beta_{-t}h_1),\beta_{-t}h_1\right),\varphi\left(t,\Gamma(\beta_{-t}h_2),\beta_{-t}h_2\right)\Big)$$
$$\leq \rho\Big(\varphi\left(t,\Gamma(\beta_{-t}h_1),\beta_{-t}h_1\right),\varphi\left(t,\Gamma(\beta_{-t}h_2),\beta_{-t}h_1\right)\Big) + \rho\Big(\varphi\left(t,\Gamma(\beta_{-t}h_2),\beta_{-t}h_1\right),\varphi\left(t,\Gamma(\beta_{-t}h_2),\beta_{-t}h_2\right)\Big)$$
$$\leq Ce^{rt}\int_H g(w)\Gamma(\beta_{-t}h_1)(dw)\|\Psi\|_\gamma|h_1-h_2|^\gamma + Ce^{-\varpi t}\rho\Big(\Gamma(\beta_{-t}h_1),\Gamma(\beta_{-t}h_2)\Big)$$
$$\leq Ce^{rt}R_{t_0}\|\Psi\|_\gamma|h_1-h_2|^\gamma + Ce^{-\varpi t}\sup_{h_1,h_2\in\mathbb{T}^n}\rho\Big(\Gamma(h_1),\Gamma(h_2)\Big)$$
$$\leq C(e^{rt}|h_1-h_2|^\gamma + e^{-\varpi t}) \leq C|h_1-h_2|^\zeta,$$

with $\zeta = \frac{\varpi\gamma}{r+\varpi}$, by applying the following lemma.

Lemma 5.2. *For $D \geq 1, \Lambda_1, \Lambda_2 > 0, \gamma \in (0,1], 0 < \delta \leq D$, one has*

$$e^{\Lambda_1 T}\delta^\gamma + e^{-\Lambda_2 T} \leq 2D^\gamma\delta^{\overline{\gamma}}$$

for $\overline{\gamma} = \frac{\Lambda_2}{\Lambda_1+\Lambda_2}\gamma$, by choosing $T = -\frac{\gamma}{\Lambda_1+\Lambda_2}\ln\delta$ for $\delta < 1$ and $T = 0$ for $\delta \geq 1$.

The proof is then complete. □

It turns out that Theorem 5.1 also implies the convergence of time averages of the transition probabilities, which is quite useful when applied to the proof of the limit theorems in the next chapter.

Proposition 5.3. *For any $(w_0, h) \in H \times \mathbb{T}^n$ and $K \in \mathbb{N}$, we have the following weak convergence of measures:*

1. $\displaystyle\frac{1}{N}\sum_{j=1}^N \mathcal{P}^*_{0,(j-1)K,h}\delta_{w_0} \to \int_{\mathbb{T}^n}\Gamma_g\lambda(dg).$
2. $\displaystyle\frac{1}{N}\sum_{j=1}^N P^*_{(j-1)K}\delta_{(w_0,h)} \to \Gamma_g(dw)\lambda(dg)$ *and* $\displaystyle\frac{1}{T}\int_0^T P^*_t\delta_{(w_0,h)}dt \to \Gamma_g(dw)\lambda(dg)$ *as well.*

Proof. For any $\phi \in \mathrm{Lip}_\rho(H)$, we have by the Monge-Kantorovich duality (2.10), the invariance and mixing of the quasi-periodic invariant measure from Theorem 5.1 that

$$\left|\left\langle \frac{1}{N}\sum_{j=1}^N\left(\mathcal{P}^*_{0,(j-1)K,h}\delta_{w_0} - \Gamma_{\beta_{(j-1)K}h}\right),\phi\right\rangle\right|$$
$$\leq \mathrm{Lip}_\rho(\phi)\frac{1}{N}\sum_{j=1}^N\rho(P^*_{0,(j-1)K,h}\delta_{w_0}, P^*_{0,(j-1)K,h}\Gamma_h) \leq C\mathrm{Lip}_\rho(\phi)\frac{1}{N}\sum_{j=1}^N e^{-\varpi(j-1)K}\rho(\delta_0,\Gamma_h),$$

which tends to 0 as $N \to \infty$. Also by Birkhoff's ergodic theorem for the irrational rotation on $\mathbb{T}^n$, one has

$$\left\langle \frac{1}{N}\sum_{j=1}^{N} \Gamma_{\beta_{(j-1)K}h}, \phi \right\rangle \to \int_{\mathbb{T}^n} \langle \Gamma_g, \phi\rangle \lambda(dg).$$

Hence the first claim of the proposition follows.

Now let $\phi \in \mathrm{Lip}_{\rho,d}(H \times \mathbb{T}^n)$, where H is equipped with the metric ρ and d is the usual distance in $\mathbb{T}^n$ induced from $\mathbb{R}^n$. Observe that

$$\left\langle \frac{1}{N}\sum_{j=1}^{N} P^*_{(j-1)K}\delta_{(w_0,h)}, \phi \right\rangle = \frac{1}{N}\sum_{j=1}^{N} \mathcal{P}_{0,(j-1)K,h}\phi(\cdot, \beta_{(j-1)K}h)(w_0)$$
$$= \frac{1}{N}\sum_{j=1}^{N} \left\langle \mathcal{P}^*_{0,(j-1)K,h}\delta_{w_0} - \mathcal{P}^*_{0,(j-1)K,h}\Gamma_h, \phi(\cdot, \beta_{(j-1)K}h) \right\rangle + \frac{1}{N}\sum_{j=1}^{N} \left\langle \mathcal{P}^*_{0,(j-1)K,h}\Gamma_h, \phi(\cdot, \beta_{(j-1)K}h) \right\rangle$$
$$:= I + II,$$

where the first term in the sum vanishes by the mixing of the quasi-periodic invariant measure since

$$\begin{aligned} |I| &\le \frac{1}{N}\sum_{j=1}^{N} \mathrm{Lip}_\rho \phi(\cdot, \sigma_{(j-1)K}h)\rho(\mathcal{P}^*_{0,(j-1)K,h}\delta_w, \mathcal{P}^*_{0,(j-1)K,h}\Gamma_h) \\ &\le \frac{C\mathrm{Lip}_{\rho,d}(\phi)}{N}\sum_{j=1}^{N} e^{-\varpi(j-1)K}\rho(\delta_{w_0}, \Gamma_h) \to 0, \end{aligned}$$

while the second term converges to the average of ϕ with respect to $\Gamma_h(dw)\lambda(dg)$ by Birkhoff's ergodic theorem for the irrational rotation β:

$$II = \frac{1}{N}\sum_{j=1}^{N} \left\langle \Gamma_{\sigma_{(j-1)K}h}, \phi(\cdot, \beta_{(j-1)K}h) \right\rangle \to \int_{H\times\mathbb{T}^n} \phi(w,g)\Gamma_g(dw)\lambda(dg),$$

since the observable $g \in \mathbb{T}^n \to \left\langle \Gamma_g, \phi(\cdot, g) \right\rangle$ is continuous. The proof for the continuous time version is similar. □

Remark. Usually the sequence of time averages of the transition probabilities always possesses a subsequence that converges to an invariant measure by the Krylov-Bogoliubov theorem, however it is not guaranteed that the whole sequence always converges to the invariant measure. While in the above proposition, even if it is the case that the homogenized Markov process is not mixing, we have the convergence of the time averages to the unique ergodic invariant measure. From the proof we see that this is a result of the mixing of the inhomogeneous Markov process along the H component, together with the unique ergodicity of the irrational rotation flow.

5.2 Exponential Mixing in Terms of Observable Functions

In this section we prove the following two implications that express the mixing property in terms of the action of the transition operators on observables. Note that the inequality (3.4) in Theorem 3.1 follows from the following Theorem 5.5 with $\mu_s = \Gamma_{\beta_s 0}$ by taking $h = \beta_s 0$ in (5.7), and applying the translation identity (2.12). We fix $h \in \mathbb{T}^n$ throughout the proof.

Corollary 5.4. *There is a constant $\eta_0 > 0$, such that for every $\eta \in (0, \eta_0]$, there exist constants $C, \varpi > 0$ such that*

$$\|\mathcal{P}_{s,t,h}\phi - \Gamma_{\beta_t h}(\phi)\|_{\rho,h,s} \le Ce^{-\varpi(t-s)}\|\phi - \Gamma_{\beta_s h}(\phi)\|_{\rho,h,s},$$

for every Fréchet differentiable $\phi : H \to \mathbb{R}$, $s \le t$ and $h \in \mathbb{T}^n$. Here

$$\|\phi\|_{\rho,h,s} = \sup_{u \ne v} \frac{|\phi(u) - \phi(v)|}{\rho(u,v)} + |\langle \Gamma_{\beta_s h}, \phi\rangle| = \mathrm{Lip}_\rho(\phi) + |\langle \Gamma_{\beta_s h}, \phi\rangle| \,. \tag{5.6}$$

Proof. By the Monge-Kantorovich duality (2.10) and the contraction on $\mathcal{P}(H)$ from Theorem 3.2, we have

$$\begin{aligned}|\mathcal{P}_{s,t,h}\phi(u) - \mathcal{P}_{s,t,h}\phi(v)| &\le \mathrm{Lip}_\rho(\phi) \sup_{\mathrm{Lip}_\rho(\varphi)\le 1} \left| \int_H \varphi(z)\mathcal{P}^*_{s,t,h}\delta_u(dz) - \int_H \varphi(z)\mathcal{P}^*_{s,t,h}\delta_v(dz)\right| \\ &= \mathrm{Lip}_\rho(\phi)\rho(\mathcal{P}^*_{s,t,h}\delta_u, \mathcal{P}^*_{s,t,h}\delta_v) \le \mathrm{Lip}_\rho(\phi)Ce^{-\gamma(t-s)}\rho(u,v),\end{aligned}$$

for any $u, v \in H$. By the invariance of the quasi-periodic invariant measure from Theorem 5.1, $\int_H (\mathcal{P}_{s,t,h}\phi - \Gamma_{\beta_t h}(\phi))\Gamma_{\beta_s h}(\phi)(du) = 0$. Therefore

$$\|\mathcal{P}_{s,t,h}\phi - \Gamma_{\beta_t h}(\phi)\|_{\rho,h,s} = \mathrm{Lip}_\rho(\mathcal{P}_{s,t,h}\phi) \le \mathrm{Lip}_\rho(\phi)Ce^{-\gamma(t-s)} = Ce^{-\gamma(t-s)}\|\phi - \Gamma_{\beta_s h}(\phi)\|_{\rho,h,s}.$$

The proof is complete. □

Theorem 5.5. *There is a constant $\eta_0 > 0$, such that for every $\eta \in (0, \eta_0]$, there exist constants $C, \varpi > 0$, such that for every $\phi \in C^1_\eta$ as in Theorem 3.1,*

$$\|\mathcal{P}_{0,t,h}\phi - \int_H \phi(w)\Gamma_{\beta_t h}(dw)\|_\eta \le Ce^{-\varpi t}\|\phi\|_\eta, \tag{5.7}$$

for any $h \in \mathbb{T}^n, t \ge 0$. In particular, the inequality (3.4) *is obtained by replacing h with $\beta_s 0$ and using the translation identity* (2.12)*, where $\mu_s := \Gamma_{\beta_s 0}$.*

The proof of the theorem will be given at the end of this section by combining Corollary 5.4 and the quasi-equivalence of $\|\cdot\|_{\rho,h,s}$ with $\|\cdot\|_\eta$ that will be proved below.

In the time homogeneous case as in [41], the exponentially mixing in an equivalent form that involves observables similar to (3.4) is given by proving the quasi-equivalence under the Markov semigroup of $\|\cdot\|_\eta$ with an appropriate norm $\|\cdot\|_\rho$ on $\mathrm{Lip}_\rho(H)$. The norm $\|\cdot\|_\rho$ is actually a combination of $\mathrm{Lip}_\rho(\phi)$ for $\phi \in \mathrm{Lip}_\rho(H)$ and the integral of ϕ with respect to the unique invariant measure. In the present time inhomogeneous setting, the quasi-periodic invariant measure depends on time and the parameter $h \in \mathbb{T}^n$. Therefore, to show the mixing property (5.7), we choose the norm $\|\cdot\|_{\rho,h,s}$ on $\mathrm{Lip}_\rho(H)$ defined as above (5.6). This is natural since $\|\cdot\|_{\rho,h,s}$ is quasi-equivalent to $\|\cdot\|_\eta$ under the transition operator $\mathcal{P}_{s,t,h}$ and $\mathcal{P}_{s,t,h}$ has a similar contraction property as in the time homogeneous case proved in Theorem 4.3 of [41]. The dependence on initial time and $h \in \mathbb{T}^n$ of the norm $\|\cdot\|_{\rho,h,s}$ can be regarded as a property that adapts to the time inhomogeneity, to yield a uniform contraction under the action of the transition operator, see Theorem 5.8 below.

To begin with, we first define a family of auxiliary norms for $r \in [0,1]$. The first involves the Lipschitz constant in terms of the metric ρ_r given in (4.1). Define

$$\|\phi\|_{\rho_r,h,s} := \mathrm{Lip}_{\rho_r}(\phi) + |\Gamma_{\beta_s h}(\phi)|, \quad s \in \mathbb{R},$$

where $\mathrm{Lip}_{\rho_r}(\phi) = \sup_{u\neq v} \frac{|\phi(u)-\phi(v)|}{\rho_r(u,v)}$. When $r=1$, it is $\mathrm{Lip}_\rho(\phi)$. The second one is a norm weighted by the Lyapunov function $V(w) = e^{\eta\|w\|^2}$, which was introduced in [41].

$$\|\phi\|_{V^r} := \sup_{w\in H} \frac{|\phi(w)| + \|\nabla\phi(w)\|}{V^r(w)}. \tag{5.8}$$

Note when $r=1$, $\|\cdot\|_{V^r} = \|\cdot\|_\eta$. We first show that $\|\cdot\|_{\rho_r,h,s}$ can be bounded by $\|\cdot\|_{V^r}$ from both sides with different values of r.

Proposition 5.6. *There is a constant $C>0$ such that*

$$C^{-1}\|\phi\|_{V^{\kappa r}} \le \|\phi\|_{\rho_r,h,s} \le C\|\phi\|_{V^r}, \tag{5.9}$$

for $r \in [r_0,1]$, $s\in\mathbb{R}$ and $\phi \in C^1(H)$, where the constants $0<r_0<1$ and $\kappa>1$ are taken from the Lyapunov structure in Proportion 4.1.

Before giving the proof, we need a lemma that connects the norm $\|\cdot\|_{\rho_r,h,s}$ with the derivative part of $\|\cdot\|_{V^r}$.

Lemma 5.7. *For every $\phi \in C^1(H)$, we have*

$$\|\phi\|_{\rho_r,h,s} = \sup_{w\in H}\frac{\|\nabla\phi(w)\|}{V^r(w)} + |\Gamma_{\beta_s h}(\phi)|, \quad s\in\mathbb{R}.$$

Proof. We first claim that for $v \in H$,

$$\lim_{\varepsilon \to 0} \sup_{u:\|u-v\|<\varepsilon} \frac{|\phi(u)-\phi(v)|}{\rho_r(u,v)} = \frac{\|\nabla\phi(v)\|}{V^r(v)}.$$

By definition of ρ_r as in (4.1),

$$\rho_r(u,v) \le \int_0^1 V^r\big((1-\tau)u+\tau v\big)\|u-v\|d\tau,$$

hence we have

$$\sup_{u:\|u-v\|<\varepsilon} \frac{|\phi(u)-\phi(v)|}{\rho_r(u,v)} \ge \sup_{u:\|u-v\|<\varepsilon} \frac{|\phi(u)-\phi(v)|}{\|u-v\|} \left(\sup_{u:\|u-v\|<\varepsilon} \int_0^1 V^r\big((1-\tau)u+\tau v\big)d\tau \right)^{-1}.$$

Therefore by taking the limit,

$$\lim_{\varepsilon \to 0} \sup_{u:\|u-v\|<\varepsilon} \frac{|\phi(u)-\phi(v)|}{\rho_r(u,v)} \ge \frac{\|\nabla\phi(v)\|}{V^r(v)}. \tag{5.10}$$

Next we prove the reverse inequality of (5.10). For fixed $v \in H$ and any u satisfying $\|u-v\| < \varepsilon$, let $R > 0$ large such that $u, v \in B_R(0)$, the ball in $(H, \|\cdot\|)$ with radius R centered at 0. For any $w_1, w_2 \in B_R(0)$, one has the equivalence of metrics

$$\|w_1 - w_2\| \le \rho_r(w_1, w_2) \le V(R)\|w_1 - w_2\|.$$

Let $K = V(R)$. Then for any $\delta > 0$, there exists a differentiable path γ connecting u, v such that

$$\rho_r(u,v) \le \int_0^1 V^r(\gamma(\tau))\|\dot\gamma(\tau)\|d\tau \le \rho_r(u,v) + \delta \le K\varepsilon + \delta.$$

Now for any $t \in [0,1]$,

$$\|\gamma(t) - v\| = \left\| \int_0^t \dot\gamma(\tau)d\tau \right\| \le \int_0^1 V^r(\gamma(\tau))\|\dot\gamma(\tau)\|d\tau \le K\varepsilon + \delta,$$

which means that $\gamma(t)$ never leaves the ball of radius $K\varepsilon + \delta$ centered at v. Therefore,

$$\rho_r(u,v) \ge \int_0^1 V^r(\gamma(\tau))\|\dot\gamma(\tau)\|d\tau - \delta \ge \inf_{w:\|w-v\|\le K\varepsilon+\delta} V^r(w)\|u-v\| - \delta.$$

Hence by taking $\delta = \varepsilon\|u-v\|$ above, we have

$$\frac{|\phi(u)-\phi(v)|}{\rho_r(u,v)} \le \frac{|\phi(u)-\phi(v)|}{\|u-v\|} \left(\inf_{w:\|w-v\|\le (K+\varepsilon)\varepsilon} V^r(w) - \varepsilon \right)^{-1}.$$

By taking limit we have

$$\lim_{\varepsilon \to 0} \sup_{u:\|u-v\|<\varepsilon} \frac{|\phi(u)-\phi(v)|}{\rho_r(u,v)} \le \frac{\|\nabla\phi(v)\|}{V^r(v)},$$

which finishes the proof of the claim.

It then follows from the claim that

$$\mathrm{Lip}_{\rho_r}(\phi) = \sup_{u \ne v} \frac{|\phi(u)-\phi(v)|}{\rho_r(u,v)} \ge \sup_{w \in H} \frac{\|\nabla\phi(w)\|}{V^r(w)}. \tag{5.11}$$

Hence $\|\phi\|_{\rho_r,h,s} \geq \sup_{w\in H} \frac{\|\nabla\phi(w)\|}{V^r(w)} + |\Gamma_{\beta_s h}(\phi)|$. It remains to show the reverse inequality of (5.11). Without loss of generality we can assume $\phi(0) = 0$ and $\mathrm{Lip}_{\rho_r}(\phi) = 1$. There is nothing to show if for some w, $1 \leq \frac{\|\nabla\phi(w)\|}{V^r(w)}$. So we assume $\|\nabla\phi(w)\| \leq V^r(w)$ for all w. Then for any $w_1, w_2 \in H$,

$$|\phi(w_1) - \phi(w_2)| = \int_0^1 \langle \nabla\phi(\gamma(\tau)), \dot\gamma(\tau)\rangle \, d\tau \leq \sup_{w\in H} \frac{\|\nabla\phi(w)\|}{V^r(w)} \int_0^1 V^r(\gamma(\tau))\|\dot\gamma(\tau)\| d\tau.$$

By taking infimum over all differentiable γ connecting w_1, w_2, we have

$$\frac{|\phi(w_1) - \phi(w_2)|}{\rho_r(w_1, w_2)} \leq \sup_{w\in H} \frac{\|\nabla\phi(w)\|}{V^r(w)}.$$

The proof is then complete. □

Proof of Proposition 5.6. By Theorem 5.1, there is a constant $C > 0$, independent of the initial time s such that

$$|\Gamma_{\beta_s h}(\phi)| \leq \int_H V^r(w) \frac{|\phi(w)|}{V^r(w)} \Gamma_{\beta_s h}(dw) \leq \|\phi\|_{V^r} \Gamma_{\beta_s h}(V^r) \leq C\|\phi\|_{V^r}. \tag{5.12}$$

Combining (5.12) with Lemma 5.7 and the definition of the norm $\|\cdot\|_{V^r}$ as (5.8), we have

$$\|\phi\|_{\rho_r,h,s} \leq \sup_{w\in H} \frac{\|\nabla\phi(w)\|}{V^r(w)} + C\|\phi\|_{V^r} \leq \widetilde{C}\|\phi\|_{V^r}.$$

To show the first inequality of (5.9), we fix ϕ with $\|\phi\|_{\rho_r,h,s} = 1$. Then by Proposition 4.1,

$$|\phi(w) - \phi(0)| \leq \mathrm{Lip}_{\rho_r}(\phi)\rho_r(w, 0) \leq \int_0^1 V^r(\tau w)\|w\| d\tau \leq \|w\| V^r(w) \leq C V^{\kappa r}(w).$$

Also by Theorem 5.1 and noting that $r \leq 1$, we have

$$\int_H \rho_r(w, 0)\Gamma_{\beta_s h}(dw) \leq \int_H \rho(w, 0)\Gamma_{\beta_s h}(dw) \leq \int_H C V^\kappa(w)\Gamma_{\beta_s h}(dw) \leq \widetilde{C}.$$

Hence

$$\begin{aligned} |\phi(0)| &\leq \left| \int_H \phi(w)\Gamma_{\beta_s h}(dw) - \phi(0) \right| + \left| \int_H \phi(w)\Gamma_{\beta_s h}(dw) \right| \\ &\leq \int_H |\phi(w) - \phi(0)|\Gamma_{\beta_s h}(dw) + \left| \int_H \phi(w)\Gamma_{\beta_s h}(dw) \right| \leq \widetilde{C} + 1, \end{aligned}$$

where $\left|\int_H \phi(w)\Gamma_{\beta_s h}(dw)\right| \leq 1$ since $\|\phi\|_{\rho_r,h,s} = 1$. It then follows that

$$|\phi(w)| \leq |\phi(0)| + |\phi(w) - \phi(0)| \leq \widetilde{C} V^{\kappa r}(w).$$

Note that $\|\phi\|_{\rho_r,h,s} = 1$ also implies that $\sup_{w\in H} \frac{\|\nabla\phi(w)\|}{V^r(w)} \leq 1$, therefore

$$\|\phi\|_{V^{\kappa r}} \leq \sup_{w\in H} \frac{\|\nabla\phi(w)\|}{V^{\kappa r}(w)} + \sup_{w\in H} \frac{|\phi(w)|}{V^{\kappa r}(w)} \leq 1 + \widetilde{C} \leq C\|\phi\|_{\rho_r,h,s}.$$

The proof is complete. □

The following result shows that the transition operator has a contraction property under the norms $\|\cdot\|_{V^r}$ and $\|\cdot\|_{\rho_r,h,s}$.

Theorem 5.8. *There are constants $C, \gamma > 0$ such that*

$$\|\mathcal{P}_{s,s+t,h}\phi\|_{V^{r(t)}} \leq Ce^{\gamma t}\|\phi\|_{V^r}, \tag{5.13}$$

$$\|\mathcal{P}_{s,s+t,h}\phi\|_{\rho_{r(t)},h,s} \leq Ce^{\gamma t}\|\phi\|_{\rho_{r,h,s+t}}, \tag{5.14}$$

for $r \in [r_0, 2\kappa]$, $s \in \mathbb{R}, t > 0$ and $\phi \in C^1(H)$, where $r(t) = \max\{r\alpha(t), r_0\}$ and $\alpha(t)$ is from Proposition 4.1 .

Proof. We first prove the inequalities for $t \in [0,1]$. By Proposition 4.1, we have

$$\begin{aligned}\|\nabla\mathcal{P}_{s,s+t,h}\phi(w)\| &\leq \mathbf{E}\|\nabla\phi(\Phi_{s,s+t,h}(w))\|\|\nabla\Phi_{s,s+t,h}(w)\| \\ &\leq \sup_{w\in H}\frac{\|\nabla\phi(w)\|}{V^r(w)}\mathbf{E}V^r(\Phi_{s,s+t,h}(w))\|\nabla\Phi_{s,s+t,h}(w)\| \\ &\leq C\sup_{w\in H}\frac{\|\nabla\phi(w)\|}{V^r(w)}V^{r\alpha(t)}(w) \leq C\|\phi\|_{V^r}V^{r\alpha(t)}(w). \end{aligned} \tag{5.15}$$

It follows from the penultimate step that

$$\sup_{w\in H}\frac{\|\nabla\mathcal{P}_{s,s+t,h}\phi(w)\|}{V^{r\alpha(t)}(w)} \leq C\sup_{w\in H}\frac{\|\nabla\phi(w)\|}{V^r(w)}. \tag{5.16}$$

Also

$$|\mathcal{P}_{s,s+t,h}\phi(w)| \leq \mathbf{E}|\phi|(\Phi_{s,s+t,h}(w)) \leq \sup_{w\in H}\frac{|\phi(w)|}{V^r(w)}\mathbf{E}V^r(\Phi_{s,s+t,h}(w)) \leq C\|\phi\|_{V^r}V^{r\alpha(t)}(w). \tag{5.17}$$

Combining the above two estimates (5.15) and (5.17), we see that $\|\mathcal{P}_{s,s+t,h}\phi\|_{V^{r(t)}} \leq C\|\phi\|_{V^r}$ with $r(t) = r\alpha(t)$. Lemma 5.7 and the invariance of the unique periodic invariant measure, together with inequality (5.16) imply that

$$\begin{aligned}\|\mathcal{P}_{s,s+t,h}\phi\|_{\rho_{r(t)},h,s} &= \sup_{w\in H}\frac{\|\nabla\mathcal{P}_{s,s+t,h}\phi(w)\|}{V^{r(t)}(w)} + \left|\int_H \mathcal{P}_{s,s+t,h}\phi(w)\Gamma_{\beta_s h}(dw)\right| \\ &\leq C\sup_{w\in H}\frac{\|\nabla\phi(w)\|}{V^r(w)} + \left|\int_H \phi(w)\Gamma_{\beta_{s+t}h}(dw)\right| \leq C\|\phi\|_{\rho_r,h,s+t}, \quad t\in[0,1].\end{aligned}$$

The case for $t > 1$ follows by iteration. From Proposition 4.1, for $n \in \mathbb{N}$, one has $\alpha(n) = \alpha(1)^n$, so that $r(1) = \max\{r\alpha(1)^n, r_0\}$. By induction, we can show that

$$\|\mathcal{P}_{s,s+n,h}\phi\|_{V^{r(n)}} \leq C^n\|\phi\|_{V^r}. \tag{5.18}$$

Indeed, the base case for $n = 1$ has been proved. In particular, replacing s by $s + k$, it follows that

$$\|\mathcal{P}_{s+k,s+(k+1),h}\phi\|_{V^{r(1)}} \leq C\|\phi\|_{V^r}. \tag{5.19}$$

Assume that for $n = k$, inequality (5.18) is true for all $s \in \mathbb{R}$ and $r \in [r_0, 2\kappa]$. Then since

$r(1) \in [r_0, 2\kappa]$, one has

$$\|\mathcal{P}_{s,s+k,h}\phi\|_{V^{\max\{r(1)\alpha(1)^k, r_0\}}} \le C^k \|\phi\|_{V^{r(1)}}. \tag{5.20}$$

It follows from (5.19), (5.20) and the evolution property of the transition operator that

$$\begin{aligned}
&\|\mathcal{P}_{s,s+(k+1),h}\phi\|_{V^{\max\{r(1)\alpha(1)^k, r_0\}}} = \|\mathcal{P}_{s,s+k.h}\mathcal{P}_{s+k,s+(k+1),h}\phi\|_{V^{\max\{r(1)\alpha(1)^k, r_0\}}} \\
&\le C^k \|\mathcal{P}_{s,s+k,h}\phi\|_{V^{r(1)}} \le C^{k+1}\|\phi\|_{V^r}.
\end{aligned}$$

Since $r(1) = \max\{r\alpha(1), r_0\}$, we find that $r(k+1) = \max\{r\alpha(1)^{k+1}, r_0\} \ge \max\{r(1)\alpha(1)^k, r_0\}$ always holds. Hence

$$\|\mathcal{P}_{s,s+n,h}\phi\|_{V^{r(k+1)}} \le \|\mathcal{P}_{s,s+(k+1),h}\phi\|_{V^{\max\{r(1)\alpha(1)^k, r_0\}}} \le C^{k+1}\|\phi\|_{V^r}.$$

This completes the induction step. Hence (5.18) is true for all $n \in \mathbb{N}$. For any $t \ge 1$, there are unique $k \in \mathbb{N}$ and $\beta \in [0, 1)$ such that $t = k + \beta$. Since $r(\beta) \in [r_0, 2\kappa]$, it follows from (5.18) that

$$\|\mathcal{P}_{s,s+k,h}\phi\|_{V^{\max\{r(\beta)\alpha(1)^k, r_0\}}} \le C^k \|\phi\|_{V^{r(\beta)}}. \tag{5.21}$$

Combining (5.21) with (5.13) for $\beta \in [0, 1)$, and the fact that

$$r(t) = \max\{r\alpha(\beta)\alpha(1)^k, r_0\} \ge \max\{r(\beta)\alpha(1)^k, r_0\},$$

we obtain (5.13) for $t \ge 1$,

$$\begin{aligned}
\|\mathcal{P}_{s,s+t,h}\phi\|_{V^{r(t)}} &\le \|\mathcal{P}_{s,s+k,h}\mathcal{P}_{s+k,s+k+\beta,h}\phi\|_{\max\{r(\beta)\alpha(1)^k, r_0\}} \\
&\le C^k \|\mathcal{P}_{s+k,s+k+\beta,h}\phi\|_{V^{r(\beta)}} \\
&\le C^{k+1}\|\phi\|_{V^r} \le Ce^{\gamma t}\|\phi\|_{V^r},
\end{aligned}$$

by choosing appropriate constants $C, \gamma > 0$ since $k = t - \beta$. The proof for (5.14) in the case $t \ge 1$ is similar.

□

Corollary 5.9. *There exist* $m > 0$, $C = C(m) > 0$ *such that*

$$\|\mathcal{P}_{s,s+m,h}\phi\|_{V^r} \le C\|\phi\|_{\rho_r,h,\tau}$$

for all $\phi \in C^1(H)$, $r \in (1 - \alpha(1), 1]$ *and* $s, \tau \in \mathbb{R}$.

Proof. Let $r_n = r_0 + \alpha(1)^n \kappa r$, where r_0 is from Proposition 4.1 and can be chosen to be arbitrarily close to 0. Since $\alpha(1) < 1$, we can choose a large m such that $\alpha(1)^m \kappa r < 1$. Fix such an m. Then we can choose r_0 small such that $r_m \le r$ and $r_0 < \alpha(1)^m \kappa r$. As a result, we have by Theorem 5.8

and Proposition 5.6, that for $r \in (1-\alpha(1), 1]$,

$$\|\mathcal{P}_{s,s+m,h}\phi\|_{V^r} \le \|\mathcal{P}_{s,s+m,h}\phi\|_{V^{rm}} \le \|\mathcal{P}_{s,s+m,h}\phi\|_{V^{\alpha(1)^m\kappa r}} \le C(m)\|\phi\|_{\kappa r} \le C(m)\|\phi\|_{\rho_r,h,\tau},$$

where in the penultimate step we use inequality (5.13) with

$$r(t) = \max\{\kappa r\alpha(m), r_0\} = \max\{\alpha(1)^m\kappa r, r_0\} = \alpha(1)^m\kappa r.$$

□

We are now in a position to prove Theorem 5.5.

Proof of Theorem 5.5. By Corollary 5.4, and Proposition 5.6, we have

$$\|\mathcal{P}_{s,s+t,h}\phi - \Gamma_{\beta_{s+t}h}(\phi)\|_{\rho,h,s} \le Ce^{-\gamma t}\|\phi - \Gamma_{\beta_s h}(\phi)\|_{\rho,h,s} \le Ce^{-\gamma t}\|\phi - \Gamma_{\beta_s h}(\phi)\|_{\eta},$$

for $s \in \mathbb{R}$ and $t \ge 0$. By Corollary 5.9, there exists $m > 0$ such that

$$\|\mathcal{P}_{s,s+m,h}\phi\|_{\eta} \le C(m)\|\phi\|_{\rho,h,\tau}.$$

Replacing ϕ by $\mathcal{P}_{s+m,s+m+t,h}\phi - \Gamma_{\beta_{s+m+t}h}(\phi)$ and letting $\tau = s+m$ on the right hand side of the above inequality, we have

$$\begin{aligned}\|\mathcal{P}_{s,s+m+t,h}\phi - \Gamma_{\beta_{s+m+t}h}(\phi)\|_{\eta} &\le C(m)\|\mathcal{P}_{s+m,s+m+t,h}\phi - \Gamma_{\beta_{s+m+t}h}(\phi)\|_{\rho,h,s+m} \\ &\le C(m)e^{-\gamma t}\|\phi - \Gamma_{\beta_{s+m}h}(\phi)\|_{\eta}.\end{aligned}$$

Combining the above estimate with (5.12), one has for $t \ge m$,

$$\|\mathcal{P}_{s,s+t,h}\phi - \Gamma_{\beta_{s+t}h}(\phi)\|_{\eta} \le Ce^{-\gamma(t-m)}\left(\|\phi\|_{\eta} + |\Gamma_{\beta_{s+m}h}(\phi)|\right) \le Ce^{-\gamma t}\|\phi\|_{\eta},$$

where C depends on m.

By Theorem 5.8, we have for all $t > 0$,

$$\|\mathcal{P}_{s,s+t,h}\phi\|_{\eta} = \|\mathcal{P}_{s,s+t,h}\phi\|_{V^1} \le \|\mathcal{P}_{s,s+t,h}\phi\|_{V^{\max\{\alpha(t),r_0\}}} \le Ce^{\gamma t}\|\phi\|_{\eta}.$$

So for $0 \le t \le m$, $\|\mathcal{P}_{s,s+t,h}\phi\|_{\eta} \le Ce^{\gamma m}\|\phi\|_{\eta}$. Replacing ϕ by $\phi - \Gamma_{\beta_{s+t}h}(\phi)$, we have

$$\|\mathcal{P}_{s,s+t,h}\phi - \Gamma_{\beta_{s+t}h}(\phi)\|_{\eta} \le Ce^{\gamma m}\|\phi - \Gamma_{\beta_{s+t}h}(\phi)\|_{\eta} \le Ce^{\gamma m}\|\phi\|_{\eta} \le Ce^{-\gamma t}\|\phi\|_{\eta},$$

by choosing the last constant C larger. The proof is complete. □

Chapter 6. Limit Theorems

In this chapter, we establish the limit theorems as well as the corresponding convergence rates as given in Theorem 3.3 and Theorem 3.4 for the time inhomogeneous solution process of the Navier-Stokes equation (2.5). In fact, we will prove these results for a more general class of observable functions that will be given below. The proof is based on a martingale approximation and the limit theorems from the martingale theory.

Due to the time inhomogeneity, it is not obvious to derive a martingale approximation for the inhomogeneous solution process. We also note that the homogenized process X_t is not mixing in the usual sense since $\mathbf{E}\phi(X_t) - \int_{H\times\mathbb{T}^n}\phi(w,h)\Gamma_h(dw)\lambda(dh)$ does not decay to 0, essentially because the irrational rotation on the torus is not mixing. Here $\Gamma_h(dw)\lambda(dh)$ is the unique invariant measure for X_t. Therefore the usual martingale approximation cannot be applied directly. Yet the exponentially mixing quasi-periodic invariant measure enables us to center the observation along the solution process in an appropriate way, which gives us a chance to have a martingale approximation. Indeed, by mixing (5.1), the transition probabilities are exponentially attracted by the quasi-periodic invariant measure, so the expectation of the observation $\phi(w_{0,t}(w_0))$ along the inhomogeneous process is attracted by the quasi-periodic path $\int_H \phi(w)\mu_t(dw)$. Hence $\phi(w_{0,t}(w_0)) - \int_H \phi(w)\mu_t(dw)$ forms a family of "asymptotically centered" random variables that is expected to have the asymptotic behavior described by the limit theorems. Note that if we let $\widetilde{\phi}(w,h) = \phi(w) - \langle\phi,\Gamma_h\rangle$, then

$$\phi(w_{0,t}(w_0)) - \int_H \phi(w)\mu_t(dw) = \widetilde{\phi}(X_t(w_0,0)),$$

which is the observation of the centered observable function along the homogenized process $X_t(w_0,0)$. This indicates that the homogenized process is mixing when acting on the observables centered by the quasi-periodic invariant measure, which enables us to derive a martingale approximation.

Section 6.1 below is devoted to the study of this particular martingale approximation. Then the strong law of large numbers and the central limit theorem are direct conclusions of the corresponding theorems established in martingale theory, which merely requires the weak convergence of the average of the transition probabilities established in Proposition 5.3 and the bounds on the approximating martingale. The proof is presented in Section 6.2. The rate of convergence for the limit theorems will be given in Section 6.3, which requires more detailed analyses on regularity of observables and the convergence rate for the moments of the time average of the observations centered by the quasi-periodic invariant measure.

We now define the space of observable functions and state the main results of this section. For $\gamma \in (0,1]$, let $C^{\gamma}_{\eta,H}(H\times\mathbb{T}^n)$ be the space of Hölder continuous functions with finite norms weighted by the Lyapunov function $e^{\eta\|w\|^2}$,

$$C^{\gamma}_{\eta,H}(H\times\mathbb{T}^n) := \left\{\phi \in C(H\times\mathbb{T}^n) : \|\phi\|_{\gamma,\eta,H} < \infty\right\}, \tag{6.1}$$

where

$$\|\phi\|_{\gamma,\eta,H} := \sup_{(w,h)\in H\times\mathbb{T}^n} \frac{|\phi(w,h)|}{e^{\eta\|w\|^2}} + \sup_{\substack{h\in\mathbb{T}^n\\ 0<\|w_1-w_2\|\le 1}} \frac{|\phi(w_1,h)-\phi(w_2,h)|}{\|w_1-w_2\|^{\gamma}\left(e^{\eta\|w_1\|^2}+e^{\eta\|w_2\|^2}\right)}. \tag{6.2}$$

Let also $C^{\gamma}_{\eta,\mathbb{T}^n}(H\times\mathbb{T}^n)$ be the space of functions that are Hölder continuous on $\mathbb{T}^n-$component and uniformly on bounded subset of H

$$C^{\gamma}_{\eta,\mathbb{T}^n}(H\times\mathbb{T}^n) := \left\{\phi \in C(H\times\mathbb{T}^n) : \|\phi\|_{\gamma,\eta,\mathbb{T}^n} < \infty\right\}, \tag{6.3}$$

where

$$\|\phi\|_{\gamma,\eta,\mathbb{T}^n} := \sup_{(w,h)\in H\times\mathbb{T}^n} \frac{|\phi(w,h)|}{e^{\eta\|w\|^2}} + \sup_{\substack{w\in H\\ 0<\|h_1-h_2\|\le 1}} \frac{|\phi(w,h_1)-\phi(w,h_2)|}{e^{\eta\|w\|^2}|h_1-h_2|^{\gamma}}.$$

Recall that ρ is the geodesic distance on H weighted by $e^{\eta\|w\|^2}$ and therefore depends on η. Let

$$[\phi]_{\gamma,\eta,H_\rho} = \sup_{\substack{h\in\mathbb{T}^n\\ 0<\rho(w_1,w_2)\le 1}} \frac{|\phi(w_1,h)-\phi(w_2,h)|}{\rho(w_1,w_2)^{\gamma}},$$

be the Hölder semi-norm under the metric ρ, and set $C^{\gamma}_{\eta,H_\rho}(H\times\mathbb{T}^n)$ as the space of bounded Hölder continuous (with respect to the metric ρ on H) functions

$$C^{\gamma}_{\eta,H_\rho}(H\times\mathbb{T}^n) = \left\{\phi\in C(H\times\mathbb{T}^n) : \|\phi\|_{\gamma,\eta,H_\rho} := \sup_{(w,h)\in H\times\mathbb{T}^n} |\phi(w,h)| + [\phi]_{\gamma,\eta,H_\rho} < \infty\right\}. \tag{6.4}$$

Remark. It is straightforward to verify that for $0<\delta\le\gamma$, the following inclusion holds:

$$\left\{\phi\in C(H\times\mathbb{T}^n) : [\phi]_{\gamma,\eta,H_\rho} < \infty\right\} \subset C^{\delta}_{2\eta,H}(H\times\mathbb{T}^n).$$

In the case when the deterministic force $f(t,x)$ vanishes and the noise is degenerate as in our work, the weak law of large numbers and central limit theorem were proved in [50] for Lipschitz observable functions ϕ with $[\phi]_{1,\eta,H_\rho} < \infty$. In view of the above inclusion and the fact that the following Theorem 6.1 and Theorem 6.2 are valid for $\phi \in C^{\gamma}_{\eta,H}(H \times \mathbb{T}^n)$, our limit theorems in the case when the deterministic force $f = 0$ can be considered as an improvement of those in [50].

For any $\phi \in C^{\gamma}_{\eta,H}(H \times \mathbb{T}^n)$, we set $\widetilde{\phi}$ as the associated function obtained by normalizing ϕ with the quasi-periodic invariant measure,

$$\widetilde{\phi}(w,h) = \phi(w,h) - \langle \Gamma_h, \phi(\cdot,h)\rangle. \tag{6.5}$$

It is clear that $\widetilde{\phi} \in C^{\gamma}_{\eta,H}(H \times \mathbb{T}^n)$ as well. Note that $X_t(w_0,h_0) = (w_{0,t,h_0}(w_0), \beta_t h_0)$, hence

$$\widetilde{\phi}(X_t(w_0,h_0)) = \phi(w_{0,t,h_0}(w_0), \beta_t h_0) - \langle \Gamma_{\beta_t h_0}, \phi(\cdot, \beta_t h_0)\rangle.$$

Recall from Theorem 5.1 that when $h_0 = 0$, $\Gamma_{\beta_t 0} = \mu_t$ is the unique quasi-periodic invariant measure of the Navier-Stokes system (2.5) with the deterministic force $f(t,x)$. And when ϕ is an observable function on H, we have $\widetilde{\phi}(X_t(w_0,0)) = \phi(w_{0,t,h_0}(w_0)) - \langle \mu_t, \phi\rangle$, which is the observation along the solution process normalized by the quasi-periodic invariant measure. In particular, Theorem 3.3 and Theorem 3.4 with $s = 0$ are obtained from the following Theorem 6.1-6.4 by taking the observable function $\phi \in C^{\gamma}_{\eta}(H) \subset C^{\gamma}_{\eta,H}(H \times \mathbb{T}^n)$ and $h_0 = 0$. For simplicity, we prove Theorem 6.1-6.4 for initial time $s = 0$, while the proof applies to $s \neq 0$ without any change. Therefore Theorem 3.3 and Theorem 3.4 hold for any initial time $s \in \mathbb{R}$.

Let η_0 be the constant from (A.1). The first result is the strong law of large numbers with its rate of convergence.

Theorem 6.1 (SLLN)**.** *For any $\eta \in (0, 2^{-4}\eta_0]$, $\phi \in C^{\gamma}_{\eta,H}(H \times \mathbb{T}^n)$, $(w_0,h_0) \in H \times \mathbb{T}^n$, and $\varepsilon > 0$,*

$$\lim_{T\to\infty} T^{-\frac{1}{2}-\varepsilon} \int_0^T \widetilde{\phi}(X_t(w_0,h_0))dt = 0, \quad \mathbf{P} - \text{a.s.} \tag{6.6}$$

The second result is the central limit theorem.

Theorem 6.2 (CLT)**.** *For any $\eta \in (0, 2^{-6}\eta_0]$, $\phi \in C^{\gamma}_{\eta,H}(H \times \mathbb{T}^n)$, and $(w_0,h_0) \in H \times \mathbb{T}^n$, one has*

$$\lim_{T\to\infty} \frac{1}{\sqrt{T}} \int_0^T \widetilde{\phi}(X_t(w_0,h_0))dt \overset{\mathcal{D}}{=} N(0,\sigma^2), \tag{6.7}$$

where $N(0,\sigma^2)$ is the standard normal random variable with variance σ^2 and $\mathcal{D}$ represents the convergence in distribution. In fact

$$\sigma^2 = \sigma^2_{\phi} = \lim_{T\to\infty} \frac{1}{T}\mathbf{E}\left(\int_0^T \widetilde{\phi}\left(X_t(w_0,h_0)\right)dt\right)^2.$$

The third result of this section is an estimate of the rate of convergence for the strong law of large numbers.

Theorem 6.3. *Let $\varepsilon > 0$, for every integer $p \geq 3$ satisfying $2^p > 1/\varepsilon$, every $\eta \in (0, 2^{-p-1}\eta_0]$, and every $\phi \in C^\gamma_{\eta,H}(H \times \mathbb{T}^n)$, $(w_0, h_0) \in H \times \mathbb{T}^n$, there is an almost surely finite random time $T_0(\omega) \geq 1$, depending on $p, \varepsilon, \|\phi\|_{\gamma,\eta,H}, \|w_0\|, h_0$ such that for all $T > T_0$, we have*

$$\left| \frac{1}{T} \int_0^T \Big(\phi(X_t(w_0, h_0)) - \big\langle \Gamma_{\beta_t h}, \phi(\cdot, \beta_t h) \big\rangle \Big) dt \right| \leq C T^{-\frac{1}{2}+\varepsilon},$$

where $C > 0$ is a constant that does not depend on the above parameters. Moreover, for every $0 < \ell < \min\{2^p\varepsilon - 1, 2^{p-2} - 1\}$, there is a constant $C_p = C_p(\|\phi\|_{\gamma,\eta,H}, \ell, \varepsilon)$ such that

$$\mathbf{E} T_0^\ell \leq C_p e^{2^{p+1}\eta\|w_0\|^2}.$$

The last result is on the convergence rate of the central limit theorem.

Theorem 6.4. *Assume $\Psi \in C^\gamma(\mathbb{T}^n, H)$ and the frequency α satisfies the Diophantine condition* (2.2) *with constant A and dimension n. Let $\Lambda = \frac{\varpi}{5(2-\gamma)}$, $\zeta = \frac{\varpi\gamma}{r+\varpi}$ and $\overline{\gamma}_0 = \frac{\Lambda\zeta}{5(\Lambda+r)(2-\gamma)}$, where ϖ is the mixing rate from Theorem 3.1 and $r = 64c_0^6\eta^{-3}\nu^{-5} + \eta C(f, \mathcal{B}_0)$ is the constant from* (A.3).

1. For any integer $p \geq 2$, $\eta \in (0, 2^{-p-1}\eta_0]$, and $\phi \in C^\gamma_{\eta,H}(H \times \mathbb{T}^n)$ with $\sigma_\phi^2 > 0$, and $(w_0, h_0) \in H \times \mathbb{T}^n$, there are constants $C_p = C_p(\|\phi\|_{\gamma,\eta,H}, \|\phi\|_{\gamma,\eta,\mathbb{T}^n}, \|w_0\|) > 0$ and $T_0 > 0$ such that for all $T \geq T_0$,

$$\sup_{z\in\mathbb{R}} \left| \mathbf{P}\left\{ \frac{1}{\sqrt{T}} \int_0^T \widetilde{\phi}(X_t(w_0, h_0)) dt \leq z \right\} - \Phi_{\sigma_\phi}(z) \right| \leq C_p \left(T^{-\frac{1}{4}} + T^{-\frac{2^{p-2}}{2^p+1}} + T^{-\frac{2^{p-1}\overline{\gamma}_0}{(2^p+1)(A+n)}} \right),$$

2. For $\eta \in (0, 2^{-7}\eta_0]$ and $\phi \in C^\gamma_{\eta,H}(H \times \mathbb{T}^n)$ such that $\sigma_\phi^2 = 0$, and $(w_0, h_0) \in H \times \mathbb{T}^n$, there is a constant $C = C(\|\phi\|_{\gamma,\eta,H}, \|\phi\|_{\gamma,\eta,\mathbb{T}^n}, \|w_0\|) > 0$ such that for all $T \geq 1$,

$$\sup_{z\in\mathbb{R}} (|z| \wedge 1) \left| \mathbf{P}\left\{ \frac{1}{\sqrt{T}} \int_0^T \widetilde{\phi}(X_t(w_0, h_0)) dt \leq z \right\} - \Phi_0(z) \right| \leq C \left(T^{-\frac{1}{4}} + T^{-\frac{\overline{\gamma}_0}{2(A+n)}} \right).$$

6.1 The Martingale Approximation

We first give several properties of the spaces of observable functions defined above in the following Proposition 6.5. Then we prove a mixing result (as a consequence of Theorem 5.1) in terms of the observable functions in Theorem 6.6, which is crucial in deriving the martingale approximation. Proposition 6.7 gives the definition of the corrector and its properties that will be used to construct the martingale approximation as given in (6.14).

Proposition 6.5. *Let η_0 be the constant from estimate* (A.1). *For $\eta \in (0, \eta_0/2]$, and any $0 < \gamma \leq 1$, P_t maps $C^\gamma_{\eta,H}(H \times \mathbb{T}^n)$ into $C^\gamma_{2\eta,H}(H \times \mathbb{T}^n)$; If we further assume that $\Psi \in C^\gamma(\mathbb{T}^n, H)$, then P_t*

maps $C^\gamma_{\eta,H}(H\times\mathbb{T}^n)\cap C^\gamma_{\eta,\mathbb{T}^n}(H\times\mathbb{T}^n)$ *into* $C^\gamma_{2\eta,H}(H\times\mathbb{T}^n)\cap C^\gamma_{2\eta,\mathbb{T}^n}(H\times\mathbb{T}^n)$.

Proof. Let $\phi\in C^\gamma_{\eta,H}(H\times\mathbb{T}^n)$. It follows from (A.1) and (A.15) that for $\eta\in(0,\eta_0/2]$,

$$\begin{aligned}
&|P_t\phi(w_1,h)-P_t\phi(w_2,h)|\\
&=|\mathbf{E}\phi(w_{0,t,h}(w_1),\beta_th)-\mathbf{E}\phi(w_{0,t,h}(w_2),\beta_th)|\le\mathbf{E}\,|\phi(w_{0,t,h}(w_1),\beta_th)-\phi(w_{0,t,h}(w_2),\beta_th)|\\
&\le\|\phi\|_{\gamma,\eta,H}\mathbf{E}\|w_{0,t,h}(w_1)-w_{0,t,h}(w_2)\|^\gamma\left(e^{\eta\|w_{0,t,h}(w_1)\|^2}+e^{\eta\|w_{0,t,h}(w_2)\|^2}\right)\\
&\le C\|\phi\|_{\gamma,\eta,H}\left(\mathbf{E}\|w_{0,t,h}(w_1)-w_{0,t,h}(w_2)\|^2\right)^{\frac{\gamma}{2}}\left(\mathbf{E}\left(e^{2\eta\|w_{0,t,h}(w_1)\|^2}+e^{2\eta\|w_{0,t,h}(w_2)\|^2}\right)\right)^{\frac12}\\
&\le C\|\phi\|_{\gamma,\eta,H}\|w_1-w_2\|^\gamma e^{\frac{\gamma r}{2}t}e^{\frac{\gamma}{2}\eta\|w_1\|^2}\left(e^{\eta\|w_1\|^2}+e^{\eta\|w_2\|^2}\right),
\end{aligned}$$

where $r=64c_0^6\eta^{-3}\nu^{-5}+\eta C\left(f,\mathcal{B}_0\right)$ is the constant from (A.4). Hence we have

$$|P_t\phi(w_1,h)-P_t\phi(w_2,h)|\le C\|\phi\|_{\gamma,\eta,H}\|w_1-w_2\|^\gamma e^{\frac{\gamma r}{2}t}\left(e^{2\eta\|w_1\|^2}+e^{2\eta\|w_2\|^2}\right),\tag{6.8}$$

which shows that P_t maps $C^\gamma_{\eta,H}(H\times\mathbb{T}^n)$ into $C^\gamma_{2\eta,H}(H\times\mathbb{T}^n)$.

Now assume $\Psi\in C^\gamma(\mathbb{T}^n,H)$ and let $\phi\in C^\gamma_{\eta,H}(H\times\mathbb{T}^n)\cap C^\gamma_{\eta,\mathbb{T}^n}(H\times\mathbb{T}^n)$. Then

$$\begin{aligned}
&|P_t\phi(w,h_1)-P_t\phi(w,h_2)|=|\mathbf{E}\phi(w_{0,t,h_1}(w),\beta_th_1)-\mathbf{E}\phi(w_{0,t,h_2}(w),\beta_th_2)|\\
&\le\mathbf{E}\,|\phi(w_{0,t,h_1}(w),\beta_th_1)-\phi(w_{0,t,h_1}(w),\beta_th_2)|+\mathbf{E}\,|\phi(w_{0,t,h_1}(w),\beta_th_2)-\phi(w_{0,t,h_2}(w),\beta_th_2)|\\
&\le\|\phi\|_{\gamma,\eta,\mathbb{T}^n}|h_1-h_2|^\gamma\mathbf{E}e^{\eta\|w_{0,t,h_1}(w)\|^2}+\|\phi\|_{\gamma,\eta,H}\mathbf{E}\|w_{0,t,h_1}(w)-w_{0,t,h_2}(w)\|^\gamma\left(e^{\eta\|w_{0,t,h_1}(w)\|^2}+e^{\eta\|w_{0,t,h_2}(w)\|^2}\right).
\end{aligned}$$

It follows from (A.1) and (A.3) that

$$\begin{aligned}
&\mathbf{E}\|w_{0,t,h_1}(w)-w_{0,t,h_2}(w)\|^\gamma\left(e^{\eta\|w_{0,t,h_1}(w)\|^2}+e^{\eta\|w_{0,t,h_2}(w)\|^2}\right)\\
&\le\left(\mathbf{E}\|w_{0,t,h_1}(w)-w_{0,t,h_2}(w)\|^2\right)^{\frac{\gamma}{2}}\left(\mathbf{E}\left(e^{2\eta\|w_{0,t,h_1}(w)\|^2}+e^{2\eta\|w_{0,t,h_2}(w)\|^2}\right)\right)^{\frac12}\\
&\le Ce^{\frac{r\gamma}{2}t}e^{2\eta\|w\|^2}\|\Psi\|^\gamma_\gamma|h_1-h_2|^\gamma.
\end{aligned}$$

Hence

$$|P_t\phi(w,h_1)-P_t\phi(w,h_2)|\le C(\|\phi\|_{\gamma,\eta,\mathbb{T}^n}+\|\phi\|_{\gamma,\eta,H}\|\Psi\|^\gamma_\gamma)e^{\frac{r\gamma}{2}t}e^{2\eta\|w\|^2}|h_1-h_2|^\gamma,\tag{6.9}$$

which shows that P_t maps $C^\gamma_{\eta,H}(H\times\mathbb{T}^n)\cap C^\gamma_{\eta,\mathbb{T}^n}(H\times\mathbb{T}^n)$ into $C^\gamma_{2\eta,H}(H\times\mathbb{T}^n)\cap C^\gamma_{2\eta,\mathbb{T}^n}(H\times\mathbb{T}^n)$. □

The following theorem shows that the homogenized process is mixing over the family of observables normalized by the quasi-periodic invariant measure as in (6.5), although it is not mixing in the usual sense (by centering the observables with the unique ergodic invariant measure of the homogenized process).

Theorem 6.6. *For any $\gamma \in (0, 1]$, $\kappa \geq 2$, $\eta \in (0, \eta_0/\kappa]$ and $\phi \in C^\gamma_{\eta,H}(H \times \mathbb{T}^n)$ we have*

$$|P_t\widetilde{\phi}(w, h)| \leq C\|\phi\|_{\gamma,\eta,H} e^{\kappa\eta\|w\|^2} e^{-\Lambda t}, \quad \forall t \geq 0, (w, h) \in H \times \mathbb{T}^n, \tag{6.10}$$

where $\Lambda = \frac{\varpi}{5(2-\gamma)}$, the mixing rate ϖ is from Theorem 5.1, and C is a positive constant independent of (w, h) and η_0 is the constant from (A.1).

Proof. For any $R > 0$, let $\chi_R : H \to \mathbb{R} \in C^\gamma_{\eta,H}(H \times \mathbb{T}^n)$ satisfying $0 \leq \chi_R \leq 1$, with $\chi_R(w) = 1$ for $\|w\| \leq R$ and $\chi_R(w) = 0$ for $\|w\| \geq R + 1$. We can actually choose a χ_R such that $\|\chi_R\|_{\gamma,\eta,H} \leq 2$. Assume without loss of generality that $\|\phi\|_{\gamma,\eta,H} = 1$. Also denote by $\overline{\chi}_R = 1 - \chi_R$. Then

$$\begin{aligned}
|P_t\widetilde{\phi}(w, h)| &= |P_t\left(\chi_R\phi + \overline{\chi}_R\phi\right)(w, h) - \langle \Gamma_{\beta_t h}, (\chi_R\phi)(\cdot, \beta_t h) + (\overline{\chi}_R\phi)(\cdot, \beta_t h)\rangle| \\
&\leq |P_t(\chi_R\phi)(w, h) - \langle \Gamma_{\beta_t h}, (\chi_R\phi)(\cdot, \beta_t h)\rangle| + |P_t(\overline{\chi}_R\phi)(w, h) - \langle \Gamma_{\beta_t h}, (\overline{\chi}_R\phi)(\cdot, \beta_t h)\rangle| \\
&:= I_1 + I_2.
\end{aligned} \tag{6.11}$$

It is straightforward to show that $\chi_R\phi \in C^\gamma_{\eta,H_\rho}(H \times \mathbb{T}^n)$. Indeed, since $\chi_R\phi$ vanishes outside of the ball $\|w\| \leq R + 1$, in view of the definition (6.2), one has

$$\sup_{(w,h)\in H\times\mathbb{T}^n} |\chi_R(w)\phi(w, h)| \leq \sup_{\substack{(w,h)\in H\times\mathbb{T}^n \\ \|w\|\leq R+1}} |\phi(w, h)| \leq \|\phi\|_{\gamma,\eta,H} e^{\eta(R+1)^2}. \tag{6.12}$$

Let

$$S = \{(w_1, w_2) \in H^2 : \|w_1\| \leq R + 1, \|w_2\| \geq R + 1, 0 < \|w_1 - w_2\| \leq 1\},$$

and

$$S_\rho = \{(w_1, w_2) \in H^2 : \|w_1\| \leq R + 1, \|w_2\| \geq R + 1, 0 < \rho(w_1, w_2) \leq 1\}.$$

It is clear that $S_\rho \subset S$ since $\|w_1 - w_2\| \leq \rho(w_1, w_2)$. It follows from $\chi_R(w_2) = 0$, $\|w_2\| \leq R + 2$ and (6.12) that

$$\begin{aligned}
\sup_{(w_1,w_2)\in S_\rho} \frac{|\chi_R(w_1)\phi(w_1, h) - \chi_R(w_2)\phi(w_2, h)|}{\rho(w_1, w_2)^\gamma} &= \sup_{(w_1,w_2)\in S_\rho} \frac{|\chi_R(w_1)\phi(w_1, h) - \chi_R(w_2)\phi(w_1, h)|}{\rho(w_1, w_2)^\gamma} \\
&\leq 2e^{\eta(R+1)^2} \sup_{(w_1,w_2)\in S} \frac{|\chi_R(w_1) - \chi_R(w_2)|}{\|w_1 - w_2\|^\gamma} \\
&\leq 4e^{2\eta(R+2)^2}.
\end{aligned}$$

Now let

$$S^R = \{(w_1, w_2) \in H^2 : \|w_1\|, \|w_2\| \leq R + 1, 0 < \|w_1 - w_2\| \leq 1\},$$

and

$$S_\rho^R = \{(w_1, w_2) \in H^2 : \|w_1\|, \|w_2\| \le R+1, 0 < \rho(w_1, w_2) \le 1\}.$$

First note that for $\|w_1\|, \|w_2\| \le R+1$,

$$|\chi_R(w_1)\phi(w_1, h) - \chi_R(w_2)\phi(w_2, h)| \le 2e^{\eta(R+1)^2}\Big(|\chi_R(w_1) - \chi_R(w_2) + |\phi(w_1, h) - \phi(w_2, h)|\Big).$$

Hence

$$\begin{aligned}
&\sup_{(w_1,w_2)\in S_\rho^R} \frac{|\chi_R(w_1)\phi(w_1, h) - \chi_R(w_2)\phi(w_2, h)|}{\rho(w_1, w_2)^\gamma} \\
&\le 2e^{\eta(R+1)^2}\Big(\sup_{(w_1,w_2)\in S^R} \frac{|\chi_R(w_1) - \chi_R(w_2)|}{\|w_1 - w_2\|^\gamma} + \sup_{(w_1,w_2)\in S^R} \frac{|\phi(w_1, h) - \phi(w_2, h)|}{\|w_1 - w_2\|^\gamma}\Big) \\
&\le 4e^{2\eta(R+2)^2}.
\end{aligned}$$

It then follows that $\chi_R\phi \in C^\gamma_{\eta,H_\rho}(H \times \mathbb{T}^n)$ and

$$\|\chi_R\phi\|_{\gamma,\eta,H_\rho} \le Ce^{2\eta(R+2)^2},$$

where C is a constant that does not depend on R.

It is known that the dual Hölder metric on $\mathcal{P}(H)$ is bounded by the Wasserstein metric (see Proposition 1.2.6 in [48] for example):

$$\sup_{\varphi\in C_b^\gamma(H),\|\varphi\|_\gamma\le 1} |\langle \mu_1, \varphi\rangle - \langle \mu_2, \varphi\rangle| \le 5(\rho(\mu_1, \mu_2))^{\frac{1}{2-\gamma}}, \quad \forall \mu_1, \mu_2 \in \mathcal{P}(H), \tag{6.13}$$

where $C_b^\gamma(H)$ is the space of bounded $\gamma-$Hölder continuous functions on H endowed with the metric ρ, and

$$\|\varphi\|_\gamma = \sup_{w\in H} |\varphi(w)| + \sup_{0<\rho(w_1,w_2)\le 1} \frac{|\varphi(w_1) - \varphi(w_2)|}{\rho(w_1, w_2)^\gamma}.$$

Combining this fact with Theorem 5.1, it follows that the first term in (6.11) satisfies

$$\begin{aligned}
I_1 &= |P_t(\chi_R\phi)(w, h) - \langle \Gamma_{\beta_t h}, (\chi_R\phi)(\cdot, \beta_t h)\rangle| \\
&= |\langle \mathcal{P}^*_{0,t,h}\delta_w, (\chi_R\phi)(\cdot, \beta_t h)\rangle - \langle \mathcal{P}^*_{0,t,h}\Gamma_h, (\chi_R\phi)(\cdot, \beta_t h)\rangle| \\
&\le 5\|\chi_R\phi\|_{\gamma,\eta,H_\rho}(\rho(\mathcal{P}^*_{0,t,h}\delta_w, \mathcal{P}^*_{0,t,h}\Gamma_h))^{\delta_0} \le Ce^{2\eta(R+2)^2}e^{-\delta_0\varpi t}e^{\kappa\eta\|w\|^2}.
\end{aligned}$$

Here $\delta_0 = \frac{1}{2-\gamma} \le 1, \kappa \ge 2$ and ϖ are positive constants.

To estimate the second term in (6.11), observe that

$$\begin{aligned}|P_t(\overline{\chi}_R\phi)(w,h)| &\le \mathbf{E}_{(w,h)}(|\overline{\chi}_R\phi|)(X_t) \le (\mathbf{E}_{(w,h)}\overline{\chi}_R(X_t))^{1/2}(\mathbf{E}_{(w,h)}|\phi(X_t)|^2)^{1/2}\\ &\le \Big(\mathbf{P}\left(\|w_{0,t,h}(w)\| \ge R\right)\Big)^{1/2}\Big(\mathbf{E}e^{2\eta\|w_{0,t,h}(w)\|^2}\Big)^{1/2}\\ &\le Ce^{\eta\|w\|^2}e^{-\eta R^2}(\mathbf{E}e^{2\eta\|w_{0,t,h}(w)\|^2})^{1/2} \le Ce^{2\eta\|w\|^2}e^{-\eta R^2},\end{aligned}$$

where we used the fact that $\|\phi\|_{\gamma,\eta,H} = 1$ and the estimate (A.1) with a smaller η_0 (which is the one in (A.1) divided by 2).

In a similar fashion, note that

$$\begin{aligned}|\langle\Gamma_{\beta_t h},(\overline{\chi}_R\phi)(\cdot,\beta_t h)\rangle| &\le \Big(\int_H \overline{\chi}_R(w)\Gamma_{\beta_t h}(dw)\Big)^{1/2}\Big(\int_H |\phi(w,\beta_t h)|^2\,\Gamma_{\beta_t h}(dw)\Big)^{1/2}\\ &\le \Big(\Gamma_{\beta_t h}(\|w\|\ge R)\Big)^{1/2}\left(\int_H e^{2\eta\|w\|^2}\Gamma_{\beta_t h}(dw)\right)^{1/2}\\ &\le Ce^{-\eta R^2}\left(\int_H e^{2\eta\|w\|^2}\Gamma_{\beta_t h}(dw)\right)^{1/2} \le Ce^{-\eta R^2}.\end{aligned}$$

As a result, we have the following estimate on (6.11),

$$|P_t\widetilde{\phi}(w,h)| \le Ce^{\kappa\eta\|w\|^2}(e^{4\eta R^2-\delta_0\varpi t}+e^{-\eta R^2}).$$

By choosing $R^2 = \frac{\delta_0\varpi t}{5\eta}$, we obtain (6.10) with $\Lambda = \frac{\delta_0\varpi}{5} = \frac{\varpi}{5(2-\gamma)}$. The proof is complete. □

We now define the corrector that will be used in the martingale approximation procedure.

Proposition 6.7 (The corrector)**.** *For $\phi \in C^\gamma_{\eta,H}(H\times\mathbb{T}^n)$, define*

$$\chi(w,h) := \int_0^\infty P_t\widetilde{\phi}(w,h)dt = \int_0^\infty \mathcal{P}_{0,t,h}\phi(\cdot,\beta_t h)(w) - \langle\Gamma_{\beta_t h},\phi(\cdot,\beta_t h)\rangle dt, \quad (w,h)\in H\times\mathbb{T}^n.$$

(1). For $\eta\in(0,\eta_0/2]$, $\gamma\in(0,1]$, the corrector $\chi\in C^{\gamma_0}_{2\eta,H}(H\times\mathbb{T}^n)$ as long as $\phi\in C^\gamma_{\eta,H}(H\times\mathbb{T}^n)$, where $\gamma_0 = \frac{\Lambda\gamma}{\Lambda+r}$, with Λ from Theorem 6.6 and $r = 64c_0^6\eta^{-3}\nu^{-5}+\eta C(f,\mathcal{B}_0)$ is the constant from (6.8)*.*

(2). If we assume $\Psi\in C^\gamma(\mathbb{T}^n,H)$ and $\phi\in C^\gamma_{\eta,H}(H\times\mathbb{T}^n)\cap C^\gamma_{\eta,\mathbb{T}^n}(H\times\mathbb{T}^n)$, then the associated corrector $\chi\in C^{\gamma_0}_{2\eta,H}(H\times\mathbb{T}^n)\cap C^{\gamma_1}_{2\eta,\mathbb{T}^n}(H\times\mathbb{T}^n)$, where $\gamma_1 = \frac{\Lambda\zeta}{5(\Lambda+r)(2-\gamma)}$ with Λ, r as above and ζ from Theorem 5.1. In particular, we have

$$\chi\in C^{\overline{\gamma}_0}_{2\eta,H}(H\times\mathbb{T}^n)\cap C^{\overline{\gamma}_0}_{2\eta,\mathbb{T}^n}(H\times\mathbb{T}^n),$$

where $\overline{\gamma}_0 = \min\{\gamma_0,\gamma_1\} = \gamma_1$.

Proof. The function χ is well defined in view of Theorem 6.6. We begin with proving the first item.

Let $\phi \in C^{\gamma}_{\eta,H}(H \times \mathbb{T}^n)$. It follows from Theorem 6.6 and inequality (6.8) that, for $\eta \in (0, \eta_0/2]$,

$$\begin{aligned}|\chi(w_1,h) - \chi(w_2,h)| &\leq \int_0^T |P_t\widetilde{\phi}(w_1,h) - P_t\widetilde{\phi}(w_2,h)|dt + \int_T^\infty |P_t\widetilde{\phi}(w_1,h)| + |P_t\widetilde{\phi}(w_2,h)|dt \\ &\leq \int_0^T |P_t\phi(w_1,h) - P_t\phi(w_2,h)|dt + C\|\phi\|_{\gamma,\eta,H}e^{-\Lambda T}\left(e^{2\eta\|w_1\|^2} + e^{2\eta\|w_2\|^2}\right) \\ &\leq C\|\phi\|_{\gamma,\eta,H}\left(\|w_1 - w_2\|^\gamma e^{\frac{\gamma r}{2}T} + e^{-\Lambda T}\right)\left(e^{2\eta\|w_1\|^2} + e^{2\eta\|w_2\|^2}\right) \\ &\leq C\|\phi\|_{\gamma,\eta,H}\left(\|w_1 - w_2\|^\gamma e^{rT} + e^{-\Lambda T}\right)\left(e^{2\eta\|w_1\|^2} + e^{2\eta\|w_2\|^2}\right)\end{aligned}$$

where r is from inequality (6.8). In view of Lemma 5.2, we have for any $0 < \|w_1 - w_2\| \leq 1$,

$$|\chi(w_1,h) - \chi(w_2,h)| \leq C\|\phi\|_{\gamma,\eta,H}\left(e^{4\eta\|w_1\|^2} + e^{4\eta\|w_2\|^2}\right)\|w_1 - w_2\|^{\gamma_0},$$

with $\gamma_0 = \frac{\Lambda\gamma}{\Lambda + r}$. This also indicates that $\chi(w,h)$ is continuous in w uniformly for h. The continuity of $\chi(w,h)$ in h for fixed w follows from the fact that $\chi_T(w,h) := \int_0^T P_t\widetilde{\phi}(w,h)dt$ is continuous in h and $\chi_T(w,h) \to \chi(w,h)$ uniformly for h. Thus $\chi \in C(H \times \mathbb{T}^n)$. It also follows from Theorem 6.6 that $|\chi(w,h)| \leq C\|\phi\|_{\gamma,\eta,H}e^{2\eta\|w\|^2}$. Hence $\chi \in C^{\gamma_0}_{2\eta,H}(H \times \mathbb{T}^n)$.

We now prove the second assertion. Let $\phi \in C^{\gamma}_{\eta,H}(H \times \mathbb{T}^n) \cap C^{\gamma}_{\eta,\mathbb{T}^n}(H \times \mathbb{T}^n)$ and assume $\Psi \in C^\gamma(\mathbb{T}^n, H)$. Then

$$\begin{aligned}|\chi(w,h_1) - \chi(w,h_2)| &\leq \int_0^T |P_t\widetilde{\phi}(w,h_1) - P_t\widetilde{\phi}(w,h_2)|dt + \int_T^\infty |P_t\widetilde{\phi}(w,h_1)| + |P_t\widetilde{\phi}(w,h_2)|dt \\ &:= I + II.\end{aligned}$$

Note that by Theorem 6.6, the second term can be estimated as

$$II \leq C\|\phi\|_{\gamma,\eta,H}e^{2\eta\|w\|^2}e^{-\Lambda T}.$$

For the first term, observe that

$$\begin{aligned}I &\leq \int_0^T |P_t\phi(w,h_1) - P_t\phi(w,h_2)|dt + \int_0^T |\langle\Gamma_{\beta_t h_1}, \phi(\cdot,\beta_t h_1)\rangle - \langle\Gamma_{\beta_t h_2}, \phi(\cdot,\beta_t h_2)\rangle|\, dt \\ &:= I_1 + I_2.\end{aligned}$$

It then follows from the estimate (6.9) that

$$\begin{aligned}I_1 &\leq C(\|\phi\|_{\gamma,\eta,\mathbb{T}^n} + \|\phi\|_{\gamma,\eta}\|\Psi\|^\gamma_\gamma)e^{2\eta\|w\|^2}|h_1 - h_2|^\gamma \int_0^T e^{\frac{r\gamma}{2}t}dt \\ &\leq C(\|\phi\|_{\gamma,\eta,\mathbb{T}^n} + \|\phi\|_{\gamma,\eta}\|\Psi\|^\gamma_\gamma)e^{2\eta\|w\|^2}|h_1 - h_2|^\gamma e^{rT}.\end{aligned}$$

To estimate I_2, noting that as in the proof of Theorem 6.6, especially by (6.13) and Theorem 5.1

(ζ-Hölder continuity of Γ), we have

$$
\begin{aligned}
I_{21} :=& |\langle \Gamma_{\beta_t h_1}, \phi(\cdot, \beta_t h_1)\rangle - \langle \Gamma_{\beta_t h_2}, \phi(\cdot, \beta_t h_1)\rangle| \\
&\leq |\langle \Gamma_{\beta_t h_1} - \Gamma_{\beta_t h_2}, \chi_R \phi(\cdot, \beta_t h_1)\rangle| + |\langle \Gamma_{\beta_t h_1}, (\overline{\chi}_R \phi)(\cdot, \beta_t h_1)\rangle| + |\langle \Gamma_{\beta_t h_2}, (\overline{\chi}_R \phi)(\cdot, \beta_t h_1)\rangle| \\
&\leq C\|\phi\|_{\gamma,\eta,H} e^{4\eta R^2} \|\Gamma\|_\zeta^{\frac{1}{2-\gamma}} |h_1 - h_2|^{\frac{\zeta}{2-\gamma}} + C\|\phi\|_{\gamma,\eta,H} e^{-\eta R^2} \\
&\leq C\|\phi\|_{\gamma,\eta,H} (\|\Gamma\|_\zeta^{\frac{1}{2-\gamma}} + 1) \left(e^{4\eta R^2} |h_1 - h_2|^{\frac{\zeta}{2-\gamma}} + e^{-\eta R^2} \right) \\
&\leq C\|\phi\|_{\gamma,\eta,H} (\|\Gamma\|_\zeta^{\frac{1}{2-\gamma}} + 1) |h_1 - h_2|^{\frac{\zeta}{5(2-\gamma)}}
\end{aligned}
$$

by Lemma 5.2. Also note that by Theorem 5.1,

$$
\begin{aligned}
I_{22} :=& |\langle \Gamma_{\beta_t h_2}, \phi(\cdot, \beta_t h_1)\rangle - \langle \Gamma_{\beta_t h_2}, \phi(\cdot, \beta_t h_2)\rangle| \\
&\leq \langle \Gamma_{\beta_t h_2}, |\phi(\cdot, \beta_t h_1) - \phi(\cdot, \beta_t h_2)|\rangle \\
&\leq \|\phi\|_{\gamma,\eta,\mathbb{T}^n} |h_1 - h_2|^\gamma \int_H e^{\eta\|w\|^2} \Gamma_{\beta_t h_2}(dw) \leq C\|\phi\|_{\gamma,\eta,\mathbb{T}^n} |h_1 - h_2|^\gamma.
\end{aligned}
$$

As a result,

$$
I_2 \leq \int_0^T I_{21} + I_{22} dt \leq CT \left(\|\phi\|_{\gamma,\eta,H} (\|\Gamma\|_\zeta^{\frac{1}{2-\gamma}} + 1) + \|\phi\|_{\gamma,\eta,\mathbb{T}^n} \right) |h_1 - h_2|^{\frac{\zeta}{5(2-\gamma)}}.
$$

Therefore by Lemma 5.2 again,

$$
\begin{aligned}
|\chi(w, h_1) - \chi(w, h_2)| &\leq C e^{2\eta\|w\|^2} \left(e^{rT} |h_1 - h_2|^{\frac{\zeta}{5(2-\gamma)}} + e^{-\Lambda T} \right) \\
&\leq C e^{2\eta\|w\|^2} |h_1 - h_2|^{\gamma_1},
\end{aligned}
$$

with $\gamma_1 = \frac{\Lambda\zeta}{5(\Lambda+r)(2-\gamma)}$. Hence $\chi \in C^{\gamma_1}_{2\eta,\mathbb{T}^n}(H \times \mathbb{T}^n)$. This completes the proof of this proposition. $\square$

We are now in a position to give the martingale approximation. For $T \geq 0$, let

$$
\int_0^T \widetilde{\phi}(X_t) dt = \int_0^N \widetilde{\phi}(X_t) dt + \int_N^T \widetilde{\phi}(X_t) dt = M_N + R_{N,T}, \tag{6.14}
$$

where N is the integer part of T,

$$
M_N = \chi(X_N) - \chi(X_0) + \int_0^N \widetilde{\phi}(X_t) dt
$$

is the Dynkin martingale (formally) and

$$
R_{N,T} = -\chi(X_N) + \chi(X_0) + \int_N^T \widetilde{\phi}(X_t) dt
$$

is the reminder term. Let $Z_N = M_N - M_{N-1}$ for $N \geq 1$ be the associated martingale difference. In what follows, we will show that M_N is indeed a martingale and $R_{N,T}$ is a negligible term that vanishes as $T \to \infty$. This will reduce the proof of Theorem 6.1-6.2 to the proof of the limit theorems for the associated martingale sequence M_N. Let $M_T = \chi(X_T) - \chi(X_0) + \int_0^T \widetilde{\phi}(X_t)dt$ for $T \geq 0$.

Lemma 6.8. *$\{M_T\}_{T\geq 0}$ is a zero mean martingale w.r.t the filtration $\{\mathcal{F}_T\}$.*

Proof. The martingale property follows from the Markov property of the homogenized process X_t as follows.

$$\mathbf{E}[M_T|\mathcal{F}_s] = \mathbf{E}[\chi(X_T)|\mathcal{F}_s] - \chi(X_0) + \int_0^s \mathbf{E}[\widetilde{\phi}(X_u)|\mathcal{F}_s]du + \int_s^T \mathbf{E}[\widetilde{\phi}(X_u)|\mathcal{F}_s]du.$$

Since X_u is $\mathcal{F}_s$ measurable for $0 \leq u \leq s$, it follows that

$$\int_0^s \mathbf{E}[\widetilde{\phi}(X_u)|\mathcal{F}_s]du = \int_0^s \widetilde{\phi}(X_u)du.$$

Moreover, by the Markov property,

$$\int_s^T \mathbf{E}[\widetilde{\phi}(X_u)|\mathcal{F}_s]du = \int_s^\infty P_{u-s}\widetilde{\phi}(X_s)du - \int_T^\infty P_{u-T}(P_{T-s}\widetilde{\phi})(X_s)du = \chi(X_s) - \mathbf{E}[\chi(X_T)|\mathcal{F}_s].$$

Hence $\mathbf{E}[M_T|\mathcal{F}_s] = M_s$.

It is zero mean since

$$\begin{aligned} M_T &= \chi(X_T) - \chi(X_0) + \int_0^T \widetilde{\phi}(X_t)dt \\ &= \int_T^\infty P_{t-T}\widetilde{\phi}(X_T)dt - \int_0^\infty P_t\widetilde{\phi}(X_0)dt + \int_0^T \widetilde{\phi}(X_t)dt \\ &= \int_T^\infty \mathbf{E}_{(w,h)}\left[\widetilde{\phi}(X_t)|\mathcal{F}_T\right]dt - \int_0^\infty P_t\widetilde{\phi}(w,h)dt + \int_0^T \widetilde{\phi}(X_t)dt, \end{aligned}$$

which implies $\mathbf{E}_{(w,h)}M_T = 0$. □

The following lemma gives estimates on the even order moments of the martingale M_N and its associated martingale difference.

Lemma 6.9 (Bounds on the martingale)**.** *For integer $p \geq 1$, $\eta \in (0, 2^{-p-1}\eta_0]$ and $\phi \in C^\gamma_{\eta,H}(H \times \mathbb{T}^n)$, one has*

$$\mathbf{E}_{(w,h)}|M_T|^{2^p} \leq C(T^{2-2^{-p}} + 1)e^{2^{p+1}\eta\|w\|^2}, \quad \mathbf{E}_{(w,h)}|Z_N|^{2^p} \leq Ce^{2^{p+1}\eta\|w\|^2},$$

for $T \geq 0$ and $N \geq 1$. Also with a larger constant C,

$$P_t\mathbf{E}_{(w,h)}|M_T|^{2p} \leq C(T^{2-2^{-p}} + 1)e^{2^{p+1}\eta\|w\|^2}, \quad \forall t \geq 0.$$

Proof. By Proposition 6.7, we know that $\chi \in C^{\gamma_0}_{2\eta,H}(H \times \mathbb{T}^n)$. Hence $\chi^{2^p} \in C^{\gamma_0}_{2^{p+1}\eta,H}(H \times \mathbb{T}^n)$. Besides, since $\phi \in C^\gamma_{\eta,H}(H \times \mathbb{T}^n)$, $|\widetilde{\phi}|^{2^p} \in C^\gamma_{2^{p+1}\eta,H}(H \times \mathbb{T}^n)$. It follows from estimate (A.1) that for

$\eta \in (0, 2^{-p-1}\eta_0]$, and any $t \geq 0$,

$$\begin{aligned}
\mathbf{E}_{(w,h)}|M_T|^{2^p} &\leq C\left(\mathbf{E}_{(w,h)}|\chi(X_T)|^{2^p} + |\chi(w,h)|^{2^p} + T^{1-2^{-p}}\int_0^T \mathbf{E}_{(w,h)}|\widetilde{\phi}(X_t)|^{2^p}dt\right) \\
&\leq C\left(\mathbf{E}e^{2^{p+1}\eta\|w_{0,T,h}(w)\|^2} + e^{2^{p+1}\eta\|w\|^2} + T^{1-2^{-p}}\int_0^T \mathbf{E}e^{2^{p+1}\eta\|w_{0,t,h}(w)\|^2}dt\right) \\
&\leq C(T^{2-2^{-p}}+1)e^{2^{p+1}\eta\|w\|^2}.
\end{aligned}$$

Similarly, one can show that for any $N \geq 1$,

$$\mathbf{E}_{(w,h)}|Z_N|^{2^p} \leq Ce^{2^{p+1}\eta\|w\|^2},$$

where C does not depend on N, h. It follows from (A.1) that

$$P_t\mathbf{E}_{(w,h)}|M_T|^{2^p} \leq C(T^{2-2^{-p}}+1)\mathbf{E}e^{2^{p+1}\eta\|w_{0,t,h}(w)\|^2} \leq C(T^{2-2^{-p}}+1)e^{2^{p+1}\eta\|w\|^2}. \qquad \square$$

The following estimate on the remainder term shows that the proof of Theorem 6.1-6.2 can be reduced to the proof of the corresponding limit theorems for the associated martingale sequence.

Lemma 6.10. *Let $R_{N,T}$ be as in (6.14). Then for any initial condition $X_0 = (w,h)$, $\eta \in (0, 2^{-4}\eta_0]$ and $\phi \in C^\gamma_{\eta,H}(H \times \mathbb{T}^n)$,*

$$\lim_{T\to\infty} \frac{1}{\sqrt{T}} R_{N,T} = 0, \quad \mathbf{P}-\text{a.s.} \tag{6.15}$$

Proof. Since N is the integer part of T, it suffices to show

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} \sup_{N\leq t\leq N+1} R_{N,t} = 0, \quad \mathbf{P}-\text{a.s.}$$

By Proposition 6.7, we have that

$$|\chi(X_N)| \leq Ce^{2\eta\|w_{0,N,h}(w)\|^2}.$$

Since $\phi \in C^\gamma_{\eta,H}(H\times\mathbb{T}^n)$, it also holds that

$$\sup_{N\leq t\leq N+1}\left|\int_N^t \widetilde{\phi}(X_s)ds\right| \leq C \sup_{N\leq t\leq N+1} e^{2\eta\|w_{0,t,h}(w)\|^2}. \tag{6.16}$$

It then follows from the Markov inequality, estimates (A.2) and (A.1) that for any $K > 0$,

$$\mathbf{P}\left(\sup_{N\leq t\leq N+1} e^{2\eta\|w_{0,t,h}(w)\|^2} > K\right) \leq Ce^{2^4\eta\|w\|^2}K^{-8}.$$

Hence

$$\begin{aligned}
&\sum_{N=1}^\infty \mathbf{P}\left(\sup_{N\leq t\leq N+1}\left(|\chi(X_N)| + \chi(w,h) + \left|\int_N^t \widetilde{\phi}(X_s)ds\right|\right) \geq N^{\frac{1}{4}}\right) \\
&\leq \sum_{N=1}^\infty \mathbf{P}\left(C\sup_{N\leq t\leq N+1} e^{2\eta\|w_{0,t,h}(w)\|^2} \geq N^{\frac{1}{4}}\right) \leq Ce^{2^4\eta\|w\|^2}\sum_{N=1}^\infty N^{-2} < \infty,
\end{aligned}$$

By the Borel-Cantelli lemma, there is an almost surely finite random integer time $N_0(\omega)$ such that for $N > N_0(\omega)$,

$$\sup_{N \le t \le N+1} R_{N,t} \le N^{1/4}, \tag{6.17}$$

which implies (6.15). □

6.2 The Limit Theorems

Based on the martingale approximation given above, we will prove Theorem 6.1 and Theorem 6.2 in this section by showing the limit results for the corresponding martingales. As remarked earlier, compared with the analysis in the next section the rate of convergence, the proof of the limit theorems is quite straightforward and merely requires certain moment bounds as given in Lemma 6.9 and ergodicity properties given in Proposition 5.3. We first prove the strong law of large numbers, which is based on the following Kolmogorov's criterion for martingales.

Theorem 6.11 ([39, 48]). *Let $\{M_N\}_{N\ge 1}$ be a zero mean square integrable martingale and let $\{c_N\}$ be an increasing sequence going to ∞ such that*

$$\sum_{N=1}^{\infty} c_N^{-2} \mathbf{E} Z_N^2 < \infty,$$

where $Z_N = M_N - M_{N-1}$ and $M_0 = 0$. Then

$$\lim_{N\to\infty} c_N^{-1} M_N \to 0, \quad \mathbf{P} - \text{a.s.}$$

Proof of Theorem 6.1. In view of the martingale approximation (6.14) and Theorem 6.10, to show (6.6), it suffices to prove that

$$\lim_{N\to\infty} N^{-\frac{1}{2}-\varepsilon} M_N = 0, \quad \mathbf{P} - \text{a.s.}$$

Lemma 6.9 with $p = 1$ ensures the condition of Theorem 6.11 with $c_N = N^{1/2+\varepsilon}$. Hence the desired convergence follows from Theorem 6.11. □

The rest of this section is devoted to the proof of Theorem 6.2. We first prove the existence of the asymptotic variance.

Proposition 6.12 (The asymptotic variance). *For any $(w, h) \in H \times \mathbb{T}^n$, $\eta \in (0, \eta_0/16]$ and*

$\phi \in C^{\gamma}_{\eta,H}(H \times \mathbb{T}^n)$, *we have*

$$\lim_{T\to\infty} \frac{1}{T}\mathbf{E}\left(\int_0^T \widetilde{\phi}\left(X_t(w,h)\right)dt\right)^2 = 2\int_{H\times\mathbb{T}^n} \widetilde{\phi}(w,h)\chi(w,h)\Gamma_h(dw)\lambda(dh) := \sigma^2_{\phi}. \tag{6.18}$$

Proof. By the Markov property of the homogenized process, one has

$$\begin{aligned}
\frac{1}{T}\mathbf{E}\left(\int_0^T \widetilde{\phi}\left(X_t(w,h)\right)dt\right)^2 &= \frac{2}{T}\mathbf{E}\int_0^T\int_s^T \widetilde{\phi}\left(X_t(w,h)\right)\widetilde{\phi}\left(X_s(w,h)\right)dtds \\
&= \frac{2}{T}\int_0^T \mathbf{E}\left[\widetilde{\phi}\left(X_s(w,h)\right)\int_s^T \mathbf{E}\left[\widetilde{\phi}\left(X_t(w,h)\right)|\mathcal{F}_s\right]dt\right]ds \\
&= \frac{2}{T}\int_0^T \left\langle P_s^*\delta_{(w,h)}, \widetilde{\phi}\int_0^{T-s} P_t\widetilde{\phi}dt\right\rangle ds.
\end{aligned}$$

In view of the weak convergence in Proposition 5.3 and the definition of the corrector in Proposition 6.7, we expect that as $T\to\infty$,

$$\begin{aligned}
&\Big|\frac{2}{T}\int_0^T \left\langle P_s^*\delta_{(w,h)}, \widetilde{\phi}\int_0^{T-s} P_t\widetilde{\phi}dt\right\rangle ds - \frac{2}{T}\int_0^T \left\langle P_s^*\delta_{(w,h)}, \widetilde{\phi}\chi\right\rangle ds\Big| \\
&\qquad\qquad = \left|\frac{2}{T}\int_0^T \left\langle P_s^*\delta_{(w,h)}, \widetilde{\phi}(\chi-\chi_{T-s})\right\rangle ds\right| \to 0,
\end{aligned} \tag{6.19}$$

where $\chi_{T-s} = \int_0^{T-s} P_t\widetilde{\phi}dt$. Indeed, it follows from Theorem 6.6 with $\kappa = 2$ that

$$|\chi - \chi_{T-s}| \leq \int_{T-s}^{\infty}\left|P_t\widetilde{\phi}(w,h)\right|dt \leq Ce^{2\eta\|w\|^2}e^{-\Lambda(T-s)}.$$

Since $\widetilde{\phi} \in C^{\gamma}_{\eta,H}(H\times\mathbb{T}^n)$, and $e^{2\eta\|\cdot\|^2} \in C^{\gamma}_{4\eta,H}(H\times\mathbb{T}^n)$, it follows that $\widetilde{\phi}e^{2\eta\|\cdot\|^2} \in C^{\gamma}_{8\eta,H}(H\times\mathbb{T}^n)$. Hence by estimate (A.1), we have for $\eta \in (0,\eta_0/8]$, and any $s \geq 0$,

$$\left\langle P_s^*\delta_{(w,h)}, |\widetilde{\phi}|e^{2\eta\|\cdot\|^2}\right\rangle \leq C\left\langle P_s^*\delta_{(w,h)}, e^{8\eta\|\cdot\|^2}\right\rangle = C\mathbf{E}e^{8\eta\|w_{0,s,h}(w)\|^2} \leq Ce^{8\eta\|w\|^2}.$$

Hence

$$\left|\frac{1}{T}\int_0^T \left\langle P_s^*\delta_{(w,h)}, \widetilde{\phi}(\chi-\chi_{T-s})\right\rangle ds\right| \leq Ce^{8\eta\|w\|^2}\frac{1}{T}\int_0^T e^{-\Lambda(T-s)}ds \to 0,$$

which implies the limit (6.19).

By Proposition 6.7, $\chi \in C^{\gamma_0}_{2\eta,H}(H\times\mathbb{T}^n)$. Hence $|\widetilde{\phi}\chi|^2 \in C^{\gamma_0}_{8\eta,H}(H\times\mathbb{T}^n)$. Then by estimate (A.1), for $\eta \in (0,\eta_0/8]$,

$$\limsup_{T\to\infty}\frac{1}{T}\int_0^T \left\langle P_s^*\delta_{(w,h)}, \left|\widetilde{\phi}\chi\right|^2\right\rangle ds \leq C\limsup_{T\to\infty}\frac{1}{T}\int_0^T \mathbf{E}e^{8\eta\|w_{0,s,h}(w)\|^2}ds < \infty.$$

Combining this moment bound with the weak convergence in Proposition 5.3, we obtain the desired convergence

$$\lim_{T\to\infty}\frac{2}{T}\int_0^T \left\langle P_s^*\delta_{(w,h)}, \widetilde{\phi}\chi\right\rangle ds = 2\int_{H\times\mathbb{T}^n}\widetilde{\phi}(w,h)\chi(w,h)\Gamma_h(dw)\lambda(dh),$$

which combined with (6.19) implies the desired (6.18). □

The following proposition gives further properties related to the asymptotic variance. In particular, the Hölder regularity of the particular observable function F as below, plays an important role when estimating the rate of convergence in the central limit theorem.

Proposition 6.13. *For $\gamma \in (0,1]$, $\eta \in (0, 2^{-5}\eta_0]$, $\phi \in C^{\gamma}_{\eta,H}(H \times \mathbb{T}^n) \cap C^{\gamma}_{\eta,\mathbb{T}^n}(H \times \mathbb{T}^n)$, and $X_0 = (w,h)$, let*

$$Y(w,h) = \mathbf{E}_{(w,h)} M_1^2 = \mathbf{E}_{(w,h)} \left(\chi(X_1) - \chi(X_0) + \int_0^1 \widetilde{\phi}(X_t)dt \right)^2 .$$

Assume $\Psi \in C^{\gamma}(\mathbb{T}^n, H)$. Then $Y \in C^{\overline{\gamma}_0}_{2^5\eta,H}(H \times \mathbb{T}^n) \cap C^{\overline{\gamma}_0}_{2^5\eta,\mathbb{T}^n}(H \times \mathbb{T}^n)$. Furthermore, the function

$$F(h) := \int_H Y(w,h)\Gamma_h(dw) = \langle \Gamma_h, Y(\cdot,h)\rangle$$

is in $C^{\overline{\gamma}_0}(\mathbb{T}^n, \mathbb{R})$. Here $\overline{\gamma}_0$ is taken from Proposition 6.7. We also have

$$\sigma^2 = \int_{\mathbb{T}^n} F(h)\lambda(dh) = \int_{H\times\mathbb{T}^n} Y(w,h)\Gamma_h(dw)\lambda(dh). \tag{6.20}$$

Proof. It follows from the Markov property that

$$\begin{aligned} Y(w,h) = \chi^2(w,h) + P_1\chi^2(w,h) - 2\chi(w,h)P_1\chi(w,h) + 2\int_0^1 P_t(\widetilde{\phi}P_{1-t}\chi)(w,h)dt \qquad (6.21) \\ - 2\chi(w,h)\int_0^1 P_t\widetilde{\phi}(w,h)dt + 2\int_0^1\int_0^t P_\tau\left(\widetilde{\phi}P_{t-\tau}\widetilde{\phi}\right)(w,h)d\tau dt. \end{aligned}$$

By Proposition 6.7 we know that $\chi \in C^{\overline{\gamma}_0}_{2\eta,H}(H \times \mathbb{T}^n) \cap C^{\overline{\gamma}_0}_{2\eta,\mathbb{T}^n}(H \times \mathbb{T}^n)$. Hence for $\eta \in (0, \eta_0/2]$,

$$\chi^2 \in C^{\overline{\gamma}_0}_{4\eta,H}(H \times \mathbb{T}^n) \cap C^{\overline{\gamma}_0}_{4\eta,\mathbb{T}^n}(H \times \mathbb{T}^n).$$

Since $\overline{\gamma}_0 < \gamma$, it follows from Proposition 6.5 that for $\eta \in (0, 2^{-3}\eta_0]$, one has

$$P_1\chi \in C^{\overline{\gamma}_0}_{4\eta,H}(H \times \mathbb{T}^n) \cap C^{\overline{\gamma}_0}_{4\eta,\mathbb{T}^n}(H \times \mathbb{T}^n),$$

and for $t \in [0,1]$,

$$\widetilde{\phi}P_{1-t}\chi, \chi P_1\chi, P_1\chi^2 \in C^{\overline{\gamma}_0}_{2^3\eta,H}(H \times \mathbb{T}^n) \cap C^{\overline{\gamma}_0}_{2^3\eta,\mathbb{T}^n}(H \times \mathbb{T}^n).$$

It then follows from (6.8) and (6.9) that for $\eta \in (0, 2^{-4}\eta_0]$,

$$\int_0^1 P_t(\widetilde{\phi}P_{1-t}\chi)(w,h)dt \in C^{\overline{\gamma}_0}_{2^4\eta,H}(H \times \mathbb{T}^n) \cap C^{\overline{\gamma}_0}_{2^4\eta,\mathbb{T}^n}(H \times \mathbb{T}^n).$$

In a similar way, one can deduce that the remaining two integrals in (6.21) also belong to the same function space. This shows that $Y \in C^{\overline{\gamma}_0}_{2^4\eta,H}(H \times \mathbb{T}^n) \cap C^{\overline{\gamma}_0}_{2^4\eta,\mathbb{T}^n}(H \times \mathbb{T}^n)$.

Note that

$$|F(h_1) - F(h_2)| \le |\langle \Gamma_{h_1}, Y(\cdot, h_1)\rangle - \langle \Gamma_{h_2}, Y(\cdot, h_1)\rangle| + |\langle \Gamma_{h_2}, Y(\cdot, h_1)\rangle - \langle \Gamma_{h_2}, Y(\cdot, h_2)\rangle|$$
$$:= I_1 + I_2.$$

And using the same functions $\chi_R, \overline{\chi}_R$ as in the proof of Theorem 6.6, together with the Hölder continuity of Γ_h and the fact (6.13), we have for $\eta \in (0, 2^{-5}\eta_0]$,

$$\begin{aligned} I_1 &\le |\langle \Gamma_{h_1} - \Gamma_{h_2}, (\chi_R Y)(\cdot, h_1)\rangle| + |\langle \Gamma_{h_1} - \Gamma_{h_2}, (\overline{\chi}_R Y)(\cdot, h_1)\rangle| \\ &\le C\|Y\|_{\overline{\gamma}_0, 2^4\eta, H} e^{2^6\eta R^2} (\rho(\Gamma_{h_1}, \Gamma_{h_2}))^{\frac{1}{2-\overline{\gamma}_0}} + C\|Y\|_{\overline{\gamma}_0, 2^4\eta, H} e^{-2^4\eta R^2} \\ &\le C\|Y\|_{\overline{\gamma}_0, 2^4\eta, H} \left(e^{2^6\eta R^2} |h_1 - h_2|^{\frac{\gamma}{2-\overline{\gamma}_0}} + e^{-2^4\eta R^2} \right), \end{aligned}$$

where we used the uniform integrability $\int_H e^{2^5\eta\|w\|^2}\Gamma_h(dw) \le C$ in the second inequality, which is a consequence of Theorem 5.1 by taking $\kappa = 2^4$. It then follows from Lemma 5.2 that

$$I_1 \le C\|Y\|_{\overline{\gamma}_0, 2^4\eta, H} |h_1 - h_2|^{\frac{\gamma}{5(2-\overline{\gamma}_0)}}.$$

Also note that

$$\begin{aligned} I_2 &\le \langle \Gamma_{h_2}, |Y(\cdot, h_1) - Y(\cdot, h_2)|\rangle \\ &\le \|Y\|_{\overline{\gamma}_0, 2^4\eta, \mathbb{T}^n} |h_1 - h_2|^{\overline{\gamma}_0} \int_H e^{2^4\eta\|w\|^2} \Gamma_{h_2}(dw) \le C\|Y\|_{\overline{\gamma}_0, 2^4\eta, \mathbb{T}^n} |h_1 - h_2|^{\overline{\gamma}_0}. \end{aligned}$$

Since $\overline{\gamma}_0 \le \frac{\gamma}{5(2-\overline{\gamma}_0)}$, we deduce that

$$|F(h_1) - F(h_2)| \le C \left(\|Y\|_{\overline{\gamma}_0, 2^4\eta, \mathbb{T}^n} + \|Y\|_{\overline{\gamma}_0, 2^4\eta, H} \right) |h_1 - h_2|^{\overline{\gamma}_0}, \quad \forall h_1, h_2 \in \mathbb{T}^n.$$

Hence $F \in C^{\overline{\gamma}_0}(\mathbb{T}^n, \mathbb{R})$.

Equation (6.20) follows from the invariance property of the invariant measure $\Gamma_h(dw)\lambda(dh)$ and the decomposition (6.21). Indeed, letting $m(dwdh) = \Gamma_h(dw)\lambda(dh)$ and $\chi_t = \int_0^t P_r \widetilde{\phi} dr$, then by the invariance of m under P_t, one has

$$\int_0^1 \int_0^t \left\langle m, P_\tau \left(\widetilde{\phi} P_{t-\tau} \widetilde{\phi} \right) \right\rangle d\tau dt = \int_0^1 \int_0^t \left\langle m, \widetilde{\phi} P_\tau \widetilde{\phi} \right\rangle d\tau dt = \int_0^1 \left\langle m, \widetilde{\phi} \chi_t \right\rangle dt$$
$$\int_0^1 \left\langle m, P_t(\widetilde{\phi} P_{1-t}\chi) \right\rangle dt = \int_0^1 \left\langle m, \widetilde{\phi} P_t \chi \right\rangle dt = \left\langle m, \widetilde{\phi}\chi \right\rangle - \int_0^1 \left\langle m, \widetilde{\phi}\chi_t \right\rangle dt,$$

where we used the fact that $P_t\chi = \chi - \chi_t$. Hence from (6.21) we have

$$
\begin{aligned}
&\int_{H\times\mathbb{T}^n} Y(w,h)\Gamma_h(dw)\lambda(dh) = \langle m, Y\rangle \\
&= 2\left\langle m,\chi^2\right\rangle - 2\left\langle m,\chi P_1\chi\right\rangle + 2\left\langle m,\widetilde{\phi}\chi\right\rangle - 2\int_0^1 \left\langle m,\widetilde{\phi}\chi_t\right\rangle dt - 2\left\langle m,\chi\chi_1\right\rangle + 2\int_0^1 \left\langle m,\widetilde{\phi}\chi_t\right\rangle dt \\
&= 2\left\langle m,\chi^2\right\rangle - 2\left\langle m,\chi(\chi-\chi_1)\right\rangle + 2\left\langle m,\widetilde{\phi}\chi\right\rangle - 2\left\langle m,\chi\chi_1\right\rangle \\
&= 2\left\langle m,\widetilde{\phi}\chi\right\rangle = \sigma^2
\end{aligned}
$$

as in (6.18). The proof is complete. □

We now proceed to prove the central limit theorem in Theorem 6.2. This is done by utilizing the following martingale central limit theorem to the approximating martingale M_N in (6.14). It is notable that the theorem only requires a weak form of law of large numbers for the martingale difference and Lindeberg type negligible conditions. Basically these conditions can be derived from the ergodic properties of the homogenized process as in Proposition 5.3 and certain moment bounds on the martingale from Lemma 6.9. Although Theorem 6.2 is a direct consequence of Theorem 6.4 that will be proved in the next subsection, we supply the proof for Theorem 6.2 below since it does not require a deep analysis of the convergence of the conditioned martingale difference to the asymptotic variance.

Theorem 6.14 ([50])**.** *Assume the martingale M_N, its quadratic variation $[M]_N$ and the associated martingale difference $Z_N = M_N - M_{N-1}$ (with $M_0 = 0$) satisfy the following*

1.(The Lindeberg type conditions) For every $\varepsilon > 0$,

$$\lim_{N\to+\infty} \frac{1}{N}\sum_{j=0}^{N-1} \mathbf{E}\left[Z_{j+1}^2, |Z_{j+1}| \geq \varepsilon\sqrt{N}\right] = 0, \tag{6.22}$$

$$\lim_{K\to\infty}\limsup_{\ell\to\infty} \frac{1}{\ell K}\sum_{m=1}^{\ell}\sum_{j=(m-1)K}^{mK-1} \mathbf{E}\left[1 + Z_{j+1}^2, \left|M_j - M_{(m-1)K}\right| \geq \varepsilon\sqrt{\ell K}\right] = 0. \tag{6.23}$$

2. (Law of large numbers for the conditioned martingale difference) There exists $\sigma \geq 0$ such that

$$\lim_{K\to\infty}\limsup_{\ell\to\infty} \frac{1}{\ell}\sum_{m=1}^{\ell} \mathbf{E}\left|\frac{1}{K}\mathbf{E}\left[[M]_{mK} - [M]_{(m-1)K} \,\middle|\, \mathfrak{F}_{(m-1)K}\right] - \sigma^2\right| = 0, \tag{6.24}$$

with uniformly square integrable condition $\sup_{n\geq 1}\mathbf{E}Z_n^2 < +\infty$.

Then one has

$$\lim_{N\to\infty}\frac{\mathbf{E}[M]_N}{N} = \sigma^2,$$

and

$$\lim_{N\to\infty}\mathbf{E}e^{i\theta M_N/\sqrt{N}} = e^{-\sigma^2\theta^2/2}, \quad \forall\theta\in\mathbb{R}.$$

Before proving Theorem 6.2 by verifying the conditions of the above Theorem 6.14, we recall a useful lemma.

Lemma 6.15. *[50] If $\{\mu_N\}_{N\geq 1}\subset\mathcal{P}(H\times\mathbb{T}^n)$ converges to μ weakly, $\{F_N\}_{N\geq 1}\subset C(H\times\mathbb{T}^n)$ converges to* 0 *uniformly on compact sets and there is $\eta>0$ such that* $\limsup_{N\to\infty}\langle\mu_N,|F_N|^{1+\eta}\rangle<\infty$*, then* $\lim_{N\to\infty}\langle\mu_N,F_N\rangle=0$.

Proof of Theorem 6.2. We first prove the The Lindeberg type conditions. From the Markov property of the homogenized process, the left hand side of (6.22) can be rewritten as

$$\frac{1}{N}\sum_{j=1}^{N}\mathbf{E}\left[Z_j^2;|Z_j|\geq\varepsilon\sqrt{N}\right]=\frac{1}{N}\sum_{j=1}^{N}\left\langle P_{j-1}^*\delta_{(w,h)},G_N\right\rangle,$$

where $G_N(u,g)=\mathbf{E}_{(u,g)}[M_1^2;|M_1|\geq\varepsilon\sqrt{N}]$ for $(u,g)\in H\times\mathbb{T}^n$. By the Markov inequality and Lemma 6.9, one has for $\eta\in(0,2^{-3}\eta_0]$,

$$G_N\leq\left(\mathbf{E}_{(u,g)}|M_1|^4\right)^{\frac{1}{2}}\mathbf{P}\left(|M_1|\geq\varepsilon\sqrt{N}\right)^{\frac{1}{2}}\leq\left(\mathbf{E}_{(u,g)}|M_1|^4\right)^{\frac{1}{2}}\left(\frac{\mathbf{E}_{(u,g)}|M_1|^4}{\varepsilon^4N^2}\right)^{\frac{1}{2}}\leq C\varepsilon^{-2}N^{-1}e^{2^4\eta\|w\|^2}.$$

Hence $G_N\to 0$ uniformly on any compact set. Also by Lemma 6.9, it follows that

$$\frac{1}{N}\sum_{j=1}^{N}\left\langle P_{j-1}^*\delta_{(w,h)},G_N^2\right\rangle\leq\frac{1}{N}\sum_{j=1}^{N}P_{j-1}\mathbf{E}_{(w,h)}|M_1|^4\leq Ce^{2^4\eta\|w\|^2},$$

where C is independent of N. Therefore

$$\limsup_{N\to\infty}\frac{1}{N}\sum_{j=1}^{N}\left\langle P_{j-1}^*\delta_{(w,h)},G_N^2\right\rangle<\infty.$$

The first Lindeberg condition (6.22) then follows from Proposition 5.3 and Lemma 6.15.

Again by the Markov property we can rewrite the left hand side formula in the limit of (6.23) as

$$\frac{1}{\ell K}\sum_{m=1}^{\ell}\sum_{j=(m-1)K}^{mK-1}\mathbf{E}\left[1+Z_{j+1}^2,\left|M_j-M_{(m-1)K}\right|\geq\varepsilon\sqrt{\ell K}\right]=\frac{1}{K}\sum_{j=0}^{K-1}\left\langle\frac{1}{\ell}\sum_{m=1}^{\ell}P_{(m-1)K}^*\delta_{(w,h)},F_{\ell,j}\right\rangle,$$

where $F_{\ell,j}(u,g) = \mathbf{E}_{(u,g)}\left[1 + Z_{j+1}^2, |M_j| \geq \varepsilon\sqrt{\ell K}\right]$. It follows from the Markov inequality that

$$\begin{aligned}
F_{\ell,j}(u,g) &\leq \left(\mathbf{E}_{(u,g)}\left(1 + Z_{j+1}^2\right)^2\right)^{\frac{1}{2}} \mathbf{P}\left(|M_j|(u,g) \geq \varepsilon\sqrt{\ell K}\right)^{\frac{1}{2}} \\
&\leq \left(\mathbf{E}_{(u,g)}\left(1 + Z_{j+1}^2\right)^2\right)^{\frac{1}{2}} \frac{\left(\mathbf{E}_{(u,g)}|M_j|^4\right)^{\frac{1}{2}}}{\left(\varepsilon\sqrt{\ell K}\right)^2} \\
&\leq C(\varepsilon)\frac{1 + \mathbf{E}_{(u,g)}|M_{j+1}|^4 + \mathbf{E}_{(u,g)}|M_j|^4}{\ell K}.
\end{aligned}$$

In view of Lemma 6.9, we know that for any $R > 0$, and $0 \leq j \leq K$, there is a constant C independent of ℓ such that

$$\sup_{(u,g)\in B_R(0)\times\mathbb{T}^n} F_{\ell,j}(u,g) \leq \frac{C}{(\ell K)^{1/2}}.$$

Hence $F_{\ell,j} \to 0$ as $\ell \to \infty$ uniformly on bounded (in particular compact) sets. Again by Lemma 6.9,

$$\left\langle \frac{1}{\ell}\sum_{m=1}^{\ell} P_{(m-1)K}^*\delta_{(w,h)}, F_{\ell,j}^2 \right\rangle \leq C\left(1 + \frac{1}{\ell}\sum_{m=1}^{\ell} P_{(m-1)K}\mathbf{E}_{(w,h)}\left(|M_j|^4 + |M_{j+1}|^4\right)\right) \leq C(K,w).$$

Therefore

$$\limsup_{\ell\to\infty} \left\langle \frac{1}{\ell}\sum_{m=1}^{\ell} P_{(m-1)K}^*\delta_{(w,h)}, F_{\ell,j}^2 \right\rangle < \infty.$$

It then follows from Proposition 5.3 and Lemma 6.15 that

$$\lim_{\ell\to\infty} \left\langle \frac{1}{\ell}\sum_{m=1}^{\ell} P_{(m-1)K}^*\delta_{(w,h)}, F_{\ell,j} \right\rangle = 0, \quad j = 0, \cdots, K-1,$$

which completes the proof of (6.23).

Now we proceed to show the law of large numbers for the conditioned martingale difference as in (6.24). The finiteness of $\sup_{n\geq 1} \mathbf{E}_{(w,h)} Z_n^2$ follows from Lemma 6.9. By the Markov property, we have

$$\frac{1}{\ell}\sum_{m=1}^{\ell} \mathbf{E}\left|\frac{1}{K}\mathbf{E}\left[[M]_{mK} - [M]_{(m-1)K}|\mathcal{F}_{(m-1)K}\right] - \sigma^2\right| = \frac{1}{\ell}\sum_{m=1}^{\ell} \left\langle P_{(m-1)K}^*\delta_{(w,h)}, |H_K|\right\rangle, \tag{6.25}$$

where

$$H_K(u,g) = \mathbf{E}_{(u,g)}\left[\frac{1}{K}[M]_K - \sigma^2\right] = \mathbf{E}_{(u,g)}\left[\frac{1}{K}M_K^2 - \sigma^2\right] = \frac{1}{K}\sum_{j=0}^{K-1} P_j Y_0(u,g),$$

with $Y_0(u,g) = \mathbf{E}_{(u,g)}[M]_1 - \sigma^2 = \mathbf{E}_{(u,g)}M_1^2 - \sigma^2$. In view of the decomposition (6.21) for $\mathbf{E}_{(u,g)}M_1^2$ and Proposition 6.5, together with Proposition 6.7, we see that for $\eta \in (0, 2^{-4}\eta_0]$, $Y_0 \in C_{2^4\eta,H}^{\overline{\gamma}_0}(H \times \mathbb{T}^n)$ where $\overline{\gamma}_0$ is as in Proposition 6.7. Note that we do not require the Hölder continuity of Ψ here.

Also by Lemma 6.9, we have

$$\frac{1}{\ell}\sum_{m=1}^{\ell}\left\langle P^*_{(m-1)K}\delta_{(w,h)}, |H_K|^2\right\rangle \le C(\sigma,K)\left(1+\frac{1}{\ell}\sum_{m=1}^{\ell}P_{(m-1)K}\mathbf{E}_{(w,h)}M_K^4\right)$$
$$\le C(\sigma,K,w),$$

which is independent of ℓ. Hence

$$\limsup_{\ell\to\infty}\frac{1}{\ell}\sum_{m=1}^{\ell}\left\langle P^*_{(m-1)K}\delta_{(w,h)}, |H_K|^2\right\rangle < \infty.$$

Now since $\frac{1}{\ell}\sum_{m=1}^{\ell}P^*_{(m-1)K}\delta_{(w,h)}$ converges weakly to $\Gamma_g(du)\lambda(dg)$ by Proposition 5.3, it follows that

$$\lim_{\ell\to\infty}\frac{1}{\ell}\sum_{m=1}^{\ell}\left\langle P^*_{(m-1)K}\delta_{(w,h)}, |H_K|\right\rangle = \int_{H\times\mathbb{T}^n}|H_K(u,g)|\Gamma_g(du)\lambda(dg).$$

Since the homogenized process is uniquely ergodic with invariant measure $\Gamma_g(du)\lambda(dg)$, it follows from the Birkhoff ergodic theorem for stationary ergodic process, that for an appropriate initial condition X_0 with law $\Gamma_g(du)\lambda(dg)$,

$$\lim_{K\to\infty}\int_{H\times\mathbb{T}^n}|H_K(u,g)|\Gamma_g(du)\lambda(dg) = \lim_{K\to\infty}\mathbf{E}\left|\frac{1}{K}\sum_{j=0}^{K-1}P_jY_0(X_0)\right| = \lim_{K\to\infty}\mathbf{E}\left|\frac{1}{K}\sum_{j=0}^{K-1}Y_0(X_j(X_0))\right|$$
$$= \left|\int_{H\times\mathbb{T}^n}Y_0(u,g)\Gamma_g(du)\lambda(dg)\right| = 0,$$

provided that $\int_{H\times\mathbb{T}^n}Y_0(u,g)\Gamma_g(du)\lambda(dg)=0$, i.e., $\sigma^2=\int_{H\times\mathbb{T}^n}\mathbf{E}_{(u,g)}M_1^2\Gamma_g(du)\lambda(dg)$, which follows from (6.20). This completes the proof of (6.24). Hence the central limit theorem (6.7) follows from Theorem 6.14, the martingale approximation (6.14) and Theorem (6.10). □

6.3 The Rate of Convergence in the Limit Theorems

The aim of this section is to show the desired rate of convergence in the limit theorems as in Theorem 6.3 and Theorem 6.4. We begin with a result that gives a convergence rate for the moments of the time average of the observations centered by the quasi-periodic invariant measure. It will be useful when estimating the rate of convergence of the conditioned martingale difference to the variance.

Proposition 6.16. *For any integer $p\ge 1$, $\eta\in(0,\frac{\eta_0}{4p}]$, $\gamma\in(0,1]$ and $\phi\in C^\gamma_{\eta,H}(H\times\mathbb{T}^n)$, we have*

$$\mathbf{E}_{(w,h)}\left|\frac{1}{N}\sum_{k=1}^{N}\Big(\phi(X_{k-1})-\left\langle\Gamma_{\beta_{k-1}h},\phi(\cdot,\beta_{k-1}h)\right\rangle\Big)\right|^{2p} \le C_p e^{4p\eta\|w\|^2}\|\phi\|^p_{\gamma,\eta,H}N^{-p}, \tag{6.26}$$

for all $N\ge 1,(w,h)\in H\times\mathbb{T}^n$. The same result also holds if we replace the summation by *integration:*

$$\mathbf{E}_{(w,h)}\left|\frac{1}{T}\int_0^T \Big(\phi(X_t)-\big\langle\Gamma_{\beta_t h},\phi(\cdot,\beta_t h)\big\rangle\Big)dt\right|^{2p}\le C_p e^{4p\eta\|w\|^2}\|\phi\|^p_{\gamma,\eta,H}T^{-p}, \tag{6.27}$$

for any $T\ge 1$ and $(w,h)\in H\times\mathbb{T}^n$.

To show this proposition, we give a lemma first.

Lemma 6.17. *For any real numbers $\{x_i\}_{i\ge 1}$ and any integer $m\ge 1, p\ge 1$, let $S_m=\sum_{i=1}^m x_i$. Then one has*

$$|S_m|^{2p}=\left|\sum_{i=1}^m x_i\right|^{2p}=\sum_{i=1}^m\sum_{j=i}^m f_{2p-2,j}(x_1,x_2,\cdots,x_{i-1},x_i)x_ix_j,$$

where

$$f_{2p-2,j}(x_1,x_2,\cdots,x_{i-1},x_i):=\begin{cases}\displaystyle\sum_{k=0}^{2p-2}(k+1)S_{i-1}^kS_i^{2p-2-k}, & \text{when } j=i,\\ \displaystyle 2p\sum_{k=0}^{2p-2}S_{i-1}^kS_i^{2p-2-k}, & \text{when } j>i.\end{cases}$$

Proof. By the multinomial formula, one has for $j=i$,

$$\begin{aligned}f_{2p-2,i}(x_1,x_2,\cdots,x_{i-1},x_i)&=\sum_{k_1+k_2+\cdots+k_i=2p-2}\frac{(2p)!}{k_1!k_2!\cdots k_{i-1}!(k_i+2)!}x_1^{k_1}x_2^{k_2}\cdots x_i^{k_i}\\&=\sum_{k_i=0}^{2p-2}C_{2p}^{k_i+2}S_{i-1}^{2p-2-k_i}x_i^{k_i}\\&=x_i^{-2}\left(S_i^{2p}-S_{i-1}^{2p}-2pS_{i-1}^{2p-1}x_i\right)\\&=x_i^{-2}\left(S_i\left(S_i^{2p-1}-S_{i-1}^{2p-1}\right)-(2p-1)S_{i-1}^{2p-1}x_i\right)\\&=\sum_{k=0}^{2p-2}(k+1)S_{i-1}^kS_i^{2p-2-k}.\end{aligned}$$

And similarly for $j>i$,

$$\begin{aligned}f_{2p-2,j}(x_1,x_2,\cdots,x_{i-1},x_i)&=\sum_{k_1+k_2+\cdots+k_i=2p-2}\frac{(2p)!}{k_1!k_2!\cdots k_{i-1}!(k_i+1)!}x_1^{k_1}x_2^{k_2}\cdots x_i^{k_i}\\&=2p\sum_{k_i=0}^{2p-2}C_{2p-1}^{k_i+1}S_{i-1}^{2p-2-k_i}x_i^{k_i}\\&=2px_i^{-1}\left(S_i^{2p-1}-S_{i-1}^{2p-1}\right)\\&=2p\sum_{k=0}^{2p-2}S_{i-1}^kS_i^{2p-2-k}.\end{aligned}$$

□

Proof of Proposition 6.16. Recall that $\widetilde{\phi}(w,h) = \phi(w,h) - \langle \Gamma_h, \phi(\cdot,h)\rangle$, then the summands in inequality (6.26) is $\xi_k := \widetilde{\phi}(X_{k-1})$. Let

$$S_m = \sum_{i=1}^{m} \xi_i, \quad s_N = \sup_{1\le m\le N} \mathbf{E}_{(w,h)}|S_m|^{2p}.$$

Let $g(\xi_i,\xi_j) = \xi_i \mathbf{E}_{(w,h)}\left[\xi_j|\mathcal{F}_{i-1}\right]$ and $g_p(w,h) = \left(\mathbf{E}_{(w,h)}\left|g(\xi_i,\xi_j)\right|^p\right)^{1/p}$. It follows from Lemma 6.17 and the Hölder inequality that

$$\begin{aligned}
&\mathbf{E}_{(w,h)}|S_m|^{2p} = \mathbf{E}_{(w,h)} \sum_{i=1}^{m}\sum_{j=i}^{m} f_{2p-2,j}(\xi_1,\xi_2,\cdots,\xi_{i-1},\xi_i)\xi_i\xi_j \\
&= \sum_{i=1}^{m}\sum_{j=i}^{m} \mathbf{E}_{(w,h)}\Big[f_{2p-2,j}(\xi_1,\xi_2,\cdots,\xi_{i-1},\xi_i)\xi_i \mathbf{E}_{(w,h)}\left[\xi_j|\mathcal{F}_{i-1}\right]\Big] \\
&\le \sum_{i=1}^{m}\sum_{j=i}^{m} \Big(\mathbf{E}_{(w,h)}\left|f_{2p-2,j}(\xi_1,\xi_2,\cdots,\xi_{i-1},\xi_i)\right|^{\frac{p}{p-1}}\Big)^{\frac{p-1}{p}} g_p(w,h) \\
&\le \sum_{i=1}^{m}\sum_{j=i}^{m} 2p\left((2p-1)^{\frac{1}{p-1}}\sum_{k=0}^{2p-2}\mathbf{E}_{(w,h)}\Big(|S_{i-1}|^k|S_i|^{2p-2-k}\Big)^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} g_p(w,h) \\
&\le \sum_{i=1}^{m}\sum_{j=i}^{m} 2p(2p-1)^{\frac{1}{p}}\left(\sum_{k=0}^{2p-2}\mathbf{E}_{(w,h)}\left(\frac{k}{2p-2}|S_{i-1}|^{2p-2}+\frac{2p-2-k}{2p-2}|S_i|^{2p-2}\right)^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} g_p(w,h) \\
&\le \sum_{i=1}^{m}\sum_{j=i}^{m} 2p(2p-1)\Big(\mathbf{E}_{(w,h)}\left[|S_{i-1}|^{2p}+|S_i|^{2p}\right]\Big)^{\frac{p-1}{p}} g_p(w,h).
\end{aligned}$$

Taking the supremum for $1\le m\le N$, one has

$$s_N \le 4p(2p-1)s_N^{\frac{p-1}{p}}\sum_{i=1}^{N}\sum_{j=i}^{N}\left(\mathbf{E}_{(w,h)}\left|g(\xi_i,\xi_j)\right|^p\right)^{1/p}.$$

Hence by letting $C_p = (4p(2p-1))^p$,

$$s_N \le C_p\left(\sum_{i=1}^{N}\sum_{j=i}^{N}\left(\mathbf{E}_{(w,h)}\left|g(\xi_i,\xi_j)\right|^p\right)^{1/p}\right)^p.$$

Note that by Theorem 6.6 with $\kappa = 2$,

$$\begin{aligned}
\mathbf{E}_{(w,h)}\left|g(\xi_i,\xi_j)\right|^p &= \mathbf{E}_{(w,h)}\left|\widetilde{\phi}(X_{i-1})\mathbf{E}_{(w,h)}\left[\widetilde{\phi}(X_{j-1})|\mathcal{F}_{i-1}\right]\right|^p \\
&\le \mathbf{E}_{(w,h)}\left|\widetilde{\phi}(X_{i-1})P_{j-i}\widetilde{\phi}(X_{i-1})\right|^p \\
&\le C^p\|\phi\|^p_{\gamma,\eta,H}e^{-p\Lambda(j-i)}\mathbf{E}_{(w,h)}e^{4p\eta\|w_{0,i-1,h}(w)\|^2} \\
&\le C^p\|\phi\|^p_{\gamma,\eta,H}e^{-p\Lambda(j-i)}e^{4p\eta\|w\|^2},
\end{aligned}$$

for $\eta \in (0, \frac{\eta_0}{4p}]$ by estimate (A.1). Therefore

$$s_N \le C C_p \|\phi\|^p_{\gamma,\eta,H} e^{4p\eta\|w\|^2} N^p,$$

where $C > 0$ does not depend on p. Dividing both sides of the above inequality by N^{2p} completes the proof of the first estimate (6.26) in Proposition 6.16.

The second inequality follows from the same argument (see also [61] for the case in the time homogeneous setting). Let $\xi(t) = \phi(X_t) - \langle \Gamma_{\beta_t h}, \phi(\cdot, \beta_t h)\rangle$, $I_r = \int_0^r \xi(t)dt$ and $\mathcal{I}_T = \sup_{0\le r\le T} \mathbf{E}_{(w,h)} |I_r|^{2p}$. We first note that

$$\begin{aligned} I_r^{2p} &= \int_{[0,r]^{2p}} \xi(t_1)\xi(t_2)\cdots\xi(t_{2p})dt_1dt_2\cdots dt_{2p} \\ &= (2p)! \int_{0\le t_1\le t_2\le\cdots\le t_{2p}\le r} \xi(t_1)\xi(t_2)\cdots\xi(t_{2p})dt_1dt_2\cdots dt_{2p}. \end{aligned}$$

For $r_1 \le r_2$, denote $\varphi(r_1, r_2) = \xi(r_1)\mathbf{E}_{(w,h)}[\xi(r_2)|\mathcal{F}_{r_1}]$. Then

$$\begin{aligned} &\mathbf{E}_{(w,h)} |I_r|^{2p} = (2p)!\mathbf{E}_{(w,h)} \int_{0\le t_1\le t_2\le\cdots\le t_{2p}\le r} \xi(t_1)\xi(t_2)\cdots\xi(t_{2p-2})\varphi(t_{2p-1}, t_{2p})dt_1dt_2\cdots dt_{2p} \\ &= (2p)!\mathbf{E}_{(w,h)} \left(\int_0^r \int_0^{t_{2p}} \varphi(t_{2p-1}, t_{2p}) \left(\int_{0\le t_1\le\cdots\le t_{2p-1}} \xi(t_1)\cdots\xi(t_{2p-2})dt_1\cdots dt_{2p-1} \right) dt_{2p-1}dt_{2p} \right) \\ &\le 2p(2p-1) \int_0^r \int_0^{t_{2p}} \left(\mathbf{E}_{(w,h)} |\varphi(t_{2p-1}, t_{2p})|^p\right)^{\frac{1}{p}} \left(\mathbf{E}_{(w,h)} \left|I_{t_{2p-1}}\right|^{2p}\right)^{\frac{p-1}{p}} dt_{2p-1}dt_{2p}. \end{aligned}$$

Taking the supremum w.r.t r over $[0, T]$, we have

$$\mathcal{I}_T \le 2p(2p-1) (\mathcal{I}_T)^{\frac{p-1}{p}} \int_0^T \int_0^{t_2} \left(\mathbf{E}_{(w,h)} |\varphi(t_1, t_2)|^p\right)^{\frac{1}{p}} dt_1dt_2.$$

Like in the proof of (6.26), one has for $\eta \in (0, \frac{\eta_0}{4p})$,

$$\left(\mathbf{E}_{(w,h)} |\varphi(t_1, t_2)|^p\right)^{\frac{1}{p}} \le C\|\phi\|_{\gamma,\eta,H} e^{-\Lambda(t_2-t_1)} e^{4\eta\|w\|^2}.$$

Therefore

$$\mathcal{I}_T \le C_p \|\phi\|^p_{\gamma,\eta,H} T^p e^{4p\eta\|w\|^2}$$

and the inequality (6.27) follows by dividing both sides with T^{2p}. □

The rate of convergence in SLLN. The convergence rate of SLLN is a consequence of (6.27), the error estimate in Lemma 6.10 and the Borel-Cantelli lemma. We now give the details.

Proof of Theorem 6.3. For any $\varepsilon > 0$, let $E_N = \left\{\omega \in \Omega : \left|\frac{1}{N} M_N\right| > N^{-(1/2-\varepsilon)}\right\}$. From Proposition 6.7, we know that for $\eta \in (0, \eta_0/2]$, $\chi \in C^{\gamma_0}_{2\eta,H}(H \times \mathbb{T}^n)$. Thus $\chi^{2^p} \in C^{\gamma_0}_{2^{p+1}\eta,H}(H \times \mathbb{T}^n)$. Hence from estimate (6.27) and (A.1) and Markov's inequality, we have for $\eta \in (0, 2^{-(p+1)}\eta_0]$

$$
\begin{aligned}
\mathbf{P}(E_N) &\le N^{2^p(-1/2-\varepsilon)}\mathbf{E}_{(w,h)}M_N^{2^p} \\
&\le C_p N^{2^p(-1/2-\varepsilon)}\mathbf{E}_{(w,h)}\left(\chi(X_N)^{2^p} + \chi(w,h)^{2^p} + \left(\int_0^N \widetilde{\phi}(X_t)dt\right)^{2^p}\right) \\
&\le C_p N^{2^p(-1/2-\varepsilon)}\left(\|\chi\|_{\gamma_0,2^{p+1}\eta,H}\mathbf{E}e^{2^{p+1}\eta\|w_{0,N,h}(w)\|^2} + \|\chi\|_{\gamma_0,2^{p+1}\eta,H}e^{2^{p+1}\eta\|w\|^2} + e^{2^{p+1}\eta\|w\|^2}\|\phi\|^{2^{p-1}}_{\gamma,\eta,H}N^{2^{p-1}}\right) \\
&\le C_p(\|\phi\|_{\gamma,\eta,H})N^{-2^p\varepsilon}e^{2^{p+1}\eta\|w\|^2}.
\end{aligned}
$$

For any $\varepsilon > 0$, and every integer p such that $2^p\varepsilon > 1$,

$$\sum_{N=1}^{\infty}\mathbf{P}(E_N) < \infty.$$

By the Borel-Cantelli lemma, there is an almost surely finite random time $N_1(\omega)$, such that for all $N > N_1(\omega)$,

$$\left|\frac{1}{N}M_N\right| \le N^{-(1/2-\varepsilon)}.$$

Note that for $\ell > 0$,

$$
\begin{aligned}
\mathbf{E}N_1^\ell = \sum_{k=1}^{\infty}\mathbf{P}(N_1 = k)k^\ell &\le \sum_{k=1}^{\infty}\mathbf{P}(E_k)k^\ell \\
&\le \sum_{k=1}^{\infty}C_p(\|\phi\|_{\gamma,\eta,H})k^{\ell-2^p\varepsilon}e^{2^{p+1}\eta\|w\|^2} \\
&\le C_p(\|\phi\|_{\gamma,\eta,H},\ell,\varepsilon)e^{2^{p+1}\eta\|w\|^2}
\end{aligned}
$$

as long as $\ell < 2^p\varepsilon - 1$. In a similar fashion we can estimate the moments of the random time $N_0(\omega)$ in (6.17). Let $\ell > 0$, then for $\eta \in (0, 2^{-p-1}\eta_0]$,

$$
\begin{aligned}
\mathbf{E}N_0^\ell &= \sum_{k=1}^{\infty}\mathbf{P}(N_0 = k)k^\ell \\
&\le \sum_{k=1}^{\infty}\mathbf{P}\left(\sup_{k\le t\le k+1}R_{k,t} > k^{1/4}\right)k^\ell \le \sum_{k=1}^{\infty}\mathbf{E}\left(\sup_{k\le t\le k+1}R_{k,t}^{2^p}\right)k^{\ell-2^{p-2}} \\
&\le C_p(\|\phi\|_{\gamma,\eta,H})\sum_{k=1}^{\infty}\mathbf{E}\sup_{k\le t\le k+1}\exp(2^{p+1}\eta\|w_{0,t,h}\|^2)k^{\ell-2^{p-2}} \\
&\le C_p(\|\phi\|_{\gamma,\eta,H})\sum_{k=1}^{\infty}e^{2^{p+1}\eta\|w\|^2}k^{\ell-2^{p-2}} = C_p(\|\phi\|_{\gamma,\eta,H},\ell)e^{2^{p+1}\eta\|w\|^2},
\end{aligned}
$$

provided that $\ell < 2^{p-2} - 1$. The conclusion of Theorem 6.3 then follows from the above estimates, the martingale approximation (6.14) and Lemma 6.10. □

The rate of convergence in CLT. We recall a Berry-Esseen type estimate for martingales from [39], which will be used in the proof of the rate of convergence in the central limit theorem.

Theorem 6.18 (Theorem 3.10 of [39])**.** *Let* $M_N = \sum_{j=1}^{N} Z_j$ *be a zero mean martingale and* $\sigma_k^2 = \sum_{j=1}^{k} \mathbf{E} Z_j^2$. *If* $q > \frac{1}{2}$, *and*

$$\max_{j \le N} \frac{1}{\sigma_N^{4q}} \mathbf{E}|Z_j|^{4q} \le \frac{M}{N^{2q}}, \tag{6.28}$$

for a constant $M > 0$. *Then there exists a constant* C *depending only on* M *and* q *such that whenever*

$$N^{-q} + \mathbf{E}\left|\frac{1}{\sigma_N^2} \sum_{j=1}^{N} \mathbf{E}\left[Z_j^2 | \mathcal{F}_{j-1}\right] - 1\right|^{2q} \le 1, \tag{6.29}$$

one has

$$\sup_{z \in \mathbb{R}} \left| \mathbf{P}\left(\frac{M_N}{\sigma_N} \le z\right) - \Phi(z) \right| \le C \left(N^{-q} + \mathbf{E}\left|\frac{1}{\sigma_N^2} \sum_{j=1}^{N} \mathbf{E}\left[Z_j^2 | \mathcal{F}_{j-1}\right] - 1\right|^{2q} \right)^{1/(4q+1)},$$

where $\Phi(z)$ *is the distribution function of the standard normal distribution.*

We are now in a position to give an estimate of the convergence rate in the central limit theorem.

Proof of Theorem 6.4. Recall from (6.20) that the asymptotic variance

$$\sigma^2 = \int_{H \times \mathbb{T}^n} Y(w,h)\Gamma_h(dw)\lambda(dh) = \int_{\mathbb{T}^n} \langle \Gamma_h, Y(\cdot,h)\rangle \lambda(dh).$$

Since $Y \in C^{\overline{\gamma}_0}_{2^5\eta,H}(H \times \mathbb{T}^n)$ by Proposition 6.13, thus Proposition 6.16 implies that for $\eta \in (0, 2^{-7}p^{-1}\eta_0]$, and $N \ge 1$,

$$\mathbf{E}_{(w,h)} \left| \frac{1}{N} \sum_{k=1}^{N} \Big(Y(X_{k-1}) - \big\langle \Gamma_{\beta_{k-1}h}, Y(\cdot, \beta_{k-1}h)\big\rangle \Big) \right|^{2p} \le C_p e^{2^7 p\eta \|w\|^2} \|Y\|^p_{\overline{\gamma}_0, 2^5\eta, H} N^{-p}. \tag{6.30}$$

In addition, since $F = \langle \Gamma_h, Y(\cdot,h)\rangle \in C^{\overline{\gamma}_0}(\mathbb{T}^n, \mathbb{R})$ by Proposition 6.13, it follows from Theorem 3 in [43] that for $N \ge 1$,

$$\left| \frac{1}{N} \sum_{k=1}^{N} \big\langle \Gamma_{\beta_{k-1}h}, Y(\cdot, \beta_{k-1}h)\big\rangle - \sigma^2 \right| \le C\|F\|_{\overline{\gamma}_0} N^{-\frac{\overline{\gamma}_0}{A+n}}, \tag{6.31}$$

where A is the constant from the Diophantine condition (2.2) and n is the dimension of the torus. Therefore, by the Markov property, inequality (6.30) with $p = 1$, and inequality (6.31), it follows that for $\eta \in (0, 2^{-7}\eta_0]$, and $N \geq 1$,

$$\left|\frac{1}{N}\sigma_N^2 - \sigma^2\right| = \left|\frac{1}{N}\sum_{k=1}^{N}\mathbf{E}_{(w,h)}Z_k^2 - \sigma^2\right|$$

$$\leq \left|\frac{1}{N}\sum_{k=1}^{N}\Big(\mathbf{E}_{(w,h)}\left[\mathbf{E}_{(w,h)}\left[Z_k^2|\mathcal{F}_{k-1}\right]\right] - \left\langle\Gamma_{\beta_{k-1}h}, Y(\cdot,\beta_{k-1}h)\right\rangle\Big)\right|$$

$$+ \left|\frac{1}{N}\sum_{k=1}^{N}\left\langle\Gamma_{\beta_{k-1}h}, Y(\cdot,\beta_{k-1}h)\right\rangle - \sigma^2\right|$$

$$\leq \mathbf{E}_{(w,h)}\left|\frac{1}{N}\sum_{k=1}^{N}\Big(Y(X_{k-1}) - \left\langle\Gamma_{\beta_{k-1}h}, Y(\cdot,\beta_{k-1}h)\right\rangle\Big)\right| + \left|\frac{1}{N}\sum_{k=1}^{N}\left\langle\Gamma_{\beta_{k-1}h}, Y(\cdot,\beta_{k-1}h)\right\rangle - \sigma^2\right|$$

$$\leq Ce^{2^6\eta\|w\|^2}\|Y\|_{\overline{\gamma}_0,2^5\eta,H}^{1/2}N^{-1/2} + C\|F\|_{\overline{\gamma}_0}N^{-\frac{\overline{\gamma}_0}{A+n}}. \tag{6.32}$$

This shows that $\frac{1}{N}\sigma_N^2$ converges to σ^2. To estimate the convergence rate for the central limit theorem of the approximating martingale sequence, we first deal with the case when $\sigma^2 > 0$. Set $q = 2^{p-2}$ for integer $p \geq 2$ in Theorem 6.18 and choose $N_1 > 0$ such that for all $N \geq N_1$, inequality (6.29) is satisfied and $\frac{1}{N}\sigma_N^2 \in [\sigma^2/2, 3\sigma^2/2]$. It then follows from Lemma (6.9) that the condition (6.28) holds for $p \geq 2$ and $\eta \in (0, 2^{-p-1}\eta_0]$.

Observe that

$$\mathbf{E}_{(w,h)}\left|\frac{1}{\sigma_N^2}\sum_{j=1}^{N}\mathbf{E}_{(w,h)}\left[Z_j^2|\mathcal{F}_{j-1}\right] - 1\right|^{2^{p-1}}$$

$$\leq C_p\left(\mathbf{E}_{(w,h)}\left|\frac{1}{\sigma_N^2}\sum_{k=1}^{N}\Big(Y(X_{k-1}) - F(\beta_{k-1}h)\Big)\right|^{2^{p-1}} + \left|\frac{1}{\sigma_N^2}\sum_{k=1}^{N}F(\beta_{k-1}h) - 1\right|^{2^{p-1}}\right)$$

$$\leq C_p\left(e^{2^{p+5}\eta\|w\|^2}\|Y\|_{\overline{\gamma}_0,2^5\eta,H}^{2^{p-2}}N^{-2^{p-2}} + \left|\frac{1}{N}\sum_{k=1}^{N}F(\beta_{k-1}h) - \sigma^2\right|^{2^{p-1}} + \left|\frac{1}{N}\sigma_N^2 - \sigma^2\right|^{2^{p-1}}\right)$$

$$\leq C_p\left(e^{2^{p+5}\eta\|w\|^2}\|Y\|_{\overline{\gamma}_0,2^5\eta,H}^{2^{p-2}}N^{-2^{p-2}} + \|F\|_{\overline{\gamma}_0}^{2^{p-1}}N^{-\frac{2^{p-1}\overline{\gamma}_0}{A+n}}\right).$$

Therefore by Theorem 6.18 we have

$$\sup_{z\in\mathbb{R}}\left|\mathbf{P}\left(\frac{M_N}{\sigma_N} \leq z\right) - \Phi(z)\right| \leq C\left(N^{-2^{p-2}} + \mathbf{E}\left|\frac{1}{\sigma_N^2}\sum_{j=1}^{N}\mathbf{E}\left[Z_j^2|\mathcal{F}_{j-1}\right] - 1\right|^{2^{p-1}}\right)^{1/(2^p+1)}$$

$$\leq Ce^{2^5\eta\|w\|^2}\left(N^{-\frac{2^{p-2}}{2^p+1}} + N^{-\frac{2^{p-1}\overline{\gamma}_0}{(2^p+1)(A+n)}}\right).$$

As a result,

$$\begin{aligned}\sup_{z\in\mathbb{R}}\left|\mathbf{P}\left(\frac{M_N}{\sqrt{N}}\le z\right)-\Phi_\sigma(z)\right| &\le \sup_{z\in\mathbb{R}}\left|\mathbf{P}\left(\frac{M_N}{\sqrt{N}}\le z\right)-\Phi\left(\frac{\sqrt{N}}{\sigma_N}z\right)\right|+\sup_{z\in\mathbb{R}}\left|\Phi\left(\frac{\sqrt{N}}{\sigma_N}z\right)-\Phi_\sigma(z)\right|\\ &\le \sup_{z\in\mathbb{R}}\left|\mathbf{P}\left(\frac{M_N}{\sigma_N}\le z\right)-\Phi\left(z\right)\right|+C\left|\frac{\sigma_N}{\sqrt{N}}-\sigma\right|\\ &\le Ce^{2^5\eta\|w\|^2}\left(N^{-\frac{2^{p-2}}{2^p+1}}+N^{-\frac{2^{p-1}\overline{\gamma}_0}{(2^p+1)(A+n)}}\right).\end{aligned}$$

When $\sigma=0$, we note that

$$\begin{aligned}&(|z|\wedge 1)\left|\mathbf{P}\left(\frac{M_N}{\sqrt{N}}\le z\right)-\Phi_0(z)\right|\le(|z|\wedge 1)\left|\mathbf{P}\left(\left|\frac{M_N}{\sqrt{N}}\right|\ge |z|\right)\right|\\ &\le(|z|\wedge 1)|z|^{-1}N^{-1/2}\mathbf{E}_{(w,h)}|M_N|\le N^{-1/2}\left(\mathbf{E}_{(w,h)}|M_N|^2\right)^{1/2}=\left(\frac{1}{N}\sum_{k=1}^N Z_k^2\right)^{1/2}\\ &\le Ce^{2^5\eta\|w\|^2}\left(N^{-1/4}+N^{-\frac{\overline{\gamma}_0}{2(A+n)}}\right)\end{aligned}$$

by estimate (6.32).

To pass the estimates to continuous times, we apply the following lemma from [61].

Lemma 6.19. *Let R_1, R_2 be real random variables. Then for any $\sigma\ge 0$ and $\varepsilon>0$ we have*

$$\sup_{z\in\mathbb{R}}|\Delta_\sigma(R_1,z)|\le\sup_{z\in\mathbb{R}}|\Delta_\sigma(R_2,z)|+\mathbf{P}(|R_1-R_2|>\varepsilon)+c_\sigma\varepsilon,$$

where c_σ is a constant depending only on σ and $\Delta_\sigma(R,z)$ for random variable R is defined as

$$\Delta_\sigma(R,z):=\begin{cases}\mathbf{P}(R\le z)-\Phi_\sigma(z)\ , & \sigma>0,\\ (|z|\wedge 1)\left(\mathbf{P}(R\le z)-\Phi_0(z)\right)\ , & \sigma=0.\end{cases}$$

Recall that N is the integer part of T in the martingale approximation (6.14). It follows from the approximation that

$$\left|\frac{1}{\sqrt{T}}\int_0^T\widetilde{\phi}(X_t)dt-\frac{M_N}{\sqrt{N}}\right|\le\frac{1}{\sqrt{NT}}\left|\int_0^T\widetilde{\phi}(X_t)dt\right|+\frac{R_{N,T}}{\sqrt{N}}.$$

The expectation of the remainder term $R_{N,T}$ obtained in the proof of Lemma 6.10 together with the Markov inequality yields

$$\mathbf{P}\left(\frac{R_{N,T}}{\sqrt{N}}>\varepsilon/2\right)\le Ce^{2\eta\|w\|^2}N^{-4}\varepsilon^{-8}.$$

It follows from (6.27) with $p=1$ that

$$\mathbf{E}\left|\int_0^T\widetilde{\phi}(X_t)dt\right|\le\left(\mathbf{E}\left(\int_0^T\widetilde{\phi}(X_t)dt\right)^2\right)^{\frac{1}{2}}\le T^{\frac{1}{2}}Ce^{2\eta\|w\|^2}.$$

Now applying Lemma 6.19 with $R_1 = \frac{1}{\sqrt{T}} \int_0^T \widetilde{\phi}(X_t)dt$ and $R_2 = \frac{M_N}{\sqrt{N}}$, and the Markov inequality we have

$$\begin{aligned}
\sup_{z\in\mathbb{R}} |\Delta_\sigma(R_1, z)| &\le \sup_{z\in\mathbb{R}} |\Delta_\sigma(R_2, z)| + \mathbf{P}(|R_1 - R_2| > \varepsilon) + c_\sigma\varepsilon \\
&\le \sup_{z\in\mathbb{R}} |\Delta_\sigma(R_2, z)| + 2\varepsilon^{-1}N^{-\frac{1}{2}}\mathbf{E}\left|T^{-\frac{1}{2}}\int_0^T \widetilde{\phi}(X_t)dt\right| + Ce^{2\eta\|w\|^2}N^{-4}\varepsilon^{-8} + c_\sigma\varepsilon \\
&\le \sup_{z\in\mathbb{R}} |\Delta_\sigma(R_2, z)| + Ce^{2\eta\|w\|^2}N^{-\frac{1}{4}}
\end{aligned}$$

by taking $\varepsilon = N^{-\frac{1}{4}}$. □

Chapter 7. Large Viscosity Implies Trivial Dynamics

In this chapter, we will show the existence of a unique stable quasi-periodic solution of (2.5) when the viscosity is large as in Theorem 3.5. Recall that $G = \sqrt{\|f\|_\infty^2/\nu^4 + \mathcal{B}_0/\nu^3}$, and c_0 is the constant from (3.9). Let $\delta_0 = \nu - c_0^2\nu^{-2}\left(\|f\|_\infty^2\nu^{-1} + \mathcal{B}_0\right)$. Then $\delta_0 > 0$ is equivalent to $G < 1/c_0$. Fix a lattice $\mathbb{Z}_\varkappa = \{\varkappa m : m \in \mathbb{Z}\}$ where $\varkappa > 0$ is a real number. For each $h \in \mathbb{T}^n$, denote by $(2.5)_h$ the Navier-Stokes equation (2.5) with $f(t,x) = \Psi(\beta_t 0, x)$ replaced by $\Psi(\beta_t h, x)$. Let $n_1 \in \mathbb{Z}_\varkappa$ and

$$w_{n_1+N,t,h}(0) = w(n_1 + N, t, h, \omega, 0), \quad N \in \mathbb{Z}_\varkappa^-, t \ge n_1$$

be the sequence of solutions of $(2.5)_h$. In exactly the same way as in [52], one can show the following

Theorem 7.1. *Assume $\delta_0 > 0$ and let $\delta \in (0, \delta_0)$. Then for each $h \in \mathbb{T}^n$ there is a full measure subset $\Omega_{h,\varkappa}$ of Ω such that:*

1. There is a complete random trajectory $w^*(\cdot, h, \cdot) : \mathbb{R}\times\Omega \to H$, which is a strong solution of $(2.5)_h$. For each $n_1 \in \mathbb{Z}_\varkappa$ and $\omega \in \Omega_{h,\varkappa}$, $w^*(\cdot, h, \omega)$ is the limit of $\{w_{n_1+N,t,h}(0)\}_{N\in\mathbb{Z}_\varkappa^-}$ in $C([n_1,\infty), H)$ equipped with the supremum norm.

2. For any $s \in \mathbb{R}$, there exist positive random times $n^*(s, \delta, h, \varkappa)$ and $n_*(s, \delta, h, \varkappa)$ having all moments finite such that

$$\sup_{w_0\in B_r(w^*(s,h,\omega))} \|w_{s,t,h}(\omega, w_0) - w^*(t, h, \omega)\|^2 \le r^2 e^{-\delta(t-s)}, \tag{7.1}$$

$$\sup_{w_0\in B_r(w^*(\tau,h,\omega))} \|w_{\tau,s,h}(\omega, w_0) - w^*(s, h, \omega)\|^2 \le r^2 e^{-\delta(s-\tau)}, \tag{7.2}$$

for all $r > 0$, $t > s + n^*$ and $\tau < s - n^*$. Here $B_r(w)$ is the ball centered at w with radius r in H with the norm $\|\cdot\|$.

Proof. It follows from (A.12) (with $\eta = 1$, $\tau = s$ and $a = 1$) that

$$\frac{c_0^2}{\nu}\int_s^t \|w_{s,r,h}\|_1^2 dr \leq \frac{c_0^2}{\nu^2}\left(\|w_0\|^2 + M(s,t) + (t-s)\left(\frac{\|f\|_\infty^2}{\nu} + \mathcal{B}_0\right)\right),$$

where

$$M(s,t) = 2\int_s^t \langle w_{s,r,h}, GdW_r\rangle.$$

It also follows from the proof of (A.4) that

$$\|\mathfrak{e}_t\|^2 \leq \|\mathfrak{e}_s\|^2 \exp\left(-\nu(t-s) + \frac{c_0^2}{\nu}\int_s^t \|w_{s,r,h}\|_1^2 dr\right),$$

where $e_t = \Phi_{s,t,h}(w_0) - \Phi_{s,t,h}(\widetilde{w}_0)$ is the difference of the two solutions starting from different initial conditions. Hence

$$\|\mathfrak{e}_t\|^2 \leq \|\mathfrak{e}_s\|^2 e^{-(t-s)(\delta_0 - \Gamma(s,t))}, \tag{7.3}$$

where

$$\Gamma(s,t) = \frac{c_0^2}{\nu^2(t-s)}\left(\|w_0\|^2 + M(s,t)\right).$$

The strategy of the proof is to show that the average $\Gamma(s,t)$ can be small for large time, so that we obtain a contraction in (7.3) as long as $\delta_0 > 0$, i.e., when the viscosity ν is large. The proof consists of two steps.

Step 1. Fix a $\delta \in (0, \delta_0)$ and $t_1 \in \mathbb{Z}_\varkappa$. We claim that for any $\varepsilon > 0$, there exists a $\mathbb{Z}_\varkappa^+$ valued random time $n^*(\varepsilon, \delta, t_1, h, \varkappa)$, such that with probability one, for any $\tau \geq 0$ and $n_1, n_2 \in \mathbb{Z}_\varkappa$,

$$n_1, n_2 < t_1 - n^* \implies \|w_{n_1,t_1+\tau,h}(0) - |w_{n_2,t_1+\tau,h}(0)\|^2 \leq \varepsilon e^{-\delta\tau}. \tag{7.4}$$

Assume $t_1 = 0$ without loss of generality. We also assume that $n \in \mathbb{Z}_\varkappa^-$. Since $w_{n-\varkappa,n,h}(0)$ starts from 0 for each n, estimate (A.1) implies that for each $p > 0$ there is a constant $C > 0$ independent of n such that $\mathbf{E}\|w_{n-\varkappa,n,h}(0)\|^p \leq C$. Then Lemma A.1 from [53] gives a $\mathbb{Z}_\varkappa^+$ valued random time $N_0(\omega) = N_0(\varepsilon, \delta, h, \kappa, \omega)$ with all moments finite, such that for any n with $|n| > N_0(\omega)$, one has

$$\|w_{n-\varkappa,n,h}(0)\|^2 < \varepsilon\delta^2|n|. \tag{7.5}$$

Set $\delta' = \delta_0 - \delta$. Consider the solution $\{w_{n,t,h}(0)\}_{t\geq n}$ that starts from 0 at time $n \in \mathbb{Z}_\varkappa^-$. By

(7.3) it follows that

$$\|w_{n,n+\tau,h}(0) - w_{n,n+\tau,h}(w_{n-\varkappa,n,h}(0))\|^2 \leq \|w_{n-\varkappa,n,h}(0)\|^2 e^{-\tau(\delta_0 - \Gamma(n,n+\tau))},$$

where $\Gamma(n, n+\tau) = \frac{c_0^2}{\nu^2 \tau} M(n, n+\tau)$. Note that the quadratic variation process of $M(n, n+\tau)$ is

$$[M](n, n+\tau) = 4 \int_n^{n+\tau} \sum_{k=1}^{d} \langle w_{n,r,h}, g_k \rangle^2 dr \leq 4\mathcal{B}_0 \int_n^{n+\tau} \|w_{n,r,h}\|^2 dr.$$

Therefore by Doob's L^p maximal inequality and the Burkholder-Davis-Gundy inequality, we find that for $\delta_1 := \frac{\nu^2 \delta'}{2c_0^2}$,

$$\begin{aligned}\mathbf{P}\left(\sup_{n \leq t \leq n+\tau} |M(n,t)| \geq \delta_1 \tau \right) &\leq C_p \frac{\mathbf{E}|M(n, n+\tau)|^{2p}}{\delta_1^{2p} \tau^{2p}} \leq C_p \frac{\mathbf{E}[M](n, n+\tau)^p}{\delta_1^{2p} \tau^{2p}} \\ &\leq C_p \frac{\tau^{p-1} \mathbf{E} \int_n^{n+\tau} \|w_{n,r,h}\|^{2p} dr}{\delta_1^{2p} \tau^{2p}} \leq \frac{C}{\delta_1^{2p} \tau^p},\end{aligned}$$

where in the last inequality we use the fact that there is a constant C independent of n, τ such that $\mathbf{E} \sup_{n \leq r \leq n+\tau} \|w_{n,r,h}\|^{2p} \leq C$, which is derived from estimate (A.1) and the fact that $w_{n,r,h}$ has initial condition $w(n) = 0$. In particular, for $m \in \mathbb{Z}_\varkappa^+$ we have

$$\mathbf{P}\left(\sup_{n+m-\varkappa \leq t \leq n+m} |M(n,t)| \geq \delta_1 m \right) \leq \frac{C}{\delta_1^{2p} m^p}. \tag{7.6}$$

Again by Lemma A.1 from [53], there exists a $\mathbb{Z}_\varkappa^+$ valued random time

$$N_1(n, \omega) = N_1(\delta_1, n, h, \varkappa, \omega) \geq \varkappa$$

with all moments finite, such that for all $m \in \mathbb{Z}_\varkappa^+$ with $m > N_1(n, \omega)$,

$$\sup_{n+m-\varkappa \leq t \leq n+m} |M(n,t)| \leq \delta_1 m.$$

Note that for any $\tau > N_1(n, \omega)$, there exists $m \in \mathbb{Z}_\varkappa^+$ such that $n + \tau \in [n+m-\varkappa, n+m]$. Hence

$$\frac{1}{\tau} |M(n, n+\tau)| \leq \delta_1 \frac{m}{\tau} \leq \delta_1 \frac{\tau + \varkappa}{\tau} \leq 2\delta_1,$$

which in turn shows that $\Gamma(n, n+\tau) \leq \frac{2c_0^2}{\nu^2} \delta_1 = \delta'$. Therefore for $\tau > N_1(n, \omega)$,

$$\|w_{n,n+\tau,h}(0) - w_{n,n+\tau,h}(w_{n-\varkappa,n,h}(0))\|^2 \leq \|w_{n-\varkappa,n,h}(0)\|^2 e^{-\tau(\delta_0 - \delta')} = \|w_{n-\varkappa,n,h}(0)\|^2 e^{-\tau\delta}. \tag{7.7}$$

Observe that the inequality (7.6) gives a bound that does not depend on n, therefore

$$\begin{aligned}\mathbf{E}N_1(n,\omega)^q &= \sum_{m\in\mathbb{Z}^+_\varkappa} m^q\mathbf{P}(N_1(n,\omega)=m)\\ &\le \sum_{m\in\mathbb{Z}^+_\varkappa} m^q\mathbf{P}\left(\sup_{n+m-\varkappa\le t\le n+m}|M(n,t)|\ge\delta_1 m\right)\\ &\le \sum_{m\in\mathbb{Z}^+_\varkappa}\frac{C}{\delta_1^{2p}m^{p-q}}\le C,\end{aligned}$$

for every $q<p-1$ and C is independent of n. Since the estimate in the last inequality is valid for $p>1$, we have $\mathbf{E}N_1(n,\omega)^q\le C$ for every $q>0$. Again by Lemma A.1 from [53], it follows that there exsits a $\mathcal{T}\mathbb{Z}^+$ valued random time $N_2(\omega)=N_2(\delta_1,h,\varkappa,\omega)$, with all moments finite, such that for n satisfying $|n|>N_2(\omega)$, one has $N_1(n,\omega)\le|n|$.

Now let $n^*(\varepsilon,\delta,h,\varkappa)=\max\{N_0,N_2\}$, which has all moments finite as N_0 and N_2 do. Then for those $n\in\mathbb{Z}^-_\varkappa$ with $|n|>n^*$, and $\tau>N_1(n,\omega)$, one has from (7.5) and (7.7) that

$$\begin{aligned}\|w_{n,n+\tau,h}(0)-w_{n,n+\tau,h}(w_{n-\varkappa,n,h}(0))\|^2 &\le \|w_{n-\varkappa,n,h}(0)\|^2e^{-\tau(\delta_0-\delta')}\\ &=\|w_{n-\varkappa,n,h}(0)\|^2e^{-\tau\delta}\le\varepsilon\delta^2|n|e^{-\delta\tau}.\end{aligned}$$

In particular, if $\tau=|n|$, then $|n|>N_1(n,\omega)$ for $|n|>n^*$, hence by the evolution property of stochastic flow,

$$\|w_{n,0,h}(0)-w_{n-\varkappa,0,h}(0)\|^2=\|w_{n,0,h}(0)-w_{n,0,h}(w_{n-\varkappa,n,h}(0))\|^2\le\varepsilon\delta^2|n|e^{-\delta|n|}.$$

This also implies that for any $\tau\ge0$, as long as $|n|>n^*$,

$$\|w_{n,\tau,h}(0)-w_{n-\varkappa,\tau,h}(0)\|^2=\|w_{n,\tau,h}(0)-w_{n,\tau,h}(w_{n-\varkappa,n,h}(0))\|^2\le\varepsilon\delta^2|n|e^{-\delta(\tau-n)}=\varepsilon\delta^2|n|e^{-\delta(\tau+|n|)}.$$

As a result, for any $n_1,n_2\in\mathbb{Z}_\varkappa$ and $n_1,n_2<-n^*<0\le\tau$, we have

$$\begin{aligned}\|w_{n_1,\tau,h}(0)-w_{n_2,\tau,h}(0)\| &\le \sum_{n\in\mathbb{Z}_\varkappa,n<-n*}\|w_{n,\tau,h}(0)-w_{n-\varkappa,\tau,h}(0)\|\\ &\le\sum_{n\in\mathbb{Z}_\varkappa,n<-n*}\sqrt{\varepsilon|n|}\delta e^{-\frac{\delta}{2}(\tau+|n|)}\le\sqrt{\varepsilon}\delta e^{-\delta\tau/2}\int_0^\infty\sqrt{x}e^{-\delta x/2}dx.\end{aligned}$$

This completes the proof of the claim (7.4).

Step 2. Let $n_1\in\mathbb{Z}_\varkappa$. Consider the sequence of solutions $\{w_{n_1-n,t,h}(0)\}_{n\in\mathbb{Z}^+_\varkappa}$ for $t\ge n_1$. The claim (7.4) in *Step 1* states that there exists a random time $n^*(\varepsilon,\delta,n_1,h,\kappa)$ such that for every $t\ge n_1$ and every $m_1,m_2\in\mathbb{Z}_\varkappa$ satisfying $m_1,m_2>n^*$, one has

$$\|w_{n_1-m_1,t,h}(0)-w_{n_1-m_2,t,h}(0)\|^2\le\varepsilon e^{-\delta(t-n_1)}.$$

This implies that $\{w_{n_1-n,\cdot,h}(0)\}_{n\in\mathbb{Z}_\varkappa^+}$ is a Cauchy sequence in $C([n_1,\infty),H)$, which is complete with the norm $\|w\|_\infty := \sup_{t\geq n_1}\|w(t)\|$. Define $w^*(t,h,\omega)$ to be the limit for $t\geq n_1$. Since n_1 is arbitrary, we obtain a process $w^*(t,h,\omega)$ defined for $t\in\mathbb{R}$. For any fixed $T>n_1$, it is well known (see [10, 26] for example) that there exists a random variable $K(\omega,T)$ such that $\limsup_n \|w_{n_1-n,t,h}(0)\|_1 < K(\omega,T)$ almost surely for all $t\in[n_1,T]$. Therefore $w^*\in C([n_1,\infty],H_1)$, and $\{w_{n_1-n,\cdot,h}(0)\}_{n\in\mathbb{Z}_\varkappa^+}$ converges to w^* weakly in H_1. Hence Lemma B.6 from [53] shows that w^* is a strong solution of equation $(2.5)_h$. This proves the first part of Theorem 7.1.

It remains to show that $w^*(t,h,\omega)$ has the attraction property. Note that each $w_{n_1-n,t,h}(0)$ starts from 0, hence estimate (A.1) implies that for $p>0$, $\mathbf{E}\|w_{n_1-n,t,h}(0)\|^p\leq C$ for some constant C independent of n and t. This in turn shows that $\mathbf{E}\|w^*(t,h)\|^p\leq C$ for $t\in\mathbb{R}$. Let $\varepsilon=\delta_0-\delta$. Define the random time $\tau_1=\frac{2c_0^2}{\varepsilon\nu^2}\|w^*(s)\|^2$, which has all moments finite. Note that for $\tau>\tau_1$, we have $\frac{c_0^2}{\nu^2\tau}\|w^*(s)\|^2<\frac{\varepsilon}{2}$, which controls the first term in $\Gamma(s,s+\tau)$ from (7.3). Based on the same reasoning as in *Step 1*, we find that there exists a $\mathbb{Z}_\varkappa^+$ valued random time $n_1(s,\delta,h,\varkappa)$ such that for all $\tau>n_1$, one has

$$\frac{c_0^2}{\nu^2\tau}|M(s,s+\tau)|<\frac{\varepsilon}{2}.$$

The estimate (7.1) then follows by taking $n^*=\max\{\tau_1,n_1\}$. For $t>0$, inequality (7.3) states that

$$\|\mathfrak{e}_s\|^2\leq\|\mathfrak{e}_{s-t}\|^2e^{-t(\delta_0-\Gamma(s-t,s))},$$

where $\Gamma(s-t,s)=\frac{c_0^2}{\nu^2t}\left(\|w^*(s-t)\|^2+M(s-t,s)\right)$. Now s is fixed, so $M(s-t,s)$ is not a martingale since it runs backwards in time. Nonetheless, we can still use the same reasoning as above. By replacing the Doob L^p maximal inequality with the backwards maximal inequality (see Lemma A.6 in [53]), one obtains a $\mathbb{Z}_\varkappa^+$ valued random time $n_2(s,\delta,h,\varkappa)$ with all moments finite such that for all $t>n_2$,

$$\frac{c_0^2}{\nu^2t}|M(s-t,s)|<\frac{\varepsilon}{2}.$$

To estimate the term $\frac{1}{t}\|w^*(s-t)\|^2$, noting that Theorem 3.13 in [53] remains true in our setting, hence there exists a random time $\tau_2>0$, such that for all $t>\tau_2$, we have $\frac{c_0^2}{\nu^2t}\|w^*(s-t)\|^2<\frac{\varepsilon}{2}$. Therefore the inequality (7.2) holds by taking $n_*=\max\{\tau_2,n_2\}$.

The proof is complete. □

The following proposition shows that there is a continuous modification of the random field

$w^*(0,h,\omega)$, which is proved by applying the Kolmogorov continuity theorem. The desired quasi-periodic solution will be constructed from this random field.

Proposition 7.2. *Assume $\Psi \in C^\gamma(\mathbb{T}^n, H)$, $Gc_0 \le \sqrt{1/2}$. For any $p \ge 1$, if*

$$\nu^3 > 8pc_0^2\mathcal{B}_0, \tag{7.8}$$

then there is a constant $C > 0$ such that

$$\mathbf{E}\|w^*(0,h_1,\cdot) - w^*(0,h_2,\cdot)\|^{2p} \le C|h_1 - h_2|^{p\gamma}, \quad \forall h_1, h_2 \in \mathbb{T}^n. \tag{7.9}$$

In particular, for any $\overline{\eta} > 0$, if condition (7.8) holds for $p = \frac{n+\overline{\eta}}{\gamma}$, then the random field $w^*(0,h,\omega)$ has a continuous (with respect to h) modification, which is η-Hölder continuous for all $0 < \eta < \frac{\overline{\eta}\gamma}{2(n+\overline{\eta})}$.

Proof. For $N \in \mathbb{Z}_\varkappa^-$ and $h_1, h_2 \in \mathbb{T}^n$, let $\mathcal{R}_r = w_{N,r,h_1}(0) - w_{N,r,h_2}(0)$. To show (7.9), we first prove that the same inequality holds for $\mathcal{R}_0$ and then letting $N \to -\infty$. Let $\delta > 0$ be a constant whose value will be determined later. Since $w_{N,r,h_i}(0)$ is the solution to $(2.5)_{h_i}$ starting from initial position $0 \in H$ at initial time $s = N$, it follows that (see also the proof of (A.3))

$$\begin{aligned}
\partial_r\|\mathcal{R}_r\|^2 &= \langle \nu\Delta\mathcal{R}_r, 2\mathcal{R}_r\rangle - \langle B\left(\mathcal{K}\mathcal{R}_r, w_{N,r,h_1}(0)\right), 2\mathcal{R}_r\rangle + \langle \Psi(\beta_r h_1) - \Psi(\beta_r h_2), 2\mathcal{R}_r\rangle \\
&\le -2\nu\|\mathcal{R}_r\|_1^2 + 2c_0\|\mathcal{R}_r\|\|w_{N,r,h_1}(0)\|_1\|\mathcal{R}_r\|_1 + 2\|\Psi(\beta_r h_1) - \Psi(\beta_r h_2)\|\|\mathcal{R}_r\| \\
&\le -2\nu\|\mathcal{R}_r\|_1^2 + c_0^2\nu^{-1}\|\mathcal{R}_r\|^2\|w_{N,r,h_1}(0)\|_1^2 + \nu\|\mathcal{R}_r\|_1^2 + 4\delta^{-1}\|\Psi\|_\gamma|h_1 - h_2|^\gamma + \frac{\delta}{4}\|\mathcal{R}_r\|^2 \\
&\le -(\nu - \delta/4)\|\mathcal{R}_r\|^2 + c_0^2\nu^{-1}\|\mathcal{R}_r\|^2\|w_{N,r,h_1}(0)\|_1^2 + 4\delta^{-1}\|\Psi\|_\gamma|h_1 - h_2|^\gamma.
\end{aligned}$$

By the Gronwall's inequality and Hölder's inequality, and noting that $\mathcal{R}_N = 0$, we have for $p, q > 0$ with $1/p + 1/q = 1$,

$$\begin{aligned}
\|\mathcal{R}_0\|^{2p} &\le 4^p\delta^{-p}\|\Psi\|_\gamma^p|h_1 - h_2|^{p\gamma}\left(\int_N^0 \exp\left(\int_z^0 -\left(\nu - \frac{\delta}{4}\right) + c_0^2\nu^{-1}\|w_{N,r,h_1}(0)\|_1^2 dr\right)dz\right)^p \\
&\le 4^p\delta^{-p}\|\Psi\|_\gamma^p|h_1 - h_2|^{p\gamma}\left(\int_N^0 e^{q\int_z^0 -\frac{\delta}{4}dr}dz\right)^{\frac{p}{q}}\int_N^0 e^{p\int_z^0 -\left(\nu-\frac{\delta}{2}\right)+c_0^2\nu^{-1}\|w_{N,r,h_1}(0)\|_1^2 dr}dz \\
&\le 4^{2p-1}q^{-\frac{p}{q}}\delta^{-2p+1}\|\Psi\|_\gamma^p|h_1 - h_2|^{p\gamma}\int_N^0 e^{p\int_z^0 -\left(\nu-\frac{\delta}{2}\right)+c_0^2\nu^{-1}\|w_{N,r,h_1}(0)\|_1^2 dr}dz.
\end{aligned}\tag{7.10}$$

We would like to have a bound on the expectation of $\exp\left(\int_z^0 pc_0^2\nu^{-1}\|w_{N,r,h_1}(0)\|_1^2 dr\right)$. It follows from Ito's formula that for $N \le z \le t \le 0$,

$$
\begin{aligned}
&\|w_{N,t,h_1}\|^2 + \nu \int_z^t \|w_{N,r,h_1}\|_1^2\,dr - \mathcal{B}_0(t-z) \\
&\qquad = \|w_{N,z,h_1}\|^2 + 2\int_z^t \langle w_{N,r,h_1}, GdW(r)\rangle - \nu \int_z^t \|w_{N,r,h_1}\|_1^2\,dr + 2\int_z^t \langle w_{N,r,h_1}, \Psi(\beta_r h_1)\rangle\,dr \\
&\qquad \le \|w_{N,z,h_1}\|^2 + 2\int_z^t \langle w_{N,r,h_1}, GdW(r)\rangle + \frac{\|f\|_\infty^2}{\nu}(t-z).
\end{aligned}
$$

Therefore for $\varepsilon > 0$,

$$
\frac{(1+\varepsilon)pc_0^2}{\nu}\int_z^t \|w_{N,r,h_1}\|_1^2\,dr - \frac{(1+\varepsilon)pc_0^2}{\nu^2}\left(\mathcal{B}_0 + \frac{\|f\|_\infty^2}{\nu}\right)(t-z) \le M(z,t),
$$

where $M(z,t) := \frac{(1+\varepsilon)pc_0^2}{\nu^2}\|w_{N,z,h_1}\|^2 + \frac{2(1+\varepsilon)pc_0^2}{\nu^2}\int_z^t \langle w_{N,r,h_1}, GdW(r)\rangle$ is a continuous square integrable martingale whose quadratic variation $[M](z,t)$ satisfies

$$
[M](z,t) = \frac{4(1+\varepsilon)^2p^2c_0^4}{\nu^4}\int_z^t \sum_{k=1}^d \langle w_{N,r,h_1}, g_k\rangle^2 dr \le \frac{4(1+\varepsilon)^2p^2c_0^4\mathcal{B}_0}{\nu^4}\int_z^t \|w_{N,r,h_1}\|_1^2 dr.
$$

As a consequence, one has

$$
\begin{aligned}
\frac{pc_0^2}{\nu}\int_z^t \|w_{N,r,h_1}\|_1^2\,dr - \frac{(1+\varepsilon)pc_0^2}{\nu^2}\left(\mathcal{B}_0 + \frac{\|f\|_\infty^2}{\nu}\right)(t-z) &\le M(z,t) - \frac{\varepsilon pc_0^2}{\nu}\int_z^t \|w_{N,r,h_1}\|_1^2\,dr \\
&\le M(z,t) - \frac{\varepsilon\nu^3}{4(1+\varepsilon)^2pc_0^2\mathcal{B}_0}[M](z,t).
\end{aligned}
$$

Let $b = \frac{\varepsilon\nu^3}{2(1+\varepsilon)^2pc_0^2\mathcal{B}_0}$. It then follows from the exponential supermartingale inequality that for $K \ge 0$,

$$
\begin{aligned}
&\mathbf{P}\left(\sup_{t\ge z}\exp\left(\frac{pc_0^2}{\nu}\int_z^t \|w_{N,r,h_1}\|_1^2\,dr - \frac{(\varepsilon+1)pc_0^2}{\nu^2}\left(\mathcal{B}_0 + \frac{\|f\|_\infty^2}{\nu}\right)(t-z)\right) \ge e^K \middle| \mathcal{F}_z\right) \\
&= \mathbf{P}\left(\sup_{t\ge z}\left(\frac{pc_0^2}{\nu}\int_z^t \|w_{N,r,h_1}\|_1^2\,dr - \frac{(\varepsilon+1)pc_0^2}{\nu^2}\left(\mathcal{B}_0 + \frac{\|f\|_\infty^2}{\nu}\right)(t-z)\right) \ge K \middle| \mathcal{F}_z\right) \\
&\le \mathbf{P}\left(\sup_{t\ge z}\left(M(z,t) - \frac{b}{2}[M](z,t)\right) \ge K \middle| \mathcal{F}_z\right) \\
&= \mathbf{P}\left(\sup_{t\ge z}\exp\left(bM(z,t) - \frac{b^2}{2}[M](z,t)\right) \ge e^{bK} \middle| \mathcal{F}_z\right) \le \exp\left(\frac{(\varepsilon+1)pc_0^2}{\nu^2}\|w_{N,z,h_1}\|^2\right)e^{-bK}.
\end{aligned}
$$

If $b = \frac{\varepsilon\nu^3}{2(1+\varepsilon)^2pc_0^2\mathcal{B}_0} > 1$, then the same argument as we derive (A.11) implies that

$$
\begin{aligned}
&\mathbf{E}\sup_{t\ge z}\exp\left(\frac{pc_0^2}{\nu}\int_z^t \|w_{N,r,h_1}\|_1^2\,dr - \frac{(\varepsilon+1)pc_0^2}{\nu^2}\left(\mathcal{B}_0 + \frac{\|f\|_\infty^2}{\nu}\right)(t-z)\right) \\
&\le \mathbf{E}\exp\left(\frac{(\varepsilon+1)pc_0^2}{\nu^2}\|w_{N,z,h_1}\|^2\right)\frac{4}{1-2^{1-b}}.
\end{aligned}
\tag{7.11}
$$

To bound the expectation on the right, we note that the solution w_{N,r,h_1} starts from 0, so (A.10) implies that

$$
\mathbf{P}\left(\|w_{N,z,h_1}\|^2 - \frac{C(f,\mathcal{B}_0)}{\nu} > \frac{K}{\alpha}\right) \le e^{-K},
$$

where $\alpha = \frac{(1-a)\nu}{2\mathcal{B}_0}$, $C(f,\mathcal{B}_0) = \frac{\|f\|_\infty^2}{2a} + \mathcal{B}_0$ and $a \in (0,1)$. Therefore

$$\mathbf{P}\left(\exp\left(\frac{(\varepsilon+1)pc_0^2}{\nu^2}\|w_{N,z,h_1}\|^2 - \frac{(\varepsilon+1)pc_0^2 C(f,\mathcal{B}_0)}{\nu^3}\right) > e^{\frac{(\varepsilon+1)pc_0^2 K}{\alpha\nu^2}}\right) \le e^{-K},$$

Again by the same argument as in (A.11) with $c = \frac{\alpha\nu^2}{(\varepsilon+1)pc_0^2}$, one has

$$\mathbf{E}\exp\left(\frac{(\varepsilon+1)pc_0^2}{\nu^2}\|w_{N,z,h_1}\|^2 - \frac{(\varepsilon+1)pc_0^2 C(f,\mathcal{B}_0)}{\nu^3}\right) \le \frac{4}{1-2^{1-c}}, \tag{7.12}$$

provided that $c = \frac{\alpha\nu^2}{(\varepsilon+1)pc_0^2} = \frac{(1-a)\nu^3}{2(\varepsilon+1)pc_0^2\mathcal{B}_0} > 1$. We take $a = \frac{1}{1+\varepsilon}$, which yields $b = c$.

Now it follows from (7.10)-(7.12) that

$$\begin{aligned}
\mathbf{E}\|\mathcal{R}_0\|^{2p} &\le 4^{2p-1}q^{-\frac{p}{q}}\delta^{-2p+1}\|\Psi\|_\gamma^p|h_1-h_2|^{p\gamma}\int_N^0 e^{p\left(\nu-\frac{\delta}{2}\right)z}\mathbf{E}e^{p\int_z^0 c_0^2\nu^{-1}\|w_{N,r,h_1}(0)\|_1^2dr}dz \\
&\le C(\delta,\varepsilon)|h_1-h_2|^{p\gamma}\int_N^0 e^{p\left(\nu-\frac{\delta}{2}\right)z}e^{-\frac{(1+\varepsilon)pc_0^2}{\nu^2}\left(\mathcal{B}_0+\frac{\|f\|_\infty^2}{\nu}\right)z}dz \\
&= C(\delta,\varepsilon)|h_1-h_2|^{p\gamma}\int_N^0 e^{pz\left(\nu-\frac{\delta}{2}-\frac{(1+\varepsilon)c_0^2}{\nu^2}\left(\mathcal{B}_0+\frac{\|f\|_\infty^2}{\nu}\right)\right)}dz
\end{aligned} \tag{7.13}$$

where $C(\delta,\varepsilon) = \left(\frac{1}{1-2^{1-c}}\right)^2 4^{2p+1}q^{-\frac{p}{q}}\delta^{-2p+1}\|\Psi\|_\gamma^p\exp\left(\frac{(1+\varepsilon)pc_0^2C(f,\mathcal{B}_0)}{\nu^3}\right)$. Since

$$\delta_0 = \nu - \frac{c_0^2}{\nu^2}\left(\mathcal{B}_0 + \frac{\|f\|_\infty^2}{\nu}\right) > 0,$$

one has

$$\delta_\varepsilon := \nu - \frac{(1+\varepsilon)c_0^2}{\nu^2}\left(\mathcal{B}_0 + \frac{\|f\|_\infty^2}{\nu}\right) > 0, \quad \text{as long as} \quad \varepsilon < \varepsilon_0 := (Gc_0)^{-2} - 1.$$

Keep in mind that we need to ensure inequalities (7.11) and (7.12) hold, which amounts to showing the existence of $\varepsilon \in (0,\varepsilon_0)$ such that $c = \frac{\varepsilon\nu^3}{2(1+\varepsilon)^2pc_0^2\mathcal{B}_0} > 1$. This can be achieved if we let $(Gc_0)^2 \le \frac{1}{2}$ (thus $\varepsilon_0 \ge 1$), and

$$\frac{2pc_0^2\mathcal{B}_0}{\nu^3} < \sup_{\varepsilon\in(0,\varepsilon_0)}\frac{\varepsilon}{(1+\varepsilon)^2} = \frac{1}{4}.$$

Therefore, under the conditions given in Proposition 7.2, there is an $\overline{\varepsilon} \in (0,\varepsilon_0)$, such that $\delta_{\overline{\varepsilon}} > 0$. And by choosing $\delta = \delta_{\overline{\varepsilon}}$, we have from (7.13) that

$$\mathbf{E}\|\mathcal{R}_0\|^{2p} \le C|h_1-h_2|^{p\gamma}, \quad \text{with} \quad C = \frac{2C(\delta_{\overline{\varepsilon}},\overline{\varepsilon})}{p\delta_{\overline{\varepsilon}}}.$$

By Theorem 7.1, we know that for $i = 1,2$, as $N \to -\infty$, $w_{N,r,h_i}(0)$ converges to $w^*(0,h_i,\omega)$ almost surely. Hence it follows from Fatou's lemma that

$$\mathbf{E}\|w^*(0,h_1,\cdot) - w^*(0,h_2,\cdot)\|^{2p} \le \liminf_{N\to-\infty}\mathbf{E}\|\mathcal{R}_0\|^{2p} \le C|h_1-h_2|^{p\gamma}. \tag{7.14}$$

Since the dimension of the torus is n, by taking $p = \frac{n+\overline{\eta}}{\gamma}$ for any $\overline{\eta} > 0$, we have

$$\mathbf{E}\|w^*(0,h_1,\cdot) - w^*(0,h_2,\cdot)\|^{\frac{2(n+\overline{\eta})}{\gamma}} \leq C|h_1 - h_2|^{n+\overline{\eta}}, \quad \forall h_1, h_2 \in \mathbb{T}^n.$$

Hence by Kolmogorov's continuity test, the random field $w^*(0,h,\omega)$ has a continuous modification. Moreover, with probability one, the modification is η-Hölder continuous for all $0 < \eta < \frac{\overline{\eta}\gamma}{2(n+\overline{\eta})}$. □

Note that for $h = 0 \in \mathbb{T}^n$, the associated equation $(2.5)_0$ is actually (2.5). We will show that the complete trajectory $w^*(t,0,\omega)$ given in Theorem 7.1 is a quasi-periodic solution in the sense of Definition 2.5.

Proposition 7.3. *The process $w^*(t,0,\omega)$ is a quasi-periodic solution of* (2.5).

Proof. The invariance property in Definition 2.5 follows from the stochastic flow property and Theorem 7.1. Note that for any $\tau \geq 0$ and $s \in \mathbb{R}$,

$$\begin{aligned} w^*(s+\tau,0,\omega) &= \lim_{\substack{N\to\infty \\ N\in\mathbb{Z}_\varkappa}} w(n_1 - N, s+\tau, 0, \omega, 0) \\ &= \lim_{\substack{N\to\infty \\ N\in\mathbb{Z}_\varkappa}} \Phi(s+\tau,\omega; s, w(n_1 - N, s, 0, \omega, 0)) \\ &= \Phi(s+\tau,\omega; s, \lim_{\substack{N\to\infty \\ N\in\mathbb{Z}_\varkappa}} w(n_1 - N, s, 0, \omega, 0)) = \Phi(s+\tau,\omega; s, w^*(s,0,\omega)), \quad \mathbf{P} - \text{a.s.} \end{aligned}$$

For $t \in \mathbb{Z}_\varkappa$, it follows that

$$\begin{aligned} w^*(t,0,\omega) &= \lim_{\substack{N\to\infty \\ N\in\mathbb{Z}_\varkappa}} w(n_1 - N, t, 0, \omega, 0) \\ &= \lim_{\substack{N\to\infty \\ N\in\mathbb{Z}_\varkappa}} w(n_1 - N - t, 0, \beta_t 0, \theta_t\omega, 0) = w^*(0, \beta_t 0, \theta_t \omega), \quad \mathbf{P} - \text{a.s.} \end{aligned}$$

In particular, for each $\varkappa \in \mathbb{R}$, there is a subset $\Omega_\varkappa \subset \Omega$ of full measure such that

$$w^*(\varkappa, 0, \omega) = w^*(0, \beta_\varkappa 0, \theta_\varkappa \omega), \quad \forall \omega \in \Omega_\varkappa.$$

Therefore for each $t \in \mathbb{R}$,

$$w^*(t, 0, \theta_{-t}\omega) = w^*(0, \beta_t 0, \omega), \quad \mathbf{P} - \text{a.s.}$$

By Proposition 7.2, we can choose a continuous version of $w^*(0,h,\omega)$ such that $w^*(t,0,\theta_{-t}\omega)$ is a random quasi-periodic function in the sense of Definition 2.5. This completes the proof. □

Appendix A. Various Estimates of The Solution

Several estimates about the solution $w_{s,t,h}(w_0)$ of the stochastic Navier-Stokes equation (2.5) with time symbol $h \in \mathbb{T}^n$ are collected in the following Lemma A.1. Note that for any $h \in \mathbb{T}^n$, we have

$$\sup_{t\in\mathbb{R}} \|\Psi(\beta_t h)\| = \sup_{t\in\mathbb{R}} \|\Psi(\beta_t 0)\| = \sup_{t\in\mathbb{R}} \|f(t)\| := \|f\|_\infty,$$

therefore the constant C in bounds on the solution does not depend on h.

Lemma A.1. *Let $a \in (0,1)$, $c \in (1,\infty)$, $\eta_0 = \frac{(1-a)\nu}{2c\mathcal{B}_0}$ and $C(f,\mathcal{B}_0) = \frac{\|f\|_\infty^2}{a\nu} + \mathcal{B}_0$. For solutions of the Navier-Stokes equation (2.5), we have*

(i) For every $t > s, h \in \mathbb{T}^n$ and every $\eta \in (0,\eta_0]$, we have

$$\mathbf{E}\exp\left(\eta \|w_{s,t,h}\|^2\right) \le C\exp\left(\eta e^{-\nu(t-s)}\|w_0\|^2\right), \tag{A.1}$$

where $C = \frac{4}{1-2^{1-c}}\exp\left(\frac{\eta_0 C(f,\mathcal{B}_0)}{\nu}\right)$.

(ii) The following inequality

$$\mathbf{E}\exp\left(\eta\sup_{t\ge\tau}\left(\|w_{s,t,h}\|^2 + \nu\int_\tau^t \|w_{s,r,h}\|_1^2\,dr - C(f,\mathcal{B}_0)(t-\tau)\right)\right) \le C\exp\left(\eta e^{-\nu(\tau-s)}\|w_0\|^2\right) \tag{A.2}$$

holds for every $\tau \ge s$ and $\eta \in (0,\eta_0]$, where $C = \frac{16}{(1-2^{1-c})^2}\exp\left(\frac{\eta_0 C(f,\mathcal{B}_0)}{\nu}\right)$.

(iii) For any $h_1, h_2 \in \mathbb{T}^n$, and every $\eta \in (0,\eta_0]$, with $r = 64c_0^6\eta^{-3}\nu^{-5} + \eta C(f,\mathcal{B}_0)$ and $\overline{C} = \frac{16(r\nu)^{-1}}{(1-2^{1-c})^2}\exp\left(\frac{\eta_0 C(f,\mathcal{B}_0)}{\nu}\right)$, we have

$$\mathbf{E}\|w_{s,t,h_1} - w_{s,t,h_2}\|^2 \le Ce^{r(t-s)}\exp\left(\eta\|w_0\|^2\right)\sup_{t\in\mathbb{R}}\|\Psi(\beta_t h_1) - \Psi(\beta_t h_2)\|^2, \tag{A.3}$$

for every $t \ge s$, where for $i = 1,2$, w_{s,t,h_i} is the solution to equation (2.5) with $f = \Psi(\beta_t 0)$ replaced by $\Psi(\beta_t h_i)$ and with initial condition (s, w_0).

(iv) For any $h \in \mathbb{T}^n$, and every $\eta \in (0,\eta_0]$, we have

$$\mathbf{E}\|w_{s,s+t,h}(w_1) - w_{s,s+t,h}(w_2)\|^2 \le C\|w_1 - w_2\|^2 e^{\eta\|w_1\|^2 + r(t-s)}, \tag{A.4}$$

for every $s \in \mathbb{R}$ and $t \ge 0$. Here $C = 64(1-2^{1-c})^{-3/2}\exp\left(\frac{\eta_0 C(f,\mathcal{B}_0)}{\nu}\right)$ and $r = 64c_0^6\eta^{-3}\nu^{-5} + \eta C(f,\mathcal{B}_0)$.

(v) There exist constants $\eta_1, a, \gamma > 0$, depending on $f, \mathcal{B}_0, \nu, \eta_0$, such that

$$\mathbf{E} \exp\left(\eta \sum_{n=0}^{N} \|w_{s,s+n,h}\|^2 - \gamma N\right) \leq \exp\left(a\eta \|w_0\|^2\right) \tag{A.5}$$

holds for every integer $N > 0$, every $\eta \leq \eta_1$, $h \in \mathbb{T}^n$, $s \in \mathbb{R}$ and every initial condition $w_0 \in H$.

(vi) For every $\eta > 0$, $h \in \mathbb{T}^n$ and $t \geq \tau \geq s$, there exists a constant $C = C(\nu, \eta) > 0$ such that the Jacobian $J_{\tau,t,h}$ as defined in (4.20) satisfies almost surely

$$\|J_{\tau,t,h}\| \leq \exp\left(\eta \int_\tau^t \|w_{s,r,h}\|_1^2 \, dr + C(t-\tau)\right). \tag{A.6}$$

(vii) For every $\eta > 0$, $h \in \mathbb{T}^n$ and every $p > 0$, there exists $C = C(f, \mathcal{B}_0, \nu, \eta, p) > 0$ such that the Hessian as defined in (4.22) satisfies

$$\|K_{\tau,t,h}\|^p \leq C \exp\left(p\eta \int_\tau^{\tau+1} \|w_{s,r,h}\|_1^2 \, dr\right) \tag{A.7}$$

for every $\tau \geq s$, $h \in \mathbb{T}^n$ and $t \in (\tau, \tau+1)$.

(viii) For any integer $k \geq 0$, set

$$\mathcal{E}_{w_0}(k,t,s) = (t-s)^k \|w_{s,t}\|_k^2 + \nu \int_s^t (r-s)^k \|w_{s,r}\|_{k+1}^2 dr.$$

Suppose that in equation (2.5), $f \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}, H_k)$ and $g_i \in H_k$ for $i = 1, \cdots, d$. Then for any $m \geq 1$, $\eta > 0$ and $T > s$, there is a constant $C = C(k, m, T-s, \nu, \|f\|_{\mathrm{L}^2([s,T],H_k)}, \mathcal{B}_k, \eta) > 0$, such that

$$\mathbf{E} \sup_{s \leq t \leq T} \mathcal{E}_{w_0}(k,t,s)^m \leq C \exp\left(\eta \|w_0\|^2\right). \tag{A.8}$$

Proof. (i) Applying Ito's formula to the functional $F(t,w) = e^{\nu(t-s)}\|w\|^2$ and noting the fact that $\langle B(\mathcal{K}w, w), w\rangle = 0$, we have

$$\begin{aligned} e^{\nu(t-s)} \|w_{s,t,h}\|^2 = &- \|w_0\|^2 - 2\nu \int_s^t e^{\nu(r-s)} \|w_{s,r,h}\|_1^2 \, dr + 2 \int_s^t e^{\nu(r-s)} \left\langle w_{s,r,h}, \sum_{k=1}^d g_k dW_k(r) \right\rangle \\ &+ \nu \int_s^t e^{\nu(r-s)} \|w_{s,r,h}\|^2 \, dr + 2 \int_s^t e^{\nu(r-s)} \langle w_{s,r,h}, \Psi(\beta_r h)\rangle \, dr + \frac{e^{\nu(t-s)} - 1}{\nu} \mathcal{B}_0. \end{aligned}$$

Given $0 < a < 1$, let $C(f, \mathcal{B}_0) = \frac{\|f\|_\infty^2}{a\nu} + \mathcal{B}_0$. By the inequality $\|w\|_1^2 \geq \|w\|^2$ and Young's product inequality, it follows that

$$\begin{aligned} &\|w_{s,t,h}\|^2 - e^{-\nu(t-s)}\|w_0\|^2 - \frac{C(f, \mathcal{B}_0)}{\nu} \\ &\quad \leq -(1-a)\nu \int_s^t e^{-\nu(t-r)} \|w_{s,r,h}\|^2 dr + 2 \int_s^t e^{\nu(r-s)} \left\langle w_{s,r,h}, \sum_{k=1}^d g_k dW_k(r) \right\rangle. \end{aligned}$$

Let $M_t = 2\int_s^t \left\langle w_{s,r,h}, \sum_{k=1}^d g_k dW_k(r)\right\rangle$ whose quadratic variation is

$$[M]_t = 4\int_s^t \sum_{k=1}^d \langle w_{s,r,h}, g_k\rangle^2 \, dr.$$

Observe that

$$-(1-a)\nu\|w\|^2 \leq -\frac{(1-a)\nu}{\mathcal{B}_0}\sum_{k=1}^d \langle w, g_k\rangle^2 = -\frac{\alpha}{2}\sum_{k=1}^d 4\langle w, g_k\rangle^2 \tag{A.9}$$

where $\alpha = \frac{(1-a)\nu}{2\mathcal{B}_0}$. Hence

$$\|w_{s,t,h}\|^2 - e^{-\nu(t-s)}\|w_0\|^2 - \frac{C(f,\mathcal{B}_0)}{\nu} \leq \int_s^t e^{-\nu(t-r)} dM_r - \frac{\alpha}{2}\int_s^t e^{-\nu(t-r)} d[M]_r.$$

Then Lemma A.1 from [54] implies that

$$\mathbf{P}\left(\|w_{s,t,h}\|^2 - e^{-\nu(t-s)}\|w_0\|^2 - \frac{C(f,\mathcal{B}_0)}{\nu} > \frac{K}{\alpha}\right) \leq e^{-K}, \tag{A.10}$$

which is, for any $c > 1$ and with $\eta_0 = \frac{\alpha}{c}$, equivalent to

$$\mathbf{P}\left(\exp\left(\eta_0\|w_{s,t,h}\|^2 - \eta_0 e^{-\nu(t-s)}\|w_0\|^2 - \frac{\eta_0 C(f,\mathcal{B}_0)}{\nu}\right) > e^{\frac{K}{c}}\right) \leq e^{-K}.$$

Now if a random variable X satisfies $\mathbf{P}(X \geq C) \leq \frac{1}{C^c}$ for every $C \geq 1$, then

$$\begin{aligned}\mathbf{E}X = \int_\Omega X d\mathbf{P} &\leq \int_{\{0\leq X\leq 1\}} X d\mathbf{P} + \int_{\{X\geq 1\}} X d\mathbf{P} \leq 1 + \sum_{n=0}^\infty \int_{\{2^n\leq X\leq 2^{n+1}\}} X d\mathbf{P} \\ &\leq 1 + \sum_{n=0}^\infty 2^{n+1}\frac{1}{2^{cn}} \leq \frac{4}{1-2^{1-c}}.\end{aligned} \tag{A.11}$$

Therefore we have for $\eta \in (0, \eta_0]$, by Hölder's inequality,

$$\begin{aligned}&\mathbf{E}\exp\left(\eta\|w_{s,t,h}\|^2 - \eta e^{-\nu(t-s)}\|w_0\|^2 - \frac{\eta C(f,\mathcal{B}_0)}{\nu}\right) \\ &\leq \left(\mathbf{E}\exp\left(\eta_0\|w_{s,t,h}\|^2 - \eta_0 e^{-\nu(t-s)}\|w_0\|^2 - \frac{\eta_0 C(f,\mathcal{B}_0)}{\nu}\right)\right)^{\eta/\eta_0} \leq \frac{4}{1-2^{1-c}}.\end{aligned}$$

Hence we arrive at (A.1) with $C = \frac{4}{1-2^{1-c}}\exp\left(\frac{\eta_0 C(f,\mathcal{B}_0)}{\nu}\right)$ and $\eta_0 = \frac{\alpha}{c} = \frac{(1-a)\nu}{2c\mathcal{B}_0}$.

(ii) Again apply Ito's formula, for any $\eta > 0$, $s \leq \tau < t$,

$$\begin{aligned}&\eta\|w_{s,t,h}\|^2 + \eta\nu\int_\tau^t \|w_{s,r,h}\|_1^2\, dr - \eta\mathcal{B}_0(t-\tau) \\ &= \eta\|w_{s,\tau,h}\|^2 + 2\eta\int_\tau^t \langle w_{s,r,h}, GdW(r)\rangle - \eta\nu\int_\tau^t \|w_{s,r,h}\|_1^2\, dr + 2\eta\int_\tau^t \langle w_{s,r,h}, \Psi(\beta_r h)\rangle\, dr \\ &\leq \eta\|w_{s,\tau,h}\|^2 + 2\eta\int_\tau^t \langle w_{s,r,h}, GdW(r)\rangle - \eta\nu(1-a)\int_\tau^t \|w_{s,r,h}\|_1^2\, dr + \frac{\eta}{a\nu}\|f\|_\infty^2 (t-\tau).\end{aligned} \tag{A.12}$$

Therefore

$$\eta \left\| w_{s,t,h} \right\|^2 + \eta\nu \int_\tau^t \left\| w_{s,r,h} \right\|_1^2 dr - \eta C(f,\mathcal{B}_0)(t-\tau)$$

$$\leq \eta \left\| w_{s,\tau,h} \right\|^2 + 2\eta \int_\tau^t \left\langle w_{s,r,h}, GdW(r) \right\rangle - \eta\nu(1-a) \int_\tau^t \left\| w_{s,r,h} \right\|_1^2 dr$$

$$\leq \eta \left\| w_{s,\tau,h} \right\|^2 + 2\eta \int_\tau^t \left\langle w_{s,r,h}, GdW(r) \right\rangle - \frac{\eta\alpha}{2} \int_\tau^t \sum_{k=1}^d 4 \left\langle w_{s,r,h}, g_k \right\rangle^2 dr. \qquad \text{(A.13)}$$

Again the last inequality is obtained by choosing $\alpha = \frac{(1-a)\nu}{2\mathcal{B}_0}$ as in (A.9). Setting $M(\tau,t) = \eta \left\| w_{s,\tau,h} \right\|^2 + 2\eta \int_\tau^t \left\langle w_{s,r,h}, GdW(r) \right\rangle$, then the right hand side of inequality (A.13) is $M(\tau,t) - \frac{\alpha}{2\eta}[M](\tau,t)$, where $[M](\tau,t)$ is the quadratic variation of the continuous L^2-martingale M. Hence by the exponential supermartingale inequality, it follows that

$$\mathbf{P}\left(\sup_{t\geq\tau} \left(M(\tau,t) - \frac{\alpha}{2\eta}[M](\tau,t) \right) \geq K \,\middle|\, \mathcal{F}_\tau \right) \leq \exp\left(\eta \left\| w_{s,\tau,h} \right\|^2 - \frac{\alpha K}{\eta} \right)$$

for all $\tau \geq s$. As a consequence,

$$\mathbf{P}\Big(\eta \sup_{t\geq\tau} \Big(\left\| w_{s,t,h} \right\|^2 + \nu \int_\tau^t \left\| w_{s,r,h} \right\|_1^2 dr - C(f,\mathcal{B}_0)(t-\tau) \Big) \geq K \,\Big|\, \mathcal{F}_\tau \Big) \qquad \text{(A.14)}$$

$$\leq \exp\Big(\eta \left\| w_{s,\tau,h} \right\|^2 - \frac{\alpha K}{\eta} \Big).$$

In view of (A.13) and (A.11), we deduce that

$$\mathbf{E}\exp\left(\frac{\alpha}{c} \sup_{t\geq\tau} \left(\|w_{s,t,h}\|^2 + \nu \int_\tau^t \left\| w_{s,r,h} \right\|_1^2 dr - C(f,\mathcal{B}_0)(t-\tau) \right) \right) \leq \frac{4}{1-2^{1-c}} \mathbf{E}\exp\left(\eta \|w_{s,\tau,h}\|^2 \right).$$

The conclusion follows from (A.1) by taking $\eta_0 = \frac{\alpha}{c}$.

(iii) Let $\mathcal{R}_t = w_{s,t,h_1} - w_{s,t,h_2}$ then

$$\partial_t \mathcal{R}_t = \nu\Delta\mathcal{R}_t + B\left(\mathcal{K}w_{s,t,h_2}, w_{s,t,h_2}\right) - B\left(\mathcal{K}w_{s,t,h_1}, w_{s,t,h_1}\right) + \Psi(\beta_t h_2) - \Psi(\beta_t h_1)$$

$$= \nu\Delta\mathcal{R}_t + B\left(\mathcal{K}\mathcal{R}_t, \mathcal{R}_t\right) - B\left(\mathcal{K}\mathcal{R}_t, w_{s,t,h_1}\right) - B\left(\mathcal{K}w_{s,t,h_1}, \mathcal{R}_t\right) + \Psi(\beta_t h_2) - \Psi(\beta_t h_1).$$

Therefore from the inequality $|\langle B(\mathcal{K}u,v),w\rangle| \leq c_0\|u\|\|v\|_1\|w\|_{1/2}$ and the interpolation inequality $\|w\|_{1/2}^2 \leq \varepsilon\|w\|_1^2 + \varepsilon^{-2}\|w\|^2$ for $\varepsilon > 0$, we have

$$\partial_t\|\mathcal{R}_t\|^2 = \langle \nu\Delta\mathcal{R}_t, 2\mathcal{R}_t\rangle - \left\langle B\left(\mathcal{K}\mathcal{R}_t, w_{s,t,h_1}\right), 2\mathcal{R}_t\right\rangle + \langle \Psi(\beta_t h_2) - \Psi(\beta_t h_1), 2\mathcal{R}_t\rangle$$

$$\leq -2\nu\|\mathcal{R}_t\|_1^2 + 2c_0\|\mathcal{R}_t\|\|w_{s,t,h_1}\|_1\|\mathcal{R}_t\|_{1/2} + 2\sup_{t\in\mathbb{R}}\|\Psi(\beta_t h_2) - \Psi(\beta_t h_1)\|\|\mathcal{R}_t\|$$

$$\leq -2\nu\|\mathcal{R}_t\|_1^2 + \frac{4c_0^2}{\eta\nu}\|\mathcal{R}_t\|_1^2 + \eta\nu\|w_{s,t,h_1}\|_1^2\|\mathcal{R}_t\|^2 + \nu\|\mathcal{R}_t\|_1^2 + \frac{1}{\nu}\sup_{t\in\mathbb{R}}\|\Psi(\beta_t h_2) - \Psi(\beta_t h_1)\|^2$$

$$\leq \left(C(\eta,\nu) + \eta\nu\|w_{s,t,h_1}\|_1^2\right)\|\mathcal{R}_t\|^2 + \frac{1}{\nu}\sup_{t\in\mathbb{R}}\|\Psi(\beta_t h_2) - \Psi(\beta_t h_1)\|^2,$$

where $C(\eta,\nu) = 64c_0^6\eta^{-3}\nu^{-5}$. By Gronwall's inequality, and estimate (A.2), we have

$$\begin{aligned}\mathbf{E}\|\mathcal{R}_t\|^2 &\leq \frac{1}{\nu}\sup_{t\in\mathbb{R}}\|\Psi(\beta_t h_2)-\Psi(\beta_t h_1)\|^2\int_s^t \exp\Big(C(\eta,\nu)(t-\tau)\Big)\mathbf{E}\left[\exp\left(\int_\tau^t \eta\nu\|w_{s,r,h_1}\|_1^2\right)dr\right]d\tau \\ &\leq \frac{C}{\nu}\sup_{t\in\mathbb{R}}\|\Psi(\beta_t h_2)-\Psi(\beta_t h_1)\|^2\int_s^t \exp\left(r(t-\tau)\right)\exp\left(\eta e^{-\nu(\tau-s)}\|w_0\|^2\right)d\tau \\ &\leq \overline{C}e^{r(t-s)}\exp\left(\eta\|w_0\|^2\right)\sup_{t\in\mathbb{R}}\|\Psi(\beta_t h_2)-\Psi(\beta_t h_1)\|^2,\end{aligned}$$

with $r = 64c_0^6\eta^{-3}\nu^{-5}+\eta C(f,\mathcal{B}_0)$ and $\overline{C} = \frac{16(r\nu)^{-1}}{(1-2^{1-c})^2}\exp\left(\frac{\eta_0 C(f,\mathcal{B}_0)}{\nu}\right)$.

(iv) We now compare solutions that start from different initial positions. Let $w_0, \widetilde{w}_0 \in H$, and $\mathfrak{e}_t = w_{s,t,h}(w_0) - \widetilde{w}_{s,t,h}(\widetilde{w}_0)$, where $w_{s,t,h}(w_0), \widetilde{w}_{s,t,h}(\widetilde{w}_0)$ are the solutions starting from $w_0, \widetilde{w}_0$. In view of equation (2.5) with symbol h, we see that $\mathfrak{e}_t$ solves the following equation

$$\begin{aligned}\partial_t\mathfrak{e}_t &= \nu\Delta\mathfrak{e}_t - B(\mathcal{K}w_{s,t,h},w_{s,t,h}) + B(\mathcal{K}\widetilde{w}_{s,t,h},\widetilde{w}_{s,t,h}) \\ &= \nu\Delta\mathfrak{e}_t - B(\mathcal{K}w_{s,t,h},w_{s,t,h}) + B(\mathcal{K}(w_{s,t,h}-\mathfrak{e}_t),w_{s,t,h}-\mathfrak{e}_t) \\ &= \nu\Delta\mathfrak{e}_t - B(\mathcal{K}w_{s,t,h},\mathfrak{e}_t) + B(\mathcal{K}\mathfrak{e}_t,w_{s,t,h}) + B(\mathcal{K}\mathfrak{e}_t,\mathfrak{e}_t).\end{aligned}$$

From the fact $\langle B(\mathcal{K}w,v),v\rangle = 0$ and the basic estimates of the nonlinear term as in the proof of (A.3), we have

$$\begin{aligned}\partial_t\|\mathfrak{e}_t\|^2 = 2\langle\mathfrak{e}_t,\partial_t\mathfrak{e}_t\rangle &= -2\nu\|\mathfrak{e}_t\|_1^2 + 2\langle B(\mathcal{K}\mathfrak{e}_t,w_{s,t,h}),\mathfrak{e}_t\rangle \\ &\leq -2\nu\|\mathfrak{e}_t\|_1^2 + 2c_0\|\mathfrak{e}_t\|\|w_{s,t,h}\|_1\|\mathfrak{e}_t\|_{1/2} \\ &\leq \left(C(\eta,\nu)+\eta\nu\|w_{s,t,h_1}\|_1^2\right)\|\mathfrak{e}_t\|^2,\end{aligned}$$

where $C(\eta,\nu) = 64c_0^6\eta^{-3}\nu^{-5}$. Hence by Gronwall's inequality,

$$\|\mathfrak{e}_t\|^2 \leq \|\mathfrak{e}_s\|^2\exp\left(C(\eta,\nu)(t-s)+\eta\nu\int_s^t\|w_{s,r,h_1}\|_1^2 dt\right).$$

From the estimate (A.2), we have

$$\mathbf{E}\|w_{s,t,h}(w_0)-\widetilde{w}_{s,t,h}(\widetilde{w}_0)\|^2 \leq C\|w_0-\widetilde{w}_0\|^2 e^{\eta\|w_0\|^2+r(t-s)}, \tag{A.15}$$

where $C = \frac{16}{(1-2^{1-c})^2}\exp\left(\frac{\eta_0 C(f,\mathcal{B}_0)}{\nu}\right)$ and $r = 64c_0^6\eta^{-3}\nu^{-5}+\eta C(f,\mathcal{B}_0)$.

(v) The proof of the inequality (A.5) is the same as that in [40], hence we omit it here.

(vi) For any $\tau \geq s$ and initial condition $\xi \in H$, the evolution of $\xi_t := J_{\tau,t,h}\xi$ is given by equation (4.20), which is a PDE with random coefficients. Taking H inner product with ξ_t and using the fact that $\langle B(\mathcal{K}w,\xi),\xi\rangle = 0$, we have

$$\partial_t \|\xi_t\|^2 = -2\nu \|\nabla\xi_t\|^2 - 2\langle B(\mathcal{K}\xi_t, w_{s,t,h}),\xi_t\rangle.$$

Then note

$$\begin{aligned}|2\langle B(\mathcal{K}\xi_t, w_{s,t,h}),\xi_t\rangle| &\le C\|w_{s,t,h}\|_1 \|\xi_t\| \|\xi_t\|_{\frac{1}{2}} \le C_\eta \|\xi_t\|^2_{\frac{1}{2}} + \frac{\eta}{2}\|w_{s,t,h}\|_1^2 \|\xi_t\|^2 \\ &\le C_\eta\left(\varepsilon\|\xi_t\|^2 + \frac{1}{\varepsilon^2}\|\xi_t\|_1^2\right) + \frac{\eta}{2}\|w_{s,t,h}\|_1^2\|\xi_t\|^2 = C\|\xi_t\|^2 + \nu\|\xi_t\|_1^2 + \frac{\eta}{2}\|w_{s,t,h}\|_1^2\|\xi_t\|^2,\end{aligned}$$

by choosing $\varepsilon = \sqrt{C_\eta/\nu}$, where C depends on η,ν. Therefore

$$\partial_t \|\xi_t\|^2 \le -\nu\|\nabla\xi_t\|^2 + C\|\xi_t\|^2 + \frac{\eta}{2}\|w_{s,t,h}\|_1^2\|\xi_t\|^2 \le C\|\xi_t\|^2 + \frac{\eta}{2}\|w_{s,t,h}\|_1^2\|\xi_t\|^2.$$

And the result follows by the Gronwall's inequality.

(vii) Define $\zeta_t = \|\xi_t\|^2 + \nu(t-\tau)\|\xi_t\|_1^2$. From the equation (4.20) for the Jacobian ξ_t, one has

$$\begin{aligned}\partial_t\left(\nu(t-\tau)\|\xi_t\|_1^2\right) &= \nu\|\xi_t\|_1^2 - 2\nu^2(t-\tau)\|\xi_t\|_2^2 + 2\nu(t-\tau)\langle\widetilde{B}(w_{s,t,h},\xi_t), -\Delta\xi_t\rangle \\ &\le \nu\|\xi_t\|_1^2 - 2\nu^2(t-\tau)\|\xi_t\|_2^2 + 2C\nu(t-\tau)\|\xi_t\|_1\|w_{s,t,h}\|_1\|\xi_t\|_{3/2}.\end{aligned}$$

Therefore

$$\partial_t\zeta_t \le C\|\xi_t\|^2 + \frac{\eta}{2}\|w_{s,t,h}\|_1^2\|\xi_t\|^2 - 2\nu^2(t-\tau)\|\xi_t\|_2^2 + 2C\nu(t-\tau)\|\xi_t\|_1\|w_{s,t,h}\|_1\|\xi_t\|_{3/2}.$$

From interpolation inequalities, one has

$$C\|w\|_1\|\xi\|_1\|\xi\|_{3/2} \le C_\eta\|\xi\|_{3/2}^2 + \frac{\eta}{2}\|w\|_1^2\|\xi\|_1^2 \le \nu\|\xi\|_2^2 + C\|\xi\|_1^2 + \frac{\eta}{2}\|w\|_1^2\|\xi\|_1^2.$$

As a consequence,

$$\partial_t\zeta_t \le C\|\zeta_t\|^2 + \frac{\eta}{2}\|w_{s,t,h}\|_1^2\|\zeta_t\|^2.$$

Hence by Gronwall's inequality, for $s \le \tau < t \le s+T$,

$$\|\xi_t\|_1^2 \le \frac{C}{t-\tau}\exp\left(\eta\int_\tau^t \|w_{s,r,h}\|_1^2\,dr\right)\|\xi\|^2, \tag{A.16}$$

where C depends on ν,η,T. From basic Sobolev inequalities and interpolation inequalities, one has

$$\begin{aligned}\|\widetilde{B}(u,w)\| &\le C\left(\|u\|_{1/2}\|w\|_1 + \|u\|_1\|w\|_{1/2}\right) \\ &\le C\left(\|u\|^{1/2}\|u\|_1^{1/2}\|w\|_1 + \|w\|^{1/2}\|w\|_1^{1/2}\|u\|_1\right).\end{aligned}$$

It then follows from the definition of $K_{\tau,t}$ in (4.22) that for $t \in (\tau, \tau+1)$,

$$\begin{aligned}\|K_{\tau,t,h}\| &\leq C \int_\tau^t \|J_{r,t,h}\| \, \|J_{\tau,r,h}\|^{\frac{1}{2}} \, \|J_{\tau,r,h}\|_1^{\frac{3}{2}} \, dr \\ &\leq \exp\left(\eta \int_\tau^{\tau+1} \|w_{s,r,h}\|_1^2 \, dr + C\right) \exp\left(\frac{3}{4}\eta \int_\tau^{\tau+1} \|w_{s,r,h}\|_1^2 \, dr\right) \int_\tau^{\tau+1} \frac{\tilde{C}}{(r-\tau)^{3/4}} dr \\ &\leq C \exp\left(\eta \int_\tau^{\tau+1} \|w_{s,r,h}\|_1^2 \, dr\right),\end{aligned}$$

where we used inequalities (A.6) and (A.16) in the second step. The proof is complete.

(viii) The proof proceeds by induction as in [48]. Let $L = -\Delta$, $F_k(w) = (t-s)^k \|w\|_k^2 = (t-s)^k \langle L^k w, w\rangle$. We first prove the base case when $k = 0$. Applying Ito's formula to the functional $F_0(w_{s,t})$, and noting $\langle B(\mathcal{K}w, w), w\rangle = 0$, we have for $s \leq t \leq T$,

$$\begin{aligned}\|w_{s,t}\|^2 &= \|w_0\|^2 - 2\nu \int_s^t \|w_{s,r}\|_1^2 dr + 2\int_s^t \langle w_{s,r}, f(r)\rangle dr + \mathcal{B}_0(t-s) + M_t \\ &\leq \|w_0\|^2 - 2\nu \int_s^t \|w_{s,r}\|_1^2 dr + \frac{\nu}{2}\int_s^t \|w_{s,r}\|^2 dr + \frac{2}{\nu}\|f\|^2_{\mathrm{L}^2([s,T],H)} + \mathcal{B}_0(T-s) + M_t\end{aligned}$$

where $M_t = 2\int_s^t \langle w_{s,r}, \sum_{i=1}^d g_i dW_i(r)\rangle$. Note that the quadratic variation process of M_t satisfies

$$[M]_t = \int_s^t 4\sum_{i=1}^d \langle w_{s,r}, g_i\rangle^2 dr \leq 4\mathcal{B}_0 \int_s^t \|w_{s,r}\|^2 dr.$$

If we let $\sigma_0 = \frac{1}{4\mathcal{B}_0}$, $C_0 = \frac{2}{\nu}\|f\|^2_{\mathrm{L}^2([s,T],H)} + \mathcal{B}_0(T-s)$, then it follows that

$$\mathcal{E}_{w_0}(0,t,s) \leq \|w_0\|^2 + C_0 + M_t - \frac{\sigma_0\nu}{2}[M]_t.$$

By the supermartingale inequality, one has for any $K > 0$

$$\mathbf{P}\left(\sup_{t\geq s}\left(M_t - \frac{\sigma_0\nu}{2}[M]_t\right) \geq K\right) \leq e^{-\sigma_0\nu K}.$$

Therefore

$$\mathbf{P}\left(\sup_{t\geq s}\left(\mathcal{E}_{w_0}(0,t,s) - C_0 - \|w_0\|^2\right) \geq K\right) \leq e^{-\sigma_0\nu K}.$$

Note for non-negative random variables a and b, one has

$$\mathbf{E}a^m \leq 2^m\left(\mathbf{E}\,(a-b)^m\,\mathbb{I}_{\{a>b\}} + \mathbf{E}b^m\right) = 2^m \int_0^\infty \mathbf{P}\{a-b > \lambda^{1/m}\} d\lambda + 2^m \mathbf{E}b^m.$$

Therefore

$$\begin{aligned}\mathbf{E}\sup_{t\geq s}\mathcal{E}_{w_0}(0,t,s)^m &\leq 2^m \int_0^\infty e^{-\sigma_0\nu\lambda^{1/m}} d\lambda + 2^m \mathbf{E}(C_0 + \|w_0\|^2)^m \\ &\leq C\exp\left(\eta\|w_0\|^2\right). \qquad \text{(A.17)}\end{aligned}$$

This finishes the proof of the base case.

We now assume that $k = n \geq 1$ and that for $k \leq n-1$ the inequality has been proved. Applying Ito's formula to $F_n(w_{s,t,h})$ we find that

$$\begin{aligned}
&F_n(w_{s,t})\\
&= \int_s^t n(r-s)^{n-1}\|w_{s,r}\|_n^2 + 2(r-s)^n\langle L^n w_{s,r}, -\nu L w_{s,r} - B(\mathcal{K}w_{s,r}, w_{s,r}) + f(r)\rangle dr\\
&+ \frac{(t-s)^{n+1}}{n+1}\mathcal{B}_n + \int_s^t 2(r-s)^n\langle L^n w_{s,r}, \sum_{i=1}^d g_i dW_i(r)\rangle.
\end{aligned}$$

Note the quadratic variation process $[M]_t$ of the martingale

$$M_t := \int_s^t 2(r-s)^n\langle L^n w_{s,r}, \sum_{i=1}^d g_i dW_i(r)\rangle$$

satisfies

$$[M]_t = \sum_{i=1}^d \int_s^t 4(r-s)^{2n}\langle L^n w_{s,r}, g_i\rangle^2 dr \leq 4\mathcal{B}_n\,(T-s)^n \int_s^t (r-s)^n\,\|w_{s,r,h}\|_n^2 dr.$$

Also note

$$\begin{aligned}
\int_s^t 2(r-s)^n\langle L^n w_{s,r}, f(r)\rangle dr &\leq \int_s^t 2(r-s)^n\|w_{s,r}\|_n\|f(r)\|_n dr\\
&\leq \int_s^t (r-s)^{n-1}\|w_{s,r}\|_n^2 + (r-s)^{n+1}\|f(r)\|_n^2 dr.
\end{aligned}$$

Applying the inequality $\|w\|_n \leq \|w\|_{n+1}$, and combining these estimates, it follows that

$$\begin{aligned}
F_n(w_{s,t}) \leq &(n+1)\int_s^t (r-s)^{n-1}\|w_{s,r}\|_n^2 dr - \frac{3\nu}{2}\int_s^t (r-s)^n\|w_{s,r}\|_{n+1}^2 dr\\
&- 2\int_s^t (r-s)^n\langle L^n w_{s,r}, B(\mathcal{K}w_{s,r}, w_{s,r})\rangle dr + C_n + N_n(t),
\end{aligned}$$

where $C_n = (T-s)^{n+1}\left(\|f\|^2_{\mathrm{L}^2([s,T],H_n)} + \frac{\mathcal{B}_n}{n+1}\right)$, $N_n(t) = M_t - \frac{\nu\sigma_n}{2}[M]_t$ and $\sigma_n = \frac{1}{4\mathcal{B}_n(T-s)^n}$.
When $n = 1$, the nonlinear term has the following bounds by the interpolation inequality,

$$2\langle Lw, B(\mathcal{K}w, w) \leq C\|w\|_{1/2}\|w\|_1\|w\|_2 \leq \frac{\nu}{2}\|w\|_2^2 + C\|w\|^{10}.$$

Then one has

$$\begin{aligned}
&\sup_{s\leq t\leq T} \mathcal{E}_{w_0}(1,t,s)\\
&\leq \frac{2}{\nu}\sup_{s\leq t\leq T}\mathcal{E}_{w_0}(0,t,s) + C\sup_{s\leq t\leq T}\int_s^t (r-s)\|w_{s,r}\|^{10} dr + C_1 + \sup_{t\geq s} N_1(t)\\
&\leq \frac{2}{\nu}\sup_{s\leq t\leq T}\mathcal{E}_{w_0}(0,t,s) + C(T-s)^2\sup_{s\leq t\leq T}\mathcal{E}_{w_0}(0,t,s)^{10} + C_1 + \sup_{t\geq s} N_1(t).
\end{aligned}$$

And the result then follows from the supermartingale inequality and the same argument as in the case $n = 0$. If $n \geq 2$, we use the following inequality to bound the nonlinear term

$$|\langle L^n w, B(\mathcal{K}w, w)\rangle| \leq C_n \|w\|_{n+1}^{\frac{4n-1}{2n}} \|w\|_1^{\frac{n+1}{2n}} \|w\|^{\frac{1}{2}} \leq \frac{\nu}{2}\|w\|_{n+1}^2 + C\|w\|_1^{2(n+1)}\|w\|^{2n},$$

which can be proved as Lemma 2.1.20 in [48]. It then follows that

$$\begin{aligned}
&\sup_{s\leq t\leq T} \mathcal{E}_{w_0}(n,t,s) \\
&\leq \frac{n+1}{\nu}\sup_{s\leq t\leq T} \mathcal{E}_{w_0}(n-1,t,s) + C\sup_{s\leq t\leq T}\int_s^t (r-s)^n \|w_{s,r}\|^{2n}\|w_{s,r}\|_1^{2(n+1)}dr + C_n + \sup_{t\geq s} N_n(t) \\
&\leq \frac{n+1}{\nu}\sup_{s\leq t\leq T} \mathcal{E}_{w_0}(n-1,t,s) + \frac{C}{\nu}\sup_{s\leq t\leq T} \mathcal{E}_{w_0}(1,t,s)^n \sup_{s\leq t\leq T} \mathcal{E}_{w_0}(0,t,s)^{n+1} + C_n + \sup_{t\geq s} N_n(t).
\end{aligned}$$

Then the desired result follows by induction hypothesis and the same reasoning as above. □

Appendix B. Approximate Controllability and Topological Irreducibility

The approximate controllability of the Navier-Stokes system by a degenerate (low modes, or finite dimensional) force was first proved by Agrachev and Sarychev [1, 2] in the case when there is no fixed external force. It was later realized that their proof still works if one add an additional body force f, whether it is time dependent or not. However, since we cannot locate any existing literature regarding the proof for the case that f is time dependent, we supply a proof here. The idea is taken from [35], where the case when f is time independent was proved. When the fixed body force is time dependent, the system becomes non-autonomous, and the notion of semigroups in [35] needs to be replaced by the evolution solution operators that depend on the initial time. And modifications are needed to adapt the proofs in [35] to the current non-autonomous setting. The key ideas of scaling and saturation are exactly the same as that in [35].

Consider the controlled Navier-Stokes equation

$$\partial_t w(x,t) - \nu\Delta w(x,t) + B(\mathcal{K}w, w)(x,t) = f(x,t) + \sum_{k=1}^{d} c_k(t) g_k, \quad w(x,s) = w_0(x), \tag{B.1}$$

where $t \geq s \geq 0$ and $\{g_k\}_{k=1}^d$ is from (2.4). Denote the solution of the above equation by $w_{s,t}(w_0, c\cdot g)$, where $c = (c_1, c_2, \cdots, c_d) : \mathbb{R} \to \mathbb{R}^d$ is piecewise constant and $g = (g_1, g_2, \cdots, g_d) \in H^d$. The approximate controllability of equation (B.1) means that for any $v_1, v_2 \in H$, $t > s$ and $\varepsilon > 0$, there is a piecewise constant control $c : [s,t] \to \mathbb{R}^d$ such that

$$\|w_{s,t}(v_1, c \cdot g) - v_2\| < \varepsilon. \tag{B.2}$$

We assume the initial time $s = 0$ but keep in mind that the equation (B.1) is non-autonomous since $f(x,t)$ depends on t. Denote the solution operator of equation (B.1) by

$$\Phi_{s,t}^{c \cdot g} w_0 := w_{s,t}(w_0, c \cdot g).$$

It follows from the uniqueness that

$$\Phi_{s,t}^{c \cdot g} w_0 = \Phi_{u,t}^{c \cdot g} \Phi_{s,u}^{c \cdot g} w_0, \quad \forall s \leq u \leq t.$$

Also for $v_1, v_2 \in H$ define $R_{s,t}^{v_2} v_1$ as the ray starting from v_1

$$R_{s,t}^{v_2} v_1 = v_1 + (t-s) v_2, \quad s \leq t. \tag{B.3}$$

Let

$$F_0 = \{\Phi_{s,t}^{c \cdot g} : c \in \mathbb{R}^d, 0 \leq s < t\} \tag{B.4}$$

and define the accessibility sets

$$A_{F_0}(v, t_0, t) := \{\Phi_{t_{m-1},t}^{m} \cdots \Phi_{t_0,t_1}^{1} v : \Phi^{\ell} \in F_0 \text{ for } 1 \leq \ell \leq m \text{ and } 0 \leq t_0 < t_1 < \cdots < t_{m-1} < t\}. \tag{B.5}$$

Then equation (B.1) is approximate controllable (B.2) if $\overline{A_{F_0}(v,t)} = H$ for every $v \in H$ and $t > 0$. Here $\overline{A_{F_0}(v,t)}$ is the closure of $A_{F_0}(v,t)$ in H.

B.1 Two Scaling Estimates

The following two scaling limits play an important role in establishing the approximate controllability of system (B.1). The first one indicates that we can approach the set of points of the form $w_0 + tc \cdot g, t \geq 0$, in a very short amount of time. The second scaling limit shows that one can generate new directions by pushing the control directions obtained in the first scaling into the system through the nonlinear term, to approach points in the form $w_0 - tB(\mathcal{K}c \cdot g, c \cdot g), t \geq 0$. The approximate controllability will follow from iterating the two scaling arguments to generate a much richer collection of new directions and a saturating process that will be given in the next section.

Proposition B.1. *For any $w_0 \in H$, $c \in \mathbb{R}^d$, and $s, t > 0$, one has*

$$\lim_{\lambda \to +\infty} \left\| \Phi^{\lambda c \cdot g}_{s, s+\frac{t}{\lambda}} w_0 - R^{c \cdot g}_{0,t} w_0 \right\| = 0, \tag{B.6}$$

$$\lim_{\lambda \to +\infty} \left\| R^{-\lambda^2 c \cdot g}_{0,\frac{1}{\lambda}} \Phi^0_{s, s+\frac{t}{\lambda^2}} R^{\lambda^2 c \cdot g}_{0,\frac{1}{\lambda}} w_0 - R^{-B(\mathcal{K} c \cdot g, c \cdot g)}_{0,t} w_0 \right\| = 0. \tag{B.7}$$

Proof. Let π_N be the orthogonal projection on the set of Fourier modes $\{e_k\}$ with $|k| \leq N$ and $R_N = R^{c \cdot g}_{0,t} \pi_N w_0$, $w_\lambda = \Phi^{\lambda c \cdot g}_{s, s+\frac{t}{\lambda}} w_0$. Note that $S = w_\lambda - R_N$ satisfies the equation

$$\partial_t S = \frac{1}{\lambda} \left(\nu \Delta S - B(\mathcal{K} S, S) + f(s+t) + \nu \Delta R_N + B(\mathcal{K} R_N, R_N) - B(\mathcal{K} w_\lambda, R_N) - B(\mathcal{K} R_N, w_\lambda) \right).$$

Taking H inner product with $2S$, and using standard estimates for the nonlinear term, we have

$$\begin{aligned} \lambda \partial_t \|S\|^2 &\leq -2\nu \|S\|_1^2 + 2\|f\|_\infty \|S\| + 2\nu \|R_N\|_1 \|S\|_1 + C\|R_N\|_1^2 \|S\| + C\|w_\lambda\|_1 \|R_N\|_1 \|S\| \\ &\leq \|w_\lambda\|_1^2 \|S\|^2 + C(\|f\|_\infty^2 + \|R_N\|_1^4 + 1) \\ &\leq \|w_\lambda\|_1^2 \|S\|^2 + CN^4(\|w_0\|^4 + t^4 \|c \cdot g\|_1^4 + 1). \end{aligned}$$

It then follows from Gronwall's inequality that

$$\begin{aligned} \|S(t)\|^2 \leq & \|w_0 - \pi_N w_0\|^2 \exp\left(\int_0^t \frac{1}{\lambda} \|w_\lambda(r)\|_1^2 dr \right) \\ & + \frac{CN^4}{\lambda} \int_0^t (\|w_0\|^4 + r^4 \|c \cdot g\|_1^4 + 1) \exp\left(\int_r^t \frac{1}{\lambda} \|w_\lambda(\tau)\|_1^2 d\tau \right) dr. \end{aligned} \tag{B.8}$$

Standard energy estimates [48] yield

$$\int_0^t \frac{1}{\lambda} \|w_\lambda(r)\|_1^2 dr \leq C \left(\frac{1}{\lambda} + \|c \cdot g\| + \|w_0\| \right),$$

where $C := C(\|f\|_\infty, t)$ is independent of λ. Now choosing N large to make the first term in the sum in (B.10) small and then letting $\lambda \to \infty$ to make the second term small, we find that $\|S(t)\| \to 0$ as $\lambda \to \infty$. The limit (B.6) follows once we note that

$$\left\| \Phi^{\lambda c \cdot g}_{s, s+\frac{t}{\lambda}} w_0 - R^{c \cdot g}_{0,t} w_0 \right\| \leq \|S(t)\| + \|w_0 - \pi_N w_0\|.$$

The proof of (B.7) is similar. Let $\mathcal{R}_N = R^{-B(\mathcal{K} c \cdot g, c \cdot g)}_{0,t} \pi_N w_0$, $\mathcal{W}_\lambda = R^{-\lambda^2 c \cdot g}_{0,\frac{1}{\lambda}} \Phi^0_{s, s+\frac{t}{\lambda^2}} R^{\lambda^2 c \cdot g}_{0,\frac{1}{\lambda}} w_0$ and $\mathcal{S} = \mathcal{W}_\lambda - \mathcal{R}_N$. Note that $\mathcal{S}$ solves the equation

$$\begin{aligned} \partial_t \mathcal{S} =& \frac{1}{\lambda^2} \Big(\nu \Delta \mathcal{S} - B(\mathcal{K} \mathcal{S}, \mathcal{S}) + \nu \Delta \mathcal{R}_N + B(\mathcal{K} \mathcal{R}_N, \mathcal{R}_N) - B(\mathcal{K} \mathcal{W}_\lambda, \mathcal{R}_N) - B(\mathcal{K} \mathcal{R}_N, \mathcal{W}_\lambda) + f(s+t) \Big) \\ & - \frac{1}{\lambda} \Big(B(\mathcal{K} \mathcal{S}, c \cdot g) + B(\mathcal{K} c \cdot g, \mathcal{S}) + B(\mathcal{K} \mathcal{R}_N, c \cdot g) + B(\mathcal{K} c \cdot g, \mathcal{R}_N) \Big). \end{aligned} \tag{B.9}$$

Taking H inner product with $2\mathcal{S}$, we find through standard estimates on the nonlinear term that

$$
\begin{aligned}
\partial_t\|\mathcal{S}\|^2 \le& \frac{1}{\lambda^2}\Big(-2\nu\|\mathcal{S}\|_1^2 + 2\nu\|\mathcal{S}\|_1\|\mathcal{R}_N\|_1 + C\|\mathcal{S}\|\|\mathcal{R}_N\|_1^2 + C\|\mathcal{S}\|\|\mathcal{W}_\lambda\|_1\|\mathcal{R}_N\|_1 + \|f\|_\infty\|\mathcal{S}\|\Big) \\
&+ \frac{C}{\lambda}\Big(\|c\cdot g\|_2\|S\|^2 + \|c\cdot g\|_2\|\mathcal{R}_N\|_2\|S\|\Big) \\
\le& \frac{C}{\lambda^2}\Big(1 + \|\mathcal{R}_N\|_1^4 + \|f\|_\infty^2\Big) + \frac{1}{\lambda^2}\|\mathcal{S}\|^2\|\mathcal{W}_\lambda\|_1^2 + \frac{C}{\lambda}(1+\|\mathcal{R}_N\|_2^2)\|S\|^2 \\
\le& \frac{CN^4}{\lambda^2}\left(1 + t^4 + \|w_0\|^4\right) + \frac{1}{\lambda^2}\|\mathcal{S}\|^2\|\mathcal{W}_\lambda\|_1^2 + \frac{C}{\lambda}(1 + N^4\|w_0\|^2 + t^2)\|S\|^2 \\
\le& \left(\frac{1}{\lambda^2}\|\mathcal{W}_\lambda\|_1^2 + \frac{C(t)N^4}{\lambda}\right)\|\mathcal{S}\|^2 + \frac{C(t)N^4}{\lambda^2},
\end{aligned}
$$

where $C(t)$ is continuously increasing in t but does not depend on λ, N. Again by Gronwall's inequality, one has

$$
\begin{aligned}
\|\mathcal{S}(t)\|^2 \le& \|w_0 - \pi_N w_0\|^2 \exp\left(\int_0^t \left(\frac{1}{\lambda^2}\|\mathcal{W}_\lambda(r)\|_1^2 + \frac{C(r)N^4}{\lambda}\right) dr\right) \\
&+ \frac{N^4}{\lambda^2}\int_0^t C(r)\exp\left(\int_r^t \left(\frac{1}{\lambda^2}\|\mathcal{W}_\lambda(\tau)\|_1^2 + \frac{C(\tau)N^4}{\lambda}\right) d\tau\right) dr. \qquad \text{(B.10)}
\end{aligned}
$$

Note that $\mathcal{W}_\lambda$ satisfies the equation

$$
\begin{aligned}
\partial_t\mathcal{W}_\lambda = \frac{1}{\lambda^2}\Big(&\nu\Delta\mathcal{W}_\lambda - B(\mathcal{K}\mathcal{W}_\lambda, \mathcal{W}_\lambda) - \lambda B(\mathcal{K}\mathcal{W}_\lambda, c\cdot g) - \lambda B(\mathcal{K}c\cdot g, \mathcal{W}_\lambda) + \nu\Delta c\cdot g \\
&- \lambda^2 B(\mathcal{K}c\cdot g, c\cdot g) + f(s+t)\Big).
\end{aligned}
$$

Taking H inner product with $\mathcal{W}_\lambda$ and using standard estimates on the nonlinear term we find

$$
\begin{aligned}
\partial_t\|\mathcal{W}_\lambda\|^2 \le& \frac{1}{\lambda^2}\Big(-2\nu\|\mathcal{W}_\lambda\|_1^2 + \lambda C\|c\cdot g\|_2\|\mathcal{W}_\lambda\|^2 + \nu\|c\cdot g\|_2\|\mathcal{W}_\lambda\| + \lambda^2 C\|c\cdot g\|_2^2\|\mathcal{W}_\lambda\| + \|f\|_\infty\|\mathcal{W}_\lambda\|\Big) \\
\le& \frac{-2\nu}{\lambda^2}\|\mathcal{W}_\lambda\|_1^2 + \frac{C}{\lambda^2}(\lambda^2 + \lambda + 1)\|\mathcal{W}_\lambda\|^2. \qquad \text{(B.11)}
\end{aligned}
$$

Gronwall's inequality implies that

$$
\|\mathcal{W}_\lambda(t)\|^2 \le e^{\frac{Ct}{\lambda^2}(\lambda^2+\lambda+1)}\|w_0\|^2.
$$

Using this estimate and integrating (B.11), it follows that

$$
\frac{1}{\lambda^2}\int_0^t \|\mathcal{W}_\lambda(r)\|_1^2 dr \le \|w_0\|^2 + \frac{C}{\lambda^2}(\lambda^2 + \lambda + 1)\int_0^t e^{\frac{Cr}{\lambda^2}(\lambda^2+\lambda+1)}\|w_0\|^2 dr,
$$

which remains bounded as $\lambda \to \infty$ with other parameters remaining fixed. Therefore we can choose N large but less than λ to make the first term in the sum in (B.10) small and then choose some $\lambda_0 > N$ such that the second term is small for all $\lambda > \lambda_0$, which implies

$$
\lim_{\lambda\to\infty}\|\mathcal{S}(t)\| = 0.
$$

The conclusion of (B.7) then follows once we note that

$$\left\| R^{-\lambda^2 c\cdot g}_{0,\frac{1}{\lambda}} \Phi^0_{s,s+\frac{t}{\lambda^2}} R^{\lambda^2 c\cdot g}_{0,\frac{1}{\lambda}} w_0 - R^{-B(\mathcal{K}c\cdot g, c\cdot g)}_{0,t} w_0 \right\| \le \|S(t)\| + \|w_0 - \pi_N w_0\|.$$

The proof is complete. □

B.2 Saturation

The saturation argument was introduced in [35] to deal with the multiple time scales when iteratively using the previous two scaling arguments. Recall the definition of the set F_0 as in (B.3). Define the time relaxed set of accessible points as

$$A_{F_0}(v, t_0, \le t) = \{\Phi^m_{t_{m-1},t_m} \cdots \Phi^1_{t_0,t_1} v : \Phi^\ell \in F_0 \text{ for } 1 \le \ell \le m \text{ and } 0 \le t_0 < \cdots < t_{m-1} < t_m \le t\}. \tag{B.12}$$

Let $X_0 = \{c \cdot g : c \in \mathbb{R}^d\}$ and define inductively for $k \ge 1$

$$X_k = \operatorname{span}\Big\{ X_{k-1} \cup \{B(\mathcal{K}\mathfrak{g}, \mathfrak{g}) : \mathfrak{g} \in X_{k-1}\} \Big\}, \text{ and } X_\infty = \cup_{k\ge 1} X_k. \tag{B.13}$$

Also let $G_0 = \Big\{ R^{\mathfrak{g}}_{s,t} : \mathfrak{g} \in X_0, s \le t \Big\} \cup F_0$, where $R^{\mathfrak{g}}_{s,t}$ is defined as in (B.3). Then define for $k \ge 1$

$$G_k = \Big\{ R^{\mathfrak{g}}_{s,t} : \mathfrak{g} \in X_k, s \le t \Big\} \cup F_0. \tag{B.14}$$

Let $S = \{\Psi : \Psi \in G_k \text{ for some } k \text{ or } \Psi \in F_0\}$. Given $F, G \in S$, G is said to subsume F, denoted by $F \preccurlyeq G$, if

$$\overline{A_F(v, t_F, \le t_F + t)} \subset \overline{A_G(v, t_G, \le t_G + t)}, \text{ for all } v \in H, t > 0, t_F, t_G \ge 0.$$

They are called equivalent if both $F \preccurlyeq G$ and $G \preccurlyeq F$, which we denote by $F \sim G$. The following lemma taken from [35] gives a useful characterization of subsuming relations. For any sequence of reals $t_i > 0$, we denote $t^{(k)} := \sum_{i=1}^k t_i$ and make the convention that $t^{(0)} = 0$.

Lemma B.2. *Let $F, G \subset S$, then $F \preccurlyeq G$ if and only if for any given $\Psi \in F$, $v \in H$, and $\varepsilon > 0, t > 0, t_F, t_G \ge 0$, there exists $\Phi^1, \cdots, \Phi^m \in G$ and positive times t_i such that $t^{(m)} \le t$ and*

$$\left\| \Phi^m_{t_G + t^{(m-1)}, t_G + t^{(m)}} \cdots \Phi^1_{t_G, t_G + t^{(1)}} v - \Psi_{t_F, t_F + t} v \right\| < \varepsilon. \tag{B.15}$$

Furthermore, for any family $F^i \subset S$ such that $F^i \preccurlyeq G$ for each i, one has $G \sim \cup_i F^i \cup G$.

Proof. If $F \preccurlyeq G$ then (B.15) follows from the definition. We now assume that the characterization (B.15) is true and to show $A_F(v, t_F, \le t_F + t) \subset \overline{A_G(v, t_G, \le t_G + t)}$ for any $v \in H$, $t > 0$, and $t_G, t_F \ge 0$. Let $u \in A_F(v, t_F, \le t_F + t)$, then there are $\Phi^1, \cdots, \Phi^m \in G$ and $t_i > 0, i = 1, \cdots, m$

with $t^{(m)} \le t$ such that

$$u = \Phi^m_{t_F+t^{(m-1)},t_F+t^{(m)}} \cdots \Phi^1_{t_F,t_F+t^{(1)}} v =: \prod_{i=1}^m \Phi^i_{t_F+t^{(i-1)},t_F+t^{(i)}} v.$$

The proof proceeds by induction on $m \ge 1$. If $m = 1$, then by (B.15), for any $\varepsilon > 0$, $t_G \ge 0$, there exist $\Psi^1, \cdots, \Psi^n \in G$ and positive times $\tau_i, 1 \le i \le n$ with $\tau^{(n)} \le t^{(1)}$ such that

$$\left\| \prod_{i=1}^n \Psi^i_{t_G+\tau^{(i-1)},t_G+\tau^{(i)}} v - u \right\| = \left\| \prod_{i=1}^n \Psi^i_{t_G+\tau^{(i-1)},t_G+\tau^{(i)}} v - \Phi_{t_F,t_F+t^{(1)}} \right\| < \varepsilon.$$

Hence $u \in \overline{A_G(v, t_F, \le t_F + t^{(1)})}$. Suppose that $m \ge 2$. By continuity of $\Phi^m_{t_F+t^{(m-1)},t_F+t^{(m)}}$, one has the existence of $\delta > 0$ such that for $w \in H$, if

$$\left\| w - \prod_{i=1}^{m-1} \Phi^i_{t_F+t^{(i-1)},t_F+t^{(i)}} v \right\| < \delta,$$

then

$$\begin{aligned} \left\| \Phi^m_{t_F+t^{(m-1)},t_F+t^{(m)}} w - u \right\| &= \left\| \Phi^m_{t_F+t^{(m-1)},t_F+t^{(m)}} w - \Phi^m_{t_F+t^{(m-1)},t_F+t^{(m)}} \prod_{i=1}^{m-1} \Phi^i_{t_F+t^{(i-1)},t_F+t^{(i)}} v \right\| \\ &< \varepsilon/2. \end{aligned} \tag{B.16}$$

By the induction hypothesis, we know that

$$\prod_{i=1}^{m-1} \Phi^i_{t_F+t^{(i-1)},t_F+t^{(i)}} v \in \overline{A_G(v, t_G, \le t_G + t^{(m-1)})}.$$

So there exist $\overline{\Psi}^1, \cdots, \overline{\Psi}^K \in G$ and positive times $r_i, 1 \le i \le K$ with $r^{(K)} \le t^{(m-1)}$ such that

$$\left\| \prod_{i=1}^K \overline{\Psi}^i_{t_G+r^{(i-1)},t_G+r^{(i)}} v - \prod_{i=1}^{m-1} \Phi^i_{t_F+t^{(i-1)},t_F+t^{(i)}} v \right\| < \delta. \tag{B.17}$$

Also from the characterization (B.15), one has the existence of $\underline{\Psi}^1, \cdots, \underline{\Psi}^J \in G$ and positive times $s_i, 1 \le i \le J$ with $s^{(J)} \le t^{(m)} - t^{(m-1)}$ such that

$$\left\| \prod_{i=1}^J \underline{\Psi}^i_{t_G+r^{(K)},t_G+r^{(K)}+s^{(i)}} \prod_{i=1}^K \overline{\Psi}^i_{t_G+r^{(i-1)},t_G+r^{(i)}} v - \Phi^m_{t_F+t^{(m-1)},t_F+t^{(m)}} \prod_{i=1}^K \overline{\Psi}^i_{t_G+r^{(i-1)},t_G+r^{(i)}} v \right\| < \varepsilon/2.$$

By setting $r^{(K+i)} = r^{(K)} + s^{(i)}$ and $\overline{\Psi}^{K+i} = \underline{\Psi}^i$ for $1 \le i \le J$, we have $r^{(J+K)} \le t^{(m)}$ and

$$\left\| \prod_{i=1}^{J+K} \overline{\Psi}^i_{t_G+r^{(i-1)},t_G+r^{(i)}} v - \Phi^m_{t_F+t^{(m-1)},t_F+t^{(m)}} \prod_{i=1}^K \overline{\Psi}^i_{t_G+r^{(i-1)},t_G+r^{(i)}} v \right\| < \varepsilon/2.$$

Now by the triangle inequality, estimates (B.16) and (B.17),

$$\begin{aligned}\left\|\prod_{i=1}^{J+K}\overline{\Psi}^{i}_{t_G+r^{(i-1)},t_G+r^{(i)}}v-u\right\| &\le \left\|\prod_{i=1}^{K}\overline{\Psi}^{i}_{t_G+r^{(i-1)},t_G+r^{(i)}}v-u\right\| \\ &\quad+\left\|\prod_{i=1}^{J+K}\overline{\Psi}^{i}_{t_G+r^{(i-1)},t_G+r^{(i)}}v-\Phi^{m}_{t_F+t^{(m-1)},t_F+t^{(m)}}\prod_{i=1}^{K}\overline{\Psi}^{i}_{t_G+r^{(i-1)},t_G+r^{(i)}}v\right\| \\ &< \varepsilon/2+\varepsilon/2=\varepsilon.\end{aligned}$$

Hence $u\in\overline{A_G(v,t_G,\le t_G+t)}$ and $F\preccurlyeq G$ as desired.

If $F^i\subset S$ such that $F^i\preccurlyeq G$ for each i, then $G\sim\cup_i F^i\cup G$ by noting that $G\preccurlyeq\cup_i F^i\cup G$ since G is a subset and $\cup_i F^i\cup G\preccurlyeq G$ by characterization (B.15). The proof is complete. □

Corollary B.3. *For G_k and X_k as defined in* (B.14) *and* (B.13)*, we have*

$$\left\{R^{\mathfrak{g}}_{s,t}:\mathfrak{g}\in X_\infty, s\le t\right\}\preccurlyeq F_0. \tag{B.18}$$

Proof. Note that the ray semigroups $R^{\mathfrak{g}}_{s,t}$ as defined in (B.3) are time homogeneous $R^{\mathfrak{g}}_{s,t}=R^{\mathfrak{g}}_{0,t-s}$. Hence we are free to choose any initial time we want. Now it follows from the scaling (B.6) and Lemma B.2 that $G_0\preccurlyeq F_0$. It also follows from the second scaling (B.7) and Lemma B.2 that $G_k\preccurlyeq G_{k-1}$ for each $k\ge 1$. Therefore $G_k\preccurlyeq F_0$ for each $k\ge 0$. This implies that $\cup_{k\ge 0}G_k\preccurlyeq F_0$. Hence

$$\left\{R^{\mathfrak{g}}_{s,t}:\mathfrak{g}\in X_\infty, s\le t\right\}\preccurlyeq\left\{R^{\mathfrak{g}}_{s,t}:\mathfrak{g}\in X_\infty, s\le t\right\}\cup F_0\preccurlyeq F_0.$$

The proof is complete. □

The following lemma, which is essentially Lemma 3.7 in [35], allows us to pass the time relaxed accessible points $A_{F_0}(v,t_0,\le t)$ in (B.12) to exact time accessible points $A_{F_0}(v,t_0,t)$ as in (B.5).

Lemma B.4. *Suppose $F\subset S$ and $V\subset H$ is open with the property that*

$$V\subset\overline{A_F(v,t_0,\le t_0+t)},\ \text{for all } v\in V, t_0\ge 0, t>0.$$

Then

$$V\subset\overline{A_F(v,t_1,t_1+t)},\ \text{for all } v\in V, t_1\ge 0, t>0.$$

Proof. For simplicity we assume that $t_1=0$ and show that $u\in\overline{A_F(v,0,t)}$ for any given $u,v\in V$ and $t>0$. Fix $\varepsilon>0$ such that $B(u,\varepsilon)\subset V$. Pick any $\Phi\in F\subset S$. By continuity we can choose

$0 < \varepsilon' < \varepsilon$ such that the first exit time

$$T := \inf_{w \in B(u,\varepsilon')} \Big\{ \inf \big\{ t > 0 : \|\Phi_{t_0,t_0+t} w - u\| > \varepsilon \big\} \Big\} > 0. \tag{B.19}$$

T is independent of any initial time $t_0 > 0$ due to the fact that Φ_{t_0,t_0+t} is either a ray semigroup as in (B.3) which is time homogeneous, or $\Phi_{t_0,t_0+t} w$ is a solution to (B.1) (with a time independent $c(t)$) that is continuous in (t, w) independent of t_0, which follows from the uniform boundedness of $f(x, t)$, i.e., $\sup_{t\in\mathbb{R}} \|f(t)\| < \infty$. By assumption, there exist $\Phi^{0,1}, \cdots, \Phi^{0,m_0}$ and positive times $r_{0,i}, 1 \le i \le m_0$ with $r_0^{(m_0)} \le t$ such that

$$u_0 := \prod_{i=1}^{m_0} \Phi^{0,i}_{r_0^{(i-1)}, r_0^{(i)}} v \in B(u, \varepsilon'),$$

where $r_0^{(k)} = \sum_{i=1}^{k} r_{0,i}$. If $r_0^{(m_0)} + T \ge t$, then in view of (B.19) we see that

$$\Phi_{r_0^{(m_0)},t} \prod_{i=1}^{m_0} \Phi^{0,i}_{r_0^{(i-1)}, r_0^{(i)}} v = \Phi_{r_0^{(m_0)}, r_0^{(m_0)}+t-r_0^{(m_0)}} \prod_{i=1}^{m_0} \Phi^{0,i}_{r_0^{(i-1)}, r_0^{(i)}} v \in B(u, \varepsilon),$$

since $t - r_0^{(m_0)} \le T$. This shows $u \in \overline{A_F(v, 0, t)}$ as desired.

If $r_0^{(m_0)} + T < t$, then there exists an integer $n \ge 1$ such that

$$r_0^{(m_0)} + nT < t \le r_0^{(m_0)} + (n+1)T.$$

Let $\delta := t - r_0^{(m_0)} - nT \le T$. Note that $u_0 \in B(u, \varepsilon') \subset V$, hence by (B.19) we find that

$$w_1 := \Phi_{r_0^{(m_0)}, r_0^{(m_0)}+T} u_0 \in B(u, \varepsilon) \subset V.$$

By assumption there exist $\Phi^{1,1}, \cdots, \Phi^{1,m_1} \in F$ and positive times $r_{1,i}, 1 \le i \le m_1$ with $r_1^{(m_1)} \le \delta/n$ such that

$$u_1 := \prod_{i=1}^{m_1} \Phi^{1,i}_{r_0^{(m_0)}+T+r_1^{(i-1)}, r_0^{(m_0)}+T+r_1^{(i)}} w_1 \in B(u, \varepsilon').$$

Iterating this process, for $2 \le k \le n$, define

$$w_k := \Phi_{\sum_{j=0}^{k-1} r_j^{(m_j)}+(k-1)T, \sum_{j=0}^{k-1} r_j^{(m_j)}+kT} u_{k-1} \in B(u, \varepsilon) \subset V.$$

Then by assumption, there are $\Phi^{k,1}, \cdots, \Phi^{k,m_k} \in F$ and positive times $r_{k,i}, 1 \le i \le m_k$ with $r_k^{(m_k)} \le \delta/n$ such that

$$u_k := \prod_{i=1}^{m_k} \Phi^{k,i}_{\sum_{j=0}^{k-1} r_j^{(m_j)}+kT+r_k^{(i-1)}, \sum_{j=0}^{k-1} r_j^{(m_j)}+kT+r_k^{(i)}} w_k \in B(u, \varepsilon').$$

Now observe that $u_n \in B(u, \varepsilon')$, so by (B.19) one has

$$\Phi_{\sum_{j=0}^{n} r_j^{(m_j)}+nT, t} u_n \in B(u, \varepsilon),$$

since

$$0 \le t - \left(\sum_{j=0}^{n} r_j^{(m_j)} + nT \right) < T.$$

This shows that $u \in \overline{A_F(v, 0, t)}$ and completes the proof. □

B.3 Approximate Controllability and Topological Irreducibility

In this section we show how the scaling and saturation results in the previous two sections yield the approximate controllability of the equation (B.1), which in turn implies the topological irreducibility of the system (2.5).

For each $t > 0$, let $\Omega_t = \{\omega \in C([0, t], \mathbb{R}^d) : \omega(0) = 0\}$ equipped with the supremum norm, be the restricted Wiener space. For any $T > 0$ and $V \in \Omega_T$, it can be shown that the solution $w_{0,t}(w_0, V)$ of the equation

$$w(x, t) - w_0(x) - \int_0^t \nu \Delta w(x, t)dt + \int_0^t B(\mathcal{K}w, w)(x, t)dt = \int_0^t f(x, t)dt + g \cdot V, \tag{B.20}$$

where $g = (g_1, \cdots, g_d) \in H^d$, is continuous in V. And if we replace $c(t)$ in (B.1) with $\partial_t \int_0^t c(t)dt$ and set initial data as $w(0, x) = w_0(x)$, then the solution $\Phi_{0,t}^{c \cdot g} w_0 = w_{0,t}(w_0, V)$ with $V(t) = \int_0^t c(t)dt$.

Proposition B.5 (Approximate Controllability)**.** *If $A_\infty = H$ (see (3.1) for the definition), then for any $\varepsilon > 0$, $w_0, w_1 \in H$ and $t > 0$, there exists $V \in \Omega_t$ such that*

$$\|w_{0,t}(w_0, V) - w_1\| < \varepsilon. \tag{B.21}$$

Proof. It follows from [35] that $A_\infty = H$ implies $\overline{X_\infty} = H$, where X_∞ is given as in (B.13). It then follows from Corollary B.3 that

$$\overline{A_F(v, t_0, \le t_0 + t)} = H, \text{ for all } v \in H, t_0 \ge 0, t > 0.$$

Lemma B.4 then yields

$$\overline{A_F(v, 0, t)} = H, \text{ for all } v \in H, t > 0.$$

In view of the Definition B.5 and B.4, we find that for given $w_0, w_1 \in H$ and $\varepsilon > 0$, there are $\Phi^1, \cdots, \Phi^m \in F_0$ and $0 = t_0 < t_1 < t_2 < \cdots < t_{m-1} < t_m = t$ such that

$$\|\Phi_{t_{m-1},t_m}^{c_m} \cdots \Phi_{t_0,t_1}^{c_1} w_0 - w_1\| < \varepsilon,$$

which is equivalent to saying that the solution $\Phi_{0,t}^{c\cdot g} w_0$ with $c(t) = \sum_{i=1}^m c_i \mathbb{I}_{[t_{i-1},t_{i+1})}(t)$ satisfies

$$\|\Phi_{0,t}^{c\cdot g} w_0 - w_1\| < \varepsilon.$$

Setting $V(t) = \int_0^t c(t)dt$, we arrive at (B.21). □

A direct consequence is the following topological irreducibility.

Corollary B.6 (Topological Irreducibility)**.** *If $A_\infty = H$, then the transition operator $\mathcal{P}_{0,t}$ (see* (2.6)*) of equation* (2.5) *satisfies*

$$\mathcal{P}_{0,t}(w_0, B_\delta(w_1)) > 0,$$

for all $w_0, w_1 \in H$, $\delta > 0$ and $t > 0$.

For a proof of the corollary, see Lemma 4.7 in [35].

Bibliography

[1] Agrachev, A. A., Sarychev, A. V.: Navier–Stokes equations: controllability by means of low modes forcing. Journal of Mathematical Fluid Mechanics, 7(1), 108-152 (2005)

[2] Agrachev, A. A., Sarychev, A. V.: Controllability of 2D Euler and Navier-Stokes equations by degenerate forcing. Communications in mathematical physics, 265(3), 673-697 (2006)

[3] Bricmont, J., Kupiainen, A., Lefevere, R.: Ergodicity of the 2D Navier-Stokes equations with random forcing. Comm. Math. Phys. 224(1), 65–81 (2001)

[4] Bricmont, J., Kupiainen, A., Lefevere, R.: Exponential mixing of the 2D stochastic Navier-Stokes dynamics. Comm. Math. Phys. 230(1), 87–132 (2002)

[5] Bogachev, V. I., Krylov, N. V., Röckner, M., Shaposhnikov, S. V.: Fokker-Planck-Kolmogorov Equations (Vol. 207). American Mathematical Society (2015)

[6] Babin, A.V.: Attractors of Navier-Stokes equations. Handbook of mathematical fluid dynamics, Vol. II, 169–222, North-Holland, Amsterdam (2003)

[7] Bakhtin, Y., Mattingly, J.C.: Malliavin calculus for infinite-dimensional systems with additive noise. J. Funct. Anal. 249(2), 307–353 (2007)

[8] Butkovs ky, O., Kulik, A., Scheutzow, M.: Generalized couplings and ergodic rates for SPDEs and other Markov models. Ann. Appl. Probab. 30(1), 1–39 (2020)

[9] Constantin, P., Glatt-Holtz, N., Vicol, V.: Unique ergodicity for fractionally dissipated, stochastically forced 2D Euler equations. Comm. Math. Phys. 330(2), 819–857 (2014)

[10] Crauel, H., Debussche, A., Flandoli, F.:Random attractors. J. Dynam. Differential Equations. 9(2), 307–341 (1997)

[11] Cheban, D., Schmalfuss, B.: Invariant manifolds, global attractors, almost automrphic and almost periodic solutions of non-autonomous differential equations, J. Math. Anal. Appl. 340(1), 374–393 (2008)

[12] Chen, M.: From Markov chains to non-equilibrium particle systems. World Scientific (2004)

[13] Constantin, P., Foias, P.: Navier-Stokes Equations. University of Chicago Press, Chicago (1988)

[14] Chepyzhov, V.V., Vishik, M.I.: Attractors for Equations of Mathematical Physics, volume 49 of AMS Colloquium Publications. AMS (2002)

[15] Chepyzhov, V.V., Vishik, M.I.: Non-autonomous 2D Navier–Stokes system with a simple global attractor and some averaging problems. El. J. ESAIM: COCV, 8, 467–487 (2002)

[16] Da Prato, G., Debussche, A.: 2D stochastic Navier–Stokes equations with a time-periodic forcing term. J. Dynam. Differential Equations. 20(2), 301-335 (2008)

[17] Dobrushin, R.: Central limit theorems for non-stationary Markov chains I,II. Theory of Probab. and its Appl. 1,65-80& 329-383 (1956)

[18] Doeblin,W.: Sur les propriétés asymptotiques de mouvement régis par certains types de chaînes simples. Bull. Math. Soc. Roum. Sci. 39(1), 57-115 (1937)

[19] Da Prato, G., Röckner, M.: Dissipative stochastic equations in Hilbert space with time dependent coefficients. Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl. 17(4), 397-403 (2006)

[20] Dolgopyat, D., Sarig, O.: Local limit theorems for inhomogeneous Markov chains. (2021)

[21] Dynkin, E.B.: Markov processes, volume 121 of Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, Berlin-Göttingen-Heidelberg (1965)

[22] Dynkin, E.B.: Entrance and exit spaces for a Markov process, volume 2 of Actes du Congrès International des Mathématiciens. Gauthier-Villars, Paris (1971)

[23] Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions, volume 152 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge (2014)

[24] E, W., Mattingly, J. C., Sinai, Y. G.: Gibbsian dynamics and ergodicity for the stochastically forced Navier-Stokes equation. Comm. Math. Phys. 224(1), 83–106 (2001)

[25] Eckmann J.P., Hairer, M.: Uniqueness of the invariant measure for a stochastic PDE driven by degenerate noise. Comm. Math. Phys. 219(3), 523–565 (2001)

[26] Flandoli, F.: Dissipativity and invariant measures for stochastic Navier–Stokes equations. NoDEA Nonlinear Differ. Equ. Appl. 1(4), 403–423 (1994)

[27] Flandoli, F., Maslowski B.: Ergodicity of the 2-D Navier-Stokes equation under random perturbations. Comm. Math. Phys. 172(1), 119–141 (1995)

[28] Ferrario, B.: Ergodic results for stochastic Navier-Stokes equation. Stochastics and Stochastics Reports. 60(3-4), 271–288 (1997)

[29] Földes, J., Glatt-Holtz, N., Richards, G., Thomann, E.: Ergodic and mixing properties of the Boussinesq equations with a degenerate random forcing. J. Funct. Anal. 269(8), 2427-2504 (2015)

[30] Foias, C., Manley, O., Rosa, R., Temam, R.: Navier-Stokes equations and turbulence (Vol. 83). Cambridge University Press (2001)

[31] Feng, C., Qu, B., Zhao, H. Random quasi-periodic paths and quasi-periodic measures of stochastic differential equations. Journal of Differential Equations, 286, 119-163 (2021)

[32] Frisch, U.: Turbulence: The Legacy of A. N. Kolmogorov. Cambridge University Press, Cambridge (1995)

[33] Feng, C., Zhao, H., Zhou, B.: Pathwise random periodic solutions of stochastic differential equations. J. Differential Equations. 251(1), 119–149 (2011)

[34] Glatt-Holtz, N., Mattingly, J.C.: Richards, G.: On unique ergodicity in nonlinear stochastic partial differential equations. J. Stat. Phys. 166(3-4), 618-649 (2017)

[35] Glatt-Holtz, N. E., Herzog, D. P., Mattingly, J. C.: Scaling and saturation in infinite-dimensional control problems with applications to stochastic partial differential equations. Annals of PDE, 4(2), 1-103 (2018)

[36] Gordin, M. I.: The central limit theorem for stationary processes. Dokl. Akad. Nauk SSSR. 188(4), 739–741 (1969)

[37] Harris, T. E.: The existence of stationary measures for certain Markov processes, Volume 2 of Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Univ. California Press, Berkeley. 113-124 (1956)

[38] Hairer, M.: Exponential mixing properties of stochastic PDEs through asymptotic coupling. Probab. Theory Related Fields. 124(1), 345–380 (2002)

[39] Hall, P., Heyde, C. C.: Martingale limit theory and its application. Academic press. (2014)

[40] Hairer, M., Mattingly, J.C.: Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing. Ann. of Math. (2) 164 (3), 993–1032 (2006)

[41] Hairer, M., Mattingly, J.C.: Spectral gaps in Wasserstein distances and the 2D stochastic Navier–Stokes equations. Annal. Probab. 36(6), 2050-2091 (2008)

[42] Hairer, M., Mattingly, J.C.: A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs. Electron. J. Probab. 16(23), 658-738 (2011)

[43] Klein, S., Liu, X., Melo, A.: Uniform convergence rate for Birkhoff means of certain uniquely ergodic toral maps. Ergodic Theory and Dynamical Systems. 1-26 (2019).

[44] Kipnis, C.; Varadhan, S.R.S.: Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. Comm. Math. Phys. 104(1), 1–19 (1986)

[45] Kuksin, S., Shirikyan, A.: Stochastic dissipative PDEs and Gibbs measures. Comm. Math. Phys. 213(2), 291–330 (2000)

[46] Kuksin, S., Shirikyan, A.: A coupling approach to randomly forced nonlinear PDE's. I. Comm. Math. Phys. 221(2), 351–366 (2001)

[47] Kuksin, S.: Ergodic theorems for 2D statistical hydrodynamics. Rev. Math. Phys. 14(6), 585–600 (2002)

[48] Kuksin, S., Shirikyan, A.: Mathematics of Two-Dimensional Turbulence, Volum 194 of Cambridge Tracts in Math. Cambridge University Press (2012)

[49] Kulik, A., Ergodic Behavior of Markov Processes: With Applications to Limit Theorems. Vol. 67 of Walter de Gruyter GmbH & Co KG (2017)

[50] Komorowski, T., Walczuk, A.: Central limit theorem for Markov processes with spectral gap in the Wasserstein metric. Stochastic Process. Appl. 122(5), 2155–2184 (2012)

[51] Marchioro, C.: An example of absence of turbulence for any Reynolds number: II. Communications in mathematical physics, 108(4), 647-651(1987)

[52] Mattingly, J. C.: The stochastically forced Navier-Stokes equations: energy estimates and phase space contraction. Ph.D.thesis, Princeton University (1998)

[53] Mattingly, J. C.: Ergodicity of 2D Navier-Stokes equations with random forcing and large viscosity. Comm. Math. Phys. 206(2), 273–288 (1999)

[54] Mattingly, J.C.: The dissipative scale of the stochastics Navier-Stokes equation: regularization and analyticity. J. Statist. Phys. 108(5-6), 1157–1179 (2002)

[55] Mattingly, J. C.: Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics. Comm. Math. Phys. 230(3), 421–462 (2002)

[56] Mattingly, J. C., Suidan, T. M.: The small scales of the stochastic Navier-Stokes equations under rough forcing. J. Stat. Phys. 118(1-2), 343–364 (2005)

[57] Mattingly, J. C., Pardoux, É.: Malliavin calculus for the stochastic 2d Navier–Stokes equation. Commun. Pure Appl. Math. 59(12), 1742–1790 (2006)

[58] Masmoudi, N., Young, L.-S.: Ergodic theory of infinite dimensional systems with applications to dissipative parabolic PDEs. Comm.Math. Phys. 227(3), 461–481 (2002)

[59] Nualart, D.: The Malliavin calculus and related topics. Probability and its Applications (New York), 2nd edn. Springer, Berlin (2006)

[60] Punshon-Smith, S., Bedrossian, J.: Chaos in stochastic 2d Galerkin-Navier-Stokes. arXiv preprint arXiv:2106.13748. (2021)

[61] Shirikyan, A.: Law of large numbers and central limit theorem for randomly forced PDE's. Probab. Theory Related Fields 134 (2), 215–247 (2006)

[62] Sethuraman, S., Varadhan, S.: A martingale proof of Dobrushin's theorem for non-homogeneous Markov chains. Electronic Journal of Probability. 10, 1221-1235 (2005)

[63] Stroock, D. W. Elements of Stochastic Calculus and Analysis. Springer (2018)

[64] Vishik, M. I., Fursikov, A. V.: Mathematical problems in statistical hydrodynamics. Kluwer, Dordrecht (1988)

[65] Villani, C.: Optimal Transport, Old and New, Volume 338 of Springer Science and Business Media (2008)