Intermittency in Turbulent Diffusion Models with a Mean Gradient

Andrew J Majda and Xin T Tong

Department of Mathematics, and Center for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA

E-mail: jonjon@cims.nyu.edu, tong@cims.nyu.edu

Abstract. This paper provides a rigorous, self contained analysis of the intermittent behavior of turbulent diffusion models with a mean gradient. The intermittency can be described as large spikes randomly occurring in the time sequence of a passive tracer or exponential like fat tails in the probability density function. This type of passive tracer intermittency is subtle and occurs without any positive Lyapunov exponents in the system. Observations of such passive tracers in nature also show such intermittency. By exploiting an intrinsic conditional Gaussian structure, the enormous fluctuation in conditional variance of the passive tracer is found to be the source of intermittency in these models. An intuitive physical interpretation of such enormous fluctuation can be described through the random resonance between Fourier modes of the turbulent velocity field and the passive tracer. This intuition can be rigorously proved in a long time slow varying limit, where the limiting distribution of the passive tracer is computed through an integral formula. This leads to rigorous predictions of various types of intermittency. Numerical experiments are conducted in different dynamical regimes to verify and supplement all the theoretical results. All the proofs in this paper are elementary and essentially self contained.

AMS classification scheme numbers: 37N10, 76F25, 76E30, 62G32

Submitted to: *Nonlinearity*

1. Introduction

Turbulent diffusion is the transportation of the mass of a substance, also called a passive tracer, through the joint effect of turbulence advection and diffusion. Its application ranges from the spread of hazardous plumes and mixing properties of turbulent combustion, to the dynamics of anthropogenic gas in climate change science [1, 2, 3]. There is often additional uniform damping in many applications in environmental science. The physical law of turbulent diffusion can be described through the dynamics of a passive tracer:

$$\frac{\partial T_t}{\partial t} + \vec{v} \cdot \nabla T = -d_T T_t + \kappa \Delta T, \tag{1.1}$$

where $\kappa > 0$ is molecular diffusion and the turbulence flow $\vec{v}$ is assumed to be incompressible. Statistical properties of solutions of (1.1) such as their energy spectrum, probability density of $T_t(x)$ (PDF) and spatial covariance, $\text{cov}(T_t(x), T_t(y))$, are all critical physical quantities of the underlying turbulent flow $\vec{v}$ [4, 5].

One key feature of turbulent diffusion with a mean background gradient is its intermittency. This can be described by large spikes randomly occurring in the time sequence of $T_t(x)$, or an exponential like heavy tail in the probability density. These phenomena are observed through laboratory observations such as classical Rayleigh-Benard convection experiments [6, 7, 8], observation data of the atmosphere [3] and numerical simulations of judiciously chosen idealized models [4, 9, 10, 11, 12, 5]. On the other hand, intermittency generally exists in various turbulence models [13, 14, 12, 15, 16, 17]. It is extremely important to understand intermittency, since such extreme events are crucial in many areas of environmental science as well as other disciplines. However, the rigorous study of the mechanisms behind intermittency is challenging.

This paper aims to give a simple but rigorous explanation of the source of intermittency in turbulent diffusion models with a mean gradient. In these models the two dimensional incompressible velocity field $\vec{v}$ is assumed to consist of a zonal cross sweep and a meridional random shear flow:

$$\vec{v_t}(x, y) = (u_t, v(x, t)).$$

The passive tracer field is assumed to have a background mean gradient:

$$T_t(x, y) = T'_t(x) + \alpha y.$$

Previous works [18, 19, 9, 10, 20, 5] have shown that this simplification preserves key features for various inertial range statistics of turbulent diffusion, including the intermittent phenomena described above. On the other hand, this simplification produces explicit formulations for the solutions which will be very convenient to apply stochastic analysis tools.

One important feature that will be repeatedly exploited in this paper is that the tracer field $T'_t(x)$ is conditionally Gaussian given the realization of the cross sweep $u_{s\le t}$. This will be derived carefully in Section 3.2 as Proposition 3.2. In other words, the conditional distribution of $T'_t(x)$ can be written as $\mathcal{N}(0, \Sigma(u_{s\le t}))$. Using the law of total expectation, the PDF of $T'_t(x)$ can be formally computed through a path based integral, and will be a Gaussian mixture of different variances. The key point here is the conditional variance $\Sigma(u_{s\le t})$ is a random variable that has enormous fluctuation, with peak value often being around 100 times the magnitude of its mean. This implies that the occasional realization of $u_{s\le t}$ that maximizes $\Sigma(u_{s\le t})$, will likely produces a spike in $T'_t(x)$ of 10 times the size of standard deviation, which is nearly impossible in a Gaussian prediction. Moreover, this enormous conditional variance dictates the tail of $T'_t(x)$'s distribution, which will be exponential like in the tail of 10^{-2}-10^{-5} range. This fits very well with the intermittent phenomenon described above.

A deeper question that naturally follows is when and why does the conditional variance reach its peak. By analyzing the integral formula, one possible cause is found as the hidden random resonance between Fourier modes of the shear flow v and passive tracer field T. This can be described as the moment when the imaginary phase speed of v and T are equal. In Section 4, this intuitive mechanism is rigorously proved in a long time slow varying limit of the original model, with the limiting conditional variance being a simple explicit function of u_t. With this explicit result, simple intuition can be developed on the cross sweep's effect over passive tracer T, while the PDF can be computed through a simple quadrature.

The remainder of this paper will be organized as follow. Section 2 formulates the turbulent diffusion model and its special solutions. Section 3 presents the fundamental properties of the turbulent diffusion models with detailed derivation and explanation, including geometric ergodicity and rigorous general bounds hinting at intermittency. In Section 4, the long time slow varying limit of turbulent diffusion models are presented with rigorous proof of the connection with intermittent random resonances. Numerical simulations with different dynamical regimes are conducted in Section 5 verifying previous mathematical claims and enhancing the intuition. Generalizations to general non-Gaussian cross sweeps and infinitely many Fourier modes are made in Section 6. A summary and discussion of our results is the final section of this paper. In order to maintain the focus of our discussion, some straightforward parts of the proofs are relegated to the Appendix.

2. Turbulent diffusion models with mean gradient formulation

In this paper, we assume the velocity field $\vec{v}$ is a special incompressible flow given by a stochastic zonal cross sweep and a meridional shear flow:

$$\vec{v}(x,t) = (u_t, v(x,t)).$$

We decompose the cross sweep u_t into its mean $\bar{u}$ and fluctuation: $u_t = \bar{u}+U_t$. For simplicity, U_t is modeled as a mean zero Ornstein-Uhlenbeck (O.U.) process, with dynamics that reads:

$$dU_t = -\gamma_U U_t \mathrm{d}t + \sigma_U \mathrm{d}W_t,$$

Here W_t is a standard real Wiener process. In Section 6.1, we will relax this O.U. formulation to general diffusion processes, but until then we will focus on this simple setting. The dynamics for the shear flow v can be described through a general formulation:

$$\frac{\partial v}{\partial t} = U_t R_1\left(\frac{\partial}{\partial x}\right)v + R_2\left(\frac{\partial}{\partial x}\right)v - \gamma_v\left(\frac{\partial}{\partial x}\right)v + \dot{W}_v(x,t). \tag{2.1}$$

Here γ_v is a positive definite linear operator that represents damping and dissipation. R_1, R_2 are linear operators that represent the internal effect of U_t upon v_t such as advection. They are defined through their image on the Fourier modes:

$$R_1\left(\frac{\partial}{\partial x}\right)e^{\mathrm{i}kx} = \mathrm{i}a_k e^{\mathrm{i}kx}, \quad R_2\left(\frac{\partial}{\partial x}\right)e^{\mathrm{i}kx} = \mathrm{i}b_k e^{\mathrm{i}kx}, \quad \gamma_v\left(\frac{\partial}{\partial x}\right)e^{\mathrm{i}kx} = \gamma_{v,k} e^{\mathrm{i}kx}.$$

Since (2.1) is linear, the Fourier expansion $v(x,t) = \sum_{k\in N} \hat{v}_{k,t} e^{\mathrm{i}kx}$ has for each wavenumber the following dynamics:

$$\mathrm{d}\hat{v}_{k,t} = -\gamma_{v,k}\hat{v}_{k,t}\mathrm{d}t + \mathrm{i}(a_k U_t + b_k)\hat{v}_{k,t}\mathrm{d}t + \sigma_{v,k}\mathrm{d}B_{k,t}.$$

Here $B_{k,t}$ are independent complex Wiener processes, i.e. $B_{k,t} = \frac{1}{\sqrt{2}}(B^1_{k,t} + \mathrm{i}B^2_{k,t})$ with $B^i_{k,t}$ being independent real Wiener processes. Thus the Gaussian random field $W_v(x,t)$ in (2.1) is explicitly given by $\sum_{k\in N} B_{k,t} e^{\mathrm{i}kx}$. On the other hand, in order to keep v real valued, we require the wavenumber set N consists of pairs of opposite wavenumbers, while $\hat{v}_{k,t} = \hat{v}^*_{-k,t}$, which can be enforced through the relation [20, 5]:

$$\gamma_{v,k} = \gamma_{v,-k}, \quad a_k = -a_{-k}, \quad b_k = -b_{-k}, \quad B_{k,t} = B^*_{-k,t}.$$

This simple formulation of cross sweep actually includes a wide range of turbulence models. Here are a few examples that are common in the literature:

(i) In a random cross sweep model, which is a stochastic version of the deterministic model in [10], $R_1 = R_2 = 0$ and $\gamma_{v,k}$ is the sum of dissipation and damping:

$$\gamma_{v,k} = d_v + \nu k^2, \quad a_k = b_k = 0.$$

(ii) In the engineering community, non-dispersive waves with selective damping [12] are commonly considered, they can be formulated using (2.1) with

$$\gamma_{v,k} = d_v + \nu k^2, \quad a_k = 0, \quad b_k = -ck.$$

(iii) Baroclinic Rossby waves without a mean flow [21, 12] can be formulated using (2.1) with

$$\gamma_{v,k} = d_v + F_s k^2, \quad a_k = 0, \quad b_k = \frac{\beta k}{k^2 + F}.$$

where F_s is the stratification constant and β is the constant for β-plane approximation.

(iv) β-plane Q-G baroclinic 1.5 layer flows [20, 5] can be formulated using (2.1) with

$$\gamma_{v,k} = d_v + \nu k^2, \quad a_k = \frac{-k^3}{k^2 + F}, \quad b_k = \frac{\beta k}{k^2 + F}.$$

Here $F = L_R^{-2}$ with L_R being the deformation radius of Rossby waves.

For detailed physical interpretation and explicit derivation of these parameters, we refer to the citations in each example. With the dynamics of the velocity field well described, we can now turn to the passive tracer field $T(t)$. Following the examples of [10, 5], we assume the turbulent tracer has a mean meridional gradient α:

$$T_t(x,y) = T_t(x) + \alpha y.$$

Inserting this relation into (1.1), the full physical model now reads:

$$\begin{aligned} \mathrm{d}U_t &= -\gamma_U U_t \mathrm{d}t + \sigma_U \mathrm{d}W_t; \\ \frac{\partial v}{\partial t} &= U_t R_1\left(\frac{\partial}{\partial x}\right) v + R_2\left(\frac{\partial}{\partial x}\right) v - \gamma_v\left(\frac{\partial}{\partial x}\right) v + \dot{W}_v(x,t); \\ \frac{\partial T_t}{\partial t} &+ (U_t + \bar{u})\frac{\partial T_t}{\partial x} = -d_T T_t + \kappa \Delta T_t - \alpha v_t. \end{aligned} \tag{2.2}$$

Since this equation is also linear for T_t, the Fourier expansion $T_t(x) = \sum_{k\in N} \widehat{T}_{k,t} e^{\mathrm{i}kx}$ has the following dynamics for each wavenumber:

$$\mathrm{d}\widehat{T}_{k,t} = -(\gamma_{T,k} + \mathrm{i}k(U_t + \bar{u}))\widehat{T}_{k,t}\mathrm{d}t - \alpha \hat{v}_{k,t}\mathrm{d}t,$$

where $\gamma_{T,k} = d_T + \kappa k^2$. In summary, an explicit solution of model (2.2) can be described through the Fourier modes:

$$\begin{aligned} \mathrm{d}U_t &= -\gamma_U U_t \mathrm{d}t + \sigma_U \mathrm{d}W_t; \\ \mathrm{d}\hat{v}_{k,t} &= (-\gamma_{v,k} + \mathrm{i}\omega_{v,k}(t))\hat{v}_{k,t}\mathrm{d}t + \sigma_{v,k}\mathrm{d}B_{k,t}; \\ \mathrm{d}\widehat{T}_{k,t} &= (-\gamma_{T,k} + \mathrm{i}\omega_{T,k}(t))\widehat{T}_{k,t}\mathrm{d}t - \alpha \hat{v}_{k,t}\mathrm{d}t; \\ \omega_{v,k}(t) &= a_k U_t + b_k, \quad \omega_{T,k}(t) = -k(U_t + \bar{u}). \end{aligned} \tag{2.3}$$

For simplicity in exposition, most part of this paper will assume there are only finitely many pairs of wavenumbers in the Fourier expansion, while a generalization to infinite dimensions is given in Section 6.2.

3. General mathematical properties

The turbulent diffusion model (2.2) we constructed is a stochastic process with many desirable mathematical properties, which are very useful for analysis and simulation purposes.

3.1. Geometric ergodicity

Geometric ergodicity guarantees that the distribution of a stochastic process X_t will converge to a unique equilibrium measure π exponentially fast in time, i.e. there are strictly positive C_μ and β such that

$$\|\mathbb{P}_t^\mu - \pi\| \leq C_\mu e^{-\beta t}.$$

Here $\mathbb{P}_t^\mu$ denotes the law of X_t given that $X_0 \sim \mu$ and the convergence above is measured in the total variation norm, which is defined between any two measures μ and ν through

$$\|\mu - \nu\| := \sup_{|f|\leq 1} \int f(x)\mu(\mathrm{d}x) - \int f(x)\nu(\mathrm{d}x).$$

The supreme here goes through all measurable functions bounded by 1. Geometric ergodicity ensures the uniqueness of invariant measure, which enables the Birkhoff ergodic theorem to be applied. This means in order to measure an integrable statistic f of the equilibrium measure π, it suffices to compute the time average of f for one realization of X_t for sufficiently long time, because by Birkhoff ergodic theorem:

$$\frac{1}{T}\int_0^T f(X_t)\mathrm{d}t \overset{T\to\infty}{\longrightarrow} \int f(x)\pi(\mathrm{d}x) \quad a.s.$$

This is extremely convenient for simulation purposes, as one long enough simulation is sufficient to represent the equilibrium measure, while in a standard Monte Carlo simulation, in principle one has to run the same simulation thousands of times. On the other hand, we can show the turbulent diffusion system (2.3) is geometrically ergodic using hypoellipticity:

Theorem 3.1. *Let the turbulent diffusion system* (2.3) *consist of conjugating Fourier modes in a finite wavenumber set N, then the joint process $(U_t, \hat{v}_{k,t}, \widehat{T}_{k,t}, k \in N \cap \mathbb{Z}^+)$ is geometrically ergodic under the total variation norm. Specifically, for any initial measure μ of $(U_0, \hat{v}_{k,0}, \widehat{T}_{k,0}, k \in N \cap \mathbb{Z}^+)$, there is a unique invariant measure π, and constants $C, \beta > 0$ such that the following holds:*

$$\|\mathbb{P}_t^\mu - \pi\| \le Ce^{-\beta t}\mathbb{E}^\mu(1 + |U_0|^2 + \sum_{k\in N}|\hat{v}_{k,0}|^2 + |\widehat{T}_{k,0}|^2).$$

The reason that we consider only the positive wavenumbers is because their negative counterparts are conjugate to them.

Proof. The proof is a standard verification of the theoretical framework established in [22, 23]. We attach the details of the proof in Section Appendix A.1. □

3.2. Conditional Gaussian structure

The simple formulation of (2.3) provides us two crucial features: first, the only correlation between Fourier modes of wavenumbers $|k| \neq |j|$ is through the realization of $U_{s\ge 0}$; second, with the realization of $U_{s\ge 0}$ fixed, the dynamics of $\hat{v}_{k,t}$ and $\widehat{T}_{k,t}$ will be linear. There are no positive Lyapunov exponents in this system. Thus by conditioning on the realization $U_{s\ge 0}$, $\hat{v}_{k,t}, \widehat{T}_{k,t}$ can be treated independently for each wavenumber and as simple linear processes. This is known as the conditional Gaussian structure [24], which can be exploited for filtering and prediction purposes [25, 26]. Here, we will use it to compute the distribution of $\widehat{T}_{k,t}$, which leads to the hidden source of intermittency developed in the next subsection.

Since U_t is an O.U. process, its invariant measure is a zero mean real Gaussian distribution:

$$\pi_U = \mathcal{N}(0, E_U), \quad E_U := \frac{\sigma_U^2}{2\gamma_U}.$$

By virtue of Theorem 3.1, without loss of generality, we can assume that U_0 is distributed as π_U while $\hat{v}_{k,0}$ and $\widehat{T}_{k,0}$ are initialized from point 0. Thus by Duhamel's formula, we have

$$\hat{v}_{k,s} = \int_0^s \exp(-\gamma_{v,k}(s-r) + \mathrm{i}\omega_{v,k}[r,s])\sigma_{v,k}\mathrm{d}B_{k,r}.$$

Here and in the following, we use the expression $X[r,s]$ to denote the integral of a process X_t on the interval $[r,s]$, i.e. $X[r,s] := \int_r^s X_u du$. We can easily split the real and imaginary parts of $\hat{v}_{k,s}$ into:

$$\mathrm{Re}(\hat{v}_{k,s}) = \frac{\sigma_{v,k}}{\sqrt{2}}\int_0^s \exp(-\gamma_{v,k}(s-r))\cos(\omega_v[r,s])\mathrm{d}B^1_{k,r} - \exp(-\gamma_{v,k}(s-r))\sin(\omega_v[r,s])\mathrm{d}B^2_{k,r},$$

$$\mathrm{Im}(\hat{v}_{k,s}) = \frac{\sigma_{v,k}}{\sqrt{2}}\int_0^s \exp(-\gamma_{v,k}(s-r))\sin(\omega_v[r,s])\mathrm{d}B^1_{k,r} - \exp(-\gamma_{v,k}(s-r))\cos(\omega_v[r,s])\mathrm{d}B^2_{k,r}.$$

Using the Itô isometry, we immediately find that $\mathrm{Re}(\hat{v}_{k,s}), \mathrm{Im}(\hat{v}_{k,s})$ are independent mean zero Gaussian random variables with the same variance. In the engineering literature [27], especially for signal processing, it is customary to call a random variable $Z = X + Y\mathrm{i}$ circularly symmetric complex Gaussian, denoted by $Z \sim \mathcal{CN}(\mu, \Gamma)$ if (X,Y) are jointly real Gaussian while

$$\mathbb{E}X = \mathrm{Re}(\mu), \quad \mathbb{E}Y = \mathrm{Im}(\mu), \quad \mathrm{var}(X) = \mathrm{var}(Y) = \tfrac{1}{2}\Gamma, \quad \mathrm{cov}(X,Y) = 0.$$

Since all the complex random variables in this paper are of this type, we will call them complex Gaussian for simplicity. For example, $\hat{v}_{k,s}$ here is evidently complex Gaussian with zero mean and variance:

$$\mathbb{E}|\hat{v}_{k,s}|^2 = 2\mathbb{E}|\mathrm{Re}(\hat{v}_{k,s})| = \frac{\sigma^2_{k,v}}{2\gamma_{k,v}}(1 - \exp(-2\gamma_{v,k}s)). \tag{3.1}$$

With the long time limit $s \to \infty$, we see $\hat{v}_{k,s}$ converges to its invariant measure $\pi_{v,k} = \mathcal{CN}(0, E_{k,v})$, $E_{k,v} = \sigma^2_{k,v}/2\gamma_{k,v}$.

Likewise, the Fourier modes of the passive tracer $\widehat{T}_{k,t}$, by Duhamel's formula and Fubini's theorem for stochastic integral [28], can be written as:

$$\begin{aligned}\widehat{T}_{k,t} &= \alpha\int_0^t \exp(-\gamma_{T,k}(t-s) + \mathrm{i}\omega_{T,k}[s,t])\hat{v}_{k,s}\mathrm{d}s \\ &= \int_0^t \left(\int_r^t \alpha\sigma_{v,k}\exp(-\gamma_{T,k}(t-s) - \gamma_{v,k}(s-r) + \mathrm{i}\omega_{T,k}[s,t] + \mathrm{i}\omega_{v,k}[r,s])ds\right)\mathrm{d}B_{k,r}.\end{aligned}$$

Notice that $B_{k,r}$ are complex Wiener processes independent of U_t, so we immediately have the following conclusion.

Proposition 3.2. *Let $U_t, \hat{v}_{k,t}, \widehat{T}_{k,t}$ be as in system (2.3) with $U_0 \sim \pi_U, \hat{v}_{k,0} = \widehat{T}_{k,0} = 0$, then conditioned on the realization of $U_{s\geq 0}$, $\widehat{T}_{k,t}$ is a complex Gaussian random variable with zero mean and variance:*

$$\Sigma_{k,t|U} = \mathbb{E}(|\widehat{T}_{k,t}|^2|U_{s\geq 0}) = \alpha^2\sigma^2_{v,k}\int_0^t \exp(-2\gamma_{v,k}(t-r))\left|\int_r^t \exp(\gamma_{R,k}s + \mathrm{i}\omega_{R,k}[s,t])\right|^2\mathrm{d}r,$$

with $\omega_{R,k}(t) := \omega_{T,k}(t) - \omega_{v,k}(t)$ being the phase speed difference and $\gamma_{R,k} = \gamma_{T,k} - \gamma_{v,k}$. By the law of total expectation, the probability density of $\widehat{T}_{k,t}$ is a mixture of zero mean complex Gaussian distributions:

$$\mathbb{P}(Re(\widehat{T}_{k,t}) \in \mathrm{d}x, Im(\widehat{T}_{k,t}) \in \mathrm{d}y) = \int \frac{1}{\pi\Sigma} \exp\bigg(- \frac{x^2+y^2}{\Sigma}\bigg) \mathbb{P}(\Sigma_{k,t|U} \in d\Sigma). \tag{3.2}$$

Although formula (3.2) admits no closed form expression, we can establish an upper bound for the tail distribution and variance of $\widehat{T}_{k,t}$ through elementary computations:

Proposition 3.3. *The conditional variance of $\widehat{T}_{k,t}$ given $U_{s\le t}$ is bounded $\mathbb{P}$-a.s. by*

$$\overline{\Sigma}_{k,t} = \frac{\sigma_{v,k}^2\alpha^2}{\gamma_{R,k}|\gamma_{R,k}|}\bigg[\frac{1-\exp(-2\gamma_{v,k}t)}{2\gamma_{v,k}} - \frac{1-\exp(-2\gamma_{T,k}t)}{2\gamma_{T,k}}\bigg], \quad \gamma_{R,k} := \gamma_{T,k} - \gamma_{v,k}.$$

Therefore the tail distribution of $\widehat{T}_{k,t}$ is bounded by a zero mean Gaussian distribution with variance $\overline{\Sigma}_{k,t}$. More strictly speaking, for any $\lambda_1, \lambda_2 \in \mathbb{R}$

$$\mathbb{E}\exp(\lambda_1 Re(\widehat{T}_{k,t}) + \lambda_2 Im(\widehat{T}_{k,t})) \le \exp((\lambda_1^2+\lambda_2^2)\overline{\Sigma}_{k,t}/4).$$

On the other hand, the variance of $\widehat{T}_{k,t}$ which is also the mean of $\Sigma_{k,t|U}$ is bounded from above by

$$\mathbb{E}|\widehat{T}_{k,t}|^2 = \mathbb{E}\Sigma_{k,t|U} \le \frac{|\gamma_{R,k}|\overline{\Sigma}_{k,t}}{\sqrt{\gamma_{R,k}^2 + (b_k + k\bar{u})^2}}.$$

Proof. The proof is based on elementary bounds and Fourier transformation of a Gaussian distribution. See Section Appendix A.2 for the complete details. □

3.3. Intermittency and resonance

At first sight, the conditional Gaussian structure and Gaussian tail bounds from Proposition 3.3 may seem contradictory to intermittent phenomena, as Gaussian distributions rarely exhibit intermittent behavior. However, if the conditional variance process $\Sigma_{k,t|U}$ has large fluctuations, the distribution of $\widehat{T}_{k,t}$ generated through formula (3.2) may possess a much heavier tail comparing to its Gaussian fit. For example, if the variance of $\widehat{T}_{k,t}$ is $\bar{\Sigma}$ and $\mathbb{P}(\Sigma_{k,t|U} > 9\bar{\Sigma}) \ge 1\%$, then

$$\mathbb{P}(|\hat{T}_{k,t}| > 3\bar{\Sigma}^{\frac{1}{2}}) \ge 2\mathbb{P}(\Sigma_{k,t|U} > 9\bar{\Sigma})\Phi(1) \ge 3\%,$$

while in a complex Gaussian distribution with variance $\bar{\Sigma}$, the probability of $\mathbb{P}(|X| > 3\bar{\Sigma}^{\frac{1}{2}})$ is about 0.3%. In other words, the large fluctuation in conditional variance makes extreme events such as large spikes much more frequent than the Gaussian fit of the distribution, and thus produces the intermittency seen in the corresponding time sequence. Moreover, the results of Proposition 3.3 actually imply that $\Sigma_{k,t|U}$ has large fluctuations; since in many

situations the damping rates $\gamma_{v,k}, \gamma_{T,k}$ are small numbers, the ratio between the upper bound of $\Sigma_{k,t|U}$ and its mean is a large quantity:

$$\frac{\overline{\Sigma}_{k,t}}{\mathbb{E}\Sigma_{k,t|U}} \geq \frac{\sqrt{(\gamma_{T,k}-\gamma_{v,k})^2+(b_k+k\bar{u})^2}}{|\gamma_{T,k}-\gamma_{v,k}|}.$$

On the other hand, we can have a glimpse of the possible cause of this large fluctuation based on the formula of $\Sigma_{k,t|U}$ in Proposition 3.2:

$$\Sigma_{k,t|U} = \alpha^2\sigma_{v,k}^2 \int_0^t \exp(-2\gamma_{T,k}(t-r))\left|\int_r^t \exp(\gamma_{R,k}s + \mathrm{i}\omega_{R,k}[s,t])\mathrm{d}s\right|^2 \mathrm{d}r,$$

where the resonant phase speed is defined by

$$\omega_{R,k}(t) = \omega_{T,k}(t) - \omega_{v,k}(t) = -(a_k+k)U_t - (b_k+k\bar{u}).$$

In a general scenario, $\omega_{R,k}$ is away from 0, so $\exp(\gamma_{R,k}s + \mathrm{i}\omega_{R,k}[s,t])$ will be oscillating in the complex plane with nonzero phase speed, this inevitably cancels a large portion of its integral, reducing the size of $\Sigma_{k,t|U}$. However, when $\omega_{R,k}(s)$ is close to 0 for a significant period of time, the integrand has no phase speed and the integral will be significantly larger. Looking back at the definition of $\omega_{R,k}(t)$, we find that it is zero when the phase speed $\omega_{T,k}$ and $\omega_{v,k}$ are equal, in other words, the advection effect of the cross sweep U_t over $\hat{v}_{k,t}$ and $\widehat{T}_{k,t}$ are at resonance. This is the reason why we call $\omega_{R,k}$ the resonance phase speed and will focus on its visit to zero in the following.

Although the intuition is clear, rigorous verification is difficult due to the lack of an explicit formula for $\Sigma_{k,t|U}$. One way is to produce a lower bound for $\Sigma_{k,t|U}$ given that $\omega_{R,k}(t)$ is around 0 under some special dynamical regime, which then explicitly bounds $\Sigma_{k,t|U}$ away from $\mathbb{E}\Sigma_{k,t|U}$ using its upper bound in Proposition 3.3. The following result is one such example, which in principle will work for small γ_U:

Proposition 3.4. *For any fixed $t_0 < \gamma_U^{-1} \wedge t$ the following holds:*

$$\mathbb{E}(\Sigma_{k,t|U}|U_t) \geq \underline{\Sigma}_{k,t} := \alpha^2\sigma_{v,k}^2\left[\frac{1-\exp(-2\gamma_{m,k}t_0)}{2\gamma_{m,k}|\gamma_{R,k}-\mathrm{i}\omega_{R,k}(t)|^2} - \frac{\frac{\gamma_{m,k}}{\gamma_{T,k}+\gamma_{v,k}} + C(1-\exp(-2\gamma_{m,k}t_0))}{\gamma_{m,k}|\gamma_{R,k}||\gamma_{R,k}-\mathrm{i}\omega_{R,k}(t)|}\right],$$

where we denote $\gamma_{m,k} := \gamma_{v,k} \wedge \gamma_{T,k}$ and

$$C = 2 - 2\exp(-\tfrac{\gamma_U}{3}(a_k+k)^2E_Ut_0^3) + \frac{\gamma_U}{2}t_0^2|\omega_{R,k}(t)+(b_k+k\bar{u})|.$$

Therefore by the Markov inequality, $\mathbb{P}(\Sigma_{k,t|U} \geq \frac{1}{2}\underline{\Sigma}_{k,t}|U_t) \geq \frac{1}{2}$, $\mathbb{P}$-a.s. As a special case, when $\gamma_{U,k} \to 0$, one can take $t_0 = t$, therefore

$$\underline{\Sigma}_{k,t} \to \alpha^2\sigma_{v,k}^2\left[\frac{1-\exp(-2\gamma_{m,k}t)}{2\gamma_{m,k}|\gamma_{R,k}-\mathrm{i}\omega_{R,k}(t)|^2} - \frac{1}{(\gamma_{v,k}+\gamma_{T,k})|\gamma_{R,k}||\gamma_{R,k}-\mathrm{i}\omega_{R,k}(t)|}\right],$$

which closely approximates the upper bound $\overline{\Sigma}_{k,t}$ given by Proposition 3.3 when $\gamma_{v,k}$ and $\gamma_{T,k}$ are far apart and $\omega_{R,k}(t)$ is close to 0.

Proof. The proof is based on elementary calculations and inequalities. See Section Appendix A.3 for details. □

For example, consider the dynamical regime for $k = 10$ with

$$\gamma_{T,k} = 0.2, \quad \gamma_{v,k} = 1, \quad \gamma_U = 10^{-5}, \quad a_k = b_k = 0, \quad \bar{u} = 10, \quad E_U = E_{v,k} = 100. \tag{3.3}$$

If we pick $t_0 = 10$, we will find $\Sigma_{k,t} \geq 8\mathbb{E}\Sigma_{k,t|U}$ when $\omega_{R,k}(t) = 0$ using the upper bound in Proposition 3.3. This explicitly shows the large fluctuations in $\Sigma_{k,t|U}$. Note that here $\gamma_{T,k} \ll \gamma_{v,k}$, which is opposite of what will be studied in Section 4, hinting that intermittency should generally exists in many dynamical regimes. On the other hand, the lower bound given by Proposition 3.4 in general settings may be too small or even negative, and probably works only when γ_U is very small. We present it here as a prototype tool for intermittency analysis and we illustrate this with a numerical example in Section 5 (see Figure 5.6). The other rigorous way to access intermittency is through a rescaling limit, which will be presented in the next section.

4. Rigorous intermittency in a random resonance regime

4.1. A long time slow varying rescaling limit

Here we consider a scenario where the velocity field (u_t, v_t) is varying at a slower time scale than the advection and diffusion process. This prolongs the visit of the resonance phase ω_R near 0, and provides us a magnified view of the corresponding effect over the turbulent tracer. This can be modeled by external forcing for $(u_t, \hat{v}_t)$ are of order ϵ, while keeping the dynamics of T_t (1.1) the same. More explicitly, we reformulate model (2.2) into

$$\begin{aligned} \mathrm{d}U_t^\epsilon &= -\epsilon\gamma_U U_t^\epsilon \mathrm{d}t + \epsilon^{\frac{1}{2}}\sigma_U \mathrm{d}W_t, \\ \frac{\partial v_t^\epsilon}{\partial t} &= R_1\left(\frac{\partial}{\partial x}\right)v_t^\epsilon + R_2\left(\frac{\partial}{\partial x}\right)v_t^\epsilon - \epsilon\gamma_v\left(\frac{\partial}{\partial x}\right)v_t^\epsilon + \epsilon^{\frac{1}{2}}\dot{W}_v(x,t), \\ \frac{\partial T_t^\epsilon}{\partial t} &+ (U_t^\epsilon + \bar{u})\frac{\partial T_t^\epsilon}{\partial x} = -d_T T_t^\epsilon + \kappa\Delta T_t^\epsilon - \alpha v_t^\epsilon. \end{aligned}$$

Notice that there is no rescaling over R_1, R_2 as they represent the forcing U_t exerts on v_t, such as the advection. By applying Fourier transformation as in Section 2, the joint dynamics of $(U_t^\epsilon, v_t^\epsilon, T_t^\epsilon)$ can be described through the dynamics of each wavenumber:

$$\begin{aligned} \mathrm{d}U_t^\epsilon &= -\epsilon\gamma_U U_t^\epsilon \mathrm{d}t + \epsilon^{\frac{1}{2}}\sigma_U \mathrm{d}W_U(t), \\ \mathrm{d}\hat{v}_{k,t}^\epsilon &= [-\epsilon\gamma_{v,k} + \mathrm{i}\omega_{v,k}(t)]\hat{v}_{k,t}^\epsilon \mathrm{d}t + \epsilon^{\frac{1}{2}}\sigma_{v,k}\mathrm{d}B_{k,t}, \\ \mathrm{d}\widehat{T}_{k,t}^\epsilon &= [-\gamma_{T,k} + \mathrm{i}\omega_{T,k}(t)]\widehat{T}_{k,t}^\epsilon \mathrm{d}t - \alpha\hat{v}_{k,t}\mathrm{d}t, \\ \omega_{v,k}(t) &= a_k U_t^\epsilon + b_k, \quad \omega_{T,k}(t) = -k(U_t^\epsilon + \bar{u}). \end{aligned}$$

With this formulation of dynamics, it is natural to look at the process at the long time scale ϵ^{-1}. With an abuse of notation, we denote the long time rescaling by:

$$(U_t, \hat{v}_{k,t}, \widehat{T}_{k,t}) = (U_{\epsilon t}^\epsilon, \hat{v}_{k,\epsilon t}^\epsilon, \widehat{T}_{k,\epsilon t}^\epsilon).$$

The corresponding dynamics is then given by the following:

$$\begin{aligned} \mathrm{d}U_t &= -\gamma_U \mathrm{d}t + \sigma_U \mathrm{d}W_t; \\ \mathrm{d}\hat{v}_{k,t} &= [-\gamma_{v,k} + \mathrm{i}\epsilon^{-1}\omega_{v,k}(t)]\hat{v}_{k,t}\mathrm{d}t + \sigma_{k,v}\mathrm{d}B_{k,t}; \\ \mathrm{d}\widehat{T}_{k,t} &= \epsilon^{-1}[(-\gamma_{T,k} + \mathrm{i}\omega_{T,k}(t))\widehat{T}_{k,t}\mathrm{d}t - \alpha\hat{v}_{k,t}]\mathrm{d}t; \\ \omega_{k,V}(t) &= a_k U_t + b_k, \quad \omega_{k,T}(t) = -k(U_t + \bar{u}). \end{aligned} \tag{4.1}$$

Immediately, we find that U_t no longer depends on ϵ anymore, and so does the distribution of $\hat{v}_{k,t}$, since phase speeds play no role in its variance, as we can see in formula(3.1). Hence we can fix the realization of $U_{s\le t}$, and focus on the conditional distribution of $T_{k,v}(t)$ and $T_t(x)$ with ϵ tends to 0.

4.2. Single Fourier mode

In the $\epsilon \to 0$ rescaling limit of (4.1), the distribution of the passive tracer can be written down explicitly. As the first step, we analyze the distribution of one fixed Fourier mode. The result is given by the following theorem:

Theorem 4.1. *For any fixed $t > 0$, let $(U_t, \hat{v}_{k,t}, \widehat{T}_{k,t})$ follow (4.1). Assume the joint process is initialized with $U_0 \sim \mathcal{N}(0, E_U), V_0 = 0, T_0 = 0$. Then given the realization of $U_{s\le t}$, the conditional distribution of $\widehat{T}_{k,t}$ is a complex Gaussian distribution with zero mean and variance $\Sigma^{\epsilon}_{k,t|U}$, which converges in L^1 to*

$$\tilde{\Sigma}_{k,t}(U_t) := \frac{\alpha^2 E_{v,k}(1 - \exp(-2\gamma_{v,k}t))}{\gamma_{T,k}^2 + \omega_{R,k}^2(t)}, \quad \omega_{R,k}(t) = \omega_{T,k}(t) - \omega_{v,k}(t).$$

More specifically, we have $\mathbb{E}|\Sigma^{\epsilon}_{k,t|U} - \tilde{\Sigma}_{k,t}(U_t)| \to 0$. The distribution of $\widehat{T}_{k,t}$ thus converges weakly to a mixture of complex Gaussian distributions with density given by:

$$\mathbb{P}(Re(\widehat{T}_{k,t}) \in \mathrm{d}x, Im(\widehat{T}_{k,t}) \in \mathrm{d}y) = \int_{-\infty}^{\infty} \frac{\exp(-\frac{u^2}{2E_U} - \frac{x^2+y^2}{\tilde{\Sigma}_{k,t}(c_k u + d_k)})}{\pi\sqrt{2\pi E_U}\tilde{\Sigma}_{k,t}(c_k u + d_k)} du, \tag{4.2}$$

where $c_k = -(a_k + k), d_k = -(b_k + k\bar{u})$. $\tilde{\Sigma}_t$ has a limit for large t:

$$\tilde{\Sigma}_k(\omega) = \frac{\alpha^2 E_{v,k}}{\gamma_{T,k}^2 + (c_k\omega + d_k)^2}.$$

This can be used for computation of the equilibrium measure for $\widehat{T}_{k,t}$.

In other words, the dependence of $\widehat{T}_{k,t}$'s distribution over the path of $U_{s\le t}$ concentrates on its end value through the resonance phase speed $\omega_{R,k}(t)$. Heuristically speaking, this is because $U_{s\le t}$ as a slow process will be of constant value around time t, while $\widehat{T}$ has its distribution converge quickly to the equilibrium distribution with $U_{s\le t}$ being a constant process.

The limiting conditional covariance $\tilde{\Sigma}_{k,t}(U_t)$ obviously reaches its peak value when $\omega_{R,k}(t) = 0$. Moreover, by applying Proposition 3.3 to the rescaled system (4.1), we can find that:

$$\mathbb{E}\Sigma_{k,t|U} \leq \frac{\alpha^2 E_{v,k}(1-\exp(-2\gamma_v t))}{\gamma_{T,k}\sqrt{\gamma_{T,k}^2+(b_k+k\bar{u})^2}} \quad \Rightarrow \quad \tilde{\Sigma}_{k,t} \geq \frac{\gamma_{T,k}\sqrt{\gamma_{T,k}^2+(b_k+k\bar{u})^2}}{\gamma_{T,k}^2+\omega_{R,k}^2(t)}\mathbb{E}\Sigma_{k,t|U}.$$

This rigorously shows that in a general scenario where the damping rate γ_T is small compared to $|b_k + k\bar{u}|$, the conditional variance, which can be approximated by $\tilde{\Sigma}_{k,t}$, will be a large multiple of its mean around resonance phase $\omega_{R,k} = 0$, and will likely create intermittency as discussed in Section 3.3. In Figure 5.1, a numerical simulation will validate our claim here with intermittent time sequence plots.

In a deterministic version of 2.2, [10, 5] studied the large Péclet number limit of

$$\frac{\partial T}{\partial t} + \mathrm{Pe}(\vec{v}\cdot\nabla T) = \Delta T,$$

where the velocity field is $(u_t, v(x,t))$ with the process U_t being a deterministic sinusoidal function. The results there show phase transition type of phenomenon when U_t visits zero, the streamlines of velocity field will be open against the general situations where the streamlines are blocked, and the passive tracer' signal produces strong spikes as consequence. The asymptotic limit there is derived through stationary phase methods. Our results here produce a stochastic counterpart in a slightly different rescaling limit. And because U_t is a stochastic process with non smooth path, the proof is based on different intuition.

Proof of Theorems 4.1. As there is only one wavenumber k being involved, we will suppress its appearance in subscripts. Moreover, since we can always divide $\widehat{T}_t$ by $\alpha\sigma_v$, so without loss of generality we assume that $\alpha\sigma_v = 1$. Based on Proposition 3.2, $\Sigma_{t|U}^{\epsilon} = \int_0^t \left| \int_r^t \sigma_{r,s}^{\epsilon} \mathrm{d}s \right|^2 \mathrm{d}r$ with

$$\sigma_{r,s}^{\epsilon} = \epsilon^{-1}\exp(-\epsilon^{-1}\gamma_T(t-s) - \gamma_v(s-r) + \epsilon^{-1}\mathrm{i}\omega_R[s,t]). \tag{4.3}$$

Consider the following approximation of $\sigma_{r,s}^{\epsilon}$:

$$\delta_{r,s}^{\epsilon} = \epsilon^{-1}\exp(-\epsilon^{-1}\gamma_T(t-s) - \gamma_v(s-r) + \mathrm{i}\epsilon^{-1}\omega_R(t)(t-s)).$$

Since $\delta_{r,s}^{\epsilon}$ is an exponential that depends on s linearly, its integral can be explicitly written down:

$$\int_r^t \delta_{r,s}^{\epsilon}\mathrm{d}s = \frac{\exp(-\gamma_v(t-r))}{\gamma_T - \epsilon\gamma_v - \mathrm{i}\omega_R(t)}[1-\exp(-\epsilon^{-1}\gamma_T(t-r) + \gamma_v(t-r) + \mathrm{i}\epsilon^{-1}\omega_R(t)(t-r_\epsilon))].$$

The first step of our proof is showing this quantity converges to $\tilde{\Sigma}_t(U_t)$. Notice that for

$\epsilon \le \gamma_T/2\gamma_v$,

$$\begin{aligned}
&\left| \int_r^t \delta_{r,s}^\epsilon \mathrm{d}s - \frac{\exp(-\gamma_v(t-r))}{\gamma_T - \mathrm{i}\omega_R(t)} \right| \\
&\le \left| \frac{\exp(-\gamma_v(t-r))}{\gamma_T - \epsilon\gamma_v - \mathrm{i}\gamma_R(t)} - \frac{\exp(-\gamma_v(t-r))}{\gamma_T - \mathrm{i}\gamma_R(t)} \right| + \frac{\exp(-\epsilon^{-1}\gamma_T(t-r))}{|\gamma_T - \epsilon\gamma_v - \mathrm{i}\omega_R(t)|} \\
&\le \frac{2\epsilon\gamma_v \exp(-\gamma_v(t-r))}{\gamma_T^2} + \frac{2\exp(-\epsilon^{-1}\gamma_T(t-r))}{\gamma_T} \le \frac{4\exp(-\gamma_v(t-r))}{\gamma_T}.
\end{aligned}$$

Therefore by the relation $|a|^2 - |b|^2 = (|a| - |b|)^2 + 2|b|(|a| - |b|)$,

$$\left| \left| \int_r^t \delta_{r,s}^\epsilon \mathrm{d}s \right|^2 - \frac{\exp(-2\gamma_v(t-r))}{\gamma_T^2 + \omega_R^2(t)} \right| \le \frac{12\exp(-\gamma_v(t-r))}{\gamma_T} \left[\frac{\epsilon\gamma_v \exp(-\gamma_v(t-r))}{\gamma_T^2} + \frac{\exp(-\epsilon^{-1}\gamma_T(t-r))}{\gamma_T} \right].$$

Since

$$\tilde{\Sigma}_t(U_t) = \int_0^t \frac{\exp(-2\gamma_v(t-r))}{\gamma_T^2 + \omega_R^2(t)} \mathrm{d}r,$$

Therefore

$$\left| \int_0^t \left| \int_r^t \delta_{r,s}^\epsilon \mathrm{d}s \right|^2 \mathrm{d}r - \tilde{\Sigma}_t(U_t) \right| \le \frac{12\epsilon}{\gamma_T^3} \left(1 + \frac{1}{\gamma_T} \right), \tag{4.4}$$

which goes to zero when ϵ does. The second step of our proof is showing that the integral of $\sigma_{r,s}^\epsilon$ converges to the integral of $\delta_{r,s}^\epsilon$. We will show the following quantity is small when ϵ is small:

$$\mathbb{E} \int_0^t \left| \int_r^t (\sigma_{r,s}^\epsilon - \delta_{r,s}^\epsilon) \mathrm{d}s \right|^2 \mathrm{d}r.$$

For this purpose, with any fixed $\delta > 0$, there is an M such that the following holds for all $\gamma_T \ge d_T$:

$$\exp(-\gamma_T M) \le \frac{\delta}{\gamma_T}.$$

Denote $r_\epsilon = (t - M\epsilon) \vee r$. Notice that

$$\int_r^{r_\epsilon} (\sigma_{r,s}^\epsilon - \delta_{r,s}^\epsilon) \mathrm{d}s \le \int_r^{r_\epsilon} |\sigma_{r,s}^\epsilon| + |\delta_{r,s}^\epsilon| \mathrm{d}s \le \frac{2\delta}{\gamma_T}.$$

For the integral between r_ϵ and t, using the relation $|e^{ai} - e^{bi}| \le |a - b|$ and Cauchy Schwartz:

$$\begin{aligned}
\left(\int_{r_\epsilon}^t |\sigma_{r,s}^\epsilon - \delta_{r,s}^\epsilon| \mathrm{d}s \right)^2 &\le \epsilon^{-2} \left(\int_{r_\epsilon}^t \exp(-\epsilon^{-1}\gamma_T(t-s)) |\epsilon^{-1}\omega_R[t,s] - \epsilon^{-1}(t-s)\omega_R(t)| \mathrm{d}s \right)^2 \\
&\le \epsilon^{-2} \int_{r_\epsilon}^t \exp(-2\epsilon^{-1}\gamma_T(t-s)) \mathrm{d}s \int_{r_\epsilon}^t [\epsilon^{-1}[\omega_R[t,s] - \omega_R(t)(t-s)]]^2 \mathrm{d}s \\
&\le \frac{2}{\epsilon^3\gamma_T} \left[\int_{r_\epsilon}^t [\omega_R[t,s] - \omega_R(t)(t-s)]^2 \mathrm{d}s \right].
\end{aligned}$$

Since $\omega_R(t)$ is a linear function of U_t, it suffices to show for all $r \le t$, there is a constant D

$$\mathbb{E}\int_{r_\epsilon}^{t}[\int_{s}^{t} U_r \mathrm{d}r - (t-s)U_t]^2 \mathrm{d}s \le D\epsilon^5.$$

This can be proved through Lemma Appendix A.4 and using the Markov property of U_t, and it holds for general diffusion process U_t. Then by Young's inequality

$$\int_0^t \mathbb{E}\left|\int_r^t (\sigma^\epsilon_{r,s} - \delta^\epsilon_{r,s})\mathrm{d}s\right|^2 \mathrm{d}r \le 2\int_0^t \frac{\delta^2}{\gamma_T^2} + \mathbb{E}\left|\int_{r_\epsilon}^t (\sigma^\epsilon_{r,s} - \delta^\epsilon_{r,s})\mathrm{d}s\right|^2 \mathrm{d}r \le 2t\left[\frac{\delta^2}{\gamma_T^2} + \frac{2\epsilon^2 D}{\gamma_T}\right].$$

Next, using a particular version of Cauchy Schwartz inequality, which is detailed in Lemma Appendix A.3,

$$\begin{aligned}\mathbb{E}\left|\int_0^t \left|\int_r^t \sigma^\epsilon_{r,s}\mathrm{d}s\right|^2 \mathrm{d}r - \int_0^t \left|\int_r^t \delta^\epsilon_{r,s}\mathrm{d}s\right|^2 \mathrm{d}r\right| \\ \le 2\frac{\sqrt{C}}{\gamma_T}\sqrt{2t\left[\frac{\delta^2}{\gamma_T^2} + \frac{2\epsilon^2 D}{\gamma_T}\right]} + 8t\left[\frac{\delta^2}{\gamma_T^2} + \frac{2\epsilon^2 D}{\gamma_T}\right].\end{aligned}$$

Here C is a constant such that

$$\int_0^t \hat{v}_{\vec{k},\pm}(s)e^{\mathrm{i}k\cdot\vec{x}}ds \le \frac{C}{\gamma_T^2},$$

while by inequality (4.4) we know such C does exist. So combining these estimates and putting back $\alpha^2\sigma_v^2$

$$\begin{aligned}&\mathbb{E}|\Sigma^\epsilon_{t|U} - \tilde{\Sigma}_t(\omega_R(t))| \\ &\le \alpha^2\sigma_v^2\left[\frac{12\epsilon}{\gamma_T^3}\left(1+\frac{1}{\gamma_T}\right) + 2\frac{\sqrt{C}}{\gamma_T}\sqrt{2t\left[\frac{\delta^2}{\gamma_T^2} + \frac{2\epsilon^2 D}{\gamma_T}\right]} + 8t\left[\frac{\delta^2}{\gamma_T^2} + \frac{2\epsilon^2 D}{\gamma_T}\right]\right]. \qquad (4.5)\end{aligned}$$

Since δ can be any small number, this quantity goes to 0 as $\epsilon \to 0$.

To see this implies that $\widehat{T}$'s distribution converges weakly to (4.2), consider the following Gaussian operator,

$$K_\Sigma : f(x,y) \mapsto \int \frac{\exp(-(x^2+y^2)/\Sigma)}{\pi\Sigma} f(x,y)\mathrm{d}x\mathrm{d}y.$$

Then $\mathbb{E}f(\mathrm{Re}(\widehat{T}), \mathrm{Im}(\widehat{T})) = \mathbb{E}K_{\Sigma^\epsilon_{t|U}}f$. Notice that $K_\Sigma f$ depends continuously on $\Sigma > 0$ for any fixed bounded f, and the law of $\Sigma^\epsilon_{t|U}$ converges towards the one of $\tilde{\Sigma}_t$ by L^1 convergence, so $\mathbb{E}K_{\Sigma^\epsilon_{t|U}}f \to \mathbb{E}K_{\tilde{\Sigma}_t}f$. On the other hand, $\mathbb{E}K_{\tilde{\Sigma}_t}f$ is clearly given by the integration of f under density (4.2), so by standard properties of weak convergence, see for example [29], $\widehat{T}$'s distribution converges weakly towards (4.2). □

4.3. Finitely many Fourier modes

With a clear understanding of each Fourier mode, the distribution of the passive tracer field $T_t(x)$ can be easily studied by summing all the Fourier modes together. Here for simplicity, we assume the wavenumber set N is finite.

Theorem 4.2. *Assume that in the Fourier expansion $T_t(x) = \sum_{k\in N} \widehat{T}_{k,t} e^{ikx}$, N consists of finitely many pairs of Fourier modes, and each follows (4.1), then conditioned on each realization of $U_{s\le t}$, $T_t(x)$ is a Gaussian random field on $[0, 2\pi]$ with mean 0, while its conditional covariance converges in L^1 to*

$$\tilde{\Sigma}_t(U_t) = \sum_{k\in N} \frac{\alpha^2 E_{v,k}(1-\exp(-2\gamma_{v,k}t))}{\gamma_{T,k}^2 + \omega_{R,k}^2(t)}, \quad \omega_{R,k}(t) = -(a_k + k)U_t - (b_k + k\bar{u}).$$

Moreover the conditional covariance between two points $x, y \in [0, 2\pi]$ converges in L^1 to:

$$\tilde{C}_t(U_t) = \sum_{k\in N} \frac{\alpha^2 E_{v,k}\cos(k(x-y))(1-\exp(-2\gamma_{v,k}t))}{\gamma_{T,k}^2 + \omega_{R,k}^2(t)}.$$

Thus the pdf of $T_t(x)$ will converge to a density

$$\mathbb{P}(T_t(x) \in \mathrm{d}\lambda) = \int \frac{1}{2\pi\sqrt{\tilde{\Sigma}_t(u)E_U}} \exp\left(-\frac{\lambda^2}{2\tilde{\Sigma}_t(u)} - \frac{u^2}{2E_U}\right)\mathrm{d}u. \tag{4.6}$$

Unlike the limiting conditional variance of a single Fourier mode, $\tilde{\Sigma}_t(u)$ may have many peaks. Heuristically speaking, each pair of conjugating Fourier modes will contribute one peak of $\tilde{\Sigma}_t(U_t)$, which is close to the time when $\omega_{R,k}(t)$ reaches 0. In other words, as U_t crosses different resonant phases

$$\omega_k := -\frac{b_k + k\bar{u}}{a_k + k},$$

the Fourier modes of wavenumber $\pm k$ will be excited and contribute a large spike of intermittency. These phenomena are explicitly verified by simulations in Figures 5.3-5.5 in the next section.

Proof for Theorem 4.2. Recall that the complete tracer field is given by:

$$T_t(x) = \sum_{k\in N} \widehat{T}_{k,t} e^{\mathrm{i}kx}, \quad \widehat{T}_{-k,t} = \widehat{T}_{k,t}^*.$$

Conditioned on the realization of $U_{s\le t}$, $(\hat{v}_{k,t}, \widehat{T}_{k,t})$ is independent of $(\hat{v}_{j,t}, \widehat{T}_{j,t})$ for $|j| \neq |k|$, and $\widehat{T}_{k,t} \sim \mathcal{CN}(0, \Sigma^{\epsilon}_{k,t|U})$. Thus conditioned on $U_{s\le t}$, $T_t(x)$ is also Gaussian with mean 0 and covariance

$$\mathbb{E}[T_t(x)|U_{s\le t}]^2 = \frac{1}{4}\mathbb{E}\left[\sum_{k\in N}(\widehat{T}_{k,t}e^{\mathrm{i}kx} + \widehat{T}_{-k,t}e^{-\mathrm{i}kx})\middle| U_{s\le t}\right]^2 = \sum_{k\in N}\mathbb{E}[|\widehat{T}_{k,t}|^2|U_{s\le t}] = \sum_{k\in N}\Sigma^{\epsilon}_{k,t|U}.$$

Likewise, the conditional covariance is given by

$$\mathbb{E}[T_t(x)T_t(y)|U_{s\le t}] = \mathbb{E}\Big[\sum_{j,k\in N}\widehat{T}_{j,t}\widehat{T}_{k,t}e^{\mathrm{i}(kx+jy)}\Big|U_{s\le t}\Big]^2 = \mathbb{E}\Big[\sum_{k\in N}|\widehat{T}_{k,t}|^2e^{\mathrm{i}k(x-y)}\Big|U_{s\le t}\Big]^2$$
$$= \sum_k \cos(k(x-y))\Sigma^{\epsilon}_{k,t|U}.$$

In order to show our claim it suffices to show

$$\mathbb{E}\sum_{k\in N}|\Sigma^{\epsilon}_{k,t|U} - \tilde{\Sigma}_{k,t}(U_t)| \xrightarrow{\epsilon\to 0} 0.$$

Since N is a finite set, one can simply apply the proof of Theorem 4.1 to each individual k and see the convergence holds. □

5. Intermittency in various regimes: theory and numerical experiments

In this section we use elementary numerical experiments to provide evidence for our theorems in Section 3 and 4. We also show how the theory leads to a wide variety of intermittent behavior in the turbulent signals. We begin with the β-plane Q-G flow model from [5]:

$$\gamma_{v,k} = d_v + \nu k^2, \quad \gamma_{T,k} = d_T + \kappa k^2, \quad a_k = \frac{-k^3}{k^2+F}, \quad b_k = \frac{\beta k}{k^2+F}.$$

where the parameters are the same ones used in [5]:

$$d_T = 0.1 \quad \kappa = 0.001, \quad d_v = 0.6, \quad \nu = 0.1, \quad \alpha = 1, \quad \beta = 8.91, \quad F = 16, \quad E_U = 0.5.$$

For their physical background, we refer to [5, 20]. We will also test the example described by (3.3) and the random cross sweep model based on [10], i.e. $a_k = b_k = 0$, which gives us connections with previously known results. Since we will simulate a multiscale system (4.1), we use an explicit Euler scheme with exponential integrator to avoid possible stiffness. In other words, system (4.1) is simulated as:

$$U_{t+\Delta} = (1-\gamma_U\Delta)U_t + \sigma_U\sqrt{\Delta}w_t;$$
$$\hat{v}_{k,t+\Delta} = \exp(-\Delta\gamma_{v,k} + \mathrm{i}\epsilon^{-1}\Delta\omega_{v,k}(t))\hat{v}_{k,t} + \sigma_{v,k}\sqrt{\Delta}\frac{b^1_t + \mathrm{i}b^2_t}{\sqrt{2}};$$
$$\widehat{T}_{k,t+\Delta} = \exp(-\epsilon^{-1}\gamma_{T,k}\Delta + \epsilon^{-1}\mathrm{i}\Delta\omega_{T,k}(t))\widehat{T}_{k,t} - \alpha\Delta\hat{v}_t.$$

Here w_t, b^i_t are sequences of standard normal random variables. In order to avoid numerical error, the time step Δ is chosen to satisfy:

$$\Delta\epsilon^{-1}(\mathbb{E}\omega_{T,k} + \sqrt{\mathrm{var}(\omega_{T,k})}) \le 0.2, \quad \Delta\epsilon^{-1}(\mathbb{E}\omega_{v,k} + \sqrt{\mathrm{var}(\omega_{v,k})}) \le 0.2.$$

By the virtue of geometric ergodicity, Theorem 3.1, we can represent the equilibrium distribution of $\widehat{T}_{k,t}$ and $T_t(x)$ with the histogram of the time sequence $\widehat{T}_{k,n\Delta}, T_{n\Delta}(x)$, as long as the time span T_{max} is long enough. Since ϵ is the time scale of $\widehat{T}_{k,t}$, here we pick $T_{max} = 10^6\epsilon$ for $\epsilon = 0.1, 0.01, 0.001$, which is at least 1000 times longer than the decorrelation time of any variable.

5.1. Various types of intermittency in a single mode of the Q-G model

First we focus on a model with only one single mode. We pick the first Fourier mode $k = 1$ of the β-plane Q-G model. Since $\widehat{T}_{k,t}$ has its real and imaginary parts symmetric, it suffices for us to focus on its real part. By Theorem 4.1, as $\epsilon \to 0$, for sufficiently large t, the distribution of $\mathrm{Re}(\widehat{T}_{k,t})$ can be approximated by:

$$\mathbb{P}(\mathrm{Re}(\widehat{T}_{1,t}) \in \mathrm{d}x) = \int_{-\infty}^{\infty} \frac{\exp(-\frac{u^2}{2E_U} - \frac{x^2}{\tilde{\Sigma}_1(c_1u+d_1)})}{\pi\sqrt{2E_U\tilde{\Sigma}_1(c_1u+d_1)}}\mathrm{d}u, \quad \tilde{\Sigma}_1(\omega) = \frac{\alpha^2 E_{v,1}}{\gamma_{T,1}^2 + \omega^2}. \tag{5.1}$$

First, we will fix the stochastic energy of $\hat{v}_{1,t}$ to be $E_{v,1} = 1$, and U_t to be the O.U. process $\mathrm{d}U_t = -U_t\mathrm{d}t + \mathrm{d}W_t$.

First, we validate the limiting behavior predicted by Theorem 4.1. In subplots a), b) and c) of Figure 5.1, the histograms of $\widehat{T}_{v,1}(t)$ in (4.1) are plotted with $\epsilon = 0.1, 0.01$ and 0.001, while the limiting density given by formula (5.1) is also given for comparison. The convergence of histograms towards the limiting distribution is evident.

Moreover, it is worth noticing that both the histogram and the limiting distribution have exponential like tails in a range roughly between $10^{-2} - 10^{-4}$. In a rough sense, the tail distribution of (5.1) has its shape determined by a Gaussian tail with peak conditional covariance

$$\sup_u \tilde{\Sigma}_{1,t}(u) = \frac{\alpha^2 E_{v,1}}{\gamma_{T,1}^2},$$

which is approximately 100 in our parameters setting. Therefore if we focus on the tail $\mathbb{P}(\mathrm{Re}\widehat{T}_{1,t} \in d\lambda)$ for $\lambda \in [5, 15]$, which is a range of practical interest, then the curve $\exp(-\lambda^2/200)$ looks very much like an exponential one. In other words, the claims of Propositions 3.2 and 3.3 do not contradict with the exponential fat tails observed through experiments or simulations, but rather fit quite well.

In a time sequential perspective, Theorem 4.1 also claims that $\mathrm{Re}(\widehat{T}_{1,t})$ should be distributed roughly as $\mathcal{N}(0, \frac{1}{2}\Sigma_{1,t|U})$. Since the maximum conditional covariance is reached at $\omega_{1,R}(t) = 0$, this indicates that the spikes of $\widehat{T}_{1,t}$ should generally occur when $\omega_{1,R}(t) = 0$. In the time sequence subplots e) and f) of Figure 5.1, the snapshot of $\mathrm{Re}(\widehat{T}_{1,t})$ and $\omega_{1,R}(t)$ with $\epsilon = 0.001$ are presented within time $[0, 100]$. We can see clearly that the spikes are correlated with $\omega_{1,R}(t)$'s visit to 0. A more rigorous test of our claim will be checking the Gaussianality of the data $\mathrm{Re}(\widehat{T}_{1,t})/\sqrt{\frac{1}{2}\Sigma_{1,t|U}}$, which is carried out by the Quantile-Quantile subplot d) of Figure 5.1. We see the distribution of $\mathrm{Re}(\widehat{T}_{1,t})/\sqrt{\frac{1}{2}\Sigma_{1,t|U}}$ is Gaussian up to 0.1% error.

On the other hand, Theorem 4.1 also provides us a general guideline for the type of intermittency created by the cross sweep U_t, which depends on the stochastic energy E_U. The impact of U_t over $\widehat{T}_{k,t}$ is through the resonance phase $\omega_{R,k}(t) = c_kU_t + d_k$, with

$c_k = -a_k - k, d_k = -b_k - k\bar{u}$. Therefore the density function of $\omega_{k,R}(t)$ is clearly:

$$\mathbb{P}(\omega_{R,k}(t) \in \mathrm{d}\lambda) = \frac{1}{c_k\sqrt{2E_U\pi}} \exp\left(-\frac{(\lambda - d_k)^2}{2c_k^2 E_U} \right).$$

Since the resonance condition, $\omega_{R,k}(t) = 0$, is the source of the spikes, by the maximum likelihood principle, we would expect E_U that maximizes the density above at $\lambda = 0$ would produce stronger and more frequent intermittency in $\widehat{T}_{k,t}$. By taking the gradient, one can find the density function at $\lambda = 0$ is maximized at

$$E_U^* = \frac{m^2}{2q^2}.$$

We verify this rule of thumb in Figure 5.2 by plotting the limiting distribution (5.1) for $E_U = 0.1E_U^*, E_U = E_U^*, E_U = 10E_U^*$. As shown by subplots a) and b), the $E_U = E_U^*$ has the fattest tail. We also simulate system (4.1) with different σ_U in different energy regimes. It can be seen that $E_U = 0.1E_U^*$ rarely gives any spikes, $E_U = E_U^*$ gives frequent and long lasting spikes, $E_U = 10E_U^*$ has frequent spikes but rather short lived.

5.2. Q-G models with multiple Fourier modes

Here we consider the combined effect of multiple Fourier modes upon the tracer field $T_t(x)$. As a simple demonstration, we simulate the system (4.1) with $k \in \{-5, \ldots, 5\}$ and $\epsilon = 0.001$. Since we assume $\int T_t(x)\mathrm{d}x = 0$ after a normalization, the 0-th mode is constantly 0. We keep U_t to be the O.U. process $\mathrm{d}U_t = -U_t\mathrm{d}t + \mathrm{d}W_t$. The other system parameters are the same as in the beginning of this section. We set the energy spectrum of the shear flow $\hat{v}_{k,t}$ to be either equipartition or following the Kolmogorov spectrum:

$$E_{k,v} = 1 \quad \text{or} \quad k^{-\frac{5}{3}}.$$

The result with equipartition energy is presented in Figure 5.3. Through the histogram of $T_t(x)$ and the Q-Q plot of $T_t(x)/\tilde{\Sigma}_t(U_t)$, we find the limiting distribution of $T_t(x)$ is clearly given by the conditional Gaussian distribution $\mathcal{N}(0, \tilde{\Sigma}_t(U_t))$ with more than 99% accuracy, validating Theorem 4.2. The combination of various Fourier modes produces a much richer dynamics, since different Fourier modes are excited at different resonance conditions:

$$\omega_{R,k}(t) = 0 \quad \Leftrightarrow \quad U_t = \omega_k := -\frac{b_k + k\bar{u}}{a_k + k} = -\frac{1}{F}(k^2\bar{u} + F\bar{u} + \beta).$$

This can first be seen in the plot of $\tilde{\Sigma}_t(u)$, where we see that there are 5 peaks of similar heights corresponds to 5 different modes. We can also check the time sequence of $T_t(x)$ and U_t. We immediately see that the spikes of the $T_t(x)$ are correlated with U_t's visits to ω_k's. Moreover, as each wavenumber has the same energy, visits to different ω_k create the same amplitudes of spikes. For example, in subplot e) around time $t = 90$, U_t crosses several resonance phases ω_k, creating a long lasting burst in $T_t(x)$.

In Figure 5.4, the case with $E_{v,k}$ being the Kolmogorov spectrum is studied. The situation is very similar to the one of equipartition. The only difference is that different resonances create spikes of very different size, as reflected by the subplot b) of Figure 5.4. Therefore U_t's visits to the resonance phase of the first Fourier mode creates much larger intermittency in the system. Notice in subplot e) around time $t = 20$, U_t crosses several resonance phases ω_k, but the resulting burst in $T_t(x)$ is of mediocre size, while the strongest bursts in $T_t(x)$ are created by visits of U_t to ω_1 only. This is very different from Figure 5.3.

5.3. Synchronized excitation in random cross sweep model

Here we consider the random cross sweep model, which is a stochastic version of the model in [10] and mentioned earlier in Section 2. The damping and advection parameters are simply

$$\gamma_{v,k} = d_v + \nu k^2, \quad a_k = b_k = 0.$$

We also set the energy spectrum of v to be equipartition. Unlike the Q-G model we considered in the last subsection, the resonance phase on all Fourier modes are the same and synchronized automatically:

$$\omega_k = -\frac{b_k + k\bar{u}}{a_k + k} = -\bar{u}.$$

In other words, all the Fourier modes will be excited with large conditional variance when $U_t = -\bar{u}$. The combined effect on $T_t(x)$ is the possibility of enormous spikes. Yet these spikes are relatively short lived comparing to the case of the Q-G model, since $T_t(x)$ will decay quickly to its normal value when U_t passes $-\bar{u}$. This is exactly what we see in subplot d) of Figure 5.5. Notice the y-axis here scales differently from the one in Figure 5.3.

5.4. Random cross sweep model in a non asymptotic regime

Proposition 3.4 produces a conditional lower bound for $\Sigma_{k,t|U}$, especially when γ_U is small. In a special dynamical regime of the random cross sweep model given by (3.3):

$$k = 10, \quad \gamma_{T,k} = 0.2, \quad \gamma_{v,k} = 1, \quad \gamma_U = 10^{-5}, \quad \bar{u} = 10, \quad E_U = E_{v,k} = 100,$$

direct computation using Proposition 3.4 shows that conditioned on $\omega_{R,k}(t) = 0$, $\Sigma_{k,t|U}$ is more than 8 times of its mean value. We numerically verify the predicted phenomenon in Figure 5.6. As we can clearly see in the time sequence, there is obvious intermittency in $\widehat{T}_{k,t}$, which is correlated with $\omega_{R,k}$'s visits to 0. Note that since γ_U is very small, the simulation time span has been extended to $[0, 10^5]$.

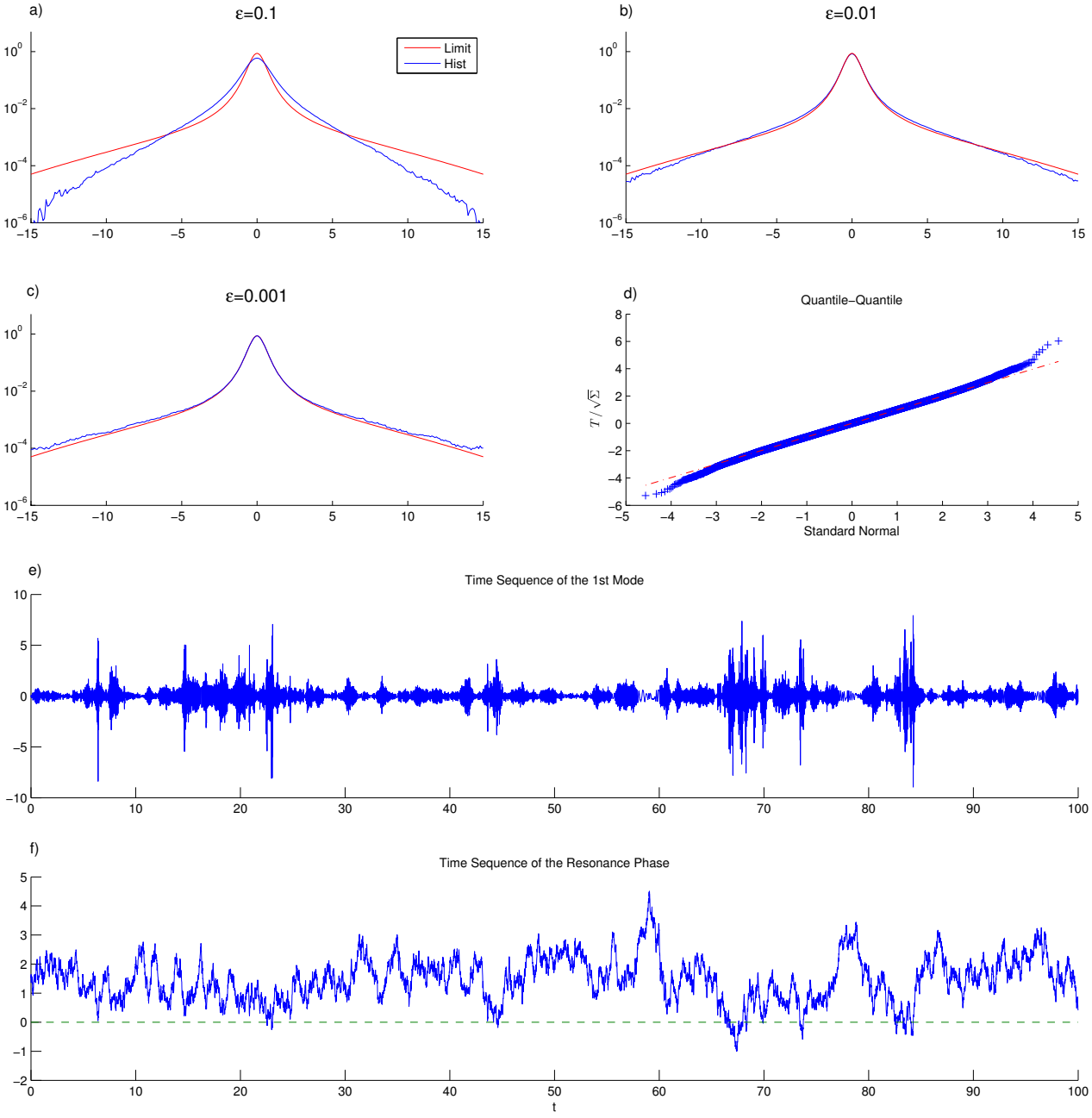

Figure 5.1: The first Fourier mode of system (4.1). In subplots a), b) and c) we have the histograms of $\mathrm{Re}(\widehat{T}_{1,t})$ for $\epsilon = 0.1, 0.01, 0.001$ with the limiting distribution (5.1) as reference. Subplot d) is the Q-Q plot of $\mathrm{Re}(\widehat{T}_{1,t})/\sqrt{\frac{1}{2}\Sigma_{1,t|U}}$ v.s. standard Gaussian for $\epsilon = 0.001$. e) is the time sequence of $\mathrm{Re}(\widehat{T}_{1,t})$ and f) is the time sequence of $\omega_{R,1}(t)$ with $t \in [0, 100]$.

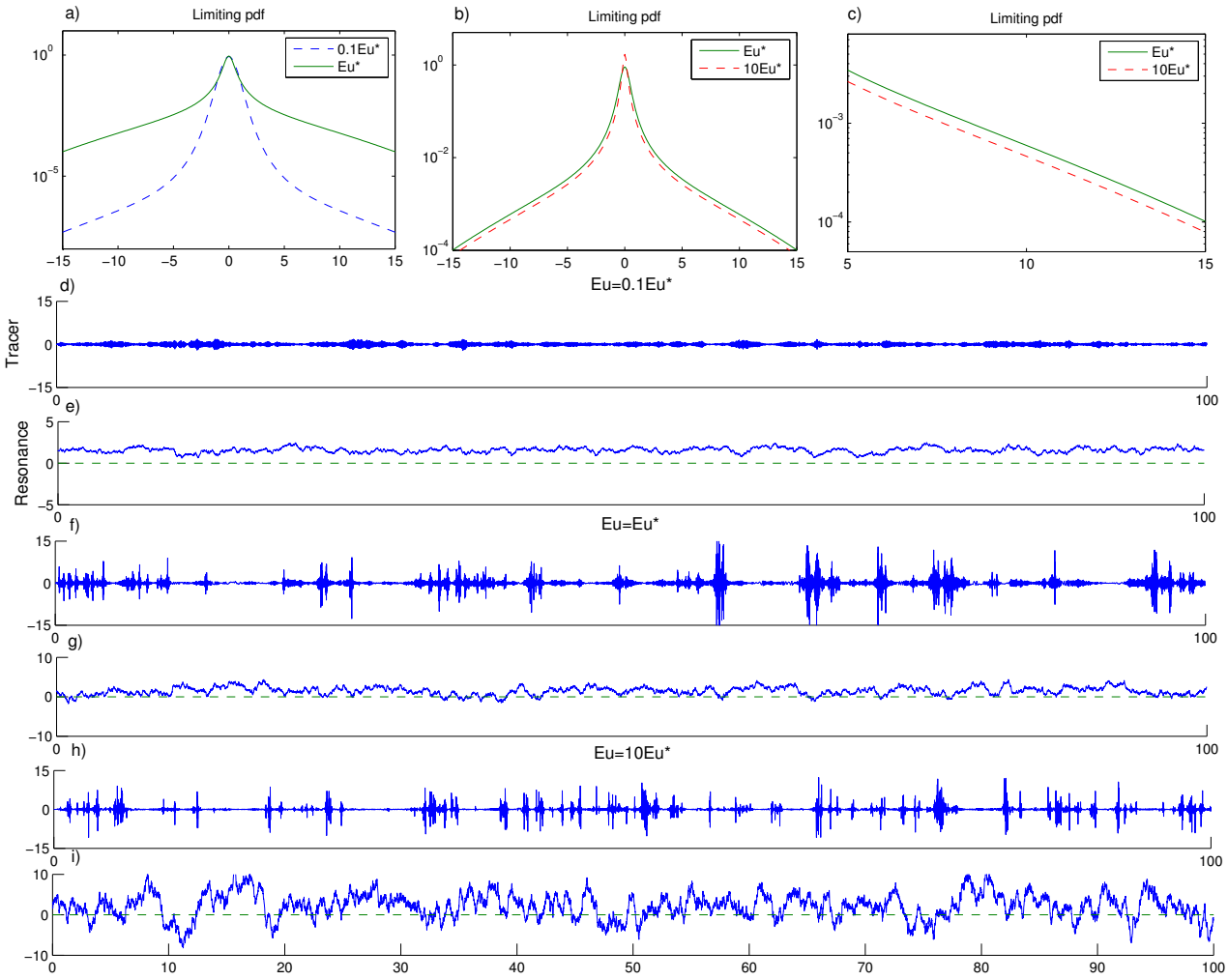

Figure 5.2: $\mathrm{Re}(\widehat{T_t})$ with different stochastic energy E_U. Subplots a) and b) are the limiting distributions of $T_t(x)$ for three different energy regime, while c) is the magnification of b)'s tail between $[5, 10]$. Subplots d), f), h) are time sequences of $T_t(x)$ when the cross sweep U_t has stochastic energy $E_U = 0.1E_U^*, E_U^*, 10E_U^*$. Subplots e), g), i) are time sequences of U_t in corresponding energy regime.

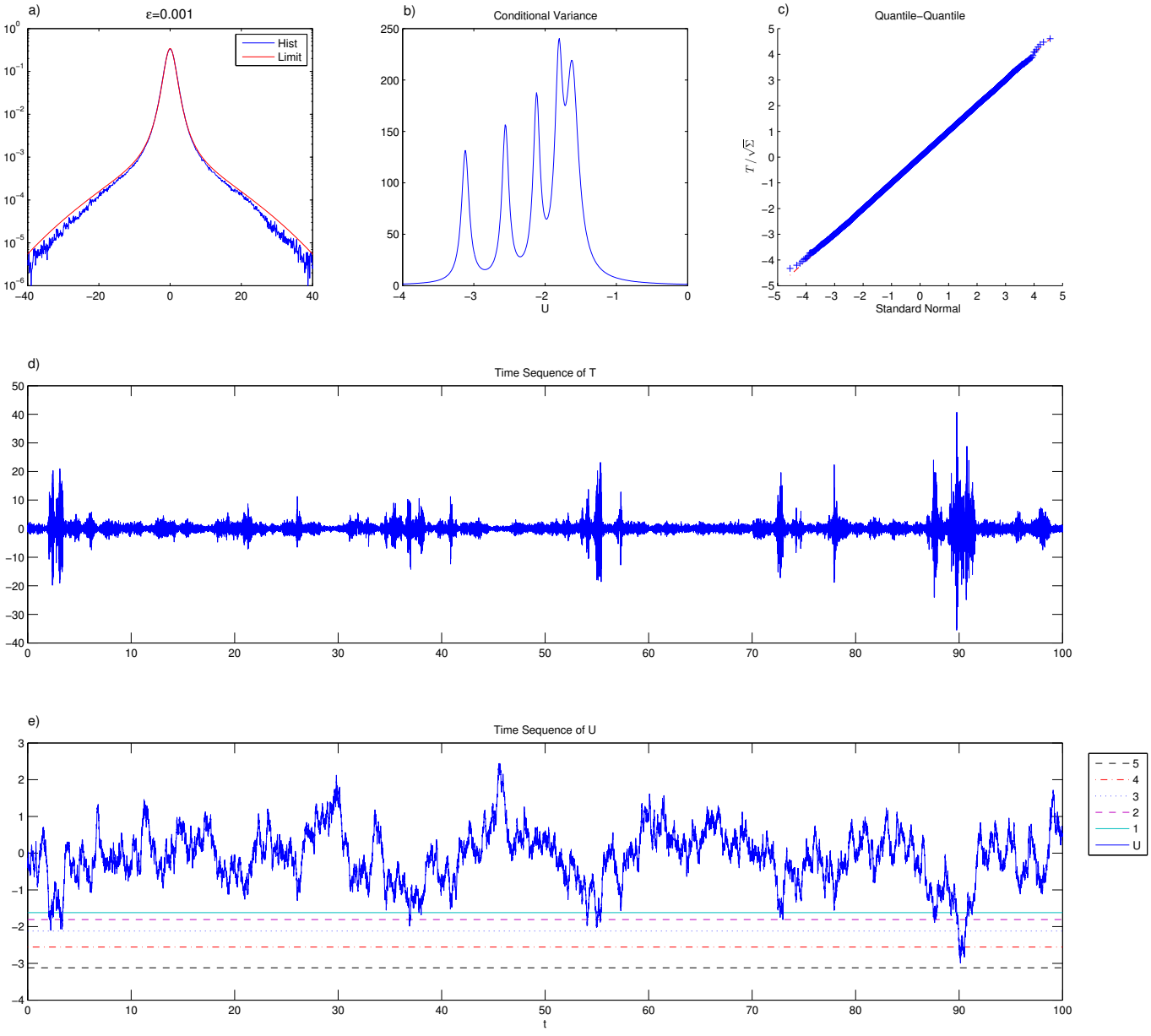

Figure 5.3: Distribution of $T_t(0)$ in Q-G model with $\epsilon = 0.001$ for equipartition energy in $E_{v,k}$. Subplot a) compares the histogram of $T_t(0)$ with the limiting distribution produced by Theorem 4.2. Subplot b) demonstrates the dependence of limiting conditional variance of $T_t(x)$ over U. Subplot c) is the Q-Q plot of $T_t(x)/\sqrt{\tilde{\Sigma}_t(U_t)}$ against standard normal. e) is the time sequence of $\widehat{T}_t(0)$ and f) is the time sequence of U_t with $t \in [0, 100]$, the reference lines in f) indicates the resonance phase ω_k.

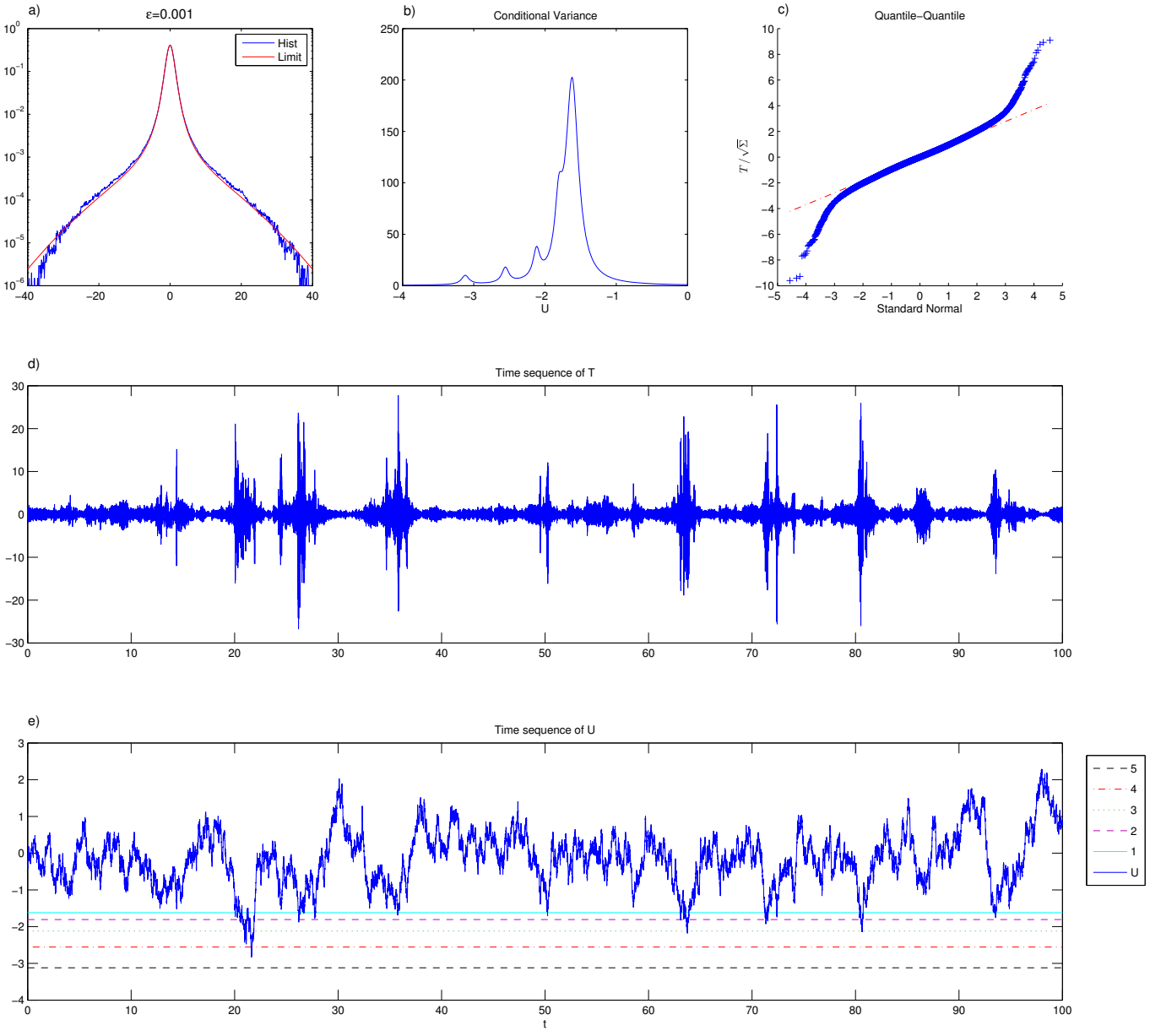

Figure 5.4: Distribution of $T_t(x)$ in Q-G model with $\epsilon = 0.001$ for Kolmogorov spectrum in $E_{v,k}$. Subplot a) compares the histogram of $T_t(0)$ with the limiting distribution produced by Theorem 4.2. Subplot b) demonstrates the dependence of limiting conditional variance of $T_t(x)$ over U. Subplot c) is the Q-Q plot of $T_t(x)/\sqrt{\tilde{\Sigma}_t(U_t)}$ against standard normal. e) is the time sequence of $\widehat{T}_t(0)$ and f) is the time sequence of U_t with $t \in [0, 100]$, the reference lines in f) indicates the resonance phase ω_k.

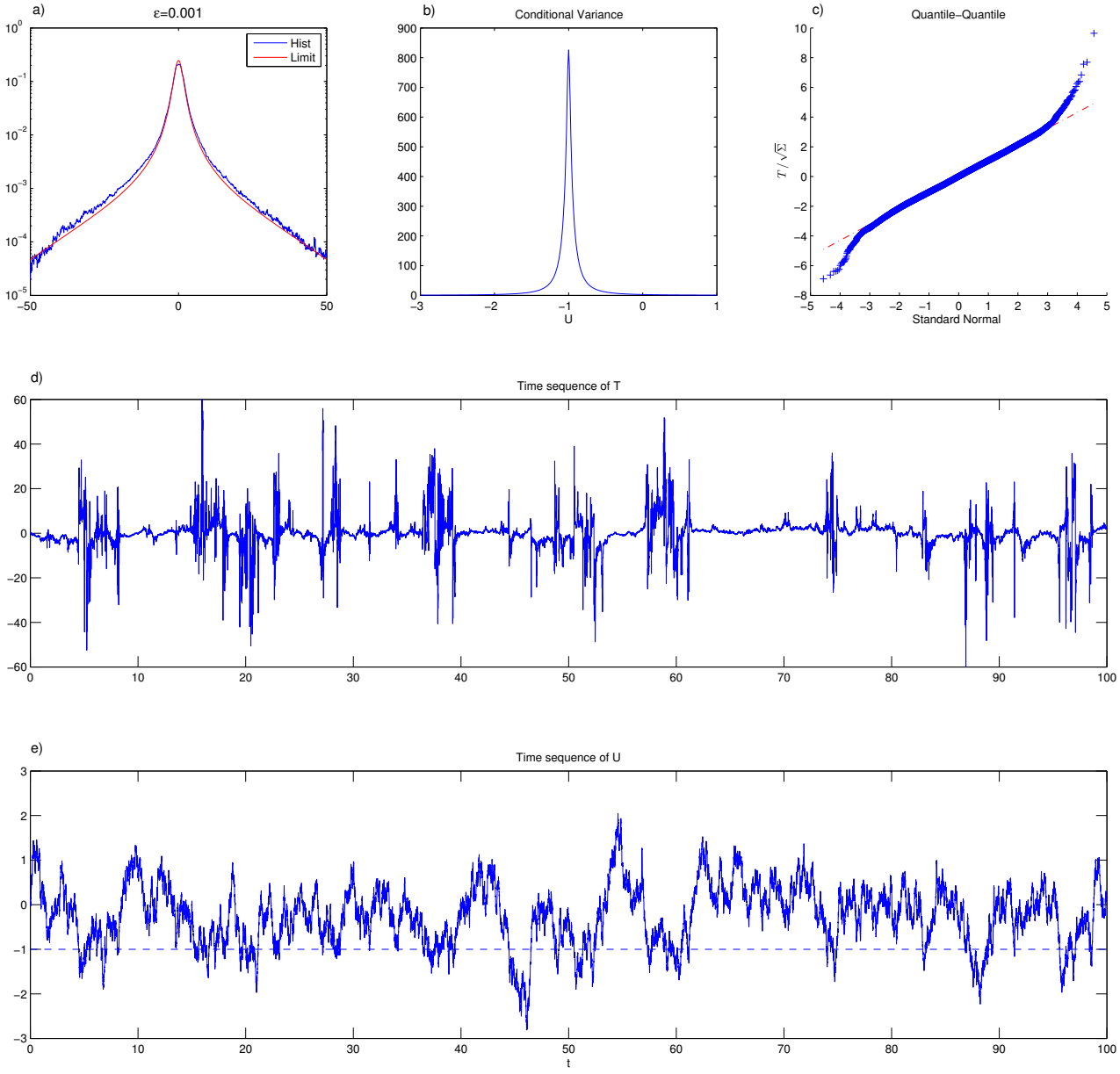

Figure 5.5: Tracer field $T_t(x)$ in random cross sweep model with $\epsilon = 0.001$ for equipartition spectrum in $E_{v,k}$. Subplot a) compares the histogram of $T_t(0)$ with the limiting distribution produced by Theorem 4.2. Subplot b) demonstrates the dependence of limiting conditional variance of $T_t(x)$ over U. Subplot c) is the Q-Q plot of $T_t(x)/\sqrt{\tilde{\Sigma}_t(U_t)}$ against standard normal. e) is the time sequence of $\widehat{T}_t(0)$ and f) is the time sequence of U_t with $t \in [0, 100]$, the reference lines in f) indicates the resonance phase ω_k.

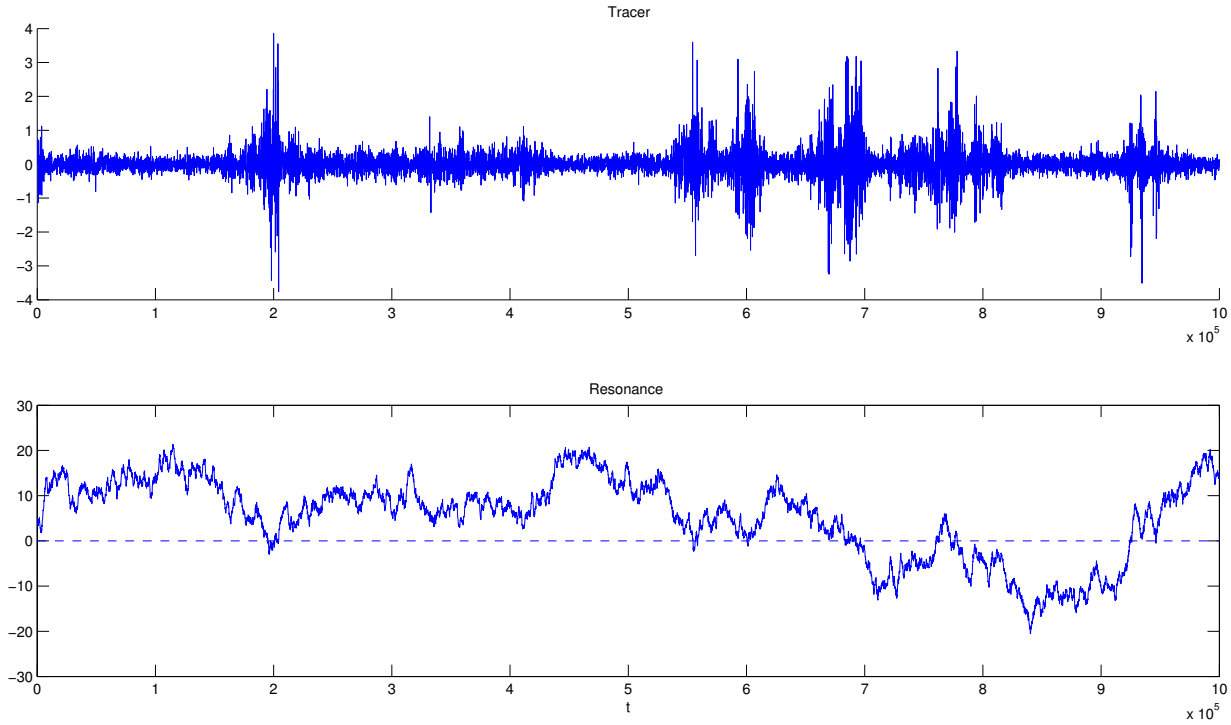

Figure 5.6: $k = 10$-th mode in a random cross sweep model with parameters given by (3.3). The upper subplot is the time sequence of $\mathrm{Re}(\widehat{T}_{k,t})$, while the lower subplot is the resonance phase $\omega_{R,k}(t)$.

6. Generalizations

6.1. Nonlinear cross sweeps

The requirement that the cross sweep U_t is an O.U. process is not necessary for our claims to hold. Actually, we can assume U_t to be a general diffusion process:

$$\mathrm{d}U_t = f_U(U_t)\mathrm{d}t + \sigma_U(U_t)\mathrm{d}W_t. \tag{6.1}$$

An interesting concrete example is the canonical scalar model with cubic nonlinearity [13, 14] with formulation:

$$\mathrm{d}U_t = [F + aU_t + bU_t^2 - cU_t^3]\mathrm{d}t + \sqrt{(A - BU_t)^2 + \sigma^2}\mathrm{d}W_t, \quad c, \sigma > 0. \tag{6.2}$$

Such a model arises naturally in low frequency reductions of large-scale climate models [13]. Such a process has a wide variety of different regimes including intermittency of various types, which are described with details in [13, 14]. In order to include these cross sweeps into our framework, we apply classical results [22, 23] and make the following assumption:

Assumption 6.1. *There exists a positive function $\mathcal{E}$, known as a Lyapunov function, such that:*

- *For any $M > 0$, the sub-level set $\{u : \mathcal{E}(u) \leq M\}$ is compact.*
- *$\mathcal{E}$ is locally dissipative, i.e. there are constants $\gamma, k_v > 0$ such that:*

$$f_U(u)\frac{\partial \mathcal{E}(u)}{\partial u} + \frac{1}{2}\sigma_U^2(u)\frac{\partial^2 \mathcal{E}(u)}{\partial u^2} \leq -\gamma \mathcal{E}(u) + k_v.$$

- *There exists a constant M such that $|f_U(u)|^2, |\sigma_U(u)|^2 \leq M\mathcal{E}(u)$.*
- *The stochastic forcing is non-degenerate, i.e. $|\sigma_U|$ is strictly positive.*

For an O.U. process $\mathrm{d}U_t = -\gamma_U U_t \mathrm{d}t + \sigma_U \mathrm{d}W_t$, it is easy to see $\mathcal{E}(u) = u^2$ is sufficient. For the concrete example (6.2), one can verify $\mathcal{E}(u) = u^6$ is a proper Lyapunov function using Hölder's inequality. A short treatment of Lyapunov function is given in Section Appendix A.5 with detailed verification of the claims above. With Assumption 6.1, Theorem 2.3 of [22] shows that U_t will be geometrically ergodic. The density of the invariant measure π_U can be formally computed through the Fokker Plank equation, which is the solution to

$$\frac{\partial}{\partial u}(f_U \pi_U) = \frac{1}{2}\frac{\partial^2}{\partial u^2}(\sigma_U^2 \pi_U).$$

As shown in [13, 14], the invariant measure of (6.2) has an explicit density:

$$\pi_U(u) = \frac{N_0}{((Bu - A)^2 + \sigma_u^2)^{a_1}} \exp\left(d_1 \arctan\left(\frac{Bu - A}{\sigma_u}\right)\right) \exp\left(-\frac{c_1 u^2 + b_1 u}{B^4}\right),$$

where $N_0, a_1, b_1, c_1, d, d_1, d_2$ are constants determined by the parameterization of (6.2). One thing worth noticing is the tail of π_U is Gaussian like.

Since we can always move the mean of U_t into the term $\bar{u}$, without loss of generality we assume that $\int u\pi_U(\mathrm{d}u) = 0$ here. In this general setting, all of our major claims actually remain true. The dynamics of the joint process can generally be written as:

$$\begin{aligned} \mathrm{d}U_t &= -f_U(U_t)\mathrm{d}t + \sigma_U(U_t)\mathrm{d}W_t; \\ \mathrm{d}\hat{v}_{k,t} &= [-\gamma_{v,k} + \mathrm{i}\epsilon^{-1}\omega_{v,k}(t)]\hat{v}_{k,t}\mathrm{d}t + \sigma_{k,v}\mathrm{d}B_{k,t}; \\ \mathrm{d}\widehat{T}_{k,t} &= \epsilon^{-1}[(-\gamma_{T,k} + \mathrm{i}\omega_{T,k}(t))\widehat{T}_{k,t}\mathrm{d}t - \alpha\hat{v}_{k,t}]\mathrm{d}t; \\ \omega_{k,V}(t) &= a_k U_t + b_k, \quad \omega_{k,T}(t) = -k(U_t + \bar{u}). \end{aligned} \tag{6.3}$$

The first generalization is the geometric ergodicity of (6.3):

Theorem 6.2. *For any fixed $\epsilon > 0$, let the turbulent diffusion system* (6.3) *consist of Fourier modes in a finite wavenumber set N, then the joint process $(U_t, \hat{v}_{k,t}, \widehat{T}_{k,t}, k \in N \cap \mathbb{Z}^+)$ is geometrically ergodic under the total variation norm. Specifically, for any initial measure μ of $(U_0, \hat{v}_{k,0}, \widehat{T}_{k,0}, k \in N \cap \mathbb{Z}^+)$, there is a unique invariant measure π, and constants $C, \beta > 0$ such that the following holds:*

$$\|\mathbb{P}_t^\mu - \pi\| \leq Ce^{-\beta t}\mathbb{E}^\mu(1 + V(U_0) + \sum_{k\in N} |\hat{v}_{k,0}|^2 + |\widehat{T}_{k,0}|^2).$$

Proof. This is a standard application of the framework set up in [22, 23]. The details are given in Section Appendix A.1. □

Our second result is a generalization of Theorems 4.1 and 4.2:

Theorem 6.3. *For any fixed $t > 0$, let $(U_t, \hat{v}_{k,t}, \widehat{T}_{k,t})$ follow* (6.3) *satisfying Assumption 6.1. Assume the joint process is initialized with $U_0 \sim \mathcal{N}(0, E_U), \hat{v}_k = 0, \widehat{T}_k = 0$. Then conditioned on $\mathbb{P}$-a.s. realization of $U_{s\leq t}$, as $\epsilon \to 0$, the claims of conditional variance's convergence in Theorem 4.1 and 4.2 hold, while* (4.2) *is replaced by:*

$$\mathbb{P}(Re(\widehat{T}_{k,t}) \in \mathrm{d}x, Im(\widehat{T}_{k,t}) \in \mathrm{d}y) = \int_{-\infty}^{\infty} \pi_U(\mathrm{d}u) \frac{\exp(-\frac{x^2+y^2}{\tilde{\Sigma}_{k,t}(c_k u + d_k)})}{\pi\tilde{\Sigma}_{k,t}(c_k u + d_k)}.$$

And formula (4.6) *is replaced by:*

$$\mathbb{P}(T_t(x) \in \mathrm{d}\lambda) = \int \pi_U(\mathrm{d}u) \frac{1}{\sqrt{2\pi\tilde{\Sigma}_t(u)}} \exp\Big(-\frac{\lambda^2}{2\tilde{\Sigma}_t(u)}\Big).$$

Here $\tilde{\Sigma}_{k,t}$, $\tilde{\Sigma}_t$ are the limits of conditional variance given by Theorems 4.1 and 4.2.

Proof. The proof is identical to the one of Theorems 4.1 and 4.2, since the only part that uses the formulation of U_t is proved under our general conditions here by Lemma Appendix A.4. □

6.2. Infinite Fourier modes

The theoretical framework we established can be easily generalized to models with countably many Fourier modes, e.g. $T_t(x) = \sum_{k\in N} e^{\mathrm{i}kx}\widehat{T}_{k,t}$ with $N = \mathbb{Z}$, as long as certain summability conditions are met. First, in order for the Birkhoff ergodic theorem to work, it suffices to show the uniqueness of invariant measure for the joint system $(U_t, \hat{v}_{k,t}, \widehat{T}_{k,t}, k \in N \cap \mathbb{Z}^+)$. However, for any two invariant measures μ and ν of the joint system, and any finite subset $M \subset N \cap \mathbb{Z}^+$, consider the restriction of μ and ν on the σ-field

$$\mathcal{F}_M = \sigma\{U_t, \hat{v}_{k,t}, \widehat{T}_{k,t}, k \in M\}.$$

Since the truncation $(U_t, \hat{v}_{k,t}, \widehat{T}_{k,t}, k \in M)$ is itself a Markov process, by Theorem 3.1 or 6.2 and the invariance of μ and ν, we find

$$\|\mu|_{\mathcal{F}_M} - \nu|_{\mathcal{F}_M}\| = \left\|\mathbb{P}_t^{\mu|_{\mathcal{F}_M}} - \mathbb{P}_t^{\nu|_{\mathcal{F}_M}}\right\| \xrightarrow{t\to\infty} 0.$$

In other words, μ and ν are identical on any finite Galerkin truncation of the full process, therefore they must the same measure, thus we have

Theorem 6.4. *Under the conditions of Theorem 6.2, the joint system* $(U_t, \hat{v}_{k,t}, \widehat{T}_{k,t}, k \in N \cap \mathbb{Z}^+)$ *has a unique invariant measure, i.e. it is ergodic.*

Infinite dimensional systems with their finite dimensional truncation converging to a corresponding distribution under total variation norm, such as the turbulent diffusion system here, are generally known as locally mixing [30] or locally ergodic [31]. One point worth noticing is that Theorem 6.4 does not require any summable conditions, since it separates each Fourier mode into a different dimension. So even systems with $\sum_k |\hat{v}_{k,t}|^2 = \infty$ are in the range of our discussion, though $v(t,x) = \sum e^{\mathrm{i}kx}\hat{v}_{k,t}$ will be ill defined. On the other hand, if one poses certain summability condition on the parameters, the joint process will likely converges to its invariant measure under certain Wasserstein norms [32, 33, 34]. But this is beyond the scope of the present paper.

The other result that requires the finiteness of the wavenumber set N in this paper is Theorem 4.2, since we need to check that $\sum_{k\in N} \mathbb{E}|\Sigma^\epsilon_{k,t} - \tilde{\Sigma}_{k,t}(\omega_{R,k}(t))| \to 0$ when $\epsilon \to 0$. Thanks to the careful treatment in the proof of Theorem 4.1 with a special version of Cauchy Schwartz, it suffices for us to show the summability of term (4.5) among all k. This leads to many different summability conditions that all generalize Theorem 4.2; the following version is the simplest and does hold for all examples mentioned in Section 2:

Theorem 6.5. *Suppose* $\gamma_{T,k}$ *is non-decreasing for* $k \in \mathbb{Z}^+$*, moreover assume that*

$$\sum_{k\in N} E_{v,k} < \infty, \quad \sum_{k\in N} \frac{\sigma^2_{v,k}}{\gamma_{T,k}} < \infty,$$

then under same conditions of Theorem 4.2 and 6.3, the same conclusions hold with countable set N *of Fourier modes.*

7. Conclusion and discussion

Intermittency is an important feature that generally exists in turbulent flows. It can be described as exceptionally large spikes in the time sequence, or an exponential like fat tail in the probability density. This paper studies a passive tracer with a mean gradient in turbulence models that consists of a zonal cross sweep and a meridional shear flow [10, 20, 5]. The intermittency in these models is especially subtle, since there are no positive Lyapunov exponents in these models, yet they exhibit realistic intermittency. By transforming the dynamics into the Fourier domain, we find the solution is conditionally Gaussian given the realization of the cross sweep U_t. Therefore the behavior of the passive tracer is controlled by $U_{s\le t}$ through the conditional covariance, which is given by the integral formula in Proposition 3.2. The randomness in $U_{s\le t}$ makes this conditional covariance highly fluctuating, as the peak value could be 100 times larger than the mean value in some dynamical regimes, according to the analysis of Proposition 3.3. This potentially can generate the large spikes in the time sequence of a passive tracer and fat tails in the distribution. On the other hand, by analyzing the formula of the conditional covariance, we found that the large fluctuation is caused by random resonance effects between the phase speed of the shear flow and the passive tracer. In Theorems 4.1 and 4.2, such intuition is rigorously proved in a long time slow varying rescaling limit of the original model, while the limiting conditional covariance has an explicit and concise dependence on the cross sweep U_t. Through this result, it is found that U_t's visits to different resonance phases excites different Fourier modes of the passive tracer field, which in turn produce intermittent spikes. Numerical simulations are presented to verify our theoretical results. They also show very different intermittent behaviors in various dynamical regimes, which all can be well explained by the theory presented here. Moreover, the simulations here can be done for only one long enough realization, since Theorem 3.1 guarantees that the turbulent diffusion model here is geometrically ergodic. Two generalizations are made at the end for cross sweeps with general formulations and the case where there are infinitely many Fourier modes.

Although this paper gives an intuitive, rigorous and simple explanation of the intermittency, there are many related questions which remain unanswered. Here are a list of directions that wait further exploration:

(i) This paper hinges heavily on the conditional Gaussian structure of the special turbulent diffusion model. Yet there are many intermittent systems that lack such structure, e.g. the canonical scalar model with cubic nonlinearity (6.2). It will be very interesting if we can apply the conditional covariance fluctuation argument presented here to other models. Actually, a broader question will be: how can we characterize intermittent behaviors without looking at the time sequence or tail distribution.

(ii) On the other hand, conditional Gaussian structure could be exploited in a wide range of models as well. It can be seen as a middle land between nonlinear dynamics and linear dynamics, where it is very possible to apply many well understood tools in linear

dynamics to study nonlinear systems, e.g. [25, 26]. It is a general question that many features of these linear tools will change in these applications. Another general question is how can we detect whether a system has a conditional Gaussian structure based on its phenomenon or data.

(iii) It seems the estimate of conditional covariance in this paper can be improved for non-asymptotic settings. In order to show large fluctuations in conditional covariance, especially showing the peak values is many times the mean, we need a lower bound for the peak value and a upper bound for the mean. Proposition 3.4 is a prototype lower bound, since it is designed to work assuming the cross sweep is slow varying. The reason that we often assume the cross sweep is slow varying is because then the conditional covariance depends highly on the end value of the cross sweep, of which the distribution is well known. The authors conjecture that a more general non asymptotic result can be produced if we focus on U_t's path in an interval rather than on the end point.

(iv) One important application problem of stochastic system is filtering and prediction. Intermittent systems are very difficult to predict in general for their enormous spikes. Yet, the prediction of these rare events are very important in real world applications. The results of this paper may shed some light on this direction.

Acknowledgments

This research is supported by the MURI award grant N-000-1412-10912, where A.J.M. is the principal investigator, while X.T.T. is supported as a postdoctoral fellows.

Appendix A. Proofs for the theorems

Appendix A.1. geometric ergodicity

Proof of Theorems 3.1 and 6.2. Since Theorem 6.2 has a more general condition, it suffices to prove it only. By the results of [23, 22], for a general diffusion process in $\mathbb{R}^d$

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \Sigma(X_t)\mathrm{d}W_t$$

to be geometrically ergodic, it suffices to show the following three properties:

- There is a Lyapunov function V and constants $\gamma, k_v > 0$ so that V has all it sub-level sets being compact and

$$\mathcal{L}V(x) = b(x)\cdot\nabla V(x) + \frac{1}{2}\mathrm{tr}[\Sigma(x)(HV)\Sigma^t(x)] \le -\gamma V(x) + k_v.$$

  Here HV is the Hessian matrix of V.
- Denote the columns of Σ as $\Sigma_1, \ldots, \Sigma_m$, and L_0 as the Lie algebra generated by $\{\Sigma_i, [\Sigma_i, b], i = 1, \ldots, m\}$ with $[\ ,\ ]$ denotes the Lie bracket. Then L_0 spans $\mathbb{R}^d$.

- There is an $x^* \in \mathbb{R}^d$ so that for any fixed $x \in \mathbb{R}^d, \epsilon > 0$, there is a smooth process $w_{s\le t}$ and time t_0, such the solution to

$$\mathrm{d}x_s = b(x_s)\mathrm{d}s + \Sigma(x_s)w_s\mathrm{d}s, \quad x_0 = x,$$

  satisfies $|x_{t_0} - x^*| < \epsilon$.

First of all, we construct a Lyapunov function by letting

$$V_t = \mathcal{E}(U_t) + \sum_{k\in N\cap\mathbb{Z}^+} \frac{|\hat{v}_{k,t}|^2}{\gamma_{v,k}} + \frac{\gamma_{T,k}}{\alpha^2}|\widehat{T}_{k,t}|^2.$$

Then the infinitesimal generator is bounded by some constants M_k and k_v:

$$\begin{aligned}\mathcal{L}V_t &= 2uf_U(U_t) + \sigma_U\sigma_U^*(U_t) + \sum_{k\in N\cap\mathbb{Z}^+}\left[-2|\hat{v}_{k,t}|^2 + \frac{\sigma_{v,k}^2}{\gamma_{v,k}} - \frac{2\gamma_{T,k}^2}{\alpha^2}|\widehat{T}_{k,t}|^2 + \frac{\gamma_{T,k}}{\alpha^2}\sigma_{T,k}^2 - \frac{2\gamma_{T,k}}{\alpha}\mathrm{Re}(\widehat{T}_{k,t}\hat{v}_{k,t})\right]\\ &\le -\gamma\mathcal{E}(U_t) - \sum_{k\in N\cap\mathbb{Z}^+}\left[|\hat{v}_{k,t}|^2 + \frac{\gamma_{T,k}^2}{\alpha^2}|\widehat{T}_{k,t}|^2 - M_k\right] \le -\tilde{\gamma}V_t + k_v.\end{aligned}$$

where we used the Young's inequality:

$$\frac{2\gamma_{T,k}}{\alpha}|\widehat{T}_{k,t}\hat{v}_{k,t}| \le |\hat{v}_{k,t}|^2 + \frac{\gamma_{T,k}^2}{\alpha^2}|\widehat{T}_{k,t}|^2$$

and $\tilde{\gamma} = \min\{\gamma, \gamma_{v,k}, \gamma_{T,k}, k \in N\} > 0$.

To verify the Lie bracket condition, we split the real and imaginary parts: $\hat{v}_{k,t} = \hat{v}_{k,t}^1 + i\hat{v}_{k,t}^2$, $\widehat{T}_{k,t} = \widehat{T}_{k,t}^1 + i\widehat{T}_{k,t}^2$. The joint dynamics can be rewritten as:

$$\mathrm{d}\begin{bmatrix} U_t \\ \vdots \\ \hat{v}_{k,t}^1 \\ \hat{v}_{k,t}^2 \\ \widehat{T}_{k,t}^1 \\ \widehat{T}_{k,t}^2 \end{bmatrix} = L\mathrm{d}t + \sigma_U(U_t)M\mathrm{d}W_t + \sum_k \frac{\sigma_{v,k}}{\sqrt{2}}(M_k^1\mathrm{d}B_{k,t}^1 + M_k^2\mathrm{d}B_{k,t}^2), \tag{A.1}$$

with

$$L = \begin{bmatrix} f_U(U_t) \\ \vdots \\ -\gamma_v\hat{v}_{k,t}^1 - \omega_v\hat{v}_{k,t}^2 \\ -\gamma_v\hat{v}_{k,t}^2 + \omega_v\hat{v}_{k,t}^1 \\ -\gamma_T\widehat{T}_{k,t}^1 - \omega_T\widehat{T}_{k,t}^2 - \alpha\hat{v}_{k,t}^1 \\ -\gamma_T\widehat{T}_{k,t}^2 + \omega_T\widehat{T}_{k,t}^1 - \alpha\hat{v}_{k,t}^2 \end{bmatrix}, \quad M = \begin{bmatrix} 1 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad M_k^1 = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad M_k^2 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Then it suffices to note that

$$[L, M_k^1] = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ -\alpha \\ 0 \end{bmatrix}, \quad [L, M_k^2] = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ -\alpha \end{bmatrix}.$$

So $\{M_k^1, M_k^2, [L, M_k^1], [L, M_k^2]\}$ obvious spans the linear space for the k-th Fourier modes, and M spans the dimension for U_t.

For the last reachability condition, we can simply let $x_0 = \vec{0}$ and let the control process be

$$w_s = [-\sigma_U(U_s)^{-1}(f_U(U_s) + U_s), 0, \ldots, 0].$$

In other words, we only control W_t in (A.1) and let $B_{k,t}^i$ simply be 0. The driven process simply follows:

$$\mathrm{d}U_s = -U_s \mathrm{d}s, \quad \mathrm{d}\hat{v}_{k,t} = [-\gamma_{v,k} + \mathrm{i}\omega_{v,k}]\hat{v}_{k,t}\mathrm{d}t, \quad \mathrm{d}\widehat{T}_{k,t} = [-\gamma_{T,k} + \mathrm{i}\omega_{T,k}]\widehat{T}_{k,t}\mathrm{d}t - \alpha\hat{v}_{k,t}\mathrm{d}t.$$

It's easy to see that U_s and $\hat{v}_{k,t}$ converge to 0 exponentially fast, then so does $\widehat{T}_{k,t}$ since by Duhamel's formula

$$\widehat{T}_{k,t} = \exp(-\gamma_{T,k}t + \mathrm{i}\omega_{T,k}[0,t])\widehat{T}_{k,0} - \alpha\int_0^t \exp(-\gamma_{T,k}(t-s) + \mathrm{i}\omega_{T,k}[s,t])\hat{v}_{k,s}\mathrm{d}s.$$

Its norm can be bounded by

$$|\widehat{T}_{k,t}| \leq \exp(-\gamma_{T,k}t)|\widehat{T}_{k,0}| + \alpha\int_0^t \exp(-\gamma_T(t-s))|\hat{v}_{k,s}|\mathrm{d}s.$$

So apparently it will converge to 0 as well. □

Appendix A.2. Bounds for sub-gaussian and variance

Proof of Proposition 3.3. As we focus on one wavenumber k at the moment, its appearance in the subscript will be suppressed. Denote:

$$\sigma_{r,s}^t = \alpha\sigma_{v,k}\exp(-\gamma_{T,k}(t-s) - \gamma_{v,k}(s-r) + \mathrm{i}\omega_{T,k}[s,t] + \mathrm{i}\omega_{v,k}[r,s]), \tag{A.2}$$

then $\Sigma_{t|U} = \int_0^t[\int_r^t |\sigma_{r,s}^t|ds]^2$. The upper bound for conditional covariance can be easily verified through the observation that $|\sigma_{r,s}^t| \leq \alpha\sigma_v \exp(-\gamma_T t + \gamma_v r)\exp(\gamma_R s)$, where we defined $\gamma_R = \gamma_T - \gamma_v$

$$\begin{aligned}\Sigma_{t|U} &\leq \sigma_v^2\alpha^2\int_0^t \exp(-2\gamma_T t + 2\gamma_v r)\left(\frac{\exp(2\gamma_R t) + \exp(2\gamma_R r)}{|\gamma_R|\gamma_R}\right)\mathrm{d}r \\ &\leq \frac{\sigma_v^2\alpha^2}{\gamma_R|\gamma_R|}\left[\frac{1-\exp(-2\gamma_v t)}{2\gamma_v} - \frac{1-\exp(-2\gamma_T t)}{2\gamma_T}\right].\end{aligned}$$

A side note here is that this bound is not optimal, but we choose it in order to make a better comparison with the second claim. Moreover the optimal upper bound will be of the same order anyway. This also bounds the Laplace transformation, since by Laplace transformation of Gaussian distribution:

$$\begin{aligned}\mathbb{E}\exp(\lambda_1\mathrm{Re}(\widehat{T}_t)+\lambda_2\mathrm{Im}(\widehat{T}_t)) &= \mathbb{E}\mathbb{E}(\exp(\lambda_1\mathrm{Re}(\widehat{T}_t)+\lambda_2\mathrm{Im}(\widehat{T}_t))|U_{s\le t})\\ &= \mathbb{E}\exp((\lambda_1^2+\lambda_2^2)\Sigma_{t|U}/4) \le \exp((\lambda_1^2+\lambda_2^2)\overline{\Sigma}_t/4).\end{aligned}$$

For the second claim, notice that

$$\left|\int_r^t \sigma_{r,s}^t \mathrm{d}s\right|^2 = \alpha^2\sigma_{v,k}^2\exp(-2\gamma_T t+2\gamma_v r)\int_r^t\int_r^t \exp(\gamma_R(s_1+s_2)+\mathrm{i}\omega_R[s_1,s_2])\mathrm{d}s_1\mathrm{d}s_2.$$

Here if $s_1>s_2$ then $\omega_R[s_1,s_2]=-\omega_R[s_2,s_1]$. Then as $\omega_R(t)=c_kU_t+d_k$ where

$$c_k=-(a_k+k),\quad d_k=-(b_k+k\bar{u}), \tag{A.3}$$

we easily find that $\mathbb{E}\omega_R(t)=d_k$ and $\mathrm{cov}(\omega_R(t),\omega_R(s))=\exp(-\gamma_U|s-t|)c_k^2E_U$. Thus,

$$\omega_R[s_1,s_2]\sim\mathcal{N}\bigg(d_k(s_2-s_1),\frac{c_k^2E_U}{\gamma_U^2}(\gamma_U|s_2-s_1|-1+e^{-\gamma_U|s_2-s_1|})\bigg).$$

Since for a Gaussian random variable X, $\mathbb{E}\exp(\mathrm{i}X)=\exp(\mathrm{i}\mathbb{E}X-\frac{1}{2}\mathrm{var}(X))$, we have

$$\mathbb{E}\exp(\mathrm{i}\omega_R[s_1,s_2])=\exp\bigg(\mathrm{i}d_k(s_2-s_1)-\frac{c_k^2E_U}{2\gamma_U^2}(\gamma_U|s_2-s_1|-1+e^{-\gamma_U|s_2-s_1|})\bigg).$$

Thus

$$\begin{aligned}&\int_0^t \mathbb{E}\exp((\gamma_T-\gamma_v)(s_2+s_1)+\mathrm{i}\omega_R[s_1,s_2])\mathrm{d}s_2\\ &=\exp(2(\gamma_T-\gamma_v)s_1)\int_{r-s_1}^{t-s_1}\exp((\gamma_T-\gamma_v)s-\mathrm{i}d_ks)\exp\bigg(-\frac{c_k^2E_U}{2\gamma_U^2}(\gamma_U|s|-1+\exp(-\gamma_U|s|))\bigg)\mathrm{d}s.\end{aligned}$$

We will apply elementary Lemma Appendix A.1 presented below and give an upper bound. Denote

$$f(x)=\exp\bigg(-\frac{c_k^2E_U}{2\gamma_U^2}(\gamma_Ux-1+\exp(-\gamma_Ux))\bigg).$$

By checking its first derivative, clearly $f(x)$ is positive and decreasing on $\mathbb{R}_+$, therefore by Lemma Appendix A.1 below

$$\begin{aligned}&\left|\int_r^t \mathbb{E}\exp(\gamma_R(s_2+s_1)+\mathrm{i}\omega_R[s_1,s_2])\mathrm{d}s_2\right|\\ &=\exp(2\gamma_Rs_1)\left|\int_0^{t-s_1}\exp(\gamma_Rs-\mathrm{i}d_ks)f(s)\mathrm{d}s+\int_0^{s_1-r}\exp(-\gamma_Rs+\mathrm{i}d_ks)f(s)\mathrm{d}s\right|\\ &\le\frac{2\exp(\gamma_Rt+\gamma_Rs_1)+4\exp(2\gamma_Rs_1)+2\exp(\gamma_Rr+\gamma_Rs_1)}{\sqrt{\gamma_R^2+d_k^2}}.\end{aligned}$$

Integrating both sides with $s_1 \in [r,t]$ gives us

$$\mathbb{E}\int_r^t\int_r^t \exp((\gamma_T-\gamma_v)(s_1+s_2)+\mathrm{i}\omega_R[s_1,s_2])\mathrm{d}s_1\mathrm{d}s_2 \leq \frac{4(\exp(2\gamma_R t)-\exp(2\gamma_R r))}{\gamma_R\sqrt{\gamma_R^2+d_k^2}}.$$

Integrating both sides with $r \in [0,t]$ gives us:

$$\mathbb{E}\Sigma_{t|U} \leq \frac{\alpha^2\sigma_v^2}{\gamma_R\sqrt{\gamma_R^2+d_k^2}}\left[\frac{1-\exp(-2\gamma_v t)}{2\gamma_v}-\frac{1-\exp(-2\gamma_T t)}{2\gamma_T}\right].$$

□

Lemma Appendix A.1. *Let $f(s)$ be a positive decreasing C^1 function and a,b to be two real constants, then*

$$\left|\int_r^t e^{(a+\mathrm{i}b)s}f(s)\mathrm{d}s\right| \leq 2\frac{\exp(ar)+\exp(at)}{\sqrt{a^2+b^2}}f(r).$$

Proof. This can directly be obtained through integration by part:

$$\begin{aligned}\left|\int_r^t e^{(a+\mathrm{i}b)s}f(s)\mathrm{d}s\right| &= \left|\frac{e^{(a+\mathrm{i}b)s}}{a+\mathrm{i}b}f(s)\right|_r^t - \int_r^t \frac{e^{(a+\mathrm{i}b)s}}{a+\mathrm{i}b}\dot{f}(s)\mathrm{d}s\right| \\ &\leq \frac{\exp(ar)+\exp(at)}{\sqrt{a^2+b^2}}f(r) + \int_r^t \frac{e^{as}}{|a+\mathrm{i}b|}|\dot{f}(s)|\mathrm{d}s \\ &\leq \frac{\exp(ar)+\exp(at)}{\sqrt{a^2+b^2}}f(r) + \frac{\exp(ar)+\exp(at)}{\sqrt{a^2+b^2}}\int_r^t |\dot{f}(s)|\mathrm{d}s \\ &\leq 2\frac{\exp(ar)+\exp(at)}{\sqrt{a^2+b^2}}f(r).\end{aligned}$$

□

Appendix A.3. Non asymptotic lower bound

Proof of Proposition 3.4. As we focus on one wavenumber k at the moment, its appearance in the subscript will be suppressed. Using the notation (A.2), $\Sigma_{t|U} = \int_0^t |\int_r^t \sigma_{r,s}^t \mathrm{d}s|^2\mathrm{d}r$ with

$$\left|\int_r^t \sigma_{r,s}^t ds\right| = \exp(-\gamma_v(t-r))\left|\int_0^{t-r}\exp(-\gamma_R s + i\tilde{\omega}_R[0,s])ds\right|$$

where $\tilde{\omega}_R(s) = \omega_R(t-s)$ is the reverse process. $\tilde{\omega}_R$ is also an OU process, which follows:

$$d\tilde{\omega}_R(s) = -\gamma_U\tilde{\omega}_R(s)ds + \gamma_U d_k\mathrm{d}s + c_k\sigma_U\mathrm{d}\tilde{W}_s.$$

where c_k, d_k are given by (A.3) and $\tilde{W}_s$ is a Wiener process. This is known as the reversibility of O.U. processes. One easy way to see it is checking the time covariance function of $\tilde{\omega}_R(s)$,

which is the same as $\omega_R(s)$, while both of them are Gaussian processes. Thus by Duhamel's formula:

$$\tilde{\omega}_R(s) = e^{-\gamma_U s}\tilde{\omega}_R(0) + (1 - e^{-\gamma_U s})d_k + c_k\sigma_U \int_0^s e^{-\gamma_U(s-u)}\mathrm{d}\tilde{W}_u.$$

We will write the conditional expectation with $\tilde{\omega}_R(0)$, or equivalently $\omega_R(t)$, given as $\mathbb{E}_0$, then

$$\mathbb{E}_0\tilde{\omega}_R(s) = e^{-\gamma_U s}\tilde{\omega}_R(0) + (1 - e^{-\gamma_U s})d_k,$$

while

$$\begin{aligned}\mathrm{cov}_0(\tilde{\omega}_R(s), \tilde{\omega}_R(t)) &= \mathbb{E}_0\tilde{\omega}_R(s)\tilde{\omega}_R(t) - \mathbb{E}_0\tilde{\omega}_R(s)\mathbb{E}_0\tilde{\omega}_R(t)\\ &= c_k^2\sigma_U^2 \int_0^{s\wedge t} e^{-\gamma_U(s+t-2u)}\mathrm{d}u\\ &= E_c(e^{-\gamma_U|s-t|} - e^{-\gamma_U(s+t)}), \quad E_c := c_k^2 E_U.\end{aligned}$$

So conditioned on $\tilde{\omega}_R(0)$, $\tilde{\omega}_R[s,t]$ is normal distributed with conditional mean

$$\begin{aligned}m_{s,t} &= d_k(t-s) + \frac{e^{-\gamma_U s} - e^{-\gamma_U t}}{\gamma_U}(\tilde{\omega}_R(0) - d_k)\\ &= \tilde{\omega}_R(0)(t-s) + \left[\frac{e^{-\gamma_U s} - e^{-\gamma_U t}}{\gamma_U} - (t-s)\right](\tilde{\omega}_R(0) - d_k)\\ &=: \tilde{\omega}_R(0)(t-s) + n_{s,t};\end{aligned}$$

and conditional variance

$$V_{s,t} = E_c\left[\frac{2|t-s|}{\gamma_U} - \frac{2 - 2e^{-\gamma_U|t-s|}}{\gamma_U^2} - \frac{(e^{-\gamma_U s} - e^{-\gamma_U t})^2}{\gamma_U^2}\right].$$

Thus by Fourier transform of Gaussian variables: $\mathbb{E}\exp(\mathrm{i}X) = \exp(\mathrm{i}\mathbb{E}X - \frac{1}{2}\mathrm{var}(X))$, we have

$$\mathbb{E}_0\left|\int_{t-r}^t \sigma_{t-r,s}^t \mathrm{d}s\right|^2 = \exp(-2\gamma_v r)\int_0^r \exp(-\gamma_R s_1)\int_0^r \mathbb{E}_0\exp(-\gamma_R s_2 + \mathrm{i}\tilde{\omega}_R[s_1,s_2])\mathrm{d}s_2\mathrm{d}s_1. \tag{A.4}$$

By Lemma Appendix A.2 established below, the inner layer of integral can be written as

$$\begin{aligned}&\int_0^r \mathbb{E}_0\exp(-\gamma_R s_2 + \mathrm{i}\tilde{\omega}_R[s_1,s_2])\mathrm{d}s_2\\ &= \int_0^r \exp(-\gamma_R s + \mathrm{i}m_{s_1,s} - \frac{1}{2}V_{s_1,s})\mathrm{d}s\\ &= \int_0^r \exp(-\gamma_R s + \mathrm{i}\tilde{\omega}_R s)\exp(-\frac{1}{2}V_{s_1,s} + \mathrm{i}m_{s_1,s})\mathrm{d}s\\ &= \frac{\exp(-\frac{1}{2}V_{s_1,0} + \mathrm{i}m_{s_1,0})}{\gamma_R - \mathrm{i}\tilde{\omega}_R(0)} - \frac{\exp(-\gamma_R r + \mathrm{i}m_{s_1,r} - \frac{1}{2}V_{s_1,s_2})}{\gamma_R - \mathrm{i}\tilde{\omega}_R(0)} + Q_{s_1,r},\end{aligned} \tag{A.5}$$

with the residual term bounded by

$$|Q_{s_1,r}| \le \frac{1 \vee e^{-\gamma_R r}}{|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|}(\mathrm{Vr}_{[0,r]} \exp(-\tfrac{1}{2}V_{s_1,x}) + \mathrm{Vr}_{[0,r]} n_{s_1,x}).$$

Here $\mathrm{Vr}_{[0,r]} f(x)$ denotes the total variation of function f inside interval $[0,r]$. For function $n_{s_1,x}$, it is easy to check that it is increasing on $[0,s_1]$ and decreasing on $[s_1,r]$, if $\tilde{\omega}_R(0) > 0$ or vice versa. Moreover, for $s,t \le r < \gamma_U^{-1}$, since

$$\frac{e^{-\gamma_U s} - e^{-\gamma_U t}}{\gamma_U} = \frac{e^{-\gamma_U s}(1 - e^{-\gamma_U(t-s)})}{\gamma_U} \ge (1 - \gamma_U s)(t - s - \frac{1}{2}\gamma_U (t-s)^2).$$

Therefore

$$|n_{s,t}| \le [\frac{1}{2}\gamma_U (t-s)^2 + s(t-s)\gamma_U]|\tilde{\omega}_R(0) - d_k| = \frac{1}{2}\gamma_U (t^2 - s^2)|\tilde{\omega}_R(0) - d_k|.$$

Thus

$$\mathrm{Vr}_{[0,r]} n_{s_1,x} \le \gamma_U r^2 |\tilde{\omega}_R(0) - d_k|. \tag{A.6}$$

For function $\exp(-\frac{1}{2}V_{s_1,x})$, notice that $V_{s_1,x}$ is decreasing in $[0,s_1]$ and increasing on $[s_1,r]$. Moreover, for $s,t \le r \le \gamma_U^{-1}$ an upper bound of $V_{s,t}$ can be obtained through

$$1 - e^{-\gamma_U|t-s|} \ge \gamma_U |t-s| - \frac{1}{2}\gamma_U^2 |t-s|^2$$

$$(e^{-\gamma_U s} - e^{-\gamma_U t})^2 = e^{-2\gamma_U s}(1 - e^{-\gamma_U(t-s)})^2 \ge (1 - 2\gamma_U s)\gamma_U^2 (t-s)^2$$

Thus

$$V_{s,t} \le E_c \left[\frac{2|t-s|}{\gamma_U} - \frac{2|t-s|}{\gamma_U} + |t-s|^2 - (1 - 2\gamma_U s)(t-s)^2 \right] = 2E_c \gamma_U s (t-s)^2, \tag{A.7}$$

and

$$\mathrm{Vr}_{[0,r]} \exp(-\tfrac{1}{2}V_{s_1,x}) \le 2 - 2\exp(-E_c \gamma_U r^3).$$

Based on our definition of C, when $r \le 1$, $|Q_{s,r}| \le |\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|^{-1} C(1 \vee e^{-\gamma_R r})$.

Note that the claimed lower bound in this proposition is actually negative if $\gamma_R = 0$, which then will trivially hold. So it suffices for us to work in the general case that $\gamma_R \neq 0$. First suppose that $\gamma_R > 0$, we will treat the first item in (A.5) as the main item, so we write

$$\text{(A.5)} = \frac{\exp(-\frac{1}{2}V_{s_1,0} + \mathrm{i}m_{s_1,0})}{-\gamma_R + \mathrm{i}\tilde{\omega}_R(0)} + R_{s_1,r}, \quad \text{with} \quad |R_{s_1,r}| \le \frac{\exp(-\gamma_R r) + C}{|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|}.$$

Thus the second layer of integral in (A.4) is bounded below by:

$$\begin{aligned} &\left| \int_0^r \exp(-\gamma_R s_1) \int_0^r \mathbb{E}_0 \exp(-\gamma_R s_2 + \mathrm{i}\tilde{\omega}_R[s_1,s_2]) \mathrm{d}s_2 \mathrm{d}s_1 \right| \\ &\ge \frac{|\int_0^r \exp(-\gamma_R s_1 - \frac{1}{2}V_{s_1,0} + \mathrm{i}m_{s_1,0}) ds_1|}{|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|} - \frac{(1 - \exp(-\gamma_R r))(\exp(-\gamma_R r) + C)}{\gamma_R |\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|}. \end{aligned} \tag{A.8}$$

We apply Lemma Appendix A.2 to the integral in the first item, following the steps in (A.5), since (A.6) and (A.7) are bounds for functions $n_{x,0}$ and $V_{x,0}$ as well, we can find:

$$\left| \int_0^r \exp(-\gamma_R s_1 + \mathrm{i}\tilde{\omega}_R(0)s_1 - \frac{1}{2}V_{s_1,0} + \mathrm{i}n_{s_1,0})ds_1 \right| \geq \frac{1}{|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|} - \frac{\exp(-\gamma_R r) + C}{|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|}.$$

Thus in combine

$$(\mathrm{A}.8) \geq \frac{1}{|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|^2} - \frac{2C + 2\exp(-\gamma_R r)}{\gamma_R|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|},$$

$$\mathbb{E}_0\Sigma_{t|U} \geq \mathbb{E}_0 \int_0^{t_0} (\mathrm{A}.4)\mathrm{d}r \geq \frac{1 - \exp(-2\gamma_v t_0)}{2\gamma_v|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|^2} - \frac{\frac{\gamma_v}{\gamma_T+\gamma_v} + C(1 - \exp(-2\gamma_v t_0))}{\gamma_v|\gamma_R||\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|}.$$

This is our claim for $\gamma_R > 0$. The $\gamma_R < 0$ case is completely symmetric. We take the second term in (A.5) as the main term, and rewrite (A.5) as

$$\int_0^r \mathbb{E}_0 \exp(-\gamma_R s_2 + \mathrm{i}\tilde{\omega}_R[s_1, s_2])\mathrm{d}s_2 = \frac{\exp(-\gamma_R r - \frac{1}{2}V_{s_1,s_2} + \mathrm{i}m_{s_1,s_2})}{-\gamma_R + \mathrm{i}\tilde{\omega}_R(0)} + R_{s_1,r},$$

while the residual:

$$|R_{s_1,r}| \leq \left| \frac{\exp(-\frac{1}{2}V_{s_1,0} + \mathrm{i}m_{s_1,0})}{\gamma_R - \mathrm{i}\tilde{\omega}_R(0)} \right| + |Q_{s_1,r}| \leq \frac{1 + C\exp(-\gamma_R r)}{|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|}.$$

Hence the second layer of integral (A.4) is bounded by

$$\begin{aligned} &\left| \int_0^r \mathrm{d}s_1 \exp(-\gamma_R s_1) \int_0^r \mathbb{E}_0 \exp(-\gamma_R s_2 + \mathrm{i}\tilde{\omega}_R[s_1, s_2])\mathrm{d}s_2 \right| \\ &\geq \left| \int_0^r \frac{\exp(-\gamma_R(s_1 + r) - \frac{1}{2}V_{s_1,r} + \mathrm{i}m_{s_1,r})\mathrm{d}s_1}{-\gamma_R + \mathrm{i}\tilde{\omega}_R(0)} \right| - \frac{\exp(-\gamma_R r) + C\exp(-2\gamma_R r)}{|\gamma_R||\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|}. \end{aligned}$$

Then we apply Lemma Appendix A.2 to the integral in the first item, and find

$$\left| \int_0^r \exp(-\gamma_R(s_1 + r) - \frac{1}{2}V_{s_1,r} + \mathrm{i}m_{s_1,r})\mathrm{d}s_1 \right| \geq \frac{\exp(-2\gamma_R r) - C\exp(-2\gamma_R r) - \exp(-\gamma_R r)}{|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|}.$$

So in combine

$$(\mathrm{A}.4) \geq \frac{\exp(-2\gamma_T r)}{|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|^2} - \frac{2\exp(-(\gamma_T + \gamma_v)r) + 2C\exp(-2\gamma_T r)}{|\gamma_R||\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|}$$

. Therefore our claim for $\gamma_R < 0$ is obtained through

$$\mathbb{E}_0\Sigma_{t|U} \geq \int_0^{t_0} (\mathrm{A}.4)\mathrm{d}r \geq \frac{1 - \exp(-2\gamma_T t_0)}{2\gamma_T|\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|^2} - \frac{\frac{\gamma_T}{\gamma_T+\gamma_v} + C(1 - \exp(-2\gamma_T t_0))}{\gamma_T|\gamma_R||\gamma_R - \mathrm{i}\tilde{\omega}_R(0)|}.$$

□

Lemma Appendix A.2. *Let $f(s)$ be a positive C^1 function, $\theta(s)$ be a C^1 function with a, b being two real constants, then*

$$\left| \int_0^t e^{(a+ib)s} f(s) e^{i\theta(s)} \mathrm{d}s - \frac{e^{(a+\mathrm{i}b)s}}{a+\mathrm{i}b} e^{\mathrm{i}\theta(s)} f(s) \Big|_0^t \right| \leq \frac{1 \vee e^{at}}{|a+bi|} [\, Vr_{[0,t]}(f) + Vr_{[0,t]}(\theta)].$$

A direct consequence is:

$$\left| \int_0^t e^{(a+ib)s} f(s) e^{i\theta(s)} \mathrm{d}s \right| \geq \frac{|f(0) - e^{at} f(t)|}{|a+\mathrm{i}b|} - \frac{1 \vee e^{at}}{|a+bi|} [\, Vr_{[0,t]}(f) + Vr_{[0,t]}(\theta)].$$

Here $Vr_{[0,t]} f$ denotes the total variation of f on the interval $[0, t]$.

Proof. The first claim is obtained through integration by part formula, since

$$\int_0^t e^{(a+ib)s} e^{\mathrm{i}\theta(s)} f(s) ds = \frac{e^{(a+\mathrm{i}b)s}}{a+\mathrm{i}b} e^{\mathrm{i}\theta(s)} f(s) \Big|_0^t - \int_0^t \frac{e^{(a+\mathrm{i}b)s}}{a+\mathrm{i}b} \dot{f}(s) e^{\mathrm{i}\theta(s)} ds - \mathrm{i} \int_0^t \frac{e^{(a+\mathrm{i}b)s}}{a+\mathrm{i}b} f(s) e^{\mathrm{i}\theta(s)} \dot{\theta}(s) ds,$$

where the two integrals on the right hand side can be bounded by:

$$\left| \int_0^t \frac{e^{(a+\mathrm{i}b)s}}{a+\mathrm{i}b} \dot{f}(s) e^{\mathrm{i}\theta(s)} \mathrm{d}s \right| \leq \int_0^t \frac{|e^{as}||\dot{f}(s)|}{|a+\mathrm{i}b|} \mathrm{d}s \leq \frac{1 \vee e^{at}}{|a+\mathrm{i}b|} \mathrm{Vr}_{[0,t]} f,$$

$$\left| \int_0^t \frac{e^{(a+\mathrm{i}b)s}}{a+\mathrm{i}b} f(s) e^{\mathrm{i}\theta(s)} \dot{\theta}(s) \mathrm{d}s \right| \leq \left| \int_0^t \frac{|e^{as}|}{|a+\mathrm{i}b|} |f(s)||\dot{\theta}(s)| \mathrm{d}s \leq \frac{1 \vee e^{at}}{|a+\mathrm{i}b|} \mathrm{Vr}_{[0,t]} \theta. \right.$$

The consequence is a direct one by the relation $|x - y| \leq ||x| - |y||$:

$$\left| \frac{e^{(a+\mathrm{i}b)s}}{a+\mathrm{i}b} e^{\mathrm{i}\theta(s)} f(s) \Big|_0^t \right| \leq \left| \frac{|e^{at}| f(t)}{|a+\mathrm{i}b|} - \frac{f(0)}{|a+\mathrm{i}b|} \right|.$$

□

Appendix A.4. Proofs for the asymptotic results

Lemma Appendix A.3. *The following holds by Cauchy Schwartz and Young's inequality*

$$\begin{aligned} \left(\int_0^t (Y_s^2 - X_s^2) \mathrm{d}s \right)^2 &\leq \left(\int_0^t (X_s + Y_s)^2 \mathrm{d}s \right) \left(\int_0^t (X_s - Y_s)^2 \mathrm{d}s \right) \\ &\leq 2 \left(\int_0^t X_s^2 \mathrm{d}s \right) \left(\int_0^t (X_s - Y_s)^2 \mathrm{d}s \right) + 8 \left(\int_0^t (X_s - Y_s)^2 \mathrm{d}s \right)^2. \end{aligned}$$

Therefore the following holds by checking the square of both sides

$$\left| \int_0^t (Y_s^2 - X_s^2) \mathrm{d}s \right| \leq 2 \sqrt{\int_0^t X_s^2 \mathrm{d}s} \sqrt{\int_0^t (X_s - Y_s)^2 \mathrm{d}s} + 4 \int_0^t (X_s - Y_s)^2 \mathrm{d}s.$$

And so does its expectation form:

$$\mathbb{E} \left| \int_0^t (Y_s^2 - X_s^2) \mathrm{d}s \right| \leq 2 \sqrt{\mathbb{E} \int_0^t X_s^2 \mathrm{d}s} \sqrt{\mathbb{E} \int_0^t (X_s - Y_s)^2 \mathrm{d}s} + 4 \mathbb{E} \int_0^t (X_s - Y_s)^2 \mathrm{d}s.$$

Lemma Appendix A.4. *Let $dU_t = f_U(U_t)\mathrm{d}t + \sigma_U(U_t)\mathrm{d}W_t$ be a diffusion process such that Assumption 6.1 holds, assume U_0 is initialized with the invariant distribution of U_t, then there exists a constant M such that:*

$$\mathbb{E}\int_0^t \left[\int_s^t U_r \mathrm{d}r - (t-s)U_t\right]^2 \mathrm{d}s \le Mt^5.$$

Proof. By definition, U_t can be written as

$$U_t = U_s + \int_s^t f(U_r)\mathrm{d}r + \int_s^t \sigma_U(U_r)\mathrm{d}W_r.$$

By Fubini's theorem of stochastic integrals,

$$\int_s^t U_r \mathrm{d}r = (t-s)U_s + \int_s^t (t-r)f(U_r)\mathrm{d}r + \int_s^t (t-r)\sigma_U(U_r)\mathrm{d}W_r.$$

Therefore for $s \le t$,

$$\begin{aligned}
\mathbb{E}\left|\int_s^t U_r \mathrm{d}r - (t-s)U_t\right|^2 &= \mathbb{E}\left|\int_s^t (t-r)f(U_r)\mathrm{d}r + \int_s^t (t-r)\sigma_U(U_r)\mathrm{d}W_r\right|^2 \\
&\le 2\mathbb{E}\left[\int_s^t (t-r)f(U_r)\mathrm{d}r\right]^2 + 2\mathbb{E}\left[\int_s^t (t-r)\sigma_U(U_r)\mathrm{d}W_r\right]^2 \\
&\le 2t^3\mathbb{E}\int_0^t f^2(U_s)\mathrm{d}s + 2t^3\mathbb{E}\int_0^t \sigma_U^2(U_s)\mathrm{d}s,
\end{aligned}$$

which implies:

$$\mathbb{E}\int_0^t \left[\int_s^t U_r \mathrm{d}r - (t-s)U_t\right]^2 \mathrm{d}s \le 3t^4\mathbb{E}\int_0^t [f_U^2(U_s) + \sigma_U^2(U_s)]\mathrm{d}s.$$

Recall that $f^2(U_s), \sigma^2(U_s) \le \mathcal{E}(U_s)$, and since $\mathcal{E}$ is a Lyapunov function, the expectation under equilibrium measure, $\mathbb{E}\mathcal{E}(U_s)$, is bounded using its property (A.9) shown in the next section. Hence the quantity above is bounded by Mt^5 for a proper M. □

Appendix A.5. Lyapunov functions

Lyapunov functions are commonly used in probability theory to stabilize a stochastic process X_t. One way to define it is through the infinitesimal generator

$$\mathcal{L}f(x) := \lim_{t\to 0}\frac{1}{t}[f(\mathcal{E}(X_t)|X_0 = x) - f(x)],$$

that is requiring for some strictly positive constants γ, k_v the following holds:

$$\mathcal{L}\mathcal{E}(x) \le -\gamma\mathcal{E}(x) + k_v,$$

while $\mathcal{E}$ should be strictly positive with its sub level sets being compact. The central role of Lyapunov function is that it can bound the behavior of stochastic process X_t, since using Dynkin's formula, we have

$$\mathbb{E}\mathcal{E}(X_t) = \mathcal{E}(X_0) + \mathbb{E}\int_0^t \mathcal{L}\mathcal{E}(X_s)\mathrm{d}s \leq \mathcal{E}(x) + \mathbb{E}\int_0^t \mathcal{L}\mathcal{E}(X_s)\mathrm{d}s.$$

Then the time derivative is bounded by $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\mathcal{E}(X_t) \leq -\gamma\mathbb{E}\mathcal{E}(X_t) + k_v$, which by Gronwall inequality gives us

$$\mathbb{E}\mathcal{E}(X_t) \leq e^{-\gamma t}\mathcal{E}(X_0) + \frac{k_v}{\gamma}. \tag{A.9}$$

As for a dimensional one diffusion process $\mathrm{d}U_t = b(U_t)\mathrm{d}t + \sigma(U_t)\mathrm{d}W_t$, it's a standard result through Fokker Plank equation that the infinitesimal generator is given by:

$$\mathcal{L}f(u) = b(u)\frac{\partial f}{\partial u} + \frac{1}{2}\frac{\partial^2 f}{\partial u^2}.$$

For details proof of the claims above, we refer to classic works in this direction [22, 23]. To see that $\mathcal{E}(u) = u^2$ is a Lyapunov function for $\mathrm{d}U_t = -\gamma_U U_t\mathrm{d}t + \sigma_U\mathrm{d}W_t$, simply apply the previous formula:

$$\mathcal{L}u^2 = -2\gamma_U u^2 + \sigma_U^2,$$

so $\gamma_U = 2\gamma$ and $k_v = \sigma_U^2$ applies. For the canonical cubic scalar model (6.2), direct computation shows:

$$\mathcal{L}u^6 = -6u^8 + 6u^5[F + au + bu^2] + 15u^4[(A - Bu)^2 + \sigma^2].$$

Since the right hand side is a polynomial of u, with the highest order term being $-6u^8$. Then by Hölder's inequality, for any fixed $\gamma > 0$, one can find a k_v such that $\mathcal{L}u^6 \leq -\gamma u^6 + k_v$.

References

[1] D Frierson. Robust increases in midlatitude static stability in global warming simulation. *Geophysical Res. Lett.*, 33(24):L24816, 2006.

[2] D Frierson. Midlatitude static stability in simple and comprehensive general circulation models. *J. Atmos. Sci.*, 65:1049–1062, 2008.

[3] J Neelin, B Lintner, B Tian, Q Li, L Zhang, Chahine M Patra, P, and S Stechmann. Long tails in deep columns of natural and anthropogenic tropospheric tracers. *Geophysical Res. Lett.*, 37:L05804, 2010.

[4] A J Majda and P Kramer. Simplied models for turbulent diffusion: theory, numerical modeling, and physical phenomena. *Phys. Rep.*, 314:237–574, 1999.

[5] A J Majda and B Gershgorin. Elementary models for turbulent diffusion with complex physical features: eddy diffusivity, spectrum, and intermittency. *Phil. Trans. Roy. Soc.*, 371(1982), 2013.

[6] B. Castaing, G Guanratne, F Heslot, L Kadanoff, A Libchaber, S Thomae, X Wu, S Zaleski, and G Zanetti. Scaling of hard thermal turbulence in rayleigh-bénard convection. *J. Fluid Mech.*, 204(1):1–30, 1989.

[7] J P Gollub, J Clarke, M Gharib, B Lane, and O N Mesquita. Fluctuations and transport in a stirred fluid with a mean gradient. *Phys. Rev. Lett.*, 67:3507–3510, 1991.

[8] Jayesh and A Warhaft. Probability distribution, conditional dissipation, and transport of passive temperature fluctuations in grid-generated turbulence. *Phys. Fluids A.*, 4:2292, 1991.

[9] M Avellaneda and A J Majda. Simple exmaples with features of renormalization for turbulent transport. *Phil. Trans. Roy. Soc.*, 346:205–233, 1994.

[10] A Bourlioux and A J Majda. Elementary models with probability distribtuion function intermittency for passive scalars with a mean gradient. *Phys. Fluids*, 14(2), 2002.

[11] R. Camassa, N. Martinsen-Burrell, and R. M. Mclaughlin. Dynamics of probability density functions for decaying passive scalars in periodic velocity fields. *Phys. Fluids*, 19(117104), 2007.

[12] A J Majda and J Harlim. *Filtering complex turbulent systems.* Cambridge University Press, Cambridge, UK, 2012.

[13] A J Majda, C Franzke, and D T Crommelin. Normal forms for reduced stochastic climate models. *Proc. Natl. Acad. Sci.*, 16(10):3649–3653, 2009.

[14] A J Majda and M Branicki. Lessons in uncertainty quantification for turbulent dynamical system. *Discrete and Continuous Dynamical Systems*, 32(9):3133–3221, 2013.

[15] M. Mohamad and T. Sapsis. Probabilistic description of extreme events in interminttently unstable systems excited by correlated stochastic process. submitted to SIAM J. of U. Q.

[16] W. Cousins and T. Sapsis. Extreme events in 1d nonlinear dispersive waves. *Physica D*, 280-281:48–58, 2014.

[17] A J Majda and Y Lee. Conceptual dynamical model for turbulence. *Proc. Natl. Acad. Sci.*, 111(18):6548–6553, 2014.

[18] M Avellaneda and A J Majda. Mathematical models with exact renormalization for turbulent transport. *Commun. Math. Phys.*, 131:381–429, 1990.

[19] M Avellaneda and A J Majda. Mathematical models with exact renormalization for turbulent transport, ii: fractal interfaces, non-gaussian statistics and the sweeping effect. *Commun. Math. Phys.*, 146:139–204, 1992.

[20] B Gershgorin and A J Majda. Filtering a statistically exactly solvable test model for turbulent tracers from partial observations. *J. Comput. Phys*, 230:1602–1638, 2011.

[21] A J Majda and B Gershgorin. Quantifying uncertainty in climate change science through emperiical information theory. *Proc. Natl Acad. Sci.*, 107:14958–14963, 2010.

[22] J C Mattingly and A M Stuart. Geometric ergodicity of some hypo-elliptic diffustion for particle motions. *Markov Process. Related Fields*, 8(2):199–214, 2002.

[23] J C Mattingly, A M Stuart, and D J Higham. Ergodicity for sdes and approximations: Locally lipschitz vector fields and degenerate noise. *Stochastic processes and their applications*, 101(2):185–232, 2002.

[24] R S Liptser and A N Shiryaev. *Statistics of random processes. I, II,*, volume 5 of *Applications of Mathematics*. Springer-Verlag, 2001.

[25] N Chen, A J Majda, and X T Tong. Information barriers for noisy lagrangian tracers in filtering random incompressible flows. *Nonlinearity*, 27:2133–2163, 2014.

[26] N Chen, A J Majda, and X T Tong. Noisy lagrangian tracers for filtering random rotating compressible flows. submitted to J. Non. Sci., 2014.

[27] B Picinbono. Second-order complex random vectors and normal distributions. *IEEE Trans on Signal Processing*, 44, 1996.

[28] P E Protter. *Stochastic Integration and Differential Equations.* Stochastic modeling and applied probability. Springer, 2005.

[29] P Billingsley. *Convergence of Probability Measure.* Wiley Series in Probability and Statistics. Wiley-Interscience, 1999.

[30] H Föllmer. Tail structure of markov chains on infinite product space. *Z. Wahrsch. Verw. Gebiete*, 50:273–285, 1979.

[31] X T Tong and R. Van Handel. Conditional ergodicity in inifite dimension. *Annals of Probability*, 42(6):2243–2313, 2014.

[32] M Hairer and J C Mattingly. Ergodicity of the 2d navier-stokes equations with degnerate stochastic

forcing. *Annals of Mathematics*, 164:993–1032, 2006.

[33] M Hairer and J C Mattingly. A theory of hypoellipticity and unique ergodicity for semilinear stochastic pdes. *Electron. Commun. Probab.*, 16(23):658–738, 2011.

[34] M Hairer, J C Mattingly, and M Scheutzow. Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields*, 149(1-2):223–259, 2011.