Quasi-Markovian property of strong wave turbulence

Wonjung Lee [*]
Department of Mathematics, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong

(Received 3 February 2021; accepted 7 April 2021; published 3 May 2021)

This paper is concerned with the reduced-order modeling of the strongly nonlinear wave turbulence system. The motivation for such an attempt comes from the utility of the probabilistic coarse-grained model in facilitating the theoretical and numerical analysis of the true dynamical system model. One typical practice of simplifying the complex physical model is, in the spirit of Brownian motion, to replace the nonlinear interactions by white noise forcing and linear dissipation. For the case of slowly varying longwave, the resulting Markov process is an accurate approximate model. However, this conventional scheme is highly inappropriate for the description of shortwaves because the rapidly varying turbulent signal acquires a significantly non-Markovian character resulting from the poor timescale separation between the relevant mode and the environmental wave field. To resolve the issue, we discuss a simplification technique for which the central concept is the quasi-Markovian property; a non-Markov stochastic process is referred to as quasi-Markovian if it can be represented as a component of Markovian system made by adding a finite number of auxiliary variables. Our contribution in this work is to single out the nontrivial and near resonances from the nonlinear interactions in search of the auxiliary variable. We perform a comparison analysis of the autocorrelation matrices of the true and approximate models, and numerically demonstrate the effectiveness of our Markovian formulation of the inherently non-Markov turbulent signal.

DOI: 10.1103/PhysRevE.103.052101

I. INTRODUCTION

A. Overview

A Hamiltonian dynamics typically arises when one takes into consideration all the state variables involved in the evolution of the entire physical system. The relevant nature phenomenon is mathematically described by the Hamilton's canonical equations. In principle, by studying this deterministic and Markovian law of motion, one can explain and predict the true dynamical system behavior. In practice, however, the full resolution of the detailed trajectories of the canonical variables in the phase space presents a great computational challenge especially when the underlying system accounting for turbulence possesses tremendously many degrees of freedom. In many cases, concern is confined to a tiny fraction of the whole physical system and such a straightforward numerical integration implies too much waste of resources. To circumvent this computational difficulty and to seek an economy in the description, the theory of stochastic processes has been initiated [1,2]. The basic idea is to design an empirical law of motion for the relevant variables while masking the irrelevant variables from observation. This coarse-graining procedure naturally induces the uncertainty to compensate the lost information regarding the initial condition of the whole system and endows the outcome dynamics with the probabilistic nature.

Brownian motion is one representative example in which a physical process is successfully modeled as stochastic process [2,3]. There it is important to note that the reduced-order modeling task yields a Markov model primarily because the surrounding fine molecules move significantly faster than the Brownian particle. In general, however, this is not the case and the future evolution of the coarse-grained model is not only determined by the information at that instant but by the history due to the poor timescale separation between the resolved target variable and the remaining unresolved ones. One intricacy here is that a non-Markov model is generally less tractable than Markov models for the purpose of the theoretical and numerical analysis. In this connection, the capability of simultaneously achieving a drastic simplification of the complex physical model and maintaining the advantageous Markovian property of the original Hamiltonian system is truly valuable, and the implementation will give us enormous benefits in the study of turbulent signals generated by the true dynamical system model. The present paper is devoted to the development of such a desired marginalization scheme for the system of nonlinear dispersive waves in the strong coupling limit.

B. Problem statement

Consider the Hamiltonian

$$\mathcal{H} = \sum_k \chi\omega_k |a_k|^2 + \frac{1}{2}\sum_{k_{1234}} W_{34}^{12}\delta_{34}^{12} a_1 a_2 a_3^* a_4^*, \tag{1}$$

where $k_{1234} = k_1, k_2, k_3, k_4$ and a_j is shorthand notation of a_{k_j} for $j = 1, 2, 3, 4$. The parameter $\chi\,(=1 \text{ or } 0)$ determines the existence and nonexistence of the quadratic terms. Here $\delta_{34}^{12} = \delta_{k_1+k_2-k_3-k}$ is Kronecker delta function and the upper

[*]lee.wonjung@cityu.edu.hk

$*$ signifies complex conjugation. The Hamilton's equation of motion governing the canonical variable a_k is given by

$$\begin{aligned} \dot{a}_k &= -i\frac{\delta \mathcal{H}}{\delta a_k^*} \\ &= -i\chi\omega_k a_k - i\sum_{k_{123}} W_{3k}^{12}\delta_{3k}^{12} a_1 a_2 a_3^*, \end{aligned} \tag{2}$$

where the upper dot denotes the time derivative.

In the case of $\chi = 1$, Eq. (2) is one prototype of the wave turbulence (WT) system, for which ω_k is the linear dispersion relation and W_{34}^{12} is the four-wave interaction tensor [4,5]. The WT theory provides a generic approach to the statistical description of the family of nonlinear dispersive waves. Note the theory is applicable to the dynamical system (2) only if (i) $\chi = 1$ so that the coefficient ω_k exists and (ii) the wave-wave interactions are in the weakly nonlinear limit. Importantly, such conditions are not compulsory in our treatment. The WT system under our consideration does not necessarily possess the linear dispersive characteristics, that is, the case of $\chi = 0$ can be addressed by our development. Besides, our discussion throughout the paper is targeted to the situations where the nonlinear coupling is strong and dominant in the course of time.

The goal of this work is to come up with a low-dimensional Markovian model characterized by a close resemblance with the true dynamical system model (1) and (2) at the statistical level. The main analytic tool for the rigorous derivation is the Mori-Zwanzig (MZ) projection theory sketched in the next subsection.

C. Background on the Mori-Zwanzig theory

Let $\mathbf{A}(t) = \mathbf{A}(\{a_k(t), a_k^*(t)\})$ be an arbitrary vector-valued variable determined by some of the canonical variables. Recall that its dynamic evolution obeys the Liouville equation $\dot{\mathbf{A}} = \mathcal{L}\mathbf{A}$ where $\mathcal{L} = \sum_k -i\frac{\delta\mathcal{H}}{\delta a_k^*}\partial_{a_k} + i\frac{\delta\mathcal{H}}{\delta a_k}\partial_{a_k^*}$ is the Liouville operator. The MZ formalism [3,6–8] performs the *exact* rearrangement of the Liouville equation, which is in the form

$$\dot{\mathbf{A}}(t) = -i\mathbf{\Omega}\mathbf{A}(t) - \int_0^t ds\, \mathbf{\Gamma}(t-s)\mathbf{A}(s) + \mathbf{R}(t). \tag{3}$$

The first term of the right-hand side of (3) is obtained from the projection of $\dot{\mathbf{A}}(t)$ onto the linear space spanned by $\mathbf{A}$. More precisely, defining the inner product $\langle \mathbf{X}, \mathbf{Y}\rangle \equiv \langle \mathbf{X}\mathbf{Y}^\mathsf{T}\rangle$ for which the angle bracket denotes the statistical average against the stationary distribution and the upper T denotes conjugate transpose, and defining the projection operator by $\mathcal{P}\mathbf{X} \equiv \langle \mathbf{X}, \mathbf{A}\rangle\langle \mathbf{A}, \mathbf{A}\rangle^{-1}\mathbf{A}$, the projection coefficient is given by $-i\mathbf{\Omega} = \langle \dot{\mathbf{A}}, \mathbf{A}\rangle\langle \mathbf{A}, \mathbf{A}\rangle^{-1}$.

The second and third terms form the complement of the projection part. With $\mathbf{I}$ denoting the identity operator, $\mathbf{R}(t) = e^{t(\mathbf{I}-\mathcal{P})\mathcal{L}}(\mathbf{I}-\mathcal{P})\mathcal{L}\mathbf{A}$ is related with the memory kernel via

$$\mathbf{\Gamma}(t) = \frac{\langle \mathbf{R}_t, \mathbf{R}_0\rangle}{\langle \mathbf{A}, \mathbf{A}\rangle}, \tag{4}$$

where the reciprocal signifies matrix inverse. The MZ framework regards $\mathbf{R}_t$ as the random noise, and refers to (3) as the generalized Langevin equation (GLE) equipped with the fluctuation-dissipation (FD) theorem (4).

It follows from Eq. (3) and the identity $\langle \mathbf{R}_t, \mathbf{A}_0\rangle = 0$ that the autocorrelation matrix evolves according to

$$\langle \dot{\mathbf{A}}_t, \mathbf{A}_0\rangle = -i\mathbf{\Omega}\langle \mathbf{A}_t, \mathbf{A}_0\rangle - \int_0^t ds\, \mathbf{\Gamma}(t-s)\langle \mathbf{A}_s, \mathbf{A}_0\rangle. \tag{5}$$

Defining the Fourier-Laplace transform of $\phi(t)$ as $\phi[w] = \mathfrak{F}\{\phi\}(\omega) \equiv \int_0^\infty dt\, e^{-i\omega t}\phi(t)$, Eq. (5) takes the form

$$\frac{\mathfrak{F}\{\langle \mathbf{A}_t, \mathbf{A}_0\rangle\}(\omega)}{\langle \mathbf{A}, \mathbf{A}\rangle} = \frac{1}{i(\omega\mathbf{I} + \mathbf{\Omega}) + \mathbf{\Gamma}[\omega]} \tag{6}$$

for which $\dot{\phi}[\omega] = i\omega\phi[\omega] - \phi(0)$ is used. Equations (4) and (6) reveal the connection between the autocorrelations of $\mathbf{A}$ and $\mathbf{R}$.

D. Organization of the paper

The remainder of this paper is organized as follows. In Secs. II and III we develop two reduced-order models: Trio model (TRM) and resonant duet model (RDM). We perform numerical experiments in Sec. IV, and give the summary of our work in Sec. V. Some calculations and arguments are gathered in the Appendices in order not to disturb the main storyline of the paper.

II. TRIO MODEL (TRM)

Section II A discusses the effective dispersion relationship of the Hamiltonian dynamics (2). In Sec. II B we define two additional variables denoted by A_k, A'_k and associate them with the system variable a_k. We devote Sec. II C to the development of the statistical model governing the triple variables (a_k, A_k, A'_k).

A. Trivial resonances and the mean-field equation

Let us consider the case of $\mathbf{A} = a_k$, and write the corresponding GLE (3) by

$$\dot{a}_k(t) = -i\Omega_k a_k(t) - \int_0^t ds\, \Gamma_a(t-s)a_k(s) + R_a(t), \tag{7}$$

where

$$\begin{aligned} \Omega_k &= i\frac{\langle \dot{a}_k, a_k\rangle}{\langle a_k, a_k\rangle} \\ &= \chi\omega_k + \sum_{k_{123}} W_{3k}^{12}\delta_{3k}^{12}\frac{\langle a_1 a_2 a_3^* a_k^*\rangle}{\langle |a_k|^2\rangle}. \end{aligned} \tag{8}$$

In general, the normal mode frequencies of the relevant variables constituting $\mathbf{A}$ are determined not only by the matrix $-i\mathbf{\Omega}$ but also by some part of the kernel $\mathbf{\Gamma}$ [9]. Specifically, in view of (7), the knowledge we can gain from the MZ formalism is at most that the coefficient Ω_k serves as an effective dispersion relation of a_k provided that this canonical variable is slowly varying and behaves as if it is a Markov process [10]. Surprisingly, however, the direct numerical simulations of a variety of WT systems have shown that the accuracy of Ω_k in describing the effective dispersion relation of a_k remains high across the whole wave-number domain [11–16]. Supported by several numerical evidences, we here proceed by tentatively concluding that the effective dispersion relation of the canonical variables of the Hamiltonian system (1) and (2) is well approximated by the quantity (8).

This wave phenomenon can be explained by virtue of the collective effect of the trivial resonances [11,12]. On one hand, the random phase argument from the WT theory is invoked to approximate the four-point function visible in (8) by the product of two two-point functions and to obtain

$$\Omega_k \doteq \chi\omega_k + \sum_{k'} \left(W_{k'k}^{k'k} + W_{k'k}^{kk'}\right)\langle|a_{k'}|^2\rangle. \tag{9}$$

On the other hand, denoting the nonlinearity of the original dynamics (2) by N_k, we consider the decomposition $N_k = T_k + N_k^{\text{eff}}$ where

$$\begin{aligned} T_k &\equiv -i \sum_{k_{123}, k_{12}=k} W_{3k}^{12}\delta_{3k}^{12} a_1 a_2 a_3^* \\ &= -i\left[\sum_{k'} \left(W_{k'k}^{k'k} + W_{k'k}^{kk'}\right)|a_{k'}|^2\right] a_k, \end{aligned} \tag{10}$$

$$N_k^{\text{eff}} \equiv -i \sum_{k_{123}, k_{12}\neq k} W_{3k}^{12}\delta_{3k}^{12} a_1 a_2 a_3^*. \tag{11}$$

Here T_k is the collection of the terms corresponding to $k_1 = k$ or $k_2 = k$, called the trivial resonances. The result of averaging the coefficient of a_k in (10) is given by the mean-field equation of motion

$$\dot{a}_k \doteq -i\Omega_k a_k + N_k^{\text{eff}} \tag{12}$$

for which N_k^{eff} operates as the effective nonlinearity. Here the re-emergence of Ω_k ensures that T_k takes part in the control of the oscillation, neither forcing nor dissipation, giving rise to a wavelike dynamics.

B. Nontrivial resonances

The comparison between (7) and (12) provides one principled way of building a simplified stochastic model, that is, to replace the effective nonlinearity N_k^{eff} by the forcing and dissipation terms constrained via the FD theorem (4). In [17,18] we apply this skill to develop a number of reduced-order models governing the sole variable a_k. Some of them will be revisited in Sec. III C.

In this work our discussion of a simplification technique begins with the decomposition of the effective nonlinearity (11), i.e., $N_k^{\text{eff}} = A_k + A_k'$, where

$$A_k \equiv -i \sum_{k_{123}, k_{12}\neq k, |\Omega_{3k}^{12}|<\varepsilon} W_{3k}^{12}\delta_{3k}^{12} a_1 a_2 a_3^*, \tag{13a}$$

$$A_k' \equiv -i \sum_{k_{123}, k_{12}\neq k, |\Omega_{3k}^{12}|>\varepsilon} W_{3k}^{12}\delta_{3k}^{12} a_1 a_2 a_3^*, \tag{13b}$$

so that Eq. (12) is in the form

$$\dot{a}_k \doteq -i\Omega_k a_k + A_k + A_k'. \tag{14}$$

Here $\Omega_{3k}^{12} = \Omega_{k_1} + \Omega_{k_2} - \Omega_{k_3} - \Omega_k$ and $0 < \varepsilon \ll 1$ is a small number. The WT theory asserts that, for the system with a cubic nonlinearity, the long-term statistical behavior is often controlled by the resonant quartets and the dominant exchange of energy among the modes occurs via the sets of four waves whose wave numbers $\{(k_1, k_2), (k_3, k)\}$ and dispersion relations $\Omega(\cdot)$ satisfy

$$k_1 + k_2 = k_3 + k, \tag{15a}$$

$$\Omega(k_1) + \Omega(k_2) = \Omega(k_3) + \Omega(k). \tag{15b}$$

Of particular importance is the nontrivial $(k_1, k_2 \neq k)$ resonances. Looking at the resonance condition (15), it is immediate to recognize that our construction of (13) is intended for the variable a_k to mainly interact with A_k while the interaction between a_k and A_k' is insignificant in a long time.

We draw specific attention to the use of the effective dispersion relation Ω_k, not the linear dispersion relation ω_k, in determining the variables in (13). Very interestingly, due to the form of Ω_k in (9), what is happening is that the nontrivial resonances can be created by the pure wave-wave interaction mechanism independently of the linear dispersive characteristics [11,12]. As will be exemplified later in Sec. IV B, this physics picture enables us to construct the variable A_k for many nonlinear dynamical systems which had been traditionally believed not subject to the four-wave resonance framework so that it is impossible to do so.

Notice that, in addition to the exact nontrivial resonances ($k_{12} \neq k$, $\Omega_{3k}^{12} = 0$), we involve the near resonances ($\Omega_{3k}^{12} \doteq 0$) for the construction of A_k. This is because such terms are as important as the exact wave resonances in the energy-momentum exchange [5]. Here arises the question of how to select a suitable value of the parameter ε so that A_k is well defined for a given wave-number k. To answer this, in Appendix C we establish one useful criterion that prevents the two different approximations of the stationary spectrum of A_k from being notably different.

C. Model development

Together with the variables in (13), we now enter a new paradigm for the reduced-order modeling of the WT system (1) and (2). Specifically, we no longer seek a closed equation for the single variable a_k. We instead design a multivariate statistical model governing $\mathbf{A}_3 = (a_k, iA_k, iA_k')^T$ where T denotes transpose.

To this end, we consider the case of $\mathbf{A} = \mathbf{A}_3$ and represent the corresponding GLE (3) by

$$\dot{\mathbf{A}}_3 = -i\mathbf{\Omega}_3\mathbf{A}_3 - \mathbf{\Gamma}_3 * \mathbf{A}_3 + \mathbf{R}_3, \tag{16}$$

where $*$ signifies the convolution in time. First, the theoretical analysis performed in Appendix A yields the approximations $\langle \mathbf{A}_3, \mathbf{A}_3\rangle \doteq \text{diag}(n_k, N_k, N_k')$ and

$$-i\mathbf{\Omega}_3 \doteq -i\begin{pmatrix} \Omega_k & 1 & 1 \\ \mathcal{A}_k & \tilde{\Omega}_k & 0 \\ \mathcal{A}_k' & 0 & \Omega_k' \end{pmatrix}. \tag{17}$$

Here $n_k = \langle|a_k|^2\rangle$, $N_k = \langle|A_k|^2\rangle$, and $N_k' = \langle|A_k'|^2\rangle$, and the ratios between the stationary spectra are denoted by $\mathcal{A}_k = N_k/n_k$ and $\mathcal{A}_k' = N_k'/n_k$. As for $\tilde{\Omega}_k$ and Ω_k', their definitions can be found in Appendix A but the sufficient information at this moment is that $\nu_I \equiv \tilde{\Omega}_k - \Omega_k$ is close to zero, whereas $\nu_I' \equiv \Omega_k' - \Omega_k$ is away from zero. To be more precise, the value of ν_I vanishes if A_k is exclusively made up of the exact resonances and the slight departure from zero is caused by the near resonances. Second, the reasoning provided in Appendix B leads us to suggest the approximations:

$$\mathbf{\Gamma}_3(t) \doteq \begin{pmatrix} 0 & 0 & 0 \\ 0 & \nu_R & 0 \\ 0 & 0 & \nu_R' \end{pmatrix} \delta_+(t), \quad \mathbf{R}_3 \doteq \begin{pmatrix} 0 \\ \Sigma \dot{W}_A \\ \Sigma' \dot{W}_{A'} \end{pmatrix}, \tag{18}$$

where $\delta_+(t)$ is a normalized Dirac delta function satisfying $\int_0^\infty dt\, \delta_+(t) = 1$, and $\dot{W}$. is complex-valued white noise.

Making use of (16), (17), and (18), we build the model equation

$$(\mathbf{TRM}) \quad \begin{cases} \dot{a}_k &= -i\Omega_k a_k + A_k + A_k', \\ \dot{A}_k &= -\mathcal{A} a_k - (i\Omega_k + \nu) A_k + \Sigma \dot{W}_A, \\ \dot{A}_k' &= -\mathcal{A}' a_k - (i\Omega_k + \nu') A_k' + \Sigma' \dot{W}_{A'}, \end{cases} \tag{19}$$

where $\nu = \nu_R + i\nu_I$ and $\nu' = \nu_R' + i\nu_I'$ are complex numbers. Equation (19) is called trio model (TRM) and can be transformed into

$$(\mathbf{TRM}) \quad \begin{pmatrix} \dot{b}_k \\ \dot{B}_k \\ \dot{B}_k' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ -\mathcal{A} & -\nu & 0 \\ -\mathcal{A}' & 0 & -\nu' \end{pmatrix} \begin{pmatrix} b_k \\ B_k \\ B_k' \end{pmatrix} + \begin{pmatrix} 0 \\ \Sigma \dot{W}_B \\ \Sigma' \dot{W}_{B'} \end{pmatrix} \tag{20}$$

in terms of $b_k = e^{i\Omega_k t} a_k$, $B_k = e^{i\Omega_k t} A_k$, and $B_k' = e^{i\Omega_k t} A_k'$. With regard to the drift matrix of (20), the real parts of the diagonal elements originate from $-\mathbf{\Gamma}_3$, and the off-diagonal elements and the imaginary parts of the diagonal elements are the contributions by $-i\mathbf{\Omega}_3$. The stationary spectrum of TRM is given by $\mathrm{diag}(n_k, N_k, N_k')$, and $N_k = \mathcal{A} n_k$ and $N_k' = \mathcal{A}' n_k$ hold [2]. We impose the constraints $\Sigma = \sqrt{2\nu_R N_k}$ and $\Sigma' = \sqrt{2\nu_R' N_k'}$ to fulfill the FD theorem (4).

III. RESONANT DUET MODEL (RDM)

In Sec. III A we analyze TRM (20) to get a clear understanding of how B_k and B_k' affect the evolution of b_k. In Sec. III B we simplify TRM to obtain a Markov model governing the duet (b_k, B_k). In Sec. III C we discuss the connection between the new and existing reduced-order models and the advancement of our development.

A. Marginalization

The dynamic equation for B_k' in TRM (20) is integrated to yield the expression

$$B_k'(t) = e^{-\nu' t} B_k'(0) - \int_0^t ds\, \mathcal{A}' e^{-\nu'(t-s)} b_k(s) + \Sigma' \int_0^t ds\, e^{-\nu'(t-s)} \dot{W}_{B'}. \tag{21}$$

The first term of the right-hand side of (21) can be dropped without causing any statistically meaningful error, and the last term denoted by $R_{\nu'}$ is the Ornstein-Uhlenbeck (OU) process satisfying $\langle R_{\nu'}(t), R_{\nu'}(0)\rangle / n_k = \mathcal{A}' e^{-\nu' t}$. Hence Eq. (21) can be rephrased as

$$B_k' \triangleq -\Gamma_{\nu'} * b_k + R_{\nu'}, \quad \Gamma_{\nu'}(t) \equiv \mathcal{A}' e^{-\nu' t} \tag{22}$$

and TRM (20) can be marginalized into

$$\begin{aligned} \dot{b}_k &= B_k - \Gamma_{\nu'} * b_k + R_{\nu'}, \\ \dot{B}_k &= -\mathcal{A} b_k - \nu B_k + \Sigma \dot{W}_B. \end{aligned} \tag{23}$$

Likewise, we can verify

$$B_k \triangleq -\Gamma_\nu * b_k + R_\nu, \quad \Gamma_\nu(t) \equiv \mathcal{A} e^{-\nu t}, \tag{24}$$

where R_ν is the OU process characterized by $\langle R_\nu(t), R_\nu(0)\rangle / n_k = \Gamma_\nu(t)$, and Eq. (23) can be marginalized into

$$\begin{aligned} \dot{b}_k &= B_k + B_k' \\ &= (-\Gamma_\nu * b_k + R_\nu) + (-\Gamma_{\nu'} * b_k + R_{\nu'}). \end{aligned} \tag{25}$$

This univariate form of TRM clearly reveals that the dynamics of b_k is independently driven by B_k and B_k', and that both B_k and B_k' force and dissipate the motion of b_k in the non-Markovian manner.

B. Timescale separation limit

We turn our attention to the true dynamical system model. In terms of

$$B_k = e^{i\Omega_k t} A_k = -i \sum_{k_{123}, k_{12} \neq k, |\Omega_{3k}^{12}| < \varepsilon} W_{3k}^{12} \delta_{3k}^{12} b_1 b_2 b_3^* e^{-i\Omega_{3k}^{12} t},$$

$$B_k' = e^{i\Omega_k t} A_k' = -i \sum_{k_{123}, k_{12} \neq k, |\Omega_{3k}^{12}| > \varepsilon} W_{3k}^{12} \delta_{3k}^{12} b_1 b_2 b_3^* e^{-i\Omega_{3k}^{12} t},$$

Eqs. (12) and (14) can be written as

$$\dot{b}_k \doteq e^{i\Omega_k} N_k^{\mathrm{eff}} = B_k + B_k'. \tag{26}$$

Equation (26) confirms that the form of (25) complies with our prior knowledge concerning the effect of N_k^{eff} on the system variable. This consistency with the argument provided in the first paragraph of Sec. II B obviously supports the plausibility of the approximations in (18). In fact, our modeling of N_k^{eff} by means of (22) and (24) is distinguished from the classical ones made without the decomposition of N_k^{eff} and hence the statistical model (25) is conceptually a more refined version of the univariate model than the existing ones presented in Sec. III C. The task is admittedly not so meaningful unless the timescales of B_k and B_k' are very different from one another. Our key intuition with regard to this issue is that, such as in the case of Brownian motion, the influence of the far-from resonances on the target system variable is at most transient and B_k' is significantly faster than the typical evolution of b_k.

We return now to TRM and show that the direct consequence of this sharp timescale separation is the approximation of (22) by

$$B_k' \doteq -\gamma b_k + \sigma \dot{W}_b, \tag{27}$$

where $\gamma = \mathcal{A}'/\nu_R'$ and $\sigma = \sqrt{2\gamma n_k}$. That B_k' is rapidly varying corresponds to a tiny relaxation time of B_k' and the real part of $\nu' = \nu_R' + i\nu_I'$ is a large number. As $\nu_R' \to \infty$, one has $\Gamma_{\nu'}[\omega] = \frac{\mathcal{A}'}{i\omega + \nu'} \to \gamma$ since $\mathcal{A}' = \gamma \nu_R'$ and hence $\Gamma_{\nu'}(t) \to \gamma \delta_+(t)$. Moreover, in this limit, one has $\langle R_{\nu'}(t), R_{\nu'}(s)\rangle \to \sigma^2 \delta(t-s)$ and $R_{\nu'} \to \sigma \dot{W}_b$ where $\delta(t)$ is Dirac delta function.

We are tempted by (27) to eliminate the fast variable B_k' of TRM (20). This can be achieved by taking the timescale separation limit $\nu_R' \to \infty$ to (23), and such an operation gives rise to the Markov model

$$(\textbf{RDM})\quad \begin{pmatrix} \dot{b}_k \\ \dot{B}_k \end{pmatrix} = \begin{pmatrix} -\gamma & 1 \\ -\mathcal{A} & -\nu \end{pmatrix} \begin{pmatrix} b_k \\ B_k \end{pmatrix} + \begin{pmatrix} \sigma \dot{W}_b \\ \Sigma \dot{W}_B \end{pmatrix} \tag{28}$$

named as resonant duet model (RDM). Equation (28) is transformed into

$$(\textbf{RDM})\quad \begin{cases} \dot{a}_k = -(i\Omega_k + \gamma) a_k + A_k + \sigma \dot{W}_a, \\ \dot{A}_k = -\mathcal{A} a_k - (i\Omega_k + \nu) A_k + \Sigma \dot{W}_A, \end{cases} \tag{29}$$

in terms of the genuine variables (a_k, A_k). This removal of one degree of freedom is indeed advantageous as the dimension reduction essentially gives no harm to the Markovian character of TRM. The stationary spectrum of RDM is given by $\mathrm{diag}(n_k, N_k)$, and $N_k = \mathcal{A} n_k$ holds. The relations $\sigma = \sqrt{2\gamma n_k}$ and $\Sigma = \sqrt{2\nu_R N_k}$ are satisfied according to the FD theorem.

C. Reduction to the univariate models

Together with RDM (29), here we repeat the same procedures with those applied to TRM, i.e., performing the marginalization and taking the timescale separation limit. In the first step, as in Sec. III A, we obtain the univariate form of RDM

$$(\textbf{ARM})\quad \dot{a}_k = -i\Omega_k a_k - \gamma a_k - \int_0^t ds\, \mathcal{A} e^{-(i\Omega_k + \nu)(t-s)} a_k(s) + \sigma \dot{W}_a + e^{-i\Omega_k t} R_\nu. \tag{30}$$

This non-Markov model is called autoregressive model (ARM). In the second step, as in Sec. III B, Eq. (30) is approximated by the Markov model

$$(\textbf{MSM})\quad \dot{a}_k = -i\Omega_k a_k - \bar{\gamma} a_k + \bar{\sigma} \dot{W}_a, \tag{31}$$

under the situation of $\nu_R \gg 1$. Here $\bar{\gamma}$ and $\bar{\sigma} = \sqrt{2\bar{\gamma} n_k}$ are used in order to avoid confusion with γ and σ. Equation (31) is referred to as mean stochastic model or MSM for short.

In fact, ARM and MSM are the existing models; they were proposed and studied by the author in [17,18]. Specifically, following the recipe that can be found in the first paragraph of Sec. II B, we have devised these univariate models from substituting $\Gamma_a(t) = \gamma \delta_+(t) + \mathcal{A} e^{-(i\Omega_k + \nu)t}$ and $\Gamma_a(t) = \bar{\gamma} \delta_+(t)$ into the GLE (7). Unfortunately, the development is rather heuristic and carried out with no serious justification for this particular choice of the kernel functions. By contrast, their construction by the multivariate modeling approach is significantly more systematic and the argument using the GLE (16) enables a comprehensive explanation for the non-Markovian form of the memory kernel of ARM.

We emphasize that the advantage of the current derivation is not limited to the theoretical rigor. It also brings the practical benefit in clarifying the validity regime of RDM. More precisely, the non-Markovian process (30) was shown to be an accurate model when k is large and a_k is a rapidly varying shortwave. In this case, it is desired to use ARM or RDM in describing the true turbulent signal. When k is small and a_k is a slowly varying longwave, however, MSM is already a decent approximation and it is redundant to introduce the variable A_k.

TABLE I. MSM and RDM are Markov models; ARM is a non-Markov model.

Governing variables	Longwave (small k)	Shortwave (large k)
a_k	MSM (31)	ARM (30)
(a_k, A_k)		RDM (29)

As a summary, Table I provides an overview of the differences between the simplified models.

Moreover, the new framework offers a means to strengthen the numerical validation of ARM. Though ARM can be reformulated as RDM along with the formal definition of A_k that is essentially given by the one in (24), RDM is distinguished from ARM in the aspect that the auxiliary variable A_k is identified in terms of the canonical system variables. The consequence is that it is possible to simulate a single trajectory of $A_k(t)$ as an addition to that of $a_k(t)$ from a direct numerical integration of the true turbulence model and, as will be seen in the next section, we can utilize this plenty of data to give further depth to the plausibility of the reduced-order models.

So far our discussion is made while keeping in mind the situation of the strong nonlinearity. We comment that the applicability of our formalism continues to hold independently of the strength of nonlinearity. Nonetheless, for the WT system with a weak nonlinearity, the concept of the multivariate models like TRM and RDM has no great merit because in such cases MSM (or its variant created from replacing Ω_k by ω_k) is capable of imitating with reasonable accuracy the true underlying signal that changes in a Markovian fashion.

IV. MODEL VALIDATION

In order to validate the proposed multivariate models, we here resort to the numerical simulations. Section IV A presents the testbed. Section IV B explains in detail the settings for the generation of the true turbulent signal. Section IV C discusses the training of the reduced-order models. Section IV D provides the simulation results.

A. Testbed

We examine the effectiveness of our reduced-order modeling framework in the context of the generalized Majda-McLaughlin-Tabak (MMT) model [11,12,19]. A finite dimensional approximation of the MMT model is given by the Hamiltonian system (2), along with $\omega_k = |\hat{k}|^\alpha$ and $W_{3k}^{12} = |\hat{k}_1 \hat{k}_2 \hat{k}_3 \hat{k}|^{\frac{\beta}{4}}$ where $\hat{k} \equiv k\pi/N$ and N is the total number of Fourier modes [18]. Here δ_{3k}^{12} is equal to unity if $k_1 + k_2 - k_3 - k$ is a multiple of N, and zero otherwise. Accordingly, Eq. (15a) is to be understood in the sense of the modulus N. The generalized MMT model possesses three parameters: χ and $\alpha, \beta (>0)$. When $\chi = 1$, the system reduces to the original MMT model introduced for the numerical study of the wave turbulence theory [19]. In this case, the MMT model includes a number of familiar wave systems as specific cases. For example, if $\alpha = 2$ and $\beta = 0$, the system corresponds to the nonlinear Schrödinger equation, while the case of $\alpha = 1/2$, $\beta = 3$ mimics the scaling present in water waves [20]. When $\chi = 0$, the model equation has a certain similarity with the

Clebsch-variables formulation of the 3D Euler equations in fluid mechanics [21].

B. True turbulent signal

The MMT model allows for an excellent case study not only because of the coverage of a wide scope of dynamical systems ranging from nonlinear waves to fluid motion, but also because of the occurrence of effective nontrivial four-wave resonances. More precisely, Eq. (9) becomes

$$\Omega_k \doteq \chi|\hat{k}|^\alpha + C_\beta|\hat{k}|^{\frac{\beta}{2}}, \quad C_\beta \equiv 2\sum_{k'}|\hat{k}'|^{\frac{\beta}{2}}\langle|a_{k'}|^2\rangle$$

in case of the MMT system. With regard to the resonance condition (15), the linear dispersion relation $\Omega(k) = \chi|\hat{k}|^\alpha$ admits the nontrivial resonances only if $\chi = 1$ and $\alpha < 1$. However, due to the scalings of the effective dispersion relation $\Omega(k) = \chi|\hat{k}|^\alpha + C_\beta|\hat{k}|^{\frac{\beta}{2}}$, the nontrivial resonances can occur if $\frac{\beta}{2} < 1$, irrespective of the value of χ and α [12].

Here we pick up two instances of the MMT system for which the parameters are given by (i) $\chi = 1, \alpha = 1/2, \beta = 3$, and $C_\beta = 3/10$, and (ii) $\chi = 0, \beta = 1$, and $C_\beta = 9/10$. Notice that C_β characterizes the MMT system in thermal equilibrium and the selected values give rise to the strong nonlinearities [18]. The Fourier index $k \in \mathbb{Z}$ is in the range of $-N/2 < k \leqslant N/2$ with $N = 512$, and we take $k = 192$ for a representative case of the high k and shortwave signal a_k.

According to the criterion provided in Appendix C, we determine the variable A_k in the way that the relative error between the two different approximations of $N_k = \langle|A_k|^2\rangle$ given by (C1) and (C2) is less than 1 percent. The designated tolerance ensures that, in both cases of $\chi = 1$ and $\chi = 0$, the variable A_k for $k = 192$ consists of about 37 percent of the total number of terms constituting the effective nonlinearity $N^{\rm eff}$. We draw attention to the fact that the application requires the knowledge of the equilibrium spectrum n_k for all wavenumber k. For this we use $n_k = \theta/\Omega_k$ where $\theta = \langle a_k^*\delta\mathcal{H}/\delta a_k^*\rangle$ is the temperature. This theoretical prediction is the generalization of the classical Rayleigh-Jeans distribution $n_k = \theta/\omega_k$ for the weakly nonlinear WT system, and is numerically validated in [11,12].

C. Model training

We first train ARM (or equivalently RDM) using the two-point function in time. We recall that the autocorrelation of $b_k(t)$ is given by

$$\frac{\langle b_k(t)b_k^*(0)\rangle}{n_k} = e^{-\frac{\tau}{2}t}\left[\cos\left(\frac{\mathcal{R}t}{2}\right) - \frac{\gamma-\nu}{\mathcal{R}}\sin\left(\frac{\mathcal{R}t}{2}\right)\right],$$

where $\tau = \gamma + \nu$ and $\mathcal{R} = \sqrt{4\mathcal{A} - (\gamma-\nu)^2}$ [17]. This statistical quantity is numerically approximated using the single trajectory of the true turbulent signal under the ergodicity assumption, and the associated parameters are learned via the curve fitting using the autocorrelation functions.

We next train TRM by capturing our assumption of the timescale separation, i.e., by taking the value of ν_R' large enough. The remaining set of parameters are kept or suitably chosen in a way that the autocorrelation matrix of (b_k, B_k) indicated by TRM are almost the same with that by RDM.

D. Autocorrelation matrices

Having trained TRM (20), we calculate the associated autocorrelation matrix $\langle\mathbf{B}_3(t), \mathbf{B}_3(0)\rangle$ where $\mathbf{B}_3 = (b_k, B_k, B_k')^T$. The theoretical predictions in the physical space and Fourier space are obtained using Eqs. (5) and (6). The elements are drawn as a function of time in both cases of $\chi = 1$ (Figs. 1 and 2) and $\chi = 0$ (Figs. 3 and 4). We also depict the corresponding numerically measured quantities of the MMT model. Despite the elementary and simple training of TRM based on the partial information matching between the correlation matrices, the remaining elements show qualitatively analogous behaviors. Because the drift matrix determines the autocorrelation matrix as can be seen in Eq. (6), this simulation result concerning the two-point functions in time lets us to conclude that the approximations in (18) and (27) are reasonable and that both TRM and RDM are good approximate models.

V. SUMMARY AND FUTURE WORKS

Writing the total Hamiltonian (1) as $\mathcal{H} = \mathcal{H}_2 + \mathcal{H}_4$, consider the decomposition of the Hamiltonian responsible for the nonlinear interactions: $\mathcal{H}_4 = \mathcal{H}_{\rm TR} + \mathcal{H}_{\rm NR}$. Here $\mathcal{H}_{\rm TR}$ gathers the terms corresponding to the trivial resonances (TR), and $\mathcal{H}_{\rm NR}$ gathers the terms corresponding to the nonlinearities remainder (NR). In our prior work [11,12] we made use of the Mori-Zwanzig theory to demonstrate that $\mathcal{H}_{\rm TR}$ controls the oscillation of a_k. Defining $\mathcal{H}_2^{\rm eff} = \mathcal{H}_2 + \overline{\mathcal{H}}_{\rm TR}$ where the upper bar denotes the mean-field approximation leading to

$$\frac{\delta\mathcal{H}_2^{\rm eff}}{\delta a_k^*} = \Omega_k a_k,$$

the effective Hamiltonian $\mathcal{H}^{\rm eff} = \mathcal{H}_2^{\rm eff} + \mathcal{H}_{\rm NR}$ provides us with the comprehensible framework that captures accurately the realistic behavior exhibited by the strongly nonlinear wave turbulence system.

Our achievement in this work is the further decomposition of the Hamiltonian responsible for the effective nonlinearity: $\mathcal{H}_{\rm NR} = \mathcal{H}_{\rm NTR} + \mathcal{H}_{\rm FFR}$. Here NTR stands for the nontrivial (and near) resonances, and FFR stands for the far-from resonances. In terms of the variables in (13), these two Hamiltonians are defined through

$$\frac{\delta\mathcal{H}_{\rm NTR}}{\delta a_k^*} = iA_k,$$

$$\frac{\delta\mathcal{H}_{\rm FFR}}{\delta a_k^*} = iA_k',$$

and we once again make use of the Mori-Zwanzig theory to demonstrate that both $\mathcal{H}_{\rm NTR}$ and $\mathcal{H}_{\rm FFR}$ independently control the forcing and dissipation of a_k. In short, breaking the quartic Hamiltonian into smaller pieces, i.e., $\mathcal{H}_4 = \mathcal{H}_{\rm TR} + \mathcal{H}_{\rm NTR} + \mathcal{H}_{\rm FFR}$, we perform the detailed analysis of each of the individuals to identify their precise roles in determining the motion of a_k.

Moreover, with the aid of the wave turbulence theory, we build a physical insight into a possibly sharp separation between the timescales induced by $\mathcal{H}_{\rm NTR}$ and $\mathcal{H}_{\rm FFR}$, and numerically demonstrate this is indeed the case when k is

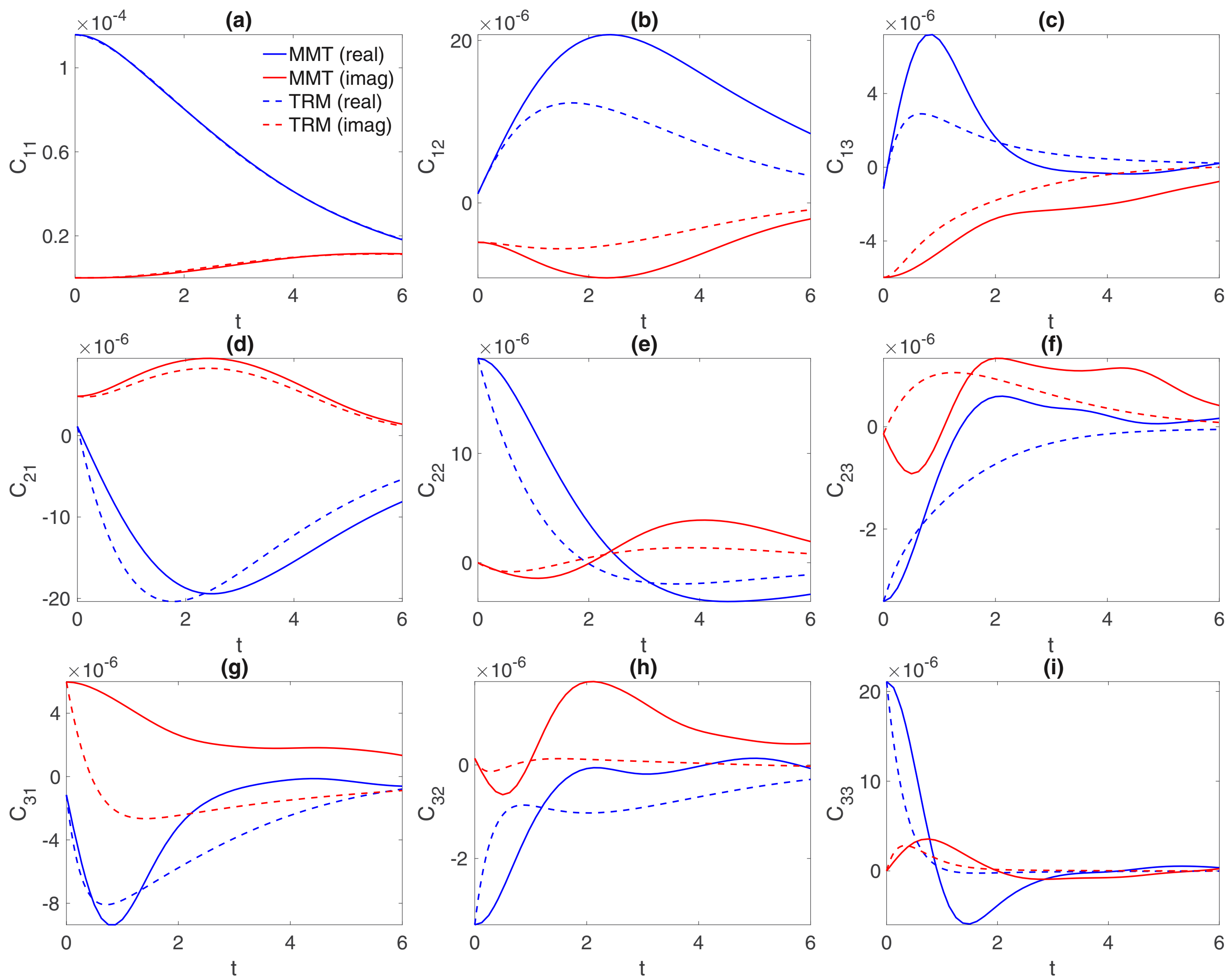

FIG. 1. The elements of $\mathbf{C} = \langle \mathbf{B}_3(t), \mathbf{B}_3(0) \rangle$ are drawn as a function of time. The MMT system is parametrized by $\chi = 1$, $\alpha = \frac{1}{2}$, and $\beta = 3$, and the wave-numbers range $-256 < k \leqslant 256$. The autocorrelation matrix of the true signal a_k is numerically measured when $k = 192$. The counterparts of TRM are depicted for comparison purposes. (a), (b), (c), (e), and (i) The solid blue (upper) and red (lower) lines are, respectively, for the real and imaginary parts of MMT; the dashed blue (upper) and red (lower) lines for the real and imaginary parts of TRM. (d), (f), (g), and (h) The solid blue (lower) and red (upper) lines are, respectively, for the real and imaginary parts of MMT; the dashed blue (lower) and red (upper) lines for the real and imaginary parts of TRM.

large and a_k is a rapidly varying shortwave. This discrepancy between the typical evolutions of A_k and A'_k guides us to statistically project the extremely complex motion of the single wave-profile a_k onto the two-dimensional space spanned by (a_k, A_k) in order to suppress the time-lag effect from which the conventional projection onto the one-dimensional space spanned by a_k inevitably suffers and to successfully address the quasi-Markovian property of the strongly nonlinear wave turbulence signal in the high wave-number domain. Our task is highlighted by the trade-off between the Markovian nature of the bivariate model called RDM and the lower dimension of the non-Markov univariate model called ARM.

The present work focuses on the rigorous derivation of the multivariate stochastic models out of the true turbulence model, and in the future we will naturally proceed to the investigation of their practical utilities. For this, the research direction currently being pursued by the author is twofold. One way is to make use of the suggested reduced-order models as the platform for the theoretical characterization of the strong wave turbulence system. The other way is to solve the uncertainty quantification problems associated with the true dynamical system model, such as probabilistic filtering, for which the simplified models are employed as stochastic emulator.

ACKNOWLEDGMENT

This work is supported by the Strategic Research Grant of City University of Hong Kong (No. 7005592).

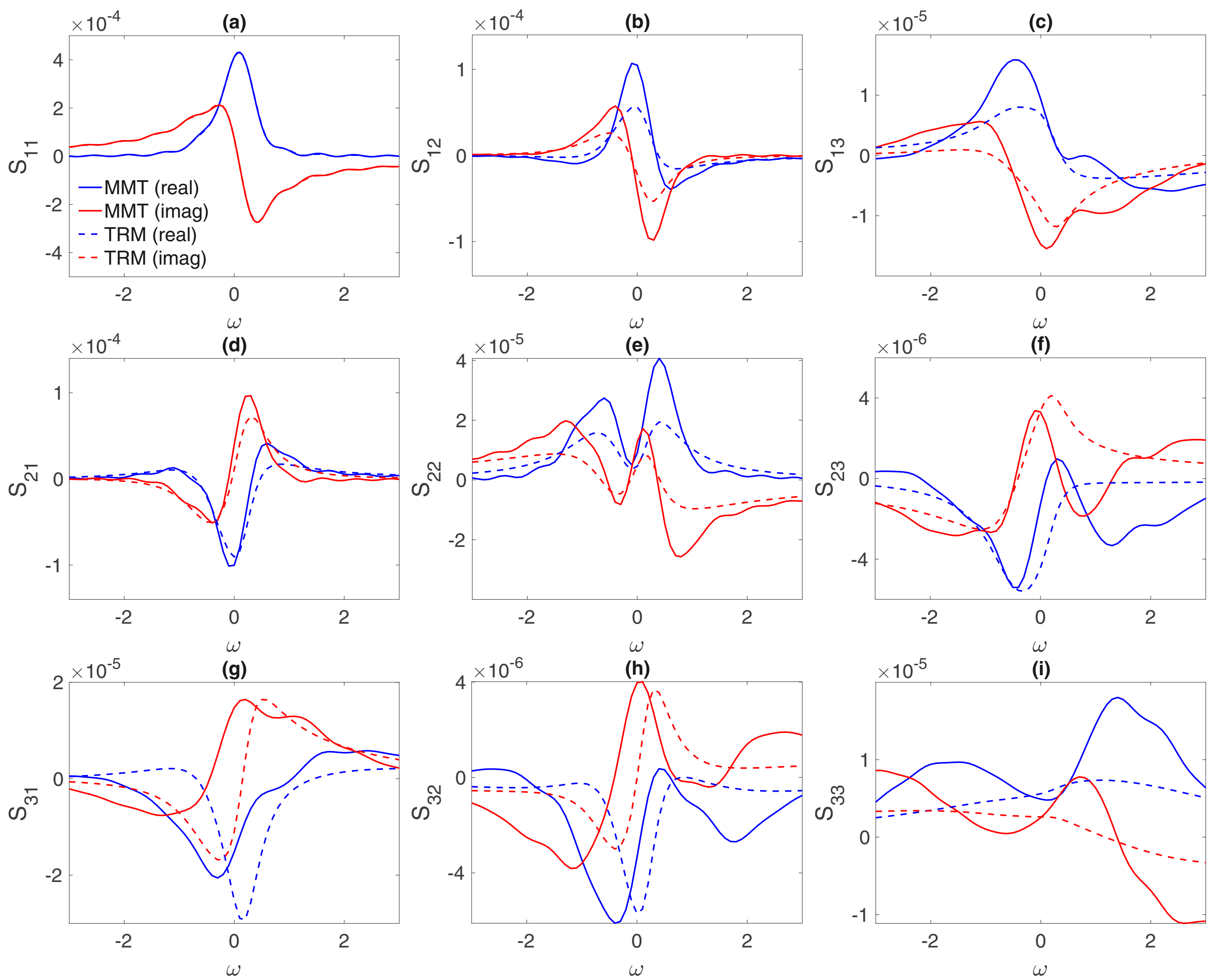

FIG. 2. The elements of $\mathbf{S} = \mathfrak{F}\{\langle \mathbf{B}_3(t), \mathbf{B}_3(0)\rangle\}$ are drawn as a function of time. The setting is the same with Fig. 1. The line specifications are the same with Fig. 1.

APPENDIX A: PROJECTION MATRIX $\boldsymbol{\Omega}_3$

First we calculate the stationary spectrum of $\mathbf{A}_3 = (a_k, iA_k, iA'_k)^T$ and obtain the approximation $\langle \mathbf{A}_3, \mathbf{A}_3\rangle \doteq \text{diag}(n_k, N_k, N'_k)$ for the true dynamical system model (1) and (2). Recall that, invoking the random phase argument (RPA), the four-point and six-point correlation functions can be approximated by

$$\langle a_1 a_2 a_3^* a_k^*\rangle \doteq n_1 n_2 (\delta_{13}\delta_{2k} + \delta_{1k}\delta_{23}),$$
$$\langle a_1 a_2 a_3^* a_{1'}^* a_{2'}^* a_{3'}\rangle \doteq n_1 n_2 n_3 \delta_{33'}(\delta_{11'}\delta_{22'} + \delta_{12'}\delta_{21'}), \tag{A1}$$

respectively [4]. Here n_j is shorthand notation for $n_{k_j} = \langle |a_{k_j}|^2\rangle$. In this RPA approximation, the elements of $\mathbf{A}_3$ are mutually orthogonal, i.e., the cross-correlations

$$\langle a_k, A_k\rangle = -i \sum_< W_{3k}^{12}\delta_{3k}^{12}\langle a_1 a_2 a_3^* a_k^*\rangle \doteq 0,$$

$$\langle a_k, A'_k\rangle = -i \sum_> W_{3k}^{12}\delta_{3k}^{12}\langle a_1 a_2 a_3^* a_k^*\rangle \doteq 0,$$

$$\langle A_k, A'_k\rangle = \sum_< W_{3k}^{12}\delta_{3k}^{12} \sum_{|\Omega_{3'k}^{1'2'}|>\varepsilon} W_{3'k}^{1'2'}\delta_{3'k}^{1'2'} \times \langle a_1 a_2 a_3^* a_{1'}^* a_{2'}^* a_{3'}\rangle \doteq 0, \tag{A2}$$

vanish simply because there are no survival terms. Here $\sum_<$ and $\sum_>$ abbreviate the summation notations shown in (13).

Next, we calculate the matrix $\langle \dot{\mathbf{A}}_3, \mathbf{A}_3\rangle$.

(1) The elements of the first row are approximated by

$$\langle \dot{a}_k, a_k\rangle = -i\Omega_k\langle a_k, a_k\rangle,$$
$$\begin{aligned}\langle \dot{a}_k, A_k\rangle &= \Big\langle -i\Big(\chi\omega_k a_k + \sum W_{3k}^{12}\delta_{3k}^{12} a_1 a_2 a_3^*\Big) \\ &\quad - i\sum_< W_{3k}^{12}\delta_{3k}^{12} a_1 a_2 a_3^*\Big\rangle \\ &\doteq \sum_< W_{3k}^{12}\delta_{3k}^{12}\Omega_k\langle a_1^* a_2^* a_3 a_k\rangle + \langle A_k + A'_k, A_k\rangle \\ &\doteq \langle A_k, A_k\rangle,\end{aligned}$$

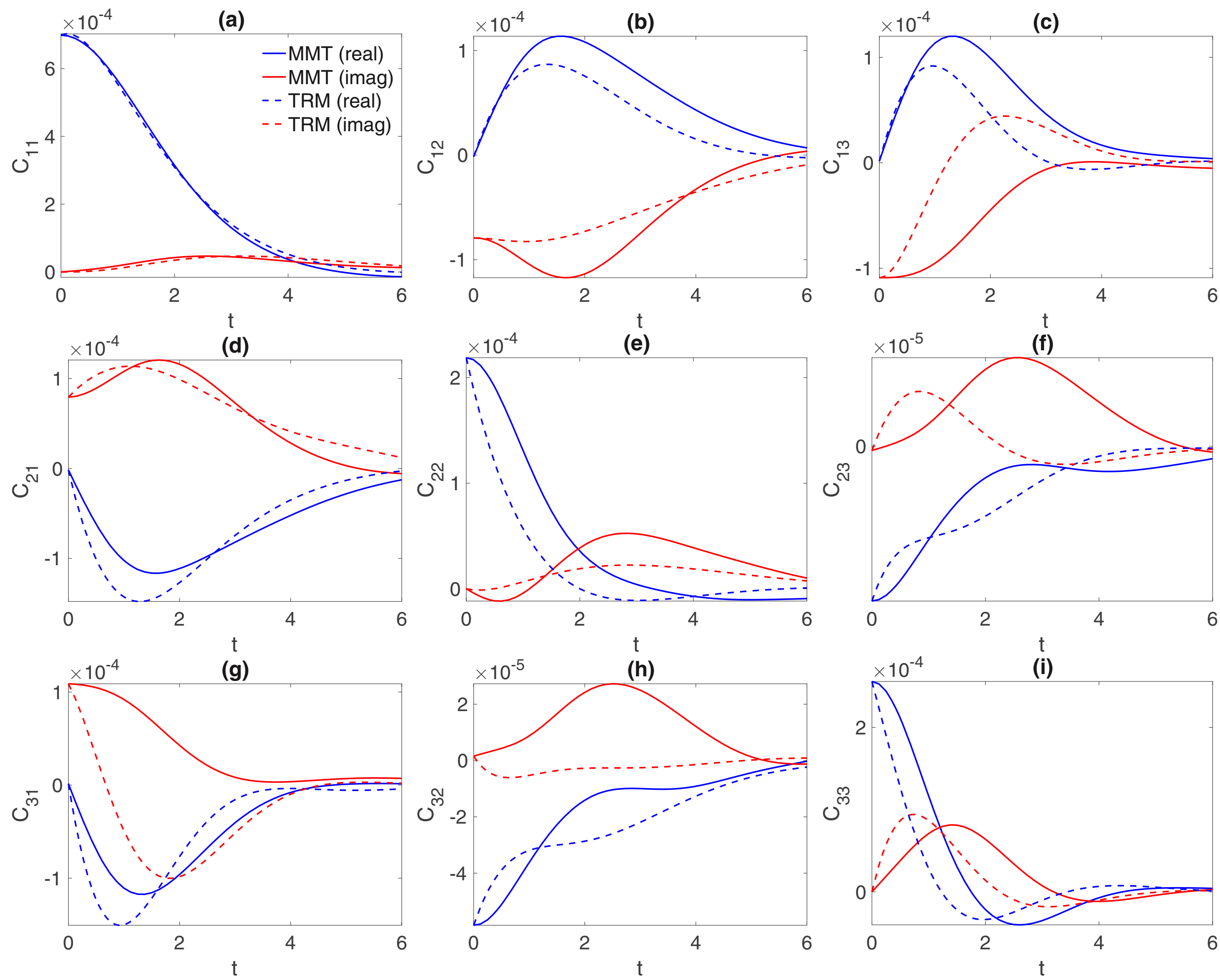

FIG. 3. The elements of $\mathbf{C} = \langle \mathbf{B}_3(t), \mathbf{B}_3(0)\rangle$ are drawn as a function of time. The setting is the same with Fig. 1, except that the MMT system is parametrized by $\chi = 0$ and $\beta = 1$. The line specifications are the same with Fig. 1.

$$
\begin{aligned}
\langle \dot{a}_k, A'_k\rangle &\doteq \sum_{>} W^{12}_{3k}\delta^{12}_{3k}\Omega_k\langle a_1^* a_2^* a_3 a_k\rangle + \langle A_k + A'_k, A'_k\rangle \\
&\doteq \langle A'_k, A'_k\rangle,
\end{aligned}
$$

for which the use is made of Eqs. (8), (9), and (13). Notice that the same results can be easily obtained from using (14), the linear expansion of $\dot{a}_k$ in terms of $\mathbf{A}_3$, and the mutual orthogonality of the elements of $\mathbf{A}_3$. Henceforth, for simplicity of the calculation of the correlation functions, we make use of the mean-field equation (14) instead of the original dynamic equation (2).

(2) The elements of the second row are approximated by

$$
\langle \dot{A}_k, a_k\rangle = -\langle A_k, \dot{a}_k\rangle \doteq -\langle A_k, A_k\rangle, \tag{A3a}
$$

$$
\begin{aligned}
\langle \dot{A}_k, A_k\rangle &\doteq -\left\langle \sum_{<} W^{12}_{3k}\delta^{12}_{3k}(\Omega_1 + \Omega_2 - \Omega_3)a_1 a_2 a_3^*, A_k\right\rangle \\
&\equiv -i\tilde{\Omega}_k\langle A_k, A_k\rangle,
\end{aligned}\tag{A3b}
$$

$$
\langle \dot{A}_k, A'_k\rangle \doteq -i\tilde{\Omega}_k\langle A_k, A'_k\rangle \doteq 0. \tag{A3c}
$$

As for (A3a), recall that the Liouville operator is skew Hermitian. As for (A3b), the three terms τ_k, $\mathcal{T}_k$, and $\tilde{\mathcal{T}}_k$ comprising the evolution of A_k are defined from

$$
\begin{aligned}
\dot{A}_k &= -i\sum_{<} W^{12}_{3k}\delta^{12}_{3k}(\dot{a}_1 a_2 a_3^* + a_1\dot{a}_2 a_3^* + a_1 a_2\dot{a}_3^*) \\
&\doteq -\sum_{<} W^{12}_{3k}\delta^{12}_{3k}(\Omega_1 + \Omega_2 - \Omega_3)a_1 a_2 a_3^* \\
&\quad - i\sum_{<} W^{12}_{3k}\delta^{12}_{3k}(A_1 a_2 a_3^* + a_1 A_2 a_3^* + a_1 a_2 A_3^*) \\
&\quad - i\sum_{<} W^{12}_{3k}\delta^{12}_{3k}(A'_1 a_2 a_3^* + a_1 A'_2 a_3^* + a_1 a_2 A_3'^*) \\
&\equiv \tau_k + \mathcal{T}_k + \tilde{\mathcal{T}}_k,
\end{aligned}\tag{A4}
$$

where (14) is used. Suppose that τ_k contains only the terms corresponding to the exact nontrivial resonances, then the quantity equals to $-i\Omega_k A_k$ because the coefficient $\Omega_1 + \Omega_2 - \Omega_3 = \Omega_k$ due to the resonance quartet condition (15b) is independent of the summation index. The presence of the near resonances leads us to introduce $\tilde{\Omega}_k$, which is close to Ω_k and is defined by $\langle \tau_k, A_k\rangle \equiv -i\tilde{\Omega}_k\langle A_k, A_k\rangle$. This statistical

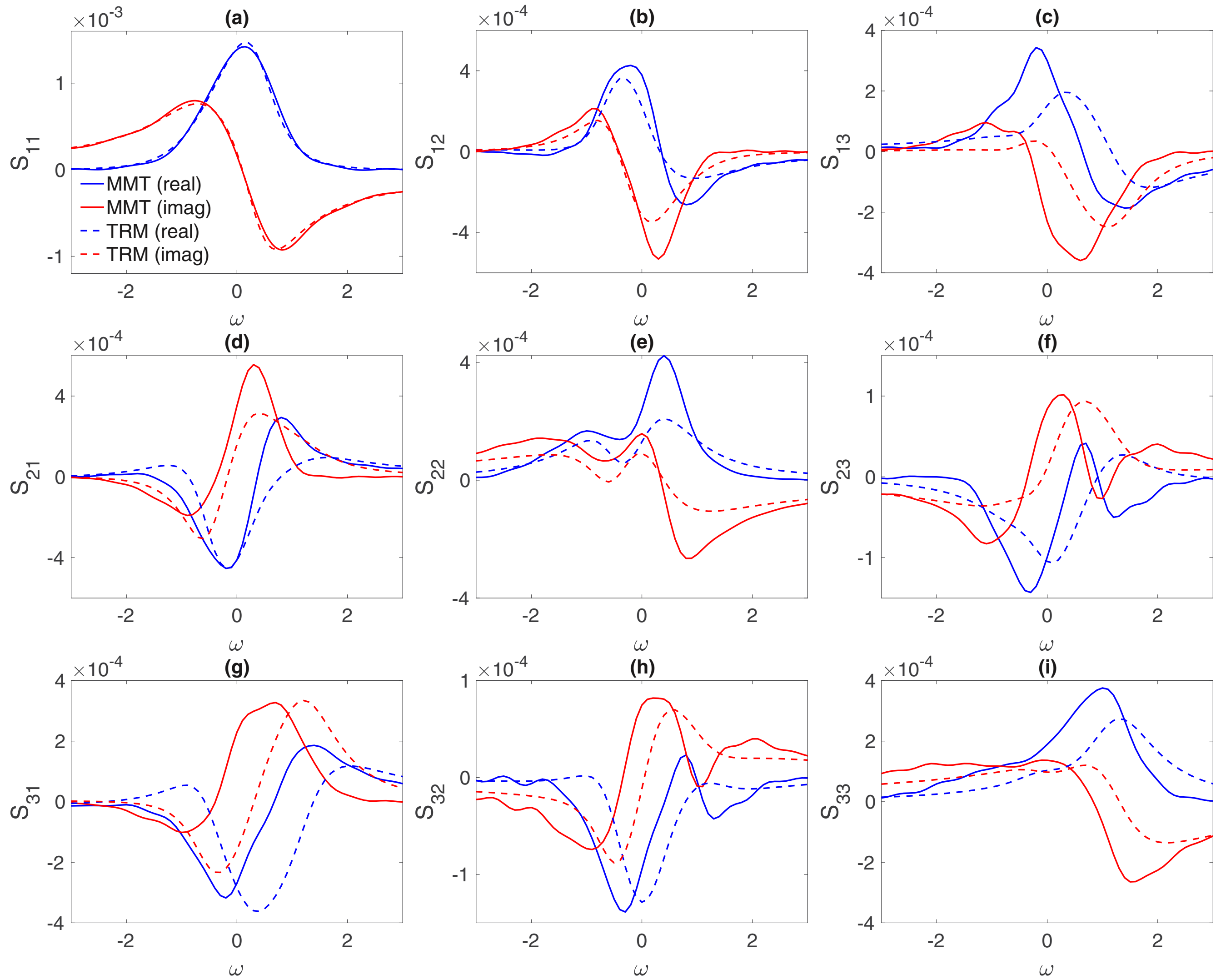

FIG. 4. The elements of $\mathbf{S} = \mathfrak{F}\{\langle \mathbf{B}_3(t), \mathbf{B}_3(0)\rangle\}$ are drawn as a function of time. The setting is the same with Fig. 3. The line specifications are the same with Fig. 1.

approximation is symbolically represented by

$$\tau_k \simeq -i\tilde{\Omega}_k A_k. \tag{A5}$$

Because the amplitude of a_k is small (this is true when k is large as discussed in Sec. IV B, and the reason why this case is under our consideration can be found in Sec. III C) and both $\mathcal{T}_k$ and $\tilde{\mathcal{T}}_k$ are of higher order than τ_k, the inner products between $\mathcal{T}_k$, $\tilde{\mathcal{T}}_k$, and A_k are significantly smaller than $\langle \tau_k, A_k\rangle$. Thus $\langle \mathcal{T}_k, A_k\rangle$ and $\langle \tilde{\mathcal{T}}_k, A_k\rangle$ are ignored in obtaining (A3b). As for (A3c), the same argument is applied to the inner products between $\mathcal{T}_k$, $\tilde{\mathcal{T}}_k$, and A'_k.

(3) The elements of the third row are approximated by

$$\begin{aligned}
\langle \dot{A}'_k, a_k\rangle &= -\langle A'_k, \dot{a}_k\rangle \doteq -\langle A'_k, A'_k\rangle,\\
\langle \dot{A}'_k, A_k\rangle &= -\langle A'_k, \dot{A}_k\rangle \doteq -i\tilde{\Omega}_k\langle A'_k, A_k\rangle \doteq 0,\\
\langle \dot{A}'_k, A'_k\rangle &\doteq -\left\langle \sum_{>} W^{12}_{3k}\delta^{12}_{3k}(\Omega_1+\Omega_2-\Omega_3)a_1a_2a_3^*, A'_k\right\rangle\\
&\equiv -i\Omega'_k\langle A'_k, A'_k\rangle.
\end{aligned}$$

The calculations are similar to the ones for the second row. Like (A4), the three terms τ'_k, $\mathcal{T}'_k$, and $\tilde{\mathcal{T}}'_k$ are defined from $\dot{A}'_k \equiv \tau'_k + \mathcal{T}'_k + \tilde{\mathcal{T}}'_k$ for which the summation is with respect to $\sum_>$ and over the far-from resonances. Introducing Ω'_k to represent the approximation $\tau'_k \simeq -i\Omega'_k A'_k$, analogous to (A5), we add a remark that the coefficient Ω'_k can deviate substantially from Ω_k due to the far-from resonances. The argument using the small magnitude is again applied to ignore the higher order terms than those of $\langle \tau'_k, A_k\rangle$ and $\langle \tau'_k, A'_k\rangle$.

Last, Eq. (17) follows from

$$\begin{aligned}
-i\mathbf{\Omega}_3 &= \frac{\langle \dot{\mathbf{A}}_3, \mathbf{A}_3\rangle}{\langle \mathbf{A}_3, \mathbf{A}_3\rangle}\\
&\doteq -i\begin{pmatrix} \Omega_k n_k & N_k & N'_k\\ N_k & \tilde{\Omega}_k N_k & 0\\ N'_k & 0 & \Omega'_k N'_k\end{pmatrix}\begin{pmatrix} n_k & 0 & 0\\ 0 & N_k & 0\\ 0 & 0 & N'_k\end{pmatrix}^{-1}.
\end{aligned}$$

APPENDIX B: MEMORY KERNEL $\boldsymbol{\Gamma}_3(t)$

In view of (14), (16), and (17), the first component of $\mathbf{R}_3$ is zero and the random noise can be written by $\mathbf{R}_3 = (0, R_A, R_{A'})^T$. The FD theorem ensures that the elements in the first row and first column of $\boldsymbol{\Gamma}_3(t)$ vanish. Here our reasoning about the form of the memory kernel is that (i) the remaining elements of $\boldsymbol{\Gamma}_3(t)$ are well approximated by the constant multiplication of the Dirac delta function, i.e.,

$$\boldsymbol{\Gamma}_3(t) \doteq \begin{pmatrix} 0 & 0 & 0 \\ 0 & \nu_R & \mu \\ 0 & \mu' & \nu_R' \end{pmatrix} \delta_+(t), \tag{B1}$$

and (ii) the off-diagonal elements of μ and μ' are minor and negligible. In a word, R_A and $R_{A'}$ are originally correlated in a nonwhite manner, but can be approximated by the independent white noises with no significant loss of accuracy. From which, alongside the FD theorem, the approximations in (18) are produced.

First, we discuss the form of (B1). It basically follows from the statistical approximations

$$\mathcal{T}_k \simeq -\mathcal{A}_k a_k - \nu_R A_k + \Sigma \dot{W}_A, \tag{B2a}$$

$$\tilde{\mathcal{T}}_k \simeq -\tilde{\mathcal{A}}_k a_k - \mu A_k' + \tilde{\Sigma} \dot{W}_{A'}, \tag{B2b}$$

and

$$\mathcal{T}_k' \simeq -\tilde{\mathcal{A}}_k' a_k - \mu' A_k + \tilde{\Sigma}' \dot{W}_A, \tag{B3a}$$

$$\tilde{\mathcal{T}}_k' \simeq -\mathcal{A}_k' a_k - \nu_R' A_k' + \Sigma' \dot{W}_{A'}, \tag{B3b}$$

which should be interpreted in the sense of the correlation agreement as in (A5). The argument in support of our demonstration that Eqs. (B2) and (B3) are reasonable is motivated from the modeling

$$N^{\text{eff}} = -i \sum_{k_{123}, k_{12} \neq k} W_{3k}^{12} \delta_{3k}^{12} a_1 a_2 a_3^* \simeq -\bar{\gamma} a_k + \bar{\sigma} \dot{W}_a. \tag{B4}$$

Note that MSM (31) can be obtained from applying (B4) to (12), and becomes an accurate model when the cubic nonlinearity in (B4) is weak. Here we refer to a set of nonlinear interactions as weak if it can be well approximated by the white-noise forcing and linear dissipation due to a sharp timescale separation. As discussed in Sec. III C, the effective nonlinearity is not always weak because ARM (30) constructed from using

$$N^{\text{eff}} \simeq -\gamma a_k - \int_0^t ds\, \mathcal{A} e^{-(i\Omega_k + \nu)(t-s)} a_k(s) + \sigma \dot{W}_a + e^{-i\Omega_k t} R_\nu$$

is in some cases a good approximate model. Nonetheless, our belief is that $\mathcal{T}_k$, $\tilde{\mathcal{T}}_k$ and $\mathcal{T}_k'$, $\tilde{\mathcal{T}}_k'$ are weak because they are of higher order and the magnitudes are significantly smaller than N^{eff}. Paying attention to the similarity between the cubic nonlinearity terms visible in (A4) and (B4), we are encouraged to perform the formal approximations in (B2) and (B3). In doing so, we take into account the style for A_k (or A_k') to be involved in $\mathcal{T}_k$, $\mathcal{T}_k'$ (or $\tilde{\mathcal{T}}_k$, $\tilde{\mathcal{T}}_k'$) and make the critical difference from (B4) in the aspect that the damping part of (B2) and (B3) is replaced by the linear combination of a_k and A_k (or A_k'). Explaining below the reason why the coefficients of a_k in (B2a) and (B3b) are given by $-\mathcal{A}_k$ and $-\mathcal{A}_k'$, and why the terms in (B2b) and (B3a) can be ignored, the approximations in (B2) and (B3) agree with the ones obtained from substituting (17) and (B1) into (16).

Second, we discuss the off-diagonal elements of (B1). Let us consider the evolution of the autocorrelation matrix, obtained from substituting (17) and (B1) into (5). After comparing the resulting equation with

$$\langle \dot{A}_k(t), a_k(0) \rangle \doteq -i\tilde{\Omega}_k \langle A_k(t), a_k(0) \rangle - \mathcal{A}_k \langle a_k(t), a_k(0) \rangle - \nu_R \langle A_k(t), a_k(0) \rangle, \tag{B5}$$

we are allowed to ignore the contribution of μ. Equation (B5) results from the approximations

$$\langle \tau_k(t), a_k(0) \rangle \doteq -i\tilde{\Omega}_k \langle A_k(t), a_k(0) \rangle, \tag{B6a}$$

$$\langle \mathcal{T}_k(t), a_k(0) \rangle \doteq -\mathcal{A}_k \langle a_k(t), a_k(0) \rangle - \nu_R \langle A_k(t), a_k(0) \rangle, \tag{B6b}$$

$$\langle \tilde{\mathcal{T}}_k(t), a_k(0) \rangle \doteq -\tilde{\mathcal{A}}_k \langle a_k(t), a_k(0) \rangle - \mu \langle A_k'(t), a_k(0) \rangle, \tag{B6c}$$

calculated as follows.

(1) Equation (B6a) is obtained from (A5) and this two-point function is nonvanishing unless $t = 0$.

(2) For the first term of the right-hand side of (B6b), we can perform a similar calculation with (C3), provided in Appendix C, and show $\langle \mathcal{T}_k(t), a_k(0) \rangle \doteq -\mathcal{A}_k \langle a_k(t), a_k(0) \rangle$ where $\mathcal{A}_k = \sum_< 2W_{3k}^{12} \delta_{3k}^{12} (W_{21}^{k3} n_2 n_3 + W_{12}^{3k} n_1 n_3 - W_{k3}^{21} n_1 n_2)$. One can see that this coefficient is a good approximation of N_k/n_k from (C4a), and that $\langle \mathcal{T}_k, a_k \rangle \doteq -N_k = -\mathcal{A}_k \langle a_k, a_k \rangle$ holds from (A3a). Hence the coefficient of a_k in (B2a) should be $-\mathcal{A}_k$. For the second term we need to seek the expression for ν_R by studying the higher order approximation than the RPA level.

(3) For the first term of the right-hand side of (B6c), we apply the RPA and perform a similar calculation with (A2) to obtain $\langle \tilde{\mathcal{T}}_k(t), a_k(0) \rangle \doteq 0$ and $\tilde{\mathcal{A}}_k \doteq 0$. This implies that $\tilde{\mathcal{A}}_k$ is significantly smaller than $\mathcal{A}_k$ and can be ignored. The symmetry in (B2) gives a strong indication that μ and $\tilde{\Sigma}$ are also significantly smaller than ν_R and Σ, respectively. Hence we believe that $\langle \tilde{\mathcal{T}}_k(t), a_k(0) \rangle \doteq 0$ for all time and the contribution by $\tilde{\mathcal{T}}_k$ is minor.

For the case of μ', considering $\langle \dot{A}_k'(t), a_k(0) \rangle$ as in (B5), we demonstrate that $\langle \mathcal{T}_k'(t), a_k(0) \rangle$ is minor and can be ignored compared with $\langle \tilde{\mathcal{T}}_k'(t), a_k(0) \rangle$ and that the coefficient $\mathcal{A}_k'$ in (B3b) equals to the one in (C4b).

APPENDIX C: EQUILIBRIUM SPECTRUM OF A_k

Here we derive the approximations

$$N_k \doteq \sum_< 2W_{3k}^{12} \delta_{3k}^{12} \big(W_{21}^{k3} n_2 n_3 + W_{12}^{3k} n_1 n_3 - W_{k3}^{21} n_1 n_2 \big) n_k \tag{C1}$$

and

$$N_k \doteq \sum_< 2 \big| W_{3k}^{12} \big|^2 \delta_{3k}^{12} n_1 n_2 n_3. \tag{C2}$$

It is demanded that two quantities (C1) and (C2) are close to one another, allowing for A_k and A'_k in (13) to be determined in an unambiguous fashion.

As for (C1), we use (A3a) and (A4) to obtain

$$\begin{aligned} N_k &= \langle A_k, A_k \rangle \doteq -\langle \dot{A}_k, a_k \rangle \\ &\doteq i \left\langle \sum_{<} W_{3k}^{12} \delta_{3k}^{12} (A_1 a_2 a_3^* + a_1 A_2 a_3^* + a_1 a_2 A_3^*), a_k \right\rangle \\ &\doteq \sum_{<} 2W_{3k}^{12} \delta_{3k}^{12} \big(W_{21}^{k3} n_2 n_3 + W_{12}^{3k} n_1 n_3 - W_{k3}^{21} n_1 n_2 \big) n_k, \end{aligned} \tag{C3}$$

where we have used

$$\begin{aligned} &\sum_{<} W_{3k}^{12} \delta_{3k}^{12} \langle A_1 a_2 a_3^* a_k^* \rangle \\ &= -i \sum_{<} W_{3k}^{12} \delta_{3k}^{12} \sum_{|\Omega_{3'4'}^{12'}| < \varepsilon} W_{2'1}^{4'3'} \delta_{3'4'}^{12'} \langle a_{3'} a_{4'} a_{2'}^* a_2 a_3^* a_k^* \rangle \\ &\doteq -i \sum_{<} 2W_{3k}^{12} W_{21}^{k3} \delta_{3k}^{12} n_2 n_3 n_k, \end{aligned}$$

and so on. A similar result holds for N'_k, and as the byproduct, the approximations

$$\mathcal{A}_k \equiv \frac{N_k}{n_k} \doteq \sum_{<} 2W_{3k}^{12} \delta_{3k}^{12} \big(W_{21}^{k3} n_2 n_3 + W_{12}^{3k} n_1 n_3 - W_{k3}^{21} n_1 n_2 \big), \tag{C4a}$$

$$\mathcal{A}'_k \equiv \frac{N'_k}{n_k} \doteq \sum_{>} 2W_{3k}^{12} \delta_{3k}^{12} \big(W_{21}^{k3} n_2 n_3 + W_{12}^{3k} n_1 n_3 - W_{k3}^{21} n_1 n_2 \big) \tag{C4b}$$

are obtained.

As for (C2), we obtain

$$\mathcal{A}_k \equiv \frac{N_k}{n_k} \doteq \frac{\sum_{<} 2 \big| W_{3k}^{12} \big|^2 \delta_{3k}^{12} n_1 n_2 n_3}{n_k} \tag{C5}$$

from directly applying the RPA (A1) to $N_k = \langle A_k, A_k \rangle$. We here remark that there is no analog of (C5) for $\mathcal{A}'_k$. We also remark that, from equating (C4a) and (C5), the (near) resonance quartets satisfying (15) are essentially constrained by

$$\frac{W_{3k}^{12}}{n_k} + \frac{W_{k3}^{21}}{n_3} - \frac{W_{12}^{3k}}{n_2} - \frac{W_{21}^{k3}}{n_1} = 0. \tag{C6}$$

Equation (C6) is the condition characterizing the thermal equilibrium of the wave turbulence system (see [22] and references therein).

[1] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications* (Springer, New York, 2010).

[2] C. Gardiner, *Handbook of Stochastic Methods* (Springer, New York, 1985), Vol. 3.

[3] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics: Nonequilibrium Statistical Mechanics* (Springer, New York, 1992), Vol. 2.

[4] V. E. Zakharov, V. S. L'vov, and G. Falkovich, *Kolmogorov Spectra of Turbulence I: Wave Turbulence* (Springer Science & Business Media, New York, 2012).

[5] S. Nazarenko, *Wave Turbulence* (Springer Science & Business Media, New York, 2011), Vol. 825.

[6] H. Mori, Prog. Theor. Phys. **33**, 423 (1965).

[7] R. Zwanzig, J. Stat. Phys. **9**, 215 (1973).

[8] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, Oxford, 2001).

[9] R. Kubo, Rep. Prog. Phys. **29**, 255 (1966).

[10] S. Lepri, Phys. Rev. E **58**, 7165 (1998).

[11] W. Lee, G. Kovacic, and D. Cai, Phys. Rev. Lett. **103**, 024502 (2009).

[12] W. Lee, G. Kovačič, and D. Cai, Proc. Natl. Acad. Sci. USA **110**, 3237 (2013).

[13] B. Gershgorin, Y. V. Lvov, and D. Cai, Phys. Rev. Lett. **95**, 264302 (2005).

[14] B. Gershgorin, Y. V. Lvov, and D. Cai, Phys. Rev. E **75**, 046603 (2007).

[15] Shi-xiao W. Jiang, H. H. Lu, D. Zhou, and D. Cai, Phys. Rev. E **90**, 032925 (2014).

[16] S. W. Jiang, H. Lu, D. Zhou, and D. Cai, New J. Phys. **18**, 083028 (2016).

[17] W. Lee, Phys. Rev. E **98**, 022137 (2018).

[18] W. Lee, J. Nonlinear Sci. **29**, 1865 (2019).

[19] A. J. Majda, D. McLaughlin, and E. Tabak, J. Nonlinear Sci. **7**, 9 (1997).

[20] G. Whitham, *Linear and Nonlinear Waves* (Wiley-Interscience, New York, 2011), Vol. 42.

[21] V. Yakhot and V. Zakharov, Physica D **64**, 379 (1993).

[22] W. Lee, G. Kovačič, and D. Cai, Phys. Rev. E **97**, 062140 (2018).