DISCRETE APPROXIMATIONS WITH ADDITIONAL CONSERVED QUANTITIES: DETERMINISTIC AND STATISTICAL BEHAVIOR [*]

RAFAIL V. ABRAMOV[†] AND ANDREW J. MAJDA[† ‡]

Abstract. Discrete numerical approximations with additional conserved quantities are developed here both for barotropic geophysical flows generalizing the 2D incompressible fluid equations and suitable discretizations of the Burgers-Hopf equation. Mathematical, numerical, and statistical properties of these approximations are studied below in various different contexts through the symbiotic interaction of mathematical theory and scientific computing. The new contributions include an explicit concrete discussion of the sine-bracket spectral truncation with many conserved quantities for 2D incompressible flow, a theoretical and numerical comparison with the standard spectral truncation, and a rigorous proof of convergence to suitable weak solutions in the limit as the number of Fourier modes increases. Systematic discretizations of the Burgers-Hopf equation are developed, which conserve linear momentum and a non-linear energy; careful numerical experiments regarding the statistical behavior of these models indicate that they are ergodic and strongly mixing but do not have equipartition of energy in Fourier space. Furthermore, the probability distribution of the values at a single grid point can be highly non-Gaussian with two strong isolated peaks in this distribution. This contrasts with earlier results for statistical behavior of difference schemes which conserve a quadratic energy. The issues of statistically relevant conserved quantities are introduced through a new case study for the Galerkin-truncated Burgers-Hopf model.

1. Introduction. Stan Osher has made numerous world class contributions to numerical analysis and scientific computing including the numerical analysis of discrete initial boundary value problems, systematic new difference schemes for compressible fluid dynamics, level set methods, and novel applications in image processing, computer vision, etc. Stan and the second author had an intense scientific collaboration in the mid 1970's with research topics spanning hyperbolic mixed problems, numerical analysis of linear hyperbolic problems with rough initial data [13, 14], and the systematic design of difference schemes for nonlinear conservation laws that avoid nonlinear instability and satisfy a discrete form of the entropy condition [15, 16]. These last works were one byproduct of a seminar at UCLA in 1976-1977 as well as a graduate course taught by the second author on the research topic of nonlinear hyperbolic conservation laws inspired by Peter Lax's research program [6, 9]. Jim Ralston, Bjorn Engquist, and Michael E. Taylor were all significant contributors to this seminar besides Osher and the second author. It is a pleasure to contribute a paper to a volume honoring Stan Osher on his 60th birthday discussing discrete approximations as models for important contemporary problems involving deterministic and statistical behavior for problems with additional conserved quantities. Even though the scientific emphasis here is completely different, the results below make some contact with the older joint work of Majda and Osher mentioned above and the style blends discrete numerical analysis and scientific computing in a manner which Osher both advocates and appreciates.

In many challenging nonlinear problems of interest in contemporary science such as climate modeling or protein folding, the questions of interest are inherently statistical and involve quantifying the potential for enhanced predictability at large scales, in a given problem with multiple scales of behavior and many degrees of freedom (see [8, 12, 17, 18], and the extensive references therein). The second author, in joint

---

work with Timofeyev and others, has been utilizing simple energy conserving discrete approximations to the Burgers-Hopf equation such as the Fourier-Galerkin truncation and the Kruskal-Zabusky finite difference scheme as elementary models with intrinsic stochastic behavior with many degrees of freedom where such statistical issues can be addressed in a concise scientific fashion with firm mathematical underpinnings as a guideline [2, 8, 17, 18]. In the model discrete approximations for the truncated Burgers-Hopf equation, there is at least one additional conserved quantity, an energy, besides the discrete linear momentum which is essential to allow energy transfer from the small scales back to the large scales and simultaneously yield ergodicity with strong mixing. Several new discrete models of this sort, motivated by Tadmor's significant systematic generalization [23] of ideas roughly originating in joint work of Osher and the second author [15, 16], are developed in Section 4 of this paper. New computational results on the statistical behavior of these models as well as a brief discussion of accessible new mathematical issues are presented in Section 4.

In another different but related topic, there has been an intense debate over the last twenty-five years regarding the relative importance of the formally infinite list of conserved quantities for geophysical flows for the coarse-grained statistical behavior at large scales [4, 7, 11, 12, 20, 22]. The simplest geophysical model of this sort is barotropic two-dimensional flow with topography in periodic geometry, which is described by the equations

$$q = \Delta\psi + h, \tag{1a}$$

$$\vec{v} = \nabla^{\perp}\psi = \begin{pmatrix} -\psi_y \\ \psi_x \end{pmatrix}. \tag{1b}$$

$$\frac{\partial q}{\partial t} + \vec{v} \cdot \nabla q = 0, \tag{1c}$$

In (1), q is the potential vorticity, $\omega = \Delta\psi$ is the relative vorticity, $\vec{v}$ is the incompressible fluid velocity, ψ is the stream function, and h is the prescribed topography. When h is zero, (1) become the equations for 2D incompressible flow. Non-zero topography often has profound impact on the large scale flow [4, 7, 12, 22]. Here and below we assume a 2π-periodic geometry in both the x and y variables which are also denoted by x_1, x_2 whenever convenient. The equations in (1) conserve kinetic energy,

$$E = \frac{1}{2}\int_{T^2}|\vec{v}|^2 = -\frac{1}{2}\int_{T^2}\psi\omega = -\frac{1}{2}\int_{T^2}\psi(q-h), \tag{2}$$

as well as the infinite number of conserved quantities,

$$Q_p(q) = \int_{T^2}q^p, \quad p = 1, 2, 3, \ldots. \tag{3}$$

In (2), (3), $\int_{T^2} \cdot$ denotes integration over the period domain in periodic geometry, where the total circulation, Q_1, satisfies $Q_1 \equiv 0$. The quadratic conserved quantity, the enstrophy

$$\mathcal{E} = \frac{1}{2}Q_2(q), \tag{4}$$

is singled out in some statistical theories for large scale flow [4, 7, 22] as having special significance, while the higher generalized enstrophies, $Q_p(q)$, for $p \geq 3$ are ignored in these theories. Other researchers [20] claim that the entire infinite list of conserved quantities in (3) is statistically significant for describing the coarse-grained features at large scales. These various statistical theories, their relative strength and weakness, new statistical approaches and potential applications [11] are discussed by Majda and Wang [12] in a recent monograph which contains many additional references and discussion. In particular, the recent contributions of Turkington [24] are especially significant. One way to address these issues is to study the statistical behavior of discrete approximations to (1) conserving the energy in (2), and in addition other discrete approximations to the generalized enstrophies in (3). This is the main topic in Sections 2 and 3 of this paper. In Section 2 we review the standard pseudo-spectral method for (1) and present a concrete self-contained introduction to the sine-bracket truncation, developed independently in an abstract form by Zeitlin [25] and in the appendix of [20]. The sine-bracket truncation is a discrete spectral approximation of (1) with many additional discrete conserved quantities, mimicking (3), in addition to energy. Algorithms for time discretization and simple numerical experiments contrasting the basic properties of these schemes are also developed in Section 2. In Section 3, it is proved that the numerical approximations discussed in Section 2 converge to a weak solution of (1) with finite enstrophy as the number of Fourier modes increases; this answers a question raised in [20] regarding the convergence of these methods. Finally, the authors [3] have utilized the numerical methods developed in Section 2 to show that a suitable discrete approximation to $Q_3(q)$ is a statistically relevant conserved quantity for the large scale behavior besides the energy and enstrophy in a suitable regime for the sine-bracket truncation with over 500 degrees of freedom. Once again the pseudo-spectral approximation to the Burgers-Hopf equation provides an important simplified model for the issues of statistically relevant conserved quantities; see the recent paper of the authors [2] for a detailed discussion as well as a collection of accessible mathematical issues. As an introduction to the issues of statistically relevant and irrelevant conserved quantities, in Section 5 we present new simulations in regimes for the Galerkin-truncated Burgers-Hopf model, where the Hamiltonian is statistically relevant, besides the energy, and show that the failure of a linear tilt in partition of energy coincides with the failure of strong mixing at large scales.

2. Spectral Truncations of Quasi-Geostrophic Flow with Additional Conserved Quantities. Here we develop the explicit mathematical properties of two different finite-dimensional truncations for the quasi-geostrophic equations in (1) in Fourier space with both conservation of energy as in (2) and conservation of quadratic enstrophy in (4). For the traditional spectral truncation these are typically the only conserved quantities while for the sine-bracket spectral truncation [20, 25] there is a much larger family of suitable discrete approximations to the higher order enstrophies in (3) that is conserved. The properties of the sine-bracket truncation are developed explicitly below. Second order accurate symplectic time integrators are developed following McLachlan's method [19] and Strang splitting. Simple numerical experiments are presented below to demonstrate the explicit conservation properties of the two numerical methods and various discrete time integrators.

2.1. Some Basic Facts about Hamiltonian Systems. We consider the equations for a finite-dimensional Hamiltonian system

$$\frac{d\mathbf{y}}{dt} = \mathbf{J}(\mathbf{y}) \nabla H(\mathbf{y}), \tag{5}$$

where $\mathbf{y}$ is the column vector of dynamical variables, H is a *Hamiltonian* function and $\mathbf{J}$ is a *symplectic matrix* which is skew-symmetric and satisfies the Jacobi identity

$$\sum_l (J_{pl}\partial_l J_{qr} + J_{rl}\partial_l J_{pq} + J_{ql}\partial_l J_{rp}) = 0 \quad \forall p, q, r \leq N, \tag{6}$$

where N is the dimension of the system. The Jacobi identity essentially tells us that the system can be represented in its canonical action-angle variables (see [21] and references therein). We will be proving the Jacobi identity here for a few cases below; however, as the reader will see, in this work we never really take any advantage of the Jacobi identity in the numerical methods.

We denote the *Poisson bracket* for $F(\mathbf{y})$ and $G(\mathbf{y})$ by

$$\{F, G\} = (\nabla F, \mathbf{J}\nabla G), \tag{7}$$

where $(\cdot, \cdot)$ denotes the scalar product of two vectors. Due to the skew-symmetry of $\mathbf{J}$ we have

$$\{F, G\} = -\{G, F\}. \tag{8}$$

In this notation the time evolution for any $F(\mathbf{y})$ can be written in the form

$$\frac{\mathrm{d}F}{\mathrm{d}t} = (\nabla F, \dot{\mathbf{y}}) = (\nabla F, \mathbf{J}\nabla H) = \{F, H\}, \tag{9}$$

thus from (8) and (9) we have

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \{H, H\} = -\{H, H\} = 0, \tag{10}$$

i.e. H is the *conserved quantity* of the system.

Generally, the conserved quantities of the system (5) depend on both $\mathbf{J}$ and H. Let us take the basis of the kernel $\ker(\mathbf{J})$ to be spanned by $\mathbf{K}_j(\mathbf{y})$, $0 < j \leq \dim(\ker(\mathbf{J}))$. If for some $\mathbf{K}_j(\mathbf{y})$ (let us denote it as $\mathbf{K}_{j_k}(\mathbf{y})$) we can find $C_k(\mathbf{y})$ such that

$$\mathbf{K}_{j_k}(\mathbf{y}) = \nabla C_k(\mathbf{y}), \tag{11}$$

then

$$\mathbf{J}\nabla C_k(\mathbf{y}) \equiv 0.$$

This condition automatically guarantees

$$\frac{\mathrm{d}C_k}{\mathrm{d}t} = \{C_k, H\} = -\{H, C_k\} = (\nabla H, \mathbf{J}\nabla C_k) = 0,$$

i.e. C_k is conserved in time regardless of the Hamiltonian and, therefore is a property of the symplectic structure $\mathbf{J}$ alone. These conserved quantities are called the *Casimir invariants*. For simple Poisson brackets the Casimir invariants are easy to find, and if we are looking for the conserved quantities of a system, then the Casimir invariants are the best candidates to be found quickly.

2.2. The Equations for Barotropic Flow in Fourier space. The two-dimensional vorticity equation in (1) on a double-periodic plane transformed through two-dimensional Fourier series takes the form of an infinite system of ODE's:

$$\frac{\mathrm{d}\hat{q}_{\mathbf{k}}}{\mathrm{d}t} = \sum_{k_1', k_2' = -\infty}^{\infty} \frac{\mathbf{k} \times \mathbf{k}'}{|\mathbf{k}'|^2} \hat{q}_{\mathbf{k}+\mathbf{k}'}(\hat{q}_{-\mathbf{k}'} - \hat{h}_{-\mathbf{k}'}), \tag{12}$$

Here $\hat{q}_{\mathbf{k}}$ denotes the spectral coefficient associated with the two-dimensional wave vector $\mathbf{k}$, whose components are integers. The skew-symmetric scalar product $\mathbf{k} \times \mathbf{k}'$ is $k_1 k_2' - k_2 k_1'$, and the norm $|\mathbf{k}|$ is $\sqrt{k_1^2 + k_2^2}$. Since q is real, then $\hat{q}_{\mathbf{k}}^* = \hat{q}_{-\mathbf{k}}$. In order to derive (12), first we look at the equation in (1a) and conclude that in Fourier space it becomes

$$\hat{q}_{\mathbf{k}} = -|\mathbf{k}|^2 \hat{\psi}_{\mathbf{k}} + \hat{h}_{\mathbf{k}}, \tag{13}$$

so that

$$\hat{\psi}_{\mathbf{k}} = -\frac{1}{|\mathbf{k}|^2}(\hat{q}_{\mathbf{k}} - \hat{h}_{\mathbf{k}}). \tag{14}$$

Now, let us write ψ_x, ψ_y, q_x, q_y in terms of Fourier coefficients:

$$\psi_x = \sum_{\mathbf{k}} (ik_1)\hat{\psi}_{\mathbf{k}} e^{i(k_1 x + k_2 y)},$$

$$\psi_y = \sum_{\mathbf{k}} (ik_2)\hat{\psi}_{\mathbf{k}} e^{i(k_1 x + k_2 y)},$$

$$q_x = \sum_{\mathbf{k}} (ik_1)\hat{q}_{\mathbf{k}} e^{i(k_1 x + k_2 y)},$$

$$q_y = \sum_{\mathbf{k}} (ik_2)\hat{q}_{\mathbf{k}} e^{i(k_1 x + k_2 y)}.$$

Then we write the dynamic equation (1c) in terms of Fourier coefficients,

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{\mathbf{k}} \hat{q}_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}} = \sum_{\mathbf{k}'} \sum_{\mathbf{k}''} (k_1' k_2'' - k_2' k_1'') \hat{\psi}_{\mathbf{k}'} \hat{q}_{\mathbf{k}''} e^{i(\mathbf{k}'+\mathbf{k}'')\mathbf{x}}.$$

Now we substitute $\mathbf{k} = \mathbf{k}' + \mathbf{k}''$ and collect terms with equal exponents:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{q}_{\mathbf{k}} = -\sum_{\mathbf{k}'} (k_2' k_1 - k_1' k_2) \hat{\psi}_{\mathbf{k}'} \hat{q}_{\mathbf{k}-\mathbf{k}'}.$$

Changing $\mathbf{k}' \to -\mathbf{k}'$ yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{q}_{\mathbf{k}} = -\sum_{\mathbf{k}'} (\mathbf{k} \times \mathbf{k}') \hat{q}_{\mathbf{k}+\mathbf{k}'} \hat{\psi}_{-\mathbf{k}'},$$

and, substituting the expression in (14) for $\hat{\psi}_{\mathbf{k}}$, we obtain (12).

It is well-known that the quasi-geostrophic equations in (1) are a Hamiltonian system in physical space with the Hamiltonian given by the energy in (2), while the higher order enstrophies in (3) are Casimir invariants [21, 22]. Here we record the corresponding formal Hamiltonian structure in Fourier space in order to motivate the

finite-dimensional spectral approximations developed below. The infinite-dimensional equation (12) has the Hamiltonian structure

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{q}_{\mathbf{k}} = \mathbf{J}_{\mathbf{kk'}}\frac{\partial H}{\partial \hat{q}_{\mathbf{k}}} \tag{15}$$

with

$$\mathbf{J}_{\mathbf{kk'}} = -(\mathbf{k}\times\mathbf{k'})\hat{q}_{\mathbf{k+k'}}. \tag{16}$$

The Hamiltonian H is given by the energy since

$$H = E = \frac{1}{2}\sum_{\mathbf{k}}|\mathbf{k}|^2|\hat{\psi}_{\mathbf{k}}|^2 = \frac{1}{2}\sum_{\mathbf{k}}\frac{|\hat{q}_{\mathbf{k}}-\hat{h}_{\mathbf{k}}|^2}{|\mathbf{k}|^2}, \tag{17a}$$

$$\frac{\partial H}{\partial \hat{q}_{\mathbf{k}}} = \frac{1}{|\mathbf{k}|^2}(\hat{q}_{-\mathbf{k}}-\hat{h}_{-\mathbf{k}}), \tag{17b}$$

The fact that $\mathbf{J}_{\mathbf{kk'}}$ is skew-symmetric follows from $\mathbf{J}_{\mathbf{kk'}} = -\mathbf{J}_{\mathbf{k'k}}$. The formal proof of the Jacobi identity follows directly from the definition in (6).

$$
\begin{aligned}
\sum_{\mathbf{l}}(\mathbf{J}_{\mathbf{il}}\partial_{\mathbf{l}}\mathbf{J}_{\mathbf{jk}} &+ \mathbf{J}_{\mathbf{kl}}\partial_{\mathbf{l}}\mathbf{J}_{\mathbf{ij}} + \mathbf{J}_{\mathbf{jl}}\partial_{\mathbf{l}}\mathbf{J}_{\mathbf{ki}}) = \\
&= \sum_{\mathbf{l=j+k}}(\mathbf{i}\times\mathbf{l})(\mathbf{j}\times\mathbf{k})\hat{q}_{\mathbf{i+j+k}}+ \\
&\quad + \sum_{\mathbf{l=i+j}}(\mathbf{k}\times\mathbf{l})(\mathbf{i}\times\mathbf{j})\hat{q}_{\mathbf{i+j+k}}+ \\
&\quad + \sum_{\mathbf{l=k+i}}(\mathbf{j}\times\mathbf{l})(\mathbf{k}\times\mathbf{i})\hat{q}_{\mathbf{i+j+k}} = \\
&= \sum_{\mathbf{l=j+k}}(\mathbf{i}\times(\mathbf{j}+\mathbf{k}))(\mathbf{j}\times\mathbf{k})\hat{q}_{\mathbf{i+j+k}}+ \\
&\quad + \sum_{\mathbf{l=i+j}}(\mathbf{k}\times(\mathbf{i}+\mathbf{j}))(\mathbf{i}\times\mathbf{j})\hat{q}_{\mathbf{i+j+k}}+ \\
&\quad + \sum_{\mathbf{l=k+i}}(\mathbf{j}\times(\mathbf{k}+\mathbf{i}))(\mathbf{k}\times\mathbf{i})\hat{q}_{\mathbf{i+j+k}} = \\
&= \sum_{\mathbf{l=j+k}}(\mathbf{i}\times\mathbf{j})(\mathbf{j}\times\mathbf{k})-(\mathbf{k}\times\mathbf{i})(\mathbf{j}\times\mathbf{k})\hat{q}_{\mathbf{i+j+k}}+ \\
&\quad + \sum_{\mathbf{l=i+j}}(\mathbf{k}\times\mathbf{i})(\mathbf{i}\times\mathbf{j})-(\mathbf{j}\times\mathbf{k})(\mathbf{i}\times\mathbf{j})\hat{q}_{\mathbf{i+j+k}}+ \\
&\quad + \sum_{\mathbf{l=k+i}}(\mathbf{j}\times\mathbf{k})(\mathbf{k}\times\mathbf{i})-(\mathbf{i}\times\mathbf{j})(\mathbf{k}\times\mathbf{i})\hat{q}_{\mathbf{i+j+k}} = 0.
\end{aligned}
\tag{18}
$$

The system (12) also has Casimir invariants of the form

$$\hat{C}_N = \sum_{\sum_{j=1}^N \mathbf{k}_j = 0}\left(\prod_{j=1}^N \hat{q}_{\mathbf{k}_j}\right), \quad \text{arbitrary } N \tag{19}$$

which are the Fourier space analogues of (3). The proof of conservation is also a direct calculation:

$$
\begin{aligned}
\sum_{\mathbf{j}} \mathbf{J}_{\mathbf{jk}} \frac{\partial C_N}{\partial \hat{q}_{\mathbf{j}}} &= \sum_{\mathbf{j}} \mathbf{J}_{\mathbf{jk}} \left(N \sum_{\sum_l^{N-1} \mathbf{m}_l = -\mathbf{j}} \prod_l^{N-1} \hat{q}_{\mathbf{m}_l} \right) = \\
&= N \sum_{\mathbf{j}} \sum_{\sum_l^{N-1} \mathbf{m}_l = -\mathbf{j}} (\mathbf{j} \times \mathbf{k}) \hat{q}_{\mathbf{j}+\mathbf{k}} \prod_l^{N-1} \hat{q}_{\mathbf{m}_l} = \\
&= N \sum_{\mathbf{j}} \sum_{\sum_l^{N-1} \mathbf{m}_l = -\mathbf{j}} ((\mathbf{j} + \mathbf{k}) \times \mathbf{k}) \hat{q}_{\mathbf{j}+\mathbf{k}} \prod_l^{N-1} \hat{q}_{\mathbf{m}_l} = \\
&= N \sum_{\mathbf{i}} \sum_{\sum_l^{N-1} \mathbf{m}_l = \mathbf{k}-\mathbf{i}} (\mathbf{i} \times \mathbf{k}) \hat{q}_{\mathbf{i}} \prod_l^{N-1} \hat{q}_{\mathbf{m}_l} = \\
&= N! \sum_{\sum_l^{N} \mathbf{m}_l = \mathbf{0}} (\mathbf{k} \times \mathbf{k}) \prod_l^{N} \hat{q}_{\mathbf{m}_l} = \mathbf{0}
\end{aligned}
$$

since $\mathbf{k} \times \mathbf{k} \equiv 0$. The conservation of potential vorticity (which is the first Casimir invariant) now becomes trivial since

$$
\frac{\mathrm{d}Q_1(q)}{\mathrm{d}t} = \frac{\mathrm{d}\hat{q}_{\mathbf{0}}}{\mathrm{d}t} = 0.
$$

The enstrophy is essentially the second Casimir invariant,

$$
\mathcal{E} = \frac{1}{2} \sum_{\mathbf{k}} |\hat{q}_{\mathbf{k}}|^2.
$$

2.3. The Traditional spectral truncation of the equations for barotropic flow. We truncate the quasi-geostrophic equations in (1) by projecting them onto the finite $(2M+1) \times (2M+1)$ set of Fourier modes. With help of the projection operator

$$
P_\varepsilon f(\vec{x}) = \sum_{|k_1|,|k_2| \leq M} \hat{f}_{\mathbf{k}} e^{i\vec{x}\cdot\mathbf{k}}, \quad \varepsilon = \frac{2\pi}{2M+1}, \tag{20}
$$

where we adopt the notation

$$
P_\varepsilon f \equiv f^\varepsilon,
$$

we write the standard Galerkin projection of the quasi-geostrophic equations in (1) as

$$
\frac{\partial q^\varepsilon}{\partial t} = -P_\varepsilon \left[\mathrm{div}\left(\vec{v}^\varepsilon q^\varepsilon\right) \right], \tag{21a}
$$

$$
q^\varepsilon = \Delta \psi^\varepsilon + h^\varepsilon, \tag{21b}
$$

$$
\vec{v}^\varepsilon = \nabla^\perp \psi^\varepsilon. \tag{21c}
$$

The following three integrals of motion survive the truncation: the projected spatial integral of the potential vorticity Q_1^ε, the projected energy E^ε, and the projected enstrophy $\mathcal{E}^\varepsilon$. In the framework of (20) and (21) the proofs of conservation are trivial: taking into account the fact that the standard inner product $(g^\varepsilon, f) = (g, f^\varepsilon)$ for all g, f, we write

- Vorticity

$$\frac{\partial Q_1^\varepsilon}{\partial t} = \int_{T^2} \frac{\partial q^\varepsilon}{\partial t} = - \int_{T^2} \operatorname{div} P_\varepsilon(\vec{v}^\varepsilon q^\varepsilon) = 0, \tag{22}$$

- Energy

$$\frac{\partial E^\varepsilon}{\partial t} = - \int_{T^2} \psi^\varepsilon \frac{\partial q^\varepsilon}{\partial t} = \int_{T^2} \psi^\varepsilon P_\varepsilon \left[\operatorname{div}(\vec{v}^\varepsilon q^\varepsilon) \right] =$$
$$= \int_{T^2} \psi^\varepsilon \operatorname{div}(\vec{v}^\varepsilon q^\varepsilon) = \int_{T^2} \operatorname{div}(\psi^\varepsilon \vec{v}^\varepsilon q^\varepsilon) = 0, \tag{23}$$

- Enstrophy

$$\frac{\partial \mathcal{E}^\varepsilon}{\partial t} = \int_{T^2} q^\varepsilon \frac{\partial q^\varepsilon}{\partial t} = - \int_{T^2} q^\varepsilon P_\varepsilon \left[\operatorname{div}(\vec{v}^\varepsilon q^\varepsilon) \right] =$$
$$= - \int_{T^2} q^\varepsilon \operatorname{div}(\vec{v}^\varepsilon q^\varepsilon) = -\frac{1}{2} \int_{T^2} \operatorname{div}(q^\varepsilon \vec{v}^\varepsilon q^\varepsilon) = 0. \tag{24}$$

We note that the Galerkin truncation in (21) of the quasi-geostrophic equations in (1) can be interpreted as the Fourier space equations in (12) limited to the finite $(2M + 1) \times (2M + 1)$ domain of the Fourier coefficients:

$$\frac{d\hat{q}_k}{dt} = \sum_{k_1', k_2' = -M}^{M} \frac{\mathbf{k} \times \mathbf{k}'}{|\mathbf{k}'|^2} \hat{q}_{\mathbf{k}+\mathbf{k}'} (\hat{q}_{-\mathbf{k}'} - \hat{h}_{-\mathbf{k}'}), \tag{25}$$

provided that we extend the finite domain of the Fourier coefficients up to $(4M+1) \times (4M + 1)$ by using the periodicity rule

$$\hat{q}_{\mathbf{k}} = \hat{q}_{(\mathbf{k}+2M+1)}. \tag{26}$$

In this fashion, the equations in (21) can be reinterpreted as a finite truncation of (12). However, the traditional spectral truncation (25) does not retain the Hamiltonian structure of the original Hamiltonian system from 2.2, because the Jacobi identity fails for straightforward truncation of (16).

2.4. The Sine-Bracket Truncation with Many Additional Conserved Quantities. Here we consider the sine-bracket truncation as an approximation to the quasi-geostrophic dynamics in (1) through the spectral representation in Section 2.2. This finite dimensional set of equations for the Fourier coefficients is given by

$$\frac{d}{dt}\hat{q}_{\mathbf{k}} = \sum_{k_1', k_2' = -M}^{M} \frac{1}{\varepsilon} \frac{\sin(\varepsilon \mathbf{k} \times \mathbf{k}')}{|\mathbf{k}'|^2} \hat{q}_{\mathbf{k}+\mathbf{k}'} (\hat{q}_{-\mathbf{k}'} - \hat{h}_{-\mathbf{k}'}), \quad \varepsilon = \frac{2\pi}{2M + 1}, \tag{27}$$

so that the summation occurs on the $(2M + 1) \times (2M + 1)$ domain of the Fourier coefficients, and, as in Section 2.3, the coefficient $\hat{q}_{\mathbf{k}}$ have period $2M+1$ in $\mathbf{k}$. The name

of the truncation follows, first, from the presence of a sine function, and, second, due to the fact that the truncation can be written in its "bracket" form as a commutator defined on the group of unitary $(2M + 1) \times (2M + 1)$ matrices, which we will show later. Note that for $M \to \infty$ and given $\mathbf{k}$ and $\mathbf{k}'$, $\varepsilon^{-1} \sin(\varepsilon \mathbf{k} \times \mathbf{k}') = \mathbf{k} \times \mathbf{k}' + O(\varepsilon^2)$, so that the equations in (27) are formally consistent with the spectral form of the quasi-geostrophic equations discussed in 2.2. This truncation possesses a Hamiltonian structure with the symplectic operator

$$\mathbf{J_{kk'}} = -\frac{1}{\varepsilon} \sin(\varepsilon \mathbf{k} \times \mathbf{k}') \hat{q}_{\mathbf{k+k'}} \tag{28}$$

and Hamiltonian

$$H = E = \frac{1}{2} \sum_{\mathbf{k}} |\mathbf{k}|^2 |\hat{\psi}_{\mathbf{k}}|^2 = \frac{1}{2} \sum_{\mathbf{k}} \frac{|\hat{q}_{\mathbf{k}} - \hat{h}_{\mathbf{k}}|^2}{|\mathbf{k}|^2}, \tag{29a}$$

$$\frac{\partial H}{\partial \hat{q}_{\mathbf{k}}} = \frac{1}{|\mathbf{k}|^2} (\hat{q}_{-\mathbf{k}} - \hat{h}_{-\mathbf{k}}). \tag{29b}$$

There is no need in demonstrating the direct proof of the energy conservation in the sine-bracket truncation (27) like we did for the traditional truncation (25) in Section 2.3, since the energy *is* the Hamiltonian for (27), and the conservation of the Hamiltonian is the general property of a Hamiltonian system, which is shown in (10).

Due to the sine function in (28) the symplectic matrix is skew-symmetric and the Jacobi identity also holds: $\forall \mathbf{i}, \mathbf{j}, \mathbf{k}$ we have

$$\sum (\mathbf{J_{il}} \partial_{\mathbf{l}} \mathbf{J_{jk}} + \mathbf{J_{kl}} \partial_{\mathbf{l}} \mathbf{J_{ij}} + \mathbf{J_{jl}} \partial_{\mathbf{l}} \mathbf{J_{ki}}) =$$
$$= \sum [\sin(\varepsilon \mathbf{i} \times (\mathbf{j} + \mathbf{k})) \sin(\varepsilon \mathbf{j} \times \mathbf{k}) + \sin(\varepsilon \mathbf{k} \times (\mathbf{i} + \mathbf{j})) \sin(\varepsilon \mathbf{i} \times \mathbf{j}) +$$
$$+ \sin(\varepsilon \mathbf{j} \times (\mathbf{k} + \mathbf{i})) \sin(\varepsilon \mathbf{k} \times \mathbf{i})] \hat{q}_{\mathbf{i+j+k}} = [\sin(\varepsilon \mathbf{i} \times \mathbf{j}) \cos(\varepsilon \mathbf{i} \times \mathbf{k}) \sin(\varepsilon \mathbf{j} \times \mathbf{k}) -$$
$$- \cos(\varepsilon \mathbf{i} \times \mathbf{j}) \sin(\varepsilon \mathbf{k} \times \mathbf{i}) \sin(\varepsilon \mathbf{j} \times \mathbf{k}) + \sin(\varepsilon \mathbf{k} \times \mathbf{i}) \cos(\varepsilon \mathbf{k} \times \mathbf{j}) \sin(\varepsilon \mathbf{i} \times \mathbf{j}) -$$
$$- \cos(\varepsilon \mathbf{k} \times \mathbf{i}) \sin(\varepsilon \mathbf{j} \times \mathbf{k}) \sin(\varepsilon \mathbf{i} \times \mathbf{j}) + \sin(\varepsilon \mathbf{j} \times \mathbf{k}) \cos(\varepsilon \mathbf{j} \times \mathbf{i}) \sin(\varepsilon \mathbf{k} \times \mathbf{i}) -$$
$$- \cos(\varepsilon \mathbf{j} \times \mathbf{k}) \sin(\varepsilon \mathbf{i} \times \mathbf{j}) \sin(\varepsilon \mathbf{k} \times \mathbf{i})] \hat{q}_{\mathbf{i+j+k}} = 0.$$

As shown below, the sine-bracket truncation in (27) conserves $2M$ independent Casimir invariants of the form

$$C_N = \sum_{Z^N} \hat{q}_{\mathbf{i}_1} \dots \hat{q}_{\mathbf{i}_N} \cos[\varepsilon A(\mathbf{i}_1, \dots, \mathbf{i}_N)],$$
$$Z^N = \{(\mathbf{i}_1, \dots, \mathbf{i}_N), \sum_{j=1}^{N} \mathbf{i}_j = \mathbf{0}\}, \quad 1 \le N \le 2M, \tag{30}$$

where

$$A(\mathbf{i}_1, \dots, \mathbf{i}_N) = \mathbf{i}_2 \times \mathbf{i}_1 + \mathbf{i}_3 \times (\mathbf{i}_1 + \mathbf{i}_2) + \dots + \mathbf{i}_N \times (\mathbf{i}_1 + \dots + \mathbf{i}_{N-1}).$$

The Casimir invariant C_2 is a multiple of the enstrophy since $A(\mathbf{i}, -\mathbf{i}) = 0$. However, the higher Casimir invariants C_N for $3 \le N \le 2M$ are suitable regularizations of those in (12), or, equivalently, (3).

In order to prove the existence of the Casimir invariants in (30), let us consider the family of $(2M + 1) \times (2M + 1)$ matrices $\mathbf{D_n}$ which have the form

$$\mathbf{D_n} = \sigma^{n_1 n_2/2} g^{n_1} h^{n_2},$$

where the unitary $(2M + 1) \times (2M + 1)$ matrices h and g satisfy

$$hg = \sigma gh \quad \text{and} \quad g^{2M+1} = h^{2M+1} = I. \tag{31}$$

We choose $(2M + 1)$ to be prime and

$$\sigma = \exp\left(\frac{4\pi i}{2M + 1}\right).$$

For our choice of σ the matrices g and h are

$$g = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \sigma & 0 & \cdots & 0 \\ 0 & 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma^{2M} \end{pmatrix}, \quad h = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

We can see that if h is applied to the given matrix from the left, it shifts the given matrix upwards, and if it is applied from the right, it shifts the given matrix to the right. Taking this into account, (31) can be easily proven. First, $g^{2M+1} = I$ is obvious for our choice of σ. Second, $h^{2M+1} = I$ because h shifts itself $2M$ times into the identity position. Then,

$$hg = \begin{pmatrix} 0 & \sigma & 0 & \cdots & 0 \\ 0 & 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma^{2M} \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix} \quad \text{and} \quad \sigma gh = \begin{pmatrix} 0 & \sigma & 0 & \cdots & 0 \\ 0 & 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma^{2M} \\ \sigma^{2M+1} & 0 & 0 & \cdots & 0 \end{pmatrix}$$

and since $\sigma^{2M+1} = 1$ we obtain $hg = \sigma gh$.

The set of $\mathbf{D_n}$'s is the group of unitary matrices which satisfy the following relations:

$$\mathbf{D_n^\dagger} = \mathbf{D_{-n}} = (\mathbf{D_n})^{-1}, \tag{32a}$$

$$\mathbf{D_n D_m} = \exp(i\varepsilon(\mathbf{m} \times \mathbf{n}))\mathbf{D_{n+m}}, \tag{32b}$$

$$\frac{1}{2M + 1}\text{Tr}(\mathbf{D_m}) = \delta_{\mathbf{m}}^{\mathbf{0}}, \tag{32c}$$

where $\varepsilon = 2\pi/(2M + 1)$. We present the proof for the three statements in (32) below.

- Proof of (32a):

$$\mathbf{D_n^{-1}} = \exp\left(-\frac{2\pi i n_1 n_2}{2M+1}\right)(g^{n_1}h^{n_2})^{-1} = \exp\left(-\frac{2\pi i n_1 n_2}{2M+1}\right)(h^{n_2})^{-1}(g^{n_1})^{-1} =$$

$$= \exp\left(-\frac{2\pi i n_1 n_2}{2M+1}\right)(h^{n_2})^\dagger(g^{n_1})^\dagger = \exp\left(-\frac{2\pi i n_1 n_2}{2M+1}\right)(g^{n_1}h^{n_2})^\dagger = \mathbf{D_n^\dagger}.$$

$$\mathbf{D_n^{-1}} = \exp\left(-\frac{2\pi i n_1 n_2}{2M+1}\right)(g^{n_1}h^{n_2})^{-1} = \exp\left(\frac{2\pi i n_1 n_2}{2M+1}\right)\sigma^{-n_1 n_2}h^{-n_2}g^{-n_1} =$$

$$= \exp\left(\frac{2\pi i n_1 n_2}{2M+1}\right)g^{-n_1}h^{-n_2} = \mathbf{D_{-n}}.$$

- Proof of (32b):

$$\mathbf{D_n D_m} = \exp\left(\frac{2\pi i(n_1 n_2 + m_1 m_2)}{2M+1}\right)g^{n_1}h^{n_2}g^{m_1}h^{m_2} =$$

$$\exp\left(\frac{2\pi i(n_1 n_2 + m_1 m_2)}{2M+1}\right)\sigma^{m_1 k}g^{n_1}h^{n_2-k}g^{m_1}h^{m_2+k} =$$

$$\exp\left(\frac{2\pi i(n_1 n_2 + m_1 m_2 + 2m_1 n_2)}{2M+1}\right)g^{m_1+n_1}h^{m_2+n_2} =$$

$$\exp\left(\frac{2\pi i[(m_1+m_2)(n_1+n_2) + (m_1 n_2 - m_2 n_1)]}{2M+1}\right)g^{m_1+n_1}h^{m_2+n_2} =$$

$$= \exp(i\varepsilon(\mathbf{m}\times\mathbf{n}))\mathbf{D_{n+m}}.$$

- Proof of (32c): first note that

$$\mathrm{Tr}(\mathbf{D_k}) = \sigma^{k_1 k_2/2}\mathrm{Tr}(g^{k_1}h^{k_2}).$$

Then, $\mathrm{Tr}(g^j) = 0$ for $j \neq (2M+1)s_1$, $s_1 \in \mathcal{Z}$, because

$$\sum_{p=0}^{2M}\exp\left(\frac{4\pi i j p}{2M+1}\right) = 0, \quad j \neq (2M+1)s_1.$$

Now we can see that $\mathrm{Tr}(g^j h^k) = 0$ for $k \neq (2M+1)s_2$, $s_2 \in \mathcal{Z}$, because h^k shifts non-zero entries of g^j off the diagonal position. Otherwise, we obtain $\mathrm{Tr}(g^j h^k) = \mathrm{Tr}(I) = 2M+1$ for $j = (2M+1)s_1$ and $k = (2M+1)s_2$. Taking into account that our indices are between $-M$ and M we have

$$\mathrm{Tr}(\mathbf{D_k}) = \begin{cases} 2M+1, & \mathbf{k} = \mathbf{0}, \\ 0, & \mathbf{k} \neq \mathbf{0}, \end{cases}$$

which gives the third statement of (32).

From (32) immediately follows

$$[\mathbf{D_n}, \mathbf{D_m}] = 2i\sin(\varepsilon(\mathbf{m}\times\mathbf{n}))\mathbf{D_{n+m}},$$

$$[\mathbf{D_n}, \mathbf{D_m}]_+ = 2\cos(\varepsilon(\mathbf{m}\times\mathbf{n}))\mathbf{D_{n+m}},$$

(33)

where $[\cdot, \cdot]$ and $[\cdot, \cdot]_+$ are the commutation and anti-commutation operators, respectively defined as follows:

$$[x, y] = xy - yx, \quad [x, y]_+ = xy + yx.$$

Now we will show how to write the sine-bracket equation in its "bracket" form via the commutation rule (33). We denote

$$Q = \sum_{\mathbf{n}} \hat{q}_{\mathbf{n}} \mathbf{D_n} \quad \text{and} \quad F = \sum_{\mathbf{n}} f_{\mathbf{n}} \mathbf{D_n},$$

where

$$f_{\mathbf{k}} = \frac{\partial H}{\partial \hat{q}_{\mathbf{k}}},$$

and H is an arbitrary Hamiltonian. We write the analogue of the sine-bracket truncation (27) with the arbitrary Hamiltonian in the form

$$\dot{Q} = \frac{i}{2\varepsilon}[Q, F]. \tag{34}$$

In order to show that (34) is true, we first write down the commutator $[Q, F]$,

$$\frac{i}{2\varepsilon}[Q, F] = \frac{i}{2\varepsilon} \left(\sum_{\mathbf{n}} \hat{q}_{\mathbf{n}} \mathbf{D_n} \right) \left(\sum_{\mathbf{m}} f_{\mathbf{m}} \mathbf{D_m} \right) - \left(\sum_{\mathbf{m}} f_{\mathbf{m}} \mathbf{D_m} \right) \left(\sum_{\mathbf{n}} \hat{q}_{\mathbf{n}} \mathbf{D_n} \right) =$$

$$= \frac{i}{2\varepsilon} \sum_{\mathbf{n},\mathbf{m}} \hat{q}_{\mathbf{n}} f_{\mathbf{m}} (\mathbf{D_n D_m} - \mathbf{D_m D_n}) = \frac{1}{\varepsilon} \sum_{\mathbf{n},\mathbf{m}} \sin(\varepsilon(\mathbf{m} \times \mathbf{n})) \hat{q}_{\mathbf{n}} f_{\mathbf{m}} \mathbf{D_{m+n}} =$$

$$= -\frac{1}{\varepsilon} \sum_{\mathbf{k},\mathbf{m}} \sin(\varepsilon(\mathbf{k} \times \mathbf{m})) \hat{q}_{\mathbf{k+m}} f_{-\mathbf{m}} \mathbf{D_k}.$$

Here we denote $\mathbf{k} = \mathbf{m} + \mathbf{n}$ and then change $\mathbf{m}$ to $-\mathbf{m}$. Now we write $\dot{Q}$ as

$$\dot{Q} = \sum_{\mathbf{k}} \dot{\hat{q}}_{\mathbf{k}} \mathbf{D_k},$$

and (34) becomes

$$\sum_{\mathbf{k}} \dot{\hat{q}}_{\mathbf{k}} \mathbf{D_k} = -\frac{1}{\varepsilon} \sum_{\mathbf{k},\mathbf{m}} \sin(\varepsilon(\mathbf{k} \times \mathbf{m})) \hat{q}_{\mathbf{k+m}} f_{-\mathbf{m}} \mathbf{D_k}. \tag{35}$$

The matrices $\mathbf{D_k}$ are linearly independent; therefore the equality (35) must hold separately for each $\mathbf{D_k}$, hence (34) is equivalent to (27).

We expect Casimirs to be traces of Q^k. To show this, first note that from the commutation rule in (33), we can see that $\mathrm{Tr}[\mathbf{D_n}, \mathbf{D_m}] = 0$. From (32), Q^k is a linear combination of $\mathbf{D}$'s so that time-derivative of Q^k is a set of $\mathbf{D}$-commutators, therefore

the traces, $\text{Tr}(Q^k)$, are conserved in time with an arbitrary Hamiltonian. Explicitly

$$
\begin{aligned}
\frac{1}{2M+1}\text{Tr}(Q^k) &= \frac{1}{2M+1}\text{Tr}\left(\sum_{\mathbf{i}}\hat{q}_{\mathbf{i}}\mathbf{D}_{\mathbf{i}}\right)^k = \\
&= \frac{1}{2M+1}\sum_{\mathbf{i}_1,\ldots,\mathbf{i}_k}\hat{q}_{\mathbf{i}_1},\ldots,\hat{q}_{\mathbf{i}_k}\text{Tr}(\mathbf{D}_{\mathbf{i}_1},\ldots,\mathbf{D}_{\mathbf{i}_k}) = \\
&= \sum_{\sum_j^k \mathbf{i}_j=0}\hat{q}_{\mathbf{i}_1},\ldots,\hat{q}_{\mathbf{i}_k}\exp\{i\varepsilon[\mathbf{i}_2\times\mathbf{i}_1+\mathbf{i}_3\times(\mathbf{i}_1+\mathbf{i}_2)+ \\
&\qquad +\ldots+\mathbf{i}_{k-1}\times(\mathbf{i}_1+\ldots+\mathbf{i}_{k-2})]\} = \\
&= \sum_{\sum_j^k \mathbf{i}_j=0}\hat{q}_{\mathbf{i}_1},\ldots,\hat{q}_{\mathbf{i}_k}\cos\{\varepsilon[\mathbf{i}_2\times\mathbf{i}_1+\mathbf{i}_3\times(\mathbf{i}_1+\mathbf{i}_2)+ \\
&\qquad +\ldots+\mathbf{i}_{k-1}\times(\mathbf{i}_1+\ldots+\mathbf{i}_{k-2})]\},
\end{aligned}
\tag{36}
$$

i.e. (36) coincides with (30).

The fact that the sine-bracket truncation conserves the same low-order quantities as the traditional truncation makes it a very convenient model to study the dynamics of a truncation for quasi-geostrophic flow with many conserved quantities in comparison with the traditional energy-enstrophy conserving truncation. A systematic statistical study is presented by the authors in [1] and [3].

In the Section 2.6 we will show that the standard numerical time integrators preserve energy, but do not preserve enstrophy and higher order Casimir invariants. The special type of integrator which allows us to preserve all Casimir invariants with machine precision is presented in the next section.

2.5. The Symplectic Numerical Time Integrator for the Truncation with Many Conserved Quantities. In this section we describe the special kind of time integrator which allows us to preserve the Poisson bracket. This integrator was suggested by R. I. McLachlan [19]. If the symplectic structure is preserved explicitly, then the Casimir invariants associated with corresponding Poisson bracket will be conserved with machine precision. The integrator we describe here uses the symplectic splitting method, the idea of which is that if the Poisson bracket is linear with respect to $\hat{q}_{\mathbf{k}}$ (also called *Lie-Poisson bracket*), then, splitting the Hamiltonian in a special way, we can make the system linear. Of course, splitting the Hamiltonian calls for an approximate solution, and here we describe the first-order symplectic integrator followed by the simple second-order accurate extension. The integrator can as well be used to integrate the traditional truncation (25) despite the fact that it lacks the Hamiltonian structure and the family of Casimir invariants (30). In that case it is shown below that only the enstrophy is preserved with machine precision.

Our symplectic matrix $\mathbf{J}_{\mathbf{kk}'}$ is linear with respect to $\hat{q}_{\mathbf{k}}$,

$$
\mathbf{J}_{\mathbf{kk}'} = -\frac{1}{\varepsilon}\sin(\varepsilon\mathbf{k}\times\mathbf{k}')\hat{q}_{\mathbf{k}+\mathbf{k}'}.
$$

Let us define the special subset σ of the two-dimensional space of integers such that

$$
\sigma\subset\mathcal{Z}^2,\ \mathbf{J}_{\mathbf{kk}'}=0\ \ \forall\mathbf{k},\mathbf{k}'\in\sigma.
$$

If the Hamiltonian is

$$H(\hat{q}_{\mathbf{k}_1}, \ldots, \hat{q}_{\mathbf{k}_n}) = H(\sigma) \tag{37}$$

for some $\mathbf{k}_1, \ldots, \mathbf{k}_n \in \sigma$ then the system

$$\frac{\mathrm{d}}{\mathrm{d}t} \hat{q}_{\mathbf{k}} = \mathbf{J}_{\mathbf{k}\mathbf{k}'} \frac{\partial H}{\partial \hat{q}_{\mathbf{k}'}}$$

is linear. This happens because the variables $\hat{q}_{\mathbf{k}'}$, which belong to the Hamiltonian, are conserved in time. For the linear system we can find an analytic solution.

The Hamiltonian in (23) does not belong to such a class. However note that the Hamiltonian is a sum of small pieces, and each piece depends on a single spectral coefficient $\hat{q}_{\mathbf{k}}$. We split the Hamiltonian into smaller pieces such that each of them separately has the structure in (37):

$$H = \sum_j H_j(\sigma_j), \ [\sigma_j \subset \mathcal{Z}^2, \ \mathbf{J}_{\mathbf{k}\mathbf{k}'} = 0 \ \ \forall \mathbf{k}, \mathbf{k}' \in \sigma_j] \ \forall j.$$

Now for a small time step we can integrate the system sequentially for all $H_j(\sigma_j)$ which will give us a first-order approximation to the exact solution. What we essentially do is that we analytically integrate the same Poisson bracket with different Hamiltonians, therefore the Casimirs (which are the property of Poisson bracket alone) are preserved with machine precision.

In our case with the sine-bracket equation for any $\mathbf{k} = (k_1, k_2)$ we denote

$$\sigma(\mathbf{k}) = \{n\mathbf{k} : -M \le n \le M\}.$$

To split the Hamiltonian we need the set $\mathbf{K}$ which consists of several vectors $\mathbf{k}_j$ to satisfy the following: let us denote $\sigma(\mathbf{K}) = \cup_j \sigma(\mathbf{k}_j)$, then $\sigma(\mathbf{K})$ should be such, that $\forall m_1, m_2 \in [-M, M]$ we would have $\mathbf{m} \in \sigma(\mathbf{K})$. Also $\mathbf{K}$ should satisfy $\sigma(\mathbf{k}_1) \cap \sigma(\mathbf{k}_2) = \emptyset \ \forall \mathbf{k}_1, \mathbf{k}_2 \in \mathbf{K}$. In other words, $\mathbf{K}$ should be such that $\sigma(\mathbf{K})$ would be a disjoint set which covers the whole lattice. When $(2M+1)$ is a prime number, one possible choice of $\mathbf{K}$ is

$$\mathbf{K} = \{(0, 1)\} \cup \{(1, m) : -M \le m \le M\}. \tag{38}$$

We can see that for this choice $\sigma(\mathbf{K})$ is not completely disjoint, but the point at the origin never evolves in time, so (38) suffices. Then for the Hamiltonian we have

$$H = \frac{1}{2} \sum_{n_1, n_2 = -N}^{N} \frac{|\hat{q}_{\mathbf{n}} - \hat{h}_{\mathbf{n}}|^2}{|\mathbf{n}|^2} = \sum_{\mathbf{k} \in \mathbf{K}} H_{\mathbf{k}},$$

$$H_{\mathbf{k}} = \frac{1}{2} \sum_{n = -N}^{N} \frac{(\hat{q}_{n\mathbf{k}} - \hat{h}_{n\mathbf{k}})(\hat{q}_{-n\mathbf{k}} - \hat{h}_{-n\mathbf{k}})}{|n\mathbf{k}|^2}. \tag{39}$$

In the above equation we should take into account that the expression $n\mathbf{k}$ in the denominator of $H_{\mathbf{k}}$ is aliased, i.e. whenever $n\mathbf{k}$ is out of bounds, the value is aliased from the opposite side of the domain by the periodicity rule (26).

Now we need to solve $\dot{\mathbf{q}} = \mathbf{J}\nabla H_{\mathbf{k}}$ (sequentially for all $\mathbf{k} \in \mathbf{K}$, of course). For the aforementioned $\mathbf{K}$ we can write (27) in the form

$$\dot{z}_m = \sum_{n=-M}^{M} a_n z_{m-n}, \tag{40}$$

where $z_m = \hat{q}_{\mathbf{j}+m\mathbf{k}}$, $\mathbf{j} = (0,j), 0 < j \le M$ for $\mathbf{k} = (1,0)$ and $\mathbf{j} = (j,0), 0 < j \le M$ for the rest of $\mathbf{k}$'s $\in \mathbf{K}$ (Here we took into account the fact that $\hat{q}_{\mathbf{m}} = \hat{q}^*_{-\mathbf{m}}$, so that we can solve only for the half of the lattice). Also

$$a_n = -\frac{\sin(\varepsilon n \mathbf{j} \times \mathbf{k})}{\varepsilon |n\mathbf{k}|^2}(\hat{q}_{n\mathbf{k}} - \hat{h}_{n\mathbf{k}}).$$

The system (40) is linear and therefore can be solved analytically. The following trick helps us to diagonalize the system. Note that (40) is essentially discretized convolution, hence using the discrete Fourier transform yields the disjoint set of equations:

$$\dot{\hat{z}}_p = \sum_{m=0}^{2M+1} z_m e^{-2\pi i p m/(2M+1)} = \sum_{m=0}^{2M+1} \sum_{n=-M}^{M} a_n z_{m-n} e^{-2\pi i p m/(2M+1)} =$$

$$= \sum_{n=-M}^{M} \sum_{s=-n}^{2M+1-n} a_n z_s e^{-2\pi i p(n+s)/(2M+1)} =$$

$$= \sum_{n=-M}^{M} a_n e^{-2\pi i p n/(2M+1)} \sum_{s=-n}^{2M+1-n} z_s e^{-2\pi i p s/(2M+1)} = \hat{a}_p \hat{z}_p.$$

Thus (40) in Fourier space has the form

$$\dot{\hat{z}}_i = \Lambda_{ij}\hat{z}_j, \quad \Lambda_{ij} = \left\{ \begin{array}{ll} \hat{a}_i: & i = j \\ 0: & i \ne j \end{array} \right. , \text{ where } \hat{\mathbf{a}} = FFT(\mathbf{a})$$

which is explicitly integrable:

$$\hat{z}_p(t + \Delta t) = e^{\Delta t \Lambda_{pp}}\hat{z}_p(t),$$

and (40) can be written in the explicit form

$$\mathbf{z}(t + \Delta t) = FFT^{-1}e^{\Delta t \Lambda}FFT\mathbf{z}(t).$$

Thus, each equation $\dot{\mathbf{q}} = \mathbf{J}\nabla H_{\mathbf{k}}$ can be solved *exactly* using fast Fourier transform $3M$ times (forward, backward and for a_n). This means that the conservation of symplectic matrix properties (such as Casimirs) should happen close to round-off error. Taking into account that we have $2(M+1)$ $\mathbf{k}$'s $\in \mathbf{K}$, the whole amount of FFT's is $6M(M+1)$ for each step of integration.

Second order symplectic integrator. The Casimir invariants (30), including enstrophy, should be conserved with machine precision with the above method of symplectic integration. However, the energy (23) is not a Casimir invariant, and

therefore the extent to which the energy is conserved is determined by the order of approximation provided by the time integrator. So far we described a symplectic integrator of the first order, which requires rather small time steps to conserve energy. A simple extension of the first order symplectic integrator to second order accuracy can be obtained with almost complete absence of additional complexity via Strang splitting.

When we sequentially integrate the system for the Hamiltonians $H_j(\sigma_j)$ we use the fact that

$$e^{\Delta t(A+B)} = e^{\Delta t A} e^{\Delta t B} + \mathcal{O}(\Delta t^2),$$

where A, B are matrices, that is why we have the first order accurate integrator. Using the fact that

$$e^{\Delta t(A+B)} = e^{\Delta t A/2} e^{\Delta t B} e^{\Delta t A/2} + \mathcal{O}(\Delta t^3),$$

or, for many matrices,

$$e^{\Delta t(A_1+\ldots+A_n)} = e^{\Delta t A_1/2} \ldots e^{\Delta t A_{n-1}/2} e^{\Delta t A_n} e^{\Delta t A_{n-1}/2} \ldots e^{\Delta t A_1/2} + \mathcal{O}(\Delta t^3),$$

we readily construct a second order accurate symplectic integrator. The second order symplectic integrator is twice as slow as the first order symplectic integrator, but it reduces the relative errors in energy by roughly four orders of magnitude for typical time steps.

2.6. Numerical Experiments with the Sine-Bracket and Traditional Spectral Truncations. In this section first we present the results of a numerical validation study of the second order symplectic integrator derived in 2.5 for the two truncations. We check that the traditional truncation preserves the energy (23) and the enstrophy (24), and does not preserve any other Casimir invariants from (30), and that the sine-bracket truncation preserves the energy, enstrophy, and also these Casimir invariants. Moreover, we check that neither truncation preserves the exactly truncated versions of the infinite-dimensional Casimir invariants in (19), which we here refer to as $\hat{C}_N$. In addition, we show that the standard numerical time integrators do not preserve the enstrophy and the high order Casimir invariants with small time steps.

2.6.1. Conservation of the energy and the Casimir invariants. For the all numerical experiments we choose random initial conditions with fixed energy and enstrophy. We take an 11×11 square domain of Fourier coefficients and integrate the system for the time $T = 5200$ with the time step $\Delta t = 0.01$. We pick the energy to be $E = 7$, and enstrophy $\mathcal{E} = 20$.

We expect both the traditional (25) and sine-bracket (27) truncations to preserve energy and enstrophy when used in combination with the second order symplectic integrator. The results of numerical validations for energy and enstrophy are shown in Figure 1 for the traditional truncation, and Figure 2 for the sine-bracket truncation. As we can see, the energy is preserved within a $4 \cdot 10^{-5}$ relative error for the traditional truncation, and $1.2 \cdot 10^{-6}$ for the sine-bracket truncation. The enstrophy is preserved with machine precision for both truncations.
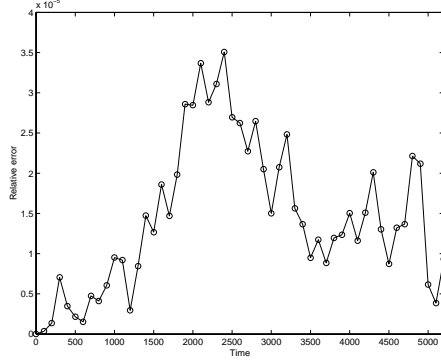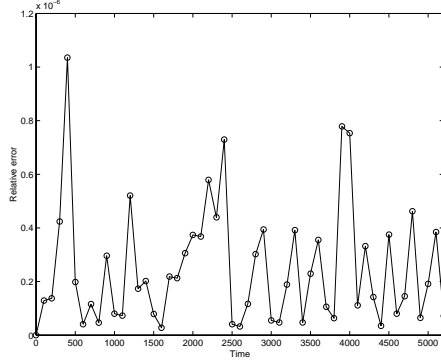
The relative error of energy for the traditional truncation. Number of modes 11×11, integration time $T = 5200$, time step $\Delta t = 0.01$, scale 10^{-5}.
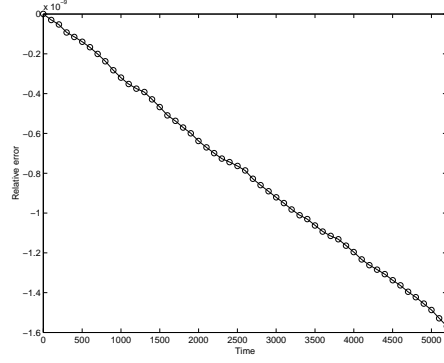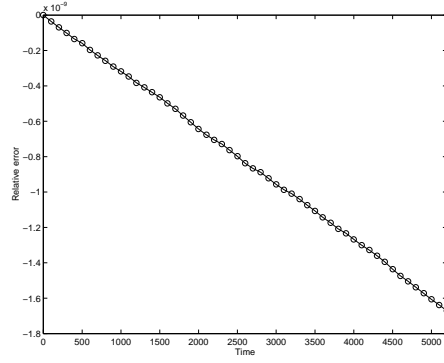
The relative error of enstrophy for the traditional truncation. Number of modes 11×11, integration time $T = 5200$, time step $\Delta t = 0.01$, scale 10^{-9}.

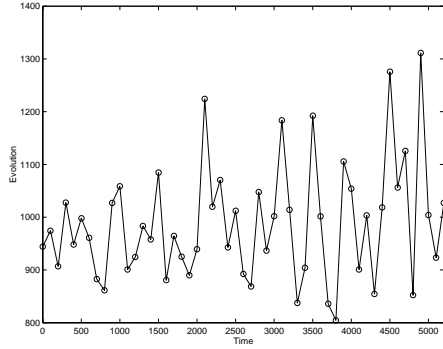FIG. 1. *The conservation of energy and enstrophy for the traditional truncation*

The relative error of energy for the sine-bracket truncation. Number of modes 11×11, integration time $T = 5200$, time step $\Delta t = 0.01$, scale 10^{-6}.
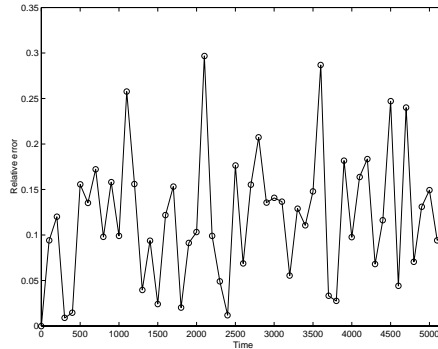
The relative error of enstrophy for the sine-bracket truncation. Number of modes 11×11, integration time $T = 5200$, time step $\Delta t = 0.01$, scale 10^{-9}.

FIG. 2. *The conservation of energy and enstrophy for the sine-bracket truncation*

Next, we show that the sine-bracket truncation preserves the Casimir invariants (30), and the traditional truncation does not preserve them. In addition, we also make sure that neither truncation preserves the truncated versions of the infinite-dimensional Casimir invariants in (19). The results are shown in Figures 3 for the traditional truncation, and 4 for the sine-bracket truncation. Figure 3 shows the time evolution of $\hat{C}_4$ and the relative error of C_4 conservation for the traditional truncation. The quantity $\hat{C}_4$ jumps within 800 and 1300, which means that the relative error for $\hat{C}_4$ is about 50%. The relative error for C_4 for the traditional truncation is 30%, thus neither C_4 or $\hat{C}_4$ is conserved for the traditional truncation. Figure 4 shows the time evolution of $\hat{C}_4$ and the relative error of C_4 conservation for the sine-bracket truncation. The $\hat{C}_4$, again, jumps within 800 and 1300, which means that the relative error for $\hat{C}_4$ is about 50%, and the sine-bracket truncation does not conserve $\hat{C}_4$. However, the relative error for the Casimir C_4 is within $3.5 \cdot 10^{-9}$ for the sine-bracket truncation, which means that C_4 is conserved within machine precision by the sine-bracket truncation.
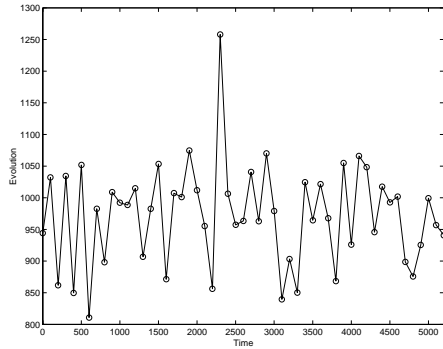
The evolution of $\hat{C}_4$ for the traditional truncation. Number of modes 11×11, integration time $T = 5200$, time step $\Delta t = 0.01$.
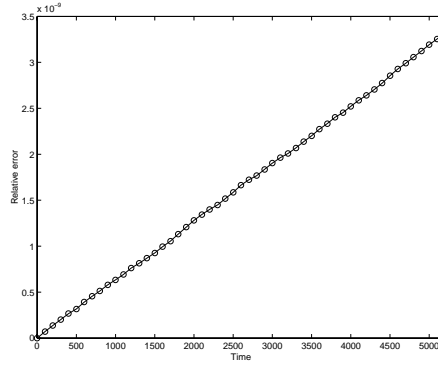
The absolute value of relative error of C_4 for the traditional truncation. Number of modes 11×11, integration time $T = 5200$, time step $\Delta t = 0.01$.

FIG. 3. *The time evolution of $\hat{C}_4$ and C_4 for the traditional truncation*

The evolution of $\hat{C}_4$ for the sine-bracket truncation. Number of modes 11×11, integration time $T = 5200$, time step $\Delta t = 0.01$.

The absolute value of relative error of C_4 for the sine-bracket truncation. Number of modes 11×11, integration time $T = 5200$, time step $\Delta t = 0.01$. The vertical scale is 10^{-9}.

FIG. 4. *The time evolution of $\hat{C}_4$ and the conservation of C_4 for the sine-bracket truncation*

2.6.2. Unstable behavior of standard numerical schemes. One can ask the following question: is there any real need to use the complicated symplectic integrator while there is plenty of standard high-order methods of integration like Runge-Kutta or the Adams-Bashforth family? In this section we check whether the two standard methods of integration can be used to solve at least the traditional truncation with two conserved quantities. It happens that both methods preserve the energy quite well, but the enstrophy is not conserved by both methods.

We perform the computations with the traditional truncation (25) using standard methods of integration. We calculate the sum on the right-hand side of (25) using the Fast Fourier Transform, which yields $\sim N^2 \log(N)$ floating point operations per time step. To exclude the modes which exceed the range of the lattice we use the twice-extended in both directions array for the Fast Fourier Transform. For the integration in time we use two methods:
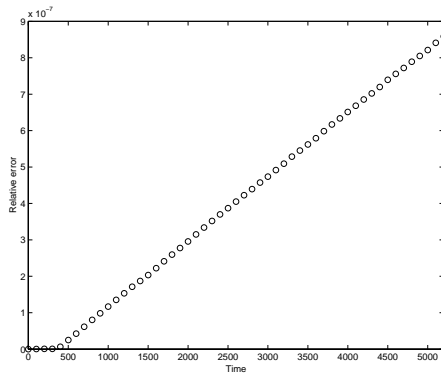
The relative error of energy for the Runge-Kutta 4th order method, $\mathcal{E} = 20$, scale 10^{-7}.
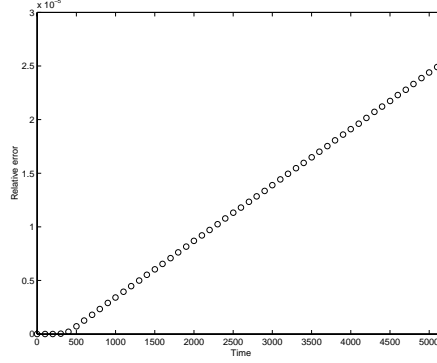
The relative error of energy for the Adams-Bashforth-Moulton 4th order predictor-corrector method, $\mathcal{E} = 20$, scale 10^{-5}.

FIG. 5. *The conservation of energy for the two standard time integrators, 11×11 truncation size*

- the standard Runge-Kutta method of 4th order,
- the Adams-Bashforth-Moulton predictor-corrector method of 4th order.

We compute the traditional truncation (25) for the 11×11 size of truncation for the time $T = 5200$ and the time step $\Delta t = 0.001$. The same values of energy and enstrophy ($E = 7$, $\mathcal{E} = 20$) were used for this numerical experiment. As we can see, the conservation of energy is satisfactory for this integration time, as shown in Figure 5. However, the enstrophy is conserved only up to the time $T \approx 400$, and then its relative error goes sky-high, beyond 1000%, both for both Runge-Kutta and Adams-Bashforth-Moulton methods, as shown in Figure 6.

This blow up effect for enstrophy is the well-known nonlinear instability of spectral methods for fluid flow with standard time integrators. The usual remedy is a spectral filter in Fourier space proposed originally by Majda and Osher [13]. See the paper by E and Shu [5] for an important application. We have shown above that this standard nonlinear instability does not occur with the second order accurate symplectic integrator and the traditional truncation. However, this integrator is much more expensive to use, while the spectral filter destroys the exact conservation of enstrophy.

3. Convergence of the Traditional and Sine-Bracket Truncations to Weak Solutions of the Quasi-Geostrophic Equations. Here we sketch a proof of the fact that both the traditional spectral truncation discussed in 2.3 and the sine-bracket truncation described in 2.4 converge to weak solutions of the quasi-geostrophic equations with finite enstrophy, $Q_2(q(\vec{x}, t)) < \infty$, provided that the initial data $q_0(\vec{x})$ satisfies $Q_2(q_0(\vec{x})) < \infty$. The proof relies on the key fact established in Section 2 that these truncations conserve both the energy and the enstrophy of the initial data in time combined with a straightforward adaptation of the arguments in Chapter 10 of the book by Majda and Bertozzi [10] to weak solutions formulated in terms of the vorticity-stream equations in (1) rather then the primitive variable formulation. Clearly, the existence of weak solutions in the vorticity-stream formulation imposes more stringent requirements beyond the primitive variable formulation. To avoid needless repetition, the material below is presented as if the reader has Chapter 10 of [10] readily available.
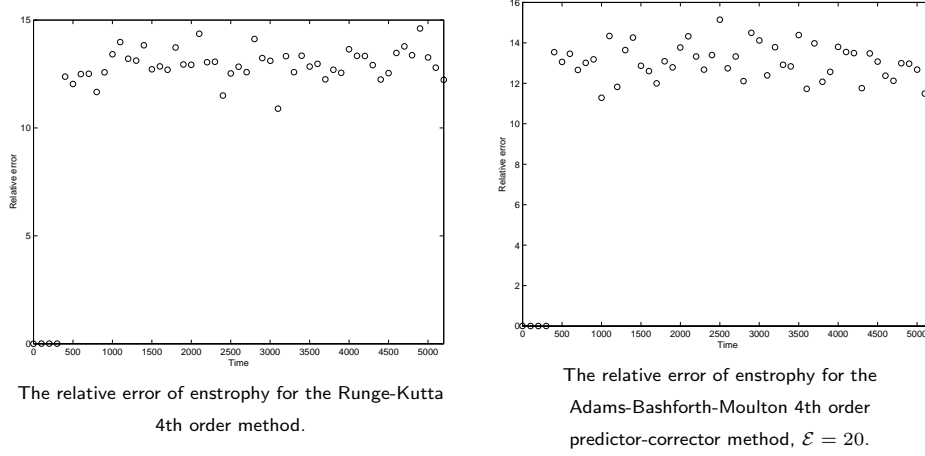
With (1), we formulate the following

The relative error of enstrophy for the Runge-Kutta
4th order method.

The relative error of enstrophy for the
Adams-Bashforth-Moulton 4th order
predictor-corrector method, $\mathcal{E} = 20$.

FIG. 6. *The conservation of enstrophy for the two standard time integrators, 11×11 truncation size*

DEFINITION 1. *The potential vorticity, $q = \omega + h$, defines a weak solution of the quasi-geostrophic equations with initial data $q_0(\vec{x}) \in L^1(T^2)$, provided that*

(i) $q \in L^1([0, T_*] \times T^2)$ *for any* $T_* > 0$,
(ii) $\vec{v}$ *satisfies*

$$\mathrm{curl}\,\vec{v} = -(v_1)_y + (v_2)_x = q - h \tag{41a}$$

$$\mathrm{div}\,\vec{v} = 0 \tag{41b}$$

in the sense of distributions,
(iii) $\vec{v}q \in L^1([0, T_*] \times T^2)$,
(iv) *for any test function* $\phi \in C^1([0, T_*] \times T^2)$

$$\int_{T^2} \phi(\vec{x}, T_*) q \, \mathrm{d}\vec{x} - \int_{T^2} \phi(\vec{x}, 0) q_0 \, \mathrm{d}\vec{x} =$$
$$= \int_0^{T_*} \int_{T^2} (\phi_t q + \nabla\phi \cdot (\vec{v}q)) \, \mathrm{d}\vec{x} \, \mathrm{d}t.$$

DEFINITION 2. *As discussed in 10.2.3 of [10], a sequence of smooth functions, q^ε and $\vec{v}^\varepsilon$, is an* approximate solution sequence with L^p-vorticity control *for the quasi-geostrophic equations provided that*

(i) $\displaystyle\max_{0 \leq T \leq T_*} \|q^\varepsilon\|_p(t) \leq C(T_*)$, $\displaystyle\max_{0 \leq T \leq T_*} \|\vec{v}^\varepsilon\|_2(t) \leq C(T_*)$, $\forall\, T_* > 0$.
(ii) $\mathrm{curl}\,\vec{v}^\varepsilon = q^\varepsilon$, $\langle \vec{v}^\varepsilon = 0 \rangle$, $\mathrm{div}\,\vec{v}^\varepsilon = 0$, $\langle q^\varepsilon \rangle = 0$.

$$\tag{42}$$

(iii) $q^\varepsilon(t)$ *is uniformly Lipschitz continuous in some negative normed Sobolev space $H^{-L}(T^2)$ for a fixed L.*
(iv) Weak consistency*: for any smooth test function, $\phi(\vec{x}, t)$*

$$\int_0^{T_*} \int_{T^2} (\phi_t q^\varepsilon + \nabla\phi \cdot (\vec{v}^\varepsilon q^\varepsilon)) \, \mathrm{d}\vec{x} \, \mathrm{d}t \to 0 \ \ as \ \varepsilon \to 0.$$

Here and below, $\|f\|_p = \left(\int_{T^2} |f|^p\right)^{1/p}$ denotes the standard L^p-norm on the torus with $\langle f \rangle$ the mean of f over the period.

EXAMPLE: SMOOTHING THE INITIAL DATA. Consider any initial data, $q_0 = \omega_0 + h$ with $h \in L^p$ and $\omega_0 \in L^p$ and approximate ω_0, h by smooth functions $\omega_0^\varepsilon, h^\varepsilon$ which converge in L^p to ω_0, h as $\varepsilon \to \infty$. Let $q^\varepsilon(\vec{x}, t) = \omega^\varepsilon(\vec{x}, t) + h^\varepsilon$ denote the global smooth solution of (1) with this initial data (see Chapter 3 or 4 of [10] for a proof), then $q^\varepsilon, \vec{v}^\varepsilon$ defines an approximate solution sequence with L^p vorticity control (see Chapter 10 in [10]).

PROPOSITION 1. *Both the traditional spectral and sine-bracket truncations presented above in Sections 2.3 and 2.4 define approximate solution sequences $(q^\varepsilon, \vec{v}^\varepsilon)$ with L^2 vorticity control for the quasi-geostrophic equations provided that the initial data for these two finite-dimensional spectral methods is given by the Fourier coefficients of an initial potential vorticity, $q_0 \in L^2(T^2)$ with*

$$q_0 = \omega_0 + h, \quad \omega_0 \in L^2(T^2), \quad h_0 \in L^2(T^2). \tag{43}$$

Here, as in Sections 2.3 and 2.4 above, ε is given by $\varepsilon = 2\pi/(2M + 1)$ in defining $q^\varepsilon, \vec{v}^\varepsilon$, where M denotes the number of Fourier modes in x, y, i.e. $|k_j| \leq M$, $j = 1, 2$.

Proposition 1 is intuitively clear given the conservation properties for the traditional and sine-bracket truncations established in Sections 2.3, 2.4 above; we postpone a detailed proof until the end of the section.

The main result of this section in the following

THEOREM 1. *Assume the initial potential vorticity satisfies $q_0 = \omega_0 + h$ for $\omega_0 \in L^p(T^2)$ and $h \in L^p(T^2)$. If $p > 4/3$, then there is a weak solution of the quasi-geostrophic equations with this initial data. This weak solution can be constructed by taking the global smooth solutions with the smoothed initial data $\omega_0^\varepsilon, \vec{v}_0^\varepsilon$ from the Example above and passing to the limit. Furthermore, for $p > 4/3$, any approximate solution sequence with L^p-vorticity control converges (by passing to subsequences) to a weak solution of the vorticity-stream formulation which inherits the a priori bounds in (42)(i),(ii),(iii), as well as the stronger $L^q(T^2)$ velocity bound in (48) below.*

By combining Proposition 1 and Theorem 1 we have the immediate

COROLLARY 1. *Consider initial data for the potential vorticity, $q_0 = \omega_0 + h$ with $\omega_0 \in L^2(T^2)$, $q_0 \in L^2(T^2)$. Then both the traditional spectral truncation and the sine-bracket truncation converge (by passing to subsequences) as $M \to \infty$ to weak solutions of the quasi-geostrophic equations in vorticity-stream form with both finite energy and finite enstrophy.*

An extremely interesting open mathematical problem is to establish for initial data q_0 with $Q_p(q_0) < \infty$ for all large p, that the additional conserved quantities in the sine-bracket truncation, C_N, for $N = 1, \ldots, M$, which only weakly approximate various Q_p as $M \to \infty$ nevertheless yield additional regularity beyond Corollary 1 for the weak solutions which are limits of this truncation. Next we present the

Sketch of the proof of Theorem 1. From Theorem 10.1 in [10], under the above hypothesis we already know by passing to a subsequence that there exist $\vec{v}, \omega$ with

$$\operatorname{div} \vec{v} = 0, \quad \operatorname{curl} \vec{v} = \omega \tag{44}$$

and

A) ω^ε converging weakly to ω in $L^\infty([0, T_*], L^p(T^2))$,

$$\text{(45)}$$

B) $\displaystyle\int_0^{T_*} \|\vec{v}^\varepsilon - \vec{v}\|_1(s)\,\mathrm{d}s \to 0$ for any $T_* > 0$.

The key fact needed to complete the proof of the theorem is to establish that $\vec{v}\omega \in L^1([0, T_*] \times T^2)$ with

$$\lim_{\varepsilon \to 0} \int_0^{T_*} \int_{T^2} \phi(\vec{v}^\varepsilon \omega^\varepsilon - \vec{v}\omega) = 0 \tag{46}$$

for any test function ϕ and any $T_* > 0$ for either the specific approximate solution sequence from the Example or a more general sequence satisfying Definition 1. Since the period domain, T^2, is compact, it is no loss of generality to consider p with $4/3 < p < 2$. Then with (ii) from (42), elliptic regularity, and the Sobolev inequality on page 400 of [10], we have

$$\|\vec{v}^\varepsilon(t)\|_{p'} \le C\|\omega^\varepsilon(t)\|_p, \quad p' = \frac{2p}{2-p}, \tag{47}$$

where that constant C depends only on p. From (i) of (42) we obtain the additional estimates

$$\max_{0 \le t < T_*} \|\vec{v}^\varepsilon(t)\|_{\frac{2p}{2-p}} \le C, \tag{48a}$$

$$\vec{v}(t) \in L^\infty([0, T_*], L^{\frac{2p}{2-p}}(T^2)). \tag{48b}$$

Since $p > 4/3$, we have $p/(p-1) < 2p/(2-p)$, so the Holder inequality and (48) give

$$\|\omega\vec{v}\|_1(t) \le \|\omega(t)\|_p\|\vec{v}(t)\|_{\frac{p}{p-1}} \le C\|\omega(t)\|_p\|\vec{v}(t)\|_{\frac{2p}{2-p}}, \tag{49}$$

so that $\omega\vec{v} \in L^1$. Also, (48) and the convergence in (45B) combine with standard interpolation (pg. 399 of [10]) to guarantee that for $p > 4/3$

$$\int_0^{T_*} \|\vec{v}^\varepsilon - \vec{v}\|_{\frac{p}{p-1}}(s)\,\mathrm{d}s \to 0 \quad \text{as} \quad \varepsilon \to 0. \tag{50}$$

Thus, we have

$$\int_0^{T_*} \int_{T^2} \phi(\vec{v}^\varepsilon \omega^\varepsilon - \vec{v}\omega) = \int_0^{T_*} \int_{T^2} \phi(\vec{v}^\varepsilon - \vec{v})\omega^\varepsilon + \int_0^{T_*} \int_{T^2} \phi\vec{v}(\omega^\varepsilon - \omega), \tag{51}$$

and the first term in (51) tends to zero by (50), (42)(i) and the same Holder estimate as utilized in (49); the second term in (51) tends to zero because $\phi\vec{v} \in L^\infty([0, T_*], L^{\frac{p}{p-1}}(T^2))$ and $\omega^\varepsilon - \omega$ tends to zero weakly in $L^\infty([0, T_*], L^p(T^2))$. This completes the sketch of the proof of Theorem 1.

Sketch of proof of Proposition 1. As established in 2.3 and 2.4 above, both the traditional spectral and sine-bracket truncations conserve energy and enstrophy so that for any t with $0 \le t < \infty$,

$$\|q^\varepsilon(t)\|_2 \le \|q_0\|_2, \quad \|\vec{v}^\varepsilon(t)\|_2 \le \|\vec{v}_0\|_2, \tag{52}$$

while (42)(ii) is satisfied automatically by these approximations. The bounds in (52) and (49), applied to $\omega^\varepsilon, \vec{v}^\varepsilon$, guarantee for either approximation

$$\|\omega^\varepsilon \vec{v}^\varepsilon\|_1(t) \leq C(\|\omega_0\|_2 + \|h\|_2)^2. \tag{53}$$

First we consider the standard spectral truncation from 2.3 which can be written as

$$\frac{\partial q^\varepsilon}{\partial t} = -P_\varepsilon \left[\operatorname{div}(\vec{v}^\varepsilon q^\varepsilon)\right], \quad P_\varepsilon q^\varepsilon = q^\varepsilon, \tag{54}$$

with P_ε the projection operator given by (20). Since, by Sobolev's Lemma, L^1 is continuously injected into $H^{-s}(T^2)$ for $s > 1/2$, and P_ε is a uniformly bounded operator from H^{-R} to H^{-R}, (53) and (54) guarantee that $\partial q / \partial t$ is uniformly bounded in $H^{-(s+1)}(T^2)$ for $s > 1/2$ for any interval of time; thus, (42)(iii) is satisfied. With (54), the check of weak consistency reduces to the fact that

$$\lim_{\varepsilon \to 0} \int_0^{T_*} \int_{T^2} \nabla (I - P_\varepsilon)(\phi)(\vec{v}^\varepsilon q^\varepsilon) \, d\vec{x} \, dt = 0. \tag{55}$$

Since for a smooth test function ϕ

$$\max_{0 \leq t \leq T_*} |\nabla (I - P_\varepsilon)\phi|_\infty(t) \to 0$$

and $\vec{v}^\varepsilon q^\varepsilon$ is uniformly bounded in L^1 by (53), the requirement in (55) is clearly satisfied and the proof is finished for the standard truncation.

Next, we check weak consistency for the sine-bracket truncation. First, since the sine-bracket truncation satisfies both the bounds in (52) and (53), by approximating a smooth test function in the C^1 norm in space-time by finite trigonometric polynomials in $\vec{x}$, we can repeat the reasoning utilized in (53) to establish that it is sufficient to check the weak consistency in (42)(iv) for test functions

$$\phi = \sum_{|\mathbf{k}| \leq L_0} \hat{\phi}_{\mathbf{k}}(t) e^{i\vec{x} \cdot \mathbf{k}} \tag{56}$$

with L_0 fixed. Interpreting (42)(iv) Fourier component-wise via the Plancherel formula, the check of weak consistency for the test function in (56) is satisfied provided that with $\varepsilon = 2\pi/(2M + 1)$ the sum

$$\sum_{\mathbf{k}'} \left(\frac{\sin(\varepsilon(\mathbf{k} \times \mathbf{k}'))}{\varepsilon(\mathbf{k} \times \mathbf{k}')} - 1\right) (\mathbf{k} \times \mathbf{k}') \hat{q}^\varepsilon_{\mathbf{k}+\mathbf{k}'}(t) |\mathbf{k}'|^{-2} \hat{q}^\varepsilon_{-\mathbf{k}'}(t) \tag{57}$$

tends to zero for a fixed $\mathbf{k}$ with $|\mathbf{k}| \leq L_0$. We split $\mathbf{k}'$ into two groups of wavenumbers. For wavenumber $\mathbf{k}'$ with $|\mathbf{k}'| \leq \delta^{-1}$ and fixed $\mathbf{k}$, we have

$$\left| \frac{\sin(\varepsilon(\mathbf{k} \times \mathbf{k}'))}{\varepsilon(\mathbf{k} \times \mathbf{k}')} - 1 \right| \leq C\frac{\varepsilon}{\delta},$$

so that

$$\left| \sum_{|\mathbf{k}'| \leq \delta^{-1}} \cdots \right| \leq C\frac{\varepsilon}{\delta}\|q^\varepsilon(t)\|_2^2 \leq C\frac{\varepsilon}{\delta}\|q_0\|_2^2. \tag{58}$$

On the other hand, for fixed $\mathbf{k}$, the decay of (57) with $|\mathbf{k}'|$ implies that

$$\left| \sum_{|\mathbf{k}'| \geq \delta^{-1}} \ldots \right| \leq C\delta \|q^\varepsilon(t)\|_2^2 \leq C\delta \|q_0\|_2^2. \tag{59}$$

Thus, by choosing $\delta = \varepsilon^{1/2}$, (58) and (59) guarantee the fact needed in (57) to check the weak consistency for the sine-bracket truncation. The proof of the uniform Lipschitz bound in (42)(iii) for this truncation follows directly from the estimate for the Fourier coefficients similar to (59) with $\delta = 1$ which yields

$$\left| \frac{\partial \hat{q}_{\mathbf{k}}^\varepsilon}{\partial t} \right| (t) \leq C|\mathbf{k}|(\|q^\varepsilon(t)\|_2^2 + \|h\|_2^2) \leq C|\mathbf{k}|(\|q_0\|_2 + \|h\|_2)^2. \tag{60}$$

This guarantees the uniform Lipschitz continuity in $H^{-(s+1)}(T^2)$ for $s > 1/2$. The proof of Proposition 1 is now complete.

4. The Statistical Behavior of Discrete Approximations to the Burgers-Hopf Equation with a Nonlinear Energy. The statistical behavior of finite difference approximations to the Burgers-Hopf equation

$$u_t + \frac{1}{2}\left(u^2\right)_x = 0 \tag{61}$$

which conserve a discrete form of the momentum

$$M = \int u \, \mathrm{d}x \tag{62}$$

as well as the standard quadratic energy

$$E = \frac{1}{2} \int u^2 \, \mathrm{d}x \tag{63}$$

is studied recently by Timofeyev and the second author [18] through a combination of mathematical theory and numerical experiments. Two finite difference schemes are studied in [18], the Kruskal-Zabusky scheme and a five point energy conserving method besides the Galerkin spectral truncation to the Burgers-Hopf equation. Numerical experiments indicate that

> *The Kruskal-Zabusky scheme has equipartition of energy and strong mixing through exponential decay of correlations.*

$$\tag{64}$$

The five-point difference scheme has equipartition of energy but the failure of strong mixing at a few isolated Fourier modes due to resonance in an interesting fashion [18]. Furthermore, as expected from the equilibrium statistical predictions in [18] and confirmed below,

> *The statistical behavior of the Kruskal-Zabusky scheme at different grid points is essentially uncorrelated and the probability density function (PDF) of the values at a single grid point is a Gaussian distribution.*

$$\tag{65}$$

Here we build discrete approximations to (61) that conserve both linear momentum and a discrete version of the nonlinear energy

$$E_{a,b} = \frac{a}{2} \int u^2 \, \mathrm{d}x + \frac{b}{4} \int u^4 \, \mathrm{d}x \tag{66}$$

with $a+b=1$, $0 \le a \le 1$, so that the special case with $a=1$ and $b=0$ coincides with the Kruskal-Zabusky scheme. These discrete approximations are built through the elegant systematic procedure of Tadmor (Theorem 4.1 in [23]) which has some of its ancestral roots in earlier work by Osher and the second author on nonlinear stability and entropy inequalities [15, 16]. We utilize careful numerical experiments to analyze the statistical behavior of the difference schemes with the conserved quantities in (62) and (66). The results reported below for $b > 0$ contrast completely with those in (64) and (65) for the Kruskal-Zabusky scheme with $b = 0$. In fact, the statistical behavior of the difference schemes conserving (66) reported below is the following:

> *For $b > 0$, the difference scheme has statistical behavior that is ergodic and mixing but*
>
> (1) *The quadratic energy in (63) in Fourier space is not equipartitioned among the individual Fourier modes*
>
> $$\tag{67}$$
>
> (2) *The probability distribution of values at a single grid point is non-Gaussian with two peaks, increasing in strength as b increases.*
> (3) *For $b \ne 0$, the statistical behavior at different grid points is much more strongly correlated than for $b = 0$ and increasing with b.*

The models presented below should be very useful for predictability studies with non-Gaussian variables [8, 12]. Next, we follow Tadmor [23] and build the discrete approximations for (61) which conserve both (62) and (66).

4.1. Discrete Approximations to the Burgers-Hopf equation Conserving a Nonlinear Energy. The nonlinear functions, $au^2/2 + bu^4/4$, $0 \le a \le 1$, $a + b = 1$, are additional convex conserved quantities for the Burgers-Hopf equation in (61). In Theorem 4.1 of [23], Tadmor gives a recipe for designing a conservative three-point difference scheme which also preserves any prescribed convex entropy for a system of conservation laws. Applying this procedure to the Burgers-Hopf equation with the convex conserved quantities $au^2/2 + bu^4/4$ results in the following family of discrete approximations

$$\frac{\partial u_j}{\partial t} + \left(F_{j+1/2} - F_{j-1/2} \right) = 0 \tag{68}$$

with the flux functions, $F_{j+1/2}$ and $F_{j-1/2}$ given by

$$F_{j+1/2}(u_{j+1}, u_j) = \frac{1}{6\Delta x} \frac{a(u_{j+1}^2 + u_{j+1}u_j + u_j^2)}{a + b(u_{j+1}^2 + u_{j+1}u_j + u_j^2)} +$$
$$+ \frac{3}{10\Delta x} \frac{b(u_{j+1}^4 + u_{j+1}^3 u_j + u_{j+1}^2 u_j^2 + u_{j+1}u_j^3 + u_j^4)}{a + b(u_{j+1}^2 + u_{j+1}u_j + u_j^2)}, \tag{69a}$$

$$F_{j-1/2} = F_{j+1/2}(u_j, u_{j-1}). \tag{69b}$$

By design, this difference scheme conserves both the discrete linear momentum,

$$\sum_j u_j, \tag{70}$$

and the nonlinear discrete energy,

$$\sum_j \left(\frac{1}{2} a u_j^2 + \frac{1}{4} b u_j^4 \right). \tag{71}$$

While (70) is obvious, it is a tedious exercise for the reader to check explicitly that (71) is conserved by (68) with (69a) and (69b). Here and below, we consider solutions of (68) with discrete periodic boundary conditions on the interval $[0, 2\pi]$. Thus, we assume that there are $2N + 1$ distinct discrete points with $(2N + 1)\Delta x = 2\pi$ and discrete values $u_j(t) = u(j\Delta x, t)$ where 2π-periodicity requires

$$u_{j+2N+1} = u_j. \tag{72}$$

Unlike the Kruskal-Zabusky scheme, *the difference approximations in (68) do not satisfy the Liouville property* [12, 17, 18, 22] *in the u_j-coordinates.* It is an interesting mathematical problem to determine if there is a nonlinear change of coordinates so that the equations transformed from (68) satisfy the Liouville property. If this is true the invariant probability measure for (68) with $b > 0$ would be explicit and allow for statistical predictions which might explain the behavior in (67) for $b > 0$ which we establish below. Also, are there difference schemes which conserve (71) and directly satisfy the Liouville property?

4.2. Numerical experiments for the Statistical Behavior. In this section we present the results of the numerical study of the finite difference scheme (68) with (69a) and (69b). The numerical experiments are performed for the three different regimes of the numerical scheme – $\{a = 1, b = 0\}$, $\{a = 0.5, b = 0.5\}$, and $\{a = 0, b = 1\}$. For the time integration we use the standard fourth order Runge-Kutta integrator with the discrete time step $\Delta t = 2.5 \cdot 10^{-4}$. The numerical scheme (68) is used with the $N = 200$ discretization points in physical space. The initial data for the numerical simulations is $u_j(0) = 2\sin(2\pi j/N)$. The overall time used for statistical computations is $T = 10000$ units. During the computations we check that the relative error in the conservation of both the momentum (62) and the nonlinear energy (66) does not exceed 10^{-6} at any time of integration for all $\{a, b\}$ regimes presented below. In contrast to the situation in Section 2, no symplectic integrator is needed here to track the conservation of (71) accurately [2, 17, 18].
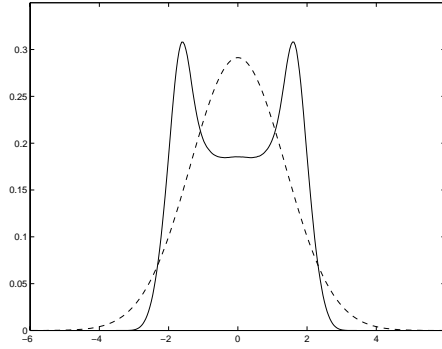
Probability density functions. First we present the probability density functions of the numerical solutions of (48) in physical space, for the three numerical regimes: $\{a = 1, b = 0\}$, $\{a = 0.5, b = 0.5\}$, and $\{a = 0, b = 1\}$. The probability density functions (PDFs) are computed using the "bin counting" procedure, which is described in [2]. The results are presented on the Figure 7. The equilibrium statistical theory [18] for the Kruskal-Zabusky scheme predicts Gaussian probability density functions. Each picture in Figure 7 features the numerically computed PDF (solid line) and the analytical Gaussian PDF with the same mean and variance (dashed line). We can see that for (68) in the case $\{a = 1, b = 0\}$ (Kruskal-Zabusky finite difference scheme, Liouville property satisfied) the numerically computed probability density function coincides with its Gaussian fit. However, the PDFs for the other
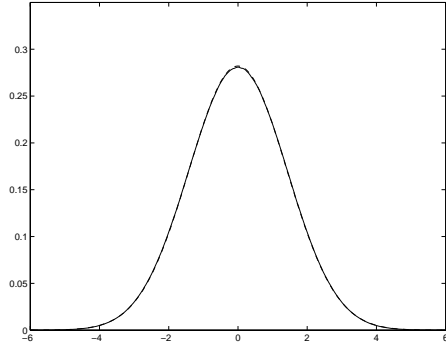
The numerical PDF (solid line) and the analytical
Gaussian (dashed line) for the regime $\{a = 1, b = 0\}$.
The PDF mean is $-5.647 \cdot 10^{-7}$ and variance is 2
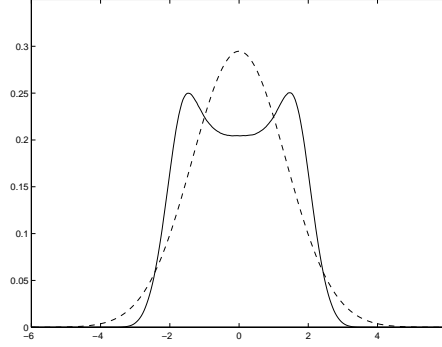
The numerical PDF (solid line) and the analytical
Gaussian (dashed line) for the regime
$\{a = 0.5, b = 0.5\}$. The PDF mean is
$-3.808 \cdot 10^{-7}$ and variance is 1.832

The numerical PDF (solid line) and the analytical
Gaussian (dashed line) for the regime $\{a = 0, b = 1\}$.
The PDF mean is $1.523 \cdot 10^{-7}$ and variance is 1.875

FIG. 7. *The probability density functions for the different numerical regimes of (68): $\{a = 1, b = 0\}$ (first picture), $\{a = 0.5, b = 0.5\}$ (second picture), $\{a = 0, b = 1\}$ (third picture). Solid line – numerical PDF, dashed line – analytical Gaussian with the same mean and variance. The numerical PDF and analytical Gaussian coincide on the first picture.*

regimes are strongly non-Gaussian, and each PDF possesses two distinct peaks which increase with b. While the means of the probability density functions stay very small for all regimes, the variances for the non-Gaussian regimes are slightly less than that for the Gaussian regime (1.832 and 1.875 versus 2, relative difference 6.25-8.4%).

Spatial correlations. Here we present the spatial correlation functions for the numerical scheme (68) with (69a) and (69b). It is clear that all the physical nodes are statistically equivalent due to the translational invariance of the numerical scheme in physical space, however it is unclear how fast the solutions decorrelate in the neighboring physical space nodes. The spatial correlation functions

$$C(j\Delta x) = \frac{\int_{T_0}^{T_0+T} \sum_{k=1}^{N} u(k\Delta x, t)u((k+j)\Delta x, t)dt}{\int_{T_0}^{T_0+T} \sum_{k=1}^{N} [u(k\Delta x, t)]^2 dt} \tag{73}$$

Node, j	Absolute values of spatial correlations		
	$\{a = 1, b = 0\}$	$\{a = 0.5, b = 0.5\}$	$\{a = 0, b = 1\}$
0	1	1	1
1	0.007	0.03	0.06
2	0.005	0.08	0.08
3	0.005	0.05	0.06
4	0.005	0.035	0.05
5	0.005	0.015	0.02
6	0.004	0.01	0.02
7	0.005	0.005	0.009

TABLE 1
The absolute values of the spatial correlation functions for different regimes of the numerical scheme

show the extent to which the physical space nodes, separated by $j\Delta x$, are correlated. In Figure 8 we can see how differently the spatial correlations behave for the different regimes: for the regime $\{a = 1, b = 0\}$ the physical space nodes are completely decorrelated (or δ-correlated), for the regime $\{a = 0.5, b = 0.5\}$ the spatial correlations extend up to 6 physical space nodes, and for the regime $\{a = 0, b = 1\}$ the spatial correlations extend up to 9 physical space nodes. Table 1 shows the absolute values of the spatial correlation function for several physical space nodes for all regimes. We observe that the absolute values of spatial correlations for the finite difference scheme (68) with $\{a = 0.5, b = 0.5\}$ and $\{a = 0, b = 1\}$, for the neighboring physical space nodes can be an order of magnitude greater than those for the Kruskal-Zabusky finite-difference scheme.

Energy partition and mixing. The equilibrium statistical theory for the Kruskal-Zabusky finite difference scheme predicts the equipartition of the energy (63) in Fourier space:

$$E_k = \langle |\hat{u}_k|^2 \rangle = \frac{2E}{N}, \tag{74}$$

where E_k denotes the energy per Fourier mode, and $\langle \cdot \rangle$ is time-averaging. Here we observe the non-equipartition of the energy (63) for the regimes $\{a = 0.5, b = 0.5\}$ and $\{a = 0, b = 1\}$ (*which is not a conserved quantity for these regimes*), shown on the Figure 9. Each of the two pictures on the Figure 9 features the numerically measured partition of the energy (63) (circles), and the equipartition of energy predicted by the equilibrium statistical theory for the Kruskal-Zabusky finite difference scheme. As we can see, the partition of the energy (63) for the regimes $\{a = 0.5, b = 0.5\}$ and $\{a = 0, b = 1\}$ possesses non-uniform structure, peaked at the Fourier wavenumbers $k = \pm 70$ for $\{a = 0.5, b = 0.5\}$ and $k = \pm 75$ for $\{a = 0, b = 1\}$, and having the smallest value at the Fourier wavenumbers $k = \pm 100$ for both regimes (we do not mention the wavenumber $k = 0$ which does not contribute to the dynamics and is conserved). We also observe that the non-uniform energy structure retains the general shape as a and b vary, and the general trend is that the non-equipartition of energy is more severe for greater b.

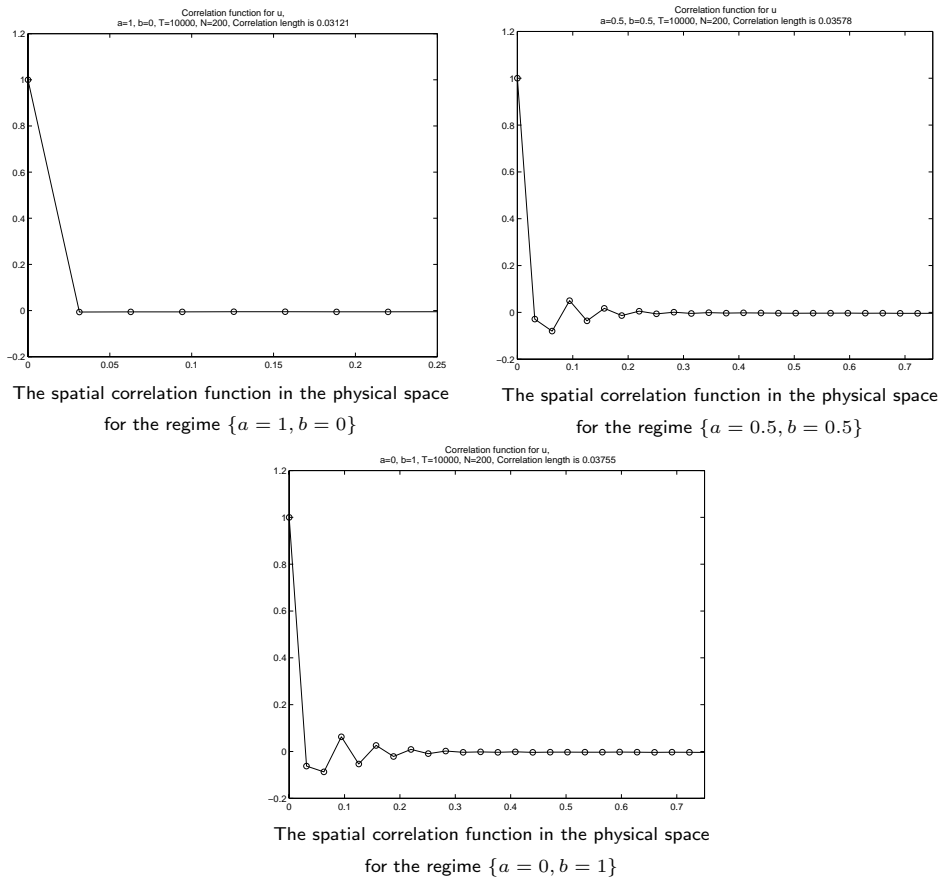Additionally, we check the correlation time scaling for the Fourier modes. The

The spatial correlation function in the physical space
for the regime $\{a = 1, b = 0\}$

The spatial correlation function in the physical space
for the regime $\{a = 0.5, b = 0.5\}$

The spatial correlation function in the physical space

for the regime $\{a = 0, b = 1\}$

FIG. 8. *The spatial correlations for the different numerical regimes of (68):* $\{a = 1, b = 0\}$ *(first picture),* $\{a = 0.5, b = 0.5\}$ *(second picture),* $\{a = 0, b = 1\}$ *(third picture). The solutions in the physical space nodes for the Kruskal-Zabusky scheme are δ-correlated*

The energy partition in the Fourier space for the
regime $\{a = 0.5, b = 0.5\}$, circles – numerical
simulation, dashed line – Kruskal-Zabusky
equipartition prediction for the same initial data

The energy partition in the Fourier space for the
regime $\{a = 0, b = 1\}$, circles – numerical simulation,
dashed line – Kruskal-Zabusky equipartition prediction
for the same initial data
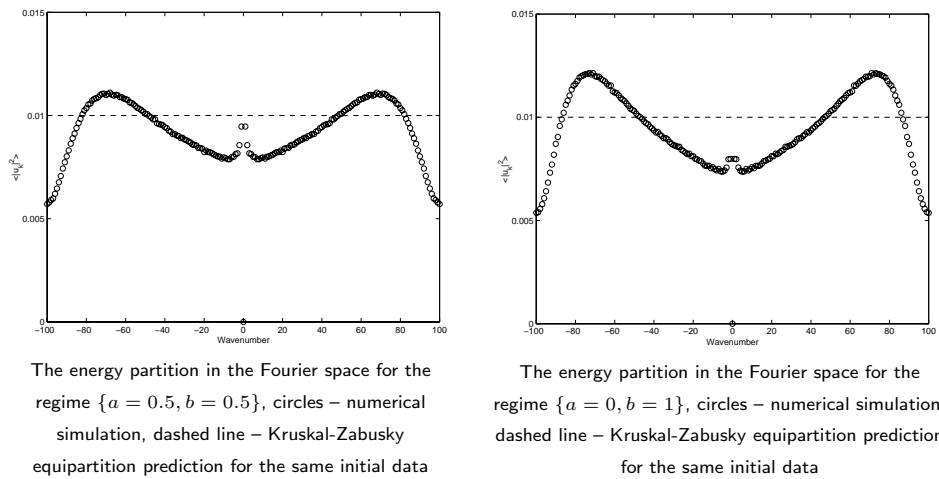
FIG. 9. *The energy partition for the Fourier modes, regimes* $\{a = 0.5, b = 0.5\}$ *and* $\{a = 0, b = 1\}$.

The spatial time correlation scaling in the Fourier space for the regime $\{a = 0.5, b = 0.5\}$

The spatial time correlation scaling in the Fourier space for the regime $\{a = 0, b = 1\}$
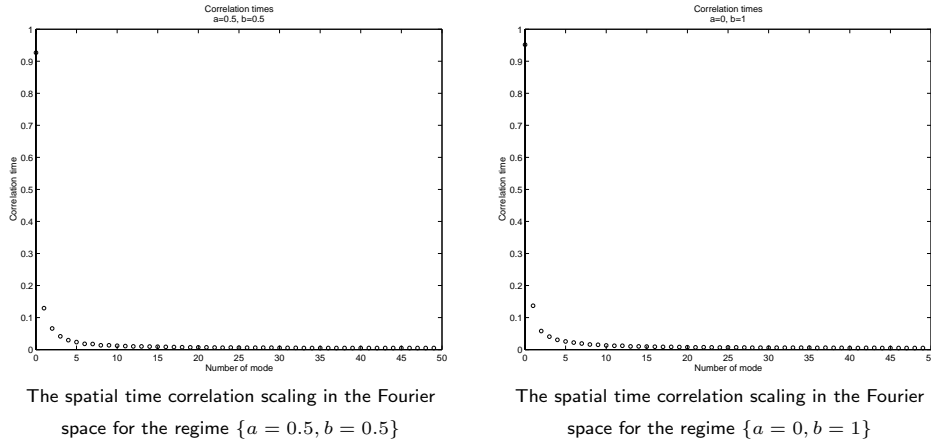
FIG. 10. *The time correlation scaling for the Fourier modes, regimes* $\{a = 0.5, b = 0.5\}$ *and* $\{a = 0, b = 1\}$.

time correlation function of a Fourier mode $\hat{u}_k$ is defined by

$$c_k(\tau) = \frac{1}{TE_k} \int_{T_0}^{T_0+T} \hat{u}_k(t)\hat{u}_k^*(t+\tau)\,\mathrm{d}t, \tag{75}$$

where the *averaging window* T is large, $T = 10000$, and the skipping time $T_0 = 1000$. The correlation time of a Fourier mode $\hat{u}_k$ is defined by

$$T_k = \int |c_k|\,\mathrm{d}\tau. \tag{76}$$

The correlation time scaling for the Galerkin-truncated Burgers-Hopf equation has been studied in [17, 18], and confirmed for the statistically relevant values of Hamiltonian in [2]. The results of the numerical study of the correlation time scaling show non-uniform time correlation structure for the Fourier modes with the large scale Fourier modes possessing longer correlation times. Decay of these correlation functions provides convenient numerical evidence for strong mixing in this dynamical system. Here in Figure 10 we present the correlation time scaling for the numerical scheme (68) for the regimes $\{a = 0.5, b = 0.5\}$ and $\{a = 0, b = 1\}$. Apparently, the different conservation regimes do not visibly affect the time correlation structure in Fourier space. Also we observe the same trend as for the Galerkin-truncated Burgers-Hopf equation, with large scale modes having longer correlation times, which means that the Galerkin truncation and the truncation (68) have similar mixing properties.

5. The Spectral Tilt in the Truncated Burgers-Hopf Model − A Case Study in Statistically Relevant Conserved Quantities. The Fourier-Galerkin truncation of the Burgers-Hopf equation (TBH),

$$(u_\Lambda)_t + \frac{1}{2}P_\Lambda\left(u_\Lambda^2\right)_x = 0, \tag{77}$$

where P_Λ is defined as in (20) with Λ denoting the number of degrees of freedom, has been introduced recently as a prototype model with solutions exhibiting intrinsic stochasticity and a wide range of correlation scaling behavior which can be predicted

successfully by simple scaling arguments [17, 18]. Solutions of the truncated Burgers-Hopf equations possess the three conserved quantities

$$\int u_\Lambda dx = M, \quad \text{Momentum,}$$

$$\int P(u_\Lambda^2) dx = E, \quad \text{Energy,}$$

$$\int P(u_\Lambda^3) dx = H, \quad \text{Hamiltonian.}$$

The equilibrium statistical theory, based of the maximum entropy principle, predicts the equipartition of the energy in the Fourier space under the constraint of fixed energy. However, it was first established by Kovačič and the authors in [2], that the truncated Burgers-Hopf model is a Hamiltonian system with Hamiltonian given by the integral of the third power. This additional conserved quantity, beyond the energy, has been ignored in previous statistical mechanics studies of this equation. The question of the statistical significance of the Hamiltonian, beyond that of the energy, has been studied in [2]. For statistically relevant values of the Hamiltonian, the direct numerical simulations showed a surprising spectral tilt rather than equipartition of energy. This spectral tilt was predicted and confirmed independently by Monte-Carlo simulations based on equilibrium statistical mechanics together with a heuristic formula for the spectral tilt,

$$\frac{1}{2}\langle|\hat{u}_k|^2\rangle = \frac{E}{\Lambda}\left(1 + \frac{8H^2}{5E^3}\frac{(\Lambda+1)/2 - k}{\Lambda}\right), \tag{78}$$

where Λ denotes the number of degrees of freedom, E is energy, and H is Hamiltonian.

However, Kovačič and the authors in [2] studied the problem when the dimensionless ratio H^2/E^3 was very small, so a linear fit like the one in (78) could be expected. Here we show that such a spectral tilt remains linear as long as the statistical theory is itself valid, as we increase the values of the cubic Hamiltonian in the TBH, while keeping the energy fixed. In the current study we address four distinct objectives:

  (i) The values of the Hamiltonian for which the spectral tilt becomes nonlinear;
 (ii) The values of the Hamiltonian for which the formula in (78) for the tilt fails;
(iii) We establish that the failure of strong mixing (i.e. the failure of exponential decay of correlations) occurs whenever (i) takes place, which means that the spectral tilt is linear whenever the direct numerical simulations yield a mixed statistical equilibrium, and vice versa;
 (iv) We also show that the failure of the spectral tilt formula (78) occurs just before strong mixing fails, as the Hamiltonian increases.

We performed the direct numerical simulations for the fixed energy $E = 0.1$, the truncation sizes ranging from $\Lambda = 10$ to $\Lambda = 100$, and the values of the Hamiltonian ranging from $H = 0.01$ to $H = 0.15$. The typical integration time in the numerical simulations was $T = 20000$, and the discrete integration time step was $\Delta t = 2 \cdot 10^{-4}$. The relative errors in the conservation of the energy and the Hamiltonian did not exceed the values of 10^{-12} and 10^{-11}, respectively, for all attempted numerical simulations at any times of integration. We show the figures with the spectral tilt and the time correlation mixing (75) for the truncation sizes $\Lambda = 20$ and $\Lambda = 100$. Table 2 contains the summary of the data from all experiments.

Spectral Tilt, Linearity, Fit, and Mixing				
Λ	H^2/E^3	Linear	Fit (78)	Mixing
10	0.1	Yes	Yes	Good
10	0.4	Yes	No	Marginal
10	1.6	No	No	Fails
20	0.1	Yes	Yes	Good
20	0.4	Yes	No	Marginal
20	1.6	No	No	Fails
20	4.9	No	No	Fails
50	0.1	Yes	Yes	Good
50	0.169	Yes	Yes	Good
50	0.289	Yes	No	Good
50	0.4	Yes	No	Marginal
50	1.6	No	No	Fails
50	4.9	No	No	Fails
50	10	No	No	Fails
100	0.1	Yes	Yes	Good
100	0.4	Yes	No	Marginal
100	1.6	No	No	Fails
100	4.9	No	No	Fails
100	10	No	No	Fails
100	22.5	No	No	Fails

TABLE 2

The energy-Hamiltonian H^2/E^3 ratio, tilt linearity, prediction, and mixing for the truncation sizes $\Lambda = 10, 20, 50,$ and 100

5.1. The Linear Spectral Tilt and the Failure of Strong Mixing for Large Values of the Hamiltonian. Here we concentrate on the connection of the nonlinearity in the spectral tilt and the failure of mixing. We present the figures with the spectral tilts and the time correlation functions of the largest Fourier mode for the two cases $\Lambda = 20$ and $\Lambda = 100$.

For the case with $\Lambda = 20$ we observe the following trends. We begin with the small value of the Hamiltonian $H = 0.01$, and for this small value the tilt is linear and the mixing is strong, as shown in Figure 11 and Table 2, where all time correlations decay very quickly, within time $T = 40$. For the next, larger value of the Hamiltonian, $H = 0.02$, the spectral tilt is still linear, and the mixing is still good, as shown in Figure 12 and Table 2, however the time correlations in the first Fourier mode become somewhat longer (never decay to zero within time $T = 200$). For the values of the Hamiltonian $H = 0.04$ and $H = 0.07$ the spectral tilt is nonlinear and the failure of mixing is observed, as shown in Figures 13, 14 and Table 2.

For the case with $\Lambda = 100$ we observe similar trends. We begin with the small value of the Hamiltonian $H = 0.01$, and for this small value the tilt is linear and the mixing is strong; as shown in Figure 15 and Table 2, all time correlations decay very quickly, within time $T = 20$. For the next, larger value of the Hamiltonian, $H = 0.02$, the spectral tilt is still linear, and the mixing is still good, as shown in Figure 12 and Table 2, however the tine correlations in the first Fourier mode become somewhat longer (within time $T = 100$). For the values of the Hamiltonian $H = 0.04, 0.07, 0.1,$ and 0.15 the spectral tilt is nonlinear and the failure of mixing is full-scale, as shown

in Figures 16, 17, 18, 19, 20 and Table 2.

As we can see, the nonlinearity in the spectral tilt occurs whenever mixing fails. Since the linear spectral tilt is predicted by the equilibrium statistical mechanics, and the failure of mixing draws a bound of validity for the equilibrium statistical mechanics, we conclude that the linearity of the spectral tilt is valid whenever the statistical predictions are applicable. We do not present the plots for the cases with $\Lambda = 10$ and $\Lambda = 50$, but the trends for them are very similar to the cases with $\Lambda = 20$ and $\Lambda = 100$, and are summarized in Table 2.

5.2. The Validity of the Heuristic Formula Predicting the Spectral Tilt.
Above we established the direct logical connection between the linearity of the spectral tilt, and the applicability of equilibrium statistical mechanics which predicts this tilt. However, the next question we address here is the bounds of applicability of the heuristic formula (78) for the spectral tilt. The formula in (78) was valid for all values of Hamiltonian studied in [2], however $H = 0.01$ was the largest value considered there. Here we start with the largest Hamiltonian considered in [2], $H = 0.01$, and increase its value. We find that this value of the Hamiltonian, $H = 0.01$, is exactly where the bound of applicability of the heuristic spectral tilt formula (78) lies. In the case with $\Lambda = 20$ we can see that for the Hamiltonian $H = 0.01$, the spectral tilt formula (78) is the perfect match. Figure 11 shows that the predicted tilt (dot-dashed line) coincides with the solid line of least squares fit to the numerical data. However, as the Hamiltonian increases, the predicted tilt slope is steeper than the actual (dot-dashed line goes steeper than the solid line). The difference between the actual and predicted tilt slopes becomes more severe as the value of the Hamiltonian increases (Figures 12, 13, and 14). The same tilt failure of the predicted linear tilt is observed for the truncation size $\Lambda = 100$. The predicted tilt coincides with the actual one for the Hamiltonian $H = 0.01$, and fails to predict anything for any other value of the Hamiltonian above $H = 0.01$ (Figures 15, 16, 17, 18, 19, and 20). We do not plot the results for the rest of the truncation sizes, instead summarizing them in Table 2. Overall we can see that the linearity of the tilt and mixing generally last longer than the tilt slope prediction (up to $H = 0.02$ and $H = 0.01$, respectively).
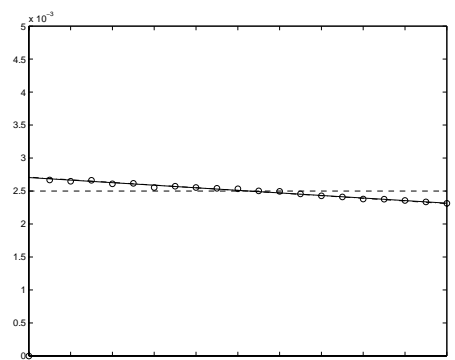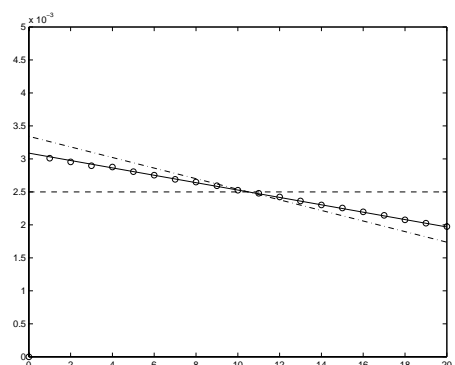
The energy spectrum for the simulation with $\Lambda = 20$, $H = 0.01$. Solid line − linear fit found by least squares, dash-dotted line − linear fit (78), dashed line shows equipartition
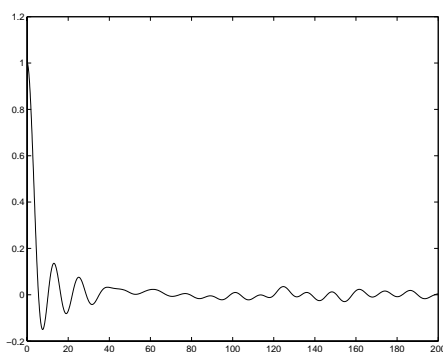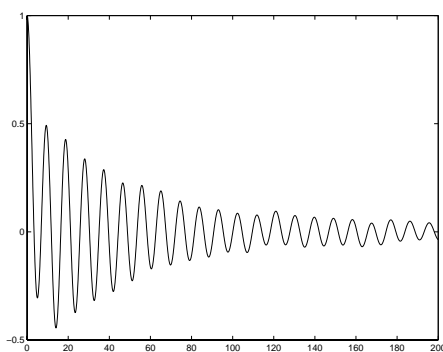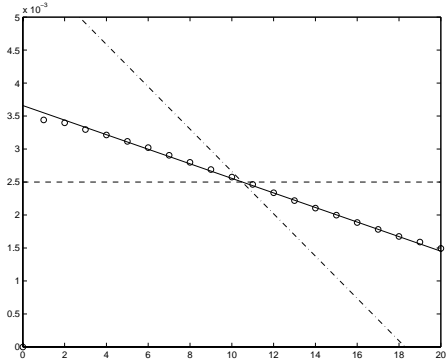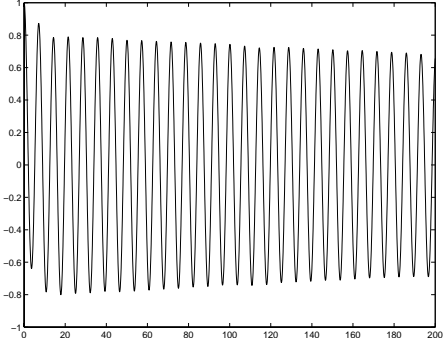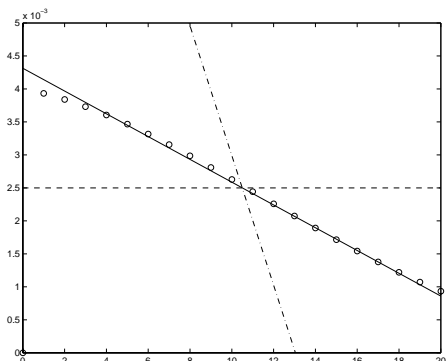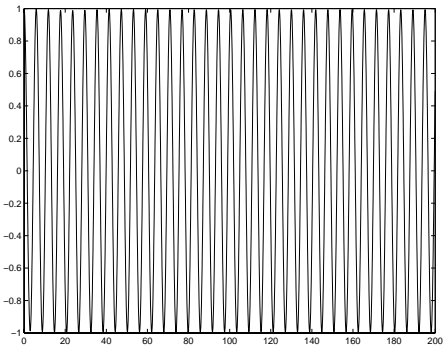
The time correlation function for the Fourier mode $k = 1$, $\Lambda = 20$, $H = 0.01$

FIG. 11. *The spectral tilt and the time correlation function for the case $\Lambda = 20$, $H = 0.01$. The correlation function confirms mixing, the spectral tilt is linear, and the predicted tilt coincides with the actual slope*

The energy spectrum for the simulation with $\Lambda = 20$, $H = 0.02$. Solid line − linear fit found by least squares, dash-dotted line − linear fit (78), dashed line shows equipartition

The time correlation function for the Fourier mode $k = 1$, $\Lambda = 20$, $H = 0.02$

FIG. 12. *The spectral tilt and the time correlation function for the case $\Lambda = 20$, $H = 0.02$. The weak mixing is still present, the spectral tilt is still linear, but the predicted tilt is steeper than the actual slope*

The energy spectrum for the simulation with $\Lambda = 20$, $H = 0.04$. Solid line — linear fit found by least squares, dash-dotted line — linear fit (78), dashed line shows equipartition

The time correlation function for the Fourier mode $k = 1$, $\Lambda = 20$, $H = 0.04$

FIG. 13. *The spectral tilt and the time correlation function for the case $\Lambda = 20$, $H = 0.04$. Long correlations show failure of mixing, the spectral tilt is nonlinear and the predicted tilt is much more steeper than the actual slope*

The energy spectrum for the simulation with $\Lambda = 20$, $H = 0.07$. Solid line — linear fit found by least squares, dash-dotted line — linear fit (78), dashed line shows equipartition

The time correlation function for the Fourier mode $k = 1$, $\Lambda = 20$, $H = 0.07$

FIG. 14. *The spectral tilt and the time correlation function for the case $\Lambda = 20$, $H = 0.07$. Long correlations show failure of mixing, the spectral tilt is nonlinear and the predicted tilt is much more steeper than the actual slope*

The energy spectrum for the simulation with $\Lambda = 100$,
$H = 0.01$. Solid line – linear fit found by least
squares, dash-dotted line – linear fit (78), dashed line
shows equipartition

The time correlation function for the Fourier mode
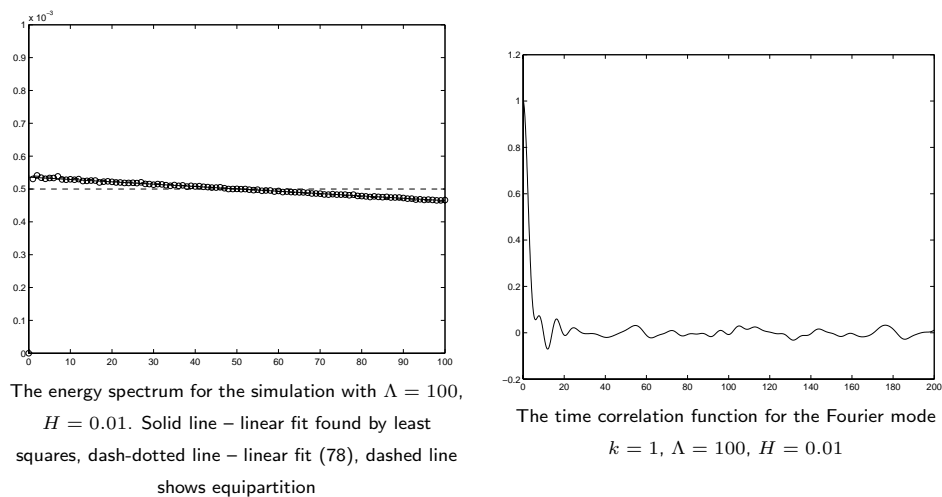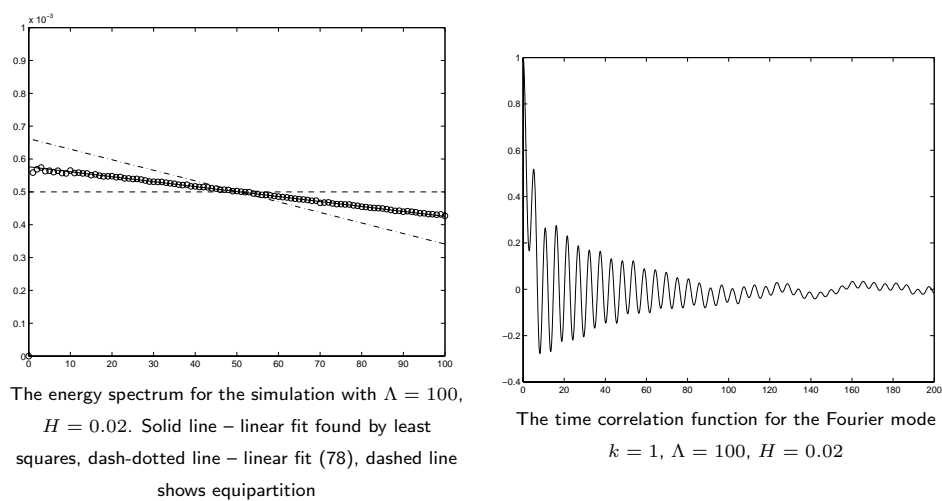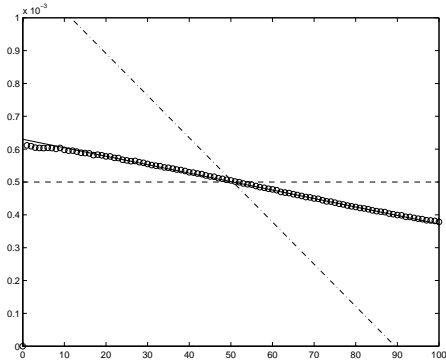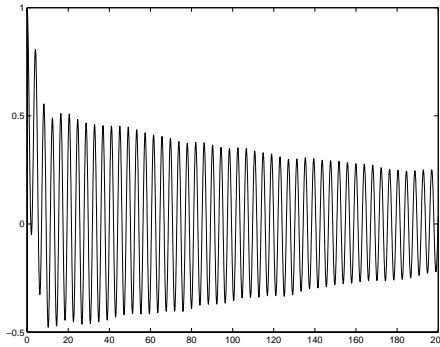$k = 1$, $\Lambda = 100$, $H = 0.01$

FIG. 15. *The spectral tilt and the time correlation function for the case $\Lambda = 100$, $H = 0.01$. The correlation function confirms mixing, the spectral tilt is linear, and the predicted tilt coincides with the actual slope*

The energy spectrum for the simulation with $\Lambda = 100$,
$H = 0.02$. Solid line – linear fit found by least
squares, dash-dotted line – linear fit (78), dashed line
shows equipartition

The time correlation function for the Fourier mode
$k = 1$, $\Lambda = 100$, $H = 0.02$

FIG. 16. *The spectral tilt and the time correlation function for the case $\Lambda = 100$, $H = 0.02$. The weak mixing is still present, the spectral tilt is still linear, but the predicted tilt is steeper than the actual slope*
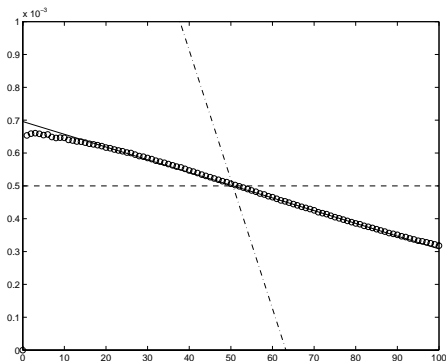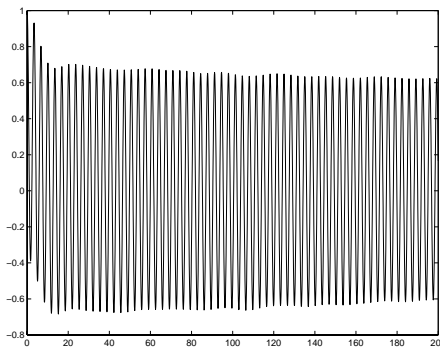
The energy spectrum for the simulation with $\Lambda = 100$, $H = 0.04$. Solid line – linear fit found by least squares, dash-dotted line – linear fit (78), dashed line shows equipartition

The time correlation function for the Fourier mode $k = 1$, $\Lambda = 100$, $H = 0.04$

FIG. 17. *The spectral tilt and the time correlation function for the case $\Lambda = 100$, $H = 0.04$. Long correlations show failure of mixing, the spectral tilt is nonlinear and the predicted tilt is much more steeper than the actual slope*
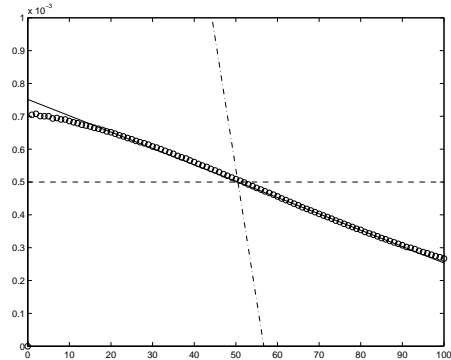

The energy spectrum for the simulation with $\Lambda = 100$, $H = 0.07$. Solid line – linear fit found by least squares, dash-dotted line – linear fit (78), dashed line shows equipartition
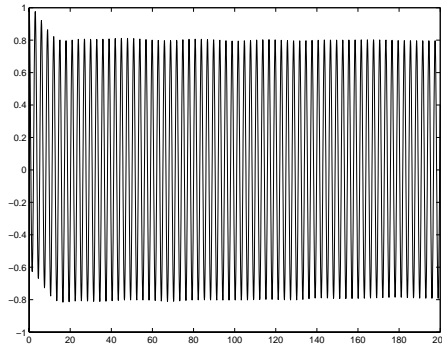
The time correlation function for the Fourier mode $k = 1$, $\Lambda = 100$, $H = 0.07$

FIG. 18. *The spectral tilt and the time correlation function for the case $\Lambda = 100$, $H = 0.07$. Long correlations show failure of mixing, the spectral tilt is nonlinear and the predicted tilt is much more steeper than the actual slope*
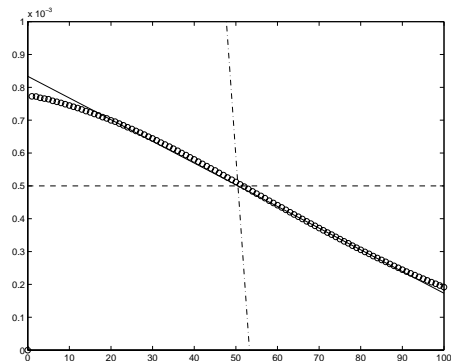
The energy spectrum for the simulation with $\Lambda = 100$, $H = 0.1$. Solid line − linear fit found by least squares, dash-dotted line − linear fit (78), dashed line shows equipartition
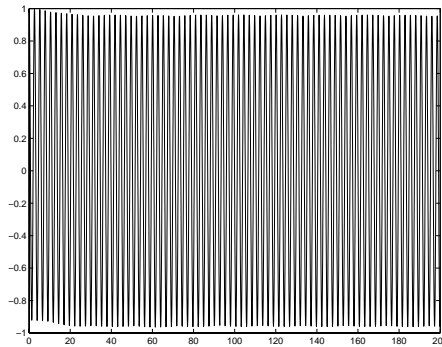
The time correlation function for the Fourier mode $k = 1$, $\Lambda = 100$, $H = 0.1$

FIG. 19. *The spectral tilt and the time correlation function for the case $\Lambda = 100$, $H = 0.1$. Long correlations show failure of mixing, the spectral tilt is nonlinear and the predicted tilt is much more steeper than the actual slope*

The energy spectrum for the simulation with $\Lambda = 100$, $H = 0.15$. Solid line − linear fit found by least squares, dash-dotted line − linear fit (78), dashed line shows equipartition

The time correlation function for the Fourier mode $k = 1$, $\Lambda = 100$, $H = 0.15$

FIG. 20. *The spectral tilt and the time correlation function for the case $\Lambda = 100$, $H = 0.15$. Long correlations show failure of mixing, the spectral tilt is nonlinear and the predicted tilt is much more steeper than the actual slope*

REFERENCES

[1] R. ABRAMOV, *Statistically Relevant and Irrelevant Conserved Quantities for the Equilibrium Statistical Description of the Truncated Burgers-Hopf Equation and the Equations for Barotropic Flow*, PhD thesis, RPI, June 2002. Supervised by A. Majda and G. Kovačič.

[2] R. ABRAMOV, G. KOVAČIČ, AND A. MAJDA, *Hamiltonian Structure and Statistically Relevant Conserved Quantities for the Truncated Burgers-Hopf Equation*, Comm. Pure Appl. Math., (2002).

[3] R. ABRAMOV AND A. MAJDA, *Statistically relevant conserved quantities for truncated quasi-geostrophic flow*, Proc. Natl. Acad. Sci., (2003), in preparation.

[4] G. F. CARNEVALE AND J. S. FREDERIKSEN, *Nonlinear stability and statistical mechanics of flow over topography*, Journal of Fluid Mechanics, 175 (1987), pp. 153–181.

[5] W. E AND C. W. SHU, *Small scale structures in Boussinesq convection*, Phys. Fluids, 6 (1994), pp. 49–58.

[6] A. HARTEN, J. M. HYMAN, AND P. D. LAX, *On finite difference approximations and entropy conditions for shocks*, Comm. Pure Appl. Math., 29 (1976), pp. 297–322.

[7] G. HOLLOWAY, *Eddies, waves, circulation, and mixing: Statistical geofluid mechanics*, Annu. Rev. Fluid Mech., 18 (1986), pp. 49–65.

[8] R. KLEEMAN, A. MAJDA, AND I. TIMOFEYEV, *Quantifying predictability in a model with statistical features of the atmosphere*, Proc. Nat. Acad. Sci., (2002), in press.

[9] P. D. LAX, *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*, volume 11, SIAM Regional Conference Lectures in Applied Mathematics, 1972.

[10] A. MAJDA AND A. BERTOZZI, *Vorticity and Incompressible Flow*, Cambrige University Press, 2002.

[11] A. MAJDA, B. TURKINGTON, K. HAVEN, AND M. DIBATTISTA, *Statistical equilibrium predictions of jets and spots on Jupiter*, Proc. Natl. Acad. Sci., 98 (2001), pp. 12346–12350.

[12] A. MAJDA AND X. WANG, *Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flow*, Cambridge University Press, 2003, in press.

[13] A. J. MAJDA, J. MCDONOUGH, AND S. OSHER, *The Fourier method for non-smooth initial data*, Math. Comp., 32 (1978), pp. 1041–1081.

[14] A. J. MAJDA AND S. OSHER, *Propagation of error into regions of smoothness for accurate difference approximations to hyperbolic equations*, Comm. Pure Appl. Math., 30 (1977), pp. 671–705.

[15] A. J. MAJDA AND S. OSHER, *A systematic approach for correcting nonlinear instabilities*, Numer. Math., 30 (1978), pp. 429–452.

[16] A. J. MAJDA AND S. OSHER, *Numerical viscosity and the entropy condition*, Comm. Pure Appl. Math., 32 (1979), pp. 1–42.

[17] A. J. MAJDA AND I. TIMOFEYEV, *Remarkable statistical behavior for truncated Burgers-Hopf dynamics*, Proc. Natl. Acad. Sci. USA, 97:23 (2000), pp. 12413–12417.

[18] A. J. MAJDA AND I. TIMOFEYEV, *Statistical mechanics for truncations of the Burgers-Hopf equation: a model for intrinsic stochastic behavior with scaling*, Milan Journal of Mathematics, 70 (2002), pp. 39–96.

[19] R. I. MCLACHLAN, *Explicit Lie-Poisson integration and the Euler equations.*, Phys. Rev. Lett., 71 (1993), pp. 3043–3046.

[20] J. MILLER, P. WEICHMAN, AND M. C. CROSS, *Statistical mechanics, Euler's equations, and Jupiter's Red Spot*, Phys. Rev., A45 (1992), pp. 2328–2359.

[21] P. J. MORRISON, *Hamiltonian description of the ideal fluid*, Reviews of Modern Physics, 70 (1998), pp. 467–521.

[22] R. SALMON, *Lectures on Geophysical Fluid Dynamics*, Oxford Press, New York, 1998.

[23] E. TADMOR, *The numerical viscosity of entropy stable schemes for systems of conservation laws*, Math. Comp., 49 (1987), pp. 91–103.

[24] B. TURKINGTON, *Statistical equilibrium measures and coherent states in two-dimensional turbulence*, Comm. Pure Appl. Math., 52 (1999), pp. 781–809.

[25] V. ZEITLIN, *Finite-mode analogues of 2D ideal hydrodynamics: Coadjoint orbits and local canonical structure*, Physica D, 49 (1991), pp. 353–362.