Commun. Math. Phys. 244, 187–208 (2004)
Digital Object Identifier (DOI) 10.1007/s00220-003-0978-2

Communications in
**Mathematical
Physics**

A Statistical Approach to the Asymptotic Behavior of a Class of Generalized Nonlinear Schrödinger Equations

Richard S. Ellis[1,⋆], **Richard Jordan**[2,⋆⋆], **Peter Otto**[1,⋆], **Bruce Turkington**[1,⋆⋆⋆]

[1] Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003, USA.
E-mail: rsellis@math.umass.edu; otto@math.umass.edu; turk@math.umass.edu
[2] Dynamics Technology, Inc., 1555 Wilson Boulevard, Suite 320, Arlington, VA 22209, USA.
E-mail: rjordan@dynatec.com

Received: 4 April 2003 / Accepted: 23 July 2003
Published online: 13 November 2003 –

Abstract: A statistical relaxation phenomenon is studied for a general class of dispersive wave equations of nonlinear Schrödinger-type which govern non-integrable, non-singular dynamics. In a bounded domain the solutions of these equations have been shown numerically to tend in the long-time limit toward a Gibbsian statistical equilibrium state consisting of a ground-state solitary wave on the large scales and Gaussian fluctuations on the small scales. The main result of the paper is a large deviation principle that expresses this concentration phenomenon precisely in the relevant continuum limit. The large deviation principle pertains to a process governed by a Gibbs ensemble that is canonical in energy and microcanonical in particle number. Some supporting Monte-Carlo simulations of these ensembles are also included to show the dependence of the concentration phenomenon on the properties of the dispersive wave equation, especially the high frequency growth of the dispersion relation. The large deviation principle for the process governed by the Gibbs ensemble is based on a large deviation principle for Gaussian processes, for which two independent proofs are given.

1. Introduction

Many dynamical models of physical systems governed by nonlinear partial differential equations exhibit a typical long-time behavior in which coherent structures organize on the large spatial scales while turbulent fluctuations dominate the small scales [19]. Perhaps the most familiar setting for this behavior is two-dimensional or quasi-geostrophic

⋆ This research was supported in part by grants from the Department of Energy (DE-FG02-99ER25376) and from the National Science Foundation (NSF-DMS-0202309).

⋆⋆ This research was partially supported by a Mathematical Sciences Postdoctoral Research Fellowship from the National Science Foundation.

⋆⋆⋆ This research was supported in part by grants from the Department of Energy (DE-FG02-99ER25376) and from the National Science Foundation (NSF-DMS-0207064).

fluid turbulence, where the coherent structures are large-scale steady motions, such as shear flows or vortices, and the turbulent background is a vorticity field that fluctuates on the small scales [12, 32]. The zonal jets and embedded vortical spots in the active weather layer of Jupiter are especially persistent and conspicuous examples of this phenomenon [13, 29]. These coherent structures have been shown to be realizable as the equilibrium states in a statistical model of the geophysical fluid dynamical system [8, 33]. Another physical system exhibiting this behavior is two-dimensional magnetohydrodynamics. Long-time numerical simulations of MHD turbulence show that the systems end in states in which the magnetic and velocity fields fluctuate on small scales around a steady mean state on the large scales [4]. Equilibrium statistical models have also been able to capture these coherent structures [20, 24]. This statistical relaxation phenomenon is shared by certain dispersive wave systems, for which the generic coherent structures are solitary waves that interact with a disorganized background of wave radiation [9, 10, 15, 34]. For instance, non-integrable, focusing, nonlinear Schrödinger equations in a bounded domain organize after a long evolution into a single solitary wave coupled with small-scale fluctuations [22, 23]. This self-organization behavior has been shown to be consistent with relaxation to a statistical equilibrium state, both qualitatively and quantitatively [25].

Motivated by this fundamental phenomenon exhibited by many complex systems, we devote the present paper to a detailed analysis of the equilibrium statistical behavior of a particular class of dispersive wave systems. Specifically, we consider a class of generalized nonlinear Schrödinger (GNLS) equations on a bounded domain D in $\mathbb{R}^d$ with appropriate boundary conditions. These systems govern the dynamics of a complex field $\psi(x,t)$, $x \in \mathbb{R}^d$, $t \in \mathbb{R}$, via the equation

$$i\psi_t + L\psi + f(|\psi|^2)\psi = 0. \tag{1.1}$$

L denotes an unbounded linear operator on the complex Hilbert space $L^2_c(\rho)$ of square-integrable functions on D with respect to a measure ρ; $\langle\cdot,\cdot\rangle$ denotes the inner product on $L^2_c(\rho)$. It is assumed that L is symmetric and that the spectrum of $-L$ consists of positive eigenvalues λ_k satisfying $\sum_{k=1}^{\infty} 1/\lambda_k < \infty$. In addition, the corresponding eigenfunctions e_k are assumed to be real functions that form an orthonormal basis of $L^2_c(\rho)$. We choose to focus our analysis on these systems for two reasons. First, they are widely considered to be prototypes for dynamical systems that exhibit organization of coherent structures within turbulence, and accordingly there is a rich literature on their phenomonology [9, 10, 15, 34]. Second, they are simple enough to be amenable to a complete and rigorous analysis by the methods of equilibrium statistical mechanics.

In one space dimension ($d = 1$) the basic example of this class is $L = \partial^2/\partial x^2$ on $D = [0,\ell]$ with Lebesgue measure ρ on $[0,\ell]$ and with homogeneous Dirichlet boundary conditions, where $\ell < \infty$. In this case, we refer to (1.1) as the basic NLS equation. Our analysis also applies to the operator $L = \partial^2/\partial x^2 + U(x)$, where $U(x)$ is a suitable potential, and to other boundary conditions such as ℓ-periodic conditions. For the basic NLS equation and for this wider class of NLS equations in one dimension, the eigenvalues λ_k grow like k^2 as $k \to \infty$.

In (1.1) we restrict our attention to smooth nonlinearities f that satisfy

$$f(0) = 0, \qquad \sup_{a\in[0,\infty)} |f(a)| < \infty\,; \tag{1.2}$$

e.g., $f(|\psi|^2) = b|\psi|^2/(1+|\psi|^2)$ with scale factor b. Nonlinearities with these properties arise in physical applications as large-amplitude corrections to the cubic NLS equation, and they are referred to as bounded or saturated nonlinearities [31]. Our analysis applies to both the focusing GNLS, for which $f'(a) > 0$, and the defocusing GNLS, for which $f'(a) < 0$. Our main interest, however, is on the focusing case since the formation of coherent solitary waves is the dominant mechanism in that case. The restriction to bounded nonlinearities excludes blow-up of solutions and the collapse of waves due to self-focusing. We impose this restriction because our goal is to analyze statistical equilibrium ensembles of regular solutions that model the long-time average behavior of the system. Accordingly, we choose GNLS equations for which solutions exist and are regular for all time.

An interesting generalization included in our analysis is to pseudo-differential operators L whose eigenvalues λ_k grow like k^{α} as $k \to \infty$ with $\alpha > 1$. The GNLS equation (1.1) then resembles the equation introduced by Majda, McLaughlin and Tabak (MMT) in their study of weak turbulence closure theories [28]. From the standpoint of equilibrium statistical mechanics, we are interested in how the phenomenon of concentration into a coherent structure depends on α. In contrast to the MMT equations, we restrict our analysis to bounded nonlinearities in ψ itself; the MMT equations pertain to homogeneous, cubic nonlinearities in $M\psi$, where M is another pseudo-differential operator with eigenvalues that grow like $k^{-\sigma}$. While it would be possible to study a broader class of such equations, the class of GNLS equations (1.1) is sufficiently broad to exhibit the typical behavior of the statistical equilibrium states and to show how this behavior depends upon the linear frequencies of the dispersive wave system.

The object of our analysis is the statistical equilibrium description of the complex dynamical system (1.1) via classical Gibbsian statistics. Our choice of distribution on phase space is a mixed Gibbs ensemble that is canonical with respect to the energy invariant and microcanonical with respect to the particle number invariant. For the GNLS equation (1.1), the Hamiltonian, or energy functional, is

$$H(\psi) \doteq -\frac{1}{2}\langle L\psi, \psi\rangle - \frac{1}{2}\int_D F(|\psi|^2)\, d\rho, \tag{1.3}$$

where F is related to the nonlinearity f by $F(a) = \int_0^a f(s)\, ds$. The particle number, or wave action, is half the $L^2_c(\rho)$-norm squared:

$$Q(\psi) \doteq \frac{1}{2}\int_D |\psi|^2\, d\rho.$$

The resulting statistical description rests on these two exact invariants of the GNLS equation (1.1), together with the conservation of phase volume under the Hamiltonian dynamics. In order to keep our development concise, we intentionally suppress the momentum invariant by breaking the x-translation invariance of the system. To this end, we consider (1.1) in a bounded domain D with homogeneous boundary conditions $\psi = 0$ imposed on ∂D. Alternatively, we could consider an operator L with a potential under periodic boundary conditions to obtain similar results.

While the statistical equilibrium NLS equation has been the focus of several analyses, including [1, 7, 27, 30, 35], our approach and our results differ fundamentally from those investigations. In particular, previous investigators have constructed Gibbs distributions that are Wiener-type measures having infinite mean energy. Our interest, on the other hand, centers on modeling the ensemble-average behavior of regular solutions to (1.1) from initial conditions having given, finite, mean energy $\langle H(\psi^0)\rangle = E$ and given particle number $Q(\psi^0) = N$. Our motivation derives from numerical studies of the underlying GNLS equation which show that from generic initial conditions, such as a field of waves emerging from a modulational instability, the dynamics approximately realize a Gibbs ensemble after a sufficient time [22, 23]. A spectral analysis of these numerical solutions identifies an approximate dimension $n = n(T)$ of the phase space that supports the Gibbs distribution after a long, but finite, time T. Moreover, a continuum limit is achieved as $T \to \infty$, in the sense that $n(T)$ goes to infinity at a definite rate with T. This observed behavior of regular solutions to (1.1) strongly suggests that the relevant continuum limit for a statistical equilibrium theory is the one obtained from the Gibbs states of the spectrally-truncated dynamics on n eigenmodes with fixed mean energy E and fixed particle number N as $n \to \infty$. Accordingly, this limit is the focus of our analysis. While we establish rigorous results about these statistical equilibrium states, we do not address the theoretical problem of proving ergodicity of this dynamics. Rather we accept the ergodic hypothesis on the basis of convincing numerical evidence [22, 23].

Our results pertain to a continuum limit $n \to \infty$ of a sequence of Gibbs distributions on n-dimensional phase spaces corresponding to spectrally-truncated, Hamiltonian dynamics having a finite number of degrees of freedom n. As has been noted elsewhere [25, 27], these Gibbs ensembles are necessarily microcanonical in Q, since a Gibbs canonical ensemble with respect to both H and Q can be divergent for a focusing nonlinearity. We therefore use a mixed ensemble in which the microcanonical condition $Q = N$ is imposed on the canonical distribution in H with an inverse temperature β that is rescaled by n so that the mean energy, $\langle H \rangle = E$, remains finite as $n \to \infty$. The study of the limiting behavior of these mixed ensembles is perfectly suited to analysis by large deviation techniques. Our main result is a large deviation principle demonstrating that the ground-state solitary waves are the most probable macroscopic states in the relevant continuum limit, in the sense that the mixed ensembles concentrate on these ground states in the $L^2_c(\rho)$-norm. This large deviation principle may be considered as a mathematically rigorous statement of the explanation of the observed formation of large-scale coherent structures within small-scale wave turbulence given in [25], where an asymptotically exact mean-field theory was developed and compared with direct numerical simulations.

The outline of the paper is as follows. In Sect.2, we construct the statistical equilibrium model based on a spectral truncation of the GNLS dynamics and motivate the mixed ensemble based upon the invariants H and Q. In Sect.3 we state the main theorem, a large deviation principle for a sequence of finite-dimensional fields with respect to the mixed ensemble introduced in Sect.2. The main theorem is proved in Sect.4 using a basic large deviation theorem for Gaussian processes, for which two independent proofs are given. Both of these proofs require that the linear frequencies λ_k satisfy $\sum_{k=1}^{\infty} 1/\lambda_k < \infty$ (see Cond. 2.1). Finally, in Sect.5 we display the results of some Monte-Carlo simulations of the mixed ensemble in one space dimension. Besides demonstrating the concentration phenomenon numerically when $\sum_{k=1}^{\infty} 1/\lambda_k < \infty$, these simulations exhibit the change in behavior when this growth condition does not hold.

2. Statistical Equilibrium Description of GNLS Dynamics

The GNLS equation (1.1) is considered on a bounded domain D in $\mathbb{R}^d$. The nonlinearity is bounded, in the sense that f satisfies the conditions (1.2). The operator L defining the linear part of the GNLS equation is assumed to satisfy the following condition:

Condition 2.1. *L is a symmetric operator on $L^2_c(\rho)$. The spectrum of $-L$ consists of positive eigenvalues λ_k satisfying $\sum_{k=1}^{\infty} 1/\lambda_k < \infty$. The corresponding eigenfunctions e_k are real functions that form an orthonormal basis of $L^2_c(\rho)$.*

A number of important examples underlie the general theory.

Example 2.2. (a) The basic example is $L = \partial^2/\partial x^2$ on $D = [0, \ell]$ with homogeneous Dirichlet boundary conditions and ρ Lebesgue measure on $[0, \ell]$, where $\ell < \infty$. In this case, for each $k \in \mathbb{N}$ $\lambda_k = (k\pi/\ell)^2$ and $e_k = \sqrt{2/\ell}\sin(k\pi x/\ell)$.

(b) Let p be a $\mathcal{C}^2$ function on $D = [0, \ell]$ satisfying $\inf_{x\in[0,\ell]} p(x) > 0$, q a negative continuous function on $[0, \ell]$, and ρ Lebesgue measure on $[0, \ell]$. For $\xi \in L^2_c(\rho)$ we define

$$L\xi \doteq \frac{d}{dx}\left(p(x)\frac{d\xi}{dx}\right) - q(x)\xi(x)$$

with homogeneous Dirichlet boundary conditions. By standard Sturm-Louiville theory, L satisfies Condition 2.1. As in the basic example given in part (a), the eigenvalues λ_k of L grow like k^2 as $k \to \infty$ [3, Thm. 10.9].

(c) Let D be any bounded domain in $\mathbb{R}^d$, ρ any measure on D, $\{\lambda_k\}_{k=1}^{\infty}$ any positive sequence satisfying $\sum_{k=1}^{\infty} 1/\lambda_k < \infty$, and $\{e_k\}_{k=1}^{\infty}$ any real orthonormal basis of $L^2_c(\rho)$. We denote by $\langle\cdot,\cdot\rangle$ the inner product on $L^2_c(\rho)$. The operator L defined for any $\xi \in L^2_c(\rho)$ by

$$L\xi \doteq -\sum_{k=1}^{\infty} \lambda_k \langle\xi, e_k\rangle e_k$$

satisfies Condition 2.1. Such operators L include a class of pseudodifferential operators that arise in weak turbulence theory [28], for which the boundary conditions are periodic, the Fourier basis functions e_k are trigonometric, and eigenvalues λ_k are powers k^α. Condition 2.1 limits the power to $\alpha > 1$.

We proceed with the definition of the probabilistic model, which pertains to a sequence of finite-dimensional approximations to the GNLS equation (1.1). For these approximations we choose a spectral truncation of the GNLS-dynamics [7, 25, 35]. The same ideas can be applied to other discrete approximations such as finite-difference [26]. With respect to the basis $\{e_k\}$ of $L^2_c(\rho)$, let $W_n \doteq \text{span}\,\{e_1, \ldots, e_n\}$ be the n-dimensional subspace consisting of functions

$$\psi^{(n)}(x) = u^{(n)}(x) + iv^{(n)}(x) \doteq \sum_{k=1}^{n} \psi_k e_k(x), \tag{2.1}$$

with arbitrary complex coefficients $\psi_k = u_k + iv_k$. For each fixed n, the field $\psi^{(n)}$ takes values in $L^2_c(\rho)$ and corresponds to an n-dimensional microstate for the model; that is, a point $\psi = (\psi_1, \ldots, \psi_n)$ in the phase space $\Gamma_n \doteq \mathbb{C}^n$ or equivalently $\Gamma_n \doteq \mathbb{R}^{2n}$. The microscopic dynamics for this model is governed by

$$i\psi_t^{(n)} + L\psi^{(n)} + P^{(n)}\left(f(|\psi^{(n)}|^2)\psi^{(n)}\right) = 0, \tag{2.2}$$

where $P^{(n)}$ denotes the orthogonal projection that maps $L^2_c(\rho)$ onto W_n. This spectral truncation of the GNLS equation (1.1) is equivalent to a system of ordinary differential equations for the real Fourier coefficients u_k and v_k, $k = 1, \ldots, n$, having a canonical Hamiltonian form; namely,

$$\frac{du_k}{dt} = \lambda_k v_k - \int_D f((u^{(n)})^2 + (v^{(n)})^2) v^{(n)} e_k \, d\rho = \frac{\partial H_n}{\partial v_k},$$
$$\frac{dv_k}{dt} = -\lambda_k u_k + \int_D f((u^{(n)})^2 + (v^{(n)})^2) u^{(n)} e_k \, d\rho = -\frac{\partial H_n}{\partial u_k}$$

with Hamiltonian

$$\begin{aligned} H_n(\psi) &= H_n(u_1, v_1, \ldots, u_n, v_n) \\ &\doteq \frac{1}{2}\sum_{k=1}^{n} \lambda_k |\psi_k|^2 \; - \; \frac{1}{2}\int_D F(|\psi^{(n)}|^2)\, d\rho \\ &= -\frac{1}{2}\langle L\psi^{(n)}, \psi^{(n)}\rangle \; - \; \frac{1}{2}\int_D F(|\psi^{(n)}|^2)\, d\rho \\ &\equiv D_n(\psi) \; + \; \Phi_n(\psi). \end{aligned} \tag{2.3}$$

The functions D_n and Φ_n are defined by this display, and

$$\Phi(\psi) \doteq -\frac{1}{2}\int_D F(|\psi|^2)\, d\rho; \tag{2.4}$$

hence for $\psi \in W_n$, $\Phi_n(\psi) = \Phi(\psi^{(n)})$. Clearly, $H_n(\psi)$ equals $H(\psi^{(n)})$, the restriction to W_n of the functional H defined in (1.3). The spectrally truncated particle number, $Q_n(\psi) \doteq Q(\psi^{(n)})$, is also an invariant of the microscopic dynamics (2.2) and is given by

$$Q_n(\psi) = Q_n(u_1, v_1, \ldots, u_n, v_n) \doteq \frac{1}{2}\int_D |\psi^{(n)}|^2 \, d\rho \; = \; \frac{1}{2}\sum_{k=1}^{n} |\psi_k|^2. \tag{2.5}$$

We define the statistical equilibrium model by a Gibbs ensemble on the $2n$-dimensional phase space Γ_n, in which H_n is treated canonically and Q_n is treated microcanonically. We refer to this ensemble as the mixed ensemble, and we denote it by $P_\beta^N(d\psi)$, where $N \in [0, \infty)$ is a given particle number and $\beta > 0$ is a given inverse temperature. Formally, the mixed ensemble is the probability distribution

$$P_{n,\beta}^N(d\psi) \doteq \frac{1}{Z_n(\beta, N)} \exp(-\beta H_n(\psi))\, \delta(Q_n(\psi) - N)\, V_n(d\psi), \tag{2.6}$$

where $Z_n(\beta, N)$ is the normalizing constant

$$Z_n(\beta, N) \doteq \int_{\{Q_n = N\}} \exp(-\beta H_n(\psi))\, ds(\psi).$$

Here $V_n(d\psi) \doteq \prod_{k=1}^n du_k dv_k$ is the phase volume on Γ_n, and $ds(\psi)$ is the hypersurface area on the sphere $Q_n(\psi) = N$, which is the support of the distribution (2.6).

This choice of ensemble can be motivated intuitively from the known dynamical behavior of numerical solutions to (1.1). Long-time simulations of the dynamics show that, while the energy H is sensitive to the fluctuations that develop on the small scales, the particle number Q depends on the coherent structure on the large scale. This phenomenon is related to the phenomenological description of weak turbulence in which there is a flux of energy to small scales and of particle number to large scales. Physical reasoning then suggests that the appropriate ensemble be canonical in H, since energy is in contact with a bath of turbulent small-scale waves, and that it be microcanonical in Q, since the particle number is contained in the coherent large-scale waves which are isolated from the turbulent bath.

Let us define this mixed ensemble precisely as a conditional probability measure. We return to the decomposition $H_n(\psi) = D_n(\psi) + \Phi_n(\psi)$ given in (2.3). An easy calculation given in part (a) of Proposition 4.3 shows that for any bounded nonlinearity f, $\sigma > 0$ can be chosen sufficiently large so that $\Phi(\xi) + \sigma Q(\xi) \geq 0$ for all $\xi \in L^2_c(\rho)$. It follows that for all $n \in \mathbb{N}$ and $\psi \in W_n$,

$$\Phi_n(\psi) + \sigma Q_n(\psi) = \Phi(\psi) + \sigma Q(\psi) \geq 0.$$

Since $D_n(\psi) = -\frac{1}{2}\langle L\psi^{(n)}, \psi^{(n)}\rangle \geq 0$, we have for all $\psi \in W_n$

$$H_n(\psi) + \sigma Q_n(\psi) = D_n(\psi) + \Phi_n(\psi) + \sigma Q_n(\psi) \geq 0.$$

It is worth noting that such a σ also exists for a wider class of nonlinearities f; e.g., unbounded, but subcritical nonlinearities. We then construct the following σ-regularized canonical measure:

$$P_{n,\beta}(d\psi) \doteq \frac{1}{Z_n(\beta)} \exp(-\beta[H_n(\psi) + \sigma Q_n(\psi)])\, V_n(d\psi), \tag{2.7}$$

which exists and is normalizable. The normalizing constant $Z_n(\beta)$ is given by

$$Z_n(\beta) \doteq \int_{\Gamma_n} \exp(-\beta[H_n(\psi) + \sigma Q_n(\psi)])\, V_n(d\psi).$$

By contrast, when $\sigma = 0$, it is known that $Z_n(\beta)$ diverges for certain focusing nonlinearities since H_n goes to $-\infty$ in some directions of the phase space Γ_n [25, 27]. Ideally, we would like to define the mixed ensemble to be

$$P^N_{n,\beta}(d\psi) \doteq P_{n,\beta}(d\psi \mid Q_n(\psi) = N); \tag{2.8}$$

namely, a regular conditional distribution given the microcanonical constraint $Q_n = N$. In this formulation, the mixed ensemble (2.8) is independent of the choice of σ and coincides with the formal expression (2.6).

In order to avoid technicalities involving regular conditional distributions, we will consider, in place of (2.8), the conditional measure

$$P^{N,\varepsilon}_{n,\beta}(d\psi) \doteq P_{n,\beta}(d\psi \mid Q_n(\psi) \in [N-\varepsilon, N+\varepsilon]), \tag{2.9}$$

where ε is a positive parameter defining the thickened shell $[N-\varepsilon, N+\varepsilon]$. For suitable values of N, all sufficiently large n, and all $\varepsilon > 0$, $P_{n,\beta}\{Q_n(\psi) \in [N-\varepsilon, N+\varepsilon]\} > 0$, and so the conditional probability $P^{N,\varepsilon}_{n,\beta}(d\psi)$ is well defined [see (4.11)]. The main theorem in this paper, stated in Theorem 3.1, is the large deviation principle on $L^2_c(\rho)$ for $\psi^{(n)}$ with respect to $P^{N,\varepsilon}_{n,n\beta}$ in the continuum limit $n \to \infty$ followed by $\varepsilon \to 0$; N is kept fixed while β has been replaced by the mean-field scaling $n\beta$. With this scaling the ensemble mean energy $\langle H_n\rangle$ tends to a finite limit E. In contrast to the definition (2.8), the conditional measures $P^{N,\varepsilon}_{n,n\beta}$ are no longer independent of σ because of the presence of ε in the definition (2.9). However, the rate function in Theorem 3.1 is independent of σ.

3. Statement of Main Theorem: LDP for Mixed Ensemble

Earlier investigations in [22, 25] give theoretical and numerical evidence that, for the basic NLS equation in which $L = \partial^2/\partial x^2$ and ρ is Lebesgue measure on $[0, \ell]$, the random field $\psi^{(n)}$ defined in (2.1) concentrates on the set of ground states $e^{i\theta}\varphi(x)$ of the NLS equation in the continuum limit $n \to \infty$ with fixed E and N. In [25] a mean-field approximation is developed to explain this phenomenon, and long-time numerical simulations of the freely-evolving dynamics support the theory [22]. Alternatively, this phenomenon of concentration on the set of ground states can be demonstrated by implementing Monte-Carlo simulations of the mixed Gibbs ensemble in the continuum limit; this approach is used in Sect.5 of the present paper. Our main goal in the present paper is to formulate and prove a large deviation principle that holds for $\psi^{(n)}$ with respect to the mixed ensemble $P^{N,\varepsilon}_{n,n\beta}$ defined in (2.9) and that is valid for general operators L satisfying Condition 2.1. This large deviation principle constitutes a precise and rigorous statement of the concentration phenomenon that occurs in the continuum limit.

We start with two definitions. Let $\mathcal{X}$ be a Hilbert space, J a function mapping $\mathcal{X}$ into $[0, \infty]$, $\{\mu_n, n \in \mathbb{N}\}$ a sequence of probability measures on $\mathcal{X}$, $\{\mu_n^\varepsilon, n \in \mathbb{N}, \varepsilon > 0\}$ a family of probability measures on $\mathcal{X}$, and a_n a positive sequence tending to ∞. J is called a rate function if for each $M < \infty$ the set $\{\xi \in \mathcal{X} : J(\xi) \le M\}$ is compact. For A a subset of $\mathcal{X}$ we write $J(A)$ for $\inf\{J(\xi) : \xi \in A\}$. The sequence μ_n is said to satisfy a large deviation principle (LDP) on $\mathcal{X}$ with the scaling constants a_n and the rate function J if for each closed subset F of $\mathcal{X}$

$$\limsup_{n\to\infty} \frac{1}{a_n} \log \mu_n\{F\} \le -J(F)$$

and for each open subset G of $\mathcal{X}$

$$\liminf_{n\to\infty} \frac{1}{a_n} \log \mu_n\{G\} \ge -J(G).$$

Similarly, as $n \to \infty$ and $\varepsilon \to 0$, the collection μ_n^ε is said to satisfy an LDP on $\mathcal{X}$ with the scaling constants a_n and the rate function J if for each closed subset F of $\mathcal{X}$

$$\limsup_{\varepsilon\to 0}\, \limsup_{n\to\infty} \frac{1}{a_n} \log \mu_n^\varepsilon\{F\} \le -J(F)$$

and for each open subset G of $\mathcal{X}$

$$\liminf_{\varepsilon\to 0}\, \liminf_{n\to\infty} \frac{1}{a_n} \log \mu_n^\varepsilon\{G\} \ge -J(G).$$

The main result in this paper is the LDP stated in Theorem 3.1. We first indicate the form of the rate function. For $\xi = \sum_{k=1}^\infty \langle \xi, e_k\rangle e_k \in L^2_c(\rho)$, the Hamiltonian introduced in (1.3) can be written as $H(\xi) = D(\xi) + \Phi(\xi)$, where $\Phi(\xi)$ is defined in (2.4), and

$$D(\xi) \doteq \frac{1}{2} \sum_{k=1}^\infty \lambda_k |\langle \xi, e_k\rangle|^2 .$$

In terms of the square root of the positive, symmetric operator $-L$,

$$D(\xi) = \begin{cases} \frac{1}{2}\|\sqrt{-L}\xi\|^2 & \text{if } \xi \in \mathrm{dom}(\sqrt{-L}) \\ \infty & \text{if } \xi \in L^2_c(\rho) \setminus \mathrm{dom}(\sqrt{-L}) \,, \end{cases}$$

where

$$\sqrt{-L}\xi \doteq \sum_{k=1}^{\infty} \sqrt{\lambda_k}\langle \xi, e_k\rangle e_k \, .$$

For $N \in [0, \infty)$ we also introduce

$$\bar{E}(N) \doteq \inf\{H(\xi) : \xi \in L^2_c(\rho),\, Q(\xi) = N\}. \tag{3.1}$$

We call $\bar{E}$ the coherent energy function since $\bar{E}(N)$ is the energy of the coherent structure with particle number N. The function $\bar{E}$ is lower semicontinuous and bounded below; indeed, by (4.12) $\bar{E}$ differs by a constant from the function $\tilde{E}$ defined in (4.10), which by part (a) of Proposition 4.5 is nonnegative and lower semicontinuous. For $N \in [0, \infty)$ and $\xi \in L^2_c(\rho)$, the rate function in Theorem 3.1 is defined to be

$$J^N(\xi) \doteq \begin{cases} H(\xi) - \bar{E}(N) & \text{if } Q(\xi) = N \\ \infty & \text{otherwise.} \end{cases} \tag{3.2}$$

Theorem 3.1 will be proved in the next section.

Theorem 3.1. *The $L^2_c(\rho)$–valued process $\psi^{(n)}$ is defined in (2.1) and the mixed ensemble $P^{N,\varepsilon}_{n,n\beta}$ in* (2.9)*. We fix $\beta > 0$, take $N \in [0, \infty)$, and assume Condition* 2.1*. Then as $n \to \infty$ and $\varepsilon \to 0$, the $P^{N,\varepsilon}_{n,n\beta}$-distributions of $\psi^{(n)}$ satisfy the LDP on $L^2_c(\rho)$ with the scaling constants $n\beta$ and the rate function J^N defined in* (3.2).

Heuristically, the LDP means that the elements $\bar{\xi} \in L^2_c(\rho)$ that minimize H subject to the constraint $Q = N$ are the overwhelmingly most probable states with respect to the mixed ensemble in the continuum limit $n \to \infty$ followed by $\varepsilon \to 0$. This set of constrained minimizers is the set of equilibrium macrostates or ground states; we denote it by $\mathcal{E}^N$. For an equilibrium macrostate $\bar{\xi}$ we have $J^N(\bar{\xi}) = 0$, while for any $\xi \in L^2_c(\rho)$ that is not an equilibrium macrostate we have $J^N(\xi) > 0$. We now consider, for any $r > 0$, the complement of an r-neighborhood of the equilibrium set $\mathcal{E}^N$ and define

$$j(r) \doteq \inf\{J^N(\xi) : \operatorname{dist}(\xi, \mathcal{E}^N) \geq r > 0\},$$

the distance being taken in the $L^2_c(\rho)$-norm. Then $j(r) > 0$. From the large deviation upper bound in Theorem 3.1, we infer that

$$P^{N,\varepsilon}_{n,n\beta}\{\operatorname{dist}(\psi^{(n)}, \mathcal{E}^N) \geq r > 0\} \;\leq\; e^{-n\beta j(r)/2} \;\to 0 \quad \text{as } n \to \infty\,,\, \varepsilon \to 0.$$

Thus any set of $\xi \in L^2_c(\rho)$ that lies a positive distance from the equilibrium set $\mathcal{E}^N$ has an exponentially small probability of being observed for sufficiently large n and sufficiently small $\varepsilon > 0$. This property of $\mathcal{E}^N$ justifies calling it the set of equilibrium macrostates.

An LDP can be viewed as an exponential-order refinement of the law of large numbers [11, 16]. From this viewpoint, we might expect the random field $\psi^{(n)}(x)$ to satisfy an LDP in the continuum limit because it is the sum of component fields $\psi_k e_k(x)$ that are asymptotically independent. In essence, this insight is the basis for the mean-field approximation used in [25], which relies on the smallness of the fluctuations of $\psi^{(n)}(x)$ in the $L^2_c(\rho)$-norm. As we will see in the next section, the proof of the LDP for $\psi^{(n)}$ depends crucially on the continuity of the functionals Q and Φ with respect to the $L^2_c(\rho)$-topology [Prop. 4.3]. These properties that are needed to prove the LDP are intimately related to the properties used to derive the mean-field theory.

4. Proof of Theorem 3.1

Given $\beta > 0$ we introduce the following Gaussian measures on the phase space $\Gamma_n = \mathbb{R}^{2n}$:

$$G_{n,\beta}(d\psi) \doteq \frac{1}{C_{n,\beta}} \exp(-\beta D_n(\psi))\, V_n(d\psi), \tag{4.1}$$

where $C_{n,\beta}$ is the normalizing constant

$$C_{n,\beta} \doteq \int_{\mathbb{R}^{2n}} \exp(-\beta D_n(\psi))\, V_n(d\psi).$$

The proof of Theorem 3.1 is based on the LDP of $\psi^{(n)}$ with respect to the measures $G_{n,n\beta}$, where β in (4.1) has been replaced by $n\beta$. The motivation for introducing these measures is that the canonical measures $P_{n,n\beta}$ can be expressed in terms of $G_{n,n\beta}$ [see (4.8)], and hence the LDP for $P_{n,n\beta}$ follows directly from that for $G_{n,n\beta}$ [Thm. 4.1]. In turn, the LDP for the $P^{N,\varepsilon}_{n,n\beta}$-distributions of $\psi^{(n)}$ stated in Theorem 3.1 is derived from the LDP for the $P_{n,n\beta}$-distributions of $\psi^{(n)}$.

The LDP for the $G_{n,n\beta}$-distributions of $\psi^{(n)}$ is stated in the next theorem and is proved using a corollary of Baldi's Theorem stated in [11, Cor. 4.5.27]. After giving this proof, we sketch a second proof using an LDP for Gaussian processes proved by Bolthausen [5].

The next theorem states the LDP on the complex Hilbert space of square-integrable functions $L^2_c(\rho)$; any $\xi \in L^2_c(\rho)$ can be written as $\xi^1 + i\xi^2$, where ξ^1 and ξ^2 are elements of the corresponding real Hilbert space $L^2(\rho)$. Since both proofs of Theorem 4.1 are based on results formulated for real spaces, we will prove an equivalent LDP replacing $L^2_c(\rho)$ by the topologically equivalent Hilbert space $L^2(\rho) \times L^2(\rho)$; this equivalence is defined by the correspondence $\xi = \xi^1 + i\xi^2 \in L^2_c(\rho) \leftrightarrow (\xi^1, \xi^2) \in L^2(\rho) \times L^2(\rho)$.

Theorem 4.1. *The $L^2_c(\rho)$-valued process $\psi^{(n)}$ is defined in* (2.1) *and the Gaussian measures $G_{n,\beta}$ in* (4.1). *We fix $\beta > 0$ and assume Condition* 2.1. *Then as $n \to \infty$, the $G_{n,n\beta}$-distributions of $\psi^{(n)}$ satisfy the LDP on $L^2_c(\rho)$ with the scaling constants $n\beta$ and the rate function*

$$I(\xi) = I(\xi^1 + i\xi^2) \doteq \frac{1}{2}\sum_{k=1}^{\infty} \lambda_k |\xi^1_k + i\xi^2_k|^2 = \frac{1}{2}\sum_{k=1}^{\infty} \lambda_k \left((\xi^1_k)^2 + (\xi^2_k) \right), \tag{4.2}$$

where $\xi^1_k + i\xi^2_k \doteq \langle \xi^1, e_k \rangle + i\langle \xi^2, e_k \rangle$. Alternatively, the rate function $I(\xi)$ equals

$$D(\xi) = \begin{cases} \frac{1}{2}\|\sqrt{-L}\xi\|^2 & \text{if } \xi \in \operatorname{dom}(\sqrt{-L}) \\ \infty & \text{if } \xi \in L^2_c(\rho) \setminus \operatorname{dom}(\sqrt{-L}). \end{cases} \tag{4.3}$$

Proof. The equality of the quantities defined in (4.2) and in (4.3) is immediate. The function $\psi^{(n)}$ defined in (2.1) can be written as

$$\psi^{(n)} = \sum_{k=1}^{n} u_k e_k(x) + i \sum_{k=1}^{n} v_k e_k(x) \equiv \psi^{(n),1} + i\psi^{(n),2}.$$

Setting $\psi_k^1 \doteq u_k$ and $\psi_k^2 \doteq v_k$, we have $\psi^{(n),1} = \sum_{k=1}^n \psi_k^1 e_k$ and $\psi^{(n),2} = \sum_{k=1}^n \psi_k^2 e_k$. Because $L_c^2(\rho)$ and $L^2(\rho) \times L^2(\rho)$ are topologically equivalent, proving an LDP for the $G_{n,n\beta}$-distributions of $\psi^{(n)}$ on $L_c^2(\rho)$ is equivalent to proving an LDP for the $G_{n,n\beta}$-distributions of $(\psi^{(n),1}, \psi^{(n),2})$ on $L^2(\rho) \times L^2(\rho)$. The inner product on $L^2(\rho) \times L^2(\rho)$ is

$$\langle(\xi^1, \xi^2), (\theta^1, \theta^2)\rangle \doteq \langle\xi^1, \theta^1\rangle + \langle\xi^2, \theta^2\rangle = \sum_{\alpha=1}^{2}\sum_{k=1}^{\infty} \xi_k^\alpha \theta_k^\alpha,$$

where for $\alpha = 1, 2$ $\xi_k^\alpha \doteq \langle\xi^\alpha, e_k\rangle$ and $\theta_k^\alpha \doteq \langle\theta^\alpha, e_k\rangle$.

We begin the proof by computing, for $\varphi \in L^2(\rho) \times L^2(\rho)$,

$$\begin{aligned} c(\varphi) &\doteq \lim_{n\to\infty} \frac{1}{n\beta} \log \int_{\mathbb{R}^n \times \mathbb{R}^n} \exp\Big(n\beta \Big\langle \varphi, \psi^{(n)} \Big\rangle\Big)\, G_{n,n\beta}(d\psi) \\ &= \lim_{n\to\infty} \frac{1}{n\beta} \log \left(\frac{1}{C_{n,n\beta}} \prod_{\alpha=1}^{2}\prod_{k=1}^{n} \int_{\mathbb{R}} \exp\Big(n\beta \left(\varphi_k^\alpha \psi_k^\alpha - \frac{1}{2}\lambda_k (\psi_k^\alpha)^2 \right)\Big) d\psi_k^\alpha \right) \\ &= \lim_{n\to\infty} \frac{1}{n\beta} \log \prod_{k=1}^{n} \exp\left(\frac{n\beta}{2} \frac{(\varphi_k^1)^2 + (\varphi_k^2)^2}{\lambda_k} \right) \\ &= \frac{1}{2} \sum_{k=1}^{\infty} \frac{(\varphi_k^1)^2 + (\varphi_k^2)^2}{\lambda_k}. \end{aligned} \tag{4.4}$$

By Condition 2.1 $\lambda_k > 0$ and $\lambda_k \to \infty$; hence $0 \le c(\varphi) \le \text{const} \cdot \|\varphi\|^2 < \infty$.

Because of the relatively simple form of c, it is elementary to check that c is Gateaux differentiable and is weakly continuous on $L^2(\rho) \times L^2(\rho)$. Because c is a sum of quadratic terms, it is also straightforward to calculate its Legendre-Fenchel transform. For $\xi = (\xi^1, \xi^2) \in L^2(\rho) \times L^2(\rho)$, this function is given by

$$\begin{aligned} I(\xi) &\doteq \sup_{\varphi \in L^2(\rho)\times L^2(\rho)} \{\langle\varphi, \xi\rangle - c(\varphi)\} \\ &= \sum_{k=1}^{\infty} \sup_{\varphi_k^1 \in \mathbb{R}} \left\{ \varphi_k^1 \xi_k^1 - \frac{1}{2}\frac{(\varphi_k^1)^2}{\lambda_k} \right\} + \sum_{k=2}^{\infty} \sup_{\varphi_k^2 \in \mathbb{R}} \left\{ \varphi_k^2 \xi_k^2 - \frac{1}{2}\frac{(\varphi_k^2)^2}{\lambda_k} \right\} \\ &= \frac{1}{2} \sum_{k=1}^{\infty} \lambda_k \left((\xi_k^1)^2 + (\xi_k^2)^2 \right). \end{aligned} \tag{4.5}$$

The function $I(\xi)$ calculated in the preceding display coincides with the function $I(\xi)$ defined in (4.2). By Corollary 4.5.27 in [11], we will be able to conclude that the $G_{n,n\beta}$-distributions of $(\psi^{(n),1}, \psi^{(n),2})$ on $L^2(\rho) \times L^2(\rho)$—and thus the $G_{n,n\beta}$-distributions of $\psi^{(n)}$ on $L_c^2(\rho)$—satisfy the LDP with rate function $I(\xi)$ after we show that the $G_{n,n\beta}$-distributions of $(\psi^{(n),1}, \psi^{(n),2})$ are exponentially tight; i.e., for any $K < \infty$ there exists a compact set A such that

$$\limsup_{n\to\infty} \frac{1}{n\beta} \log G_{n,n\beta}\{(\psi^{(n),1}, \psi^{(n),2}) \in A^c\} < -K.$$

In order to prove the exponential tightness, we define for $M < \infty$ the level sets

$$A_M \doteq \{\xi \in L^2(\rho) \times L^2(\rho) : I(\xi) \leq M\}.$$

We first prove that the sets A_M are compact by showing that any sequence $\xi^{(n)} = (\xi^{(n),1}, \xi^{(n),2})$ in A_M has a subsequence converging to an element of A_M. Since $\xi^{(n)} \in A_M$, we have for all $k \in \mathbb{N}$, $n \in \mathbb{N}$, and $\alpha = 1, 2$ $\frac{1}{2}\lambda_k(\xi_k^{(n),\alpha})^2 \leq M$; thus for each k and α $\xi_k^{(n),\alpha}$ has a convergent subsequence. For $\alpha = 1, 2$ let $\xi_1^{(n_1),\alpha}$ be the convergent subsequence of $\xi_1^{(n),\alpha}$, and for each k and α let $\xi_{k+1}^{(n_{k+1}),\alpha}$ be the convergent subsequence of $\xi_k^{(n_k),\alpha}$. For each k and α there exists ξ_k^α such that $\lim_{n_k \to \infty} \xi_k^{(n_k),\alpha} = \xi_k^\alpha$. A diagonalization argument yields a subsequence $\hat{n} \in \mathbb{N}$ such that $\lim_{\hat{n} \to \infty} \xi_k^{(\hat{n}),\alpha} = \xi_k^\alpha$ for each k and α. The quantity $\xi \doteq \sum_{k=1}^\infty (\xi_k^1, \xi_k^2) e_k$ is an element of A_M. Indeed, by Fatou's lemma

$$\frac{1}{2} \sum_{\alpha=1}^{2} \sum_{k=1}^{\infty} \lambda_k (\xi_k^\alpha)^2 \leq \liminf_{\hat{n} \to \infty} \frac{1}{2} \sum_{\alpha=1}^{2} \sum_{k=1}^{\infty} \lambda_k (\xi_k^{(\hat{n}),\alpha})^2 \leq M.$$

For each $\hat{n}$ we define $\xi^{(\hat{n})} \doteq \sum_{k=1}^\infty (\xi_k^{(\hat{n}),1}, \xi_k^{(\hat{n}),2}) e_k$. In order to complete the proof that A_M is compact, we show that $\xi^{(\hat{n})} \to \xi$ in $L^2(\rho) \times L^2(\rho)$. Since $\xi^{(\hat{n})}$ and ξ are in A_M, the finiteness of $\sum_{k=1}^\infty 1/\lambda_k$ assumed in Condition 2.1 implies that uniformly over $\hat{n}$

$$\begin{aligned} \sum_{\alpha=1}^{2} \sum_{k=1}^{\infty} (\xi_k^{(\hat{n}),\alpha} - \xi_k^\alpha)^2 &\leq 2 \sum_{\alpha=1}^{2} \sum_{k=1}^{\infty} \left[(\xi_k^{(\hat{n}),\alpha})^2 + (\xi_k^\alpha)^2 \right] \\ &\leq 4 \sum_{k=1}^{\infty} \frac{M}{\lambda_k} < \infty. \end{aligned}$$

Hence by the dominated convergence theorem

$$\lim_{\hat{n} \to \infty} \|\xi^{(\hat{n})} - \xi\|^2 = \lim_{\hat{n} \to \infty} \sum_{\alpha=1}^{2} \sum_{k=1}^{\infty} (\xi_k^{(\hat{n}),\alpha} - \xi_k^\alpha)^2 = 0.$$

This concludes the proof that A_M is compact in $L^2(\rho) \times L^2(\rho)$.

We complete the proof of Theorem 4.1 by showing that the $G_{n,n\beta}$-distributions of $(\psi^{(n),1}, \psi^{(n),2})$ are exponentially tight. For any $M < \infty$ Chebyshev's inequality yields

$$\begin{aligned} &G_{n,n\beta}\{(\psi^{(n),1}, \psi^{(n),2}) \in A_M^c\} \\ &\quad = G_{n,n\beta} \left\{ \frac{1}{2} \sum_{k=1}^{n} \lambda_k \left((\psi_k^1)^2 + \psi_k^2)^2 \right) > M \right\} \\ &\quad \leq \exp\left(-\frac{n\beta}{2} M \right) \frac{1}{C_{n,n\beta}} \int_{\mathbb{R}^n \times \mathbb{R}^n} \exp\left(\frac{n\beta}{4} \sum_{k=1}^{n} \lambda_k \left((\psi_k^1)^2 + \psi_k^2)^2 \right) \right) G_{n,n\beta}(d\psi) \end{aligned}$$

$$
\begin{aligned}
&= \exp\left(-\frac{n\beta}{2}M\right) \prod_{k=1}^{n} \prod_{\alpha=1}^{2} \frac{\int_{\mathbb{R}} \exp\left[-\frac{n\beta}{4}\lambda_k (\psi_k^\alpha)^2\right] d\psi_k^\alpha}{\int_{\mathbb{R}} \exp\left[-\frac{n\beta}{2}\lambda_k (\psi_k^\alpha)^2\right] d\psi_k^\alpha} \\
&= \exp\left(-\frac{n\beta}{2}M\right) 2^n.
\end{aligned}
$$

It follows that

$$
\limsup_{n\to\infty} \frac{1}{n\beta} \log G_{n,n\beta}\{(\psi^{(n),1}, \psi^{(n),2}) \in A_M^c\} \le -\frac{M}{2} + \frac{\log 2}{\beta}.
$$

Since A_M is compact and M can be taken arbitrarily large, the proof is complete. □

Before proving Theorem 3.1 we sketch a second proof of Theorem 4.1 using an LDP for Gaussian processes proved by Bolthausen [5]. Let $(\Omega, \mathcal{F}, \Pi)$ be a probability space on which is defined a doubly indexed sequence of independent, $N(0, 1)$ Gaussian random variables g_k^α indexed by $k \in \mathbb{N}$ and $\alpha = 1, 2$. In terms of the eigenvalues and eigenfunctions of L introduced in Condition 2.1, we define for $n \in \mathbb{N}$, $\omega \in \Omega$, and $x \in D$ the independent mean-0 Gaussian processes

$$
\begin{aligned}
y^{(n),1} &= y^{(n),1}(\omega, x) \doteq \sum_{k=1}^{n} g_k^1(\omega)\, \frac{e_k(x)}{\sqrt{\lambda_k}} \quad \text{and} \\
y^{(n),2} &= y^{(n),2}(\omega, x) \doteq \sum_{k=1}^{n} g_k^2(\omega)\, \frac{e_k(x)}{\sqrt{\lambda_k}}.
\end{aligned} \tag{4.6}
$$

These processes take values in $L^2(\rho)$, and $y^{(n),1} + iy^{(n),2}$ take values in $L_c^2(\rho)$.

The basic NLS equation is defined by $L = \partial^2/\partial x^2$ on $[0, \ell]$. Inserting into (4.6) the eigenvalues and eigenfunctions given in Example 2.2(a), we have for $\alpha = 1, 2$ and $x \in [0, \ell]$,

$$
y^{(n),\alpha}(\omega, x) = \frac{\sqrt{2\ell}}{\pi} \sum_{k=1}^{n} g_k^\alpha(\omega) \frac{\sin(k\pi x/\ell)}{k}.
$$

It is well known that with probability 1, as $n \to \infty$ these processes converge in $L^2(dx)$ to independent Brownian bridges on $[0, \ell]$. The processes $y^{(n),\alpha}$ are closely related to the processes used by Wiener in his construction of Brownian motion [21, pp. 21-22]. The probability–1 convergence of $y^{(n),\alpha}$ in the general case of (4.6) is the basis of the second proof of Theorem 4.1. We will prove the convergence in a moment.

We need the following lemma relating the distributions of $(y^{(n),1} + iy^{(n),2})/\sqrt{\beta}$ and $\psi^{(n)}$. The routine proof is omitted.

Lemma 4.2. *Fix $\beta > 0$. Then as measures on $L_c^2(\rho)$, the Π-distributions of $(y^{(n),1} + iy^{(n),2})/\sqrt{\beta}$ and the $G_{n,\beta}$-distributions of $\psi^{(n)}$ are equal. In particular, replacing β by $n\beta$, we see that the Π-distributions of $(y^{(n),1} + iy^{(n),2})/\sqrt{n\beta}$ and the $G_{n,n\beta}$-distributions of $\psi^{(n)}$ are equal.*

This lemma allows us to prove Theorem 4.1 by showing that the Π-distributions of $(y^{(n),1} + iy^{(n),2})/\sqrt{n\beta}$ on $L^2_c(\rho)$—equivalently, the Π-distributions of $(y^{(n),1}, y^{(n),2})/\sqrt{n\beta}$ on $L^2(\rho) \times L^2(\rho)$—satisfy the LDP with the scaling constants $n\beta$ and the rate function $I(\xi)$ defined in (4.2).

We first prove that as $n \to \infty$, with Π-probability 1 the Gaussian processes $y^{(n),1}$ and $y^{(n),2}$ defined in (4.6) converge in $L^2(\rho)$ to the independent Gaussian processes

$$y^1 \doteq \sum_{k=1}^{\infty} g_k^1 \frac{e_k}{\sqrt{\lambda_k}} \text{ and } y^2 \doteq \sum_{k=1}^{\infty} g_k^2 \frac{e_k}{\sqrt{\lambda_k}}.$$

On the product space $\Omega \times D$, we define for $n \in \mathbb{N}$ and $\alpha = 1, 2$

$$\overline{y}^{(n),\alpha} \doteq \sum_{k=1}^{n} |g_k^\alpha| \frac{|e_k|}{\sqrt{\lambda_k}} \quad \text{and} \quad \overline{y}^{\alpha} \doteq \sum_{k=1}^{\infty} |g_k^\alpha| \frac{|e_k|}{\sqrt{\lambda_k}}.$$

Since the g_k^α are independent, $N(0, 1)$ Gaussian random variables and the e_k are orthonormal in $L^2(\rho)$,

$$\begin{aligned}
\int_{\Omega \times D} (\overline{y}^\alpha)^2 \, d\Pi \times d\rho &= \lim_{n\to\infty} \int_{\Omega \times D} (\overline{y}^{(n),\alpha})^2 \, d\Pi \times d\rho \\
&= \sum_{k=1}^{\infty} \int_D \left(\int_\Omega (g_k^\alpha)^2 \frac{(e_k)^2}{\lambda_k} d\Pi \right) d\rho \\
&= \sum_{k=1}^{\infty} \frac{1}{\lambda_k}.
\end{aligned}$$

Since this sum is finite by Condition 2.1 and $|y^\alpha| \le |\overline{y}^\alpha|$, we have $\overline{y}^\alpha \in L^2(\rho)$ Π-a.e. and thus $y^\alpha \in L^2(\rho)$ Π-a.e. The bound $|y^\alpha - y^{(n),\alpha}|^2 \le (2\overline{y}^\alpha)^2 \in L^1(\rho)$ Π-a.e. allows us to apply the dominated convergence theorem, which yields the desired limit:

$$\int_D |y^\alpha - y^{(n),\alpha}|^2 \, d\rho \to 0 \qquad \Pi\text{−a.e.}$$

This completes the proof that with Π-probability 1 $y^{(n),1}$ and $y^{(n),2}$ converge in $L^2(\rho)$ to y^1 and y^2.

The probability–1 convergence just proved implies the weak convergence

$$\begin{aligned}
&\Pi\{(y^{(n),1}, y^{(n),2})/\sqrt{\beta} \in d\xi^1 \times d\xi^2\} \\
&\quad \Rightarrow \Pi\{(y^1, y^2)/\sqrt{\beta} \in d\xi^1 \times d\xi^2\} \text{ on } L^2(\rho) \times L^2(\rho).
\end{aligned} \tag{4.7}$$

By Theorem 2 in [5] and the discussion following that theorem, we conclude that the Π-distributions of $(y^{(n),1}, y^{(n),2})/\sqrt{n\beta}$ on $L^2(\rho) \times L^2(\rho)$—and thus the $G_{n,n\beta}$-distributions of $\psi^{(n)}$ on $L^2_c(\rho)$—satisfy the LDP with scaling constants $n\beta$ and rate function

$$\frac{1}{\beta} h_\beta(\xi) \doteq \frac{1}{\beta} \cdot \sup_{\varphi \in L^2(\rho)\times L^2(\rho)} \{\langle \varphi, \xi \rangle - \log M_\beta(\varphi)\},$$

where

$$M_\beta(\varphi) \doteq \int_\Omega \exp\langle \varphi, (y^1, y^2)/\sqrt{\beta} \rangle \, \Pi(d\varphi).$$

In order to calculate $M_\beta(\varphi)$, we use the bound

$$\sup_{n\in\mathbb{N}} \int_\Omega \exp(t\|(y^{(n),1}, y^{(n),2})\|)\, d\Pi < \infty \quad \text{for all } t > 0,$$

which follows from the weak convergence (4.7) [5, p. 427]. Applying Lemma 4.2 and calculating as in (4.4), we find that for $\varphi = (\varphi^1, \varphi^2) \in L^2(\rho) \times L^2(\rho)$

$$\begin{aligned} M_\beta(\varphi) &= \lim_{n\to\infty} \int_\Omega \exp\langle \varphi, (y^{(n),1}, y^{(n),2})/\sqrt{\beta}\rangle\, d\Pi \\ &= \lim_{n\to\infty} \int_{\mathbb{R}^{2n}} \exp\langle \varphi, \psi^{(n)}\rangle\, G_{n,\beta}(d\psi) \\ &= \exp\left(\frac{1}{2\beta} \sum_{k=1}^{\infty} \frac{(\varphi_k^1)^2 + (\varphi_k^2)^2}{\lambda_k} \right). \end{aligned}$$

Via a calculation as in (4.5), we conclude that for $\xi = (\xi^1, \xi^2) \in L^2(\rho) \times L^2(\rho)$

$$\frac{1}{\beta} h(\xi) = \frac{1}{2} \sum_{k=1}^{\infty} \lambda_k (\xi_k^1)^2 + (\xi_k^2)^2).$$

Since this equals the function $I(\xi)$ defined in (4.2), our sketch of the second proof of Theorem 4.1 is complete.

We now turn to the proof of Theorem 3.1, which states the LDP for the distributions of $\psi^{(n)}$ with respect to the conditional measures $P_{n,n\beta}^{N,\varepsilon}$ defined in (2.9). Before proving this theorem, we must establish several properties of the functionals Q and F appearing in $P_{n,n\beta}^{N,\varepsilon}$. We recall that for $\xi \in L_c^2(\rho)$

$$Q(\xi) \doteq \frac{1}{2} \int_D |\xi|^2\, d\rho \quad \text{and} \quad \Phi(\xi) \doteq -\frac{1}{2} \int_D F(|\xi|^2)\, d\rho$$

and that for $a \geq 0$ $F(a) \doteq \int_0^a f(s)\, ds$.

Proposition 4.3. *The following properties are valid.*

(a) *For any* $\sigma > \|f\|_\infty$ *we have* $\Phi(\xi) + \sigma Q(\xi) \geq 0$ *for all* $\xi \in L_c^2(\rho)$.
(b) *Both* Q *and* Φ *are continuous functionals on* $L_c^2(\rho)$.

Proof. (a) For any $\xi \in L_c^2(\rho)$,

$$\Phi(\xi) + \sigma Q(\xi) = -\frac{1}{2} \int_D F(|\xi|^2)\, d\rho + \frac{\sigma}{2} \int_D |\xi|^2\, d\rho \geq -\frac{\|f\|_\infty}{2} \|\xi\|^2 + \frac{\sigma}{2} \|\xi\|^2.$$

The last expression is positive provided $\sigma > \|f\|_\infty$.

(b) The continuity of Q is obvious. To prove the continuity of Φ, we note that for $0 \le a \le b < \infty$, $|F(b) - F(a)| \le \|f\|_\infty |b-a|$. Hence for ξ and ζ in $L^2_c(\rho)$,

$$\begin{aligned}
&|\Phi(\xi) - \Phi(\zeta)| \\
&\le \frac{\|f\|_\infty}{2}\left[\int_{\{|\xi|\ge|\zeta|\}} \left(|\xi|^2 - |\zeta|^2\right) d\rho + \int_{\{\xi|<|\zeta|\}} \left(|\zeta|^2 - |\xi|^2\right) d\rho\right] \\
&\le \frac{\|f\|_\infty}{2} \int_D \left| \, |\xi|^2 - |\zeta|^2 \, \right| d\rho \\
&\le \frac{\|f\|_\infty}{2} \int_D |\xi - \zeta| \cdot (|\xi| + |\zeta|) \, d\rho \\
&\le \frac{\|f\|_\infty}{2} \left(\int_D |\xi - \zeta|^2 d\rho\right)^{1/2} \left(\int_D (|\xi| + |\zeta|)^2 \, d\rho\right)^{1/2} \\
&\le \frac{\|f\|_\infty}{2} \|\xi - \zeta\| \left(2\|\xi\|^2 + 2\|\zeta\|^2\right)^{1/2}.
\end{aligned}$$

This yields the continuity of Φ. □

For the remainder of this section we choose $\sigma > \|f\|_\infty$ so that $\Phi(\xi) + \sigma Q(\xi) \ge 0$ for any $\xi \in L^2_c(\rho)$ [Prop. 4.3(a)]. We are now ready to prove Theorem 3.1, which states the LDP for the distributions of $\psi^{(n)}$ with respect to the conditional measures $P^{N,\varepsilon}_{n,n\beta}$. We first express the measures $P_{n,n\beta}$ in terms of $G_{n,n\beta}$:

$$\begin{aligned}
P_{n,n\beta}(d\psi) &\doteq \frac{1}{Z_n(n\beta)} \exp(-n\beta[H_n(\psi) + \sigma Q_n(\psi)]) \, V_n(d\psi) \\
&= \frac{1}{\hat{Z}_n(n\beta)} \exp(-n\beta[\Phi(\psi^{(n)}) + \sigma Q(\psi^{(n)})]) \, G_{n,n\beta}(d\psi), \qquad (4.8)
\end{aligned}$$

where $\hat{Z}_n(n\beta)$ denotes the normalizing constant

$$\hat{Z}_n(n\beta) \doteq \int_{\Gamma_n} \exp(-n\beta[\Phi(\psi^{(n)}) + \sigma Q(\psi^{(n)})]) \, G_{n,n\beta}(d\psi).$$

In order to prove the LDP for the measures $P_{n,n\beta}$, we need a definition. Let J be a rate function on $L^2_c(\rho)$. A sequence of measure μ_n on $L^2_c(\rho)$ is said to satisfy the Laplace principle on $L^2_c(\rho)$ with the scaling constants $n\beta$ and the rate function J if for all bounded continuous functions h,

$$\lim_{n\to\infty} \frac{1}{n\beta} \log \int_{L^2_c(\rho)} \exp(-n\beta h) \, d\mu_n = - \inf_{\xi \in L^2_c(\rho)} \{h(\xi) + J(\xi)\}.$$

As proved in Theorems 1.2.3 and 1.2.5 in [14], μ_n satisfies the Laplace principle on $L^2_c(\rho)$ with the rate function J if and only if μ_n satisfies the LDP on $L^2_c(\rho)$ with the rate function J.

The measures $P_{n,n\beta}$ defined in (2.7) have the form of a canonical ensemble with interaction function $\Phi + \sigma Q$. We prove the LDP for the $P_{n,n\beta}$-distributions of $\psi^{(n)}$ by proving the Laplace principle for these distributions. By Theorem 4.1 the $G_{n,n\beta}$-distributions of $\psi^{(n)}$ satisfy the LDP on $L^2_c(\rho)$ with rate function $D(\xi)$. Since $\Phi(\xi) + \sigma Q(\xi) \ge 0$ for any $\xi \in L^2_c(\rho)$ [Prop. 4.3 (a)], for any bounded continuous function h , $\Phi + \sigma Q + h$

is bounded below. Hence the Laplace principle for the $P_{n,n\beta}$-distributions of $\psi^{(n)}$ is a consequence of Theorem 1.3.4 in [14]. A straightforward calculation (see, e.g., the proof of Thm. 3.1 in [6]) allows us to express the rate function in terms of the Hamiltonian $H(\xi) \doteq D(\xi) + \Phi(\xi)$:

$$\begin{aligned} J(\xi) &= D(\xi) + \Phi(\xi) + \sigma Q(\xi) \\ &\quad - \inf\{D(\xi) + \Phi(\xi) + \sigma Q(\xi) : \xi \in L_c^2(\rho)\} \\ &= H(\xi) + \sigma Q(\xi) - \inf\{H(\xi) + \sigma Q(\xi) : \xi \in L_c^2(\rho)\}. \end{aligned} \tag{4.9}$$

We have proved the following result.

Theorem 4.4. *As $n \to \infty$, the sequence $P_{n,n\beta}(\psi^{(n)} \in d\xi)$ satisfies the LDP on $L_c^2(\rho)$ with the scaling constants $n\beta$ and the rate function J defined in* (4.9).

We now address the question of when the conditional measure $P_{n,n\beta}^{N,\varepsilon}$ is well defined. To this end, we introduce

$$\tilde{E}(N) \doteq \inf\{J(\xi) : \xi \in L_c^2(\rho), Q(\xi) = N\}, \tag{4.10}$$

which is nonnegative and finite for all $N \in [0, \infty)$. Because Q is a continuous functional on $L_c^2(\rho)$, the preceding theorem and the contraction principle imply that $\tilde{E}$ is a rate function—and thus is lower semicontinuous—and also yield the LDP stated in part (a) of the next proposition [11, Thm. 4.2.1]. If one applies the large deviation lower bound to the open set $(N-\varepsilon, N+\varepsilon) \subset [N-\varepsilon, N+\varepsilon]$ and uses the bound $\tilde{E}((N-\varepsilon, N+\varepsilon)) \leq \tilde{E}(N)$, then one obtains part (b). Part (b) implies that for any $N \in [0, \infty)$, all sufficiently large n, and all $\varepsilon > 0$ we have

$$P_{n,n\beta}\{Q(\psi^{(n)}) \in [N - \varepsilon, N + \varepsilon]\} > 0. \tag{4.11}$$

Hence for these values of N, n, and ε, the conditional measure $P_{n,n\beta}^{N,\varepsilon}$ is well defined.

Proposition 4.5. (a) *As $n \to \infty$, the sequence $P_{n,n\beta}\{Q(\psi^{(n)}) \in dx\}$ satisfies the LDP on $\mathbb{R}$ with the scaling constants $n\beta$ and the rate function $\tilde{E}$. In particular, $\tilde{E}$ is nonnegative and lower semicontinuous on $\mathbb{R}$.*
(b) *For $N \in [0, \infty)$ and any $\varepsilon > 0$,*

$$\liminf_{n\to\infty} \frac{1}{n\beta} \log P_{n,n\beta}\{Q(\psi^{(n)}) \in [N - \varepsilon, N + \varepsilon]\} \geq -\tilde{E}(N) > -\infty.$$

We indicate other expressions for $\tilde{E}$. Substituting into the definition of $\tilde{E}$ the formula (4.9) for J, we obtain

$$\begin{aligned} \tilde{E}(N) &\doteq \inf\{J(\xi) : \xi \in L_c^2(\rho), Q(\xi) = N\} \\ &= \inf\{H(\xi) + \sigma Q(\xi) : \xi \in L_c^2(\rho), Q(\xi) = N\} \\ &\quad - \inf\{H(\xi) + \sigma Q(\xi) \ : \ \xi \in L_c^2(\rho)\} \\ &= \inf\{H(\xi) : \xi \in L^2(\rho), Q(\xi) = N\} + \sigma N \\ &\quad - \inf\{H(\xi) + \sigma Q(\xi) : \xi \in L_c^2(\rho)\}. \end{aligned}$$

Recalling the function $\bar{E}(N) \doteq \inf\{H(\xi) : \xi \in L_c^2(\rho), Q(\xi) = N\}$ introduced in (3.1), we see that

$$\tilde{E}(N) = \bar{E}(N) + \sigma N - \inf\{H(\xi) + \sigma Q(\xi) : \xi \in L_c^2(\rho)\}. \tag{4.12}$$

Since $\tilde{E}$ is nonnegative and lower semicontinuous, it follows that $\bar{E}$ is bounded below and lower semicontinuous.

We now complete the proof of Theorem 3.1, which states the LDP for the $P_{n,n\beta}^{N,\varepsilon}$-distributions of $\psi^{(n)}$. This is carried out by proving that as $n \to \infty$ and $\varepsilon \to 0$ the sequence $P_{n,n\beta}^{N,\varepsilon}(\psi^{(n)} \in d\xi)$ satisfies the LDP on $L_c^2(\rho)$ with the scaling constants $n\beta$ and the rate function J^N defined in (3.2). The function Q defining the conditioning in this measure is continuous. If it were also bounded, then the desired LDP would be a consequence of Theorem 3.2 in our paper [17]. However, a quick examination of the proof of that theorem reveals that only the continuity of Q is required, not its boundedness (specifically, in the application of the contraction principle in the proof of Prop. 3.1 in [17]). For $N \in [0, \infty)$, Theorem 3.2 in [17], with its proof modified as just described, shows that as $n \to \infty$ and $\varepsilon \to 0$ the sequence $P_{n,n\beta}^{N,\varepsilon}(\psi^{(n)} \in d\xi)$ satisfies the LDP on $L_c^2(\rho)$ with the scaling constants $n\beta$ and the rate function

$$\tilde{J}^N(\xi) \doteq \begin{cases} J(\xi) - \tilde{E}(N) & \text{if } Q(\xi) = N \\ \infty & \text{otherwise.} \end{cases}$$

Substituting the formula (4.9) for J into the definition of $\tilde{J}^N$ and using (4.12) to relate $\tilde{E}$ and $\bar{E}$, we see that $\tilde{J}^N$ equals the function J^N defined in (3.2). This yields the desired LDP for $P_{n,n\beta}^{N,\varepsilon}(\psi^{(n)} \in d\xi)$. The proof of Theorem 3.1 is concluded.

5. Monte-Carlo Simulations

Here we summarize some numerical computations that display the variety of behaviors that are exhibited by the mixed Gibbs ensembles $P_{n,n\beta}^{N,\varepsilon}(d\psi^{(n)})$ in the continuum limit as $n \to \infty$. In particular, we implement a Monte-Carlo sampling procedure to probe the dependence of the concentration phenomenon on Condition 2.1. To this end we consider operators L on $L_c^2[0, \pi]$ with Dirichlet boundary conditions and with eigenfunctions $e_k(x) = \sqrt{2}\sin(kx)$, and we choose the corresponding eigenvalues to be $\lambda_k = -k^\alpha$ for $0 < \alpha < +\infty$. Our main concern is to distinguish the case when $\alpha > 1$, for which Condition 2.1 holds and hence our LDP applies, from the case when $\alpha \leq 1$, for which a concentration behavior may or may not occur.

Our Monte-Carlo procedure is a modification of the standard Metropolis algorithm [2] appropriate to the mixed ensemble, which is canonical with respect to the energy H_n and microcanonical with respect to the particle number Q_n. The microcanonical constraint is enforced exactly at each step of the Markov chain that defines the Metropolis algorithm; at each step two components of the random state $\psi \in \Gamma_n$ are updated in a manner that preserves the spherical constraint, $\sum_1^n |\psi_k|^2 = 2N$. To improve the sampling properties of the algorithm, a form of simulated annealing is used. That is, the sampling procedure is implemented in two stages: first at a small β (high temperature) and then at the prescribed β (given temperature).

To exhibit the concentration behavior predicted by the LDP for $\psi^{(n)}$, we sample the mixed Gibbs ensemble $P_{n,n\beta}^{N,\varepsilon}$ (with $\varepsilon \to 0+$) for the three values $n = 16, 64, 256$ with β fixed. In all our computations the underlying GNLS equation has the saturated nonlinearity

$$f(|\psi|^2) = \frac{b\,|\psi|^2}{(1 + |\psi|^2)} \tag{5.1}$$

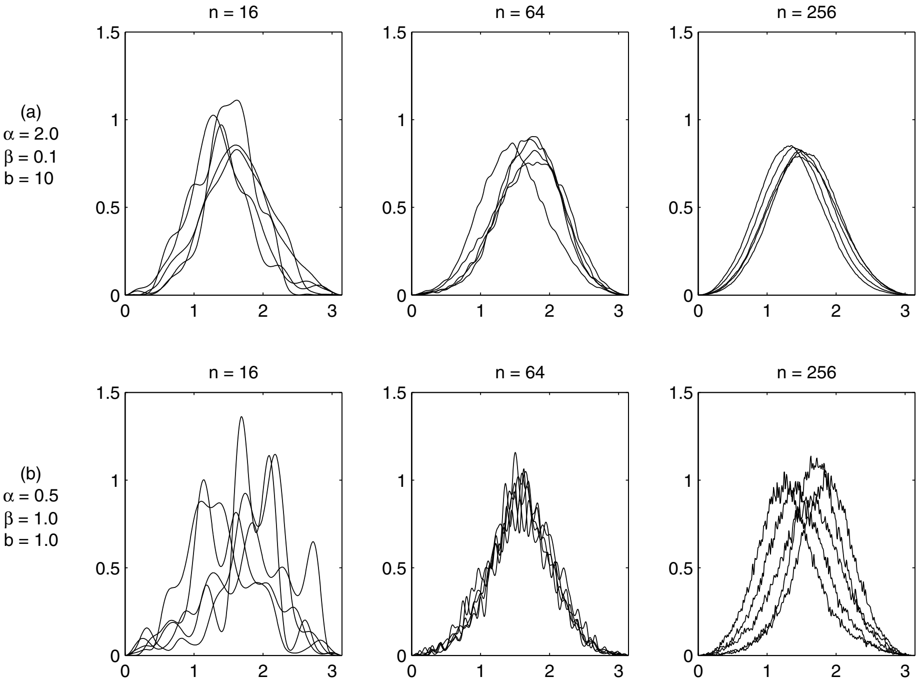

Fig. 1a,b. Samples from mixed Gibbs ensemble for $\alpha = 2.0$ and $\alpha = 0.5$

with scale factor b. First we set $\alpha = 2.0$, $\beta = 0.1$, and $b = 10$. These parameters are chosen to yield a ground state of approximately unit amplitude and unit width. Figure 1a displays the three plots for this sequence, each plot composed of five representative samples of $|\psi|$ drawn from the Monte-Carlo sampled ensemble. The expected behavior under Condition 2.1 is clearly demonstated; namely, the fluctuations visible for $n = 16$ decrease for $n = 64$ and almost disappear for $n = 256$, for which all five samples remain close to the ground state. Next we set $\alpha = 0.5$, $\beta = 1.0$, and $b = 1.0$. This choice of α furnishes an example of the behavior of the mixed Gibbs ensemble in the continuum limit when Condition 2.1 is violated. As in Fig. 1a, the parameters β and b are fixed so that the ground state has approximately unit amplitude and unit width. The sequence of three plots displayed in Fig.1b shows greater fluctuations than the corresponding plots in Fig.1a. As n increases, both the spatial scale and the magnitude of the fluctuations decreases, so that for $n = 256$ each of the five samples is a near-ground state on the large scale having small fluctuations on the small scales.

Figure 1b appears to show a concentration around the ground state even in the case when $\alpha = 0.5$, but possibly with a slower rate of convergence than for $\alpha > 1$, and possibly in a weaker norm than the $L^2_c(\rho)$-norm. We do not know whether Condition 2.1 is necessary as well as sufficient for an LDP for the process $\psi^{(n)}$ in the $L^2_c(\rho)$-topology. On the one hand, the displayed computations for $\alpha < 1$ and other computed results for $\alpha < 1$ not included here suggest that a weaker condition may be sufficient for such an LDP. On the other hand, the proof of Theorem 4.1, the basic LDP for Gaussian processes on which our main Theorem 3.1 is based, is not valid for $\alpha < 1$. This suggests that the concentration property breaks down when $\alpha < 1$. A more exhaustive numerical investigation might help to resolve this question, but such a computationally intensive study will not be pursued here.

A second set of numerical computations is shown in Fig. 2. Here we investigate the dependence of the concentration phenomenon and the associated ground states on the strength of the nonlinearity in the GNLS equation. Specifically, we vary the parameter $b > 0$ in the saturated nonlinearity given in (5.1). Since the GNLS equation is focusing, an increase of b is expected to localize and intensify the ground state solitary wave. In Fig. 2a this effect is displayed for $b = 10, 20, 100$, with $\alpha = 2.0$, $\beta = 10$ and $n = 64$; in Fig. 2b it is displayed for $b = 1, 2, 10$, with $\alpha = 0.5$, $\beta = 100$ and $n = 64$. As in Fig. 1, these plots exhibit five samples of $|\psi|$ for each ensemble. Since β is relatively large in each of these cases, each sample has the shape of a ground state. For large b, whether or not the summability of $1/\lambda_k$ in Condition 2.1 holds, we observe a ground state that is a highly localized solitary wave; an extreme case of this is given in Fig. 2b for $b = 10$.

From the point of view of the present paper, the most noteworthy effect displayed in Fig. 2 is the presence of approximate translates of the exact ground state among the Monte Carlo samples. Indeed, for high β (low temperature) the energy of all the displayed samples is close to the ground state energy, and the samples themselves are often close to being translates of the exact ground state. This effect is straightforwardly explained by the fact that, for relatively localized states, the energy H is only slightly different among all translations of a ground state that are not too close to the boundary. According to our main LDP in Theorem 3.1, the rate function is simply the energy difference between a candidate state and the ground state(s) [see (3.2)]. Thus, while we have adopted Dirichlet

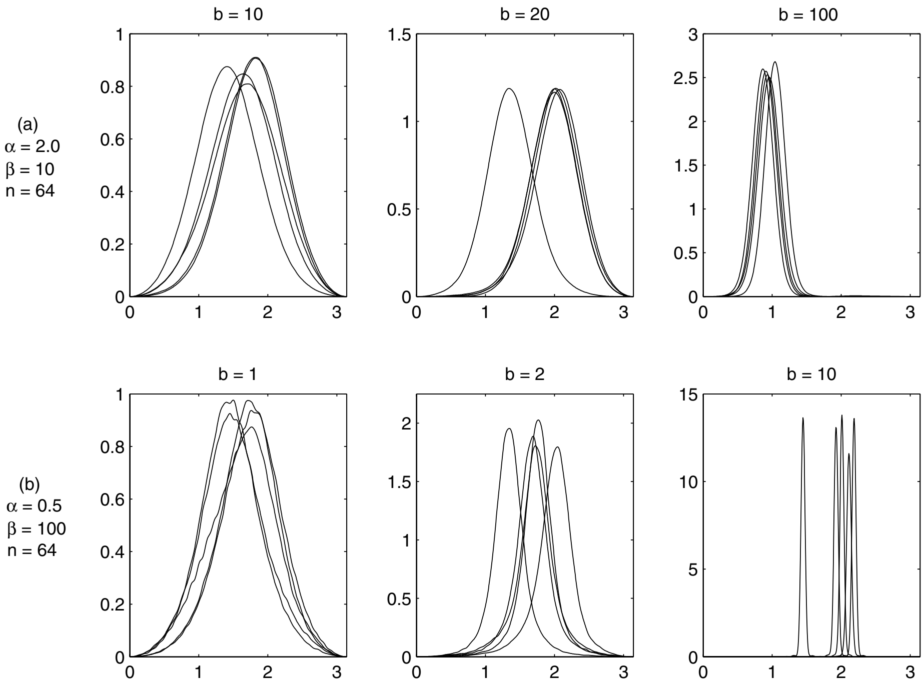

Fig. 2a,b. Samples from mixed Gibbs ensemble for increasing values of b

boundary conditions to reduce the translational invariance of the GNLS equation and thereby to simplify the presentation of our results, we have found that in the strongly focusing regime an approximate translational invariance persists. Of course, this effect diminishes as the Monte Carlo simulations are carried out for increasing n.

In summary, the concentration phenomenon that is precisely expressed in our main LDP is definitely borne out by numerical sampling of the mixed ensembles over a wide range of parameters, even though this phenomenon is somewhat complicated by the near translation-invariance of localized ground states. In fact, the range of parameters for which the LDP holds may be wider than the range covered by our main theorem.

Acknowledgement. The authors thank Adam Eisner, who provided the Monte-Carlo method and code used in Sect. 5 to carry out the simulations of the mixed ensemble.

References

1. Bidegaray, B.: Invariant measures for some partial differential equations. Physica D **82**, 340–364 (1995)
2. Binder, K., Heermann, D.W.: *Monte Carlo Simulation in Statistical Physics.* Fourth edition. Springer Series in Solid-State Sciences, Vol. **80**, Berlin: Springer-Verlag, 2002
3. Birkhoff, G., Rota, G.-C.: *Ordinary Differential Equations.* Second edition. Waltham: Blaisdell Publishing Co., 1969
4. Biskamp, D.: *Nonlinear Magnetohydrodynamics.* Cambridge Monographs in Plasma Physics. Cambridge: Cambridge Univ. Press,1993
5. Bolthausen, E.: On the probability of large deviations in Banach spaces. Ann. Probab. **12**, 427–435 (1984)
6. Boucher, C., Ellis, R.S., Turkington, B.: Derivation of maximum entropy principles in two-dimensional turbulence via large deviations. J. Stat. Phys **98**, 1235–1278 (2000)
7. Bourgain, J.: Periodic nonlinear Schrödinger equation and invariant measures. Commun. Math. Phys. **166**, 1–26 (1994)
8. Bouchet, F., Sommeria, J.: Emergence of intense jets and Jupiter's Great Red Spot as maximum-entropy structures. J. Fluid Mech. **464**, 165–207 (2002)
9. Cai, D., Majda, A.J., McLaughlin, D.W., Tabak, E.G.: Spectral bifurcations in dispersive wave turbulence. Proc. Nat. Acad. Sci. **96**, 14216–14221 (1999)
10. Cai, D., McLaughlin, D.W.: Chaotic and turbulent behavior of unstable 1D nonlinear dispersive waves. J. Math. Phys. **41**, 4125–4153 (2000)
11. Dembo, A., Zeitouni, O.: *Large Deviations Techniques and Applications*. Second edition. New York: Spring-Verlag, 1998
12. DiBattista, M.T., Majda, A.J., Grote, M.J.: Meta-stability of equilibrium statistical structures for prototype geophysical flows with damping and driving. Physica D **151**, 271–304 (2001)
13. Dowling, T.E.: Dynamics of Jovian atmospheres. Ann. Rev. Fluid Mech. **27**, 293–334 (1995)
14. Dupuis, P., Ellis, R.S.: *A Weak Convergence Approach to the Theory of Large Deviations*. New York: John Wiley & Sons, 1997
15. Dyachenko, S., Zakharov, V.E., Pushkarev, A.N., Shvets, V.F., Yan'kov, V.V.: Soliton turbulence in nonintegrable wave systems. Soviet Phys. JETP **69**, 1144–1147 (1989)
16. Ellis, R.S.: *Entropy, Large Deviations and Statistical Mechanics*. New York: Springer-Verlag, 1985
17. Ellis, R.S., Haven, K., Turkington, B.: Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles. J. Stat. Phys. **101**, 999–1064 (2000)
18. Gikhman, I.I., Skorohod, A.V.: *The Theory of Stochastic Processes I.* Trans. by S. Kotz, Berlin: Springer-Verlag, 1974
19. Hasegawa, A.: Self-organization processes in continuous media. Adv. Phys. **34**, 1–42 (1985)
20. Isichenko, M.B., Gruzinov, A.V.: Isotopological relaxation, coherent structures, and Gaussian turbulence in two-dimensional magnetohydrodynamics. Phys. Plasmas **1**, 1802–1816 (1994)
21. Itô, K., McKean, H.P.: *Diffusion Processes and Their Sample Paths*. New York/Berlin: Academic Press/Springer Verlag, 1965
22. Jordan, R., Josserand, C.: Self-organization in nonlinear wave turbulence. Phys. Rev. E **61**, 1527–1539 (2000)
23. Jordan, R., Josserand, C.: Statistical equilibrium states for the nonlinear Schrödinger equation. Math. Comp. Simulation **55**, 433–447 (2001)

24. Jordan, R., Turkington, B.: Ideal magnetofluid turbulence in two dimensions J. Stat. Phys. **87**, 661–695 (1997)
25. Jordan, R., Turkington, B., Zirbel, C.L.: A mean-field statistical theory for the nonlinear Schrödinger equation. Physica D **137**, 353–378 (2000)
26. Kevrekidis, P.G., Rasmussen, K.O., Bishop, A.R.: The discrete nonlinear Schrödinger equation: A survey of recent results. Int. J. Mod. Phys. B. **15**, 2833–2900 (2001)
27. Lebowitz, J.L., Rose, H.A., Speer, E.R.: Statistical mechanics of a nonlinear Schrödinger equation. J. Stat. Phys., **50**, 657–687 (1988)
28. Majda, A.J., McLaughlin, D.W., Tabak, E.G.: A one-dimensional model for dispersive wave turbulence. J. Nonlinear Sci. **7**, 9–44 (1997)
29. Marcus, P.S.: Jupiter's Great Red Spot and other vortices. Annual Rev. Astronomy and Astrophys. **31**, 523–573 (1993)
30. McKean, H.P.: Statistical mechanics of nonlinear wave equations IV. Cubic Schrödinger. Commun. Math. Phys.**168**, 479–491 (1995)
31. Rasmussen, J.J., Rypdal, K.: Blow-up in nonlinear Schroedinger equations–I: A general review. Physica Scripta **33**, 481–504 (1986)
32. Segre, E., Kida, S.: Late states of incompressible 2d decaying vorticity fields. Fluid Dyn. Res. **23**, 89–112 (1998)
33. Turkington, B., Majda, A.J., Haven, K., DiBattista, M.: Statistical equilibrium predictions of jets and spots on Jupiter. Proc. Nat. Acad. Sci. USA **98**, 12346–12350 (2001)
34. Zakharov, V.E., Pushkarev, A.N., Shvets, V.F., Yan'kov, V.V.: Soliton turbulence. JETP Lett. **48**, 83–86 (1988)
35. Zhidkov, P.E.: On an invariant measure for a nonlinear Schrödinger equation. Soviet Math. Dokl. **43**, 431–434 (1991)

Communicated by P. Constantin