

# A Statistical Dynamical Model to Predict Extreme Events and Anomalous Features in Shallow Water Waves with Abrupt Depth Change

Andrew J. Majda<sup>a,1</sup>, M. N. J. Moore<sup>b</sup>, and Di Qi<sup>a,1</sup>

<sup>a</sup>Department of Mathematics and Center for Atmosphere and Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, NY 10012;

<sup>b</sup>Department of Mathematics and Geophysical Fluid Dynamics Institute, Florida State University, Tallahassee, FL 32306

This manuscript was compiled on November 28, 2018

**Understanding and predicting extreme events and their anomalous statistics in complex nonlinear systems is a grand challenge in climate, material, and neuroscience, as well as for engineering design. Recent laboratory experiments in weakly turbulent shallow water reveal a remarkable transition from Gaussian to anomalous behavior as surface waves cross an abrupt depth change (ADC). Downstream of the ADC, PDFs of surface displacement exhibit strong positive skewness, accompanied by an elevated level of extreme events. Here we develop a statistical dynamical model to explain and quantitatively predict the above anomalous statistical behavior as experimental control parameters are varied. The first step is to use incoming and outgoing truncated Korteweg-de Vries (TKdV) equations matched in time at the ADC. The TKdV equation is a Hamiltonian system which induces incoming and outgoing statistical Gibbs invariant measures. The statistical matching of the known nearly Gaussian incoming Gibbs state at the ADC completely determines the predicted anomalous outgoing Gibbs state, which can be calculated by a simple sampling algorithm, verified by direct numerical simulations, and successfully captures key features of the experiment. There is even an analytic formula for the anomalous outgoing skewness. The strategy here should be useful for predicting extreme anomalous statistical behavior in other dispersive media in different settings (21, 22).**

extreme anomalous event | statistical TKdV model | matching Gibbs measures | surface wave displacement and slope

**U**nderstanding and predicting extreme events and their anomalous statistics in complex nonlinear systems is a grand challenge in climate, material and neuroscience as well as for engineering design. This is a very active contemporary topic in applied mathematics with qualitative and quantitative models (1–7) and novel numerical algorithms which overcome the curse of dimensionality for extreme event prediction in large complex systems (2, 8–11). The occurrence of Rogue waves as extreme events within different physical settings of deep water (12–16) and shallow water (17–19) is an important practical topic.

Recent laboratory experiments in weakly turbulent shallow water reveal a remarkable transition from Gaussian to anomalous behavior as surface waves cross an abrupt depth change (ADC). A normally-distributed incoming wave train, upstream of the ADC, transitions to a highly non-Gaussian outgoing wave train, downstream of the ADC, that exhibits large positive skewness of the surface height and more frequent extreme events (20). Here we develop a statistical dynamical model to explain and quantitatively predict this anomalous behavior as experimental control parameters are varied. The first step is to use incoming and outgoing truncated Korteweg-de Vries (TKdV) equations matched in time at the ADC. The TKdV

equation is a Hamiltonian system which induces incoming and outgoing Gibbs invariant measures. The statistical matching of the known nearly Gaussian incoming Gibbs state at the ADC completely determines the predicted anomalous outgoing Gibbs state, which can be calculated by a simple MCMC algorithm, verified by direct numerical simulations, and successfully captures key features of the experiment. There is even an analytic formula for the anomalous outgoing skewness. The strategy here should be useful for predicting extreme anomalous statistical behavior in other dispersive media in different settings (21, 22).

## 1. Experiments showing anomalous wave statistics induced by an abrupt depth change

Controlled laboratory experiments were carried out in (20) to examine the statistical behavior of surface waves crossing an ADC. In these experiments, nearly unidirectional waves are generated by a paddle wheel and propagate through a long, narrow wave tank. Midway through, the waves encounter a step in the bottom topography, and thus abruptly transition to a shallower depth. The paddle wheel is forced with a pseudo-random signal intended to mimic a Gaussian random sea upstream of the ADC. In particular, the paddle angle is

### Significance Statement

**Understanding and predicting extreme events and their anomalous statistics in complex nonlinear systems is a grand challenge in applied sciences as well as for engineering design. Recent controlled laboratory experiments in weakly turbulent shallow water with abrupt depth change exhibit a remarkable transition from nearly Gaussian statistics to extreme anomalous statistics with large positive skewness of the surface height. We develop a statistical dynamical model to explain and quantitatively predict the anomalous statistical behavior. The incoming and outgoing waves are modeled by the truncated Korteweg-de Vries equations statistically matched at the depth change. The statistical matching of the known nearly Gaussian incoming Gibbs state completely determines the predicted anomalous outgoing Gibbs state, and successfully captures key features of the experiment.**

A.J.M. designed the research. A.J.M., M.N.J.M., and D.Q. performed the research. A.J.M., M.N.J.M., and D.Q. wrote the paper.

The authors declare no conflict of interest.

<sup>1</sup>To whom correspondence should be addressed. E-mail: qidi@cims.nyu.edu, jonjon@cims.nyu.edu

125 specified as

$$126 \quad \theta(t) = \theta_0 + \Delta\theta \sum_{n=1}^N a_n \cos(\omega_n t + \delta_n), \quad E \sim (\Delta\theta)^2 \sum_n a_n^2 \omega_n^2.$$

129 where the weights  $a_n$  are Gaussian in spectral space with  
 130 peak frequency  $\omega_p$  and the phases  $\delta_n$  are uniformly distributed  
 131 random variables. The peak frequency gives rise to a char-  
 132 acteristic wavelength  $\lambda_c$  which can be estimated from the  
 133 dispersion relation. The energy  $E$  injected into the system  
 134 is determined by the angle amplitude  $\Delta\theta$ , which is the main  
 135 control parameter varied in (20). Optical measurements of the  
 136 free surface reveal a number of surprising statistical features:  
 137

- 138 • Distinct statistics are found between the incoming and  
 139 outgoing wave disturbances: incoming waves display near-  
 140 Gaussian statistics, while outgoing waves skew strongly  
 141 towards positive displacement.
- 142 • The degree of non-Gaussianity present in the outgoing  
 143 waves depends on the injected energy  $E$ : larger energies  
 144 generate stronger skewness in the surface displacement  
 145 PDFs and more extreme events.
- 146 • Compared to the incoming wave train, the power spectrum  
 147 of the outgoing wave field decays more slowly, which  
 148 indicates that the anomalous behavior is associated with  
 149 an elevated level of high frequencies.

## 152 2. Surface wave turbulence modeled by truncated KdV 153 equation with depth dependence

154 The Korteweg-de Vries (KdV) equation is a one-dimensional,  
 155 deterministic model capable of describing surface wave tur-  
 156 bulence. More specifically, KdV is leading-order approxima-  
 157 tion for surface waves governed by a balance of nonlinear  
 158 and dispersive effects, valid in an appropriate far-field limit  
 159 (23). Moreover, KdV has been adapted to describe waves  
 160 propagating over variable depth (23). Here, we consider the  
 161 variable-depth KdV equation truncated at wavenumber  $\Lambda$   
 162 (with  $J = 2\Lambda + 1$  grid points) in order to generate weakly  
 163 turbulent dynamics. The surface displacement is described  
 164 by the state variable  $u_\Lambda^\pm(x, t)$  with superscript ‘-’ for the  
 165 incoming waves and ‘+’ for the outgoing waves. The Galerkin  
 166 truncated variable  $u_\Lambda = \sum_{1 \leq |k| \leq \Lambda} \hat{u}_k(t) e^{ikx}$  is normalized  
 167 with zero mean  $\hat{u}_0 = 0$  and unit energy  $2\pi \sum_{k=1}^\Lambda |\hat{u}_k|^2 = 1$ ,  
 168 which are conserved quantities. Here,  $u_\Lambda \equiv \mathcal{P}_\Lambda u$  denotes the  
 169 subspace projection. The evolution of  $u_\Lambda^\pm$  is governed by the  
 170 truncated KdV equation with depth change  $D_\pm$

$$171 \quad \frac{\partial u_\Lambda^\pm}{\partial t} + \frac{D_\pm^{-3/2}}{2} E_0^{1/2} L_0^{-3/2} \frac{\partial}{\partial x} \mathcal{P}_\Lambda (u_\Lambda^\pm)^2 + D_\pm^{1/2} L_0^{-3} \frac{\partial^3 u_\Lambda^\pm}{\partial x^3} = 0. \quad [1]$$

172 Equation [1] is non-dimensionalized on the periodic domain  $x \in$   
 173  $[-\pi, \pi]$ . The depth is assumed to be unit  $D_- = 1$  before the  
 174 ADC and  $D_+ < 1$  after the ADC. The conserved Hamiltonian  
 175 can be decomposed as

$$176 \quad \mathcal{H}_\Lambda^\pm = D_\pm^{-3/2} E_0^{1/2} L_0^{-3/2} H_3(u_\Lambda^\pm) - D_\pm^{1/2} L_0^{-3} H_2(u_\Lambda^\pm),$$

$$177 \quad H_3(u) = \frac{1}{6} \int_{-\pi}^{\pi} u^3 dx, \quad H_2(u) = \frac{1}{2} \int_{-\pi}^{\pi} \left( \frac{\partial u}{\partial x} \right)^2 dx.$$

178 where we refer to  $H_3$  as the cubic term and  $H_2$  the quadratic  
 179 term. We introduce parameters  $(E_0, L_0, \Lambda)$  based on the fol-  
 180 lowing assumptions:

- 181 • The wavenumber truncation  $\Lambda$  is fixed at a moderate  
 182 value for generating weakly turbulent dynamics. 187  
 183
- 184 • The state variable  $u_\Lambda^\pm$  is normalized with zero mean and  
 185 unit energy,  $\mathcal{M}(u_\Lambda) = 0, \mathcal{E}(u_\Lambda) = 1$ , which are conserved  
 186 during evolution. Meanwhile,  $E_0$  characterizes the total  
 187 energy injected into the system based on the driving  
 188 amplitude  $\Delta\theta$ . 189
- 189 • The length scale of the system  $L_0$  is chosen so that the  
 190 resolved scale  $\Delta x = 2\pi L_0/J$  is comparable to the the  
 191 characteristic wave length  $\lambda_c$  from the experiments. 192

192 Some intuition for how equation [1] produces different dynam-  
 193 ics on either side of the ADC can be gained by considering  
 194 the relative contributions of  $H_3$  and  $H_2$  in the Hamiltonian  
 195  $\mathcal{H}_\Lambda^\pm$ . The depth change,  $D_+ < 1$ , increases the weight of  $H_3$   
 196 and decreases that of  $H_2$ , thereby promoting the effects of  
 197 nonlinearity over dispersion and creating conditions favorable  
 198 for extreme events. Since  $\frac{\partial u}{\partial x}$  is the slope of the wave height,  
 199  $H_2(u)$  measures the wave slope energy. 200

200 A *deterministic matching condition* is applied to the surface  
 201 displacement  $u_\Lambda^\pm$  to link the incoming and outgoing wave trains.  
 202 Assuming the abrupt depth change is met at  $t = T_{\text{ADC}}$ , the  
 203 matching condition is given by 204

$$204 \quad u_\Lambda^-(x, t) |_{t=T_{\text{ADC}}} = u_\Lambda^+(x, t) |_{t=T_{\text{ADC}}} +,$$

205 Equation [1] is not designed to capture the short scale changes  
 206 in rapid time. Rather, since we are interested in modeling  
 207 statistics before and after the ADC, we will examine the long-  
 208 time dynamics of large-scale structures. 209

### 210 Interpreting experimental parameters in the dynamical model.

211 The model parameters  $(E_0, L_0, \Lambda)$  in [1] can be directly linked  
 212 to the basic scales from the physical problem. The important  
 213 parameters that characterize the experiments (20) include:  
 214  $\epsilon = \frac{a}{H_0}$  the wave amplitude  $a$  to (upstream) water depth  $H_0$   
 215 ratio;  $\delta = \frac{H_0}{\lambda_c}$  the water depth to wavelength  $\lambda_c$  ratio; and  
 216  $D_0 = \frac{d}{H_0}$  the depth ratio with upstream value  $d = H_0$  and  
 217 downstream value  $d < H_0$ . By comparing the characteristic  
 218 physical scales, the normalized TKdV equation [1] can be  
 219 linked to these experimental parameters via

$$220 \quad L_0 = 6^{\frac{1}{3}} \left( M \epsilon^{\frac{1}{2}} \delta^{-1} \right), \quad E_0 = \frac{27}{2} \gamma^{-2} \left( M \epsilon^{\frac{1}{2}} \delta^{-1} \right), \quad [2]$$

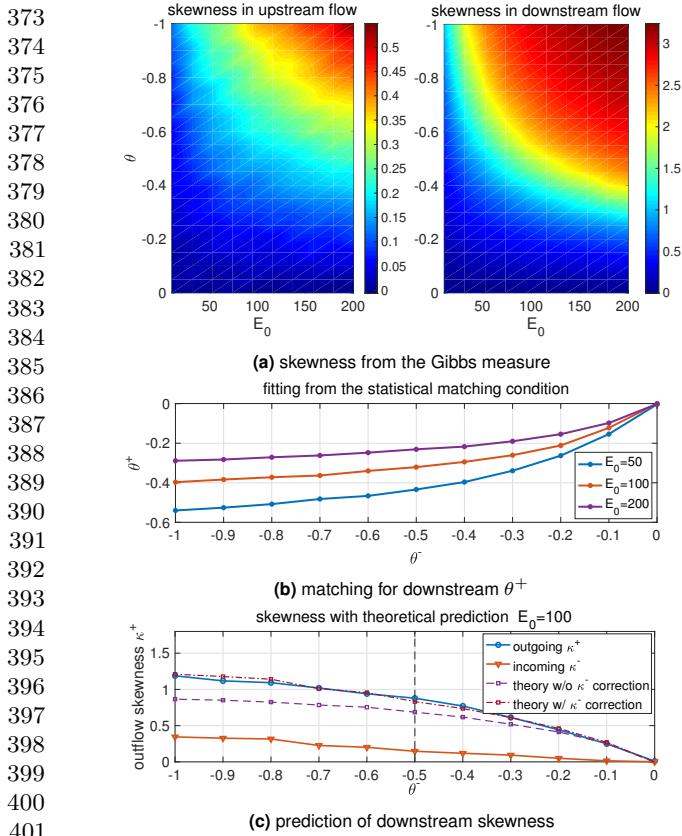
221 where  $M$  defines the computational domain size  $M\lambda_c$  as  $M$ -  
 222 multiple of the characteristic wavelength  $\lambda_c$ , and  $\gamma = \frac{U}{a}$  with  
 223  $U$  a scaling factor for the state variable  $u_\Lambda$  normalizes the  
 224 total energy of the system to one. 225

225 Consider a spatial discretization with  $J = 2\Lambda + 1$  grid  
 226 points, so that the smallest resolved scale is comparable to  
 227 the characteristic wavelength 228

$$228 \quad \Delta x = \frac{2\pi M \lambda_c}{J} \lesssim \lambda_c \Rightarrow M = \frac{J}{2\pi} \sim 5, \quad J = 32. \quad [2]$$

229 In the practical numerical simulations, we select  $M = 5$  and  
 230 let  $\gamma$  vary in the range  $[0.5, 1]$ . Using reference values from the  
 231 experiments, (20), we find  $\epsilon \in [0.0024, 0.024]$ ,  $\delta \sim 0.22$ , and  
 232  $D_0$  changes from 1 to 0.24 at the ADC. These values yield  
 233 the following ranges for the model parameters:  $L_0 \in [2, 6]$  and  
 234  $E_0 \in [50, 200]$ . These are the values we will test in the direct  
 235 numerical simulations. See details about the derivation from  
 236 scale analysis in *SI Appendix, A*. 237

<p>249 <b>3. Equilibrium statistical mechanics for generating the</b> 250 <b>stationary invariant measure</b></p> <p>252 Since the TKdV equation satisfies the Liouville property, the 253 equilibrium invariant measure can be described by a statistical 254 formalism (24–26) based on a Gibbs measure with the 255 conserved energy <math>\mathcal{E}_\Lambda</math> and Hamiltonian <math>\mathcal{H}_\Lambda</math>. The equilibrium 256 invariant measure is dictated by the conservation laws in the 257 TKdV equation. In the case of fixed energy <math>E_0</math>, this is the 258 <i>mixed Gibbs measure</i> with microcanonical energy and canonical 259 Hamiltonian ensembles (24)</p> <p>260 <math>\mathcal{G}_\theta^\pm(u_\Lambda^\pm; E_0) = C_\theta^\pm \exp(-\theta^\pm \mathcal{H}(u_\Lambda^\pm)) \delta(\mathcal{E}(u_\Lambda^\pm) - E_0), [3]</math></p> <p>262 where <math>\theta</math> represents the “inverse temperature”. The distinct 263 statistics in the upstream and downstream waves can be con- 264 trolled <math>\theta</math>. As shown below, we find that negative temperature, 265 <math>\theta^\pm &lt; 0</math>, is the appropriate regime to predict the experiments. 266 In the incoming wave field, <math>\theta^-</math> is chosen so that <math>\mathcal{G}_\theta^-</math> has nearly 267 Gaussian statistics. Using the above invariant measures [3], 268 the expectation of any functional <math>F(u)</math> can be computed as</p> <p>269</p> <p>271 <math>\langle F \rangle_{\mathcal{G}_\theta} \equiv \int F(u) \mathcal{G}_\theta(u) du.</math></p> <p>273</p> <p>274 The value of <math>\theta</math> in the invariant measure is specified from <math>\langle H_\Lambda \rangle_{\mathcal{G}_\theta}</math> 275 (24, 26).</p> <p>276 In addition to producing equilibrium PDFs of <math>u_\Lambda</math>, the in- 277 variant measure can be used to predict the equilibrium energy 278 spectrum without the direct simulation of TKdV. Direct sim- 279 ulation, however, is required to recover transient statistics of 280 <math>u_\Lambda</math> and time autocorrelations.</p> <p>281</p> <p>282 <b>Statistical matching condition of the invariant measures be-</b> 283 <b>fore and after the abrupt depth change.</b> The Gibbs measures 284 <math>\mathcal{G}_\theta^\pm</math> are defined based on the different inverse temperatures <math>\theta^\pm</math> 285 on the two sides of the solutions</p> <p>286</p> <p>287 <math>\mu_t^-(u_\Lambda^-; D_-), \quad u_\Lambda _{t=T_{ADC}-} = u_0, \quad t &lt; T_{ADC};</math> 288 <math>\mu_t^+(u_\Lambda^+; D_+), \quad u_\Lambda _{t=T_{ADC}+} = u_0, \quad t &gt; T_{ADC},</math></p> <p>289</p> <p>290 where <math>u_0</math> represents the deterministic matching condition be- 291 tween the incoming and outgoing waves. The two distributions, 292 <math>\mu_t^-, \mu_t^+</math> should be matched at <math>T_{ADC}</math> giving</p> <p>293</p> <p>294 <math>\mu_{t=T_{ADC}}^-(u_\Lambda) = \mu_{t=T_{ADC}}^+(u_\Lambda).</math></p> <p>295</p> <p>296 In matching the flow statistics before and after the abrupt 297 depth change, we first use the conservation of the determinis- 298 tic Hamiltonian <math>H_\Lambda^+</math> after the depth change. Then assuming 299 ergodicity (24, 25), the statistical expectation for the Hamil- 300 tonian <math>\langle H_\Lambda^+ \rangle</math> is conserved in time after the depth change at 301 <math>t = T_{ADC}</math> and should remain at this value as the system ap- 302 proaches equilibrium as <math>t \rightarrow \infty</math>. The final statistical matching 303 condition to get the outgoing flow statistics with parameter 304 <math>\theta^+</math> can be found by</p> <p>305</p> <p>306 <math>\langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^+} = \langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^-}, [4]</math></p> <p>307</p> <p>308 with the outgoing flow Hamiltonian <math>H_\Lambda^+</math> and the Gibbs mea- 309 sures <math>\mathcal{G}_\theta^\pm</math> before and after the abrupt depth change.</p>	<p><b>4. The nearly Gaussian incoming statistical state</b></p> <p>For the parameters explored in (20), the incoming wave field is always characterized by a near-Gaussian distribution of the surface displacement. It is found that a physically consistent Gibbs measure should take negative values in the inverse temperature parameter <math>\theta &lt; 0</math>, where a proper distribution function and a decaying energy spectrum are generated (see (26) and <i>SI Appendix, B.1</i> for the explicit simulation results). The upstream Gibbs measure <math>\mathcal{G}_\theta^-</math> with <math>D_- = 1</math> displays a wide parameter regime in <math>(\theta^-, E_0)</math> with near-Gaussian statistics. In the left panel of Figure 1 (a), the inflow skewness <math>\kappa_3^-</math> varies only slightly with changing values of <math>E_0</math> and <math>\theta^-</math>. The incoming flow PDF then can be determined by picking the proper parameter value <math>\theta^-</math> in the near Gaussian regime with small skewness. In contrast, the downstream Gibbs measure <math>\mathcal{G}_\theta^+</math> with <math>D_+ = 0.24</math> shown in the right panel of Figure 1 (a) generates much larger skewness <math>\kappa_3^+</math> as the absolute value of <math>\theta^+</math> and the total energy level <math>E_0</math> increases. The solid lines in Figure 1 (c) offer a further confirmation of the transition from near-Gaussian statistics with small <math>\kappa_3^-</math> to a strongly skewed distribution <math>\kappa_3^+</math> after the depth change.</p> <p>In the next step, the value of the downstream <math>\theta^+</math> is determined based on the matching condition [4]. The expectation <math>\langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^-}</math> about the incoming flow Gibbs measure can be calculated according to the predetermined parameter values of <math>\theta^-</math> as well as <math>E_0</math> from the previous step. For the direct numerical experiments shown later in Figure 2, we pick proper choices of test parameter values as <math>L_0 = 6, E_0 = 100</math> and <math>\theta^- = -0.1, -0.3, -0.5</math>. More test cases with different system energy <math>E_0</math> can be found in <i>SI Appendix, B.2</i> where similar transition from near Gaussian symmetric PDFs to skewed PDFs in the flow state <math>u_\Lambda^\pm</math> can always be observed.</p> <p><b>Direct numerical model simulations.</b> Besides the prediction of equilibrium statistical measures from the equilibrium statistical approach, another way to predict the downstream model statistics is through running the dynamical model [1] directly. The TKdV equation is found to be ergodic with proper mixing property as measured by the decay of autocorrelations as long as the system starts from a negative inverse temperature state as described before. For direct numerical simulations of the TKdV equations, a proper symplectic integrator is required to guarantee the Hamiltonian and energy are conserved in time. It is crucial to use the symplectic scheme to guarantee the exact conservation of the energy and Hamiltonian since they are playing the central role in generating the invariant measure and the statistical matching. The symplectic schemes used here for the time integration of the equation is the 4th-order midpoint method (27). Details about the mixing properties from different initial states and the numerical algorithm are described in <i>SI Appendix, C</i>.</p> <p><b>5. Predicting extreme anomalous behavior after the ADC by statistical matching</b></p> <p>With the inflow statistics well described and the numerical scheme set up, we are able to predict the downstream anomalous statistics starting from the near-Gaussian incoming flow going through the abrupt depth change from <math>D_- = 1</math> to <math>D_+ = 0.24</math>. First, we consider the statistical prediction in the downstream equilibrium measure directly from the matching</p>
--	--



**Fig. 1.** First row: skewness from the Gibbs measures in incoming and outgoing flow states with different values of total energy  $E_0$  and inverse temperature  $\theta$  (notice the different scales in the incoming and outgoing flows); Second row: outgoing flow parameter  $\theta^+$  as a function of the incoming flow  $\theta^-$  computed from the statistical matching condition with three energy level  $E_0$ ; Last row: skewness in the outgoing flow with the matched value of  $\theta^+$  as a function of the inflow parameter  $\theta^-$  (the theoretical predictions using [5] are compared).

condition. The downstream parameter value  $\theta^+$  is determined by solving the nonlinear equation [4] as a function of  $\theta^+$ ,  $F(\theta^+) = \langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^+}(\theta^+) - \langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^-} = 0$ . In the numerical approach, we adopt a modified secant method avoiding the stiffness in the parameter regime (see the *SI Appendix, B.2* for the algorithm). The fitted solution is plotted in Figure 1 (b) as a function of the proposed inflow  $\theta^-$ . A nonlinear  $\theta^- - \theta^+$  relation is discovered from the matching condition. The downstream inverse temperature  $\theta^+$  will finally saturate at some level. The corresponding downstream skewness of the wave displacement  $u_\Lambda$  predicted from the statistical matching of Gibbs measures is plotted in Figure 1 (c). In general, a large positive skewness for outgoing flow  $\kappa_3^+$  is predicted from the theory, while the incoming flow skewness  $\kappa_3^-$  is kept in a small value in a wide range of  $\theta^-$ . Note that with  $\theta^- \sim 0$  (that is, using the microcanonical ensemble only with energy conservation), the outflow statistics are also near Gaussian with weak skewness. The skewness in the outflow statistics grows as the inflow parameter value  $\theta^-$  increases in amplitude.

For a second approach, we can use direct numerical simulations starting from the initial state sampled from the incoming flow Gibbs measure  $\mathcal{G}_\theta^-$  and check the transient changes in the model statistics. Figure 2 illustrates the change of statistics as the flow goes through the abrupt depth change. The first

row plots the changes in the skewness and kurtosis for the state variable  $u_\Lambda$  after the depth change at  $t = 0$ . The PDFs in the incoming and outgoing flow states are compared with three different initial inverse temperatures  $\theta^-$ . After the depth changes to  $D_0 = 0.24$  abruptly at  $t = 0$ , both the skewness and kurtosis jump to a much larger value in a short time, implying the rapid transition to a highly skewed non-Gaussian statistical regime after the depth change. Further from Figure 2, different initial skewness (but all relatively small) is set due to the various values of  $\theta^-$ . With small  $\theta^- = -0.1$ , the change in the skewness is not very obvious (see the second row of Figure 2 for the incoming and outgoing PDFs of  $u_\Lambda$ ). In comparison, if the incoming flow starts from the initial parameter  $\theta^- = -0.3$  and  $\theta^- = -0.5$ , much larger increase in the skewness is induced from the abrupt depth change. Furthermore, in the detailed plots in the third row of Figure 2 for the downstream PDFs under logarithmic scale, fat tails towards the positive direction can be observed, which represent the extreme events in the downstream flow (see also Figure 3 for the time-series of  $u_\Lambda$ ).

As a result, the downstream statistics in final equilibrium predicted from the direct numerical simulations here agree with the equilibrium statistical mechanics prediction illustrated in Figure 1. The prediction from these two different approaches confirm each other.

## 6. Analytic formula for the upstream skewness after the ADC

A statistical link between the upstream and downstream energy spectra can be found for an analytical prediction of the skewness in the flow state  $u$  after the ADC. The skewness of the state variable  $u_j$  at one spatial grid point is defined as the ratio between the third and second moments

$$\kappa_3 = \langle u_j^3 \rangle_\mu / \langle u_j^2 \rangle_\mu^{3/2}.$$

Now we introduce mild assumptions on the distribution functions:

- The upstream equilibrium measure  $\mu_-$  has a relatively small skewness so that

$$\langle H_3 \rangle_{\mu_-} = \frac{1}{6} \int_{-\pi}^{\pi} \langle u^3 \rangle_{\mu_-} dx \equiv \epsilon;$$

- The downstream equilibrium measure  $\mu_+$  is homogeneous at each physical grid point, so that the second and third moments are invariant at each grid point

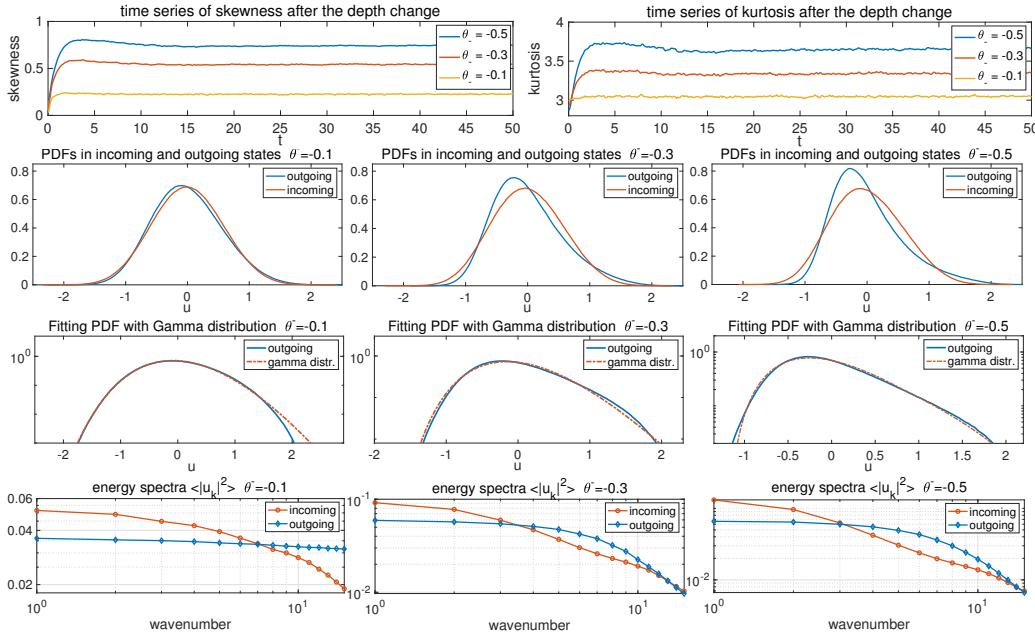
$$\langle u_j^2 \rangle_{\mu_+} = \sigma^2 = \pi^{-1}, \quad \langle u_j^3 \rangle_{\mu_+} = \sigma^3 \kappa_3 = \pi^{-3/2} \kappa_3^+.$$

Then the skewness of the downstream state variable  $u_\Lambda^+$  after the ADC is given by the difference between the inflow and outflow wave slope energy of  $u_x$

$$\kappa_3^+ = \frac{3}{2} \pi^{1/2} L_0^{-3/2} E_0^{-1/2} D_+^2 \int_{-\pi}^{\pi} [\langle u_x^2 \rangle_{\mu_+} - \langle u_x^2 \rangle_{\mu_-}] dx + 3\pi^{1/2} \epsilon.$$

The detailed derivation is shown in *SI Appendix, B.2*. In particular, the downstream skewness with near-Gaussian inflow statistics  $\epsilon \ll 1$  is positive if and only if the difference of the

497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558



**Fig. 2.** Changes in the statistics of the flow state going through the abrupt depth change. The initial ensemble is set with the incoming flow Gibbs measure with different inverse temperature  $\theta^-$ . First row: time evolution of the skewness and kurtosis. The abrupt depth change is taking place at  $t = 0$ ; Second row: inflow and outflow PDFs of  $u_A$ ; Third row: the downstream PDFs fitted with Gamma distributions with consistent variance and skewness (in log coordinate in  $y$ ); Last row: energy spectra in the incoming and outgoing flows.

incoming and outgoing wave slope energy is positive. This means that there is more small scale wave slope energy in the outgoing state. As an evidence, in the last row of Figure 2 in all the weak and strong skewness cases, the outflow energy spectrum always has a slower decay rate than the inflow energy spectrum which possesses stronger energy in larger scales and weaker energy in the smaller scales.

In Figure 1 (c), we compare the accuracy of the theoretical estimation [5] with numerical tests. In the regime with small incoming inverse temperature  $\theta^-$ , the theoretical formula offers a quite accurate approximation of the third-order skewness using only information from the second-order moments of the wave-slope spectrum.

## 7. Key features from experiments captured by the statistical dynamical model

In this final section, we emphasize the crucial features generated by the statistical dynamical model [1] by making comparison with the experimental observations in (20). As from the scale analysis displayed in Section 2, the theory is set in the same parameter regime as the experimental setup.

- The transition from near-Gaussian to skewed non-Gaussian distribution as well as the jump in both skewness and kurtosis observed in the experiment observations (Fig. 1 of (20)) can be characterized by the statistical model simulation results (see the first and second row of Figure 2). Notice that the difference in the decay of third and fourth moments in the far end of the downstream regime from the experimental data is due to the dissipation effect in the flow from the wave absorbers that is not modeled in the statistical model here. The model simulation time-series plotted in Figure 3 can be compared with the observed time sequences from experiments (Fig. 1 of

(20)). The downstream simulation generates waves with strong and frequent intermittency towards the positive displacement, while the upstream waves show symmetric displacements in two directions with at most small peaks in slow time. Even in the time-series at a single location  $x = 0$ , the long-time variation displays similar structures.

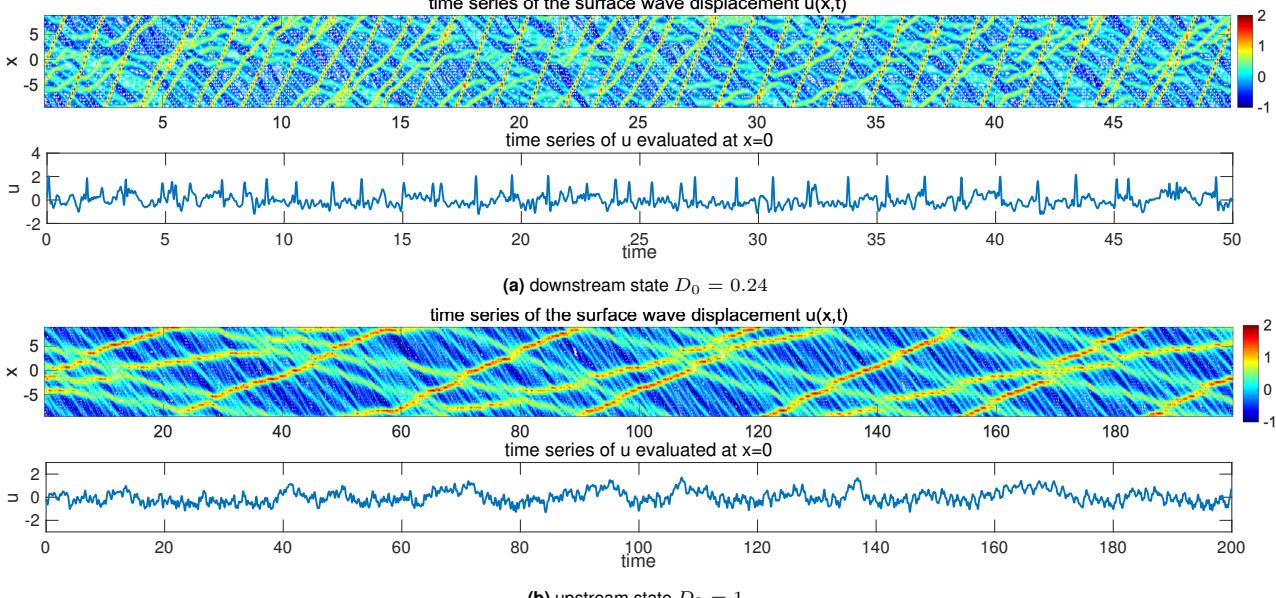
- The downstream PDFs in experimental data are estimated with a Gamma distribution in Fig. 2 of (20). Here in the same way, we can fit the highly skewed outgoing flow PDFs from the numerical results with the Gamma distribution

$$\rho(u; k, \alpha) = \frac{e^{-k} \alpha^{-1}}{\Gamma(k)} (k + \alpha^{-1} u)^{k-1} e^{-\alpha^{-1} u}.$$

The parameters  $(k, \alpha)$  in the Gamma distribution are fitted according to the measured statistics in skewness and variance, that is,  $\sigma^2 = k\alpha^2$ ,  $\kappa_3 = 2/\sqrt{k}$ . And the excess kurtosis of the Gamma distribution can be recovered as  $\kappa_4 = 6/k$ . As shown in the third row of Figure 2, excellent agreement in the PDFs with the Gamma distributions is reached in consistency with the experimental data observations. The accuracy with this approximation increases as the initial inverse temperature  $\theta^-$  increases in value to generate more skewed distribution functions.

- Experimental measurements of the power spectra (Fig. 4 of (20)) reveal the downstream measurements to contain more energy at small scales, i.e. a relatively slower decay rate of the spectrum. This result is also observed in the direct numerical simulations here (detailed results shown in *SI Appendix, C.2*), as the outgoing state contains more energetic high frequencies.

559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620



**Fig. 3.** Realization of the downstream and upstream flow solutions  $u_A^{\pm}$ . Note the larger vertical scale in the downstream time-series plot.

## 8. Concluding discussion

We have developed a statistical dynamical model to explain and predict extreme events and anomalous features of shallow water waves crossing an abrupt depth change. The theory is based on the dynamical modeling strategy consisting of the TKdV equation matched at the abrupt depth change with conservation of energy and Hamiltonian. Predictions can be made of the extreme events and anomalous features by matching incoming and outgoing statistical Gibbs measures before and after the abrupt depth transition. The statistical matching of the known nearly Gaussian incoming Gibbs state completely

determines the predicted anomalous outgoing Gibbs state, which can be calculated by a simple sampling algorithm, verified by direct numerical simulations, and successfully captures key features of the experiment. An analytic formula for the anomalous outgoing skewness is also derived. The strategy here should be useful for predicting extreme statistical events in other dispersive media in different settings.

**ACKNOWLEDGMENTS.** This research of A. J. M. is partially supported by the Office of Naval Research through MURI N00014-16-1-2161. D. Q. is supported as a postdoctoral fellow on the grant. M.N.J.M would like to acknowledge support from Simons grant 524259.

1. Majda AJ, Branicki M (2012) Lessons in uncertainty quantification for turbulent dynamical systems. *Discrete & Continuous Dynamical Systems-A* 32(9):3133–3221.
2. Mohamad MA, Sapsis TP (2018) A sequential sampling strategy for extreme event statistics in nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* 115(44):11138–11143.
3. Qi D, Majda AJ (2016) Predicting fat-tailed intermittent probability distributions in passive scalar turbulence with imperfect models through empirical information theory. *Communications in Mathematical Sciences* 14(6):1687–1722.
4. Majda AJ, Tong XT (2015) Intermittency in turbulent diffusion models with a mean gradient. *Nonlinearity* 28(11):4171–4208.
5. Majda AJ, Chen N (2018) Model error, information barriers, state estimation and prediction in complex multiscale systems. *Entropy* 20(9):644.
6. Majda AJ, Tong XT (2018) Simple nonlinear models with rigorous extreme events and heavy tails. *arXiv preprint arXiv:1805.05615*.
7. Thual S, Majda AJ, Chen N, Stechmann SN (2016) Simple stochastic model for el nino with westerly wind bursts. *Proceedings of the National Academy of Sciences* 113(37):10245–10250.
8. Chen N, Majda AJ (2018) Efficient statistically accurate algorithms for the Fokker-Planck equation in large dimensions. *Journal of Computational Physics* 354:242–268.
9. Chen N, Majda AJ (2017) Beating the curse of dimension with accurate statistics for the Fokker-Planck equation in complex turbulent systems. *Proceedings of the National Academy of Sciences* 114(49):12864–12869.
10. Chen N, Majda AJ (2018) Conditional Gaussian systems for multiscale nonlinear stochastic systems: Prediction, state estimation and uncertainty quantification. *Entropy* 20(7):509.
11. Qi D, Majda AJ (2018) Predicting extreme events for passive scalar turbulence in two-layer baroclinic flows through reduced-order stochastic models. *Communications in Mathematical Sciences* 16(1):17–51.
12. Adcock TA, Taylor PH (2014) The physics of anomalous ('rogue') ocean waves. *Reports on Progress in Physics* 77(10):105901.
13. Cousins W, Sapsis TP (2015) Unsteady evolution of localized unidirectional deep-water wave groups. *Physical Review E* 91(6):063204.
14. Farazmand M, Sapsis TP (2017) Reduced-order prediction of rogue waves in two-dimensional deep-water waves. *Journal of Computational Physics* 340:418–434.
15. Onorato M, Osborne AR, Serio M, Bertone S (2001) Freak waves in random oceanic sea states. *Physical Review Letters* 86(25):5831–5834.
16. Dematteis G, Grafke T, Vanden-Eijnden E (2018) Rogue waves and large deviations in deep sea. *Proceedings of the National Academy of Sciences* 115(5):855–860.
17. Sergeeva A, Pelinovsky E, Talipova T (2011) Nonlinear random wave field in shallow water: variable Korteweg-de Vries framework. *Natural Hazards and Earth System Sciences* 11(2):323–330.
18. Trulsen K, Zeng H, Gramstad O (2012) Laboratory evidence of freak waves provoked by non-uniform bathymetry. *Physics of Fluids* 24(9):097101.
19. Viotti C, Dias F (2014) Extreme waves induced by strong depth transitions: Fully nonlinear results. *Physics of Fluids* 26(5):051705.
20. Bolles CT, Speer K, Moore MNJ (2018) Anomalous wave statistics induced by abrupt depth change. *arXiv preprint arXiv:1808.07958*.
21. Solli D, Ropers C, Koonath P, Jalali B (2007) Optical rogue waves. *Nature* 450(7172):1054.
22. Höhmann R, Kuhl U, Stöckmann HJ, Kaplan L, Heller E (2010) Freak waves in the linear regime: A microwave study. *Physical review letters* 104(9):093901.
23. Johnson RS (1997) *A modern introduction to the mathematical theory of water waves*. (Cambridge university press) Vol. 19.
24. Abramov RV, Kováčik G, Majda AJ (2003) Hamiltonian structure and statistically relevant conserved quantities for the truncated Burgers-Hopf equation. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences* 56(1):1–46.
25. Majda A, Wang X (2006) *Nonlinear dynamics and statistical theories for basic geophysical flows*. (Cambridge University Press).
26. Bajars J, Frank J, Leimkuhler B (2013) Weakly coupled heat bath models for Gibbs-like invariant states in nonlinear wave equations. *Nonlinearity* 26(7):1945–1973.
27. McLachlan R (1993) Symplectic integration of hamiltonian wave equations. *Numerische Mathematik* 66(1):465–492.