

# 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

This manuscript was compiled on November 25, 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 controlled laboratory experiments in weakly turbulent shallow water with abrupt depth change (ADC) exhibit a remarkable transition from nearly Gaussian statistics in incoming wave trains before the ADC to outgoing waves trains after the ADC with extreme anomalous statistics with large positive skewness of the surface height. 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 which are statistically matched at the ADC. 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.**

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 dimension 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 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 which are statistically matched at the ADC. 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.

If you are interested, we could cite papers on ‘Optical rogue waves’ and/or rogue waves in microwave systems here.

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

A series of experiments are carried out (20) studying the anomalous statistical behaviors in surface water waves going through an abrupt depth transition. The unidirectional waves propagate along a water tank over a step in the bottom topography, and the surface displacements of the wave levels are

### 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 measured at several upstream and downstream locations. The  
126 wave field is excited by a paddle wheel forcing with angle  
127

$$128 \quad \theta(t) = \theta_0 + \Delta\theta \sum_{n=1}^N a_n \cos(\omega_n t + \delta_n), \quad E \sim (\Delta\theta)^2 \sum_n a_n^2 \omega_n^2.$$

130 The total energy  $E$  in the system is determined by the angle  
131 amplitude  $\Delta\theta$ . Several observations are found in the anomalous  
132 wave statistics:

- 134 • Distinct statistics are found between the incoming and  
135 outgoing wave disturbances: the incoming waves display  
136 near-Gaussian statistics, while the outgoing waves show  
137 skewness toward the positive displacement.
- 138 • The non-Gaussian statistics is related with the total en-  
139 ergy contained in the system: larger driving amplitude  
140  $\Delta\theta$  will generate stronger skewness in the PDFs.
- 141 • The waves also show different characteristic peak wave-  
142 lengths in incoming and outgoing flows.

## 145 **2. Surface wave turbulence modeled by truncated KdV 146 equation with depth dependence**

147 The surface wave turbulence is modeled by a one-dimensional  
148 deterministic dynamical model. The Korteweg-de Vries (KdV)  
149 equation (21) is a leading-order approximation of the surface  
150 waves that are determined by the balance of nonlinear and  
151 dispersive effects in an appropriate far-field limit. Here, the  
152 KdV equation is truncated in the first  $\Lambda$  modes (with  $J = 2\Lambda + 1$  grid points) to generate weakly turbulent dynamics.  
153 Therefore, the surface disturbance is modeled by the state  
154 variable  $u_\Lambda^\pm(x, t)$  with superscript ‘–’ for the incoming waves  
155 and ‘+’ for the outgoing waves. The Galerkin truncated  
156 variable  $u_\Lambda = \sum_{1 \leq |k| \leq \Lambda} \hat{u}_k(t) e^{ikx}$  is normalized with zero  
157 mean  $\hat{u}_0 = 0$  and unit energy  $2\pi \sum_{k=1}^\Lambda |\hat{u}_k|^2 = 1$ , which are  
158 conserved quantities, and  $u_\Lambda \equiv \mathcal{P}_\Lambda u$  denotes the subspace  
159 projection. The motion is governed by the truncated KdV  
160 equation with depth change  $D_\pm$  about the state variable  $u_\Lambda^\pm$   
161

$$162 \quad \frac{\partial u_\Lambda^\pm}{\partial t} + \frac{D_\pm^{-3/2}}{2} E_0^{1/2} L_0^{-3/2} \frac{\partial}{\partial x} \mathcal{P}_\Lambda (u_\Lambda^\pm)^2 + D_\pm^{1/2} L_0^{-3} \frac{\partial^3 u_\Lambda^\pm}{\partial x^3} = 0, \quad [1]$$

166 on the normalized periodic domain  $x \in [-\pi, \pi]$  with the  
167 conserved Hamiltonian decomposed into the difference of two  
168 components containing the cubic and quadratic terms  
169

$$170 \quad \mathcal{H}_\Lambda^\pm = D_\pm^{-3/2} E_0^{1/2} L_0^{-3/2} H_3(u_\Lambda^\pm) - D_\pm^{1/2} L_0^{-3} H_2(u_\Lambda^\pm),$$

$$171 \quad H_3(u) = \frac{1}{6} \int_{-\pi}^{\pi} u^3 dx, \quad H_2(u) = \frac{1}{2} \int_{-\pi}^{\pi} \left( \frac{\partial u}{\partial x} \right)^2 dx.$$

174 The model [1] is non-dimensionalized in the periodic domain.  
175 The depth is assumed to be unit  $D_- = 1$  before the abrupt  
176 depth change and becomes  $D_+ < 1$  for the flows after the  
177 change. We introduce the model parameters  $(E_0, L_0, \Lambda)$  based  
178 on the following model assumptions:

- 179 • The wavenumber truncation  $\Lambda$  is fixed in a moderate  
180 value for generating weakly turbulent dynamics;
- 181 • The state variable  $u_\Lambda^\pm$  is normalized with zero mean and  
182 unit energy,  $\mathcal{M}(u_\Lambda) = 0, \mathcal{E}(u_\Lambda) = 1$ , conserved during  
183 the evolution, while  $E_0$  characterizes the total energy  
184 injected in the system based on the paddle amplitude  
185  $(\Delta\theta)^2$ ;

- 187 • The length scale of the system is defined by  $L_0$ . The  
188 value is chosen so that the resolved scale  $\Delta x = 2\pi L_0/J$   
189 is comparable with the the characteristic wave length  $\lambda_c$   
190 found from the experiments.

191 The intuition for the distinct model dynamics comes from the  
192 balance between the cubic and quadratic terms in the Hamil-  
193 tonian  $\mathcal{H}_\Lambda^\pm$ . After the depth change,  $D_+ < 1$ , more weight is  
194 added in the cubic term,  $H_3$ , for stronger nonlinearity and  
195 weaker dispersion for the third-order derivative term reflected  
196 by the smaller coefficient for  $H_2$  in the Hamiltonian. Since  
197  $\frac{\partial u}{\partial x}$  is the slope of the wave height,  $H_2(u)$  measures the wave  
198 slope energy.

199 A *deterministic matching condition* is given for the surface  
200 displacement  $u_\Lambda^\pm$  agreeing at the locations before and after the  
201 abrupt depth change  $T_{ADC}$

$$203 \quad u_\Lambda^-(x, t)|_{t=T_{ADC}-} = u_\Lambda^+(x, t)|_{t=T_{ADC}+},$$

205 assuming the abrupt depth change is met at  $t = T_{ADC}$ . Equation  
206 [1] is not designed to capture the short scale changes in  
207 rapid time. On the other hand, we are interested in the model  
208 statistical transition before and after the depth change, so it  
209 is reasonable to observe the suitable slow-time performance in  
210 the large scale structures.

### 212 **Interpreting experimental parameters in the dynamical model.**

213 The model parameters  $(E_0, L_0, \Lambda)$  in [1] can be directly linked  
214 with the basic scales from the physical problem. The important  
215 characterizing parameters measured from the experiments  
216 include:  $\epsilon = \frac{a}{H_0}$  the wave amplitude  $a$  to water depth  $H_0$   
217 ratio;  $\delta = \frac{H_0}{\lambda_c}$  the water depth to wavelength scale  $\lambda_c$  ratio;  
218 and  $D_0 = \frac{d}{H_0}$  the normalized wave depth ratio with incoming  
219 flow depth  $d = H_0$  to the outgoing flow depth  $d < H_0$ . The  
220 interpretations and reference values of these model parameters  
221 are based on the experimental setup (20). By comparing the  
222 characteristic physical scales, the normalized TKdV equation  
223 [1] can be linked directly with the measured non-dimensional  
224 quantities by

$$226 \quad L_0 = 6^{\frac{1}{3}} \left( M \epsilon^{\frac{1}{2}} \delta^{-1} \right), \quad E_0 = \frac{27}{2} \gamma^{-2} \left( M \epsilon^{\frac{1}{2}} \delta^{-1} \right), \quad [2]$$

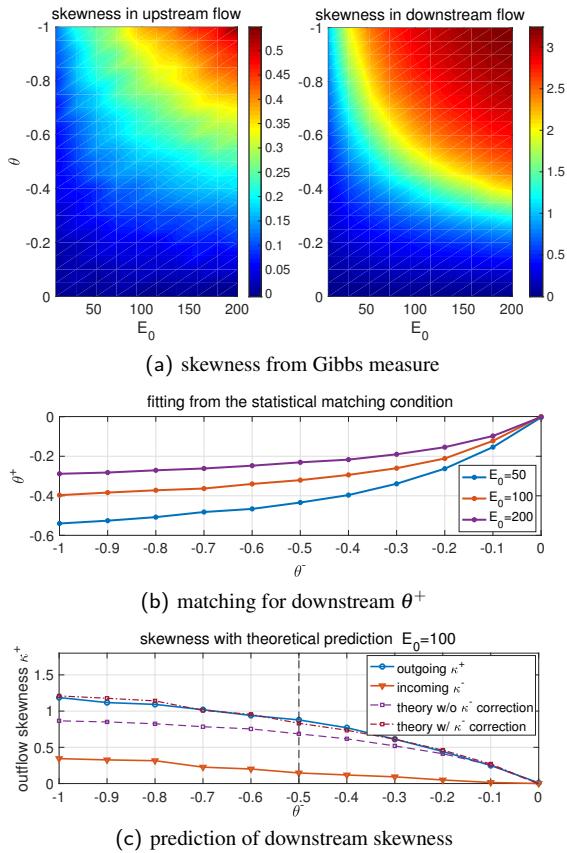
227 where  $M$  defines the computational domain size  $L_d = M \lambda_c$  as  
228  $M$ -multiple of the characteristic wavelength  $\lambda_c$ , and  $\gamma = \frac{U}{a}$   
229 represents the factor to normalize the total energy in the state  
230 variable  $u_\Lambda$  to one.

231 Consider the spatial discretization  $J = 2\Lambda + 1$  so that the  
232 smallest resolved scale is comparable with the characteristic  
233 wavelength

$$234 \quad \Delta x = \frac{2\pi M \lambda_c}{J} \lesssim \lambda_c \Rightarrow M = \frac{J}{2\pi} \sim 5, \quad J = 32.$$

235 Therefore in the practical numerical simulations, we pick  
236  $M = 5$  and  $\gamma$  varies in the range  $[0.5, 1]$ . Using the reference  
237 experimental measurements (20),  $\epsilon \in [0.0024, 0.024]$ ,  $\delta \sim 0.22$ ,  
238 and  $D_0$  changes from 1 to 0.24 before and after the depth  
239 change. The reference values for the model scales can be esti-  
240 mated in the range  $L_0 \in [2, 6]$  and  $E_0 \in [50, 200]$ . These are  
241 the values we will test in the direct numerical simulations. See  
242 details about the derivation from scale analysis in *SI Appendix*,  
243 *A*.

<p>249 <b>3. Equilibrium statistical mechanism for generating</b> 250 <b>the stationary invariant measure</b></p> <p>252 Since the TKdV equation satisfies the Liouville property, the 253 equilibrium invariant measure can be described by an equi- 254 librium statistical formulism (22–24) using a Gibbs measure 255 with the conserved energy <math>\mathcal{E}_\Lambda</math> and Hamiltonian <math>\mathcal{H}_\Lambda</math>. The equi- 256 librium invariant measure is dictated by the conservation laws 257 in the TKdV equation. In the case with fixed total energy <math>E_0</math>, 258 this is the <i>mixed Gibbs measure</i> in the truncated model with 259 microcanonical energy and canonical Hamiltonian ensembles (22)</p> <p>261 <math display="block">\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), \quad [3]</math></p> <p>262 with <math>\theta</math> representing the “inverse temperature”. The distinct 263 statistics in the upstream and downstream waves can be con- 264 trolled by the parameter value of the inverse temperature. 265 The negative temperature regime, <math>\theta^\pm &lt; 0</math>, is the appropriate 266 regime to predict the experiments as shown below. In the 267 incoming flow field, the inverse temperature <math>\theta^-</math> is chosen so 268 that <math>\mathcal{G}_\theta^-</math> has Gaussian statistics. Using the above invariant 269 measures [3], the expectation of any functional <math>F(u)</math> can be 270 computed based on the Gibbs measure</p> <p>272</p> <p>273 <math display="block">\langle F \rangle_{\mathcal{G}_\theta} \equiv \int F(u) \mathcal{G}_\theta(u) du.</math></p> <p>274</p> <p>275 The value of <math>\theta</math> in the invariant measure is specified from <math>\langle H_\Lambda \rangle_{\mathcal{G}_\theta}</math> (22, 24). The invariant measure also predicts an equilibrium 276 energy spectrum without running the TKdV equation directly. 277 On the other hand, the time autocorrelation and transient 278 statistics about the state variable <math>u_\Lambda</math> cannot be recovered from 279 the statistical theory.</p> <p>282 <b>Statistical matching condition in invariant measures before</b> 283 <b>and after the abrupt depth change.</b> The Gibbs measures <math>\mathcal{G}_\theta^\pm</math> 284 are defined based on the different inverse temperatures <math>\theta^\pm</math> on 285 the two sides of the solutions</p> <p>287 <math display="block">\mu_t^-(u_\Lambda^-; D_-), \quad u_\Lambda _{t=T_{ADC}-} = u_0, \quad t &lt; T_{ADC};</math></p> <p>288 <math display="block">\mu_t^+(u_\Lambda^+; D_+), \quad u_\Lambda _{t=T_{ADC}+} = u_0, \quad t &gt; T_{ADC},</math></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 also be matched at the depth change location 293 <math>T_{ADC}</math>, so that</p> <p>295 <math display="block">\mu_\infty^-(u_\Lambda) = \mu_{t=T_{ADC}}^-(u_\Lambda) = \mu_{t=T_{ADC}}^+(u_\Lambda).</math></p> <p>296</p> <p>297 In matching the flow statistics before and after the abrupt 298 depth change, first we use the conservation of the determinis- 299 tic Hamiltonian <math>H_\Lambda^+</math> after the depth change. Then assuming 300 ergodicity (22, 23), the statistical expectation for the Hamil- 301 tonian <math>\langle H_\Lambda^+ \rangle</math> is conserved in time after the depth change at 302 <math>t = T_{ADC}</math> and should stay in the same value as the system ap- 303 proaches equilibrium as <math>t \rightarrow \infty</math>. The final statistical matching 304 condition to get the outgoing flow statistics with parameter 305 <math>\theta^+</math> can be found by</p> <p>306</p> <p>307 <math display="block">\langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^+} = \langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^-}, \quad [4]</math></p> <p>308</p> <p>309 with the outgoing flow Hamiltonian <math>H_\Lambda^+</math> and the Gibbs mea- 310 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>The incoming flow is always characterized by a near-Gaussian distribution in the wave 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 (24) 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 within small amplitudes among 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> with highly skewed PDFs as the absolute value of <math>\theta^+</math> and the total energy level <math>E_0</math> increase in amplitude. The solid lines in Figure 1 (c) offer a further confirmation of the transition from near-Gaussian statistics with tiny <math>\kappa_3^-</math> to 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 (25). 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 mechanism 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

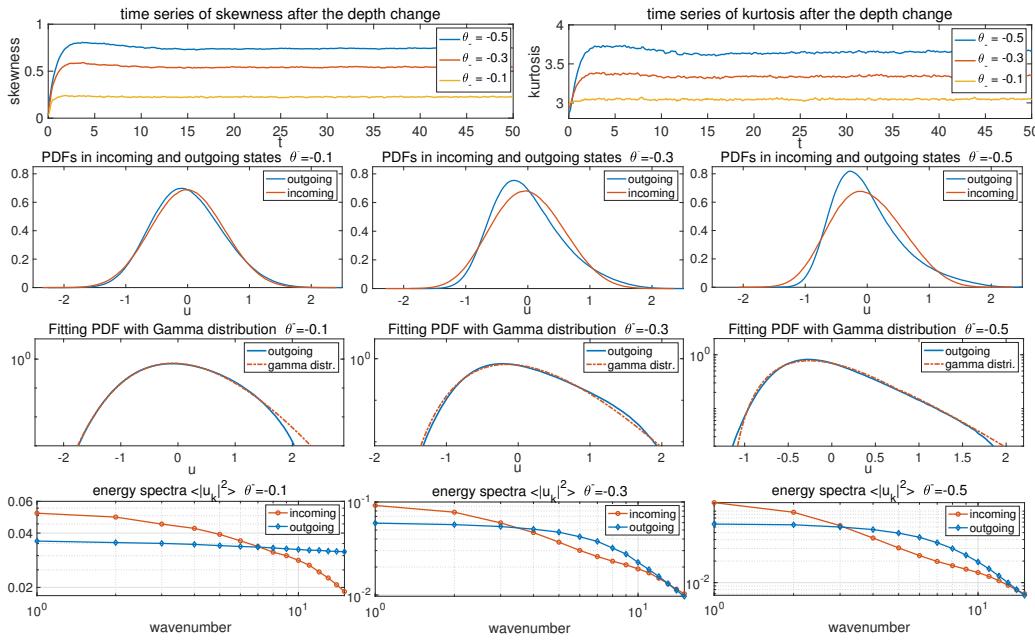
$$\langle u_j^2 \rangle_{\mu_+} = \sigma^2 = \pi^{-1}, \quad \langle u_j^3 \rangle_{\mu_+} = \sigma^3 \kappa_3 = \pi^{-3/2} \kappa_3^+.$$

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  $u_x$

$$\kappa_3^+ = \frac{3}{2} \pi^{1/2} L_0^{-3/2} E_0^{-1/2} D_+^2 \int_{-\pi}^{\pi} [\langle u_x^2 \rangle_{\mu_+} - \langle u_x^2 \rangle_{\mu_-}] dx + 3\pi^{1/2} \epsilon. \quad [5]$$

The detailed derivation is shown in *SI Appendix, B.2*. In particular, the downstream skewness with near-Gaussian inflow

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  $u_\Lambda$ ; 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.

statistics  $\epsilon \ll 1$  is positive if and only if the difference of the 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

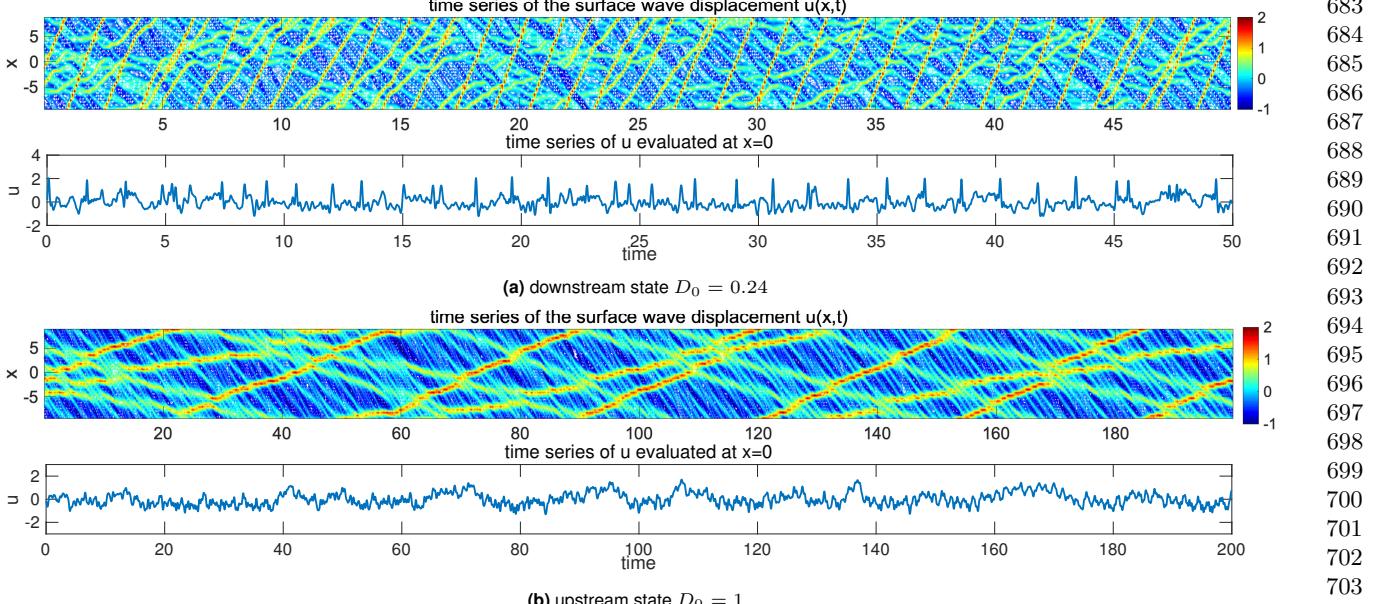
time-series plotted in Figure 3 can be compared with the observed time sequences from experiments (Fig. 1 of (20)). The downstream simulation generates waves with strong and frequent intermittency towards the positive displacement, while the upstream waves show symmetric displacements in two directions 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.

- The experiments also have the up and down stream power spectra in time (Fig. 4 of (20)), which shows more energy at small time scales, i.e., a relatively slower decay rate in the downstream compared with the upstream case. This is also observed in the direct numerical simulations here (detailed results shown in *SI Appendix, C.2*). The



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

downstream state contains more energetic high frequencies. The peak frequency illustrates the occurrence of the transporting waves along the water tank.

I do not understand the last sentence above. Why does the peak frequency illustrate the occurrence of transporting water waves?

## 8. Concluding discussion

We have developed a statistical dynamical model to explain and predict extreme events and anomalous features in shallow water waves with 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 second 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. Johnson RS (1997) *A modern introduction to the mathematical theory of water waves*. (Cambridge university press) Vol. 19.
22. Abramov RV, Kovačić 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.
23. Majda A, Wang X (2006) *Nonlinear dynamics and statistical theories for basic geophysical flows*. (Cambridge University Press).
24. 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.
25. McLachlan R (1993) Symplectic integration of hamiltonian wave equations. *Numerische Mathematik* 66(1):465–492.