

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

Andrew J. Majda^{a,1}, M. N. J. Moore^b, and Di Qi^{a,1}

^aDepartment of Mathematics and Center for Atmosphere and Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, NY 10012;

^bDepartment of Mathematics and Geophysical Fluid Dynamics Institute, Florida State University, Tallahassee, FL

This manuscript was compiled on November 27, 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. 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

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. 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.

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

- 125 specified as 187
 126 $\theta(t) = \theta_0 + \Delta\theta \sum_{n=1}^N a_n \cos(\omega_n t + \delta_n)$, $E \sim (\Delta\theta)^2 \sum_n a_n^2 \omega_n^2$. 188
 127 where the weights a_n are Gaussian in spectral space with 189
 128 peak frequency ω_p and the phases δ_n are uniformly distributed 190
 129 random variables. The peak frequency gives rise to a characteristic wavelength λ_c which can be estimated from the 191 dispersion relation. The energy E injected into the system 192 is determined by the angle amplitude $\Delta\theta$, which is the main 193 control parameter varied in (20). Optical measurements of the 194 free surface reveal a number of surprising statistical features: 195
- Distinct statistics are found between the incoming and outgoing wave disturbances: incoming waves display near-Gaussian statistics, while outgoing waves skew strongly towards positive displacement.
 - The degree of non-Gaussianity present in the outgoing waves depends on the injected energy E : larger energies generate stronger skewness in the surface displacement PDFs and more extreme events.
 - Compared to the incoming wave train, the power spectrum of the outgoing wave field decays more slowly, which indicates that the anomalous behavior is associated with an elevated level of high frequencies.
- 152 **2. Surface wave turbulence modeled by truncated KdV 200**
 153 **equation with depth dependence** 201
- 154 The surface wave turbulence is modeled by a one-dimensional 202 deterministic dynamical model. 203
- 155 This first sentence seems awkward to me. 204
- 156 The Korteweg-de Vries (KdV) equation (23) is a leading-order 205 approximation of the surface waves that are determined 206 by the balance of nonlinear and dispersive effects in an appropriate 207 far-field limit. Here, the KdV equation is truncated 208 in the first Λ modes (with $J = 2\Lambda + 1$ grid points) to generate 209 weakly turbulent dynamics. The surface displacement 210 is described by the state variable $u_\Lambda^\pm(x, t)$ with superscript 211 ‘–’ for the incoming waves and ‘+’ for the outgoing waves. 212 The Galerkin truncated variable $u_\Lambda = \sum_{1 \leq |k| \leq \Lambda} \hat{u}_k(t) e^{ikx}$ 213 is normalized with zero mean $\hat{u}_0 = 0$ and unit energy 214 $2\pi \sum_{k=1}^\Lambda |\hat{u}_k|^2 = 1$, which are conserved quantities, and 215 $u_\Lambda \equiv \mathcal{P}_\Lambda u$ denotes the subspace projection. The evolution of 216 u_Λ^\pm is governed by the truncated KdV equation with depth 217 change D_\pm
- 157
$$\frac{\partial u_\Lambda^\pm}{\partial t} + \frac{D_\pm^{-3/2}}{2} E_0^{1/2} L_0^{-3/2} \frac{\partial}{\partial x} \mathcal{P}_\Lambda (u_\Lambda^\pm)^2 + D_\pm^{1/2} L_0^{-3} \frac{\partial^3 u_\Lambda^\pm}{\partial x^3} = 0, \quad [1]$$
 218
- 158 on the normalized periodic domain $x \in [-\pi, \pi]$. The conserved 219 Hamiltonian can be decomposed into a cubic and a quadratic 220 term
- 159
$$\mathcal{H}_\Lambda^\pm = D_\pm^{-3/2} E_0^{1/2} L_0^{-3/2} H_3(u_\Lambda^\pm) - D_\pm^{1/2} L_0^{-3} H_2(u_\Lambda^\pm),$$
 221
- 160
$$H_3(u) = \frac{1}{6} \int_{-\pi}^{\pi} u^3 dx, \quad H_2(u) = \frac{1}{2} \int_{-\pi}^{\pi} \left(\frac{\partial u}{\partial x} \right)^2 dx.$$
 222
- 161 Equation [1] is non-dimensionalized in the periodic domain. 223 The depth is assumed to be unit $D_- = 1$ before the ADC and 224 becomes $D_+ < 1$ after the ADC. We introduce parameters 225 (E_0, L_0, Λ) based on the following assumptions:
- 162 • The wavenumber truncation Λ is fixed in a moderate value for generating weakly turbulent dynamics; 188
- 163 • The state variable u_Λ^\pm is normalized with zero mean and unit energy, $\mathcal{M}(u_\Lambda) = 0, \mathcal{E}(u_\Lambda) = 1$, which are conserved during evolution. Meanwhile, E_0 characterizes the total energy injected into the system based on the paddle amplitude $(\Delta\theta)^2$; 194
- 164 • The length scale of the system is denoted L_0 . The value is chosen so that the resolved scale $\Delta x = 2\pi L_0/J$ is comparable to the the characteristic wave length λ_c in the experiments. 199
- 165 Some intuition for how Eq. (1) produces different dynamics on either side of the ADC can be gained by considering the balance of the cubic and quadratic terms in the Hamiltonian \mathcal{H}_Λ^\pm . The depth change, $D_+ < 1$, increases the weight of H_3 and decreases that of H_2 , thus promoting the effects of nonlinearity and reducing dispersion. Since $\frac{\partial u}{\partial x}$ is the slope of the wave height, $H_2(u)$ measures the wave slope energy. 200
- 166 A *deterministic matching condition* is applied to the surface displacement u_Λ^\pm to link the incoming and outgoing wave trains. 201 Assuming the abrupt depth change is met at $t = T_{\text{ADC}}$, the 202 matching condition is given by 203
- 167
$$u_\Lambda^-(x, t)|_{t=T_{\text{ADC}}-} = u_\Lambda^+(x, t)|_{t=T_{\text{ADC}}+}, \quad [2]$$
 204
- 168 Equation [1] is not designed to capture the short scale changes 205 in rapid time. Rather, since we are interested in modeling 206 statistics before and after the depth change, we will examine 207 the long-time dynamics of large-scale structures. 208
- 169 I modified the last sentence above. Feel free to revert to 209 your original sentence if you do not like mine. (the change 210 looks good) 211
- 170 **Interpreting experimental parameters in the dynamical model.** 212 The model parameters (E_0, L_0, Λ) in [1] can be directly linked 213 to the basic scales from the physical problem. The important 214 characterizing parameters measured from the experiments 215 include: $\epsilon = \frac{a}{H_0}$ the wave amplitude a to water depth H_0 216 ratio; $\delta = \frac{\Lambda}{\lambda_c}$ the water depth to wavelength scale λ_c ratio; 217 and $D_0 = \frac{d}{H_0}$ the normalized wave depth ratio with incoming 218 flow depth $d = H_0$ to the outgoing flow depth $d < H_0$. The 219 interpretations and reference values of these model parameters 220 are based on the experimental setup (20). By comparing the 221 characteristic physical scales, the normalized TKdV equation 222 [1] can be linked directly with the measured non-dimensional 223 quantities by 224
- 171
$$L_0 = 6^{\frac{1}{3}} \left(M \epsilon^{\frac{1}{2}} \delta^{-1} \right), \quad E_0 = \frac{27}{2} \gamma^{-2} \left(M \epsilon^{\frac{1}{2}} \delta^{-1} \right), \quad [2]$$
 225
- 172 where M defines the computational domain size $M \lambda_c$ as M -multiple 226 of the characteristic wavelength λ_c , and $\gamma = \frac{U}{a}$ represents the factor to normalize the total energy in the state 227 variable u_Λ to one. 228
- 173 Consider the spatial discretization $J = 2\Lambda + 1$ so that the 229 smallest resolved scale is comparable with the characteristic 230 wavelength 231
- 174
$$\Delta x = \frac{2\pi M \lambda_c}{J} \lesssim \lambda_c \Rightarrow M = \frac{J}{2\pi} \sim 5, \quad J = 32. \quad [2]$$
 232

249 Therefore in the practical numerical simulations, we pick
 250 $M = 5$ and γ varies in the range $[0.5, 1]$. Using the reference
 251 experimental measurements (20), $\epsilon \in [0.0024, 0.024]$, $\delta \sim 0.22$,
 252 and D_0 changes from 1 to 0.24 before and after the depth
 253 change. The reference values for the model scales can be esti-
 254 mated in the range $L_0 \in [2, 6]$ and $E_0 \in [50, 200]$. These are
 255 the values we will test in the direct numerical simulations. See
 256 details about the derivation from scale analysis in *SI Appendix,*
 257 *A*.

258 3. Equilibrium statistical mechanics for generating the 259 stationary invariant measure

260 Since the TKdV equation satisfies the Liouville property, the
 261 equilibrium invariant measure can be described by an equi-
 262 librium statistical formulism (24–26) using a Gibbs measure
 263 with the conserved energy \mathcal{E}_Λ and Hamiltonian \mathcal{H}_Λ . The equi-
 264 librium invariant measure is dictated by the conservation laws
 265 in the TKdV equation. In the case with fixed total energy E_0 ,
 266 this is the *mixed Gibbs measure* in the truncated model with
 267 microcanonical energy and canonical Hamiltonian ensembles
 268 (24)

$$269 \quad \mathcal{G}_\theta^\pm(u_\Lambda^\pm; E_0) = C_\theta^\pm \exp(-\theta^\pm \mathcal{H}(u_\Lambda^\pm)) \delta(\mathcal{E}(u_\Lambda^\pm) - E_0), \quad [3]$$

270 with θ representing the “inverse temperature”. The distinct
 271 statistics in the upstream and downstream waves can be con-
 272 trolled by the value of θ . Negative temperature, $\theta^\pm < 0$, is
 273 the appropriate regime to predict the experiments as shown
 274 below. In the incoming flow field, the inverse temperature θ^-
 275 is chosen so that \mathcal{G}_θ^- has Gaussian statistics. Using the above
 276 invariant measures [3], the expectation of any functional $F(u)$
 277 can be computed based on the Gibbs measure

$$278 \quad \langle F \rangle_{\mathcal{G}_\theta} \equiv \int F(u) \mathcal{G}_\theta(u) du.$$

279 The value of θ in the invariant measure is specified from $\langle H_\Lambda \rangle_{\mathcal{G}_\theta}$
 280 (24, 26). The invariant measure also predicts an equilibrium
 281 energy spectrum without running the TKdV equation directly.
 282 On the other hand, the time autocorrelation and transient
 283 statistics about the state variable u_Λ cannot be recovered from
 284 the statistical theory.

285 **Statistical matching condition in invariant measures before**
 286 **and after the abrupt depth change.** The Gibbs measures \mathcal{G}_θ^\pm
 287 are defined based on the different inverse temperatures θ^\pm on
 288 the two sides of the solutions

$$289 \quad \begin{aligned} \mu_t^-(u_\Lambda^-; D_-), \quad u_\Lambda |_{t=T_{\text{ADC}}^-} &= u_0, \quad t < T_{\text{ADC}}; \\ 290 \quad \mu_t^+(u_\Lambda^+; D_+), \quad u_\Lambda |_{t=T_{\text{ADC}}^+} &= u_0, \quad t > T_{\text{ADC}}, \end{aligned}$$

291 where u_0 represents the deterministic matching condition be-
 292 tween the incoming and outgoing waves. The two distributions,
 293 μ_t^-, μ_t^+ should also be matched at the depth change location
 294 T_{ADC} , so that,

$$295 \quad \mu_{t=T_{\text{ADC}}}^-(u_\Lambda) = \mu_{t=T_{\text{ADC}}}^+(u_\Lambda).$$

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 H_Λ^+ after the depth change. Then assuming
 299 ergodicity (24, 25), the statistical expectation for the Hamil-
 300 tonian $\langle H_\Lambda^+ \rangle$ is conserved in time after the depth change at

311 $t = T_{\text{ADC}}$ and should stay in the same value as the system ap-
 312 proaches equilibrium as $t \rightarrow \infty$. The final statistical matching
 313 condition to get the outgoing flow statistics with parameter
 314 θ^+ can be found by

$$315 \quad \langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^+} = \langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^-}, \quad [4]$$

316 with the outgoing flow Hamiltonian H_Λ^+ and the Gibbs mea-
 317 sures \mathcal{G}_θ^\pm before and after the abrupt depth change.

318 4. The nearly Gaussian incoming statistical state

319 The incoming flow is always characterized by a near-Gaussian
 320 distribution in the wave displacement. It is found that a
 321 physically consistent Gibbs measure should take negative val-
 322 ues in the inverse temperature parameter $\theta < 0$, where a
 323 proper distribution function and a decaying energy spectrum
 324 are generated (see (26) and *SI Appendix, B.1* for the explicit
 325 simulation results). The upstream Gibbs measure \mathcal{G}_θ^- with
 326 $D_- = 1$ displays a wide parameter regime in (θ^-, E_0) with
 327 near-Gaussian statistics. In the left panel of Figure 1 (a), the
 328 inflow skewness κ_3^- varies only slightly with changing values of
 329 E_0 and θ^- . The incoming flow PDF then can be determined
 330 by picking the proper parameter value θ^- in the near Gaus-
 331 sian regime with small skewness. In contrast, the downstream
 332 Gibbs measure \mathcal{G}_θ^+ with $D_+ = 0.24$ shown in the right panel
 333 of Figure 1 (a) generates much larger skewness κ_3^+ as the
 334 absolute value of θ^+ and the total energy level E_0 increases.
 335 The solid lines in Figure 1 (c) offer a further confirmation of
 336 the transition from near-Gaussian statistics with small κ_3^- to
 337 a strongly skewed distribution κ_3^+ after the depth change.

338 In the next step, the value of the downstream θ^+ is deter-
 339 mined based on the matching condition [4]. The expectation
 340 $\langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^-}$ about the incoming flow Gibbs measure can be cal-
 341 culated according to the predetermined parameter values of
 342 θ^- as well as E_0 from the previous step. For the direct nu-
 343 matical experiments shown later in Figure 2, we pick proper
 344 choices of test parameter values as $L_0 = 6, E_0 = 100$ and
 345 $\theta^- = -0.1, -0.3, -0.5$. More test cases with different system
 346 energy E_0 can be found in *SI Appendix, B.2* where similar
 347 transition from near Gaussian symmetric PDFs to skewed
 348 PDFs in the flow state u_Λ^\pm can always be observed.

349 **Direct numerical model simulations.** Besides the prediction of
 350 equilibrium statistical measures from the equilibrium statisti-
 351 cal approach, another way to predict the downstream model
 352 statistics is through running the dynamical model [1] directly.
 353 The TKdV equation is found to be ergodic with proper mixing
 354 property as measured by the decay of autocorrelations as long
 355 as the system starts from a negative inverse temperature state
 356 as described before. For direct numerical simulations of the
 357 TKdV equations, a proper symplectic integrator is required to
 358 guarantee the Hamiltonian and energy are conserved in time.
 359 It is crucial to use the symplectic scheme to guarantee the
 360 exact conservation of the energy and Hamiltonian since they
 361 are playing the central role in generating the invariant measure
 362 and the statistical matching. The symplectic schemes used
 363 here for the time integration of the equation is the 4th-order
 364 midpoint method (27). Details about the mixing properties
 365 from different initial states and the numerical algorithm are
 366 described in *SI Appendix, C*.

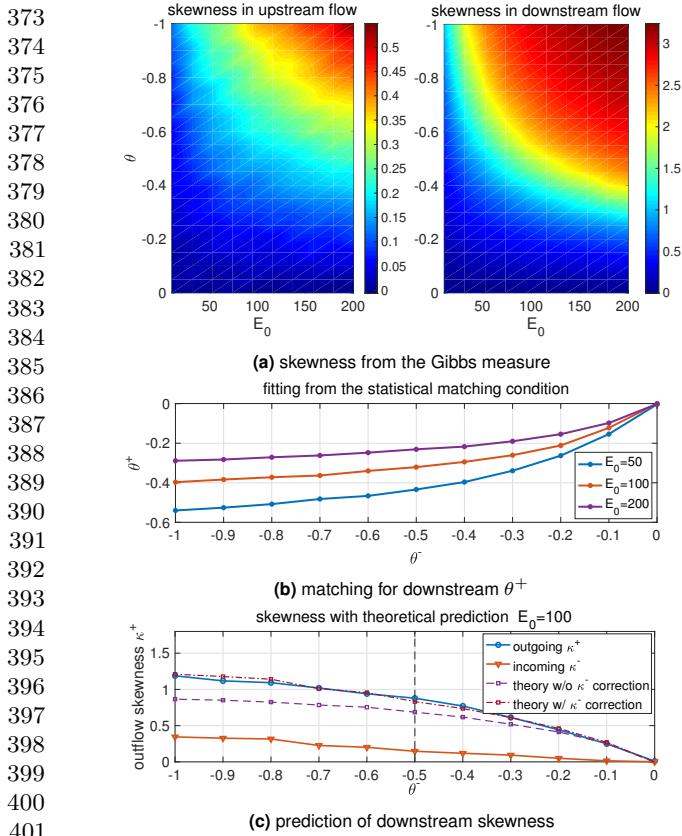


Fig. 1. First row: skewness from the Gibbs measures in incoming and outgoing flow states with different values of total energy E_0 and inverse temperature θ (notice the different scales in the incoming and outgoing flows); Second row: outgoing flow parameter θ^+ as a function of the incoming flow θ^-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

5. Predicting extreme anomalous behavior after the ADC by statistical matching

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 $D_- = 1$ to $D_+ = 0.24$. First, we consider the statistical prediction in the downstream equilibrium measure directly from the matching condition. The downstream parameter value θ^+ is determined by solving the nonlinear equation [4] as a function of θ^+ , $F(\theta^+) = \langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^+}(\theta^+) - \langle H_\Lambda^+ \rangle_{\mathcal{G}_\theta^-} = 0$. 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 θ^- . A nonlinear $\theta^- - \theta^+$ relation is discovered from the matching condition. The downstream inverse temperature θ^+ will finally saturate at some level. The corresponding downstream skewness of the wave displacement u_Λ predicted from the statistical matching of Gibbs measures is plotted in Figure 1 (c). In general, a large positive skewness for outgoing flow κ_3^+ is predicted from the theory, while the incoming flow skewness κ_3^- is kept in a small value in a wide range of θ^- . 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 θ^- 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}_θ^- 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_Λ after the depth change at $t = 0$. The PDFs in the incoming and outgoing flow states are compared with three different initial inverse temperatures θ^- . 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 θ^- . 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_Λ). 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_Λ).

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

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

Now we introduce mild assumptions on the distribution functions:

- The upstream equilibrium measure μ_- 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 μ_+ is homogeneous at each physical grid point, so that the second and third moments are invariant at each grid point

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

Then the skewness of the downstream state variable u_Λ^+ after the ADC is given by the difference between the inflow and

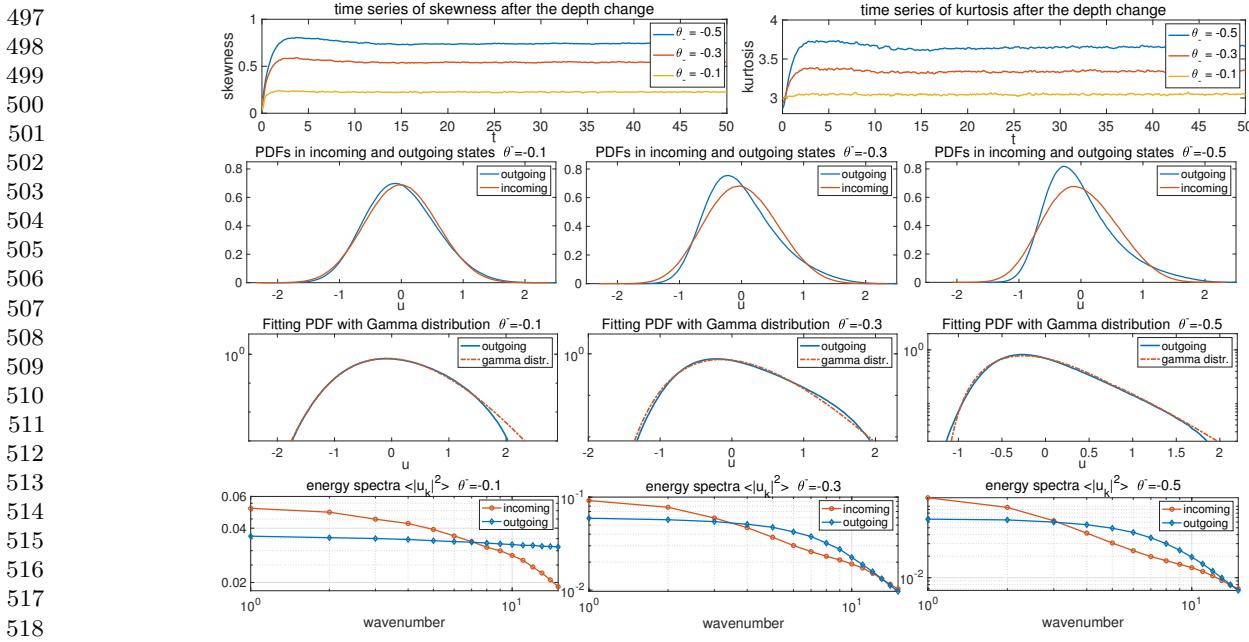


Fig. 2. Changes in the statistics of the flow state going through the abrupt depth change. The initial ensemble is set with the incoming flow Gibbs measure with different inverse temperature θ^- . First row: time evolution of the skewness and kurtosis. The abrupt depth change is taking place at $t = 0$; Second row: inflow and outflow PDFs of u_Λ ; 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.

outflow wave slope energy of u_x

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

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 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 θ^- , 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, α) 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 θ^- increases in value to generate more skewed distribution functions.

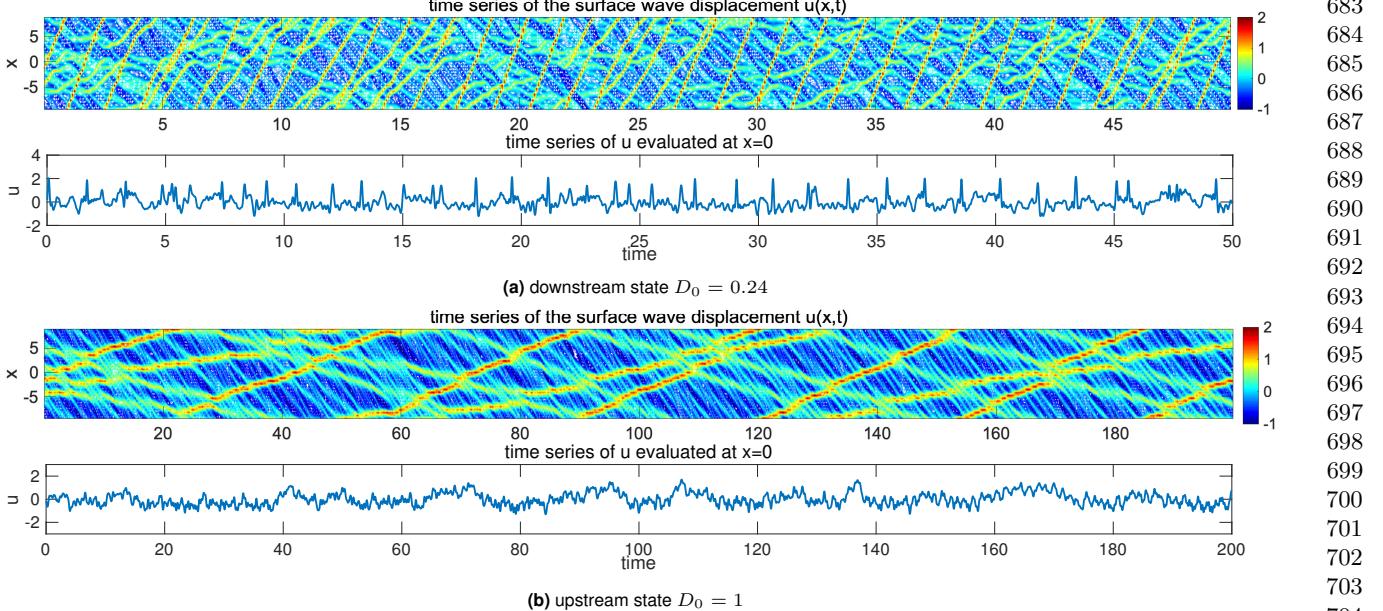


Fig. 3. Realization of the downstream and upstream flow solutions u_A^\pm . Note the larger vertical scale in the downstream time-series plot.

- 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 downstream state contains more energetic high frequencies. **The peak frequency illustrates the characteristic time scale of the transporting wave trains 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 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 second grant. M.N.J.M would like to acknowledge support from Simons grant 524259.

- Majda AJ, Branicki M (2012) Lessons in uncertainty quantification for turbulent dynamical systems. *Discrete & Continuous Dynamical Systems-A* 32(9):3133–3221.
- 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.
- 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.
- Majda AJ, Tong XT (2015) Intermittency in turbulent diffusion models with a mean gradient. *Nonlinearity* 28(11):4171–4208.
- Majda AJ, Chen N (2018) Model error, information barriers, state estimation and prediction in complex multiscale systems. *Entropy* 20(9):644.
- Majda AJ, Tong XT (2018) Simple nonlinear models with rigorous extreme events and heavy tails. *arXiv preprint arXiv:1805.05615*.
- 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.
- Chen N, Majda AJ (2018) Efficient statistically accurate algorithms for the Fokker–Planck equation in large dimensions. *Journal of Computational Physics* 354:242–268.
- 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.
- Chen N, Majda AJ (2018) Conditional Gaussian systems for multiscale nonlinear stochastic systems: Prediction, state estimation and uncertainty quantification. *Entropy* 20(7):509.
- Qin, Majda AJ (2019) Reducing extreme events for passive scalar turbulence in two-layer baroclinic flows through reduced-order stochastic models. *Communications in Mathematical Sciences* 16(1):17–51.
- Adcock TA, Taylor PH (2014) The physics of anomalous ('rogue') ocean waves. *Reports on Progress in Physics* 77(10):105901.
- Farazmand M, Sapsis TP (2017) Reduced-order prediction of rogue waves in two-dimensional deep-water waves. *Journal of Computational Physics* 340:418–434.
- Onorato M, Osborne AR, Serio M, Bertone S (2001) Freak waves in random oceanic sea states. *Physical Review Letters* 86(25):5831–5834.
- 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.
- 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.
- Trulsen K, Zeng H, Gramstad O (2012) Laboratory evidence of freak waves provoked by non-uniform bathymetry. *Physics of Fluids* 24(9):097101.
- Viotti C, Dias F (2014) Extreme waves induced by strong depth transitions: Fully nonlinear results. *Physics of Fluids* 26(5):051705.
- Bolles CT, Speer K, Moore MNJ (2018) Anomalous wave statistics induced by abrupt depth change. *arXiv preprint arXiv:1808.07958*.
- Solli D, Ropers C, Koonath P, Jalali B (2007) Optical rogue waves. *Nature* 450(7172):1054.
- 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.
- Johnson RS (1997) *A modern introduction to the mathematical theory of water waves*. (Cambridge university press) Vol. 19.
- Abramov RV, Kovačič 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.
- Majda A, Wang X (2006) *Nonlinear dynamics and statistical theories for basic geophysical flows*. (Cambridge University Press).
- 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.
- McLachlan R (1993) Symplectic integration of hamiltonian wave equations. *Numerische*