Supporting Information

Dematteis et al. 10.1073/pnas.1710670115

1. Construction of the Initial Condition and Dynamical Consistency Check

Our procedure requires to specify the statistics of the (complex) envelope at initial time, $u_0(x)$, whereas the experimental spectrum is for the surface elevation $\eta(x)$ which is related to $u_0(x)$ as

$$\eta(x) = \Re\big(u_0(x)e^{ik_0x}\big), \qquad \text{[S1]}$$

To construct the initial $u_0(x)$, we introduce the auxiliary variable $\zeta(x)$, 5

$$\zeta(x) = \Im\big(u_0(x)e^{ik_0x}\big), \qquad \text{[S2]}$$

which we treat as a field independent of $\eta(x)$, with the same statistics. It is easy to see from Eqs. S1 and S2 that the envelope $u_0(x)$ can then be expressed as

$$u_0(x) = \big(\eta(x) + i\zeta(x)\big)e^{-ik_0x}. \qquad \text{[S3]}$$

Assuming that both $\eta(x)$ and $\zeta(x)$ are independent Gaussian fields with covariance $\mathbb{E}(\eta(x)\eta(x')) = \mathbb{E}(\zeta(x)\zeta(x')) = C_\eta(x - x')$, the envelope $u_0(x)$ is also Gaussian, with covariance $C_u(x - x') = \mathbb{E}(u(x)\bar{u}(x'))$ given by

$$C_u(x - x') = 2C_\eta(x - x')e^{-ik_0(x-x')}. \qquad \text{[S4]}$$

This relation implies that

$$\hat{C}_u(k) = 2\hat{C}_\eta(k + k_0). \qquad \text{[S5]}$$

where we defined

$$\hat{C}_u(k) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-ikx}C_u(x)\,dx, \qquad \hat{C}_\eta(k) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-ikx}C_\eta(x)\,dx. \qquad \text{[S6]}$$

Recalling that k_0 is defined as the wave vector at which the spectrum of $\eta(x)$ should be centered, if we take a Gaussian ansatz for this spectrum, we should pick

$$\hat{C}_\eta(k) = \hat{C}_\eta(0)e^{-|k-k_0|^2/(2\Delta^2)}. \qquad \text{[S7]}$$

As a result,

$$\hat{C}_u(k) = 2\hat{C}_\eta(0)e^{-k^2/(2\Delta^2)}. \qquad \text{[S8]}$$

The spectrum for $u_0(x)$ used in the work is a discretized version of the one above, with $A = (2\pi/L)^2\hat{C}_\eta(0)$.

The results reported in the main text require us to evolve the field $u(t,x)$ from its initial condition $u_0(x)$. As explained in the main text, through this evolution, the probabilities $P_t(z) = \mathbb{P}(\max_x|u(t,x)|)$ change with time t until they converge to some limit value. It is interesting to ask how much this evolution changes the prior information we used to construct the initial $u_0(x)$; that is, it is interesting to look at the spectrum of $u(t,x)$ and see how much it differs from that of $u_0(x)$. The results of this calculation are shown in Fig. S1, and they indicate that the spectrum stays essentially constant in time over 100 min. This justifies our choice of prior: indeed, from the viewpoint of this prior, the time evolution of $u(t,x)$ leads to no significant changes. Of course, some features of $u(t,x)$ change, as apparent from the evolution of other observables such as $P_t(z) = \mathbb{P}(\max_x|u(t,x)|)$. Detecting the trace of these changes in the spectrum requires one to look at much finer energy scales: This can be seen in Fig. S1, *Right*, where we plot the energy contained in modes above $k > 0$ for increasing values of k.

2. Influence of the Size of the Domain and of the Observation Window

In this section, we investigate the influence of the size of the domain and/or that of the observation window on our results. To this end, we conduct experiments in domains of size $L = L_0 = 40\pi$ (the domain size used in the main text, which is $L_0 \approx 4.53 \times 10^3$ m in dimensional units), and compare with $L = 2L_0$, $L = 4L_0$, and $L = 8L_0$. The base domain size L_0 was chosen to be as small as possible for computational efficiency, but still large enough that the influence of the periodic boundary conditions be negligible (as checked below). Consequently, the results below can be interpreted either by thinking of $L \geq L_0$ as the actual domain size, or as the size of the observation window in an even larger domain (including one that could be infinite). We also stress that our results are numerically converged and consistent in terms of numerical resolution, in the sense that we doubled both the number of grid points in the domain and the number of modes in the initial data each time we doubled the domain size. In particular, we used 2^{12} grid points and $M = 23$ initial modes ($-11 \leq n \leq 11$) in the domain of size L, 2^{13} grid points and $M = 47$ initial modes ($-23 \leq n \leq 23$) in the domain of size $2L$, etc.

We begin by checking that the domain of size $L_0 = 40\pi$ is already large enough to render negligible the effect of the boundary conditions. To this end, let us consider a different observable than the one in the main text, namely, the probability that $|u(t,x)|$ be above a certain threshold at a given location x_0 in the domain,

$$P_0^L(t,z) = \mathbb{P}(|u(t,x_0)| > z), \qquad x_0 \in [0,L]. \qquad \text{[S9]}$$

By translational invariance, $P_0^L(t,z)$ is independent of x_0. As $L \to \infty$ this probability converges to a limiting value, $P_0^L(t,z) \to P_0(t,z)$, which makes it useful to consider here. As can be seen from Fig. S2, convergence is already achieved for $L = L_0$, $P_0^{L_0}(t,z) \approx P_0(t,z)$. The results shown in Fig. S2 are for $t = 15$ min, when the probability has converged to that on the invariant measure already. A similar conclusion can be made at intermediate times: Fig. S3 shows that doubling the domain size makes no significant difference, i.e., $P_0^{2L_0}(t,z) \approx P_0^{L_0}(t,z)$, both in the results from Monte Carlo sampling and in those from our large deviation approach. The same invariance is also observed in the trajectories obtained by optimization in the large deviation approach (Fig. S4). Note that these results are not surprising since L_0 is already much larger than the correlation length of the initial field, $L_0 \simeq 10L_c$—this is in fact why this value of L_0 was chosen to begin with.

Coming back to the quantity investigated in the main text, let us denote

$$P_{\max}^L(t,z) = \mathbb{P}\left(\max_{x\in[0,L]}|u(t,x)| > z\right). \qquad \text{[S10]}$$

Unlike $P_0^L(t,z)$, the probability $P_{\max}^L(t,z)$ does depend on L—the larger L, the higher $P_{\max}^L(t,z)$. We can actually estimate this growth explicitly. To see how, consider a domain of size NL that can be partitioned into $N \geq 1$ subdomains of size L, each large enough to be roughly statistically independent of the others. Then, we have

$$1 - P_{\max}^{NL}(t,z) = \Big(1 - P_{\max}^L(t,z)\Big)^N, \qquad N \geq 1 \qquad \text{[S11]}$$

since in order for the maximum of $|u|$ to be less than z in the larger domain of size NL, it must be less than z in each of

the (roughly independent) subdomains of size L. Eq. **S11** is the fundamental equation used in extreme value statistics. We confirmed its applicability for $L = L_0 = 40\pi$ in our system via direct estimation of $P_{\max}^{NL_0}(t,z)$ for $N = 1, 2, 4, 8$ by Monte Carlo sampling. These results are reported in Fig. S5.

Since $L_0 = 40\pi$ is already large enough for Eq. **S11** to hold, we can rewrite this equation as

$$1 - P_{\max}^{L}(t,z) = \left(1 - P_{\max}^{L_0}(t,z)\right)^{L/L_0}, \qquad L \geq L_0 \quad \textbf{[S12]}$$

Note that this equation implies that, at fixed z, $P_{\max}^{L}(t,z)$ increases with L since $1 - P_{\max}^{L_0}(t,z) < 1$ and therefore $1 - P_{\max}^{L}(t,z) = \left(1 - P_{\max}^{L_0}(t,z)\right)^{L/L_0} \leq 1 - P_{\max}^{L_0}(t,z)$ for $L \geq L_0$. Intuitively, this increase in $P_{\max}^{L}(t,z)$ stems from the fact that multiple large values of $|u|$ are expected to arise simultaneously in different subdomains since they are statistically independent—this is usually referred to as an entropic effect, and it can be seen in the typical realizations from the Monte Carlo sampling shown in Fig. S6 for $L = L_0$ and $L = 8L_0$. Of course, this effect is properly accounted for by Eq. **S11**. Indeed, realizations like those shown in Fig. S6 are those from which the probabilities shown in Fig. S5 were calculated.

It is also important to stress that this entropic effect cannot be accounted for directly by our large deviation approach. The solution obtained by optimization becomes independent of L for L large enough (which is the case already for $L = L_0$). This implies that, without correction, the results of the large deviation approach will deteriorate with increasing L. Eq. **S12** shows that this issue can be easily fixed, however: Indeed, this formula indicates how the large deviation results at $L = L_0$ (i.e., in a domain that is large enough to not be influenced by the boundary condition, but small enough that the entropic effects remain negligible) can be extended to larger L.

3. The Case of the NLS and the Role of the Peregrine soliton

For completeness, we redid all of our calculations in the context of the standard NLS equation instead of the MNLS equation. Using the same nondimensional variables as in MNLS, NLS reads

$$\partial_t u + \frac{i}{8}\partial_{xx} u + \frac{i}{2} u|u|^2 = 0. \quad \textbf{[S13]}$$

Fig. S7 shows the distributions for the spatial maximum of the envelope $|u|$ at different times calculated by both direct Monte Carlo sampling and minimization using our large deviation approach, using the same random initial conditions as in MNLS. As can be seen, here, too, the approach based on LDT does an excellent job at capturing these PDFs.

The advantage of using NLS is that it permits us to assess the relevance of the Peregrine soliton (PS), which is an exact solution of NLS (although not of MNLS) that has been invoked as prototype mechanism for rogue waves creation (1–6)—recent experimental results in the context of water waves (7–9), plasmas (10), and fiber optics (11–13) have lent support to this hypothesis. The PS reads

$$u(t,x) = U_i e^{-it/T_{\rm nl}}\left(\frac{4(1 - 2it/T_{\rm nl})}{1 + 4\left(t/T_{\rm nl}\right)^2 + 4\left(x/L_{\rm nl}\right)^2} - 1\right), \quad \textbf{[S14]}$$
$$T_{\rm nl} = \frac{2}{U_i^2}, \quad L_{\rm nl} = \frac{1}{4}\sqrt{T_{\rm nl}} = \frac{\sqrt{2}}{4U_i},$$

where $U_i > 0$ is a free parameter. It can be checked that this solution reaches its maximal amplitude $|u(0,0)| = 3U_i$ at $(t,x) = (0,0)$ and decays both forward and backward in time to $\lim_{t\to\pm\infty}|u(t,x)| = U_i$.

To compare the PS to our results, we translated t in Eq. **S14** to make the time at which this solution reaches its maximal amplitude coincide with the time at which a prescribed value of the wave elevation is observed in either our minimization procedure or in the Monte Carlo sampling. By adjusting U_i so that the maximal amplitude of the PS also coincides with this prescribed value of the amplitude, we can then verify how well the PS reproduces our instanton as well as the mean and variance of the solutions observed in the Monte Carlo sampling. These results are reported in Fig. S8. As can be seen, the PS captures the shape of the instanton at final time (i.e., when the rogue wave occurs) reasonably well, at least near the location $x = 0$ where the maximum amplitude is observed (focusing region). The PS also does a reasonably good job at tracking the evolution of the solution that led to this extreme event. In particular, the focusing time scale of the optimized solutions (which we interpret to also describe the convergence time of the a priori distribution to the invariant distribution) is in rough agreement with the effective focusing time scale of the PS starting from a pulse of size L_i (11, 13). This time scale is given by $\tau_c = \sqrt{T_{\rm nl}\, T_{\rm lin}}$, where the nonlinear time $T_{\rm nl}$ is defined in Eq. **S14**, and the linear time $T_{\rm lin} = 8L_i^2$ is that associated with group velocity dispersion of the initial pulse—in dimensional units, these are $T_{\rm nl} = \left(\frac{1}{2}\omega_0 k_0^2 U_i^2\right)^{-1}$ and $T_{\rm lin} = 8\omega_0^{-1} k_0^2 L_i^2$.

The relative agreement both in shape and timescale between the optimized solution and the PS suggests that the main physical phenomenon responsible for the focusing in the NLS equation is the gradient catastrophe (14), which fosters a very unique evolution pathway as the point of maximum focusing is approached in space-time. Still, it should be stressed that the discrepancies between the PS and the actual solution we observe become more and more pronounced backward in time. These differences can also be observed in Fig. S9, where we plot the amplitude of u for a more extreme event that is too rare to be observed by Monte Carlo sampling. In this figure, we show the optimized solutions obtained for two different spectral widths Δ, whose shapes are slightly different from one another: Clearly, these differences cannot be captured by the PS since this solution is completely specified by the final amplitude, which is the same for both sets.

For completeness, we also compared the PS with the solutions obtained in the context of MNLS. These results are reported in Fig. S10 and show similar types of agreement, in particular in term of the shape of the rogue wave near its maximum and the time scale of its emergence. Note the discrepancies between the PS and our solutions is even more pronounced in this case, which is to be expected since PS is an exact solution of NLS, but not of MNLS.

To summarize, while the PS can explain some features of the rogue waves, in particular their shape as well as the focusing time scale over which these waves evolve from a large initial pulse, it does not capture the details of the formation of these waves—indeed, there is no reason why it should, since different sets of random initial conditions lead to waves with different shapes (and whose amplitudes have different statistics), and this information is not seen by the PS. In particular, the instanton solution for the initial data chosen here depends on two parameters, the significant wave height H_s and the BFI, while the PS only allows a single parameter U_i. Additionally, and more importantly, the PS does not allow the estimation of the probability of observing rogue waves of given amplitude since this solution per se is devoid of a probabilistic framework.

1. Onorato M, Residori S, Bortolozzo U, Montina A, Arecchi F (2013) Rogue waves and their generating mechanisms in different physical contexts. *Phys Rep* 528:47–89.
2. Peregrine DH (1983) Water waves, nonlinear Schrödinger equations and their solutions. *ANZIAM J* 25:16–43.
3. Akhmediev N, Ankiewicz A, Taki M (2009) Waves that appear from nowhere and disappear without a trace. *Phys Lett A* 373:675–678.
4. Shrira VI, Geogjaev VV (2010) What makes the peregrine soliton so special as a prototype of freak waves?*J Eng Math* 67:11–22.
5. Akhmediev N, Dudley JM, Solli D, Turitsyn S (2013) Recent progress in investigating optical rogue waves. *J Opt* 15:060201.
6. Toenger S, et al. (2015) Emergent rogue wave structures and statistics in spontaneous modulation instability. *Sci Rep* 5:10380.
7. Chabchoub A, Hoffmann N, Akhmediev N (2011) Rogue wave observation in a water wave tank. *Phys Rev Lett* 106:204502.
8. Chabchoub A, Hoffmann N, Onorato M, Akhmediev N (2012) Super rogue waves: Observation of a higher-order breather in water waves. *Phys Rev X* 2:011015.
9. Chabchoub A (2016) Tracking breather dynamics in irregular sea state conditions. *Phys Rev Lett* 117:144103.
10. Bailung H, Sharma S, Nakamura Y (2011) Observation of peregrine solitons in a multicomponent plasma with negative ions. *Phys Rev Lett* 107: 255005.
11. Tikan A, et al. (2017) Universality of the peregrine soliton in the focusing dynamics of the cubic nonlinear Schrödinger equation. *Phys Rev Lett* 119: 033901.
12. Kibler B, et al. (2010) The peregrine soliton in nonlinear fibre optics. *Nat Phys* 6: 790–795.
13. Suret P, et al. (2016) Single-shot observation of optical rogue waves in integrable turbulence using time microscopy. *Nat Commun* 7:13136.
14. Bertola M, Tovbis A (2013) Universality for the focusing nonlinear Schrödinger equation at the gradient catastrophe point: Rational breathers and poles of the Tritronquée solution to Painlevé I. *Commun Pure Appl Math* 66:678–752.

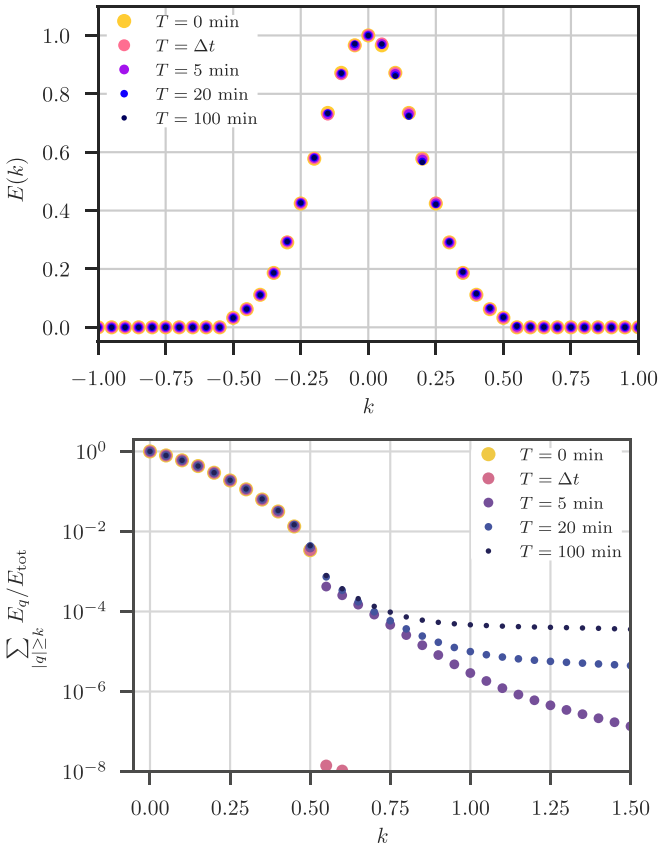


Fig. S1. Evolution of the spectrum of $u(t, x)$. *Upper* shows that this spectrum stays essentially constant in time over 100 min, which justifies our choice of prior: Indeed, from the viewpoint of this prior, the time evolution of $u(t, x)$ leads to no changes. Of course, some features of $u(t, x)$ change, as apparent from the evolution of other observables such as $P_t(z) = \mathbb{P}(\max_x |u(t, x)|)$: These changes can be detected in the spectrum, but they require us to look at much finer energy scales, as shown in *Lower*, where we plot the energy contained in modes above $k > 0$ as k increases.

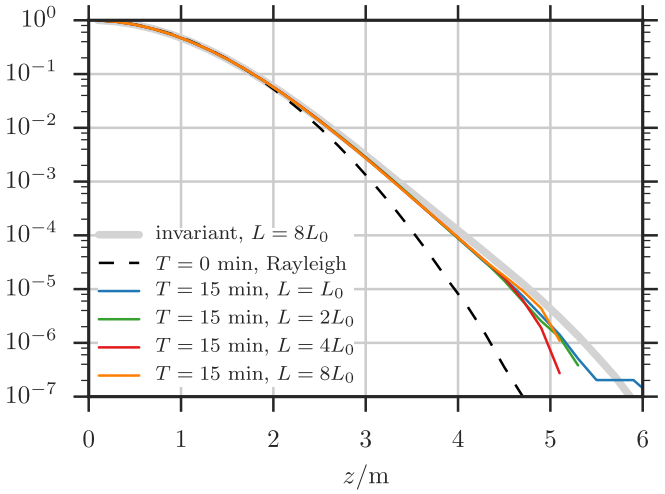


Fig. S2. Numerical verification of the invariance $P_0^L(z) = \lim_{t\to\infty} P_0^L(t, z)$ for $L \geq L_0$. The limiting value $P_0(z)$ (gray curve) was calculated by propagating 1500 samples up to time of 3000 min in the largest domain with $L = 8L_0$). Note that this also shows that $P_0^{L_0}(t, z)$ in the Monte Carlo sampling has essentially converged to the invariant $P_0(z)$ after only 15 min.

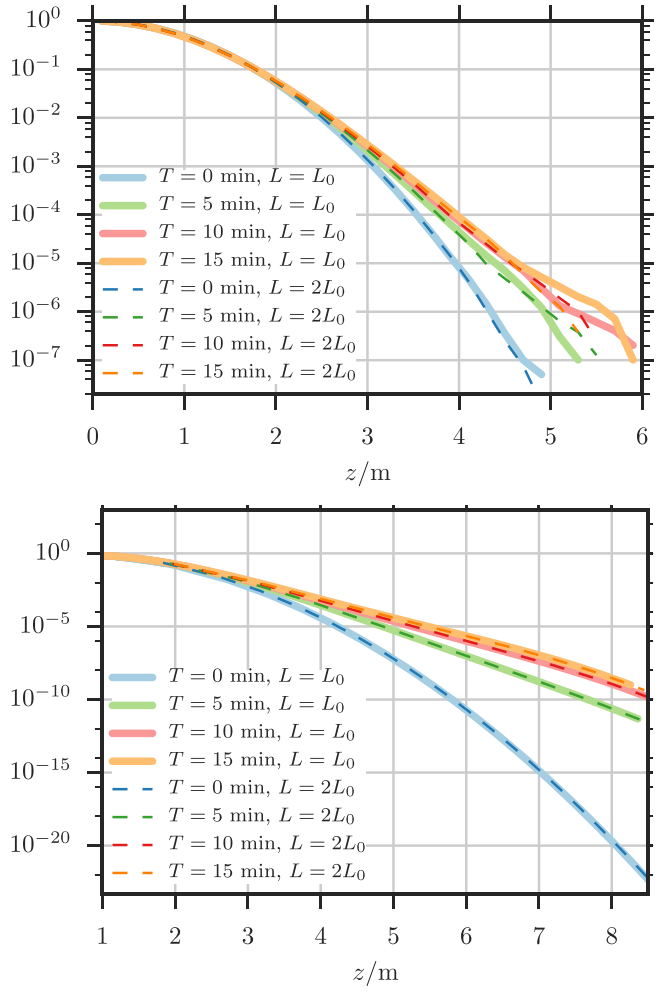


Fig. S3. (*Upper*) $P_0^L(t, z) = \mathbb{P}(|u(t, x_0)| > z)$ at a fixed location x_0 and different times t in domains of size $L = L_0$ and $L = 2L_0$ obtained by Monte Carlo sampling. (*Lower*) Same, obtained by optimization using our large deviation approach and a larger range of values for z (such large values cannot easily be reached by Monte Carlo). As can be seen, the probability distribution functions (PDFs) essentially lay on top of each other for the two different domains, confirming that the domain size L_0 is large enough for the periodic boundary conditions to not affect the results.

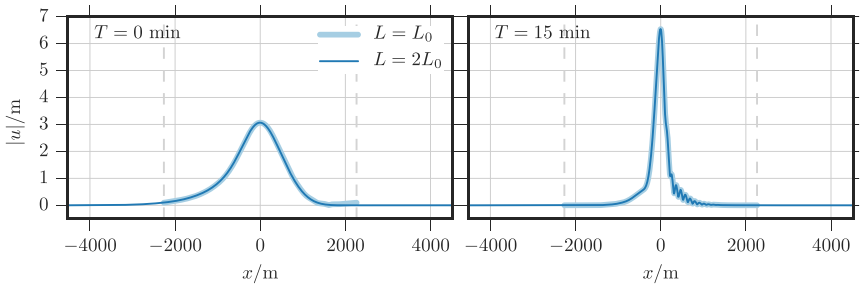


Fig. S4. Optimal trajectories calculated in the domains of size $L = L_0$ (thick line) and $L = 2L_0$ (thin line). As can be seen, the periodicity of the domain does not affect significantly the shape of the instanton inside this domain.

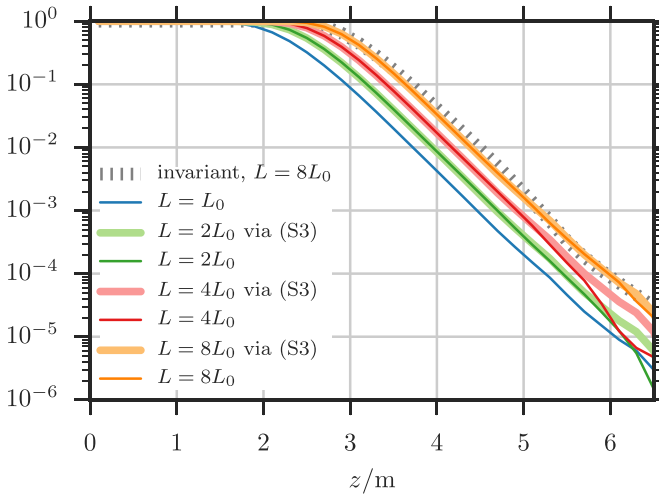


Fig. S5. Numerical verification of Eq. **S11** for $L = L_0 = 40\pi$. These results confirm that adjacent boxes of size L_0 can be considered statistically independent. The probability $P^L_{\max}(z) = \lim_{t\to\infty} P^L_{\max}(t, z)$ for $L = 8L_0$, is also shown, indicating that this quantity can be estimated accurately from $P^{L_0}_{\max}$ at 15 min using Eq. **S11**.

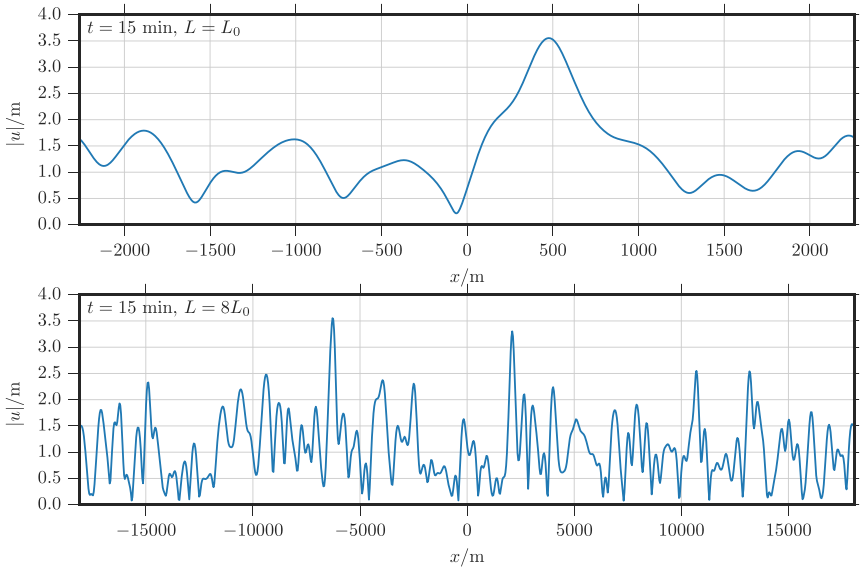


Fig. S6. Typical realizations from the Monte Carlo sampling such that $\max_x |u(t, x)| \geq 3.5$ m at $t = 15$ min in the domains of size $L = L_0$ (*Upper*) and $L = 8L_0$ (*Lower*). As can be seen, as the domain size increases, it becomes increasingly likely to observe more than one large value of $u(t, x)|$ in the domain.

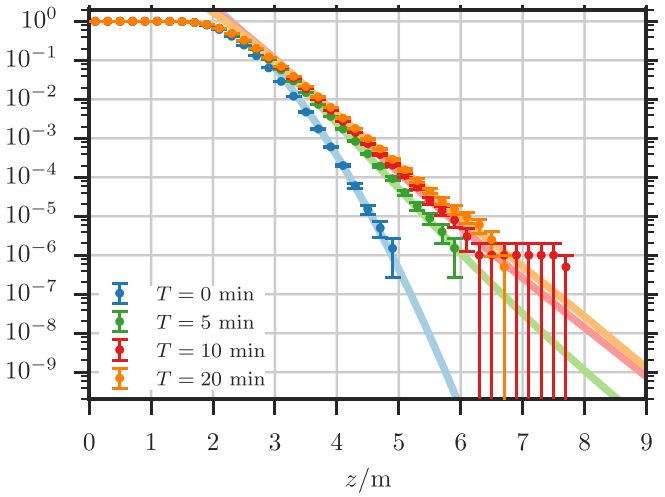


Fig. S7. $\mathbb{P}(\max_x|u(t, x)| \geq z)$ for $u(t, x)$ solution of NLS at different times calculated by Monte Carlo sampling using 10^6 realizations and compared with the results obtained via optimization in our large deviation approach.

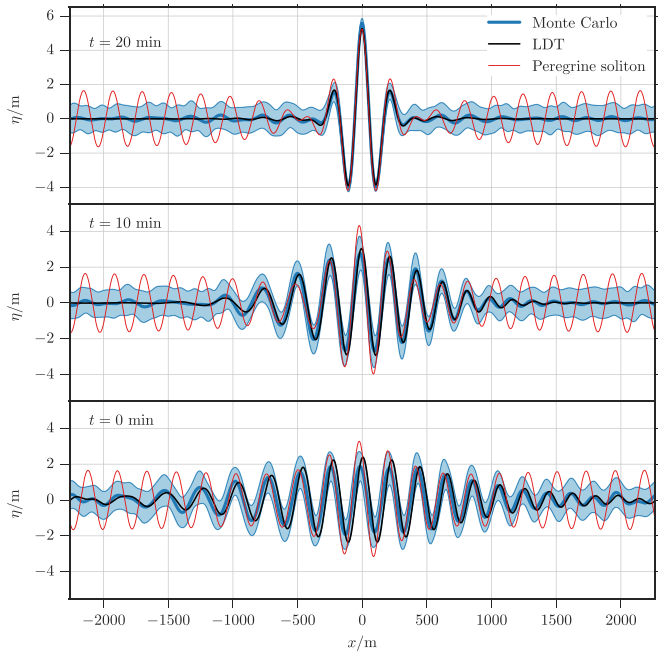


Fig. S8. Comparison of the optimized solution, the mean and SD of the Monte Carlo realizations, and the PS reaching the same maximal surface elevation at $T = 20$ min. From bottom to top, the figures are at 0, 10, and 20 minutes respectively, and these results are for NLS.

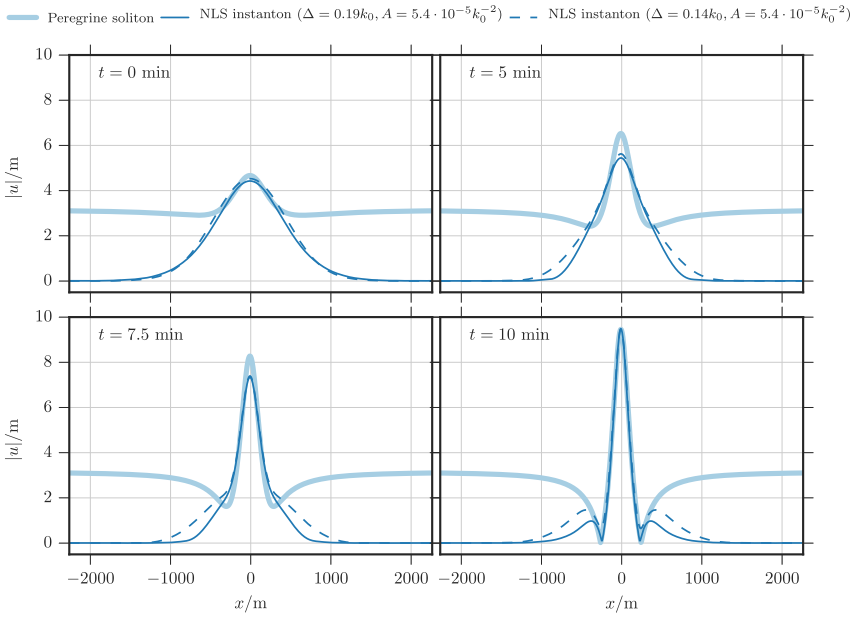


Fig. S9. Comparison between the optimized solution for a very extreme surface elevation and the PS reaching the same final height (after $T = 10$ min). Comparison with realization from the Monte Carlo sampling is impossible due to the extreme rareness of such event on the ensemble of the initial conditions. The evolution is shown at times 0, 5, 7.5, and 10 min, respectively. These results are for NLS.

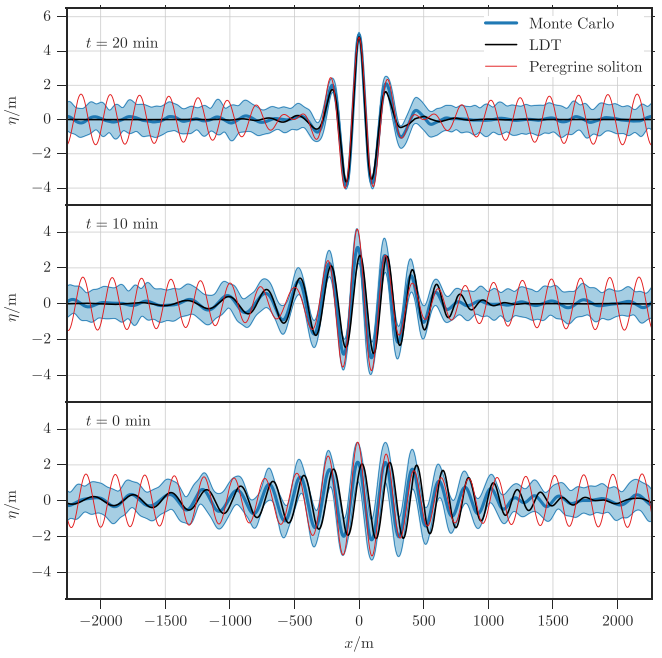
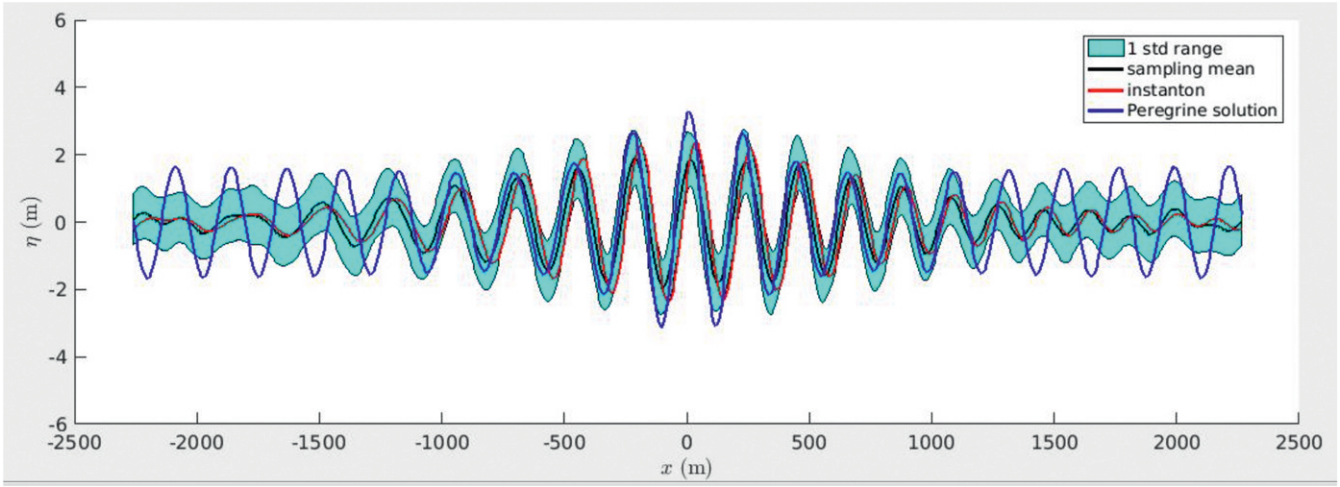
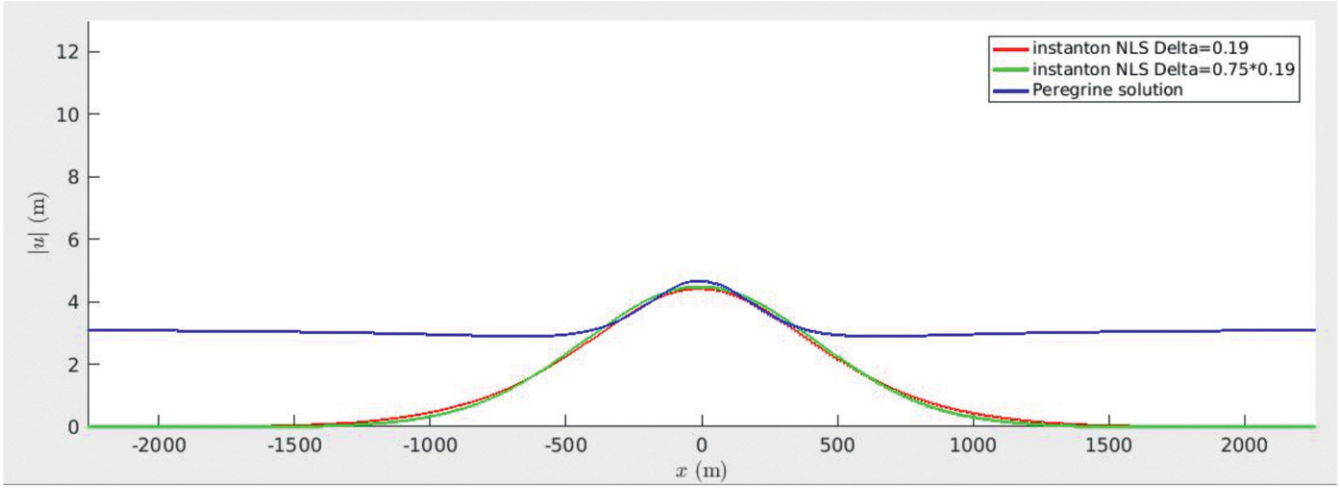


Fig. S10. Same as in Fig. S8 for MNLS.



Movie S1. Time evolution of the surface elevation of the optimized solution and the PS reaching the same maximal amplitude $\max_x|u(T,x)| = 5.25$ m at $T = 20$ min, compared with that of the mean and SD of the trajectories sampled by Monte Carlo that reach $\max_x|u(T,x)| \geq 5.25$ m. These calculations were performed in the context of the NLS equation, for which the PS is an exact solution.

Movie S1



Movie S2. Comparison between two instantons and the PS reaching the same maximal amplitude $\max_x|u(T,x)|$ m at $T = 10$ min. The two instantons are optimized solutions for two statistical states of the sea with a different spectral width Δ. These calculations were performed in the context of the NLS equation, for which the PS is an exact solution. For the event shown here, the extreme size and rareness make comparison with the Monte Carlo sampling impossible in practice.

Movie S2