The inverse scattering of the Zakharov-Shabat system solves the weak noise theory of the Kardar-Parisi-Zhang equation

Alexandre Krajenbrink, Pierre Le Doussal

The inverse scattering of the Zakharov-Shabat system solves the weak noise theory of the Kardar-Parisi-Zhang equation

Alexandre Krajenbrink*
SISSA and INFN, via Bonomea 265, 34136 Trieste, Italy

Pierre Le Doussal†
Laboratoire de Physique de l'École Normale Supérieure, CNRS, ENS & PSL University, Sorbonne Université, Université de Paris, 75005 Paris, France

(Dated: May 21, 2021)

We solve the large deviations of the Kardar-Parisi-Zhang (KPZ) equation in one dimension at short time by introducing an approach which combines field theoretical, probabilistic and integrable techniques. We expand the program of the weak noise theory, which maps the large deviations onto a non-linear hydrodynamic problem, and unveil its complete solvability through a connection to the integrability of the Zakharov-Shabat system. Exact solutions, depending on the initial condition of the KPZ equation, are obtained using the inverse scattering method and a Fredholm determinant framework recently developed. These results, explicit in the case of the droplet geometry, open the path to obtain the complete large deviations for general initial conditions.

Large deviation rate functions characterize rare events and play a key role in non-equilibrium statistical physics, as generalizations of the thermodynamic potentials [1–3]. They have been much studied for interacting particle models in one dimension. For diffusive systems, the macroscopic fluctuation theory (MFT) [4] provides a powerful framework to calculate the large deviation of the density, of the current and of other fluctuating quantities, in agreement with the available exact solutions [5]. For driven diffusive systems, however, such as the asymmetric exclusion process (ASEP) [6], there is not yet a general approach to calculate the large deviations for all geometries. Exact results from the matrix product ansatz [7] and the Bethe ansatz are available in special cases, for instance in a stationary regime, either on a finite ring, where rate functions are found to exhibit a universal shape, for the TASEP [8, 9], the ASEP [10, 11] and the Kardar-Parisi-Zhang (KPZ) equation [12, 13], or in open geometries [14–16].

The KPZ equation is a prominent example of the driven diffusive class. It allows for a few exact solutions valid for all times [17–31], which exhibit at large times the universal typical fluctuations common to systems in the KPZ class [32–34]. Much recent attention has shifted to its large deviation properties, at late times [35–46], and also at short times [37, 38, 47–64] where two main approaches were developed. The first one uses the aforementioned exact solutions for all times, obtained from a mapping of KPZ observables to the integrable (replica) delta Bose gas. This allowed to obtain the short time large deviations in a few cases [37, 38, 47–50]. A more versatile approach, closer in spirit to the MFT, is the weak noise theory (WNT) [53–63, 65]. It is a saddle point method on the dynamical field theory, which is exact at short time. It leads to a system of two coupled non-linear partial differential equations, which determine the "optimal" KPZ height field and noise producing the rare fluctuation. Until now however these equations have been solved only numerically, except in some limits where useful but approximate solutions were found. Although the existence of multi-soliton solutions was noted [55], no exact solution allowing for the full calculation of the large deviations was obtained.

In this Letter we construct the exact solution to the weak noise theory of the KPZ equation. Through the integrability of the Zakharov-Shabat (ZS) system, originally introduced to solve the non-linear Schrodinger equation (NLS) [66], we show that the full space time dependence of the optimal height and noise fields admit representations in terms of Fredholm determinants. We provide an explicit formula for the KPZ droplet initial condition, and give the general form for a large class of initial conditions.

The KPZ equation [67] describes the stochastic growth in time τ of the height field $h(y,\tau)$ of an interface, here in one space dimension $y \in \mathbb{R}$

$$\partial_\tau h(y,\tau) = \nu \partial_y^2 h(y,\tau) + \frac{\lambda_0}{2}(\partial_y h(y,\tau))^2 + \sqrt{D}\eta(y,\tau) \tag{1}$$

where $\eta(y,\tau)$ is a standard space time white noise, i.e. $\overline{\eta(y,\tau)\eta(y',\tau')} = \delta(\tau-\tau')\delta(y-y')$, where $\overline{\cdots}$ denotes averages over the noise. We choose units such that $D = \lambda_0 = 2$, $\nu = 1$ [68]. We consider the probability $P(H,T)$ to observe the value $h(0,T) = H - H_0$ at time $\tau = T$, where H_0 is a constant chosen below. At short time, although the typical height fluctuations are Gaussian with Edwards-Wilkinson scaling $\delta H \sim T^{1/4}$, the KPZ non-linearity leads to non-trivial and non-perturbative tails for $P(H,T)$, describing rare events. For $T \ll 1$, it takes the large deviation form

$$P(H,T) \sim \exp(-\Phi(H)/\sqrt{T}) \tag{2}$$

where the exact rate function $\Phi(H)$ was obtained for droplet, Brownian and flat initial height profiles, from the exact solutions [47, 48, 50, 58].

We now explain how to obtain such rate function from the WNT: we first derive the WNT equations in a way leading directly to the so-called $\{P,Q\}$ system, which we then analyze. To that aim, it is useful to define the rescaled time and space variables as $t = \tau/T$, $x = y/\sqrt{T}$, where T, the observation time, is fixed. Through the Cole-Hopf map the KPZ field is equivalently described introducing $Z(x,t) = e^{h(y,\tau)+H_0}$, which satisfies the (rescaled) stochastic heat equation (SHE) in the Ito sense

$$\partial_t Z(x,t) = \partial_x^2 Z(x,t) + \sqrt{2}T^{1/4}\tilde{\eta}(x,t)Z(x,t) \tag{3}$$

where $\tilde{\eta}(x,t)$ is another standard space time white noise. This equation is now studied for $t \in [0,1]$. The noise amplitude is now of order $T^{1/4}$, hence a short observation time $T \ll 1$ corresponds to a weak noise. Our convenient choice is $H_0 = \frac{1}{2}\log T$ [69]. It is convenient to study the following generating function which admits a large deviation principle at short time $T \ll 1$, with $z \geqslant 0$

$$\overline{\exp(-ze^H/\sqrt{T})} \sim \exp(-\Psi(z)/\sqrt{T}) \tag{4}$$

Inserting (2) into the l.h.s., we see that for $T \ll 1$, $\Psi(z)$ and $\Phi(H)$ are related through a Legendre transform

$$\Psi(z) = \min_H(ze^H + \Phi(H)) \tag{5}$$

Here we aim to calculate $\Psi(z)$ and $\Phi(H)$ using the WNT, for an initial condition of the form $e^{h(y,0)} = \frac{1}{\sqrt{T}}Z_0(y/\sqrt{T})$ where $Z_0(x)$ is given, an example being the droplet initial condition $Z_0(x) = \delta(x)$.

Any average of the form (4) can be represented using the dynamical field theory associated to the rescaled SHE (3) as $\overline{e^{\frac{1}{\sqrt{T}}\iint \mathrm{d}t\mathrm{d}x j(x,t)Z(x,t)}} = \int \mathcal{D}Z\mathcal{D}\tilde{Z}e^{-\frac{1}{\sqrt{T}}S[\tilde{Z},Z,j]}$ with the dynamical action

$$S[\tilde{Z},Z,j] = \int_0^{+\infty}\mathrm{d}t\int_{\mathbb{R}}\mathrm{d}x[\tilde{Z}(\partial_t - \partial_x^2)Z - \tilde{Z}^2Z^2 - jZ] \tag{6}$$

where $\tilde{Z}/\sqrt{T}$ is the response field. In (4) the source field is $j(x,t) = -z\delta(x)\delta(t-1)$. For $T \ll 1$, the action is evaluated by a saddle point method. Defining $\tilde{Z} = -zP$, $Q = Z$ and $g = -z$, the saddle point equations of the WNT, $\frac{\delta S}{\delta \tilde{Z}} = 0$ and $\frac{\delta S}{\delta Z} = 0$, take the form of the $\{P,Q\}$ system

$$\begin{aligned}\partial_t Q &= \partial_x^2 Q + 2gPQ^2\\ -\partial_t P &= \partial_x^2 P + 2gP^2Q\end{aligned} \tag{7}$$

a close cousin of the NLS equation [70], which was also discussed in [55]. While the $\{P,Q\}$ system is interesting in its own right, we will apply its study to the following mixed boundary conditions, of interest for the WNT

$$Q(x,0) = Q_0(x), \quad P(x,1) = \delta(x) \tag{8}$$

The source j imposes this form for P at $t = 1$ [71, 72], while Q is specified at $t = 0$ from the initial height of the KPZ equation, i.e. $Q_0(x) = Z_0(x)$. The function $\Psi(z)$ in (4) is obtained from the action S in (6) at the saddle point. Using the first equation in (7) it can be written in the form (5) allowing to identify $\Phi(H) \equiv g^2\int_0^1 \mathrm{d}t\int_{\mathbb{R}}\mathrm{d}xP^2Q^2$, with $H = H_z^* := \arg\min_H(ze^H + \Phi(H))$, in agreement with [53] (see also [72]). The "optimal shape" $h_{\rm opt}(y,\tau)$ of the KPZ height field from the WNT, i.e. the most probable one realizing the value $h_{\rm opt}(0,T) = H - H_0$ at $\tau = T$ and $y = 0$, is obtained from the solution $Q(x,t)$ of (7) for $t \in [0,1]$ as $e^{h_{\rm opt}(y,\tau)} = \frac{1}{\sqrt{T}}Q(\frac{y}{\sqrt{T}},\frac{\tau}{T})$.

Let us first analyze the $\{P,Q\}$ system (7) for general initial conditions, and return to the WNT later. Remarkably, (7) belongs to the AKNS class of integrable non-linear problems [73], for which there exists a Lax pair, i.e. a pair of linear differential equations whose compatibility conditions are equivalent to (7). Here the system reads $\partial_x\vec{v} = U_1\vec{v}$, $\partial_t\vec{v} = U_2\vec{v}$ where $\vec{v} = (v_1,v_2)^\intercal$ is a two component vector (depending on x,t,k) where

$$U_1 = \begin{pmatrix} -\mathbf{i}k/2 & -gP(x,t)\\ Q(x,t) & \mathbf{i}k/2\end{pmatrix} \quad , \quad U_2 = \begin{pmatrix}\mathsf{A} & \mathsf{B}\\ \mathsf{C} & -\mathsf{A}\end{pmatrix} \tag{9}$$

where $\mathsf{A} = k^2/2 - gPQ$, $\mathsf{B} = g(\partial_x - \mathbf{i}k)P$, $\mathsf{C} = (\partial_x + \mathbf{i}k)Q$. One can check that the compatibility condition, $\partial_t U_1 - \partial_x U_2 + [U_1,U_2] = 0$, recovers the system (7) which we solve through the following scattering problem. Let $\vec{v} = e^{k^2 t/2}\phi$ with $\phi = (\phi_1, \phi_2)^\intercal$ and $\vec{v} = e^{-k^2 t/2}\bar{\phi}$ be two independent solutions of the linear problem such that at $x \to -\infty$, $\phi \simeq (e^{-\mathrm{i}kx/2}, 0)^\intercal$ and $\bar{\phi} \simeq (0, -e^{\mathrm{i}kx/2})^\intercal$. Assuming from now on that P, Q vanish at infinity, the $x \to +\infty$ behavior of these solutions defines scattering amplitudes

$$\phi \underset{x\to+\infty}{\simeq} \begin{pmatrix} a(k,t)e^{-\frac{\mathrm{i}kx}{2}} \\ b(k,t)e^{\frac{\mathrm{i}kx}{2}} \end{pmatrix} \;,\; \bar{\phi} \underset{x\to+\infty}{\simeq} \begin{pmatrix} \tilde{b}(k,t)e^{-\frac{\mathrm{i}kx}{2}} \\ -\tilde{a}(k,t)e^{\frac{\mathrm{i}kx}{2}} \end{pmatrix} \tag{10}$$

Plugging this form into the ∂_t equation of the Lax pair at $x \to +\infty$, one finds a very simple time dependence, $a(k,t) = a(k)$ and $b(k,t) = b(k)e^{-k^2 t}$, $\tilde{a}(k,t) = \tilde{a}(k)$ and $\tilde{b}(k,t) = \tilde{b}(k)e^{k^2 t}$. Another relation is obtained from the Wronskian of the two solutions , $a(k)\tilde{a}(k) + b(k)\tilde{b}(k) = 1$ [74].

Before providing explicit formula for these scattering amplitudes let us show how to obtain from them the solution for the $\{P, Q\}$ system, i.e. how to construct the inverse scattering transform. The spatial part of the Lax pair is a 1D Dirac equation called the ZS system, originally introduced to solve the NLS equation [66, 75], and extended by AKNS [73]. It was shown very recently [76, 77] that the inverse scattering problem can be solved by the means of Fredholm determinants (FD). Introducing the two reflection coefficients $r(k) = b(k)/a(k)$ and $\tilde{r}(k) = \tilde{b}(k)/(g\tilde{a}(k))$ one defines two functions [78]

$$A_t(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} r(k) e^{\mathrm{i}kx - k^2 t}, B_t(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} \tilde{r}(k) e^{k^2 t - \mathrm{i}kx} \tag{11}$$

and two linear operators from $\mathbb{L}^2(\mathbb{R}^+)$ to $\mathbb{L}^2(\mathbb{R}^+)$ with respective kernels

$$\mathcal{A}_{xt}(v, v') = A_t(x+v+v') \;,\; \mathcal{B}_{xt}(v, v') = B_t(x+v+v') \tag{12}$$

Note that these functions and kernels obey the simple heat equation in space time, and we assume that $A_t(x)$, $B_t(x)$ vanish fast enough for $x \to +\infty$. The solutions P, Q [76, 77] are reconstructed as

$$\begin{aligned} Q(x,t) &= \langle\delta|\, \mathcal{A}_{xt}(I + g\mathcal{B}_{xt}\mathcal{A}_{xt})^{-1} \,|\delta\rangle \\ P(x,t) &= \langle\delta|\, \mathcal{B}_{xt}(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1} \,|\delta\rangle \end{aligned} \tag{13}$$

where $|\delta\rangle$ is the vector with component $\delta(v)$ so that $\langle\delta|\, \mathcal{O} \,|\delta\rangle = \mathcal{O}(0,0)$ for any operator $\mathcal{O}$. The product PQ, which is a conserved charge, i.e. $\partial_t \int_{\mathbb{R}} \mathrm{d}x PQ = 0$ as easily verified from (7), can be expressed from a FD as $gPQ = \partial_x^2 \log \mathrm{Det}(I + g\mathcal{B}_{xt}\mathcal{A}_{xt})$. The formula (13) thus provides the general solution of the $\{P, Q\}$ system, parameterized by the two functions A_t and B_t, equivalently, by the scattering amplitudes. Although these are in one-to-one correspondence with the $\{P, Q\}$ boundary data, making it explicit is non-trivial, and is our aim below. Particular cases are such that $\mathcal{A}_{xt}$ and $\mathcal{B}_{xt}$ are operators of finite ranks, leading to solitonic type solutions [72]. The simplest one leads to $gPQ = \frac{(\kappa+\mu)^2}{4\cosh^2 \frac{1}{2}(\kappa+\mu)(x - x_0(t))}$. In the context of WNT this soliton has been used as an approximate solution for $H \to +\infty$, and another rank one family was noticed in [55]. However this is insufficient to obtain the full rate function $\Phi(H)$ which requires the (infinite-rank) general solution obtained in this work.

Let us now apply this to the WNT, i.e. for the boundary data in (8), and characterize the scattering amplitudes. Integrating the ∂_x equation of the Lax pair at $t = 1$ for $\bar{\phi}$ and ϕ using (8) allows to obtain [72] that $\tilde{b}(k) = ge^{-k^2}$. In addition, if the initial condition $Q(x, 0)$ is even in x (which we assume from now on) then $\tilde{a}(k) = a(-k) = (a(k))^*$ and $b(k)$ is real and even. From the Wronskian this leads to the form

$$a(k) = e^{-\mathrm{i}\varphi(k)} \sqrt{1 - gb(k)e^{-k^2}} \tag{14}$$

where we still have two unknown functions, a phase $\varphi(k)$, which is odd $\varphi(k) = -\varphi(-k)$, and $b(k)$.

To determine them, we have derived from (12)-(13), and the boundary data (8) at $t = 1$, the following integral equation which the functions $A_{t=1}, B_{t=1}$ must obey, see [72]

$$B_1(x) = \delta(x) + g\Theta(-x) \int_0^{+\infty} \mathrm{d}v B_1(x+v) A_1(v) \tag{15}$$

where $A_1(v) = (p * A_0)(v) := \int_{\mathbb{R}} \mathrm{d}y p(v - y) A_0(y)$ denoting the heat kernel at unit time $p(z) := \frac{e^{-z^2/4}}{\sqrt{4\pi}}$.

Droplet initial condition. Let us now specialize to $Q_0(x) = \delta(x)$. The solution of (7) then satisfies the symmetry $Q(x,t) = P(x, 1-t)$. This in turn implies that $A_t(x) = B_{1-t}(x)$ and $r(k) = e^{k^2}\tilde{r}(-k)$, also implying $b(k) = 1$, which we use below. Hence in (15) we can replace $A_1(v)$ by $(p * B_1)(v)$ and we obtain a closed *non-linear* integral equation for the function B_1 (which equals A_0). This equation still looks formidable, however, for readers familiar with random walks, it has a flavor of another famous integral equation, the Hopf-Ivanov (HI) equation [79, 80], which however is *linear*, and reads

$$B_1(x) = \delta(x) + g\Theta(-x) \int_{-\infty}^{0} \mathrm{d}y p(x - y) B_1(y) \tag{16}$$

Amazingly, we found that these two equations, (16) and (15) are *equivalent*. This can be tested in perturbation in g, and is shown to all orders in [72]. The HI equation arises in survival probabilities of random walks [81–84]. Indeed writing $B_1(x)$ as a series, $B_1(x) = \delta(x) + \sum_{n=1}^{+\infty} g^n B_{1,n}(x)$, and inserting in (16), leads to the recursion $B_{1,n}(x) = \int_{y<0} p(x-y) B_{1,n-1}(y)$. The interpretation is then straightforward. Consider $X(j) \in \mathbb{R}$ a discrete time random walk, $X(j+1) = X(j) + z_j$, with z_j i.i.d. with jump probability $p(z)$. Then $B_{1,n}(x)$ is the probability that the walk starting at $X(0) = 0$ arrives at $X(n) = x$ in $n \geqslant 1$ steps, while remaining negative, $\{X(j) \leqslant 0\}_{j=0,\dots,n}$ (and $\int_{-\infty}^{0} \mathrm{d}x B_{1,n}(x) = \binom{2n}{n} 2^{-2n}$ is given by the universal Sparre-Andersen theorem [85, 86]).

To show that the solution of (16) also solves (15) then amounts to split the walk into two independent parts, upon crossing the level x for the last time, [72]. Introducing the Laplace transform $\hat{B}_1(s) = \int_{-\infty}^{0} e^{sx} B_1(x)$, the solution of the HI equation is known to be [80]

$$\hat{B}_1(s) = \exp\Big(-\int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \frac{s}{s^2+q^2} \log(1 - g\tilde{p}(q))\Big) \quad (17)$$

where $\tilde{p}(k)$ is the Fourier transform of $p(z)$, here $\tilde{p}(k) = e^{-k^2}$. Going from Laplace to Fourier, from (11) one finds $r(k) = e^{k^2}\tilde{r}(-k) = \hat{B}_1(s)|_{s=-\mathbf{i}k+0^+}$. Using $\frac{1}{s+\mathbf{i}q} \to PV(\frac{\mathbf{i}}{k-q}) + \pi\delta(k-q))$ we obtain from (17), the reflection coefficient $r(k)$ and its phase $\varphi(k)$

$$r(k) = \frac{e^{\mathbf{i}\varphi(k)}}{\sqrt{1-ge^{-k^2}}}\ ,\ \varphi(k) = \int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \frac{k\log(1-ge^{-q^2})}{q^2-k^2} \quad (18)$$

which, together with (11)-(13), completes the solution of (7) for droplet IC. Plots of the optimal height, $\log Q(x,t)$, are shown in Fig.1, and [72]. Note that for $t=1$ a simpler formula, $Q(x,1) = A_1(|x|)$ holds.

To extract $\Phi(H) \equiv g^2 \int_{\mathbb{R}} \int_0^1 \mathrm{d}x\mathrm{d}t\, P^2 Q^2$ from our solution requires the computation of a difficult integral. This is overcome by relating it, as well as $\Psi(z)$, to conserved quantities. We use the construction of ZS [66] to generate all conserved quantities C_n for the $\{P,Q\}$ system, see [72]. We find $C_1 = g\int_{\mathbb{R}} \mathrm{d}xPQ$, $C_3 = g(\int_{\mathbb{R}} \mathrm{d}xP\partial_x^2 Q + gP^2Q^2)$. The *values* $C_n(g)$ taken by these conserved charges can be retrieved from the Laurent expansion of $\log a(k) = \sum_{n\geqslant 1} \frac{C_n(g)}{(\mathbf{i}k)^n}$. Until now this is general for any initial condition of the $\{P,Q\}$ system. Now recall that for the droplet IC we obtained $\log a(k) = -\mathbf{i}\varphi(k) + \frac{1}{2}\log(1-ge^{-k^2})$. Since the second term has vanishing Laurent expansion, we find that $-\mathbf{i}\varphi(k) = \sum_{n\geqslant 1} \frac{C_n(g)}{(\mathbf{i}k)^n}$. Expanding in powers of $1/k$ in (18)

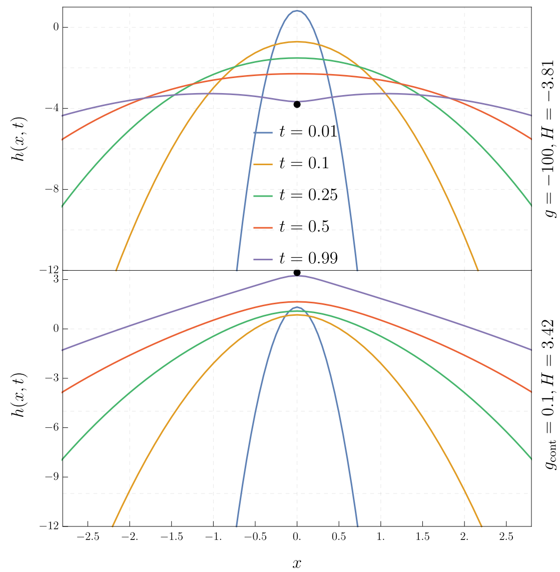

Figure 1. The optimal height $h(x,t) = \log Q(x,t)$ for droplet initial condition at various times t for final values (black dot) $H = -3.81$ (**top**) and $H = 3.42$ (**bottom**).

we obtain [87]

$$C_1(g) = \frac{1}{\sqrt{4\pi}} \mathrm{Li}_{\frac{3}{2}}(g)\ ,\ C_3(g) = \frac{-1}{\sqrt{16\pi}} \mathrm{Li}_{\frac{5}{2}}(g) \quad (19)$$

Since $C_1 = g\int_{\mathbb{R}} \mathrm{d}xPQ$ is time independent, evaluated at $t=1$ it leads to $C_1(g) = gQ(0,1) = ge^H$. On the other hand, differentiating the Legendre transform in (5) w.r.t. z gives $\Psi'(z) = e^H$. This implies that $C_1(-z) = -z\Psi'(z)$, and by integration

$$\Psi(z) = \Psi_0(z) := -\frac{1}{\sqrt{4\pi}} \mathrm{Li}_{5/2}(-z) \quad (20)$$

which allows to determine $\Phi(H)$ parametrically as $\Phi(H) = \Psi(z) - z\Psi'(z)$, $e^H = \Psi'(z)$ (it can also be obtained from $C_3(g)$, see [72]). Our WNT result (20) agrees with [47], without relying on an exact solution of the KPZ equation.

Until now we assumed $z = -g \geqslant 0$ corresponding to $H \leqslant \hat{H}_0 = -\frac{1}{2}\log(4\pi)$, the most probable value of H such that $\Phi'(\hat{H}_0) = 0$ [88]. However (7) also holds for any $H > \hat{H}_0$ [55, 72], corresponding to the attractive regime $g > 0$ of the $\{P,Q\}$ system. Indeed, $\Psi(z)$ can be analytically continued to $z < 0$, allowing to determine $\Phi(H)$ for any H [47]. For $H \in (-\infty, H_c]$, (20) holds, with $z = -g$ varying from $+\infty$ down to $z = -1$. For $H > H_c = \log \frac{\zeta(3/2)}{\sqrt{4\pi}}$, a second continuation is needed, $\Psi(z) = \Psi_0(z) + \Delta(z)$, with $\Delta(z) = \frac{4}{3}(\log(-\frac{1}{z}))^{3/2}$ with $z \in [-1, 0)$ as $H \in [H_c, +\infty)$. These continuations correspond to two branches of solutions of the $\{P, Q\}$ system for $0 < g \leqslant 1$. One finds [72] that the second branch corresponds to the spontaneous generation of a solitonic part in the solution, of rapidity κ_0 with $g = e^{-\kappa_0^2}$, which dominates the large deviations for $H \to +\infty$. It is described by $A_t(x) = A_t(x)|_{\phi(k)\to\phi(k)+\Delta\phi(k)} + 2\kappa_0 e^{-\kappa_0 x + \kappa_0^2 t + \mathrm{i}\varphi(\mathrm{i}\kappa_0)}$, where $\Delta\varphi(k) = 2\arctan(\frac{\kappa_0}{k})$ and $B_t(x) = A_{1-t}(x)$. The values of the odd conserved charges are increased by $\Delta C_n(g) = \frac{2}{n}\kappa_0^n$, which for $n = 1$ induces the additional part $\Delta(z)$.

General initial condition. For general even $Q_0(x)$, the only difference is that $b(k)$ is non-trivial, with $r(k) = b(k)e^{k^2}\tilde{r}(-k)$, and now $A_t(x) = (\hat{b} * B_{1-t})(x)$ where $\hat{b}(x)$ denotes the Fourier transform of $b(k)$ and $*$ the convolution. Equation (15), replacing $A_1(v) = (p * \hat{b} * B_1)(v)$, is again equivalent to a linear HI equation for $B_1(x)$, obtained by simply replacing p by $p * \hat{b}$ in (16), with the same random walk interpretation for a new jump probability $p(z) \to (p * \hat{b})(z)$. Thus this leads to $r(k) = \frac{b(k)e^{\mathrm{i}\varphi(k)}}{\sqrt{1-gb(k)e^{-k^2}}}$ and $\varphi(k) = \fint_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \frac{k\log(1-gb(q)e^{-q^2})}{q^2-k^2}$. One now finds $C_1(g) = \int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \mathrm{Li}_1(gb(q)e^{-q^2})$ leading to [89]

$$z\Psi'(z) = \int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \log(1 + zb(q)e^{-q^2}) \qquad (21)$$

Interestingly, such a form was observed to describe all known exact solutions [90], e.g. flat IC [91]. Thus, for a general initial condition we reduced the problem to computing a single unknown function $b(k)$, and relating it to $Q_0(x)$, a question left for the future. In [72] we give a formula relating $b(k)$, $Q_0(x)$ and $P(x, 1)$ allowing for expansions around the droplet solution.

In conclusion our solution allows to calculate the optimal height and noise for arbitrary values of H, previously inaccessible. The Fredholm approach provides a novel analytical and numerical scheme for the solution of the integrable $\{P, Q\}$ system as shown in Fig. 1, see [72]. The present work demonstrates that inverse scattering methods can successfully address optimal fluctuation theory of stochastic systems, leading to analytic results and interesting phenomena such as spontaneous soliton generation.

Acknowledgments. We thank T. Bothner, S. N. Majumdar, B. Meerson, G. Schehr and N. R. Smith for discussions on related topics. AK acknowledges support from ERC under Consolidator grant number 771536 (NEMO). PLD acknowledges support from the ANR grant ANR-17-CE30-0027-01 RaMaTraF.

* alexandre.krajenbrink@sissa.it
† ledou@lpt.ens.fr

[1] Daniel Nickelsen and Hugo Touchette. *Anomalous scaling of dynamical large deviations.* Physical review letters, 121(9):090602, (2018). Hugo Touchette. *The large deviation approach to statistical mechanics.* Physics Reports, 478(1-3):1–69, (2009).

[2] Juan P. Garrahan *Aspects of non-equilibrium in classical and quantum systems: slow relaxation and glasses, dynamical large deviations, quantum non-ergodicity, and open quantum dynamics* Physica A 504, 130, (2018).

[3] Yongjoo Baek and Yariv Kafri. *Singularities in large deviation* functions. Journal of Statistical Mechanics: Theory and Experiment, 2015(8):P08026, (2015).

[4] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio and C. Landim, *Macroscopic fluctuation theory.* Reviews of Modern Physics, 87(2):593, (2015).

[5] B. Derrida, *Non equilibrium steady states: fluctuations and large deviations of the density and of the current* Journal of Statistical Mechanics: Theory and Experiment, 2007(07):P07023, (2007).

[6] Derrida B, *An exactly soluble non-equilibrium system: the asymmetric simple exclusion process.* Physics Reports, 301(1-3):65–83, (1998).

[7] Derrida B, Evans M R, Hakim V and Pasquier V, *Exact solution of a 1d asymmetric exclusion model using a matrix formulation.* Journal of Physics A: Mathematical and General, 26(7):1493, (1993).

[8] Derrida B and Lebowitz J L, *Exact large deviation function in the asymmetric exclusion process.* Physical review letters, 80(2):209, (1998).

[9] S. Prolhac, *Current fluctuations and large deviations for periodic TASEP on the relaxation scale.* Journal of Statistical Mechanics: Theory and Experiment, 2015(11):P11028, (2015).

[10] D. S. Lee, D. Kim, *Large deviation function of the partially asymmetric exclusion process.* Physical Review E, 59(6):6476, (1999).
A. E. Derbyshev, A. M. Povolotsky, V. B. Priezzhev, *Emergence of jams in the generalized totally asymmetric simple exclusion process.* Physical Review E, 91(2):022125, (2015).
T. C. Dorlas, A. M. Povolotsky, V. B. Priezzhev, *From vicious walkers to TASEP.* Journal of Statistical Physics, 135(3):483–517, (2009). D. S. Lee,

D. Kim, *Universal fluctuation of the average height in the early-time regime of one-dimensional Kardar–Parisi–Zhang-type growth.* Journal of Statistical Mechanics: Theory and Experiment, 2006(08):P08014, (2006).

[11] S. Prolhac *Finite-time fluctuations for the totally asymmetric exclusion process*, Physical Review Letters **116**:090601, (2016).

[12] Derrida B and Appert C, *Universal large deviation function of the Kardar–Parisi–Zhang equation in one dimension*, J. Stat. Phys. 94 1, (1999)

[13] E. Brunet, B. Derrida, *Probability distribution of the free energy of a directed polymer in a random medium.* Physical Review E, 61(6):6789, (2000). *Ground state energy of a non-integer number of particles with δ attractive interactions.* Physica A: Statistical Mechanics and its Applications, 279(1-4):398–407, (2000).

[14] Derrida B, Lebowitz J L and Speer E R, *Exact large deviation functional of a stationary open driven diffusive system: the asymmetric exclusion process*, J. Stat. Phys. 110 775, (2003).

[15] M. Gorissen, A. Lazarescu, K. Mallick and C. Vanderzande *Exact Current Statistics of the Asymmetric Simple Exclusion Process with Open Boundaries.* Physical Review Letters **109**:170601, (2012)

[16] N. Crampe, R.I. Nepomechie, *Equivalent T-Q relations and exact results for the open TASEP*, Journal of Statistical Mechanics: Theory and Experiment, 2018(10):103105, (2018).

[17] P. Calabrese, P. Le Doussal, A. Rosso, *Free-energy distribution of the directed polymer at high temperature*, Europhys. Lett. **90**, 20002, (2010).

[18] V. Dotsenko, *Bethe ansatz derivation of the Tracy-Widom distribution for one-dimensional directed polymers* EPL **90**, 20003 (2010); *Replica Bethe ansatz derivation of the Tracy-Widom distribution of the free energy fluctuations in one-dimensional directed polymers*, J. Stat. Mech. P07010, (2010); V. Dotsenko and B. Klumov, *Bethe ansatz solution for one-dimensional directed polymers in random media* J. Stat. Mech. P03022, (2010).

[19] T. Sasamoto, H. Spohn, *One-dimensional Kardar-Parisi-Zhang equation: an exact solution and its universality*, Phys. Rev. Lett. **104**, 230602, (2010).

[20] G. Amir, I. Corwin, J. Quastel, *Probability distribution of the free energy of the continuum directed random polymer in 1 + 1 dimensions*, Comm. Pure and Appl. Math. **64**, 466, (2011).

[21] P. Calabrese, M. Kormos and P. Le Doussal, *From the sine-Gordon field theory to the Kardar-Parisi-Zhang growth equation*, arXiv:1405.2582, EPL 107 10011, (2014).

[22] P. Calabrese, P. Le Doussal, *Exact solution for the Kardar-Parisi-Zhang equation with flat initial conditions*, Phys. Rev. Lett. **106**, 250603, (2011); *The KPZ equation with flat initial condition and the directed polymer with one free end*, J. Stat. Mech. P06001, (2012).

[23] J. Ortmann, J. Quastel, D. Remenik, *Exact formulas for random growth with half-flat initial data*, arXiv:1407.8484, Annals of Applied Probability, Vol. 26, No. 1, 507-548, (2016).

[24] T. Imamura, T. Sasamoto, *Exact solution for the stationary Kardar-Parisi-Zhang equation*, Phys. Rev. Lett. **108**, 190603, (2012).

[25] T. Imamura, T. Sasamoto, *Stationary correlations for the 1D KPZ equation*, J. Stat. Phys. **150**, 908-939 , (2013).

[26] A. Borodin, I. Corwin, P. L. Ferrari. B. Veto, *Height fluctuations for the stationary KPZ equation*, Math. Phys. Anal. Geom. **18**, 20, (2015).

[27] T. Gueudré, P. Le Doussal, *Directed polymer near a hard wall and KPZ equation in the half-space*, Europhys. Lett. **100**, 26006, (2012).

[28] A. Borodin, A. Bufetov, I. Corwin, *Directed random polymers via nested contour integrals.* Annals of Physics, **368** 191–247, (2016).

[29] G. Barraquand, A. Borodin, I. Corwin, M. Wheeler, *Stochastic six-vertex model in a half-quadrant and half-line open ASEP*, arXiv:1704.04309, (2017).

[30] A. Krajenbrink, P. Le Doussal, *Replica Bethe Ansatz solution to the Kardar-Parisi-Zhang equation on the half-line.* SciPost Phys. **8** 035, (2019).

[31] G. Barraquand, A. Krajenbrink, P. Le Doussal. *Half-Space Stationary Kardar–Parisi–Zhang Equation*, J. Stat. Phys. **181** 1149–1203, (2020).

[32] I. Corwin, *The Kardar-Parisi-Zhang equation and universality class*, Random Matrices: Theory Appl. **01** 1130001, (2012).

[33] J. Quastel and H. Spohn, The one-dimensional KPZ equation and its universality class, J. Stat.Phys. 160, 965–984, (2015).

[34] K. A. Takeuchi. *An appetizer to modern developments on the Kardar–Parisi–Zhang universality class.* Physica A: Statistical Mechanics and its Applications, 504:77– 105, (2018).

[35] P. Le Doussal, S. N. Majumdar, G. Schehr, *Large deviations for the height in 1D Kardar-Parisi-Zhang growth at late times*, Europhys. Lett. **113**, 60004, (2016).

[36] P. Sasorov, B. Meerson, S. Prolhac, *Large deviations of surface height in the 1+1 dimensional Kardar-Parisi-Zhang equation: exact long-time results for* $\lambda H < 0$, J. Stat. Mech. 063203, (2017).

[37] A. Krajenbrink, P. Le Doussal, *Simple derivation of the* $(-\lambda H)^{5/2}$ *large deviation tail for the 1D KPZ equation*, J. Stat. Mech. 063210, (2018).

[38] A. Krajenbrink, P. Le Doussal, S. Prolhac, *Systematic time expansion for the Kardar-Parisi-Zhang equation, linear statistics of the GUE at the edge and trapped fermions.* Nuclear Physics B, **936** 239–305, (2018).

[39] I. Corwin, P. Ghosal, *Lower tail of the KPZ equation.* arXiv:1802.03273, (2018).

[40] I. Corwin, P. Ghosal, A Krajenbrink, P. Le Doussal, L-C Tsai, *Coulomb-Gas Electrostatics Controls Large Fluctuations of the Kardar-Parisi-Zhang*

Equation Phys. Rev. Lett. **121**, 060201, (2018).

[41] L.-C. Tsai. *Exact lower tail large deviations of the KPZ equation.* arXiv:1809.03410, (2018).

[42] A. Krajenbrink, P. Le Doussal, *Linear statistics and pushed Coulomb gas at the edge of the β-random matrices: four paths to large deviations.* Europhysics Letters **125** 20009. Supplementary materials available at arXiv:1811.00509, (2018)

[43] I. Corwin and P. Ghosal. *KPZ equation tails for general initial data.* arXiv:1810.07129, (2018).

[44] Pierre Le Doussal. *Large deviations for the Kardar–Parisi–Zhang equation from the Kadomtsev–Petviashvili equation.* Journal of Statistical Mechanics: Theory and Experiment, 2020(4):043201, 2020.

[45] Mattia Cafasso and Tom Claeys. *A Riemann-Hilbert approach to the lower tail of the KPZ equation.* arXiv:1910.02493, (2019).

[46] S. Prolhac, *Riemann surfaces for KPZ with periodic boundaries*, SciPost Phys. 8, 008 (2020).

[47] P. Le Doussal, S. N. Majumdar, A. Rosso, G. Schehr, *Exact short-time height distribution in 1D KPZ equation and edge fermions at high temperature*, Phys. Rev. Lett. **117**, 070403, (2016).

[48] A. Krajenbrink, P. Le Doussal, *Exact short-time height distribution in the one-dimensional Kardar-Parisi-Zhang equation with Brownian initial condition*, Phys. Rev. E **96**, 020102, (2017).

[49] A. Krajenbrink and P. Le. Doussal, *Large fluctuations of the KPZ equation in a half-space.* SciPost Phys. **5**, 032, (2018)

[50] Alexandre Krajenbrink. *Beyond the typical fluctuations: a journey to the large deviations in the Kardar-Parisi-Zhang growth model.* PhD thesis, PSL Research University, 2019.

[51] A. K. Hartmann, P. Le Doussal, S. N. Majumdar, A. Rosso, G. Schehr, *High-precision simulation of the height distribution for the KPZ equation*, *Europhys. Lett.* **121**, 67004 (2018).

[52] A. Hartmann, A. Krajenbrink, P. Le Doussal, *Probing the large deviations of the Kardar-Parisi-Zhang equation with an importance sampling of directed polymers in random media.* Phys. Rev. E **101** 012134, (2019).

[53] B. Meerson, E. Katzav, A. Vilenkin, *Large Deviations of Surface Height in the Kardar-Parisi-Zhang Equation*, Physical Review Letters **116**, 070601, (2016).

[54] A. Kamenev, B. Meerson, P. V. Sasorov, *Short-time height distribution in 1D KPZ equation: starting from a parabola*, Phys. Rev. E **94**, 032108, (2016).

[55] M. Janas, A. Kamenev, B. Meerson. *Dynamical phase transition in large-deviation statistics of the Kardar-Parisi-Zhang equation*, Phys. Rev. E **94**, 032133, (2016).

[56] B. Meerson, J. Schmidt, *Height distribution tails in the Kardar-Parisi-Zhang equation with Brownian initial conditions*, J. Stat. Mech. 103207, (2017).

[57] N. R. Smith, A. Kamenev, B. Meerson, *Landau theory of the short-time dynamical phase transition of the Kardar-Parisi-Zhang interface* , arXiv:1802.07497, (2018).

[58] N. R. Smith and B. Meerson, *Exact short-time height distribution for the flat Kardar-Parisi-Zhang interface* , arXiv:1803.04863, (2018).

[59] T. Asida, E. Livne, and B. Meerson. *Large fluctuations of a Kardar-Parisi-Zhang interface on a half-line: the height statistics at a shifted point.* Phys. Rev. E 99, 042132, arXiv:1901.07608, (2019).

[60] B. Meerson and A. Vilenkin. *Large fluctuations of a Kardar-Parisi-Zhang interface on a half line.* Physical Review E, 98(3):032145, (2018).

[61] N. R. Smith, B. Meerson, and P. Sasorov. *Finite-size effects in the short-time height distribution of the Kardar–Parisi–Zhang equation.* Journal of Statistical Mechanics: Theory and Experiment,2018(2):023202, (2018).

[62] N. R. Smith, B. Meerson, and A. Vilenkin. *Time-averaged height distribution of the Kardar-Parisi-Zhang interface.* arXiv:1902.08110, (2019).

[63] Yier Lin and Li-Cheng Tsai. Short time large deviations of the KPZ equation. arXiv:2009.10787, (2020).

[64] A. K. Hartmann, B. Meerson, and P. Sasorov. *Optimal paths of non-equilibrium stochastic fields: the Kardar-Parisi-Zhang interface as a test case.* arXiv:1907.05677, *Physical Review Research*, 1(3):032043, 2019.

[65] I. V. Kolokolov, S. E. Korshunov, *Explicit solution of the optimal fluctuation problem for an elastic string in random potential.* Phys. Rev. E **80**, 031107, (2009); *Universal and non-universal tails of distribution functions in the directed polymer and KPZ problems.* Phys. Rev. B **78**, 024206, (2008); *Optimal fluctuation approach to a directed polymer in a random medium.* Phys. Rev. B **75**, 140201, (2007).

[66] Shabat, A., and V. Zakharov. *Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media.* Soviet physics JETP 34.1 (1972).

[67] M. Kardar, G. Parisi and Y-C. Zhang, *Dynamic Scaling of Growing Interfaces*, Phys. Rev. Lett. **56**, 889, (1986).

[68] This is equivalent to use everywhere the following units of space, time and heights $y^* = (2\nu)^3/(D\lambda_0^2)$, $\tau^* = 2(2\nu)^5/(D^2\lambda_0^4)$ and $h^* = \frac{2\nu}{\lambda_0}$.

[69] so that H remains $\mathcal{O}(1)$ as $T \to 0$. Indeed from Ito one has $\overline{\exp(h(y,\tau))} = 1/\sqrt{4\pi\tau}$ for e.g. droplet IC, and more generally $\overline{\exp(h(y,\tau))} = \frac{1}{\sqrt{T}} \int_{\mathbb{R}} \frac{\mathrm{d}x'}{\sqrt{4\pi t}} e^{-(x-x')^2/(4t)} Q_0(x')$.

[70] If $\Psi(x,\tau)$ solves NLS then $Q = \Psi(x,\tau)|_{\tau\to -\mathrm{i}t}$ and $P = \Psi^*(x,\tau)|_{\tau\to -\mathrm{i}t}$ solves the $\{P,Q\}$ system.

[71] The equation for P should thus be solved backward in time.

[72] See Supplemental material.

[73] Ablowitz, M. J., Kaup, D. J., Newell, A. C., and Segur, H. *The inverse scattering transform-Fourier*

analysis for nonlinear problems. Studies in Applied Mathematics, 53(4), 249-315, (1974).

[74] The Wronskian $W = \phi_1\bar{\phi}_2 - \phi_2\bar{\phi}_1$ is space and time independent since $\partial_x W = \mathrm{Tr}(U_1)W = 0$ and $\partial_t W = \mathrm{Tr}(U_2)W = 0$. It is $W = -1$ at $x = -\infty$ and $W = -a(k)\tilde{a}(k) - b(k)\tilde{b}(k)$ from (10) at $x = +\infty$.

[75] Vladimir E Zakharov and Alexey Borisovich Shabat. *Integration of nonlinear equations of mathematical physics by the method of inverse scattering. II.* Functional Analysis and Its Applications, 13(3):166–174, 1979.

[76] A. Krajenbrink. *From Painlevé to Zakharov-Shabat and beyond: Fredholm determinants and integro-differential hierarchies*, Journal of Physics A: Mathematical and Theoretical, 54(3):035001, (2020).

[77] T. Bothner, *On Riemann-Hilbert methods in the analysis of Fredholm determinants.* In preparation.

[78] For the present application to WNT, these functions are real.

[79] Hopf, E., *Mathematical problems of radiative equilibrium*, Cambridge University Press, Cambridge (1934).

[80] VV Ivanov. *Resolvent method: exact solutions of half-space transport problems by elementary means.* Astronomy and Astrophysics, 286:328–337, (1994).

[81] F. Pollaczek, *Fonctions caracteristiques de certaines repartitions definies au moyen de la notion d'ordre - application a la theorie des attentes.* Comptes rendus hebdomadaires des seances de l'academie des sciences, 234(24):2334–2336, (1952).

[82] F. Spitzer, *A combinatorial lemma and its application to probability theory.* Transactions of the American Mathematical Society, 82(2):323–339, (1956). *The Wiener-Hopf equation whose kernel is a probability density.* Duke Mathematical Journal, 24(3):327–343, (1957).

[83] Satya N Majumdar. *Universal first-passage properties of discrete-time random walks and Lévy flights on a line: Statistics of the global maximum and records.* Physica A: Statistical Mechanics and its Applications, 389(20):4299–4316, 2010.

[84] Philippe Mounaix, Satya N Majumdar, and Grégory Schehr. *Asymptotics for the expected maximum of random walks and Lévy flights with a constant drift.* Journal of Statistical Mechanics: Theory and Experiment, 2018(8):083201, (2018).

[85] E. Sparre Andersen, *On the fluctuations of sums of random variables.* Mathematica Scandinavica, pages 263–285, (1953). *On the fluctuations of sums of random variables II.* Mathematica Scandinavica, pages 195–223, (1954).

[86] Note that in Ref. [38] the Sparre-Andersen theorem also appears in the short time large deviation expansion, using a different route (based on the exact droplet solution at all times).

[87] Note that all even conserved quantities $C_{2p} = 0$ by parity of the initial condition $Q_0(x) = \delta(x)$.

[88] From (5) one has $\Psi'(z) = \overline{e^H}$ and in particular $\Psi'(0) = e^{\hat{H}_0}$ where $\hat{H}_0$ is the most probable value of H, i.e. such that $\Phi'(\hat{H}_0) = 0$. Since one has, to $\mathcal{O}(z)$, $\Psi(z) \simeq z\overline{e^H}$, it implies from [69] that $\hat{H}_0 = \log \overline{e^H} = \log \int_{\mathbb{R}} \frac{\mathrm{d}x'}{\sqrt{4\pi}} e^{-(x')^2/(4)} Q_0(x')$ for a general initial condition, and $\hat{H}_0 = -\frac{1}{2}\log(4\pi)$ for the droplet IC.

[89] where we recall that $b(q)$ depends on z.

[90] See [50, Table 7.1], and also [38].

[91] Deriving it from the WNT requires to extend our method to a non-decaying $Q_0(x)$, a work in progress.

[92] Note that with our notations $H_{z=0}$ equals $\hat{H}_0$ and not H_0.

[93] P. C. Martin, E. D. Siggia, and H. A. Rose, *Statistical dynamics of classical systems.* Physical Review A, 8(1):423, (1973).
C de Dominicis. *Technics of field renormalization and dynamics of critical phenomena.* In J. Phys.(Paris), Colloq, pages C1–247, (1976).
H. K. Janssen, *On a lagrangean for classical field dynamics and renormalization group calculations of dynamical critical properties.* Zeitschrift für Physik B Condensed Matter, 23(4):377–380, (1976).

[94] Alex Kamenev. *Field theory of non-equilibrium systems.* Cambridge University Press, 2011.

[95] Täuber, Uwe C. *Critical dynamics: a field theory approach to equilibrium and non-equilibrium scaling behavior.* Cambridge University Press, 2014.

[96] Folkmar Bornemann. *On the numerical evaluation of Fredholm determinants.* Mathematics of Computation, 79(270):871–915, 2010.

.

Supplementary Material for

The inverse scattering of the Zakharov-Shabat system solves the weak noise theory of the Kardar-Parisi-Zhang equation

We give the principal details of the calculations described in the main text of the Letter. We also give additional information about the results displayed in the text.

S-A. Large deviation rate functions, analytic continuations, and $\{P, Q\}$ system

We review here the general properties of the large deviation rate functions $\Phi(H)$ and $\Psi(z)$, and the explicit results in the case of the droplet initial condition. These functions are defined in the text from the probability $P(H,t)$ of the (shifted) KPZ height $H = h(0,T) + H_0$, which takes the following leading asymptotics for $H = \mathcal{O}(1)$, $z = \mathcal{O}(1)$ and $T \ll 1$

$$P(H,T) \sim e^{-\frac{\Phi(H)}{\sqrt{T}}} \quad , \quad \int_{\mathbb{R}} \mathrm{d}H\, P(H,t) e^{-\frac{z}{\sqrt{T}} e^H} \sim e^{-\frac{\Psi(z)}{\sqrt{T}}} \tag{S22}$$

Inserting the first estimate into the integral in the second term, we see that the rate functions $\Psi(z)$ and $\Phi(H)$ are related through the Legendre transform

$$\Psi(z) = \min_{H\in\mathbb{R}} (\Phi(H) + z e^H) = (\Phi(H) + z e^H)|_{H=H_z} \quad , \quad H = H_z \quad \Leftrightarrow \quad \Phi'(H) = -z e^H \tag{S23}$$

Since it is expected that $\Phi(H)$ is convex, with $\Phi(\pm\infty) = \infty$, there is a unique solution to the minimization condition $\Phi'(H) = -z e^H$, denoted H_z^* in the text, and simply H_z here for convenience [92]. From $\Phi(H)$ one thus obtains $\Psi(z)$ as in (S23). However in the known solvable cases, it is $\Psi(z)$ which can be computed, and $\Phi(H)$ is obtained from it. Formally this is done as follows. Taking a derivative of (S23) one obtains $\Psi'(z) = e^{H_z}$, and one obtains $\Phi(H)$ in a parametric form

$$e^H = \Psi'(z) \quad , \quad \Phi(H) = \Psi(z) - z\Psi'(z) \tag{S24}$$

However, the definition (S22) of $\Psi(z)$ is obviously valid only for $\Re(z) \geqslant 0$, since otherwise the integral over H diverges. For $z \geqslant 0$, $\Phi'(H_z) = -z e^{H_z} \leqslant 0$, hence H_z varies in $(-\infty, \hat{H}_0]$, where $\hat{H}_0$ is the most probable value of H defined by $\Phi'(\hat{H}_0) = 0$ [92]. Thus $z \geqslant 0$ corresponds to the left side of $P(H,t)$. However, the function $\Psi(z)$ is the generating function of the cumulants of e^H, and can be analytically continued for $z < 0$. This allows to obtain $\Phi(H)$ for all H, including $H > \hat{H}_0$, using (S24). Let us now explain this point further on a specific example.

These functions are known for a very limited number of initial conditions, see [50, Table 7.1]. For the droplet initial condition, as found in [47] and obtained here by a completely different method, one has

$$\Psi(z) = \Psi_0(z) := -\frac{1}{\sqrt{4\pi}} \mathrm{Li}_{5/2}(-z) \tag{S25}$$

where we recall that $\mathrm{Li}_a(x) = \sum_{n\geqslant 1} \frac{1}{n^a} x^n$ is the polylogarithm function. The function $\Psi_0(z)$ is analytic in z except for a branch cut on the real axis for $z < -1$. The most probable value corresponds to $z = 0$ and from (S24) reads $e^{\hat{H}_0} = \Psi'(0) = 1/\sqrt{4\pi}$, hence $z \leqslant 0$ corresponds to $H = H_z \leqslant \hat{H}_0 = -\frac{1}{2}\log(4\pi)$. The function $\Psi(z)$ can be analytically continued to $z \in [-1, 0)$. Similarly to the logarithm function, on top of the principal value of the polylogarithm, analytic continuations can also be defined on higher Riemann sheets by adding a jump function. Here in order to obtain $\Phi(H)$ for $H \in (-\infty, +\infty)$ we need the two branches

$$\Psi(z) = \Psi_0(z) \quad , \quad \Psi(z) = \Psi_0(z) + \Delta(z) \quad , \quad \Delta(z) = \frac{4}{3}(\log(-\frac{1}{z}))^{3/2} \tag{S26}$$

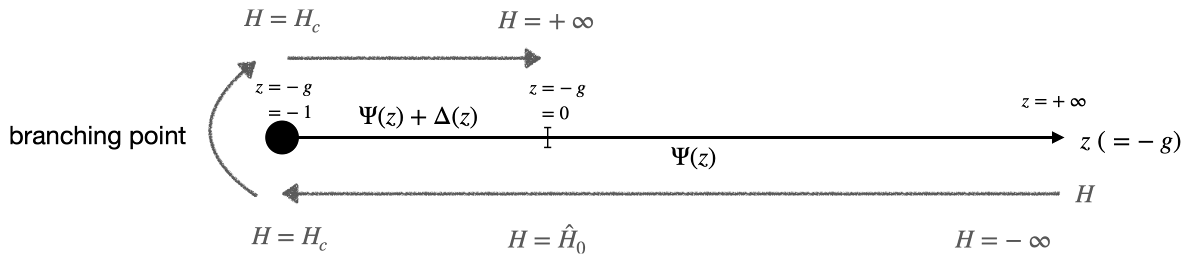

Figure S2. Schematic representation of the parametric solution of the optimization problem. For $H < H_c$ one uses the function Ψ to invert the Legendre transform taking the parameter z to decrease from $+\infty$ to -1. At $H = H_c$ or $z = -1$, one needs to turn around the branching point and replace Ψ by its continuation $\Psi + \Delta$ to determine all $H > H_c$ by increasing the parameter z from -1 to 0.

It is then possible to invert the Legendre transform (S23) in the form

$$\Phi(H) = \begin{cases} \max_{z\in[-1,+\infty[}(\Psi_0(z) - ze^H) & , \quad H < H_c \\ \min_{z\in[-1,0]}(\Psi_0(z) + \Delta(z) - ze^H) & , \quad H > H_c \end{cases} \tag{S27}$$

which leads to the same parametric determination (S24) upon inserting either $\Psi(z) = \Psi_0(z)$ (first line in (S27)) or $\Psi(z) = \Psi_0(z) + \Delta(z)$ (second line in (S27)). Here $H_c = \log \frac{\zeta(3/2)}{\sqrt{4\pi}}$ corresponds to the value $H = H_z$ where $z = -1$, i.e. $e^{H_c} = \Psi'(-1) = \Psi_0'(-1)$. Indeed $\Delta'(-1) = 0$ which ensures continuity at the turning point, and in fact all derivatives of $\Phi(H)$ at H_c are found to be continuous. We will refer to $\Psi_0(z)$ as the first branch, or main branch, and $\Psi_0(z) + \Delta(z)$ as the second branch.

This leads to the following asymptotic behaviors for $\Phi(H)$

$$\Phi(H) \simeq \begin{cases} c_-|H|^{5/2} & , \quad H \to -\infty \\ c_2(H - \hat{H}_0)^2 & , \quad |H - \hat{H}_0| \ll 1 \\ c_+ H^{3/2} & , \quad H \to +\infty \end{cases} \tag{S28}$$

where for the droplet initial condition we have [47, 54]

$$c_- = \frac{4}{15\pi}, \quad c_+ = \frac{4}{3}, \quad c_2 = \frac{1}{\sqrt{2\pi}} \tag{S29}$$

The asymptotics (S28) are in fact quite general, i.e. valid for other initial conditions with different values of the amplitudes $c_\pm, c_2$. For H around the most probable value $\hat{H}_0$, inserting the estimate (S28) into (S22) gives the typical fluctuations, which are Gaussian and obey the Edwards-Wilkinson scaling $H - \hat{H}_0 \sim T^{1/4}$. However in the tails the non trivial exponents 5/2 (left) and 3/2 emerge (right). To derive (S28) from (S27) one uses the asymptotics of the polylogs $\mathrm{Li}_n(-z) \underset{z\to+\infty}{\sim} -\frac{1}{\Gamma(1+n)}(\log z)^n + \mathcal{O}((\log z)^{n-2})$, see details in Refs. [47, 50].

The existence of H_c and of two branches in $\Psi(z)$ is quite general (not special to the droplet initial condition), as was found in Refs. [48–50]. It is useful for the following to indicate a general recipe to obtain the jump function $\Delta(z)$. Assume one can write the function $\Psi(z)$ to be continued as

$$\Psi(z) = \int_I \mathrm{d}y F(y) \frac{z}{y(y+z)}\,, \tag{S30}$$

where I is some interval. In that case one has [50]

$$\Delta(z) = 2\mathbf{i}\pi F(-z)\,. \tag{S31}$$

This reproduces the above result for the droplet initial condition. We will return to this procedure below in Section S-L. Let us now discuss this structure in terms of the $\{P, Q\}$ system (7). The $\{P, Q\}$ system was derived in the text, as the saddle point equations in the field variables $Z, \tilde{Z}$ of the dynamical action $S[\tilde{Z}, Z, j]$ in Eq. (6). More precisely it was shown that the large deviation function $\Psi(z)$ defined in (S22)

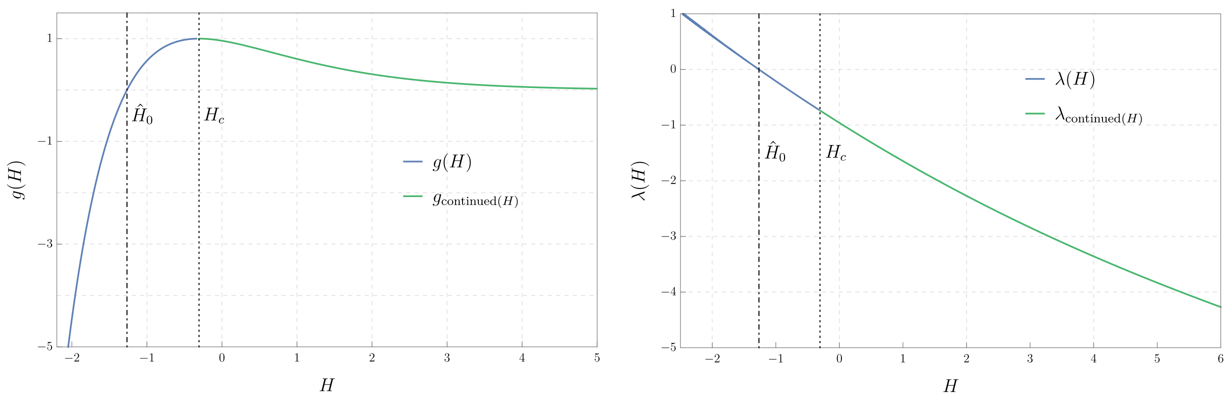

Figure S3. **Left.** Plot of the coupling g versus H, where g is the coupling of the $\{P, Q\}$ system (7) with boundary data (8) which should be used to obtain the large deviation rate function $\Phi(H)$ for a given H. For $0 < g \leqslant 1$ two values of H correspond to the same g, leading to the two branches of solutions discussed in the text. **Right.** Plot of the parameter λ versus H, where the $\{P, Q\}$ system (7) should now be solved with boundary data (S36). The function $\lambda(H)$ is monotonous. In this parametrization, the two branches of solutions in the left plot can be seen as a single branch. The notation *continued* is used to describe the second branch of the functions.

can be written from the saddle point values of the fields, as

$$\Psi(z) = S[\tilde{Z}, Z, j]|_{Z_{\rm sp}=Q, \tilde{Z}_{\rm sp}=gP, j=g\delta(x)\delta(t-1), g=-z} = \int_0^1 \int_{\mathbb{R}} \mathrm{d}t\mathrm{d}x [gP(\partial_t - \partial_x^2)Q - g^2 P^2 Q^2] - gQ(0,1)|_{g=-z} \tag{S32}$$

and where P, Q are the solutions of (7) with $g = -z$ and the boundary data (8). Using $(\partial_t - \partial_x^2)Q = 2gP^2Q^2$ this leads to

$$\Psi(z) = \Phi(H) + ze^H|_{H=H_z} = g^2 \int_0^1 \int_{\mathbb{R}} \mathrm{d}t\mathrm{d}x P^2 Q^2 - gQ(0,1)|_{g=-z} \tag{S33}$$

The middle expression is from (S23). The last term can be identified with ze^{H_z} with $Q(0,1) = Z_{\rm sp}(0,1) = e^H|_{H=H_z}$, i.e. the saddle point (i.e. "optimal") value of the solution of the SHE at $x = 0$ and $t = 1$, $Z(0,1) = e^H$, where there H denotes the fluctuating KPZ variable. From (S33) we see that we can thus identify the rate function $\Phi(H)$ as

$$\Phi(H)|_{H=H_z} = g^2 \int_0^1 \int_{\mathbb{R}} \mathrm{d}t\mathrm{d}x P^2 Q^2|_{g=-z} \tag{S34}$$

Hence, assuming that we know the solution P, Q of the $\{P, Q\}$ system (7) with the boundary data $P(x,1) = \delta(x)$ and $Q(x,0) = Q_0(x)$, upon setting $g = -z$, the above method allows to obtain $\Psi(z)$ and $\Phi(H)$. $\Psi(z)$ can be obtained e.g. from the value of $Q(0,1)$ as a function of z, since $\Psi'(z) = e^{H_z} = Q(0,1)$. Alternatively $\Phi(H_z)$ can be obtained computing the integral (S34). For the droplet initial condition, the first is done in the text, the second in Section S-K, in both cases using conserved charges. To obtain $\Phi(H)$ for $H > H_c$ one uses an analytical continuation in z, with two branches, as explained above in the case of the droplet initial condition. It is illustrated in Fig. S3 (left) where the relation g versus H is plotted, obtained by inverting the relation $H = \log(\Psi'(-g))$, with $\Psi(z)$ given in Eq. (S26). As we see on the figure, for $0 < g \leqslant 1$ there are two values of H corresponding to the same g. This corresponds to two distinct solutions of the $\{P, Q\}$ system in the attractive regime: they are different since the values of $Q(0,1) = e^H$ are different. As discussed in the text and below in Section S-L the second branch contains an additional solitonic part, and for $H \to +\infty$, although $g \to 0$ the solution remains non trivial.

To see this phenomenon in a different light, instead of using z as a parameter, it is possible to use H. More precisely, we can consider the ensemble where the parameter conjugated to H, called λ, is fixed (while z is the parameter conjugated to e^H). This amount to consider the generating function $G(\lambda)$ and its associated rate function $\psi(\lambda)$ defined as

$$G(\lambda) = \int_{\mathbb{R}} \mathrm{d}H P(H,T) e^{-\frac{\lambda}{\sqrt{T}}H} \simeq e^{-\frac{\psi(\lambda)}{\sqrt{T}}} \quad , \quad \psi(\lambda) = \min_{H\in\mathbb{R}} (\Phi(H) + \lambda H) \tag{S35}$$

The relation between λ and H is now $\Phi'(H) = -\lambda$. Since $\Phi(H)$ is convex, $\lambda(H)$ is now monotonous, and there is a unique $H \in (-\infty, \infty)$ for each $\lambda \in (-\infty, +\infty)$. In Fig. S3 (right) we have plotted $\lambda(H)$ for the droplet initial condition, obtained parametrically as $\lambda = ze^H = z\Psi'(z)$, $H = \log(\Psi'(z))$. The two branches for $\Psi(z)$ can now be seen as a single branch for $\lambda(H)$. As is detailed in the next section, working at fixed λ corresponds to solving the $\{P,Q\}$ system (7) with boundary data

$$Q(x,0) = Q_0(x) \quad , \quad gP(x,1)Q(x,1) = -\lambda\delta(x) \tag{S36}$$

Note that g can be absorbed in a redefinition $gP \to P$, hence it can be set to e.g. $g = -1$, so now the only parameter is λ, which varies on the whole real axis. In this parametrization (λ instead of $g = -z$), the solution of the $\{P,Q\}$ system with boundary data (S36) can be seen as a single branch. Although this parameterization looks conceptually simpler, it hides the fact that something non trivial happens at $H = H_c$, i.e. the emergence of a solitonic part in the solution, which can be seen as the emergence of a bound state in the inverse scattering method.

S-B. KPZ dynamical action and and $\{P,Q\}$ system in λ parametrization

In this section we derive the $\{P,Q\}$ system as the WNT on the dynamical action associated to the KPZ field h. It complements the short derivation given in the text using the dynamical action associated to the SHE field Z and, being very similar, it provides the necessary details for that derivation as well. In addition it shows that the $\{P,Q\}$ system allows to solve the large deviation problem for any value of H, as claimed in the text. The derivation is a variation on the one of [53, 55], but here we work in the fixed λ ensemble (see previous section), which allows to give a clear interpretation of the parameter Λ used there (analogous to our λ, see below).

Let us consider the KPZ equation (1) of the text, and rewrite it in the rescaled coordinates $x = y/\sqrt{T}$ and $t = \tau/T$, where we recall that T is a fixed observation time. It is convenient to define the shifted field $h(y,\tau) + H_0$, where H_0 is a constant (defined as $H_0 = \frac{1}{2}\log T$). We will slightly abuse notations and use the same name for the new function, i.e. $h(x,t) = h(y,\tau) + H_0$. It satisfies

$$\partial_t h = \partial_x^2 h + (\partial_x h)^2 + \sqrt{2} T^{1/4} \tilde{\eta}(x,t) \tag{S37}$$

where $\tilde{\eta}(x,t)$ is a standard space-time white noise. Let us denote $H = h(0,1)$. The generating function $\overline{\exp(-\frac{\lambda}{\sqrt{T}}H)}$, where $\overline{\ldots}$ denotes the average over the noise, of measure $\mathcal{D}\tilde{\eta}\exp(-1/2\iint_{xt}\tilde{\eta}(x,t)^2)$, can be written using the standard MSR method [93–95]. One expresses the equation of motion (S37) using delta functions, and their integral representations by multiplying (S37) by the response field $\frac{1}{\sqrt{T}}\tilde{h}(x,t)$, leading to

$$\begin{aligned} G(\lambda) &= \overline{\exp(-\frac{\lambda}{\sqrt{T}}H)} \\ &= \int \mathcal{D}\tilde{h}\mathcal{D}h\mathcal{D}\tilde{\eta} \exp\left(-\frac{1}{\sqrt{T}}[S_0[\tilde{h},h] + \lambda H - \int_0^{+\infty}\int_{\mathbb{R}} \mathrm{d}t\mathrm{d}x(\sqrt{2}T^{1/4}\tilde{h}(x,t)\tilde{\eta}(x,t) - \frac{\sqrt{T}}{2}\tilde{\eta}(x,t)^2)]\right) \\ &= \int \mathcal{D}\tilde{h}\mathcal{D}h \exp\left(-\frac{1}{\sqrt{T}}(S[\tilde{h},h] + \lambda H)\right) \end{aligned} \tag{S38}$$

where

$$S_0[\tilde{h},h] = \int_0^{+\infty}\int_{\mathbb{R}} \mathrm{d}t\mathrm{d}x \tilde{h}(\partial_t h - \partial_x^2 h - (\partial_x h)^2) \quad , \quad S[\tilde{h},h] = S_0[\tilde{h},h] - \int_0^{+\infty}\int_{\mathbb{R}} \mathrm{d}t\mathrm{d}x \tilde{h}^2 \tag{S39}$$

One has $G(0)=1$ automatically since the MSR path integral has unit normalisation as is checked using the proper time discretization (here we use a fixed initial condition $h(x,0)$).

For $T \ll 1$ this can be evaluated by the saddle point (SP) method, leading to the SP equations obtained by functional derivatives w.r.t. $\tilde{h}(x,t)$ and $h(x,t)$ of (S38) respectively

$$\partial_t h = \partial_x^2 h + (\partial_x h)^2 + 2\tilde{h} \tag{S40}$$

$$-\partial_t \tilde{h} = \partial_x^2 \tilde{h} - 2\partial_x(\tilde{h}\partial_x h) - \lambda\delta(x)\delta(t-1) \tag{S41}$$

and the saddle point solution obeys $\tilde{h}(x,t)=0$ for $t>1$. The functional derivative w.r.t. $h(x,t)$ of the term in (S38) containing λ produces a source term in the second equation. By integrating over time $t \in [1^-,1^+]$ in (S41) and using that $\tilde{h}(x,1^+)=0$, we see that it can be removed if we declare that the equations are to be studied on $t\in[0,1]$ and supplemented by the boundary condition at $t=1$

$$\tilde{h}(x,t=1) = -\lambda\delta(x) \tag{S42}$$

the boundary condition for $\tilde{h}$ being free at $t=0$. The boundary condition is given for h at $t=0$ and is free at $t=1$. The solution of the SP equation depends on a single parameter λ. Let us denote this solution $h_\lambda(x,t)$ and $\tilde{h}_\lambda(x,t)$, and insert them back into Eq. (S38). One obtains

$$G(\lambda) = \overline{\exp(-\frac{\lambda}{\sqrt{T}}H)} \sim \exp\left(-\frac{1}{\sqrt{T}}(\int_0^1\int_{\mathbb{R}} \mathrm{d}t\mathrm{d}x\tilde{h}_\lambda^2 + \lambda h_\lambda(0,1))\right) \tag{S43}$$

Hence, from (S35), this implies

$$\psi(\lambda) = \min_{H\in\mathbb{R}}(\Phi(H)+\lambda H) = \int_0^1\int_{\mathbb{R}} \mathrm{d}t\mathrm{d}x\tilde{h}_\lambda^2 + \lambda h_\lambda(0,1) \tag{S44}$$

Let us denote $H_\lambda = h_\lambda(0,1)$. Taking a derivative on each side and using that it is a saddle-point, we obtain that H_λ is also equal to $H_\lambda = \mathrm{argmin}(\Phi(H)+\lambda H)$, that is the solution of $\Phi'(H_\lambda)=-\lambda$, Hence, we also have the identity

$$\Phi(H_\lambda) = \int_0^1\int_{\mathbb{R}} \mathrm{d}t\mathrm{d}x\tilde{h}_\lambda^2 \tag{S45}$$

Equivalence with the $\{P,Q\}$ system. The saddle point equations (S40), (S41) can be equivalently written as the $\{P,Q\}$ system, see e.g. [55]. Let us define

$$Q(x,t) = e^{h(x,t)} \quad , \quad gP(x,t)Q(x,t) = \tilde{h}(x,t) \tag{S46}$$

Substituting into (S40), (S41) we obtain the P,Q system equation of the text (7). However, from (S42) and (S46), the boundary data are now given by (S36), and the solution depends on the single parameter λ (since g can be absorbed in P simultaneously in the equations (7) and in the boundary data (S36)). The $\{P,Q\}$ system is thus naturally obtained here in the λ parametrization. As discussed in the previous section, it is equivalent to the parametrization using $g=-z$. Indeed, one can identify $H_z = H_\lambda$, $\lambda = ze^{H_z}$, and $\Phi(H_\lambda)$ from (S45) with $\Phi(H_z)$ from (S34). The relation between λ and $g=-z$ however is not bijective and, for the droplet initial condition, can be read parametrically from Fig. S3, by eliminating H.

Remark. In the text we used Ito time discretization to define the SHE (3) as is customary as a proper definition of the KPZ equation via the Cole-Hopf mapping [32]. Going from $Z(x,t)$ to $h(x,t)=\log Z(x,t)$ adds a constant term in the KPZ equation (S37). For a smoothed noise correlator $\overline{\tilde{\eta}(x,t)\tilde{\eta}(x',t')} = R(x-x')\delta(t-t')$ this constant term is $R(0)$ (which diverges for white noise). We have ignored this term in the above calculation. Alternatively one could instead work using Stratonovich.

S-C. Comparison with the notations used in Meerson et al. works

It is useful to give a dictionary with the common notations used in the WNT works of Refs. [53–62]. In these works the WNT equations involve the KPZ height field $\mathsf{h}(x,t)$ and the noise field $\varrho(x,t)$. As here,

x, t are the rescaled space time variables, with $t \in [0,1]$ and T denotes the observation time. We denote h the KPZ field h used in these works, H its value at $x=0$, $t=1$ ($t=T$ in the unrescaled original time), and $\mathsf{S} = \frac{1}{2}\iint \mathrm{d}x\mathrm{d}t\varrho^2$ the saddle point action there. It is a function $\mathsf{S}(\mathsf{H})$ of the observed KPZ height. The correspondence with our notations goes as follows

$$h(x,t) = -\mathsf{h}(x,t)/2 \quad , \quad \tilde{h}(x,t) = -\varrho(x,t)/4 \quad , \quad H = -\mathsf{H}/2 \quad , \quad \Phi(H) = \frac{1}{8}\mathsf{S}(\mathsf{H}) \quad , \quad \lambda = \frac{1}{4}\Lambda \tag{S47}$$

The saddle point equation $\Phi'(H) = -\lambda$ is thus equivalent to $\mathsf{S}'(\mathsf{H}) = \Lambda$. Note that the definition of H in these works includes a shift (analogous to our H_0) which, however we have not attempted to pin down. The centering of $\Phi(H)$ around the most probable value allows to make the correspondence in each case.

S-D. General solutions of the $\{P, Q\}$ system and the solitonic (finite-rank) solutions

As mentioned in the text the general solution of the $\{P, Q\}$ system (7) are of the form [76, 77]

$$Q(x,t) = \langle\delta|\,\mathcal{A}_{xt}(I + g\mathcal{B}_{xt}\mathcal{A}_{xt})^{-1}\,|\delta\rangle \quad , \quad P(x,t) = \langle\delta|\,\mathcal{B}_{xt}(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1}\,|\delta\rangle \tag{S48}$$

in terms of two space-time dependent linear operators $\mathcal{A}_{xt}$, $\mathcal{B}_{xt}$ acting from $\mathbb{L}^2(\mathbb{R}^+)$ to $\mathbb{L}^2(\mathbb{R}^+)$, where $\mathcal{O}^{-1}$ denotes the inverse of the operator $\mathcal{O}$. The kernel of these operators have the form, with $v, v' \geqslant 0$

$$\mathcal{A}_{xt}(v,v') = A_t(x+v+v')\ ,\ \mathcal{B}_{xt}(v,v') = B_t(x+v+v') \tag{S49}$$

where the functions $A_t(x)$ and $B_t(x)$ are two solutions of the standard heat equation, i.e. forward $\partial_t A_t(x) = \partial_x^2 A_t(x)$, and backward $-\partial_t B_t(x) = \partial_x^2 B_t(x)$. These operators thus act on functions $f(v)$ as $(\mathcal{A}_{xt}f)(v) = \int_0^{+\infty}\mathrm{d}v' A_t(x+v+v')f(v') = \langle\delta_v|\mathcal{A}_{xt}|f\rangle$. We often use the bra-ket notation $f(v) = \langle\delta_v|\,f\rangle$, where $|\delta_w\rangle$ is the vector with component $\delta(v-w)$ so that $\langle\delta_v|\,\mathcal{O}\,|\delta_{v'}\rangle = \mathcal{O}(v,v')$, and in (S48) we denote $|\delta\rangle = |\delta_0\rangle$.

One can check, e.g. by expansion in g that the $\{P, Q\}$ system (7) is obeyed order by order: at lowest order g^0 the functions $P(x,t), Q(x,t)$ are simply identical to $B_t(x), A_t(x)$.

This operator construction requires that the functions $A_t(x), B_t(x)$ vanish sufficiently fast at $x \to +\infty$ so that all integrals over the variables v, v' converge. In the limit $x \to +\infty$ the operators $\mathcal{A}_{xt}$, $\mathcal{B}_{xt}$ thus become "small" and $P(x,t), Q(x,t)$ are thus asymptotically equivalent to $B_t(x), A_t(x)$. Hence Eq. (S48) can only describe solutions $P(x,t), Q(x,t)$ which decay at $x \to +\infty$ (there are no restrictions however on their behavior for $x \to -\infty$). Note that a mirror-image construction can also be performed exchanging the roles of $x > 0$ and $x < 0$.

The product operator $\mathcal{B}_{xt}\mathcal{A}_{xt}$ which occurs in (S48) thus reads in components $(\mathcal{B}_{xt}\mathcal{A}_{xt})(v,v') = \int_0^{+\infty}\mathrm{d}v'' B_t(x+v+v'')A_t(x+v''+v')$. Taking a derivative w.r.t. x we see that the integrand becomes a total derivative w.r.t. v'' which upon integration gives simply $-B_t(x+v)A_t(x+v')$, which is a projector. This leads to the important identities used in the text

$$\partial_x(\mathcal{B}_{xt}\mathcal{A}_{xt}) = -\mathcal{B}_{xt}\,|\delta\rangle\,\langle\delta|\,\mathcal{A}_{xt} \quad , \quad \partial_x(\mathcal{A}_{xt}\mathcal{B}_{xt}) = -\mathcal{A}_{xt}\,|\delta\rangle\,\langle\delta|\,\mathcal{B}_{xt}\,. \tag{S50}$$

It allows to express the product gPQ in terms of a Fredholm determinant. Indeed one has, using (S50)

$$\begin{aligned}
\partial_x \log \mathrm{Det}(I + g\mathcal{B}_{xt}\mathcal{A}_{xt})_{\mathbb{L}^2(\mathbb{R}_+)} &= \partial_x \mathrm{Tr}\log(I + g\mathcal{B}_{xt}\mathcal{A}_{xt}) \\
&= -g\mathrm{Tr}[(I + g\mathcal{B}_{xt}\mathcal{A}_{xt})^{-1}\mathcal{B}_{xt}\,|\delta\rangle\,\langle\delta|\,\mathcal{A}_{xt}] \\
&= -g\,\langle\delta|\,\mathcal{A}_{xt}(I + g\mathcal{B}_{xt}\mathcal{A}_{xt})^{-1}\mathcal{B}_{xt}\,|\delta\rangle \\
&= -g\,\langle\delta|\,(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1}\mathcal{A}_{xt}\mathcal{B}_{xt}\,|\delta\rangle \\
&= \langle\delta|\,((I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1} - I)\,|\delta\rangle
\end{aligned} \tag{S51}$$

Taking another derivative, using that

$$\partial_x(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1} = -g(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1}\partial_x(\mathcal{A}_{xt}\mathcal{B}_{xt})(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1} \tag{S52}$$

and using again (S50) we obtain

$$\begin{aligned}\partial_x^2 \log \mathrm{Det}(I + g\mathcal{B}_{xt}\mathcal{A}_{xt})_{\mathbb{L}^2(\mathbb{R}_+)} &= g \langle\delta| (I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1}\mathcal{A}_{xt} |\delta\rangle \langle\delta| \mathcal{B}_{xt}(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1} |\delta\rangle \\ &= gQ(x,t)P(x,t)\end{aligned} \tag{S53}$$

where we rearranged the operators to obtain the last equality. Note that, since the product $C_1 = \int_{\mathbb{R}} \mathrm{d}x gPQ$ is a conserved quantity, i.e. $\partial_t C_1 = 0$ as easily verified from (7), C_1 equals to minus the value reached by the last term in (S51) at $x = -\infty$ (since it vanishes for $x = +\infty$). As a consequence, this value is finite and time-independent.

In general the inversion of the operators in (S48) is a non-trivial task. However there exists solutions for which this inversion is possible explicitly, these are the (multi) solitonic solutions. They correspond to the case where at least one of the operators $\mathcal{A}_{xt}$ and $\mathcal{B}_{xt}$ is a finite-rank operator.

Rank-one. The simplest solution is when $\mathcal{A}_{xt}$ and $\mathcal{B}_{xt}$ are both rank-one operators. Since they are of the form (S49) and must obey the heat equation, they can be written as

$$\mathcal{A}_{xt}(v,v') = q_\kappa(x,t)e^{-\kappa(v+v')} \quad , \quad \mathcal{B}_{xt}(v,v') = p_\mu(x,t)e^{-\mu(v+v')} \tag{S54}$$

where we denote $q_\kappa = q_\kappa(x,t) = \tilde{q}e^{-\kappa x+\kappa^2 t}$ and $p_\mu = p_\mu(x,t) = \tilde{p}e^{-\mu x-\mu^2 t}$ two (evanescent) plane wave solutions of the heat equation, where $\tilde{q}, \tilde{p}$ are amplitudes. One can equivalently write $\mathcal{A}_{xt} = q_\kappa(x,t) |\kappa\rangle \langle\kappa|$ and $\mathcal{B}_{xt} = p_\mu(x,t) |\mu\rangle \langle\mu|$ were we denote $\langle\delta_v| \kappa\rangle = e^{-\kappa v}$, $\langle\delta_v| \mu\rangle = e^{-\mu v}$. In order for the operator product $\mathcal{A}_{xt}\mathcal{B}_{xt}$ to be convergent we must impose that $\Re(\kappa+\mu) > 0$. Here we will focus on the case where κ, μ are real and the product $g\tilde{p}\tilde{q} > 0$, where the solutions for P, Q can be defined for all times. It is possible to choose these parameter complex and obtain solutions on finite time intervals but we will not consider it here. The scalar product is then $\langle\kappa|\mu\rangle = \frac{1}{\kappa+\mu}$. Since $\mathcal{A}_{xt}\mathcal{B}_{xt} = \frac{q_\kappa p_\mu}{\kappa+\mu}|\kappa\rangle\langle\mu|$ is also rank-one, and similarly for $\mathcal{B}_{xt}\mathcal{A}_{xt}$ the operator inversion in (S48) is easy to perform using the Sherman Morrison formula $(I+|a\rangle\langle b|)^{-1} = I - \frac{|a\rangle\langle b|}{1+\langle b|a\rangle}$ and one obtains the simplest solution of the $\{P, Q\}$ system

$$Q(x,t) = \frac{\tilde{q}e^{-\kappa x+\kappa^2 t}}{1 + g\frac{\tilde{p}\tilde{q}}{(\kappa+\mu)^2}e^{-(\kappa+\mu)x+(\kappa^2-\mu^2)t}} \quad , \quad P(x,t) = \frac{\tilde{p}e^{-\mu x-\mu^2 t}}{1 + g\frac{\tilde{p}\tilde{q}}{(\kappa+\mu)^2}e^{-(\kappa+\mu)x+(\kappa^2-\mu^2)t}} \tag{S55}$$

For κ, μ real this corresponds to the standard one-soliton solution with

$$gPQ = \frac{(\kappa+\mu)^2}{4\cosh^2\left(\frac{1}{2}(\kappa+\mu)(x - x_0(t))\right)} \quad , \quad x_0(t) = (\kappa-\mu)t + y_0 \tag{S56}$$

with $y_0 = \frac{2}{\kappa+\mu}\log\frac{\sqrt{g\tilde{p}\tilde{q}}}{\kappa+\mu}$. Note that the matrix determinant lemma for a rank-one perturbation also gives $\mathrm{Det}(I+g\mathcal{B}_{xt}\mathcal{A}_{xt}) = \mathrm{Det}(I+g\frac{q_\kappa p_\mu}{\kappa+\mu}|\mu\rangle\langle\kappa|) = 1+g\frac{q_\kappa(x,t)p_\mu(x,t)}{(\kappa+\mu)^2}$. We see that (S56) then agrees with the general formula (S53). In particular, integrating over space, one finds that the conserved charge for the soliton is $C_1 = \int_{\mathbb{R}} \mathrm{d}x gPQ = \kappa+\mu$. Hence C_1 does not depend on g, which shows the non-perturbative nature of such solutions. The higher-order conserved quantities read

$$C_2 = \frac{1}{2}(\mu^2-\kappa^2), \quad C_3 = \frac{1}{3}(\mu^3+\kappa^3), \quad C_n = \frac{1}{n}(\mu^n - (-\kappa)^n). \tag{S57}$$

Note that soliton solutions of the $\{P, Q\}$ system with $\kappa+\mu < 0$ also exist, and can be deduced from the present ones by $x \to -x$, However in terms of operators, they require the mirror-image construction.

General finite-rank. Consider now the case where $\mathcal{A}_{xt}$ and $\mathcal{B}_{xt}$ are rank n_1 and n_2 operators, respectively, i.e. $\mathcal{A}_{xt} = \sum_{j=1}^{n_1} q_{\kappa_j}|\kappa_j\rangle\langle\kappa_j|$ and $\mathcal{B}_{xt} = \sum_{i=1}^{n_2} p_{\mu_i}|\mu_i\rangle\langle\mu_i|$ and $q_{\kappa_j} = q_{\kappa_j}(x,t) = \tilde{q}_j e^{-\kappa_j x+\kappa_j^2 t}$ and $p_{\mu_i} = p_{\mu_i}(x,t) = \tilde{p}_i e^{-\mu_i x-\mu_i^2 t}$ are plane waves. An elementary calculation, not reproduced here, using either the Woodbury matrix identity, or resummation of the expansion in g, gives

$$Q(x,t) = \sum_{i,j=1}^{n_1} q_{\kappa_i}(I + g\sigma\gamma)^{-1}_{ij} \quad , \quad P(x,t) = \sum_{i,j=1}^{n_2} p_{\mu_i}(I + g\sigma^T\gamma^T)^{-1}_{ij} \tag{S58}$$

$$\gamma_{ij} = \frac{p_{\mu_i} q_{\kappa_j}}{\mu_i + \kappa_j} \quad , \quad \sigma_{ij} = \frac{1}{\kappa_i + \mu_j} \quad , \quad q_{\kappa_j} = \tilde{q}_j e^{-\kappa_j x + \kappa_j^2 t} \quad , \quad p_{\mu_j} = \tilde{p}_j e^{-\mu_j x - \mu_j^2 t} \tag{S59}$$

For small ranks it is easy to check with Mathematica that these obey the $\{P, Q\}$ system. Using the matrix determinant lemma one also has

$$\mathrm{Det}(I + g\mathcal{B}_{xt}\mathcal{A}_{xt}) = \mathrm{Det}_{n_1 \times n_1}(I + g\sigma\gamma) = \mathrm{Det}_{n_2 \times n_2}(I + g\sigma^T\gamma^T) = \mathrm{Det}(I + g\mathcal{A}_{xt}\mathcal{B}_{xt}) \tag{S60}$$

and one can check explicitly again the relation (S53). Hence the determinant which appears in these finite-rank soliton solutions is, more explicitly

$$\mathrm{Det}_{1 \leqslant i,j \leqslant n_1}(\delta_{i,j} + g \sum_{k=1}^{n_2} \frac{1}{\kappa_i + \mu_k} \frac{p_{\mu_k}(x,t) q_{\kappa_j}(x,t)}{\mu_k + \kappa_j}) = \mathrm{Det}_{1 \leqslant i,j \leqslant n_2}(\delta_{i,j} + g \sum_{k=1}^{n_1} \frac{1}{\mu_i + \kappa_k} \frac{q_{\kappa_k}(x,t) p_{\mu_j}(x,t)}{\kappa_k + \mu_j}) \tag{S61}$$

One can check that the family of solutions pointed out in [55, Appendix B] using the Hirota method is equivalent to our particular family $n_1 = N - 1$, $n_2 = 1$ and considering $q_{\kappa_1}, \dots, q_{\kappa_{N-1}}, \frac{1}{p_\mu}$ that is $\kappa_j = c_j$, $j = 1, N-1$ and $c_N = -\mu$. So it is a rank-one family.

S-E. More on the structure of the $\{P, Q\}$ solutions

In complete generality, the operator structure (S48) as a solution of the $\{P, Q\}$ system (7) can be verified from

- either linear algebra, employing the following identities:

$$\begin{aligned} \partial_t \mathcal{A}_{xt} = \partial_x^2 \mathcal{A}_{xt} \quad &, \quad \partial_t \mathcal{B}_{xt} = -\partial_x^2 \mathcal{B}_{xt} \\ \partial_x(\mathcal{A}_{xt}\mathcal{B}_{xt}) &= -\mathcal{A}_{xt} \left|\delta\right\rangle \left\langle\delta\right| \mathcal{B}_{xt} \\ \partial_t(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1} &= -g(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1}\partial_t(\mathcal{A}_{xt}\mathcal{B}_{xt})(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1} \\ \partial_x(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1} &= -g(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1}\partial_x(\mathcal{A}_{xt}\mathcal{B}_{xt})(I + g\mathcal{A}_{xt}\mathcal{B}_{xt})^{-1} \end{aligned} \tag{S62}$$

  as well as

$$\begin{aligned} \partial_t(\mathcal{A}_{xt}\mathcal{B}_{xt}) &= \partial_x \left[(\partial_x \mathcal{A}_{xt})\mathcal{B}_{xt} - \mathcal{A}_{xt}\partial_x \mathcal{B}_{xt})\right] \\ &= -\partial_x \mathcal{A}_{xt} \left|\delta\right\rangle \left\langle\delta\right| \mathcal{B}_{xt} + \mathcal{A}_{xt} \left|\delta\right\rangle \left\langle\delta\right| \partial_x \mathcal{B}_{xt} \end{aligned} \tag{S63}$$

  and similarly with $\mathcal{A}_{xt}$, $\mathcal{B}_{xt}$ exchanged.

- or a mapping of the Lax pair (9) of the $\{P, Q\}$ system to a Riemann-Hilbert problem uniquely solvable by the means of a Fredholm determinant, see [76, 77] for recent works on this approach.

S-F. Scattering problem at $t = 1$

Here we show the relations obtained in the text for the scattering amplitudes for the WNT with a general initial condition. Although $Q(x, 0)$ is as yet unspecified, we can use the data at $t = 1$, $P(x, 1) = \delta(x)$ from Eq. (8).

Equation for $\bar{\phi}$ at $t = 1$. This equation allows to determine $\tilde{b}(k)$ and gives some relation involving $\tilde{a}(k)$. We call $\bar{\phi}_{1,2}(x,t)$ the two components of $\bar{\phi}$. Let us recall that at $x \to -\infty$, $\bar{\phi} \simeq (0, -e^{\mathbf{i}kx/2})^\intercal$. The first equation of the Lax pair given in the text, $\partial_x \vec{v} = U_1 \vec{v}$ with $\vec{v} = e^{-k^2 t/2} \bar{\phi}$ and U_1 given in (9) , reads in components at $t = 1$, using that $P(x, 1) = \delta(x)$ from (8),

$$\partial_x \bar{\phi}_1 = -\mathbf{i}\frac{k}{2}\bar{\phi}_1 - g\delta(x)\bar{\phi}_2 \quad , \quad \partial_x \bar{\phi}_2 = \mathbf{i}\frac{k}{2}\bar{\phi}_2 + Q(x,1)\bar{\phi}_1 \tag{S64}$$

Let us integrate the first equation. Since $\bar{\phi}_1$ vanishes at $-\infty$ it gives

$$\bar{\phi}_1(x, 1) = -g e^{-\mathbf{i}\frac{k}{2}x} \Theta(x) \bar{\phi}_2(0, 1) \tag{S65}$$

Taking the limit $x \to +\infty$, we thus obtain, from the asymptotics (10) that

$$\tilde{b}(k, t = 1) = -g\bar{\phi}_2(0, 1) \tag{S66}$$

To determine $\bar{\phi}_2(0, 1)$ we can integrate the second equation in (S64), which gives, using (S65) and (S66)

$$\begin{cases} e^{-\mathbf{i}\frac{k}{2}x}\bar{\phi}_2(x,1) = \bar{\phi}_2(0,1) + \tilde{b}(k,1)\int_0^x \mathrm{d}x' Q(x',1)e^{-\mathbf{i}kx'}, \quad x > 0 \\ \bar{\phi}_2(x,1) = -e^{\mathbf{i}\frac{k}{2}x}, \quad x < 0 \end{cases} \tag{S67}$$

where in the second equation we have used that $\bar{\phi}_2(x, 1) \simeq -e^{\mathbf{i}\frac{k}{2}x}$ for $x \to -\infty$. Assuming continuity of $\bar{\phi}_2(x, 1)$ at $x = 0$, this leads to $\bar{\phi}_2(0, 1) = -1$ and to

$$\tilde{b}(k, t = 1) = g \quad \Rightarrow \quad \tilde{b}(k) = ge^{-k^2} \tag{S68}$$

since we recall that $\tilde{b}(k, t) = \tilde{b}(k)e^{k^2 t}$, see text. Taking the $x \to +\infty$ limit of (S67) and using the asymptotics (10) we also obtain the relation

$$\tilde{a}(k, 1) = \tilde{a}(k) = 1 - g\int_0^{+\infty} \mathrm{d}x' Q(x', 1)e^{-\mathbf{i}kx'} \tag{S69}$$

Equation for ϕ at $t = 1$. This equation allows to obtain some relation involving $a(k)$ and $b(k)$. The first equation of the Lax pair, $\partial_x \vec{v} = U_1\vec{v}$ with $\vec{v} = e^{k^2 t/2}\phi$ as given in the text now reads, in components and at $t = 1$, using that $P(x, 1) = \delta(x)$ from (8),

$$\partial_x\phi_1 = -\mathbf{i}\frac{k}{2}\phi_1 - g\delta(x)\phi_2 \quad , \quad \partial_x\phi_2 = \mathbf{i}\frac{k}{2}\phi_2 + Q(x, 1)\phi_1 \tag{S70}$$

i.e. the same equations as (S64) but the boundary conditions are different. Let us recall that at $x \to -\infty$, $\phi \simeq (e^{-\mathbf{i}kx/2}, 0)^\intercal$. The equations (S70) can be rewritten as

$$[e^{\mathbf{i}\frac{k}{2}x}\phi_1(x,1)]' = -g\delta(x)\phi_2(x,1)e^{\mathbf{i}\frac{k}{2}x}, \qquad [e^{-\mathbf{i}\frac{k}{2}x}\phi_2(x,1)]' = Q(x,1)\phi_1(x,1)e^{-\mathbf{i}\frac{k}{2}x} \tag{S71}$$

Integrating these two equations, and using the asymptotics (10) at $x \to +\infty$ we obtain

$$\begin{aligned} &\phi_1(x,1) = e^{-\mathbf{i}\frac{k}{2}x}(\Theta(-x) + a(k)\Theta(x)), \quad a(k) - 1 = -g\phi_2(0,1) \\ &\phi_2(x,1) = e^{\mathbf{i}\frac{k}{2}x}\int_{-\infty}^x \mathrm{d}x' Q(x',1)e^{-\mathbf{i}kx'}(\Theta(-x') + a(k)\Theta(x')) \end{aligned} \tag{S72}$$

where we used that $a(k, t) = a(k)$, see the main text. Setting $x = 0$ in the second equation we obtain the relation

$$\phi_2(0,1) = \int_{-\infty}^0 \mathrm{d}x' Q(x',1)e^{-\mathbf{i}kx'}, \quad a(k) = 1 - g\int_{-\infty}^0 \mathrm{d}x' Q(x',1)e^{-\mathbf{i}kx'} \tag{S73}$$

Note that integrating the second equation in (S71) for $\phi_2(x, 1)e^{-\mathbf{i}\frac{k}{2}x}$ between 0 and $+\infty$ and using the asymptotics (10) leads to an expression for $b(k)$, however this expression is equivalent to the one obtained from the relation $a(k)\tilde{a}(k) + b(k)\tilde{b}(k) = 1$ obtained from the Wronskian (see text) together with the above results for $\tilde{b}(k), \tilde{a}(k), a(k)$.

As noted in the text, at this stage $Q(x, 1)$ is unknown. If the initial condition $Q(x, 0)$ is even in x, then so is $Q(x, 1)$ and (S69) and (S73) imply that $\tilde{a}(k) = a(-k) = a^*(k^*)$. From the Wronskian relation and (S68) one thus gets $b(k)ge^{-k^2} = 1 - a(k)a(-k) = 1 - |a(k)|^2$, hence $b(k)$ is real and even in k. Alternatively one sees that $|a(k)|$ is fixed by $b(k)$ so one can write

$$a(k) = e^{-\mathbf{i}\varphi(k)}\sqrt{1 - gb(k)e^{-k^2}} \tag{S74}$$

where $\varphi(k)$ is a real and odd function $\varphi(k) = -\varphi(-k)$.

S-G. Scattering problem: general t

Here we obtain some relations from the scattering problem at any t always valid for the $\{P, Q\}$ system (for arbitrary boundary conditions, i.e. beyond the WNT). In the case of the WNT with the boundary conditions (8), when specified to $t = 1$ they lead to the same results as in the previous section. When specified to $t = 0$ they give some (useful but not too explicit) formula for $b(k)$ for general initial condition. For the droplet initial condition they give $b(k) = 1$.

Let us return to the ∂_x equation of the Lax pair for $\phi(x, t)$, which we write at a fixed t, in the form (here we also indicate the dependence in k)

$$\partial_x[e^{\mathbf{i}\frac{k}{2}x}\phi_1(x,t,k)] = -gP(x,t)\phi_2(x,t,k)e^{\mathbf{i}\frac{k}{2}x}, \qquad \partial_x[e^{-\mathbf{i}\frac{k}{2}x}\phi_2(x,t,k)] = Q(x,t)\phi_1(x,t,k)e^{-\mathbf{i}\frac{k}{2}x} \tag{S75}$$

Integrating the first equation of (S75) from $x = -\infty$ to $x = +\infty$, and using that $a(k, t) = a(k)$ we have

$$a(k) - 1 = -g\int_{-\infty}^{+\infty} \mathrm{d}x\, P(x,t)\phi_2(x,t,k)e^{\mathbf{i}\frac{k}{2}x} \tag{S76}$$

and from $-\infty$ and a value x

$$e^{\mathbf{i}\frac{k}{2}x}\phi_1(x,t,k) = 1 - g\int_{-\infty}^{x} \mathrm{d}x'\, P(x',t)\phi_2(x',t,k)e^{\mathbf{i}\frac{k}{2}x'} \tag{S77}$$

Integrating the second equation of (S75) between $-\infty$ and $+\infty$, we have

$$b(k)e^{-k^2t} = \int_{-\infty}^{+\infty} \mathrm{d}xQ(x,t)\phi_1(x,t,k)e^{-\mathbf{i}\frac{k}{2}x} \tag{S78}$$

and from $-\infty$ and a value x

$$e^{-\mathbf{i}\frac{k}{2}x}\phi_2(x,t,k) = \int_{-\infty}^{x} \mathrm{d}x'\, Q(x',t)\phi_1(x',t,k)e^{-\mathbf{i}\frac{k}{2}x'} \tag{S79}$$

Inserting, this gives two integral equations for $\phi_{1,2}$

$$\begin{aligned} e^{\mathbf{i}\frac{k}{2}x}\phi_1(x,t,k) &= 1 - g\int_{-\infty}^{x}\mathrm{d}x'\int_{-\infty}^{x'}\mathrm{d}x''\, P(x',t)Q(x'',t)\phi_1(x'',t,k)e^{\mathbf{i}kx'-\mathbf{i}\frac{k}{2}x''} \\ e^{-\mathbf{i}\frac{k}{2}x}\phi_2(x,t,k) &= \int_{-\infty}^{x}\mathrm{d}x'\, e^{-\mathbf{i}kx'}Q(x',t)\left(1 - g\int_{-\infty}^{x'}\mathrm{d}x''\, P(x'',t)\phi_2(x'',t,k)e^{\mathbf{i}\frac{k}{2}x''}\right) \end{aligned} \tag{S80}$$

Iteration of these equations gives a series representation for $b(k)$ as a sum of alternating products of terms $Q(QP)^{n-1}$, $n \geqslant 1$, integrated over ordered sectors as

$$\begin{aligned} & b(k)e^{-k^2t} = \\ & \sum_{n=0}^{\infty}(-g)^n\int_{x_{2n+1}<x_{2n}<\dots<x_1}\prod_{j=1}^{2n+1}\mathrm{d}x_j e^{\mathbf{i}k(\sum_{j=1}^{n}x_{2j}-\sum_{j=0}^{n}x_{2j+1})}Q(x_1,t)P(x_2,t)\dots P(x_{2n},t)Q(x_{2n+1},t) \end{aligned} \tag{S81}$$

as well as a a series representation for $a(k)$ as a sum of alternating products of terms $(QP)^n$, $n \geqslant 0$, integrated over ordered sectors as

$$a(k) = \sum_{n=0}^{\infty}(-g)^n\int_{x_{2n}<x_{2n}<\dots<x_1}\prod_{j=1}^{2n}\mathrm{d}x_j e^{\mathbf{i}k\sum_{j=1}^{n}(x_{2j+1}-x_{2j})}P(x_1,t)Q(x_2,t)\dots P(x_{2n},t)Q(x_{2n},t) \tag{S82}$$

These relations are valid for any t, and any boundary condition for the $\{P, Q\}$ system (beyond its application to WNT).

Let us apply them to the WNT with the boundary conditions (8). Setting $t = 1$ in (S82) we see that $P(x, 1) = \delta(x)$ implies that only the first two terms, $n = 0$ and $n = 1$, survive in the series. Indeed for $n \geqslant 2$ the δ function implies that in the integral all odd $x_{2n+1} = 0$ and the integration over the even x_{2n} will be restricted to a vanishing small interval, leading to a vanishing result since $Q(x,1)$ is a smooth function. Hence, from $P(x,1)=\delta(x)$ we obtain

$$a(k) = 1 - g\int_{x_2<x_1} \mathrm{d}x_1\mathrm{d}x_2 P(x_1,1)Q(x_2,1)e^{\mathbf{i}k(x_1-x_2)} = 1 - g\int_{x_2<0}\mathrm{d}x_2 Q(x_2,1)e^{-\mathbf{i}kx_2} \tag{S83}$$

which is precisely (S73). For the droplet initial condition $Q(x,0)=\delta(x)$, setting $t=0$ in (S81), by the same argument we see that only the first term $n=0$ remains, leading to

$$b(k) = \int_{\mathbb{R}} \mathrm{d}x_1 e^{-\mathbf{i}kx_1}Q(x_1,0) = 1 \tag{S84}$$

Of course this result can be also obtained as in the previous section considering ϕ at $t=0$. Although we will not present it here, by considering the equation for $\bar{\phi}$ one arrives at similar equations, and similar series expansions for $\tilde{a}(k)$ (of the form $(QP)^{n-1}$) and $\tilde{b}(k)$ (of the form $P(QP)^{n-1}$). Specified at $t=1$ they reproduce the results of the previous section, i.e. $\tilde{b}(k)=ge^{-k^2}$ and (S69) for $\tilde{a}(k)$.

Expansion near the droplet IC. Although complicated looking, formula (S81) allows for an expansion around the droplet initial condition. Let us define, as in the text, the Fourier transform $\hat{b}(x)=\int_{\mathbb{R}}\frac{\mathrm{d}k}{2\pi}b(k)e^{ikx}$. Consider (S81) at $t=0$. It implies

$$\hat{b}(x) = Q_0(x) + \int_{x_3<x_2<x_1}\delta(x+x_2-x_1-x_3)Q_0(x_1)P(x_2,0)Q_0(x_3) + \dots \tag{S85}$$

Suppose now that $Q_0(x)=\delta(x)+\epsilon f(x)$ where ϵ is small. To first order in ϵ one finds that all terms with more integrations than in the second term vanish, hence we have formally

$$\hat{b}(x) = Q_0(x) + \epsilon\int \mathrm{d}x_2 \left(\Theta(-x)\Theta(-x_2)+\Theta(x)\Theta(x_2)\right) f(x+x_2)P^{\mathrm{drop}}(x_2,0) + \mathcal{O}(\epsilon^2) \tag{S86}$$

where the superscript "drop" denotes the value for the droplet initial condition. Note that $P^{\mathrm{drop}}(x_2,0) = Q^{\mathrm{drop}}(x_2,1) = A_1(|x_2|)$, hence to this order we obtain an explicit formula in terms of known quantities. If we choose $f(x)$ even we further obtain the (even) result

$$\begin{aligned}\hat{b}(x) &= \delta(x) + \epsilon\left(f(x) + \int_0^{+\infty}\mathrm{d}y f(x+y)A_1^{\mathrm{drop}}(y)\right) + \mathcal{O}(\epsilon^2)\\ &= \delta(x) + \epsilon\left(f(x) + \int_{-\infty}^{0}\mathrm{d}z A_0^{\mathrm{drop}}(z)\tilde{f}(z;x)\right) + \mathcal{O}(\epsilon^2)\end{aligned} \tag{S87}$$

where $\tilde{f}(z;x)=\int_0^{+\infty}\mathrm{d}y f(x+y)p(z-y)$, $p(z)=\frac{1}{\sqrt{4\pi}}e^{-z^2}$ and $A_t^{\mathrm{drop}}(x)$ is given in (S123).

S-H. Derivation of the non-linear integral equation (15) for A, B

Here we give the details of the derivation of (15) in the text. Let us set $t=1$ and ignore the time index. Differentiating the vector $|\beta_x\rangle := (I+g\mathcal{A}_x\mathcal{B}_x)^{-1}|\delta\rangle$, with respect to x and using the identity $\partial_x(\mathcal{A}_x\mathcal{B}_x) = -\mathcal{A}_x|\delta\rangle\langle\delta|\mathcal{B}_x$ we obtain

$$\partial_x|\beta_x\rangle = gP(x,t=1)(I+g\mathcal{A}_x\mathcal{B}_x)^{-1}\mathcal{A}_x|\delta\rangle \tag{S88}$$

Since $P(x,t=1)=\delta(x)$, integrating one finds

$$|\beta_x\rangle = |\delta\rangle - g\Theta(-x)(I+g\mathcal{A}_0\mathcal{B}_0)^{-1}\mathcal{A}_0|\delta\rangle \tag{S89}$$

Hence, for $x>0$, $|\beta_x\rangle=|\delta\rangle$ which, inserted in the definition of $|\beta_x\rangle$, i.e. $(I+g\mathcal{A}_x\mathcal{B}_x)|\beta_x\rangle=|\delta\rangle$, leads to $g\mathcal{A}_x\mathcal{B}_x|\delta\rangle=0$ for $x>0$. Taking another derivative w.r.t x one obtains $A_1(x')B_1(x)=0$ for $x'>x>0$, which implies that $B_1(x>0)=0$, hence that the operator $\mathcal{B}_{x1}$ is zero for $x>0$. This in turns yields $Q(x>0,1)=A_1(x)$, which for Q even implies the interesting result $Q(x,1)=A_1(|x|)$. Next, inserting (S89) for general x into the above definition of $|\beta_x\rangle$, taking a derivative w.r.t x, and finally using the fact that

$\mathcal{A}_{x1}\mathcal{B}_{x1}\mathcal{A}_{x1}\left|\delta\right\rangle = 0$ as a consequence of $B_1(x>0)=0$, we obtain $B_1(x) = \delta(x) + g\Theta(-x)\left\langle\delta\right|\mathcal{B}_{x,1}\mathcal{A}_{0,1}\left|\delta\right\rangle$. This is the sought for integral equation, written in the more explicit form (15) in the text

S-I. From a non-linear integral equation to the Hopf-Ivanov linear equation

Consider the non-linear integral equation (15) in the text. For the droplet initial condition one has $A_0(x) = B_1(x)$. This equation then becomes a closed equation for the function $B_1(x)$

$$B_1(x) = \delta(x) + g\Theta(-x)\int_0^{+\infty} \mathrm{d}v B_1(x+v)\int_{-\infty}^{+\infty} \mathrm{d}y p(v-y)B_1(y) \tag{S90}$$

Here we have defined $p(z) := \frac{e^{-z^2/4}}{\sqrt{4\pi}}$, although the present considerations extend to any continuous symmetric function $p(z) = p(-z)$ normalized to unity. Now consider the linear Hopf-Ivanov (HI) equation (16), which reads

$$B_1(x) = \delta(x) + g\Theta(-x)\int_{-\infty}^{0} \mathrm{d}y p(x-y)B_1(y) \tag{S91}$$

Let us separate, in both cases, the delta function part and write

$$B_1(x) = \delta(x) + \Theta(-x)\tilde{B}_1(x) \tag{S92}$$

where $\tilde{B}_1(x)$ is smooth. It is defined only for $x \leqslant 0$, hence all equations below are understood for $x \in \mathbb{R}^-$. We then obtain from (S90) the non-linear integral equation to be solved for $\tilde{B}_1(x)$

$$\tilde{B}_1(x) = gp(x) + g\int_{-\infty}^{0} \mathrm{d}y p(x+y)\tilde{B}_1(y) + g\int_0^{-x} \mathrm{d}v p(v)\tilde{B}_1(x+v) + g\int_0^{-x} \mathrm{d}v \tilde{B}_1(x+v)\int_{-\infty}^{0} dy p(v-y)\tilde{B}_1(y) \tag{S93}$$

while Hopf-Ivanov's equation (S91) leads instead to

$$\tilde{B}_1(x) = gp(x) + g\int_{-\infty}^{0} \mathrm{d}y p(x-y)\tilde{B}_1(y) \tag{S94}$$

Note that we will use the symmetry $p(-z) = p(z)$ in several places. To show that the solutions of (S94) are also solutions of (S93) we rewrite these two equations as recursion relations by expanding $\tilde{B}_1(x)$ in series of g

$$\tilde{B}_1(x) = \sum_{n=1}^{+\infty} g^n B_{1,n}(x) \tag{S95}$$

We already note that for both equations

$$B_{1,1}(x) = p(x) \tag{S96}$$

The recursion relation for the $B_{1,n}$ from (S93) is, for $n \geqslant 2$

$$B_{1,n}(x) = \int_{-\infty}^{0} \mathrm{d}y p(x+y)B_{1,n-1}(y) + \int_x^0 \mathrm{d}y p(x-y)B_{1,n-1}(y) \tag{S97}$$

$$+\sum_{m=1}^{n-2}\int_x^0 dy_1 \int_{-\infty}^{0} dy_2 p(y_1-y_2-x)B_{1,m}(y_1)B_{1,n-1-m}(y_2) \tag{S98}$$

where we recall $x \leqslant 0$, while Hopf-Ivanov's recursion is

$$B_{1,n}(x) = \int_{-\infty}^{0} \mathrm{d}y p(x-y)B_{1,n-1}(y) \tag{S99}$$

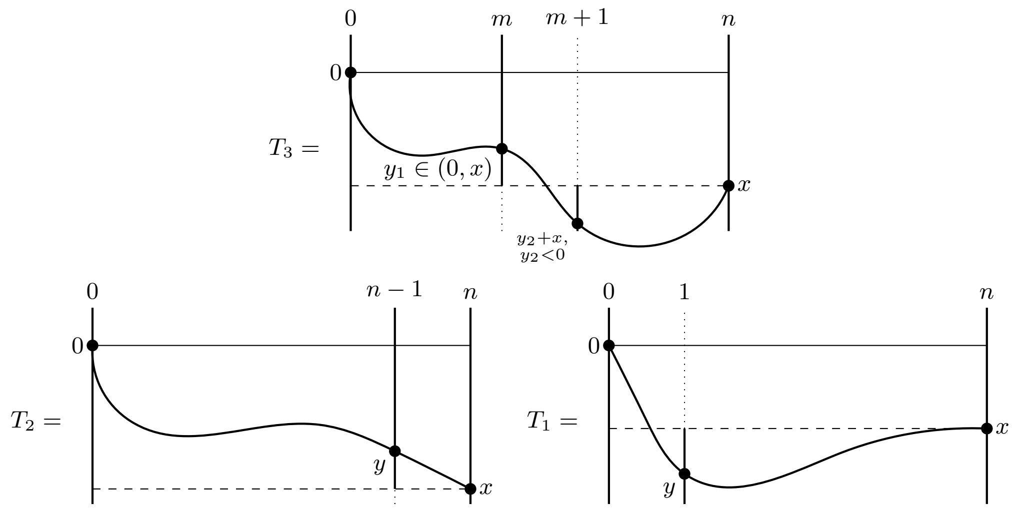

Figure S4. Illustration of the three terms T_1, T_2, T_3 which enter into the recurrence (S101) for $B_{1,n}(x) = T_1 + T_2 + T_3$. We recall that $B_{1,n}(x)$ is the probability that a random walk of n steps starting at 0 arrives at $x < 0$ and has all points negative. One defines $m = 0, \dots, n-1$ the step such that the level x is crossed *for the last time* between m and $m+1$. Then the term T_3 is the generic term, the term T_2 corresponds to $m = n-1$ and the term T_1 corresponds to $m = 0$.

We can already check the agreement of the two recursions to order 2, indeed, from (S97) we have

$$B_{1,2}(x) = \int_{-\infty}^{0} \mathrm{d}y p(x+y)p(y) + \int_{x}^{0} \mathrm{d}y p(x-y)p(y) = \int_{-\infty}^{-x} \mathrm{d}y p(x+y)p(y) = \int_{-\infty}^{0} \mathrm{d}y p(y)p(y-x) \quad \text{(S100)}$$

which is exactly the result for $B_{1,2}(x)$ from (S99), using (S96). Note that this agreement is independent on precise shape of $p(z)$, provided $p(-z) = p(z)$.

To show the coincidence to all orders, it is convenient to recall the random walk interpretation of the Hopf-Ivanov equation, as mentionned in the text. Consider $X(j) \in \mathbb{R}$ a discrete time random walk, $X(j+1) = X(j) + z_j$, with z_j i.i.d. with a symmetric jump probability $p(z)$. The recursion relation (S99) simply expresses that $B_{1,n}(x)$ is the probability that the walk starting at $X(0) = 0$ arrives at $X(n) = x$ in $n \geq 1$ steps, while remaining negative, $\{X(j) \leqslant 0\}_{j=0,\dots,n}$. Let us now rewrite the recurrence (S97) as a sum of three terms, $B_{1,n}(x) = T_1 + T_2 + T_3$, with

$$T_1 = \int_{-\infty}^{0} \mathrm{d}y p(x+y) B_{1,n-1}(y) = \int_{-\infty}^{x} \mathrm{d}y p(y) B_{1,n-1}(y-x) \quad \text{(S101)}$$

$$T_2 = \int_{x}^{0} \mathrm{d}y p(x-y) B_{1,n-1}(y) \quad \text{(S102)}$$

$$T_3 = \sum_{m=1}^{n-2} \int_{x}^{0} \mathrm{d}y_1 \int_{-\infty}^{0} \mathrm{d}y_2 p(y_1 - y_2 - x) B_{1,m}(y_1) B_{1,n-1-m}(y_2) \quad \text{(S103)}$$

This recursion and the Hopf-Ivanov recursion (S99) become identical if one notices that m in T_3 is actually the step just before the last time that the level x is crossed by the walk, before reaching x at its end. Indeed then one can decompose the walk in three parts, (i) before and up to m, (ii) the single step between m and $m+1$ which must then necessarily cross level x, hence it must be above x at m and below x at $m+1$. And (iii) the third part is then by reflection exactly the same as going from zero (the endpoints being shifted by x) to $y - x$ with $y < x$). This is further explained in Fig. S4.

S-J. Conserved charges

The generating function of the conserved charges is $\Gamma(x,t,k)$. It is defined from a particular solution of the Lax pair $\partial_x \vec{v} = U_1\vec{v}$, $\partial_t \vec{v} = U_2\vec{v}$ with U_1, U_2 given in (9). More precisely it is defined as $\Gamma = \frac{v_2}{v_1} = \frac{\phi_2}{\phi_1}$, where $\vec{v} = (v_1, v_2)^\intercal$ is the solution of the Lax pair of the form $\vec{v} = e^{k^2t/2}\phi$ where $\phi = (\phi_1, \phi_2)^\intercal$ is as defined in the text, i.e. such that at $x \to -\infty$, $\phi \simeq (e^{-\mathbf{i}kx/2}, 0)^\intercal$ and that it behaves as (10) for $x \to +\infty$. The ∂_x equation gives

$$\partial_x v_1 = -\mathbf{i}\frac{k}{2}v_1 - gPv_2 \quad , \quad \partial_x v_2 = \mathbf{i}\frac{k}{2}v_2 + Qv_1 \tag{S104}$$

which implies that Γ obeys a Ricatti equation

$$\partial_x \Gamma = \mathbf{i}k\Gamma + Q + gP\Gamma^2 \tag{S105}$$

From (9) using $\vec{v} = e^{k^2t/2}\phi$ we see that $\log\phi_1$ satisfies

$$\partial_x \log\phi_1 = -\mathbf{i}\frac{k}{2} - gP\Gamma \quad , \quad \partial_t \log\phi_1 = \mathsf{A} - \frac{k^2}{2} + \mathsf{B}\Gamma \tag{S106}$$

where the second equation comes from the first component of the ∂_t equation of the Lax pair, i.e. $\partial_t(e^{k^2t/2}\phi) = U_2 e^{k^2t/2}\phi$ where U_2 is given in (9) and $\mathsf{A} = k^2/2 - gPQ$, $\mathsf{B} = g(\partial_x - \mathbf{i}k)P$. Taking the cross derivatives in (S106) we obtain the conservation equation

$$\partial_t(-gP\Gamma) = \partial_x(\mathsf{A} - \frac{k^2}{2} + \mathsf{B}\Gamma) \tag{S107}$$

To obtain the conserved charges and the associated currents one expands Γ in a Laurent series $\Gamma = \sum_{n\geqslant 1}\frac{\Gamma_n(x,t)}{(\mathbf{i}k)^n}$ as mentioned in the text. The conserved quantities C_n are then

$$C_n = \int_{\mathbb{R}} \mathrm{d}x\, \tilde{C}_n \quad , \quad \tilde{C}_n = -gP\Gamma_n \quad , \quad \partial_t\tilde{C}_n = \partial_x J_n \tag{S108}$$

where the associated currents are obtained from $\mathsf{A} - \frac{k^2}{2} + \mathsf{B}\Gamma = \sum_n \frac{J_n(x,t)}{(\mathbf{i}k)^n}$, leading to

$$J_n = g(\partial_x P)\Gamma_n - gP\Gamma_{n+1} \tag{S109}$$

Now the Γ_n can be obtained recursively as follows. From the Riccatti equation (S105) one obtains the following recurrence for $n \geqslant 1$

$$\Gamma_{n+1} = \partial_x\Gamma_n - gP\sum_{p=1}^{n-1}\Gamma_p\Gamma_{n-p} \quad , \quad \Gamma_1 = -Q \tag{S110}$$

Explicit calculation then gives the conserved quantities

$$\begin{aligned} &\Gamma_1 = -Q \quad , \quad C_1 = g\int_{\mathbb{R}}\mathrm{d}xPQ \quad , \quad \Gamma_2 = -\partial_x Q \quad , \quad C_2 = g\int_{\mathbb{R}}\mathrm{d}xP\partial_x Q \\ &C_3 = \int_{\mathbb{R}}\mathrm{d}x(gP\partial_x^2 Q + g^2P^2Q^2) \quad , \quad C_4 = \int_{\mathbb{R}}\mathrm{d}xgP\left(gQ^2\partial_x P + 4gPQ\partial_x Q + \partial_x^3 Q\right) \end{aligned} \tag{S111}$$

and the associated currents

$$\begin{aligned} J_1 &= g(P\partial_x Q - \partial_x PQ) \quad , \quad J_2 = g(P\partial_x^2 Q - \partial_x P\partial_x Q) + g^2P^2Q^2 \\ J_3 &= g(P\partial_x^3 Q - \partial_x P\partial_x^2 Q) + 4g^2P^2Q\partial_x Q \end{aligned} \tag{S112}$$

Note that the conservation equations $\frac{d}{\mathrm{d}t}C_n = 0$ holds if the current J_n vanishes at infinity, which is the case here, since we assume that the functions P, Q both vanish at infinity.

Finally, there is a relation between the scattering amplitude $a(k)$ and the *values* taken by the conserved quantities. Indeed, integrating the equation (S106) for $\partial_x \log \phi_1$ from $x = -\infty$ to $x = +\infty$ gives

$$\log a(k) = [\log(\phi_1 e^{\mathbf{i}\frac{k}{2}x})]_{-\infty}^{+\infty} = -g \int_{\mathbb{R}} \mathrm{d}x P(x,t)\Gamma(x,t,k) = \sum_{n\geqslant 1} \frac{C_n}{(\mathbf{i}k)^n} \tag{S113}$$

where we have used that $\phi_1 \simeq e^{-\mathbf{i}kx/2}$ as $x \to -\infty$, and (10) as $x \to -\infty$. Thus if one knows the solution for $a(k)$ one can expand in a Laurent series in $\mathbf{i}k$ to obtain the values of the C_n, as done in the text where they are denoted $C_n(g)$. Alternatively the Laurent series of $a(k)$ can be reconstructed from the knowledge of the values taken by the conserved quantities.

S-K. Calculation of $\Phi(H)$ from C_3

Here we calculate the space time integral of the solution of the P, Q system

$$g^2 \int_0^1 \mathrm{d}t \int_{\mathbb{R}} \mathrm{d}x\, P^2 Q^2 \equiv \Phi(H) \tag{S114}$$

where the $\equiv$ means that it also defines $\Phi(H)$ if the system is parameterized by H, which is equal to $\Phi(H_z)$ if the system is parameterized by $z = -g$, as in the text. We recall that H_z denotes the value of H where the minimum in $\Psi(z) = \min_H(\Phi(H) + ze^H)$ is attained. Here we show that for the droplet initial condition the resulting $\Phi(H)$ agrees with the Legendre transform of $\Psi(z)$ obtained in (20).

We first note that the third conserved quantity can also be rewritten as

$$C_3 = \int_{\mathbb{R}} \mathrm{d}x(g^2 P^2 Q^2 + gP\partial_x^2 Q) = \frac{g^2}{2}\int_{\mathbb{R}} \mathrm{d}x P^2 Q^2 + \frac{1}{2}\int_{\mathbb{R}} \mathrm{d}x J_2 \tag{S115}$$

where J_2 is the current given in (S112), where we used integration by part to replace the term $-\int_{\mathbb{R}} \mathrm{d}x \partial_x P \partial_x Q$ by $\int_{\mathbb{R}} \mathrm{d}x P \partial_x^2 Q$. Now since $\frac{d}{\mathrm{d}t} C_3 = 0$ we can integrate this equation over $t \in [0,1]$

$$C_3 = \frac{g^2}{2}\int_0^1 \mathrm{d}t \int_{\mathbb{R}} \mathrm{d}x P^2 Q^2 + \frac{1}{2}\int_0^1 \mathrm{d}t \int_{\mathbb{R}} \mathrm{d}x J_2 \tag{S116}$$

The last integral can be transformed using an integration by part and the conservation law (S108)

$$\int_0^1 \mathrm{d}t \int_{\mathbb{R}} \mathrm{d}x J_2 = -\int_0^1 \mathrm{d}t \int_{\mathbb{R}} \mathrm{d}x x \partial_x J_2 = -\int_0^1 \mathrm{d}t \int_{\mathbb{R}} \mathrm{d}x x \partial_t \tilde{C}_2 = g\int_{\mathbb{R}} \mathrm{d}x x P Q'|_{t=0} \tag{S117}$$

where we have used that $\tilde{C}_2 = gPQ'$ from (S111), and that $\int_{\mathbb{R}} \mathrm{d}x x P Q'|_{t=1} = 0$ since $P(x, t = 1) = \delta(x)$ (and $Q(x, t = 1)$ is a smooth function). Performing an additional integration by part we obtain

$$\int_0^1 \mathrm{d}t \int_{\mathbb{R}} \mathrm{d}x J_2 = -g\int_{\mathbb{R}} \mathrm{d}x Q(x,0)(xP'(x,0) + P(x,0)) \tag{S118}$$

Until now this is valid for any initial condition since we used only that $P(x, t = 1) = \delta(x)$.

Let us now specify to the droplet initial condition, i.e. $Q(x,0) = \delta(x)$. Since $P(x,0)$ is smooth we obtain

$$\int_0^1 \mathrm{d}t \int_{\mathbb{R}} \mathrm{d}x J_2 = -g\int_{\mathbb{R}} \mathrm{d}x Q(x,0)P(x,0) = -C_1 \tag{S119}$$

Putting together (S119) and (S115) we obtain that for the droplet initial condition

$$g^2 \int_0^1 \mathrm{d}t \int_{\mathbb{R}} \mathrm{d}x\, P^2 Q^2 = 2C_3(g) + C_1(g) \tag{S120}$$

In the text we have obtained

$$C_3(g = -z) = \frac{\mathrm{Li}_{5/2}(-z)}{\sqrt{16\pi}} = \frac{1}{2}\Psi(z) \quad , \quad C_1(g = -z) = \frac{1}{\sqrt{4\pi}}\mathrm{Li}_{3/2}(-z) = -z\Psi'(z) \tag{S121}$$

and we note that $-2g\partial_g C_3(g) = C_1(g)$. Hence we now obtain, with $g = -z$

$$g^2 \int_0^1 \mathrm{d}t \int_{\mathbb{R}} \mathrm{d}x \, P^2 Q^2 = \Psi(z) - z\Psi'(z) = \Phi(H = H_z) \tag{S122}$$

The last equality is the parametric solution of the Legendre transform $\Psi(z) = \min_H(\Phi(H) + ze^H)$, which is inverted by applying d/dz leading to $\Psi'(z) = e^{H_z}$ where H_z is the value of H where the minimum is attained.

S-L. Solutions of the WNT for general H, and solitonic solutions of the $\{P, Q\}$ system in the attractive case

In the text, we described the solutions of the $\{P, Q\}$ system applied to the WNT, i.e. with boundary conditions (8). It has the form (13), (12), in terms of the two functions $A_t(x)$, $B_t(x)$. In (11) we assumed that the integral over k is on the real axis. This is correct for $H < H_c$, and we recall (see Section S-A) that as H varies from $-\infty$ to H_c, the coupling g varies from $-\infty$ to $g = 1$. We recall below the expressions and properties of these functions $A_t(x)$, $B_t(x)$ in that regime.

For $H > H_c$ however the solution of the $\{P, Q\}$ system admits solitonic contributions. The existence of solitonic contributions is a known scenario in the inverse scattering method [66], which occurs when $r(k)$, has a pole for $\Im k > 0$, or $\tilde{r}(k)$, has a pole for $\Im k < 0$. This corresponds in the (v_1, v_2) system to a discrete part in the spectrum which is interpreted as a bound state of the Dirac operator. In this case we have to return to the formula for the $A_t(x), B_t(x)$ functions (11) in the text. This formula, where k is on the real axis, can be violated in the attractive case $g > 0$, as we find here for $H > H_c$. In this case, the functions $A_t(x), B_t(x)$ are the sum of two parts, one which is analogous to (11) with an integral over k on the real axis, usually called the radiative part, and the other being a solitonic part (i.e. of finite-rank). We derive below these two parts.

Droplet initial condition

Let us start with the droplet initial condition, and first recall the solution for $H < H_c$ obtained in the text. In this case $Q(x, 0) = Q_0(x) = \delta(x)$ and the symmetry $Q(x, t) = P(x, 1-t)$ holds for $t \in [0, 1]$. In addition, $P(x, t)$ and $Q(x, t)$ are both even functions.

Solution for $H < H_c$. Let us first recall the expressions for the functions $A_t(x), B_t(x)$ for $H < H_c$

$$A_t(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{\mathbf{i}kx - k^2 t + \mathbf{i}\varphi(k)} \frac{1}{\sqrt{1 - ge^{-k^2}}} \quad , \quad B_t(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{-\mathbf{i}kx - k^2(1-t) - \mathbf{i}\varphi(k)} \frac{1}{\sqrt{1 - ge^{-k^2}}} \tag{S123}$$

where $g = g(H)$ is obtained by inverting $H = \log(\Psi_0'(-g))$, with $\Psi_0(z)$ given in (S26), see Fig. S3, and

$$\varphi(k) = \fint_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \frac{k \log(1 - ge^{-q^2})}{q^2 - k^2} \tag{S124}$$

The function $\varphi(k)$ is odd and is plotted in Fig. S5. It decays as $1/k$ at large k, and expanding (S124) in k and g it is easy to see that it admits the following series representation in $1/k$ at large $|k|$

$$\varphi(k) = \frac{1}{2\pi} \sum_{m \geqslant 0} \frac{1}{k^{1+2m}} \Gamma(\frac{1}{2} + m) \mathrm{Li}_{\frac{3}{2}+m}(g) \tag{S125}$$

Using that $-\mathbf{i}\varphi(k) = \sum_{n \geqslant 1} \frac{C_n(g)}{(\mathbf{i}k)^n}$ we obtain the values of all the conserved charge C_n defined in the text and in Section S-J. One finds $C_n(g) = 0$ for even n, and for odd $n = 2m + 1$ with $m \geqslant 0$

$$C_{2m+1}(g) = (-1)^m \frac{1}{2\pi} \Gamma(\frac{1}{2} + m) \mathrm{Li}_{\frac{3}{2}+m}(g) \tag{S126}$$

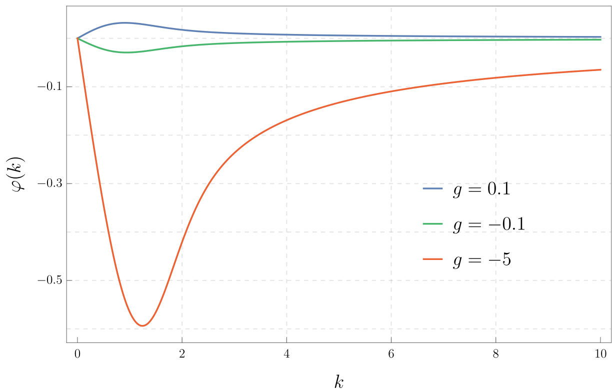

Figure S5. The phase $\varphi(k)$ defined in (S124) plotted versus k for various values of g

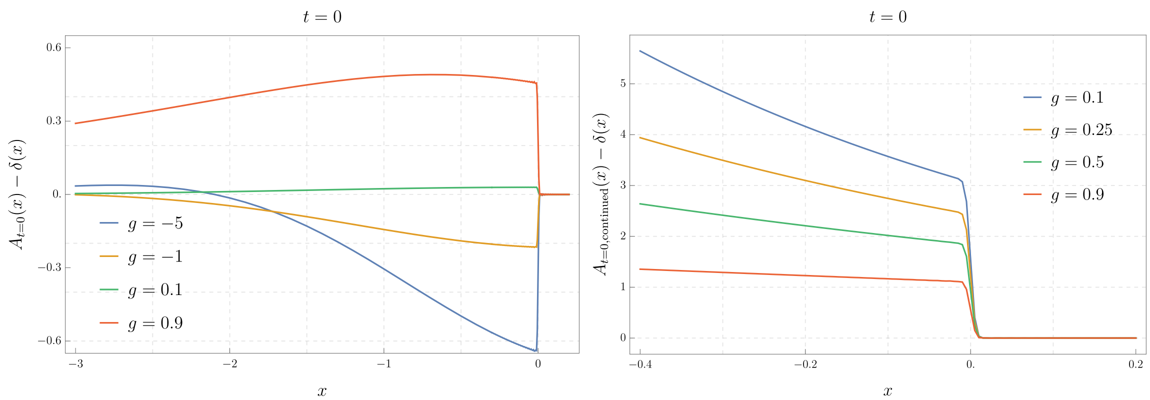

Figure S6. **Left:** Plot of the smooth part $A_0(x) - \delta(x)$ of $A_0(x)$ from (S123) plotted versus x for various values of g. One sees that it vanishes for $x > 0$. **Right:** the same plot for the second branch (S138) containing a solitonic contribution discussed below. The vanishing property at $x > 0$ still holds for this branch.

which for $m = \{0, 1\}$ is the result given in the text in (19). One can also obtains its small k behavior which is

$$\varphi(k) = kp_1(g) + k|k|p_2(g) + \mathcal{O}(k^3) \quad , \quad p_1(g) = \int_2^{+\infty} \frac{du}{2\pi u^{3/2}} \log\left(\frac{1 - ge^{-u}}{1 - g}\right) \quad p_2(g) = \frac{g(\text{arcsinh}(1) - \sqrt{2})}{\pi(g-1)} \tag{S127}$$

The functions $A_t(x)$ and $B_t(x)$ in (S123) are forward and backward solutions of the standard heat equation. They are related through $A_t(x) = B_{1-t}(x)$, a property specific to the droplet initial condition. For $g = 0$ they are simply equal to the heat kernel, i.e. $A_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}$, but at $g \neq 0$ they are non trivial, and very asymmetric in x. As found in the text (see also Section S-I), $A_{t=0}(x)$ is the sum of a delta function and of a smooth part, i.e. $A_0(x) = \delta(x) + \Theta(-x)\tilde{A}_0(x)$. Remarkably, the smooth part, which is $\mathcal{O}(g)$ at small g, vanishes for $x > 0$ for all g. The same property holds for $B_{t=1}(x)$ at $t = 1$. It is easy to see on the integral representation (S123) that there should be a delta function part: since $\varphi(k)$ vanishes at large k, the integrand at $t = 0$ becomes simply $e^{\mathbf{i}kx}$ at large $|k|$ (i.e. the reflection amplitude $r(k) \to 1$ at large $|k|$ in (11)). However it is non-trivial to see on the Fourier representations that these functions vanish for all $x > 0$. These functions are plotted for various t and g in the Figures S6 (left) and S7. In Fig. S6 (left) the smooth part at $t = 0$ has been plotted (i.e. the $\delta(x)$ piece was substracted) and we see that indeed it takes the form $\Theta(-x)\tilde{A}_0(x)$, i.e. it exhibits a jumps at $x = 0$ and it vanishes for

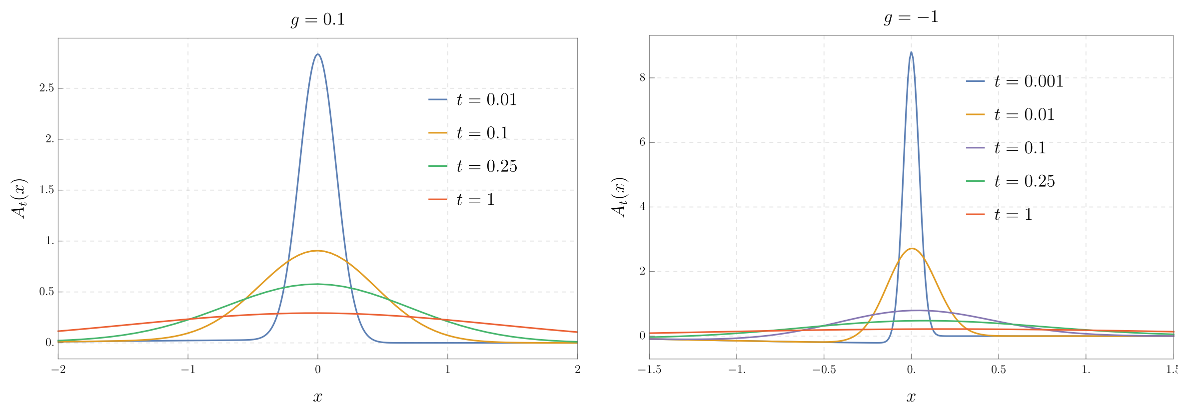

Figure S7. Plots of $A_t(x)$ from (S123) for various t and for some values of g. We see the broadening at $t > 0$ of the delta peak present at $t = 0$.

$x > 0$, a quite non trivial check. At $t > 0$ as we can see in Fig. S7, the delta peak in $A_t(x)$ broadens diffusively.

Solution for $H > H_c$. We now construct the second branch of solutions for $0 < g \leqslant 1$, which corresponds to $H > H_c$. Let us recall that for that branch the relation $g = g(H)$ is obtained by inverting $H = \log(\Psi'(-g))$, with $\Psi(z) = \Psi_0(z) + \Delta(z)$ given in (S26), see also (S27) and Fig. S3. We start again with the formula (17) for the Laplace transform $\hat{B}_1(s) = \int_{-\infty}^0 \mathrm{d}x e^{sx} B_1(x)$, which we now use to probe the structure of $r(k)$ in the complex plane for k. It reads

$$\hat{B}_1(s) = \exp(-f(s)) \quad , \quad f(s) := \int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \frac{s}{s^2+q^2} \log(1 - g e^{-q^2}) \tag{S128}$$

It is convenient to consider the following derivative of $f(s)$, obtained as

$$g\partial_g f(s) = 2\int_0^{+\infty} \frac{\mathrm{d}q}{2\pi} \frac{s}{s^2+q^2} \frac{-g e^{-q^2}}{1 - g e^{-q^2}} \tag{S129}$$

Upon the change of variable $y = e^{q^2}$, with $\mathrm{d}y/y = 2q\mathrm{d}q$, we can rewrite it as

$$g\partial_g f(s) = -\int_1^{\infty} \frac{\mathrm{d}y}{2\pi} \frac{s}{s^2+q(y)^2} \frac{g}{y} \frac{1}{q(y)} \frac{1}{y-g} \tag{S130}$$

This integral has the form (S30) with $z = -g$ which, according to (S31), leads to a jump $\Delta f(s)$ which obeys

$$g\partial_g \Delta f(s) = \mathbf{i} \frac{s}{s^2+q(g)^2} \frac{1}{q(g)} \tag{S131}$$

where $q(g)$ is the solution of $g = e^{q^2}$. Using that $g\partial_g = 1/(2q)\partial_q$, this is integrated into

$$\Delta f(s) = 2\mathbf{i} \arctan(\frac{q(g)}{s}) \tag{S132}$$

where, to ensure continuity at the branching point $g = 1$, we required that $\Delta f(s)|_{g=1} = 0$, using that $q(g=1) = 0$. The new branch thus corresponds to the following continuation of the Laplace transform of B_1, which contains an additional rational factor as compared the first branch.

$$\hat{B}_1(s) \to \hat{B}_1(s) e^{-2\mathbf{i}\arctan(\frac{q(g)}{s})} = \hat{B}_1(s) \frac{s - \mathbf{i}q(g)}{s + \mathbf{i}q(g)} = \hat{B}_1(s) \frac{s+\kappa_0}{s-\kappa_0} \tag{S133}$$

where $q(g) = \mathbf{i}\kappa_0$, for $0 < g \leqslant 1$, in terms of the real positive parameter κ_0, which satisfies $g = e^{-\kappa_0^2}$ and has the interpretation of a rapidity (see below). One goes from Laplace to Fourier by comparing the definition $\hat{B}_1(s) = \int_{-\infty}^0 \mathrm{d}x e^{sx} B_1(x) = \int_{\mathbb{R}} \mathrm{d}x e^{sx} B_1(x)$ with the inverse Fourier transform of (11), i.e.

$\tilde{r}(-k)e^{-k^2} = \int_{\mathbb{R}} \mathrm{d}x e^{-\mathbf{i}kx} B_1(x)$, leading to $r(k) = e^{k^2}\tilde{r}(-k) = \hat{B}_1(s)|_{s=-\mathbf{i}k+0^+}$. To obtain the continuation of $\tilde{r}(k)$ in the complex plane we must thus change $k \to -k$, i.e. substitute $s = \mathbf{i}k + 0^+$ in (S133), leading to

$$\tilde{r}(k) \to \tilde{r}(k)\frac{k - \mathbf{i}\kappa_0}{k + \mathbf{i}\kappa_0} \tag{S134}$$

We see that this continuation has a pole at $k = -\mathbf{i}\kappa_0$ in the lower half plane. According to the general theory [66] it implies that the function $B_t(x)$ has two parts,

$$B_t(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{-\mathbf{i}kx - k^2(1-t) - \mathbf{i}\varphi(k)} \frac{k - \mathbf{i}\kappa_0}{k + \mathbf{i}\kappa_0} \frac{1}{\sqrt{1 - ge^{-k^2}}} + 2\kappa_0 e^{-\kappa_0 x + \kappa_0^2(1-t) - \mathbf{i}\varphi(-\mathbf{i}\kappa_0)}, \quad g = e^{-\kappa_0^2} \tag{S135}$$

To obtain the first term (radiative part) we have used (S134) and the formula for the first branch of $\tilde{r}(k)$ given in the text and in (S123). The second term originates from the pole of $\tilde{r}(k)$ from (S134) in the complex plane at $k = -\mathbf{i}\kappa_0$, and has the (rank-one) solitonic form $e^{-\kappa_0 x + \kappa_0^2(1-t)}$ described in Section S-D. To obtain its amplitude we can focus on $t = 1$ in (S123) and compute the residue

$$-\mathrm{Res}_{k=-\mathbf{i}\kappa_0} \tilde{r}(k) e^{k^2} \frac{k - \mathbf{i}\kappa_0}{k + \mathbf{i}\kappa_0} = -\mathrm{Res}_{k=-\mathbf{i}\kappa_0} \hat{B}_1(s = \mathbf{i}k + 0^+) = 2\kappa_0 e^{-\mathbf{i}\varphi(-\mathbf{i}\kappa_0)} \tag{S136}$$

$$\mathbf{i}\varphi(-\mathbf{i}\kappa_0) = f(s = \kappa_0) = \int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \frac{\kappa_0}{\kappa_0^2 + q^2} \log(1 - ge^{-q^2}) \tag{S137}$$

Note the extra minus sign due to the fact that the contour around $-\mathbf{i}\kappa_0$ must be clockwise. Note also that the factor $1/\sqrt{1 - ge^{-k^2}}$ is only present for k on the real axis, hence it is not present in the residue of the pole at $k = -\mathbf{i}\kappa_0$. Similarly we obtain $r(k) \to r(k)\frac{k+\mathbf{i}\kappa_0}{k-\mathbf{i}\kappa_0}$ and

$$A_t(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{\mathbf{i}kx - k^2 t + \mathbf{i}\varphi(k)} \frac{k + \mathbf{i}\kappa_0}{k - \mathbf{i}\kappa_0} \frac{1}{\sqrt{1 - ge^{-k^2}}} + 2\kappa_0 e^{-\kappa_0 x + \kappa_0^2 t + \mathbf{i}\varphi(\mathbf{i}\kappa_0)} \tag{S138}$$

with the relation $A_t(x) = B_{1-t}(x)$, which is consistent with the symmetry $Q(x,t) = P(x, 1-t)$ of the solution for the droplet initial condition.

The function $A_t(x)$ for the second branch of solutions, given by (S138), is plotted in Fig. S8. On the left it is plotted for $g = 0.95$ for various t. For finite t one distinguishes the soliton shape, but as $t \to 0$ and for $x > 0$ it becomes again peaked around $x = 0$. Eventually, as we have checked in Fig. S6 (right), at $t = 0$ it vanishes for $x > 0$, which again is non trivial to see on the integral representation (S138). Again $A_0(x)$ is the sum of a delta function (with unit coefficient) and a continuous part, and we see that for $x > 0$, the positive part of the soliton exactly cancels with the radiative part.

In Fig. S8 (right) the two branches are plotted for $t = 0.25$ around the turning point at $g = 1$. One sees that the transition occurs smoothly. As $g \to 1$ along the first branch (i.e. $H \to H_c^-$) the function $A_t(x)$ decays more slowly for $x \to -\infty$, and for $g = 1$ is goes to a constant. This coincides with the generation of the soliton along the second branch, the function $A_t(x)$ is now growing as $x \to -\infty$.

Conserved quantities. Let us now evaluate the values of the conserved quantities associated with this new branch of solutions, which can be written $C_n(g) + \Delta C_n(g)$, where $C_n(g)$ are the values along the first branch. As explained in the text and in Section S-J they are obtained from Laurent expansion of $\log a(k)$ at large k. For the droplet initial condition, one has $b(k) = 1$, $r(k) = 1/a(k)$, $\log a(k) = -\mathbf{i}\varphi(k) + \frac{1}{2}\log(1 - ge^{-k^2})$. From the above, the continued version is obtained by replacing $\varphi(k) \to \varphi(k) + \Delta\varphi(k)$ with

$$\Delta\varphi(k) = 2\,\arctan(\frac{\kappa_0}{k}) \tag{S139}$$

where we recall $g = e^{-\kappa_0^2}$. The values of the additional contribution $\Delta C_n(g)$ to the conserved quantities C_n given in (S111) are obtained from the Laurent expansion (S139) of $\Delta\varphi(k)$ at large k, since $-\mathbf{i}\Delta\varphi(k) = \sum_{n\geqslant 1} \frac{\Delta C_n(g)}{(\mathbf{i}k)^n}$. It reads, for general odd integer n, and for $n = 1, 3$

$$\Delta C_n(g) = \frac{2}{n}\kappa_0^n \quad , \quad \Delta C_1(g) = 2\kappa_0 \quad , \quad \Delta C_3(g) = \frac{2}{3}\kappa_0^3 \quad , \quad \kappa_0 = (\log\frac{1}{g})^{1/2} \tag{S140}$$

Note that $\Delta C_n(g) = 0$ for even integer n. Since from (S121) one has $C_3(g) = \frac{1}{2}\Psi(-g)$, this result is consistent

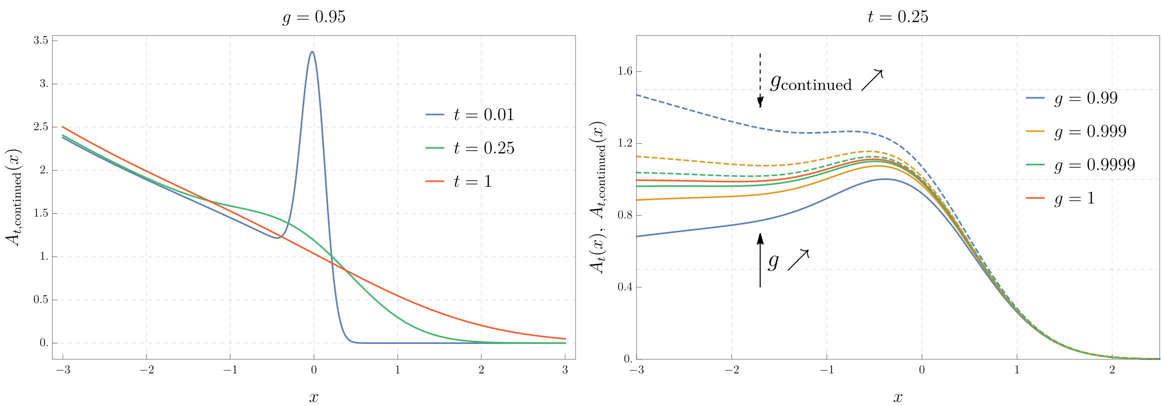

Figure S8. **Left:** the function $A_t(x)$ for various t and $g = 0.95$ ($H = 0.03564$ for a critical height $H_c = -0.3052$). **Right:** Both branches of solutions for $A_t(x)$ around the turning point $g = 1$ (for clarity the second branch is indicated as $A_t^{\text{continued}}(x)$) and the related curves are dashed).

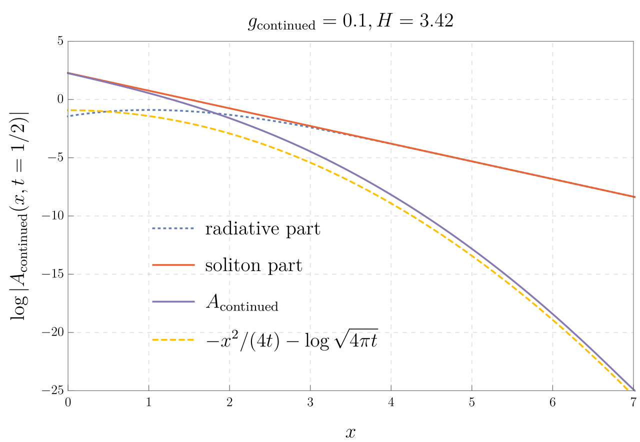

Figure S9. Plot of the logarithm of the absolute values of the radiative part, of the soliton part, and of the sum of the two terms $A_t(x)$ in (S138) for $t = 1/2$ and $g = 0.1$. We see that although the soliton exponential decay dominates at small x, at large x the two parts cancel and the remainder (the sum) is a fast decaying Gaussian, well fitted by $-x^2/(4t)$ in log scale.

with the known result for the jump in the second branch of $\Psi(z)$, see Eq. (S26)

$$\Delta(z) = 2\Delta C_3(g) = \frac{4}{3}\kappa_0^3 = \frac{4}{3}(\log(\frac{-1}{z}))^{3/2} \tag{S141}$$

Note the relation $-2g\partial_g \Delta C_3(g) = \frac{1}{\kappa_0}\partial_{\kappa_0}\Delta C_3 = 2\kappa_0 = \Delta C_1$.

Limit $H \to +\infty$

We now study the limit $H \to +\infty$ of our exact solution for the WNT for the droplet initial condition, and show that one recovers the approximate solution obtained in Ref. [54] in that limit. This limit correspond to the second branch of solutions with g vanishing as $g \simeq 2H^{1/2}e^{-H}$. Indeed, for $H \in [H_c, +\infty)$ the relation between g and H is determined by $\Psi'(z) = e^H$ with $\Psi(z) = -\frac{1}{\sqrt{4\pi}}\text{Li}_{5/2}(z) + \frac{4}{3}(\log(\frac{-1}{z}))^{3/2}$, from Ref. [47], as recalled in (S26). Setting $z = -g$, this leads to $e^H \simeq \frac{2}{g}(\log(\frac{1}{g}))^{1/2}$ as $g \to 0^+$.

In that limit the solitonic part in (S138) plays a dominant role. The rapidity parameter κ_0 is given by

$\kappa_0^2 = \log(\frac{1}{g}) \simeq H - \frac{1}{2}\log(4H)$, hence $\kappa_0 \to +\infty$. To approach the problem, let us first write the solution of the $\{P, Q\}$ system if one neglects the radiative part in (S138) (for any κ_0). This soliton only solution corresponds to the rank-one soliton described in Section S-D, more precisely (S54), with $\kappa = \mu = \kappa_0$, $\tilde{q} = 2\kappa_0 e^{\mathbf{i}\phi(\mathbf{i}\kappa_0)}$ and $\tilde{p} = 2\kappa_0 e^{\kappa_0^2} e^{-\mathbf{i}\phi(-\mathbf{i}\kappa_0)}$. Using (S55) we obtain this soliton-only solution, which we denote $P_s(x,t), Q_s(x,t)$, as

$$Q_s(x,t) = \frac{2\kappa_0 e^{\mathbf{i}\phi(\mathbf{i}\kappa_0)} e^{-\kappa_0 x + \kappa_0^2 t}}{1 + g e^{\kappa_0^2} e^{2\mathbf{i}\phi(\mathbf{i}\kappa_0)} e^{-2\kappa_0 x}} = \frac{\kappa_0 e^{\kappa_0^2 t}}{\cosh(\kappa_0(x - y_0))} \quad , \quad g P_s(x,t) Q_s(x,t) = \frac{\kappa_0^2}{\cosh^2(\kappa_0(x - y_0))} \tag{S142}$$

where we have used that $g = e^{-\kappa_0^2}$, and where $\kappa_0 y_0 = \frac{1}{2}\log(g e^{\kappa_0^2} e^{2\mathbf{i}\phi(\mathbf{i}\kappa_0)}) = \mathbf{i}\phi(\mathbf{i}\kappa_0)$, a quantity which is real and given in (S137) (recalling that $\varphi(k)$ is an odd function). Until now this is valid for any κ_0, i.e. $P_s(x,t), Q_s(x,t)$ are exact solutions of the $\{P, Q\}$ system, which however do not obey the boundary data conditions for WNT (8). These are ensured only if the radiative part is added. In the limit $\kappa_0 \simeq H^{1/2} \to +\infty$, and in a certain range of x and t, the radiative part is comparatively smaller and $P_s(x,t), Q_s(x,t)$ are very good approximations of the exact solution.

This can be compared with the solution in Ref. [54] in the "boundary layer", for $\varrho(x,t) = \varrho_{\rm bl}(x)$ given in Eq. (50) there, and $\mathsf{h}(x,t) = \mathsf{h}_{\rm bl}(x,t)$ given there in Eq. (53). Using our equation (S46) and the dictionary (S47), we see that $\varrho_{\rm bl}(x)$ is identical with $-4gP_s(x,t)Q_s(x,t)$ if one identifies $c = 2\kappa_0^2$. This identification is correct since $c = -\mathsf{H}$ (below (53) there) and $-\mathsf{H} = 2H \simeq 2\kappa_0^2$ from the dictionary (S47). Note that we have neglected the shift y_0 in (S142) since $\kappa_0 y_0 = \mathbf{i}\phi(\mathbf{i}\kappa_0) \to 0$ in the limit $\kappa_0 \to +\infty$ from (S137). We also find that $\mathsf{h}_{\rm bl}(x,t) = 2\log\cosh(\sqrt{c/2}x) - ct$ identifies (up to a constant term $-2\log\kappa_0$) with $-2h_s(x,t)$ where $h_s(x,t) = \log Q_s(x,t)$ and $Q_s(x,t)$ given in (S142).

We have checked numerically that the exact solution for large positive H is indeed very close to the above soliton-only solution for $x = \mathcal{O}(1/\kappa_0)$, i.e. in the boundary layer of Ref. [54]. In that region it is time independent except very near $t = 0, 1$. This can be seen in Fig. S11 already for a moderately large value of H. In addition, there is a second region, for x outside of this boundary layer where the decay of $Q(x,t)$ is Gaussian $h(x,t) \sim -x^2/(4t)$, in agreement with the arguments in Ref. [54], see Eq. (57) there (with $L = 0$ there). The way it arises here is also non trivial and as follows. We recall that $Q(x,t) \simeq A_t(x)$ as $x \to +\infty$, from (13) since in $A_x(t)$ and $B_x(t)$ vanish in that limit. We note that the decay at large x of the soliton part in (S138) is exponential, i.e. much slower than Gaussian. However, what happens is that at large positive x the first term in (S138) (the radiative term) cancels the soliton, and what remains is indeed decaying as a Gaussian in x. This is illustrated in Fig. S9.

General initial condition

Let us now consider a general initial condition $Q_0(x)$, which we take even in x for simplicity. As we have seen in the text, the only difference with the droplet case is that the amplitude $b(k)$ is now a non-trivial even function of k. Some of its properties are obtained in Sections S-F and S-G. One has now $r(k) = b(k) e^{k^2} \tilde{r}(-k)$, and now $A_t(x) = (\hat{b} * B_{1-t})(x)$ where $\hat{b}(x)$ denote the Fourier transform of $b(k)$ and $*$ the convolution.

Principal branch. As shown in the text, one obtains

$$A_t(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{\mathbf{i}kx - k^2 t + \mathbf{i}\varphi(k)} \frac{b(k)}{\sqrt{1 - g b(k) e^{-k^2}}} \quad , \quad B_t(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{-\mathbf{i}kx - k^2(1-t) - \mathbf{i}\varphi(k)} \frac{1}{\sqrt{1 - g b(k) e^{-k^2}}} \tag{S143}$$

in agreement with the general relation $A_t(x) = (\hat{b} * B_{1-t})(x)$, where now the phase reads

$$\varphi(k) = \int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \frac{k}{q^2 - k^2} \log(1 - g b(q) e^{-q^2}) \tag{S144}$$

We stress that both $b(k)$ and $\varphi(k)$ depend on g, and the solution for the droplet initial condition is recovered setting $b(k) = 1$. The associated conserved charges are obtained as before from the Laurent series $-\mathbf{i}\varphi(k) = \sum_{n\geqslant 1} \frac{C_n(g)}{(\mathbf{i}k)^n}$ which leads to

$$C_{2m+1}(g) = (-1)^{m-1} \int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} q^{2m} \log(1 - gb(q)e^{-q^2}) \tag{S145}$$

which, we can check, recovers (S126) for $b(q) = 1$. This leads to $C_1(g) = \int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \mathrm{Li}_1(gb(q)e^{-q^2})$ since $\mathrm{Li}_1(y) = -\log(1-y)$. From the relation $C_1(g) = g\int_{\mathbb{R}} \mathrm{d}x PQ|_{t=1} = gQ(0,1) = ge^{H} = -z\Psi'(z)$, with $g = -z$, one obtains

$$-z\Psi'(z) = \int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \mathrm{Li}_1(-zb(q)e^{-q^2}) \tag{S146}$$

If $b(q)$ is z independent one can integrate this relation using $z\partial_z \mathrm{Li}_n = \mathrm{Li}_{n-1}$, which leads to forms similar to the one displayed in [50, Table 7.1]. However, this is not the case in general.

Until now this is the principal branch. Since $\mathrm{Li}_1(y)$ has a branch cut on the real axis for $y > 1$, the discussion of the possible other branches depends on the behavior of the function $gb(q)e^{-q^2}$ as a function of q (recalling that $b(q)$ is an even function). Let us denote $M(g) = \max_{q\in\mathbb{R}} gb(q)e^{-q^2}$. The main branch terminates at $g = g_c$ such that $M(g_c) = 1$.

Second branch. We perform the same analysis as for the droplet case. The Laplace transform $\hat{B}_1(s) = e^{-f(s)}$ with now $f(s) := \int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \frac{s}{s^2+q^2} \log(1 - gb(q)e^{-q^2})$. We will assume that $b(q)e^{-q^2}$ is a decreasing function of q^2. In that case the branching point is at $g_c = 1/b(0)$. Taking the derivative $g\partial_g$ it is then possible to perform the change of variable $y = e^{q^2}/b(q)$, with $\mathrm{d}y/y = (2q - \frac{b'(q)}{b(q)})\mathrm{d}q$, leading to

$$g\partial_g f(s) = -2\int_1^\infty \frac{\mathrm{d}y}{2\pi} \frac{s}{s^2 + q(y)^2} \frac{g}{y} \frac{1}{(2q(y) - \frac{b'(q(y))}{b(q(y))})} \frac{1}{y-g} (1 + \frac{g\partial_g b(q(y))}{b(q(y))}) \tag{S147}$$

Since the integral has the form (S30) we can apply (S31) to obtain $g\partial_g \Delta f(s)$ and, upon integration using that $g\partial_g = (1 + g\frac{\partial_g b(q(g))}{b(q(g))})/(2q - \frac{b'(q(g))}{b(q(g))})\partial_q$ we obtain the jump

$$\Delta f(s) = 2\mathbf{i} \arctan(\frac{q(g)}{s}) \tag{S148}$$

where $q(g)$ is the solution of $gb(q)e^{-q^2} = 1$. Similar manipulations as in the droplet case lead to

$$B_t(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{-\mathbf{i}kx - k^2(1-t) - \mathbf{i}\varphi(k)} \frac{k - \mathbf{i}\kappa_0}{k + \mathbf{i}\kappa_0} \frac{1}{\sqrt{1 - gb(k)e^{-k^2}}} + 2\kappa_0 e^{-\kappa_0 x + \kappa_0^2(1-t) - \mathbf{i}\varphi(-\mathbf{i}\kappa_0)} \tag{S149}$$

and

$$A_t(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{\mathbf{i}kx - k^2 t + \mathbf{i}\varphi(k)} \frac{k + \mathbf{i}\kappa_0}{k - \mathbf{i}\kappa_0} \frac{b(k)}{\sqrt{1 - gb(k)e^{-k^2}}} + 2\kappa_0 b(\mathbf{i}\kappa_0) e^{-\kappa_0 x + \kappa_0^2 t + \mathbf{i}\varphi(\mathbf{i}\kappa_0)} \tag{S150}$$

where κ_0 is now the solution of the equation

$$g = \frac{e^{-\kappa_0^2}}{b(\mathbf{i}\kappa_0)} \tag{S151}$$

This is again in agreement with the general relation $A_t(x) = (\hat{b} * B_{1-t})(x)$. Since one still has $\Delta\varphi(k) = 2\arctan(\frac{\kappa_0}{k})$ the shift in the values of the conserved quantities are again $\Delta C_n(g) = \frac{2}{n}\kappa_0^n$, i.e. $\Delta C_1(g) = 2\kappa_0$ and $\Delta C_3(g) = \frac{2}{3}\kappa_0^3$.

Note that if $b(q)e^{-q^2}$ is not a decreasing function of q^2 the analysis of the various branches of solutions may be more complicated.

S-M. Numerical calculation of P, Q, h and Fredholm inversions

In this Section we evaluate numerically the solutions $P(x,t)$ and $Q(x,t)$ of the $\{P,Q\}$ system from the operator inversion formula (13) in the text. In the first part we explain the numerical algorithm to handle formula (13) for given functions $A_t(x)$, $B_t(x)$. In the second part we apply it to the WNT boundary data (8) in the case of the droplet initial condition. In that case the operators $\mathcal{A}_{xt}$ and $\mathcal{B}_{xt}$ and their associated functions $A_t(x)$, $B_t(x)$ are known from the Section S-L. We plot the optimal height field $h(x,t) = -\mathsf{h}(x,t)/2$ and the optimal noise field $2\tilde{h}(x,t) = 2gP(x,t)Q(x,t) = -\varrho(x,t)/2$ for various values of g and H.

Algorithm for the operator inversion and the numerical evaluation of the solutions of the $\{P,Q\}$ system

All operators considered in this work act on $\mathbb{L}^2(\mathbb{R}_+)$. From a numerical analysis point of view, we will approximate these operators as acting on a discrete weighted space defined by a prescribed quadrature similarly to Bornemann's method to evaluate Fredholm determinants of operators acting on continuous space [96]. Concretely, since all functions considered have an exponential decay towards $+\infty$, any non-trivial integral over $\mathbb{R}_+$ will be approximated as

$$\int_{\mathbb{R}_+} \mathrm{d}x\, f(x) \simeq \sum_{j=1}^{m} w_j f(y_j) \tag{S152}$$

where the set $\{w_j, y_j\}_{j=1}^m$ defines a quadrature rule over the interval $[0, M]$ with $M \gg 1$, see [96] for more details. Recall that the solution $Q(x,t)$ reads

$$Q(x,t) = \langle\delta|\, \mathcal{A}_{xt} \frac{I}{I + g\mathcal{B}_{xt}\mathcal{A}_{xt}} \,|\delta\rangle \tag{S153}$$

we see that two types of integrals are considered. Either we multiply two operators together, e.g. $\mathcal{B}_{xt}\mathcal{A}_{xt}$, or we multiply an operator with a bra or a ket, e.g. $\langle\delta|\, \mathcal{A}_{xt}$.

- The products of type $\langle\delta|\, \mathcal{A}_{xt}$ are considered as trivial since

$$(\langle\delta|\, \mathcal{A}_{xt})(v) = \int_{\mathbb{R}_+} \mathrm{d}u \delta(u) A_t(x+u+v) = A_t(x+v) \tag{S154}$$

  we only have to enforce the left variable of $\mathcal{A}_{xt}$ to be equal to zero.

- Operator products are considered as non-trivial and the quadratures will be inserted there as

$$\begin{aligned}(\mathcal{B}_{xt}\mathcal{A}_{xt})(u,v) &= \int_{\mathbb{R}_+} \mathrm{d}r\, B_t(x+u+r)A_t(x+r+v) \\ &\simeq \sum_{j=1}^{m} w_j B_t(x+u+y_j)A_t(x+y_j+v) \\ &\simeq \sum_{j=1}^{m} (B_t(x+u+y_j)\sqrt{w_j})(\sqrt{w_j}A_t(x+y_j+v))\end{aligned} \tag{S155}$$

  where we split the weights to keep the operators symmetric.

- In a nutshell, if an operator is not involved in a product with a bra or a ket, its approximation will read e.g.

$$\mathcal{A}_{xt} = \left(\sqrt{w_i}A_t(x+y_i+y_j)\sqrt{w_j}\right)_{i,j=1}^m := \sqrt{w}A_{xt}\sqrt{w} \tag{S156}$$

  and similarly for $\mathcal{B}_{xt}$. We can interpret w as being an $m \times m$ diagonal matrix with entries w_i and A_{xt} as being a symmetric $m \times m$ matrix with elements $A_t(x+y_i+y_j)$. If an operator is involved in a product with a bra or a ket, its approximation will read e.g.

$$\langle\delta|\,\mathcal{A}_{xt} = \left(A_t(x+y_j)\sqrt{w_j}\right)_{j=1}^{m} \tag{S157}$$

Lifting these results from the operators $\{\mathcal{A}_{xt},\mathcal{B}_{xt}\}$ to the solutions of the $\{P,Q\}$ system is now simple. Rewriting $Q(x,t)$ as

$$Q(x,t) = \langle\delta|\,\mathcal{A}_{xt}\,|\delta\rangle - g\,\langle\delta|\,\mathcal{A}_{xt}\frac{I}{I+g\mathcal{B}_{xt}\mathcal{A}_{xt}}\mathcal{B}_{xt}\mathcal{A}_{xt}\,|\delta\rangle \tag{S158}$$

we approximate it as

$$Q(x,t) \simeq \tilde{\tilde{Q}}(x,t) = A_t(x) - g\,\langle\delta|\,A_{xt}\sqrt{w}\frac{I}{I_m+g\sqrt{w}B_{xt}wA_{xt}\sqrt{w}}\sqrt{w}B_{xt}wA_{xt}\,|\delta\rangle \tag{S159}$$

With this approximation, the inverse $(I+g\mathcal{B}_{xt}\mathcal{A}_{xt})^{-1}$ has been transformed into the inverse of an $m\times m$ dimensional matrix. In Eq. (S158), we completed the numerator to avoid the evaluation of the inverse against the ket $|\delta\rangle$. For the purpose of the numerical evaluations of this work, the typical values $m=20$ and $M=10$ are sufficient to obtain solutions of the $\{P,Q\}$ system up to $\sim 10^{-5}$ precision. We present below the Mathematica implementation of this method used to produce the numerical results of this work. We first present the quadrature implementation for the kernel $K(v,v') = \int_{\mathbb{R}_+} \mathrm{d}r B_t(x+v+r)A_t(x+v'+r)$ and then the one for the solution $Q(x,t)$.

```
Needs["NumericalDifferentialEquationAnalysis`"]
M = 10;
m = 20;

(* we assume the functions A and B to be already defined,
x is the space, t is the time and g is the coupling constant *)

A[x_,t_,g_]:=A[x,t,g];
B[x_,t_,g_]:=B[x,t,g];

K[x_, y_, t_, g_] := Module[{r, w},
  {r, w} = Transpose[GaussianQuadratureWeights[m, 0, M]];
  Flatten[Outer[B, x + r, {t}, {g}] w ].Flatten[Outer[A, y + r, {t}, {g}]]
  ];

Q[x_, t_, g_] := Module[{u, w},
    {u, w} = Transpose[GaussianQuadratureWeights[m, 0, M]];
    w = Sqrt[w];
    A[x, t, g] - g Flatten[ Outer[A, u + x, {t}, {g}] w].
        Inverse[IdentityMatrix[m] + g  Outer[Times, w, w] ArrayFlatten[Outer[K, u + x, u + x, {t}, {g}]]].
        Flatten[w Outer[K, u + x, {x}, {t}, {g}]]
    ];
```

Results for the droplet initial condition

We have computed the functions $P(x,t)$ and $Q(x,t)$ using the above numerical scheme for the functions $A_t(x)$, $B_t(x)$ defined in Eq. (S123) for the main branch $H<H_c$, and in (S135) and (S138) for the second branch $H>H_c$. Note that for $t=1$, there is a simpler formula, $Q(x,t=1)=A_1(|x|)$, see section S-H.

A first numerical check of our precision is to verify that the functions P,Q are even in x, which is quite a non-trivial check given that the formula a very asymmetric in $x\to -x$. We have verified this fact with the same numerical precision as mentioned in the previous section.

In Fig. S10 we have plotted the optimal height $h(x,t) = \log Q(x,t)$ for various values of g, H and time t. The black dot on the figure is the value of $h(0,1) = H$ reached at the observation time $t = 1$. Far from the origin the behavior is the parabola $-x^2/(4t)$, but an important deformation arises at smaller x.

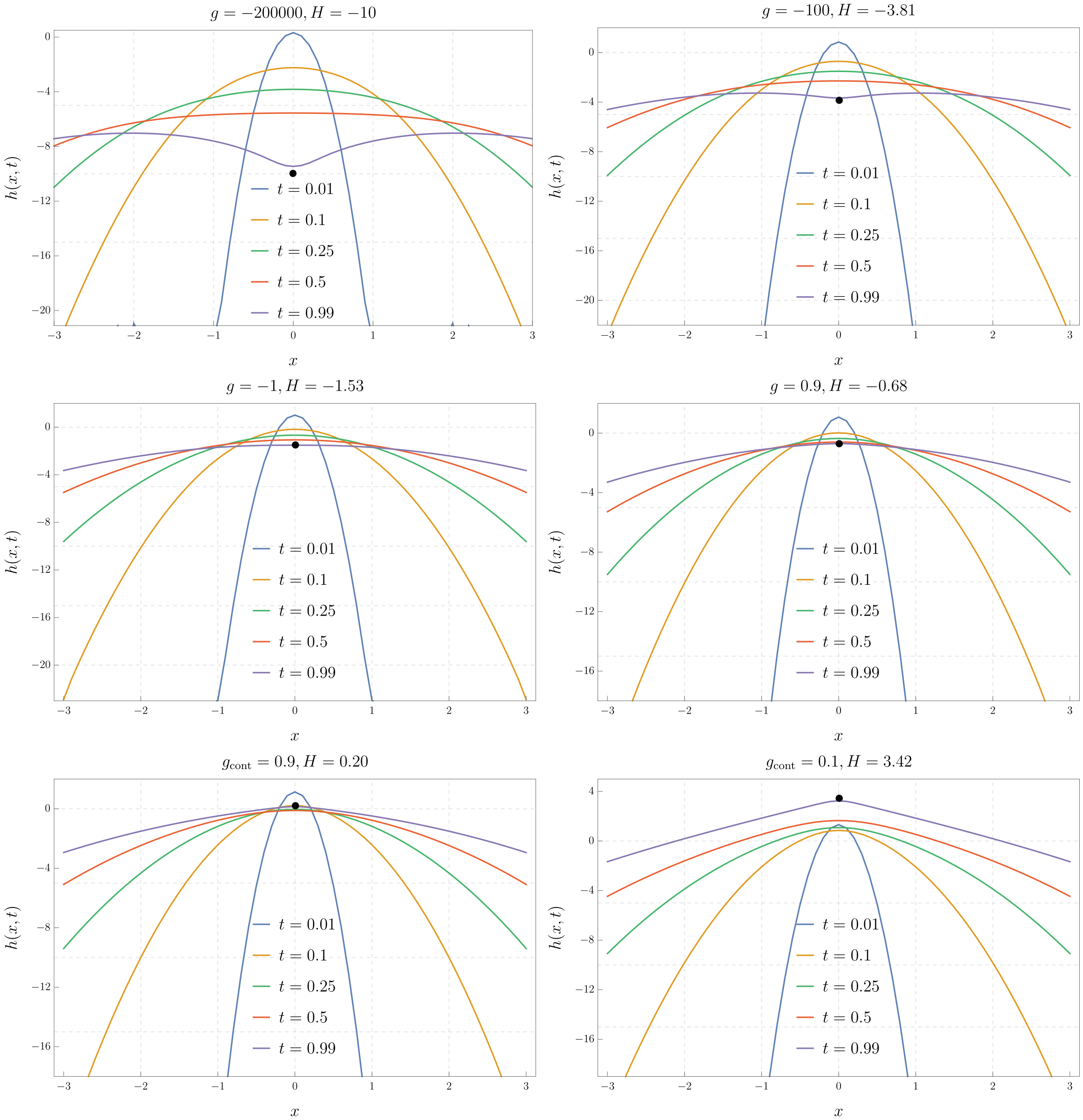

Figure S10. Optimal height $h(x,t)$ evaluated at various times and various final height at the origin H. We observe that at short time the height evolves from a parabolic profile to a non-trivial profile at final time $t = 1$ which behavior depends on whether $H \gg 1$ or $H \ll 1$.

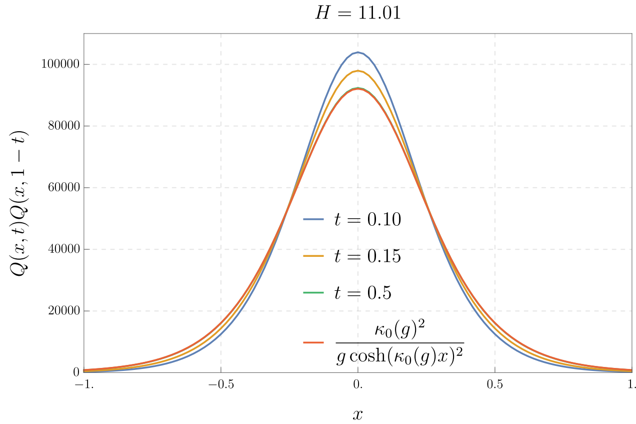

Figure S11. Optimal noise profile $\tilde{h}(x,t)/g = Q(x,t)Q(x,1-t)$ as a function of x for $H = 11.01$ and various times t. A weak time dependence can be seen. For $t = 1/2$ (in green) it already overlaps almost perfectly with the "soliton only" (time independent) solution, valid in the boundary layer $x \sim 1/\kappa_0$ for $H \to +\infty$ (in red)

In Fig. S11 we have plotted the optimal noise. The optimal noise can be identified by comparing the right hand side of the saddle point equation (S40) with the KPZ equation itself (S37), leading to $\sqrt{2}T^{1/4}\tilde{\eta}_{\rm opt}(x,t) = 2\tilde{h}(x,t)$. For convenience we denote $2\tilde{h}(x,t)$ at the saddle point as the optimal noise, i.e. the field $-\varrho(x,t)/2$ in [54]. We would like to point out that for the droplet initial condition it is directly related to the optimal height by the relation

$$2\tilde{h}(x,t) = 2gP(x,t)Q(x,t) = 2gQ(x,t)Q(x,1-t) = 2ge^{h(x,t)+h(x,1-t)} \tag{S160}$$

In the Fig. S11 we have plotted $Q(x,t)Q(x,1-t)$ for various t and for $g = 0.0001$ which corresponds to $H = 11.01$. Since it is a moderately large positive value for H, we compare it with the soliton only solution given in (S142), i.e. $Q_s(x,t)Q_s(x,1-t) = \frac{\kappa_0^2}{g\cosh(\kappa_0 x)^2}$, where $g = e^{-\kappa_0^2}$. This soliton-only solution is time independent and is expected to hold for $H \to +\infty$ inside the boundary layer $x \sim 1/\kappa_0$ except very near $t = 0$ or $t = 1$ [54]. For the moderately large value of H considered here, we still some small time dependence, while for $t = 1/2$ we already see excellent agreement with the soliton-only solution.

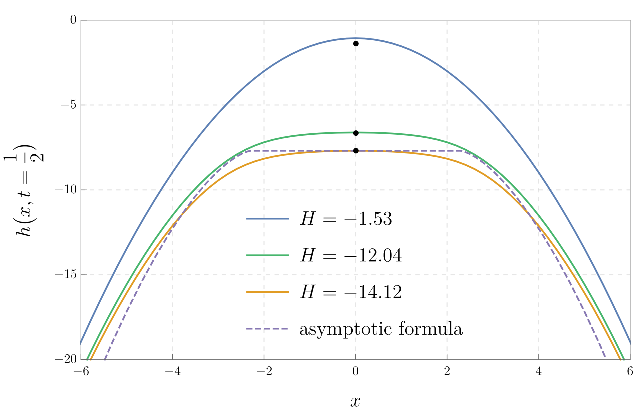

Figure S12. Optimal height $h(x,t=1/2)$ as a function of x for three negative values of the field (indicated by the black dots), i.e. in the left tail regime. As H decreases we can see a plateau forming near the origin. The dashed line is the parametric function $x(u) = \frac{2}{\pi}\sqrt{|H-\hat{H}_0|}(1+u\ \arctan u)$, $h(u) = \frac{H+\hat{H}_0}{2}[1+\frac{2}{\pi}(u+(u^2-1)\arctan u]$ for $u \geqslant 0$, based on the scaling form predicted to hold in the limit $H \to -\infty$ in Ref. [64]

In the Fig. S12 we have plotted $h(x,t=1/2)$ for three different negative values for H (left tail region,

$g < 0$). The back dot indicates the value of $h(0,1) = H$ reached at the observation time $t = 1$. The prediction from [64] is that for $H \to -\infty$ the optimal height profile develops at $t = 1/2$ a plateau of height $h(x, t = 1/2) = \frac{H+\hat{H}_0}{2}$ in a band $|x| < \frac{2}{\pi}\sqrt{|H - \hat{H}_0|}$. Furthermore, away from this band the height reaches the parabolic form $-x^2/(4t)$. As we can see for the most negative value of H, $H = -14.12$, one clearly sees on Fig. S12 the plateau developing. We have also plotted for comparison the prediction of Ref. [64] for the crossover function between the plateau regime and the parabola regime. For this value of H the asymptotic regime is not yet reached, however its main features are visible.

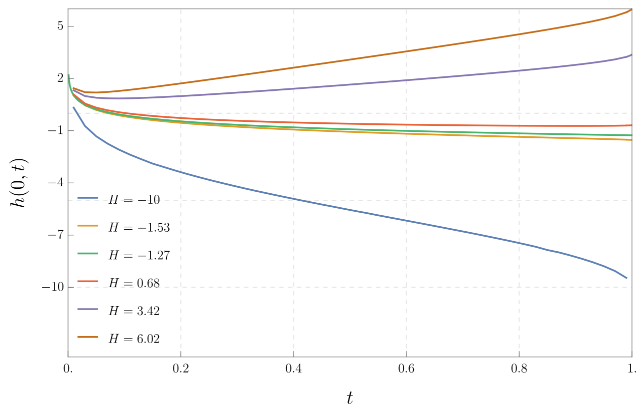

Figure S13. Optimal height at the origin, $h(0,t)$, as a function of t for various values of H. The green curve corresponds to the typical height, i.e. $H = \hat{H}_0 = -1.27$, see text.

Finally in Fig. S13 we have plotted the optimal height at the origin, $h(0,t)$, as a function of t for various values of H. This shows how the value $h(0,1) = H$ is reached for various H as the time increases. The green curve is the typical height $h(0,t) = -\log(\sqrt{4\pi t})$ and corresponds to $g = 0$ and $H = \hat{H}_0 = -\log(\sqrt{4\pi}) = -1.27$. All the curves are asymptotic to that one in the very small time regime, $t \to 0$ (not shown in details).