LIPSCHITZ METRIC FOR THE NOVIKOV EQUATION

HONG CAI, GENG CHEN, ROBIN MING CHEN, AND YANNAN SHEN

ABSTRACT. We consider the Lipschitz continuous dependence of solutions for the Novikov equation with respect to the initial data. In particular, we construct a Finsler type optimal transport metric which renders the solution map Lipschitz continuous on bounded sets of $H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})$, although it is not Lipschitz continuous under the natural Sobolev metric from energy law due to the finite time gradient blowup. By an application of Thom's transversality theorem, we also prove that when the initial data is in an open dense subset of $H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})$, the solution is piecewise smooth. This generic regularity result helps us extend the Lipschitz continuous metric to the general weak solutions. Our method of constructing the metric can be used to treat other kinds of quasi-linear equations, provided a good knowledge about the energy concentration.

CONTENTS

1. INTRODUCTION

Many evolutionary partial differential equations (PDEs) have the general form

$$u_t + Lu = 0, \quad u(0) = u_0.$$

Classical well-posedness theory suggests the existence of a continuous semigroup of solutions, at least for a short time. For a large class of semi-linear PDEs, basic techniques such as Picard and Duhamel iteration can be applied to obtain Lipschitz continuity of the semigroup, that is, for any pair of solutions u, v, it holds

$$\frac{d}{dt}\|u(t) - v(t)\| \lesssim \|u(t) - v(t)\| \tag{1.1}$$

1

for a suitable (Sobolev) norm. Typical examples include the Korteweg-de Vries (KdV) equation, the nonlinear Schrödinger (NLS) equation, the semi-linear wave equation, and so on.

On the other hand, a noteworthy exception is provided by many *quasi-linear* equations. Due to the dominating nonlinearity, the initial information of the data can determine the later dynamics in a substantial way. In particular, solutions with smooth initial data can lose regularity in finite time, and one cannot in general expect (1.1) to hold true under the natural energy norms.

In this paper, we would like to address the issue of the Lipschitz continuity of the flow map using the following quasi-linear equation, namely the Novikov equation

$$u_t - u_{xxt} + 4u^2 u_x = 3uu_x u_{xx} + u^2 u_{xxx}. \tag{1.2}$$

This equation was derived by Novikov [21] in a symmetry classification of nonlocal PDEs with cubic nonlinearity, and can in some sense be related to the well-known Camassa-Holm equation [8, 12]. In fact writing the Novikov equation (1.2) in the weak form

$$u_t + u^2 u_x + \partial_x(1 - \partial_x^2)^{-1}\left(u^3 + \frac{3}{2}uu_x^2\right) + (1 - \partial_x^2)^{-1}\left(\frac{1}{2}u_x^3\right) = 0, \tag{1.3}$$

one may recognize the similarities with the Camassa-Holm equation, which, in a nonlocal form, reads

$$u_t + uu_x + \partial_x(1 - \partial_x^2)^{-1}\left(u^2 + \frac{1}{2}u_x^2\right) = 0. \tag{1.4}$$

Analytically, the Novikov equation also shares many properties in common with the Camassa-Holm equation, among which the two most remarkable features are the breaking waves and peakons. From examining the weak formulation (1.3), a transport theory can be applied to derive the blow-up criterion which asserts that singularities are caused by the focusing of characteristics. This in combined with the H^1-conservation of solutions indicates that the exact blow-up scenario is in the sense of *wave-breaking*, i.e., the solution remains bounded but its slope becomes infinite in finite time. Some results on this issue can be found, for instance, in [11, 18].

As an example, the wave-breaking phenomenon can also be manifested by the so-called *multi-peakon solutions*. The Novikov multi-peakon solution takes the form

$$u(t, x) = \sum_{i=1}^{N} p_i(t)e^{|x - q_i(t)|} \tag{1.5}$$

subject to the following equations of motion for the peak positions $q_i(t)$ and amplitudes $p_i(t)$ [16, 17]

$$\begin{cases} \dot{q}_i = \sum_{j,k=1}^{N} p_j p_k e^{-|q_i - q_j| - |q_i - q_k|}, \\ \dot{p}_i = p_i \sum_{j,k=1}^{N} p_j p_k \mathrm{sgn}(q_i - q_j)e^{-|q_i - q_j| - |q_i - q_k|}. \end{cases} \tag{1.6}$$

It can be seen that a single peakon always travels to the right since $\dot{q}_i = u(q_i)^2$. Therefore for a Novikov peakon pair, only over-taking collisions may take place [19], which differs from the Camassa-Holm case where head-on collisions are also possible. In a typical situation when two peakons of anti-strength ($p_1 p_2 < 0$) cross each other at time t_*, then as $t \to t_*^-$,

$$\begin{aligned} p_1(t) \to +\infty \text{ (or } -\infty), \quad p_2(t) \to -\infty \text{ (or } +\infty), \quad p_1(t) + p_2(t) \to p_*, \\ q_1(t), q_2(t) \to q_*, \quad q_1(t) < q_2(t) \text{ for } t < t_*^-, \end{aligned} \tag{1.7}$$

for some $p_*, q_* \in \mathbb{R}$; see [19]. In particular we have $\|u_x\|_{L^\infty} \to \infty$. In this case, one needs to be careful with the meaning of continuing a solution beyond a collision.

In a recent work [9], we managed to find a way to uniquely extend the Novikov solution beyond the point of collision, such as multi-peakon solutions, while keeping the "total energy" conserved.

Moreover, the result in [9] applies to the general case of continuing solution after wave-breaking. Compared to the Camassa-Holm equation, the strong nonlinearity and nonlocal effects in the Novikov equation necessitates the need to work in a higher regularity space. Thanks to a higher order conservation law (2.7), which will serve as the "total energy", one can close all the estimates in $H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})$. This extra regularity enhancement also results in the exact conservation of the H^1 norm of solutions for all time, which is in contrast to the Camassa-Holm case, where there are still possible concentration of $u_x^2 dx$ [3, 4, 13]. In fact, the energy concentration in the Novikov equation occurs at the level of $u_x^4 dx$; see Section 8 for a more detailed discussion.

1.1. Main result.

Having established the existence and uniqueness of conservative solutions, our main goal here is to show that these solutions form a continuous semigroup. Because of the conservation laws (2.6)-(2.7), the $H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})$ norm seems to be a natural metric for the stability theory. However, the previous multi-peakon interaction reveals the opposite. From the result of [9, Theorem 1.1], for each time t, we introduce the measure ν_t whose absolutely continuous part with respect to Lebesgue measure has density $u^4 + 2u^2 u_x^2 - \frac{1}{3} u_x^4$. Consider a time t_* when a pair of peakons collide according to (1.7). Then as $t \to t_*^-$, ν_t converges weakly to some ν_{t_*} which contains a Dirac mass at the point q_*. The energy conservation implies that

$$\int_{\mathbb{R}} \left(u^2 + u_x^2\right)(t_*)\ dx = \lim_{t \to t_*^-} \int_{\mathbb{R}} \left(u^2 + u_x^2\right)(t)\ dx,$$

$$\int_{\mathbb{R}} \left(u^4 + 2u^2 u_x^2 - \frac{1}{3} u_x^4\right)(t_*)\ dx - \nu_{t_*}(\{q_*\}) = \lim_{t \to t_*^-} \int_{\mathbb{R}} \left(u^4 + 2u^2 u_x^2 - \frac{1}{3} u_x^4\right)(t)\ dx.$$

Therefore if we choose as from above a sequence of two-peakons of anti-strength u^ϵ defined by $u^\epsilon = u(t - \epsilon, x)$, then

$$\lim_{\epsilon \to 0} \left(\|u(0) - u^\epsilon(0)\|_{H^1} + \|u(0) - u^\epsilon(0)\|_{W^{1,4}}\right) = 0,$$

but

$$\begin{aligned}
\lim_{\epsilon \to 0} &\left(\frac{1}{3}\|u_x^\epsilon(t_*)\|_{L^4}^4 - \frac{1}{3}\|u_x(t_*)\|_{L^4}^4\right) \\
&= \nu_{t_*}(\{q_*\}) - \lim_{\epsilon \to 0} \int_{\mathbb{R}} \left[\left((u^\epsilon)^4 + 2(u^\epsilon)^2(u^\epsilon)_x^2\right) - \left(u^4 + 2u^2 u_x^2\right)\right](t_*)\ dx = \nu_{t_*}(\{q_*\}) > 0.
\end{aligned} \tag{1.8}$$

Therefore the flow is clearly discontinuous with respect to the $W^{1,4}$ norm. Numerical evidence verifying (1.8) is given in the Section 8.

A further message one can take from this example is that the main obstacle in establishing the Lipschitz continuous dependence on initial data lies in the concentration phenomenon of certain energy density due to possible focusing of wave fronts; see Figure 1.
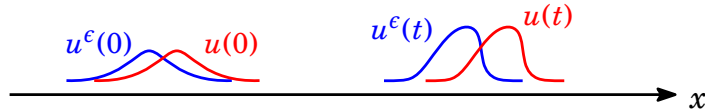

$u^\epsilon(0)$ $u(0)$ $u^\epsilon(t)$ $u(t)$ x

FIGURE 1. A sketch of the solution u to the Novikov equation and a small perturbation u^ϵ. Here u^ϵ is a backward shift of u. At time $t \to t_*^-$, both u_x^4 and $(u_x^\epsilon)^4$ approach to the delta functions, and hence the $W^{1,4}$ distance becomes large. A nature choice is a transport metric!

Similar situation occurs in a number of other quasi-linear equations, including the Hunter-Saxton equation, the Camassa-Holm equation, and the nonlinear variational wave equation. A class of optimal transport metrics was constructed to determine the minimum cost to transport an energy

measure from one solution to the other, such that the corresponding flow map remains uniformly Lipschitz on bounded time intervals for this new geodesic distance; see [2, 6, 7, 14, 15]. Compared with the previous literature, the major difficulty in dealing with the Novikov equation lies in the fact that equation (1.2) exhibits cubic nonlinearity, which requires one to work in a stronger regularity space, and therefore changes the energy concentration nature. Hence one needs to accordingly adjust terms in the metric of the infinitesimal tangent vectors. On the other hand, two complicated non-local convolution terms make the Lipschitz estimate fairly subtle.

The main theorem in this paper can be stated as:

Theorem 1.1. *The geodesic distance d, defined in Definition 5.3, renders Lipschitz continuous the flow generated by the equation (2.1). In particular, consider two initial data $u_0(x)$ and $\tilde{u}_0(x)$ which are absolute continuous and belong to $H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})$. Then for any $T > 0$, the corresponding solutions $u(t,x)$ and $\tilde{u}(t,x)$ satisfy*

$$d\big(u(t), \tilde{u}(t)\big) \leq C \ d(u_0, \tilde{u}_0),$$

when $t \in [0,T]$, where the constant C depends only on T and $H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})$–norm of initial data.

1.2. **Structure of the article.** We construct the metric d and prove that the solution map is Lipschitz continuous under this metric in several steps.

1. We construct a Lipschitz metric for smooth solution (Sections 3).
2. We prove by an application of the Thom's Transversality theorem that the piecewise smooth solutions with only generic singularities are dense in $H^1 \cap W^{1,4}$ (Section 4).
3. We first extend the Lipschitz metric to piecewise smooth solutions with only generic singularities, and then to general weak solution by the "density" result established in Step 2 (Section 5).
4. We compare our metric with other metrics, such as the Sobolev metric and Kantorovich-Rubinstein or Wasserstein metric (Section 6).

A more detailed explanation is given below.

Step 1: Metric for smooth solutions. We begin, in Section 2, by collecting several important local and global conservation laws of equation (1.2), which motivates the definition of an energy conservative solution. Then we recall the main theorem on the existence and uniqueness of such solutions established in [9], cf. Theorem 2.1.

To keep track of the cost of the energy transportation, we are led to construct the geodesic distance. That is, for two given solution profiles $u(t)$ and $u^\epsilon(t)$, we consider all possible smooth deformations/paths $\gamma^t : \theta \mapsto u^\theta(t)$ for $\theta \in [0,1]$ with $\gamma^t(0) = u(t)$ and $\gamma^t(1) = u^\epsilon(t)$, and then measure the length of these paths through integrating the norm of the tangent vector $d\gamma^t/d\theta$; see Figure 2 (a). The distance between u and u^ϵ will be calculated by the optimal path length

$$d\left(u(t), u^\epsilon(t)\right) = \inf_{\gamma^t} \|\gamma^t\| := \inf_{\gamma^t} \int_0^1 \|v^\theta(t)\|_{u^\theta(t)} \ d\theta, \quad \text{where } v^\theta(t) = \frac{d\gamma^t}{d\theta}.$$

Here the subscript $u^\theta(t)$ emphasizes the dependence of the norm on the flow u.

The next step is to define a Finsler norm for an infinitesimal tangent vector, by measuring the cost (with energy density μ) in shifting from one solution to the other one. From Figure 3, in measuring the cost of transporting u to u^ϵ on the x-u plane, we notice that the tangent flow v only measures the vertical displacement (dashed arrow) between two solutions. In order to provide enough freedom for planar transports, one needs to add a quantity, named as w, to measure the (horizontal) shift on x. It is nature to consider both vertical and horizontal shifts (solid arrows) in the energy space to estimate the cost of transport.

(a) Method of homotopy   (b) Loss of regularity
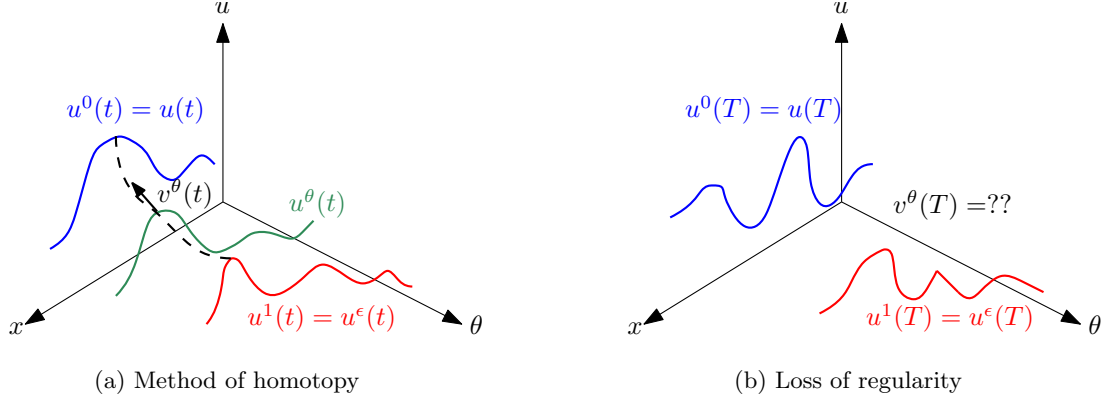
FIGURE 2. Compare two solutions $u(x)$ and $u^\epsilon(x)$ at a given time t.

(a)   (b)
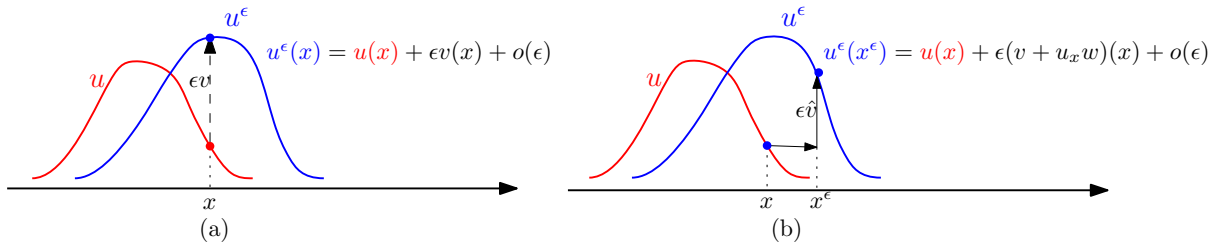
FIGURE 3. A sketch of how to deform from u to u^ϵ: (a) a vertical shift ϵv; (b) a horizontal shift $\epsilon u_x w$ followed by a vertical displacement ϵv. We denote the total shift as $\epsilon \hat{v} = \epsilon(v + u_x w)$ as in (1.10). Here $x^\epsilon := x + \epsilon w(x)$.

With this in mind,, the cost function basically includes

$$\|(w, \hat{v})\|_u = \int_{\mathbb{R}} \{[\text{horizontal shift}] + [\text{vertical shift}] + [\text{change in } u]\}\, d\mu$$
$$+ \int_{\mathbb{R}} [\text{change in the base measure with density } \mu]\, dx, \tag{1.9}$$

where, for example, as shown in Figure 3, it follows from Taylor's expansion that

$$[\text{change in } u] = o(\epsilon) \text{ order of } (u^\epsilon(x^\epsilon) - u(x)) = \hat{v} = v(x) + u_x(x)w(x). \tag{1.10}$$

A more detailed description is given in Section 3.

This way for an infinitesimal tangent vector v, we can accordingly define its Finsler norm to be

$$\|v\|_u = \inf_{w \in \mathcal{A}} \|(w, \hat{v})\|_u, \quad \text{with} \quad \hat{v} = v + u_x w,$$

where the admissible set $\mathcal{A}$ (defined in (3.7)) involves information of the characteristics of the equation that helps select the "reasonable" transports.

The explicit definition of the Finsler norm is provided in (3.8) and (3.9). The next thing we do is to investigate how this norm of the tangent vector propagates in time along any solution. The key estimate on the growth of the norm of the tangent vector is given in Lemma 3.1.

The implication of Lemma 3.1 is that, for any $T > 0$, if all solutions $u^\theta(t, \cdot)$ are sufficiently regular for $t \in [0, T]$, such that the tangent vectors are well-defined on γ^t, then it holds that

$$\|\gamma^t\| \le C\|\gamma^0\| \quad \forall\, t \in [0, T], \tag{1.11}$$

where C only depends on the total energy of the initial data. Therefore it is natural to define the geodesic distance

$$d\left(u, u^\epsilon\right) := \inf\left\{\|\gamma^t\|; \ \gamma^t: \ [0,1] \to H^1 \cap W^{1,4}, \ \gamma^t(0) = u, \ \gamma^t(1) = u^\epsilon\right\},$$

and one can expect from (1.11) that the solution map is Lipschitz continuous under this metric.

Step 2: Generic regularity. However, smooth solutions do not always remain smooth for all time. In fact for the Novikov equation (1.2), the gradient of the solution can blow up in finite time, and therefore a smooth path of initial data γ^0 may lose regularity at later time T so that the tangent vector $d\gamma^T/dt$ may not be well-defined (see Figure 2 (b)); even if it does exist, it is not obvious that the estimate (1.11) should remain valid. The idea is to show that there are sufficiently many paths which are initially smooth, and remain (piecewise) smooth later in time. We prove in Section 4

Theorem 1.2. *Let $T > 0$ be given, then there exists an open dense set*

$$\mathcal{D} \subset \left(\mathcal{C}^3(\mathbb{R}) \cap H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})\right),$$

such that, for $u_0 \in \mathcal{D}$, the conservative solution $u = u(t,x)$ of (2.1) is twice continuously differentiable in the complement of finitely many characteristic curves, within the domain $[0,T] \times \mathbb{R}$.

The generic regularity is itself of great interest because it shows very detailed structures of singularities. As a consequence, in Corollary 4.1, we prove the existence of regular paths connecting any two solutions of Theorem 1.2.

Similar generic regularity result was first established for the variational wave equation [1] and used in establishing a Lipschitz metric in [2]. This idea was applied to the Camassa-Holm equation in [20]. The proof of Theorem 1.2 relies on an application of Thom's Transversality Theorem.

Step 3: Metric for general weak solutions. For general weak solutions, we first use the semi-linear system established in [9] to extend the Lipschitz metric from smooth solutions to piecewise smooth solutions, cf. Theorem 5.1, Definition 5.1 and Definition 5.2. In the second step, we generalize this metric on the space $H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})$ and prove the Lipschitz property. This step is now very nature because the set of regular paths, that are roughly speaking piecewise smooth solutions, is a dense set of the solution space.

Step 4: Compare with other metrics. See details in Section 6.

Remark 1.1. *Currently, there are several parallel techniques in constructing the Lipschitz metrics. These include: the Finsler distance used in this paper (first established in [1, 2] for variational wave equation); method in [14, 15] for Camassa-Holm equation, which uses semi-linear equations on Lagrangian coordinates; and a direct transport method in [6] combing a convergence argument from multi-peakon solutions.*

Notice that, since we mainly work on the original coordinates and do not need the existence of sufficiently many special solutions such as multi-peakon solutions, our way to construct the metric is quite transparent and robust, as long as one has good knowledge about the energy concentration. In Section 7, we also apply this method to the Camassa-Holm equation, and the computation for the key estimates is greatly simplified.

We also point out at the end of the article that the semi-linear system introduced in [9] can be successfully implemented to simulate the peakon interactions. This is usually considered to be a quite challenging problem. Most of the current methods treat directly the ODEs governing the dynamics of peakons. Yet to the authors' knowledge, no schemes have been proposed to realize the energy concentration exactly at the collision time. This is largely due to the blow-up of the momentum density when two peakons collide. Here we take advantage of the semi-linear system in [9] which makes use of a set of new variables that dilates the interacting characteristics and in particular, allows one to trace the behavior of peakons even at the instant of collision. We present

in Section 8 several numerical experiments to both showcase explicitly the energy concentration of a peakon-antipeakon interaction, and the merits of using our scheme. Our method can be easily employed to other integrable system to analyze the peakon interactions.

2. Preliminary results

In this section we recall some useful estimates and the existence and uniqueness of global conservative solution of the Novikov equation. For more details the reader can refer to [9].

First we rewrite the Cauchy problem for the Novikov equation (1.2) as

$$\begin{cases} u_t + u^2 u_x + \partial_x P_1 + P_2 = 0, \\ u(0, x) = u_0(x), \end{cases} \tag{2.1}$$

where

$$P_1 := p * \left(\frac{3}{2} u u_x^2 + u^3 \right), \quad P_2 := \frac{1}{2} p * \left(u_x^3 \right), \tag{2.2}$$

with $p = \frac{1}{2} e^{-|x|}$ being the Green's function for $(1 - \partial_x^2)^{-1}$ on $\mathbb{R}$.

For smooth solutions, differentiating equation (2.1) with respect to x we have

$$u_{xt} + u^2 u_{xx} + \frac{1}{2} u u_x^2 - u^3 + P_1 + \partial_x P_2 = 0. \tag{2.3}$$

Further manipulation leads to the following local conservation laws

$$\left(\frac{u^2 + u_x^2}{2} \right)_t + \left(\frac{u^2 u_x^2}{2} + u P_1 + u \partial_x P_2 \right)_x = 0, \tag{2.4}$$

$$\left(u^4 + 2u^2 u_x^2 - \frac{1}{3} u_x^4 \right)_t + \left(2u^4 u_x^2 - \frac{1}{3} u^2 u_x^4 + \frac{4}{3} u^3 (P_1 + \partial_x P_2) \right)_x$$
$$+ \frac{4}{3} \left((P_1 + \partial_x P_2)^2 - (P_2 + \partial_x P_1)^2 \right)_x = 0, \tag{2.5}$$

which indicates two conserved quantities

$$\mathcal{E}(t) = \int_{\mathbb{R}} (u^2 + u_x^2)(t, x) \, dx = \mathcal{E}(0), \tag{2.6}$$

$$\mathcal{F}(t) = \int_{\mathbb{R}} (u^4 + 2u^2 u_x^2 - \frac{1}{3} u_x^4)(t, x) = \mathcal{F}(0) \, dx. \tag{2.7}$$

Therefore we can bound

$$\|u\|_{L^\infty}^2 \leq \|u\|_{H^1}^2 = \mathcal{E}(0),$$

$$\|u_x\|_{L^4}^4 = 3 \int_{\mathbb{R}} (u^4 + 2u^2 u_x^2) \, dx - 3\mathcal{F}(t) \leq 3 \left(2\mathcal{E}(0)^2 - \mathcal{F}(0) \right),$$

which in turn implies that

$$\|u_x\|_{L^3}^3 \leq \sqrt{3\mathcal{E}(0) \left(2\mathcal{E}(0)^2 - \mathcal{F}(0) \right)} =: K. \tag{2.8}$$

Now, we are able to bound $P_i(t)$ and the derivatives ∂P_i for $i = 1, 2$ as follows.

$$
\begin{aligned}
\|P_1(t)\|_{L^\infty}, \|\partial_x P_1(t)\|_{L^\infty} &\leq \|p\|_{L^\infty}\|\frac{3}{2}uu_x^2 + u^3\|_{L^1} \leq \frac{3}{4}\mathcal{E}(0)^{\frac{3}{2}}, \\
\|P_1(t)\|_{L^2}, \|\partial_x P_1(t)\|_{L^2} &\leq \|p\|_{L^2}\|\frac{3}{2}uu_x^2 + u^3\|_{L^1} \leq \frac{3}{2\sqrt{2}}\mathcal{E}(0)^{\frac{3}{2}}, \\
\|P_2(t)\|_{L^\infty}, \|\partial_x P_2(t)\|_{L^\infty} &\leq \frac{1}{2}\|p\|_{L^\infty}\|u_x^3\|_{L^1} \leq \frac{1}{4}K, \\
\|P_2(t)\|_{L^2}, \|\partial_x P_2(t)\|_{L^2} &\leq \frac{1}{2}\|p\|_{L^2}\|u_x^3\|_{L^1} \leq \frac{1}{2\sqrt{2}}K.
\end{aligned}
\tag{2.9}
$$

Now we state the theorem on the existence and uniqueness of conservative solutions to the Cauchy problem (2.1).

Theorem 2.1 ([9]). *Let $u_0 \in H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})$ be an absolute continuous function. Then the Cauchy problem (2.1) admits a unique energy **conservative** solution $u(t,x)$ defined for all $(t,x) \in \mathbb{R}^+ \times \mathbb{R}$ in the following sense.*

(i) For any fixed $t \geq 0$, $u(t,\cdot) \in H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})$. The map $t \to u(t,\cdot)$ is Lipschitz continuous under the $L^4(\mathbb{R})$ metric.

(ii) The solution $u = u(t,x)$ satisfies the initial condition of (2.1) in $L^4(\mathbb{R})$, and

$$
\iint_\Lambda \left\{ -u_x \left(\phi_t + u^2 \phi_x \right) + \left(-\frac{3}{2}uu_x^2 - u^3 + P_1 + \partial_x P_2 \right) \phi \right\} dx\, dt + \int_{\mathbb{R}} u_{0,x}\phi(0,x)\, dx = 0 \tag{2.10}
$$

for every test function $\phi \in C_c^1(\Lambda)$ with $\Lambda = \left\{ (t,x) \,\middle|\, t \in [0,\infty), x \in \mathbb{R} \right\}$.

(iii) There exists a family of Radon measures $\{\mu_t,\ t \in \mathbb{R}^+\}$, depending continuously on time and w.r.t the topology of weak convergence of measures. For every $t \in \mathbb{R}^+$, the absolutely continuous part of μ_t with respect to the Lebesgue measure has density $u_x^4(t,\cdot)$, which provides a measure-valued solution to the balance law

$$
\int_{\mathbb{R}^+} \left\{ \int (\phi_t + u^2\phi_x)d\mu_t + \int \left(4u^3u_x^3 - 4u_x^3(P_1 + \partial_x P_2) \right)\phi\, dx \right\} dt - \int_{\mathbb{R}} u_{0,x}^4\phi(0,x)dx = 0, \tag{2.11}
$$

for every test function $\phi \in C_c^1(\Lambda)$.

Moreover, the solution also satisfies the following properties.

(1) $u(t,x)$ is Hölder continuous with exponent $3/4$ on both t and x.

(2) The first energy density $u^2 + u_x^2$ is conserved for any time $t \geq 0$, i.e.

$$
\mathcal{E}(t) = \|u(t)\|_{H^1}^2 = \|u_0\|_{H^1}^2 \quad \text{for any} \quad t \geq 0; \tag{2.12}
$$

(3) The second energy density $u^4 + 2u^2u_x^2 - \frac{1}{3}u_x^4$ is conserved in the following sense.

(a) An energy inequality is satisfied:

$$
\mathcal{F}(t) = \int_{\mathbb{R}} \left(u^4 + 2u^2u_x^2 - \frac{1}{3}u_x^4 \right)(t,x)\, dx \geq \mathcal{F}(0) \quad \text{for any} \quad t \geq 0. \tag{2.13}
$$

(b) Denote a family of Radon measures $\{\nu_t,\ t \in \mathbb{R}^+\}$, such that

$$
\nu_t(\mathcal{A}) = \int_{\mathcal{A}} \left(u^4 + 2u^2u_x^2 \right)(t,x)\, dx - \frac{1}{3}\mu_t(\mathcal{A})
$$

for any Lebesgue measurable set $\mathcal{A}$ in $\mathbb{R}$. Then for any $t \in \mathbb{R}^+$,

$$
\nu_t(\mathbb{R}) = \nu_0(\mathbb{R}) = \mathcal{F}(0) = \int_{\mathbb{R}} \left(u^4 + 2u^2u_x^2 - \frac{1}{3}u_x^4 \right)(0,x)\, dx.
$$

For any $t \in \mathbb{R}^+$, the absolutely continuous part of ν_t with respect to Lebesgue measure has density $u^4 + 2u^2 u_x^2 - \frac{1}{3} u_x^4$. For almost every $t \in \mathbb{R}^+$, the singular part of ν_t is concentrated on the set where $u = 0$.

(4) *A continuous dependence result holds. Consider a sequence of initial data u_{0n} such that $\|u_{0n} - u_0\|_{H^1 \cap W^{1,4}} \to 0$, as $n \to \infty$. Then the corresponding solutions $u_n(t,x)$ converge to $u(t,x)$ uniformly for (t,x) in any bounded sets.*

3. THE NORM OF TANGENT VECTOR FOR SMOOTH SOLUTIONS

To illustrate the ideas on how to construct the Lipschitz metric, in this section we first consider smooth solutions to (2.1). We take a family of perturbed solutions $u^\epsilon(x)$ to (2.1), which can be written as

$$u^\epsilon(x) = u(x) + \epsilon v(x) + o(\epsilon). \tag{3.1}$$

Here in this section, we make an abuse of notation of using $f(t)$ or $f(x)$ to denote a function $f(t,x)$.

A straightforward calculation yields that the first order perturbation v satisfies

$$
\begin{aligned}
v_t + u^2 v_x + 2uvu_x &+ \frac{1}{2} \left(\int_x^\infty - \int_{-\infty}^x \right) e^{-|x-y|} \left(3u_y v_y u + \frac{3}{2} u_y^2 v + 3u^2 v \right) dy \\
&+ \frac{3}{4} \int_{-\infty}^\infty e^{-|x-y|} u_y^2 v_y \, dy = 0.
\end{aligned}
\tag{3.2}
$$

Differentiating (3.2) with respect to x, one obtains

$$
\begin{aligned}
v_{xt} + u^2 v_{xx} + uu_x v_x &+ \frac{1}{2} u_x^2 v + 2uvu_{xx} - 3u^2 v + \frac{1}{2} \int_\mathbb{R} e^{-|x-y|} \left(3u_y v_y u + \frac{3}{2} u_y^2 v + 3u^2 v \right) dy \\
&+ \frac{3}{4} \left(\int_x^\infty - \int_{-\infty}^x \right) e^{-|x-y|} u_y^2 v_y \, dy = 0.
\end{aligned}
\tag{3.3}
$$

As explained in Section 1.2, to obtain enough freedom in measuring the shift from one solution to the other one, we need to add a quantity $w(t,x)$ measuring the horizontal shift:

$$x^\epsilon := x + \epsilon w(x) + o(\epsilon). \tag{3.4}$$

To focus only on reasonable transports between two solutions, we select $w(t,x)$ by propagating along characteristics the shifts $w_0(x)$ as the initial data. That is, we require that when $x(t)$ is a characteristic emanating from x_0 then $x^\epsilon(t)$ is also a characteristic emanating from x_0^ϵ, so

$$\frac{d}{dt} x^\epsilon(t) = (u^\epsilon(x^\epsilon))^2 \qquad \text{when} \qquad \frac{d}{dt} x(t) = u^2(x).$$

Thus, using (3.1), (3.4) and taking the limit as $\epsilon \to 0$, we have

$$w_t + u^2 w_x = 2u(v + u_x w). \tag{3.5}$$

So the Finsler norm on the space of infinitesimal tangent vector v takes the form

$$\|v\|_u := \inf_{w \in \mathcal{A}} \|(w, \hat{v})\|_u, \quad \text{with} \quad \hat{v} = v + u_x w, \tag{3.6}$$

where the admissible set is defined as

$$\mathcal{A} = \{\text{solutions } w(t,x) \text{ of } (3.5) \text{ with smooth initial data } w_0(x)\}. \tag{3.7}$$

Note that u is always smooth in this section, hence in (3.5) w can be solved with a given initial data $w(0,x) = w_0(x)$.

To motivate the explicit construction of $\|(w,v)\|_u$, we consider a reference solution u together with a perturbation u^ϵ, as shown in Figure 3. Recall that we are interested in determining the cost

of transporting the energy (with density $\mu \approx u_x^4$) from u to u^ϵ. In fact we will choose $\mu = (1 + u_x^2)^2$, and the cost should account for the following

$$
\begin{aligned}
\|(w, \hat{v})\|_u &= \int_{\mathbb{R}} \left\{[\text{change in } x] + [\text{change in } u] + [\text{change in } \arctan u_x]\right\} (1 + u_x^2)^2 e^{-|x|} \, dx \\
&\quad + \int_{\mathbb{R}} [\text{change in the base measure with density } (1 + u_x^2)^2] e^{-|x|} \, dx \\
&=: I_1 + I_2 + I_3 + I_4 \,.
\end{aligned}
\tag{3.8}
$$

More precisely, we have

$$
\begin{aligned}
\|v\|_u &= \inf_{w \in \mathcal{A}} \int_{\mathbb{R}} \left\{|w|(1 + u_x^2)^2 + |v + u_x w|(1 + u_x^2)^2 + |v_x + u_{xx} w|(1 + u_x^2) \right. \\
&\qquad\qquad\quad \left. + |4(u_x + u_x^3)(v_x + u_{xx} w) + (1 + u_x^2)^2 w_x|\right\} e^{-|x|} \, dx \\
&=: \inf_{w \in \mathcal{A}} (I_1 + I_2 + I_3 + I_4) \,.
\end{aligned}
\tag{3.9}
$$

We note that when $u \in L^1$, the $e^{-|x|}$ term is not necessary. Below we briefly explain how to obtain (3.9), using (3.1) and (3.4).

- For [change in x] in I_1,
$$
\frac{1}{\epsilon}[x^\epsilon - x] = w + o(\epsilon).
$$

- For [change in u] in I_2,
$$
\frac{1}{\epsilon}[u^\epsilon(x^\epsilon) - u(x)] = v(x) + u_x(x)w(x) + o(\epsilon).
$$

- For [change in $\arctan u_x$] in I_3,
$$
\begin{aligned}
\arctan u_x^\epsilon(x^\epsilon) &= \arctan\left[u_x(x^\epsilon) + \epsilon v_x(x^\epsilon) + o(\epsilon)\right] \\
&= \arctan\left[u_x(x) + \epsilon w(x)u_{xx}(x) + \epsilon v_x(x) + o(\epsilon)\right] \\
&= \arctan u_x(x) + \epsilon \frac{v_x(x) + w(x)u_{xx}(x)}{1 + u_x^2(x)} + o(\epsilon).
\end{aligned}
$$

- For [change in the base measure with density $(1 + u_x^2)^2$] in I_4, using the following identities
$$
\begin{aligned}
(u_x^\epsilon(x^\epsilon))^2 &= u_x^2(x^\epsilon) + 2\epsilon u_x(x^\epsilon)v_x(x^\epsilon) + o(\epsilon) \\
&= u_x^2(x) + 2\epsilon w(x)u_x(x)u_{xx}(x) + 2\epsilon v_x(x)u_x(x) + o(\epsilon),
\end{aligned}
$$

we obtain that

$$
\left(1 + (u_x^\epsilon(x^\epsilon))^2\right)^2 dx^\epsilon - \left(1 + u_x^2(x)\right)^2 dx = \left(4\epsilon(v_x + w u_{xx})(u_x + u_x^3) + \epsilon w_x(1 + u_x^2)^2 + o(\epsilon)\right) dx.
$$

Next we want to understand how this norm of the tangent vector changes in time. The main result of this section is the following.

Lemma 3.1. *Let $T > 0$ be given, and $u(t, x)$ be a smooth solution to (2.1) when $t \in [0, T]$. Assume that the first order perturbation v satisfies equation (3.2). Then it follows that*

$$
\|v(t)\|_{u(t)} \leq \mathcal{C}(T)\|v(0)\|_{u(0)},
\tag{3.10}
$$

for some constant $\mathcal{C}(T)$ depending only on initial total energy $\mathcal{E}(0)$, $\mathcal{F}(0)$ and T.

Proof. It suffices to show that

$$
\frac{d}{dt}\|(w(t), v(t))\|_{u(t)} \leq C \|(w(t), v(t))\|_{u(t)} \,,
\tag{3.11}
$$

for any v and w satisfying (3.2), (3.3) and (3.5). Here C is a constant depending only on the initial energy $\mathcal{E}(0)$ and $\mathcal{F}(0)$. In rest of this paper, unless specified, we will use C to denote a constant depending on the initial energy $\mathcal{E}(0)$, $\mathcal{F}(0)$.

To prove (3.11), first, notice that for any smooth function f, we have

$$\frac{d}{dt} \int_{\mathbb{R}} |f| e^{-|x|} \, dx$$
$$= \int_{\mathbb{R}} (|f| e^{-|x|})_t + (u^2 |f| e^{-|x|})_x \, dx$$
$$= \int_{\mathbb{R}} \text{sign}(f)[f_t + (u^2 f)_x] e^{-|x|} \, dx - \int_0^\infty u^2 |f| e^{-x} dx + \int_{-\infty}^0 u^2 |f| e^x dx$$
$$\leq \int_{\mathbb{R}} |f_t + (u^2 f)_x| e^{-|x|} \, dx + \|u\|_{L^\infty}^2 \int_{\mathbb{R}} |f| e^{-|x|} dx.$$

Therefore in the following we only have to show that $|f_t + (u^2 f)_x| e^{-|x|}$ is bounded by the integrands in $I_1 + I_2 + I_3 + I_4$, with f being $w(1 + u_x^2)^2$, $(v + u_x w)(1 + u_x^2)^2$, $(v_x + u_{xx} w)(1 + u_x^2)$ and $4(u_x + u_x^3)(v_x + u_{xx} w) + (1 + u_x^2)^2 w_x$, respectively.

To simplify the computation, we drop all the $e^{-|x|}$ terms (this can be thought of as assuming that the solution has compact support). In the general case, the same results remain valid when the factor $e^{-|x|}$ is inserted back.

1. For I_1 in (3.9), it follows from a direct computation that

$$\left((1 + u_x^2)^2\right)_t + \left(u^2(1 + u_x^2)^2\right)_x = (u_x^3 + u_x)(4u^3 - 4(P_1 + \partial_x P_2) + 2u). \tag{3.12}$$

By (3.5) and (3.12), we have

$$\left(w(1 + u_x^2)^2\right)_t + \left(u^2 w(1 + u_x^2)^2\right)_x$$
$$= (w_t + u^2 w_x)(1 + u_x^2)^2 + w\left[\left((1 + u_x^2)^2\right)_t + \left(u^2(1 + u_x^2)^2\right)_x\right] \tag{3.13}$$
$$= 2u(v + u_x w)(1 + u_x^2)^2 + w(u_x^3 + u_x)\left(4u^3 - 4(P_1 + \partial_x P_2) + 2u\right).$$

This in turn yields the estimate

$$\frac{dI_1}{dt} = \frac{d}{dt} \int_{\mathbb{R}} |w|(1 + u_x^2)^2 \, dx \leq C \int_{\mathbb{R}} |v + u_x w|(1 + u_x^2)^2 \, dx + C \int_{\mathbb{R}} |w|(1 + u_x^2)^2 \, dx$$
$$\leq C(I_1 + I_2). \tag{3.14}$$

2. The estimates for the second and third terms are much more delicate. For I_2, from (2.3), (3.2), (3.5) and (3.12), we obtain

$$\left((v + u_x w)(1 + u_x^2)^2\right)_t + \left(u^2(v + u_x w)(1 + u_x^2)^2\right)_x$$

$$= \left[v_t + u^2 v_x + u_x(w_t + u^2 w_x) + w\left(u_{xt} + u^2 u_{xx}\right)\right](1 + u_x^2)^2$$

$$+ (v + u_x w)\left[\left((1 + u_x^2)^2\right)_t + \left(u^2(1 + u_x^2)^2\right)_x\right]$$

$$= \left[-\frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right)e^{-|x-y|}\left(3u_y v_y u + \frac{3}{2}u_y^2 v + 3u^2 v\right)dy - \frac{3}{4}\int_{-\infty}^\infty e^{-|x-y|}u_y^2 v_y\, dy\right.$$

$$\left. -2uvu_x + 2uu_x(v + u_x w) + w\left(-\frac{uu_x^2}{2} + u^3 - P_1 - \partial_x P_2\right)\right](1 + u_x^2)^2$$

$$+ (v + u_x w)(u_x^3 + u_x)\left(4u^3 - 4(P_1 + \partial_x P_2) + 2u\right)$$

$$= \left[-\frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right)e^{-|x-y|}\left(3u_y v_y u + \frac{3}{2}u_y^2 v + 3u^2 v\right)dy - \frac{3}{4}\int_{-\infty}^\infty e^{-|x-y|}u_y^2 v_y\, dy\right.$$

$$\left. +\frac{3}{2}uu_x^2 w + w(u^3 - P_1 - \partial_x P_2)\right](1 + u_x^2)^2 + (v + u_x w)(u_x^3 + u_x)(4u^3 - 4(P_1 + \partial_x P_2) + 2u)$$

$$=: (I_{21} + I_{22} + I_{23} + I_{24})(1 + u_x^2)^2 + (v + u_x w)(u_x^3 + u_x)(4u^3 - 4(P_1 + \partial_x P_2) + 2u). \tag{3.15}$$

For the local term $I_{23} = \frac{3}{2}uu_x^2 w(x)$, we have

$$I_{23} = -\frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right)\left(e^{-|x-y|}\frac{3}{2}uu_y^2 w\right)_y dy$$

$$= -\frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right)e^{-|x-y|}\left(3uu_y u_{yy}w + \frac{3}{2}u_y^3 w + \frac{3}{2}uu_y^2 w_y\right)dy$$

$$+ \frac{3}{4}\int_{-\infty}^\infty e^{-|x-y|}uu_y^2 w\, dy.$$

Hence

$$I_{21} + I_{23}$$

$$= -\frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right)e^{-|x-y|}\left[3uu_y(v_y + u_{yy}w) + \frac{3}{2}uu_y^2 w_y + \frac{3}{2}u_y^2(v + u_y w) + 3u^2 v\right]dy$$

$$+ \frac{3}{4}\int_{-\infty}^\infty e^{-|x-y|}uu_y^2 w\, dy$$

$$= -\frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right)e^{-|x-y|}\left\{\frac{3}{2}u\left[4u_y(v_y + u_{yy}w) + (1 + u_y^2)w_y\right] - 3uu_y(v_y + u_{yy}w)\right.$$

$$\left. -\frac{3}{2}(uw)_y + \frac{3}{2}u_y w + \frac{3}{2}u_y^2(v + u_y w) + 3u^2 v\right\}dy + \frac{3}{4}\int_{-\infty}^\infty e^{-|x-y|}uu_y^2 w\, dy$$

$$= -\frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right)e^{-|x-y|}\left\{\frac{3}{2}u\left[4u_y(v_y + u_{yy}w) + (1 + u_y^2)w_y\right] - 3uu_y(v_y + u_{yy}w)\right.$$

$$\left. +\frac{3}{2}u_y w + \frac{3}{2}u_y^2(v + u_y w) + 3u^2 v\right\}dy - \frac{3}{4}\left(\int_x^\infty - \int_{-\infty}^x\right)\frac{\partial e^{-|x-y|}}{\partial y}uw\, dy$$

$$-\frac{3}{2}uw(x) + \frac{3}{4}\int_{-\infty}^\infty e^{-|x-y|}uu_y^2 w\, dy$$

$$= -\frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right)e^{-|x-y|}\left\{\frac{3}{2}u\left[4u_y(v_y + u_{yy}w) + (1 + u_y^2)w_y\right] - 3uu_y(v_y + u_{yy}w)\right.$$

$$+ \frac{3}{2} u_y w + \frac{3}{2} u_y^2 (v + u_y w) + 3u^2 v \bigg\} \, dy + \frac{3}{4} \int_{-\infty}^{\infty} e^{-|x-y|} (uw + uu_y^2 w) \, dy - \frac{3}{2} uw(x).$$

Note that

$$3u^2 v(y) = 3u^2(v + u_y w) - 3u^2 u_y w,$$

which, together with the previous equality and the Sobolev inequality, implies that

$$\left| \int_{\mathbb{R}} (I_{21} + I_{23})(1 + u_x^2)^2 \, dx \right|$$

$$\leq C \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-|x-y|} (1 + u_x^2)^2 \, dx \right) \left[\left| \frac{3}{2} u \left(4u_y(v_y + u_{yy}w) + (1 + u_y^2)w_y \right) \right| \right.$$

$$+ |3uu_y(v_y + u_{yy}w)| + \frac{3}{2} |u_y w| + \frac{3}{2} |u_y^2 (v + u_y w)| + 3|u^2(v + u_y w)| + 3|u^2 u_y w|$$

$$\left. + \frac{3}{4} |uw| + \frac{3}{4} |uu_y^2 w| \right] \, dy + C \int_{\mathbb{R}} |w|(1 + u_x^2)^2 \, dx \tag{3.16}$$

$$\leq C \int_{\mathbb{R}} \left| 4u_x(v_x + u_{xx}w) + (1 + u_x^2)w_x \right| (1 + u_x^2) \, dx + C \int_{\mathbb{R}} |w|(1 + u_x^2)^2 \, dx$$

$$+ C \int_{\mathbb{R}} |v_x + u_{xx}w|(1 + u_x^2) \, dx + C \int_{\mathbb{R}} |v + u_x w|(1 + u_x^2)^2 \, dx$$

$$\leq C(I_4 + I_1 + I_3 + I_2),$$

where we have used the fact that

$$\int_{\mathbb{R}} e^{-|x-y|} (1 + u_x^2)^2 \, dx \leq \int_{\mathbb{R}} e^{-|x-y|} \, dx + 2 \int_{\mathbb{R}} u_x^2 \, dx + \int_{\mathbb{R}} u_x^4 \, dx \leq C.$$

For the second term I_{22}, we have

$$3u_y^2 v_y = 3u_y^2(v_y + u_{yy}w) - 3u_y^2 u_{yy}w = 3u_y^2(v_y + u_{yy}w) - (u_y^3 w)_y + u_y^3 w_y.$$

The last term can be bounded as

$$u_y^3 w_y \leq |(1 + u_y^2)u_y w_y| \leq |4u_y^2(v_y + u_{yy}w) + (1 + u_y^2)u_y w_y| + |4u_y^2(v_y + u_{yy}w)|$$

$$\leq C|4u_y(v_y + u_{yy}w) + (1 + u_y^2)w_y|(1 + u_y^2) + |4u_y^2(v_y + u_{yy}w)|. \tag{3.17}$$

Then it holds that

$$\left| \frac{3}{4} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-|x-y|} u_y^2 v_y \, dy (1 + u_x^2)^2 \, dx \right|$$

$$\leq C \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-|x-y|} (1 + u_x^2)^2 \, dx \right) \left(|u_y^2(v_y + u_{yy}w)| + |4u_y(v_y + u_{yy}w) + (1 + u_y^2)w_y|(1 + u_y^2) \right) \, dy$$

$$+ C \left| \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-|x-y|} (u_y^3 w)_y \, dy \right) (1 + u_x^2)^2 \, dx \right| \tag{3.18}$$

$$\leq C \int_{\mathbb{R}} |v_x + u_{xx}w|(1 + u_x^2) \, dx + C \int_{\mathbb{R}} \left| 4u_x(v_x + u_{xx}w) + (1 + u_x^2)w_x \right| (1 + u_x^2) \, dx$$

$$+ C \int_{\mathbb{R}} \int_{\mathbb{R}} \left| \frac{\partial e^{-|x-y|}}{\partial y} \right| (1 + u_x^2)^2 \, dx \cdot |u_y^3 w| \, dy$$

$$\leq C(I_3 + I_4 + I_1).$$

Plugging (3.16) and (3.18) into (3.15), we can conclude that

$$\frac{dI_2}{dt} = \frac{d}{dt} \int_{\mathbb{R}} |v + u_x w|(1 + u_x^2)^2 \, dx \leq C(I_1 + I_2 + I_3 + I_4). \tag{3.19}$$

3. To estimate the time derivative of I_3, using (2.3), (3.3) and (3.5) we obtain

$$
\begin{aligned}
&\big((v_x + u_{xx}w)(1 + u_x^2)\big)_t + \big(u^2(v_x + u_{xx}w)(1 + u_x^2)\big)_x \\
&= \big[v_{xt} + u^2 v_{xx} + u_{xx}(w_t + u^2 w_x) + w(u_{xxt} + u^2 u_{xxx})\big](1 + u_x^2) \\
&\quad + (v_x + u_{xx}w)\big[(1 + u_x^2)_t + \big(u^2(1 + u_x^2)\big)_x\big] \\
&= \Big[-uu_x v_x - \frac{1}{2}u_x^2 v - 2uvu_{xx} + 3u^2 v - \frac{1}{2}\int_{-\infty}^{\infty} e^{-|x-y|}\big(3u_y v_y u + \frac{3}{2}u_y^2 v + 3u^2 v\big)\,dy \\
&\qquad - \frac{3}{4}\Big(\int_x^{\infty} - \int_{-\infty}^x\Big)e^{-|x-y|}u_y^2 v_y\,dy + 2uu_{xx}(v + u_x w) + w(3u^2 u_x - 3uu_x u_{xx} \\
&\qquad - \partial_x P_1 - P_2)\Big](1 + u_x^2) + (v_x + u_{xx}w)\big[2u^3 u_x - 2u_x(P_1 + \partial_x P_2) + uu_x^3 + 2uu_x\big] \\
&= \Big[-\frac{1}{2}u_x^2 v + 3u^2(v + u_x w) - \frac{1}{2}\int_{-\infty}^{\infty} e^{-|x-y|}\big(3u_y v_y u + \frac{3}{2}u_y^2 v + 3u^2 v\big)\,dy \\
&\qquad - \frac{3}{4}\Big(\int_x^{\infty} - \int_{-\infty}^x\Big)e^{-|x-y|}u_y^2 v_y\,dy - w(\partial_x P_1 + P_2)\Big](1 + u_x^2) \\
&\quad + (v_x + u_{xx}w)\big[2u^3 u_x - 2u_x(P_1 + \partial_x P_2) + uu_x\big].
\end{aligned}
\tag{3.20}
$$

Note that

$$
\begin{aligned}
-\frac{1}{2}u_x^2 v &= -\frac{1}{2}u_x^2(v + u_x w) + \frac{1}{2}u_x^3 w \\
&= -\frac{1}{2}u_x^2(v + u_x w) - \frac{1}{2}\Big(\int_x^{\infty} - \int_{-\infty}^x\Big)\Big(e^{-|x-y|}\frac{1}{2}u_y^3 w(y)\Big)_y\,dy \\
&= -\frac{1}{2}u_x^2(v + u_x w) + \frac{1}{4}\int_{-\infty}^{\infty} e^{-|x-y|}u_y^3 w\,dy \\
&\quad - \frac{1}{4}\Big(\int_x^{\infty} - \int_{-\infty}^x\Big)e^{-|x-y|}\big(3u_y^2 u_{yy}w + u_y^3 w_y\big)\,dy.
\end{aligned}
$$

From the above inequality and (3.17), the first and fourth terms in the last equality of (3.20) can be estimated as

$$
\begin{aligned}
&\left|\int_{\mathbb{R}}\left\{-\frac{1}{2}u_x^2 v - \frac{3}{4}\Big(\int_x^{\infty} - \int_{-\infty}^x\Big)e^{-|x-y|}u_y^2 v_y(y)\,dy\right\}(1 + u_x^2)\,dx\right| \\
&= \left|-\frac{1}{2}\int_{\mathbb{R}}u_x^2(v + u_x w)(1 + u_x^2)\,dx + \frac{1}{4}\int_{\mathbb{R}}\int_{\mathbb{R}}e^{-|x-y|}u_y^3 w\,dy(1 + u_x^2)\,dx\right. \\
&\qquad \left.- \frac{1}{4}\int_{\mathbb{R}}\Big(\int_x^{\infty} - \int_{-\infty}^x\Big)e^{-|x-y|}\big[3u_y^2(v_y + u_{yy}w) + u_y^3 w_y\big]\,dy(1 + u_x^2)\,dx\right| \\
&\leq C\int_{\mathbb{R}}|v + u_x w|(1 + u_x^2)^2\,dx + \frac{1}{4}\int_{\mathbb{R}}\Big(\int_{\mathbb{R}}e^{-|x-y|}(1 + u_x^2)\,dx\Big)|u_y^3 w|\,dy \\
&\quad + C\int_{\mathbb{R}}\Big(\int_{\mathbb{R}}e^{-|x-y|}(1 + u_x^2)\,dx\Big)\big[|v_y + u_{yy}w|(1 + u_y^2) \\
&\qquad + |4u_y(v_y + u_{yy}w) + (1 + u_y^2)w_y|(1 + u_y^2) + |4u_y^2(v_y + u_{yy}w)|\big]\,dy \\
&\leq C\int_{\mathbb{R}}|v + u_x w|(1 + u_x^2)^2\,dx + C\int_{\mathbb{R}}|v_x + u_{xx}w|(1 + u_x^2)\,dx \\
&\quad + C\int_{\mathbb{R}}|w|(1 + u_x^2)^2\,dx + C\int_{\mathbb{R}}|4u_x(v_x + u_{xx}w) + (1 + u_x^2)w_x|(1 + u_x^2)\,dx \\
&\leq C(I_2 + I_3 + I_1 + I_4).
\end{aligned}
\tag{3.21}
$$

On the other hand, for the third term in the last equality of (3.20), we notice that

$$3u_y v_y u = 3uu_y(v_y + u_{yy}w) - 3uu_y u_{yy}w$$

$$= 3uu_y(v_y + u_{yy}w) - \frac{3}{2}\left(uu_y^2 w\right)_y + \frac{3}{2}u_y^3 w + \frac{3}{2}uu_y^2 w_y, \tag{3.22}$$

and

$$\frac{3}{2}uu_y^2 w_y \le \frac{3}{4}\left|u(1+u_y^2)u_y w_y\right|$$

$$\le \frac{3}{4}|u| \cdot \left|(1+u_y^2)u_y w_y + 4u_y^2(v_y + u_{yy}w)\right| + 3|u| \cdot \left|u_y^2(v_y + u_{yy}w)\right| \tag{3.23}$$

$$\le C|u| \cdot \left|(1+u_y^2)w_y + 4u_y(v_y + u_{yy}w)\right|(1+u_y^2) + 3|u| \cdot \left|u_y^2(v_y + u_{yy}w)\right|.$$

Thus, by (3.22), (3.23) and integration by parts, we bound the third term in the following way.

$$\frac{1}{2}\left|\int_{\mathbb{R}}\left(\int_{\mathbb{R}} e^{-|x-y|}\left(3u_y v_y u + \frac{3}{2}u_y^2 v + 3u^2 v\right)dy\right)(1+u_x^2)\,dx\right|$$

$$\le C\int_{\mathbb{R}}\left(\int_{\mathbb{R}} e^{-|x-y|}(1+u_x^2)\,dx\right)\left\{3|uu_y(v_y + u_{yy}w)| + 3|u||u_y^2(v_y + u_{yy}w)|\right.$$

$$+ |u|\left|(1+u_y^2)w_y + 4u_y(v_y + u_{yy}w)\right|(1+u_y^2) + \frac{3}{2}|u_y^2(v + u_y w)|$$

$$\left.+3|u^2(v + u_y w)| + 3|u^2 u_y w|\right\}\,dy + C\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\frac{\partial e^{-|x-y|}}{\partial y}\right|(1+u_x^2)\,dx \cdot |uu_y^2 w|\,dy$$

$$\le C\int_{\mathbb{R}}|v_x + u_{xx}w|(1+u_x^2)\,dx + C\int_{\mathbb{R}}|w|(1+u_x^2)^2\,dx \tag{3.24}$$

$$+ C\int_{\mathbb{R}}\left|4u_x(v_x + u_{xx}w) + (1+u_x^2)w_x\right|(1+u_x^2)\,dx + C\int_{\mathbb{R}}|v + u_x w|(1+u_x^2)^2\,dx$$

$$\le C(I_3 + I_1 + I_4 + I_2).$$

Plugging (3.21) and (3.24) into (3.20), it holds that

$$\frac{dI_3}{dt} = \frac{d}{dt}\int_{\mathbb{R}}|v_x + u_{xx}w|(1+u_x^2)\,dx \le C(I_1 + I_2 + I_3 + I_4). \tag{3.25}$$

4. To estimate the time derivative of I_4, differentiating (3.5) with respect to x we get

$$w_{tt} + u^2 w_{xx} = 2u_x(v + u_x w) + 2u(v_x + u_{xx}w). \tag{3.26}$$

Using (2.3), (3.3), (3.5), (3.12) and (3.26), it holds that

$$\left(4(u_x + u_x^3)(v_x + u_{xx}w) + (1+u_x^2)^2 w_x\right)_t + \left(u^2\left(4(u_x + u_x^3)(v_x + u_{xx}w) + (1+u_x^2)^2 w_x\right)\right)_x$$

$$= 4\left[(1+3u_x^2)(u_{xt} + u^2 u_{xx}) + 2uu_x^2(1+u_x^2))\right](v_x + u_{xx}w) + 4\left[v_{xt} + u^2 v_{xx} + u_{xx}(w_t + u^2 w_x)\right.$$

$$\left.+w(u_{xxt} + u^2 u_{xxx})\right](u_x + u_x^3) + (1+u_x^2)^2(w_{xt} + u^2 w_{xx}) + w_x\left[((1+u_x^2)^2)_t + \left(u^2(1+u_x^2)^2\right)_x\right]$$

$$= 4\left[(1+3u_x^2)\left(-\frac{1}{2}uu_x^2 + u^3 - P_1 - \partial_x P_2\right) + 2uu_x^2(1+u_x^2))\right](v_x + u_{xx}w)$$

$$+ 4(u_x + u_x^3)\left[-uu_x v_x - \frac{1}{2}u_x^2 v - 2uv u_{xx} + 3u^2 v - \frac{1}{2}\int_{\mathbb{R}} e^{-|x-y|}\left(3u_y v_y u + \frac{3}{2}u_y^2 v + 3u^2 v\right)dy\right.$$

$$-\frac{3}{4}\left(\int_x^\infty - \int_{-\infty}^x\right)e^{-|x-y|}u_y^2 v_y\,dy + 2uu_{xx}(v + u_x w) + w\left(3u^2 u_x - 3uu_x u_{xx} - \partial_x P_1 - P_2\right)\bigg]$$

$$+ (1+u_x^2)^2\left[2u_x(v + u_x w) + 2u(v_x + u_{xx}w)\right] + w_x(u_x^3 + u_x)\left[4u^3 - 4(P_1 + \partial_x P_2) + 2u\right]$$

$$= \left[6uu_x^2 + 2u + 4(1+3u_x^2)\left(u^3 - P_1 - \partial_x P_2\right)\right](v_x + u_{xx}w) + 2u_x(1+u_x^2)\left[v + u_x(1+u_x^2)w\right]$$

$$- 3(u_x + u_x^3) \left(\int_x^\infty - \int_{-\infty}^x \right) e^{-|x-y|} u_y^2 v_y \, dy + 4(u_x + u_x^3) \left[3u^2(v + u_x w) - w(\partial_x P_1 + P_2) \right.$$

$$\left. - \frac{1}{2} \int_{\mathbb{R}} e^{-|x-y|} (3u_y v_y u + \frac{3}{2} u_y^2 v + 3u^2 v) \, dy \right] + w_x(u_x^3 + u_x) \left[4u^3 - 4(P_1 + \partial_x P_2) + 2u \right].$$

For the second and third terms in the last equality, we have

$$2u_x(1 + u_x^2) \left[v + u_x(1 + u_x^2)w \right] - 3(u_x + u_x^3) \left(\int_x^\infty - \int_{-\infty}^x \right) e^{-|x-y|} u_y^2 v_y \, dy$$

$$= -u_x(1 + u_x^2) \left(\int_x^\infty - \int_{-\infty}^x \right) \left[\left(e^{-|x-y|} \left(v + u_y(1 + u_y^2)w \right) \right)_y + 3e^{-|x-y|} u_y^2 v_y \right] dy$$

$$= -u_x(1 + u_x^2) \left(\int_x^\infty - \int_{-\infty}^x \right) e^{-|x-y|} \left(v_y + u_{yy}w + 3u_y^2(v_y + u_{yy}w) + u_y(1 + u_y^2)w_y \right) dy$$

$$+ u_x(1 + u_x^2) \int_{\mathbb{R}} e^{-|x-y|} \left(v + u_y(1 + u_y^2)w \right) dy \tag{3.27}$$

$$= -u_x(1 + u_x^2) \left(\int_x^\infty - \int_{-\infty}^x \right) e^{-|x-y|} \left[v_y + u_{yy}w + 4u_y^2(v_y + u_{yy}w) + u_y(1 + u_y^2)w_y \right.$$

$$\left. - u_y^2(v_y + u_{yy}w) \right] dy + u_x(1 + u_x^2) \int_{\mathbb{R}} e^{-|x-y|} \left(v + u_y w + u_y^3 w \right) dy.$$

While for the last term,

$$w_x(u_x^3 + u_x)(4u^3 - 4(P_1 + \partial_x P_2) + 2u)$$

$$= \left(4u^3 - 4(P_1 + \partial_x P_2) + 2u \right) \left[w_x(u_x^3 + u_x) + 4u_x^2(v_x + u_{xx}w) - 4u_x^2(v_x + u_{xx}w) \right]. \tag{3.28}$$

Finally, the term $-2(u_x + u_x^3) \int_{-\infty}^\infty e^{-|x-y|} (3u_y v_y u + \frac{3}{2} u_y^2 v + 3u^2 v)(y) \, dy$ can be estimated in a similar way as is done to obtain (3.24). We thus conclude from all of the above that

$$\frac{dI_4}{dt} = \frac{d}{dt} \int_{\mathbb{R}} \left| 4(u_x + u_x^3)(v_x + u_{xx}w) + (1 + u_x^2)^2 w_x \right| dx \leq C(I_1 + I_2 + I_3 + I_4). \tag{3.29}$$

Combining the inequalities (3.14), (3.19), (3.25) and (3.29) together, we obtain the desired inequality (3.11), and hence (3.10). □

4. Generic regularity for the Novikov equation: Proof of Theorem 1.2.

In this section, we use the Thom's transversality Theorem to prove generic regularity result given in Theorem 1.2.

4.1. The semi-linear system on new coordinates.
As a start, we first briefly review the semi-linear system introduced in [9], which will be used later. Please find detail calculations and derivations in [9].

Following the idea in [9], we introduce new coordinates (t, Y), such that

$$Y \equiv Y(t, x) := \int_0^{x^c(0; t, x)} (1 + u_x^2(0, x'))^2 \, dx',$$

where $a \mapsto x^c(a; t, x)$ denotes a characteristic passing through the point (t, x). Here the equation of characteristic is

$$\frac{dx(t)}{dt} = u^2(t, x(t)).$$

In fact, $Y = Y(t, x)$ is a characteristic coordinate, which satisfies $Y_t + u^2 Y_x = 0$ for any $(t, x) \in \mathbb{R}^+ \times \mathbb{R}$. We also denote

$$\alpha = 2 \arctan u_x \quad \text{and} \quad \xi = \frac{(1 + u_x^2)^2}{Y_x}, \tag{4.1}$$

then one can derive a semi-linear system

$$
\begin{cases}
u_t(t,Y) = -\partial_x P_1 - P_2, \\
\alpha_t(t,Y) = -u\sin^2\frac{\alpha}{2} + 2u^3\cos^2\frac{\alpha}{2} - 2\cos^2\frac{\alpha}{2}(P_1 + \partial_x P_2), \\
\xi_t(t,Y) = \xi[(2u^3 + u) - 2(P_1 + \partial_x P_2)]\sin\alpha.
\end{cases}
\tag{4.2}
$$

Here the initial conditions given as

$$
\begin{cases}
u(0,Y) = u_0(x(0,Y)), \\
\alpha(0,Y) = 2\arctan u_{0,x}(x(0,Y)), \\
\xi(0,Y) = 1,
\end{cases}
\tag{4.3}
$$

where

$$
\begin{aligned}
P_1(Y) &= \frac{1}{2}\int_{-\infty}^{\infty} e^{-|\int_Y^{\tilde Y}(\xi\cos^4\frac{\alpha}{2})(t,\tilde Y)\,d\tilde Y|}\left(\frac{3}{8}u\sin^2\alpha + u^3\cos^4\frac{\alpha}{2}\right)\xi(t,\bar Y)\,d\bar Y, \\
\partial_x P_1(Y) &= \frac{1}{2}\left(\int_Y^{\infty} - \int_{-\infty}^Y\right) e^{-|\int_Y^{\tilde Y}(\xi\cos^4\frac{\alpha}{2})(t,\tilde Y)\,d\tilde Y|}\left(\frac{3}{8}u\sin^2\alpha + u^3\cos^4\frac{\alpha}{2}\right)\xi(t,\bar Y)\,d\bar Y, \\
P_2(Y) &= \frac{1}{8}\int_{-\infty}^{\infty} e^{-|\int_Y^{\tilde Y}(\xi\cos^4\frac{\alpha}{2})(t,\tilde Y)\,d\tilde Y|}\left(\xi\sin\alpha\sin^2\frac{\alpha}{2}\right)(t,\bar Y)\,d\bar Y, \\
\partial_x P_2(Y) &= \frac{1}{8}\left(\int_Y^{\infty} - \int_{-\infty}^Y\right) e^{-|\int_Y^{\tilde Y}(\xi\cos^4\frac{\alpha}{2})(t,\tilde Y)\,d\tilde Y|}\left(\xi\sin\alpha\sin^2\frac{\alpha}{2}\right)(t,\bar Y)\,d\bar Y.
\end{aligned}
\tag{4.4}
$$

By expressing the solution $u(t,Y)$ in terms of the original variables (t,x), one obtains a weak solution of the Cauchy problem (2.1) as stated in Theorem 2.1. Furthermore, this solution is proved to be the unique conservative solution in [9]. Indeed, the following result is also proved in [9].

Lemma 4.1. *Let $(x,u,\alpha,\xi)(t,Y)$ be the solution to the system (4.2)–(4.3) with $\xi > 0$. Then the set of points*

$$
\mathrm{Graph}(u) = \{(t,x(t,Y),u(t,Y)):\quad (t,Y)\in\mathbb{R}^+\times\mathbb{R}\}
\tag{4.5}
$$

is the graph of a unique conservative solution to the Novikov equation (2.1).

4.2. **Families of perturbed solutions.** To begin with, we state the following technical lemma (cf. [20]) which will be used later.

Lemma 4.2. *Consider an ODE system*

$$
\frac{d}{dt}\mathbf{u}^\varepsilon = f(\mathbf{u}^\varepsilon),\quad \mathbf{u}^\varepsilon(0) = \mathbf{u}_0 + \varepsilon_1\mathbf{v}_1 + \cdots + \varepsilon_m\mathbf{v}_m,
$$

where $\mathbf{u}^\varepsilon(t):\mathbb{R}\to\mathbb{R}^n$, f is a Lipschitz continuous function. The system is well-posed in $[0,t^*)$. Let the matrix

$$
D_\varepsilon\mathbf{u}^\varepsilon(0) = (\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_m)\in\mathbb{R}^{n\times m},
$$

have rank $\mathbf{rank}\,(D_\varepsilon\mathbf{u}^\varepsilon(0)) = k$. Then for any $t\in[0,t^*)$,

$$
\mathbf{rank}\,(D_\varepsilon\mathbf{u}^\varepsilon(t)) = k.
$$

Now, for a fixed solution of (4.2), we are going to construct several families of perturbed solutions.

Lemma 4.3. *Let (u,α,ξ) be a smooth solution of the semi-linear system (4.2) and let a point $(t_0,Y_0)\in\mathbb{R}^+\times\mathbb{R}$ be given.*

(1) If $(\alpha,\alpha_Y,\alpha_{YY})(t_0,Y_0) = (\pi,0,0)$, then there exists a 3–parameter family of smooth solutions $(u^\vartheta,\alpha^\vartheta,\xi^\vartheta)$ of (4.2), depending smoothy on $\vartheta\in\mathbb{R}^3$, such that the following holds.

(i) when $\vartheta = 0\in\mathbb{R}^3$, one recovers the original solution, namely $(u^0,\alpha^0,\xi^0) = (u,\alpha,\xi)$.

(ii) At the point (t_0,Y_0), when $\vartheta = 0$ one has

$$
\mathbf{rank}\,D_\vartheta(\alpha^\vartheta,\alpha_Y^\vartheta,\alpha_{YY}^\vartheta) = 3.
\tag{4.6}
$$

(2) *If $(\alpha, \alpha_Y, \alpha_t)(t_0, Y_0) = (\pi, 0, 0)$, then there exists a 3–parameter family of smooth solutions $(u^\vartheta, \alpha^\vartheta, \xi^\vartheta)$ of (4.2), depending smoothy on $\vartheta \in \mathbb{R}^3$, satisfying (i)–(ii) as above, with (4.6) replaced by*

$$\mathbf{rank}\ D_\vartheta(\alpha^\vartheta, \alpha_Y^\vartheta, \alpha_t^\vartheta) = 3. \tag{4.7}$$

Proof. Let (u, α, ξ) be a smooth solution of the semi-linear system (4.2). Given the point (t_0, Y_0). Taking derivatives to the equations of α and ξ in (4.2), we have

$$\begin{aligned}
\frac{\partial}{\partial t}\alpha_Y(t, Y) &= u_Y(6u^2 \cos^2 \frac{\alpha}{2} - \sin^2 \frac{\alpha}{2}) - \alpha_Y \sin \alpha(\frac{1}{2}u + u^3 - P_1 - \partial_x P_2) \\
&\quad - 2\cos^2 \frac{\alpha}{2}\xi\big(\cos^4 \frac{\alpha}{2}(\partial_x P_1 + P_2) - \frac{1}{4}\sin^2 \frac{\alpha}{2}\sin \alpha\big),
\end{aligned} \tag{4.8}$$

$$\begin{aligned}
\frac{\partial}{\partial t}\alpha_{YY}(t, Y) &= u_{YY}(6u^2 \cos^2 \frac{\alpha}{2} - \sin^2 \frac{\alpha}{2}) - \alpha_Y u_Y \sin \alpha(1 + 6u^2) \\
&\quad - (\alpha_{YY} \sin \alpha + \alpha_Y^2 \cos \alpha)(\frac{1}{2}u + u^3 - P_1 - \partial_x P_2) \\
&\quad + \alpha_Y \xi \sin \alpha\big(4\cos^4 \frac{\alpha}{2}(\partial_x P_1 + P_2) - \frac{1}{2}\sin^2 \frac{\alpha}{2}\sin \alpha\big) \\
&\quad - 2\cos^2 \frac{\alpha}{2}\xi_Y\big(\cos^4 \frac{\alpha}{2}(\partial_x P_1 + P_2) - \frac{1}{4}\sin^2 \frac{\alpha}{2}\sin \alpha\big) \\
&\quad - 2\cos^6 \frac{\alpha}{2}\xi^2\big(\cos^4 \frac{\alpha}{2}(P_1 + \partial_x P_2 - u^3) - \frac{3}{8}u\sin^2 \alpha\big) \\
&\quad + \frac{1}{4}\cos^2 \frac{\alpha}{2}\alpha_Y \xi\big(\sin^2 \alpha + 2\sin^2 \frac{\alpha}{2}\cos \alpha\big),
\end{aligned} \tag{4.9}$$

$$\begin{aligned}
\frac{\partial}{\partial t}\alpha_t(t, Y) &= (\partial_x P_1 + P_2)(\sin^2 \frac{\alpha}{2} - 6u^2 \cos^2 \frac{\alpha}{2}) - 2\cos^2 \frac{\alpha}{2}(\partial_t P_1 + \partial_t \partial_x P_2) \\
&\quad - \sin \alpha\big(-u\sin^2 \frac{\alpha}{2} + 2u^3 \cos^2 \frac{\alpha}{2} - 2\cos^2 \frac{\alpha}{2}(P_1 + \partial_x P_2)\big)(\frac{1}{2}u + u^3 - P_1 - \partial_x P_2),
\end{aligned} \tag{4.10}$$

$$\begin{aligned}
\frac{\partial}{\partial t}\xi_Y(t, Y) &= (\xi_Y \sin \alpha + \xi\alpha_Y \cos \alpha)\big(2u^3 + u - 2(P_1 + \partial_x P_2)\big) \\
&\quad + \xi \sin \alpha\big(6u^2 u_Y + u_Y - 2\xi \cos^4 \frac{\alpha}{2}(\partial_x P_1 + P_2) + \frac{1}{2}\xi \sin \alpha \sin^2 \frac{\alpha}{2}\big),
\end{aligned} \tag{4.11}$$

with

$$u_Y(t, Y) = \frac{1}{2}\xi \sin \alpha \cos^2 \frac{\alpha}{2}, \quad u_{YY} = \frac{1}{2}\xi_Y \sin \alpha \cos^2 \frac{\alpha}{2} + \frac{1}{2}\xi\alpha_Y \cos \alpha \cos^2 \frac{\alpha}{2} - \frac{1}{4}\xi\alpha_Y \sin^2 \alpha.$$

Combining (4.2), (4.8), (4.9) and (4.11), we obtain a complete system. Now, we construct a family of solutions $(\bar{u}^\vartheta, \bar{\alpha}^\vartheta, \bar{\xi}^\vartheta)$ to the complete system (4.2), (4.8), (4.9) and (4.11) with initial data being the perturbations of (4.3)

$$\begin{cases}
\bar{u}^\vartheta(0, Y) = u(0, Y) + \displaystyle\sum_{i=1,2,3} \vartheta_i U_i(Y), \\
\bar{\alpha}^\vartheta(0, Y) = \alpha(0, Y) + \displaystyle\sum_{i=1,2,3} \vartheta_i V_i(Y), \\
\bar{\xi}^\vartheta(0, Y) = \xi(0, Y) + \displaystyle\sum_{i=1,2,3} \vartheta_i \zeta_i(Y),
\end{cases} \tag{4.12}$$

for some suitable functions $U_i, V_i, \zeta_i \in \mathcal{C}_0^\infty(\mathbb{R})$. That is, consider the system

$$\frac{\partial}{\partial t} \begin{pmatrix} \bar{u}^\vartheta \\ \bar{\alpha}^\vartheta \\ \bar{\xi}^\vartheta \\ \bar{\alpha}_Y^\vartheta \\ \bar{\alpha}_{YY}^\vartheta \\ \bar{\xi}_Y^\vartheta \end{pmatrix} = \begin{pmatrix} f_1^\vartheta \\ f_2^\vartheta \\ f_3^\vartheta \\ f_4^\vartheta \\ f_5^\vartheta \\ f_6^\vartheta \end{pmatrix},$$

with the initial data (4.12), where $f_1^\vartheta, \cdots, f_6^\vartheta$ are the perturbations of the right-hand side of (4.2), (4.8), (4.9) and (4.11). Thanks to Lemma 4.2, it remains to prove the Lipschitz continuity of $f_1^\vartheta, \cdots, f_6^\vartheta$. Observe that (u, α, ξ) is smooth, so we just need to consider the Lipschitz continuity of P_j and $\partial_x P_j$ for $j = 1, 2$. From the definition of the source term P_j at (4.4), it follows

$$\left| \frac{\partial P_1}{\partial u}(t, Y) \right| = \frac{1}{2} \left| \int_{-\infty}^\infty e^{-|\int_{\bar{Y}}^Y (\xi \cos^4 \frac{\alpha}{2})(t, \tilde{Y}) d\tilde{Y}|} \left(\frac{3}{8} \sin^2 \alpha + 3u^2 \cos^4 \frac{\alpha}{2} \right) \xi(t, \bar{Y}) d\bar{Y} \right|$$

$$\leq \frac{1}{2} \int_{-\infty}^\infty e^{-|x - \bar{x}|} \left(\frac{3}{2} u_x^2 + 3u^2 \right)(t, \bar{x}) d\bar{x} \leq \frac{3}{2} \mathcal{E}(0).$$

Similarly, by (2.6) and (4.4), we can obtain

$$\left| \frac{\partial P_1}{\partial \alpha}(t, Y) \right| \leq \frac{1}{2} \int_{-\infty}^\infty e^{-|\int_{\bar{Y}}^Y (\xi \cos^4 \frac{\alpha}{2})(t, \tilde{Y}) d\tilde{Y}|} \left| \int_{\bar{Y}}^Y (\xi \cos^2 \frac{\alpha}{2} \sin \alpha)(t, \tilde{Y}) d\tilde{Y} \right|$$

$$\cdot \left| \frac{3}{8} u \sin^2 \alpha + u^3 \cos^4 \frac{\alpha}{2} \right) \left| \xi(t, \bar{Y}) d\bar{Y} \right.$$

$$+ \frac{1}{2} \int_{-\infty}^\infty e^{-|\int_{\bar{Y}}^Y (\xi \cos^4 \frac{\alpha}{2})(t, \tilde{Y}) d\tilde{Y}|} \left| \frac{3}{4} u \sin \alpha \cos \alpha - u^3 \cos^2 \frac{\alpha}{2} \sin \alpha \right| \xi(t, \bar{Y}) d\bar{Y}$$

$$\leq \frac{1}{2} \int_{-\infty}^\infty e^{-|x - \bar{x}|} \left| 2 \int_{\bar{x}}^x u_x(t, \tilde{x}) d\tilde{x} \right| \left| \left(\frac{3}{2} u u_x^2 + u^3 \right)(t, \bar{x}) \right| d\bar{x}$$

$$+ \frac{1}{2} \int_{-\infty}^\infty e^{-|x - \bar{x}|} \left| \left(\frac{3}{2} u u_x - \frac{3}{2} u u_x^3 - 2u^3 u_x \right)(t, \bar{x}) \right| d\bar{x}$$

$$\leq \int_{-\infty}^\infty e^{-|x - \bar{x}|} |x - \bar{x}|^{\frac{1}{2}} \mathcal{E}(0)^{\frac{1}{2}} \left| \left(\frac{3}{2} u u_x^2 + u^3 \right)(t, \bar{x}) \right| d\bar{x} + \frac{3}{4} \|u\|_{L^2} \|u_x\|_{L^2}$$

$$+ \frac{3}{4} \|u\|_{L^\infty} \|u_x\|_{L^3} + \|u\|_{L^\infty}^2 \|u\|_{L^2} \|u_x\|_{L^2}$$

$$\leq \frac{3\sqrt{2}}{2} \mathcal{E}(0)^{\frac{3}{2}} + \frac{3}{4} \mathcal{E}(0) + \frac{3}{4} \mathcal{E}(0)^{\frac{1}{2}} K^{\frac{1}{3}} + 2\mathcal{E}(0)^2,$$

where K is defined in (2.8). In a similar fashion as the above two estimates and the fact that ξ is bounded, it is easy to verify the boundedness of $|\frac{\partial P_2}{\partial u}|, |\frac{\partial P_2}{\partial \alpha}|, \frac{\partial P_j}{\partial \xi}|, |\partial_u \partial_x P_j|, |\partial_\alpha \partial_x P_j|, |\partial_\xi \partial_x P_j|$ for $j = 1, 2$. This completes the proof of the Lipschitz continuity of $f_1^\vartheta, \cdots, f_6^\vartheta$.

Thus, by choosing suitable perturbations V_i, $i = 1, 2, 3$, at the point (t_0, Y_0) and $\vartheta = 0$, using Lemma 4.2, we can get

$$\mathbf{rank}\ D_\vartheta \begin{pmatrix} \alpha \\ \alpha_Y \\ \alpha_{YY} \end{pmatrix} = 3.$$

On the other hand, (4.2), (4.8) and (4.10) form a complete system. Similar to the proof of (4.6), we can choose suitable perturbations V_i, $i = 1, 2, 3$, such that, at the point (t_0, Y_0) and $\vartheta = 0$, we

have

$$\mathbf{rank}\ D_\vartheta \begin{pmatrix} \alpha \\ \alpha_Y \\ \alpha_t \end{pmatrix} = 3.$$

This completes the proof of Lemma 4.3.   □

4.3. **Proof of Theorem 1.2.** Now, we will use Lemma 4.3 together with Transversality argument to study the smooth solutions to the semi-linear system (4.2), and hence determine the generic structure of the level sets $\{(t,Y); \alpha(t,Y) = \pi\}$. The proof of the following lemma is similar to [20], and we omit it here for brevity.

Lemma 4.4. *Let a compact domain*

$$\Omega := \{(t,Y);\ 0 \le t \le T, |Y| \le M\},$$

and define S to be the family of all $\mathcal{C}^2$ solutions (u, α, ξ) to the semi-linear system (4.2), with $\xi > 0$ for all $(t,Y) \in \mathbb{R}^+ \times \mathbb{R}$. Moreover, define $S' \subset S$ to be the subfamily of all solutions (u, α, ξ), such that for $(t,Y) \in \Omega$, none of the following values is attained:

$$(\alpha, \alpha_Y, \alpha_{YY}) = (\pi, 0, 0), \quad (\alpha, \alpha_Y, \alpha_t) = (\pi, 0, 0). \tag{4.13}$$

Then S' is a relatively open and dense subset of S, in the topology induced by $\mathcal{C}^2(\Omega)$.

With the help of Lemma 4.4, we can now prove Theorem 1.2 for the generic regularity of conservative solutions to the Novikov equation (2.1).

Proof of Theorem 1.2. First, we denote

$$\mathcal{N} := \mathcal{C}^3(\mathbb{R}) \cap H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R}),$$

with norm

$$\|u_0\|_\mathcal{N} := \|u_0\|_{\mathcal{C}^3} + \|u_0\|_{H^1} + \|u_0\|_{W^{1,4}}.$$

Given initial data $\hat{u}_0 \in \mathcal{N}$ and denote the open ball

$$B_\delta := \{u_0 \in \mathcal{N};\ \|u_0 - \hat{u}_0\|_\mathcal{N} < \delta\}.$$

Now, we will prove our theorem by showing that, for any $\hat{u}_0 \in \mathcal{N}$, there exists a radius $\delta > 0$ and an open dense subset $\tilde{\mathcal{D}} \subset B_\delta$, with the following property: for every initial data $u_0 \in \tilde{\mathcal{D}}$, the conservative solution $u = u(t,x)$ of (2.1) is twice continuously differentiable in the complement of finitely many characteristic curves within the domain $[0,T] \times \mathbb{R}$.

1. Since $u_0 \in \mathcal{N}$, by the definition of the space $\mathcal{N}$, we have

$$u_0(x) \to 0 \quad \text{and} \quad u_{0,x}(x) \to 0, \quad \text{as } |x| \to \infty.$$

Thus, we can choose $r > 0$ sufficiently large, such that $u_0(x)$ and $u_{0,x}(t,x)$ being uniformly bounded for all $|x| \ge r$. On a domain of the form $\{(t,x);\ t \in [0,T], |x| \ge r+T\|u\|_{L^\infty}^2\}$, a standard comparison argument yields that the partial derivative u_x remains uniformly bounded. This implies that the singularities of $u(t,x)$ in the set $[0,T] \times \mathbb{R}$ only appear on the compact set $\mathcal{M} := [0,T] \times [-r - T\|u\|_{L^\infty}^2, r + T\|u\|_{L^\infty}^2]$.

Next, for any $u_0 \in B_\delta$, let Λ be the map $(t,Y) \mapsto \Lambda(t,Y) := (t, x(t,Y))$, and let Ω be a domain as in Lemma 4.4. Choosing M large enough and by possibly shrinking the radius δ, we can obtain the inclusion $\mathcal{M} \subset \Lambda(\Omega)$.

Now, the subset $\tilde{\mathcal{D}} \subset B_\delta$ is defined as follows. $u_0 \in \tilde{\mathcal{D}}$ if $u_0 \in B_\delta$ and for the corresponding solution (u, α, ξ) of (4.2), the values (4.13) are never attained for any (t,Y) such that $(t, x(t,Y)) \in \mathcal{M}$. In the next two steps, we examine $\tilde{\mathcal{D}}$ is an open dense set.

2. To prove that the set $\tilde{\mathcal{D}}$ is open in the topology of $\mathcal{C}^3$, we choose a sequence of initial data $(u_0^\nu)_{\nu \geq 1}$ such that the sequence converges to u_0, with $u_0^\nu \notin \tilde{\mathcal{D}}$. By the definition of $\tilde{\mathcal{D}}$, there exist points (t^ν, Y^ν) at which the corresponding solutions $(u^\nu, \alpha^\nu, \xi^\nu)$ satisfy

$$(\alpha^\nu, \alpha_Y^\nu, \alpha_{YY}^\nu)(t^\nu, Y^\nu) = (\pi, 0, 0), \quad (t^\nu, x^\nu(t^\nu, Y^\nu)) \in \mathcal{M},$$

for all $\nu \geq 1$. Observe that the domain $\mathcal{M}$ is compact, thus we can choose a subsequence, denote still by (t^ν, Y^ν), such that $(t^\nu, Y^\nu) \to (\bar{t}, \bar{Y})$ for some point $(\bar{t}, \bar{Y})$. By continuity,

$$(\alpha, \alpha_Y, \alpha_{YY})(\bar{t}, \bar{Y}) = (\pi, 0, 0), \quad (t, x(\bar{t}, \bar{Y})) \in \mathcal{M},$$

which implies $u_0 \notin \tilde{\mathcal{D}}$. Repeating the same procedure when $(\alpha, \alpha_Y, \alpha_T) = (\pi, 0, 0)$, it follows that $\tilde{\mathcal{D}}$ is open.

3. Now, we will prove the set $\tilde{\mathcal{D}}$ is dense in B_δ. Given $u_0 \in B_\delta$, by a small perturbation, we can assume that $u_0 \in \mathcal{C}^\infty$.

From Lemma 4.4, we can construct a sequence of solutions $(u^\nu, \alpha^\nu, \xi^\nu)$ of (4.2), such that, for every $\nu \geq 1, (t, Y) \in \Omega$, the values in (4.13) are never attained, and the $\mathcal{C}^k, k \geq 1$ norm satisfies

$$\lim_{\nu \to \infty} \|(u^\nu - u, \alpha^\nu - \alpha, \xi^\nu - \xi, x^\nu - x)\|_{\mathcal{C}^k(\Gamma)} = 0,$$

for every bounded set $\Gamma \subset [0, T] \times \mathbb{R}$. Thus, for $t = 0$, the corresponding sequence of initial values satisfies

$$\lim_{\nu \to \infty} \|u_0^\nu - u_0\|_{\mathcal{C}^k(I)} = 0, \tag{4.14}$$

for every bounded set $I \subset \mathbb{R}$.

Consider a cutoff function $\psi(x) \in \mathcal{C}_0^\infty$, such that

$$\psi(x) = 1, \qquad \text{if } |x| \leq l,$$
$$\psi(x) = 0, \qquad \text{if } |x| \geq l + 1,$$

where $l \gg r + T\|u\|_{L^\infty}^2$ is large enough. Then for every $\nu \geq 1$, define the following initial data

$$\tilde{u}_0^\nu := \psi u_0^\nu + (1 - \psi)u_0,$$

which together with (4.14) implies

$$\lim_{\nu \to \infty} \|\tilde{u}_0^\nu - u_0\|_{\mathcal{N}} = 0.$$

Further enlarge l such that for every $(t, x) \in \mathcal{M}$,

$$\tilde{u}^\nu(t, x) = u^\nu(t, x).$$

Notice that $\tilde{u}^\nu(t, x)$ remains $\mathcal{C}^2$ in the outer domain $\{(t, x); \ t \in [0, T], |x| \geq r + T\|u\|_{L^\infty}^2\}$. Thus, we have proved for every $\nu \geq 1$ sufficiently large, $\tilde{u}_0^\nu \in \tilde{D}$, which means that $\tilde{D}$ is dense in B_δ.

4. Finally, we will show that for every initial data $u_0 \in \tilde{D}$, the corresponding solution $u(t, x)$ of (2.1) is piecewise $\mathcal{C}^2$ on the domain $[0, T] \times \mathbb{R}$. Indeed, we know that u is $\mathcal{C}^2$ in the outer domain $\{(t, x); \ t \in [0, T], |x| \geq r + T\|u\|_{L^\infty}^2\}$ by the previous argument. So we need to study the singularity of u in the inner domain $\mathcal{M}$.

Recall from step 1, every point in $\mathcal{M}$ is contained in the image of the domain Ω. Thus, for every point $(t_0, Y_0) \in \Omega$, we have two cases.

Case I. $\alpha(t_0, Y_0) \neq \pi$. From the coordinate change $x_Y = \xi \cos^4 \frac{\alpha}{2}, t = t$, we know that the map $(t, Y) \mapsto (t, x)$ is locally invertible in a neighborhood of (t_0, Y_0). Therefore, the function u is $\mathcal{C}^2$ in a neighborhood of $(t_0, x(t_0, Y_0))$.

Case II. $\alpha(t_0, Y_0) = \pi$. Since $u_0 \in \tilde{D}$, so by the definition of $\tilde{D}$, we have $\alpha_t(t_0, Y_0) \neq 0$ or $\alpha_Y(t_0, Y_0) \neq 0$.

5. By continuity, there exists an $\eta > 0$, such that the values in (4.13) are never attained in the open neighborhood

$$\Omega' := \{(t, Y); \ 0 \leq t \leq T, |Y| \leq M + \eta\}.$$

Applying the implicit function theorem, we derive that the set

$$S^\alpha := \{(t, Y) \in \Omega'; \ \alpha(t, Y) = \pi\}$$

is a one-dimensional embedded manifold of class $\mathcal{C}^2$.

Now, we claim that the number of connected components of S^α that intersect the compact set Ω is finite. Assume, by contradiction, that $P_1, P_2, \cdots$ is a sequence of points in $S^\alpha \cap \Omega$ belonging to distinct components. Thus, we can choose a subsequence P_i, such that $P_i \to \bar{P}$ for some $\bar{P} \in S^\alpha \cap \Omega$. By assumption, $(\alpha_t, \alpha_Y)(\bar{P}) \neq (0, 0)$. Hence, by the implicit function theorem, there is a neighborhood $\mathcal{U}$ of $\bar{P}$ such that $\gamma := S^\alpha \cap \mathcal{U}$ is a connected $\mathcal{C}^2$ curve. Thus, $P_i \in \gamma$ for i large enough, providing a contradiction.

6. To complete the proof, we need to study in more detail the image of the singular set S^α, since the set of points (t, x) where u is singular coincides with the image of the set S^α under the $\mathcal{C}^2$ map $(t, Y) \mapsto \Lambda(t, Y) = (t, x(t, Y))$.

By the argument in step 5, inside the compact set Ω, there are only finitely many points where $\alpha = \pi, \alpha_Y = 0, \alpha_t \neq 0$, say $P_i = (t_i, Y_i), i = 1, \cdots, m$, and also where $\alpha = \pi, \alpha_t = 0, \alpha_Y \neq 0$, say $Q_\jmath = (t'_\jmath, Y'_\jmath), \jmath = 1, \cdots, n$.

By the analysis in step 5, the set $S^\alpha \backslash \{P_1, \cdots, P_m, Q_1, \cdots, Q_n\}$ has finitely many connected components which intersect Ω. Consider any one of these components. This is a connected curve, say γ_j, such that $\alpha = \pi, \alpha_Y \neq 0$ for any $(t, Y) \in \gamma_j$. Thus, this curve can be expressed in the form

$$\gamma_j = \{(t, Y): \ Y = \phi_j(t), a_j < t < b_j\},$$

for a suitable function ϕ_j.

At this stage, we claim that the image $\Lambda(\gamma_j)$ is a $\mathcal{C}^2$ curve in the t–x plane. Indeed, it suffices to show that, on the open interval (a_j, b_j), the differential of the map $t \mapsto (t, x(t, \phi_j(t)))$ does not vanish. This is true, because

$$\frac{d}{dt} x(t, \phi_j(t)) = 1 + x_Y \phi'_j = 1 > 0,$$

since $x_Y = \xi \cos^4 \frac{\alpha}{2} = 0$ when $\alpha = \pi$. Hence, the singular set $\Lambda(S^\alpha)$ is the union of the finitely many points $p_i = \Lambda(P_i), i = 1, \cdots, m$, $q_\jmath = \Lambda(Q_\jmath), \jmath = 1, \cdots, n$, together with finitely many $\mathcal{C}^2$–curve $\Lambda(\gamma_j)$. This completes the proof of Theorem 1.2. $\qquad \square$

4.4. **One-parameter families of solutions.** In this subsection, we study families of conservative solutions $u^\theta = u(t, x, \theta)$ of (2.1) with initial data $u(0, x, \theta) = u_0(x, \theta)$, depending smoothly on an additional parameter $\theta \in [0, 1]$, More precisely, these paths of initial data will lie in the space

$$\mathcal{X} := \mathcal{C}^3\big(\mathbb{R} \times [0, 1]\big) \cap L^\infty\big([0, 1]; H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})\big).$$

Now, we have the following generic regularity for one-parameter family of solution. Roughly speaking, for a one-parameter family of initial data $\theta \mapsto \hat{u}_0^\theta$, with $\theta \in [0, 1]$, it can be uniformly approximated by a second path of initial data $\theta \mapsto u_0^\theta$, such that the corresponding solutions $u^\theta = u^\theta(t, x)$ of (2.1) are piecewise smooth in the domain $[0, T] \times \mathbb{R}$. The proof is similar to [1], and hence we omit it here for brevity. Note that this is the step where Thom's Transversality Theorem is used. We refer the reader to [1] for the Thom's Transversality Theorem.

Theorem 4.1. *Let $T > 0$ be given, then for any one-parameter family of initial data $\hat{u}_0^\theta \in \mathcal{X}$ and any $\varepsilon > 0$, there exists a perturbed family $(x, \theta) \mapsto u_0(x, \theta) =: u_0^\theta(x)$ such that*

$$\|u_0^\theta - \hat{u}_0^\theta\|_{\mathcal{X}} < \varepsilon,$$

and moreover, for all except at most finitely many $\theta \in [0, 1]$, the conservative solution $u^\theta = u(t, x, \theta)$ of (2.1) is smooth in the complement of finitely many points and finitely many $\mathcal{C}^2$ curves in the domain $[0, T] \times \mathbb{R}$.

In accordance with the previous argument, we give the following definitions.

Definition 4.1. *We say that a solution $u = u(t, x)$ of (2.1) has generic singularities for $t \in [0, T]$ if it admits a representation of the form (4.5), where*
   (i) *the functions $(x, u, \alpha, \xi)(t, Y)$ are C^{∞},*
   (ii) *for $t \in [0, T]$, the generic condition*

$$\alpha = \pi, \alpha_Y = 0 \implies \alpha_t \neq 0, \alpha_{YY} \neq 0, \tag{4.15}$$

holds.

Definition 4.2. *We say that a path of initial data $\gamma^0 : \theta \mapsto u_0^{\theta}$, $\theta \in [0, 1]$ is a piecewise regular path if the following conditions hold*
   (i) *There exists a continuous map $(Y, \theta) \mapsto (x, u, \alpha, \xi)$ such that the semi-linear system (4.2)–(4.3) holds for $\theta \in [0, 1]$, and the function $u^{\theta}(x, t)$ whose graph is*

$$Graph\ (u^{\theta}) = \{(t, x(t, Y, \theta), u(t, Y, \theta)); \ (t, Y) \in \mathbb{R}^+ \times \mathbb{R}\}$$

provides the conservation solution of (2.1) with initial data $u^{\theta}(0, x) = u_0^{\theta}(x)$.
   (ii) *There exist finitely many values $0 = \theta_0 < \theta_1 < \cdots < \theta_N = 1$ such that the map $(Y, \theta) \mapsto (x, u, \alpha, \xi)$ is C^{∞} for $\theta \in (\theta_{i-1}, \theta_i), i = 1, \cdots, N$, and the solution $u^{\theta} = u^{\theta}(t, x)$ has only generic singularities at time $t = 0$.*
   In addition, if for all $\theta \in [0, 1] \backslash \{\theta_1, \cdots, \theta_N\}$, the solution u^{θ} has generic singularities for $t \in [0, T]$, then we say that the path of solution $\gamma^t : \theta \mapsto u^{\theta}$ is **piecewise regular** *for $t \in [0, T]$.*

An application of Theorem 4.1 gives the following density result.

Corollary 4.1. *Given $T > 0$, let $\theta \mapsto (x^{\theta}, u^{\theta}, \alpha^{\theta}, \xi^{\theta}), \theta \in [0, 1]$, be a smooth path of solutions to the system (4.2)–(4.3). Then there exists a sequence of paths of solutions $\theta \mapsto (x_n^{\theta}, u_n^{\theta}, \alpha_n^{\theta}, \xi_n^{\theta})$, such that*
   (i) *For each $n \geq 1$, the path of the corresponding solution of (2.1) $\theta \mapsto u_n^{\theta}$ is regular for $t \in [0, T]$ in the sense of Definition 4.2.*
   (ii) *For any bounded domain Ω in the (t, Y) space, the functions $(x_n^{\theta}, u_n^{\theta}, \alpha_n^{\theta}, \xi_n^{\theta})$ converge to $(x^{\theta}, u^{\theta}, \alpha^{\theta}, \xi^{\theta})$ uniformly in $C^k([0, 1] \times \Omega)$, for every $k \geq 1$, as $n \to \infty$.*

5. Metric for general weak solutions

In this section, we will first extend the Lipschitz metric for smooth solutions in Section 3 to piecewise smooth solutions, and then to general weak solutions using the generic regularity result established in Theorem 1.2. Through this Theorem 1.1 can be achieved.

5.1. **Tangent vectors in transformed coordinates.** For a reference solution $u(t, x)$ of (2.1) and a family of perturbed solutions $u^{\varepsilon}(t, x)$, we assume that, in the (t, Y) coordinates, these define a family of smooth solutions of (4.2)–(4.3), denoted by $(x^{\varepsilon}, u^{\varepsilon}, \alpha^{\varepsilon}, \xi^{\varepsilon})$.

Consider the perturbed solutions of the form

$$(x^{\varepsilon}, u^{\varepsilon}, \alpha^{\varepsilon}, \xi^{\varepsilon}) = (x, u, \alpha, \xi) + \varepsilon(X, U, A, \zeta) + o(\varepsilon). \tag{5.1}$$

Since the coefficients of (4.2)–(4.3) are smooth, we have that the first order perturbations satisfy a linearized system and are well defined for $(t, Y) \in \mathbb{R}^+ \times \mathbb{R}$. In the following, we will express the quantities in (3.9) in terms of (X, U, A, ζ).
   (1) The shift in x is computed by

$$w = \lim_{\varepsilon \to 0} \frac{x^{\varepsilon}(t, Y^{\varepsilon}) - x(t, Y)}{\varepsilon} = X + x_Y \cdot \left. \frac{\partial Y^{\varepsilon}}{\partial \varepsilon} \right|_{\varepsilon = 0}. \tag{5.2}$$

   (2) The change in u is

$$v + u_x w = \lim_{\varepsilon \to 0} \frac{u^{\varepsilon}(t, Y^{\varepsilon}) - u(t, Y)}{\varepsilon} = U + u_Y \cdot \left. \frac{\partial Y^{\varepsilon}}{\partial \varepsilon} \right|_{\varepsilon = 0}. \tag{5.3}$$

(3) For the change in $\arctan u_x$, it holds that

$$v_x + u_{xx}w = \frac{d}{d\varepsilon}\tan\frac{\alpha^\varepsilon(t, Y^\varepsilon)}{2}\Big|_{\varepsilon=0} = \frac{1}{2}\sec^2\frac{\alpha}{2}\left(A + \alpha_Y \cdot \frac{\partial Y^\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0}\right). \tag{5.4}$$

(4) Finally, we will derive an expression for the term I_4 in (3.9).

$$\frac{d}{d\varepsilon}(\xi^\varepsilon(t, Y^\varepsilon) + \xi x_Y Y_x^\varepsilon)\Big|_{\varepsilon=0} = \zeta + \xi_Y \cdot \frac{\partial Y^\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0} + \xi x_Y \cdot \frac{\partial Y_x^\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0}, \tag{5.5}$$

where the change in the base measure with density $(1 + u_x^2)^2$ is given by

$$\frac{d}{d\varepsilon}\xi^\varepsilon\Big|_{\varepsilon=0} = \lim_{\varepsilon\to 0}\frac{\xi^\varepsilon(t, Y^\varepsilon) - \xi(t, Y)}{\varepsilon} = \zeta + \xi_Y \cdot \frac{\partial Y^\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0}.$$

Notice that

$$(1 + u_x^2)^2\, dx = \xi\, dY.$$

In light of (5.2)–(5.5), the weighted norm of a tangent vector (3.9) can be written as

$$\|(w, v)\|_u = \sum_{\ell=1}^{4}\int_{\mathbb{R}}|J_\ell(t, Y)|\, dY, \tag{5.6}$$

where

$$J_1 = \left(X + x_Y \cdot \frac{\partial Y^\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0}\right)\xi e^{-|x(t,Y)|}, \qquad J_2 = \left(U + u_Y \cdot \frac{\partial Y^\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0}\right)\xi e^{-|x(t,Y)|},$$

$$J_3 = \frac{1}{2}\left(A + \alpha_Y \cdot \frac{\partial Y^\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0}\right)\xi e^{-|x(t,Y)|}, \quad J_4 = \left(\zeta + \xi_Y \cdot \frac{\partial Y^\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0} + \xi x_Y \cdot \frac{\partial Y_x^\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0}\right)e^{-|x(t,Y)|}.$$

Since $Y^\varepsilon(t) = Y^\varepsilon(0)$ along the characteristic, it is easy to verify that the integrands J_ℓ are smooth, for $\ell = 1, 2, 3, 4$.

5.2. **Length of piecewise regular paths.** Now we define the length of the piecewise regular path by optimizing over all transportation plans, namely,

Definition 5.1. *The length $\|\gamma^t\|$ of the piecewise regular path $\gamma^t : \theta \mapsto u^\theta(t)$ is defined as*

$$\|\gamma^t\| = \inf_{\gamma^t}\int_0^1\sum_{\ell=1}^{4}\int_{\mathbb{R}}|J_\ell^\theta(t, Y)|\, dY\, d\theta,$$

where the infimum is taken over all piecewise regular paths.

Our main result in this section is stated as follows.

Theorem 5.1. *Given any $T > 0$, consider a path of solutions $\theta \mapsto u^\theta$ of (2.1), which is piecewise regular for $t \in [0, T]$. Moreover, the total energy $\|u^\theta\|_{H^1}^2$ and the norm $\|u_x^\theta\|_{L^4}^4$ is less than some constants $E_1 > 0$ and $E_2 > 0$, respectively. Then there exists some constant $C > 0$, such that the length satisfies*

$$\|\gamma^t\| \le C\|\gamma^0\|, \tag{5.7}$$

where the constant C depends only on T, E_1 and E_2.

Proof. By the definition of piecewise regular paths, we know that u^θ has generic regularities for every $\theta \in [0, 1]\backslash\{\theta_1, \cdots, \theta_N\}$. So the solution u^θ is smooth in the t-Y coordinates and piecewise smooth in the t-x coordinates, thus the existence of the tangent vector is obvious.

1. We claim that, for $\theta \in [0,1] \backslash \{\theta_1, \cdots, \theta_N\}$, if we have

$$\|v^\theta(t)\|_{u^\theta(t)} \le e^{C_1 t} \|v^\theta(0)\|_{u^\theta(0)}, \tag{5.8}$$

then (5.7) holds. Here the constant C_1 depends only on the upper bound of the total energies.

Indeed, according to the definition of the length of a piecewise regular path, fix $\epsilon > 0$, there exists some Y, such that, at time $t = 0$,

$$\int_0^1 \sum_{\ell=1}^4 \int_{\mathbb{R}} |J_\ell^\theta(0, Y)| \, dY \, d\theta \le \|\gamma^0\| + \epsilon.$$

Thus, integrating (5.8) over the interval $\theta \in [0,1]$, it holds that

$$\|\gamma^t\| \le C_2(\|\gamma^0\| + \epsilon),$$

which implies (5.7), since $\epsilon > 0$ is arbitrary.

2. Therefore in the following, we will concentrate on proving (5.8). If u^θ is smooth in the (t, x) variables, in light of (3.10), (5.8) follows directly. Thus, it suffices to show that the same result can be applied if u^θ is piecewise smooth with generic singularities.

We first note that there exist at most finitely many points $Q_j = (t_j, Y_j), j = 1, \cdots, N$ on which $\alpha^\theta = \pi$, $\alpha_Y^\theta = 0$, when $t \in [0, T]$. In fact, using the implicit function theorem and the generic condition (4.15), we know that each point Q with $\alpha^\theta = \pi$ and $\alpha_Y^\theta = 0$ is isolated from other $Q's$ in a small neighborhood. Then it follows that there exist at most finite many points Q_j by applying a finite covering argument on the bounded region including all possible singularities (the existence of such a bounded region is shown in Part 1 of the proof of Theorem 1.2 in Section 4.3.).

Now, for each time $t = t_j$ corresponding to the point Q_j, the map

$$t \mapsto \int_0^1 \sum_{\ell=1}^4 \int_{\mathbb{R}} |J_\ell^\theta(t, Y)| \, dY \, d\theta$$

is continuous, Since the number of t_j is finite, the metric will not be affected.

At time $t \ne t_j$, on the other hand, we claim that the generic singularity does not affect the estimate (3.10), that is, the effect of the generic singularity to the time derivative

$$\frac{d}{dt} \sum_{\ell=1}^4 \int_{\mathbb{R}} |J_\ell^\theta(t, Y)| \, dY$$

is negligible.

We know that

$$\alpha^\theta = \pi, \qquad \alpha_Y^\theta \ne 0 \tag{5.9}$$

at the singularities. For a fixed time τ, let the point $(t_\varepsilon, Y_\varepsilon)$ be the intersection of $\Gamma_{\tau-\varepsilon} = \{(t, Y); t = \tau - \varepsilon\}$ and $\{(t, Y); \alpha^\theta(t, Y) = \pi\}$, and the point $(t_\varepsilon', Y_\varepsilon')$ be the intersection of $\Gamma_{\tau+\varepsilon} = \{(t, Y); t = \tau + \varepsilon\}$ and $\{(t, Y); \alpha^\theta(t, Y) = \pi\}$. Then denote

$$\begin{cases} \Lambda_\varepsilon^+ := \Gamma_{\tau+\varepsilon} \cap \{(t, Y); Y \in [Y_\varepsilon', Y_\varepsilon]\}, \\ \Lambda_\varepsilon^- := \Gamma_{\tau-\varepsilon} \cap \{(t, Y); Y \in [Y_\varepsilon', Y_\varepsilon]\}. \end{cases}$$

Thus, we have

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \Big(\int_{\Lambda_\varepsilon^+} - \int_{\Lambda_\varepsilon^-} \Big) \sum_{\ell=1}^4 |J_\ell^\theta(t, Y)| \, dY = 0,$$

since each integrand is continuous and $|Y_\varepsilon - Y_\varepsilon'| = O(\varepsilon)$ because of (5.9). This means that the curve $\{(t, Y); \alpha^\theta(t, Y) = \pi\}$ has non-horizontal tangent line at the singularity. This implies that the estimate (3.10) remains valid in the presence of singular curve where $\alpha^\theta = \pi$. Hence we complete the proof of Theorem 5.1. □

Remark 5.1. *We want to point out that the assumption that the solution path in Corollary 4.1 is regular, i.e. the solution has only generic singularities, except at finitely many θ values, is very crucial. Because of this property we show that there are only finitely many points Q_j with $\alpha = \pi$, $\alpha_Y = 0$. However at time $t = t_j$ we cannot treat these Q_j's directly by the method used for $t \neq t_j$ since now $|Y_\varepsilon - Y'_\varepsilon| = O(1)$.*

5.3. **Construction of the geodesic distance.** In light of Theorem 1.2, there exists an open dense set $\mathcal{D} \subset \left(\mathcal{C}^3(\mathbb{R}) \cap H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R}) \right)$, such that, for $u_0 \in \mathcal{D}$, the solution of (2.1) has only generic singularities. Now, on $\mathcal{D}^\infty := \mathcal{C}_0^\infty \cap \mathcal{D}$, we construct a geodesic distance, defined as the infimum among the weighted lengths of all piecewise regular paths connecting two given points.

Consider two solutions $u, \tilde{u} \in \mathcal{D}^\infty$. Denote their total energies as

$$\mathcal{E}(u) := \int_\mathbb{R} (u^2 + u_x^2)(x)\, dx, \quad \mathcal{E}(\tilde{u}) := \int_\mathbb{R} (\tilde{u}^2 + \tilde{u}_x^2)(x)\, dx,$$

respectively. For $E_1 > 0$ and $E_2 > 0$ denote the set

$$\Sigma := \{ u \in H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R}); \ \mathcal{E}(u) \leq E_1, \|u_x\|_{L^4}^4 \leq E_2 \}.$$

Definition 5.2. *For solutions with initial data in $\mathcal{D}^\infty \cap \Sigma$, we define the geodesic distance $d(u, \tilde{u})$ as the infimum among the weighted lengths of all piecewise regular paths, which connect u with $\tilde{u}$, that is, for any time t,*

$$d(u, \tilde{u}) := \inf \left\{ \|\gamma^t\| : \ \gamma^t \text{ is a piecewise regular path}, \gamma^t(0) = u, \gamma^t(1) = \tilde{u}, \right.$$

$$\left. \mathcal{E}(u^\theta) \leq E_1, \|u_x^\theta\|_{L^4}^4 \leq E_2, \ \text{for all } \theta \in [0,1] \right\}.$$

Finally we can define the metric for the general weak solutions.

Definition 5.3. *Let u_0 and $\tilde{u}_0$ in $H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})$ be two absolute continuous initial data as required in the existence and uniqueness Theorem 2.1. Denote u and $\tilde{u}$ to be the corresponding global weak solutions, then we define, for any time t,*

$$d(u, \tilde{u}) := \lim_{n \to \infty} d(u^n, \tilde{u}^n),$$

for any two sequences of solutions u^n and $\tilde{u}^n$ in $\mathcal{D}^\infty \cap \Sigma$ with

$$\|u^n - u\|_{H^1 \cap W^{1,4}} \to 0, \quad \text{and} \quad \|\tilde{u}^n - \tilde{u}\|_{H^1 \cap W^{1,4}} \to 0.$$

The limit in the definition is independent of the choice of sequences, because the solution flows are Lipschitz in $\mathcal{D}^\infty \cap \Sigma$. Since the concatenation of two piecewise regular paths is still a piecewise regular path (after a suitable re-parameterization), it is clear that $d(\cdot, \cdot)$ is a distance. This way the metric is well-defined.

Note that when

$$\|u_0^n - u_0\|_{H^1 \cap W^{1,4}} \to 0,$$

it is easy to show that the corresponding solutions satisfy, for any $t > 0$,

$$\|u^n - u\|_{H^1 \cap W^{1,4}} \to 0$$

by the semi-linear equations (4.2). So it is clear that the Lipschitz property in Theorem 5.1 can be extended to the general solutions, and so we conclude to obtain Theorem 1.1.

6. COMPARISON WITH OTHER METRICS

The purpose of this section is to compare the distance $d(\cdot, \cdot)$ with other types of metrics.

Proposition 6.1 (Comparison with the Sobolev metric). *For any solutions $u, \tilde{u} \in \Sigma$ to (2.1), there exists some constant C depends only on E_1 and E_2, such that,*

$$d(u, \tilde{u}) \leq C \left(\|u - \tilde{u}\|_{H^1} + \|(u - \tilde{u})e^{-|x|}\|_{L^1} + \|(u_x - \tilde{u}_x)e^{-|x|}\|_{L^1} + \|u_x - \tilde{u}_x\|_{L^4} \right).$$

Proof. Without loss of generality, we only consider the solutions in $\mathcal{D}^\infty \cap \Sigma$.

For $\theta \in [0,1]$, consider the interpolated data u^θ as

$$u^\theta = \theta\tilde{u} + (1-\theta)u. \tag{6.1}$$

Then the total energy of u^θ satisfies

$$
\begin{aligned}
\int_{\mathbb{R}} \Big((u^\theta)^2 + (u_x^\theta)^2 \Big)(t,x)\, dx &= \int_{\mathbb{R}} \Big([\theta\tilde{u} + (1-\theta)u]^2 + [\theta\tilde{u}_x + (1-\theta)u_x]^2 \Big)\, dx \\
&\leq \int_{\mathbb{R}} \Big([\theta^2 + \theta(1-\theta)](\tilde{u}^2 + \tilde{u}_x^2) + [(1-\theta)^2 + \theta(1-\theta)](u^2 + u_x^2) \Big)\, dx \\
&\leq \max\{\mathcal{E}(u), \mathcal{E}(\tilde{u})\} \leq E_1.
\end{aligned}
\tag{6.2}
$$

Also, for the L^4 norm of u_x^θ, it holds that

$$\int_{\mathbb{R}} (u_x^\theta)^4(t,x)\, dx \leq \int_{\mathbb{R}} \big(\theta\tilde{u}_x^4 + (1-\theta)u_x^4 \big)\, dx \leq \max\left\{ \int_{\mathbb{R}} u_x^4\, dx, \int_{\mathbb{R}} \tilde{u}_x^4\, dx \right\} \leq E_2. \tag{6.3}$$

Now, we will estimate the weighted length of the path $\gamma^t : \theta \mapsto u^\theta$ in (6.1). The goal is to show that

$$\|\gamma^t\| \leq C\Big(\|u - \tilde{u}\|_{H^1} + \|(u-\tilde{u})e^{-|x|}\|_{L^1} + \|(u_x - \tilde{u}_x)e^{-|x|}\|_{L^1} + \|u_x - \tilde{u}_x\|_{L^4} \Big)$$

for some constant C depends only on E_1 and E_2. First, from (6.1), we obtain

$$v^\theta = \frac{du^\theta}{d\theta} = \tilde{u} - u, \quad u_x^\theta = \theta\tilde{u}_x + (1-\theta)u_x. \tag{6.4}$$

To derive an upper bound for the weighted length $\|\gamma^t\|$, we can choose the shift $w = 0$ in (3.9). Indeed, by (6.2)–(6.4) and the definition of the weighted length of the path γ^t, we have

$$
\begin{aligned}
\|\gamma^t\| &= \int_0^1 \|v^\theta\|_{u^\theta}\, d\theta \\
&= \int_0^1 \int_{\mathbb{R}} \Big(|v^\theta|(1 + (u_x^\theta)^2)^2 + |v_x^\theta|(1 + (u_x^\theta)^2) + 4\big|\big(u_x^\theta + (u_x^\theta)^3\big)v_x^\theta\big| \Big)e^{-|x|}\, dx\, d\theta \\
&\leq \|(\tilde{u} - u)e^{-|x|}\|_{L^1} + C\|\tilde{u} - u\|_{L^\infty} \int_0^1 (\|u_x^\theta\|_{L^2}^2 + \|u_x^\theta\|_{L^4}^4)\, d\theta + \|(\tilde{u}_x - u_x)e^{-|x|}\|_{L^1} \\
&\quad + \|\tilde{u}_x - u_x\|_{L^2} \int_0^1 \|u_x^\theta\|_{L^4}^2\, d\theta + C\|\tilde{u}_x - u_x\|_{L^2} \int_0^1 \|u_x^\theta\|_{L^2}\, d\theta \\
&\quad + C\|\tilde{u}_x - u_x\|_{L^4} \int_0^1 \|u_x^\theta\|_{L^4}^3\, d\theta \\
&\leq C\Big(\|u - \tilde{u}\|_{H^1} + \|(u-\tilde{u})e^{-|x|}\|_{L^1} + \|(u_x - \tilde{u}_x)e^{-|x|}\|_{L^1} + \|u_x - \tilde{u}_x\|_{L^4} \Big).
\end{aligned}
$$

Here, C denotes a generic positive constant depends only on E_1 and E_2. This completes the proof of Proposition 6.1. $\qquad\square$

In the following two propositions, we compare the distance with the L^1 distance and the Kantorovich-Rubinstein or Wasserstein distance.

Proposition 6.2 (Comparison with L^1 metric). *For any solutions $u, \tilde{u} \in H^1(\mathbb{R}) \cap W^{1,4}(\mathbb{R})$ to (2.1), there exists some constant C depends only on E_1 and E_2, such that,*

$$\|(u - \tilde{u})e^{-|x|}\|_{L^1} \leq C \cdot d(u, \tilde{u}). \tag{6.5}$$

Proof. Assume that $\gamma^t : \theta \mapsto u^\theta$ is a regular path connecting u with $\tilde{u}$.

1. Notice that

$$|v| = |v + u_x w - u_x w| \leq |v + u_x w| + |u_x w| \leq |v + u_x w|(1 + u_x^2)^2 + |w|(1 + u_x^2)^2.$$

Thus, by the above inequality, the definition 5.2, (3.9) and (6.4), for some constants $C_3, C_4 > 0$, we have

$$d(u, \tilde{u}) \geq C_3 \inf_{\gamma^t} \int_0^1 \int_{\mathbb{R}} |v^\theta| e^{-|x|} \, dx \, d\theta = C_3 \inf_{\gamma^t} \int_0^1 \int_{\mathbb{R}} |\frac{du^\theta}{d\theta}| e^{-|x|} \, dx \, d\theta \geq C_4 \|(u - \tilde{u}) e^{-|x|}\|_{L^1},$$

which implies (6.5). $\qquad\square$

Proposition 6.3 (Comparison with the Kantorovich-Rubinstein metric). *Consider the same assumptions as in Proposition 6.2, and further assume that $u, \tilde{u} \in L^1(\mathbb{R})$. Then one could drop the $e^{-|x|}$ term in (3.9) to define a new distance d^*, such that*

$$\|(u - \tilde{u})\|_{L^1} \leq C \cdot d^*(u, \tilde{u}), \quad \text{and} \tag{6.6}$$

$$\sup_{\|f\|_{\mathcal{C}^1} \leq 1} \left| \int f \, d\mu - \int f \, d\tilde{\mu} \right| \leq d^*(u, \tilde{u}), \tag{6.7}$$

where $\mu, \tilde{\mu}$ are the measures with densities $(1 + (u_x)^2)^2$ and $(1 + (\tilde{u}_x)^2)^2$ w.r.t the Lebesgue measure, respectively.

Proof. First, repeating the same proof as in Proposition 6.2 without the term $e^{-|x|}$, one can very easily obtain (6.6).

Then, for any function f, such that $\|f\|_{\mathcal{C}^1} \leq 1$, denote μ^θ to be the measure with density $(1 + (u_x^\theta)^2)^2$ w.r.t the Lebesgue measure, then the following holds

$$\left| \int_0^1 \frac{d}{d\theta} \int f \, d\mu^\theta \, d\theta \right| \leq \int_0^1 \int_{\mathbb{R}} |f'| \cdot |w^\theta|(1 + (u_x^\theta)^2)^2 \, dx \, d\theta$$
$$+ \int_0^1 \int_{\mathbb{R}} |f| \cdot |4(u_x^\theta + (u_x^\theta)^3)(v_x^\theta + u_{xx}^\theta w^\theta) + (1 + u_x^\theta)^2)^2 w_x^\theta| \, dx \, d\theta, \tag{6.8}$$

where the two integrals on the right hand side of (6.8) are exactly I_1 and I_4 of (3.9). Hence, we get (6.7) immediately. This completes the proof of Proposition 6.2. $\qquad\square$

The metric (6.7) is usually called a Kantorovich-Rubinstein distance, which is equivalent to a Wasserstein distance by a duality theorem [22].

7. Application to the Camassa-Holm equation

Our method of constructing the geodesic distance can be systematically applied to many other quasi-linear equations. Here we give an example of the Camassa-Holm (CH) equation. The energy measure associated to the CH equation has density u_x^2. Here in this section we will only outline the construction of the Finsler norm for the infinitesimal tangent vector for smooth solutions. The issues on the generic regularity (already given in [20]) and the construction of the geodesic distance can be treated in a similar way as is done for Novikov equation in this paper.

The CH equation reads

$$\begin{cases} u_t + \left(\frac{u^2}{2}\right)_x + P_x = 0, \\ u(0, x) = u_0(x), \end{cases} \tag{7.1}$$

where

$$P := \frac{1}{2} e^{-|x|} * \left(u^2 + \frac{u_x^2}{2}\right).$$

Let $u(x)$ be a smooth solution to (7.1) and consider a family of perturbed solutions

$$u^\epsilon(x) = u(x) + \epsilon v(x) + o(\epsilon).$$

A direct calculation yields that the first order perturbation v satisfies

$$v_t + uv_x + vu_x + \frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right)e^{-|x-y|}(2uv + u_y v_y)(y)\,dy = 0. \tag{7.2}$$

$$v_{xt} + uv_{xx} + u_x v_x + vu_{xx} - 2uv + \frac{1}{2}\int_{\mathbb{R}} e^{-|x-y|}(2uv + u_y v_y)(y)\,dy = 0. \tag{7.3}$$

Similar as before, we introduce a horizontal shift w satisfying

$$x^\epsilon = x + \epsilon w + o(\epsilon)$$

Again the shift component of the tangent vector must propagate along characteristic. Namely

$$w_t + uw_x = v + u_x w. \tag{7.4}$$

Now we can define a Finsler norm

$$\|v\|_u = \inf_{w\in\mathcal{A}} \|(w,\hat{v})\|_u, \quad \text{with} \quad \hat{v} = v + u_x w,$$

where

$$\mathcal{A} = \{\text{solutions } w(t,x) \text{ of } (7.4) \text{ with smooth initial data } w_0(x)\}, \quad \text{and}$$

$$\|(w,\hat{v})\|_u = \int_{\mathbb{R}} \{[\text{change in } x] + [\text{change in } u]\}(1 + u_x^2)\,dx$$
$$+ \int_{\mathbb{R}} [\text{change in the base measure with density } u_x^2]\,dx.$$

More precisely, we have

$$\|v\|_u = \inf_w \int_{\mathbb{R}} \{|w|(1 + u_x^2) + |v + u_x w|(1 + u_x^2) + |2u_x(v_x + u_{xx}w) + u_x^2 w_x|\}e^{-|x|}\,dx$$
$$= \inf_w (I_1 + I_2 + I_3). \tag{7.5}$$

The goal of the forthcoming computations is to validate the estimate

$$\frac{d}{dt}\|v(t)\|_{u(t)} \le C\|v(t)\|_{u(t)}, \tag{7.6}$$

for some constant C depending only on the total energy. As same as before, in the following calculation, we drop the $e^{-|x|}$ terms.

1. To estimate the time derivative of I_1, by (7.1) and (7.4), we have

$$\left(w(1 + u_x^2)\right)_t + \left(uw(1 + u_x^2)\right)_x = (w_t + uw_x)(1 + u_x^2) + w[(1 + u_x^2)_t + (u(1 + u_x^2))_x]$$
$$= (v + u_x w)(1 + u_x^2) + w(u_x + 2u^2 u_x - 2u_x P).$$

This yields the estimate

$$\frac{dI_1}{dt} = \frac{d}{dt}\int_{\mathbb{R}} |w|(1 + u_x^2)\,dx \le \int_{\mathbb{R}} |v + u_x w|(1 + u_x^2)\,dx + C\int_{\mathbb{R}} |w|(1 + u_x^2)\,dx$$
$$\le C(I_2 + I_1). \tag{7.7}$$

2. To estimate the time derivative of I_2, recalling (7.1), (7.2) and (7.4), we obtain

$$
\big((v + u_x w)(1 + u_x^2)\big)_t + \big(u(v + u_x w)(1 + u_x^2)\big)_x
$$
$$
= \left[v_t + u v_x + u_x(w_t + u w_x) + w(u_{xt} + u u_{xx})\right](1 + u_x^2) + (v + u_x w)\left[(1 + u_x^2)_t + (u(1 + u_x^2))_x\right]
$$
$$
= \left[-v u_x - \frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right) e^{-|x-y|}(2uv + u_y v_y)\, dy + u_x(v + u_x w)\right.
$$
$$
\left. + w\left(-\frac{u_x^2}{2} + u^2 - P\right)\right](1 + u_x^2) + (v + u_x w)(u_x + 2u^2 u_x - 2u_x P)
$$
$$
= \left[-\frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right) e^{-|x-y|}(2uv + u_y v_y)\, dy + \frac{1}{2}u_x^2 w + w(u^2 - P)\right](1 + u_x^2)
$$
$$
+ (v + u_x w)(u_x + 2u^2 u_x - 2u_x P).
$$

Note that
$$
2uv = 2u(v + u_y w) - 2u u_y w,
$$

and
$$
\frac{1}{2}u_x^2 w = -\frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right)\left(e^{-|x-y|}\frac{1}{2}u_y^2 w\right)_y dy
$$
$$
= -\frac{1}{2}\left(\int_x^\infty - \int_{-\infty}^x\right) e^{-|x-y|}(u_y u_{yy} w + \frac{1}{2}u_y^2 w_y)\, dy + \frac{1}{2}\int_{\mathbb{R}} e^{-|x-y|}\frac{1}{2}u_y^2 w\, dy.
$$

We thus conclude
$$
\frac{dI_2}{dt} = \frac{d}{dt}\int_{\mathbb{R}} |v + u_x w|(1 + u_x^2)\, dx
$$
$$
\leq C\left(\int_{\mathbb{R}} |2u_x(v_x + u_{xx}w) + u_x^2 w_x|\, dx + \int_{\mathbb{R}} |v + u_x w|(1 + u_x^2)\, dx + \int_{\mathbb{R}} |w|(1 + u_x^2)\, dx\right) \quad (7.8)
$$
$$
\leq C(I_3 + I_2 + I_1).
$$

3. To estimate the time derivative of I_3, using (7.1), (7.3) and (7.4) to get

$$
\big(2u_x(v_x + u_{xx}w) + u_x^2 w_x\big)_t + \big(u(2u_x(v_x + u_{xx}w) + u_x^2 w_x)\big)_x
$$
$$
= 2(u_{xt} + u u_{xx})(v_x + u_{xx}w) + 2u_x\left[v_{xt} + (u v_x)_x\right] + 2u_x u_{xx}(w_t + u w_x)
$$
$$
+ 2u_x w\left[u_{xxt} + (u u_{xx})_x\right] + 2u_x w_x(u_{xt} + u u_{xx}) + u_x^2\left[w_{xt} + (u w_x)_x\right]
$$
$$
= 2\left(-\frac{u_x^2}{2} + u^2 - P\right)(v_x + u_{xx}w + u_x w_x) - 2u_x\left(v u_{xx} - 2uv + \frac{1}{2}\int_{\mathbb{R}} e^{-|x-y|}(2uv + u_y v_y)\, dy\right)
$$
$$
+ 2u_x u_{xx}(v + u_x w) + 2u_x w(-u_x u_{xx} + 2u u_x - P_x) + u_x^2(v_x + u_x w_x + u_{xx}w)
$$
$$
= 2(u^2 - P)(v_x + u_{xx}w + u_x w_x) + 4u u_x(v + u_x w) - u_x\int_{\mathbb{R}} e^{-|x-y|}(2uv + u_y v_y)\, dy - 2u_x w P_x.
$$
$$
\tag{7.9}
$$

The first term in the last equality can be estimated as
$$
2\int_{\mathbb{R}} (u^2 - P)(v_x + u_{xx}w + u_x w_x)\, dx
$$
$$
= 2\int_{\mathbb{R}} (u^2 - P)(v + u_x w)_x\, dx = -2\int_{\mathbb{R}} (u^2 - P)_x(v + u_x w)\, dx \quad (7.10)
$$
$$
\leq C\int_{\mathbb{R}} |v + u_x w|(1 + u_x^2).
$$

For the third term, observe that

$$2uv + u_y v_y = 2u(v + u_y w) - 2u u_y w + \frac{1}{2}\left[2u_y(v_y + u_{yy}w) + u_y^2 w_y\right] - \frac{1}{2}(u_y^2 w)_y. \tag{7.11}$$

We have

$$\left|\frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{-|x-y|}(u_y^2 w)_y \, dy\, dx\right| \le C\int_{\mathbb{R}}\left(\int_{\mathbb{R}}|(e^{-|x-y|})_y|\, dx\right)|w|u_y^2\, dy$$
$$\le C\int_{\mathbb{R}}|w|(1 + u_x^2)\, dx. \tag{7.12}$$

In light of (7.9)–(7.12), we obtain

$$\frac{dI_3}{dt} = \frac{d}{dt}\int_{\mathbb{R}}|2u_x(v_x + u_{xx}w) + u_x^2 w_x|\, dx$$
$$\le C\left(\int_{\mathbb{R}}|2u_x(v_x + u_{xx}w) + u_x^2 w_x|\, dx + \int_{\mathbb{R}}|v + u_x w|(1 + u_x^2)\, dx + \int_{\mathbb{R}}|w|(1 + u_x^2)\, dx\right) \tag{7.13}$$
$$\le C(I_3 + I_2 + I_1).$$

Combining the inequalities (7.7), (7.8) and (7.13) together, we obtain the desired inequality (7.6).

8. INTERACTION OF TWO PEAKONS.

In this last section, we use numeric method to study the interaction of two peakons, in the form of (1.5), for the Novikov equation. Especially, in this example, we show the energy concentration, which indicates that the $W^{1,4}(\mathbb{R})$ space is ill-suited for studying the Lipschitz continuous dependence. This also motivates the necessity of using the transport metric.

As an example, we consider the following two-peakon initial data

$$u(0, x) = \sum_{i=1}^{2} p_i(0)e^{|x - q_i(0)|} \tag{8.1}$$

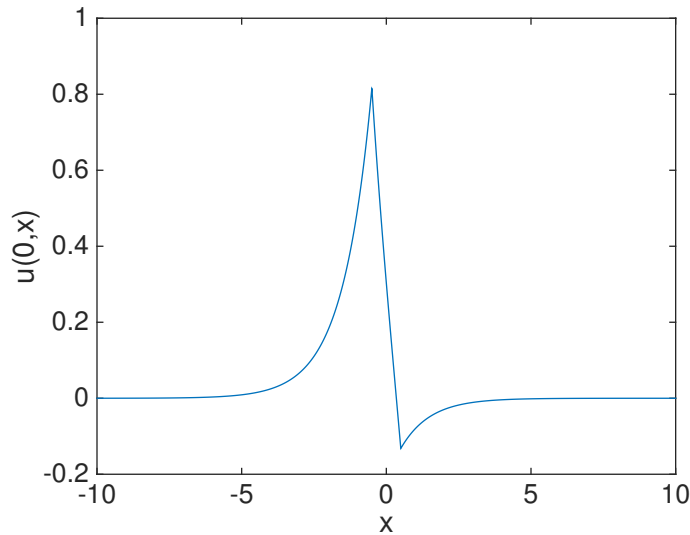where $p_1 = 1$, $p_2 = -0.5$, $q_1 = -0.5$ and $q_2 = 0.5$, as shown in Figure 4.

FIGURE 4. Two peakons: Initial data.

By studying the dynamical system (1.6) of p_i and q_i in a numeric way, we can clearly see, from Figure 5, that two peakons will interact. Similar simulation can be found in [19].

In order to know whether there is any energy concentration when two peakons interact, one needs to study the energies $\mathcal{E}$ and $\mathcal{F}$ in the interval $[q_1(t), q_2(t)]$ near and at the time when two peakons collide. Here $x = q_1(t)$ and $x = q_2(t)$ are two characteristics starting from the tips of two initial peakons, respectively. However, the analysis becomes very hard at the collision if one only uses the system (1.6), because of the blowup of p_i.

Instead, in this paper, we use the semi-linear system established in [9], which dilates the interacting characteristics in the new (t, Y)-coordinates. In fact, integrating (4.2), one has

$$
\begin{cases}
u(t, Y) = u(0, Y) - \int_0^t \partial_x P_1 + P_2 \ dt, \\
\alpha(t, Y) = \alpha(0, Y) + \int_0^t 2u^3 \cos^2 \frac{\alpha}{2} - u \sin^2 \frac{\alpha}{2} - 2 \cos^2 \frac{\alpha}{2}(P_1 + \partial_x P_2) \ dt, \\
\xi(t, Y) = \xi(o, Y) + \int_0^t \xi[(2u^3 + u) - 2(P_1 + \partial_x P_2)] \sin \alpha \ dt.
\end{cases}
\tag{8.2}
$$

We use an iteration:

$$
\begin{cases}
u_{n+1}(t, Y) = u(0, Y) - \int_0^t \partial_x P_1(u_n, \alpha_n, \xi_n) + P_2(u_n, \alpha_n, \xi_n) \ dt, \\
\alpha_{n+1}(t, Y) = \alpha(0, Y) + \int_0^t 2u_n^3 \cos^2 \frac{\alpha_n}{2} - u_n \sin^2 \frac{\alpha_n}{2} - 2 \cos^2 \frac{\alpha_n}{2}(P_1 + \partial_x P_2)(u_n, \alpha_n, \xi_n) \ dt, \\
\xi_{n+1}(t, Y) = \xi(0, Y) + \int_0^t \xi[(2u_n^3 + u_n) - 2(P_1 + \partial_x P_2)(u_n, \alpha_n, \xi_n)] \sin \alpha_n \ dt.
\end{cases}
$$

$$
\tag{8.3}
$$

with initial data

$$
u(0, Y), \quad \alpha(0, Y) \quad \text{and} \quad \xi(0, Y)
$$

calculated by (8.1) and (4.1). Here n is the iteration index. The graphs of u and α are given in Figures 6 and 7, respectively.

The convergence of this algorithm can be proved by a similar method as the one used in [9]. This is one of the greatest advantages of our algorithm.

Then we study the energies $\mathcal{E}$ and $\mathcal{F}$, defined in (2.6)–(2.7), between two characteristics $x = q_1(t)$ and $x = q_2(t)$ starting from the tips of two initial peakons, respectively.

We introduce several notations: First, we denote

$$
Y_1 = Y(0, q_1(0)), \qquad Y_2 = Y(0, q_2(0)),
$$

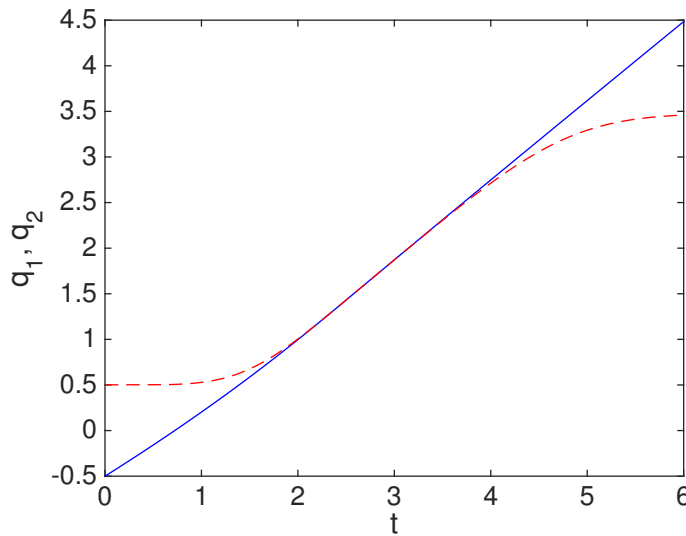FIGURE 5. Interaction of two peakons on (t, x)-plane. The blue line is for the characteristics $q_1(t)$ and the red dashed line is for the characteristics $q_2(t)$.

which are constant on two characteristics, respectively. And we denote $t = t_*$ to be the time when two peakons first interact and

$$q_* := q_1(t_*) = q_2(t_*).$$

By (4.1), (2.6) and (2.7), before the interaction of two peakons, we have

$$\mathcal{F}[q_1(t), q_2(t)] = \int_{[q_1(t), q_2(t)]} (u^4 + 2u^2 u_x^2 - \frac{1}{3} u_x^4)(t, x)\, dx$$

$$= \int_{[Y_1, Y_2]} \left(u^4 \cos^4 \frac{\alpha}{2} + 2u^2 \cos^2 \frac{\alpha}{2} \sin^2 \frac{v}{2} - \frac{1}{3} \sin^4 \frac{\alpha}{2} \right) \xi\, dY$$
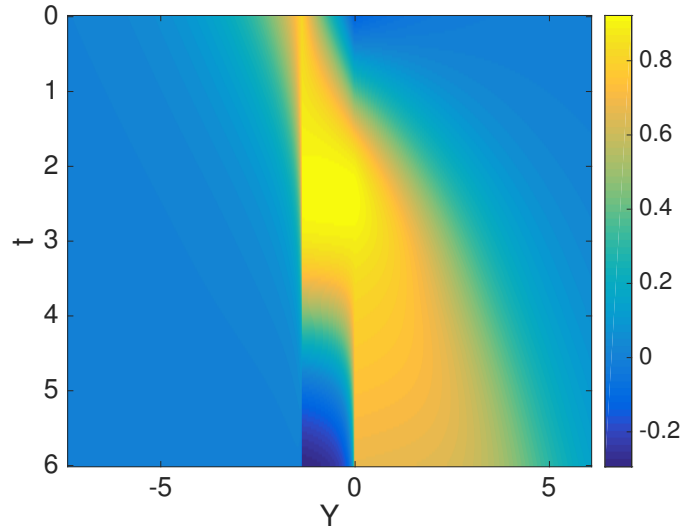
FIGURE 6. Solution $u(t, Y)$.
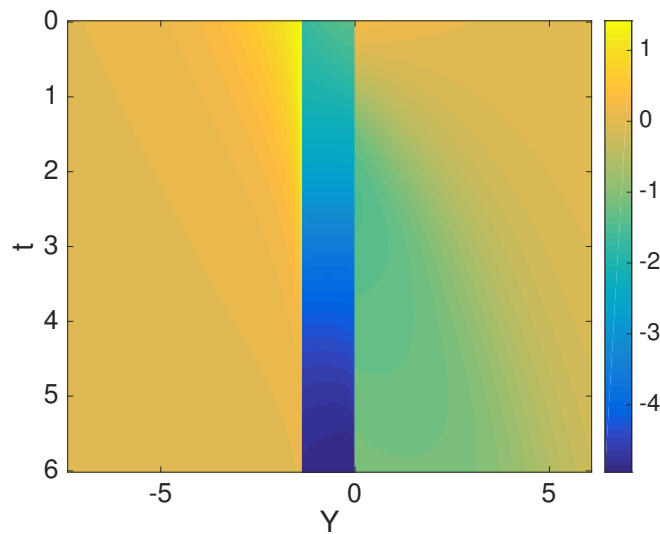
FIGURE 7. Function $\alpha(t, Y)$. When $\alpha = (2k+1)\pi$, with any integer k, $|u_x|$ blowup (when two peakons collide). In this example, α attains $-\pi$.

and

$$\mathcal{L}[q_1(t), q_2(t)] := \int_{[p_1(t), p_2(t)]} u_x^4(t, x) \, dx = \int_{[Y_1, Y_2]} (\sin^4 \frac{\alpha}{2}) \xi \, dY.$$

From Figure 8, we see that $\mathcal{L}[q_1(t), q_2(t)]$ is always positive. And when $t = t_*$,

$$\mathcal{F}[\{q_*\}] = \nu_{t_*}(\{q_*\}) = \int_{[Y_1, Y_2]} \left(u^4 \cos^4 \frac{\alpha}{2} + 2u^2 \cos^2 \frac{\alpha}{2} \sin^2 \frac{v}{2} - \frac{1}{3} \sin^4 \frac{\alpha}{2} \right) \xi(t_*, Y) \, dY$$

and

$$\mathcal{L}[\{q_*\}] = \int_{[Y_1, Y_2]} (\sin^4 \frac{\alpha}{2}) \xi(t_*, Y) \, dY$$

are also nonzero. This means that there exists fourth-order-energy concentration at the point q_* when peakons interact in our example. This exactly shows the failure of applying the natural Sobolev norm from energy law in studying the stability or Lipschitz continuous dependence of solutions, as shown in Figure 1.

Here, we note that by studying the solution on the (t, Y)-coordinates, we can avoid the difficulties caused by the blowup of $|u_x|$. One can clearly see that above integrals in the interval $[Y_1, Y_2]$ are ordinary integrals instead of improper ones.

Similarly, before the interaction of two peakons,

$$\mathcal{E}[q_1(t), q_2(t)] = \int_{[q_1(t), q_2(t)]} (u^2 + u_x^2)(t, x) \, dx = \int_{[Y_1, Y_2]} \left(u^2 \cos^2 \frac{\alpha}{2} + \sin^2 \frac{\alpha}{2} \right) \xi \cos^2 \frac{\alpha}{2} \, dY.$$

However, it is clear that when two peakons interact at time $t = t_*$, $\cos \frac{\alpha}{2} = 0$, hence,

$$\mathcal{E}\{q_*\} = \int_{[Y_1, Y_2]} \left(u^2 \cos^2 \frac{\alpha}{2} + \sin^2 \frac{\alpha}{2} \right) \xi \cos^2 \frac{\alpha}{2}(t_*, Y) \, dY = 0.$$

This tells that there exists no second-order-energy ($\mathcal{E}$) concentration for any weak solutions.

We also provide a picture of $u(t, x)$ in the original (t, x)-coordinates in Figure 9.

References

[1] A. Bressan and G. Chen, Generic regularity of conservative solutions to a nonlinear wave equation, Ann. I. H. Poincaré–AN (2016), http://dx.doi.org/10.1016/j.anihpc.2015.12.004.

[2] A. Bressan and G. Chen, Lipschitz metric for a class of nonlinear wave equations, submitted, available at arXiv: 1506.06310.

[3] A. Bressan, G. Chen, and Q. Zhang, Uniqueness of conservative solutions to the Camassa-Holm equation via characteristics, Discr. Cont. Dynam. Syst., **35** (2015), 25–42.

[4] A. Bressan and A. Constantin, Global conservative solutions to the Camassa-Holm equation, Arch. Rat. Mech. Anal., **183** (2007), 215-239.

[5] A. Bressan and A. Constantin, Global solutions to the Hunter-Saxton equations, SIAM J. Math. Anal., **37** (2005), 996–1026.

[6] A. Bressan and M. Fonte, An optimal transportation metric for solutions of the Camassa-Holm equation, Methods and Applications of Analysis, **12** (2005), 191–220.

[7] A. Bressan, H. Holden, and X. Raynaud, Lipschitz metric for the Hunter-Saxton equation, J. Math. Pures Appl., **94** (2010), 68–92.

[8] R. Camassa and D. D. Holm, An integrable shallow water equation with peaked solitons, Phys. Rev. Lett., **71** (1993), 1661–1664.

[9] G. Chen, R. M. Chen, and Y. Liu, Existence and uniqueness of the global conservative weak solutions for the integrable Novikov equation, arXiv:1509.08569.
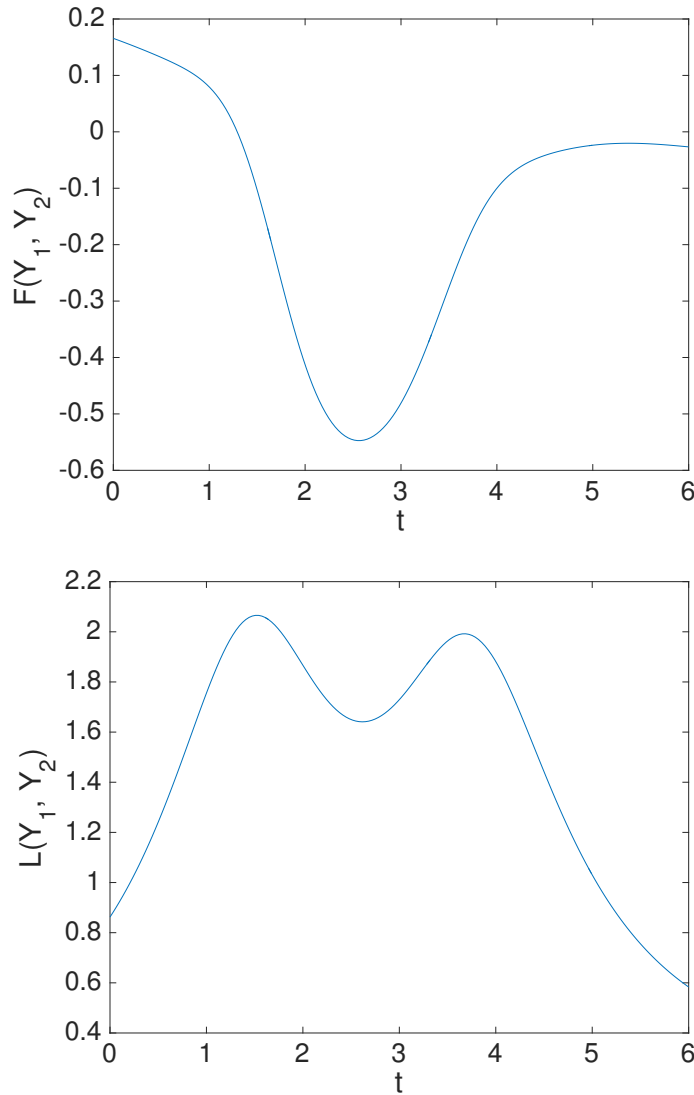
FIGURE 8. The functions $\mathcal{F}[q_1(t), q_2(t)]$ and $\mathcal{L}[q_1(t), q_2(t)]$.

[10] G. CHEN AND Y. SHEN, EXISTENCE AND REGULARITY OF SOLUTIONS IN NONLINEAR WAVE EQUATIONS *Discr. Cont. Dynam. Syst.*, **35** (2015), 3327-3342.

[11] R. M. CHEN, F. GUO, Y. LIU, AND C. Z. QU, ANALYSIS ON THE BLOW-UP OF SOLUTIONS TO A CLASS OF INTEGRABLE PEAKON EQUATIONS, *J. Funct. Anal.*, **270** (2016), 2343–2374.

[12] B. FUCHSSTEINER AND A. S. FOKAS, SYMPLECTIC STRUCTURES, THEIR BÄCKLUND TRANSFORMATIONS AND HEREDITARY SYMMETRIES, *Physica D*, **4** (1981/1982), 47–66.

[13] H. HOLDEN AND X. RAYNAUD, GLOBAL CONSERVATIVE SOLUTIONS OF THE CAMASSA-HOLM EQUATION–LAGRANGIAN POINT OF VIEW, *Comm. Partial Differential Equations*, **32** (2007), 1511–1549.

[14] K. GRUNERT, H. HOLDEN, AND X. RAYNAUD, LIPSCHITZ METRIC FOR THE PERIODIC CAMASSA-HOLM EQUATION, *J. Differential Equations*, **250** (2011), 1460–1492.

[15] K. GRUNERT, H. HOLDEN, AND X. RAYNAUD, LIPSCHITZ METRIC FOR THE CAMASSA-HOLM EQUATION ON THE LINE, *Discrete Contin. Dyn. Syst.*, **33** (2013), 2809–2827.

[16] A. N. W. HONE, H. LUNDMARK, AND J. SZMIGIELSKI, EXPLICIT MULTIPEAKON SOLUTIONS OF NOVIKOVS CUBICALLY NONLINEAR INTEGRABLE CAMASSA-HOLM TYPE EQUATION, *Dynamics of PDEs*, **6** (2009), 253–289.

[17] A. N. W. HONE AND J. WANG, INTEGRABLE PEAKON EQUATIONS WITH CUBIC NONLINEARITY, *J. Phys. A*, **37** (2008) 372002.
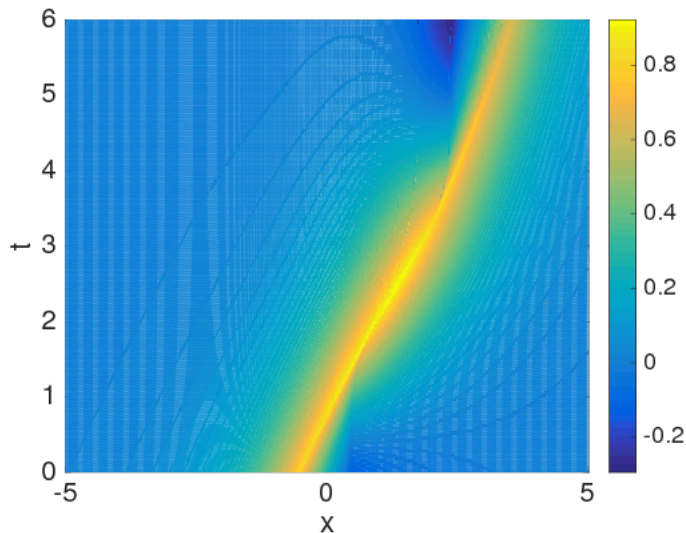
FIGURE 9. Solution $u(t, x)$.

[18] Z. JIANG AND L. NI, BLOW-UP PHENOMENON FOR THE INTEGRABLE NOVIKOV EQUATION, *preprint.*

[19] M. KARDELL, PEAKON-ANTIPEAKON SOLUTIONS OF THE NOVIKOV EQUATION, THESIS.

[20] M. J. LI AND Q. T. ZHANG, GENERIC REGULARITY OF CONSERVATIVE SOLUTIONS TO CAMASSA–HOLM TYPE EQUATIONS. 2015. HAL-01202927.

[21] V. NOVIKOV, GENERALIZATIONS OF THE CAMASSA-HOLM TYPE EQUATION, *J. Phys. A*, **42** (2009) 342002.

[22] C. VILLANI, *Topics in Optimal Transportation.* AMERICAN MATHEMATICAL SOCIETY, PROVIDENCE, 2003.

HONG CAI
DEPARTMENT OF MATHEMATICS, QINGDAO UNIVERSITY OF SCIENCE AND TECHNOLOGY, QINGDAO, SHANDONG, P.R. CHINA, 266061.
E-mail address: caihong19890418@163.com

GENG CHEN
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF KANSAS, LAWRENCE, KS 66045
E-mail address: gengchen@ku.edu

ROBIN MING CHEN
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PITTSBURGH, PITTSBURGH PA 15260
E-mail address: mingchen@pitt.edu

YANNAN SHEN
DEPARTMENT OF MATHEMATICS, CALIFORNIA SATE UNIVERSITY, NORTHRIDGE, CA 91330
E-mail address: yannan.shen@csun.edu