The most general scalar-tensor theory: its generalization and nonlinear wave phenomena

Norihiro Tanahashi*
Department of Physics, Osaka University, Toyonaka, Osaka 560-0043, Japan

A vast number of scalar-tensor theories have been proposed and studied in various contexts of physics, hence it is desirable to have the most general theory that can be used as a framework to treat those theories in a unified manner. In the single scalar field case, the most general scalar-tensor theory called the Horndeski theory played that role, while there were no such general theories in more generic cases where, e.g., more degrees of freedom are present in the theory. To make a progress in this situation, in this essay we propose an extension of the most general scalar-tensor theory with multiple scalar fields, and also examine nonlinear wave phenomena in this theory to elucidate its fundamental properties. The goal of this essay is threefold: i) construct the most general scalar-tensor theory with second-order equations of motion in the two scalar field case, ii) examine propagation of gravitational wave and scalar field wave to clarify the causal structure in this theory, and iii) examine nonlinear effects in wave propagation focusing on the shock formation phenomenon in this theory. Combining findings obtained from these studies, we aim to obtain insights into universal features of various scalar-tensor theories and also to open up brand-new avenues of future investigations.

CONTENTS

* Email: tanahashi“at”het.phys.sci.osaka-u.ac.jp

I. INTRODUCTION

The general relativity (GR), Einstein's theory of gravitation, was proposed as a theory that reduces to Newtonian theory in the limit of weak gravity and slow motion. However, GR is not the unique theory with this property, and actually there are numerous theories of gravitation where GR is modified by introducing additional degrees of freedom such as scalar fields, or by correcting the theory with higher-order corrections. Such modifications to GR typically appear in top-down models where the gravitation theory emerges as a part of the low-energy effective theory. In this kind of construction, various fields other than the gravitational one naturally appear in the theory and we are lead to consider their gravitational effects.

Scalar-tensor theories, in which GR is modified by additional scalar fields, are the simplest among those modified theories of gravitation. Such scalar fields naturally arise in top-down models, and also they play key roles in phenomenological studies since a lot of interesting phenomena can be realized using them. Just to list some of examples, the inflation scenario describing our universe in the earliest era [1] is typically based on a scalar field called inflaton that is coupled to gravity. Also, many attempts have been made to explain the accelerated expansion of our universe at late time [2] by modifying gravitation at cosmologically large scale rather than by attributing it to the cosmological constant, whose origin it yet to be known. Black holes dressed with scalar field cloud (see e.g. [3, 4]) are attracting interest since they could be important targets of various astrophysical observations and particularly the gravitational wave observations, which was recently realized by LIGO [5]. Even in high energy physics and the string theory, black holes with scalar fields plays a crucial role in the context of the gauge-gravity correspondence [6] and their various applications (see e.g. [7–10]).

A plethora of scalar-tensor theories have been proposed by now (see [11] for a review), hence it is desirable to have a theory that could be used as a framework to treat them in a unified manner. Such a theory should be free from pathological behaviors including ghost instability, and also it should encompass a wide variety of scalar-tensor theories as its subclass. In such a context, the Galileon theory [12] in flat spacetime was proposed as a scalar field theory that is free from the Ostrogradsky ghost instability [13, 14] although its Lagrangian contains higher derivative terms. Later on, this theory was covariantized to include dynamical gravity and also was generalized to incorporate more parameters into the theory by [15, 16]. The resultant theory was dubbed the generalized Galileon theory, which was shown by [17] to be equivalent to the Horndeski theory [18] constructed in 1970's. This theory has been utilized in studies on cosmology and gravitational physics, e.g. by [17] which pursued inflation scenario based on it.

The Horndeski theory is the most general covariant scalar-tensor theory with a single scalar field whose Euler-Lagrange equation has derivatives of the metric and the scalar field only up to second order. There are a several ways to extend this theory further. One of the simplest extensions is to incorporate multiple scalar fields into the theory, which can describe inflation models with multiple fields (see [19] for a review). Extensions of this type was first realized by the generalized multi-Galileon theory [20]. Later on it was realized that, however, the multi-DBI inflation models [21–28] are not included in this theory [29], hence this theory is not the most general theory for multiple scalar fields.

This finding tempts us to construct the truly most general scalar-tensor theory with multiple scalar fields, and also to clarify properties of the general theory so that we can make predictions about various multi-field theories incorporated in that general theory. These are the main motivation in this essay, and for the latter half of it we will focus on the nonlinear wave phenomena in the most general theory. We will explain more details below.

Let us state the problem for the first half of the above motivation more precisely. We will try to construct the most general scalar-tensor theory with multiple scalar fields that is described by a covariant Lagrangian and whose equations of motion are up to second order in derivatives of the dynamical variables, which are the metric and the scalar fields. For simplicity, we consider only the case with two scalar fields in four-dimensional spacetime, which may be viewed as a minimal extension of the original Horndeski theory for a single scalar field. We call this theory the bi-Horndeski theory in this work. This construction was attempted in Ref. [30], and in the first part of this essay we will review its outcome.

The goal of the program addressed in this essay is not only to construct the most general scalar-tensor theory, but also to use such a theory as a framework to study the properties of various theories in a unified manner. Some of such properties will be model-dependent, in which case they can be studied only after specifying the model. If there are properties shared universally by various theories, however, they should be found by studying the most general theory as well. Clarification of such universal properties of the most general theory is another goal of this essay.

For this purpose, we will focus on the wave propagation in the most general theory. Wave propagation is one of the most fundamental dynamical processes in a theory, since it governs perturbative dynamics of the theory and also it defines the causal structure upon which the future and past regions in the spacetime are defined. The properties of the wave propagation become important particularly when we consider black hole horizons, which are defined as the boundaries of causal contact for an observer outside the black hole region. One of the aim of this work is to clarify such causal structure and properties of black hole horizons in the most general theory.

Another focus of this essay is the nonlinear effects in wave propagation. In the most general scalar-tensor theory, the propagation speeds of the gravitational wave and scalar field wave may depend on the background and also the amplitude of the waves themselves, unlikely to GR and the canonical scalar field for which the waves just propagate at the light speed. Once we take into account such a nonlinear effect in wave propagation, the waveform may be distorted in time and eventually it may form a shock. The simplest example of such shock formation induced by the nonlinear interaction is found in the wave obeying Burgers' equation $u_{,t} + u\, u_{,x} = 0$, where the nonlinear term $u\, u_{,x}$ causes the waveform distortion and the shock formation (see Fig. 1). Such shock formation takes place also for fields obeying other nonlinear equations [31, 32], compressible perfect fluids [33] and gravitational wave in Lovelock theories in higher dimensions [34].

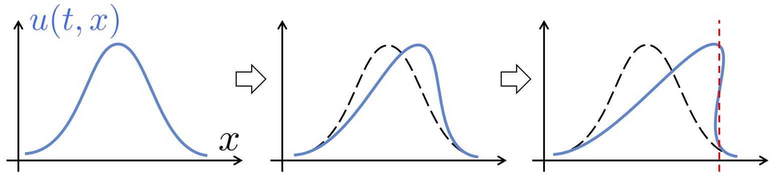


FIG. 1: Schematic of waveform distortion and shock formation in a solution of Burgers' equation. Due to the nonlinear term $u\, u_{,x}$, the waveform is distorted and eventually becomes double-valued, which corresponds to a shock formation.

Since the derivative of the dynamical variable diverges at the shock, this shock formation process can be regarded as a singularity formation process. It may lead to breakdown of the theory unless the singularity is resolved by some mechanisms, and also it could be an important target of astrophysical measurements including the gravitational wave observations. Motivated by these interests, we will examine if the scalar field wave and also the gravitational wave in the most general scalar-tensor theory suffer from this type of effect. Such formation of shock (or caustics) was studied for a probe scalar field with Horndeski-type action on flat spacetime in [35–37].[1] Our aim is to extend these studies by taking into account nontrivial backgrounds and also the dynamics of gravity, aiming to obtain implications to gravitational wave physics. Combining insights obtained from this study and also from the analysis on the wave propagation, we try to elucidate various features of scalar-tensor theories as much and general as possible.

A. Outline and Summary

As stated above, the aim of this essay is to clarify universal features shared by a broad variety of scalar-tensor theories with multiple scalar fields. For this purpose, we first construct the most general scalar-tensor theory with two scalar fields, which we call the bi-Horndeski theory, and then analyze the wave phenomena realized within such a theory. For the latter, we focus on the (linear) wave propagation

[1] Properties of caustics were studied also by [38] in the DBI-type scalar field theories and by [39] in other theories.

and the causality defined by it, and also the nonlinear effects in wave propagation such as the shock formation. Below, we outline our strategies to tackle these problems and summarize the main results. Based on them, we conclude this essay in **section V: Outlook** with discussions on possible future directions. Lengthy formulae used in the main text are summarized in the appendices to make this essay self-contained.

- **Construction of the bi-Horndeski theory** (section II)

  To construct the bi-Horndeski theory, we follow the original construction in the single scalar field case done by Horndeski [18]. We first formulate the problem and the construction strategy in section II A. As explained there, the construction is divided into **Step 1: construction of the most general equations of motion** (sections II B, II C) that is compatible with the covariance of the theory, and **Step 2: construction of the Lagrangian** (section II D) that generates the equations of motion as its Euler-Lagrange equations.

  The key idea to accomplish **Step 1** is to use the generalized Bianchi identity (9), which follows from the general covariance of the theory. We use this identity in two ways as explained at the beginning of section II B. Correspondingly, the construction procedure is divided into **Step 1–a** and **Step 1–b**. Following these steps, we accomplish the construction of the most general equations of motion, where the resultant equation is shown in Eq. (39). We examine their physical properties in section II C, and confirm that the theory defined by those equations of motion is more general than the generalized multi-Galileon theory also the multi-DBI models, as we aimed.

  Having the most general equations of motion at hand, the next task is to tackle **Step 2** for the construction of the Lagrangian. In the case of the Horndeski theory for a single scalar field, it turned out that the trace of the equations of motion is intimately related to the Lagrangian, and based on this fact the full expression of the Lagrangian was correctly found. We follow this procedure in section II D, and obtain a set of partial differential equations (46) and (47) for the arbitrary functions in the candidate Lagrangian. We have to solve these differential equations to construct the Lagrangian, but unfortunately it has not accomplished yet thus far. However, results up to here would serve as a stepping stone for further trials for the theory construction, and also it may be possible even to solve those equations if we take into account the integrability conditions for Lagrangian, as we mention in section V. Such trials toward the Lagrangian construction is one of the future tasks in this project.

- **Wave propagation in the bi-Horndeski theory** (section III)

  Based only on the equations of motion of the bi-Horndeski theory, we can study the dynamics in this theory without directly referring to the Lagrangian, which is yet to be constructed. Taking advantage of this fact, we study the wave propagation in the bi-Horndeski theory aiming to elucidate some properties that are shared universally by the theories incorporated in the bi-Horndeski theory.

  Our analysis is based on the method of characteristics [40], which is a mathematical tool to find the maximum propagation speed allowed in the theory, or in other words the propagation speed in the short wavelength limit. After introducing this method for a generic equation of motion in section III A, we apply it to the following two problems, both of which are related to the causal structure implemented in this theory.

Black hole horizon and causal edge (section III B 1): In GR with a minimally-coupled canonical scalar field, the gravitational wave and the scalar field wave propagate at the speed of light, and then it is ensured a black hole horizon in the metric becomes an event horizon for those waves. However, once the theory is modified, propagation speeds of the waves may be modified as well. When they becomes faster than the light speed, the event horizon for these waves may differ from the metric horizon, which is defined in accord with the light propagation. Motivated by this point, we examine properties of Killing horizons, such as black hole horizons in stationary spacetimes, in section III B 1 based on the method of characteristics. As argued around Eqs. (64) and (66), it turns out that the Killing horizon becomes an event horizon for gravitational and scalar field waves if the scalar fields share the symmetry of the background spacetime.[2] These conditions will provide a useful way to clarify causal structures of black hole solutions realized within the bi-Horndeski theory and various multi-scalar-tensor theories incorporated in this theory.

Wave propagation on plane wave solution (section III B 2): To study the wave propagation and causal structure in this theory in a more general situation, we study another example of nontrivial background solution. For this purpose we focus on the plane wave solution [45], which is an exact solution describing gravitational wave and scalar field wave propagating in a common null direction. Unfortunately, this solution had been constructed only in the (single scalar) Horndeski theory that is invariant against a constant shift in the scalar field (see Eq. (67) for definition). Hence we will limit our analysis to this subclass of the bi-Horndeski theory below. In section III B 2, after briefly introducing the Horndeski theory that is shift-symmetric in scalar field, we analyze the propagation speeds and causal structure for fluctuations on the plane wave solution background. We will find that effective metrics (87), which differ from the physical metric, can be defined in this case, and the propagation of the gravitational wave and scalar field wave obey those effective metrics. Hence the cones of the wave propagation surfaces differ from the light cone, and it turns out that they form a nested cones that are aligned to the background null direction (Fig. 4). Hence, on the plane wave background, the causality can be defined in a usual manner if we use the largest cone as the causal boundary instead of the light cone.

- **Shock formation in the Horndeski theory** (section IV)

  In the above studies on the wave propagation, we were basically focusing on the behavior of linear waves. In this section, we aim to go beyond by taking into account the nonlinear effects in the wave propagation. As mentioned earlier, we particularly focus on the shock formation phenomenon that is caused by such nonlinear effects.

  In **section IV A: Formalism for shock formation**, we summarize the formalism to study shock formation for a generic equations of motion, which was originally introduced in [46] for relativistic hydrodynamics and utilized by [34] to study shock formation of gravitational wave in Lovelock theories. This analysis focuses on the transport of the weak discontinuity introduced to second derivatives of dynamical variables, as depicted in Fig. 5. It turns out that the transport equation (Eq. (96)) is given by a nonlinear equation of the amplitude of the discontinuity ($\Pi(x^\mu)$ of Eq. (90)). The nonlinear

[2] This result is intimately related to the properties of Lovelock theories of gravity [41] and scalar-tensor theories with a non-minimally coupled scalar field studied in Refs. [42–44]. In Lovelock theories, it was shown that the superluminal propagation of the gravitational wave is prohibited on a Killing horizon [42, 43], hence it becomes an event horizon automatically. In the scalar-tensor theories with non-minimal coupling, it was shown that the scalar field must obey the background symmetry so that a Killing horizon to be an event horizon [44].

term is proportional to a coefficient $\mathcal{N}$ (Eq. (94)), which is given by a derivative of principal symbol P (essentially the coefficients of the kinetic term) of the equations of motion. Roughly speaking, this quantity governs the dependence of the propagation speed on the wave's amplitude. When $\mathcal{N}$ is nonzero, two nearby wave trajectories cross each other, and the amplitude of the wave diverges there. It corresponds to a shock formation. Such scenario is depicted by the solution (97) of the transport equation, from which we can find that the amplitude of the weak discontinuity diverges within finite time when $\mathcal{N} \neq 0$. [3]

In this essay, we apply this formalism to the shift-symmetric Horndeski theory, which was studied in section III B 2 as well. As explained above, some explicit solutions are known in this subclass of the most general theory, and they can be used as the background solutions to study the shock formation. The results of this analysis is summarized in **section IV B: Shock formation in the shift-symmetric Horndeski theory**. It turns out that the so-called k-essence model [47] coupled to GR can be studied without specifying the background solutions, and it is found that the gravitational wave in this theory is always free from the shock formation, while the scalar field wave suffers from it unless the model reduces to the scalar field version of the DBI model [48]. For a probe scalar field on flat spacetime, this property of the scalar DBI model was found by Refs. [35–37] and also in earlier works [49]. We confirm that this property of the scalar field wave persists even when we take the gravity sector into account. Also, the gravitational wave was not studied in these works, and it is shown for the first time by our work that it is free from the shock formation in this class of the theory.

For theories more general than the above, we need to specify the background solution to make progress. Such studies on some given background solutions are summarized in **section IV C: Examples of shock formation**, where we focus on the plane wave solution and also the dynamical solution with symmetries in angular directions in sections IV C 1 and IV C 2, respectively. The background solution of this type can describe, e.g., the spatially homogeneous and isotropic universe, that is, the Friedmann-Robertson-Walker (FRW) universe, and also spherically-symmetric dynamical stars (see Fig. 6). It turns out that, in any of these examples, the gravitational wave is always free from the shock formation, while the scalar field wave generically suffers from it unless the theory is set to the scalar DBI model coupled to GR. This result may be suggesting that, at least on the backgrounds with some symmetries, the gravitational wave in the most general scalar-tensor theory is more well-behaved compared to the scalar field wave, in the sense that it is protected against the shock formation.

Notation. Before starting our discussion, we summarize the notations used throughout this paper. We consider a four-dimensional spacetime with a metric g_{ab} and two scalar fields ϕ^I with $I = 1, 2$. Following [18], partial derivatives of g_{ab} and ϕ^I with respect to x^a are denoted as

$$g_{ab,c} \equiv \frac{\partial g_{ab}}{\partial x^c}, \qquad \phi^I_{,a} \equiv \frac{\partial \phi^I}{\partial x^a}, \tag{1}$$

while the covariant derivative of ϕ^I and its scalar product are denoted as

$$\phi^I_{|a} \equiv \nabla_a \phi^I, \qquad X^{IJ} \equiv -\frac{1}{2}\phi^I_{|a}\phi^{J|a}, \tag{2}$$

[3] There are certain theories for which $\mathcal{N}$ identically vanishes and the transport equation becomes linear, in which case the theory is called exceptional or linearly degenerate [46]. For example, GR coupled to a canonical scalar field is an exceptional theory, hence the gravitational wave and scalar field wave in this theory are guaranteed to be free from the shock formation.

where X^{IJ} is symmetric in I and J. We use a strike " $|$ " also as a separator in (anti-)symmetrization unless it causes ambiguity. For example, $[I|JK,L|M]$ stands for anti-symmetrization of I and M. Partial derivatives of a function $A^{a\dots b}(g,\partial g,\partial^2 g,\phi^I,\partial\phi^I,\partial^2\phi^I)$ are expressed as

$$
\begin{aligned}
A^{a\dots b;cd} &\equiv \frac{\partial A^{a\dots b}}{\partial g_{cd}}, \quad A^{a\dots b;cd,e} \equiv \frac{\partial A^{a\dots b}}{\partial g_{cd,e}}, \quad A^{a\dots b;cd,ef} \equiv \frac{\partial A^{a\dots b}}{\partial g_{cd,ef}}, \\
A^{a\dots b;}{}_{I} &\equiv \frac{\partial A^{a\dots b}}{\partial \phi^I}, \quad A^{a\dots b;c}{}_{I} \equiv \frac{\partial A^{a\dots b}}{\partial \phi^I_{,c}}, \quad A^{a\dots b;cd}{}_{I} \equiv \frac{\partial A^{a\dots b}}{\partial \phi^I_{,cd}},
\end{aligned}
$$

and partial derivatives of a function $A(\phi^I, X^{JK})$ are expressed as

$$
A_{,I} \equiv \frac{\partial A}{\partial \phi^I}, \qquad A_{,IJ} \equiv \frac{1}{2}\left(\frac{\partial A}{\partial X^{IJ}} + \frac{\partial A}{\partial X^{JI}}\right). \tag{3}
$$

In the equations of motion and the Lagrangian, we use the generalized Kronecker delta defined by

$$
\delta^{a_1\dots a_n}_{b_1\dots b_n} \equiv n!\,\delta^{a_1}_{[b_1}\dots\delta^{a_n}_{b_n]}, \qquad \delta^{IK}_{JL} \equiv 2\delta^{I}_{[J}\delta^{K}_{L]}. \tag{4}
$$

Repeated indices are summed over $a = 0,1,2,3$ and $I = 1,2$. We also use various indices summarized in Table I for the analysis in sections III and IV, since we need to decompose the spacetime according to the wave propagation surface and also the background symmetry directions. See also the definitions in each section for more details.

TABLE I: Indices used in this work.

$a,b,c,d,\dots,\ q,r,s,t$	Four-dimensional indices for $x^{a=0,1,2,3}$
$\mu,\nu,\dots$	Three-dimensional indices for $x^{\mu=1,2,3}$ on the hypersurface Σ at $x^0=0$
$\alpha,\beta,\dots$	Two-dimensional spatial indices for $x^{\alpha=2,3}$ in section III, and for angular directions in section IV C 2
i,j	Two-dimensional spatial indices of the null basis in section III B 2
$A,B,\dots$	Two-dimensional indices for $x^A=\tau,\chi$ in section IV C 2
$I,J,\dots$	Scalar field indices of the bi-Horndeski theory $\phi_{I=1,2}$; used also as generic indices (e.g. in Eq. (48))

II. CONSTRUCTION OF THE BI-HORNDESKI THEORY

We summarize the method to construct the most general scalar-tensor theory with two scalar fields whose Euler-Lagrange equations are up to second order in the derivatives. We will closely follow the method of Ref. [18] for the single scalar field case, hence the procedure below reduces to [18] once the number of the scalar field is reduced from two to one. In this section, we show how this method can be generalized into the two scalar field case.

A. Assumptions and construction strategy

The problem we consider is defined as follows. We consider a generally-covariant theory described by a Lagrangian scalar density $\mathcal{L}$, which is a function of a metric g_{ab}, two scalar fields $\phi^{I=1,2}$, and their derivatives of arbitrary order. Hence $\mathcal{L}$ in our problem is given by

$$\mathcal{L} = \mathcal{L}\left(g_{ab}\,, g_{ab,c}\,, g_{ab,cd}\,, \ldots\,, \phi^I\,, \phi^I_{,a}\,, \phi^I_{,ab}\,, \ldots\right). \tag{5}$$

We also demand that the equations of motion derived as the Euler-Lagrange equations for $\mathcal{L}$ are given by derivatives of the field variables up to second order:

$$0 = \frac{\delta\mathcal{L}}{\delta g_{ab}} = \sqrt{-g}\,\mathcal{G}^{ab}\left(g_{cd}\,, g_{cd,e}\,, g_{cd,ef}\,, \phi^J\,, \phi^J_{,c}\,, \phi^J_{,cd}\right), \tag{6}$$

$$0 = \frac{\delta\mathcal{L}}{\delta \phi^I} = \sqrt{-g}\,\mathcal{E}_I\left(g_{ab}\,, g_{ab,c}\,, g_{ab,cd}\,, \phi^J\,, \phi^J_{,a}\,, \phi^J_{,ab}\right). \tag{7}$$

Then the problem we try to solve is to construct the most general Lagrangian $\mathcal{L}$ (5) that fulfills the requirements (6) and (7).

In the single scalar field case, such a theory was constructed by Horndeski [18]. The construction procedure is divided into the following two steps:

Step 1. Construct the most general equations of motion that obey the identity following from the general covariance of the theory.

Step 2. Construct the Lagrangian that generates the equations of motion obtained in Step 1.

In this work, we employ the same construction strategy, generalizing it for the two scalar field case.

B. Construction of the most general equations of motion

To accomplish Step 1 in the above, we utilize the identity following from the general covariance of the theory. Since the theory we consider is generally covariant, the action must be invariant under a coordinate transformation $x^a \to x^a + \xi^a$, that is,

$$0 = \delta\int d^4x\,\mathcal{L} = 2\int d^4x\sqrt{-g}\left(\nabla_b\mathcal{G}^{ab} - \frac{1}{2}\mathcal{E}_I\nabla^a\phi^I\right)\xi_a. \tag{8}$$

Since this equation must be satisfied for any ξ_a, it implies an identity given by

$$\nabla_b\mathcal{G}^{ab} = \frac{1}{2}\mathcal{E}_I\nabla^a\phi^I. \tag{9}$$

This equation is a generalization of the contracted Bianchi identity in GR, $\nabla_a G^{ab} = 0$ for the Einstein tensor $G_{ab} \equiv R_{ab} - \frac{1}{2}Rg_{ab}$.

Identity (9) constraints the structure of $\mathcal{G}_{ab}$ and $\mathcal{E}_I$ in the following two ways. First, it forbids $\nabla_b\mathcal{G}^{ab}$ to depend on the third derivatives of the fields, because the right-hand side is manifestly up to second order in derivatives. Since $\nabla_b\mathcal{G}^{ab}$ may have such third derivative terms in general as we can see from Eq. (6), this property gives a nontrivial constraint on $\mathcal{G}_{ab}$. Second, identity (9) demands that the divergence $\nabla_b\mathcal{G}^{ab}$ is proportional to a gradient of scalar field $\nabla^a\phi^I$ and the components in any other directions must vanish. This property gives further constraints on the structure of $\mathcal{G}^{ab}$. Hence, it is useful to subdivide Step 1 of the construction procedure into the following two steps:

Step 1–a. Construct the most general two-tensor $\mathcal{G}_{ab}$ of the form of Eq. (6) whose divergence is up to second order in derivatives.

Step 1–b. Require that $\nabla_b\mathcal{G}^{ab}$ is proportional to $\nabla^a\phi^I$, to restrict further the form of $\mathcal{G}_{ab}$.

Below, we will follow these steps to construct the most general $\mathcal{G}_{ab}$.

Step 1–a. *The most general tensor whose divergence is second order in derivatives*

Step 1–a of the construction is based on the property that $\nabla_b \mathcal{G}^{ab}$ must be free from third derivatives of g_{ab} and ϕ^I, which may be expressed as

$$\frac{\partial \nabla_b \mathcal{G}^{ab}}{\partial g_{cd,efg}} = 0, \qquad \frac{\partial \nabla_b \mathcal{G}^{ab}}{\partial \phi^I_{,cde}} = 0. \tag{10}$$

Applying the chain rule to the derivatives of $\mathcal{G}_{ab}$ and $\mathcal{E}_I$, which are given by Eqs. (6) and (7), we can show that the above condition is equivalent to

$$\mathcal{G}^{ab;cd,ef} + \mathcal{G}^{ae;cd,fb} + \mathcal{G}^{af;cd,be} = 0, \qquad \mathcal{G}^{ab;cd}_{\ \ I} + \mathcal{G}^{ac;db}_{\ \ I} + \mathcal{G}^{ad;bc}_{\ \ I} = 0. \tag{11}$$

Adding to this one, we have another type of symmetry in the indices that follows from the fact that $\mathcal{L}$ behaves as a scalar density [50, 51]

$$\mathcal{G}^{ab;cd,ef} + \mathcal{G}^{ab;ce,fd} + \mathcal{G}^{ab;cf,de} = 0. \tag{12}$$

We also have obvious symmetries such as $\mathcal{G}^{ab;cd}_{\ \ I} = \mathcal{G}^{ba;cd}_{\ \ I} = \mathcal{G}^{ab;dc}_{\ \ I}$. Then, combining (11) and (12), we can show that $\mathcal{G}^{ab;cd,ef}$ and $\mathcal{G}^{ab;cd}_{\ \ I}$ are symmetric under swapping of the sets of two indices:

$$\mathcal{G}^{ab;cd,ef} = \mathcal{G}^{cd;ab,ef} = \mathcal{G}^{ef;cd,ab}, \qquad \mathcal{G}^{ab;cd}_{\ \ I} = \mathcal{G}^{cd;ab}_{\ \ I}. \tag{13}$$

Due to this symmetry and the property (11), in four-dimensional spacetime a derivative of $\mathcal{G}^{ab}$ with respect to $g_{cd,ef}$ and ϕ^I_{cd} identically vanishes once it has nine or more tensor indices. It is because one of the indices appears three times at least in that case, and then the property (11) or (12) asserts that the derivative must vanish. This fact implies the following identities:

$$\frac{\partial}{\partial g_{cd,ef}} \frac{\partial}{\partial g_{ij,kl}} \mathcal{G}^{ab} = 0, \tag{14}$$

$$\frac{\partial}{\partial g_{cd,ef}} \frac{\partial}{\partial \phi^I_{,ij}} \frac{\partial}{\partial \phi^J_{,kl}} \mathcal{G}^{ab} = 0, \tag{15}$$

$$\frac{\partial}{\partial \phi^I_{,cd}} \frac{\partial}{\partial \phi^J_{,ef}} \frac{\partial}{\partial \phi^K_{,ij}} \frac{\partial}{\partial \phi^L_{,kl}} \mathcal{G}^{ab} = 0. \tag{16}$$

Identities (14)–(16) can be integrated to give $\mathcal{G}^{ab}$ as follows. First, integrating Eq. (14) yields[4]

$$\mathcal{G}^{ab} = \tilde{\xi}^{abcdef} g_{cd,ef} + \tilde{\xi}^{ab} = \xi^{abcdef} R_{cdef} + \xi^{ab}, \tag{17}$$

where ξ^{abcdef} and ξ^{ab} are functions of $g_{ab}, g_{ab,c}, \phi^I, \phi^I_{,a}, \phi^I_{,ab}$ that have the symmetries same as Eqs. (11)–(13), that is, they are symmetric with respect to swapping of sets of two indices and vanish if three indices are symmetrized. Substituting Eq. (17) into Eq. (15), and integrating it, we obtain

$$\mathcal{G}^{ab} = \xi^{abcdefgh}_I R_{cdef} \phi^I_{|gh} + \xi^{abcdef} R_{cdef} + \xi^{ab}, \tag{18}$$

where $\xi^{abcdefgh}$ and $\xi^{abcdefgh}_I$ are functions of $g_{ab}, g_{ab,c}, \phi^I, \phi^I_{,a}$, while ξ^{ab} is a function of g_{ab}, $g_{ab,c}$, ϕ^I, $\phi^I_{,a}$, $\phi^I_{,ab}$. It can be also seen that all the ξ tensors have the symmetries of Eqs. (11)–(13). Repeating the same procedure for Eq. (16), we find

$$\mathcal{G}^{ab} = \xi^{abcdefgh}_I R_{cdef} \phi^I_{|gh} + \xi^{abcdefgh}_{IJK} \phi^I_{|cd} \phi^J_{|ef} \phi^K_{|gh} + \xi^{abcdef} R_{cdef} + \xi^{abcdef}_{IJ} \phi^I_{|cd} \phi^J_{|ef} + \xi^{abcd}_I \phi^I_{|cd} + \xi^{ab}, \tag{19}$$

[4] The second equality of Eq. (17) can be shown using the identity $\tilde{\xi}^{abcdef} g_{cd,ef} = \frac{2}{3} \tilde{\xi}^{abcdef} R_{ecdf} + \bar{\xi}^{ab}$ for some tensor $\bar{\xi}^{ab}$, which follows from the definition of the Riemann tensor and the fact that $\tilde{\xi}^{abcdef}$ has the symmetries of Eqs. (11)–(13).

where all of the above ξ tensors are composed of $g_{ab}, g_{ab,c}, \phi^I, \phi^I_{,a}$, and have the symmetries of Eqs. (11)–(13).

The tensor (19) contains the ξ tensors, which appeared as integration constants in the above derivation. They are functions of $g_{ab}, g_{ab,c}, \phi^I, \phi^I_{,a}$, and have the symmetries of Eqs. (11)–(13). The next task to accomplish Step 1–a is to construct those ξ tensors with the most generality, using which we can find the explicit form of the most general two-tensor $\mathcal{G}^{ab}$ with the desired properties.

To construct such ξ tensors, we can use $\phi^{I|a}$, g^{ab}, and the totally antisymmetric tensor ε^{abcd} as the building blocks. For the simplest ξ tensor, ξ^{ab}, it is easy to find the most general form of it:

$$\xi^{ab} = a(\phi^I, X^{JK}) g^{ab} + b_{IJ}(\phi^I, X^{JK}) \phi^{I|a}\phi^{J|b}, \tag{20}$$

where $a(\phi^I, X^{JK})$ and $b^{IJ}(\phi^I, X^{JK})$ are arbitrary functions of ϕ^I and X^{JK} satisfying $b^{IJ} = b^{JI}$. Here, we have used the assumption that the number of the scalar fields is two for the first time in the derivation.[5] The explicit forms of $\xi^{abcd}_I, \xi^{abcdef}, \xi^{abcdef}_{IJ}, \xi^{abcdefgh}$, and $\xi^{abcdefgh}_I$ for the bi-scalar case are given in appendix A 1. Substituting all the ξ tensors into Eq. (19) and rearranging it, we arrive at the most general tensor $\mathcal{G}^{ab}$ that remains to be second order in derivatives after taking a gradient:

$$\begin{aligned}
\mathcal{G}^a_b = {} & A\delta^a_b + B_{IJ}\phi^{I|a}\phi^J_{|b} + C_I \delta^{ac}_{bd}\phi^{I|d}_{|c} + D_{IJK}\delta^{ace}_{bdf}\phi^I_{|c}\phi^{J|d}\phi^{K|f}_{|e} + E_{IJKLM}\delta^{aceg}_{bdfh}\phi^I_{|c}\phi^{J|d}\phi^K_{|e}\phi^{L|f}\phi^{M|h}_{|g} \\
& + F_{IJKLM}\delta^{aceg}_{bdfh}\left(\varepsilon_{cepq}\phi^{I|p}\phi^{J|q}\phi^{K|d}\phi^{L|f} + \phi^I_{|c}\phi^J_{|e}\varepsilon^{dfpq}\phi^K_{|p}\phi^L_{|q}\right)\phi^{M|h}_{|g} + G_{IJ}\delta^{ace}_{bdf}\phi^{I|d}_{|c}\phi^{J|f}_{|e} \\
& + H_{IJKL}\delta^{aceg}_{bdfh}\phi^I_{|c}\phi^{J|d}\phi^{K|f}_{|e}\phi^{L|h}_{|g} + I\delta^{ace}_{bdf}R_{ce}{}^{df} + J_{IJ}\delta^{aceg}_{bdfh}\phi^I_{|c}\phi^{J|d}R_{eg}{}^{fh} + K_I\delta^{aceg}_{bdfh}\phi^{I|d}_{|c}R_{eg}{}^{fh} \\
& + L_{IJK}\delta^{aceg}_{bdfh}\phi^{I|d}_{|c}\phi^{J|f}_{|e}\phi^{K|h}_{|g},
\end{aligned} \tag{21}$$

where $A, B_{IJ}, C_I, D_{IJK}, E_{IJKLM}, F_{IJKLM}, G_{IJ}, H_{IJKL}, I, J_{IJ}, K_I$, and L_{IJK} are arbitrary functions of ϕ^I and X^{IJ} satisfying

$$\begin{aligned}
& B_{IJ} = B_{JI}, \quad D_{IJK} = D_{JIK}, \quad E_{IJKLM} = -E_{KJILM} = -E_{ILKJM} = E_{JILKM}, \\
& G_{IJ} = G_{JI}, \quad H_{IJKL} = H_{JIKL} = H_{IJLK}, \quad F_{IJKLM} = -F_{JIKLM} = -F_{IJLKM} = F_{JILKM}, \\
& J_{IJ} = J_{JI}, \quad L_{IJK} = L_{JIK} = L_{IKJ}.
\end{aligned} \tag{22}$$

Step 1–b. *Imposing* $\nabla_b\mathcal{G}^{ab} \propto \nabla^a\phi^I$

Step 1–b of the construction procedure is to demand that the tensor (21) obeys Eq. (9), that is, the divergence $\nabla_b\mathcal{G}^{ab}$ calculated from Eq. (21) becomes proportional to the gradient of the scalar field $\nabla^a\phi^I$. This condition is not satisfied in general, and to satisfy it the arbitrary functions in (21) must be related with each other in a certain way. We will find constraints on the form of $\mathcal{G}_{ab}$ in this manner below.

Strategy to find the constraints from Eq. (9) is to calculate $\nabla_b\mathcal{G}^{ab}$ from Eq. (21) first, and then to project it onto vectors orthogonal to $\mathcal{E}_I\nabla^a\phi^I$ next. Such projected components must vanish if $\nabla_b\mathcal{G}^{ab}$ is proportional to $\nabla^a\phi^I$, which become the constraints to be imposed. What we do below is to find those constraints starting from Eq. (21).

From Eq. (21), $\nabla_b\mathcal{G}^{ab}$ is calculated as

$$\nabla^b\mathcal{G}^a_b = \mathcal{Q}_I\phi^{I|a} + \alpha_{IJ}\delta^{ace}_{bdf}\phi^{I|d}_{|c}\phi^{J|l}R_{el}{}^{bf} + \beta_{IJ}\delta^{ace}_{bdf}\phi^I_{|l}\phi^{J|lb}R_{ce}{}^{df} + \gamma_{IJKL}\delta^{ace}_{bdf}\phi^K_{|l}\phi^{L|lb}\phi^I_{|c}\phi^{J|m}R_{em}{}^{df}$$

[5] If the number of the scalar fields were greater than two, we would have for example the term such as $c_{IJKL}\left(\varepsilon^{acde}\phi^I_{|c}\phi^J_{|d}\phi^K_{|e}\phi^{L|b} + \varepsilon^{bcde}\phi^I_{|c}\phi^J_{|d}\phi^K_{|e}\phi^{L|a}\right)$ in ξ^{ab}, where c_{IJKL} is arbitrary functions of ϕ^I and X^{IJ}, and totally anti-symmetric in I, J and K.

$$
\begin{aligned}
&+\epsilon_{IJK}\delta^{acel}_{dfhm}\phi^{K|d}_{|c}\phi^{I|f}_{|e}\phi^{J|g}R_{gl}{}^{hm}+\mu_I\delta^{ac}_{bd}R_{cl}{}^{bd}\phi^{I|l}+\nu_{IJKL}\delta^{ace}_{bdf}\phi^{K}_{|l}\phi^{L|lb}\phi^{I|d}_{|c}\phi^{J|f}_{|e}\\
&+\omega_{IJK}\delta^{ac}_{bd}\phi^{J}_{|l}\phi^{K|lb}\phi^{I|d}_{|c}+\xi_{IJ}\phi^{I|l}\phi^{J|a}_{|l}+\zeta_{I[JK]}\delta^{ac}_{bd}\phi^{I}_{|c}\phi^{J|l}\phi^{K|m}R_{lm}{}^{bd}+\iota_{IJK}\delta^{aceg}_{bdfh}\phi^{J}_{|l}\phi^{K|lb}\phi^{I|d}_{|c}R_{eg}{}^{fh}\\
&+2\eta_{I[J|K|L]}\delta^{ace}_{bfh}\phi^{I}_{|c}\phi^{J|g}\phi^{K|f}_{|e}\phi^{L|l}R_{gl}{}^{bh}+\left(\lambda_{IJKLM}-\lambda_{ILMJK}\right)\delta^{ac}_{bd}\phi^{L}_{|e}\phi^{M|eb}\phi^{I}_{|c}\phi^{J|f}\phi^{K|d}_{|f}\\
&+\sigma_{IJKLMN}\delta^{acg}_{bfh}\phi^{M}_{|l}\phi^{N|lb}\phi^{I}_{|c}\phi^{J|e}\phi^{K|f}_{|e}\phi^{L|h}_{|g}+\frac{3}{2}\tau_{IJKLM}\delta^{aceg}_{bdfh}\phi^{L|l}\phi^{M|b}_{|l}\phi^{I|d}_{|c}\phi^{J|f}_{|e}\phi^{K|h}_{|g}\\
&+2F_{IJKLM}\left(\varepsilon_{bdfh}\phi^{I|a}\phi^{J|g}\phi^{K|d}\phi^{L|f}+\varepsilon^{aceg}\phi^{I}_{|c}\phi^{J}_{|e}\phi^{K}_{|b}\phi^{L}_{|h}\right)\phi^{M|l}R_{gl}{}^{bh}\\
&+4\varepsilon_{bdfh}\left(F_{IJKLM}\phi^{I|a}\phi^{J|g}\right)^{|b}\phi^{K|d}\phi^{L|f}\phi^{M|h}_{|g}+4\varepsilon^{aceg}\left(F_{IJKLM}\phi^{I}_{|c}\phi^{J}_{|e}\phi^{K}_{|b}\phi^{L}_{|h}\right)^{|b}\phi^{M|h}_{|g},
\end{aligned}\tag{23}
$$

where the coefficients $\alpha_{IJ},\beta_{IJ},\ldots$ are functions of ϕ^I, X^{JK} made from the arbitrary functions in Eq. (21). Their expressions and also the definitions of $\mathcal{Q}_I$ are given in appendices A 2 and A 3, respectively.

As the next step, we need to project (23) to vectors orthogonal to $\nabla^a\phi^I$, and find necessary and sufficient conditions on the coefficient functions to make them vanishing for any configuration of the scalar fields ϕ^I and the spacetime curvature R^{ab}_{cd}. For this purpose, we classify the terms in (23) according to the numbers of $\phi^{I|a}_{|b}$ and R^{ab}_{cd} appearing in those terms. For example, the terms of highest power in $\phi^{I|a}_{|b}$ and R^{ab}_{cd} are given by

$$
\nabla^b\mathcal{G}^a_b=\epsilon_{IJK}\delta^{acel}_{dfhk}\phi^{K|d}{}_{|c}\phi^{I|f}{}_{|e}\phi^{J|g}R_{gl}{}^{hk}+\iota_{IJK}\delta^{aceg}_{bdfh}\phi^{J}{}_{|l}\phi^{K|lb}\phi^{I|d}{}_{|c}R_{eg}{}^{fh}+\cdots\,,\tag{24}
$$

where the ellipses denote the terms with less numbers of $\phi^{I|a}_{|b}$ or R^{ab}_{cd}. Then, we can extract the coefficient of $\phi^A{}_{,mn}\phi^B{}_{,op}g_{qr,st}$ out of this expression by taking its derivative with respect to $\phi^A{}_{,mn}\phi^B{}_{,op}g_{qr,st}$ as

$$
\begin{aligned}
\left(\nabla^b\mathcal{G}^a_b\right)^{;mn;op;qr,st}_{\;A\;\;\;B}=&\,2\epsilon_{(A|J|B)}\delta^{acel}_{dfhk}\delta^{(m}_c g^{n)d}\delta^{(o}_e g^{p)f}\left(\phi^{J|(q}g^{r)k}\delta^{(s}_l g^{t)h}+g^{h(q}\delta^{r)}_l\phi^{J|(s}g^{t)k}\right)\\
&+2\iota_{IJK}\delta^{aceg}_{bdfh}\left(\delta^K_A\delta^I_B\phi^{J|(m}g^{n)b}\delta^{(o}_c g^{p)d}+\delta^I_A\delta^K_B\phi^{J|(o}g^{p)b}\delta^{(m}_c g^{n)d}\right)\delta^{(q}_e g^{r)h}\delta^{(s}_g g^{t)f},
\end{aligned}\tag{25}
$$

where the terms in the ellipses of Eq. (24) drop out by this operation. Equation (25) must vanish when contracted with a vector Y_a such that $Y_a\phi^{I|a}=0$, because Eq. (9) implies that

$$
Y_a\left(\nabla^b\mathcal{G}^a_b\right)^{;mn;op;qr,st}_{\;A\;\;\;B}=Y_a\left(\frac{1}{2}\mathcal{E}_I\phi^{I|a}\right)^{;mn;op;qr,st}_{\;A\;\;\;B}=\frac{1}{2}Y_a\phi^{I|a}\mathcal{E}_I{}^{;mn;op;qr,st}_{\;A\;\;\;B}=0.\tag{26}
$$

Thus, we obtain a constraint equation given by

$$
\begin{aligned}
&2\epsilon_{(A|J|B)}Y_a\delta^{acel}_{dfhk}\delta^{(m}_c g^{n)d}\delta^{(o}_e g^{p)f}\left(\phi^{J|(q}g^{r)k}\delta^{(s}_l g^{t)h}+g^{h(q}\delta^{r)}_l\phi^{J|(s}g^{t)k}\right)\\
&\quad+2\iota_{IJK}Y_a\delta^{aceg}_{bdfh}\left(\delta^K_A\delta^I_B\phi^{J|(m}g^{n)b}\delta^{(o}_c g^{p)d}+\delta^I_A\delta^K_B\phi^{J|(o}g^{p)b}\delta^{(m}_c g^{n)d}\right)\delta^{(q}_e g^{r)h}\delta^{(s}_g g^{t)f}=0.
\end{aligned}\tag{27}
$$

This constraint has eight free indices (m,n,o,p,q,r,s,t), and its any components must vanish. To find the conditions on ϵ_{AJB} and ι_{IJK} from it, we first look at the trace part by contracting with $g_{mn}g_{op}g_{qr}$:

$$
-8\epsilon_{(A|J|B)}\phi^{J(s}Y^{t)}+8\iota_{(A|J|B)}\phi^{J(s}Y^{t)}=-8\epsilon_{(A|J|B)}\phi^{J(s}Y^{t)}=0,\tag{28}
$$

where we have used $\iota_{(A|J|B)}=0$ that follows from the definition (A7). This equation must be satisfied for any $\phi^{J|a}$, and therefore it is necessary to impose $\epsilon_{(A|J|B)}\,(=\epsilon_{AJB})=0$. Then, we can plug it back in to Eq. (27) to obtain an equation on ι_{IJK}. It turns out that this equation implies $\iota_{IJK}=0$, as we can confirm as follows. Let us project Eq. (27) with $\epsilon_{AJB}=0$ imposed to the basis vectors $Y^a,\tilde{Y}^a,\phi^{I|a}$ $(I=1,2)$ such that

$$
Y_aY^a=\tilde{Y}_a\tilde{Y}^a=1,\qquad Y_a\tilde{Y}^a=0,\qquad Y_a\phi^{I|a}=\tilde{Y}_a\phi^{I|a}=0.\tag{29}
$$

Contracting Eq. (27) with $Z_m W_n V_o \phi^C_{|p}\phi^D_{|q}\phi^E_{|r}\phi^F_{|s}\phi^G_{|t}$, where Z^a, W^a, V^a are either Y^a or $\tilde{Y}^a$, we find

$$0 = 4\iota_{AJB}\delta^{ac}_{bd}Y_a V^b\big(Z_c W^d + W_c Z^d\big)X^{JC}\left(X^{(D|(F}X^{G)|E)} - X^{DE}X^{FG}\right). \tag{30}$$

We may focus on the $\iota_{AJB}X^{JC}$ part of this equation, and it must vanish identically for any X^{JC}. Then we can see that $\iota_{AJB} = 0$ must be imposed.

Applying this procedure to any other second derivative terms in Eq. (23) one by one from higher to lower power ones, we find

$$\begin{aligned}
&\alpha_{IJ} = -2\beta_{JI}, \quad \gamma_{IJKL} = -4\eta_{I[J|L|K]}, \quad \alpha_{AI}\left[\delta_{I(C}X^{-1}_{D)B} - \delta_{IB}X^{-1}_{CD}\right] = 4\left(\eta_{(CD)AB} - \eta_{(C|BA|D)}\right), \\
&\epsilon_{IJK} = \iota_{IJK} = \omega_{IJK} = \lambda_{(IJ)KLM} - \lambda_{(I|LM|J)K} = \mu_I = \zeta_{I[JK]} = \xi_{IJ} = \tau_{IJKLM} = F_{IJKLM} = 0,
\end{aligned} \tag{31}$$

$$\nu_{ACDB}X^{-1}_{EF} - \nu_{AB(E|C}X^{-1}_{|F)D} - 2\sigma_{(EF)CADB} = 0, \qquad \nu_{B[A|K|C]} = \sigma_{EF[CA]DB} = 0, \tag{32}$$

must be imposed, where X^{-1}_{IJ} is the inverse matrix of X^{IJ}. Expressions in terms of the functions appearing in Eq. (21), the constraints (31) are expressed as

$$\begin{aligned}
&B_{IJ} = -2\left(\mathcal{F} + 2\mathcal{W}\right)_{,I,J} + A_{,IJ} + 2D_{(I|K|J),L}X^{KL} - 16E_{K(I|MN|J),L}X^{KL}X^{MN} - 8\left(J_{K(I,J),L} - J_{KL,I,J}\right)X^{KL}, \\
&C_I = -2\left(\mathcal{F} + 2\mathcal{W}\right)_{,I} + 2\left(D_{JKI} + 8J_{J[K,I]}\right)X^{JK} - 8E_{JKLMI}X^{JK}X^{LM}, \\
&F_{IJKLM} = 0, \qquad G_{IJ} = 2J_{IJ} - 2K_{(I,J)} + 4J_{K(I,J)L}X^{KL}, \qquad H_{IJKL} = 2J_{IJ,KL}, \\
&K_{[I,J]} = -2J_{K[I,J]L}X^{KL}, \qquad K_{I,JK} = K_{J,IK}, \qquad L_{IJK} = \frac{2}{3}K_{(I,JK)}, \qquad I = \frac{1}{2}\mathcal{F} + \mathcal{W},
\end{aligned} \tag{33}$$

where $\mathcal{W} = \mathcal{W}(\phi^I)$, and $\mathcal{F} = \mathcal{F}(\phi^I, X^{JK})$ is a function satisfying $\mathcal{F}_{,IJ} = G_{IJ}$, which can be integrated as

$$\mathcal{F} = \int G_{IJ}\, dX^{IJ} = \int \left(2J_{IJ} - 2K_{I,J} + 4J_{KI,JL}X^{KL}\right) dX^{IJ}. \tag{34}$$

The conditions $\zeta_{I[JK]} = \omega_{IJK} = \lambda_{(IJ)KLM} - \lambda_{(I|LM|J)K} = 0$ in Eq. (31) imply

$$D_{I[JK]} = -4J_{I[J,K]} + 8E_{LMI[JK]}X^{LM}, \tag{35}$$

$$D_{I(JK)} = \frac{1}{2}C_{J,IK} + G_{JK,I} + \left(-D_{LMJ,IK} + 4H_{[I|LJK,|M]}\right)X^{LM} + 8E_{(J|ILM|K)}X^{LM} + 4E_{LMNOJ,IK}X^{LM}X^{NO}, \tag{36}$$

$$\begin{aligned}
0 = &\frac{1}{2}D_{IJK,LM} - \frac{1}{2}D_{(I|LM,|J)K} + H_{IJKM,L} - H_{(I|LMK,|J)} \\
&- 2\left(E_{ML(IJ)K} - E_{K(IJ)LM}\right) - 4\left(E_{NO(IJ)K,LM} - E_{NO(I|LM,|J)K}\right)X^{NO},
\end{aligned} \tag{37}$$

and Eq. (32) implies

$$G_{I[J,K]L} = 0, \quad H_{IJK[L,M]N} = 0, \quad G_{(IJ,KL)} = 3H_{L(IJK)} + 2H_{LM(IJ,KN)}X^{MN} - 2K_{(I,JK),L}\,. \tag{38}$$

Among these, $G_{I[J,K]L} = 0$ implies $\mathcal{F}_{,IJ,KL} = \mathcal{F}_{,KL,IJ}$, which is nothing but the integrability condition that guarantees the integral (34) to exist.

C. Properties of the most general equations of motion

Imposing the constraints (33) to Eq. (21), we finally obtain the most general second-order field equations of the bi-scalar-tensor theory $\mathcal{G}^a_b = 0$, where

$$\begin{aligned}\mathcal{G}^a_b =\ & A\delta^a_b + \left[-2\mathcal{F}_{,I} - 4\mathcal{W}_{,I} + 2\left(D_{JKI} + 8J_{J[K,I]}\right)X^{JK} - 8E_{JKLMI}X^{JK}X^{LM}\right]\delta^{ac}_{bd}\phi^{I|d}_{|c} \\ & + \left(-2\mathcal{F}_{,I,J} - 4\mathcal{W}_{,I,J} + A_{,IJ} + 2D_{IKJ,L}X^{KL} - 16E_{KIMNJ,L}X^{KL}X^{MN} - 16J_{K[I,L],J}X^{KL}\right)\phi^{(I|a}\phi^{J)}_{|b} \\ & + D_{IJK}\delta^{ace}_{bdf}\phi^I_{|c}\phi^{J|d}\phi^{K|f}_{|e} + E_{IJKLM}\delta^{aceg}_{bdfh}\phi^I_{|c}\phi^{J|d}\phi^K_{|e}\phi^{L|f}\phi^{M|h}_{|g} + \left(\frac{1}{2}\mathcal{F} + \mathcal{W}\right)\delta^{ace}_{bdf}R_{ce}{}^{df} + \mathcal{F}_{,IJ}\delta^{ace}_{bdf}\phi^{I|d}_{|c}\phi^{J|f}_{|e} \\ & + J_{IJ}\delta^{aceg}_{bdfh}\phi^I_{|c}\phi^{J|d}R_{eg}{}^{fh} + 2J_{IJ,KL}\delta^{aceg}_{bdfh}\phi^I_{|c}\phi^{J|d}\phi^{K|f}_{|e}\phi^{L|h}_{|g} + K_I\delta^{aceg}_{bdfh}\phi^{I|d}_{|c}R_{eg}{}^{fh} + \frac{2}{3}K_{I,JK}\delta^{aceg}_{bdfh}\phi^{I|d}_{|c}\phi^{J|f}_{|e}\phi^{K|h}_{|g}.\end{aligned} \tag{39}$$

Note that one can eliminate $\mathcal{W}(\phi^I)$ from the above equation by redefining $\mathcal{F} \to \hat{\mathcal{F}}(\phi^I, X^{JK}) = \mathcal{F} + 2\mathcal{W}$. We can see that Eqs. (36)–(38) do not reduce the number of the arbitrary functions, because these are the relations between derivatives of the functions, and then they do not affect the structure of the field equations (39) directly. From (9), the scalar-field equations of motion are found as

$$0 = \mathcal{E}_I = 2\mathcal{Q}_I + \delta^{cegl}_{bdhm}\left(-\gamma_{JIKL}\phi^{K|b}\phi^J_{|c}\phi^{L|d}_{|e}R_{gl}{}^{hm} + \frac{2}{3}\sigma_{JIKLMN}\phi^J_{|c}\phi^{M|b}\phi^{K|d}_{|e}\phi^{L|h}_{|g}\phi^{N|m}_{|l}\right). \tag{40}$$

We can confirm these equations reduce to that for the single scalar field derived by [18] once the number of the scalar fields is reduced from two to one [30].

In [29], cosmological perturbations was studied in the multi-field DBI galileon theories [21–28], and as a result it was noticed that the double-dual Riemann term

$$\mathcal{L} = \sqrt{-g}\,\delta_{IJ}\delta_{KL}\delta^{aceg}_{bdfh}\phi^I_{|a}\phi^{J|b}\phi^K_{|c}\phi^{L|d}R_{eg}{}^{fh} \tag{41}$$

is not included in the generalized multi-Galileon theory [20]. We can check that this term is actually contained in our theory. It is straightforward to derive the field equations from Eq. (41):

$$E^{ab}(\mathcal{L}) = 4g^{l(a}\delta^{b)ceg}_{ldfh}X^{I[I}\phi^{J]}_{|c}\phi^{J|d}R_{eg}{}^{fh} + 8g^{l(a}\delta^{b)ceg}_{ldfh}\delta_{I[J}\delta_{K]L}\phi^I_{|c}\phi^{J|d}\phi^{K|f}_{|e}\phi^{L|h}_{|g}\ . \tag{42}$$

This is reproduced by setting $J_{IJ} = 2\left(\delta_{IJ}\delta_{KL} - \delta_{IK}\delta_{JL}\right)X^{KL}$ in our field equations. It can be generalized further, since J_{IJ} can be an arbitrary function of ϕ^I and X^{IJ} in our theory. In addition, the E_{IJKLM} term in the equations of motion (39) is another new term which was not incorporated in the generalized multi-Galileon theory nor in the Horndeski theory, since this term identically vanishes when the number of the scalar fields is reduced from two to one.

To conclude, by our construction we successfully reproduced the terms missing in the candidate of the most general theory proposed earlier [20] and also found an additional new term.

D. Construction of Lagrangian and challenges thereof

As mentioned in section II A, Step 2 of the construction procedure is to construct the Lagrangian from which the most general equations of motion (39) is obtained as its Euler-Lagrange equations. This would be accomplished by "integrating" the equations of motion following the inverse of the calculus of variation. Unfortunately there are no systematic ways to do it in general, and it is difficult particularly for complicated equations of motion such as Eqs. (39) and (40). Below, we introduce our attempt to construct the Lagrangian based on the strategy for [18] in the single scalar field case.

In the single scalar field case, the expression of the Lagrangian was guessed from the trace of the equations of motion, and the correct expression was found based on it. Motivated by this fact, we examine the trace of the equations of motion (39) in the two scalar field case. It turns out that the terms appearing in the trace can be classified according to their tensorial structure as

$$\mathcal{L}_1 = \sqrt{-g}\, M^{(1)}{}_I \phi^{I|c}_{|c}, \tag{43a}$$

$$\mathcal{L}_2 = \sqrt{-g}\left(M^{(2)}\delta^{ce}_{df} R_{ce}{}^{df} + 2M^{(2)}_{,IJ}\delta^{ce}_{df}\phi^{I|d}_{|c}\phi^{J|f}_{|e}\right), \tag{43b}$$

$$\mathcal{L}_3 = \sqrt{-g}\, M^{(3)}_{IJK}\delta^{ce}_{df}\phi^{I}_{|c}\phi^{J|d}\phi^{K|f}_{|e}, \tag{43c}$$

$$\mathcal{L}_4 = \sqrt{-g}\left(M^{(4)}_I \delta^{ceg}_{dfh}\phi{I}^{|d}_{|c} R_{eg}{}^{fh} + \frac{2}{3}M^{(4)}_{I,JK}\delta^{ceg}_{dfh}\phi^{I|d}_{|c}\phi^{J|f}_{|e}\phi^{K|h}_{|g}\right), \tag{43d}$$

$$\mathcal{L}_5 = \sqrt{-g}\left(M^{(5)}_{IJ}\delta^{ceg}_{dfh}\phi^{I}_{|c}\phi^{J|d}R_{eg}{}^{fh} + 2M^{(5)}_{IJ,KL}\delta^{ceg}_{dfh}\phi^{I}_{|c}\phi^{J|d}\phi^{K|f}_{|e}\phi^{L|h}_{|g}\right), \tag{43e}$$

$$\mathcal{L}_6 = \sqrt{-g}\, M^{(6)}, \tag{43f}$$

$$\mathcal{L}_7 = \sqrt{-g}\, M^{(7)}_{IJKLM}\delta^{ceg}_{dfh}\phi^{I}_{|c}\phi^{J|d}\phi^{K}_{|e}\phi^{L|f}\phi^{M|h}_{|g}, \tag{43g}$$

where $M^{(1)}, M^{(2)}, M^{(3)}_{IJK}, M^{(4)}_I, M^{(5)}_{IJ}, M^{(6)}$, and $M^{(7)}_{IJKLM}$ are arbitrary functions of ϕ^I and X^{IJ} satisfying

$$M^{(3)}_{IJK} = M^{(3)}_{JIK}, \quad M^{(5)}_{IJ} = M^{(5)}_{JI}, \quad M^{(7)}_{IJKLM} = -M^{(7)}_{KJILM} = -M^{(7)}_{ILKJM} = M^{(7)}_{JILKM}, \tag{44}$$

$$M^{(1)}_{[I,J]K} = M^{(2)}_{,I[J,K]L} = M^{(3)}_{IJ[K,L]M} = M^{(4)}_{[I,J]K} = M^{(4)}_{I,J[K,L]M} = M^{(5)}_{IJ,K[L,M]N} = M^{(7)}_{IJKL[M,N]O} = 0. \tag{45}$$

The condition (45) is crucial to keep the equations of motion second order in derivatives; the Euler-Lagrange equations derived from (43) contains higher derivative terms such as $\nabla_a\nabla_b\nabla_c\phi^I$ multiplied by derivatives of $M^{(1,\dots,7)}$, and they can be eliminated only when (45) is imposed.

The Euler-Lagrange equations for the Lagrangian densities (43) (listed in appendix A 4) must reproduce the most general equations of motion (39) if (43) defines the most general theory for two scalar fields. Comparing Eq. (39) with equations in appendix A 4 term by term, we find it is the case if

$$\begin{aligned} A = {} & M^{(1)}_{I,J}X^{IJ} + 4M^{(2)}_{,I,J}X^{IJ} - 2M^{(3)}_{IJK,L}X^{IJ}X^{KL} - 8\left(M^{(5)}_{IJ,K,L} - M^{(5)}_{IL,J,K}\right)X^{IJ}X^{KL} \\ & + \frac{1}{2}M^{(6)} + 8M^{(7)}_{IJKLM,N}X^{IJ}X^{KL}X^{MN}, \end{aligned} \tag{46a}$$

$$\begin{aligned} D_{IJK} = {} & -\frac{1}{2}M^{(1)}_{(I,J)K} - 4M^{(2)}_{,(I,J)K} + \frac{3}{2}M^{(3)}_{(IJK)} + \left(M^{(3)}_{L(IJ),MK} + M^{(3)}_{(I|LM,|J)K} - M^{(3)}_{IJL,KM}\right)X^{LM} \\ & - 2M^{(4)}_{K,I,J} + 2\left(2M^{(5)}_{K(I,J)} - M^{(5)}_{IJ,K}\right) + 4\left(2M^{(5)}_{L(I,J),MK} - M^{(5)}_{IJ,L,MK}\right)X^{LM} \\ & + 12M^{(7)}_{IJ(KLM)}X^{LM} + 8M^{(7)}_{IJNLM,OK}X^{LM}X^{NO}, \end{aligned} \tag{46b}$$

$$\begin{aligned} E_{IJKLM} = {} & \frac{1}{8}\left(\delta^{PQ}_{IK}\delta^{RS}_{JL} + \delta^{PQ}_{JL}\delta^{RS}_{IK}\right)\Big[-\frac{1}{2}M^{(3)}_{PRQ,SM} - 4M^{(5)}_{PR,Q,SM} + M^{(7)}_{PRQSM} + M^{(7)}_{PRMSQ} \\ & + M^{(7)}_{PRQMS} - \left(M^{(7)}_{PRQSN,OM} - 2M^{(7)}_{PRNSQ,OM} - 2M^{(7)}_{PNQSO,RM}\right)X^{NO}\Big], \end{aligned} \tag{46c}$$

$$\mathcal{F} + 2\mathcal{W} = M^{(2)} - 2M^{(2)}_{,IJ}X^{IJ} - 2M^{(4)}_{I,J}X^{IJ} + 2M^{(5)}_{IJ}X^{IJ} + 4M^{(5)}_{IK,JL}X^{IJ}X^{KL}, \tag{46d}$$

$$J_{IJ} = -\frac{1}{2}M^{(2)}_{,IJ} - M^{(4)}_{(I,J)} + M^{(5)}_{IJ} + \left(2M^{(5)}_{K(I,J)L} - M^{(5)}_{IJ,KL}\right)X^{KL}, \tag{46e}$$

$$K_I = -M^{(4)}_{J,KI} X^{JK}. \tag{46f}$$

In addition, comparing the $\phi^{(I|a}\phi^{J)}{}_{|b}$ and $\delta^{ac}_{bd}\phi^{I|d}_{\ \ |c}$ terms, we see that the following two conditions must be satisfied:

$$\begin{aligned}
&-2\left(\mathcal{F}+2\mathcal{W}\right)_{,I,J} + A_{,IJ} + 2D_{K(IJ),L}X^{KL} + 16E_{KMN(IJ),L}X^{KL}X^{MN} - 8\left(J_{K(I,J),L} - J_{KL,I,J}\right)X^{KL} \\
&= M^{(1)}_{(I,J)} + 2M^{(2)}_{,I,J} - \left(M^{(3)}_{IJK,L} + 2M^{(3)}_{KL(I,J)} - 2M^{(3)}_{K(IJ),L}\right)X^{KL} - 4\left(M^{(5)}_{IJ,KL} + M^{(5)}_{KL,IJ} - 2M^{(5)}_{IK,JL}\right)X^{KL} \\
&\quad + \frac{1}{2}M^{(6)}_{,IJ} + 8\left(M^{(7)}_{MNKL(I,J)} - 2M^{(7)}_{MN(KI)J,L} + M^{(7)}_{MN(IJ)K,L}\right)X^{KL}X^{MN},
\end{aligned} \tag{47a}$$

$$\begin{aligned}
&-2\left(\mathcal{F}+2\mathcal{W}\right)_{,K} + 2\left(D_{IJK} + 8J_{I[J,K]}\right)X^{IJ} - 8E_{IJLMK}X^{IJ}X^{LM} \\
&= -M^{(1)}_{I,JK}X^{IJ} - 2\left(M^{(2)}_{,K} + 2M^{(2)}_{I,JK}X^{IJ}\right) + 3M^{(3)}_{(IJK)}X^{IJ} + 2M^{(3)}_{IJM,LK}X^{IL}X^{JM}.
\end{aligned} \tag{47b}$$

If we could solve these equations for $M^{(1,\dots,7)}$ to express them in terms of the arbitrary functions $A, \dots, K_I$ appearing in (39), then the Lagrangian density would be given by a summation of $\mathcal{L}_{1,\dots,7}$ of Eq. (43). In the single scalar field case, the first step was to integrate the equation corresponding to Eq. (46f), hence this equation would be the starting point even in our two scalar field case. It have not been achieved yet, however, and also it should be kept in mind that the terms we considered $(M^{(1)}, \dots, M^{(7)})$ may be insufficient to construct the true Lagrangian. In the single scalar field case, the terms corresponding to Eq. (43) happened to be enough to reproduce the most general equations of motion. However, there is no guarantee that they are enough also in the two scalar field case. Albeit these difficulties, we believe that above results are useful to build the Lagrangian for the most general theory.

III. WAVE PROPAGATION IN THE BI-HORNDESKI THEORY

In section II, we have successfully constructed the most general equations of motion in the bi-Horndeski theory. Although we still lack the Lagrangian that gives those equations, some physical properties of the theory can be studied using only the equations of motion. An example is the classical dynamics of wave in those theories, which is governed by the equations of motion and does not depend on the Lagrangian directly.

Wave propagation is one of the most basic physical process ongoing in a theory. It governs the dynamics of small fluctuations in the theory, and also the causal structure will be implemented in accord with the maximum propagation speed. Motivated by these facts, as a first investigation on the physical properties of the bi-Horndeski theory, we study the maximum propagation speed in this theory and examine the causal structure based upon it.

A. Method of characteristics

As a preparation, let us briefly review the method of characteristics for a generic equation of motion. This method is particularly useful when we study the maximum speed of propagation. Suppose that a vector of dynamical variables v_I obeys field equations

$$E_I(v, \partial v, \partial^2 v) = 0. \tag{48}$$

To describe the time evolution based on this equation, we introduce a three-dimensional hypersurface Σ and a coordinate system $(x^a) = (x^0, x^\mu)$, where Σ is at $x^0 = 0$ and x^μ lies on Σ. We use the notation that Latin indices $(a, b, \ldots)$ denote all the four dimensions and Greek indices $(\mu, \nu, \ldots)$ denote only the three dimensions on Σ. Now let us assume that E_I is linear in $\partial_0^2 v$, which is the case in the bi-Horndeski theory. Then Eq. (48) is expressed as

$$\frac{\partial E_I}{\partial v_{J,00}} v_{J,00} + \cdots = 0, \tag{49}$$

where the ellipses denote terms up to first order in derivatives with respect to x^0. Equation (49) can be solved to determine $v_{J,00}$ in terms of quantities with lower-order x^0 derivatives as long as the coefficients of the $v_{J,00}$ term, $\partial E_I/\partial v_{J,00}$, is invertible as a matrix acting on the vector v_J. If it is not invertible, on the other hand, the value of $v_{J,00}$ cannot be fixed by Eq. (49), and particularly $v_{J,00}$ can be discontinuous there. We call Σ a characteristic surface in such a case.

A characteristic surface Σ can be interpreted as a wavefront in the high frequency limit, and also it defines the maximum propagation speed in the theory and thus Σ becomes a boundary of causal domain. These properties can be seen as follows [40] (see also [42, 43]). Suppose that we have discontinuity in $v_{I,00}$ across Σ, while v_a and $v_{a,0}$ are continuous there. Then this surface must be characteristic, because otherwise $v_{I,00}$ cannot be discontinuous as we argued above. It implies that the discontinuity propagates on the characteristic surface Σ. Then, interpreting the discontinuity as wave in high-frequency limit, we may regard Σ as the wavefront for the high-frequency wave.

Now let us consider time evolution from an initial time slice σ, and focus on a finite part σ_0 on it (see Fig. 2). In the region enclosed by σ_0 and the characteristic surface emanating from its edge, the time evolution will be uniquely specified by the initial data on σ_0. It is because we can solve the equations of motion (49) to fix the time evolution in such a region by taking x^0-constant surface that is not characteristic (say Σ' in the figure). On the other hand, the solution outside this region cannot be fixed only by the initial data on σ_0, because disturbances outside σ_0 can propagate in this outer region along characteristic surfaces. In this sense, the boundary of the causal domain for σ_0 is given by the characteristic surface emanating from its edge. In other words, the maximum propagation speed in a theory is determined by characteristic surfaces.

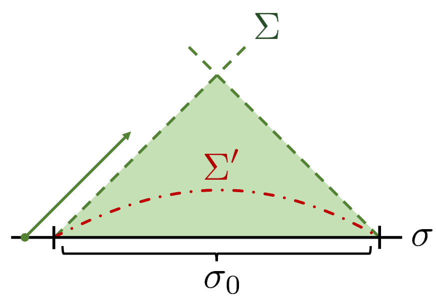


FIG. 2: Initial time slice σ and the causal domain for the region σ_0. In the green shaded region enclosed by σ_0 and the characteristic surface Σ from its edge (green dashed line), the solution is uniquely fixed by the initial data on σ_0 by solving the field equations using non-characteristic surfaces (say Σ' represented by the red dot-dashed curve).

To express the coefficient matrix $\left(\partial E_I/\partial v_{J,00}\right)$ covariantly, we introduce a normal vector of Σ, $\xi_a \equiv$

$(dx^0)_a$, with which we can express the equation to determine characteristic surfaces as

$$P(x,\xi)\cdot r \equiv \xi_s\xi_t \frac{\partial E_I}{\partial v_{J,st}} r_J = 0. \tag{50}$$

Σ is characteristic if Eq. (50) has a nontrivial solution r_J, which is realized when $\det P(\xi) = 0$. The eigenvector r_J for a vanishing eigenvalue of P corresponds to a mode propagating on Σ. P is called the principal symbol of the field equation (49), and $\det P(\xi) = 0$ is called the characteristic equation. In this essay we call also $P\cdot r = 0$ a characteristic equation, which should be understood in the above sense.

B. Wave propagation in the bi-Horndeski theory

Based on the the method of characteristics introduced above, we analyze the wave propagation and causal structure in the bi-Horndeski theory in this section. The first step is to express the equations of motion (39), (40) in the form of Eq. (49), which are given as[6]

$$\mathcal{G}^{\mu\nu}(\mathcal{L}) = \mathcal{A}^{\mu\nu,\rho\sigma} g_{\rho\sigma,00} + \mathcal{B}^{\mu\nu}_I \phi_{I,00} + \mathcal{C}^{\mu\nu}, \qquad \mathcal{E}_I = \widetilde{\mathcal{A}}^{\mu\nu}_I g_{\mu\nu,00} + \widetilde{\mathcal{B}}_{IJ}\phi_{J,00} + \widetilde{\mathcal{C}}_I, \tag{51}$$

where $\mathcal{A}^{\mu\nu,\rho\sigma}, \mathcal{B}^{\mu\nu}_I$ and $\mathcal{C}^{\mu\nu}$ in the metric equation $\mathcal{G}^{\mu\nu}$ are given by

$$\mathcal{A}^{\mu\nu,\rho\sigma} = -\left(\mathcal{F}+2\mathcal{W}\right) g^{p(\mu}\delta^{\nu)0(\rho}_{pmf} g^{\sigma)f} g^{0m} - 2J_{IJ} g^{p(\mu}\delta^{\nu)c0(\rho}_{pdmf} g^{\sigma)f} g^{0m} \phi^I_{|c}\phi^{J|d} - 2K_I g^{p(\mu}\delta^{\nu)c0(\rho}_{pdmf} g^{\sigma)f} g^{0m}\phi^{I|d}_{|c}, \tag{52}$$

$$\begin{aligned}\mathcal{B}^{\mu\nu}_I =& \tilde{B}_I g^{l(\mu}\delta^{\nu)0}_{lm} g^{0m} + D_{JKI} g^{l(\mu}\delta^{\nu)c0}_{ldm} g^{0m}\phi^J_{|c}\phi^{K|d} + E_{JKLMI} g^{l(\mu}\delta^{\nu)ce0}_{ldfm} g^{0m}\phi^J_{|c}\phi^{K|d}\phi^L_{|e}\phi^{M|f} + 2\mathcal{F}_{,IJ} g^{l(\mu}\delta^{\nu)c0}_{ldm} g^{0m}\phi^{J|d}_{|c} \\ &+ 4J_{JK,LI} g^{l(\mu}\delta^{\nu)ce0}_{ldfm} g^{0m}\phi^J_{|c}\phi^{K|d}\phi^{L|f}_{|e} + K_I g^{l(\mu}\delta^{\nu)0ce}_{lmdf} g^{0m} R_{ce}{}^{df} + 2K_{I,JK} g^{l(\mu}\delta^{\nu)ce0}_{ldfm} g^{0m}\phi^{J|d}_{|c}\phi^{K|f}_{|e},\end{aligned} \tag{53}$$

$$\mathcal{C}^{\mu\nu} = \mathcal{C}^{\mu\nu}\big(g_{\rho\sigma}, g_{\rho\sigma,0}, g_{\rho\sigma,\kappa}, g_{\rho\sigma,\kappa 0}, g_{\rho\sigma,\kappa\lambda}, \phi_I, \phi_{I,0}, \phi_{I,\rho}, \phi_{I,0\rho}, \phi_{I,\rho\sigma}\big). \tag{54}$$

$\tilde{B}_I$ in Eq. (53) is defined as

$$\tilde{B}_I \equiv -2\mathcal{F}_{,I} - \mathcal{W}_{,I} + 2\left(D_{JKI} + 8J_{J[K,I]}\right) X^{JK} - 8E_{JKLMI} X^{JK} X^{LM}. \tag{55}$$

Expressions of $\widetilde{\mathcal{A}}^{\mu\nu}_I, \widetilde{\mathcal{B}}_{IJ}, \widetilde{\mathcal{C}}_I$ of the scalar equation in (51) are rather lengthy, as shown in appendix B 1. Using the expressions above, the equations of motion are written as $P\cdot v_{,00} = S$, where

$$P = \begin{pmatrix} \mathcal{A}^{\mu\nu,\rho\sigma} & \mathcal{B}^{\mu\nu}_J \\ \widetilde{\mathcal{A}}^{\rho\sigma}_I & \widetilde{\mathcal{B}}_{IJ} \end{pmatrix}, \qquad v = \begin{pmatrix} g_{\rho\sigma} \\ \phi_J \end{pmatrix}, \qquad S = \begin{pmatrix} \mathcal{C}^{\mu\nu} \\ \widetilde{\mathcal{C}}_I \end{pmatrix}. \tag{56}$$

Then the characteristics are found by solving

$$P\cdot r = 0, \tag{57}$$

where $r = (r_{ab}, r_J)$ is a vector made of a symmetric tensor r_{ab} and a vector of scalars with two components r_J. The characteristic equation is given by $\det P = 0$, and eigenvectors for vanishing eigenvalues correspond to the modes propagating on the characteristic surface Σ, as we argued in section III A.

[6] It turns out that, in the bi-Horndeski theory, the (00) and (0μ) components of the gravitational equations do not contain second derivatives of the dynamical variables with respect to x^0, and only the $(\mu\nu)$ components have them. It can be confirmed also that $g_{00,00}$ and $g_{0\mu,00}$ do not appear in those equations and also in the scalar field equations. Hence, it suffices to look at the $(\mu\nu)$ components of field equations in which only $g_{\rho\sigma,00}$ appears.

The integrability conditions for equations of motion, Eq. (135) we discuss later in section V, guarantee that the matrix P is symmetric, i.e.,

$$\mathcal{A}^{\mu\nu,\rho\sigma} = \mathcal{A}^{\rho\sigma,\mu\nu}, \quad \mathcal{B}_I^{\mu\nu} = \widetilde{\mathcal{A}}_I^{\mu\nu}, \quad \widetilde{\mathcal{B}}_{IJ} = \widetilde{\mathcal{B}}_{JI}. \tag{58}$$

The integrability conditions have not been imposed to the field equations (39), (40), and then the principal symbol P defined by Eq. (57) does not have the symmetry (58) in general. We proceed without imposing these conditions in the analysis below, and we leave the full analysis with these conditions for future work. In the next section, we find that some properties of the causal structure in this theory can be read out despite this restriction.

1. *Black hole horizon and causal edge*

In GR, a null surface is always a characteristic surface and hence it gives a boundary of causal domain. This property may be lost in the bi-Horndeski theory. Particularly, a Killing horizon, a null hypersurface at which the spacetime is locally stationary, may not be a characteristic surface in this theory, which means that the black hole region defined based on the metric may not be an event horizon, because there may be superluminal modes that enables communication beyond the horizon in the metric.

In this section, as a first application of the formalism developed in the previous section, we clarify conditions for a null hypersurface to be characteristic. We will find that, for a Killing horizon to become a boundary of causal domain, the scalar fields must obey the symmetry of the background spacetime.

Null hypersurface. Let us check if a null hypersurface is characteristic in the bi-Horndeski theory. It is guaranteed to be so in GR, while it is nontrivial in the bi-Horndeski theory due to the modifications made to the theory. We assume that a x^0 = constant surface Σ is null, which is expressed in terms of the metric components as

$$g^{00} = 0, \quad g^{0\alpha} = 0, \quad g_{11} = 0, \quad g_{1\alpha} = 0, \tag{59}$$

hence only g^{01}, g_{01} remain nonzero, where x^1 is the null coordinate lying on Σ and x^α ($\alpha = 2, 3$) are the other spatial coordinates along Σ (see Fig. 3).

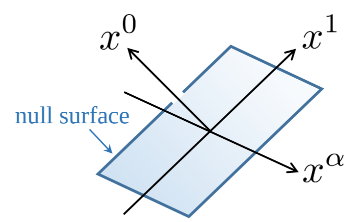


FIG. 3: Coordinates adapted to a null hypersurface.

Under the conditions (59), we can evaluate P of Eq. (57) and show that some of its components vanishes:

$$\mathcal{A}^{11,11} = \mathcal{A}^{11,1\alpha} = \mathcal{A}^{1\alpha,11} = 0, \tag{60}$$

while any other component remain nonzero. Hence, the characteristic equation (57) has the following structure on the null hypersurface:

$$0 = P \cdot r = \begin{pmatrix} \mathcal{A}^{11,11} & 2\mathcal{A}^{11,1\gamma} & \mathcal{A}^{11,\gamma\delta} & \mathcal{B}_J^{11} \\ \mathcal{A}^{1\alpha,11} & 2\mathcal{A}^{1\alpha,1\gamma} & \mathcal{A}^{1\alpha,\gamma\delta} & \mathcal{B}_J^{1\alpha} \\ \mathcal{A}^{\alpha\beta,11} & 2\mathcal{A}^{\alpha\beta,1\gamma} & \mathcal{A}^{\alpha\beta,\gamma\delta} & \mathcal{B}_J^{\alpha\beta} \\ \widetilde{\mathcal{A}}_I^{11} & 2\widetilde{\mathcal{A}}_I^{1\gamma} & \widetilde{\mathcal{A}}_I^{\gamma\delta} & \widetilde{\mathcal{B}}_{IJ} \end{pmatrix} \begin{pmatrix} r_{11} \\ r_{1\gamma} \\ r_{\gamma\delta} \\ r_J \end{pmatrix} = \begin{pmatrix} 0 & 0 & \mathcal{A}^{11,\gamma\delta} & \mathcal{B}_J^{11} \\ 0 & -\mathcal{A}^{11,\alpha\gamma} & \mathcal{A}^{1\alpha,\gamma\delta} & \mathcal{B}_J^{1\alpha} \\ \mathcal{A}^{11,\alpha\beta} & 2\mathcal{A}^{\alpha\beta,1\gamma} & \mathcal{A}^{\alpha\beta,\gamma\delta} & \mathcal{B}_J^{\alpha\beta} \\ \widetilde{\mathcal{A}}_I^{11} & 2\widetilde{\mathcal{A}}_I^{1\gamma} & \widetilde{\mathcal{A}}_I^{\gamma\delta} & \widetilde{\mathcal{B}}_{IJ} \end{pmatrix} \begin{pmatrix} r_{11} \\ r_{1\gamma} \\ r_{\gamma\delta} \\ r_J \end{pmatrix}. \tag{61}$$

Even after imposing Eq. (60), P is invertible unless some of the remaining components happen to vanish, hence a null hypersurface is not characteristic in general. It implies that the scalar field wave and gravitational wave typically propagate at subluminal or superluminal speed in the bi-Horndeski theory.

Killing horizon with additional conditions. Since a null hypersurface is not characteristic in general, next let us examine what happens if more conditions are imposed upon it. One important example of such a null hypersurface is the Killing horizon, which is a null hypersurface aligned to the symmetry direction of the spacetime. For example, black hole horizons in stationary spacetimes are Killing horizons. Since a Killing horizon is aligned to the spacetime symmetry direction, the derivatives of the metric in that direction vanish:

$$\partial_1 g_{\alpha\beta} = 0, \qquad \partial_1^2 g_{\alpha\beta} = 0, \qquad \partial_1\partial_\gamma g_{\alpha\beta} = 0, \tag{62}$$

which implies that the following components of the Riemann tensor vanish on the null surface Σ:

$$R_{1\alpha\beta\gamma} = R_{1\alpha1\beta} = 0. \tag{63}$$

In Gauss-Bonnet and Lovelock theories of gravity, it was confirmed that a Killing horizon becomes characteristic once these conditions are imposed [42, 43]. In the Horndeski-type theory, however, it was noticed that these conditions are insufficient and some additional conditions on the scalar field must be imposed to make the Killing horizon characteristic [44].

What happens in the bi-Horndeski theory is similar to that for the Horndeski theory. By explicit calculations, we can show that the conditions (62) make some terms with curvature tensors in P vanishing. Even when this occurs, however, no components in P vanish completely and hence the Killing horizon Σ will not be characteristic in general.

In Ref. [44], it was noticed that a Killing horizon in the Horndeski theory becomes characteristic if the scalar field satisfies certain conditions. We consider generalizations of these conditions[7] given by

$$\partial_1\phi_I = 0, \qquad \partial_1^2\phi_I = 0, \qquad \partial_1\partial_\alpha\phi_I = 0. \tag{64}$$

A physical interpretation of these conditions are that the scalar fields share the symmetry of the background metric on the Killing horizon. With these conditions imposed, we can check that the components of P vanish except for

$$\begin{aligned}\mathcal{A}^{11,\alpha\beta} &= \mathcal{A}^{\alpha\beta,11} = -2\mathcal{A}^{1\alpha,1\beta}\\ &= (g^{01})^2\left[(\mathcal{F}+2\mathcal{W})\,g^{\alpha\beta} - 2J_{IJ}\left(2X^{IJ}g^{\alpha\beta} + \phi^{I|\alpha}\phi^{J|\beta}\right) + 2K_I\left(\phi^{I|\sigma}_{|\sigma}g^{\alpha\beta} - \phi^{I|\alpha\beta}\right)\right],\end{aligned} \tag{65a}$$

$$\begin{aligned}\mathcal{B}^{11}_I = (g^{01})^2\Big[&-\tilde{B}_I + 2D_{JKI}X^{JK} - 8E_{JKLMI}X^{JK}X^{ML} - 2\mathcal{F}_{,IJ}\phi^{J|\gamma}_{|\gamma}\\ &- 4J_{JK,LI}\delta^{\gamma\hat{e}}_{\delta\hat{f}}\phi^{J}_{|\gamma}\phi^{K|\delta}\phi^{L|\hat{f}}_{|\hat{e}} - 2K_{I,JK}\delta^{\gamma\hat{e}}_{\delta\hat{f}}\phi^{J|\delta}_{|\gamma}\phi^{K|\hat{f}}_{|\hat{e}} - 2K_I R_{\gamma\delta}{}^{\gamma\delta}\Big],\end{aligned} \tag{65b}$$

$$\begin{aligned}\widetilde{\mathcal{A}}^{11}_I = 2(g^{01})^2\Big[&\left(D_{IJK} - 8J_{J[K,I]} - 8E_{LKJIM}X^{LM}\right)X^{JK} - D_{IJK}X^{JK} - 8E_{LJKIM}X^{JL}X^{KM}\\ &- 2J_{IL,JK}\delta^{\gamma\hat{e}}_{\delta\hat{f}}\phi^{J}_{|\gamma}\phi^{L|\delta}\phi^{K|\hat{f}}_{|\hat{e}} + 2I_{,I} - 2\left(J_{IJ} - K_{(I,J)} + 2J_{K(I,J)L}X^{KL}\right)\phi^{J|\beta}_{|\beta}\\ &- K_{J,IK}\delta^{\gamma\hat{e}}_{\delta\hat{f}}\phi^{J|\delta}_{|\gamma}\phi^{K|\hat{f}}_{|\hat{e}} - K_I R_{\gamma\delta}{}^{\gamma\delta}\Big].\end{aligned} \tag{65c}$$

[7] In some class of scalar-tensor theories, these conditions on scalar fields are automatically satisfied if spacetime is stationary [52, 53].

Then, the characteristic equation (57) simplifies in this case to

$$0 = P \cdot r = \begin{pmatrix} 0 & 0 & \mathcal{A}^{11,\gamma\delta} & \mathcal{B}^{11}_J \\ 0 & -\mathcal{A}^{11,\alpha\gamma} & 0 & 0 \\ \mathcal{A}^{11,\alpha\beta} & 0 & 0 & 0 \\ \widetilde{\mathcal{A}}^{11}_I & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r_{11} \\ r_{1\gamma} \\ r_{\gamma\delta} \\ r_J \end{pmatrix} . \tag{66}$$

The principal symbol P is degenerate, and it gives d conditions in total on an eigenvector r with vanishing eigenvalue. Then the eigenvector will have $\frac{1}{2}d(d-1) + 2 - d = \frac{1}{2}d(d-3) + 2$ degrees of freedom, which coincides with the number of physical degrees of freedom of the bi-scalar-tensor theory. Therefore, the additional conditions (64) on the scalar fields guarantee the Killing horizon Σ to be characteristic for all of the physical modes. In more physical terms, we have found the sufficient condition for the Killing horizon becomes an event horizon, and the conditions were that the gradients of scalar fields are aligned to the symmetry direction of the background spacetime. A trivial background with flat spacetime and constant scalar fields satisfies the above condition, hence the gravitational wave and the scalar field wave propagate at the light speed on such a trivial background.

2. *Wave propagation on plane wave solution*

We focused on the properties of the Killing horizon and event horizon in the above, and in this section let us focus on a more general background solution and the causal structure realized therein. It is desirable to do this analysis in the bi-Horndeski theory, but unfortunately it is not straightforward to do it at this moment since no explicit solutions have been constructed in this theory. We can still make progress by focusing on some theories included in the bi-Horndeski theory for which some solutions are known. A particularly important theory of this sort is the (single-scalar) Horndeski theory that is invariant when the scalar field is shifted by a constant, which is called the shift-symmetric Horndeski theory. In this theory, we have an exact solution describing a composite of plane wave of scalar field and gravitational waves propagating in the same direction [45]. We will study the propagation of linear, small amplitude wave on this background in this section.

For this purpose, we employ a covariant formalism developed for the Lovelock theories of gravity in Ref. [43], which is equivalent to an analysis in the harmonic gauge. Using this formalism, we will find that, on the plane wave solution, we can define effective metrics that differ from the physical metric and characteristic surfaces are given by null hypersurfaces with respect to the effective metrics. Later in section IV C 1, we will study shock formation phenomena on the plane wave solution later in section IV C 1 based on the results in this section.

Shift-symmetric Horndeski theory. Let us first introduce the subset of the bi-Horndeski theory we focus on in the following part of this essay. We assume that the theory has a single scalar field, and also that the theory is invariant under a constant shift of the scalar field $\phi \to \phi + c$. Then, the most general theory whose equations of motion is up to second order in derivatives is given by the Horndeski theory with its arbitrary functions $K, G_{3,4,5}$ depending only on $X \equiv -\frac{1}{2}\phi_{|a}\phi^{|a}$ but not explicitly on ϕ. Hence the Lagrangian of the theory we consider is given by $L=\sum_{n=2}^{5} L_n$, where

$$
\begin{aligned}
&L_2=K(X), \quad L_3=-G_3(X)\Box\phi, \quad L_4=G_4(X)R+G_{4X}(X)\delta^{b_1b_2}_{a_1a_2}\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}, \\
&L_5=G_5(X)G_{ab}\phi^{|ab}-\frac{1}{6}G_{5X}(X)\delta^{b_1b_2b_3}_{a_1a_2a_3}\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}_{|b_3}.
\end{aligned}
\tag{67}
$$

The metric equations $\mathcal{G}_{ab}=0$ and the scalar field equation $\mathcal{E}_\phi=0$ in this theory can be derived from this Lagrangian, and also they can be reproduced from the equations of the bi-Horndeski theory by reducing the number of the scalar field from two to one. For convenience of our analysis, we define the "trace-reversed" metric equations of motion $\tilde{\mathcal{G}}_{ab}=0$ by

$$
\tilde{\mathcal{G}}_{ab}\equiv\mathcal{G}_{ab}-\frac{1}{2}\mathcal{G}^c{}_c\,g_{ab}. \tag{68}
$$

Later, we will see that this formulation is useful particularly when we examine GR.

From these equations of motion, we can derive the characteristic equation (57) as we did around Eqs. (51), or equivalently by taking derivatives of the equations of motion with respect to the second derivatives of the variables $\xi_s\xi_t g_{qr,st}$ and $\xi_s\xi_t\phi_{,st}$ as we did in Eq. (50). We summarize the explicit expressions of these derivatives in appendix B 2. Using them, the principal symbol $\tilde{P}$ based on the trace-reversed metric equation (68) and the scalar equation is constructed as

$$
\tilde{P}(x,\xi)\cdot r=\begin{pmatrix}\big(\tilde{P}(x,\xi)\cdot r\big)_{ab}\\ \big(\tilde{P}(x,\xi)\cdot r\big)_{\phi}\end{pmatrix}=\xi_s\xi_t\begin{pmatrix}\frac{\partial\tilde{\mathcal{G}}_{ab}}{\partial g_{qr,st}} & \frac{\partial\tilde{\mathcal{G}}_{ab}}{\partial\phi_{,st}}\\ \frac{\partial\tilde{\mathcal{E}}_{\phi}}{\partial g_{qr,st}} & \frac{\partial\tilde{\mathcal{E}}_{\phi}}{\partial\phi_{,st}}\end{pmatrix}\begin{pmatrix}r_{qr}\\ r_\phi\end{pmatrix}, \tag{69}
$$

and the characteristic equation is given by $\tilde{P}(x,\xi)\cdot r=0$ using this expression.

Gauge symmetry and physical modes. Since we are now studying a scalar-tensor theory that is covariant, we have gauge invariance that leads to non-dynamical gauge modes besides the physical propagating modes. One way to remove such non-dynamical modes from the analysis is to introduce a time coordinate and look only at the spatial components of the equations of motion, as we did in section III B. Below, we show how to study the physical modes based on our covariant formalism.

For this purpose it is useful to notice that, as we can confirm using the expressions in appendix B 2, the mode $r=\big(\xi_{(a}X_{b)},0\big)$ for any vector X_a is annihilated identically by acting $\tilde{P}(x,\xi)$, that is,

$$
\tilde{P}(x,\xi)\cdot r=\xi_s\xi_t\begin{pmatrix}\frac{\partial\tilde{\mathcal{G}}_{ab}}{\partial g_{qr,st}}\xi_{(q}X_{r)}\\ \frac{\partial\tilde{\mathcal{E}}_{\phi}}{\partial g_{qr,st}}\xi_{(q}X_{r)}\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}. \tag{70}
$$

This property originates from the Bianchi identity (9), which becomes $\nabla^a\mathcal{G}_{ab}=\frac{1}{2}\phi_{|b}\mathcal{E}_\phi$ in this theory, and the fact that the left-hand side of this identity cannot not have third derivatives since $\mathcal{E}_\phi$ is up to second order in derivatives. This property (70) implies that $\tilde{P}\cdot r$ is invariant under the transformation

$$
r_{ab}\to r_{ab}+\xi_{(a}X_{b)}, \tag{71}
$$

as manifested by Eq. (70). Hence, we can regard $r=\big(\xi_{(a}X_{b)},0\big)$ as the gauge mode that does not affect the characteristic equation (50).

One way to eliminate the gauge mode from consideration is to impose the following condition or r:

$$
\xi^a r_{ab}-\frac{1}{2}\xi_b r^a{}_a=0. \tag{72}
$$

We call this the transverse condition. Identity (70) implies that the gravitational part of the principal symbol is transverse:

$$\xi^a (\tilde{P} \cdot r)_{ab} - \frac{1}{2} \xi_b (\tilde{P} \cdot r)^a{}_a = 0. \tag{73}$$

Hence, $\tilde{P}$ can be regarded as a map acting on transverse tensors. Below, we consider the space of the physical modes r satisfying the transverse condition (72), and examine the characteristic equation.

Characteristic surfaces in GR. Before studying the wave propagation in the Horndeski theory, let us see how the above formalism works in the GR, which is reproduced by setting $K = G_{3,5} = 0$ and G_4 constant in (67). The characteristic equation $\tilde{P} \cdot r = 0$ in this case becomes, as we can obtain from Eq. (B15),

$$-\frac{1}{2} G_4 \left(\xi^2 r_{ab} - 2\xi_c r^c{}_{(a} \xi_{b)} + \xi_a \xi_b r^c{}_c \right) = 0. \tag{74}$$

When ξ is not null, this equation implies $r_{ab} = \xi_{(a} X_{b)}$ for some vector X_a. This r is pure gauge and does not correspond to a physical propagating mode. When ξ is null ($\xi^2 = 0$), Eq. (74) is satisfied for r obeying the transverse condition (72). Also, the gauge mode $r_{ab} = \xi_{(a} X_{b)}$ is transverse as well when $\xi^2 = 0$. Hence, the number of physical modes will be two: among the ten components of r_{ab}, four of them are constrained by the transverse condition (72), and in the remainder four components correspond to the gauge mode $r_{ab} = \xi_{(a} X_{b)}$. The remaining two components propagate on Σ that is null, and they correspond to the usual gravitational wave in GR.

Characteristic surfaces on the plane wave solution. Now let us apply the above covariant formalism to the plane wave solution in the shift-symmetric Horndeski theory [45]. This solution is based on the ansatz adapted to the plane wave of metric and scalar field:

$$ds^2 = F(u, x, y) du^2 + 2dudv + dx^2 + dy^2, \qquad \phi = \phi(u). \tag{75}$$

Constant-u lines are null in this ansatz, and the gravitational and scalar field waves are propagating in that direction. To describe this solution, we introduce a null basis ℓ^a, n^a and $m_i{}^a$ with $i = x, y$ satisfying

$$\ell_a = (du)_a, \quad \ell^a n_a = 1, \quad m_{ia} m_j{}^a = \delta_{ij}, \quad \ell^2 = n^2 = \ell^a m_{ia} = n^a m_{ia} = 0. \tag{76}$$

In this frame, the non-vanishing components of the Riemann tensor are given by

$$R_{uiuj} = -\frac{1}{2} F_{,ij} \,, \qquad R_{uu} = -\frac{1}{2} \Delta F, \tag{77}$$

where $F_{,ij} \equiv \partial^2 F / \partial x^i \partial x^j$ and $\Delta F \equiv F_{,ii}$. Adding to that, the derivatives of ϕ are proportional to ℓ ($\phi_{|a} = \phi'(u) \ell_a$, $\phi_{|ab} = \phi''(u) \ell_a \ell_b$) and hence $X = -\frac{1}{2} \phi_{|a} \phi^{|a}$ becomes zero. These properties drastically simplify the equations of motion, and it turns out that they are satisfied if $K(0) = 0$ and $F(u, x, y)$ obeys

$$\Delta F = -\frac{K_X(0)}{G_4(0)} \phi'^2 \equiv -\kappa \, \phi'^2, \tag{78}$$

while $\phi = \phi(u)$ remains arbitrary. For our study, we consider a solution of Eq. (78) given by

$$F = -\frac{1}{4} \kappa \phi'^2 \left(x^2 + y^2 \right) + a_{ij}(u) x^i x^j, \tag{79}$$

where $a_{ij}(u)$ is a symmetric traceless matrix.

Now let us take the plane wave solution (75) with (79) imposed as the background solution and consider propagation of wave upon it. The maximum propagation speed can be found from the characteristic surface Σ, which is fixed by solving the characteristic equation $\tilde{P}(x,\xi)\cdot r = 0$ using Eq. (69). This is an eigenvalue problem where the matrix $\tilde{P}$ depends on the normal vector ξ of Σ. For an appropriately chosen ξ, an eigenvalue of $\tilde{P}$ becomes zero and there will be a nontrivial eigenvector r corresponding to the vanishing eigenvalue.

In the basis (76), we may expand $r = (r_{ab}, r_\phi)$ in general as

$$r = (r_{ab}, r_\phi) = \left(2\ell_{(a}X_{b)} + 2n_{(a}Y_{b)} + (\hat{r}_{ij} + \alpha\,\delta_{ij})m_{ia}m_{jb}, r_\phi\right), \tag{80}$$

where X^a and Y^a are arbitrary vectors and $\hat{r}_{ij}$ is a symmetric traceless tensor. The first term in r_{ab}, $\ell_{(a}X_{b)}$, corresponds to the aforementioned gauge mode. By plugging (80) into the characteristic equation and solving it for ξ, it turns out that nontrivial solution is obtained only when $Y^a = \alpha = 0$ (see [54] for more details). Then, imposing further the transverse condition (72), we find that the physical modes are described by

$$r = (r_{ab}, r_\phi) = \left(2r_{\ell n}\ell_{(a}n_{b)} + \hat{r}_{ij}m_{ia}m_{jb}, r_\phi\right), \tag{81}$$

and the nontrivial components of the characteristic equation reduce to

$$0 = \begin{pmatrix} \big(\tilde{P}\cdot r\big)_{\ell n} \\ \big(\tilde{P}\cdot r\big)_{ij} \\ \big(\tilde{P}\cdot r\big)_{\phi} \end{pmatrix} = \begin{pmatrix} -\frac{G_4}{2}\,\xi^2 r_{\ell n} + \left(-\frac{1}{2}G_{3X}\phi'^2 + G_{4X}\phi''\right)(\xi\cdot\ell)^2 r_\phi \\ -\frac{G_4}{2}\,\xi^2\,\hat{r}_{ij} - \frac{G_{4X}}{2}\phi'^2(\xi\cdot\ell)^2\hat{r}_{ij} \\ -K_X\xi^2 r_\phi \ + \left(K_{XX}\phi'^2 - 2G_{3X}\phi'' - G_{4X}\Delta F\right)(\xi\cdot\ell)^2 r_\phi \end{pmatrix}, \tag{82}$$

where G_5 does not appear in the above expression although it appears in $\tilde{P}$ on the general background, as shown in appendix B 2. The ℓn component of this equation just fixes $r_{\ell n}$ in terms of r_ϕ. The other components can be solved by the following two ways. First way is to impose the following on ξ and r:

$$-\frac{G_4}{2}\xi^2\hat{r}_{ij} - \frac{G_{4X}}{2}\phi'^2(\xi\cdot\ell)^2\hat{r}_{ij} = 0, \qquad (r_{\ell n}, \hat{r}_{ij}, r_\phi) = (0, \hat{r}_{ij}, 0). \tag{83}$$

Another way is to impose

$$-K_X\xi^2 r_\phi \ + \left(K_{XX}\phi'^2 - 2G_{3X}\phi'' - G_{4X}\Delta F\right)(\xi\cdot\ell)^2 r_\phi = 0, \qquad (r_{\ell n}, \hat{r}_{ij}, r_\phi) = (\tilde{r}_{\ell n}, 0, r_\phi), \tag{84}$$

where

$$\tilde{r}_{\ell n} = \frac{2K_X}{G_4}\,\frac{-\frac{1}{2}G_{3X}\phi'^2 + G_{4X}\phi''}{K_{XX}\phi'^2 - 2G_{3X}\phi'' - G_{4X}\Delta F}\,r_\phi. \tag{85}$$

The first solution (83) has two modes because $\hat{r}_{ij}$ is a traceless symmetric tensor. We call the first solution (83) and the second one (84) the tensor and scalar modes, respectively.

The characteristic surfaces Σ, or the wave propagation surfaces for these modes, are obtained as surfaces normal to ξ_a that obey the first equations in Eqs. (83) and (84). These equations can be equivalently expressed as

$$0 = \left(g^{ab} + \omega\,\ell^a\ell^b\right)\xi_a\xi_b \equiv G_\omega^{ab}\xi_a\xi_b, \qquad \omega = \begin{cases} \frac{G_{4X}}{G_4}\phi'^2 & \text{(tensor)} \\ -\left(\frac{K_{XX}}{K_X} + \frac{G_{4X}}{G_4}\right)\phi'^2 + \frac{2G_{3X}}{K_X}\phi'' & \text{(scalar)} \end{cases}, \tag{86}$$

where we used Eq. (78) to simplify ω for the scalar mode. Hence, a characteristic surface Σ is obtained as a surface whose normal is null with respect to G_ω^{ab}, and in this sense G_ω^{ab} can be regarded as the inverse effective metric for the tensor and scalar modes. We can also construct the effective metrics for tangent vectors of Σ as

$$G_{ab}^{\omega} = g_{ab} - \omega \ell_a \ell_b, \tag{87}$$

which is shifted from the physical metric g_{ab} by an additional term $-\omega \ell_a \ell_b$.

The light cone, i.e., the characteristic cone for matter field minimally coupled to the physical metric g_{ab} such as the maxwell fields, is defined based on g_{ab}. From the effective metric (87), we can see that the tensor and scalar modes in this setup propagate differently. The characteristic cones for the tensor and scalar modes do not coincide with each other in general, while they always do along ℓ, because ℓ is null with respect to G_ω^{ab} for any ω. Therefore, the characteristic surfaces form nested cones that touch with each other along ℓ, as shown in Fig. 4. We can see also that ω vanishes when $\phi' = \phi'' = 0$, and then all the modes propagate at the speed of light. This situation is realized not only on the flat background but also on the purely gravitational plane wave background, for which ϕ is constant but $F = a_{ij}(u)x^i x^j$ is nontrivial. The effect of nontrivial F appears only in the physical metric and spacetime curvature, but not in the deviation of the effective metrics from the physical one.

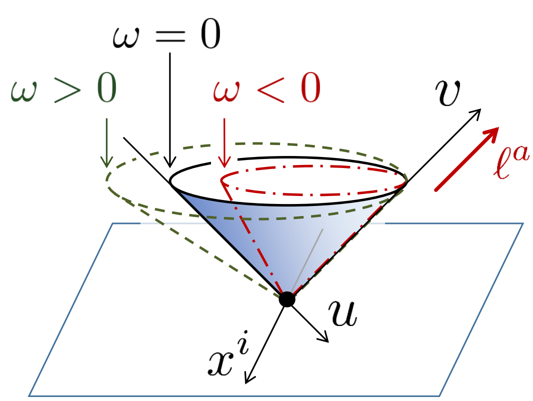


FIG. 4: A schematic of characteristic cones for waves on the plane wave solution at $F(u, x, y) = 0$, where $-u$ and v are taken toward the future direction. The null cone with respect to the physical metric (75) is given by solid black line, and the characteristic cones obtained from the effective metrics (87) are shown by green dashed curve and red dot-dashed curve for $\omega > 0$ and $\omega < 0$, respectively. All the cones are aligned in the direction parallel to ℓ^a and they split in other directions if $\omega \neq 0$, hence the characteristic surfaces form a nested set of cones in general.

As long as ω in Eqs. (87) is finite, propagation speeds of waves are finite and we can define causality as usual, despite the propagation may become superluminal if $\omega > 0$. The only difference from the canonical scalar field coupled to GR is that the causality is not defined with respect to the light cone but to the largest characteristic cone, which is realized for the largest ω. Since the effective metrics (87) are always Lorentzian, the hyperbolicity of the field equation is maintained and the initial value problem is guaranteed to be well-posed.

IV. SHOCK FORMATION IN THE HORNDESKI THEORY

The target of the previous section was the maximum propagation speed of the wave in the most general scalar-tensor theory, and we used the method of characteristics to study it. This analysis clarified the structure of the wave propagation surface and also the causality implemented in this theory, but it did not tell us much about what happens in the dynamics inside the characteristic cones. This raises a natural question to ask: what happens to the waves as they propagates?

In GR and also for the canonical scalar field, the answer is simple: the characteristic surfaces are always null, and correspondingly the gravitational wave and the scalar field wave in those theories propagate at the speed of light. In more general theories, however, it is no more true and generically the propagation speeds of waves depend on the background and also their own amplitudes. This property results in the waveform distortion and shock formation, as we can see a simple example in the time evolution based on Burgers' equation (see Fig. 1). The aim of this section is to examine if a similar shock formation phenomenon could occur in the scalar-tensor theory.

For this purpose, we focus on the propagation of discontinuity in second derivatives of the scalar field and also the metric. This method enables us to study the nonlinear effects in the wave propagation, such as the waveform distortion and shock formation, in a relatively simple manner. This method was applied to Lovelock theories of gravity in [34], and it was found that the gravitational wave suffers from shock formation due to the deviation of this theory from GR. The Horndeski theory has a structure similar to Lovelock theories, hence in principle a similar phenomenon could take place even in the Horndeski theory. Below, we examine this possibility using the formalism of [34].

We first review the formalism of shock formation for a generic equation of motion in section IV A, following Ref. [34]. We will apply this formalism to the shift-symmetric Horndeski theory, and examine conditions to avoid the shock formation for a subclass of this theory without specifying the background solution in section IV B. In more general cases, we have to specify the background solution to make progress. We do it in sections IV C 1 and IV C 2 using the plane wave solution and also the solutions that have certain symmetries, respectively. The latter includes simple background solutions such as the FRW universe and also the spherically symmetric solutions. For these background solutions, we will find that the gravitational wave is always free from shock formation, while the scalar field wave suffers from it generically.

A. Formalism for shock formation

In this section, we introduce a formalism to treat the propagation of discontinuity in second derivatives based on a generic equation of motion (48). This formalism was introduced in Ref. [46] for relativistic fluids and was employed by Ref. [34] to analyze shock formation process in Lovelock theories. We reproduce a part of their formulation to get our analysis oriented and to fix the notation.

We will employ the coordinates $(x^a) = (x^0, x^\mu)$ introduced in section III A, where a characteristic surface Σ lies on $x^0 = 0$. We assume that the equation of motion has the following structure:

$$P_{IJ}(v_{,\mu\nu}\,, v_{,0}\,, v_{,\mu}\,, v\,, x)v_{J,00} + b_I(v_{,0\mu}\,, v_{,\mu\nu}\,, v_{,0}\,, v_{,\mu}\,, v\,, x) = 0. \tag{88}$$

Here we assumed that P_{IJ} is independent of $v_{,0\mu}$, which is the case in the Horndeski theory. On the characteristic surface Σ, $\det P = 0$ is satisfied and hence there is an eigenvector of P with vanishing eigenvalue:

$$r_I P_{IJ} = P_{IJ} r_J = 0, \tag{89}$$

where we assumed that P_{IJ} is symmetric in its indices, hence the left and right eigenvectors of P coincide with each other. This symmetry is guaranteed if the equations of motion (88) is derived from a Lagrangian, hence it holds even in the Horndeski theory.

Now let us consider time evolution from an initial time slice that intersects with $x^0 = 0$, and assume that the dynamical variable v has a discontinuity in its second derivative with respect to x^0 at the locus of $x^0 = 0$ on the initial time slice (see Fig. 5). The discontinuity propagates on Σ, and the solution on the past side of Σ will not be influenced by the discontinuity. Hence, we may regard the wave of the discontinuity to propagate into the "background solution", which is a solution (before introducing the discontinuity) of the equation of motion.

Since the discontinuous part of (88) is given by $P_{IJ}\,[v_{J,00}] = 0$, where the quantity with square brackets denotes its discontinuous part, comparing with Eq. (89) we find that $[v_{I,00}]$ must be proportional to an eigenvector r_I and hence

$$[v_{I,00}] = \Pi(x^\mu)\, r_I, \tag{90}$$

where $\Pi(x^i)$ is the proportional factor, which may be regarded as the amplitude of the discontinuity. Below, we focus on the time evolution of this amplitude Π.

The evolution equation of Π can be constructed by firstly taking x^0 derivative of the equation of motion (88), acting r^I on it to remove third derivatives with respect to x^0, and finally picking up discontinuous part of the resultant equation. With some calculations, we can show that this operation results in

$$\mathcal{K}^\mu \Pi_{,\mu} + \mathcal{M}\,\Pi + \mathcal{N}\,\Pi^2 = 0, \tag{91}$$

where

$$\mathcal{K}^\mu = r_I r_J \frac{\partial b_I}{\partial v_{J,0\mu}} \tag{92}$$

$$\mathcal{M} = r_I \left\{ \frac{\partial b_I}{\partial v_{J,0\mu}} r_{J,\mu} + \left(\frac{\partial P_{IJ}}{\partial v_{K,\mu\nu}} v_{K,0\mu\nu} + \frac{\partial P_{IJ}}{\partial v_{K,\mu}} v_{K,0\mu} + \frac{\partial P_{IJ}}{\partial v_K} v_{K,0} + \frac{\partial P_{IJ}}{\partial x^0} + \frac{\partial b_I}{\partial v_{J,0}} \right) r_J + 2 \frac{\partial P_{IJ}}{\partial v_{K,0}} \left(v_{(J|,00} \right)^- r_{|K)} \right\} \tag{93}$$

$$\mathcal{N} = r_I r_J r_K \frac{\partial P_{IJ}}{\partial v_{K,0}}, \tag{94}$$

and $(v_{J,00})^- \equiv \lim_{x^0 \to -0} v_{J,00}$. The coefficients in Eq. (91) depend only on the field values at $x^0 \to -0$, that is, the background solution on the past side of the characteristic surface. The discontinuity propagates along the integral curve generated by $\mathcal{K}^\mu$, which can be found by integrating

$$\frac{dx^\mu}{ds} = \mathcal{K}^\mu(x^\nu), \tag{95}$$

where we have introduced a parameter s along the integral curve $x^\mu = x^\mu(s)$ and set $s = 0$ at the initial time slice. Then, denoting $\dot{\Pi} \equiv d\Pi(s)/ds$, Eq. (91) may be written as

$$\dot{\Pi} + \mathcal{M}\,\Pi + \mathcal{N}\,\Pi^2 = 0. \tag{96}$$

An equation equivalent to Eq. (96) is obtained also for propagation of weakly-nonlinear high frequency waves, whose frequency is sufficiently large compared to the background time dependence [34, 46, 55–57].

As we can see from the following analysis, the coefficient of the nonlinear term $\mathcal{N}$ plays a key role in the shock formation process. Time evolution of the amplitude $\Pi(s)$ is described by the general solution of Eq. (96), which is given by

$$\Pi(s) = \frac{\Pi(0)e^{-\Phi(s)}}{1+\Pi(0)\int_0^s \mathcal{N}(s')e^{-\Phi(s')}ds'}\,, \tag{97}$$

where $\Phi(s) \equiv \int_0^s \mathcal{M}(s')ds'$. When $\mathcal{N}=0$, $\Pi(s)$ remain finite unless $\mathcal{M}(s)$ diverges. Since $\mathcal{M}(s)$ is determined only by information of the background solution and the characteristic surface, divergence of $\mathcal{M}$ occurs only when the background solution or Σ is not regular.[8] When $\mathcal{N}\neq 0$, $\Pi(s)$ diverges even when the background is regular and $\Phi(s)$ remains finite, since the denominator of (97) may vanish as s increases from zero, particularly for sufficiently large initial amplitude $\Pi(0)$.

At the moment when $\Pi(s)$ diverges, the second derivatives on the future side become infinite hence the first derivatives become discontinuous at Σ. We call it a shock formation in this work. This phenomenon originates from the nonlinear effect due to nonzero $\mathcal{N}$ as we observed above, and also it can be shown that it occurs when two different characteristic surfaces collide with each other in time evolution [31–34]. See also [46] for more details on this shock formation process.

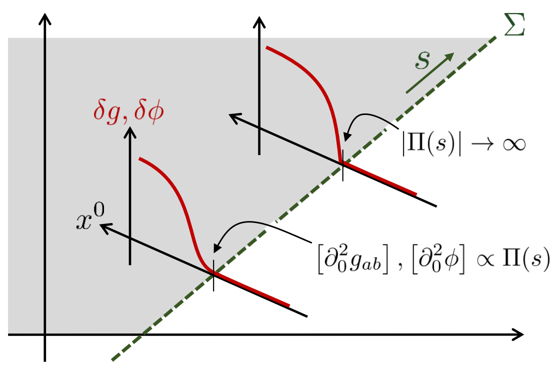


FIG. 5: Propagation of discontinuity in second derivatives of the metric and scalar field, where red curves show profile of difference between the "background solution" without discontinuity and the solution to which the discontinuity is added. Solution in the future region of Σ (shaded region) is influenced by the discontinuity, while that in the past region is not. The discontinuity propagates on Σ, and its amplitude $\Pi(s)$ evolves as it propagates. When $\mathcal{N}\neq 0$, $\Pi(s)$ and hence the second derivatives of fields generically diverge at finite s, which corresponds to a shock formation.

B. Shock formation in the shift-symmetric Horndeski theory

As shown in the previous section, shock formation in second derivatives may occur when $\mathcal{N}$ defined by Eq. (94) does not vanish. In the following, we examine properties of $\mathcal{N}$ and shock formation process

[8] Divergence of $\Phi(s)$ happens when characteristic curves on Σ form a caustic on it by crossing with each other, where the amplitude of wave may diverge due to the focusing effect. To distinguish it from the shock generated by nonlinear effect associated with nonzero $\mathcal{N}$, sometimes this type of shock formation is called a linear shock [46].

in the shift-symmetric Horndeski theory. We first show the expression of $\mathcal{N}$ in this theory on general background, and then examine sufficient conditions for $\mathcal{N}=0$ for a subclass of this theory.

Expression of $\mathcal{N}$ on general background. As shown in Eq. (94), the coefficient of the nonlinear term $\mathcal{N}$ is given by a derivative of the principal symbol of the field equation. In the shift-symmetric Horndeski theory, it is given by

$$\mathcal{N}=\left(r_{ef}\frac{\partial}{\partial g_{ef,0}}+r_\phi\frac{\partial}{\partial\phi_{,0}}\right)\left\{\begin{pmatrix} r^{ab} & r_\phi\end{pmatrix}\begin{pmatrix}\frac{\partial\mathcal{G}_{ab}}{\partial g_{cd,00}} & \frac{\partial\mathcal{G}_{ab}}{\partial\phi_{,00}}\\ \frac{\partial\mathcal{E}_\phi}{\partial g_{cd,00}} & \frac{\partial\mathcal{E}_\phi}{\partial\phi_{,00}}\end{pmatrix}\begin{pmatrix} r_{cd}\\ r_\phi\end{pmatrix}\right\}, \tag{98}$$

where we have taken $\xi_a=(dx^0)_a$ and expressed $\xi_c\xi_d\frac{\partial}{\partial g_{ab,cd}}\equiv\frac{\partial}{\partial g_{ab,00}}$ and $\xi_c\xi_d\frac{\partial}{\partial\phi_{,cd}}\equiv\frac{\partial}{\partial\phi_{,00}}$. The derivatives with respect to $g_{ef,0}$ and $\phi_{,0}$ act only on the components of the two-by-two matrix in Eq. (98). Calculating the terms in this expression, it turns out that that the following relation holds:

$$r_{ef}\frac{\partial^2\mathcal{E}_\phi}{\partial g_{ef,0}\partial\phi_{,00}}=\frac{\partial^2\mathcal{G}_{ab}}{\partial\phi_{,0}\partial\phi_{,00}}r^{ab}=\frac{\partial^2\mathcal{E}_\phi}{\partial\phi_{,0}\partial g_{cd,00}}r_{cd}\,,\quad r_{ef}\frac{\partial^2\mathcal{G}_{ab}}{\partial g_{ef,0}\partial\phi_{,00}}r^{ab}=r_{ef}\frac{\partial^2\mathcal{E}_\phi}{\partial g_{ef,0}\partial g_{cd,00}}r_{cd}=\frac{\partial^2\mathcal{G}_{ab}}{\partial\phi_{,0}\partial g_{cd,00}}r^{ab}r_{cd}\,. \tag{99}$$

Using it, $\mathcal{N}$ is expressed as

$$\mathcal{N}=r^{ab}r_{cd}r_{ef}\frac{\partial}{\partial g_{ef}}\frac{\partial\mathcal{G}_{ab}}{\partial g_{cd,00}}+{r_\phi}^3\frac{\partial}{\partial\phi_{,0}}\frac{\partial\mathcal{E}_\phi}{\partial\phi_{,00}}+3\left(r_\phi r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial\mathcal{G}_{ab}}{\partial\phi_{,00}}r^{ab}+{r_\phi}^2r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial\mathcal{E}_\phi}{\partial\phi_{,00}}\right), \tag{100}$$

where the explicit expressions of the terms of (100) are shown in appendix C 1. This $\mathcal{N}$ does not vanish in general, hence the shock formation occurs generically in this theory, while there are some special cases where $\mathcal{N}=0$, as we see some examples below.

Sufficient conditions for $\mathcal{N}=0$. In this section, let us examine sufficient conditions for $\mathcal{N}$ to vanish. Among the theories incorporated in the shift-symmetric Horndeski theory, we find that the k-essence [47] coupled to GR emerges as the special theory where the scalar sector decouples from the metric sector, and also that the scalar field version of the DBI model [48] turns out to be the unique nontrivial theory to make $\mathcal{N}=0$ on the general background.

For the k-essence coupled to GR ($G_3=G_5=0$, $G_4=$ const. with general $K(X)$), non-vanishing parts of the (trace-reversed) principal symbol can be obtained from expressions in appendix B 2 as

$$\xi_s\xi_t\frac{\partial\mathcal{E}_\phi}{\partial\phi_{,st}}=-K_X\left(g^{ab}-\frac{K_{XX}}{K_X}\phi^{|a}\phi^{|b}\right)\xi_a\xi_b,\quad \xi_s\xi_t r_{qr}\frac{\partial\tilde{\mathcal{G}}_{ab}}{\partial g_{qr,st}}=\frac{G_4}{2}\left(-\xi^2r_{ab}+2\xi_c r^c_{(a}\xi_{b)}-r^c_c\xi_a\xi_b\right). \tag{101}$$

There are no terms that mix the scalar and metric parts, and in this sense the scalar part decouples from the metric part in this analysis. Characteristic surfaces can be found by equating these expressions with zero and solving for ξ.

The metric part in (101) is the same as that of GR, which was already discussed at Eq. (74). The result was that there are two physical modes of the form $r=(r_{ab},r_\phi)=(r_{ab},0)$ for which ξ is null. For the scalar part, we can see that the characteristic equation is given by

$$\left(g^{ab}-\frac{K_{XX}}{K_X}\phi^{|a}\phi^{|b}\right)\xi_a\xi_b=0, \tag{102}$$

where the expression in the parenthesis can be regarded as the effective metric for the scalar mode. A surface whose normal ξ satisfies this equation is characteristic, and the eigenvector is simply given by $r=(r_{ab},r_\phi)=(0,r_\phi)$.

Let us evaluate $\mathcal{N}$ for these modes next. For the k-essence coupled to GR, only (C11) contributes among the terms appearing in $\mathcal{N}$, hence

$$\mathcal{N} = {r_\phi}^3 \frac{\partial}{\partial \phi_{,0}} \frac{\partial \mathcal{E}_\phi^2}{\partial \phi_{,00}} = 3\phi^{|0} K_{XX} \xi^2 - (\phi^{|0})^3 K_{XXX}, \tag{103}$$

where $\phi^{|0} = \xi_a \phi^{|a}$. For the tensor modes, this $\mathcal{N}$ vanishes identically, and hence the shock formation does not occur. For the scalar mode, normalizing $r_\phi = 1$ and using (102), we find

$$\mathcal{N} = \frac{\left(\phi^{|0}\right)^3}{K_X} \left(3{K_{XX}}^2 - K_X K_{XXX}\right). \tag{104}$$

The condition $\mathcal{N} = 0$ to avoid the shock formation is equivalent to $3{K_{XX}}^2 - K_X K_{XXX} = 0$ (assuming $\phi^{|0} \neq 0$), which may be viewed as a differential equation of $K(X)$. A trivial solution is $K \propto X$, which corresponds to a canonical scalar field. The general solution other than this one is given by

$$K = -\lambda\sqrt{c \pm X} + \Lambda, \tag{105}$$

where λ, c, Λ are constants. This is the Lagrangian of the scalar DBI model. Hence, among the theories described by the k-essence coupled to GR, the scalar DBI model (105) is singled out as the nontrivial theory free from the shock formation.

This behavior is the same as that of plane-symmetric simple wave solutions for probe scalar field studied by Refs. [35, 36], where the scalar DBI model turned out to be the theory free from caustics formation.[9] As discussed above, the scalar and metric part decouples in our setup if $G_{3,5}(X)$ and non-constant part of $G_4(X)$ are set to zero. Having this decoupling, it seems natural that the result obtained from our setup coincides with those for the probe scalar field, although the metric sector is taken into account in our setup.

For more general theories, it seems difficult to find characteristics and eigenvectors since the scalar and metric parts do not decouple in the characteristic equation. However, we can still find a sufficient condition to realize $\mathcal{N} = 0$,[10] that is to set

$$\phi^{|0} = \Gamma^0_{\alpha\beta} = 0 \tag{106}$$

on Σ. As we can see from the expressions in appendix C 1, all the terms appearing in $\mathcal{N}$ vanish identically once this condition is imposed.

Γ^0_{ab} is proportional to extrinsic curvature on Σ, that is essentially the derivative in the normal direction of the induced metric on Σ, $\partial_0 g_{\mu\nu}$. Hence, the metric and also the scalar field are locally constant in the x^0 direction. It seems that the shock formation is suppressed when the background has this local homogeneity. The flat spacetime with constant ϕ, and also the Killing horizon with $\phi^{|0} = 0$ imposed additionally are examples in which the condition (106) is satisfied and then the shock formation is suppressed.

[9] See [49] for earlier discussions on exceptional theories for a scalar field on flat spacetime, in which the canonical scalar and the scalar DBI were found to be exceptional, that is, $\mathcal{N}$ vanishes identically and the shock formation does not occur for them. Also the DBI model for a probe vector field is shown to be exceptional [58].

[10] It is sufficient that $\Gamma^0_{\alpha\beta}$ vanishes, since Γ^0_{ab} is always contracted with $\xi_{a'}\xi_{b'}$ and a generalized Kronecker delta in $\mathcal{N}$ (see appendix C 1), and due to this structure Γ^0_{00} and $\Gamma^0_{0\alpha}$ are eliminated and only $\Gamma^0_{\alpha\beta}$ is left in $\mathcal{N}$.

C. Examples of shock formation

To study the properties of shock formation for more general choices of $G_n(X)$ on generic background, we need to look into explicit examples of background solutions and study wave propagation on them. Such a study using explicit background solutions is the main subject of this section.

The first example is the plane wave solution, for which we follow the analysis of Ref. [34] for shock formation on the plane wave solution in Lovelock theories. We introduce coordinates adapted to geodesics in this setup, and examine shock formation using them. We find that the tensor modes or the gravitational wave does not suffer from the shock formation, while the scalar field wave forms a shock in general.

Another example of the background is the dynamical solutions that are symmetric in the two-dimensional angular directions. Some simple solutions appearing in the studies on cosmology and astrophysics, such as the FRW solutions and spherically-symmetric dynamical star (see Fig. 6) fall into this class of solutions, for example. On this background, we again find that the gravitational wave is protected from the shock formation, while the scalar field wave is generically not. We also show an example of nontrivial theory for which even the scalar field wave, not only the gravitational wave, is protected against the shock formation on the FRW universe background.

1. Shock formation on plane wave solution

As the first example, let us examine the shock formation on the plane wave background, for which we studied the propagation of perturbations in section III B 2. We first introduce the coordinates adapted to the wave propagation direction, and then derive the transport equation (96) to check if the amplitude $\Pi(s)$ of the discontinuity diverges or not.

Characteristic surfaces on the plane wave background is given by null hypersurfaces with respect to effective metric (87), which can be transformed as

$$G^{\omega}_{\mu\nu}dx^{\mu}dx^{\nu} = (F-\omega)du^2 + 2dudv + dx^2 + dy^2 = Fdu^2 + 2dudv' + dx^2 + dy^2, \tag{107}$$

where we introduced a new coordinate $v' \equiv v - \frac{1}{2}\omega\, u$. For simplicity, we consider plane-fronted wave propagating from a surface $u = v' = 0$, and focus on the propagation in the (negative) u direction, which is opposite from the direction of ℓ^a in Fig. 4. Such wave propagates along the null geodesics of the effective metric (107), which can be found by solving the geodesic equation for the effective metric $\frac{d^2}{d\lambda^2}x^a + \Gamma^a_{bc}\frac{dx^b}{d\lambda}\frac{dx^c}{d\lambda} = 0$. Its components are given by

$$\ddot{u} = 0, \qquad \ddot{v}' + \frac{1}{2}\dot{u}\left(\dot{u}F_{,u} + 2\dot{x}^{\alpha}F_{,x^{\alpha}}\right) = 0, \qquad \ddot{x}^{\alpha} - \frac{1}{2}\dot{u}\,F_{,x^{\alpha}} = 0, \tag{108}$$

where $x^{\alpha} = x, y$ and a dot denotes a derivative with respect to the affine parameter λ. The first of these implies that we may take $u = \lambda$. Below, for simplicity we consider a background given by

$$\phi(u) = \phi'\,u, \qquad F = -\frac{\kappa}{2}\phi'^2\left(x^2+y^2\right) + A\left(x^2-y^2\right), \tag{109}$$

where ϕ' and A are constants. It corresponds to setting $a_{xx} = -a_{yy} = A$ and $a_{xy} = 0$ in Eq. (79), and making the gradient of the scalar field constant. Then the last of Eq. (108) can be solved by

$$x^{\alpha} = \eta^{\alpha}\cosh\left(\sqrt{A_{\alpha}}\lambda\right) \qquad \left(A_x = A - \frac{1}{4}\kappa\phi'^2,\ \ A_y = -A - \frac{1}{4}\kappa\phi'^2\right), \tag{110}$$

where η^I is the initial position x^I of a geodesic at $\lambda = 0$. Assuming $\kappa > 0$,[11] it is guaranteed that one of A_x and A_y becomes negative at least for any choice of A. We proceed with assuming $A < 0$ (hence $A_x < 0$ at least) below. Then the second of Eq. (108) can be integrated as

$$v' = -\frac{1}{4} \sum_{\alpha=x,y} (\eta^\alpha)^2 \sinh\!\left(2\sqrt{A_\alpha}\lambda\right). \tag{111}$$

Introducing the coordinates adapted to the geodesics by

$$\eta^\alpha = \frac{x^\alpha}{\cosh\!\left(\sqrt{A_\alpha}\lambda\right)}, \qquad x^0 = v - \frac{1}{2}\omega\, u + \frac{1}{4} \sum_{\alpha=x,y} \sqrt{A_\alpha}\, (\eta^\alpha)^2 \sinh\!\left(2\sqrt{A_\alpha}u\right), \tag{112}$$

the physical metric (75) becomes

$$ds^2 = \omega du^2 + 2du dx^0 + \sum_{\alpha=x,y} \cosh^2\!\left(\sqrt{A_\alpha}u\right) (d\eta^\alpha)^2 \,. \tag{113}$$

The characteristic surface is at $x^0 = 0$, and its normal is given by $\xi = dx^0$. This metric becomes singular at $u = \pm\frac{\pi}{2\sqrt{-A_\alpha}}$ for $A_\alpha < 0$, which corresponds to a caustic of the null geodesics. The only nonzero components of $\Gamma^0_{\alpha\beta}$ on $x^0 = 0$ are

$$\Gamma^0_{\alpha\alpha} = -\frac{1}{2}\sqrt{A_\alpha} \sinh\!\left(2\sqrt{A_\alpha}u\right) \qquad \text{(no sum on } \alpha). \tag{114}$$

$\mathcal{N}$ on the plane wave solution can be obtained by plugging the background solution (109), (113), (114) into the general expression (100), whose expressions are given in appendix C 1. We summarize the explicit formula obtained from this procedure in appendix C 2.

For the tensor mode (83), the only term that could contribute to $\mathcal{N}$ is the pure metric term (C3):

$$\mathcal{N} = r^{ab} r_{cd} r_{ef} \frac{\partial}{\partial g_{ef,0}} \frac{\partial \mathcal{G}^5_{ab}}{\partial g_{cd,00}} = \frac{\phi^{|0}}{4} X G_{5X} \delta^{bb_1b_2b_3}_{aa_1a_2a_3} \xi^a \xi_b r^{a_1}_{b_1} r^{a_2}_{b_2} r^{a_3}_{b_3}. \tag{115}$$

This term becomes zero because $X = 0$ for the plane wave solution and also $\delta^{bb_1b_2b_3}_{aa_1a_2a_3} \xi^a \xi_b r^{a_1}_{b_1} r^{a_2}_{b_2} r^{a_3}_{b_3}$ vanishes identically if Eq. (83) is plugged in. Hence, $\mathcal{N}$ vanishes for the tensor modes on the plane wave background, or in other words gravitational wave on this background does not suffer from the shock formation.

For the scalar mode (84), $\mathcal{N}$ is given by

$$\mathcal{N} = C_+\, t_+(u) + C_-\, t_-(u) + C_0, \tag{116}$$

where

$$t_\pm(u) \equiv \sqrt{A_x} \tanh\!\left(\sqrt{A_x}u\right) \pm \sqrt{A_y} \tanh\!\left(\sqrt{A_y}u\right), \tag{117}$$

and $C_{\pm,0}$ are constants given by

$$C_+ = \phi'^2 \left\{ -\frac{1}{2G_4} \left(2G_{3X}G_{4X} + K_X G_{5X}\right) + \frac{2K_{XX}G_{3X}}{K_X} - G_{3XX} \right\} \tag{118}$$

$$C_- = -2A\, G_{5X} \tag{119}$$

---

[11] $\kappa = K_X(0)/G_4(0)$ is defined by Eq. (78), and it will be positive if we require $K_X(0), G_4(0) > 0$, which is the stability conditions for the canonical scalar field minimally coupled to GR.

$$C_0 = \phi'^3 \left\{ \frac{3}{G_4} \left(-G_{3X}{}^2 - K_X G_{4XX} + K_{XX} G_{4X} \right) + \frac{3K_{XX}{}^2}{K_X} - K_{XXX} \right\}. \tag{120}$$

We have also taken the normalization $r_\phi = 1$. For a generic choice of K and $G_{3,4,5}$, $\mathcal{N}$ defined by Eq. (116) becomes nonzero and hence the shock formation occur within finite time, as we showed in section IV A.

Actually, in this case we can show more than that, since we know the detailed structure of the characteristic surface Σ and the properties of $\mathcal{N}$. From the background metric (113), we can evaluate the entire part of the transport equation (96), not only the coefficient $\mathcal{N}$ of the nonlinear term. Based on the full transport equation, we can give an estimate on the time at which the shock formation occurs. Leaving the details to appendix C 3, let us just summarize the main results below. Since $\mathcal{N}$ defined by (116) diverges at $u = \frac{-\pi}{2\sqrt{-A_x}} \equiv u_*$, it can be shown that $\Pi(u)$ diverges at $|u| < |u_*|$ even if the initial amplitude of the discontinuity $\Pi(0)$ is arbitrarily small, and the moment of the shock formation $u = u_0$ is roughly estimated as $u_0/u_* \sim 1 - \left(\frac{\pi(C_+ + C_-)}{2K_X} \Pi(0) \right)^2$ for small enough $\Pi(0)$. Hence, the shock formation occurs for arbitrarily small initial amplitude $\Pi(0)$ on this background, and the shock formation time becomes earlier for larger $\Pi(0)$.

2. Shock formation on dynamical solutions with symmetries

We now focus on another type of dynamical background, for which the two-dimensional angular part of the spacetime is maximally symmetric. For example, the FRW universe, and also spherically-symmetric dynamical solutions fall into this class of background. We will see that the conditions for shock formation can be derived without making much assumptions on the free parameters ($K(X)$, $G_n(X)$) of the theory.

Two-dimensionally maximally-symmetric dynamical solutions. Dynamical solutions with metric whose two-dimensional spatial part is maximally symmetric can be expressed in general as

$$ds^2 = f(\tau,\chi)\left(-d\tau^2 + d\chi^2\right) + \rho^2(\tau,\chi)\gamma_{\alpha\beta}dx^\alpha dx^\beta \equiv f\eta_{AB}dx^A dx^B + \rho^2\gamma_{\alpha\beta}dx^\alpha dx^\beta, \quad \phi = \phi(\tau,\chi), \tag{121}$$

where $x^A = \tau, \chi$, and $\gamma_{\alpha\beta}$ is the metric of the two-dimensional subspace with constant curvature $k = 0, \pm 1$ spanned by x^α. We also assumed that the scalar field shares the symmetry of the spacetime. Due to the symmetry, components of the curvature tensor vanish except for

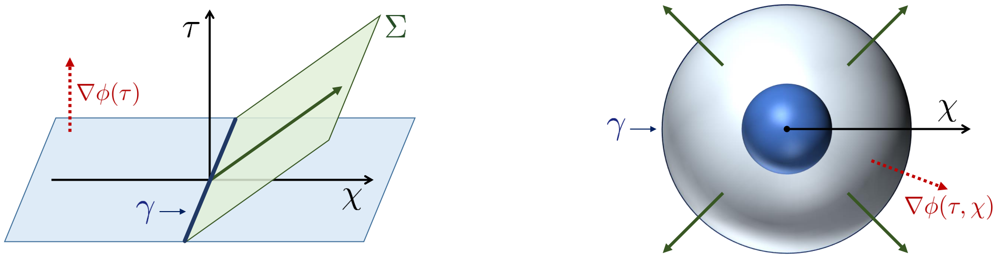
$$R^{A_1A_2}_{B_1B_2} = \frac{\left(\partial_\tau^2 - \partial_\chi^2\right)\log f}{2f}\delta^{A_1A_2}_{B_1B_2} \equiv R_1\delta^{A_1A_2}_{B_1B_2}, \quad R^{A_1\alpha_2}_{B_1\beta_2} \equiv R_2{}^{A_1}_{B_2}\delta^{\alpha_2}_{\beta_2}, \quad R^{\alpha_1\alpha_2}_{\beta_1\beta_2} = \frac{k\,f + \rho_{,\tau}^2 - \rho_{,\chi}^2}{f\rho^2} \equiv R_3\delta^{\alpha_1\alpha_2}_{\beta_1\beta_2}, \tag{122}$$

where $R_2{}^A_B$ is a two-dimensional tensor defined by Eq. (C35). Also, the components of $\phi_{|ab}$ other than $\phi_{|AB}$ and $\phi_{|\alpha\beta}$ vanish, as shown in Eq. (C36). These properties simplifies the evaluation of $\mathcal{N}$ drastically.

Below, we focus on wave whose wavefront shares the symmetry with the background spacetime, that is, we assume that the wavefront is given by a χ-constant surface and ξ_μ has only τ, χ components. For example, plane wave in the flat FRW universe, and spherically-symmetric wavefront around a spherically-symmetric dynamical star fulfill this assumption (see Fig. 6).

Gravitational wave. Characteristic surfaces on the background (121) can be found by solving the characteristic equation (57) following the procedure of section III B. To accomplish it for the gravitational wave, we focus on the vector r given by

$$r = (r_{ab}, r_\phi) = (r^{(\mathrm{T})}_{ab}, 0), \tag{123}$$



(a) FRW universe **(b)** Spherically-symmetric solution

FIG. 6: Examples of two-dimensionally maximally-symmetric dynamical solutions and wave propagation therein. Panel (a) shows a flat FRW universe and wave with plane-symmetric wavefront, and panel (b) shows a spherically-symmetric background and spherically-symmetric wavefront. The ansatz (121) can describe more general solutions than these examples.

where $r^{(\mathrm{T})}_{ab}$ is a traceless tensor which has components only in the angular directions, that is, $r^{(\mathrm{T})}_{AB} = 0$. By explicit calculations, we can check that this vector actually solves the characteristic equation as follows. The scalar and mixed parts of $\tilde{P}\cdot r$ vanish for this r, hence only the metric part shown in appendix B 2 remains nontrivial and is given by

$$\xi_s\xi_t r^{(\mathrm{T})}_{qr}\frac{\partial\tilde{\mathcal{G}}^{4b}{}_{a}}{\partial g_{qr,st}} = -\frac{1}{2}\left(G_4 - 2XG_{4X}\right)\xi^2 r^{(\mathrm{T})b}{}_{a} + \frac{1}{2}G_{4X}\,\delta^{B_1B_2}_{A_1A_2}\xi^{A_1}\xi_{B_1}\phi^{|A_2}\phi_{|B_2}\,r^{(\mathrm{T})b}{}_{a}, \tag{124}$$

$$\xi_s\xi_t r^{(\mathrm{T})}_{qr}\frac{\partial\tilde{\mathcal{G}}^{5b}{}_{a}}{\partial g_{qr,st}} = -\frac{1}{2}XG_{5X}\,\delta^{B_1B_2}_{A_1A_2}\xi^{A_1}\xi_{B_1}\phi^{|A_2}_{|B_2}\,r^{(\mathrm{T})b}{}_{a}. \tag{125}$$

From these expressions, we find that the characteristic equation $\tilde{P}\cdot r = 0$ is satisfied if

$$\left\{\left(G_4 - 2XG_{4X}\right)\delta^B_A - \delta^{BB_1}_{AA_1}\left(G_{4X}\phi^{|A_1}\phi_{|B_1} - XG_{5X}\phi^{|A_1}_{|B_1}\right)\right\}\xi^A\xi_B = 0, \tag{126}$$

where the expression in the curly brackets is the effective metric for gravitational wave on the background (121).[12]

Let us check if shock could form for gravitational wave (123) by evaluating $\mathcal{N}$. The only term that could be nonzero is $r^{ab}r_{cd}r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial\mathcal{G}^5_{ab}}{\partial g_{cd,00}}$ in Eq. (C3), which is proportional to $\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b r^{(\mathrm{T})a_1}{}_{b_1} r^{(\mathrm{T})a_2}{}_{b_2} r^{(\mathrm{T})a_3}{}_{b_3}$. However, $r^{(\mathrm{T})}$ is a traceless tensor living in the two-dimensional angular part of the spacetime, and then Eq. (C3) identically vanishes since it involves three $r^{(\mathrm{T})}$ tensors contracted with a single generalized Kronecker delta. Hence, $\mathcal{N}$ is zero and shock formation does not occur for gravitational wave on the background (121).

Scalar field wave. Next, we study wave that involves the scalar field and examine if it suffers from the shock formation. The first step is to find the characteristic surface and the eigenmode corresponding to the scalar field wave. On the background (121) and for the wavefront that inherits the symmetry of the background, we may assume that the eigenvector r has the same symmetry and is given by

$$r = (r_{ab}, r_\phi), \qquad r_{AB} = r_\eta\,\eta_{AB} + \frac{2r_\gamma}{\xi^2}\xi_A\xi_B, \quad r_{A\alpha} = 0, \quad r_{\alpha\beta} = r_\gamma\,\gamma_{\alpha\beta}, \tag{127}$$

[12] This effective metric for gravitational wave coincides with that of [59], though G_5 was not taken into account in their analysis. Also, the propagation speed obtained from (126) coincides with that derived in [17] when the background solution is set to the FRW universe.

where r_ϕ, r_η and r_γ are functions of τ, χ. The term involving $\xi_A\xi_B$ is the gauge part added so that r satisfies the transverse condition (72).

For the ansatz (127), we may parameterize the eigenvector by $(r_\eta, r_\gamma, r_\phi)$, and then we can solve the characteristic equation (57) to fix the characteristic surface with normal ξ_a and the eigenvector $(r_\eta, r_\gamma, r_\phi)$. Once they are found, the propagation speed of the scalar mode will be given by $c_S = |\xi_\tau/\xi_\chi|$. It turns out that the physical eigenvector is uniquely fixed by this process, although its expression typically becomes lengthy and not illuminating. Below, we show an example for which we can obtain relatively compact expressions for it.

Scalar field wave propagation in the FRW universe. An example is when the background is given by the FRW universe described by

$$f = \rho^2 = a(\tau)^2, \qquad \phi = \phi(\tau), \tag{128}$$

where τ in this case is the conformal time from which a standard time coordinate may be defined by $dt = a(\tau)d\tau$. We will use the Hubble parameter in terms of t, $H(t) \equiv a_{,t}/a$, and also the notation $\dot\phi \equiv \phi_{,t}$ and $X = \dot\phi^2/2$ below. tt and $\chi\chi$ components of the background equation of motion (B4) give the modified Friedmann equations

$$-K + 2XK_X + 6HX\dot\phi G_{3X} - 6H^2\left(G_4 - 4XG_{4X} + 4X^2G_{4XX}\right) + 2H^3X\dot\phi\left(5G_{5X} + 2XG_{5XX}\right) = 0, \tag{129}$$

$$\begin{aligned}\mathcal{E} \equiv K + 2\big(3H^2 + 2\dot H\big)\left(G_4 - 2XG_{4X}\right) - 4H\big(H^2 + \dot H\big)X\dot\phi G_{5X} \\ - 2\ddot\phi\left[XG_{3X} + 2H\dot\phi\left(G_{4X} + 2XG_{4XX}\right) + H^2X\left(3G_{5X} + 2XG_{5XX}\right)\right] = 0.\end{aligned} \tag{130}$$

Properties of the background solution (128) and its perturbations are studied by Ref. [17]. Particularly, propagation speed of the scalar mode is given by $c_S = \sqrt{\mathcal{F}_S/\mathcal{G}_S}$, where

$$\Sigma \equiv X\mathcal{E}_{,X} + \frac{1}{2}H\mathcal{E}_{,H}, \quad \Theta \equiv -\frac{1}{6}\mathcal{E}_{,H}, \quad \mathcal{F}_S \equiv \frac{1}{a}\frac{d}{dt}\left(\frac{a}{\Theta}\mathcal{G}_T^2\right) - \mathcal{F}_T, \quad \mathcal{G}_S \equiv \frac{\Sigma}{\Theta^2}\mathcal{G}_T^2 + 3\mathcal{G}_T. \tag{131}$$

This c_s coincides with the propagation speed obtained from the eigenvalue obtained above once the Friedmann equations (129), (130) are imposed.

Expressions of the propagation speed and the eigenvector r become lengthy in general. One exception is the case discussed in section IV B, where $G_{3,4,5}$ are constants while $K(X)$ is kept general. In this case it follows that

$$(r_{ab}, r_\phi) = (0, r_\phi), \qquad c_S^2 = \frac{K_X}{K_X + 2XK_{XX}}, \tag{132}$$

and also we can check that $\mathcal{N} = 0$ is achieved by the scalar DBI model coupled to GR, as we discussed in section IV B. Another case that gives relatively simple c_S and r is when

$$K = -\lambda\sqrt{1 + c\,X}, \quad G_4 = a_4\sqrt{1 + \tilde c\,X}, \quad G_{3,5} = 0, \tag{133}$$

where $\lambda, c, a_4, \tilde c$ are constants that satisfies $c \neq \tilde c$.[13] In this case, using the Friedmann equations (129), (130) we can simplify the propagation speed and the eigenvector as

$$c_S{}^2 = 1 + c\,X, \qquad (r_\eta, r_\gamma, r_\phi) \propto \left(\frac{\tilde c\,(1 + c\,X)^{3/4}}{c\,\dot\phi}, \frac{\tilde c\,\dot\phi}{(1 + c\,X)^{1/4}}, \sqrt{\frac{6a_4}{\lambda}}\,(1 + \tilde c\,X)^{1/4}\right), \tag{134}$$

[13] When $c = \tilde c$, $\mathcal{F}_S$ and $\mathcal{G}_S$ vanish and then the quadratic Lagrangian for scalar perturbation shown in [17] vanishes identically, which indicates that the theory is in the strong coupling regime.

and we can check that $\mathcal{N}$ identically vanishes even in this case. It can be confirmed by direct calculations that other choices of G_4 such as $G_4 \propto (1+X)^{-1/2}$ typically result in $\mathcal{N} \neq 0$. Hence, it seems that the choice (133) is special among other choices of $K(X)$ and $G_4(X)$ in the sense that it leads to a cosmological solution free from shock formation.

For more general choices of the arbitrary functions $K(X)$, $G_{3,4,5}(X)$, the expressions of the propagation speed and the eigenvector become more lengthy and complicated. We can still make progress in such a case by evaluating those expressions and $\mathcal{N}$ with the aid of computer algebra. In general, we can confirm that $\mathcal{N}$ becomes nonzero hence the scalar field wave suffers from the shock formation.

V. OUTLOOK

The main results and implications obtained in this essay are summarized in section I A. In this section, let us conclude this essay with comments on open questions and possible future directions of the study. Let us categorize them according to the three main topics addressed in this essay.

- **Construction of the bi-Horndeski theory** (section II)

  Construction of Lagrangian using integrability conditions. In this essay, we have successfully constructed the most general equations of motion compatible with the covariance of the theory, while we are still missing the Lagrangian that gives rise those equations as its Euler-Lagrange equations. We can study the classical dynamics in this theory using only the equations of motion, as we exemplified in sections III and IV, but it is better to have the Lagrangian at hand, particularly when we consider quantum effects in this theory or when we apply it to studies in the cosmology context (see e.g. Ref. [17]).

  One of the conditions we have not used in our construction was the integrability condition that guarantees the existence of the Lagrangian

$$\frac{\delta\sqrt{-g}\,\mathcal{G}^{ab}(x)}{\delta g_{cd}(y)} - \frac{\delta\sqrt{-g}\,\mathcal{G}^{cd}(y)}{\delta g_{ab}(x)} = \frac{\delta\sqrt{-g}\,\mathcal{G}^{ab}(x)}{\delta \phi^I(y)} - \frac{\delta\sqrt{-g}\,\mathcal{E}_I(y)}{\delta g_{ab}(x)} = \frac{\delta\sqrt{-g}\,\mathcal{E}_I(x)}{\delta \phi^J(y)} - \frac{\delta\sqrt{-g}\,\mathcal{E}_J(y)}{\delta \phi^I(x)} = 0, \quad (135)$$

  where $\delta/\delta A$ denotes variation with respect to a field A. Explicit formula obtained from this condition can be derived following [60], and it is summarized in [54]. In the single scalar field case, the Lagrangian was constructed without using these conditions. In the multi-scalar field case, however, they may restrict the form of the equations of motion, and then some clues to the construction of the Lagrangian may be obtained from them. Such a trial employing the integrability conditions will be one of the future tasks.

  Construction of Lagrangian using simple solutions/ansatz. A possible approach toward construction of the Lagrangian may be to focus on simpler solutions and construct the Lagrangian that generates them. For example, we could take the mini-superspace approach where we focus only on the FRW-type solutions. The solutions with some symmetries we studied in section IV C 2 may be useful as well. The equations of motion and also the candidate Lagrangian are simplified for these simpler solutions, hence it may be fruitful to examine these subclass of the general solutions first before trying the construction for the general solutions. Also, these simpler solutions can be useful also for applications of this theory to various problems in the context of astrophysics and cosmology, such as studies on cosmological solutions or black hole solutions realized in this theory.

Another possible approach is to introduce ansatz for the Lagrangian and try to find the correct choice of the parameters in the ansatz by comparing the field equations (39) to the Euler-Lagrange equations obtained from the ansatz. The approach taken in section II D is one example of this approach, where we have used the trace of the field equations as the ansatz. There are some theories with multiple scalar fields proposed by different methods (see e.g. [61–63]), and it would be fruitful to examine those theories to see if we can extract any hints for the correct form of the most general Lagrangian.

- **Wave propagation in the bi-Horndeski theory** (section III)

  Phenomenology of super/subluminal propagations. In section III we formulated the method of characteristics for the gravitational wave and scalar field wave in the most-general scalar tensor theory. Using them, we can study the conditions for those waves to propagate at super/subluminal speeds. It would be fruitful to study the phenomenology of those super/subluminal propagation of the waves. When the waves propagate subluminally, they can be emitted as Cherenkov radiation from rapidly-moving particles. Such Cherenkov radiation could be used to constrain the theory observationally [64]. On the other hand, superluminal propagation of waves could be an obstacle for the UV completion of a low-energy effective theory [65], although there are some subtleties about it [66, 67]. In [59], anomalous propagation speed of the gravitational wave and its importance in the future multi-messenger gravitational wave astronomy was investigated. It deserves to re-examine these issues for the general scalar-tensor theories based on the formalism given in this essay. Also, it may be interesting to study the implications of the new terms in the equations of motion of the bi-Horndeski theory, which was discussed in section II C, in the light of the phenomenology of super/subluminal propagations.

  Black hole horizon and causal structure of spacetime. In section III B 1, we showed that a Killing horizon, a null hypersurface at which the spacetime is locally stationary, is not a causal edge in general in the bi-Horndeski theory. We found also that a superluminal propagation across the horizon is shut off if the scalar fields share the Killing symmetry on the horizon. This property may be regarded as a natural generalization of that found for scalar-tensor theories with a non-minimally coupled scalar field studied in Ref. [44]. In [68] the causal structure in Einstein-aether and Hořava–Lifshitz theories were studied, and it was noticed that the universal horizon, which acts as the causal edge for any mode, is orthogonal to the background vector field in those theories. This property seems similar to that found in this work, if the gradient of the background scalar field is identified as such a vector field. It is tempting to conjecture that the property of this kind is universal in gravitational theories with additional degrees of freedom, that is, a Killing horizon becomes an event horizon in any theory if the fields are aligned to the background symmetry direction. It would be interesting to pursue such a possibility and examine its implications.

  Extensions to more general theories. In this essay, we focused on the (bi-)Horndeski theory, which is the most general covariant theory whose Euler-Lagrange equations are up to second order in derivatives. Rather recently, some scalar-tensor theories that encompass the Horndeski theory as their subclass were proposed in the single scalar field case. In these theories, the Euler-Lagrange equations contain higher-derivative terms, unlikely to the Horndeski theory, while the number of the degrees of freedom is unchanged (one plus two from the scalar field and graviton) due to nontrivial constraints built in those theories. A theory in this class was first introduced in [69–71], where the new theory was related to the Horndeski theory by disformal transformation. This theory was dubbed the beyond-Horndeski theory, and later on it was extended further to incorporate higher order terms in Lagrangian without re-introducing the ghost instability [72–83]. There are similar extended theories involving vector degrees of freedom and also derivatives of spacetime curvature [84, 85], which have been attracting wide interest recently. It would be interesting to check if the formalism used in this essay could be extended for these theories. The subtleties are that the equations of motion contain higher order derivatives, and also that nontrivial constraints are present in those theories. The second-order formalism in this essay must be generalized to incorporate them, or it might be helpful to use the first-order formalism employed in [86] for nonlinear massive gravity theory, which has a nontrivial constraint built in the theory.

- **Shock formation in the Horndeski theory** (section IV)

  Gravitational wave and shock formation. In any example studied in this essay, the propagation modes corresponding to gravitational wave were always free from shock formation. If this feature persists on any background solution, we may conclude that the gravitational wave in this class of theory is more well-behaved compared to the scalar field, which typically suffers from shock formation. On the other hand, if we could find some background solutions or nontrivial shape of wavefront for which shock formation occurs even for gravitational wave, it might give interesting implications to gravitational wave observations which was realized recently by LIGO [5]. Such studies on shock formation in gravitational wave with nontrivial wavefront on nontrivial background would be interesting.

  In this essay, we analyzed the shock formation only in the shift-symmetric Horndeski theory for simplicity. It is straightforward to extend it to more general theories including the bi-Horndeski theory. One of the problems in this approach is that not so many solutions have been constructed in these theories, and then we lack the background solutions to use in the analysis. Hence it may be better first to look at simpler theories, such as the generalized multi-Galileon theory [20], and seek for background solutions which can be used for the analysis.

  In Lovelock theories of gravity in higher dimensions, it was shown that the shock formation occurs in gravitational wave [34]. In this theory, the Lagrangian and the equation of motion contain products of the spacetime curvature tensors, and also the gravitational wave have more degrees of freedom compared to the four-dimensional case. These ingredients might have been the source of the shock formation in this theory, because the shock formation is caused by nonlinear self interaction of the wave. In the bi-Horndeski theory in four dimensions, it seems that these ingredients are missing, and maybe it implies that shock formation does not occur even in this theory and also in other multi-field scalar field theories related to it. It would be interesting to examine this expectation.

  Time evolution after shock formation. When a shock forms, derivatives of fields diverge there and the theory would break down unless higher-order corrections to the theory ameliorate the singular behavior, or unless we accept such a field configuration as a weak solution of the theory. Also, a shock formation would correspond to a naked singularity formation unless it is covered by an event horizon or resolved by corrections to the theory, as argued in [34, 36]. It would be interesting to study time evolution after shock formation by taking higher-order corrections into account or by regarding a shock as a weak solution, as we usually do in dynamical simulations of compressible fluids.

To treat a shock as a weak solution in the theory, we need use the junction conditions to impose at the shock surface, which correspond to the Rankine-Hugoniot conditions for shock waves in usual fluids. Such junction conditions have been derived and applied in the Horndeski theory [87, 88]. Using them, it may be possible to simulate a shock formation and time evolution after that starting from smooth initial data. It may be fruitful to conduct such numerical or analytical analysis on phenomena that involve shock formation and to examine their implications in, e.g., dynamical phenomena in astrophysics such as binary star motion or gravitational collapse of matter field.

Nonlinear time evolution with shock formations. As discussed above, the phenomena discussed in this work could be important in time evolution that involves nonlinear dynamics of gravity and scalar fields. In this work, we focused only on the evolution of weak discontinuity, and found that its shock formation is controlled by the parameter $\mathcal{N}$ (94), which is given by a derivative of the principal symbol P of the equations of motion. The importance and implications of this quantity in the full nonlinear dynamics within scalar-tensor theories should be investigated further. Such a study may provide some insights that help us when we apply the scalar-tensor theories to various problems in physics, including those in astrophysics and cosmology.

ACKNOWLEDGMENTS

This work was supported in part by JSPS KAKENHI Grant Numbers 15H03658 and 16H06932.

Appendix A: Formulae for construction of the bi-Horndeski theory

1. The ξ tensors

We summarize the explicit forms of the ξ tensors introduced in section II B to construct the most general equations of motion. As described in section II B, the ξ tensors are constructed by taking the products and linear combinations of $\phi^{I|a}$, g^{ab} and ε^{abcd} so that the resultant ξ tensors have the symmetries of Eqs. (11)–(13). We need to find all such ξ tensors to construct the most general second-order tensor whose divergence remains of second order. In doing so, we find that not all of the ξ tensors give nontrivial contributions to Eq. (19), because there are some degeneracies and the identical terms appear from more than one ξ tensor. The ξ tensors that give nontrivial contributions to Eq. (19) are summarized as

$$\begin{aligned}
\xi_I^{abcd} &= a_I\left(g^{ac}g^{bd} + g^{ad}g^{bc} - 2g^{ab}g^{cd}\right) \\
&\quad + b_{IJK}\left[g^{ac}\phi^{J|b}\phi^{K|d} + g^{ad}\phi^{J|b}\phi^{K|c} + g^{bd}\phi^{J|a}\phi^{K|c} + g^{bc}\phi^{J|a}\phi^{K|d} - 2\left(g^{ab}\phi^{J|c}\phi^{K|d} + g^{cd}\phi^{J|a}\phi^{K|b}\right)\right] \\
&\quad + c_{IJKLM}\left(\phi^{J|a}\phi^{K|c}\phi^{L|b}\phi^{M|d} + \phi^{J|a}\phi^{K|d}\phi^{L|b}\phi^{M|c} - 2\phi^{J|a}\phi^{K|b}\phi^{L|c}\phi^{M|d}\right) \\
&\quad + d_{IJKLM}\left(\phi^{J|a}\phi^{K|c}\varepsilon^{bdef}\phi^L_{|e}\phi^M_{|f} + \phi^{J|a}\phi^{K|d}\varepsilon^{bcef}\phi^L_{|e}\phi^M_{|f}\right), \qquad \text{(A1)} \\
\xi^{abcdef} &= \hat{a}_{IJ}\Big(\epsilon^{Iace}\epsilon^{Jbdf} + \epsilon^{Ibce}\epsilon^{Jadf} + \epsilon^{Iade}\epsilon^{Jbcf} + \epsilon^{Ibde}\epsilon^{Jacf}\Big)
\end{aligned}$$

$$+\hat{b}\,g_{gh}\Big(\varepsilon^{aceg}\varepsilon^{bdfh}+\varepsilon^{bceg}\varepsilon^{adfh}+\varepsilon^{adeg}\varepsilon^{bcfh}+\varepsilon^{bdeg}\varepsilon^{acfh}+\varepsilon^{acfg}\varepsilon^{bdeh}+\varepsilon^{bcfg}\varepsilon^{adeh}+\varepsilon^{adfg}\varepsilon^{bceh}+\varepsilon^{bdfg}\varepsilon^{aceh}\Big), \tag{A2}$$

$$\begin{aligned}\xi^{abcdef}_{IJ} &= \hat{a}_{IJKL}\Big(\epsilon^{Kace}\epsilon^{Lbdf}+\epsilon^{Kbce}\epsilon^{Ladf}+\epsilon^{Kade}\epsilon^{Lbcf}+\epsilon^{Kbde}\epsilon^{Lacf}\Big)\\ &\quad+\hat{b}_{IJ}\,g_{gh}\Big(\varepsilon^{aceg}\varepsilon^{bdfh}+\varepsilon^{bceg}\varepsilon^{adfh}+\varepsilon^{adeg}\varepsilon^{bcfh}+\varepsilon^{bdeg}\varepsilon^{acfh}+\varepsilon^{acfg}\varepsilon^{bdeh}+\varepsilon^{bcfg}\varepsilon^{adeh}+\varepsilon^{adfg}\varepsilon^{bceh}+\varepsilon^{bdfg}\varepsilon^{aceh}\Big),\end{aligned} \tag{A3}$$

$$\xi^{abcdefgh} = \hat{a}\Big(\varepsilon^{aceg}\varepsilon^{bdfh}+\varepsilon^{bceg}\varepsilon^{adfh}+\varepsilon^{adeg}\varepsilon^{bcfh}+\varepsilon^{bdeg}\varepsilon^{acfh}+\varepsilon^{acfg}\varepsilon^{bdeh}+\varepsilon^{bcfg}\varepsilon^{adeh}+\varepsilon^{adfg}\varepsilon^{bceh}+\varepsilon^{bdfg}\varepsilon^{aceh}\Big), \tag{A4}$$

where a_I, b_{IJK}, c_{IJKLM}, d_{IJKLM}, $\hat{a}$, $\hat{a}_{IJ}$, $\hat{a}_{IJKL}$, $\hat{b}$ and $\hat{b}_{IJ}$ are arbitrary functions of ϕ^I and X^{JK} satisfying

$$b_{IJK}=b_{IKJ},\quad c_{IJKLM}=c_{IKJLM}=c_{IJKML}=c_{ILMJK},\quad d_{IJKLM}=-d_{IKJLM}=-d_{IJKML}=d_{ILMJK}, \tag{A5}$$

$$\hat{a}_{IJ}=\hat{a}_{JI},\qquad \hat{a}_{IJKL}=\hat{a}_{IJLK}. \tag{A6}$$

We have also defined $\epsilon^{Iabc}=\varepsilon^{abcd}\phi^I_{|d}$.

2. Coefficient functions of $\nabla^b\mathcal{G}^a_b$

By direct calculations, the coefficients appearing in $\nabla^b\mathcal{G}^a_b$ (23) can be derived and expressed in terms of the coefficients of $\mathcal{G}^a_b$ (21) as

$$\begin{aligned}\alpha_{IJ} &= G_{IJ}-2J_{IJ}+2K_{I,J}-2H_{KLIJ}X^{KL},\qquad \beta_{IJ}=-I_{,IJ}+J_{IJ}-K_{J,I}+2J_{KL,IJ}X^{KL},\\ \gamma_{IJKL} &= -2J_{IJ,KL}+H_{IKJL},\qquad \epsilon_{IJK}=K_{(I,K)J}-\frac{3}{2}L_{KIJ},\\ \mu_I &= \frac{1}{2}C_I+2I_{,I}-\left(D_{JKI}+8J_{J[K,I]}\right)X^{JK}+4E_{JKLMI}X^{JK}X^{LM},\\ \nu_{IJKL} &= -G_{IJ,KL}+3H_{K(IJL)}+2H_{MNIJ,KL}X^{MN}-3L_{LIJ,K},\\ \omega_{IJK} &= -C_{I,JK}+2D_{J(IK)}-2G_{IK,J}+2\left(D_{LMI,JK}-4H_{[J|LIK,|M]}\right)X^{LM}\\ &\quad-16E_{(I|JLM|K)}X^{LM}-8E_{LMNOI,JK}X^{LM}X^{NO},\\ \xi_{IJ} &= -A_{,IJ}+B_{IJ}-C_{J,I}-4D_{K[I|J,|L]}X^{KL}-8E_{KLMNJ,I}X^{KL}X^{MN}+16E_{KIMNJ,L}X^{KL}X^{MN},\\ \zeta_{IJK} &= -\frac{1}{2}D_{IJK}-2J_{IJ,K}+4E_{LMIJK}X^{LM},\qquad \eta_{IJKL}=\frac{1}{2}H_{IJKL},\\ \lambda_{IJKLM} &= \frac{1}{2}D_{IJK,LM}+H_{IJKM,L}-2E_{MLIJK}-4E_{NOIJK,LM}X^{NO},\\ \sigma_{IJKLMN} &= H_{IJKL,MN}-H_{IMNL,JK},\qquad \tau_{IJKLM}=-L_{[I|JK,L|M]},\qquad \iota_{IJK}=-K_{[I,K]J}.\end{aligned} \tag{A7}$$

3. Explicit form of $\mathcal{Q}_I$

The explicit form of $\mathcal{Q}_I$ appearing in Eq. (23) and also in the scalar field equations of motion (40) is given by

$$\mathcal{Q}_I\equiv\mathcal{Q}_I^{(A)}+\mathcal{Q}_I^{(B)}+\mathcal{Q}_I^{(C)}+\mathcal{Q}_I^{(D)}+\mathcal{Q}_I^{(E)}+\mathcal{Q}_I^{(G)}+\mathcal{Q}_I^{(H)}+\mathcal{Q}_I^{(I)}+\mathcal{Q}_I^{(J)}+\mathcal{Q}_I^{(K)}+\mathcal{Q}_I^{(L)}, \tag{A8}$$

where

$$\mathcal{Q}_I^{(A)} = A_{,I}, \tag{A9}$$

$$\mathcal{Q}_I^{(B)} = -2B_{IJ,K}X^{JK} - B_{IJ,KL}\phi^K_{|c}\phi^{L|cb}\phi^J_{|b} + B_{IJ}\phi^{J|b}_{|b}, \tag{A10}$$

$$\mathcal{Q}_I^{(C)} = C_{J,I}\phi^{J|c}_{|c}, \tag{A11}$$

$$\begin{aligned}\mathcal{Q}_I^{(D)} &= D_{JKL,I}\delta^{ce}_{df}\phi^J_{|c}\phi^{K|d}\phi^{L|f}_{|e} + 2D_{IJK,L}X^{JL}\phi^{K|c}_{|c} + D_{IJK,L}\phi^{J|c}\phi^K_{|cd}\phi^{L|d} \\ &+ D_{IJK,LM}\delta^{ce}_{bf}\phi^L_{|d}\phi^{M|db}\phi^J_{|c}\phi^{K|f}_{|e} - D_{IJK}\delta^{ce}_{bf}\phi^{J|b}_{|c}\phi^{K|f}_{|e} - \frac{1}{2}D_{IJK}\delta^{ce}_{bf}\phi^J_{|c}\phi^{K|l}R_{el}{}^{bf},\end{aligned} \tag{A12}$$

$$\begin{aligned}\mathcal{Q}_I^{(E)} &= E_{JKLMN,I}\delta^{ceg}_{dfh}\phi^J_{|c}\phi^{K|d}\phi^L_{|e}\phi^{M|f}\phi^{N|h}_{|g} - 2E_{LJKIM,N}\delta^{ceg}_{dfh}\phi^L_{|c}\phi^{J|d}\phi^K_{|e}\phi^{N|f}\phi^{M|h}_{|g} \\ &- 2E_{LJKIM,NO}\delta^{ceg}_{bfh}\phi^N_{|l}\phi^{O|lb}\phi^L_{|c}\phi^K_{|e}\phi^{J|f}\phi^{M|h}_{|g} + 4E_{LJKIM}\delta^{ceg}_{bfh}\phi^{L|b}_{|c}\phi^K_{|e}\phi^{J|f}\phi^{M|h}_{|g} \\ &+ E_{LJKIM}\delta^{ceg}_{bfh}\phi^L_{|c}\phi^K_{|e}\phi^{J|f}\phi^{M|l}R_{gl}{}^{bh},\end{aligned} \tag{A13}$$

$$\mathcal{Q}_I^{(G)} = G_{JK,I}\delta^{ce}_{df}\phi^{J|d}_{|c}\phi^{K|f}_{|e}, \tag{A14}$$

$$\begin{aligned}\mathcal{Q}_I^{(H)} &= H_{JKLM,I}\delta^{ceg}_{dfh}\phi^J_{|c}\phi^{K|d}\phi^{L|f}_{|e}\phi^{M|h}_{|g} - H_{IJKL}\delta^{ceg}_{bfh}\phi^{J|b}_{|c}\phi^{K|f}_{|e}\phi^{L|h}_{|g} \\ &+ 2H_{IJKL,M}X^{JM}\delta^{eg}_{fh}\phi^{K|f}_{|e}\phi^{L|h}_{|g} + 2H_{IJKL,M}\delta^{cg}_{fh}\phi^J_{|c}\phi^{M|e}\phi^{K|f}_{|e}\phi^{L|h}_{|g} \\ &- H_{IJKL}\delta^{ceg}_{bfh}\phi^J_{|c}\phi^{K|f}_{|e}\phi^{L|l}R_{gl}{}^{bh} + H_{IJKL,MN}\delta^{ceg}_{bfh}\phi^M_{|l}\phi^{N|lb}\phi^J_{|c}\phi^{K|f}_{|e}\phi^{L|h}_{|g},\end{aligned} \tag{A15}$$

$$\mathcal{Q}_I^{(I)} = I_{,I}\delta^{ce}_{df}R_{ce}{}^{df}, \tag{A16}$$

$$\begin{aligned}\mathcal{Q}_I^{(J)} &= J_{JK,I}\delta^{ceg}_{dfh}\phi^J_{|c}\phi^{K|d}R_{eg}{}^{fh} + 2J_{IJ,K}X^{JK}\delta^{eg}_{fh}R_{eg}{}^{fh} + 2J_{IJ,K}\delta^{cg}_{fh}\phi^J_{|c}\phi^{K|e}R_{eg}{}^{fh} \\ &+ J_{IJ,KL}\delta^{ceg}_{bfh}\phi^K_{|d}\phi^{L|db}\phi^J_{|c}R_{eg}{}^{fh} - J_{IJ}\delta^{ceg}_{bfh}\phi^{J|b}_{|c}R_{eg}{}^{fh},\end{aligned} \tag{A17}$$

$$\mathcal{Q}_I^{(K)} = K_{J,I}\delta^{ceg}_{dfh}\phi^{J|d}_{|c}R_{eg}{}^{fh} - \frac{1}{2}K_{J,IK}\delta^{cegl}_{dfhm}\phi^{K|d}_{|c}\phi^{J|f}_{|e}R_{gl}{}^{hm} - \frac{1}{8}K_I\delta^{cegl}_{dfhm}R_{ce}{}^{df}R_{gl}{}^{hm}, \tag{A18}$$

$$\mathcal{Q}_I^{(L)} = L_{JKL,I}\delta^{ceg}_{dfh}\phi^{J|d}_{|c}\phi^{K|f}_{|e}\phi^{L|h}_{|g} - \frac{1}{4}L_{LJK,IM}\delta^{cegl}_{dfhm}\phi^{M|d}_{|c}\phi^{L|f}_{|e}\phi^{J|h}_{|g}\phi^{K|m}_{|l}. \tag{A19}$$

These expressions appear also in the (re-organized) scalar field equations in Eq. (51) and used to derive the principal symbol P in section III B.

4. Euler-Lagrange Equations of the candidate Lagrangian

The explicit form of the Euler-Lagrange equations of the candidate Lagrangian $\mathcal{L}_1$–$\mathcal{L}_7$ defined by Eq. (43) are given by

$$E^{ab}(\mathcal{L}_1) = M^{(1)}_{I,J}\left(\phi^{I|(a}\phi^{J|b)} + g^{ab}X^{IJ}\right) - \frac{1}{2}M^{(1)}_{I,JK}\left(g^{l(b}\delta^{a)ce}_{ldf}\phi^I_{|c}\phi^{J|d}\phi^{K|f}_{|e} + 2X^{IJ}g^{l(b}\delta^{a)c}_{ld}\phi^{K|d}_{|c}\right), \tag{A20}$$

$$\begin{aligned}E^{ab}(\mathcal{L}_2) &= 2M^{(2)}_{,I,J}\left(\phi^{I|(a}\phi^{J|b)} + 2g^{ab}X^{IJ}\right) - 2\left(M^{(2)}_{,I} + 2M^{(2)}_{,K,IJ}X^{JK}\right)g^{l(a}\delta^{b)c}_{ld}\phi^{I|d}_{|c} \\ &- \frac{1}{2}M^{(2)}_{,IJ}g^{l(a}\delta^{b)ceg}_{ldfh}\phi^I_{|c}\phi^{J|d}R_{eg}{}^{fh} + \left(\frac{1}{2}M^{(2)} - M^{(2)}_{,IJ}X^{IJ}\right)g^{l(a}\delta^{b)ce}_{ldf}R_{ce}{}^{df} \\ &- \left(M^{(2)}_{,IJ} + 2M^{(2)}_{,IJ,KL}X^{KL}\right)g^{l(a}\delta^{b)ce}_{ldf}\phi^{I|d}_{|c}\phi^{J|f}_{|e} - M^{(2)}_{,IJ,KL}g^{l(a}\delta^{b)ceg}_{ldfh}\phi^I_{|c}\phi^{J|d}\phi^{K|f}_{|e}\phi^{L|h}_{|g} \\ &- 4M^{(2)}_{,I,JK}g^{l(a}\delta^{b)ce}_{ldf}\phi^I_{|c}\phi^{J|d}\phi^{K|f}_{|e},\end{aligned} \tag{A21}$$

$$\begin{aligned}E^{ab}(\mathcal{L}_3) &= -2M^{(3)}_{IJK,L}X^{IJ}X^{KL}g^{ab} - M^{(3)}_{IJK,L}\left(X^{KL}\phi^{I|(a}\phi^{J|b)} + 2X^{IJ}\phi^{K|(a}\phi^{L|b)} + 2X^{L(I}\phi^{J)|(a}\phi^{K|b)}\right) \\ &+ \frac{3}{2}M^{(3)}_{(IJK)}\left(g^{l(a}\delta^{b)ce}_{ldf}\phi^I_{|c}\phi^{J|d}\phi^{K|f}_{|e} + 2X^{IJ}g^{l(a}\delta^{b)c}_{ld}\phi^{K|d}_{|c}\right) - \frac{1}{2}M^{(3)}_{IJK,LM}g^{l(a}\delta^{b)ceg}_{ldfh}\phi^I_{|c}\phi^{J|d}\phi^K_{|e}\phi^{L|f}\phi^{M|h}_{|g}\end{aligned}$$

$$
+\left(M^{(3)}_{ILM,JK}-2M^{(3)}_{I[JM],KL}\right)X^{LM}\,g^{l(a}\delta^{b)cg}_{lfh}\phi^{I}_{|c}\phi^{J|f}\phi^{K|h}_{|g}+2M^{(3)}_{IJK,LM}X^{IL}X^{JK}\,g^{l(a}\delta^{b)c}_{ld}\phi^{M|d}_{|c},
\tag{A22}
$$

$$
\begin{aligned}
E^{ab}\left(\mathcal{L}_4\right)=&-M^{(4)}_{I,JK}X^{IJ}\,g^{n(a}\delta^{b)egl}_{nfhm}\phi^{K|f}_{|e}R_{gl}{}^{hm}-M^{(4)}_{I,J}X^{IJ}\,g^{l(a}\delta^{b)eg}_{lfh}R_{eg}{}^{fh}-M^{(4)}_{I,J}\,g^{l(a}\delta^{b)ceg}_{ldfh}\phi^{I}_{|c}\phi^{J|d}R_{eg}{}^{fh}\\
&-\frac{2}{3}\left(M^{(4)}_{I,JK,LM}X^{IJ}+M^{(4)}_{K,LM}\right)g^{n(a}\delta^{b)egl}_{nfhm}\phi^{K|f}_{|e}\phi^{L|h}_{|g}\phi^{M|m}_{|l}-2M^{(4)}_{I,J,K}\,g^{l(a}\delta^{b)ce}_{ldf}\phi^{J}_{|c}\phi^{K|d}\phi^{I\,|f}_{|e}\\
&-2\left(M^{(4)}_{I,J}+M^{(4)}_{K,IJ,L}X^{KL}\right)g^{l(a}\delta^{b)ce}_{ldf}\phi^{I\,|d}_{|c}\phi^{J\,|f}_{|e}-2M^{(4)}_{I,K(J,L)}\,g^{l(a}\delta^{b)ceg}_{ldfh}\phi^{J}_{|c}\phi^{L|d}\phi^{I\,|f}_{|e}\phi^{K|h}_{|g},
\end{aligned}
\tag{A23}
$$

$$
\begin{aligned}
E^{ab}\left(\mathcal{L}_5\right)=&\left(M^{(5)}_{IJ}+2M^{(5)}_{IK,JL}X^{KL}\right)X^{IJ}\,g^{l(a}\delta^{b)eg}_{lfh}R_{eg}{}^{fh}\\
&+\left[M^{(5)}_{IJ}+\left(2M^{(5)}_{IK,JL}-M^{(5)}_{IJ,KL}\right)X^{KL}\right]g^{l(a}\delta^{b)ceg}_{ldfh}\phi^{I}_{|c}\phi^{J|d}R_{eg}{}^{fh}\\
&+2\left[M^{(5)}_{IJ}+\left(M^{(5)}_{KL,IJ}+4M^{(5)}_{IK,JL}\right)X^{KL}+2M^{(5)}_{MN,KL,IJ}X^{KM}X^{LN}\right]g^{l(a}\delta^{b)ce}_{ldf}\phi^{I\,|d}_{|c}\phi^{J\,|f}_{|e}\\
&+2\left[2M^{(5)}_{IK,JL}+\left(2M^{(5)}_{IM,JN,KL}-M^{(5)}_{IJ,KL,MN}\right)X^{MN}\right]g^{l(a}\delta^{b)ceg}_{ldfh}\phi^{I}_{|c}\phi^{J|d}\phi^{K|f}_{|e}\phi^{L\,|h}_{|g}\\
&+2\left[2M^{(5)}_{KI,J}-M^{(5)}_{IJ,K}+2\left(2M^{(5)}_{IL,J,MK}-M^{(5)}_{IJ,L,MK}\right)X^{LM}\right]g^{l(a}\delta^{b)ce}_{ldf}\phi^{I}_{|c}\phi^{J|d}\phi^{K|f}_{|e}\\
&-16M^{(5)}_{I[J,L],K}X^{IJ}X^{KL}g^{ab}-4\left(M^{(5)}_{IJ,K,L}+M^{(5)}_{KL,I,J}-2M^{(5)}_{IK,J,L}\right)X^{KL}\phi^{I|(a}\phi^{J|b)}\\
&-4M^{(5)}_{IJ,K,LM}g^{l(a}\delta^{b)ceg}_{ldfh}\phi^{I}_{|c}\phi^{J|d}\phi^{K}_{|e}\phi^{L|f}\phi^{M|h}_{|g},
\end{aligned}
\tag{A24}
$$

$$
E^{ab}\left(\mathcal{L}_6\right)=\frac{1}{2}M^{(6)}g^{ab}+\frac{1}{2}M^{(6)}_{,IJ}\phi^{I|(a}\phi^{J|b)},
\tag{A25}
$$

$$
\begin{aligned}
E^{ab}\left(\mathcal{L}_7\right)=&\left(M^{(7)}_{IJKLM}+M^{(7)}_{IJMLK}+M^{(7)}_{IJKML}\right)g^{l(a}\delta^{b)ceg}_{ldfh}\phi^{I}_{|c}\phi^{J|d}\phi^{K}_{|e}\phi^{L|f}\phi^{M|h}_{|g}\\
&+4\left(M^{(7)}_{IJKLM}+M^{(7)}_{KLMIJ}+M^{(7)}_{KLIMJ}\right)X^{IJ}\,g^{l(a}\delta^{b)ce}_{ldf}\phi^{K}_{|c}\phi^{L|d}\phi^{M|f}_{|e}\\
&+\left(-M^{(7)}_{IJKLN,OM}+2M^{(7)}_{IJNLK,OM}+2M^{(7)}_{INKLO,JM}\right)X^{NO}\,g^{l(a}\delta^{b)ceg}_{ldfh}\phi^{I}_{|c}\phi^{J|d}\phi^{K}_{|e}\phi^{L|f}\phi^{M|h}_{|g}\\
&+8M^{(7)}_{NLIJM,KO}X^{LM}X^{NO}\,g^{l(a}\delta^{b)ce}_{ldf}\phi^{I}_{|c}\phi^{J|d}\phi^{K|f}_{|e}+8M^{(7)}_{IJKLM,N}X^{IJ}X^{KL}X^{MN}\,g^{ab}\\
&+8\left(M^{(7)}_{IJKLM,N}-2M^{(7)}_{IJ(NK)M,L}+M^{(7)}_{IJMNK,L}\right)X^{IJ}X^{KL}\phi^{M|(a}\phi^{N|b)}.
\end{aligned}
\tag{A26}
$$

The partial differential equations (46) and (47) on the arbitrary functions $M^{(1,\ldots,7)}$ of the candidate Lagrangian (43) are obtained by the above equations to the most general equations of motion (39) of the bi-Horndeski theory.

Appendix B: Formulae for analysis on wave propagation in the bi-Horndeski theory

1. Coefficients of the bi-Horndeski theory field equations

In section III B, the equations of motion of the bi-Horndeski theory (39) and (40) is re-expressed in the form of Eq. (51) to find the principal symbol P from it by Eq. (56). The coefficients $\mathcal{A}^{\mu\nu,\rho\sigma},\mathcal{B}^{\mu\nu}_{I},\mathcal{C}^{\mu\nu}$ of the metric equation $\mathcal{G}^{\mu\nu}$ are already given by Eqs. (52)–(54). The coefficients $\widetilde{\mathcal{A}}^{\mu\nu}_{I},\widetilde{\mathcal{B}}_{IJ},\widetilde{\mathcal{C}}_{I}$ of the scalar equation are given by

$$
\begin{aligned}
\widetilde{\mathcal{A}}^{\mu\nu}_{I}=&\left(D_{IJK}-8J_{J[K,I]}-8E_{LKJIM}X^{LM}\right)\delta^{c0(\mu}_{dml}g^{\nu)l}\phi^{J}_{|c}\phi^{K|d}g^{m0}+2D_{IJK}X^{JK}\delta^{0(\mu}_{ml}g^{\nu)l}g^{m0}\\
&+2E_{LJKIM}\delta^{ce0(\mu}_{dfmh}g^{\nu)h}\phi^{L}_{|c}\phi^{J|d}\phi^{K}_{|e}\phi^{M|f}g^{m0}+4J_{IL,JK}g^{l(\mu}\delta^{\nu)ce0}_{ldfh}\phi^{J}_{|c}\phi^{L|d}\phi^{K|f}_{|e}g^{h0}\\
&-4I_{,I}g^{d(\mu}\delta^{\nu)0}_{df}g^{f0}+4\left(J_{IJ}-K_{(I,J)}+2J_{K(I,J)L}X^{KL}\right)\delta^{c0(\mu}_{bhf}g^{\nu)f}\phi^{J|b}_{|c}g^{h0}
\end{aligned}
$$

$$+ 2K_{J,IK}\delta^{ce0(\mu}_{dfmh}g^{\nu)h}\phi^{J|d}_{|c}\phi^{K|f}_{|e}g^{m0} + K_I\delta^{ce0(\mu}_{dfmh}g^{\nu)h}R_{ce}{}^{df}g^{m0}, \tag{B1}$$

$$\begin{aligned}
\widetilde{\mathcal{B}}_{IJ} &= 2B_{IJ}g^{00} - 2B_{IK,LJ}\phi^{K|0}\phi^{L|0} + 2C_{J,I}g^{00} + 4D_{K[I|J,|L]}\left(2X^{KL}g^{00} + \phi^{K|0}\phi^{L|0}\right) \\
&+ 2D_{IKJ,LM}\delta^{c0}_{bf}\phi^K_{|c}\phi^{L|d}\phi^{M|b}_{|d}g^{f0} + 2D_{IKL,MJ}\delta^{ce}_{bf}\phi^K_{|c}\phi^{L|f}_{|e}\phi^{M|0}g^{b0} - 4D_{I(JK)}\delta^{c0}_{bf}\phi^{K|b}_{|c}g^{f0} \\
&+ 2\left(E_{KLMNJ,I} - 2E_{KLMIJ,N}\right)\delta^{ce0}_{dfh}\phi^K_{|c}\phi^{L|d}\phi^M_{|e}\phi^{N|f}g^{h0} \\
&- 4E_{KLMIJ,NO}\delta^{ce0}_{bfh}\phi^K_{|c}\phi^{L|f}\phi^M_{|e}\phi^{N|l}\phi^{O|b}_{|l}g^{h0} - 4E_{KLMIN,OJ}\delta^{ceg}_{bfh}\phi^K_{|c}\phi^{L|f}\phi^M_{|e}\phi^{N|h}_{|g}\phi^{O|0}g^{b0} \\
&+ 16E_{(K|LMI|J)}\delta^{ce0}_{bfh}\phi^{K|b}_{|c}\phi^M_{|e}\phi^{L|f}g^{h0} + 4G_{JK,I}\delta^{c0}_{df}g^{f0}\phi^{K|d}_{|c} \\
&+ 8\left(J_{KL,MJ,I} - J_{IK,MJ,L}\right)\delta^{ce0}_{dfh}\phi^K_{|c}\phi^{L|d}\phi^{M|f}_{|e}g^{h0} - 12J_{I(J,KL)}\delta^{ce0}_{bfh}\phi^{K|b}_{|c}\phi^{L|f}_{|e}g^{h0} \\
&- 4J_{KM,LN,IJ}\delta^{ceg0}_{dfhm}\phi^K_{|c}\phi^{M|d}\phi^{L|f}_{|e}\phi^{N|h}_{|g}g^{m0} - 8J_{IK,LM,NJ}X^{KM}\delta^{eg0}_{fhm}g^{m0}\phi^{L|f}_{|e}\phi^{N|h}_{|g} \\
&- 2J_{KL,IJ}\delta^{ceg0}_{dfhm}g^{m0}\phi^K_{|c}\phi^{L|d}R_{eg}{}^{fh} \\
&- 2\left(-K_{(I,J)} + J_{IJ} + 2J_{K(I,J)L}X^{KL}\right)\delta^{0eg}_{dfh}g^{0d}R_{eg}{}^{fh} - 2K_{J,IK}\delta^{0egl}_{dfhm}g^{0d}\phi^{K|f}_{|e}R_{gl}{}^{hm} \\
&+ 4K_{J,KL,I}\delta^{ce0}_{dfh}\phi^{K|d}_{|c}\phi^{L|f}_{|e}g^{h0} - \frac{4}{3}K_{L,JK,IM}\delta^{ceg0}_{dfhm}\phi^{M|d}_{|c}\phi^{L|f}_{|e}\phi^{K|h}_{|g}g^{m0},
\end{aligned} \tag{B2}$$

$$\widetilde{\mathcal{C}}_I = \widetilde{\mathcal{C}}_I\big(g_{\rho\sigma}, g_{\rho\sigma,0}, g_{\rho\sigma,\kappa}, g_{\rho\sigma,\kappa 0}, g_{\rho\sigma,\kappa\lambda}, \phi_I, \phi_{I,0}, \phi_{I,\rho}, \phi_{I,0\rho}, \phi_{I,\rho\sigma}\big). \tag{B3}$$

2. Equations of motion and principal symbol of the shift-symmetric Horndeski theory

Metric equation following from the Lagrangian (67) is given by

$$\mathcal{G}_{ab} = \sum_{n=2}^{5}\mathcal{G}^n_{ab} = 0, \tag{B4}$$

where

$$\mathcal{G}^{2b}_{\ a} = -\frac{1}{2}K_X\phi_{|a}\phi^{|b} - \frac{1}{2}K\delta^b_a \tag{B5}$$

$$\mathcal{G}^{3b}_{\ a} = -\frac{1}{2}G_{3X}\left(\delta^{bb_1b_2}_{aa_1a_2}\phi^{|a_1}_{|b_1}\phi^{|a_2}\phi_{|b_2} + 2X\delta^{bb_1}_{aa_1}\phi^{|a_1}_{|b_1}\right) \tag{B6}$$

$$\begin{aligned}
\mathcal{G}^{4b}_{\ a} &= \frac{1}{2}\left(G_{4X} + 2XG_{4XX}\right)\delta^{bb_1b_2}_{aa_1a_2}\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2} + \frac{1}{2}G_{4XX}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}\phi_{|b_3} \\
&\quad - \frac{1}{4}\left(G_4 - 2XG_{4X}\right)\delta^{bb_1b_2}_{aa_1a_2}R^{a_1a_2}_{b_1b_2} + \frac{1}{4}G_{4X}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}R^{a_1a_2}_{b_1b_2}\phi^{|a_3}\phi_{|b_3}
\end{aligned} \tag{B7}$$

$$\mathcal{G}^{5b}_{\ a} = -\frac{1}{6}\left(G_{5X} + XG_{5XX}\right)\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}_{|b_3} - \frac{1}{4}XG_{5X}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}R^{a_1a_2}_{b_1b_2}\phi^{|a_3}_{|b_3}. \tag{B8}$$

The scalar equation of motion is given by

$$\mathcal{E}_\phi = \sum_{n=2}^{5}\mathcal{E}^n_\phi = 0, \tag{B9}$$

where

$$\mathcal{E}^2_\phi = -K_X\Box\phi + K_{XX}\phi_{|a_1}\phi^{|a_1}_{|a_2}\phi^{|a_2} \tag{B10}$$

$$\mathcal{E}^3_\phi = \left(G_{3X} + XG_{3XX}\right)\delta^{b_1b_2}_{a_1a_2}\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2} + \frac{1}{2}G_{3XX}\delta^{b_1b_2b_3}_{a_1a_2a_3}\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}\phi_{|b_3} - G_{3X}R^{a_1}_{b_1}\phi_{|a_1}\phi^{b_1} \tag{B11}$$

$$\mathcal{E}^4_\phi = -\frac{1}{6}\left(3G_{4XX} + 4XG_{4XXX}\right)\delta^{b_1b_2b_3}_{a_1a_2a_3}\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}_{|b_3} - \frac{1}{3}G_{4XXX}\delta^{b_1b_2b_3b_4}_{a_1a_2a_3a_4}\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}_{|b_3}\phi^{|a_4}\phi_{|b_4}$$

$$-\frac{1}{2}\left(G_{4X}+2XG_{4XX}\right)\delta^{b_1b_2b_3}_{a_1a_2a_3}\phi^{|a_1}_{|b_1}R^{a_2a_3}_{b_2b_3}-\frac{1}{2}G_{4XX}\delta^{b_1b_2b_3b_4}_{a_1a_2a_3a_4}\phi^{|a_1}_{|b_1}R^{a_2a_3}_{b_2b_3}\phi^{|a_4}\phi_{|b_4} \tag{B12}$$

$$\mathcal{E}^5_\phi=\frac{1}{12}\left(2G_{5XX}+XG_{5XXX}\right)\delta^{b_1b_2b_3b_4}_{a_1a_2a_3a_4}\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}_{|b_3}\phi^{|a_4}_{|b_4}$$
$$+\frac{1}{4}\left(G_{5X}+XG_{5XX}\right)\delta^{b_1b_2b_3b_4}_{a_1a_2a_3a_4}R^{a_1a_2}_{b_1b_2}\phi^{|a_3}_{|b_3}\phi^{|a_4}_{|b_4}+\frac{1}{16}XG_{5X}\delta^{b_1b_2b_3b_4}_{a_1a_2a_3a_4}R^{a_1a_2}_{b_1b_2}R^{a_3a_4}_{b_3b_4}. \tag{B13}$$

The above equations hold in any dimensions except for $\mathcal{G}^5_{\mu\nu}$ and $\mathcal{E}^5_\phi$, which are simplified by using the Schouten identities $\delta^{bb_1b_2b_3b_4}_{aa_1a_2a_3a_4}R^{a_1a_2}_{b_1b_2}\phi^{|a_3}_{|b_3}\phi^{|a_4}\phi_{|b_4}=\delta^{bb_1b_2b_3b_4}_{aa_1a_2a_3a_4}\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}_{|b_3}\phi^{|a_4}\phi_{|b_4}=0$ that hold only in four or lower dimensions.

In the main text, we considered the trace-reversed equations of motion defined by deforming the expressions above following Eq. (68), and then we defined the principal symbol $\tilde{P}$ (69) by taking its derivatives with respect to $\xi_s\xi_t g_{qr,st}$ and $\xi_s\xi_t\phi_{,st}$. We show their expressions below. For the derivation we use the fact that the Riemann tensor is given by $R_{abcd}=-2g_{[a|[c,d]|b]}+\cdots$, hence its derivative with respect to $\xi_s\xi_t g_{qr,st}$ is $\frac{\partial R^{b_1b_2}_{a_1a_2}}{\partial g_{qr,st}}r_{qr}\xi_s\xi_t=-2r^{[b_1}_{[a_1}\xi_{a_2]}\xi^{b_2]}$.

The pure metric part is given by

$$\xi_s\xi_t r_{qr}\frac{\partial\tilde{\mathcal{G}}^{2b}{}_a}{\partial g_{qr,st}}=\xi_s\xi_t r_{qr}\frac{\partial\tilde{\mathcal{G}}^{3b}{}_a}{\partial g_{qr,st}}=0 \tag{B14}$$

$$\xi_s\xi_t r_{qr}\frac{\partial\tilde{\mathcal{G}}^{4b}{}_a}{\partial g_{qr,st}}=\frac{G_4-2XG_{4X}}{2}\left(\delta^{bb_1b_2}_{aa_1a_2}-\delta^b_a\delta^{b_1b_2}_{a_1a_2}\right)\xi^{a_1}\xi_{b_1}r^{a_2}_{b_2}-\frac{G_{4X}}{2}\left(\delta^{bb_1b_2b_3}_{aa_1a_2a_3}-\frac{1}{2}\delta^b_a\delta^{b_1b_2b_3}_{a_1a_2a_3}\right)\xi^{a_1}\xi_{b_1}r^{a_2}_{b_2}\phi^{|a_3}\phi_{|b_3} \tag{B15}$$

$$\xi_s\xi_t r_{qr}\frac{\partial\tilde{\mathcal{G}}^{5b}{}_a}{\partial g_{qr,st}}=\frac{1}{2}XG_{5X}\left(\delta^{bb_1b_2b_3}_{aa_1a_2a_3}-\frac{1}{2}\delta^b_a\delta^{b_1b_2b_3}_{a_1a_2a_3}\right)\xi^{a_1}\xi_{b_1}r^{a_2}_{b_2}\phi^{|a_3}_{|b_3}, \tag{B16}$$

and the mixed part between the metric and scalar parts is given by

$$\xi_s\xi_t r_{qr}\frac{\partial\mathcal{E}^2_\phi}{\partial g_{qr,st}}=0 \tag{B17}$$

$$\xi_s\xi_t r_{qr}\frac{\partial\mathcal{E}^3_\phi}{\partial g_{qr,st}}=-\frac{1}{2}G_{3X}\left(\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b r^{b_1}_{a_1}\phi^{|a_2}\phi_{|b_2}+2X\delta^{aa_1}_{bb_1}\xi^a\xi_b r^{a_2}_{b_2}\right) \tag{B18}$$

$$\xi_s\xi_t r_{qr}\frac{\partial\mathcal{E}^4_\phi}{\partial g_{qr,st}}=\left(G_{4X}+2XG_{4XX}\right)\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b r^{a_1}_{b_1}\phi^{|a_2}_{|b_2}+G_{4XX}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b r^{a_1}_{b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}\phi_{|b_3} \tag{B19}$$

$$\xi_s\xi_t r_{qr}\frac{\partial\mathcal{E}^5_\phi}{\partial g_{qr,st}}=-\frac{1}{2}\left(G_{5X}+XG_{5XX}\right)\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b r^{a_1}_{b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}_{|b_3}-\frac{1}{4}XG_{5X}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b r^{a_1}_{b_1}R^{a_2a_3}_{b_2b_3}. \tag{B20}$$

Finally, the scalar part is given by (where $\xi^2\equiv\xi_a\xi^a$ and $\xi\cdot\phi\equiv\xi_a\phi^{|a}$)

$$\xi_s\xi_t\frac{\partial\mathcal{E}^2_\phi}{\partial\phi_{,st}}=-K_X\xi^2+K_{XX}(\xi\cdot\phi)^2 \tag{B21}$$

$$\xi_s\xi_t\frac{\partial\mathcal{E}^3_\phi}{\partial\phi_{,st}}=2\left(G_{3X}+XG_{3XX}\right)\delta^{bb_1}_{aa_1}\xi^a\xi_b\phi^{|a_1}_{|b_1}+G_{3XX}\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b\phi^{|a_1}_{|b_1}\phi^{|a_2}\phi_{|b_2} \tag{B22}$$

$$\xi_s\xi_t\frac{\partial\mathcal{E}^4_\phi}{\partial\phi_{,st}}=-\left(3G_{4XX}+2XG_{4XXX}\right)\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}-G_{4XXX}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}\phi_{|b_3}$$
$$-\frac{1}{2}\left(G_{4X}+2XG_{4XX}\right)\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b R^{a_1a_2}_{b_1b_2}-\frac{1}{2}G_{4XX}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b R^{a_1a_2}_{b_1b_2}\phi^{|a_3}\phi_{|b_3} \tag{B23}$$

$$\xi_s\xi_t\frac{\partial\mathcal{E}^5_\phi}{\partial\phi_{,st}}=\frac{2G_{5XX}+XG_{5XXX}}{3}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}_{|b_3}+\frac{G_{5X}+XG_{5XX}}{2}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b\phi^{|a_1}_{|b_1}R^{a_2a_3}_{b_2b_3}. \tag{B24}$$

Expressions with non-contracted indices is obtained just by removing one of r^a_b from the expressions above.

The other mixed part between the metric and scalar parts, $\xi_s\xi_t\frac{\partial\tilde{\mathcal{G}}^b_a}{\partial\phi_{,st}}$. can be obtained as follows. Since the equations of motion (B4) and (B9) of the shift-symmetric Horndeski theory are derived from the Lagrangian (67), they automatically satisfy the integrability conditions (135) as we argued at Eq. (58) and also at Eq. (89). It implies that the principal symbol P (before trace-reversing) becomes symmetric, that is,

$$\xi_c\xi_d\frac{\partial\mathcal{G}^{ab}}{\partial\phi_{,cd}}=\xi_c\xi_d\frac{\partial\mathcal{E}_\phi}{\partial g_{ab,cd}}. \tag{B25}$$

Using this symmetry, $\xi_s\xi_t\frac{\partial\tilde{\mathcal{G}}^b_a}{\partial\phi_{,st}}$ is obtained just by trace-reversing $\xi_c\xi_d\frac{\partial\mathcal{E}_\phi}{\partial g_{ab,cd}}$ in accord with Eq. (68). Then, the components of the principal symbol (69) can be obtained by summing the above expressions over $n=2,3,4,5$.

Appendix C: Formulae for analysis on shock formation in the Horndeski theory

1. $\mathcal{N}$ of the shift-symmetric Horndeski theory

We summarize the explicit formula of $\mathcal{N}$ in the shift-symmetric Horndeski theory discussed in section IV B. This quantity is derived by taking derivatives of the principal symbol obtained in section III B and appendix B 2. In this appendix, we use the coordinate x^0 where $\xi_a=\left(dx^0\right)_a$ is a normal of a x^0-constant surface. We also utilize the following formula for derivatives by $\xi_c\frac{\partial}{\partial g_{ab,c}}=\frac{\partial}{\partial g_{ab,0}}$ and $\xi_a\frac{\partial}{\partial\phi_{,a}}=\frac{\partial}{\partial\phi_{,0}}$:

$$\begin{aligned}
&r_{ab}\frac{\partial R^{a_1a_2}_{b_1b_2}}{\partial g_{ab,0}}=2\Gamma^{0[a_1}_{\ [b_1}r^{a_2]}_{b_2]},\qquad \frac{\partial\phi^2}{\partial\phi_{,0}}=2\phi^{|0},\qquad \frac{\partial X}{\partial\phi_{,0}}=-\phi^{|0},\qquad \frac{\partial\phi_a}{\partial\phi_{,0}}=\delta^0_a=\xi_a,\\
&\frac{\partial\phi_{ab}}{\partial\phi_{,0}}=-\Gamma^{0a}_{\ \ b},\qquad r_{cd}\frac{\partial\phi^a_b}{\partial g_{cd,0}}=\frac{1}{2}\left(\phi^{|0}r^a_b-\xi^a r_{bc}\phi^{|c}-\xi_b r^a_c\phi^{|c}\right)\simeq\frac{1}{2}\phi^{|0}r^a_b.
\end{aligned} \tag{C1}$$

The last equality ($\simeq$) holds only when $\frac{\partial\phi^a_b}{\partial g_{cd,0}}$ is contracted with the generalized Kronecker delta multiplied by $\xi^a\xi_b$, by which the terms involving ξ vanish identically.

We summarize the terms appearing in $\mathcal{N}$ (100) in general spacetime below. First, the pure metric terms proportional to ${r_{ab}}^3$ are given by

$$r^{ab}r_{cd}r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial\mathcal{G}^{2,3,4}_{ab}}{\partial g_{cd,00}}=0 \tag{C2}$$

$$r^{ab}r_{cd}r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial\mathcal{G}^{5}_{ab}}{\partial g_{cd,00}}=\frac{\phi^{|0}}{4}XG_{5X}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b r^{a_1}_{b_1}r^{a_2}_{b_2}r^{a_3}_{b_3}. \tag{C3}$$

Next, the mixed terms proportional to ${r_{ab}}^2r_\phi$ are given by

$$r^{ab}r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial\mathcal{G}^{2,3}_{ab}}{\partial\phi_{,00}}=0 \tag{C4}$$

$$r^{ab}r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial\mathcal{G}^{4}_{ab}}{\partial\phi_{,00}}=\frac{\phi^{|0}}{2}\left(G_{4X}+2XG_{4XX}\right)\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b r^{a_1}_{b_1}r^{a_2}_{b_2}+\frac{\phi^{|0}}{2}G_{4XX}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b r^{a_1}_{b_1}r^{a_2}_{b_2}\phi^{|a_3}\phi_{|b_3} \tag{C5}$$

$$r^{ab}r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial \mathcal{G}^5_{ab}}{\partial \phi_{,00}} = -\frac{\phi^{|0}}{2}\left(G_{5X}+XG_{5XX}\right)\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b r^{a_1}_{b_1}r^{a_2}_{b_2}\phi^{|a_3}_{|b_3} - \frac{XG_{5X}}{2}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b r^{a_1}_{b_1}r^{a_2}_{b_2}\Gamma^{0a_3}_{\ \ b_3}, \tag{C6}$$

where the overall factor r_ϕ is omitted. The other mixed terms proportional to $r_{ab}{r_\phi}^2$ are given by

$$r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial \mathcal{E}^2_\phi}{\partial \phi_{,00}} = 0 \tag{C7}$$

$$r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial \mathcal{E}^3_\phi}{\partial \phi_{,00}} = \phi^{|0}\left(G_{3X}+XG_{3XX}\right)\delta^{bb_1}_{aa_1}\xi^a\xi_b r^{a_1}_{b_1} + \frac{\phi^{|0}}{2}G_{3XX}\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b r^{a_1}_{b_1}\phi^{|a_2}\phi_{|b_2} \tag{C8}$$

$$\begin{aligned} r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial \mathcal{E}^4_\phi}{\partial \phi_{,00}} &= -\phi^{|0}\xi^a\xi_b\phi^{|a_1}_{|b_1}r^{a_2}_{b_2}\left\{\left(3G_{4XX}+2XG_{4XXX}\right)\delta^{bb_1b_2}_{aa_1a_2} + G_{4XXX}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\phi^{|a_3}\phi_{|b_3}\right\} \\ &\quad - \left(G_{4X}+2XG_{4XX}\right)\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b r^{a_1}_{b_1}\Gamma^{0a_2}_{\ \ b_2} - G_{4XX}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b r^{a_1}_{b_1}\Gamma^{0a_2}_{\ \ b_2}\phi^{|a_3}\phi_{|b_3} \end{aligned} \tag{C9}$$

$$\begin{aligned} r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial \mathcal{E}^5_\phi}{\partial \phi_{,00}} &= \frac{1}{4}\left(G_{5X}+XG_{5XX}\right)\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b\left(\phi^{|0}r^{a_1}_{b_1}R^{a_2a_3}_{b_2b_3} + 4\phi^{|a_1}_{|b_1}\Gamma^{0a_2}_{\ \ b_2}r^{a_3}_{b_3}\right) \\ &\quad + \frac{\phi^{|0}}{2}\left(2G_{5XX}+XG_{5XXX}\right)\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}r^{a_3}_{b_3}. \end{aligned} \tag{C10}$$

Last, the pure scalar terms proportional to ${r_\phi}^3$ are given by

$$\frac{\partial}{\partial \phi_{,0}}\frac{\partial \mathcal{E}^2_\phi}{\partial \phi_{,00}} = 3\phi^{|0}K_{XX}\xi^2 - (\phi^{|0})^3K_{XXX} \tag{C11}$$

$$\begin{aligned} \frac{\partial}{\partial \phi_{,0}}\frac{\partial \mathcal{E}^3_\phi}{\partial \phi_{,00}} &= -\phi^{|0}\left\{2\left(2G_{3XX}+XG_{3XXX}\right)\delta^{bb_1}_{aa_1}\xi^a\xi_b\phi^{|a_1}_{|b_1} + G_{3XXX}\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b\phi^{|a_1}_{|b_1}\phi^{|a_2}\phi_{|b_2}\right\} \\ &\quad - 2\left(G_{3X}+XG_{3XX}\right)\delta^{bb_1}_{aa_1}\xi^a\xi_b\Gamma^{0a_1}_{\ \ b_1} - G_{3XX}\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b\Gamma^{0a_1}_{\ \ b_1}\phi^{|a_2}\phi_{|b_2} \end{aligned} \tag{C12}$$

$$\begin{aligned} \frac{\partial}{\partial \phi_{,0}}\frac{\partial \mathcal{E}^4_\phi}{\partial \phi_{,00}} &= \frac{\phi^{|0}}{2}\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b\left\{2\left(5G_{4XXX}+2XG_{4XXXX}\right)\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2} + \left(3G_{4XX}+2XG_{4XXX}\right)R^{a_1a_2}_{b_1b_2}\right\} \\ &\quad + \frac{\phi^{|0}}{2}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b\phi^{|a_1}\phi_{|b_1}\left(2G_{4XXXX}\phi^{|a_2}_{|b_2}\phi^{|a_3}_{|b_3} + G_{4XXX}R^{a_2a_3}_{b_2b_3}\right) \\ &\quad + 2\left(3G_{4XX}+2XG_{4XXX}\right)\delta^{bb_1b_2}_{aa_1a_2}\xi^a\xi_b\phi^{|a_1}_{|b_1}\Gamma^{0a_2}_{\ \ b_2} + 2G_{4XXX}\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b\Gamma^{0a_1}_{\ \ b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}\phi_{|b_3} \end{aligned} \tag{C13}$$

$$\begin{aligned} \frac{\partial}{\partial \phi_{,0}}\frac{\partial \mathcal{E}^5_\phi}{\partial \phi_{,00}} &= -\delta^{bb_1b_2b_3}_{aa_1a_2a_3}\xi^a\xi_b\bigg\{\left(2G_{5XX}+XG_{5XXX}\right)\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\Gamma^{0a_3}_{\ \ b_3} + \frac{1}{2}\left(G_{5X}+XG_{5XX}\right)\Gamma^{0a_1}_{\ \ b_1}R^{a_2a_3}_{b_2b_3} \\ &\quad + \frac{\phi^{|0}}{3}\left(3G_{5XXX}+XG_{5XXXX}\right)\phi^{|a_1}_{|b_1}\phi^{|a_2}_{|b_2}\phi^{|a_3}_{|b_3} + \frac{\phi^{|0}}{2}\left(2G_{5XX}+XG_{5XXX}\right)\phi^{|a_1}_{|b_1}R^{a_2a_3}_{b_2b_3}\bigg\}. \end{aligned} \tag{C14}$$

Then the terms appearing in Eq. (100) are obtained by summing the above expressions over $n = 2,3,4,5$.

2. $\mathcal{N}$ on the plane wave solution

In this appendix, we show the explicit form of $\mathcal{N}$ on the plane wave solution, which can be obtained by plugging the background solution shown in section III B 2 into expressions in appendix C 1. We also need to use the expressions of the eigenvectors (83), (84) of the tensor and scalar modes to find $\mathcal{N}$ for each mode.

For the tensor mode $r = (r_{ab}^{(\mathrm{T})}, 0)$, the only nontrivial term is (C3), but this term is zero for the plane wave solution since $X = 0$. Hence $\mathcal{N} = 0$ for the tensor mode on the plane wave solution.

Let us move on to the scalar mode (84), whose eigenvector is given by

$$r = (r_{ab}, r_\phi) = \left(2\tilde{r}_{\ell n}\ell_{(a}n_{b)}, r_\phi\right), \quad \tilde{r}_{\ell n} = \frac{2K_X}{G_4}\,\frac{-\frac{1}{2}G_{3X}\phi'^2 + G_{4X}\phi''}{K_{XX}\phi'^2 - 2G_{3X}\phi'' - G_{4X}\Delta F}\, r_\phi. \tag{C15}$$

When this eigenvector is contracted with the generalized Kronecker delta multiplied by $\xi_a\xi^b$, it simplifies to (by dropping the ξ terms)

$$r_a^b = r_{\ell n}\left(\ell_a n^b + \ell^b n_a\right) \simeq \omega r_{\ell n}\ell_a\ell^b. \tag{C16}$$

Below, we evaluate the terms appearing in $\mathcal{N}$, whose general expression is summarized in appendix C 1, for the plane wave solution using the eigenvector given by (C15). It is useful to use the following formula for the evaluation:

$$\ell_a = (du)_a, \quad \ell^a = (\partial_v)^a = \left(\ell^u, \ell^0\right) = (0, 1), \quad \xi_a = \left(dx^0\right)_a, \quad \xi^a = (\xi^u, \xi^0) = (1, -\omega), \tag{C17}$$

$$\ell\cdot\xi = g^{0u} = 1, \quad \xi^2 = g^{00} = -\omega, \quad n_a = \frac{\omega}{2}\ell_a + \xi_a, \quad n\cdot\xi = -\frac{\omega}{2}, \quad \phi^{|0} = \xi_a\phi^{|a} = \phi'\xi\cdot\ell = \phi'. \tag{C18}$$

First, we find that the pure metric terms and also the mixed terms proportional to ${r_{ab}}^2 r_\phi$ identically vanish:

$$r^{ab}r_{cd}r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial \mathcal{G}_{ab}^{2,3,4,5}}{\partial g_{cd,00}} = r^{ab}r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial \mathcal{G}_{ab}^{2,3,4,5}}{\partial \phi_{,00}} = 0. \tag{C19}$$

The other mixed terms proportional to $r_{ab}{r_\phi}^2$ are given by

$$r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial \mathcal{E}_\phi^{2,5}}{\partial \phi_{,00}} = 0, \quad r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial \mathcal{E}_\phi^{3}}{\partial \phi_{,00}} = \phi' G_{3X}\delta^{bb_1}_{aa_1}\xi^a\xi_b r^{a_1}_{b_1}, \quad r_{ef}\frac{\partial}{\partial g_{ef,0}}\frac{\partial \mathcal{E}_\phi^{4}}{\partial \phi_{,00}} = -G_{4X}\delta^{b\beta_1 b_2}_{a\alpha_1 a_2}\xi^a\xi_b\Gamma^{0\alpha_1}_{\ \ \beta_1}r^{a_2}_{b_2}, \tag{C20}$$

where $x^\alpha = x, y$ as we defined in section IV C 1. Last, the pure scalar terms are given by

$$\frac{\partial}{\partial\phi_{,0}}\frac{\partial \mathcal{E}_\phi^2}{\partial\phi_{,00}} = -3\omega\phi' K_{XX} - \phi'^3 K_{XXX} \tag{C21}$$

$$\frac{\partial}{\partial\phi_{,0}}\frac{\partial \mathcal{E}_\phi^3}{\partial\phi_{,00}} = -4\phi'\phi'' G_{3XX}\delta^{bb_1}_{aa_1}\xi^a\xi_b\ell^{a_1}\ell_{b_1} - 2G_{3X}\delta^{b\beta_1}_{a\alpha_1}\xi^a\xi_b\Gamma^{0\alpha_1}_{\ \ \beta_1} - \phi'^2 G_{3XX}\delta^{b\beta_1 b_2}_{a\alpha_1 a_2}\xi^a\xi_b\Gamma^{0\alpha_1}_{\ \ \beta_1}\ell^{a_2}\ell_{b_2} \tag{C22}$$

$$\frac{\partial}{\partial\phi_{,0}}\frac{\partial \mathcal{E}_\phi^4}{\partial\phi_{,00}} = -3\phi' G_{4XX}\delta^{b\beta_1 b_2}_{a\alpha_1 a_2}\xi^a\xi_b F^{,\alpha_1}_{,\beta_1}\ell^{a_2}\ell_{b_2} + 6G_{4XX}\phi''\delta^{bb_1\beta_2}_{aa_1\alpha_2}\xi^a\xi_b\ell^{a_1}\ell_{b_1}\Gamma^{0\alpha_2}_{\ \ \beta_2} \tag{C23}$$

$$\frac{\partial}{\partial\phi_{,0}}\frac{\partial \mathcal{E}_\phi^5}{\partial\phi_{,00}} = G_{5X}\delta^{b\beta_1\beta_2 b_3}_{a\alpha_1\alpha_2 a_3}\xi^a\xi_b\Gamma^{0\alpha_1}_{\ \ \beta_1}F^{,\alpha_2}_{,\beta_2}\ell^{a_3}\ell_{b_3}. \tag{C24}$$

Summing up the above terms, we find $\mathcal{N}$ for the scalar mode on the plane wave solution to be given by

$$\begin{aligned}\mathcal{N} = {}& 3\omega{r_\phi}^2\tilde{r}_{\ell n}\left(\phi' G_{3X}\delta^{bb_1}_{aa_1}\xi^a\xi_b\ell^{a_1}\ell_{b_1} - G_{4X}\delta^{b\beta_1 b_2}_{a\alpha_1 a_2}\xi^a\xi_b\Gamma^{0\alpha_1}_{\ \ \beta_1}\ell^{a_2}\ell_{b_2}\right)\\ &+ {r_\phi}^3\Big\{-3\omega K_{XX}\phi' - K_{XXX}\phi'^3 - 4\phi'\phi'' G_{3XX}\delta^{bb_1}_{aa_1}\xi^a\xi_b\ell^{a_1}\ell_{b_1}\\ &\quad - 2G_{3X}\delta^{b\beta_1}_{a\alpha_1}\xi^a\xi_b\Gamma^{0\alpha_1}_{\ \ \beta_1} + \left(-\phi'^2 G_{3X} + 6\phi'' G_{4XX}\right)\delta^{b\beta_1 b_2}_{a\alpha_1 a_2}\xi^a\xi_b\Gamma^{0\alpha_1}_{\ \ \beta_1}\ell^{a_2}\ell_{b_2}\\ &\quad - 3\phi' G_{4XX}\delta^{b\beta_1 b_2}_{a\alpha_1 a_2}\xi^a\xi_b F^{,\alpha_1}_{,\beta_1}\ell^{a_2}\ell_{b_2} + G_{5X}\delta^{b\beta_1 b_2\beta_3}_{a\alpha_1 a_2\alpha_3}\xi^a\xi_b\Gamma^{0\alpha_1}_{\ \ \beta_1}\ell^{a_2}\ell_{b_2}F^{,\alpha_3}_{,\beta_3}\Big\}.\end{aligned} \tag{C25}$$

If we further assume $\phi'' = 0$ and F to be given by Eq. (109), we can introduce the metric (113) adapted to geodesics, with which we can evaluate Eq. (C25) more explicitly to arrive at Eq. (116) by setting $r_\phi = 1$.

3. Details of shock formation analysis on the plane wave background

In this appendix, we show the details of the shock formation analysis on the plane wave background discussed in section IV C 1. For this purpose, we need to derive the full expression of the transport equation (96). In principle, we can derive it based on the definition (91)–(94), but it is cumbersome. Since we have already found geodesics and the coordinates adapted to it, as summarized in section IV C 1, there is an alternative method to derive $\mathcal{N}$ and also the other parts of the transport equation. The key is to assume that the field variables are given by

$$g_{ab} = \bar{g}_{ab} + \frac{1}{2}\,(x^0)^2\,\Theta(x^0)\,\Pi\big(u,\eta^\alpha\big)\,r_{ab}\,, \qquad \phi = \bar{\phi} + \frac{1}{2}\,(x^0)^2\,\Theta(x^0)\,\Pi\big(u,\eta^\alpha\big)\,r_\phi\,, \tag{C26}$$

where $\bar{g}_{ab}, \bar{\phi}$ are the background solutions and $\Theta(x^0)$ is a step function. The above g_{ab}, ϕ correctly give the discontinuities in their second derivatives at $x^0 = 0$ as prescribed by Eq. (90), although they do not satisfy the equations of motion in $x_0 > 0$. However, the quantities appearing in the transport equation depend only on the values of g_{ab}, ϕ for $x_0 \to -0$, hence the transport equation (91) is correctly obtained by evaluating the equation of motion at $x^0 = 0$ using Eq. (C26) and contracting it with r. See [34] for more details on the construction of the transport equation by this method.

With the aid of computer algebra, we can follow the above procedure to find the transport equation of $\Pi(u,\eta^\alpha)$ as

$$\mathcal{K}\,\Pi_{,u} + \mathcal{M}\,\Pi + \mathcal{N}\,\Pi^2 = 0, \tag{C27}$$

where $\mathcal{N}$ is given by Eq. (116), and the other coefficients are given by

$$\mathcal{K} = -2K_X, \qquad \mathcal{M} = -K_X t_+(u). \tag{C28}$$

Then, the transport equation takes the form of Eq. (96) once the parameter s along the characteristic curve is introduced following (95) as

$$s = -\frac{1}{2K_X}\,u\,. \tag{C29}$$

We assume $A < 0$, as we defined below Eq. (110). It implies that $A_x < 0$ and $|A_x| > |A_y|$. Then Eq. (117) becomes

$$t_\pm = \sqrt{-A_x}\tan\left(\frac{\pi}{2}\frac{s}{s_*}\right) \mp \sqrt{A_y}\tanh\left(\frac{\pi}{2}\frac{s}{s_*}\right), \tag{C30}$$

where $s_* \equiv \frac{\pi}{4K_X\sqrt{-A_x}}$, and also $e^{-\Phi}$ appearing in the general solution of $\Pi(s)$ (97) is calculated as

$$e^{-\Phi} = \left\{\cos\left(\frac{\pi}{2}\frac{s}{s_*}\right)\cosh\left(2K_X\sqrt{A_y}s\right)\right\}^{-1/2}, \tag{C31}$$

which diverges for $s \to s_*-$ like $\propto (s_* - s)^{-1}$. Also, $\mathcal{N}e^{-\Phi}$ behaves for $s \to s_*-$ as

$$\mathcal{N}e^{-\Phi} \simeq \frac{(C_+ + C_-)\sqrt{-A_x}}{\cosh^{1/2}\left(\frac{\pi}{2}\sqrt{\frac{A_y}{-A_x}}\right)}\left(\frac{s_* - s}{s_*}\right)^{-3/2} \equiv \mathcal{N}_0\left(\frac{s_* - s}{s_*}\right)^{-3/2}. \tag{C32}$$

This quantity diverges for $s \to s_*-$, hence the denominator of $\Pi(s)$ given by Eq. (97) can be zero at finite $s = s_0$ (such that $0 < s_0 < s_*$) no matter how small $|\Pi(0)|$ is, only if the sign of $\Pi(0)$ is appropriately chosen. Hence, $\Pi(s)$ can diverge and the shock formation occurs at $s = s_0$ for an arbitrarily small $\Pi(0)$ in this example.

Assuming that $\mathcal{N}e^{-\Phi}$ is well approximated by Eq. (C32), s_0 may be estimated as follows. Using Eq. (C32) the integral in the denominator of (97) is estimated as

$$\int_0^s \mathcal{N}(s')e^{-\Phi(s')}ds' \simeq 2s_*\mathcal{N}_0 \left(\frac{s_* - s}{s_*}\right)^{-1/2}. \tag{C33}$$

If this approximation is valid, s_0 will be approximated by

$$s_0 \simeq \left\{1 - (2\Pi(0)s_*\mathcal{N}_0)^2\right\} s_*. \tag{C34}$$

When $C_+ = C_- = 0$ but $C_0 \neq 0$, the denominator of (97) does not vanish for $0 < s < s_*$ if the initial amplitude $\Pi(0)$ is made sufficiently small. Even in this case, if $\Pi(0)$ is sufficiently large so that $\Pi(0)C_0 s = -1$ at $s < s_*$, the denominator vanishes at $s < s_*$ hence shock formation occur at that s.

In this example, $\Pi(s)$ diverges at $s = s_*$ even when $\mathcal{N}$ happens to vanish, because $\mathcal{N} = 0$ implies that $\Pi(s) = \Pi(0)e^{-\Phi}$ and then it diverges due to $e^{-\Phi} \propto (s * -s)^{-1/2} \to \infty$ for $s \to s_*-$. This divergence occurs at the caustics of geodesics on Σ due to a focusing effect, and sometime called a linear shock (see footnote 8).

4. Formulae for the two-dimensionally maximally-symmetric dynamical solutions

Section IV C 2 is dealing with the dynamical background solutions that is symmetric in the two-dimensional angular directions. Based on the ansatz (121), we can derive the curvature tensor as (122), where the two-dimensional tensor $R_2{}^A_B$ in it is given by

$$\left(R_2{}^A_B\right) = \frac{1}{2f^2\rho}\begin{pmatrix} -\left(f_{,\tau}\rho_{,\tau} + f_{,\chi}\rho_{,\chi} - 2f\rho_{,\tau\tau}\right) & -2\left(f_{,(\tau}\rho_{,\chi)} - f\rho_{,\tau\chi}\right) \\ 2\left(f_{,(\tau}\rho_{,\chi)} - f\rho_{,\tau\chi}\right) & f_{,\tau}\rho_{,\tau} + f_{,\chi}\rho_{,\chi} - 2f\rho_{,\chi\chi} \end{pmatrix}. \tag{C35}$$

Also, the nonzero components of $\phi_{|ab}$ are given by

$$\left(\phi_{|AB}\right) = -\frac{1}{2f}\begin{pmatrix} f_{,\tau}\phi_{,\tau} + f_{,\chi}\phi_{,\chi} - 2f\phi_{,\tau\tau} & 2\left(f_{,(\tau}\phi_{,\chi)} - f\phi_{,\tau\chi}\right) \\ 2\left(f_{,(\tau}\phi_{,\chi)} - f\phi_{,\tau\chi}\right) & f_{,\tau}\phi_{,\tau} + f_{,\chi}\phi_{,\chi} - 2f\phi_{,\chi\chi} \end{pmatrix}, \quad \phi_{|\alpha\beta} = \frac{\rho}{f}\left(-\rho_{,\tau}\phi_{,\tau} + \rho_{,\chi}\phi_{,\chi}\right)\gamma_{\alpha\beta}. \tag{C36}$$

[1] A. A. Starobinsky, Phys. Lett. **B91**, 99 (1980); K. Sato, Mon. Not. Roy. Astron. Soc. **195**, 467 (1981); A. H. Guth, Phys. Rev. **D23**, 347 (1981).

[2] B. P. Schmidt *et al.* (Supernova Search Team), Astrophys. J. **507**, 46 (1998), arXiv:astro-ph/9805200 [astro-ph]; A. G. Riess *et al.* (Supernova Search Team), Astron. J. **116**, 1009 (1998), arXiv:astro-ph/9805201 [astro-ph]; S. Perlmutter *et al.* (Supernova Cosmology Project), Astrophys. J. **517**, 565 (1999), arXiv:astro-ph/9812133 [astro-ph].

[3] C. A. R. Herdeiro and E. Radu, *Proceedings, 7th Black Holes Workshop 2014: Aveiro, Portugal, December 18-19, 2014*, Int. J. Mod. Phys. **D24**, 1542014 (2015), arXiv:1504.08209 [gr-qc].

[4] H. Yoshino and H. Kodama, Class. Quant. Grav. **32**, 214001 (2015), arXiv:1505.00714 [gr-qc].

[5] B. P. Abbott *et al.* (Virgo, LIGO Scientific), Phys. Rev. Lett. **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].

[6] J. M. Maldacena, Int. J. Theor. Phys. **38**, 1113 (1999), [Adv. Theor. Math. Phys.2,231(1998)], arXiv:hep-th/9711200 [hep-th].

[7] S. A. Hartnoll, C. P. Herzog, and G. T. Horowitz, Phys. Rev. Lett. **101**, 031601 (2008), arXiv:0803.3295 [hep-th].

[8] V. E. Hubeny and M. Rangamani, Adv. High Energy Phys. **2010**, 297916 (2010), arXiv:1006.3675 [hep-th].

[9] P. Bizon and A. Rostworowski, Phys. Rev. Lett. **107**, 031102 (2011), arXiv:1104.3702 [gr-qc].

[10] V. E. Hubeny, Class. Quant. Grav. **32**, 124010 (2015), arXiv:1501.00007 [gr-qc].

[11] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, Phys. Rept. **513**, 1 (2012), arXiv:1106.2476 [astro-ph.CO].

[12] A. Nicolis, R. Rattazzi, and E. Trincherini, Phys. Rev. **D79**, 064036 (2009), arXiv:0811.2197 [hep-th].

[13] M. Ostrogradsky, Mém. de l'acad. de St. Pétersbourg **VI**, 385 (1850).

[14] R. P. Woodard, Scholarpedia **10**, 32243 (2015), arXiv:1506.02210 [hep-th].

[15] C. Deffayet, G. Esposito-Farese, and A. Vikman, Phys. Rev. **D79**, 084003 (2009), arXiv:0901.1314 [hep-th].

[16] C. Deffayet, X. Gao, D. A. Steer, and G. Zahariade, Phys. Rev. **D84**, 064039 (2011), arXiv:1103.3260 [hep-th].

[17] T. Kobayashi, M. Yamaguchi, and J. Yokoyama, Prog. Theor. Phys. **126**, 511 (2011), arXiv:1105.5723 [hep-th].

[18] G. W. Horndeski, Int. J. Theor. Phys. **10**, 363 (1974).

[19] J.-O. Gong, Int. J. Mod. Phys. **D26**, 1740003 (2016), arXiv:1606.06971 [gr-qc].

[20] A. Padilla, P. M. Saffin, and S.-Y. Zhou, JHEP **12**, 031 (2010), arXiv:1007.5424 [hep-th]; JHEP **01**, 099 (2011), arXiv:1008.3312 [hep-th]; K. Hinterbichler, M. Trodden, and D. Wesley, Phys. Rev. **D82**, 124018 (2010), arXiv:1008.1305 [hep-th]; A. Padilla, P. M. Saffin, and S.-Y. Zhou, Phys. Rev. **D83**, 045009 (2011), arXiv:1008.0745 [hep-th]; A. Padilla and V. Sivanesan, JHEP **04**, 032 (2013), arXiv:1210.4026 [gr-qc].

[21] D. A. Easson, R. Gregory, D. F. Mota, G. Tasinato, and I. Zavala, JCAP **0802**, 010 (2008), arXiv:0709.2666 [hep-th].

[22] M.-x. Huang, G. Shiu, and B. Underwood, Phys. Rev. **D77**, 023511 (2008), arXiv:0709.3299 [hep-th].

[23] D. Langlois and S. Renaux-Petel, JCAP **0804**, 017 (2008), arXiv:0801.1085 [hep-th].

[24] D. Langlois, S. Renaux-Petel, D. A. Steer, and T. Tanaka, Phys. Rev. Lett. **101**, 061301 (2008), arXiv:0804.3139 [hep-th].

[25] D. Langlois, S. Renaux-Petel, D. A. Steer, and T. Tanaka, Phys. Rev. **D78**, 063523 (2008), arXiv:0806.0336 [hep-th].

[26] D. Langlois, S. Renaux-Petel, and D. A. Steer, JCAP **0904**, 021 (2009), arXiv:0902.2941 [hep-th].

[27] S. Renaux-Petel, JCAP **0910**, 012 (2009), arXiv:0907.2476 [hep-th].

[28] K. Koyama, G. W. Pettinari, S. Mizuno, and C. Fidler, Class. Quant. Grav. **31**, 125003 (2014), arXiv:1303.2125 [astro-ph.CO].

[29] T. Kobayashi, N. Tanahashi, and M. Yamaguchi, Phys. Rev. **D88**, 083504 (2013), arXiv:1308.4798 [hep-th].

[30] S. Ohashi, N. Tanahashi, T. Kobayashi, and M. Yamaguchi, JHEP **07**, 008 (2015), arXiv:1505.06029 [gr-qc].

[31] A. Majda, *Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables*, Applied Mathematical Sciences (Springer New York, 1984).

[32] F. John and L. University, *Nonlinear Wave Equations, Formation of Singularities*, Dimacs Series in Discrete Mathematics and Theoretical Comput (American Mathematical Soc.).

[33] D. Christodoulou, *The Formation of Shocks in 3-dimensional Fluids*, EMS monographs in mathematics (European Mathematical Society, 2007).

[34] H. S. Reall, N. Tanahashi, and B. Way, Phys. Rev. **D91**, 044013 (2015), arXiv:1409.3874 [hep-th].

[35] E. Babichev, JHEP **04**, 129 (2016), arXiv:1602.00735 [hep-th].

[36] S. Mukohyama, R. Namba, and Y. Watanabe, Phys. Rev. **D94**, 023514 (2016), arXiv:1605.06418 [hep-th].

[37] C. de Rham and H. Motohashi, Phys. Rev. **D95**, 064008 (2017), arXiv:1611.05038 [hep-th].

[38] G. N. Felder, L. Kofman, and A. Starobinsky, JHEP **09**, 026 (2002), arXiv:hep-th/0208019 [hep-th]; N. Barnaby, JHEP **07**, 025 (2004), arXiv:hep-th/0406120 [hep-th]; U. D. Goswami, H. Nandan, and M. Sami, Phys.

Rev. **D82**, 103530 (2010), arXiv:1006.3659 [hep-th].

[39] J. D. Bekenstein, Phys. Rev. **D70**, 083509 (2004), [Erratum: Phys. Rev.D71,069901(2005)], arXiv:astro-ph/0403694 [astro-ph]; C. R. Contaldi, T. Wiseman, and B. Withers, Phys. Rev. **D78**, 044034 (2008), arXiv:0802.1215 [gr-qc]; S. Mukohyama, JCAP **0909**, 005 (2009), arXiv:0906.5069 [hep-th]; M. R. Setare and D. Momeni, Int. J. Theor. Phys. **51**, 198 (2012), arXiv:1009.0918 [hep-th].

[40] R. Courant and D. Hilbert, Methods of mathematical physics, Vol. 2 (Wiley, 1962).

[41] D. Lovelock, J. Math. Phys. **12**, 498 (1971).

[42] K. Izumi, Phys. Rev. **D90**, 044037 (2014), arXiv:1406.0677 [gr-qc].

[43] H. Reall, N. Tanahashi, and B. Way, Class. Quant. Grav. **31**, 205005 (2014), arXiv:1406.3379 [hep-th].

[44] M. Minamitsuji, Phys. Lett. **B743**, 272 (2015).

[45] E. Babichev, Phys. Rev. **D86**, 084037 (2012), arXiv:1207.4764 [gr-qc].

[46] A. Anile, *Relativistic Fluids and Magneto-fluids: With Applications in Astrophysics and Plasma Physics*, Cambridge Monographs on Mathem (Cambridge University Press, 2005); P. D. Lax, Ann. Math. Stud. **33**, 211 (1954); Comm. Pure Appl. Math. **10**, 537 (1957); A. Jeffrey, *London, Pitman Publishing, Ltd. (Research Notes in Mathematics, No. 5), 1976. 230 p.* (1976); G. Boillat, Journal of Mathematical Physics **10**, 452 (1969), http://dx.doi.org/10.1063/1.1664860.

[47] C. Armendariz-Picon, T. Damour, and V. F. Mukhanov, Phys. Lett. **B458**, 209 (1999), arXiv:hep-th/9904075 [hep-th]; C. Armendariz-Picon, V. F. Mukhanov, and P. J. Steinhardt, Phys. Rev. **D63**, 103510 (2001), arXiv:astro-ph/0006373 [astro-ph].

[48] M. Born, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **143**, 410 (1934); M. Born and L. Infeld, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **144**, 425 (1934); P. A. M. Dirac, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **268**, 57 (1962).

[49] G. Boillat, T. Ruggeri, C. Dafermos, P. Lax, and T. Liu, *Recent Mathematical Methods in Nonlinear Wave Propagation: Lectures given at the 1st Session of the Centro Internazionale Matematico Estivo (C.I.M.E.), held in Montecatini Terme, Italy, May 23-31, 1994*, Lecture Notes in Mathematics (Springer Berlin Heidelberg, 2006); T. Taniuti, Progress of Theoretical Physics Supplement **9**, 69 (1959); G. B. Whitham, "Exact solutions; interacting solitary waves," in *Linear and Nonlinear Waves* (John Wiley & Sons, Inc., 1999) pp. 577–620; B. Barbashov and N. Chernikov, Sov. Phys. JETP **23**, 861 (1966); G. Boillat and A. Muracchini, "Isocline curves and variational scalar field," in *Asymptotic Methods in Nonlinear Wave Phenomena* (WORLD SCIENTIFIC, 2012) pp. 24–34; S. Deser, J. G. McCarthy, and O. Sarioglu, Class. Quant. Grav. **16**, 841 (1999), arXiv:hep-th/9809153 [hep-th].

[50] H. Rund, in *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, Vol. 29 (Springer, 1966) pp. 243–262.

[51] J. C. Du Plessis, Tensor **20**, 347 (1969).

[52] I. Smolić, Class. Quant. Grav. **32**, 145010 (2015), arXiv:1501.04967 [gr-qc].

[53] A. A. H. Graham and R. Jha, Phys. Rev. **D90**, 041501 (2014), arXiv:1407.6573 [gr-qc].

[54] N. Tanahashi and S. Ohashi, (2017), arXiv:1704.02757 [hep-th].

[55] Y. Choquet-Bruhat and Chrusciel, J. Maths. Pures et App. **48**, 117 (1969).

[56] J. K. Hunter and J. B. Keller, Commun. Pure Appl. Math. **36**, 547 (1983).

[57] Y. Choquet-Bruhat, *General Relativity and the Einstein Equations*, Oxford Mathematical Monographs (OUP Oxford, 2009).

[58] G. Boillat, Journal of Mathematical Physics **11**, 941 (1970), http://dx.doi.org/10.1063/1.1665231.

[59] D. Bettoni, J. M. Ezquiaga, K. Hinterbichler, and M. Zumalacárregui, Phys. Rev. **D95**, 084029 (2017), arXiv:1608.01982 [gr-qc].

[60] C. Deffayet, A. E. Gümrükçüoğlu, S. Mukohyama, and Y. Wang, JHEP **04**, 082 (2014), arXiv:1312.6690 [hep-th].

[61] A. Padilla, D. Stefanyszyn, and M. Tsoukalas, Phys. Rev. **D89**, 065009 (2014), arXiv:1312.0975 [hep-th].

[62] E. Allys, (2016), arXiv:1612.01972 [hep-th].
[63] S. Akama and T. Kobayashi, Phys. Rev. **D95**, 064011 (2017), arXiv:1701.02926 [hep-th].
[64] R. Kimura and K. Yamamoto, JCAP **1207**, 050 (2012), arXiv:1112.4284 [astro-ph.CO].
[65] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis, and R. Rattazzi, JHEP **10**, 014 (2006), arXiv:hep-th/0602178 [hep-th].
[66] R. Geroch, AMS/IP Stud. Adv. Math. **49**, 59 (2011), arXiv:1005.1614 [gr-qc].
[67] G. Papallo and H. S. Reall, JHEP **11**, 109 (2015), arXiv:1508.05303 [gr-qc].
[68] E. Barausse, T. Jacobson, and T. P. Sotiriou, Phys. Rev. **D83**, 124043 (2011), arXiv:1104.2889 [gr-qc].
[69] M. Zumalacarregui and J. Garcia-Bellido, Phys. Rev. **D89**, 064046 (2014), arXiv:1308.4685 [gr-qc].
[70] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, Phys. Rev. Lett. **114**, 211101 (2015), arXiv:1404.6495 [hep-th].
[71] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, JCAP **1502**, 018 (2015), arXiv:1408.1952 [astro-ph.CO].
[72] X. Gao, Phys. Rev. **D90**, 081501 (2014), arXiv:1406.0822 [gr-qc].
[73] C. Deffayet, G. Esposito-Farese, and D. A. Steer, Phys. Rev. **D92**, 084013 (2015), arXiv:1506.01974 [gr-qc].
[74] G. Domenech, S. Mukohyama, R. Namba, A. Naruko, R. Saitou, and Y. Watanabe, Phys. Rev. **D92**, 084027 (2015), arXiv:1507.05390 [hep-th].
[75] D. Langlois and K. Noui, JCAP **1602**, 034 (2016), arXiv:1510.06930 [gr-qc].
[76] D. Langlois and K. Noui, JCAP **1607**, 016 (2016), arXiv:1512.06820 [gr-qc].
[77] M. Crisostomi, K. Koyama, and G. Tasinato, JCAP **1604**, 044 (2016), arXiv:1602.03119 [hep-th].
[78] J. Ben Achour, D. Langlois, and K. Noui, Phys. Rev. **D93**, 124005 (2016), arXiv:1602.08398 [gr-qc].
[79] J. M. Ezquiaga, J. Garcia-Bellido, and M. Zumalacarregui, Phys. Rev. **D94**, 024005 (2016), arXiv:1603.01269 [hep-th].
[80] H. Motohashi, K. Noui, T. Suyama, M. Yamaguchi, and D. Langlois, JCAP **1607**, 033 (2016), arXiv:1603.09355 [hep-th].
[81] C. de Rham and A. Matas, JCAP **1606**, 041 (2016), arXiv:1604.08638 [hep-th].
[82] J. Ben Achour, M. Crisostomi, K. Koyama, D. Langlois, K. Noui, and G. Tasinato, JHEP **12**, 100 (2016), arXiv:1608.08135 [hep-th].
[83] D. Langlois, M. Mancarella, K. Noui, and F. Vernizzi, (2017), arXiv:1703.03797 [hep-th].
[84] L. Heisenberg, R. Kase, and S. Tsujikawa, Phys. Lett. **B760**, 617 (2016), arXiv:1605.05565 [hep-th].
[85] A. Naruko, D. Yoshida, and S. Mukohyama, Class. Quant. Grav. **33**, 09LT01 (2016), arXiv:1512.06977 [gr-qc].
[86] K. Izumi and Y. C. Ong, Class. Quant. Grav. **30**, 184008 (2013), arXiv:1304.0211 [hep-th].
[87] A. Padilla and V. Sivanesan, JHEP **08**, 122 (2012), arXiv:1206.1258 [gr-qc].
[88] S. Nishi, T. Kobayashi, N. Tanahashi, and M. Yamaguchi, JCAP **1403**, 008 (2014), arXiv:1401.1045 [hep-th].