INVERSE PROBLEM FOR EINSTEIN-SCALAR FIELD EQUATIONS

YAROSLAV KURYLEV, MATTI LASSAS, LAURI OKSANEN, AND GUNTHER UHLMANN

ABSTRACT. The paper introduces a method to solve inverse problems for hyperbolic systems where the leading order terms are non-linear. We apply the method to the coupled Einstein-scalar field equations and study the question whether the structure of spacetime can be determined by making active measurements near the world line of an observer. We show that such measurements determine the topological, differential and conformal structure of the spacetime in the optimal chronological diamond type set containing the world line. In the case when the unknown part of the spacetime is vacuum, we can also determine the metric itself. We exploit the non-linearity of the equation to obtain a rich set of propagating singularities, produced by a non-linear interaction of singularities that propagate initially as for linear wave equations. This non-linear effect is then used as a tool to solve the inverse problem for the non-linear system. The method works even in cases where the corresponding inverse problems for linear equations remain open, and it can potentially be applied to a large class of inverse problems for non-linear hyperbolic equations encountered in practical imaging problems.

CONTENTS

1. Introduction and main results

In this paper we introduce a new method to solve inverse problems for hyperbolic systems where the leading order terms depend non-linearly on the solution. Earlier, we have considered a related inverse problem for scalar wave equations where the leading order terms are given by the linear wave operator with a fixed metric [49]. Non-linear hyperbolic systems are used in several applications, but the models of the non-linear terms are often quite heuristic. In order to have a canonical non-linearity, we consider a measurement model for Einstein's equations coupled with scalar fields.

In this context several difficulties appear. In particular, detection of singularities of waves becomes non-trivial: A well-known physical example of this is that only two polarisations of gravitational waves can be observed even though the solutions to Einstein's equations take values in the space of symmetric 4×4 matrices that corresponds to a 10-dimensional space of polarisations (see Section 2 for further discussion). Moreover, the definition of measurements becomes non-trivial as any source changes the space-time itself in the future of its support.

The method we develop utilises the non-linearity as a tool and it enables us to solve inverse problems for non-linear equations even in cases where the inverse problem for the corresponding linear system remains open. Indeed, the existing uniqueness results for linear hyperbolic equations require restrictive geometric assumptions such as stationarity, whereas our result is applicable to any globally hyperbolic spacetime. We review the existing literature on inverse problems in detail in Section 1.2.

Outside the context of inverse problems, there are many results on non-linear interaction of waves, starting from the studies of Bony [11], Melrose and Ritter [60, 64] and Rauch and Reed, [73]. However, these studies are different from the present paper since they assume that the geometrical setting, in particular the

locations and types of caustics, is a priori known, whereas in our inverse problem wave propagation is studied on an unknown manifold and we allow for caustics of arbitrary, including unstable, type. This causes further difficulties in the analysis of the non-linear interaction.

In physical terms, we study the question: Can an observer determine the structure of the surrounding spacetime by doing measurements near its world line? The conservation law for Einstein's equations dictates, roughly speaking, that any source in the equation must take energy from some fields in order to increase energy in other fields. The scalar fields are included in the model that we consider to facilitate this. They correspond to spin zero particles.

Let us now give the mathematical formulation of the problem. Let M be a smooth $1 + 3$-dimensional manifold. Einstein's equations for a Lorentzian metric g on M are

$$\mathrm{Ein}(g) = T,$$

where $\mathrm{Ein}(g)$ is the Einstein tensor associated to g and T is a stress-energy tensor. In vacuum $T = 0$. The measurement model we consider is

$$(1) \qquad \mathrm{Ein}(g) = T, \quad T = \mathbb{T}(g, \phi) + \mathcal{F}^1,$$

$$(2) \qquad \Box_g \phi_l - \mathcal{V}'_l(\phi) = \mathcal{F}^2_l, \quad l = 1, 2, ..., L.$$

Here $\mathcal{V}'_l(\phi) = \partial_{\phi_l} \mathcal{V}(\phi)$, $\phi = (\phi_l)$, $l = 1, 2, \ldots, L$, are the scalar fields, and

$$\mathcal{F} = (\mathcal{F}^1, \mathcal{F}^2) = (\mathcal{F}^1, \mathcal{F}^2_1, \ldots, \mathcal{F}^2_L),$$

models a source in the measurements. The physical interpretation of $\mathcal{F}$ is discussed in Appendix C. The coupling $\mathbb{T} = (\mathbb{T}_{jk})$ is given by

$$\mathbb{T}_{jk}(g, \phi) = \sum_{l=1}^{L} \left(\partial_j \phi_l \partial_k \phi_l - \frac{1}{2} g_{jk} g^{pq} \partial_p \phi_l \partial_q \phi_l \right) - \mathcal{V}(\phi) g_{jk},$$

the potential $\mathcal{V}$ is assumed to be a smooth function $\mathbb{R}^L \to \mathbb{R}$, and $\Box_g$ is the wave operator associated to g. For example, a typical model is

$$\mathcal{V}(\phi) = \frac{1}{2} \sum_{l=1}^{L} m_l^2 \phi_l^2,$$

where $m_l \geq 0$ are constants.

We say that a smooth Lorentzian manifold $(M, \widehat{g})$ and a function $\widehat{\phi}$ are background spacetime and scalar fields if they satisfy (1)–(2) with $\mathcal{F} = 0$ and if $(M, \widehat{g})$ is globally hyperbolic. The definition of global hyperbolicity is recalled below, see Definition 1.7. A globally hyperbolic manifold is isometric to a product manifold $\mathbb{R} \times N$ with the Lorentzian metric given by

$$\widehat{g} = -\beta(t, y) dt^2 + \kappa(t, y),$$

where $\beta : \mathbb{R} \times N \to \mathbb{R}^+$ is smooth and κ is a Riemannian metric on N depending smoothly on t, see [9]. Without loss of generality, we make the standing assumption that $M = \mathbb{R} \times N$, and consider equations (1)–(2) with the initial conditions

$$(3) \qquad\qquad g = \widehat{g}, \ \phi = \widehat{\phi}, \quad \text{in } (-\infty, 0) \times N.$$

The source $\mathcal{F}$ in (1)–(2) can not be arbitrary since the Bianchi identities imply that that the $\text{Ein}(g)$ is divergence free, that is, $\text{div}_g \text{Ein}(g) = 0$. Hence the stress energy tensor T needs to satisfy the conservation law

$$(4) \qquad\qquad \text{div}_g T = 0.$$

This again implies the compatibility condition

$$(5) \qquad\qquad \text{div}_g \mathcal{F}^1 + \sum_{l=1}^{L} \mathcal{F}_l^2 d\phi_l = 0,$$

see Corollary 5.8 for a detailed discussion. In order to guarantee that (1)–(3) admits a rich set of sources in a neighbourhood of $\mathcal{F} = 0$, we assume the following generic condition, that we call the *non-degeneracy condition* for the background scalar fields,

(ND) $d\widehat{\phi}_l(x)$, $l = 1, 2, 3, 4$, are linearly independent for all $x \in U$,

where $U \subset M$ is a set containing the support of the source $\mathcal{F}$ in (1)–(2).

Observe that the system (1)–(3) is invariant with respect to isometries and that the metric tensor g begins to change as soon as $\mathcal{F}$ becomes non-zero. In order to take this invariance into account, we model local measurements using Fermi coordinates. Let $x \in \{0\} \times N$ and let $\xi \in T_x M$ be timelike and future-pointing. Define $\mu_{\widehat{g}}(s) = \exp_x^{\widehat{g}}(s\xi)$, where $\exp^{\widehat{g}}$ is the exponential map on $(M, \widehat{g})$. Let X_j, $j = 0, 1, 2, 3$, be a basis of $T_x M$, with $X_0 = \xi$, and consider the following Fermi coordinates around the world line $\mu_{\widehat{g}}$,

$$(6) \qquad\qquad \Phi_{\widehat{g}}(s, y^1, y^2, y^3) = \exp_{\mu_{\widehat{g}}(s)}^{\widehat{g}} \left(\sum_{j=1}^{3} y^j Y_j^{\widehat{g}} \right),$$

where $Y_j^{\widehat{g}}$ is the parallel transport of X_j along $\mu_{\widehat{g}}$, $j = 1, 2, 3$. Here the parallel transport is taken with respect to $\widehat{g}$, and

$$s \in [0, 1], \quad (y^1, y^2, y^3) \in \overline{B_0},$$

where $B_0 \subset \mathbb{R}^3$ is a small enough neighbourhood of the origin so that the Fermi coordinates (6) are well-defined in $[0, 1] \times \overline{B_0}$. We write

$$V = (0, 1) \times B_0.$$

Let $\widehat{T} > 0$ satisfy $\Phi_{\widehat{g}}(\overline{V}) \subset (-\infty, \widehat{T}) \times N$, and consider a Lorentzian metric g on $(-\infty, \widehat{T}) \times N$ such that the corresponding Fermi coordinates $\Phi_g : V \to M$ are well-defined. Here Φ_g is defined as $\Phi_{\widehat{g}}$ above except that now the world line μ_g,
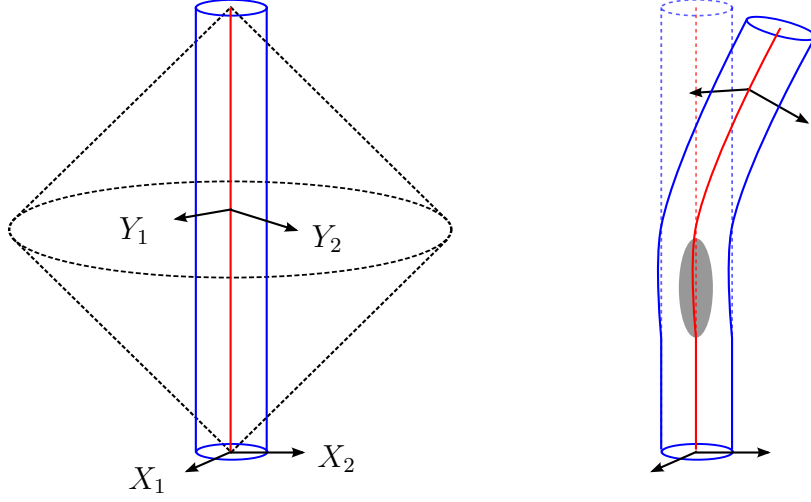
FIGURE 1. *Left.* The setting of Theorem 1.1. The solid red line depicts the worldline μ along which the frame $Y_j = Y_j^{\widehat{g}}$, $j = 1, 2, 3$, defines the Fermi coordinates (the third direction is suppressed in the picture). The blue cylinder is the coordinate neighbourhood $\Phi(V)$ and the dashed double cone is the chronological diamond $I(\mu(0), \mu(1))$. *Right.* Fermi coordinates for a perturbed metric g. As the source $\mathcal{F}$ becomes non-zero, the Fermi coordinates start to change. The support of $\mathcal{F}$ is depicted as the grey area.

the exponential map and the parallel transports are defined with respect to g. We often use the shorthand notation $\mu = \mu_{\widehat{g}}$, $\Phi = \Phi_{\widehat{g}}$ and

$$\widehat{M} = (0, \widehat{T}) \times N. \tag{7}$$

Then we model local measurements near the world line μ by the following data set

$$\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi}) = \{(\Phi_g^* g, \Phi_g^* \phi, \Phi_g^* \mathcal{F}) \in C^k(V) \; ; \tag{8}$$

$$(g, \phi, \mathcal{F}) \text{ satisfies } (1)\text{--}(2) \text{ on } \widehat{M}, \; (3) \text{ holds,}$$

$$\operatorname{supp}(\mathcal{F}) \subset \Phi_g(V), \; \left\| \Phi_g^* \mathcal{F} \right\|_{C^k(V)} < r \},$$

where $r > 0$, $k \in \mathbb{N}$, and Φ_g^* is the pullback under $\Phi_g : V \to M$. This set corresponds to the graph of the response operator used in earlier studies, see (14) below.

We emphasize that in the above model, the representations of the fields g and ϕ in the Fermi coordinates are measured only in a small neighbourhood $\Phi_g(V)$ of the world line μ_g. The neighbourhood $\Phi_{\widehat{g}}(V)$ is depicted as the blue cylinder Figure 1. Let us also remark that $(g, \phi, \mathcal{F})$ in the defining property of $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$ satisfies (5) since it satisfies (1)–(2). In particular, for fixed (g, ϕ), the sources $\mathcal{F}$ in (8) depend on (g, ϕ) via the constraint (5).

As signals in Einstein's equations propagate with finite speed, the data $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$ can not determine $\widehat{g}$ on the whole manifold M. The fastest signals, for example gravitational waves, propagate at the speed of light, and an observer moving along μ is able to probe a region U in M only if signals can reach U from μ and if signals from U can return to μ.

We will define next the set where we are able to gain information on $\widehat{g}$. This diamond type set is close to optimal. Indeed, if the chronological future and past are replaced by the causal future and past in its definition, then it is not possible to obtain information on $\widehat{g}$ in the complement of this slightly larger set, independently of the choice of the Fermi coordinates. We give the definitions of causal future and past in Section 1.1 below, and recall here only the concepts that are needed to formulate our main result. For $p, q \in M$, $p \ll q$ in $(M, \widehat{g})$ means that p and q can be joined by future pointing timelike curves in $(M, \widehat{g})$. The *chronological future* and *past* of $p \in M$ are

$$(9) \quad I^+(p) = \{q \in M \; ; \; p \ll q \text{ in } (M, \widehat{g})\}, \quad I^-(p) = \{q \in M \; ; \; q \ll p \text{ in } (M, \widehat{g})\}.$$

We will obtain information on $\widehat{g}$ on the chronological diamond type set

$$(10) \qquad\qquad I(p, q) = I^+(p) \cap I^-(q),$$

containing the world line μ in the sense that $p = \mu(0)$ and $q = \mu(1)$.

Theorem 1.1. *Suppose that smooth $\widehat{g}$ and $\widehat{\phi}$ satisfy (1)–(2) on M with $\mathcal{F} = 0$, and that $(M, \widehat{g})$ is globally hyperbolic. Consider the Fermi coordinates $\Phi : V \to M$ around a world line μ as defined in (6), and suppose that $\widehat{\phi}$ satisfies (ND) with $U = \overline{\Phi(V)}$. Then for arbitrary small $r > 0$ and large $k \in \mathbb{N}$, the data set $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$ determines the topology, differentiable structure and conformal class of the metric $\widehat{g}$ in the chronological diamond $I(\mu(0), \mu(1))$.*

In the case of a vacuum spacetime we can recover also the conformal factor.

Corollary 1.2. *Let $\widehat{g}$, $\widehat{\phi}$, Φ and μ be as in Theorem 1.1. Suppose, furthermore, that $W \subset I(\mu(0), \mu(1))$ is vacuum, i.e. Ricci-flat, and that any point in W can be connected by a curve, lying completely in W, to a point in $\Phi(V)$. Then for arbitrary small $r > 0$ and large $k \in \mathbb{N}$, the data set $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$ determines the topology and differentiable structure of W and isometry class of $\widehat{g}$ in W.*

Remark 1.3. The techniques considered in this paper can be used also to study inverse problems for non-linear hyperbolic systems encountered in applications. In fact, one of the motivations of this paper is to develop methods that will transfer tools of General Relativity to medical and seismic imaging. For instance, in medical imaging, in the the recently developed Ultrasound Elastography imaging technique the elastic material parameters are reconstructed by sending (s-polarized) elastic waves that are imaged using (p-polarized) elastic waves, see e.g. [41, 59]. This imaging method uses interaction of waves and is based on the non-linearity of the system.

Remark 1.4. The question, what is the topology of our Universe, can be considered as an inverse problem for the coupled Einstein and matter field equations. This fundamental question has been studied in [56, 77]. In addition to the topology, in this paper we consider also the determination of the conformal type of the Lorentzian metric.

Remark 1.5. Theorem 1.1 implies that, if we have two non-conformal spacetimes, a generic measurement gives different results on these manifolds. In particular, this implies that perfect spacetime cloaking, in sense of light rays, see [23, 58], is not possible with a smooth metric in a globally hyperbolic universe.

Remark 1.6. The non-degeneracy assumption (ND) could be replaced by the assumption that $d\widehat{\phi}_l(x)$, $l = 1, \ldots, L$, span T_x^*M for all $x \in \overline{\Phi(V)}$. We make the stronger assumption just to avoid cumbersome notation in the linearization stability proofs below.

1.1. **Notations and definitions.** Analogously to the chronological future and past (9), we define the causal future and past of a point $p \in M$ on a Lorentzian manifold (M, g) by

$$J_{(M,g)}^+(p) = \{q \in M \; ; \; p \leq q \text{ in } (M,g)\}, \quad J_{(M,g)}^-(p) = \{q \in M \; ; \; q \leq p \text{ in } (M,g)\}.$$

where $p \leq q$ means that p and q can be joined by a future pointing causal curve in (M, g) or $p = q$. We write also $J_{(M,g)}(p, q) = J_{(M,g)}^+(p) \cap J_{(M,g)}^-(q)$.

Definition 1.7 ([10]). *A Lorentzian manifold (M, g) is* globally hyperbolic *if there are no closed causal paths in M, and the set $J_{(M,g)}(p, q)$ is compact for any pair of points $p, q \in M$.*

We denote by ∇_g, div_g and $\mathrm{Ric}(g)$ the covariant derivative, divergence and the Ricci tensor with respect to g. Using the shorthand notation

$$(d\phi \otimes d\phi)_{jk} = \sum_{\ell=1}^{L} \partial_j \phi_\ell \partial_k \phi_\ell, \quad (\mathcal{V}'(\phi))_\ell = \partial_{\phi_\ell} \mathcal{V}(\phi).$$

where $j, k = 0, 1, 2, 3$, $\ell = 1, \ldots, L$, equation (1) can be written as

(11) $$\mathrm{Ric}(g) - d\phi \otimes d\phi - \mathcal{V}(\phi)g = F^1, \quad F^1 = I_g \mathcal{F}^1,$$

where I_g is the involution

$$I_g F^1 = F^1 - \frac{\mathrm{tr}_g F^1}{2} g.$$

Here $\mathrm{tr}_g F^1 = g^{jk} F_{jk}^1$ is the contraction with g. We will use also the shorthand notation

$$I_g \mathcal{F} = (I_g \mathcal{F}^1, \mathcal{F}^2).$$

The bundle of symmetric 2-tensors over M is denoted by

$$\mathrm{Sym}^2 = \mathrm{Sym}(T^*M \otimes T^*M).$$

For vector bundles E and F over M, we denote by $\text{Hom}(E, F)$ the vector bundle whose fibre at $x \in M$ is the space of linear maps from the fibre E_x of E to the fibre F_x of F. We write also $\text{End}(E) = \text{Hom}(E, E)$.

1.2. **Earlier results.** Inverse problems for partial differential equations is a much studied topic, however, the present theory is largely confined to the case of linear equations. For a majority of the few existing results on non-linear equations, e.g. [44, 79, 82], non-linearity is an obstruction rather than a helpful feature. On the contrary, our method to solve the inverse problem for Einstein's equations is based on the non-linear interaction of gravitational waves, and only global hyperbolicity is assumed on the underlying spacetime. In an earlier paper [49] we considered a similar method in the case of a semi-linear wave equation, and we discuss this result in more detail below. We emphasize that the present study of Einstein's equations can not use the techniques in [49] in a straightforward manner due to the much more complicated structure of the non-linearity. In particular, contrary to [49], the non-linearity appears also in the leading order terms.

As our approach uses the non-linearity in an essential way, it differs from the classical methods used to study inverse problems for linear wave models, that is, the Boundary Control method [46] and the geometric optics based approach [72, 81], that have been used to solve various inverse problems for linear wave equations. However, these methods fail when the equations have general time-dependent coefficients. For the former method this is because the sharp unique continuation result [85] does not hold for wave equations with coefficients that are smooth, but not real analytic, in the time variable [1]; and for the latter method because it is not known if the light ray transform is injective on a globally hyperbolic Lorentzian manifold. In the case of time-independent coefficients, the light ray transform reduces to the geodesic ray transform. The injectivity of the latter transform has been studied extensively in the Riemannian context [26, 43, 68, 69, 70, 80, 86].

The inverse problem for linearized Einstein's equations is open, and the same is true for other similar linear wave equations. We solve the inverse problem for the physical, non-linear Einstein's equations.

In the case of linear equations, it is now classical to model active measurements by using the Dirichlet-to-Neumann map [84]. In the case of Einstein's equations, modelling active measurements is more subtle since the spacetime can not be observed from outside, and (8) appears to be the first model for active measurements in this case. Let us now recall from [49] the model of active measurements for the non-linear wave equation

$$(12) \qquad \begin{aligned} &\Box_{\widehat{g}} u(x) + a(x)\, u(x)^2 = f(x) \quad \text{on } M, \\ &u(x) = 0, \quad \text{for } x \in M_0 \setminus J^+_{(M,\widehat{g})}(\text{supp}\,(f)), \end{aligned}$$

where a is non-vanishing and smooth, and $(M, \widehat{g})$ is globally hyperbolic. Letting μ be as in Theorem 1.1, and letting $V \subset M$ be a neighbourhood of $\mu([0, 1])$, the active measurements can be modelled by the set

$$(13) \qquad \mathcal{D}_{r,k}^{\text{wave}}(\widehat{g}) = \{(u|_V, f) \; ; \; (u, f) \text{ satisfies (12) on } \widehat{M},$$
$$\operatorname{supp}(f) \subset V, \; \|f\|_{C^k(V)} < r\}.$$

The set (13) is the analogue of (8) but simpler, since the complications related to the coordinate invariance of the Einstein's equations do not arise in this case. Let us remark that in [49] the results were actually formulated by using the response operator

$$(14) \qquad\qquad\qquad L_V f = u|_V$$

For small enough $r > 0$ and large enough $k \in \mathbb{N}$, the set (13) is the graph of L_V. It was shown in [49] that the set (13) determines the conformal class of the metric $\widehat{g}$ in $I(\mu(0), \mu(1))$. This is the analogue of Theorem 1.1 for (12).

Remark 1.8. The result in [49] deals with an inverse problem for "near field" measurements, modelled by (13) or the corresponding response operator. In the case of inverse problems for linear equations, near field measurements are typically modelled by the Dirichlet-to-Neumann map, and they are equivalent to scattering measurements or "far field" information [8]. Analogous equivalence for non-linear equations has not yet been studied but it is plausible that the theory of near field measurements is useful when studying inverse scattering problems also in this case. On related inverse scattering problems, see [27, 45, 62, 78].

Finally, even though in this paper we consider near-filed measurements, the techniques we use to study non-linear interaction of waves takes its inspiration from microlocal scattering theory [3, 4, 29, 32, 35, 36, 63, 87]. In the case of the near field measurements, the inverse problems for Einstein manifolds have been studied in the Riemannian context in [78].

2. Outline of the proof

In order to give a rough idea of the proof, let us consider here only the simple model case (12) with $\widehat{g} = -dt^2 + dx^2 + dy^2 + dz^2$, the Minkowski metric, and $a = 1$ identically. The idea is to use the source f to generate propagating singularities that interact via the non-linearity. The propagating singularities will be constructed so that their interaction produces a new singularity that behaves like a point source. Recording the singularities produced by each such "artificial point source" in $I(\mu(0), \mu(1))$ allows for a reduction to the purely geometric problem solved in [49, Th. 1.2]. Let us also mention that the geometric step has been generalized for manifolds with boundary in [34]. Another generalization involving broken geodesics is given in [51].
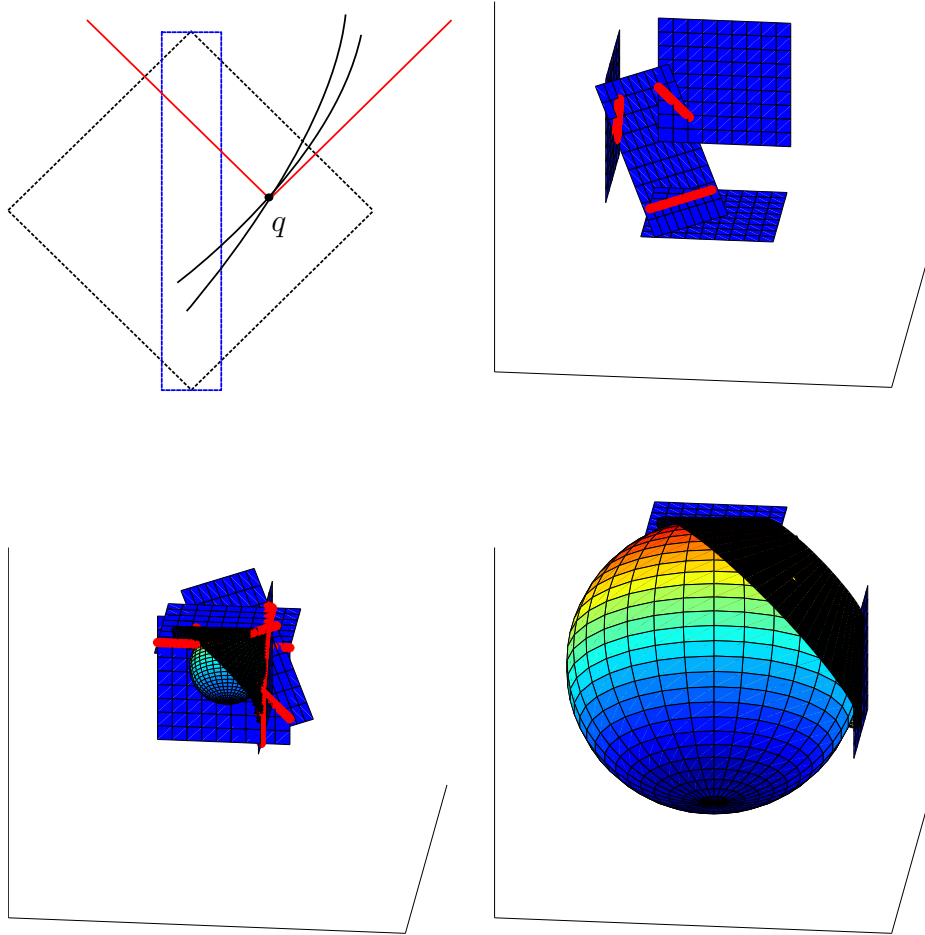
FIGURE 2.   *Top left.* A schematic picture of outline of the proof
of Theorem 1.1 as discussed in Section 2. Waves propagating near
geodesics (depicted as black curves), sent from $\Phi(V)$, interact at the
point $q \in I(\mu(0), \mu(1))$. Via a non-linear interaction, they can pro-
duce propagating singularities (in red) analogous to those generated
by a point source at q. *Rest of the panels.* Interactions of pieces
of the four plane waves u_j as time passes. *Top right.* The pieces
interact pairwise on the red line segments, no new propagating sin-
gularities appear. *Bottom left.* Interactions of triples produce conic
singularities and the interaction of all the four pieces of waves pro-
duces a spherical singularity. Only one of the four conic singularities
(in black) is shown in the picture. *Bottom right.* The spherical sin-
gularity expands as time passes. It can be detected far away from
the point where the interaction happened.

To explain this in more detail, but still roughly, let $f = \epsilon h$, $\epsilon > 0$, and write the asymptotic expansion of the solution of (12),

$$u = \epsilon w_1 + \epsilon^2 w_2 + \epsilon^3 w_3 + \epsilon^4 w_4 + \mathcal{O}(\epsilon^5),$$

where, using the notation $Q = \Box_{\widehat{g}}^{-1}$ for the causal inverse,

$$(15) \qquad w_1 = Qh, \quad w_2 = -Q(w_1^2), \quad w_3 = -2Q(w_1 w_2),$$
$$w_4 = -Q(w_2^2) - 2Q(w_1 w_3).$$

We say, for example, that w_3 results from the interaction of w_1 and w_2.

Let us make a linear change of coordinates in the Minkowski space, such that in the new coordinates $x^j, j = 1, 2, 3, 4$, the hyperplanes $K_j = \{x^j = 0\}$ are lightlike, that is, $T_x K_j$ contains a lightlike vector for all $x \in \mathbb{R}^4$. The plane waves $u_j(x) = (x^j)_+^m$, where $m > 0$ and

$$(16) \qquad (x^j)_+^m = \begin{cases} (x^j)^m, & x^j \geq 0, \\ 0, & \text{otherwise,} \end{cases}$$

are solutions to the wave equation $\Box_{\widehat{g}} u = 0$. They are singular on the hyperplanes K_j. Moreover, the singularity on K_j has a specific direction: the wave front set of u_j is the conormal bundle of K_j, that is, $\mathrm{WF}(u_j) = N^* K_j \setminus 0$, see Section 3.3 for this notation.

Interactions of the waves u_j produce new singularities, for instance, the product $u_1 u_2$ is singular on $U \cup X$ where $U = N^* K_1 \cup N^* K_2$ and $X = N^*(K_1 \cap K_2)$. The set X contains covectors that are not in U, however, in this case of the pairwise interaction, all the lightlike covectors in X are also in U. As the solution operator Q propagates only lightlike singularities, no new propagating singularities are produced in pairwise interactions.

When three waves interact, say u_1, u_2, u_3, new propagating singularities are produced since $N^*(K_1 \cap K_2 \cap K_3)$ contains lightlike covectors that are not in any $N^* K_j$, $j = 1, 2, 3$. The interaction of three waves has been analysed in [74, 61, 65]. A conic singularity comes out of the line of intersection, and the geometry of the singularity is similar to shock discontinuities that happen at the interface between subsonic and supersonic speeds, see Figure 2.

The intersection of the four hyperplanes K_j is the origin, and its conormal bundle is the fibre of the cotangent space at the origin. In particular, the conormal bundle contains the light-cone $L_0^+ \mathbb{R}^4$, and hence the interaction of the four waves u_j produces a new singularity that corresponds to a spherical wave, in the sense that the singularity coincides with that generated by a point source at the origin. Summarizing, the four wave interactions can produce artificial point sources.

The proof of Theorem 1.1 is based on an analysis of the non-linear interaction of four solutions to the linearized Einstein-scalar field equations. Of course, as the background metric $\widehat{g}$ can be very general, and as the sources $\mathcal{F}$ in (8) need to

satisfy support constraints, we can not use explicit solutions like (16). These are replaced by more general conormal distributions in the actual proof below.

We begin by showing that the Einstein-scalar field equations admit a rich set of solutions with propagating singularities. That is, in Section 3 we show that (1)–(2) can be linearized along families of solutions so that the thus obtained solutions to the corresponding linearized equations have propagating singularities, and that the geometric structure of the singularities is constrained only by a linearized, microlocal version of the conservation law (4), see Definition 3.4.

In Section 4 we study the analogue of the interaction terms (15) for the Einstein-scalar field equations. Although on the level of the geometry of propagation of singularities the analysis of the interaction terms is very similar with the above model case, it is more complicated to show that interaction terms do not vanish, say, due to them cancelling each other. We show this by considering the principal symbols of the interaction terms, viewed as conormal distributions. Let us also point out that the analysis of the principal symbols is not analogous to that in [49], in particular, the asymptotic approach in [49, Section 3.4] does not seem to work in our setting.

Finally in Section 5 we deal with the global geometric aspects of the proof and finish the reduction to [49, Th. 1.2]. Also we treat there the complications related to the coordinate invariance of the Einstein's equations, that are manifest both when sending and receiving propagating singularities.

To get a flavour of the complications when trying to detect singularities in g, say gravitational waves, observe that the Fermi coordinates Φ_g in the definition of the data set (8) may be non-smooth when g is non-smooth. The use of non-smooth coordinates Φ_g can hide singularities in g. However, if g is a conormal distribution on a submanifold of codimension one, certain singularities are generically visible in the coordinates Φ_g. More precisely, the principal symbol of $\Phi_g^* g$ is non-vanishing generically on a rank 2 subbundle of the rank 10 bundle of symmetric 2-tensors. Indeed, as the Fermi coordinates are used, $\Phi_g^* g$ has 6 degrees of freedom to begin with, and on top of this, the conservation law (4) imposes 4 additional constraints. Definition 3.4 below gives the analogue (4) for principal symbols. In physical terms, if g is a gravitational wave, the rank 2 subbundle corresponds to the two polarizations of that can be observed, see [57]. More detailed physical explanations on the used measurement model can be found from the earlier extended preprint version of this paper [50].

3. Construction of solutions with propagating singularities

In this section we will show that Einstein's equations coupled with scalar fields (1)–(2) admit a rich set of solutions with propagating singularities. These solutions are constructed for sources $\mathcal{F}$ having small support. The locally generated propagating singularities are then used in later sections to probe the background spacetime $(M, \widehat{g})$.

3.1. **Einstein's equations as a system of quasilinear wave equations.** In order to treat (1)–(2) as a hyperbolic system, we need to fix a suitable gauge. We follow the presentation in [76, pp. 152-154]. Let $\widehat{g}$ be a smooth Lorentzian metric tensor on M. Let $\widehat{\Gamma}^r_{pq}$ be the Christoffel symbols of $\widehat{g}$ and define $\widehat{H}_k = g_{kr}g^{pq}\widehat{\Gamma}^r_{pq}$. Define also

$$\Gamma_{pkq} = \frac{1}{2}(\partial_p g_{kq} + \partial_q g_{pk} - \partial_k g_{pq}), \quad \Gamma_k = g^{pq}\Gamma_{pkq}.$$

As the difference of two connections is a tensor, we see that the local expression $\widehat{H}_k - \Gamma_k$ defines a tensor, that we denote by $H_{\widehat{g}}(g)$. That is,

$$(17) \qquad\qquad H_{\widehat{g}}(g)_k = \widehat{H}_k - \Gamma_k, \quad k = 0, 1, 2, 3.$$

We define the reduced Ricci tensor

$$\mathrm{Ric}_{\widehat{g}}(g) = \mathrm{Ric}(g) + \nabla_g H_{\widehat{g}}(g),$$

and consider the reduced Einstein's equations coupled with scalar fields

$$(18) \qquad \mathrm{Ric}_{\widehat{g}}(g) - d\phi \otimes d\phi - \mathcal{V}(\phi)g = F^1, \quad \text{in } \widehat{M},$$

$$\Box_g \phi - \mathcal{V}'(\phi) = F^2, \quad \text{in } \widehat{M},$$

$$g = \widehat{g}, \ \phi = \widehat{\phi}, \quad \text{in } (-\infty, 0) \times N.$$

This is a quasilinear wave equation. Indeed, locally [76, eq. (14.3)],

$$\mathrm{Ric}_{\widehat{g}}(g)_{jk} = -\frac{1}{2}g^{pq}\partial_p\partial_q g_{jk} + \nabla_j \widehat{H}_k + g^{pq}g^{rs}(\Gamma_{prj}\Gamma_{qsk} + \Gamma_{prj}\Gamma_{qks} + \Gamma_{prk}\Gamma_{qjs}),$$

where only the first term contains second order derivatives of g. Here ∇_j is the covariant derivative ∇_g in the coordinate direction ∂_j.

It follows from a typical fixed point argument that (18) has a unique solution (g, ϕ) when (F^1, F^2) is in a small neighbourhood of the origin in $C^k_0(U; \mathrm{Sym}^2 \oplus \mathbb{R}^L)$, with $U \subset \widehat{M}$ bounded and $k \in \mathbb{N}$ large, see [42] or [47] for such an argument and Appendix B for further discussion.

Let (g, ϕ) solve (18) and set $\mathcal{F} = I_g F$ where $F = (F^1, F^2)$. Then $(g, \phi, \mathcal{F})$ solves (1)–(2) if and only if the following gauge condition holds

$$(19) \qquad\qquad H_{\widehat{g}}(g) = 0.$$

This is equivalent with the compatibility condition (5), see Corollary 5.8 in Appendix A. In the next section we will give a construction of sources $\mathcal{F}$ that satisfy (5), and that generate a rich set of propagating singularities, as demonstrated later in Section 3.5.

It will be occasionally useful that the subprincipal part of the second equation in (18) simplifies under the condition (19). Indeed, in local coordinates (19) is equivalent with $g^{pq}\Gamma^j_{pq} = g^{pq}\widehat{\Gamma}^j_{pq}$, $j = 0, 1, 2, 3$, and then

$$(20) \qquad\qquad \Box_g \phi = g^{-1}(D, D)\phi - g^{pq}\widehat{\Gamma}^j_{pq}\partial_{x^j}\phi,$$

where $g^{-1}(D, D)\phi = g^{pq}\partial_p\partial_q\phi$. In particular, the subprincipal part of the second equation in (18) does not contain derivatives of g.

3.2. **Linearization stability.** We will next formulate a linearization stability result for the system (1)–(3). Linearization stability has been studied extensively in the context of the initial value problem for (1)–(2), i.e. when $\mathcal{F} = 0$, see e.g. [12, 13, 17, 22, 25]. The distinctive feature of our result is that the source $\mathcal{F}$ can be non-zero but its support is localized. The localization is essential when considering inverse problems, in particular, it allows us to satisfy the constraint $\mathrm{supp}\,(\mathcal{F}) \subset \Phi_g(V)$ in (8).

Proposition 3.1. *Let $U \subset \widehat{M}$ be open, and suppose that $L \geq 4$ and that $\widehat{\phi}$ satisfies the non-degeneracy condition (ND). If $(f^1, f^2) \in C^k(\widehat{M}; \mathrm{Sym}^2 \oplus \mathbb{R}^L)$ is compactly supported in U, k is large enough, and if*

$$(21) \qquad\qquad div_{\widehat{g}} f^1 + f^2 \cdot d\widehat{\phi} = 0,$$

then there is a neighbourhood $\mathcal{E} \subset \mathbb{R}$ of the origin and a family of solutions $(g(\epsilon), \phi(\epsilon), \mathcal{F}(\epsilon))$, $\epsilon \in \mathcal{E}$, to (1)–(3) that satisfies

$$(22) \qquad \partial_\epsilon \mathcal{F}^j(\epsilon)|_{\epsilon=0} = f^j, \; j = 1, 2, \qquad \bigcup_{j=1,2} supp\,(\mathcal{F}^j(\epsilon)) \subset \bigcup_{j=1,2} supp\,(f^j).$$

Here $f^2 \cdot d\widehat{\phi}$ is the inner product on $\mathbb{R}^L$, that is, $f^2 \cdot d\widehat{\phi} = \sum_{l=1}^L f_l^2 d\widehat{\phi}_l$.

In the context of the initial value problem for (1)–(2), there are counterexamples to linearization stability, that is, there are solutions to the linearization of (1)–(2) that do not arise as the derivative of a one-parameter family of solutions to the original non-linear problem. In fact, linearization stability for the initial value problem for Einstein's equations in vacuum is characterized by the absence of Killing fields [2]. The generic condition (ND) can be viewed as an analogue of this property in the present setting.

We emphasize that by (22), the support of the source functions $\mathcal{F}(\epsilon)$ can be chosen to be small if the same is true for the supports of f^1 and f^2. This is crucial from the point of view of the inverse problem that we consider, since it allows us to choose sources supported near the world line μ. If the sources $\mathcal{F}(\epsilon)$ are interpreted as being produced by a measurement device, see Appendix C, $\mathcal{F}(\epsilon)$ having a small support means that the device can be located in a small neighbourhood of μ.

In what follows, we will use only a microlocal version of Proposition 3.1, but the above version might be of independent interest and we include its proof. The proof is based on the following lemma.

Lemma 3.2. *Let $U \subset \widehat{M}$ be open and suppose that $\widehat{\phi}$ satisfies the non-degeneracy condition (ND). Then there are vector bundle homomorphisms*

$$A_\phi^1 \in C^\kappa(U; \mathrm{Hom}(T^*U, \mathbb{R}^L)), \quad A_\phi^2 \in C^\kappa(U; \mathrm{Hom}(\mathbb{R}^L, \mathbb{R}^L)),$$

that satisfy the following:

(i) $A^1_\phi(v) \cdot d\phi = v$ for all $v \in C^\infty(U; T^*U)$.
(ii) $A^2_\phi(w) \cdot d\phi = 0$ for all $w \in C^\infty(U; \mathbb{R}^L)$.
(iii) $A^1_\phi(w \cdot d\phi) + A^2_\phi(w) = w$ for all $w \in C^\infty(U; \mathbb{R}^L)$.
(iv) The map $\phi \mapsto (A^1_\phi, A^2_\phi)$ is C^∞-smooth in a neighbourhood of $\widehat{\phi}$, and depends only on $d\phi$.

We postpone the proof of the lemma in Appendix A.

Proof of Proposition 3.1. Using Lemma 3.2, define

$$(23) \qquad \mathcal{F}^1(\epsilon) = \epsilon f^1, \quad \mathcal{F}^2(\epsilon) = \epsilon A^1_\phi(-\mathrm{div}_g f^1) + \epsilon A^2_\phi(f^2).$$

By the discussion in Appendix B, the system (18), with F being replaced by $I_g \mathcal{F}(\epsilon)$, has a unique solution $(g(\epsilon), \phi(\epsilon))$ on $\widehat{M}$ when $\epsilon \in \mathcal{E}$ and $\mathcal{E} \subset \mathbb{R}$ is small enough neighbourhood of the origin. The compatibility condition (5) holds by Lemma 3.2. Indeed, by the properties (i) and (ii), we have $\mathcal{F}^2 \cdot d\phi = -\mathrm{div}_g \mathcal{F}^1$. Moreover, as f^1 and f^2 satisfy the linearized compatibility condition (21), it follows from property (iii) of Lemma 3.2 that

$$\partial_\epsilon \mathcal{F}^2(\epsilon)|_{\epsilon=0} = A^1_{\widehat{\phi}}(f^2 \cdot d\widehat{\phi}) + A^2_{\widehat{\phi}}(f^2) = f^2.$$

$\square$

3.3. **Notations for analysis of singularities.** We refer to [39] for detailed exposition of the concepts that we use for analysis of singularities. We use the notation $\mathrm{WF}(u)$ for the wave front set of a distribution $u \in \mathcal{D}'(M)$. The wave front is a subset of $T^*M \setminus 0$, the cotangent bundle with the zero section removed, and its projection on the base space M is called the singular support, $\mathrm{singsupp}(u)$. Fibres $T^*_x M$ are occasionally identified with the sets $\{(x, \xi) \in T^*M; \ \xi \in T^*_x M\}$, $x \in M$.

For a conic Lagrangian submanifold $\Lambda \subset T^*M \setminus 0$ and a vector bundle E over M, we denote by $I^p(\Lambda; E)$ the space of Lagrangian distributions of order $p \in \mathbb{R}$ associated to Λ, and taking values in E. We write also $I(\Lambda; E) = \bigcup_{p \in \mathbb{R}} I^p(\Lambda; E)$, and sometimes omit writing E if it is clear from the context. Recall that

$$\mathrm{WF}(u) \subset \Lambda, \quad u \in I(\Lambda; E).$$

For a submanifold $K \subset M$, we define the space of conormal distributions on K by $I(K; E) = I(N^*K \setminus 0; E)$, where N^*K is the conormal bundle of K. The singleton $\{x\}$, $x \in M$, is considered to be a zero dimensional manifold with $N^*\{x\} = T^*_x M$. We follow the convention that conormal bundles contain the zero section, but the wave front sets do not.

The principal symbol of $u \in I(\Lambda; E)$ is denoted by $\sigma[u](x, \xi)$, $(x, \xi) \in \Lambda$, see [39, Th. 25.1.9] for the definition, and we omit writing out the half density and Maslov factors. Moreover, we denote the lift of the bundle E to Λ still by E, see Definition

3.3 for a more detailed discussion in the particular case that is important for us. The spaces of symbols are denoted by $S^p(\Lambda; E)$, and we write also

$$S(\Lambda; E) = \bigcup_{p \in \mathbb{R}} S^p(\Lambda; E).$$

When $\Lambda \subset T^*(M \times M) \setminus 0$ the distributions in $I(\Lambda; \mathrm{End}(E))$ correspond to kernels of Fourier integral operators acting on sections of E. We denote by Δ the twisted diagonal of $T^*(M \times M) \setminus 0$, that is,

$$\Delta = \{(x, \xi, x, -\xi); \ (x, \xi) \in T^*M \setminus 0\}.$$

Then $I(\Delta; \mathrm{End}(E))$ are the pseudodifferential operators, and in this case we consider the symbol to be defined on $T^*M \setminus 0$.

For two conic Lagrangian submanifolds $\Lambda_0, \Lambda_1 \subset T^*M \setminus 0$ intersecting cleanly, that is,

$$T_\lambda(\Lambda_0) \cap T_\lambda(\Lambda_1) = T_\lambda(\Lambda_0 \cap \Lambda_1), \quad \lambda \in \Lambda_0 \cap \Lambda_1,$$

the space of paired Lagrangian distributions associated to Λ_0, Λ_1 is denoted by $I^{p,l}(\Lambda_0, \Lambda_1; E)$, $p, l \in \mathbb{R}$. Microlocally away from the intersection $\Lambda_0 \cap \Lambda_1$, it holds that $u \in I^{p+l}(\Lambda_0 \setminus \Lambda_1; E)$ and $u \in I^p(\Lambda_1 \setminus \Lambda_0; E)$, see [66]. We write also $I(\Lambda_0, \Lambda_1; E) = \bigcup_{p,l \in \mathbb{R}} I^{p,l}(\Lambda_0, \Lambda_1; E)$. Let us point out that [66] discusses only the case where E is a trivial line bundle, but the generalization for arbitrary vector bundles is straightforward. We refer also to [31] in relation to theory of paired Lagrangian distributions. The first use of paired Lagrangian distributions in the context of cosmology was [30].

For us, the primary example of a paired Lagrangian distribution is the causal inverse of the wave operator. This is discussed in detail in next section.

3.4. **Propagation of singularities for linear wave equations.** Let $(M, \widehat{g})$ be a background spacetime. We denote by $\Box = \Box_{\widehat{g}}$ the wave operator associated to $\widehat{g}$, and by Σ and $L_xM \subset T_x^*M$, $x \in M$, the characteristic variety of $\Box$ and light cones of $(M, \widehat{g})$. We follow the convention that the characteristic variety does not contain the zero section but each light cone contain the corresponding vertex, that is,

$$(24) \qquad\qquad \Sigma = \bigcup_{x \in M} L_xM \setminus 0.$$

For a set $S \subset T^*(M \times M)$ we define its twist by

$$S' = \{(x, \xi, y, -\eta); \ (x, \xi, y, \eta) \in S\}.$$

Observe that $(S')' = S$.

Let V be a first order differential operator acting on sections of a vector bundle E over M. As $(M, \widehat{g})$ is globally hyperbolic, the causal inverse

$$Q = (\Box + V)^{-1}$$

is well-defined and its Schwartz kernel is in $I(\Delta, \Lambda; \mathrm{End}(E))$ where Λ is the twist of the canonical relation Λ' of $\square$ defined by

$$(25) \qquad \Lambda' = \{(x, \xi, y, \eta) \in \Sigma^2; \ (y, \eta) = \beta(s; x, \xi), \ s \in \mathbb{R}\}.$$

Here the scalar wave operator $\square$ acts diagonally and $\beta(\cdot; x, \xi)$ is the bicharacteristic of $\square$ satisfying $\beta(0; x, \xi) = (x, \xi)$. That is,

$$\beta(s; x, \xi) = (\gamma_{x,\xi}(s), \dot{\gamma}^*_{x,\xi}(s)),$$

where $\gamma_{x,\xi}$ is the geodesic of $(M, \widehat{g})$ with the initial data (x, ξ_*), and we use the notation $\xi^j_* = \widehat{g}^{jk}\xi_k$, $\xi \in T^*_x M$, and $v^*_j = \widehat{g}_{jk}v^k$, $v \in T_x M$, for isomorphism between the tangent and cotangent vectors.

Recall that Q propagates singularities along $\Lambda' \setminus \Delta$ in the following sense: if $(x, \xi, y, \eta) \in \Lambda' \setminus \Delta$ then there is linear isomorphism $\widetilde{q}(x, \xi, y, \eta)$ between the fibres E_x and E_y of E such that

$$(26) \qquad \sigma[Qf](y, \eta) = \widetilde{q}(x, \xi, y, \eta)\sigma[Qf](x, \xi),$$

whenever $\beta(\cdot; x, \xi)$ does not intersect $\mathrm{WF}(f)$ between (y, η) and (x, ξ). To be more precise, suppose that $(y, \eta) = \beta(s; x, \xi)$ with $s > 0$. Then (26) holds if $\beta(s'; x, \xi) \notin \mathrm{WF}(f)$ for $s' \in [0, s]$. The case $s < 0$ is analogous.

The linear map $\widetilde{q}(x, \xi, y, \eta)$ is obtained by solving the transport equation

$$(27) \qquad -i\mathscr{L}_H \widetilde{q} + c\widetilde{q} = 0$$

along the bicharacteristic $\beta(\cdot; x, \xi)$, with the initial condition given by the identity map, see [21, Th. 5.3.1]. Here H is the Hamilton vector field corresponding to the principal symbol $-\widehat{g}(\xi, \xi)$ of $\square$, c is the subprincipal symbol of $\square + V$, and $\mathscr{L}_H$ is the Lie derivative along H. Solving the transport equation in the opposite direction gives the inverse of $\widetilde{q}(x, \xi, y, \eta)$.

We define for a set $S \subset T^*M$,

$$(28) \qquad \Lambda(S) = \{(y, \eta) \in \Sigma; \ (x, \xi, y, \eta) \in \Lambda', \ (x, \xi) \in S\},$$

and say that S Λ-invariant if

$$(29) \qquad S \subset \Sigma \quad \text{and} \quad \Lambda(S) = S.$$

Moreover, we say that a smooth submanifold $K \subset M$ is Λ-invariant if $N^*K \setminus 0$ is.

Let Y be a submanifold of M. We assume that each lightlike geodesic intersects Y at most once, and that the intersection is transversal. If $f \in I(Y; E)$, then $u = Qf$ is in $I(N^*Y \setminus 0, \Lambda(N^*Y); E)$. A more precise statement can be formulated by using the conic Lagrangian submanifold with boundary, the flowout N^*Y in the future direction,

$$\Lambda^+(N^*Y) = \{(y, -\eta) \in \Sigma; \ (y, \eta) = \beta(s; x, \xi), \ s \in \mathbb{R}, \ (x, \xi) \in \Sigma \cap N^*Y, \ x \leq y\}.$$

Taking into account the causality, it holds that $u \in I(N^*Y \setminus 0, \Lambda^+(N^*Y); E)$, see [66]. There the authors give also a relation between $\sigma[u]$, as a symbol in

$S(\Lambda^+(N^*Y); E)$, and $\sigma[f]$, as a symbol in $S(Y; E)$, at points $(x, \xi) \in \partial\Lambda^+(N^*Y)$. The equation (6.7) in [66] defines a map

$$\sigma[f](x, \xi) \mapsto \sigma[u](x, \xi), \quad (x, \xi) \in \partial\Lambda^+(N^*Y), \tag{30}$$

that is a linear automorphism on the fibre E_x. We denote the composition of $\widetilde{q}$ and (30) by q. Then $q = q(x, \xi, y, \eta)$ gives a linear isomorphism

$$q : E_x \to E_y, \quad q : \sigma[f](x, \xi) \mapsto \sigma[u](y, \eta), \quad (y, \eta) = \beta(s; x, \xi), \ s \in \mathbb{R}. \tag{31}$$

3.5. **Microlocal linearization stability.** In this section we give a microlocal version of Proposition 3.1. We begin by considering the principal part of the compatibility condition (21).

Definition 3.3. *For a submanifold $Y \subset M$, Sym^2_Y is the lift of*

$$\mathrm{Sym}^2 = \mathrm{Sym}(T^*M \otimes T^*M)$$

on $N^*Y \setminus 0$. That is, Sym^2_Y is a vector bundle over $N^*Y \setminus 0$ and at a point $(x, \xi) \in N^*Y \setminus 0$ its fibre is

$$(\mathrm{Sym}^2_Y)_{(x,\xi)} = \mathrm{Sym}(T^*_x M \otimes T^*_x M).$$

When considering the space of principal symbols of conormal disributions on Y, taking values on Sym^2_Y, we write simply $S(Y; \mathrm{Sym}^2)$.

Definition 3.4. *Let $(M, \widehat{g})$ be a background spacetime, and let $Y \subset M$ be a submanifold. We define a subbundle $\mathbb{S}_Y$ of Sym^2_Y over $N^*Y \setminus 0$ by the following equation for $h \in \mathrm{Sym}^2_Y$,*

$$h(\xi_*, v) - \frac{\mathrm{tr}_{\widehat{g}} h}{2} \widehat{g}(\xi_*, v) = 0, \quad \text{for all } (x, \xi) \in N^*Y \setminus 0, \ v \in T_x M, \tag{32}$$

where ξ_* is the vector $\xi^j_* = \widehat{g}^{jk}\xi_k$. We denote by $S(Y; \mathbb{S})$ the subspace of symbols $S(Y; \mathrm{Sym}^2)$ taking values on $\mathbb{S}_Y$, that is,

$$S(Y; \mathbb{S}) = \{s \in S(Y; \mathrm{Sym}^2); \ s(x, \xi) \in \mathbb{S}_Y, \ (x, \xi) \in N^*Y \setminus 0\}. \tag{33}$$

Lemma 3.5. *The bundle $\mathbb{S}_Y$ has rank 6.*

Proof. Equation (32) is equivalent with $\widetilde{h} = I_{\widehat{g}} h$ and

$$\widetilde{h}(\xi_*, v) = 0, \quad \text{for all } (x, \xi) \in N^*Y \setminus 0, \ v \in T_x M. \tag{34}$$

Write $e_1 = (1, 0, 0, 0)$, $(\widetilde{h}\xi_*)_j = \widetilde{h}_{jk}\xi^k_*$, and $\widehat{\xi} = \xi_*/|\xi_*|$ where $|\cdot|$ is the Euclidean norm in the coordinates. Choose a neighbourhood W of (x, ξ) and a smooth map $R : W \to O(4)$ such that $R(\widehat{\xi})\widehat{\xi} = e_1$ for $(x, \xi) \in W$. Define for $a = (a_1, \ldots, a_6)$,

$$\mathcal{H}_{ker} : \mathbb{R}^6 \to \mathrm{Sym}(\mathbb{R}^4 \otimes \mathbb{R}^4), \quad \mathcal{H}_{ker}(a) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a_1 & a_2 & a_3 \\ 0 & a_2 & a_4 & a_5 \\ 0 & a_3 & a_5 & a_6 \end{pmatrix}. \tag{35}$$

Now $\widetilde{h}\xi_* = |\xi_*|\widetilde{h}R(\widehat{\xi})^T e_1$, whence, in local coordinates, (34) is equivalent with

$$(36) \qquad R(\widehat{\xi})\widetilde{h}R(\widehat{\xi})^T \in \mathcal{H}_{ker}(\mathbb{R}^6), \quad \text{for all } (x,\xi) \in N^*Y \setminus 0.$$

Now for each $(x,\xi) \in W$, the map $a \mapsto I_{\widehat{g}}R(\widehat{\xi})^T \mathcal{H}_{ker}(a)R(\widehat{\xi})$ gives a parametrization of the fibre $(\mathbb{S}_Y)_{(x,\xi)}$, and thus $\mathbb{S}_Y$ has rank 6. $\qquad \square$

Proposition 3.6. *Let $(M,\widehat{g})$ be a background spacetime, let $Y \subset M$ be a submanifold, let $U \subset M$ be a neighbourhood of $\overline{Y}$, and suppose that $\widehat{\phi} \in C^\infty(M;\mathbb{R}^L)$ satisfies the non-degeneracy condition (ND). Write $E = \mathrm{Sym}^2 \oplus \mathbb{R}^L$ and*

$$(37) \qquad I_Y^p = \{f \in I^p(Y;E); \ \sigma[f] \in S^{p+1}(Y; \mathbb{S} \oplus \mathbb{R}^L)\}.$$

Then there are $p_0 \in \mathbb{R}$ and $k_0 \in \mathbb{N}$ such that for $p \leq p_0$ and $k \geq k_0$ there are maps

$$A_{(g,\phi)} : I_Y^p \to I^p(Y;E), \quad (g,\phi) \in C^k(M;E),$$
$$\mathbb{F} : S^{p+1}(Y; \mathbb{S} \oplus \mathbb{R}^L) \to I_Y^p,$$

with the following properties for each $s \in S^{p+1}(Y; \mathbb{S} \oplus \mathbb{R}^L)$,

   (A1) $f = \mathbb{F}(s)$ *satisfies the support condition $\mathrm{supp}(f) \subset U$, $A_{(g,\phi)}$ is a differential operator, and $\mathcal{F} = A_{(g,\phi)}(f)$ satisfies the compatibility condition (5).*
   (A2) $F = I_{\widehat{g}}A_{(\widehat{g},\widehat{\phi})}(f)$ *satisfies $\sigma[F] = s$.*
   (A3) *The map $(g,\phi) \mapsto A_{(g,\phi)}(f)$ is C^∞-smooth in a neighbourhood of $(\widehat{g},\widehat{\phi})$, and depends on the derivatives of g and ϕ only up to the first order.*

Proof. Observe that for each $(x,\xi) \in N^*Y \setminus 0$ the map

$$\iota_\xi h = I_{\widehat{g}}h(\xi_*, \cdot)$$

is surjective from $(\mathrm{Sym}_Y^2)_{(x,\xi)}$ to T_x^*M. Indeed, the kernel of ι_ξ is given by the equation (34), and as seen in the proof of Lemma 3.5, the kernel is of dimension 6. As the rank of Sym^2 is 10, the map ι_ξ has 4 dimensional range. Let us now choose a right inverse r_ξ of ι_ξ such that $r_\xi \in C^\infty(N^*Y \setminus 0; \mathrm{Hom}(T^*M, \mathrm{Sym}^2))$ and that r_ξ maps $S^{p+1}(Y;T^*M)$ to $S^p(Y;\mathrm{Sym}^2)$.

Let $s = (s^1, s^2) \in S^{p+1}(Y; \mathbb{S} \oplus \mathbb{R}^L)$. As the principal symbol map is an isomorphism

$$I^p(Y;E)/I^{p-1}(Y;E) \to S^{p+1}(Y;E)/S^p(Y;E),$$

see e.g. [39, Th. 25.1.9], we can choose

$$\mathbb{F}_{pri}^1 : S^{p+1}(Y;\mathbb{S}) \to I^p(Y;\mathrm{Sym}^2), \quad \mathbb{F}^2 : S^{p+1}(Y;\mathbb{R}^L) \to I^p(Y;\mathbb{R}^L),$$
$$\mathbb{F}_{sub}^1 : S^p(Y;\mathrm{Sym}^2) \to I^{p-1}(Y;\mathrm{Sym}^2),$$

so that $\sigma[\mathbb{F}_{pri}^1(s^1)] = s^1$, $\sigma[\mathbb{F}^2(s^2)] = s^2$, $\sigma[\mathbb{F}_{sub}^1(s^3)] = s^3$, and that $\mathbb{F}_{pri}^1(s^1)$, $\mathbb{F}^2(s^2)$ and $\mathbb{F}_{sub}^1(s^3)$ are all supported in U. As $I_{\widehat{g}}$ is an involution, by (32) it holds for $s^1 \in S^{p+1}(Y;\mathbb{S})$ that the principal symbol of $\mathrm{div}_{\widehat{g}}I_{\widehat{g}}\mathbb{F}_{pri}^1(s^1)$ vanishes. In other words,

$$\sigma[\mathrm{div}_{\widehat{g}}I_{\widehat{g}}\mathbb{F}_{pri}^1(s^1)] = \iota_\xi \sigma[\mathbb{F}_{pri}^1(s^1)] \in S^{p+1}(Y;T^*M).$$

Thus we can define the map

$$z : S^{p+1}(Y; \mathbb{S}) \times S^{p+1}(Y; \mathbb{R}^L) \to S^p(Y; \mathrm{Sym}^2),$$

$$z(s^1, s^2) = r_\xi \left(-\iota_\xi \sigma[\mathbb{F}^1_{pri}(s^1)] - s^2 \cdot d\widehat{\phi} \right),$$

and also the maps

(38)
$$\mathbb{F}(s^1, s^2) = \left(\mathbb{F}^1_{pri}(s^1) + \mathbb{F}^1_{sub}(z(s^1, s^2)), \mathbb{F}^2(s^2) \right),$$

$$A_{(g,\phi)}(f^1, f^2) = (I_g f^1, A^1_\phi(-\mathrm{div}_g I_g f^1) + A^2_\phi(f^2)).$$

Observe that the definition of $A_{(g,\phi)}$ is closely related to the choice in (23).

Now the the compatibility condition (5) follows from (i) and (ii) in Lemma 3.2. Setting $(f^1, f^2) = \mathbb{F}(s^1, s^2)$ and $(F^1, F^2) = I_{\widehat{g}} A_{(\widehat{g},\widehat{\phi})}(f^1, f^2)$, we have

$$\sigma[F^1] = \sigma[I_{\widehat{g}}^2 f^1] = \sigma[f^1] = \sigma[\mathbb{F}^1_{pri}(s^1)] = s^1.$$

Moreover,

$$\sigma[\mathrm{div}_{\widehat{g}} I_{\widehat{g}} f^1]) = \iota_\xi \sigma[\mathbb{F}^1_{pri}(s^1) + \mathbb{F}^1_{sub}(z(s^1, s^2))]$$

$$= \iota_\xi \sigma[\mathbb{F}^1_{pri}(s^1)] + \iota_\xi r_\xi \left(-\iota_\xi \sigma[\mathbb{F}^1_{pri}(s^1)] - s^2 \cdot d\widehat{\phi} \right) = -s^2 \cdot d\widehat{\phi},$$

and therefore using (iii) in Lemma 3.2 it holds that

$$\sigma[F^2] = A^1_\phi(-\sigma[\mathrm{div}_{\widehat{g}} I_{\widehat{g}} f^1]) + A^2_\phi(\sigma[f^2]) = A^1_\phi(s^2 \cdot d\widehat{\phi}) + A^2_\phi(s^2) = s^2.$$

The smoothness follows from (iv) in Lemma 3.2. $\qquad\square$

Consider $s \in S^p(Y; \mathbb{S} \oplus \mathbb{R}^L)$ with negative enough $p \in \mathbb{R}$ and define the family $F(\epsilon) = \epsilon I_g A_{(g,\phi)}(E(s))$, $\epsilon \in \mathbb{R}$. Now the system (18) has a unique solution (g, ϕ) on $\widehat{M}$ when $\epsilon \in \mathcal{E}$ and $\mathcal{E} \subset \mathbb{R}$ is small enough neighbourhood of the origin. Moreover, setting $\mathcal{F} = I_g F$, we have that the family $(g, \phi, \mathcal{F})$, $\epsilon \in \mathcal{E}$, satisfies (1)–(2). As $\sigma[\partial_\epsilon F^j|_{\epsilon=0}] = s$, we say that the equation (1)–(2) has microlocal linearization stability for principal symbols in $S^p(Y; \mathbb{S} \oplus \mathbb{R}^L)$.

3.6. **Propagation of singularities for linearized Einstein's equations.** Let $\mathcal{E} \subset \mathbb{R}$ be a neighbourhood of the origin, and let us consider a family

(39)
$$(g(\epsilon), \phi(\epsilon), F(\epsilon)) \in C^1(\mathcal{E}; C^2(\widehat{M}; (\mathrm{Sym}^2 \oplus \mathbb{R}^L)^2))$$

of solutions to (18) near a background spacetime and scalar fields $(M, \widehat{g})$ and $\widehat{\phi}$ in the sense that $g(0) = \widehat{g}$, $\phi(0) = \widehat{\phi}$ and $F(0) = 0$. We write

(40)
$$v(\epsilon) = (g(\epsilon) - \widehat{g}, \phi(\epsilon) - \widehat{\phi}).$$

Then $\dot{v} = \partial_\epsilon v|_{\epsilon=0}$ satisfies a linear differential equation that we denote by

(41)
$$\Box_{\widehat{g}, \widehat{\phi}} \dot{v} = \dot{F} \quad \text{on } \widehat{M},$$

where $\dot{F} = \partial_\epsilon F|_{\epsilon=0}$. We stress that (41) is coordinate invariant since (18) is, and express next (41) in local coordinates.

To simplify the notation, we rescale the first equation in (18) by -2. Then (18) can be written locally in the form

$$(42) \qquad g^{-1}(D, D)v + A(x, v, Dv) = F,$$

where, for a local trivialization $U \subset \mathbb{R}^4 \times \mathbb{R}^{10+L}$ of the bundle $\mathrm{Sym}^2 \oplus \mathbb{R}^L$,

$$(x, y, z) \mapsto A(x, y, z) : U \times \mathbb{R}^{4(10+L)} \to \mathbb{R}^{10+L}$$

is a smooth map. Differentiating (42) with respect to ϵ at $\epsilon = 0$, we obtain

$$(43) \qquad \widehat{g}^{-1}(D, D)\dot{v} + D_z A(x, 0, 0) D\dot{v} + D_y A(x, 0, 0)\dot{v} = \dot{F}.$$

Thus $\Box_{\widehat{g}, \widehat{\phi}}$ is of the form $\Box + V$ where $\Box = \Box_{\widehat{g}}$ and V is a first order differential operator on the vector bundle $\mathrm{Sym}^2 \oplus \mathbb{R}^L$.

As the linearized equation (41) is of the form discussed in Section 3.4, we know how it propagates singularities. However, we still need to understand how the constraint (32) is propagated. Put differently, we analyse next what is the image of the rank 6 subbundle $\mathbb{S}_Y$ of Sym^2_Y under the isomorphism q defined by (31).

Assume now that the family (39) is constructed as in the end of Section 3.5. Then $\sigma[\dot{F}^1]$ takes values in $\mathbb{S}_Y$, and g satisfies the gauge condition (19). Differentiating (19) with respect to ϵ at $\epsilon = 0$, and writing the result in local coordinates, gives

$$\widehat{g}^{pq}(\partial_p \dot{g}_{kq} + \partial_q \dot{g}_{pk} - \partial_k \dot{g}_{pq}) = r_k \dot{g}, \quad k = 0, 1, 2, 3,$$

where $\dot{g} = \partial_\epsilon g|_{\epsilon=0}$ and r_k is a zeroth order operator. Taking the principal symbol of this expression yields

$$\widehat{g}^{pq}(\xi_p \sigma[\dot{g}_{kq}] + \xi_q \sigma[\dot{g}_{pk}] - \xi_k \sigma[\dot{g}_{pq}]) = 0, \quad k = 0, 1, 2, 3,$$

or equivalently, writing $h = \sigma[\dot{g}]$,

$$(44) \qquad \xi^p_* h_{pk} - \frac{\mathrm{tr}_{\widehat{g}} h}{2}\xi_k = 0, \quad k = 0, 1, 2, 3.$$

Supposing that $N^*K \setminus 0 \subset \Lambda^+(N^*Y)$ for a submanifold $K \subset M$, these four conditions for h coincide with the definition of the rank 6 (equivalently, codimension 4) subbundle $\mathbb{S}_K$ on N^*K, see Definition 3.4 with Y being replaced by K. As q is a linear isomorphism, we see that the map

$$(45) \qquad \sigma[\dot{F}](x, \xi) \mapsto \sigma[\dot{v}](y, \eta), \quad (y, \eta) = \beta(s; x, \xi), \ s \in \mathbb{R},$$

gives an isomorphism between the fibres $(\mathbb{S}_Y \oplus \mathbb{R}^L)_{(x,\xi)}$ and $(\mathbb{S}_K \oplus \mathbb{R}^L)_{(y,\eta)}$. Here (x, ξ) is in $\partial\Lambda^+(N^*Y)$ and $\beta(\cdot; x, \xi)$ is assumed to intersect $\partial\Lambda^+(N^*Y)$ only at (x, ξ).

Recall that the gauge condition (19) implies that (20) holds. This again implies that, writing $\dot{v} = (\dot{g}, \dot{\phi})$, the second term in (43) is of the form

$$D_z A(x, 0, 0) D\dot{v} = \begin{pmatrix} a_{11}(x) & a_{12}(x) \\ 0 & a_{22}(x) \end{pmatrix} \begin{pmatrix} D\dot{g} \\ D\dot{\phi} \end{pmatrix}, \quad a_{22}(x)D\dot{\phi} = -\widehat{g}^{pq}\widehat{\Gamma}^j_{pq}\partial_{x^j}\dot{\phi}.$$

In particular, the subprincipal symbol of $\square_{\widehat{g},\widehat{\phi}}$, which appears as c in (27), is of the above upper triangular form. This observation leads to the following remark.

Remark 3.7. The restriction of the map (45) on the subspace $(\mathbb{S}_Y)_{(x,\xi)}$ of the fibre $(\mathbb{S}_Y \oplus \mathbb{R}^L)_{(x,\xi)}$ is an isomorphism between the two subspaces of dimension 6,

$$(46) \qquad (\mathbb{S}_Y)_{(x,\xi)} \to (\mathbb{S}_K)_{(y,\eta)}.$$

4. INTERACTION OF SINGULAR SOLUTIONS

In this section we show how a non-linear interaction of singular solutions to (41) can behave, on the microlocal level, as a point source. In order to have a product calculus for principal symbols, we consider the case where the singular solutions are conormal distributions.

As the singularities of a solution to (41) are invariant under the map Λ, see (28), we are restricted to the case where the conormal distributions are associated to submanifolds of codimension one in M. This again implies that we need four such submanifolds to get a single point as their intersection. For this reason, we begin by studying the linearization of (18) with respect to a four dimensional family of solutions. We note that the interaction of two waves having very strong, impulse type singularities has been studied in [55].

4.1. Fourfold linearization.
Let $\epsilon \in \mathcal{E}$, a neighbourhood of the origin in $\mathbb{R}^4$, and consider a family

$$(47) \qquad (g(\epsilon), \phi(\epsilon), F(\epsilon)) \in C^4(\mathcal{E}; C^2(\widehat{M}; (\mathrm{Sym}^2 \oplus \mathbb{R}^L)^2))$$

of solutions to (18) near a background spacetime and scalar fields $(M, \widehat{g})$ and $\widehat{\phi}$. Define v as in (40). Analogously to (41), we see that

$$(48) \qquad v_j = \partial_{\epsilon_j} v|_{\epsilon=0}, \quad j = 1, 2, 3, 4,$$

satisfies the wave equation

$$(49) \qquad \square_{\widehat{g},\widehat{\phi}} v_j = f_j,$$

where $f_j = \partial_{\epsilon_j} F|_{\epsilon=0}$. Let us now consider interactions of linearized solutions v_j. To simplify the computations we assume that

$$(50) \qquad \partial_{\epsilon_j}\partial_{\epsilon_k} F|_{\epsilon=0} = \partial_{\epsilon_j}\partial_{\epsilon_k}\partial_{\epsilon_l} F|_{\epsilon=0} = \partial_{\epsilon_1}\partial_{\epsilon_2}\partial_{\epsilon_3}\partial_{\epsilon_4} F|_{\epsilon=0} = 0$$

for all distinct j, k and l.

Recall that $A(x, y, z)$ is defined by (42). We use the shorthand notation

$$A_j^k(x, y, z) = D_y^j D_z^k A(x, y, z)$$

for the derivatives of A, that we consider as multilinear forms, and omit writing (x, y, z) as a parameter for A below. Note that A is a quadratic polynomial in z, and therefore $A_j^k = 0$ when $k > 2$. We write also

$$(51) \qquad v_{j_1 \ldots j_n} = \partial_{\epsilon_{j_1}} \ldots \partial_{\epsilon_{j_n}} v|_{\epsilon=0},$$

denote by S_n the permutations on $\{1, 2, \ldots, n\}$, and define

$$P(u_1, u_2, \ldots, u_n) = \widehat{g}^{-1} g_1 \widehat{g}^{-1} g_2 \ldots \widehat{g}^{-1} g_n \widehat{g}^{-1}(D, D),$$

where $u_j = (g_j, \phi_j) \in C^2(\widehat{M}; \mathrm{Sym}^2 \oplus \mathbb{R}^L)$.

Differentiating (42) with respect to ϵ_j and ϵ_k, $j \neq k$, we see that v_{jk} satisfies the wave equation

$$(52) \qquad \Box_{\widehat{g}, \widehat{\phi}} v_{jk} = \mathscr{I}_2(v_j, v_k) + \mathscr{R}_2(v_j, v_k),$$

where

$$\mathscr{I}_2(u_1, u_2) = \sum_{\tau \in S_2} \widetilde{\mathscr{I}_2}(u_{\tau(1)}, u_{\tau(2)}), \quad \widetilde{\mathscr{I}_2}(u_1, u_2) = P(u_1)u_2 - \frac{1}{2} A^2(Du_1, Du_2),$$

$$\mathscr{R}_2(u_1, u_2) = \sum_{\tau \in S_2} \widetilde{\mathscr{R}_2}(u_{\tau(1)}, u_{\tau(2)}), \quad \widetilde{\mathscr{R}_2}(u_1, u_2) = -A_1^1(Du_1, u_2) - \frac{1}{2} A_2(u_1, u_2).$$

Observe that $\mathscr{I}_2$ and $\mathscr{R}_2$ are bilinear differential operators of order 2 and 1, respectively. Here the order is the total number of derivatives.

Differentiating (42) with respect to ϵ_j, ϵ_k and ϵ_l, with j, k and l distinct, we see that v_{jkl} satisfies the wave equation

$$(53) \qquad \Box_{\widehat{g}, \widehat{\phi}} v_{jkl} = \sum_{\tau \in S(j,k,l)} \Big(\mathscr{I}_3^1\big(v_{\tau(1)}, v_{\tau(2)}, v_{\tau(3)}\big) + \mathscr{I}_3^2\big(v_{\tau(1)\tau(2)}, v_{\tau(3)}\big)$$

$$+ \mathscr{R}_3^1\big(v_{\tau(1)}, v_{\tau(2)}, v_{\tau(3)}\big) + \mathscr{R}_3^2\big(v_{\tau(1)\tau(2)}, v_{\tau(3)}\big) \Big),$$

where $S(j, k, l) = \{\tau : \{1, 2, 3\} \to \{j, k, l\}; \ \tau \text{ bijection}\}$, and $\mathscr{I}_3^1$ and $\mathscr{I}_3^2$ are the following trilinear and bilinear second order differential operators

$$\mathscr{I}_3^1(u_1, u_2, u_3) = -P(u_1, u_2)u_3 - \frac{1}{2} A_1^2(Du_1, Du_2, u_3),$$

$$\mathscr{I}_3^2(u_1, u_2) = \frac{1}{2} \sum_{\tau \in S_2} \left(P(u_1)u_2 - \frac{1}{2} A^2(Du_1, Du_2) \right).$$

Here $\mathscr{R}_3^1$ and $\mathscr{R}_3^2$ are trilinear and bilinear first order differential operators of similar structure. We define

$$\mathscr{I}_3(u_1, u_2, u_3) = \sum_{\tau \in S_3} \Big(\mathscr{I}_3^1(u_{\tau(1)}, u_{\tau(2)}, u_{\tau(3)}) + \mathscr{I}_3^2(\Box_{\widehat{g}, \widehat{\phi}}^{-1} \mathscr{I}_2(u_{\tau(1)}, u_{\tau(2)}), u_{\tau(3)}) \Big),$$

and use an analogous definition for $\mathscr{R}_3$. Then (53) can be rewritten in the form

$$(54) \qquad \Box_{\widehat{g}, \widehat{\phi}} v_{jkl} = \mathscr{I}_3(v_j, v_k, v_l) + \mathscr{R}_3(v_j, v_k, v_l).$$

Recall that Σ is the characteristic variety of the wave operator $\Box_{\widehat{g}, \widehat{\phi}}$. In the symbol computations in Sections 4.3 and 4.4 below, we will consider the causal inverse $\Box_{\widehat{g}, \widehat{\phi}}^{-1}$ on the microlocal region $T^*M \setminus \Sigma$ where it is a pseudodifferential operator of order

-2. In this region $\mathscr{I}_3$ and $\mathscr{R}_3$ are trilinear pseudodifferential operators of orders 2 and 1, respectively.

Differentiating (42) again, we see that v_{1234} satisfies the wave equation

$$(55) \qquad \Box_{\widehat{g},\widehat{\phi}} v_{1234} = \mathscr{I}_4(v_1, v_2, v_3, v_4) + \mathscr{R}_4(v_1, v_2, v_3, v_4),$$

where

$$(56) \qquad \mathscr{I}_4(v_1, v_2, v_3, v_4) = \sum_{\tau \in S_4} \widetilde{\mathscr{I}_4}(v_{\tau(1)}, v_{\tau(2)}, v_{\tau(3)}, u_4),$$

and writing $\mathscr{Q}_j = \Box_{\widehat{g},\widehat{\phi}}^{-1} \mathscr{I}_j$, $j = 2, 3$,

$$
\begin{aligned}
\widetilde{\mathscr{I}_4}(u_1, \ldots, u_4) ={}& P(u_1, u_2, u_3)u_4 - \frac{1}{2}P(u_1, \mathscr{Q}_2(u_2, u_3))u_4 - \frac{1}{2}P(\mathscr{Q}_2(u_1, u_2), u_3)u_4 \\
&+ \frac{1}{4}P(\mathscr{Q}_2(u_1, u_2))\mathscr{Q}_2(u_3, u_4) - \frac{1}{2}P(u_1, u_2)\mathscr{Q}_2(u_3, u_4) + \frac{1}{6}P(\mathscr{Q}_3(u_1, u_2, u_3))u_4 \\
&+ \frac{1}{6}P(u_1)\mathscr{Q}_3(u_2, u_3, u_4) - \frac{1}{4}A^2(D\mathscr{Q}_2(u_1, u_2), D\mathscr{Q}_2(u_3, u_4)) \\
&- \frac{1}{6}A^2(D\mathscr{Q}_3(u_1, u_2, u_3), Du_4) - \frac{1}{2}A_1^2(D\mathscr{Q}_2(u_1, u_2), Du_3, u_4) \\
&- \frac{1}{4}A_1^2(Du_1, Du_2, \mathscr{Q}_2(u_3, u_4)) - \frac{1}{4}A_2^2(Du_1, Du_2, u_3, u_4).
\end{aligned}
$$

Also, in the microlocal region $T^*M \setminus \Sigma$, $\mathscr{R}_4$ is a 4-linear pseudodifferential operator of order 1 of similar structure.

We will next consider the case that v_j, $j = 1, 2, 3, 4$, are conormal distributions associated to a Λ-invariant submanifolds of codimension one. We derive first an upper bound, in the sense of the inclusion relation, for the outgoing singular support, $\mathrm{singsupp}(v_{1234})$, assuming that not all the four incoming singular supports, $\mathrm{singsupp}(v_j)$, $j = 1, 2, 3, 4$, intersect, see Lemma 4.6. Then we proceed to study the principal symbol of the fourfold interaction terms $\mathscr{I}_4 + \mathscr{R}_4$. We consider the terms where $\mathscr{Q}_2$ appears at most once in Lemma 4.8 and the terms where $\mathscr{Q}_2$ appears twice in Lemma 4.9.

4.2. **Upper bounds for wave front sets and singular supports.** We omit the proof of the following elementary lemma.

Lemma 4.1. *Let M be a smooth manifold and let $K_1, K_2 \subset M$ be transversal submanifolds. Then for all $x \in K_1 \cap K_2$ it holds that $N_x^*(K_1 \cap K_2) = N_x^*K_1 \oplus N_x^*K_2$.*

Recall that Λ-invariance is defined by (29).

Lemma 4.2 (Twofold interaction). *Let $u_1, u_2 \in \mathcal{D}'(M)$ and suppose that $WF(u_j)$ is Λ-invariant for $j = 1, 2$. Let $Q \in I(\Delta, \Lambda)$. The following holds*
  (i) *If $(x, \xi) \in WF(u_1)$ implies $(x, -\xi) \notin WF(u_2)$, then*

$$\mathrm{singsupp}(Q(u_1 u_2)) \subset \mathrm{singsupp}(u_1) \cup \mathrm{singsupp}(u_2).$$

(ii) *If $u_j \in I(K_j)$, $j = 1, 2$, for transversal submanifolds K_1 and K_2 in M, then*

$$WF(Q(u_1 u_2)) \subset N^* K_1 \cup N^* K_2 \cup N^*(K_1 \cap K_2).$$

Proof. By the assumption in (i) and [40, Th. 8.2.10],

$$\mathrm{WF}\,(u_1 u_2) \subset \{(x, \xi + \eta) \in T^* M \setminus 0;\ (x, \xi) \in \mathrm{WF}\,(u_1) \text{ or } \xi = 0, \text{ and}$$
$$(x, \eta) \in \mathrm{WF}\,(u_2) \text{ or } \eta = 0\}.$$

As the sum of two lightlike vectors is lightlike if and only if it is a scalar multiple of one of the summands, see e.g. [67, Lem. 27, p. 141], claim (i) follows from [28, Prop. 2.1] due to the Λ-invariance of WF (u_j), $j = 1, 2$.

Let us now consider (ii). Due to the transversality, we can apply again [40, Th. 8.2.10], and analogously with the above,

(57)
$$L_x M \cap N^*_x(K_1 \cap K_2) = N^*_x K_1 \cup N^*_x K_2.$$

Claim (ii) follows again from [28, Prop. 2.1]. □

Remark 4.3. The non-linear interaction of waves may cause extraordinary singularities. For example, M. Beals showed for the wave equation

$$\Box u(t, y) + b(t, y) u(t, y)^3 = 0$$

in $\mathbb{R}^{1+3}$ that there are solutions for which the singular support of the Cauchy data $(u|_{t=0}, \partial_t u|_{t=0})$ is the point $\{0\}$, but the singular support of u contains the entire solid cone $\{(t, y) \in \mathbb{R}^4;\ |y| < t\}$, see [6, Thm. 2.10] and [5]. This is related to the case when the condition in the claim (i) of Lemma 4.2 is violated.

Definition 4.4. *We say that submanifolds $K_1, \ldots, K_n \subset M$ are transversal if for any disjoint subsets $I, J \subset \{1, \ldots, n\}$ the manifolds $\bigcap_{i \in I} K_i$ and $\bigcap_{j \in J} K_j$ are transversal.*

Observe that if $K_1, \ldots, K_4 \subset M$ are transversal submanifolds of codimension 1, then Lemma 4.1 implies for any $x \in \bigcap_{j=1}^4 K_j$ that

$$T^*_x M = N^*_x K_1 \oplus \cdots \oplus N^*_x K_4.$$

Lemma 4.5 (Threefold interaction). *Let $K_1, K_2, K_3 \subset M$ be transversal Λ-invariant submanifolds of codimension 1. Let $u_j \in I(K_j)$, $j = 1, 2, 3$, and $Q_1, Q_2 \in I(\Delta, \Lambda)$. Then*

$$WF(Q_2(u_3 Q_1(u_2 u_1))) \subset \mathcal{X} \cup \mathcal{L} \cup \mathcal{U}$$

where

(58)
$$\mathcal{X} = N^*(K_1 \cap K_2 \cap K_3), \quad \mathcal{L} = \Lambda(\mathcal{X}),$$
$$\mathcal{U} = \bigcup_{\tau \in S_3} \left(N^* K_{\tau(1)} \cup N^*(K_{\tau(2)} \cap K_{\tau(3)}) \right).$$

Proof. By Lemma 4.2, and using [40, Th. 8.2.10] as in its proof,

$$\mathrm{WF}\,(u_3 Q_1(u_2 u_1)) \subset \mathcal{X} \cup \mathcal{U}.$$

By (57) it holds that $\Lambda(\mathcal{U} \cap \Sigma) \subset \mathcal{U}$, and the claim follows as in Lemma 4.2.   □

We denote by $\pi : T^*M \to M$ the canonical projection to the base manifold. The equation (57) implies that the sets in the previous lemma satisfy

$$(59) \qquad \pi(\mathcal{X} \cup \mathcal{L} \cup \mathcal{U}) = \pi(\mathcal{L}) \cup K_1 \cup K_2 \cup K_3.$$

Lemma 4.6. *Let* $K_j \subset M$, $j = 1, 2, 3, 4$, *be transversal* Λ*-invariant submanifolds of codimension 1, and define for* $\tau \in S_4$ *the sets*

$$(60) \qquad \mathcal{K} = \bigcup_{\tau \in S_4} \pi(\mathcal{L}_\tau) \cup \bigcup_{j=1}^4 K_j, \quad \mathcal{L}_\tau = \Lambda(\mathcal{X}_\tau), \quad \mathcal{X}_\tau = N^*(\bigcap_{j=1}^3 K_{\tau(j)}).$$

Suppose that $\bigcap_{j=1}^4 K_j = \emptyset$ *and that* $v_j \in I(K_j; \mathrm{Sym}^2 \oplus \mathbb{R}^L)$. *Then for a solution* v_{1234} *of (55) it holds that*

$$\mathrm{singsupp}(v_{1234}) \subset \mathcal{K}.$$

Proof. We consider only the singularities caused by the intersection

$$X = K_1 \cap K_2 \cap K_3,$$

the other cases are similar. We use the notation from Lemma 4.5. As the four manifolds K_j do not all intersect, the distribution v_4 is smooth near X. Then, near X, all the singularities of

$$v_{1234} = \Box_{\widehat{g},\widehat{\phi}}^{-1} \mathscr{I}_4(v_1, v_2, v_3, v_4) + \Box_{\widehat{g},\widehat{\phi}}^{-1} \mathscr{R}_4(v_1, v_2, v_3, v_4)$$

are due to terms of the form $Q_2(u_3 Q_1(u_2 u_1))$ where Q_1 and Q_2 are in $I(\Delta, \Lambda)$, and $u_j \in I(K_{\tau'(j)})$, $j = 1, 2, 3$, for some $\tau' \in S_3$. Therefore Lemma 4.5 and (59) imply that $\mathrm{singsupp}(v_{1234}) \cap X$ is a subset of $\mathcal{K}$.

There may be interactions also away from X since $\pi(\mathcal{L})$ may intersect K_4. These interactions represented by a sum of terms like

$$w_1 = P(v_{123})v_4, \quad w_2 = P(v_4)v_{123}, \quad w_3 = -A^2(Dv_{123}, Dv_4)$$

in (56). Lemma 4.5 implies that $\mathrm{singsupp}(v_{123}) \subset \pi(\mathcal{L}) \cup K_1 \cup K_2 \cup K_3$, and hence by Lemma 4.2

$$\mathrm{singsupp}(w_j) \subset \mathrm{singsupp}(v_{123}) \cup \mathrm{singsupp}(v_4) \subset \pi(\mathcal{L}) \cup K_1 \cup \cdots \cup K_4 \subset \mathcal{K},$$

as long as $(x', \xi') \in \mathrm{WF}\,(v_{123})$ implies $(x', -\xi') \notin \mathrm{WF}\,(v_4)$. Note that $(x', \pm\xi')$ lie on the image β of the same bicharacteristic. If $(x', \xi') \in \mathcal{L}$ then $\pi(\beta)$ intersects X, and if $(x', -\xi') \in N^*K_4$ then $\beta \subset N^*K_4$ by the Λ-invariance. But this implies that K_4 intersects X which is a contradiction with the assumption $\bigcap_{j=1}^4 K_j = \emptyset$.   □

4.3. **Principal symbols of fourfold interaction terms.** We refer to [52, Lem. 3.3] for a proof of the following lemma. We ignore the 2π factors that do not play any role in the analysis below.

Lemma 4.7. *Let M be a smooth manifold and let $K_1, K_2 \subset M$ be transversal submanifolds. Let $u_j \in I(K_j)$, $j = 1, 2$, and let χ be a pseudodifferential operator of order zero that vanishes in a conical neighbourhood of $N^*K_1 \cup N^*K_2$. Then $\chi(u_1 u_2) \in I(K_1 \cap K_2)$ and for all $(x, \xi) \in N^*(K_1 \cap K_2)$ it holds that*

$$\sigma[\chi(u_1 u_2)](x, \xi) = \sigma[\chi](x, \xi)\, \sigma[u_1](x, \xi_1)\, \sigma[u_2](x, \xi_2),$$

where $\xi = \xi_1 + \xi_2$, $\xi_j \in N_x^* K_j$.

Lemma 4.8 (Type $1+3$ fourfold interaction). *Let $K_1, \ldots, K_4 \subset M$ be transversal Λ-invariant submanifolds of codimension 1. Let $u_j \in I(K_j)$, $Q_1, Q_2 \in I(\Delta, \Lambda)$ and let*

$$(61) \qquad x \in \bigcap_{j=1}^{4} K_j \setminus \pi(\mathcal{L} \setminus \mathcal{X}),$$

where $\mathcal{X}$ and $\mathcal{L}$ are defined by (58). Let $\xi \in L_x M \subset N_x^* K_1 \oplus \cdots \oplus N_x^* K_4$ be non-zero, and define $\eta_\ell = \sum_{j=1}^{\ell} \xi_j$, $\ell = 1, 2, 3, 4$, where $\xi_j \in N_x^* K_j$ give the direct sum decomposition $\xi = \sum_{j=1}^{4} \xi_j$. Suppose that

$$(62) \qquad \xi_j \neq 0, \;\; j = 1, 2, 3, 4.$$

Write $s_j = \sigma[u_j]$, $j = 1, 2, 3, 4$, and $q_j(y, \eta) = \sigma[Q_j](y, \eta)$, $j = 1, 2$, for $y \in M$ and $\eta \in T_y^* M \setminus L_y M$. Then there are $h \in I(\{x\})$ and $r \in \mathcal{D}'(M)$ such that

$$(63) \qquad u_4 Q_2(u_3 Q_1(u_2 u_1)) = h + r, \quad (x, \xi) \notin WF(r),$$

$$\sigma[h](x, \xi) = s_4(x, \xi_4) q_2(x, \eta_3) s_3(x, \xi_3) q_1(x, \eta_2) s_2(x, \xi_2) s_1(x, \xi_1).$$

Above it is assumed that the product in (63) is a well defined distribution. This is the case if the conormal distributions u_j have negative enough orders, or if N^*K_4 and $\mathcal{L}$ are transversal. Note that Q_j is a pseudodifferential operator on $\Delta \setminus \Lambda$ and thus we can consider $\sigma[Q_j]$ to be defined on $T^*M \setminus 0$. Observe also that the condition (61) guarantees that propagating wave fronts from earlier threefold interactions do not interfere with the fourfold interaction at the point x.

Proof. Observe that $\eta_3 \notin L_x M$. Indeed, $\xi = \eta_3 + \xi_4 \in L_x M$ and $\xi_4 \in L_x M$ by the Λ-invariance of K_4. If also $\eta_3 \in L_x M$ then by the transversality of $K_1, \ldots, K_4$ it holds that $\xi_4 = 0$ or $\eta_3 = 0$, both of which lead to a contradiction with (62). Also $\eta_2 \notin L_x M$ as the sum of $\xi_j \in L_x M$, $j = 1, 2$.

We recall that Σ is the characteristic variety defined by (24). For $\ell = 2, 3, 4$, let χ_ℓ be a microlocalization near (x, η_ℓ), that is, χ_ℓ is a pseudodifferential operator of order zero such that $1 - \chi_\ell$ is smoothing in a conical neighbourhood of (x, η_ℓ). The microlocalizations can be chosen so that χ_ℓ is smoothing outside a conical neighbourhood U_ℓ of (x, η_ℓ) satisfying

(i) $U_\ell \cap \Sigma = \emptyset$, $\ell = 2, 3$,

(ii) $U_\ell \cap \left(N^* K_\ell \cup N^* (\bigcap_{j=1}^{\ell-1} K_j) \right) = \emptyset$, $\ell = 2, 3, 4$.

Indeed, (i) can be fulfilled since $\eta_\ell \notin L_x M$, $\ell = 2, 3$, and (ii) since

$$\eta_\ell \notin N_x^* K_\ell \cup N_x^* (\bigcap_{j=1}^{\ell-1} K_j), \quad \ell = 2, 3, 4.$$

The latter holds since $\eta_\ell \in N_x^* K_\ell$ leads to the contradiction $\xi_j = 0$, $j = 1, \dots, \ell-1$, and since $\eta_\ell \in N_x^* (\bigcap_{j=1}^{\ell-1} K_j)$ leads to the contradiction $\xi_\ell = 0$.

Moreover, the microlocalization χ_4 can be chosen to be of the form

$$\chi_4(x, D) = \chi_{42}(x, D)\chi_{41}(x),$$

where $\chi_{41} \in C_0^\infty(M)$ satisfies

(iii) if $y \in \text{supp}\,(\chi_{41})$ and $(y, \eta) \in N^* K_4$ then $(y, -\eta) \notin \mathcal{L}$.

Indeed, for a small enough neighbourhood $U \subset M$ of x, the assumption (61) implies that it is enough to consider the part $\mathcal{L}_U$ of $\mathcal{L}$ that flows out from U, that is, the part

$$\mathcal{L}_U = \Lambda(N^*(U \cap K_1 \cap K_2 \cap K_3)).$$

It follows from the transversality of $K_1, \dots, K_4$ that, for small enough U,

$$(y, \eta) \in N^* K_4 \quad \text{implies} \quad (y, -\eta) \notin \mathcal{L}_U.$$

Taking $\chi_{41} \in C_0^\infty(U)$, we have (iii).

We omit writing the point x as a parameter below. Define

$$h = \chi_4(u_4 Q_2 \chi_3(u_3 Q_1 \chi_2(u_2 u_1))).$$

As $U_2 \cap (N^* K_1 \cup N^* K_2) = \emptyset$, Lemma 4.7 implies that

$$\chi_2(u_2 u_1) \in I(K_1 \cap K_2), \quad \sigma[\chi_2(u_2 u_1)](\eta_2) = s_2(\xi_2) s_1(\xi_1).$$

As $U_2 \cap \Sigma = \emptyset$, Q_1 is a pseudodifferential operator in U_2. Therefore

$$Q_1 \chi_2(u_2 u_1) \in I(K_1 \cap K_2), \quad \sigma[Q_1 \chi_2(u_1 u_2)](\eta_2) = q_1(\eta_2) \sigma[\chi_2(u_1 u_2)](\eta_2).$$

Analogously, as $U_3 \cap (N^* K_3 \cup N^*(K_1 \cap K_2) \cup \Sigma) = \emptyset$,

$$Q_2 \chi_3(u_3 Q_1 \chi_2(u_2 u_1)) \in I(K_1 \cap K_2 \cap K_3),$$

$$\sigma[Q_2 \chi_3(u_3 Q_1 \chi_2(u_2 u_1))](\eta_3) = q_2(\eta_3) s_3(\xi_3) \sigma[Q_1 \chi_2(u_1 u_2)](\eta_2),$$

and as $U_4 \cap (N^* K_4 \cup N^*(K_1 \cap K_2 \cap K_4)) = \emptyset$,

$$h \in I(K_1 \cap K_2 \cap K_3 \cap K_4), \quad \sigma[h](\xi) = s_4(\xi_4) \sigma[Q_2 \chi_3(u_3 Q_1 \chi_2(u_2 u_1))](\eta_3).$$

Observe that $u_4 Q_2(u_3 Q_1(u_2 u_1)) = h + r_1 + r_2 + r_3$, where

$$r_1 = (1 - \chi_4)(u_4 Q_2(u_3 Q_1(u_2 u_1))), \quad r_2 = \chi_4(u_4 Q_2(1 - \chi_3)(u_3 Q_1(u_2 u_1))),$$

$$r_3 = \chi_4(u_4 Q_2 \chi_3(u_3(1 - \chi_2)Q_1(u_2 u_1))).$$

It remains to show that $(x, \xi) \notin \text{WF}\,(r_j)$, $j = 1, 2, 3$. This is immediate for r_1.

Let us consider r_2. We use the shorthand notation $w = Q_2(1 - \chi_3)(u_3 Q_1(u_2 u_1))$ and $\mathrm{WF}_x(w) = \mathrm{WF}(w) \cap T_x M$. Observe that $\mathcal{U} \cap T_x M \subset \mathcal{X}$ where $\mathcal{U}$ is given by (58). This together with Lemma 4.5 and $x \notin \pi(\mathcal{L} \setminus \mathcal{X})$ imply that $\mathrm{WF}_x(w) \subset \mathcal{X}$. Using the transversality, we see that

$$(y, \eta) \in N^* K_4 \setminus 0 \quad \text{implies} \quad (y, -\eta) \notin \mathcal{X} \cup \mathcal{U}.$$

The above, together with (iii) and [40, Th. 8.2.10], imply that any $\zeta \in \mathrm{WF}_x(u_4 w)$ has a direct sum decomposition of the form $\zeta = \zeta_4 + \omega_3$ where

(64) $\quad \zeta_4 \in \mathrm{WF}_x(u_4) \cup 0 \subset N_x^* K_4, \quad \omega_3 \in \mathrm{WF}_x(w) \cup 0 \subset N_x^* K_1 \oplus N_x^* K_2 \oplus N_x^* K_3.$

As $\eta_3 \notin L_x M$, it holds that $\eta_3 \notin \mathrm{WF}_x(Q_2(1 - \chi_3)u)$ for all $u \in \mathcal{D}'(M)$. In particular, $\eta_3 \notin \mathrm{WF}_x(w)$. Thus

$$\xi = \xi_4 + \eta_3 \notin \mathrm{WF}_x(u_4 w) \supset \mathrm{WF}_x(r_2).$$

Let us consider r_3. Define $w_1 = (1 - \chi_2)Q_1(u_2 u_1)$ and $w_2 = Q_2 \chi_3(u_3 w_1)$. Analogously to (64), any $\zeta \in \mathrm{WF}_x(u_4 w_2)$ decomposes as $\zeta = \zeta_4 + \omega_3$ where $\zeta_4 \in N_x^* K_4$ and $\omega_3 \in \mathrm{WF}_x(w_2) \cup 0$. Suppose now that $\omega_3 \notin L_x M$. Then $\omega_3 \in \mathrm{WF}(u_3 w_1)$, and Lemma 4.2 and [40, Th. 8.2.10] imply that $\omega_3 = \zeta_3 + \omega_2$ where $\zeta_3 \in N_x^* K_3$ and $\omega_2 \in \mathrm{WF}_x(w_1) \cup 0$. As $\eta_3 \notin L_x M$, $\eta_2 \notin \mathrm{WF}_x(w_1)$ and $\eta_2 \neq 0$, it holds that

$$\xi = \xi_4 + \eta_3 = \xi_4 + \xi_3 + \eta_2 \notin \mathrm{WF}_x(u_4 w_2) \supset \mathrm{WF}_x(r_3).$$

$\square$

An analogous proof gives the following.

Lemma 4.9 (Type $2 + 2$ fourfold interaction)**.** *Let $K_1, \ldots, K_4 \subset M$ be transversal Λ-invariant submanifolds of codimension 1. Let $u_j \in I(K_j)$ and $Q_1, Q_2 \in I(\Delta, \Lambda)$. Let $(x, \xi) \in T^* M$, and use the notation from Lemma 4.8. Suppose that (62) holds. Then there are $h \in I(\{x\})$ and $r \in \mathcal{D}'(M)$ such that*

(65) $\quad Q_2(u_4 u_3)Q_1(u_2 u_1) = h + r, \quad (x, \xi) \notin WF(r),$

$\quad \sigma[h](x, \xi) = q_2(x, \xi_3 + \xi_4)s_4(x, \xi_4)s_3(x, \xi_3)q_1(x, \xi_1 + \xi_2)s_2(x, \xi_2)s_1(x, \xi_1).$

4.4. **Principal symbol of the fourfold interaction.** Let $(x, \xi) \in T^* M \setminus 0$, $f \in \mathcal{D}'(M)$, and suppose that there are $h \in I(\{x\})$ and $r \in \mathcal{D}'(M)$ such that $f = h + r$ and $(x, \xi) \notin \mathrm{WF}(r)$. Then we define $\sigma[f](x, \xi) = \sigma[h](x, \xi)$.

Let $K_1, \ldots, K_4 \subset M$ be transversal Λ-invariant submanifolds of codimension 1, $x \in \bigcap_{j=1}^4 K_j$, $\xi \in L_x M \setminus 0$ and

(66) $\qquad\qquad \xi = \xi_1 + \xi_2 + \xi_3 + \xi_4 \quad \text{where} \quad \xi_j \in N_x^* K_j.$

Suppose that the covectors ξ_j satisfy (62) and

(67) $\qquad\qquad x \in \bigcap_{j=1}^4 K_j \setminus \bigcup_{\tau \in S_4} \pi(\mathcal{L}_\tau \setminus \mathcal{X}_\tau),$

where $\mathcal{X}_\tau$ and $\mathcal{L}_\tau$ are as in (60). Note that (67) is a version of the global geometric condition (61) that is symmetric with respect to the indices $1, 2, 3, 4$.

We denote by $L_x^+ M$ the future-pointing covectors in $L_x M$. If $b_1, \ldots, b_4 \in L_x^+ M$ are such that

$$(68) \qquad\qquad N_x^* K_j = \mathbb{R} b_j, \quad j = 1, 2, 3, 4,$$

then for fixed (x, ξ), the covectors ξ_j defined by (66) can be viewed as functions of $\boldsymbol{b} = (b_1, \ldots, b_4)$, and we write

$$(69) \qquad\qquad \xi_j = \xi_j(\boldsymbol{b}).$$

Furthermore, in view of the constraint (32) for the symbols of solutions to (41), we write

$$\mathbb{S}_j(\boldsymbol{b}) = (\mathbb{S}_{K_j})_{(x, \xi_j(\boldsymbol{b}))}.$$

Supposing that $v_j \in I(K_j; \mathrm{Sym}^2 \oplus \mathbb{R}^L)$ have negative enough orders, the interaction terms $\mathscr{I}_4$ and $\mathscr{R}_4$ in (56) are well-defined. By applying Lemmas 4.8 and 4.9 componentwisely, we see that there are $r_1, r_2 \in \mathcal{D}'(M)$ such that $(x, \xi) \notin \mathrm{WF}(r_j)$, and $\mathscr{I}_4(v_1, v_2, v_3, v_4) - r_1$ and $\mathscr{R}_4(v_1, v_2, v_3, v_4) - r_2$ are in $I(\{x\}; \mathrm{Sym}^2 \oplus \mathbb{R}^L)$. Moreover, $\mathscr{R}_4(v_1, v_2, v_3, v_4) - r_2$ is one degree smoother as a conormal distribution than $\mathscr{I}_4(v_1, v_2, v_3, v_4) - r_1$. Thus

$$(70) \qquad\qquad \sigma[\Box_{\widehat{g}, \widehat{\phi}} v_{1234}](x, \xi) = \sigma[\mathscr{I}_4(v_1, v_2, v_3, v_4)](x, \xi).$$

The principal symbol at (x, ξ) can be viewed as the map

$$(71) \qquad\quad \mathscr{S}(\boldsymbol{b}, \boldsymbol{s}) = \sigma[\mathscr{I}_4(v_1, v_2, v_3, v_4)](x, \xi), \quad (\boldsymbol{b}, \boldsymbol{s}) \in X_1,$$

where $\boldsymbol{s} = (s_1, s_2, s_3, s_4)$, $s_j = \sigma[v_j](x, \xi_j)$, v_j is as in (48), ξ_j is as in (69), and

$$X_1 = \{(\boldsymbol{b}, s_1, s_2, s_3, s_4); \; s_j \in \mathbb{S}_j(\boldsymbol{b}), \; \boldsymbol{b} \in X_0\},$$
$$X_0 = \{\boldsymbol{b} \in (L_x^+ M)^4; \; \boldsymbol{b} \text{ is a basis of } \mathbb{R}^4 \text{ and}$$
$$\xi_j = \xi_j(\boldsymbol{b}), \, j = 1, 2, 3, 4, \text{ satisfy } (62)\}.$$

Note that transversality of $K_1, \ldots, K_4$ implies that $\boldsymbol{b}$ is a basis. Conversely, if $\boldsymbol{b} \in (L_x^+ M)^4$ is a basis then there are transversal Λ-invariant submanifolds of codimension 1, such that (68) holds. An explicit construction is given in the proof of Theorem 5.6 below.

Proposition 4.10. *Let* $(x, \xi) \in \Sigma$. *There is an open and dense subset* $\mathcal{U}$ *of* X_1^6 *such that for all* $(\boldsymbol{b}^n, \boldsymbol{s}^n)_{n=1}^6 \in \mathcal{U}$ *the span of* $\mathscr{S}(\boldsymbol{b}^n, \boldsymbol{s}^n)$, $n = 1, \ldots, 6$, *is a* 6-*dimensional subspace of* $\mathrm{Sym}_x^2 \subset (\mathrm{Sym}^2 \oplus \mathbb{R}^L)_x$.

The proof of Proposition 4.10 is divided in two steps. We will first show that there is a point $(\boldsymbol{b}^n, \boldsymbol{s}^n)_{n=1}^6 \in X_1^6$ such that the span of $\mathscr{S}(\boldsymbol{b}^n, \boldsymbol{s}^n)$, $n = 1, \ldots, 6$, is 6-dimensional. Then we use an argument based on the fact that $\mathscr{S}$ can be viewed as a meromorphic function.

4.4.1. *A choice of incoming directions and principal symbols.* We use normal coordinates at x with respect to $\widehat{g}$ in the computations below. In these coordinates, $\widehat{g}$ is the Minkowski metric and $\widehat{\Gamma}^r_{pq} = 0$ at x, see e.g. [67, Prop. 33, p. 73]. Moreover, we choose the coordinates so that $\xi = (1, 1, 0, 0)$.

We begin by choosing the symbols $\boldsymbol{s}^n = (s_1, \dots, s_4)$ so that they do not depend on n, and write

$$s_j = (g_{(j)}, \phi_{(j)}), \quad j = 1, 2, 3, 4,$$

where $s_j = \sigma[v_j](x, \xi_j)$, v_j is as in (48), and the parenthesis are to emphasize that the indices are not tensor indices. We will use an analogous notation for the higher order derivatives in ϵ, that is,

$$\sigma[v_{jk}](x, \eta_{jk}) = (g_{(jk)}, \phi_{(jk)}), \quad \sigma[v_{jkl}](x, \eta_{jkl}) = (g_{(jkl)}, \phi_{(jkl)}),$$

where $\eta_{jk} = \xi_j + \xi_k$ and $\eta_{jkl} = \xi_j + \xi_k + \xi_l$. Here v_{jk} and v_{jkl} are defined by (51), and we emphasize again that the indices are not tensorial.

We choose $g_{(j)} = 0$ for all $j = 1, 2, 3, 4$, and

$$\phi_{(1)} = \phi_{(2)} = (1, 0, \dots, 0) \in \mathbb{R}^L, \quad \phi_{(3)} = \phi_{(4)} = (0, 1, 0, \dots, 0) \in \mathbb{R}^L.$$

Observe that then $s_j \in \mathbb{S}_j(\boldsymbol{b})$ for any $\boldsymbol{b}$. Let us use the shorthand notations

$$q(\eta) = 1/\widehat{g}(\eta, \eta), \quad \xi \widehat{\otimes} \eta = \xi \otimes \eta + \eta \otimes \xi.$$

Differentiating (18) with respect to ϵ_j and ϵ_k, $j \neq k$, and evaluating the principal symbol at $\epsilon = 0$, we see that when $x \notin \operatorname{supp}(F)$,

$$g_{(jk)} = -2q(\eta_{jk}) \xi_j \widehat{\otimes} \xi_k, \quad jk = 12, 34.$$

Moreover, $g_{(jk)} = 0$ when $jk \neq 12, 34$, and $\phi_{(jk)} = 0$ for all pairs jk.

We write

$$\Gamma_{\alpha\gamma\beta}(\eta)g = \frac{1}{2}(\eta_\alpha g_{\gamma\beta} + \eta_\beta g_{\alpha\gamma} - \eta_\gamma g_{\alpha\beta}), \quad \Gamma_{\alpha\gamma\beta} \, g_{(jk)} = \Gamma_{\alpha\gamma\beta}(\eta_{jk}) \, g_{(jk)},$$

and define also

$$p(g_{(jk)})\phi_l = \widehat{g}g_{(jk)}\widehat{g}(\eta_l, \eta_l)\phi_l, \quad p(g_{(jk)})g_{(lm)} = \widehat{g}g_{(jk)}\widehat{g}(\eta_{lm}, \eta_{lm})g_{(lm)}.$$

Recalling (20), we differentiate (18) with respect to ϵ_j, ϵ_k, ϵ_l, where j, k and l are distinct, and evaluate the principal symbol at $\epsilon = 0$. This yields for $x \notin \operatorname{supp}(F)$,

$$\phi_{(12j)} = q(\eta_{12j})p(g_{(12)})\phi_{(j)}, \quad \phi_{(34k)} = q(\eta_{34k})p(g_{(34)})\phi_{(k)},$$

where $j = 3, 4$, $k = 1, 2$. Moreover, $\phi_{(jkl)} = 0$ for the other triplets jkl, and $g_{(jkl)} = 0$ for all triplets.

Analogously, we see that the scalar field components of $\sigma[\mathscr{I}_4(v_1, v_2, v_3, v_4)](x, \xi)$ vanish and the metric tensor part is

$$\begin{aligned}
(72) \quad & p(g_{(12)})g_{(34)} + p(g_{(34)})g_{(12)} - 2\eta_{123}\widehat{\otimes}\xi_4 - 2\eta_{124}\widehat{\otimes}\xi_3 - 2\eta_{341}\widehat{\otimes}\xi_2 - 2\eta_{342}\widehat{\otimes}\xi_1 \\
& + 2\widehat{g}^{pq}\widehat{g}^{rs}(\Gamma_{prj}g_{(12)} \, \Gamma_{qsk}g_{(34)} + \Gamma_{prj}g_{(12)} \, \Gamma_{qks}g_{(34)} + \Gamma_{prk}g_{(12)} \, \Gamma_{qjs}g_{(34)}) \\
& + 2\widehat{g}^{pq}\widehat{g}^{rs}(\Gamma_{prj}g_{(34)} \, \Gamma_{qsk}g_{(12)} + \Gamma_{prj}g_{(34)} \, \Gamma_{qks}g_{(12)} + \Gamma_{prk}g_{(34)} \, \Gamma_{qjs}g_{(12)}).
\end{aligned}$$

Let us now choose the incoming directions $\boldsymbol{b}^n = (b_1^n, \ldots, b_4^n)$. We will choose each b_j^n to be in the following set of six Pythagorean quadruples

$$B = \{(p^2 + m^2 + n^2, 2mp, 2np, p^2 - (m^2 + n^2));\ m = 1, 2, 3,\ n = 1, 2,\ p = 1\}.$$

Then we consider the following choices

$$\mathbb{B} = \{\boldsymbol{b} = (b_1, b_2, b_3, b_4);\ b_j \in B,\ \boldsymbol{b} \in X_0\}.$$

The set $\mathbb{B}$ contains 9 elements, up to reindexing. Indeed, out of the 15 subsets of B with 4 elements, 6 fail to be bases or satisfy (62). A straightforward computation shows that the principal symbols (72) corresponding to the incoming directions in $\mathbb{B}$ span a 6-dimensional subspace of the fibre Sym_x^2. We have verified this using a computer algebra software. All the computations were performed using rational arithmetic. Thus there is a point $(\boldsymbol{b}^n, \boldsymbol{s}^n)_{n=1}^6 \in X_1^6$ such that the span of $\mathscr{S}(\boldsymbol{b}^n, \boldsymbol{s}^n)$, $n = 1, \ldots, 6$, is 6-dimensional and contained in Sym_x^2. This completes the first step of the proof of Proposition 4.10.

4.4.2. *Changing incoming directions.* We choose local coordinates and identify $T_x^* M$ with $\mathbb{R}^4$. Consider $\boldsymbol{b} \in X_0$ and write

(73) $$\xi = \xi_1 + \xi_2 + \xi_3 + \xi_4, \quad \xi_j = r_j(\boldsymbol{b}) b_j.$$

As the coefficients $r_j(\boldsymbol{b})$ can be obtained by solving a system of linear equations (for example, using Cramer's rule), we have that

$$r_j(\boldsymbol{b}) = \frac{p_j(\boldsymbol{b})}{q(\boldsymbol{b})}, \quad j = 1, 2, 3, 4,$$

where p_j and q are polynomials in $\boldsymbol{b} \in (\mathbb{R}^4)^4$ that do not vanish at any point on X_0. We can also consider p_j, $j = 1, 2, 3, 4$, and q as functions on $\widetilde{X}_0 = (L_+ M)^4$. Here $\widetilde{X}_0$ is a connected, real-analytic manifold, and p_j and q are real-analytic functions on $\widetilde{X}_0$. Observe also that the equation (32) is linear in h and of full rank when $\xi \in L_x^+ M$, and therefore the manifold

$$\widetilde{X}_1 = \{(\boldsymbol{b}, s_1, s_2, s_3, s_4);\ s_j \in \mathbb{S}_j(\boldsymbol{b}),\ \boldsymbol{b} \in \widetilde{X}_0\}$$

is real-analytic, see e.g. [48, Prop. 1.9.2].

The expression (56) is of the form $P_0(\boldsymbol{\xi}, \boldsymbol{s}) / Q_0(\boldsymbol{\xi})$ where P_0 and Q_0 are polynomials in $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3, \xi_4) \in (\mathbb{R}^4)^4$. Here P_0 takes values on $(\mathrm{Sym}^2 \oplus \mathbb{R}^L)_x = \mathbb{R}^{10+L}$ and Q_0 on $\mathbb{R}$. Moreover, for $\boldsymbol{\xi}$ as in (73), with $\boldsymbol{b} \in X_0$, it holds that Q_0 is non-zero. It follows that

$$\mathscr{S}(\boldsymbol{b}, \boldsymbol{s}) = \frac{P_1(\boldsymbol{b}, \boldsymbol{s})}{Q_1(\boldsymbol{b})},$$

where P_1 and Q_1 are real-analytic on $\widetilde{X}_1$ and $\widetilde{X}_0$, respectively, and Q_1 is nowhere vanishing on $X_0 \subset \widetilde{X}_0$.

In the previous section, we showed that there is $(\boldsymbol{b}^n, \boldsymbol{s}^n)_{n=1}^6 \in X_1^6$ such that the span of $\mathscr{S}(\boldsymbol{b}^n, \boldsymbol{s}^n)$, $n = 1, \ldots, 6$, is 6-dimensional subspace of the fibre $(\mathrm{Sym}^2 \oplus \mathbb{R}^L)_x$. We set $e_n = \mathscr{S}(\boldsymbol{b}^n, \boldsymbol{s}^n)$, $n = 1, \ldots, 6$, and choose

$$e_n \in (\mathrm{Sym}^2 \oplus \mathbb{R}^L)_x, \quad n = 7, \ldots, 10 + L,$$

such that the vectors e_n, $n = 1, \ldots, 10 + L$ form a basis of $(\mathrm{Sym}^2 \oplus \mathbb{R}^L)_x$. Let $[e_1, \ldots, e_{L+10}]$ be the matrix with the columns e_n, and consider the map

$$(\boldsymbol{b}^n, \boldsymbol{s}^n)_{n=1}^6 \mapsto \det([\mathscr{S}(\boldsymbol{b}^1, \boldsymbol{s}^1), \ldots, \mathscr{S}(\boldsymbol{b}^6, \boldsymbol{s}^6), e_7, \ldots, e_{L+10}]).$$

This can be written as

$$\frac{\det([P_1(\boldsymbol{b}^1, \boldsymbol{s}^1), \ldots, P_1(\boldsymbol{b}^6, \boldsymbol{s}^6), e_7, \ldots, e_{L+10}])}{\prod_{n=1}^6 Q_1(\boldsymbol{b}^n)} = \frac{P_2((\boldsymbol{b}^n, \boldsymbol{s}^n)_{n=1}^6)}{Q_2((\boldsymbol{b}^n, \boldsymbol{s}^n)_{n=1}^6)}$$

where P_2 and Q_2 are real-analytic functions on the connected, real-analytic manifold $\widetilde{X}_1^6$ that do not vanish at the point $(\boldsymbol{b}^n, \boldsymbol{s}^n)_{n=1}^6$ considered in the previous section. Thus the unique continuation principle for real-analytic functions on real-analytic manifolds, see e.g. [33, Lem. VI.4.3], implies that there is an open and dense set $\mathcal{U} \subset \widetilde{X}_1^6$ such that both Q_2 and P_2 are nowhere vanishing on $\mathcal{U}$. This finishes to proof of Proposition 4.10. □

5. Construction of light observation sets

In this section we prove the main theorem formulated in Section 1. We begin by relating the singular solutions constructed in Section 3 with the data set (8), and by showing that singularities due to non-linear interactions of singular solutions, as studied in Section 4, can be detected from (8). Then we relate the detected singularities with geometric objects on $(M, \widehat{g})$, the earliest light observation sets. This allows us to reduce the reconstruction of the Lorentzian structure to a purely geometric problem solved in [49].

5.1. **Sending singularities.** Let us show how to use the data set (8) to find families of sources parametrized by $\epsilon \in \mathcal{E}$ as in (47). Recall that (8) gives sources $\mathcal{F}$ in the Fermi coordinates Φ_g that depend on the source itself. For this reason, we need to construct first common coordinates in order to parametrize the sources smoothly.

We will use the wave maps. For a metric tensor g close to $\widehat{g}$, the wave map operator is defined in local coordinates by

$$\Box_{g,\widehat{g}} \Psi^p = g^{jk} \left(\frac{\partial}{\partial x^j} \frac{\partial}{\partial x^k} \Psi^p - \Gamma_{jk}^n(x) \frac{\partial}{\partial x^n} \Psi^p + \widehat{\Gamma}_{BC}^p(\Psi(x)) \frac{\partial}{\partial x^j} \Psi^B \frac{\partial}{\partial x^k} \Psi^C \right),$$

where $\widehat{\Gamma}_{BC}^p$ and Γ_{kl}^j are the Christoffel symbols of the metrics $\widehat{g}$ and g, respectively. This operator invariant under diffeomorphisms in the following sense: if $\Box_{g,\widehat{g}} \Psi = 0$ and if ψ and μ are diffeomorphisms, then setting $\widetilde{\Psi} = \mu \circ \Psi \circ \psi$, it holds that $\Box_{\psi^* g, \mu_* \widehat{g}} \widetilde{\Psi} = 0$, where $\mu_* = (\mu^{-1})^*$, the pushforward by μ, see e.g. [15].

Lemma 5.1. *Let $(\widehat{g}, \widehat{\phi})$ be as in Theorem 1.1 and write $\widehat{g}' = \Phi_{\widehat{g}}^* \widehat{g}$. Let $p \in V$, let $B \subset V$ be a neighbourhood of p satisfying $\overline{B} \subset V$, and let $q \in V$ satisfy $J_{(V,\widehat{g}')}(p, q) \subset B$. If $k \in \mathbb{N}$ is large and $r > 0$ is small, then for all $(g', \phi', \mathcal{F}')$ in $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$ satisfying $\operatorname{supp}(\mathcal{F}') \subset J_{(V,g')}^+(p)$, the equation*

$$(74) \qquad\qquad \Box_{g',\widehat{g}'} \Psi = 0 \quad \text{in } J_{(V,g')}(p, q),$$
$$\Psi = \Phi_{\widehat{g}}^{-1} \circ \Phi_g, \quad \text{in } B \setminus J_{(V,g')}^+(p),$$

has a unique solution that coincides with $\Phi_{\widehat{g}}^{-1} \circ \Phi_g$ on $J_{(V,\widehat{g}')}(p, q)$. Here g is a metric that satisfies

  (G1) $(g', \phi', \mathcal{F}') = \Phi_g^*(g, \phi, \mathcal{F})$,
  (G2) $(g, \phi, \mathcal{F})$ *satisfies (1)–(2) on $\widehat{M}$, and (3) holds*,
  (G3) $\operatorname{Ric}_{\widehat{g}}(g) = \operatorname{Ric}(g)$.

Moreover, for small enough B, the equation (74) can be solved given the data set $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$.

Proof. As $(g', \phi', \mathcal{F}') \in \mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$ there is $(\widetilde{g}, \widetilde{\phi}, \widetilde{\mathcal{F}})$ on $\widehat{M}$ such that (G1) and (G2) hold for this triple. Let $\widetilde{T} > \widehat{T}$ and write $\widetilde{M} = (0, \widetilde{T}) \times N$. For small r and large k, $(\widetilde{g}, \widetilde{\phi}, \widetilde{\mathcal{F}})$ can extended so that it solves the equation (1)–(2) on $\widetilde{M}$ with $\widetilde{\mathcal{F}}$ vanishing on $(\widehat{T}, \widetilde{T}) \times N$. Then, for small r and large k, the solution $\widetilde{\Psi}$ of

$$(75) \qquad\qquad \Box_{\widetilde{g},\widehat{g}} \widetilde{\Psi} = 0 \quad \text{in } \widetilde{M},$$
$$\widetilde{\Psi}(x) = x, \quad \text{in } (-\infty, 0) \times N,$$

is a diffeomorphism and $\widehat{M} \subset \widetilde{\Psi}(\widetilde{M})$. Moreover, the reduced Ricci tensor of $g = \widetilde{\Psi}_* \widetilde{g}$ coincides with the actual Ricci tensor in the sense of (G3), see e.g. [16, Sec. VI.7.2 and Th. 4.2 in Appendix III]. Setting $(g, \phi, \mathcal{F}) = \widetilde{\Psi}_*(\widetilde{g}, \widetilde{\phi}, \widetilde{\mathcal{F}})$, we have that $(g, \phi, \mathcal{F})$ satisfies (G1)-(G3). Note that (G1) holds since the Fermi coordinates coincide for isometric manifolds satisfying the initial condition (3). In other words, the Fermi coordinates have the naturality property described the following commuting diagram:

Properties (G2) and (G3) imply that

$$(76) \qquad\qquad \Box_{g,\widehat{g}} \operatorname{Id} = 0, \quad \text{in } \widehat{M},$$

where $\operatorname{Id}(x) = x$. Indeed, in local coordinates,

$$\Box_{g,\widehat{g}} \operatorname{Id}^p = g^{jk}(-\Gamma_{jk}^n \delta_n^p + \widehat{\Gamma}_{BC}^p \delta_j^B \delta_k^C) = -g^{pq}(\widehat{H}_q - \Gamma_q),$$

where $\widehat{H}_q - \Gamma_q$ are the components of the tensor $H_{\widehat{g}}(g)$, see (17). Moreover, the equality $\mathrm{Ric}_{\widehat{g}}(g) = \mathrm{Ric}(g)$ implies that

$$\mathrm{div}_g \, \mathrm{Ein}_{\widehat{g}}(g) = \mathrm{div}_g \, \mathrm{Ein}(g) = 0,$$

where the reduced Einstein tensor is defined by $\mathrm{Ein}_{\widehat{g}}(g) = I_g \, \mathrm{Ric}_{\widehat{g}}(g)$. As $(g, \phi, \mathcal{F})$ satisfies the initial condition (3) we may apply Lemma 5.7, when $k \in \mathbb{N}$ is large and $r > 0$ is small, to conclude that $H_{\widehat{g}}(g) = 0$.

Now (76), together with the diffeomorphism invariance of the wave map operator, imply that $\square_{g',\widehat{g}'}(\Phi_{\widehat{g}}^{-1} \circ \Phi_g) = 0$. Hence $\Phi_{\widehat{g}}^{-1} \circ \Phi_g$ is the solution of (74).

Let us now show how to solve (74) given $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$. We begin by showing that $\widehat{g}'$ is determined by $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$, which again follows from the fact that $\mathcal{F}' = 0$ implies $g' = \widehat{g}'$. Indeed, in this case also $\mathcal{F} = 0$, and (G2) and (G3) imply that $(g, \phi, 0)$ satisfies (18). But this quasilinear wave equation has a unique solution for small r and large k, and therefore $g = \widehat{g}$. Thus we find $\widehat{g}'$ as the first component the unique triple in $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$ with the third component vanishing.

As $\widehat{g}'$ is known, we can evaluate the wave map operator $\square_{g',\widehat{g}'} \Psi$ for any g' on $J_{(V,g')}(p, q)$ and any Ψ close to the identity. It remains to show that the initial condition in (74) is determined by $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$. This follows from the identity

$$(77) \qquad \exp_p^{\widehat{g}'} \circ (\exp_p^{g'})^{-1} = \Phi_{\widehat{g}}^{-1} \circ \Phi_g, \quad \text{in } B \setminus J_{(V,g')}^+(p),$$

that we will show next. Here $\exp^{g'}$ denotes the exponential map with respect to the metric g'.

The assumption $\mathrm{supp}\,(\mathcal{F}') \subset J_{(V,g')}^+(p)$ implies that $g = \widehat{g}$ on $\widehat{M} \setminus J_{(M,\widehat{g})}^+(\Phi(p))$, using again the the fact that (18) has a unique solution. Here we used the notation $\Phi = \Phi_{\widehat{g}}$. In particular, it holds that $\Phi(p) = \Phi_g(p)$, $d\Phi = d\Phi_g$ at p, and

$$\mathrm{Id} = \exp_{\Phi(p)}^{\widehat{g}} \circ (\exp_{\Phi(p)}^g)^{-1} \quad \text{in } U_1 \setminus J_{(M,\widehat{g})}^+(\Phi(p)),$$

for a suitable neighbourhood $U_1 \subset \widehat{M}$ of $\Phi(p)$. Equation (77) is obtained by combining the previous identity with the naturality of the exponential map, that is, with the fact that the following two diagrams commute:

$$
\begin{array}{ccc}
U_2 & \xrightarrow{\;d\Phi\;} & U_3 \\
{\scriptstyle \exp_p^{\widehat{g}'}} \downarrow & & \downarrow {\scriptstyle \exp_{\Phi(p)}^{\widehat{g}}} \\
(V, \widehat{g}') & \xrightarrow{\;\Phi_{\widehat{g}}\;} & (\widehat{M}, \widehat{g})
\end{array}
\qquad
\begin{array}{ccc}
U_2 & \xrightarrow{\;d\Phi\;} & U_3 \\
{\scriptstyle \exp_p^{g'}} \downarrow & & \downarrow {\scriptstyle \exp_{\Phi(p)}^{g}} \\
(V, g') & \xrightarrow{\;\Phi_g\;} & (\widehat{M}, g)
\end{array}
$$

where $U_2 \subset T_p V$ and $U_3 \subset T_{\Phi(p)}\widehat{M}$ are suitable neighbourhoods of the respective origins. Indeed, in $B \setminus J_{(V,g')}^+(p)$

$$\exp_p^{\widehat{g}'} \circ (\exp_p^{g'})^{-1} = \Phi_{\widehat{g}}^{-1} \circ \exp_{\Phi(p)}^{\widehat{g}} \circ d\Phi \circ d\Phi^{-1} \circ (\exp_{\Phi(p)}^g)^{-1} \circ \Phi_g = \Phi_{\widehat{g}}^{-1} \circ \Phi_g.$$

$\square$

5.2. **Receiving singularities.** The only difficulty in receiving of singularities is that the Fermi coordinates Φ_g can be non-smooth when g has singularities. This problem is solved by using the following two lemmas: the first one shows that at least some singularities of the fourfold linearization of g are visible when the principal symbol of the fourfold linearization is non-vanishing in a large enough subbundle and suitable coordinates are used, and the second one shows that changing from the Fermi coordinates to normal coordinates gives suitable coordinates.

In the first lemma, we denote by $\mathrm{Sym}^2(K)$ the rank 6 subbundle of Sym^2 over a submanifold $K \subset M$ of codimension one, that consists of the symmetric 2-tensors on K. Later, we will use this lemma when K is a Λ-invariant submanifold, in fact, in the proof of Theorem 5.6 below, the submanifold K will be contained in the image of $L_x^+ M$ under the exponential map. Here x is as in Theorem 5.6, and there are no cut points on the geodesic segment joining x and y, with y as in the lemma.

Lemma 5.2. *Consider a family* $g \in C^4(\mathcal{E}; C^2(M; \mathrm{Sym}^2))$ *where* $\mathcal{E}$ *is a neighbourhood of the origin in* $\mathbb{R}^4$, *and let* $K \subset M$ *be a submanifold of codimension one. Let* $y \in K$ *and let* $U \subset M$ *be a small neighbourhood of* y. *Suppose that for* $\alpha = (1,1,1,1)$ *it holds that*

$$(78) \qquad \partial_\epsilon^\alpha g|_{\epsilon=0} \in I(K; \mathrm{Sym}^2),$$

$$(79) \qquad \partial_\epsilon^\beta g|_{\epsilon=0} \in C^\infty(U; \mathrm{Sym}^2), \quad \textit{for all } \beta < \alpha.$$

Here the $\beta < \alpha$ *means the partial order in the sense of multi-indices. Suppose, moreover, that the restriction of* $\sigma[\partial_\epsilon^\alpha g|_{\epsilon=0}](y, \cdot)$ *on* $\mathrm{Sym}^2(K)$ *does not vanish. Let* $V \subset \mathbb{R}^4$ *be open and let* $\Psi \in C^4(\mathcal{E}; C^2(V; U))$. *Suppose that* $\Psi(\epsilon)$ *gives a system of coordinates near* y *and that*

$$(80) \qquad \partial_\epsilon^\beta \Psi^* g|_{\epsilon=0} \in C^\infty(V; \mathrm{Sym}^2), \quad \textit{for all } \beta < \alpha.$$

Then $\partial_\epsilon^\alpha \Psi^* g|_{\epsilon=0}$ *is not smooth at* $\Psi^{-1}(y)$. *Here* $(\Psi^* g)(\epsilon, \cdot) = \Psi(\epsilon, \cdot)^* g(\epsilon, \cdot)$.

Proof. For small enough U, we can choose such smooth coordinates (x^0, x^1, x^2, x^3) that $K \cap U = \{x \in U; x^0 = 0\}$. Then the curvature $(0,4)$-tensor can be expressed in the coordinates as follows

$$R_{ik\ell m} = \frac{1}{2}\left(\frac{\partial^2 g_{im}}{\partial x^k \partial x^\ell} + \frac{\partial^2 g_{k\ell}}{\partial x^i \partial x^m} - \frac{\partial^2 g_{i\ell}}{\partial x^k \partial x^m} - \frac{\partial^2 g_{km}}{\partial x^i \partial x^\ell}\right)$$
$$+ g_{np}\left(\Gamma^n_{\ k\ell}\Gamma^p_{\ im} - \Gamma^n_{\ km}\Gamma^p_{\ i\ell}\right),$$

where $\Gamma^m_{\ ij}$ are the Christoffel symbols of g. We write $g^{(\alpha)} = \partial_\epsilon^\alpha g|_{\epsilon=0}$. As $g^{(\alpha)} \in I^p(K; \mathrm{Sym}^2)$ and as its principal symbol does not vanish, there is the maximal smoothness index $s = -p - 1$ such that $g^{(\alpha)}$ is in the Besov type space $^\infty H_{(s)}(U; \mathrm{Sym}^2)$, see [38, Def. 18.2.6]. Moreover, it holds for $i, m = 1, 2, 3$ that

$$\frac{\partial^2 g_{k\ell}^{(\alpha)}}{\partial x^i \partial x^m} \in {}^\infty H_{(s)}(U), \qquad \frac{\partial^2 g_{i\ell}^{(\alpha)}}{\partial x^k \partial x^m} \in {}^\infty H_{(s-1)}(U), \qquad \frac{\partial^2 g_{km}^{(\alpha)}}{\partial x^i \partial x^\ell} \in {}^\infty H_{(s-1)}(U).$$

Thus, writing $R^{(\alpha)} = \partial_\epsilon^\alpha R|_{\epsilon=0}$,

$$R_{i00m}^{(\alpha)} = \frac{1}{2}\frac{\partial^2 g_{im}^{(\alpha)}}{\partial x^0 \partial x^0} + r, \quad i, m = 1, 2, 3,$$

where $r \in {}^\infty H_{(s-1)}(U)$. We used here the fact that all the terms in $R^{(\alpha)}$, where the four derivatives with respect to ϵ do not act on the same instance of g, are smooth. This follows from (79).

Consider a family of curves $\gamma \in C^4(\mathcal{E}; C^1(-1, 1; U))$ and suppose that $\gamma(\epsilon)$ is a geodesic with respect to $g(\epsilon)$. Write $\widehat{\gamma} = \gamma|_{\epsilon=0}$ and suppose, moreover, that $\widehat{\gamma}$ intersects K transversally at $y = \widehat{\gamma}(0)$. We can consider $g_{im}^{(\alpha)} \circ \widehat{\gamma}$ as a conormal distribution associated to $\{0\} \subset (-1, 1)$. As $\sigma[g^{(\alpha)}](\widehat{\gamma}(0), \cdot)$ does not vanish on $\mathrm{Sym}^2(K)$, there are $i, m = 1, 2, 3$ such that

$$(81) \qquad \sigma[R_{i00m}^{(\alpha)} \circ \widehat{\gamma}](0, \tau) = -\frac{1}{2}\xi_0^2\, \sigma[g_{im}^{(\alpha)} \circ \widehat{\gamma}](0, \tau),$$

does not vanish for some $\tau \in \mathbb{R} \setminus 0$.

By differentiating the geodesic equation $\nabla_{\dot\gamma}\dot\gamma = 0$ with respect to ϵ we see that the path $\gamma_{(\alpha)} = \partial_\epsilon^\alpha \gamma|_{\epsilon=0}$ satisfies

$$(82) \quad \partial_t \dot\gamma_{(\alpha)}^k + (\widehat{\Gamma}_{ij}^k \circ \widehat{\gamma})\dot{\widehat{\gamma}}^j \dot\gamma_{(\alpha)}^i + (\widehat{\Gamma}_{ij}^k \circ \widehat{\gamma})\dot\gamma_{(\alpha)}^j \dot{\widehat{\gamma}}^i + \gamma_{(\alpha)}^\ell(\partial_{x^\ell}\widehat{\Gamma}_{ij}^k \circ \widehat{\gamma})\dot{\widehat{\gamma}}^j \dot{\widehat{\gamma}}^i = h^k + r^k,$$

where $k = 0, 1, 2, 4$, $h^k = -(\partial_\epsilon^\alpha \Gamma_{ij}^k|_{\epsilon=0} \circ \widehat{\gamma})\dot{\widehat{\gamma}}^j\dot{\widehat{\gamma}}^i$ and r^k is smooth due to (79). This is a linear system of differential equations for $\gamma_{(\alpha)}$ with smooth coefficients. There is a maximal smoothness index $s' \in \mathbb{R}$ such that

$$g_{jk}^{(\alpha)} \circ \widehat{\gamma} \in {}^\infty H_{(s')}(-1, 1), \quad h^k \in {}^\infty H_{(s'-1)}(-1, 1).$$

Thus (82) implies that $\gamma_{(\alpha)}^k, \dot\gamma_{(\alpha)}^k \in {}^\infty H_{(s'-1)}(-1, 1)$ for $k = 0, 1, 2, 3$.

As the intersection of $\widehat{\gamma}$ and K is transversal, there are C^4-smooth maps $t(\epsilon)$ and $y(\epsilon)$, mapping a neighbourhood of the origin in $\mathcal{E}$ to small neighbourhoods of the origin in $\mathbb{R}$ and y in M, respectively, such that $\gamma(\epsilon, t(\epsilon)) = y(\epsilon) \in K$. Let $Z(\epsilon, t), V(\epsilon, t)$ and $W(\epsilon, t)$ be the parallel transports of $\partial_{x^0}, \partial_{x^i}, \partial_{x^m} \in T_{y(\epsilon)}M$ along $\gamma(\epsilon)$ with respect to $g(\epsilon)$. The vector field $Z_{(\alpha)} = \partial_\epsilon^\alpha Z|_{\epsilon=0}$ satisfies an equation similar to (82), and analogously with the above, we see that its components satisfy $Z_{(\alpha)}^k \in {}^\infty H_{(s'-1)}(-1, 1)$. The same is true for V and W. It follows, using again (79), that the principal symbol of $\partial_\epsilon^\alpha \langle R(Z, V)Z, W \rangle|_{\epsilon=0}$ is (81). As the principal symbol does not vanish, we see that the function

$$(83) \qquad \partial_\epsilon^\alpha \langle R(Z, V)Z, W \rangle|_{\epsilon=0}$$

is not in $C^\infty(-t_0, t_0)$ for any $t_0 > 0$.

To get a contradiction suppose that there is a neighbourhood V' of $\Psi^{-1}(y)$ such that $\Psi^* \partial_\epsilon^\alpha g|_{\epsilon=0}$ is smooth in V'. Now (80) together with the smoothness of $\Psi^* \partial_\epsilon^\alpha g|_{\epsilon=0}$ implies that the function (83) is smooth at $t = 0$. But this is a contradiction. $\qquad\square$

In the proof of Theorem 5.6 we will apply the above lemma in a setting where Ψ is given as the composition of the Fermi coordinates Φ_g and certain normal coordinates. The next lemma captures the essential features of the setting. We will apply it in the case $\Phi(\epsilon, y) = \Phi_{g(\epsilon)}(y)$, $y \in V$.

Lemma 5.3. *Let $g \in C^4(\mathcal{E}; C^3(M; \mathrm{Sym}^2))$, with $\mathcal{E}$ a neighbourhood of the origin in $\mathbb{R}^4$, let $\Phi \in C^4(\mathcal{E}; C^2(V; M))$, and suppose that $\Phi(\epsilon, \cdot)$ gives local coordinates on M. Write $g' = \Phi^* g$, that is, $(\Phi^* g)(\epsilon, \cdot) = \Phi(\epsilon, \cdot)^* g(\epsilon, \cdot)$. Let $y \in V$, let X_j, $j = 0, 1, 2, 3$, be a basis of $T_y V$, and consider the normal coordinates*

$$(84) \qquad \Xi(\epsilon, z) = \exp_y^{g'(\epsilon)} \left(\sum_{j=0}^3 z^j X_j \right).$$

Define $\Psi(\epsilon, z) = \Phi(\epsilon, \Xi(\epsilon, z))$. Then there is a neighbourhood V' of y such that $\Psi \in C^4(\mathcal{E}; C^2(V'; U))$. Furthermore, if g satisfies (79) then $\Psi^* g$ satisfies (80) with $V = V'$.

Proof. Naturality of the exponential map implies that

$$\Psi(\epsilon, z) = \exp_{\Phi(\epsilon, y)}^{g(\epsilon)} \left(\sum_{j=0}^3 z^j (\partial_{y^j} \Phi^k)(\epsilon, y) X_k \right).$$

For small enough neighbourhood V' of y, the smoothness $\Psi \in C^4(\mathcal{E}; C^2(V'; U))$ follows from $g \in C^4(\mathcal{E}; C^3(M; \mathrm{Sym}^2))$. As the exponential map is obtained by a solving system of ordinary differential equations, while varying the initial conditions, the assumption (79) implies that $\partial_\epsilon^\beta \Psi|_{\epsilon=0} \in C^\infty(V'; \mathrm{Sym}^2)$ for any $\beta < \alpha$ where $\alpha = (1, 1, 1, 1)$. Thus (80) holds with $V' = V$. $\square$

5.3. **Concepts from global Lorentzian geometry.** Up to this point the global geometry of $(M, \widehat{g})$ has played a minor role. When we combine sending, interaction and receiving of singularities in the next section, we need to take into account certain global aspects.

Let $p, q \in M$. We write $p < q$ when p and q can be joined by a future pointing causal curve in $(M, \widehat{g})$. When this is the case, we define the time separation function $\tau(p, q) \in [0, \infty)$ to be the supremum of the lengths

$$L(\alpha) = \int_0^1 \sqrt{-\widehat{g}(\dot{\alpha}(s), \dot{\alpha}(s))} \, ds$$

of the piecewise smooth causal curves $\alpha : [0, 1] \to M$ from p to q. If the condition $p < q$ does not hold, we define $\tau(p, q) = 0$.

When $(x, \xi) \in \Sigma$, we define $\mathcal{T}(x, \xi)$ to be the length of the maximal interval on which $\gamma_{x,\xi} : [0, \mathcal{T}(x, \xi)) \to M$ is defined. Recall that $\gamma_{x,\xi}$ is the geodesic of $(M, \widehat{g})$ with the initial data (x, ξ_*) where ξ_* is the vector $\xi_*^j = \widehat{g}^{jk} \xi_k$. When ξ is future pointing, we define the modified cut locus functions, c.f. [7, Def. 9.32],

$$(85) \qquad \rho(x, \xi) = \sup\{s \in [0, \mathcal{T}(x, \xi)); \ \tau(x, \gamma_{x,\xi}(s)) = 0\},$$
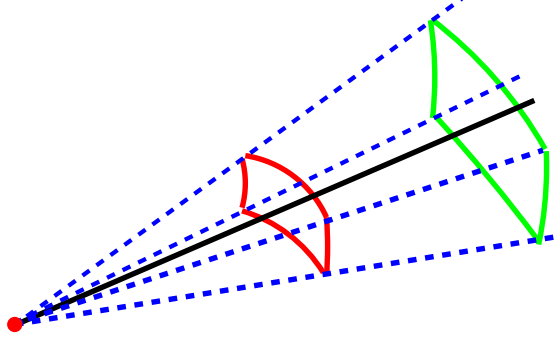
FIGURE 3. Schematic of $K(x, \xi; \delta)$ in the Minkowski space. The projection of $K(x, \xi; \delta)$ at a fixed time is a piece of a sphere. The spherical piece propagates near the lightlike geodesic $\gamma_{x,\xi}$ and is singular on a subset of the light cone emanating from x. The black line segment is the projection of $\gamma_{x,\xi}$ on $\mathbb{R}^3$. Two projections of $K(x, \xi; \delta)$ on $\mathbb{R}^3$, corresponding to two different times, are drawn in red and green.

Using [7, Th. 9.33], we see that the function $\rho(x, \xi)$ is lower semi-continuous on the globally hyperbolic Lorentzian manifold $(M, \widehat{g})$.

5.4. **Probing the background geometry with propagating singularities.** We will next combine all the steps studied above: sending, interaction and receiving of singularities. This is done in Theorem 5.6, that relates analytical information on propagating singularities to purely geometrical information on lightlike geodesics and the causal structure of $(M, \widehat{g})$. After establishing the relation between analytical and geometrical information, the final step of the proof of Theorem 1.1 is purely geometrical and coincides with the geometric part of the proof in [49].

Consider the Fermi coordinates $\Phi = \Phi_{\widehat{g}}$ defined by (6). Let $x \in V$ and $\xi \in L_x^+ V$, where the light cone is defined with respect to $\widehat{g}' = \Phi^* \widehat{g}$. Let $\delta > 0$, denote by $B_x(\xi, \delta) \subset T_x^* V$ the ball of radius δ with respect to a fixed auxiliary Riemannian metric on V, and define

(86) $$K(x, \xi; \delta) = \pi(\Lambda(B_x(\xi, \delta))) \setminus J_{(M,\widehat{g})}^-(x).$$

Here $\pi : T^* M \to M$ is the canonical projection to the base manifold, Λ is the flowout defined by (28), and we have extended the definition as follows

$$\Lambda(S) = \Lambda(\{(\Phi(x), \Phi_* \xi); \ (x, \xi) \in S\}), \quad S \subset T^* V.$$

Observe that $K(x, \xi; \delta)$ is a 3-dimensional Λ-invariant manifold away from the cut locus of x. Moreover, using the notation,

$$\gamma_{x,\xi} = \gamma_{\Phi(x), \Phi_* \xi}, \quad \rho(x, \xi) = \rho(\Phi(x), \Phi_* \xi), \quad (x, \xi) \in T^* V,$$

it contains the lightlike geodesic $\gamma_{x,\xi}((0, \rho(x, \xi)))$.

For $z_j \in V$, $\zeta_j \in L^+_{z_j} V$, $j = 1, 2, 3, 4$, we write $\boldsymbol{z} = (z_j)^4_{j=1}$, $\boldsymbol{\zeta} = (\zeta_j)^4_{j=1}$ and define

$$C^+(\boldsymbol{z}, \boldsymbol{\zeta}) = \bigcup_{j=1}^4 J^+_{(M,\widehat{g})}(\gamma_{z_j,\zeta_j}(\rho(z_j, \zeta_j))),$$

the causal future of the first cut points on the geodesics γ_{z_j,ζ_j}. Below, it is enough to consider the complement of $C^+(\boldsymbol{z}, \boldsymbol{\zeta})$, since the cut points are treated in the purely geometric step.

Denote by $\mathcal{K}(\boldsymbol{z}, \boldsymbol{\zeta}, \delta)$ the set $\mathcal{K}$ defined by formula (60) with

$$K_j = K(z_j, \zeta_j; \delta), \quad \delta > 0, \ j = 1, 2, 3, 4.$$

Recall that $\mathcal{K}(\boldsymbol{z}, \boldsymbol{\zeta}, \delta)$ is the union of the sets K_j together with a set giving an upper bound for the singular supports of the waves produced by the threefold interactions of waves with singular supports on the sets K_j. When $\delta \to 0$, the set $\mathcal{K}(\boldsymbol{z}, \boldsymbol{\zeta}, \delta)$ tends to the set

$$\mathcal{K}_0(\boldsymbol{z}, \boldsymbol{\zeta}) = \bigcap_{\delta > 0} \mathcal{K}(\boldsymbol{z}, \boldsymbol{\zeta}, \delta),$$

that is of Hausdorff dimension at most two. Below, we consider $\mathcal{K}_0(\boldsymbol{z}, \boldsymbol{\zeta})$ as a small exceptional set.

We define

$$\mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta}) = \bigcap_{j=1}^4 \gamma_{z_j,\zeta_j}((0, \rho(z_j, \zeta_j))).$$

Note that $\mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta}) \neq \emptyset$ if and only if all the four geodesics γ_{z_j,ζ_j} intersect before their first cut points. A piece of information that we want to recover given the data $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$ is whether $\mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta}) \neq \emptyset$ or not. Moreover, in the case $\mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta}) \neq \emptyset$, the main piece of information that we want to recover is if $x \in \mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta})$ can be joined to a point $y \in \Phi(V)$ by a lightlike geodesic segment without cut points. More precisely, we want to determine if $y \in \mathcal{E}_V(x)$ where

$$(87) \qquad \mathcal{E}_V(x) = V \cap \mathcal{E}(x), \quad \mathcal{E}(x) = \{\gamma_{x,\xi}(t); \ t \in [0, \rho(x, \xi)], \ \xi \in L^+_x M\}.$$

We denote by $B((\boldsymbol{z}, \boldsymbol{\zeta}), \delta)$, the ball of radius $\delta > 0$ with respect to a fixed auxiliary Riemannian metric on $(T^*V)^4$, and use also the shorthand notation

$$E = \text{Sym}^2 \oplus \mathbb{R}^L.$$

The following two definitions form the basis of a test that allows us to extract geometric information given the data (8).

Definition 5.4. *Let $p, q \in V$ and suppose that there is open $B \subset \mathbb{R}^4$ such that $J_{(V,\widehat{g}')}(p, q) \subset B$ and $\overline{B} \subset V$. Write $V^{in} = J_{(V,\widehat{g}')}(p, q)$ and define*

$$L_4 V^{in} = \{(z_1, \ldots, z_4, \zeta_1, \ldots, \zeta_4) \in (T^*V^{in})^4; \ \zeta_j \in L^+_{z_j} V, \ z_j \notin J^+_{(V,\widehat{g}')}(z_k), \ j \neq k\}.$$

Let $\mathcal{E} \subset \mathbb{R}^4$ be a neighbourhood of the origin. We say that a family of measurements

$$u(\epsilon) = (g'(\epsilon), \phi'(\epsilon), \mathcal{F}'(\epsilon)) \in \mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi}), \quad \epsilon \in \mathcal{E},$$

sends singularities of width $\delta > 0$ from $(\widetilde{\boldsymbol{z}}, \widetilde{\boldsymbol{\zeta}}) \in L_4 V^{in}$ to $y \in V$ if $\mathcal{F}'(0) = 0$ and the following conditions hold:

(In 1) Writing $(g^{in}(\epsilon), \phi^{in}(\epsilon), \mathcal{F}^{in}(\epsilon)) = u^{in}(\epsilon) = \Psi^\epsilon_* u(\epsilon)$ where Ψ^ϵ is the solution of (74) with $g' = g'(\epsilon)$, $\epsilon \in \mathcal{E}$, it holds that $supp\,(\Psi^\epsilon_* \mathcal{F}'(\epsilon)) \subset V^{in}$.

(In 2) The family u^{in} is C^4-smooth with respect to ϵ, (50) holds with $F = \mathcal{F}^{in}$, and writing $f_j = \partial_{\epsilon_j} \mathcal{F}^{in}(\epsilon)|_{\epsilon=0}$, it holds that

$$supp\,(f_j) \cap J^+_{(V,\widehat{g}')}(supp\,(f_k)) = \emptyset, \quad j \neq k.$$

(In 3) There are submanifolds $Y'_j \subset V^{in}$ of codimension one such that the geodesic γ_{z_j,ζ_j} intersects $\Phi(Y'_j)$ transversally, and that

$$f_j \in I(Y'_j \cap K'_j; E), \quad j = 1, 2, 3, 4,$$

where $K'_j = \Phi^{-1}(K(\widetilde{z}_j, \widetilde{\zeta}_j; \delta))$, $\widetilde{\boldsymbol{z}} = (\widetilde{z}_j)^4_{j=1}$ and $\widetilde{\boldsymbol{\zeta}} = (\widetilde{\zeta}_j)^4_{j=1}$.

(Out) Writing $g^{out} = \Xi^* g'$ where Ξ are the normal coordinates defined by (84) and $\alpha = (1, 1, 1, 1)$, it holds that $\partial^\alpha_\epsilon g^{out}$ is not smooth at the origin.

Definition 5.5. Let $r > 0$ and $k \in \mathbb{N}$, and let $L_4 V^{in}$ be as in Definition 5.4. We define a singularity detection function

$$D_{r,k} : V \times L_4 V^{in} \to \{0, 1\}$$

by $D_{r,k}(y, \boldsymbol{z}, \boldsymbol{\zeta}) = 1$ if and only if there are $k_0 \in \mathbb{N}$ and $r_0, \delta_0 > 0$ such that $k_0 \geq k$, $r_0 \leq r$ and that for all $0 < \delta < \delta_0$ there are $(\widetilde{\boldsymbol{z}}, \widetilde{\boldsymbol{\zeta}}) \in L_4 V^{in} \cap B((\boldsymbol{z}, \boldsymbol{\zeta}), \delta)$, a neighbourhood $\mathcal{E} \subset \mathbb{R}^4$ of the origin and a family of measurements

$$u(\epsilon) = (g'(\epsilon), \phi'(\epsilon), \mathcal{F}'(\epsilon)) \in \mathcal{D}_{r_0,k_0}(\widehat{g}, \widehat{\phi}), \quad \epsilon \in \mathcal{E},$$

that sends singularities of width δ from $(\widetilde{\boldsymbol{z}}, \widetilde{\boldsymbol{\zeta}})$ to y. We say that singularities are detected in a stable way at the point y if $D_{r,k}(y, \boldsymbol{z}, \boldsymbol{\zeta}) = 1$.

Theorem 5.6. Suppose that (ND) holds with $U = \overline{\Phi(V)}$. Let $r > 0$, $k \in \mathbb{N}$ and let $D_{r,k}$ be the singularity detection function. Then the closure $S(\boldsymbol{z}, \boldsymbol{\zeta})$ of the set $\{\Phi(y); \ y \in V, \ D_{r,k}(y, \boldsymbol{z}, \boldsymbol{\zeta}) = 1\}$ has the following two properties:

(i) If $\mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta}) \neq \emptyset$ then $\mathcal{E}_V(x) \subset S(\boldsymbol{z}, \boldsymbol{\zeta}) \subset J^+_{(M,\widehat{g})}(x)$.

(ii) If $\mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta}) = \emptyset$ then $S(\boldsymbol{z}, \boldsymbol{\zeta}) \subset C^+(\boldsymbol{z}, \boldsymbol{\zeta}) \cup \mathcal{K}_0(\boldsymbol{z}, \boldsymbol{\zeta})$.

Furthermore the function $D_{r,k}(y, \boldsymbol{z}, \boldsymbol{\zeta})$ is determined by the data $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$.

Before entering in the rather long proof, let us briefly describe its outline. We send four waves propagating near the four geodesics γ_{z_j, ζ_j}. If those geodesics intersect at x, we can make a small perturbation in z_j, ζ_j, so that the four geodesics intersect still at x, and choose the linearized sources f_j, as in Definition 5.4, so that the fourfold interaction of the corresponding linearized solutions at x has a

principal symbol, in the sense of Proposition 4.10, that does not vanish in the direction towards the point y. Moreover, by varying the symbols of f_j, we can produce a six dimensional space of singularities at y, and at least some of those singularities are observed in the normal coordinates Ξ.

Proof. It follows from Lemma 5.1 that the data $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$ determines whether a family a family of measurements

$$(88) \qquad u(\epsilon) = (g'(\epsilon), \phi'(\epsilon), \mathcal{F}'(\epsilon)) \in \mathcal{D}_{r_0,k_0}(\widehat{g}, \widehat{\phi}), \quad \epsilon \in \mathcal{E},$$

with $r_0 \le r$ and $k_0 \ge$, satisfies (In 1)–(In 3). Clearly, also the validity of (Out) can be verified by using the data. Thus the function $D_{r,k}(y, \boldsymbol{z}, \boldsymbol{\zeta})$ is determined by the data $\mathcal{D}_{r,k}(\widehat{g}, \widehat{\phi})$.

We note that since $\mathcal{E}_0(x) = \{\gamma_{x,\xi}(t); \ t \in (0, \rho(x,\xi)), \ \xi \in L_x^+ M\}$ is a smooth 3-dimensional manifold and the Hausdorff dimension of $\mathcal{K}_0(\boldsymbol{z}, \boldsymbol{\zeta})$ is at most 2, the set $\mathcal{E}_0(x) \setminus \mathcal{K}_0(\boldsymbol{z}, \boldsymbol{\zeta})$ is dense in $\mathcal{E}_V(x)$.

The properties (i) and (ii) follow, by passing to the closure, after we show that for any $y \in V$ satisfying

$$(89) \qquad \Phi(y) \notin C^+(\boldsymbol{z}, \boldsymbol{\zeta}) \cup \mathcal{K}_0(\boldsymbol{z}, \boldsymbol{\zeta})$$

the following two implications hold: (1) if $\mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta}) = \emptyset$ or (a) then $D(y, \boldsymbol{z}, \boldsymbol{\zeta}) = 0$; (2) if (b) then $D(y, \boldsymbol{z}, \boldsymbol{\zeta}) = 1$, where

(a) For any point $x \in \mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta})$ it holds that $\Phi(y) \notin J_{(M,\widehat{g})}^+(x)$.

(b) There are $x \in \mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta})$, $\xi \in L_x^+ M$ and $t \in (0, \rho(x,\xi))$ such that $\Phi(y) = \gamma_{x,\xi}(t)$.

To establish the implication (1), we show for large $k_0 \in \mathbb{N}$ and small $\delta_0, r_0 > 0$, that if the family (88) satisfies (In 1)–(In 3), with $(\widetilde{\boldsymbol{z}}, \widetilde{\boldsymbol{\zeta}}) \in L_4 V^{\text{in}} \cap B((\boldsymbol{z}, \boldsymbol{\zeta}), \delta)$ and $0 < \delta < \delta_0$, and $\mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta}) = \emptyset$ or (a) holds, then (Out) does not hold.

We use the shorthand notation $\gamma_j = \gamma_{z_j, \zeta_j}$ and

$$\mathcal{N} = \widehat{M} \setminus \left(C^+(\boldsymbol{z}, \boldsymbol{\zeta}) \cup \bigcup_{j=1}^4 J_{(M,\widehat{g})}^-(\Phi(z_j)) \right).$$

Note that the geodesics γ_j, γ_k with $j \ne k$ can intersect only once in $\mathcal{N}$. Indeed, if γ_j and γ_k intersect at points $x_1 < x_2$, then the existence of the broken causal geodesic between z_j and x_2, obtained by switching from γ_j to γ_k at x_1, implies that $\tau(z_j, x_2) > 0$. This is a contradiction with the fact that $x_2 = \gamma_j(s)$ for some $s < \rho(z_j, \zeta_j)$.

An analogous argument shows that γ_j can not have any self-intersections in $\mathcal{N}$. In particular, if $\widetilde{\mathcal{N}} \subset \mathcal{N}$ is open and if its closure is contained in $\mathcal{N}$, then for small δ, writing $K_j = K(z_j, \zeta_j; \delta)$, the intersection $K_j \cap \widetilde{\mathcal{N}}$ is a smooth manifold. The same is true for $\widetilde{K}_j = K(\widetilde{z}_j, \widetilde{\zeta}_j; \delta)$ when $(\widetilde{z}_j, \widetilde{\zeta}_j) \in B((z_j, \zeta_j), \delta)$ for small δ. Hence the results in Section 4 can be applied on $\widetilde{\mathcal{N}}$. Observe also that when considering

a fourfold interaction at $x \in \widetilde{\mathcal{N}}$ the condition (67) holds, since there can not exist an earlier threefold interaction.

For large k_0 and small r_0, Lemma 5.1 implies that the family (88) can be also written as

$$u(\epsilon) = \Phi^*_{g(\epsilon)}(g(\epsilon), \phi(\epsilon), \mathcal{F}(\epsilon)), \quad \Psi^\epsilon_* \mathcal{F}'(\epsilon) = \Phi^* \mathcal{F}(\epsilon),$$

where $(g(\epsilon), \phi(\epsilon), I_{g(\epsilon)} \mathcal{F}(\epsilon))$ solves (18). We define $v(\epsilon)$ by (40), and use the notation $Y' = \bigcup_{j=1}^4 Y'_j$ and $Y = \Phi(Y')$ where Y'_j is as in (In 3). We write also $\alpha = (1, 1, 1, 1)$ and

$$\mathbb{K}_\beta = \mathrm{singsupp}(\partial^\beta_\epsilon v|_{\epsilon=0}).$$

Now (In 3) implies that for small δ,

$$v_j = \partial_{\epsilon_j} v|_{\epsilon=0} \in I(\widetilde{K}_j; E) \quad \text{in } \widetilde{\mathcal{N}} \setminus Y.$$

Moreover, Lemmas 4.2 and 4.5 imply that $\mathbb{K}_\beta \subset \mathcal{K}(\widetilde{\boldsymbol{z}}, \widetilde{\boldsymbol{\zeta}}, \delta)$ for $\beta < \alpha$.

Let $x \in \mathcal{N}$ be a point where a pair of geodesics γ_j, γ_k, $j \neq k$, intersect. There is at most one such a point, since any pair of geodesics can intersect only once in $\mathcal{N}$. If $\mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta}) = \emptyset$, then for small δ, one of the distributions v_j, $j = 1, 2, 3, 4$, is smooth near x. Thus Lemma 4.6 implies that $\mathbb{K}_\alpha \subset \mathcal{K}(\widetilde{\boldsymbol{z}}, \widetilde{\boldsymbol{\zeta}}, \delta)$. As y satisfies (89) it holds for small enough δ that $\Phi(y) \notin \mathbb{K}_\alpha$. On the other hand if (a) holds, then for small δ it holds again that $\Phi(y) \notin \mathbb{K}_\alpha$ since $\mathbb{K}_\alpha$ lies near

$$J^+_{(M,\widehat{g})}(x) \cup \bigcup_{j=1}^4 \gamma_{z_j, \zeta_j}(\mathbb{R}),$$

and $\bigcup_{j=1}^4 \gamma_{z_j, \zeta_j}((0, \infty)) \subset \mathcal{K}_0(\boldsymbol{z}, \boldsymbol{\zeta})$.

In particular, in both the cases $\mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta}) = \emptyset$ and (a), it holds for small δ that

$$\Phi(y) \notin \mathrm{singsupp}(\partial^\beta_\epsilon g|_{\epsilon=0}), \quad \beta \leq \alpha.$$

Using the naturality of the exponential map, analogously with the proof of Lemma 5.3, we see that $\partial^\alpha_\epsilon g^{\mathrm{out}}$ is smooth near the origin. This concludes the proof of the implication (1).

Let us now turn to the implication (2) and suppose that (b) holds. We choose spacelike submanifolds

$$Y'_j = Y'(z_j, \zeta_j) \subset V^{\mathrm{in}}, \quad j = 1, 2, 3, 4,$$

of codimension one satisfying the following:

(S1) γ_j intersects Y_j transversally where $Y_j = \Phi(Y'_j)$,
(S2) $z_j < z$ with respect to $\widehat{g}'$ for all $z \in Y'_j$,
(S3) $Y'_j \cap J^+_{(V,\widehat{g}')}(z) = \emptyset$ for all $z \in Y'_k$ and $j \neq k$.

Let

(90) $$\sigma_j \in S^p(Y_j \cap K_j; \mathbb{S} \oplus \mathbb{R}^L), \quad j = 1, 2, 3, 4,$$

and define the map

$$(91) \qquad F(\epsilon, g, \phi) = I_g A_{(g,\phi)}(\mathbb{F} \sum_{j=1}^{4} \epsilon_j \sigma_j), \quad \epsilon = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4),$$

where $\mathbb{F}$ and $A_{(g,\phi)}$ are as in Proposition 3.6. We choose the neighbourhood U in Proposition 3.6 so that $U = U_1 \cup \cdots \cup U_4$ where $\overline{Y}_j \subset U_j$, $U \subset \Phi(V^{\mathrm{in}})$, and that the following perturbed version of (S3) holds:

(S3') $U_j \cap J^+_{(V,\widehat{g}')}(z) = \emptyset$ for all $z \in U_k$ and $j \neq k$.

For any $k_0 \in \mathbb{N}$, the quasilinear wave equation (18) with the right-hand side $(F^1, F^2) = F(\epsilon, g, \phi)$ has a unique solution

$$(g(\epsilon), \phi(\epsilon)) \in C^\infty(\mathcal{E}; C^{k_0}(\widehat{M}; \mathrm{Sym}^2 \otimes \mathbb{R}^L))$$

when $p \in \mathbb{R}$ is negative enough and $\mathcal{E} \subset \mathbb{R}^4$ is a small neighbourhood of the origin. Note that the neighbourhood $\mathcal{E}$ depends on the symbols σ_j, $j = 1, 2, 3, 4$.

We define

$$\mathcal{F}(\epsilon) = I_{g(\epsilon)} F(\epsilon, g(\epsilon), \phi(\epsilon)).$$

By (A1) of Proposition 3.6, the family $(g(\epsilon), \phi(\epsilon), \mathcal{F}(\epsilon))$ satisfies the compatibility condition (5), and therefore it solves (1)–(3). Moreover, the support condition in (A1) implies that $\mathrm{supp}\,(\mathcal{F}(\epsilon)) \subset U \subset \Phi(V^{\mathrm{in}})$. By (A3) the source $\mathcal{F}(\epsilon)$ depends smoothly on $\epsilon \in \mathcal{E}$. For any $r_0 > 0$ there is a neighbourhood $\mathcal{E}_{r_0} \subset \mathcal{E}$ of the origin so that $u(\epsilon) = \Phi^*_{g(\epsilon)}(g(\epsilon), \phi(\epsilon), \mathcal{F}(\epsilon)) \in \mathcal{D}_{r_0, k_0}(\widehat{g}, \widehat{\phi})$ when $\epsilon \in \mathcal{E}_{r_0}$.

Lemma 5.1 implies that

$$(92) \qquad u^{\mathrm{in}}(\epsilon) = \Psi^\epsilon_* u(\epsilon) = \Phi^*(g(\epsilon), \phi(\epsilon), \mathcal{F}(\epsilon)).$$

Using again the support condition in (A1) we see that $u^{\mathrm{in}}(\epsilon)$ satisfies (In 1). The equation (92) implies also that u^{in} is smooth with respect to ϵ. Linearity of $A_{(g,\phi)}$ implies that

$$(93) \qquad \partial_{\epsilon_j} \mathcal{F}(\epsilon)|_{\epsilon=0} = A_{(\widehat{g}, \widehat{\phi})}(E\sigma_j) \in I(Y_j \cap K_j; E).$$

To see that (In 2) is satisfied, we use (S3') together with arguments analogous to Lemma 5.9 in the below appendix. The condition (In 3) follows from (92), (93) and (S1).

We will finish the proof of the implication (2) by showing that for all small $\delta > 0$ there are $(\widetilde{z}_j, \widetilde{\zeta}_j) \in B((z_j, \zeta_j), \delta)$ and symbols $\sigma_j \in S^p(Y_j \cap K_j; \mathbb{S} \oplus \mathbb{R}^L)$ such that (Out) holds. By (b) there is $x \in \mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta})$, and by perturbing the geodesics γ_j slightly we get the following. For any small $\delta > 0$, there are $(\widetilde{z}^n_j, \widetilde{\zeta}^n_j) \in B((z_j, \zeta_j), \delta)$ such that the geodesics $\gamma^n_j = \gamma_{\widetilde{z}^n_j, \widetilde{\zeta}^n_j}$, $j = 1, 2, 3, 4$, intersect at the point x and the corresponding manifolds K^n_j satisfy $(\boldsymbol{b}^n, \boldsymbol{s}^n)^6_{n=1} \in \mathcal{U}$ for some $\boldsymbol{s}^n$ where $\boldsymbol{b}^n = (b^n_1, b^n_2, b^n_3, b^n_4)$, $N^*_x K^n_j = \mathbb{R}b^n_j$ and $\mathcal{U}$ is the set in Proposition 4.10. For small δ, the geodesic γ^n_j intersects Y_j and we denote by z^n_j the point of intersection. Moreover, we write $\xi = \xi^n_1 + \cdots + \xi^n_4$ where $\xi^n_j \in \mathbb{R}b^n_j \setminus 0$. Then there is a unique covector

$\zeta_j^n \in N_{z_j^n}^*(Y_j \cap K_j^n)$ that lies on the bicharacteristic of $\square_{\widehat{g},\widehat{\phi}}$ through (x, ξ_j^n). As discussed in Section 3.6, taking $\sigma_j = \sigma_j^n$ in (90), the map

$$\sigma_j^n(z_j^n, \zeta_j^n) \mapsto \sigma[v_j](x, \xi_j^n)$$

is an isomorphism between the fibres $(\mathbb{S}_{Y_j \cap K_j^n} \oplus \mathbb{R}^L)_{(z_j^n, \zeta_j^n)}$ and $(\mathbb{S}_{K_j^n} \oplus \mathbb{R}^L)_{(x, \xi_j^n)}$. Thus we can choose the symbols σ_j^n so that $\boldsymbol{s}^n = (\sigma[v_j](x, \xi_j^n))_{j=1}^4$, $n = 1, \ldots, 6$.

We write

$$\widetilde{E}_x = span\{\mathscr{S}(\boldsymbol{b}^n, \boldsymbol{s}^n); \ n = 1, \ldots, 6\},$$

and $K_5 = K(x, \xi; \delta')$, $\delta' > 0$. Observe that K_5 is a smooth manifold near $\Phi(y)$ since $t < \rho(x, \xi)$. There is a unique covector $\eta \in N_y^* K_5$ that lies on the bicharacteristic of $\square_{\widehat{g},\widehat{\phi}}$ through (x, ξ). Recalling (70), the discussion in Section 3.4 implies that the map

$$\sigma[\mathscr{I}_4(v_1, v_2, v_3, v_4)](x, \xi) \mapsto \sigma[v_{1234}](y, \eta)$$

is an isomorphism between the fibres E_x and E_y. Hence it maps the subspace $\widetilde{E}_x$ onto a 6-dimensional subspace $\widetilde{E}_y$ of E_y. Moreover, by Remark 3.7, the subspace $\widetilde{E}_y$ is contained in Sym_y^2 since $\widetilde{E}_x$ is contained in Sym_x^2. The intersection of $\widetilde{E}_y$ and the 6-dimensional subspace $Sym^2(K_5)_y \subset Sym_y^2$ is at least of dimension 2. Now Lemmas 5.2 and 5.3 imply that there is $n = 1, \ldots, 6$ such that (Out) holds when $\sigma_j = \sigma_j^n$ in (90). $\qquad\square$

Let us summarize the above considerations. When the small (2-dimensional) sets $\mathcal{K}_0(\boldsymbol{z}, \boldsymbol{\zeta})$ are not considered, Theorem 5.6 essentially states the following: First, if the geodesics γ_{z_j, ζ_j}, $j = 1, 2, 3, 4$, intersect before their first cut points and the intersection point is x, then the set $S(\boldsymbol{z}, \boldsymbol{\zeta}) \subset \Phi(V)$ where singularities can be detected in a stable way satisfies $\mathcal{E}_V(x) \subset S(\boldsymbol{z}, \boldsymbol{\zeta}) \subset J^+(x)$. The set $\mathcal{E}_V(x)$, defined in (87), is called the earliest light observation set corresponding to the source point x. Second, if the geodesics γ_{z_j, ζ_j} do not intersect before their first cut points, then singularities are not detected in a stable way outside the sets $J^+(p_j)$ where $p_j = \gamma_{z_j, \zeta_j}(\rho(z_j, \zeta_j))$, $j = 1, 2, 3, 4$, are the first cut points.

Roughly speaking, the above means that when we send singular waves along the geodesics γ_{z_j, ζ_j}, and the geodesics intersect at a point x before their cut points, the non-linear interaction creates an artificial point source at the point x. This point source sends singularities along the light cone emanating from the point x and the earliest arriving singularities are detected in $\Phi(V)$. Thus the inverse problem in Theorem 1.1 is reduced to the geometric problem of determining the conformal type of $I(\mu(0), \mu(1))$ when we are given the sets $S(\boldsymbol{z}, \boldsymbol{\zeta})$ for all initial directions $(\boldsymbol{z}, \boldsymbol{\zeta}) \in L_4 V^{in}$. This geometric problem is solved [49] where the inverse problem for the passive observations of the light cones emanating from the point sources is studied. Thus the rest of the proof of Theorem 1.1 is purely geometrical and coincides with the geometric part of the proof in [49]. As our notations differ slightly from those in [49], we spell out the relation to [49] in the short proof below.

Proof of Theorem 1.1. Similarly to Definition 4.2 of [49] we define the set $\mathcal{S}_e(\boldsymbol{z}, \boldsymbol{\zeta}) \subset S(\boldsymbol{z}, \boldsymbol{\zeta})$ of the earliest points where singularities are detected in a stable way. Lemma 4.4 of [49] shows that if $x \in \mathbb{X}(\boldsymbol{z}, \boldsymbol{\zeta})$ then $\mathcal{S}_e(\boldsymbol{z}, \boldsymbol{\zeta})$ coincides with the earliest observation set $\mathcal{E}_V(x)$ corresponding to the point x and the observation set $\Phi(V)$. Theorem 5.6 above, together with Theorem 4.5 and Lemma 4.4 of [49], implies that the data set (8) determines uniquely the collection of the earliest light observations sets

$$\Phi^{-1}(\mathcal{E}_V(W)) = \{\Phi^{-1}(\mathcal{E}_V(x)); \ x \in W\}$$

with source points in the chronological diamond $W = I(\mu(0), \mu(1))$. By [49, Th. 1.2] the collection $\mathcal{E}_V(W)$, or equivalently, the collection $\Phi^{-1}(\mathcal{E}_V(W))$, determines the conformal type of the set $(W, \widehat{g})$. $\qquad\square$

Corollary 1.2 follows in the same way as Corollary 1.3 of [49].

APPENDIX A: PROOFS FOR LINEARIZATION STABILITY

The next lemma is well-known but we give a short proof for the convenience of the reader.

Lemma 5.7. *Let $(M, \widehat{g})$ be a background spacetime, and suppose that a Lorentzian metric $g \in C^3(M; \mathrm{Sym}^2)$ satisfies the initial condition (3), (M, g) is globally hyperbolic, and that $\Sigma_0 = \{0\} \times N$ is a Cauchy surface with respect to g. Define the reduced Einstein tensor by $\mathrm{Ein}_{\widehat{g}}(g) = I_g \mathrm{Ric}_{\widehat{g}}(g)$. Then $\mathrm{div}_g \mathrm{Ein}_{\widehat{g}}(g) = 0$ implies that $H_{\widehat{g}}(g) = 0$.*

Proof. We use the shorthand notation $H = H_{\widehat{g}}(g)$, $\nabla = \nabla_g$, $\widehat{\nabla} = \nabla_{\widehat{g}}$, $\nabla_j = \nabla_{\partial_j}$ and $\widehat{\nabla}_j = \widehat{\nabla}_{\partial_j}$, and define also the operator $\flat_g^s$ that lowers indices and symmetrizes two tensors S as follows,

$$(\flat_g^s S)_{jk} = \frac{1}{2}(g_{jp} S_k^p + g_{kp} S_j^p).$$

In local coordinates, $\mathrm{tr}_g \flat_g^s \widehat{\nabla} H = \widehat{\nabla}_p H^p$, whence

$$\mathrm{Ein}(g)_{jk} - \mathrm{Ein}_{\widehat{g}}(g)_{jk} = \frac{1}{2}(g_{jp}\widehat{\nabla}_k H^p + g_{kp}\widehat{\nabla}_j H^p - g_{jk}\widehat{\nabla}_p H^p).$$

As also $\mathrm{Ein}(g)$ has vanishing divergence,

$$0 = g^{qj}\nabla_q(g_{jp}\widehat{\nabla}_k H^p + g_{kp}\widehat{\nabla}_j H^p - g_{jk}\widehat{\nabla}_p H^p)$$
$$= g_{kp}g^{qj}\nabla_q\widehat{\nabla}_j H^p + L_{kp}H^p,$$

where $L_{kp}H^p = \nabla_p\widehat{\nabla}_k H^p - \nabla_k\widehat{\nabla}_p H^p$. Note that L_{kp} is a linear first order differential operator. We see that H satisfies the linear second order hyperbolic system

$$(94) \quad \begin{cases} g^{qj}\nabla_q\widehat{\nabla}_j H^r + g^{rk}L_{kp}H^p = 0, & \text{in } M, \\ H^r = 0, & \text{in } (-\infty, 0) \times N, \end{cases} \quad r = 1, 2, 3, 4,$$

whose principal part coincides with the principal part of $\square_g$. As (M,g) is globally hyperbolic and Σ_0 is a Cauchy surface for g, $H = 0$ on M, see e.g. [76, Cor. 12.14]. $\qquad\square$

Corollary 5.8. *Let $(M,\widehat{g})$ and $g \in C^3(M; \mathrm{Sym}^2)$ be as in Lemma 5.7. Let*

$$\phi \in C^2(M; \mathbb{R}^L), \quad F^1 \in C^1(M; \mathrm{Sym}^2), \quad F^2 \in C^0(M; \mathbb{R}^L).$$

Suppose that $S = (g, \phi, F^1, F^2)$ satisfies (18). Then S satisfies (1)–(2) if and only if the compatibility condition (5) holds with $\mathcal{F}^1 = I_g F^1$ and $\mathcal{F}^2 = F^2$.

Proof. If S satisfies (1)–(2) then a direct computation shows that (5) holds. Suppose now that (5) holds. We apply I_g on the first equation in (18), and get

$$\mathrm{Ein}_{\widehat{g}}(g) = \mathbb{T}(g, \phi) + \mathcal{F}^1.$$

Now a computation shows that (18) and (5) imply

$$\mathrm{div}_g \, \mathrm{Ein}_{\widehat{g}}(g) = (\square_g - \mathcal{V}'(\phi))\phi \cdot d\phi + \mathrm{div}_g \mathcal{F}^1 = \mathrm{div}_g \mathcal{F}^1 + \mathcal{F}^2 \cdot d\phi = 0.$$

Lemma 5.7 implies $H_{\widehat{g}}(g) = 0$, and therefore

$$\mathrm{Ein}(g) = \mathrm{Ein}_{\widehat{g}}(g) = T(g, \phi) + \mathcal{F}^1.$$

$\qquad\square$

Proof of Lemma 3.2. We omit writing ϕ as a subscript in the proof. We denote by e_ℓ, $\ell = 1, \dots, L$, the constant frame corresponding to the standard basis of $\mathbb{R}^L$, and define a vector bundle homomorphism $A^1 \in C^\kappa(U; \mathrm{Hom}(T^*U, \mathbb{R}^L))$ by

$$A^1(d\phi_j) = e_j, \quad j = 1, 2, 3, 4.$$

Then $A^1(d\phi_j) \cdot d\phi = e_j \cdot d\phi = d\phi_j$, and A^1 satisfies (i).

We define

$$A^2(e_\ell) = 0, \qquad\qquad \ell = 1, 2, 3, 4,$$

$$A^2(e_\ell) = e_\ell - \sum_{j=1}^{4} c_{\ell j} e_j, \quad \ell = 5, \dots, L,$$

where $c_{\ell j} \in C^\kappa(U; \mathbb{R})$ are the coefficients of $d\phi_\ell$ in the frame $d\phi_j$, $j = 1, 2, 3, 4$. Note that for $\ell = 5, \dots, L$,

$$A^2(e_\ell) \cdot d\phi = d\phi_\ell - \sum_{j=1}^{4} c_{\ell j} d\phi_j = 0.$$

Hence A^2 satisfies (ii).

Let us now turn to the property (iii). Suppose that $w = \sum_{\ell=1}^{L} a_\ell e_\ell$. Then

$$A^1(w \cdot d\phi) = \sum_{\ell=1}^{L} a_\ell A^1(d\phi_\ell) = \sum_{j=1}^{4} a_j e_j + \sum_{\ell=5}^{L} \sum_{j=1}^{4} a_\ell c_{\ell j} e_j,$$

$$A^2(w) = \sum_{\ell=5}^{L} a_\ell e_\ell - \sum_{\ell=5}^{L} a_\ell \sum_{j=1}^{4} c_{\ell j} e_j,$$

and thus $A^1(w \cdot d\phi) + A^2(w) = w$, that is, (iii) holds. $\qquad \square$

Note that $A^1_\phi(v) = 0$ and $A^2_\phi(w) = 0$ outside the supports of v and w, respectively, simply because these maps are vector bundle homomorphisms. We need a similar property also for the derivatives of A^j_ϕ, $j = 1, 2$, with respect to ϕ. We will give a proof only in the case of $j = 1$, the other case being similar.

Lemma 5.9. *Let $\widehat{\phi}$ be as in Lemma 3.2. Let $\mathcal{E} \subset \mathbb{R}$ be a neighbourhood of the origin, and let $\phi(\epsilon) \in C^\infty(\mathcal{E}; C^{\kappa+1}(U; \mathbb{R}^L))$ satisfy $\phi(0) = \widehat{\phi}$. Then $\partial_\epsilon A^1_\phi|_{\epsilon=0}(v) = 0$ if the supports of $\partial_\epsilon \phi|_{\epsilon=0}$ and v do not intersect.*

Proof. Observe that in local coordinates,

$$A^1_{\phi(\epsilon)} = \begin{pmatrix} \Phi(\epsilon)^{-1} \\ 0 \end{pmatrix}$$

where $\Phi(\epsilon)$ is the matrix that has the columns $d\phi_j$, $j = 1, 2, 3, 4$. Thus, writing $\widehat{\Phi} = \Phi(0)$, we have

$$\partial_\epsilon A^1_\phi|_{\epsilon=0} = \begin{pmatrix} -\widehat{\Phi}^{-1} \partial_\epsilon \Phi|_{\epsilon=0} \widehat{\Phi}^{-1} \\ 0 \end{pmatrix},$$

and the claim follows. $\qquad \square$

APPENDIX B: LOCAL EXISTENCE OF SOLUTIONS

In this section we outline briefly the results on quasilinear wave equations that we need. The reduced Einstein's equations coupled with scalar fields (18) can be written in the abstract form (42), that is,

$$(95) \qquad\qquad g^{-1}(D, D)v + A(x, v, Dv) = F.$$

Here $v = (g - \widehat{g}, \phi - \widehat{\phi})$, and when considering sources of the form (91) we allow also F to be a function of v and Dv, that is, $F = F(x, v, Dv)$. Recall that a globally hyperbolic Lorentzian manifold is isometric to a product manifold $\mathbb{R} \times N$ with metric of the form

$$\widehat{g} = -\beta(t, y)dt^2 + \kappa(t, y).$$

Then (95) can be written in the form

$$(96) \qquad \beta(t,y;v)\partial_t^2 v = \sum_{j,k=1}^{3} \kappa^{jk}(t,y;v)\partial_{y^j}\partial_{y^k} v + A(t,y;v,Dv) - F(t,y;v,Dv),$$

where $g(t,y) = -\beta(t,y;v)dt^2 + \kappa(t,y;v)$. Well-posedness of equations of this form was established in [42]. The results there are not directly applicable in our case since [42] considers perturbations of the initial condition whereas we consider perturbations via a small compactly supported source F. However, the fixed point argument there can be adapted to our case, and we refer to Appendix B in [50] for the details.

Let $\widehat{T} > 0$ and consider the spaces

$$E^k = \bigcap_{j=0}^{k} C^j(0,\widehat{T};H^{k-j}(N;\mathrm{Sym}^2 \oplus \mathbb{R}^L)), \quad s \in \mathbb{N}.$$

Let $K \subset (0,\widehat{T}) \times N$ be compact. Suppose that for all large $k \in \mathbb{N}$, there is a neighbourhood U_k of the origin in E^k such that for $v \in U_k$ the function $f(t,y) = F(t,y;v,Dv)$ satisfies $\mathrm{supp}\,(f) \subset K$. Then for all large $k \in \mathbb{N}$ there is $r > 0$ such that if $\|f\|_{E^k} < r$ for all $v \in U_k$ then there is a unique solution $v \in E^k$ of (96) satisfying the initial condition (3).

Observe that the above result is local in the sense that only a fixed time interval is considered and the source is assumed to be supported in a fixed compact set.

In the case of Cauchy problem for Einstein's equations, there are also global results available, see e.g. [18, 37, 54, 75, 83]. Above we paid also no attention to the size of the smoothness index k. For results on the Cauchy problem for the Einstein and the Eistein-scalar field systems in low regularity settings see [20, 19, 76, 77].

APPENDIX C: ON ACTIVE MEASUREMENTS

By an active measurement we mean a model where we can control some of the physical fields and the other physical fields adapt to the changes in all the fields so that the conservation law holds. Roughly speaking, we can consider measurement devices as a complicated process that changes one energy form to other forms of energy, like a system of explosives that transform some potential energy to kinetic energy. This process creates a perturbation of the metric, a gravitational wave, and it also perturbs the matter fields. We can then observe these perturbations in a subset of the spacetime.

There are several proposals on generation of large enough gravitational waves so that they can in principle be detected [14, 24, 71, 88]. Also, gravitational waves from two black holes, in close orbit around each other, were recently observed [53], however, no artificially generated gravitational waves have been measured at the

present time. In this paper, our aim has been to consider a mathematical model that can be rigorously analyzed.

Let us take $\mathcal{V}_\ell(\phi) = \frac{1}{2}m^2\phi_\ell^2$ for simplicity, and let g and ϕ satisfy Einstein's equations coupled with scalar fields

$$(97) \qquad \mathrm{Ein}_{\widehat{g}}(g) = P + \mathbb{T}(g, \phi),$$

$$\Box_g \phi_\ell - m^2 \phi_\ell = S_\ell \quad \text{in } \widehat{M}, \quad \ell = 1, 2, 3, \ldots, L,$$

$$S_\ell = Q_\ell + \mathcal{S}_\ell^{2nd}(g, \phi, Q, P), \quad \text{in } \widehat{M},$$

$$g = \widehat{g}, \quad \phi_\ell = \widehat{\phi}_\ell, \quad \text{in } (-\infty, 0) \times N.$$

Here, in comparison to (1)–(2), we have simply written $P = \mathcal{F}^1$ and $\mathcal{S}_\ell^{2nd} = \mathcal{F}_\ell^2 - Q_\ell$, where Q_ℓ is arbitrary. This is to emphasize that P and Q_ℓ are considered as primary sources produced by the measurement devices, and the functions $\mathcal{S}_\ell^{2nd}(g, \phi, Q, P)$ are secondary sources that transform energy and momentum from the ϕ fields to the sources Q_ℓ and P. Mathematically speaking the secondary sources cause an effect that guarantees the conservation law (4) is satisfied.

We allow $\mathcal{S}_\ell^{2nd}$ to depend also on the first derivatives of (g, ϕ, Q, P) but hide this in the notation. The precise construction of $\mathcal{S}_\ell^{2nd}$ is quite technical, see (38), but these functions can be viewed as the instructions on how to build a device that can be used to measure the structure of the spacetime far away. Put differently they can be seen as recipe to transform energy between the gravitational field and the scalar fields. We note that if the primary sources satisfy $P = \epsilon f_1$ and $Q = \epsilon f_2$ and the pair (f_1, f_2) satisfies the linearized conservation law (21), then the first ϵ derivative of the secondary source is zero.

Above, the source term P can be written in the form

$$(98) \qquad P_{jk}(x) = \sum_{\kappa=1}^K \mu_\kappa(x) v_j^\kappa(x) v_k^\kappa(x),$$

where v^κ are timelike covector fields, μ_κ real-valued functions and $K \geq 10$. When $\mu_\kappa \geq 0$, this decomposition corresponds to non-interacting fluids with densities μ_k, 4-velocity vectors v^κ, and zero pressures (i.e. particle or dust flows). Recall that above we have considered the case when the source field P in (97) vanishes for the background solution $(\widehat{g}, \widehat{\phi})$. This condition can be relaxed, for instance, we can assume that $(\widehat{g}, \widehat{\phi})$ satisfy (97) with background source $\widehat{P}$ for which the conservation law $\mathrm{div}_{\widehat{g}}(\widehat{P}) = 0$ holds. When we assume that $\widehat{P}$ is written using suitable timelike fields $\widehat{v}^\kappa$, $\kappa = 1, 2, \ldots, K$, $K \geq 10$ and densities $\widehat{\mu}_\kappa$ that are positive in V, any small symmetric perturbation P of $\widehat{P}$ in the set V can be obtained using the fluids $(\mu_\kappa, \widehat{v}_\kappa)$ such that the perturbed densities μ_κ are positive. The inverse problem studied in Theorem 1.1 can be modified to the setting where the background source $\widehat{P}$ in given in the form (98) with fluids $(\widehat{\mu}_\kappa, \widehat{v}_\kappa)$ that are non-vanishing near the

geodesic μ and the measurements are implemented by perturbing these fluids, but the detailed considerations of this model are outside the scope of this paper.

References

[1] S. Alinhac. Non-unicité du problème de Cauchy. *Ann. of Math. (2)*, 117(1):77–108, 1983.

[2] J. M. Arms and J. E. Marsden. The absence of Killing fields is necessary for linearization stability of Einstein's equations. *Indiana Univ. Math. J.*, 28(1):119–125, 1979.

[3] D. Baskin and A. Sá Barreto. Radiation fields for semilinear wave equations. *Trans. Amer. Math. Soc.*, 367(6):3873–3900, 2015.

[4] D. Baskin, A. Vasy, and J. Wunsch. Asymptotics of radiation fields in asymptotically Minkowski space. *Amer. J. Math.*, 137(5):1293–1364, 2015.

[5] M. Beals. Self-spreading and strength of singularities for solutions to semilinear wave equations. *Ann. of Math. (2)*, 118(1):187–214, 1983.

[6] M. Beals. *Propagation and interaction of singularities in nonlinear hyperbolic problems*, volume 3 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Boston, Inc., Boston, MA, 1989.

[7] J. K. Beem, P. E. Ehrlich, and K. L. Easley. *Global Lorentzian geometry*, volume 202 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, second edition, 1996.

[8] Y. M. Berezanski. The uniqueness theorem in the inverse problem of spectral analysis for the Schrödinger equation. *Trudy Moskov. Mat. Obšč.*, 7:1–62, 1958.

[9] A. N. Bernal and M. Sánchez. Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes. *Comm. Math. Phys.*, 257(1):43–50, 2005.

[10] A. N. Bernal and M. Sánchez. Globally hyperbolic spacetimes can be defined as 'causal' instead of 'strongly causal'. *Classical Quantum Gravity*, 24(3):745–749, 2007.

[11] J.-M. Bony. Second microlocalization and propagation of singularities for semilinear hyperbolic equations. In *Hyperbolic equations and related topics (Katata/Kyoto, 1984)*, pages 11–49. Academic Press, Boston, MA, 1986.

[12] D. Brill, O. Reula, and B. Schmidt. Local linearization stability. *J. Math. Phys.*, 28(8):1844–1847, 1987.

[13] D. R. Brill and S. Deser. Instability of closed spaces in general relativity. *Comm. Math. Phys.*, 32:291–304, 1973.

[14] G. F. Chapline, J. Nuckolls, and L. L. Wood. Gravitational-radiation production using nuclear explosions. *Physical Review D*, 10(4):1064–1065, aug 1974.

[15] Y. Choquet-Bruhat. Global wave maps on curved space times. In *Mathematical and quantum aspects of relativity and cosmology (Pythagoreon, 1998)*, volume 537 of *Lecture Notes in Phys.*, pages 1–29. Springer, Berlin, 2000.

[16] Y. Choquet-Bruhat. *General relativity and the Einstein equations*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2009.

[17] Y. Choquet-Bruhat and S. Deser. On the stability of flat space. *Ann. Physics*, 81:165–178, 1973.

[18] D. Christodoulou and S. Klainerman. *The global nonlinear stability of the Minkowski space*, volume 41 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993.

[19] M. Dafermos. Stability and instability of the Cauchy horizon for the spherically symmetric Einstein-Maxwell-scalar field equations. *Ann. of Math. (2)*, 158(3):875–928, 2003.

[20] M. Dafermos. The evolution problem in general relativity. In *Current developments in mathematics, 2008*, pages 1–66. Int. Press, Somerville, MA, 2009.

[21] J. J. Duistermaat and L. Hörmander. Fourier integral operators. II. *Acta Math.*, 128(3-4):183–269, 1972.

[22] A. E. Fischer and J. E. Marsden. Linearization stability of the Einstein equations. *Bull. Amer. Math. Soc.*, 79:997–1003, 1973.

[23] M. Fridman, A. Farsi, Y. Okawachi, and A. L. Gaeta. Demonstration of temporal cloaking. *Nature*, 481(7379):62–65, Jan 2012.

[24] A. Füzfa. How current loops and solenoids curve spacetime. *Physical Review D*, 93(2), jan 2016.

[25] J. Girbau and L. Bruna. *Stability by linearization of Einstein's field equation*, volume 58 of *Progress in Mathematical Physics*. Birkhäuser Verlag, Basel, 2010.

[26] C. R. Graham, C. Guillarmou, P. Stefanov, and G. Uhlmann. X-ray transform and boundary rigidity for asymptotically hyperbolic manifolds. *Preprint arXiv:1709.05053*, 2017.

[27] C. R. Graham and M. Zworski. Scattering matrix in conformal geometry. *Invent. Math.*, 152(1):89–118, 2003.

[28] A. Greenleaf and G. Uhlmann. Recovering singularities of a potential from singularities of scattering data. *Comm. Math. Phys.*, 157(3):549–572, 1993.

[29] C. Guillarmou, M. Salo, and L. Tzou. Inverse scattering at fixed energy on surfaces with Euclidean ends. *Comm. Math. Phys.*, 303(3):761–784, 2011.

[30] V. Guillemin. *Cosmology in $(2+1)$-dimensions, cyclic models, and deformations of $M_{2,1}$*, volume 121 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1989.

[31] V. Guillemin and G. Uhlmann. Oscillatory integrals with singular symbols. *Duke Math. J.*, 48(1):251–267, 1981.

[32] A. Hassell, R. Melrose, and A. Vasy. Microlocal propagation near radial points and scattering for symbolic potentials of order zero. *Anal. PDE*, 1(2):127–196, 2008.

[33] S. Helgason. *Differential geometry and symmetric spaces*. Pure and Applied Mathematics, Vol. XII. Academic Press, New York, 1962.

[34] P. Hintz and G. Uhlmann. Reconstruction of lorentzian manifolds from boundary light observation sets. *To appear in International Mathematics Research Notices, Preprint arXiv:1705.01215*, May 2017.

[35] P. Hintz and A. Vasy. Semilinear wave equations on asymptotically de Sitter, Kerr–de Sitter and Minkowski spacetimes. *Anal. PDE*, 8(8):1807–1890, 2015.

[36] P. Hintz and A. Vasy. Global analysis of quasilinear wave equations on asymptotically Kerr–de Sitter spaces. *Int. Math. Res. Not. IMRN*, (17):5355–5426, 2016.

[37] P. Hintz and A. Vasy. The global non-linear stability of the kerr-de sitter family of black holes. *Preprint arXiv:1606.04014*, 2016.

[38] L. Hörmander. *The analysis of linear partial differential operators. III*, volume 274 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1985.

[39] L. Hörmander. *The analysis of linear partial differential operators. IV*, volume 275 of *Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, Berlin, 1985.

[40] L. Hörmander. *The analysis of linear partial differential operators. I*, volume 256 of *Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, Berlin, 1990.

[41] P. R. Hoskins. Principles of ultrasound elastography. *Ultrasound*, 20(1):8–15, 2012.

[42] T. J. R. Hughes, T. Kato, and J. E. Marsden. Well-posed quasi-linear second-order hyperbolic systems with applications to nonlinear elastodynamics and general relativity. *Arch. Rational Mech. Anal.*, 63(3):273–294 (1977), 1976.

[43] J. Ilmavirta and M. Salo. Broken ray transform on a Riemann surface with a convex obstacle. *Comm. Anal. Geom.*, 24(2):379–408, 2016.

[44] V. Isakov. On uniqueness in inverse problems for semilinear parabolic equations. *Arch. Rational Mech. Anal.*, 124(1):1–12, 1993.

[45] M. S. Joshi and A. Sá Barreto. Inverse scattering on asymptotically hyperbolic manifolds. *Acta Math.*, 184(1):41–86, 2000.

[46] A. Katchalov, Y. Kurylev, and M. Lassas. *Inverse boundary spectral problems*, volume 123 of *Monographs and Surveys in Pure and Applied Mathematics.* Chapman & Hall/CRC, Boca Raton, FL, 2001.

[47] T. Kato. Quasi-linear equations of evolution, with applications to partial differential equations. pages 25–70. Lecture Notes in Math., Vol. 448, 1975.

[48] S. G. Krantz and H. R. Parks. *A primer of real analytic functions*, volume 4 of *Basler Lehrbücher.* Birkhäuser Verlag, Basel, 1992.

[49] Y. Kurylev, M. Lassas, and G. Uhlmann. Inverse problems for Lorentzian manifolds and non-linear hyperbolic equations. *Preprint arXiv:1405.3386*, 2014.

[50] Y. Kurylev, M. Lassas, and G. Uhlmann. Inverse problems in spacetime i: Inverse problems for Einstein equations - extended preprint version. *Preprint arXiv:1405.4503*, 2014.

[51] E. Larsson. Broken causal lens rigidity and sky shadow rigidity of lorentzian manifolds. *Preprint arXiv:1512.05200*, 2015.

[52] M. Lassas, G. Uhlmann, and Y. Wang. Inverse problems for semilinear wave equations on Lorentzian manifolds. *To appear in Comm. Math. Phys. Preprint arXiv:1606.06261*, 2016.

[53] LIGO Scientific Collaboration and Virgo Collaboration. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, Feb 2016.

[54] H. Lindblad and I. Rodnianski. Global existence for the Einstein vacuum equations in wave coordinates. *Comm. Math. Phys.*, 256(1):43–110, 2005.

[55] J. Luk and I. Rodnianski. Nonlinear interaction of impulsive gravitational waves for the vacuum einstein equations. *To appear in Cambridge Journal of Math. Preprint arXiv:1301.1072*, 2013.

[56] J.-P. Luminet, J. R. Weeks, A. Riazuelo, R. Lehoucq, and J.-P. Uzan. Dodecahedral space topology as an explanation for weak wide-angle temperature correlations in the cosmic microwave background. *Nature*, 425:593 EP –, Oct 2003.

[57] M. Maggiore. *Gravitational waves, Volume 1: Theory and Experiments.* Oxford Univ. Press, Oxford [u.a.], 2008.

[58] M. W. McCall, A. Favaro, P. Kinsler, and A. Boardman. A spacetime cloak, or a history editor. *Journal of Optics*, 13(2):024003, 2011.

[59] J. McLaughlin and D. Renzi. Shear wave speed recovery in transient elastography and supersonic imaging using propagating fronts. *Inverse Problems*, 22(2):681–706, 2006.

[60] R. Melrose and N. Ritter. Interaction of nonlinear progressing waves for semilinear wave equations. *Ann. of Math. (2)*, 121(1):187–213, 1985.

[61] R. Melrose and N. Ritter. Interaction of nonlinear progressing waves for semilinear wave equations. *Ann. of Math. (2)*, 121(1):187–213, 1985.

[62] R. Melrose, A. Sá Barreto, and A. Vasy. Asymptotics of solutions of the wave equation on de Sitter-Schwarzschild space. *Comm. Partial Differential Equations*, 39(3):512–529, 2014.

[63] R. Melrose and J. Wunsch. Propagation of singularities for the wave equation on conic manifolds. *Invent. Math.*, 156(2):235–299, 2004.

[64] R. B. Melrose and N. Ritter. Interaction of progressing waves for semilinear wave equations. II. *Ark. Mat.*, 25(1):91–114, 1987.

[65] R. B. Melrose, A. Sá Barreto, and M. Zworski. Semi-linear diffraction of conormal waves. *Astérisque*, (240):vi+132 pp. (1997), 1996.

[66] R. B. Melrose and G. A. Uhlmann. Lagrangian intersection and the Cauchy problem. *Comm. Pure Appl. Math.*, 32(4):483–519, 1979.

[67] B. O'Neill. *Semi-Riemannian geometry*, volume 103 of *Pure and Applied Mathematics*. Academic Press Inc., New York, 1983.

[68] G. P. Paternain. Inverse problems for connections. In *Inverse problems and applications: inside out. II*, volume 60 of *Math. Sci. Res. Inst. Publ.*, pages 369–409. Cambridge Univ. Press, Cambridge, 2013.

[69] G. P. Paternain, M. Salo, and G. Uhlmann. The attenuated ray transform for connections and Higgs fields. *Geom. Funct. Anal.*, 22(5):1460–1489, 2012.

[70] G. P. Paternain, M. Salo, and G. Uhlmann. Tensor tomography on surfaces. *Invent. Math.*, 193(1):229–247, 2013.

[71] M. Portilla and R. Lapiedra. Generation of high frequency gravitational waves. *Physical Review D*, 63(4), jan 2001.

[72] Rakesh and W. W. Symes. Uniqueness for an inverse problem for the wave equation. *Comm. Partial Differential Equations*, 13(1):87–96, 1988.

[73] J. Rauch and M. C. Reed. Singularities produced by the nonlinear interaction of three progressing waves; examples. *Comm. Partial Differential Equations*, 7(9):1117–1133, 1982.

[74] J. Rauch and M. C. Reed. Singularities produced by the nonlinear interaction of three progressing waves; examples. *Comm. Partial Differential Equations*, 7(9):1117–1133, 1982.

[75] H. Ringström. Future stability of the Einstein-non-linear scalar field system. *Invent. Math.*, 173(1):123–208, 2008.

[76] H. Ringström. *The Cauchy problem in general relativity*. ESI Lectures in Mathematics and Physics. European Mathematical Society (EMS), Zürich, 2009.

[77] H. Ringström. *On the topology and future stability of the universe*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2013.

[78] A. Sá Barreto. Radiation fields, scattering, and inverse scattering on asymptotically hyperbolic manifolds. *Duke Math. J.*, 129(3):407–480, 2005.

[79] M. Salo and X. Zhong. An inverse problem for the p-Laplacian: boundary determination. *SIAM J. Math. Anal.*, 44(4):2474–2495, 2012.

[80] P. Stefanov, G. Uhlmann, and A. Vasy. Boundary rigidity with partial data. *J. Amer. Math. Soc.*, 29(2):299–332, 2016.

[81] P. D. Stefanov. Uniqueness of the multi-dimensional inverse scattering problem for time dependent potentials. *Math Z*, 201(4):541–559, dec 1989.

[82] Z. Sun. On a quasilinear inverse boundary value problem. *Math. Z.*, 221(2):293–305, 1996.

[83] C. Svedberg. Future stability of the Einstein-Maxwell-scalar field system. *Ann. Henri Poincaré*, 12(5):849–917, 2011.

[84] J. Sylvester and G. Uhlmann. A global uniqueness theorem for an inverse boundary value problem. *Ann. of Math. (2)*, 125(1):153–169, 1987.

[85] D. Tataru. Unique continuation for solutions to PDE's; between Hörmander's theorem and Holmgren's theorem. *Comm. Partial Differential Equations*, 20(5-6):855–884, 1995.

[86] G. Uhlmann and A. Vasy. The inverse problem for the local geodesic ray transform. *Invent. Math.*, 205(1):83–120, 2016.

[87] A. Vasy. Some recent advances in microlocal analysis. In *Proceedings of ICM 2014 Seoul*, pages 915–939.

[88] J. Weber. Detection and generation of gravitational waves. *Phys. Rev.*, 117(1):306–313, jan 1960.

Department of Mathematics, University College London, Gower Street, London UK, WC1E 6BT.
   E-mail address: y.kurylev@ucl.ac.uk

Department of Mathematics and Statistics, University of Helsinki, Box 68, Helsinki, 00014, Finland
   E-mail address: Matti.Lassas@helsinki.fi

Department of Mathematics, University College London, Gower Street, London UK, WC1E 6BT.
   E-mail address: l.oksanen@ucl.ac.uk

Department of Mathematics, University of Washington, Box 354350 Seattle, Washington 98195, USA

Institute for Advanced Study, Hong Kong University of Science and Technology, Hong Kong SAR
   E-mail address: gunther@math.washington.edu