# The Area Theorem

Piotr T. Chruściel* Erwann Delay†
Département de Mathématiques
Faculté des Sciences
Parc de Grandmont
F37200 Tours, France

Gregory J. Galloway‡
Department of Mathematics and Computer Science
University of Miami
Coral Gables FL 33124, USA

Ralph Howard§
Department of Mathematics
University of South Carolina
Columbia S.C. 29208, USA

January 3, 2000

**Abstract**

We prove that the area of sections of future event horizons in space–times satisfying the null energy condition is non–decreasing towards the future under any one of the following circumstances: 1) the horizon is future geodesically complete; 2) the horizon is a black hole event horizon in a globally hyperbolic space–time and there exists a conformal completion with a "$\mathcal{H}$–regular" $\mathscr{I}^+$; 3) the horizon is a black hole event horizon in a space–time which has a globally hyperbolic conformal completion. (Some related results under less restrictive hypotheses are also established.) This

*Supported in part by KBN grant # 2 P03B 130 16. *E–mail*: Chrusciel@Univ-Tours.fr
†Current adress: Department of Mathematics, Royal Institute of Technology, S–10044 Stockholm. *E–mail*: Delay@gargan.math.Univ-Tours.fr
‡Supported in part by NSF grant # DMS-9803566. *E–mail*: galloway@math.miami.edu
§Supported in part by DoD Grant # N00014-97-1-0806 *E–mail*: howard@math.sc.edu

extends a theorem of Hawking, in which piecewise smoothness of the event horizon seems to have been assumed. No assumptions about the cosmological constant or its sign are made. We prove smoothness or analyticity of the relevant part of the event horizon when equality in the area inequality is attained — this has applications to the theory of stationary black holes, as well as to the structure of compact Cauchy horizons. In the course of the proof we establish several new results concerning the differentiability properties of horizons.

# Contents



# 1 Introduction

The thermodynamics of black holes rests upon *Hawking's area theorem* [37] which asserts that in appropriate space–times the area of cross–sections of black hole horizons is non–decreasing towards the future. In the published proofs of this result [37, 64] there is considerable vagueness as to the hypotheses of differentiability of the event horizon (see, however, [35]). Indeed, it is known that black hole horizons can be pretty rough [15], and it is not immediately clear that the area of their cross–sections can even be defined. The reading of the proofs given in [37, 64] suggests that those authors have assumed the horizons under consideration to be piecewise $C^2$. Such a hypothesis is certainly incompatible with the examples constructed in [15] which are *nowhere* $C^2$. The aim of this paper is to show that the monotonicity theorem holds without any further differentiability hypotheses on the horizon, in appropriate space–times, for a large class of cross–sections of the horizon. More precisely we show the following:

**Theorem 1.1 (The area theorem)** *Let $\mathcal{H}$ be a black hole event horizon in a smooth space–time $(M, g)$. Suppose that either*

*a) $(M, g)$ is globally hyperbolic, and there exists a conformal completion $(\bar{M}, \bar{g}) = (M \cup \mathscr{I}^+, \Omega^2 g)$ of $(M, g)$ with a $\mathcal{H}$–regular[1] $\mathscr{I}^+$. Further the null energy condition holds on the past $I^-(\mathscr{I}^+; \bar{M}) \cap M$ of $\mathscr{I}^+$ in $M$, or*

*b) the generators of $\mathcal{H}$ are future complete and the null energy condition holds on $\mathcal{H}$, or*

*c) there exists a globally hyperbolic conformal completion $(\bar{M}, \bar{g}) = (M \cup \mathscr{I}^+, \Omega^2 g)$ of $(M, g)$, with the null energy condition holding on $I^-(\mathscr{I}^+; \bar{M}) \cap M$.*

*Let $\Sigma_a$, $a = 1, 2$ be two achronal spacelike embedded hypersurfaces of $C^2$ differentiability class, set $S_a = \Sigma_a \cap \mathcal{H}$. Then:*

1. *The area* $\mathrm{Area}(S_a)$ *of $S_a$ is well defined.*

2. *If*

$$S_1 \subset J^-(S_2) ,$$

*then the area of $S_2$ is larger than or equal to that of $S_1$. (Moreover, this is true even if the area of $S_1$ is counted with multiplicity of generators provided that $S_1 \cap S_2 = \varnothing$, see Theorem 6.1.)*

Point 1. of Theorem 1.1 is Proposition 3.3 below, see also Proposition 3.4. Point 2. above follows immediately from Proposition 4.2, Corollary 4.14 and Proposition 4.18, as a special case of the first part of Theorem 6.1 below.[2] The question of how to define the area of sections of the horizon is discussed in detail in Section 3. It is suggested there that a notion of area, appropriate for the identification of area with the entropy, should include the multiplicity of generators of the horizon.

We stress that we are *not* assuming that $\mathscr{I}^+$ is null — in fact it could be even changing causal type from point to point — in particular Theorem 1.1 also

[1] See Section 4.1 for definitions.

[2] The condition $S_1 \cap S_2 = \varnothing$ of Theorem 6.1 is needed for monotonicity of "area with multiplicity", but is not needed if one only wants to compare standard areas, see Remark 6.2.

applies when the cosmological constant does not vanish. In point c) global hyperbolicity of $(\bar{M}, \bar{g})$ should be understood as that of a manifold with boundary, *cf.* Section 4.1. Actually in points a) and c) of Theorem 1.1 we have assumed global hyperbolicity of $(M, g)$ or that of $(\bar{M}, \bar{g})$ for simplicity only: as far as Lorentzian causality hypotheses are concerned, the assumptions of Proposition 4.1 are sufficient to obtain the conclusions of Theorem 1.1. We show in Section 4.1 that those hypotheses will hold under the conditions of Theorem 1.1. Alternative sets of causality conditions, which do not require global hyperbolicity of $(M, g)$ or of its conformal completion, are given in Propositions 4.8 and 4.10.

It seems useful to compare our results to other related ones existing in the literature [37, 47, 64]. First, the hypotheses of point a) above are fulfilled under the conditions of the area theorem of [64] (after replacing the space–time $M$ considered in [64] by an appropriate subset thereof), and (disregarding questions of differentiability of $\mathcal{H}$) are considerably weaker than the hypotheses of [64]. Consider, next, the original area theorem of [37], which we describe in some detail in Appendix B for the convenience of the reader. We note that we have been unable to obtain a proof of the area theorem without some condition of causal regularity of $\mathscr{I}^+$ (*e.g.* the one we propose in point a) of Theorem 1.1), and we do not know[3] whether or not the area theorem holds under the original conditions of [37] without the modifications indicated above, or in Appendix B; see Appendix B for some comments concerning this point.

Let us make a few comments about the strategy of the proof of Theorem 1.1. It is well known that event horizons are Lipschitz hypersurfaces, and the examples constructed in [8, 15] show that much more cannot be expected. We start by showing that horizons are *semi–convex*[4]. This, together with Alexandrov's theorem concerning the regularity of convex functions shows that they are twice differentiable in an appropriate sense (*cf.* Proposition 2.1 below) almost everywhere. This allows one to define notions such as the divergence $\theta_{\mathcal{A}l}$ of the generators of the horizon $\mathcal{H}$, as well as the divergence $\theta_{\mathcal{A}l}^{\mathcal{H}\cap\mathcal{S}}$ of sections $\mathcal{H}\cap\mathcal{S}$ of $\mathcal{H}$. The existence of the second order expansions at Alexandrov points leads further to the proof that, under appropriate conditions, $\theta_{\mathcal{A}l}$ or $\theta_{\mathcal{A}l}^{\mathcal{H}\cap\Sigma}$ have the right sign. Next, an approximation result of Whitney type, Proposition 6.6, allows one to embed certain subsets of the horizon into $C^{1,1}$ manifolds. The area theorem follows then from the change–of–variables theorem for Lipschitz maps proved in [23]. We note that some further effort is required to convert the information that $\theta_{\mathcal{A}l}$ has the correct sign into an inequality concerning the Jacobian that appears in the change of variables formula.

[3] The proof of Proposition 9.2.1 in [37] would imply that the causal regularity condition assumed here holds under the conditions of the area theorem of [37]. However, there are problems with that proof (this has already been noted by Newman [53]).

[4] Actually this depends upon the time orientation: *future* horizons are *semi–convex*, while *past* horizons are *semi–concave*.

Various authors have considered the problem of defining black holes in settings more general than standard asymptotically flat spacetimes or spacetimes admitting a conformal infinity; see especially [34, 47, 49] and references cited therein. It is likely that proofs of the area theorem given in these more general settings, for horizons assumed to be piecewise $C^2$, which are based on establishing the positivity of the (classically defined) expansion $\theta$ of the null generators can be adapted, using the methods of Section 4 (*cf.* especially Proposition 4.1 and Lemma 4.15), to obtain proofs which do not require the added smoothness. In all situations which lead to the positivity of $\theta$ in the weak Alexandrov sense considered here, the area theorem follows from Theorem 6.1 below.

It is of interest to consider the equality case: as discussed in more detail in Section 7, this question is relevant to the classification of stationary black holes, as well as to the understanding of compact Cauchy horizons. Here we prove the following:

**Theorem 1.2** *Under the hypotheses of Theorem 1.1, suppose that the area of $S_1$ equals that of $S_2$. Then*

$$(J^+(S_1) \setminus S_1) \cap (J^-(S_2) \setminus S_2)$$

*is a* smooth *(analytic if the metric is analytic) null hypersurface with vanishing null second fundamental form. Moreover, if $\gamma$ is a null generator of $\mathcal{H}$ with $\gamma(0) \in S_1$ and $\gamma(1) \in S_2$, then the curvature tensor of $(M, g)$ satisfies $R(X, \gamma'(t))\gamma'(t) = 0$ for all $t \in [0, 1]$ and $X \in T_{\gamma(t)}\mathcal{H}$.*

Theorem 1.2 follows immediately from Corollary 4.14 and Proposition 4.18, as a special case of the second part of Theorem 6.1 below. The key step of the proof here is Theorem 6.18, which has some interest in its own. An application of those results to stationary black holes is given in Theorem 7.1, Section 7.

As already pointed out, one of the steps of the proof of Theorem 1.1 is to establish that a notion of divergence $\theta_{\mathcal{A}l}$ of the generators of the horizon, or of sections of the horizon, can be defined almost everywhere, and that $\theta_{\mathcal{A}l}$ so defined is positive. We note that $\theta_{\mathcal{A}l}$ coincides with the usual divergence $\theta$ for horizons which are twice differentiable. Let us show, by means of an example, that the positivity of $\theta$ might fail to hold in space–times $(M, g)$ which do not satisfy the hypotheses of Theorem 1.1: Let $t$ be a standard time coordinate on the three dimensional Minkowski space–time $\mathbb{R}^{1,2}$, and let $K \subset \{t = 0\}$ be an open conditionally compact set with smooth boundary $\partial K$. Choose $K$ so that the mean curvature $H$ of $\partial K$ has changing sign. Let $M = I^-(K)$, with the metric conformal to the Minkowski metric by a conformal factor which is one in a neighborhood of $\partial\mathcal{D}^-(K; \mathbb{R}^{1,2}) \setminus K$, and which makes $\partial J^-(K, \mathbb{R}^{1,2})$ into $\mathscr{I}^+$ in the completion $\bar{M} \equiv M \cup \partial J^-(K, \mathbb{R}^{1,2})$. We have $M \setminus J^-(\mathscr{I}^+; \bar{M}) = \mathcal{D}^-(K; \mathbb{R}^{1,2}) \neq \varnothing$, thus $M$ contains a black hole region, with the event horizon being the Cauchy horizon $\partial\mathcal{D}^-(K; \mathbb{R}^{1,2}) \setminus K$. The generators of the event horizon coincide with the generators of $\partial\mathcal{D}^-(K;\mathbb{R}^{1,2})$, which are null geodesics normal to $\partial K$. Further for $t$ negative and close to zero the divergence $\theta$ of those generators is well defined in a classical sense (since the horizon is smooth there) and approaches, when $t$ tends to zero along the generators, the mean curvature $H$ of $\partial K$. Since the conformal factor equals one in a neighborhood of the horizon the null energy condition holds there, and $\theta$ is negative near those points of $\partial K$ where $H$ is negative. This implies also the failure of the area theorem for some (local) sections of the horizon. We note that condition b) of Theorem 1.1 is not satisfied because the generators of $\mathcal{H}$ are not future geodesically complete. On the other hand condition a) does not hold because $g$ will not satisfy the null energy condition throughout $I^-(\mathscr{I}^+;\bar{M})$ whatever the choice of the conformal factor[5].

This paper is organized as follows: In Section 2 we show that future horizons, as defined there, are always *semi–convex* (Theorem 2.2). This allows us to define such notions as the *Alexandrov divergence* $\theta_{\mathcal{A}l}$ of the generators of the horizon, and their *Alexandrov* null second fundamental form. In Section 3 we consider sections of horizons and their geometry, in particular we show that sections of horizons have a well defined area. We also discuss the ambiguities which arise when defining the area of those sections when the horizon is not *globally smooth*. In fact, those ambiguities have nothing to do with "very low" differentiability of horizons and arise already for piecewise smooth horizons. In Section 4 we prove positivity of the Alexandrov divergence of generators of horizons — in Section 4.1 this is done under the hypothesis of existence of a conformal completion satisfying a regularity condition, together with some global causality assumptions on the space–time; in Section 4.2 positivity of $\theta_{\mathcal{A}l}$ is established under the hypothesis that the generators of the horizon are future complete. In Section 5 we show that Alexandrov points *"propagate to the past"* along the generators of the horizon. This allows one to show that the optical equation holds on "almost all" generators of the horizon. We also present there a theorem (Theorem 5.6) which shows that "almost all generators are Alexandrov"; while this theorem belongs naturally to Section 5, its proof uses methods which are developed in Section 6 only, therefore it is deferred to Appendix C. In Section 6 we prove our main result – the monotonicity theorem, Theorem 6.1. This is done under the assumption that $\theta_{\mathcal{A}l}$ is non–negative. One of the key elements of the proof is a new (to us) extension result of Whitney type (Proposition 6.6), the hypotheses of which are rather different from the usual ones; in particular it seems to be much easier to work with in some situations. Section 7 discusses the relevance of the rigidity part of Theorem 6.1 to the theory of black holes and to the differentiability of compact Cauchy horizons. Appendix A reviews the geometry of $C^2$ horizons, we also prove there a new result concerning the relationship between the (classical) differentiability of a horizon *vs* the (classical) differentiability of sections thereof, Proposition A.3. In Appendix B some comments on the area theorem of [37] are

[5]This follows from Theorem 1.1.

made.

## 2 Horizons

Let $(M, g)$ be a smooth spacetime, that is, a smooth paracompact Hausdorff time-oriented Lorentzian manifold, of dimension $n+1 \geq 3$, with a smooth Lorentzian metric $g$. Throughout this paper hypersurfaces are assumed to be embedded. A hypersurface $\mathcal{H} \subset M$ will be said to be *future null geodesically ruled* if every point $p \in \mathcal{H}$ belongs to a future inextensible null geodesic $\Gamma \subset \mathcal{H}$; those geodesics will be called *the generators* of $\mathcal{H}$. We emphasize that the generators are allowed to have past endpoints on $\mathcal{H}$, but no future endpoints. *Past null geodesically ruled* hypersurfaces are defined by changing the time orientation. We shall say that $\mathcal{H}$ is *a future (past) horizon* if $\mathcal{H}$ is an *achronal, closed, future (past) null geodesically ruled topological hypersurface.* A hypersurface $\mathcal{H}$ will be called a *horizon* if $\mathcal{H}$ is a future or a past horizon. Our terminology has been tailored to the black hole setting, so that a future black hole event horizon $\partial J^-(\mathscr{I}^+)$ is a future horizon in the sense just described [37, p. 312]. The terminology is somewhat awkward in a Cauchy horizon setting, in which a *past Cauchy horizon* $\mathcal{D}^-(\Sigma)$ of an achronal edgeless set $\Sigma$ is a *future* horizon in our terminology [56, Theorem 5.12].

Let $\dim M = n+1$ and suppose that $\mathcal{O}$ is a domain in $\mathbb{R}^n$. Recall that a continuous function $f\colon \mathcal{O} \to \mathbb{R}$ is called semi–convex iff each point $p$ has a convex neighborhood $\mathcal{U}$ in $\mathcal{O}$ so that there exists a $C^2$ function $\phi\colon \mathcal{U} \to \mathbb{R}$ such that $f + \phi$ is convex in $\mathcal{U}$. We shall say that the graph of $f$ is a semi–convex hypersurface if $f$ is semi–convex. A hypersurface $\mathcal{H}$ in a manifold $M$ will be said to be semi–convex if $\mathcal{H}$ can be covered by coordinate patches $\mathcal{U}_\alpha$ such that $\mathcal{H} \cap \mathcal{U}_\alpha$ is a semi–convex graph for each $\alpha$. The interest of this notion stems from the specific differentiability properties of such hypersurfaces:

**Proposition 2.1 (Alexandrov [26, Appendix E])** *Let $B$ be an open subset of $\mathbb{R}^p$ and let $f\colon B \to \mathbb{R}$ be semi–convex. Then there exists a set $B_{\mathcal{A}l} \subset B$ such that:*

1. *the $p$ dimensional Lebesgue measure $\mathfrak{L}^p(B \setminus B_{\mathcal{A}l})$ of $B \setminus B_{\mathcal{A}l}$ vanishes.*

2. *$f$ is differentiable at all points $x \in B_{\mathcal{A}l}$,* i.e.,

$$\begin{aligned} &\forall\, x \in B_{\mathcal{A}l}\ \exists x^* \in (\mathbb{R}^p)^* \text{ such that} \\ &\quad \forall\, y \in B \quad f(y) - f(x) = x^*(y-x) + r_1(x,y)\ , \end{aligned} \tag{2.1}$$

   *with $r_1(x,y) = o(|x-y|)$. The linear map $x^*$ above will be denoted by $df(x)$.*

3. *$f$ is twice–differentiable at all points $x \in B_{\mathcal{A}l}$ in the sense that*

$$\begin{aligned} &\forall\, x \in B_{\mathcal{A}l}\ \exists\, Q \in (\mathbb{R}^p)^* \otimes (\mathbb{R}^p)^* \text{ such that } \forall\, y \in B \\ &\quad f(y) - f(x) - df(x)(y-x) = Q(x-y, x-y) + r_2(x,y)\ , \end{aligned} \tag{2.2}$$

*with $r_2(x,y) = o(|x-y|^2)$. The symmetric quadratic form $Q$ above will be denoted by $\frac{1}{2}D^2 f(x)$, and will be called the* second Alexandrov derivative *of $f$ at $x$.*[6]

The points $q$ at which Equations (2.1)–(2.2) hold will be called *Alexandrov points* of $f$, while the points $(q, f(q))$ will be called *Alexandrov points of the graph of $f$* (it will be shown in Proposition 2.5 below that if $q$ is an Alexandrov point of $f$, then $(q, f(q))$ will project to an Alexandrov point of any graphing function of the graph of $f$, so this terminology is meaningful).

We shall say that $\mathcal{H}$ is *locally achronal* if for every point $p \in \mathcal{H}$ there exists a neighborhood $\mathcal{O}$ of $p$ such that $\mathcal{H} \cap \mathcal{O}$ is achronal in $\mathcal{O}$. We have the following:

**Theorem 2.2** *Let $\mathcal{H} \neq \varnothing$ be a locally achronal future null geodesically ruled hypersurface. Then $\mathcal{H}$ is semi–convex.*

**Remark 2.3** An alternative proof of Theorem 2.2 can be given using a variational characterization of horizons (compare [4, 33, 57]).

**Remark 2.4** We note that while the notion of convexity of a function and its epigraph is coordinate dependent, that of semi–convexity is not; indeed, the proof given below is based on the equivalence of semi–convexity and of existence of lower support hypersurfaces with locally uniform one side Hessian bounds. The latter is clearly independent of the coordinate systems used to represent $f$, or its graph, as long as the relevant orientation is preserved (changing $y^{n+1}$ to $-y^{n+1}$ transforms a lower support hypersurface into an upper one). Further, if the future of $\mathcal{H}$ is represented as a graph in two different ways,

$$\begin{aligned} J^+(\mathcal{H}) &= \{x^{n+1} \geq f(x^1, \ldots, x^n)\} \\ &= \{y^{n+1} \geq g(y^1, \ldots, y^n)\} , \end{aligned} \tag{2.3}$$

then semi–convexity of $f$ is equivalent to that of $g$. This follows immediately from the considerations below. Thus, the notion of a semi–convex hypersurface is not tied to a particular choice of coordinate systems and of graphing functions used to represent it.

Proof: Let $\mathcal{O}$ be as in the definition of local achronality; passing to a subset of $\mathcal{O}$ we can without loss of generality assume that $\mathcal{O}$ is globally hyperbolic. Replacing the space–time $(M, g)$ by $\mathcal{O}$ with the induced metric we can without loss of generality assume that $(M, g)$ is globally hyperbolic. Let $t$ be a time function on $\mathcal{O}$ which induces a diffeomorphism of $\mathcal{O}$ with $\mathbb{R} \times \Sigma$ in the standard

[6] Caffarelli and Cabrè [9] use the term "second punctual differentiability of $f$ at $x$" for (2.2); Fleming and Soner [26, Definition 5.3, p. 234] use the name "point of twice–differentiability" for points at which (2.2) holds.

way [32, 59], with the level sets $\Sigma_\tau \equiv \{p \mid t(p) = \tau\}$ of $t$ being Cauchy surfaces. As usual we identify $\Sigma_0$ with $\Sigma$, and in the identification above the curves $\mathbb{R}\times\{q\}$, $q \in \Sigma$, are integral curves of $\nabla t$. Define

$$\Sigma_{\mathcal{H}} = \{q \in \Sigma \mid \mathbb{R} \times \{q\} \text{ intersects } \mathcal{H}\} \ . \tag{2.4}$$

For $q \in \Sigma_{\mathcal{H}}$ the set $(\mathbb{R} \times \{q\}) \cap \mathcal{H}$ is a point by achronality of $\mathcal{H}$, we shall denote this point by $(f(q), q)$. Thus an achronal hypersurface $\mathcal{H}$ in a globally hyperbolic space–time is a graph over $\Sigma_{\mathcal{H}}$ of a function $f$. The map which to a point $p \in \mathcal{H}$ assigns $q \in \Sigma$ such that $p = (f(q), q)$ is injective, so that the hypothesis that $\mathcal{H}$ is a topological hypersurface together with the invariance of the domain theorem (*cf., e.g.,* [18, Prop. 7.4, p. 79]) imply that $\Sigma_{\mathcal{H}}$ is open. We wish to use [1, Lemma 3.2][7] to obtain semi–convexity of the (local) graphing function $f$, this requires a construction of lower support hypersurfaces at $p$. Let $\Gamma$ be a generator of $\mathcal{H}$ passing through $p$ and let $p_+ \in \Gamma \cap J^+(p)$. By achronality of $\Gamma$ there are no points on $J^-(p_+) \cap J^+(p)$ which are conjugate to $p_+$, *cf.* [37, Prop. 4.5.12, p. 115] or [5, Theorem 10.72, p. 391]. It follows that the intersection of a sufficiently small neighborhood of $p$ with the past light cone $\partial J^-(p_+)$ of $p_+$ is a smooth hypersurface contained in $J^-(\mathcal{H})$. This provides the appropriate lower support hypersurface at $p$. In particular, for suitably chosen points $p_+$, these support hypersurfaces have null second fundamental forms (see Appendix A) which are locally (in the point $p$) uniformly bounded from below. This in turn implies that the Hessians of the associated graphing functions are locally bounded from below, as is needed to apply Lemma 3.2 in [1]. □

Theorem 2.2 allows us to apply Proposition 2.1 to $f$ to infer twice Alexandrov differentiability of $\mathcal{H}$ almost everywhere, in the sense made precise in Proposition 2.1. Let us denote by $\mathfrak{L}^n_{h_0}$ the $n$ dimensional Riemannian measure[8] on $\Sigma \equiv \Sigma_0$, where $h_0$ is the metric induced on $\Sigma_0$ by $g$, then there exists a set $\Sigma_{\mathcal{H}_{\mathcal{A}l}} \subset \Sigma_{\mathcal{H}}$ on which $f$ is twice Alexandrov differentiable, and such that $\mathfrak{L}^n_{h_0}(\Sigma_{\mathcal{H}} \setminus \Sigma_{\mathcal{H}_{\mathcal{A}l}}) = 0$. Set

$$\mathcal{H}_{\mathcal{A}l} \equiv \text{graph of } f \text{ over } \Sigma_{\mathcal{H}_{\mathcal{A}l}} \ . \tag{2.5}$$

Point 1 of Proposition 2.1 shows that $\mathcal{H}_{\mathcal{A}l}$ has a tangent space at every point $p \in \mathcal{H}_{\mathcal{A}l}$. For those points define

$$k(p) = k_\mu(p) dx^\mu = -dt + df(q) \ , \qquad p = (f(q), q), \qquad q \in \Sigma_{\mathcal{H}_{\mathcal{A}l}} \ . \tag{2.6}$$

A theorem of Beem and Królak shows that $\mathcal{H}$ is differentiable precisely at those points $p$ which belong to exactly one generator $\Gamma$ [6, 14], with $T_p\mathcal{H} \subset T_pM$ being the null hyper-plane containing $\dot\Gamma$. It follows that

$$K \equiv g^{\mu\nu} k_\mu \partial_\nu \tag{2.7}$$

[7]The result in that Lemma follows actually from [2, Lemma 2.15] together with [20, Prop. 5.4, p. 24].

[8]In local coordinates, $d\mathfrak{L}^n_{h_0} = \sqrt{\det h_0}\, d^n x$.

is null, future pointing, and tangent to the generators of $\mathcal{H}$, wherever defined.

We define the generators of $\mathcal{H}_{\mathcal{A}l}$ as the intersections of the generators of $\mathcal{H}$ with $\mathcal{H}_{\mathcal{A}l}$. Point 2 of Proposition 2.1 allows us to define the *divergence* $\theta_{\mathcal{A}l}$ of those generators, as follows: Let $e_i$, $i = 1, \ldots, n$ be a basis of $T_p\mathcal{H}$ such that

$$g(e_1, e_1) = \cdots = g(e_{n-1}, e_{n-1}) = 1 \ , \qquad g(e_i, e_j) = 0 \ , i \neq j \ , \qquad e_n = K \ . \quad (2.8)$$

We further choose the $e_a$'s, $a = 1, \ldots, n-1$ to be orthogonal to $\nabla t$. It follows that the vectors $e_a = e^\mu_a \partial_\mu$'s have no $\partial/\partial t$ components in the coordinate system used in the proof of Theorem 2.2, thus $e^0_a = 0$. Using this coordinate system for $p \in \mathcal{H}_{\mathcal{A}l}$ we set

$$i, j = 1, \ldots, n \qquad \nabla_i k_j = D^2_{ij} f - \Gamma^\mu_{ij} k_\mu \ , \quad (2.9)$$

$$\theta_{\mathcal{A}l} = (e^i_1 e^j_1 + \cdots + e^i_{n-1} e^j_{n-1}) \nabla_i k_j \ . \quad (2.10)$$

It is sometimes convenient to set $\theta_{\mathcal{A}l} = 0$ on $\mathcal{H} \setminus \mathcal{H}_{\mathcal{A}l}$. The function $\theta_{\mathcal{A}l}$ so defined on $\mathcal{H}$ will be called the divergence (towards the future) of both the generators of $\mathcal{H}_{\mathcal{A}l}$ and of $\mathcal{H}$. It coincides with the usual divergence of the generators of $\mathcal{H}$ at every set $\mathcal{U} \subset \mathcal{H}$ on which the horizon is of $C^2$ differentiability class: Indeed, in a space–time neighborhood of $\mathcal{U}$ we can locally extend the $e_i$'s to $C^1$ vector fields, still denoted $e_i$, satisfying (2.8). It is then easily checked that the set of Alexandrov points of $\mathcal{U}$ is $\mathcal{U}$, and that the divergence $\theta$ of the generators as defined in [37, 64] or in Appendix A coincides on $\mathcal{U}$ with $\theta_{\mathcal{A}l}$ as defined by (2.10).

The *null second fundamental form* $B_{\mathcal{A}l}$ of $\mathcal{H}$, or of $\mathcal{H}_{\mathcal{A}l}$, is defined as follows: in the basis above, if $X = X^a e_a$, $Y = Y^b e_b$ (the sums are from 1 to $n-1$), then at Alexandrov points we set

$$B_{\mathcal{A}l}(X, Y) = X^a Y^b e^i_a e^j_b \nabla_i k_j. \quad (2.11)$$

with $\nabla_i k_j$ defined by (2.9). This coincides[9] with the usual definition of $B$ as discussed *e.g.* in Appendix A on any subset of $\mathcal{H}$ which is $C^2$. In this definition of the null second fundamental form $B$ with respect to the null direction $K$, $B$ measures expansion as positive and contraction as negative.

The definitions of $\theta_{\mathcal{A}l}$ and $B_{\mathcal{A}l}$ given in Equations (2.10)–(2.11) have, so far, been only given for horizons which can be globally covered by an appropriate coordinate system. As a first step towards a globalization of those notions one needs to find out how those objects change when another representation is chosen. We have the following:

**Proposition 2.5** *1. Let $f$ and $g$ be two locally Lipschitz graphing functions representing $\mathcal{H}$ in two coordinate systems $\{x^i\}_{i=1,\ldots,n+1}$ and $\{y^i\}_{i=1,\ldots,n+1}$,*

[9] More precisely, when $\mathcal{H}$ is $C^2$ equation (2.11) defines, in local coordinates, a tensor field which reproduces $b$ defined in Appendix A when passing to the quotient $\mathcal{H}/K$.

*related to each other by a $C^2$ diffeomorphism $\phi$. If $(x_0^1, \dots x_0^n)$ is an Alexandrov point of $f$ and*

$$(y_0^1, \dots, y_0^n, g(y_0^1, \dots, y_0^n)) = \phi(x_0^1, \dots, x_0^n, f(x_0^1, \dots, x_0^n)) ,$$

*then $(y_0^1, \dots, y_0^n)$ is an Alexandrov point of $g$.*

2. *The null second fundamental form $B_{\mathcal{A}l}$ of Equation (2.11) is invariantly defined modulo a point dependent multiplicative factor. In particular the sign of $\theta_{\mathcal{A}l}$ defined in Equation (2.10) does not depend upon the choice of the graphing function used in (2.10).*

**Remark 2.6** Recall that the standard divergence $\theta$ of generators is defined up to a multiplicative function (constant on the generators) only, so that (essentially) the only geometric invariant associated to $\theta$ is its sign.

**Remark 2.7** We emphasize that we do not assume that $\partial/\partial x^{n+1}$ and/or $\partial/\partial y^{n+1}$ are timelike.

PROOF: 1. Let $\vec{y} = (y^1, \dots, y^n)$, $\vec{x}_0 = (x_0^1, \dots, x_0^n)$, *etc.*, set $f_0 = f(\vec{x}_0)$, $g_0 = g(\vec{y}_0)$. Without loss of generality we may assume $(\vec{x}_0, f(\vec{x}_0)) = (\vec{y}_0, f(\vec{y}_0)) = 0$. To establish (2.2) for $g$, write $\phi$ as $(\vec{\phi}, \phi^{n+1})$, and let

$$\vec{\psi}(\vec{x}) := \vec{\phi}(\vec{x}, f(\vec{x})) .$$

As $(id, f)$ and $(id, g)$, where $id$ is the identity map, are bijections between neighborhoods of zero and open subsets of the graph, invariance of domain shows that $\vec{\psi}$ is a homeomorphism from a neighborhood of zero to its image. Further, $\vec{\psi}$ admits a Taylor development of order two at the origin:

$$\vec{\psi}(\vec{x}) = L_1\vec{x} + \alpha_1(\vec{x}, \vec{x}) + o(|\vec{x}|^2) , \tag{2.12}$$

where

$$L_1 = D_{\vec{x}}\vec{\phi}(0) + D_{x^{n+1}}\vec{\phi}(0)\, Df(0) .$$

Let us show that $L_1$ is invertible: suppose, for contradiction, that this is not the case, let $\vec{x} \neq 0$ be an element of the kernel $\ker L_1$, and for $|s| << 1$ let $\vec{y}_s := \vec{\psi}(s\vec{x}) = O(s^2)$. As $\vec{y}_s = \vec{\phi}(s\vec{x}, f(s\vec{x}))$ we have

$$g(\vec{y}_s) = \phi^{n+1}(s\vec{x}, f(s\vec{x})) = s[D_{\vec{x}}\phi^{n+1}(0) + D_{x^{n+1}}\phi^{n+1}(0)Df(0)]\vec{x} + o(s) .$$

The coefficient of the term linear in $s$ does not vanish, otherwise $(\vec{x}, Df(0)\vec{x})$ would be a non–zero vector of ker $D\phi(0)$. We then obtain $|g(\vec{y}_s)|/|\vec{y}_s| \geq C/|s| \to_{s\to 0} \infty$, which contradicts the hypothesis that $g$ is locally Lipschitz.

To finish the proof, for $\vec{y}$ close to the origin let $\vec{x} := \psi^{-1}(\vec{y})$, thus $\vec{y} = \psi(\vec{x})$, and from (2.12) we infer

$$\vec{x} = L_1^{-1}\vec{y} - L_1^{-1}\alpha_1(L_1^{-1}\vec{y}, L_1^{-1}\vec{y}) + o(|\vec{y}|^2) \ . \tag{2.13}$$

Equation (2.2) and twice differentiability of $\phi$ show that $\phi^{n+1}(\cdot, f(\cdot))$ has a second order Alexandrov expansion at the origin:

$$\phi^{n+1}(\vec{x}, f(\vec{x})) = L_2\vec{x} + \alpha_2(\vec{x}, \vec{x}) + o(|\vec{x}|^2) \ . \tag{2.14}$$

Finally from $g(\vec{y}) = \phi^{n+1}(\vec{x}, f(\vec{x}))$ and Equations (2.13) and (2.14) we get

$$g(\vec{y}) = L_2L_1^{-1}\vec{y} + (\alpha_2 - L_2L_1^{-1}\alpha_1)(L_1^{-1}\vec{y}, L_1^{-1}\vec{y}) + o(|\vec{y}|^2) \ .$$

2. The proof of point 1. shows that under changes of graphing functions the Alexandrov derivatives $D_{ij}f$ transform into the $D_{ij}g$'s exactly as they would if $f$ and $g$ were twice differentiable. The proof of point 2. reduces therefore to that of the analogous statement for $C^2$ functions, which is well known. □

Proposition 2.5 shows that Equation (2.11) defines an equivalence class of tensors $B_{\mathcal{A}l}$ at every Alexandrov point of $\mathcal{H}$, where two tensors are identified when they are proportional to each other with a positive proportionality factor. Whenever $\mathcal{H}$ can not be covered by a global coordinate system required in Equation (2.11), one can use Equation (2.11) in a local coordinate patch to define a representative of the appropriate equivalence class, and the classes so obtained will coincide on the overlaps by Proposition 2.5. From this point of view $\theta_{\mathcal{A}l}(p)$ can be thought as the assignment, to an Alexandrov point of the horizon $p$, of the number $0, \pm 1$, according to the sign of $\theta_{\mathcal{A}l}(p)$.

In some situations it is useful to renormalize the vector field $K$ defined almost everywhere by (2.7); this leads to a different point of view concerning $B_{\mathcal{A}l}$, not an equivalence class of tensors, but an object defined with respect to an appropriately scaled field of vectors $K$. For example, the change of coordinates discussed in Proposition 2.5 will change the graphing function of $\mathcal{H}$, which will lead to a new vector field $K$, equal to the old vector field $K$ multiplied by a function $\phi$; yet another normalization of $K$ is used in the proof of Proposition A.5, Appendix A. If $\phi$ is differentiable almost everywhere (which will be the case if *e.g.* $\phi$ is Lipschitz), at the points of differentiability of $\phi$ one can formally use the chain rule in (2.11) to see that $B_{\mathcal{A}l}$ defined with respect to $\phi K$ is $\phi$ times the $B_{\mathcal{A}l}$ defined with respect to $K$. Whatever the differentiability properties of $\phi$, one can use this as *the definition of $B_{\mathcal{A}l}$ defined with respect to $\phi K$*; such a definition is then consistent with the usual properties of differentiable objects whenever those properties apply.

The generators of a horizon $\mathcal{H}$ play an important part in our results. The following shows that most points of a horizon are on just one generator.

**Proposition 2.8** *The set of points of a horizon $\mathcal{H}$ that are on more than one generator has vanishing $n$ dimensional Hausdorff measure.*

PROOF: Let $\mathcal{T}$ be the set of points of $\mathcal{H}$ that are on two or more generators of $\mathcal{H}$. By a theorem of Beem and Królak [6, 14] a point of $\mathcal{H}$ is differentiable if and only if it belongs to exactly one generator of $\mathcal{H}$. Therefore no point of $\mathcal{T}$ is differentiable and thus no point of $\mathcal{T}$ is an Alexandrov point of $\mathcal{H}$. Whence Theorem 2.2 and Proposition 2.1 imply $\mathcal{H}^n(\mathcal{T}) = 0$. □

**Remark 2.9** For a future horizon $\mathcal{H}$ in a spacetime $(M, g)$ let $\mathcal{E}$ be the set of past endpoints of generators of $\mathcal{H}$. Then the set of points $\mathcal{T}$ of $\mathcal{H}$ that are on two or more generators is a subset of $\mathcal{E}$. From the last Proposition we know that $\mathfrak{H}^n(\mathcal{T}) = 0$ and it is tempting to conjecture $\mathfrak{H}^n(\mathcal{E}) = 0$. However this seems to be an open question. We note that there are examples [15] of horizons $\mathcal{H}$ so that $\mathcal{E}$ is dense in $\mathcal{H}$.

# 3 Sections of horizons

Recall that in the standard approach to the area theorem [37] one considers sections $\mathcal{H}_\tau$ of a black hole event horizon $\mathcal{H}$ obtained by intersecting $\mathcal{H}$ with the level sets $\Sigma_\tau$ of a time function $t$,

$$\mathcal{H}_\tau = \mathcal{H} \cap \Sigma_\tau \ . \tag{3.1}$$

We note that Theorem 2.2 and Proposition 2.1 do not directly yield any information about the regularity of all the sections $\mathcal{H}_\tau$, since Proposition 2.1 gives regularity away from a set of zero measure, but those sections are expected to be of zero measure anyway. Now, because the notion of semi–convexity is independent of the choice of the graphing function, and it is preserved by taking restrictions to lower dimensional subsets, we have:

**Proposition 3.1** *Let $\Sigma$ be an embedded $C^2$ spacelike hypersurface and let $\mathcal{H}$ be a horizon in $(M, g)$, set*

$$S = \Sigma \cap \mathcal{H} \ .$$

*There exists a semi–convex topological submanifold $S_{reg} \subset S$ of co–dimension two in $M$ which contains all points $p \in S$ at which $\mathcal{H}$ is differentiable.*

**Remark 3.2** "Most of the time" $S_{\rm reg}$ has full measure in $S$. There exist, however, sections of horizons for which this is not the case; *cf.* the discussion at the end of this section, and the example discussed after Equation (3.12) below.

PROOF: Let $p \in S$ be such that $\mathcal{H}$ is differentiable at $p$, let $\mathcal{O}_p$ be a coordinate neighborhood of $p$ such that $\Sigma \cap \mathcal{O}_p$ is given by the equation $t = 0$, and such that $\mathcal{H} \cap \mathcal{O}_p$ is the graph $t = f$ of a semi–convex function $f$. If we write $p = (f(q), q)$, then $f$ is differentiable at $q$. We wish, first, to show that $S$ can be locally represented as a Lipschitz graph, using the Clarke's implicit function theorem [17, Chapter 7]. Consider the Clarke differential $\partial f$ of $f$ at $q$ (*cf.* [17, p. 27]); by [17, Prop. 2.2.7] and by standard properties of convex (hence of semi–convex) functions we have $\partial f(q) = \{df(q)\}$. By a rigid rotation of the coordinate axes we may assume that $df(q)(\partial/\partial x^n) \neq 0$. If we write $q = (0, q^1, \ldots, q^n)$, then Clarke's implicit function theorem [17, Corollary, p. 256] shows that there exists a neighborhood $\mathcal{W}_q \subset \{(t, x^1, \ldots, x^{n-1}) \in \mathbb{R}^n\}$ of $(0, q^1, \ldots, q^{n-1})$ and a Lipschitz function $g$ such that, replacing $\mathcal{O}_p$ by a subset thereof if necessary, $\mathcal{H} \cap \mathcal{O}_p$ can be written as a graph of $g$ over $\mathcal{W}_q$. Remark 2.4 establishes semi–convexity of $g$. Now, semi–convexity is a property which is preserved when restricting a function to a smooth lower dimensional submanifold of its domain of definition, which shows that $h \equiv g|_{t=0}$ is semi–convex. It follows that the graph $\mathcal{V}_q \subset \mathcal{H}$ of $h$ over $\mathcal{W}_q \cap \{t = 0\}$ is a semi–convex topological submanifold of $\mathcal{H}$. The manifold

$$S_{\rm reg} \equiv \cup_q \mathcal{V}_q \ ,$$

where the $q$'s run over the set of differentiability points of $\mathcal{H}$ in $\Sigma$, has all the desired properties. □

Let $S$, $S_{\rm reg}$, $\mathcal{H}$ and $\Sigma$ be as in Proposition 3.1. Now $S_{\rm reg}$ and $\mathcal{H}$ are both semi–convex, and have therefore their respective Alexandrov points; as $S_{\rm reg} \subset \mathcal{H}$, it is natural to inquire about the relationship among those. It should be clear that if $p \in S$ is an Alexandrov point of $\mathcal{H}$, then it also is an Alexandrov point of $S$. The following example shows that the converse is not true: Let $\mathcal{H} \subset R^{1,2}$ be defined as the set $\mathcal{H} = \{(t, x, y) \mid t = |x|\}$, and let $\Sigma_\tau$ be the level sets of the Minkowskian time in $\mathbb{R}^{1,2}$, $\Sigma_\tau = \{(t, x, y) \mid t = \tau\}$. Each section $S_\tau = \mathcal{H} \cap \Sigma_\tau$ is a smooth one dimensional submanifold of $R^{1,2}$ (not connected if $\tau > 0$, empty if $\tau < 0$). In particular all points of $S_0 = \{t = x = 0\}$ are Alexandrov points thereof, while none of them is an Alexandrov point of $\mathcal{H}$. In this example the horizon $\mathcal{H}$ is not differentiable on any of the points of $S_0$, which is a necessary condition for being an Alexandrov point of $\mathcal{H}$. It would be of some interest to find out whether or not Alexandrov points of sections of $\mathcal{H}$ which are also points of differentiability of $\mathcal{H}$ are necessarily Alexandrov points of $\mathcal{H}$.

Following [29], we shall say that a differentiable embedded hypersurface $\mathcal{S}$ meets $\mathcal{H}$ *properly transversally* if for each point $p \in \mathcal{S} \cap \mathcal{H}$ for which $T_p\mathcal{H}$ exists the tangent hyperplane $T_p\mathcal{S}$ is transverse to $T_p\mathcal{H}$. If $\mathcal{S}$ is spacelike and intersects $\mathcal{H}$ proper transversality will always hold; on the other hand if $\mathcal{S}$ is timelike this might, but does not have to be the case. If $\mathcal{H}$ and $\mathcal{S}$ are $C^1$ and $\mathcal{S}$ is either spacelike or timelike intersecting $\mathcal{H}$ transversely then $\mathcal{H} \cap \mathcal{S}$ is a $C^1$ spacelike submanifold. Therefore when $\mathcal{S}$ is timelike there is a spacelike $\mathcal{S}_1$ so that $\mathcal{H} \cap \mathcal{S} = \mathcal{H} \cap \mathcal{S}_1$.

It would be interesting to know if this was true (even locally) for the transverse intersection of general horizons $\mathcal{H}$ with timelike $C^2$ hypersurfaces $\mathcal{S}$.

Let $\mathcal{S}$ be any spacelike or timelike $C^2$ hypersurface in $M$ meeting $\mathcal{H}$ properly transversally.[10] Suppose, first, that $\mathcal{S}$ is covered by a single coordinate patch such that $\mathcal{S}\cap\mathcal{H}$ is a graph $x^n = g(x^1,\ldots,x^{n-1})$ of a semi–convex function $g$. Let $\mathbf{n}$ be the field of unit normals to $\mathcal{S}$; at each point $(x^1,\ldots,x^{n-1})$ at which $g$ is differentiable and for which $T_p\mathcal{H}$ exists, where

$$p \equiv (x^1,\ldots,x^{n-1}, g(x^1,\ldots,x^{n-1})) \ ,$$

there is a unique number $a(p) \in \mathbb{R}$ and a unique future pointing null vector $K \in T_p\mathcal{H}$ such that

$$\langle K - a\mathbf{n}, \cdot\rangle = dx^n - dg$$

We set

$$k_\mu dx^\mu = \langle K, \cdot\rangle \in T^*_pM \ .$$

Assume moreover that $(x^1,\ldots,x^{n-1})$ is an Alexandrov point of $g$, we thus have the Alexandrov second derivatives $D^2 g$ at our disposal. Consider $e_i$ — a basis of $T_p\mathcal{H}$ as in Equation (2.8), satisfying further $e_i \in T_p\mathcal{S}\cap T_p\mathcal{H}$, $i = 1,\ldots,n-1$. In a manner completely analogous to (2.9)–(2.10) we set

$$i,j = 1,\ldots,n-1 \qquad \nabla_i k_j = D^2_{ij} g - \Gamma^\mu_{ij} k_\mu \ , \tag{3.2}$$

$$\theta^{\mathcal{S}\cap\mathcal{H}}_{\mathcal{A}l} = (e^i_1 e^j_1 + \cdots + e^i_{n-1} e^j_{n-1})\nabla_i k_j \ . \tag{3.3}$$

(Here, as in (2.9), the $\Gamma^\mu_{ij}$'s are the Christoffel symbols of the space–time metric $g$.) For $X, Y \in T_p\mathcal{S}\cap T_p\mathcal{H}$ we set, analogously to (2.11),

$$B^{\mathcal{S}\cap\mathcal{H}}_{\mathcal{A}l}(X,Y) = X^a Y^b e^i_a e^j_b \nabla_i k_j. \tag{3.4}$$

Similarly to the definitions of $\theta_{\mathcal{A}l}$ and $B_{\mathcal{A}l}$, the vector $K$ with respect to which $\theta^{\mathcal{S}\cap\mathcal{H}}_{\mathcal{A}l}$ and $B^{\mathcal{S}\cap\mathcal{H}}_{\mathcal{A}l}$ have been defined has been tied to the particular choice of coordinates used to represent $\mathcal{S}\cap\mathcal{H}$ as a graph. In order to globalize this definition it might be convenient to regard $B^{\mathcal{S}\cap\mathcal{H}}_{\mathcal{A}l}(p)$ as an equivalence class of tensors defined up to a positive multiplicative factor. Then $\theta^{\mathcal{S}\cap\mathcal{H}}_{\mathcal{A}l}(p)$ can be thought as the assignment to a point $p$ of the number $0, \pm 1$, according to the sign of $\theta^{\mathcal{S}\cap\mathcal{H}}_{\mathcal{A}l}(p)$. This, together with Proposition 2.5, can then be used to define $B^{\mathcal{S}\cap\mathcal{H}}_{\mathcal{A}l}$ and $\theta^{\mathcal{S}\cap\mathcal{H}}_{\mathcal{A}l}$ for $\mathcal{S}$ which are not globally covered by a single coordinate patch.

[10] Our definitions below makes use of a unit normal to $\mathcal{S}$, whence the restriction to spacelike or timelike, properly transverse $\mathcal{S}$'s. Clearly one should be able to give a definition of $\theta^{\mathcal{S}\cap\mathcal{H}}_{\mathcal{A}l}$, *etc.*, for any hypersurface $\mathcal{S}$ intersecting $\mathcal{H}$ properly transversally. Now for a smooth $\mathcal{H}$ and for smooth properly transverse $\mathcal{S}$'s the intersection $\mathcal{S}\cap\mathcal{H}$ will be a smooth spacelike submanifold of $M$, and it is easy to construct a spacelike $\mathcal{S}'$ so that $\mathcal{S}'\cap\mathcal{H} = \mathcal{S}\cap\mathcal{H}$. Thus, in the smooth case, no loss of generality is involved by restricting the $\mathcal{S}$'s to be spacelike. For this reason, and because the current setup is sufficient for our purposes anyway, we do not address the complications which arise when $\mathcal{S}$ is allowed to be null, or to change type.

We mention that in the current setup one can multiply $K$ by a function $\phi$ so that the one–form $\phi(dx^n - dg)$ has unit length with respect to the metric induced by $g$ on $\mathcal{S}$, and redefine $\theta_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}$ and $B_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}$ by multiplying the previous definitions by $\phi$ (see the discussion at the end of Section 2, which also applies here). Such an approach presents some advantages: one can tie $B_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}$ to the second fundamental form of $\mathcal{S}\cap\mathcal{H}$ in $\mathcal{S}$ and to the second fundamental form of $\mathcal{S}$ in $M$. Further, using Proposition 2.5 one can obtain objects which are globally defined on $\mathcal{S}\cap\mathcal{H}$, without assuming that $\mathcal{S}$ can be covered by a global coordinate chart. We note, however, that if $\mathcal{S}\cap\mathcal{H}$ is represented as $\mathcal{S}'\cap\mathcal{H}$, for a second hypersurface $\mathcal{S}'$ satisfying our hypotheses, then (whether or not normalized as above) the new $\theta_{\mathcal{A}l}^{\mathcal{S}'\cap\mathcal{H}}$ and $B_{\mathcal{A}l}^{\mathcal{S}'\cap\mathcal{H}}$ will in general differ from $\theta_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}$ and $B_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}$ by a (positive) multiplicative factor, so it is more natural in any case to think of $B_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}$ as of an equivalence class of tensors defined at each point (almost everywhere) up to a (positive) multiplicative factor.

As already pointed out, if $p\in\mathcal{S}\cap\mathcal{H}$ is an Alexandrov point of $\mathcal{H}$ then $p$ will also be an Alexandrov point of $\mathcal{S}\cap\mathcal{H}$. In such a case the equivalence class of $B_{\mathcal{A}l}$, defined at $p$ by Equation (2.11), will coincide with that defined by (3.4), when (2.11) is restricted to vectors $X, Y\in T_p\mathcal{S}\cap T_p\mathcal{H}$. Similarly the sign of $\theta_{\mathcal{A}l}(p)$ will coincide with that of $\theta_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}(p)$, and $\theta_{\mathcal{A}l}(p)$ will vanish if and only if $\theta_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}(p)$ does.

Let us turn our attention to the question, how to define the area of sections of horizons. The monotonicity theorem we prove in Section 6 uses the Hausdorff measure, so let us start by pointing out the following:

**Proposition 3.3** *Let $\mathcal{H}$ be a horizon and let $\mathcal{S}$ be an embedded hypersurface in $M$. Then $\mathcal{S}\cap\mathcal{H}$ is a Borel set, in particular it is $\nu$–Hausdorff measurable for any $\nu\in\mathbb{R}^+$.*

PROOF: Let $\sigma$ be any complete Riemannian metric on $M$, we can cover $\mathcal{S}$ by a countable collection of sets $\mathcal{O}_i\subset\mathcal{S}$ of the form $\mathcal{O}_i = B_\sigma(p_i, r_i)\cap\mathcal{S}$, where the $B_\sigma(p_i, r_i)$ are open balls of $\sigma$–radius $r_i$ centered at $p_i$ with compact closure. We have $\mathcal{O}_i = \left(\overline{B_\sigma(p_i, r_i)\cap\mathcal{S}}\right)\setminus\partial B_\sigma(p_i, r_i)$, which shows that the $\mathcal{O}_i$'s are Borel sets. Since $\mathcal{S}\cap\mathcal{H} = \cup_i(\mathcal{O}_i\cap\mathcal{H})$, the Borel character of $\mathcal{S}\cap\mathcal{H}$ ensues. The Hausdorff measurability of $\mathcal{S}\cap\mathcal{H}$ follows now from the fact that Borel sets are Hausdorff measurable[11]. $\Box$

Proposition 3.3 is sufficient to guarantee that a notion of area of sections of horizons — namely their $(n-1)$–dimensional Hausdorff area — is well defined. Since the Hausdorff area is not something very handy to work with in practice,

[11] Let $(X, d)$ be a metric space. Then an outer measure $\mu$ defined on the class of all subsets of $X$ is a *metric outer measure* iff $\mu(A\cup B) = \mu(A) + \mu(B)$ whenever the distance (*i.e.* $\inf_{a\in A, b\in B} d(a, b)$) is positive. For a metric outer measure, $\mu$, the Borel sets are all $\mu$-measurable (*cf.* [36, p. 48 Prob. 8] or [39, p. 188 Exercise 1.48]). The definition of the Hausdorff outer measures implies they are metric outer measures (*cf.* [39, p. 188 Exercise 1.49]). See also [61, p. 7] or [19, p. 147].

it is convenient to obtain more information about regularity of those sections. Before we do this let us shortly discuss how area can be defined, depending upon the regularity of the set under consideration. When $S$ is a piecewise $C^1$, $(n-1)$ dimensional, paracompact, orientable submanifold of $M$ this can be done by first defining

$$d^{n-1}\mu = e_1 \wedge \cdots \wedge e_{n-1} \ ,$$

where the $e_a$'s form an oriented orthonormal basis of the cotangent space $T^*S$; obviously $d^{n-1}\mu$ does not depend upon the choice of the $e_a$'s. Then one sets

$$\mathrm{Area}(S) = \int_S d^{n-1}\mu \ . \tag{3.5}$$

Suppose, next, that $S$ is the image by a Lipschitz map $\psi$ of a $C^1$ manifold $N$. Let $h_S$ denote any complete Riemannian metric on $N$; by [61, Theorem 5.3] for every $\epsilon > 0$ there exists a $C^1$ map $\psi_\epsilon$ from $N$ to $M$ such that

$$\mathfrak{L}^{n-1}_{h_S}(\{\psi \neq \psi_\epsilon\}) \leq \epsilon \ . \tag{3.6}$$

Here and throughout $\mathfrak{L}^{n-1}_{h_S}$ denotes the $(n-1)$ dimensional Riemannian measure associated with a metric $h_S$. One then sets $S_\epsilon = \psi_\epsilon(N)$ and

$$\mathrm{Area}(S) = \lim_{\epsilon\to 0} \mathrm{Area}(S_\epsilon) \ . \tag{3.7}$$

(It is straightforward to check that Area($S$) so defined is independent of the choice of the sequence $\psi_\epsilon$. In particular if $\psi$ is $C^1$ on $N$ one recovers the definition (3.5) using $\psi_\epsilon = \psi$ for all $\epsilon$.)

It turns out that for general sections of horizons some more work is needed. Throughout this paper $\sigma$ will be an auxiliary Riemannian metric such that $(M, \sigma)$ is a complete Riemannian manifold. Let $S$ be a Lipschitz $(n-1)$ dimensional submanifold of $M$ such that $S \subset \mathcal{H}$; recall that, by Rademacher's theorem, $S$ is differentiable $\mathfrak{H}^{n-1}_\sigma$ almost everywhere. Let us denote by $\mathfrak{H}^s_\sigma$ the $s$ dimensional Hausdorff measure [24] defined using the distance function of $\sigma$. Recall that $S$ is called *$n$ rectifiable* [24] iff $S$ is the image of a bounded subset of $\mathbb{R}^n$ under a Lipschitz map. A set is *countably $n$ rectifiable* iff it is a countable union of $n$ rectifiable sets. (*Cf.* [24, p. 251].) (Instructive examples of countably rectifiable sets can be found in [52].) We have the following result [42] (The proof of the first part of Proposition 3.4 is given in Remark 6.13 below):

**Proposition 3.4** *Let $S$ be as in Proposition 3.1, then $S$ is countably $(n-1)$ rectifiable. If $S$ is compact then it is $(n-1)$ rectifiable.*

Consider, then, a family of sets $\mathcal{V}_{q_i}$ which are Lipschitz images and form a partition of $S$ up to a set of $\mathfrak{H}^{n-1}_h$ Hausdorff measure zero, where $h$ is the metric on $\Sigma$ induced by $g$. One sets

$$\mathrm{Area}(S) = \sum_i \mathrm{Area}(\mathcal{V}_{q_i}) \ . \tag{3.8}$$

We note that $A$ so defined again does not depend on the choices made, and reduces to the previous definitions whenever applicable. Further, $\mathrm{Area}(S)$ so defined is[12] precisely the $(n-1)$ dimensional Hausdorff measure $\mathfrak{H}_h^{n-1}$ of $S$:

$$\mathrm{Area}(S) = \int_S d\mathfrak{H}_h^{n-1}(p) = \mathfrak{H}_h^{n-1}(S) \ . \tag{3.9}$$

As we will see, there is still another quantity which appears naturally in the area theorem, Theorem 6.1 below: the area *counting multiplicities*. Recall that the *multiplicity* $N(p)$ of a point $p$ belonging to a horizon $\mathcal{H}$ is defined as the number (possibly infinite) of generators passing through $p$. Similarly, given a subset $S$ of $\mathcal{H}$, for $p \in \mathcal{H}$ we set

$$\begin{aligned} N(p,S) = {} & \text{the number of generators of } \mathcal{H} \text{ passing through} \\ & p \text{ which meet } S \text{ when followed to the (causal) future} \ . \end{aligned} \tag{3.10}$$

Note that $N(p,S) = N(p)$ for $p \in S$. Whenever $N(p,S_2)$ is $\mathfrak{H}_{h_1}^{n-1}$ measurable on the intersection $S_1$ of $\mathcal{H}$ with a spacelike hypersurface $\Sigma_1$ we set

$$\mathrm{Area}_{S_2}(S_1) = \int_{S_1} N(p,S_2)\, d\mathfrak{H}_{h_1}^{n-1}(p) \ , \tag{3.11}$$

where $h_1$ is the metric induced by $g$ on $\Sigma_1$. Note that $N(p,S_2) = 0$ at points of $S_1$ which have the property that the generators through them do not meet $S_2$; thus the area $\mathrm{Area}_{S_2}(S_1)$ only takes into account those generators that are seen from $S_2$. If $S_1 \subset J^-(S_2)$, as will be the case *e.g.* if $S_2$ is obtained by intersecting $\mathcal{H}$ with a Cauchy surface $\Sigma_2$ lying to the future of $S_1$, then $N(p,S_2) \geq 1$ for all $p \in S_1$ (actually in that case we will have $N(p,S_2) = N(p)$), so that

$$\mathrm{Area}_{S_2}(S_1) \geq \mathrm{Area}(S_1) \ .$$

Let us show, by means of an example, that the inequality can be strict in some cases. Consider a black hole in a three dimensional space–time, suppose that its section by a spacelike hypersurface $t = 0$ looks as shown in Figure 1. As we are in three dimensions area should be replaced by length. (A four dimensional analogue of Figure 1 can be obtained by rotating the curve from Figure 1 around a vertical axis.) When a slicing by spacelike hypersurfaces is appropriately chosen, the behavior depicted can occur when two black holes merge together[13]. When

[12] The equality (3.9) can be established using the area formula. In the case of $C^1$ submanifolds of Euclidean space this is done explicitly in [61, p. 48]. This can be extended to countable $n$ rectifiable sets by use of more general version of the area theorem [61, p. 69] (for subsets of Euclidean space) or [23, Theorem 3.1] (for subsets of Riemannian manifolds).

[13] The four dimensional analogue of Figure 1 obtained by rotating the curve from Figure 1 around a *vertical* axis can occur in a time slicing of a black–hole space–time in which a "temporarily toroidal black hole" changes topology from toroidal to spherical. Ambiguities related

Figure 1: A section of the event horizon in a $2+1$ dimensional space–time with "two black holes merging".

measuring the length of the curve in Figure 1 one faces various options: a) discard the middle piece altogether, as it has no interior; b) count it once; c) count it twice — once from each side. The purely differential geometric approach to area, as given by Equation (3.5) does not say which choice should be made. The Hausdorff area approach, Equation (3.9), counts the middle piece once. The prescription (3.11) counts it twice. We wish to argue that the most reasonable prescription, from an entropic point of view, is to use the prescription (3.11). In order to do that, let $(M,g)$ be the three dimensional Minkowski space–time and consider a thin long straw $R$ lying on the $y$ axis in the hypersurface $t=0$:

$$R = \{t = x = 0, y \in [-10, 10]\} \ . \tag{3.12}$$

Set $K = J^+(R) \cap \{t=1\}$, then $K$ is a convex compact set which consists of a strip of width 2 lying parallel to the straw with two half–disks of radius 1 added at the ends. Let $\hat{M} = M \setminus K$, equipped with the Minkowski metric, still denoted by $g$. Then $\hat{M}$ has a black hole region $\mathcal{B}$ which is the past domain of dependence of $K$ in $M$ (with $K$ removed). The sections $\mathcal{H}_\tau$ of the event horizon $\mathcal{H}$ defined as $\mathcal{H} \cap \{t=\tau\}$ are empty for $\tau < 0$ and $\tau \geq 1$. Next, $\mathcal{H}_0$ consists precisely of the straw. Finally, for $t \in (0,1)$ $\mathcal{H}_t$ is the union of a strip of width $2t$ lying parallel to the straw with two half–disks of radius $t$ added at the ends of the strip. Thus

$$\mathrm{Area}(\mathcal{H}_t) = \int_{\mathcal{H}_t} d\mathfrak{H}^1_h = \begin{cases} 0 \ , & t < 0 \ , \\ 10 \ , & t = 0 \ , \\ 20 + 2\pi t \ , & 0 < t < 1 \ , \end{cases} \tag{3.13}$$

while

$$\int_{\mathcal{H}_t} N(p)\, d\mathfrak{H}^1_h = \begin{cases} 0 \ , & t < 0 \ , \\ 20 \ , & t = 0 \ , \\ 20 + 2\pi t \ , & 0 < t < 1 \ . \end{cases} \tag{3.14}$$

(The end points of the straw, at which the multiplicities are infinite, do not contribute to the above integrals with $t=0$, having vanishing measure.) We see

to the definition of area and/or area discontinuities do occur in this example. On the other hand, the four dimensional analogue of Figure 1 obtained by rotating the curve from Figure 1 around a *horizontal* axis can occur in a time slicing of a black–hole space–time in which two black holes merge together. There are neither obvious ambiguities related to the definition of the area nor area discontinuities in this case.

that Area($\mathcal{H}_t$) jumps once when reaching zero, and a second time immediately thereafter, while the "area counting multiplicities" (3.14) jumps only once, at the time of formation of the black hole. From an entropy point of view the existence of the first jump can be explained by the formation of a black hole in a very unlikely configuration: as discussed below events like that can happen only for a negligible set of times. However, the second jump of Area($\mathcal{H}_t$) does not make any sense, and we conclude that (3.14) behaves in a more reasonable way. We note that the behavior seen in Figure 1 is obtained by intersecting $\mathcal{H}$ by a spacelike surface which coincides with $t = 0$ in a neighborhood of the center of the straw $R$ and smoothly goes up in time away from this neighborhood.

As already mentioned it footnote 13, this example easily generalizes to 3 + 1 dimensions: To obtain a 3 + 1 dimensional model with a similar discontinuity in the cross sectional area function, for each fixed $t$, rotate the curve $\mathcal{H}_t$ about a *vertical* axis in $x$–$y$–$z$ space. This corresponds to looking at the equi–distant sets to a disk. The resulting model is a flat space model for the Hughes *et al.* temporarily toroidal black hole [43], *cf.* also [30, 44, 60].

We now show that the behavior exhibited in Figure 1, where there is a set of points of positive measure with multiplicity $N(p) \geq 2$, can happen for at most a *negligible set of times.*

**Proposition 3.5** *Let $M$ be a spacetime with a global $C^1$ time function $\tau\colon M \to \mathbb{R}$ and for $t \in \mathbb{R}$ let $\Sigma_t = \{p \in M : \tau(p) = t\}$ be a level set of $\tau$. Then for any past or future horizon $\mathcal{H} \subset M$*

$$\mathfrak{H}^{n-1}_{h_t}(\{p \in \Sigma_t : N(p) \geq 2\}) = 0$$

*for almost all $t \in \mathbb{R}$. (Where $h_t$ is the induced Riemannian metric on $\Sigma_t$.) For these $t$ values*

$$\int_{\mathcal{H}\cap\Sigma_t} N(p)\, d\mathfrak{H}^{n-1}(p) = \mathfrak{H}^{n-1}_{h_t}(\mathcal{H} \cap \Sigma_t) \ ,$$

*so that for almost all $t$ the area of $\mathcal{H} \cap \Sigma_t$ counted with multiplicity is the same as the $n-1$ dimensional Hausdorff measure of $\mathcal{H} \cap \Sigma_t$.*

PROOF: Let $\mathcal{H}_{\text{sing}}$ be the set of points of $\mathcal{H}$ that are on more than one generator. By Proposition 2.8 $\mathfrak{H}^n_\sigma(\mathcal{H}_{\text{sing}}) = 0$ (where $\sigma$ is an auxiliary complete Riemannian metric $\sigma$ on $M$). The Hausdorff measure version of Fubini's theorem, known as the "co–area formula" [61, Eq. 10.3, p. 55], gives

$$\int_{\mathbb{R}} \mathfrak{H}^{n-1}_\sigma(\mathcal{H}_{\text{sing}} \cap \Sigma_t)\, dt = \int_{\mathbb{R}} \mathfrak{H}^{n-1}_\sigma(\tau|_{\mathcal{H}}^{-1}[t])\, dt = \int_{\mathcal{H}_{\text{sing}}} J(\tau|_{\mathcal{H}})\, d\mathfrak{H}^n_\sigma = 0 \ . \quad (3.15)$$

Here $J(\tau|_{\mathcal{H}})$ is the Jacobian of the function $\tau$ restricted to $\mathcal{H}$ and the integral over $\mathcal{H}_{\text{sing}}$ vanishes as $\mathfrak{H}^n_\sigma(\mathcal{H}_{\text{sing}}) = 0$.

Equation (3.15) implies that for almost all $t \in \mathbb{R}$ we have $\mathfrak{H}^{n-1}_\sigma(\mathcal{H} \cap \Sigma_t) = 0$. But for any such $t$ we also have $\mathfrak{H}^{n-1}_{h_t}(\mathcal{H} \cap \Sigma_t) = 0$. (This can be seen by noting that the identity map between $\Sigma_t$ with the metric of $\sigma$ restricted to $\Sigma_t$ and $(\Sigma_t, h_t)$ and is locally Lipschitz. And locally Lipschitz maps send sets of $n-1$ dimensional Hausdorff measure zero to sets with $n-1$ dimensional Hausdorff measure zero.) This completes the proof. $\Box$

We note that the equation $\mathfrak{H}^{n-1}_\sigma(\mathcal{H}_{\rm sing} \cap \Sigma_t) = 0$ for almost all $t$'s shows that the set $(S_t)_{\rm reg} \subset S_t \equiv \Sigma_t \cap \mathcal{H}$ given by Proposition 3.1 has full measure in $S_t$ for almost all $t$'s. It is not too difficult to show (using the Besicovitch covering theorem) that $(S_t)_{\rm reg}$ is countably rectifiable which gives a proof, alternative to Proposition 3.4, of countable rectifiablity (up to a negligible set) of $S_t$ for almost all $t$'s. We emphasize, however, that Proposition 3.4 applies to all sections of $\mathcal{H}$.

# 4 Positivity of $\theta_{\mathcal{A}l}$

The proof of the area theorem consists of two rather distinct steps: the first is to show the non–negativity of the divergence of the generators of event horizons under appropriate conditions, the other is to use this result to conclude that the area of sections is nondecreasing towards the future. In this section we shall establish positivity of $\theta_{\mathcal{A}l}$.

## 4.1 Causally regular conformal completions

A pair $(\bar{M}, \bar{g})$ will be called a *conformal completion* of $(M, g)$ if $\bar{M}$ is a manifold with boundary such that $M$ is the interior of $\bar{M}$. The boundary of $\bar{M}$ will be called $\mathscr{I}$. We shall further suppose that there exists a function $\Omega$, positive on $M$ and differentiable on $\bar{M}$, which vanishes on $\mathscr{I}$, with $d\Omega$ *nowhere vanishing* on $\mathscr{I}$. We emphasize that no assumptions about the causal nature of Scri are made. We shall also require that the metric $\Omega^{-2}g$ extends by continuity to $\mathscr{I}$, in such a way that the resulting metric $\bar{g}$ on $\bar{M}$ is differentiable. We set

$$\mathscr{I}^+ = \{p \in \mathscr{I} \mid I^-(p; \bar{M}) \cap M \neq \varnothing\} \ .$$

In the results presented below $\mathscr{I}^+$ is not required to be connected.

Recall that a space–time is said to satisfy the *null energy condition*, or the *null convergence condition* if

$$\mathrm{Ric}(X, X) \geq 0 \tag{4.1}$$

for all null vectors $X$. The following Proposition spells out some conditions which guarantee positivity of the Alexandrov divergence of the generators of $\mathcal{H}$; as already pointed out, we do not assume that $\mathscr{I}^+$ is null:

**Proposition 4.1** *Let $(M,g)$ be a smooth space–time with a conformal completion $(\bar M,\bar g) = (M\cup \mathscr{I}^+,\Omega^2 g)$ of $(M,g)$ and suppose that the null energy condition holds on the past $I^-(\mathscr{I}^+;\bar M)\cap M$ of $\mathscr{I}^+$ in $M$. Set*

$$\mathcal{H} = \partial J^-(\mathscr{I}^+;\bar M)\cap M \ .$$

*Suppose that there exists a neighborhood $\mathcal{O}$ of $\mathcal{H}$ with the following property: for every compact set $C\subset\mathcal{O}$ that meets $I^-(\mathscr{I}^+;\bar M)$ there exists a future inextendible (in $M$) null geodesic $\eta\subset\partial J^+(C;M)$ starting on $C$ and having future end point on $\mathscr{I}^+$. Then*

$$\theta_{\mathcal{A}l}\ge 0 \ .$$

*Further, if $\mathcal{S}$ is any $C^2$ spacelike or timelike hypersurface which meets $\mathcal{H}$ properly transversally, then*

$$\theta_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}\ge 0 \quad \text{on } \mathcal{S}\cap\mathcal{H} \ .$$

We remark that Królak [49] has obtained a related result, but with the null expansion $\theta$ assumed to be defined classically, in a black hole setting of a more general nature. He makes an assumption analogous to the assumption of the existence of the null geodesic generator $\eta$ of $\partial J^+(C;M)$, which he refers to as *weak regularity* of the event horizon. Conditions ensuring the existence of such an $\eta$ are described later in this section. Some further results concerning positivity of the classical $\theta$ can be found in [48].

PROOF: Suppose that the result does not hold. Let, first, $p_0$ be an Alexandrov point of $\mathcal{H}$ with $\theta_{\mathcal{A}l}<0$ at $p_0$ and consider a neighborhood $\mathcal{N}$ of $p_0$ of the form $\Sigma\times\mathbb{R}$, constructed like the set $\mathcal{O}$ (not to be confused with the set $\mathcal{O}$ in the statement of the present proposition) in the proof of Theorem 2.2, so that $\mathcal{H}\cap\mathcal{N}$ is a graph over $\Sigma$ of a function $f\colon\Sigma\to\mathbb{R}$, define $x_0$ by $p_0=(x_0,f(x_0))$. By point 3. of Proposition 2.1 in a coordinate neighborhood $\mathcal{U}\subset\Sigma$ the function $f$ can be written in the form

$$f(x) = f(x_0)+df(x_0)(x-x_0)+\frac12 D^2f(x_0)(x-x_0,x-x_0)+o(|x-x_0|^2) \ . \quad (4.2)$$

After a translation, a rotation and a rescaling we will have

$$x_0=0\ , \quad f(x_0)=0\ , \quad df(x_0)=dx^n \ . \quad (4.3)$$

Let $B^{n-1}(\delta)\subset\mathbb{R}^{n-1}$ be the $(n-1)$ dimensional open ball of radius $\delta$ centered at the origin, for $q\in B^{n-1}(\delta)$, set $x=(q,0)$ and

$$f_{\epsilon,\eta}(q) = f(x_0)+df(x_0)(x-x_0)+\frac12 D^2f(x_0)(x-x_0,x-x_0)+\epsilon|x-x_0|^2+\eta \ . \quad (4.4)$$

Define

$$S_{\epsilon,\eta,\delta} = \{\ \text{graph over } (B^{n-1}(\delta)\times\{0\})\cap\mathcal{U} \text{ of } f_{\epsilon,\eta}\} \ . \quad (4.5)$$

If its parameters are chosen small enough, $S_{\epsilon,\eta,\delta}$ will be a smooth spacelike submanifold of $M$ of co–dimension two. Let $\theta_{\epsilon,\eta,\delta}$ be the mean extrinsic curvature of $S_{\epsilon,\eta,\delta}$, from (4.4) when all the parameters are smaller in absolute value than some thresholds one will have

$$|\theta_{\epsilon,\eta,\delta}(p_0) - \theta_{\mathcal{A}l}(p_0)| \leq C\epsilon$$

for some constant $C$, so that if $\epsilon$ is chosen small enough we will have

$$\theta_{\epsilon,\eta,\delta} < 0 \tag{4.6}$$

at $p_0$; as $\theta_{\epsilon,\eta,\delta}$ is continuous Equation (4.6) will hold throughout $S_{\epsilon,\eta,\delta}$ if $\delta$ is chosen small enough, for all sufficiently small $\epsilon$'s and $\eta$'s. If $\delta$ is small enough and $\eta = 0$ we have

$$\overline{S_{\epsilon,\eta=0,\delta}} \subset J^+(\mathcal{H}) \ , \qquad \overline{S_{\epsilon,\eta=0,\delta}} \cap \mathcal{H} = \{p_0\} \ .$$

Here $\overline{S_{\epsilon,\eta,\delta}}$ denotes the closure of $S_{\epsilon,\eta,\delta}$. It follows that for all sufficiently small strictly negative $\eta$'s we will have

$$S_{\epsilon,\eta,\delta} \cap J^-(\mathscr{I}^+) \neq \varnothing \ , \tag{4.7}$$

$$\overline{S_{\epsilon,\eta,\delta}} \setminus S_{\epsilon,\eta,\delta} \subset J^+(\mathcal{H}) \ , \tag{4.8}$$

Making $\delta$ smaller if necessary so that $S_{\epsilon,\eta,\delta} \subset \mathcal{O}$, our condition on $\mathcal{O}$ implies that there exists a null geodesic $\Gamma\colon [0,1] \to \bar{M}$ such that $\Gamma(0) \in \overline{S_{\epsilon,\eta,\delta}}$, $\Gamma|_{[0,1)} \subset \partial J^+(S_{\epsilon,\eta,\delta}; M)$, and $\Gamma(1) \in \mathscr{I}^+$. The behavior of null geodesics under conformal rescalings of the metric (*cf., e.g.,* [37, p. 222]) guarantees that $\Gamma|_{[0,1)}$ is a complete geodesic in $M$. Suppose that $\Gamma(0) \in \overline{S_{\epsilon,\eta,\delta}} \setminus S_{\epsilon,\eta,\delta}$, then, by Equation (4.8), $\Gamma$ would be a causal curve from $J^+(\mathcal{H})$ to $\mathscr{I}^+$, contradicting the hypothesis that $\mathcal{H} = \partial J^-(\mathscr{I}^+)$. It follows that $\Gamma(0) \in S_{\epsilon,\eta,\delta}$. Equation (4.6) and standard results [55, Theorem 51, p. 298] imply that $\Gamma$ enters $I^+(S_{\epsilon,\eta,\delta})$ and therefore cannot lie in $\partial J^+(S_{\epsilon,\eta,\delta}; M)$, which gives a contradiction, and positivity of $\theta_{\mathcal{A}l}$ follows.

The argument to establish positivity of $\theta_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}$ is an essentially identical (and somewhat simpler) version of the above. Suppose, thus, that the claim about the sign of $\theta_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}$ is wrong, then there exists a point $q_0 \in \mathcal{S}$ which is an Alexandrov point of $\mathcal{S} \cap \mathcal{H}$ and at which $\theta_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}}$ is negative. Consider a coordinate patch around $q_0 = 0$ such that $\mathcal{S} \cap \mathcal{H}$ is a graph $x^n = g(x^1, \ldots, x^{n-1})$ of a semi–convex function $g$. Regardless of the spacelike/timelike character of $\mathcal{S}$ we can choose the coordinates, consistently with semi–convexity of $g$, so that $J^-(\mathcal{H})$ lies under the graph of $g$ . The definition (4.4) is replaced by

$$f_{\epsilon,\eta}(q) = g(q_0) + dg(q_0)(q-q_0) + \frac{1}{2}D^2g(q_0)(q-q_0, q-q_0) + \epsilon|q-q_0|^2 + \eta \ , \tag{4.9}$$

while (4.5) is replaced by

$$S_{\epsilon,\eta,\delta} = \{ \text{ graph over } B^{n-1}(\delta) \text{ of } f_{\epsilon,\eta} \} \ . \tag{4.10}$$

The other arguments used to prove that $\theta_{\mathcal{A}l} \geq 0$ go through without modifications. $\Box$

Proposition 4.1 assumes the existence of a neighborhood of the horizon with some precise global properties, and it is natural to look for global conditions which will ensure that such a neighborhood exists. The simplest way to guarantee that is to assume that the conformal completion $(\bar{M}, \bar{g})$ is globally hyperbolic, perhaps as a manifold with boundary. To be precise, we shall say that a manifold $(\bar{M}, \bar{g})$, with or without boundary, is called globally hyperbolic if there exists a smooth time function $t$ on $\bar{M}$, and if $J^+(p) \cap J^-(q)$ is compact for all $p, q \in \bar{M}$ (compare [30]). For example, Minkowski space–time with the standard $\mathscr{I}^+$ attached is a globally hyperbolic manifold with boundary. (However, if both $\mathscr{I}^+$ and $\mathscr{I}^-$ are attached, then it is not. Note that if $\mathscr{I}^+$ and $\mathscr{I}^-$ and $i^0$ are attached to Minkowski space–time, then it is not a manifold with boundary any more. Likewise, the conformal completions considered in [64] are not manifolds with boundary.) Similarly Schwarzschild space–time with the standard $\mathscr{I}^+$ attached to it is a globally hyperbolic manifold with boundary. Further, the standard conformal completions of de–Sitter, or anti–de Sitter space–time [34] as well as those of the Kottler space–times [46] (sometimes called Schwarzchild – de Sitter and Schwarzchild – anti–de Sitter space–times) and their generalizations considered in [7] are globally hyperbolic manifolds with boundary.

It is well known that in globally hyperbolic manifolds the causal hypotheses of Proposition 4.1 are satisfied, therefore we have proved:

**Proposition 4.2** *Under the condition c) of Theorem 1.1, the conclusions of Proposition 4.1 hold.*

The hypothesis of global hyperbolicity of $(\bar{M}, \bar{g})$ is esthetically unsatisfactory, as it mixes conditions concerning the physical space–time $(M, g)$ together with conditions concerning an artificial boundary one attaches to it. On the other hand it seems sensible to treat on a different footing the conditions concerning $(M, g)$ and those concerning $\mathscr{I}^+$. We wish to indicate here a possible way to do this. In order to proceed further some terminology will be needed:

**Definition 4.3** 1. *A point $q$ in a set $A \subset B$ is said to be a* past *point*[14] *of $A$ with respect to $B$ if $J^-(q; B) \cap A = \{q\}$.*

2. *We shall say that $\mathscr{I}^+$ is $\mathcal{H}$–*regular *if there exists a neighborhood $\mathcal{O}$ of $\mathcal{H}$ such that for every compact set $C \subset \mathcal{O}$ satisfying $I^+(C; \bar{M}) \cap \mathscr{I}^+ \neq \varnothing$ there exists a past point* with respect to *$\mathscr{I}^+$ in $\partial I^+(C; \bar{M}) \cap \mathscr{I}^+$.*

[14] A very similar notion has already been used in [11, p. 102].

We note, as is easily verified[15], that $\partial I^+(C;\bar M)\cap \mathscr{I}^+ = \partial(I^+(C;\bar M)\cap\mathscr{I}^+)$, where the boundary on the right hand side is meant as a subset of $\mathscr{I}^+$.

**Remark 4.4** In the null and timelike cases the purpose of the condition is to exclude pathological situations in which the closure of the domain of influence of a compact set $C$ in $M$ contains points which are arbitrarily far in the past on $\mathscr{I}^+$ in an uncontrollable way. A somewhat similar condition has been first introduced in [31] for null $\mathscr{I}^+$'s in the context of topological censorship, and has been termed "the $i^0$ avoidance condition" there. We shall use the term "$\mathcal{H}$–regular" instead, to avoid the misleading impression that we assume existence of an $i^0$.

**Remark 4.5** When $\mathscr{I}^+$ is *null* throughout the condition of $\mathcal{H}$*–regularity* is equivalent to the following requirement: there exists a neighborhood $\mathcal{O}$ of $\mathcal{H}$ such that for every compact set $C\subset\mathcal{O}$ satisfying $\overline{I^+(C;\bar M)}\cap\mathscr{I}^+\neq\varnothing$ there exists at least one generator of $\mathscr{I}^+$ which intersects $\overline{I^+(C;\bar M)}$ and leaves it when followed to the past[16]. This condition is satisfied by the standard conformal completions of Minkowski space–time, or of the Kerr–Newman space–times. In fact, in those examples, when $\mathcal{O}$ is suitably chosen, for every compact set $C\subset\mathcal{O}$ satisfying $\overline{I^+(C;\bar M)}\cap\mathscr{I}^+\neq\varnothing$ it holds that *every* generator of $\mathscr{I}^+$ intersects $\overline{I^+(C;\bar M)}$ and leaves it when followed to the past.

**Remark 4.6** When $\mathscr{I}^+$ is *timelike* throughout the condition of $\mathcal{H}$*–regularity* is equivalent to the following requirement: there exists a neighborhood $\mathcal{O}$ of $\mathcal{H}$ such that for every compact set $C\subset\mathcal{O}$ satisfying $\overline{I^+(C;\bar M)}\cap\mathscr{I}^+\neq\varnothing$, there is a point $x$ in $\overline{I^+(C;\bar M)\cap\mathscr{I}^+}$ such that every past inextendible causal curve in $\mathscr{I}^+$ from $x$ leaves $\overline{I^+(C;\bar M)}$. (The standard conformal completions of anti–de Sitter space–time [34] as well as those of the negative $\Lambda$ Kottler space–times [46] and their generalizations considered in [7] all satisfy this condition.) Under this condition the existence of a past point in $\partial I^+(C;\bar M)\cap\mathscr{I}^+$ can be established as follows: Let $p\in\partial I^+(C;\bar M)\cap J^-(x;\mathscr{I}^+)$, let $\gamma$ be a causal curve entirely contained in $\partial I^+(C;\bar M)\cap\mathscr{I}^+$ through $p$; if no such curves exist then $p$ is a past point of $\partial I^+(C;\bar M)\cap\mathscr{I}^+$ and we are done; if such curves exist, by a standard construction that involves Zorn's lemma we can without loss of generality assume that $\gamma$ is

[15]The inclusion "$\subset$" makes use of the fact that if $q\in\mathscr{I}^+$, then for any $p\in M$ near $q$ there exists $p'\in\mathscr{I}^+\cap I^+(p;\bar M)$ near $q$. If $\mathscr{I}^+$ is null (and hence possibly type changing) at $q$ one sees this as follows. From the definition of $\mathscr{I}^+$, there exists at $q$ a future directed outward pointing timelike vector $X_0$. This can be extended in a neighborhood $U$ of $q$ to a future directed timelike vector field $X$ everywhere transverse to $\mathscr{I}^+$. By flowing along the integral curves of $X$, we see that any point $p\in M$ sufficiently close to $q$ is in the timelike past of a point $p'\in\mathscr{I}^+$ close to $q$.

[16]Those points at which the generators of $\mathscr{I}^+$ exit $\overline{I^+(C;\bar M)}$ are past points of $\partial I^+(C;\bar M)\cap\mathscr{I}^+$.

past inextendible *in* $\partial I^+(C;\bar M)\cap\mathscr{I}^+$. Let $\Gamma$ be a past inextendible causal curve *in* $\mathscr{I}^+$ which contains $\gamma$. The current causal regularity condition on $\mathscr{I}^+$ implies that $\Gamma$ has an end point $q$ on $\partial I^+(C;\bar M)\cap\mathscr{I}^+$. If $q$ were not a past point of $\partial I^+(C;\bar M)\cap\mathscr{I}^+$ one could extend $\gamma$ as a causal curve in $\partial I^+(C;\bar M)\cap\mathscr{I}^+$, which would contradict the maximality of $\gamma$. Hence $q$ is a past point of $\partial I^+(C;\bar M)\cap\mathscr{I}^+$.

**Remark 4.7** When $\mathscr{I}^+$ is *spacelike* throughout the condition of $\mathcal{H}$*–regularity* is equivalent to the requirement of existence of a neighborhood $\mathcal{O}$ of $\mathcal{H}$ such that for every compact set $C\subset\mathcal{O}$ satisfying $I^+(C;\bar M)\cap\mathscr{I}^+\neq\varnothing$, the set $\overline{I^+(C;\bar M)}$ does not contain all of $\mathscr{I}^+$. We note that the standard conformal completions of de Sitter space–time, as well as those of the positive–$\Lambda$ generalized Kottler space–times [34], satisfy our $\mathcal{H}$–regularity condition.

Let us present our first set of conditions which guarantees that the causal hypotheses of Proposition 4.1 hold:

**Proposition 4.8** *Let $(M,g)$ be a spacetime with a $\mathcal{H}$–regular $\mathscr{I}^+$, and suppose that there exists in $M$ a partial Cauchy surface $\Sigma$ such that*

*(i)* $I^+(\Sigma;M)\cap I^-(\mathscr{I}^+;\bar M)\subset D^+(\Sigma;M)$, *and*

*(ii)* $\mathscr{I}^+\subset I^+(\Sigma;\bar M)$.

*Then there is a neighborhood $\mathcal{O}$ of $\mathcal{H}$ such that if $C$ is a compact subset of $J^+(\Sigma;M)\cap\mathcal{O}$ that meets $I^-(\mathscr{I}^+;\bar M)$, there exists a future inextendible (in $M$) null geodesic $\eta\subset\partial I^+(C;M)$ starting at a point on $C$ and having future end point on $\mathscr{I}^+$.*

**Remark 4.9** The conditions (i) and (ii) in Proposition 4.8 form a version of *"asymptotic predictability"*. They *do not* imply that global hyperbolicity extends to the horizon or $\mathscr{I}^+$. These conditions are satisfied in the set-up of [37].

Proof: Choose $\mathcal{O}$ as in the definition of $\mathcal{H}$–regularity; then for compact $C\subset\mathcal{O}$ satisfying $I^+(C;\bar M)\cap\mathscr{I}^+\neq\varnothing$ there is a past point $q$ on $\partial I^+(C;\bar M)\cap\mathscr{I}^+$. By [64, Theorem 8.1.6, p. 194] (valid in the present context) there exists a causal curve $\eta\subset\partial I^+(C;\bar M)$ with future end point $q$ which either is past inextendible in $\bar M$ or has a past end point on $C$. As $q$ is a past point with respect to $\mathscr{I}^+$, $\eta$ must meet $M$. Note that, due to the potentially unusual shape of $\mathscr{I}^+$, in principle $\eta$ may meet $\mathscr{I}^+$ after $q$ when followed backwards in time even infinitely often in a finite interval. The following argument avoids this difficulty altogether: we consider any connected component $\eta_0$ of $\eta\cap M$; $\eta_0$ is a null geodesic with future end point on $\mathscr{I}^+$. Now, conditions (i) and (ii) imply that $\eta_0$ enters the interior of $D^+(\Sigma;M)$. If $\eta_0$ is past inextendible in $M$, then, following $\eta_0$ into the past, $\eta_0$ must meet and enter the timelike past of $\Sigma$. This leads to an achronality violation of $\Sigma$: Choose $p \in \eta_0 \cap I^-(\Sigma; M)$. Then, since $I^+(\partial I^+(C;M);M) \subset I^+(C;M)$, by moving slightly to the future of $p$, we can find a point $p' \in I^+(C;M) \cap I^-(\Sigma;M) \subset I^+(\Sigma;M) \cap I^-(\Sigma;M)$. We conclude that $\eta_0$ has a past end point on $C$. $\square$

The conditions of the proposition that follows present an alternative to those of Proposition 4.8; they form a version of *"strong asymptotic predictability"*:

**Proposition 4.10** *Let $(M,g)$ be a spacetime with a $\mathcal{H}$–regular $\mathscr{I}^+$, and suppose that $M$ contains a causally simple domain $V$ such that*

$$\overline{I^-(\mathscr{I}^+;\bar{M})} \cap M \subset V \ .$$

*Then there is a neighborhood $\mathcal{O}$ of $\mathcal{H}$ such that if $C$ is a compact subset of $V \cap \mathcal{O}$ that meets $I^-(\mathscr{I}^+;\bar{M})$, there exists a future inextendible (in $M$) null geodesic $\eta \subset \partial I^+(C;V)$ starting at a point on $C$ and having future end point on $\mathscr{I}^+$.*

**Remark 4.11** Recall, an open set $V$ in $M$ is causally simple provided for all compact subsets $K \subset V$, the sets $J^\pm(K;V)$ are closed in $V$. Causal simplicity is implied by global hyperbolicity; note, however, that the latter is not a sensible assumption in the timelike $\mathscr{I}^+$ case: For example, anti–de Sitter space is not globally hyperbolic; nevertheless it is causally simple.

**Remark 4.12** Proposition 4.10 assumes that causal simplicity extends to the horizon, but not $\mathscr{I}^+$. Its hypotheses are satisfied in the set-up of [64] and in the (full) set-up of [37]. (In both those references global hyperbolicity of $V$ is assumed.) We note that replacing $M$ by $V$ we might as well assume that $M$ is causally simple.

**Remark 4.13** If $C$ is a smooth spacelike hypersurface-with-boundary, with $\partial C = S$, then (i) $\eta$ meets $C$ at a point on $S$, (ii) $\eta$ meets $S$ orthogonally, and (iii) $\eta$ is outward pointing relative to $C$.

PROOF: Choose $\mathcal{O} \subset V$ as in the definition of $\mathcal{H}$–regularity, and let $q$ be a past point of $\partial I^+(C;\bar{M}) \cap \mathscr{I}^+$ with respect to $\mathscr{I}^+$. Then $q \in \partial I^+(C;\bar{V})$, where $\bar{V} = V \cup \mathscr{I}^+$. Arguing as in Proposition 4.8 we obtain a causal curve $\eta \subset \partial I^+(C;\bar{V})$ with future end point $q \in \mathscr{I}^+$, which is past inextendible in $\partial I^+(C;\bar{V})$. As before, $\eta$ meets $M$, and we let $\eta_0$ be a component of $\eta \cap M$. Then $\eta_0$ is a null geodesic in $V$ with future end point on $\mathscr{I}^+$ which is past inextendible in $\partial I^+(C;V)$. Since $V$ is causally simple, $\partial I^+(C;V) = J^+(C;V) \setminus I^+(C;V)$, which implies that $\eta_0$ has past end point on $C$. $\square$

If $M$ itself is globally hyperbolic, taking $V = M$ in Proposition 4.10 one obtains:

**Corollary 4.14** *Under the hypotheses of point a) of Theorem 1.1, the conclusions of Proposition 4.1 hold.*

## 4.2 Complete generators

Let us turn our attention to horizons $\mathcal{H}$ the generators of which are future complete. We emphasize that $\mathcal{H}$ is not necessarily an event horizon, and the space–time does not have to satisfy any causality conditions. We start with some terminology: consider a $C^2$ spacelike manifold $S \subset M$ of co–dimension two, let $p_0 \in S$. We can choose coordinates in a globally hyperbolic neighborhood $\mathcal{O}$ of $p_0$ such that the paths $s \to (x^1, \dots, x^n, x^{n+1} = s)$ are timelike and such that

$$S \cap \mathcal{O} = \{x^n = x^{n+1} = 0\} \ . \tag{4.11}$$

Suppose further that

$$p_0 \in S \cap \mathcal{H} \ , \qquad T_{p_0} S \subset T_{p_0} \mathcal{H} \ , \tag{4.12}$$

and that $p_0$ is an Alexandrov point of $\mathcal{H}$. Passing to a subset of $\mathcal{O}$ if necessary we may assume that $\mathcal{O} \cap \mathcal{H}$ is a graph of a function $f$, with $p_0 = (x_0, f(x_0))$. We shall say that $S$ *is second order tangent to* $\mathcal{H}$ if the above conditions hold and if in the coordinate system of Equation (4.11) we have

$$\forall \ X, Y \in T_{p_0} S \qquad D^2 f(X, Y) = 0 \ . \tag{4.13}$$

Here $D^2 f$ is the Alexandrov second derivative of $f$ at $x_0$, as in Equation (2.2). The notion of $S \subset \mathcal{S}$ being *second order tangent to a section* $\mathcal{S} \cap \mathcal{H}$ at an Alexandrov point of this section is defined in an analogous way.

We note the following result:

**Lemma 4.15** *Let $S$ be a $C^2$ spacelike manifold of co–dimension two which is tangent to a horizon $\mathcal{H}$ at $p_0 \in \mathcal{H} \cap S$. Suppose that*

1. *$p_0$ is an Alexandrov point of $\mathcal{H}$, and $S$ is second order tangent to $\mathcal{H}$ there, or*

2. *$p_0$ is an Alexandrov point of $\mathcal{S} \cap \mathcal{H}$, with $\mathcal{S} = \{x^n = 0\}$ in the coordinate system of (4.11), and $S$ is second order tangent to $\mathcal{S} \cap \mathcal{H}$ there, or*

3. *$p_0$ is an Alexandrov point of $\mathcal{S} \cap \mathcal{H}$, with $\mathcal{S} = \{x^{n+1} = 0\}$ in the coordinate system of (4.11), and $S$ is second order tangent to $\mathcal{S} \cap \mathcal{H}$ there.*

*Let $\Gamma$ be a null geodesic containing a generator $\gamma$ of $\mathcal{H}$ through $p_0$, with $\Gamma(0) = p_0$. If $\Gamma(1)$ is a focal point of $S$, then there exists $a \in [0, 1]$ such that $\Gamma(a)$ is an endpoint of $\gamma$ on $\mathcal{H}$.*

**Remark 4.16** The result is still true for more general $\mathcal{S}$'s than those considered in points 2. and 3. above, but the above results are sufficient for our purposes.

**Remark 4.17** We stress that we haven't assumed anything about the time orientation of $\Gamma$.

PROOF: If $\Gamma(0)$ is an end point of $\mathcal{H}$ there is nothing to prove, so we can suppose that $p_0$ is an interior point of the generator $\gamma$. Suppose that $\Gamma(1)$ is a focal point of $S$, it is well known[17] that for any $b > 1$ there exists a one parameter family of timelike paths $\Gamma_v: [0, b] \to M$, $|v| \le v_0$, such that

$$\Gamma_v(0) \in S \ , \qquad \Gamma_v(b) = \Gamma(b) \ ,$$
$$\forall u \in [0, b] \qquad d_\sigma(\Gamma_v(u), \Gamma(u)) \le C|v| \ , \tag{4.14}$$
$$g\Big(\frac{\partial \Gamma_v}{\partial u}, \frac{\partial \Gamma_v}{\partial u}\Big) \le -Cv^2 \ , \tag{4.15}$$

with some constant $C$. Here $d_\sigma$ denotes the distance on $M$ measured with respect to some auxiliary complete Riemannian metric $\sigma$.

Let $X$ be any smooth vector field which vanishes at $\Gamma(b)$ and which equals $\partial/\partial x^{n+1}$ in a neighborhood of $p_0$, where $x^{n+1}$ is the $(n+1)$st coordinate in the coordinate system of Equation (4.11). Let $\phi_s$ denote the (perhaps locally defined) flow of $X$, for $|s| \le 1$ we have the straightforward estimate

$$\Big| g\Big(\frac{\partial(\phi_s(\Gamma_v(u)))}{\partial u}, \frac{\partial(\phi_s(\Gamma_v(u)))}{\partial u}\Big) - g\Big(\frac{\partial \Gamma_v(u)}{\partial u}, \frac{\partial \Gamma_v(u)}{\partial u}\Big)\Big| \le C_1|s| \ , \tag{4.16}$$

for some constant $C_1$, whenever $\phi_s(\Gamma_v)(u)$ is defined. Set

$$\eta(\delta) = \sup_{q \in B^{n-1}(0,\delta)} \frac{|f(q)|}{(d_\sigma(q,0))^2} \ ,$$

where $f$ is again the graphing function of $\mathcal{H} \cap \mathcal{O}$, and $B^{n-1}(0,\delta)$ is as in the proof of Proposition 4.1; $d_\sigma(q, 0)$ should be understood in the obvious way. Under the hypotheses 1. and 2., by Equations (2.2) and (4.13) we have

$$\eta(\delta) \to_{\delta \to 0} 0 \ . \tag{4.17}$$

On the other hand, suppose that 3. holds, let $g$ be the graphing function of $\mathcal{S} \cap \mathcal{H}$ in the coordinate system of (4.11), the hypothesis that $S$ is second order tangent to $\mathcal{S} \cap \mathcal{H}$ gives

$$g(x) = o(|x|^2) \ ,$$

where $x$ is a shorthand for $(x^1, \ldots, x^{n-1})$. Now $(x, g(x)) \in \mathcal{H}$ so that $f(x, g(x)) = 0$, hence

$$|f(x, 0)| = |f(x, 0) - f(x, g(x))| \le L|g(x)| = o(|x|^2) \ ,$$

[17] The only complete proof of this fact known to us is to be found in [28, Section 2]. For the case at hand one should use the variation defined there with $c_\ell = 0$. We note the following misprints in [28, Section 2]: In Equation (2.7) $\eta_0'$ should be replaced by $N_0$ and $\eta_\ell'$ should be replaced by $N_\ell$; in Equation (2.10) $N_0$ should be replaced by $\eta_0'$ and $N_\ell$ should be replaced by $\eta_\ell'$.

where $L$ is the Lipschitz continuity constant of $f$ on $B^n(0,\delta)$. Thus Equation (4.17) holds in all cases.

Equation (4.14) shows that for $|v| \le \delta/C_2$, for some constant $C_2$, we will have $\Gamma_v(0) \in B^{n-1}(0,\delta) \times \{0\} \times \{0\} \subset S \cap \mathcal{O}$, so that $\Gamma_v(0)$ can be written as

$$\Gamma_v(0) = (x(v), x^{n+1} = 0) \ , \qquad x(v) = (q(v), x^n = 0) \ , \qquad q(v) \in B^{n-1}(0,\delta) \ .$$

Consider the point $p(v) \equiv (x(v), f(x(v)))$; by definition of $X$ we have

$$p(v) = \phi_{f(x(v))}(\Gamma_v(0)) \in \mathcal{H} \ .$$

From Equations (4.14), (4.15) and (4.16), together with the definition of $\eta$ we obtain

$$\forall u \in [0,b] \qquad g\Big(\frac{\partial(\phi_{f(x(v))}(\Gamma_v(u)))}{\partial u}, \frac{\partial(\phi_{f(x(v))}(\Gamma_v(u)))}{\partial u}\Big) \le -Cv^2 + C_1 C^2 \eta(\delta) v^2 \ .$$

This and Equation (4.17) show that for $\delta$ small enough $\phi_s(\Gamma_v)$ will be a timelike path from $p(v) \in \mathcal{H}$ to $\Gamma_v(b)$. Achronality of $\mathcal{H}$ implies that $\Gamma_v(b) \notin \mathcal{H}$, thus $\Gamma$ leaves $\mathcal{H}$ somewhere on $[0,b)$. As $b$ is arbitrarily close to 1, the result follows. □

As a Corollary of Lemma 4.15 one immediately obtains:

**Proposition 4.18** *Under the hypothesis b) of Theorem 1.1, we have*

$$\theta_{\mathcal{A}l} \ge 0 \ .$$

*Further, if $\mathcal{S}$ is any spacelike or timelike $C^2$ hypersurface that meets $\mathcal{H}$ properly transversally, then*

$$\theta_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}} \ge 0 \ .$$

PROOF: Suppose that $\theta_{\mathcal{A}l}(p_0) < 0$, for $\delta$ small enough let $S = S(\delta)$ be the manifold $S_{0,0,\delta}$ defined in Equation (4.5). Consider the future directed maximally extended null geodesic $\Gamma$ normal to $S$ which coincides, for some values of its parameters, with a generator $\gamma$ of $\mathcal{H}$ through $p_0$. Now, $\gamma$ is complete to the future by hypothesis, thus so must be the case with $\Gamma$. By well known results[17] [55, Theorem 43, p. 292], $S$ has a focal point along $\Gamma$ at finite affine distance, say 1. By Lemma 4.15 the generator $\gamma$ has a future end point, which contradicts the definition of a future horizon. Finally, is $\mathcal{S}$ is a spacelike hypersurface, or a timelike hypersurface properly transverse to $\mathcal{H}$, then the same argument with $S_{0,0,\delta}$ given by Equation (4.10) establishes $\theta_{\mathcal{A}l}^{\mathcal{S}\cap\mathcal{H}} \ge 0$. □

# 5 Propagation of Alexandrov points along generators, optical equation

The aim of this section is to show that, roughly speaking, Alexandrov points *"propagate to the past"* along the generators. Recall, now, that the Weingarten map $b$ of a smooth null hypersurface $S$ satisfies a Ricatti equation

$$b' + b^2 + R = 0. \tag{5.1}$$

Here $'$ and $R$ are defined in Appendix A, see Equation (A.3) and the paragraph that follows there. Equation (5.1) leads to the well known Raychaudhuri equation in general relativity: by taking the trace of (5.1) we obtain the following formula for the derivative of the null mean curvature $\theta = \theta(s)$ along $\eta$,

$$\theta' = -\mathrm{Ric}(\eta', \eta') - \sigma^2 - \frac{1}{n-2}\theta^2, \tag{5.2}$$

where $\sigma$, the shear scalar[18], is the trace of the square of the trace free part of $b$. Equation (5.2) is the well-known Raychaudhuri equation (for an irrotational null geodesic congruence) of relativity. This equation shows how the Ricci curvature of spacetime influences the null mean curvature of a null hypersurface. We will refer to Equation (5.1) as the *optical equation.*

Let $\sigma$ be any auxiliary complete smooth Riemannian metric on $M$ and define $UM \subset TM$ to be the bundle of $\sigma$–unit vectors tangent to $M$. Following [15] for $p \in \mathcal{H}$ we define $\mathcal{N}_p \subset U_pM$ as the collection of $\sigma$–unit, future directed vectors tangent to generators of $\mathcal{H}$ through $p$. Those vectors are necessarily lightlike and will be called *semi–tangents* to $\mathcal{H}$. We set

$$\mathcal{N} = \cup_{p\in\mathcal{H}}\mathcal{N}_p \ .$$

The main result of this section is the following:

**Theorem 5.1** *Let $\Gamma_{\mathrm{int}}$ denote the set of interior points of a generator $\Gamma$ of $\mathcal{H}$ (*i.e.*, $\Gamma$ without its end–point, if any). If $p_0 \in \Gamma_{\mathrm{int}}$ is an Alexandrov point of a section $\mathcal{S} \cap \mathcal{H}$, where $\mathcal{S}$ is an embedded spacelike or timelike $C^2$ hypersurface that intersects $\mathcal{H}$ properly transversally, then*

$$\Gamma_{\mathrm{int}} \cap J^-(p_0) \setminus \{p_0\} \subset \mathcal{H}_{Al} \ . \tag{5.3}$$

*Further the Alexandrov derivative $D^2 f$ of any graphing function of $\mathcal{H}$ varies smoothly over $\Gamma_{\mathrm{int}} \cap J^-(p_0)$, and the null Weingarten map $b_{Al}$ constructed out of $D^2 f$ in a way analogous to that presented in Appendix A satisfies the optical equation (5.1) and the Raychaudhuri equation (5.2) there.*

[18]This is one of the very few occurrences of the shear scalar (traditionally denoted by $\sigma$ in the physics literature) in our paper, we hope that this conflict of notation with the auxiliary Riemannian metric also denoted by $\sigma$ will not confuse the reader.

**Remark 5.2** It would be of interest to find out whether or not the inclusion (5.3) can be strengthened to $\Gamma_{\rm int} \subset \mathcal{H}_{\mathcal{A}l}$. It follows from Theorem 5.6 below that this last inclusion will hold for $\mathfrak{H}^{n-1}_\sigma$ almost all generators passing through any given section $\mathcal{S} \cap \mathcal{H}$ of $\mathcal{H}$, when the generators are counted by counting the points at which they meet $\mathcal{S} \cap \mathcal{H}$.

**Remark 5.3** There exist horizons containing generators on which no points are Alexandrov. As an example, let $\mathcal{H}$ be the boundary of the future of a square lying in the $t = 0$ hypersurface of three dimensional Minkowski space–time. Then $\mathcal{H}$ is a union of portions of null planes orthogonal to the straight segments lying on the boundary of the square together with portions of light cones emanating from the corners of the square. Consider those generators of $\mathcal{H}$ at which the null planes meet the light cones: it is easily seen that no point on those generators is an Alexandrov point of $\mathcal{H}$.

As constructing support functions (or support hypersurfaces) to the horizon $\mathcal{H}$ is generally easier than doing analysis directly on $\mathcal{H}$ we start by giving a criteria in terms of upper and lower support functions for a function to have a second order expansion (2.2).

**Lemma 5.4** *Let $U \subset \mathbb{R}^n$ be an open neighborhood of $x_0$ and $f: U \to \mathbb{R}$ a function. Assume for $\ell = 1, 2, \ldots$ that there are open neighborhoods $U^+_\ell$ and $U^-_\ell$ of $x_0$ and $C^2$ functions $f^\pm_\ell : U^\pm_\ell \to \mathbb{R}$ so that $f^-_\ell \le f$ on $U^-_\ell$ and $f \le f^+_\ell$ on $U^+_\ell$, with $f(x_0) = f^\pm_\ell(x_0)$. Also assume that there is a symmetric bilinear form $Q$ so that*

$$\lim_{\ell\to\infty} D^2 f^+_\ell(x_0) = \lim_{\ell\to\infty} D^2 f^-_\ell(x_0) = Q \ .$$

*Then $f$ has a second order expansion at $x_0$:*

$$f(x) = f(x_0) + df(x_0)(x - x_0) + \frac{1}{2} Q(x - x_0, x - x_0) + o(|x - x_0|^2) \ ,$$

*where $df(x_0) = df^\pm_\ell(x_0)$ (this is independent of $\ell$ and the choice of $+$ or $-$). Thus the Alexandrov second derivative of $f$ exists at $x = x_0$ and is given by $D^2 f(x_0) = Q$.*

**Remark 5.5** The extra generality will not be needed here, but we remark that the proof shows that the hypothesis the functions $f^+_\ell$ are $C^2$ can be weakened to only requiring that they all have a second order expansion at $x_0$ with no regularity being needed away from $x = x_0$.

PROOF: Without loss of generality we may assume that $x_0 = 0$. By replacing $U^+_\ell$ and $U^-_\ell$ by $U_\ell := U^+_\ell \cap U^-_\ell$ we assume for each $\ell$ that $f^+_\ell$ and $f^-_\ell$ have the same domain. For any $k, \ell$ we have $f^+_\ell - f^-_k \ge 0$ has a local minimum at $x_0 = 0$

and thus a critical point there. Whence $df_\ell^+(0) = df_k^-(0)$ for all $k$ and $l$. This shows the linear functional $df(0) = df_\ell^+(0)$ is well defined. We now replace $f$ by $x \mapsto f(x) - df(0)x - \frac{1}{2}Q(x,x)$ and $f_\ell^\pm$ by $x \mapsto f_\ell^\pm(x) - df(0)x - \frac{1}{2}Q(x,x)$. Then $df(0) = df_\ell^\pm(0) = 0$, $\lim_{\ell\to 0} D^2 f_\ell^\pm(0) = 0$ and to finish the proof it is enough to show that $f(x) = o(|x|^2)$. That is we need to find for every $\varepsilon > 0$ a $\delta_\varepsilon > 0$ so that $|x| < \delta_\varepsilon$ implies $|f(x)| \le \varepsilon|x|^2$. To do this choose any $r_0 > 0$ so that $B(0,r_0) \subset U$. Then choose an $\ell$ large enough that

$$-\varepsilon|x|^2 \le D^2 f_\ell^-(0)(x,x)\ , \qquad D^2 f_\ell^+(0)(x,x) \le \varepsilon|x|^2\ ,$$

for all $x \in B(0,r_0)$. Now choose $r_1 \le r_0$ so that $B(0,r_1) \subset U_\ell$. For this $\ell$ we can use the Taylor expansions of $f_\ell^+$ and $f_\ell^-$ at 0 to find a $\delta_\varepsilon > 0$ with $0 < \delta_\varepsilon \le r_1$ so that

$$\Big| f_\ell^\pm(x) - \frac{1}{2} D f_\ell^\pm(0)(x,x) \Big| \le \frac{1}{2}\varepsilon|x|^2$$

for all $x$ with $|x| < \delta_1$. Then $f \le f_\ell^+$ implies that if $|x| < \delta_\varepsilon$ then

$$f(x) \le \frac{1}{2} D^2 f_\ell^+(0)(x,x) + \Big( f_\ell^+(x) - \frac{1}{2} D^2 f_\ell^+(0)(x,x) \Big) \le \frac{\varepsilon}{2}|x|^2 + \frac{\varepsilon}{2}|x|^2 = \varepsilon|x|^2\ ,$$

with a similar calculation, using $f_\ell^- \le f$, yielding $-\varepsilon|x|^2 \le f(x)$. This completes the proof. □

Proof: This proof uses, essentially, the same geometric facts about horizons that are used in the proof of Proposition A.1. Let $p_0 \in \Gamma_{\mathrm{int}}$ be an Alexandrov point of a section $\mathcal{S} \cap \mathcal{H}$, where $\mathcal{S}$ is a spacelike or timelike $C^2$ hypersurface that intersects $\mathcal{H}$ properly transversally, and with $p_0 \in \Gamma_{\mathrm{int}}$. By restricting to a suitable neighborhood of $p_0$ we can assume without loss of generality that $M$ is globally hyperbolic, that $\Gamma$ maximizes the distance between any two of its points, and that $\mathcal{H}$ is the boundary of $J^+(\mathcal{H})$ and $J^-(\mathcal{H})$ (that is $\mathcal{H}$ divides $M$ into two open sets, its future $I^+(\mathcal{H})$ and its past $I^-(\mathcal{H})$).

We will simplify notation a bit and assume that $\mathcal{S}$ is spacelike, only trivial changes are required in the case $\mathcal{S}$ is timelike. We can find $C^2$ coordinates $x^1, \dots, x^n$ on $\mathcal{S}$ centered at $p_0$ so that in these coordinates $\mathcal{S} \cap \mathcal{H}$ is given by a graph $x^n = h(x^1, \dots, x^{n-1})$. By restricting the size of $\mathcal{S}$ we can assume that these coordinates are defined on all of $\mathcal{S}$ and that their image is $B^{n-1}(r) \times (-\delta, \delta)$ for some $r, \delta > 0$ and that $h\colon B^{n-1}(r) \to (-\delta, \delta)$. As $p_0$ is an Alexandrov point of $\mathcal{S} \cap \mathcal{H}$ and $h(0) = 0$ the function $h$ has a second order expansion

$$h(\vec{x}) = dh(0)\vec{x} + \frac{1}{2} D^2 h(0)(\vec{x},\vec{x}) + o(|\vec{x}|^2)\ ,$$

where $\vec{x} = (x^1, \dots, x^{n-1})$. By possibly changing the sign of $x^n$ we may assume that

$$\{(\vec{x}, y) \in \mathcal{S} \mid y \ge h(\vec{x})\} \subset J^+(\mathcal{H})\ , \qquad \{(\vec{x}, y) \in \mathcal{S} \mid y \le h(\vec{x})\} \subset J^-(\mathcal{H})\ .$$

For $\varepsilon \geq 0$ define

$$h^\pm_\varepsilon(\vec{x}) = dh(0)\vec{x} + \frac{1}{2}D^2h(0)(\vec{x},\vec{x}) \pm \frac{\varepsilon}{2}|\vec{x}|^2 \ .$$

Then $h^\pm_\varepsilon$ is a $C^2$ function on $B^{n-1}(r)$, and for $\varepsilon = 0$ the function $h_0 = h^+_0 = h^-_0$ is just the second order expansion of $h$ at $\vec{x} = 0$. For each $\varepsilon > 0$ the second order Taylor expansion for $h$ at 0 implies that there is an open neighborhood $V_\varepsilon$ of 0 in $B^{n-1}(r)$ so that

$$h^-_\varepsilon \leq h \leq h^+_\varepsilon \quad \text{on} \quad V_\varepsilon \ .$$

Set

$$N := \{(\vec{x}, h_0(\vec{x})) \mid \vec{x} \in B^{n-1}(r)\} \ , \qquad N^\pm_\varepsilon := \{(\vec{x}, h^\pm_\varepsilon(\vec{x})) \mid \vec{x} \in V_\varepsilon\} \ .$$

Let $\mathbf{n}$ be the future pointing timelike unit normal to $\mathcal{S}$ and let $\eta\colon (a,b) \to \Gamma_{\mathrm{int}}$ be the affine parameterization of $\Gamma_{\mathrm{int}}$ with $\eta(0) = p_0$ and $\langle \eta'(0), \mathbf{n}(p_0)\rangle = -1$ (which implies that $\eta$ is future directed). Let $\mathbf{k}$ be the unique $C^1$ future directed null vector field along $N$ so that $\mathbf{k}(p_0) = \eta'(0)$ and $\langle \mathbf{k}, \mathbf{n}\rangle = -1$. Likewise let $\mathbf{k}^\pm_\varepsilon$ be the $C^1$ future directed null vector field along $N^\pm_\varepsilon$ with $\mathbf{k}(p_0) = \eta'(0)$ and $\langle \mathbf{k}^\pm_\varepsilon, \mathbf{n}\rangle = -1$.

Let $p$ be any point of $\Gamma_{\mathrm{int}} \cap J^-(p_0) \setminus \{p_0\}$. Then $p = \eta(t_0)$ for some $t_0 \in (a,b)$. To simplify notation assume that $t_0 \leq 0$. By Lemma 4.15 $N$ has no focal points in $\Gamma_{\mathrm{int}}$ and in particular no focal points on $\eta|_{(a,0]}$. Therefore if we fix a $t_1$ with $t_1 < t_0 < b$ then there will be an open neighborhood $W$ of 0 in $B^{n-1}(r)$ so that

$$\widetilde{\mathcal{H}} := \{\exp(t\mathbf{k}(\vec{x}, h_0(\vec{x}))) \mid \vec{x} \in W,\ t \in (t_1, 0)\}$$

is an embedded null hypersurface of $M$. By Proposition A.3 the hypersurface is of smoothness class $C^2$. The focal points depend continuously on the second fundamental form so there is an $\varepsilon_0 > 0$ so that if $\varepsilon < \varepsilon_0$ then none of the submanifolds $N^\pm_\varepsilon$ have focal points along $\eta|_{[t_1,0]}$. Therefore if $0 < \varepsilon < \varepsilon_0$ there is an open neighborhood $W_\varepsilon$ and such that

$$\widetilde{\mathcal{H}}^\pm_\varepsilon := \{\exp(t\mathbf{k}(\vec{x}, h(\vec{x}))) \mid \vec{x} \in W_\varepsilon,\ t \in (t_1, 0)\}$$

is a $C^2$ embedded hypersurface of $M$.

We now choose smooth local coordinates $y^1, \ldots, y^{n+1}$ for $M$ centered at $p = \eta(t_0)$ so that $\partial/\partial y^{n+1}$ is a future pointing timelike vector field and the level sets $y^{n+1} =$ Const. are spacelike. Then there will be an open neighborhood $U$ of 0 in $\mathbb{R}^n$ so that near $p$ the horizon $\mathcal{H}$ is given by a graph $y^{n+1} = f(y^1, \ldots, y^n)$. Near $p$ the future and past of $\mathcal{H}$ are given by $J^+(\mathcal{H}) = \{y^{n+1} \geq f(y^1, \ldots, y^n)\}$ and $J^-(\mathcal{H}) = \{y^{n+1} \leq f(y^1, \ldots, y^n)\}$. There will also be open neighborhoods $U_0$ and $U^\pm_\varepsilon$ of 0 in $\mathbb{R}^n$ and functions $f_0$ and $f^\pm_\varepsilon$ defined on these sets so that near $p$

$$\begin{aligned}
\widetilde{\mathcal{H}} &= \{(y^1, \ldots, y^n, f_0(y^1, \ldots, y^n)) \mid (y^1, \ldots, y^n) \in U_0\} \ , \\
\widetilde{\mathcal{H}}^\pm_\varepsilon &= \{(y^1, \ldots, y^n, f^\pm_\varepsilon(y^1, \ldots, y^n)) \mid (y^1, \ldots, y^n) \in U^\pm_\varepsilon\} \ .
\end{aligned}$$

The hypersurfaces $\widetilde{\mathcal{H}}$ and $\widetilde{\mathcal{H}}_\varepsilon^\pm$ are $C^2$ which implies the functions $f_0$ and $f_\varepsilon^\pm$ are all $C^2$.

Since $N_\varepsilon^- \subset J^-(\mathcal{H})$, a simple achronality argument shows that $\widetilde{\mathcal{H}}_\varepsilon^- \subset J^-(\mathcal{H})$. (This uses the properties of $\mathcal{H}$ described in the first paragraph of the proof.) By choosing $N_\varepsilon^+$ small enough, we can assume that $\widetilde{\mathcal{H}}_\varepsilon^+$ is achronal. We now show that, relative to some neighborhood of $p$, $\mathcal{H} \subset J^-(\widetilde{\mathcal{H}}_\varepsilon^+)$. Let $\mathcal{O}_\varepsilon$ be a neighborhood of $p$, disjoint from $\mathcal{S}$, such that $\widetilde{\mathcal{H}}_\varepsilon^+ \cap \mathcal{O}_\varepsilon$ separates $\mathcal{O}_\varepsilon$ into the disjoint open sets $I^+(\widetilde{\mathcal{H}}_\varepsilon^+ \cap \mathcal{O}_\varepsilon; \mathcal{O}_\varepsilon)$ and $I^-(\widetilde{\mathcal{H}}_\varepsilon^+ \cap \mathcal{O}_\varepsilon; \mathcal{O}_\varepsilon)$. Now, by taking $\mathcal{O}_\varepsilon$ small enough, we claim that $\mathcal{H}$ does not meet $I^+(\widetilde{\mathcal{H}}_\varepsilon^+ \cap \mathcal{O}_\varepsilon; \mathcal{O}_\varepsilon)$. If that is not the case, there is a sequence $\{p_\ell\}$ such that $p_\ell \in \mathcal{H} \cap I^+(\widetilde{\mathcal{H}}_\varepsilon^+ \cap \mathcal{O}_\varepsilon; \mathcal{O}_\varepsilon)$ and $p_\ell \to p$. For each $\ell$, let $\Gamma_\ell$ be a future inextendible null generator of $\mathcal{H}$ starting at $p_\ell$. Since $p$ is an interior point of $\Gamma$, the portions of the $\Gamma_\ell$'s to the future of $p_\ell$ must approach the portion of $\Gamma$ to the future of $p$. Hence for $\ell$ sufficiently large, $\Gamma_\ell$ will meet $\mathcal{S}$ at a point $q_\ell \in \mathcal{S} \cap \mathcal{H}$, say, such that $q_\ell \to p_0$. For such $\ell$, the segment of $\Gamma_\ell$ from $p_\ell$ to $q_\ell$ approaches the segment of $\Gamma$ from $p$ to $p_0$ uniformly as $\ell \to \infty$. Since $\widetilde{\mathcal{H}}_\varepsilon^+$ separates a small neighborhood of the segment of $\Gamma$ from $p$ to $p_0$, it follows that for $\ell$ large enough, the segment of $\Gamma_\ell$ from $p_\ell$ to $q_\ell$ will meet $\widetilde{\mathcal{H}}_\varepsilon^+$. But this implies for such $\ell$ that $p_\ell \in I^-(N_\varepsilon^+)$, which contradicts the achronality of $\widetilde{\mathcal{H}}_\varepsilon^+$. We conclude, by choosing $\mathcal{O}_\varepsilon$ small enough, that $\mathcal{H} \cap I^+(\widetilde{\mathcal{H}}_\varepsilon^+ \cap \mathcal{O}_\varepsilon; \mathcal{O}_\varepsilon) = \varnothing$, and hence that $\mathcal{H} \cap \mathcal{O}_\varepsilon \subset J^-(\widetilde{\mathcal{H}}_\varepsilon^+ \cap \mathcal{O}_\varepsilon; \mathcal{O}_\varepsilon)$. Shrinking $U_\varepsilon^+$ if necessary, this inclusion, and the inclusion $\widetilde{\mathcal{H}}_\varepsilon^- \subset J^-(\mathcal{H})$ imply that on $U_\varepsilon^-$ we have $f_\varepsilon^- \leq f$ and that on $U_\varepsilon^+$ we have $f \leq f_\varepsilon^+$. Therefore if we can show that $\lim_{\varepsilon \searrow 0} D^2 f_\varepsilon^\pm(0) = D^2 f_0(0)$ then Lemma 5.4 implies that 0 is an Alexandrov point of $f$ with Alexandrov second derivative given by $D^2 f(0) = D^2 f_0(0)$.

To see that $\lim_{\varepsilon \searrow 0} D^2 f_\varepsilon^\pm(0) = D^2 f_0(0)$ let $b_0$ and $b_\varepsilon^\pm$ be the Weingarten maps of $\widetilde{\mathcal{H}}$ and $\widetilde{\mathcal{H}}_\varepsilon^\pm$ along $\eta$. By Proposition A.1 they all satisfy the same Ricatti equation (A.2). The initial condition for $b_0$ is calculated algebraically from the second fundamental form of $N$ at $p_0$ (*cf.* Section 3) and likewise the initial condition for $b_\varepsilon^\pm$ is calculated algebraically from the second fundamental form of $N_\varepsilon^\pm$ at $p_0$. From the definitions we clearly have

$$\text{Second Fundamental of } N_\varepsilon^\pm \text{ at } p_0 \quad \longrightarrow \quad \text{Second Fundamental of } N \text{ at } p_0$$

as $\varepsilon \searrow 0$. Therefore continuous dependence of solutions to ODE's on initial conditions implies that $\lim_{\varepsilon \searrow 0} b_\varepsilon^\pm = b_0$ at all points of $\eta|_{[0,t_0]}$. As $D^2 f_\varepsilon^\pm(0)$ and $D^2 f_0(0)$ are algebraic functions of $b_\varepsilon^\pm$ and $b_0$ at the point $p = \eta(t_0)$ this implies that $\lim_{\varepsilon \searrow 0} D^2 f_\varepsilon^\pm(0) = D^2 f_0(0)$ and completes the proof that $p$ is an Alexandrov point of $\mathcal{H}$.

Finally it follows from the argument that at all points of $\Gamma_{\mathrm{int}}$ the null Weingarten map $b$ for $\mathcal{H}$ is the same as the Weingarten for the $C^2$ null hypersurface $\widetilde{\mathcal{H}}$. So $b$ will satisfy (A.2) by Proposition A.1. □

We end this section with one more regularity result, Theorem 5.6 below. Its proof requires some techniques which are introduced in Section 6 only — more precisely, an appropriate generalization of Lemma 6.11 is needed; this, in turn, relies on an appropriate generalization of Lemma 6.9. For this reason we defer the proof of Theorem 5.6 to an appendix, Appendix C.

**Theorem 5.6** *Let $\mathcal{S}$ be any $C^2$ hypersurface intersecting $\mathcal{H}$ properly transversally, and define*

$$\begin{aligned} S_0 &= \{q \in \mathcal{S} \cap \mathcal{H} :\ q \text{ is an interior point of a generator of } \mathcal{H}\} \ , \quad (5.4) \\ S_1 &= \{q \in S_0 :\ \text{all interior points of the generator} \\ &\qquad \text{through } q \text{ are Alexandrov points of } \mathcal{H}\} \ , \quad (5.5) \\ S_2 &= \{q \in S_0 :\ \ q \text{ is an Alexandrov point of } \mathcal{H}\} \ , \quad (5.6) \end{aligned}$$

*Then $S_1$ and $S_2$ have full $n-1$ dimensional Hausdorff measure in $S_0$.*

**Remark 5.7** $S_0$ does not have to have full measure in $\mathcal{S} \cap \mathcal{H}$, it can even be empty. This last case occurs indeed when $\mathcal{S} = \{t = 0\}$ in the example described around Equation (3.12) in Section 3. Note, however, that if $\mathcal{S}$ is a level set of a properly transverse foliation $\mathcal{S}(t)$, then (as already mentioned) for almost all $t$'s the sets $S(t)_0$ (as defined in (5.4) with $\mathcal{S}$ replaced by $\mathcal{S}(t)$) will have full measure in $\mathcal{S}(t)$. We shall call a generator an *Alexandrov generator* if all its interior points are Alexandrov points. It follows that for generic sections (in the measure sense above) for almost all points through those sections the corresponding generators will be Alexandrov. The discussion here thus gives a precise meaning to the statement that *almost all generators of a horizon are Alexandrov generators.*

# 6 Area monotonicity

In this section we shall show the monotonicity of the area, assuming that the Alexandrov divergence $\theta_{\mathcal{A}l}$ of the generators of $\mathcal{H}$, or that of a section of $\mathcal{H}$, is non–negative. The result is local in the sense that it only depends on the part of the event horizon $\mathcal{H}$ that is between the two sections $S_1$ and $S_2$ whose area we are trying to compare. We consider a spacetime $(M, g)$ of dimension $n+1$. We have the following:

**Theorem 6.1** *Suppose that $(M, g)$ is an $(n+1)$ dimensional spacetime with a future horizon $\mathcal{H}$. Let $\Sigma_a$, $a = 1, 2$ be two embedded achronal spacelike hypersurfaces of $C^2$ differentiability class, set $S_a = \Sigma_a \cap \mathcal{H}$. Assume that*

$$S_1 \subset J^-(S_2) \ , \qquad S_1 \cap S_2 = \varnothing \ , \tag{6.1}$$

*and that either*

1. *the divergence $\theta_{\mathcal{A}l}$ of $\mathcal{H}$ defined ($\mathfrak{H}^n_\sigma$–almost everywhere) by Equation (2.10) is non–negative on $J^+(S_1) \cap J^-(S_2)$, or*

2. *the divergence $\theta^{S_2}_{\mathcal{A}l}$ of $S_2$ defined ($\mathfrak{H}^{n-1}_\sigma$–almost everywhere) by Equation (3.3) is non–negative, with the null energy condition (4.1) holding on $J^+(S_1) \cap J^-(S_2)$.*

*Then*

1. $$\mathrm{Area}_{S_2}(S_1) \leq \mathrm{Area}(S_2) \ , \tag{6.2}$$
   *with* $\mathrm{Area}_{S_2}(S_1)$ *defined in Equations (3.10)–(3.11) and* $\mathrm{Area}(S_2)$ *defined in Equation (3.9). This implies the inequality*
   $$\mathrm{Area}(S_1) \leq \mathrm{Area}(S_2) \ .$$

2. *If equality holds in (6.2), then $(J^+(S_1) \setminus S_1) \cap (J^-(S_2) \setminus S_2)$ (which is the part of $\mathcal{H}$ between $S_1$ and $S_2$) is a* smooth totally geodesic *(analytic if the metric is analytic) null hypersurface in $M$. Further, if $\gamma$ is a null generator of $\mathcal{H}$ with $\gamma(0) \in S_1$ and $\gamma(1) \in S_2$, then the curvature tensor of $(M, g)$ satisfies $R(X, \gamma'(t))\gamma'(t) = 0$ for all $t \in [0, 1]$ and $X \in T_{\gamma(t)}\mathcal{H}$.*

**Remark 6.2** Note that if $S_1 \cap S_2 \neq \varnothing$ then the inequality $\mathrm{Area}_{S_2}(S_1) \leq \mathrm{Area}(S_2)$ need not hold. (The inequality $\mathrm{Area}(S_1) \leq \mathrm{Area}(S_2)$ will still hold.) For example, if $\mathrm{Area}_{S_1}(S_1) > \mathrm{Area}(S_1)$, then letting $S_2 = S_1$ gives a counterexample. If $S_1 \cap S_2 \neq \varnothing$ then the correct inequality is $\mathrm{Area}_{S_2}(S_1 \setminus S_2) \leq \mathrm{Area}(S_2 \setminus S_1)$.

Proof: Let us start with an outline of the proof, without the technical details — these will be supplied later in the form of a series of lemmas. Let $\mathcal{N}_{\mathcal{H}}(S_1)$ be the collection of generators of $\mathcal{H}$ that meet $S_1$. Let $A \subseteq S_2$ be the set of points that are of the form $S_2 \cap \gamma$ with $\gamma \in \mathcal{N}_{\mathcal{H}}(S_1)$; replacing $\Sigma_2$ by an appropriate submanifold thereof if necessary, $A$ will be a closed subset of $S_2$ (Proposition 6.3). The condition (6.1) together with achronality of $\mathcal{H}$ imply that every $q \in A$ is on exactly one of the generators $\gamma \in \mathcal{N}_{\mathcal{H}}(S_1)$, we thus have a well defined function $\phi: A \to S_1$ given by

$$\phi(q) = S_1 \cap \gamma \quad \text{where } \gamma \in \mathcal{N}_{\mathcal{H}}(S_1) \text{ and } q = S_2 \cap \gamma. \tag{6.3}$$

Let us for simplicity assume that the affine distance from $S_1$ to $S_2$ on the generators passing through $S_1$ is bounded from below, and that the affine existence time of those generators to the future of $S_2$ is also bounded from below (in what follows we will see how to reduce the general case to this one). In this case $A$ can be embedded in a $C^{1,1}$ hypersurface $N$ in $\Sigma_2$ (Lemma 6.9) and $\phi$ can be extended to a locally Lipschitz function $\widehat{\phi}$, from $N$ to $\Sigma_1$ (Lemma 6.11). $A$ is

$\mathfrak{H}^{n-1}_{h_2}$ measurable (closed subsets of manifolds are Hausdorff measurable[11]), $\phi$ is a locally Lipschitz function (Lemma 6.11), and so we can apply the generalization of the change–of–variables theorem known as the area formula [23, Theorem 3.1] (with $m$ and $k$ there equal to $n-1$) to the extension $\widehat{\phi}$ of $\phi$ to get

$$\int_{S_1} \#\phi^{-1}[p]\, d\mathfrak{H}^{n-1}_{h_1}(p) = \int_A J(\phi)(q)\, d\mathfrak{H}^{n-1}_{h_2}(q) \ , \tag{6.4}$$

where $J(\phi)$ is the restriction of the Jacobian[19] of $\widehat{\phi}$ to $A$, and $h_a$, $a=1,2$, denotes the metric induced on $\Sigma_a$ by $g$. We observe that

$$\#\phi^{-1}[p] = N(p, S_2)$$

for all $q \in S_1$. Indeed, if $\gamma$ is a null generator of $\mathcal{H}$ and $p \in \gamma \cap S_1$ then the point $q \in S_2$ with $q \in \gamma$ satisfies $\phi(q) = p$. Thus there is a bijective correspondence between $\phi^{-1}[p]$ and the null generators of $\mathcal{H}$ passing through $p$. (This does use that $S_1 \cap S_2 = \varnothing$.) This implies that

$$\int_{S_1} \#\phi^{-1}[p]\, d\mathfrak{H}^{n-1}_{h_1}(p) = \int_{S_1} N(p, S_2)\, d\mathfrak{H}^{n-1}_{h_1}(p) \ ,$$

which can be combined with the area formula (6.4) to give that

$$\mathrm{Area}_{S_2}(S_1) = \int_A J(\phi)(q)\, d\mathfrak{H}^{n-1}_{h_2}(q). \tag{6.5}$$

We note that [24, Theorem 3.2.3, p. 243] also guarantees[20] that $A \ni p \to N(p, S_2)$ is measurable, so that $\mathrm{Area}_{S_2}(S_1)$ is well defined. Now a calculation, that is straightforward when $\mathcal{H}$ is smooth (Proposition A.5, Appendix A), shows that having the null mean curvatures $\theta_{\mathcal{A}l}$ – or $\theta^{S_2}_{\mathcal{A}l}$ nonnegative together with the null energy condition – implies that $J(\phi) \le 1$ almost everywhere with respect to $\mathfrak{H}^{n-1}_{h_2}$ on $A$ — this is established in Proposition 6.16. Using this in (6.5) completes the proof.

Our first technical step is the following:

**Proposition 6.3** *Let the setting be as in Theorem 6.1. Then there exists an open submanifold $\Sigma_2'$ of $\Sigma_2$ such that $A$ is a closed subset of $S_2' = \Sigma_2' \cap \mathcal{H}$. Replacing $\Sigma_2$ with $\Sigma_2'$ we can thus assume that $A$ is closed in $S_2$.*

---

[19] It should be pointed out that the Jacobian $J(\widehat{\phi})$ is *not* the usual Jacobian which occurs in the change–of–variables theorem for Lebesgue measure on $\mathbb{R}^n$, but contains the appropriate $\sqrt{\det g_{ij}}$ factors occurring in the definition of the measure associated with a Riemannian metric, see [23, Def. 2.10, p. 423]. A clear exposition of the Jacobians that occur in the area and co–area formulas can be found in [22, Section 3.2] in the flat $\mathbb{R}^n$ case. See also [41, Appendix p. 66] for the smooth Riemannian case.

[20] The result in [24, Theorem 3.2.3, p. 243] is formulated for subsets of $\mathbb{R}^q$, but the result generalizes immediately to manifolds by considering local charts, together with a partition of unity argument.

PROOF: Let $W_2$ be the subset of $S_2$ consisting of all points $p \in S_2$ for which there exists a semi–tangent $X$ at $p$ of $\mathcal{H}$ such that the null geodesic $\eta$ starting at $p$ in the direction $X$ does not meet $\Sigma_1$ when extended to the past (whether or not it remains in $\mathcal{H}$). Let $\{p_k\}$ be a sequence of points in $W_2$ such that $p_k \to p \in S_2$. We show that $p \in W_2$, and hence that $W_2$ is a closed subset of $S_2$. For each $k$, let $X_k$ be a semi–tangent at $p_k$ such that the null geodesic $\eta_k$ starting at $p_k$ in the direction $X_k$ does not meet $\Sigma_1$ when extended to the past. Since the collection of $\sigma$-normalized null vectors is locally compact, by passing to a subsequence if necessary, we may assume that $X_k \to X$, where $X$ is a future pointing null vector at $p$. Let $\eta$ be the null geodesic starting at $p$ in the direction $X$. Since $(p_k, X_k) \to (p, X)$, $\eta_k \to \eta$ in the strong sense of geodesics. Since, further, the $\eta_k$'s remain in $\mathcal{H}$ to the future and $\mathcal{H}$ is closed, it follows that $\eta$ is a null generator of $\mathcal{H}$ and $X$ is a semi-tangent of $\mathcal{H}$ at $p$. Since $\eta_k \to \eta$, if $\eta$ meets $\Sigma_1$ when extended to the past then so will $\eta_k$ for $k$ large enough. Hence, $\eta$ does not meet $\Sigma_1$, so $p \in W_2$, and $W_2$ is a closed subset of $S_2$. Then $S_2' := S_2 \setminus W_2$ is an open subset of $S_2$. Thus, there exists an open set $\Sigma_2'$ in $\Sigma_2$ such that $S_2' = S_2 \cap \Sigma_2' = \Sigma_2' \cap \mathcal{H}$. Note that $p \in S_2'$ iff for each semi-tangent $X$ of $\mathcal{H}$ at $p$, the null geodesic starting at $p$ in the direction $X$ meets $\Sigma_1$ when extended to the past. In particular, $A \subset S_2'$. To finish, we show that $A$ is a closed subset of $S_2'$. Let $\{p_k\}$ be a sequence in $A$ such that $p_k \to p \in S_2'$. Let $\eta_k$ be the unique null generator of $\mathcal{H}$ through $p_k$. Let $X_k$ be the $\sigma$-normalized tangent to $\eta_k$ at $p_k$. Let $q_k$ be the point where $\eta_k$ meets $S_1$. Again, by passing to a subsequence if necessary, we may assume $X_k \to X$, where $X$ is a semi–tangent of $\mathcal{H}$ at $p$. Let $\eta$ be the null geodesic at $p$ in the direction $X$. We know that when extended to the past, $\eta$ meets $\Sigma_1$ at a point $q$, say. Since $\eta_k \to \eta$ we must in fact have $q_k \to q$ and hence $q \in S_1$. It follows that $\eta$ starting from $q$ is a null generator of $\mathcal{H}$. Hence $p \in A$, and $A$ is closed in $S_2'$. $\Box$

We note that in the proof above we have also shown the following:

**Lemma 6.4** *The collection $\mathcal{N} = \cup_{p \in \mathcal{H}} \mathcal{N}_p$ of semi–tangents is a closed subset of $TM$.*

Recall that we have fixed a complete Riemannian metric $\sigma$ on $M$. For each $\delta > 0$ let

$$\begin{aligned} A_\delta : \ = \ & \{p \in A \mid \sigma\text{-dist}(p, S_1) \geq \delta\,, \text{ and the generator } \gamma \text{ through} \\ & p \text{ can be extended at least a } \sigma\text{–distance } \delta \text{ to the future}\}\,. \end{aligned} \tag{6.6}$$

We note that if the $\sigma$ distance from $S_1$ to $S_2$ on the generators passing through $S_1$ is bounded from below, and if the length of the portions of those generators which lie to the future of $S_2$ is also bounded from below (which is trivially fulfilled when the generators are assumed to be future complete), then $A_\delta$ will coincide with $A$ for $\delta$ small enough.

**Lemma 6.5** *Without loss of generality we may assume*

$$\overline{S}_1 \cap S_2 = \varnothing \tag{6.7}$$

*and*

$$A = \cup_{\delta>0} A_\delta \ . \tag{6.8}$$

PROOF: We shall show how to reduce the general situation in which $S_1 \cap S_2 = \varnothing$ to one in which Equation (6.7) holds. Assume, first, that $\Sigma_1$ is connected, and introduce a complete Riemannian metric on $\Sigma_1$. With respect to this metric $\Sigma_1$ is a complete metric space such that the closed distance balls $\overline{B}(p,r)$ (closure in $\Sigma_1$) are compact in $\Sigma_1$, and hence also in $M$. Choose a $p$ in $\Sigma_1$ and let

$$\Sigma_{1,i} = B(p,i)$$

($B(p,i)$ – open balls). Then $\Sigma_1 = \cup_i \Sigma_{1,i} = \cup_i \overline{\Sigma}_{1,i})$ (closure either in $\Sigma_1$ or in $M$) is an increasing union of compact sets. The $\Sigma_{1,i}$'s are spacelike achronal hypersurfaces which have the desired property

$$\overline{S}_{1,i} \cap S_2 \subset \Sigma_1 \cap \Sigma_2 = \varnothing$$

(closure in M), $S_{1,i} \equiv \Sigma_{1,i} \cap \mathcal{H}$. Suppose that we have shown that point 1 of Theorem 6.1 is true for $S_{1,i}$, thus

$$\mathrm{Area}_{S_2}(S_{1,i}) \le \mathrm{Area}(S_2) \ .$$

As $S_{1,i} \subset S_{1,i+1}$, $S_1 = \cup S_{1,i}$, the monotone convergence theorem gives

$$\lim_{i\to\infty} \mathrm{Area}_{S_2}(S_{1,i}) = \mathrm{Area}_{S_2}(S_1) \ ,$$

whence the result.

If $\Sigma_1$ is not connected, one can carry the above procedure out on each component (at most countably many), obtaining a sequence of sets $\Sigma_{1,i}$ for each component of $\Sigma_1$. The resulting collection of sets is countable, and an obvious modification of the above argument establishes that (6.7) holds.

But if Equation (6.7) holds, that is if $\overline{S}_1 \cap S_2 = \varnothing$, we have $\sigma\text{-dist}(p, S_1) = \sigma\text{-dist}(p, \overline{S}_1) > 0$ for all $p \in S_2$. Therefore (6.8) holds. This completes the proof. □

In order to continue we need the following variation of the Whitney extension theorem. As the proof involves very different ideas than the rest of this section we postpone the proof to Appendix D.

**Proposition 6.6** *Let $A \subset \mathbb{R}^n$ and $f\colon A \to \mathbb{R}$. Assume there is a constant $C > 0$ and for each $p \in A$ there is a vector $a_p \in \mathbb{R}^n$ so that the inequalities*

$$f(p) + \langle x-p, a_p\rangle - \frac{C}{2}\|x-p\|^2 \le f(x) \le f(p) + \langle x-p, a_p\rangle + \frac{C}{2}\|x-p\|^2 \tag{6.9}$$

*hold for all $x \in A$. Also assume that for all $p, q \in A$ and all $x \in \mathbb{R}^n$ the inequality*

$$f(p) + \langle x - p, a_p \rangle - \frac{C}{2}\|x - p\|^2 \le f(q) + \langle x - q, a_q \rangle + \frac{C}{2}\|x - q\|^2 \qquad (6.10)$$

*holds. Then there is a function $F\colon \mathbb{R}^n \to \mathbb{R}$ of class $C^{1,1}$ so that $f$ is the restriction of $F$ to $A$.*

**Remark 6.7** Following [9, Prop. 1.1 p. 7] the hypothesis (6.9) of Proposition 6.6 will be stated by saying $f$ has *global upper and lower support paraboloids of opening $C$* in $K$. The condition (6.10) can be expressed by saying *the upper and lower support paraboloids of $f$ are disjoint.*

**Remark 6.8** Unlike most extension theorems of Whitney type this result does not require that the set being extend from be closed (here $A$ need not even be measurable) and there is no continuity assumption on $f$ or on the map that sends $p$ to the vector $a_p$ (cf. [62, Chap. VI Sec. 2], [40, Vol. 1, Thm 2.3.6, p. 48] where to get a $C^{1,1}$ extension the mapping $p \mapsto a_p$ is required to be Lipschitz.) The usual continuity assumptions are replaced by the "disjointness" condition (6.10) which is more natural in the geometric problems considered here.

Proposition 6.6 is a key element for the proof of the result that follows [42]:

**Lemma 6.9** *For any $\delta > 0$ there is a $C^{1,1}_{\mathrm{loc}}$ hypersurface $N_\delta$ of $\Sigma_2$ (thus $N_\delta$ has co–dimension two in $M$) with*

$$A_\delta \subseteq N_\delta \ .$$

**Remark 6.10** We note that $N_\delta$ does not have to be connected.

PROOF: The strategy of the proof is as follows: We start by showing that all points $q$ in $A_\delta$ possess space–time neighborhoods in which all the generators of $\mathcal{H}$ passing through $A_\delta$ are contained in a $C^{1,1}$ hypersurface in $M$. This hypersurface is not necessarily null, but is transverse to $\Sigma_2$ on $A_\delta$. $N_\delta$ will be obtained by a globalization argument that uses the intersections of those locally defined hypersurfaces with $\Sigma_2$.

By the local character of the arguments that follow, one can assume without loss of generality that $M$ is globally hyperbolic; otherwise, where relevant, one could restrict to a globally hyperbolic neighborhood and define the pasts, futures, *etc,* with respect to this neighborhood.

For $q \in A_\delta$ define $q^\pm \in J^\pm(q) \cap \mathcal{H}$ as those points on the generator $\gamma_q$ of $\mathcal{H}$ through $q$ which lie a $\sigma$–distance $\delta$ away from $q$. Achronality of $\mathcal{H}$ implies that $\gamma_q$ has no conjugate points, hence the hypersurfaces $\partial J^\mp(q^\pm)$ are smooth in a neighborhood of $q$, tangent to $\mathcal{H}$ there.

Let $q_0$ in $A_\delta$ and choose a basis $B = \{E_1, \dots, E_n, E_{n+1}\}$ of $T_{q_0^+}M$ such that $g(E_i, E_j) = \delta_{ij}$ and $g(E_i, E_{n+1}) = 0$ for $i, j = 1, \dots, n$, while $g(E_{n+1}, E_{n+1}) = -1$. We shall denote by $(y_1, \dots, y_{n+1})$ the normal coordinates associated with this basis.

We note that $A_{\delta'} \subset A_\delta$ for $\delta < \delta'$ so that it is sufficient to establish our claims for $\delta$ small. Choosing $\delta$ small enough the cone

$$\{y_{n+1} = -\sqrt{y_1^2 + \cdots + y_n^2} \mid y_i \in \mathbb{R}, i = 1, \dots, n\}$$

coincides with $\partial J^-(q_0^+)$ in a neighborhood of $q_0$. Reducing $\delta$ again if necessary we can suppose that $q_0^-$ belongs to a normal coordinate neighborhood of $q_0^+$. Let $y_{n+1}^0$ be the $(n+1)^{st}$ coordinate of $q_0$ and let $y_{n+1}^{0-}$ be the $(n+1)^{st}$ coordinate of $q_0^-$. Denote by $C^-$ the smooth hypersurface with compact closure obtained by intersecting the previous cone by the space–time slab lying between the hypersurfaces $y_{n+1} = \frac{1}{2}y_{n+1}^0$ and $y_{n+1} = \frac{1}{2}(y_{n+1}^0 + y_{n+1}^{0-})$. Let us use the symbol $\omega$ to denote any smooth local parameterization $\omega \longrightarrow \mathbf{n}(\omega) \in T_{q_0^+}M$ of this hypersurface. We then obtain a smooth parameterization by $\omega$ of $\partial J^-(q_0^+)$ in a neighborhood of $q_0$ by using the map $\omega \longrightarrow \exp_{q_0^+}(\mathbf{n}(\omega))$. Reducing $\delta$ again if necessary, without loss of generality we may assume that there are no conjugate points on $\exp_{q_0^+}(\mathbf{n}(\omega)) \subset C^-$.

Using parallel transport of $B$ along radial geodesics from $q_0^+$, we can then use $\omega$ to obtain a smooth parameterization of $\partial J^-(m)$ in a neighborhood of $q_0$, for all $m$ in a neighborhood of $q_0^+$. For all $m$ in this neighborhood let us denote by $\mathbf{n}_m(\omega)$ the corresponding vector in $T_mM$. The map $(m, \omega) \longrightarrow \exp_m(\mathbf{n}_m(\omega))$ is smooth for $m$ close to $q_0^+$ and for $\omega \in C^-$. This implies the continuity of the map which to a couple $(m, \omega)$ assigns the second fundamental form $I\!I(m, \omega)$, defined with respect to an auxiliary Riemannian metric $\sigma$, of $\partial J^-(m)$ at $\exp_m(\mathbf{n}_m(\omega))$. If we choose the $m$'s in a sufficiently small compact coordinate neighborhood of $q_0^+$, compactness of $C^-$ implies that the $\sigma$ norm of $I\!I$ can be bounded in this neighborhood by a constant, independently of $(m, \omega)$.

Let, now $(x_1, \dots, x_{n+1})$ be a coordinate system covering a globally hyperbolic neighborhood of $q_0$, centered at $q_0$, and of the form $\mathcal{O} = B^n(2r) \times (-a, a)$. We can further require that $g(\partial_{n+1}, \partial_{n+1}) < C < 0$ over $\mathcal{O}$. Transversality shows that, reducing $r$ and $a$ if necessary, the hypersurfaces $\exp_m(\mathbf{n}_m(\omega))$ are smooth graphs above $B^n(2r)$, for $m$ close to $q_0^+$; we shall denote by $f_m^-$ the corresponding graphing functions. Now, the first derivatives of the $f_m^-$'s can be bounded by an $m$–independent constant: the vectors $V_i = \partial_i + D_i f_m^- \partial_{n+1}$ are tangent to the graph, hence non–timelike, and the result follows immediately from the equation $g(V_i, V_i) \geq 0$. Next, the explicit formula for the second fundamental form of a graph (*cf., e.g.*, [2, Eqns. (3.3) and (3.4), p. 604]) gives a uniform bound for the second derivatives of the $f_m^-$'s over $B^n(2r)$. We make a similar argument, using future light cones issued from points m near to $q_0^-$, and their graphing functions over $B^n(2r)$ denoted by $f^+_m$. Let $f$ be the graphing function of $\mathcal{H}$ over $B^n(2r)$; we can choose a constant $C$ such that, for all $p$ in

$$G_\delta := \{x \in \Gamma \mid \Gamma \text{ is a generator of } \mathcal{H} \text{ passing through } A_\delta\} \cap \mathcal{O} \;, \tag{6.11}$$

the graph of the function

$$f_p(x) = f(x_p) + Df(x_p)(x - x_p) - C||x - x_p||^2$$

lies under the graph of $f^-_{p^+}$, hence in the past of $p^+$. Here we write $x_p$ for the space coordinate of $p$, thus $p = (x_p, f(x_p))$, *etc.* Similarly, for all $q$ in $G_\delta$ the graph of

$$f_q(x) = f(x_q) + Df(x_q)(x - x_q) + C||x - x_q||^2$$

lies above the graph of $f^+_{q^-}$, thus is to the future of $q^-$. Achronality of the horizon implies then that the inequality (6.10) holds for $x \in B^n(2r)$. Increasing $C$ if necessary, Equation (6.10) will hold for all $x_p, x_q \in B^n(r)$ and $x \in \mathbb{R}^n$: Indeed, let us make $C$ large enough so that

$$(C-1)r^2 \geq \mathrm{Lip}^2_{B^n(r)}(f) - \inf_{B^n(r)}(f) + \sup_{B^n(r)}(f) \;,$$

where $\mathrm{Lip}_{B^n(r)}(f)$ is the Lipschitz continuity constant of $f$ on $B^n(r)$. For $x \notin B^n(2r)$ we then have

$$\begin{aligned}
f(x_q) - f(x_p) + \frac{C}{2}(||x - x_q||^2 + ||x - x_p||^2) & \\
+\langle x - x_q, a_q\rangle - \langle x - x_p, a_p\rangle \;&\geq \\
f(x_q) - f(x_p) + \frac{C}{2}(||x - x_q||^2 + ||x - x_p||^2) & \\
-\frac{1}{2}(||x - x_q||^2 + ||x - x_p||^2 + ||a_q||^2 + ||a_p||^2) \;&\geq \\
\inf_{B^n(r)}(f) - \sup_{B^n(r)}(f) + (C-1)r^2 - \mathrm{Lip}^2_{B^n(r)}(f) \;&\geq\; 0 \;.
\end{aligned}$$

Here $a_p = df(x_p)$, $a_q = df(x_q)$, and we have used $\mathrm{Lip}_{B^n(r)}(f)$ to control $||a_p||$ and $||a_q||$.

Let the set $A$ of Proposition 6.6 be the projection on $B^n(r)$ of $G_\delta$; by that proposition there exists a $C^{1,1}$ function from $B^n(r)$ to $(-a, a)$, the graph of which contains $G_\delta$. (It may be necessary to reduce $r$ and $\mathcal{O}$ to obtain this). Note that this graph contains $A_\delta \cap \mathcal{O}$ and is transverse to $\Sigma_2$ there. From the fact that a transverse intersection of a $C^2$ hypersurface with a $C^{1,1}$ hypersurface is a $C^{1,1}$ manifold we obtain that $A_\delta \cap \mathcal{O}$ is included in a $C^{1,1}$ submanifold of $\Sigma_2$, which has space–time co–dimension two.

Now $A_\delta$ is a closed subset of the manifold $\Sigma_2'$ defined in Proposition 6.3, and can thus be covered by a countable union of compact sets. Further, by definition of $A_\delta$ any point thereof is an interior point of a generator. Those facts and the arguments of the proof of Proposition 3.1 show that $A_\delta$ can be covered by a countable locally finite collection of relatively compact coordinate neighborhoods $\mathcal{U}_i$ of the form $\mathcal{U}_i = B^{n-1}(\epsilon_i) \times (-\eta_i, \eta_i)$ such that $\mathcal{U}_i \cap \mathcal{H}$ is a graph of a semi–convex function $g_i$:

$$\mathcal{U}_i \cap \mathcal{H} = \{x^n = g_i(x^A)\,, (x^A) \equiv (x^1, \ldots, x^{n-1}) \in B^{n-1}(\epsilon_i)\}\ . \tag{6.12}$$

The $\epsilon_i$'s will further be restricted by the requirement that there exists a $C^{1,1}$ function

$$h_i\colon B^{n-1}(\epsilon_i) \to \mathbb{R} \tag{6.13}$$

such that the graph of $h_i$ contains $\mathcal{U}_i \cap \mathcal{A}_\delta$. The $h_i$'s are the graphing functions of the $C^{1,1}$ manifolds just constructed in a neighborhood of each point of $A_\delta$.

At those points at which $S_2$ is differentiable let $\mathbf{m}$ denote a $h_2$–unit vector field normal to $S_2$ pointing towards $J^-(\mathcal{H}) \cap \Sigma_2$. Choosing the orientation of $x^n$ appropriately we can assume that at those points at which $\mathbf{m}$ is defined we have

$$\mathbf{m}(x^n - g_i) > 0\ . \tag{6.14}$$

In order to globalize this construction we use an idea of [42]. Let $\phi_i$ be a partition of unity subordinate to the cover $\{\mathcal{U}_i\}_{i\in\mathbb{N}}$, define

$$\chi_i(q) = \begin{cases} (x^n - h_i(x^A))\phi_i(q)\ , & q = (x^A, x^n) \in \mathcal{U}_i\ , \\ 0\ , & \text{otherwise,} \end{cases} \tag{6.15}$$

$$\chi_\delta = \sum_i \chi_i\ . \tag{6.16}$$

Define a hypersurface $N_\delta \subset \cup_i \mathcal{U}_i$ via the equations

$$N_\delta \cap \mathcal{U}_i = \{\chi_\delta = 0\ , d\chi_\delta \neq 0\}\ . \tag{6.17}$$

(We note that $N_\delta$ does not have to be connected, but this is irrelevant for our purposes.) If $q \in A_\delta$ we have $\chi_i(q) = 0$ for all $i$'s hence $A_\delta \subset \{\chi_\delta = 0\}$. Further, for $q \in A_\delta \cap \mathcal{U}_i$ we have $\mathbf{m}(\chi_i) = \phi_i \mathbf{m}(x^n - g_i) \geq 0$ from Equation (6.14), and since for each $q \in A_\delta$ there exists an $i$ at which this is strictly positive we obtain

$$\mathbf{m}(\chi_\delta) = \sum_i \mathbf{m}(\chi_i) > 0\ .$$

It follows that $A_\delta \subset N_\delta$, and the result is established. □

**Lemma 6.11** *Under the condition (6.8), the map $\phi\colon A \to S_1$ is locally Lipschitz.*

PROOF: Let $q_0 \in A$, we need to find a neighborhood $\mathcal{U}_A \equiv \mathcal{U} \cap A$ in $A$ so that $\phi$ is Lipschitz in this neighborhood. By the condition (6.8) there exists $\delta > 0$ such that $q_0 \in A_\delta$. By lower semi–continuity of the existence time of geodesics we can choose a neighborhood $\mathcal{U}$ of $q_0$ in $\Sigma_2$ small enough so that

$$\mathcal{U} \cap A_\delta = \mathcal{U} \cap A . \tag{6.18}$$

Denote by $N$ the $C^{1,1}_{\mathrm{loc}}$ hypersurface $N_\delta$ (corresponding to the chosen small value of $\delta$) given by Lemma 6.9, so that

$$A \cap \mathcal{U} \subseteq N \cap \mathcal{U} .$$

Let $\mathbf{n}$ be the future pointing unit normal to $\Sigma_2$. (So $\mathbf{n}$ is a unit timelike vector.) Let $\mathbf{m}$ be the unit normal to $N$ in $\Sigma_2$ that points to $J^-(\mathcal{H}) \cap \Sigma_2$. Let $\mathbf{k} := -(\mathbf{n} + \mathbf{m})$. Then $\mathbf{k}$ is a past pointing null vector field along $N$ and if $q \in A_\delta$ then $\mathbf{k}(q)$ is tangent to the unique generator of $\mathcal{H}$ through $q$. As $\Sigma_2$ is $C^2$ then the vector field $\mathbf{n}$ is $C^1$, while the $C^{1,1}_{\mathrm{loc}}$ character of $N$ implies that $\mathbf{m}$ is locally Lipschitz. It follows that $\mathbf{k}$ is a locally Lipschitz vector field along $N$. Now for each $q \in A \cap \mathcal{U}$ there is a unique positive real number $r(q)$ so that $\phi(q) = \exp(r(q)\mathbf{k}(q)) \in S_1$. Lower semi–continuity of existence time of geodesics implies that, passing to a subset of $\mathcal{U}$ if necessary, for each $q \in N \cap \mathcal{U}$ there is a unique positive real number $\hat{r}(q)$ so that $\widehat{\phi}(q) = \exp(\hat{r}(q)\mathbf{k}(q)) \in \Sigma_1$. As $\Sigma_1$ is a Lipschitz hypersurface, Clarke's implicit function theorem [17, Corollary, p. 256] implies that $q \mapsto \hat{r}(q)$ is a locally Lipschitz function near $q_0$. Thus $\widehat{\phi}$ is Lipschitz near $q_0$ and the restriction of $\widehat{\phi}$ to $A_\delta$ is $\phi$. □

**Corollary 6.12** *The section $S_1$ is countably $(n-1)$ rectifiable.*

**Remark 6.13** By starting with an arbitrary section $S = S_1$ of the form $S = \mathcal{H} \cap \Sigma$, where $\Sigma$ is a $C^2$ spacelike hypersurface or timelike hypersurface that meets $\mathcal{H}$ properly transversely, and then constructing a section $S_2 = \mathcal{H} \cap \Sigma_2$ for a $C^2$ spacelike hypersurface $\Sigma_2$ so that the hypotheses of Theorem 6.1 hold, one obtains that every such section $S$ is countably $(n-1)$ rectifiable. In the case that $\Sigma_1 \equiv \Sigma$ is spacelike a more precise version of this is given in [42] (where it is shown that $S$ is countably $(n-1)$ rectifiable of class $C^2$).

PROOF: From Lemma 6.9 it is clear for each $\delta > 0$ that $A_\delta$ is countably $(n-1)$ rectifiable. By Lemma 6.11 the map $\phi|_{A_\delta} \colon A_\delta \to S_1$ is locally Lipschitz and therefore $\phi[A_\delta]$ is locally countable $(n-1)$ rectifiable. But $S_1 = \bigcup_{k=1}^\infty \phi[A_{1/k}]$ so $S_1$ is a countable union of countably $(n-1)$ rectifiable sets and therefore is itself countably $(n-1)$ rectifiable. □

It follows from the outline given at the beginning of the proof of Theorem 6.1 that we have obtained:

**Proposition 6.14** *The formula (6.5) holds.*

**Corollary 6.15** *For any section $S = \mathcal{H}\cap\Sigma$ where $\Sigma$ is a $C^2$ spacelike or timelike hypersurface that intersects $\mathcal{H}$ properly transversely the set of points of $S$ that are on infinitely many generators has vanishing $(n-1)$ dimensional Hausdorff measure.*

PROOF: For any point $p$ of $S$ choose a globally hyperbolic neighborhood $\mathcal{O}$ of the point and then choose a neighborhood $S_1$ of $p$ in $S$ small enough that the closure $\overline{S}_1$ is compact, satisfies $\overline{S}_1 \subset S\cap\mathcal{O}$, and so that there is a $C^\infty$ Cauchy hypersurface $\Sigma_2$ of $\mathcal{O}$ such that $\overline{S}_1 \subset I^-(\Sigma_2;\mathcal{O})$. Let $S_2 = \mathcal{H}\cap\Sigma_2$ and let $A$ be the set of points of $S_2$ that are of the form $S_2\cap\gamma$ where $\gamma$ is a generator of $\mathcal{H}$ that meets $\overline{S}_1$. Compactness of $\overline{S}_1$ together with the argument of the proof of Proposition 6.3 show that the set of generators of $\mathcal{H}$ that meet $\overline{S}_1$ is a compact subset of the bundle of null geodesic rays of $M$, and that $A$ is a compact subset of $S_2$. Then Lemma 6.9 implies that $A$ is a compact set in a $C^{1,1}$ hypersurface of $\Sigma_2$ and so $\mathfrak{H}^{n-1}(A) < \infty$. Lemma 6.11 and the compactness of $A$ yields that $\phi\colon A \to \overline{S}_1$ given by (6.3) (with $\overline{S}_1$ replacing $S_1$) is Lipschitz. Therefore the Jacobian $J(\phi)$ is bounded on $A$. By Proposition 6.14 $\int_{\overline{S}_1} N(p,S_2)\, d\mathfrak{H}^{n-1}(p) = \int_A J(\phi)(q)\, d\mathfrak{H}^{n-1}_{h_2}(q) < \infty$ and so $N(p,S_2) < \infty$ except on a set of $\mathfrak{H}^{n-1}$ measure zero. But $N(p,S_2)$ is the number of generators of $\mathcal{H}$ through $p$ so this implies the set of points of $S_1$ that are on infinitely many generators has vanishing $(n-1)$ dimensional Hausdorff measure. Now $S$ can be covered by a countable collection of such neighborhoods $S_1$ and so the set of points on $S$ that are on infinitely many generators has vanishing $(n-1)$ dimensional Hausdorff measure. This completes the proof. □

To establish (6.2) it remains to show that $J(\phi) \le 1$ $\mathfrak{H}^{n-1}_{h_2}$-almost everywhere. To do this one would like to use the classical formula that relates the Jacobian of $\phi$ to the divergence $\theta$ of the horizon, *cf.* Proposition A.5, Appendix A below. However, for horizons which are not $C^2$ we only have the Alexandrov divergence $\theta_{\mathcal{A}l}$ at our disposal, and it is not clear whether or not this formula holds with $\theta$ replaced by $\theta_{\mathcal{A}l}$ for general horizons. The proof below consists in showing that this formula remains true after such a replacement for generators passing through almost all points of $A$.

**Proposition 6.16** $J(\phi) \le 1$ $\mathfrak{H}^{n-1}_{h_2}$*–almost everywhere on $A$.*

PROOF: The argument of the paragraph preceding Equation (6.18) shows that it is sufficient to show $J(\phi) \le 1$ $\mathfrak{H}^{n-1}_{h_2}$–almost everywhere on $A_\delta$ for each $\delta > 0$. Let $N_\delta$ be the $C^{1,1}$ manifold constructed in Lemma 6.9, and let $\mathcal{U} \subset \Sigma_2$ be a coordinate neighborhood of the form $\mathcal{V}\times(a,b)$, with $\mathcal{V} \subset \mathbb{R}^{n-1}$ and $a, b, \in \mathbb{R}$, in which $\mathcal{U}\cap N_\delta$ is the graph of a $C^{1,1}$ function $g\colon \mathcal{V} \to \mathbb{R}$, and in which $\mathcal{H}\cap\mathcal{U}$ is the graph of a semi–convex function $f\colon \mathcal{V} \to \mathbb{R}$. By [24, Theorem 3.1.5, p. 227] for every $\epsilon > 0$ there exists a twice differentiable function $g_{1/\epsilon}: \mathcal{V} \to \mathbb{R}$ such that

$$\mathfrak{L}^{n-1}_{h_\mathcal{V}}(\{g \neq g_{1/\epsilon}\}) < \epsilon \ . \tag{6.19}$$

Here $\mathfrak{L}^{n-1}_{h_\mathcal{V}}$ denotes the $(n-1)$ dimensional Riemannian measure on $\mathcal{V}$ associated with the pull–back $h_\mathcal{V}$ of the space–time metric $g$ to $\mathcal{V}$. Let $\mathrm{pr}A_\delta$ denote the projection on $\mathcal{V}$ of $A_\delta \cap \mathcal{U}$, thus $A_\delta \cap \mathcal{U}$ is the graph of $g$ over $\mathrm{pr}A_\delta$. For $q \in \mathcal{V}$ let $\theta^{n-1}_*(\mathfrak{L}^{n-1}_{h_\mathcal{V}}, \mathrm{pr}A_\delta, q)$ be the density function of $\mathrm{pr}A_\delta$ in $\mathcal{V}$ with respect to the measure $\mathfrak{L}^{n-1}_{h_\mathcal{V}}$, defined as in [61, page 10] using geodesic coordinates centered at $q$ with respect to the metric $h_\mathcal{V}$. Define

$$B = \{q \in \mathrm{pr}A_\delta \mid \theta^{n-1}_*(\mathfrak{L}^{n-1}_{h_\mathcal{V}}, \mathrm{pr}A_\delta, q) = 1\} \subset \mathcal{V} \ . \tag{6.20}$$

By [52, Corollary 2.9] or [24, 2.9.12] the function $\theta^{n-1}_*(\mathfrak{L}^{n-1}_{h_\mathcal{V}}, \mathrm{pr}A_\delta, \cdot)$ differs from the characteristic function $\chi_{\mathrm{pr}A_\delta}$ of $\mathrm{pr}A_\delta$ by a function supported on a set of vanishing measure, which implies that $B$ has full measure in $\mathrm{pr}A_\delta$. Let

$$B_{1/\epsilon} = B \cap \{g = g_{1/\epsilon}\} \ ;$$

Equation (6.19) shows that

$$\mathfrak{L}^{n-1}_{h_\mathcal{V}}(B \setminus B_{1/\epsilon}) < \epsilon \ . \tag{6.21}$$

We define $\tilde{B}_{1/\epsilon} \subset B_{1/\epsilon} \subset \mathrm{pr}A_\delta \subset \mathcal{V}$ as follows:

$$\tilde{B}_{1/\epsilon} = \{q \in B_{1/\epsilon} \mid \theta^{n-1}_*(\mathfrak{L}^{n-1}_{h_\mathcal{V}}, B_{1/\epsilon}, q) = 1\} \ . \tag{6.22}$$

Similarly the set $\tilde{B}_{1/\epsilon}$ has full measure in $B_{1/\epsilon}$, hence

$$\mathfrak{L}^{n-1}_{h_\mathcal{V}}(B \setminus \tilde{B}_{1/\epsilon}) < \epsilon \ . \tag{6.23}$$

Let $(\mathrm{pr}A_\delta)_{\mathcal{A}l}$ denote the projection on $\mathcal{V}$ of the set of those Alexandrov points of $\mathcal{H} \cap \Sigma_2$ which are in $A_\delta \cap \mathcal{U}$; $(\mathrm{pr}A_\delta)_{\mathcal{A}l}$ has full measure in $\mathrm{pr}A_\delta$. Let, further, $\mathcal{V}_{\mathrm{Rad}}$ be the set of points at which $g$ is twice differentiable; by Rademacher's theorem (*cf., e.g.,* [22, p. 81]) $\mathcal{V}_{\mathrm{Rad}}$ has full measure in $\mathcal{V}$. We set

$$\hat{B}_{1/\epsilon} = \tilde{B}_{1/\epsilon} \cap (\mathrm{pr}A_\delta)_{\mathcal{A}l} \cap \mathcal{V}_{\mathrm{Rad}} \ , \tag{6.24}$$

$$\hat{B} = \bigcup_{i \in \mathbb{N}} \hat{B}_i \tag{6.25}$$

It follows from (6.23) that $\hat{B}$ has full measure in $\mathrm{pr}A_\delta$. Since $g$ is Lipschitz we obtain that the graph of $g$ over $\hat{B}$ has full $\mathfrak{H}^{n-1}_{h_2}$–measure in $A_\delta \cap \mathcal{U}$.

Consider any $x_0 \in \hat{B}$, then $x_0$ is an Alexandrov point of $f$ so that we have the expansion

$$f(x) = f(x_0) + df(x_0)(x - x_0) + \tfrac{1}{2} D^2 f(x_0)(x - x_0, x - x_0) + o(|x - x_0|^2) \tag{6,26}$$

Next, $g$ is twice differentiable at $x_0$ so that we also have

$$g(x) = g(x_0) + dg(x_0)(x - x_0) + \tfrac{1}{2}D^2g(x_0)(x - x_0, x - x_0) + o(|x - x_0|^2), \tag{6.27}$$

$$dg(x) - dg(x_0) = D^2g(x_0)(x - x_0, \cdot) + o(|x - x_0|) \ . \tag{6.28}$$

Further there exists $j \in \mathbb{N}$ such that $x_0 \in \hat{B}_j \subset B_j$; by definition of $B_j$ we then have

$$g_j(x_0) = g(x_0) = f(x_0) \ . \tag{6.29}$$

We claim that the set of directions $\vec{n} \in B^{n-1}(1) \subset T_{x_0}\mathcal{V}$ for which there exists a sequence of points $q_i = x_0 + r_i\vec{n}$ with $r_i \to 0$ and with the property that $f(q_i) = g_j(q_i)$ is dense in $B^{n-1}(1)$. Indeed, suppose that this is not the case, then there exists $\epsilon > 0$ and an open set $\Omega \subset B^{n-1}(1) \subset T_{x_0}\mathcal{V}$ of directions $\vec{n} \in B^{n-1}(1)$ such that the solid cone $K_\epsilon = \{x_0 + r\vec{n} \mid \vec{n} \in \Omega \, , r \in [0, \epsilon]\}$ contains no points from $B_j$. It follows that the density function $\theta_*^{n-1}(\mathfrak{L}_{h_\mathcal{V}}^{n-1}, B_j, x_0)$ is strictly smaller than one, which contradicts the fact that $x_0$ is a density point of $\mathrm{pr}A_\delta$ ($x_0 \in \hat{B}_j \subset B$, *cf.* Equation (6.20)), and that $x_0$ is a density point of $B_j$ ($x_0 \in \hat{B}_j \subset \tilde{B}_j$, *cf.* Equation (6.22)).

Equation (6.29) leads to the following Taylor expansions at $x_0$:

$$g_j(x) = f(x_0) + dg_j(x_0)(x - x_0) + \tfrac{1}{2}D^2g_j(x_0)(x - x_0, x - x_0) + o(|x - x_0|^2), \tag{6.30}$$

$$dg_j(x) - dg_j(x_0) = D^2g_j(x_0)(x - x_0, \cdot) + o(|x - x_0|) \ . \tag{6.31}$$

Subtracting Equation (6.26) from Equation (6.30) at points $q_i = x_0 + r_i\vec{n}$ at which $f(q_i) = g_j(q_i)$ we obtain

$$dg_j(x_0)(\vec{n}) - df(x_0)(\vec{n}) = O(r_i) \ . \tag{6.32}$$

Density of the set of $\vec{n}$'s for which (6.32) holds implies that

$$dg_j(x_0) = df(x_0) \ . \tag{6.33}$$

Again comparing Equation (6.26) with Equation (6.30) at points at which $f(q_i) = g_j(q_i)$ it now follows that

$$D^2g_j(x_0)(\vec{n}, \vec{n}) - D^2f(x_0)(\vec{n}, \vec{n}) = 2\frac{g_j(x_0 + r_i\vec{n}) - f(x_0 + r_i\vec{n})}{r_i^2} + o(1) = o(1) \ , \tag{6.34}$$

and density of the set of $\vec{n}$'s together with the polarization identity gives

$$D^2g_j(x_0) = D^2f(x_0) \ . \tag{6.35}$$

Similarly one obtains

$$dg(x_0) = df(x_0) \ , \tag{6.36}$$

$$D^2g(x_0) = D^2f(x_0) \ . \tag{6.37}$$

Define $S_j$ to be the graph (over $\mathcal{V}$) of $g_j$; Equations (6.33) and (6.36) show that both $S_j$ and $N_\delta$ are tangent to $\mathcal{H} \cap \Sigma_2$ at $p_0 \equiv (x_0, f(x_0))$:

$$T_{p_0}(\mathcal{H} \cap \Sigma_2) = T_{p_0} N_\delta = T_{p_0} S_j \ . \tag{6.38}$$

Let $\mathbf{n}_j$ denote the ($C^1$) field of $\sigma$–unit future directed null normals to $S_j$ such that

$$\mathbf{n}_j(p_0) = \mathbf{n}(p_0) \ ,$$

where $\mathbf{n}(p_0)$ is the semi–tangent to $\mathcal{H}$ at $p_0$. Let $\phi_j : S_j \to \Sigma_1$ be the map obtained by intersecting the null geodesics passing through points $q \in S_j$ with tangent parallel to $\mathbf{n}_j(q)$ there. Equation (6.38) shows that

$$\phi_j(p_0) = \phi(p_0) = \widehat{\phi}(p_0) \ . \tag{6.39}$$

The lower semi–continuity of the existence time of geodesics shows that, passing to a subset of $\mathcal{U}$ if necessary, $\phi_j$ is well defined on $S_j$. By an argument similar to the one leading to (6.33), it follows from Equations (6.28), (6.31), (6.33) and (6.35)–(6.37) that the derivatives of $\phi_j$ and of $\widehat{\phi}$ coincide at $p_0$, in particular

$$J(\phi)(p_0) \equiv J(\widehat{\phi})(p_0) = J(\phi_j)(p_0) \ . \tag{6.40}$$

Equation (6.35) further shows that $S_j$ is second order tangent to $\mathcal{H}$ at $p_0$ in the sense defined before Lemma 4.15, we thus infer from that lemma that there are no focal points of $S_j$ along the segment of the generator $\Gamma$ of $\mathcal{H}$ passing through $p_0$ which lies to the future of $\Sigma_1$. Consider the set $\mathcal{H}_j$ obtained as the union of null geodesics passing through $S_j$ and tangent to $\mathbf{n}_j$ there; standard considerations show that there exists a neighborhood of $\Gamma \cap I^+(\Sigma_1) \cap I^-(\Sigma_2)$ in which $\mathcal{H}_j$ is a $C^1$ hypersurface. It is shown in Appendix A that 1) $\mathcal{H}_j$ is actually a $C^2$ hypersurface, *cf.* Proposition A.3, and 2) the null Weingarten map $b = b_{\mathcal{H}_j}$ of $\mathcal{H}_j$ satisfies the Ricatti equation

$$b' + b^2 + R = 0 \ . \tag{6.41}$$

Here a prime denotes a derivative with respect to an affine parameterization $s \mapsto \eta(s)$ of $\Gamma$ that makes $\Gamma$ future directed. Theorem 5.1 implies that $\mathcal{H}$ has a null Weingarten map $b_{\mathcal{A}l}$ defined in terms of the Alexandrov second derivatives of $\mathcal{H}$ on all of the segment $\Gamma \cap I^+(\Sigma_1) \cap I^-(\Sigma_2)$ and that this Weingarten map also satisfies the Ricatti equation (6.41). As the null Weingarten map can be expressed in terms of the first and second derivatives of the graphing function of a section Equations (6.33) and (6.35) imply that $b_{\mathcal{H}_j}(p_0) = b_{\mathcal{A}l}(p_0)$. Therefore uniqueness of solutions to initial value problems implies that $b_{\mathcal{H}_j} = b_{\mathcal{A}l}$ on all of $\Gamma \cap I^+(\Sigma_1) \cap I^-(\Sigma_2)$. But the divergence (or null mean curvature) of a null hypersurface is the trace of its null Weingarten map and thus on the segment $\Gamma \cap I^+(\Sigma_1) \cap I^-(\Sigma_2)$ we have $\theta_{\mathcal{H}_j} = \operatorname{trace} b_{\mathcal{H}_j} = \operatorname{trace} b_{\mathcal{A}l} = \theta_{\mathcal{A}l}$.

We now finish the proof under the first of hypothesis of Theorem 6.1, that is that the divergence the $\theta_{\mathcal{A}l} \geq 0$ on $J^+(S_1) \cap J^-(S_2)$. Then $\theta_{\mathcal{H}_j} = \theta_{\mathcal{A}l} \geq 0$ and Proposition A.5 implies that $J(\phi)(p_0) = J(\phi_j)(p_0) \leq 1$ as required.

The other hypothesis of Theorem 6.1 is that $\theta^{S_2}_{\mathcal{A}l} \geq 0$ and the null energy condition holding on $J^+(S_1) \cap J^-(S_2)$. Recall that $p_0 = \Gamma \cap S_2$ is an Alexandrov point of $\mathcal{H}$; thus Theorem 5.1 applies and shows that $\theta = \theta_{\mathcal{A}l}$ exists along the segment $\Gamma \cap I^+(\Sigma_1) \cap I^-(\Sigma_2)$, and satisfies the Raychaudhuri equation

$$\theta' = -\mathrm{Ric}(\eta', \eta') - \sigma^2 - \frac{1}{n-2}\theta^2 .$$

Here $\sigma^2$ is the norm squared of the shear (and should not be confused with the auxiliary Riemannian metric which we have also denoted by $\sigma$). But the null energy condition implies $\mathrm{Ric}(\eta', \eta') \geq 0$ so this equation and $\theta^{S_2}_{\mathcal{A}l} \geq 0$ implies $\theta = \theta_{\mathcal{A}l} \geq 0$ on $\Gamma \cap I^+(\Sigma_1) \cap I^-(\Sigma_2)$. Then the equality $\theta_{\mathcal{H}_j} = \theta_{\mathcal{A}l}$ yields $\theta_{\mathcal{H}_j} \geq 0$ and again we can use Proposition A.5 to conclude $J(\phi)(p_0) = J(\phi_j)(p_0) \leq 1$. This completes the proof. □

To finish the proof of Theorem 6.1, we need to analyze what happens when

$$\mathrm{Area}_{S_2}(S_1) = \mathrm{Area}(S_2) . \tag{6.42}$$

In this case Equation (6.4) together with Proposition 6.16 show that $A$ has full measure in $S_2$, that $N(p, S_2) = 1$ $\mathfrak{H}^{n-1}_{h_1}$–almost everywhere on $S_1$, and that $J(\phi) = 1$ $\mathfrak{H}^{n-1}_{h_2}$–almost everywhere on $S_2$. Next, the arguments of the proof of that Proposition show that

$$\theta_{\mathcal{A}l} = 0 \tag{6.43}$$

$\mathfrak{H}^n_\sigma$–almost everywhere on $J^-(A) \cap J^+(S_1) = J^-(S_2) \cap J^+(S_1)$. The proof of Proposition 6.16 further shows that $\mathrm{Ric}(\eta', \eta') = 0$ $\mathfrak{H}^n_\sigma$–almost everywhere on $\mathcal{H} \cap J^+(S_1) \cap J^-(S_2)$ (*cf.* Equation (6.41)), hence everywhere there as the metric is assumed to be smooth and the distribution of semi–tangents is a closed set (Lemma 6.4). Here $\eta'$ is any semi–tangent to $\mathcal{H}$. We first note the following observation, the proof of which borrows arguments from [6, Section IV]:

**Lemma 6.17** *Under the hypotheses of Theorem 6.1, suppose further that the equality (6.42) holds. Then there are no end points of generators of $\mathcal{H}$ on*

$$\Omega \equiv (J^+(S_1) \setminus S_1) \cap (J^-(S_2) \setminus S_2) . \tag{6.44}$$

Proof: Suppose that there exists $q \in \Omega$ which is an end point of a generator $\Gamma$ of $\mathcal{H}$, set $\{p\} = \Gamma \cap S_2$, extending $\Gamma$ beyond its end point and parameterizing it appropriately we will have

$$\Gamma(0) = q , \quad \Gamma(1) = p , \quad \Gamma(a) \in I^-(\mathcal{H}) ,$$

for any $a < 0$ for which $\Gamma(a)$ is defined. Now $p$ is an interior point of a generator, and semi–tangents at points in a sufficiently small neighborhood of $p$ are arbitrarily close to the semi–tangent $X_p$ at $p$. Since $I^-(\mathcal{H})$ is open it follows from continuous dependence of solutions of ODE's upon initial values that there exists a neighborhood $\mathcal{V} \subset S_2$ of $p$ such that every generator of $\mathcal{H}$ passing through $\mathcal{V}$ leaves $\mathcal{H}$ before intersecting $\Sigma_1$ when followed backwards in time from $S_2$, hence $A \cap \mathcal{V} = \varnothing$, and $A$ does not have full measure in $S_2$. $\square$

To finish the proof we shall need the following result, which seems to be of independent interest:

**Theorem 6.18** *Let $\Omega$ be an open subset of a horizon $\mathcal{H}$ which contains no end points of generators of $\mathcal{H}$, and suppose that the divergence $\theta_{\mathcal{A}l}$ of $\mathcal{H}$ defined ($\mathfrak{H}^n_\sigma$–almost everywhere) by Equation (2.10) vanishes $\mathfrak{H}^n_\sigma$–almost everywhere. Then $\Omega$ is a smooth submanifold of $M$ (analytic if the metric is analytic).*

**Remark 6.19** The condition on $\Omega$ is equivalent to $\Omega$ being a $C^1$ hypersurface, *cf.* [6].

Proof: Let $p_0 \in \Omega$ and choose a smooth local foliation $\{\Sigma_\lambda \mid -\varepsilon < \lambda < \varepsilon\}$ of an open neighborhood $\mathcal{U}$ of $p_0$ in $M$ by spacelike hypersurfaces so that $\overline{\mathcal{U} \cap \Omega} \subset \Omega$ and so that $p_0 \in \Sigma_0$. Letting $\sigma$ be a the auxiliary Riemannian metric, by possibly making $\mathcal{U}$ smaller we can assume that the $\sigma$-distance of $\overline{\mathcal{U} \cap \Omega}$ to $\overline{\Omega} \setminus \Omega$ is $< \delta$ for some $\delta > 0$. Let $\Omega_{\mathcal{A}l}$ be the set of Alexandrov points of $\Omega$ and let $\mathcal{B} = \Omega \setminus \Omega_{\mathcal{A}l}$ be the set of points of $\Omega$ where the Alexandrov second derivatives do not exist. We view $\lambda$ as a function $\lambda : \mathcal{U} \to \mathbb{R}$ in the natural way. Now $\lambda$ is smooth on $\mathcal{U}$ and by Remark 6.19 $\Omega$ is a $C^1$ manifold so the restriction $\lambda|_\Omega$ is a $C^1$ function. Letting $h_\sigma$ be the pull back of our auxiliary Riemannian metric $\sigma$ to $\Omega$ we apply the co-area formula to $\lambda|_\Omega$ and use that by Alexandrov's theorem $\mathfrak{H}^n_{h_\sigma}(\mathcal{B}) = 0$ to get

$$\int_{-\varepsilon}^{\varepsilon} \mathfrak{H}^{n-1}_{h_\sigma}(\mathcal{B} \cap \Sigma_\lambda)\, d\lambda = \int_{\mathcal{B}} J(\lambda|_\Omega)\, d\mathfrak{H}^n_{h_\sigma} = 0 .$$

This implies that for almost all $\lambda \in (-\varepsilon, \varepsilon)$ that $\mathfrak{H}^{n-1}_{h_\sigma}(\mathcal{B} \cap \Sigma_\lambda) = 0$. Therefore we can choose a $\lambda$ just a little bigger than 0 with $\mathfrak{H}^{n-1}_{h_\sigma}(\mathcal{B} \cap \Sigma_\lambda) = 0$ and so that $p_0 \in J^-(\Omega \cap \Sigma_\lambda)$. To simplify notation we denote $\Sigma_\lambda$ by $\Sigma$. Then from the choice of $\Sigma$ we have that $\mathfrak{H}^{n-1}_{h_\sigma}$ almost every point of $\Sigma$ is an Alexandrov point of $\Omega$.

By transversality and that $\Omega$ is $C^1$ the set $\Sigma \cap \Omega$ is a $C^1$ submanifold of $\Sigma$. Recalling that $\delta$ is less than the $\sigma$-distance of $\overline{\mathcal{U} \cap \Omega}$ to $\overline{\Omega} \setminus \Omega$ we see that for any $p \in \Sigma$ that the unique (because $\Omega$ is $C^1$) generator $\Gamma$ of $\Omega$ through $p$ extends in $\Omega$ a $\sigma$-distance of at least $\delta$ both to the future and to the past of $\Sigma$. Letting $A = \Sigma \cap \Omega$ and using the notation of Equation (6.6), this implies that $A = A_\delta$. As $A = \Sigma \cap \Omega$ is already a $C^1$ submanifold of $\Sigma$ Lemma 6.9 implies that $\Sigma \cap \Omega$ is a $C^{1,1}$ hypersurface in $\Sigma$. Let $g$ by any Lipschitz local graphing function of $A$ in $\Sigma$.

From Rademacher's theorem (*cf., e.g.,* [22, p. 81]) it follows that $C^{1,1} = W^{2,\infty}_{\rm loc}$, further the Alexandrov second derivatives of $g$ coincide with the classical ones almost everywhere. By [22, p. 235] the second distributional derivatives of $g$ equal the second classical derivatives of $g$ almost everywhere. It follows that the equation

$$\theta_{\mathcal{A}l} = 0 \tag{6.45}$$

can be rewritten, by freezing the coefficients of the second derivatives at the solution $g$, as a linear elliptic weak (distributional) equation with Lipschitz continuous coefficients for the graphing function $g \in W^{2,\infty}_{\rm loc}$. Elliptic regularity shows that $g$ is, locally, of $C^{2,\alpha}$ differentiability class for any $\alpha \in (0,1)$. Further, Equation (6.45) is a quasi–linear elliptic equation for $g$ (*cf., e.g.,* [29]), a standard bootstrap argument shows that $g$ is smooth (analytic if the metric is analytic) and it easily follows that $\Omega$ in a neighborhood of $\Sigma \cap \Omega$ containing $p_0$ is smooth (or analytic). As $p_0$ was an arbitrary point of $\Omega$ this completes the proof. $\square$

Returning to the proof of Theorem 6.1, we note that Lemma 6.17 shows that all points of $\Sigma \cap \Omega$, where $\Omega$ is given by Equation (6.44), are interior points of generators of $\mathcal{H}$. Simple arguments together with the invariance of the domain theorem (*cf., e.g.,* [18, Prop. 7.4, p. 79]) show that $\Omega$ is an open submanifold of $\mathcal{H}$, and Equation (6.43) shows that we can use Theorem 6.18 to conclude. $\square$

# 7 Conclusions

Let us present here some applications of Theorems 1.2 and 6.18, proved above. The first one is to the theory of stationary black holes (*cf., e.g.*, [12, 38] and references therein): in that theory the question of differentiability of event horizons arises at several key places. Recall that smoothness of event horizons has been established in 1) static [10, 63] and 2) [10] stationary–axisymmetric space–times. However, staticity or stationarity–axisymmetry are often not known *a priori* — that is indeed the case in *Hawking's rigidity theorem* [37][21]. Now, the rigidity theorem asserts that a certain class of stationary black holes have axi–symmetric domains of outer communication; its hypotheses include that of analyticity of the metric *and of the event horizon*. The examples of black holes (in analytic vacuum space–times) the horizons of which are nowhere $C^2$ constructed in [15] show that the hypothesis of analyticity of the event horizon and that of analyticity of the metric are logically independent. It is thus of interest to note the following result, which is a straightforward corollary of Theorem 1.2 and of the fact that isometries preserve area:

[21] This theorem is actually wrong as stated in [37]; a corrected version, together with a proof, can be found in [13].

**Theorem 7.1** *Let $\phi$ be an isometry of a black–hole space–time $(M, g)$ satisfying the hypotheses of Theorem 1.2. If $\phi$ maps $\mathcal{H}$ into $\mathcal{H}$, then for every spacelike hypersurface $\Sigma$ such that*

$$\phi(\Sigma \cap \mathcal{H}) \subset J^+(\Sigma \cap \mathcal{H}) \tag{7.1}$$

*the set*

$$\big(J^-(\phi(\Sigma \cap \mathcal{H})) \setminus \phi(\Sigma \cap \mathcal{H})\big) \cap \big(J^+(\Sigma \cap \mathcal{H}) \setminus (\Sigma \cap \mathcal{H})\big) \subset \mathcal{H} \tag{7.2}$$

*is a smooth (analytic if the metric is analytic) null submanifold of $M$ with vanishing null second fundamental form.*

As already pointed out, the application we have in mind is that to stationary black holes, where $\phi$ actually arises from a one parameter group of isometries $\phi_t$. We note that in such a setting the fact that isometries preserve the event horizon, as well as the existence of hypersurfaces $\Sigma$ for which (7.1) holds with $\phi$ replaced by $\phi_t$ (for some, or for all $t$'s), can be established under various standard conditions on the geometry of stationary black holes, which are of no concern to us here.

Recall, next, that the question of differentiability of *Cauchy* horizons often arises in considerations concerning *cosmic censorship* issues (*cf., e.g.,* [3, 58]). An interesting result in this context, indicating non–genericity of occurrence of *compact* Cauchy horizons, is the Isenberg–Moncrief theorem, which asserts that *analytic compact Cauchy horizons with periodic generators* in *analytic, electro–vacuum* space-times are *Killing* horizons, for a Killing vector field defined on a neighborhood of the Cauchy horizon [45]. We note that if all the generators of the horizon are periodic, then the horizon has no end–points, and analyticity follows[22] from Theorem 6.18. Hence the hypothesis of analyticity of the event horizon is not needed in [45]. We also note that there exists a (partial) version of the Isenberg–Moncrief theorem, due to Friedrich, Rácz and Wald [27], in which the hypotheses of analyticity of [45] are replaced by those of smoothness both of the metric and of the Cauchy horizon. Theorem 6.18 again shows that the hypothesis of smoothness of the Cauchy horizon is not necessary in [27].

To close this section let us note an interesting theorem of Beem and Królak [6, Section IV], which asserts that if a compact Cauchy horizon in a space–time satisfying the null energy condition contains, roughly speaking, an open dense subset $\mathcal{O}$ which is a $C^2$ manifold, then there are no end points of the generators of the event horizon, and the divergence of the event horizon vanishes. Theorem 6.18 again applies to show that the horizon must be as smooth as the metric allows. Our methods here could perhaps provide a proof of a version of the Beem–Królak theorem in which the hypothesis of existence of the set $\mathcal{O}$ will not be needed; this remains to be seen.

[22] The proof proceeds as follows: Theorem 5.1 shows that the optical equations hold on almost all generators of the Cauchy horizon; periodicity of the generators together with the Raychaudhuri equation shows then that $\theta_{\mathcal{A}l} = 0$ almost everywhere, hence Theorem 6.18 applies.

# A The Geometry of $C^2$ Null Hypersurfaces

In this appendix we prove a result concerning the regularity of null hypersurfaces normal to a $C^k$ submanifold in space–time. We also review some aspects of the geometry of null hypersurfaces, with the presentation adapted to our needs. We follow the exposition of [29].

Let $(M, g)$ be a spacetime, *i.e.*, a smooth, paracompact time-oriented Lorentzian manifold, of dimension $n+1 \geq 3$. We denote the Lorentzian metric on $M$ by $g$ or $\langle\, ,\rangle$. A *($C^2$) null hypersurface* in $M$ is a $C^2$ co-dimension one embedded submanifold $\mathcal{H}$ of $M$ such that the pullback of the metric $g$ to $\mathcal{H}$ is degenerate. Each such hypersurface $\mathcal{H}$ admits a $C^1$ non-vanishing future directed null vector field $K \in \Gamma T\mathcal{H}$ such that the normal space of $K$ at a point $p \in \mathcal{H}$ coincides with the tangent space of $\mathcal{H}$ at $p$, *i.e.*, $K_p^\perp = T_p\mathcal{H}$ for all $p \in \mathcal{H}$. (If $\mathcal{H}$ is $C^2$ the best regularity we can require for $K$ is $C^1$.) In particular, tangent vectors to $\mathcal{H}$ not parallel to $K$ are spacelike. It is well-known that the integral curves of $K$, when suitably parameterized, are null geodesics. These integral curves are called the *null geodesic generators* of $\mathcal{H}$. We note that the vector field $K$ is unique up to a positive scale factor.

Since $K$ is orthogonal to $\mathcal{H}$ we can introduce the null Weingarten map and null second fundamental form of $\mathcal{H}$ with respect $K$ in a manner roughly analogous to what is done for spacelike hypersurfaces or hypersurfaces in a Riemannian manifold, as follows: We start by introducing an equivalence relation on tangent vectors: for $X, X' \in T_p\mathcal{H}$, $X' = X \bmod K$ if and only if $X' - X = \lambda K$ for some $\lambda \in \mathbb{R}$. Let $\overline{X}$ denote the equivalence class of $X$. Simple computations show that if $X' = X \bmod K$ and $Y' = Y \bmod K$ then $\langle X', Y'\rangle = \langle X, Y\rangle$ and $\langle \nabla_{X'} K, Y'\rangle = \langle \nabla_X K, Y\rangle$, where $\nabla$ is the Levi-Civita connection of $M$. Hence, for various quantities of interest, components along $K$ are not of interest. For this reason one works with the tangent space of $\mathcal{H}$ modded out by $K$, *i.e.*, $T_p\mathcal{H}/K = \{\overline{X} \mid X \in T_p\mathcal{H}\}$ and $T\mathcal{H}/K = \cup_{p\in\mathcal{H}} T_p\mathcal{H}/K$. $T\mathcal{H}/K$ is a rank $n-1$ vector bundle over $\mathcal{H}$. This vector bundle does not depend on the particular choice of null vector field $K$. There is a natural positive definite metric $h$ in $T\mathcal{H}/K$ induced from $\langle , \rangle$: For each $p \in \mathcal{H}$, define $h\colon T_p\mathcal{H}/K \times T_p\mathcal{H}/K \to \mathbb{R}$ by $h(\overline{X}, \overline{Y}) = \langle X, Y\rangle$. From remarks above, $h$ is well-defined.

The *null Weingarten map* $b = b_K$ of $\mathcal{H}$ with respect to $K$ is, for each point $p \in \mathcal{H}$, a linear map $b\colon T_p\mathcal{H}/K \to T_p\mathcal{H}/K$ defined by $b(\overline{X}) = \overline{\nabla_X K}$. It is easily verified that $b$ is well-defined and, as it involves taking a derivative of $K$, which is $C^1$ the tensor $b$ will be $C^0$ but no more regularity can be expected. Note if $\widetilde{K} = fK$, $f \in C^1(\mathcal{H})$, is any other future directed null vector field tangent to $\mathcal{H}$, then $\nabla_X \widetilde{K} = f\nabla_X K \bmod K$. Thus $b_{fK} = f b_K$. It follows that the Weingarten map $b$ of $\mathcal{H}$ is unique up to positive scale factor and that $b$ at a given point $p \in \mathcal{H}$ depends only on the value of $K$ at $p$ when we keep $\mathcal{H}$ fixed but allow $K$ to vary while remaining tangent to the generators of $\mathcal{H}$.

A standard computation shows, $h(b(\overline{X}),\overline{Y}) = \langle \nabla_X K, Y\rangle = \langle X, \nabla_Y K\rangle = h(\overline{X}, b(\overline{Y}))$. Hence $b$ is self-adjoint with respect to $h$. The *null second fundamental form* $B = B_K$ of $\mathcal{H}$ with respect to $K$ is the bilinear form associated to $b$ via $h$: For each $p \in \mathcal{H}$, $B\colon T_p\mathcal{H}/K \times T_p\mathcal{H}/K \to \mathbb{R}$ is defined by $B(\overline{X},\overline{Y}) = h(b(\overline{X}),\overline{Y}) = \langle \nabla_X K, Y\rangle$. Since $b$ is self-adjoint, $B$ is symmetric. In a manner analogous to the second fundamental form for spacelike hypersurfaces, a null hypersurface is totally geodesic if and only if $B$ vanishes identically [51, Theorem 30].

The *null mean curvature* of $\mathcal{H}$ with respect to $K$ is the continuous scalar field $\theta \in C^0(\mathcal{H})$ defined by $\theta = \operatorname{tr} b$; in the general relativity literature $\theta$ is often referred to as the *convergence* or *divergence* of the horizon. Let $e_1, e_2, \ldots, e_{n-1}$ be $n-1$ orthonormal spacelike vectors (with respect to $\langle , \rangle$) tangent to $\mathcal{H}$ at $p$. Then $\{\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_{n-1}\}$ is an orthonormal basis (with respect to $h$) of $T_p\mathcal{H}/K$. Hence at $p$,

$$\begin{aligned} \theta &= \operatorname{tr} b = \textstyle\sum_{i=1}^{n-1} h(b(\overline{e}_i), \overline{e}_i) \\ &= \textstyle\sum_{i=1}^{n-1} \langle \nabla_{e_i} K, e_i\rangle. \end{aligned} \tag{A.1}$$

Let $\Sigma$ be the intersection, transverse to $K$, of a hypersurface in $M$ with $\mathcal{H}$. Then $\Sigma$ is a $C^2$ $(n-1)$ dimensional spacelike submanifold of $M$ contained in $\mathcal{H}$ which meets $K$ orthogonally. From Equation (A.1), $\theta|_\Sigma = \operatorname{div}_\Sigma K$, and hence the null mean curvature gives a measure of the divergence of the null generators of $\mathcal{H}$. Note that if $\widetilde{K} = fK$ then $\widetilde{\theta} = f\theta$. Thus the null mean curvature inequalities $\theta \geq 0$, $\theta \leq 0$, are invariant under positive scaling of $K$. In Minkowski space, a future null cone $\mathcal{H} = \partial I^+(p) - \{p\}$ (respectively, past null cone $\mathcal{H} = \partial I^-(p) - \{p\}$) has positive null mean curvature, $\theta > 0$ (respectively, negative null mean curvature, $\theta < 0$).

The null second fundamental form of a null hypersurface obeys a well-defined comparison theory roughly similar to the comparison theory satisfied by the second fundamental forms of a family of parallel spacelike hypersurfaces (*cf.* Eschenburg [21], which we follow in spirit).

Let $\eta\colon (a,b) \to M$, $s \to \eta(s)$, be a future directed affinely parameterized null geodesic generator of $\mathcal{H}$. For each $s \in (a,b)$, let

$$b(s) = b_{\eta'(s)} : T_{\eta(s)}\mathcal{H}/\eta'(s) \to T_{\eta(s)}\mathcal{H}/\eta'(s)$$

be the Weingarten map based at $\eta(s)$ with respect to the null vector $K = \eta'(s)$. Recall that the null Weingarten map $b$ of a smooth null hypersurface $\mathcal{H}$ satisfies a Ricatti equation (*cf.* [5, p. 431]; for completeness we indicate the proof below).

$$b' + b^2 + R = 0. \tag{A.2}$$

Here $'$ denotes covariant differentiation in the direction $\eta'(s)$, with $\eta$ – an affinely parameterized null geodesic generator of $\mathcal{H}$; more precisely, if $X = X(s)$ is a vector field along $\eta$ tangent to $\mathcal{H}$, then[23]

$$b'(\overline{X}) = b(\overline{X})' - b(\overline{X'}). \tag{A.3}$$

Finally $R\colon T_{\eta(s)}\mathcal{H}/\eta'(s) \to T_{\eta(s)}\mathcal{H}/\eta'(s)$ is the curvature endomorphism defined by $R(\overline{X}) = \overline{R(X,\eta'(s))\eta'(s)}$, where $(X,Y,Z) \to R(X,Y)Z$ is the Riemann curvature tensor of $M$ (in our conventions, $R(X,Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$).

We indicate the proof of Equation (A.2). Fix a point $p = \eta(s_0)$, $s_0 \in (a,b)$, on $\eta$. On a neighborhood $U$ of $p$ in $\mathcal{H}$ we can scale the null vector field $K$ so that $K$ is a geodesic vector field, $\nabla_K K = 0$, and so that $K$, restricted to $\eta$, is the velocity vector field to $\eta$, *i.e.*, for each $s$ near $s_0$, $K_{\eta(s)} = \eta'(s)$. Let $X \in T_pM$. Shrinking $U$ if necessary, we can extend $X$ to a smooth vector field on $U$ so that $[X,K] = \nabla_X K - \nabla_K X = 0$. Then, $R(X,K)K = \nabla_X\nabla_K K - \nabla_K\nabla_X K - \nabla_{[X,K]}K = -\nabla_K\nabla_K X$. Hence along $\eta$ we have, $X'' = -R(X,\eta')\eta'$ (which implies that $X$, restricted to $\eta$, is a Jacobi field along $\eta$). Thus, from Equation (A.3), at the point $p$ we have,

$$\begin{aligned} b'(\overline{X}) &= \overline{\nabla_X K}\,' - b(\overline{\nabla_K X}) = \overline{\nabla_K X}\,' - b(\overline{\nabla_X K}) \\ &= \overline{X''} - b(b(\overline{X})) = -\overline{R(X,\eta')\eta'} - b^2(\overline{X}) \\ &= -R(\overline{X}) - b^2(\overline{X}), \end{aligned} \tag{A.4}$$

which establishes Equation (A.2).

Equation (A.2) leads to the well known Raychaudhuri equation for an irrotational null geodesic congruence in general relativity: by taking the trace of (A.2) we obtain the following formula for the derivative of the null mean curvature $\theta = \theta(s)$ along $\eta$,

$$\theta' = -\mathrm{Ric}(\eta',\eta') - \sigma^2 - \frac{1}{n-2}\theta^2, \tag{A.5}$$

where $\sigma$, the shear scalar, is the trace of the square of the trace free part of $b$. This equation shows how the Ricci curvature of spacetime influences the null mean curvature of a null hypersurface. We note the following:

**Proposition A.1** *Let $\mathcal{H}$ be a $C^2$ null hypersurface in the $(n+1)$ dimensional spacetime $(M,g)$ and let $b$ be the one parameter family of Weingarten maps along an affine parameterized null generator $\eta$. Then the covariant derivative $b'$ defined by Equation (A.3) exists and satisfies Equation (A.2).*

[23] Here $b(\overline{X})$ is an equivalence class of vectors, so it might be useful to give a practical prescription how its derivative $b(\overline{X})'$ can be calculated. Let $s \to c(s)$ be a null generator of $\mathcal{H}$. Let $s \to V(s)$ be a $T\mathcal{H}/K$–vector field along $c$, *i.e.*, for each $s$, $V(s)$ is an element of $T_{c(s)}\mathcal{H}/K$. Say $s \to V(s)$ is smooth if (at least locally) there is a smooth — in the usual sense — vector field $s \to Y(s)$ along $c$ such that $V(s) = \bar{Y}(s)$ for each $s$. Then define the covariant derivative of $s \to V(s)$ along $c$ by: $V'(s) = \overline{Y'(s)}$, where $Y'$ is the usual covariant derivative. It is easily shown, using the fact that $\nabla_K K$ is proportional to $K$, that $V'$ so defined is independent of the choice of $Y$. This definition applies in particular to $b(\bar{X})$.

**Remark A.2** When $\mathcal{H}$ is smooth this is a standard result, proved by the calculation (A.4). However when $\mathcal{H}$ is only $C^2$ all we know is that $b$ is a $C^0$ tensor field so that there is no reason *a priori* that the derivative $b'$ should exist. A main point of the proposition is that it does exist and satisfies the expected differential equation. As the function $s \mapsto R_{\eta(s)}$ is $C^\infty$ then the Riccati equation implies that actually the dependence of $b_{\eta(s)}$ on $s$ is $C^\infty$. This will be clear from the proof below for other reasons.

PROOF: Let $\eta\colon (a,b) \to \mathcal{H}$ be an affinely parameterized null generator of $\mathcal{H}$. To simplify notation we assume that $0 \in (a,b)$ and choose a $C^\infty$ spacelike hypersurface $\Sigma$ of $M$ that passes through $p = \eta(0)$ and let $N = \mathcal{H} \cap \Sigma$. Then $N$ is a $C^2$ hypersurface in $\Sigma$. Now let $\tilde{N}$ be a $C^\infty$ hypersurface in $\Sigma$ so that $\tilde{N}$ has second order contact with $N$ at $p$. Let $\tilde{K}$ be a smooth null normal vector field along $\tilde{N}$ such that at $p$, $\tilde{K} = \eta'(0)$. Consider the hypersurface $\tilde{\mathcal{H}}$ obtained by exponentiating normally along $\tilde{N}$ in the direction $\tilde{K}$; by Lemma 4.15 there are no focal points along $\eta$ as long as $\eta$ stays on $\mathcal{H}$. Passing to a subset of $\tilde{N}$ if necessary to avoid cut points, $\tilde{\mathcal{H}}$ will then be a $C^\infty$ null hypersurface in a neighborhood of $\eta$. Let $B(s)$ and $\tilde{B}(s)$ be the null second fundamental forms of $\mathcal{H}$ and $\tilde{\mathcal{H}}$, respectively, at $\eta(s)$ in the direction $\eta'(s)$. We claim that $\tilde{B}(s) = B(s)$ for all $s \in (a,b)$. Since the null Weingarten maps $\tilde{b} = \tilde{b}(s)$ associated to $\tilde{B} = \tilde{B}(s)$ satisfy Equation (A.2), this is sufficient to establish the lemma.

We first show that $\tilde{B}(s) = B(s)$ for all $s \in [0,c]$ for some $c \in (0,b)$. By restricting to a suitable neighborhood of $p$ we can assume without loss of generality that $M$ is globally hyperbolic. Let $X \in T_p\Sigma$ be the projection of $\eta'(0) \in T_pM$ onto $T_p\Sigma$. By an arbitrarily small second order deformation of $\tilde{N} \subset \Sigma$ (depending on a parameter $\epsilon$ in a fashion similar to Equation (4.4)) we obtain a $C^\infty$ hypersurface $\tilde{N}^+_\epsilon$ in $\Sigma$ which meets $N$ only in the point $p$ and lies to the side of $N$ into which $X$ points. Similarly, we obtain a $C^\infty$ hypersurface $\tilde{N}^-_\epsilon$ in $\Sigma$ which meets $N$ only in the point $p$ and lies to the side of $N$ into which $-X$ points. Let $\tilde{K}^\pm_\epsilon$ be a smooth null normal vector field along $\tilde{N}^\pm_\epsilon$ which agrees with $\eta'(0)$ at $p$. By exponentiating normally along $\tilde{N}^\pm_\epsilon$ in the direction $\tilde{K}^\pm_\epsilon$ we obtain, as before, in a neighborhood of $\eta|_{[0,c]}$ a $C^\infty$ null hypersurface $\tilde{\mathcal{H}}^\pm_\epsilon$, for some $c \in (0,b)$. Let $\tilde{B}^\pm_\epsilon(s)$ be the null second fundamental form of $\tilde{\mathcal{H}}^\pm_\epsilon$ at $\eta(s)$ in the direction $\eta'(s)$.

By restricting the size of $\Sigma$ if necessary we find open sets $W$, $W^\pm_\epsilon$ in $\Sigma$, with $W^-_\epsilon \subset W \subset W^+_\epsilon$, such that $N \subset \partial_\Sigma W$ and $\tilde{N}^\pm_\epsilon \subset \partial_\Sigma W^\pm_\epsilon$. Restricting to a sufficiently small neighborhood of $\eta|_{[0,c]}$, we have $\mathcal{H} \cap J^+(\Sigma) \subset \partial J^+(W)$ and $\tilde{\mathcal{H}}^\pm_\epsilon \cap J^+(\Sigma) \subset \partial J^+(W^\pm_\epsilon)$. Since $J^+(\overline{W^-_\epsilon}) \subset J^+(\overline{W}) \subset J^+(\overline{W^+_\epsilon})$, it follows that $\tilde{\mathcal{H}}^-_\epsilon$ is to the future of $\mathcal{H}$ near $\eta(s)$ and $\mathcal{H}$ is to the future of $\tilde{\mathcal{H}}^+_\epsilon$ near $\eta(s)$, $s \in [0,c]$. Now if two null hypersurfaces $\mathcal{H}_1$ and $\mathcal{H}_2$ are tangent at a point $p$, and $\mathcal{H}_2$ is to the future of $\mathcal{H}_1$, then the difference of the null second fundamental forms $B_2 - B_1$ is positive semidefinite at $p$. We thus obtain $\tilde{B}^-_\epsilon(s) \geq B(s) \geq \tilde{B}^+_\epsilon(s)$. Letting $\epsilon \to 0$, (i.e., letting the deformations go to zero), we obtain $\tilde{B}(s) = B(s)$ for all $s \in [0,c]$. A straightforward continuation argument implies, in fact, that

$\tilde{B}(s) = B(s)$ for all $s \in [0, b)$. A similar argument establishes equality for $s \in (a, 0]$. $\square$

In the last result above the hypersurface $\mathcal{H}$ had to be of at least $C^2$ differentiability class. Now, in our applications we have to consider hypersurfaces $\mathcal{H}$ obtained as a collection of null geodesics normal to a $C^2$ surface. A naive inspection of the problem at hand shows that such $\mathcal{H}$'s could in principle be of $C^1$ differentiability only. Let us show that one does indeed have $C^2$ differentiability of the resulting hypersurface:

**Proposition A.3** *Consider a $C^{k+1}$ spacelike submanifold $N \subset M$ of co–dimension two in an $(n+1)$ dimensional spacetime $(M, g)$, with $k \geq 1$. Let $\mathbf{k}$ be a non-vanishing $C^k$ null vector field along $N$, and let $\mathcal{U} \subseteq \mathbb{R} \times N \to M$ be the set of points where the function*

$$f(t,p) := \exp_p(t\mathbf{k}(p))$$

*is defined. If $f_{(t_0,p_0)*}$ is injective then there is an open neighborhood $\mathcal{O}$ of $(t_0, p_0)$ so that the image $f[\mathcal{O}]$ is a $C^{k+1}$ embedded hypersurface in $M$.*

**Remark A.4** In our application we only need the case $k = 1$. This result is somewhat surprising as the function $p \mapsto \mathbf{k}(p)$ used in the definition of $f$ is only $C^k$. We emphasize that we are *not* assuming that $f$ is injective. We note that $f$ will not be of $C^{k+1}$ differentiability class in general, which can be seen as follows: Let $t \to r(t)$ be a $C^{k+1}$ curve in the $x$-$y$ plane of Minkowski 3-space which is *not* of $C^{k+2}$ differentiability class. Let $t \to \mathbf{n}(t)$ be the spacelike unit normal field along the curve in the $x$-$y$ plane, then $t \to \mathbf{n}(t)$ is $C^k$ and is *not* $C^{k+1}$. Let $T = (0, 0, 1)$ be the unit normal to the $x$-$y$ plane. Then $K(t) = \mathbf{n}(t) + T$ is a $C^k$ normal null field along $t \to r(t)$. The normal exponential map $f \colon R^2 \to \mathbb{R}^3$ in the direction $K$ is given by $f(s,t) = r(t) + s[\mathbf{n}(t) + T]$, and hence $df/dt = r'(t) + sn'(t)$, showing explicitly that the regularity of $f$ can be no greater than the regularity of $\mathbf{n}(t)$, and hence no greater than the regularity of $r'(t)$.

PROOF: This result is local in $N$ about $p_0$ so there is no loss of generality, by possibly replacing $N$ by a neighborhood of $p_0$ in $N$, in assuming that $N$ is a embedded submanifold of $M$. The map $f$ is of class $C^k$ and the derivative $f_{(t_0,p_0)*}$ is injective so the implicit function theorem implies $f[\mathcal{U}]$ is a $C^k$ hypersurface near $f(t_0, p_0)$. Let $\eta$ be any nonzero timelike $C^\infty$ vector field on $M$ defined near $p_0$ (some restrictions to be put on $\eta$ shortly) and let $\Phi_s$ be the flow of $\eta$. Then for sufficiently small $\varepsilon$ the map $\tilde{f} \colon (-\varepsilon, \varepsilon) \times N \to M$ given by

$$\tilde{f}(s,p) := \Phi_s(p)$$

is injective and of class $C^{k+1}$. Extend $\mathbf{k}$ to any $C^k$ vector field $\tilde{\mathbf{k}}$ along $\tilde{f}$. (It is not assumed that the extension $\tilde{k}$ is null.) That is $\tilde{\mathbf{k}} \colon (-\varepsilon, \varepsilon) \times N \to TM$ is a $C^k$

map and $\tilde{\mathbf{k}}(s,p) \in T_{\tilde{f}(s,p)}M$. Note that we can choose $\tilde{\mathbf{k}}(s,p)$ so that the covariant derivative $\frac{\nabla \tilde{\mathbf{k}}}{\partial s}(0,p_0)$ has any value we wish at the one point $(0,p_0)$. Define a map $F\colon (t_0-\varepsilon, t_0+\varepsilon) \times (-\varepsilon,\varepsilon) \times N \to M$ by

$$F(t,s,p) = \exp(t\tilde{\mathbf{k}}(s,p)).$$

We now show that $F$ can be chosen to be a local diffeomorphism near $(t_0,0,p_0)$. Note that $F(t,0,p) = f(t,p)$ and by assumption $f_{*(t_0,p_0)}$ is injective. Therefore the restriction of $F_{*(t_0,0,p_0)}$ to $T_{(t_0,p_0)}(\mathbb{R}\times N) \subset T_{(t_0,0,p_0)}(\mathbb{R}\times\mathbb{R}\times N)$ is injective. Thus by the inverse function theorem it is enough to show that $F_{*(t_0,0,p_0)}(\partial/\partial s)$ is linearly independent of the subspace $F_{*(t_0,0,p_0)}[T_{(t_0,p_0)}(\mathbb{R}\times N)]$. Let

$$V(t) = \frac{\partial F}{\partial s}(t,s,p_0)\Big|_{s=0} .$$

Then $V(t_0) = F_{*(t_0,0,p_0)}(\partial/\partial s)$ and our claim that $F$ is a local diffeomorphism follows if $V(t_0) \notin F_{*(t_0,0,p_0)}[T_{(t_0,p_0)}(\mathbb{R}\times N)]$. For each $s,p$ the map $t \mapsto F(s,t,p)$ is a geodesic and therefore $V$ is a Jacobi field along $t \mapsto F(0,t,p_0)$. (Those geodesics might change type as $s$ is varied at fixed $p_0$, but this is irrelevant for our purposes.) The initial conditions of this geodesic are

$$V(0) = \frac{\partial}{\partial s}F(0,s,p_0)\Big|_{s=0} = \frac{\partial}{\partial s}\Phi_s(p_0)\Big|_{s=0} = \eta(p_0)$$

and

$$\frac{\nabla V}{\partial t}(0) = \frac{\nabla}{\partial t}\frac{\nabla}{\partial s}F(t,s,p_0)\Big|_{s=0,t=0} = \frac{\nabla}{\partial s}\frac{\nabla}{\partial t}F(t,s,p_0)\Big|_{s=0,t=0} = \frac{\nabla \tilde{\mathbf{k}}}{\partial s}(0,p_0) .$$

From our set up we can choose $\eta(p_0)$ to be any timelike vector and $\frac{\nabla \tilde{\mathbf{k}}}{\partial s}(0,p_0)$ to be any vector. As the linear map from $T_{p_0}M \times T_{p_0}M \to T_{f(t_0,p_0)}M$ which maps the initial conditions $V(0)$, $\frac{\nabla V}{\partial t}(0)$ of a Jacobi field $V$ to its value $V(t_0)$ is surjective[24] it is an open map. Therefore we can choose $\eta(p_0)$ and $\frac{\nabla \tilde{\mathbf{k}}}{\partial s}(0,p_0)$ so that $V(t_0)$ is not in the nowhere dense set $F_{*(t_0,0,p_0)}[T_{(t_0,p_0)}(\mathbb{R}\times N)]$. Thus we can assume $F$ is a local $C^k$ diffeomorphism on some small neighborhood $\mathcal{A}$ of $(t_0,0,p_0)$ onto a small neighborhood $\mathcal{B} := F[\mathcal{A}]$ of $F(t_0,0,p_0)$ as claimed.

Consider the vector field $F_*(\partial/\partial t) = \partial F/\partial t$ along $F$. Then the integral curves of this vector field are the geodesics $t \mapsto F(t,s,p) = \exp(t\tilde{\mathbf{k}}(s,p))$. (This is true even when $F$ is not injective on its entire domain.) These geodesics and their velocity vectors depend smoothly on the initial data. In the case at hand the initial data is $C^k$ so $\partial F/\partial t$ is a $C^k$ vector field along $F$. Therefore the one form $\alpha$ defined by $\alpha(X) := \langle X, \partial F/\partial t\rangle$ on the neighborhood $\mathcal{B}$ of $q_0$ is $C^k$. The

[24] If $v \in T_{f(t_0,p_0)}N$ there is a Jacobi field with $V(t_0) = v$ and $\frac{\nabla V}{\partial t}(t_0) = 0$, which implies subjectivity.

definition of $F$ implies that $f(t,p) = F(t,0,p)$ and therefore the vector field $\partial F/\partial t$ is tangent to $f[\mathcal{O}]$ and the null geodesics $t \mapsto f(t,p) = F(t,0,p)$ rule $f[\mathcal{O}]$ so that $f[\mathcal{O}]$ is a null hypersurface. Therefore for any vector $X$ tangent to $f[\mathcal{O}]$ we have $\alpha(X) = \langle X, \partial F/\partial t\rangle = 0$. Thus $f[\mathcal{O}]$ is an integral submanifold for the distribution $\{X \mid \alpha(X) = 0\}$ defined by $\alpha$. But, as is easily seen by writing out the definitions in local coordinates, an integral submanifold of a $C^k$ distribution is a $C^{k+1}$ submanifold. (Note that in general there is no reason to believe that the distribution defined by $\alpha$ is integrable. However, we have shown directly that $f[\mathcal{O}]$ is an integral submanifold of that distribution.) $\square$

We shall close this appendix with a calculation, needed in the main body of the paper, concerning Jacobians. Let us start by recalling the definition of the Jacobian needed in our context. Let $\phi\colon M \to N$ be a $C^1$ map between Riemannian manifolds, with $\dim M \le \dim N$. Let $n = \dim M$ and let $e_1, \ldots, e_n$ be an orthonormal basic of $T_pM$ then the Jacobian of $\phi$ at $p$ is $J(\phi)(p) = \|\phi_{*p}e_1 \wedge \phi_{*p}e_2 \wedge \cdots \wedge \phi_{*p}e_n\|$. When $\dim M = \dim N$ and both $M$ and $N$ are oriented with $\omega_M$ being the volume form on $M$, and $\omega_N$ being the volume form on $N$, then $J(\phi)$ can also be described as the positive scalar satisfying: $\phi^*(\omega_N) = \pm J(\phi)\,\omega_M$.

Let $S$ be a $C^2$ co-dimension two acausal spacelike submanifold of a smooth spacetime $M$, and let $K$ be a past directed $C^1$ null vector field along $S$. Consider the normal exponential map in the direction $K$, $\Phi\colon \mathbb{R} \times S \to M$, defined by $\Phi(s,x) = \exp_x sK$. ($\Phi$ need not be defined on all of $\mathbb{R} \times S$.) Suppose the null geodesic $\eta\colon s \to \Phi(s,p)$ meets a given acausal spacelike hypersurface $\Sigma$ at $\eta(1)$. Then there is a neighborhood $W$ of $p$ in $S$ such that each geodesic $s \to \Phi(s,x)$, $x \in W$ meets $\Sigma$, and so determines a $C^1$ map $\phi\colon W \to \Sigma$, which is the projection into $\Sigma$ along these geodesics. Let $J(\phi)$ denote the Jacobian determinant of $\phi$ at $p$. $J(\phi)$ may be computed as follows. Let $\{X_1, X_2, \ldots, X_k\}$ be an orthonormal basis for the tangent space $T_pS$. Then ,

$$J(\phi) = \|\phi_{*p}X_1 \wedge \phi_{*p}X_2 \wedge \cdots \wedge \phi_{*p}X_k\|.$$

Suppose there are no focal points to $S$ along $\eta|_{[0,1]}$. Then by shrinking $W$ and rescaling $K$ if necessary, $\Phi\colon [0,1] \times W \to M$ is a $C^1$ embedded null hypersurface $N$ such that $\Phi(\{1\} \times W) \subset \Sigma$. Extend $K$ to be the $C^1$ past directed null vector field, $K = \Phi_*(\frac{\partial}{\partial s})$ on $N$. Let $\theta = \theta(s)$ be the null mean curvature of $N$ with respect to $-K$ along $\eta$. For completeness let us give a proof of the following, well known result:

**Proposition A.5** *With $\theta = \theta(s)$ as described above,*

1. *If there are no focal points to $S$ along $\eta|_{[0,1]}$, then*

$$J(\phi) = \exp\Big(-\int_0^1 \theta(s)ds\Big)\ . \tag{A.6}$$

2. *If $\eta(1)$ is the first focal point to $S$ along $\eta|_{[0,1]}$, then*

$$J(\phi) = 0\ .$$

**Remark A.6** In particular, if $N$ has nonnegative null mean curvature with respect to the future pointing null normal, *i.e.*, if $\theta \geq 0$, we obtain that $J(\phi) \leq 1$.

**Remark A.7** Recall that $\theta$ was only defined when a normalization of $K$ has been chosen. We stress that in (A.6) that normalization is so that $K$ is tangent to an affinely parameterized geodesic, with $s$ being an affine distance along $\eta$, and with $p$ corresponding to $s = 0$ and $\phi(p)$ corresponding to $s = 1$.

PROOF: 1. To relate $J(\phi)$ to the null mean curvature of $N$, extend the orthonormal basis $\{X_1, X_2, \ldots, X_k\}$ to Lie parallel vector fields $s \to X_i(s)$, $i = 1, \ldots, k$, along $\eta$, $\mathcal{L}_K X_i = 0$ along $\eta$. Then by a standard computation,

$$\begin{aligned} J(\phi) &= \|\phi_{*p}X_1 \wedge \phi_{*p}X_2 \wedge \cdots \wedge \phi_{*p}X_k\| \\ &= \|X_1(1) \wedge X_2(1) \wedge \cdots \wedge X_k(1)\| \\ &= \sqrt{g}\,\Big|_{s=1}\ , \end{aligned}$$

where $g = \det[g_{ij}]$, and $g_{ij} = g_{ij}(s) = \langle X_i(s), X_j(s)\rangle$. We claim that along $\eta$,

$$\theta = -\frac{1}{\sqrt{g}}\frac{d}{ds}\sqrt{g}\ .$$

The computation is standard. Set $b_{ij} = B(\overline{X}_i, \overline{X}_j)$, where $B$ is the null second fundamental form of $N$ with respect to $-K$, $h_{ij} = h(\overline{X}_i, \overline{X}_j) = g_{ij}$, and let $g^{ij}$ be the $i,j$th entry of the inverse matrix $[g_{ij}]^{-1}$. Then $\theta = g^{ij}b_{ij}$. Differentiating $g_{ij}$ along $\eta$ we obtain,

$$\begin{aligned} \frac{d}{ds}g_{ij} = K\langle X_i, X_j\rangle &= \langle \nabla_K X_i, X_j\rangle + \langle X_i, \nabla_K X_j\rangle \\ &= \langle \nabla_{X_i} K, X_j\rangle + \langle X_i, \nabla_{X_j} K\rangle \\ &= -(b_{ij} + b_{ji}) = -2b_{ij}\ . \end{aligned}$$

Thus,

$$\theta = g^{ij}b_{ij} = -\frac{1}{2}g^{ij}\frac{d}{ds}g_{ij} = -\frac{1}{2}\frac{1}{g}\frac{dg}{ds} = -\frac{1}{\sqrt{g}}\frac{d}{ds}\sqrt{g}\,,$$

as claimed. Integrating along $\eta$ from $s = 0$ to $s = 1$ we obtain,

$$J(\phi) = \sqrt{g}\,\Big|_{s=1} = \sqrt{g}\,\Big|_{s=0} \cdot \exp\Big(-\int_0^1 \theta\, ds\Big) = \exp\Big(-\int_0^1 \theta\, ds\Big)\,.$$

2. Suppose now that $\eta(1)$ is a focal point to $S$ along $\eta$, but that there are no focal points to $S$ along $\eta$ prior to that. Then we can still construct the $C^1$ map $\Phi\colon [0,1]\times W \to M$, with $\Phi(\{1\}\times W)\subset \Sigma$, such that $\Phi$ is an embedding when restricted to a sufficiently small open set in $[0,1]\times W$ containing $[0,1)\times\{p\}$. The vector fields $s\to X_i(s)$, $s\in[0,1)$, $i=1,..,k$, may be constructed as above, and are Jacobi fields along $\eta|_{[0,1)}$, which extend smoothly to $\eta(1)$. Since $\eta(1)$ is a focal point, the vectors $\phi_* X_1 = X_1(1)$, …, $\phi_* X_k = X_k(1)$ must be linearly dependent, which implies that $J(\phi)=0$. □

# B Some comments on the area theorem of Hawking and Ellis

In this appendix we wish to discuss the status of our $\mathcal{H}$–regularity condition with respect to the conformal completions considered by Hawking and Ellis [37] in their treatment of the area theorem. For the convenience of the reader let us recall here the setting of [37]. One of the conditions of the Hawking–Ellis area theorem [37, Proposition 9.2.7, p. 318] it is assumed that spacetime $(\mathscr{M},g)$ is *weakly asymptotically simple and empty* ("WASE", [37, p. 225]). This means that there exists an open set $\mathscr{U}\subset\mathscr{M}$ which is isometric to $\mathscr{U}'\cap\mathscr{M}'$, where $\mathscr{U}'$ is a neighborhood of null infinity in an asymptotically simple and empty (ASE) spacetime $(\mathscr{M}',g')$ [37, p. 222]. It is further assumed that $\mathscr{M}$ admits a partial Cauchy surface $\mathscr{S}$ with respect to which $\mathscr{M}$ is *future asymptotically predictable* ([37], p. 310). This is defined by the requirement that $\mathscr{I}^+$ is contained in the closure of the future domain of dependence $\mathcal{D}^+(\mathscr{S};\mathscr{M})$ of $\mathscr{S}$, where the closure is taken in the conformally completed manifold $\bar{\mathscr{M}}=\mathscr{M}\cup\mathscr{I}^+\cup\mathscr{I}^-$, with both $\mathscr{I}^+$ and $\mathscr{I}^-$ being null hypersurfaces. Next, one says that $(\mathscr{M},g)$ is *strongly future asymptotically predictable* ([37], p. 313) if it is future asymptotically predictable and if $J^+(\mathscr{S})\cap\bar{J}^-(\mathscr{I}^+;\bar{\mathscr{M}})$ is contained in $\mathcal{D}^+(\mathscr{S};\mathscr{M})$. Finally ([37], p. 318), $(\mathscr{M},g)$ is said to be a *regular predictable space* if $(\mathscr{M},g)$ is strongly future asymptotically predictable and if the following three conditions hold:

($\alpha$) $\mathscr{S}\cap\bar{J}^-(\mathscr{I}^+;\bar{\mathscr{M}})$ is homeomorphic to $\mathbb{R}^3\backslash$(an open set with compact closure).

($\beta$) $\mathscr{S}$ is simply connected.

($\gamma$) the family of hypersurfaces $\mathscr{S}(\tau)$ constructed in [37, Proposition 9.2.3, p. 313] has the property that for sufficiently large $\tau$ the sets $\mathscr{S}(\tau)\cap\bar{J}^-(\mathscr{I}^+;\bar{\mathscr{M}})$ are contained in $\bar{J}^+(\mathscr{I}^-;\bar{\mathscr{M}})$.

It is then asserted in [37, Proposition 9.2.7, p. 318] that the area theorem holds for regular predictable spaces satisfying the null energy condition.

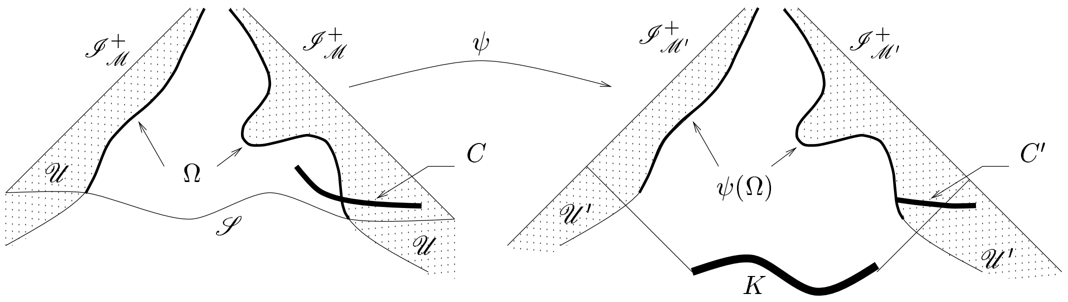


Figure 2: The set $\Omega \equiv J^+(\mathscr{S};\mathscr{M}) \cap \partial\mathscr{U}$ and its image under $\psi$.

Now in the proof of [37, Proposition 9.2.1, p. 311] (which is one of the results used in the proof of [37, Proposition 9.2.7, p. 318]) Hawking and Ellis write: "This shows that if $\mathscr{W}$ is any compact set of $\mathscr{S}$, every generator of $\mathscr{I}^+$ leaves $J^+(\mathscr{W};\bar{\mathscr{M}})$." The justification of this given in [37] is wrong[25]. If one is willing to impose this as a supplementary hypothesis, then this condition can be thought of as the Hawking–Ellis equivalent of our condition of $\mathcal{H}$–regularity of $\mathscr{I}^+$. When such a condition is imposed in addition to the hypothesis of strong asymptotic predictability and weak asymptotic simplicity ("WASE") of $(\mathscr{M}, g)$, then the hypotheses of Proposition 4.1 hold, and the conclusions of our version of the area theorem, Theorem 1.1, apply.

An alternative way to guarantee that the hypotheses of Proposition 4.1 will hold in the "future asymptotically predictable WASE" set–up of [37] (for those sets $C$ which lie to the future of $\mathscr{S}$) is to impose some mild additional conditions on $\mathscr{U}$ and $\mathscr{S}$. There are quite a few possibilities, one such set of conditions is as follows: Let $\psi : \mathscr{U} \to \mathscr{U}' \cap \mathscr{M}'$ denote the isometry arising in the definition of the WASE spacetime $\mathscr{M}$. First, we require that $\psi$ can be extended by continuity to a continuous map, still denoted by $\psi$, defined on $\overline{\mathscr{U}}$. Next, suppose there exists a compact set $K \subset \mathscr{M}'$ such that,

$$\psi(J^+(\mathscr{S};\mathscr{M}) \cap \partial\mathscr{U}) \subset J^+(K;\mathscr{M}') \,, \tag{B.1}$$

see Figure 2. Let us show that, under the future asymptotically predictable WASE conditions together with (B.1), for every compact set $C \subset J^+(\mathscr{S};\mathscr{M})$

[25]In the proof of [37, Proposition 9.2.1, p. 311] it is claimed that "... Then $\mathscr{S}' \setminus \mathscr{U}'$ is compact...". This statement is incorrect in general, as shown by the example $(\mathscr{M}, g) = (\mathscr{M}', g') = (\mathbb{R}^4, \mathrm{diag}(-1,+1,+1,+1))$, $\mathscr{S} = \mathscr{S}' = \{t = 0\}$, $\mathscr{U} = \mathscr{U}' = \{t \neq 0\}$. This example does not show that the claim is wrong, but that the proof is; we do not know whether the claim in Proposition 9.2.1 is correct as stated under the hypothesis of future asymptotic predictability of $(\mathscr{M}, g)$ made there. Let us note that the conditions $(\alpha)$–$(\gamma)$ do not seem to be used anywhere in the proof of Proposition 9.2.7 as presented in [37], and it is conceivable that the authors of [37] had in mind some use of those conditions in the proof of Propostion 9.2.1. We have not investigated in detail whether or not the assertion made there can be justified if the supplementary hypothesis that $(\mathscr{M}, g)$ is a regular predictable space is made, as the approach we advocate in Section 4.1 allows one to avoid the "WASE" framework altogether.

that meets $I^-(\mathscr{I}^+;\bar{\mathscr{M}})$ there exists a future inextendible (in $\mathscr{M}$) null geodesic $\eta \subset \partial J^+(C;\mathscr{M})$ starting on $C$ and having future end point on $\mathscr{I}^+$. First, we claim that

$$\psi(J^+(C;\mathscr{M})\cap\mathscr{U}) \subset J^+(K\cup C';\mathscr{M}') , \tag{B.2}$$

where $C' = \psi(C\cap\bar{\mathscr{U}})$. Indeed, let $p \in \psi(J^+(C;\mathscr{M})\cap\mathscr{U})$, therefore there exists a future directed causal curve $\gamma$ from $C$ to $\psi^{-1}(p) \in \mathscr{U}$. If $\gamma \subset \mathscr{U}$, then $p \in J^+(C';\mathscr{M}')$. If not, then $\gamma$ exits $\mathscr{U}$ when followed from $\psi^{-1}(p)$ to the past at some point in $J^+(\mathscr{S};\mathscr{M})\cap\partial\mathscr{U}$, and thus $p \in \psi(J^+(\mathscr{S};\mathscr{M})\cap\partial\mathscr{U}) \subset J^+(K;\mathscr{M}')$, which establishes (B.2). Since $K\cup C'$ is compact, by Lemma 4.5 and Proposition 4.13 in [54], each generator of $\mathscr{I}^+_{\mathscr{M}'}$ meets $\partial J^+(K\cup C';\bar{\mathscr{M}}')$ exactly once. It follows that, under the natural identification of $\mathscr{I}^+_{\mathscr{M}}$ with $\mathscr{I}^+_{\mathscr{M}'}$, the criteria for $\mathcal{H}$-regularity discussed in Remark 4.5 are satisfied. Hence, we may apply Proposition 4.8 to obtain the desired null geodesic $\eta$.

There exist several other proposals how to modify the WASE conditions of [37] to obtain better control of the space–times at hand [16, 50, 53], but we have not investigated in detail their suitability to the problems considered here.

# C Proof of Theorem 5.6

For $q \in S_0$ let $\Gamma_q \subset \mathcal{H}$ denote the generator of $\mathcal{H}$ passing through $q$. Throughout this proof all curves will be parameterized by signed $\sigma$–distance from $S_0$, with the distance being negative to the past of $S_0$ and positive to the future. We will need the following Lemma:

**Lemma C.1** *$S_0$ is a Borel subset of $\mathcal{S}$, in particular $S_0$ is $\mathfrak{H}^{n-1}_\sigma$ measurable.*

PROOF: For each $\delta > 0$ let

$$\begin{aligned} A_\delta : &= \{p\in S_0 \mid \text{the domain of definition of } \Gamma_p \\ &\qquad\qquad \text{contains the interval } [-\delta,\delta]\} . \end{aligned}\tag{C.1}$$

$$\begin{aligned} B_\delta : &= \{p\in S_0 \mid \text{the domain of definition of the inextendible geodesic} \\ &\qquad\qquad \gamma_p \text{ containing } \Gamma_p \text{ contains the interval } [-\delta,\delta]\} . \end{aligned}\tag{C.2}$$

Lower semi–continuity of existence time of geodesics shows that the $B_\delta$'s are open subsets of $\mathcal{S}$. Clearly $A_{\delta'} \subset B_\delta$ for $\delta' \ge \delta$. We claim that for $\delta' \ge \delta$ the sets $A_{\delta'}$ are closed subsets of the $B_\delta$'s. Indeed, let $q_i \in A_{\delta'}\cap B_\delta$ be a sequence such that $q_i \to q_\infty \in B_\delta$. Since the generators of $\mathcal{H}$ never leave $\mathcal{H}$ to the future, and since $q_\infty \in B_\delta$, it immediately follows that the domain of definition of $\Gamma_{q_\infty}$ contains the interval $[0,\delta]$. Suppose, for contradiction, that $\Gamma_{q_\infty}(s_-)$ is an endpoint on $\mathcal{H}$ with $s_- \in (-\delta, 0]$, hence there exists $s' \in (-\delta,0]$ such that $\gamma_{q_\infty}(s') \in I^-(\mathcal{H})$. As $q_\infty$ is an interior point of $\Gamma_{q_\infty}$, the $\sigma$–unit tangents to the $\Gamma_{q_i}$'s at $q_i$ converge to the $\sigma$–unit tangent to $\Gamma_{q_\infty}$ at $q_\infty$. Now $I^-(\mathcal{H})$ is open, and continuous dependence of ODE's upon initial data shows that $\gamma_{q_i}(s') \in I^-(\mathcal{H})$ for $i$ large enough, contradicting the fact that $q_i \in A_{\delta'}$ with $\delta' \geq \delta$. It follows that $q_\infty \in A_{\delta'}$, and that $A_{\delta'}$ is closed in $B_\delta$. But a closed subset of an open set is a Borel set, hence $A_{\delta'}$ is Borel in $\mathcal{S}$. Clearly

$$S_0 = \cup_i A_{1/i} \ ,$$

which implies that $S_0$ is a Borel subset of $\mathcal{S}$. The $\mathfrak{H}^{n-1}_\sigma$–measurability of $S_0$ follows now from [25, p. 293] or [19, p. 147][11]. □

Returning to the proof of Theorem 5.6, set

$$\begin{aligned} \Gamma_q^+ &= \Gamma_q \cap J^+(q) \ , \\ \mathcal{H}_{\rm sing} &= \mathcal{H} \setminus \mathcal{H}_{\mathcal{A}l} \ , \\ \Omega &= \cup_{q\in S_0} \Gamma_q^+ \ , \\ \Omega_{\rm sing} &= \Omega \cap \mathcal{H}_{\rm sing} \ . \end{aligned}$$

By definition we have $\Omega_{\rm sing} \subset \mathcal{H}_{\rm sing}$ and completeness of the Hausdorff measure[26] together with $\mathfrak{H}^n_\sigma(\mathcal{H}_{\rm sing}) = 0$ implies that $\Omega_{\rm sing}$ is $n$–Hausdorff measurable, with

$$\mathfrak{H}^n_\sigma(\Omega_{\rm sing}) = 0 \ . \tag{C.3}$$

Let $\phi\colon \Omega \to S_0$ be the map which to a point $p \in \Gamma_q^+$ assigns $q \in S_0$. The arguments of the proofs of Lemmata 6.9 and 6.11 show that $\phi$ is locally Lipschitz. This, together with Lemma C.1, allows us to use the co–area formula [23, Theorem 3.1] to infer from (C.3) that

$$0 = \int_{\Omega_{\rm sing}} J(\phi) d\mathfrak{H}^n_\sigma = \int_{S_0} \mathfrak{H}^1_\sigma(\Omega_{\rm sing} \cap \phi^{-1}(q)) d\mathfrak{H}^{n-1}_\sigma(q) \ ,$$

where $J(\phi)$ is the Jacobian of $\phi$, *cf.* [23, p. 423]. Hence

$$\mathfrak{H}^1_\sigma(\Omega_{\rm sing} \cap \phi^{-1}(q)) = 0 \tag{C.4}$$

for almost all $q$'s in $S_0$. A chase through the definitions shows that (C.4) is equivalent to

$$\mathfrak{H}^1_\sigma(\Gamma_q^+ \setminus \mathcal{H}_{\mathcal{A}l}) = 0 \ , \tag{C.5}$$

for almost all $q$'s in $S_0$. Clearly the set of Alexandrov points of $\mathcal{H}$ is dense in $\Gamma_q^+$ when (C.5) holds. Theorem 5.1 shows, for such $q$'s, that all interior points of $\Gamma_q$ are Alexandrov points of $\mathcal{H}$, hence points $q$ satisfying (C.5) are in $S_1$. It follows that $S_1$ is $\mathfrak{H}^{n-1}_\sigma$ measurable, and has full $(n-1)$–Hausdorff measure in $S_0$. This establishes our claim about $S_1$. The claim about $S_2$ follows now from the inclusion $S_1 \subset S_2$. □

[26] A measure is *complete* iff all sets of outer measure zero are measurable. Hausdorff measure is constructed from an outer measure using Carathéodory's definition of measurable sets [24, p. 54]. All such measures are complete [24, Thm 2.1.3 pp. 54–55].

# D Proof of Proposition 6.6

Because of the identities

$$\begin{aligned}&\Big(f(p)+\langle x-p,a_p\rangle-\frac{C}{2}\|x-p\|^2\Big)+\frac{C}{2}\|x\|^2\\&=\Big(f(p)+\frac{C}{2}\|p\|^2\Big)+\langle x-p,a_p+Cp\rangle\ ,\end{aligned}$$

and

$$\begin{aligned}&\Big(f(q)+\langle x-q,a_q\rangle+\frac{C}{2}\|x-q\|^2\Big)+\frac{C}{2}\|x\|^2\\&=\Big(f(q)+\frac{C}{2}\|q\|^2\Big)+\langle x-q,a_q+Cq\rangle+C\|x-q\|^2\ ,\end{aligned}$$

we can replace $f$ by $x\mapsto f(x)+C\|x\|^2/2$ and $a_p$ by $a_p+Cp$ and assume that for all $p,x\in A$ we have

$$f(p)+\langle x-p,a_p\rangle\le f(x)\le f(p)+\langle x-p,a_p\rangle+C\|x-p\|^2\ ,\tag{D.1}$$

and for all $p,q\in A$ and $x\in\mathbb{R}^n$

$$f(p)+\langle x-p,a_p\rangle\le f(q)+\langle x-q,a_q\rangle+C\|x-q\|^2\ .\tag{D.2}$$

These inequalities can be given a geometric form that is easier to work with. Let $P:=\{(x,y)\in\mathbb{R}^n\times\mathbb{R}:y>C\|x\|^2\}$. Then $P$ is an open convex solid paraboloid of $\mathbb{R}^{n+1}$. We will denote the closure of $P$ by $\overline{P}$. From the identity

$$f(q)+\langle x-q,a_q\rangle+C\|x-q\|^2=(f(q)-(4C)^{-1}\|a_q\|^2)+C\|x-(q-(2C)^{-1}a_q)\|^2$$

it follows that the solid open paraboloids $\{(x,y)\in\mathbb{R}^n\times\mathbb{R}:y>f(q)+\langle x-q,a_q\rangle+C\|x-q\|^2\}$ are translates in $\mathbb{R}^{n+1}$ of $P$. Let $G[f]:=\{(x,y)\in\mathbb{R}^n\times\mathbb{R}:x\in A,\ y=f(x)\}$ be the graph of $f$. The inequalities (D.1) and (D.2) imply that for each $p\in A$ there is an affine hyperplane $H_p=\{(x,y)\in\mathbb{R}^n\times\mathbb{R}:y=f(p)+\langle x-p,a_p\rangle\}$ of $\mathbb{R}^{n+1}$ and a vector $b_p\in\mathbb{R}^{n+1}$ so that

$$(p,f(p))\in H_p,\tag{D.3}$$

$$(P+b_p)\cap G[f]=\varnothing\quad\text{but}\quad(p,f(p))\in b_p+\overline{P},\tag{D.4}$$

and for all $p,q\in A$

$$H_p\cap(b_q+P)=\varnothing.\tag{D.5}$$

As the paraboloids open up, this last condition implies that each $b_q+P$ lies above all the hyperplanes $H_p$.

Let

$$Q:=\text{Convex Hull}\Big(\bigcup_{p\in A}(b_p+P)\Big).$$

Because $P$ is convex if $\alpha_1, \ldots, \alpha_m \geq 0$ satisfy $\sum_{i=1}^m \alpha_i = 1$ then for any $v_1, \ldots, v_m \in \mathbb{R}^{n+1}$

$$\alpha_1(v_1 + P) + \alpha_2(v_2 + P) + \cdots + \alpha_m(v_m + P) = (\alpha_1 v_1 + \cdots + \alpha_m v_m) + P.$$

Therefore if

$$B := \text{Convex Hull}\,\{b_p : p \in A\}$$

then

$$Q = \bigcup_{v \in B} (v + P)$$

so that $Q$ is a union of translates of $P$. Thus $Q$ is open. Because $P$ is open we have that if $\lim_{\ell\to\infty} v_\ell = v$ then $v + P \subseteq \bigcup_\ell (v_\ell + P)$. So if $\overline{B}$ is the closure of $B$ in $\mathbb{R}^{n+1}$ then we also have

$$Q = \bigcup_{v \in \overline{B}} (v + P). \tag{D.6}$$

For each $p$ the open half space above $H_p$ is an open convex set and $Q$ is the convex hull of a subset of this half space. Therefore $Q$ is contained in this half space. Therefore $Q \cap H_p = \varnothing$ for all $p$.

We now claim that for each point $z \in \partial Q$ there is a supporting paraboloid for $z$ in the sense that there is a vector $v \in \overline{B}$ with $v + P \subset Q$ and $z \in v + \overline{P}$. To see this note that as $z \in \partial Q$ there is a sequence $\{b_\ell\}_{\ell=1}^\infty \subset B$ and $\{w_\ell\}_{\ell=1}^\infty \subset P$ so that $\lim_{\ell\to\infty}(b_\ell + w_\ell) = z$. Fixing a $p_0 \in A$ and using that all the sets $b_\ell + P$ are above the hyperplane $H_{p_0}$ we see that both the sequences $b_\ell$ and $w_\ell$ are bounded subsets of $\mathbb{R}^{n+1}$ and by going to a subsequence we can assume that $v := \lim_{\ell\to\infty} b_\ell$ and $w := \lim_{\ell\to\infty} w_\ell$ exist. Then $z = v + w$, $v \in \overline{B}$ and $w \in \overline{P}$. Then (D.6) implies $v + P \subset Q$ and $w \in \overline{P}$ implies $z \in v + \overline{P}$. Thus we have the desired supporting paraboloid.

We also claim that the graph $G[f]$ satisfies $G[f] \subset \partial Q$. This is because for $p \in A$ the point then $(p, f(p)) \in H_p$ and $H_p$ is disjoint from $Q$. Thus $(p, f(p)) \notin Q$. But from (D.4) $(p, f(p)) \in b_p + \overline{P}$ and as $b_p + \overline{P} \subset \overline{Q}$ this implies $(p, f(p)) \in \overline{Q}$. Therefore $(p, f(p)) \in \partial Q$ as claimed.

Let $F\colon \mathbb{R}^n \to \mathbb{R}$ be the function that defines $\partial Q$. Explicitly

$$F(x) = \inf\{y : (x, y) \in Q\}.$$

This is a function defined on all of $\mathbb{R}^n$ and $G[f] \subset \partial Q$ implies that this function extends $f$. For any $x_0 \in \mathbb{R}^n$ the convexity of $Q$ implies there is a supporting hyperplane $H$ for $Q$ at its boundary point $(x_0, F(x_0))$ and we have seen there also a supporting paraboloid $v + P$ for $\partial Q = G[F]$ at $(x_0, F(x_0))$. Expressing $H$ and $\partial(v + P)$ as graphs over $\mathbb{R}^n$ these geometric facts yield that there is a vector $a_{x_0}$ so that the inequalities

$$F(x_0) + \langle x - x_0, a_{x_0}\rangle \leq F(x) \leq F(x_0) + \langle x - x_0, a_{x_0}\rangle + C\|x - x_0\|^2$$

hold for all $x \in \mathbb{R}^n$. Because the function $F$ is defined on all of $\mathbb{R}^n$ (rather than just the subset $A$), it follows from [9, Prop. 1.1 p. 7] that $F$ is of class $C^{1,1}$. $\square$

**Acknowledgments** P.T.C. acknowledges useful discussions with or comments from Guy Barles, Piotr Hajłasz, Tom Ilmanen, Bernd Kirchheim, Andrzej Królak, Olivier Ley and Laurent Véron. R.H. benefited from discussions and/or correspondence with Joe Fu, Sergei Konyagin, and Ron DeVore. We are grateful to Lars Andersson for bibliographical advice.

# References

[1] L. Andersson, G.J. Galloway, and R. Howard, *The cosmological time function*, Class. Quantum Grav. **15** (1998), 309–322, gr-qc/9709084.

[2] ———, *A strong maximum principle for weak solutions of quasi-linear elliptic equations with applications to Lorentzian and Riemannian geometry*, Comm. Pure Appl. Math. **51** (1998), 581–624.

[3] L. Andersson and V. Moncrief, *The global existence problem in general relativity*, 1999, Proceedings of the Besse Seminar on Lorentzian Geometry, Nancy, June 1998.

[4] F. Antonacci and P. Piccione, *A Fermat principle on Lorentzian manifolds and applications*, Appl. Math. Lett. **9** (1996), 91–95.

[5] J. K. Beem, P. E. Ehrlich, and K. L. Easley, *Global Lorentzian geometry*, 2 ed., Pure and Applied Mathematics, vol. 202, Marcel Dekker, New York, 1996.

[6] J.K. Beem and A. Królak, *Cauchy horizon endpoints and differentiability*, Jour. Math. Phys. **39** (1998), 6001–6010, gr-qc/9709046.

[7] D. Brill, J. Louko, and P. Peldan, *Thermodynamics of (3+1)-dimensional black holes with toroidal or higher genus horizons*, Phys. Rev. **D56** (1997), 3600–3610, gr-qc/9705012.

[8] R. Budzyński, W. Kondracki, and A. Królak, *On the differentiability of Cauchy horizons*, Jour. Math. Phys. **40** (1999), 5138–5142.

[9] L.A. Caffarelli and X. Cabré, *Fully nonlinear elliptic equations*, Colloqium Publications, vol. 43, AMS, Providence, RI, 1995.

[10] B. Carter, *Killing horizons and orthogonally transitive groups in space–time*, Jour. Math. Phys. **10** (1969), 70–81.

[11] Y. Choquet-Bruhat and J. York, *The Cauchy problem*, General Relativity (A. Held, ed.), Plenum Press, New York, 1980.

[12] P.T. Chruściel, *Uniqueness of black holes revisited*, (N. Straumann, Ph. Jetzer, and G. Lavrelashvili, eds.), vol. 69, Helv. Phys. Acta, May 1996, Proceedings of Journés Relativistes 1996, gr-qc/9610010, pp. 529–552.

[13] ______, *On rigidity of analytic black holes*, Commun. Math. Phys. **189** (1997), 1–7, gr-qc/9610011.

[14] ______, *A remark on differentiability of Cauchy horizons*, Class. Quantum Grav. (1998), 3845–3848, gr-qc/9807059.

[15] P.T. Chruściel and G.J. Galloway, *Horizons non–differentiable on dense sets*, Commun. Math. Phys. **193** (1998), 449–470, gr-qc/9611032.

[16] C.J.S. Clarke and F. de Felice, *Globally noncausal space-times*, Jour. Phys. A **15** (1982), 2415–2417.

[17] F.H. Clarke, *Optimization and nonsmooth analysis*, second ed., Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.

[18] A. Dold, *Lectures on algebraic topology*, Springer, Berlin, Heidelberg and New York, 1972.

[19] G. A. Edgar, *Measure, topology, and fractal geometry*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1990.

[20] I. Ekeland and R. Temam, *Convex analysis and variational problems*, Studies in Math. and its Appl., vol. 1, North Holland, Amsterdam, 1976.

[21] J.-H. Eschenburg, *Comparison theorems and hypersurfaces*, Manuscripta Math. **59** (1987), no. 3, 295–323.

[22] L.C. Evans and R.F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992.

[23] H. Federer, *Curvature measures*, Trans. Amer. Math. Soc. **93** (1959), 418–491.

[24] ______, *Geometric measure theory*, Springer Verlag, New York, 1969, (Die Grundlehren der mathematischen Wissenschaften, Vol. 153).

[25] ______, *Colloquium lectures on geometric measure theory*, Bull. Amer. Math. Soc. **84** (1978), no. 3, 291–338.

[26] W.H. Fleming and H. Mete Soner, *Controlled Markov processes and viscosity solutions*, Applications of Mathematics, vol. 25, Springer–Verlag, New York, Heidelberg, 1993.

[27] H. Friedrich, I. Rácz, and R.M. Wald, *On the rigidity theorem for space-times with a stationary event horizon or a compact Cauchy horizon*, Commun. Math. Phys. **204** (1999), 691–707, gr-qc/9811021.

[28] G.J. Galloway, *A "finite infinity" version of the FSW topological censorship*, Class. Quantum Grav. **13** (1996), 1471–1478.

[29] ______, *Maximum principles for null hypersurfaces and null splitting theorems*, Stockholm preprint, 1999.

[30] G.J. Galloway, K. Schleich, D.M. Witt, and E. Woolgar, *Topological censorship and higher genus black holes*, (1999), gr-qc/9902061.

[31] G.J. Galloway and E. Woolgar, *The cosmic censor forbids naked topology*, Class. Quantum Grav. **14** (1996), L1–L7, gr-qc/9609007.

[32] R. Geroch, *Domain of dependence*, Jour. Math. Phys. **11** (1970), 437–449.

[33] F. Giannoni, A. Masiello, and P. Piccione, *A variational theory for light rays in stably causal Lorentzian manifolds: regularity and multiplicity results*, Commun. Math. Phys. **187** (1997), 375–415.

[34] G.W. Gibbons and S.W. Hawking, *Cosmological event horizons, thermodynamics, and particle creation*, Phys. Rev. **D15** (1977), 2738–2751.

[35] D. Giulini, *Is there a general area theorem for black holes?*, Jour. Math. Phys. **39** (1998), 6603–6606.

[36] P.R. Halmos, *Measure Theory*, D. Van Nostrand Company, Inc., New York, N. Y., 1950.

[37] S.W. Hawking and G.F.R. Ellis, *The large scale structure of space-time*, Cambridge University Press, Cambridge, 1973.

[38] M. Heusler, *Black hole uniqueness theorems*, Cambridge University Press, Cambridge, 1996.

[39] E. Hewitt and K. Stromberg, *Real and abstract analysis*, Springer-Verlag, New York, 1975, Graduate Texts in Mathematics, No. 25.

[40] L. Hörmander, *The analysis of linear partial differential operators*, Springer, Berlin, 1985.

[41] R. Howard, *The kinematic formula in Riemannian homogeneous spaces*, Mem. Amer. Math. Soc. **106** (1993), no. 509, vi+69.

[42] ———, *The boundary structure of sets satisfying a locally uniform inner ball condition*, preprint, 1999.

[43] S.A. Hughes, C.R. Keeton, S.T. Shapiro, S.A. Teukolsky, P. Walker, and K. Walsh, *Finding black holes in numerical spacetimes*, Phys. Rev. D **49** (1994), 4004–4015.

[44] S. Husa and J. Winicour, *The asymmetric merger of black holes*, Phys. Rev. **D60** (1999), 084019 (13 pp.), gr-qc/9905039.

[45] J. Isenberg and V. Moncrief, *Symmetries of cosmological Cauchy horizons with exceptional orbits*, Jour. Math. Phys. **26** (1985), 1024–1027.

[46] F. Kottler, *Über die physikalischen Grundlagen der Einsteinschen Gravitationstheorie*, Annalen der Physik **56** (1918), 401–462.

[47] A. Królak, *Definitions of black holes without use of the boundary at infinity*, Gen. Rel. Grav. **14** (1982), 793–801.

[48] ———, *Black holes and the strong cosmic censorship*, Gen. Rel. Grav. **16** (1984), 121–130.

[49] ———, *Black holes and the weak cosmic censorship*, Gen. Rel. Grav. **16** (1984), 365–373.

[50] ———, *The existence of regular partially future asymptotically predictable space-times*, Jour. Math. Phys. **29** (1988), 1786–1788.

[51] D.N. Kupeli, *On null submanifolds in spacetimes*, Geom. Dedicata **23** (1987), 33–51.

[52] F. Morgan, *Geometric measure theory, a beginner's guide*, Academic Press, San Diego, 1995.

[53] R.P.A.C. Newman, *Cosmic censorship and curvature growth*, Gen. Rel. Grav. **15** (1983), 641–353.

[54] ———, *The global structure of simple space-times*, Commun. Math. Phys. **123** (1989), 17–52.

[55] B. O'Neill, *Semi–Riemannian geometry*, Academic Press, New York, 1983.

[56] R. Penrose, *Techniques of differential topology in relativity*, SIAM, Philadelphia, 1972, (Regional Conf. Series in Appl. Math., vol. 7).

[57] V. Perlick, *On Fermat's principle in general relativity: I. The general case*, Class. Quantum Grav. **7** (1990), 1319–1331.

[58] A. Rendall, *Local and global existence theorems for the Einstein equations*, Living Reviews in Relativity **1** (1998), URL http://www.livingreviews.org.

[59] H.J. Seifert, *Smoothing and extending cosmic time functions*, Gen. Rel. Grav. **8** (77), 815–831.

[60] S.T. Shapiro, S.A. Teukolsky, and J. Winicour, *Toroidal black holes and topological censorship*, Phys. Rev. D **52** (1995), 6982–6987.

[61] L. Simon, *Lectures on geometric measure theory*, Proceedings of the CMA, vol. 3, Australian University Press, Canberra, 1983.

[62] E.M. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press, Princeton, N.J., 1970, Princeton Mathematical Series, No. 30.

[63] C.V. Vishveshwara, *Generalization of the "Schwarzschild surface" to arbitrary static and stationary metrics*, Jour. Math. Phys. **9** (1968), 1319–1322.

[64] R.M. Wald, *General relativity*, University of Chicago Press, Chicago, 1984.