DISSERTATION | DOCTORAL THESIS

Titel | Title

Singularities and rigidity in smooth and non-smooth spacetimes

verfasst von | submitted by

Argam Ohanyan BSc BSc MSc MSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 796 605 405
Dissertationsgebiet lt. Studienblatt | Field of study as it appears on the student record sheet:	Mathematik
Betreut von | Supervisor:	ao. Univ.-Prof. Mag. Dr. Roland Steinbauer

Abstract

This thesis deals with questions of geometric rigidity and the existence of singularities in spacetimes. First, we extend the Hawking–Penrose singularity theorem to spacetime metrics of regularity C^1. In the cosmological case, Bartnik's famous splitting conjecture asserts the rigidity of the Hawking–Penrose theorem, and has been shown to be equivalent to the existence of CMC Cauchy surfaces. We extend a construction of Bartnik to obtain a large class of cosmological spacetimes which, although timelike geodesically incomplete, contain no CMC Cauchy surfaces. Lastly, we present a generalization of the Lorentzian splitting theorem to the synthetic context of Lorentzian length spaces.

Zusammenfassung

Diese Dissertation befasst sich mit Fragen zur geometrischer Rigidität und der Existenz von Singularitäten in Raumzeiten. Wir beginnen mit einer Erweiterung des Hawking–Penrose Singularitätentheorems auf Raumzeitmetriken der Regularität C^1. Im kosmologischen Fall behauptet die bekannte Spaltungsvermutung von Bartnik die Rigidität des Hawking–Penrose Theorems, und diese Vermutung ist bekanntlich äquivalent zur Existenz von CMC Cauchyflächen. Wir erweitern eine Konstruktion von Bartnik, um eine große Klasse an kosmologischen Raumzeiten zu erhalten die, obwohl zeitartig geodätisch unvollständig, keine CMC Cauchyflächen enthalten. Zum Schluss präsentieren wir eine Verallgemeinerung des Lorentzschen Spaltungstheorems auf den synthetischen Kontext von Lorentzlängenräumen.

Acknowledgements

Let me start by thanking my PhD supervisor Roland Steinbauer and my de facto second PhD supervisor Michael Kunzinger. The atmosphere that they have cultivated in the DIANA research group, both from a professional and from a personal perspective, is nothing short of phenomenal. I cannot express how grateful I am for having met and worked with such great people, hopefully we will work together for many more years to come. By extension, I would like to thank all of the current and former members of the DIANA research group (most of whom I have already collaborated with) for contributing to the positive atmosphere, in particular Tobias Beran, Matteo Calisti, Melanie Graf, Felix Rott and Clemens Sämann.

This thesis would have certainly not come together in this way without all of my collaborators, to whom I am deeply grateful and from whom I have learned a lot, whether from talks, chats at conferences, or discussions online. I would like to direct particular thanks to three of them: Mathias Braun, Robert McCann and Nicola Gigli. Ever since I had the pleasure to meet them at the "Thematic Program" on Riemannian and Lorentzian geometry at the Fields Institute back in 2022, I have been able to learn so much from them, both related and unrelated to mathematics.

On the personal side, I would like to thank my wife, Emma, and my mother, Anush, who support me in all of my endeavors and without whose solid backing I could not dedicate so much of my time and effort to doing research.

Introduction

This thesis deals with some fundamental questions in spacetime geometry, the geometric setting Einstein's theory of General Relativity (GR). Classically, GR has been studied in the standard differential-geometric setting of smooth manifolds equipped with smooth Lorentzian metric tensors. This approach, however, does not satisfactorily describe important physical situations such as shock waves, conical singularities or impulsive gravitational waves [36, 41, 21, 38, 37], which are inherently non-smooth. There are different main ways to pursue studies of non-smoothness in spacetime geometry: One could consider spacetimes where the metric g is non-smooth (usually less than C^2, since for most global geometric questions in GR the metric does not need to be differentiated more than twice), or one could do away with manifolds and tensors entirely and pursue a synthetic approach analogous to metric (or length) spaces as synthetic variants of Riemannian manifolds.

Spacetimes with non-smooth metrics have been studied in quite great detail over the last decade, in particular with focus on singularity theorems [29, 30, 31, 32, 20, 19, 27]. In these settings, curvature is generally only distributional, so the energy conditions necessary to prove singularity theorems have to be stated as distributional inequalities (which, incidentally, imply automatic measure-theoretic regularity for these tensors). These distributional energy conditions then imply surrogate conditions for smooth approximating metrics, for which classical Lorentz-geometric arguments can be applied. In this context, one essentially requires approximations which respect the causality [13].

The synthetic approach to Lorentzian geometry was initiated in [28] in the form of Lorentzian length spaces, building on the earlier theory of causal sets [26]. Lorentzian length spaces are meant to play to role that metric length spaces play in positive definite signature [10]. In [28], the authors put a particular emphasis on curvature bounds, which are formulated via timelike (or causal) triangle comparison, in the same spirit as curvature in Alexandrov spaces [10]. In the smooth setting, such notions of curvature bounds correspond to bounds on the timelike sectional curvature, pioneered in [22] in the setting of spacetimes. A general theory of sectional curvature bounds on semi-Riemannian manifolds is available [1], but it has to take into account spacelike sectional curvatures as well, which is unnatural from the point of view of GR. Surprisingly, the full equivalence of Lorentzian curvature bounds via triangle comparison (or other synthetic methods) and timelike sectional curvature bounds was only recently established [5].

More recently, synthetic notions of timelike Ricci curvature bounds (such as the strong energy condition) have been studied on synthetic Lorentzian spaces. Such bounds are formulated via convexity properties of entropy functionals defined on certain spaces of probability measures. In these spaces, geodesics are obtained by way of solving an optimal transport problem whose cost is a power of the time separation function (see e.g. [11, 7, 8, 9]).

The aim of this thesis is to study singularities and rigidity in spacetimes, mostly in the non-smooth setting. Let us now go into more detail on each of the papers included while highlighting their interrelationship.

The first paper we include [27] gives a generalization of the Hawking–Penrose singularity theorem to spacetime metric which are C^1-regular. Recall that singularity theorems always have the following general form [39]: If (M, g) is a spacetime satisfying an energy condition, a causality condition, and an initial or boundary condition, then there exist incomplete causal geodesics in M. The most sophisticated of the classical singularity theorems is the Hawking–Penrose singularity theorem [25], which makes only weak assumptions on the causality and allows for a wide array of initial and boundary conditions. The key difference to other singularity theorems such as the one due to Hawking describing a Big Bang scenario [23] or the one due to Penrose describing gravitational collapse [35] is that there is no assumption of initial geodesic convergence by way of mean curvature or similar quantities (see [24] for more on singularity theorems). Rather, one imposes a global gravitational nontriviality condition called *genericity*, which states that the tidal force operator $R(\cdot, \gamma')\gamma'$ is nontrivial at some parameter along every inextendible causal geodesic γ. Our main innovation in [27] was to find a suitable distributional formulation of this condition for metrics $g \in C^1$ in order to obtain the necessary approximate conditions on g_ε, where g_ε is a suitable approximation of g by smooth spacetime metrics. Since $g \in C^1$, the geodesic equation has C^0 right hand side, thus by Peano's theorem, it is solvable for any initial data, but no longer uniquely so. This leads to complications when trying to approximate g-geodesics by g_ε-ones. Thus, in addition to (suitable reformulations of) the usual assumptions of the Hawking–Penrose theorem, we had to assume that maximizing causal geodesics do not exhibit branching. The latter can be interpreted as an observer or light ray “splitting into two”, which is physically equally as catastrophic as a incompleteness.

Let us turn to *cosmological spacetimes*, i.e. globally hyperbolic spacetimes with compact Cauchy surfaces satisfying the strong energy condition. Such spacetimes naturally satisfy the technical assumptions of the Hawking–Penrose singularity theorem, so under the assumption of genericity, they contain incomplete causal geodesics. This conclusion is clearly wrong without the assumption of genericity: Indeed, let (S, h) be any compact Riemannian manifold with nonnegative Ricci curvature, then the product $(\mathbb{R} \times S, -dt^2 + h)$ is a geodesically complete cosmological spacetime. Bartnik [2] conjectured that these are the only non-generic complete examples, i.e. he conjectured the rigidity of the cosmological Hawking–Penrose theorem in its assumption of genericity. Bartnik's cosmological splitting conjecture remains one of the most fundamental open problems in mathematical GR and spacetime geometry to date.

Bartnik observed that the conclusion of his conjecture is equivalent to the existence of CMC (constant mean curvature) Cauchy surfaces. Moreover, he showed via a spacetime gluing construction that there exist cosmological spacetimes which, albeit incomplete, have no CMC Cauchy surfaces. The rough procedure is as follows: Take one half of the maximally extended Schwarzschild spacetime and glue it to an FLRW spacetime which is spatially T^3, then extend it to the Schwarzschild event horizon and attach there a time inverted copy of the resulting spacetime. A similar vacuum example is obtained via initial data gluing in [14]. In the second paper included in this thesis [33], we expand the known class of such "no-CMC" cosmological spacetimes by generalizing Bartnik's spacetime gluing construction. It requires an investigation of Tolman–Bondi metrics and implicit solutions of certain ODEs (which are explicit in Bartnik's case), allowing for more general spatial topologies on FLRW, as well as arbitrary higher spacetime dimensions. The hope is that our work will contribute to a better understanding of the relationship between CMC Cauchy surfaces and (in)completeness in the context of Bartnik's conjecture.

The final paper included deals with a more classical rigidity problem, namely the splitting of a space of nonnegative curvature given the existence of a global distance-realizing curve (i.e. a *line*). Splitting theorems are available in various parts of geometry, such as for Riemannian manifolds with nonnegative sectional [40] and Ricci [12] curvature or $\mathrm{RCD}(0, N)$-spaces (i.e. metric measure spaces which have synthetic Ricci curvature bounded below by 0 and dimension above by N) with $N < \infty$ [17, 18]. In the context of Lorentzian geometry, the splitting theorem for globally hyperbolic spacetimes with nonpositive timelike sectional curvature containing a complete timelike line was obtained in [3, 4], and various versions under the strong energy condition in [15, 16, 34]. Our work [6] generalizes [3, 4] to the setting of Lorentzian length spaces. The fundamental difficulty we had to overcome was the absence of a notion of spacelike distance on hypersurfaces. To deal with this issue, we introduced the notion of parallelity for timelike lines: Two timelike lines are parallel if they can be mapped via a bijective and time separation preserving map onto two vertical lines in 2-dimensional Minkowski space. It turns out that their Euclidean cross-distance is a well-defined distance between these lines. Fixing a parametrization, one can thus endow a spatial cross section of lines constructed as asymptotes to the original one with the structure of a metric space, the splitting then follows suit. The additional techniques we use are inspired by the angle and triangle comparison methods already employed in [3, 4].

List of included papers

1. M. Kunzinger, A. Ohanyan, B. Schinnerl, R. Steinbauer.
   The Hawking–Penrose singularity theorem for C^1-Lorentzian metrics.
   Comm. Math. Phys. 391 (2022), no.3, 1143–1179.
   DOI: https://link.springer.com/article/10.1007/s00220-022-04335-8

2. E. Ling, A. Ohanyan.
   Examples of cosmological spacetimes without CMC Cauchy surfaces.
   Lett. Math. Phys. 114 (2024), no.4, Paper No. 96.
   DOI: https://link.springer.com/article/10.1007/s11005-024-01843-7

3. T. Beran, A. Ohanyan, F. Rott, D. A. Solis.
   The splitting theorem for globally hyperbolic Lorentzian length spaces with non-negative timelike curvature.
   Lett. Math. Phys. 113 (2023), no.2, Paper No. 48, 47 pp.
   DOI: https://link.springer.com/article/10.1007/s11005-023-01668-w

References

[1] S. B. Alexander and R. L. Bishop. Lorentz and semi-Riemannian spaces with Alexandrov curvature bounds. *Comm. Anal. Geom.*, 16(2):251–282, 2008.

[2] R. Bartnik. Remarks on cosmological spacetimes and constant mean curvature surfaces. *Communications in mathematical physics*, 117:615–624, 1988.

[3] J. K. Beem, P. E. Ehrlich, S. Markvorsen, and G. J. Galloway. Decomposition theorems for Lorentzian manifolds with nonpositive curvature. *J. Diff. Geom.*, 22(1), 1985.

[4] J. K. Beem, P. E. Ehrlich, S. Markvorsen, and G. J. Galloway. A Toponogov splitting theorem for Lorentzian manifolds. In *Lecture Notes in Mathematics: Global Differential Geometry and Global Analysis 1984*, volume 1156, pages 1–13. Springer, Berlin, Heidelberg, 1985.

[5] T. Beran, M. Kunzinger, A. Ohanyan, and F. Rott. The equivalence of smooth and synthetic notions of timelike sectional curvature bounds. *arXiv preprint arXiv:2403.18077*, 2024.

[6] T. Beran, A. Ohanyan, F. Rott, and D. A. Solis. The splitting theorem for globally hyperbolic Lorentzian length spaces with non-negative timelike curvature. *Lett. Math. Phys.*, 113(2):Paper No. 48, 47, 2023.

[7] M. Braun. Rényi's entropy on Lorentzian spaces. Timelike curvature-dimension conditions. *J. Math. Pures Appl. (9)*, 177:46–128, 2023.

[8] M. Braun and R. J. McCann. Causal convergence conditions through variable timelike Ricci curvature bounds. *arXiv preprint arXiv:2312.17158*, 2023.

[9] M. Braun and S.-i. Ohta. Optimal transport and timelike lower Ricci curvature bounds on Finsler spacetimes. *Trans. Amer. Math. Soc.*, 377(5):3529–3576, 2024.

[10] D. Burago, Y. Burago, and S. Ivanov. *A Course in Metric Geometry, Graduate Studies in Mathematics*, volume 33. American Mathematical Society, 2001.

[11] F. Cavalletti and A. Mondino. Optimal transport in Lorentzian synthetic spaces, synthetic timelike Ricci curvature lower bounds and applications, 2020.

[12] J. Cheeger and D. Gromoll. The splitting theorem for manifolds of non-negative Ricci curvature. *J. Diff. Geom.*, 6(1):119–128, 1971.

[13] P. T. Chruściel and J. D. Grant. On Lorentzian causality with continuous metrics. *Classical and Quantum Gravity*, 29(14):145001, 2012.

[14] P. T. Chruściel, J. Isenberg, and D. Pollack. Initial data engineering. *Communications in mathematical physics*, 257(1):29–42, 2005.

[15] J.-H. Eschenburg. The splitting theorem for space-times with strong energy condition. *Journal of differential geometry*, 27(3):477–491, 1988.

[16] G. J. Galloway. The lorentzian splitting theorem without the completeness assumption. *Journal of differential geometry*, 29(2):373–387, 1989.

[17] N. Gigli. The splitting theorem in non-smooth context, 2013. https://arxiv.org/abs/1302.5555.

[18] N. Gigli. An overview of the proof of the splitting theorem in spaces with non-negative Ricci curvature. *Anal. Geom. Metr. Spaces*, 2(1):169–213, 2014.

[19] M. Graf. Singularity theorems for C^1-Lorentzian metrics. *Communications in Mathematical Physics*, 378(2):1417–1450, 2020.

[20] M. Graf, J. D. E. Grant, M. Kunzinger, and R. Steinbauer. The Hawking–Penrose Singularity Theorem for $C^{1,1}$-Lorentzian Metrics. *Comm. Math. Phys.*, 360(3):1009–1042, 2018.

[21] J. B. Griffiths and J. Podolský. *Exact space-times in Einstein's general relativity.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 2009.

[22] S. G. Harris. A triangle comparison theorem for Lorentz manifolds. *Indiana Univ. Math. J.*, 31(3):289–308, 1982.

[23] S. W. Hawking. The occurrence of singularities in cosmology. iii. causality and singularities. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 300(1461):187–201, 1967.

[24] S. W. Hawking and G. F. R. Ellis. *The large scale structure of space-time*, volume No. 1 of *Cambridge Monographs on Mathematical Physics.* Cambridge University Press, London-New York, 1973.

[25] S. W. Hawking and R. Penrose. The singularities of gravitational collapse and cosmology. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 314(1519):529–548, 1970.

[26] E. H. Kronheimer and R. Penrose. On the structure of causal spaces. *Proc. Cambridge Philos. Soc.*, 67(2):481–501, 1967.

[27] M. Kunzinger, A. Ohanyan, B. Schinnerl, and R. Steinbauer. The Hawking-Penrose singularity theorem for C^1-Lorentzian metrics. *Comm. Math. Phys.*, 391(3):1143–1179, 2022.

[28] M. Kunzinger and C. Sämann. Lorentzian length spaces. *Ann. Glob. Anal. Geom.*, 54(3):399–447, 2018. https://doi.org/10.1007/s10455-018-9633-1.

[29] M. Kunzinger, R. Steinbauer, and M. Stojković. The exponential map of a $C^{1,1}$-metric. *Differential Geom. Appl.*, 34:14–24, 2014.

[30] M. Kunzinger, R. Steinbauer, M. Stojković, and J. A. Vickers. A regularisation approach to causality theory for $C^{1,1}$-Lorentzian metrics. *Gen. Relativity Gravitation*, 46(8):Art. 1738, 18, 2014.

[31] M. Kunzinger, R. Steinbauer, M. Stojković, and J. A. Vickers. Hawking's singularity theorem for $C^{1,1}$-metrics. *Classical Quantum Gravity*, 32(7):075012, 19, 2015.

[32] M. Kunzinger, R. Steinbauer, and J. A. Vickers. The Penrose singularity theorem in regularity $C^{1,1}$. *Classical Quantum Gravity*, 32(15):155010, 12, 2015.

[33] E. Ling and A. Ohanyan. Examples of cosmological spacetimes without CMC Cauchy surfaces. *Lett. Math. Phys.*, 114(4):Paper No. 96, 2024.

[34] R. P. A. C. Newman. A proof of the splitting conjecture of S.-T. Yau. *J. Differential Geom.*, 31(1):163–184, 1990.

[35] R. Penrose. Gravitational collapse and space-time singularities. *Physical Review Letters*, 14(3):57, 1965.

[36] R. Penrose. The geometry of impulsive gravitational waves. In *General relativity (papers in honour of J. L. Synge)*, pages 101–115. Clarendon Press, Oxford, 1972.

[37] J. Podolský, C. Sämann, R. Steinbauer, and R. Švarc. The global uniqueness and C^1-regularity of geodesics in expanding impulsive gravitational waves. *Classical Quantum Gravity*, 33(19):23, 2016. https://arxiv.org/abs/1602.05020.

[38] J. Podolský, C. Sämann, R. Steinbauer, and R. Švarc. The global existence, uniqueness and C^1-regularity of geodesics in nonexpanding impulsive gravitational waves. *Classical Quantum Gravity*, 32(2):025003, 23, 2015.

[39] J. M. Senovilla. Singularity theorems and their consequences. *General Relativity and Gravitation*, 30(5):701–848, 1998.

[40] V. Toponogov. The metric structure of Riemannian spaces with nonnegative curvature which contain straight lines. *Amer. Math. Soc. Transl.*, 70:225–239, 1968.

[41] J. A. Vickers and J. P. Wilson. Generalized hyperbolicity in conical spacetimes. *Classical Quantum Gravity*, 17(6):1333–1360, 2000.

Commun. Math. Phys. 391, 1143–1179 (2022)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-022-04335-8

Communications in
**Mathematical
Physics**

The Hawking–Penrose Singularity Theorem for C^1-Lorentzian Metrics

Michael Kunzinger, Argam Ohanyan, Benedict Schinnerl, Roland Steinbauer

Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria.
E-mail: michael.kunzinger@univie.ac.at; argam.ohanyan@univie.ac.at; benedict.schinnerl@univie.ac.at; roland.steinbauer@univie.ac.at

Received: 18 October 2021 / Accepted: 18 January 2022
Published online: 26 February 2022

Abstract: We extend both the Hawking–Penrose theorem and its generalisation due to Galloway and Senovilla to Lorentzian metrics of regularity C^1. For metrics of such low regularity, two main obstacles have to be addressed. On the one hand, the Ricci tensor now is distributional, and on the other hand, unique solvability of the geodesic equation is lost. To deal with the first issue in a consistent way, we develop a theory of tensor distributions of finite order, which also provides a framework for the recent proofs of the theorems of Hawking and of Penrose for C^1-metrics (Graf in Commun Math Phys 378(2):1417–1450, 2020). For the second issue, we study geodesic branching and add a further alternative to causal geodesic incompleteness to the theorem, namely a condition of maximal causal non-branching. The genericity condition is re-cast in a distributional form that applies to the current reduced regularity while still being fully compatible with the smooth and $C^{1,1}$-settings. In addition, we develop refinements of the comparison techniques used in the proof of the $C^{1,1}$-version of the theorem (Graf in Commun Math Phys 360:1009–1042, 2018). The necessary results from low regularity causality theory are collected in an appendix.

1. Introduction

The classical singularity theorems[1] of General Relativity (GR) collect sufficient and physically reasonable conditions that lead to causal geodesic incompleteness of spacetime, hence to the occurrence of a singularity in the spirit of Roger Penrose's Nobel Prize-winning approach [40]. Being rigorous statements in pure Lorentzian geometry, they were first formulated within the smooth category. As they form a body of essential results in GR, the quest for low regularity versions of the theorems is eminent and was already explicitly discussed in [14, Sec. 8.4]. In fact, the (smooth) theorems can be read as predicting a mere drop of the differentiability of the metric below C^2, rather than incompleteness.

[1] See [14, Ch. 8], [21, Ch. 9], [46–48] for extensive treatments.

However, only with the rather recent advent of systematic studies in low regularity GR [6] and, in particular, causality theory [5,24,31,41], a rigorous extension of the singularity theorems for metrics below the C^2-class became feasible, cf. [46, Sec. 6.2]. This applies first of all to metrics of regularity $C^{1,1}$, which model situations where the matter variables possess finite jumps. Indeed the Hawking, the Penrose, and, finally, the Hawking-Penrose theorem were proven in this regularity [8,25,26]. These results rule out that a smooth spacetime that is incomplete by the classical theorems, can be extended to a complete $C^{1,1}$-spacetime. Also, they imply that the spacetime either is incomplete or the differentiability of the metric is below $C^{1,1}$, and consequently the curvature becomes unbounded.

Technically, in $C^{1,1}$ the exponential map is still available [23,31] and the curvature is still locally bounded, which allows for a natural extension of the energy and the genericity conditions. However, the curvature tensor is only defined almost everywhere, which forbids the use of Jacobi fields and conjugate points, both essential tools in the classical proofs. Instead one uses regularisation techniques which allow one to derive weakened versions of the energy conditions for approximating smooth metrics with controlled causality. Using careful comparison techniques for the (matrix) Riccati equation, it is then possible to show that these (still) lead to the occurrence of conjugate or focal points along causal geodesics for the approximating metrics. This in turn forces the geodesics of the $C^{1,1}$-metric to stop maximising. Similarly, a natural extension of the initial and the energy conditions leads to the formation of a trapped set. This, together with an extension of the causal parts of the classical proofs, allows one to finally derive the results.

The next step in lowering the regularity assumptions was undertaken by Graf [7], who extended the Hawking and the Penrose theorems to C^1-metrics, with the Gannon-Lee theorem following in [43]. C^1-regularity is at the moment the lowest possible class where classical singularity theorems have been established, and in this paper we complete this effort by proving the most refined of these statements, namely the Hawking-Penrose theorem.

In this regularity class one faces the following added severe complications arising from the fact that the Levi-Civita connection is only continuous:

(a) The curvature is no longer locally bounded, but is merely a distribution *of order one*.
(b) The initial value problem for the geodesic equation is solvable, but *not uniquely* so.
(c) The exponential map is no longer defined.

The first item is especially relevant when formulating the energy conditions. While the strong energy condition can be extended in a straightforward manner, a formulation of the null energy condition is more subtle, cf. [7, Sec. 5], see also Sect. 2.3, below. The genericity condition needed for the Hawking-Penrose theorem is still more delicate. Actually it turns out to be necessary to apply the Ricci tensor to vector fields constructed via parallel transport, which means that they are merely of C^1-regularity, see Definition 2.15. This can only be done since the curvature is a distribution of order one, and we present the required distributional setting in Sect. 2.1, below. Whereas in the $C^{1,1}$-context the use of distributional methods could be avoided, it becomes essential in the present work, and at the same time has the benefit of providing a global framework also for the earlier results by Graf [7] on which we build.

Then, while we can use the weakened versions of the strong and null energy conditions for approximating smooth metrics derived in [7] (employing a refined version of the Friedrichs Lemma), we have to derive from the distributional genericity condition an appropriate weakened version for smooth approximations that allows us to employ an extension of the Riccati comparison techniques developed in [8], see Lemmas 2.17, and 2.18 . The key technique of regularisation by smooth metrics with adapted causality as put forward in [5] is recalled in Sect. 2.2.

The first issue connected to item (b) is that it necessitates a decision on how to extend the notion of geodesic completeness, and following [7] we say that a spacetime is complete if *every* inextendible solution to the geodesic equation is complete (rather than merely demanding only one complete solution for any choice of initial data).

Next, item (b) together with item (c) forbids the use of an essential argument in the context of approximating maximising causal geodesics by maximising causal geodesics of the approximating metrics. This issue has been solved in [7, Sec. 2] in the *globally hyperbolic* case. In the context of the present work, however, we have to establish more general results, see Sect. 3. In fact, we have to introduce an additional assumption, namely a non-branching condition for maximising causal geodesics (Definition 3.1), to derive the corresponding approximation results in Proposition 3.4. Here it is also essential to apply a classical ODE result (Proposition 2.5), which we reformulate for our purpose in Corollary 2.6. Moreover, since the interrelation between causal geodesics (i.e., solutions of the geodesic equation) and maximising curves becomes more subtle below $C^{1,1}$ [16,42], we rely in particular on the fact that maximising causal curves are indeed (unbroken) geodesics and hence C^2-curves [27,43].

The new non-branching assumption is well motivated by similar conditions used in metric geometry, where due to the lack of a differentiable structure the geodesic equation is not available. Also a null version of the condition was essential in the proof of the C^1-Gannon-Lee theorem, see [43, Sec. 3]. Moreover it adds a novel facet to the interpretation of the C^1-version of the theorem: It predicts either geodesic incompleteness or branching of maximising causal geodesics, with both alternatives physically signifying a catastrophic event for the observer corresponding to the geodesic. Of course, there is the alternative that the regularity of the metric drops below C^1, which renders the curvature a distribution of higher order. Again, the result also forbids the extension of the spacetime to a complete one of regularity C^1 without (maximal causal) geodesic branching. This once more complements the recent C^0-inextendibility results of [9,45].

Next we briefly discuss the further prospects of low-regularity singularity theorems. Concerning the causality part of the results it is expected that they extend to $C^{0,1}$-metrics[2], while some features of causality theory fail below this class [5,12]. Nevertheless, the causal core of some of the singularity theorems has been established in the very general setting of closed cone structures, see [35, Sec. 2.15].

On the analytic side, already in $C^{0,1}$ one faces the problem that the right hand side of the geodesic equation is merely locally bounded and one would have to resort to non-classical solution concepts. Also another analytical technique at the core of the arguments, i.e. the Friedrichs Lemma which is used to go from the distributional energy conditions of the singular metric to useful surrogate conditions for the smooth approximations, appears to be sharp, see [7, Lem. 4.8] for the currently most advanced version. This is actually a long way from the largest possible class where the curvature can be (stably) defined in an analytical (distributional) way, i.e. $H^1 \cap L^\infty$ locally [11,28,50]. However, the quest is there to at least go into the direction of regularity classes more closely linked to the classical existence results for the field equations, H^s locally for $s > 5/2$, or to current formulations of cosmic censorship, i.e. the connection lying locally in L^2.

[2] A class for which Hawking and Ellis [14, p. 268, 287] still speculate the singularity theorems to hold.

Very recently, also synthetic formulations of the singularity theorems have appeared: In [1], Lorentzian length spaces that are warped products are studied and a Hawking singularity theorem is derived from suitable sectional curvature bounds (implying Ricci curvature bounds in such geometries), based on triangle comparison. Moreover, Cavalletti and Mondino prove a version of Hawking's singularity theorem in Lorentzian metric measure spaces, where Ricci curvature bounds are implemented using methods from optimal transport [4]. However, it remains unclear to date how these results precisely relate to the analytical approach to low regularity pursued in this work.

The Hawking-Penrose theorem classically comes in a causal version [17, p. 538, Thm.], which asserts the incompatibility of the following three conditions for a C^2-spacetime (M, g): chronology, the fact that every inextendible causal geodesic stops maximising, and the existence of a (future or past) trapped set. The analytic main result is then a corollary [17, Sec. 3, Cor.] which collects sufficient conditions (energy and genericity conditions, and initial conditions) to derive the above incompatible items under the assumption of causal geodesic completeness. We will generally follow this path, but prove a more general version of the theorem due to Galloway and Senovilla [10], which adds the existence of a trapped submanifold of arbitrary codimension to the list of initial conditions.

Our version of the causal result is Theorem 6.2 and, for the sake of completeness and for the convenience of the reader we collect all results from C^1-causality theory[3] needed here (and elsewhere) in Appendix A. The main result is then Theorem 6.3. The results needed to infer from its analytic conditions (the distributional energy and genericity conditions) and the non-branching assumptions the inexistence of lines are proven in Sect. 4. The existence of a trapped set is derived from the various sets of initial conditions (and the energy conditions) in Sect. 5. Finally, some technical results on extending vector fields in C^1-spacetimes are collected in Appendix B. They allow us to put to use the distributional genericity condition in the context of regularisations.

To conclude this introduction, we fix the notations and conventions used throughout the paper. By a manifold M we will mean a smooth, connected, second countable Hausdorff manifold of dimension n (with $n \geq 3$). When such an M is endowed with a Lorentzian metric g of regularity C^1 of signature $(-, + \dots, +)$ and is time-oriented via a smooth timelike vector field we call it a C^1-spacetime (hence lowering the regularity will always apply to the spacetime metric only, not to the underlying manifold). The Levi-Civita connection of a smooth or C^1-spacetime will be denoted by ∇ (for its distributional definition in the C^1-case, see the next section). We will always assume that M is also endowed with a smooth (complete, background) Riemannian metric h with associated Riemannian distance function d_h and norm $\| \, \|_h$ which we use to estimate tensor fields. Since all such estimates will be on compact sets only, they are independent of the choice of h.

A curve $\gamma : I \to M$ is called timelike (causal, null, future or past directed) if it is locally Lipschitz and $\dot\gamma(t)$ is timelike (causal, null, future or past directed) almost everywhere. Concerning causality theory we use standard notation, cf. e.g., [34,39]. Thus $p \ll q$ (resp. $p \leq q$) means that there exists a future directed timelike (resp. causal) curve from p to q, $I^+(A) := \{q \in M : p \ll q \text{ for some } p \in A\}$ and $J^+(A) := \{q \in M : p \leq q \text{ for some } p \in A\}$. These definitions do not change if one defines the causality relations via piecewise smooth curves [24,31]. The Lorentzian distance (or time-separation function) associated to a Lorentzian metric g will be denoted by d_g. A C^1-spacetime (M, g) is globally hyperbolic if it is causal (i.e., contains no closed

[3] It should be mentioned that these results all follow from the more general approaches of [5,35,41].

causal curves) and $J(p,q) := J^+(p) \cap J^-(q)$ is compact for all $p, q \in M$. A Cauchy hypersurface in (M, g) is a closed acausal set that is met by each inextendible causal curve.

For the Riemann curvature tensor we use the convention $R(X,Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z$ and the Ricci tensor is given by $\mathrm{Ric}(X,Y) = \sum_{i=1}^n \langle E_i, E_i\rangle\langle R(E_i, X)Y, E_i\rangle$ (again we refer to the next section for a distributional interpretation in the case of C^1-metrics). Here $(E_i)_{i=1}^n$ is a (local) orthonormal frame field, while by $(e_i)_{i=1}^n$ we will denote frames in individual tangent spaces T_pM. For an embedded submanifold N of M of codimension m we define the second fundamental form by $\mathrm{II}(V, W) = \mathrm{nor}(\nabla_V W)$ and the shape operator derived from a normal unit field ν by $S(X) = \nabla_X \nu$. Using (2.13) below it follows that for $g \in C^1$ and $V, W \in C^1$, $\mathrm{II}(V, W)$ is continuous.

For g smooth (resp. C^1), $1 < m < n$ and S a smooth (resp. C^2) spacelike $(n-m)$-dimensional submanifold let $e_1(q), \ldots, e_{n-m}(q)$ be an orthonormal basis for T_qS, varying smoothly (resp. C^1) with q in a neighborhood (in S) of $p \in S$. Then $H_S := \frac{1}{n-m}\sum_{i=1}^{n-m} \mathrm{II}(e_i, e_i)$ denotes the mean curvature vector field of S, and $\mathbf{k}_S(v) := g(H_S, v)$ the convergence of $v \in TM|_S$. For g in C^1, both H_S and $\mathbf{k}_S$ are continuous. A closed spacelike submanifold S is called *(future) trapped* if for any future-directed null vector $\nu \in TS^\perp$ the convergence $\mathbf{k}_S(\nu)$ is positive, or equivalently that the mean curvature vector field H_S is past pointing timelike on S.

One final convention we shall require is that of tensor classes (cf. [21, Sec. 4.6.3]): Given a causal geodesic γ in a C^1-spacetime (M, g), we set $[\dot\gamma(t)]^\perp := (\dot\gamma(t))^\perp/\mathbb{R}\dot\gamma(t)$. For γ null, $[\dot\gamma(t)]^\perp$ is an $(n-2)$-dimensional subspace of the hypersurface $(\dot\gamma(t))^\perp$. On the other hand, for γ timelike, $[\dot\gamma(t)]^\perp$ equals $(\dot\gamma(t))^\perp$. To obtain a unified notation, throughout the paper we will denote the dimension of $[\dot\gamma(t)]^\perp$ by d, which amounts to setting $d = n-2$ in the null case and $d = n-1$ in the timelike case. Furthermore, we set $[\dot\gamma]^\perp = \bigcup_t [\dot\gamma(t)]^\perp$. Then for every normal tensor field A along γ we obtain a well-defined tensor class $[A]$ along γ, as well as a well-defined tensor class representing the induced covariant derivative $\nabla_{\dot\gamma}$. We then set $[\dot A] = [\nabla_{\dot\gamma} A]$. The metric $g|_{[\dot\gamma]^\perp}$ is positive definite in both the null and the timelike case. For smooth metrics, the curvature (or tidal force) operator is given by $[R](t) : [\dot\gamma(t)]^\perp \to [\dot\gamma(t)]^\perp$, $[v] \mapsto [R(v, \dot\gamma(t))\dot\gamma(t)]$.

2. Distributional Curvature of C^1-Spacetimes

2.1. Curvature tensors of C^1-metrics. In what follows we discuss the general distributional framework in which to understand the curvature quantities associated to metric tensors (of arbitrary signature) of regularity below C^2, as is required for our further analysis. The general mathematical framework goes back to [29], while [11] provided the first study specific to GR. We mainly follow [28], cf. also [13,49], but put a special emphasis on metrics of regularity C^1 and distributions of finite order.

As in [13, Sec. 3.1], for $k \in \mathbb{N}_0 \cup \{\infty\}$ we denote by $\mathrm{Vol}(M)$ the volume bundle over M, and by $\Gamma_c^k(M, \mathrm{Vol}(M))$ the space of compactly supported one-densities on M (i.e., sections of $\mathrm{Vol}(M)$) that are k-times continuously differentiable. Then the space of distributions of order k on M is the topological dual of $\Gamma_c^k(M, \mathrm{Vol}(M))$,

$$\mathcal{D}'^{(k)}(M) := \Gamma_c^k(M, \mathrm{Vol}(M))'.$$

For $k = \infty$ we omit the superscript (k). Clearly there are topological embeddings $\mathcal{D}'^{(k)}(M) \hookrightarrow \mathcal{D}'^{(k+1)}(M) \hookrightarrow \mathcal{D}'(M)$ for all k. As we shall see, all curvature quantities of C^1-metrics are distributional tensor fields of order 1. The space of distributional (r, s)-tensor fields of order k is defined as

$$\mathcal{D}'^{(k)}\mathcal{T}^r_s(M) \equiv \mathcal{D}'^{(k)}(M, T^r_s M) := \Gamma^k_c(M, \mathcal{T}^s_r(M) \otimes \mathrm{Vol}(M))'. \tag{2.1}$$

By [13, 3.1.15],

$$\mathcal{D}'\mathcal{T}^r_s(M) \cong \mathcal{D}'(M) \otimes_{C^\infty(M)} \mathcal{T}^r_s(M) \cong L_{C^\infty(M)}(\Omega^1(M)^r \times \mathfrak{X}(M)^s; \mathcal{D}'(M)). \tag{2.2}$$

This algebraic isomorphism ultimately rests on the fact that the $C^\infty(M)$-module of C^∞-sections $\Gamma(M, F)$ in any smooth vector bundle $F \to M$ is finitely generated and projective, which allows one to apply [3, Ch. II §4.2, Prop. 2]. These properties persist if we consider $\Gamma_{C^k}(M, F)$ as a $C^k(M)$-module for any finite k, so we also have

$$\begin{aligned} \mathcal{D}'^{(k)}\mathcal{T}^r_s(M) &\cong \mathcal{D}'^{(k)}(M) \otimes_{C^k(M)} (\mathcal{T}^r_s)_{C^k}(M) \\ &\cong L_{C^k(M)}(\Omega^1_{C^k}(M)^r \times \mathfrak{X}_{C^k}(M)^s; \mathcal{D}'^{(k)}(M)). \end{aligned} \tag{2.3}$$

In other words, distributional tensor fields of order k act as C^k-balanced multilinear maps on one forms and vector fields of regularity C^k to give a scalar distribution of order k. For finite k, the multiplication $C^k(M) \times \mathcal{D}'^{(k)}(M) \to \mathcal{D}'^{(k)}(M)$ is jointly continuous (w.r.t. the strong topology), whereas for $k = \infty$ it is only hypocontinuous [20, p. 362], but still jointly sequentially continuous [37, p. 233] since $C^\infty(M)$ is Fréchet, hence barrelled. The corresponding continuity properties therefore also hold for the tensor operations introduced above. The isomorphisms (2.2), (2.3) are not topological, but still bornological by [36, Thm. 15] (which remains valid for spaces of C^k-sections with k finite).

Smooth, and in fact even L^1_{loc}-tensor fields are continuously and densely embedded via

$$\begin{aligned} &\mathcal{T}^r_s(M) \hookrightarrow \mathcal{D}'^{(k)}\mathcal{T}^r_s(M) \\ &\quad t \mapsto [(\theta_1, \ldots, \theta_r, X_1, \ldots, X_s) \mapsto [\omega \mapsto \int_M t(\theta_1, \ldots, \theta_r, X_1, \ldots, X_s)\omega]]. \end{aligned}$$

Here, ω is a one-density. If M is orientable, densities can be canonically identified with n-forms. The fact that $\mathcal{T}^r_s(M)$ is dense in $\mathcal{D}'^{(k)}\mathcal{T}^r_s(M)$ uniquely fixes all the operations on distributional tensor fields to be introduced below in a way compatible with smooth pseudo-Riemannian geometry.

For any $t \in \mathcal{T}^r_s(M)$ there is a unique extension that accepts one distributional argument. E.g., if $\tilde\theta_1 \in \mathcal{D}'\mathcal{T}^0_1(M)$, then since $t(\,.\,, \theta_2, \ldots, X_s) \in \mathfrak{X}(M)$ we may set

$$t(\tilde\theta_1, \theta_2, \ldots, X_s) := \tilde\theta_1(t(\,.\,, \theta_2, \ldots, X_s)) \in \mathcal{D}'(M), \tag{2.4}$$

and analogously for the other slots.

Definition 2.1. A *distributional connection* is a map $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathcal{D}'\mathcal{T}^1_0(M)$ satisfying for $X, X', Y, Y' \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$ the usual computational rules: $\nabla_{fX+X'}Y = f\nabla_X Y + \nabla_{X'}Y$, $\nabla_X(Y+Y') = \nabla_X Y + \nabla_X Y'$, $\nabla_X(fY) = X(f)Y + f\nabla_X Y$.

Using (2.4) any distributional connection can be extended to the entire tensor algebra, i.e., to a map $\nabla : \mathfrak{X}(M) \times \mathcal{T}^r_s(M) \to \mathcal{D}'\mathcal{T}^r_s(M)$, e.g. for $\Theta \in \Omega^1(M)$ we set

$$(\nabla_X \Theta)(Y) := X(\Theta(Y)) - \Theta(\nabla_X Y) \quad (X, Y \in \mathfrak{X}(M)). \tag{2.5}$$

Let $\mathcal{G}$ be any of the spaces C^k $(0 \leq k)$ or L^p_{loc} $(1 \leq p)$, then we call a distributional connection a *$\mathcal{G}$-connection* if $\nabla_X Y$ is a $\mathcal{G}$-vector field for any $X, Y \in \mathfrak{X}(M)$. A particularly important case are L^2_{loc}-connections since they form the largest class that allows for a stable definition of the curvature tensor in distributions, cf. [11,28,49].

Using a local frame one easily sees that by virtue of the respective computational rules any $\mathcal{G}$-connection can be (uniquely) extended to a map $\nabla : \mathfrak{X}_{\mathcal{F}}(M) \times \mathfrak{X}_{\mathcal{H}}(M) \to \mathfrak{X}_{\mathcal{G}}(M)$, provided that $\mathcal{F}$ and $\mathcal{H}$ are function spaces such that $\mathcal{F} \cdot \mathcal{G} \subseteq \mathcal{G}$, $D\mathcal{H} \subseteq \mathcal{G}$, and $\mathcal{H} \cdot \mathcal{G} \subseteq \mathcal{G}$. Hence, in particular, each L^2_{loc}-connection extends in this way to a map

$$\nabla : \mathfrak{X}_{C^0}(M) \times \mathfrak{X}_{C^1}(M) \to \mathfrak{X}_{L^2_{\text{loc}}}(M). \tag{2.6}$$

Moreover, by a similar reasoning, $\mathcal{G}$-connections allow one to insert even less regular vector fields in the first and second slot at the price of a less regular outcome. In particular, any L^2_{loc}-connection can be extended to a map

$$\nabla : \mathfrak{X}(M) \times \mathfrak{X}_{L^2_{\text{loc}}}(M) \to \mathcal{D}'\mathcal{T}^1_0(M). \tag{2.7}$$

Using this extension, we have [28, Def. 3.3]:

Definition 2.2. The distributional Riemann tensor of an L^2_{loc}-connection ∇ is the map $R : \mathfrak{X}(M)^3 \to \mathcal{D}'\mathcal{T}^1_0(M)$,

$$R(X, Y, Z)(\theta) \equiv (R(X, Y)Z)(\theta) := (\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z)(\theta)$$

for $X, Y, Z \in \mathfrak{X}(M)$ and $\theta \in \Omega^1(M)$,

If F_i is a local frame in $\mathfrak{X}(U)$ and $F^j \in \Omega^1(U)$ its dual frame, then the Ricci tensor corresponding to ∇ is given by

$$\text{Ric}(X, Y) := (R(X, F_i)Y)(F^i) \in \mathcal{D}'(U) \qquad (X, Y \in \mathfrak{X}(U)). \tag{2.8}$$

Recall the Koszul formula for the Levi-Civita connection of a smooth metric g on M:

$$\begin{aligned} 2g(\nabla_X Y, Z) &= X(g(Y, Z)) + Y(g(Z, X)) - Z(g(X, Y)) \\ &\quad - g(X, [Y, Z]) + g(Y, [Z, X]) + g(Z, [X, Y]) =: F(X, Y, Z) \end{aligned} \tag{2.9}$$

Focusing now on the case at hand in this work, suppose that g is merely C^1 (still allowing arbitrary signature). Then $F(X, Y, Z)$ is a well-defined element of $C^0(M)$, and

$$\nabla^\flat_X Y := Z \mapsto \frac{1}{2} F(X, Y, Z) \in \Omega^1_{C^0}(M) \subseteq \mathcal{D}'\mathcal{T}^0_1(M) \tag{2.10}$$

defines the *distributional Levi-Civita connection* of g [28, Def. 4.2]. Note that this is not (yet) a distributional connection in the sense of Definition 2.1 since it is of order $(0, 1)$ instead of $(1, 0)$. In addition to the standard product rules it also possesses the properties (corresponding to [39, Thm. 3.11])

$$\begin{aligned} &\nabla^\flat_X Y - \nabla^\flat_Y X = [X, Y]^\flat, \text{ i.e., } (\nabla^\flat_X Y - \nabla^\flat_Y X)(Z) = g([X, Y], Z), \\ &X(g(Y, Z)) = (\nabla^\flat_X Y)(Z) + (\nabla^\flat_X Z)(Y) \end{aligned} \tag{2.11}$$

for all $X, Y, Z \in \mathfrak{X}(M)$. These equalities hold distributionally, hence also in C^0.

In order to obtain an L^2_{loc}-connection from $\nabla^\flat$ (which in turn will allow us to define the curvature tensors we require) we need to raise the index via g, setting

$$g(\nabla_X Y, Z) := (\nabla^\flat_X Y)(Z) \qquad (X, Y, Z \in \mathfrak{X}(M)). \tag{2.12}$$

This defines a continuous vector field $\nabla_X Y$, hence even a C^0-connection. More generally, by [28, Sec. 4] this procedure yields an L^2_{loc}-connection even for the significantly larger Geroch-Traschen class of metrics ($g \in H^1_{\mathrm{loc}} \cap L^\infty_{\mathrm{loc}}$ and locally uniformly non-degenerate).

The Riemann tensor of $g \in C^1$ is now given by Definition 2.2. Furthermore, since in this case ∇ is a C^0-connection, it follows that $(R(X, Y)Z)(\theta) \in \mathcal{D}'^{(1)}(M)$, and consequently, $R \in \mathcal{D}'^{(1)}\mathcal{T}^1_3(M)$. Indeed by the same reasoning that led to (2.6), (2.7) we may extend the C^0-Levi-Civita connection of g to a map

$$\nabla : \mathfrak{X}_{C^0}(M) \times \mathfrak{X}_{C^1}(M) \to \mathfrak{X}_{C^0}(M), \text{ and } \nabla : \mathfrak{X}_{C^1}(M) \times \mathfrak{X}_{C^0}(M) \to \mathcal{D}'^{(1)}\mathcal{T}^1_0(M), \tag{2.13}$$

respectively. Here, the second equation shows that the term $([\nabla_X, \nabla_Y]Z)(\theta)$ yields an order one distribution, while $(\nabla_{[X,Y]}Z)(\theta) \in C^0$ since ∇ is a C^0-connection.

For $W, X, Y, Z \in \mathfrak{X}(M)$ we define $R(W, X, Y, Z) \in \mathcal{D}'^{(1)}(M)$ by

$$\begin{aligned} R(W, X, Y, Z) := {} & X(g(W, \nabla_Y Z)) - Y(g(W, \nabla_X Z)) \\ & - g(\nabla_X W, \nabla_Y Z) + g(\nabla_Y W, \nabla_X Z) - g(W, \nabla_{[X,Y]} Z). \end{aligned}$$

Using (2.4), (2.7) and (2.11) it is straightforward to check that cf. [28, Rem. 4.5]

$$R(W, Z, X, Y) = g(W, R(X, Y)Z).$$

Alternatively, one may verify this identity in local coordinates using the fact that there is a well-defined multiplication of distributions of first order with C^1-functions.

The Ricci tensor of a C^1 metric g is given by (2.8), hence, in particular, is again a distributional tensor field of order 1. Note that by (2.3) this allows us to apply it to C^1-vector fields, which will be essential in the context of the genericity condition below. Similarly, since the Riemann tensor, again by (2.3), acts on C^1-vector fields, we may in particular use g-orthonormal frames (which generically are only C^1) to calculate the Ricci tensor, which is also vital below. Explicitly, (2.8) holds for F_i a g-orthonormal frame and F^i its g-dual. All these facts can alternatively be derived directly from the extension of ∇ in (2.13).

Again, by the same observations, we have that the scalar curvature S is a distribution of order 1 (this answers a concern from [7, Sec. 3.2]) and that the standard local formulae hold in $\mathcal{D}'^{(1)}$:

$$\begin{aligned} R^m_{ijk} &= \partial_j \Gamma^m_{ik} - \partial_k \Gamma^m_{ij} + \Gamma^m_{js}\Gamma^s_{ik} - \Gamma^m_{ks}\Gamma^s_{ij}, \\ \mathrm{Ric}_{ij} &= R^m_{imj}, \\ S &= \mathrm{Ric}^i_i. \end{aligned}$$

This shows that the local definitions given in [7, (3.2),(3.3)] are compatible with the global approach developed here.

2.2. Regularisation. As in previous works on the generalisation of singularity theorems to metrics of differentiability below C^2 [7,24–26], an essential tool in our approach will be suitable regularisations, both of the metrics themselves and of the derived distributional curvature quantities. A general regularisation scheme based on chart-wise convolution with a mollifier $\rho \in \mathcal{D}(B_1(0))$, $\int \rho = 1$, $\rho \geq 0$ (cf. [13, 3.2.10], [24, Sec. 2], [7, Sec. 3.3]) is as follows: Cover M by a countable and locally finite family of relatively compact chart neighbourhoods (U_i, ψ_i) $(i \in \mathbb{N})$. Let $(\zeta_i)_i$ be a subordinate partition of unity with $\operatorname{supp}(\zeta_i) \subseteq U_i$ for all i and choose a family of cut-off functions $\chi_i \in \mathscr{D}(U_i)$ with $\chi_i \equiv 1$ on a neighbourhood of $\operatorname{supp}(\zeta_i)$. Finally, for $\varepsilon \in (0, 1]$ set $\rho_\varepsilon(x) := \varepsilon^{-n}\rho\left(\frac{x}{\varepsilon}\right)$. Then denoting by f_* (respectively f^*) push-forward (resp. pull-back) of distributions under a diffeomorphism f, for any $\mathcal{T} \in \mathcal{D}'\mathcal{T}^r_s(M)$ consider the expression

$$\mathcal{T} \star_M \rho_\varepsilon(x) := \sum_i \chi_i(x)\, \psi_i^*\Big(\big(\psi_{i\,*}(\zeta_i \cdot \mathcal{T})\big) * \rho_\varepsilon\Big)(x). \tag{2.14}$$

Here, $\psi_{i\,*}(\zeta_i \cdot \mathcal{T})$ is viewed as a compactly supported distributional tensor field on $\mathbb{R}^n$, so componentwise convolution with ρ_ε yields a smooth field on $\mathbb{R}^n$. The cut-off functions χ_i secure that $(\varepsilon, x) \mapsto \mathcal{T} \star_M \rho_\varepsilon(x)$ is a smooth map on $(0, 1] \times M$. For any compact set $K \Subset M$ there is an ε_K such that for all $\varepsilon < \varepsilon_K$ and all $x \in K$ (2.14) reduces to a finite sum with all $\chi_i \equiv 1$ (hence to be omitted from the formula), namely when ε_K is less than the distance between the support of $\zeta_i \circ \psi_i^{-1}$ and the boundary of $\psi_i(U_i)$ for all i with $U_i \cap K \neq \emptyset$.

Just as is the case for smoothing via convolution in the local setting, $\mathcal{T} \star_M \rho_\varepsilon$ converges to $\mathcal{T}$ as $\varepsilon \to 0$ in $\mathcal{D}'\mathcal{T}^r_s$, and indeed in C^k_{loc} or $W^{k,p}_{\text{loc}}$ $(p < \infty)$ if $\mathcal{T}$ is contained in these spaces [7, Prop. 3.5].

For a given C^1 Lorentzian metric g we will in addition require concrete regularisations that are adapted to the causal structure induced by g. This construction goes back to Chrusciel and Grant [5, Prop. 1.2], cf. also [24, Prop. 2.5]. The version we will use is [7, Lem. 4.2, Cor. 4.3]:

Lemma 2.3. *Let (M, g) be a C^1-spacetime. Then for any $\varepsilon > 0$ there exist smooth Lorentzian metrics $\check{g}_\varepsilon$, $\hat{g}_\varepsilon$ on M, time orientable by the same timelike vector field as g, and satisfying:*

(i) *$\check{g}_\varepsilon \prec g \prec \hat{g}_\varepsilon$ for all ε.*
(ii) *$\check{g}_\varepsilon$ and $\hat{g}_\varepsilon$ converge to g in C^1_{loc} as $\varepsilon \to 0$.*
(iii) *$\check{g}_\varepsilon - g \star_M \rho_\varepsilon \to 0$ in C^∞_{loc} and for any $K \Subset M$ there exist $c_K > 0$ and $\varepsilon_0(K)$ such that $\|\check{g}_\varepsilon - g \star_M \rho_\varepsilon\|_{\infty,K} \leq c_K \varepsilon$ and $\|g - g \star_M \rho_\varepsilon\|_{\infty,K} \leq c_K \varepsilon$ for all $\varepsilon < \varepsilon_0$. An analogous statement holds for $\hat{g}_\varepsilon$ as well as for the inverse metrics g^{-1}, $(\check{g}_\varepsilon)^{-1}$ and $(\hat{g}_\varepsilon)^{-1}$.*

In this work we will generally use $\check{g}_\varepsilon$, $\hat{g}_\varepsilon$ to denote the regularisations of Lemma 2.3, which are of utmost importance throughout.

Next we recall the following essential result on convergence of the curvature under regularisations, which is a special case of several so-called stability results, see [11, Thm. 2], [28, Thm. 4.6], [50, Thm. 5.1].

Lemma 2.4. *Let (M, g) be a C^1-spacetime, and let g_ε be either $\check{g}_\varepsilon$ or $\hat{g}_\varepsilon$. Then we have for the Riemann and the Ricci tensor*

$$R[g_\varepsilon] \to R[g] \quad \textit{and} \quad \operatorname{Ric}[g_\varepsilon] \to \operatorname{Ric}[g] \quad \textit{distributionally.} \tag{2.15}$$

In particular, the Riemann and the Ricci tensor possess the usual symmetries.

We will frequently need the following result from ODE theory (cf. [15, Ch.2, Thm. 3.2]), which we will mainly use in the context of regularisation given below in Corollary 2.6.

Proposition 2.5. *Let f and f_m, $m \in \mathbb{N}$ be continuous functions on an open set $E \subseteq \mathbb{R} \times \mathbb{R}^n$ with $f_m \to f$ in C^0_{loc}. Let $y_m = y_m(t)$ be a solution to the initial value problem*

$$y_m'(t) = f_m(t, y_m(t)), \quad y_m'(t_m) = y_{m0}$$

with data $(t_m, y_{m0}) \in E$ and maximal interval of existence (ω_m^-, ω_m^+). Suppose that the sequence of initial data converges,

$$(t_m, y_{m0}) \to (t_0, y_0) \in E.$$

Then there exist a solution $y = y(t)$ for the initial value problem

$$y'(t) = f(t, y(t)), \quad y(t_0) = y_0,$$

defined on a maximal interval (ω^-, ω^+), and a subsequence m_l with the property that for any $s_1 < s_2$ with $(s_1, s_2) \subseteq (\omega^-, \omega^+)$ we have $(s_1, s_2) \subseteq (\omega_{m_l}^-, \omega_{m_l}^+)$ for all large l, and $y_{m_l} \to y$ in $C^1_{\text{loc}}(\omega^-, \omega^+)$.[4] *In particular,*

$$\limsup_m \omega_m^- \leq \omega^- < \omega^+ \leq \liminf_m \omega_m^+.$$

We will mostly use the previous result to conclude that g_ε-geodesics, where $g_\varepsilon \in \{\check{g}_\varepsilon, \hat{g}_\varepsilon\}$, converge uniformly on compact intervals to a g-geodesic:

Corollary 2.6. *Let (M, g) be a C^1-spacetime and let $g_m = \check{g}_{\varepsilon_m}$ or $g_m = \hat{g}_{\varepsilon_m}$ with $\varepsilon_m \to 0$. Let $\gamma_m : (a_m, b_m) \to M$ be inextendible g_m-geodesics. Suppose that $\gamma_m(t_0) \to p \in M$ and $\dot{\gamma}_m(t_0) \to v \in T_pM$, where $t_0 \in \bigcap(a_m, b_m)$. Then there exist an inextendible g-geodesic $\gamma : (a, b) \to M$ with $\gamma(t_0) = p$, $\dot{\gamma}(t_0) = v$ and a subsequence m_l such that whenever $(s_1, s_2) \subseteq (a, b)$, we have $(s_1, s_2) \subseteq (a_{m_l}, b_{m_l})$ for all large l and γ_{m_l} converges in $C^2_{\text{loc}}(a, b)$ to γ. Moreover,*

$$\limsup_m a_m \leq a < b \leq \liminf_m b_m.$$

Proof. As this is a local statement, we may assume $M = \mathbb{R}^n$. We rewrite the geodesic equations for g_m as a first order initial value problem (with Γ_m the Christoffel symbols of g_m) and $1 \leq i \leq n$:

$$\begin{aligned}
\dot{\gamma}_m^i(t) &= \eta_m^i(t), \\
\dot{\eta}_m^i &= -(\Gamma_m)^i_{jk}(\gamma_m(t))\eta_m^j(t)\eta_m^k(t), \\
\gamma_m^i(t_0) &= p_m^i, \\
\eta_m^i(t_0) &= v_m^i.
\end{aligned}$$

We define the continuous functions $f_m : \mathbb{R}^{2n} \to \mathbb{R}$, $f_m(x, y) := (y, -(\Gamma_m)^i_{jk}(x)y^j y^k)$, and similarly $f : \mathbb{R}^{2n} \to \mathbb{R}$, $f(x, y) := (y, -\Gamma_{jk}(x)y^j y^k)$. Then $f_m \to f$ uniformly on compact subsets of $\mathbb{R}^{2n}$, and by assumption $(p_m, v_m) \to (p, v) \in \mathbb{R}^{2n}$. The existence of an inextendible g-geodesic sublimit γ of the sequence γ_m now follows from Proposition 2.5, and the C^2_{loc}-convergence follows suit. $\square$

[4] In [15, Ch.2, Thm. 3.2], convergence is only stated in C^0_{loc}, but inserting this into the equivalent integral equations immediately yields the claim as given here.

In the case of only future or past inextendibility, an obvious analogue of Corollary 2.6 holds.

Remark 2.7. (Parallel transport) Under the assumptions of Corollary 2.6, let $w_m \in T_{\gamma_m(t_0)}M$, $w_m \to w \in T_pM$. Then since the ODEs governing parallel transport are linear, there are unique smooth vector fields W_m along γ_m that are global solutions of the corresponding initial value problem, and by the standard results on continuous dependence of solutions on the right hand side and the data, they converge in C^1_{loc} to the unique C^1-vector field W along γ resulting from g-parallel transport of w along γ.

2.3. Energy conditions. Let us briefly recall the timelike and null energy conditions for C^1-spacetimes. They are formulated in a way so that they imply corresponding conditions for the curvatures of approximating metrics. If the metric is $C^{1,1}$ or better, then the energy conditions presented here are equivalent to the usual pointwise (or pointwise a.e.) energy conditions. We follow [7, Sec. 3-5], to which we also refer for proofs.
We start with the timelike energy condition, whose distributional generalisation is straightforward, cf. [7, Def. 3.3].

Definition 2.8 (*Distributional timelike energy condition*). Let (M, g) be a C^1-spacetime. We say that (M, g) satisfies the *distributional timelike energy condition* if for any timelike $X \in \mathfrak{X}(M)$, $\mathrm{Ric}(X, X) \in \mathcal{D}'^{(1)}(M)$ is a nonnegative scalar distribution on M.

Recall that a scalar distribution $u \in \mathcal{D}'(M)$ is nonnegative, $u \geq 0$, if $u(\omega) \equiv \langle u, \omega\rangle \geq 0$ for all nonnegative test densities $\omega \in \Gamma_c(M, \mathrm{Vol}(M))$. A nonnegative distribution is always a measure [19, Thm. 2.1.7] and hence a distribution of order 0. Moreover, nonnegativity is stable with respect to regularisation [19, Thm. 2.1.9]. Finally, for $u, v \in \mathcal{D}'(M)$ we write $u \geq v$ if $u - v \geq 0$.

Remark 2.9. We note that an equivalent formulation of the distributional timelike energy condition consists in imposing that, for any $U \subseteq M$ open and any $X \in \mathfrak{X}(U)$ timelike, $\mathrm{Ric}(X, X) \geq 0$ in $\mathcal{D}'^{(1)}(U)$. Indeed, given a timelike $X \in \mathfrak{X}(U)$, let $V \Subset U$ be open and choose a cut-off function $\chi \in \mathcal{D}(U)$ that equals 1 on a neighbourhood of $\overline{V}$. By time-orientability of M there exists some timelike $T \in \mathfrak{X}(M)$ of the same time-orientation as X, so $Y := \chi X + (1 - \chi)T$ is a global timelike vector field on M. Then for any non-negative $\omega \in \Gamma_c(M, \mathrm{Vol}(M))$ with support in V we have $0 \leq \langle \mathrm{Ric}(Y, Y), \omega\rangle = \langle \mathrm{Ric}(X, X), \omega\rangle$, and by the sheaf property of $\mathcal{D}'^{(1)}$, we obtain $\mathrm{Ric}(X, X) \geq 0$ in $\mathcal{D}'^{(1)}(U)$.

The usefulness of Definition 2.8 is highlighted by its implications for the curvatures of approximating metrics, cf. [7, Lem. 4.1, Lem. 4.6].

Lemma 2.10. *Let (M, g) be a C^1-spacetime satisfying the distributional timelike energy condition and let K be compact in M. Then*

$$\begin{aligned}&\forall C > 0\ \forall \delta > 0\ \forall \kappa < 0\ \exists \varepsilon_0 > 0\ \forall \varepsilon < \varepsilon_0\ \forall X \in TM|_K\\ &\textit{with } g(X, X) \leq \kappa \textit{ and } \|X\|_h \leq C :\ \mathrm{Ric}[\check{g}_\varepsilon](X, X) > -\delta.\end{aligned}$$

The following is an immediate consequence of the previous result.

Corollary 2.11. *Let (M, g) be a C^1-spacetime satisfying the distributional timelike energy condition and let K be compact in M. Then*

$$\forall \delta > 0\ \exists \varepsilon_0 > 0\ \forall \varepsilon < \varepsilon_0\ \forall X \in TM|_K \textit{ with } \check{g}_\varepsilon(X, X) = -1 :\ \mathrm{Ric}[\check{g}_\varepsilon](X, X) > -\delta.$$

Let us now turn to the proper distributional formulation of the null energy condition. The naive approach of defining it in analogy to Definition 2.8 does not yield analogues of Lemma 2.10 and Corollary 2.11 because vector fields that are $\check{g}_\varepsilon$-null are only almost g-null. This necessitates a definition of the following form.

Definition 2.12 (*Distributional null energy condition*, [7, Def. 5.1]). Let (M, g) be a C^1-spacetime. We say (M, g) satisfies the *distributional null energy condition* if for any compact set K and any $\delta > 0$ there exists $\varepsilon = \varepsilon(\delta, K) > 0$ such that $\mathrm{Ric}(X, X) > -\delta$ (as distributions) for any local smooth vector field $X \in \mathfrak{X}(U)$ ($U \subseteq K$ open) with $\|X\|_h = 1$ and $|g(X, X)| < \varepsilon$ on U.

The above requirement $|g(X, X)| < \varepsilon$ may equivalently be replaced by $\|X - N\|_h < \varepsilon$, where N is a C^1 g-null vector field on U, cf. [7, Lem. 5.2]. Yet another useful equivalent reformulation is the following rescaled version, cf. [7, Def. 5.3].

Lemma 2.13. *A C^1-spacetime (M, g) satisfies the distributional null energy condition if and only if for any compact $K \subseteq M$, any $c_1, c_2 > 0$ and any $\delta > 0$ there is $\varepsilon = \varepsilon(\delta, K, c_1, c_2) > 0$ such that $\mathrm{Ric}(X, X) > -\delta$ (as distributions) for any local smooth vector field $X \in \mathfrak{X}(U)$, $U \subseteq K$, with $0 < c_1 \leq \|X\|_h \leq c_2$ and $|g(X, X)| < \varepsilon$ on U.*

The following analogue of Lemma 2.10 and Corollary 2.11 shows that the above definition of the null energy condition is the correct one in the C^1-case (see [7, Lem. 5.5]).

Lemma 2.14. *Let (M, g) be a C^1-spacetime satisfying the distributional null energy condition. Let $K \subseteq M$ be compact and let $c_1, c_2 > 0$. Then for all $\delta > 0$ there is $\varepsilon_0 = \varepsilon_0(\delta, K, c_1, c_2) > 0$ such that*

$$\forall \varepsilon < \varepsilon_0 \; \forall X \in TM|_K \text{ with } 0 < c_1 \leq \|X\|_h \leq c_2 \text{ and}$$
$$\check{g}_\varepsilon(X, X) = 0 : \; \mathrm{Ric}[\check{g}_\varepsilon](X, X) > -\delta.$$

The above distributional versions of the timelike and null energy conditions were successfully used in [7] to prove the singularity theorems of Hawking and of Penrose for C^1-spacetimes. For our C^1-proof of the Hawking-Penrose singularity theorem, we need a distributional version of the genericity condition along geodesics.

Definition 2.15 (*Distributional genericity condition*). Let (M, g) be a C^1-spacetime and let $\gamma : I \to M$ be a causal geodesic. We say that the *distributional genericity condition* holds at $\gamma(t_0)$, $t_0 \in I$, if there is a neighbourhood U of $\gamma(t_0)$, C^1-vector fields X, V on U with the property that $(X \circ \gamma)(t) = \dot{\gamma}(t)$, $(V \circ \gamma)(t) \perp \dot{\gamma}(t)$ for all $t \in I$ with $\gamma(t) \in U$, and there exist $c > 0$ and $\delta > 0$ such that for all C^1-vector fields $\tilde{X}, \tilde{V}$ on U with $\|X - \tilde{X}\|_h < \delta$ and $\|V - \tilde{V}\|_h < \delta$ we have

$$g(R(\tilde{X}, \tilde{V})\tilde{V}, \tilde{X}) > c \quad \text{in } \mathcal{D}'^{(1)}(U). \tag{2.16}$$

This condition extends the one used in the $C^{1,1}$-case in [8, Def. 2.2] as we shall discuss in Remark 2.19, below. Again it allows us to derive an approximate genericity condition for regularisations. Note, however, that it will turn out to be essential to use C^1-vector fields (and their approximations), which can be inserted into $R \in \mathcal{D}'^{(1)}\mathcal{T}^1_3(M)$, by (2.3).

Also the distributional genericity condition appropriately restricts to smaller sets, a technical point which we clarify first.

Remark 2.16. In Definition 2.15 we may shrink the neighbourhood U while retaining the constants c and δ. More precisely, let U, X, V, c, and δ be as in the definition and let $W \Subset U$, then (2.16) holds for all C^1-vector fields $\tilde{X}_W$, $\tilde{V}_W$ on W which satisfy $\|X|_W - \tilde{X}_W\|_h < \delta$ and $\|V|_W - \tilde{V}_W\|_h < \delta$. Indeed, by the sheaf property of distributions, it suffices to establish that (2.16) then holds locally around each point $x \in W$. To establish this let $\chi \in \mathcal{D}(W)$ be a plateau function that equals 1 in a neighbourhood W' of x and set $\tilde{X} = X + \chi(\tilde{X}_W - X)$ and likewise for $\tilde{V}$. Then $\tilde{X}$ and $\tilde{V}$ are C^1-vector fields on U with $\|X - \tilde{X}\|_h < \delta$ and $\|V - \tilde{V}\|_h < \delta$ (2.16). So we obtain on W'

$$c < g(R(\tilde{X}, \tilde{V})\tilde{V}, \tilde{X}) = g(R(\tilde{X}_W, \tilde{V}_W)\tilde{V}_W, \tilde{X}_W). \tag{2.17}$$

Lemma 2.17. *Let (M, g) be a C^1-spacetime and let g_ε be either $\check{g}_\varepsilon$ or $\hat{g}_\varepsilon$. Suppose γ is a g-causal g-geodesic and assume that the distributional genericity condition holds at $\gamma(t_0)$ (with neighbourhood U, vector fields X, V and constants c, δ). Suppose that X_ε, V_ε are C^1-vector fields on U with $X_\varepsilon, V_\varepsilon \to X, V$ in $C^1_{\text{loc}}(U)$. Then for any compact subset $K \Subset U$ there is $\tilde{\varepsilon} > 0$ such that for all $\varepsilon < \tilde{\varepsilon}$:*

$$g_\varepsilon(R[g_\varepsilon](X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon) > \frac{c}{2}$$

on K.

Proof. We may take U to be relatively compact. By assumption, for all C^1-vector fields $\tilde{X}$, $\tilde{V}$ on U with $\|X - \tilde{X}\|_h < \delta$ and $\|V - \tilde{V}\|_h < \delta$, we have

$$g(R(\tilde{X}, \tilde{V})\tilde{V}, \tilde{X}) > c \quad \text{in } \mathcal{D}'^{(1)}(U).$$

Since $X_\varepsilon \to X$ and $V_\varepsilon \to V$ in C^1_{loc}, there is $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$, $\|X_\varepsilon - X\|_h < \delta$ and $\|V_\varepsilon - V\|_h < \delta$. Let from now on $\varepsilon < \varepsilon_0$. Hence, the assumption gives

$$g(R(X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon) > c \quad \text{in } \mathcal{D}'^{(1)}(U).$$

Since $\star_M$-convolution with a non-negative mollifier ρ_ε respects positivity, we have

$$g(R(X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon) \star_M \rho_\varepsilon > c.$$

We claim that

$$g(R(X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon) \star_M \rho_\varepsilon - (g \star_M \rho_\varepsilon)(R[g \star_M \rho_\varepsilon](X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon) \to 0 \quad \text{in } C^0_{\text{loc}}(U).$$

This is seen by first showing that

$$g(R(X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon) \star_M \rho_\varepsilon - g((R[g] \star_M \rho_\varepsilon)(X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon) \to 0 \quad \text{in } C^0_{\text{loc}}(U).$$

The latter locally corresponds to the question whether given a net h_ε of C^1-functions that converges in C^1_{loc} to $h \in C^1$ and a first order distribution T we can show $(h_\varepsilon T) * \rho_\varepsilon - h_\varepsilon(T * \rho_\varepsilon) \to 0$ in C^0_{loc}. This is shown by arguing along the lines of [7, Lem. 4.8] and using that T can locally be written as a first order distributional derivative of a C^0-function. By a similar argument (noting the fast convergence speed of $g \star_M \rho_\varepsilon \to g$), the right hand side above is C^0_{loc}-equivalent (i.e. the difference goes to 0 in C^0_{loc}) to

$$(g \star_M \rho_\varepsilon)(R[g] \star_M \rho_\varepsilon(X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon).$$

This is C^0_{loc}-equivalent to

$$(g \star_M \rho_\varepsilon)(R[g \star_M \rho_\varepsilon](X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon)$$

by the same arguments as in [7, Lem. 4.6] and uniform boundedness of X_ε, V_ε.

Next, we show that

$$(g \star_M \rho_\varepsilon)(R[g \star_M \rho_\varepsilon](X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon) - g_\varepsilon(R[g_\varepsilon](X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon) \to 0$$

in $C^0_{\mathrm{loc}}(U)$. Upon writing everything out in coordinates, we see that only terms of the form

$$(g \star_M \rho_\varepsilon)\partial\Gamma[g \star_M \rho_\varepsilon]X_\varepsilon V_\varepsilon V_\varepsilon X_\varepsilon - g_\varepsilon \partial\Gamma[g_\varepsilon]X_\varepsilon V_\varepsilon V_\varepsilon X_\varepsilon$$

are of interest, because for any of the other terms in the difference, both summands converge in $C^0_{\mathrm{loc}}(U)$ to the respective expressions for g. Since the X_ε, V_ε are uniformly bounded, we may omit them in our estimates without loss of information. But for terms as above, we only need to consider those expressions where we have second derivatives of the metrics, since $\partial_j(g \star_M \rho_\varepsilon)^{mk} \to \partial_j g^{mk}$ in $C^0_{\mathrm{loc}}(U)$ and similarly for $\partial_j(g_\varepsilon)^{mk}$. Hence, we work in a chart, fix a compact set K and are left with estimating the following expression on K:

$$\begin{aligned}
&|(g \star_M \rho_\varepsilon)_{bc}(g \star_M \rho_\varepsilon)^{mk}\partial_s\partial_r(g \star_M \rho_\varepsilon)_{ij} - (g_\varepsilon)_{bc}(g_\varepsilon)^{mk}\partial_s\partial_r(g_\varepsilon)_{ij}| \\
&\leq |(g \star_M \rho_\varepsilon)_{bc}(g \star_M \rho_\varepsilon)^{mk}| \cdot |\partial_s\partial_r(g \star_M \rho_\varepsilon)_{ij} - \partial_s\partial_r(g_\varepsilon)_{ij}| + II.
\end{aligned}$$

The first term goes to 0 uniformly on K by Lemma 2.3(iii). We continue estimating the remaining term II:

$$\begin{aligned}
II &= |(g \star_M \rho_\varepsilon)_{bc}(g \star_M \rho_\varepsilon)^{mk}\partial_s\partial_r(g_\varepsilon)_{ij} - (g_\varepsilon)_{bc}(g_\varepsilon)^{mk}\partial_s\partial_r(g_\varepsilon)_{ij}| \\
&\leq |\partial_s\partial_r(g_\varepsilon)_{ij}| \cdot |(g \star_M \rho_\varepsilon)_{bc}(g \star_M \rho_\varepsilon)^{mk} - (g_\varepsilon)_{bc}(g_\varepsilon)^{mk}|.
\end{aligned}$$

To show that this goes to 0 uniformly on K, it suffices to show that the second factor can be estimated by a constant times ε by the same argument as in the proof of [7, Lem. 4.5]. Since $\|g_\varepsilon - (g \star_M \rho_\varepsilon)\|_{\infty,K}$ and $\|g_\varepsilon^{-1} - (g \star_M \rho_\varepsilon)^{-1}\|_{\infty,K}$ can both be estimated by a constant times ε again by Lemma 2.3(iii), this is easily seen to hold.

Hence, there is $\varepsilon_1 > 0$ such that for all $\varepsilon < \min(\varepsilon_0, \varepsilon_1)$,

$$g_\varepsilon(R[g_\varepsilon](X_\varepsilon, V_\varepsilon)V_\varepsilon, X_\varepsilon) > \frac{c}{2} \quad \text{on } K.$$

□

Next we derive from the distributional genericity condition an estimate on the tidal force operator for the regularised metrics. It is this result which allows us to enter into the Riccati comparison techniques in Sect. 4.

Lemma 2.18. *Let the assumptions of Lemma 2.17 hold and let γ_ε be g_ε-causal g_ε-geodesics, whose g_ε-causal character is the same as the g-causal character of γ and which converge in C^1_{loc} to γ. Then there are $\tilde{c} > 0$, $r > 0$, $\varepsilon_0 > 0$ and a neighbourhood U of $\gamma(t_0)$ such that for any $0 < \varepsilon < \varepsilon_0$ there exist vector fields E_i^ε on U such that $E_i^\varepsilon \circ \gamma_\varepsilon$ constitutes a g_ε-orthonormal frame along γ_ε and such that $E_i^\varepsilon \to E_i$ in $C^1_{\mathrm{loc}}(U)$, where $E_i \circ \gamma$ is an orthonormal frame along γ. Further, there exists some $C = C(\varepsilon) > 0$ such* *that along γ_ε the ε-tidal force operator* $[R_\varepsilon](t) := [R_\varepsilon(., \dot\gamma_\varepsilon(t))\dot\gamma_\varepsilon(t)] : [\dot\gamma_\varepsilon(t)]^\perp \to [\dot\gamma_\varepsilon(t)]^\perp$ *fulfills*

$$[R_\varepsilon](t) > \operatorname{diag}(\tilde c, -C, \dots, -C) \text{ on } [t_0 - r, t_0 + r] \tag{2.18}$$

in terms of the frame E_i^ε and where we have used the shorthand R_ε for $R[g_\varepsilon]$.

Proof. Let w.l.o.g. $t_0 = 0$. We deal with the case of γ being null or timelike simultaneously and construct a g-orthonormal frame for $[\dot\gamma]^\perp$ along γ. First note that, if γ is timelike, $V \circ \gamma$ is nowhere proportional to $\dot\gamma$, but the same may not be true in the null case. First we will show that we can replace V by a vector field whose restriction to γ is parallel in $[\dot\gamma]^\perp$ and still retain (2.16).

There must exist some t such that $V \circ \gamma(t) \notin \operatorname{span}(\dot\gamma(t))$. Indeed, suppose to the contrary that $V(\gamma(t)) = f(t) \cdot X(\gamma(t))$ for some real-valued C^1 function f and all t. Then extending both $X \circ \gamma$ and $V \circ \gamma$ in a cylindrically constant fashion as given by Lemma B.1, we obtain vector fields $\tilde V, \tilde X$ on a neighbourhood W of $\gamma(0)$ that are proportional and can be made arbitrarily close to V resp. X in C^0 when shrinking W since they restrict to the same vector fields on γ. According to Definition 2.15 they therefore satisfy (2.16) close to $\gamma(0)$. However, by the symmetry properties of R (cf. Lemma 2.4) and the proportionality of $\tilde V$ and $\tilde X$ we must have $R(\tilde X, \tilde V)\tilde V = 0$ in $\mathcal{D}'(W)$, a contradiction.

Hence there exists some $v := V(t_1)$ with $v \in [\dot\gamma(t_1)]^\perp$. Defining V_1 as the parallel translate of v along γ we obtain $V_1 \in [\dot\gamma]^\perp$ along all of γ. By the same continuity argument as above, (2.16) then still holds for the vector fields V_1 and X (note that a t_1 as above must exist arbitrarily close to 0). In the timelike case we may assume that γ is parametrised to unit speed, and we set $e_n := \dot\gamma(0)$. If γ is null we can write $\dot\gamma(0) = e_{n-1} + e_n$ for orthonormal vectors e_{n-1}, e_n with e_n timelike.

In the null case, extend $e_1 := V_1(0) \in [\dot\gamma(0)]^\perp$ to a g-orthonormal basis $\{e_i\}_{i=1}^n$ (with e_{n-1}, e_n as above) and define E_i to be the g-parallel translate of e_i along γ (so that $E_1 = V_1$ and $X = E_{n-1} + E_n$ along γ). For γ timelike we also construct an orthonormal frame $E_1 = V_1, E_2, \dots, E_n = X$ along γ via parallel transport. As $\gamma_\varepsilon \to \gamma$ in C^1_{loc}, we can find a g_ε orthonormal basis $\{e_i^\varepsilon\}$ at $\gamma_\varepsilon(0)$ with $e_i^\varepsilon \to e_i$ in TM, and such that $\dot\gamma_\varepsilon(0) = e_n^\varepsilon$ if γ_ε is g_ε-timelike, while $\dot\gamma_\varepsilon(0) = e_{n-1}^\varepsilon + e_n^\varepsilon$ with e_n^ε timelike if γ_ε is g_ε-null. For every ε let E_i^ε be the g_ε-parallel translate of e_i^ε along γ_ε. By Remark 2.7, $E_i^\varepsilon \to E_i$ in C^1_{loc} (i.e. in $T(TM)$, uniformly on compact time intervals) and we can use Lemma B.2 to extend E_i^ε to vector fields on U, converging in $C^1_{\mathrm{loc}}(U)$ to the extensions of E_i to U.

As we have now found vector fields E_i^ε that restrict to a g_ε-frame along γ_ε and converge to the g-frame E_i along γ, we are in a position to apply Lemma 2.17 and hence estimate the g_ε-tidal force operator along γ_ε.

For the remainder of the proof we roughly follow the layout of the proof of [8, Prop. 3.6]. There, for a $C^{1,1}$-metric g the statement is first shown for the g-tidal force operator along γ and then carried over for small ε by an approximation argument. However, in the C^1-setting this is no longer possible, as R is now merely distributional and for a matrix with distributional entries eigenvalues are not defined. Also, a direct approximation procedure of that form is no longer feasible. Instead we use Lemma 2.17, and argue separately for any ε, making the procedure (and hence the constant C) dependent on ε[5]. More precisely:

Let c be the constant given by the genericity condition and let $\tilde c := \frac{c}{2}$ and $0 < c_1 < \tilde c$. Set $R_{ij}^\varepsilon := \langle R_\varepsilon(E_i^\varepsilon, X_\varepsilon)X_\varepsilon, E_j^\varepsilon\rangle$, $i, j = 1, \dots, d$, where $X_\varepsilon = E_n^\varepsilon$ in the timelike case

[5] This is, however, sufficient in our proof of the main result.

and $X_\varepsilon = E^\varepsilon_{n-1} + E^\varepsilon_n$ in the null case. As in the proof of (3.5) in [8, Lemma 3.7], we choose $C(\varepsilon) > 0$ such that $\lambda^\varepsilon_{\min}(x) \geq \frac{\|(R^\varepsilon_{1j})_j\|_e}{\tilde{c}-c_1}$ for all $x \in U$, where $\lambda^\varepsilon_{\min}(x)$ is the smallest eigenvalue of $(R^\varepsilon_{ij})^d_{i,j=2} + C(\varepsilon)\mathrm{Id}_{d-1}$ at $x \in U$. By Lemma 2.17 there exists some $\tilde{\varepsilon} > 0$ such that $R^\varepsilon_{11} > \tilde{c}$ for all $0 < \varepsilon < \tilde{\varepsilon}$ on U (shrinking U once more if necessary). For these ε it then follows that $R^\varepsilon_{ij} - \mathrm{diag}(c_1, -C(\varepsilon), \ldots - C(\varepsilon))$ is positive definite on U. Then picking $r > 0$ such that $\gamma_\varepsilon([-r, r]) \subseteq U$ for all ε small, evaluation along γ_ε gives the claim. □

Remark 2.19. Note that while Definition 2.15 may at first sight appear to be stronger than a straightforward generalisation of the genericity condition given for $C^{1,1}$-metrics in [8], due to the local boundedness of R in that regularity it is actually equivalent:

For $C^{1,1}$-metrics Definition 2.15 clearly implies Definition 2.2 of [8]. For the reverse implication, let X, V be continuous vector fields in a neighbourhood of $\gamma(t_0)$ such that $X(\gamma(t)) = \dot{\gamma}(t)$ and $V(\gamma(t)) \in (\dot{\gamma}(t))^\perp$ for all t and such that $\langle R(V, X)X, V\rangle > c > 0$. Arguing exactly as in the proof of Lemma 2.18 we can then use parallel transport and cylindrically constant extension to find C^1-vector fields that are uniformly close to X and V. It therefore suffices to show that the genericity condition from [8] is stable under locally uniform perturbation. Thus let $\tilde{V}, \tilde{X}$ be C^1 vector fields that are uniformly close to V, X. Then one can locally in a coordinate system estimate $|\tilde{V}_i - V_i| \leq \delta$, $|\tilde{X}_i - X_i| \leq \delta$ for some small δ. Consequently,

$$\langle R(\tilde{V}, \tilde{X})\tilde{X}, \tilde{V}\rangle = g_{ij} R^i_{klm} \tilde{V}^l \tilde{X}^m \tilde{X}^k \tilde{V}^j \geq g_{ij} R^i_{klm} V^l X^m X^k V^j + \mathcal{O}(\delta)$$

where the last term collects all δ-contributions and takes into account the local boundedness of R. Consequently, choosing δ small enough (and appropriately shrinking the neighbourhood of $\gamma(t_0)$) we can secure that the sum still remains $> c/2$.

3. Branching

In metric geometry, where geodesics are defined as (local) minimisers of the length functional, local uniqueness of geodesics is expressed in the form of a *non-branching* condition. Here, a branch point is defined as an element of a minimiser at which the curve splits into two minimisers that on some positive parameter interval do not have another point in common, cf., e.g., [44,51]. Similarly, in the synthetic Lorentzian setting [22], the role of causal geodesics is taken on by maximising causal curves, and non-branching is formulated analogously. In both cases, lower synthetic sectional curvature bounds (formulated via triangle comparison in constant curvature model spaces) imply non-branching [22,44]. Synthetic Ricci curvature bounds in metric measure spaces (curvature-dimension conditions), on the other hand, do not imply non-branching of geodesics [38]. Similarly, in recent work on synthetic Ricci curvature bounds in Lorentzian pre-length spaces [4], a timelike non-branching condition is required in addition to a timelike-curvature dimension condition to obtain a version of the Hawking singularity theorem.

In the C^1-setting we are concerned with in this paper, the coincidence between causal local maximisers and geodesics that is familiar from smooth Lorentzian geometry ceases to hold (cf. Example 3.2 below), although it is still true that causal maximisers are geodesics (Lemma A.4). In addition, one generically cannot expect unique local solvability of the geodesic initial value problem. Nevertheless, on physical grounds, it seems reasonable to assign a privileged role to causal geodesics that *are* locally maximising.

We therefore introduce a (weak notion of a) non-branching condition that is intended to preclude locally maximising geodesics from branching (where we do not require the second branch to be maximising as well):

Definition 3.1 (*Non-branching conditions*). Let (M, g) be a C^1-spacetime. A geodesic $\gamma : [a, b] \to M$ branches at $t_0 \in (a, b)$ if there exist $\varepsilon > 0$ and some geodesic σ with $\gamma|_{[t_0-\varepsilon,t_0]} \subseteq \sigma$, but $\gamma|_{(t_0,t_0+\varepsilon)} \cap \sigma = \emptyset$.
(M, g) is called maximally causally (resp. timelike, resp. null) non-branching (MCNB, MTNB, MNNB), if no maximal causal (resp. timelike, resp. null) geodesic branches in the above sense.

Example 3.2. A C^1-spacetime in which maximal causal branching occurs can be constructed from the second Riemannian example given in [16]. There, a $C^{1,\alpha}$-Riemannian metric ($\alpha < 1$) in the (u, v)-plane is given with non-unique solutions to the geodesic initial value problem starting at $\{v = 0\}$. Further it is shown that the geodesic boundary value problem for geodesics γ starting at the surface $\{v = 0\}$ is uniquely solvable and hence such geodesics are at least initially minimizing: If they were not, by properties of any C^1-Riemannian manifold as a locally compact length space, there would exist a minimiser σ between two points on γ. However in C^1-Riemannian manifolds minimisers are geodesics and by uniqueness (of the boundary value problem) $\gamma = \sigma$ (cf. [42] for proofs and references).

In [18] it is shown that causal geodesics in static spacetimes are maximising if and only if their Riemannian parts are minimizing and so one can easily construct both maximising timelike and null geodesics that branch.

Lemma 3.3. *If the C^1-spacetime (M, g) is MCNB (resp. MTNB, MNNB), then any two maximal causal (resp. timelike, null) geodesics that meet tangentially at an interior point must coincide on their maximal domain of definition.*

Proof. We only show this for MCNB spacetimes, the other cases are proven in exactly the same way by replacing all occurrences of “causal" with “timelike” resp. “null”. Let γ, σ be two maximising, causal geodesics meeting tangentially at p, which is not an endpoint of either curve. W.l.o.g. we may assume that $\gamma(0) = \sigma(0) = p$. It suffices to show that there is $\varepsilon > 0$ such that $\gamma|_{[0,\varepsilon)} = \sigma|_{[0,\varepsilon)}$. Indeed since then γ and σ meet tangentially at $t = \varepsilon$, the maximal such interval coincides with the maximal domain of definition (to the right and analogously to the left).

Assume to the contrary that $\gamma|_{[0,\varepsilon)} \neq \sigma|_{[0,\varepsilon)}$ for any ε. As the curves both satisfy the geodesic equation at p and their tangents agree, so do their second derivatives. Thus γ and $\gamma|_{(-\varepsilon,0]} \cup \sigma|_{[0,\varepsilon)}$ are both (unbroken) geodesics and by MCNB γ cannot branch at $t = 0$. So for every $\varepsilon > 0$ we have $\gamma|_{(0,\varepsilon)} \cap \sigma|_{(0,\varepsilon)} \neq \emptyset$. Hence there exist $t_k \downarrow 0$ such that $\gamma(t_k) = \sigma(t_k)$, and by our indirect assumption there exists $s_k \downarrow 0$ such that $\gamma(s_k) \neq \sigma(s_k)$. For any k set

$$\eta_k := \sup\{a \in [0, s_k] : \gamma|_{(s_k-a,s_k]} \cap \sigma|_{(s_k-a,s_k]} = \emptyset\}.$$

Further choose some $\tilde{\eta}_k > 0$ such that for $I_k := (s_k-\eta_k, s_k+\tilde{\eta}_k)$ we have $\gamma|_{I_k} \cap \sigma|_{I_k} = \emptyset$. Since $t_k \downarrow 0$ and $\gamma(t_k) = \sigma(t_k)$, we know that $s_k > \eta_k$. However, $\gamma(s_k-\eta_k) = \sigma(s_k-\eta_k)$ by our choice of η_k. Set $s_k - \eta_k = r_k$ and for simplicity of notation assume $r_k < t_k$. If $\dot{\gamma}(r_k) \neq \dot{\sigma}(r_k)$, the curve $\alpha := \gamma|_{[0,r_k]} \cup \sigma|_{[r_k,t_k]}$ is a broken geodesic from p to $\gamma(t_k)$ and hence not maximising by [43, Lem. 3.2], i.e., $d(p, \gamma(t_k)) > L(\alpha)$. Since γ and σ are maximisers we have $d(p, \gamma(r_k)) = L(\gamma|_{[0,r_k]}) = L(\sigma|_{[0,r_k]})$ Consequently,

$$d(p, \sigma(t_k)) = L(\sigma|_{[0,t_k]}) = L(\sigma|_{[0,r_k]}) + L(\sigma|_{[r_k,t_k]}) = L(\alpha) < d(p, \sigma(t_k)),$$

a contradiction, and hence $\dot\gamma(r_k) = \dot\sigma(r_k)$.

Again employing the geodesic equation, also the second derivatives of γ and σ coincide at r_k, and so the curves $\gamma|_{[0,s_k]}$ and $\gamma|_{[0,r_k]} \cup \sigma|_{[r_k,s_k]}$ display maximal causal branching, a contradiction. □

The following result shows that in maximally causally non-branching spacetimes, maximal causal geodesics can always be approximated by geodesics for the regularised spacetimes $\hat g_\varepsilon$ and $\check g_\varepsilon$. Whether such a result is true in general is questionable and very likely regularisation dependent.

Proposition 3.4. *Let (M, g) be a C^1-spacetime that is MCNB.*

(i) *Suppose that M is globally hyperbolic and let $\varepsilon_k \searrow 0$ $(k \to \infty)$. Set $g_k := \hat g_{\varepsilon_k}$ or $g_k := \check g_{\varepsilon_k}$ for all k. If $\gamma : [0, a] \to M$ is a maximising, timelike g-geodesic then for any small $\delta > 0$ there exists a subsequence g_{k_l} of g_k and g_{k_l}-maximising, timelike geodesics γ_l converging in C^1 to $\gamma|_{[0,a-\delta]}$.*

(ii) *Let $g_k = \check g_{\varepsilon_k}$ for all k. If M is causal and $\gamma : [0, a] \to M$ is a g-maximizing null geodesic, then for any small $\delta > 0$, there exists a subsequence g_{k_l} of g_k, $t_l \downarrow 0$ and g_{k_l}-null geodesics $\gamma_l : [t_l, a-\delta] \to M$ contained in $\partial I^+_l(\gamma(0))$ (hence in particular g_{k_l}-maximising), which converge in C^1_{loc} to $\gamma|_{[0,a-\delta]}$.*

If, furthermore, M is strongly causal and S is an acausal set such that $\gamma \subseteq E^+(S)$, then even $\gamma_l : [t_l, a-\delta] \to E^+_l(S)$ and $\gamma_l(t_l) \in S$ with $\gamma_l(t_l) \to \gamma(0)$.

Proof. Fix $\delta \in (0, a)$ and set $p := \gamma(0)$, $q := \gamma(a)$, and $q_\delta := \gamma(a - \delta)$.

(i) Since γ is timelike, $q_\delta \in I^+(p)$ and without loss of generality we may assume that also $q_\delta \in I^+_k(p)$ for all k. Since g is globally hyperbolic we can also assume w.l.o.g. that g_k are globally hyperbolic as well (see [26, Prop. 2.3 (iv)], which remains valid for $g \in C^1$). Thus there exist g_k-maximising, timelike geodesics γ_k from p to q_δ [41, Prop. 6.5]. By Corollary 2.6 an affinely reparametrised subsequence, without loss of generality γ_k itself, converges in C^1_{loc} to a g-causal geodesic σ. By Lemma A.13 σ is a g-maximising timelike geodesic from p to q_δ, in particular $d(p, q_\delta) = L(\gamma|_{[0,a-\delta]}) = L(\sigma)$. We cannot yet apply Lemma 3.3 since it is not clear whether σ is maximising beyond q_δ. However denote by $\eta := \sigma \cup \gamma|_{[a-\delta,a]}$ the concatenation of σ and γ. Since $d(p, q) = L(\gamma) = L(\gamma|_{[0,a-\delta]}) + L(\gamma|_{[a-\delta,a]}) = L(\sigma) + L(\gamma|_{[a-\delta,a]}) = L(\eta)$, it is a maximising, timelike geodesic from p to q. We can now apply Lemma 3.3 to the curves η and γ, to obtain $\eta = \gamma$ and hence $\sigma = \gamma|_{[0,a-\delta]}$.

(ii) A similar statement was shown in [43, Cor. 3.5], but we need to adapt some of the main points of the argument to our assumptions.
Choose a sequence of points $q^k_\delta \in \partial I^+_{g_k}(p)$ converging to q_δ. There exist future directed, g_k-null maximising geodesics γ_k ending at q^k_δ and contained in $\partial I^+_{g_k}(p)$. W.l.o.g we may assume their terminal velocities to be h-normalized, so we can apply Corollary 2.6 to obtain a subsequence of γ_k (denoted in the same way) converging to a g-null geodesic σ. Since $\partial I^+_{g_k}(p) \subseteq \overline{I^+(p)}$ (recall that $g_k = \check g_{\varepsilon_k}$) we know that $\sigma \subseteq \partial I^+(p)$, since otherwise $q_\delta \in I^+(p)$, hence σ is maximising (cf. Lemma A.15). Note that γ and σ meet tangentially at q_δ, because otherwise $\eta := \sigma \cup \gamma|_{[a-\delta,a]}$ would not be maximising by Lemma A.4 and so $q \in I^+(\sigma) \subseteq I^+(\overline{I^+(p)}) \subseteq I^+(p)$, a contradiction. Lemma 3.3 now shows that $\gamma = \eta$ and so $\sigma = \gamma$, wherever both curves are defined. There are two possibilities for σ, either it is past inextendible or it reaches p. However, if it is past inextendible, we must

have $\sigma \supseteq \gamma|_{[0,a-\delta]}$. Hence in both cases the γ_k converge to γ in C^1_{loc}. Further this means that there are $t_k \in [0, a-\delta]$ with $\gamma_k(t_k) \to p$, and we have $t_k \to t_0 = 0$ since otherwise $\gamma(t_0) = p = \gamma(0)$, contradicting the causality of M. Now assume M to be strongly causal and $\gamma \subseteq E^+(S)$ for S acausal, then there exists a g-causally convex, g-globally hyperbolic neighbourhood U of p by [32, Lem. 3.21] (this result holds also for C^1-spacetimes). In such neighbourhoods it holds that $I^+_U(S) = I^+(S) \cap U$, $J^+_U(S) = J^+(S) \cap U$ and hence also $E^+_U(S) = \partial I^+_U(S) = \partial I^+(S) \cap U$. Clearly U is also g_k-causally convex and g_k-globally hyperbolic and so these equalities also hold for g_k. Note that in the same fashion as for the case of a single point p one can show that $\gamma_k \to \gamma|_{[0,a-\delta]}$. If γ_k reaches S, we immediately obtain $\gamma_k \subseteq E^+_k(S)$. In the other case $\gamma_k \subseteq \partial I^+_k(S) \setminus J^+_k(S)$ is past inextendible. However, as they converge to $\gamma|_{[0,a-\delta]}$, there must be $p_k \in \gamma_k$, with $p_k \to p$. Hence, w.l.o.g. $p_k \in \partial I^+_k(S) \cap U = E^+_{U,k}(S)$ and so $p_k \in J^+_k(S)$, a contradiction. We have shown that γ_k reach S, and it only remains to show that they do so close to 0. Again let $\gamma_k(t_k) \in S$ with $\gamma_k(t_k) \to p$. If $t_k \to t_0$, then $\gamma(t_0) \in S$ and by acausality of S we must have $t_0 = 0$. $\square$

Remark 3.5. (i) As the proof of Proposition 3.4 shows, for (i) it suffices to assume (M, g) to be MTNB, while for (ii) MNNB would suffice.

(ii) We will make use of Proposition 3.4(ii) in two different scenarios: Thm. 4.3 and Proposition 5.5 below. In the first case (M causal) all we need to ensure is that γ_l are maximizing for long enough, i.e. $a - \delta - t_l$ is close to a. In the second (M strongly causal) it is important to also obtain that γ_l reaches S, in order to be able to use the initial conditions assumed on S.

4. No Lines

In this section we will prove that, under suitable causality and energy conditions, complete causal geodesics stop being maximising also in (MCNB) C^1-spacetimes. To this end we first prove the existence of conjugate points for causal geodesics in smooth spacetimes under weakened versions of the energy conditions. More precisely we will invoke the energy conditions derived for regularisations in Lemma 2.10 and in Lemma 2.14 from the distributional energy conditions, as well as in Lemma 2.18 from the distributional genericity condition. The following is a strengthened version of [8, Prop. 4.2]:

Lemma 4.1. *Let (M, g) be a smooth spacetime. Given some $c > 0$ and $0 < r < \frac{\pi}{4\sqrt{c}}$, there exist $\delta(c, r) > 0$ and $T(c, r) > 0$ such that any causal geodesic γ defined on $[-T, T]$ for which the following hold*

(i) *$Ric(\dot\gamma, \dot\gamma) \geq -\delta$ on $[-T, T]$ and*
(ii) *there exists a smooth parallel orthonormal frame for $[\dot\gamma]^\perp$ and some $C > 0$ such that w.r.t. this frame the tidal force operator satisfies $[R](t) > \mathrm{diag}(c, -C, \ldots, -C)$ on $[-r, r]$,*

possesses a pair of conjugate points on $[-T, T]$.

Proof. This actually follows by a more careful choice of constants in the proof of [8, Prop. 4.2]. To see this we briefly recall the main steps of that proof. The argument proceeds indirectly, assuming that for any $\delta > 0$ and $T > 0$ there is some γ satisfying (i) and (ii) without conjugate points in $[-T, T]$. Denote by $[A]$ the unique Jacobi tensor class along γ with $[A](-T) = 0$ and $[A](0) = \mathrm{id}$. Taking $[E_1](t), \ldots, [E_d](t)$ as in (ii), linear endomorphisms of $[\dot{\gamma}]^{\perp}$ are written as matrices in this basis, and we set $[\tilde{R}](t) := \mathrm{diag}(c, -C, \ldots, -C)$ with C as in (ii). Then by (ii), $[\tilde{R}](t) < [R](t)$ on $[-r, r]$.

The self-adjoint operator $[B] := [\dot{A}] \cdot [A]^{-1}$ satisfies the matrix Riccati equation

$$[\dot{B}] + [B]^2 + [R] = 0, \tag{4.1}$$

and we denote by $[\tilde{B}]$ the solution to (4.1), with $[\tilde{R}]$ instead of $[R]$ and initial value prescribed at some $t_1 \in [-r, r]$. We may even choose t_1 in $[-r, 0]$ and $[\tilde{B}](t_1) := \tilde{\beta}(t_1) \cdot \mathrm{id}$, where $\tilde{\beta}(t_1)$ is greater or equal than the largest eigenvalue of $[B](t_1)$. More precisely, examining the Raychaudhuri equation

$$\dot{\theta} + \frac{1}{d}\theta^2 + \mathrm{tr}(\sigma^2) + \mathrm{tr}([R]) = 0, \tag{4.2}$$

for the expansion $\theta = \mathrm{tr}([B])$ (with $\sigma = [B] - \frac{1}{d}\theta \cdot \mathrm{id}$), it follows that one can pick $\tilde{\beta}(t_1) := f(\nu, \delta, r)$ and $[\tilde{B}](t_1) := f(\nu, \delta, r) \cdot \mathrm{id}$ to indeed achieve that $[B](t) \le [\tilde{B}](t)$ on $[t_1, r]$. Here, $f(\nu, \delta, r) = \sqrt{\frac{2\nu}{r} + \delta} + \frac{\nu}{d}$ with $\nu = 4d/T$.

Moreover, due to the fact that both $[\tilde{R}]$ and $[\tilde{B}](t_1)$ are diagonal, the Riccati equation for $[\tilde{B}]$ decouples. Indeed it can be explicitly solved by

$$[\tilde{B}](t) = \frac{1}{d}\mathrm{diag}(H_{c,f}(t), H_{-C,f}(t), \ldots, H_{-C,f}(t)),$$

where

$$H_{c,f}(t) = d\sqrt{c}\cot(\sqrt{c}(t - t_1) + \mathrm{arccot}(f/\sqrt{c})),$$

and

$$H_{-C,f}(t) = d\sqrt{C}\tanh\left(\sqrt{C}(t - t_1) + \mathrm{artanh}(f/\sqrt{C})\right).$$

From this point on, the proof of [8, Prop. 4.2] proceeds by exclusively analysing the function $H_{c,f}$, and the only requirement on r turns out to be $4r\sqrt{c} < \pi$. Once this is granted, δ and T can be chosen depending only on c to arrive at the desired contradiction. $\square$

Recall that a line is an inextendible causal geodesic maximising the Lorentzian distance between any of its points.

Theorem 4.2. (No timelike lines). *Let (M, g) be a globally hyperbolic, MTNB C^1-spacetime satisfying the distributional timelike energy condition and the distributional genericity condition along any inextendible timelike geodesic. Then there is no complete timelike line in M.*

Proof. Suppose $\gamma : \mathbb{R} \to M$ is a complete timelike line. We may assume that distributional genericity holds at $t_0 = 0$ along γ. We approximate g by $\check{g}_\varepsilon \prec g$, which are hence also globally hyperbolic. By Lemma 2.18 there exist $c > 0$, $0 < r < \frac{\pi}{4\sqrt{c}}$ with the following property: For any C^1-approximation γ_ε of γ by $\check{g}_\varepsilon$-timelike geodesics, there exists $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$, there is $C = C(\varepsilon) > 0$ so that (ii) in Lemma 4.1 is satisfied for R_ε. Now for the above choice of c and r, pick $\delta > 0$ and $T > 0$ as in Lemma 4.1, and let $\tilde{T} > T$. By Proposition 3.4(i), for the curve $\gamma|_{[-T,\tilde{T}]}$

and $\delta < \tilde{T} - T$, we obtain a subsequence $\check{g}_{\varepsilon_k}$ and $\check{g}_{\varepsilon_k}$-maximal timelike geodesics γ_k from $\gamma(-T)$ to $\gamma(\tilde{T})$ converging to $\gamma|_{[-T,\tilde{T}]}$ in C^1.

Let K be a compact neighbourhood of $\gamma([-T, T])$. Since $\gamma_{\varepsilon_k} \to \gamma$ in $C^1([-T, T])$, there are $k_0 \in \mathbb{N}$, $\tilde{C} > 0$ and $\kappa < 0$ such that for all $k \geq k_0$, we have $\gamma_{\varepsilon_k}([-T, T]) \subseteq K$, $\|\dot{\gamma}_{\varepsilon_k}\|_h \leq \tilde{C}$ and $\check{g}_{\varepsilon_k}(\dot{\gamma}_{\varepsilon_k}, \dot{\gamma}_{\varepsilon_k}) < \kappa$ on $[-T, T]$. Hence by [7, Lem. 4.6] (and the remark preceding it) we have $\mathrm{Ric}_{\varepsilon_k}(\dot{\gamma}_{\varepsilon_k}, \dot{\gamma}_{\varepsilon_k}) \geq -\delta$ on $[-T, T]$ for large k. Therefore, also (i) in Lemma 4.1 is satisfied for γ_{ε_k}, yielding the existence of conjugate points for any γ_{ε_k} (k large) in $[-T, T]$, a contradiction since they are maximising even in the strictly larger interval $[-T, \tilde{T}]$. □

Theorem 4.3. *(No null lines) Let (M, g) be a causal, MNNB C^1-spacetime that satisfies the distributional null energy condition and the distributional genericity condition along any inextendible null geodesic. Then there is no complete null line in M.*

Proof. Suppose $\gamma : \mathbb{R} \to M$ is a complete null line. We will again make use of the approximation $\check{g}_\varepsilon$. Suppose without loss of generality that the distributional genericity condition holds along γ at $t_0 = 0$. By Lemma 2.18 there exist $c > 0$ and $0 < r < \frac{\pi}{4\sqrt{c}}$ satisfying: For any C^1-approximation γ_ε of γ by $\check{g}_\varepsilon$-null geodesics, there exists $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$, there is $C = C(\varepsilon) > 0$ so that (ii) in Lemma 4.1 is satisfied for R_ε. Choose $\delta > 0$ and $T > 0$ as in Lemma 4.1 for this pair (c, r) and let $\tilde{T} > T$. Using Proposition 3.4(ii) for the curve $\gamma|_{[-T,\tilde{T}]}$ and $S = \gamma(-T)$ (cf. Remark 3.5(ii)), there exists a subsequence $\check{g}_{\varepsilon_k}$ and $\check{g}_{\varepsilon_k}$-maximal null geodesics $\gamma_k : [-T, \tilde{T}] \to M$ converging to $\gamma|_{[-T,\tilde{T}]}$ in C^1.

Since $\gamma_k \to \gamma$ in $C^1([-T, T])$, once we choose a compact neighbourhood K of $\gamma([-T, T])$, there are $k_0 \in \mathbb{N}$ and $\tilde{C}_2 > \tilde{C}_1 > 0$ such that for all $k \geq k_0$, we have $\gamma_{\varepsilon_k}([-T, T]) \subseteq K$, $\tilde{C}_1 < \|\dot{\gamma}_k\|_h < \tilde{C}_2$ in $[-T, T]$. Hence, Lemma 2.14 implies that $\mathrm{Ric}_{\varepsilon_k}(\dot{\gamma}_k, \dot{\gamma}_k) > -\delta$ in $[-T, T]$. But then Lemma 4.1 may be invoked to give the existence of conjugate points on $\gamma_k|_{[-T,T]}$ for large k, a contradiction since they were supposed to be maximising even on $[-T, \tilde{T}]$. □

5. Initial Conditions

The classical Hawking-Penrose theorem for spacetimes of dimension 4 assumes three alternative initial conditions: the existence of a compact, spacelike hypersurface, a trapped (2-)surface, or a "trapped point", i.e., a point from where all future (or past) null geodesics converge. The second condition was later generalised in [10] to trapped submanifolds of arbitrary codimension m with $1 < m < n = \dim(M)$.

While the first condition does not need any special attention here we will start by generalising the trapped submanifold-case to C^1-spacetimes. As for $C^{1,1}$-spacetimes (see [8, Sec. 6.2]), trapped submanifolds are defined in the support sense. However, to show that normal null geodesics emanating from them stop being maximising is more delicate now due to the lack of an exponential map, cf. (the proof of) Proposition 5.5. At the end of this section we will deal with the case of a "trapped point".

Definition 5.1 (*Support submanifolds*). Let (M, g) be a C^1-spacetime and let $S, \tilde{S} \subseteq M$ be submanifolds. We say that $\tilde{S}$ is a *future support submanifold for S at $q \in S$* if $\dim S = \dim \tilde{S}$, $q \in \tilde{S}$, and there is a neighbourhood U of q in M such that $\tilde{S} \cap U \subseteq J^+_U(S)$.

Definition 5.2 (Trapped C^0-submanifolds). Let (M, g) be a C^1-spacetime of dimension n. We say that a C^0-submanifold $S \subseteq M$ of codimension $1 \leq m < n$ is a *future trapped submanifold* if it is compact without boundary and for any $p \in S$ there exists a neighbourhood U of p such that $U \cap S$ is achronal in U and moreover S has past pointing timelike mean curvature in the sense of support submanifolds, i.e. for any $q \in S$ there exists a future C^2-support submanifold $\tilde{S}$ for S at q whose mean curvature vector at q is past-pointing timelike.

Now, to show that lightrays from trapped submanifolds stop maximising, we reduce the problem to a question about smooth approximating metrics, which are understood much better. The essential results that deal with the corresponding situations for smooth metrics are [8, Lem. 6.4] and [7, Lem. 5.6]. To give a precise formulation we first introduce some notation.

Suppose $S \subseteq M$ is a spacelike C^2-submanifold of codimension $1 < m < n$, $p \in S$ and $\nu \in T_pM$ is a future null normal to S. Let γ be a geodesic with $\gamma(0) = p$, $\dot{\gamma}(0) = \nu$. Let $e_1, \ldots, e_{n-m}$ be an ON-basis on S around p. Further let $E_1, \ldots, E_{n-m}$ be the parallel translates of $e_1(p), \ldots, e_{n-m}(p)$ along γ, which are C^1 since they satisfy the parallel transport equation. We will use this notation for the following results. It will be clear from the context which surface the E_i refer to.

Lemma 5.3. *Let (M, g) be a smooth spacetime and let S be a codimension-m ($1 < m < n$) spacelike C^2 submanifold of M. Let γ be a geodesic starting in S such that $\nu := \dot{\gamma}(0) \in TM|_S$ is a future null normal to S. Suppose that $c := \mathbf{k}_S(\nu) > 0$ and let $b > 1/c$. Then there is $\delta = \delta(b, c) > 0$ such that, if for E_i as described above*

$$\sum_{i=1}^{n-m} g(R(E_i, \dot{\gamma})\dot{\gamma}, E_i) \geq -\delta$$

along γ, then $\gamma|_{[0,b]}$ is not maximising from S, provided that γ exists up to $t = b$.

Lemma 5.4. *Let (M, g) be a smooth spacetime and let S be a codimension-2 spacelike C^2 submanifold of M. Let γ be a geodesic starting in S such that $\nu := \dot{\gamma}(0)$ is a future null normal to S. Suppose that $c := \mathbf{k}_S(\nu) > 0$ and let $b > 1/c$. Then there is $\delta = \delta(b, c) > 0$ such that, if*

$$\mathrm{Ric}(\dot{\gamma}, \dot{\gamma}) \geq -\delta$$

along γ, then $\gamma|_{[0,b]}$ is not maximising from S, provided that γ exists up to $t = b$.

The following result is the C^1-analogue of [8, Prop. 6.5]. As was the case for the distributional genericity condition, cf. Definition 2.15, also here we have to assume that the distributional curvature condition at hand is stable under C^0-perturbations of the vector fields involved. This ensures that we can derive the necessary curvature conditions for the smooth approximating metrics in order to be able to use Lemmas 5.3 and 5.4 .

Proposition 5.5 *(Light rays from a submanifold). Let (M, g) be a strongly causal, MNNB C^1-spacetime and let $\tilde{S} \subseteq M$ be a C^2-spacelike submanifold of codimension $1 < m < n$. Suppose $\mathbf{k}_{\tilde{S}}(\nu) = g(\nu, H_p) > c > 0$ and let $b > 1/c$. Suppose there is a null geodesic γ with $\dot{\gamma}(0) = \nu$, a neighbourhood U of $\gamma|_{[0,b]}$ and C^1-extensions $\overline{E}_i$ of E_i and $\overline{N}$ of $\dot{\gamma}$ to U such that for each $\delta > 0$ there exists $\eta > 0$ such that for all collections of C^1-vector* *fields* $\{\tilde{E}_1, \ldots, \tilde{E}_{n-m}, \tilde{N}\}$ *on* U *with* $\|\tilde{E}_i - \overline{E}_i\|_h < \eta$ *for all* i *and* $\|\tilde{N} - \overline{N}\|_h < \eta$, *we have*

$$\sum_{i=1}^{n-m} g(R(\tilde{E}_i, \tilde{N})\tilde{N}, \tilde{E}_i) \geq -\delta \quad in\, \mathcal{D}'^{(1)}(U). \tag{5.1}$$

Then $\gamma|_{[0,b]}$ *is not maximising from* $\tilde{S}$.

Note that condition (5.1) localises in a similar manner as the genericity condition. For the latter see Remark 2.16.

Proof. Let $g_\varepsilon = \check{g}_\varepsilon$, $R_\varepsilon = R[\check{g}_\varepsilon]$. Clearly, $\mathbf{k}_{\tilde{S}}$ is continuous and $\mathbf{k}_{\tilde{S},\varepsilon} \to \mathbf{k}_{\tilde{S}}$ uniformly on compact sets. Hence, there is a neighbourhood V of ν in $TM|_{\tilde{S}}$ and ε_0 such that $\forall \varepsilon \leq \varepsilon_0$ and $\forall v \in V$: $\mathbf{k}_{\tilde{S},\varepsilon}(v) > c$. Note that g-spacelikeness implies g_ε-spacelikeness. Hence $U \cap \tilde{S}$ is g_ε-spacelike and we may assume that $W := \pi(V)$ is contained in $U \cap \tilde{S}$.

Suppose now to the contrary that $\gamma|_{[0,b]}$ maximises the distance from $\overline{U} \cap \tilde{S}$. Let $1/c < b' < b'' < b$ and let $q = \gamma(b'')$. Find $q_k \in \partial J^+_k(\overline{U} \cap \tilde{S})$ (where $\partial J_k := \partial J_{g_{\varepsilon_k}}$ and $\varepsilon_k \to 0$) with $q_k \to q$. By Corollary A.12 there are g_{ε_k}-null geodesics $\gamma_k : I_k \to M$ (in future directed parametrisation) with $\gamma_k(b'') = q_k$ either intersecting $\overline{U} \cap \tilde{S}$ or past inextendible. We may assume that $\|\dot{\gamma}_k(b'')\|_h$ are all equal to $\|\dot{\gamma}(b'')\|_h$, hence we may assume that the $\dot{\gamma}_k(b'')$ converge to a g-null vector v and the geodesics converge to a g-null geodesic γ_v in C^2_{loc} by Corollary 2.6. Strong causality allows us to invoke Lemma A.30, so we can find a neighbourhood W of $\gamma(0)$ with $W \subseteq U$ such that $W \cap \tilde{S}$ is acausal in M. By (the proof of) Proposition 3.4(ii) (see also Remark 3.5(i)) we obtain that $\gamma = \gamma_v$, and that (up to picking a subsequence) γ_k reaches $\tilde{S}$ for k large at $t_k \downarrow 0$.

Since any of the g_{ε_k}-geodesics γ_k reaches $\tilde{S}$ for all $\varepsilon_k < \varepsilon_0$ (for some suitable $\varepsilon_0 > 0$) at points $\gamma_k(t_k) \in \tilde{S}$, by C^1_{loc} convergence of γ_k to γ we also have $\dot{\gamma}_k(t_k) \in V$. Note that, due to $\gamma_k(t_k) \to \gamma(0)$, we can pick g_{ε_k}-orthonormal bases for $T_{\gamma_k(t_k)}\tilde{S}$ converging to a g-orthonormal basis for $T_{\gamma(0)}\tilde{S}$. Denote by $E_i^{\varepsilon_k}$ and E_i the g_{ε_k}- and g-parallel transports of these bases along γ_k and γ, respectively. By Lemma B.2 there exist C^1-extensions $\tilde{E}_i^{\varepsilon_k}$ and $\tilde{E}_i$ to the (possibly shrunk) neighbourhood U. By a similar argument one can extend the velocity vector fields along γ_k and γ to all of U, they will be denoted by $\tilde{N}^\varepsilon$ and $\tilde{N}$, respectively.

Assume also without loss of generality that U is relatively compact and that $\gamma_k([t_k, b]) \subseteq U$ for large k. Pick $\delta = \delta(b', c)$ according to Lemma 5.3. Since they restrict to the same vector fields along γ, by shrinking U and by continuity of $\overline{E_i}$ and $\tilde{E}_i$ resp. $\overline{N}$ and $\tilde{N}$, we can assume these vector fields to be arbitrarily close in $\|\,.\,\|_h$ on U. Moreover, by the convergence of $\tilde{E}_i^{\varepsilon_k}$ and $\tilde{N}^{\varepsilon_k}$ to $\tilde{E}_i$ and $\tilde{N}$, respectively, also these vector fields can be made arbitrarily $\|\,.\,\|_h$-close to $\overline{E_i}$ and $\overline{N}$, respectively, on U for large k. The assumption of the Proposition therefore implies that

$$\sum_{i=1}^{n-m} g(R(\tilde{E}_i^{\varepsilon_k}, \tilde{N}^{\varepsilon_k})\tilde{N}^{\varepsilon_k}, \tilde{E}_i^{\varepsilon_k}) \geq -\delta/2$$

in $\mathcal{D}'(U)$ for k large. Since $\star_M \rho_{\varepsilon_k}$ respects inequalities, we conclude that also

$$\sum_{i=1}^{n-m} g(R(\tilde{E}_i^{\varepsilon_k}, \tilde{N}^{\varepsilon_k})\tilde{N}^{\varepsilon_k}, \tilde{E}_i^{\varepsilon_k}) \star_M \rho_{\varepsilon_k} \geq -\delta/2$$

on U for large k.

By the same reasoning as in Lemma 2.17 we obtain (once again writing R_{ε_k} for $R[g_{\varepsilon_k}]$)

$$g(R(\tilde{E}_i^{\varepsilon_k}, \tilde{N}^{\varepsilon_k})\tilde{N}^{\varepsilon_k}, \tilde{E}_i^{\varepsilon_k}) \star_M \rho_{\varepsilon_k} - g_{\varepsilon_k}(R_{\varepsilon_k}(\tilde{E}_i^{\varepsilon_k}, \tilde{N}^{\varepsilon_k})\tilde{N}^{\varepsilon_k}, \tilde{E}_i^{\varepsilon_k}) \to 0$$

uniformly on U as $\varepsilon_k \to 0$, $i = 1, \dots, n-m$. But then, for large k,

$$\sum_{i=1}^{n-m} g_{\varepsilon_k}(R_{\varepsilon_k}(\tilde{E}_i^{\varepsilon_k}, \tilde{N}^{\varepsilon_k})\tilde{N}^{\varepsilon_k}, \tilde{E}_i^{\varepsilon_k}) \geq -\delta.$$

In particular, the above holds along γ_k, so

$$\sum_{i=1}^{n-m} g_{\varepsilon_k}(R_{\varepsilon_k}(E_i^{\varepsilon_k}(t), \dot{\gamma}_k(t))\dot{\gamma}_k(t), E_i^{\varepsilon_k}(t)) \geq -\delta$$

for all $t \in [t_k, b']$, where we may assume that k is so large that all γ_k are defined on $[t_k, b']$.

Due to our choice of δ, Lemma 5.3 implies that, for k sufficiently large, each γ_k stops maximising at parameter $b' + t_k$ at the latest (if γ_k is not g_ε-normal to $\tilde{S}$, then γ_k stops maximising the distance immediately, see Remark 5.9 below). By construction the γ_k maximize from t_k to b'', hence if k is so large that $b' + t_k < b''$ we obtain a contradiction. □

Remark 5.6. By arguments analogous to those given for the genericity condition, cf. Remark 2.19, namely essentially by boundedness of R on compact sets, one can see that if $g \in C^{1,1}$, then the curvature condition (5.1) in Proposition 5.5 is equivalent to the one given in [8, Prop. 6.5].

Proposition 5.7 *(The case of trapped surfaces). Let (M, g) be a strongly causal, MNNB C^1-spacetime satisfying the distributional null energy condition. Let $\tilde{S} \subseteq M$ be a codimension-2 C^2-spacelike submanifold such that $\mathbf{k}_{\tilde{S}}(\nu) > c > 0$, where ν is a null normal to $\tilde{S}$, and let $b > 1/c$. If γ is a null geodesic with $\dot{\gamma}(0) = \nu$, then $\gamma|_{[0,b]}$ is not maximising from $\tilde{S}$.*

Proof. The proof is completely analogous to that of Proposition 5.5, with the sole difference that one needs to use Lemma 5.4 instead of Lemma 5.3. Let us give a brief outline: Approximate g again by $g_{\varepsilon_k} \equiv \check{g}_{\varepsilon_k}$ and the compact segment $\gamma|_{[0,b]}$ as before by g_{ε_k}-null geodesics γ_k. Then by Lemma 2.14 (see also the proof of [7, Thm. 5.7]), for all $\delta > 0$ and large k we have $\mathrm{Ric}_{\varepsilon_k}(\dot{\gamma}_k, \dot{\gamma}_k) > -\delta$. Using this, combined with Lemma 5.4 and proceeding as in Proposition 5.5, one arrives at a contradiction as before. □

Remark 5.8. The corresponding $C^{1,1}$-version of Proposition 5.7 was dealt with in [8, Remark 6.6(i)], however, a similar line of reasoning is not available in our case of C^1-metrics: While it is still possible to extend an ON-frame along a curve to a neighbourhood, this does not suffice to derive condition (5.1) of Proposition 5.5, as the needed rigidity of the condition cannot be guaranteed. More precisely: For any ON-frame $\{\bar{E}_i, \bar{N}\}$ on U the equality (as before $\bar{N}$ denotes the extension of γ' in the frame)

$$\sum_{i=1}^{n-m} g(R(\bar{E}_i, \bar{N})\bar{N}, \bar{E}_i) = \mathrm{Ric}[g](\bar{N}, \bar{N}) \geq -\delta$$

still holds and the final inequality is due to the distributional null energy condition. However, to carry out a proof similar to the one of Proposition 5.5, this estimate would need to hold for all collections of vector fields close to the frame in the sense given in the Proposition. This cannot be guaranteed for C^1-metrics anymore as R is, in general, unbounded. So instead, the slightly different focusing result used in Proposition 5.7 is required.

Remark 5.9. If, in the situation of Proposition 5.5, ν is null but not a null normal, then geodesics in that direction immediately enter I^+: This is shown in the same way as in the smooth case, e.g. by using ideas from [39, Lem. 10.45, Lem. 10.50]: Let c be a geodesic in direction ν, hence in particular C^2. By [43, Lem. 3.1] for a (piecewise) C^1-vector field X along c with $g(X', \dot{c}) < 0$, any variation c_s with variation field X is timelike and longer than c (for small s).

Now for any vector v at p not a null normal to a spacelike surface S, w.l.o.g. there exists $y \in T_pS$ with $g(y, v) > 0$. Define $V(t) = (1 - \frac{t}{b})Y(t)$, where Y is the parallel translate of y along c. A straightforward calculation shows $g(V', \dot{c}) < 0$. Now choose a variation c_s of c with variational vector field V such that $c_s(0) \in S$. We obtain that c_s is timelike from S to $c(b)$ and as b was arbitrary the statement follows.

The main result on trapped submanifolds in C^1-spacetimes now is the following:

Proposition 5.10. *Let (M, g) be a strongly causal, MNNB C^1-spacetime and let $S \subseteq M$ be a trapped C^0-submanifold of codimension m, $1 < m < \dim(M) = n$. If $m = 2$, suppose that (M, g) satisfies the distributional null energy condition, and if $m \neq 2$, suppose that any support submanifold $\tilde{S}$ of S satisfies the condition in Proposition 5.5 in any null normal direction ν and for any null geodesic γ with $\dot{\gamma}(0) = \nu$. Then $E^+(S)$ is compact or (M, g) is null geodesically incomplete.*

Proof. Suppose (M, g) is null geodesically complete.

$E^+(S)$ is relatively compact: Assume, to the contrary, that there are $q_k \in E^+(S)$ with $q_k \to \infty$. Lemma A.5 guarantees the existence of future directed maximising null geodesics $\gamma_k : [0, t_k] \to M$ connecting $p_k \in S$ to q_k, which we may assume to be parametrised by h-unit speed. By compactness we may further assume that $\gamma_k(0) = p_k \to p \in S$, and we write $\gamma_k(t_k) = q_k$. Due to $q_k \to \infty$, only finitely many γ_k are contained in a fixed neighbourhood U of p. So we may apply Theorem A.10, implying that (a subsequence, not relabeled, of) γ_k converges to a g-null curve $\gamma : [0, \infty) \to M$ (in h-unit speed parametrisation) from p. Since the γ_k are S-maximising, so is γ. In particular, γ is a g-null pre-geodesic. Reparametrising γ as a geodesic, it is still defined on $[0, \infty)$ due to null geodesic completeness and so $\gamma : [0, \infty) \to M$ is a future complete g-null S-ray. Let $\tilde{S}$ be a future support submanifold of S at p, then γ maximises the distance from $\tilde{S}$ everywhere, because otherwise there would be some T such that $\gamma(T) \in I^+(\tilde{S}) \subseteq I^+(S)$, contradicting the fact that γ is an S-ray. As this cannot happen due to Proposition 5.5 resp. Proposition 5.7, $E^+(S)$ is relatively compact.

$E^+(S)$ is closed: Let $q_k \in E^+(S)$, $q_k \to q$. Let $\gamma_k : [0, t_k] \to M$ be null pre-geodesics (in h-unit speed parametrisation) from some $p_k \in S$ to q_k. By compactness of S, we may assume that $p_k \to p \in S$. If $q = p$, then we are done. Otherwise, there are two possibilities: If the t_k are bounded, then by going over to subsequences, $t_k \to t \in (0, \infty)$. Since $q \neq p$, we are in a position to invoke Theorem A.11 to get a maximising causal limit geodesic $\gamma : [0, t] \to M$ connecting p to q. γ cannot be timelike, as this would imply $q \in I^+(S)$ and hence $q_k \in I^+(S)$ for large k. Hence $q \in J^+(S) \setminus I^+(S) = E^+(S)$. The other possibility is that the t_k are unbounded, i.e. w.l.o.g. $t_k \to \infty$. Since (M, g) is

strongly causal, this means that the γ_k leave every compact set, i.e., $\gamma_k(t_k) \to \infty$ (after possibly relabelling t_k). However, as $E^+(S)$ is contained in some compact set, this is not possible. $\square$

Corollary 5.11. *Let the assumptions be as in Proposition 5.10. Then $E^+(S) \cap S$ is achronal, and $E^+(E^+(S) \cap S)$ is compact or M is null geodesically incomplete.*

Proof. This is shown as in the smooth case, see [10, Prop. 4] or [46, Prop. 4.3] and using that for any $p \in S$ there is some neighbourhood U such that $S \cap U$ is achronal in U. $\square$

We now start with our discussion of "trapped points". In [8, Sec. 6.3], a faithful way to generalise the classical condition, which uses Jacobi tensor classes—a tool no longer available already for $C^{1,1}$-metrics—is presented: It uses the mean curvature of spacelike 2-surfaces given as the level sets of the exponential map that generate the light cone. While the latter tool presently is not at our disposal, the very formulation of the corresponding condition again uses support manifolds and can be adopted verbatim. However, it is now more delicate to derive that the horismos of a trapped point is a trapped set.

Definition 5.12 (*Trapped points*).
A point $p \in M$ is *future trapped* if for any future pointing null vector $v \in T_pM$ and for any null geodesic γ with $\gamma(0) = p$, $\dot\gamma(0) = v$, there exists a parameter t and a spacelike C^2-submanifold of codimension $m = 2$ with $\tilde S \subseteq J^+(p)$, $\gamma(t) \in \tilde S$ and $\mathbf{k}_{\tilde S}(\dot\gamma(t)) > 0$.

Proposition 5.13. *Let (M, g) be a strongly causal, MNNB C^1-spacetime satisfying the distributional null energy condition. If $p \in M$ is a future trapped point and (M, g) is null geodesically complete then $E^+(p)$ is compact.*

Proof. Since trapped points are defined by means of codimension-2 submanifolds $\tilde S$, we will use Proposition 5.7 for which (only) the distributional null energy condition is needed. The proof is analogous to that of Proposition 5.10.

$E^+(p)$ is relatively compact: Suppose $q_j \in E^+(p)$, $q_j \to \infty$. Let $\gamma_j : [0, t_j] \to M$ be maximising null geodesics connecting p to q_j, cf. Lemma A.5. Then, up to a subsequence, the γ_j converge to a g-null ray $\gamma : [0, \infty) \to M$ from p (the domain is $[0, \infty)$ even in g-affine parametrisation since (M, g) is null geodesically complete). By assumption, there is a C^2-spacelike submanifold $\tilde S$ of codimension 2 and a parameter t such that $\mathbf{k}_{\tilde S}(\dot\gamma(t)) > c > 0$ and we let $b > 1/c$. γ is indeed an $\tilde S$-ray: If not, it would enter $I^+(\tilde S) \subseteq I^+(p)$, a contradiction. But by the previous results in this section, γ cannot maximise from $\tilde S$ past b, a contradiction since it is a ray.

$E^+(p)$ is closed: This is shown in exactly the same way as closedness was shown in the proof of Proposition 5.10. $\square$

6. The Main Result

Given the results established so far, the general mechanics of the proof of the Hawking–Penrose theorem remains the same as in the smooth case. There do, however, remain notable differences in the details in this lower regularity (e.g. the use of [7, Prop. 2.13] in the proof of Theorem 6.2), so we include the full argument.

Lemma 6.1. *Let (M, g) be a strongly causal C^1-spacetime such that no inextendible null geodesic in M is maximising. Let $A \subseteq M$ be achronal such that $E^+(A)$ is compact, and let γ be a future inextendible timelike curve contained in $D^+(E^+(A))^\circ$. Then $F := E^+(A) \cap \overline{J^-(\gamma)}$ is achronal and $E^-(F)$ is compact.*

Proof. Due to Lemma A.16 we may assume that $A = \overline{A}$. Since $E^+(A)$ is always achronal and $F \subseteq E^+(A)$, it is also achronal. By compactness of $E^+(A)$, F is compact. Note that $E^-(F) \subseteq F \cup \partial J^-(\gamma)$ by similar arguments as in [21, Lem. 9.3.4], so to show compactness of $E^-(F)$, it suffices to show compactness of $E^-(F) \cap \partial J^-(\gamma)$.

Suppose now that $v \in TM|_F$ is past pointing causal and let $c_v : [0, a) \to M$ be any past inextendible, past directed geodesic with $\dot{c}_v(0) = v$ (recall that in C^1-spacetimes we no longer have unique solvability of the geodesic initial value problem). Then $c_v \subseteq \overline{J^-(\gamma)}$, and $\overline{J^-(\gamma)} = I^-(\gamma) \cup \partial J^-(\gamma)$. We show that c_v meets $I^-(\gamma)$: Suppose, to the contrary, that c_v is a null geodesic entirely contained in $\partial J^-(\gamma)$. Since γ is a future inextendible timelike curve, we have $J^-(\gamma) = I^-(\gamma)$, hence c_v lies entirely in $\partial J^-(\gamma) \setminus J^-(\gamma)$. By Corollary A.12, there is a future directed, future inextendible null geodesic λ starting at $c_v(0)$ entirely in $\partial J^-(\gamma)$ (γ is closed in M by Corollary A.33 and Lemma A.32). This means that either $c_v\lambda$ is an inextendible broken null geodesic, hence not maximising by Lemma A.4, or an inextendible unbroken maximising null geodesic. By assumption such a geodesic cannot exist, hence in either case $c_v\lambda$ cannot lie entirely in $\partial J^-(\gamma)$ and so c_v must enter $I^-(\gamma)$.

Thus we have shown that for any v and any geodesic c_v as above there is $t_v = t_v(v, c_v)$ (for which c_v is still defined) such that $c_v(t_v) \in I^-(\gamma)$. It remains to show that $E^-(F) \cap \partial J^-(\gamma)$ is compact.

$E^-(F) \cap \partial J^-(\gamma)$ *is relatively compact:* Suppose there are $q_k \in E^-(F) \cap \partial J^-(\gamma)$ with $q_k \to \infty$. Then there are past-directed maximising null pre-geodesics $c_k : [0, t_k] \to M$ in h-unit speed parametrisation with $c_k(0) = p_k \in F$ and $c_k(t_k) = q_k$. Since $q_k \to \infty$, also $t_k \to \infty$ By compactness, we may assume $p_k \to p \in F$. Since $q_k \to \infty$, there is a neighbourhood U which almost all c_k leave and we are in a position to apply the Theorem A.10. We get a past-directed inextendible null limit curve $c : [0, \infty) \to M$ from p and a subsequence of the c_k (w.l.o.g. c_k itself) which converges to c locally uniformly. It is easy to see that all c_k have to lie entirely in $\partial J^-(\gamma)$, hence so does the limit c. In particular, c is everywhere maximising and may be reparametrised as an inextendible geodesic $c : [0, a) \to M$ entirely contained in $\partial J^-(\gamma)$. This is a contradiction since we showed before that each past directed null geodesic with initial velocity in $TM|_F$ has to enter $I^-(\gamma)$.

$E^-(F) \cap \partial J^-(\gamma)$ *is closed*: Let $q_j \in E^-(F) \cap \partial J^-(\gamma)$, $q_j \to q$, where we may assume $q \notin F$, otherwise there is nothing to prove. Connect $F \ni p_j \to q_j$ via maximising null geodesics. Up to a subsequence, their limit has to be a maximising null geodesic from p to q, hence $q \in E^-(F) \cap \partial J^-(\gamma)$, because otherwise one would again obtain an inextendible maximising null geodesic entirely in $\partial J^-(\gamma)$, which would lead to the same contradiction as before. □

Now we are ready to give the C^1-version of the causal Hawking-Penrose theorem. It is the direct analog of [17, p. 538, Thm.] and [8, Thm. 7.4], respectively.

Theorem 6.2. *Let (M, g) be a C^1-spacetime. Then the following conditions cannot all hold:*

(i) *(M, g) is causal.*
(ii) *No inextendible timelike geodesic in an open globally hyperbolic subset is everywhere maximising.*
(iii) *No inextendible null geodesic is everywhere maximising.*
(iv) *There is an achronal set $A \subseteq M$ such that $E^+(A)$ or $E^-(A)$ is compact.*

Proof. Suppose all of the above conditions hold. Then by Lemma A.31, (i) and (iii) imply that (M, g) is strongly causal. W.l.o.g. we assume that $E^+(A)$ is compact. By

Lemma A.35 there exists a future inextendible (in M) timelike curve $\gamma \subseteq D^+(E^+(A))^\circ$. Set $F := E^+(A) \cap \overline{J^-(\gamma)}$, which by Lemma 6.1 is achronal and $E^-(F)$ is compact. Once more by Lemma A.35 there is some past-inextendible timelike curve $\lambda \subseteq D^-(E^-(F))^\circ$. Exactly as in the smooth case, see e.g. the second paragraph of [8, Proof of Thm. 7.4], one can show $\gamma \subseteq D^+(E^-(F))^\circ$, by invoking Proposition A.20 and Lemma A.24. So $\gamma, \lambda \subseteq D(E^-(F))^\circ$ which by Proposition A.28 is globally hyperbolic. Now we pick sequences $p_k = \lambda(s_k)$ and $q_k = \gamma(t_k)$ with $t_k, s_k \nearrow \infty$, leaving every compact set, which is possible because γ and λ are inextendible curves in M contained in a strongly causal set (Lemma A.32). Also because $\lambda \subseteq J^-(E^-(F)) \subseteq J^-(F) \subseteq J^-(\overline{J^-(\gamma)})$ and because γ is timelike, we can choose $p_k \in I^-(q_k)$.

By [7, Prop. 2.13] there exist maximising timelike geodesics $\tilde{\gamma}_k$ in $D(E^-(F))^\circ$ from p_k to q_k, which must intersect $E^-(F)$, say at r_k. By compactness of $E^-(F)$ we can assume $r_k \to r$ and by Corollary 2.6 there exists an inextendible, causal limit geodesic $\tilde{\gamma}$, which is also maximising (as a limit of maximisers).

By (iii), $\tilde{\gamma}$ cannot be null, so it must be timelike. As a limit of curves in $D(E^-(F))^\circ$ it is contained in $\overline{D(E^-(F))}$. Again using Proposition A.20 and Lemma A.24 one can now proceed as in the last step in [8, Proof of Thm. 7.4] to arrive at the desired contradiction (to (ii)). $\square$

Finally, we formulate and prove the main result of this work, the analytical Hawking-Penrose theorem for maximally causally non-branching C^1-spacetimes.

Theorem 6.3. *(The Hawking-Penrose singularity theorem for C^1-metrics)*
Let (M, g) be a C^1-spacetime of dimension n with the following properties.

(i) (M, g) *is causal.*
(ii) (M, g) *satisfies the distributional timelike and null energy conditions.*
(iii) (M, g) *satisfies the distributional genericity condition along any inextendible causal geodesic.*
(iv) (M, g) *is maximally causally non-branching.*
(v) (M, g) *contains one of the following.*
  (a) *a compact, achronal, edgeless set.*
  (b) *a future trapped point.*
  (c) *a future trapped C^0-submanifold of codimension* 2.
  (d) *a future trapped C^0-submanifold of codimension* $2 < m < n$ *such that its support submanifolds satisfy the condition in Proposition 5.5.*

Then (M, g) is causally geodesically incomplete.

Proof. Assuming causal geodesic completeness, we would like to deduce a contradiction by using Theorem 6.2.

Note that due to Theorems 4.2 and 4.3 the assumptions (ii) and (iii) of Theorem 6.2 are satisfied and it remains to establish the existence of a future or past trapped set, i.e. an achronal set with compact future or past horismos. This will be implied by (v), but we have to distinguish the different cases:

By Proposition A.36, (a) implies the existence of a trapped set. Note that by Lemma A.31 (which can be applied by Theorem 4.3) (M, g) is strongly causal and hence Proposition 5.13 covers the case (b), whereas Proposition 5.10 covers cases (c) and (d). $\square$

Acknowledgements This work was supported by project P 33594 of the Austrian Science Fund FWF and by the Uni:Docs program of the University of Vienna. We thank Eduard A. Nigsch for helpful discussions.

Funding Open access funding provided by University of Vienna.

A. Appendix: Causality Results in C^1-Spacetimes

In this section, we collect various results on the causality of C^1-Lorentzian metrics that will be needed in the proof of the main result. Unless stated otherwise, throughout this section (M, g) consists of a connected smooth (Hausdorff and second countable) manifold M, and a C^1-Lorentzian metric g that is time-orientable (i.e. there exists a smooth timelike vector field on M). Many of the results in this section hold in even greater generality, namely either for Lorentzian metrics that are merely Lipschitz or even in the general setting of closed cone structures [35]. Although we will occasionally note this, in proving the results in this section we will stay close to smooth causality theory (mainly based on the comprehensive reference work [34]) and only make those changes from classical proofs that are required due to the absence of certain tools (e.g., convex normal neighbourhoods) in the C^1-setting.

We begin with two elementary results, namely the openness of timelike futures and pasts and the push-up property. Both of these hold for causally plain spacetimes, which include spacetimes with Lipschitz continuous metrics [5, Cor. 1.17].

Lemma A.1. *For any $A \subseteq M$, the sets $I^{\pm}(A)$ are open in M.*

Proof. See [5, Prop. 1.21]. □

Lemma A.2. *Let $p, q, r \in M$ such that $p \le q \ll r$ or $p \ll q \le r$. Then $p \ll r$.*

Proof. See [5, Lem. 1.22]. □

Remark A.3. Note that from this one can show as in the smooth case, e.g. [39, Lem. 14.03][6], that the timelike relation is open, i.e., if $p \ll q$ then there are neighbourhoods U_p, U_q of p and q respectively, such that for all $x \in U_p$ and all $y \in U_q$, we have $x \ll y$.

Lemma A.4. *Let (M, g) be a C^1-spacetime and let $\gamma : [a, b] \to M$ be a maximising causal curve. Then γ is a (reparametrisation of a) causal geodesic. In particular, it is a C^2-curve and has a causal character.*

Proof. See [27, Thm. 1.1] or [43, Thm. 3.3]. □

For any point $q \in M$, we denote by $E^+(q) := J^+(q) \setminus I^+(q)$ its *future horismos*. Its *past horismos* $E^-(q)$ is defined analogously.

Lemma A.5. *If $p \in E^+(q)$ then there exists a maximising null geodesic segment from q to p.*

[6] the use of convex sets is not necessary once openness of I^+ in is established

Proof. Let c be a future causal curve from q to p. Since $p \notin I^+(q)$, c is null and maximising from q to p (because of push-up), hence it is (a reparametrisation of) a maximising null geodesic by Lemma A.4. □

Lemma A.6. *For any subset $A \subseteq M$, we have $I^{\pm}(A) = J^{\pm}(A)^{\circ}$ and $\partial I^{\pm}(A) = \partial J^{\pm}(A)$. Moreover, the sets $E^{\pm}(A)$ and $\partial J^{\pm}(A)$ are achronal.*

Proof. This follows by the same proofs as in the smooth case (cf. [34, Thm. 2.27], [39, Cor. 14.27]). □

Definition A.7. (Edge)
Let $A \subseteq M$ be achronal. The *edge* of A, denoted by $\mathrm{edge}(A)$, is the set of all $x \in \overline{A}$ such that for any neighbourhood U of x, there is a timelike curve from $I^-_U(x)$ to $I^+_U(x)$ that does not meet A.

Lemma A.8. *Let (M, g) be a C^1-spacetime. An achronal set $A \subseteq M$ is a topological hypersurface if and only if $A \cap \mathrm{edge}(A) = \emptyset$. Moreover, A is a closed topological hypersurface if and only if $\mathrm{edge}(A) = \emptyset$.*

Proof. The proofs for smooth metrics (cf. [39, Prop. 14.25, Cor. 14.26]) still hold in C^1. □

Lemma A.9. *For any $A \subseteq M$, $\partial J^+(A)$ is a (topologically) closed, achronal topological hypersurface.*

Proof. This follows from the fact that $\partial J^+(A)$ is achronal and $\mathrm{edge}(\partial J^+(A)) = \emptyset$, which is proven in precisely the same way as in the smooth case. □

Recall the notion of a limit maximising sequence of causal curves: A sequence $\gamma_k : [a_k, b_k] \to M$ of future directed causal curves is called *limit maximising* if there is $\varepsilon_k \to 0$ such that

$$L(\gamma_k) \geq d_g(\gamma(a_k), \gamma(b_k)) - \varepsilon_k,$$

where d_g is the time separation function induced by g. In particular a sequence of maximising curves is limit maximising. We shall employ the following version of the limit curve theorem:

Theorem A.10. *(Limit Curve Theorem)*
Let h be a complete Riemannian metric on M and let $\gamma_k : [a_k, b_k] \to M$ be future directed, h-arc length parametrised causal curves with $0 \in [a_k, b_k]$ such that the sequence $\{\gamma_k(0)\}$ has an accumulation point y. Suppose there are $a \leq 0$ and $b \geq 0$ such that $a_k \to a$, $b_k \to b$. If there is a neighbourhood U of y such that almost all γ_k leave U, then a subsequence of γ_k converges h-uniformly on compact sets to a future directed causal Lipschitz curve $\gamma : [a, b] \to M$. γ is future resp. past inextendible iff $b = \infty$ resp. $a = -\infty$. If $\{\gamma_k\}$ is limit maximising, then γ is maximising.

Proof. This follows from [30, Thm. 3.1] and [5, Thm. 1.6] by the same proof as in [8, Thm. A.6]. □

For a much more general version of the limit curve theorem, valid in closed cone structures, we refer to [35, Thm. 2.14]. We shall also require the following variant of the limit curve theorem for a sequence of curves with converging past and future endpoints, cf. [30, Thm. 3.1(2)].

Theorem A.11 *(Limit Curve Theorem: two converging endpoints).* *Let $\gamma_k : [0, b_k] \to M$ be a sequence of future-directed, causal curves in h-arc length parametrisation connecting x_k to z_k, with $x_k \to x$ and $z_k \to z$. Suppose that $b_k \to b \in (0, \infty)$. If there is a neighbourhood U of x such that only a finite number of γ_k are entirely contained in U, then there is a Lipschitz future-directed, causal curve $\gamma : [0, b] \to M$ connecting x to z such that a subsequence of the γ_k converges to γ h-uniformly on compact sets. Moreover, if $\{\gamma_k\}$ is limit maximising, then γ is maximising.*

Proof. This is contained in the statement of [30, Thm. 3.1(2)], whose proof is easily seen to hold for C^1-metrics once Theorem A.10 is established. □

Corollary A.12. *Let $A \subseteq M$. For any point $x \in \partial J^+(A) \setminus \overline{A}$, there is a causal curve $\gamma \subseteq \partial J^+(A)$ with future endpoint x that is either past inextendible and does not meet $\overline{A}$ or has a past endpoint in $\overline{A}$. It is a (reparametrisation of a) maximising null geodesic. If A is closed and $x \notin J^+(A)$, then $\gamma \subseteq \partial J^+(A) \setminus J^+(A)$ and it is past inextendible.*

Proof. Using Theorem A.10 and Lemma A.5, this follows as in [8, Prop. A.7]. □

Lemma A.13. *Let (M, g) be a C^1-spacetime and let $\{g_k\}$ be either the sequence $\{\hat{g}_{\varepsilon_k}\}$ or $\{\check{g}_{\varepsilon_k}\}$ with $\varepsilon_k \downarrow 0$. Let $p, q \in M$ and let γ_k be a g_k-maximising curve from p to q, so γ_k is in particular a g_k-geodesic. Suppose that γ_k converge to a curve γ in C^1_{loc}. If $\{g_k\} = \{\check{g}_{\varepsilon_k}\}$, suppose in addition that (M, g) is globally hyperbolic and $p \ll_g q$. Then γ is a g-maximising geodesic from p to q.*

Proof. The case $\{g_k\} = \{\hat{g}_{\varepsilon_k}\}$ follows immediately from [41, Prop. 6.5], so we will only consider the case $\{g_k\} = \{\check{g}_{\varepsilon_k}\}$ together with the additional assumptions of (M, g) being globally hyperbolic and $p \ll_g q$ in this case. Due to global hyperbolicity, the space $C(p, q)$ of future directed causal curves from p to q (considered up to orientation preserving reparametrisations) is compact with respect to its natural topology (see [41, Thm. 3.2]). Consequently, the h-speeds of curves in $C(p, q)$ are uniformly bounded by some constant $C_1 > 0$, where h is some complete Riemannian metric on M.

Since (M, g) is globally hyperbolic, there is an L_g-maximising g-geodesic $c : [0, 1] \to M$ with $c(0) = p$, $c(1) = q$ [41, Prop. 6.4]. Also, $p \ll_g q$ implies that c is g-timelike. By compactness, there is a constant $C_2 > 0$ such that $g(\dot{c}, \dot{c}) < -C_2$ on $[0, 1]$, and hence $g_k(\dot{c}, \dot{c}) < 0$ for k large, implying that c is g_k-timelike for such k. If we set $\delta_k := d_h(g, g_k)$ (the h-distance of g, g_k as in [5, (1.6)]), then

$$\begin{aligned} L_g(c) = \int_0^1 \sqrt{-g(\dot{c}, \dot{c})}dt &\leq \int_0^1 \sqrt{-g_k(\dot{c}, \dot{c}) + C_1^2 \delta_k dt} \leq L_{g_k}(c) + C_1\sqrt{\delta_k} \\ &\leq d_{g_k}(p, q) + C_1\sqrt{\delta_k} = L_{g_k}(\gamma_k) + C_1\sqrt{\delta_k} \to L_g(\gamma) \qquad (k \to \infty), \end{aligned}$$

so $L_g(\gamma) \geq L_g(c)$, which proves that also γ is maximising from p to q. □

Lemma A.14. *Let $\gamma_k \subseteq \overline{J^+(p)}$ be causal curves converging locally uniformly to a causal curve γ with future endpoint q. If $d_g(p, q) = 0$, then γ is a maximising null geodesic.*

Proof. Clearly $\gamma \subseteq \overline{J^+(p)}$. If γ would meet $I^+(p)$, then there would exist some $r \in I^+(p) \cap J^-(q)$ and using push-up we could conclude that $q \in I^+(p)$, contradicting $d_g(p, q) = 0$. Hence $\gamma \subseteq \partial J^+(p)$, which is achronal and γ is a maximising null geodesic by Lemma A.4. □

Lemma A.15. *Let $g_k = \check{g}_{\varepsilon_k}$ and let $\gamma_k \subseteq \overline{J^+(p)}$ be g_k-null geodesics converging in C^1_{loc} to a g-geodesic γ that ends in q. Assume that $d_g(p, q) = 0$. Then γ is a maximising g-null geodesic.*

Proof. This follows immediately from Lemma A.14 upon noting that any g_k-null geodesic is g causal. $\square$

Lemma A.16. *Let A be achronal. Then so is $\overline{A}$. Moreover, if $E^+(A)$ is closed, then $E^\pm(A) = E^\pm(\overline{A})$.*

Proof. This follows by the same proof as in [8, Lem. A.8]. $\square$

Lemma A.17. *Let (M, g) be a C^1-spacetime. Then every point in M has a neighbourhood base of globally hyperbolic neighbourhoods.*

Proof. Fix any smooth, time-orientable Lorentzian metric $\hat{g}$ with $g \prec \hat{g}$. For $x \in M$, [34, Thm. 2.7] gives a neighbourhood base V_m of neighbourhoods of x that are globally hyperbolic with respect to $\hat{g}|_{V_m}$. Any Cauchy hypersurface for $\hat{g}$ in V_m is also a Cauchy hypersurface for g, so the claim follows from [41, Thm. 5.7]. $\square$

Lemma A.18. *Let S be a spacelike (C^2-) hypersurface and let $p \in S$. Then there exists a neighbourhood V of p in M such that $V \cap S$ is a Cauchy hypersurface in V.*

Proof. The proof of the analogous result [25, Lemma A.25] also works in the C^1-setting. $\square$

Definition A.19. (Cauchy development and Cauchy horizon)
Let A be an achronal set. The *future Cauchy development* of A is the set

$$D^+(A) := \{x \in M : \text{ every past inextendible causal curve through } x \text{ meets } A\}.$$

The *future Cauchy horizon* of A is defined as

$$H^+(A) := \overline{D^+(A)} \setminus I^-(D^+(A)).$$

The *past Cauchy development and past Cauchy horizon* are defined analogously.

The first part of the following result is partly a C^1-analogue of [39, Lem. 14.51], where it is proved using convex normal neighbourhoods, a tool that we do not have at our disposal in the C^1-setting. The use of such neighbourhoods can, however, be avoided, as our arguments below illustrate.

Proposition A.20. *Let A be a closed, achronal set. Then*

$$\overline{D^+(A)} = \{x \in M : \textit{every p.i. t.l. curve through } x \textit{ meets } A\}$$

and

$$\partial D^+(A) = A \cup H^+(A).$$

Proof. Let T be the set on the right hand side in the first equation.
$\overline{D^+(A)} \subseteq T$: Suppose there is a point $p \in \overline{D^+(A)} \setminus T$. Then there is some past inextendible timelike (future directed) curve $\alpha : (a, 0] \to M$ with $\alpha(0) = p$ not meeting A. Let $p_k \in D^+(A)$, $p_k \to p$ and let U_k be open neighbourhoods of p with $U_k \to \{p\}$. By choosing subsequences, we may assume that $p_k \in I^+_{U_k}(\alpha(-1/k))$ and $\alpha|_{[-1/k,0]} \subseteq U_k$ for all k.
Construct now past inextendible causal curves α_k by connecting $\alpha(-1/k)$ to p_k in a timelike way such that these curves stay in U_k, and let the rest of α_k (i.e. the part to the past of $\alpha(-1/k)$) be the original curve α. Then the α_k are all past inextendible timelike, and since $p_k \in D^+(A)$ they must meet A at some a_k. Since α does not meet A, the

a_k have to lie on the replaced piece. But these pieces are entirely in U_k, which go to $\{p\}$, so we see that $a_k \to p$. Since A is closed, we have $p \in A$, which is absurd since $p \notin T \supseteq A$ by assumption.

$\overline{D^+(A)} \supseteq T$: Suppose $q \notin \overline{D^+(A)}$, and let $r \in I^-_{M \setminus \overline{D^+(A)}}(q)$. Since in particular $r \notin D^+(A)$, there is a past inextendible causal curve α from r missing A. Precisely as in the smooth case [39, Lem. 14.30(1)], noting that pushup still holds in C^1), there is a past inextendible timelike curve from q not meeting A. Thus $q \notin T$.

Finally, $\partial D^+(A) = A \cup H^+(A)$ follows exactly as in the smooth case (cf. [39, Lem. 14.52]). □

We note that the previous result in fact remains true even for locally Lipschitz proper cone structures [35, Thm. 2.35, 2.36].

Lemma A.21. *Let A be achronal. Then $H^+(A)$ is a closed, achronal set.*

Proof. The proof of [34, Prop. 3.15] carries over unchanged to C^1 metrics. □

Lemma A.22. *Let A be closed and achronal. If $x \in D^+(A) \setminus H^+(A)$, then every past inextendible causal curve through x meets $I^-(A)$. Moreover, if $x \in D(A)^\circ$, then every future (resp. past) inextendible future (resp. past) directed causal curve emanating from x intersects $I^+(A)$ (resp. $I^-(A)$).*

Proof. This follows from [34, Prop. 3.27] and [34, Prop. 3.42], which still hold for C^1-metrics. □

Lemma A.23. *Let A be closed, achronal. Let $x \in J^+(A) \setminus D^+(A)$ or $x \in I^+(A) \setminus D^+(A)^\circ$. Then every causal curve from x to A meets $H^+(A)$.*

Proof. The proof for $C^{1,1}$-metrics [8, Lem. A.12] still holds in C^1 because of Lemma A.1, Lemma A.2, and Proposition A.20. □

Using these results, one establishes the following formula for the timelike future of the Cauchy horizon of a closed, achronal set precisely as in the $C^{1,1}$-case, cf. [8, Lem. A.13].

Lemma A.24. *Let A be closed and achronal. Then $I^+(H^+(A)) = I^+(A) \setminus \overline{D^+(A)}$.*

Lemma A.25. *Let A be closed and achronal. Then* $\mathrm{edge}(H^+(A)) \subseteq \mathrm{edge}(A)$.

Proof. Based on Lemmas A.23 and A.24 , the proof of [8, Lem. A.14] carries over to the C^1-setting. □

Lemma A.26. *Let A be an achronal set, then $H^+(\partial J^+(A))$ is a closed, achronal, topological hypersurface.*

Proof. $H^+(\partial J^+(A))$ is closed and achronal by Lemma A.21, and by the previous result, combined with Lemmas A.8 and A.9 , $\mathrm{edge}(H^+(\partial J^+(A))) \subseteq \mathrm{edge}(\partial J^+(A)) = \emptyset$, so the claim follows from Lemma A.8. □

Lemma A.27. *Let A be closed, achronal. Then $H^+(\overline{E^+(A)}) \subseteq H^+(\partial J^+(A))$.*

Proof. The proof of [8, Lem. A.16] carries over, using Corollary A.12, Proposition A.20 and Lemmas A.23, A.24. □

Proposition A.28. *Let (M, g) be a C^1-spacetime and let $A \subseteq M$ be a closed, achronal set. Then the interior of its Cauchy development, i.e. the set $D(A)^\circ = (D^+(A) \cup D^-(A))^\circ$, if nonempty, is a globally hyperbolic C^1-spacetime when considered with the induced metric.*

Proof. Based on Theorem A.10 and Lemma A.22, this follows from the same proof as in [34, Thm. 3.45]. □

Definition A.29. (Strong causality)
A C^1-spacetime (M, g) is called *strongly causal at* $p \in M$ if for any neighbourhood U of p there is a neighbourhood V of p with $V \subseteq U$ such that any causal curve starting and ending in V is contained in U. (M, g) is called *strongly causal* if it is strongly causal at every point.

Lemma A.30. *(M, g) is strongly causal at p if and only if for any neighbourhood U of p there is a neighbourhood V of p with $V \subseteq U$ such that no causal curve leaving V ever returns.*

Proof. This follows from [33, Lem. 3.21], which continues to hold in the C^1-case. □

Lemma A.31. *Let (M, g) be chronological and suppose there are no inextendible maximising null geodesics in M. Then (M, g) is strongly causal.*

Proof. Using Theorem A.10, this follows by the same argument as [8, Lem. A.19]. □

Lemma A.32. *A strongly causal C^1-spacetime (M, g) is non-totally and non-partially imprisoning: No future or past inextendible causal curve is contained in a compact set and no future or past inextendible causal curve returns to a compact set infinitely often.*

Proof. The classical proof for the smooth case (see [39, Lem. 14.13]) holds even for continuous metrics. □

Corollary A.33. *Let (M, g) be a non-totally and non-partially imprisoning C^1-spacetime. Then the image of every inextendible causal curve is closed in M.*

Proof. Let γ be any inextendible causal curve. Since M is a manifold, the topology on M is compactly generated. Let $K \subseteq M$ be any compact subset. By non-total and non-partial imprisonment, $\gamma \cap K$ is a finite union of closed segments of γ, hence $\gamma \cap K$ is closed in K. Since K was arbitrary, γ is closed in M. □

Lemma A.34. *Let (M, g) be a strongly causal C^1-spacetime and let $A \subseteq M$ be a closed, achronal subset. If $H^+(\overline{E^+(A)})$ is nonempty, then it is noncompact.*

Proof. The proof of the smooth case goes through, because all the necessary preliminary results continue to hold in C^1 (see the outline in [8, Lem. 7.1]). □

Lemma A.35. *Let (M, g) be a strongly causal C^1-spacetime. Let $A \subseteq M$ be an achronal subset such that $E^+(A)$ is compact. Then there is a future inextendible timelike curve γ contained in $D^+(E^+(A))^\circ$.*

Proof. The proof of the smooth case goes through in C^1, see e.g. [17, Lem. 2.12], or also [21, Lem. 9.3.3], combined with Lemma A.16. □

Proposition A.36. *Let (M, g) be a C^1-spacetime and let $A \subseteq M$ be achronal and edgeless. Then $E^+(A) = A$, in particular if A is compact so is $E^+(A)$.*

Proof. This was shown e.g. in [34, Cor. 2.145] and the proof carries over to C^1 spacetimes. □

B. Appendix: Extending Vector Fields in C^1-Spacetimes Uniformly

When applying the genericity condition (Definition 2.15) we need to extend vectors or vector fields along curves to neighbourhoods, due to the fact that distributional (or also L^∞) curvature quantities are not well defined on points or along curves (as these are sets of measure zero). It is rather straightforward to extend a vector field along a curve to a neighbourhood. For C^1-metrics one can still use parallel transport along certain curves spanning an open set. Alternatively, one can also use "cylindrically constant" (see below) extensions in coordinates around a (part of a) curve.

In our case, we need to simultaneously extend vector fields along a sequence of converging curves. To be more precise, let γ_k be a sequence converging in C^1 to some γ and let V_k be vector fields along γ_k and V along γ such that $V_k \to V$ in C^1 as $k \to \infty$. Then given a point on γ, does it possess a neighbourhood U such that all V_k can be extended to vector fields on U and such that their extensions converge in C^1 to the extension of V?

To begin with, we consider the situation of extending vector fields around a point of a curve to a neighbourhood by cylindrically constant extension in coordinates.

Lemma B.1 *(Cylindrically constant extension of vector fields on curves). Let $\gamma : I \to M$ be a C^1-curve which is regular at 0, i.e. $\dot\gamma(0) \neq 0$ and set $p := \gamma(0)$. Then there exists a chart $(U, (x^1, \dots, x^n))$ around p with the following property: For any C^1 vector field V along γ there exists a C^1 extension $\tilde V$ of V to U such that, in these coordinates, $\tilde V$ is independent of $x^2, \dots, x^n$.*

Proof. Since the claim is local we may assume that $M = \mathbb{R}^n$, that $p = 0$, and that $\gamma = (\gamma^1, \dots, \gamma^n)$ with $\gamma_1'(t) > c > 0$ for some c and all $t \in (a, b)$, where $a < 0 < b$. Thus γ^1 is a C^1-diffeomorphism from (a, b) onto some interval J. Then we set $U := J \times B_R^{n-1}(0)$, where $R > 0$ is such that $\gamma((a, b)) \subseteq U$. Given a C^1-vector field V along γ, i.e., a C^1-map $V : (a, b) \ni t \mapsto (\gamma(t), v(t))$ with $v \in C^1((a, b), \mathbb{R}^n)$, set $\tilde V : U \to U \times \mathbb{R}^n$,

$$\tilde V(x) := (x, v((\gamma^1)^{-1}(x_1)))$$

to obtain the desired extension. □

Now that we have set up a way of extending vector fields in a cylindrically constant fashion, we can implement an analogous procedure uniformly on a sequence of curves.

Lemma B.2. *Let $\gamma_k : [-1, 1] \to M$ be a sequence of C^1-curves converging in $C^1([-1, 1])$ to the regular C^1-curve $\gamma : [-1, 1] \to M$ and let V_k and V be C^1 vector fields along γ_k and γ, respectively. Further, let $V_k \to V$ in C^1, i.e., both $V_k \to V$ and $V_k' \to V'$ in TM, uniformly on $[-1, 1]$. Then there exists some open neighbourhood U of $\gamma(0)$ and C^1 extensions $\tilde V_k$ of V_k and $\tilde V$ of V to U such that $\tilde V_k \to \tilde V$ in $C^1_{\mathrm{loc}}(U)$.*

Proof. Again we may assume that $M = \mathbb{R}^n$. As in the previous Lemma we write $\gamma = (\gamma^1, \dots, \gamma^n)$ and can assume without loss of generality that $\gamma(0) = 0$, and that $(\gamma^1)'(t) > c > 0$ for some c and all $t \in (a, b)$, where $a < 0 < b$. Due to the assumption on the C^1-convergence of the γ_k, we may additionally suppose that the same inequality holds on (a, b) for each $(\gamma_k^1)'$. Thus each γ_k^1 and γ^1 itself are C^1-diffeomorphisms from (a, b) onto their respective image. In addition (restricting to n large if necessary), there exists a nontrivial interval J around 0 that is contained in $\bigcap_k \gamma_k^1((a, b)) \cap \gamma((a, b))$. Then we set $U := J \times B_R^{n-1}(0)$, where $R > 0$ is chosen such that $\bigcup_k \gamma_k((a,b)) \cup \gamma((a,b)) \subseteq U$. We can write V in the form $t \mapsto (\gamma(t), v(t))$ with $v \in C^1((a,b), \mathbb{R}^n)$, and analogously $V_k(t) = (\gamma_k(t), v_k(t))$, with $v_k \to v$ in C^1. For $x \in U$ we set

$$\tilde{V}(x) := (x, v((\gamma^1)^{-1}(x^1))) \qquad \tilde{V}_k(x) := (x, v_k((\gamma_k^1)^{-1}(x^1))),$$

giving C^1-extensions of V resp. V_k to U. By [2, Cor. 1] we have that $(\gamma_k^1)^{-1}$ converges to $(\gamma^1)^{-1}$ locally uniformly, and that the same is true for the first derivatives. Consequently, $\tilde{V}_k \to \tilde{V}$ in $C^1_{\text{loc}}(U)$. $\square$

As the proof shows, if the V_k converge to V merely in C^0_{loc} then one can still find C^1-extensions to U that also converge in C^0_{loc}.

References

1. Alexander, S., Graf, M., Kunzinger, M., Sämann, C.: Generalized cones as Lorentzian length spaces: causality, curvature, and singularity theorems, Comm. Anal. Geom. (to appear)
2. Barvinek, E., Daler, I., Francocircu, J.: Convergence of sequences of inverse functions. Arch. Math. (Brno) **27B**, 201–204 (1991)
3. Bourbaki, N.: Elements of Mathematics. Algebra, Part I: Chapters 1–3. Hermann, Paris (1974)
4. Cavalletti, F., Mondino, A.: Optimal transport in Lorentzian synthetic spaces, synthetic timelike Ricci curvature lower bounds and applications. arXiv: 2004.08934
5. Chruściel, P.T., Grant, J.D.E.: On Lorentzian causality with continuous metrics. Classical Quantum Gravity **29**(14), 145001, 32 (2012)
6. Chruściel, P.T., Grant, J.D.E., Kunzinger M, Minguzzi, E.: Non-regular spacetime geometry. J. Phys. Conf. Ser. 968 (2018)
7. Graf, M.: Singularity theorems for C^1-Lorentzian metrics. Commun. Math. Phys. **378**(2), 1417–1450 (2020)
8. Graf, M., Grant, J.D.E., Kunzinger, M., Steinbauer, R.: The Hawking–Penrose singularity theorem for $C^{1,1}$-Lorentzian metrics. Commun. Math. Phys. **360**, 1009–1042 (2018)
9. Galloway, G., Ling, E., Sbierski, J.: Timelike completeness as an obstruction to C^0-extensions. Commun. Math. Phys. **359**(3), 937–949 (2018)
10. Galloway, G., Senovilla, J.: Singularity theorems based on trapped submanifolds of arbitrary codimension. Classical Quantum Gravity **27**(15), 152002,10 (2010)
11. Geroch, R., Traschen, J.: Strings and other distributional sources in general relativity. Phys. Rev. D **36**, 1017–1031 (1987)
12. Grant, J.D.E., Kunzinger, M., Sämann, C., Steinbauer, R.: The future is not always open. Lett. Math. Phys. **110**(1), 83–103 (2020)
13. Grosser, M., Kunzinger, M., Oberguggenberger, M., Steinbauer, R.: Geometric Theory of Generalized Functions. Kluwer, Dordrecht (2001)
14. Hawking, S.W., Ellis, G.F.R.: The Large Scale Structure of Space-Time. Cambridge University Press, Cambridge (1973)
15. Hartman, P.: Ordinary Differential Equations, Volume 38 of Classics in Applied Mathematics. SIAM (2002)
16. Hartman, P., Wintner, A.: On the problems of geodesics in the small. Am. J. Math. **73**, 132–148 (1951)
17. Hawking, S.W., Penrose, R.: The singularities of gravitational collapse and cosmology. Proc. Roy. Soc. Lond. A **314**, 529–548 (1970)
18. Hedicke, J., Minguzzi, E., Schinnerl, B.: Steinbauer, R., Suhr, S.: Causal simplicity and (maximal) null pseudoconvexity, submitted. arXiv:2105.08998 [gr-qc]
19. Hörmander, L.: The Analysis of Linear Partial Differential Operators I. Springer, Berlin (1990)
20. Horváth, J.: Topological Vector Spaces and Distributions, vol. I. Addison-Wesley Publishing Co. Reading, Boston (1966)
21. Kriele, M.: Spacetime. Springer, Berlin (2001)
22. Kunzinger, M., Sämann, C.: Lorentzian length spaces. Ann. Global Anal. Geom. **54**(3), 399–447 (2018)
23. Kunzinger, M., Steinbauer, R., Stojković, M.: The exponential map of a $C^{1,1}$-metric. Differ. Geom. Appl. **34**, 14–24 (2014)
24. Kunzinger, M., Steinbauer, R., Stojković, M., Vickers, J.A.: A regularisation approach to causality theory for $C^{1,1}$-Lorentzian metrics. Gen. Relativ. Gravit. **46**, 1738, 18 (2014)

25. Kunzinger, M., Steinbauer, R., Stojković, Vickers, J.A.: Hawking's singularity theorem for $C^{1,1}$-metrics. Classical Quantum Gravity **32**, 075012, 19 (2015)
26. Kunzinger, M., Steinbauer, R., Vickers, J.A.: The Penrose singularity theorem in regularity $C^{1,1}$. Classical Quantum Gravity **32**, 155010 (2015)
27. Lange, C., Lytchak, A., Sämann C.: Lorentz meets Lipschitz. Adv. Theor. Math. Phys. (to appear)
28. LeFloch, P.G., Mardare, C.: Definition and stability of Lorentzian manifolds with distributional curvature. Port. Math. (N.S.) **64**(4), 535–573 (2007)
29. Marsden, J.E.: Generalized Hamiltonian mechanics: a mathematical exposition of non-smooth dynamical systems and classical Hamiltonian mechanics. Arch. Rational Mech. Anal. **28**, 323–361 (1967/68)
30. Minguzzi, E.: Limit curve theorems in Lorentzian geometry. J. Math. Phys. **49**(9), 092501 (2008)
31. Minguzzi, E.: Convex neighborhoods for Lipschitz connections and sprays. Monatsh. Math. **177**(4), 569–625 (2015)
32. Minguzzi, E., Sanchez, M.: The causal hierarchy of spacetimes, recent developments in pseudo Riemannian geometry. ESI Lect. Math. Phys. **299–358**, 0609119 (2008)
33. Minguzzi, E., Sánchez, M.: The causal hierarchy of spacetimes. In: Recent Developments in Pseudo-Riemannian Geometry. ESI Lectures in Mathematics and Physics, pp. 299–358. European Mathematical Society, Zurich (2008)
34. Minguzzi, E.: Lorentzian causality theory. In: Living Reviews in Relativity, vol. 22:3, pp. 1–202 Springer (2019)
35. Minguzzi, E.: Causality theory for closed cone structures with applications. Rev. Math. Phys. **31**(5), 1930001, 139 (2019)
36. Nigsch, E.A.: Bornologically isomorphic representations of distributions on manifolds. Monatsh. Math. **170**, 49–63 (2013)
37. Oberguggenberger, M.: Über Folgenkonvergenz in lokalkonvexen Räumen. Math. Nachr. **129**, 219–234 (1986)
38. Ohta, S.: Examples of spaces with branching geodesics satisfying the curvature-dimension condition. Bull. Lond. Math. Soc. **46**(1), 19–25 (2014)
39. O'Neill, B.: Semi-Riemannian Geometry with Applications to Relativity. Pure and Applied Mathematics, vol. 103. Academic Press, New York (1983)
40. Penrose, R.: Gravitational collapse and space-time singularities. Phys. Rev. Lett. **14**, 57–59 (1965)
41. Sämann, C.: Global hyperbolicity for spacetimes with continuous metrics. Ann. Henri Poincaré **17**(6), 1429–1455 (2016)
42. Sämann, C., Steinbauer, R.: On geodesics in low regularity. J. Phys. Conf. Ser. **968**, 012010, 14 (2018)
43. Schinnerl, B., Steinbauer, R.: A Note on the Gannon-Lee theorem. Lett. Math. Phys. **111**, 142 (2021)
44. Shiohama, K.: An Introduction to the Geometry of Alexandrov Spaces. Lecture Notes Series, vol. 8. Seoul National University (1993)
45. Sbierski, J.: The C^0-inextendibility of the Schwarzschild spacetime and the spacelike diameter in Lorentzian geometry. J. Differ. Geom. **108**, 319–378 (2018)
46. Senovilla, J.M.M.: Singularity theorems and their consequences. Gen. Relat. Gravit. **30**(5), 701–848 (1998)
47. Senovilla, J.M.M.: A critical appraisal of the singularity theorems. arXiv:2108.07296 [gr-qc] (2021)
48. Senovilla, J.M.M., Garfinkle, D.: The 1965 Penrose singularity theorem. Classical Quantum Gravity **32**(12), 124008, 45 (2015)
49. Steinbauer, R.: A note on distributional semi-Riemannian geometry. Novi Sad. J. Math. **38**(3), 189–199 (2008)
50. Steinbauer, R., Vickers, J.: On the Geroch-Traschen class of metrics. Classical Quantum Gravity **26**(6), 065001, 19 (2009)
51. Villani, C.: Optimal Transport. Old and New, Grundlehren der Mathematischen Wissenschaften. Springer, Berlin (2009)

Communicated by P. Chrusciel

Letters in Mathematical Physics (2024) 114:96
https://doi.org/10.1007/s11005-024-01843-7

Examples of cosmological spacetimes without CMC Cauchy surfaces

Eric Ling[1] · **Argam Ohanyan**[2]

Received: 23 February 2024 / Revised: 20 May 2024 / Accepted: 21 June 2024

Abstract

CMC (constant mean curvature) Cauchy surfaces play an important role in mathematical relativity as finding solutions to the vacuum Einstein constraint equations is made much simpler by assuming CMC initial data. However, Bartnik (Commun Math Phys 117(4):615–624, 1988) constructed a cosmological spacetime without a CMC Cauchy surface whose spatial topology is the connected sum of two three-dimensional tori. Similarly, Chruściel et al. (Commun Math Phys 257(1):29–42, 2005) constructed a vacuum cosmological spacetime without CMC Cauchy surfaces whose spatial topology is also the connected sum of two tori. In this article, we enlarge the known number of spatial topologies for cosmological spacetimes without CMC Cauchy surfaces by generalizing Bartnik's construction. Specifically, we show that there are cosmological spacetimes without CMC Cauchy surfaces whose spatial topologies are the connected sum of any compact Euclidean or hyperbolic three-manifold with any another compact Euclidean or hyperbolic three-manifold. Analogous examples in higher spacetime dimensions are also possible. We work with the Tolman–Bondi class of metrics and prove gluing results for variable marginal conditions, which allows for smooth gluing of Schwarzschild to FLRW models.

Keywords Cosmological spacetimes · CMC Cauchy surfaces · Tolman–Bondi metrics

Mathematics Subject Classification 83C20 · 53B30 · 53C50

Dedicated to the late Robert Bartnik.

✉ Eric Ling
el@math.ku.dk

Argam Ohanyan
argam.ohanyan@univie.ac.at

1 Copenhagen Centre for Geometry and Topology (GeoTop), Department of Mathematical Sciences, University of Copenhagen, 2100 Copenhagen, Denmark

2 Department of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria

Published online: 09 July 2024

Contents

1 Introduction

In this section, we present the sources of interest for "no-CMC" cosmological spacetimes, i.e., cosmological spacetimes without any CMC Cauchy surfaces. The flow of the presentation follows [12], where we refer to for many more details, especially regarding historical developments. Let us begin by recalling the cosmological version of the Hawking–Penrose singularity theorem.[1]

Theorem 1.1 (Cosmological Hawking–Penrose singularity theorem). *Let (M, g) be a globally hyperbolic spacetime, with compact Cauchy surfaces, satisfying the strong energy condition. (Such spacetimes are sometimes called* cosmological.*) If (M, g) satisfies the generic condition, i.e., along every inextendible causal geodesic γ, there exists a parameter t such that the tidal force operator*

$$R : T_{\gamma(t)}M \to T_{\gamma(t)}M, \quad v \mapsto R(v, \gamma'(t))\gamma'(t)$$

is not identically zero, then there are incomplete causal geodesics in M.

Observe that the assumption of genericity cannot be dropped: Indeed, take (S, h) to be any compact Riemannian manifold with $\mathrm{Ric} \geq 0$, then $(\mathbb{R} \times S, -dt^2 + h)$ satisfies all of the assumptions except for genericity and is causally geodesically complete. In 1988, Bartnik [2] conjectured that this theorem is rigid in the genericity assumption, i.e., that Lorentzian products are the only nongeneric counterexamples. More precisely, the conjecture states the following:

Conjecture 1.2 (Bartnik's splitting conjecture). Let (M, g) be a globally hyperbolic spacetime with compact Cauchy surfaces satisfying the strong energy condition. If (M, g) is timelike geodesically complete, then (M, g) splits isometrically as a product $(\mathbb{R} \times S, -dt^2 + h)$, where (S, h) is a compact Riemannian manifold.

While this conjecture has been proved over the years under various additional assumptions (we refer to [12] for a detailed discussion of these developments), the version as stated remains open to this day and is one of the most significant open problems in mathematical General Relativity.

In connection with Conjecture 1.2, Bartnik proved the following:

[1] The general form of the Hawking–Penrose theorem makes significantly fewer assumptions on the causality of the spacetime, see, e.g., [3, Thm. 12.47].

Theorem 1.3 [2] *Let (M, g) be a globally hyperbolic spacetime, with compact Cauchy surfaces, satisfying the strong energy condition. If (M, g) is timelike geodesically complete, then (M, g) splits isometrically as a product $(\mathbb{R} \times S, -dt^2 + h)$ if and only if there exists a constant mean curvature (CMC) Cauchy surface in M.*

In [10], Dilts and Holst review the issue of the existence of CMC Cauchy surfaces in globally hyperbolic spacetimes with compact Cauchy surfaces and raise the question: When do spacetimes have CMC Cauchy surfaces? Motivated by this question, in [13], Galloway and the first named author proved the existence of a CMC Cauchy surface under the assumptions of compact Cauchy surfaces, future timelike geodesic completeness and nonpositive timelike sectional curvature. Clearly the latter assumption implies the strong energy condition. Therefore, they proposed the following conjecture:

Conjecture 1.4 [13] Let (M, g) be a globally hyperbolic, future timelike geodesically complete spacetime with compact Cauchy surfaces satisfying the strong energy condition. Then M contains a CMC Cauchy surface.

As discussed in [10], most existence results rely on barrier methods, see [14] for an application. However, a well-known example of Bartnik [2] shows that not all spacetimes with compact Cauchy surfaces satisfying the strong energy condition contain CMC Cauchy surfaces. Specifically, Bartnik constructs a timelike geodesically incomplete spacetime satisfying the remaining assumptions of Conjecture 1.2 that does not contain any CMC Cauchy surfaces. The rough procedure is as follows: Consider flat FLRW with T^3 spacelike slices, cut out a small ball out of T^3 and glue the spatial cylinder of Schwarzschild to it. Then extend Schwarzschild to one-half of its maximal extension, and glue a time-inverted copy of the resulting spacetime across one of the event horizons in Schwarzschild. The gluing of Schwarzschild to flat FLRW is done in the framework of Tolman–Bondi metrics, reducing the gluing of spacetimes to the smooth interpolation of functions. Using initial data gluing methods, Chruściel, Isenberg and Pollack construct a similar (vacuum) example obtained in [9]. To our knowledge, these are the only such examples which appear in the literature. The goal of this article is to provide additional ones, by generalizing Bartnik's spacetime gluing construction.

Our generalization of Bartnik's example is twofold: First, we extend the gluing of Tolman–Bondi metrics to variable marginal conditions using ODE existence and uniqueness arguments (see Theorem 2.9). While certainly interesting, this gluing procedure may result in spacetimes which are not globally hyperbolic due to the smallness of the gluing region. To remedy this, we analyze maximality of the ODE solutions at hand and are able to generalize Bartnik's construction by gluing Schwarzschild to flat and hyperbolic FLRW models with compact spacelike slices in such a way that the resulting spacetime is globally hyperbolic (see Sect. 4 and Theorem 4.3). This gives a large class of reasonable spacetimes which do not contain any CMC Cauchy surfaces. Note that if Conjecture 1.2 is true, then any such example must be incomplete. While Bartnik's construction as well as our generalizations are manifestly incomplete, that incompleteness is a consequence of how the spacetime was constructed. It would be interesting to find similar constructions where the incompleteness is less obvious and more so a consequence of general principles. Such an example would help to gain insight into Conjecture 1.2. In addition, it would be interesting to determine whether or not the example in [9] is timelike geodesically complete.

For convenience, let us state the main results (Theorem 2.9 and Theorem 4.3) informally:

Theorem 1.5 *Let $(M_1 = D_1 \times S^2, g_1)$ and $(M_2 = D_2 \times S^2, g_2)$ be Tolman–Bondi spacetimes satisfying Einstein's equations for dust (with nonnegative energy density), where g_i are of the form*

$$g_i = -dt^2 + X_i(t,r)^2 dr^2 + Y_i(t,r)^2 d\Omega^2, \quad 0 < X_i, Y_i \in C^\infty(D_i), \tag{1.1}$$

$D_i \subset \mathbb{R}^2$ open and connected. If certain monotonicity conditions are satisfied, then these spacetimes may be (smoothly) glued across an r-interval (where we write $(t,r) \in D_i$) to a larger Tolman–Bondi spacetime (M, g) which also satisfies Einstein's equations for dust with nonnegative energy density.

Theorem 1.6 *Given compact Riemannian quotients $Q, \tilde{Q}$ of $\mathbb{R}^3$ or H^3, we glue Schwarzschild spacetime to a dust-filled FLRW spacetime with spacelike slices Q, then glue the resulting spacetime to a gluing of Schwarzschild to FLRW with spacelike slices $\tilde{Q}$, but with opposite time orientation, across a Schwarzschild event horizon. The resulting spacetime, (M, g), is globally hyperbolic with compact Cauchy surfaces of topology $Q\#\tilde{Q}$ and satisfies the strong and dominant energy conditions. Moreover, (M, g) does not contain any CMC Cauchy surfaces. Analogous examples can be constructed in higher spacetime dimensions by considering higher-dimensional Schwarzschild and FLRW spacetimes.*

The paper is organized as follows: In Sect. 2, we discuss important properties of Tolman–Bondi spacetimes, where, for convenience, we give full derivations of results we could only find in older literature. We discuss how to describe Schwarzschild and FLRW spacetimes as Tolman–Bondi spacetimes. The section is rounded off by the general gluing result for Tolman–Bondi spacetimes (Theorem 2.9). Section 3 is dedicated to a thorough discussion of Bartnik's construction. We review both proofs of Bartnik as to why the constructed spacetime has no CMC Cauchy surfaces. One is of topological nature, relying on the form of the topology of the spacelike slice of FLRW, while the other is more elementary and Lorentz-geometric, using Hawking's singularity theorem and CMC foliation arguments to achieve a contradiction to the existence of CMCs. We describe in Sect. 4 how to generalize Bartnik's construction by gluing Schwarzschild to flat or hyperbolic FLRW with arbitrary compact Riemannian quotients of $\mathbb{R}^3$ or H^3 as spacelike slices. One of our proofs makes use of the positive resolution of the surface subgroup conjecture [16]. We also describe how to obtain the analogous constructions in higher spacetime dimensions. Finally, we give a conclusion and an outlook on possible relevant future lines of research in Sect. 5.

1.1 Notation and conventions

Let us collect here some notation and conventions we will use throughout the paper. $A \subset B$ denotes (not necessarily strict) inclusion. A *spacetime* is a connected, time-oriented Lorentzian manifold whose metric has the signature $(-, +, \dots, +)$. By *Cauchy surface*, we will always mean a smooth, spacelike hypersurface that is met uniquely by each future-directed C^1-causal curve $\gamma : (a, b) \to M$ that is inextendible, i.e., the limits $\lim_{t\to a} \gamma(t)$ and $\lim_{t\to b} \gamma(t)$ do not exist.[2] Ric denotes the Ricci tensor and R denotes the scalar curvature. For functions $f(t, r)$ depending on a time variable t and a radial variable r, we will write $\dot{f} = \partial_t f$ and $f' = \partial_r f$. We write T^n for the n-dimensional torus and H^n for n-dimensional hyperbolic space. Given two smooth manifolds M_1 and M_2, $M_1 \# M_2$ denotes their connected sum. Given groups G_1 and G_2, $G_1 * G_2$ denotes their free product. Given any spacelike hypersurface S in a spacetime M, $H_S \equiv H(S) = \mathrm{Tr}(\nabla N)$ denotes its (future) mean curvature (where N is its future unit normal). We say a spacelike hypersurface S has constant mean curvature (or is CMC) if H_S is a constant. We say it is maximal if $H_S = 0$. We will write Pr_i for the projection map from a given product space to the i-th factor. By a *Riemannian quotient* Q of a Riemannian manifold N, we mean a quotient arising from the smooth, free, proper and isometric action of a *discrete* Lie group Γ on N, i.e., $Q = N/\Gamma$, equipped with the unique Riemannian metric such that the quotient map $N \to Q$ is a normal Riemannian covering; in particular, Q is locally isometric to N and $\dim N = \dim Q$.

2 Tolman–Bondi spacetimes

Let us begin with a discussion of Tolman–Bondi metrics, which are a class of metrics in four spacetime dimensions with S^2-symmetry and which present the setting for the constructions of cosmological spacetimes without CMC Cauchy surfaces. The advantage in using them for our setting is that they allow for purely analytic gluing constructions. We follow the derivations in [4] and [11], but write out all of the proofs for convenience.

Definition 2.1 (Tolman–Bondi spacetimes). A *Tolman–Bondi (TB) spacetime* is a four-dimensional Lorentzian manifold $M = D \times S^2$ with metric

$$g = -dt^2 + X(t, r)^2 dr^2 + Y(t, r)^2 d\Omega^2, \tag{2.1}$$

where D is an open connected subset of $\mathbb{R}^2$, and X and Y are smooth positive functions on D. Here (and everywhere else), $d\Omega^2$ is the standard round metric on S^2. Moreover, (M, g) is understood to be time-oriented via $\frac{\partial}{\partial t}$.

Observe that a Tolman–Bondi spacetime is a warped product of the form $D \times_Y S^2$, where S^2 is equipped with the usual round metric and D is a two-dimensional spacetime with the Lorentzian metric $-dt^2 + X(t, r)dr^2$.

[2] This is the "correct" notion of smooth, spacelike Cauchy surface, cf. [17, Rem. 3.31].

We will be interested in Tolman–Bondi spacetimes satisfying Einstein's equations for dust:

Definition 2.2 (Dust-filled Tolman–Bondi spacetimes). We say a Tolman–Bondi spacetime $(M = D \times S^2, g)$ is *dust-filled* if it satisfies the Einstein equations for dust, i.e., there exists $\rho \in C^\infty(D)$ such that

$$G := \mathrm{Ric} - \frac{1}{2}Rg = 8\pi\rho\, dt^2. \tag{2.2}$$

Recall that a spacetime (M, g) is said to satisfy the *strong energy condition (SEC)* if $\mathrm{Ric}(V, V) \geq 0$ for all timelike vector fields V on M. It satisfies the *dominant energy condition (DEC)* if $G(V, W) \geq 0$ for future-directed causal vector fields V, W on M.

Lemma 2.3 (Dust-filled TB spacetimes satisfy SEC and DEC). *If (M, g) is a dust-filled Tolman–Bondi spacetime with ρ nonnegative, then it satisfies the strong and dominant energy conditions.*

Proof We first show the SEC. Complete ∂_t to an orthonormal frame of M and take the g-trace of (2.2) with respect to that frame to get

$$R - 2R = -8\pi\rho,$$

thus $R = 8\pi\rho$. Reinserting this into (2.2), we get

$$\mathrm{Ric} = 8\pi\rho dt^2 + 4\pi\rho g.$$

Thus, if V is any unit timelike vector field on M, the reverse Cauchy–Schwarz inequality together with $\rho \geq 0$ implies

$$\mathrm{Ric}(V, V) \;=\; 8\pi\rho\, g(\nabla t, V)^2 - 4\pi\rho \;\geq\; 8\pi\rho - 4\pi\rho \;=\; 4\pi\rho \geq 0,$$

thus establishing the SEC.

Showing the DEC is easier: Let V and W be future-directed causal vectors. Then the component $V^t = dt(V)$ is nonnegative since V is future directed. Likewise $W^t \geq 0$. Therefore, $G(V, W) = 8\pi\rho V^t W^t \geq 0$. □

Proposition 2.4 *Let (M, g) be a dust-filled Tolman–Bondi spacetime. Then $\partial_t(\frac{1}{X}Y') = 0$ and $\partial_t(\rho Y' Y^2) = 0$.*

Proof Let $G = \mathrm{Ric} - \frac{1}{2}Rg$ denote the Einstein tensor. Since the Einstein equations hold with dust, we have $G(\partial_t, \partial_r) = 0$. A calculation shows

$$G(\partial_t, \partial_r) \;=\; \frac{2}{XY}(\dot{X}Y' - X\dot{Y}'). \tag{2.3}$$

Therefore, $\dot{X}Y' - X\dot{Y}' = 0$ which implies

$$X^2\partial_t(Y'/X) = 0.$$

Since X is never zero, $\partial_t(Y'/X) = 0$.

To prove that $\rho Y'Y^2$ is independent of t, we investigate $G(\partial_t, \partial_t)$ and $G(\partial_r, \partial_r)$: An explicit calculation yields

$$G(\partial_t, \partial_t) = \frac{1}{X^3Y^2}\Big(X^3\dot{Y}^2 + 2X^2\dot{X}Y\dot{Y} + X^3 - 2XYY'' - XY'^2 + 2X'YY'\Big) \tag{2.4}$$

$$G(\partial_r, \partial_r) = -\frac{1}{Y^2}\Big(2X^2Y\ddot{Y} + X^2\dot{Y}^2 + X^2 - Y'^2\Big). \tag{2.5}$$

We introduce the quantity

$$S := Y\left(1 + \dot{Y}^2 - \frac{Y'^2}{X^2}\right). \tag{2.6}$$

It is elementary to check that

$$G(\partial_t, \partial_t)Y' = \frac{1}{Y^2}S', \tag{2.7}$$

$$G(\partial_r, \partial_r)\dot{Y} = -\frac{X^2}{Y^2}\dot{S}. \tag{2.8}$$

Now, $\dot{S} = 0$ since $G(\partial_r, \partial_r) = 0$. Therefore, $\dot{S}' = 0$; hence,

$$\partial_t\big(Y^2Y'G(\partial_t, \partial_t)\big) = 0.$$

Since $G(\partial_t, \partial_t) = 8\pi\rho$, the claim follows. □

The next result is a key property of dust-filled Tolman–Bondi spacetimes, as it allows one to view the coefficient functions X and Y as solutions of ODEs.

Proposition 2.5 *Let (M, g) be a dust-filled Tolman–Bondi spacetime. Then the function S defined in Eq. (2.6) is independent of t and is an antiderivative of $8\pi\rho Y'Y^2$. (The latter is independent of t due to Proposition 2.4.) Moreover, writing $W := Y'/X$, it holds that*

$$\dot{Y}(t, r)^2 = W(r)^2 - 1 + \frac{1}{Y(t, r)}S(r). \tag{2.9}$$

Proof Let S be the function introduced in Eq. (2.6). We have already shown in the proof of Proposition 2.4 that $\dot{S} = 0$. (This is a consequence of Eq. (2.8).) By Eq. (2.7), and noting that $G(\partial_t, \partial_t) = 8\pi\rho$, we get

$$S' = 8\pi\rho Y' Y^2.$$

Inserting the definition of S, we get

$$S = Y\left(1 + \dot{Y}^2 - \frac{Y'^2}{X^2}\right),$$

solving for $\dot{Y}^2$ gives the claim. □

Definition 2.6 (Marginal condition and mass function). Let (M, g) be a dust-filled Tolman Bondi spacetime,

$$g = -dt^2 + X^2 dr^2 + Y^2 d\Omega^2.$$

We call the associated functions $W := Y'X^{-1}$ and $S := Y(1+\dot{Y}^2 - Y'^2 X^{-2})$ *marginal condition* and *mass function*,[3] respectively.

Next, we discuss some important examples of Tolman–Bondi spacetimes, namely Schwarzschild and FLRW. These will be the building blocks for constructing cosmological spacetimes without CMC Cauchy surfaces.

Example 2.7 (Schwarzschild spacetime). The standard (nonextended) Schwarzschild spacetime of mass $m > 0$ is $(R \neq 0, 2m)$

$$g_S := -\left(1 - \frac{2m}{R}\right) d\tau^2 + \frac{1}{1 - \frac{2m}{R}} dR^2 + R^2 d\Omega^2.$$

We introduce Lemaître coordinates t, r via

$$dt = d\tau + \sqrt{\frac{2m}{R}} \frac{dR}{1 - \frac{2m}{R}},$$
$$dr = d\tau + \sqrt{\frac{R}{2m}} \frac{dR}{1 - \frac{2m}{R}}.$$

Writing the metric in these coordinates yields

$$g_S = -dt^2 + \left(\frac{2m}{\frac{3}{2}(r - t)}\right)^{\frac{2}{3}} dr^2 + (2m)^{\frac{2}{3}} \left(\frac{3}{2}(r - t)\right)^{\frac{4}{3}} d\Omega^2.$$

This is in Tolman–Bondi form defined on $M := \{(t, r) \in \mathbb{R}^2 : t < r\} \times S^2$ and describes one half of the maximal analytic extension of the Schwarzschild spacetime.

[3] $S/2$ can justifiably be called *mass* of the Tolman–Bondi spacetime (in Schwarzschild: $S \equiv 2m$). We prefer to work with S and have thus chosen the name *mass function* to avoid confusion.

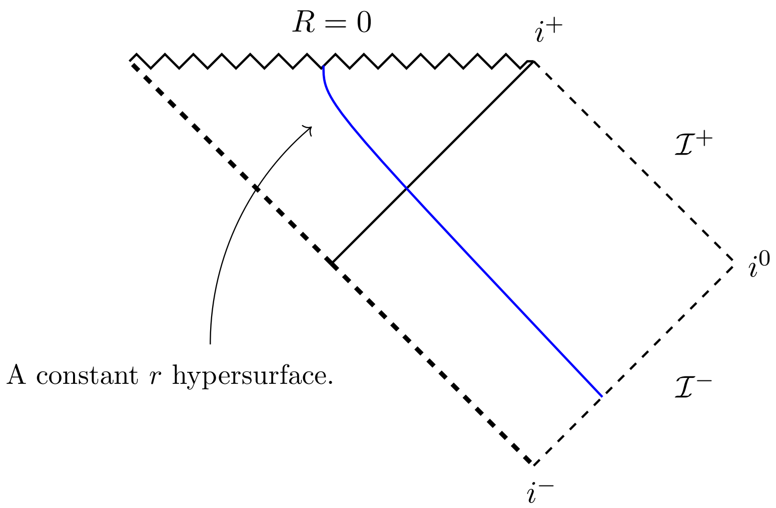


Fig. 1 The Lemaître coordinates (t, r) cover one-half of the maximally extended Schwarzschild spacetime. The hypersurfaces of constant r are timelike and foliate the spacetime; they become null as they approach $\mathcal{I}^-$. As $r \to -\infty$, the hypersurfaces approach the (totally geodesic) null hypersurface represented by the thick dashed lines, i.e., the hypersurface $T = -X$ in Kruskal coordinates

The hypersurfaces $\{r = const\}$ are timelike and, for $r \to -\infty$, converge to the event horizon separating the two halves of the maximal extension (see Fig. 1). For later use, let us note that

$$W(r) = 1,$$
$$S(r) = 2m.$$

Example 2.8 (FLRW). We will now describe the FLRW spacetimes with any of the constant spatial sectional curvatures $k = +1, 0, -1$ as Tolman–Bondi metrics. We will refer to these as *spherical, flat and hyperbolic* FLRW, respectively. We refer to [18, Sec. 12.5] for more details than what is presented below. Note that we work with negative time intervals for technical reasons, as this will later allow us to glue FLRW spacetimes to Schwarzschild using Theorem 2.9 (resp. Corollary 2.12) in conjunction with Propositions 3.1 and 4.1.

(i) Flat FLRW: For any parameter $\mathcal{M} > 0$, the spacetime $(-\infty, 0) \times_{a_f} \mathbb{R}^3$, with

$$a_f(t) := (6\pi\mathcal{M})^{\frac{1}{3}} t^{\frac{2}{3}},$$

satisfies Einstein's equations for dust, with energy density $\rho(t) = \mathcal{M} a_f(t)^{-3}$, where $\mathcal{M}$ is the Friedmann mass parameter. Writing the three-dimensional Euclidean metric in spherical coordinates as $dr^2 + r^2 d\Omega^2$, the metric for flat FLRW can be written in Tolman–Bondi form

$$g_f = -dt^2 + a_f(t)^2 dr^2 + a_f(t)^2 r^2 d\Omega^2$$

on the manifold $(-\infty, 0) \times (0, \infty) \times S^2$. We calculate W and S for this metric:

$$W(r) = 1,$$
$$S(r) = \frac{8\pi \mathcal{M}}{3} r^3.$$

(ii) Hyperbolic FLRW: Here, the spacetime is $(-\infty, 0) \times_{a_h} H^3$. The function a_h cannot be written explicitly, but it satisfies Friedmann's equation

$$\dot{a}_h^2 - 1 = \frac{8\pi \mathcal{M}}{3a_h}$$

and can be written in dependence on a parameter $\eta < 0$ (together with time) as

$$t(\eta) = \frac{4\pi \mathcal{M}}{3}(\sinh(\eta) - \eta),$$
$$a_h(\eta) = \frac{4\pi \mathcal{M}}{3}(\cosh(\eta) - 1).$$

Writing the hyperbolic metric on H^3 as $dr^2 + \sinh(r)^2 d\Omega^2$ (on $(0, \infty) \times S^2$), the hyperbolic FLRW metric can be expressed on the manifold $(-\infty, 0) \times (0, \infty) \times S^2$ in Tolman–Bondi form as

$$g_h = -dt^2 + a_h(t)^2 dr^2 + a_h(t)^2 \sinh(r)^2 d\Omega^2.$$

The energy density is again $\rho(t) = \mathcal{M} a_h(t)^{-3}$, and moreover,

$$W(r) = \cosh(r),$$
$$S(r) = \frac{8\pi \mathcal{M}}{3} \sinh(r)^3.$$

(iii) Spherical FLRW: Similarly to before, consider $\left(-\frac{8\pi^2 \mathcal{M}}{3}, 0\right) \times_{a_s} S^3$, where a_s satisfies

$$\dot{a_s}^2 + 1 = \frac{8\pi \mathcal{M}}{3a_s}.$$

As in the hyperbolic case, a_s cannot be written explicitly, but can be described in dependence on a parameter $\theta \in (-2\pi, 0)$ (together with time) via

$$t(\theta) = \frac{4\pi \mathcal{M}}{3}(\theta - \sin(\theta)),$$
$$a_s(\theta) = \frac{4\pi \mathcal{M}}{3}(1 - \cos(\theta)).$$

Writing the metric on S^3 as $dr^2 + \sin(r)^2 d\Omega^2$ (on $(0,\pi) \times S^2$), the spherical FLRW metric can be written on the manifold $\left(-\frac{8\pi^2\mathcal{M}}{3}, 0\right) \times (0,\pi) \times S^2$ as

$$g_s = -dt^2 + a_s(t)^2 dr^2 + a_s(t)^2 \sin(r)^2 d\Omega^2.$$

Here, $\rho(t) = \mathcal{M} a_s(t)^{-3}$ and

$$W(r) = \cos(r),$$
$$S(r) = \frac{8\pi\mathcal{M}}{3}\sin(r)^3.$$

The next result shows that, under certain monotonicity conditions, we can always glue dust-filled Tolman–Bondi spacetimes across intervals of r by solving the ODE for Y given in Proposition 2.5.

Theorem 2.9 (Gluing Tolman–Bondi spacetimes). *Let $(M_1 = D_1 \times S^2, g_1)$ and $(M_2 = D_2 \times S^2, g_2)$ be dust-filled Tolman–Bondi spacetimes with nonnegative energy densities ρ_1, ρ_2. We write $Pr_2(D_i) = (r_i^-, r_i^+)$, $i = 1, 2$. Suppose $r_1 \in (r_1^-, r_1^+)$ and $r_2 \in (r_2^-, r_2^+)$, $r_1 < r_2$. Moreover, suppose there exists t_0 such that $(t_0, r_i) \in D_i$, such that either both $\dot{Y}_1(t_0, r_i) > 0$ and $\dot{Y}_2(t_0, r_i) > 0$ or both $\dot{Y}(t_0, r_1) < 0$ and $\dot{Y}(t_0, r_2) < 0$, and suppose that the following technical assumption is satisfied: There exist smooth functions $W, S, r \mapsto Y_r(t_0) \in C^\infty((r_1^-, r_2^+))$ agreeing with $W_i, S_i, r \mapsto Y_i(t_0, r)$ on their corresponding domains with S monotonically increasing, $r \mapsto Y_r(t_0)$ strictly monotonically increasing, W positive, such that on $[r_1, r_2]$,*

$$W(r)^2 - 1 + \frac{S(r)}{Y_r(t_0)} > 0.$$

Then there exists a dust-filled Tolman–Bondi spacetime $(M = D \times S^2, g)$ with a nonnegative energy density ρ such that $Pr_2(D) = (r_1^-, r_2^+)$, $g = g_1$ for $r \leq r_1$ and $g = g_2$ for $r \geq r_2$.

Proof Let us assume for definiteness that $\dot{Y}_1(t_0, r_1) < 0$ and $\dot{Y}_2(t_0, r_2) < 0$. We are interested in solving the following ODE initial value problem for fixed $r \in [r_1, r_2]$:

$$\begin{cases} \dot{Y}(t,r) = -\sqrt{W(r)^2 - 1 + \frac{S(r)}{Y(t,r)}}, \\ Y(t_0, r) = Y_r(t_0). \end{cases}$$

First, the fact that $\dot{Y}_i(\cdot, r_i) < 0$ implies that for all t close to t_0,

$$\dot{Y}_i(t, r_i) = -\sqrt{W_i(r_i)^2 - 1 + \frac{S_i(r_i)}{Y_i(t, r_i)}}.$$

Let S, W, $Y_r(t_0)$ be as in the assumption. Due to positivity, the right-hand side of the initial value problem in question depends smoothly on $Y(\cdot, r)$, for every fixed

$r \in [r_1, r_2]$, hence admits a unique solution $Y(t, r)$ on a maximal domain of definition (t_r^-, t_r^+). We may extract a common domain of definition (t^-, t^+) such that all $Y(t, r)$, $r \in [r_1, r_2]$, are defined for $t \in (t^-, t^+)$. Moreover, since the initial data $r \mapsto Y_r(t_0)$ depends smoothly on r, so does the solution. We may assume that (t^-, t^+) is small enough so that $Y, Y' > 0$, $\dot{Y} < 0$ for all $(t, r) \in (t^-, t^+) \times [r_1, r_2]$.

Let us now construct the spacetime: We set $D := \tilde{D}_1 \cup ((t^-, t^+) \times [r_1, r_2]) \cup \tilde{D}_2$, where

$$\tilde{D}_1 := \{(t, r) \in D_1 : r < r_1\},$$
$$\tilde{D}_2 := \{(t, r) \in D_2 : r > r_2\}.$$

For $(t, r) \in (t^-, t^+) \times [r_1, r_2]$, we let Y be the solution obtained above and $X := Y'/W$. Then $(t, r) \in D \mapsto X(t, r), Y(t, r)$ are positive, smooth functions on D. Moreover, by construction, the corresponding Tolman–Bondi metric is dust-filled with energy density given for $(t, r) \in (t^-, t^+) \times [r_1, r_2]$ by

$$\rho := \frac{S'(r)}{8\pi Y^2 Y'} \geq 0.$$

To elucidate, Eqs. (2.3)–(2.8) in the proof of Proposition 2.4 are true for any Tolman–Bondi metric. Therefore, $G(\partial_t, \partial_r) = 0$ since $X = Y'/W$ and $G(\partial_r, \partial_r) = 0$ since $\dot{S} = 0$. Thus, the Tolman–Bondi metric is dust-filled. □

Remark 2.10 In many cases, it is elementary to satisfy the technical assumption of Theorem 2.9: Indeed, if

(i) either $\dot{Y}_1(t_0, r_1) > 0$ and $\dot{Y}_2(t_0, r_2) > 0$ or $\dot{Y}_1(t_0, r_1) < 0$ and $\dot{Y}_2(t_0, r_2) < 0$,
(ii) $0 < S_1(r_1) \leq S_2(r_2)$,
(iii) $Y_1(t_0, r_1) < Y_2(t_0, r_2)$, $\quad Y_1'(t_0, r_1) > 0$, $\quad Y_2'(t_0, r_2) > 0$,
(iv) $W_1(r_1)^2 \geq 1$ and $W_2(r_2)^2 \geq 1$, $W_1 > 0$ and $W_2 > 0$,

then one can simply choose monotonically increasing (resp. strictly monotonically increasing) connecting functions S, $Y_r(t_0)$ as well as any interpolating function W which increases resp. decreases from $W_1(r_1)$ to $W_2(r_2)$ (if $W_1(r_1) \leq W_2(r_2)$ resp. $W_1(r_1) \geq W_2(r_2)$).

Remark 2.11 (Gluing vacuum TB spacetimes). Note that if (M_1, g_1) and (M_2, g_2) are vacuum Tolman–Bondi spacetimes, i.e., $\rho_1 = 0$ and $\rho_2 = 0$, then S_1, S_2 are constants. If $S_1 = S_2$ and the remaining assumptions of Theorem 2.9 are satisfied, the proof shows that, by choosing the connecting smooth function S to be that same constant, the glued spacetime can be constructed to be vacuum.

Let us note the following useful consequence of Theorem 2.9 regarding the gluing of Schwarzschild spacetime to (fairly) arbitrary Tolman–Bondi spacetimes.

Corollary 2.12 (Gluing TB spacetimes to Schwarzschild). *Let (M_1, g_1) be Schwarzschild spacetime of mass m in Lemaître coordinates as described in Example 2.7. Let $(M_2 = D_2 \times S^2, g_2)$ be any dust-filled Tolman–Bondi spacetime with nonnegative energy density ρ_2. Let $r_1 \in \mathbb{R}$. Suppose there is an $r_2 > r_1$ and $t_0 < r_1$ such that $(t_0, r_2) \in D_2$. Moreover, suppose the following technical assumptions are satisfied:*

(i) $\dot{Y}_2(t_0, r_2) < 0$.
(ii) $2m \leq S_2(r_2)$.
(iii) $(2\,m)^{\frac{1}{3}}\left(\frac{3}{2}(r_1 - t_0)\right)^{\frac{2}{3}} < Y_2(t_0, r_2)$, $Y_2'(t_0, r_2) > 0$.

Then there exists a dust-filled Tolman–Bondi spacetime $(M = D \times S^2, g)$ *with a nonnegative energy density* ρ *such that* $Pr_2(D) = (-\infty, \sup Pr_2(D_2))$, $g|_{(-\infty, r_1]} = g_1$, *and* $g|_{[r_2, \sup Pr_2(D_2))} = g_2$.

Proof We need to verify the technical condition in Theorem 2.9. Indeed, pick any smooth, monotonically increasing connector S for S_1, S_2 and strictly monotonically increasing connector $Y_r(t_0)$ for $Y_1(t_0, \cdot)$ and $Y_2(t_0, \cdot)$ on the interval $[r_1, r_2]$. Note that $W_1 \equiv 1$, and $W_2(r_2) = \frac{Y_2'(t_0,r_2)}{X_2(t_0,r_2)} > 0$ by assumption. If $W_2(r_2) \geq 1$, we are done by Remark 2.10, so suppose $0 < W_2(r_2) < 1$. Moreover,

$$W_2(r_2)^2 - 1 + \frac{S_2(r_2)}{Y_2(t_0, r_2)} > 0$$

since $\dot{Y}_2(t_0, r_2) < 0$. Let $\delta > 0$ be such that

$$\frac{S_2(r_2)}{Y_2(t_0, r_2)} > 1 - W_2(r_2)^2 + \delta.$$

Let $\varepsilon > 0$ be such that

$$\frac{S(r)}{Y_r(t_0)} > 1 - W_2(r_2)^2 + \frac{\delta}{2}$$

on $[r_2 - \varepsilon, r_2]$. Now let W be a smooth function connecting W_1 to W_2 as follows: $W \equiv 1$ on $[r_1, r_2 - \varepsilon]$ and smoothly and monotonically decreases on $[r_2 - \varepsilon, r_2]$ to $W_2(r_2)$. Then on $[r_2 - \varepsilon, r_2]$,

$$\frac{S(r)}{Y_r(t_0)} > 1 - W_2(r_2)^2 + \frac{\delta}{2} \geq 1 - W(r)^2 + \frac{\delta}{2}.$$

Since $W = 1$ on $[r_1, r_2 - \varepsilon]$, we have thus shown that the technical condition in Theorem 2.9 is satisfied. This concludes the proof. $\square$

Remark 2.13 While the preceding results show that the gluing of Tolman–Bondi spacetimes can in principle be done under fairly general circumstances, the resulting construction will in general not be globally hyperbolic since the glued region may be very small. This can be seen in examples where a Tolman–Bondi spacetime is glued to Schwarzschild (cf. Corollary 2.12): If the glued region does not extend to past timelike infinity i^- (i.e., if $t_- \neq -\infty$ in the proof of Theorem 2.9), then one may find timelike curves emanating from i^- which are unable to cross into the glued region. Thus, the glued spacetime does not contain any Cauchy hypersurfaces.

To ensure global hyperbolicity, an investigation into the maximal interval of existence of the ODE in Proposition 2.5 is necessary.

3 Bartnik's "no-CMC" example

In this section, we present the cosmological spacetime without CMC Cauchy surfaces constructed by Bartnik [2]. An essential ingredient of the construction is the gluing Schwarzschild spacetime to flat FLRW. In this case, the ODE given in Proposition 2.9 may be solved explicitly, allowing for a simple gluing argument in terms of functions.

Proposition 3.1 (Explicit TB metrics). *Let $(D \times S^2, g)$ be a dust-filled Tolman–Bondi spacetime with marginal condition $W(r) \equiv 1$ and suppose the mass function is always positive, $S > 0$. Then, writing $M(r) = S(r)/2$, there exists a smooth function $t_0(r)$ such that*[4]

$$X(t,r) = \frac{M'(r)(t_0(r) - t) + 2M(r)t_0'(r)}{[6M(r)^2(t_0(r) - t)]^{\frac{1}{3}}}, \tag{3.1}$$

$$Y(t,r) = \left(\frac{9}{2}M(r)(t_0(r) - t)^2\right)^{\frac{1}{3}}. \tag{3.2}$$

Moreover, the metric can be extended smoothly to $\tilde{D} \times S^2$, where $\tilde{D} = \{(t,r) \in \mathbb{R}^2 : r \in \{M > 0\}, t \in (-\infty, t_0(r))\}$, provided that

$$\frac{d}{dr}\left(M(r)(t_0(r) - t)^2\right) > 0. \tag{3.3}$$

Conversely, given smooth functions $M(r)$ and $t_0(r)$ defined on $I \subset \mathbb{R}$ with $M > 0$ satisfying (3.3) (this is in particular the case if t_0 is strictly monotonically increasing and M is monotonically increasing, or vice versa), Eqs. (3.1) and (3.2) define a Tolman–Bondi spacetime on $\tilde{D} \times S^2$, where $\tilde{D}$ is given as before ($I = \{M > 0\}$).

Proof Since $W = 1$ and $S > 0$, by Eq. (2.9) $\dot{Y}$ is either always positive or always negative, w.l.o.g. let us assume that $\dot{Y} < 0$.[5] Then, by the same reference,

$$\dot{Y}(t,r) = -\sqrt{\frac{S(r)}{Y(t,r)}}.$$

By assumption $S(r) > 0$, so fixing some s_0 such that $(s_0, r) \in D$, $Y(t,r)$ has to agree with the unique solution $\tilde{Y}(t,r)$ of the following r-parameter family of ODEs:

$$\begin{cases} \dot{\tilde{Y}}(t,r) = -\sqrt{\frac{S(r)}{\tilde{Y}(t,r)}}, \\ \tilde{Y}(s_0, r) = Y(s_0, r). \end{cases}$$

4 We use M instead of S here to be consistent with the sources [4, 11] and [2]. The function $t_0(r)$ is an integration constant depending on r which arises from solving the ODE for Y.

5 This is the case for Schwarzschild and flat FLRW, cf. Example 2.7 and Example 2.8.

This ODE can be integrated explicitly and yields Eq. (3.2) with $M := S/2$ and

$$t_0(r) = s_0 + \frac{\sqrt{2}}{3\sqrt{M(r)}} Y(s_0, r)^{\frac{3}{2}}.$$

Evidently, the maximal existence interval for $Y(t,r)$ is $t \in (-\infty, t_0(r))$. The condition (3.3) guarantees that $X = Y' > 0$ on $\tilde{D}$. □

Let us now describe the construction: Let (M_1, g_1) be the following portion of Schwarzschild spacetime in Lemaître coordinates with mass $m = 1$, as described in Example 2.7: $M_1 = \{(t,r) \in \mathbb{R}^2 : t < r \leq 1 - \varepsilon\} \times S^2$ (for some small $\varepsilon > 0$). Since Schwarzschild has marginal condition $W = 1$, the functions X, Y are given in terms of functions $M(r)$ and $t_0(r)$ as described in Proposition 3.1. It is easily checked that

$$M(r) = 1, \quad t_0(r) = r. \tag{3.4}$$

Next, consider T^3 as a Riemannian quotient of $\mathbb{R}^3$ with the group action induced by $3\mathbb{Z}^3$, hence we can consider $(-\infty, 1) \times_{a_f} T^3$ as a Lorentzian quotient of flat FLRW $(-\infty, 1) \times_{a_f} \mathbb{R}^3$ with Friedmann mass parameter $\mathcal{M} = \frac{3}{4\pi}$. Note that we shift the upper limit of the time parameter by 1 for monotonicity reasons which will become clear in a moment. Let $x_0 \in T^3$ be arbitrary and $0 < 2\varepsilon < 1$. Then $B^{T^3}_{1+2\varepsilon}(x_0)$ is isometric to $B^{\mathbb{R}^3}_{1+2\varepsilon}(y_0) \subset \mathbb{R}^3$ (for any y_0 in the preimage of x_0 under the quotient map); hence, the spacetime portion $(-\infty, 1) \times_{a_f} (B^{T^3}_{1+2\varepsilon}(x_0) \backslash B^{T^3}_{1+\varepsilon}(x_0))$ can be written in Tolman–Bondi form, identical to flat FLRW. Let (M_2, g_2) be the spacetime $(-\infty, 1) \times_{a_f} (T^3 \backslash B^{T^3}_{1+\varepsilon}(x_0))$. Since flat FLRW also has marginal condition $W = 1$, it is also given explicitly according to Proposition 3.1 with functions

$$M(r) = r^3, \quad t_0(r) = 1. \tag{3.5}$$

We now glue (M_1, g_1) to (M_2, g_2) across the r-interval $[1-\varepsilon, 1+\varepsilon]$: Indeed, due to the explicit form of the functions M, t_0 in the two spacetimes, we may connect them smoothly across $[1-\varepsilon, 1+\varepsilon]$ while satisfying (on $(1-\varepsilon, 1+\varepsilon)$)[6]

$$t_0' > 0 \text{ and } M' > 0. \tag{3.6}$$

(This is the reason for shifting the time parameter of FLRW by 1, so that $t_0' > 0$ can be achieved.) As we argued in the proof of Proposition 3.1, $t_0' > 0$ and $M' > 0$ imply $Y, Y' = X > 0$, as well as $\rho = M'/4\pi Y^2 Y' \geq 0$. We denote the resulting spacetime by (M^+, g^+).

Lemma 3.2 *(M^+, g^+) is globally hyperbolic.*

[6] Bartnik [2] demands in addition that $t_0'' \leq 0$, $M'' \geq 0$ and $t_0^2 + M^2 > 0$. The latter is automatic, and the convexity/concavity conditions (while certainly achievable) do not appear to be necessary.

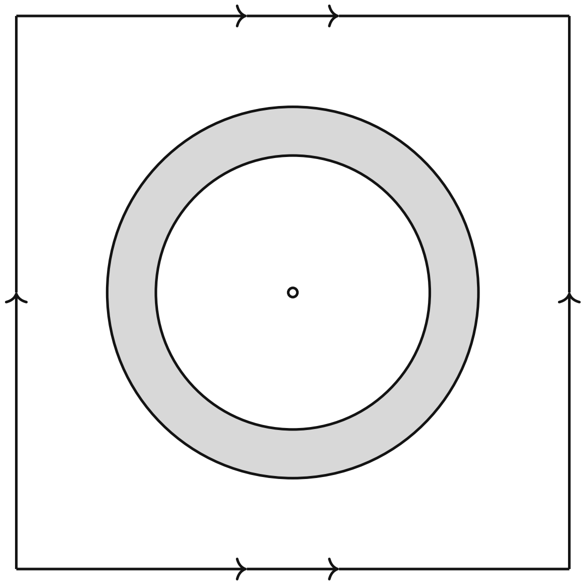


Fig. 2 A two-dimensional spatial cross section of (M^+, g^+). The inner and outer circles are the spheres $r = 1 - \varepsilon$ and $r = 1 + \varepsilon$, respectively. For $r < 1 - \varepsilon$, the spacetime is exactly given by the Schwarzschild spacetime. For $r > 1 + \varepsilon$, the spacetime is exactly given by a $k = 0$ FLRW dust spacetime; at large values r ceases to be a coordinate since we identify the sides. The shaded region between the spheres represents the gluing region. The point at the center represents $r = -\infty$

We omit the proof of this Lemma, since it is very similar to the one we give in the generalization to the hyperbolic case in Lemma 4.2.

Now let (M^-, g^-) be a time-reversed copy of (M^+, g^+). We extend them both to $r = -\infty$ (which corresponds to extending Schwarzschild to its event horizon) and glue them there[7]. The resulting Lorentzian manifold (M, g) is smooth and time-orientable due to the properties of the maximal extension of Schwarzschild. Observe that there exists a time-inverting isometry $\varphi : (M, g) \to (M, g)$ mapping (M^+, g^+) to (M^-, g^-) and vice versa. Now, since both (M^+, g^+) and (M^-, g^-) are globally hyperbolic, so is (M, g). (This can be seen by looking at the corresponding Penrose diagram, see [2, Fig. 2].) A Cauchy surface can be constructed by connecting the bifurcate horizon smoothly to one of the spatial slices in the each of the FLRW regions. [18, Cor. 14.54] can be used to show that this is in fact a Cauchy surface.

We will give the two proofs of Bartnik as to why (M, g) has no CMC Cauchy surfaces. The first is of topological nature and requires the following fundamental result of Schoen and Yau on the relationship between scalar curvature and topology on a Riemannian three-manifold.

Theorem 3.3 [19, Thm. 5.2] *Let N be a compact, oriented three-manifold. Suppose that one of the following holds:*

(i) *$\pi_1(N)$ contains a finitely generated, noncyclic Abelian subgroup.*
(ii) *$\pi_1(N)$ contains a subgroup which is isomorphic to the fundamental group of a surface of genus $g > 1$.*

Then N admits no Riemannian metric of positive scalar curvature. Moreover, every Riemannian metric on N of nonnegative scalar curvature is flat.

The second proof follows from much simpler Lorentz-geometric arguments and is more generally applicable, see the subsequent section. It requires the following CMC foliation theorem due to Bartnik.

[7] One could have also started out with two sides of the Schwarzschild horizon in the Kruskal extension and then glued FLRW to both of them in Lemaître and time-reversed Lemaître coordinates, and it would have amounted to the same construction.

Theorem 3.4 [2] *Let $(\tilde{M}, \tilde{g})$ be a globally hyperbolic spacetime, with compact Cauchy surfaces, satisfying the strong energy condition. Suppose $S \subset \tilde{M}$ is a compact maximal Cauchy surface. Then a neighborhood of S is foliated with CMC Cauchy surfaces. Moreover, if every such foliation is by maximal surfaces, then the spacetime is isometric to an open subset of $(\mathbb{R} \times S, -dt^2 + h)$, where h is the induced Riemannian metric on S. In particular, the spacetime is static.*

Let us now give the announced proofs:

Theorem 3.5 [2] *The constructed spacetime (M, g) is globally hyperbolic, timelike geodesically incomplete and satisfies the strong energy condition. Its Cauchy surfaces are topologically $T^3 \# T^3$ and it does not contain any CMC Cauchy surfaces.*

Proof The only part left to prove is the nonexistence of CMC Cauchy surfaces.

(i) *Topological proof*: Let $\varphi : (M, g) \to (M, g)$ be the time-inverting isometry. Suppose $S \subset M$ is a compact CMC Cauchy surface. If $H_S \neq 0$, since $H_{\varphi(S)} = -H_S$, we may assume w.l.o.g. that $H_S > 0$. We claim that $\varphi(S) \subset I^+(S)$. Indeed: Suppose $\varphi(S)$ is not a subset of $I^+(S)$. Then there is a point $p \in \varphi(S)$ such that $p \notin I^+(S)$. An inextendible future timelike curve through p must intersect S, say at q. Let γ be a maximizing future timelike geodesic from p to q. By the Brill–Flaherty uniqueness result (see [2, Eq. (2)]), we have

$$H_{\varphi(S)}(p) - H_S(q) \geq \int \mathrm{Ric}(\gamma', \gamma').$$

The right-hand side is ≥ 0, but the left-hand side is < 0, which is a contradiction. So $\varphi(S) \subset I^+(S)$. Now we invoke the maximum principle for prescribed mean curvature [1, 15] to conclude the existence of a maximal Cauchy surface $S_0 \subset I^+(S) \cap I^-(\varphi(S))$. Since $\pi_1(S_0) \cong \pi_1(T^3 \# T^3) \cong \mathbb{Z}^3 * \mathbb{Z}^3$, it contains the subgroup $\mathbb{Z}^2$ which is noncyclic, finitely generated and Abelian. Thus, by Theorem 3.3, the induced metric h on S_0 can only have nonnegative scalar curvature if it is flat. But by the constraint equation for scalar curvature (see [8, Eq. (G.2)]),

$$R_h = 16\pi\rho + |K|_h^2 - H_{S_0}^2,$$

where K is the second fundamental form. By maximality, $H_{S_0} = 0$, so $R_h \geq 0$. We conclude that h is flat, which implies $R_h = 0$. But ρ is nonzero, e.g., in the FLRW parts of the spacetime, which gives the desired contradiction.

(ii) *Lorentz-geometric proof*: Suppose $S \subset M$ is a compact CMC Cauchy surface, and suppose for the moment that $H_S \neq 0$. Without loss of generality, we may assume that $H_S > 0$. By Hawking's singularity theorem, each past inextendible timelike curve starting in S has finite Lorentzian arclength. However, this is contradicted, e.g., by the "Lemaître-vertical" curves in Schwarzschild (one could also argue via the infinitely long verticals in the FLRW parts): Consider $\gamma : t \mapsto (t, r_0, p_0)$ (with $r_0 < 1 - \varepsilon$, $p_0 \in S^2$ arbitrary) in the (M^+, g^+)-part of (M, g): We may extend γ so far into the future that it meets S, but already the Schwarzschild part has infinite Lorentzian arclength, a contradiction. So the CMC Cauchy surface S must be maximal, $H_S = 0$. By Theorem 3.4, a neighborhood of S is foliated by CMC Cauchy surfaces, all of which must again be maximal. Hence, by the same reference, (M, g) is static, which is manifestly not the case.

□

Remark 3.6 Note that the Lorentz-geometric proof does not require the FLRW part to be spatially T^3. In fact, any compact Riemannian quotient of $\mathbb{R}^3$ does the job. Nor is the application of the "side-switch" isometry really necessary: Indeed, if $H_S > 0$, we use the infinitely long past curves in the (M^+, g^+)-side to arrive at a contradiction to Hawking's singularity theorem, and if $H_S < 0$, in a similar manner we use the infinitely long future curves in the (M^-, g^-)-side.

Remark 3.7 In [9], Chruściel, Isenberg and Pollack use their results on the gluing of initial data sets to construct a globally hyperbolic spacetime with $\mathrm{Ric} = 0$ and compact Cauchy surfaces of topology $T^3 \# T^3$, that does not contain any CMC Cauchy surfaces. While Bartnik's example is manifestly timelike geodesically incomplete, this is unclear in their example. However, null geodesic incompleteness has been established, cf. [5, 6]. It would be interesting to see whether similar generalizations as the ones we give in the next section could be achieved using initial data gluing methods.

4 Generalizations of Bartnik's "no-CMC" example

In this section, we generalize Bartnik's construction by gluing Schwarzschild to hyperbolic FLRW models which have compact spacelike slices. Here, the ODE in Proposition 2.9 (with $W > 1$) is not explicitly solvable, but implicit integration will turn out to be sufficient for our purposes. The properties are very similar to the flat case ($W = 1$). It turns out that one cannot do this construction with spherical FLRW, because if $W < 1$ the ODE develops finite time past singularities, giving an obstruction to global hyperbolicity.

Proposition 4.1 *Let $(M = D \times S^2, g)$ be a dust-filled Tolman–Bondi spacetime with $\dot{Y} < 0$, monotonically increasing marginal condition $W \geq 1$ and monotonically increasing positive mass function $S > 0$. Then there exists a smooth function $t_0(r)$ such that the metric can be extended smoothly to $\tilde{D} \times S^2$, where $\tilde{D} = \{(t, r) \in \mathbb{R}^2 : r \in \{S > 0\}, t \in (-\infty, t_0(r))\}$. Conversely, for any choice of strictly monotonically increasing smooth function $t_0(r)$, monotonically increasing positive functions $S(r)$ and $W(r)$ (with $W(r) \geq 1$), the unique solution of the ODE family $\dot{Y} = -\sqrt{W^2 - 1 + S/Y}$ with initial condition determined by t_0 defines a Tolman–Bondi spacetime (together with $X := Y'/W$) on $\tilde{D} \times S^2$. In particular, $Y' > 0$ on $\tilde{D}$. Whenever $W(r) = 1$, the solution is given explicitly as described in Proposition 3.1, and whenever $W(r) > 1$, the solution is described implicitly by*

$$t_0(r) - t = F(W(r), S(r), Y(t, r)).$$

In this case, $F = F(x, y, z)$ is the function

$$F(x, y, z) = \frac{\sqrt{x^2 - 1 + \frac{y}{z}}\,z}{x^2 - 1} - \frac{y \operatorname{arcoth}\left(\sqrt{1 + \frac{y}{z(x^2-1)}}\right)}{(x^2 - 1)^{\frac{3}{2}}}.$$

Proof By Proposition 2.5, Y satisfies

$$\dot{Y}(t, r) = -\sqrt{W(r)^2 - 1 + \frac{S(r)}{Y(t, r)}}.$$

This ODE is uniquely solvable upon fixing some initial data $Y(s_0, r)$ for every $r \in \{S > 0\}$. If $W(r) = 1$, we are in the situation of Proposition 3.1, so we may assume $W > 1$. Let $(a, b) = (a(r), b(r))$ be the maximal existence interval for t when considering Y as a solution of this ODE. Clearly $a = -\infty$ for every r, since Y is monotonically decreasing in t and $\dot{Y}$ is bounded for $Y \to +\infty$. To determine $b = b(r)$, we separate variables in the above ODE and integrate:

$$t_0(r) - t = \frac{\sqrt{W(r)^2 + \frac{S(r)}{Y(t,r)}}\,Y(t, r)}{W(r)^2 - 1} - \frac{S(r) \operatorname{arcoth}\left(\sqrt{1 + \frac{S(r)}{Y(t,r)(W(r)^2-1)}}\right)}{(W(r)^2 - 1)^{\frac{3}{2}}}. \quad (*)$$

Here, $t_0(r)$ is an integration constant depending on r, which can be determined by evaluating $(*)$ at the initial time $t = s_0$. Note that by monotonicity, $Y(t, r) \to 0$ as $t \to b(r)$. Using that $\operatorname{arcoth}(x) \to 0$ as $x \to \infty$, the right-hand side of $(*)$ tends to 0 as $t \to b(r)$. So, necessarily, $b(r) = t_0(r)$.

It is useful to view the right-hand side of $(*)$ as $F(W(r), S(r), Y(t, r))$, where $F = F(x, y, z)$ is a function in three variables defined above and which satisfies, in light of the integration of the ODE,

$$\partial_z F(x, y, z) = \frac{1}{\sqrt{x^2 - 1 + \frac{y}{z}}}.$$

In particular, $\partial_z F(W(r), S(r), Y(t, r)) > 0$. Differentiating $(*)$ with respect to r and solving for Y' yields

$$Y'(t, r) = \frac{1}{\partial_z F}\left(t_0'(r) - \partial_x F \cdot W'(r) - \partial_y F \cdot S'\right).$$

Since $t_0(r) = F(W(r), S(r), Y(s_0, r)) + s_0$, it follows that

$$Y'(t, r) = \frac{1}{\partial_z F}((\partial_x F|_{s_0} - \partial_x F|_t)W' + (\partial_y F|_{s_0} - \partial_y F|_t)S' + \partial_z F|_{s_0} Y'(s_0, r)).$$

Since $Y'(s_0, r) > 0$ (by positivity of W), it follows that $(\partial_z F|_{s_0})Y'(s_0, r) > 0$. We claim that the other summands are positive as well (recall that $W', S' \geq 0$), as long as

$t < s_0$: Indeed,

$$\partial_t(\partial_x F|_t) = \partial_x \partial_z F|_t \dot{Y}(t, x),$$
$$\partial_t(\partial_y F|_t) = \partial_y \partial_z F|_t \dot{Y}(t, x).$$

One checks that

$$\partial_x \partial_z F(x, y, z) = -x \left(x^2 - 1 + \frac{y}{z}\right)^{-\frac{3}{2}} < 0,$$
$$\partial_y \partial_z F(x, y, z) = -\frac{1}{2z} \left(x^2 - 1 + \frac{y}{z}\right)^{-\frac{3}{2}} < 0.$$

Hence, together with $\dot{Y} < 0$ we get $\partial_t(\partial_x F(W(r), S(r), Y(t, r))) > 0$ and $\partial_t(\partial_y F(W(r), S(r), Y(t, r))) > 0$. Moreover, one can check explicitly that $\partial_x F \to 0$ and $\partial_y F \to 0$ as $z \to 0$, so if the initial condition satisfies

$$\partial_z F|_{s_0} Y'(s_0, r) + \partial_x F|_{s_0} W' + \partial_y F|_{s_0} S' > 0,$$

or equivalently $t_0' > 0$, then $Y'(t, r) > 0$ for all r and for all $t \in (-\infty, t_0(r))$. □

Let us now generalize the construction in the previous section (up to appropriate choices of parameters) to the gluing of arbitrary (compactified) flat or hyperbolic FLRW spacetimes to Schwarzschild. Since for general marginal conditions W, dust-filled Tolman–Bondi spacetimes cannot be written as explicitly as in the case $W = 1$ (cf. Proposition 3.1), we have to rely on our general gluing result (Theorem 2.9) in conjunction with the above analysis on the maximal solution intervals of the ODE in Proposition 2.5.

Let us now dive into the details: Let N be either $\mathbb{R}^3$ or H^3, let Q be a compact Riemannian quotient manifold of N, $q : N \to Q$ the quotient map. Let $x_0 \in Q$ and $0 < 3r_2$ be such that $B^Q_{3r_2}(x_0) \subset Q$ is isometric to $B^N_{3r_2}(y_0)$ (for any $y_0 \in q^{-1}(x_0)$). Let (M_2, g_2) be the spacetime $(-\infty, r_2) \times_a (Q \backslash B^Q_{r_2}(x_0))$ with $a = a_f$ resp. $a = a_h$ if $N = \mathbb{R}^3$ resp. $N = H^3$. (We again shift the time parameter for monotonicity reasons.) Then the portion $(-\infty, r_2) \times_a (B^Q_{3r_2}(x_0) \backslash B^Q_{r_2}(x_0))$ can be written in Tolman–Bondi form, identical to flat or hyperbolic FLRW. Let (M_1, g_1) be Schwarzschild spacetime in Lemaître coordinates up to a radial coordinate value $r \leq r_1 < r_2$. For fixed Schwarzschild mass $m > 0$, in the flat case the Friedmann mass $\mathcal{M}$ needs to be chosen such that

$$S_1(r_1) = 2m \leq \frac{8\pi \mathcal{M}}{3} r_2^3 = S_2(r_2),$$

and in the hyperbolic case, we need

$$S_1(r_1) = 2m \leq \frac{8\pi \mathcal{M}}{3} \sinh(r_2)^3.$$

Then it is possible to glue (M_1, g_1) to (M_2, g_2) along the interval (r_1, r_2) by prescribing monotonically increasing connecting functions W, S, and a strictly monotonically increasing connecting function t_0 for the corresponding functions on either side, cf. Proposition 4.1. We call the resulting spacetime (M^+, g^+).

Lemma 4.2 *(M^+, g^+) is globally hyperbolic.*

Proof For simplicity, we give the argument for global hyperbolicity if the gluing procedure is done without any quotient operation on the FLRW part in order to use global Tolman–Bondi spacetime arguments, and the compactified case can be proved along similar lines.[8]

Keeping this in mind, (M^+, g^+) is a warped product $D \times_Y S^2$, so it suffices to check global hyperbolicity of the two-dimensional spacetime $(D, -dt^2 + X^2dr^2)$, where $D = \{(t, r) \in \mathbb{R}^2 : t < t_0(r)\}$ (cf. [3, Thm. 3.68]). For small enough $\delta > 0$, we claim that $\Sigma := \{t = t_0(r) - \delta\}$ is a Cauchy surface. It is clearly a smooth, spacelike hypersurface. Let $\gamma(s) := (t(s), r(s))$, $s \in (0, 1)$, be an arbitrary C^1-causal curve. We will show that if γ does not meet Σ, its limits at $s = 0$ or $s = 1$ exist. The acausality of Σ can be proven along similar lines and will be omitted.

Suppose first that $t(s) < t_0(r(s)) - \delta$. By monotonicity, $t(s)$ has a limit T as $s \to 1$, with $T \le r_2 - \delta$ (since $t_0 \le r_2$ by construction). Again by causality,

$$|\dot r(s)| \le \frac{\dot t(s)}{X(t(s), r(s))}.$$

Let $R^+ := \limsup_{s\to 1} r(s)$. Suppose that $R^+ = +\infty$, and let $s_{2,n}$ be a realizing sequence, i.e., $r(s_{2,n}) \to +\infty$. Then the points $(t(s_{2,n}), (r(s_{2,n})))$ are eventually in the hyperbolic FLRW part of M^+, where $X(t, r) = a_h(t)$. Let s_1 be such that this is true for all $s \in (s_1, s_{2,n})$. Using that $a_h' \ge 1$ by Friedmann's equation, we get

$$\dot r(s) \le |\dot r(s)| \le \frac{\dot t(s)}{a_h(t(s))} \le \frac{\dot t(s)}{a_h(t(s))} a_h'(t(s)) = \partial_s \log(a_h(t(s))).$$

Integrating from s_1 to $s_{2,n}$, we see that $r(s_{2,n}) \to +\infty$ is impossible. Thus, $R^+ < +\infty$ and due to $T \le t_0(R^+) - \delta < t_0(R^+)$, $(T, R^+) \in D$. Now let $R^- := \liminf_{s\to 1} r(s)$. Suppose that $R^- = -\infty$ and let $s_{1,n}$ be a realizing sequence. In this case, $(t(s_{1,n}), r(s_{1,n}))$ is eventually in the Schwarzschild part, so one can argue just as in the case below to show $R^- > -\infty$. To show that $R^+ = R^-$, observe that $\min_{s\in[s_{1,n}, s_{2,n}]} W(r(s))^{-1}$ is bounded below by some constant $C > 0$ independently

[8] Indeed, if the spatial FLRW slice is a quotient Q of H^3, consider the universal cover which is partitioned by fundamental domains. Using analogous arguments as the ones appearing in this proof, it can be shown that the future end of the lift to the spacetime universal cover of any inextendible causal curve lies in just one fundamental domain, likewise for the past end. Since the fundamental domains are portions of a Tolman–Bondi spacetime, the arguments given in this proof apply.

of n. Then, using causality and the fact that $\partial_t Y < 0$,

$$(*) \qquad t(s_{2,n}) - t(s_{1,n}) \geq \int_{s_{1,n}}^{s_{2,n}} \frac{\partial_r Y(t(s), r(s))}{W(r(s))} |\dot{r}(s)| \mathrm{d}s$$
$$\geq C(Y(t(s_{2,n}, r(s_{2,n}))) - Y(t(s_{1,n}, r(s_{1,n})))).$$

Taking first $s_{2,n} \to 1$ and then $s_{1,n} \to 1$ gives $T - T = 0$ on the left-hand side, thus giving

$$Y(T, R^-) \geq Y(T, R^+).$$

Since $\partial_r Y > 0$, we get $R^+ = R^-$. Thus, γ is extendible to $s = 1$.

The other possibility is that $t(s) > t_0(r(s)) - \delta$ for all $s \in (0, 1)$. We will show that γ is extendible to $s = 0$. Again by monotonicity, $t(s)$ has a limit T as $s \to 0$ (a priori $T \geq -\infty$). Suppose that indeed $T = -\infty$. Let $R^- := \liminf_{s \to 0} r(s)$. It is impossible that $R^- > -\infty$, since $t_0(r(s)) - \delta < t(s) < t_0(r(s))$. So $R^- = -\infty$. In particular, we may choose a realizing sequence $r(s_n) \to -\infty$ for $s_n \to 0$. Then $\gamma(s_n)$ is eventually in the Schwarzschild region of (M^+, g^+). Thus, causality together with $t(s) > t_0(r(s)) - \delta$ imply the estimate (using the explicit form of the Tolman–Bondi metric for Schwarzschild, cf. Example 2.7)

$$\dot{t}(s_n)^2 \geq \left(\frac{2m^{1/3}}{(6\delta)^{1/3}} \right)^2 \dot{r}(s_n)^2.$$

Note that everything on the right-hand side is bounded independently of γ. If δ was chosen small enough, then $t(s_n) \to -\infty$ much faster than $r(s_n) \to -\infty$; thus, the inequality $t(s_n) > t_0(r(s_n)) - \delta = r(s_n) - \delta$ cannot be maintained. The conclusion is that $T > -\infty$. From here, one can use $(*)$ and the arguments in the preceding case to see that $+\infty > R^+ := \limsup_{s \to 0} r(s) = R^- =: R$. Moreover, the function $t(s) - t_0(r(s))$ is easily seen to be monotonically increasing in s because $t_0'(r(s)) \leq X(t(s), r(s)) = Y'(t(s), r(s))/W(r(s))$ and $\gamma(s)$ is causal. Hence, $T = t_0(R)$ is impossible. So $T < t_0(R)$; thus, $(T, R) \in D$. This shows extendibility of $\gamma(s)$ to $s = 0$. □

Similarly, let $\tilde{Q}$ be a compact quotient of either $\mathbb{R}^3$ or H^3, and construct the analogous globally hyperbolic spacetime (M^-, g^-), but with inverted time orientation. We extend both to $r = -\infty$ and glue them there to obtain a spacetime (M, g), which is also globally hyperbolic.

Theorem 4.3 (Examples of "no-CMC" spacetimes). *Let Q, $\tilde{Q}$ be compact Riemannian quotients[9] of $\mathbb{R}^3$ or H^3. Let (M, g) be the spacetime constructed above by gluing Schwarzschild (of any mass m) to flat or hyperbolic FLRW which is spatially Q (with Friedmann mass $\mathcal{M}_1 = \mathcal{M}_1(m)$), and then attaching a time-inverted gluing*

[9] All combinations are allowed: Q, $\tilde{Q}$ both quotients of $\mathbb{R}^3$, both quotients of H^3, or one a quotient of $\mathbb{R}^3$ and the other a quotient of H^3.

of Schwarzschild to flat or hyperbolic FLRW which is spatially $\tilde{Q}$ (with Friedmann mass $\mathcal{M}_2 = \mathcal{M}_2(m)$) along the Schwarzschild event horizon. Then (M, g) is globally hyperbolic with topologically $\tilde{Q}\#Q$ Cauchy surfaces, is timelike geodesically incomplete, satisfies the strong and dominant energy conditions, and does not contain any CMC Cauchy surfaces.

Proof Evidently, (M^+, g^+) and (M^-, g^-) contain infinitely long past and future timelike curves, respectively, so the proof may be carried out just like in the Lorentz-geometric proof of Theorem 3.5.

In the symmetric case of $Q = \tilde{Q}$ being a quotient of either $\mathbb{R}^3$ or H^3 (i.e., (M^-, g^-) is just a time-reversed copy of (M^+, g^+)), as well as Q oriented, a proof based on Theorem 3.3 may also be given. (In the hyperbolic case, we utilize the positive resolution of the surface subgroup conjecture [16].)

If Q is a quotient of $\mathbb{R}^3$, then $\pi_1(Q)$ contains a subgroup which is isomorphic to the fundamental group of a surface of genus $g = 1$. If Q is a quotient of H^3, then $\pi_1(Q)$ contains a subgroup which is isomorphic to the fundamental group of a surface of genus $g \geq 2$. In either case, $\pi_1(Q\#Q) \cong \pi_1(Q) * \pi_1(Q)$ contains a subgroup which is isomorphic to the fundamental group of a genus $g \geq 1$ surface. One proceeds as in the topological proof of Theorem 3.5, but uses Theorem 3.3(i) in the $\mathbb{R}^3$ case and Theorem 3.3(ii) in the hyperbolic case. □

Remark 4.4 (Higher spacetime dimensions). We have done our analysis of Tolman–Bondi spacetimes and constructed spacetimes without CMC Cauchy surfaces in spacetime dimension $d = 4$ in order to be consistent with the literature [2, 4, 11]. However, all of our gluing and ODE arguments continue to hold for d-dimensional Tolman–Bondi spacetimes $D \times S^{d-2}$, where the metric is of the form

$$g = -dt^2 + X(t, r)^2 dr^2 + Y(t, r)^2 d\Omega^2_{d-2}.$$

Here $d\Omega^2_{d-2}$ denotes the standard round metric on S^{d-2}. Examples of this metric are FLRW spacetimes in d dimensions, as well as the d-dimensional Schwarzschild–Tangherlini spacetime of mass $m > 0$ ($R \neq 0, 2m$)

$$g_S = -\left(1 - \frac{2m}{R^{d-3}}\right) d\tau^2 + \frac{1}{1 - \frac{2m}{R^{d-3}}} dR^2 + R^2 d\Omega^2_{d-2}.$$

Just like in $d = 4$ spacetime dimensions, one can introduce Lemaître coordinates (t, r) via

$$dt = d\tau + \sqrt{\frac{2m}{R^{d-3}}} dR^2,$$
$$dr = d\tau + \sqrt{\frac{R^{d-3}}{2m}} dR^2,$$

in which the metric takes the form

$$g_S = -dt^2 + (2m)^{\frac{2}{d-1}} \left(\frac{d-1}{2}(r-t) \right)^{\frac{6-2d}{d-1}} dr^2 + (2m)^{\frac{2}{d-1}} \left(\frac{d-1}{2}(r-t) \right)^{\frac{4}{d-1}} d\Omega^2_{d-2}.$$

One can then proceed to glue d-dimensional spatially compact flat or hyperbolic FLRW models to Schwarzschild–Tangherlini (where, as before, the Friedmann mass parameter $\mathcal{M}$ has to be chosen accordingly in dependence on the Schwarzschild mass m), and attach two such gluings (one of them time-reversed) along the event horizon. The resulting spacetime can be shown to have no CMC Cauchy surfaces via the same Lorentz-geometric arguments as the ones used in the proofs of Theorems 3.5 and 4.3. Note that the topological proof is not applicable, as that relies on Theorem 3.3 which is a result for three-manifolds. Let us summarize these observations in following corollary.

Corollary 4.5 (Higher-dimensional "no-CMC" examples). *Let $d \geq 4$ and let $Q, \tilde{Q}$ be compact Riemannian quotients of $\mathbb{R}^{d-1}$ or H^{d-1}. Let (M, g) be the spacetime constructed by gluing d-dimensional Schwarzschild–Tangherlini (of mass $m > 0$) to flat or hyperbolic d-dimensional FLRW which is spatially Q (with Friedmann mass $\mathcal{M}_1 = \mathcal{M}_1(m)$), and then attaching a time-inverted gluing of d-dimensional Schwarzschild–Tangherlini to flat or hyperbolic d-dimensional FLRW which is spatially $\tilde{Q}$ (with Friedmann mass $\mathcal{M}_2 = \mathcal{M}_2(m)$) along the Schwarzschild–Tangherlini event horizon. Then (M, g) is globally hyperbolic,* $\dim M = d$*, with Cauchy surfaces of topology $\tilde{Q}\#Q$, is timelike geodesically incomplete, satisfies the strong and dominant energy conditions and does not contain any CMC Cauchy surfaces.*

Remark 4.6 (The spherical case). When gluing Schwarzschild to spherical FLRW, the function W is eventually < 1, so $W^2 - 1 < 0$. This causes the solution of the ODE

$$\dot{Y}(t, r) = -\sqrt{W(r)^2 - 1 + \frac{S(r)}{Y(t, r)}}$$

to develop a past singularity in finite time. So the glued spacetime cannot be globally hyperbolic, as one may simply take inextendible curves emerging from Schwarzschild past timelike infinity i^- which cannot cross into the glued region. Of course, one can remove portions of the Schwarzschild region to make it globally hyperbolic, but in this case, the removed portion would be so large that the resulting spacetime could no longer contain timelike curves with infinite length *and* be a neighborhood of the event horizon ($T = -X$ in Kruskal coordinates). Indeed, if both could be achieved, then we could glue the resulting spacetime with a time-inverted copy of itself to produce a globally hyperbolic spacetime with spatial topology $S^3\#S^3 \cong S^3$ that would contain timelike curves with infinite length. Since the resulting spacetime is spherically symmetric, this contradicts a result of Burnett [7, Thm. 1].

On the other hand, one could do the full time-inverted gluing in the spherical case without first removing a portion of the Schwarzschild spacetime to obtain a spacetime (M, g) which is not globally hyperbolic, then restrict attention to the Cauchy development $(\tilde{M}, \tilde{g})$ of some spacelike hypersurface of topology $S^3\#S^3 \cong S^3$, thus resolving the issue of global hyperbolicity. Of course the methods of proof used so far do not apply to show that $(\tilde{M}, \tilde{g})$ has no CMC Cauchy surfaces, as they rely on the existence of infinitely long timelike curves, so it would be interesting to determine if $(\tilde{M}, \tilde{g})$ contains a CMC Cauchy surface or not.

Also, let (M^+, g^+) be a toroidal FLRW glued to Schwarzschild, and let (M^-, g^-) be a time-inverted copy of a spherical FLRW glued to Schwarzschild. One can imagine gluing these spacetimes along the event horizon. However, like above, to achieve global hyperbolicity, portions of the manifold would need to be removed. The resulting spacetime has toroidal spatial topology $T^3 \cong T^3 \# S^3$; however, given the remarks above, we cannot conclude that the spacetime has no CMC Cauchy surface since the Lorentz-geometric proof does not carry over: Based on the construction, there are either timelike curves that have infinite length to the past or to the future but not both. It would be interesting to find cosmological spacetimes without CMC Cauchy surfaces whose spatial topologies are T^3, or prove that none can exist.

5 Conclusion and outlook

In this work, we enlarge the number of known spatial topologies for cosmological spacetimes (i.e., spacetimes with compact Cauchy surfaces and satisfying the strong energy condition) without CMC Cauchy surfaces. Specifically, we show that if Q is any compact Euclidean or hyperbolic three-manifold and $\tilde{Q}$ is any other compact Euclidean or hyperbolic three-manifold, then there are cosmological spacetimes of that type with spatial topologies $Q \# \tilde{Q}$. (All of these examples are manifestly timelike incomplete, see Conjecture 1.4.) To obtain our examples, we generalize a gluing construction of Bartnik [2]. We glue general Tolman–Bondi spacetimes with variable marginal conditions and use that to construct cosmological spacetimes without CMC Cauchy surfaces. Bartnik's original construction comes from gluing a Schwarzschild Tolman–Bondi spacetime to a flat FLRW spacetime; the gluing procedure in this case is simplified since the marginal condition satisfies $W = 1$ and so the Einstein equations can be integrated, see Proposition 3.1. For the hyperbolic case, the marginal condition satisfies $W \geq 1$, and the Einstein equations can no longer be integrated; however, maximal solutions can still be constructed implicitly, see Proposition 4.1.

Our main result, Theorem 4.3, utilizes two different arguments (both due to Bartnik) to establish the nonexistence of a CMC Cauchy surface. The first is a topological argument using well-known results from Schoen and Yau which forbid certain three-manifolds from having nonnegative scalar curvature. This argument is purely at the initial data level. The second argument is more Lorentz-geometric and uses Hawking's cosmological singularity theorem, but it requires global knowledge of the spacetime. The topological argument only works when $Q \cong \tilde{Q}$, and we rely on the positive resolution of the surface subgroup conjecture [16] when Q is a hyperbolic three-manifold. The Lorentz-geometric argument works in all cases, but it relies on knowledge of the global spacetime. We describe in Remark 4.4 and Corollary 4.5 how our arguments generalize to produce analogous examples in arbitrary spacetime dimensions. The vacuum cosmological spacetime with spatial topology $T^3 \# T^3$ constructed by Chruściel, Isenberg and Pollack in [9] makes use of the topological proof alluded here. In light of the positive resolution of the surface subgroup conjecture, it would be interesting if the example in [9] generalizes to the more general topologies considered here.

Possible future directions of research include Tolman–Bondi metrics with more general stress–energy tensors (e.g., radiation models), compatibility with results obtained by initial data gluing methods, as well as attempts to find "no-CMC" cosmological spacetimes where the timelike incompleteness is less obvious (so as to gain more insight into Conjecture 1.4 resp. Conjecture 1.2). Also, models with only axial symmetry could be of interest. Such models may have a chance to produce spacetimes with non-CMC Cauchy surfaces provided Burnett's result [7, Thm. 1] does not generalize to the axially symmetric setting. Lastly, as alluded to at the end of Remark 4.6, our methods cannot be used to establish the nonexistence of CMC Cauchy surfaces in cosmological spacetimes with toroidal spatial topology T^3. It would be interesting to find such examples, if any exist.

Acknowledgements Eric Ling is supported by Carlsberg Foundation CF21-0680 and Danmarks Grundforskningsfond CPH-GEOTOP-DNRF151. Argam Ohanyan is funded by the ÖAW-DOC scholarship of the Austrian Academy of Sciences and by the project P-33594 of the Austrian Science Fund FWF. He is grateful for the hospitality of the Mathematics Department of the University of Copenhagen, where parts of this work were conducted. The authors appreciate the hospitality of the Fields Institute and the ESI during the "Thematic Program on Nonsmooth Riemannian and Lorentzian Geometry" and the "Conference on Non-Regular Spacetime Geometry," respectively. They thank Piotr Chruściel, Gregory Galloway, Michael Kunzinger and Roland Steinbauer for helpful discussions and valuable feedback.

This research was funded in part by the Austrian Science Fund (FWF) [Grant DOI 10.55776/P-33594]. For open access purposes, the authors have applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

Funding Open access funding provided by Copenhagen University.

References

1. Bartnik, R.: Existence of maximal surfaces in asymptotically flat spacetimes. Commun. Math. Phys. **94**(2), 155–175 (1984)
2. Bartnik, R.: Remarks on cosmological spacetimes and constant mean curvature surfaces. Commun. Math. Phys. **117**(4), 615–624 (1988)
3. Beem, J., Ehrlich, P., Easley, K.: Global Lorentzian Geometry. Marcel Dekker (1996)
4. Bondi, H.: Spherically symmetrical models in general relativity. Mon. Not. R. Astron. Soc. **107**(5–6), 410–425 (1947)
5. Burkhart, M., Lesourd, M., Pollack, D.: Null geodesic incompleteness of spacetimes with no CMC Cauchy surfaces. Pure Appl. Math. Q. **15**(3), 839–849 (2019)
6. Burkhart, M., Pollack, D.: Causal geodesic incompleteness of spacetimes arising from imp gluing. Gen. Relativ. Gravit. **51**, 1–19 (2019)
7. Burnett, G.A.: Lifetimes of spherically symmetric closed universes. Phys. Rev. D **51**(4), 1621 (1995)
8. Chruściel, P.T.: Geometry of black holes, vol. 169. Oxford University Press (2020)

9. Chruściel, P.T., Isenberg, J., Pollack, D.: Initial data engineering. Commun. Math. Phys. **257**(1), 29–42 (2005)
10. Dilts, J., Holst, M.: When do spacetimes have constant mean curvature slices? (2017). arXiv preprint arXiv:1710.03209
11. Eardley, D.M., Smarr, L.: Time functions in numerical relativity: marginally bound dust collapse. Phys. Rev. D **19**(8), 2239 (1979)
12. Galloway, G.J.: Existence of CMC Cauchy surfaces and spacetime splitting. Pure Appl. Math. Q. **15**(2), 667–682 (2019)
13. Galloway, G.J., Ling, E.: Existence of CMC Cauchy surfaces from a spacetime curvature condition. Gen. Relativ. Gravit. **50**, 1–7 (2018)
14. Galloway, G.J., Ling, E.: Remarks on the existence of CMC Cauchy surfaces. In: Developments in Lorentzian Geometry, pp. 93–104. Springer International Publishing, Cham (2022)
15. Gerhardt, C.: H-surfaces in Lorentzian manifolds. Commun. Math. Phys. **89**, 523–553 (1983)
16. Kahn, J., Markovic, V.: Immersing almost geodesic surfaces in a closed hyperbolic three manifold. Ann. Math. 1127–1190 (2012)
17. Minguzzi, E.: Lorentzian causality theory. Living Rev. Relativ. **22**(1), 1–202 (2019)
18. O'Neill, B.: Semi-Riemannian geometry with applications to relativity. Academic press, London (1983)
19. Schoen, R., Yau, S.-T.: Existence of incompressible minimal surfaces and the topology of three dimensional manifolds with non-negative scalar curvature. Ann. Math. **110**(1), 127–142 (1979)

Letters in Mathematical Physics (2023) 113:48
https://doi.org/10.1007/s11005-023-01668-w

The splitting theorem for globally hyperbolic Lorentzian length spaces with non-negative timelike curvature

Tobias Beran[1] · Argam Ohanyan[1] · Felix Rott[1] · Didier A. Solis[2]

Received: 10 November 2022 / Revised: 20 March 2023 / Accepted: 28 March 2023 /
Published online: 18 April 2023

Abstract

In this work, we prove a synthetic splitting theorem for globally hyperbolic Lorentzian length spaces with global non-negative timelike curvature containing a complete timelike line. Inspired by the proof for smooth spacetimes (Beem et al. in Global differential geometry and global analysis 1984, Springer, pp. 1–13, 1985), we construct complete, timelike asymptotes which, via triangle comparison, can be shown to fit together to give timelike lines. To get a control on their behaviour, we introduce the notion of parallelity of timelike lines in the spirit of the splitting theorem for Alexandrov spaces as proven in Burago et al. (A course in metric geometry, vol 33, American Mathematical Society, Providence, 2001) and show that asymptotic lines are all parallel. This helps to establish a splitting of a neighbourhood of the given line. We then show that this neighbourhood has the *timelike completeness* property and is hence inextendible by a result in Grant et al. (Ann Glob Anal Geom 55(1):133–147, 2019), which globalises the local result.

Keywords Lorentzian length spaces · Synthetic curvature bounds · Splitting theorems

✉ Didier A. Solis
didier.solis@correo.uady.mx

Tobias Beran
tobias.beran@univie.ac.at

Argam Ohanyan
argam.ohanyan@univie.ac.at

Felix Rott
felix.rott@univie.ac.at

[1] Department of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria

[2] Facultad de Matemáticas, Universidad Autónoma de Yucatán, Anillo Periférico, Tablaje 13615, Mérida, Mexico

Mathematics Subject Classification 53C50 · 53C23 · 53B30

Contents

1 Introduction

As originally formulated in the context of Riemannian geometry, splitting theorems are a class of results which establish that a Riemannian manifold (M, g) subject to certain hypotheses on its curvature and its geodesic structure actually *splits*, i.e. it is isometric to a product. In most cases, the curvature assumption involves a weak inequality (usually on sectional or Ricci curvature) while the geodesic assumption is related to the existence of a *line*, that is, a globally distance realising complete geodesic. Since lines are conjugate point free and positive curvature promotes the appearance of conjugate points, both features are expected to hold only under very special circumstances.

The very first attempt to establish the metric product structure of a Riemannian manifold with non-negative curvature and having a line is attributed to Cohn–Vossen [27], who proved that a non-compact complete manifold of dimension 2 with sectional curvature $K \geq 0$ is either diffeomorphic to $\mathbb{R}^2$ or flat. However, the proof of this result relied on the Gauss-Bonnet Theorem and therefore was not suitable for generalisation to arbitrary dimensions. The breakthrough that allowed such a generalisation was due to Toponogov, who in the mid 1960s proved his celebrated Triangle Comparison Theorem, which asserts that "the angles of an arbitrary triangle in a Riemannian space made up of minimising arcs are no greater than the corresponding angles of the plane Euclidean triangle with sides of the same length" [62]. Armed with this tool, Toponogov established the archetype of a splitting theorem [61, 62]:

Theorem 1.1 *Let (M, g) be a complete Riemannian manifold with $K \geq 0$ containing a line, then (M, g) is isometric to $(\mathbb{R}^k \times N, g_0 \oplus h)$ where $(\mathbb{R}^k, g_0)$ is the standard Euclidean space and (N, h) is complete and does not have any lines.*

Shortly after Toponogov's splitting theorem was published, Cheeger and Gromoll made substantial progress by generalising it under the less stringent condition $\mathrm{Ric} \geq 0$

[24]. Although their original motivation was rooted in the investigation of topological obstructions on manifolds of non-negative curvature [25], their result sparked the interest in the study of splitting theorems in many different contexts, like in the theory of orbifolds [14] or in the study of curvature inequalities relating other tensors, (like the curvature operator [56] or Bakry–Emery tensor [29, 64]). Furthermore, some more general versions of the Cheeger–Gromoll splitting theorem have been proven. For example, the curvature condition can be replaced by averaged Ricci inequalities [31], or by almost positivity of the Ricci tensor [21, 57]. There are also some local versions in which the splitting occurs either on tubular neighbourhoods or the complement of a compact set disjoint from a line [20, 22].

Remarkably, Yau posed in 1982 the problem of establishing a Lorentzian analogue of the Cheeger–Gromoll splitting theorem [65], thus inaugurating a very active field in which different versions of the Lorentzian splitting theorem were established [10, 11, 28, 30, 32, 55],[1] and some important problems have remained unsolved even to this day, like the famous Bartnik Conjecture [8, 33]. In analogy to its Riemannian counterparts, the first Lorentzian splitting theorem was proven for globally hyperbolic spacetimes with non-positive sectional curvature on timelike planes. In Beem et al. [10, 11] the authors proved what can be thought as the Lorentzian analogue of Toponogov's splitting theorem:

Theorem 1.2 *Let (M, g) be a spacetime of dimension $n \geq 2$ that satisfies the following conditions:*

(i) *(M, g) is globally hyperbolic.*
(ii) *$K(\Pi) \leq 0$ for all timelike planes Π.*
(iii) *M has a timelike line.*

Then (M, g) splits isometrically as $(M, g) \simeq (\mathbb{R} \times N, -dt^2 \oplus h)$, where (N, h) is a complete Riemannian manifold.

Notice that global hyperbolicity can be seen as a more suitable analogue of Riemannian completeness than timelike geodesic completeness, since the latter fails to guarantee connectability by distance realising geodesic segments. The original proof of this result not only uses techniques related to Toponogov's theorem—a Lorentzian triangle comparison due to Harris [43]—, but also tools used in Cheeger-Gromoll's theorem, like Busemann functions and Hessian estimates.

Some of the key techniques used in Toponogov's original proof can be abstracted into an axiomatic setting rather than derived, thus expanding its range of applications to non-smooth contexts, particularly, to the realm of metric length spaces. Recall that a length space (X, d) is a metric space whose distance is intrinsic, in other words, the distance $d(p, q)$ can be recovered as the infimum of the lengths of curves joining p to q. This is the setting for the so-called synthetic methods in geometry, that have proven instrumental in the development of recent mathematical landmarks such as the study of geometry in the large [41, 42], precompactness theorems [17, 18], geometric flows [6, 39] and optimal transport [60, 63]. An *Alexandrov space with curvature bounded from below by k*—or $\mathrm{Alex}(k)$ for short—is a locally compact, complete and path connected

[1] Refer to chapter 14 in [9] for a detailed account.

(metric) length space (X, d) on which the triangle comparison theorem holds [17]. For Alex(0) spaces an analogue of Theorem 1.1 was first proved by Milka [53]. In this result, instead of the existence of a line, the existence of an m-affine function, (i.e. a function $g: M \to \mathbb{R}$ that when restricted to any unit speed geodesics satisfies the differential equation $g'' + mg = 0$) is assumed. A warped product $I \times_g F$ with g an m-affine function is called a *cone* [18]. Cones naturally have m-affine functions, and conversely, an Alex(k) space with a non-constant m-affine function splits as a cone, provided that a boundary condition is met [3, 51]. In the particular case of a complete Riemannian manifold, Innami [46] showed that a 0-affine function exists if and only if M is isometric to a product $\mathbb{R} \times N$. Moreover, under the hypothesis of Theorem 1.1 the Busemann functions associated to a line are affine and thus Innami's theorem implies Toponogov's splitting theorem. Finally, let us note that in [17] there is a proof of Milka's splitting theorem for Alexandrov spaces that resembles more closely the original works of Toponogov. The precise statement is as follows:

Theorem 1.3 *Let (X, d) be an* Alex(0) *space containing a line. Then X splits as a metric product $\mathbb{R} \times Y$, where Y is an* Alex(0) *space.*

In their seminal work [48] Kunzinger and Sämann set the foundations for a synthetic approach to Lorentzian geometry. Their novel notion of Lorentzian (pre-)length space is suited to accommodate several different non-smooth scenarios such as cones [5, 54], spacetimes with C^0 metrics [26, 50], contact structures [45] or causal boundaries [1]. Moreover, in [48] the authors also introduced a notion of timelike curvature bounds in the same spirit of Alex(k) spaces. Among the recent developments in this fast growing field we have detailed analyses of the causal structure [2, 36], extendibility [35, 40], convergence [49], gluing techniques [12, 58] and the basis for a comparison theory [7, 13].

In this work, we aim at proving a splitting theorem for Lorentzian length spaces in the spirit of Theorems 1.1, 1.2 and 1.3. In precise terms, we establish the following splitting result for Lorentzian length spaces:

Theorem 1.4 (Splitting) *Let $(X, d, \ll, \leq, \tau)$ be a connected, regularly localisable, globally hyperbolic Lorentzian length space with proper metric d and global non-negative timelike curvature satisfying timelike geodesic prolongation and containing a complete timelike line $\gamma : \mathbb{R} \to X$. Then there is a τ and causality preserving homeomorphism $f : \mathbb{R} \times S \to X$, where S is a proper, strictly intrinsic metric space of Alexandrov curvature ≥ 0.*

The paper is organised as follows: In Sect. 2, we collect some well-known facts about Lorentzian (pre-)length spaces while focusing on (versions of) results best suited for our needs. We review the basic theory of Lorentzian (pre-)length spaces and discuss concepts like timelike curvature bounds, angles and extensions. Section 3 is dedicated to the study of product Lorentzian pre-length spaces of the form $\mathbb{R} \times X$, where (X, d) is a metric space. Section 4 is the heart of the proof and deals with the construction of co-rays and asymptotes, where we follow [11]. In order to have a control on the behaviour of lines formed as asymptotes to a given line, we introduce the notion of parallelity of timelike lines which is motivated by the splitting theorem for Alexandrov spaces as treated in [17] and show that asymptotic lines are always timelike, of infinite τ-length, and parallel to each other. In Sect. 5 we first establish a local splitting result by endowing a cross section S of the parallel asymptotes spanning $I(\gamma)$ with a natural metric and then showing that $I(\gamma)$ splits as $\mathbb{R} \times S$. Then we show that $I(\gamma)$ has the timelike completeness (TC) property, from which, via an inextendibility argument, $I(\gamma) = X$ follows. Finally, in Sect. 6 we note some classes of spacetimes which naturally satisfy the technical assumptions in our results, and hence split synthetically. We then give an outlook on open problems in the context of synthetic splitting theorems that may be addressed in future projects.

1.1 Notation and conventions

Let us collect some notation and conventions that will be used throughout the paper.

$A \subset B$ means A is a subset of B (not necessarily a proper one). $\mathbb{R}^{1,1}$ is two dimensional Minkowski space with the Lorentzian metric $-dt^2+dx^2$ and coordinates (t, x). $\overline{\tau}$ denotes the time separation on $\mathbb{R}^{1,1}$. A *proper* metric space (X, d) is a metric space such that all closed balls are compact. A metric space is called *length space* or *intrinsic* if the distance between two points is the infimum of lengths of curves running between them, and *strictly intrinsic* if that infimum is a minimum, i.e. between any two points there exists a distance-realising curve.

We denote an open ball with radius R around a point x in a metric space by $B_R(x)$, and the corresponding closed ball by $\bar{B}_R(x)$. Although in general this ball is not the closure $\overline{B_R(x)}$ of the open ball, notice that for length spaces and proper metric spaces we do have the equality $\bar{B}_R(x) = \overline{B_R(x)}$.

2 Basic theory of Lorentzian (pre-)length spaces

2.1 Lorentzian (pre-)length spaces

We give a brief review of the theory of Lorentzian (pre-)length spaces. We focus on the specific results that we need and give proofs if they cannot be found in the literature, otherwise we give precise references. In particular, we refer to [48] for a detailed treatment. We start with a few basic definitions.

Definition 2.1 (*Lorentzian pre-length spaces*) A *Lorentzian pre-length space* $(X, d, \ll, \leq, \tau)$ consists of a metric space (X, d), a reflexive and transitive relation $\leq$ (the *causal relation*), a transitive relation $\ll$ (the *timelike relation*) contained in $\leq$, and a lower semi-continuous map (*time separation*) $\tau : X \times X \to [0, \infty]$ with the following properties: $\tau(x, y) = 0$ if $x \not\leq y$, and $\tau(x, y) > 0$ if and only if $x \ll y$. Moreover, if $x \leq y \leq z$, then the following *reverse triangle inequality* holds:

$$\tau(x, z) \geq \tau(x, y) + \tau(y, z).$$

In the synthetic theory, we employ the usual nomenclature and well-known notation from Lorentzian geometry, such as $I^{\pm}(x)$, $J^{\pm}(x)$ for timelike and causal pasts and futures, as well as $I(x, y) := I^+(x) \cap I^-(y)$ and $J(x, y) := J^+(x) \cap J^-(y)$ for timelike and causal diamonds, respectively.

Lemma 2.2 (Openness of timelike futures and push-up) *Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space and $x \ll y \leq z$ or $x \leq y \ll z$. Then $x \ll z$. In particular, for any $x \in X$, the sets $I^\pm(x)$ are open. Moreover, the relation $\ll$ is open in $X \times X$.*

Proof See [48, Lem. 2.10, Lem. 2.12, Prop. 2.13]. □

Definition 2.3 (*Causal curves*) Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space. A non-constant curve $\gamma : I \to X$, I an interval, is called *future directed causal* (resp. *future directed timelike*) if it is locally Lipschitz and for all $t_1, t_2 \in I$ with $t_1 < t_2$ we have that $\gamma(t_1) \leq \gamma(t_2)$ (resp. $\gamma(t_1) \ll \gamma(t_2)$). It is called *future directed null* if it is future directed causal and no two points on the curve are $\ll$-related. *Past directed causal/timelike/null* curves are defined dually. From now on, unless explicitly stated otherwise, we assume all causal curves to be future directed.

We will sometimes deal with causal curves which are not locally Lipschitz. For these curves, some important results (e.g. the limit curve theorem) will not hold in general, however, many important properties (especially those that are purely causal-theoretic in nature) will continue to be true.

Definition 2.4 (*Continuous causal curves*) Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space. A non-constant curve $\gamma : I \to X$, I an interval, is called a *future directed continuous causal curve* (resp. *future directed continuous timelike curve*) if it is continuous and for all $t_1, t_2 \in I$ with $t_1 < t_2$ we have that $\gamma(t_1) \leq \gamma(t_2)$ (resp. $\gamma(t_1) \ll \gamma(t_2)$). It is called a *future directed continuous null curve* if it is a future directed continuous causal curve and no two points on the curve are $\ll$-related. Past versions of these notions are defined dually. From now on, unless explicitly stated otherwise, we assume all continuous causal curves to be future directed.

We emphasise that a *causal/timelike/null curve* is always understood to be locally Lipschitz. If *continuous* causal/timelike/null curves are meant, we will state that explicitly. Wherever possible, we will state definitions and results for continuous causal curves instead of (locally Lipschitz) causal curves for greater generality. In most cases, the proofs given in the literature for these results apply word for word to the continuous case.

Definition 2.5 (*Extensions of causal curves*) Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space and let $\gamma : I \to X$ be a continuous causal curve, where I is any interval. We say that γ is *(continuously) extendible* if there exists a (continuous) causal curve $\tilde{\gamma} : J \to X$, J an interval with $J \supsetneq I$, such that $\tilde{\gamma}|_I = \gamma$. Otherwise, we say γ is *(continuously) inextendible*. We refer to (in)extendibility to points in the future resp. past directions of γ as *future/past (in)extendibility*.[2]

[2] Geodesics defined on intervals $[a, b]$ with $a = -\infty$ or $b = \infty$ should be thought of as properly reparametrised.

Definition 2.6 (τ*-length*) Let $\gamma : [a, b] \to X$ be a future directed continuous causal curve. Then its τ*-length* is defined by

$$L_\tau(\gamma) := \inf \left\{ \sum_{i=0}^{N-1} \tau(\gamma(t_i), \gamma(t_{i+1})) \mid a = t_0 < t_1 < \cdots < t_N = b \right\}.$$

If γ is past directed causal, then $L_\tau(\gamma)$ is defined analogously. For continuous causal curves defined on half-open intervals, e.g. $[a, b)$, one takes the limit of $L_\tau(\gamma|_{[a,c]})$ as $c \to b$, and similarly in the other cases.

The τ-length of a continuous causal curve is invariant under reparametrisations. We will have more to say on these topics later.

Definition 2.7 (*Maximising causal curves*) A future directed (continuous) causal curve $\gamma : [a, b] \to X$ is called *maximising* (or τ*-maximising*) if $\tau(\gamma(a), \gamma(b)) = L_\tau(\gamma)$, and analogously for past directed (continuous) causal curves. We also refer to such curves as *(continuous) distance realisers*.

Lorentzian pre-length spaces where between each pair of causally related points there is a maximising (locally Lipschitz) causal curve are referred to as *strictly intrinsic* or *geodesic*.

Next, we define some of the steps on the causal ladder for Lorentzian pre-length spaces.

Definition 2.8 (*Causality conditions*) A Lorentzian pre-length space $(X, d, \ll, \leq, \tau)$ is called

(i) *non-totally imprisoning* if for every compact set $K \subset X$ there is $C > 0$ such that the d-lengths of all causal curves in K is bounded by C,
(ii) *strongly causal* if the Alexandrov topology generated by the subbasis of timelike diamonds $\{I(x, y) \mid x, y \in X\}$ coincides with the metric topology,
(iii) *globally hyperbolic* if it is non-totally imprisoning and all *causal diamonds* $J(x, y)$ are compact.

Lemma 2.9 *Let $(X, d, \ll, \leq, \tau)$ be a strongly causal Lorentzian pre-length space. Then each $x \in X$ has a neighbourhood basis consisting of causally convex*[3] *neighbourhoods. Even more is true: X has a topological basis of causally convex neighbourhoods.*

Proof This is trivial since by definition finite intersections of timelike diamonds are a basis for the topology, but these sets are causally convex. □

Definition 2.10 (*Causal path-connectedness*) A Lorentzian pre-length space X is called *causally path-connected* if for all $x < y$ there is a causal curve from x to y and for all $x \ll y$ there is a timelike curve from x to y.

[3] A set $V \subset X$ is called *causally convex* if for all $p, q \in V$ we have $J(p, q) \subset V$.

For the next definition, we use the version given in [2, Def. 2.16] rather than [48, Def. 3.4] as it more closely resembles the case of smooth spacetimes. Before doing so, note that on causally path-connected Lorentzian pre-length spaces X, we can introduce a *local causal relation*: Let $U \subset X$ be open and $x, y \in U$, then we say $x \leq_U y$ if there exists a future causal curve from x to y entirely contained in U.

Definition 2.11 (*Local causal closedness*) Let X be causally path-connected and $U \subset X$ be open. We say that U is *causally closed* if for all sequences p_n, q_n in U converging to $p, q \in U$ and $p_n \leq_U q_n$, we have that $p \leq_U q$. We say that X is *locally causally closed* if every point in X has a causally closed neighbourhood.

Definition 2.12 (*Global causal closedness*) A Lorentzian pre-length space X is called *globally causally closed* if $\leq$ is a closed relation.

Recall that in the theory of smooth spacetimes, a globally causally closed and causal spacetime is *causally simple*. Such spacetimes are one step below global hyperbolicity in the classical causal ladder.

Definition 2.13 (*d-compatibility*) A Lorentzian pre-length space $(X, d, \ll, \leq, \tau)$ is called *d-compatible* if for every $x \in X$ there is a neighbourhood U of x and a constant $C > 0$ such that the d-lengths of all causal curves contained in U are bounded above by C.

Lemma 2.14 *Let $(X, d, \ll, \leq, \tau)$ be a locally causally closed, d-compatible, globally hyperbolic Lorentzian pre-length space. Then no compact set in X contains an inextendible causal curve.*

Proof See [48, Cor. 3.15]. □

Definition 2.15 (*Localisability*) A Lorentzian pre-length space $(X, d, \ll, \leq, \tau)$ is called *localisable* if for each $x \in X$ there exists a neighbourhood (called *localising neighbourhood*) Ω_x containing x with the following properties:

(i) Causal curves in Ω_x have uniformly bounded d-length, i.e., Ω_x is d-compatible.
(ii) For each $y \in \Omega_x$, $I^{\pm}(y) \cap \Omega_x \neq \emptyset$.
(iii) There is a continuous map (*local time separation*) $\omega_x : \Omega_x \times \Omega_x \to [0, \infty)$ such that Ω_x is a Lorentzian pre-length space upon restricting $d, \ll, \leq$.
(iv) For all $p, q \in \Omega_x$ with $p < q$ there exists a causal curve γ_{pq} from p to q entirely in Ω_x with maximal τ-length among all causal curves from p to q contained in Ω_x, as well as $L_\tau(\gamma_{pq}) = \omega_x(p, q)$.

Note that this necessarily means that ω_x is of the form

$$\omega_x(p, q) := \max\{L_\tau(\gamma) \mid \gamma \text{ is a causal curve from } p \text{ to } q \text{ in } \Omega_x.\}. \tag{2.1}$$

If in addition, for $p, q \in \Omega_x$ with $p \ll q$ the curve γ_{pq} is timelike and strictly longer than any causal curve from p to q in Ω_x containing a null segment, then Ω_x is called a *regular localising neighbourhood*, and X *regularly localisable* if any point has a regular localising neighbourhood. Moreover, X is called *strongly localisable* if each point has a neighbourhood base of localising neighbourhoods and *SR-localisable* if each point has a neighbourhood base of regular localising neighbourhoods.

Lemma 2.16 *Let $(X, d, \ll, \leq, \tau)$ be a strongly causal Lorentzian pre-length space.*

(i) *If X is localisable, then it is strongly localisable.*
(ii) *If X is regularly localisable, then it is SR-localisable.*

Proof By Lemma 2.9, each point in X has a neighbourhood basis of causally convex neighbourhoods. Now suppose X is localisable, take $x \in X$ and let Ω_x be a localising neighbourhood of x and let $U \subset \Omega_x$ be any neighbourhood of x. Then there is a causally convex neighbourhood V of x such that $V \subset U$. We are done if we show that V is a localising neighbourhood. But this is easy to see since any causal curve starting and ending in V is entirely contained in there. If X is regularly localisable, then V is even a regular localising neighbourhood. □

Regularly localisable spaces have the following important property:

Theorem 2.17 (Causal character of maximising causal curves) *In a regularly localisable Lorentzian pre-length space, maximising causal curves are either timelike or null.*

Proof See [48, Thm. 3.18]. □

Note that in localisable Lorentzian pre-length spaces, the sets $I^{\pm}(y)$ are never empty for any $y \in X$.

Definition 2.18 (*Geodesic*) Let X be a localisable Lorentzian pre-length space and $\gamma : I \to X$ a (continuous) causal curve. Then γ is a (future-directed causal) *(continuous) geodesic* if for each $t_0 \in I$ there exists a localising neighbourhood Ω of $\gamma(t_0)$ and a neighbourhood $[c, d] \subset I$ of t_0 such that $\gamma([c, d]) \subset \Omega$ and $\gamma|_{[c,d]}$ is maximising in Ω with respect to the local time separation.

We will sometimes need the notion of *geodesic (in)extendibility* which is defined in terms of the existence of a geodesic extending a given geodesic. Note that inextendibility implies geodesic inextendibility but the converse need not hold in general.

For many arguments in the proof of the splitting theorem, the property that geodesics can always be extended to open domains will be essential. While this is a consequence of the ODE theory of the geodesic equation in the case of smooth spacetimes, we will need to assume it here:

Definition 2.19 (*Timelike geodesic prolongation*) A localisable Lorentzian pre-length space X is said to have the *timelike geodesic prolongation* property if any maximising timelike segment $\gamma : [a, b] \to X$ can be extended to a timelike geodesic with an open domain, i.e. there is an $\varepsilon > 0$ and a timelike geodesic $\bar{\gamma} : (a - \varepsilon, b + \varepsilon) \to X$ with $\bar{\gamma}|_{[a,b]} = \gamma$.

It is clear that this is satisfied in smooth strongly causal spacetimes considered as a Lorentzian pre-length space. One easily sees that spacetimes with timelike boundary do not satisfy this.

Definition 2.20 (*Locally quasiuniformly maximising*) A localisable Lorentzian pre-length space X is called *locally quasiuniformly maximising* if for each $x \in X$ there exists a localisable neighbourhood U containing x such that each causal geodesic entirely contained in U is maximising in U with respect to the local time separation.

It is unclear to us how the property of being locally quasiuniformly maximizing is related to other notions for Lorentzian pre-length spaces, such as being intrinsic, strongly causal or (locally) causally closed.

Proposition 2.21 (Upper semicontinuity of Lorentzian arclength) *Let $(X, d, \ll, \leq, \tau)$ be a strongly causal and localisable Lorentzian pre-length space. Then L_τ is upper semi-continuous, i.e. if $\gamma_n : [a, b] \to X$ are causal curves converging d-uniformly to a causal curve $\gamma : [a, b] \to X$, then*

$$L_\tau(\gamma) \geq \limsup_{n\to\infty} L_\tau(\gamma_n).$$

Proof See [48, Prop. 3.17]. □

Proposition 2.22 *Let X be a strongly causal, localisable, locally quasiuniformly maximising Lorentzian pre-length space and let $\gamma : [a, b) \to X$ be a causal geodesic. Then γ is geodesically extendible to $[a, b]$ if and only if it is extendible as a continuous curve to $[a, b]$.*

Proof See [40, Prop. 4.6] and note that the proof really requires the space to be locally quasiuniformly maximising. □

We now state the limit curve theorem in the case of strongly causal Lorentzian pre-length spaces, which will be sufficient for our needs. Note that in this case, by [2, Prop. 2.21], Definition 2.11 is equivalent to [48, Def. 3.4].

Theorem 2.23 (Limit curve theorem) *Let $(X, d, \ll, \leq, \tau)$ be a locally causally closed, localisable, strongly causal, d-compatible Lorentzian pre-length space with proper metric d. Let $\gamma_n : [0, L_n] \to X$ be a sequence of causal curves parametrised by d-arclength, and suppose that $L_n = L^d(\gamma_n) \to \infty$. If the sequence $\gamma_n(0)$ has an accumulation point $x \in X$, then some subsequence of the γ_n converges locally uniformly to a future inextendible causal curve $\gamma : [0, \infty) \to X$. Moreover, if the γ_n are maximising, then so is γ. In this case, for any $T > 0$, $L_\tau(\gamma|_{[0,T]}) = \lim_n L_\tau(\gamma_n|_{[0,T]})$.*

Proof The statement about the existence of the limit curve is shown in [48, Thm. 3.14]. Now suppose that γ_n are maximising and w.l.o.g. let γ_n itself converge locally uniformly to γ. Consider any $T > 0$. Then by upper semicontinuity of L_τ and lower semicontinuity of τ,

$$\begin{aligned}\limsup_n L_\tau(\gamma_n|_{[0,T]}) &\leq L_\tau(\gamma|_{[0,T]}) \leq \tau(\gamma(0), \gamma(T)) \\ &\leq \liminf_n \tau(\gamma_n(0), \gamma_n(T)) = \liminf_n L_\tau(\gamma_n|_{[0,T]}).\end{aligned}$$

Thus, equality holds everywhere. In particular, γ is maximising and $\lim_n L_\tau(\gamma_n|_{[0,T]}) = L_\tau(\gamma|_{[0,T]})$. □

In the context of the splitting theorem, we are interested in Lorentzian pre-length spaces with significantly more structure, which we will discuss now.

Definition 2.24 (*Lorentzian length spaces*) A locally causally closed, causally path-connected and localisable Lorentzian pre-length space $(X, d, \ll, \leq, \tau)$ is called *Lorentzian length space* if for all $x \leq y$, $x \neq y$,

$$\tau(x, y) = \sup\{L_\tau(\gamma) \mid \gamma \text{ future directed causal from } x \text{ to } y\}.$$

Note that since by definition $L_\tau(\gamma) \leq \tau(p, q)$ for any (even continuous) causal curve γ from p to q, one can equivalently take the supremum over continuous causal curves to obtain the time separation in the case of Lorentzian length spaces.

Proposition 2.25 (Causality of Lorentzian length spaces)

(i) *A Lorentzian length space is strongly causal if and only if each point has a neighbourhood base of causally convex neighbourhoods.*
(ii) *Strongly causal Lorentzian length spaces are non-totally imprisoning.*
(iii) *Globally hyperbolic Lorentzian length spaces are strongly causal. Moreover, τ is continuous and finite and for any two causally related points there exists a maximising causal curve connecting them.*

Proof See [48, Thm. 3.26, Thm. 3.28, Thm. 3.30]. □

Lemma 2.26 *Let $(X, d, \ll, \leq, \tau)$ be a strongly causal Lorentzian pre-length space such that for each $x \in X$ there exists a timelike curve $\gamma : [-\varepsilon, \varepsilon] \to X$ with $\gamma(0) = x$. Then $\{I(x, y) \mid x, y \in X\}$ is a base for the topology on X. Moreover, for each $x \in X$, each timelike curve $\gamma : [-\varepsilon, \varepsilon] \to X$ with $\gamma(0) = x$ and parameters $0 < s_n, t_n \to 0$, $\{I(\gamma(-s_n), \gamma(t_n)) \mid n \in \mathbb{N}\}$ is a neighbourhood base at x.*

Proof Let $x \in X$ with $x \in U$, $U \subset X$ open. By the definition of the Alexandrov topology, finite intersections of the sets $I(y, z)$ with $y, z \in X$ are a base. Thus, there exist $x_1, \ldots, x_n, y_1, \ldots, y_n \in X$ such that

$$x \in V := I(x_1, y_1) \cap \cdots \cap I(x_n, y_n) \subset U.$$

Now let $\gamma : [-\varepsilon, \varepsilon] \to X$ be a timelike curve with $\gamma(0) = x$. By continuity, maybe after making ε smaller, we may assume that $\gamma([-\varepsilon, \varepsilon]) \subset V$. Then $W := I(\gamma(-\varepsilon), \gamma(\varepsilon))$ is an open set containing x and we claim that $W \subset V$. To see this, let $z \in W$, then for any $i = 1, \ldots, n$ we have $x_i \ll \gamma(-\varepsilon) \ll z$, hence $z \in I^+(x_i)$. Similarly, $z \ll \gamma(\varepsilon) \ll y_i$, hence $z \in I^-(y_i)$. Thus $z \in V$. This proves that $\{I(x, y) \mid x, y \in X\}$ is a base for the topology of X, and the second claim then easily follows from the arguments above. □

Lorentzian pre-length spaces where each point lies in the interior of a timelike curve are future and past approximating in the sense of [19, Def. 2.17] (see also [19, Lem. 2.18])

Lemma 2.27 (Geodesics in strongly causal Lorentzian length spaces) *Let $(X, d, \ll, \leq, \tau)$ be a strongly causal Lorentzian length space. A (continuous) causal curve $\gamma : I \to X$ is a (continuous) geodesic if and only if for each $t_0 \in I$ there exists a subinterval $[c, d] \subset I$ with $t_0 \in [c, d]$ such that $\gamma|_{[c,d]}$ is τ-maximising.*

Proof This follows from [40, Lem. 4.3]. □

To conclude this subsection, we discuss d-arclength and τ-arclength reparametrisations of causal curves.

Remark 2.28 (On d-arclength parametrisations) Let $(X, d, \ll, \leq, \tau)$ be a globally hyperbolic Lorentzian length space with proper metric d and $\gamma : [a, b) \to X$ a causal curve that is inextendible at b. Then γ has infinite d-length: By Lemma 2.14, γ has to leave every compact set, in particular it leaves each compact ball $\overline{B}_n(\gamma(a))$, which shows that $L^d(\gamma) = \infty$. If we reparametrise γ by d-arclength (note that this is in general not possible for continuous causal curves), it is thus defined on $[0, \infty)$. Similarly, if we have a doubly inextendible causal curve $\tilde{\gamma} : (a, b) \to X$, then upon reparametrising it by d-arclength, it is defined on $\mathbb{R}$, so that $L^d(\tilde{\gamma}|_{[c,d]}) = |d - c|$ (see also [19, Rem. 2.5]). Note that d-arclength parametrisations are globally 1-Lipschitz.

Lemma 2.29 (τ-arclength parametrisations for inextendible curves) *Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space with continuous time separation τ which satisfies $\tau(x, x) = 0$ for all $x \in X$. Let $\gamma : [a, b) \to X$ be a future inextendible continuous timelike curve defined on a half-open interval with $L_\tau(\gamma|_{[a,c]}) < \infty$ for each $c \in [a, b)$ (notice we might have $L_\tau(\gamma) = \infty$). Then the map $\phi : [a, b) \to [0, L_\tau(\gamma))$, $t \mapsto L_\tau(\gamma|_{[a,t]})$ is continuous and strictly increasing. Moreover, $\tilde{\gamma} := \gamma \circ \phi^{-1} : [0, \infty) \to X$ is a weak reparametrisation of γ with respect to τ-arclength. Similarly, a doubly inextendible continuous timelike curve $\gamma : (a, b) \to X$ can be weakly parametrised by τ-arclength and is then defined on $\big(-L_\tau(\gamma|_{(a,t_0]}), L_\tau(\gamma|_{[t_0,b)})\big)$, where $t_0 \in (a, b)$ is arbitrary.*

Proof This can be seen by applying [48, Lem. 3.33, Lem. 3.34] to each $\gamma|_{[a,\tilde{b}]}$ with $\tilde{b} < \infty$. □

Note that, in general, the τ-arclength parametrisation results in a continuous timelike curve, even if the original curve is locally Lipschitz.

In the following, whenever τ-arclength parametrised curves are mentioned, it is always implicitly understood that those curves have finite τ-length on compact subintervals. In localisable spaces this is anyway the case.

2.2 Timelike curvature bounds

Establishing an analogue to sectional curvature bounds in semi-Riemannian manifolds á la [4] was one of the main goals in the first work on Lorentzian length spaces [48]. Historically, splitting theorems always used some type of curvature condition, either Ricci or sectional curvature. By M_k we denote the two-dimensional Lorentzian model space of constant sectional curvature k and, in this subsection only, $\bar{\tau}$ denotes the time separation in M_k (in general $\bar{\tau}$ will only be used as the time separation in $M_0 = \mathbb{R}^{1,1}$). For more details on this, see [48, Definition 4.5]. We will always assume that all mentioned triangles satisfy size bounds for M_k, cf. [48, Lemma 4.6]. However, since we are mainly working with $k = 0$, we do not have to worry about this.

We call three points $x_1, x_2, x_3 \in X$ that are timelike related together with maximisers between them a *timelike triangle*. We denote such triangles by $\Delta(x_1, x_2, x_3)$.

Definition 2.30 Let X be a Lorentzian pre-length space and $\Delta(x_1, x_2, x_3)$ be a timelike triangle. Then a timelike triangle $\Delta(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ in M_k such that $\tau(x_i, x_j) = \bar{\tau}(\bar{x}_i, \bar{x}_j)$ is called a *comparison triangle* for $\Delta(x_1, x_2, x_3)$. It always exists (if the size bounds are satisfied) and is unique up to isometries of M_k.

Definition 2.31 (*Timelike curvature bounds*) A Lorentzian pre-length space X is said to satisfy a *timelike curvature bound from below by* k, if each point in X has a so-called comparison neighbourhood U, satisfying the following:

(i) $\tau|_{U\times U}$ is finite and continuous.
(ii) For all $x, y \in U$ with $x \ll y$ there exists a causal maximiser γ from x to y which is entirely contained in U.
(iii) Let $\Delta(x, y, z)$ be a triangle in U and let $\Delta(\bar{x}, \bar{y}, \bar{z})$ be its comparison triangle in M_k. Let $p, q \in \Delta(x, y, z)$ and let $\bar{p}, \bar{q}$ be the corresponding comparison points in $\Delta(\bar{x}, \bar{y}, \bar{z})$. Then $\tau(p, q) \leq \bar{\tau}(\bar{p}, \bar{q})$.

For curvature bounds from above, the inequality in (iii) is reversed, but we will not need these types of curvature bounds. If $k = 0$, we also say that X has *non-negative timelike curvature* resp. *non-positive timelike curvature*.

Evidently, curvature bounds are only formulated locally as of yet. There is, however, a very natural way of globalising this concept.

Definition 2.32 (*Global curvature bound*) A Lorentzian pre-length space X is said to have *global curvature bounded below by* k if X itself is a comparison neighbourhood in the sense of Definition 2.31.

Remark 2.33 (Continuous triangles vs. Lipschitz triangles) If X is a globally hyperbolic Lorentzian length space with global nonnegative timelike curvature, then in fact curvature comparison even holds for timelike triangles where the maximisers are continuous: Indeed, suppose $\Delta := \Delta_{C^0}(x, y, z)$ is a continuous timelike triangle, and let $p, q \in \Delta$. Due to the properties of X, we find Lipschitz maximisers from the endpoints of Δ to p, q, respectively. The concatenations at p and q, respectively, of those maximisers are again maximisers because the sides on Δ are (continuous) maximisers. Hence we have realised p, q on a Lipschitz triangle. If X is in addition regularly localisable, then in fact every continuous timelike maximiser is Lipschitz reparametrisable [52], so the above discussion is superfluous in this case.

One of the most commonly used implications of lower (timelike) curvature bounds is the prohibition of branching of distance-realisers. A formulation of this result for Lorentzian pre-length spaces was first given in [48, Theorem 4.12]. However, with the introduction of hyperbolic angles in [13] it was possible to generalise this result by omitting some of the additional assumptions:

Theorem 2.34 (Timelike non-branching) *Let X be a strongly causal Lorentzian pre-length space with timelike curvature bounded below by $k \in \mathbb{R}$. Then timelike distance realisers cannot branch, i.e., if $\alpha, \beta : [-\varepsilon, \varepsilon] \to X$ are timelike distance realisers such that there exists $t_0 \in (-\varepsilon, \varepsilon)$ with $\alpha|_{[-\varepsilon, t_0]} = \beta|_{[-\varepsilon, t_0]}$, then $\alpha = \beta$.*

Proof See [13, Theorem 4.7]. □

The non-branching of timelike distance realisers is a key property of spaces with lower curvature bounds and will appear in various forms in our proof of the splitting theorem.

2.3 Angles and comparison angles

Hyperbolic angles in Lorentzian pre-length spaces were introduced in [7, 13], where the former puts a bigger focus on comparison results. We will follow the conventions of the latter reference.

Lemma 2.35 (The law of cosines ($k = 0$)) *Let $X = \mathbb{R}^{1,1}$ be $2D$-Minkowski space and x_1, x_2, x_3 be three points which are timelike related (an (unordered) timelike triangle). Let $a_{ij} = \max(\tau(x_i, x_j), \tau(x_j, x_i))$ (note one of these is zero anyway). Let ω be the hyperbolic angle between the straight lines x_1x_2 and x_2x_3 at x_2. Set $\sigma = 1$ if x_2 is not a time endpoint of the triangle (i.e., $x_1 \ll x_2 \ll x_3$ or $x_3 \ll x_2 \ll x_1$) and $\sigma = -1$ if x_2 is a time endpoint of the triangle (i.e., $x_2 \ll x_1, x_3$ or $x_1, x_3 \ll x_2$). Then we have:*

$$a_{13}^2 = a_{12}^2 + a_{23}^2 + \sigma a_{12} a_{23} \cosh(\omega).$$

In particular, when only changing one side-length, the angle ω is a monotonously increasing function of the longest side-length and monotonously decreasing in the other side-lengths.

Proof See [13, Appendix A]. □

Note that for $a_{ij} > 0$ satisfying a reverse triangle inequality and choosing an appropriate $\sigma = \pm 1$, we can always solve this equation for ω.

Definition 2.36 (*Comparison angles*) Let X be a Lorentzian pre-length space and x_1, x_2, x_3 three timelike related points. Let $\bar{x}_1, \bar{x}_2, \bar{x}_3 \in \mathbb{R}^{1,1}$ be a comparison triangle for x_1, x_2, x_3. We define the *comparison angle* $\tilde{\measuredangle}_{x_2}(x_1, x_3)$ as the hyperbolic angle between the straight lines $\bar{x}_1\bar{x}_2$ and $\bar{x}_2\bar{x}_3$ at $\bar{x}_2$. It can be calculated with the law of cosines using $a_{ij} = \max(\tau(x_i, x_j), \tau(x_j, x_i))$ and σ, where we set $\sigma = 1$ if x_2 is not a time endpoint of the triangle and $\sigma = -1$ if x_2 is a time endpoint of the triangle. σ is called the *sign* of the comparison angle (even though we always have $\tilde{\measuredangle}_{x_2}(x_1, x_3) > 0$). For a reduction of the number of case distinctions we also define the *signed comparison angle* $\tilde{\measuredangle}^{\mathrm{S}}_{x_2}(x_1, x_3) = \sigma \tilde{\measuredangle}_{x_2}(x_1, x_3)$.

Definition 2.37 (*Angles*) Let X be a Lorentzian pre-length space and $\alpha, \beta : [0, \varepsilon) \to X$ be two timelike curves (future or past directed or one of each) with $x := \alpha(0) =$

$\beta(0)$. Then we define the *upper angle*

$$\measuredangle_x(\alpha, \beta) = \limsup_{(s,t)\in D, s,t\to 0} \tilde{\measuredangle}_x(\alpha(s), \beta(t)),$$

where $D = \{(s, t) : s, t > 0,\ \alpha(s), \beta(t)$ timelike related$\}$. If the limes superior is a limit and finite, we say *the angle exists* and call it an *angle*.

Note that the sign of the comparison angle is independent of $(s, t) \in D$. We define the *sign* of the (upper) angle σ to be that sign, and define the *signed (upper) angle* to be $\measuredangle_x^{\mathrm{S}}(\alpha, \beta) = \sigma \measuredangle_x(\alpha, \beta)$.

If maximisers between any two timelike related points are unique (as e.g. in $\mathbb{R}^{1,1}$), then we simply write $\measuredangle_p(x, y)$ for the angle at p between the maximisers from p to x and p to y.

We give an alternative definition of timelike curvature bounds in the case $k = 0$ following Definition 4.9 in [13]:

Definition 2.38 Let X be a Lorentzian pre-length space. We say that X has *timelike curvature bounded below by* $k = 0$ *(bounded above by* $k = 0$*) in the sense of monotonicity comparison* if every point in X has a neighbourhood U such that

(i) $\tau|_{U\times U}$ is finite and continuous.
(ii) U is *timelike geodesically connected*, i.e. whenever $x, y \in U$ with $x \ll y$, there exists a future-directed timelike distance realiser α in U from x to y, and any distance realiser from x to y contained in U is timelike.
(iii) Whenever $\alpha : [0, a] \to U$, $\beta : [0, b] \to U$ are timelike distance realisers in U with $x := \alpha(0) = \beta(0)$, we define the function $\theta : (0, a] \times (0, b] \supset D \to \mathbb{R}$ by $\theta(s, t) := \tilde{\measuredangle}_x^{\mathrm{S}}(\alpha(s), \beta(t))$, where $(s, t) \in D$ precisely when $\alpha(s), \beta(t)$ are timelike related. We require this to be monotonically increasing (decreasing) in s and t. Additionally, in the case of non-positive timelike curvature, we consider $0 < s' \leq s$ and $0 < t' \leq t$ and require that if $\theta(s', t')$ is not defined (i.e. $\alpha(s'), \beta(t')$ are not timelike related) but $\theta(s, t)$ is, then also the comparison points for $\alpha(s'), \beta(t')$ in a comparison triangle for $\Delta(x, \alpha(s), \beta(t))$ are not timelike related.

We call the curvature bound *global* if X is a comparison neighbourhood in the above sense.

Theorem 2.39 (Timelike curvature: Equivalence of definitions) *Let X be a locally timelike geodesically connected Lorentzian pre-length space. Then X has non-negative (non-positive) timelike curvature in the sense of Definition 2.31 if and only if it has non-negative (non-positive) timelike curvature in the sense of monotonicity comparison (see Definition 2.38).*

Proof See [13, Theorem 4.12]. □

Remark 2.40 Note that if X is *globally timelike geodesically connected*, i.e. between any two timelike related points there is a timelike distance realiser connecting them, and any distance realiser connecting them is timelike (which is certainly true if X is a regularly localisable and globally hyperbolic Lorentzian length space), then X has global non-negative (non-positive) timelike curvature in the sense of Definition 2.32 if and only if it has global non-negative (non-positive) timelike curvature in the sense of monotonicity comparison.

The timelike geodesic prolongation property plays a technical role in our proof of the splitting theorems mainly because of the following result.

Proposition 2.41 (Continuity of angles in spaces with timelike curvature bounded below) *Let X be a strongly causal, localisable, timelike geodesically connected and locally causally closed Lorentzian pre-length space with global non-negative timelike curvature and which satisfies timelike geodesic prolongation. Let $\alpha_n, \alpha, \beta_n, \beta$ be future or past directed timelike geodesics all starting at $\alpha_n(0) = \alpha(0) = \beta_n(0) = \beta(0) =: x$ and with $\alpha_n \to \alpha$ and $\beta_n \to \beta$ pointwise (in particular, α_n is future directed if and only if α is, and similarly for β_n and β). Then*

$$\measuredangle_x(\alpha, \beta) = \lim_n \measuredangle_x(\alpha_n, \beta_n)$$

Proof See [13, Proposition 4.14]. □

The following two propositions will be proven together.

Proposition 2.42 (Alexandrov lemma: across version) *Let X be a Lorentzian pre-length space. Let $\Delta := \Delta(x, y, z)$ be a timelike triangle (in particular the distance realisers between the endpoints exist). Let p be a point on the side xz with $p \ll y$, such that the distance realiser between p and y exists. Then we can consider the smaller triangles $\Delta_1 := \Delta(x, p, y)$ and $\Delta_2 := \Delta(p, y, z)$. We construct a comparison situation consisting of a comparison triangle $\bar{\Delta}_1$ for Δ_1 and $\bar{\Delta}_2$ for Δ_2, with $\bar{x}$ and $\bar{z}$ on different sides of the line through $\bar{p}\bar{y}$ and a comparison triangle $\tilde{\Delta}$ for Δ with a comparison point $\tilde{p}$ for p on the side xz. This contains the subtriangles $\tilde{\Delta}_1 := \Delta(\tilde{x}, \tilde{y}, \tilde{p})$ and $\tilde{\Delta}_2 := \Delta(\tilde{p}, \tilde{y}, \tilde{z})$, see Fig. 1.*

Then the situation $\bar{\Delta}_1$, $\bar{\Delta}_2$ is convex (concave) at p (i.e. $\tilde{\measuredangle}_p(x, y) = \measuredangle_{\bar{p}}(\bar{x}, \bar{y}) \geq \measuredangle_{\bar{p}}(\bar{y}, \bar{z}) = \tilde{\measuredangle}_p(y, z)$ (or $\leq$)) if and only if $\tau(p, y) \leq \bar{\tau}(\bar{p}, \bar{y})$ (or $\geq$). If this is the case, we have that

- *each angle in the triangle $\bar{\Delta}_1$ is $\geq$ (or $\leq$) than the corresponding angle in the triangle $\tilde{\Delta}_1$,*
- *each angle in the triangle $\bar{\Delta}_2$ is $\geq$ (or $\leq$) than the corresponding angle in the triangle $\tilde{\Delta}_2$.*

In any case, we have that

- $\measuredangle_{\bar{y}}(\bar{x}, \bar{z}) \geq \measuredangle_{\tilde{x}}(\tilde{x}, \tilde{z}) = \tilde{\measuredangle}_y(x, z)$.

The same is true if p is a point on the side xz such that $y \ll p$. Note that if X has non-negative (non-positive) timelike curvature and Δ is within a comparison neighbourhood, the condition is satisfied, i.e. $\tau(p, y) \leq \bar{\tau}(\tilde{p}, \tilde{y})$ (or $\geq$).

Fig. 1 A concave situation in the across version

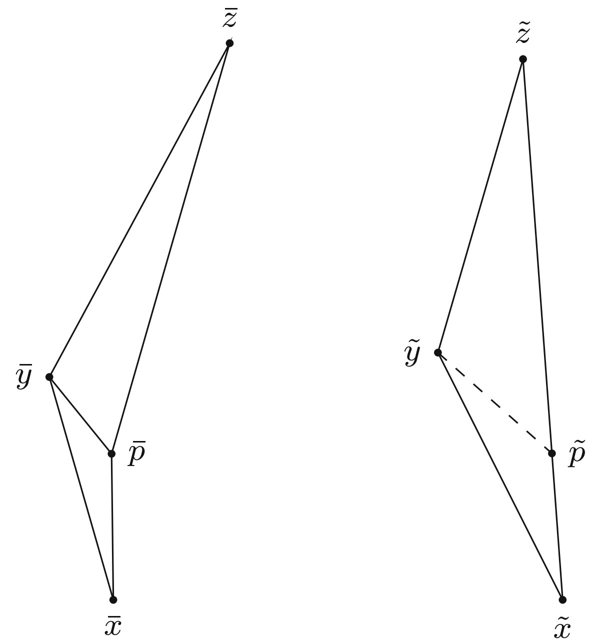
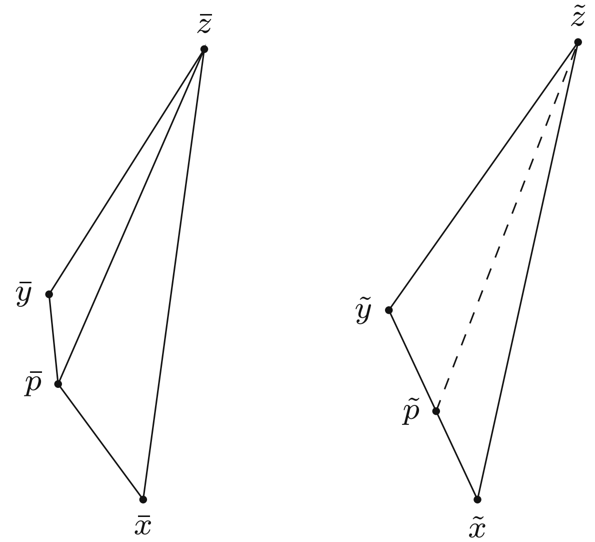


Fig. 2 A convex situation in the future version



Proposition 2.43 (Alexandrov lemma: future version) *Let X be a Lorentzian pre-length space. Let $\Delta := \Delta(x, y, z)$ be a timelike triangle (in particular the distance realisers between the endpoints exist). Let p be a point on the side xy, such that the distance realiser between p and z exists. Then we can consider the smaller triangles $\Delta_1 := \Delta(x, p, z)$ and $\Delta_2 := \Delta(p, y, z)$. We construct a comparison situation consisting of a comparison triangle $\bar{\Delta}_1$ for Δ_1 and $\bar{\Delta}_2$ for Δ_2, with $\bar{x}$ and $\bar{y}$ on different sides of the line through $\bar{p}\bar{z}$ and a comparison triangle $\tilde{\Delta}$ for Δ with a comparison point $\tilde{p}$ for p on the side xy. This contains the subtriangles $\tilde{\Delta}_1 := \Delta(\tilde{x}, \tilde{p}, \tilde{z})$ and $\tilde{\Delta}_2 := \Delta(\tilde{p}, \tilde{y}, \tilde{z})$, see Fig. 2.*

Then the situation $\bar{\Delta}_1, \bar{\Delta}_2$ is convex (concave) at p (i.e. $\measuredangle_{\bar{p}}(\bar{y}, \bar{z}) \geq \measuredangle_{\bar{p}}(\bar{x}, \bar{z})$ (or $\leq$)) if and only if $\tau(p, z) \leq \bar{\tau}(\bar{p}, \bar{z})$ (or $\geq$). If this is the case, we have that

- *each angle in the triangle $\bar{\Delta}_1$ is $\geq$ (or $\leq$) than the corresponding angle in the triangle $\tilde{\Delta}_1$,*

- *each angle in the triangle $\bar{\Delta}_2$ is $\leq$ (or $\geq$) than the corresponding angle in the triangle, $\tilde{\Delta}_2$.*

In any case, we have that

- $\measuredangle_{\bar{z}}(\bar{x}, \bar{y}) \leq \measuredangle_{\tilde{z}}(\tilde{x}, \tilde{y}) = \tilde{\measuredangle}_z(x, y)$.

Note that if X has non-negative (non-positive) timelike curvature and Δ is within a comparison neighbourhood, the condition is satisfied, i.e. $\tau(p, z) \leq \bar{\tau}(\tilde{p}, \tilde{z})$ (or $\geq$).

Proof (Proof for both Alexandrov lemmas) It is sufficient to show the "if" part of the statement. Indeed, under the assumption that the triangles Δ_1 and Δ_2 are non-degenerate, the "if" part holds with strict inequalities, which then also implies the "only if" part of the statement. The degenerate cases are trivial.

For the triangles $\bar{\Delta}_1$ vs. $\tilde{\Delta}_1$, only one side-length changes, and it is not the longest side of the triangle in both versions. Thus, the monotonicity statement in the law of cosines, cf. Lemma 2.35, implies that in the across version, the relation between the desired angle $\measuredangle_{\bar{p}}(\bar{x}, \bar{y})$ in $\bar{\Delta}_1$ to the corresponding angle $\measuredangle_{\tilde{p}}(\tilde{x}, \tilde{y})$ in $\tilde{\Delta}_1$ is pointing in the opposite direction than the relation of the changing side-length $\tau(p, y) = \bar{\tau}(\bar{p}, \bar{y})$ to the side-length $\bar{\tau}(\tilde{p}, \tilde{y})$. In the future version, we obtain that the relation between the desired angle $\measuredangle_{\bar{p}}(\bar{x}, \bar{z})$ in $\bar{\Delta}_1$ to the corresponding angle $\measuredangle_{\tilde{p}}(\tilde{x}, \tilde{z})$ in $\tilde{\Delta}_1$ is pointing in the other direction than the relation of the changing side-length $\tau(p, z) = \bar{\tau}(\bar{p}, \bar{z})$ to the side-length $\bar{\tau}(\tilde{p}, \tilde{z})$.

Similarly, for the triangles $\bar{\Delta}_2$ vs. $\tilde{\Delta}_2$, only one side-length changes. In the across version, it is not the longest side, and for the future version, it is. Thus, in the future version, the monotonicity statement in the law of cosines yields that the relation between the angle $\measuredangle_{\bar{p}}(\bar{y}, \bar{z})$ in $\bar{\Delta}_1$ to the corresponding angle $\measuredangle_{\tilde{p}}(\tilde{y}, \tilde{z})$ in $\tilde{\Delta}_1$ is pointing in the same direction as the relation of the changing side-length $\bar{\tau}(\bar{p}, \bar{y})$ to the side-length $\bar{\tau}(\tilde{p}, \tilde{y})$. Similarly, in the across version, we get that the relations between the angles $\measuredangle_{\bar{p}}(\bar{y}, \bar{z})$ in $\bar{\Delta}_1$ and $\measuredangle_{\tilde{p}}(\tilde{y}, \tilde{z})$ in $\tilde{\Delta}_1$ points in the opposite direction as the relation of the changing side-length $\bar{\tau}(\bar{p}, \bar{y})$ to the side-length $\bar{\tau}(\tilde{p}, \tilde{y})$. Note that $\tilde{\Delta}_1$ and $\tilde{\Delta}_2$ together form the triangle $\tilde{\Delta}$, so we have $\measuredangle_{\tilde{p}}(\tilde{x}, \tilde{y}) = \measuredangle_{\tilde{p}}(\tilde{y}, \tilde{z})$ in the across version and $\measuredangle_{\tilde{p}}(\tilde{x}, \tilde{z}) = \measuredangle_{\tilde{p}}(\tilde{y}, \tilde{z})$ in the future version. From this, the desired inequalities follow.

For the "split" angle, we use the triangle equality along lines and the reverse triangle inequality on the broken side, giving $\bar{\tau}(\bar{x}, \bar{z}) \geq \bar{\tau}(\bar{x}, \bar{p}) + \bar{\tau}(\bar{p}, \bar{z}) = \tau(x, p) + \tau(p, z) = \tau(x, z) = \bar{\tau}(\tilde{x}, \tilde{z})$ in the across statement and $\bar{\tau}(\bar{x}, \bar{y}) \geq \bar{\tau}(\bar{x}, \bar{p}) + \bar{\tau}(\bar{p}, \bar{y}) = \tau(x, p) + \tau(p, y) = \tau(x, y) = \bar{\tau}(\tilde{x}, \tilde{y})$ in the future statement. For the triangles $\bar{\Delta} = \Delta(\bar{x}, \bar{y}, \bar{z})$ and $\tilde{\Delta} = \Delta(\tilde{x}, \tilde{y}, \tilde{z})$, only one side-length changes. In the across version it is the longest side of the triangle, and in the future version, it is not. Thus, in the future version, the monotonicity statement in the law of cosines implies that the relation between the angle $\measuredangle_{\bar{z}}(\bar{x}, \bar{y})$ in $\bar{\Delta}$ to the corresponding angle $\measuredangle_{\tilde{z}}(\tilde{x}, \tilde{y})$ in $\tilde{\Delta}$ is pointing in the opposite direction than the relation of the changing side-lengths $\tau(x, y) = \bar{\tau}(\bar{x}, \bar{y})$ and $\bar{\tau}(\tilde{x}, \tilde{y})$. Similarly, in the across version, we obtain that the relation of the angle $\measuredangle_{\bar{y}}(\bar{x}, \bar{z})$ in $\bar{\Delta}$ and $\measuredangle_{\tilde{y}}(\tilde{x}, \tilde{z})$ in $\tilde{\Delta}$ points in the same direction as the inequality between $\tau(x, z) = \bar{\tau}(\bar{x}, \bar{z})$ and $\bar{\tau}(\tilde{x}, \tilde{z})$.

Finally, note that all monotonicity arguments work the same if the timelike relation between y and p is reversed in the across version. $\square$

2.4 Extensions of Lorentzian length spaces

As in [11], in the proof of the splitting theorem we will first show that $I(\gamma) := I^+(\gamma) \cap I^-(\gamma)$ splits, and then we will use an inextendibility argument to show that $I(\gamma) = X$. To this end, let us recall the notion of extensions of Lorentzian (pre-)length spaces as introduced in [40, Def. 3.1].

Definition 2.44 (*Extensions*) Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space. A Lorentzian pre-length space $(\tilde{X}, \tilde{d}, \tilde{\ll}, \tilde{\leq}, \tilde{\tau})$ is called an *extension* of $(X, d, \ll, \leq, \tau)$ if

(i) $(\tilde{X}, \tilde{d})$ is connected,
(ii) there exists an isometric embedding $\iota : (X, d) \to (\tilde{X}, \tilde{d})$,
(iii) $\iota(X) \subsetneq \tilde{X}$ is open,
(iv) ι preserves $\leq$ and $\ll$,
(v) ι preserves τ-lengths of causal curves.

Remark 2.45 (Convex extensions) An important class of examples (and the only one we will need) is the case of open, causally convex subsets of a suitable Lorentzian length space. In this case, ι even preserves τ.

Next, we recall the definition of *timelike completeness* for (localisable) Lorentzian pre-length spaces, cf. [40, Def. 5.1]. However, since continuous extendibility, extendibility as a causal curve and geodesic extendibility are in general inequivalent concepts, we will define the timelike completeness property precisely in such a way that it is sufficient for the subsequent inextendibility result.

Definition 2.46 (*Timelike completeness (TC)*) Let $(X, d, \ll, \leq, \tau)$ be a localisable Lorentzian pre-length space. We say X satisfies the *timelike completeness (TC) property* if each (future or past directed) timelike geodesic $\alpha : [a, b) \to X$ that is inextendible to $[a, b]$ as a continuous curve has infinite τ-length.

Theorem 2.47 (Inextendibility of (TC) spaces) *Let $(X, d, \ll, \leq, \tau)$ be a strongly causal Lorentzian length space satisfying the (TC) property. Then X is inextendible as a regularly localisable Lorentzian length space.*

Proof See [40, Thm. 5.3]. □

3 Lorentzian products

This section is dedicated to the treatment of Lorentzian product spaces. In the context of Lorentzian pre-length spaces, (warped) products were first introduced in [5]. On the one hand, this investigation is too general for our purposes, we only need products instead of warped products. On the other hand, the formulation we use also works for non-intrinsic spaces.

Definition 3.1 (*Lorentzian product*) Let (X, d) be a metric space. Define the space $Y := \mathbb{R} \times X$. Equip it with the product metric

$$D : Y \times Y \to \mathbb{R}, \ D((s, x), (t, y)) := \sqrt{|t - s|^2 + d(x, y)^2}.$$

Define the timelike relation as $(s, x) \ll (t, y) :\iff t - s > d(x, y)$ and the causal relation as $(s, x) \leq (t, y) :\iff t - s \geq d(x, y)$. Define the product time separation $\tau : Y \times Y \to \mathbb{R}$ via

$$\tau((s, x), (t, y)) := \sqrt{(t - s)^2 - d(x, y)^2}$$

if $(s, x) \leq (t, y)$, and 0 otherwise.

Then $(Y, D, \ll, \leq, \tau)$ is called the Lorentzian product of X with $\mathbb{R}$. If it is clear that the Lorentzian product is meant, we simply denote it by $\mathbb{R} \times X$.

The following is more or less immediate from the definition:

Proposition 3.2 (Properties of Lorentzian products) *Let (X, d) be a metric space. Then the Lorentzian product $(Y, D, \ll, \leq, \tau)$ is a Lorentzian pre-length space. Moreover, τ is continuous and $\leq$ is closed on $Y \times Y$.*

Proof τ is continuous by definition, so in particular, lower semi-continuous. The inclusion of $\ll$ in $\leq$ is clear from the definition as well. The transitivity of both relations as well as the reverse triangle inequality of τ for $\leq$-related triples follow (via an elementary calculation) from the triangle inequality for d. The reflexivity of $\leq$ is clear from the definition. Finally, the definition of $\ll$ is a reformulation of $\tau > 0$.

Note that for converging real sequences $n_i \to n, m_i \to m$ such that $n_i \leq m_i$ for all i, we have $n \leq m$. The closedness of $\leq$ on $Y \times Y$ follows immediately from this fact. □

It is easily seen that generalised cones in the sense of [5] with warping function $f \equiv 1$ are Lorentzian products in the sense of Definition 3.1. Indeed, in this case we have, in the notation of [5, Def. 3.9], that $m_{s,t} := \min_{r \in [s,t]} f(r) = 1$ for all $s, t \in \mathbb{R}$. Then [5, Lem. 3.10] shows that a causal curve in the sense of [5, Def. 3.2] is a causal curve in the usual sense in our setting, with timelikeness being inherited as well. Moreover, the variational length of [5, Def. 3.9] is precisely the τ-length of a causal curve in the sense of Definition 2.6. It is easily seen that if X is strictly intrinsic, then the product $\mathbb{R} \times X$ in our sense can be canonically identified with the warped product $\mathbb{R} \times_f X$ with $f \equiv 1$.

We now show that Lorentzian products are always strongly causal and non-totally imprisoning.

Proposition 3.3 (Diamonds form basis) *Let $Y := \mathbb{R} \times X$ be a Lorentzian product. Then $\{I(p, q) \mid p, q \in Y\}$ forms a basis for the topology. In particular, Y is strongly causal.*

Proof Consider the open set $O := (a, c) \times B_R(x)$ for $a, c \in \mathbb{R}$, $R > 0$, $x \in X$ and let $(b, y) \in O$. Choose $\varepsilon = \min(b - a, c - b, R - d(x, y)) > 0$, then $p := (b - \varepsilon, y)$, $q := (b + \varepsilon, y) \in \bar{O}$. Then we have that $I(p, q) \subseteq O$: If $(s, z) \in I(p, q)$, then $s \in (b - \varepsilon, b + \varepsilon) \subseteq (a, c)$. In particular, at least one of $|s - (b - \varepsilon)| \leq \varepsilon$ and $|s - (b + \varepsilon)| \leq \varepsilon$ holds. By definition of $\ll$, we then also have $d(x, z) \leq d(x, y) + d(y, z) < R$, so $y \in B_R(x)$. Thus, $(s, z) \in O$. □

Proposition 3.4 (Non-total imprisonment of Lorentzian products) *Any Lorentzian product $Y = \mathbb{R} \times X$, where (X, d) is a metric space, is non-totally imprisoning.*

Proof For any two points $(s, x) \leq (t, y) \in Y$ we have $d(x, y) \leq t - s$, so $D((s, x), (t, y)) = \sqrt{(t - s)^2 + d(x, y)^2} \leq \sqrt{2}(t - s)$. In particular, for any future directed causal curve $\gamma : [a, b] \to Y$ starting at $\gamma(a) = (s, x)$ and ending at $\gamma(b) = (t, y)$ we have $L_D(\gamma) \leq \sqrt{2}(t - s)$: In any partition of $[a, b]$, the above bound in any of the subintervals forms a telescopic sum, making $\sqrt{2}(t - s)$ an upper bound on the length of the polygon approximation of γ corresponding to the partition. Let now K be compact, then we can enclose it in a bounded set: $K \subseteq [s, t] \times \bar{B}_R(x)$ for some $s < t$, $R > 0$ and $x \in X$. We know that the D-length of future directed causal curves contained in $[s, t] \times \bar{B}_R(x)$ is bounded by $\sqrt{2}(t - s)$, so we have that Y is non-totally imprisoning. □

The next result characterises distance realisers in $Y = \mathbb{R} \times X$. Note that we make use of the following: A distance minimiser in a metric space can always be parametrised by unit speed, which is evidently a Lipschitz parametrisation, cf. [17, Proposition 2.5.9].

Proposition 3.5 (Characterisation of distance realisers) *Let (X, d) be a metric space and let $Y := \mathbb{R} \times X$ the Lorentzian product. Then a continuous causal curve $\gamma = (\alpha, \beta) : [a, b] \to Y$ is a distance realiser if and only if either $\beta = const$ or $\beta : [a, b] \to X$ is a (metric) distance realiser and when reparametrising γ such that β is unit speed parametrised, $\alpha : [a, b] \to \mathbb{R}$ is affine, i.e. of the form $\alpha(t) = ct + d$ with $c = 1$ (which implies that γ future directed null) or $c > 1$ (which implies that γ is future directed timelike). In particular, γ has causal character and a reparametrisation as a (Lipschitz) causal curve. If in addition Y is localisable, then the same is true for continuous geodesics.*

Proof γ is a distance realiser if and only if for all $r < s < t$ we have

$$\tau(\gamma(r), \gamma(t)) = \tau(\gamma(r), \gamma(s)) + \tau(\gamma(s), \gamma(t)). \qquad (\star)$$

We denote $\gamma(r) = (t_1, x_1)$, $\gamma(s) = (t_2, x_2)$ and $\gamma(t) = (t_3, x_3)$, then $(\star)$ reads:

$$\sqrt{(t_3 - t_1)^2 - d(x_1, x_3)^2} = \sqrt{(t_2 - t_1)^2 - d(x_1, x_2)^2} + \sqrt{(t_3 - t_2)^2 - d(x_2, x_3)^2}$$

We now define $\lambda = \frac{t_2 - t_1}{t_3 - t_1}$ and the function $f(x) = \sqrt{1 - x^2}$. Then $(\star)$ simplifies to

$$f\left(\frac{d(x_1, x_3)}{t_3 - t_1}\right) = \lambda f\left(\frac{d(x_1, x_2)}{\lambda(t_3 - t_1)}\right) + (1 - \lambda) f\left(\frac{d(x_2, x_3)}{(1 - \lambda)(t_3 - t_1)}\right). \qquad (\star\star)$$

We use that f is concave and monotonously decreasing on $[0, 1]$ and $d(x_1, x_2) + d(x_2, x_3) \geq d(x_1, x_3)$ on the right hand side of $(\star\star)$:

$$\lambda f\left(\frac{d(x_1, x_2)}{\lambda(t_3 - t_1)}\right) + (1 - \lambda) f\left(\frac{d(x_2, x_3)}{(1 - \lambda)(t_3 - t_1)}\right)$$
$$\leq f\left(\frac{d(x_1, x_2) + d(x_2, x_3)}{(t_3 - t_1)}\right) \leq f\left(\frac{d(x_1, x_3)}{(t_3 - t_1)}\right).$$

So we get that $(\star\star)$ is equivalent to having equality here. As f is strictly concave, the first inequality has equality if and only if

$$\frac{d(x_1, x_2)}{t_2 - t_1} = \frac{d(x_2, x_3)}{t_3 - t_2}, \tag{3.1}$$

where we got rid of λ again for generality. As f is strictly monotonously decreasing, the second inequality has equality if and only if

$$d(x_1, x_2) + d(x_2, x_3) = d(x_1, x_3). \tag{3.2}$$

We translate this back to what happens for general $r < s < t$: Eq. 3.2 reads $d(\beta(r), \beta(s)) + d(\beta(s), \beta(t)) = d(\beta(r), \beta(t))$, i.e., it holds if and only if β is a distance realiser. Equation 3.1 reads $\frac{d(\beta(r),\beta(s))}{\alpha(s)-\alpha(r)} = \frac{d(\beta(s),\beta(t))}{\alpha(t)-\alpha(s)}$, i.e. that quantity is a constant $\frac{1}{c}$ (if $\frac{1}{c}$ is 0 we get that β is constant, in this case the result is trivial). In particular, if we reparametrise γ such that β is unit speed (i.e. $d(\beta(s), \beta(t)) = t - s$), $\alpha(t) - \alpha(s) = c(t - s)$, i.e. α is affine.

Now we look at what happens for different parameters c: If $c > 1$, we easily see that $f(c) > 0$, so $\tau(\gamma(r), \gamma(t)) = (\alpha(t) - \alpha(r)) f(c) > 0$ and the curve is timelike. For $c = 1$, $f(c) = f(1) = 0$, so $\alpha(t) - \alpha(r) = d(\beta(r), \beta(t))$ and the curve is null.

Now if Y is localisable and γ is only a geodesic, we can cover the domain with intervals J_i that are open in $[a, b]$ where it is distance realising. Then on each of these J_i we apply the above. As the J_i overlap, the constant c agrees, and we get that $\beta|_{J_i}$ is distance realising, thus β is a geodesic. The converse follows equivalently. We even get that the subintervals of the domain where the curves are distance realising agree. □

Corollary 3.6 (τ-arclength parametrisations of distance realisers) *Any continuous timelike maximiser $\gamma : [a, b] \to \mathbb{R} \times X$ has a Lipschitz τ-arclength parametrisation, which is of the form (α, β), with $\alpha : [0, L_\tau(\gamma)] \to \mathbb{R}$, $\alpha(t) = ct + d$, and $\beta : [0, L_\tau(\gamma)] \to X$ a constant speed minimiser of speed $\sqrt{c^2 - 1}$. If $\mathbb{R} \times X$ is localisable, then the same is true for timelike geodesics (with β constant speed geodesic).*

Note that the results above continue to hold for maximisers resp. geodesics defined on any interval I, not just a closed one.

To finish this section, we show that global hyperbolicity of the Lorentzian product $Y = \mathbb{R} \times X$ is equivalent to the metric properness of X.

Proposition 3.7 (Global hyperbolicity and properness) *A Lorentzian product $Y := \mathbb{R} \times X$ is globally hyperbolic if and only if X is a proper metric space.*

Proof Since products are always non-totally imprisoning by Proposition 3.4, we only need to check that causal diamonds in Y are compact if and only if (X, d) is a proper metric space. First, suppose that Y is globally hyperbolic. Let $(t, x) \in Y$ and fix $R > 0$. Consider the points $p := (t - R, x)$, $q := (t + R, x)$. Then $(t, x) \in J(p, q)$. By definition of $\leq$, any point (t, y) with the same $\mathbb{R}$-coordinate as (t, x) is in $J(p, q)$ if and only if $d(x, y) \leq R$. Thus, the set $J(p, q) \cap \{(t, x) \mid x \in X\}$, which is compact by assumption, can be identified with the closed ball $\bar{B}_R(x)$ in X of radius R around x, establishing the fact that closed balls in X are compact.

Conversely, suppose that X is proper and let $p = (r, x), q = (t, z) \in Y$ with $p \leq q$. By definition of $\leq$, if $(s, y) \in J(p, q)$, then $r \leq s \leq t$. Moreover, we have $|s| \leq |r| + |t|$. Set $R := 2|r| + 2|t|$. Then for $(s, y) \in J(p, q)$, we also have $d(x, y) \leq |s - r| \leq |s| + |r| \leq R$. Thus, $y \in \bar{B}_R(x)$, which is compact by assumption. In total, we conclude that $J(p, q) \subseteq [r, t] \times \bar{B}_R(x)$, which is a compact set as the Cartesian product of two compact sets. The fact that $J(p, q)$ is closed follows immediately from the closedness of $\leq$, so $J(p, q)$ is compact as well. □

4 Rays, lines, co-rays and asymptotes

4.1 Rays, lines, co-rays, asymptotes, timelike co-ray condition

In this subsection, we study causal rays and lines and show how to obtain them as limits of causal maximisers. Moreover, we analyse triangles where one side is a segment on a timelike line and show that the angles adjacent to the line are equal to their comparison angles. This in fact follows from the more general principle that one can stack triangle comparisons of nested triangles with two endpoints on a timelike line. The latter situation arises in the construction of asymptotes, and via the stacking principle, one can show any future directed and any past directed asymptote from a common point to a given timelike line fit together to give a (timelike) asymptotic line. In constructing co-rays and asymptotes we follow [11], whereas the stacking principle and equality of angles can be viewed as Lorentzian analogues of [17, Lem. 10.5.4].

Definition 4.1 (*Rays, Lines*) Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space. A *future directed causal ray* is a future inextendible, future directed causal curve $c : I \to X$ that maximises the time separation between any of its points, where I is either a closed interval $[a, b]$ or a half-open interval $[a, b)$. A *future directed causal line* is a (doubly) inextendible, future directed causal curve $\gamma : I \to X$ that maximises the time separation between any of its points. Here I can in general be open, closed, or half-open. More generally, for $S \subset X$, a *future directed causal S-ray* is a future directed, future inextendible causal curve starting in S and satisfying $\tau(S, c(t)) := \sup_{p \in S} \tau(p, c(t)) = L_\tau(c|_{[0,t]})$. Past directed rays and lines are defined analogously. A future directed causal ray $c : I \to X$ is called *complete* if $L_\tau(c) = \infty$. Similarly, a future directed causal line γ is called *complete* if there is $t_0 \in I$ such that the past and future directed rays obtained from dividing γ at $\gamma(t_0)$ are complete.

Unless explicitly stated otherwise, all causal rays and lines are understood to be future directed.

Remark 4.2 (i) Clearly, any ray c is a $\{c(0)\}$-ray. Conversely, if $c : [a, b) \to X$ is a future directed S-ray for some $S \subset X$, then $\tau(c(a), c(t)) \leq \tau(S, c(t)) = L_\tau(c|_{[a,t]}) \leq \tau(c(a), c(t))$, so c is a ray. A similar statement is true for past directed rays.

(ii) If X is regularly localisable, then any causal ray/line c is either timelike or null by Theorem 2.17.

(iii) If X is localisable and causally path-connected, then any ray c is defined on a half-open interval and any line γ is defined on an open interval.

In the following, unless we specify the assumptions on X, we always take X to be as in the splitting theorem.

Proposition 4.3 *Let $z_n \to z$ in X. Let $p_n \in I^+(z_n)$ and let $c_n : [0, a_n] \to X$ be a sequence of future directed, maximising causal curves from z_n to p_n in d-arclength parametrisation. If $\tau(z_n, p_n) \to \infty$, then there is a limit causal ray $c : [0, \infty) \to X$ from z.*

Proof We only have to show $a_n \to \infty$, the rest follows from the limit curve theorem. Suppose $a_n \to a < \infty$. This implies that all p_n are contained in a common compact set K: Indeed, if we assume this to be false, then for each $n > 0$ there would be a p_k, $k = k(n)$, such that $p_k \notin \overline{B}_n(z) = \{x \in X : d(x, z) \leq n\}$ (recall that (X, d) is proper), which in particular means that

$$a_{k(n)} = L^d(c_{k(n)}) \geq d(z_{k(n)}, p_{k(n)}) \geq d(p_{k(n)}, z) - d(z_{k(n)}, z) \geq n - d(z_{k(n)}, z),$$

which is a contradiction since the right hand side tends to ∞. So, up to a choice of subsequence, we may assume that $\gamma_n(a_n) = p_n \to p$. But then, by continuity of τ,

$$\tau(z, p) = \lim_n \tau(z_n, p_n) = \infty,$$

a contradiction to the finiteness of τ on all of $X \times X$. $\square$

Let now $c : [0, \infty) \to X$ be a complete timelike S-ray, and fix $z \in I^-(c) \cap I^+(S)$. Let $z_n \to z$ and $p_n := c(r_n)$ for some sequence of parameters $r_n \to \infty$. Then it is easily seen that $\tau(z_n, p_n) \to \infty$: Indeed, let $r > 0$ be large and assume w.l.o.g. that all z_n are in $I^-(c(r))$, then

$$\tau(z_n, p_n) \geq \tau(z_n, c(r)) + \tau(c(r), c(r_n)) \to \infty.$$

Hence, by the above result, we may construct causal rays as follows: Let μ_n be maximising timelike curves from z_n to p_n and let μ be a limit causal ray from z whose existence we just proved.

Definition 4.4 (*Co-rays, asymptotes*) Any causal ray μ constructed by the above method is called a *co-ray* to the ray c at z. If $z_n = z$ for all n, then μ is called an *asymptote* to c at z.

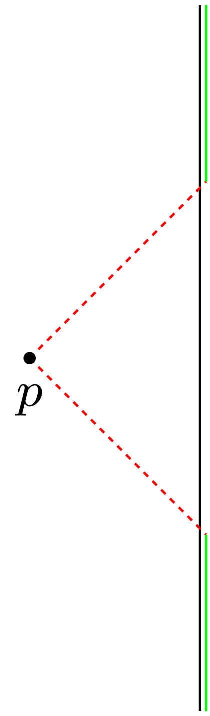


Fig. 3 The domain where the y_i can lie in is in the double line

Since maximising causal curves have causal character, a co-ray is either timelike or null.

Definition 4.5 (*Timelike co-ray condition*) Let $c : [0, \infty) \to X$ be a complete timelike S-ray and let $p \in I^-(c) \cap I^+(S)$. We say the *timelike co-ray condition (TCRC)* holds at p if every co-ray to c at p is timelike.

We turn to the treatment of triangles adjacent to lines and begin with the aforementioned stacking principle. It will be an essential technical tool for controlling the behaviour of asymptotes to a line. The following two results are true for rather general Lorentzian pre-length spaces, and the exact conditions are specified.

Proposition 4.6 (Comparison situations stack along a geodesic) *Let X be a timelike geodesically connected Lorentzian pre-length space with global timelike curvature bounded below by* 0 *and $\gamma : \mathbb{R} \to X$ be a complete timelike line. Let $p \in X$ be a point not on γ. Let $t_1 < t_2 < t_3$ such that all $y_i := \gamma(t_i)$ are timelike related to p, see Fig. 3. Let $\bar{\Delta}_{12} := \Delta(\bar{p}, \bar{y}_1, \bar{y}_2)$ be a comparison triangle for $\Delta_{12} := \Delta(p, y_1, y_2)$ and extend the side $\bar{p}, \bar{y}_2$ to a comparison triangle $\bar{\Delta}_{23} := \Delta(\bar{p}, \bar{y}_2, \bar{y}_3)$ for $\Delta_{23} := \Delta(p, y_2, y_3)$. We choose it in such a way that $\bar{y}_1$ and $\bar{y}_3$ lie on opposite sides of the line through $\bar{y}_1, \bar{y}_2$. Then $\bar{y}_1, \bar{y}_2, \bar{y}_3$ are collinear. That makes $\bar{\Delta}_{13} := \Delta(\bar{p}, \bar{y}_1, \bar{y}_3)$ a comparison triangle for $\Delta_{13} := \Delta(p, y_1, y_3)$.*

Proof We set $s_- = \sup(\gamma^{-1}(I^-(p)))$ and $s_+ = \inf(\gamma^{-1}(I^+(p)))$. Then the set of s where $\gamma(s)$ is timelike related to p is $(-\infty, s_-) \cup (s_+, +\infty)$. We assume $p \ll y_2$, the other case $y_2 \ll p$ can be reduced to this one by flipping the time orientation of the space.

We create comparison situations for an Alexandrov situation: We take the triangles $\bar{\Delta}_{12} = \Delta(\bar{p}, \bar{y}_1, \bar{y}_2)$ and $\bar{\Delta}_{23} = \Delta(\bar{p}, \bar{y}_2, \bar{y}_3)$ as in the statement. We also create a comparison triangle $\tilde{\Delta}_{13} = \Delta(\tilde{p}, \tilde{y}_1, \tilde{y}_3)$ for (p, y_1, y_3) and get a comparison point $\tilde{y}_2$ for y_2 on the side $y_1 y_3$. Then we can apply curvature comparison to get $\bar{\tau}(\tilde{p}, \tilde{y}_2) \geq \tau(p, y_2) = \bar{\tau}(\bar{p}, \bar{y}_2)$. By Lemmas 2.42 and 2.43,[4] this means the situation is convex, i.e.

$$\tilde{\measuredangle}_{y_2}(p, y_1) \leq \tilde{\measuredangle}_{y_2}(p, y_3). \tag{4.1}$$

If $p \ll y_1$ we also get

$$\tilde{\measuredangle}_{y_1}(p, y_2) \leq \tilde{\measuredangle}_{y_1}(p, y_3) \tag{4.2}$$

and if $y_1 \ll p$, this inequality flips.

Now we apply Eq. (4.2) to the situation where we fix y_2 and move y_3 and y_1: We get $\tilde{\measuredangle}_{y_2}(p, \gamma(t_3))$ is monotonously increasing in t_3. Similarly, the time-reversed situation gives that if $p \ll \gamma(t_1)$, $\tilde{\measuredangle}_{y_2}(p, \gamma(t_1))$ is monotonously increasing in t_1 and if $\gamma(t_1) \ll p$ it is also monotonously increasing in t_1 (note it is also increasing when switching from one case to the other). Then remember $\tilde{\measuredangle}_{y_2}(p, \gamma(t_1)) \leq \tilde{\measuredangle}_{y_2}(p, \gamma(t_3))$ by the convexity of the Alexandrov situation (Eq. (4.1)). Note that the left hand side is decreasing with decreasing t_1 and the right hand side is increasing with increasing t_3.

Claim: For each t_1 and t_3, Eq. (4.1) has equality.

Then: The comparison situation is straight, i.e. $\bar{y}_1, \bar{y}_2, \bar{y}_3$ are collinear, and by triangle equality along straight lines we have $\tau(y_1, y_3) = \tau(y_1, y_2) + \tau(y_2, y_3) = \bar{\tau}(\bar{y}_1, \bar{y}_2) + \bar{\tau}(\bar{y}_2, \bar{y}_3) = \bar{\tau}(\bar{y}_1, \bar{y}_3)$, i.e., $\bar{p}, \bar{y}_1, \bar{y}_3$ is a comparison triangle for p, y_1, y_3.

Proof: We indirectly assume there is a t_1^0, t_3^0 such that the comparison situation $\bar{\Delta}_{12}, \bar{\Delta}_{23}$ is strictly convex. We realise this situation such that $\bar{p} = 0$, $\bar{y}_2$ is to the right of the t-axis, such that the side $\bar{y}_2 \bar{y}_3^0$ slopes to the left and $\bar{y}_1^0 \bar{y}_2$ slopes to the right (i.e., $\frac{x(\bar{y}_3^0 - \bar{y}_2)}{t(\bar{y}_3^0 - \bar{y}_2)} < 0$, $\frac{x(\bar{y}_2 - \bar{y}_1^0)}{t(\bar{y}_2 - \bar{y}_1^0)} > 0$). When we vary t_1 and t_3, the comparison situation is chosen such that $\bar{p}$ and $\bar{y}_2$ stay fixed. As the comparison angle $\tilde{\measuredangle}_{y_2}(p, \gamma(t_3))$ is monotonously increasing in t_3, we get that for $t_3 \geq t_3^0$ the slope $\frac{x(\bar{y}_3 - \bar{y}_2)}{t(\bar{y}_3 - \bar{y}_2)}$ is smaller than the slope of $\bar{y}_2 \bar{y}_3$. In particular, $\bar{y}_3$ lies to the left of the extension of the side $\bar{y}_2 \bar{y}_3^0$. Thus, for large enough t_3, $\bar{y}_3$ lies to the left of the t-axis, and for a certain t_3^* $\bar{y}_3^*$ lies on it. Similarly, we find a t_1^* such that $\bar{y}_1^*$ lies on the t-axis. See Fig. 4for a visualisation of the construction. But then

$$\begin{aligned} \tau(y_1^*, p) + \tau(p, y_3^*) &= \bar{\tau}(\bar{y}_1^*, \bar{p}) + \bar{\tau}(\bar{p}, \bar{y}_3^*) > \bar{\tau}(\bar{y}_1^*, \bar{y}_2) + \bar{\tau}(\bar{y}_2, \bar{y}_3^*) \\ &= \tau(y_1^*, y_2) + \tau(y_2, y_3^*) = \tau(y_1^*, y_3^*) \end{aligned}$$

in contradiction to the reverse triangle inequality. Thus we get the claim. □

Remark 4.7 The proof of the statement can also be used in more general situations: Let X be a locally timelike geodesically connected Lorentzian pre-length space with local timelike curvature bounded below by 0, and let γ be a (possibly extendible) timelike

[4] If $p \ll y_1 \ll y_2 \ll y_3$, we use Lemma 2.43 and if $y_1 \ll p \ll y_2 \ll y_3$ we use Lemma 2.42.

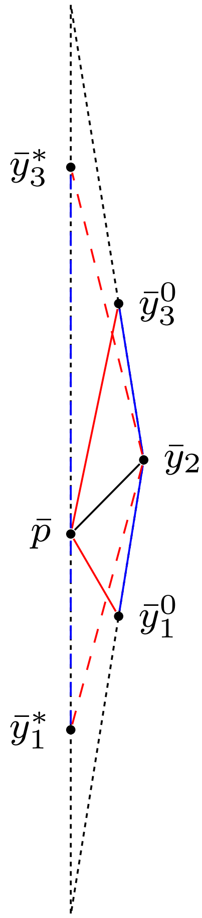


Fig. 4 We assume that the path $\bar{y}_1^0\bar{y}_2\bar{y}_3^0$ is longer than the path $\bar{y}_1^0\bar{p}\bar{y}_3^0$. If the angle at $\bar{y}_2$ is not straight, we extend some lines and as points further from $\bar{p}$ are more on the left, we find some t_1^* and t_3^* such that $\bar{y}_1^*\bar{p}\bar{y}_3^*$ are collinear. But then the path $\bar{y}_1^0\bar{y}_2\bar{y}_3^0$ is shorter than $\bar{y}_1^0\bar{p}\bar{y}_3^0$, which then yields a contradiction to γ being maximising

maximiser. Let $p \ll x = \gamma(t_2)$ be points in a comparison neighbourhood. Assume the statement is not true (with certain $t_1 < t_2 < t_3$). Then we get $\omega_1 = \tilde{\measuredangle}_{y_2}(p, y_1)$, $\omega_3 = \tilde{\measuredangle}_{y_2}(p, y_3)$. Let $d_1 + d_2 = \omega_3 - \omega_1$ and $e = \tau(p, x)\sinh(d_1 + \omega_1)$. Then for the parameters $t_1 = -\frac{e}{\sinh(d_1)}$ and $t_3 = \frac{e}{\sinh(d_3)}$ we have one of the following:

- γ is not defined at/inextendible to one of them,
- γ stops being distance realising between t_1^* and t_3^* (and we have found a longer curve),
- there is no curvature comparison neighbourhood containing both $\gamma(t_i^*)$ $(i = 1, 3)$.

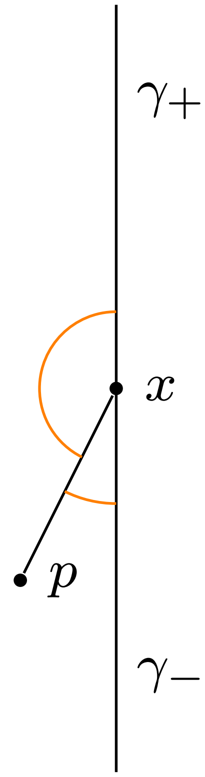
For this, realise the above with $\measuredangle_{\bar{y}_2}(\partial_t, \bar{y}_3) = d_3$ and $\measuredangle_{\bar{y}_2}(\partial_t, \bar{y}_1) = d_1$. Then e is the difference of the x-coordinates of p and x.

Note the similarity of this result and classical arguments for conjugate points along geodesics in Riemannian/Lorentzian geometry.

Proposition 4.8 (Angle = comparison angle) *Let X be a timelike geodesically connected Lorentzian pre-length space with global timelike curvature bounded below by 0 and $\gamma : \mathbb{R} \to X$ be a complete timelike line and $x := \gamma(t_0)$ a point on it. We split γ into the future part $\gamma_+ = \gamma|_{[t_0,+\infty)}$ and the past part $\gamma_- = \gamma|_{(-\infty,t_0]}$. Let $p \in X$ be a point not on γ with x and p timelike related and $\alpha : x \rightsquigarrow p$ a connecting distance realiser. Then for all $s \neq t$ such that p and $\gamma(s)$ are timelike related, we have:*

$$\tilde{\measuredangle}_x(p, \gamma(s)) = \measuredangle_x(\alpha, \gamma_+) = \measuredangle_x(\alpha, \gamma_-),$$

Fig. 5 Illustration: These angles are the same, and have the same value as if they are considered as comparison angles



i.e. the comparison angle is equal to the angle, and the same in both directions (see Fig. 5).

Proof First, we check that the comparison angle $\tilde{\measuredangle}_x(p, \gamma(s))$ is constant in s: For s_1 and s_2 for which this is defined, we have three parameters on γ involved: s_1, s_2, t_0. The previous result (Proposition 4.6) tells us we can construct a comparison situation for all three triangles at once. As the comparison situations have the comparison angle $\tilde{\measuredangle}_x(p, \gamma(s_1))$ resp. $\tilde{\measuredangle}_x(p, \gamma(s_2))$ as the angle in $\bar{x}$, they are equal.

Now we look at the angle $\measuredangle_x(\alpha, \gamma_\pm)$: We assume $p \ll x$. We already know $\tilde{\measuredangle}_x(\alpha(s), \gamma(t))$ is constant in t. We now have to look at its dependence on s: By Theorem 2.39, $\tilde{\measuredangle}_x^{\mathrm{S}}(\alpha(s), \gamma(t))$ is monotonously increasing in s. Note now that for $t < t_0$, the sign of this angle is $\sigma = -1$, and for $t > t_0$, the sign of this angle is $\sigma = +1$. So choose some $t_- < t_0$ and $t_+ > t_0$ for which all the necessary angles exist (i.e. $p \ll \gamma(t_-)$ and $t_+ > t_0$) we have that $\tilde{\measuredangle}_x(\alpha(s), \gamma(t_-)) = \tilde{\measuredangle}_x(\alpha(s), \gamma(t_+))$ is both a monotonously decreasing and increasing function in s. Thus it is constant, and $\tilde{\measuredangle}_x(\alpha(s), \gamma(t)) = \measuredangle_x(\alpha, \gamma_-) = \measuredangle_x(\alpha, \gamma_+)$, which includes the desired equalities. $\square$

Corollary 4.9 *Let X be a timelike geodesically connected, globally causally closed Lorentzian pre-length space satisfying $J^\pm(p) = \overline{I^\pm(p)}$ for all $p \in X$.*[5] *and having global timelike curvature bounded below by 0 and let $\gamma : \mathbb{R} \to X$ be a complete timelike line. Then for any point $p \in X$ and two points $x_1 = \gamma(t_1)$ and $x_2 = \gamma(t_2)$ which are timelike related to p, we get a timelike triangle $\Delta = \Delta(x_1, x_2, p)$. Let q_1, q_2 be any points on Δ, one of them lying on the side x_1x_2. We form a comparison triangle $\bar{\Delta}$ for Δ and find comparison points $\bar{q}_1, \bar{q}_2$ for q_1, q_2. Then $q_1 \leq q_2$ if and*

5 Approximating and globally causally closed Lorentzian pre-length spaces satisfy this, cf. [19, Def. 2.17].

only if $\bar{q}_1 \leq \bar{q}_2$ *and* $\tau(q_1, q_2) = \bar{\tau}(\bar{q}_1, \bar{q}_2)$*, where* $\bar{\tau}$ *denotes the time separation of* $2D$*-Minkowski space.*

Proof First assume $q_1 \leq q_2$. As global curvature is bounded below by 0, we get that $\tau(q_1, q_2) \leq \bar{\tau}(\bar{q}_1, \bar{q}_2)$, and a continuity argument then shows that $\bar{q}_1 \leq \bar{q}_2$.

Conversely, let $\bar{q}_1 \leq \bar{q}_2$. We distinguish which sides the q_i lie on: Note we assumed one of them is on the side x_1x_2. We only prove the case where q_1 is on the side α connecting x_1x_2 (say $q_1 = \alpha(s)$) and q_2 is on the side β connecting x_1, p (say $q_2 = \beta(t)$), the proof of the other cases is easily adapted.

Consider now the hinge $(\alpha|_{[0,s]}, \beta|_{[0,t]})$, which has base-point x_1 and two tips at q_1, q_2. By the previous Proposition 4.8, we get that $\omega := \measuredangle_{x_1}(\alpha, \beta) = \tilde{\omega} := \tilde{\measuredangle}_{x_1}(p, x_2)$. We now have two comparison situations: $\bar{\Delta}$ has angle $\tilde{\measuredangle}_{x_1}(p, x_2)$ at x_1, and a comparison hinge $(\tilde{\alpha}, \tilde{\beta})$ with tips $\tilde{q}_1, \tilde{q}_2$. But note that the angles and distances at $\bar{x}_1$ resp. $\tilde{x}_1$ are the same, so we have that $\bar{\tau}(\bar{q}_1, \bar{q}_2) = \bar{\tau}(\tilde{q}_1, \tilde{q}_2)$. But for the comparison hinge, [13, Cor. 4.11] gives that $\tau(q_1, q_2) \geq \bar{\tau}(\tilde{q}_1, \tilde{q}_2) = \bar{\tau}(\bar{q}_1, \bar{q}_2)$. In total, this gives that $\tau(q_1, q_2) = \bar{\tau}(\bar{q}_1, \bar{q}_2)$, as desired. A similar continuity argument then gives that $q_1 \leq q_2$, thus completing the proof. $\square$

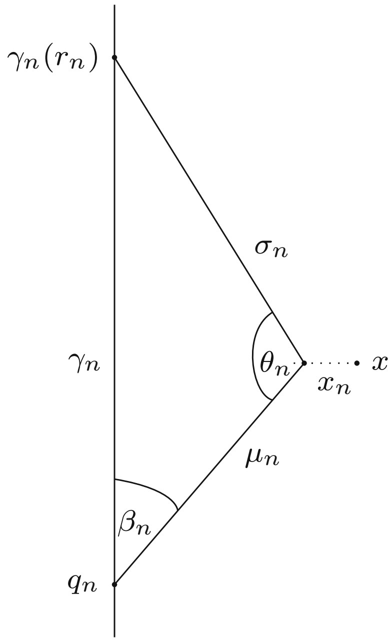
We return to the situation of the splitting theorem. Let $\gamma : \mathbb{R} \to X$ be the complete timelike line whose existence we assume, and let it be in any locally Lipschitz parametrisation (e.g. parametrisation by d-arclength) defined on $\mathbb{R}$. We call any co-ray to any of its past or future directed subrays a co-ray to the line γ, so in particular, we can construct past and future directed co-rays from all points on $I(\gamma) := I^+(\gamma) \cap I^-(\gamma)$. The next result shows that the timelike co-ray condition holds on $I(\gamma)$, i.e. all co-rays from all points in $I(\gamma)$ are timelike. For our purposes it would be sufficient to show this for asymptotes, since we only work with those in the proof.

Proposition 4.10 *The timelike co-ray condition holds on* $I(\gamma)$*.*

Proof Suppose there is a point $x \in I(\gamma)$ such that the TCRC does not hold at x, so w.l.o.g. there is a sequence $x_n \to x$, $r_n \to \infty$ and maximal timelike curves σ_n from x_n to $\gamma(r_n)$ such that σ_n converge to a future directed null ray σ. Choose some $q \in \gamma \cap I^-(x)$ and let μ_n be maximal timelike curves from q to x_n (assuming n to be large enough, $q \ll x_n$). Suitably pre- and post-composing μ_n and then applying the limit curve theorem, it is easily seen that (up to a choice of subsequence) the μ_n converge locally uniformly to a maximising limit causal curve μ from q to x, which is timelike since $q \ll x$. Denote by γ_n the piece of γ that runs between q and $\gamma(r_n)$. Set $a_n := L_\tau(\mu_n)$, $b_n := L_\tau(\sigma_n)$ and $c_n := L_\tau(\gamma_n)$. Let $\beta_n := \measuredangle_q(\mu_n, \gamma_n)$ and $\theta_n := \measuredangle_{x_n}(\mu_n, \sigma_n)$ (see Fig. 6). Then $a_n \to a := \tau(q, x)$ and by the continuity of angles (cf. Proposition 2.41), $\beta_n \to \beta$, where β is the angle between μ and γ. Consider the comparison triangle $(\overline{\mu}_n, \overline{\sigma}_n, \overline{\gamma}_n)$ in $\mathbb{R}^{1,1}$, then the angle $\overline{\beta}_n$ between $\overline{\mu}_n$ and $\overline{\gamma}_n$ satisfies $\overline{\beta}_n \leq \beta_n$ and similarly $\overline{\theta}_n \geq \theta_n$. Since $\beta_n \to \beta$, there is some $C > 0$ such that $\overline{\beta}_n \leq C$ for all n. The hyperbolic law of cosines in $\mathbb{R}^{1,1}$ (see Lemma 2.35) gives

$$b_n^2 = a_n^2 + c_n^2 - 2a_nc_n \cosh(\overline{\beta}_n),$$
$$c_n^2 = a_n^2 + b_n^2 + 2a_nb_n \cosh(\overline{\theta}_n).$$

Fig. 6 The construction of maximal curves along $I(\gamma)$



Using these two equations and solving for $\cosh(\overline{\theta}_n)$, we get

$$\cosh(\overline{\theta}_n) = -\frac{a_n}{b_n} + \frac{c_n}{b_n}\cosh(\overline{\beta}_n).$$

By the initial equation for b_n, it is easy to see that $b_n/c_n \to 1$ and $b_n \to \infty$ (using that $a_n \to a$ and $\overline{\beta}_n \leq C$), hence there is some constant $\tilde{C} > 0$ such that $\cosh(\overline{\theta}_n) \leq \tilde{C}$ and thus also $\theta_n \leq C'$ for some constant C'. Using the monotonicity condition, we get that $\theta_n = \measuredangle_{x_n}(\sigma_n, \mu_n) \geq \overline{\measuredangle}_{x_n}(\sigma_n(s), \mu_n(t))$) for each s and t, so this is bounded too. We will see that this is incompatible with σ_n "getting more and more null": Note that we get the following estimate for some constant C'':

$$C'' \geq \cosh(\overline{\measuredangle}_{x_n}(\sigma_n(s), \mu_n(t))) = \frac{\tau(\mu_n(t), \sigma_n(s))^2 - \tau(\mu_n(t), x_n)^2 - \tau(x_n, \sigma_n(s))^2}{2\tau(x_n, \sigma_n(s))\tau(\mu_n(t), x_n)}.$$

Since the denominator goes to 0 for $n \to \infty$ (as $\tau(x_n, \sigma_n(s)) \to \tau(x, \sigma(s)) = 0$ and $\tau(\mu_n(t), x_n) \to \tau(\mu(t), x) > 0$), the numerator has to go to 0 as well, in particular this means

$$\tau(\mu(t), \sigma(s)) = \tau(\mu(t), x)$$

for all s, t. This implies that running along μ from $\mu(t)$ to x and then along σ to $\sigma(s)$ gives a maximiser, but this curve has a timelike and a null piece, a contradiction to Theorem 2.17. □

Next, we show that any asymptote to γ (which we now know to be timelike) is complete.

Proposition 4.11 *Let $x \in I(\gamma)$ and let η be a timelike asymptote to γ at x. Then $L_\tau(\eta) = \infty$.*

Proof W.l.o.g. let η be future-directed. Suppose $L := L_\tau(\eta) < \infty$. By construction, $\eta : [0, \infty) \to X$ arises as a locally uniform limit of timelike maximisers $\eta_n : [0, a_n] \to X$ from x to $\gamma(t_n)$, where $t_n \to \infty$. By continuity of angles (see Proposition 2.41), $\measuredangle_x(\eta, \eta_n) \to 0$ for $n \to \infty$. Let $\varepsilon > 0$ and let $N \in \mathbb{N}$ be such that for $n \geq N$, $\measuredangle_x(\eta, \eta_n) < \varepsilon$ (in fact, any finite bound would suffice here, we do not necessarily need an arbitrarily small one). Since $\tau(x, \gamma(t_n)) \to \infty$, we may assume that $\tau(x, \gamma(t_n)) \geq 3L\cosh(\varepsilon)$ (upon possibly choosing a larger N). Now note that for any t_n with $n \geq N$, $\partial J^-(\gamma(t_n)) \cap \eta = (J^-(\gamma(t_n)) \backslash I^-(\gamma(t_n))) \cap \eta$ is non-empty: Certainly, $x \in I^-(\gamma(t_n))$, so if this intersection were empty, then η would be imprisoned in the compact set $J^+(x) \cap J^-(\gamma(t_n))$, which cannot happen. So we find a point $y_0 \in \eta$ that is null-related to $\gamma(t_n)$, i.e. $y_0 < \gamma(t_n)$ and $\tau(y_0, \gamma(t_n)) = 0$. By continuity, there is y on η near y_0 such that $0 < \tau(y, \gamma(t_n)) < 3L\cosh(\varepsilon)/2$. Let now ν be the part of η from x to y with length $L_\tau(\nu) =: a$, σ a timelike maximiser from y to $\gamma(t_n)$ with length $b := L_\tau(\sigma) = \tau(y, \gamma(t_n))$. Moreover, we write $c := L_\tau(\eta_n)$. Then (ν, σ, η_n) forms a timelike triangle with angle $\beta := \measuredangle_x(\eta, \eta_n) < \varepsilon$ at x, consider a corresponding comparison triangle $(\overline{\nu}, \overline{\sigma}, \overline{\eta}_n)$ with angle $\overline{\beta}$ at $\overline{x}$. By the law of cosines, we get

$$b^2 = a^2 + c^2 - 2ac\cosh(\overline{\beta}). \tag{4.3}$$

Our curvature assumption gives that $\overline{\beta} \leq \beta < \varepsilon$. Moreover, as we have argued, $c \geq 3L\cosh(\varepsilon)$ and $b < c/2$ and $a < L$. Inserting all of this, we get

$$\begin{aligned} b^2 &\geq a^2 + c^2 - 2ac\cosh(\varepsilon) \geq a^2 + c^2 - 2Lc\cosh(\varepsilon) \geq c^2 - 2Lc\cosh(\varepsilon) \\ &= c^2\left(1 - \frac{2L\cosh(\varepsilon)}{c}\right) \geq c^2/3, \end{aligned}$$

contradicting $b < c/2$. □

To conclude this subsection, we show that any future directed and any past directed (timelike) asymptote to γ from a common point fit together to give a timelike line.

Proposition 4.12 (Asymptotic lines) *Let $p \in I(\gamma)$ and consider any future and past rays $\sigma^+ : [0, \infty) \to X$ and $\sigma^- : (-\infty, 0] \to X$ from p to γ. Then $\sigma := \sigma^-\sigma^+ : \mathbb{R} \to X$ is a complete timelike line.*

Proof Let σ_n^+ and σ_n^- be two sequences of timelike maximisers from p to $\gamma(r_n)$ and $\gamma(-r_n)$, respectively, such that σ^+ and σ^- arise as limits of these sequences as $r_n \to \infty$. To show that σ is a line, it is sufficient to show that for any $t > 0$, $\tau(\sigma(-t), \sigma(t)) = L_\tau(\sigma|_{[-t,t]})$. To see this, let $q_+ := \sigma(t)$ and $q_- := \sigma(-t)$, and $q_+^n := \sigma_n^+(t)$, $q_-^n := \sigma_n^-(-t)$. Then $q_\pm = \lim_n q_\pm^n$. Consider the triangle going from x via σ_n^+ to the endpoint of σ_n^+, then following γ down to the endpoint of σ_n^-, and finally running the latter up to x again. Consider a comparison triangle with points $\overline{q_\pm^n}$ corresponding to $q_\pm^n$. Sending $n \to \infty$, we see that the stacked comparison triangles in $\mathbb{R}^{1,1}$ converge to a vertical line (here we use Proposition 4.6 and the completeness of γ), hence our curvature bound gives

$$\begin{aligned}\tau(q_-, q_+) &= \lim_{n\to\infty} \tau(q_-^n, q_+^n) \le \lim_{n\to\infty} \overline{\tau}(\overline{q_-^n}, \overline{q_+^n}) = \lim_n L_\tau(\sigma_n^- \sigma_n^+|_{[-t,t]}) \\ &= L_\tau(\sigma|_{[-t,t]}),\end{aligned}$$

which is what we wanted to show, as the other inequality is trivial. □

4.2 Parallel lines

This subsection introduces the notion of parallelity for complete timelike lines. This approach is better suited for the synthetic case as it circumvents the analysis of Busemann functions. In the following, we call a map $f : Y_1 \to Y_2$ between Lorentzian pre-length spaces $(Y_1, d_1, \ll_1, \le_1, \tau_1)$ and $(Y_2, d_2, \ll_2, \le_2, \tau_2)$ *τ-preserving* if for all $p, q \in Y_1$, $\tau_1(p, q) = \tau_2(f(p), f(q))$, and we call f *causality preserving* if $p \le_1 q$ if and only if $f(p) \le_2 f(q)$.

Definition 4.13 (*Parallel lines*) Let α, β be two complete timelike lines in a Lorentzian pre-length space X defined on open intervals, w.l.o.g. on $\mathbb{R}$. They are called *parallel* if there exists a τ- and causality preserving map $f : (\alpha(\mathbb{R}) \cup \beta(\mathbb{R})) \to \mathbb{R}^{1,1}$ such that $f(\alpha(\mathbb{R}))$ and $f(\beta(\mathbb{R}))$ are parallel timelike lines in Minkowski (in the sense of parallel lines in affine spaces). We call such a map f a *parallel realisation* of α and β.

Remark 4.14 Note that by post-composing this by an isometry of Minkowski space, we can always achieve that $f(\alpha(\mathbb{R})) = \{(t, 0) : t \in \mathbb{R}\} \subseteq \mathbb{R}^{1,1}$ and $f(\beta(\mathbb{R})) = \{(t, c) : t \in \mathbb{R}\}$ for some $c \ge 0$. In this form, if α and β are parallel and both parametrized by τ-arclength (this is possible if τ is continuous and $\tau(x, x) = 0$ for all $x \in X$, cf. Lemma 2.29), we get that $f(\alpha(t)) = (t + a, 0)$ and $f(\beta(t)) = (t + b, c)$. By doing a shift in Minkowski, we can make a to be 0 (changing b to $b - a$).

Lemma 4.15 (Properties of parallel realisations) *Let X be a strongly causal Lorentzian pre-length space with continuous time separation, let $\alpha, \beta : \mathbb{R} \to X$ be two complete timelike lines and let $f : \alpha(\mathbb{R}) \cup \beta(\mathbb{R}) \to \mathbb{R}^{1,1}$ be a parallel realisation. Then f is a homeomorphism onto its image.*

Proof We first show that f is injective. Certainly, due to the fact that f preserves time separations, $f|_{\alpha(\mathbb{R})}$ and $f|_{\beta(\mathbb{R})}$ are injective: Indeed suppose that e.g. $f(\alpha(t)) = f(\alpha(s)) = (r, x)$ for some $s < t$, then $0 < \tau(\alpha(s), \alpha(t)) = \overline{\tau}((r, x), (r, x)) = 0$, a contradiction. Now suppose that the lines $f(\alpha(\mathbb{R}))$ and $f(\beta(\mathbb{R}))$ in $\mathbb{R}^{1,1}$ are different, then we are done. Otherwise, by nature of lines in Minkowski space, they have to be equal if they intersect. In that case, each point on that line is reached by one point on α and one point on β via f. So suppose that $(r, x) = f(\alpha(t)) = f(\beta(s))$. Since f is τ-preserving and τ is continuous, for each $\varepsilon > 0$ there are $\delta, \tilde{\delta} > 0$ such that $f(\beta(s + \delta)) = (r + \varepsilon, x)$ and $f(\beta(s - \tilde{\delta})) = (r - \varepsilon, x)$. But this means that $\alpha(t) \in \bigcap_{\delta,\tilde{\delta}\downarrow 0} I(\beta(s - \tilde{\delta}), \beta(s + \delta))$. By strong causality, these sets are a neighbourhood basis at $\beta(s)$, hence $\alpha(t) = \beta(s)$. This concludes the proof that f is injective.

Clearly, $\alpha(\mathbb{R}) \cup \beta(\mathbb{R})$ is a strongly causal Lorentzian pre-length space with the restriction of the structure of X, and similarly for their images in Minkowski space.

Moreover, f maps the timelike diamonds in $\alpha(\mathbb{R}) \cup \beta(\mathbb{R})$ to the timelike diamonds in $f(\alpha(\mathbb{R})) \cup f(\beta(\mathbb{R}))$. As these form topological bases due to strong causality, we conclude that f is a homeomorphism onto its image. □

Note that in the above setting, any parallel realisation f gives us a τ-arclength parametrisation of the parallel lines α, β: Let $\tilde{\alpha}$ be a τ-arclength parametrisation of the line $f(\alpha(\mathbb{R}))$ in Minkowski, then $f^{-1} \circ \tilde{\alpha}$ is τ-arclength parametrised, similarly for β.

Definition 4.16 (*Synchronised parallel lines*) Let X be a Lorentzian pre-length space with continuous time separation τ satisfying $\tau(x, x) = 0$ for all $x \in X$. Let $\alpha : \mathbb{R} \to X$, $\beta : \mathbb{R} \to X$ be two τ-arclength parametrised, parallel and complete timelike lines. They are called *synchronised parallel*[6] if the parallel realisation $f : (\alpha(\mathbb{R}) \cup \beta(\mathbb{R})) \to \mathbb{R}^{1,1}$ can be chosen to be of the form $f(\alpha(t)) = (t, 0)$ and $f(\beta(t)) = (t, c)$ for some $c \geq 0$. The constant c is a well-defined property of (α, β) and is called the *distance* of the parallel lines. For any two parallel lines α and β parametrised by τ-arclength, one can shift β such that $(\alpha, \beta \circ (t \mapsto t - a))$ is synchronised parallel.

Lemma 4.17 (The c-criterion for parallel lines) *Let X be a Lorentzian pre-length space with continuous time separation τ satisfying $\tau(x, x) = 0$ for all $x \in X$. Let $\alpha : \mathbb{R} \to X$ and $\beta : \mathbb{R} \to X$ be two τ-arclength parametrised complete timelike lines. We define the following partial functions:*

- $c_{\alpha\beta}(s, t) = \sqrt{(t - s)^2 - \tau(\alpha(s), \beta(t))^2} (\in \mathbb{C})$ *if* $\alpha(s) \leq \beta(t)$ *(otherwise undefined),*
- $c_{\beta\alpha}(s, t) = \sqrt{(s - t)^2 - \tau(\beta(t), \alpha(s))^2} (\in \mathbb{C})$ *if* $\beta(t) \leq \alpha(s)$ *(otherwise undefined),*
- $c^N_{\alpha+}(s) = \inf\{t - s : \alpha(s) \leq \beta(t)\}$,
- $c^N_{\beta+}(s) = \inf\{t - s : \beta(s) \leq \alpha(t)\}$.

These are all constant where defined, have the same value c and the infima are minima if and only if α and β are synchronised parallel. This constant c is the distance between α and β.

If X is additionally globally causally closed and $\alpha \cap I^+(\beta) \neq \emptyset$ and $\beta \cap I^+(\alpha) \neq \emptyset$, then the infima in $c^N_{\alpha+}$ and $c^N_{\beta+}$ are automatically minima and the conditions that $c^N_{\alpha+} = c^N_{\beta+} = c$ are automatically satisfied if $c_{\alpha\beta} = c_{\beta\alpha} = c$. In that case, all of the constants agree.

Proof We define f as in the definition of synchronised parallel with distance c: $f(\alpha(s)) := (s, 0)$, $f(\beta(s)) := (s, c)$. We will prove that f is causality preserving if and only if $c^N_{\alpha+}$ and $c^N_{\beta+}$ are constantly c, and under the assumption that this is the case f is τ-preserving if and only if $c_{\alpha\beta}$ and $c_{\beta\alpha}$ are constantly c wherever defined.

As α and β are future directed τ-arclength parametrised timelike lines, it is clear that f is causality and τ-preserving along α and along β, i.e. we only have to check the conditions on $f(\alpha(s)) \leq f(\beta(t))$, their τ-distance and the same for α, β reversed.

[6] This agrees with the usual notion of synchronising clocks in the sense that this is the parametrisation coming from synchronising the clocks on the particles α and β and parametrising α and β by their clock values.

So first, for causality preserving: We know that $f(\alpha(s)) \leq f(\beta(t)) \Leftrightarrow (s, 0) \leq (t, c) \Leftrightarrow (t - s) \geq c$. On the other hand, $\alpha(s) \leq \beta(t) \Leftrightarrow t - s \geq c^N_{\alpha+}(s)$ if this is a minimum and $\alpha(s) \leq \beta(t) \Leftrightarrow t - s > c^N_{\alpha+}(s)$ if it is only an infimum. In particular, these conditions are the same precisely when $c^N_{\alpha+}(s) = c$ and it is a minimum. Thus, $\alpha(s) \leq \beta(t) \Leftrightarrow f(\alpha(s)) \leq f(\beta(t))$ if and only if $c^N_{\alpha+}$ is constantly c and is always a minimum. Analogously, we get that $\beta(s) \leq \alpha(t) \Leftrightarrow f(\beta(s)) \leq f(\alpha(t))$ if and only if $c^N_{\beta+}$ is constantly c and is always a minimum.

We already proved that f is causality preserving. Next, we show that it preserves τ. On the Minkowski side, we have: $\bar{\tau}(f(\alpha(s)), f(\beta(t))) = \sqrt{(t - s)^2 - c^2}$ (if $f(\alpha(s)) \leq f(\beta(t))$). On the X side, we solve the defining equation for $c_{\alpha\beta}$ for $\tau(\alpha(s), \beta(t))$ to get: $\tau(\alpha(s), \beta(t)) = \sqrt{(t - s)^2 - c_{\alpha\beta}(s, t)^2}$ (if $\alpha(s) \leq \beta(t)$). Now note that as f is causality preserving, the side conditions when these equations can be applied match up (if they are not applied, $\tau(\alpha(s), \beta(t)) = 0 = \bar{\tau}(f(\alpha(s)), f(\beta(t)))$ anyway). If the side conditions are satisfied, we note that the equation differ only in one place: $\tau(\alpha(s), \beta(t))$ uses $c_{\alpha\beta}(s, t)$ and $\bar{\tau}(f(\alpha(s)), f(\beta(t)))$ uses c. Thus, we easily see that $\tau(\alpha(s), \beta(t)) = \bar{\tau}(f(\alpha(s)), f(\beta(t)))$ if and only if $c_{\alpha\beta}(s, t) = c$.

Analogously, we get that $\tau(\beta(t), \alpha(s)) = \bar{\tau}(f(\beta(t)), f(\alpha(s)))$ if and only if $c_{\beta\alpha}(s, t) = c$.

For the additional claim, fix s. If X satisfies the additional assumptions, $\{t : \alpha(s) \leq \beta(t)\}$ is closed and bounded (as is easily seen from the fact that $c_{\alpha\beta} \in \mathbb{R}$ when defined due to the reverse triangle inequality), thus it has a minimum automatically. We consider t such that $\alpha(s) \ll \beta(t)$, that is $t \in (t_s, +\infty) = \beta^{-1}(I^+(\alpha(s)))$. We get that $\tau(\alpha(s), \beta(t)) \to 0$ as $t \searrow t_s$. We transform the equation $c = \sqrt{(t - s)^2 - \tau(\alpha(s), \beta(t))^2}$ to $\tau(\alpha(s), \beta(t)) = \sqrt{(t - s)^2 - c^2}$. By continuity of τ, this still holds in the limit $t \searrow t_s$, i.e. $\sqrt{(t_s - s)^2 - c^2} = \tau(\alpha(s), \beta(t_s)) = 0$, so $t_s - s = c$ and $\{t : \alpha(s) \leq \beta(t)\} = [t_s, +\infty)$ as claimed. □

Lemma 4.18 (The strong causality trick) *Let X be a strongly causal Lorentzian pre-length space with continuous τ and $\tau(x, x) = 0$ for all $x \in X$. Let $\alpha, \beta : [0, b) \to \mathbb{R}$ be two τ-arclength parametrised timelike distance realisers with $x := \alpha(0) = \beta(0)$. Assume that for all s, t such that $\alpha(s)$ and $\beta(t)$ are timelike related, the comparison angle is $\tilde{\measuredangle}_x(\alpha(s), \beta(t)) = 0$. Then $\alpha = \beta$.*

Proof We set $f_+(s, t) = \tau(\alpha(s), \beta(t))$ and $f_-(s, t) = \tau(\beta(s), \alpha(t))$. They are both monotonically increasing in t and continuous. We want to describe the set where $f_+ > 0$ resp. $f_- > 0$. We set $t^+_s = \inf\{t : f_+(s, t) > 0\}$, then $\lim_{t \searrow t^+_s} f_+(s, t) = 0$. Thus in the law of cosines (Lemma 2.35), we get $\lim_{t \searrow t^+_s} \cosh(\tilde{\measuredangle}_x(\alpha(s), \beta(t))) = 1 = \lim_{t \searrow t^+_s} \frac{s^2 + t^2 - f_+(s,t)}{2st} = \frac{s^2 + (t^+_s)^2}{2st^+_s}$, so $s = t^+_s$. Analogously, we get $s = \inf\{t : f_-(s, t) > 0\}$, in total $f_\pm(s, t) > 0$ for $s < t$.

That means that whenever $s_- < t < s_+$ we have that $\alpha(t) \in I(\beta(s_-), \beta(s_+))$. By Lemma 2.26, the $I(\beta(s_-), \beta(s_+))$ form a neighbourhood basis of the point $\beta(t)$, and $\alpha(t)$ is inside of all of these neighbourhoods. As X is Hausdorff, we get that $\alpha(t) = \beta(t)$ for all $t \in (0, b)$, so $\alpha = \beta$. □

Lemma 4.19 (Parallel lines are unique) *Let X be a strongly causal, timelike geodesically connected Lorentzian pre-length space with continuous τ, $\tau(x,x) = 0$ for all $x \in X$. Suppose that X has global timelike curvature bounded below by 0, let $\alpha : \mathbb{R} \to X$ be a complete timelike line and $p \in X$ a point. Then there is (up to reparametrisation) at most one parallel line to α through p.*

Proof We indirectly assume there are two parallel lines to α through p, namely $\beta : \mathbb{R} \to X$ and $\tilde{\beta} : \mathbb{R} \to X$. We assume all three curves are parametrised by τ-arclength, such that $\beta(0) = \tilde{\beta}(0) = p$ and in such a way that α, β are synchronised parallel. Then also $\alpha, \tilde{\beta}$ are parametrised in a way that they are synchronised parallel, and both α, β and $\alpha, \tilde{\beta}$ have the same distance: set $s_+ = \inf \alpha^{-1}(I^+(p))$ and $s_- = \sup \alpha^{-1}(I^-(p))$ the boundary of the points on α timelike related to p. Then the distance of α, β and that of $\alpha, \tilde{\beta}$ can be calculated as $\frac{s_+ - s_-}{2}$, so they are the same, and the necessary shift to make β and $\tilde{\beta}$ synchronised parallel to α can be calculated by $\frac{s_+ + s_-}{2}$, so they are the same too. As α, β are assumed to be synchronised already, this shift is 0.

We construct a comparison situation for all three lines at once: We choose the parallel realisation f for α and β given by $f(\alpha(s)) = (s, 0)$ and $f(\beta(s)) = (s, c)$, and similarly we choose $\tilde{f}$ for α and $\tilde{\beta}$ given by $\tilde{f}(\tilde{\beta}(s)) = (s, -c)$ and $\tilde{f}(\alpha(s)) = (s, 0)$, i.e. we realise β and $\tilde{\beta}$ on opposite sides of α.

We calculate the angle $\omega := \measuredangle_p(\beta|_{[0,+\infty)}, \tilde{\beta}|_{[0,+\infty)})$: By Proposition 4.8, this is equal to any comparison angle $\tilde{\measuredangle}_p(\beta(s), \tilde{\beta}(t))$ as long as $\beta(s) \ll \tilde{\beta}(t)$ or conversely. We know that $\beta(s) \ll \alpha(s + c + \varepsilon) \ll \tilde{\beta}(s + 2c + 2\varepsilon)$ and set $t = s + 2c + 2\varepsilon$. We now look at the law of cosines to estimate the comparison angle $\omega = \tilde{\measuredangle}_p(\beta(s), \tilde{\beta}(t))$: $\cosh(\omega) = \frac{s^2 + t^2 - \tau(\beta(s), \tilde{\beta}(t))^2}{2st} \leq \frac{2t^2}{2s^2}$ which converges to 1 as $s \to +\infty$. In particular, $\omega = 0$. We can now apply the strong causality trick to get $\beta = \tilde{\beta}$ (so they are even equal in this parametrisation), so there is only one line parallel to α through p. $\square$

Let us now show for X as in the splitting theorem that asymptotic lines to γ constructed in Proposition 4.12 are parallel to γ.

Lemma 4.20 *Let $\alpha : \mathbb{R} \to X$ be a complete timelike asymptotic line to γ. Then α, γ are parallel.*

Proof By construction, α arises as the limit of timelike maximising segments α_n^+ from $p := \alpha(0)$ to $\gamma(r_n)$ and α_n^- from p to $\gamma(-r_n)$, where $r_n \to \infty$. We will use the stacking principle (cf. Proposition 4.6) to show that α and γ are parallel. Indeed, the triangles $\Delta(\gamma(-r_n), p, \gamma(r_n))$ stack in $\mathbb{R}^{1,1}$, hence by orienting the side corresponding to the segment on γ vertically (see Fig. 7), we may assume that for $n \to \infty$ the comparison triangles converge to two vertical lines in $\mathbb{R}^{1,1}$.

We now argue that the map sending $\alpha(\mathbb{R})$ and $\gamma(\mathbb{R})$ to the corresponding limit lines $\bar{\alpha}$ and $\bar{\gamma}$ is a parallel realisation. For any t, s, consider $\tau(\alpha(t), \gamma(s)) = \lim_n \tau(\alpha_n(t), \gamma(s))$. For n so large that $-r_n < s < r_n$, $\gamma(s)$ is part of the triangle $\Delta(\gamma(-r_n), p, \gamma(r_n))$. From Corollary 4.9 we conclude that $\alpha_n(t) \ll \gamma(s)$ if and only if $\overline{\alpha}_n(t) \ll \overline{\gamma}(s)$, and $\tau(\alpha_n(t), \gamma(s)) = \overline{\tau}(\overline{\alpha}_n(t), \overline{\gamma}(s))$, where the bars denote the corresponding points on the comparison side. This shows that the realisation of α as $\bar{\alpha}$ and γ as $\bar{\gamma}$ is a parallel realisation. $\square$

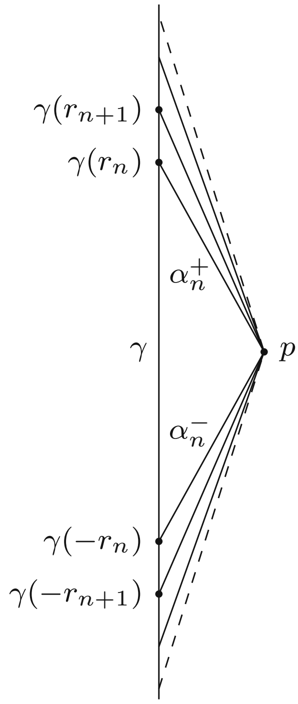


Fig. 7 Stacking comparison triangles

Lemma 4.21 (Shifting asymptotes) *Let $\alpha : \mathbb{R} \to X$ be a complete timelike asymptote to γ through $\alpha(0)$. Then the asymptote to γ through $\alpha(s)$ is $\alpha(\cdot - s)$ after parametrising both by τ-arclength.*

Proof Let $\tilde{\alpha}$ be the asymptotic line through $\alpha(s)$. Then the previous result shows that both γ and α as well as γ and $\tilde{\alpha}$ are parallel and they meet at $\alpha(s) = \tilde{\alpha}(0)$. Thus, as parallel lines are unique (see Lemma 4.19), we have that (after synchronising) $\alpha = \tilde{\alpha}$. To get the correct parameters before shifting, we know $\alpha(s) = \tilde{\alpha}(0)$, which fixes the shift parameter. □

Corollary 4.22 (Asymptotes stay in $I(\gamma)$) *Asymptotic lines to γ from points in $I(\gamma)$ stay in $I(\gamma)$.*

Proof Since asymptotic lines to γ are parallel to γ by Lemma 4.20, this readily follows. □

Now that we have established the parallelity of γ and its asymptotes, we may reparametrise all of them by τ-arclength and fix a shift on each asymptote so that they are synchronised to γ. Note that we fix a τ-arclength parametrisation of γ, namely the one keeping $\gamma(0)$ the same.

Definition 4.23 (*Busemann parametrisation*) Let $\alpha : \mathbb{R} \to X$ be a complete timelike asymptote to γ. Then we call the (unique) τ-arclength parametrisation of α that synchronises it to γ its *Busemann parametrisation*.

Remark 4.24 (Busemann functions) Up to this point, we have avoided the use of Busemann functions in our treatment. Though it will not really be necessary in what follows, let us discuss them briefly here. Let X be as in the splitting theorem and γ the given timelike line in τ-arclength parametrisation. Then we define the (*future*) *Busemann* *function* of γ as $b^+ : I(\gamma) \to \mathbb{R} \cup \{\pm\infty\}$ via $b^+(x) := \lim_{t\to\infty}(t - \tau(x, \gamma(t)))$. If $\alpha : \mathbb{R} \to X$ is any asymptotic line to γ in τ-arclength parametrisation and we choose the parallel realisation between α and γ so that $\gamma(t)$ is mapped to $(t, 0) \in \mathbb{R}^{1,1}$ (which we always do) and $\alpha(s)$ is mapped to $(s + S, c) \in \mathbb{R}^{1,1}$, then

$$b^+(\alpha(s)) = \lim_{t\to\infty}(t - \tau(\alpha(s), \gamma(t))) = \lim_{t\to\infty}(t - \sqrt{(t - (s + S))^2 - c^2}) = s + S.$$

Hence, b^+ indicates the shift in the time parameter along any τ-arclength parametrised asymptote. In particular, b^+ is finite-valued on all of $I(\gamma)$ since any point lies on an asymptote, and in fact it is not hard to show that b^+ is continuous on $I(\gamma)$ in a similar fashion as [11, Lem. 3.3], but we will not need this. Note that if α is in Busemann parametrisation, then $b^+(\alpha(s)) = s$ for all $s \in \mathbb{R}$, hence the name.

Lemma 4.25 (Parallelity is weakly transitive) *Let X be a strongly causal, timelike geodesically connected Lorentzian pre-length space with continuous τ, $\tau(x, x) = 0$ for all x, and global non-negative timelike curvature. Let $\alpha, \beta, \gamma : \mathbb{R} \to X$ be timelike lines such that α, β and β, γ are parallel. Assume that there is a point p on γ such that there is a parallel line to α through p. Then this parallel line is already γ.*

Proof We assume the parallel lines are even synchronised parallel in a suitable τ-arclength parametrisation. Let $\tilde{\gamma}$ be the synchronised parallel line to α through p. Let a be the distance between α, β, b be the distance between β, γ and c be the distance between $\alpha, \tilde{\gamma}$. We have two lines through p: γ and $\tilde{\gamma}$. We assume the parameters for p are $p = \gamma(t_0) = \tilde{\gamma}(\tilde{t}_0)$. By Proposition 4.8, we know that $\tilde{\measuredangle}_p(\gamma(s), \tilde{\gamma}(t))$ is constant in s and t as long as $\gamma(s) \ll \tilde{\gamma}(t)$ (or conversely). We know that $\gamma(s) \ll \beta(s+b+\varepsilon) \ll \alpha(s + a + b + 2\varepsilon) \ll \tilde{\gamma}(s + a + b + c + 3\varepsilon)$ and set $t = s + a + b + c + 3\varepsilon$. We now look at the law of cosines to estimate the comparison angle $\omega = \tilde{\measuredangle}_p(\gamma(s), \tilde{\gamma}(t))$: $\cosh(\omega) = \frac{(s-t_0)^2+(t-\tilde{t}_0)^2-\tau(\gamma(s),\tilde{\gamma}(t))}{2(s-t_0)(t-\tilde{t}_0)} \leq \frac{2(t-\tilde{t}_0)^2}{2(s-t_0)^2}$ which converges to 0 as $s \to +\infty$. In particular, $\omega = 0$. We can now apply the strong causality trick to get that γ and $\tilde{\gamma}$ are just shifts of each other, so α and γ are parallel.

By doing the same argument backwards as well, i.e. $\alpha(s - a - b - 2\varepsilon) \ll \gamma(s) \ll \alpha(s + a + b + 2\varepsilon)$, we get that the synchronised version only differs by a shift of at most $a + b$ (upon taking $\varepsilon \to 0$). □

Lemma 4.26 (Verticality in Minkowski space) *Let $a = (0, 0) \ll b$ be points in Minkowski space $\mathbb{R}^{1,1}$. Let the t-coordinate difference be $\Delta t = t(b) - t(a)$. Let c_n be points with $b \ll c_n$, the t-coordinate $t(c_n) \to +\infty$, c_n above the straight line through a, b and $\tau(a, c_n) - \tau(b, c_n) \to \Delta t$ as $n \to +\infty$. Then $\frac{x(c_n)}{t(c_n)} \to 0$, or equivalently, $\frac{c_n}{\|c_n\|} \to \partial_t$.*[7]

Proof We show that $\forall\varepsilon > 0$ there are constants $T > 0$ and $\delta > 0$ such that for all p with $\tau(a, p) \geq T$ and $|\tau(a, p) - \tau(b, p) - (t_b - t_a)| < \delta$, we have that the angle of ap with the t-axis satisfies $\measuredangle_a(\partial_t, p) < \varepsilon$.

We set $\omega_1 = \measuredangle_a(\partial_t, b)$ the angle between the vertical and ab, and $\omega_2 = \measuredangle_a(p, b)$ the angle between the vertical and ap. We will prove $|\omega_1 - \omega_2| < \varepsilon$. By assumption,

[7] So c_n converges to the point $[\partial_t]$ on the limit sphere.

p and ∂_t lie on the same side of the straight line through ab, so by angle additivity in the plane we also get that $\measuredangle_a(\partial_t, p) < \varepsilon$.

We apply the law of cosines:

$$\tau(b, p)^2 - \tau(a, p)^2 = \tau(a, b)^2 - 2\tau(a, b)\tau(a, p)\cosh(\omega_2)$$

and use that $\cosh(\omega_1)\tau(a, b) = t_b - t_a$ to get:

$$\tau(a, p) - \tau(b, p) = (t_b - t_a)\underbrace{\frac{(2\tau(a, p)\cosh(\omega_2) - 1)}{\cosh(\omega_1)(\tau(b, p) + \tau(a, p))}}_{=:F}.$$

We now have to check that whenever $|\omega_1 - \omega_2| \geq \varepsilon$, the factor F is bounded away from 1.

We have two cases to cover: First, we indirectly assume $\omega_1 > \omega_2 + \varepsilon$ (then $1 - \frac{\cosh(\omega_2)}{\cosh(\omega_1)} > 1 - \cosh(\varepsilon)$). (Assuming b lies to the right of the t-axis, this is the case where p also lies on the right.) We will show $F < 1 - \delta$ (if $\tau(a, p)$ is large enough and $\tau(a, p), \tau(b, p)$ are close enough). This is equivalent to:

$$(2\tau(a, p)\cosh(\omega_2) - 1) < (1 - \delta)\cosh(\omega_1)(\tau(b, p) + \tau(a, p))$$

Dividing by $\tau(a, p)$, this is obvious: $\frac{\tau(b,p)}{\tau(a,p)} \to 1$ as $\tau(a, p) \to +\infty$ and the difference is bounded. We can e.g. choose

$$\delta < \cosh(\varepsilon) - 1$$

and T large enough.

Now we indirectly assume $\omega_1 < \omega_2 - \varepsilon$ (then $\frac{\cosh(\omega_2)}{\cosh(\omega_1)} - 1 > \cosh(\varepsilon) - 1$) (Assuming b lies to the right of the t-axis, this is the case where p lies on the left.) We will show $F > 1 + \delta$ (if $\tau(a, p)$ is large enough and $\tau(a, p), \tau(b, p)$ are close enough). We need to check:

$$(2\tau(a, p)\cosh(\omega_2) - 1) > (1 + \delta)\cosh(\omega_1)(\tau(b, p) + \tau(a, p))$$

Dividing by $\tau(a, p)$, this is obvious: $\frac{\tau(b,p)}{\tau(a,p)} \to 1$ as $\tau(a, p) \to +\infty$ and the difference is bounded. We can e.g. choose

$$\delta < \cosh(\varepsilon) - 1$$

and T large enough. □

We now run into a subtlety: being parallel is in general only weakly transitive, but being synchronised parallel is probably not. Luckily, for X as in the splitting theorem, the fact that we consider asymptotic lines to γ simplifies the situation:

Lemma 4.27 (Two asymptotes are parallel) *Let $x, y \in I(\gamma)$ and α, β the asymptotes to γ through x resp. y in Busemann parametrisation. Then α, β are synchronised parallel.*

Proof We dive into the construction of asymptotes: We first assume $x \ll y$. We set $s_0 = b^+(x)$ and $t_0 = b^+(y)$, then there are maximisers α_n from x to $\gamma(T_n)$ parametrised in τ-arclength such that $\alpha_n(s_0) = x$ and β_n from y to $\gamma(T_n)$ parametrised in τ-arclength such that $\beta_n(t_0) = y$, which converge (pointwise) to the upper parts of α and β in Busemann parametrisation for $T_n \to \infty$, respectively. Note that $\alpha_n(s_0 + \tau(x, \gamma(T_n))) = \beta_n(t_0 + \tau(y, \gamma(T_n))) = \gamma(T_n)$.

We try to prove the c-criterion (Lemma 4.17): Clearly, due to our assumptions on X, we do not need to consider the null functions. We look at $c_{\alpha\beta}(s_0, t_0)$ and $c_{\alpha\beta}(s', t')$ or $c_{\beta\alpha}(s', t')$ for $s_0 \leq s'$, $t_0 \leq t'$: We require $a := x = \alpha(s_0) \ll b := y = \beta(t_0)$. We get points converging to these parameters: $a'_n := \alpha_n(s') \to a' := \alpha(s')$ and $b'_n := \beta_n(t') \to b' := \beta(t')$ (note that $\alpha_n(s_0) = a$ and $\beta_n(t_0) = b$ anyway). We get a timelike triangle $\Delta_n = \Delta(a, b, c_n := \gamma(T_n))$ containing the points (a'_n, b'_n). We form a comparison situation for this in Minkowski space $\mathbb{R}^{1,1}$: $\bar{\Delta}_n = \Delta(\bar{a}, \bar{b}, \bar{c}_n)$ with comparison points $\bar{a}'_n, \bar{b}'_n$. As X has global curvature bounded below by 0, we get that $\tau(a'_n, b'_n) \leq \bar{\tau}(\bar{a}'_n, \bar{b}'_n)$ and $\tau(b'_n, a'_n) \leq \bar{\tau}(\bar{b}'_n, \bar{a}'_n)$.

We would now like to let $T_n \to +\infty$, so we need control over what the comparison triangle $\bar{\Delta}_n$ converges to. For this, we select which way to realise $\bar{\Delta}_n$ in $\mathbb{R}^{1,1}$. We choose:

- $\bar{a} = (s_0, 0)$ constant in n, i.e. at the right t-coordinate and with x-coordinate 0,
- $\bar{b} = (t_0, c_{\alpha\beta}(s_0, t_0))$ constant in n. Note this has the right τ-distance to $\bar{a}$ by definition of $c_{\alpha\beta}$.
- $\bar{c}_n$ above the straight line $\bar{a}$ and $\bar{b}$ lie on.

We claim that in the limit $T_n \to +\infty$, the line $\bar{a}\bar{c}_n$ becomes vertical (the line $\bar{b}\bar{c}_n$ also becomes vertical). For this, we look at what $\bar{\tau}(\bar{a}, \bar{c}_n) - \bar{\tau}(\bar{b}, \bar{c}_n)$ does as $T_n \to +\infty$: Using Landau small-oh notation, we get that $\bar{\tau}(\bar{a}, \bar{c}_n) = T_n - b^+(a) + o(1)$ and $\bar{\tau}(\bar{b}, \bar{c}_n) = T_n - b^+(b) + o(1)$ as $T_n \to +\infty$, so the difference is $b^+(b) - b^+(a) + o(1)$, which approaches the t-coordinate-difference $t_0 - s_0$ of $\bar{a}$ and $\bar{b}$. We can now apply Lemma 4.26 to get that $\frac{x(c_n)}{t(c_n)} \to 0$ and $t(c_n) \to +\infty$, thus these lines become vertical.

We have proven that the comparison points converge, so we have limit comparison points: $\bar{a} = (s_0, 0)$ and $\bar{b} = (t_0, c_{\alpha\beta}(s_0, t_0))$ stay fixed anyway, $\bar{a}'_n \to \bar{a}' = (s', 0)$, $\bar{b}'_n \to \bar{b}' = (t', c_{\alpha\beta}(s_0, t_0))$. We also set the comparison sides $\bar{\alpha}(s) = (s, 0)$ and $\bar{\beta}(t) = (t, c_{\alpha\beta}(s_0, t_0))$. By continuity of τ and $\bar{\tau}$, our curvature assumption gives $\tau(a', b') \leq \bar{\tau}(\bar{a}', \bar{b}')$ in the limit (similarly with arguments flipped). Now we compare the definition of $c_{\alpha\beta}(s', t')$ resp. $c_{\beta\alpha}(s', t')$ with the c-functions for $\bar{\alpha}$ and $\bar{\beta}$ at the same parameters:

$$c_{\alpha\beta}(s', t') = \sqrt{(t' - s')^2 - \tau(a', b')^2},$$

$$c_{\bar{\alpha}\bar{\beta}}(s', t') = \sqrt{(t' - s')^2 - \bar{\tau}(\bar{a}', \bar{b}')^2} = x(\bar{b}') - x(\bar{a}') = c_{\alpha\beta}(s_0, t_0),$$

whenever defined. The last equality holds because we know $\bar{\alpha}$ and $\bar{\beta}$ are synchronised parallel with distance $c_{\alpha\beta}(s_0, t_0)$. Note that $a' \leq b'$ implies $\bar{a}' \leq \bar{b}'$ by the curvature bound and a simple continuity argument, so whenever the first line is defined so is the second. Notice these equations only differ in the τ term, and we know $\tau(a', b') \leq \bar{\tau}(\bar{a}', \bar{b}')$, so we get $c_{\alpha\beta}(s', t') \geq c_{\alpha\beta}(s_0, t_0)$ whenever the former is defined. Similarly, if $b' \leq a'$ we get $c_{\beta\alpha}(s', t') \geq c_{\alpha\beta}(s_0, t_0)$ for all $s' \geq s_0$ and $t' \geq t_0$. We can also do this if $a \gg b$, giving $c_{\alpha\beta}(s', t') \geq c_{\beta\alpha}(s_0, t_0)$ and $c_{\beta\alpha}(s', t') \geq c_{\beta\alpha}(s_0, t_0)$.

Doing the same construction towards the past, we similarly get $c_{\alpha\beta}(s', t') \geq c_{\alpha\beta}(s_0, t_0)$ and $c_{\beta\alpha}(s', t') \geq c_{\alpha\beta}(s_0, t_0)$ resp. $c_{\alpha\beta}(s', t') \geq c_{\beta\alpha}(s_0, t_0)$ and $c_{\beta\alpha}(s', t') \geq c_{\beta\alpha}(s_0, t_0)$ (depending on whether $a \ll b$ or $b \ll a$) for all $s' \leq s_0$ and $t' \leq t_0$.

Now we use Lemma 4.21 to get that the (Busemann parametrised) asymptote to γ through $\alpha(s)$ is α, and similarly the asymptote to γ through $\beta(t)$ is β. In particular, we can use the above argument again for $\alpha(s)$ instead of x and $\beta(t)$ instead of y. We know that $b^+(\alpha(s)) = s$ and $b^+(\beta(t)) = t$. The above (to the future, $a \ll b$) then gives: $c_{\alpha\beta}(s', t') \geq c_{\alpha\beta}(s, t)$ for $s \leq s', t \leq t'$ such that $a \ll b$ and $a' \ll b'$, extending continuously, $c_{\alpha\beta}$ is monotonously increasing where defined. On the other hand, the above (to the past, $a \ll b$) gives: $c_{\alpha\beta}(s', t') \geq c_{\alpha\beta}(s, t)$ for $s \geq s', t \geq t'$ such that $a \ll b$ and $a' \ll b'$, extending continuously, $c_{\alpha\beta}$ is monotonously decreasing where defined. Via a two-step process, we see that $c_{\alpha\beta}$ is constant where defined. Similarly, we get that also $c_{\beta\alpha}$ is constant where defined, and that they have the same value.

Thus, the c-criterion (Lemma 4.17) yields that α and β are synchronised parallel. $\square$

This result establishes the transitivity of synchronised parallel lines: Any two asymptotes to γ in Busemann parametrisation are synchronised parallel.

5 Proof of the main result

Let us summarise what we have shown so far: Let $(X, d, \ll, \leq, \tau)$ be a connected, regularly localisable, globally hyperbolic Lorentzian length space satisfying timelike geodesic prolongation, with proper metric d and global non-negative timelike curvature containing a complete timelike line $\gamma : \mathbb{R} \to X$. Then from each point in $I(\gamma) = I^+(\gamma) \cap I^-(\gamma)$ we can construct (unique) asymptotic rays to γ, all of which are timelike with infinite τ-length. Future directed and past directed rays from a common point fit together to give a timelike line which is parallel to γ. We fix a τ-arclength parametrisation of γ fixing $\gamma(0)$, and always consider it to be mapped to the t-axis in $\mathbb{R}^{1,1}$ in any parallel realisation. Then, by way of the Busemann parametrisation, all asymptotic lines become synchronised parallel to γ and to each other. This synchronisation selects a “spacelike” slice (namely $\{b^+ = 0\}$) in X which will provide the metric part of the splitting.

In this section, we complete the proof of the splitting theorem in two steps: First, we show that $I(\gamma)$ splits, then we show that $I(\gamma)$ has the (TC) property, which implies $I(\gamma) = X$ due to Theorem 2.47.

Definition 5.1 (*Spacelike slice*) We call the set of all $\alpha(0)$, where α is a Busemann parametrised timelike asymptotic line to γ, the *spacelike slice* and denote it by S. For $\alpha(0), \beta(0) \in S$, we define $d_S(\alpha(0), \beta(0))$ to be the distance between α and β in the sense of parallel lines.

Lemma 5.2 (S, d_S) *is a metric space.*

Proof As asymptotes to γ are parallel to γ and parallel lines are unique, d_S is well-defined.

Let $p, q \in S$. It is obvious from the definition of the distance of two parallel lines that $d_S(p, q) \geq 0$. If $d_S(p, q) = 0$, we get that the asymptotes through p and q are the same curve α. As the Busemann parametrisation is unique, we get that $p = q$.

For the triangle inequality, let $p, q, r \in S$. We get asymptotes α through p, β through q and γ through r, and by Lemma 4.27, they are pairwise synchronised parallel. Let $d_1 = d_S(p, q)$ and $d_2 = d_S(q, r)$, then $\alpha(0) \leq \beta(d_1)$ by the $c^N_{\alpha+}(0)$ criterion (see Lemma 4.17) for α, β and $\beta(d_1) \leq \gamma(d_1 + d_2)$ by the $c^N_{\beta+}(d_1)$ criterion for β, γ. Thus, we get that $\alpha(0) \leq \gamma(d_1 + d_2)$, giving $d_S(p, r) \leq d_1 + d_2$ by the $c^N_{\alpha+}(0)$ criterion for α, γ. □

Definition 5.3 (*The splitting map*) We define $f : \mathbb{R} \times S \to I(\gamma) \subset X$ by $f(s, p) = \alpha_p(s)$, where α_p is the asymptote to γ through p in Busemann parametrisation, that is $\alpha_p(0) = p$.

Proposition 5.4 (Local splitting) *Let X be a connected, regularly localisable, globally hyperbolic Lorentzian length space with proper metric d and global non-negative timelike curvature satisfying timelike geodesic prolongation and containing a complete timelike line $\gamma : \mathbb{R} \to X$. Then $I(\gamma) \subset X$ is a causally convex open set that is itself a path-connected, regularly localisable, globally hyperbolic Lorentzian length space of global non-negative timelike curvature with the metric, relations and time separation induced from X. Moreover, the spacelike slice S is a proper (hence complete), strictly intrinsic metric space, the Lorentzian product $\mathbb{R} \times S$ is a path-connected, regularly localisable, globally hyperbolic Lorentzian length space and the splitting map $f : \mathbb{R} \times S \to I(\gamma)$ is a τ- and causality preserving homeomorphism.*

Proof First, it is clear that $I(\gamma)$ is path-connected, causally convex in X and has global non-negative timelike curvature. It is hence causally path-connected since X is and it is trivially locally causally closed. Moreover, if $x \in I(\gamma)$ and U is a regular localising neighbourhood of x in X, then $U \cap I(\gamma)$ is a regular localising neighbourhood of x in $I(\gamma)$, hence $I(\gamma)$ is regularly localisable. By causal convexity, the time separation between causally related points in $I(\gamma)$ is achieved as the supremum of lengths of causal curves running between them which have to stay inside $I(\gamma)$. The causal diamonds in $I(\gamma)$ are precisely those in X, since they must be contained in $I(\gamma)$. Finally, $I(\gamma)$ is non-totally imprisoning (it inherits this from X), thus we have shown all the claims on $I(\gamma)$.

Next, we argue that f is τ- and causality preserving: Let $p, q \in S$. Let α be the asymptote to γ through p and let β be the asymptote to γ through q. Then α and β are synchronised parallel with distance $d_S(p, q)$, so there is a parallel realisation $\tilde{f} : \alpha(\mathbb{R}) \cup \beta(\mathbb{R}) \to \mathbb{R}^{1,1}$ defined by $\tilde{f}(\alpha(s)) = (s, 0)$ and $\tilde{f}(\beta(t)) = (t, d_S(p, q))$. In particular, $\alpha(s) \leq_X \beta(t) \Leftrightarrow t - s \geq d_S(p, q)$ and if this is true, $\tau(\alpha(s), \beta(t)) = \sqrt{(t - s)^2 - d_S(p, q)^2}$. But that is just the definition of $(s, p) \leq_{\mathbb{R}\times S} (t, q)$ resp. $\tau_{\mathbb{R}\times S}((s, p), (t, q))$ in $\mathbb{R} \times S$. This is true for all $p, q \in S$ and $s, t \in \mathbb{R}$, so f is τ- and causality preserving. f is injective as asymptotes to γ are parallel to γ and parallel lines are unique, and surjective since any point in $I(\gamma)$ lies on an asymptote.

From the discussion above, it is easy to see that f maps timelike (and causal) diamonds in $I(\gamma)$ to timelike (and causal) diamonds in $\mathbb{R}\times S$. As both sides are strongly causal, this implies that f is a continuous open bijection, hence a homeomorphism. Since $\mathbb{R} \times S$ is always non-totally imprisoning (cf. Proposition 3.4) and its causal diamonds are compact (as continuous images of compact causal diamonds in $I(\gamma)$), we conclude that $\mathbb{R}\times S$ is globally hyperbolic, hence S is proper by Proposition 3.7. To see that S is a strictly intrinsic space, fix $p, q \in S$ and connect any two timelike related points on the corresponding asymptotes by a distance realiser in $I(\gamma)$. The image of that distance realiser under f is a continuous distance realiser in $\mathbb{R} \times S$, hence by Proposition 3.5 the projection onto S gives a distance minimiser in S between p and q.

Finally, we need to prove the remaining claimed properties of $\mathbb{R} \times S$. Path-connectedness is inherited from $I(\gamma)$ via f, and products are always globally causally closed (cf. Proposition 3.2). Note that $I(x, y)$ for $x, y \in I(\gamma)$ are regular localising neighbourhoods in $I(\gamma)$. Since $f(I(x, y)) = I(f(x), f(y))$ and the d-lengths of causal curves in timelike diamonds can always be uniformly bounded in products (cf. the proof of Proposition 3.4), timelike diamonds in $\mathbb{R}\times S$ are in fact (regular) localising neighbourhoods: For the local time separation, take the restriction of $\tau_{\mathbb{R}\times S}$, and note that maximisers in $I(\gamma)$ (or in any $I(x, y) \subset I(\gamma)$) map to continuous maximisers in $\mathbb{R} \times S$, which are always Lipschitz reparametrisable and are hence causal curves (cf. Proposition 3.5). □

Proposition 5.5 (Global splitting) *Let $(X, d, \ll, \leq, \tau)$ satisfy the assumptions in Proposition 5.4. Then $I(\gamma) = X$, i.e. the splitting in Proposition 5.4 is global.*

Proof Suppose $\alpha : [a, b) \to I(\gamma)$ is a future directed timelike geodesic with finite τ-length. We reparametrise α by τ-arclength and denote the reparametrised curve again by α, it is then defined on $[0, L)$, where $L := L_\tau(\alpha) < \infty$. Then $f^{-1} \circ \alpha$ is a continuous timelike geodesic in $\mathbb{R} \times S$, where f is the splitting map. We decompose it into time- and space-component: $f^{-1} \circ \alpha = (\alpha^t, \alpha^x)$. If α^x is constantly p, we see that $\alpha = \alpha_p$, a contradiction since asymptotes have infinite τ-length. Note that since the τ-lengths of $f^{-1} \circ \alpha$ and α agree, $f^{-1} \circ \alpha$ also has finite τ-length in $\mathbb{R} \times S$. By Corollary 3.6, α^x has constant speed and domain $[0, L)$, thus finite d_S-length.

α^t can certainly be extended to L. As S is a complete metric space, we can extend α^x continuously to L. Mapping this extension back to $I(\gamma)$ via f, we get a continuous extension in $I(\gamma)$ of α to b (in its original parametrisation). This shows that $I(\gamma)$ has the (TC) property, thus $I(\gamma) = X$ by Theorem 2.47. This concludes the proof of the splitting in Theorem 1.4. □

Recall the notion of Cauchy sets and (Cauchy) time functions on Lorentzian pre-length spaces from [19, Sec. 5.1]: A Cauchy set is any subset that is met exactly once by doubly inextendible causal curves, and a Cauchy time function is a continuous function $t : X \to \mathbb{R}$ such that $x < y$ implies $t(x) < t(y)$ and the image under t of any doubly inextendible causal curve is all of $\mathbb{R}$.

Corollary 5.6 *Let $(X, d, \ll, \leq, \tau)$ satisfy the assumptions of Theorem 1.4 and let $f : \mathbb{R} \times S \to X$ be the splitting. Then the sets $S_t := f(\{t\} \times S)$ are Cauchy sets in X that* *are all homeomorphic to S. Moreover, the map $pr_1 \circ f^{-1}$ is a Cauchy time function. Moreover, all Cauchy sets in X are homeomorphic to S.*

Proof Let $\alpha : (a, b) \to X$ be any doubly inextendible causal curve in X meeting any S_t twice, then its image under f^{-1} meets $\{t\} \times S$ twice, which cannot happen since $\{t\} \times S$ is acausal in $\mathbb{R} \times S$. Next we argue that any such α meets S_t: For simplicity let $t = 0$, and suppose that α does not meet S, so w.l.o.g. we may assume that $\alpha \subset I^+(S)$ (this is because $X = I^-(S) \cup S \cup I^+(S)$ due to the splitting, and this union is disjoint). Let $t_0 \in (a, b)$, then $\alpha((a, t_0]) \subset J^-(\alpha(t_0)) \cap J^+(S)$ which is easily seen to be a compact set by considering the corresponding situation in $\mathbb{R} \times S$. But this is a contradiction, since X is non-totally imprisoning. Hence S (and any S_t) is a Cauchy set. They are homeomorphic to S since X has the topology of $\mathbb{R} \times S$.

We now argue that $pr_1 \circ f^{-1}$ is a Cauchy time function for X. It is clearly a time function, as it is continuous and the time separation on $\mathbb{R} \times S$ is more or less a reformulation of that. Now let $\alpha : (a, b) \to X$ be a doubly inextendible causal curve, we need to show that $pr_1 \circ f^{-1} \circ \alpha((a, b)) = \mathbb{R}$. Suppose not, so there is some time value t_0 that is not attained. W.l.o.g. suppose $t_0 \geq 0$, so the image is contained in $(-\infty, t_0]$. Similarly to before, this would imply that $\alpha|_{[t_0,b)}$ is contained in the compact set $J^+(\alpha(t_0)) \cap J^-(S_{t_0})$, a contradiction.

Finally, let C be any Cauchy set in X. Then the projection $C \to S$ (via the splitting) is continuous. Its inverse is given by sending each $p \in S$ to the unique point on α_p meeting C. This is a continuous map since C is achronal. This shows that C is homeomorphic to S. □

Note that in general, Cauchy sets in globally hyperbolic Lorentzian pre-length spaces need not be homeomorphic, as [19, Ex. 5.7, Ex. 5.8] show.

Corollary 5.7 *Let $(X, d, \ll, \leq, \tau)$ satisfy the assumptions of Theorem 1.4. Then (S, d_S) has Alexandrov curvature ≥ 0.*

Proof Due to the splitting, the space $\mathbb{R} \times S$ has global nonnegative timelike curvature (cf. Remark 2.33; the splitting maps Lipschitz triangles in $\mathbb{R} \times S$ to continuous triangles in X, but this does not lead to any issues), and then it follows from [5, Thm. 5.7] that S has Alexandrov curvature ≥ 0. □

6 Applications and outlook

Let us now give several reformulations of our splitting result for spacetimes. We start with smooth spacetimes, where we get a weaker version of the smooth splitting result proven in [11].

Corollary 6.1 (Smooth spacetimes) *Let (M, g) be a smooth, globally hyperbolic spacetime with non-positive timelike sectional curvature containing a complete timelike line $\gamma : \mathbb{R} \to M$. Then the Lorentzian length space induced by (M, g) splits in the sense of Theorem 1.4.*

Proof Fix any complete Riemannian metric h on M and let d_h be the induced distance, then $(M, d_h, \ll, \leq, \tau)$ is a connected, regularly localisable (via convex neighbourhoods) Lorentzian length space and d_h is a proper metric (note that the τ-length of causal curves agrees with the usual notion of length in (M, g) due to [48, Prop. 2.32]). The fact that (M, g) has global non-negative timelike curvature in the sense of Definition 2.32 is a consequence of the Lorentzian Toponogov theorem (see [43, p. 303]). The remaining technical assumptions are easily seen to hold. Hence the Lorentzian length space induced by (M, g) can be split according to Theorem 1.4. □

Remark 6.2 In the setting of Corollary 6.1, once one has shown that the spacelike slice S is a spacelike hypersurface in M and the normal asymptotic distance d_S is the Riemannian distance on S (e.g. via an analysis of the gradient of the Busemann function as in [11]), the Hawking–King–McCarthy theorem [44], which states that any τ-preserving homeomorphism between strongly causal spacetimes is a smooth isometry, provides a somewhat alternative proof of the smooth splitting result proven in [11].

Next, we turn to $C^{1,1}$-spacetimes. First note that the Lorentzian pre-length space induced by any strongly causal Lipschitz spacetime is in fact a Lorentzian length space by [48, Thm. 5.12]. For $C^{1,1}$-metrics, the notion of sectional curvature is not available any more, so we have to assume synthetic curvature bounds. However, most of the other technical assumptions are easily seen to be satisfied due to the existence of convex neighbourhoods. For details on low regularity spacetimes, see e.g. the recent review article [59] and the references therein.

Corollary 6.3 ($C^{1,1}$-spacetimes) *Let (M, g) be a globally hyperbolic $C^{1,1}$-spacetime with global non-negative timelike curvature containing a complete timelike line $\gamma : \mathbb{R} \to M$. Then the Lorentzian length space induced by (M, g) splits in the sense of Theorem 1.4.*

Finally let us mention the case of C^1-spacetimes, which are substantially different. Notably, the Lorentzian length space induced by a globally hyperbolic C^1-spacetime need not satisfy the timelike geodesic prolongation property. This is because solutions of the g-geodesic equation need not maximise locally, nor does a maximising segment necessarily extend as a local maximiser beyond its endpoints. However, points in C^1-spacetimes still have neighbourhood bases consisting of globally hyperbolic, causally convex neighbourhoods, so regular localisability is satisfied. Assuming the necessary synthetic properties as well, we have the following (note that geodesics in the synthetic sense are locally maximising curves, which are necessarily C^2-solutions of the geodesic equation even if the metric is just C^1):

Corollary 6.4 (C^1-spacetimes) *Let (M, g) be a globally hyperbolic C^1-spacetime with global non-negative timelike curvature containing a complete timelike line $\gamma : \mathbb{R} \to M$. Suppose that each maximising timelike geodesic segment extends to a locally maximising timelike geodesic on an open domain. Then the Lorentzian length space induced by (M, g) splits in the sense of Theorem 1.4.*

In this work, we have established a synthetic splitting result for Lorentzian length spaces using triangle comparison. Ideas for future projects include the generalisation to Lorentzian length spaces lower on the causal ladder (e.g. (TC) spaces instead of globally hyperbolic ones, an approach to this that seems to be adaptable to the synthetic situation is [34]), an analysis of the aforementioned splittings of low regularity spacetimes (regarding regularity of the splitting, level set, etc.), as well as an investigation into synthetic timelike Ricci curvature bounds introduced in [23] and expanded in [15]. In the case of low regularity spacetimes, it is not yet clear how these different notions of curvature bounds relate to each other (some recent progress in this direction includes [16, 47]). For metric measure spaces, it is known that the CD-condition is not sufficient and the full RCD-condition is needed for a splitting result (see [37, 38]). An analogue of the RCD-condition is currently unavailable for Lorentzian length spaces and will probably have to be developed before a splitting theorem with synthetic timelike Ricci curvature bounds can be considered.

Acknowledgements Argam Ohanyan and Felix Rott were supported by project P 33594 of the Austrian Science Fund FWF. Argam Ohanyan was also supported by the ÖAW-DOC scholarship of the Austrian Academy of Sciences. Didier Solis acknowledges the support of Conacyt under grant SNI 38368 and UADY under program FMAT-PTA2022, as well as the hospitality of the University of Vienna, where parts of this work were conducted. Tobias Beran acknowledges the support of University of Vienna. The authors would like to thank the Erwin Schrödinger Institute for support and hospitality during the programme *Non-regular Spacetime Geometry* where part of this work was undertaken. The authors would like to thank Matteo Calisti, Gregory Galloway, Melanie Graf, Michael Kunzinger, Clemens Sämann, Benedict Schinnerl and Roland Steinbauer for helpful discussions and valuable feedback. The authors would like to acknowledge the thorough work of the reviewer, whose suggestions helped improving this article.

Availability of data and materials No data was collected nor produced in this work.

Declarations

Conflict of interest/Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Ake Hau, L., Burgos, S., Solis, D.A.: Causal completions as Lorentzian pre-length spaces. Gen. Relativ. Gravit. **54**(9), 1–20 (2022)
2. Ake Hau, L., Cabrera Pacheco, A.J., Solis, D.A.: On the causal hierarchy of Lorentzian length spaces. Class. Quantum Gravity **37**(21), 215013 (2020)
3. Alexander, S., Bishop, R.: A cone splitting theorem for Alexandrov spaces. Pac. J. Math. **218**, 1–15 (2005)
4. Alexander, S.B., Bishop, R.L.: Lorentz and semi-Riemannian spaces with Alexandrov curvature bounds. Commun. Anal. Geom. **16**(2), 251–282 (2008)
5. Alexander, S. B., Graf, M., Kunzinger, M., Sämann, C.: Generalized cones as Lorentzian length spaces: Causality, curvature, and singularity theorems. arXiv:1909.09575 (2019)
6. Ambrosio, L., Gigli, N., Savare, G.: Gradient flows in metric spaces and in the space of probability measures. Birkhauser (2008)
7. Barrera, W., Montes de Oca, L., Solis, D. A.: Comparison theorems for Lorentzian length spaces with lower timelike curvature bounds. arXiv:2204.09612 (2022)
8. Bartnik, R.: Remarks on cosmological spacetimes and constant mean curvature surfaces. Commun. Math. Phys. **117**(4), 615–624 (1988)
9. Beem, J., Ehrlich, P., Easley, K.: Global Lorentzian Geometry. Marcel Dekker (1996)

10. Beem, J., Ehrlich, P., Markvorsen, S., Galloway, G.: Decomposition theorems for Lorentzian manifolds with nonpositive curvature. J. Differ. Geom. **22**, 29–42 (1985)
11. Beem, J., Ehrlich, P., Markvorsen, S., Galloway, G.: A Toponogov splitting theorem for Lorentzian manifolds. In: Global Differential Geometry and Global Analysis 1984, pp. 1–13. Springer (1985)
12. Beran, T., Rott, F.: Gluing constructions for Lorentzian length spaces. arXiv:2201.09695 (2022)
13. Beran, T., Sämann, C.: Hyperbolic angles in Lorentzian length spaces and timelike curvature bounds. J. Lond. Math. Soc. **2**, 1–58 (2023)
14. Borzellino, J., Zhu, S.: The splitting theorem for orbifolds. Ill. J. Math. **38**, 679–691 (1994)
15. Braun, M.: Rényi's entropy on Lorentzian spaces. Timelike curvature-dimension conditions. arXiv:2206.13005 (2022)
16. Braun, M., Calisti, M.: Timelike Ricci bounds for low regularity spacetimes by optimal transport. arXiv:2209.03802 (2022)
17. Burago, D., Burago, Y., Ivanov, S.: A Course in Metric Geometry, vol. 33. American Mathematical Society, Providence, RI (2001)
18. Burago, Y., Gromov, M., Perelman, G.: AD Alexandrov spaces with curvatures bounded below I. Russian Math. Surv. **47**, 1–58 (1992)
19. Burtscher, A., García-Heveling, L.: Time functions on Lorentzian length spaces. arXiv:2108.02693 (2021)
20. Cai, M.: Ends of Riemannian manifolds with nonnegative Ricci curvature outside of a compact set. Bull. Am. Math. Soc. **24**, 371–377 (1991)
21. Cai, M.: A splitting theorem for manifolds of almost nonnegative Ricci curvature. Ann. Glob. Anal. Geom. **11**, 373–385 (1993)
22. Cai, M., Galloway, G., Liu, Z.: Local splitting theorems for Riemannian manifolds. Proc. Am. Math. Soc. **120**, 1231–1239 (1994)
23. Cavalletti, F., Mondino, A.: Optimal transport in Lorentzian synthetic spaces, synthetic timelike Ricci curvature lower bounds and applications. arXiv:2004.08934 (2020)
24. Cheeger, J., Gromoll, D.: The splitting theorem for manifolds of nonnegative Ricci curvature. J. Differ. Geom. **6**(1), 119–128 (1971)
25. Cheeger, J., Gromoll, D.: On the structure of complete manifolds of nonnegative curvature. Ann. Math. **96**, 413–443 (1972)
26. Chruściel, P.T., Grant, J.D.E.: On Lorentzian causality with continuous metrics. Class. Quantum Gravity **29**(14), 145001 (2012)
27. Cohn-Vossen, S.: Totalkrümmung und geodätische Linien auf einfach zusammenhängenden, offenen, vollstandingen Flächenstucken. Math. Sb. **43**, 139–163 (1936)
28. Eschenburg, J.-H.: The splitting theorem for space–times with strong energy condition. J. Differ. Geom. **27**(3), 477–491 (1988)
29. Fang, F., Li, X., Zhang, Z.: Two generalizations of Cheeger–Gromoll splitting theorem via Bakry–Emery Ricci curvature. Ann. Inst. Fourier **59**, 563–573 (2009)
30. Galloway, G.: Splitting theorems for spatially closed space–times. Commun. Math. Phys. **96**(4), 423–429 (1984)
31. Galloway, G.: A generalization of the Cheeger–Gromoll splitting theorem. Arch. Math. **47**, 372–375 (1986)
32. Galloway, G.: The Lorentzian splitting theorem without the completeness assumption. J. Differ. Geom. **29**, 373–387 (1989)
33. Galloway, G.: Existence of CMC Cauchy surfaces and spacetime splitting. Pure Appl. Math. Q. **15**, 667–682 (2019)
34. Galloway, G., Horta, A.: Regularity of Lorentzian Busemann functions. Trans. Am. Math. Soc. **348**(5), 2063–2084 (1996)
35. Galloway, G., Ling, E., Sbierski, J.: Timelike completeness as an obstruction to C^0-extensions. Commun. Math. Phys. **359**(3), 937–949 (2018)
36. García-Heveling, L.: Causality theory of spacetimes with continuous Lorentzian metrics revisited. Class. Quantum Gravity **38**(14), 145028 (2021)
37. Gigli, N.: The splitting theorem in non-smooth context. arXiv:1302.5555 (2013)
38. Gigli, N.: An overview of the proof of the splitting theorem in spaces with non-negative Ricci curvature. Anal. Geom. Metric Spaces **2**(1) (2014)
39. Gigli, N., Pasqualetto, E.: Lectures on Non-smooth Differential geometry. Springer (2018)

40. Grant, J.D.E., Kunzinger, M., Sämann, C.: Inextendibility of spacetimes and Lorentzian length spaces. Ann. Glob. Anal. Geom. **55**(1), 133–147 (2019)
41. Gromov, M.: Synthetic geometry in Riemannian manifolds. Proc. Int. Congress Math. **1**, 415–419 (1978)
42. Gromov, M.: Metric Structures for Riemannian and Non-Riemannian Spaces. Birkhauser (1999)
43. Harris, S.: A triangle comparison theorem for Lorentz manifolds. Indiana Univ. Math. J. **31**(3), 289–308 (1982)
44. Hawking, S.W., King, A.R., McCarthy, P.J.: A new topology for curved space-time which incorporates the causal, differential, and conformal structures. J. Math. Phys. **17**(2), 174–181 (1976)
45. Hedicke, J.: Lorentzian distance functions in contact geometry. J. Topol. Anal. 1–21 (2022)
46. Innami, N.: Splitting theorems of Riemannian manifolds. Comput. Math. **47**, 237–247 (1982)
47. Kunzinger, M., Oberguggenberger, M., Vickers, J. A.: Synthetic versus distributional lower Ricci curvature bounds. arXiv:2207.03715 (2022)
48. Kunzinger, M., Sämann, C.: Lorentzian length spaces. Ann. Glob. Anal. Geom. **54**(3), 399–447 (2018)
49. Kunzinger, M., Steinbauer, R.: Null distance and convergence of Lorentzian length spaces. Ann. Henri Poincare 1424 (2022)
50. Ling, E.: Aspects of C^0 causal theory. Gen. Relativ. Gravit. **52**(57) (2020)
51. Mashiko, Y.: A splitting theorem for Alexandrov spaces. Pac. J. Math. **204**, 445–458 (2002)
52. McCann, R. J.: Private communication (2023)
53. Milka, A.D.: Metric structure of a certain class of spaces that contain straight lines. Ukrain. Geometr. Sb. Vyp. **4**, 43–48 (1967)
54. Minguzzi, E.: Causality theory for closed cone structures with applications. Rev. Math. Phys. **31**(05), 1930001 (2019)
55. Newman, R. P.: A proof of the splitting conjecture of S.-T. Yau. J. Differ. Geom. **31**(1):163–184 (1990)
56. Noronha, M.: A splitting theorem for complete manifolds with non-negative curvature operator. Proc. Am. Math. Soc. **105**, 979–985 (1989)
57. Paeng, S.-H.: Manifolds with non-negative Ricci curvature almost everywhere. J. Korean Math. Soc. **36**(1), 125–137 (1999)
58. Rott, F.: Gluing of Lorentzian length spaces and the causal ladder. arXiv:2209.06894 (2022)
59. Steinbauer, R.: The singularity theorems of General Relativity and their low regularity extensions. arXiv:2206.05939 (2022)
60. Sturm, K.T.: On the geometry of measure metric spaces. Acta Math. **196**, 65–131 (2006)
61. Toponogov, V.: Riemannian spaces with straight lines. AMS Transl. **37**, 287–290 (1964)
62. Toponogov, V.: The metric structure of Riemannian spaces with non-negative curvature which contain straight lines. Am. Math. Soc. Transl. **70**, 225–239 (1968)
63. Villani, C.: Optimal Transport, Old and New. Springer (2009)
64. Wang, L.: A splitting theorem for the weighted measure. Ann. Glob. Anal. Geom. **42**, 79–89 (2012)
65. Yau, S.T.: Problem section. In: Annals of Mathematics Studies. Princeton University Press (1982)