Communications in
**Mathematical
Physics**

A Quasi-Local Mass

Aghil Alaee[1,2] ⃝, **Marcus Khuri**[3], **Shing-Tung Yau**[4,5]

1 Department of Mathematics, Clark University, Worcester, MA 01610, USA.
  E-mail: aalaeekhangha@clarku.edu; aghil.alaee@cmsa.fas.harvard.edu
2 Center of Mathematical Sciences and Applications, Harvard University, Cambridge, MA 02138, USA
3 Department of Mathematics, Stony Brook University, Stony Brook, NY 11794, USA.
  E-mail: marcus.khuri@stonybrook.edu
4 Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China
5 Yanqi Lake Beijing Institute of Mathematical Sciences and Applications, Beijing 101408, China.
  E-mail: styau@tsinghua.edu.cn; yau@math.harvard.edu

Abstract: We define a new gauge independent quasi-local mass and energy, and show its relation to the Brown–York Hamilton–Jacobi analysis. A quasi-local proof of the positivity, based on spacetime harmonic functions, is given for admissible closed space-like 2-surfaces which enclose an initial data set satisfying the dominant energy condition. Like the Wang-Yau mass, the new definition relies on isometric embeddings into Minkowski space, although our notion of admissibility is different from that of Wang and Yau. Rigidity is also established, in that vanishing energy implies that the 2-surface arises from an embedding into Minkowski space, and conversely the mass vanishes for any such surface. Furthermore, we show convergence to the ADM mass at spatial infinity, and provide the equation associated with optimal isometric embedding.

1. Introduction

In 1982, Penrose [25] listed a set of major open problems in which one was to *find a suitable quasi-local definition of energy-momentum in general relativity*. As is well-known, the fundamental difficulty is that there is no natural notion of energy density for the gravitational field, due to Einstein's principle of equivalence. This has led to a plethora of different mathematical formulations including: the localization of ADM mass by Bartnik [2], twistor and spinor approaches of Dougan and Mason [8], Lott [21], Ludvigsen and Vickers [22], Penrose [26], and Zhang [41], the Hawking mass [13], as well as Hamilton–Jacobi methods employed by Booth and Mann [3], Brown and York [6], Epp [10], Hawking and Horowitz [14], Kijowski [18], Liu and Yau [20], Tsang [36], and Wang and Yau [38]. For a detailed account of the various quasi-local masses, see the survey paper of Szabados [35].

There are several desirable properties that a quasi-local mass should have. For instance, like the positive mass theorem it should be nonnegative for a large class of

surfaces in the presence of the dominant energy condition, and exhibit rigidity in the sense that it vanishes if and only if the surface arises from Minkowski space. Appropriate asymptotics are also important, in that the ADM and Bondi masses should be recovered in the large sphere limits at spatial and null infinities. Moreover, gauge independence and relation with a Hamilton–Jacobi analysis are preferable, in order to aid with physical relevance and interpretations. While many other properties of the quasi-local mass itself may be added to the list, a potentially advantageous characteristic of a proof of nonnegativity is that it be quasi-local. This refers to a proof strategy that appeals solely to a compact initial data set enclosed by the spacelike 2-surface in question, as opposed to the use of asymptotically flat (or other) extensions and the positive mass theorem. The latter strategy of extensions is essential in the work of Shi and Tam [32] on the Brown–York mass, as well as for the Liu-Yau [20] and Wang-Yau [37,38] masses, and is even contained within the definition of the Bartnik mass [2]. In fact, Schoen asked in [30] whether it is possible to *find a quasi-local proof of nonnegativity for the Brown–York and related quasi-local masses*. An affirmative answer to this question for the Brown–York definition has recently been put forward by Montiel [23], utilizing a spinorial approach.

In the current paper, we introduce a new gauge independent expression for quasi-local mass with a quasi-local proof of nonnegativity. This mass also comes with a Hamilton–Jacobi interpretation, and although it shares some similarities with the Wang-Yau definition, the new notion has a completely separate derivation and involves a different range of applicability. The derivation and basic motivation comes from certain integral expressions associated with spacetime harmonic functions [15]. In [33], level set techniques for harmonic maps to S^1 were introduced into the study of scalar curvature on closed 3-dimensional Riemannian manifolds. Inspired by this, a proof of the Riemannian positive mass theorem [5] was given based on level sets of asymptotically linear harmonic functions. Spacetime harmonic functions were then introduced [15] to treat the asymptotically flat spacetime version of the positive mass theorem, and have subsequently been used to prove the corresponding theorem in asymptotically (locally) hyperbolic settings [1,4], as well as for comparison theorems in Riemannian geometry [16,17]; surveys of some recent advancements may be found in [4,34]. Throughout this work, unless specified otherwise, all manifolds will be assumed to be connected, oriented, and smooth.

Let Σ be a closed (compact without boundary) spacelike 2-surface having induced metric σ in spacetime $N^{3,1}$ with metric $\langle \cdot, \cdot \rangle$ of signature $(-+++)$. Let $\{e_3, e_4\}$ be an orthonormal frame for the (time oriented) normal bundle consisting of a spacelike and future directed timelike vector respectively; note that this bundle is topologically trivial since the classifying space for $SO_+(1, 1) \cong \mathbb{R}$ is a point. In this gauge, the normal bundle connection 1-form and mean curvature vector are given by

$$\alpha_{e_3}(\cdot) = \langle \nabla^N_{(\cdot)} e_3, e_4 \rangle, \qquad \vec{H} = (\mathrm{div}_\sigma e_3)e_3 - (\mathrm{div}_\sigma e_4)e_4. \tag{1.1}$$

Consider an isometric embedding $\iota : \Sigma \hookrightarrow \mathbb{R}^{3,1}$, and choose Cartesian coordinates $(\mathbf{t}, \mathbf{x}^i)$, $i = 1, 2, 3$ for the target Minkowski space. One may then obtain a function $u_a = \iota^* \left(-\mathbf{t} + a_i \mathbf{x}^i \right)$ on Σ for any set of constants $\mathbf{a} = (a_1, a_2, a_3)$, which will be restricted to satisfy $|\mathbf{a}|^2 = \sum_i a_i^2 = 1$. Recall that in [37,38], Wang-Yau parameterize a class of isometric embeddings into Minkowski space with a time function τ on Σ satisfying the convexity condition

$$\left(1 + |\nabla_\partial \tau|^2 \right) K_{\tilde{\sigma}} = K_\sigma + \frac{\det(\nabla^2_\partial \tau)}{(\det \sigma) \left(1 + |\nabla_\partial \tau|^2 \right)} > 0, \tag{1.2}$$

where K_σ and $K_{\tilde\sigma}$ denote the Gaussian curvatures of σ and $\tilde\sigma = \sigma + d\tau^2$, and ∇_∂ is the connection with respect to σ. By the classical theorem of Nirenberg [24] and Pogorelov [27] there exists a unique isometric embedding up to rigid motion into $\mathbb{R}^3$, and from this one obtains an isometric embedding into $\mathbb{R}^{3,1}$. An alternative method to produce isometric embeddings, which does not rely on (1.2), is to embed Σ into a (hyperboloid) hyperbolic space $\mathbb{H}^3_{-\kappa} \subset \mathbb{R}^{3,1}$ with large $\kappa > 0$ to aid with ellipticity. For instance, every metric on the 2-sphere admits such an isometric embedding [28,29], and a related result of Gromov [12, Section 3.2.4] shows that any closed 2-surface admits an isometric embedding into some complete 3-dimensional Riemannian target space of constant negative sectional curvature.

The quintuple $(\sigma, \vec{H}, \alpha, \iota, u_a)$ will be used to build the quasi-local energy, and may be referred to as a *quasi-local data set* for Σ. If the mean curvature vector is spacelike, then for any constant $\varepsilon > 0$, there exists a unique orthonormal frame $\{\bar{e}_3, \bar{e}_4\}$ for the normal bundle of Σ such that

$$\langle \vec{H}, \bar{e}_3 \rangle > 0, \qquad \langle \vec{H}, \bar{e}_4 \rangle = \frac{-\Delta_\partial u_a}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}. \tag{1.3}$$

We define the quasi-local energy with respect to the observer determined by the pair (ι, u_a) to be

$$E(\Sigma, \iota, u_a) = \lim_{\varepsilon \to 0} \frac{1}{8\pi} \int_\Sigma \left(\mathcal{H}_0(\hat{e}_3, u_a) - \mathcal{H}(\bar{e}_3, u_a) \right) dA, \tag{1.4}$$

where the *quasi-local Hamiltonian density* is

$$\mathcal{H}(\bar{e}_3, u_a) = \sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}\, \langle \vec{H}, \bar{e}_3 \rangle + \alpha_{\bar{e}_3}(\nabla_\partial u_a) \tag{1.5}$$

and $\mathcal{H}_0$ represents the same quantity for $\iota(\Sigma) \subset \mathbb{R}^{3,1}$ with $\{\hat{e}_3, \hat{e}_4\}$ denoting the frame given by equations (1.3) in the Minkowski setting. In analogy with special relativity, the quasi-local mass is then set to be the infimum of energy over all admissible observers

$$\mathrm{mass}(\Sigma) = \inf_{(\iota, u_a)} E(\Sigma, \iota, u_a). \tag{1.6}$$

Due to the null trajectory of the observers, this mass may be interpreted as the difference of energy and norm of linear momentum from the 4-momentum vector as opposed to its Lorentz length; see the remark in [7, p. 1496] for a related discussion of this discrepancy.

A pair (ι, u_a) will be called *admissible* for $\Sigma \subset N^{3,1}$ if the mean curvature vector of $\iota(\Sigma)$ is spacelike, and the regular level sets of u_a are maximal in the following sense. Namely, there exists a compact spacelike hypersurface $\hat{\Omega} \subset \mathbb{R}^{3,1}$ with $\partial\hat{\Omega} = \iota(\Sigma)$, such that if any regular s-level set of u_a has n components then the Euler characteristic of the s-level fill-in with respect to $\hat{\Omega}$ satisfies $\chi(\hat{\Sigma}_s) = n$. Here, $\hat{\Sigma}_s$ denotes the set of points within $\hat{\Omega}$ satisfying the level set equation $s = -\mathbf{t} + a_i\mathbf{x}^i$. Furthermore, we will say that the *dominant energy condition* holds for a spacetime if the Einstein equations are enforced with stress-energy tensor satisfying $T(\mathbf{v}, \mathbf{w}) \geq 0$ for all future-pointing causal vectors $\mathbf{v}$ and $\mathbf{w}$.

Theorem 1.1. *Let Σ be a closed spacelike 2-surface in a 4-dimensional spacetime satisfying the dominant energy condition, and assume that Σ has an outward pointing spacelike mean curvature vector and bounds a compact spacelike hypersurface Ω with trivial second homology $H_2(\Omega; \mathbb{Z}) = 0$. The following statements hold with $\mathbf{a} \in \mathbb{R}^3$ satisfying $|\mathbf{a}| = 1$.*

1. *If an isometric embedding $\iota : \Sigma \hookrightarrow \mathbb{R}^{3,1}$ together with function u_a is admissible for Σ, then the quasi-local energy limit exists, is finite, and is nonnegative: $E(\Sigma, \iota, u_a) \geq 0$.*
2. *Under the same hypotheses, if $E(\Sigma, \iota, u_a) = 0$ for all $\mathbf{a}$ then $(\Sigma, \sigma, \vec{H}, \alpha)$ arises from Minkowski space.*
3. *If $\Sigma \subset \mathbb{R}^{3,1}$ is such that the inclusion map is admissible with some u_a, then the quasi-local mass vanishes: mass$(\Sigma) = 0$.*

Remark 1.2. Nonnegativity of the energy holds in the more general circumstance of a disconnected Σ and without the homology assumption on Ω, when a suitable modification of the admissibility condition is enforced. This is discussed in more detail below at the end of Sect. 4.

A model situation for an admissible pair (ι, u_a) occurs when Σ is a topological 2-sphere, and its isometric image in Minkowski space bounds a 3-ball whose generic intersection with the null hyperplanes associated to u_a is in the form of a disc. Therefore, this theorem yields positivity under the dominant energy condition for a large class of surfaces. Moreover, because this admissibility condition is not based on a convexity property such as (1.2) which appears in [38, Definition 5.1], it is applicable beyond spheres to surfaces of positive genus. Beyond this difference with the Wang-Yau mass in terms of the admissibility conditions and range of applicability, there are two other immediate distinguishing characteristics between these two quasi-local masses. First, as explained in Sect. 3, a Hamilton–Jacobi analysis shows that energy (1.4) is measured by a null observer while the Wang-Yau energy is measured by a timelike observer. Secondly, while the proof of nonnegativity for the Wang-Yau mass relies on asymptotically flat extensions and ultimately the positive mass theorem, the proof here requires only quasi-local information.

A consequence of the proof of Theorem 1.1 is that the limit may be evaluated explicitly in definition (1.4). In particular, if $\tilde{\Sigma}$ denotes the open subset of Σ on which $|\nabla_\partial u_a| \neq 0$ then we have

$$E(\Sigma, \iota, u_a) = \frac{1}{8\pi} \int_{\tilde{\Sigma}} \left(\sqrt{|\vec{H}_0|^2 |\nabla_\partial u_a|^2 + (\Delta_\partial u_a)^2} + \nabla_\partial f_0 \cdot \nabla_\partial u_a + \alpha_{\frac{\vec{H}_0}{|\vec{H}_0|}} (\nabla_\partial u_a) \right) dA$$
$$- \frac{1}{8\pi} \int_{\tilde{\Sigma}} \left(\sqrt{|\vec{H}|^2 |\nabla_\partial u_a|^2 + (\Delta_\partial u_a)^2} + \nabla_\partial f \cdot \nabla_\partial u_a + \alpha_{\frac{\vec{H}}{|\vec{H}|}} (\nabla_\partial u_a) \right) dA,$$
(1.7)

where $\vec{H}_0$ is the mean curvature vector of the isometric embedding $\iota(\Sigma)$ and

$$f_0 = \sinh^{-1} \left(\frac{\Delta_\partial u_a}{|\vec{H}_0| |\nabla_\partial u_a|} \right), \qquad f = \sinh^{-1} \left(\frac{\Delta_\partial u_a}{|\vec{H}| |\nabla_\partial u_a|} \right).$$
(1.8)

In addition to the attributes of positivity and rigidity, the new definition also behaves well with respect to asymptotic limits at spatial infinity. Let (M, g, k) be a 3-dimensional initial data set for the Einstein equations, with g denoting a Riemannian metric and k a symmetric 2-tensor representing the extrinsic curvature in spacetime. These objects satisfy the constraint equations

$$\mu = \frac{1}{2} \left(R_g + (\mathrm{Tr}_g k)^2 - |k|^2 \right), \qquad J = \mathrm{div}_g \left(k - (\mathrm{Tr}_g k) g \right),$$
(1.9)

where R_g is the scalar curvature and μ and J represent the energy and momentum density of matter fields multiplied by 8π. The data are asymptotically flat (with one end) if outside a compact set, M is diffeomorphic to the compliment of a ball in Euclidean space $\mathbb{R}^3 \setminus B_1$, and in the coordinates provided by this diffeomorphism

$$|\partial^l(g_{ij} - \delta_{ij})(x)| = O(|x|^{-\tau-l}) \quad l = 0, 1, 2, 3,$$
$$|\partial^l k_{ij}(x)| = O(|x|^{-\tau-1-l}) \quad l = 0, 1, \tag{1.10}$$

for some $\tau > \frac{1}{2}$. An extra third derivative of g is included for control of isometric embeddings in the next result. The energy and momentum densities will be taken to be integrable $\mu, J \in L^1(M)$ so that the ADM energy and linear momentum are well-defined and given by

$$\mathcal{E} = \lim_{r\to\infty} \frac{1}{16\pi} \int_{S_r} \sum_i \left(g_{ij,i} - g_{ii,j}\right) \upsilon^j dA,$$

$$\mathcal{P}_i = \lim_{r\to\infty} \frac{1}{8\pi} \int_{S_r} \left(k_{ij} - (\mathrm{Tr}_g k)g_{ij}\right) \upsilon^j dA, \tag{1.11}$$

where υ is the unit outer normal to the coordinate sphere S_r of radius $r = |x|$ and dA denotes its area element. The ADM mass is the Lorentz length of the ADM energy-momentum vector $(\mathcal{E}, \mathcal{P})$. If the dominant energy condition is satisfied which implies that $\mu \geq |J|$, then the spacetime positive mass theorem [9,15,31,40] asserts that the ADM energy-momentum is nonspacelike, and characterizes Minkowski space as the unique spacetime having asymptotically flat initial data with vanishing mass; see [19] for a detailed account.

Theorem 1.3. *Let (M, g, k) be an asymptotically flat initial data set for the Einstein equations, and let $\iota_r : S_r \hookrightarrow \mathbb{R}^{3,1}$ denote the (unique up to Euclidean motion) isometric embedding of the r-coordinate sphere into a constant time slice $\mathbb{R}^3 \subset \mathbb{R}^{3,1}$. Then for any $\mathbf{a}$, the asymptotic limit of quasi-local energies is given in terms of the ADM energy and linear momentum by*

$$\lim_{r\to\infty} E(S_r, \iota_r, u_a) = \mathcal{E} - \langle \mathbf{a}, \mathcal{P} \rangle. \tag{1.12}$$

This paper is organized as follows. In the next section the main ideas behind the derivation of the energy will be explained, while in Sect. 3 a physical interpretation will be given in terms of a Hamilton–Jacobi analysis. Section 4 is dedicated to the proof of nonnegativity, and Sect. 5 deals with the rigidity statement of Theorem 1.1. Moreover, the asymptotic behavior of Theorem 1.3 will be established in Sect. 6, while the equation associated with optimal isometric embedding is obtained in Sect. 7.

2. Derivation of the Energy

A motivation for the quasi-local energy comes from the level set technique and certain integral formulae involving *spacetime harmonic functions* [15]. Consider a compact initial data set (Ω, g, k). Recall that a function $u \in C^2(\Omega)$ is spacetime harmonic if it satisfies the equation

$$\Delta_g u + (\mathrm{Tr}_g k)|\nabla u| = 0, \tag{2.1}$$

and note that this arises as the trace of the *spacetime Hessian*

$$\overline{\nabla}_{ij}u := \nabla_{ij}u + k_{ij}|\nabla u|. \tag{2.2}$$

The following inequality was established in [1, Proposition 3.1].

Proposition 2.1. *Let* (Ω, g, k) *be a 3-dimensional compact initial data set with smooth boundary* $\partial\Omega$, *having outward unit normal* v. *Let* $u : \Omega \to \mathbb{R}$ *be a spacetime harmonic function which lies in* $C^{2,\varsigma}(\Omega)$, $0 < \varsigma < 1$, *and consider the open subset* $\bar{\partial}\Omega$ *of the boundary on which* $|\nabla_\partial u| \neq 0$, *where* $\nabla_\partial u$ *is the projection of the full gradient onto the boundary tangent space. If* $\bar{u}$ *and* $\underline{u}$ *are the maximum and minimum values of* u *and* $\Sigma_s = u^{-1}(s)$, *then*

$$\int_{\partial\Omega} (k(\nabla_\partial u, v) - |\nabla u|H - v(u)\mathrm{Tr}_{\partial\Omega}k)\,dA$$
$$+ \int_{\bar{\partial}\Omega} \frac{|\nabla_\partial u|}{|\nabla u|} \nabla_\partial u \left(\frac{v(u)}{|\nabla_\partial u|} \right) dA + 2\pi \int_{\underline{u}}^{\bar{u}} \chi(\Sigma_s)ds \tag{2.3}$$
$$\geq \int_\Omega \left(\frac{1}{2} \frac{|\overline{\nabla}^2 u|^2}{|\nabla u|} + \mu|\nabla u| + J(\nabla u) \right) dV,$$

where $\chi(\Sigma_s)$ *denotes the Euler characteristic, and* H *is the mean curvature of the boundary with respect to* v.

The right-hand side of (2.3) is nonnegative if the dominant energy condition holds, and this suggests investigating the boundary terms in relation to a quasi-local energy. In order to better interpret the boundary expression with regards to spacetime geometry, assume that the data arise from a spacetime $(\Omega, g, k) \hookrightarrow N^{3,1}$ and let $\mathbf{n}$ be the associated unit timelike future directed normal vector field. Consider the frame $\{v, \mathbf{n}\}$ for the $SO(1, 1)$ normal bundle over $\partial\Omega$, and observe that the mean curvature vector and an auxiliary vector field are given by

$$\vec{H} = Hv - (\mathrm{Tr}_{\partial\Omega}k)\mathbf{n}, \qquad \vec{w} = |\nabla u|v + v(u)\mathbf{n}, \tag{2.4}$$

with $|\vec{w}|^2 = |\nabla_\partial u|^2$. Therefore, two of the boundary terms combine to form

$$\langle \vec{H}, \vec{w} \rangle = |\nabla u|H + v(u)\mathrm{Tr}_{\partial\Omega}k. \tag{2.5}$$

The remaining two boundary terms may be interpreted as follows. In the gauge determined by the current frame, the connection 1-form for the normal bundle is

$$\alpha_v(X) = -k(X, v) = \langle \nabla_X^N v, \mathbf{n} \rangle, \tag{2.6}$$

where X is any tangent vector field to the boundary surface. Notice that with a change of gauge to the frame determined by $e_3 = av + b\mathbf{n}$, $e_4 = bv + a\mathbf{n}$ with $a = \cosh f$, $b = \sinh f$ yields

$$\alpha_{e_3}(X) = \langle \nabla_X^N e_3, e_4 \rangle = a^2 \langle \nabla_X^N v, \mathbf{n} \rangle + b^2 \langle \nabla_X^N \mathbf{n}, v \rangle + X(a)b - X(b)a$$
$$= \alpha_v(X) - X(f), \tag{2.7}$$

for any function $f \in C^1(\partial\Omega)$. Below, for simplicity of the discussion, we will *assume that all calculations occur away from critical points* of u restricted to the boundary. Then choosing $X = \nabla_\partial u$ and $f = \sinh^{-1}(v(u)/|\nabla_\partial u|)$ produces

$$X(f) = \nabla_\partial u \left(\sinh^{-1} \left(\frac{v(u)}{|\nabla_\partial u|} \right) \right) = \frac{|\nabla_\partial u|}{|\nabla u|} \nabla_\partial u \left(\frac{v(u)}{|\nabla_\partial u|} \right). \tag{2.8}$$

Hence, combining (2.6), (2.7), and (2.8) shows that

$$\alpha_{e_3}(\nabla_\partial u) = \alpha_v(\nabla_\partial u) - \nabla_\partial u \left(\sinh^{-1} \left(\frac{v(u)}{|\nabla_\partial u|} \right) \right)$$

$$= -k(\nabla_\partial u, v) - \frac{|\nabla_\partial u|}{|\nabla u|} \nabla_\partial u \left(\frac{v(u)}{|\nabla_\partial u|} \right). \tag{2.9}$$

Observe that the vector $\vec{w}$ is proportional to a special case (when $\varepsilon = 0$) of the first member of the *level set frame*

$$e_3' = \frac{\sqrt{|\nabla u|^2 + \varepsilon^2}}{\sqrt{|\nabla_\partial u|^2 + \varepsilon^2}} v + \frac{v(u)}{\sqrt{|\nabla_\partial u|^2 + \varepsilon^2}} \mathbf{n},$$

$$e_4' = \frac{v(u)}{\sqrt{|\nabla_\partial u|^2 + \varepsilon^2}} v + \frac{\sqrt{|\nabla u|^2 + \varepsilon^2}}{\sqrt{|\nabla_\partial u|^2 + \varepsilon^2}} \mathbf{n}. \tag{2.10}$$

Here, the parameter $\varepsilon > 0$ is included to avoid the technical issue of critical points for the restriction of u to $\partial\Omega$. The computations (2.5) and (2.9) then motivate us to define the quasi-local Hamiltonian density as

$$\mathcal{H}(e_3', u) = \sqrt{|\nabla_\partial u|^2 + \varepsilon^2} \, \langle \vec{H}, e_3' \rangle + \alpha_{e_3'}(\nabla_\partial u), \tag{2.11}$$

with respect to the level set frame. The expression for quasi-local energy $E(\Sigma, \iota, u_a)$ in (1.4) follows by choosing an 'optimal' frame, and comparing to an appropriate Hamiltonian density in the ground state.

3. A Hamilton–Jacobi Interpretation of the Energy

In this section we will describe the relationship between the quasi-local energy defined in the introduction, and a Hamilton–Jacobi analysis. Recall that Brown and York [6] and Hawking and Horowitz [14] derived a Hamiltonian for closed spacelike 2-surfaces $\Sigma \hookrightarrow N^{3,1}$ which enclose a compact initial data set (Ω, g, k). As before let $\mathbf{n}$ be the unit timelike future directed normal to the data, and let v be the unit outer normal to Σ with respect to Ω. If $T = \varphi \mathbf{n} + Y$ is a timelike vector field along Σ, representing an observer with lapse φ and shift Y, then the surface Hamiltonian takes the form

$$\mathbf{H}(\Sigma, T, \mathbf{n}) = -\frac{1}{8\pi} \int_\Sigma \left(\varphi H - k(v, Y) + (\mathrm{Tr}_g k) g(v, Y) \right) dA, \tag{3.1}$$

where dA is the area element on Σ. As pointed out by Wang and Yau [37], this Hamiltonian may be reexpressed with the aid of the vector field

$$P = H\mathbf{n} + k(v) - (\mathrm{Tr}_g k)v, \tag{3.2}$$

so that

$$\mathbf{H}(\Sigma, T, \mathbf{n}) = \frac{1}{8\pi} \int_{\Sigma} \langle P, T \rangle dA. \tag{3.3}$$

Note that P is perpendicular to the the mean curvature vector $\vec{H} = H\nu - (\mathrm{Tr}_{\Sigma}k)\mathbf{n}$. The associated energy is then defined by choosing a reference Hamiltonian, determined by an isometric embedding $\iota : \Sigma \hookrightarrow \mathbb{R}^{3,1}$ and corresponding vector fields T_0 and $\mathbf{n}_0$ along the image in Minkowski space, namely

$$E(\Sigma) = \mathbf{H}(\Sigma, T, \mathbf{n}) - \mathbf{H}_0(\iota(\Sigma), T_0, \mathbf{n}_0). \tag{3.4}$$

This notion of energy depends on the choices of T, $\mathbf{n}$ and T_0, $\mathbf{n}_0$, as well as on the isometric embedding. Observe that if the isometric embedding lands in a time slice $\mathbb{R}^3 \subset \mathbb{R}^{3,1}$ with normal $\mathbf{n}_0$, and the choices $T = \mathbf{n}$ and $T_0 = \mathbf{n}_0$ are made, then the typical expression for the the Brown–York mass is recovered; it depends on the initial data Ω and hence on $\mathbf{n}$. The Liu-Yau mass is obtained with the same prescription, except that $\mathbf{n}$ is taken to satisfy $\langle \vec{H}, \mathbf{n} \rangle = 0$. Furthermore, given an admissible time function τ on Σ associated with an isometric embedding, the Wang-Yau energy is produced by setting

$$T_0 = \sqrt{1 + |\nabla_{\partial}\tau|^2}\,\mathbf{n}_0 - \nabla_{\partial}\tau, \tag{3.5}$$

and choosing $\mathbf{n}_0$ so that $\{\nu_0, \mathbf{n}_0\}$ is the unique frame for the normal bundle of $\iota(\Sigma)$ satisfying

$$\langle \vec{H}_0, \nu_0 \rangle > 0, \qquad \langle \vec{H}_0, \mathbf{n}_0 \rangle = \frac{-\Delta_{\partial}\tau}{\sqrt{1 + |\nabla_{\partial}\tau|^2}}, \tag{3.6}$$

while T is set to have the same lapse and shift as T_0 and $\mathbf{n}$ is chosen to satisfy conditions analogous to (3.6) in $N^{3,1}$.

In order to apply these considerations to the quasi-local energy introduced in Sect. 1, let $(\mathbf{t}, \mathbf{x}^i)$, $i = 1, 2, 3$ be coordinates for the reference Minkowski space, and pull back a linear null function $u_a = \iota^*(-\mathbf{t} + a_i\mathbf{x}^i)$ to Σ. For each $\varepsilon > 0$ we then set

$$T_0 = \sqrt{|\nabla_{\partial}u_a|^2 + \varepsilon^2}\,\mathbf{n}_0 + \nabla_{\partial}u_a, \tag{3.7}$$

and choose $\mathbf{n}_0$ so that $\{\nu_0, \mathbf{n}_0\}$ is the unique frame for the normal bundle of $\iota(\Sigma)$ satisfying

$$\langle \vec{H}_0, \nu_0 \rangle > 0, \qquad \langle \vec{H}_0, \mathbf{n}_0 \rangle = \frac{-\Delta_{\partial}u_a}{\sqrt{|\nabla_{\partial}u_a|^2 + \varepsilon^2}}. \tag{3.8}$$

Moreover, T is set to have the same lapse and shift as T_0, and $\mathbf{n}$ is chosen to satisfy conditions analogous to (3.8) in $N^{3,1}$. Note that T and T_0 are timelike with $|T|^2 = |T_0|^2 = -\varepsilon^2$, and are approaching null vectors as $\varepsilon \to 0$. Observe that writing $\langle \vec{H}, \nu \rangle = H$ and $\langle \vec{H}, \mathbf{n} \rangle = \mathrm{Tr}_{\Sigma}k$ gives rise to

$$\begin{aligned}
\mathbf{H}^{\varepsilon}(\Sigma, T, \mathbf{n}) &:= \frac{1}{8\pi} \int_{\Sigma} \langle P, T \rangle dA \\
&= \frac{1}{8\pi} \int_{\Sigma} \left(-H\sqrt{|\nabla_{\partial}u_a|^2 + \varepsilon^2} + k(\nu, \nabla_{\partial}u_a) \right) dA \\
&= \frac{1}{8\pi} \int_{\Sigma} \left(-\sqrt{|\nabla_{\partial}u_a|^2 + \varepsilon^2}\langle \vec{H}, \nu \rangle - \alpha_{\nu}(\nabla_{\partial}u_a) \right) dA \\
&= -\frac{1}{8\pi} \int_{\Sigma} \mathcal{H}(\nu, u_a) dA,
\end{aligned} \tag{3.9}$$

and similarly for the reference Hamiltonian. Therefore, the new quasi-local energy arises from Hamiltonians by taking a limit as the observer approaches a null direction

$$E(\Sigma, \iota, u_a) = \lim_{\varepsilon \to 0} \left(\mathbf{H}^\varepsilon (\Sigma, T, \mathbf{n}) - \mathbf{H}_0^\varepsilon (\iota(\Sigma), T_0, \mathbf{n}_0) \right), \tag{3.10}$$

where in terms of the notation of the introduction we have $\{\bar{e}_3, \bar{e}_4\} = \{\nu, \mathbf{n}\}$ and $\{\hat{e}_3, \hat{e}_4\} = \{\nu_0, \mathbf{n}_0\}$.

4. Proof of Nonnegativity

The purpose of the current section is to establish the inequality portion of Theorem 1.1. As before, given a linear null function u_a on a spacelike 2-surface Σ in spacetime $N^{3,1}$, consider the normal bundle frame $\{\bar{e}_3, \bar{e}_4\}$ defined by

$$\langle \vec{H}, \bar{e}_3 \rangle > 0, \qquad \langle \vec{H}, \bar{e}_4 \rangle = \frac{-\Delta_\partial u_a}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}, \tag{4.1}$$

for $\varepsilon > 0$. We begin with a preliminary result, similar to [38, Proposition 2.1] for the Wang-Yau quasi-local energy, which demonstrates how this may be interpreted as an optimal frame.

Lemma 4.1. *If the mean curvature vector $\vec{H}$ of Σ is spacelike, and $\{e_3, e_4\}$ is any frame for the normal bundle of Σ with the properties that e_3 is spacelike and $\langle \vec{H}, e_3 \rangle > 0$, then*

$$\int_\Sigma \mathcal{H}(e_3, u_a) dA \geq \int_\Sigma \mathcal{H}(\bar{e}_3, u_a) dA \tag{4.2}$$

for all $\varepsilon > 0$.

Proof. Consider the following functional which sends normal bundle frames to the real numbers

$$\{e_3, e_4\} \mapsto \int_\Sigma \mathcal{H}(e_3, u_a) dA = \int_\Sigma \left(\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2} \, \langle \vec{H}, e_3 \rangle + \alpha_{e_3}(\nabla_\partial u_a) \right) dA. \tag{4.3}$$

Since the normal bundle of Σ is rank 2 with structure group $SO(1, 1)$, each frame may be given by a hyperbolic angle function. In particular, because $\vec{H}$ is spacelike and $\langle \vec{H}, e_3 \rangle > 0$, we may express the frame defined by the mean curvature vector as

$$\bar{e}_3 := \frac{\vec{H}}{|\vec{H}|} = (\cosh f) e_3 + (\sinh f) e_4, \qquad \bar{e}_4 := (\sinh f) e_3 + (\cosh f) e_4, \tag{4.4}$$

for some $f \in C^\infty(\Sigma)$. It follows that

$$\langle \vec{H}, e_3 \rangle = |\vec{H}| \cosh f, \qquad \langle \vec{H}, e_4 \rangle = -|\vec{H}| \sinh f. \tag{4.5}$$

Furthermore, using Eqs. (2.7) and (4.4) produces

$$\alpha_{e_3}(\nabla_\partial u_a) = \alpha_{\bar{e}_3}(\nabla_\partial u_a) + \nabla_\partial u_a \cdot \nabla_\partial f. \tag{4.6}$$

Therefore, the functional of (4.3) can be rewritten as

$$
\begin{aligned}
f \mapsto & \int_{\Sigma} \left(\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2} |\vec{H}| \cosh f + \nabla_\partial u_a \cdot \nabla_\partial f + \alpha_{\tilde{e}_3}(\nabla_\partial u_a) \right) dA \\
= & \int_{\Sigma} \left(\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2} |\vec{H}| \cosh f - f \Delta_\partial u_a + \alpha_{\tilde{e}_3}(\nabla_\partial u_a) \right) dA.
\end{aligned}
\tag{4.7}
$$

Since $|\vec{H}| > 0$ this functional is convex, and it may be easily checked that the minimum occurs when

$$
|\vec{H}| \sinh f = \frac{\Delta_\partial u_a}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}.
\tag{4.8}
$$

Hence, (4.5) shows that the minimum is achieved at the frame $\{\bar{e}_3, \bar{e}_4\}$. $\square$

Let (Ω, g, k) be the initial data set for a compact spacelike hypersurface in $N^{3,1}$ which is enclosed by Σ, and let $u_a \in C^\infty(\Sigma)$. Consider the unique solution of the spacetime harmonic Dirichlet problem

$$
\Delta_g u + (\mathrm{Tr}_g k)|\nabla u| = 0 \quad \text{in } \Omega, \qquad u = u_a \quad \text{on } \partial\Omega = \Sigma.
\tag{4.9}
$$

The existence of a unique solution $u \in C^{2,\varsigma}(\Omega)$ for any $\varsigma \in (0, 1)$, follows in a straightforward manner from the results of [15, Section 4.1]. We may then define a level set frame for the normal bundle of Σ by

$$
\begin{aligned}
e_3' &= \frac{\sqrt{|\nabla u|^2 + \varepsilon^2}}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \nu + \frac{\nu(u)}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \mathbf{n}, \\
e_4' &= \frac{\nu(u)}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \nu + \frac{\sqrt{|\nabla u|^2 + \varepsilon^2}}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \mathbf{n},
\end{aligned}
\tag{4.10}
$$

for each $\varepsilon > 0$ and where $\{\nu, \mathbf{n}\}$ is the normal bundle frame determined by Ω in which ν is the outer normal to $\partial\Omega$ and $\mathbf{n}$ is future directed timelike. Note that since the mean curvature vector of Σ is outward pointing spacelike, we have

$$
H = \langle \vec{H}, \nu \rangle > |\langle \vec{H}, \mathbf{n} \rangle| = |\mathrm{Tr}_\Sigma k|
\tag{4.11}
$$

and therefore

$$
\begin{aligned}
\langle \vec{H}, e_3' \rangle &= \frac{\sqrt{|\nabla u|^2 + \varepsilon^2}}{\sqrt{|\nabla_\partial \hat{u}_a|^2 + \varepsilon^2}} H + \frac{\nu(u)}{\sqrt{|\nabla_\partial \hat{u}_a|^2 + \varepsilon^2}} \mathrm{Tr}_\Sigma k \\
&> \frac{\sqrt{|\nabla u|^2 + \varepsilon^2}}{\sqrt{|\nabla_\partial \hat{u}_a|^2 + \varepsilon^2}} |\mathrm{Tr}_\Sigma k| + \frac{\nu(u)}{\sqrt{|\nabla_\partial \hat{u}_a|^2 + \varepsilon^2}} \mathrm{Tr}_\Sigma k \\
&\geq \frac{|\nu(u)|}{\sqrt{|\nabla_\partial \hat{u}_a|^2 + \varepsilon^2}} |\mathrm{Tr}_\Sigma k| + \frac{\nu(u)}{\sqrt{|\nabla_\partial \hat{u}_a|^2 + \varepsilon^2}} \mathrm{Tr}_\Sigma k \\
&\geq 0.
\end{aligned}
\tag{4.12}
$$

This allows for an application of Lemma 4.1 to conclude that

$$
\int_{\Sigma} \mathcal{H}(e_3', u_a) \, dA \geq \int_{\Sigma} \mathcal{H}(\bar{e}_3, u_a) \, dA.
\tag{4.13}
$$

Inequality (4.13), together with the following estimate, will be used together to show nonnegativity of the quasi-local energy.

Lemma 4.2. *Let (Ω, g, k) be initial data for a compact spacelike hypersurface with boundary $\partial\Omega = \Sigma$ in spacetime $N^{3,1}$, and let $u_a \in C^\infty(\Sigma)$ and $\varepsilon > 0$. If $u \in C^{2,\varsigma}(\Omega)$ is the spacetime harmonic solution of (4.9) with associated level set frame $\{e'_3, e'_4\}$, then*

$$\lim_{\varepsilon \to 0} \int_\Sigma \mathcal{H}(e'_3, u_a) dA + \int_\Omega \left(\frac{1}{2} \frac{|\nabla^2 u + k|\nabla u||^2}{|\nabla u|} + \mu|\nabla u| + J(\nabla u) \right) dV$$

$$\leq 2\pi \int_{\underline{u}}^{\overline{u}} \chi(\Sigma_s) ds, \tag{4.14}$$

where $\Sigma_s = u^{-1}(s)$ and $\overline{u}, \underline{u}$ represent the maximum and minimum values of u. Furthermore, the limit in this expression exists and is finite.

Proof. Let $f_\varepsilon \in C^\infty(\Sigma)$ be such that

$$\cosh f_\varepsilon = \frac{\sqrt{|\nabla u|^2 + \varepsilon^2}}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}, \qquad \sinh f_\varepsilon = \frac{\nu(u)}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}, \tag{4.15}$$

then (2.6) and (2.7) imply

$$\alpha_{e'_3}(\nabla_\partial u_a) = \alpha_\nu(\nabla_\partial u_a) - \nabla_\partial u_a \cdot \nabla_\partial f_\varepsilon = -k(\nabla_\partial u_a, \nu) - \nabla_\partial u_a \cdot \nabla_\partial f_\varepsilon. \tag{4.16}$$

It follows that

$$\mathcal{H}(e'_3, u_a) = \sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2} \langle \vec{H}, e'_3 \rangle + \alpha_{e'_3}(\nabla_\partial u_a)$$

$$= \sqrt{|\nabla u|^2 + \varepsilon^2} H + \nu(u) \text{Tr}_\Sigma k$$

$$- k(\nabla_\partial u_a, \nu) - \nabla_\partial u_a \left(\sinh^{-1} \frac{\nu(u)}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \right) \tag{4.17}$$

$$= \sqrt{|\nabla u|^2 + \varepsilon^2} H + \nu(u) \text{Tr}_\Sigma k$$

$$- k(\nabla_\partial u_a, \nu) - \frac{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}{\sqrt{|\nabla u|^2 + \varepsilon^2}} \nabla_\partial u_a \left(\frac{\nu(u)}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \right).$$

Since the last term in this expression may be estimated by

$$\frac{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}{\sqrt{|\nabla u|^2 + \varepsilon^2}} \left| \frac{\nabla_\partial u_a \cdot \nabla_\partial (\nu(u))}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} - \frac{\nu(u)\nabla_\partial^2 u_a(\nabla_\partial u_a, \nabla_\partial u_a)}{(|\nabla_\partial u_a|^2 + \varepsilon^2)^{3/2}} \right|$$

$$\leq |\nabla^2 u| + |II||\nabla_\partial u_a| + |\nabla_\partial^2 u_a|, \tag{4.18}$$

where II is the second fundamental form of Σ as a submanifold of (Ω, g), we may apply the dominated convergence theorem to conclude that the relevant limit exists, is finite, and satisfies

$$\lim_{\varepsilon \to 0} \int_\Sigma \mathcal{H}(e'_3, u_a) dA = \int_{\partial\Omega} (|\nabla u| H + \nu(u) \text{Tr}_\Sigma k - k(\nabla_\partial u_a, \nu)) dA$$

$$- \int_{\bar{\partial}\Omega} \frac{|\nabla_\partial u_a|}{|\nabla u|} \nabla_\partial u_a \left(\frac{\nu(u)}{|\nabla_\partial u_a|} \right) dA. \tag{4.19}$$

Here $\bar{\partial}\Omega$ denotes the open set of points within $\partial\Omega$ on which $|\nabla_\partial u_a| \neq 0$. The desired result now follows from Proposition 2.1. $\qquad\square$

Proof of Theorem 1.1: Nonnegativity. A direct application of (4.13), Lemmas 4.2, and 5.1 (in the next section) yields

$$
\begin{aligned}
E(\Sigma, \iota, u_a) &= \lim_{\varepsilon \to 0} \frac{1}{8\pi} \int_\Sigma \left(\mathcal{H}_0(\hat{e}_3, u_a) - \mathcal{H}(\bar{e}_3, u_a) \right) dA \\
&\geq \lim_{\varepsilon \to 0} \frac{1}{8\pi} \int_\Sigma \left(\mathcal{H}_0(\hat{e}_3, u_a) - \mathcal{H}(e'_3, u_a) \right) dA \\
&\geq \int_\Omega \left(\frac{1}{2} \frac{|\nabla^2 u + k|\nabla u||^2}{|\nabla u|} + \mu |\nabla u| + J(\nabla u) \right) dV \\
&\quad + \frac{1}{4} \int_{\underline{u}_a}^{\bar{u}_a} \chi(\hat{\Sigma}_s) ds - \frac{1}{4} \int_{\underline{u}}^{\bar{u}} \chi(\Sigma_s) ds,
\end{aligned}
\tag{4.20}
$$

where $\hat{\Sigma}_s$ are the level sets of the relevant null linear function in Minkowski space restricted to the fill-in $\hat{\Omega}$ of $\iota(\Sigma)$. By the maximum principle for (4.9) we find that $\bar{u} = \bar{u}_a$ and $\underline{u} = \underline{u}_a$. Furthermore, analogous arguments to those used in [15, Proposition 5.2] show that the trivial homology hypothesis $H_2(\Omega; \mathbb{Z}) = 0$ guarantees that each component of any regular level Σ_s for u must intersect the boundary $\partial\Omega$. Hence, $\chi(\Sigma_s) \leq n$ where n is the number of components of the s-level set for $u_a = u|_{\partial\Omega}$. On the other hand, the admissibility condition ensures that $\chi(\hat{\Sigma}_s) = n$, so the difference of Euler characteristic integrals in (4.20) is nonnegative. The dominant energy condition $\mu \geq |J|$ then gives $E(\Sigma, \iota, u_a) \geq 0$.

It remains to show that the limit defining the quasi-local energy exists and is finite. From the proof of Lemma 4.1 it follows that

$$
\begin{aligned}
\mathcal{H}(\bar{e}_3, u_a) = &\sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2} + \alpha_{\bar{e}_3}(\nabla_\partial u_a) \\
&+ \nabla_\partial u_a \cdot \nabla_\partial \left(\sinh^{-1} \frac{\Delta_\partial u_a}{|\vec{H}|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \right).
\end{aligned}
\tag{4.21}
$$

The last term in this expression may be estimated by

$$
\begin{aligned}
&\left| \frac{\nabla_\partial u_a(\Delta_\partial u_a) - (\Delta_\partial u_a)\nabla_\partial u_a(\log|\vec{H}|) - (\Delta_\partial u_a)\nabla_\partial^2 u_a(\nabla_\partial u_a, \nabla_\partial u_a)(|\nabla_\partial u_a|^2 + \varepsilon^2)^{-1}}{\sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2}} \right| \\
&\leq |\vec{H}|^{-1}|\nabla_\partial^3 u_a| + |\nabla_\partial u_a||\nabla_\partial \log|\vec{H}|| + |\nabla_\partial^2 u_a|.
\end{aligned}
\tag{4.22}
$$

We may then apply the dominated convergence theorem to find that the limit exists, is finite, and satisfies

$$
\begin{aligned}
\lim_{\varepsilon \to 0} \int_\Sigma \mathcal{H}(\bar{e}_3, u_a) dA = &\int_\Sigma \left(\sqrt{|\vec{H}|^2|\nabla_\partial u_a|^2 + (\Delta_\partial u_a)^2} + \alpha_{\bar{e}_3}(\nabla_\partial u_a) \right) dA \\
&+ \int_{\tilde{\Sigma}} \nabla_\partial u_a \cdot \nabla_\partial \left(\sinh^{-1} \frac{\Delta_\partial u_a}{|\vec{H}||\nabla_\partial u_a|} \right) dA,
\end{aligned}
\tag{4.23}
$$

where $\tilde{\Sigma}$ denotes the open subset of Σ on which $|\nabla_\partial u_a| \neq 0$. Similar arguments hold for the limit of the reference Hamiltonian. $\qquad\square$

Nonnegativity with a disconnected surface. In Remark 1.2, it was stated that nonnegativity of the energy holds in the more general circumstance of a disconnected Σ and without the homology assumption on Ω, when a suitable modification of the admissibility condition is enforced. More precisely, in this setting we may define a pair (ι, u_a) to be *admissible* for $\Sigma \subset N^{3,1}$ if the mean curvature vector of $\iota(\Sigma)$ is outward pointing spacelike, and there exist compact spacelike hypersurfaces $\hat{\Omega} \subset \mathbb{R}^{3,1}$, $\Omega \subset N^{3,1}$ with $\partial\hat{\Omega} = \iota(\Sigma)$, $\partial\Omega = \Sigma$ such that

$$\int_{\underline{u}_a}^{\overline{u}_a} \left(\chi(\hat{\Sigma}_s) - \chi(\Sigma_s) \right) ds \geq 0, \tag{4.24}$$

where the level sets $\hat{\Sigma}_s$, Σ_s are defined as above. Since (4.13), Lemmas 4.2, and 5.1 continue to hold under the more general hypotheses presented here, inequality (4.20) again implies that $E(\Sigma, \iota, u_a) \geq 0$. □

5. Proof of Rigidity

The purpose of the current section is to establish the rigidity statement of Theorem 1.1. We will first compute the surface Hamiltonian for surfaces in Minkowski space. Let $\hat{\Sigma} = \iota(\Sigma) \subset \mathbb{R}^{3,1}$ be a spacelike 2-surface, and assume that (ι, u_a) is admissible. There is then a compact spacelike hypersurface $(\hat{\Omega}, \hat{g}, \hat{k})$ in Minkowski space with boundary $\partial\hat{\Omega} = \hat{\Sigma}$. Let $\{\hat{v}, \hat{\mathbf{n}}\}$ be a normal bundle frame for $\hat{\Sigma}$, where $\hat{v}$ is the outer normal with respect to $\hat{\Omega}$ and $\hat{\mathbf{n}}$ is future directed timelike. For $\varepsilon > 0$ consider the level set frame

$$\hat{e}_3' = \frac{\sqrt{|\nabla\hat{u}|^2 + \varepsilon^2}}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \hat{v} + \frac{\hat{v}(\hat{u})}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \hat{\mathbf{n}},$$

$$\hat{e}_4' = \frac{\hat{v}(\hat{u})}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \hat{v} + \frac{\sqrt{|\nabla\hat{u}|^2 + \varepsilon^2}}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \hat{\mathbf{n}}, \tag{5.1}$$

where $\hat{u} = (-\mathbf{t} + a_i\mathbf{x}^i)|_{\hat{\Omega}}$ is the restriction of the null linear function to the hypersurface, u_a is the restriction of $\hat{u}$ to $\hat{\Sigma}$, with ∇ and ∇_∂ denoting the connections on $\hat{\Omega}$ and $\hat{\Sigma}$ respectively. Note that since the spacetime gradient of the linear function is null it holds that $\hat{\mathbf{n}}(-\mathbf{t} + a_i\mathbf{x}^i) = -|\nabla\hat{u}|$. Moreover, linearity of this function implies that

$$0 = \hat{\Box}(-\mathbf{t} + a_i\mathbf{x}^i) = \left(\hat{\nabla}_{\hat{v}\hat{v}} + \hat{\nabla}_{\hat{\mathbf{n}}\hat{\mathbf{n}}} + \vec{H}_0 + \Delta_\partial \right) (-\mathbf{t} + a_i\mathbf{x}^i)$$

$$= \left(\hat{H}\hat{v} - (\mathrm{Tr}_{\hat{\Sigma}}\hat{k})\hat{\mathbf{n}} + \Delta_\partial \right) (-\mathbf{t} + a_i\mathbf{x}^i) \tag{5.2}$$

$$= \hat{H}\hat{v}(\hat{u}) + (\mathrm{Tr}_{\hat{\Sigma}}\hat{k})|\nabla\hat{u}| + \Delta_\partial u_a,$$

where $\hat{\Box}$ and $\hat{\nabla}$ represent the wave operator and connection on Minkowski space, and $\vec{H}_0 = \hat{H}\hat{v} - (\mathrm{Tr}_{\hat{\Sigma}}\hat{k})\hat{\mathbf{n}}$ is the mean curvature vector of $\hat{\Sigma}$. Since $\hat{\mathbf{n}}$ is future pointing timelike, the null condition for the linear function also gives $|\nabla\hat{u}| > 0$ on $\hat{\Omega}$. Therefore, an examination of the proof for Proposition 2.1 shows that the inequality of (2.3) is in fact an equality. This and (4.19), combined with the observation that $\hat{u}$ has vanishing

spacetime Hessian [4, Section 5] (see also [15, Section 3]), and using that $\mu = |J| = 0$ in Minkowski space, yield a computation of the surface Hamiltonian

$$
\lim_{\varepsilon \to 0} \int_{\hat{\Sigma}} \mathcal{H}_0(\hat{e}'_3, u_a) d\hat{A} = \int_{\hat{\Sigma}} \left(|\nabla \hat{u}| \hat{H} + \hat{v}(\hat{u}) \mathrm{Tr}_{\hat{\Sigma}} \hat{k} - \hat{k}(\nabla_\partial u_a, \hat{v}) \right) d\hat{A}
$$
$$
- \int_{\partial \hat{\Omega}} \frac{|\nabla_\partial u_a|}{|\nabla \hat{u}|} \nabla_\partial u_a \left(\frac{\hat{v}(\hat{u})}{|\nabla_\partial u_a|} \right) d\hat{A} \tag{5.3}
$$
$$
= 2\pi \int_{\underline{u}_a}^{\overline{u}_a} \chi(\hat{\Sigma}_s) ds,
$$

where $\hat{\Sigma}_s$ and $\overline{u}_a$, $\underline{u}_a$ denote the level sets, maximum, and minimum of $\hat{u}$ respectively. The next result shows that the same value is achieved by evaluating at the optimal frame, which is uniquely determined by

$$
\langle \vec{H}_0, \hat{e}_3 \rangle > 0, \qquad \langle \vec{H}_0, \hat{e}_4 \rangle = \frac{-\Delta_\partial u_a}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}. \tag{5.4}
$$

Lemma 5.1. *Let (ι, u_a) be an admissible pair for a spacelike 2-surface Σ, then the reference Hamiltonian satisfies*

$$
\lim_{\varepsilon \to 0} \int_{\Sigma} \mathcal{H}_0(\hat{e}_3, u_a) dA = 2\pi \int_{\underline{u}_a}^{\overline{u}_a} \chi(\hat{\Sigma}_s) ds. \tag{5.5}
$$

Proof. For convenience we will remove extraneous notation, including the subscript on the mean curvature vector and the hat notation from all objects except the optimal frame. Consider the frame associated with Ω, namely

$$
\tilde{e}_3 = \frac{\vec{H}}{|\vec{H}|} = \frac{H}{|\vec{H}|} v - \frac{(\mathrm{Tr}_\Sigma k)}{|\vec{H}|} \mathbf{n} := (\cosh \ell) v - (\sinh \ell) \mathbf{n},
$$
$$
\tilde{e}_4 = -(\sinh \ell) v + (\cosh \ell) \mathbf{n}. \tag{5.6}
$$

This may be used as an intermediary frame to compute the relation between optimal frame $\{\hat{e}_3, \hat{e}_4\}$ and the Ω frame $\{v, \mathbf{n}\}$. In particular, since

$$
\hat{e}_3 = (\cosh f_\varepsilon) \tilde{e}_3 - (\sinh f_\varepsilon) \tilde{e}_4, \qquad \hat{e}_4 = -(\sinh f_\varepsilon) \tilde{e}_3 + (\cosh f_\varepsilon) \tilde{e}_4 \tag{5.7}
$$

where

$$
\cosh f_\varepsilon = \frac{\sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2}}{|\vec{H}| \sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}, \qquad \sinh f_\varepsilon = \frac{\Delta_\partial u_a}{|\vec{H}| \sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}, \tag{5.8}
$$

it follows that

$$
\hat{e}_3 = (\cosh q_\varepsilon) v - (\sinh q_\varepsilon) \mathbf{n}, \qquad \hat{e}_4 = -(\sinh q_\varepsilon) v + (\cosh q_\varepsilon) \mathbf{n} \tag{5.9}
$$

with $q_\varepsilon = f_\varepsilon + \ell$ and

$$\cosh q_\varepsilon = \frac{\sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2}}{|\vec{H}|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \cdot \frac{H}{|\vec{H}|} + \frac{\Delta_\partial u_a}{|\vec{H}|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \cdot \frac{\mathrm{Tr}_\Sigma k}{|\vec{H}|}$$

$$\sinh q_\varepsilon = -\frac{\Delta_\partial u_a}{|\vec{H}|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \cdot \frac{H}{|\vec{H}|} - \frac{\sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2}}{|\vec{H}|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \cdot \frac{\mathrm{Tr}_\Sigma k}{|\vec{H}|}.$$

$$(5.10)$$

The Hamiltonian density function may then be rewritten with respect to the Ω frame utilizing (2.6) and (2.7) to obtain

$$\mathcal{H}_0(\hat{e}_3, u_a) = \sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}\langle \vec{H}, \hat{e}_3 \rangle + \alpha_{\hat{e}_3}(\nabla_\partial u_a)$$
$$= \sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}\,(H \cosh q_\varepsilon + (\mathrm{Tr}_\Sigma k) \sinh q_\varepsilon) - k(\nabla_\partial u_a, \nu) - \nabla_\partial u_a \cdot \nabla_\partial q_\varepsilon.$$

$$(5.11)$$

By employing the fact that the spacetime Hessian of u vanishes on Ω, or rather solving for the Laplacian in (5.2), we find that

$$\Delta_\partial u_a = -H\nu(u) - (\mathrm{Tr}_\Sigma k)|\nabla u|. \tag{5.12}$$

This expression may then be inserted into (5.10) to produce

$$\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}H \cosh q_\varepsilon \to H|\nabla u|,$$
$$\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}(\mathrm{Tr}_\Sigma k) \sinh q_\varepsilon \to (\mathrm{Tr}_\Sigma k)\nu(u), \tag{5.13}$$

as $\varepsilon \to 0$. Moreover, similarly to (4.21) and (4.22) we can estimate

$$|\nabla_\partial u_a \cdot \nabla_\partial q_\varepsilon| \le 2\left(\frac{|H| + |\mathrm{Tr}_\Sigma k|}{|\vec{H}|}\right)(|\nabla_\partial^3 u_a| + |\nabla_\partial^2 u_a| + |\nabla_\partial \log |\vec{H}|||\nabla_\partial u_a|)$$
$$+ 2\left(|\nabla_\partial(|\vec{H}|^{-1}H)| + |\nabla_\partial(|\vec{H}|^{-1}\mathrm{Tr}_\Sigma k)|\right)|\nabla_\partial u_a|. \tag{5.14}$$

Therefore the dominated convergence theorem may be employed to yield

$$\lim_{\varepsilon \to 0}\int_\Sigma (\nabla_\partial u_a \cdot \nabla_\partial q_\varepsilon)\,dA = \int_{\bar\partial\Omega} \frac{|\nabla_\partial u_a|}{|\nabla u|}\nabla_\partial u_a\left(\frac{\nu(u)}{|\nabla_\partial u_a|}\right)dA. \tag{5.15}$$

The desired result now follows from the second equality in (5.3), together with (5.11), (5.13), and (5.15). $\square$

Proof of Theorem 1.1: Rigidity. To establish part *(3)* of this of this theorem, it suffices to observe that if the inclusion map of a spacelike 2-surface in Minkowski space is admissible with some u_a, then the physical and reference Hamiltonian densities are the same, so that the associated energy is zero. It then follows from nonnegativity in part *(1)* that the infimum of energies, and hence the mass of this surface, is zero.

Consider now part *(2)*. Suppose that for some spacelike 2-surface the pair (ι, u_a) is admissible, and $E(\Sigma, \iota, u_a) = 0$ for all **a**. Let (Ω, g, k) be a compact initial data set satisfying the dominant energy condition, with trivial second homology, and boundary $\partial\Omega = \Sigma$. The spacetime harmonic function on Ω with boundary values u_a, will also be

denoted by u_a. We will follow the general strategy of [15, Section 7]. First observe that for each **a** the inequality (4.20) implies

$$\nabla^2 u_a + k|\nabla u_a| = 0, \qquad \mu = |J| = 0 \qquad \text{on } \Omega, \tag{5.16}$$

whenever $|\nabla u_a| \neq 0$. Moreover, the Hopf lemma applied to a maximum point $x_0 \in \Sigma$ of u_a shows that $|\nabla u_a(x_0)| \neq 0$. Let $\gamma \subset \Omega$ be a curve emanating from x_0 parameterized by arclength, and observe that since

$$|\nabla|\nabla u_a|| \leq |\nabla^2 u_a| \leq |k||\nabla u_a| \tag{5.17}$$

holds away from critical points, we have

$$|(\log |\nabla u_a| \circ \gamma)'| \leq |\nabla \log |\nabla u_a|| \circ \gamma \leq C \tag{5.18}$$

away from critical points, for some constant C. By integrating along arbitrary γ, it follows that there is a constant C_1 such that

$$C_1^{-1}|\nabla u_a(x_0)| \leq |\nabla u_a(x)| \leq C_1|\nabla u_a(x_0)|, \qquad x \in \Omega. \tag{5.19}$$

Hence, $|\nabla u_a|$ does not vanish globally for any **a**.

We will now choose three spacetime harmonic functions u_1, u_2, and u_3 on Ω in the following way. Let u_3 be the spacetime harmonic function associated with $\mathbf{a}_3 = (0, 0, 1)$, and consider a point $p \in \Sigma$ at which u_3 achieves its maximum; at this point $|\nabla_\partial u_3(p)| = 0$. In Minkowski space, the isometric image $\iota(\Sigma)$ is tangent at $\iota(p)$ to a hyperplane $-\mathbf{t} + \mathbf{x}^3 = const$ and lies to one side of it. By computing the projection of Minkowski space gradients for the functions $-\mathbf{t} + a_i \mathbf{x}^i$ onto the tangent space $T_{\iota(p)}\iota(\Sigma)$, it is possible to find two choices for **a** such that the corresponding spacetime harmonic functions u_1, u_2 have nonvanishing boundary gradients at p and satisfy $\nabla_\partial u_1(p) \perp \nabla_\partial u_2(p)$. It follows that the vector fields $\nabla u_l, l = 1, 2, 3$ are linearly independent on Ω.

Consider the quantities

$$\varphi = |\nabla u_1| + |\nabla u_2| + |\nabla u_3|, \qquad Y = \nabla u_1 + \nabla u_2 + \nabla u_3, \tag{5.20}$$

and build the stationary spacetime $(\mathbb{R} \times \Omega, \bar{g})$ where

$$\bar{g} = -(\varphi^2 - |Y|^2)dt^2 + 2Y_i dx^i dt + g. \tag{5.21}$$

This is the Kiling development of $(\Omega, g, k, \varphi, Y)$ with Killing initial data lapse-shift (φ, Y) decomposing the Killing vector $\partial_t = \varphi \mathbf{n} + Y$, where **n** is the future pointing unit normal to the constant time slices. As is shown in [15, proof of Theorem 7.3] the initial data for these slices is (Ω, g, k), and as a consequence of the vanishing spacetime Hessians (5.16) the function $\varphi^2 - |Y|^2 =: c^2$ is constant. Furthermore, observe that

$$\varphi^2 - |Y|^2 = 2 \sum_{l<m} (|\nabla u_l||\nabla u_m| - \nabla u_l \cdot \nabla u_m)$$

$$= \sum_{l<m} |\nabla u_l||\nabla u_m| \left| \frac{\nabla u_l}{|\nabla u_l|} - \frac{\nabla u_m}{|\nabla u_m|} \right|^2. \tag{5.22}$$

Since $|\nabla u_l|$ never vanishes, if $c = 0$ then $\nabla u_l \parallel \nabla u_m$ for all l, m. In particular, there exist constants $c_{lm} > 0$ such that at a given point $x_1 \in \Omega$ it holds that

$$\nabla u_l(x_1) - c_{lm} \nabla u_m(x_1) = 0. \tag{5.23}$$

We claim that these relations hold at all points. To see this, note that (5.16) implies

$$|\nabla|\nabla(u_l - c_{lm}u_m)|| \leq |\nabla^2(u_l - c_{lm}u_m)|$$
$$\leq |k|||\nabla u_l| - c_{lm}|\nabla u_m|| \leq |k|||\nabla(u_l - c_{lm}u_m)|. \quad (5.24)$$

Then integrating along curves emanating from x_1 produces

$$C_2^{-1}|\nabla(u_l - c_{lm}u_m)(x_1)| \leq |\nabla(u_l - c_{lm}u_m)(x)| \leq C_2|\nabla(u_l - c_{lm}u_m)(x_1)| \quad (5.25)$$

for some constant $C_2 > 0$ and all $x \in \Omega$, yielding the desired claim. However, (5.23) cannot hold at $p \in \Sigma$ due to the properties of the boundary gradients $\nabla_\partial u_l(p)$. We conclude that the constant $c \neq 0$, and hence

$$\bar{g} = -\left(cdt - c^{-1}Y_i dx^i\right)^2 + (g_{ij} + c^{-2}Y_i Y_j)dx^i dx^j = -d\bar{t}^2 + (g + d\mathbf{u}^2) \quad (5.26)$$

where $\bar{t} = ct - c^{-1}\mathbf{u}$ and $\mathbf{u} = u_1 + u_2 + u_3$.

We will now show that the Killing development is isometric to a portion of Minkowski space. Consider the null vector fields

$$X_l = \nabla\tilde{u}_l + |\nabla\tilde{u}_l|\mathbf{n}, \qquad l = 1, 2, 3, \quad (5.27)$$

where the functions $\tilde{u}_l$ are spacetime harmonic functions on Ω extended trivially in the t-direction to all of $\mathbb{R} \times \Omega$. These functions are chosen in the following way. Let $p_1 \in \Sigma$ be a maximum point for $u_1 + u_2 + u_3$ on Σ. Then in Minkowski space, the isometric image $\iota(\Sigma)$ is tangent at $\iota(p_1)$ to a hyperplane $-\mathbf{t} + b_i\mathbf{x}^i = const$ and lies to one side of it, where $|\mathbf{b}| \leq 1$. We may rotate this to a null hyperplane which is still tangent to $\iota(\Sigma)$ at $\iota(p_1)$, and which is defined by $-\mathbf{t} + \tilde{a}_i\mathbf{x}^i = const$ with $|\tilde{\mathbf{a}}| = 1$. We then set $\tilde{u}_3$ to be the associated spacetime harmonic function, and note that $|\nabla_\partial\tilde{u}_3(p_1)| = 0$. As above, we may also find two additional spacetime harmonic functions $\tilde{u}_1, \tilde{u}_2$ having nonvanishing boundary gradients at p_1 and satisfying $\nabla_\partial u_1(p_1) \perp \nabla_\partial u_2(p_1)$. Since

$$\sum_l d_l X_l + d_4\partial_t = \sum_l d_l\nabla\tilde{u}_l + d_4\nabla(u_1 + u_2 + u_3)$$

$$+ \left(\sum_l d_l|\nabla\tilde{u}_l| + d_4(|\nabla u_1| + |\nabla u_2| + |\nabla u_3|)\right)\mathbf{n} \quad (5.28)$$

for any constants $d_1, \ldots, d_4$, we find that setting this quantity to zero implies $d_1 = d_2 = 0$ by evaluating on $T_{p_1}\Sigma$. Moreover, since X_3 is null and ∂_t is timelike, it follows that $\{X_1, X_2, X_3, \partial_t\}$ is linearly independent. Additionally, note that from (5.16) the functions $\tilde{u}_l$ have vanishing spacetime Hessians, which as in [15, proof of Theorem 7.3] implies that each of these four vector fields is covariantly constant in spacetime. Hence $(\mathbb{R} \times \Omega, \bar{g})$ is flat, and consequently the Riemannian manifold $(\Omega, g + d\mathbf{u}^2)$ is flat. It remains to show that this manifold is isometric to a domain in Euclidean 3-space.

The vanishing quasi-local energy, the admissiblity condition, and inequality (4.20) show that the Euler characteristics agree $\chi(\Sigma_s) = \chi(\hat{\Sigma}_s)$ for regular values s. Here Σ_s is the s-level set of an arbitrary spacetime harmonic function u_a, and $\hat{\Sigma}_s$ is the s-level

set of the corresponding null linear function in $\hat{\Omega} \subset \mathbb{R}^{3,1}$ where $\hat{\Omega}$ is a spacelike fill-in for $\iota(\Omega)$. It follows that Ω is diffeomorphic to $\hat{\Omega}$. Consider now the dual 1-forms to the vector fields X_l on spacetime, and restrict them to the $\bar{t} = 0$ slice to obtain 1-forms $\omega^l, l = 1, 2, 3$ on $(\Omega, g + d\mathbf{u}^2)$. Since the ω^l are covariantly constant, they are closed. We claim that they are also exact. To see this, it will be shown that they integrate to zero on any closed curve $\mathbf{c} \subset \Omega$. Indeed, such a $\mathbf{c}$ is homologous to a closed curve $\tilde{\mathbf{c}}$ inside the $\bar{t} = -c^{-1}\mathbf{u}$ slice, which coincides with the initial data (Ω, g, k). Furthermore, according to (5.27) the restriction of the X_l dual 1-forms to this hypersurface is exact, and hence integrates to zero along $\tilde{\mathbf{c}}$. Hence, there exist functions $\bar{u}^l \in C^{2,\varsigma}(\Omega), l = 1, 2, 3$ such that $\omega^l = d\bar{u}^l$. Since the ω^l are covariantly constant, by applying a Gram-Schmidt procedure we may assume that they are orthonormal, and in addition the Hessians of the $\bar{u}^l$ vanish. Therefore

$$g + d\mathbf{u}^2 = (d\bar{u}^1)^2 + (d\bar{u}^2)^2 + (d\bar{u}^3)^2, \tag{5.29}$$

and $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$ can be used as a system of global coordinates on the $\bar{t} = 0$ slice. It follows that the Killing development of (Ω, g, k) is isometric to a portion of Minkowski space. $\square$

6. Asymptotics of the Energy

In this section we will establish Theorem 1.3, which shows that the quasi-local energy asymptotes to the appropriate ADM quantity along coordinate spheres in an asymptotically flat end. Let (M, g, k) be an asymptotically flat initial data set for the Einstein equations, and consider a coordinate sphere $S_r \subset M$. According to [11, Lemma 2.1] the mean and Gauss curvatures of these spheres satisfy the following expansions

$$H = \frac{2}{r} + O(r^{-1-\tau}), \qquad K = \frac{1}{r^2} + O(r^{-2-\tau}), \tag{6.1}$$

where $\tau > \frac{1}{2}$ is the asymptotic flatness parameter of (1.10). It follows that for sufficiently large r the Gauss curvature is positive, and hence from [24, p. 353] (see also [11, (2.18)]) there exists an isometric embedding into a constant time slice of Minkowski space $\iota_r : S_r \hookrightarrow \mathbb{R}^3 \subset \mathbb{R}^{3,1}$ such that

$$|\nabla_\partial^l (\iota_r - \mathrm{id}_r)| = O(r^{1-\tau-l}), \qquad l = 0, 1, 2, \tag{6.2}$$

where id_r is the identity map on the sphere of radius r in $\mathbb{R}^3$. It follows that the null linear function pullback, used to define the quasi-local energy, may be approximated by a linear combination of asymptotically flat coordinates. In particular, if (x^1, x^2, x^3) are coordinates from (1.10) in the asymptotic end of M and $u_a = \iota_r^*(-\mathbf{t} + a_i \mathbf{x}^i)$ then

$$|\nabla_\partial^l (u_a - u)| = O(r^{1-\tau-l}), \qquad l = 0, 1, 2, \tag{6.3}$$

where $u = a_i x^i$. We will also use the notation $\hat{u} = -\mathbf{t} + a_i \mathbf{x}^i = a_i \mathbf{x}^i$ on $\mathbb{R}^3$.

The pre-limit energy of coordinate spheres may be expressed using (4.7) as

$$E_\varepsilon(S_r, \iota_r, u_a) = \frac{1}{8\pi} \int_{S_r} \left(\sqrt{|\vec{H}_0|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2} - f_0 \Delta_\partial u_a + \alpha_{\frac{\vec{H}_0}{|\vec{H}_0|}}(\nabla_\partial u_a) \right) dA$$

$$- \frac{1}{8\pi} \int_{S_r} \left(\sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2} - f_\varepsilon \Delta_\partial u_a + \alpha_{\frac{\vec{H}}{|\vec{H}|}}(\nabla_\partial u_a) \right) dA, \tag{6.4}$$

where $\vec{H} = Hv - (\mathrm{Tr}_{S_r} k)\mathbf{n}$ and $\vec{H}_0 = \hat{H}\hat{v}$ are the mean curvature vectors of $S_r \subset M$ and $\iota_r(S_r) \subset \mathbb{R}^{3,1}$, and

$$f_\varepsilon = \sinh^{-1}\left(\frac{\Delta_\partial u_a}{|\vec{H}|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}\right), \qquad f_0 = \sinh^{-1}\left(\frac{\Delta_\partial u_a}{|\vec{H}_0|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}\right).$$

(6.5)

In the decomposition of the mean curvature vectors, the normals v and $\hat{v}$ are tangent to M and the constant time slice of Minkowski space respectively. From (1.10), (5.2), and (6.2) we have

$$\Delta_\partial u_a = -\hat{H}\hat{v}(\hat{u}) = -|\vec{H}_0|\hat{v}(\hat{u}) = -|\vec{H}|\hat{v}(\hat{u}) + O(r^{-1-\tau}) = -H\hat{v}(\hat{u}) + O(r^{-1-\tau}),$$

(6.6)

since $\hat{v}(\hat{u}) = O(1)$. In particular, observing that $|\nabla_\partial u_a|^2 + \hat{v}(\hat{u})^2 = 1$ produces

$$|\nabla_\partial u_a|^2 + |\vec{H}|^{-2}(\Delta_\partial u_a)^2 = 1 + O(r^{-\tau}), \qquad |\nabla_\partial u_a|^2 + |\vec{H}_0|^{-2}(\Delta_\partial u_a)^2 = 1.$$

(6.7)

With the help of $|\nabla_\partial u_a| = O(1)$ it follows that

$$\frac{\left(|\vec{H}_0| + |\vec{H}|\right)(|\nabla_\partial u_a|^2 + \varepsilon^2)}{\sqrt{|\vec{H}_0|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2} + \sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2}}$$
$$= |\nabla_\partial u_a|^2 + O(r^{-\tau} + \varepsilon^2),$$

(6.8)

and therefore

$$\int_{S_r}\left(\sqrt{|\vec{H}_0|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2} - \sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2}\right)dA$$
$$= \int_{S_r}\left(|\vec{H}_0| - |\vec{H}|\right)\frac{\left(|\vec{H}_0| + |\vec{H}|\right)(|\nabla_\partial u_a|^2 + \varepsilon^2)}{\sqrt{|\vec{H}_0|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2} + \sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2}}dA \quad (6.9)$$
$$= \int_{S_r}\left(|\vec{H}_0| - |\vec{H}|\right)|\nabla_\partial u_a|^2 dA + O(r^{1-2\tau} + \varepsilon^2 r^{1-\tau}).$$

Consider now the terms involving f_ε and f_0. Notice that by setting $\zeta = |\vec{H}|^{-1}|\vec{H}_0| - 1 = O(r^{-\tau})$ and applying the mean value theorem to the following function of ζ we find

$$\sinh^{-1}\left(\frac{\Delta_\partial u_a}{|\vec{H}|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}\right)$$

$$= \sinh^{-1}\left(\frac{\Delta_\partial u_a}{|\vec{H}_0|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \cdot (1 + \zeta)\right)$$

$$= \sinh^{-1}\left(\frac{\Delta_\partial u_a}{|\vec{H}_0|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}\right) + \left[1 + \left(\frac{(1 + O(r^{-\tau}))\Delta_\partial u_a}{|\vec{H}_0|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}\right)^2\right]^{-1/2} \frac{\Delta_\partial u_a}{|\vec{H}_0|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \cdot \zeta$$

$$= \sinh^{-1}\left(\frac{\Delta_\partial u_a}{|\vec{H}_0|\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}}\right) + \frac{\Delta_\partial u_a}{|\vec{H}_0|}\left(\frac{|\vec{H}_0|}{|\vec{H}|} - 1\right)\left(1 + O(r^{-\tau} + \varepsilon^2)\right),$$

(6.10)

where (6.7) has also been used. Hence (6.6) yields

$$\int_{S_r} (f_\varepsilon - f_0)\Delta_\partial u_a \, dA = \int_{S_r} \hat{v}(\hat{u})^2 \left(|\vec{H}_0| - |\vec{H}|\right) dA + O(r^{1-2\tau} + \varepsilon^2 r^{1-\tau}),$$

(6.11)

and combining this with (6.9) gives

$$\int_{S_r}\left(\sqrt{|\vec{H}_0|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2} - f_0\Delta_\partial u_a\right)dA$$

$$- \int_{S_r}\left(\sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2} - f\Delta_\partial u_a\right)dA$$

$$= \int_{S_r}\left(|\vec{H}_0| - |\vec{H}|\right)dA + O(r^{1-2\tau} + \varepsilon^2 r^{1-\tau})$$

$$= 8\pi\mathcal{E} + o(1) + O(r^{1-2\tau} + \varepsilon^2 r^{1-\tau}).$$

(6.12)

Here $\mathcal{E}$ is the ADM energy, and in the last step we utilize the fact that the Liu-Yau and Brown–York energy have the same large sphere limit [39, proof of Theorem 3.1], together with the convergence of the Brown–York energy to the ADM energy [11, Theorem 1.1].

Lastly, consider the connection 1-forms within (6.4). According (2.6) we find that the reference connection 1-form vanishes, and with (2.7) as well as (6.3) it follows that

*-12pt

$$\alpha_{\frac{\vec{H}}{|\vec{H}|}}(\nabla_\partial u_a) = -k(\nabla_\partial u_a, \nu) - \nabla_\partial u_a \cdot \nabla_\partial h$$

$$= -k(\nabla_\partial u, \nu) + O(r^{-1-2\tau}) - \nabla_\partial u_a \cdot \nabla_\partial h$$

$$= -\left(k - (\mathrm{Tr}_g k)g\right)(\nabla u, \nu) - (\mathrm{Tr}_{S_r} k)\nu(u) + h\Delta_\partial u_a - \mathrm{div}_\partial (h\nabla_\partial u_a) + O(r^{-1-2\tau}),$$

(6.13)

where

$$h = -\sinh^{-1}\left(\frac{\mathrm{Tr}_{S_r} k}{|\vec{H}|}\right) = -\left(1 + O(r^{-\tau})\right)\frac{\mathrm{Tr}_{S_r} k}{|\vec{H}|}$$

(6.14)

with the mean value theorem being used in the last equality. Furthermore (6.2) implies that $\hat{v}(\hat{u}) = v(u) + O(r^{-\tau})$, which together with (6.6) yields

$$\Delta_\partial u_a = -|\vec{H}|v(u) + O(r^{-\tau}), \tag{6.15}$$

and hence

$$- (\mathrm{Tr}_{S_r} k)v(u) + h\Delta_\partial u_a = (\mathrm{Tr}_{S_r} k)v(u) \cdot O(r^{-\tau}) = O(r^{-1-2\tau}). \tag{6.16}$$

Moreover, since $\nabla u = a_i \partial_{x^i} + O(r^{-\tau})$ we obtain

$$\int_{S_r} \left(\alpha_{\frac{\vec{H}_0}{|\vec{H}_0|}} (\nabla_\partial u_a) - \alpha_{\frac{\vec{H}}{|\vec{H}|}} (\nabla_\partial u_a) \right) dA = -8\pi \langle \mathbf{a}, \mathcal{P} \rangle + o(1) + O(r^{1-2\tau}), \tag{6.17}$$

where $\mathcal{P}$ is the ADM linear momentum. By combining (6.4), (6.12), and (6.17) the desired result is achieved

$$\lim_{r \to \infty} E(S_r, \iota_r, u_a) = \lim_{r \to \infty} \lim_{\varepsilon \to 0} E_\varepsilon(S_r, \iota_r, u_a) = \mathcal{E} - \langle \mathbf{a}, \mathcal{P} \rangle. \tag{6.18}$$

7. First Variation of the Energy

The purpose of this section is to derive the Euler-Lagrange equation for the function u_a at a critical point of the energy, under ideal conditions. In particular, it will be assumed that the critical pair (ι, u_a) is admissible and has the following properties. The set of critical points for u_a is sufficiently mild to allow for the interchange of limits as $\varepsilon \to 0$ with integration and variational differentiation, and the Euler characteristics $\chi(\hat{\Sigma}_s)$ of regular level sets within the reference space fill-in $\hat{\Omega}$ take only the value 1. We will use the notation δ to denote the operation of variation.

Consider the quantity

$$E_\varepsilon(\Sigma, \iota, u_a) = \frac{1}{4} \int_{\underline{u}_a}^{\overline{u}_a} \chi(\hat{\Sigma}_s) ds$$
$$- \frac{1}{8\pi} \int_\Sigma \left(\sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2} + \nabla_\partial f_\varepsilon \cdot \nabla_\partial u_a + \alpha_{\frac{\vec{H}}{|\vec{H}|}} (\nabla_\partial u_a) \right) dA, \tag{7.1}$$

where

$$f_\varepsilon = \sinh^{-1} \left(\frac{\Delta_\partial u_a}{|\vec{H}| \sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} \right). \tag{7.2}$$

Observe that (4.21) together with Lemma 5.1 and the coarea formula show

$$E(\Sigma, \iota, u_a) = \lim_{\varepsilon \to 0} E_\varepsilon(\Sigma, \iota, u_a). \tag{7.3}$$

Direct calculations yield

$$\delta \int_{\underline{u}_a}^{\overline{u}_a} \chi(\hat{\Sigma}_s) ds = \delta \overline{u}_a - \delta \underline{u}_a, \tag{7.4}$$

$$\delta \int_\Sigma \sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2} dA$$

$$= \int_\Sigma \frac{|\vec{H}|^2 \nabla_\partial u_a \cdot \nabla_\partial \delta u_a + \Delta_\partial u_a \Delta_\partial \delta u_a}{\sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2}} dA, \tag{7.5}$$

and

$$\delta \int_\Sigma \alpha_{\frac{\vec{H}}{|\vec{H}|}}(\nabla_\partial u_a) dA = \int_\Sigma \alpha_{\frac{\vec{H}}{|\vec{H}|}}(\nabla_\partial \delta u_a) dA. \tag{7.6}$$

Moreover since

$$\delta f_\varepsilon = \frac{\Delta_\partial \delta u_a - (|\nabla u_a|^2 + \varepsilon^2)^{-1}(\Delta_\partial u_a)\nabla_\partial u_a \cdot \nabla_\partial \delta u_a}{\sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2}}, \tag{7.7}$$

we have

$$\delta \int_\Sigma \nabla_\partial f_\varepsilon \cdot \nabla_\partial u_a dA = -\delta \int_\Sigma f_\varepsilon \Delta_\partial u_a dA$$

$$= -\int_\Sigma \left((\delta f_\varepsilon)\Delta_\partial u_a + f_\varepsilon \Delta_\partial \delta u_a \right) dA$$

$$= \int_\Sigma \left(\frac{-\Delta_\partial u_a \Delta_\partial \delta u_a + (|\nabla u_a|^2 + \varepsilon^2)^{-1}(\Delta_\partial u_a)^2 \nabla_\partial u_a \cdot \nabla_\partial \delta u_a}{\sqrt{|\vec{H}|^2(|\nabla_\partial u_a|^2 + \varepsilon^2) + (\Delta_\partial u_a)^2}} \right) dA \tag{7.8}$$

$$+ \int_\Sigma \nabla_\partial f_\varepsilon \cdot \nabla_\partial \delta u_a dA.$$

It follows that

$$\delta E_\varepsilon(\Sigma, \iota, u_a) = \frac{1}{4}(\delta \overline{u}_a - \delta \underline{u}_a)$$

$$- \frac{1}{8\pi} \int_\Sigma \left(|\vec{H}|(\cosh f_\varepsilon) \frac{\nabla_\partial u_a \cdot \nabla_\partial \delta u_a}{\sqrt{|\nabla_\partial u_a|^2 + \varepsilon^2}} + \nabla_\partial f_\varepsilon \cdot \nabla_\partial \delta u_a + \alpha_{\frac{\vec{H}}{|\vec{H}|}}(\nabla_\partial \delta u_a) \right) dA. \tag{7.9}$$

Under the ideal conditions mentioned at the beginning of this section, we may interchange limit and variational derivative, and apply the dominated convergence theorem as in (4.23) to obtain

$$\delta E(\Sigma, \iota, u_a) = \lim_{\varepsilon \to 0} \delta E_\varepsilon(\Sigma, \iota, u_a)$$

$$= \frac{1}{4}(\delta \overline{u}_a - \delta \underline{u}_a) \tag{7.10}$$

$$- \frac{1}{8\pi} \int_\Sigma \left(|\vec{H}|(\cosh f) \frac{\nabla_\partial u_a \cdot \nabla_\partial \delta u_a}{|\nabla_\partial u_a|} + \nabla_\partial f \cdot \nabla_\partial \delta u_a + \alpha_{\frac{\vec{H}}{|\vec{H}|}}(\nabla_\partial \delta u_a) \right) dA,$$

where

$$f = \sinh^{-1}\left(\frac{\Delta_\partial u_a}{|\vec{H}||\nabla_\partial u_a|} \right). \tag{7.11}$$

If the variations δu_a are plentiful enough so as to include all smooth functions, then we find that the critical isometric embedding pair gives rise to a weak solution of the 4th order equation

$$\text{div}_\sigma \left(|\vec{H}|(\cosh f)\frac{\nabla_\partial u_a}{|\nabla_\partial u_a|} + \nabla_\partial f + V \right) = 2\pi(\boldsymbol{\delta}_- - \boldsymbol{\delta}_+), \qquad (7.12)$$

in which V is the dual vector field to the connection 1-form $\alpha_{\frac{\vec{H}}{|\vec{H}|}}$ and $\boldsymbol{\delta}_\pm$ are Dirac delta distributions at the max and min points.

References

1. Alaee, A., Khuri, M., Hung, P.-K.: The positive energy theorem for asymptotically hyperboloidal initial data sets with toroidal infinity and related rigidity results. Commun. Math. Phys. **396**(2), 451–480 (2022)
2. Bartnik, R.: New defnition of quasi-local mass. Phys. Rev. Lett. **62**(20), 2346–2348 (1989)
3. Booth, I., Mann, R.: Moving observers, nonorthogonal boundaries, and quasilocal energy. Phys. Rev. D **59**(6), 064021 (1999)
4. Bray, H., Hirsch, S., Kazaras, D., Khuri, M., Zhang, Y.: Spacetime harmonic functions and applications to mass. In: Gromov, M., Lawson, H. (eds.), Perspectives Scalar Curvature, vol. 2. World Scientific Publishing, pp. 593–639 (2023)
5. Bray, H., Kazaras, D., Khuri, M., Stern, D.: Harmonic functions and the mass of 3-dimensional asymptotically flat Riemannian manifolds. J. Geom. Anal **32**(6), 184 (2022)
6. Brown, J., York, J.: Quasilocal energy and conserved charges derived from the gravitational action. Phys. Rev. D. **47**(4), 1407–1419 (1993)
7. Chen, P.-N., Hung, P.-K., Wang, M.-T., Yau, S.-T.: The rest mass of an asymptotically anti-de Sitter spacetime. Ann. Henri Poincaré **18**(5), 1493–1518 (2017)
8. Dougan, A., Mason, L.: Quasilocal mass constructions with positive energy. Phys. Rev. Lett. **67**(16), 2119–2122 (1991)
9. Eichmair, M., Huang, L.-H., Lee, D., Schoen, R.: The spacetime positive mass theorem in dimensions less than eight. J. Eur. Math. Soc. (JEMS) **18**(1), 83–121 (2016)
10. Epp, R.: Angular momentum and an invariant quasilocal energy in general relativity. Phys. Rev. D **62**(12), 124108 (2000)
11. Fan, X.-Q., Shi, Y., Tam, L.-F.: Large-sphere and small-sphere limits of the Brown-York mass. Commun. Anal. Geom. **17**(1), 37–72 (2009)
12. Gromov, M.: Partial Differential Relations, Ergeb. Math. Grenzgeb. **9** [Results in Mathematics and Related Areas] Springer-Verlag, Berlin, (1986), 363 pp
13. Hawking, S.: Gravitational radiation in an expanding universe. J. Math. Phys. **9**(4), 598–604 (1968)
14. Hawking, S., Horowitz, G.: The gravitational Hamiltonian, action, entropy and surface terms. Class. Quantum Grav. **13**(6), 1487–1498 (1996)
15. Hirsch, S., Kazaras, D., Khuri, M.: Spacetime harmonic functions and the mass of 3-dimensional asymptotically flat initial data for the Einstein equations. J. Differ. Geom. **122**(2), 223–258 (2022)
16. Hirsch, S., Kazaras, D., Khuri, M., Zhang, Y.: Spectral Torical band inequalities and generalizations of the Schoen-Yau black hole existence theorem. Int. Math. Res. Not. IMRN (2024), no. 4, 3139–3175. arXiv:2301.08270
17. Hirsch, S., Kazaras, D., Khuri, M., Zhang, Y.: Rigid comparison geometry for Riemannian bands and open incomplete manifolds. arXiv:2209.12857 (2022)
18. Kijowski, J.: A simple derivation of canonical structure and quasi-local Hamiltonians in general relativity. Gen. Relativ. Gravit. **29**(3), 307–343 (1997)
19. Lee, D.: Geometric Relativity. Graduate Studies in Mathematics, Vol. 201 (2019)

20. Liu, C.-C., Yau, S.-T.: Positivity of quasi-local mass II. J. Am. Math. Soc. **19**(1), 181–204 (2006)
21. Lott, J.: A spinorial quasi-local mass. arXiv:2212.09619 (2023)
22. Ludvigsen, M., Vickers, J.: Momentum, angular momentum and their quasi-local null surface extensions. J. Phys. A **16**, 1155–1168 (1983)
23. Montiel, S.: Compact approach to the positivity of Brown–York mass and rigidity of manifolds with mean-convex boundaries in flat and spherical contexts. arXiv:2209.07762 (2022)
24. Nirenberg, L.: The Weyl and Minkowski problems in differential geometry in the large. Commun. Pure Appl. Math. **6**, 337–394 (1953)
25. Penrose, R.: Some unsolved problems in classical general relativity. Ann. Math. Stud. **102**, 631–668 (1982)
26. Penrose, R.: Quasi-local mass and angular momentum in general relativity. Proc. R. Soc. Lond. Ser. A **381**(1780), 53–63 (1982)
27. Pogorelov, A.: Regularity of a convex surface with given Gaussian curvature. (Russian) Mat. Sbornik N.S. **31**(73), 88–103 (1952)
28. Pogorelov, A.: Some results on surface theory in the large. Adv. Math. **1**(2), 191–264 (1964)
29. Pogorelov, A.: The extrinsic geometry of convex surfaces, (Russian) Izdat. Nauka, Moscow, 1969, 759 pp. (English) Translations of Mathematical Monographs, vol. 35. American Mathematical Society, Providence, RI (1973)
30. Schoen, R.: Talk at Conference on Geometric Analysis and General Relativity, University of Tokyo, November (2019)
31. Schoen, R., Yau, S.-T.: Proof of the positive mass theorem II. Commun. Math. Phys. **79**, 231–260 (1981)
32. Shi, Y., Tam, L.-F.: Positive mass theorem and the boundary behaviors of compact manifolds with non-negative scalar curvature. J. Differ. Geom. **62**(1), 79–125 (2002)
33. Stern, D.: Scalar curvature and harmonic maps to S^1. J. Differ. Geom. **122**(2), 259–269 (2022)
34. Stern, D.: Level set methods in the study of scalar curvature. In: Gromov, M., Lawson, H.(eds.), Perspectives in Scalar Curvature, Vol. 1. World Scientific Publishing , pp. 577–610 (2023)
35. Szabados, L.: Quasi-local energy-momentum and angular momentum in general relativity. Living Rev. Relat. **12**, 1–163 (2009)
36. Tsang, T.-Y.: On a spacetime positive mass theorem with corners (2022). arXiv:2109.11070
37. Wang, M.-T., Yau, S.-T.: Quasi-local mass in general relativity. Phys. Rev. Lett. **102**(2), 021101 (2009)
38. Wang, M.-T., Yau, S.-T.: Isometric embeddings into the Minkowski space and new quasi-local mass. Commun. Math. Phys. **288**(3), 919–942 (2009)
39. Wang, M.-T., Yau, S.-T.: Limit of quasilocal mass at spatial infinity. Commun. Math. Phys. **296**(1), 271–283 (2010)
40. Witten, E.: A simple proof of the positive energy theorem. Commun. Math. Phys. **80**(3), 381–402 (1981)
41. Zhang, X.: A new quasi-local mass and positivity. Acta Math. Sin. (Engl. Ser.) **24**(6), 881–890 (2008)