A JANG EQUATION APPROACH TO THE PENROSE INEQUALITY FOR PERTURBATIONS OF SCHWARZSCHILD SPACETIME

Hollis Williams

School of Engineering, University of Warwick, Coventry CV4 7AL, UK

We show that solutions exist to a simplified version of the system of equations obtained by coupling Bray's conformal flow of metrics and the generalised Jang equation. This would establish the Penrose inequality for a class of conformally flat perturbations of Schwarzschild spacetime, provided that one can additionally prove that existence of a suitable approximate solution to the Jang equation implies the Penrose inequality.

1. Introduction

A Cauchy data for the Einstein field equations is a triple (M, g, k), where M is a Riemannian 3-manifold, g is a Riemannian metric, and k is a symmetric 2-tensor field, which together satisfies the equations of constraint

$$16\pi\mu \;=\; R + (\operatorname{Tr} k)^2 - |k|^2 \,, \tag{1a}$$

$$8\pi J \;=\; \operatorname{div}(k + (\operatorname{Tr} k)g) \,, \tag{1b}$$

where μ is the energy density of the matter fields, J is the momentum density, and R is the scalar curvature of g. We assume the dominant energy condition throughout, $\mu \geq |J|$. It is important in the rigorous PDE formulation of the Einstein equations that one satisfies these equations of constraint, as technically speaking any metric tensor would satisfy the field equations otherwise.

Given the above hypotheses on the initial data, an open question in general relativity is the Penrose conjecture, which states the following bound (known as the Penrose inequality) on the ADM energy for a black hole spacetime:

$$E_{\text{ADM}} \geq \sqrt{\frac{A_{\min}}{16\pi}} \,, \tag{2}$$

where $A_{\min}$ is the minimum area required to enclose the boundary ∂M, which consists of an outermost apparent horizon [1]. Equality holds for (2) if and only if the Cauchy triple (M, g, k) is for a spacelike hypersurface in Schwarzschild spacetime with outer-minimizing boundary. An important special case for the Penrose inequality occurs when the spacelike hypersurface has a vanishing trace of the second fundamental form: this is known as the Riemannian Penrose inequality. There are various survey articles which the reader can consult on the Penrose inequality and other geometric inequalities for quasi-local masses in general relativity (see, for example, [2, 3]).

The Riemannian Penrose inequality was proved using flows from geometric analysis (the inverse mean curvature flow in the case of the Huisken–Ilmanen proof and the conformal flow of metrics in the case of Bray's proof) [4, 5]. Given this, one might wonder if one could use the Jang equation to reduce back to the special case of a maximal spacelike hypersurface, as was done in the original proof of Schoen and Yau [6, 7]. Unfortunately, the Jang equation is not suitable for studying the Penrose inequality. Roughly speaking, the reason for this is that the Jang equation is so optimized towards proving the positive mass theorem that it cannot be used to prove the original Penrose inequality. In [8], Bray and Khuri proposed a generalised version of the Jang equation, where as before one searches for a hypersurface Σ given by a graph $t = f(x)$ in the product space $M \times \mathbb{R}$, but now the product metric is a warped product metric $\bar{g} = -\phi^2 \mathrm{d}f^2 + g$, where ϕ is a nonnegative function defined on M.

There are many possible choices for ϕ and we use the one outlined by Han and Khuri (more information on this will follow) [9]. This choice for ϕ gives a warping factor which is strictly positive. As in the classical case, the Jang surface Σ should satisfy a PDE of the form of

$$H_\Sigma = \mathrm{Tr}_\Sigma K \,, \tag{3}$$

where H_Σ is the mean curvature of the hypersurface and $\mathrm{Tr}_\Sigma K$ is the trace of the second fundamental form extended to the product $M \times \mathbb{R}$ and taken over Σ. However, unlike the classical case studied by Schoen and Yau, the extension of k to $M \times \mathbb{R}$ is nontrivial, and is in fact given explicitly by

$$
\begin{aligned}
K(\partial_{x^i}, \partial_{x^j}) &= K(\partial_{x^j}, \partial_{x^i}) = k(\partial_{x^i}, \partial_{x^j}) && \text{for} \quad 1 \le i, j \le 3 \,, && \text{(4a)} \\
K(\partial_{x^i}, \partial_t) &= K(\partial_t, \partial_{x^i}) = 0 && \text{for} \quad 1 \le i \le 3 \,, && \text{(4b)} \\
K(\partial_t, \partial_t) &= \frac{\phi^2 g(\nabla f, \nabla \phi)}{\sqrt{1 + \phi^2 |\nabla f|^2}} \,, &&&& \text{(4c)}
\end{aligned}
$$

where x^i are local coordinates on M. When k is extended in this way, the

generalised Jang equation in local coordinates is

$$\left(g^{ij} - \frac{\phi^2 f^i f^j}{1 + \phi^2 |\nabla f|^2} \right) \left(\frac{\phi \nabla_{ij} f + \phi_i f_j + \phi_j f_i}{\sqrt{1 + \phi^2 |\nabla f|^2}} - k_{ij} \right) = 0 \,. \tag{5}$$

One can confirm that this reduces back to the Jang equation studied by Schoen and Yau in the case where $\phi = 1$. The main technical difficulty with using this equation to prove the Penrose inequality is that one must couple it to one of the flows mentioned previously, which produces a system of equations to solve rather than one single equation.

Proving the Penrose inequality in full generality is likely to be extremely difficult, so recent results in the literature have focused on proving important special cases. We investigate here the possibility of approximate solutions to the Jang/CFM coupling given a number of other assumptions on the perturbations, where the solutions correspond to conformally flat perturbations of Schwarzschild spacetime. To be specific, the main result of our article is as follows:

Theorem 1.1 *An approximate solution exists to the system of PDEs furnished by the coupling of Bray's conformal flow of metrics to the generalised Jang metric provided that the 3-metric g is the metric corresponding to small conformally flat linear maximal perturbations of the standard Schwarzschild slice. Furthermore, we assume the following conditions on the perturbation and the warping factor ϕ:*

$$2c(e_1)\frac{\partial \sigma}{\partial x} \;=\; 2c(e_2)\frac{\partial \sigma}{\partial y} = -c(e_3)\frac{\partial \sigma}{\partial z} \,, \tag{6a}$$

$$\frac{\partial \sigma}{\partial x_0} \;\leq\; \left(1 + \frac{M}{2r} \right)^{-3/2} - \left(1 + \frac{M}{2r} \right)^{-11/2} \sigma \,, \tag{6b}$$

$$F_1 \sigma \;=\; 0 \,, \qquad F_2 \frac{\partial \sigma}{\partial t} = 0 \,, \qquad \frac{\partial \sigma}{\partial x_k} \frac{\partial f}{\partial x_k} = 0 \,, \qquad \frac{\partial \phi}{\partial x_k} \frac{\partial f}{\partial x_k} = 0 \,, \tag{6c}$$

for F_1 a function which decays faster than e^{-2t} and F_2 a function which decays as e^{-t}. Besides this, we assume that any quantities of the order of $\mathcal{O}(1/r^{10})$ can be neglected and that the size of the spinor ψ_t^- is equal to the size of the spinor $\tilde{\psi}_t$, where ψ_t^- and $\tilde{\psi}_t$ are defined in the text.

Our reason for neglecting terms of $\mathcal{O}(1/r^{10})$ is that such terms decay with such speed that they are physically inadmissible, and so terms of this order can be dropped after we have plugged in the solution which we will find.

Regarding the previous results obtained with regard to the Penrose inequality for perturbed metrics, we will mention that Alexakis used a perturbation argument around a spherically symmetric null hypersurface in Schwarzschild spacetime to prove the null Penrose inequality for perturbations of Schwarzschild exterior spacetime [10]. Note that this Penrose inequality differs from ours, as it is expressed in terms of the Bondi energy rather than the ADM energy. Kopiński and Tafel adapted the conformal method for solving constraints to prove the Penrose inequality for Schwarzschild initial data under an addition of the axially symmetric traceless exterior curvature, assuming the initial metric to be conformally flat [11]. This proof was later extended by the same authors to nonmaximal perturbations of Schwarzschild initial data [12]. Finally, Roszkowski and Malec used different techniques to prove the Penrose inequality for Schwarzschild spacetime with axially symmetric linear perturbations [13]. Note that there is a missing step in our argument which means that we have not shown that the Penrose inequality is implied although we show existence of solutions (we will discuss this in more detail shortly).

2. Jang coupling to the conformal flow of metrics

We will begin by quoting the system of equations which we propose to use to prove the inequality. It was shown by Han and Khuri that this is one possible system for which an existence result would imply the Penrose inequality. In the interest of keeping the article self-contained, we will give some exposition of how this result is derived [8, 9]. The first equation in the system is the generalised Jang equation (equation (5)). The conformal flow of metrics which we need is given by $g_t = u_t^4 g$, where

$$\frac{\mathrm{d}}{\mathrm{d}t} u_t = v_t u_t \,, \tag{7a}$$

$$\Delta_{g_t} v_t = 0 \quad \text{on} \quad M_t \,, \qquad v_t|_{\partial M_t} = 0 \,, \tag{7b}$$

$$v_t(x) \to -1 \text{ as } |x| \to \infty \,, \qquad u_t(x) \to \mathrm{e}^{-t} \text{ as } |x| \to \infty \,, \tag{7c}$$

where M_t is the region outside of ∂M_t and $|x| = \sqrt{x^2 + y^2 + z^2}$. This flow was originally introduced by Bray to prove the Riemannian Penrose inequality and adapted by Han and Khuri so that it can be applied to the Penrose inequality for general initial data sets [8, 9].

Since the conformal flow of metrics was used to prove the Riemannian Penrose inequality and the Jang surface $(\Sigma, \bar{g})$ can be viewed as a deformation of the initial data set (M, g, k) with weakly nonnegative scalar curvature, the intuitive idea is to apply the same techniques which were used to prove the Riemannian case, but to apply them in the Jang surface. This

leads to a coupling between the generalised Jang equation and either the inverse mean curvature flow or the conformal flow of metrics, where the coupling is implemented via a particular choice for the warping factor ϕ which appears in the Jang metric. The technical obstacle to doing this is that the choice which we arrive at for the warping factor could be very problematic and difficult to work with (the warping factor could be negative, for example, or lack regularity). In fact, this is the main problem with trying to prove the Penrose inequality by coupling the generalised Jang equation to the inverse mean curvature flow, since the warping factor can vanish when the weak inverse mean curvature flow undergoes discontinuous jumps, which causes the generalised Jang equation to become degenerate [9]. In [9], it was shown that the expression for ϕ which one obtains by coupling the Jang equation to the conformal flow of metrics yields an improved system of equations, since ϕ is now positive on the interior of M and has better regularity properties.

In the Jang coupling to the conformal flow of metrics, the second equation in the system is the equation for the warping factor ϕ

$$\phi = 2 \int\limits_0^\infty \chi_t \left(\left(w_t^- \right)^2 |\psi_t^-|^2 + \left(w_t^+ \right)^2 |\psi_t^+|^2 \right) u_t^2 \, \mathrm{d}t \,, \tag{8}$$

where χ_t is an indicator function on Σ_t, the region outside the outermost minimal surface $\partial \Sigma_t$ (Σ_0 being a solution of the Jang equation whose metric is then evolved under the conformal flow of metrics) and the other quantities are defined shortly. This choice for the warping factor might seem mysterious, but follows naturally when one performs the conformal flow of metrics on the Jang surface [9]. If we take a conformal flow ($g_t = u_t^{-4}\bar{g}$, Σ_t), then the scalar curvature evolves by

$$R_{g_t} = u_t^{-4} \bar{R} \,. \tag{9}$$

Since the ADM energies of (M, g, k) and $(\Sigma, \bar{g})$ must agree, one obtains by a computation the following inequality:

$$E_{\mathrm{ADM}} - \sqrt{\frac{|\partial \Sigma_0|}{16\pi}} \geq \int\limits_0^\infty \tilde{E}_{\mathrm{ADM}}(t) \, \mathrm{d}t \,, \tag{10}$$

where $\tilde{M}_t = M_t^- \cup M_t^+$, the doubled manifold reflected across the outermost minimal surface ∂M_t in (M, g_t). The metric on $\tilde{M}_t$ is piecewise defined $\tilde{g}_t = g_t^- \cup g_t^+$, where $g_t^\pm = (w_t^\pm)^4 g_t$. $\tilde{E}_{\mathrm{ADM}}$ denotes the total ADM energy of $(\tilde{\Sigma}_t, g_t^- \cup g_t^+)$ and is given by

$$4\pi\tilde{E}_{\text{ADM}}(t) = \int\limits_{\Sigma_t} \left[\left(w_t^- u_t\right)^6 \left|\nabla^-\psi_t^-\right|^2 + \left(w_t^+ u_t\right)^6 \left|\nabla^+\psi_t^+\right|^2 \right] d\omega_{\bar{g}}$$

$$+ \frac{1}{4} \int\limits_{\Sigma_t} \left[\left(w_t^-\right)^2 \left|\psi_t^-\right|^2 + \left(w_t^+\right)^2 \left|\psi_t^+\right|^2 \right] u_t^2 \bar{R}\, d\omega_{\bar{g}}, \qquad (11)$$

where $\tilde{\psi}_t$ is a harmonic spinor defined on $w_t^\pm = (1 \pm v_t)/2$ and $\psi_t^\pm$ is the corresponding restriction of $\tilde{\psi}_t$.

The central issue is that one needs positive scalar curvature for the PDE approach to work when trying to prove the Penrose inequality. One cannot conclude that $\tilde{E}_{\text{ADM}} \geq 0$ because the scalar curvature for the Jang metric does not need to be nonnegative. In [8], it is shown that the scalar curvature $\bar{R}$ is given by

$$\bar{R} = 16\pi(\mu - J(w)) + |h - K|_{\Sigma}|^2 + 2|q|^2 - 2\phi^{-1}\overline{\text{div}}(\phi q), \qquad (12)$$

where h is the second fundamental form of Σ, $\overline{\text{div}}$ is the divergence operation associated with the induced metric for the graph $\bar{g} = g + \phi^2 df^2$, and q and w are 1-forms defined as

$$w_i = \frac{\phi f_i}{\sqrt{1 + \phi^2|\nabla f|^2}}, \qquad (13a)$$

$$q_i = \frac{\phi f^j}{\sqrt{1 + \phi^2|\nabla f|^2}} \left(h_{ij} - (K|_{\Sigma})_{ij}\right). \qquad (13b)$$

The obstruction to nonnegative scalar curvature is the divergence term, which can be integrated away if we prescribe ϕ to take the form given above (the dominant energy condition implies that the right-hand side of (12) is manifestly nonnegative except for the troublesome divergence term). A computation with Fubini's theorem, the boundary condition for ϕ, and the formula for $\bar{R}$ shows that

$$\int\limits_0^\infty \int\limits_{\Sigma_t} \left[\left(w_t^-\right)^2 \left|\psi_t^-\right|^2 + \left(w_t^+\right)^2 \left|\psi_t^+\right|^2 \right] u_t^2 \bar{R}\, d\omega_{\bar{g}}\, dt \geq \int\limits_{\partial\Sigma_0} \phi|q|^2\, d\omega_{\bar{g}}, \qquad (14)$$

where it has been assumed that the 1-form q is uniformly bounded and decays very quickly as one goes out to spatial infinity. This equation can be combined with the formula for $\bar{R}$ and equations (10) and (11) to obtain the Penrose inequality (using the fact that $\bar{g}$ measures areas to be at least as large as those measured by g).

The third and final equation in the system is the Dirac equation for harmonic spinors

$$\mathcal{D}\tilde{\psi}_t = 0, \qquad (15)$$

where the Dirac operator is given by

$$\mathcal{D} = \sum_{i=1}^{3} c(e_i) \nabla_{e_i} \,. \tag{16}$$

The use of harmonic spinors is not essential and is mainly motivated by the fact that all the manifolds which interest us in this application are automatically spin manifolds, but one can derive a similar alternative expression for ϕ which does not use spinors [9, 14]. The appearance of the Dirac equation is simply the statement of existence of a harmonic spinor (which is assumed by Han and Khuri): these spinors are then used to obtain equation (8). As we have said, the idea of Bray and Khuri is to couple the conformal flow of metrics to the Jang equation by performing the flow on the Jang surface itself, so the reader must bear in mind that even though we are using a notation such as g_t, depending on the context this might be $u_t^4 \bar{g}$, not $u_t^4 g$.

The spin connection is defined by

$$\nabla_{e_i} = e_i + \frac{1}{4} \sum_{j,l=1}^{3} \Gamma_{ij}^{l} c(e_j) c(e_l) \,, \tag{17}$$

where Γ_{ij}^{l} are the Levi-Civita connection cofficients for $\bar{g}$. e_1, e_2, and e_3 are orthonormal frame fields for $\bar{g}$ and $c : Cl(T\tilde{\Sigma}_t) \to \mathrm{End}(C)$ is the standard representaton of the Clifford algebra on the spinor bundle (C) such that

$$c(X)c(Y) + c(Y)c(X) = -2g(X, Y) \,. \tag{18}$$

The boundary condition on the harmonic spinor is the standard one that it should converge to a constant spinor of unit norm at spatial infinity and it is also required that $\phi(x) \to 1$ as x goes to spatial infinity. Existence of a harmonic spinor with the necessary rate of decay was assumed in Witten's spin geometric proof of the positive mass theorem and demonstrated by Parker and Taubes using estimates of Nirenberg, Walker, and Cantor [15–18]. Furthermore, assuming that as we have done that ϕ is positive on the interior and that it vanishes at ∂M_0 implies the following boundary condition for the generalised Jang equation:

$$\tilde{H}_{\partial \Sigma_0} = 0 \,, \tag{19}$$

where $\tilde{H}_{\partial \Sigma_0}$ is the mean curvature of the Jang surface Σ which encloses the region outside the outermost minimal area enclosure of M^3. In our case, we will take the perturbations which we use to be maximal such that this boundary condition is satisfied. We will require a number of assumptions on

the derivatives of the perturbation. In general, any term which decays very quickly as $\mathcal{O}(1/r^{10})$ or faster as one goes to spatial infinity will be assumed to vanish.

For a conformal perturbation given by a single function σ, we will further require that

$$\frac{\partial \sigma}{\partial x_0} \leq \left(1 + \frac{M}{2r}\right)^{-3/2} - \left(1 + \frac{M}{2r}\right)^{-11/2} \sigma\,. \qquad (20)$$

This condition seems restrictive, but testing with some numerical examples shows that it holds generically as long as the perturbation σ is small. We will also require a proportionality relation of the below type or similar:

$$2c(e_1)\frac{\partial \sigma}{\partial x} = 2c(e_2)\frac{\partial \sigma}{\partial y} = -c(e_3)\frac{\partial \sigma}{\partial z}\,. \qquad (21)$$

The assumptions which we use seem quite strong, but are typical of those which other authors have used to make progress in the literature. At the time of writing, very specific restrictions are necessary until a wider body of theory is introduced to deal with a system such as the one which we are considering. For example, the generalised Jang equation was used to prove the Penrose inequality for spherically symmetric spacetimes only by prescribing the warping factor ϕ to have an extremely specific form, which limits the proof of Bray and Khuri to cases where ϕ is prescribed by hand to take this form [19]. Besides this, in the spherically symmetric setting, the generalised Jang equation decouples from the other equations in the system. This is no longer the case in our setting where we must consider the full coupled system, which is even more unamenable to analysis.

3. Existence of solutions for perturbations of Schwarzschild spacetime

In general relativity (and mathematical physics more generally), there are two main ways to obtain information about solutions of PDEs. One is to do qualitative analysis and prove that the solutions have desirable properties (usually with *a priori* estimates) [20]. As far as we know, attempts to use this approach with the system we have given to prove the Penrose inequality for arbitrary perturbations of Schwarzschild spacetime have so far failed. The second approach is to use some kind of approximation and set certain quantities to zero so that one has a simplified system which approximates the original system. This is the approach which we will use here, but note that existence of approximate solutions does not necessarily prove the Penrose inequality because one has to check that nonnegative scalar curvature is not lost for the specific class of perturbations being considered.

Our first task is to assume certain perturbations of Schwarzschild space-time and see if the equations can be simplified. Throughout this article, we will always assume linear perturbations and neglect any higher-order terms which appear. We are working in adapted coordinates with ∂_0 orthogonal to the spacelike hypersurface, so the induced metric is simply the perturbed Schwarzschild metric $g_{\mu\nu}$ restricted to the spatial coordinates

$$g_{ij} = \left(1 + \frac{M}{2r}\right)^4 \delta_{ij} + h_{ij} \,, \tag{22}$$

where the size of the components of h is small in a suitable norm. Note that δ_{ij} is the Euclidean 3-metric and that we are using the conformal description for the Schwarzschild metric.

We will also assume that the perturbation is conformally flat. In this case, the induced metric is

$$g_{ij} = \left(\left(1 + \frac{M}{2r}\right)^4 + \sigma\right) \delta_{ij} \,. \tag{23}$$

The extrinsic curvature k_{ij} follows from a computation, where k is written with respect to the future unit normal vector n

$$k_{ij} = \langle n, \nabla_i \partial_j \rangle = n^\alpha \Gamma_{\alpha ij} \,. \tag{24}$$

The future-pointing unit normal vector in this case is

$$n = \frac{1}{\sqrt{-g^{00}}} g^{0\alpha} \partial_\alpha \,. \tag{25}$$

It follows that

$$k_{ij} = \frac{\Gamma^0_{ij}}{\sqrt{-g^{00}}} \,. \tag{26}$$

We require two standard results on perturbed metrics from perturbation theory in general relativity

$$g^{ab} = g^{ab}_{\text{SC}} - h^{ab} \,, \tag{27a}$$

$$\Gamma^c_{ab} = \frac{1}{2} g^{cd}_{\text{SC}} (\partial_a h_{bd} + \partial_b h_{ad} - \partial_d h_{ab}) \,, \tag{27b}$$

where g_{SC} is the unperturbed metric. From the second formula, one has

$$\Gamma^0_{ij} = -\frac{1}{2} g^{00}_{\text{SC}} \frac{\partial h_{ij}}{\partial x_0} \,, \tag{28}$$

when $i = j$, and $\Gamma_{ij}^0 = 0$ otherwise. Plugging this back in, we get

$$k_{ij} = \frac{-g_{\text{SC}}^{00}}{2\sqrt{-g_{\text{SC}}^{00} + h^{00}}} \frac{\partial h_{ij}}{\partial t} = \frac{1}{2} \left(\frac{1 - \frac{M}{2r}}{1 + \frac{M}{2r}} \right)^{-2} \frac{\partial h_{ij}}{\partial x_0}, \qquad (29)$$

where we have used the fact that we are working to linear order. Note that if the perturbation were not time-dependent, the second fundamental form would vanish. If the perturbation did not depend on the spatial coordinates, it would be identically zero due to asymptotic flatness.

We substitute g_{ij} and k_{ij} into equation (5) and use assumptions (6c) and the hypothesis that terms of the order of $\mathcal{O}(1/r^{10})$ can be neglected to drop terms where possible, we end up with a simplified version of the generalised Jang equation

$$\frac{\left(\left(1 + \frac{M}{2r}\right)^{-4} - \left(1 + \frac{M}{2r}\right)^{-8} \sigma \right) \phi \Delta f}{\sqrt{1 + \phi^2 \left(\left(1 + \frac{M}{2r}\right)^{-4} - \left(1 + \frac{M}{2r}\right)^{-8} \sigma \right) \left(\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2 \right)}}$$
$$= \frac{3}{2} \left(1 - \frac{M}{2r} \right)^{-2} \left(1 + \frac{M}{2r} \right)^{-2} \frac{\partial \sigma}{\partial x_0}. \qquad (30)$$

Next, we consider the Dirac equation. This simplifies in a similar way once suitable terms are neglected under the same assumptions as those used for the Jang equation

$$\left(c(e_i)e_i + \frac{1}{8} \left(1 + \frac{M}{2r} \right)^{-4} c(e_i) \frac{\partial \sigma}{\partial x_i} \right) \tilde{\psi}_t = 0. \qquad (31)$$

Finally, we consider the equation for the warping factor. Given that the perturbation σ must decay to zero at spatial infinity, it is possible to give an explicit solution of the conformal flow of metrics by

$$u_t = \mathrm{e}^{-t} + \mathrm{e}^{-2t}\sigma, \qquad v_t = -1 - \mathrm{e}^{-t}\sigma, \qquad (32)$$

after using the assumption that one can neglect the time derivative of σ when it is multiplied by a function which decays exponentially with time. Equation (8) then becomes

$$\phi = 2 \int_0^\infty \left(\left(\frac{2 + \mathrm{e}^{-t}\sigma}{2} \right)^2 |\psi_t^-|^2 - \left(\frac{\mathrm{e}^{-t}\sigma}{2} \right)^2 |\psi_t^+|^2 \right) \left(\mathrm{e}^{-t} + \mathrm{e}^{-2t}\sigma \right)^2 \, \mathrm{d}t. \quad (33)$$

Neglecting terms which are higher-order in σ and using our assumption that a product of σ with a function which decays with time faster than e^{-2t} can be neglected, we end up with

$$\phi = 2 \int_0^\infty |\psi_t^-|^2 \mathrm{e}^{-2t}\,\mathrm{d}t = |\psi_t^-|^2 \tag{34}$$

since $\tilde{\psi}_t$ does not depend on time.

It now remains to find a solution for this system (it is not necessary to provide a uniqueness result). Given our assumptions on the spatial derivatives of σ, we may freely specify a solution to (23), which then prescribes ϕ. We will make the choice

$$\tilde{\psi}_t = \left(1 + \frac{M}{2r}\right)^{1/4}, \tag{35}$$

which determines

$$\phi = \sqrt{1 + \frac{M}{2r}}, \tag{36}$$

when we use our assumption on the sizes of the spinors that $|\psi_t^-| = |\tilde{\psi}_t|$.

We will then take the equation for the Jang graph $f(x,y,z)$ to be

$$f = \frac{1}{(x^2 + y^2 + z^2)^{3/2}}. \tag{37}$$

Solutions of the generalised Jang equation are studied in detail in [21] (interestingly, blow-up solutions are allowed for this equation). Substituting ϕ and f into the Jang equation, we obtain after re-arranging

$$\frac{\partial \sigma}{\partial x_0} = \frac{4\left(1 - \frac{M}{2r}\right)^2 r^{-5}\left(\left(1 + \frac{M}{2r}\right)^{-3/2} - \left(1 + \frac{M}{2r}\right)^{-11/2}\sigma\right)}{\sqrt{1 + 9\left(\left(1 + \frac{M}{2r}\right)^{-3} - \left(1 + \frac{M}{2r}\right)^{-7}\sigma\right)r^{-3}}}. \tag{38}$$

By our earlier assumption, the left-hand side is less than or equal to the right-hand side. Furthermore, ϕ and $\tilde{\psi}_t$ meet the required boundary conditions, so we have shown that the boundary value problem as a whole admits a solution for the particular perturbations which we have chosen.

Note that our result is not quite a proof of the Penrose inequality because we have found an approximate solution to a simplified system, whereas one needs a full solution of the system in order to imply the Penrose inequality. Although we currently do not have a rigorous proof, we believe that one

can rigorously show that the existence of certain approximate solutions also implies the Penrose inequality. This would be an interesting avenue for future research, as it might allow for a proof for a wider class of perturbations of Schwarzschild spacetime, not just the ones which we have considered. To see why one might expect this to be true, note that the main purpose of the Jang equation is to deal with troublesome terms in the generalised Schoen–Yau identity which do not have a definite sign. If one linearises the Jang equation and solves it to leading order, one could plug the solution to leading order back into the generalised Schoen–Yau identity to confirm that the scalar curvature is still nonnegative. This would then show that the Penrose inequality is satisfied up to leading order, which is what is required when proving the inequality for perturbations about Schwarzschild spacetime. We do not see any reason why the scalar curvature should cease to be nonnegative using this approach, although the implication is not immediate and would require a proof.

Another criticism of our use of an approximate solution is that the choices which we have imposed on the warping factor ϕ and the graph f are somewhat *ad hoc*. As emphasised earlier, other authors in the literature who have worked with our system of equations (or alternatively, the system given by coupling the generalised Jang equation to the inverse mean curvature flow) have also imposed extremely specific and equally *ad hoc* choices for the warping factor ϕ [19]. Such a choice is currently necessary in order to obtain an existence result for our system, even though a restriction of this kind does not have an obvious physical interpretation and might be regarded as unphysical.

We will make a few more comments to compare the perturbations we have studied here with those considered by the authors mentioned in Introduction. Kopiński and Tafel showed that in the conformally flat case the following Penrose inequality is specifically preserved under a small addition of the axially symmetric traceless exterior curvature to the Schwarzschild initial data

$$M \geq \sqrt{\frac{|S_\mathrm{h}|}{16\pi}}, \tag{39}$$

where $|S_\mathrm{h}|$ is the surface area of the event horizon in the Kerr metric [11]. Although this class of perturbations seems restrictive, the main advantage of this result is that in their approach the internal horizon is not assumed to be a minimal surface, in contrast to the Jang equation approach [9]. The initial data considered by Kopiński and Tafel were assumed to be maximal (much as we have assumed), but they subsequently generalised the same result to nonmaximal initial data [12]. Roszkowski and Malec, on the other

hand, demonstrated that the Penrose inequality

$$M \geq \sqrt{\frac{A_{\mathrm{h}}}{16\pi}} \tag{40}$$

holds for linear axially symmetric perturbations of maximal slices of Schwarzschild initial data, where A_{h} is the area of the apparent horizon [13]. As with our result, they assume that terms of the order of ϵ^2 or higher can be neglected, where ϵ is a 'smallness parameter' such that when $\epsilon = 0$ the background geometry emerges (the standard Schwarzschild slice). In our case, however, rather than a parameter ϵ, we have a function σ which is small in a suitable norm, such that terms of the order of σ^2 or higher can be neglected. In their case, there is no requirement that the horizon be a minimal surface. The geometric approach taken by these authors is very different and requires demonstrating that there exist certain convex foliations of the initial value surface, following a suggestion of Malec, Mars and Simon for proving the general Penrose inequality [22].

4. Conclusions

In summary, we have shown that an approximate solution exists to the system of equations given by the coupling between the generalised Jang equation and the conformal flow of metrics. It only remains to show how and why approximate solutions to the system can yield the Penrose inequality to establish the inequality for spacetimes which are certain specific conformally flat perturbations of Schwarzschild spacetime. In working with this system, we have made a number of other assumptions which makes our class of possible perturbations somewhat restrictive. For this reason, it is still not clear whether the approach of coupling the conformal flow of metrics to the generalised Jang equation is likely to be a viable method for proving the Penrose inequality, even if we restrict only to arbitrary perturbations of Schwarzschild spacetime. It would be a very significant advance if one could prove the Penrose inequality for an arbitrary perturbation of Schwarzschild spacetime by linearising the Jang/conformal flow system and solving it with a suitable implicit function theorem on Banach spaces, but the analysis and estimates involved in such a task are currently intractable as far as we know. This gives one an idea of how intractable the analysis would be in the general case, given that even the linearised system does not appear to be amenable to current analysis techniques.

It is possible to couple the generalised Jang equation to other geometric flows such as the inverse mean curvature flow, but one then loses the more attractive features of the conformal flow of metrics (improved regularity and strict positivity of the warping factor, for example) [8, 9]. However, hope

is not lost regarding the prospect of proving the general Penrose inequality with current techniques, as other authors have started to consider the possibility of a spin geometric proof (similar in some ways to Witten's spin geometry proof of the positive energy theorem) [23]. This approach offers an alternative possibility for proving the general Penrose inequality at some point in the future, although such a proof would itself also involve some delicate analytic issues.

REFERENCES

[1] R. Penrose, «Naked singularities», *Ann. N.Y. Acad. Sci.* **224**, 125 (1973).

[2] M. Mars, «Present status of the Penrose inequality», *Class. Quantum Grav.* **26**, 193001 (2009).

[3] H. Williams, «Recent Developments in the Penrose Conjecture», `arXiv:2111.00106 [gr-qc]`.

[4] G. Huisken, T. Ilmanen, «The Inverse Mean Curvature Flow and the Riemannian Penrose Inequality», *J. Differ. Geom.* **59**, 353 (2001).

[5] H.L. Bray, «Proof of the Riemannian Penrose Inequality Using the Positive Mass Theorem», *J. Differ. Geom.* **59**, 177 (2001).

[6] R. Schoen, S.-T. Yau, «On the proof of the positive mass conjecture in general relativity», *Commun. Math. Phys.* **65**, 45 (1979).

[7] R. Schoen, S.-T. Yau, «Proof of the positive mass theorem. II», *Commun. Math. Phys.* **79**, 231 (1981).

[8] H.L. Bray, M. Khuri, «P.D.E.'s Which Imply the Penrose Conjecture», *Asian J. Math.* **15**, 557 (2011).

[9] Q. Han, M. Khuri, «The Conformal Flow of Metrics and the General Penrose Inequality», *Adv. Math. Phys.* **2018**, 7390148 (2018).

[10] S. Alexakis, «The Penrose inequality on perturbations of the Schwarzschild exterior», `arXiv:1506.06400 [gr-qc]`.

[11] J. Kopiński, J. Tafel, «The Penrose inequality for perturbations of the Schwarzschild initial data», *Class. Quantum Grav.* **37**, 015012 (2020).

[12] J. Kopiński, J. Tafel, «The Penrose inequality for nonmaximal perturbations of the Schwarzschild initial data», *Class. Quantum Grav.* **37**, 105006 (2020).

[13] K. Roszkowski, E. Malec, «The Penrose inequality in perturbed Schwarzschild geometries», *Acta Phys. Pol. B* **36**, 2931 (2005).

[14] P. Petersen, «Riemannian Geometry», *Springer*, New York 2006.

[15] E. Witten, «A new proof of the positive energy theorem», *Commun. Math. Phys.* **80**, 381 (1981).

[16] T. Parker, C.H. Taubes, «On Witten's proof of the positive energy theorem», *Commun. Math. Phys.* **84**, 223 (1982).

[17] L. Nirenberg, H.F. Walker, «The Null Spaces of Elliptic Partial Differential Operators in $\mathbb{R}^n$», *J. Math. Anal. Appl.* **42**, 271 (1973).

[18] M. Cantor, «Spaces of Functions with Asymptotic Conditions on $\mathbb{R}^n$», *Indiana Univ. Math. J.* **24**, 897 (1975).

[19] H.L. Bray, M. Khuri, «A Jang Equation Approach to the Penrose Inequality», *Discrete Cont. Dyn. Syst.* **27**, 741 (2010).

[20] D. Gilbarg, N.S. Trudinger, «Elliptic Partial Differential Equations of Second Order», *Springer Berlin Heidelberg*, Berlin, Heidelberg 2001.

[21] Q. Han, M. Khuri, «Existence and Blow-Up Behavior for Solutions of the Generalized Jang Equation», *Commun. Part. Differ. Eqs.* **38**, 2199 (2013).

[22] E. Malec, M. Mars, W. Simon, «On the Penrose inequality for General Horizons», *Phys. Rev. Lett.* **88**, 121102 (2002).

[23] J. Kopiński, J.A.V. Kroon, «New spinorial approach to mass inequalities for black holes in general relativity», *Phys. Rev. D* **103**, 024057 (2021).