

Construction of Cauchy data of vacuum Einstein field equations evolving to black holes

By JUNBIN LI and PIN YU

Abstract

We show the existence of complete, asymptotically flat Cauchy initial data for the vacuum Einstein field equations, free of trapped surfaces, whose future development must admit a trapped surface. Moreover, the datum is exactly a constant time slice in Minkowski space-time inside and exactly a constant time slice in Kerr space-time outside.

The proof makes use of the full strength of Christodoulou's work on the dynamical formation of black holes and Corvino-Schoen's work on the construction of initial data sets.

1. Introduction

1.1. *Earlier works.* Black holes are the central objects of study in general relativity. The presence of a black hole is usually detected through the existence of a trapped surface, namely, a two-dimensional space-like sphere whose outgoing and incoming expansions are negative. The celebrated Penrose singularity theorem states that under suitable assumptions, if the space-time has a trapped surface, then the space-time is future causally geodesically incomplete; i.e., it must be singular, at least in some weak sense. On the other hand, the weak cosmic censorship conjecture (WCC) asserts that singularities have to be hidden from an observer at infinity by the event horizon of a black hole. Thus, assuming WCC, the theorem of Penrose predicts the existence of black holes, via the exhibition of a trapped surface. This is precisely the reason why the trapped surfaces are intimately related to the understanding of the mechanism of gravitational collapse.

JL is partially supported by NSF-China Grant 11271377. PY is supported by NSF-China Grant 11101235.

© 2015 Department of Mathematics, Princeton University.

In [2], Christodoulou discovered a remarkable mechanism responsible for the dynamical formation of trapped surfaces. He proved that a trapped surface can form, even in vacuum space-time, from completely dispersed initial configurations (i.e., free of any trapped surfaces) and purely by the focusing effect of gravitational waves. Christodoulou described an open set of initial data (so-called *short pulse ansatz*) on an outgoing null hypersurfaces without any symmetry assumptions. Based on the techniques he and Klainerman developed in [3] proving the global stability of the Minkowski space-time, he managed to understand the cancellations among the different components of connection and curvature. This eventually enabled him to obtain a complete picture of how the various geometric quantities propagate along the evolution. Christodoulou also proved a version of the above result for data prescribed at the past null infinity. He showed that strongly focused gravitational waves, coming in from past null infinity, lead to a trapped surface. More precisely, he showed that if the incoming energy per unit solid angle in each direction in an advanced time interval $[0, \delta]$ at null infinity is sufficiently large (and yet sufficiently dispersed so that no trapped surfaces are present), the focusing effect will lead to gravitational collapses. Besides its important physical significance, from the point of view of partial differential equations, it establishes the first result on the long time dynamics in general relativity for general initial data which are not necessarily close to the Minkowski space-time.

In [11], Klainerman and Rodnianski extended Christodoulou's result. They introduced a parabolic scaling and studied a broader class of initial data. The new scaling allowed them to capture the hidden smallness of the nonlinear interactions in Einstein equations. Another key observation in their paper is that, if one enlarges the admissible set of initial conditions, the corresponding propagation estimates are much easier to derive. Based on this idea, they gave a significant simplification of Christodoulou's result. At the same time, their relaxation of the propagation estimates were just enough to guarantee the formation of a trapped surface. Based on the geometric sharp trace theorems, which they have introduced earlier in [6] and applied to local well posedness for vacuum Einstein equations in [7],[8], [10] and [9], they were also able to reduce the number of derivatives needed in the estimates from two derivatives on curvature (in Christodoulou's proof) to just one. The price for such a simpler proof, with a larger set of data, is that the natural propagation estimates, consistent with the new scaling, are weaker than those of Christodoulou's. Nevertheless, once the main existence results are obtained, improved propagation estimates can be derived by assuming more conditions on the initial data, such as those consistent with Christodoulou's assumptions. This procedure allows one to recover Christodoulou's stronger results in a straightforward manner. We remark that, from a purely analytic point of view, the main difficulty of all

the aforementioned results on dynamical formation of black holes is the proof of the long time existence results. The work [11] overcomes this difficulty in an elegant way.¹

When matter fields are present, some black-hole formation results have been established much earlier, under additional symmetry assumptions. The most important such work is by Christodoulou (see [1]). He studied the evolutionary formation of singularities for the Einstein-scalar field equations, under the assumption of spherical symmetry. (Notice that according to Birkhoff theorem, this assumption is not interesting in vacuum.) The incoming energy of the scalar field was the main factor leading to the gravitational collapse. We note also that the work [1] provides a much more precise picture for the large scale structure of the space-time than that available in the general case.

In all the aforementioned works, the initial data are prescribed on null hypersurfaces. It is however natural to study the question of formation of trapped surfaces for Cauchy initial data. We first recall that Cauchy data, i.e., the data defined on a space-like hypersurface, must satisfy a system of partial differential equations, namely, the constraint equations (1.1). The main advantage of using characteristic initial data is that one has complete freedom in specifying data without any constraint. In [2], Christodoulou also proposed an idea for initial data prescribed on the *complete* past null infinity. This approach leads to asymptotically flat Cauchy data by a limiting argument. We remark that, in such a way, the behavior of the data near the space-like is not clear. It is natural to look for a constructive approach to this problem. In this connection, we mention an interesting piece of work [15] by Schoen and Yau. They showed that on a space-like hypersurface, there exists a trapped surface when the matter field is condensed in a small region; see also [16] for an improvement. Their proof analyzed the constraint equations and made use of their earlier work on positive mass theorem (especially the resolution of Jang's equation). We remark that their work is not evolutionary and matter fields are essential to the existence of trapped surfaces.

The goal of the current paper is to exhibit Cauchy initial data for vacuum Einstein field equations with a precise asymptotic behavior at space-like infinity, free of trapped surfaces, that lead to trapped surfaces along the evolution. We give the precise statement of the result in next section. We remark that the precise behavior of the data at infinity in this paper is also related to the conjecture that the final state of a vacuum dynamical black hole is a Kerr black hole.

¹We remark that Klainerman and Rodnianski only considered the problem on a finite region. We also note that the problem from past null infinity has also been studied by Reiterer and Trubowitz [13], by a different approach.

1.2. *Main result.* Let Σ be a three-dimensional differentiable manifold diffeomorphic to \mathbb{R}^3 and (x_1, x_2, x_3) be the standard Cartesian coordinate system. We also use $|x| = \sqrt{x_1^2 + x_2^2 + x_3^2}$ to denote the usual radius function. Let $r_0 > 1$ be a given number, $\delta > 0$ a small positive number and $\varepsilon_0 > 0$ another small positive number. We divide Σ into four concentric regions $\Sigma = \Sigma_M \cup \Sigma_C \cup \Sigma_S \cup \Sigma_K$, where

$$\begin{aligned}\Sigma_M &= \{x \mid |x| \leq r_0\}, & \Sigma_C &= \{x \mid r_0 \leq |x| \leq r_1\}, \\ \Sigma_S &= \{x \mid r_1 \leq |x| \leq r_2\}, & \Sigma_K &= \{x \mid |x| \geq r_2\}.\end{aligned}$$

The numbers r_1, r_2 will be fixed in the sequel such that $r_1 - r_0 = O(\delta)$ and $r_2 - r_1 = O(\varepsilon_0)$.

A Cauchy initial datum for vacuum Einstein field equations on Σ consists of a Riemannian metric \bar{g} and a symmetric two tensor \bar{k} (as the second fundamental form of Σ embedded in the space-time to be solved) subject to the following constraint equations:

$$(1.1) \quad \begin{aligned}R(\bar{g}) - |\bar{k}|^2 + \bar{h}^2 &= 0, \\ \operatorname{div}_{\bar{g}} \bar{k} - d\bar{h} &= 0,\end{aligned}$$

where $R(\bar{g})$ is the scalar curvature of the metric \bar{g} and $\bar{h} \triangleq \operatorname{tr}_{\bar{g}} \bar{k}$ is the mean curvature.

In order to state the main theorem, we also need to specify a mass parameter $m_0 > 0$ that will be defined in an explicit way in the course of the proof. Our main result is as follows:

MAIN THEOREM. *For any sufficiently small $\varepsilon > 0$, there is a Riemannian metric \bar{g} and a symmetric two tensor \bar{k} on Σ satisfying (1.1), such that*

1. Σ_M is a constant time slice in Minkowski space-time; in fact, $(\bar{g}, \bar{k}) = (\bar{g}_E, 0)$ where \bar{g}_E is the Euclidean metric.
2. Σ_K is isometric to a constant time slice in Boyer-Lindquist coordinate² all the way up to space-like infinity in a Kerr space-time with mass m and angular momentum \mathbf{a} . Moreover, $|m - m_0| + |\mathbf{a}| \lesssim \varepsilon$.
3. Σ is free of trapped surfaces.
4. Trapped surfaces will form in the future domain of dependence of Σ .

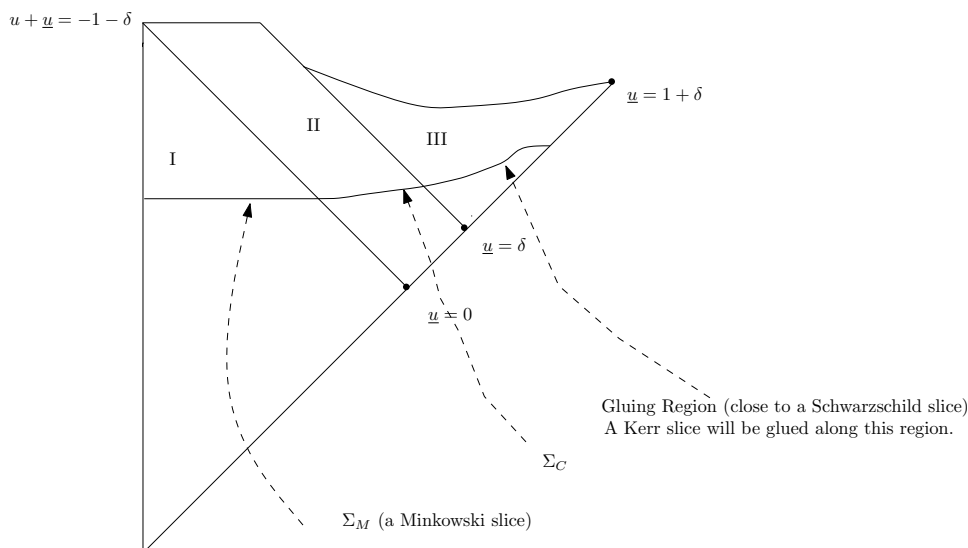
²We recall the form of the Kerr metric in the Boyer-Lindquist coordinates:

$$\begin{aligned}g_K &= \left(-1 + \frac{2mr}{r^2 + a^2 \cos^2 \theta}\right) dt^2 - \frac{2mra \sin^2 \theta}{r^2 + a^2 \cos^2 \theta} dt d\varphi + \frac{r^2 + a^2 \cos^2 \theta}{r^2 - 2mr + a^2} dr^2 \\ &\quad + (r^2 + a^2 \cos^2 \theta) d\theta^2 + \sin^2 \theta \left(r^2 + a^2 + \frac{2mra^2 \sin^2 \theta}{r^2 + a^2 \cos^2 \theta}\right) d\varphi^2.\end{aligned}$$

Remark 1. The mass parameter m_0 reflects the amount of incoming gravitational energy that we inject into the Minkowski space-time through an outgoing null hypersurface. It can be computed explicitly from the initial conditions as follows (see Section 2 for definitions):

$$m_0 = \frac{1}{4} \int_0^\delta |u_0|^2 |\widehat{\chi}(\underline{u}, u_0, \theta)|_g^2 d\underline{u}.$$

Remark 2. The regions Σ_C and Σ_S do not appear in the theorem. In the proof, we shall see that Σ_S will be ε -close to a constant time slice in a Schwarzschild space-time with mass m_0 and Σ_C will be constructed from Christodoulou’s short pulse ansatz.



We now sketch the heuristics of the proof with the help the above picture. We start by describing initial data on a truncated null cone (represented by the outgoing segment between the bottom vertex of the cone and the point $\underline{u} = 1 + \delta$ in the picture).

The part of the cone between the vertex and $\underline{u} = 0$ will be a light cone in Minkowski space-time; thus its future development, i.e., the region I in the picture, is flat.

For data between $\underline{u} = 0$ and $\underline{u} = \delta$, we use Christodoulou’s initial data (from which a trapped surface will appear on the right-top vertex of the region II) with an additional condition. This condition requires that the incoming energy through this part of the cone be spherically symmetric. The main consequence of imposing this condition is that the incoming hypersurface from $\underline{u} = \delta$ (represented by segment between the region II and the region III in the picture) will be close to an incoming null cone in Schwarzschild space-time.

For data between $\underline{u} = \delta$ and $\underline{u} = 1 + \delta$, we require that its shear be identically zero. Together with the data on $\underline{u} = \delta$, one can show that this part of the data will be close to an outgoing null cone in Schwarzschild space-time.

We construct the region III (pictured above) by means of solving the vacuum Einstein equation using the initial data described in the previous two paragraphs. Thanks to the consequences of our additional condition, this region is close to a region in Schwarzschild space-time.

We then can choose a smooth space-like hypersurface such that it coincides with a constant time slice in Minkowski space-time in the region I (noted as Σ_M in the picture) and it is also close to a constant time slice in Schwarzschild space-time (noted as the gluing region in the picture). Due to the closeness to Schwarzschild slice, we can do explicit computations to understand the obstruction space needed in Corvino-Schoen [5] construction, and we can eventually attach a Kerr slice to this region. In particular, this Kerr slice contains the spatial infinity i^0 (which is not shown in the above picture). Compared to the results of Christodoulou [2] or Klainerman-Rodnianski [11], this is precisely the point of the current work: we construct Cauchy initial data for vacuum Einstein equations with a precise asymptotic behavior at space-like infinity, free of trapped surfaces, that lead to trapped surfaces along the evolution.

1.3. *Comments on the proof.* We would now like to address the motivations for and difficulties in the proof.

As we stated in the Main Theorem, the ultimate goal is to obtain a Kerr slice Σ_K . To this end, we would like to use a gluing construction for the constraint equations due to Corvino and Schoen [5]. Roughly speaking, if two initial data sets are close to each other, this construction allows us to patch one to another along a gluing region without changing the data outside. In reality, because the constraint equations (1.1) is a nonlinear elliptic system, in order to realize the gluing construction, Corvino and Schoen used a sophisticated type of implicit function theorem argument in such a way that the resulting metric (and the resulting second fundamental form) also satisfies the constraint equations in the gluing region. [5] extended an earlier work of Corvino [4] which proved a parallel result for time symmetric data. Both constructions in [4] and [5] relied on the study of local deformations of the constraint maps (see Section 2.3). We will also have to make use of the deformation technique. Morally speaking, three ingredients are required to implement this approach:

1. *A reasonable amount of smoothness, say some $C^{k,\alpha}$ control on the background geometry.* This is necessary since the local deformation techniques can only be proved for a relatively regular class of data; see Section 2.3 or [5] for details.
2. *Some precise information on the background metric.* This is essential since eventually we can only glue two data sets close to each other. *A priori*, one of

them is a Kerr slice. Therefore, we expect to contract some space-time close to some Kerr slice on a given region. We call this requirement the *smallness* condition since it will be captured eventually by the smallness of ε .

3. *The gluing region must have a fixed size.* This is important because the local deformation techniques work only on a fixed open set. If the gluing region shrinks to zero, it is not clear how one can proceed.

The first ingredient on the smoothness motivates us to derive the higher order energy estimates in Section 3. In Christodoulou's work [2], he obtained energy estimates up to two derivatives on the curvature components. He also had energy estimates for higher order derivatives, but the bounds were far from sharp. For our purposes, we require not only differentiability but also better bounds.

The second ingredient turns out to be the key of the entire proof. First of all, we remark that the higher order energy estimates are also important for the second ingredient. To clarify this connection, we review a technical part of Christodoulou's proof in [2]. One of the most difficult estimates is the control on $\|\nabla \widehat{\otimes} \eta\|_{L^\infty}$. He used two mass aspect functions μ and $\underline{\mu}$ coupled with η and $\underline{\eta}$ in Hodge systems. The procedure allowed him to gain one derivative from elliptic estimates to close the bootstrap argument. However, he had to incur a loss of smallness on the third derivatives of η . We remark that the loss only happened to the top order derivatives. In fact, for lower order derivatives, we can integrate some propagation equations to obtain smallness. (But this integration loses one derivative instead!) In other words, we can afford one more derivative in exchange for the smallness. From the above, it is clear that once we have higher order energy estimates, we can expect better controls.

To obtain the second ingredient, we will show that a carefully designed data will evolve to a region that is ε -close to some Schwarzschild space-time; it is in this sense that we have precise information on the background metric. As a consequence, we will be able to explicitly write down a four-dimensional space in the kernel of the formal adjoint of the linearized constraint maps. This kernel is the obstruction space for the gluing construction, and we shall use mass and angular momentum parameters (which is of four dimensions in total) of the Kerr family to remove the obstruction. In this connection, we want to point out that in the work of Reiterer and Trubowitz [13], they also obtained such a near Schwarzschild space-time region (which to the best of our knowledge may not satisfy the requirement in our third ingredient).

To obtain an almost Schwarzschild slab, one needs a careful choice of initial data from amongst those prescribed using the short pulse ansatz of Christodoulou. We shall impose the following condition:

$$\int_0^\delta |u_0|^2 |\widehat{\chi}(\underline{u}, u_0, \theta)|_{\underline{g}}^2 d\underline{u} = \text{constant},$$

namely, independent of the θ variable. Heuristically, this says that the total incoming gravitational energy is the same for all the spherical directions. We expect that through such a pulse of incoming gravitational wave, after an advanced time interval $[0, \delta]$, we can regain certain spherical symmetry. We remark that the above condition is the main innovation of the paper and will be crucial to almost all of our estimates.

Finally, we discuss our third ingredient. One may attempt to carry out the gluing directly on the space-time constructed in Christodoulou's work. However, when the parameter δ goes to zero, one runs the risk of shrinking the gluing region to zero. To solve this difficulty, we will further extend the short pulse ansatz beyond the advanced time interval $[0, \delta]$ to $[0, 1 + \delta]$. This extension allows us to further solve Einstein equations to construct a space-time slab of a fixed size.

To close the section, we would like to discuss a way to prove the Main Theorem based on the work of Klainerman and Rodnianski [11] instead of the stronger result [2] of Christodoulou. Since we have to rely on higher order derivative estimates, we can actually start with the existence result in [11], which only has control up to first derivatives on curvature. We then can assume more on the data, as we mentioned in the introduction, to derive more on the solution. As we will do in sequel, we can do the same induction argument on the number of derivatives on curvatures. This will yield the same higher order energy estimates, and the rest of the proof remains the same.

Acknowledgements. J. L. is deeply indebted to his Ph.D advisor Xi-Ping Zhu for support on this work and to Siu-Hung Tang for discussions. He would like to thank the Mathematical Sciences Center of Tsinghua University where the work was partially done during his visit. P. Y. would like to thank Sergiu Klainerman for numerous suggestions on improving the earlier manuscripts; Shing-Tung Yau for his idea of proving the absence of trapped surfaces; and Demetrios Christodoulou, Igor Rodnianski and Richard Schoen for their interests and discussions. We would like to thank anonymous referee for his/her thorough review and highly appreciate the comments and suggestions, which significantly contributed to improving the quality (of the mathematics and the language) of our paper.

2. Preliminaries

2.1. *Preliminaries on geometry.* We follow the geometric setup in [2] and use the double null foliations for most of the paper. We use \mathcal{D} to denote the underlying space-time and use g to denote the background 3+1-dimensional Lorentzian metric. We also use ∇ to denote the Levi-Civita connection of the metric g .

Recall that we have two optical functions \underline{u} and u defined on \mathcal{D} such that

$$g(\nabla \underline{u}, \nabla \underline{u}) = g(\nabla u, \nabla u) = 0.$$

The space-time \mathcal{D} can be foliated by the level sets of \underline{u} and by the level sets of u respectively. Since the gradients of u and \underline{u} are null, we call these two foliation together a *double null foliation*. We require the functions u and \underline{u} increase towards the future, and this indeed defines a time orientation on \mathcal{D} . We use C_u to denote the outgoing null hypersurfaces generated by the level surfaces of u and use $\underline{C}_{\underline{u}}$ to denote the incoming null hypersurfaces generated by the level surfaces of \underline{u} . We also use $S_{\underline{u},u} = \underline{C}_{\underline{u}} \cap C_u$, which is a space-like two-sphere. The notation $\underline{C}_{\underline{u}}^{[u',u']}$ refers to the part of the incoming cone $\underline{C}_{\underline{u}}$ where $u' \leq u \leq u''$, and the notation $C_u^{[u',u']}$ refers to the part of the outgoing cone C_u where $u' \leq u \leq u''$.

We remark that topology of \mathcal{D} is as the product of U and a two-sphere. Here, U is an open set in \mathbb{R}^2 and has (\underline{u}, u) as coordinates. The space-time \mathcal{D} does not exist at the moment, since it will be the solution of the Einstein equations that need to be solved. However, in this section, not only do we assume \mathcal{D} is given, but we also assume it has special geometric structures. The main reason is that we will do *a priori* estimates for the solutions of Einstein equations. Once we obtain the *a priori* estimates, in reality, the continuity argument (or bootstrap argument) gives the construction of \mathcal{D} together with those geometric structures.

Following Christodoulou [2], for the initial null hypersurface C_{u_0} where $u_0 < -2$ is a fixed constant, we require that $C_{u_0}^{[u_0,0]}$ be a flat light cone in Minkowski space-time and S_{0,u_0} be the standard sphere with radius $|u_0|$. Thus, by solving the Einstein vacuum equations, we know the future domain of dependence of $C_{u_0}^{[u_0,0]}$ is flat. It is equivalent to saying that the past of \underline{C}_0 can be isometrically embedded into Minkowski space-time. In particular, the incoming null hypersurface \underline{C}_0 coincides with an incoming light cone of the Minkowski space-time. In the sequel, on the initial hypersurface C_{u_0} , we shall only specify initial data on $C_{u_0}^{[0,\delta+1]}$.

We are ready to define various geometric quantities. The positive function Ω is defined by the formula $\Omega^{-2} = -2g(\nabla \underline{u}, \nabla u)$. We then define the normalized null pair (e_3, e_4) by $e_3 = -2\Omega \nabla \underline{u}$ and $e_4 = -2\Omega \nabla u$. We also need two more null vector fields $\underline{L} = \Omega e_3$ and $L = \Omega e_4$. We remark that the flows generated by \underline{L} and L preserve the double null foliation. On a given two sphere $S_{\underline{u},u}$ we choose a local orthonormal frame (e_1, e_2) . We call (e_1, e_2, e_3, e_4) a *null frame*. As a convention, throughout the paper, we use capital Latin letters A, B, C, \dots to denote an index from 1 to 2, e.g., e_A denotes either e_1 or e_2 ;

we use little Latin letters i, j, k, \dots to denote an index from 1 to 3. Repeated indices should always be understood as summations.

Let ϕ be a tangential tensorfield on \mathcal{D} . By definition, ϕ being tangential means that ϕ is *a priori* a tensorfield defined on the space-time \mathcal{D} and all the possible contractions of ϕ with either e_3 or e_4 are zeros. We use $D\phi$ and $\underline{D}\phi$ to denote the projection to $S_{\underline{u},u}$ of usual Lie derivatives $\mathcal{L}_L\phi$ and $\mathcal{L}_{\underline{L}}\phi$. The space-time metric g induces a Riemannian metric \mathcal{g} on $S_{\underline{u},u}$, and \mathcal{f} is the volume form of \mathcal{g} on $S_{\underline{u},u}$. We use \mathcal{d} and ∇ to denote the exterior differential and covariant derivative (with respect to \mathcal{g}) on $S_{\underline{u},u}$.

Let $(\theta^A)_{A=1,2}$ be a local coordinate system on the two sphere S_{0,u_0} . We can extend θ^A 's to the entire \mathcal{D} by first setting $L(\theta^A)=0$ on C_{u_0} , and then setting $\underline{L}(\theta^A)=0$ on \mathcal{D} . Therefore, we obtain a local coordinate system $(\underline{u}, u, \theta^A)$ on \mathcal{D} . In such a coordinate system, the Lorentzian metric g takes the following form:

$$g = -2\Omega^2(d\underline{u} \otimes d\underline{u} + du \otimes du) + \mathcal{g}_{AB}(d\theta^A - b^A du) \otimes (d\theta^B - b^B du).$$

The null vectors \underline{L} and L can be computed as $\underline{L} = \partial_{\underline{u}}$ and $L = \partial_{\underline{u}} + b^A \partial_{\theta^A}$. By construction, we have $b^A(\underline{u}, u_0, \theta) = 0$. In addition, we also require $\Omega(\underline{u}, u_0, \theta) = \Omega(0, u, \theta) = 1$, which implies that \underline{u} restricted to C_{u_0} and u restricted to \underline{C}_0 are affine parameters for the null generators of C_{u_0} and \underline{C}_0 respectively.

We recall the definitions of null connection coefficients. Roughly speaking, the following quantities are Christoffel symbols of ∇ according to the null frame (e_1, e_2, e_3, e_4) :

$$\begin{aligned} \chi_{AB} &= g(\nabla_A e_4, e_B), & \eta_A &= -\frac{1}{2}g(\nabla_3 e_A, e_4), & \omega &= \frac{1}{2}\Omega g(\nabla_4 e_3, e_4), \\ \underline{\chi}_{AB} &= g(\nabla_A e_3, e_B), & \underline{\eta}_A &= -\frac{1}{2}g(\nabla_4 e_A, e_3), & \underline{\omega} &= \frac{1}{2}\Omega g(\nabla_3 e_4, e_3). \end{aligned}$$

They are all tangential tensorfields. (ω and $\underline{\omega}$ are scalar fields, which can be viewed as $(0,0)$ type tensorfields.) We define $\chi' = \Omega^{-1}\chi$, $\underline{\chi}' = \Omega^{-1}\underline{\chi}$ and $\zeta = \frac{1}{2}(\eta - \underline{\eta})$. The trace of χ and $\underline{\chi}$ will play an important role in Einstein field equations, and they are defined by $\text{tr } \chi = \mathcal{g}^{AB}\chi_{AB}$ and $\text{tr } \underline{\chi} = \mathcal{g}^{AB}\underline{\chi}_{AB}$.³ We remark that the trace is taken with respect to the metric \mathcal{g} and the indices are raised by \mathcal{g} . By definition, we can directly check the following identities: $\mathcal{d} \log \Omega = \frac{1}{2}(\eta + \underline{\eta})$, $D \log \Omega = \omega$, $\underline{D} \log \Omega = \underline{\omega}$ and $\underline{D}b = 4\Omega^2\zeta^\sharp$. Here, ζ^\sharp is the vector field dual to the 1-form ζ . In the sequel, we will suppress the sign \sharp and use the metric \mathcal{g} to identify ζ and ζ^\sharp .

³Recall that the null expansion of a surface S related to a null normal is the mean curvature of S related to the null normal. Therefore $\text{tr } \chi$ and $\text{tr } \underline{\chi}$ are two null expansions of the surface $S_{\underline{u},u}$ and $S_{\underline{u},u}$ is a trapped surface if and only if $\text{tr } \chi < 0$ and $\text{tr } \underline{\chi} < 0$.

We can also decompose the curvature tensor \mathbf{R} into null curvature components:

$$\begin{aligned} \alpha_{AB} &= \mathbf{R}(e_A, e_4, e_B, e_4), & \beta_A &= \frac{1}{2} \mathbf{R}(e_A, e_4, e_3, e_4), & \rho &= \frac{1}{4} \mathbf{R}(e_3, e_4, e_3, e_4), \\ \underline{\alpha}_{AB} &= \mathbf{R}(e_A, e_3, e_B, e_3), & \underline{\beta}_A &= \frac{1}{2} \mathbf{R}(e_A, e_3, e_3, e_4), & \sigma &= \frac{1}{4} \mathbf{R}(e_3, e_4, e_A, e_B) \ell^{AB}. \end{aligned}$$

In order to express the Einstein vacuum equations with respect to a null frame, we have to introduce some operators. For a symmetric tangential 2-tensorfield θ , we use $\widehat{\theta}$ and $\text{tr} \theta$ to denote the trace-free part and trace of θ (with respect to $\not\partial$). If θ is trace-free, $\widehat{D}\theta$ and $\underline{D}\theta$ refer to the trace-free part of $D\theta$ and $\underline{D}\theta$. Let ξ be a tangential 1-form. We define some products and operators for later use. For the products, we define $(\theta_1, \theta_2) = \not\partial^{AC} \not\partial^{BD} (\theta_1)_{AB} (\theta_2)_{CD}$ and $(\xi_1, \xi_2) = \not\partial^{AB} (\xi_1)_A (\xi_2)_B$. This also leads to the following norms: $|\theta|^2 = (\theta, \theta)$ and $|\xi|^2 = (\xi, \xi)$. We then define the contractions $(\theta \cdot \xi)_A = \theta_A^B \xi_B$, $(\theta_1 \cdot \theta_2)_{AB} = (\theta_1)_A^C (\theta_2)_{CB}$, $\theta_1 \wedge \theta_2 = \ell^{AC} \not\partial^{BD} (\theta_1)_{AB} (\theta_2)_{CD}$ and $\xi_1 \widehat{\otimes} \xi_2 = \xi_1 \otimes \xi_2 + \xi_2 \otimes \xi_1 - (\xi_1, \xi_2) \not\partial$. The Hodge dual for ξ is defined by ${}^* \xi_A = \ell_A^C \xi_C$. For the operators, we define $\text{div} \xi = \nabla^A \xi_A$, $\text{curl} \xi_A = \ell^{AB} \nabla_A \xi_B$ and $(\text{div} \theta)_A = \nabla^B \theta_{AB}$. We finally define a traceless operator $(\nabla \widehat{\otimes} \xi)_{AB} = (\nabla \xi)_{AB} + (\nabla \xi)_{BA} - \text{div} \xi \not\partial_{AB}$.

For the sake of simplicity, we will use abbreviations Γ and R to denote an arbitrary connection coefficient and an arbitrary null curvature component. We introduce a schematic way to write products. Let ϕ and ψ be arbitrary tangential tensorfields; we also use $\phi \cdot \psi$ to denote an arbitrary contraction of ϕ and ψ by $\not\partial$ and ℓ . This schematic notation only captures the quadratic nature of the product, and it will be good enough for most of the cases when we derive estimates. As an example, the notation $\Gamma \cdot R$ means a sum of products between a connection coefficient and a curvature component.

We use a null frame (e_1, e_2, e_3, e_4) to decompose the Einstein vacuum equations $\text{Ric}(g) = 0$ into components. This leads to the following null structure equations (where K is the Gauss curvature of $S_{u,u}$):⁴

$$(2.1) \quad \widehat{D}\widehat{\chi}' = -\alpha,$$

$$(2.2) \quad D \text{tr} \chi' = -\frac{1}{2} \Omega^2 (\text{tr} \chi')^2 - \Omega^2 |\widehat{\chi}'|^2,$$

$$(2.3) \quad \underline{D}\underline{\widehat{\chi}}' = -\underline{\alpha},$$

$$(2.4) \quad \underline{D} \text{tr} \underline{\chi}' = -\frac{1}{2} \Omega^2 (\text{tr} \underline{\chi}')^2 - \Omega^2 |\underline{\widehat{\chi}}'|^2,$$

$$(2.5) \quad D\eta = \Omega(\chi \cdot \underline{\eta} - \beta),$$

⁴Readers may see Section 1.2 of [2] for the derivation of these equations.

$$(2.6) \quad \underline{D}\eta = \Omega(\underline{\chi} \cdot \eta + \underline{\beta}),$$

$$(2.7) \quad D\underline{\omega} = \Omega^2(2(\eta, \underline{\eta}) - |\eta|^2 - \rho),$$

$$(2.8) \quad \underline{D}\omega = \Omega^2(2(\eta, \underline{\eta}) - |\underline{\eta}|^2 - \rho),$$

$$(2.9) \quad K = -\frac{1}{4} \operatorname{tr} \chi \operatorname{tr} \underline{\chi} + \frac{1}{2}(\widehat{\chi}, \widehat{\underline{\chi}}) - \rho,$$

$$(2.10) \quad \operatorname{div} \widehat{\chi}' = \frac{1}{2} \not{d} \operatorname{tr} \chi' - \widehat{\chi}' \cdot \eta + \frac{1}{2} \operatorname{tr} \chi' \eta - \Omega^{-1} \beta,$$

$$(2.11) \quad \operatorname{div} \widehat{\underline{\chi}}' = \frac{1}{2} \not{d} \operatorname{tr} \underline{\chi}' - \widehat{\underline{\chi}}' \cdot \underline{\eta} + \frac{1}{2} \operatorname{tr} \underline{\chi}' \underline{\eta} - \Omega^{-1} \underline{\beta},$$

$$(2.12) \quad \operatorname{curl} \eta = \sigma - \frac{1}{2} \widehat{\chi} \wedge \widehat{\underline{\chi}},$$

$$(2.13) \quad \operatorname{curl} \underline{\eta} = -\sigma + \frac{1}{2} \widehat{\underline{\chi}} \wedge \widehat{\chi},$$

$$(2.14) \quad \widehat{D}(\Omega \widehat{\underline{\chi}}) = \Omega^2 \left(\nabla \widehat{\otimes} \eta + \underline{\eta} \widehat{\otimes} \eta + \frac{1}{2} \operatorname{tr} \chi \widehat{\underline{\chi}} - \frac{1}{2} \operatorname{tr} \underline{\chi} \widehat{\chi} \right),$$

$$(2.15) \quad D(\Omega \operatorname{tr} \underline{\chi}) = \Omega^2 \left(2 \operatorname{div} \underline{\eta} + 2|\eta|^2 - (\widehat{\chi}, \widehat{\underline{\chi}}) - \frac{1}{2} \operatorname{tr} \chi \operatorname{tr} \underline{\chi} + 2\rho \right),$$

$$(2.16) \quad \widehat{D}(\Omega \widehat{\chi}) = \Omega^2 \left(\nabla \widehat{\otimes} \eta + \eta \widehat{\otimes} \eta + \frac{1}{2} \operatorname{tr} \underline{\chi} \widehat{\chi} - \frac{1}{2} \operatorname{tr} \chi \widehat{\underline{\chi}} \right),$$

$$(2.17) \quad \underline{D}(\Omega \operatorname{tr} \chi) = \Omega^2 \left(2 \operatorname{div} \eta + 2|\eta|^2 - (\widehat{\underline{\chi}}, \widehat{\chi}) - \frac{1}{2} \operatorname{tr} \chi \operatorname{tr} \underline{\chi} + 2\rho \right).$$

We also use the null frame to decompose the contracted second Bianchi identity $\nabla^\alpha \mathbf{R}_{\alpha\beta\gamma\delta} = 0$ into components. This leads the following null Bianchi equations:⁵

$$(2.18) \quad \widehat{D}\alpha - \frac{1}{2} \Omega \operatorname{tr} \underline{\chi} \alpha + 2\underline{\omega} \alpha + \Omega \{ -\nabla \widehat{\otimes} \beta - (4\underline{\eta} + \zeta) \widehat{\otimes} \beta + 3\underline{\chi} \rho + 3^* \widehat{\underline{\chi}} \sigma \} = 0,$$

$$(2.19) \quad \widehat{D}\underline{\alpha} - \frac{1}{2} \Omega \operatorname{tr} \chi \underline{\alpha} + 2\omega \underline{\alpha} + \Omega \{ \nabla \widehat{\otimes} \underline{\beta} + (4\underline{\eta} - \zeta) \widehat{\otimes} \underline{\beta} + 3\underline{\chi} \rho - 3^* \widehat{\underline{\chi}} \sigma \} = 0,$$

$$(2.20) \quad D\beta + \frac{3}{2} \Omega \operatorname{tr} \chi \beta - \Omega \widehat{\underline{\chi}} \cdot \beta - \omega \beta - \Omega \{ \operatorname{div} \alpha + (\underline{\eta} + 2\underline{\zeta}) \cdot \alpha \} = 0,$$

$$(2.21) \quad \underline{D}\underline{\beta} + \frac{3}{2} \Omega \operatorname{tr} \underline{\chi} \underline{\beta} - \Omega \widehat{\chi} \cdot \underline{\beta} - \underline{\omega} \underline{\beta} + \Omega \{ \operatorname{div} \underline{\alpha} + (\eta - 2\underline{\zeta}) \cdot \underline{\alpha} \} = 0,$$

$$(2.22) \quad \underline{D}\beta + \frac{1}{2} \Omega \operatorname{tr} \underline{\chi} \beta - \Omega \widehat{\underline{\chi}} \cdot \beta + \omega \beta \\ - \Omega \{ \not{d} \rho + {}^* \not{d} \sigma + 3\underline{\eta} \rho + 3^* \underline{\eta} \sigma + 2\underline{\chi} \cdot \underline{\beta} \} = 0,$$

$$(2.23) \quad D\underline{\beta} + \frac{1}{2} \Omega \operatorname{tr} \chi \underline{\beta} - \Omega \widehat{\chi} \cdot \underline{\beta} + \omega \underline{\beta} \\ + \Omega \{ \not{d} \rho - {}^* \not{d} \sigma + 3\underline{\eta} \rho - 3^* \underline{\eta} \sigma - 2\underline{\chi} \cdot \beta \} = 0,$$

⁵See Proposition 1.2 in [2].

$$(2.24) \quad D\rho + \frac{3}{2}\Omega \operatorname{tr} \chi \rho - \Omega \left\{ \operatorname{div} \beta + (2\underline{\eta} + \zeta, \beta) - \frac{1}{2}(\widehat{\chi}, \alpha) \right\} = 0,$$

$$(2.25) \quad \underline{D}\rho + \frac{3}{2}\Omega \operatorname{tr} \underline{\chi} \rho + \Omega \left\{ \operatorname{div} \underline{\beta} + (2\underline{\eta} - \zeta, \underline{\beta}) + \frac{1}{2}(\widehat{\chi}, \underline{\alpha}) \right\} = 0,$$

$$(2.26) \quad D\sigma + \frac{3}{2}\Omega \operatorname{tr} \chi \sigma + \Omega \left\{ \operatorname{curl} \beta + (2\underline{\eta} + \zeta, * \beta) - \frac{1}{2}\widehat{\chi} \wedge \alpha \right\} = 0,$$

$$(2.27) \quad \underline{D}\sigma + \frac{3}{2}\Omega \operatorname{tr} \underline{\chi} \sigma + \Omega \left\{ \operatorname{curl} \underline{\beta} + (2\underline{\eta} - \zeta, * \underline{\beta}) + \frac{1}{2}\widehat{\chi} \wedge \underline{\alpha} \right\} = 0.$$

To conclude this subsection, we recall how one prescribes characteristic data for Einstein vacuum equations on two transversally intersecting null hypersurfaces, say $C_{u_0}^{[0, \delta+1]} \cup \underline{C}_0$ in our current situation. In general, the initial data given on $C_{u_0} \cup \underline{C}_0$ should consist of the full metric \mathcal{g}_{0, u_0} , torsion ζ , outgoing expansion $\operatorname{tr} \chi$ and incoming expansion $\operatorname{tr} \underline{\chi}$ on the intersecting sphere S_{0, u_0} together with the conformal geometry on $C_{u_0}^{[0, \delta+1]}$ and \underline{C}_0 . As we observed, the incoming surface \underline{C}_0 is a fixed cone in the Minkowski space-time, and thus \mathcal{g}_{0, u_0} , ζ , $\operatorname{tr} \chi$ and $\operatorname{tr} \underline{\chi}$ are already fixed on S_{0, u_0} . Therefore, to specify initial data, we only need to specify the conformal geometry on $C_{u_0}^{[0, \delta+1]}$. We will see how the *short pulse* data of Christodoulou is prescribed in next subsection.

2.2. *The work of Christodoulou.* We first discuss Christodoulou's short pulse ansatz presented in [2]. As we mentioned earlier, we need to specify the conformal geometry on $C_{u_0}^{[0, \delta+1]}$. Let $\Phi_{\underline{u}}$ be the one-parameter group generated by L . We can rewrite the induced metric $\mathcal{g}|_{S_{\underline{u}, u_0}}$ uniquely as $\mathcal{g}|_{S_{\underline{u}, u_0}} = (\phi|_{S_{\underline{u}, u_0}})^2 \widehat{\mathcal{g}}|_{S_{\underline{u}, u_0}}$, where $\phi|_{S_{\underline{u}, u_0}}$ is a positive function, such that the metric $\Phi_{\underline{u}}^* \widehat{\mathcal{g}}|_{S_{\underline{u}, u_0}}$ on S_{0, u_0} has the same volume form as $\mathcal{g}|_{S_{0, u_0}}$. In this language, we only need to specify $\Phi_{\underline{u}}^* \widehat{\mathcal{g}}|_{S_{\underline{u}, u_0}}$ freely on S_{0, u_0} since it yields the conformal geometry on $C_{u_0}^{[0, \delta+1]}$.

Let $\{(U_1, (\theta_1^A)), (U_2, (\theta_2^A))\}_{A=1,2}$ be the two stereographic charts on S_{0, u_0} . Thus, the round metric $\mathcal{g}|_{S_{0, u_0}}$ is expressed as $(\mathcal{g}|_{S_{0, u_0}})_{AB}(\theta) = \frac{|u_0|^2}{(1 + \frac{1}{4}|\theta|^2)^2} \delta_{AB}$ with $\theta = \theta_1$ or θ_2 and $|\theta|^2 = |\theta^1|^2 + |\theta^2|^2$. Since we require that $\widehat{\mathcal{g}}(\underline{u}) \triangleq \Phi_{\underline{u}}^* \widehat{\mathcal{g}}|_{S_{\underline{u}, u_0}}$ on S_{0, u_0} has the same volume form as $\mathcal{g}|_{S_{0, u_0}}$, that is,

$$\det(\widehat{\mathcal{g}}(\underline{u})_{AB})(\theta) = \det(\mathcal{g}(0)_{AB})(\theta) = \frac{|u_0|^4}{(1 + \frac{1}{4}|\theta|^2)^4},$$

$\widehat{\mathcal{g}}(\underline{u})$ is given by

$$\widehat{\mathcal{g}}(\underline{u})_{AB}(\theta) = \frac{|u_0|^2}{(1 + \frac{1}{4}|\theta|^2)^2} m_{AB}(\underline{u}, \theta) = \frac{|u_0|^2}{(1 + \frac{1}{4}|\theta|^2)^2} \exp \psi_{AB}(\underline{u}, \theta),$$

where m_{AB} takes values in the set of positive definite symmetric matrices with determinant one and ψ_{AB} takes values in the set of symmetric trace-free

matrices. After this reduction, to prescribe initial data, we only need to specify a function

$$\psi : [0, \delta + 1] \times S_{0, u_0} \longrightarrow \widehat{S}_2, \quad (\underline{u}, \theta) \mapsto \psi_{AB}(\underline{u}, \theta),$$

where \widehat{S}_2 denotes the set of 2×2 symmetric trace-free matrices.

In Christodoulou's work [2], he only provided data on $C_{u_0}^{[0, \delta]}$, i.e., $\underline{u} \in [0, \delta]$. First he chose a smooth compactly supported \widehat{S}_2 -valued function $\psi_0 \in C_c^\infty((0, 1) \times S_{0, u_0})$. Then he called the following specific data,

$$(2.28) \quad \psi(\underline{u}, \theta) = \frac{\delta^{\frac{1}{2}}}{|u_0|} \psi_0\left(\frac{\underline{u}}{\delta}, \theta\right),$$

the *short pulse ansatz* and he called ψ_0 the *seed data*.

Remark 3. A key ingredient for the current work is to give further restriction on the seed data ψ_0 ; see (2.30) or (2.31). We shall further extend ψ to the whole region $[0, \delta + 1] \times S_{0, u_0}$ by zero.

To state the main theorems in [2], especially the energy estimates, we also need to define some norms. We first assign to every one of connection coefficients and curvature components two numbers r and s . The number r is assigned in the following way:

$$\begin{aligned} r(\widehat{\chi}) &= -\frac{1}{2}, \quad r\left(\operatorname{tr} \chi - \frac{2}{|u|}\right) = 0, \quad r(\eta) = r(\underline{\eta}) = 0, \\ r(\widehat{\underline{\chi}}) &= \frac{1}{2}, \quad r\left(\operatorname{tr} \underline{\chi} + \frac{2}{|u|} - \frac{2\underline{u}}{|u|^2}\right) = 1, \quad r(\omega) = -\frac{1}{2}, \quad r(\underline{\omega}) = 1, \\ r(\alpha) &= -\frac{3}{2}, \quad r(\beta) = -\frac{1}{2}, \quad r(\rho) = r(\sigma) = 0, \quad r(\underline{\beta}) = 1, \quad r(\underline{\alpha}) = \frac{3}{2}. \end{aligned}$$

The number s is assigned in the following way:

$$\begin{aligned} s(\widehat{\chi}) &= -1, \quad s\left(\operatorname{tr} \chi - \frac{2}{|u|}\right) = -2, \quad s(\eta) = s(\underline{\eta}) = -2, \\ s(\widehat{\underline{\chi}}) &= -2, \quad s\left(\operatorname{tr} \underline{\chi} + \frac{2}{|u|} - \frac{2\underline{u}}{|u|^2}\right) = -3, \quad s(\omega) = -1, \quad s(\underline{\omega}) = -3, \\ s(\alpha) &= -1, \quad s(\beta) = -2, \quad s(\rho) = s(\sigma) = -3, \quad s(\underline{\beta}) = -4, \quad s(\underline{\alpha}) = -\frac{9}{2}. \end{aligned}$$

Recall that we use Γ or R to denote an arbitrary connection coefficient or curvature component. Let $k \in \mathbb{Z}_{\geq 0}$ be a nonnegative integer. Recall that Γ or

R denotes an arbitrary choice of connection coefficients or curvature components. According the index associated to Γ or R as above, we define

$$\begin{aligned} \mathcal{R}_k[R](u) &= \delta^{-\frac{1}{2}}|u|^{-1}\|\delta^{-r(R)}|u|^{-s(R)}(|u|\nabla)^k R\|_{L^2(C_u^{[0,\delta]})}, \\ \underline{\mathcal{R}}_k[R](\underline{u}) &= \||u|^{-\frac{3}{2}}\delta^{-r(R)}|u|^{-s(R)}(|u|\nabla)^k R\|_{L^2(\underline{C}_{\underline{u}})}, \\ \mathcal{O}_{k+1}[\Gamma](u) &= \delta^{-\frac{1}{2}}|u|^{-1}\|\delta^{-r(\Gamma)}|u|^{-s(\Gamma)}(|u|\nabla)^{k+1}\Gamma\|_{L^2(C_u^{[0,\delta]})}. \end{aligned}$$

We also define, with the same k ,

$$\begin{aligned} \mathcal{R}_{k-1}^4[R](\underline{u}, u) &= |u|^{-\frac{1}{2}}\|\delta^{-r(R)}|u|^{-s(R)}(|u|\nabla)^{k-1}R\|_{L^4(S_{\underline{u},u})}, \\ \mathcal{R}_{k-2}^\infty[R](\underline{u}, u) &= \|\delta^{-r(R)}|u|^{-s(R)}(|u|\nabla)^{k-2}R\|_{L^\infty(S_{\underline{u},u})}. \end{aligned}$$

We also use the notation $\mathcal{R}_k(u)$, $\underline{\mathcal{R}}_k(\underline{u})$, $\mathcal{O}_{k+1}(u)$, $\mathcal{R}_{k-1}^4(\underline{u}, u)$ and $\mathcal{R}_{k-2}^\infty(\underline{u}, u)$ to denote the sum over all possible R or Γ of the corresponding norms. Finally, we define total norms

$$\mathcal{R}_k = \sup_u \mathcal{R}_k(u), \mathcal{O}_{k+1} = \sup_u \mathcal{O}_{k+1}(u), \mathcal{R}_{\leq k} = \sum_{j \leq k} \mathcal{R}_j, \mathcal{O}_{\leq k} = \sum_{j \leq k} \mathcal{O}_j.$$

We state the main result in [2]. Roughly speaking, it asserts that up to two derivatives on curvature, all the norms ($k \leq 2$) defined above can propagate along the evolution of Einstein vacuum equations.

CHRISTODOULOU’S MAIN ESTIMATES ([2, Th. 16.1]). *If $\delta > 0$ is sufficiently small (depending on the C^8 bound of the seed data ψ_0), there exists a unique solution for the Einstein field equations on a domain M corresponding to $0 \leq \underline{u} \leq \delta$ and $u_0 \leq u \leq -1 - \delta$. Moreover, the following total norms,*

$$\mathcal{R}_{\leq 2}, \underline{\mathcal{R}}_{\leq 2}[\underline{\alpha}], \mathcal{R}_{\leq 1}^4, \mathcal{R}_0^\infty, \mathcal{O}_{\leq 3}, \mathcal{O}_{\leq 2}^4, \mathcal{O}_{\leq 1}^\infty,$$

are bounded by a constant depending on the C^8 bound of the seed data ψ_0 .

The main consequence of this theorem is the dynamical formation of trapped surfaces. In [2], Christodoulou showed that if

$$\frac{1}{8} \int_0^1 \left| \frac{\partial \psi_0}{\partial s}(s, \theta) \right|^2 ds > 1,$$

uniformly in θ , then a trapped surface forms in the future of C_{u_0} .

2.3. *The work of Corvino-Schoen.* We follow closely the notation used by Corvino and Schoen in [5] unless there are conflicts with the current work. We first recall some definitions on a given three-dimensional space-like slice Σ in a vacuum space-time. The vacuum initial data on Σ consist of a Riemannian metric \bar{g} and a symmetric 2-tensor \bar{k} subject to the following constraints (to distinguish from the notation in four dimensions, we shall always use barred

notation in three dimensions, e.g., $\bar{\nabla}$ denotes the Levi-Civita connection associated to \bar{g}):

$$\begin{aligned} R(\bar{g}) - |\bar{k}|^2 + \bar{h}^2 &= 0, \\ \operatorname{div}_{\bar{g}} \bar{k} - d\bar{h} &= 0, \end{aligned}$$

where $\bar{h} = \operatorname{tr}_{\bar{g}} \bar{k} = \bar{g}_{ij} \bar{k}^{ij}$ denotes the mean curvature, $\operatorname{div}_{\bar{g}} \bar{k}_i = \bar{\nabla}_j \bar{k}^{ij}$, and all quantities are computed with respect to \bar{g} . One then rewrites the constraint equations by introducing the momentum tensor $\bar{\pi}_{ij} = \bar{k}_{ij} - \bar{h} \cdot \bar{g}_{ij}$. Let \mathcal{H} and Φ be the following maps:

$$\begin{aligned} \mathcal{H}(\bar{g}, \bar{\pi}) &= R_{\bar{g}} + \frac{1}{2}(\operatorname{tr} \bar{\pi})^2 - |\bar{\pi}|^2, \\ \Phi(\bar{g}, \bar{\pi}) &= (\mathcal{H}(\bar{g}, \bar{\pi}), \operatorname{div}_{\bar{g}} \bar{\pi}). \end{aligned}$$

The vacuum constraints then take the form $\Phi(\bar{g}, \bar{\pi}) = 0$.

We use $\mathcal{M}^{k,\alpha}(\Sigma)$, $\mathcal{S}^{k,\alpha}(\Sigma)$ and $\mathcal{X}^{k,\alpha}(\Sigma)$ to denote the set of Riemannian metrics, symmetric two tensors and vector fields on Σ with $C^{k,\alpha}$ regularity respectively. Thus, we have

$$\Phi : \mathcal{M}^{k+2,\alpha}(\Sigma) \times \mathcal{S}^{k+2,\alpha}(\Sigma) \rightarrow C^{k,\alpha}(\Sigma) \times \mathcal{X}^{k+1,\alpha}(\Sigma).$$

The formal L^2 -adjoint operator $D\Phi_{(\bar{g}, \bar{\pi})}^*$ of the linearization $D\Phi_{(\bar{g}, \bar{\pi})}$ is then given by

$$\begin{aligned} (2.29) \quad D\mathcal{H}_{(\bar{g}, \bar{\pi})}^*(f) &= \left((L_{\bar{g}}^* f)_{ij} + (\operatorname{tr}_{\bar{g}} \bar{\pi} \cdot \bar{\pi}_{ij} - 2\bar{\pi}_{ik} \bar{\pi}^k_j) f, (\operatorname{tr}_{\bar{g}} \bar{\pi} \cdot \bar{g}_{ij} - 2\bar{\pi}_{ij}) f \right), \\ D\operatorname{div}_{(\bar{g}, \bar{\pi})}^*(X) &= \frac{1}{2} \left(\mathcal{L}_X \bar{\pi}_{ij} + \bar{\nabla}_k X^k \pi_{ij} - (X_i (\bar{\nabla}_k \pi^k_j + X_j \bar{\nabla}_k \pi^k_i) \right. \\ &\quad \left. - (\bar{\nabla}_m X_k \pi^{km} + X_k \bar{\nabla}_m \pi^{mk}) g_{ij}, -\mathcal{L}_X g_{ij} \right), \end{aligned}$$

where $L_{\bar{g}}^* f = -\Delta_{\bar{g}} f + \bar{\nabla}^2 f - f \cdot \operatorname{Ric}(\bar{g})$ is the formal L^2 -adjoint of the linearization of the scalar curvature operator.

Let $\Omega \subset \Sigma$ be a given bounded domain with smooth boundary and ρ be a smooth positive function on Ω that near Ω decays as a power of distance to the boundary; i.e., $\rho \sim d^N$ where d is the distance to the boundary and N will be fixed later. (See the theorem below.) The weighted Hölder space $C_{\rho^{-1}}^{k,\alpha}(\Omega)$ is defined by the norm $\|f\|_{C_{\rho^{-1}}^{k,\alpha}} = \|f\rho^{-\frac{1}{2}}\|_{C^{k,\alpha}}$ in the obvious way; similarly, we can define those spaces for tensors. We are ready to state the local deformation theorem, which will play a key role in our gluing construction.

LOCAL DEFORMATION THEOREM (Theorem 2 in [5]). *Let $\zeta \in C_0^\infty(\Omega)$ be a bump function and $(g_0, \pi_0) \in \mathcal{M}^{k+4,\alpha}(\Sigma) \times \mathcal{S}^{k+3,\alpha}(\Sigma)$. Then for N sufficiently large, there is an $\varepsilon > 0$ such that for all pairs $(v, W) \in C^{k,\alpha}(\bar{\Omega}) \times \mathcal{X}^{k+1,\alpha}(\bar{\Omega})$ with the support of $(v, W) - \Phi(g_0, \pi_0)$ contained in $\bar{\Omega}$ and with $\|(v, W) - \Phi(g_0, \pi_0)\|_{C_{\rho^{-1}}^{k,\alpha}(\bar{\Omega}) \times \mathcal{X}_{\rho^{-1}}^{k+1,\alpha}(\bar{\Omega})} < \varepsilon$, there is a pair $(g, \pi) \in \mathcal{M}^{k+2,\alpha}(\Sigma) \times$*

$\mathcal{S}^{k+2,\alpha}(\Sigma)$, such that $\Phi(g, \pi) - (v, W) \in \zeta \cdot \text{Ker } D\Phi_{(g_0, \pi_0)}^*$ in Ω and $(g, \pi) = (g_0, \pi_0)$ outside Ω . Moreover, $(g, \pi) \in \mathcal{M}^{k+2,\alpha}(\Sigma) \times \mathcal{S}^{k+2,\alpha}(\Sigma)$ depends continuously on $(v, W) - \Phi(g_0, \pi_0) \in C_{\rho^{-1}}^{k,\alpha}(\bar{\Omega}) \times \mathcal{X}_{\rho^{-1}}^{k+1,\alpha}(\bar{\Omega})$.

In [5], Corvino and Schoen used this theorem to approximate asymptotically flat initial data for the vacuum Einstein field equations by solutions that agree with the original data inside a given domain and are identical to that of a suitable Kerr slice outside a compact set. We will use this theorem in a similar manner in a different situation to prove our main theorem.

2.4. *The structure of the proof.* This section is devoted to an outline of the proof. It consists of three steps.

Step 1: *Higher order energy estimates.*

As we mentioned in the introduction, since we would like to use the Corvino-Schoen construction (see the Local Deformation Theorem in Section 2.3), which requires certain regularity of the space-time, we are obliged to derive higher order energy estimates. This is accomplished in Section 3.

We will derive the estimates on M where $0 \leq \underline{u} \leq \delta$. Christodoulou’s work [2] already proved such estimates on the level of two derivatives on the curvature. (See the Christodoulou Main Estimates in Section 2.2.) His estimates are already good enough to construct the space-time. Thus, in our case, we can use an induction (on the number of derivatives) argument instead of a bootstrap argument.

Although this part should be regarded as the routine proof of the persistence of regularity for vacuum Einstein field equations, we would like to emphasize that by affording more derivatives, we can gain smallness in δ as compared to Christodoulou’s work. For example, in [2] we have

$$|\nabla \eta| \lesssim \frac{1}{|u|^3}$$

while, in the current work, since we can control the third derivatives on the curvature components, we actually have

$$|\nabla \eta| \lesssim \frac{1}{|u|^3} \delta^{\frac{1}{2}}.$$

The gain in δ will play a crucial role in Step 2. Here the notation $A \lesssim B$ should be understood as $A \leq CB$, where C is a constant depending only on the seed data.

Step 2: *Construction of a transition region close to the Schwarzschild space-time with mass m_0 .*

This step is the main innovation of the paper, and it is complete in Section 4. Roughly speaking, this step builds a bridge from Christodoulou’s work

to the Corvino-Schoen construction. By solving the vacuum Einstein field equations, we can construct a region with a fixed size (independent of the small parameter δ) that is close (measured by δ) to a region in the domain of outer communication in the Schwarzschild space-time with mass m_0 . It is precisely on this region, or more precisely a space-like hypersurface inside this region, where we can use the Corvino-Schoen construction.

As we mentioned in Remark 3, in addition to Christodoulou's short pulse ansatz described in Section 2.2, we need to impose one more condition on the initial data defined on $C_{u_0}^{[0,\delta]}$; that is, for all θ ,

$$(2.30) \quad \int_0^1 \left| \frac{\partial \psi_0}{\partial s}(s, \theta) \right|^2 ds = 16m_0.$$

This condition is the key ingredient to Step 2, and most of the estimates are directly tied to it.

Remark 4. This condition is slightly different from the condition proposed in Remark 1; namely,

$$(2.31) \quad |u_0|^2 \int_0^\delta |\widehat{\chi}(\underline{u}, u_0, \theta)|_{\underline{g}}^2 d\underline{u} = 4m_0.$$

The advantage of using condition (2.30) over (2.31) is that as δ changes, (2.30) is always valid, but (2.31) is not. However, under condition (2.30), (2.31) is valid up to an error of size δ to some power. (See Lemma 1 in the beginning of Section 4 for a precise statement.) We emphasize that, mathematically, we should impose condition (2.30) via the seed data ψ_0 in order to follow the short pulse ansatz described in Section 2.2. On the other hand, though only valid up to an error, condition (2.31) has a clear physical interpretation: it requires the incoming gravitational energy per solid angle in the advanced time interval $[0, \delta]$ to be the same for all angles θ . Roughly speaking, we impose certain spherical symmetry on the initial data.

In Section 4.1, we show that the sphere S_{δ, u_0} is δ -close to a given sphere in the Schwarzschild space-time with mass m_0 ; here, the closeness is measured in C^k norms. The proof is based on condition (2.30).

In Section 4.2, we show that the incoming cone \underline{C}_δ is δ -close to a given incoming cone in the Schwarzschild space-time with mass m_0 . This cone is rooted on the given sphere S_{δ, u_0} from Section 4.1. The proof is based on condition (2.30) as well as the higher order energy estimates derived in Section 3.

In Section 4.3, we further extend our data on $C_{u_0}^{[\delta, \delta+1]}$ smoothly by setting

$$(2.32) \quad \widehat{\chi} \equiv 0, \quad \delta \leq \underline{u} \leq \delta + 1.$$

We then show that the outgoing cone $C_{u_0}^{[\delta, \delta+1]}$ is δ -close to a given outgoing cone in the Schwarzschild space-time with mass m_0 . This cone is also rooted on the given sphere S_{δ, u_0} from Section 4.1. The proof is based on condition (2.30).

In Section 4.4, since the Sections 4.1, 4.2 and 4.3 provide a characteristic data set for the vacuum Einstein field equations, we will solve the field equations to further extend Christodoulou's solution. By virtue of the closeness to the Schwarzschild data, we show that the resulting space-time is close to a region in the domain of outer communication in the Schwarzschild space-time with mass m_0 . Furthermore, we will also see that this region has a fixed size; namely, it will not shrink to zero when we decrease the parameter δ .

Step 3: *Gluing a Kerr slice.*

This step is completed in Section 5. As a consequence of Section 4.4, we can choose a three-dimensional space-like region for gluing. Then we follow the procedure in [4] and [5]. Thanks to the closeness derived in Step 2, we can almost explicitly write down the kernel of the adjoint of the linearized operator of the constraint map. This allows one to use the Local Deformation Theorem in Section 2.3. The kernel will have four dimensions. Combined with a fixed point argument, we use a four-parameter family of Kerr space-times, varying the mass m and the angular momentum \mathbf{a} to kill the kernel. We also show that the resulting initial data is free of trapped surfaces.

3. Higher order energy estimates

In this section, we derive energy estimates for higher order derivatives on curvature components. At the same time, this also yields the control of higher order derivatives on connection coefficients. More precisely, we shall prove

THEOREM 1. *Let $k \in \mathbb{Z}_{\geq 3}$. For $N \in \mathbb{Z}_+$ sufficiently large, if $\delta > 0$ is sufficiently small depending C^{k+N} norm of the seed data ψ_0 , then the following quantities*

$$\mathcal{R}_{\leq k}, \underline{\mathcal{R}}_{\leq k}[\underline{\alpha}], \mathcal{R}_{\leq k-1}^4, \mathcal{R}_{\leq k-2}^\infty, \mathcal{O}_{\leq k+1}, \mathcal{O}_{\leq k}^4, \mathcal{O}_{\leq k-1}^\infty,$$

are bounded by a constant depending on C^{k+N} bounds of the seed data.

The integer N is chosen such that for all k , $\mathcal{R}_k(u_0)$ is bounded by a constant depending on the C^{k+N} bound of the seed data. We remark that for $k \leq 2$, the theorem was proved in [2] with $N = 6$. In fact the precise value of N is not important in the current work. Then we shall proceed by induction on k . We then make the induction assumption as follows: for $k \geq 3$,

$$(3.1) \quad \mathcal{R}_{\leq k-1} + \underline{\mathcal{R}}_{\leq k-1}[\underline{\alpha}] + \mathcal{R}_{\leq k-2}^4 + \mathcal{R}_{\leq k-3}^\infty + \mathcal{O}_{\leq k} + \mathcal{O}_{\leq k-1}^4 + \mathcal{O}_{\leq k-2}^\infty \leq F_{k-1+N},$$

where F_i depends only on the C^i bound of seed data.

3.1. Sobolev inequalities and elliptic estimates for Hodge systems. In this section, we recall the elliptic estimates for Hodge systems on $S_{\underline{u}, u}$. Together with Sobolev inequalities, this will serve as the basic tool for us to control the L^∞ norms. We refer to [2] for the proof and we shall take $\delta > 0$ to be

sufficiently small as in [2]. For the sake of simplicity, we use S to denote $S_{u,u}$, C to denote C_u and \underline{C} to denote \underline{C}_u .

We first collect all the Sobolev inequalities as follows: equation

$$(3.2) \quad \|\xi\|_{L^4(S)} \leq C \left(|u|^{\frac{1}{2}} \|\nabla \xi\|_{L^2(S)} + |u|^{-\frac{1}{2}} \|\xi\|_{L^2(S)} \right),$$

$$(3.3) \quad \|\xi\|_{L^\infty(S)} \leq C \left(|u|^{\frac{1}{2}} \|\nabla \xi\|_{L^4(S)} + |u|^{-\frac{1}{2}} \|\xi\|_{L^4(S)} \right),$$

$$(3.4) \quad \sup_{\underline{u}} (|u|^{\frac{1}{2}} \|\xi\|_{L^4(S)}) \\ \leq C \left(\|\xi\|_{L^4(S)} + \|D\xi\|_{L^2(\underline{C})}^{\frac{1}{2}} (\|\xi\|_{L^2(\underline{C})}^{\frac{1}{2}} + |u|^{\frac{1}{2}} \|\nabla \xi\|_{L^2(\underline{C})}^{\frac{1}{2}}) \right),$$

$$(3.5) \quad \sup_u (|u|^q \|\xi\|_{L^4(S)}) \leq C \left(|u_0|^q \|\xi\|_{L^4(S_{u_0, u_0})} \right. \\ \left. + \| |u|^q D\xi \|_{L^2(\underline{C})}^{\frac{1}{2}} (\| |u|^{q-1} \xi \|_{L^2(\underline{C})}^{\frac{1}{2}} + \| |u|^q \nabla \xi \|_{L^2(\underline{C})}^{\frac{1}{2}}) \right),$$

where C is a universal constant. We remark that q in the last inequality can be any real number as long as the right-hand side is finite. The disparity of the powers of $|u|$ reflects the fact that the area of $S_{u,u}$ is of order $|u|^2$. The proof of the above inequalities is routine, we refer readers to Lemma 5.1, Lemma 5.2, Proposition 10.1 and Proposition 10.2 in [2] for details.

For Hodge systems, we have the following two cases. First of all, if θ is a traceless symmetric two tensor, such that

$$\text{div} \theta = f,$$

where f is a one-form on S , then

$$(3.6) \quad \|\nabla^{k+1} \theta\|_{L^2(S)} \leq C \left(\|\nabla^k f\|_{L^2(S)} + |u|^{-1} \|\nabla^{k-1} f\|_{L^2(S)} \right. \\ \left. + |u|^{-1} \|\nabla^k \theta\|_{L^2(S)} + |u|^{-2} \|\nabla^{k-1} \theta\|_{L^2(S)} \right. \\ \left. + \sum_{i=1}^{k-1} (\|\nabla^i K\|_{L^2(S)} + |u|^{-1} \|\nabla^{i-1} K\|_{L^2(S)}) \|\nabla^{k-1-i} \theta\|_{L^\infty} \right).$$

Secondly, for a one-form ξ that satisfies

$$\text{div} \xi = f, \quad \text{curl} \xi = g,$$

we have

$$(3.7) \quad \|\nabla^{k+1} \xi\|_{L^2(S)} \leq C \left(\|\nabla^k f\|_{L^2(S)} + |u|^{-1} \|\nabla^{k-1} f\|_{L^2(S)} \right. \\ \left. + \|\nabla^k g\|_{L^2(S)} + |u|^{-1} \|\nabla^{k-1} g\|_{L^2(S)} \right. \\ \left. + |u|^{-1} (\|\nabla^k \xi\|_{L^2(S)} + |u|^{-1} \|\nabla^{k-1} \xi\|_{L^2(S)}) \right. \\ \left. + \sum_{i=1}^{k-1} (\|\nabla^i K\|_{L^2(S)} + |u|^{-1} \|\nabla^{i-1} K\|_{L^2(S)}) \|\nabla^{k-1-i} \xi\|_{L^\infty} \right).$$

In both inequalities, C is a universal constant. We refer readers to Lemmas 7.3 and 7.4 in [2] for the above estimates.

Finally, we also collect two Gronwall type estimates for later uses. They can be found in Lemma 4.6 and Lemma 4.7 in [2]. If we have

$$D\theta = \frac{\nu}{2}\Omega \operatorname{tr} \chi \theta + \gamma \cdot \theta + \xi,$$

where θ , ξ are r -covariant tangential tensorfields, γ is (r, r) type tangential tensorfield so that $|\gamma| \leq m\Omega|\widehat{\chi}|$, ν and m are constants, then for any $p \geq 1$, there is a constant $C > 0$ depending only on p , ν and m such that

$$\|\theta\|_{L^p(S_{\underline{u}, u})} \leq C \left(\|\theta\|_{L^p(S_{0, u})} + \int_0^u \|\xi\|_{L^p(S_{\underline{u}', u})} d\underline{u}' \right).$$

If we have

$$\underline{D}\underline{\theta} = \frac{\nu}{2}\Omega \operatorname{tr} \underline{\chi} \underline{\theta} + \underline{\gamma} \cdot \underline{\theta} + \underline{\xi}$$

where $\underline{\theta}$, $\underline{\xi}$ are r -covariant tangential tensorfields, $\underline{\gamma}$ is (r, r) type tangential tensorfield such that $|\underline{\gamma}| \leq m\Omega|\underline{\widehat{\chi}}|$, ν and m are constants, then for any $p \geq 1$, there is a constant $C > 0$ depending only on p , r , ν , m such that

$$|u|^{r-\nu-\frac{2}{p}} \|\underline{\theta}\|_{L^p(S_{\underline{u}, u})} \leq C \left(|u_0|^{r-\nu-\frac{2}{p}} \|\underline{\theta}\|_{L^p(S_{\underline{u}, u_0})} + \int_{u_0}^u |u'|^{r-\nu-\frac{2}{p}} \|\underline{\xi}\|_{L^p(S_{\underline{u}, u'})} d\underline{u}' \right).$$

3.2. Estimates for connection coefficients.

3.2.1. *Estimates for $\nabla^{k+1}\widehat{\chi}'$ and $\nabla^{k+1}\operatorname{tr} \chi'$.* We first consider $\nabla^{k+1}\widehat{\chi}'$ and $\nabla^{k+1}\operatorname{tr} \chi'$. We will give all the details for these two terms. In the following sections, when we estimate the other connection coefficients, i.e., $\nabla^{k+1}\underline{\widehat{\chi}}'$, $\nabla^{k+1}\operatorname{tr} \underline{\chi}'$, $\nabla^{k+1}\eta$, $\nabla^{k+1}\underline{\eta}$, $\nabla^{k+1}\omega$ and $\nabla^{k+1}\underline{\omega}$, since the proofs follow a similar pattern, we will be sketchy and only emphasize the key points.

In view of the terms on the right-hand side of (2.10), by induction hypothesis (3.1), we compute

$$\begin{aligned} \|\nabla^k(\widehat{\chi}' \cdot \eta)\|_{L^2(C_u)} &\lesssim \|\nabla^k \widehat{\chi}'\|_{L^2(C_u)} \|\eta\|_{L^\infty} \\ &\quad + \int_0^u \sum_{i=1}^{k-1} \|\nabla^i \widehat{\chi}'\|_{L^4(S_{\underline{u}', u})} \|\nabla^{k-1-i} \eta\|_{L^4(S_{\underline{u}', u})} d\underline{u}' \\ &\quad + \|\widehat{\chi}'\|_{L^\infty} \|\nabla^k \eta\|_{L^2(C_u)} \lesssim \delta^{-\frac{1}{2}} |u|^{-3-k} \cdot \delta^{\frac{1}{2}} |u|. \end{aligned}$$

We remark that the factor $\delta^{\frac{1}{2}}|u|$ comes from the norms $\|\cdot\|_{L^2(C_u)}$. Throughout this section, the notation $A \lesssim B$ should be understood as $A \leq CB$ where C is a constant depending on induction hypothesis (3.1). Similarly, we have

$$\|\nabla^k(\operatorname{tr} \chi' \eta)\|_{L^2(C_u)} \lesssim |u|^{-3-k} \cdot \delta^{\frac{1}{2}} |u|.$$

If one replaces k by $k-1$, it is obvious that the above estimates also hold. We denote $I = -\widehat{\chi}' \cdot \eta + \frac{1}{2} \operatorname{tr} \chi' \eta - \Omega^{-1} \beta$, i.e., those terms in (2.10). In view of the fact that $\nabla^i \not{d} \log \Omega = \frac{1}{2} \nabla^i (\eta + \underline{\eta})$, we have

$$(3.8) \quad \|\nabla^k I\|_{L^2(C_u)} + |u|^{-1} \|\nabla^{k-1} I\|_{L^2(C_u)} \lesssim \|\nabla^k \beta\|_{L^2(C_u)} + \delta^{-\frac{1}{2}} |u|^{-3-k} \cdot \delta^{\frac{1}{2}} |u|.$$

In view of (3.6) and (3.7), we turn to the estimates on the Gauss curvature K . For $i = 0, \dots, k-1$, by (2.9),

$$\begin{aligned} \|\nabla^i K\|_{L^2(C_u)} &\lesssim \|\nabla^i \rho\|_{L^2(C_u)} + \|\nabla^i (\widehat{\chi}, \widehat{\underline{\chi}})\|_{L^2(C_u)} \\ &\quad + \|\nabla^i (\operatorname{tr} \chi \operatorname{tr} \underline{\chi})\|_{L^2(C_u)} \lesssim |u|^{-2-i} \cdot \delta^{\frac{1}{2}} |u|. \end{aligned}$$

Therefore,

$$\begin{aligned} \sum_{i=1}^{k-1} (\|\nabla^i K\|_{L^2(C_u)} + |u|^{-1} \|\nabla^{i-1} K\|_{L^2(C_u)}) \|\nabla^{k-1-i} \widehat{\chi}'\|_{L^\infty(S)} \\ \lesssim \delta^{-\frac{1}{2}} |u|^{-2-k} \cdot \delta^{\frac{1}{2}} |u|. \end{aligned}$$

According to (2.10), (3.6) and the inequality $\int_0^u \|\cdot\|_{L^2(S_{\underline{u}',u})} \leq \delta^{\frac{1}{2}} \|\cdot\|_{L^2(C_u)}$, we obtain

$$(3.9) \quad \begin{aligned} \int_0^u \|\nabla^{k+1} \widehat{\chi}'\|_{L^2(S_{\underline{u}',u})} d\underline{u}' &\lesssim \int_0^u \|\nabla^{k+1} \operatorname{tr} \chi'\|_{L^2(S_{\underline{u}',u})} d\underline{u}' \\ &\quad + \delta^{\frac{1}{2}} \|\nabla^k \beta\|_{L^2(C_u)} + \delta^{-\frac{1}{2}} |u|^{-2-k} \cdot \delta |u|. \end{aligned}$$

We turn to (2.2); by commuting derivatives, we have

(3.10)

$$\begin{aligned} D \nabla^{k+1} \operatorname{tr} \chi' &= [D, \nabla^k] \not{d} \operatorname{tr} \chi' - \Omega \operatorname{tr} \chi \nabla^{k+1} \operatorname{tr} \chi' - 2\Omega \widehat{\chi} \otimes \nabla^{k+1} \widehat{\chi}' \\ &\quad - \sum_{i=0}^{k-1} \nabla^i \not{d} (\Omega \operatorname{tr} \chi) \otimes \nabla^{k-1-i} \not{d} \operatorname{tr} \chi' \\ &\quad - \sum_{i=0}^{k-1} \nabla^i \nabla (\Omega \widehat{\chi}) \otimes \nabla^{k-1-i} \nabla \widehat{\chi}' \\ &\quad - \sum_{\substack{i+j+h=k \\ i,j,h \geq 0}} \nabla^i \not{d} \log \Omega \otimes (\nabla^j \operatorname{tr} \chi \otimes \nabla^h \operatorname{tr} \chi + 2\nabla^j \widehat{\chi} \otimes \nabla^h \widehat{\chi}'). \end{aligned}$$

In fact, the commutator $[D, \nabla^i]$ can be written as

$$[D, \nabla^i] \theta = \sum_{j=0}^{i-1} \nabla^j \nabla (\Omega \chi) \cdot \nabla^{i-1-j} \theta,$$

so

$$[D, \nabla^k] \not{d} \operatorname{tr} \chi' = \sum_{i=0}^{k-1} \nabla^i \nabla (\Omega \chi) \cdot \nabla^{k-1-i} \not{d} \operatorname{tr} \chi'.$$

Since $\nabla^i \not{d} \log \Omega = \frac{1}{2} \nabla^i (\eta + \underline{\eta})$, we rewrite (3.10) as

$$D\nabla^{k+1} \operatorname{tr} \chi' = -\Omega \operatorname{tr} \chi \nabla^{k+1} \operatorname{tr} \chi' - 2\Omega \widehat{\chi} \otimes \nabla^{k+1} \widehat{\chi}' + \text{l.o.t.}$$

Remark 5. The notation l.o.t. here denotes the terms that contain at most k -th order derivatives of connection coefficients and $(k-1)$ -st order derivatives of curvature components. These terms can be estimated by induction hypothesis (3.1). This notation has the same meaning in this subsection about estimates for connection coefficients. We will also use this notation in other parts of this paper with slightly different meaning, which will be mentioned if necessary.

Therefore,

$$(3.11) \quad \|\nabla^{k+1} \operatorname{tr} \chi'\|_{L^2(S_{\underline{u}, u})} \lesssim \delta^{-\frac{1}{2}} |u|^{-1} \int_0^u \|\nabla^{k+1} \widehat{\chi}'\|_{L^2(S_{\underline{u}', u})} d\underline{u}' + |u|^{-3-k} \cdot |u|.$$

Combining (3.9) and (3.11), we obtain

$$\begin{aligned} \|\nabla^{k+1} \operatorname{tr} \chi'\|_{L^2(S_{\underline{u}, u})} &\lesssim \delta^{-\frac{1}{2}} |u|^{-1} \int_0^u \|\nabla^{k+1} \operatorname{tr} \chi'\|_{L^2(S_{\underline{u}', u})} d\underline{u}' \\ &\quad + |u|^{-1} \|\nabla^k \beta\|_{L^2(C_u)} + |u|^{-3-k} \cdot |u|. \end{aligned}$$

Thus,

$$(3.12) \quad \|\nabla^{k+1} \operatorname{tr} \chi'\|_{L^2(S_{\underline{u}, u})} \lesssim |u|^{-3-k} (\mathcal{R}_k[\beta] + C) \cdot |u|.$$

Finally, according to (3.8), (2.10) and (3.6), we obtain

$$(3.13) \quad \begin{aligned} \|\nabla^{k+1} \widehat{\chi}'\|_{L^2(C_u)} &\lesssim \|\nabla^{k+1} \operatorname{tr} \chi'\|_{L^2(C_u)} + \|\nabla^k \beta\|_{L^2(C_u)} + \delta^{-\frac{1}{2}} |u|^{-2-k} \cdot \delta^{\frac{1}{2}} |u| \\ &\lesssim \delta^{-\frac{1}{2}} |u|^{-2-k} (\mathcal{R}_k[\beta] + C) \cdot \delta^{\frac{1}{2}} |u|. \end{aligned}$$

3.2.2. *Smallness on $\nabla^k(\eta, \underline{\eta})$.* In order to estimate $\nabla^{k+1} \widehat{\chi}'$ and $\nabla^{k+1} \operatorname{tr} \chi'$, we must improve the estimates on $\|\nabla^k(\eta, \underline{\eta})\|_{L^2(C_u)}$. We also mentioned this in the introduction. We can actually afford one more derivative to gain this smallness in powers of δ . We also remark that at this stage, we can only prove that $\|\nabla^{k+1}(\eta, \underline{\eta})\|_{L^2(C_u)}$ is bounded without any smallness. We commute ∇^k with (2.5) to derive

$$(3.14) \quad \begin{aligned} D\nabla^k \eta &= \Omega \chi \nabla^k \underline{\eta} + \sum_{i=0}^{k-1} \left(\nabla^i \nabla(\Omega \chi) \nabla^{k-1-i} \eta \right. \\ &\quad \left. + \nabla^i \nabla(\Omega \chi) \otimes \nabla^{k-1-i} \underline{\eta} \right) - \nabla^k(\Omega \beta). \end{aligned}$$

Therefore,

$$(3.15) \quad \|\nabla^k \eta\|_{L^2(S_{\underline{u}, u})} \lesssim \delta^{-\frac{1}{2}} |u|^{-1} \int_0^u \|\nabla^k \underline{\eta}\|_{L^2(S_{\underline{u}', u})} d\underline{u}' + \delta^{\frac{1}{2}} \|\nabla^k \beta\|_{L^2(C_u)} + \delta |u|^{-3-k} \cdot |u|.$$

Similarly, according to (2.6), we have

(3.16)

$$\underline{D}\nabla^k \underline{\eta} = \Omega_{\underline{\chi}} \nabla^k \eta + \sum_{i=0}^{k-1} (\nabla^i \nabla (\Omega_{\underline{\chi}}) \nabla^{k-1-i} \underline{\eta} + \nabla^i \nabla (\Omega_{\underline{\chi}}) \otimes \nabla^{k-1-i} \eta) + \nabla^k (\Omega_{\underline{\beta}}),$$

which implies

(3.17)

$$\begin{aligned} |u|^k \|\nabla^k \underline{\eta}\|_{L^2(C_u)} &\lesssim |u_0|^k \|\nabla^k \underline{\eta}\|_{L^2(C_{u_0})} + \int_{u_0}^u |u'|^k \cdot |u'|^{-1} \|\nabla^k \eta\|_{L^2(C_{u'})} du' \\ &\quad + \int_{u_0}^u |u'|^k \|\nabla^k \underline{\beta}\|_{L^2(C_{u'})} du' + \int_{u_0}^u |u'|^k \cdot \delta |u'|^{-4-k} \cdot \delta^{\frac{1}{2}} |u'| du'. \end{aligned}$$

We then substitute (3.17) into (3.15) and take supremum on \underline{u} . As a result, we have

$$\begin{aligned} \sup_{\underline{u}} |u|^k \|\nabla^k \eta\|_{L^2(S_{\underline{u},u})} &\lesssim |u_0|^k |u|^{-1} \|\nabla^k \underline{\eta}\|_{L^2(C_{u_0})} \\ &\quad + \delta^{\frac{1}{2}} |u|^{-1} \int_{u_0}^u \sup_{\underline{u}} |u'|^{k-1} \|\nabla^k \eta\|_{L^2(S_{\underline{u},u'})} du' \\ &\quad + \delta^{\frac{3}{2}} |u|^{-4} (\mathcal{R}_k[\underline{\beta}] + C) \cdot |u| + \delta^{\frac{1}{2}} |u|^{-2} (\mathcal{R}_k[\beta] + \delta^{\frac{1}{2}} |u|^{-1} C) \cdot |u|. \end{aligned}$$

Thus,

$$|u|^k \|\nabla^k \eta\|_{L^2(S_{\underline{u},u})} \lesssim |u_0|^k |u|^{-1} \|\nabla^k \underline{\eta}\|_{L^2(C_{u_0})} + \delta^{\frac{1}{2}} |u|^{-2} (\mathcal{R}_k + C) \cdot |u|.$$

Recall that $\eta = -\underline{\eta}$ on C_{u_0} . In view of (3.15), we have

$$|u_0|^k \|\nabla^k \eta\|_{L^2(S_{\underline{u},u_0})} \lesssim \delta^{\frac{1}{2}} |u_0|^{-2} (\mathcal{R}_k[\beta] + \delta^{\frac{1}{2}} |u_0|^{-1} C) \cdot |u_0|.$$

Combining all the estimates above, we finally obtain

$$\|\nabla^k \eta\|_{L^2(S_{\underline{u},u})}, \delta^{-\frac{1}{2}} \|\nabla^k \underline{\eta}\|_{L^2(C_u)} \lesssim \delta^{\frac{1}{2}} |u|^{-2-k} (\mathcal{R}_k + C) \cdot |u|.$$

3.2.3. *Estimates for $\nabla^{k+1} \widehat{\underline{\chi}}'$ and $\nabla^{k+1} \text{tr } \underline{\chi}'$.* The estimates will be derived in a similar way as for $\nabla^{k+1} \widehat{\underline{\chi}}$ and $\nabla^{k+1} \text{tr } \underline{\chi}$. By (2.11) and the above improved smallness on $\nabla^k \underline{\eta}$, we have

$$(3.18) \quad \begin{aligned} \|\nabla^{k+1} \widehat{\underline{\chi}}'\|_{L^2(C_u)} &\lesssim \|\nabla^{k+1} \text{tr } \underline{\chi}'\|_{L^2(C_u)} \\ &\quad + \|\nabla^k \underline{\beta}\|_{L^2(C_u)} + \delta^{\frac{1}{2}} |u|^{-3-k} (\mathcal{R}_k + C) \cdot \delta^{\frac{1}{2}} |u|. \end{aligned}$$

We then commute ∇^{k+1} with (2.4) to derive

$$\underline{D}\nabla^{k+1} \text{tr } \underline{\chi}' = -\Omega \text{tr } \underline{\chi} \nabla^{k+1} \text{tr } \underline{\chi}' - 2\Omega \widehat{\underline{\chi}} \otimes \nabla^{k+1} \widehat{\underline{\chi}}' + (\text{tr } \underline{\chi})^2 \nabla^{k+1} \log \Omega + \text{l.o.t.}$$

Although $\nabla^{k+1} \log \Omega = \frac{1}{2} \nabla^k (\eta + \underline{\eta})$ is a lower order term (in derivative), this naive estimate will not yield the expected smallness of $\nabla^{k+1} \text{tr } \underline{\chi}'$. To remedy this defect, we introduce an auxiliary bootstrap assumption as follows:

AUXILIARY BOOTSTRAP ASSUMPTION. For some large constant Δ_k depending on \mathcal{R}_k ,

$$(3.19) \quad \|\nabla^{k+1} \log \Omega\|_{L^2(C_u)} \leq \delta |u|^{-3-k} \Delta_k \cdot \delta^{\frac{1}{2}} |u|.$$

Therefore,

$$\begin{aligned} |u|^{k+2} \|\nabla^{k+1} \text{tr } \underline{\chi}'\|_{L^2(C_u)} &\lesssim |u_0|^{k+2} \|\nabla^{k+1} \text{tr } \underline{\chi}'\|_{L^2(C_{u_0})} \\ &+ \int_{u_0}^u |u'|^{k+2} \cdot \delta^{\frac{1}{2}} |u'|^{-2} \|\nabla^{k+1} \text{tr } \underline{\chi}'\|_{L^2(C_{u'})} du' \\ &+ \int_{u_0}^u |u'|^{k+2} \cdot \delta^{\frac{1}{2}} |u'|^{-2} \|\nabla^k \underline{\beta}\|_{L^2(C_{u'})} du' + \delta |u|^{-2} (\mathcal{R}_k + \Delta_k + C) \cdot \delta^{\frac{1}{2}} |u|. \end{aligned}$$

According to (2.15), we have $|u_0|^{k+1} \|\nabla^{k+1} \text{tr } \underline{\chi}'\|_{L^2(C_{u_0})} \lesssim \delta |u_0|^{-3} C \cdot \delta^{\frac{1}{2}} |u_0|$, so

$$(3.20) \quad \|\nabla^{k+1} \text{tr } \underline{\chi}'\|_{L^2(C_u)} \lesssim \delta |u|^{-4-k} (\mathcal{R}_k + \Delta_k + C) \cdot \delta^{\frac{1}{2}} |u|.$$

Thanks to (3.18), we obtain

$$(3.21) \quad \|\nabla^{k+1} \underline{\chi}'\|_{L^2(C_u)} \lesssim \delta^{\frac{1}{2}} |u|^{-3-k} (\mathcal{R}_k + C + \delta^{\frac{1}{2}} |u|^{-1} \Delta_k) \cdot \delta^{\frac{1}{2}} |u|.$$

3.2.4. *Estimates for $\nabla^{k+1} \eta$ and $\nabla^{k+1} \underline{\eta}$.* The estimates in this section rely on the bootstrap assumption (3.19) and the estimates established in previous sections. In view of (2.12) and (2.13), we have two Hodge systems:

$$(3.22) \quad \text{div} \eta = -\rho + \frac{1}{2} (\widehat{\chi}, \widehat{\chi}) - \mu, \quad \text{curl} \eta = \sigma - \frac{1}{2} \widehat{\chi} \wedge \widehat{\chi},$$

$$(3.23) \quad \text{div} \underline{\eta} = -\rho + \frac{1}{2} (\widehat{\chi}, \widehat{\chi}) - \underline{\mu}, \quad \text{curl} \underline{\eta} = -\sigma + \frac{1}{2} \widehat{\chi} \wedge \widehat{\chi},$$

where $\mu, \underline{\mu}$ are the mass aspect functions. (See [2] for definitions.) To justify the names, we notice that the integrals of μ or $\underline{\mu}$ on $S_{u,u}$ are the Hawking masses of $S_{u,u}$. In reality, we just take the above equations as definitions of μ and $\underline{\mu}$. They satisfy the following propagation equations:

$$(3.24) \quad \begin{aligned} D\mu &= -\Omega \text{tr } \chi \mu - \frac{1}{2} \Omega \text{tr } \chi \underline{\mu} - \frac{1}{4} \Omega \text{tr } \chi |\widehat{\chi}|^2 \\ &+ \frac{1}{2} \Omega \text{tr } \chi |\underline{\eta}|^2 + \text{div} (2\Omega \widehat{\chi} \cdot \eta - \Omega \text{tr } \chi \underline{\eta}) \end{aligned}$$

$$(3.25) \quad \begin{aligned} \underline{D}\underline{\mu} &= -\Omega \text{tr } \underline{\chi} \underline{\mu} - \frac{1}{2} \Omega \text{tr } \underline{\chi} \mu - \frac{1}{4} \Omega \text{tr } \underline{\chi} |\widehat{\chi}|^2 \\ &+ \frac{1}{2} \Omega \text{tr } \underline{\chi} |\eta|^2 + \text{div} (2\Omega \widehat{\chi} \cdot \underline{\eta} - \Omega \text{tr } \underline{\chi} \eta). \end{aligned}$$

We commute ∇^k with (3.24) to derive

$$\begin{aligned} D\nabla^k \underline{\mu} &= -\Omega \operatorname{tr} \chi \nabla^k \underline{\mu} - \frac{1}{2} \Omega \operatorname{tr} \chi \nabla^k \underline{\mu} + 2\nabla^k \operatorname{div}(\Omega \widehat{\chi}) \cdot \eta + 2\Omega \widehat{\chi} \cdot \nabla^k \operatorname{div} \eta \\ &\quad - \nabla^k \operatorname{div}(\Omega \operatorname{tr} \chi) \cdot \underline{\eta} - \Omega \operatorname{tr} \chi \nabla^k \operatorname{div} \underline{\eta} + \text{l.o.t.} \end{aligned}$$

Since we can control $\nabla^k \operatorname{div}(\Omega \widehat{\chi})$ and $\nabla^k \operatorname{div}(\Omega \operatorname{tr} \chi)$ by (3.12) and (3.13), and we can control $\nabla^k \operatorname{div} \eta$ and $\nabla^k \operatorname{div} \underline{\eta}$ by (3.7) and (3.22), we have

(3.26)

$$\|\nabla^k \underline{\mu}\|_{L^2(S_{\underline{u}, u})} \lesssim |u|^{-1} \int_0^u \|\nabla^k \underline{\mu}\|_{L^2(S_{\underline{u}', u})} d\underline{u}' + |u|^{-3-k} (C + \delta^{\frac{1}{2}} |u|^{-1} \mathcal{R}_k) \cdot |u|.$$

Similarly, we commute ∇^k with (3.25) to derive

$$\begin{aligned} \underline{D}\nabla^k \underline{\mu} &= -\Omega \operatorname{tr} \underline{\chi} \nabla^k \underline{\mu} - \frac{1}{2} \Omega \operatorname{tr} \underline{\chi} \nabla^k \underline{\mu} + 2\nabla^k \operatorname{div}(\Omega \widehat{\underline{\chi}}) \cdot \underline{\eta} + 2\Omega \widehat{\underline{\chi}} \cdot \nabla^k \operatorname{div} \underline{\eta} \\ &\quad - \nabla^k \operatorname{div}(\Omega \operatorname{tr} \underline{\chi}) \cdot \eta - \Omega \operatorname{tr} \underline{\chi} \nabla^k \operatorname{div} \eta + \text{l.o.t.} \end{aligned}$$

In view of (3.7), (3.20), (3.21) and (3.23), we obtain

$$\begin{aligned} (3.27) \quad |u|^{k+1} \|\nabla^k \underline{\mu}\|_{L^2(C_u)} &\lesssim |u_0|^{k+1} \|\nabla^k \underline{\mu}\|_{L^2(C_{u_0})} + \int_{u_0}^u |u'|^k \|\nabla^k \underline{\mu}\|_{L^2(C_{u'})} du' \\ &\quad + |u|^{-2} (C + \mathcal{R}_k + \delta^{\frac{3}{2}} |u|^{-2} \Delta_k) \cdot \delta^{\frac{1}{2}} |u|. \end{aligned}$$

Since $\mu = -2\rho + (\widehat{\chi}, \widehat{\underline{\chi}}) - \underline{\mu}$ on C_{u_0} , (3.26) yields

$$\|\nabla^k \underline{\mu}\|_{L^2(S_{\underline{u}, u_0})} \lesssim |u_0|^{-3-k} (C + \delta^{\frac{1}{2}} |u_0|^{-1} \mathcal{R}_k) \cdot |u_0|.$$

Together with (3.26) and (3.27), we have

$$\begin{aligned} \|\nabla^k \underline{\mu}\|_{L^2(S_{\underline{u}, u})} &\lesssim |u|^{-3-k} (C + \delta^{\frac{1}{2}} |u|^{-1} \mathcal{R}_k + \delta^{\frac{5}{2}} |u|^{-3} \Delta_k) \cdot |u|, \\ \|\nabla^k \underline{\mu}\|_{L^2(C_u)} &\lesssim |u|^{-3-k} (C + \mathcal{R}_k + \delta^{\frac{3}{2}} |u|^{-2} \Delta_k) \cdot \delta^{\frac{1}{2}} |u|. \end{aligned}$$

Once again, we use (3.7), (3.22) and (3.23) to derive the final estimates

$$(3.28) \quad \|\nabla^{k+1} \eta\|_{L^2(C_u)} \lesssim |u|^{-3-k} (C + \mathcal{R}_k + \delta^{\frac{5}{2}} |u|^{-3} \Delta_k) \cdot \delta^{\frac{1}{2}} |u|,$$

$$(3.29) \quad \|\nabla^{k+1} \underline{\eta}\|_{L^2(C_u)} \lesssim |u|^{-3-k} (C + \mathcal{R}_k + \delta^{\frac{3}{2}} |u|^{-2} \Delta_k) \cdot \delta^{\frac{1}{2}} |u|.$$

3.2.5. *Estimates for $\nabla^{k+1} \underline{\omega}$ and $\nabla^{k+1} \omega$.* We now estimate the last two connection coefficients. We will first establish the estimates for $\nabla^{k+1} \underline{\omega}$ and close the auxiliary bootstrap argument based on (3.19) in this section. We first introduce two auxiliary quantities $\underline{\kappa}$ and κ as follows:⁶

$$(3.30) \quad \underline{\Delta} \underline{\omega} = \underline{\kappa} + \operatorname{div}(\Omega \underline{\beta}), \quad \underline{\Delta} \omega = \kappa - \operatorname{div}(\Omega \beta).$$

⁶See Section 6.5 and 6.6 of [2] (but a different notation).

They satisfy the following propagation equations:

$$(3.31) \quad D\underline{\kappa} + \Omega \operatorname{tr} \chi \underline{\kappa} = -2(\Omega \widehat{\chi}, \nabla^2 \underline{\omega}) + \underline{m},$$

$$(3.32) \quad \underline{D}\kappa + \Omega \operatorname{tr} \underline{\chi} \kappa = -2(\Omega \widehat{\chi}, \nabla^2 \omega) + m,$$

where

$$\begin{aligned} \underline{m} &= -2(\operatorname{div}(\Omega \widehat{\chi}), \underline{d}\omega) \\ &\quad + \frac{1}{2} \operatorname{div}(\Omega \operatorname{tr} \chi \cdot \Omega \underline{\beta}) - (\underline{d}(\Omega^2), \underline{d}\rho) - (\underline{d}(\Omega^2), {}^* \underline{d}\sigma) - \rho \underline{\Delta}(\Omega^2) \\ &\quad + \underline{\Delta}(\Omega^2(2(\underline{\eta}, \underline{\eta}) - |\underline{\eta}|^2)) + \operatorname{div}(\Omega^2(\widehat{\chi} \cdot \underline{\beta} - 2\underline{\widehat{\chi}} \cdot \underline{\beta} + 3\underline{\eta}\rho - 3{}^* \underline{\eta}\sigma)) \end{aligned}$$

and

$$\begin{aligned} m &= -2(\operatorname{div}(\Omega \widehat{\chi}), \underline{d}\omega) \\ &\quad - \frac{1}{2} \operatorname{div}(\Omega \operatorname{tr} \underline{\chi} \cdot \Omega \beta) - (\underline{d}(\Omega^2), \underline{d}\rho) + (\underline{d}(\Omega^2), {}^* \underline{d}\sigma) - \rho \underline{\Delta}(\Omega^2) \\ &\quad + \underline{\Delta}(\Omega^2(2(\underline{\eta}, \underline{\eta}) - |\underline{\eta}|^2)) - \operatorname{div}(\Omega^2(\widehat{\chi} \cdot \underline{\beta} - 2\underline{\widehat{\chi}} \cdot \underline{\beta} - 3\underline{\eta}\rho - 3{}^* \underline{\eta}\sigma)). \end{aligned}$$

We commute ∇^{k-1} with (3.31) to derive

$$D\nabla^{k-1} \underline{\kappa} + \Omega \operatorname{tr} \chi \nabla^{k-1} \underline{\kappa} = -2\Omega \widehat{\chi} \cdot \nabla^{k+1} \underline{\omega} + \nabla^{k-1} \underline{m} + \text{l.o.t.}$$

$\nabla^{k+1} \underline{\omega}$ is controlled by the definition of $\underline{\kappa}$ and elliptic estimates (3.7). For $\nabla^{k+1} \underline{m}$, the highest order terms are $\nabla^{k+1} \underline{m}$ are $\nabla^k \beta$, $\nabla^k \underline{\beta}$, $\nabla^k \rho$, $\nabla^k \sigma$ and $\nabla^{k+1} \underline{\eta}$, $\nabla^{k+1} \underline{\eta}$. They can be controlled either by (3.1), (3.28) or (3.29). Therefore

$$\|\nabla^{k-1} \underline{\kappa}\|_{L^2(S_{\underline{u}, u})} \lesssim \delta |u|^{-4-k} (C + \mathcal{R}_k + \delta^2 |u|^{-3} \Delta_k) \cdot |u|.$$

By the definition of $\underline{\kappa}$ and standard elliptic estimates (3.7), we obtain

$$(3.33) \quad \|\nabla^{k+1} \underline{\omega}\|_{L^2(C_u)} \lesssim \delta |u|^{-4-k} (C + \mathcal{R}_k + \delta^2 |u|^{-3} \Delta_k) \cdot \delta^{\frac{1}{2}} |u|.$$

We now close the bootstrap assumption (3.19) using this bound. Recall that $\underline{\omega} = \underline{D} \log \Omega$, and by commuting with ∇^{k+1} , we have

$$\underline{D} \nabla^{k+1} \log \Omega = \nabla^{k+1} \underline{\omega} + \sum_{i=0}^{k-1} \nabla^i \nabla(\Omega \underline{\chi}) \cdot \nabla^{k-1-i} \underline{d} \log \Omega.$$

By integrating, we derive

$$|u|^k \|\nabla^{k+1} \log \Omega\|_{L^2(C_u)} \leq C' \delta |u|^{-3} (C + \mathcal{R}_k + \delta^2 |u|^{-3} \Delta_k) \cdot \delta^{\frac{1}{2}} |u|.$$

We then let $\Delta_k = 4C'(C + \mathcal{R}_k)$ in (3.19) and choose δ sufficiently small such that $4C'\delta^2 \leq 1$, so that $\|\nabla^{k+1} \log \Omega\|_{L^2(C_u)} \leq \frac{1}{2} \delta |u|^{-3-k} \Delta_k \cdot \delta^{\frac{1}{2}} |u|$. Hence, the

bootstrap argument is closed. By the choice of Δ_k , we rewrite (3.20), (3.21), (3.28), (3.29) and (3.33) as

(3.34)

$$\begin{aligned} & \delta^{-1}|u|^3\|\nabla^{k+1}\operatorname{tr}\underline{\chi}'\|_{L^2(C_u)} + \delta^{-\frac{1}{2}}|u|^2\|\nabla^{k+1}\widehat{\underline{\chi}}'\|_{L^2(C_u)} + |u|^2\|\nabla^{k+1}\eta\|_{L^2(C_u)} \\ & + |u|^2\|\nabla^{k+1}\underline{\eta}\|_{L^2(C_u)} + \delta^{-1}|u|^3\|\nabla^{k+1}\underline{\omega}\|_{L^2(C_u)} \lesssim |u|^{-k-1}(C + \mathcal{R}_k) \cdot \delta^{\frac{1}{2}}|u|. \end{aligned}$$

Similarly (we now have (3.34) at our disposal), we also obtain the estimates for $\nabla^{k+1}\omega$:

$$\|\nabla^{k+1}\omega\|_{L^2(C_u)} \lesssim \delta^{-\frac{1}{2}}|u|^{-2-k}(C + \mathcal{R}_k) \cdot \delta^{\frac{1}{2}}|u|.$$

3.3. Estimates for deformation tensors. Following [2], we define the rotation vector fields O_i ($i = 1, 2, 3$) to satisfy $DO_i = 0$ on C_{u_0} and $\underline{D}O_i = 0$ on M , and let $O_i|_{S_{0,u_0}}$ be the standard rotation vector field on the round S^2 . For the sake of simplicity, we will suppress the lower index i in sequel. The deformation tensor is defined by ${}^{(O)}\pi = \mathcal{L}_O g$. We also define its null components ${}^{(O)}\not\pi_{AB} = {}^{(O)}\pi_{AB}$, and $Z^A = \Omega g^{AB} {}^{(O)}\pi_{4B}$. In particular, $Z = 0$ on C_{u_0} by construction. For a tangential tensor field ϕ , $\not\mathcal{L}_O \phi$ is defined as the projection of $\mathcal{L}_O \phi$ to $S_{u,u}$. By definition, on C_{u_0} ,

$$D \operatorname{tr} {}^{(O)}\not\pi = 2O(\Omega \operatorname{tr} \chi), \quad D {}^{(O)}\widehat{\not\pi} - \Omega \operatorname{tr} \chi {}^{(O)}\widehat{\not\pi} = -\Omega \widehat{\chi} \operatorname{tr} {}^{(O)}\not\pi + 2\not\mathcal{L}_O(\Omega \widehat{\chi}).$$

On \mathcal{D} , we also have

$$\underline{D} \operatorname{tr} {}^{(O)}\not\pi = 2O(\Omega \operatorname{tr} \underline{\chi}),$$

$$\underline{D} Z = -4\not\mathcal{L}_O(\Omega^2 \zeta), \quad \underline{D} {}^{(O)}\widehat{\not\pi} - \Omega \operatorname{tr} \underline{\chi} {}^{(O)}\widehat{\not\pi} = -\Omega \widehat{\underline{\chi}} \operatorname{tr} {}^{(O)}\not\pi + 2\not\mathcal{L}_O(\Omega \widehat{\underline{\chi}}).$$

The purpose of the section is to prove the following estimates:

$$(3.35) \quad \|\nabla^{k+1}O\| \lesssim |u|^{-k}(C + \delta^{\frac{1}{2}}|u|^{-2}\mathcal{R}_k) \cdot \delta^{\frac{1}{2}}|u|,$$

$$(3.36) \quad \begin{aligned} & \delta^{-1}|u|^2\|\nabla^k \operatorname{tr} {}^{(O)}\not\pi\| + \delta^{-\frac{1}{2}}|u|\|\nabla^k {}^{(O)}\widehat{\not\pi}\| \\ & + |u|\|\nabla^k Z\| \lesssim |u|^{-k}(C + \mathcal{R}_k) \cdot \delta^{\frac{1}{2}}|u|, \end{aligned}$$

where all the norms are taken with respect to $\|\cdot\|_{L^2(C_u)}$.

We first prove by induction that, for $i \leq k-1$, the following quantities,

$$\begin{aligned} & \|\nabla^{i+1}O\|_{L^2(C_u)}, \quad \delta^{-1}|u|^2\|\nabla^i \operatorname{tr} {}^{(O)}\not\pi\|_{L^2(C_u)}, \\ & \delta^{-\frac{1}{2}}|u|\|\nabla^i {}^{(O)}\widehat{\not\pi}\|_{L^2(C_u)}, \quad |u|\|\nabla^i Z\|_{L^2(C_u)}, \end{aligned}$$

are bounded by $|u|^{-i}C \cdot \delta^{\frac{1}{2}}|u|$ where $C = C(F_{k-1+N})$. The cases for $i = 0, 1, 2$ have been proved in [2]. We assume the case for all $j \leq i-1$ is true. We commute ∇^i with the above equations to derive (the last three equations only hold on C_{u_0})

$$(3.37) \quad \underline{D}\nabla^{i+1}O = \nabla^{i+1}(\Omega \underline{\chi}) \cdot O + \text{l.o.t.},$$

$$(3.38) \quad \underline{D}\nabla^i \operatorname{tr} {}^{(O)}\not\pi = 2\nabla^i \not\mathcal{L}_O(\Omega \operatorname{tr} \underline{\chi}) + \text{l.o.t.},$$

$$(3.39) \quad \underline{D}\nabla^{i(O)}\widehat{\not\chi} - \Omega \operatorname{tr} \underline{\chi} \nabla^{i(O)}\widehat{\not\chi} = -\Omega \widehat{\chi} \otimes \nabla^i \operatorname{tr}^{(O)}\not\chi + 2\nabla^i \not\mathcal{L}_O(\Omega \widehat{\chi}) + \text{l.o.t.},$$

$$(3.40) \quad \underline{D}\nabla^i Z = -4\nabla^i \not\mathcal{L}_O(\Omega^2 \zeta) + \text{l.o.t.}$$

$$(3.41) \quad D\nabla^{i+1}O = \nabla^{i+1}(\Omega \chi) \cdot O + \text{l.o.t.},$$

$$(3.42) \quad D\nabla^i \operatorname{tr}^{(O)}\not\chi = 2\nabla^i \not\mathcal{L}_O(\Omega \operatorname{tr} \chi) + \text{l.o.t.},$$

$$(3.43) \quad D\nabla^{i(O)}\widehat{\not\chi} - \Omega \operatorname{tr} \chi \nabla^{i(O)}\widehat{\not\chi} = -\Omega \widehat{\chi} \otimes \nabla^i \operatorname{tr}^{(O)}\not\chi + 2\nabla^i \not\mathcal{L}_O(\Omega \widehat{\chi}) + \text{l.o.t.}$$

In the current situation, l.o.t. consists of the terms containing at most $(i-1)$ -st order derivatives of deformation tensors that are bounded by induction hypothesis and the i -th order derivatives of connection coefficients that are bounded by (3.1). Hence, by (3.37) and (3.41), we conclude $\|\nabla^{i+1}O\|_{L^2(C_u)} \lesssim |u|^{-i}C \cdot \delta^{\frac{1}{2}}|u|$.

For $\phi \in \{\Omega \operatorname{tr} \underline{\chi}, \Omega \operatorname{tr} \chi, \Omega \widehat{\chi}, \Omega \widehat{\chi}, \Omega^2 \zeta\}$, we have

$$\nabla^i \not\mathcal{L}_O \phi = \not\mathcal{L}_O \nabla^i \phi + \sum_{j=0}^{i-1} \nabla^j \nabla^{(O)}\not\chi \cdot \nabla^{i-1-j} \phi = \not\mathcal{L}_O \nabla^i \phi + \nabla^{i+1}O \cdot \phi + \text{l.o.t.}$$

Since $|\not\mathcal{L}_O \nabla^i \phi| \lesssim |O| |\nabla^{i+1} \phi| + |\nabla O| |\nabla^i \phi|$, combined with (3.38), (3.42), (3.40), the induction hypothesis (3.1) and the estimates for $\nabla^{k+1}O$, we conclude

$$\delta^{-1}|u|^2 \|\nabla^i \operatorname{tr}^{(O)}\not\chi\|_{L^2(C_u)}, |u| \|\nabla^i Z\|_{L^2(C_u)} \lesssim |u|^{-i}C \cdot \delta^{\frac{1}{2}}|u|.$$

Combined with (3.39), (3.43), we also conclude

$$\delta^{-\frac{1}{2}}|u| \|\nabla^{i(O)}\widehat{\not\chi}\|_{L^2(C_u)} \lesssim |u|^{-i}C \cdot \delta^{\frac{1}{2}}|u|.$$

For the case when $i=k$, we can proceed in a same way as to the case when $i \leq k-1$. The only difference is that $\phi \in \{\Omega \operatorname{tr} \underline{\chi}, \Omega \operatorname{tr} \chi, \Omega \widehat{\chi}, \Omega \widehat{\chi}, \Omega^2 \zeta, \nabla^{k+1}\phi\}$ is not bounded by (3.1) but by (3.12), (3.13), (3.34). Thus, we have completed the proof of (3.35) and (3.36).

3.4. Estimates for curvature components. In this section, based on the induction hypothesis (3.1), we derive energy estimates for curvature components. Together with the estimates from previous sections, this will complete the induction argument. Usually, the derivation of energy estimates is based on the Bel-Robinson tensors, see [3] or [2]. We shall proceed directly by integration by parts without Bel-Robinson tensors. This is similar to [12].

3.4.1. Energy inequalities. We commute $\not\mathcal{L}_O^k$ with (2.20)-(2.27) and commute $\widehat{\not\mathcal{L}}_O^k$ with (2.18) and (2.19). From (2.18) and (2.22), we have

$$\widehat{D}\not\mathcal{L}_O^k \alpha - \frac{1}{2}\Omega \operatorname{tr} \underline{\chi} \widehat{\not\mathcal{L}}_O^k \alpha - \Omega \nabla \widehat{\otimes} \not\mathcal{L}_O^k \beta = E_3^k(\alpha),$$

$$D\not\mathcal{L}_O^k \beta - \Omega \widehat{\chi} \cdot \not\mathcal{L}_O^k \beta - \Omega \operatorname{div} \widehat{\not\mathcal{L}}_O^k \alpha = E_4^k(\beta),$$

where $E_3^k(\alpha)$ and $E_4^k(\beta)$ are error terms that will be expressed explicitly later. Therefore,

$$\begin{aligned} & \underline{D}(|\widehat{\mathcal{L}}_O^k \alpha|^2 d\mu_{\mathcal{G}}) + D(2|\mathcal{L}_O^k \beta|^2 d\mu_{\mathcal{G}}) \\ &= (4\text{div}(\Omega \widehat{\mathcal{L}}_O^k \alpha \cdot \mathcal{L}_O^k \beta) - 4\Omega \widehat{\mathcal{L}}_O^k \alpha (\mathcal{d} \log \Omega, \mathcal{L}_O^k \beta)) d\mu_{\mathcal{G}} \\ &+ ((\widehat{\mathcal{L}}_O^k \alpha, 2E_3^k(\alpha)) + 2(\mathcal{L}_O^k \beta, 2E_4^k(\beta))) d\mu_{\mathcal{G}}. \end{aligned}$$

We can integrate this identity on M to derive

$$(3.44) \quad \int_{C_u} |\widehat{\mathcal{L}}_O^k \alpha|^2 + \int_{C_u} 2|\mathcal{L}_O^k \beta|^2 \leq \int_{C_{u_0}} |\widehat{\mathcal{L}}_O^k \alpha|^2 \\ + \iint_M | -4\Omega \widehat{\mathcal{L}}_O^k \alpha (\mathcal{d} \log \Omega, \mathcal{L}_O^k \beta) + (\widehat{\mathcal{L}}_O^k \alpha, 2E_3^k(\alpha)) + 2(\mathcal{L}_O^k \beta, 2E_4^k(\beta)) |.$$

Similarly, from (2.22), (2.24) and (2.26), we have

$$\begin{aligned} \underline{D}\mathcal{L}_O^k \beta - \Omega \widehat{\chi} \cdot \mathcal{L}_O^k \beta - \Omega (\mathcal{d}\mathcal{L}_O^k \rho + * \mathcal{d}\mathcal{L}_O^k \sigma) &= E_3^k(\beta), \\ D\mathcal{L}_O^k \rho + \frac{1}{2} \Omega \text{tr} \chi \mathcal{L}_O^k \rho - \Omega \text{div} \mathcal{L}_O^k \beta &= E_4^k(\rho), \\ D\mathcal{L}_O^k \sigma + \frac{1}{2} \Omega \text{tr} \chi \mathcal{L}_O^k \sigma + \Omega \text{curl} \mathcal{L}_O^k \beta &= E_4^k(\sigma), \end{aligned}$$

which implies

$$(3.45) \quad \int_{C_u} |u|^2 |\mathcal{L}_O^k \beta|^2 + \int_{C_u} |u|^2 (|\mathcal{L}_O^k \rho|^2 + |\mathcal{L}_O^k \sigma|^2) \leq \int_{C_{u_0}} |u_0|^2 |\mathcal{L}_O^k \beta|^2 \\ + \iint_M |u|^2 \left(-\frac{2}{|u|} |\widehat{\mathcal{L}}_O^k \beta|^2 + 2\Omega (\mathcal{d} \log \Omega, \mathcal{L}_O^k \rho \mathcal{L}_O^k \beta + \mathcal{L}_O^k \sigma * \mathcal{L}_O^k \beta) \right. \\ \left. + (\mathcal{L}_O^k \beta, 2E_3^k(\beta)) + (\mathcal{L}_O^k \rho, 2E_4^k(\rho)) + (\mathcal{L}_O^k \sigma, 2E_4^k(\sigma)) \right).$$

We remark that $|u|^2$ appears as a weight. We deal with the remaining equations in a similar way. In fact, we have

$$\begin{aligned} \underline{D}\mathcal{L}_O^k \rho + \frac{1}{2} \Omega \text{tr} \underline{\chi} \mathcal{L}_O^k \rho + \Omega \text{div} \mathcal{L}_O^k \underline{\beta} &= E_3^k(\rho), \\ \underline{D}\mathcal{L}_O^k \sigma + \frac{1}{2} \Omega \text{tr} \underline{\chi} \mathcal{L}_O^k \sigma + \Omega \text{curl} \mathcal{L}_O^k \underline{\beta} &= E_3^k(\sigma), \\ D\mathcal{L}_O^k \underline{\beta} - \Omega \widehat{\chi} \cdot \mathcal{L}_O^k \underline{\beta} + \Omega (\mathcal{d}\mathcal{L}_O^k \rho - * \mathcal{d}\mathcal{L}_O^k \sigma) &= E_4^k(\underline{\beta}) \end{aligned}$$

and

$$\begin{aligned} \underline{D}\mathcal{L}_O^k \underline{\beta} - \Omega \widehat{\chi} \cdot \mathcal{L}_O^k \underline{\beta} + \Omega \text{div} \widehat{\mathcal{L}}_O^k \alpha &= E_3^k(\underline{\beta}), \\ \widehat{D}\widehat{\mathcal{L}}_O^k \alpha - \frac{1}{2} \Omega \text{tr} \chi \widehat{\mathcal{L}}_O^k \alpha + \Omega \nabla \widehat{\otimes} \mathcal{L}_O^k \underline{\beta} &= E_4^k(\alpha). \end{aligned}$$

Therefore, we have

$$(3.46) \quad \int_{C_u} |u|^4 (|\mathcal{L}_O^k \rho|^2 + |\mathcal{L}_O^k \sigma|^2) + \int_{\underline{C}_u} |u|^4 |\mathcal{L}_O^k \underline{\beta}| \leq \int_{C_{u_0}} |u_0|^4 (|\mathcal{L}_O^k \rho|^2 + |\mathcal{L}_O^k \sigma|^2) \\ + \iint_M |u|^4 \left| -\frac{4}{|u|} (|\mathcal{L}_O^k \rho|^2 + |\mathcal{L}_O^k \sigma|^2) + 2\Omega (\not{d} \log \Omega, \mathcal{L}_O^k \rho \mathcal{L}_O^k \underline{\beta} - \mathcal{L}_O^k \sigma^* \mathcal{L}_O^k \underline{\beta}) \right. \\ \left. + (\mathcal{L}_O^k \rho, 2E_3^k(\rho)) + (\mathcal{L}_O^k \sigma, 2E_3^k(\sigma)) + (\mathcal{L}_O^k \underline{\beta}, 2E_4^k(\underline{\beta})) \right|$$

and

$$(3.47) \quad \int_{C_u} 2|u|^6 |\mathcal{L}_O^k \underline{\beta}|^2 + \int_{\underline{C}_u} |u|^6 |\mathcal{L}_O^k \underline{\alpha}|^2 \leq \int_{C_{u_0}} 2|u_0|^6 |\mathcal{L}_O^k \underline{\beta}|^2 \\ + \iint_M |u|^6 \left| -\frac{12}{|u|} |\mathcal{L}_O^k \underline{\beta}|^2 + 4\Omega \widehat{\mathcal{L}}_O^k \underline{\alpha} (\not{d} \log \Omega, \mathcal{L}_O^k \underline{\beta}) \right. \\ \left. + 2(\mathcal{L}_O^k \underline{\beta}, 2E_4^k(\underline{\beta})) + (\widehat{\mathcal{L}}_O^k \underline{\alpha}, 2E_3^k(\underline{\alpha})) \right|.$$

3.4.2. *Estimates for error terms.* We turn to the bound of the error terms that have appeared in the previous sections. We first recall various commutator formulas from [2] that will be used to derive the exact expression of the error terms. For a traceless symmetric two tensor θ , we have

$$[\mathcal{L}_O, \text{div}] \theta_A = \frac{1}{2} {}^{(O)} \widehat{\not{\nabla}}^{BC} \nabla_A \theta_{BC} - \text{div}_B ({}^{(O)} \widehat{\not{\nabla}}^{BC} \theta_{AC}) - \frac{1}{2} \text{tr} {}^{(O)} \not{\nabla} (\text{div} \theta)_A, \\ [\widehat{\mathcal{L}}_O, \widehat{D}] \theta = \widehat{\mathcal{L}}_Z \theta - {}^{(O)} \widehat{\not{\nabla}} (\Omega \widehat{\chi}, \theta) + \Omega \widehat{\chi} ({}^{(O)} \widehat{\not{\nabla}}, \theta), \\ [\widehat{\mathcal{L}}_O, \widehat{\underline{D}}] \theta = -{}^{(O)} \widehat{\not{\nabla}} (\Omega \widehat{\underline{\chi}}, \theta) + \Omega \widehat{\underline{\chi}} ({}^{(O)} \widehat{\not{\nabla}}, \theta).$$

For a one-form ξ , we have

$$[\mathcal{L}_O, \text{div}] \xi = -\text{div} ({}^{(O)} \widehat{\not{\nabla}} \cdot \xi) - \frac{1}{2} \text{tr} {}^{(O)} \not{\nabla} \text{div} \xi, \\ [\mathcal{L}_O, \text{curl}] \xi = -\frac{1}{2} \text{tr} {}^{(O)} \not{\nabla} \text{curl} \xi, \\ [\widehat{\mathcal{L}}_O, \nabla \widehat{\otimes}] \xi = -2 \nabla {}^{(O)} \widehat{\not{\nabla}} \cdot \xi - {}^{(O)} \widehat{\not{\nabla}} \text{div} \xi - \frac{1}{2} ({}^{(O)} \widehat{\not{\nabla}}, \nabla \widehat{\otimes} \xi), \\ [\mathcal{L}_O, D] \xi = \mathcal{L}_Z \xi, \quad [\mathcal{L}_O, \underline{D}] \xi = 0.$$

For a function ϕ , we have

$$[\mathcal{L}_O, \not{d}] \phi = 0, \quad [\mathcal{L}_O, * \not{d}] \phi = {}^{(O)} \widehat{\not{\nabla}} \cdot * \not{d} (O_i \phi), \\ [\mathcal{L}_O, D] \phi = Z \phi, \quad [\mathcal{L}_O, \underline{D}] \phi = 0.$$

We turn to the estimates for $E_3^k(\alpha)$ and $E_4^k(\beta)$. The above formulas allow us to write

$$\begin{aligned} E_3^k(\alpha) &= \frac{1}{2} \sum_{i=0}^{k-1} \mathcal{L}_O^i \mathcal{L}_O(\Omega \operatorname{tr} \underline{\chi}) \widehat{\mathcal{L}}_O^{k-1-i} \alpha \\ &\quad + \widehat{\mathcal{L}}_O^k E_3^0(\alpha) + [\widehat{D}, \widehat{\mathcal{L}}_O^k] \alpha - (\Omega \nabla \widehat{\otimes} \mathcal{L}_O^k \beta - \widehat{\mathcal{L}}_O^k(\Omega \nabla \widehat{\otimes} \beta)), \end{aligned}$$

where $E_3^0(\alpha)$ is from (2.18). For a traceless symmetric two tensor θ , we have

$$\widehat{\mathcal{L}}_O^l \theta = \sum_{i_1 + \dots + i_l + i_{l+1} = l} \nabla^{i_1} O \cdot \nabla^{i_2} O \dots \nabla^{i_l} O \cdot \nabla^{i_{l+1}} \theta.$$

Then first term in $E_3^k(\alpha)$ can be written as

$$\sum_{i_1 + \dots + i_{k+2} = k-1} \nabla^{i_1} O \cdot \nabla^{i_2} O \dots \nabla^{i_k} O \cdot \nabla^{i_{k+1}} \nabla(\Omega \operatorname{tr} \underline{\chi}) \cdot \nabla^{i_{k+2}} \alpha.$$

We bound this term using the induction hypothesis. If $i_{k+1} = k-1$ or $i_{k+2} = k-1$, we bound the corresponding factors in $L^2(C_u)$ and the others in $L^\infty(C_u)$; if $i_j = k-1$ for $j = 1, \dots, k$, $i_{k+1} = k-2$ or $i_{k+2} = k-2$, we bound the corresponding factors in $L^4(S_{u,u})$ (note that when $k = 3$, the case $i_{k+1} = i_{k+2} = k-2$ can happen) and the others in $L^\infty(S_{u,u})$; for the other possible cases, we simply bound all factors in $L^\infty(S_{u,u})$.

For the second term in $E_3^k(\alpha)$, notice that

$$E_3^0(\alpha) = -2\underline{\omega} \alpha - \Omega(- (4\eta + \zeta) \widehat{\otimes} \beta + 3\widehat{\chi} \rho + 3^* \widehat{\chi} \sigma).$$

We use Γ to be an arbitrary connection coefficient and R to be an arbitrary curvature component.⁷ We can compute $\widehat{\mathcal{L}}_O^k E_3^0(\alpha)$ by the commutator formulas. All the terms can be bounded exactly in the same way as for the first term $E_3^k(\alpha)$ except for

$$O^k \cdot (-2\underline{\omega} \nabla^k \alpha + \Omega(4\eta + \zeta) \cdot \nabla^k \beta - 3\Omega \widehat{\chi} \cdot \nabla^k \rho - 3\Omega \widehat{\chi} \cdot \nabla^k \sigma).$$

This term can only be bounded by \mathcal{R}_k . In fact, this collection of terms contains all the terms involving k -th derivatives of the curvature components while the other terms are in the following form,

$$\sum_{i_1 + \dots + i_{k+2} = k, i_{k+2} < k} \nabla^{i_1} O \cdot \nabla^{i_2} O \dots \nabla^{i_k} O \cdot \nabla^{i_{k+1}} \Gamma \cdot \nabla^{i_{k+2}} R,$$

which can be easily estimated by the induction hypothesis.

⁷Throughout the paper, we shall use Γ and R as schematic notation. The expression $\Gamma \cdot R$ denotes a sum of products of this form.

For the third term in $E_3^k(\alpha)$, since $[\widehat{D}, \widehat{\mathcal{L}}_O^k]\alpha = \sum_{i=0}^{k-1} \widehat{\mathcal{L}}_O^{k-1-i}([\widehat{D}, \widehat{\mathcal{L}}_O]\widehat{\mathcal{L}}_O^i\alpha)$, according to the commutator formulas, this term can be rewritten as

$$\sum_{i_1+\dots+i_{k+2}=k-1} \nabla^{i_1} O \cdot \nabla^{i_2} O \dots \nabla^{i_{k-1}} O \cdot \nabla^{i_k(O)} \widehat{\mathcal{L}}_O \cdot \nabla^{i_{k+1}}(\Omega \widehat{\chi}) \cdot \nabla^{i_{k+2}} \alpha.$$

The observation is that none of terms contains k -th derivative of the curvature components. Therefore, $[\widehat{D}, \widehat{\mathcal{L}}_O^k]\alpha$ can also be estimated by induction hypothesis as above.

For the last term in $E_3^k(\alpha)$, we can write it as

$$\Omega(\nabla \widehat{\otimes} \widehat{\mathcal{L}}_O^k \beta - \widehat{\mathcal{L}}_O^k \nabla \widehat{\otimes} \beta) = \Omega \sum_{i=0}^{k-1} \widehat{\mathcal{L}}_O^{k-1-i} (\nabla \widehat{\otimes} \widehat{\mathcal{L}}_O^i \beta - \widehat{\mathcal{L}}_O \nabla \widehat{\otimes} \widehat{\mathcal{L}}_O^i \beta) + \text{l.o.t.},$$

where l.o.t. denotes the terms with at most $(k-1)$ -st derivatives on the curvature components. They can be estimated by induction hypothesis. The main term, i.e., the first term on the right-hand side of the above equation, reads as

$$\sum_{i_1+\dots+i_{k+1}=k, i_k+i_{k+1} \geq 1} \nabla^{i_1} O \cdot \nabla^{i_2} O \dots \nabla^{i_{k-1}} O \cdot \nabla^{i_k(O)} \widehat{\mathcal{L}}_O \cdot \nabla^{i_{k+1}} \beta.$$

Thus, all terms can be estimated by induction hypothesis except for

$$O^{k-1} \cdot (\nabla^{k(O)} \widehat{\mathcal{L}}_O \cdot \beta + {}^{(O)} \widehat{\mathcal{L}}_O \cdot \nabla^k \beta),$$

which should be bound by \mathcal{R}_k according to the definition of \mathcal{R}_k .

Putting the estimates together, we obtain

$$\iint_M |\widehat{\mathcal{L}}_O^k \alpha, 2E_3^k(\alpha)| \lesssim \delta^{-1} \mathcal{R}_k[\alpha](C + \mathcal{R}_k).$$

We move on to $E_4^k(\beta)$, which can be written as

$$\begin{aligned} E_4^k(\beta) &= \sum_{i=0}^{k-1} \widehat{\mathcal{L}}_O^i \widehat{\mathcal{L}}_O(\Omega \widehat{\chi}^i) \cdot \widehat{\mathcal{L}}_O^{k-1-i} \beta \\ &\quad + \widehat{\mathcal{L}}_O^k E_4^0(\beta) + [D, \widehat{\mathcal{L}}_O^k] \beta - (\Omega \text{div} \widehat{\mathcal{L}}_O^k \alpha - \widehat{\mathcal{L}}_O^k(\Omega \text{div} \alpha)). \end{aligned}$$

Since the estimates can be derived almost in the same way as before, we only emphasize the difference. In fact, the third term in $E_4^k(\beta)$ is slightly different from before because $[D, \widehat{\mathcal{L}}_O] \neq 0$. This commutator term contains $\nabla^k Z$, which should be estimated by \mathcal{R}_k . To be more precise, this term can be written as

$$\begin{aligned} [D, \widehat{\mathcal{L}}_O^k] \beta &= \sum_{i=0}^{k-1} \widehat{\mathcal{L}}_O^{k-1-i} ([D, \widehat{\mathcal{L}}_O] \widehat{\mathcal{L}}_O^i \beta) = \sum_{i=0}^{k-1} \widehat{\mathcal{L}}_O^{k-1-i} \widehat{\mathcal{L}}_Z \widehat{\mathcal{L}}_O^i \beta \\ &= \sum_{i_1+\dots+i_{k+1}=k} \nabla^{i_1} O \dots \nabla^{i_{k-1}} O \cdot \nabla^{i_k} Z \cdot \nabla^{i_{k+1}} \beta. \end{aligned}$$

All the terms can be bounded by induction hypothesis expect for $i_k = k$ and $i_{k+1} = k$. Finally, we obtain

$$\iint_M |\mathcal{L}_O^k \beta, 2E_4^k(\beta)| \lesssim \delta^{-\frac{1}{2}} \mathcal{R}_k[\beta](C + \mathcal{R}_k + \mathcal{D}_{k+1}[\text{tr } \underline{\chi}]).$$

We switch to the second group of error terms, $E_3^k(\beta) - \frac{1}{|u|} \mathcal{L}_O^k \beta$ and $E_4^k(\rho, \sigma)$. We have

$$\begin{aligned} E_3^k(\beta) &= \sum_{i=0}^{k-1} \mathcal{L}_O^i \mathcal{L}_O(\Omega \widehat{\chi}^\sharp) \cdot \mathcal{L}_O^{k-1-i} \beta + \mathcal{L}_O^k E_3^0(\beta) \\ &\quad + [D, \mathcal{L}_O^k] \beta - (\Omega(\not{d} \mathcal{L}_O^k \rho + {}^* \not{d} \mathcal{L}_O^k \sigma) - \mathcal{L}_O^k(\Omega(\not{d} \rho + {}^* \not{d} \sigma))). \end{aligned}$$

It is necessary to observe that we need the term $-\frac{1}{|u|} \mathcal{L}_O^k \beta$ to cancel the term $-\frac{1}{2} \text{tr } \underline{\chi} \mathcal{L}_O^k \beta$ in $\mathcal{L}_O^k E_3^k(\beta)$. The derivation of the estimates is almost the same as before, and finally we obtain

$$\iint_M |u|^2 |\mathcal{L}_O^k \beta, 2E_3^k(\beta) - \frac{2}{|u|} \mathcal{L}_O^k \beta| \lesssim \delta^{\frac{1}{2}} \mathcal{R}_k[\beta](C + \mathcal{R}_k).$$

For $E_3^k(\rho, \sigma)$, we need to argue more carefully because it contains some borderline terms (meaning that there is not positive power of δ in the estimates). First of all, we have

$$\begin{aligned} E_4^k(\rho) &= -\frac{1}{2} \sum_{i=0}^{k-1} \mathcal{L}_O^i \mathcal{L}_O(\Omega \text{tr } \chi) \mathcal{L}_O^{k-1-i} \rho \\ &\quad + \mathcal{L}_O^k E_4^0(\rho) + [D, \mathcal{L}_O^k] \rho + (\Omega \text{div } \mathcal{L}_O^k \beta - \mathcal{L}_O^k(\Omega \text{div } \beta)), \\ E_4^k(\sigma) &= -\frac{1}{2} \sum_{i=0}^{k-1} \mathcal{L}_O^i \mathcal{L}_O(\Omega \text{tr } \chi) \mathcal{L}_O^{k-1-i} \sigma \\ &\quad + \mathcal{L}_O^k E_4^0(\sigma) + [D, \mathcal{L}_O^k] \sigma - (\Omega \text{curl } \mathcal{L}_O^k \beta - \mathcal{L}_O^k(\Omega \text{curl } \beta)). \end{aligned}$$

The terms in $[D, \mathcal{L}_O^k] \rho$ and $[D, \mathcal{L}_O^k] \sigma$ do not have the factor $\nabla^k Z$ because ρ and σ are functions. This is slightly different from previous cases, and the estimates are even easier. The key different terms are those $\mathcal{L}_O^k(\widehat{\chi} \cdot \alpha)$ contained in $\mathcal{L}_O^k E_4^0(\rho)$ and $\mathcal{L}_O^k E_4^0(\sigma)$. They can be bounded by $\mathcal{R}_k[\alpha]$ and without any gain in δ . Thus, they contribute to borderline terms. Finally, we obtain

$$\begin{aligned} \iint_M |u|^2 |(\mathcal{L}_O^k \rho, 2E_3^k(\rho)) + (\mathcal{L}_O^k \sigma, 2E_3^k(\sigma))| \\ \lesssim \mathcal{R}_k[\rho, \sigma](C + \mathcal{R}_k[\alpha] + \delta^{\frac{1}{2}}(\mathcal{R}_k + \mathcal{D}_{k+1}[\text{tr } \underline{\chi}])). \end{aligned}$$

For the remaining error terms, namely, $E_3^k(\rho, \sigma)$, $E_4^k(\underline{\beta})$, $E_3^k(\underline{\beta})$ and $E_4^k(\underline{\alpha})$, they can be expressed as

$$\begin{aligned}
E_3^k(\rho) &= -\frac{1}{2} \sum_{i=0}^{k-1} \mathcal{L}_O^i \mathcal{L}_O(\Omega \operatorname{tr} \underline{\chi}) \mathcal{L}_O^{k-1-i} \rho \\
&\quad + \mathcal{L}_O^k E_3^0(\rho) + [\underline{D}, \mathcal{L}_O^k] \rho - (\Omega \operatorname{div} \mathcal{L}_O^k \underline{\beta} - \mathcal{L}_O^k(\Omega \operatorname{div} \underline{\beta})), \\
E_3^k(\sigma) &= -\frac{1}{2} \sum_{i=0}^{k-1} \mathcal{L}_O^i \mathcal{L}_O(\Omega \operatorname{tr} \underline{\chi}) \mathcal{L}_O^{k-1-i} \sigma \\
&\quad + \mathcal{L}_O^k E_3^0(\sigma) + [\underline{D}, \mathcal{L}_O^k] \sigma - (\Omega \operatorname{curl} \mathcal{L}_O^k \underline{\beta} - \mathcal{L}_O^k(\Omega \operatorname{curl} \underline{\beta})), \\
E_4^k(\underline{\beta}) &= \sum_{i=0}^{k-1} \mathcal{L}_O^i \mathcal{L}_O(\Omega \widehat{\chi}^\sharp) \cdot \mathcal{L}_O^{k-1-i} \underline{\beta} + \mathcal{L}_O^k E_4^0(\underline{\beta}) \\
&\quad + [D, \mathcal{L}_O^k] \underline{\beta} + (\Omega(\not\partial \mathcal{L}_O^k \rho - {}^* \not\partial \mathcal{L}_O^k \sigma) - \mathcal{L}_O^k(\Omega(\not\partial \rho - {}^* \not\partial \sigma))), \\
E_3^k(\underline{\beta}) &= \sum_{i=0}^{k-1} \mathcal{L}_O^i \mathcal{L}_O(\Omega \widehat{\chi}^\sharp) \cdot \mathcal{L}_O^{k-1-i} \underline{\beta} \\
&\quad + \widehat{\mathcal{L}}_O^k E_3^0(\underline{\beta}) + [\underline{D}, \mathcal{L}_O^k] \underline{\beta} + (\Omega \operatorname{div} \widehat{\mathcal{L}}_O^k \underline{\alpha} - \widehat{\mathcal{L}}_O^k(\Omega \operatorname{div} \underline{\alpha})), \\
E_4^k(\underline{\alpha}) &= \frac{1}{2} \sum_{i=0}^{k-1} \mathcal{L}_O^i \mathcal{L}_O(\Omega \operatorname{tr} \chi) \widehat{\mathcal{L}}_O^{k-1-i} \underline{\alpha} \\
&\quad + \widehat{\mathcal{L}}_O^k E_4^0(\underline{\alpha}) + [\widehat{D}, \widehat{\mathcal{L}}_O^k] \underline{\alpha} + (\Omega \nabla \widehat{\otimes} \mathcal{L}_O^k \underline{\beta} - \widehat{\mathcal{L}}_O^k(\Omega \nabla \widehat{\otimes} \underline{\beta})).
\end{aligned}$$

We can proceed exactly as before to derive

$$\begin{aligned}
&\iint_M |u|^4 \left| (\mathcal{L}_O^k \rho, 2E_3^k(\rho) - \frac{4}{|u|} \mathcal{L}_O^k \rho) + \left(\mathcal{L}_O^k \sigma, 2E_3^k(\sigma) - \frac{4}{|u|} \mathcal{L}_O^k \sigma \right) \right| \\
&\quad \lesssim \delta^2 \mathcal{R}_k[\rho, \sigma](C + \mathcal{R}_k + \underline{\mathcal{R}}_k[\underline{\alpha}]), \\
&\iint_M |u|^4 |(\mathcal{L}_O^k \underline{\beta}, 2E_4^k(\underline{\beta}))| \lesssim \delta^2 \mathcal{R}_k[\underline{\beta}](C + \mathcal{R}_k), \\
&\iint_M |u|^6 \left| \left(\mathcal{L}_O^k \underline{\beta}, 2E_3^k(\underline{\beta}) - \frac{6}{|u|} \mathcal{L}_O^k \underline{\beta} \right) \right| \lesssim \delta^{\frac{7}{2}} \mathcal{R}_k[\underline{\beta}](C + \mathcal{R}_k + \underline{\mathcal{R}}_k[\underline{\alpha}]), \\
&\iint_M |u|^6 |(\mathcal{L}_O^k \underline{\alpha}, 2E_4^k(\underline{\alpha}))| \lesssim \delta^3 \underline{\mathcal{R}}_k[\underline{\alpha}](C + \mathcal{R}_k[\rho, \sigma] + \delta^{\frac{1}{2}}(\mathcal{R}_k + \underline{\mathcal{R}}_k[\underline{\alpha}])).
\end{aligned}$$

3.5. Completion of the induction argument. We first close the induction argument on the L^2 level. In fact, based on the estimates derived in the previous sections, we show that if δ is sufficiently small depending on the C^{k+N} bounds of the seed data ψ_0 , then

$$\mathcal{R}_k, \underline{\mathcal{R}}_k[\alpha], \mathcal{O}_{k+1} \leq F_{k+N},$$

where F_{k+N} is a constant depending only on the C^{k+N} bounds of the seed data.

Recall that for a one-form ξ , $\mathcal{L}_O^k \xi - \nabla_O^k \xi = \sum_{i_{k+1} < k} \nabla^{i_1} O \dots \nabla^{i_k} O \cdot \nabla^{i_{k+1}} \xi$, so that

$$\int_{C_u} |u|^{2k} |\nabla^k \xi|^2 \lesssim \int_{C_u} |\mathcal{L}_O^i \xi|^2 + \int_{C_u} \sum_{i=0}^{k-1} |u|^{2i} |\nabla^i \xi|^2.$$

Similar inequalities hold for traceless symmetric two tensors and functions.

We suppose that $\mathcal{R}_k, \underline{\mathcal{R}}_k[\alpha] \leq G$ for some large constant G to be fixed later. In view of the above estimates, we multiply δ^2 to (3.44), δ^0 to (3.45), δ^{-1} to (3.46) and δ^{-3} to (3.47), then substitute the above estimates into (3.44)-(3.47) and choose $\delta > 0$ sufficiently small depending on G and F_{k-1+N} (3.1), so that we obtain the following inequalities:

$$\begin{aligned} \mathcal{R}_k[\alpha]^2 &\leq \mathcal{R}_k[\alpha](u_0)^2 + 1, \\ \mathcal{R}_k[\beta]^2 &\leq \mathcal{R}_k[\beta](u_0)^2 + C\mathcal{R}_k[\rho, \sigma](1 + \mathcal{R}_k[\alpha]) + 1, \\ \mathcal{R}_k[\rho, \sigma]^2 &\leq \mathcal{R}_k[\rho, \sigma](u_0)^2 + 1, \\ \mathcal{R}_k[\beta]^2 + \underline{\mathcal{R}}_k[\underline{\alpha}]^2 &\leq \mathcal{R}_k[\beta](u_0)^2 + C\underline{\mathcal{R}}_k[\underline{\alpha}](1 + \mathcal{R}_k[\rho, \sigma]) + 1. \end{aligned}$$

We substitute the first and the third to the second to derive

$$\mathcal{R}_k[\beta]^2 \leq \mathcal{R}_k[\beta](u_0)^2 + C\sqrt{\mathcal{R}_k[\rho, \sigma](u_0)^2 + 1}(1 + \sqrt{\mathcal{R}_k[\alpha](u_0)^2 + 1}) + 1.$$

We then substitute the third to the fourth to derive

$$\underline{\mathcal{R}}_k[\underline{\alpha}] \leq C(1 + \sqrt{\mathcal{R}_k[\rho, \sigma](u_0)^2 + 1}) + \sqrt{\mathcal{R}_k[\beta](u_0)^2 + 1}.$$

Thus, we can choose G depending on $\mathcal{R}_k(u_0)$ and hence on C^{k+N} bounds of the seed data, such that $\mathcal{R}_k, \underline{\mathcal{R}}_k[\alpha] \leq G/2$. By continuity argument and the estimates for connection coefficients in previous sections, we have obtained the estimates for $\mathcal{R}_k, \underline{\mathcal{R}}_k[\alpha], \mathcal{O}_{k+1}$.

We turn to the following estimates:

$$\mathcal{R}_{\leq k-1}^4, \mathcal{R}_{\leq k-2}^\infty, \mathcal{O}_{\leq k}^4, \mathcal{O}_{\leq k-1}^\infty \leq C(F_{k+N}).$$

We first commute ∇^{k-1} with (2.20), (2.24), (2.26) and (2.23) to derive

$$\begin{aligned} \delta^{-\frac{1}{2}} |u|^{-1} \left(\|\delta^{\frac{1}{2}} |u|^2 (\delta D)(|u| \nabla)^{k-1} \beta\|_{L^2(C_u)} + \||u|^3 (\delta D)(|u| \nabla)^{k-1} (\rho, \sigma)\|_{L^2(C_u)} \right. \\ \left. + \|\delta^{-1} |u|^4 (\delta D)(|u| \nabla)^{k-1} \underline{\beta}\|_{L^2(C_u)} \right) \leq C(F_{k+N}). \end{aligned}$$

We then apply $\widehat{\mathcal{L}}_O^{k-1}\widehat{D}$ to (2.18), $\mathcal{L}_O^{k-1}D$ to (2.22), $\widehat{\mathcal{L}}_O^{k-1}\widehat{D}$ to (2.19) and $\mathcal{L}_O^{k-1}\underline{D}$ to (2.23) to derive

$$\begin{aligned}\widehat{D}\widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha - \frac{1}{2}\Omega \operatorname{tr} \chi \widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha - \Omega \nabla \widehat{\otimes} \mathcal{L}_O^{k-1}D\beta &= E_3^{(k-1)4}(\alpha), \\ D\mathcal{L}_O^{k-1}D\beta - \Omega \widehat{\chi} \cdot \mathcal{L}_O^{k-1}D\beta - \Omega \operatorname{div} \widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha &= E_4^{(k-1)4}(\beta), \\ \underline{D}\mathcal{L}_O^{k-1}\underline{D}\beta - \Omega \widehat{\chi} \cdot \mathcal{L}_O^{k-1}\underline{D}\beta + \Omega \operatorname{div} \widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha &= E_3^{(k-1)3}(\beta), \\ \widehat{D}\widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha - \frac{1}{2}\Omega \operatorname{tr} \chi \widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha + \Omega \nabla \widehat{\otimes} \mathcal{L}_O^{k-1}\underline{D}\beta &= E_4^{(k-1)3}(\alpha).\end{aligned}$$

Therefore, we have

$$(3.48) \quad \int_{C_u} |\widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha|^2 \leq \int_{C_{u_0}} |\widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha|^2 + \iint_M \left| -4\Omega \widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha (\not{d} \log \Omega, \mathcal{L}_O^{k-1}D\beta) \right. \\ \left. + (\widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha, 2E_3^{(k-1)4}(\alpha)) + 2(\mathcal{L}_O^{k-1}D\beta, 2E_4^{(k-1)4}(\beta)) \right|$$

and

$$(3.49) \quad \int_{\underline{C}_u} |u|^8 |\mathcal{L}_O^{k-1}\underline{D}\alpha|^2 \leq \int_{C_{u_0}} 2|u_0|^8 |\mathcal{L}_O^{k-1}\underline{D}\beta|^2 \\ + \iint_M |u|^8 \left| -\frac{16}{|u|} |\mathcal{L}_O^{k-1}\underline{D}\beta|^2 + 4\Omega \widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha (\not{d} \log \Omega, \mathcal{L}_O^{k-1}\underline{D}\beta) \right. \\ \left. + 2(\mathcal{L}_O^{k-1}\underline{D}\beta, 2E_4^{(k-1)3}(\beta)) + (\widehat{\mathcal{L}}_O^{k-1}\widehat{D}\alpha, 2E_3^{(k-1)3}(\alpha)) \right|.$$

This leads to

$$\delta^{-\frac{1}{2}}|u| \|\delta^{\frac{3}{2}}|u|(\delta D)(|u|\nabla^{k-1})\alpha\|_{L^2(C_u)} \\ + \||u|^{-\frac{3}{2}}\delta^{-\frac{3}{2}}|u|^{\frac{9}{2}}(|u|\underline{D})(|u|\nabla)^{k-1}\underline{\alpha}\|_{L^2(\underline{C}_u)} \leq C(F_{k+N}).$$

The Sobolev inequalities (3.3), (3.4) and (3.5) yield $\mathcal{R}_k^4, \mathcal{R}_{k-1}^\infty \leq C(F_{k+N})$.

Similarly, we commute ∇^k with (2.1), (2.2), (2.5), (2.7), (2.14) and (2.15) to derive

$$\begin{aligned}\delta^{-\frac{1}{2}}|u|^{-1} \left(\|\delta^{\frac{1}{2}}|u|(\delta D)(|u|\nabla)^k \widehat{\chi}\|_{L^2(C_u)} + \||u|^2(\delta D)(|u|\nabla)^k \operatorname{tr} \chi\|_{L^2(C_u)} \right. \\ \left. + \|\delta^{-\frac{1}{2}}|u|^2(\delta D)(|u|\nabla)^k \widehat{\chi}\|_{L^2(C_u)} + \|\delta^{-1}|u|^3(\delta D)(|u|\nabla)^k \operatorname{tr} \chi\|_{L^2(C_u)} \right. \\ \left. + \||u|^2(\delta D)(|u|\nabla)^k \eta\|_{L^2(C_u)} + \|\delta^{-1}|u|^3(\delta D)(|u|\nabla)^k \underline{\omega}\|_{L^2(C_u)} \right) \\ \leq C(F_{k+N}).\end{aligned}$$

According to the relation $\eta + \underline{\eta} = 2\not{d} \log \Omega$, we can also derive

$$\delta^{-\frac{1}{2}}|u|^{-1} \||u|^2(\delta D)(|u|\nabla)^k \eta\|_{L^2(C_u)} \leq C(F_{k+N}).$$

Now (3.3) and (3.4) yield $\mathcal{O}_k^4, \mathcal{O}_k^\infty \leq C(F_{k+N})$ except for the component ω .

In order to bound ω , we commute ∇^k with (2.8). We also take ∇^k to both sides of the equation and we hope to bound the $L^2(\underline{C}_u)$ norm for $\underline{D}\nabla^k\omega$. The commutator $[\underline{D}, \nabla^k]\omega$ is easy to control thanks to the induction hypothesis (3.1). Because we do have an $L^4(S_{\underline{u},u})$ bound for $\nabla^k\eta$ and $\nabla^k\underline{\eta}$ at the moment, then we can bound $\| |u|^{-\frac{3}{2}}|u|^2(|u|\nabla)^k((\eta, \underline{\eta}) - |\eta|^2) \|_{L^2(\underline{C}_u)} \leq C(F_{k+N})$. In view of (3.45), we have $\| |u|^{-\frac{3}{2}}|u|^{\frac{5}{2}}(|u|\nabla)^k\rho \| \leq C(F_{k+N})$ and we can deduce that

$$\| |u|^{-\frac{3}{2}}|u|(|u|\underline{D})(|u|\nabla)^k\omega \| \leq C(F_{k+N}).$$

Finally, (3.3) and (3.5) yield $\mathcal{O}_k^4[\omega], \mathcal{O}_{k-1}^\infty[\omega] \leq C(F_{k+N})$.

We have completed the proof of Theorem 1. □

4. Construction of the transition region

We have already mentioned that, in addition to Christodoulou’s short pulse ansatz described, we need to impose one more condition on the initial data defined on $C_{u_0}^{[0,\delta]}$, that is, (2.30) or (2.31). This condition is the key to the construction in the sequel. We start from condition (2.30), which is imposed on the seed data, that is,

$$(4.1) \quad \int_0^1 \left| \frac{\partial\psi_0}{\partial s}(s, \theta) \right|^2 ds = 16m_0.$$

We show that this condition is equivalent to the more physical one (2.31) up to an error of size δ to some positive power. More precisely, we have

LEMMA 1. *Assume (2.30) holds. Then on the sphere $S_{0,\delta}$, for all k , we have*

$$|u_0|^k \left| \nabla^k \left(\int_0^\delta |u_0|^2 |\widehat{\chi}(\underline{u}, u_0)|^2 d\underline{u} - 4m_0 \right) \right| \lesssim_k \delta^{\frac{1}{2}} |u_0|^{-1}.$$

The notation $A \lesssim_k B$ means that $A \leq C_k B$, where C_k is a constant depending on the C^{k+N} norms of the seed data for sufficiently large N .

Proof. In stereographic coordinates, we have

$$\begin{aligned} 4 \int_0^\delta |\widehat{\chi}(\underline{u}, u_0)|^2 d\underline{u} &= \int_0^\delta \widehat{g}^{AC} \widehat{g}^{BD} D\widehat{g}_{AB} D\widehat{g}_{CD} d\underline{u} \\ &= \int_0^\delta m^{AC} m^{BD} \partial_{\underline{u}} m_{AB} \partial_{\underline{u}} m_{CD} d\underline{u}. \end{aligned}$$

In view of the short pulse ansatz (2.28), we have $|\partial_\theta^k(m^{AB} - \delta^{AB})| \lesssim_k \delta^{\frac{1}{2}} |u_0|^{-1}$. Now we can make use of the definition $m = \exp \psi$ to derive

$$\left| \partial_\theta^k \left(4 \int_0^\delta |u_0|^2 |\widehat{\chi}(\underline{u}, u_0)|^2 d\underline{u} - \int_0^\delta |u_0|^2 |\partial_{\underline{u}}\psi(\underline{u})|^2 d\underline{u} \right) \right| \lesssim_k \delta^{\frac{1}{2}} |u_0|^{-1}.$$

In view of (2.28) and (2.30), we have $\int_0^\delta |u_0|^2 |\partial_{\underline{u}} \psi|^2 d\underline{u} = \int_0^1 |\partial_s \psi_0|^2 ds = 16m_0$, and therefore

$$|u_0|^k \left| \nabla^k \left(\int_0^\delta |u_0|^2 |\widehat{\chi}(\underline{u}, u_0)|^2 d\underline{u} - 4m_0 \right) \right| \lesssim_k \delta^{\frac{1}{2}} |u_0|^{-1}. \quad \square$$

Thanks to the lemma, we shall not differentiate conditions (2.30) and (2.31). We turn to the geometric consequences of these conditions.

4.1. *Geometry on S_{δ, u_0} .* The purpose of this section is to prove the following lemma which, roughly speaking, says the geometry of the two sphere S_{δ, u_0} is close to the geometry of a given two sphere in the Schwarzschild space-time with mass m_0 .

LEMMA 2. *On the sphere S_{δ, u_0} , we have $\alpha \equiv 0$ and for all k ,*

$$|u_0|^k \left(\left| \nabla^k \left(\text{tr } \chi - \left(\frac{2}{|u_0|} - \frac{4m_0}{|u_0|^2} \right) \right) \right|, |\nabla^k \beta|, |u_0| \left| \nabla^k \left(\rho + \frac{2m_0}{|u_0|^3} \right) \right|, |u_0| |\nabla^k \sigma| \right) \lesssim_k \delta^{\frac{1}{2}} |u_0|^{-2}.$$

Remark 6. The connection coefficients and the curvature components appearing in the lemma have improvement on “smallness,” which means that they are controlled by $\delta^{\frac{1}{2}}$ instead of δ to some nonpositive power in [2]. This improvement comes from condition (2.31). The other connection coefficients and the curvature components, which are not mentioned in the lemma, have already been bounded by δ to some positive power in [2]. This suggests that the geometry of the sphere S_{δ, u_0} is close to the geometry of a sphere in Schwarzschild spacetime.

Proof. First of all, that $\alpha = 0$ on S_{δ, u_0} follows immediately from (2.1) and the fact that $\widehat{\chi}$ has compact support in $C_{u_0}^{(0, \delta)}$.

For $\text{tr } \chi$, on C_{u_0} , we can write (2.2) as

$$D \text{tr } \chi = -\frac{1}{2} (\text{tr } \chi)^2 - |\widehat{\chi}|^2$$

since this is where $\Omega = 1$. We integrate this equation along $[0, \delta]$ to derive

$$\text{tr } \chi(\delta, u_0) - \frac{2}{|u_0|} = -\frac{1}{2} \int_0^\delta (\text{tr } \chi(\underline{u}, u_0))^2 d\underline{u} - \int_0^\delta |\widehat{\chi}(\underline{u}, u_0)|^2 d\underline{u}.$$

Since $|u_0|^k |\nabla^k (\text{tr } \chi - \frac{2}{|u_0|})| \lesssim_k |u_0|^{-2}$ (see Chapter 2 of [2]), taking ∇^k on the above equation yields the desired estimates for $\text{tr } \chi$ on the sphere S_{δ, u_0} .

For β , on C_{u_0} , we can write (2.10) as

$$d \widehat{\chi} = \frac{1}{2} d \text{tr } \chi - \widehat{\chi} \cdot \eta + \frac{1}{2} \text{tr } \chi \eta - \beta.$$

Since $|u_0|^k |\nabla^k \eta| \lesssim_k \delta^{\frac{1}{2}} |u_0|^{-2}$ (see Chapter 2 of [2]), $\widehat{\chi} = 0$ on S_{δ, u_0} , combined with the estimates on $\text{tr } \chi$ just derived, we can apply ∇^k on the above equation to derived the desired estimates for β on S_{δ, u_0} .

For σ , on C_{u_0} , we recall (2.12):

$$\text{curl } \eta = \sigma - \frac{1}{2} \widehat{\chi} \wedge \widehat{\chi}.$$

Since $|u_0|^k |\nabla^k \eta| \lesssim_k \delta^{\frac{1}{2}} |u_0|^{-2}$ and $\widehat{\chi} = 0$ on S_{δ, u_0} , we obtain the desired estimates for σ on S_{δ, u_0} by simply applying ∇^k to the above equation.

Finally, we consider ρ . On C_{u_0} , the Gauss equation (2.9) can be written as

$$K = -\frac{1}{4} \text{tr } \chi \text{tr } \underline{\chi} + \frac{1}{2} (\widehat{\chi}, \widehat{\chi}) - \rho.$$

We rewrite this equation in the following renormalized form:

$$\left(K - \frac{1}{|u_0|^2} \right) + \frac{1}{4} \text{tr } \chi \left(\text{tr } \underline{\chi} + \frac{2}{|u_0|} \right) = - \left(\rho - \frac{1}{2|u_0|} \left(\text{tr } \chi - \frac{2}{|u_0|} \right) \right).$$

Since $|u_0|^k |\nabla^k (K - \frac{1}{|u_0|^2})| \lesssim_k \delta^{\frac{1}{2}} |u_0|^{-3}$ and $|u_0|^k |\nabla^k (\text{tr } \underline{\chi} + \frac{2}{|u_0|})| \lesssim_k \delta |u_0|^{-2}$ on C_{u_0} (see again Chapter 2 of [2]), combined with the estimates for $\text{tr } \chi$ derived above, we obtain the desired estimates for ρ by applying ∇^k to the above equation. This completes the proof of the lemma. \square

4.2. Geometry on \underline{C}_δ . We shall study the geometry of \underline{C}_δ by integrating the equations (2.16), (2.17), (2.25), (2.27), (2.22), (2.18), (2.8) along \underline{C}_δ . These equations can be viewed as ordinary differential equations for the connection coefficients and curvature components that are not bounded by δ to some positive power in [2]. We will see that they are actually bounded by δ to some positive power under the additional condition (2.31). We will regain the “smallness” from Theorem 1 and improved smallness from Lemma 2. More precisely, we prove the following lemma:

LEMMA 3. *Suppose that $\delta > 0$ has been chosen such that Theorem 1 is true for some k . Then on the incoming cone \underline{C}_δ , we have the following estimates:*

$$\begin{aligned} \|(|u|\nabla)^{k-3}(\Omega\widehat{\chi})\|_{L^\infty} &\lesssim \delta^{\frac{1}{2}}|u|^{-1} \quad \text{for } i \leq k-3, \\ |u|^{-\frac{1}{2}}\|(|u|\nabla)^{k-2}(\Omega\widehat{\chi})\|_{L^4(S_{\delta,u})} + |u|^{-1}\|(|u|\nabla)^{k-1}(\Omega\widehat{\chi})\|_{L^2(S_{\delta,u})} &\lesssim \delta^{\frac{1}{2}}|u|^{-2}; \\ \left\| (|u|\nabla)^i \left(\rho + \frac{2m_0}{|u|^3}, \sigma \right) \right\|_{L^\infty} &\lesssim \delta^{\frac{1}{2}}|u|^{-3} \quad \text{for } i \leq k-3, \\ |u|^{-\frac{1}{2}}\|(|u|\nabla)^{k-2}(\rho, \sigma)\|_{L^4(S_{\delta,u})} + |u|^{-1}\|(|u|\nabla)^{k-1}(\rho, \sigma)\|_{L^2(S_{\delta,u})} &\lesssim \delta^{\frac{1}{2}}|u|^{-3}; \\ \|(|u|\nabla)^i \beta\|_{L^\infty} &\lesssim \delta^{\frac{1}{2}}|u|^{-2} \quad \text{for } i \leq k-4, \\ |u|^{-\frac{1}{2}}\|(|u|\nabla)^{k-3}\beta\|_{L^4(S_{\delta,u})} + |u|^{-1}\|(|u|\nabla)^{k-2}\beta\|_{L^2(S_{\delta,u})} &\lesssim \delta^{\frac{1}{2}}|u|^{-2}; \end{aligned}$$

$$\begin{aligned} \|(|u|\nabla)^i \alpha\|_{L^\infty} &\lesssim \delta^{\frac{1}{2}} |u|^{-2} \quad \text{for } i \leq k - 5, \\ |u|^{-\frac{1}{2}} \|(|u|\nabla)^{k-4} \alpha\|_{L^4(S_{\delta,u})} + |u|^{-1} \|(|u|\nabla)^{k-3} \alpha\|_{L^2(S_{\delta,u})} &\lesssim \delta^{\frac{1}{2}} |u|^{-2}; \\ \left\| (|u|\nabla)^i \left(\text{tr } \chi - \left(\frac{2}{|u|} - \frac{4m_0}{|u|^2} \right), \omega - \frac{m_0}{|u|^2} + \frac{m_0}{|u_0|^2} \right) \right\|_{L^\infty} &\lesssim \delta^{\frac{1}{2}} |u|^{-3} \quad \text{for } i \leq k - 3, \\ |u|^{-\frac{1}{2}} \|(|u|\nabla)^{k-4} (\text{tr } \chi, \omega)\|_{L^4(S_{\delta,u})} & \\ + |u|^{-1} \|(|u|\nabla)^{k-3} (\text{tr } \chi, \omega)\|_{L^2(S_{\delta,u})} &\lesssim \delta^{\frac{1}{2}} |u|^{-3}. \end{aligned}$$

Remark 7. The smallness gained in the lemma comes from Lemma 2 and therefore from condition (2.31). Heuristically, the worst control for the geometric quantities on \underline{C}_u is propagated from the initial surface C_{u_0} . This is a very important technical feature of the work of Christodoulou [2]. Therefore, condition (2.31) allows one to improve the geometry on S_{δ,u_0} and hence the geometry on \underline{C}_δ .

Proof. We will derive the estimates for $\widehat{\chi}$, ρ , σ , β , α , ω and $\text{tr } \chi$ one-by-one. The proof for each component is by induction on the number of derivatives. The parameter $\delta > 0$ is chosen to be sufficiently small so that Theorem 1 holds for $k \gg 5$.

For $\widehat{\chi}$, we rewrite (2.16) as

$$(4.2) \quad \underline{D}(\Omega\widehat{\chi}) = \frac{1}{2} \Omega \text{tr } \underline{\chi}(\Omega\widehat{\chi}) + 2(\Omega\widehat{\chi}, \Omega\widehat{\chi})\not{g} + \Omega^2 \left(\nabla \widehat{\otimes} \eta + \eta \widehat{\otimes} \eta - \frac{1}{2} \text{tr } \chi \widehat{\chi} \right),$$

where the second term is the trace of $\underline{D}(\Omega\widehat{\chi})$.

Remark 8. In Christodoulou’s work, one only obtains bounds for the second derivatives of the curvature components that imply the L^∞ estimates $|\nabla \widehat{\otimes} \eta| \lesssim |u|^{-3}$. If we have estimates for the third derivatives of the curvature components, as stated in Theorem 1, we have $|\nabla \widehat{\otimes} \eta| \lesssim \delta^{\frac{1}{2}} |u|^{-3}$. Therefore, we can bound $\Omega^2(\nabla \widehat{\otimes} \eta + \eta \widehat{\otimes} \eta - \frac{1}{2} \text{tr } \chi \widehat{\chi})$ in L^∞ by δ to some positive power. So the right-hand side of the above equation can be treated as a small term. This illustrates the idea of the proof and the importance of higher order energy estimates.

We go back to the proof and integrate $\underline{D}(\Omega\widehat{\chi})$ along \underline{C}_δ . The first two terms on the right-hand side can be absorbed by Gronwall’s inequality. In view of the fact that $\widehat{\chi}(\delta, u_0) = 0$, according to Theorem 1, we have

$$|u| |\Omega\widehat{\chi}(\delta, u)| \lesssim \int_{u_0}^u |u'| \left| \Omega^2 \left(\nabla \widehat{\otimes} \eta + \eta \widehat{\otimes} \eta - \frac{1}{2} \text{tr } \chi \widehat{\chi} \right) \right| du' \lesssim \delta^{\frac{1}{2}} |u|^{-1}.$$

We then use an induction argument to obtain control for higher order derivatives. Suppose that for all $j < i \leq k - 3$, $|u|^j \|\nabla^j(\Omega\widehat{\chi})\|_{L^\infty(\delta,u)} \lesssim \delta^{\frac{1}{2}} |u|^{-2}$; then

we apply ∇^i to (4.2) to derive

$$\begin{aligned} \underline{D}\nabla^i(\Omega\widehat{\chi}) &= \frac{1}{2}\Omega \operatorname{tr} \underline{\chi} \nabla^i(\Omega\widehat{\chi}) + 2(\Omega\widehat{\chi}, \nabla^i(\Omega\widehat{\chi})) \\ &\quad + \nabla^i \left(\Omega^2(\nabla \widehat{\otimes} \eta + \eta \widehat{\otimes} \eta - \frac{1}{2} \operatorname{tr} \chi \widehat{\chi}) \right) + \text{l.o.t.} \end{aligned}$$

Therefore, the third term on the right-hand side is bounded by

$$|u|^i \left| \nabla^i \left(\Omega^2 \left(\nabla \widehat{\otimes} \eta + \eta \widehat{\otimes} \eta - \frac{1}{2} \operatorname{tr} \chi \widehat{\chi} \right) \right) \right| \lesssim \delta^{\frac{1}{2}} |u|^{-2}$$

thanks to Theorem 1. The l.o.t. is of the form

$$\sum_{j=0}^{i-1} \nabla^j \nabla (\Omega \operatorname{tr} \underline{\chi}) \cdot \nabla^{i-j-1}(\Omega\widehat{\chi}) + \sum_{j=0}^{i-1} \nabla^j \nabla (\Omega\widehat{\chi}) \cdot \nabla^{i-j-1}(\Omega\widehat{\chi}).$$

We can bound $\nabla^j \nabla (\Omega \operatorname{tr} \underline{\chi})$ and $\nabla^j \nabla (\Omega\widehat{\chi})$ by Theorem 1, and we can bound $\nabla^{i-j-1}(\Omega\widehat{\chi})$ by the induction hypothesis. By integrating the above equation and using the Gronwall's inequality to absorb the first two terms, we obtain that

$$(4.3) \quad \|(|u|\nabla)^i(\Omega\widehat{\chi})\|_{L^\infty} \lesssim \delta^{\frac{1}{2}} |u|^{-1}$$

for $i \leq k - 3$.

For $i = k - 2$, we only expect that the $L^4(S_{\delta,u})$ norm on $\widehat{\chi}$ is controlled by δ to some positive power, and for $i = k - 1$, only the $L^2(S_{\delta,u})$ norm is small. This is because we only have information on $\|(|u|\nabla)^{k-1}\eta\|_{L^4(S_{\delta,u})}$ and $\|(|u|\nabla)^k\eta\|_{L^2(S_{\delta,u})}$ from Theorem 1, but no information for higher derivatives. By a similar argument and integrating along \underline{C}_δ , we obtain

$$(4.4) \quad |u|^{-\frac{1}{2}} \|(|u|\nabla)^{k-2}(\Omega\widehat{\chi})\|_{L^4(S_{\delta,u})} + |u|^{-1} \|(|u|\nabla)^{k-1}(\Omega\widehat{\chi})\|_{L^2(S_{\delta,u})} \lesssim \delta^{\frac{1}{2}} |u|^{-2}.$$

It is easy to see that (4.3) and (4.4) also hold if we replace $\Omega\widehat{\chi}$ by $\widehat{\chi}$.

We perform a similar argument for the other components. We now bound ρ and σ . We recall (2.25) and (2.27):

$$\begin{aligned} \underline{D}\rho + \frac{3}{2}\Omega \operatorname{tr} \underline{\chi} \rho &= -\Omega \left\{ \operatorname{div} \underline{\beta} + (2\underline{\eta} - \zeta, \underline{\beta}) - \frac{1}{2}(\widehat{\chi}, \underline{\alpha}) \right\}, \\ \underline{D}\sigma + \frac{3}{2}\Omega \operatorname{tr} \underline{\chi} \sigma &= -\Omega \left\{ \operatorname{curl} \underline{\beta} + (2\underline{\eta} - \zeta, * \underline{\beta}) + \frac{1}{2}\widehat{\chi} \wedge \underline{\alpha} \right\}. \end{aligned}$$

Because $k \geq 3$, we have $|\nabla \underline{\beta}| \lesssim \delta |u|^{-5}$. Thus,

$$\begin{aligned} \left| -\Omega \left\{ \operatorname{div} \underline{\beta} + (2\underline{\eta} - \zeta, \underline{\beta}) - \frac{1}{2}(\widehat{\chi}, \underline{\alpha}) \right\} \right| \\ + \left| -\Omega \left\{ \operatorname{curl} \underline{\beta} + (2\underline{\eta} - \zeta, * \underline{\beta}) + \frac{1}{2}\widehat{\chi} \wedge \underline{\alpha} \right\} \right| \lesssim \delta |u|^{-5}. \end{aligned}$$

We renormalize the first equation in the following form:

$$\begin{aligned} \underline{D} \left(\rho + \frac{2m_0}{|u|^3} \right) + \frac{3}{2} \Omega \operatorname{tr} \underline{\chi} \left(\rho + \frac{2m_0}{|u|^3} \right) \\ = \frac{3m_0}{|u|^3} \left(\Omega \operatorname{tr} \underline{\chi} + \frac{2}{|u|} \right) - \Omega \left\{ \operatorname{div} \underline{\beta} + (2\eta - \zeta, \underline{\beta}) - \frac{1}{2} (\widehat{\chi}, \underline{\alpha}) \right\}. \end{aligned}$$

In view of the fact that $|\Omega \operatorname{tr} \underline{\chi} + \frac{2}{|u|}| \lesssim \delta |u|^{-2}$ and the improved smallness for initial data on S_{δ, u_0} in Lemma 2, we integrate the above equations along \underline{C}_δ and we obtain

$$\left| \rho + \frac{2m_0}{|u|^3} \right| \lesssim \delta^{\frac{1}{2}} |u|^{-3} \quad \text{and} \quad |\sigma| \lesssim \delta^{\frac{1}{2}} |u|^{-3}.$$

Similarly, we can also run an induction argument as above to obtain

$$(4.5) \quad \left\| (|u| \nabla)^i \left(\rho + \frac{2m_0}{|u|^3}, \sigma \right) \right\|_{L^\infty} \lesssim \delta^{\frac{1}{2}} |u|^{-3} \quad \text{for } i \leq k-3,$$

$$|u|^{-\frac{1}{2}} \| (|u| \nabla)^{k-2} (\rho, \sigma) \|_{L^4(S_{\delta, u})}, |u|^{-1} \| (|u| \nabla)^{k-1} (\rho, \sigma) \|_{L^2(S_{\delta, u})} \lesssim \delta^{\frac{1}{2}} |u|^{-3}.$$

We now bound β . We examine the terms in (2.22):

$$\underline{D}\beta + \frac{1}{2} \Omega \operatorname{tr} \underline{\chi} \beta - \Omega \widehat{\chi} \cdot \beta = -\underline{\omega} \beta + \Omega \{ \not\partial \rho + {}^* \not\partial \sigma + 3\eta \rho + 3^* \eta \sigma + 2\widehat{\chi} \cdot \beta \}.$$

If $k \geq 4$, we have the bound $|\not\partial \rho + {}^* \not\partial \sigma| \lesssim \delta^{\frac{1}{2}} |u|^{-4}$ thanks to (4.5) and the bound $|-\underline{\omega} \beta + \Omega(3\eta \rho + 3^* \eta \sigma + 2\widehat{\chi} \cdot \beta)| \lesssim \delta^{\frac{1}{2}} |u|^{-5}$ from Theorem 1. We remark that there is a loss of derivatives here because we control β in terms of derivatives of ρ and σ via Bianchi equations. We then integrate the equation along \underline{C}_δ , and combined with the estimates on the initial data on S_{δ, u_0} obtained in Lemma 2, for $i \leq k-4$, we have

$$(4.6) \quad \| (|u| \nabla)^i \beta \|_{L^\infty(S_{\delta, u})} \lesssim \delta^{\frac{1}{2}} |u|^{-2},$$

$$|u|^{-\frac{1}{2}} \| (|u| \nabla)^{k-3} \beta \|_{L^4(S_{\delta, u})} + |u|^{-1} \| (|u| \nabla)^{k-2} \beta \|_{L^2(S_{\delta, u})} \lesssim \delta^{\frac{1}{2}} |u|^{-2}.$$

Notice that we only have L^∞ estimates on $(k-4)$ -th derivatives on β instead of $(k-3)$ -rd derivatives due to the derivative loss we just mentioned.

We now bound α . We rewrite (2.18) as follows:

$$\underline{D}\alpha - \frac{1}{2} \Omega \operatorname{tr} \underline{\chi} \alpha + 2\underline{\omega} \alpha - 2(\Omega \widehat{\chi}, \alpha) \not\partial = -\Omega \{ -\nabla \widehat{\otimes} \beta - (4\eta + \zeta) \widehat{\otimes} \beta + 3\widehat{\chi} \rho + 3^* \widehat{\chi} \sigma \},$$

where $2(\Omega \widehat{\chi}, \alpha)$ is the trace of $\underline{D}\alpha$. If $k \geq 5$, by (4.6), we have $|\nabla \widehat{\otimes} \beta| \lesssim \delta^{\frac{1}{2}} |u|^{-3}$. We remark that at this point there is a loss of derivatives here because we use the derivatives of β . We must also use $|\widehat{\chi}| + |\beta| \lesssim \delta^{\frac{1}{2}} |u|^{-2}$ by (4.3) and (4.6) to get $|-\Omega \{ -\nabla \widehat{\otimes} \beta - (4\eta + \zeta) \widehat{\otimes} \beta + 3\widehat{\chi} \rho + 3^* \widehat{\chi} \sigma \}| \lesssim \delta^{\frac{1}{2}} |u|^{-3}$. Therefore, in view

of the fact that $\alpha(\delta, u_0) = 0$ on S_{δ, u_0} , by integrating the equation along \underline{C}_δ , for $i \leq k - 5$, we obtain

$$(4.7) \quad \begin{aligned} \|(|u|\nabla)^i \alpha\|_{L^\infty(S_{\delta, u})} &\lesssim \delta^{\frac{1}{2}} |u|^{-2}, \\ |u|^{-\frac{1}{2}} \|(|u|\nabla)^{k-4} \alpha\|_{L^4(S_{\delta, u})}, |u|^{-1} \|(|u|\nabla)^{k-3} \alpha\|_{L^2(S_{\delta, u})} &\lesssim \delta^{\frac{1}{2}} |u|^{-2}. \end{aligned}$$

We remark that the terms $-\frac{1}{2}\Omega \operatorname{tr} \underline{\chi} \alpha + 2\omega \alpha - 2(\Omega \widehat{\underline{\chi}}, \alpha) \not\partial$ on the left-hand side are absorbed by Gronwall's inequality. We also want to point out that we will not use the $L^2(S_{\delta, u})$ estimates on $\nabla^{k-3} \alpha$ in sequel. Notice also that we only have L^∞ estimates on $(k - 5)$ -th derivatives on α due to the derivative loss we just mentioned.

Finally, we bound $\operatorname{tr} \chi$ and ω . We renormalize (2.17) and (2.8) as follows:

$$\begin{aligned} \underline{D} \left(\Omega \operatorname{tr} \chi - \left(\frac{2}{|u|} - \frac{4m_0}{|u|^2} \right) \right) + \frac{1}{2} \Omega \operatorname{tr} \underline{\chi} \left(\Omega \operatorname{tr} \chi - \left(\frac{2}{|u|} - \frac{4m_0}{|u|^2} \right) \right) \\ = -\frac{1}{|u|} \left(1 - \frac{2m_0}{|u|} \right) \left(\Omega \operatorname{tr} \underline{\chi} + \frac{2}{|u|} \right) \\ + \left(2\Omega^2 \rho + \frac{4m_0}{|u|^3} \right) + \Omega^2 (2\operatorname{div} \eta + 2|\eta|^2 - (\widehat{\underline{\chi}}, \widehat{\underline{\chi}})), \\ \underline{D} \left(\omega - \frac{m_0}{|u|^2} \right) = - \left(\Omega^2 \rho + \frac{2m_0}{|u|^3} \right) + \Omega^2 (2(\eta, \underline{\eta}) - |\underline{\eta}|^2). \end{aligned}$$

We have already obtained the following estimates: $|\Omega \operatorname{tr} \underline{\chi} + \frac{2}{|u|}| \lesssim \delta |u|^{-2}$, $|\Omega^2 \rho + \frac{2m_0}{|u|^3}| \lesssim \delta^{\frac{1}{2}} |u|^{-3}$, $|\Omega - 1| \lesssim \delta |u|^{-2}$, $|\Omega^2 (2\operatorname{div} \eta + 2|\eta|^2 - (\widehat{\underline{\chi}}, \widehat{\underline{\chi}})) + |\Omega^2 (2(\eta, \underline{\eta}) - |\underline{\eta}|^2)| \lesssim \delta^{\frac{1}{2}} |u|^{-3}$ and $|\Omega \operatorname{tr} \chi(\delta, u_0) - (\frac{2}{|u|} - \frac{4m_0}{|u|^2})| \lesssim \delta^{\frac{1}{2}} |u|^{-2}$. In view of the fact that $\omega(\delta, u_0) = 0$, by integrating the equation along \underline{C}_δ , we obtain

$$\left\| \operatorname{tr} \chi - \left(\frac{2}{|u|} - \frac{4m_0}{|u|^2} \right) \right\|_{L^\infty} + \left\| \omega - \frac{m_0}{|u|^2} + \frac{m_0}{|u_0|^2} \right\|_{L^\infty} \lesssim \delta^{\frac{1}{2}} |u|^{-2},$$

where $\Omega \operatorname{tr} \chi$ has been replaced by $\operatorname{tr} \chi$ thanks again to $|\Omega - 1| \lesssim \delta |u|^{-2}$. Similarly, we can derive estimates for higher order derivatives. Finally, for $i \leq k - 3$, we have

$$(4.8) \quad \begin{aligned} \left\| (|u|\nabla)^i \left(\operatorname{tr} \chi - \left(\frac{2}{|u|} - \frac{4m_0}{|u|^2} \right), \omega - \frac{m_0}{|u|^2} + \frac{m_0}{|u_0|^2} \right) \right\|_{L^\infty(S_{\delta, u})} &\lesssim \delta^{\frac{1}{2}} |u|^{-3}, \\ |u|^{-\frac{1}{2}} \|(|u|\nabla)^{k-2} (\operatorname{tr} \chi, \omega)\|_{L^4(S_{\delta, u})}, |u|^{-1} \|(|u|\nabla)^{k-1} (\operatorname{tr} \chi, \omega)\|_{L^2(S_{\delta, u})} &\lesssim \delta^{\frac{1}{2}} |u|^{-3}. \end{aligned}$$

This completes the proof of the lemma. □

4.3. *Geometry on $C_{u_0}^{[\delta, \delta+1]}$.* We extend the initial data on $C_{u_0}^{[0, \delta]}$ to $C_{u_0}^{[0, \delta+1]}$ by setting $\widehat{\underline{\chi}} \equiv 0$ for $\underline{u} \in [\delta, \delta + 1]$, and we will derive estimates on $C_{u_0}^{[\delta, \delta+1]}$ for all the connection coefficients and curvature components. For the sake of simplicity, we denote $C_{u_0} = C_{u_0}^{[\delta, \delta+1]}$ in this subsection.

We first give another initial data on $C_{u_0} \cup \underline{C}_\delta$, and it is in fact an initial data set for the Schwarzschild space-time with mass m_0 . To distinguish this new data from the old one, we shall use a lower index m_0 for its connection coefficients and curvature components. The new initial data set consists of the following quantities: a metric $g_{m_0}(\delta, u_0)$ on S_{δ, u_0} that is the round metric of the sphere with radius $|u_0|$, the torsion $\zeta_{m_0}(\delta, u_0) \equiv 0$ on S_{δ, u_0} , two null expansions $\text{tr } \chi_{m_0}(\delta, u_0) \equiv \frac{2}{|u_0|} - \frac{4m_0}{|u_0|^2}$ and $\text{tr } \underline{\chi}_{m_0}(\delta, u_0) \equiv -\frac{2}{|u_0|}$ on S_{δ, u_0} , the lapse function $\Omega_{m_0} \equiv 1$ on $C_{u_0} \cup \underline{C}_\delta$ and two shears $\widehat{\chi}_{m_0} \equiv 0$ on C_{u_0} and $\widehat{\underline{\chi}}_{m_0} \equiv 0$ on \underline{C}_δ .

We also impose the condition $|u_0| > 2m_0$. This condition guarantees that the future development of the data is in the domain of outer communication of the Schwarzschild space-time. We remark that the only nonzero null components for the new data on C_{u_0} are $\text{tr } \chi_{m_0}$, $\text{tr } \underline{\chi}_{m_0}$, $\underline{\omega}_{m_0}$ and ρ_{m_0} with $\text{tr } \chi_{m_0} > 0$, and the only nonzero null components for the new data on \underline{C}_δ are $\text{tr } \chi_{m_0}$, $\text{tr } \underline{\chi}_{m_0}$, $\underline{\omega}_{m_0}$ and ρ_{m_0} .

We now restate the estimates on the sphere S_{δ, u_0} derived in Lemma 2 as well as the estimates given in Chapter 2 of [2]: we have $\widehat{\chi} \equiv 0$, $\alpha \equiv 0$, $\omega \equiv 0$, and for all k ,

$$(4.9) \quad |\nabla^k(\widehat{\chi}, \underline{\eta}, \underline{\eta}, \underline{\omega} - \underline{\omega}_{m_0}, \text{tr } \chi - \text{tr } \chi_{m_0}, \text{tr } \underline{\chi} - \text{tr } \underline{\chi}_{m_0}, \beta, \rho - \rho_{m_0}, \sigma, \underline{\beta}, \underline{\alpha})| \lesssim_k \delta^{\frac{1}{2}}.$$

Notice that we have dropped the weight $|u_0|$. We will not use the weight in the sequel since we can fix a $|u_0| > 2m_0$ now, and it is not relevant from now on. We shall prove that the above estimates actually hold on C_{u_0} :

LEMMA 4. *On the incoming cone $C_{u_0}^{[\delta, \delta+1]}$, we have $\widehat{\chi} \equiv 0$, $\alpha \equiv 0$, $\omega \equiv 0$, and for all k , we have*

$$|\nabla^k(\widehat{\chi}, \underline{\eta}, \underline{\eta}, \underline{\omega} - \underline{\omega}_{m_0}, \text{tr } \chi - \text{tr } \chi_{m_0}, \text{tr } \underline{\chi} - \text{tr } \underline{\chi}_{m_0}, \beta, \rho - \rho_{m_0}, \sigma, \underline{\beta}, \underline{\alpha})| \lesssim_k \delta^{\frac{1}{2}}.$$

Proof. By construction, we also have $\widehat{\chi} = 0$, $\alpha = 0$ and $\omega = 0$ on C_{u_0} . The idea to prove the estimate is to use null structure equations in a correct order.

We first control $\text{tr } \chi$. According to (4.9) and choosing δ sufficiently small, we have $\text{tr } \chi(\delta, u_0) > 0$ so that $\text{tr } \chi$ can be solved⁸ for $\underline{u} \in [\delta, \delta + 1]$ and $0 < \text{tr } \chi(\underline{u}, u_0) < \text{tr } \chi(\delta, u_0)$. We then rewrite (2.2) as

$$(4.10) \quad D(\text{tr } \chi - \text{tr } \chi_{m_0}) = -\frac{1}{2}(\text{tr } \chi + \text{tr } \chi_{m_0})(\text{tr } \chi - \text{tr } \chi_{m_0}).$$

⁸ Since $\widehat{\chi} \equiv 0$, $\text{tr } \chi$ satisfies the following ordinary differential equation: $D\text{tr } \chi = -\frac{1}{2}(\text{tr } \chi)^2$. A positive initial value for this equation leads to global solution. In particular, the solution exists on $[\delta, \delta + 1]$.

Because $(\text{tr } \chi + \text{tr } \chi_{m_0})$ is bounded pointwisely, thanks to Gronwall's inequality and (4.9), we have

$$|\text{tr } \chi - \text{tr } \chi_{m_0}| \lesssim \left| \text{tr } \chi(\delta, u_0) - \left(\frac{2}{|u_0|} - \frac{4}{|u_0|^2} \right) \right| \lesssim \delta^{\frac{1}{2}}.$$

We can also apply ∇^k to the above argument and use an induction argument as before to derive

$$|\nabla^k(\text{tr } \chi - \text{tr } \chi_{m_0})| \lesssim \left| \nabla^k \left(\text{tr } \chi(\delta, u_0) - \left(\frac{2}{|u_0|} - \frac{4}{|u_0|^2} \right) \right) \right| \lesssim_k \delta^{\frac{1}{2}}.$$

We now bound $\eta = -\underline{\eta}$. We eliminate β in (2.5) and (2.10) to derive

$$D\eta + \text{tr } \chi \eta = -\frac{1}{2} \text{div } \text{tr } \chi.$$

Therefore, by Gronwall's inequality and an induction argument, we obtain

$$|\nabla^k(\eta, \underline{\eta})(\underline{u}, u_0)| \lesssim_k \delta^{\frac{1}{2}}.$$

Before going to the next component, we have to digress to the estimates for the Gauss curvature K on $S_{\underline{u}, u_0}$. Recall that K satisfies the following propagation equations (see Chapter 5 of [2]):

$$DK + \Omega \text{tr } \chi K = \text{div } \text{div } (\Omega \widehat{\chi}) - \frac{1}{2} \Delta (\Omega \text{tr } \chi).$$

We rewrite the equation on C_{u_0} as

$$D(K - K_{m_0}) + \text{tr } \chi(K - K_{m_0}) = -(\text{tr } \chi - \text{tr } \chi_{m_0})K_{m_0} - \frac{1}{2} \Delta \text{tr } \chi,$$

which yields $|\nabla^k(K(\underline{u}, u_0) - K_{m_0}(\underline{u}, u_0))| \lesssim_k \delta^{\frac{1}{2}}$.

To bound $\underline{\chi}$ and $\widehat{\chi}$, we eliminate ρ in (2.15) and (2.9) to derive the following propagation equation for $\text{tr } \underline{\chi} - \text{tr } \underline{\chi}_{m_0}$:

$$D(\text{tr } \underline{\chi} - \text{tr } \underline{\chi}_{m_0}) + \text{tr } \chi(\text{tr } \underline{\chi} - \text{tr } \underline{\chi}_{m_0}) = 2(K_{m_0} - K) - \text{div } \eta + |\eta|^2,$$

and we rewrite (2.14) as

$$D\widehat{\chi} - \frac{1}{2} \text{tr } \chi \widehat{\chi} = -\nabla \widehat{\otimes} \eta + \eta \widehat{\otimes} \eta.$$

These equations yield $|\nabla^k(\text{tr } \underline{\chi}(\underline{u}, u_0) - \text{tr } \underline{\chi}_{m_0}(\underline{u}, u_0))| + |\nabla^k \widehat{\chi}(\underline{u}, u_0)| \lesssim_k \delta^{\frac{1}{2}}$.

We then use (2.10), (2.11) and the estimates above to derive $|\nabla^k(\beta, \underline{\beta})(\underline{u}, u_0)| \lesssim_k \delta^{\frac{1}{2}}$.

We rewrite (2.9) as

$$K - K_{m_0} = -\frac{1}{4}(\text{tr } \chi \text{tr } \underline{\chi} - \text{tr } \chi_{m_0} \text{tr } \underline{\chi}_{m_0}) - (\rho - \rho_{m_0}),$$

which implies that $|\nabla^k(\rho(\underline{u}, u_0) - \rho_{m_0}(\underline{u}, u_0))| \lesssim_k \delta^{\frac{1}{2}}$.

We rewrite (2.7) as

$$D(\underline{\omega} - \underline{\omega}_{m_0}) = -3|\eta|^2 - (\rho - \rho_{m_0}),$$

which implies $|\nabla^k(\underline{\omega}(\underline{u}, u_0) - \underline{\omega}_{m_0}(\underline{u}, u_0))| \lesssim_k \delta^{\frac{1}{2}}$.

Finally, (2.19) implies $|\nabla^k \underline{\alpha}(\underline{u}, u_0)| \lesssim_k \delta^{\frac{1}{2}}$. This completes the proof of the lemma. \square

4.4. *Construction of the transition region.* We consider a characteristic initial data problem on $C_{u_0}^{[\delta, \delta+1]} \cup C_{\underline{u}}^{[u_0, -1-\delta]}$ where the data are given by $\widehat{\chi} = 0$ on $C_{u_0}^{[\delta, \delta+1]}$ and the data on $C_{\underline{u}}^{[u_0, -1-\delta]}$ is induced by the solution constructed by Christodoulou in [2] under condition (2.30) with $|u_0| > 2m_0$. We continue to use the space-time metric g to denote the solution of this problem.

Adapted to this problem, we introduce the following norms:

$$\begin{aligned} \mathcal{R}_k^u(u) &= \|\nabla^k(\alpha, \beta, \rho - \rho_{m_0}, \sigma, \underline{\beta})\|_{L^2(C_u^u)}, \\ \mathcal{R}_k^{\underline{u}}(\underline{u}) &= \|\nabla^k(\beta, \rho - \rho_{m_0}, \sigma, \underline{\beta}, \underline{\alpha})\|_{L^2(C_{\underline{u}}^u)}, \\ \mathcal{O}_k(\underline{u}, u) &= \left\| \nabla^k(\widehat{\chi}, \widehat{\chi}, \eta, \underline{\eta}, \text{tr } \chi - \text{tr } \chi_{m_0}, \text{tr } \underline{\chi} - \text{tr } \underline{\chi}_{m_0}, \right. \\ &\quad \left. \omega - \omega_{m_0}, \underline{\omega} - \underline{\omega}_{m_0}) \right\|_{L^2(S_{\underline{u}, u})}. \end{aligned}$$

We use short-hand notation C_u^u referring to $C_u^{[\delta, u]}$ and $C_{\underline{u}}^u$ referring to $C_{\underline{u}}^{[u_0, u]}$. We take a small parameter ϵ to be proportional to $\delta^{\frac{1}{2}}$, then the results in Section 4.1, Section 4.2 and Section 4.3 can be summarized in the following proposition:

PROPOSITION 1. *Fix $k \gg 5$. For any $\epsilon > 0$, there exists $\delta_0 > 0$ depending only on the C^{k+N} bounds of the seed data for some sufficient large N and on ϵ , such that for all $\delta < \delta_0$ and for all $(\underline{u}, u) \in [\delta, \delta + 1] \times [u_0, -1 - \delta]$, we have,*

$$\begin{aligned} \mathcal{R}_{\leq k-2}^{\delta+1}(u_0) + \mathcal{R}_{\leq k-2}^{-1-\delta}(\delta) + \mathcal{O}_{\leq k-2}(\underline{u}, u_0) + \mathcal{O}_{\leq k-2}(\delta, u) &\leq \epsilon, \\ \|\nabla^{\leq k-4} \alpha\|_{L^4(S_{\delta, u})} + \|\nabla^{\leq k} \underline{\alpha}\|_{L^\infty(S_{\underline{u}, u_0})} &\leq \epsilon. \end{aligned}$$

Based on this proposition, we will prove the following theorem in this section:

THEOREM 2. *Let $k \gg 5$ and $\epsilon > 0$. If $\delta > 0$ is sufficiently small depending on C^{k+N} bounds of the seed data for sufficient large N and on $\epsilon > 0$, then there is a unique smooth solution (M_{ϵ_0}, g) of vacuum Einstein field equations to the characteristic initial data problem described above. The space-time M_{ϵ_0} corresponds to the region $\delta \leq \underline{u} \leq \delta + \epsilon_0$ and $u_0 \leq u \leq u_0 + \epsilon_0$. The parameter $\epsilon_0 > 0$ depends only on m_0 and u_0 and it is independent of δ when δ is sufficiently small. Moreover, the space-time (M_{ϵ_0}, g) is ϵ -close to the Schwarzschild space-time with mass m_0 in C^{k-3} norms, namely,*

$$\|g - g_{m_0}\|_{C^{k-3}(M_{\epsilon_0}, g_{m_0})} \lesssim \epsilon.$$

Remark 9. We would like to emphasize that once m_0 and u_0 are fixed, the size of ϵ_0 is also fixed. In the introduction, we mentioned that in order to use the Corvino-Schoen construction, the gluing region must have a fixed size when δ goes to zero. The size of ϵ_0 meets this requirement.

Remark 10. It is not hard to see that g_{m_0} is isometric to the Schwarzschild space-time with mass m_0 . In fact, the initial data for g_{m_0} on $C_{u_0}^{[\delta, \delta + \epsilon_0]} \cup \underline{C}_\delta$ is spherically symmetric; then g_{m_0} is also spherically symmetric. The orbits of the isometry group $SO(3)$ are simply $S_{\underline{u}, u}$. Therefore by Birkhoff's Theorem, g_{m_0} is isometric to the Schwarzschild space-time. The mass parameter m_0 can be figured out by computing the Hawking mass of S_{δ, u_0} in g_{m_0} .

Proof. The existence of local solutions is due to the classical result [14] of Rendall. We make use of a recent improvement [12] due to Luk. Rendall's original proof only gave the local existence around the intersection of two initial null hypersurface, and Luk showed that the existence region can be enlarged to a full neighborhood of initial hypersurfaces in their future development. The main advantage of [12] is that it also gives the estimates on the solution that depend only on the initial data. We now rephrase the results of [12] in our situation: there exists ϵ_0 that depends only on m_0 and u_0 and is independent of δ hence ϵ , such that we can solve the metric g for $(\underline{u}, u) \in [\delta, \delta + \epsilon_0] \times [u_0, u_0 + \epsilon_0]$. Moreover, we have the following bounds:

PROPOSITION 2. *With assumptions as in Theorem 2, we have a smooth solution (M_{ϵ_0}, g) where the M_{ϵ_0} corresponds to $(\underline{u}, u) \in [\delta, \delta + 1] \times [u_0, -1 - \delta]$. Moreover, we have the following estimates:*

$$\mathcal{R}^{\delta + \epsilon_0}(u) + \underline{\mathcal{R}}^{u_0 + \epsilon_0}(\underline{u}) + \mathcal{O}(\underline{u}, u_0) + \mathcal{O}(\delta, u) \leq C(m_0, u_0, \epsilon_0).$$

We also have the Sobolev inequalities

$$\begin{aligned} \|\phi\|_{L^4(S_{\underline{u}, u})} &\leq C(m_0, u_0, \epsilon_0)(\|\nabla\phi\|_{L^2(S_{\underline{u}, u})} + \|\phi\|_{L^2(S_{\underline{u}, u})}), \\ \|\phi\|_{L^\infty(S_{\underline{u}, u})} &\leq C(m_0, u_0, \epsilon_0) \sum_{i=0}^2 \|\nabla^i \phi\|_{L^2(S_{\underline{u}, u})}. \end{aligned}$$

For the sake of simplicity, we will suppress the lower index ≤ 3 in the norms, e.g., $\mathcal{R} = \mathcal{R}_{\leq 3}$. In the remaining part of the current section, C refers to various constants depending only on m_0 , u_0 and ϵ_0 and $A \lesssim B$ refers to $A \leq CB$.

Since we expect the solution is close to some Schwarzschild space-time, we have to derive more precise energy estimates on M_{ϵ_0} . Combined with Rendall's result [14] and usual bootstrap arguments, our estimates will be good enough to prove existence. We remark that Luk's result [12] is extremely convenient

to use at this stage: we can avoid a long bootstrap argument, and we can take the existence for granted.

We start to prove Theorem 2 by deriving the estimates only for $\mathcal{R}_i, \mathcal{B}_i$ and \mathcal{O}_i for $i = 0, 1, 2, 3$. We remark that in this case $k = 5$ (k is as in Proposition 1) is sufficient. The estimates for higher order derivatives can be obtained by a routine induction argument. Since it is similar to the higher order energy estimates derived in Section 3, we will omit the proof.

We first prove the following lemma, which claims that the connection coefficients can be controlled by the initial data and the curvature components.

LEMMA 5. *We have the following estimate:*

$$\mathcal{O}(\underline{u}, u) \lesssim \sup_{\delta \leq u' \leq \underline{u}} (\mathcal{O}(\underline{u}', u_0) + \mathcal{R}^u(\underline{u}')) + \sup_{u_0 \leq u' \leq u} (\mathcal{O}(\delta, u') + \mathcal{R}^u(u')).$$

Proof. The proof is once again by integrating the null structure equations (2.1)–(2.8). At the beginning, we remark that for a tensor ϕ , we have

$$C^{-1} \sum_{i=0}^3 \|\nabla^i(\Omega\phi)\|_{L^2(S_{\underline{u},u})} \leq \sum_{i=0}^3 \|\nabla^i\phi\|_{L^2(S_{\underline{u},u})} \leq C \sum_{i=0}^3 \|\nabla^i(\Omega\phi)\|_{L^2(S_{\underline{u},u})}.$$

This can be derived by the bounds for Ω and $\not{d}\log\Omega = (\eta + \underline{\eta})/2$. Similar inequalities also hold for $L^4(S_{\underline{u},u})$ and $L^\infty(S_{\underline{u},u})$ norms.

We start with the bounds on $\hat{\chi}$. In view of (2.1), we have

$$D|\hat{\chi}'|^2 + 2\Omega \operatorname{tr}\chi|\hat{\chi}'|^2 \leq 2|\hat{\chi}'|\alpha|.$$

We then use the L^∞ bounds for Ω and $\operatorname{tr}\chi$ in Proposition 2 as well as the Sobolev inequalities to derive

$$\begin{aligned} \|\hat{\chi}'\|_{L^\infty(S_{\underline{u},u})} &\lesssim \|\hat{\chi}'\|_{L^\infty(S_{\delta,u})} + \int_{\delta}^u \|\alpha\|_{L^\infty(S_{\underline{u}',u})} \\ &\lesssim \sum_{i=0}^2 \left(\|\nabla^i\hat{\chi}'\|_{L^2(S_{\delta,u})} + \int_{\delta}^u \|\nabla^i\alpha\|_{L^2(S_{\underline{u}',u})} \right) \lesssim \mathcal{O}(\delta, u) + \mathcal{R}^u(u). \end{aligned}$$

We now turn to the derivatives of $\hat{\chi}$. We first rewrite (2.1) as $D\hat{\chi}' = \Omega^2|\hat{\chi}'|^2\not{g} - \alpha$ and then commute with ∇ to derive

$$D\nabla\hat{\chi}' = \nabla(\Omega\chi) \cdot \hat{\chi}' + \nabla(\Omega^2)|\hat{\chi}'|^2\not{g} + \Omega^2(\hat{\chi}', \nabla\hat{\chi}')\not{g} - \nabla\alpha.$$

The last term $\nabla\alpha$ can be bounded by $\mathcal{R}^u(u)$. For the remaining terms, each of them contains a factor $|\hat{\chi}'|$. We can bound the other factors by C and bound $|\hat{\chi}'(\underline{u}, u)|$ by $\mathcal{O}(\delta, u) + \mathcal{R}^u(u)$. Thus, we obtain

$$\|\nabla\hat{\chi}'\|_{L^\infty(S_{\underline{u},u})} \lesssim \mathcal{O}(\delta, u) + \mathcal{R}^u(u).$$

The estimates on higher order derivatives are similar. For second and third derivatives of $\widehat{\chi}$, we have

$$D\mathcal{V}^k \widehat{\chi}' = \sum_{i=0}^{k-1} \mathcal{V}^i \mathcal{V}'(\Omega\chi) \cdot \mathcal{V}^{k-1-i} \widehat{\chi}' + \sum_{i=0}^k \mathcal{V}^i(\Omega^2) \cdot \mathcal{V}^{k-i} |\widehat{\chi}'|^2 \mathcal{V}' - \mathcal{V}^k \alpha.$$

When $k = 2$, for the first and the second terms, we can use Hölder's inequality by placing the factors with highest order derivatives in $L^4(S_{\underline{u},u})$ and the others in L^∞ . In this way, for each product, one factor is bounded by $\mathcal{O}(\delta, u) + \mathcal{R}^u(u)$ and the other one is bounded by C . The last term $\mathcal{V}^k \alpha$ is also bounded by $\mathcal{R}^u(u)$. When $k = 3$, there are two cases: if there are factors containing third order derivatives, we can bound it in $L^2(S_{\underline{u},u})$ and bound the other factors in L^∞ ; if not, we bound the factors containing second order derivatives in $L^4(S_{\underline{u},u})$ and bound the other in L^∞ . Therefore, we obtain

$$\|\mathcal{V}^2 \widehat{\chi}'\|_{L^4(S_{\underline{u},u})} + \|\mathcal{V}^3 \widehat{\chi}'\|_{L^2(S_{\underline{u},u})} \lesssim \mathcal{O}(\delta, u) + \mathcal{R}^u(u).$$

For $\widehat{\chi}$, in a similar manner, we can use (2.3) to derive

$$\|\widehat{\chi}\|_{L^\infty} + \|\mathcal{V} \widehat{\chi}\|_{L^\infty} + \|\mathcal{V}^2 \widehat{\chi}\|_{L^4(S_{\underline{u},u})} + \|\mathcal{V}^3 \widehat{\chi}\|_{L^2(S_{\underline{u},u})} \lesssim \mathcal{O}(\underline{u}, u_0) + \mathcal{R}^u(\underline{u}).$$

For η and $\underline{\eta}$, we consider (2.5) and (2.6) as a coupled system and we derive

$$\begin{aligned} \|\eta\|_{L^\infty(S_{\underline{u},u})} &\lesssim \|\eta\|_{L^\infty(S_{\delta,u})} + \int_{\delta}^u \|\underline{\eta}\|_{L^\infty(S_{\underline{u}',u})} d\underline{u}' + \int_{\delta}^u \|\beta\|_{L^\infty(S_{\underline{u}',u})} d\underline{u}', \\ \|\underline{\eta}\|_{L^\infty(S_{\underline{u},u})} &\lesssim \|\underline{\eta}\|_{L^\infty(S_{\underline{u},u_0})} + \int_{u_0}^u \|\eta\|_{L^\infty(S_{\underline{u},u'})} d\underline{u}' + \int_{u_0}^u \|\underline{\beta}\|_{L^\infty(S_{\underline{u},u'})} d\underline{u}'. \end{aligned}$$

We substitute the second inequality to the first one and take $\sup_{\delta \leq \underline{u}' \leq \underline{u}}$ on both sides, and we then have

$$\begin{aligned} \sup_{\delta \leq \underline{u}' \leq \underline{u}} \|\eta\|_{L^\infty(S_{\underline{u}',u})} &\lesssim \|\eta\|_{L^\infty(S_{\delta,u})} + \sup_{\delta \leq \underline{u}' \leq \underline{u}} \|\underline{\eta}\|_{L^\infty(S_{\underline{u}',u_0})} \\ &\quad + \int_{u_0}^u \sup_{\delta \leq \underline{u}' \leq \underline{u}} \|\eta\|_{L^\infty(S_{\underline{u}',u'})} d\underline{u}' \\ &\quad + \int_{u_0}^u \sup_{\delta \leq \underline{u}' \leq \underline{u}} \|\underline{\beta}\|_{L^\infty(S_{\underline{u}',u'})} d\underline{u}' + \int_{\delta}^u \|\beta\|_{L^\infty(S_{\underline{u}',u})} d\underline{u}'. \end{aligned}$$

By the Gronwall and Sobolev inequalities, we have

$$\begin{aligned} \sup_{\delta \leq \underline{u}' \leq \underline{u}} \|\eta\|_{L^\infty(S_{\underline{u}',u})} &\lesssim \|\eta\|_{L^\infty(S_{\delta,u})} + \sup_{\delta \leq \underline{u}' \leq \underline{u}} \|\underline{\eta}\|_{L^\infty(S_{\underline{u},u_0})} \\ &\quad + \int_{u_0}^u \sup_{\delta \leq \underline{u}' \leq \underline{u}} \|\beta\|_{L^\infty(S_{\underline{u}',u'})} d\underline{u}' + \int_{\delta}^u \|\beta\|_{L^\infty(S_{\underline{u}',u})} d\underline{u}' \\ &\lesssim \mathcal{O}(\delta, u) + \sup_{\delta \leq \underline{u}' \leq \underline{u}} \mathcal{O}(\underline{u}', u_0) + \sup_{\delta \leq \underline{u}' \leq \underline{u}} \mathcal{R}^u(\underline{u}') + \mathcal{R}^u(u). \end{aligned}$$

We then define $\mathcal{P}(\underline{u}, u)$ and rewrite the above estimates as

$$\begin{aligned} \sup_{\delta \leq u' \leq u} \|\underline{\eta}\|_{L^\infty(\underline{u}', u)} &\lesssim \mathcal{P}(\underline{u}, u) \\ &\triangleq \sup_{\delta \leq u' \leq u} (\mathcal{O}(\underline{u}', u_0) + \underline{\mathcal{R}}^u(\underline{u}')) + \sup_{u_0 \leq u' \leq u} (\mathcal{O}(\delta, u') + \mathcal{R}^u(u')). \end{aligned}$$

We turn to the derivatives of η and $\underline{\eta}$. Recall that we can commute derivatives with (2.5) and (2.6) to derive (3.14) and (3.16). When $k = 1$, we have

$$\begin{aligned} \|\nabla \eta\|_{L^\infty(S_{\underline{u}, u})} &\lesssim \|\nabla \eta\|_{L^\infty(S_{\delta, u})} \\ &\quad + \int_{\delta}^u \|\nabla \underline{\eta}\|_{L^\infty(S_{\underline{u}', u})} d\underline{u}' + \int_{\delta}^u \|\nabla \beta\|_{L^\infty(S_{\underline{u}', u})} d\underline{u}' + \mathcal{P}(\underline{u}, u), \\ \|\nabla \underline{\eta}\|_{L^\infty(S_{\underline{u}, u})} &\lesssim \|\nabla \underline{\eta}\|_{L^\infty(S_{\underline{u}, u_0})} \\ &\quad + \int_{u_0}^u \|\nabla \eta\|_{L^\infty(S_{\underline{u}, u'})} du' + \int_{u_0}^u \|\nabla \underline{\beta}\|_{L^\infty(S_{\underline{u}, u'})} du' + \mathcal{P}(\underline{u}, u). \end{aligned}$$

By a similar argument as above, we obtain

$$\|\nabla \eta\|_{L^\infty(S_{\underline{u}, u}), \|\nabla \underline{\eta}\|_{L^\infty(S_{\underline{u}, u})} \lesssim \mathcal{P}(\underline{u}, u).$$

Similar arguments also apply to the cases when $k = 2, 3$. We finally obtain

$$\|\nabla^2(\eta, \underline{\eta})\|_{L^4(S_{\underline{u}, u}), \|\nabla^3(\eta, \underline{\eta})\|_{L^2(S_{\underline{u}, u})} \lesssim \mathcal{P}(\underline{u}, u).$$

For ω and $\underline{\omega}$, the estimates rely on (2.7) and (2.8). We rewrite (2.7) as

$$(4.11) \quad D(\underline{\omega} - \underline{\omega}_{m_0}) = (\Omega^2 - \Omega_{m_0}^2)(2(\eta, \underline{\eta}) - |\eta|^2 - \rho) + \Omega_{m_0}^2(2(\eta, \underline{\eta}) - |\eta|^2 + \rho - \rho_{m_0}).$$

In view of the facts that $D \log \Omega = \underline{\omega}$, $|\Omega| + |\Omega^{-1}| \leq C$, we have

$$\begin{aligned} \|\Omega - \Omega_{m_0}\|_{L^\infty(S_{\underline{u}, u})} &\lesssim \|\log \Omega - \log \Omega_{m_0}\|_{L^\infty(S_{\underline{u}, u})} \\ &\lesssim \int_{u_0}^u \|\underline{\omega} - \underline{\omega}_{m_0}\|_{L^\infty(S_{\underline{u}, u'})} du' \\ &\lesssim \sup_{u_0 \leq u' \leq u} \|\underline{\omega} - \underline{\omega}_{m_0}\|_{L^\infty(S_{\underline{u}, u'})}. \end{aligned}$$

We can integrate (4.11) to derive

$$\begin{aligned} \|\underline{\omega} - \underline{\omega}_{m_0}\|_{L^\infty(S_{\underline{u}, u})} &\lesssim \|\underline{\omega} - \underline{\omega}_{m_0}\|_{L^\infty(S_{\delta, u})} \\ &\quad + \int_0^u \|\Omega - \Omega_{m_0}\|_{L^\infty(S_{\underline{u}', u})} du' + \mathcal{P}(\underline{u}, u) \\ &\lesssim \int_0^u \sup_{u_0 \leq u' \leq u} \|\underline{\omega} - \underline{\omega}_{m_0}\|_{L^\infty(S_{\underline{u}', u'})} du' + \mathcal{P}(\underline{u}, u). \end{aligned}$$

Therefore, Gronwall's inequality yields

$$\sup_{u_0 \leq u' \leq u} \|\underline{\omega} - \underline{\omega}_{m_0}\|_{L^\infty(S_{\underline{u}, u'})} \lesssim \mathcal{P}(\underline{u}, u).$$

As a byproduct, we also have

$$\|\Omega - \Omega_{m_0}\|_{L^\infty(S_{\underline{u},u})} \lesssim \mathcal{P}(\underline{u}, u).$$

We then rewrite (2.8) as

$$\underline{D}(\omega - \omega_{m_0}) = (\Omega^2 - \Omega_{m_0}^2)(2(\eta, \underline{\eta}) - |\underline{\eta}|^2 - \rho) + \Omega_{m_0}^2(2(\eta, \underline{\eta}) - |\underline{\eta}|^2 + \rho - \rho_{m_0}).$$

We can integrate this equation and use the bounds for $|\Omega - \Omega_{m_0}|$ to derive

$$\|\omega - \omega_{m_0}\|_{L^\infty(S_{\underline{u},u})} \lesssim \mathcal{P}(\underline{u}, u).$$

We turn to the derivatives of ω and $\underline{\omega}$. Since ω_{m_0} and $\underline{\omega}_{m_0}$ are constants on $S_{\underline{u},u}$, applying ∇^k on (2.7) and (2.8) will kill those terms. We also notice that ρ_{m_0} is constant on $S_{\underline{u},u}$ and $\not{d} \log \Omega = (\eta + \underline{\eta})/2$ has already been bounded by \mathcal{P} . We then take derivatives

$$\begin{aligned} \underline{D}\nabla^k \omega &= \sum_{i=0}^{k-2} \nabla^i \nabla (\Omega \underline{\chi}) \cdot \nabla^{k-2-i} \not{d}\omega + \nabla^k (\Omega^2(2(\eta, \underline{\eta}) - |\underline{\eta}|^2 - \rho)), \\ \underline{D}\nabla^k \underline{\omega} &= \sum_{i=0}^{k-2} \nabla^i \nabla (\Omega \underline{\chi}) \cdot \nabla^{k-2-i} \not{d}\underline{\omega} + \nabla^k (\Omega^2(2(\eta, \underline{\eta}) - |\underline{\eta}|^2 - \rho)). \end{aligned}$$

For $k = 1, 2, 3$, we can integrate as before to derive

$$\|\not{d}(\omega, \underline{\omega})\|_{L^\infty(S_{\underline{u},u})} + \|\nabla^2(\omega, \underline{\omega})\|_{L^4(S_{\underline{u},u})} + \|\nabla^3(\omega, \underline{\omega})\|_{L^2(S_{\underline{u},u})} \lesssim \mathcal{P}(\underline{u}, u).$$

Finally, we consider $\text{tr } \chi'$ and $\text{tr } \underline{\chi}'$. We rewrite (2.2) as

$$D(\text{tr } \chi' - \text{tr } \chi'_{m_0}) = -\frac{1}{2}(\text{tr } \chi + \text{tr } \chi_{m_0})(\Omega \text{tr } \chi' - \Omega_{m_0} \text{tr } \chi'_{m_0}) - |\widehat{\chi}|^2.$$

We can bound $|\Omega \text{tr } \chi - \Omega_{m_0} \text{tr } \chi'_{m_0}| \lesssim |\Omega - \Omega_{m_0}| + |\text{tr } \chi' - \text{tr } \chi'_{m_0}|$. We then integrate the above equation to derive

$$\begin{aligned} \|\text{tr } \chi' - \text{tr } \chi'_{m_0}\|_{L^\infty(S_{\underline{u},u})} &\lesssim \|\text{tr } \chi' - \text{tr } \chi'_{m_0}\|_{L^\infty(S_{\delta,u})} \\ &\quad + \int_{\delta}^u \|\text{tr } \chi' - \text{tr } \chi'_{m_0}\|_{L^\infty(S_{\underline{u}',u})} d\underline{u}' \\ &\quad + \sup_{u_0 \leq u' \leq u} (\mathcal{O}(\delta, u') + \mathcal{R}^u(u')). \end{aligned}$$

Thanks to Gronwall's inequality, we have

$$\|\text{tr } \chi' - \text{tr } \chi'_{m_0}\|_{L^\infty(S_{\underline{u},u})} \lesssim \sup_{u_0 \leq u' \leq u} (\mathcal{O}(\delta, u') + \mathcal{R}^u(u')).$$

In a similar manner, we have the following bound for $\text{tr } \underline{\chi}'$:

$$\|\text{tr } \underline{\chi}' - \text{tr } \underline{\chi}'_{m_0}\|_{L^\infty(S_{\underline{u},u})} \lesssim \mathcal{O}(\underline{u}, u_0) + \mathcal{R}^u(\underline{u}) + \sup_{u_0 \leq u' \leq u} (\mathcal{O}(\delta, u') + \mathcal{R}^u(u')).$$

We turn to the derivatives of $\text{tr } \chi'$ and $\text{tr } \underline{\chi}'$. We apply \not{d} to (2.2) to derive

$$D\not{d} \text{tr } \chi' = -\text{tr } \chi \not{d}(\Omega \text{tr } \chi') - \not{d}(|\widehat{\chi}|^2) = -\Omega \text{tr } \chi \not{d} \text{tr } \chi' - \text{tr } \chi \text{tr } \chi' \not{d}\Omega - \not{d}(|\widehat{\chi}|^2).$$

The last two terms can be bounded by $\mathcal{P}(\underline{u}, u)$, and the first term can be absorbed by Gronwall's inequality. Thus, we obtain

$$\|\not{d} \operatorname{tr} \chi'\|_{L^\infty(\underline{u}, u)} \lesssim \mathcal{P}(\underline{u}, u).$$

For higher order derivatives, we have

$$D\nabla^k \operatorname{tr} \chi' = -\Omega \operatorname{tr} \chi \nabla^k \operatorname{tr} \chi' + \text{l.o.t.},$$

where l.o.t. can be bounded directly by $\mathcal{P}(\underline{u}, u)$. Thanks to Gronwall's inequality, we obtain

$$\|\nabla^2 \operatorname{tr} \chi'\|_{L^4(\underline{u}, u)} + \|\nabla^3 \operatorname{tr} \chi'\|_{L^2(\underline{u}, u)} \lesssim \mathcal{P}(\underline{u}, u).$$

Similarly, we have the following estimates for $\operatorname{tr} \underline{\chi}'$:

$$\|\not{d} \operatorname{tr} \underline{\chi}'\|_{L^\infty(\underline{u}, u)} + \|\nabla^2 \operatorname{tr} \underline{\chi}'\|_{L^4(\underline{u}, u)} + \|\nabla^3 \operatorname{tr} \underline{\chi}'\|_{L^2(\underline{u}, u)} \lesssim \mathcal{P}(\underline{u}, u).$$

Finally, using the bound of $\nabla^i(\Omega - \Omega_{m_0})$ in terms of $\mathcal{P}(\underline{u}, u)$ for $i = 0, 1, 2, 3$, the above estimates for $\nabla^i(\operatorname{tr} \chi' - \operatorname{tr} \chi'_{m_0})$ and $\nabla^i(\operatorname{tr} \underline{\chi}' - \operatorname{tr} \underline{\chi}'_{m_0})$ also hold for $\nabla^i(\operatorname{tr} \chi - \operatorname{tr} \chi_{m_0})$ and $\nabla^i(\operatorname{tr} \underline{\chi} - \operatorname{tr} \underline{\chi}_{m_0})$, or $\nabla^i(\Omega \operatorname{tr} \chi - \Omega_{m_0} \operatorname{tr} \chi_{m_0})$ and $\nabla^i(\Omega \operatorname{tr} \underline{\chi} - \Omega_{m_0} \operatorname{tr} \underline{\chi}_{m_0})$ for $i = 0, 1, 2, 3$.

Therefore, we have completed the proof of Lemma 5. \square

So far, we have obtained the bounds for connection coefficients in terms of initial data and the curvature components. In what follows, we will use the null Bianchi equations to bound the curvature components. This is similar to what has been done in Section 3.4.

We rewrite (2.24) and (2.25) in terms of the renormalized quantities such as $\rho - \rho_{m_0}$. Recall that ρ_{m_0} satisfies the following equations:

$$D\rho_{m_0} + \frac{3}{2}\Omega_{m_0} \operatorname{tr} \chi_{m_0} \rho_{m_0} = 0, \quad \underline{D}\rho_{m_0} + \frac{3}{2}\Omega_{m_0} \operatorname{tr} \underline{\chi}_{m_0} \rho_{m_0} = 0.$$

We can subtract these two equations from (2.24) and (2.25) to derive

$$(4.12) \quad D(\rho - \rho_{m_0}) + \frac{3}{2}(\Omega \operatorname{tr} \chi - \Omega_{m_0} \operatorname{tr} \chi_{m_0})\rho + \frac{3}{2}\Omega_{m_0} \operatorname{tr} \chi_{m_0}(\rho - \rho_{m_0}) \\ - \Omega \left\{ \operatorname{div} \beta + (2\underline{\eta} + \zeta, \beta) - \frac{1}{2}(\widehat{\chi}, \alpha) \right\} = 0,$$

$$(4.13) \quad \underline{D}(\rho - \rho_{m_0}) + \frac{3}{2}(\Omega \operatorname{tr} \underline{\chi} - \Omega_{m_0} \operatorname{tr} \underline{\chi}_{m_0})\rho + \frac{3}{2}\Omega_{m_0} \operatorname{tr} \underline{\chi}_{m_0}(\rho - \rho_{m_0}) \\ + \Omega \left\{ \operatorname{div} \underline{\beta} + (2\eta - \zeta, \underline{\beta}) - \frac{1}{2}(\widehat{\chi}, \underline{\alpha}) \right\} = 0.$$

Because ρ_{m_0} is constant on each $S_{\underline{u},u}$, we can also rewrite (2.22) and (2.23) as

$$(4.14) \quad \underline{D}\beta + \frac{1}{2}\Omega \operatorname{tr} \underline{\chi}\beta - \Omega \widehat{\underline{\chi}} \cdot \beta + \underline{\omega}\beta \\ - \Omega \left\{ \underline{d}(\rho - \rho_{m_0}) + {}^* \underline{d}\sigma + 3\underline{\eta}\rho + 3{}^* \underline{\eta}\sigma + 2\widehat{\underline{\chi}} \cdot \underline{\beta} \right\} = 0,$$

$$(4.15) \quad D\beta + \frac{1}{2}\Omega \operatorname{tr} \chi\beta - \Omega \widehat{\chi} \cdot \beta + \omega\beta \\ + \Omega \left\{ \underline{d}(\rho - \rho_{m_0}) - {}^* \underline{d}\sigma + 3\underline{\eta}\rho - 3{}^* \underline{\eta}\sigma - 2\widehat{\underline{\chi}} \cdot \underline{\beta} \right\} = 0.$$

We call (4.12), (4.13), (4.14) and (4.15) together with the following six original Bianchi equations (2.18), (2.19), (2.20), (2.21), (2.26) and (2.27) *the renormalized Bianchi equations*. In a similar manner as we derived energy inequalities (3.44), (3.45), (3.46) and (3.47), we have the following energy inequalities:

$$\begin{aligned} & \sum_{R \in \{\alpha, \beta, \rho - \rho_{m_0}, \sigma, \beta\}} \int_{C_{\underline{u}}^u} |R|^2 + \sum_{\underline{R} \in \{\beta, \rho - \rho_{m_0}, \sigma, \beta, \alpha\}} \int_{\underline{C}_{\underline{u}}^u} |\underline{R}|^2 \\ \lesssim & \sum_{R \in \{\alpha, \beta, \rho - \rho_{m_0}, \sigma, \beta\}} \int_{C_{u_0}^u} |R|^2 + \sum_{\underline{R} \in \{\beta, \rho - \rho_{m_0}, \sigma, \beta, \alpha\}} \int_{\underline{C}_{\delta}^u} |\underline{R}|^2 \\ & + \sum_{\substack{R \in \{\alpha, \beta, \rho - \rho_{m_0}, \sigma, \beta\} \\ \underline{R} \in \{\beta, \rho - \rho_{m_0}, \sigma, \beta, \alpha\} \\ \mathcal{R}_1, \mathcal{R}_2 \in \{\alpha, \beta, \rho - \rho_{m_0}, \sigma, \beta, \alpha\}}} \iint_{M_{\underline{u},u}} |\Omega \Gamma \cdot \mathcal{R}_1 \cdot \mathcal{R}_2| \\ & + |\rho| |(\Omega \operatorname{tr} \chi - \Omega_{m_0} \operatorname{tr} \chi_{m_0})(\rho - \rho_{m_0})| \\ & + |\widehat{\chi} \cdot \alpha + \eta \cdot \beta + \widehat{\underline{\chi}} \cdot \underline{\alpha} + \underline{\eta} \cdot \underline{\beta}|, \end{aligned}$$

where $M_{\underline{u},u}$ refers to the region corresponding to $\delta \leq \underline{u}' \leq \underline{u}$, $u_0 \leq u' \leq u$. We also have

$$\begin{aligned} & \sum_{R \in \{\alpha, \beta, \rho - \rho_{m_0}, \sigma, \beta\}} \int_{C_{\underline{u}}^u} |\nabla^i R|^2 + \sum_{\underline{R} \in \{\beta, \rho - \rho_{m_0}, \sigma, \beta, \alpha\}} \int_{\underline{C}_{\underline{u}}^u} |\nabla^i \underline{R}|^2 \\ \lesssim & \sum_{R \in \{\alpha, \beta, \rho - \rho_{m_0}, \sigma, \beta\}} \int_{C_{u_0}^u} |\nabla^i R|^2 + \sum_{\underline{R} \in \{\beta, \rho - \rho_{m_0}, \sigma, \beta, \alpha\}} \int_{\underline{C}_{\delta}^u} |\nabla^i \underline{R}|^2 \\ & + \sum_{\substack{R \in \{\alpha, \beta, \rho - \rho_{m_0}, \sigma, \beta\} \\ \underline{R} \in \{\beta, \rho - \rho_{m_0}, \sigma, \beta, \alpha\} \\ \mathcal{R}_1, \mathcal{R}_2 \in \{\alpha, \beta, \rho - \rho_{m_0}, \sigma, \beta, \alpha\}}} \iint_{M_{\underline{u},u}} \sum_{j=0}^{i-1} |\nabla^j \nabla(\Omega \chi) \cdot \nabla^{i-1-j} \underline{R} \cdot \nabla^i \underline{R}| \\ & + \sum_{j=0}^{i-1} |\nabla^j \nabla(\Omega \underline{\chi}) \cdot \nabla^{i-1-j} R \cdot \nabla^i R| \\ & + \sum_{j=0}^{i-1} |\nabla^j K \cdot \nabla^{i-1-j} R \cdot \nabla^i R + \nabla^j K \cdot \nabla^{i-1-j} \underline{R} \cdot \nabla^i \underline{R}| \end{aligned}$$

$$\begin{aligned}
& + |\nabla^i((\Omega\Gamma) \cdot \mathcal{R}_1) \cdot \nabla^i \mathcal{R}_2| + |\rho| |\nabla^i(\Omega \operatorname{tr} \chi - \Omega_{m_0} \operatorname{tr} \chi_{m_0}) \cdot \nabla^i \rho| \\
& + |\rho| |\nabla^i \widehat{\chi} \cdot \nabla^i \alpha + \nabla^i \eta \cdot \nabla^i \beta + \nabla^i \underline{\widehat{\chi}} \cdot \nabla^i \underline{\alpha} + \nabla^i \underline{\eta} \cdot \nabla^i \underline{\beta}|
\end{aligned}$$

for $i = 1, 2, 3$. We remark that in terms $\mathcal{R}_1 \cdot \mathcal{R}_2$, the term $\alpha \cdot \underline{\alpha}$ do not appear. Thus, we can regard $\mathcal{R}_1 \cdot \mathcal{R}_2$ as either $R_1 \cdot R_2$ or $\underline{R}_1 \cdot \underline{R}_2$. We then have the following estimates:

$$\begin{aligned}
\sum_{j=0}^1 \iint_{M_{\underline{u}, u}} |\nabla^j(\Omega\Gamma) \cdot \nabla^{i-j} R_1|^2 & \lesssim \int_{u_0}^u \int_{C_{u'}^{\underline{u}}} \sum_{j=0}^i |\nabla^j R_1|^2 du' \quad \text{for } i \leq 3, \\
\iint_{M_{\underline{u}, u}} |\nabla^2(\Omega\Gamma) \cdot \nabla^i R_1|^2 & \lesssim \int_{u_0}^u \int_{\delta}^{\underline{u}} \|\nabla^i R_1\|_{L^4(\underline{u}', u')}^2 d\underline{u}' du' \\
& \lesssim \int_{u_0}^u \int_{\delta}^{\underline{u}} \sum_{j=0}^{i+1} \|\nabla^j R_1\|_{L^2(\underline{u}', u')}^2 d\underline{u}' du' \\
& \lesssim \int_{u_0}^u \int_{C_{u'}^{\underline{u}}} \sum_{j=0}^{i+1} |\nabla^j R_1|^2 du' \quad \text{for } i \leq 1, \\
\iint_{M_{\underline{u}, u}} |\nabla^3(\Omega\Gamma) \cdot R_1|^2 & \lesssim \int_{u_0}^u \int_{\delta}^{\underline{u}} \|R_1\|_{L^\infty(\underline{u}', u')}^2 d\underline{u}' du' \\
& \lesssim \int_{u_0}^u \int_{\delta}^{\underline{u}} \sum_{j=0}^2 \|\nabla^j R_1\|_{L^2(\underline{u}', u')}^2 d\underline{u}' du' \\
& \lesssim \int_{u_0}^u \int_{C_{u'}^{\underline{u}}} \sum_{j=0}^2 |\nabla^j R_1|^2 du'.
\end{aligned}$$

Therefore, by Cauchy-Schwartz inequality, we have

$$\sum_{i=0}^3 \iint_{M_{\underline{u}, u}} |\nabla^i(\Omega\Gamma \cdot R_1) \cdot \nabla^i R_2| \lesssim \sum_{i=0}^3 \int_{u_0}^u \int_{C_{u'}^{\underline{u}}} (|\nabla^i R_1|^2 + |\nabla^i R_2|^2) du'.$$

Similarly, we have

$$\sum_{i=0}^3 \iint_{M_{\underline{u}, u}} |\nabla^i(\Omega\Gamma \cdot \underline{R}_1) \cdot \nabla^i \underline{R}_2| \lesssim \sum_{i=0}^3 \int_{\delta}^{\underline{u}} \int_{C_{\underline{u}'}} (|\nabla^i \underline{R}_1|^2 + |\nabla^i \underline{R}_2|^2) d\underline{u}'.$$

We turn to the following terms

$$\begin{aligned}
\sum_{i=0}^2 \iint_{M_{\underline{u}, u}} |\nabla^i(\nabla(\Omega\chi) \cdot \underline{R}) \cdot \nabla \underline{R} + \nabla^i(\nabla(\Omega\underline{\chi}) \cdot R) \cdot \nabla R| \\
+ |\nabla^i(KR \cdot \nabla R + KR \cdot \underline{R})|.
\end{aligned}$$

They are treated in the same way as above by virtue of the bounds on K . They are eventually bounded by

$$\sum_{i=0}^3 \left(\int_{u_0}^u \int_{C_{u'}^u} |\nabla^i R|^2 du' + \int_{\delta}^u \int_{\underline{C}_{\underline{u}'}} |\nabla^i \underline{R}|^2 d\underline{u}' \right).$$

Finally, we consider the following terms:

$$\begin{aligned} & \sum_{i=0}^3 \iint_{M_{\underline{u}, u}} |\rho| |\nabla^i (\Omega \operatorname{tr} \chi - \Omega_{m_0} \operatorname{tr} \chi_{m_0}) \cdot \nabla^i (\rho - \rho_{m_0}) \\ & \quad + \nabla^i \widehat{\chi} \cdot \nabla^i \alpha + \nabla^i \eta \cdot \nabla^i \beta + \nabla^i \underline{\widehat{\chi}} \cdot \nabla^i \underline{\alpha} + \nabla^i \underline{\eta} \cdot \nabla^i \underline{\beta} |. \end{aligned}$$

According to Lemma 5, they are bounded by

$$\begin{aligned} & \sup_{\delta \leq u \leq \delta + \epsilon_0} \mathcal{O}(\underline{u}, u_0)^2 + \sup_{u_0 \leq u \leq -1 - u} \mathcal{O}(\delta, u)^2 \\ & + \sum_{i=0}^3 \left(\int_{u_0}^u \sup_{u_0 \leq u'' \leq u'} \int_{C_{u''}^u} |\nabla^i R|^2 du' + \int_{\delta}^u \sup_{\delta \leq u'' \leq u'} \int_{\underline{C}_{\underline{u}''}} |\nabla^i \underline{R}|^2 d\underline{u}' \right). \end{aligned}$$

We define

$$\begin{aligned} \mathcal{E}(u) &= \sum_{R \in \{\alpha, \beta, \rho - \rho_{m_0}, \sigma, \underline{\beta}\}} \sup_{u_0 \leq u' \leq u} \sum_{i=0}^3 \int_{C_{u'}} |\nabla^i R|^2, \\ \mathcal{F}(\underline{u}) &= \sum_{\underline{R} \in \{\beta, \rho - \rho_{m_0}, \sigma, \underline{\beta}, \underline{\alpha}\}} \sup_{\delta \leq \underline{u}' \leq \underline{u}} \sum_{i=0}^3 \int_{\underline{C}_{\underline{u}'}} |\nabla^i \underline{R}|^2. \end{aligned}$$

Therefore, the above estimates can be summarized as

$$\begin{aligned} \mathcal{E}(u) + \mathcal{F}(\underline{u}) &\lesssim \mathcal{E}(u_0) + \mathcal{F}(\delta) \\ & \quad + \sup_{\delta \leq u \leq \delta + \epsilon_0} \mathcal{O}(\underline{u}, u_0)^2 + \sup_{u_0 \leq u \leq -1 - u} \mathcal{O}(\delta, u)^2 \\ & \quad + \left(\int_{u_0}^u \mathcal{E}(u') du' + \int_{\delta}^u \mathcal{F}(\underline{u}') d\underline{u}' \right). \end{aligned}$$

Thanks to Gronwall's inequality, we have proved

PROPOSITION 3. *If we have the following smallness on initial data*

$$\mathcal{R}^{1+\delta}(u_0), \mathcal{R}^{-1-\delta}(\delta), \mathcal{O}(\underline{u}, u_0), \mathcal{O}(\delta, u) \leq \epsilon,$$

then for $(\underline{u}, u) \in [\delta, \delta + \epsilon_0] \times [u_0, u_0 + \epsilon_0]$, we have

$$\mathcal{R}^u(u), \mathcal{R}^u(\underline{u}), \mathcal{O}(\underline{u}, u), \mathcal{O}(\underline{u}, u) \lesssim \epsilon.$$

The smallness of the data is of course guaranteed by Proposition 1.

Equipped with Proposition 3, we are able to control C^k norms. More precisely, we shall prove that

$$(4.16) \quad \|g - g_{m_0}\|_{C^2(M_{\epsilon_0, g_{m_0}})} \lesssim \epsilon.$$

We start with the C^0 norms. We write g in canonical double null coordinates as follows:

$$g = -2\Omega^2(\underline{d}\underline{u} \otimes du + du \otimes \underline{d}\underline{u}) + \not{g}_{AB}(d\theta^A - b^A \underline{d}\underline{u}) \otimes (d\theta^B - b^B \underline{d}\underline{u}).$$

Since $\underline{D}(\Omega^2) = 2\Omega^2 \underline{\omega}$, $\underline{D}\not{g} = 2\Omega \underline{\chi}$ and $\underline{D}b = 4\Omega^2 \zeta$, by virtue of the L^∞ bounds of $\underline{\omega} - \underline{\omega}_{m_0}$, $\underline{\chi} - \underline{\chi}_{m_0}$ and ζ in Proposition 3 as well as the Sobolev inequalities, we have

$$\|g - g_{m_0}\|_{C^0(M_{\epsilon_0, g_{m_0}})} \lesssim \epsilon.$$

For C^1 norms of $g - g_{m_0}$, we use $\nabla_{m_0}(g - g_{m_0}) = (\nabla - \nabla_{m_0})g$. We also need the L^∞ bounds on $\nabla - \nabla_{m_0}$, which can be obtained by $\Gamma - \Gamma_{m_0}$ and $\not{\Gamma} - \not{\Gamma}_{m_0}$, where Γ refers to the null connection coefficients and $\not{\Gamma}$ refers to the Christoffel symbols of \not{g} . We have already derived

$$|\Gamma - \Gamma_{m_0}| \lesssim \epsilon$$

thanks to Proposition 3 and the Sobolev inequalities. To estimate $\not{\Gamma} - \not{\Gamma}_{m_0}$, we need the following propagation equation:

$$\underline{D}(\not{\Gamma} - \not{\Gamma}_{m_0})_{AB}^C = \frac{1}{2} \not{g}^{CD} (\not{\nabla}_A(\Omega \underline{\chi})_{BD} + \not{\nabla}_B(\Omega \underline{\chi})_{AD} - \not{\nabla}_D(\Omega \underline{\chi})_{AB}).$$

Since we have already obtained the L^∞ bounds on $\not{\nabla}(\Omega \underline{\chi})$, by directly integrating the above equation, we obtain $|\not{\Gamma} - \not{\Gamma}_{m_0}| \lesssim \epsilon$. Finally, we have

$$\|g - g_{m_0}\|_{C^1(M_{\epsilon_0, g_{m_0}})} \lesssim \epsilon.$$

For C^2 bounds on $g - g_{m_0}$, we write

$$\nabla_{m_0}^2(g - g_{m_0}) = (\nabla^2 - \nabla_{m_0}^2)g = \nabla((\nabla - \nabla_{m_0})g) + (\nabla - \nabla_{m_0})\nabla_{m_0}g.$$

The last term $(\nabla - \nabla_{m_0})\nabla_{m_0}g$ has already been controlled. Thus, we need the L^∞ bounds of the following quantities:

$$\begin{aligned} &D(\Gamma - \Gamma_{m_0}), \quad D(\not{\Gamma} - \not{\Gamma}_{m_0}), \quad \underline{D}(\Gamma - \Gamma_{m_0}), \\ &\underline{D}(\not{\Gamma} - \not{\Gamma}_{m_0}), \quad \not{\nabla}(\Gamma - \Gamma_{m_0}), \quad \not{\nabla}(\not{\Gamma} - \not{\Gamma}_{m_0}). \end{aligned}$$

The estimates for the first four quantities can be obtained using the null structure equations (2.1)–(2.8) and (2.14)–(2.17), and provided we have L^∞ bounds of all first derivatives of null curvature components, but we do not have control on

$$D(\omega - \omega_{m_0}), \quad \underline{D}(\underline{\omega} - \underline{\omega}_{m_0}).$$

The estimates for the fifth quantity are obtained directly by Proposition 3. The bounds for the last quantity are obtained also by the propagation equation of $\Gamma - \Gamma_{m_0}$ provided we have an L^∞ bound of

$$\nabla^2(\Omega\underline{\chi}).$$

The strategy is clear now: we have to control

$$\|(D(\omega - \omega_{m_0}))\|_{L^\infty}, \quad \|\underline{D}(\underline{\omega} - \underline{\omega}_{m_0})\|_{L^\infty} \text{ and } \|\nabla^2(\Omega\underline{\chi})\|_{L^\infty}.$$

For this purpose, we first claim that

$$\|\alpha, \beta, \rho - \rho_{m_0}, \sigma, \underline{\beta}, \underline{\alpha}\|_{L^\infty} \lesssim \epsilon.$$

This can be obtained by using a variation of (3.4) and (3.5) by setting the constants depending on m_0, ϵ_0, u_0 and the fact that $\|\nabla^{\leq 1}\alpha\|_{L^4(S_{\delta,u})} \lesssim \epsilon$, which is stated in Proposition 1. To apply (3.4) and (3.5), besides $\mathcal{R}_2, \underline{\mathcal{R}}_2 \lesssim \epsilon$, we also need $\|\nabla D\alpha\|_{L^2(C_u)} \lesssim \epsilon$ and $\|\nabla \underline{D}(\beta, \rho - \rho_{m_0}, \sigma, \underline{\beta}, \underline{\alpha})\|_{L^2(\underline{C}_u)} \lesssim \epsilon$. For $\beta, \rho - \rho_{m_0}, \sigma$ and $\underline{\beta}$ this is done by simply taking ∇ to both sides of the renormalized Bianchi equations.⁹ Estimates on $\|\nabla D\alpha\|_{L^2(C_u)}$ and $\|\nabla \underline{D}\alpha\|_{L^2(\underline{C}_u)}$ are obtained using a variation of (3.48) and (3.49) by replacing Lie derivatives by ∇ .

We then claim

$$\|\nabla(\beta, \underline{\beta})\|_{L^\infty} \lesssim \epsilon.$$

The proof is similar to the above; we first apply ∇^2 to (2.22) and (2.21) and then apply (3.5) and Sobolev inequalities on $S_{u,u}$. This yields $\|\nabla(\beta, \underline{\beta})\|_{L^\infty} \lesssim \epsilon$. As a consequence, we have

$$\|D(\omega - \omega_{m_0}), \underline{D}(\underline{\omega} - \underline{\omega}_{m_0})\|_{L^\infty} \lesssim \epsilon.$$

In fact, we can apply D to (2.8) to derive a propagation equation for $D(\omega - \omega_{m_0})$ that looks like

$$\underline{D}D(\omega - \omega_{m_0}) = D(\rho - \rho_{m_0}) + \text{l.o.t.} = \nabla\beta + \text{l.o.t.}$$

Here, l.o.t. denotes the terms without derivatives acting on connection coefficients or curvature components in (4.12) and (4.13). We then integrate this equation to conclude. Similarly, we have estimates for $\underline{D}(\underline{\omega} - \underline{\omega}_{m_0})$.

To prove $\|\nabla^2(\Omega\underline{\chi})\|_{L^\infty} \lesssim \epsilon$, using bounds of $\|\nabla^3\underline{\beta}\|_{L^2(\underline{C}_u)}$ given by Proposition 3, we couple (2.11) with (2.4) to derive bounds of $\|\nabla^4\underline{\chi}\|_{L^2(\underline{C}_u)}$ and $\|\nabla^4 \text{tr} \underline{\chi}\|_{L^2(S_{u,u})}$. Therefore, we have proved

$$\|g - g_{m_0}\|_{C^2(M_{\epsilon_0, g_{m_0}})} \lesssim \epsilon.$$

⁹Recall that the renormalized Bianchi equations refer to (4.12), (4.13), (4.14), (4.15), (2.18), (2.19), (2.20), (2.21), (2.26) and (2.27). Here we take ∇ to both sides of (4.13), (4.14), (2.21), (2.27).

For higher order estimates on C^k norms, as we remarked before, we can simply use a routine induction argument and we omit the proof. Therefore, we have completed the proof of Theorem 2. \square

5. Gluing construction

5.1. *Preparation.* We summarize some key properties of the metric g constructed in the previous sections:

- (1) In the region $(\underline{u}, u) \in [u_0, 0] \times [u_0, 0]$, g coincides with the standard Minkowski metric.
- (2) In the region where $\underline{u} \in [0, \delta]$ and $u \in [u_0, -1 - \underline{u}]$, the metric g is constructed by Christodoulou in his work [2]. In addition, we have imposed condition (2.30) with $|u_0| > 2m_0$, so that on the incoming cone \underline{C}_δ , the estimates in Section 4.2 hold. For a sufficiently small δ , $\text{tr } \chi(\delta, u_0) > 0$ and hence S_{δ, u_0} is not a trapped surface, and if $2m_0 > 1$, according to Theorem 17.1 in [2], $S_{\delta, -1-\delta}$ is a trapped surface.

We remark that, for $0 < m_0 < \frac{1}{2}$, it is straightforward from the proof (Theorem 16.1 for the existence and Theorem 17.1 for the formation of trapped surfaces) of [2] that g actually exists up to $\underline{u} + u = -2m_0$ and $S_{\delta, -2m_0-\delta}$ is a trapped surface, once we choose a sufficiently small δ .

- (3) In the region $(\underline{u}, u) \in [\delta, \delta + \epsilon_0] \times [u_0, u_0 + \epsilon_0]$, the metric g is ϵ -close (in C^{k-3} norms) to the Schwarzschild metric g_{m_0} .

We fix a sphere $S_1 = S_{\delta, u_1}$ near S_{δ, u_0} by choosing u_1 close to u_0 in such a way that $\text{tr } \chi(\delta, u_1) > 0$ and $\text{tr } \chi_{m_0}(\delta, u_1) > 0$. We emphasize that the choice of S_1 does not depend on δ if δ is sufficiently small. Because g_{m_0} is the Schwarzschild metric with mass m_0 , we can choose the time function t of g_{m_0} in the Boyer-Lindquist coordinates (we regard the Schwarzschild metric as a member in the Kerr family) such that $t(\delta, u_1) = 0$. We also regard t as a smooth function in the \underline{u} - u plane $(\underline{u}, u) \in [\delta, \delta + \epsilon_0] \times [u_0, u_0 + \epsilon_0]$.

Now in the region $(\underline{u}, u) \in [\delta, \delta + \epsilon_0] \times [u_0, u_0 + \epsilon_0]$, we choose a three-dimensional hypersurface

$$H = \bigcup_{t(\underline{u}, u)=0} S_{\underline{u}, u},$$

and we use S_2 to denote the intersection of H with $C_{u_0} \cup \underline{C}_{\delta+\epsilon_0}$. We define the following two radii, $r_1 = r|_{S_1}$ and $r_2 = r|_{S_2}$, where $r(\underline{u}, u) = r_{m_0}(\underline{u}, u)$ is the radius of $S_{\underline{u}, u}$ in Schwarzschild metric g_{m_0} . By construction, $2m_0 < r_1 < r_2$ and H is space-like with respect to g_{m_0} . Thus, if ϵ is sufficiently small, H is also space-like with respect to the metric g .

Let $(\bar{g}_{m_0}, \bar{k}_{m_0})$ be the induced metric and the second fundamental form of H as a submanifold in the Schwarzschild space-time with mass m_0 . Therefore,

for $r_1 \leq r \leq r_2$, we have

$$\bar{g}_{m_0} = \left(1 - \frac{2m_0}{r}\right)^{-1} dr^2 + r^2 d\sigma_{S^2}^2$$

and

$$\bar{k}_{m_0} \equiv 0.$$

Let (\bar{g}, \bar{k}) be the metric and the second fundamental form of H induced by g ; since g is ϵ -close to the Schwarzschild metric g_{m_0} in C^{k-3} norms, we have

$$\|\bar{g} - \bar{g}_{m_0}\|_{C^{k-3}(H, \bar{g}_{m_0})} \lesssim \epsilon$$

and

$$\|\bar{k}\|_{C^{k-4}(H, \bar{g}_{m_0})} \lesssim \epsilon.$$

We recall the form of the Kerr metric in the Boyer-Lindquist coordinates

$$\begin{aligned} g_{m,(0,0,a)} = & \left(-1 + \frac{2mr}{r^2 + a^2 \cos^2 \theta}\right) dt^2 - \frac{2mra \sin^2 \theta}{r^2 + a^2 \cos^2 \theta} dt d\varphi + \frac{r^2 + a^2 \cos^2 \theta}{r^2 - 2mr + a^2} dr^2 \\ & + (r^2 + a^2 \cos^2 \theta) d\theta^2 + \sin^2 \theta \left(r^2 + a^2 + \frac{2mra^2 \sin^2 \theta}{r^2 + a^2 \cos^2 \theta}\right) d\varphi^2, \end{aligned}$$

where $m > a \geq 0$. The lower index $(0, 0, a)$ specifies the angular momentum vector: if we use the spherical polar coordinates $x^1 = r \cos \varphi \sin \theta$, $x^2 = r \sin \varphi \sin \theta$ and $x^3 = r \cos \theta$, then the vector $(0, 0, a)$ is the same as the axis (of the rotation of the Kerr black hole) and the norm of $(0, 0, a)$ is the angular momentum a . For an arbitrary vector $\mathbf{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$ with $|\mathbf{a}| = a$ and an isometry $\Omega_{\mathbf{a}} \in SO(3)$ mapping \mathbf{a} to $(0, 0, a)$, we can also define a Kerr metric $g_{m,\mathbf{a}} = (\text{id}_{\mathbb{R}} \times \Omega_{\mathbf{a}})^* g_{m,(0,0,a)}$, where $\text{id}_{\mathbb{R}}$ is the identity map of the t axis. We remark that this definition does not depend on the choice of $\Omega_{\mathbf{a}}$.

We choose the slice H to be the $t = 0$ slice in the above Kerr space-time. We use (x^1, x^2, x^3) and (r, θ, φ) as coordinates on H . Let $(\bar{g}_{m,(0,0,a)}, \bar{k}_{m,(0,0,a)})$ be the induced metric and the second fundamental form of H ; thus, we have

$$\begin{aligned} \bar{g}_{m,(0,0,a)} = & \frac{r^2 + a^2 \cos^2 \theta}{r^2 - 2mr + a^2} dr^2 + (r^2 + a^2 \cos^2 \theta) d\theta^2 \\ & + \left(r^2 + a^2 + \frac{2mra^2 \sin^2 \theta}{r^2 + a^2 \cos^2 \theta}\right) \sin^2 \theta d\varphi \end{aligned}$$

and

$$\begin{aligned} \bar{k}_{m,(0,0,a)} = & \frac{2}{r^2 + a^2 + \frac{2mra^2 \sin^2 \theta}{r^2 + a^2 \cos^2 \theta}} \left(\frac{ma}{r^2 + a^2 \cos^2 \theta} (r^2 - a^2) dr d\varphi \right. \\ & \left. + \frac{2m^2 r^2 a^3 \sin 2\theta \sin^2 \theta}{(r^2 + a^2 \cos^2 \theta)^2} d\theta d\varphi \right). \end{aligned}$$

For an arbitrary vector $\mathbf{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$ with $|\mathbf{a}| = a$, we can also define the corresponding initial data set by $\bar{g}_{m,\mathbf{a}} = \Omega_{\mathbf{a}}^* \bar{g}_{m,(0,0,a)}$ and $\bar{k}_{m,\mathbf{a}} = \Omega_{\mathbf{a}}^* \bar{k}_{m,(0,0,a)}$. They also correspond to the $t = 0$ slices.

We now recall the definition for the constraint map Φ , that is,

$$\Phi(\bar{g}, \bar{\pi}) = (\mathcal{H}(\bar{g}, \bar{\pi}), \operatorname{div} \bar{\pi}),$$

where $\bar{\pi} = \bar{k} - \operatorname{tr} \bar{k} \bar{g}$ and $\mathcal{H}(\bar{g}, \bar{\pi}) = R(\bar{g}) + \frac{1}{2}(\operatorname{tr} \bar{\pi})^2 - |\bar{\pi}|^2$. Let $D\Phi_{(\bar{g}, \bar{\pi})}^*$ be the formal L^2 -adjoint of the linearized operator of Φ at $(\bar{g}, \bar{\pi})$. When $(\bar{g}, \bar{\pi}) = (\bar{g}_{m_0}, 0)$, we use K to denote the null space $\operatorname{Ker} D\Phi_{(\bar{g}_{m_0}, 0)}^*$. We will explicitly write down K . In (x^1, x^2, x^3) coordinates, we use Ω_i ($i = 1, 2, 3$) to denote the following vectors fields:

$$\Omega_1 = x^2 \frac{\partial}{\partial x_3} - x^3 \frac{\partial}{\partial x_2}, \quad \Omega_2 = x^3 \frac{\partial}{\partial x_1} - x^1 \frac{\partial}{\partial x_3}, \quad \Omega_3 = x^1 \frac{\partial}{\partial x_2} - x^2 \frac{\partial}{\partial x_1},$$

and in (r, θ, φ) coordinates, they can be written as

$$(5.1) \quad \Omega_1 = -\sin \varphi \partial_\theta - \cos \varphi \cot \theta \partial_\varphi, \quad \Omega_2 = \cos \varphi \partial_\theta - \sin \varphi \cot \theta \partial_\varphi, \quad \Omega_3 = \partial_\varphi.$$

Then we have the following description of K :

LEMMA 6. *K is a four-dimensional real vector space. In fact, K is spanned by the following four elements: $((1 - \frac{2m_0}{r})^{\frac{1}{2}}, 0)$, $(0, \Omega_1)$, $(0, \Omega_2)$ and $(0, \Omega_3)$.*

Proof. We first recall the form of the linear operator $D\Phi_{(\bar{g}, \bar{\pi})}^*$ (see also Section 2.3). For a smooth function f and a vector field X , we have $D\Phi_{(\bar{g}, \bar{\pi})}^*(f, X) = D\mathcal{H}_{(\bar{g}, \bar{k})}^*(f) + D\operatorname{div}_{(\bar{g}, \bar{k})}^*(X)$, where

$$\begin{aligned} D\mathcal{H}_{(\bar{g}, \bar{\pi})}^*(f) &= ((L_{\bar{g}}^* f)_{ij} + (\operatorname{tr}_{\bar{g}} \bar{\pi} \cdot \bar{\pi}_{ij} - 2\bar{\pi}_{ik} \bar{\pi}^k_j) f, (\operatorname{tr}_{\bar{g}} \bar{\pi} \cdot \bar{g}_{ij} - 2\bar{\pi}_{ij}) f), \\ D\operatorname{div}_{(\bar{g}, \bar{\pi})}^*(X) &= \frac{1}{2}(\mathcal{L}_X \bar{\pi}_{ij} + \bar{\nabla}_k X^k \bar{\pi}_{ij} - (X_i(\bar{\nabla}_k \bar{\pi}^k_j + X_j \bar{\nabla}_k \bar{\pi}^k_i) \\ &\quad - (\bar{\nabla}_m X_k \bar{\pi}^{km} + X_k \bar{\nabla}_m \bar{\pi}^{mk}) g_{ij}, \quad -\mathcal{L}_X g_{ij}), \end{aligned}$$

and $L_{\bar{g}}^*(f) = -(\Delta_{\bar{g}} f) \bar{g} + \operatorname{Hess}_{\bar{g}}(f) - f \operatorname{Ric}_{\bar{g}}$. In the current case, $(\bar{g}, \bar{\pi}) = (g_{m_0}, 0)$, and thus,

$$(5.2) \quad D\Phi_{(\bar{g}_{m_0}, 0)}^*(f, X) = \left(L_{\bar{g}_{m_0}}^* f, -\frac{1}{2} \mathcal{L}_X \bar{g}_{m_0} \right).$$

First of all, we consider the second equation $\mathcal{L}_X \bar{g}_{m_0} = 0$ of (5.2), which amounts to saying that X is a Killing vector field on constant t slice for \bar{g}_{m_0} . Thus, X is spanned by Ω_1, Ω_2 and Ω_3 , i.e., the infinitesimal rotations of the Schwarzschild space-time. We remark that $m_0 \neq 0$ in this case, otherwise we may also have the spatial translations of the Minkowski space-time.

Secondly, we consider the first equation in (5.2):

$$L_{\bar{g}_{m_0}}^* f = -(\Delta_{\bar{g}_{m_0}} f) \bar{g}_{m_0} + \operatorname{Hess}_{\bar{g}_{m_0}}(f) - f \operatorname{Ric}_{\bar{g}_{m_0}} = 0.$$

Taking the trace, we obtain $\Delta_{\bar{g}_{m_0}} f = 0$. So the above equation is equivalent to

$$(5.3) \quad \text{Hess}_{\bar{g}_{m_0}}(f) = f \text{Ric}_{\bar{g}_{m_0}}.$$

We show that f depends only on r . We consider the $(\partial_r, \partial_\theta)$ -component of (5.3), that is, $\partial_r \partial_\theta f - \frac{1}{r} \partial_\theta f = 0$. Thus, $\partial_\theta f = \alpha r$ for some $\alpha \in \mathbb{R}$. We then turn to the (∂_r, ∂_r) -component of (5.3); it can be written as

$$\partial_r^2 f + \frac{m_0}{r^2} \left(1 - \frac{2m_0}{r}\right)^{-1} \partial_r f = \left(-\frac{2m_0}{r^3} \left(1 - \frac{2m_0}{r}\right)^{-1}\right) f.$$

We can take derivative ∂_θ on both sides and then substitute $\partial_\theta f = \alpha r$ to the above equation. Those operations yield $\frac{\alpha m_0}{r^2} = -\frac{2\alpha m_0}{r^2}$. Therefore, $\alpha = 0$ since $m_0 \neq 0$. We conclude that $\partial_\theta f = 0$. We then proceed in the same manner to show $\partial_\varphi f = 0$. Hence, f depends only on r . To obtain the exact form of f , we consider the $(\partial_\theta, \partial_\theta)$ -component of (5.3), that is,

$$\partial_\theta^2 f + r \left(1 - \frac{2m_0}{r}\right) \partial_r f = \frac{m_0}{r} f.$$

Since f is simply a function of r , we have $\frac{\partial_r f}{f} = \frac{m_0}{r^2} \left(1 - \frac{2m_0}{r}\right)^{-1}$. Therefore, $f(r) = \alpha \left(1 - \frac{2m_0}{r}\right)^{\frac{1}{2}}$ for some $\alpha \in \mathbb{R}$. It is straightforward to verify that $f(r) = \alpha \left(1 - \frac{2m_0}{r}\right)^{\frac{1}{2}}$ is indeed a solution of (5.3) for all $\alpha \in \mathbb{R}$. Hence, we complete the proof of the lemma. \square

5.2. *Gluing construction.* We choose a sufficiently large k in Theorem 1 so that we can use the Local Deformation Lemma in Section 2.3. To run the gluing construction, we take $\phi = \phi(r)$ a smooth cut-off function on the three slice H such that $\phi \equiv 1$ near S_1 and $\phi \equiv 0$ near S_2 . We define an approximation to the final initial data set as follows:

$$(\tilde{g}, \tilde{k}) = (\phi \bar{g} + (1 - \phi) \bar{g}_{m, \mathbf{a}}, \phi \bar{k} + (1 - \phi) \bar{k}_{m, \mathbf{a}}).$$

At this stage, (\tilde{g}, \tilde{k}) may not satisfy the constraint equations (1.1) and we will use the techniques introduced in Section 2.3 to deform (\tilde{g}, \tilde{k}) to be a solution of the constraint equations.

We choose two free parameters $m \in \mathbb{R}_{>0}$ and $\mathbf{a} \in \mathbb{R}^3$, so that

$$|m - m_0| + |\mathbf{a}| \leq C_0 \epsilon,$$

where C_0 is large positive constant to be determined later. The constant C_0 is independent of the choice of ϵ , m and \mathbf{a} .

By the construction, it is easy to see that

$$\|\tilde{g} - \bar{g}_{m_0}\|_{C_{\rho^{-1}}^{k-4, \alpha}} \lesssim \epsilon, \quad \|\tilde{k}\|_{C_{\rho^{-1}}^{k-5, \alpha}} \lesssim \epsilon$$

for some $0 < \alpha < 1$ and ρ is a weight function decays near $\partial H = S_1 \cup S_2$. (See Section 2.3 for definitions.) We remark that the above norms are computed

with respect to \bar{g}_{m_0} . Since both (\bar{g}, \bar{k}) and $(\bar{g}_{m,\mathbf{a}}, \bar{k}_{m,\mathbf{a}})$ solve the constraint equations, it is easy to see that

$$\|\Phi(\tilde{g}, \tilde{\pi})\|_{C_{\rho^{-1}}^{k-6,\alpha}} \lesssim \epsilon,$$

where $\tilde{\pi} = \tilde{g} - \text{tr}_{\tilde{g}} \tilde{k} \tilde{g}$.

We now follow the procedure in [4] and [5] to deform $(\tilde{g}, \tilde{\pi})$. Let ζ be a bump function on H that is compactly supported between S_1 and S_2 . We define ζK to be the following set vector space:

$$\zeta K = \{(\zeta f, \zeta X) \mid (f, X) \in K\}.$$

It is the obstruction space for the gluing for data close to the Schwarzschild data. We remark that if ζ is sufficiently close to constant 1 and if the data (g, π) is sufficiently close to those of the Schwarzschild data $(g_{m_0}, 0)$, then the restriction of the operator $D\Phi_{(g,\pi)}^*$ on the orthogonal complement of ζK is still injective. In this case, we can still apply the Local Deformation Theorem (Theorem 2 of [5] or Section 2.3) to deform the image into ζK instead of K . We also notice that, in this case, if we use $V = L^2(\mathcal{S}(H)) \times L^2(\mathcal{S}(H))$ to denote the pair of square integrable symmetric two tensors on H , then $V = K \oplus (\zeta K)^\perp$. After applying the Local Deformation Lemma, there will be no component in $(\zeta K)^\perp$. We also fix a weight function ρ that behaves as $\rho \sim d(\cdot, \partial H)^N$ for sufficiently large N near the boundary. According to the Local Deformation Theorem, if ϵ is sufficiently small, there exists a pair $(h, \omega) \in \mathcal{S}^{k-4,\alpha}(H) \times \mathcal{S}^{k-5,\alpha}(H)$ such that $\Phi(\tilde{g} + h, \tilde{\pi} + \omega) \in \zeta K$, $(h, \omega) = 0$ near $\partial H = S_1 \cup S_2$. Moreover, (h, ω) satisfies the following estimates:

$$\|(h, \omega)\|_{C_{\rho^{-1}}^{k-5,\alpha}} \lesssim \epsilon.$$

We also remark that, instead of the polynomial weight ρ defined above, we can also use an exponential weight ρ , namely, $\rho \sim e^{-\frac{1}{d}}$ near the boundary, to solve (h, ω) in the smooth class instead of Hölder classes (because $\Phi(\tilde{g}, \tilde{\pi})$ is strictly supported in the interior of H). One can refer to [4] and [5] for more details.

We are going to argue that we can choose suitable parameters (m, \mathbf{a}) so that $\Phi(\tilde{g} + h, \tilde{\pi} + \omega) = 0$. According to the above discussion, in order to show this, it suffices to show the projection of $\Phi(\tilde{g} + h, \tilde{\pi} + \omega)$ to K vanishes; i.e., the following projection map \mathcal{I} hits zero:

$$\mathcal{I} : B_{C_0\epsilon} \rightarrow \mathbb{R}^4, \quad (m, \mathbf{a}) \mapsto (\mathcal{I}_0, \mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3),$$

where $B_{C_0\epsilon} = \{(m, \mathbf{a}) \in \mathbb{R}^4 : |m - m_0| + |\mathbf{a}| \leq C_0\epsilon\}$,

$$\mathcal{I}_0(m, \mathbf{a}) = \int_H \left(1 - \frac{2m_0}{r}\right)^{\frac{1}{2}} \mathcal{H}(\tilde{g} + h, \tilde{\pi} + \omega) d\mu_{\bar{g}_{m_0}},$$

and

$$\mathcal{I}_i(m, \mathbf{a}) = \int_H \operatorname{div}_{\tilde{g}+h}(\tilde{\pi} + \omega)_j \Omega_i^j d\mu_{\tilde{g}_{m_0}}$$

for $i = 1, 2, 3$.

We first analyze \mathcal{I}_0 . For $k \geq 7$, we recall that

$$\mathcal{H} : \mathcal{M}^{k-5,\alpha}(H) \times \mathcal{S}^{k-5,\alpha}(H) \rightarrow C^{k-7,\alpha}(H).$$

We consider the Taylor expansion of \mathcal{H} at the point $(\tilde{g}, \tilde{\pi}) \in \mathcal{M}^{k-5,\alpha}(H) \times \mathcal{S}^{k-5,\alpha}(H)$; thus, we have

$$\mathcal{I}_0(m, \mathbf{a}) = \int_H \left(1 - \frac{2m_0}{r}\right)^{\frac{1}{2}} (\mathcal{H}(\tilde{g}, \tilde{\pi}) + D\mathcal{H}_{(\tilde{g}, \tilde{\pi})}(h, \omega) + O(\|(h, \omega)\|_{C^{k-5,\alpha}}^2)) d\mu_{\tilde{g}_{m_0}}.$$

We observe that the second and the last terms in the above integral are of size $O(\epsilon^2)$ with respect to our small parameter ϵ . This is obvious for the last term. For the second term, recall that $(1 - \frac{2m_0}{r})^{\frac{1}{2}} \in \operatorname{Ker} D\mathcal{H}_{(\tilde{g}_{m_0}, 0)}^*$ and (h, ω) vanishes near ∂H ; therefore,

$$\int_H \left(1 - \frac{2m_0}{r}\right)^{\frac{1}{2}} D\mathcal{H}_{(\tilde{g}_{m_0}, 0)}(h, \omega) d\mu_{\tilde{g}_{m_0}} = 0.$$

We then subtract the above quantity from the second term to retrieve one more ϵ as follows:

$$\begin{aligned} & \int_H \left(1 - \frac{2m_0}{r}\right)^{\frac{1}{2}} D\mathcal{H}_{(\tilde{g}, \tilde{\pi})}(h, \omega) d\mu_{\tilde{g}_{m_0}} \\ &= \int_H \left(1 - \frac{2m_0}{r}\right)^{\frac{1}{2}} (D\mathcal{H}_{(\tilde{g}, \tilde{\pi})} - D\mathcal{H}_{(\tilde{g}_{m_0}, 0)})(h, \omega) d\mu_{\tilde{g}_{m_0}} \\ &\lesssim \|(\tilde{g} - \tilde{g}_{m_0}, \tilde{\pi})\|_{C^2} \|(h, \omega)\|_{C^2} = O(\epsilon^2). \end{aligned}$$

Finally, we consider the first term in \mathcal{I}_0 . Because $|\mathbf{a}| \leq C_0\epsilon$, we will make use of the following key observation:

$$\bar{g}_{m, \mathbf{a}} = \left(1 - \frac{2m}{r}\right)^{-1} dr^2 + r^2 d\sigma_{S^2}^2 + O(\epsilon^2).$$

This is obvious according to the formula for $\bar{g}_{m, \mathbf{a}}$ in last section. We also write

$$\bar{g} = \left(1 - \frac{2m_0}{r}\right)^{-1} dr^2 + r^2 d\sigma_{S^2}^2 + h_\epsilon,$$

where h_ϵ is a two tensor and $|h_\epsilon| \lesssim \epsilon$. Therefore, we have

$$\begin{aligned} (5.4) \quad & \int_H \left(1 - \frac{2m_0}{r}\right)^{\frac{1}{2}} \mathcal{H}(\tilde{g}, \tilde{\pi}) d\mu_{\tilde{g}_{m_0}} = \int_H R(f(r) dr^2 + r^2 d\sigma_{S^2}^2) d\mu_{\tilde{g}_0} \\ & + \int_H \left(1 - \frac{2m_0}{r}\right)^{\frac{1}{2}} D\mathcal{H}_{(\tilde{g}_{m_0}, 0)}(\phi h_\epsilon, 0) d\mu_{\tilde{g}_{m_0}} + O(\epsilon^2), \end{aligned}$$

where $f(r) = \phi(1 - \frac{2m_0}{r})^{-1} + (1 - \phi)(1 - \frac{2m}{r})^{-1}$. To compute the first term on the right-hand side, we have to compute the scalar curvature $R(f(r)dr^2 + r^2d\sigma^2)$. By direct computation,

$$\begin{aligned} R(f(r)dr^2 + r^2d\sigma^2) &= 2r^{-2}(rf^{-2}(r)f'(r) - f^{-1}(r) + 1) \\ &= -2r^{-2}\partial_r(r(f^{-1}(r) - 1)). \end{aligned}$$

By virtue of this formula, we can easily derive

$$\int_H R(f(r)dr^2 + r^2d\sigma^2)d\mu_{\bar{g}_0} = -8\pi(r(f^{-1}(r) - 1))|_{r_1}^{r_2} = 16\pi(m - m_0).$$

For the second term on the right-hand side of (5.4), we denote

$$\tilde{\epsilon}_0 \triangleq - \int_H \left(1 - \frac{2m_0}{r}\right)^{\frac{1}{2}} D\mathcal{H}_{(\bar{g}_{m_0}, 0)}(\phi h_\epsilon, 0)d\mu_{\bar{g}_{m_0}}.$$

Recall that $(1 - \frac{2m_0}{r})^{\frac{1}{2}} \in \text{Ker}D\mathcal{H}_{(\bar{g}_{m_0}, 0)}^*$, $\phi = 1$ near S_1 and $\phi = 0$ near S_2 . Therefore, $\tilde{\epsilon}_0$ can be written as an integral on S_1 , which only depends on h_ϵ on S_1 , and $|\tilde{\epsilon}_0| \lesssim \epsilon$. Finally, we summarize the above calculations for \mathcal{I}_0 as

$$(5.5) \quad \mathcal{I}_0(m, \mathbf{a}) = 16\pi(m - m_0) - \tilde{\epsilon}_0 + O(\epsilon^2).$$

For each $i = 1, 2, 3$, we analyze \mathcal{I}_i as follows:

$$\begin{aligned} \mathcal{I}_i(m, \mathbf{a}) &= \int_H \text{div}_{\bar{g}_{m_0}}(\tilde{\pi} + \omega)_j \Omega_i^j d\mu_{\bar{g}_{m_0}} + \int_H (\text{div}_{\bar{g}+h} - \text{div}_{\bar{g}_{m_0}})(\tilde{\pi} + \omega)_j \Omega_i^j d\mu_{\bar{g}_{m_0}} \\ &= \int_H \text{div}_{\bar{g}_{m_0}}(\tilde{\pi} + \omega)_j \Omega_i^j d\mu_{\bar{g}_{m_0}} + O(\epsilon^2). \end{aligned}$$

Recall that $\text{div}_{\bar{g}_{m_0}}^*(\Omega_i) = 0$, $(\tilde{g}, \tilde{\pi})$ coincides with $(\bar{g}, \bar{\pi})$ on S_1 and $(\tilde{g}, \tilde{\pi})$ coincides with $(\bar{g}_{m, \mathbf{a}}, \bar{\pi}_{m, \mathbf{a}})$ on S_2 . We can apply Stokes Theorem to derive

$$\mathcal{I}_i(m, \mathbf{a}) = \int_{S_2} (\bar{k}_{m, \mathbf{a}})_{lj} \Omega_i^j(\partial_r)^l d\mu_{r_2} - \int_{S_1} \bar{k}_{lj} \Omega_i^j(\partial_r)^l d\mu_{r_1} + O(\epsilon^2),$$

where $d\mu_r$ is the volume form on the round sphere with radius r . Let $\tilde{\epsilon}_i = \int_{S_1} \bar{k}_{lj} \Omega_i^j(\partial_r)^l d\mu_{r_1}$, which is a constant coming from the integral on the inner sphere S_1 . Since g is close to g_{m_0} , we know that $|\tilde{\epsilon}_i| \lesssim \epsilon$. We emphasize that $\tilde{\epsilon}_k$ is independent of (m, \mathbf{a}) and the constant C_0 . Thus,

$$\mathcal{I}_i(m, \mathbf{a}) = \int_{S_2} (\bar{k}_{m, \mathbf{a}})_{lj} \Omega_i^j(\partial_r)^l d\mu_{r_2} - \tilde{\epsilon}_i + O(\epsilon^2).$$

To understand the local behavior of \mathcal{I}_i , we study the differential of the map

$$J : \mathbf{a} \mapsto \left(\int_{S_2} (\bar{k}_{m, \mathbf{a}})_{lj} \Omega_1^j(\partial_r)^l d\mu_{r_2}, \int_{S_2} (\bar{k}_{m, \mathbf{a}})_{lj} \Omega_2^j(\partial_r)^l d\mu_{r_2}, \int_{S_2} (\bar{k}_{m, \mathbf{a}})_{lj} \Omega_3^j(\partial_r)^l d\mu_{r_2} \right)$$

at $\mathbf{a} = 0$, where m is viewed as a fixed parameter.

We first compute

$$dJ|_{\mathbf{a}=0}(0, 0, 1) = \frac{d}{dt} \Big|_{t=0} \left(\int_{S_2} (\bar{k}_{m,(0,0,t)})_{lj} \Omega_i^j(\partial_r)^l d\mu_{r_2} \right)_{i=1,2,3}.$$

When $i = 1$, we have

$$\begin{aligned} \int_{S_2} (\bar{k}_{m,(0,0,t)})_{lj} \Omega_1^j(\partial_r)^l d\mu_{r_2} \\ = - \int_0^\pi d\theta \int_0^{2\pi} \cos \varphi \cot \theta \bar{k}_{m,(0,0,t)}(\partial_\varphi, \partial_r)|_{r=r_2} \cdot r_2^2 \sin \theta d\varphi = 0 \end{aligned}$$

since $\bar{k}_{m,(0,0,t)}(\partial_\varphi, \partial_r)$ does not depend on φ and $\int_0^{2\pi} \cos \varphi d\varphi = 0$. When $i = 2$, similarly, we have

$$\int_{S_2} (\bar{k}_{m,(0,0,t)})_{lj} \Omega_2^j(\partial_r)^l d\mu_{r_2} = 0.$$

When $i = 3$, in view of the exact formula for $k_{m,(0,0,a)}$ in the previous section, the $drd\varphi$ component of $\bar{k}_{m,(0,0,t)}$ is of the form $\frac{2mt}{r^2} + O(t^2)$ for sufficiently small t . Thus,

$$\begin{aligned} \int_{S_2} (\bar{k}_{m,(0,0,t)})_{lj} \Omega_3^j(\partial_r)^l d\mu_{r_2} &= \int_0^\pi d\theta \int_0^{2\pi} 2mt \sin \theta d\varphi + O(t^2) \\ &= 8\pi mt + O(t^2). \end{aligned}$$

Therefore, we finally have

$$\frac{d}{dt} \Big|_{t=0} \left(\int_{S_2} (\bar{k}_{m,(0,0,t)})_{lj} \Omega_3^j(\partial_r)^l d\mu_{r_2} \right) = 8\pi m,$$

or equivalently, $dJ|_{\mathbf{a}=0}(0, 0, 1) = (0, 0, 8\pi m)$.

To compute $dJ|_{\mathbf{a}=0}(0, 1, 0)$, we use the $SO(3)$ symmetry on \mathbf{a} . We take $R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$, which maps $(0, 1, 0)$ to $(0, 0, 1)$, and use dR to denote its differential. Apparently, we have

$$dR : (\Omega_1, \Omega_2, \Omega_3) \mapsto (\Omega_1, \Omega_3, -\Omega_2).$$

Therefore, we have

$$\begin{aligned} dJ|_{\mathbf{a}=0}(0, 1, 0) &= \frac{d}{dt} \Big|_{t=0} \left(\int_{S_2} (R^* \bar{k}_{m,(0,0,t)})_{lj} \Omega_i^j(\partial_r)^l d\mu_{r_2} \right)_{i=1,2,3} \\ &= \frac{d}{dt} \Big|_{t=0} \left(\int_{S_2} (\bar{k}_{m,(0,0,t)})_{lj} (dR(\Omega_i))^j(\partial_r)^l d\mu_{r_2} \right) \\ &= (0, 8\pi m, 0). \end{aligned}$$

Similarly, we take $R = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, which maps $(1, 0, 0)$ to $(0, 0, 1)$, and we have

$$dR : (\Omega_1, \Omega_2, \Omega_3) \mapsto (\Omega_3, \Omega_2, -\Omega_1).$$

As a result, we obtain

$$dJ|_{\mathbf{a}=0} (1, 0, 0) = (8\pi m, 0, 0).$$

Combining all those computations on $dJ|_{\mathbf{a}=0}$ and the fact that $J(0) = 0$, for $(m, \mathbf{a}) \in B_{C_0\epsilon}$, we have

$$\begin{aligned} J(\mathbf{a}) &= 8\pi m \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{a} + O(\epsilon^2) \\ &= 8\pi m_0 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{a} + O(\epsilon^2). \end{aligned}$$

We now understand the local behavior of \mathcal{I} near $(m_0, 0)$:

$$(5.6) \quad \mathcal{I}(m, \mathbf{a}) = (16\pi(m_0 - m), 8\pi m_0(a_1, a_2, a_3)) - (\tilde{\epsilon}_0, \tilde{\epsilon}_1, \tilde{\epsilon}_2, \tilde{\epsilon}_3) + O(\epsilon^2),$$

where $|\tilde{\epsilon}_i| \lesssim \epsilon$ for $i = 1, 2, 3$.

We now carry out a degree argument as in [4] and [5] to show that $\mathcal{I}(m, \mathbf{a}) = 0$ for some $(m, \mathbf{a}) \in B_{C_0\epsilon}$ provided C_0 is sufficiently large. We define

$$\mathcal{I}^{(1)}(m, \mathbf{a}) = (16\pi(m - m_0), 8\pi m_0(a_1, a_2, a_3)) - (\tilde{\epsilon}_0, \tilde{\epsilon}_1, \tilde{\epsilon}_2, \tilde{\epsilon}_3),$$

and we then choose C_0 such that $16\pi C_0 \geq 2|\tilde{\epsilon}_0|$ and $8\pi m_0 C_0 \geq 2|\tilde{\epsilon}_i|$ for $i = 1, 2, 3$. Therefore, we have

$$\mathcal{I}^{(1)}\left(m_0 + \frac{\tilde{\epsilon}_0}{16\pi}, \frac{1}{8\pi m_0}(\tilde{\epsilon}_1, \tilde{\epsilon}_2, \tilde{\epsilon}_3)\right) = 0$$

and $(m_0 + \frac{\tilde{\epsilon}_0}{16\pi}, \frac{1}{8\pi m_0}(\tilde{\epsilon}_1, \tilde{\epsilon}_2, \tilde{\epsilon}_3)) \in B_{C_0\epsilon}$. Therefore, \mathcal{I}_1 is a homomorphism from $B_{C_0\epsilon}$ to another box containing 0 in \mathbb{R}^4 centered at $(\tilde{\epsilon}_0, \tilde{\epsilon}_k)$. We then define

$$\mathcal{I}(t, m, \mathbf{a}) = (16\pi(m - m_0), 8\pi m_0(a_1, a_2, a_3)) - (\tilde{\epsilon}_0, \tilde{\epsilon}_1, \tilde{\epsilon}_2, \tilde{\epsilon}_3) + tO(\epsilon^2)$$

on $[0, 1] \times B_{C_0\epsilon}$ to be a homotopy between \mathcal{I}_1 and \mathcal{I} , where the error term $tO(\epsilon^2)$ is coming from the $O(\epsilon^2)$ term of (5.6). For sufficiently small ϵ , $0 \notin \mathcal{I}([0, 1] \times \partial B_{C_0\epsilon})$, so the degree of \mathcal{I} at the value 0 is equal to the degree of \mathcal{I}_1 at the value 0, which is 1. This implies that $\mathcal{I}(m, \mathbf{a}) = 0$ for some $(m, \mathbf{a}) \in B_{C_0\epsilon}$.

Therefore, by a suitable choice of $(m, \mathbf{a}) \in B_{C_0\epsilon}$, we finally deform (\tilde{g}, \tilde{k}) to be $(\tilde{g}_S, \tilde{k}_S)$ where $\tilde{g}_S = \tilde{g} + h$ and $\tilde{k}_S = \tilde{\pi} + \omega + \frac{1}{2} \text{tr}_{\tilde{g}_S}(\tilde{\pi} + \omega)\tilde{g}_S$, so that $(\tilde{g}_S, \tilde{k}_S)$ satisfies the constraint equations on H , $(\tilde{g}_S, \tilde{k}_S) = (\tilde{g}, \tilde{k})$ near S_1 and $(\tilde{g}_S, \tilde{k}_S) = (\tilde{g}_{m,\mathbf{a}}, \tilde{k}_{m,\mathbf{a}})$ near S_2 . This completes the gluing construction.

5.3. *Proof of the Main Theorem.* Recall the choice of H ; it is only defined for $\underline{u} \geq \delta$ and it has an inner boundary at $H \cap \underline{C}_\delta$. We would like to extend H to the interior to complete it as a three manifold without boundary. This can be done by choosing a smooth function $u = f(\underline{u})$ such that

$$\begin{cases} f'(\underline{u}) = -1 & \text{for } \underline{u} \leq 0, \\ f'(\underline{u}) < 0 & \text{for } 0 \leq \underline{u} \leq \delta, \\ t(\underline{u}, f(\underline{u})) = 0 & \text{for } \underline{u} \geq \delta. \end{cases}$$

In addition, f can be chosen such that $\text{tr} \chi(\underline{u}, f(\underline{u})) > 0$ for all $0 \leq \underline{u} \leq \delta$. The space-like piece in the picture in Section 1.2 can be viewed as the graph of f in \underline{u} - u plane. Suppose that the curve $u = f(\underline{u})$ intersects the ‘‘central’’ line $\underline{u} = u$ at $(\underline{u}_{\text{cen}}, f(\underline{u}_{\text{cen}}) = \underline{u}_{\text{cen}})$. Let $H_0 = \bigcup_{u=f(\underline{u}), \underline{u}_{\text{cen}} \leq \underline{u} \leq \delta} S_{\underline{u}, u}$, which is smooth space-like hypersurface of M with (\bar{g}, \bar{k}) as the induced metric and the second fundamental form. It is clear from the construction that $(\bar{g}, \bar{k}) = (\delta_{ij}, 0)$ for $\underline{u}_{\text{cen}} \leq \underline{u} \leq 0$, where δ is the Euclidean metric on the ball of radius r_0 that is the radius of $S_{0, f(0)}$ computed in g . It is also clear that $0 < r_0 < r_1$.

We divide $\Sigma = \mathbb{R}^3$ with Cartesian coordinate (x^1, x^2, x^3) into four concentric regions $\Sigma_M, \Sigma_C, \Sigma_S$ and Σ_K as

$$\begin{aligned} \Sigma_M &= \{x : |x| \leq r_0\}, & \Sigma_C &= \{x : r_0 \leq |x| \leq r_1\}, \\ \Sigma_S &= \{x : r_1 \leq |x| \leq r_2\}, & \Sigma_K &= \{x : |x| \geq r_2\}. \end{aligned}$$

By construction, we have $r_1 - r_0 = O(\delta)$ and $r_2 - r_1 = O(\epsilon_0)$. Thus, H_0 is diffeomorphic to $\Sigma_M \cup \Sigma_C$ and H is diffeomorphic to Σ_S . These diffeomorphisms are realized through the Cartesian coordinates in an obvious way. We now define $(\bar{g}_\Sigma, \bar{k}_\Sigma)$ on Σ as follows:

$$(\bar{g}_\Sigma, \bar{k}_\Sigma) = \begin{cases} (\bar{g}, \bar{k}) & \text{on } \Sigma_M \cup \Sigma_C, \\ (\bar{g}_S, \bar{k}_S) & \text{on } \Sigma_S, \\ (\bar{g}_{m,\mathbf{a}}, \bar{k}_{m,\mathbf{a}}) & \text{on } \Sigma_K. \end{cases}$$

By construction, $(\bar{g}_\Sigma, \bar{k}_\Sigma)$ is smooth on Σ . It is *Minkowski* inside, namely, $(\bar{g}_\Sigma, \bar{k}_\Sigma) = (\delta, 0)$ on Σ_M ; similarly, it is *Kerr* on the outer region Σ_K with $|m - m_0| + |\mathbf{a}| \lesssim \epsilon$. Moreover, according to Christodoulou’s work [2], there will be a trapped surface in the future domain of dependence of $\Sigma_M \cup \Sigma_C$.

To complete the proof of the Main Theorem, we also have to show the nonexistence of trapped surfaces on Σ . For any two-sphere S embedded in Σ , we let $S_{r_S} = \{x : |x| = r_S\}$ be the sphere that is the innermost one of the spheres in the form $S_r = \{x : |x| = r\}$ containing S . Then S and S_{r_S} are tangent at some point p . The outer null expansion θ_S of S , which is the mean curvature of S related to the outer null normal vector, can be computed as $\theta_S = \text{tr}_S \bar{k}_\Sigma + H_S$, where $\text{tr}_S \bar{k}_\Sigma$ is the trace of \bar{k}_Σ computed on S and

H_S is the mean curvature of S in Σ with respect to the outer normal. By construction, $\theta_{S_r} > 0$ for all r . In fact, it is obvious that $\theta_{S_r} > 0$ for $r \leq r_0$ and $r \geq r_2$ because in these cases $S_r \subset \Sigma_M \cup \Sigma_K$, where Σ_M and Σ_K are exactly constant time slices of Minkowski and Kerr space-times. For $r_0 \leq r \leq r_1$, that $\theta_{S_r} > 0$ is ensured by our choice of f that satisfies $\text{tr} \chi(\underline{u}, f(\underline{u})) > 0$ for $0 \leq \underline{u} \leq \delta$, and for $r_1 \leq r \leq r_2$, that $\theta_{S_r} > 0$ is ensured by the choice of inner boundary $S_1 = S_{\delta, u_1}$ of the gluing region H , such that $\text{tr} \chi(\delta, u_1) > 0$ and the closeness of H to the constant time slice of Schwarzschild space-time in Boyer-Lindquist coordinate. Then at the point p , $\text{tr}_S \bar{k}_\Sigma(p) = \text{tr}_{S_{r_S}} \bar{k}_\Sigma(p)$ since S and S_{r_S} tangent at p , and $H_S(p) \geq H_{S_{r_S}}(p)$ by the maximum principle. So

$$\theta_S(p) = \text{tr}_S \bar{k}_\Sigma(p) + H_S(p) \geq \text{tr}_{S_{r_S}} \bar{k}_\Sigma(p) + H_{S_{r_S}}(p) = \theta_{S_{r_S}}(p) > 0.$$

Now assume that S is a trapped surface; then by the definition of trapped surfaces, both null expansions are negative on S . In particular, the outer null expansion θ_S should be negative on S , which contradicts that $\theta_S(p) > 0$. Therefore, there are no trapped surfaces embedded in Σ , and this completes the proof of the Main Theorem.

References

- [1] D. CHRISTODOULOU, The formation of black holes and singularities in spherically symmetric gravitational collapse, *Comm. Pure Appl. Math.* **44** (1991), 339–373. MR 1090436. Zbl 0728.53061. <http://dx.doi.org/10.1002/cpa.3160440305>.
- [2] D. CHRISTODOULOU, *The Formation of Black Holes in General Relativity*, *EMS Monogr. Math.*, European Mathematical Society (EMS), Zürich, 2009. MR 2488976. Zbl 1197.83004. <http://dx.doi.org/10.4171/068>.
- [3] D. CHRISTODOULOU and S. KLAINERMAN, *The Global Nonlinear Stability of the Minkowski Space*, *Princeton Mathematical Series* **41**, Princeton Univ. Press, Princeton, NJ, 1993. MR 1316662. Zbl 0827.53055.
- [4] J. CORVINO, Scalar curvature deformation and a gluing construction for the Einstein constraint equations, *Comm. Math. Phys.* **214** (2000), 137–189. MR 1794269. Zbl 1031.53064. <http://dx.doi.org/10.1007/PL00005533>.
- [5] J. CORVINO and R. M. SCHOEN, On the asymptotics for the vacuum Einstein constraint equations, *J. Differential Geom.* **73** (2006), 185–217. MR 2225517. Zbl 1122.58016. Available at <http://projecteuclid.org/euclid.jdg/1146169910>.
- [6] S. KLAINERMAN and I. RODNIANSKI, Sharp trace theorems for null hypersurfaces on Einstein metrics with finite curvature flux, *Geom. Funct. Anal.* **16** (2006), 164–229. MR 2221255. Zbl 1206.35081. <http://dx.doi.org/10.1007/s00039-006-0557-8>.
- [7] S. KLAINERMAN and I. RODNIANSKI, Ricci defects of microlocalized Einstein metrics, *J. Hyperbolic Differ. Equ.* **1** (2004), 85–113. MR 2052472. Zbl 1063.53051. <http://dx.doi.org/10.1142/S0219891604000056>.

- [8] S. KLAINERMAN and I. RODNIANSKI, Causal geometry of Einstein-vacuum spacetimes with finite curvature flux, *Invent. Math.* **159** (2005), 437–529. MR 2125732. Zbl 1136.58018. <http://dx.doi.org/10.1007/s00222-004-0365-4>.
- [9] S. KLAINERMAN and I. RODNIANSKI, The causal structure of microlocalized rough Einstein metrics, *Ann. of Math.* **161** (2005), 1195–1243. MR 2180401. Zbl 1089.83007. <http://dx.doi.org/10.4007/annals.2005.161.1195>.
- [10] S. KLAINERMAN and I. RODNIANSKI, Rough solutions of the Einstein-vacuum equations, *Ann. of Math.* **161** (2005), 1143–1193. MR 2180400. Zbl 1089.83006. <http://dx.doi.org/10.4007/annals.2005.161.1143>.
- [11] S. KLAINERMAN and I. RODNIANSKI, On the formation of trapped surfaces, *Acta Math.* **208** (2012), 211–333. MR 2931382. Zbl 1246.83028. <http://dx.doi.org/10.1007/s11511-012-0077-3>.
- [12] J. LUK, On the local existence for the characteristic initial value problem in general relativity, *Int. Math. Res. Not.* **2012**, no. 20, 4625–4678. MR 2989616. Zbl 1262.83011. <http://dx.doi.org/10.1093/imrn/rnr201>.
- [13] M. REITERER and E. TRUBOWITZ, Strongly focused gravitational waves, *Comm. Math. Phys.* **307** (2011), 275–313. MR 2837117. Zbl 1229.83027. <http://dx.doi.org/10.1007/s00220-011-1314-x>.
- [14] A. D. RENDALL, Reduction of the characteristic initial value problem to the Cauchy problem and its applications to the Einstein equations, *Proc. Roy. Soc. London Ser. A* **427** (1990), 221–239. MR 1032984. Zbl 0701.35149. <http://dx.doi.org/10.1098/rspa.1990.0009>.
- [15] R. SCHOEN and S. T. YAU, The existence of a black hole due to condensation of matter, *Comm. Math. Phys.* **90** (1983), 575–579. MR 0719436. Zbl 0541.53054. <http://dx.doi.org/10.1007/BF01216187>.
- [16] S. T. YAU, Geometry of three manifolds and existence of black hole due to boundary effect, *Adv. Theor. Math. Phys.* **5** (2001), 755–767. MR 1926294. Zbl 1019.53016.

(Received: July 13, 2012)

(Revised: November 12, 2013)

DEPARTMENT OF MATHEMATICS, SUN YAT-SEN UNIVERSITY, GUANGZHOU, CHINA
E-mail: lijunbin@mail.sysu.edu.cn

MATHEMATICAL SCIENCES CENTER, TSINGHUA UNIVERSITY, BEIJING, CHINA
E-mail: pin@math.tsinghua.edu.cn