LOCALIZED DEFORMATION FOR INITIAL DATA SETS WITH THE DOMINANT ENERGY CONDITION

JUSTIN CORVINO AND LAN-HSUAN HUANG

ABSTRACT. We consider localized deformation for initial data sets of the Einstein field equations with the dominant energy condition. Deformation results with the weak inequality need to be handled delicately. We introduce a modified constraint operator to absorb the first order change of the metric in the dominant energy condition. By establishing the local surjectivity theorem, we can promote the dominant energy condition to the strict inequality by compactly supported variations and obtain new gluing results with the dominant energy condition. The proof of local surjectivity is a modification of the earlier work for the usual constraint map by the first named author and R. Schoen [7] and by P. Chruściel and E. Delay [3], with some refined analysis.

1. INTRODUCTION

Deformations to obtain the strict dominant energy condition are important analytical tools in the study of initial data sets. Among various applications, the most prominent one is perhaps the proof of the Positive Mass Theorem by R. Schoen and S.-T. Yau, in which they use the strict dominant energy condition in conjunction with the stability of a minimal hypersurface (or more generally a marginally outer trapped hypersurface) to study the geometry and topology of the manifold. Their deformation results for asymptotically flat manifolds are *global* because their argument involves a conformal change of the metric, and the resulting variations of the initial data set, which satisfy an elliptic equation, cannot have compact supports, see [19] for the scalar curvature operator and [20] for the dominant energy condition with nonzero current density J. A general global deformation result for asymptotically flat initial data sets is obtained by the second named author with M. Eichmair, D. Lee, and R. Schoen as a central analytical step in the proof of the spacetime Positive Mass Theorem [11].

In contrast, if one restricts to compactly supported variations of initial data sets, so-called *localized* deformations, there is an obstruction to deform to the strict dominant energy condition. For vacuum initial data sets, the obstruction is related to whether the corresponding spacetime has Killing vector fields. More specifically, work of A. Fischer and J. Marsden [12, 13] shows that the constraint map is locally surjective if the kernel of the formal L^2 adjoint of the linearized constraint operator is trivial, a condition which V. Moncrief [18] proves is equivalent to the absence of spacetime Killing vector fields. Fischer-Marsden's proof uses the splitting theorem of the function spaces on a closed manifold in an essential way. For compact manifolds with boundary, the first named author uses a variational approach to prove a local surjectivity result for the scalar curvature

operator [5], and then with R. Schoen for the full constraint map [7]. P. Chruściel and E. Delay introduce finer weighted spaces and derive a systematic approach to localized deformations for the constraint map in various settings [3]. Localized deformations play an important role in gluing constructions (e.g. [5, 7, 3, 2, 17, 4]) and have applications to rigidity type results [5, Theorem 8], but all earlier results focus on initial data sets that are either *vacuum* or have the *strict* dominant energy condition. In this paper, we present localized deformations without assuming either condition, by introducing a new modified constraint operator, and obtain new gluing applications. The rigidity type applications will be discussed in a forthcoming paper.

For non-vacuum initial data sets, there is a serious technical detail in deforming to the strict inequality from a weak inequality. Essentially, the deformation does not seem to follow directly from local surjectivity of the constraint map, as we now discuss. (Please refer Section 2 for the relevant definitions.) Suppose the constraint map Φ is surjective at an initial data set (g, π). Given (ψ, V), suppose one solves for a small deformation (h, w) to achieve $\Phi(g + h, \pi + w) = \Phi(g, \pi) + (2\psi, V)$. That is, the mass and current densities of the deformed initial data set $(\bar{g}, \bar{\pi}) = (g + h, \pi + w)$ are $\bar{\mu} = \mu + \psi$ and $\bar{J} = J + V$. The norm of $\bar{J}$ is taken with respect to the deformed metric, and

$$|\bar{J}|_{\bar{g}} = |J + V|_{g+h} \leq |J + V|_g + \tfrac{1}{2}|h|_g|J + V|_g + O(|h|_g^2).$$

Thus for the deformed data $(\bar{g}, \bar{\pi})$

$$(1.1) \qquad \bar{\mu} - |\bar{J}|_{\bar{g}} \geq \mu - |J + V|_g + \psi - \tfrac{1}{2}|h|_g|J + V|_g + O(|h|_g^2).$$

Note that h depends on the choice of (ψ, V) and the estimates do *not* indicate that ψ can dominate the first order change involving h to promote the dominant energy condition (while we can arrange for ψ to dominate the term $O(|h|_g^2)$).

In this work, we introduce a *modified constraint map*. Given a vector field V, a metric g and a symmetric $(0, 2)$-tensor h, let $h \cdot V$ denote the vector field dual (with respect to g) to the tensor contraction of h and V. For a fixed initial data set (g, π) and a vector field W, let $\Phi^W_{(g,\pi)}$ be defined by

$$\Phi^W_{(g,\pi)}(\gamma, \tau) = \Phi(\gamma, \tau) + \left(0, \tfrac{1}{2}\gamma \cdot (J + W)\right),$$

where $J = \operatorname{div}_g \pi$ is the current density of (g, π). When $W = V$, the additional term is designed to absorb the first order change that results in the term $\tfrac{1}{2}|h|_g|J + V|_g$ from (1.1) and is motivated by the linear map introduced in [11]. We establish a sufficient condition, in terms of the modified operator $\Phi^0_{(g,\pi)}$ (setting $W = 0$), to promote to the dominant energy condition. Throughout this paper, we let $\overline{\Omega}$ be a compact connected manifold with smooth boundary, and let Ω be the manifold interior, unless otherwise indicated.

Theorem 1.1. *Let $(g, \pi) \in C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ be an initial data set. Suppose that the kernel of $D\Phi^0_{(g,\pi)}|^*_{(g,\pi)}$ is trivial on Ω. Then there is a $C^{2,\alpha}(\overline{\Omega})$ neighborhood $\mathcal{W}$ of the zero vector and constants $\epsilon > 0$, $C > 0$ such that for $(\psi, V) \in \mathcal{B}_0 \times (\mathcal{B}_1 \cap \mathcal{W})$ with $\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq \epsilon$, there exists $(h, w) \in \mathcal{B}_2 \times \mathcal{B}_2$ with $\|(h, w)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \leq C\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1}$ such that $(\bar{g}, \bar{\pi}) = (g + h, \pi + w) \in C^{2,\alpha}(\overline{\Omega}) \times C^{2,\alpha}(\overline{\Omega})$ is an initial data set and satisfies*

$$\bar{\mu} - |\bar{J}|_{\bar{g}} \geq \mu - |J + V|_g + \psi.$$

The weighted Banach spaces $\mathcal{B}_k = \mathcal{B}_k(\Omega) \subset C^{k,\alpha}_{\mathrm{loc}}(\Omega)$ for $k = 0, 1, 2$ and the respective norms are defined in Section 2.5. In particular, if (g, π) satisfies the (weak) dominant energy condition, the above theorem gives a sufficient condition to deform to the strict dominant energy condition in Ω by setting $V = 0$ and $\psi > 0$ in Ω.

Our proofs also give the following version of Theorem 1.1 that includes higher order regularity and uniformity in the neighborhood of an arbitrary initial data set.

Theorem 1.2. *Let $k \geq 0$. Let $(g_0, \pi_0) \in C^{k+4,\alpha}(\overline{\Omega}) \times C^{k+3,\alpha}(\overline{\Omega})$ be an initial data set. Suppose that the kernel of $D\Phi^0_{(g_0,\pi_0)}|^*_{(g_0,\pi_0)}$ is trivial on Ω. Then there is a $C^{k+4,\alpha}(\overline{\Omega}) \times C^{k+3,\alpha}(\overline{\Omega})$ neighborhood $\mathcal{U}$ of (g_0, π_0), a $C^{k+2,\alpha}(\overline{\Omega})$ neighborhood $\mathcal{W}$ of the zero vector, and constants $\epsilon > 0$, $C > 0$ such that for $(g, \pi) \in \mathcal{U}$ and for $(\psi, V) \in C^{k,\alpha}_c(\Omega) \times (C^{k+1,\alpha}_c(\Omega) \cap \mathcal{W})$ with $\|(\psi, V)\|_{\mathcal{B}_2 \times \mathcal{B}_1} \leq \epsilon$, there exists $(h, w) \in C^{k+2,\alpha}_c(\Omega) \times C^{k+2,\alpha}_c(\Omega)$ with $\|(h, w)\|_{C^{k+2,\alpha} \times C^{k+2,\alpha}} \leq C\|(\psi, V)\|_{C^{k,\alpha} \times C^{k+1,\alpha}}$ such that $(\bar{g}, \bar{\pi}) = (g + h, \pi + w) \in C^{k+2,\alpha}(\overline{\Omega}) \times C^{k+2,\alpha}(\overline{\Omega})$ is an initial data set that satisfies*

$$\bar{\mu} - |\bar{J}|_{\bar{g}} \geq \mu - |J + V|_g + \psi.$$

If, in addition, $(g, \pi) \in C^\infty(\overline{\Omega})$ and $(\psi, V) \in C^\infty_c(\Omega)$, then we can achieve $(\bar{g}, \bar{\pi}) \in C^\infty(\overline{\Omega})$.

As an application, we give the following gluing construction of initial data sets from interpolation. This extends the scalar curvature result of E. Delay [9], but the presence of $|J|$ adds an analytical subtlety.

Theorem 1.3. *Let $k \geq 0$. Let $(g_0, \pi_0) \in C^{k+4,\alpha}(\overline{\Omega}) \times C^{k+3,\alpha}(\overline{\Omega})$ be an initial data set. Suppose that the kernel of $D\Phi^0_{(g_0,\pi_0)}|^*_{(g_0,\pi_0)}$ is trivial on Ω. Let $0 \leq \chi \leq 1$ be a smooth function such that $\chi(1 - \chi)$ is supported on a compact subset of Ω. Then there exists a $C^{k+4,\alpha}(\overline{\Omega}) \times C^{k+3,\alpha}(\overline{\Omega})$ neighborhood $\mathcal{U}$ of (g_0, π_0) such that for $(g_1, \pi_1), (g_2, \pi_2) \in \mathcal{U}$ and for $(g, \pi) = \chi(g_1, \pi_1) + (1 - \chi)(g_2, \pi_2)$, there exists a pair of symmetric tensors (h, w) supported in Ω such that the initial data set $(\bar{g}, \bar{\pi}) = (g + h, \pi + w) \in C^{k+2,\alpha}(\overline{\Omega}) \times C^{k+2,\alpha}(\overline{\Omega})$ satisfies*

$$\bar{\mu} - |\bar{J}|_{\bar{g}} \geq \chi(\mu_1 - |J_1|_{g_1}) + (1 - \chi)(\mu_2 - |J_2|_{g_2}).$$

If, in addition, $(g_1, \pi_1), (g_2, \pi_2) \in C^\infty(\overline{\Omega})$, then we can achieve $(\bar{g}, \bar{\pi}) \in C^\infty(\overline{\Omega})$.

We will consider deformation and gluing constructions in the asymptotically flat setting, for which it is essential to make use of the modified operator $\Phi^W_{(g,\pi)}$ for W not necessarily 0. As an application, we show that for any asymptotically flat initial data set *without* assuming the no-kernel condition, one can solve for a new initial data set that interpolates to a model initial data set in a way that the dominant energy condition also interpolates. Note that at vacuum data, the modified operator $\Phi^0_{(g,\pi)}$ recovers the usual constraint map, so that in particular the adjoint operator $(D\Phi^0_{(g_\mathbb{E},0)})^*$ at the flat data has a kernel. In the asymptotic gluing, the deformation may occur far into the asymptotically flat end, so we need to take into account the finite-dimensional approximate kernel from the flat data, by employing an admissible family. This is carried out as in the vacuum case, but we remark that the admissible family in our setting can include not only the Kerr family, but also non-vacuum ones such as the Kerr-Newman family.

Let χ be a smooth cutoff function that is $\chi = 1$ on the Euclidean unit ball B_1 and $\chi = 0$ outside B_2 with $\chi(1 - \chi)$ supported on a compact subset of $B_2 \setminus \overline{B_1}$. Let $\chi_R(y) = \chi(y/R)$ be the rescaled cutoff function.

Theorem 1.4. *Let $k \geq 0$. Let $(M, g, \pi) \in C_{\mathrm{loc}}^{k+4,\alpha} \times C_{\mathrm{loc}}^{k+3,\alpha}$ be an asymptotically flat initial data set with the ADM energy-linear momentum (E, P). Given $\epsilon > 0$, there exists $R_0 > 0$ such that for any $R > R_0$, there is an initial data set $(\bar{g}, \bar{\pi}) \in C_{\mathrm{loc}}^{k+2,\alpha} \times C_{\mathrm{loc}}^{k+2,\alpha}$ with*

$$(\bar{g}, \bar{\pi}) = (g, \pi) \quad in \ B_R$$
$$(\bar{g}, \bar{\pi}) = (g^\theta, \pi^\theta) \quad in \ M \setminus B_{2R}$$

for some (g^θ, π^θ) in the admissible family of (g, π) so that $(\bar{g}, \bar{\pi})$ satisfies the dominant energy condition

$$\bar{\mu} - |\bar{J}|_{\bar{g}} \geq \chi_R (\mu - |J|_g) + (1 - \chi_R)(\mu^\theta - |J^\theta|_{g^\theta})$$

with the strictly larger ADM energy $E^\theta > E$ and

$$|P^\theta - P| < E^\theta - E < \epsilon.$$

If, in addition, $(g, \pi), (g^\theta, \pi^\theta) \in C^\infty$, then we can achieve $(\bar{g}, \bar{\pi}) \in C^\infty$.

Remark 1.5. If $E \geq |P|$ in the above theorem, then the ADM mass of (g^θ, π^θ) is strictly larger than that of (g, π), i.e. $\sqrt{(E^\theta)^2 - |P^\theta|^2} > \sqrt{E^2 - |P|^2}$, by direct manipulations.

Note that our gluing results do not fully recover the vacuum gluing results even for vacuum initial data sets, since we do not obtain the equality for the dominant energy condition. On the other hand, our gluing construction includes the feature to bring up the ADM energies by promoting the dominant energy condition. In particular, we can glue a Schwarzschild solution to a (different) Kerr solution through a region where the dominant energy condition holds, which may not be feasible by the vacuum theorem.

The paper is organized as follows. In Section 2, we introduce basic properties of the modified constraint map and some analytical preliminaries, including an improved estimate for weighted spaces (Proposition 2.10). In Section 3, we study localized deformation by the modified constraint map and prove Theorem 1.1 and Theorem 1.3 (with $k = 0$). In Section 4, we prove asymptotic gluing results to an admissible family, including Theorem 1.4 (with $k = 0$). In Section 5 and Section 6, we prove the local surjectivity theorems for the modified operator and the projected operator, respectively; the analysis follows closely that from the vacuum case, but we include the details to emphasize the uniformity of various required estimates.

2. Preliminaries

We use the Einstein summation convention, summing over repeated upper and lower indices, throughout, and we use the convention that a semicolon denotes a covariant derivative, while a comma denotes a partial derivative.

2.1. **Initial data sets.** Let $n \geq 3$. An n-dimensional *initial data set* is an n-dimensional manifold M equipped with a C_{loc}^2 Riemannian metric g and a C_{loc}^1 symmetric $(2, 0)$ tensor K. The *mass density* μ and the *current density* J are defined by

$$\mu = \frac{1}{2}(R_g - |K|_g^2 + (\mathrm{tr}_g K)^2)$$
$$J = \mathrm{div}_g K - d(\mathrm{tr}_g K)$$

where $R_g = g^{ij}R_{ij}$ is the scalar curvature of g, with R_{ij} the components of the Ricci tensor. It is convenient for us to consider the momentum $(2, 0)$ tensor

$$\pi^{ij} = K^{ij} - (\mathrm{tr}_g K)g^{ij}.$$

Abusing terminology slightly, we refer to (g, π) as an *initial data set* throughout this paper. The initial data set is said to satisfy the *dominant energy condition* if

$$\mu \geq |J|_g$$

holds everywhere in M.

The *constraint map* is defined by

$$\Phi(g, \pi) = \left(R(g) + \frac{1}{n-1}(\mathrm{tr}_g \pi)^2 - |\pi|_g^2, \; \mathrm{div}_g \pi \right) = (2\mu, J).$$

The linearization is given by the following formula (see, for example, [11, Lemma 20])

$$
\begin{aligned}
(2.1) \quad D\Phi|_{(g,\pi)}(h, w) = \Big(& L_g h - 2h_{ij}\pi_\ell^i \pi^{j\ell} - 2\pi_k^j w_j^k + \tfrac{2}{n-1}\mathrm{tr}_g \pi (h_{ij}\pi^{ij} + \mathrm{tr}_g w), \\
& (\mathrm{div}_g w)^i - \tfrac{1}{2}\pi^{jk}h_{jk;\ell}g^{\ell i} + \pi^{jk}h^i_{j;k} + \tfrac{1}{2}\pi^{ij}(\mathrm{tr}_g h)_{,j} \Big).
\end{aligned}
$$

Here all indices are raised or lowered with respect to g. The linearized scalar curvature operator $L_g(h) = -\Delta_g(\mathrm{tr}_g h) + \mathrm{div}_g \mathrm{div}_g(h) - h^{ij}R_{ij}$ appears above. The formal L^2 adjoint operator of $D\Phi|_{(g,\pi)}$ is given by

$$
\begin{aligned}
(2.2) \quad & D\Phi|^*_{(g,\pi)}(f, X) \\
& = \Big(L_g^* f + \big(\tfrac{2}{n-1}(\mathrm{tr}_g \pi)\pi_{ij} - 2\pi_{ik}\pi_j^k \big)f \\
& \quad + \tfrac{1}{2}\big(g_{i\ell}g_{jm}(L_X\pi)^{\ell m} + (X_{;k}^k)\pi_{ij} - X_i\pi_{j;k}^k - X_j\pi_{i;k}^k - X_{k;m}\pi^{km}g_{ij} - X_k\pi^{km}_{;m}g_{ij} \big), \\
& \quad -\tfrac{1}{2}(\mathcal{D}_g X)^{ij} + \big(\tfrac{2}{n-1}(\mathrm{tr}_g \pi)g^{ij} - 2\pi^{ij} \big)f \Big),
\end{aligned}
$$

where $L_g^* f = -(\Delta_g f)g + \mathrm{Hess}_g f - f\mathrm{Ric}(g)$, $L_X\pi$ is the Lie derivative, and $\mathcal{D}_g X = L_X g$ is the Lie derivative operator $(\mathcal{D}_g X)^{ij} = X^i_{;\ell}g^{\ell j} + X^j_{;\ell}g^{\ell i}$. See [7, Lemma 2.3] for $n = 3$, [11, Lemma 20] for general n.

2.2. Modified constraint map.

Let (g, π) be an initial data set and let W be a vector field. We define the *modified* constraint map $\Phi^W_{(g,\pi)}$ by

$$(2.3) \qquad \Phi^W_{(g,\pi)}(\gamma, \tau) = \Phi(\gamma, \tau) + (0, \tfrac{1}{2}\gamma \cdot (\mathrm{div}_g \pi + W)),$$

where $(\gamma \cdot Y)^i = g^{ij}\gamma_{jk}Y^k$ in local coordinates. The linearized operator at (g, π) is denoted by $D\Phi^W_{(g,\pi)} = D\Phi^W_{(g,\pi)}|_{(g,\pi)}$ and has the following expression

$$(2.4) \qquad D\Phi^W_{(g,\pi)}(h, w) = D\Phi|_{(g,\pi)}(h, w) + (0, \tfrac{1}{2}h \cdot (\mathrm{div}_g \pi + W)).$$

The formal L^2 adjoint operator has the expression

$$(2.5) \qquad (D\Phi^W_{(g,\pi)})^*(f, X) = D\Phi|^*_{(g,\pi)}(f, X) + \big(\tfrac{1}{4}[X_i(\mathrm{div}_g \pi + W)_j + X_j(\mathrm{div}_g \pi + W)_i], 0 \big),$$

where the indices are lowered by g.

2.3. **Kernel of the adjoint operators.** We include regularity results for any kernel element $(f, X) \in H^2_{\mathrm{loc}}(U) \times H^1_{\mathrm{loc}}(U)$ where $U \subset M$ is an open subset, from which we can obtain higher order regularity depending on the smoothness of the initial data sets. The analysis is similar to the scalar curvature operator in [5, Section 2.2].

Proposition 2.1. *Let $k \geq 2$, $\alpha \in (0, 1)$. Let $(g, \pi) \in C^{k,\alpha}_{\mathrm{loc}}(U) \times C^{k-1,\alpha}_{\mathrm{loc}}(U)$ be an initial data set. Suppose that $(f, X) \in H^2_{\mathrm{loc}}(U) \times H^1_{\mathrm{loc}}(U)$ satisfies $D\Phi|^*_{(g,\pi)}(f, X) = 0$ weakly. Then the following holds:*

(1) $(f, X) \in C^{k,\alpha}_{\mathrm{loc}}(U) \times C^{k-1,\alpha}_{\mathrm{loc}}(U)$.

(2) If $(g, \pi) \in C^{k,\alpha}(\overline{U}) \times C^{k-1,\alpha}(\overline{U})$, then $(f, X) \in C^{k-2,\alpha}(\overline{U}) \times C^{k-2,\alpha}(\overline{U})$.

(3) If U is connected, the space of solutions $(f, X) \in H^2_{\mathrm{loc}}(U) \times H^1_{\mathrm{loc}}(U)$ to the homogeneous equation $D\Phi|^*_{(g,\pi)}(f, X) = 0$ is finite-dimensional, and a non-trivial solution cannot vanish on any open subset of U.

Proof. Let $(f, X) \in H^2_{\mathrm{loc}}(U) \times H^1_{\mathrm{loc}}(U)$ satisfy $D\Phi|^*_{(g,\pi)}(f, X) = 0$. Taking the trace of the first component of (2.2) gives an equation for $\Delta_g f$. Using this equation, we can eliminate the term $\Delta_g f$ from the first component of the system $D\Phi|^*_{(g,\pi)}(f, X) = 0$ to obtain

$$(2.6) \qquad f_{;ij} = A_{ij}f + B_{ijk}X^k + C^\ell_{ijk}X^k_{;\ell} ,$$

where A_{ij}, B_{ijk} and C^ℓ_{ijk} are functions locally computed as polynomials in the components g_{ab}, g^{ab}, π_{ab}, $\partial_c g_{ab}$, $\partial^2_{cd} g_{ab}$, and $\partial_c \pi_{ab}$. The other components in $D\Phi|^*_{(g,\pi)}(f, X) = 0$ constitute the following system:

$$(2.7) \qquad \tfrac{1}{2}(\mathcal{D}_g X)^{ij} = \left(\tfrac{2}{n-1}(\mathrm{tr}_g \pi)g^{ij} - 2\pi^{ij} \right) f.$$

Thus $\mathcal{D}_g X \in H^1_{\mathrm{loc}}(U)$, and by commuting the order of derivatives and using the Ricci formula, we have

$$(2.8) \qquad \begin{aligned} \Big[(\mathcal{D}_g X)_{ij;k} &+ (\mathcal{D}_g X)_{ki;j} - (\mathcal{D}_g X)_{jk;i} \Big] \\ &= (X_{i;jk} + X_{i;kj}) + (X_{j;ik} - X_{j;ki}) + (X_{k;ij} - X_{k;ji}) \\ &= 2X_{i;jk} + (R^\ell_{kji} + R^\ell_{ikj} + R^\ell_{ijk})X_\ell, \end{aligned}$$

where the sign convention for the Riemannian curvature tensor is so that the Ricci tensor $R_{jk} = R^\ell_{\ell jk}$. Along with (2.7), this implies that $X \in H^2_{\mathrm{loc}}(U)$. By taking the trace of the first component of $D\Phi|^*_{(g,\pi)}(f, X) = 0$ and the divergence of the other components, (f, X) satisfies a second order elliptic linear system (see [7, Proposition 3.1] for $n = 3$ and [11, Lemma 20] for general n). The desired interior regularity $(f, X) \in C^{k,\alpha}_{\mathrm{loc}}(U) \times C^{k-1,\alpha}_{\mathrm{loc}}(U)$ follows from elliptic regularity.

Let $\gamma = \gamma(t)$ be a geodesic. Because $\nabla_{\gamma'}\gamma' = 0$, we have

$$(2.9) \qquad \begin{aligned} (f \circ \gamma)''(t) &= f_{;ij}|_{\gamma(t)} \dot{\gamma}^i(t)\dot{\gamma}^j(t) \\ \left(\frac{D^2 X(\gamma(t))}{dt^2} \right)^k &= (\nabla_{\dot\gamma} \nabla_{\dot\gamma} X)^k = X^k_{;ij}|_{\gamma(t)} \dot{\gamma}^i(t)\dot{\gamma}^j(t). \end{aligned}$$

We have the formula for $f_{;ij}$ in (2.6). The term $X^k_{;ij}$ is obtained from (2.8):

(2.10)
$$\begin{aligned}
X^i_{;jk} &= \frac{1}{2} g^{i\ell} \left[(\mathcal{D}_g X)_{\ell j;k} + (\mathcal{D}_g X)_{k\ell;j} - (\mathcal{D}_g X)_{jk;\ell} \right] - \frac{1}{2} g^{i\ell} \widetilde{C}^p_{\ell jk} X_p \\
&= \widetilde{A}^i_{jk} f + \widetilde{B}^{i\ell}_{jk} f_{;\ell} - \frac{1}{2} g^{i\ell} \widetilde{C}^p_{\ell jk} X_p,
\end{aligned}$$

where $\widetilde{A}^i_{jk}, \widetilde{B}^{i\ell}_{jk}, \widetilde{C}^\ell_{ijk}$ are locally computed as polynomials in g_{ab}, g^{ab}, π_{ab}, $\partial_c g_{ab}$, $\partial^2_{cd} g_{ab}$, and $\partial_c \pi_{ab}$. For convenience, let $\{E_i(t), i = 1, \ldots, n\}$ be a parallel orthonormal frame field along γ. Let $X(\gamma(t)) = X^i(t) E_i(t)$, and $X^0(t) = f(\gamma(t))$. Let $Z(t)$ be the column vector with the components $X^0(t), X^1(t), \ldots, X^n(t)$. Then by (2.9), (2.6), and (2.10), the vector satisfies a second-order linear system of ordinary differential equations along any geodesic γ in U:

$$Z''(t) = A(t) Z'(t) + B(t) Z(t),$$

where $A(t)$ and $B(t)$ are $(n+1) \times (n+1)$ matrix functions whose components are computed locally as polynomials in g_{ij}, g^{ij}, π_{ij}, $\partial_k g_{ij}$, $\partial^2_{k\ell} g_{ij}$, and $\partial_k \pi_{ij}$, evaluated along γ. If U is connected, then (f, X) is determined by its 1-jet at a point in U, and thus the dimension of the kernel is at most $(n+1)^2$, and any non-trivial element in the kernel cannot vanish on an open subset.

Boundary regularity for (f, X) follows from the ODE argument. By extending the initial data set (g, π) in a neighborhood of the boundary, we may assume that $\overline{U}$ is in a manifold interior. Let $q \in \partial U$ and let $2r > 0$ be the injectivity radius at q. For a point $p \in U$ with $d(p, q) < r$, we can extend (f, X) on $B_r(p) \setminus U$ along the unique geodesic in $B_r(p)$ starting at p with initial velocity $v = \exp^{-1}(x)$ that reaches $x \in B_r(p) \setminus U$. By the smooth dependence of solutions of ODE on parameters (note that $\exp^{-1}$ is $C^{k-1,\alpha}$ and the ODE system involves $\mathrm{Ric}(g)$, which accounts in part for the ensuing regularity), we have that $(f, X) \in C^{k-2,\alpha}_{\mathrm{loc}}(B_r(p)) \times C^{k-2,\alpha}_{\mathrm{loc}}(B_r(p))$.

$\square$

The above proposition also applies to the modified constraint operator because its adjoint operator $(D\Phi^W_{(g,\pi)})^*$ differs from $D\Phi|^*_{(g,\pi)}$ only by a zero-th order term. Essentially the same proof implies the following statement.

Proposition 2.2. *Let $k \geq 2$, $\alpha \in (0, 1)$. Let $(g, \pi) \in C^{k,\alpha}_{\mathrm{loc}}(U) \times C^{k-1,\alpha}_{\mathrm{loc}}(U)$ be an initial data set. Let $W \in C^{k-2,\alpha}_{\mathrm{loc}}(U)$ be a vector field. Suppose that $(f, X) \in H^2_{\mathrm{loc}}(U) \times H^1_{\mathrm{loc}}(U)$ satisfies $(D\Phi^W_{(g,\pi)})^*(f, X) = 0$ weakly. Then the following holds:*

(1) $(f, X) \in C^{k,\alpha}_{\mathrm{loc}}(U) \times C^{k-1,\alpha}_{\mathrm{loc}}(U)$.

(2) If $(g, \pi), (\gamma, \tau) \in C^{k,\alpha}(\overline{U}) \times C^{k-1,\alpha}(\overline{U})$, $W \in C^{k-2,\alpha}(\overline{U})$, then $(f, X) \in C^{k-2,\alpha}(\overline{U}) \times C^{k-2,\alpha}(\overline{U})$.

(3) If U is connected, the space of solutions $(f, X) \in H^2_{\mathrm{loc}}(U) \times H^1_{\mathrm{loc}}(U)$ to the homogeneous equation $(D\Phi^W_{(g,\pi)})^*(f, X) = 0$ is finite-dimensional, and a non-trivial solution cannot vanish on any open subset of U.

Definition 2.3. The *kernel* of $(D\Phi^W_{(g,\pi)})^*$ on U is the set $K \subset H^2_{\mathrm{loc}}(U) \times H^1_{\mathrm{loc}}(U)$ which consists of those (f, X) that satisfy $(D\Phi^W_{(g,\pi)})^*(f, X) = 0$ weakly. The kernel of $(D\Phi^W_{(g,\pi)})^*$ is said to be *trivial* on U if $K = \{0\}$.

Example 2.4. Consider the flat data $(g_{\mathbb{E}}, 0)$ on an open connected subset of $\mathbb{R}^3$ and $W = 0$. Then the modified operator is the usual constraint map, and its formal L^2 adjoint operator $D\Phi|^*_{(g_{\mathbb{E}}, 0)}(f, X) = (-(\Delta_{g_{\mathbb{E}}} f)g_{\mathbb{E}} + \mathrm{Hess}_{g_{\mathbb{E}}} f, -\frac{1}{2}\mathcal{D}_{g_{\mathbb{E}}} X)$ has a ten-dimensional kernel $K = K_0 \oplus K_1$, where

$$K_0 = \mathrm{span}\{1, x^1, x^2, x^3\}$$

$$K_1 = \mathrm{span}\left\{\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}, x \times \frac{\partial}{\partial x^1}, x \times \frac{\partial}{\partial x^2}, x \times \frac{\partial}{\partial x^3}\right\}.$$

2.4. **Weighted Sobolev spaces.** Let $d_g(x) = d_g(x, \partial\Omega)$ be the distance to the boundary with respect to g; the boundary is assumed smooth, so d is as regular as g near $\partial\Omega$. We will work with uniformly equivalent metrics in a bounded open set $\mathcal{U}_0$ in the space of $C^m(\overline{\Omega})$ ($m \geq 2$) Riemannian metrics such that $\|d_g\|_{C^m}$ is uniformly bounded near $\partial\Omega$. We will establish a framework uniformly across $\mathcal{U}_0$ in what follows.

Let $V_\Omega = \{x \in \Omega : d_g(x) < r_0$ for some $g \in \mathcal{U}_0\}$ be a thin regular collar neighborhood of $\partial\Omega$. There is $r_0 \in (0, \frac{1}{2})$ sufficiently small so that a neighborhood of V_Ω is foliated by smooth (as regular as the metric g is) level sets of d_g and that $d_g(x) \leq \frac{1}{2}$ for all $x \in V_\Omega$ and $g \in \mathcal{U}_0$.

Let $0 < r_1 < r_0$ be fixed. Define a smooth positive monotone function $\tilde{\rho} : (0, \infty) \to \mathbb{R}$ such that $\tilde{\rho}(t) = e^{-1/t}$ for $t \in (0, r_1)$ and $\tilde{\rho}(t) = 1$ for $t > r_0$. Let N be a fixed large number, chosen so that

$$(2.11) \qquad\qquad\qquad N > \max\{4(4k - 3), 4C_0\}$$

where $C_0 > 0$ is the constant appearing in (5.3). While the discussion in this section holds for all $k \leq m$, in this paper we only apply (2.11) of the case $k \leq 2$ in the variational argument in Section 5.

Notation 2.5 (Exponential weight function). Let ρ_g be the positive function on Ω defined by

$$\rho_g(x) = (\tilde{\rho} \circ d_g(x))^N.$$

Let $L^2_{\rho_g}(\Omega, g)$ be the set of functions or tensor fields u such that $|u|\rho_g^{\frac{1}{2}} \in L^2(\Omega, g)$ with the norm defined by

$$\|u\|_{L^2_{\rho_g}(\Omega, g)} = \left(\int_\Omega |u|^2 \rho_g \, d\mu_g\right)^{\frac{1}{2}}.$$

The pairing $\langle u, v\rangle_{L^2_{\rho_g}(\Omega, g)} = \langle u\rho_g^{\frac{1}{2}}, v\rho_g^{\frac{1}{2}}\rangle_{L^2(\Omega, g)}$ makes $L^2_{\rho_g}(\Omega, g)$ a Hilbert space. Let $H^k_{\rho_g}(\Omega, g)$ be the Hilbert space of functions or tensor fields whose covariant derivatives up to and including order k with respect to g are also in $L^2_{\rho_g}(\Omega, g)$ with the norm defined as (where I is a multi-index)

$$\|u\|^2_{H^k_{\rho_g}(\Omega, g)} = \sum_{|I|=0}^k \|\nabla^I_g u\|^2_{L^2_{\rho_g}(\Omega, g)} = \sum_{|I|=0}^k \int_\Omega |\nabla^I_g u|^2 \rho_g \, d\mu_g.$$

By [7, Lemma 2.1], $H^k(\Omega, g)$ (and hence $C^\infty(\overline{\Omega})$) is dense in $H^k_{\rho_g}(\Omega, g)$. We note that the tensor fields in $H^k(\Omega, g)$ are the same across g, and while we can further shrink $\mathcal{U}_0$ so that the norms are equivalent for $g \in \mathcal{U}_0$ as well, the weighted norms $H^k_{\rho_g}(\Omega, g)$ may not be equivalent as g varies in $\mathcal{U}_0$. (Note that they would be equivalent if ρ were a polynomial weight function, i.e. $\tilde{\rho}(t) = t$ near the boundary, and such a weight can be used for simplicity to establish the finite regularity results [5, 7].) We often suppress Ω and g from the notation when it is clear from the context.

It is useful to compare the norms $\|u\|_{H^k_{\rho g}}$ and $\|u\rho_g^{\frac{1}{2}}\|_{H^k}$. We shall show that $u \in H^k_{\rho g}$ implies

$$\|u\rho_g^{\frac{1}{2}}\|_{H^k} \le C\|u\|_{H^k_{\rho g}},$$

where the constant C is uniform in $\mathcal{U}_0$.

We begin with basic lemmas for which we work at a fixed metric $g \in C^k(\overline{\Omega})$ and write $d = d_g$ and $\rho = \rho_g$; we note explicit dependence of the constants in the estimates so that the estimates will hold uniformly across $\mathcal{U}_0$.

Lemma 2.6. *Let $N \ge 1$. There is a constant $C > 0$ independent of N (depending only on k, $\|\log \tilde{\rho}\|_{C^k([r_1,r_0])}$ and $\|d\|_{C^k(\overline{V_\Omega})}$) such that*

$$|\nabla^k(\rho^{\frac{1}{2}})| \le CN^k \rho^{\frac{1}{2}} d^{-2k}.$$

holds on Ω.

Proof. By direct computation,

$$(2.12) \qquad \nabla(\rho^{\frac{1}{2}})(x) = \begin{cases} \frac{1}{2}N\rho^{\frac{1}{2}}d^{-2}\nabla d & \text{if } 0 < d(x) \le r_1 \\ \frac{1}{2}N\rho^{\frac{1}{2}}d^{-2}\left(d^2(\log \tilde{\rho})'(d)\right)\nabla d & \text{if } r_1 \le d(x) \le r_0 \\ 0 & \text{if } d(x) \ge r_0 \end{cases}.$$

This implies the estimate for $k = 1$. The estimate for $k > 1$ follows from induction. □

Lemma 2.7. *Let $N \ge 1$. There is $r_2 \in (0, r_1)$ independent of N (depending only on $\|\Delta d\|_{C^0(\overline{V_\Omega})}$) such that if $0 < d(x) \le r_2$, we have*

$$(2.13) \qquad \frac{1}{2}N^2 d^{-4}\rho \le \Delta\rho.$$

Proof. By direct computation, for $0 < d(x) < r_1$,

$$\Delta\rho = N^2 d^{-4}\rho(1 + N^{-1}d^2\Delta d - 2N^{-1}d).$$

If r_2 is sufficiently small, for $0 < d(x) \le r_2$,

$$1 + N^{-1}d^2\Delta d - 2N^{-1}d \ge 1 - r_2^2\|\Delta d\|_{C^0(\overline{V_\Omega})} - 2r_2 \ge \frac{1}{2}.$$

□

In the next lemma we use a cutoff function $\xi = \xi_g$ to handle estimates near the boundary. Let $0 \le \xi_g \le 1$ be smooth with $\xi_g = 0$ on the compact subset $\{x \in \Omega : d_g(x) \ge r_2\}$ and $\xi_g = 1$ in a collar neighborhood $\{x \in \Omega : d_g(x) \le r_2/2\}$ of $\partial\Omega$ with $|\nabla\xi|_g \le 4/r_2$. In the following lemma, u can be a function or a tensor field.

Lemma 2.8. *For $j \in \mathbb{Z}_+$, and for $u \in C^j(\overline{\Omega})$, if $N \ge 4(4j - 3)$, then*

$$\int_\Omega \xi|u|^2 d^{-4j}\rho \, d\mu_g \le \left(\frac{4}{N}\right)^j \int_\Omega \xi|\nabla^j u|^2 \rho \, d\mu_g$$

$$+ \sum_{i=1}^{j}\left(\frac{4}{N}\right)^{j+1-i}\sup_\Omega(|\nabla\xi|d^{-4i+2})\|\nabla^{j-i}u\|^2_{L^2_\rho(\Omega)}.$$

Proof. Multiplying $\xi|u|^2 d^{-4j+4}$ to (2.13) and noting $|\nabla\rho| \leq Nd^{-2}\rho$ and $d \leq 1$ in the supports of ξ and $\nabla\xi$ by (2.12), we have

$$\frac{1}{2}N^2 \int_\Omega \xi|u|^2 d^{-4j}\rho \, d\mu_g \leq \int_\Omega \xi|u|^2 d^{-4j+4}\Delta\rho \, d\mu_g$$

$$\leq \int_\Omega \left[2\xi|\nabla u||u|d^{-4j+4} + (4j-4)\xi|u|^2 d^{-4j+3} + |\nabla\xi||u|^2 d^{-4j+4}\right]|\nabla\rho| \, d\mu_g$$

$$\leq N \int_\Omega \left[2\xi|\nabla u||u|d^{-4j+2} + (4j-4)\xi|u|^2 d^{-4j+1} + |\nabla\xi||u|^2 d^{-4j+2}\right]\rho \, d\mu_g$$

$$\leq N \left((4j-3)\int_\Omega \xi|u|^2 d^{-4j}\rho \, d\mu_g + \int_\Omega \xi|\nabla u|^2 d^{-4j+4}\rho \, d\mu_g + \sup_\Omega(|\nabla\xi|d^{-4j+2})\|u\|^2_{L^2_\rho}\right),$$

where we applied the AM-GM inequality in the last inequality. Absorbing the first term into the left hand side, we have

$$\int_\Omega \xi|u|^2 d^{-4j}\rho \, d\mu_g \leq \frac{4}{N}\int_\Omega \xi|\nabla u|^2 d^{-4j+4}\rho \, d\mu_g + \frac{4}{N}\sup_\Omega(|\nabla\xi|d^{-4j+2})\|u\|^2_{L^2_\rho}.$$

This proves the case $j = 1$. The case $j > 1$ follows by induction. $\qquad\square$

Corollary 2.9. *Let $u \in H^k_\rho(\Omega)$ and $N \geq 4(4k-3)$. For $j = 0, 1, \ldots, k$,*

$$\int_\Omega |\nabla^{k-j}u|^2 d^{-4j}\rho \, d\mu_g \leq C\|u\|^2_{H^k_\rho(\Omega)},$$

where C depends on N, j and r_2.

Proof. By density, it suffices to prove the estimate for $u \in C^\infty(\overline{\Omega})$. By Lemma 2.8,

$$\int_\Omega |\nabla^{k-j}u|^2 d^{-4j}\rho \, d\mu_g$$

$$= \int_\Omega \xi|\nabla^{k-j}u|^2 d^{-4j}\rho \, d\mu_g + \int_\Omega (1-\xi)|\nabla^{k-j}u|^2 d^{-4j}\rho \, d\mu_g$$

(2.14)

$$\leq \left(\frac{4}{N}\right)^j \int_\Omega \xi|\nabla^k u|^2\rho + \sum_{i=1}^j \left(\frac{4}{N}\right)^{j+1-i}\sup_\Omega(|\nabla\xi|d^{-4i+2})\|\nabla^{k-i}u\|^2_{L^2_\rho}$$

$$+ \sup_\Omega((1-\xi)d^{-4j})\|\nabla^{k-j}u\|^2_{L^2_\rho}.$$

This implies the desired inequality. $\qquad\square$

Proposition 2.10. *Let $u \in H^k_\rho(\Omega)$ and $N \geq 4(4k-3)$. Then*

$$\|u\rho^{\frac{1}{2}}\|_{H^k(\Omega)} \leq C\|u\|_{H^k_\rho(\Omega)},$$

where C depends on N, k, r_2, $\|\log\tilde{\rho}\|_{C^k([r_1,r_0])}$ and $\|d\|_{C^k(\overline{V_\Omega})}$.

Proof. Recall

$$\|u\rho^{\frac{1}{2}}\|^2_{H^k} = \sum_{j=0}^k \|\nabla^j(u\rho^{\frac{1}{2}})\|^2_{L^2}.$$

We may assume $u \in C^\infty(\overline{\Omega})$, using density. By Lemma 2.6, for some constants C_{ij},

$$|\nabla^j(u\rho^{\frac{1}{2}})| \leq |\nabla^j u|\rho^{\frac{1}{2}} + \sum_{i=1}^j C_{ij}|\nabla^{j-i}u| \, |\nabla^i(\rho^{\frac{1}{2}})|$$

$$\leq |\nabla^j u|\rho^{\frac{1}{2}} + C\sum_{i=1}^j N^i|\nabla^{j-i}u|d^{-2i}\rho^{\frac{1}{2}}.$$

Therefore,

$$\|\nabla^j(u\rho^{\frac{1}{2}})\|_{L^2}^2 = \int_\Omega |\nabla^j(u\rho^{\frac{1}{2}})|^2 \, d\mu_g$$

$$\leq 2\int_\Omega |\nabla^j u|^2 \rho \, d\mu_g + C^2 N^{2j} \sum_{i=1}^j \int_\Omega |\nabla^{j-i}u|^2 d^{-4i}\rho \, d\mu_g.$$

This implies the desired estimate by Corollary 2.9. $\qquad\qquad\square$

2.5. **Weighted Hölder spaces.** Here we follow the idea of [3] to consider weighted Hölder norms in small balls $B_{\phi(x)}(x)$ that cover Ω. The weight function $\phi = \phi_g$ satisfies the following properties with uniform estimates across g in a $C^m(\overline{\Omega})$ neighborhood $\mathcal{U}_0$. We suppress the subscript g in ϕ, d, ρ, ∇, when the context is clear. Recall the neighborhood V_Ω defined in the previous section.

Proposition 2.11. *For $g \in \mathcal{U}_0$, we define $\phi(x) = (d(x))^2$ in V_Ω. There exists a constant $C > 0$, uniform across $\mathcal{U}_0$, such that we can extend ϕ to Ω with $0 < \phi < 1$ and with the following properties.*

(1) ϕ has a positive lower bound on $\Omega \setminus V_\Omega$ uniformly in $g \in \mathcal{U}_0$, and for each x, $B_{\phi(x)}(x)$ is contained in a coordinate ball of Ω.

(2) For $x \in \Omega$ and $k \leq m$, we have

$$|\phi^k \rho^{-1}\nabla^k \rho| \leq C.$$

(3) For $x \in \Omega$ and for $y \in B_{\phi(x)}(x)$, we have

(2.15)
$$C^{-1}\rho(y) \leq \rho(x) \leq C\rho(y)$$
$$C^{-1}\phi(y) \leq \phi(x) \leq C\phi(y).$$

Proof. (1) is obvious. (2) follows by Lemma 2.6. It is clear that (3) holds for $x \in \Omega \setminus V_\Omega$ and $y \in B_{\phi(x)}(x)$, since both ρ and ϕ have positive uniform lower bounds. For $x \in V_\Omega$ and $y \in B_{\phi(x)}(x)$, the triangle inequality implies $d(x) - \phi(x) \leq d(y) \leq d(x) + \phi(x)$. The desired estimates follows since $\phi = d^2$ and $d \leq \frac{1}{2}$ in V_Ω. $\qquad\square$

Let $r, s \in \mathbb{R}$ and $\varphi = \phi^r \rho^s$. For $u \in C_{loc}^{k,\alpha}(\Omega)$, we define the weighted Hölder norms $\|u\|_{C_{\phi,\varphi}^{k,\alpha}(\Omega)}$ by

$$\|u\|_{C_{\phi,\varphi}^{k,\alpha}(\Omega)} = \sup_{x\in\Omega}\left(\sum_{j=0}^k \varphi(x)\phi^j(x)\|\nabla_g^j u\|_{C^0(B_{\phi(x)}(x))} + \varphi(x)\phi^{k+\alpha}(x)[\nabla_g^k u]_{0,\alpha;B_{\phi(x)}(x)}\right).$$

Note that ϕ is to make the norm scaling invariant with respect to the size of the ball. The weighted Hölder space $C_{\phi,\varphi}^{k,\alpha}(\Omega)$ consists of $C_{loc}^{k,\alpha}(\Omega)$ functions or tensor fields with finite $C_{\phi,\varphi}^{k,\alpha}(\Omega)$ norm. If $u \in C_{\phi,\varphi}^{k,\alpha}(\Omega)$, then u is dominated by φ^{-1} in the sense that $u = O(\varphi^{-1})$ and $\nabla^j u = O(\varphi^{-1}\phi^{-j})$ near the boundary. The norms are equivalent to those introduced in [3, Appendix A] (cf. [6]).

Note that differentiation is a continuous map from $C^{k,\alpha}_{\phi,\varphi}$ to $C^{k-1,\alpha}_{\phi,\phi\varphi}$. For $u \in C^{k,\alpha}_{\phi,\varphi}(\Omega)$, $v \in C^{k,\alpha}(\overline{\Omega})$ we have $uv \in C^{k,\alpha}_{\phi,\varphi}(\Omega)$ with

$$\|uv\|_{C^{k,\alpha}_{\phi,\varphi}(\Omega)} \leq C\|u\|_{C^{k,\alpha}_{\phi,\varphi}(\Omega)}\|v\|_{C^{k,\alpha}(\overline{\Omega})}$$

where C depends only on k. Furthermore, using Lemma 2.6, it follows that multiplication by ρ is a continuous map from $C^{k,\alpha}_{\phi,\varphi}$ to $C^{k,\alpha}_{\phi,\varphi\rho^{-1}}$.

We will use the following Banach spaces $\mathcal{B}_k(\Omega)$ (for functions or tensor fields):

$$\mathcal{B}_k(\Omega) = C^{k,\alpha}_{\phi,\phi^{4-k+\frac{n}{2}}\rho^{-\frac{1}{2}}}(\Omega) \cap L^2_{\rho^{-1}}(\Omega) \qquad \text{(for } k = 0,1,2)$$

$$\mathcal{B}_3(\Omega) = C^{3,\alpha}_{\phi,\phi^{1+\frac{n}{2}}\rho^{\frac{1}{2}}}(\Omega) \cap H^1_\rho(\Omega)$$

$$\mathcal{B}_4(\Omega) = C^{4,\alpha}_{\phi,\phi^{\frac{n}{2}}\rho^{\frac{1}{2}}}(\Omega) \cap H^2_\rho(\Omega) \,,$$

with the Banach norms:

$$\|u\|_{\mathcal{B}_k(\Omega)} = \|u\|_{C^{k,\alpha}_{\phi,\phi^{4-k+\frac{n}{2}}\rho^{-\frac{1}{2}}}(\Omega)} + \|u\|_{L^2_{\rho^{-1}}(\Omega)} \qquad \text{(for } k = 0,1,2)$$

$$\|X\|_{\mathcal{B}_3(\Omega)} = \|X\|_{C^{3,\alpha}_{\phi,\phi^{1+\frac{n}{2}}\rho^{\frac{1}{2}}}(\Omega)} + \|X\|_{H^1_\rho(\Omega)}$$

$$\|f\|_{\mathcal{B}_4(\Omega)} = \|f\|_{C^{4,\alpha}_{\phi,\phi^{\frac{n}{2}}\rho^{\frac{1}{2}}}(\Omega)} + \|f\|_{H^2_\rho(\Omega)}.$$

It is clear that these Banach spaces contain the smooth functions with compact supports in Ω, and that $\mathcal{B}_2(\Omega) \subset \mathcal{B}_1(\Omega) \subset \mathcal{B}_0(\Omega)$.

We will frequently use the product norms:

$$\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} = \|\psi\|_{\mathcal{B}_0} + \|V\|_{\mathcal{B}_1}$$

$$\|(h, w)\|_{\mathcal{B}_2 \times \mathcal{B}_2} = \|h\|_{\mathcal{B}_2} + \|w\|_{\mathcal{B}_2}$$

$$\|(f, X)\|_{\mathcal{B}_4 \times \mathcal{B}_3} = \|f\|_{\mathcal{B}_4} + \|X\|_{\mathcal{B}_3}.$$

3. Localized deformation with the no-kernel condition

In this section, we show how to locally deform an initial data set while controlling the dominant energy inequality. To do this, we employ a modified constraint operator to handle the first order change in $|J|_g$ under the perturbation.

The modified map

$$\Phi^W_{(g,\pi)}(\gamma, \tau) = \Phi(\gamma, \tau) + (0, \tfrac{1}{2}\gamma \cdot (\text{div}_g\pi + W))$$

differs from the usual constraint map only by a term of lower order in derivatives and hence has similar analytic properties as the usual constraint map. In particular, we have the following local surjectivity theorem. The proof, which is deferred to Section 5, is a straightforward modification of the proof for the constraint map in [7]. In our proof, we obtain uniform estimates in a neighborhood of an arbitrary initial data set, which sharpen the estimates in [7], for the usual constraint map at a fixed initial data set or in a neighborhood of the flat data.

Recall that we let $\overline{\Omega}$ be a compact manifold with smooth boundary, with Ω the manifold interior of $\overline{\Omega}$. For notational simplicity, we denote $D\Phi^W_{(g,\pi)} = D\Phi^W_{(g,\pi)}|_{(g,\pi)}$ and its formal L^2 adjoint operator by $(D\Phi^W_{(g,\pi)})^*$.

Theorem 3.1. *Let $(g_0, \pi_0) \in C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ be an initial data set, and let $W_0 \in C^{2,\alpha}(\overline{\Omega})$ be a vector field. Suppose that the kernel of $(D\Phi_{(g_0,\pi_0)}^{W_0})^*$ is trivial on Ω. Then there is a neighborhood $\mathcal{U}$ of (g_0, π_0) in $C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$, a neighborhood $\mathcal{W}$ of W_0 in $C^{2,\alpha}(\overline{\Omega})$, and constants $\epsilon > 0$, $C > 0$ such that for $(g, \pi) \in \mathcal{U}$, $W \in \mathcal{W}$, and for $(\psi, V) \in \mathcal{B}_0 \times \mathcal{B}_1$ with $\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq \epsilon$, there is a pair of symmetric tensors $(h, w) \in \mathcal{B}_2 \times \mathcal{B}_2$ with $\|(h, w)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \leq C\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1}$ such that the initial data set $(g + h, \pi + w)$ satisfies*

$$\Phi_{(g,\pi)}^W(g + h, \pi + w) = \Phi_{(g,\pi)}^W(g, \pi) + (2\psi, V),$$

where the weighted norms are taken on Ω with respect to g.

Remark 3.2. Rewriting the identity $\Phi_{(g,\pi)}^W(g + h, \pi + w) = \Phi_{(g,\pi)}^W(g, \pi) + (2\psi, V)$ in terms of the usual constraint map, we see that (h, w) solves

$$\Phi(g + h, \pi + w) = \Phi(g, \pi) + (2\psi, V) - \left(0, \tfrac{1}{2}h \cdot (\mathrm{div}_g\pi + W)\right).$$

The following computational lemma gives the motivation behind our definition of the modified map.

Lemma 3.3. *Let (g, π) be an initial data set. Suppose the initial data set $(\bar{g}, \bar{\pi}) = (g + h, \pi + w)$ satisfies*

$$\Phi_{(g,\pi)}^V(g + h, \pi + w) = \Phi_{(g,\pi)}^V(g, \pi) + (2\psi, V).$$

If $|h|_g \leq 3$, then

$$\bar{\mu} - |\bar{J}|_{\bar{g}} \geq \mu + \psi - |J + V|_g,$$

where $(\bar{\mu}, \bar{J})$ and (μ, J) are the mass and current densities of (g, π) and $(\bar{g}, \bar{\pi})$, respectively.

Proof. Denote by $Y = J + V$. By Remark 3.2,

$$\bar{\mu} = \mu + \psi \quad \text{and} \quad \bar{J} = Y - \frac{1}{2}h \cdot Y.$$

We compute the norm of $\bar{J}$ with respect to $\bar{g}$ below:

$$\begin{aligned}
|\bar{J}|_{\bar{g}}^2 &= (g + h)_{ij}\left(Y^i - \frac{1}{2}(h \cdot Y)^i\right)\left(Y^j - \frac{1}{2}(h \cdot Y)^j\right) \\
&= (g + h)_{ij}\left(Y^iY^j - Y^i(h \cdot Y)^j + \frac{1}{4}(h \cdot Y)^i(h \cdot Y)^j\right) \\
&= |Y|_g^2 + h_{ij}Y^iY^j - g_{ij}Y^i(h \cdot Y)^j - h_{ij}Y^i(h \cdot Y)^j \\
&\quad + \frac{1}{4}|h \cdot Y|_g^2 + \frac{1}{4}h_{ij}(h \cdot Y)^i(h \cdot Y)^j \\
&= |Y|_g^2 - \frac{3}{4}|h \cdot Y|_g^2 + \frac{1}{4}h_{ij}(h \cdot Y)^i(h \cdot Y)^j \\
&\leq |Y|_g^2.
\end{aligned}$$

$\square$

Theorem 1.1 directly follows from Theorem 3.1 (setting $W_0 = 0$) and Lemma 3.3. In particular, by choosing $V = 0$ and $\psi > 0$ on Ω with $\|\psi\|_{\mathcal{B}_2}$ sufficiently small, we have an immediate corollary.

Corollary 3.4. *Let $(\overline{\Omega}, g_0, \pi_0)$ be as in Theorem 3.1, with $W_0 = 0$. Suppose that (g_0, π_0) satisfies the dominant energy condition. Then there exists $(h, w) \in \mathcal{B}_2(\Omega) \times \mathcal{B}_2(\Omega)$ such that the initial data set $(\bar{g}, \bar{\pi}) = (g_0 + h, \pi_0 + w)$ satisfies the strict dominant energy condition $\bar{\mu} > |\bar{J}|_{\bar{g}}$ in Ω.*

Definition 3.5. In the sections which follow, we will consider estimates where some parameter fields (initial data sets, auxiliary cutoff functions) may be allowed to vary. In the case where we can choose a single constant so that the estimate holds for all nearby parameter fields (with respect to specified norms), we say that the constant *depends locally uniformly* on the fields.

We apply Lemma 3.3 to initial data sets which may not satisfy the dominant energy condition but have an error from interpolation. The following computational lemma suggests an appropriate choice of (ψ, V) to absorb the error term.

Lemma 3.6. *Let $(g_1, \pi_1), (g_2, \pi_2) \in C^{2,\alpha}(\overline{\Omega}) \times C^{2,\alpha}(\overline{\Omega})$ be initial data sets. Let $0 \le \chi \le 1$ be a smooth function such that $\chi(1 - \chi)$ is supported on a compact subset of Ω. Denote by $(g, \pi) = \chi(g_1, \pi_1) + (1 - \chi)(g_2, \pi_2)$. Let*

$$(2\psi, V) = -\Phi(g, \pi) + \chi\Phi(g_1, \pi_1) + (1 - \chi)\Phi(g_2, \pi_2) + (2\psi_0, 0)$$

for some $\psi_0 \in \mathcal{B}_0$. Then

$$\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \le C(\|\chi(1 - \chi)\|_{\mathcal{B}_0} + \|\nabla\chi\|_{\mathcal{B}_1})\|(g_1 - g_2, \pi_1 - \pi_2)\|_{C^{2,\alpha} \times C^{2,\alpha}} + \|\psi_0\|_{\mathcal{B}_0}$$

where the constant C depends locally uniformly on $(g_1, \pi_1), (g_2, \pi_2) \in C^{2,\alpha}(\overline{\Omega}) \times C^{2,\alpha}(\overline{\Omega})$ and on $\chi \in C^{2,\alpha}(\overline{\Omega})$, and the norms are taken on $\overline{\Omega}$ (with respect to a fixed metric, say g_1).
  Furthermore,

$$|J + V|_g = |\chi J_1 + (1 - \chi)J_2|_g$$
$$\le \chi|J_1|_{g_1} + (1 - \chi)|J_2|_{g_2} + \chi(1 - \chi)(|g_1 - g_2|_{g_1}|J_1|_{g_1} + |g_1 - g_2|_{g_2}|J_2|_{g_2}),$$

where J, J_1, J_2 are the current densities of $(g, \pi), (g_1, \pi_1), (g_2, \pi_2)$, respectively.

Proof. The estimate of (ψ, V) follows by Lemma A.2. The inequality for $J + V$ follows by direct computation:

(3.1)
$$\begin{aligned} |\chi J_1 + (1 - \chi)J_2|_g &\le \chi|J_1|_g + (1 - \chi)|J_2|_g \\ &\le \chi|J_1|_{g_1} + (1 - \chi)|J_2|_{g_2} \\ &\quad + \chi(1 - \chi)|g_1 - g_2|_{g_1}|J_1|_{g_1} + (1 - \chi)\chi|g_1 - g_2|_{g_2}|J_2|_{g_2}. \end{aligned}$$

$\square$

We now prove Theorem 1.3 for $k = 0$, and the version of higher k is handled later in Section 6.4. The statement can be expressed more precisely as follows.

Theorem 3.7. *Let $(g_0, \pi_0) \in C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ be an initial data set. Suppose that the kernel of $(D\Phi^0_{(g_0, \pi_0)})^*$ is trivial in Ω. Let $0 \le \chi \le 1$ be a smooth function such that $\chi(1 - \chi)$ is supported on a compact subset of Ω. Then there exists a neighborhood $\mathcal{U}$ of (g_0, π_0) in $C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ and $\epsilon > 0$ such that for $(g_1, \pi_1), (g_2, \pi_2) \in \mathcal{U}$, if we set $(g, \pi) = \chi(g_1, \pi_1) + (1 - \chi)(g_2, \pi_2)$, there exists*

$(h, w) \in \mathcal{B}_2(\Omega) \times \mathcal{B}_2(\Omega)$ with $\|(h, w)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \leq \epsilon$ such that the initial data set $(\bar{g}, \bar{\pi}) = (g + h, \pi + w)$ satisfies

$$\bar{\mu} - |\bar{J}|_{\bar{g}} \geq \chi(\mu_1 - |J_1|_{g_1}) + (1 - \chi)(\mu_2 - |J_2|_{g_2}).$$

Proof. Let $\mathcal{U}, \mathcal{W}, \epsilon$ and $C > 1$ be as in Theorem 3.1 (with $W_0 = 0$). Choose (ψ, V) below, as in Lemma 3.6,

$$(2\psi, V) = -\Phi(g, \pi) + \chi\Phi(g_1, \pi_1) + (1 - \chi)\Phi(g_2, \pi_2) + (2\psi_0, 0),$$

where $\psi_0 \in \mathcal{B}_0$ is a smooth function, positive on Ω, with $\|\psi_0\|_{\mathcal{B}_0} \leq \frac{\epsilon}{2C}$. For $\mathcal{U}$ sufficiently small, we have $V \in \mathcal{W}$ and $\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq \frac{\epsilon}{C}$ by Lemma 3.6.

Applying Theorem 3.1 gives (h, w) that satisfies

$$\Phi^V_{(g, \pi)}(g + h, \pi + w) = \Phi^V_{(g, \pi)}(g, \pi) + (2\psi, V).$$

The mass and current densities $(\bar{\mu}, \bar{J})$ of the deformed initial data set $(\bar{g}, \bar{\pi}) = (g + h, \pi + w)$ satisfy, by Lemma 3.3,

$$\bar{\mu} - |\bar{J}|_{\bar{g}} \geq \mu + \psi - |J + V|_g = \chi\mu_1 + (1 - \chi)\mu_2 + \psi_0 - |\chi J_1 + (1 - \chi)J_2|_g.$$

Applying Lemma 3.6 to estimate the last term yields

$$\bar{\mu} - |\bar{J}|_{\bar{g}} \geq \chi(\mu_1 - |J_1|_{g_1}) + (1 - \chi)(\mu_2 - |J_2|_{g_2})$$
$$+ \psi_0 - \chi(1 - \chi)(|g_1 - g_2|_{g_1}|J_1|_{g_1} + |g_1 - g_2|_{g_2}|J_2|_{g_2}).$$

Because $\psi_0 > 0$ on Ω and $\chi(1 - \chi)$ is supported on a compact subset of Ω, we have, by further shrinking $\mathcal{U}$ if necessary,

$$\psi_0 \geq \chi(1 - \chi)(|g_1 - g_2|_{g_1}|J_1|_{g_1} + |g_1 - g_2|_{g_2}|J_2|_{g_2}).$$

$\square$

Remark 3.8. It is clear from the proof that we can obtain the *strict* dominant energy condition in Ω by choosing ψ_0 strictly greater than the error term. Also, the theorem still holds if $\chi(1 - \chi)$ is supported in $\overline{\Omega}$ with an appropriate fall-off rate toward the boundary so that it can be dominated by ψ_0.

As another application, we provide a refined version of the density-type results with the dominant energy condition. A density-type theorem says that given an asymptotically flat initial data set (M, g, π), there exists a sequence of initial data sets (M, g_k, π_k) with the desired asymptotic properties, e.g. harmonic asymptotics (cf. [11, Theorem 18]), such that (g_k, π_k) converges to (g, π) in some appropriate topology. If (g, π) satisfies the no-kernel condition, then the approximate sequence can be made to be identical to (g, π) in a fixed compact set. The reason is that an initial data set interpolating between (g, π) and (g_k, π_k) would utimately satisfy the no-kernel condition in a fixed compact set for k sufficiently large, so one can perform the localized deformation to reimpose the dominant energy condition.

Corollary 3.9. *Let (M, g, π) be a $C^{4,\alpha}_{\mathrm{loc}} \times C^{3,\alpha}_{\mathrm{loc}}$ asymptotically flat initial data set with the dominant energy condition. Let B_R be a coordinate ball of radius R. Suppose the kernel of $(D\Phi^0_{(g_0, \pi_0)})^*$ is trivial in $B_{R_2} \setminus \overline{B_{R_1}}$, $R_2 > R_1$. For any sequence of asymptotically flat initial data sets (g_k, π_k)*

with the dominant energy condition that converges to (g, π) in $C^{4,\alpha}_{\text{loc}}(M) \times C^{3,\alpha}_{\text{loc}}(M)$, there exists k_0 and a sequence $(\bar{g}_k, \bar{\pi}_k) \in C^{2,\alpha}_{\text{loc}}(M) \times C^{2,\alpha}_{\text{loc}}(M)$ with the dominant energy condition for $k \geq k_0$ that converges to (g, π) in $C^{2,\alpha}_{\text{loc}}(M) \times C^{2,\alpha}_{\text{loc}}(M)$ and

$$(\bar{g}_k, \bar{\pi}_k) = (g, \pi) \qquad in \ B_{R_1}$$

$$(\bar{g}_k, \bar{\pi}_k) = (g_k, \pi_k) \qquad in \ M \setminus B_{R_2}.$$

Remark 3.10. Clearly from the proof we can reverse the roles of (g, π) and (g_k, π_k) and construct the converging sequence of initial data sets which is (g, π) outside B_{R_2} and is (g_k, π_k) in B_{R_1}.

4. GLUING IN THE ASYMPTOTICALLY FLAT REGION

In this section, we prove gluing results with the dominant energy condition for initial data sets that are arbitrarily close to the flat data. We focus on the three-dimensional case since our examples of admissible families are in three dimensions (see Section 4.2), but the analysis presented here can be generalized to higher dimensions. Please refer to Appendix B for the definitions of asymptotically flat initial data sets and the ADM integrals.

4.1. **Rescaling.** Given any pair of asymptotically flat initial data sets defined on the exterior of a ball in $\mathbb{R}^3$, they are both close to the flat data on A_R for R large, by the asymptotic flatness. Hence the error from interpolation between those two initial data sets is small enough, for R sufficiently large, so that the localized deformation is applicable. Instead of working on A_R, it is convenient to perform the analysis over a fixed region A_1 via rescaling.

Notation 4.1. Let B_R be an open ball of radius R in $\mathbb{R}^3$. Denote by $A_R = B_{2R} \setminus \overline{B_R}$ the open annulus. Let χ be a smooth cutoff function that is $\chi = 1$ on B_1 and $\chi = 0$ outside B_2 with $\chi(1-\chi)$ supported on a compact subset of A_1. Denote by $\chi_R(y) = \chi(y/R)$ the rescaled cutoff function.

We make some general remarks on rescaling. Let (g, π) be an initial data set on $\mathbb{R}^3 \setminus B$. Here we write π as a $(0, 2)$ tensor with the indices lowered by g. Let $F_R : A_1 \to A_R$ be the diffeomorphism sending $x \mapsto y = Rx$. Define the rescaled initial data set on A_1, via pullback, as follows:

$$g^R = R^{-2} F_R^* g, \qquad \pi^R = R^{-1} F_R^* \pi.$$

Noting $(F_R)_* (\partial_{x^i}) = R \partial_{y^i}$, in coordinates we have

$$g^R(\partial_{x^i}, \partial_{x^j})(x) = R^{-2} g((F_R)_*(\partial_{x^i}), (F_R)_*(\partial_{x^j}))(y) = g(\partial_{y^i}, \partial_{y^j})(y)$$

$$\pi^R(\partial_{x^i}, \partial_{x^j})(x) = R^{-1} \pi((F_R)_*(\partial_{x^i}), (F_R)_*(\partial_{x^j}))(y) = R \pi(\partial_{y^i}, \partial_{y^j})(y).$$

The constraint operator is scaling invariant in the sense that

$$\Phi(g^R, \pi^R) = R^2 F_R^* \Phi(g, \pi).$$

If (g, π) is asymptotically flat at the rate (q, q_0) with respect to y, then the rescaled initial data sets on A_1 satisfy

$$|g^R_{ij} - \delta_{ij}| \leq CR^{-q}, \qquad |\pi^R_{ij}| \leq CR^{-q}, \qquad |\Phi(g^R, \pi^R)| \leq CR^{-1-q_0}$$

and for a multi-index I with $|I| \le k$,

$$|\partial_x^I g_{ij}^R(x)| = |R^{|I|} \partial_y^I g_{ij}(y)| \le CR^{-q}$$
$$|\partial_x^I \pi_{ij}^R(x)| = |R^{1+|I|} \partial_y^I \pi_{ij}(y)| \le CR^{-q},$$

where C depends on $\|g - g_{\mathbb{E}}\|_{C_{-q}^k(A_R)}, \|\pi\|_{C_{-1-q}^k(A_R)}$.

The following computational lemma says that the interpolation between μ and J gives the interpolation between the dominant energy inequality, up to a controllable error term. The lemma will be applied to initial data sets on A_1 that come from rescaling.

Lemma 4.2. *Let* $(g_1, \pi_1), (g_2, \pi_2) \in C^{2,\alpha}(\overline{A_1}) \times C^{2,\alpha}(\overline{A_1})$ *be initial data sets on* $\overline{A_1}$. *Suppose* $C_1 > 0$ *is such that*

$$\|(g_1 - g_2, \pi_1 - \pi_2)\|_{C^{2,\alpha} \times C^{2,\alpha}} \le C_1 R^{-q}.$$

Let $(g, \pi) = \chi(g_1, \pi_1) + (1 - \chi)(g_2, \pi_2)$, *and let*

$$(2\psi, V) = -\Phi(g, \pi) + \chi\Phi(g_1, \pi_1) + (1 - \chi)\Phi(g_2, \pi_2) + (2\psi_0 R^{-1-q_0}, 0),$$

for some $\psi_0 \in \mathcal{B}_0(A_1)$. *Then*

$$(4.1) \qquad \|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \le CR^{-\min(q, 1+q_0)},$$

where C *depends on* $C_1, \|\psi_0\|_{\mathcal{B}_0}, \|\chi(1-\chi)\|_{\mathcal{B}_0}, \|\nabla\chi\|_{\mathcal{B}_1}$ *and locally uniformly on* $(g_1, \pi_1), (g_2, \pi_2) \in C^{2,\alpha}(\overline{A_1}) \times C^{2,\alpha}(\overline{A_1})$ *and* $\chi \in C^{2,\alpha}(\overline{A_1})$. *Note all norms are taken on* A_1, *with respect to a fixed metric, say* g_1.

Furthermore, if $\|J_1\|_{C^0} + \|J_2\|_{C^0} \le C_1 R^{-1-q_0}$, *and if there exists* (h, w) *so that the initial data set* $(\bar{g}, \bar{\pi}) = (g + h, \pi + w)$ *satisfies*

$$\Phi_{(g,\pi)}^V(g + h, \pi + w) = \Phi_{(g,\pi)}^V(g, h) + (2\psi, V),$$

with $|h|_g \le 3$, *then*

$$(4.2) \qquad \begin{aligned} \bar{\mu} - |\bar{J}|_{\bar{g}} &\ge \chi(\mu_1 - |J_1|_{g_1}) + (1 - \chi)(\mu_2 - |J_2|_{g_2}) \\ &\quad + \left(\psi_0 - 2C_1^2 \chi(1-\chi)R^{-q}\right) R^{-1-q_0}, \end{aligned}$$

where $(\bar{\mu}, \bar{J}), (\bar{\mu}_1, \bar{J}_1), (\bar{\mu}_2, \bar{J}_2)$ *are the mass and current densities of* $(\bar{g}, \bar{\pi}), (g_1, \pi_1), (g_2, \pi_2)$, *respectively.*

Proof. By Lemma A.2, we have the following

$$\begin{aligned} \|(2\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} &= \|\chi\Phi(g_1, \pi_1) + (1-\chi)\Phi(g_2, \pi_2) - \Phi(g, \pi)\|_{\mathcal{B}_0 \times \mathcal{B}_1} + \|2\psi_0\|_{\mathcal{B}_0} R^{-1-q_0} \\ &\le CR^{-\min(q, 1+q_0)}. \end{aligned}$$

By Lemma 3.6, we have

$$\begin{aligned} |J + V|_g &\le \chi|J_1|_{g_1} + (1-\chi)|J_2|_{g_2} \\ &\quad + \chi(1-\chi)(|g_1 - g_2|_{g_1}|J_1|_{g_1} + |g_1 - g_2|_{g_2}|J_2|_{g_2}). \end{aligned}$$

Lemma 3.3 gives $\bar{\mu} - |\bar{J}|_{\bar{g}} \ge \mu + \psi - |J + V|_g$, so that inequality (4.2) follows using

$$\bar{\mu} = \chi\mu_1 + (1-\chi)\mu_2 + \psi_0 R^{-1-q_0}.$$

$\square$

For R large, the rescaled initial data set (g^R, π^R) on A_1 is close to the flat data $(g_{\mathbb{E}}, 0)$. Hence the adjoint of the linearized operator at (g^R, π^R) has an *approximate kernel* comprised of the ten-dimensional kernel K of the flat data (see Example 2.4), in the sense that elements of unit norm in K are mapped by the adjoint operator at (g^R, π^R) to elements of small norm. Thus there is no uniform coercivity estimate for the adjoint operator at (g^R, π^R) as R tends to infinity. The following theorem, whose proof is deferred to Section 6, says that the modified constraint operator can be solved transverse to the approximate kernel.

Let K be the kernel of $(D\Phi^{W_0}_{(g_0, \pi_0)})^*$, which is finite-dimensional by Proposition 2.2. Let $\mathcal{U}_0$ be a bounded neighborhood of a Riemannian metric g_0, as in Section 2.4. We fix a smooth bump function ζ supported in $\Omega_0 \subset \Omega$, where the compact set Ω_0 is chosen so that $\rho_g \equiv 1$ on Ω_0 for all $g \in \mathcal{U}_0$. Denote by S_g the $L^2(d\mu_g)$-orthogonal complement of ζK. Let $\Pi_g : \mathcal{B}_0 \times \mathcal{B}_1 \to (\mathcal{B}_0 \times \mathcal{B}_1) \cap S_g$ be the $L^2(d\mu_g)$-orthogonal projection.

Theorem 4.3. *Let $(g_0, \pi_0) \in C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ be an initial data set, and let $W_0 \in C^{2,\alpha}(\overline{\Omega})$ be a vector field. There is a neighborhood $\mathcal{U}$ of (g_0, π_0) in $C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$, a neighborhood $\mathcal{W}$ of W_0 in $C^{2,\alpha}(\overline{\Omega})$, and constants $\epsilon > 0$, $C > 0$ such that for $(g, \pi) \in \mathcal{U}$, $W \in \mathcal{W}$, and $(\psi, V) \in \mathcal{B}_0(\Omega) \times \mathcal{B}_1(\Omega)$ with $\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq \epsilon$, there is a pair of symmetric tensors $(h, w) \in \mathcal{B}_2(\Omega) \times \mathcal{B}_2(\Omega)$ such that the initial data set $(g + h, \pi + w)$ satisfies*

$$\Pi_{g_0} \circ \Phi^W_{(g,\pi)}(g + h, \pi + w) = \Pi_{g_0} \circ \Phi^W_{(g,\pi)}(g, \pi) + \Pi_{g_0}(2\psi, V)$$

with

$$\|(h, w)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \leq C\|\Pi_{g_0}(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1}.$$

Here the weighted norms are all taken on Ω with respect to g. The solution $(h, w) \in \mathcal{B}_2 \times \mathcal{B}_2$ depends continuously on $(\psi, V) \in \mathcal{B}_0 \times \mathcal{B}_1$.

We now apply this in the setting we study here, interpolating μ and J.

Proposition 4.4. *Let (g_1, π_1) and $(g_2, \pi_2) \in C^{4,\alpha}_{\text{loc}} \times C^{3,\alpha}_{\text{loc}}$ be asymptotically flat initial data sets at the rate (q, q_0) on $\mathbb{R}^3 \setminus B$. Consider the corresponding rescaled initial data sets $(g_1^R, \pi_1^R), (g_2^R, \pi_2^R)$ on A_1. Define the initial data set*

$$(\gamma^R, \tau^R) = \chi(g_1^R, \pi_1^R) + (1 - \chi)(g_2^R, \pi_2^R).$$

Let

$$(2\psi^R, V^R) = -\Phi(\gamma^R, \tau^R) + \chi\Phi(g_1^R, \pi_1^R) + (1 - \chi)\Phi(g_2^R, \pi_2^R) + (2\psi_0 R^{-1-q_0}, 0)$$

for some $\psi_0 \in \mathcal{B}_0$. There is $R_0 > 0$ and $C > 0$, depending only on $g_{\mathbb{E}}, \chi, \|\psi_0\|_{\mathcal{B}_0}, \|(g_1 - g_{\mathbb{E}}, \pi_1)\|_{C^{2,\alpha}_{-q} \times C^{2,\alpha}_{-1-q}}, \|(g_2 - g_{\mathbb{E}}, \pi_2)\|_{C^{2,\alpha}_{-q} \times C^{2,\alpha}_{-1-q}}$, such that for each $R > R_0$, there exists a pair of symmetric tensors $(h^R, w^R) \in \mathcal{B}_2(A_1) \times \mathcal{B}_2(A_1)$ such that the initial data set $(\gamma^R + h^R, \tau^R + w^R)$ satisfies

$$\Pi_{g_{\mathbb{E}}} \circ \Phi^{V^R}_{(g_{\mathbb{E}}, 0)}(\gamma^R + h^R, \tau^R + w^R) = \Pi_{g_{\mathbb{E}}} \circ \Phi^{V^R}_{(g_{\mathbb{E}}, 0)}(\gamma^R, \tau^R) + \Pi_{g_{\mathbb{E}}}(2\psi^R, V^R)$$

with the estimate

(4.3) $$\|(h^R, w^R)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \leq CR^{-\min(q, 1+q_0)}.$$

Proof. Apply Theorem 4.3 with $(g_0, \pi_0) = (g_{\mathbb{E}}, 0)$ and $W_0 = 0$. By Lemma 4.2 and Lemma 6.1, we have for R sufficiently large $V^R \in \mathcal{W}$ and

$$\|\Pi_{g_0}(\psi^R, V^R)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \le CR^{-\min(q, 1+q_0)} < \epsilon.$$

Therefore we can apply Theorem 4.3 to solve for (h^R, w^R) with the estimate (4.3). $\qquad \square$

4.2. **An admissible family and gluing.** As observed in [7], the 10-dimensional approximate kernel in the asymptotically gluing is ultimately connected to the 10-dimensional parameter space of an *admissible family*. An improved rescaling argument to handle a more general situation is discussed in [2]. However, since neither paper states in sufficiently explicit terms the requisite conditions on an admissible family that can be used to model the asymptotics of initial data sets by asymptotic gluing, we proceed to do so now.

In order to handle initial data sets whose center of mass and angular momentum integrals may not converge, we define the center of mass and angular momentum integrals of (g, π) at the finite radius R as

$$\mathcal{C}_i^R = \frac{1}{16\pi} B_{(g,\pi)}^R(x^i, 0), \qquad \mathcal{J}_k^R = \frac{1}{8\pi} B_{(g,\pi)}^R(0, x \times \frac{\partial}{\partial x^k}).$$

It is clear that $\mathcal{C}^R, \mathcal{J}^R$ are continuous in R. If (g, π) is asymptotically flat at the rate $q = q_0 = 1$, then $(\mathcal{C}^R, \mathcal{J}^R) = O(\log R)$ as R tends to infinity. For other values of q, q_0, $(\mathcal{C}^R, \mathcal{J}^R) = O(R^{1-\min\{1, q_0, 2q-1\}})$. Note that if (g, π) satisfies the Regge-Teitelboim conditions, then $(\mathcal{C}^R, \mathcal{J}^R)$ converges to a pair of vectors (see, e.g. [7, 8, 14, 15]). We denote $a = \min\{1, q_0, 2q - 1\} \in (0, 1]$.

Definition 4.5. Let $(g, \pi) \in C_{\text{loc}}^{k+4,\alpha} \times C_{\text{loc}}^{k+3,\alpha}$ be an asymptotically flat initial data set on $\mathbb{R}^3 \setminus B_{R_0}$, at the rate (q, q_0). Let $\mathcal{S} = \{(g^\theta, \pi^\theta)\}_{\theta \in \Theta}$ be a family of $C_{\text{loc}}^{k+4,\alpha} \times C_{\text{loc}}^{k+3,\alpha}$ asymptotically flat initial data sets defined on $\mathbb{R}^3 \setminus B_{R_0}$, where the components of $\theta = (E^\theta, P^\theta, \mathcal{C}^\theta, \mathcal{J}^\theta)$ are the ADM energy, linear momentum, center of mass, and angular momentum of (g^θ, π^θ). The family $\mathcal{S}$ is said to be an *admissible family for* (g, π) if the following properties hold, with respect to a fixed asymptotically flat coordinate chart:

(1) The map $(g^\theta, \pi^\theta) \mapsto \theta \in \Theta = \Theta_1 \times \Theta_2$ is a homeomorphism where $\Theta_1 \subset \mathbb{R}^4$, $\Theta_2 \subset \mathbb{R}^6$ are open sets such that Θ_1 contains (E, P) of (g, π) and $\Theta_2 = \cup_{R \ge R_0} \Theta_2^R$ where Θ_2^R is the ball centered at $(\mathcal{C}^R, \mathcal{J}^R)$ of radius $R^{1-a} \log R$.

(2) (g^θ, π^θ) satisfies the following uniformity conditions: there is a constant $\kappa > 0$ such that for $\theta \in \Theta_1 \times \Theta_2^R$ and for all $R > R_0$,

(4.4)
$$\|(g^\theta - g_{\mathbb{E}}, \pi^\theta)\|_{C_{-q}^2(\mathbb{R}^3 \setminus B_R) \times C_{-1-q}^1(\mathbb{R}^3 \setminus B_R)} < \kappa$$
$$\|(\mu^\theta, J^\theta)\|_{C_{-3-q_0}^0(\mathbb{R}^3 \setminus B_R)} < \kappa$$

and

(4.5)
$$|B_{(g^\theta, \pi^\theta)}^R(x^k, 0) - 16\pi \mathcal{C}_k^\theta| \le \kappa |\theta|^2 R^{-1}$$
$$|B_{(g^\theta, \pi^\theta)}^R(0, x \times \frac{\partial}{\partial x^k}) - 8\pi \mathcal{J}_k^\theta| \le \kappa |\theta|^2 R^{-1}.$$

Remark 4.6. The definition of the parameter set $\Theta_1 \times \Theta_2^R$ is set up in a way so that θ can be large enough to account for the error terms from $(\mathcal{C}^R, \mathcal{J}^R)$ and from the right hand side of (4.5), which may compete with largeness of θ. This subtle balance shows up in Lemma 4.8 below in the degree

argument. We remark that in Definition 4.5, the term $|\theta|^2$ appearing in (4.5) can be replaced by $|\theta|^{\kappa_2}$ for $\kappa_2 > 0$. If $q = q_0 = 1$, the same parameter set $\Theta_1 \times \Theta_2^R$ is still valid. For other values of q, q_0, the radius of Θ_2^R needs to be modified accordingly, depending on κ_2. This can be done by tracking the exponents in the proof of Lemma 4.8.

Example 4.7. Let (g, π) be an asymptotically flat initial data set with $E > |P|$. Let $\Theta = \Theta_1 \times \mathbb{R}^6$, where $(E, P) \in \Theta_1$ is a precompact open subset of $\{(a, b) \in \mathbb{R} \times \mathbb{R}^3 : a > |b|\}$. There is an admissible *vacuum* family $\mathcal{S}_{\text{Kerr}} = \{(g^\theta, \pi^\theta)\}_{\theta \in \Theta}$ for (g, π) obtained from the Kerr spacetime. Condition (1) is shown in [3, Appendix F]. Inequalities (4.4) in Condition (2) follow from the asymptotic expansions of (g^θ, π^θ) and compactness of Θ_1. Inequalities (4.5) follows from a standard argument: by the divergence theorem, $B^R_{(g^\theta, \pi^\theta)}(x^k, 0) - 16\pi \mathcal{C}^\theta_k = \int_{M \setminus B_R} x^k \sum_{i,j} (g^\theta_{ij,ij} - g^\theta_{ii,jj}) \, dx$. The estimate then follows by using the constraint equations to rewrite the integrand and estimate the resulting terms in an obvious way. The inequality for the angular momentum can be obtained in the same fashion.

The same argument would also allow us to find an admissible *non-vacuum* family $\mathcal{S}_{\text{Kerr-Newman}} = \{(g^\theta, \pi^\theta)\}_{\theta \in \Theta}$ for (g, π) obtained from the Kerr-Newman spacetime for each fixed pair of electric and magnetic charges. $\qquad\square$

Admissible families will be used in an asymptotic gluing construction in Theorem 4.9 below, and we want to highlight how the assumptions in Definition 4.5 will appear. The condition on the parameter set Θ_2^R is to handle the scaling in the center of mass and angular momentum components of a map $\mathcal{I}^R$, in order to be able to apply a degree argument as used in the following technical lemma. The uniformity assumptions (4.4) and (4.5) will be used in the proof of the theorem to establish the desired estimate of the error term $\mathcal{I}_0^R$, as appears in the following lemma.

Lemma 4.8. *Let $\Theta_1 \times \Theta_2^R$ be as in Definition 4.5. Let $\mathcal{I}^R : \Theta_1 \times \Theta_2^R \to \mathbb{R}^{10}$ be a family of continuous maps, continuously parametrized by $R \in (R_0, \infty)$. Suppose that there is a constant $C > 0$ such that for any $\theta = (E^\theta, P^\theta, \mathcal{C}^\theta, \mathcal{J}^\theta) \in \Theta_1 \times \Theta_2^R$ and any $R \geq R_0$,*

$$\mathcal{I}^R(\theta) = (E^\theta - E, P^\theta - P, R^{-1}(\mathcal{C}^\theta - \mathcal{C}^R), R^{-1}(\mathcal{J}^\theta - \mathcal{J}^R)) + \mathcal{I}_0^R(\theta),$$

where

$$|\mathcal{C}^R| \leq C \log R$$
$$|\mathcal{J}^R| \leq C \log R$$
$$|\mathcal{I}_0^R(\theta)| \leq C(R^{-a}(\log R)^{\frac{1}{2}} + |\theta|^2 R^{-2}).$$

Then for each R sufficiently large, there exists $\theta \in \Theta_1 \times \Theta_2^R$ such that $\mathcal{I}^R(\theta) = 0$.

Proof. Let $\theta_0^R = (E, P, \mathcal{C}^R, \mathcal{J}^R)$. We first translate and rescale $\Theta_1 \times \Theta_2^R$, in order to employ a degree argument over sets centered around the same point. Suppose Θ_1 contains the ball centered at (E, P) of radius ϵ for some $\epsilon > 0$. Let $T^R : B_\epsilon(0) \times B_{R^{-a} \log R}(0) \to \Theta_1 \times \Theta_2^R$ be given by

$$T^R(v, w) = (v, Rw) + \theta_0^R.$$

The composition map satisfies

$$\mathcal{I}^R \circ T^R(v, w) = (v, w) + \mathcal{I}_0^R \circ T^R(v, w).$$

Hence, for R sufficiently large, we have

$$|\mathcal{I}^R \circ T^R(v,w) - (v,w)| \leq C(R^{-a}(\log R)^{\frac{1}{2}} + |T^R(v,w)|^2 R^{-2}) < R^{-a} \log R,$$

where we use that $\theta = T^R(v,w) \in \Theta_1 \times \Theta_2^R$ and hence $|\theta| \leq C|R^{1-a} \log R|$. The rest of proof follows from a standard degree argument (see, e.g. [16, Lemma 5.2]).

$\square$

We now restate Theorem 1.4 for $k = 0$ in a more specific form and give a proof. The version for higher k is discussed in Section 6.4

Theorem 4.9. *Let $(g,\pi) \in C_{\mathrm{loc}}^{4,\alpha} \times C_{\mathrm{loc}}^{3,\alpha}$ be an asymptotically flat initial data set at the rate (q, q_0) on $\mathbb{R}^3 \setminus B$ with the ADM energy-momentum (E, P). Let $\epsilon > 0$. There is a constant $R_0 > 0$ such that for $R > R_0$, there is an initial data set $(\bar{g}, \bar{\pi}) \in C_{\mathrm{loc}}^{2,\alpha} \times C_{\mathrm{loc}}^{2,\alpha}$ such that*

$$(\bar{g}, \bar{\pi}) = (g, \pi) \quad in \ B_R$$
$$(\bar{g}, \bar{\pi}) = (g^\theta, \pi^\theta) \quad in \ M \setminus B_{2R}$$

for some (g^θ, π^θ) in the admissible family for (g, π), and $(\bar{g}, \bar{\pi})$ satisfies the dominant energy condition

$$\bar{\mu} - |\bar{J}|_{\bar{g}} \geq \chi_R(\mu - |J|_g) + (1 - \chi_R)(\mu^\theta - |J^\theta|_{g^\theta})$$

with the strictly larger ADM energy $E^\theta > E$ and

$$|P^\theta - P| < E^\theta - E < \epsilon.$$

Proof. We prove the case $q = q_0 = 1$, as the proof for other values of q, q_0 is similar. Given $R \geq 1$ sufficiently large and $\theta \in \Theta$, we apply Proposition 4.4 with $(g_1, \pi_1) = (g, \pi)$, $(g_2, \pi_2) = (g^\theta, \pi^\theta)$ and $(\gamma^R, \tau^R) = \chi(g^R, \pi^R) + (1 - \chi)((g^\theta)^R, (\pi^\theta)^R)$. Let $\psi_0 \in \mathcal{B}_0(A_1)$ be a fixed positive function in A_1. Define (ψ^R, V^R), as in Proposition 4.4, by

$$(2\psi^R, V^R) = -\Phi(\gamma^R, \tau^R) + \chi\Phi(g^R, \pi^R) + (1 - \chi)\Phi((g^\theta)^R, (\pi^\theta)^R) + (2\psi_0 R^{-2}, 0).$$

There exists $(h^R, w^R) \in \mathcal{B}_2(A_1) \times \mathcal{B}_2(A_1)$ (with the superscript θ omitted) such that $(\bar{g}^R, \bar{\pi}^R) = (\gamma^R + h^R, \tau^R + w^R)$ solves the following projected problem, for some radial bump function ζ supported in A_1,

$$\Phi_{(\gamma^R, \tau^R)}^{V^R}(\bar{g}^R, \bar{\pi}^R) - \Phi_{(\gamma^R, \tau^R)}^{V^R}(\gamma^R, \tau^R) - (2\psi^R, V^R) \in \zeta K,$$

where K is the kernel at the flat data (see Example 2.4). Fix a constant $\lambda > 0$. We show that, for each R sufficiently large, there exists $\theta \in \Theta_1 \times \Theta_2^R$ such that

$$\mathcal{E}^R(\theta) = \Phi_{(\gamma^R, \tau^R)}^{V^R}(\bar{g}^R, \bar{\pi}^R) - \Phi_{(\gamma^R, \tau^R)}^{V^R}(\gamma^R, \tau^R) - (2\psi^R, V^R) - (\lambda\zeta R^{-2}(\log R)^{\frac{1}{2}}, 0) = 0.$$

Note that (h^R, w^R) is independent of λ. We remark that the additional term $(\lambda\zeta R^{-2}(\log R)^{\frac{1}{2}}, 0)$ from the kernel will help to bring up the ADM energy and also keep the dominant energy condition.

It suffices to show that $\mathcal{E}^R(\theta)$ is $L^2(dx)$-orthogonal to K because $K^\perp$ is transverse to ζK. Consider the L^2 projection $\mathcal{I}^R(\theta) : \Theta_1 \times \Theta_2^R \to \mathbb{R}^{10}$ onto K that sends θ to $(e, \mathbf{p}, \mathbf{c}, \mathbf{j})$ where

$\mathbf{p} = (p_1, p_2, p_3)$, $\mathbf{c} = (c_1, c_2, c_3)$, $\mathbf{j} = (j_1, j_2, j_3)$:

$$e = \frac{R}{16\pi} \int_{A_1} \mathcal{E}^R(\theta) \cdot (1, 0) \, dx$$

$$p_i = \frac{R}{8\pi} \int_{A_1} \mathcal{E}^R(\theta) \cdot (0, \frac{\partial}{\partial x^i}) \, dx$$

$$c_k = \frac{R}{16\pi} \int_{A_1} \mathcal{E}^R(\theta) \cdot (x^k, 0) \, dx$$

$$j_\ell = \frac{R}{8\pi} \int_{A_1} \mathcal{E}^R(\theta) \cdot (0, x \times \frac{\partial}{\partial x^\ell}) \, dx.$$

The map $\mathcal{I}^R$ is continuous because (h^R, w^R) depends continuously on (ψ^R, V^R), which is continuous in θ by definition.

Expressing $\mathcal{E}^R(\theta)$ in terms of the usual constraint map, we have

$$\mathcal{E}^R(\theta) = \Phi(\bar{g}^R, \bar{\pi}^R) + \mathcal{E}_2^R(\theta) - (\lambda \zeta R^{-2} (\log R)^{\frac{1}{2}}, 0),$$

where

$$\mathcal{E}_2^R(\theta) = -\chi \Phi(g^R, \pi^R) - (1 - \chi) \Phi((g^\theta)^R, (\pi^\theta)^R) - (2\psi_0 R^{-2}, \tfrac{1}{2} h^R \cdot (\operatorname{div}_{\gamma^R} \tau^R + V^R)).$$

The L^2 projection of $\Phi(\bar{g}^R, \bar{\pi}^R)$ is handled exactly as in the vacuum case; estimates of the projection are included in Lemma B.1 for which we employ the uniformity conditions, in particular (4.5). The L^2 projection of $\mathcal{E}_2^R(\theta)$ is of lower order because by using (4.4) and the estimates for (h^R, w^R) in Proposition 4.4, we obtain $\left| \mathcal{E}_2^R(\theta) \right|_{g_{\mathbb{E}}} \leq C R^{-2}$. Because ζ is radial, the L^2 projection of the last term $-(\lambda \zeta R^{-2} (\log R)^{\frac{1}{2}}, 0)$ is non-zero only onto the kernel element $(1, 0)$, and we find

$$\int_{A_1} \lambda \zeta R^{-2} (\log R)^{\frac{1}{2}} \, dx = 16\pi \tilde{\lambda} R^{-2} (\log R)^{\frac{1}{2}} > 0,$$

where $\tilde{\lambda} := (16\pi)^{-1} \int_{A_1} \lambda \zeta \, dx$. We then obtain, together with Lemma B.1, for $\theta \in \Theta_1 \times \Theta_2^R$,

$$\mathcal{I}^R(\theta) = (E^\theta - E, P^\theta - P, R^{-1}(\mathcal{C}^\theta - \mathcal{C}^R), R^{-1}(\mathcal{J}^\theta - \mathcal{J}^R)) + \mathcal{I}_1^R(\theta) - (\tilde{\lambda} R^{-1} (\log R)^{\frac{1}{2}}, 0)$$

where

$$|\mathcal{I}_1^R(\theta)| \leq C(R^{-1} + |\theta|^2 R^{-2}).$$

By Lemma 4.8 with $\mathcal{I}_0^R(\theta) = \mathcal{I}_1^R(\theta) - (\tilde{\lambda} R^{-1} (\log R)^{\frac{1}{2}}, 0)$, there is $\theta \in \Theta_1 \times \Theta_2^R$ such that $\mathcal{I}^R(\theta) = 0$ for R sufficiently large. Because the term $\tilde{\lambda} R^{-1} (\log R)^{\frac{1}{2}}$ dominates other error terms in the identity $\mathcal{I}^R(\theta) = 0$ with a favorable sign, we obtain $E < E^\theta$ and $|P^\theta - P| < E^\theta - E < \epsilon$ for R large.

Last, we verify the dominant energy condition for $(\bar{g}^R, \bar{\pi}^R)$ on A_1. We have solved

$$\Phi_{(\gamma^R, \tau^R)}^{V^R}(\bar{g}^R, \bar{\pi}^R) = \Phi_{(\gamma^R, \tau^R)}^{V^R}(\gamma^R, \tau^R) + (2\psi^R + \lambda \zeta R^{-2} (\log R)^{\frac{1}{2}}, V^R).$$

Because $\lambda > 0$, by Lemma 4.2, it suffices to show that on A_1

$$\psi_0 \geq 2 C_1^2 \chi (1 - \chi) R^{-1}.$$

Since ψ_0 is positive in A_1 and $\chi(1 - \chi)$ is supported on a compact subset of A_1, the above inequality holds for R sufficiently large.

$\square$

Another application is a version of the N-body initial data construction of [2]. Once the gluing construction Theorem 4.9 is achieved, the proof of the N-body construction with the dominant energy condition is essentially identical to [2] for the vacuum case, so we omit the proof.

Theorem 4.10. *Let* (M_k, g^k, π^k), $k = 1, \ldots, N$, *be three-dimensional asymptotically flat initial data sets that satisfy the dominant energy condition and have time-like ADM energy-momentum vector* (E_k, P_k). *Let* $U_k \subset M_k$ *be compact subsets. There is* $\epsilon_0 > 0$ *so that for* $0 < \epsilon < \epsilon_0$, *there is an initial data set* (M, g, π) *with the dominant energy condition which contains a region* U *isometric to* $\bigcup_{k=1}^{N} (U_k, g^k)$, *and the distances between distinct* U_k *are* $O(\epsilon^{-1})$, *and* $(M \setminus U, g, \pi)$ *has one asymptotically flat end identical to a Kerr initial data set, with ADM energy-momentum* (E, P) *satisfying* $\left| E - \sum_{k=1}^{N} E_k \right| < \epsilon$ *and* $\left| P - \sum_{k=1}^{N} P_k \right| < \epsilon$.

5. DEFORMING THE MODIFIED CONSTRAINT MAP

In this section, we prove Theorem 3.1. The argument is similar to that for the constraint map [7, Section 4.2]. We first solve the linearized equation via a variational approach with estimates and then use iteration for the nonlinear problem. We also pay special care to ensure uniform estimates. Throughout the section, the weighted Sobolev and Hölder norms are all taken on Ω.

5.1. **The linearized equation.** The goal is to solve the following linearized system for given (ψ, V)

$$D\Phi^W_{(g,\pi)}|_{(g,\pi)}(h, w) = (\psi, V).$$

We recall that to simplify notation, we let $D\Phi^W_{(g,\pi)} = D\Phi^W_{(g,\pi)}|_{(g,\pi)}$, the linearization at (g, π). Consider the functional $\mathcal{G}$ defined by

$$\mathcal{G}(f, X) = \int_\Omega \left(\frac{1}{2} \rho_g \left| (D\Phi^W_{(g,\pi)})^*(f, X) \right|^2_g - (\psi, V) \cdot_g (f, X) \right) d\mu_g.$$

Clearly the functional is convex. To derive the key coercivity property, we need some basic estimates. Recall $L^*_g f = -(\Delta_g f)g + \text{Hess}_g f - f \text{Ric}(g)$.

Lemma 5.1. *Let* $g_0 \in C^2(\overline{\Omega})$. *There is a* $C^2(\overline{\Omega})$ *neighborhood* $\mathcal{U}_0$ *of* g_0 *and a constant* $C > 0$ *such that for* $g \in \mathcal{U}_0$ *and for* $(f, X) \in H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)$,

$$\|f\|_{H^2_{\rho_g}} \leq C \left(\|L^*_g f\|_{L^2_{\rho_g}} + \|f\|_{L^2_{\rho_g}} \right) \tag{5.1}$$

$$\|X\|_{H^1_{\rho_g}} \leq C \left(\|\mathcal{D}_g X\|_{L^2_{\rho_g}} + \|X\|_{L^2_{\rho_g}} \right). \tag{5.2}$$

Proof. The uniform dependence of the constant C on the metric in (5.1) follows from the proof of [5, Proposition 3.1-3.2, Theorem 3]; in particular, one uses [5, Equation (13)] and the co-area formula as in the end of the proof of [5, Theorem 3], cf. [6, Proposition 3.1 and Remark 3.6].

For (5.2), it suffices to prove that there is a uniform constant C such that

$$\int_\Omega |\nabla_g X|^2 \rho_g \, d\mu_g \leq C \left(\int_\Omega |\mathcal{D}_g X|^2 \rho_g \, d\mu_g + \int_\Omega |X|^2 \rho_g \, d\mu_g \right).$$

The estimate for a fixed metric is obtained in [7, Proof of Lemma 4.1]. In their proof, an analysis of the arguments in [7, pp. 201-202 and the first paragraph of p. 203] shows that there is a neighborhood $\mathcal{U}_0$ so that for $X \in H^1_{\rho_g}(\Omega)$

(5.3)
$$\int_\Omega |\nabla_g X|^2 \rho_g \, d\mu_g$$
$$\leq C_0 \left(N^2 \int_\Omega |\mathcal{D}_g X|^2 \rho_g \, d\mu_g + \int_\Omega |X|^2 \rho_g \, d\mu_g + \int_\Omega |X|^2 d_g^{-4} \rho_g \, d\mu_g \right),$$

where C_0 is independent of $g \in \mathcal{U}_0$, and N. It is important that the coefficient of the last integral above does not depend on N. (Note that we use the exponential weight function ρ_g, instead of the polynomial weight function, so the power of the distance function in the last term is different from [7, p. 203].) The last term in the right hand side was handled by an indirect argument in [7]. Here we apply (2.14) (with $j = k = 1$) in the proof of Corollary 2.9 to the last term and derive

$$\int_\Omega |X|^2 d_g^{-4} \rho_g d\mu_g \leq \frac{4}{N} \left(\int_\Omega |\nabla_g X|^2 \rho_g \, d\mu_g + C \int_\Omega |X|^2 \rho_g \, d\mu_g \right).$$

The first term in the right hand side can be absorbed into the left hand side of (5.3) for our choice of N in (2.11).

$\square$

Lemma 5.2. *Let $(g_0, \pi_0) \in C^2(\overline{\Omega}) \times C^1(\overline{\Omega})$ be an initial data set, and let $W_0 \in C^0(\overline{\Omega})$ be a vector field. There is a neighborhood $\mathcal{U}$ of (g_0, π_0) in $C^2(\overline{\Omega}) \times C^1(\overline{\Omega})$, a neighborhood $\mathcal{W}$ of W_0 in $C^0(\overline{\Omega})$, and a constant $C > 0$ such that for $(g, \pi) \in \mathcal{U}$, $W \in \mathcal{W}$, and for $(f, X) \in H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)$, the following estimate holds:*

(5.4)
$$\|(f, X)\|_{H^2_{\rho_g} \times H^1_{\rho_g}} \leq C \left(\|(D\Phi^W_{(g,\pi)})^*(f, X)\|_{L^2_{\rho_g}} + \|(f, X)\|_{L^2_{\rho_g}} \right).$$

Proof. The terms in $(D\Phi^W_{(g,\pi)})^*(f, X)$ that have the highest order of derivatives are $L_g^* f$ and $\mathcal{D}_g X$. Hence,

$$\|L_g^* f\|_{L^2_{\rho_g}} \leq C \left(\|(D\Phi^W_{(g,\pi)})^*(f, X)\|_{L^2_{\rho_g}} + \|f\|_{L^2_{\rho_g}} + \|X\|_{H^1_{\rho_g}} \right)$$
$$\|\mathcal{D}_g X\|_{L^2_{\rho_g}} \leq C \left(\|(D\Phi^W_{(g,\pi)})^*(f, X)\|_{L^2_{\rho_g}} + \|(f, X)\|_{L^2_{\rho_g}} \right).$$

The desired inequalities follow from Lemma 5.1.

$\square$

Theorem 5.3. *Let $(g_0, \pi_0) \in C^2(\overline{\Omega}) \times C^1(\overline{\Omega})$ be an initial data set, and let $W_0 \in C^0(\overline{\Omega})$ be a vector field. Suppose that the kernel of $(D\Phi^{W_0}_{(g_0,\pi_0)})^*$ is trivial on Ω. Then there is a neighborhood $\mathcal{U}$ of (g_0, π_0) in $C^2(\overline{\Omega}) \times C^1(\overline{\Omega})$, a neighborhood $\mathcal{W}$ of W_0 in $C^0(\overline{\Omega})$, and a constant $C > 0$ such that for $(g, \pi) \in \mathcal{U}$, $W \in \mathcal{W}$, and $(f, X) \in H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)$, the following estimate holds:*

(5.5)
$$\|(f, X)\|_{H^2_{\rho_g} \times H^1_{\rho_g}} \leq C \|(D\Phi^W_{(g,\pi)})^*(f, X)\|_{L^2_{\rho_g}}.$$

Proof. The proof is a standard argument, but we include it for the reader's convenience. Let $\mathcal{U}, \mathcal{W}$ be from Lemma 5.2. By shrinking the neighborhoods if necessary, we may assume that $(D\Phi^W_{(g,\pi)})^*$ has a trivial kernel on Ω for $(g, \pi) \in \mathcal{U}$, $W \in \mathcal{W}$. Suppose there were sequences $(g_k, \pi_k) \to (g, \pi)$ in $\mathcal{U}$, $W_k \to W$ in $\mathcal{W}$, and $(f_k, X_k) \in H^2_{\rho_k} \times H^1_{\rho_k}$, for which

$$\|(f_k, X_k)\|_{H^2_{\rho_k} \times H^1_{\rho_k}} = 1$$

but with

$$\|(D\Phi^{W_k}_{(g_k,\pi_k)})^*(f_k,X_k)\|_{L^2_{\rho_k}} \to 0,$$

where $\rho_k = \rho_{g_k}$. By Proposition 2.10, the sequence $(f_k,X_k)\rho_k^{\frac{1}{2}}$ is bounded in $H^2(\Omega) \times H^1(\Omega)$. By the Rellich theorem, upon choosing a suitable subsequence and relabeling, there is $(f,X) \in H^2_{\text{loc}}(\Omega) \times H^1_{\text{loc}}(\Omega)$ such that

$$\|(f_k,X_k)\rho_k^{\frac{1}{2}} - (f,X)\rho_g^{\frac{1}{2}}\|_{H^1 \times L^2} \to 0.$$

(Since the $H^1(\Omega) \times L^2(\Omega)$ norms are equivalent for g in a neighborhood of g_0, the above convergence can be taken, for example, with respect to g.) Because ρ_k has a uniform positive lower bound on any given compact subset of Ω, it implies that (f_k,X_k) converges in L^2_{loc} to (f,X) and $(D\Phi^W_{(g,\pi)})^*(f,X) = 0$ weakly and hence (f,X) is in the kernel of $(D\Phi^W_{(g,\pi)})^*$. Thus, $(f,X) = (0,0)$ because the kernel of $(D\Phi^W_{(g,\pi)})^*$ is trivial. Thus, $(f_k,X_k)\rho_k^{\frac{1}{2}}$ converges to zero in $H^1(\Omega) \times L^2(\Omega)$. Lemma 5.2 implies $\|(f_k,X_k)\|_{H^2_{\rho_k} \times H^1_{\rho_k}} \to 0$ and contradicts our assumption. $\qquad\square$

Remark 5.4. The above theorem is essentially the only place we need to assume that the kernel is trivial. If the kernel were not trivial, the coercivity estimate would still hold transverse to the kernel. More precisely, let S be a complete linear subspace of $H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)$ such that $S \cap K = \{0\}$ where $K = \ker(D\Phi^{W_0}_{(g_0,\pi_0)})^*$. Then the above argument implies that the coercivity estimate (5.5) holds for $(f,X) \in S$. The only difference in the proof is that after showing that the sequence (f_k,X_k) converges to $(f,X) \in \ker(D\Phi^W_{(g,\pi)})^* =: K'$, one uses that K' is also transverse to S for sufficiently small neighborhoods $\mathcal{U}, \mathcal{W}$ to conclude $(f,X) = (0,0)$.

We now apply the coercivity estimate to obtain the variational solution of the linearized equation.

Theorem 5.5. *Let $(g_0,\pi_0) \in C^4(\overline{\Omega}) \times C^3(\overline{\Omega})$ be an initial data set, and let $W_0 \in C^0(\overline{\Omega})$ be a vector field. Suppose that the kernel of $(D\Phi^{W_0}_{(g_0,\pi_0)})^*$ is trivial on Ω. Let the neighborhoods $\mathcal{U}, \mathcal{W}$ and the constant C be as in Theorem 5.3. Then for $(g,\pi) \in \mathcal{U}$, $W \in \mathcal{W}$, and $(\psi,V) \in L^2_{\rho_g^{-1}}(\Omega) \times L^2_{\rho_g^{-1}}(\Omega)$, the functional $\mathcal{G}(f,X)$ has a global minimizer $(f,X) \in H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)$. Furthermore, (f,X) is the unique weak solution of the linear system*

$$(5.6) \qquad D\Phi^W_{(g,\pi)} \circ \rho_g(D\Phi^W_{(g,\pi)})^*(f,X) = (\psi,V)$$

and satisfies the estimate

$$(5.7) \qquad \|(f,X)\|_{H^2_{\rho_g} \times H^1_{\rho_g}} \le 2C\|(\psi,V)\|_{L^2_{\rho_g^{-1}} \times L^2_{\rho_g^{-1}}}.$$

Proof. Theorem 5.3 implies that the infimum of the functional $\mathcal{G}$ is bounded from below because

$$\mathcal{G}(f,X) \ge \frac{1}{2C}\|(f,X)\|^2_{H^2_{\rho_g} \times H^1_{\rho_g}} - \|(\psi,V)\|_{L^2_{\rho_g^{-1}} \times L^2_{\rho_g^{-1}}}\|(f,X)\|_{L^2_{\rho_g} \times L^2_{\rho_g}}.$$

By standard variational theory, e.g. as in [5, p. 150-152], the functional has a global minimizer $(f,X) \in H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)$. Deriving the Euler-Lagrange equation for the functional $\mathcal{G}$ yields (5.6). The estimate (5.7) follows because $\mathcal{G}(f,X) \le 0$.

$\qquad\square$

5.2. **Weighted Schauder estimates.** This section is devoted to derive the following weighted interior Schauder estimates for the linearized equation.

Theorem 5.6. *Let $(g_0, \pi_0) \in C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ be an initial data set, and let $W_0 \in C^{2,\alpha}(\overline{\Omega})$ be a vector field. Suppose that the kernel of $(D\Phi_{(g_0, \pi_0)}^{W_0})^*$ is trivial on Ω. There exists a neighborhood $\mathcal{U}$ of (g_0, π_0) in $C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$, a neighborhood $\mathcal{W}$ of W_0 in $C^{2,\alpha}(\overline{\Omega})$, and a constant $C > 0$ such that for $(g, \pi) \in \mathcal{U}$, $W \in \mathcal{W}$, and for $(\psi, V) \in \mathcal{B}_0 \times \mathcal{B}_1$, if $(f, X) \in H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)$ weakly solves the linear system*

$$D\Phi_{(g,\pi)}^W \circ \rho_g (D\Phi_{(g,\pi)}^W)^*(f, X) = (\psi, V),$$

then $(f, X) \in \mathcal{B}_4 \times \mathcal{B}_3$ and

(5.8) $$\|(f, X)\|_{\mathcal{B}_4 \times \mathcal{B}_3} \leq C\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1}.$$

Furthermore, if we set $(h, w) = \rho_g (D\Phi_{(g,\pi)}^W)^*(f, X)$, then

(5.9) $$\|(h, w)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \leq C\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1}.$$

We set up the framework of Douglis-Nirenberg [10] for the interior Schauder estimate for an elliptic system. Denote the linear system

(5.10) $$L(f, X) := \rho_g^{-1} D\Phi_{(g,\pi)}^W \rho_g (D\Phi_{(g,\pi)}^W)^*(f, X).$$

Note the leading order terms of this operator are the same as the operator with the usual constraint map. Hence L is strictly elliptic as a system of mixed order in the sense of Douglis-Nirenberg [10], cf. [3, pp. 4-5], [7, p. 207]. In local coordinates, we write $X = (X^1, \ldots, X^n)$ and set U to be the vector-valued function

$$U = (U^1, U^2, \ldots, U^{n+1}) = (f, X^1, \ldots, X^n).$$

Denote the jth component of LU by $(LU)_j$ for $j = 1, \ldots, n+1$, which we express locally as

$$(LU)_j = \sum_{k=1}^{n+1} \sum_{|\beta|=0}^{4} b_{jk}^\beta \partial^\beta U^k =: \sum_{k=1}^{n+1} L_{jk} U^k,$$

where L_{jk} is the differential operator $L_{jk} = \sum_\beta b_{jk}^\beta \partial^\beta$. Using the notation as in [10], we solve for integers $s_1, \ldots, s_{n+1}$ and $t_1, \ldots, t_{n+1}$ such that $s_j + t_k$ is the order of the differential operator L_{jk} and such that $s_1 = 0$. This gives

$$s_1 = 0, \quad t_1 = 4, \quad s_j = -1, \quad t_k = 3 \qquad (j, k = 2, \ldots, n+1).$$

By direct computation, the function b_{jk}^β is a degree one polynomial of $\rho^{-1}\nabla^\ell \rho$, $0 \leq \ell \leq s_j + t_k - |\beta|$ with coefficients depending locally uniformly in $(g, \pi) \in C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ and $W \in C^{2,\alpha}(\overline{\Omega})$. By Proposition 2.11, we have the following estimate.

Lemma 5.7. *The coefficients b_{jk}^β satisfy*

$$\|b_{jk}^\beta\|_{C_{\phi,\phi^{s_j+t_k-|\beta|}}^{-s_j,\alpha}(\Omega)} \leq C,$$

where the constant C depends locally uniformly on $(g, \pi) \in C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ and $W \in C^{2,\alpha}(\overline{\Omega})$.

We have the following interior Schauder estimate. The proof follows from a scaling argument as in [3, Appendix B] and is included in Appendix C, where we spell out the dependence of the constant C.

Theorem 5.8. *Let L be a linear differential operator on $U = (U^1, \ldots, U^{n+1})$ such that the j-th component of the operator L is defined by*

$$(LU)_j = \sum_{k=1}^{n+1} L_{jk}U^k, \qquad (j = 1, \ldots, n+1)$$

where $L_{jk} = \sum_{|\beta|=0}^{s_j+t_k} b_{jk}^\beta \partial^\beta$ is a differential operator of order $s_j + t_k$ with

$$s_1 = 0, \quad t_1 = 4$$
$$s_j = -1, \quad t_k = 3, \qquad (j, k = 2, \ldots, n+1).$$

Let $\varphi_j = \phi^{r+4-t_j}\rho^s$ for $r, s \in \mathbb{R}$. Then

(5.11)
$$\sum_{j=1}^{n+1} \|U^j\|_{C_{\phi,\varphi_j}^{t_j,\alpha}(\Omega)} \le C \left(\sum_{j=1}^{n+1} \|(LU)_j\|_{C_{\phi,\phi^{t_j+s_j}\varphi_j}^{-s_j,\alpha}(\Omega)} + \sum_{j=1}^{n+1} \|U^j\|_{L_{\phi^{-n}\varphi_j^2}^2(\Omega)} \right)$$

where C depends only on n, α, $\displaystyle\sup_{\substack{j,k=1,\ldots,n+1 \\ |\beta| \le s_j+t_k}} \|b_{jk}^\beta\|_{C_{\phi,\phi^{s_j+t_k-|\beta|}}^{-s_j,\alpha}(\Omega)}$, and the lower bound of ellipticity of the operator L.

Remark 5.9. We also note that higher order estimates take the form

$$\sum_{j=1}^{n+1} \|U^j\|_{C_{\phi,\varphi_j}^{\ell+t_j,\alpha}(\Omega)} \le C \left(\sum_{j=1}^{n+1} \|(LU)_j\|_{C_{\phi,\phi^{t_j+s_j}\varphi_j}^{\ell-s_j,\alpha}(\Omega)} + \sum_{j=1}^{n+1} \|U^j\|_{L_{\phi^{-n}\varphi_j^2}^2(\Omega)} \right),$$

where C depends only on n, α, $\displaystyle\sup_{\substack{j,k=1,\ldots,n+1 \\ |\beta| \le s_j+t_k}} \|b_{jk}^\beta\|_{C_{\phi,\phi^{s_j+t_k-|\beta|}}^{\ell-s_j,\alpha}(\Omega)}$, and the lower bound of ellipticity of the operator L.

Proof of Theorem 5.6. Applying Theorem 5.8 (with $r = \frac{n}{2}, s = \frac{1}{2}$) to the operator

$$L(f, X) = \rho_g^{-1} D\Phi_{(g,\pi)}^W \rho_g (D\Phi_{(g,\pi)}^W)^*(f, X),$$

we have

$$\|f\|_{C_{\phi,\phi^{\frac{n}{2}}\rho_g^{\frac{1}{2}}}^{4,\alpha}} + \|X\|_{C_{\phi,\phi^{1+\frac{n}{2}}\rho_g^{\frac{1}{2}}}^{3,\alpha}} = \sum_{i=1}^{n+1} \|U^i\|_{C_{\phi,\varphi_i}^{t_i,\alpha}}$$

$$\le C \left[\sum_{i=1}^{n+1} \|(LU)_i\|_{C_{\phi,\phi^{t_i+s_i}\varphi_i}^{-s_i,\alpha}} + \sum_{i=1}^{n+1} \|U^i\|_{L_{\phi^{-n}\varphi_i^2}^2} \right]$$

$$= C \left(\|\rho_g^{-1}\psi\|_{C_{\phi,\phi^{4+\frac{n}{2}}\rho_g^{\frac{1}{2}}}^{0,\alpha}} + \|\rho_g^{-1}V\|_{C_{\phi,\phi^{3+\frac{n}{2}}\rho_g^{\frac{1}{2}}}^{1,\alpha}} + \|f\|_{L_{\rho g}^2} + \|X\|_{L_{\phi^2\rho g}^2} \right)$$

$$\le C \left(\|\psi\|_{C_{\phi,\phi^{4+\frac{n}{2}}\rho_g^{-\frac{1}{2}}}^{0,\alpha}} + \|V\|_{C_{\phi,\phi^{3+\frac{n}{2}}\rho_g^{-\frac{1}{2}}}^{1,\alpha}} + \|f\|_{L_{\rho g}^2} + \|X\|_{L_{\phi^2\rho g}^2} \right).$$

The above Hölder estimate, together with the Sobolev estimate (5.7), implies (5.8).

The estimate (5.9) for (h, w) follows because differentiation is a continuous operator from $C^{k,\alpha}_{\phi,\varphi}$ to $C^{k-1,\alpha}_{\phi,\phi\varphi}$ and from $H^k_{\rho_g^{-1}}$ to $H^{k-1}_{\rho_g^{-1}}$. □

5.3. **Solving the nonlinear problem by iteration.** We complete the proof of Theorem 3.1, which can be formulated more explicitly as follows.

Theorem 5.10. *Let* $(g_0, \pi_0) \in C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ *be an initial data set, and let* $W_0 \in C^{2,\alpha}(\overline{\Omega})$ *be a vector field. Suppose that the kernel of* $(D\Phi^{W_0}_{(g_0,\pi_0)})^*$ *is trivial in* Ω. *Let* C *be the constant from Theorem 5.6. There is a neighborhood* $\mathcal{U}$ *of* (g_0, π_0) *in* $C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$, *and a neighborhood* $\mathcal{W}$ *of* W_0 *in* $C^{2,\alpha}(\overline{\Omega})$, *and* $\epsilon > 0$ *such that for* $(g, \pi) \in \mathcal{U}$, $W \in \mathcal{W}$ *and for* $(\psi, V) \in \mathcal{B}_0 \times \mathcal{B}_1$ *with* $\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq \epsilon$, *there exists* $(f, X) \in \mathcal{B}_4 \times \mathcal{B}_3$ *such that*

$$(h, w) = \rho_g(D\Phi^W_{(g,\pi)})^*(f, X)$$

satisfies $(h, w) \in \mathcal{B}_2 \times \mathcal{B}_2$ *and*

$$\Phi^W_{(g,\pi)}(g + h, \pi + w) = \Phi^W_{(g,\pi)}(g, \pi) + (\psi, V)$$

with $\|(h, w)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \leq C\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1}$.

Proof. By Theorem 5.5, there exists $(f_0, X_0) \in \mathcal{B}_4 \times \mathcal{B}_3$ such that $(h_0, w_0) = \rho_g(D\Phi^W_{(g,\pi)})^*(f_0, X_0)$ solves $D\Phi^W_{(g,\pi)}(h_0, w_0) = (\psi, V)$ and, by Theorem 5.6, the following estimates hold

$$\|(h_0, w_0)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \leq C\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1}, \qquad \|(f_0, X_0)\|_{\mathcal{B}_4 \times \mathcal{B}_3} \leq C\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1}.$$

By Lemma A.3 we obtain that

$$\|\Phi^W_{(g,\pi)}(g, \pi) + (\psi, V) - \Phi^W_{(g,\pi)}(g + h_0, \pi + w_0)\|_{\mathcal{B}_0 \times \mathcal{B}_1}$$
$$= \|Q^W_{(g,\pi)}(h_0, w_0)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq D\|(h_0, w_0)\|^2_{\mathcal{B}_2 \times \mathcal{B}_2} \leq DC^2\|(\psi, V)\|^2_{\mathcal{B}_0 \times \mathcal{B}_1}.$$

Write $(\gamma_1, \tau_1) = (g, \pi) + (h_0, w_0)$. Note that γ_1 is a Riemannian metric provided $\|(h_0, w_0)\|_{\mathcal{B}_0 \times \mathcal{B}_1}$ is sufficiently small. Fix $\delta \in (0, 1)$. We set ϵ sufficiently small so that $DC^2\epsilon^{1-\delta} \leq 1$.

We then proceed recursively as in the following lemma, whose proof is included in Appendix D.

Lemma 5.11. *Fix* (g_0, π_0) *and* $\delta \in (0, 1)$. *There exists a neighborhood* $\mathcal{U}$ *of* (g_0, π_0) *in* $C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$, *a neighborhood* $\mathcal{W}$ *of* W_0 *in* $C^{2,\alpha}(\overline{\Omega})$, *and* $\epsilon \in (0, \frac{1}{2})$ *depending only on* δ *and* Ω, *such that for* $(g, \pi) \in \mathcal{U}$ *and for* $W \in \mathcal{W}$ *the following holds. Suppose that* $m \geq 1$ *and we have constructed* $(f_0, X_0), \ldots, (f_{m-1}, X_{m-1}) \in \mathcal{B}_4(\Omega) \times \mathcal{B}_3(\Omega)$, $(h_0, w_0), \ldots, (h_{m-1}, w_{m-1}) \in \mathcal{B}_2(\Omega) \times \mathcal{B}_2(\Omega)$ *where* $(h_p, w_p) = \rho_g(D\Phi^W_{(g,\pi)})^*(f_p, X_p)$ *and* $(\gamma_1, \tau_1), \ldots, (\gamma_m, \tau_m) \in C^{2,\alpha}(\overline{\Omega}) \times C^{1,\alpha}(\overline{\Omega})$ *where* $\gamma_j = g + \sum_{p=0}^{j-1} h_p$ *and* $\tau_j = \pi + \sum_{p=0}^{j-1} w_p$. *Assume that* $\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq \epsilon$ *and that for all* $0 \leq p \leq m - 1$,

(5.12) $\|(f_p, X_p)\|_{\mathcal{B}_4 \times \mathcal{B}_3} \leq C\|(\psi, V)\|^{1+p\delta}_{\mathcal{B}_0 \times \mathcal{B}_1}$ *and* $\|(h_p, w_p)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \leq C\|(\psi, V)\|^{1+p\delta}_{\mathcal{B}_0 \times \mathcal{B}_1}$,

and that for all $1 \leq j \leq m$,

(5.13) $\|\Phi^W_{(g,\pi)}(g, \pi) + (\psi, V) - \Phi^W_{(g,\pi)}(\gamma_j, \tau_j)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq \|(\psi, V)\|^{1+j\delta}_{\mathcal{B}_0 \times \mathcal{B}_1}$.

If we define $(h_m, w_m) = \rho_g D(\Phi^W_{(g,\pi)})^*(f_m, X_m)$ *where* (f_m, X_m) *is the variational solution to*

$$D\Phi^W_{(g,\pi)}\rho_g(D\Phi^W_{(g,\pi)})^*(f_m, X_m) = \Phi^W_{(g,\pi)}(g, \pi) + (\psi, V) - \Phi^W_{(g,\pi)}(\gamma_m, \tau_m),$$

and if we set $(\gamma_{m+1}, \tau_{m+1}) = (\gamma_m, \tau_m) + (h_m, w_m)$, then the estimates (5.12) and (5.13) hold for $p = m$ and $j = m + 1$.

We obtain the series $\sum_{p=0}^{\infty}(f_p, X_p)$ converging in $\mathcal{B}_4(\Omega) \times \mathcal{B}_3(\Omega)$ to some (f, X). Let $(h, w) = \rho_g(D\Phi_{(g,\pi)}^W)^*(f, X)$, then $(g + h, \pi + w)$ satisfies the nonlinear system $\Phi_{(g,\pi)}^W(g + h, \pi + w) = \Phi_{(g,\pi)}^W(g, \pi) + (\psi, V)$. $\qquad\square$

6. Proof of Theorem 4.3 and higher regularity

The proof of Theorem 4.3 is along the same line as the proof of Theorem 3.1 and Theorem 5.10 in Section 5. Recall that $\mathcal{U}_0$ is a bounded neighborhood of a Riemannian metric g_0, as in Section 2.4, and the compact set $\Omega_0 \subset \Omega$ is chosen so that $\rho_g \equiv 1$ on Ω_0 for all $g \in \mathcal{U}_0$. Fix a smooth bump function ζ supported in Ω_0. Denote by S_g the $L^2(d\mu_g)$-orthogonal complement of ζK, where K is the kernel of $(D\Phi_{(g_0,\pi_0)}^{W_0})^*$. Let $\Pi_g : \mathcal{B}_0 \times \mathcal{B}_1 \to (\mathcal{B}_0 \times \mathcal{B}_1) \cap S_g$ be the $L^2(d\mu_g)$-orthogonal projection.

We begin with a basic lemma on the projection map. Throughout this section, the function spaces are all taken on Ω, unless otherwise indicated.

Lemma 6.1. *There is a constant $C > 0$ such that for $g \in \mathcal{U}_0$,*

$$\|\Pi_g(\psi, V)\|_{L^2_{\rho_g^{-1}}} \leq \|(\psi, V)\|_{L^2_{\rho_g^{-1}}}$$

$$\|(\psi, V)^{\perp}\|_{L^2_{\rho_g^{-1}}} \leq \|(\psi, V)\|_{L^2_{\rho_g^{-1}}}$$

$$\|\Pi_g(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq C\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1},$$

where $(\psi, V)^{\perp} = (\psi, V) - \Pi_g(\psi, V) \in \zeta K$.

Proof. Because ζ is supported in Ω_0 where $\rho_g \equiv 1$, the following weighted orthogonality holds:

$$\langle \Pi_g(\psi, V), (\psi, V)^{\perp} \rangle_{L^2_{\rho_g^{-1}}(\Omega)} = \langle \Pi_g(\psi, V), (\psi, V)^{\perp} \rangle_{L^2(\Omega_0)} = 0.$$

The first two inequalities follow by

$$\|(\psi, V)\|_{L^2_{\rho_g^{-1}}}^2 = \|\Pi_g(\psi, V)\|_{L^2_{\rho_g^{-1}}}^2 + \|(\psi, V)^{\perp}\|_{L^2_{\rho_g^{-1}}}^2.$$

To establish the last inequality, we recall that all norms of the finite-dimensional space ζK are equivalent and the $\mathcal{B}_0 \times \mathcal{B}_1$-norms on ζK are all uniformly equivalent to each other for $g \in \mathcal{U}_0$:

$$\|\Pi_g(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq \|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} + \|(\psi, V)^{\perp}\|_{\mathcal{B}_0 \times \mathcal{B}_1}$$
$$\leq \|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} + C\|(\psi, V)^{\perp}\|_{L^2}$$
$$= \|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} + C\|(\psi, V)^{\perp}\|_{L^2_{\rho_g^{-1}}}$$
$$\leq (1 + C)\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1}.$$

In the third line we use that ζ is supported where $\rho_g \equiv 1$.

$\qquad\square$

6.1. **The linearized equation.** In this section, we solve the linearized equation for the operator $\Pi_{g_0} \circ \Phi^W_{(g,\pi)}$. We again denote by $D\Phi^W_{(g,\pi)} = D\Phi^W_{(g,\pi)}|_{(g,\pi)}$ the linearization at (g,π).

Theorem 6.2. *Let $(g_0, \pi_0) \in C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ be an initial data set, and let $W_0 \in C^{2,\alpha}(\overline{\Omega})$ be a vector field. There is a neighborhood $\mathcal{U}$ of (g_0, π_0) in $C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$, a neighborhood $\mathcal{W}$ of W_0 in $C^{2,\alpha}(\overline{\Omega})$, and a constant $C > 0$, such that for $(g,\pi) \in \mathcal{U}, W \in \mathcal{W}$ and for $(\psi, V) \in L^2_{\rho_g^{-1}}$, there is a unique $(f, X) \in (H^2_{\rho_g} \times H^1_{\rho_g}) \cap S_g$ that weakly solves*

(6.1)
$$\Pi_{g_0} \circ D\Phi^W_{(g,\pi)} \circ \rho_g (D\Phi^W_{(g,\pi)})^*(f, X) = \Pi_{g_0}(\psi, V),$$

or, equivalently,

$$D\Phi^W_{(g,\pi)} \circ \rho_g (D\Phi^W_{(g,\pi)})^*(f, X) - (\psi, V) \in \zeta K.$$

Moreover, (f, X) satisfies the estimate

(6.2)
$$\|(f, X)\|_{H^2_{\rho_g} \times H^1_{\rho_g}} \leq C \|(\psi, V)\|_{L^2_{\rho_g^{-1}} \times L^2_{\rho_g^{-1}}}.$$

The proof of the theorem is a modification of the variational argument in Theorem 5.5. Note that the projection map is taken with respect to a fixed metric g_0, but the functional $\mathcal{G}$ below naturally involves g. To resolve this, we solve the linearized equation whose projection is with respect to g and look for solutions $(f, X) \in S_g$. Theorem 6.2 follows from the proposition below.

Proposition 6.3. *Let $(g_0, \pi_0) \in C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ be an initial data set, and let $W_0 \in C^{2,\alpha}(\overline{\Omega})$ be a vector field. There is a neighborhood $\mathcal{U}$ of (g_0, π_0) in $C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$, a neighborhood $\mathcal{W}$ of W_0 in $C^{2,\alpha}(\overline{\Omega})$, and a constant $C > 0$ such that for $(g,\pi) \in \mathcal{U}, W \in \mathcal{W}$, and for $(\psi, V) \in L^2_{\rho_g^{-1}}(\Omega)$, there is a unique $(f, X) \in (H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)) \cap S_g$ that weakly solves*

(6.3)
$$\Pi_g \circ D\Phi^W_{(g,\pi)} \circ \rho_g (D\Phi^W_{(g,\pi)})^*(f, X) = \Pi_g(\psi, V),$$

or equivalently,

$$D\Phi^W_{(g,\pi)} \circ \rho_g (D\Phi^W_{(g,\pi)})^*(f, X) - (\psi, V) \in \zeta K.$$

Moreover, (f, X) satisfies the estimate

(6.4)
$$\|(f, X)\|_{H^2_{\rho_g} \times H^1_{\rho_g}} \leq C \|(\psi, V)\|_{L^2_{\rho_g^{-1}} \times L^2_{\rho_g^{-1}}}.$$

Proof. For given $(\psi, V) \in L^2_{\rho_g^{-1}}(\Omega)$, let $\mathcal{G}$ be a similar functional as in Theorem 5.5, whose domain is restricted to the linear subspace $(H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)) \cap S_g$: for $(f, X) \in (H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)) \cap S_g$, let

$$\mathcal{G}(f, X) = \int_\Omega \left[\frac{1}{2} \rho_g \left| (D\Phi^W_{(g,\pi)})^*(f, X) \right|_g^2 - \Pi_g(\psi, V) \cdot_g (f, X) \right] d\mu_g.$$

By the coercivity estimate (see Remark 5.4),

$$\mathcal{G}(f, X) \geq \frac{1}{2C} \|(f, X)\|^2_{H^2_{\rho_g} \times H^1_{\rho_g}} - \|\Pi_g(\psi, V)\|_{L^2_{\rho_g^{-1}}} \|(f, X)\|_{L^2_{\rho_g}}$$

It is clear that functional is still convex when restricted on the linear subspace, so there is a unique minimizer $(f, X) \in (H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)) \cap S_g$. Furthermore, since $\mathcal{G}(f, X) \leq 0$ and by Lemma 6.1, the estimate (6.4) holds.

To see that (f, X) solves (6.3) weakly, we need to show that, for all smooth compactly supported test fields (u, Y),

$$(6.5) \qquad \int_\Omega \left(\Pi_g \circ D\Phi^W_{(g,\pi)} \circ \rho_g (D\Phi_{(g,\pi)})^* (f, X) - \Pi_g(\psi, V) \right) \cdot_g (u, Y) \, d\mu_g = 0.$$

The equality trivially holds for $(u, Y) \in \zeta K$. For $(u, Y) \in S_g$, because (f, X) is a minimizer in $(H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)) \cap S_g$, we see that

$$
\begin{aligned}
0 &= \frac{d}{dt}\bigg|_{t=0} \mathcal{G}\left((f, X) + t(u, Y) \right) \\
&= \int_\Omega \rho_g (D\Phi^W_{(g,\pi)})^* (f, X) \cdot_g (D\Phi^W_{(g,\pi)})^* (u, Y) - \Pi_g(\psi, V) \cdot_g (u, Y) \, d\mu_g \\
&= \int_\Omega \left(D\Phi^W_{(g,\pi)} \circ \rho_g (D\Phi^W_{(g,\pi)})^* (f, X) - \Pi_g(\psi, V) \right) \cdot_g (u, Y) \, d\mu_g ,
\end{aligned}
$$

which implies (6.5). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

6.2. Weighted Schauder estimates.

Proposition 6.4. *Let* $(g_0, \pi_0) \in C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$ *be an initial data set, and let* $W_0 \in C^{2,\alpha}(\overline{\Omega})$ *be a vector field. There is a neighborhood* $\mathcal{U}$ *of* (g_0, π_0) *in* $C^{4,\alpha}(\overline{\Omega}) \times C^{3,\alpha}(\overline{\Omega})$, *a neighborhood* $\mathcal{W}$ *of* W_0 *in* $C^{2,\alpha}(\overline{\Omega})$, *and a constant* $C > 0$ *such that for* $(g, \pi) \in \mathcal{U}$, $W \in \mathcal{W}$, *and for* $(\psi, V) \in \mathcal{B}_0 \times \mathcal{B}_1$, *if* $(f, X) \in (H^2_{\rho_g}(\Omega) \times H^1_{\rho_g}(\Omega)) \cap S_g$ *weakly solves the linear system*

$$\Pi_{g_0} \circ D\Phi^W_{(g,\pi)} \circ \rho_g (D\Phi^W_{(g,\pi)})^* (f, X) = \Pi_{g_0}(\psi, V),$$

then $(f, X) \in \mathcal{B}_4 \times \mathcal{B}_3$ *and*

$$\| (f, X) \|_{\mathcal{B}_4 \times \mathcal{B}_3} \leq C \| \Pi_{g_0}(\psi, V) \|_{\mathcal{B}_0 \times \mathcal{B}_1}.$$

Proof. Define

$$
\begin{aligned}
L(f, X) &= \rho_g^{-1} D\Phi^W_{(g,\pi)} \circ \rho_g (D\Phi^W_{(g,\pi)})^* (f, X) \\
U &= (f, X^1, \ldots, X^n) \quad \text{(with respect to a fixed coordinate chart)}
\end{aligned}
$$

as in the proof of Theorem 5.6. Below we denote ρ_g by ρ. We write

$$LU = \rho^{-1} \Pi_{g_0}(\rho LU) + \rho^{-1}(\rho LU)^\perp = \rho^{-1} \Pi_{g_0}(\psi, V) + (LU)^\perp ,$$

where we used the fact that $\rho_{g_0} \equiv 1$ on the support of ζ. Applying Theorem 5.8 as in the proof of Theorem 5.6, we have

$$
\begin{aligned}
\|f\|_{C^{4,\alpha}_{\phi,\phi^{\frac{n}{2}}\rho^{\frac{1}{2}}}} &+ \|X\|_{C^{3,\alpha}_{\phi,\phi^{1+\frac{n}{2}}\rho^{\frac{1}{2}}}} = \sum_{j=1}^{n+1} \|U^j\|_{C^{t_j,\alpha}_{\phi,\varphi_j}} \\
&\leq C \left[\sum_{j=1}^{n+1} \|(LU)_j\|_{C^{-s_j,\alpha}_{\phi,\phi^{t_j+s_j}\varphi_j}} + \sum_{j=1}^{n+1} \|U^j\|_{L^2_{\phi^{-n}\varphi_j^2}} \right] \\
&\leq C \left(\sum_{j=1}^{n+1} \|(\rho^{-1}\Pi_{g_0}(\psi,V))_j\|_{C^{-s_j,\alpha}_{\phi,\phi^{t_j+s_j}\varphi_j}} + \sum_{j=1}^{n+1} \|(LU)^{\perp}_j\|_{C^{-s_j,\alpha}_{\phi,\phi^{t_j+s_j}\varphi_j}} + \sum_{j=1}^{n+1} \|U^j\|_{L^2_{\phi^{-n}\varphi_j^2}} \right) \\
&\leq C \left(\sum_{j=1}^{n+1} \|(\Pi_{g_0}(\psi,V))_j\|_{C^{-s_j,\alpha}_{\phi,\phi^{t_j+s_j}\varphi_j\rho^{-1}}} + \sum_{j=1}^{n+1} \|(LU)^{\perp}_j\|_{C^{-s_j,\alpha}_{\phi,\phi^{t_j+s_j}\varphi_j}} + \sum_{j=1}^{n+1} \|U^j\|_{L^2_{\phi^{-n}\varphi_j^2}} \right).
\end{aligned}
$$

We now estimate the $(LU)^{\perp}$-term. Just as in the proof of Lemma 6.1, using the fact that all norms of a finite-dimensional space are equivalent, and that the support of $(LU)^{\perp}_j$ is contained in Ω_0, we obtain the following estimate, uniformly across $\mathcal{U}$ and $\mathcal{W}$:

$$
\|(LU)^{\perp}_j\|_{C^{-s_j,\alpha}_{\phi,\phi^{t_j+s_j}\varphi_j}(\Omega)} \leq C\|(LU)^{\perp}\|_{L^2(\Omega)} = C\|(LU)^{\perp}\|_{L^2(\Omega_0)} \leq C\|LU\|_{L^2(\Omega_0)}.
$$

To estimate $\|LU\|_{L^2(\Omega_0)}$, we note that L is a differential operator that contains four derivatives on f and three derivatives on X and that Ω_0 is compact. Then by enlarging the constant C if necessary and by interpolation, we have

$$
\begin{aligned}
\|(LU)^{\perp}_j\|_{C^{-s_j,\alpha}_{\phi,\phi^{t_j+s_j}\varphi_j}(\Omega)} &\leq C\|(f,X)\|_{C^4(\Omega_0)\times C^3(\Omega_0)} \\
&\leq \epsilon\|(f,X)\|_{C^{4,\alpha}(\Omega_0)\times C^{3,\alpha}(\Omega_0)} + C(\epsilon)\|(f,X)\|_{L^2(\Omega_0)\times L^2(\Omega_0)}. \\
&\leq C\left[\epsilon\left(\|f\|_{C^{4,\alpha}_{\phi,\phi^{\frac{n}{2}}\rho^{\frac{1}{2}}}} + \|X\|_{C^{3,\alpha}_{\phi,\phi^{1+\frac{n}{2}}\rho^{\frac{1}{2}}}} \right) + C(\epsilon)\left(\|f\|_{L^2_\rho} + \|X\|_{L^2_{\phi^2\rho}} \right) \right],
\end{aligned}
$$

where in the last inequality, we replace the norms on Ω_0 by the corresponding weighted norms on Ω, up to the multiple of a constant that is uniform in $(g,\pi) \in \mathcal{U}$ and $W \in \mathcal{W}$. Choosing ϵ sufficiently small and absorbing the weighted Hölder norm on (f,X) gives the desired Hölder estimates. The weighted Sobolev estimates on (f,X) follow by Proposition 6.3, and the fact that the solution to the linear system is unique.

$\square$

6.3. Solving the nonlinear projected problem by iteration.

We discuss the proof of Theorem 4.3 to solve for the nonlinear problem: for $(\psi,V) \in \mathcal{B}_0 \times \mathcal{B}_1$ sufficiently small, there is $(f,X) \in (\mathcal{B}_4 \times \mathcal{B}_3) \cap S_g$ such that $(h,w) = \rho_g(D\Phi^W_{(g,\pi)})^*(f,X)$ solves

$$
\Pi_{g_0} \circ \Phi^W_{(g,\pi)}(g+h,\pi+w) = \Pi_{g_0} \circ \Phi^W_{(g,\pi)}(g,\pi) + \Pi_{g_0}(\psi,V)
$$

with $\|(h,w)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \leq C\|\Pi_{g_0}(\psi,V)\|_{\mathcal{B}_0 \times \mathcal{B}_1}$.

The proof follows the same iteration scheme as in the proof of Theorem 5.10 by replacing $\Phi^W_{(g,\pi)}$ with $\Pi_{g_0} \circ \Phi^W_{(g,\pi)}$ and (ψ, V) with $\Pi_{g_0}(\psi, V)$. The initial step of the iteration is solving the following for $(f_0, X_0) \in S_g$

$$\Pi_{g_0} \circ D\Phi^W_{(g,\pi)} \circ \rho_g(D\Phi^W_{(g,\pi)})^*(f_0, X_0) = \Pi_{g_0}(\psi, V) \ ,$$

and then setting $(h_0, w_0) = \rho_g(D\Phi^W_{(g,\pi)})^*(f_0, X_0)$ and $(\gamma_1, \tau_1) = (g + h_0, \pi + w_0)$. We then solve inductively for $m \geq 0$

$$\Pi_{g_0} \circ D\Phi^W_{(g,\pi)} \circ \rho_g(D\Phi^W_{(g,\pi)})^*(f_m, X_m) = \Pi_{g_0} \circ \Phi^W_{(g,\pi)}(g, \pi) + \Pi_{g_0}(\psi, V) - \Pi_{g_0} \circ \Phi^W_{(g,\pi)}(\gamma_m, \tau_m)$$

and set $(h_m, w_m) = \rho_g(D\Phi^W_{(g,\pi)})^*(f_m, X_m)$ and $(\gamma_{m+1}, \tau_{m+1}) = (g + \sum_{p=0}^m h_p, \pi + \sum_{p=0}^m w_p)$. The essential estimates to guarantee the iteration procedure converges are the following:

$$\|(f_m, X_m)\|_{\mathcal{B}_4 \times \mathcal{B}_3} \leq C \|\Pi_{g_0} \circ \Phi^W_{(g,\pi)}(g, \pi) + \Pi_{g_0}(\psi, V) - \Pi_{g_0} \circ \Phi^W_{(g,\pi)}(\gamma_m, \tau_m)\|_{\mathcal{B}_0 \times \mathcal{B}_1}$$

$$\|(h_m, w_m)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \leq C \|(f_m, X_m)\|_{\mathcal{B}_4 \times \mathcal{B}_3}$$

and

$$\|\Pi_{g_0} \circ D\Phi^W_{(g,\pi)}|_{(\gamma,\tau)}(h, w) - \Pi_{g_0} \circ D\Phi^W_{(g,\pi)}|_{(\gamma',\tau')}(h, w)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq D \|(h, w)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \|(\gamma - \gamma', \tau - \tau')\|_{\mathcal{B}_2 \times \mathcal{B}_2}$$

$$\|\Pi_{g_0} \circ Q^W_{(g,\pi)}(h_0, w_0)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq D \|(h_0, w_0)\|^2_{\mathcal{B}_2 \times \mathcal{B}_2}.$$

The first two estimates follow by Proposition 6.4 and the fact that the differential operator is continuous between the corresponding weighted spaces. The last two estimates follow by Lemma 6.1 and the estimates for unprojected operators from Lemma A.3.

6.4. **Higher order regularity and continuous dependence.** The previous analysis, in particular Remark 5.9, implies the following version of the local surjectivity theorem with higher order regularity; continuous dependence also follows directly from the above analysis, cf. [6, Proposition 3.7]. For simplicity we state the theorem for (ψ, V) of compact support, but one can more generally pose that (ψ, V) lies in suitable weighted spaces (an infinite intersection of such spaces for the C^∞-case, for example).

Theorem 6.5. *Let $k \geq 0$. Let $(g_0, \pi_0) \in C^{k+4,\alpha}(\overline{\Omega}) \times C^{k+3,\alpha}(\overline{\Omega})$ be an initial data set, and let $W_0 \in C^{k+2,\alpha}(\overline{\Omega})$. Suppose that the kernel of $(D\Phi^{W_0}_{(g_0,\pi_0)})^*$ is K (may be trivial). Then there is a $C^{k+4,\alpha}(\overline{\Omega}) \times C^{k+3,\alpha}(\overline{\Omega})$ neighborhood $\mathcal{U}$ of (g_0, π_0), a neighborhood $\mathcal{W}$ of W_0 in $C^{k+2,\alpha}(\overline{\Omega})$, and constants $\epsilon > 0$, $C > 0$ such that for $(g, \pi) \in \mathcal{U}$ and for $(\psi, V) \in C^{k,\alpha}_c(\Omega) \times C^{k+1,\alpha}_c(\Omega)$ with $\|(\psi, V)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq \epsilon$, there is a pair of symmetric tensors $(h, w) \in C^{k+2,\alpha}_c(\Omega) \times C^{k+2,\alpha}_c(\Omega)$ with $\|(h, w)\|_{C^{k+2,\alpha} \times C^{k+2,\alpha}} \leq C \|(\psi, V)\|_{C^{k,\alpha} \times C^{k+1,\alpha}}$ such that the initial data set $(g + h, \pi + w) \in C^{k+2,\alpha}(\overline{\Omega}) \times C^{k+2,\alpha}(\overline{\Omega})$ satisfies*

$$\Pi_{g_0} \circ \Phi^W_{(g,\pi)}(g + h, \pi + w) = \Pi_{g_0} \Phi^W_{(g,\pi)}(g, \pi) + \Pi_{g_0}(\psi, V),$$

where the weighted norms are taken with respect to g. The solution $(h, w) \in \mathcal{B}_2 \times \mathcal{B}_2$ depends continuously on $(\psi, V) \in \mathcal{B}_0 \times \mathcal{B}_1$. If, in addition, $(g, \pi) \in C^\infty(\overline{\Omega})$ and $(\psi, V) \in C^\infty_c(\Omega)$, then $(h, w) \in C^\infty_c(\Omega)$.

Proof. In the proof of Theorem 4.3 for $k = 0$, we have obtained $(f, X) \in (\mathcal{B}_4 \times \mathcal{B}_3) \cap S_g$ such that

$$(h, w) = \rho_g (D\Phi^W_{(g,\pi)})^*(f, X)$$

solves the nonlinear equation. That is, (f, X) satisfies the quasi-linear elliptic system

$$\Phi^W_{(g,\pi)} \left((g, \pi) + \rho_g (D\Phi^W_{(g,\pi)})^*(f, X) \right) - \Phi^W_{(g,\pi)}(g, \pi) - (\psi, V) \in \zeta K.$$

Because elements in ζK are smooth with compact support, using the initial regularity $(f, X) \in \mathcal{B}_4 \times \mathcal{B}_3$, along with bootstrapping, one can get the higher-order estimates and boundary decay, cf. [6, Sec. 3.7]; this method produces the C^∞-regularity statement as well. $\qquad\square$

Alternatively, one could pose the smallness condition on the norm $\|(\psi, V)\|_{C^{k,\alpha} \times C^{k+1,\alpha}}$ and prove convergence of the iteration scheme in $C^{k+2,\alpha} \times C^{k+2,\alpha}$ to the limit (h, w), applying the estimates in Remark 5.9 to the linearized operator. This argument proves the finite regularity statement.

This completes the proof of Theorem 1.2, and this also gives us what is required in the proof of Theorem 1.3 and Theorem 1.4.

Appendix A. Estimates on the Taylor expansions

Consider the Taylor expansion of the constraint map Φ at an initial data set (g, π)

(A.1) $$\Phi(g + h, \pi + w) = \Phi(g, \pi) + D\Phi|_{(g,\pi)}(h, w) + Q_{(g,\pi)}(h, w).$$

In local coordinates, the first slot of $D\Phi|_{(g,\pi)}(h, w)$ is a homogeneous linear polynomial in $\partial^2_{ij} h_{kl}, \partial_i h_{kl}, h_{kl}$ and $\partial_i w^{kl}, w^{kl}$ whose coefficients are smooth functions of $\partial^2_{ij} g_{kl}, \partial_i g_{kl}, g_{kl}, \partial_i \pi^{kl}, \pi^{kl}$, and the second slot of $D\Phi|_{(g,\pi)}(h, w)$ is of the same type but contains no second derivatives of g_{kl}, h_{kl}. The remainder term $Q_{(g,\pi)}(h, w)$ is a homogeneous quadratic polynomial in $\partial^2_{ij} h_{kl}, \partial_i h_{kl}, h_{kl}, \partial_i w^{kl}, w^{kl}$ whose coefficients are smooth functions of $\partial^2_{ij} g_{kl}, \partial_i g_{kl}, g_{kl}, \partial_i \pi^{kl}, \pi^{kl}$ and $\partial^2_{ij} h_{kl}, \partial_i h_{kl}, h_{kl}, \partial_i w^{kl}, w^{kl}$, and note that the second slot of $Q_{(g,\pi)}(h, w)$ contains no second derivatives in g_{kl} and h_{kl}.

It is clear that if $(g, \pi) \in C^{k+1,\alpha}(\overline{\Omega}) \times C^{k+1,\alpha}(\overline{\Omega})$ for $k \geq 1$, we have

(A.2)
$$\|D\Phi|_{(g,\pi)}(h, w)\|_{C^{k-1,\alpha}(\overline{\Omega}) \times C^{k,\alpha}(\overline{\Omega})} \leq C\|(h, w)\|_{C^{k+1,\alpha}(\overline{\Omega}) \times C^{k+1,\alpha}(\overline{\Omega})}$$
$$\|Q_{(g,\pi)}(h, w)\|_{C^{k-1,\alpha}(\overline{\Omega}) \times C^{k,\alpha}(\overline{\Omega})} \leq C\|(h, w)\|^2_{C^{k+1,\alpha}(\overline{\Omega}) \times C^{k+1,\alpha}(\overline{\Omega})}$$

where C depends locally uniformly on $(g, \pi), (h, w) \in C^{k+2,\alpha} \times C^{\ell+1,\alpha}$. By direct analysis (with a bit more care), we have the following estimates involving the weights.

Lemma A.1. *Suppose that f is a $C^{2,\alpha}(\overline{\Omega})$ function such that ∇f is supported on a compact subset of Ω. Then*

$$\|D\Phi|_{(g,\pi)}(fh, fw) - f D\Phi|_{(g,\pi)}(h, w)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq C\|\nabla f\|_{\mathcal{B}_1}\|(h, w)\|_{C^{2,\alpha} \times C^{2,\alpha}}$$
$$\|D\Phi|_{(g+h,\pi+w)}(h, w) - D\Phi|_{(g,\pi)}(h, w)\|_{C^{0,\alpha} \times C^{1,\alpha}} \leq C\|(h, w)\|^2_{C^{2,\alpha} \times C^{2,\alpha}}$$
$$\|Q_{(g,\pi)}(fh, fw) - f^2 \widetilde{Q}_{(g,\pi)}(h, w)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq C\|\nabla f\|_{\mathcal{B}_1}\|(h, w)\|^2_{C^{2,\alpha} \times C^{2,\alpha}}$$

for some

$$\|\widetilde{Q}_{(g,\pi)}(h, w)\|_{C^{0,\alpha} \times C^{1,\alpha}} \leq C\|(h, w)\|^2_{C^{2,\alpha} \times C^{2,\alpha}},$$

where C depends locally uniformly on $(g, \pi), (h, w) \in C^{2,\alpha} \times C^{2,\alpha}$ and $f \in C^{2,\alpha}$. Note all the norms are taken on $\overline{\Omega}$.

We apply those estimates to interpolation between initial data sets.

Lemma A.2. *Let $0 \leq \chi \leq 1$ be a $C^{k+2,\alpha}(\overline{\Omega})$ bump function such that $\chi(1-\chi)$ is supported on a compact subset of Ω. Denote by $(g, \pi) = \chi(g_1, \pi_1) + (1-\chi)(g_2, \pi_2)$. Then*

(1) The following Hölder estimate holds:

$$\|\Phi(g, \pi) - \chi\Phi(g_1, \pi_1) - (1-\chi)\Phi(g_2, \pi_2)\|_{C^{k-1,\alpha}(\overline{\Omega}) \times C^{k,\alpha}(\overline{\Omega})}$$
$$\leq C\|(g_1 - g_2, \pi_1 - \pi_2)\|_{C^{k+1,\alpha}(\overline{\Omega}) \times C^{k+1,\alpha}(\overline{\Omega})},$$

where C depends locally uniformly on $(g_1, \pi_1), (g_2, \pi_2) \in C^{k+2,\alpha}(\overline{\Omega}) \times C^{k+1,\alpha}(\overline{\Omega})$ and $\chi \in C^{k+2,\alpha}(\overline{\Omega})$.

(2) The following weighted estimate holds:

$$\|\Phi(g, \pi) - \chi\Phi(g_1, \pi_1) - (1-\chi)\Phi(g_2, \pi_2)\|_{\mathcal{B}_0 \times \mathcal{B}_1}$$
$$\leq C(\|\chi(1-\chi)\|_{\mathcal{B}_0} + \|\nabla\chi\|_{\mathcal{B}_1})\|(g_1 - g_2, \pi_1 - \pi_2)\|_{C^{2,\alpha} \times C^{2,\alpha}},$$

where C depends locally uniformly on $(g_1, \pi_1), (g_2, \pi_2) \in C^{2,\alpha} \times C^{2,\alpha}$ and $\chi \in C^{2,\alpha}$.

Note that all the norms are taken on $\overline{\Omega}$ with respect to a fixed metric, say g_1.

Proof. Writing $\Phi(g, \pi) = \chi\Phi(g, \pi) + (1-\chi)\Phi(g, \pi)$, we apply Taylor expansion to the first term at (g_1, π_1) and the second term at (g_2, π_2) and derive

$$\Phi(g, \pi) - (\chi\Phi(g_1, \pi_1) + (1-\chi)\Phi(g_2, \pi_2))$$
$$= \chi D\Phi|_{(g_1, \pi_1)}((1-\chi)(g_2 - g_1, \pi_2 - \pi_1)) + \chi Q_{(g_1, \pi_1)}((1-\chi)(g_2 - g_1, \pi_2 - \pi_1))$$
$$+ (1-\chi)D\Phi|_{(g_2, \pi_2)}(\chi(g_1 - g_2, \pi_1 - \pi_2)) + (1-\chi)Q_{(g_2, \pi_2)}(\chi(g_1 - g_2, \pi_1 - \pi_2)).$$

The estimate (1) follows by (A.2).

The weighted estimate (2) follows by Lemma A.1. For example, in analyzing the preceding equation, the following term appears

$$\chi(1-\chi)\left(D\Phi|_{(g_2, \pi_2)}(g_1 - g_2, \pi_1 - \pi_2) - D\Phi|_{(g_1, \pi_1)}(g_1 - g_2, \pi_1 - \pi_2)\right)$$

and satisfies the desired estimate. $\square$

We also need to estimate the Taylor expansion of the modified constraint map for the iteration scheme. For a fixed vector field $W \in C^{k,\alpha}(\overline{\Omega})$, the Taylor expansion of the modified map $\Phi^W_{(g,\pi)}$ at (γ, τ) is

$$\Phi^W_{(g,\pi)}(\gamma + h, \tau + w) = \Phi^W_{(g,\pi)}(\gamma, \tau) + D\Phi^W_{(g,\pi)}|_{(\gamma,\tau)}(h, w) + Q^W_{(g,\pi),(\gamma,\tau)}(h, w).$$

In local coordinates, the linearized equation and the quadratic error term have a similar type of expressions as those for the usual constraint map. By direct analysis, we have the following estimates.

Lemma A.3. *There is a constant D depending locally uniformly on $(g, \pi), (\gamma, \tau), (\gamma', \tau'), (h, w) \in C^{2,\alpha}(\overline{\Omega}) \times C^{2,\alpha}(\overline{\Omega})$ and $W \in C^{1,\alpha}(\overline{\Omega})$ such that*

$$\|D\Phi^W_{(g,\pi)}|_{(\gamma,\tau)}(h, w) - D\Phi^W_{(g,\pi)}|_{(\gamma',\tau')}(h, w)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq D\|(h, w)\|_{\mathcal{B}_2 \times \mathcal{B}_2}\|(\gamma - \gamma', \tau - \tau')\|_{\mathcal{B}_2 \times \mathcal{B}_2}$$
$$\|Q^W_{(g,\pi),(\gamma,\tau)}(h, w)\|_{\mathcal{B}_0 \times \mathcal{B}_1} \leq D\|(h, w)\|^2_{\mathcal{B}_2 \times \mathcal{B}_2}.$$

Appendix B. Asymptotically flat initial data sets

Let B be a closed ball in $\mathbb{R}^3$. For every $k \in \{0, 1, \ldots\}$, $\alpha \in (0, 1)$, and $q \in \mathbb{R}$ we define the norm $C^{k,\alpha}_{-q}(\mathbb{R}^3 \setminus B)$ for $f \in C^{k,\alpha}_{\mathrm{loc}}(\mathbb{R}^3 \setminus B)$ as

$$\|f\|_{C^{k,\alpha}_{-q}(\mathbb{R}^3 \setminus B)} = \sum_{|I| \leq k} \sup_{x \in \mathbb{R}^3 \setminus B} \left| |x|^{|I|+q}(\partial^I f)(x) \right| + \sum_{|I|=k} \left[|x|^{k+q+\alpha}(\partial^I f)(x) \right]_{\alpha, \mathbb{R}^3 \setminus B}.$$

Let M be a smooth manifold such that there is a compact set $K \subset M$ and a diffeomorphism $M \setminus K \cong \mathbb{R}^3 \setminus B$. The $C^{k,\alpha}_{-q}$ norm on M is defined by taking the maximum of the $C^{k,\alpha}_{-q}(\mathbb{R}^3 \setminus B)$ norm and the $C^{k,\alpha}$ norm on the compact set K. The weighted Hölder space $C^{k,\alpha}_{-q}(M)$ is the collection of those $f \in C^{k,\alpha}_{\mathrm{loc}}(M)$ with finite $C^{k,\alpha}_{-q}(M)$ norm.

Let $q > \frac{1}{2}$, $q_0 > 0$. We say that an initial data set (M, g, π) is *asymptotically flat at the rate* (q, q_0) if there is a compact set $K \subset M$ and a diffeomorphism $M \setminus K \cong \mathbb{R}^3 \setminus B$ for a closed ball $B \subset \mathbb{R}^3$ such that

$$(g - \delta_{\mathbb{E}}, \pi) \in C^{2,\alpha}_{-q}(M) \times C^{1,\alpha}_{-q-1}(M),$$

where $\delta_{\mathbb{E}}$ is a smooth symmetric $(0, 2)$ tensor that coincides with the Euclidean metric $g_{\mathbb{E}}$ on $M \setminus K \cong \mathbb{R}^3 \setminus B$, and such that

$$\mu, J \in C^{0,\alpha}_{-3-q_0}(M).$$

Our definition focuses on the analysis of one asymptotically flat end, but can obviously accommodate M with multiple asymptotically flat ends.

For an asymptotically flat initial data set (M, g, π), one can define the following boundary integral, for a function N and a vector field X,

$$B^r_{(g,\pi)}(N, X) = \int_{|x|=r} \sum_{i,j=1}^{3} \left[N\left(g_{ij,i} - g_{ii,j}\right) - \left(N_{,i}g_{ij} - N_{,j}g_{ii}\right) + X^i \pi_{ij} \right] \nu_0^j \, d\sigma_0.$$

Here, the integrals are computed in the coordinate chart $M \setminus K \cong_x \mathbb{R}^3 \setminus B$, $\nu_0^j = x^j/|x|$, and $d\sigma_0$ is 2-dimensional Euclidean Hausdorff measure. The ADM energy E, linear momentum P, center of mass[1] $\mathcal{C}$, and angular momentum $\mathcal{J}$ are defined by

$$E = \tfrac{1}{16\pi} \lim_{r \to \infty} B^r_{(g,\pi)}(1, 0)$$

$$P_i = \tfrac{1}{8\pi} \lim_{r \to \infty} B^r_{(g,\pi)}\left(0, \frac{\partial}{\partial x^i}\right)$$

$$\mathcal{C}_i = \tfrac{1}{16\pi} \lim_{r \to \infty} B^r_{(g,\pi)}(x^i, 0)$$

$$\mathcal{J}_k = \tfrac{1}{8\pi} \lim_{r \to \infty} B^r_{(g,\pi)}\left(0, x \times \frac{\partial}{\partial x^k}\right),$$

where $i, k = 1, 2, 3$, and $x \times \frac{\partial}{\partial x^k}$ is the cross product. The well-definedness of those quantities can be found in, e.g., [1, 14, 15].

For an asymptotically flat initial data set (g, π), by Taylor expansion (A.1) at the flat data $(g_{\mathbb{E}}, 0)$,

$$\Phi(g, \pi) = D\Phi|_{(g_{\mathbb{E}}, 0)}(g - g_{\mathbb{E}}, \pi) + Q_{(g_{\mathbb{E}}, 0)}(g - g_{\mathbb{E}}, \pi).$$

---

[1]We remark that (for $E \neq 0$) $\mathcal{C}$ is sometimes written $\mathcal{C} = Ec$, where c is the center of mass in other references, e.g. [2]

For (N, X) so that $D\Phi|_{(g_\mathbb{E}, 0)}^*(N, X)$ vanishes, we have

(B.1) $$\int_{\{R_1 \leq |x| \leq R_2\}} D\Phi|_{(g_\mathbb{E}, 0)}(g - g_\mathbb{E}, \pi) \cdot (N, X) \, dx = B_{(g,\pi)}^{R_2}(N, X) - B_{(g,\pi)}^{R_1}(N, X).$$

Lemma B.1. *Let (g, π) be an asymptotically flat initial data set on $\mathbb{R}^3 \setminus B$ at the rate $q = q_0 = 1$. Let (g^θ, π^θ) be an admissible family for (g, π). Consider the initial data set $(\bar{g}^R, \bar{\pi}^R)$ on A_1 obtained in Proposition 4.4 (with $(g_1, \pi_1) = (g, \pi), (g_2, \pi_2) = (g^\theta, \pi^\theta)$):*

$$(\bar{g}^R, \bar{\pi}^R) = \chi(g^R, \pi^R) + (1 - \chi)((g^\theta)^R, (\pi^\theta)^R) + (h^R, w^R).$$

Then there is a constant C such that for $\theta \in \Theta_1 \times \Theta_2^R$ and for R sufficiently large,

$$\left| R \int_{A_1} \Phi(\bar{g}^R, \bar{\pi}^R) \cdot (1, 0) \, dx - 16\pi(E^\theta - E) \right| \leq CR^{-1}$$

$$\left| R \int_{A_1} \Phi(\bar{g}^R, \bar{\pi}^R) \cdot (0, \frac{\partial}{\partial x^i}) \, dx - 8\pi(P^\theta - P) \right| \leq CR^{-1}$$

$$\left| R \int_{A_1} \Phi(\bar{g}^R, \bar{\pi}^R) \cdot (x^k, 0) \, dx - 16\pi R^{-1}(\mathcal{C}_k^\theta - \mathcal{C}_k^R) \right| \leq C(R^{-1} + |\theta|^2 R^{-2})$$

$$\left| R \int_{A_1} \Phi(\bar{g}^R, \bar{\pi}^R) \cdot (0, x \times \frac{\partial}{\partial x^\ell}) \, dx - 8\pi R^{-1}(\mathcal{J}_\ell^\theta - \mathcal{J}_\ell^R) \right| \leq C(R^{-1} + |\theta|^2 R^{-2}).$$

Proof. By (4.4) and the estimate for (h^R, π^R) in Proposition 4.4, the quadratic error terms are estimated as follows:

$$\left| R \int_{A_1} Q_{(g_\mathbb{E}, 0)}(\bar{g}^R - g_\mathbb{E}, \bar{\pi}^R) \cdot (N, X) \, dx \right| \leq CR \|(\bar{g}^R - g_\mathbb{E}, \bar{\pi}^R)\|_{C^2 \times C^1}^2 \leq CR^{-1}.$$

We now estimate the integrals of the linearized operator. The energy and momentum integrals can be estimated similarly, so we only show the one for the energy. By (B.1),

$$R \int_{A_1} D\Phi(\bar{g}^R, \bar{\pi}^R) \cdot (1, 0) \, dx = R \left(B_{((g^\theta)^R, (\pi^\theta)^R)}^2(1, 0) - B_{(g^R, \pi^R)}^1(1, 0) \right)$$

$$= B_{(g^\theta, \pi^\theta)}^{2R}(1, 0) - B_{(g, \pi)}^R(1, 0),$$

where we note that the rescaling in the last line accounts for the factor R. It is standard to relate the boundary integral to the ADM energy at infinity by the divergence theorem. In particular, the uniformity assumption (4.4) implies that there is a constant C such that for $\theta \in \Theta_1 \times \Theta_2^R$,

$$\left| B_{(g^\theta, \pi^\theta)}^{2R}(1, 0) - 16\pi E^\theta \right| \leq CR^{-1}.$$

The proofs for the center of mass and angular momentum integrals are similar, so we only show prove the one for the center of mass. By (B.1) we have

$$R \int_{A_1} D\Phi(\bar{g}^R, \bar{\pi}^R) \cdot (x^k, 0) \, dx = R \left(B_{((g^\theta)^R, (\pi^\theta)^R)}^2(x^k, 0) - B_{(g^R, \pi^R)}^1(x^k, 0) \right)$$

$$= R^{-1} \left(B_{(g^\theta, \pi^\theta)}^{2R}(x^k, 0) - B_{(g, \pi)}^R(x^k, 0) \right),$$

where the rescaling in the last line gives an extra factor R^{-1} from the rescaling of the coordinate function x^k. To obtain the desired estimate, we see that the term $B_{(g^\theta, \pi^\theta)}^{2R}(x^k, 0)$ is estimated by the uniformity (4.5) and $B_{(g, \pi)}^R(x^k, 0) = 16\pi \mathcal{C}_k^R$ by definition. $\square$

APPENDIX C. INTERIOR SCHAUDER ESTIMATES

Let $x \in \Omega$ be fixed. Consider the ball $B_{\phi(x)}(x)$ centered at x of radius $\phi(x)$, where $\phi(x)$ is the weight function defined in Section 2.5. We blur the distinction between $B_{\phi(x)}(x)$ and its coordinate image, and we consider the diffeomorphism $F_x : B_1(0) \to B_{\phi(x)}(x)$ by $z \mapsto x + \phi(x)z = y$, where $B_1(0)$ is the unit ball in $\mathbb{R}^n$ centered at the origin. For any function f defined on $B_{\phi(x)}(x)$, let

$$\widetilde{f}(z) = F_x^*(f)(z) = f \circ F_x(z)$$

denote the pull-back of f on $B_1(0)$.

With a minor abuse of notation, we denote for $a \in (0,1]$,

$$\|f\|_{C_{\phi,\varphi}^{k,\alpha}(B_{a\phi(x)}(x))} = \sum_{j=0}^{k} \varphi(x)\phi^j(x)\|\nabla^j f\|_{C^0(B_{a\phi(x)}(x))} + \varphi(x)\phi^{k+\alpha}(x)[\nabla^k f]_{0,\alpha;B_{a\phi(x)}(x)}.$$

One can easily obtain the following lemma.

Lemma C.1. *Let f and g be functions defined on $B_{\phi(x)}(x)$. The following properties hold.*

(1) $\widetilde{f+g} = \widetilde{f} + \widetilde{g}$ and $\widetilde{fg} = \widetilde{f}\widetilde{g}$.

(2) $\widetilde{\partial_y^\beta f} = (\phi(x))^{-|\beta|}\partial_z^\beta \widetilde{f}$, where $\beta = (\beta_1, \ldots, \beta_k)$ is a multi-index, $\partial_y^\beta = \partial_{y^{i_1}}^{\beta_1} \cdots \partial_{y^{i_k}}^{\beta_k}$, $i_1, \ldots, i_k \in \{1, 2, \ldots, n\}$, and ∂_z^β is defined analogously.

(3) *For any $a \in (0,1]$,*

$$\|\varphi(x)\widetilde{f}\|_{C^{k,\alpha}(B_a(0))} = \|f\|_{C_{\phi,\varphi}^{k,\alpha}(B_{a\phi(x)}(x))}$$

$$\|\varphi(x)\widetilde{f}\|_{L^2(B_a(0))} = \|f\|_{L^2_{\phi^{-n}\varphi^2}(B_{a\phi(x)}(x))}.$$

Let $U = (U^1, U^2, U^3, \ldots, U^{n+1})$ where each U^i is a function defined on $B_{\phi(x)}(x)$. Consider the partial differential system LU whose j-th component is

$$(LU)_j = \sum_{k=1}^{n+1} \sum_{|\beta|=0}^{s_j+t_k} b_{jk}^\beta \partial_y^\beta U^k,$$

where

$$s_1 = 0, \quad t_1 = 4, \qquad s_j = -1, \quad t_k = 3 \qquad (j, k = 2, \ldots, n+1).$$

Theorem C.2. *Suppose that the operator L is strictly elliptic in the sense of* [10]. *For any $r, s \in \mathbb{R}$, let $\varphi_j = \phi^{r+4-t_j}\rho^s$. Then for each $k \in \{1, 2, \ldots, n+1\}$,*

(C.1)

$$\|U^k\|_{C_{\phi,\varphi_k}^{t_k,\alpha}(B_{\phi(x)/2}(x))}$$

$$\leq C \left(\sum_{j=1}^{n+1} \|(LU)_j\|_{C_{\phi,\phi^{t_j+s_j}\varphi_j}^{-s_j,\alpha}(B_{\phi(x)}(x))} + \sum_{j=1}^{n+1} \|U^j\|_{L^2_{\phi^{-n}\varphi_j^2}(B_{\phi(x)}(x))} \right),$$

where C depends only on n, α, $\displaystyle\sup_{\substack{j,k=1,\ldots,n+1 \\ |\beta| \leq s_j+t_k}} \|b_{jk}^\beta\|_{C_{\phi,\varphi^{s_j+t_k-|\beta|}}^{\ell-s_j,\alpha}(B_{\phi(x)}(x))}$, and the lower bound of ellipticity of the operator L.

Proof. By Lemma C.1, for $j = 1, \ldots, n+1$,

$$\widetilde{(LU)}_j = \sum_{k=1}^{n+1} \sum_{|\beta|=0}^{s_j+t_k} \widetilde{b_{jk}^\beta} (\phi(x))^{-|\beta|} \partial_z^\beta \widetilde{U^k}$$

$$= \sum_{k=1}^{n+1} \sum_{|\beta|=0}^{s_j+t_k} (\phi(x))^{-4+t_k-|\beta|} \widetilde{b_{jk}^\beta} \partial_z^\beta ((\phi(x))^{4-t_k} \widetilde{U^k}).$$

Multiplying $(\phi(x))^{4+s_j}$ to the above identity, we have

$$(\phi(x))^{4+s_j} \widetilde{(LU)}_j = \sum_{k=1}^{n+1} \sum_{|\beta|=0}^{s_j+t_k} (\phi(x))^{s_j+t_k-|\beta|} \widetilde{b_{jk}^\beta} \partial_z^\beta ((\phi(x))^{4-t_k} \widetilde{U^k}).$$

We remark the power of ϕ is specifically chosen such that, for our application to Theorem 5.6, each of the coefficients $(\phi(x))^{s_j+t_k-|\beta|} \widetilde{b_{jk}^\beta}$ is bounded in the $C^{-s_j,\alpha}(B_1(0))$ norm. Now we apply the Schauder interior estimate and obtain

$$\sum_{j=1}^{n+1} \|(\phi(x))^{4-t_j} \widetilde{U^j}\|_{C^{t_j,\alpha}(B_{\frac{1}{2}}(0))}$$

$$\leq C \left(\sum_{j=1}^{n+1} \|(\phi(x))^{4+s_j} \widetilde{L_j U}\|_{C^{-s_j,\alpha}(B_1(0))} + \sum_{j=1}^{n+1} \|(\phi(x))^{4-t_j} \widetilde{U^j}\|_{L^2(B_1(0))} \right),$$

where C depends only on n, α, $\displaystyle\sup_{\substack{j,k=1,\ldots,n+1 \\ |\beta| \leq s_j+t_k}} \|(\phi(x))^{s_j+t_k-|\beta|} \widetilde{b_{jk}^\beta}\|_{C^{-s_j,\alpha}(B_1(0))}$, and the lower bound of ellipticity of the operator L. By Lemma C.1 (3),

$$\|(\phi(x))^{s_j+t_k-|\beta|} \widetilde{b_{jk}^\beta}\|_{C^{-s_j,\alpha}(B_1(0))} = \|b_{jk}^\beta\|_{C_{\phi,\phi^{s_j+t_k-|\beta|}}^{-s_j,\alpha}(B_{\phi(x)}(x))}.$$

For $r, s \in \mathbb{R}$, let $\varphi_j = \phi^{r+4-t_j} \rho^s$. Multiplying $\phi^r(x) \rho^s(x)$ to the above inequality, we have

$$\sum_{j=1}^{n+1} \|\varphi_j(x) \widetilde{U^j}\|_{C^{t_j,\alpha}(B_{\frac{1}{2}}(0))}$$

$$\leq C \left(\sum_{j=1}^{n+1} \|(\phi(x))^{t_j+s_j} \varphi_j(x) \widetilde{(LU)}_j\|_{C^{-s_j,\alpha}(B_1(0))} + \sum_{j=1}^{n+1} \|\varphi_j(x) \widetilde{U^j}\|_{L^2(B_1(0))} \right).$$

Then Lemma C.1 (3) implies the desired estimate. $\qquad\square$

Proof of Theorem 5.8. By the definition of the weighted Hölder norm and by taking the supremum of (C.1) among $x \in \Omega$, it suffices to prove that for any u and $\varphi = \phi^r \rho^s$

$$\sup_{x \in \Omega} \|f\|_{C_{\phi,\varphi}^{k,\alpha}(B_{\phi(x)}(x))} \leq C \sup_{x \in \Omega} \|f\|_{C_{\phi,\varphi}^{k,\alpha}(B_{\phi(x)/2}(x))}$$

where C denotes a positive constant that depends only on k, α, r, s, and the constant in (2.15). The proof is a straightforward computation using (2.15). $\qquad\square$

Appendix D. Iteration scheme

We now prove Lemma 5.11. The proof follows essentially [6, Lemma 3.5].

Proof. By the induction hypothesis (5.13), (γ_m, τ_m) satisfies that $\Phi^W_{(g,\pi)}(g,\pi) + (\psi, V) - \Phi^W_{(g,\pi)}(\gamma_m, \tau_m) \in L^2_{\rho-1} \times L^2_{\rho-1}$. Use the variational result Theorem 5.5, we find (f_m, X_m) such that $(h_m, w_m) = \rho_g (D\Phi^W_{(g,\pi)}|_{(g,\pi)})^*(f_m, X_m)$ satisfies

$$D\Phi^W_{(g,\pi)}|_{(g,\pi)}(h_m, w_m) = \Phi^W_{(g,\pi)}(g,\pi) + (\psi, V) - \Phi^W_{(g,\pi)}(\gamma_m, \tau_m).$$

By Theorem 5.6 and the induction hypothesis (5.13),

$$\|(f_m, X_m)\|_{\mathcal{B}_4 \times \mathcal{B}_3} \le C\|\Phi^W_{(g,\pi)}(g,\pi) + (\psi, V) - \Phi^W_{(g,\pi)}(\gamma_m, \tau_m)\|_{\mathcal{B}_0 \times \mathcal{B}_1}$$
$$\le C\|(\psi, V)\|^{1+m\delta}_{\mathcal{B}_0 \times \mathcal{B}_1}$$
$$\|(h_m, w_m)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \le C\|\Phi^W_{(g,\pi)}(g,\pi) + (\psi, V) - \Phi^W_{(g,\pi)}(\gamma_m, \tau_m)\|_{\mathcal{B}_0 \times \mathcal{B}_1}$$
$$\le C\|(\psi, V)\|^{1+m\delta}_{\mathcal{B}_0 \times \mathcal{B}_1}.$$

This gives the desired estimates (5.12) for $p = m$.

To prove (5.13) for $j = m + 1$, we note that by Taylor expansion,

$$\Phi^W_{(g,\pi)}(\gamma_{m+1}, \tau_{m+1})$$
$$= \Phi^W_{(g,\pi)}(\gamma_m, \tau_m) + D\Phi^W_{(g,\pi)}|_{(\gamma_m, \tau_m)}(h_m, w_m) + Q^W_{(\gamma_m, \tau_m)}(h_m, w_m)$$
$$= \Phi^W_{(g,\pi)}(g,\pi) + (\psi, V) - D\Phi^W_{(g,\pi)}|_{(g,\pi)}(h_m, w_m)$$
$$\quad + D\Phi^W_{(g,\pi)}|_{(\gamma_m, \tau_m)}(h_m, w_m) + Q^W_{(\gamma_m, \tau_m)}(h_m, w_m)$$
$$= \Phi^W_{(g,\pi)}(g,\pi) + (\psi, V)$$
$$\quad + \sum_{p=0}^{m-1} \Big[D\Phi^W_{(g,\pi)}|_{(\gamma_{p+1}, \tau_{p+1})}(h_m, w_m) - D\Phi^W_{(g,\pi)}|_{(\gamma_p, \tau_p)}(h_m, w_m) \Big]$$
$$\quad + Q^W_{(\gamma_m, \tau_m)}(h_m, w_m).$$

By Lemma A.3,

$$\|\Phi^W_{(g,\pi)}(g,\pi) + (\psi, V) - \Phi^W_{(g,\pi)}(\gamma_{m+1}, \tau_{m+1})\|_{\mathcal{B}_0 \times \mathcal{B}_1}$$
$$\le D \left(\|(h_m, w_m)\|^2_{\mathcal{B}_2 \times \mathcal{B}_2} + \|(h_m, w_m)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \sum_{p=0}^{m-1} \|(\gamma_{p+1}, \tau_{p+1}) - (\gamma_p, \tau_p)\|_{\mathcal{B}_2 \times \mathcal{B}_2} \right)$$
$$\le DC^2 \left(\|(\psi, V)\|^{2+2m\delta}_{\mathcal{B}_0 \times \mathcal{B}_1} + \|(\psi, V)\|^{2+m\delta}_{\mathcal{B}_0 \times \mathcal{B}_1} \sum_{p=0}^{m-1} \|(\psi, V)\|^{p\delta}_{\mathcal{B}_0 \times \mathcal{B}_1} \right)$$
$$\le 2DC^2 \epsilon^{1-\delta}(1-\epsilon^\delta)^{-1}\|(\psi, V)\|^{1+(m+1)\delta}_{\mathcal{B}_0 \times \mathcal{B}_1}.$$

Choose $\epsilon > 0$ small enough so that $2DC^2 \epsilon^{1-\delta}(1-\epsilon^\delta)^{-1} \le 1$. $\qquad\square$

References

1. Robert Bartnik, *The mass of an asymptotically flat manifold*, Comm. Pure Appl. Math. **39** (1986), no. 5, 661–693. MR 849427 (88b:58144)

2. Piotr T. Chruściel, Justin Corvino, and James Isenberg, *Construction of N-body initial data sets in general relativity*, Comm. Math. Phys. **304** (2011), no. 3, 637–647. MR 2794541

3. Piotr T. Chruściel and Erwann Delay, *On mapping properties of the general relativistic constraints operator in weighted function spaces, with applications*, Mém. Soc. Math. Fr. (N.S.) (2003), no. 94, vi+103. MR 2031583 (2005f:83008)

4. Piotr T. Chruściel, James Isenberg, and Daniel Pollack, *Initial data engineering*, Comm. Math. Phys. **257** (2005), no. 1, 29–42. MR 2163567

5. Justin Corvino, *Scalar curvature deformation and a gluing construction for the Einstein constraint equations*, Comm. Math. Phys. **214** (2000), no. 1, 137–189. MR 1794269 (2002b:53050)

6. Justin Corvino, Michael Eichmair, and Pengzi Miao, *Deformation of scalar curvature and volume*, Math. Ann. **357** (2013), no. 2, 551–584. MR 3096517

7. Justin Corvino and Richard Schoen, *On the asymptotics for the vacuum Einstein constraint equations*, J. Differential Geom. **73** (2006), no. 2, 185–217. MR 2225517 (2007e:58044)

8. Justin Corvino and Haotian Wu, *On the center of mass of isolated systems*, Classical Quantum Gravity **25** (2008), no. 8, 085008, 18. MR 2409248 (2009e:83007)

9. Erwann Delay, *Localized gluing of Riemannian metrics in interpolating their scalar curvature*, Differential Geom. Appl. **29** (2011), no. 3, 433–439. MR 2795849 (2012f:53057)

10. Avron Douglis and Louis Nirenberg, *Interior estimates for elliptic systems of partial differential equations*, Comm. Pure Appl. Math. **8** (1955), 503–538. MR 0075417 (17,743b)

11. Michael Eichmair, Lan-Hsuan Huang, Dan Lee, and Richard Schoen, *The spacetime positive mass theorem in dimensions less than eight*, J. Eur. Math. Soc. (JEMS) **18** (2016), no. 1, 83–121. MR 3438380

12. Arthur E. Fischer and Jerrold E. Marsden, *Linearization stability of the Einstein equations*, Bull. Amer. Math. Soc. **79** (1973), 997–1003. MR 0426035 (54 #13981)

13. ———, *Deformations of the scalar curvature*, Duke Math. J. **42** (1975), no. 3, 519–547. MR 0380907

14. Lan-Hsuan Huang, *On the center of mass of isolated systems with general asymptotics*, Classical Quantum Gravity **26** (2009), no. 1, 015012, 25. MR 2470255

15. ———, *Foliations by stable spheres with constant mean curvature for isolated systems with general asymptotics*, Comm. Math. Phys. **300** (2010), no. 2, 331–373. MR 2728728 (2012a:53045)

16. Lan-Hsuan Huang, Richard Schoen, and Mu-Tao Wang, *Specifying angular momentum and center of mass for vacuum initial data sets*, Comm. Math. Phys. **306** (2011), no. 3, 785–803. MR 2825509 (2012j:83009)

17. James Isenberg, Rafe Mazzeo, and Daniel Pollack, *On the topology of vacuum spacetimes*, Ann. Henri Poincaré **4** (2003), no. 2, 369–383. MR 1985777 (2004h:53053)

18. Vincent Moncrief, *Spacetime symmetries and linearization stability of the Einstein equations. I*, J. Mathematical Phys. **16** (1975), 493–498. MR 0363398 (50 #15836)

19. Richard Schoen and Shing-Tung Yau, *On the proof of the positive mass conjecture in general relativity*, Comm. Math. Phys. **65** (1979), no. 1, 45–76. MR 526976 (80j:83024)

20. ———, *Proof of the positive mass theorem. II*, Comm. Math. Phys. **79** (1981), no. 2, 231–260. MR 612249 (83i:83045)

DEPARTMENT OF MATHEMATICS, LAFAYETTE COLLEGE, EASTON, PA 18042, USA
E-mail address: `corvinoj@lafayette.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CONNECTICUT, STORRS, CT 06269, USA
E-mail address: `lan-hsuan.huang@uconn.edu`