THE GAUGE-INvariant I-METHOD FOR YANG-MILLS

Cristian Gavrus

Abstract. We prove global well-posedness of the 3d Yang-Mills equation in the temporal gauge in $H^\sigma$ for $\sigma > \frac{5}{6}$.

Unlike related equations, Yang-Mills is not directly amenable to the method of almost conservation laws (I-method) in its Fourier and global version. We propose a modified energy which:

1. Is gauge-invariant and easy to localize
2. Provides local gauges which give control of local Sobolev norms (through an Uhlenbeck-type lemma for fractional regularities)
3. Is slightly smoother in time compared to the classical I-method energy for related systems.

The spatial smoothing is realized via the Yang-Mills heat flow instead of the multiplier $I$.

Due to the temporal condition and its finite speed of propagation, the local gauge selection is compatible with recent initial data extension results. Therefore, smoothed energy differences $\mathcal{E}(t_1, s) - \mathcal{E}(t_0, s)$ can be partitioned into local pieces whose (appropriately extended) bounds can be square summed. After revealing the null structure within the trilinear integrals, these can be estimated using known methods.

In an appendix we show how an invariant modified energy for Maxwell-Klein-Gordon can extend previous results to regularities $\sigma > \frac{5}{6}$.

Contents

1. Introduction 1
2. Preliminaries 11
3. Local well-posedness for Yang-Mills 17
4. Modified local well-posedness 20
5. Yang-Mills Heat Flow solutions 22
6. Energy estimates for the curvature and the tension field 24
7. The local change of gauge: Proof of Proposition 1.10 31
8. Local differences of heat flows which coincide on a ball 38
9. Decompositions and estimates in DeTurck’s gauge 43
10. Trilinear estimates 48
Appendix A. Modified energy for Maxwell-Klein-Gordon 53
References 56

1. Introduction

In this paper we consider the problem of low regularity global existence for the initial value problem for the Yang-Mills equations on $\mathbb{R}^{3+1}$ with a non-abelian structural group $G$. We say that the connection $A_{\alpha} : \mathbb{R}^{3+1} \to \mathfrak{g}$ solves the Yang-Mills equation if

$$D^\alpha F_{\alpha \beta} = 0 \quad (YM)$$

The author thanks Ben Dodson for suggesting the paper [22] and for discussions on the I-method, and thanks Sung-Jin Oh for discussions on his Yang-Mills papers. The author expresses gratitude to the John Hopkins University, where a large part of this work was carried.
where

\[ F_{\alpha\beta} := \partial_\alpha F_\beta - \partial_\beta F_\alpha + [A_\alpha, A_\beta] \quad \text{and} \quad D_\alpha B := \partial_\alpha B + [A_\alpha, B] \]

denote the \textit{curvature} and the \textit{covariant derivative}, while \( \mathfrak{g} \) is the associated Lie algebra. The background on Lie theory notions used is presented in section 2.0.2.

Expanding, (YM) takes the form

\[ \partial^\alpha F_{\alpha\beta} + [A^\alpha, \partial_\alpha A_\beta] = 0 \]

An important feature of the (YM) equations is its \textit{gauge invariance}. Given a \( \mathcal{G} \)-valued potential \( U \), (YM) is invariant under the following gauge transformations

\[ A_\alpha \mapsto UA_\alpha U^{-1} - \partial_\alpha UU^{-1}, \quad F_{\alpha\beta} \mapsto UF_{\alpha\beta}U^{-1}. \] (1.1)

For this reason, the Cauchy problem cannot be well-posed until a choice of gauge is made. If we impose the \textit{Temporal gauge} condition \( A_0 = 0 \) the equations become

\[ \Box A_i - \partial_i \text{div} A = \partial_j [A_i, A_j] + [\partial_j A_i, A_j] + [A_j, \partial_i A_j] - [A_j, [A_j, A_i]] \] (1.2)

These equations can be rewritten as (4.1) or (4.4).

The initial data sets consist of \( (\bar{A}_i, \bar{E}_i) = (A_i, F_{0i})(t = 0) \) for \( i = 1, 2, 3 \). Note that in the temporal gauge \( F_{0i} = \partial_t A_i \). Due to the first equation in (1.2), i.e. (YM) for \( \beta = 0 \), initial data sets have to satisfy the \textit{constraint} (or Gauss) equation

\[ \bar{D}^\ell \bar{E}_\ell := \partial^\ell \bar{E}_\ell + [\bar{A}^\ell, \bar{E}_\ell] = 0. \] (1.3)

The energy of a connection \( A_\alpha \) at \( t \) is defined by

\[ E[F_{\alpha\beta}](t) := \frac{1}{2} \sum_{\alpha<\beta} \int_{\mathbb{R}^3} (F_{\alpha\beta}(t), F_{\alpha\beta}(t)) \, dx \] (1.4)

and it is conserved for sufficiently smooth solutions of (YM).

1.1. Prior Yang-Mills results on \( \mathbb{R}^{3+1} \). The early works of Segal [46], Choquet-Bruhat and Christodoulou [3], and Eardley-Moncrief [16] establish local and global well-posedness with high regularity (e.g. \( H^2 \)).

Finite energy global well-posedness for (YM) was first proved by Klainerman and Machedon in [27]:

Theorem 1.1 ([27]). The Yang-Mills equation in the temporal gauge is globally well-posed for finite energy initial data which is locally in \( H^1 \times L^2 \).

The proof involves local Coulomb gauges based on Uhlenbeck’s lemma. Two other proofs appear in [38] and [40]: one is based on the Yang-Mills heat flow for gauge selection, the other proof rests on initial data surgery techniques. In this work we need to combine these three methods together with some new ones.

Removing the smallness assumption by shrinking the time interval is a delicate matter. This was obtained by Oh and Tătărău in [40] as a consequence of their initial data excision and extension techniques for the Gauss equation (1.3):
Theorem 1.3 ([40]). Let $\sigma > \frac{3}{4}$. The Yang-Mills equation in the temporal gauge is locally well-posed for initial data satisfying (1.3) and in $H^\sigma_{loc}(\mathbb{R}^3) \times H^{\sigma-1}_{loc}(\mathbb{R}^3)$, on a sufficiently small time interval.

By a square summability argument one can formulate this result for global Sobolev spaces as well (see section [3.1]):

Corollary 1.4 (Local well-posedness in $H^\sigma \times H^{\sigma-1}$). Let $\sigma > \frac{3}{4}$. Suppose the initial data $(\bar{A}_x, \bar{E}_x)$ is in $H^\sigma(\mathbb{R}^3) \times H^{\sigma-1}(\mathbb{R}^3)$ and satisfies (1.3). Then the local in time solution from Theorem 1.3 is in $C_t H^\sigma \cap C_t^1 H^{\sigma-1}$. For any $t$ in the time interval the solution map $(\bar{A}_x, \bar{E}_x) \mapsto A[t]$ is locally Lipschitz continuous on $H^\sigma \times H^{\sigma-1}$.

Persistence of higher regularity in $H^\sigma \times H^{\sigma-1}$ holds too ($\bar{\sigma} \geq \sigma$). Solutions solve (1.2) in the sense of distributions. Note that $[A, \partial A] \in H^{-1} \subset S^*$ due to (2.21).

We remark that extending Theorem 1.2 (and thus also 1.3 - 1.4) to $\sigma > \frac{3}{2}$ (which is the critical regularity) as in the case of Maxwell-Klein-Gordon [36] remains an open question.

1.2. Main result.

The purpose of this article is to obtain global solutions with regularities below the energy:

Theorem 1.5. Let $\sigma > \frac{5}{6}$. The Yang-Mills equation in the temporal gauge is globally well-posed in $H^\sigma(\mathbb{R}^3) \times H^{\sigma-1}(\mathbb{R}^3)$: the solutions in Corollary 1.4 extend to $C_t H^\sigma \cap C_t^1 H^{\sigma-1}(\mathbb{R}_t \times \mathbb{R}^3)$.

This result was motivated by [22] which proves a global result for the Maxwell-Klein-Gordon (MKG) equations in the Coulomb gauge for regularities $\sigma > \frac{\sqrt{3}}{2}$. Indeed, results on Yang-Mills usually follow in the footsteps of similar results for MKG, although the transition can often be highly non-trivial. The paper [22] uses the I-method which proceeds by inserting a smoothing Fourier multiplier operator $I = I_N$ (see (2.10)) into the MKG Hamiltonian (Energy) $H$. The main technical step is showing that $H[I\Phi(t)]$ is "almost conserved" - i.e. it varies very slowly in time. Letting $N \to \infty$ allows control of the size of the solution for long times. See also: [2], [23], [24], [3], [6], [7], [8], [9], [45], [14], [19]. For an introduction to the I-method for the cubic wave equation we refer to [14].

However, applying the usual I-method directly to (YM) is unlikely to work and a new approach is needed for the modified energy. We discuss the reasons below.

To obtain a global in time result for (YM) the (global or local) choices of gauge are essential. In particular, the energy $E[F](t)$ or modified energies (such as $E[I\Phi](t)$) should provide control of (global or local) Sobolev norms such as $\|\nabla_{t,x}^\alpha A_i\|_{L^2}$ or $\|I\nabla_{t,x}^\alpha A_i\|_{L^2}$. We must consider:

1. In general, the curvature $F_{ij}$ can control only the curl $\partial_i A_j - \partial_j A_i$ part of $\nabla_x A_i$. For instance, in the temporal gauge $F_{0i} = \partial_i A_i$ and $E[F](t) \approx \|\nabla_x A_{gf}^j(t)\|_{L^2}^2 + \|\partial_i A(t)\|_{L^2}^2$ so the energy gives no Sobolev control of the curl-free part $\nabla_x A_{gf}^j$. Setting the Coulomb condition $\text{div} A_x = 0$ would make $A_{gf}^j = 0$ automatically. However,

2. Global Coulomb gauges cannot be imposed for large data when $G$ is non-abelian [27], [17]. This is a fundamental difference to the MKG case [26], [22]. Local Coulomb gauges were used instead in [27] where Uhlenbeck’s lemma provided control of local $H^1$ norms.

3. Unless one fixes a global gauge (and does all the analysis in that gauge), setting the local gauges and patching the solutions together using finite speed of propagation (in the temporal gauge) requires the gauge invariance of the energy, [27], [40]. However,


(4) The classical I-method modified energy $E[F[IA]](t)$ or $E[IF](t)$ fails to be gauge-invariant or easy to localize, like the original energy. This is due to the presence of the Fourier multiplier $I$, see (2.10).

To deal with these considerations we propose a modified energy which:

I. Is gauge invariant and fairly easy to localize, addressing (3),(4)

II. Addresses issues (1),(2) through Proposition 1.10 (which is an Uhlenbeck-type lemma below $H^1$) providing gauge selections which give control of local Sobolev norms, see Remark 1.11.

III. Is smoother in time compared to the classical I-method modified energy $1$, slightly improving the numerology compared to the MKG case in [22], allowing lower regularities $\sigma > \frac{5}{4}$ rather than $\sigma > \frac{\sqrt{3}}{2}$, see Remark 1.14.

Instead of using the Fourier multiplier $I$ in the definition, the same degree of spatial smoothing will be realized geometrically using the Yang-Mills heat flow

$$F_{si} = D^s F_{ti} \quad \text{(cYMHF)}$$
and

$$F_{s\alpha} = D^s F_{t\alpha} \quad \text{(dYMHF)}$$

introduced by Donaldson [15] and developed extensively by Sung-Jin Oh in [37], [38] and in [39], [41] for the analysis of finite energy YM, along with its dynamic version (dYMHF), together with Oh’s related notion of caloric gauges for YM, paralleling the caloric gauges for Wave Maps [53] and Schrödinger Maps [49], [1] introduced by Tao. The smoothing is realized in a parabolic way, effectively implementing a covariant Littlewood-Paley decomposition, see Remark 1.12.

As an alternative to II one could address (1),(2) by opting for global in space caloric gauges like in [37], [38] instead of local ones. Unfortunately, that approach requires integration in time which would worsen the numerology and would also be more technically difficult to implement in the present context.

If $A_t : \mathbb{R} \times \mathbb{R}^3 \times [0,s_0] \to \mathfrak{g}$ is a regular connection, using the bi-invariant inner product from (2.1), we denote the energies at each $s \in [0,s_0]$ by

$$\mathcal{E}(t,s) := \frac{1}{2} \sum_{\alpha < \beta} \int_{\mathbb{R}^3} (F_{\alpha\beta}, F_{\alpha\beta})(t,s) \, dx$$

(1.5)

Definition 1.6 (Modified energy). Let $N^2 s_0 = 1$. We define the following modified energy

$$\mathcal{IE}(t) := \sup_{s \in [0,s_0]} (N^2 s)^{1-\sigma} \mathcal{E}(t,s) + \int_0^{s_0} (N^2 s)^{1-\sigma} \mathcal{E}(t,s) \frac{ds}{s}$$

(1.6)

The similarity of $\mathcal{IE}(t)$ with the classical modified energy $E[F[IA]](t)$ is discussed in Remark 1.13 based on the heat flow evolution relation to Littlewood-Paley theory (Remark 1.12).

The key to the proof of the main result is the following:

Proposition 1.7 (Almost conservation law). Let $\sigma > \frac{5}{6}$. Let $A_{t,x}$ be a global regular solution to (YM) in the temporal gauge. Suppose $A_{t,x,s}$ solves (dYMHF) on $\mathbb{R} \times \mathbb{R}^3 \times [0,s_0]$. Let $t_0$ be a time such that

$$\mathcal{IE}(t_0) \leq 2\eta^2, \quad \|A(t_0)\|_{H^{\sigma} \times H^{\sigma-1}} \leq M_0$$

1The structure of the differentiated energy effectively shifts a derivative from high to low frequency (Remark 1.14).
Then for any \( t \in [t_0, t_1] \), \( t_1 = t_0 + 1 \) and any \( s \in [0, s_0] \), \( s_0 = N^{-2} \) one has
\[
(N^2 s)^{-\sigma} |\mathcal{E}(t, s) - \mathcal{E}(t_0, s)| \lesssim \frac{s^{1/2}}{(Ns^2)^{(1-\sigma)}} \eta^2 \lesssim \frac{\eta^2}{N^{1/2}}
\]
as well as
\[
\|A[t]\|_{H^\sigma \times H^{\sigma-1}} \lesssim \eta^2/N^{1/2}.
\]
In particular, \(|\mathcal{IE}(t) - \mathcal{IE}(t_0)| \lesssim \eta^2/N^{1/2}\).

The error \( 1/N^{1/2} \) corresponds to the condition \( \sigma > \frac{5}{6} \), which appears to be the limit of this method in the current form.

Remark 1.8. A feature of the proof is that we do not prove the almost-conservation law for the whole solution at once (this is due to the fact that we only control the solution in local gauges). Instead, we partition \( \mathcal{E}(t_1, s) - \mathcal{E}(t_0, s) \) into local pieces (see (1.31)) and then we square sum the estimates obtained for the (appropriately extended) local solutions.

Finally, let
\[
w_\alpha(s) := D_\beta F_{\alpha\beta}(s)
\]
be the Yang-Mills tension field, which is a measure of the failure of \( A_\alpha(s) \) to satisfy the (YM) equation for \( s > 0 \), since the heat flow and (YM) do not exactly commute. The gauge transformation (1.1) makes \( w_\alpha \mapsto U w_\alpha U^{-1} \).

1.3. Overview of the paper and methods.

We now outline the key results which play a role in the proof and their organization.

Section 2 introduces the basic definitions, notations, properties and preliminaries.

Section 3 is devoted to local well-posedness for (YM). It begins by addressing \( H^\sigma \times H^{\sigma-1} \) approximations by regular initial data sets satisfying (1.3). Then in Prop. 3.2 we discuss square summability for fractional local Sobolev spaces \( \|u\|_{H^s}^2 \simeq \sum_j \|u_j\|^2_{W^{s,2}(B_j)} \). We continue by recalling initial data extension results from [40] before reviewing Theorem 1.3 and discussing Corollary 1.4.

Section 4 develops space-time control of temporal solutions in: hyperbolic \( X^{s,b} \) spaces for the divergence free part \( A^{df} = PA \) and in \( X^{r,\theta} \) spaces for the curl free part \( A^{cf} = P_\perp A \). As revealed in [50], small solutions can enjoy \( \frac{1}{4} \) more regularity for \( A^{cf} \) after a suitable change of gauge. Locally in time, by Proposition 4.1 one will have \( IA^{df} \in X^{1/4,1}, IA^{cf} \in H^{1/4} \times H^{1/4} \) provided one can somehow make the solution small.

Section 5 is reserved for the Yang-Mills heat flow in deTurck’s and caloric gauges including: well-posedness in \( I^{-1} H^1 \) and \( H^\rho \), control of the modified energy (1.6) from the initial data and the change between the two gauges as in [37] (deTurck’s trick).

A bound on the modified energy \( \mathcal{IE}(t) \) allows us to control higher covariant derivatives of \( F_{\alpha\beta}(s) \) in \( L^2 \). This is proved in Proposition 5.3 using covariant energy estimates for parabolic equations. The same is true for \( w_\alpha(s) \), see Section 6. To implement Remark 1.8, we show that these bounds can be localized while maintaining square summability:

**Proposition 1.9.** Let \( A_{t,x,s} \) be a regular solution to (YM) and (dYMHF) on an interval \( J \times \mathbb{R}^3 \times [0, s_0] \) satisfying \( \mathcal{IE}(t) \leq C \eta^2 \ll 1 \quad \forall \, t \in J \).
Let \((B_j)_{j \in J}\) be a finitely overlapping covering of \(\mathbb{R}^3\) by balls of radius \(\simeq 1\). Then there exist square summable coefficients \((c_j)_{j \in J}\) associated to \((B_j)_{j \in J}\) with

\[
\|(Ns^{1/2})^1-\sigma (s^{1/2}D_x)^m F_{\alpha \beta}(t, s)\|_{L^2_xL^2_t\cap L^2_{x,t}([0,s_0] \times B_j)} \leq c_j \tag{1.10}
\]

\[
\|(N^2 s^{1/2})^1-\sigma (s^{1/2}D_x)^m w_x(t, s)\|_{L^2_t L^2_x\cap L^2_{x,t}([0,s_0] \times B_j)} \leq c_j \tag{1.11}
\]

holding for all \(t \in J, j \in J, 0 \leq m \leq 8\) and

\[
\sum_{j \in J} c_j^2 \lesssim \eta^2. \tag{1.12}
\]

Once these coefficients are defined, section \(7\) shows that on each ball we can define local gauges using:

**Proposition 1.10.** Let \(B\) be a ball of radius \(\simeq 1\) and \(N^2 s_0 = 1\).

1. Let \(A_{s,t} \in (\text{YMHF})\) in the caloric gauge \(A_s = 0\) on \(B \times [0, s_0]\). There is a sufficiently small \(\delta > 0\) such that if the curvature \(F_{ij}\) obeys

\[
\sup_{0 \leq m \leq 3} \sup_{s \in [0, s_0]} (Ns^{1/2})^1-\sigma (s^{1/2}D_x)^m F(s)\|_{L^2(B)} \leq \delta \tag{1.13}
\]

then there exists a spatial field \(U : B \rightarrow G, U = U(x)\) such that the transformation

\[
\tilde{A}_i = U A_i U^{-1} - \partial_t U U^{-1} \tag{1.14}
\]

satisfies \(\tilde{A}_i(s_0) \in \delta H^1(B)\) and

\[
\delta^{-1} \tilde{A}_i(\cdot, 0) \in H^1(B) + N^{\sigma - 1} H^{\sigma}(B) \tag{1.15}
\]

2. Consider a smooth solution \(A_{t,x,s} \in (\text{YMHF})\) on \([t_0, t_1] \times B \times [0, s_0]\) in the temporal-caloric gauge \(A_0(t,x,0) = 0, A_s(t,x,s) = 0\) for which the curvature \(F_{\alpha \beta}(t_0)\) obeys \(1.13\). Apply Part (1) to \(A_{t,s_0,\cdot}\) and use \(U = U(x)\) to make the transformation \(1.14\) on \([t_0, t_1] \times B\). Then in addition to \(1.15\) one also has

\[
\delta^{-1} \partial_t \tilde{A}_i(t_0, \cdot, 0) \in L^2(B) + N^{\sigma - 1} H^{\sigma - 1}(B) \tag{1.16}
\]

**Remark 1.11.** This Proposition can be thought of as an Uhlenbeck-type lemma for fractional regularities below \(L^2\). It essentially says that if \(\|IF\|_{L^2(B)} < \delta\) then one can find gauge-equivalent \(\tilde{A} \sim A, \tilde{E} \sim E\) for which \(\|I \tilde{A}\|_{H^1(B)} \lesssim \delta\) and \(\|I\tilde{E}\|_{L^2(B)} \lesssim \delta\).

Of course, the multiplier \(I\) is not defined on a domain, so the remark above is simply a heuristic. However, once one has the bounds \(1.15, 1.16\) it is possible to use an extension result (Proposition 3.5) to extend \(\tilde{A}\) to the whole space \(\mathbb{R}^3\) such that \(I \tilde{A} \in \delta H^1, I \partial_t \tilde{A} \in \delta L^2\) while still maintaing the Gauss equation \(1.3\).

Once we restrict a solution to a ball \(B\) and perform the extension procedure, one obtains two heat flows: one for the original solution and one for the new one, which coincide at \(s = 0\) on \(B\). The need arises to compare them locally for parabolic times \(s > 0\). Due to infinite speed of parabolic propagation, they will not coincide, but one can obtain difference bounds of type \(\|\chi[F(s) - F'(s)]\|_{L^2(B)} \ll s^M\). This is proved by Proposition 8.1 to which Section 8 is devoted.

Eventually, the result is reduced to the space-time bound \((1.32)\) which represents a trilinear estimate proved in Proposition 10.1. Building upon the decompositions and estimates from Section
In the Appendix we formulate the invariant modified energy for Maxwell-Klein-Gordon and show how it can be used to increase the range of regularities for which global well-posedness is known to hold, from \( \sigma > \frac{\sqrt{3}}{2} \) in \([22]\) to \( \sigma > \frac{5}{6} \).

1.4. Reduction of Theorem 1.5 to the almost conservation law Prop. 1.7

Consider initial data sets of size \( \| (A_x, E_x) \|_{H^\sigma \times H^{\sigma-1}} \leq M \). Fix an arbitrary time \( T > 0 \). By the well-posedness statement in Corollary 1.4 it suffices to prove

\[
\| A[T] \|_{H^\sigma \times H^{\sigma-1}} \leq C_{M,T} < \infty
\]

for all such solutions which exist on \([0, T] \). Applying Corollary 1.4 again together with the approximation statement in Proposition 3.1 implies that it suffices to show (1.17) for regular solutions \( A \in C^\infty_t H^\infty_x (\mathbb{R} \times \mathbb{R}^3) \), which are known to be global.

We will choose a large number \( N = N(M, T) \) to be specified later, with \( T < M N^{0+} \) as usual with the I-method. Rescale the solution by \( A^\lambda(t, x) := \lambda^{-1} A(t/\lambda, x/\lambda) \) where we choose \( \lambda, N \) such that the rescaled initial data satisfies

\[
\| (A^\lambda, \dot{E}^\lambda) \|_{\dot{H}^\sigma \times \dot{H}^{\sigma-1}} \lesssim \lambda^{\frac{1}{2}-\sigma} M \lesssim \frac{\eta}{N^{1-\sigma}}, \quad \| \dot{A}^\lambda \|_{\dot{H}^{\frac{1}{2}+}} \lesssim \lambda^0 M \lesssim 1, \tag{1.18}
\]

so \( \lambda, N \) are related by

\[
\lambda = N^{\sigma - \frac{1}{2}} \left( \frac{CM}{\eta} \right)^{\frac{1}{2}} \tag{1.19}
\]

and \( \eta \ll 1 \) is a constant chosen small enough to overpower a number of universal constants in different estimates. Now we need to obtain a bound on \( A^\lambda[\lambda T] \).

We extend \( A^\lambda_{x,s}(t, x) \) to \( A^\lambda_{x,s}(t, x, s) \) by the heat flow using Theorem 5.9. We obtain a regular solution \( A^\lambda \in C^\infty_t H^\infty_x ([0, \lambda T] \times \mathbb{R}^3 \times [0, 1]) \) to (dYMHE) in the caloric gauge \( A^\lambda_s = 0 \), i.e.

\[
\partial_s A^\lambda_{\alpha} = D^\epsilon F^\lambda_{\epsilon \alpha}, \quad A^\lambda_\alpha(t, x, s = 0) = A^\lambda_\alpha(t, x)
\]

Let \( s_0 = N^{-2} \). Denote the energies (1.5) determined by \( A^\lambda_\alpha \) at \( (t, s) \) by \( \mathcal{E}(t, s) \) and consider the modified energy (1.7). Begin with the initial data:

By Corollary 5.7 with \( \ve \ll \eta N^{1-\sigma} \) and (1.19) we obtain \( \mathcal{E}(0) \leq \eta^2 \).

From this and \( O(\lambda T) \) application of Proposition 1.7 we obtain \( \mathcal{E}(t) \leq 2\eta^2 \) for all \( t \in [0, \lambda T] \) and \( \| A^\lambda[\lambda T] \|_{H^\sigma \times H^{\sigma-1}} < \infty \) provided that

\[
\lambda T \ll N^{\frac{1}{2}-}, \quad T \ll_M N^{\frac{1}{2} - 1 - \frac{\sigma}{1-\sigma}} = N^{0+}. \tag{1.20}
\]

It is seen that it is possible to choose \( N, \lambda \) satisfying (1.18), (1.19), (1.20) if \( \sigma > \frac{5}{6} \). Undoing the scaling we obtain (1.17).

1.5. Proof of the almost conservation Proposition 1.7

By a continuity argument one can assume that \( \mathcal{E}(t) \leq 3\eta^2 \) holds for all \( t \in [t_0, t_1] \).

Without loss of generality we prove (1.7) for \( t = t_1 \).

By Theorem 5.9 we may assume the caloric gauge condition \( A_s(t, x, s) = 0 \).
We note that by using Bianchi’s identity (2.3) and the Leibniz rule (2.2) one can write the differentiated energy at \( s \) as
\[
\frac{d}{dt} \mathcal{E}(t, s) = \sum_{\ell=1}^{3} \int_{\mathbb{R}^3} (w_\ell, F_{0\ell})(s) \, dx
\]
where \( w_\ell \) is given by (1.9) and therefore
\[
\mathcal{E}(t_1, s) - \mathcal{E}(t_0, s) = \sum_{\ell=1}^{3} \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (w_\ell(s), F_{0\ell}(s)) \, dx \, dt
\]
(1.21)
The strategy we use consists of localizing this integral and estimating it in appropriate local gauges, crucially using the fact that (1.21) is gauge-invariant.

Let \( (B_j^{(1)})_{j \in \mathcal{J}} \) be a finitely overlapping covering of \( \mathbb{R}^3 \) with balls of radius 1 centered at \( (y_j)_{j \in \mathcal{J}} \). Let \( (\chi_j)_{j \in \mathcal{J}} \) be a spatial partition of unity with \( \text{supp} \chi_j \subset B_j^{(1)} \).

Let \( (B_j)_{j \in \mathcal{J}} \) consist of balls with same centers \( (y_j)_{j \in \mathcal{J}} \) but radius 2, which are also finitely overlapping. Let \( (\mathcal{D}_j)_{j \in \mathcal{J}} \) be domains of dependency with bases \( (B_j)_{j \in \mathcal{J}} \), which are truncated cones
\[
\mathcal{D}_j = \{(t, x) \mid t - t_0 + |x - y_j| < 2, \ t \in [t_0, t_1]\}
\]
which form a covering of \([t_0, t_1] \times \mathbb{R}^3\). Note that \([t_0, t_1] \times B_j^{(1)} \subset \mathcal{D}_j\).

By Proposition 1.9, we define square summable coefficients \( (c_j)_{j \in \mathcal{J}} \) associated to \( (B_j)_{j \in \mathcal{J}} \) such that (1.10)-(1.11) and (1.12) hold.

For each \( j \in \mathcal{J} \) we invoke Propositions 1.10 on \( B_j \) with \( \delta = c_j \) which provide spatial \( U^{(j)} : B_j \rightarrow G \) and corresponding transformations on \([t_0, t_1] \times B_j \times [0, s_0]\):
\[
A^{(j)}_\alpha := U^{(j)} A_\alpha U^{(j)\dagger} - \partial_\alpha U^{(j)} U^{(j)\dagger}, \quad F^{(j)}_{\alpha\beta} = U^{(j)} F_{\alpha\beta} U^{(j)\dagger}
\]
(1.22)
with
\[
\|A^{(j)}(t_0, \cdot, s_0)\|_{H^1(B_j)} \lesssim c_j
\]
(1.23)
and
\[
\|A^{(j)}(t_0, \cdot, 0)\|_{H^1(B_j) + N^{s-1}H^s(B_j)} + \|\partial_t A^{(j)}(t_0, \cdot, 0)\|_{L^2(B_j) + N^{s-1}H^s(B_j)} \lesssim c_j
\]
Since the \( U^{(j)}(x) \) are independent of \( t \) and \( s \), the \( A^{(j)}_\alpha \)'s remain smooth solutions of \( (\text{YM}) \) and \( (\text{YM/H}) \) on \([t_0, t_1] \times B_j \times [0, s_0] \) in the temporal and caloric gauges \( A^{(j)}_\alpha(t, x, 0) = 0, \ A^{(j)}_\alpha(t, x, s) = 0 \). Due to gauge-invariance the bounds (1.10)-(1.11) hold for \( F^{(j)}_{\alpha\beta}, w^{(j)} \) too (replacing \( c_j \) by \( Cc_j \) if needed) and, moreover
\[
\chi_j(w_\ell(s), F_{0\ell}(s)) = \chi_j(w^{(j)}_\ell(s), F^{(j)}_{0\ell}(s)).
\]
(1.24)
We now use Proposition 3.5 to extend \( (A^{(j)}(t_0, \cdot, 0), \partial_t A^{(j)}(t_0, \cdot, 0)) \) from \( B_j \) to a regular \( (\tilde{a}^{(j)}, \tilde{e}^{(j)}) \) on \( \mathbb{R}^3 \) in such a way that the constrain equation (1.3) holds and we have the bound (see Lemma 2.1)
\[
\|(I\tilde{a}^{(j)}, I\tilde{e}^{(j)})\|_{H^1 \times L^2} \simeq \|\tilde{a}^{(j)}\|_{H^{1+N^{s-1}H^s}} + \|\tilde{e}^{(j)}\|_{L^2+N^{s-1}H^{s-1}} \lesssim c_j
\]
Let \( \tilde{A}^{(j)} \) be the regular Yang-Mills solution in the temporal gauge on \([t_0, t_1] \times \mathbb{R}^3 \) with initial data \( (\tilde{A}^{(j)}(t_0), \partial_t \tilde{A}^{(j)}(t_0)) = (\tilde{a}^{(j)}, \tilde{e}^{(j)}) \). By Proposition 4.1 Part (1):
\[
\|I(\tilde{A}^{(j)}(t))\|_{L^\infty_t(H^1 \times L^2)[[t_0, t_1] \times \mathbb{R}^3]} \lesssim c_j
\]
(1.25)
Due to finite speed of propagation we have that

\[ A^{(j)}(\cdot, 0)\big|_{\mathbb{D}_j} = \tilde{A}^{(j)} \]  

(1.26)

From this and (1.23) one has

\[ \|A^{(j)}[t_1]\|_{s=0}^{H^1 + N^s - 1, H^s}(B^{(j)}(t_1)) \lesssim c_j \]  

(1.27)

We now extend \( \tilde{A}^{(j)} \) from \([t_0, t_1] \times \mathbb{R}^3 \) to \([t_0, t_1] \times \mathbb{R}^3 \times [0, s_0] \) as the solution of the (dYMFF) in the caloric gauge \( \tilde{A}^{(j)} = 0 \) given by Theorem 5.9. By Proposition 5.10 and Remark in Proposition 6.3. Without loss of generality we prove (1.8) for all \( s \in [0, s_0] \), that

\[ \|\tilde{A}^{(j)}\|_{L_t^\infty L_x^{\infty}(H^1 + N^s - 1, H^s)} \lesssim c_j \]  

(1.28)

From Corollary 5.2 we control the modified energy of \( \tilde{A}^{(j)} \) as \( \mathcal{I}\mathcal{E}[\tilde{F}^{(j)}](t) \lesssim s^2 c_j \) for all \( t \in [t_0, t_1] \). From Propositions 6.3 and 6.4 we control higher derivatives of \( \tilde{F}^{(j)} \) and \( \tilde{w}^{(j)} \), which restricted to the ball \( B_j \) provide (1.10), (1.11) for \( \tilde{F}^{(j)} \) and \( \tilde{w}^{(j)} \) as well. These bounds together with (1.28) and (1.26) allows us to apply Proposition 8.1 to conclude, for all \( s \in [0, s_0] \), that

\[ (Ns^{\frac{1}{2}})^{\sigma(1-\sigma)}\|\chi_j[F^{(j)}(t, s) - \tilde{F}^{(j)}(t, s)]\|_{L_t^\infty L_x^2} \lesssim s^2 c_j \]  

(1.29)

\[ (Ns^{\frac{1}{2}})^{\sigma(1-\sigma)}\|\chi_j[w^{(j)}(t, s) - \tilde{w}^{(j)}(t, s)]\|_{L_t^\infty L_x^2} \lesssim s^2 c_j \]  

(1.30)

By (1.21) and (1.24)

\[ \mathcal{E}(t_1, s) - \mathcal{E}(t_0, s) = \sum_{j \in J} \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \chi_j(x)(w^{(j)}(s), F^{(j)}(s)) \, dx \, dt \]  

(1.31)

Using (1.29), (1.30) and (1.10), (1.11) with \( m = 0 \) for \( \tilde{F}^{(j)} \) and \( \tilde{w}^{(j)} \) to bound

\[ \chi_j(w^{(j)}(t, s) - \tilde{F}^{(j)}(t, s)) \quad \text{and} \quad \chi_j(w^{(j)}(t, s) - \tilde{w}^{(j)}(t, s)) \]  

in \( L_t^1 L_x^1 \), together with (1.12) allows us to reduce (1.7) to summing

\[ (N^2 s)^{1-\sigma} \left| \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \chi_j(x)(\tilde{w}^{(j)}(t, s), F^{(j)}(s)) \, dx \, dt \right| \lesssim s^{1\frac{-\sigma}{N^2 s}} c_j \]  

(1.32)

We now use Proposition 1.1 Part (2) to place \( \tilde{A}^{(j)}(t_0, \cdot, 0) \) in the Coulomb gauge (only at \( t = t_0 \)), as this results in the improved regularity from (1.8) for the curl-free part. We make this transformation on \([t_0, t_1] \times \mathbb{R}^3 \times [0, s_0] \) using a spatial \( U^{(j)}(x) \), so we maintain a regular caloric solution which at \( s = 0 \) is temporal, but now at \( s = 0 \) satisfying (1.7), (1.8) with \( \varepsilon = c_j \). Finally we make another change of gauge (5.29) to deTurck’s gauge using Proposition 5.10 and Remark 5.11 which leaves the solution invariant at \( s = 0 \). Due to gauge invariance (1.32) now follows from Proposition 10.1.

1.5.1. Proof of the Sobolev bound (1.8).

We apply Proposition 5.12 to obtain \( \|A^{(j)}(t_0, \cdot, s_0)\|_{H^s} \lesssim C_0(M_0) \), as we have (5.30) from Proposition 6.3. Without loss of generality we prove (1.8) for \( t = t_1 \).

Applying Lemma 2.3 for (1.22) at \( s = s_0, \; t = t_0 \), using (1.23) we obtain

\[ \|\nabla U^{(j)}\|_{H^s(B_j)} \lesssim d_j(1 + C_0(M_0)); \quad U^{(j)}, U^{-1} \in C_1(M_0)X_\sigma(B_j) \]  

(1.33)

where we denote

\[ d_j = c_j + \|A^{(j)}(t_0, \cdot, s_0)\|_{H^s(B_j)}, \quad \sum_j d_j^2 \lesssim C_0(M_0)^2 + 1 \]
The last bound follows from Proposition 3.2. From (1.27) we have
\[ \|A^{(j)}(t_1)\|_{H^s(B_j^{(1)})} + \|\partial_tA^{(j)}(t_1)\|_{H^{s-1}(B_j^{(1)})} \lesssim c_j \]
Combining this with (1.33), from (2.26) and (1.22) at \( s = 0, \ t = t_1 \) we obtain
\[ \|A(t_1)\|_{H^s(B_j^{(1)})} + \|\partial_tA(t_1)\|_{H^{s-1}(B_j^{(1)})} \lesssim M_0 c_j + d_j. \]
Applying Proposition 3.2 again we conclude \( \|A[t_1]\|_{H^s \times H^{s-1}} \leq C(M_0) < \infty. \)

1.6. Heuristic remarks and other works.

Remark 1.12 (Smoothing heuristics and connection to Littlewood-Paley theory). The heat flows solutions to (YMHP), (YMHP) satisfy the covariant parabolic equations (6.1). Therefore, in the first approximation one can think of the nonlinear \( F_{\alpha\beta}(s) \) as roughly \( F_{\alpha\beta}(s) \approx e^{s\Delta} F_{\alpha\beta}(0) \). Expanding (YMHP) it is seen that in general, the evolution of \( A_i(s) \) is only degenerately parabolic. With an appropriate choice of gauge (De Turck \( A_s = \partial^f A_i \)) one can make the \( A_i \) equations genuinely parabolic, specifically (5.1). Then the approximation \( A_i(s) \approx e^{s\Delta} A_i(0) \) is still expected.

The Gaussian multipliers \( e^{s\Delta} \) make the bulk of the frequencies to become concentrated to \( |\xi| \lesssim s^{-\frac{1}{4}} \). In connection to Littlewood-Paley projections this can be stated as \( e^{s\Delta} \approx P_{\leq k(s)} \) and \( \partial_x e^{s\Delta} \approx 2^{k(s)} P_{k(s)} \) where \( 2^{k(s)} = s^{-\frac{1}{4}} \). Since \( F_{\alpha\beta} \) contains derivatives, in a certain gauge one heuristically has \( F(s) \approx P_{k(s)} F(0) \).

Now to decompose \( A_i(0) \) one can set the caloric condition \( A_s = 0 \), turning (YMHP) into \( \partial_s A_i = D^f F_{i\ell}, \) which integrates to
\[ A_i(s) = A_i(s_0) + \int_s^{s_0} s D^f F_{i\ell} \frac{ds}{s} \]
Since \( ds \approx -ds/s, \ 2^{2k} s = 1 \) and taking \( s_0 = N^{-2} \) one can interpret (1.34) as a continuous Littlewood-Paley decomposition
\[ A(0) \approx P_{\leq N} A(0) + \int_{\log N}^{\infty} 2^{-k} P_k F(0) \, dk, \quad P_k A_i(0) \approx s D^f F_{i\ell}(s) \]

Remark 1.13 (Comparison of IE(t) with the classical I-method modified energy). The I-method suggests applying the Fourier multiplier \( I \) inside the energy and trying to prove almost conservation. This is explained in detail in [22] for the closely related MKG and in [44] for the simpler cubic wave equation.

The \( I \) operator is defined in (2.10). Roughly speaking
\[ \mathbf{E}[F[IA]](t) \approx \mathbf{E}[IF](t) \approx \|P_{\leq NF}\|_{L^2}^2 + N^{2(1-\sigma)} \sum_{2^k \geq N} \|D|^{\sigma-1} P_k F\|_{L^2}^2 \]
Recall the definition of \( I\mathcal{E}(t) \) and \( \mathcal{E}(t,s) \) in (1.6), (1.5). In light of Remark 1.12
\[ \mathcal{E}(t,s_0) \approx \|P_{\leq NF}\|_{L^2}^2, \quad \int_0^{s_0} (N^2 s)^{1-\sigma} \mathcal{E}(t,s) \frac{ds}{s} \approx \sum_{2^k \geq N} \left( \frac{N}{2^k} \right)^{2(1-\sigma)} \|P_k F\|_{L^2}^2 \]
where we denote \( F = F_{\alpha\beta}(0) \). Since \( s^{\frac{3}{4}} \approx 2^{-k} \approx |D|^{-1} \) the similarity between (1.36) and (1.37) is clear, suggesting \( I\mathcal{E}(t) \) could serve as a substitute to \( \mathbf{E}[F[IA]](t) \).
Remark 1.14 (Comparison of almost conservation laws). We have claimed above in III that $\mathcal{IE}(t)$ enjoys better almost conservation properties than $\mathbb{E}[F[I\Phi]](t)$ or $H[I\Phi](t)$ - the latter being the Maxwell-Klein-Gordon (MKG) classical $I$-method modified energy from [22] ($\Phi = (A_x, \phi)$), while the former would be the YM $I$-method modified energy in the Coulomb gauge (if it could be implemented for large data). We consider $\mathbb{E}[F[I\Phi]](t)$ and $H[I\Phi](t)$ as essentially the same, since we can ignore the $[A_i, A_j]$ commutators because they have better regularity.

The dominant part of $H[I\Phi](t_1) - H[I\Phi](t_0)$ is proved in [22] (106),(124) to be

$$
\sum_{k,k_i} \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \left( \partial_t I\Phi_k, I\mathbb{N}(\Phi_{k_1}, \Phi_{k_2}) - \mathbb{N}(I\Phi_{k_1}, I\Phi_{k_2}) \right) \, dx \, dt
$$

(1.38)

where $\mathbb{N}$ is a null form defined in (2.8), (2.9). Consider the more difficult case $N < 2^{k_1} < 2^{k_2} \simeq 2^k$, in which the commutator in (1.38) does not provide any cancelation and can be effectively written as

$$
\frac{1}{2^{k_{\min}+k_{\max}}} \left( \frac{N}{2^k} \right)^{2(1-\sigma)} \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \left( \partial_t \Phi_k, \mathbb{N}(\nabla_{x,t} \Phi_{k_1}, \nabla_{x,t} \Phi_{k_2}) \right) \, dx \, dt
$$

(1.39)

On the other hand, $\mathcal{E}(t_1, s) - \mathcal{E}(t_0, s)$ can be written as (1.21), of which (10.5) is the leading part. The similarity to (1.38) begins by noting that $w$ is also an essentially bilinear term with some cancelations, representing a "heat-wave commutator".

In (10.14), (10.15) it is revealed that the main contribution to $\mathcal{IE}(t_1) - \mathcal{IE}(t_0)$ comes roughly from terms

$$
(N^2 s)^{1-\sigma} \sum_{k,k_i} \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \left( \partial_t A_{k_i}^{df}, \mathbb{N}(\partial_t A_{k_1}^{df}, \partial_t A_{k_2}^{df}) \right) \, dx \, dt \times \text{weight}
$$

(1.40)

The weight in (1.40) is from (10.15) and in the same case $N < 2^{k_1} < 2^{k_2} \simeq 2^k$ it becomes concentrated at $2^k \simeq s^{-\frac{1}{2}}$ turning (1.40) into (with the notation $\Phi = A^{df}$)

$$
\frac{1}{2^{k_{\max}}} \left( \frac{N}{2^k} \right)^{2(1-\sigma)} \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \left( \partial_t \Phi_k, \mathbb{N}(\nabla_{x,t} \Phi_{k_1}, \nabla_{x,t} \Phi_{k_2}) \right) \, dx \, dt
$$

(1.41)

Comparing (1.41) with (1.38)-(1.39) shows that the structure of $\mathcal{IE}(t_1) - \mathcal{IE}(t_0)$ effectively shifts a derivative from a high frequency to low frequency, compared to $H[I\Phi](t_1) - H[I\Phi](t_0)$. This translates into a bound of $O(N^{-\frac{1}{2}})$ in (1.7) compared to $O(N^{-\frac{1}{2}} - \sigma)$ in [22] (45), leading to a result for lower regularities $\sigma > \frac{5}{6}$ rather than $\sigma > \frac{\sqrt{3}}{2}$.

1.6.1. Other works. Local well-posedness in the Lorenz gauge is given in [48]. [55]. Global existence for critical power Yang-Mills-Higgs equations at the energy regularity is proved in [21]. Global existence of solutions on globally hyperbolic Lorentzian manifolds, under a higher regularity assumption, is considered in [4] following the classical approach of [16].

We now mention notable results in $4+1$-dimensions, which is the energy-critical case, beginning with [31]. The global result for small energy was established in [33] building upon the high-dimensional case in [32]. The large data theory culminated with the proof of the Threshold Conjecture in [42].

2. Preliminaries

2.0.1. Space-time. We will work on the Minkowski space $\mathbb{R}^{1+3}$ equipped with the Minkowski metric, according to which indices will be raised and lowered. We adopt the Einstein summation convention. Greek letters run over the space-time variables, while latin indices are used only for spatial variables.

We use $A[t]$ as a short-hand for the pair $(A(t), \partial_t A(t))$. 


2.0.2. **Lie theory.** Let $G$ be a Lie group and $\mathfrak{g}$ its associated Lie algebra. We assume the existence of a bi-invariant inner product $(\cdot, \cdot) : \mathfrak{g} \times \mathfrak{g} \to [0, \infty)$. This means
\[(A, [B, C]) = (A, [B, C]) \quad \text{or equivalently,} \quad (UAU^{-1}, UBU^{-1}) = (A, B) \tag{2.1}\]
for all $A, B, C \in \mathfrak{g}$, $U \in G$. The Leibniz rule
\[\partial_{a}(B, C) = (\partial_{a}B, C) + (B, \partial_{a}C) \tag{2.2}\]
holds, as well as Bianchi’s identity
\[\partial_{a}F_{\beta\gamma} + \partial_{\beta}F_{\gamma\alpha} + \partial_{\gamma}F_{\alpha\beta} = 0. \tag{2.3}\]
Covariant derivatives are commuted using
\[\partial_{a}B_{\beta} - \partial_{\beta}B_{a} = [B_{a}, \cdot] \tag{2.4}\]
The covariant laplacian is denoted by $\Delta_{A} := D^{\alpha}D_{\alpha}$.
Due to (2.1), the inner-product can be used to define the $L^{p}$ norms of $\mathfrak{g}$-valued functions in a gauge-invariant way.

For concreteness, it is common to take $G$ to be a matrix group such as $SO(n, \mathbb{R})$ or $SU(n, \mathbb{C})$, with the associated Lie Algebras $\mathfrak{so}(n, \mathbb{R})$ and $\mathfrak{su}(n)$. In these cases the Lie bracket is $[A, B] = AB - BA$ given by matrix multiplication, while the bi-invariant scalar products are $(A, B) = \text{tr}(AB^{*})$, respectively $(A, B) = \text{tr}(AB^{*})$.

2.0.3. **Regular functions.** A function $f$ on $\mathbb{R}^{3}$ (such as an initial data $\hat{A}_{i}$) is called regular if $f \in H^{\infty}(\mathbb{R}^{3}) := \bigcap_{n \geq 0} H^{n}(\mathbb{R}^{3})$. Solutions $A_{\alpha}(t, x)$ or $A_{\lambda}(x, s)$ defined on intervals are called regular if $A_{\alpha} \in C_{t}^{\infty}H_{x}^{\infty}(I \times \mathbb{R}^{3})$, respectively $A_{\lambda} \in C_{x}^{\infty}H_{t}^{\infty}(\mathbb{R}^{3} \times J)$. Similarly for $A_{\alpha} \in C_{t, x}^{\infty}H_{x}^{\infty}(I \times \mathbb{R}^{3} \times J)$. Gauge transformations are called regular if $\nabla U, \nabla U^{-1} \in H^{\infty}(\mathbb{R}^{3})$.

2.0.4. **Boundedness notations.** We use $A \lesssim B$ to denote $A \leq CB$ and $A \ll B$ to denote $AC \leq B$, where $C$ is a large constant. We use $A+ \text{ or } A-$ to denote $A + \varepsilon$ and $A - \varepsilon$ where $0 < \varepsilon \ll 1$ is a large number which can depend on $\sigma > 0$. We denote $(x)^{2} = 1 + |x|^{2}$.

For injective operators $T$ on $X$ we denote by $TX$ the image space with norm $\|T^{-1}u\|_{X}$. The embedding $X \subseteq Y$ as an estimate denotes $\|u\|_{Y} \lesssim \|u\|_{X}$. The mapping $T : X \to Y$ as an estimate means $\|T u\|_{Y} \lesssim \|u\|_{X}$. We say $u \in \delta X$ if $\|u\|_{X} \lesssim \delta$. Product estimates $X \times Y \to Z$ denote $\|uv\|_{Z} \lesssim \|u\|_{X}\|v\|_{Y}$. The space of sums $X + Y$ is endowed with the norm $\|u\|_{X+Y} := \inf\{\|u_{1}\|_{X} + \|u_{2}\|_{Y} \mid u = u_{1} + u_{2}\}$.

2.0.5. **Fourier operators.** The Fourier transform of $f$ is denoted by $\hat{f}$ or $\mathcal{F} f$. Littlewood-Paley projections are denoted by $P_{k}$. Whether the frequencies $|\xi| \lesssim 1$ are included in $P_{0}$ or not will be clear from the context. Moreover, $\hat{P}_{k}$ denotes a similar multiplier such that $\hat{P}_{k}P_{k} = P_{k}$.

Linear and bilinear operators which are invariant to spatial translations are understood either through their spatial kernels, or as multipliers through their Fourier symbols. For instance: the smoothing $I$ operator is defined by $\mathcal{F} I \mathcal{F} \hat{f}$, fractional derivative operators $|D|^{p}$, $\langle D \rangle^{p}$ are defined through their symbols $|\xi|^{p}$, $\langle \xi \rangle^{p}$.

When the fields $A, B$ are $\mathfrak{g}$-valued a bilinear operator with symbol $m^{ij}(\xi_{1}, \xi_{2})$ takes the form
\[\mathcal{M}(A, B)(x) = \int_{\mathbb{R}^{3} \times \mathbb{R}^{3}} e^{ix(\xi_{1} + \xi_{2})}m^{ij}(\xi_{1}, \xi_{2})[\hat{A}_{j}(\xi_{1}), \hat{B}_{i}(\xi_{2})] \, d\xi_{1} \, d\xi_{2} \tag{2.5}\]
\[= \int_{\mathbb{R}^{3} \times \mathbb{R}^{3}} K^{ij}(x - y_{1}, x - y_{2})[A_{i}(y_{1}), B_{j}(y_{2})] \, dy_{1} \, dy_{2}\]
where $m(\xi_{1}, \xi_{2}) = \hat{K}(\xi_{1}, \xi_{2})$. 

12
A linear or bilinear multiplier operator is called disposable when its kernel is a function or measure with bounded mass. Minkowski’s inequality insures that disposable operators are bounded on translation-invariant normed spaces.

2.0.6. **Leray projections and null forms.** Denote by \( P \) and \( P^\perp \) the projections on divergence free, respectively curl-free vector fields:

\[
P_j A := \Delta^{-1} \partial^i (\partial_i A_j - \partial_j A_i), \quad P^\perp_j A := (1 - P_j) A = \Delta^{-1} \partial_j \text{div} A.
\] (2.6)

Recall the classical identities ([26], [27])

\[
P_j (\phi \nabla_x \varphi) = \Delta^{-1} \nabla^i Q_{ij} (\phi, \varphi)
\]

\[
P A^i \partial_i \phi = Q_{ij} \left( \Delta^{-1} \nabla^i A^j, \phi \right)
\] (2.7)

which follow from the definitions by simple computations.

By \( N(f,g) \) we denote a null form, meaning a linear combination of \( \Delta^{-1} \nabla^i Q_{ij}(f,g) \) or \( Q_{ij}(\Delta^{-1} \nabla^i f, g) \) or

\[
Q_{ij}(f,g) := \partial_i f \partial_j g - \partial_j f \partial_i g
\] (2.8)

where

\[
Q_{ij}(f,g) := \partial_i f \partial_j g - \partial_j f \partial_i g
\] (2.9)

In the context of \( g \)-valued fields (2.9) is replaced by (4.3).

2.0.7. **The \( I \) multiplier.** Fix a number \( N \gg 1 \) and let \( m(\xi) \) and \( I \) be a smooth positive radial symbol and the associated Fourier multiplier operator such that

\[
m(\xi) = \begin{cases} 1, & |\xi| \leq N \\ \left( \frac{N}{|\xi|} \right)^{-1-\sigma}, & |\xi| \geq 2N \end{cases}; \quad I \tilde{f}(\xi) := m(\xi) \tilde{f}(\xi)
\] (2.10)

The \( I \) operator smoothens the high frequencies and acts as the identity on relatively lower frequencies. It has an integrable kernel. It is straightforward to verify

**Lemma 2.1.** One has \( I f \in H^1 \iff f \in H^1 + N^{\sigma-1}H^\sigma \) and \( I g \in L^2 \iff g \in L^2 + N^{\sigma-1}H^{\sigma-1} \)

with equivalent norms.

The advantage of these second formulations is that they can be directly localized to a domain.

2.0.8. **Covariant Gagliardo-Nirenberg inequalities and elliptic equations.** We recall the following *diamagnetic*, or *Kato’s*, inequality. For a proof see, for example [38, Lemma 4.2] or [54, Lemma 3.9]. One has

\[
|\partial_x |\varphi|| \leq |D_x \varphi|
\]

for regular \( g \)-valued functions \( \varphi \) on \( \mathbb{R}^3 \). As a consequence one has the following covariant Gagliardo-Nirenberg and Sobolev inequalities on \( \mathbb{R}^3 \)

\[
\|\varphi\|_{L^p} \lesssim \|\varphi\|_{L^2}^{\frac{1}{2}} \|D_x \varphi\|_{L^2}^{\frac{1}{2}}, \quad \alpha = 3(2^{-1} - p^{-1})
\]

\[
\|\varphi\|_{L^3} \lesssim \|\varphi\|_{L^2}^{\frac{1}{2}} \|D_x \varphi\|_{L^2}^{\frac{1}{2}}
\]

\[
\|\varphi\|_{L^6} \lesssim \|D_x \varphi\|_{L^2}
\]

\[
\|\varphi\|_{L^\infty} \lesssim \|D_x \varphi\|_{L^2} \|D^2_x \varphi\|_{L^2} \|D^3_x \varphi\|_{L^2}
\] (2.11)

If \( \varphi \) is defined on a unit ball \( B \) instead, one has

\[
\|\varphi\|_{L^p(B)} \lesssim \|\varphi\|_{L^2(B)}^{1-\alpha} \|D_x \varphi\|_{L^2(B)} + \|\varphi\|_{L^2(B)}, \quad \frac{1}{p} = \frac{\alpha}{6} + \frac{1-\alpha}{2}.
\] (2.12)
\[ \| \varphi \|_{L^\infty(B)} \lesssim (\| D_x \varphi \|_{L^2(B)} + \| \varphi \|_{L^2(B)})^{\frac{1}{2}} (\| D_x^2 \varphi \|_{L^2(B)} + \| D_x \varphi \|_{L^2(B)})^{\frac{1}{2}} + \| \varphi \|_{L^2(B)} \]  

(2.13)

By the argument in [39, Theorem 4.1] one has

**Proposition 2.2.** Assume \( A \in \dot{H}^2(\mathbb{R}^3) \). Then the covariant elliptic equation \( \Delta_A B = F \) is solved by the \( \mathfrak{g} \)-valued function \( B \) with the bound

\[ \| B \|_{\dot{H}^\rho} \lesssim \| F \|_{\dot{H}^{\rho - 2}}, \quad \rho \in (0, 2). \]

### 2.1. Sobolev spaces.

The global Sobolev spaces \( H^s = H^s(\mathbb{R}^n) \) and \( \dot{H}^s = \dot{H}^s(\mathbb{R}^n) \) are defined using the Fourier transform, for any \( s \in \mathbb{R} \), by

\[ \| u \|_{H^s} := \| \langle \xi \rangle^s \hat{u} \|_{L^2(\mathbb{R}^n)}, \quad \| u \|_{\dot{H}^s} := \| \| \xi \|_s \hat{u} \|_{L^2(\mathbb{R}^n)} \]  

(2.14)

We denote \( H^\infty(\mathbb{R}^n) := \cap_{m \geq 0} H^m(\mathbb{R}^n) \). For \( s > 0 \) one has \( H^s = L^2 \cap \dot{H}^s \) while for \( s < 0 \) one has \( H^s = L^2 + \dot{H}^s \). By duality one may identify \( \dot{H}^{-s} \) with \((H^s)^*\).

We next define Sobolev spaces \( W^{s,2}(D) \) (also denoted \( H^s(D) \)) when \( D \) is either a ball \( B \subset \mathbb{R}^n \) of radius \( \simeq 1 \) or \( D = \mathbb{R}^n \). When \( s \in \{0, 1\} \) we use the classical definition. When \( s \in (0, 1) \) we use the norm

\[ \| u \|_{W^{s,2}(D)}^2 := \| u \|_{L^2(D)}^2 + \| |\cdot|_{W^{s,2}(D)}^2 \]  

(2.15)

where \( |\cdot|_{W^{s,2}(D)} \) denotes the Gagliardo seminorm defined by

\[ |u|_{W^{s,2}(D)}^2 := \int_D \int_D \frac{|u(x) - u(y)|^2}{|x - y|^{n+2s}} \, dx \, dy \]  

(2.16)

Due to (2.17) we will use the notations \( H^s(D) \) and \( W^{s,2}(D) \) (resp. \( \dot{H}^s(D) \) and \( \dot{W}^{s,2}(D) \)) interchangeably for both \( D = B \) and \( D = \mathbb{R}^n \).

We denote \( W^{0,2}_0(B) \) the closure of \( C^\infty_0(B) \) in the norm \( |\cdot|_{W^{s,2}(B)} \). Of course \( W^{0,2}_0(\mathbb{R}^n) = W^{s,2}(\mathbb{R}^n) \).

When \( s \in (-1, 0) \) we define \( W^{s,2}(D) \), as a subspace of distributions, by duality

\[ W^{s,2}(D) := W^{-s,2}_0(D)^*. \]

We recall some properties from [13] For \( D = \mathbb{R}^n \) the two definitions (2.14), (2.15) coincide

\[ \| u \|_{H^s} \simeq \| u \|_{W^{s,2}(\mathbb{R}^n)}, \quad \| u \|_{\dot{H}^s} \simeq \| u \|_{W^{s,2}(\mathbb{R}^n)} \]  

(2.17)

If \( B \) is a ball, then any \( u \in W^{s,2}(\mathbb{R}^n) \) has an extension \( \tilde{u} \in W^{s,2}(\mathbb{R}^n) \) with comparable norm

\[ \| u \|_{W^{s,2}(B)} \simeq \| \tilde{u} \|_{W^{s,2}(\mathbb{R}^n)} \simeq \inf\{ \| \tilde{u} \|_{W^{s,2}(\mathbb{R}^n)} \mid \tilde{u} \mid_B = u \}. \]

If \( \psi \in C^{0,1}(B) \), \( 0 \leq \psi \leq 1 \) then

\[ \| \psi \|_{W^{s,2}(B)} \lesssim \| u \|_{W^{s,2}(B)}, \quad s \in (0, 1). \]  

(2.18)
2.1.1. **Sobolev multiplication laws.**

We recall from [11], [52] that: if \( s_1 + s_2 > s + \frac{4}{5}, s_1 + s_2 \geq 0 \) and \( s \leq \min(s_1, s_2) \) then one has

\[
\|fg\|_{H^s} \lesssim \|f\|_{H^{s_1}} \|g\|_{H^{s_2}}. \tag{2.19}
\]

In a similar spirit, one has the homogeneous version [37, Lemma 3.2]

\[
\|fg\|_{\bar{H}^s} \lesssim \|f\|_{\bar{H}^{s_1}} \|g\|_{\bar{H}^{s_2}}. \tag{2.20}
\]

provided \( s_1 + s_2 = s + \frac{4}{5} \) and \(-s, s_1, s_2 < \frac{4}{5}\).

For \( r > 3/4 \) we will use the version [22, Lemma 2.1]:

\[
\|fg\|_{\bar{H}^{r-1}} \lesssim \|f\|_{\bar{H}^{r}} \|g\|_{L^2}. \tag{2.21}
\]

One also has

\[
\|I(fg)\|_{\bar{H}^{-1}} \lesssim \|If\|_{\bar{H}^1} \|Ig\|_{L^2} \tag{2.22}
\]

which follows immediately from (2.21) with \( r = \sigma \).

2.1.2. **Norms for gauge transformations.** We now define some further norms. Let \( X_\sigma = \bar{H}^1 \cap \bar{H}^{1+\sigma} \cap L^\infty(\mathbb{R}^3) \), and the similar definition on a ball, i.e.

\[
\|U\|_{X_\sigma} := \|\nabla U\|_{H^\sigma(\mathbb{R}^3)} + \|U\|_{L^\infty(\mathbb{R}^3)}, \quad \|U\|_{X_\sigma(B)} := \|\nabla U\|_{H^\sigma(B)} + \|U\|_{L^\infty(B)}. \tag{2.23}
\]

Similarly one defines

\[
\|V\|_X := \|I\nabla V\|_{H^1(\mathbb{R}^3)} + \|V\|_{L^\infty(\mathbb{R}^3)}
\]

One has the following algebra and product estimates

\[
X_\sigma \times X_\sigma \to X_\sigma, \quad H^\sigma \times X_\sigma \to H^\sigma, \quad H^{-1+\sigma} \times X_\sigma \to H^{-1+\sigma} \tag{2.24}
\]

\[
X \times X \to X, \quad \mathcal{I}^{-1} H^1 \times X \to \mathcal{I}^{-1} H^1, \quad \mathcal{I}^{-1} L^2 \times X \to \mathcal{I}^{-1} L^2 \tag{2.25}
\]

These are proved using the Littlewood-Paley trichotomy, (2.24) is [50, Eq. (10)]. By an extension argument, (2.24) holds on domains as well

\[
\|UV\|_{X_\sigma(B)} \lesssim \|U\|_{X_\sigma(B)} \|V\|_{X_\sigma(B)}, \quad \|fU\|_{H^\sigma(B)} \lesssim \|f\|_{H^\sigma(B)} \|U\|_{X_\sigma(B)} \tag{2.26}
\]

for \( r \in \{\sigma, \sigma - 1\} \).

One may obtain \( X_\sigma(B) \) control for the gauge transformation between two \( H^\sigma(B) \) fields. This is [56, Lemma 1.2] for integer regularities and it easily generalizes.

**Lemma 2.3.** Let \( U : B \to G \) be such that \( A_i = U \tilde{A}_i U^{-1} - \partial_i U U^{-1} \) on \( B \). If \( M = \|A_i\|_{H^\sigma(B)} + \|\tilde{A}_i\|_{H^\sigma(B)} \), then \( \|U\|_{X_\sigma(B)} - 1 \lesssim \|\nabla U\|_{H^\sigma(B)} \lesssim M(1 + M) \).

**Proof.** The \( L^\infty(B) \) bound is automatic, so we prove \( \|\partial_i U\|_{H^\sigma(B)} \lesssim M(1 + M) \). Write \( \partial_i U = U \tilde{A}_i - A_i U \), which implies \( \|U\|_{W^{1,p}(B)} \lesssim 1 + M \) where \( p \) is the Sobolev exponent for \( H^\sigma \). We use

\[
\|Vg\|_{H^\sigma(\mathbb{R}^3)} \lesssim \|V\|_{W^{1,p}(\mathbb{R}^3)} \|g\|_{H^\sigma(\mathbb{R}^3)} \Rightarrow \|Uf\|_{H^\sigma(B)} \lesssim \|U\|_{W^{1,p}(B)} \|f\|_{H^\sigma(B)}
\]

The first inequality follows from Littlewood-Paley theory, while the second follows from the first by an extension argument. Apply this inequality for \( U \tilde{A}_i \) and \( A_i U \) to obtain \( \|\partial_i U\|_{H^\sigma(B)} \lesssim M(1 + M) \). \( \square \)
2.2. $X^{s,b}$ spaces. We outline the main properties, for more details we refer to [47], [30]. The $X^{s,b}$ spaces associated to the wave equation are defined by the norm

$$
\| u \|_{X^{s,b}} = \| \langle \xi \rangle^s (|\tau| - |\xi|)^b F u(\tau, \xi) \|_{L_{\tau,x}^2}
$$

where $F u$ denotes the space-time Fourier transform on $\mathbb{R} \times \mathbb{R}^3$.

Following [50] we also use the $X^{r,\theta}_{\tau=0} = H^r_\tau H^\theta_t$ spaces adapted to solutions of the equation $\partial_t v = F$, with norm

$$
\| v \|_{X^{r,\theta}_{\tau=0}} = \| \langle \xi \rangle^r (\tau)^\theta F v(\tau, \xi) \|_{L_{\tau,x}^2}
$$

For a bounded interval $J$ and $X = X^{s,b}$ or $X^{r,\theta}_{\tau=0}$ one defines the restricted space $X_J$ consisting of functions or tensors $u$ defined on $J \times \mathbb{R}^3$ for which the following norm is finite:

$$
\| u \|_{X_J} = \inf \{ \| v \|_X ; v|_{J \times \mathbb{R}^3} = u \}
$$

To say $u \in X$ on $J \times \mathbb{R}^3$ means $u \in X_J$, i.e. it has an extension $v \in X$ to $\mathbb{R} \times \mathbb{R}^3$.

One has the well-known embeddings

$$
X^{s,\frac{1}{r}+} \subseteq L^\infty_t H^s_x, \quad X^{r,\frac{1}{q}+}_{\tau=0} \subseteq L^\infty_t H^r_x.
$$

Moreover, one has the embedding

$$
X^{0,0+} \subseteq L^2_t + L^\infty_t
$$

which follows by interpolation between $X^{0,0} = L^2_t L^2_x$ and $X^{0,\frac{1}{q}+} \subseteq L^\infty_t L^2_x$.

Multiplications by cutoff functions interact well with these spaces:

**Lemma 2.4.** Denote $\| u \|_{X^{1,\theta}} := \| u \|_{X^{1,0}} + \| \partial_t u \|_{X^{0,b}}$ and let $\chi$ be a bump function. Then, for $0 \leq b \leq 1$ one has $\| \chi u \|_{X^{1,\theta}} \lesssim \| u \|_{X^{1,\theta}}$.

**Proof.** It is easily seen that an equivalent norm for $X^{1,\theta}$ is

$$
\| \langle |\tau| + |\xi| \rangle (|\tau| - |\xi|)^b F u(\tau, \xi) \|_{L_{\tau,x}^2}
$$

When $b = 0$, respectively $b = 1$ this is equivalent to

$$
\| u \|_{L^2_t L^2_x} + \| \nabla_{t,x} u \|_{L^2_t L^2_x}, \quad \text{respectively} \quad \| u \|_{L^2_t L^2_x} + \| \nabla_{t,x} u \|_{L^2_t L^2_x} + \| \Box u \|_{L^2_t L^2_x}
$$

and then the property follows from the chain rule. The general case follows from these by complex interpolation, see [18] Lemma 1.4. \hfill \Box

2.2.1. $X^{s,b}$ and $X^{r,\theta}_{\tau=0}$ energy estimates. Suppose $t_0 \in J$ with $|J| \lesssim 1$ and $b, \theta \in (\frac{1}{2}, 1)$, $s, r \in \mathbb{R}$. If $\Box u = F$ then

$$
\| u \|_{X^{s,b}_{\tau=0}} + \| \partial_t u \|_{X^{s-1,b}_{\tau=0}} \lesssim \| u(t_0) \|_{H^s \times H^{s-1}} + \| F \|_{X^{s-1,b-1}_{\tau=0}}
$$

If $\partial_t = G$ then

$$
\| v \|_{X^{r,\theta}_{\tau=0}} \lesssim \| v(t_0) \|_{H^r} + \| G \|_{X^{r,\theta-1}_{\tau=0}}
$$

2.2.2. Strichartz estimates. We recall

$$
X^{s,\frac{1}{r}+} \subseteq L^2_t L^q_x
$$

when $q \in (2, \infty)$, $s \geq 0$ and $\frac{1}{q} + \frac{1}{r} \leq \frac{1}{2}$; $\frac{1}{q} + \frac{3}{r} \geq \frac{3}{2} - s$. Moreover, we will use the following Bilinear Strichartz estimate (see e.g. [31]):

$$
D^{-\left(\frac{1}{2} - \right)} \left( X^{s,\frac{1}{r}+} \times X^{s,\frac{1}{q}+} \right) \subseteq L^2_t L^2_x
$$
3. Local well-posedness for Yang-Mills

3.1. Approximation by regular initial data sets.

We begin with a $H^\sigma \times H^{\sigma - 1}$ approximation result, recalling that initial data sets for Yang-Mills have to satisfy the Gauss equation (1.3) in order to be admissible. The similar statement for $H^1 \times L^2$ appears in [27, Proposition 1.2].

**Proposition 3.1.** Any initial data $(A_i, E_i) \in H^\sigma \times H^{\sigma - 1}$ satisfying the constraint equation (1.3) can be approximated in $H^\sigma \times H^{\sigma - 1}$ by a sequence of regular initial data $(A^n_i, E^n_i) \in H^\infty \times H^\infty$ satisfying the constraint equation (1.3). Moreover, $A^n_i$ can be chosen compactly supported.

Proof. Begin with a sequence of smooth compactly supported sequence $(A^n, \tilde{E}^n)$ converging to $(A, E)$ in $H^\sigma \times H^{\sigma - 1}$. Denote the covariant derivatives by $D^n := \partial + [A^n, \cdot]$. Using (1.3) and writing, for any $f \in H^{1-\sigma}$

$$\langle D^{n,i} \tilde{E}^n_i, f \rangle = \langle E_i - \tilde{E}^n_i, \partial^n f \rangle + \langle [A^n - A, \tilde{E}^n], f \rangle + \langle [A, \tilde{E}^n - E], f \rangle$$

one obtains $D^{n,i} \tilde{E}^n_i \to 0$ in $H^{\sigma - 1} \hat{H}1 = H^{-1} + \hat{H}^{\sigma - 2}$ after using $H^\sigma \times H^{1-\sigma} \hat{H}1 \to H^{1-\sigma}$, which is a version of (2.19).

Define $E^n_i = \tilde{E}^n_i + D^n_i \phi^n$ which will satisfy (1.3) provided we choose

$$\Delta A^n \phi^n = -D^{n,i} \tilde{E}^n_i$$

It remains to show $D^n_i \phi^n \to 0$ in $H^{\sigma - 1} = L^2 + \hat{H}^{\sigma - 1}$. By Proposition 2.2 we obtain $\phi^n \to 0$ in $H^1 + \hat{H}^\sigma$ so it remains to use the following easy version of (2.19)

$$\|\|[A^n, \phi^n]\|_{H^{\sigma - 1}} \lesssim \|A^n\|_{H^\sigma} \|\phi^n\|_{\hat{H}1 + \hat{H}^\sigma} \to 0$$

\[\square\]

3.2. Square summability of fractional Sobolev spaces.

We will need the following square-summability property of $W^{s,2}$ norms.

**Proposition 3.2.** Let $(B_j)_j$ be a finitely overlapping uniform covering of $\mathbb{R}^n$ by balls of radius $\simeq 1$ and let $s \in (-1, 1)$. Then, for any $u \in W^{s,2}(\mathbb{R}^n)$ one has

$$\sum_j \|u\|^2_{W^{s,2}(B_j)} \simeq \|u\|^2_{W^{s,2}(\mathbb{R}^n)} \tag{3.1}$$

Proof. When $s = 0$ the property is obvious. Suppose first that $s \in (0, 1)$. Clearly, from (2.16)

$$\sum_j |u|^2_{W^{s,2}(B_j)} \lesssim \|u\|^2_{W^{s,2}(\mathbb{R}^n)}$$

Conversely, by splitting the $\mathbb{R}^n \times \mathbb{R}^n$ integral in (2.16) into regions where there exists $j$ such that both $x, y \in B_j$ and regions where $|x - y| > \delta$ (where we bound $|u(x) - u(y)| \leq |u(x)| + |u(y)|$ and use the $L^2(B_j)$ norms), we obtain

$$|u|^2_{W^{s,2}(\mathbb{R}^n)} \lesssim \sum_j \|u\|^2_{W^{s,2}(B_j)}$$

Now assume $s \in (-1, 0)$. Let $(\psi_j)_j$ be a partition of unity associated to $(B_j)_j$. Let $u$ be a tempered distribution and let $\phi$ be a Schwartz function. Using (2.18) and what we already proved

$$|\langle u, \phi \rangle| \lesssim \sum_j \|u\|_{W^{s,2}(B_j)} \|\psi_j \phi\|_{W_0^{\sigma,2}(B_j)} \lesssim \left( \sum_j \|u\|^2_{W^{s,2}(B_j)} \right)^{1/2} \|\phi\|_{W^{\sigma,2}(\mathbb{R}^n)}$$

17
This shows RHS $(3.1) \lesssim \text{LHS (3.1)}$. Conversely,

\[
\text{LHS (3.1)} \lesssim \sup_{\|\epsilon\|_{L^2} \leq 1} \sum_j c_j \| u \|_{W^{s,2}(B_j)} = \sup_{\|\epsilon\|_{L^2} \leq 1} \sum_j \sup_{\|\phi_j\|_{W^{-s,2}(B_j)} \leq 1} c_j \langle u, \phi_j \rangle \lesssim \sup_{\|\phi\|_{W^{-s,2}(\mathbb{R}^n)} \leq 1} \langle u, \phi \rangle = \| u \|_{W^{s,2}(\mathbb{R}^n)} = \text{RHS (3.1)} \]

\[\square\]

### 3.3. Initial data extension.

We recall the following result from [40], which we state only for a ball $B$ of radius $\approx 1$.

**Theorem 3.3** ([40] - Theorem 5.4.). Let $\rho \in (\frac{1}{2}, \frac{5}{2})$ and let $(\bar{A}_\rho, E_\rho) \in H^\rho(B) \times H^{\rho-1}(B)$ satisfy the Gauss equation \( (1.3) \). If $\| \bar{A}_\rho \|_{H^\rho(B)} \leq \epsilon$ for a sufficiently small $\epsilon > 0$, then there exist $(\bar{A}_\rho, E_\rho) \in H^\rho(\mathbb{R}^3) \times H^{\rho-1}(\mathbb{R}^3)$ which satisfies the Gauss equation \( (1.3) \), which coincides with $(\bar{A}_\rho, E_\rho)$ on $B$ such that

\[
\| (\bar{A}_\rho, E_\rho) \|_{H^\rho \times H^{\rho-1}} \lesssim \| (\bar{A}_\rho, E_\rho) \|_{H^\rho(B) \times H^{\rho-1}(B)}
\]

and such that the map $\bar{A}_\rho, E_\rho) \mapsto (\bar{A}_\rho, E_\rho)$ is locally Lipschitz continuous. If $(\bar{A}_\rho, E_\rho)$ is smooth then so is $(\bar{A}_\rho, E_\rho)$.

The proof is based from the following solvability result for the inhomogeneous Gauss equation.

**Proposition 3.4** ([40] - Prop. 4.2.). Suppose a satisfies $\| a \|_{H^\frac{1}{4}(\mathbb{R}^3)} \leq \epsilon_*$. Let $K$ be a convex domain. If $\epsilon_\ast > 0$ is small enough there exists a solution operator $T_{a}$ for the equation $D^\ell e_\ell = h$ with the following properties:

1. $\| T_{a} h \|_{W^{2,p}} \lesssim K_{r,p} \| h \|_{W^{2-1,p}}$, holds for $2 \leq p < \infty$ and $1 - \frac{3}{p} < r < \frac{3}{2}$.
2. If $h = 0$ in $\lambda K$, then $T_{a} h = 0$ in $\lambda K$.
3. If $a$ and $h$ are smooth, so is $T_{a} h$.

In the proof of the almost conservation law we need the following version of Theorem 3.3. The proof carries over almost verbatim from Theorem 5.4. from [40]. We present the argument for the sake of completeness.

**Proposition 3.5.** Suppose $(a, e)$ satisfy the constraint equation \( (1.3) \) on a ball $B \subset \mathbb{R}^3$ of radius $\approx 1$ and

\[
(a, e) \in (H^1 + N^{\sigma-1}H^\sigma)(B) \times (L^2 + N^{\sigma-1}H^{\sigma-1})(B)
\]

Suppose $\| a \|_{H^\frac{1}{4}(B)} \leq \epsilon$ for $\epsilon > 0$ small enough. Then there exist $\tilde{a}, \tilde{e}$ solving \( (1.3) \) on $\mathbb{R}^3$ which coincide with $(a, e)$ on $B$ with

\[
\| \tilde{a} \|_{H^1 + N^{\sigma-1}H^\sigma} + \| \tilde{e} \|_{L^2 + N^{\sigma-1}H^{\sigma-1}} \lesssim \| a \|_{H^1 + N^{\sigma-1}H^\sigma(B)} + \| e \|_{L^2 + N^{\sigma-1}H^{\sigma-1}(B)}.
\]

If $(a, e)$ are smooth, then $(\tilde{a}, \tilde{e})$ are regular.

**Proof.** Decompose $a = a^1 + N^{\sigma-1}a^\sigma, e = e^0 + N^{\sigma-1}e^\sigma$ where $a^1 \in H^1(B), a^\sigma \in H^\sigma(B), e^0 \in L^2(B), e^\sigma \in H^\sigma(B)$ and $e^\sigma \subset H^{\sigma-1}(B)$. One begins by applying a universal extension operator to $a^1, a^\sigma, e^0, e^\sigma$ obtaining $\tilde{a}^1, \tilde{a}^\sigma, \tilde{e}^0, \tilde{e}^\sigma$ on $\mathbb{R}^3$ with comparable norms. Let $\tilde{a} = a^1 + N^{\sigma-1}\tilde{a}^\sigma, \tilde{e} = \tilde{e}^0 + N^{\sigma-1}\tilde{e}^\sigma$.

Since the constraint equation \( (1.3) \) is violated in general outside $B$ by $(\tilde{a}, \tilde{e})$, one must correct $\tilde{e}$ by adding a term. Let $h^{-1} = (D^\delta)^{\epsilon_0} \in H^{-1}$ and $h^{-2} = (D^\delta)^{\epsilon_-1} \in H^{-2}$ be the errors of the Gauss equations, which are supported outside $B$. Now apply Proposition 3.4 to obtain $d^0 = -T_{\tilde{a}} h^{-1}$ and $d^{-1} = -T_{\tilde{a}} h^{-2}$ which solve $(D^\delta)^{d^0} = -h^{-1}$ and $(D^\delta)^{d^{-1}} = -h^{-2}$ and satisfy $\| d^0 \|_{L^2} \lesssim \| e^0 \|_{L^2}$, $\| d^{-1} \|_{H^{\sigma-1}} \lesssim \| e^{-1} \|_{H^{\sigma-1}}$.

Now let $d = d^0 + N^{\sigma-1}d^\sigma$ and define $\tilde{e} = \tilde{e} + d$. Then $(\tilde{a}, \tilde{e})$ satisfy the conditions. \[\square\]
3.4. Review of Theorem 1.3 and proof of Corollary 1.4

For the sake of completeness and to set up the notation we review the argument for Theorem 1.3 from [40] Theorem 1.27. The key technical tool is their initial data surgery result in Theorem 3.3 ([40], Theorem 5.4].

We begin with an initial data set \((\bar{A}_x, \bar{E}_x) \in H^\sigma(\mathbb{R}^3) \times H^{\sigma-1}(\mathbb{R}^3)\) satisfying (1.3). It suffices to consider \(\frac{3}{2} < \sigma < 1\) and to obtain a solution for positive times.

Rescale the initial data by \(\bar{A}^\lambda(x) = \lambda^{-1}\bar{A}(x/\lambda)\), \(\bar{E}^\lambda(x) = \lambda^{-2}\bar{E}(x/\lambda)\) so that

\[
\|((\bar{A}^\lambda, \bar{E}^\lambda))\|_{H^\sigma \times H^{\sigma-1}} \leq \lambda^{\frac{3}{2} - \sigma}\|((\bar{A}, \bar{E}))\|_{H^\sigma \times H^{\sigma-1}} \ll \varepsilon_*,
\]

where \(\varepsilon_* > 0\) is chosen small enough to be able to apply Theorem 1.2 and Theorem 3.3 to obtain modified initial data sets as follows.

Let \((B_j)_j\) be a finitely overlapping uniform covering of \(\mathbb{R}^3\) by balls of radius 2 centered at \((y_j)_j\) such that the domains of dependency \((D_j)_j\) with bases \((B_j)_j\) form a covering of \([0, 1] \times \mathbb{R}^3\). These are truncated cones \(D_j := \{(t, x) \mid t + |x - y_j| < 2, t \in [0, 1]\}\).

One may apply Theorem 3.3 for each \(B_j\) to obtain an extension \((\bar{A}^{(j)}, \bar{E}^{(j)})\) to \(\mathbb{R}^3\) satisfying the constraint equation (1.3) and which coincide with \((\bar{A}^\lambda, \bar{E}^\lambda)\) on \(B_j\) with the bounds

\[
\|((\bar{A}^{(j)}, \bar{E}^{(j)}))\|_{H^\sigma \times H^{\sigma-1}} \lesssim \|((\bar{A}^\lambda, \bar{E}^\lambda))\|_{W^{\sigma,2} \times W^{\sigma-1,2}(B_j)} \ll \varepsilon_*.
\]

The map \((\bar{A}^\lambda, \bar{E}^\lambda) \mapsto (\bar{A}^{(j)}, \bar{E}^{(j)})\) is locally Lipschitz on these spaces.

To each \((\bar{A}^{(j)}, \bar{E}^{(j)})\) one applies Theorem 1.2 (i.e. Theorem 1.1. in [50]) to obtain a Yang-Mills solution \(A^{(j)}\) in the temporal gauge on \([0, 1] \times \mathbb{R}^3\) satisfying

\[
\|A^{(j)}[t]\|_{L^\infty(H^\sigma \times H^{\sigma-1})([0, 1] \times \mathbb{R}^3)} \lesssim \|((\bar{A}^\lambda, \bar{E}^\lambda))\|_{W^{\sigma,2} \times W^{\sigma-1,2}(B_j)} \quad (3.2)
\]

One obtains a solution \(A^\lambda\) on \([0, 1] \times \mathbb{R}^3\) defined by each \(A^{(j)}\) restricted to \(D_j\) since any two of these solutions coincide on their common domain as a consequence of finite speed of propagation in the temporal gauge.

We apply the square summability estimate (3.1) at the initial time and at an arbitrary time \(t \in [0, 1]\), together with (3.2) obtaining

\[
\|A^\lambda[t]\|_{H^\sigma \times H^{\sigma-1}}^2 \simeq \sum_j \|A^{(j)}[t]\|_{W^{\sigma,2} \times W^{\sigma-1,2}(D_j \cap \{t\} \times \mathbb{R}^3)}^2 \lesssim \sum_j \|((\bar{A}^\lambda, \bar{E}^\lambda))\|_{W^{\sigma,2} \times W^{\sigma-1,2}(B_j)}^2 \simeq \|((\bar{A}^\lambda, \bar{E}^\lambda))\|_{H^\sigma \times H^{\sigma-1}}^2.
\]

Undoing the scaling one obtains a solution \(A \in C_t H^\sigma \cap C^1_t H^{\sigma-1}([0, 1] \times \mathbb{R}^3)\) given by \(A(t, x) = \lambda A^\lambda(\lambda t, \lambda x)\).

If \((\bar{A}', \bar{E}') \in H^\sigma \times H^{\sigma-1}(\mathbb{R}^3)\) is another initial data set satisfying (1.3) one has

\[
\|A^{(j)}[] - A'^{(j)}[\cdot]\|_{L^\infty(H^\sigma \times H^{\sigma-1})} \lesssim \|((\bar{A}^\lambda - \bar{A}'^\lambda, \bar{E}^\lambda - \bar{E}'^\lambda))\|_{W^{\sigma,2} \times W^{\sigma-1,2}(B_j)} \quad (3.3)
\]

By repeating the same square summability argument for differences using (3.3) and undoing the scaling one obtains Lipschitz dependence

\[
\|A[t] - A'[t]\|_{H^\sigma \times H^{\sigma-1}} \lesssim \|(\bar{A}, \bar{E}) - (\bar{A}', \bar{E}')\|_{H^\sigma \times H^{\sigma-1}}.
\]

The Lipschitz constant depends, of course, on the size of the initial data.
4. Modified local well-posedness

In this section we derive control of space-time norms that will be needed for the trilinear estimate (1.32).

Let \( A_i \) be a solution to \((\text{YM})\) in the temporal gauge \( A_0 = 0 \). We use the Leray projections (2.6) to split \( A = PA + P^\perp A \) into a divergence-free part and a curl-free part. From the definition of \( P^\perp \) and the first equation in (1.2), and by applying \( P \) to the second equation in (1.2) one obtains

\[
\begin{align*}
\partial_t P^\perp A &= \Delta^{-1} \nabla \partial_t [A_j, A_j] \\
\Box P A &= -2P[A_j, \partial_j A] + P[A, \partial_j A_j] - P[A_j, [A_j, A]]
\end{align*}
\]

(4.1)

Following [50] one proceeds to uncovering the null structure in (4.1) as follows.

Recall the definition of a null form from (2.8), (2.9). For Lie algebra valued fields, \( N(A, B) \) is understood as a linear combination of

\[
\Delta^{-1} \nabla^i Q_{ij}(A, B),
\]

and

\[
Q_{ij}(C, D) := [\partial_i C, \partial_j D] - [\partial_j C, \partial_i D]
\]

(4.2)

(4.3)

Denote \( A^{df} = PA \) and \( A^{cf} = P^\perp A \). Using the identities (2.7) one can write the system (4.1) as

\[
\begin{align*}
\partial_t A^{cf} &= \nabla^{-1} \mathcal{O}(\partial_t A, A) \\
\Box A^{df} &= N(A^{df}, A^{df}) + \mathcal{O}(A^{df}, \nabla A^{cf}) + \mathcal{O}(A^{cf}, \nabla A^{df}) + \mathcal{O}(A^{cf}, \nabla A^{cf}) + \mathcal{O}(A^3)
\end{align*}
\]

(4.4)

where \( \mathcal{O} \) denote (matrix-valued) bilinear or trilinear expressions in the values of the inputs.

Well-posedness in \( I^{-1}H^1 \times I^{-1}L^2 \) holds for this system.

**Proposition 4.1.** (1) Let \( A_x(t, x) \) be a global regular solution to \((\text{YM})\) in the temporal gauge \( A_0 = 0 \) such that

\[
\|IA_x(t_0)\|_{H^1} + \|I\partial_t A_x(t_0)\|_{L^2} \leq \varepsilon
\]

Let \( t_1 \leq t_0 + 1 \). Then, assuming \( \varepsilon \) is small enough, one has

\[
\|IA_x(t)\|_{L^\infty_t(H^1 \times L^2)([t_0, t_1] \times \mathbb{R}^3)} \leq C\varepsilon
\]

(4.5)

(2) Moreover, there exists a spatial field \( U : \mathbb{R}^3 \to G, U = U(x) \) such that the change of gauge

\[
A_i \mapsto U A_i U^{-1} - \partial_i U U^{-1}
\]

achieves \( \text{div} A_x(t_0) = 0 \), the bound (4.5) and

\[
\begin{align*}
IA^{df} &\in \varepsilon X^{1+\frac{3}{4}+}, & I\partial_t A^{df} &\in \varepsilon X^{0+\frac{3}{4}+} \\
IA^{cf} &\in \varepsilon X^{1+\frac{1}{4}+}, & P_k P_l \partial_l A^{cf} &\in \varepsilon L^\infty_t H^\frac{1}{2}
\end{align*}
\]

(4.7)

(4.8)

hold on \([t_0, t_1] \times \mathbb{R}^3\), where \( A^{df} = PA_x, A^{cf} = P^\perp A_x \).

This proposition follows from the local well-posedness theory developed in [50].
4.1. Proof of Proposition 4.1

Step 1. We begin by discussing the change of gauge that sets $A^g(t_0) = 0$, i.e. the Coulomb gauge $\text{div} A_x(t_0) = 0$, only at the initial time. Without this transformation (4.8) could not have more regularity than the initial data.

The following procedure is adapted here to $I^{-1}H^1$ from the $H^s$ case in [50, Section 3] (arxiv version), attributed there to Mark Keel. Suppose

$$\|IA_x(t_0)\|_{H^1} \leq \varepsilon, \quad \|IA^g(t_0)\|_{H^1} \leq \delta$$

for $\delta \leq \varepsilon \ll 1$. Write $A^g(t_0) = \nabla V$ for a $g$-valued field $V$ with $\|V\|_X \lesssim \delta$ where the $X$ norm is defined in (2.23) and obeys the algebra and product estimates (2.25).

Applying (4.6) with $U := \exp(V)$ one obtains a new temporal solution, which by abuse of notation we still denote by $A$, with initial data written as

$$\tilde{A}(t_0) = e^V A^g(t_0) e^{-V} + \left( e^V \nabla V - \nabla (e^V) \right) e^{-V}$$

and

$$e^V A^g(t_0) e^{-V} = A^g(t_0) + \left( e^V A^g(t_0) e^{-V} - A^g(t_0) \right)$$

Based on these equalities, Taylor expansions and (2.25) one obtains that

$$\|\tilde{A}(t_0)\|_{H^1} \leq \varepsilon + C\varepsilon \delta, \quad \|IA^g(t_0)\|_{H^1} \leq C\varepsilon \delta, \quad \|U - 1\|_{X} \lesssim \delta$$

Applying this procedure, one obtains in the limit a new temporal solution, which by abuse of notation we still denote by $A_x$, such that $A^g(t_0) = 0$. The change of gauge (4.6) is done with a $U$ with $\|U - 1\|_{X} \lesssim \varepsilon$. The new solution has

$$\|IA_x(t_0)\|_{H^1} + \|I\partial_t A_x(t_0)\|_{L^2} \lesssim \varepsilon \quad (4.9)$$

and (4.3) follows using (2.25) if we prove

$$\|IA_x(t)\|_{L^2(H^1 \times L^2)([t_0,t_1] \times \mathbb{R}^3)} \lesssim \varepsilon \quad (4.10)$$

for the new solution. Thus it remains to prove the bounds in part (2) of Prop. 4.1.

Step 2. We now consider a (YM) solution in the temporal gauge satisfying (4.9) with $A^g(t_0) = 0$. The system of equation solved by $IA$ is written schematically as

$$\Box IA^g = IN, \quad \partial_t IA^g = I\nabla^{-1}(A\partial_t A) \quad (4.11)$$

$\mathcal{N} := N(A^g, A^g) + A^g \nabla A^g + A^g \nabla A^g + A^g \nabla A^g + A^g$

For $J = [t_0, t_1]$ let

$$\|IA\|_X := \left\|IA^g\right\|_{X_{1,J}^{1/2} +} + \left\|I\partial_t A^g\right\|_{X_{0,J}^{1/2} +} + \left\|IA^g\right\|_{X_{1,J}^{1/2} +}$$

One obtains a-priori estimates using the following bounds:

$$\|\mathcal{N}(A_1, A_2)\|_{X_{0,J}^{-1/2} +} \lesssim \|IA_1\|_{X_{1,J}^{1/2} +} \|IA_2\|_{X_{1,J}^{1/2} +}$$

$$\|I(A_1 \nabla A_2)\|_{X_{0,J}^{1/2} +} + \|I(A_2 \nabla A_1)\|_{X_{0,J}^{1/2} +} \lesssim \|IA_1\|_{X_{1,J}^{1/2} +} \|IA_2\|_{X_{1,J}^{1/2} +}$$

$$\|I(A_1 \nabla A_2)\|_{X_{0,J}^{1/2} +} \lesssim \|IA_1\|_{X_{1,J}^{1/2} +} \|IA_2\|_{X_{1,J}^{1/2} +}$$

$$\|I(A_1 A_2 A_3)\|_{X_{0,J}^{1/2} +} \lesssim \prod_{i=1}^{3} \min \left( \|IA_i\|_{X_{1,J}^{1/2} +}, \|IA_i\|_{X_{1,J}^{1/2} +} \right)$$
and

\[ \| I \nabla^{-1} (A_1 \partial_t A_2) \|_{X^s_{r,0} + \frac{1}{a} +} \lesssim \| I A_1 \|_{X^s_{r,+}} \| I A_2 \|_{X^s_{r,+}} \]

\[ \| I \nabla^{-1} (A_1 \partial_t A_2) \|_{X^s_{r,0} - \frac{1}{a} +} + \| I \nabla^{-1} (A_2 \partial_t A_1) \|_{X^s_{r,0} + \frac{1}{a} +} \lesssim \| I A_1 \|_{X^s_{r,+}} \| I A_2 \|_{X^s_{r,+}} \]

\[ \| I \nabla^{-1} (A_1 \partial_t A_2) \|_{X^s_{r,0} + \frac{1}{a} +} \lesssim \| I A_1 \|_{X^s_{r,+}} \| I A_2 \|_{X^s_{r,+}} \]

holding for any \( A_1, A_2, A_3 \) on \( \mathbb{R}^{3+1} \).

These estimates are stated and proved in \([50]\) Eq (15)-(21) without the \( I \) multipliers and with \( X^{s,\frac{3}{2}+}, X^{s,-\frac{1}{2}+}, X^{s,\frac{1}{2}+}, X^{s,\frac{3}{2}-} \) instead of \( X^{\frac{3}{2}+}, X^{0,-\frac{1}{2}+}, X^{\frac{1}{2}+}, X^{1,-\frac{1}{2}+} \) for any \( s > \frac{3}{2} \). The version stated here follows immediately from those by doing basic Fourier decompositions. For details see the interpolation lemma in \([5] \) Lemma 12.1.

Combining these estimates with the energy-type inequalities \((2.28), (2.29)\) for \( X^{s,b}, X_{r,\theta} \) spaces and \( A^{cf}(t_0) = 0 \) one obtains

\[ \| I A \|_{X} \lesssim \varepsilon + \| I A \|^2_X + \| I A \|^3_X \]

from which one obtains \( \| I A \|_{X} \lesssim \varepsilon \) (if \( \varepsilon \) is small enough) by a continuity argument in \( t_1 \). This proves the first three bounds in \((4.7), (4.8)\). In particular,

\[ \| I A^{cf}(t) \|_{L_\infty^\tau(H^1 \times L^2)} + \| I A^{cf} \|_{L_\infty^\tau H^{1+\frac{1}{2}}} \lesssim \varepsilon \]

From \((4.11)\), using the inequality \((2.22)\), for any \( t \in [t_0, t_1] \) one obtains

\[ \| I \partial_t A^{cf} \|_{L^2} \lesssim \| I A \|_{H^1} \left( \| I \partial_t A^{cf} \|_{L^2} + \| I \partial_t A^{cf} \|_{L^2} \right) \lesssim \varepsilon^2 + \varepsilon \| I \partial_t A^{cf} \|_{L^2} \]

Absorbing the last term to the LHS completes the proof of \((4.10)\).

In fact, we can improve the previous inequality at high frequencies - we obtain the last bound in \((4.8)\) by the Littlewood-Paley trichotomy for the equation \( \partial_t A^{cf} = \nabla^{-1}(A \partial_t A) \) from the \( H^1 \) and \( L^2_x \) bounds already obtained in \((4.10)\) and using Bernstein’s inequality for the low frequency (input or output).

4.1.1. Improved bounds for \( \partial_t A^{cf} \).

**Proposition 4.2.** Let \( A_x \) be the solution from Proposition 4.1 Part (2) on \([t_0, t_1] \times \mathbb{R}^3\). Then

\[ \| P_k \partial_t A^{cf} \|_{L^2_t L^2_x} \lesssim \varepsilon 2^{\frac{1}{2}+k} \left( \frac{\varepsilon k}{N} \right)^{2(1-\sigma)} \]  

\[ \| P_k \partial_t A^{cf} \|_{L^2_t L^2_x} \lesssim \varepsilon 2^{-\frac{1}{2}k} \left( \frac{\varepsilon k}{N} \right)^{2(1-\sigma)} \]  

\[ \| P_k \partial_t A^{cf} \|_{L^\infty_t L^\infty_x} \leq \varepsilon 2^{(1+\sigma)k} \left( \frac{\varepsilon k}{N} \right)^{1-\sigma} \]

**Proof.** For \((4.12), (4.13)\) we use the equation

\[ \partial_t A^{cf} = \nabla^{-1}(A^{df} \partial_t A^{df} + A^{df} \partial_t A^{cf} + A^{cf} \partial_t A^{df} + A^{cf} \partial_t A^{cf}) \]

and perform a Littlewood-Paley decomposition of the RHS.

**Proof of (4.12)** First consider the case when the output is a high frequency (low-high or high-low cases). When and at least one of the inputs is \( df \) - place that input in \( L^{2+} L^{\infty} \) using Strichartz \((2.30)\) and the other term in \( L^{\infty} L^{\infty} \) (using Bernstein and \( L^{\infty} L^2 \)). When both inputs are \( cf \) use \( L^{\infty} L^{\infty} \) norms (and Bernstein with \( L^{\infty} L^2 \)), based on \((4.8)\).
In the case of high-high-to-low frequency interactions we use Bernstein first. When both inputs are \( \partial f \) - estimate both inputs in \( L^\infty L^2 \) using (4.8) and the output in \( L^2 L^1. \) When both inputs are \( \partial f \) use the bilinear Strichartz estimate (2.31) to bound the output in \( L^2 L^2. \)

Now assume one of the high-high inputs is \( \partial f \) and one is \( \partial f. \) For \( A^{\partial f} \partial f A^{\partial f} \) estimate the output in \( L^2 L^2, \) placing the \( \partial f \) term in \( L^{2+} L^\infty \) and the \( \partial f \) term in \( L^\infty L^2. \) For \( A^{\partial f} \partial f A^{\partial f} \) estimate the output in \( L^2 L^2, \) placing \( A^{\partial f} \) in \( L^\infty L^2 \) and \( \partial f A^{\partial f} \) in \( L^2 L^8 \) (by interpolating between \( L^2 L^\infty \) and energy - \( L^2 L^2 \)). This is done to ensure adequate summation even when \( \sigma \) is close to \( \frac{5}{2}. \)

**Proof of (4.14)** This inequality follows from (4.12) and from (4.7) together with Strichartz (2.30) for \( L^{2+} L^\infty \) and Hölder in \( t. \)

**Proof of (4.13)** For high-high-to-low frequency interactions we argue exactly the same as for (4.12), only the power obtained from Bernstein’s inequality changes by a factor of \( 2^{-\frac{3}{4}k}. \) The same is true for low-high and high-low \( A^{\partial f} \partial f A^{\partial f} \) interactions.

For the remaining high-low interactions place \( A \) (high frequency) in \( L^\infty L^2 \) and \( \partial f A \) (low frequency) in \( L^2 L^\infty \) using (4.14).

For low-high interactions, place \( A^{\partial f} \) (low) in \( L^2 L^\infty \) and \( \partial f A \) (high) in \( L^\infty L^2. \) It remains to deal with \( A^{\partial f} \partial f A^{\partial f} \) for \( k_1 \leq k. \) This term is responsible for the \( 2^{-\frac{3}{4}k} \) factor in (4.13), rather than \( 2^{-k} \) (which can probably be fully obtained). Place \( A^{\partial f} \partial f A^{\partial f} \) in \( L^\infty L^{12} \) using (4.8) and Bernstein and place \( \partial f A^{\partial f} \) in \( L^\infty L^{12} \) using Bernstein again.

\[
\Box
\]

5. Yang-Mills Heat Flow solutions

This section is devoted to the Yang-Mills heat flow. We consider deTurck’s and caloric gauges including: well-posedness in \( \Gamma^{-1} H^1 \) and \( H^p, \) control of the modified energy (1.6) from the initial data and the change between the two gauges.

5.1. Local well-posedness in deTurck’s gauge.

In the covariant (cYMHP) equation, setting the deTurck gauge condition \( A_s = \partial^\ell A_\ell \) one obtains the parabolic equation for \( A_\ell \)

\[
(\partial_s - \Delta)A_\ell = 2 \left[ A^\ell, \partial_\ell A_s \right] - \left[ A^\ell, \partial_s A_\ell \right] + \left[ A^\ell, [A_\ell, A_s] \right]
\]

(5.1)

In short, this equation can be written schematically as \( (\partial_s - \Delta)A = A \nabla A + A^3. \)

One can separate the linear and nonlinear parts of \( A_\ell: \)

\[
A_\ell(s) = e^{s\Delta} A_\ell(0) + A_\ell^{bil}(s)
\]

(5.2)

where the bilinear term is written schematically as

\[
A_\ell^{bil}(s) := \int_0^s e^{(s_1-s)\Delta} (A \nabla A + A^3) \, ds
\]

(5.3)

and will enjoys more favorable estimates, such as (5.14).

This equation is locally well-posed in \( \Gamma^{-1} H^1: \)

**Proposition 5.1.** Suppose the initial data satisfies \( IA_\ell \in \varepsilon H^1 \) on \( \mathbb{R}^3 \) and \( \varepsilon > 0 \) is small enough. Then there exists a solution to (5.1) on \( \mathbb{R}^3 \times [0, 1] \) with \( A_\ell(s = 0) = A_\ell \) which satisfies

\[
IA_\ell \in \varepsilon L^\infty_s H^1(\mathbb{R}^3 \times [0, 1])
\]

and is the unique limit of regular solutions.
One may use the following parabolic energy estimates, [37, Proposition 3.12]

\[ \| \int_0^{s_1} e^{(s_1-s)\Delta} (A\nabla B) \, ds \|_S \lesssim \| A \|_S \| B \|_S \]

\[ \| \int_0^{s_1} e^{(s_1-s)\Delta} (A_1 A_2 A_3) \, ds \|_S \lesssim \| A_1 \|_S \| A_2 \|_S \| A_3 \|_S \]

One may use the space

\[ \| A \|_S := \| \left\langle 2^{2k} s \right\rangle^{10} \| P_k IA(s) \|_{H^1} \| L^\infty_{x} \| L^2_{t \in \mathbb{N}} \]  

(5.7)

to check these estimates using basic Littlewood-Paley theory.

One may also use the following parabolic energy estimates, [37, Proposition 3.12]

\[ \| \int_0^{s_1} e^{(s_1-s)\Delta} N(s) \, ds \|_S \lesssim \| s N \|_{L^1_{x} \cap L^2_{x} (L^2_{t})} \]

\[ \| f \|_S := \| f \|_{L^\infty_{x} L^2_{t}} + \| s^{\frac{1}{2}} \partial_x f \|_{L^\infty_{x} \cap L^2_{x} (L^2_{t})} + \| s \partial^2_x f \|_{L^2_{x} L^2_{t}} \]

(5.8)

used in the proof of [37, Proposition 5.2] which is the corresponding well-posedness statement for \( H^1 \). Here one writes the equation as

\[ (\partial_s - \Delta) I \langle D_x \rangle A = I \langle D_x \rangle (A \nabla A + A^3) \]

(5.9)

One defines \( \| A \|_S := \| I \langle D_x \rangle A \|_S \). The needed estimates to check for this case of \( I^{-1} H^1 \), as opposed to \( H^1 \), become

\[ \| I \langle D \rangle (f \nabla g) \|_{L^2} \lesssim \| f \|_{H^{\frac{3}{2} + (1-\sigma)}_{x}} \| g \|_{H^{2}} \]

\[ \| I \langle D \rangle (f_1 f_2 f_3) \|_{L^2} \lesssim \| f_1 \|_{H^{2-\frac{\sigma}{2}}} \| f_2 \|_{H^{2-\frac{\sigma}{2}}} \| f_3 \|_{H^{2-\frac{\sigma}{2}}} \]

(5.10)

which follow easily from (2.19) or from Littlewood-Paley, by splitting into different cases depending on whether the inputs have frequencies higher or lower than \( N \).

By similar arguments for the differentiated equation one obtains persistence of regularity. We refer to [37, Proposition 5.2] for more details.

The bounds (5.5) follow from the \( \| A \|_S \) bound and the definition of \( A_s = \partial_s A_t \).

Denoting by \( N \) the nonlinear RHS of (5.1), for our solution, the bounds above show

\[ \| I \langle D_x \rangle N \|_{L^1_{x} L^2_{t}} \lesssim \varepsilon^2 \]

(5.11)

We use this to prove (5.13). The equation for \( I A_s \) is of the type \( (\partial_s - \Delta) I A_s = I \partial_s N \). Integrating the equation and taking \( L^2_{x} \) norms one obtains

\[ \| \int_{s_1}^{s_2} I \Delta A_s \, ds \|_{L^2_{x}} \leq \| I A_s(s_1) \|_{L^2_{x}} + \| I A_s(s_2) \|_{L^2_{x}} + \| I \partial_s N \|_{L^1_{x} L^2_{t}} \lesssim \varepsilon. \]

24
As a consequence one obtains gauge-invariant bounds for the curvature, such as control of the modified energy \( E_0 \).

**Corollary 5.2.** Suppose \( A_t(t, x) \) is a regular solution of the \((\text{YM})\) equation in the temporal gauge on \([t_0, t_1] \times \mathbb{R}^3\) with \((IA_t, IF_0) \in \epsilon L^\infty_t(\mathcal{L}^1 \times \mathcal{L}^2)([t_0, t_1] \times \mathbb{R}^3) \).

Suppose \( A_\alpha(t, x, s) \) solves \((\text{YMHe})\) on \([t_0, t_1] \times \mathbb{R}^3 \times [0, s_0] \) in DeTurck’s gauge \( A_\alpha = \partial' A_\ell \), where \( s_0 = N^{-2} \). Then, assuming \( \epsilon \ll 1 \), for any \( t \in [t_0, t_1] \) one has

\[
\|(Ns^{1\over 2})^{1-\sigma} F_{\alpha \beta}(s)\|_{L^2_x L^2_t} + \|((Ns^{1\over 2})^{1-\sigma} F_{\alpha \beta}(s))_{L^2_x {\partial_t}^2} \| \lesssim \epsilon \]  

(5.12)

As a consequence, for \((\text{5.10})\) one has the gauge-invariant bound \( IE(t) \lesssim \epsilon^2 \) for all \( t \in [t_0, t_1] \).

**Proof.** Fix a time \( t \in [t_0, t_1] \). For \( A_i \) we apply Proposition 5.1 and so it obeys the estimates established in that proof.

We begin with \( F_{ij}(s) \). First note that \( \|[A_i, A_j](s)\|_{L^2_x} \lesssim \epsilon^2 \) since from (5.4) and Sobolev embedding we have \( A(s) \in \epsilon L^4_x \). From the \( \|I(\partial_x^k)A(s)\|_{\tilde{S}} \) norm and interpolation one obtains \( (Ns^{1\over 2})^{1-\sigma} \|\partial_j A_i(s)\|_{L^2_x} \lesssim \epsilon \) and the same for \( \partial_j A_i \) which concludes the first term in (5.12) for \( F_{ij} \). For the second term we note that from control of (5.7) we have \( \|P_k A_i(s)\|_{H^1} \lesssim \alpha_k \left( {\epsilon^k \over s^k} \right)^{1-\sigma/(2k^2 s)} \) for all \( k \geq 0, s \in [0, s_0] \), for some sequence with \( \|\alpha_k\|_{\ell^\infty_k} \lesssim \epsilon \). We bound

\[
\|(Ns^{1\over 2})^{1-\sigma} \partial_j A_i(s)\|_{L^2_x L^2_t} \lesssim \sum_{0 \leq k \leq N} \int_0^{s_0} (N^2 s)^{-1\over 2} \alpha_k \frac{ds}{s} + \sum_{k \geq N} \alpha_k^2 \int_0^{s_0} \left( {\epsilon^k \over s^k} \right)^{1-\sigma} \frac{ds}{s}
\]

which is \( \lesssim \epsilon^2 \).

Now consider \( F_{0i}(s) \), which solves the covariant equation (6.1). Expanding it and canceling the \([A_s, F] \) term with the \([\partial' A, F] \) term one obtains equation \((6.3)\), which can be written in the form \((5.8)\) for \( B = F_{0i} \). As before, the solution is obtained by Picard iteration. One may use the parabolic energy estimate \((5.8)\) to estimate the solution of the equation

\[
(\partial_s - \Delta) IB = I(A \nabla B + \nabla A B + A^2 B), \quad IB(0) \in \epsilon L^2 \]

in the norm \( \|B\|_{L^{1+\tilde{S}}_t} = \|IB\|_{\tilde{S}} \). The estimates we need now are

\[
\|I(A \nabla B)\|_{L^2} + \|I(B \nabla A)\|_{L^2} \lesssim \|IA\|_{H^{1/2+ (1-\sigma)/2}} \|IB\|_{H^{1+}}
\]

\[
\|I(A^2 B)\|_{L^2} \lesssim \|IA\|_{H^{2-\sigma}} \|IB\|_{H^{2-\sigma}}
\]

which follow from \((2.19), (2.20)\) or from Littewood-Paley, by splitting into frequencies higher or lower than \( N \). For \( A \) we already have control of the norms from the proof of Proposition 5.1. This provides \( \|IF_{0i}\|_{\tilde{S}} \lesssim \epsilon \). One obtains control in the norm \( \|I(2k^2 s)\|_{L^2} \|IIF_{0i}(s)\|_{L^2_x L^2_t} \|_{\ell^\infty_k} \) as well. Then we conclude exactly like for \( F_{ij} \).

We record some \( L^\infty_x \) bounds separately.

**Lemma 5.3.** Suppose \( A_i \) obeys the conditions of Proposition 5.1 uniformly on an interval \( I \times \mathbb{R}^3 \times [0, s_0] \). Then

\[
\|A_i(s)\|_{L^\infty_t L^\infty_x} \leq \epsilon s^{-1/2} / (Ns^{1/2})^{1-\sigma} \]  

(5.13)

\[
\|A_i^{bl}(s)\|_{L^\infty_t L^\infty_x} \leq \epsilon / (N^2 s)^{1-\sigma} \]  

(5.14)
Proof. Fix a time $t$. The first bound \((5.13)\) follows immediately from the bounds in Proposition \ref{Proposition5.1}. Since \((s_1 - s)^{\frac{3}{2}}e^{(s_1 - s)\Delta} : L^2_x \to L^\infty_x\), we can bound using \((5.13)\)
\[
\|A_i^{p_i}(s_1)\|_{L^\infty_x} \lesssim \int_0^{s_1} (s_1 - s)^{-\frac{3}{4}} \|A\|_{L^\infty_x} \|\nabla A\|_{L^2_x} + \|A\|_{L^3_x}^3 ds
\]
\[
\lesssim \int_0^{s_1} (s_1 - s)^{-\frac{3}{4}} s^{-\frac{1}{4}}/(N s^{\frac{3}{4}})^{2(1-\sigma)} + s^{-\frac{1}{4}}/(N s^{\frac{3}{4}})^{3(1-\sigma)} ds.
\]
which proves \((5.14)\) after integration. \hfill \Box

Remark 5.4. The same argument as in Proposition \ref{Proposition5.1} proves well-posedness of \((5.1)\) in $\dot{H}^\rho$ for any $\frac{3}{2} < \rho < 1$. Instead of \((5.9)\) one considers $(\partial_s - \Delta)|D|^\rho A = |D|^{\rho}(A\nabla A + A^2)$ together with the norm $\|A\|_{S^\rho} := \||D|^\rho A\|_{\tilde{S}}$. The estimates \((5.10)\) are replaced by
\[
\| |D|^\rho (f\nabla g)\|_{L^2} \lesssim \|f\|_{\dot{H}^{\frac{3}{4}} L^\infty} \|g\|_{\dot{H}^{1+\rho}}
\]
\[
\| |D|^\rho (f_1 f_2 f_3)\|_{L^2} \lesssim \|f_1\|_{\dot{H}^{1+\frac{3}{4}}} \|f_2\|_{\dot{H}^{1+\frac{3}{4}}} \|f_3\|_{\dot{H}^{1+\frac{3}{4}}}.
\]
See \((2.20)\). Control in the norm $\|\|\langle 2^k s \rangle \|P_k A(s)\|_{\dot{H}^\rho} L^{\infty}_x\|_{L^2_k}$ holds as well. One obtains

Proposition 5.5. Let $\frac{1}{2} < \rho < 1$. Suppose the initial data satisfies $\tilde{A}_i \in \varepsilon \dot{H}^\rho(\mathbb{R}^3)$ and $\varepsilon > 0$ is small enough. Then there exists a solution to \((5.11)\) on $\mathbb{R}^3 \times [0, 1]$ with $A_i(s = 0) = \tilde{A}_i$ which is in $S^\rho$ and is the unique limit of regular solutions. In particular, one has
\[
\sup_{s \in [0,1]} s^{\frac{3}{2} - \rho} \|A_i(s)\|_{\dot{H}^{1+\rho}} + \|s^{\frac{3}{2} - \rho} A_i(s)\|_{L^2_x \dot{H}^{1+\rho}} \lesssim \varepsilon \quad \forall \rho \in (\rho, \rho + 1].
\]

Similarly to Proposition \ref{Proposition5.3} one has a statement for the equation satisfied by $B_i = F_{0i}$ in DeTurck’s gauge:

Proposition 5.6. Let $-\frac{1}{2} < \gamma < 0$. Suppose the connection coefficients $A_i$ satisfy the conditions in Proposition \ref{Proposition5.7} and Remark \ref{Remark5.4} on $\mathbb{R}^3 \times [0, 1]$. Suppose the initial data satisfies $B_i \in \varepsilon \dot{H}^{-\gamma}(\mathbb{R}^3)$ and $\varepsilon > 0$ is small enough. Then the schematic equation
\[
(\partial_s - \Delta)B = A\nabla B + B\nabla A + A^2 B
\]
has a unique solution on $\mathbb{R}^3 \times [0, 1]$ with $B_i(s = 0) = \tilde{B}_i$ and satisfies
\[
\sup_{s \in [0,1]} s^{\frac{3}{4} - \gamma} \|B_i(s)\|_{\dot{H}^{-\gamma}} + \|s^{\frac{3}{4} - \gamma} B_i(s)\|_{L^2_x \dot{H}^{-\gamma}} \lesssim \varepsilon \quad \forall \gamma \in (\gamma, \gamma + 1].
\]

Proof. As before, the solution is obtained by Picard iteration. One uses the parabolic energy estimate \((5.8)\) to estimate the solution of the equation
\[
(\partial_s - \Delta)|D|^\gamma B = |D|^\gamma (A\nabla B + \nabla AB + A^2 B)
\]
in the norm $\|B\|_{S^\gamma} := \||D|^\gamma B\|_{\tilde{S}}$. Recall \((2.20)\). The estimates we need are
\[
\|A\nabla B\|_{\dot{H}^{\gamma}} + \|B\nabla A\|_{\dot{H}^{\gamma}} \lesssim \|A\|_{\dot{H}^{\frac{3}{4}} L^\infty} \|B\|_{\dot{H}^{1+\gamma}}
\]
\[
\|A^2 B\|_{\dot{H}^{\gamma}} \lesssim \|A\|_{\dot{H}^{1+\gamma}} \|B\|_{\dot{H}^{1+\gamma}}.
\]
One obtains control in the norm $\|\|\langle 2^k s \rangle \|P_k B(s)\|_{\dot{H}^{\gamma}} L^\infty_x\|_{L^2_k}$ as well. \hfill \Box

Now we can obtain estimates for the curvature at the initial time $t = 0$. 

26
Corollary 5.7. Suppose $A_i(t, x)$ is a regular solution to the Yang-Mills equation in the temporal gauge on $[-1, 1] \times \mathbb{R}^3$ with initial data $(\dot{A}_i, \dot{E}_i) \in \varepsilon(\dot{H}^\sigma \times \dot{H}^{\sigma-1}) (\mathbb{R}^3)$ and $A_i \in \dot{H}^{\frac{3}{2}+}(\mathbb{R}^3)$.

Suppose $A_o(t, x, s)$ solves \textbf{(YMIII)} on $[-1, 1] \times \mathbb{R}^3 \times [0, 1]$ in DeTurck’s gauge $A_s = \partial^s A_t$. Then, assuming $\varepsilon \ll 1$, at $t = 0$ one has

$$
\|s^{-\frac{1}{2}} F_{\alpha \beta}(s)\|_{L^\infty_x L^4_x} + \|s^{-\frac{1}{2}} F_{\alpha \beta}(s)\|_{L^3_x L^2_x} \lesssim \varepsilon
$$

As a consequence, at $t = 0$, one has the gauge-invariant version

$$
\sup_{s \in [0, 1]} \int_{\mathbb{R}^3} (F_{\alpha \beta}, F_{\alpha \beta})(s) \, dx + \int_0^1 \int_{\mathbb{R}^3} (F_{\alpha \beta}, F_{\alpha \beta})(s) \, dx \, ds \lesssim \varepsilon^2.
$$

Proof. Consider $F_{ij}(s)$ first. For $\partial_i A_j$ and $\partial_j A_i$ we use Proposition \textbf{5.5} with $\rho = \sigma$ and $\rho_1 = 1$. For $[A_i, A_j]$ we use that from \textbf{5.16} and applying again Proposition \textbf{5.5} with $\rho = \frac{1}{2}$ one has

$$
s^{-\frac{1}{2}} \|A_i(s)\|_{H^{1-\epsilon}} \lesssim \varepsilon, \quad \|A_i(s)\|_{\dot{H}^{\frac{1}{2}+}} \lesssim 1.
$$

The bounds for $[A_i, A_j]$ follow from this with Hölder and Sobolev embeddings $H^{1-} \times \dot{H}^{\frac{3}{2}+} \subseteq L^6 \times L^{3+} \subseteq L^2$.

Now consider $F_{0j}(s)$, which solves the covariant equation \textbf{6.1}. Expanding it and canceling the $[A_s, F]$ term with the $[\partial^s A_t, F]$ term one obtains equation \textbf{5.3}, which can be written as \textbf{5.17}. Then \textbf{5.18} is concluded by Proposition \textbf{5.6} with $\gamma = -1$ and $\gamma_1 = 0$. \hfill \square

The next goal is to switch to caloric gauges which satisfy $A_s = 0$. This is accomplished by a change of gauge (DeTurck’s trick) obtained by solving the ODE \textbf{5.19}, see Proposition \textbf{5.10}.

Lemma 5.8. Let $A_i, A_s = \partial^s A_t$ be given by Proposition \textbf{5.7}. Let $U(s)$ be the solution of the following ODE

$$
\partial_t U = U A_s, \quad U(s = 0) = 1
$$

Then one has

$$
U - 1 \in \varepsilon L^\infty_x L^2_x, \quad \nabla_x U \in \varepsilon L^\infty_x L^2_x, \quad \Delta U \in \varepsilon L^\infty_x L^2_x
$$

and $U^{-1}$ obeys the same bounds. If $A_i$ is a regular solution then $U$ is a regular gauge transformation.

Proof. Write the integral form

$$
U(s_1) = 1 + \int_0^{s_1} U A_s \, ds
$$

The first two bounds in \textbf{5.20} follow from the first two bounds in \textbf{5.5} using Gronwall’s inequality. Similarly one shows $U - 1 \in \varepsilon L^\infty_x \dot{H}^{\frac{3}{2}+}$.

The third bound in \textbf{5.20} is more delicate. We use an argument from \textbf{[27, Theorem 7]} and \textbf{[37, Appendix B]}. One writes

$$
I \Delta U(s_1) = I \int_0^{s_1} \Delta U A_s + 2 \nabla U \nabla A_s + U \Delta A_s \, ds
$$

The first two terms in the integral are handled by Gronwall’s inequality and by

$$
\|I(fg)\|_{L^2} \lesssim \|I f\|_{L^2} \|g\|_{\dot{H}^{\frac{3}{2}+}},
$$

$$
\|I(fg)\|_{L^2} \lesssim \|I f\|_{H^1} \|g\|_{\dot{H}^{\frac{3}{2}+}}
$$

for $f = \Delta U, \ g = A_s$, respectively $f = \nabla U, \ g = \nabla A_s$, provided we can also bound the third term.

For the third term one plugs in $U(s)$ from \textbf{5.21} and by Fubini’s theorem write

$$
\int_0^{s_1} I \Delta A_s \, ds + I \int_0^{s_1} U(s') A_s(s') \left( \int_0^{s_1} \Delta A_s(s) \, ds \right) \, ds'
$$
Bound this in $L^2$ using (5.6) for the first integral, then use (5.22) for $g = U(s')A_s(s')$ and $f = \int_{s'}^s \Delta A_s(s) \, ds \in \varepsilon I^{-1}L^2$ by (5.6). Note that $UA_s \in \varepsilon L^1H_{2+}^2$ by using the algebra property for $H_{2+}^2$.

\[ \square \]

5.2. Caloric gauge solutions.

Setting the caloric condition $A_s = 0$ in the covariant (cYMHF) and the dynamic (dYMHF) Yang-Mills heat flows one obtains the initial value problems

\[ \partial_s A_i = D^\ell F_{\ell i}, \quad A_i(s = 0) = \bar{A}_i \]  
\[ \partial_s A_\alpha = D^\ell F_{\ell \alpha}, \quad A_\alpha(s = 0) = \bar{A}_\alpha \]

**Theorem 5.9** (Existence and uniqueness for regular solutions, [38, 43]).

1. Consider the initial value problem (5.24), i.e. the (cYMHF) in the caloric gauge $A_s = 0$, with regular initial data $\bar{A}_i \in H^\infty_0$ (in particular having finite magnetic energy). Then there exists a unique global regular solution $A_i \in C^\infty H^\infty_0(\mathbb{R}^3 \times [0, \infty))$.

2. Consider the initial value problem (5.25), i.e. the (dYMHF) in the caloric gauge with regular initial data $\bar{A}_\alpha \in C^\infty H^\infty_0(I \times \mathbb{R}^2)$ where $I$ is a time interval. Then there exists a unique regular solution $A_\alpha \in C^\infty H^\infty_0(I \times \mathbb{R}^3 \times [0, 1])$.

For part (1) see [38, Corollary 6.7] for the statement and proof in the form stated here, of the similar result due to [43] in a different context.

Part (2) is also from [38]. Like in the first part, no smallness assumption is required. We sketch the details of part (2) from [38, Thm 6.10, Lemma 6.9] and [37, Step 1 in Thm 4.8, Prop. 5.7, Lemma 6.1] for the sake of completeness:

The spatial components $A_i$ satisfy (5.24) and are obtained from part (1), so it remains to determine $A_0$. This can be done from the equation $\partial_s A_0 = D^\ell F_{\ell 0}$ if one can determine what $F_{\ell 0}$ should be, knowing that they have to solve the linear system with smooth coefficients

\[ (\partial_s - \Delta_A)B_\ell = 2[F_{\ell j}, B_j], \quad B_\ell(0) = \bar{F}_{\ell 0} \]

Proposition 5.7 in [37] provides a unique regular solution $B$, which is used to define regular $A_0$ from the initial data $\bar{A}_0$ by integrating $\partial_s A_0 = D^\ell B_\ell$. Then one obtains curvatures $F_{\ell 0} = \partial_\ell A_0 - \partial_\ell A_\ell + [A_\ell, A_0]$ which turn out to be equal to $B_\ell$, as guaranteed by [37, Lemma 6.1]. Therefore, the solution satisfies (5.25). Uniqueness is proved in [38, Lemma 6.9].

Now we transfer $I^{-1}H^1$ bounds from DeTurck’s gauge to the Caloric gauge under a smallness assumption.

**Proposition 5.10.** Let $A_i$ be the unique regular solution of (5.24) obtained in Theorem 5.9 from an initial data $\bar{A}_i$ assumed to be regular and to satisfy $\|I\bar{A}_i\|_{H^1} \leq \varepsilon$. Then, if $\varepsilon$ is small enough, one has

\[ \|I\bar{A}_i\|_{L^\infty H^1(\mathbb{R}^3 \times [0, 1])} \leq \varepsilon. \]  

Moreover, $A_i$ is equivalent, using a regular gauge transformation, to

\[ A'_i = U^{-1}A_i U - \partial_\ell U^{-1}U, \quad U(s = 0) = 1 \]

the regular solution to (cYMHF) in DeTurck’s gauge $A'_s = \partial^\ell A'_i$ with initial data $\bar{A}_i$ obtained in Proposition 5.7.

The proof follows the strategy of [37, Theorem 5.1] and consists of a version of DeTurck’s trick [12, 15], see [37, Remark 5.9].
Proof. Let \( A' \) be the regular solution to \( (\text{cYMHF}) \) in DeTurck’s gauge \( A'_s = \partial^s A'_t \) with initial data \( \bar{A}_i \) obtained in Proposition 5.1 and which satisfies \( IA'_t \in \varepsilon L^\infty_x H^1(\mathbb{R}^3 \times [0,1]) \). Making the gauge transformation
\[
\bar{A}_i = U A'_t U^{-1} - \partial_i U U^{-1}, \quad \bar{A}_s = U A'_t U^{-1} - \partial_s U U^{-1}
\]
on \( \mathbb{R}^3 \times [0,1] \) one obtains a caloric solution \( \bar{A}_i \), \( \bar{A}_s = 0 \) with \( \bar{A}_i(s = 0) = \bar{A}_i \) provided \( U(s) \) solves the following ODE pointwise in \( x \)
\[
\partial_i U = U A'_t, \quad U(s = 0) = 1
\]
Since \( A'_s \) is regular, also \( U \), and therefore \( \bar{A}_i \) are. From the uniqueness part in Theorem 5.9 we deduce that \( \bar{A}_i = A_i \).

By Lemma 5.8 \( U \) and \( U^{-1} \) satisfy the bounds (5.20), and therefore \( U - 1, U^{-1} - 1 \in \varepsilon L^\infty_x X \) where \( X \) is defined in (2.23). Using (5.28), \( IA'_t \in \varepsilon L^\infty_x H^1 \) and the product estimate (2.25) one obtains (5.20).

Remark 5.11. Suppose \( A_\alpha(t,x,s) \) is the the unique regular solution of (5.25) obtained in Theorem 5.9 Part(2) on an interval \([t_0,t_1] \times \mathbb{R}^3 \times [0,1] \) from regular and temporal initial data \( A_\alpha \in C^\infty_t H^{\infty}_x ([t_0,t_1] \times \mathbb{R}^3), \) \( \bar{A}_0 = 0 \) satisfying \( \|IA'_t\|_{L^\infty_x H^1} \leq \varepsilon \). Then we may apply Proposition 5.10 for each \( t \in [t_0,t_1] \) obtaining a regular \( U(t,x,s) \) with \( U(t,x,s = 0) = 1 \) and (5.26) holds uniformly in \( t \). In addition to (5.27) one has
\[
A'_\alpha = U^{-1} A_\alpha U - \partial_\alpha U^{-1} U, \quad F'_{\alpha\beta} = U^{-1} F_{\alpha\beta} U, \quad U(t,s = 0) = 1
\]
on \([t_0,t_1] \times \mathbb{R}^3 \times [0,1] \) where \( A'_\alpha \) is the solution to \( (\text{cYMHF}) \) in DeTurck’s gauge, with (temporal) initial data \( A_\alpha \).

Finally, one has improved well-posedness in Sobolev spaces conditional on the smallness of the (almost conserved) modified energy.

Proposition 5.12. Suppose \( A_i \) is the (caloric) regular solution of the problem (5.24) from Theorem 5.9 (1) for regular initial data \( A_i \) with \( \|A_i\|_{H^s} \leq M_0 \) and
\[
\sup_{0 \leq m \leq 3} \sup_{s \in [0,s_0]} (N s^4)^{-1-s} (s^2 D_x)^m F(s) \|L^2(\mathbb{R}^3) \| \leq 1
\]
Then one has \( \|A_i(s)\|_{H^s} \leq C(M_0) \) for all \( s \in [0,s_0], s_0 = N^{-2}. \)

Proof. First note that we have local wellposedness for the equation in DeTurck’s gauge (5.1). This means a (regular) solution \( \bar{A}_i(s) \) to (5.1) exists on \([0,s^*] \) with \( \bar{A}_i(0) = \bar{A}_i, \|\bar{A}_i(s)\|_{H^s} \leq M_0 \) for all \( s \in [0,s^*] \) with \( s^* \sim M_0^{-p_\sigma} \) where \( p_\sigma = \frac{2}{\sigma - 1/2}. \) This is proved by the argument in Proposition 5.1 and Remark 5.4, one just replaces (5.15) by their homogeneous version and one uses the smallness of the interval instead of smallness of the initial data.

We next consider the ODE (5.19) for \( \bar{A}_i \), i.e. \( \partial_s U = U \bar{A}_s, \) \( U(0) = 1 \) where \( \bar{A}_s = \partial^s \bar{A}_t, \) which defines the gauge transformation from DeTurck’s gauge to the caloric gauge. The \( A_i \) and \( \bar{A}_i \) are related by
\[
A_i = U \bar{A}_i U^{-1} - \partial_i U U^{-1}, \quad A_s = U \bar{A}_s U^{-1} - \partial_s U U^{-1} = 0
\]
Now (5.20) is replaced by: \( U \in C(M_0) L^\infty_x L^{\infty}_x \) and \( \nabla_x U \in C(M_0) L^\infty_x H^s \) on \([0,s^*], \) and the same for \( U^{-1}. \) We denote by \( C(M_0) \) a constant that can change from line to line. By (2.24) one obtains \( \|A_i(s)\|_{H^s} \leq C(M_0) \) for \( s \in [0,s^*] \).
It remains to extend this bound to \([0, s_0]\) using \((5.30)\). Let \(p \in [4, 6]\) be the Sobolev exponent such that \(H^p \subset L^p\). From \((5.24)\) and the covariant Gagliardo-Nirenberg inequality \((2.11)\) one has
\[
\| \partial_s A_i(s) \|_{L^p} \lesssim s^{-1} (Ns^2)^{\sigma - 1}.
\]
Integrating one has
\[
\| A_i(s) \|_{L^p} \lesssim \| A_i(s^*) \|_{H^\sigma} + (Ns^2)^{\frac{\sigma - 1}{2}} \lesssim C(M_0), \quad s \in [s^*, s_0].
\]
Now we let \(1/q = 1/2 - 1/p\), \(q \in [3, 4]\) and bound
\[
\| \partial_s A_i(s) \|_{H^1} \lesssim \| D_x F(s) \|_{L^2} + \| D_x^2 F(s) \|_{L^2} + \| A_i(s) \|_{L^p} \| D_x F(s) \|_{L^q}
\]
\[
\lesssim s^{-1} (Ns^{-2})^{\sigma - 1} + C(M_0) s^{-1} (Ns^{-2})^{\sigma - 1}
\]
Integrating again one has
\[
\| A_i(s) \|_{H^\sigma} \leq \| A_i(s^*) \|_{H^\sigma} + \int_{s^*}^s \| \partial_s A_i(s') \|_{H^1} \, ds' \lesssim C(M_0) (Ns^2)^{\frac{\sigma - 1}{2}} \lesssim C(M_0)
\]
for all \(s \in [s^*, s_0]\). \(\square\)

6. Energy estimates for the curvature and the tension field

The goal of this section is to obtain weighted bounds for higher covariant derivatives for the curvature \(F_{\alpha \beta}(s)\) and the tension field \(w_x(s)\), first globally in space and then locally - with square summability of the local bounds, proving Proposition \((1.9)\). All estimates in this section are covariant, independent of the choice of gauge.

Let \(A_{t,x,s}\) be a regular solution to \((dYMHF)\) on an interval \(I \times \mathbb{R}^3 \times [0, s_0]\). Then the curvature satisfies
\[
\left( D_s - D^\ell D_\ell \right) F_{\alpha \beta} = 2 \left[ F_{\alpha \ell}, F_{\ell \beta} \right]
\]
\((6.1)\)

This is obtained from Bianchi’s identity \((2.3)\), commuting the covariant derivatives by \((2.4)\) and applying Bianchi’s identity again.

For higher covariant derivatives one computes (see \((38)\) Section 5) for details:
\[
\left( D_s - D^\ell D_\ell \right) D_x^{(m)} F = \sum_{i=0}^m O \left( D_x^{(i)} F, D_x^{(m-i)} F \right)
\]
\((6.2)\)

Now let \(w_k(s)\) be the Yang-Mills tension field defined in \((1.9)\). It satisfies
\[
\left( D_s - D^\ell D_\ell \right) w_k = 2 \left[ F^\ell_k, w_\ell \right] + 2 \left[ F^{0 \ell}, D_\ell F_{0 \ell} + 2 D_\ell F_{k 0} \right]
\]
\((6.3)\)

See \((39)\) Eq. \((8.8)\), \((37)\) Appendix A. For higher covariant derivatives one has
\[
\left( D_s - D^\ell D_\ell \right) D_x^{(m)} w_k = \sum_{i=0}^m O \left( D_x^{(i)} F, D_x^{(m-i)} w_x \right) + O \left( D_x^{(i)} F, D_x^{(m+1-i)} F \right)
\]
\((6.4)\)

6.1. Covariant energy estimates and the comparison principle.

Suppose \(G\) solves the covariant heat equation
\[
\left( D_s - D^\ell D_\ell \right) G = N
\]
$^{30}$
Lemma 6.2

and the Duhamel principle, \[38, \text{Corollary 4.7}\] formulates the comparison principle:

The following inequality holds point-wise Lemma 6.1.

or, in short, denoting \( J = [s_1, s_2] \),

\[
\| s^p G \|_{L^\infty_2 L^2_x(J)} + \| s^p x^2 \|_{L^2_2 L^2_x(J)} \lesssim s^p \| G(s_1) \|_{L^2_x} + p \| s^p G \|_{L^2_2 L^2_x(J)} + \| s^p \mathcal{N} \|_{L^1_2 L^2_x(J)}
\]

This is obtained by applying the Fundamental theorem of calculus on an interval to the function \( \varphi(s) = \frac{1}{2} s^p \int_{\mathbb{R}^3} (G, G)(s) \, dx \), using the equation, integrating by parts, applying Cauchy-Schwarz and absorbing a term into the LHS. See \[38, \text{Lemma 4.5}\] for more details.

As a consequence of the diamagnetic-type inequality (\[38, \text{Lemma 4.6}\], \[54, \text{Lemma 3.9}\])

\((\partial_s - \Delta) |G| \leq |\mathcal{N}|\)

and the Duhamel principle, \[38, \text{Corollary 4.7}\] formulates the comparison principle:

**Lemma 6.1.** The following inequality holds point-wise

\[
|G(s)| \leq e^{s\Delta} |G(0)| + \int_0^s e^{(s-s')\Delta} \left| (D_s - D^\ell) G(s') \right| \, ds'.
\]

We will also use the following estimate for Duhamel terms

**Lemma 6.2 (\[38\]).** For \( p < \frac{3}{4} \) and \( s_0 \leq 1 \) the following estimate holds on \( [0, s_0] \times \mathbb{R}^3 \):

\[
\| s^p \int_0^{s_1} e^{(s_1-s)\Delta} \mathcal{N}(s) \, ds \|_{L^2_2 L^2_x} \lesssim \| s^{p+\frac{1}{2}} \mathcal{N} \|_{L^2_2 L^2_x}
\]

See \[38, \text{Lemma 4.8}\].

**6.2. Higher covariant derivatives bounds for \( F \) and \( w \).**

Denote

\[
\alpha^m(s') := \| (Ns_x^{1/2})^{1-\sigma} (s_x^{1/2})^{m} F_{\alpha \beta}(s) \|_{L^\infty_2 L^2_x([0, s'] \times \mathbb{R}^3)}
\]

\[
\beta^m(s') := \| (Ns_x^{1/2})^{1-\sigma} (s_x^{1/2})^{m} F_{\alpha \beta}(s) \|_{L^2_2 L^2_x([0, s'] \times \mathbb{R}^3)}
\]

**Proposition 6.3.** Let \( A_{t,x,s} \) be a regular solution to (\DMNHF) on an interval \( J \times \mathbb{R}^3 \times [0, s_0] \). Fix a time \( t \in J \) and denote \( F(s) = F(t, s) \). Let \( \eta \ll 1 \) and suppose

\[
\alpha^0(s_0) + \beta^0(s_0) \lesssim \eta
\]

Then, for \( N^2 s_0 = 1 \) and for all \( 0 \leq m \leq 10 \) one has

\[
\alpha^m(s_0) + \beta^{m+1}(s_0) \lesssim m \eta
\]

**Proof.** Denote by \( \mathcal{N}^m \) the RHS of (6.2) and let

\[
\gamma^m(s') := \| (Ns_x^{1/2})^{1-\sigma} (s_x^{1/2})^{m} \mathcal{N}(s) \|_{L^2_2 L^2_x([0, s'])}
\]

Applying (6.5) to (6.2) with \( p = \frac{m}{2} + \frac{1-\sigma}{2} \) on \([0, s]\) and multiplying by \( N^{1-\sigma} \) one obtains

\[
\alpha^m(s) + \beta^{m+1}(s) \lesssim \beta^m(s) + \gamma^m(s)
\]
First let $m = 0$. Using Holder’s inequality and \((2.11)\) one has
\[
\| [F,F] (s) \|_{L^2_s} \lesssim \| F(s) \|_{L^2_s}^{\frac{1}{2}} \| D_x F(s) \|_{L^2_s}^{\frac{3}{2}}
\]
which implies that one can bound
\[
\beta^1(s) \lesssim \eta + \gamma^1(s) \lesssim \eta + \eta^{\frac{1}{2}} s^{\frac{1}{4}} (\beta^1(s))^{\frac{3}{2}} (N s^2)^{1-\sigma}
\]
Since \(\lim_{s \to 0} \beta^1(s) = 0\) one can use a continuity argument to obtain \(\beta^1(s_0) \lesssim \eta\). Then one proceeds by induction on \(m\), similarly bounding
\[
\alpha^m(s) + \beta^{m+1}(s) \lesssim \beta^m(s) + \gamma^m(s) \lesssim \eta + \eta^{\frac{1}{2}} s^{\frac{1}{4}} (\beta^m(s))^{\frac{3}{2}} (N s^2)^{1-\sigma}
\]
and concluding like before. \(\square\)

Now we turn to the Yang-Mills tension field \(w_x\).

**Proposition 6.4.** Given \((\text{YM})\) solutions \(A_{t,x,s}\) and \(F_{\alpha \beta}\) satisfying the assumptions from Proposition 6.3 for \(w_x(s) := D^\alpha F_{\alpha \beta}(s)\) and \(0 \leq m \leq 8\) one has
\[
\| (N^2 s)^{1-\sigma} s^{\frac{1}{4}} (s^{\frac{1}{2}} D_x)^m w_x(s) \|_{L^\infty_t L^2_x \cap L^2_t L^3_x ([0,s_0] \times \mathbb{R}^3)} \lesssim \eta^2
\]
Moreover,
\[
(N^2 s)^{1-\sigma} \| w_x(s) \|_{L^1_x} \lesssim s^{\frac{1}{2}} \eta \quad \forall s \in [0,s_0].
\]
**Proof.** Denote by \(\delta_m\) and \(\rho_m\) the \(L^\infty_t L^2_x\), respectively the \(L^2_t L^3_x\) norms from \((6.6)\). Note that \(D_x^m w(s = 0) = 0\) because \(F\) is assumed to solve the Yang-Mills equation at \(s = 0\).

We begin by showing \(\rho_0 \lesssim \eta^2\).

We use equation \((6.3)\), the fact that \(w_x(s = 0) = 0\), the Comparison principle in Lemma 6.1, Lemma 6.2 with \(p = \frac{1}{4} + 1 - \sigma\) and the following estimates
\[
\| (N^2 s)^{1-\sigma} s^{\frac{1}{4}} [F,w_x] \|_{L^2_t L^1_x} \lesssim \| (N s^2)^{\frac{1}{4}} F \|_{L^\infty_t L^2_x} \| s^{\frac{1}{4}} (N s^2)^{1-\sigma-1} \|_{L^\infty_t} \rho_0 \lesssim \eta \delta_0^\frac{1}{4} \rho_0
\]
\[
\| (N^2 s)^{1-\sigma} s^{\frac{1}{4}} [F,D_x F] \|_{L^2_t L^1_x} \lesssim \| (N s^2)^{\frac{1}{4}} F \|_{L^\infty_t L^2_x} \| (N s^2)^{1-\sigma} s^{\frac{1}{4}} D_x F \|_{L^2_t L^3_x} \lesssim \eta^2
\]
to obtain
\[
\rho_0 \lesssim \eta \delta_0^\frac{1}{4} \rho_0 + \eta^2.
\]
and then we absorb the small \(\rho_0\) term to the left side since \(\eta \ll 1\).

The same estimate for \([F,D_x F]\), but with a different factor of \(s\) and bounding in \(L^1_t\), together with the Comparison principle in Lemma 6.1 will prove \((6.7)\). Denoting \(M_1 := \| w_x \|_{L^\infty_t L^1_x [0,s_1]}\), now one absorbs the following term to the LHS:
\[
\| [F,w_x] \|_{L^1_t L^1_x} \lesssim M_1 \int_0^{s_1} \| F(s) \|_{L^\infty_x} \| w_x \|_{L^1_x} ds \lesssim M_1 s_1^{\frac{1}{4}} / (N^2 s)^{1-\sigma} \eta \ll M_1.
\]

Now we proceed by induction. We assume we have showed \(\rho_i \lesssim \eta^2\) for all \(0 \leq i \leq m\) and we prove
\[
\delta_m + \rho_{m+1} \lesssim \rho_m + \| (N^2 s)^{1-\sigma} s^{\frac{1}{4}} N_m(s) \|_{L^1_t L^2_x} \lesssim \eta^2
\]
where \(N_m(s)\) is the RHS of \((6.4)\). The first inequality is the energy estimate \((6.5)\). We will use
\[
\| D_x^i F \|_{L^\infty_x} \lesssim \| D_x^{i+1} F \|_{L^1_x}^{1/2} \| D_x^{i+2} F \|_{L^2_x}^{1/2}
\]
which follows from \( \text{(2.11)} \). We bound the first component of \( \mathcal{N}_m(s) \)
\[
\|(N^2 s)^{1-\sigma} s^{\frac{1}{2}} s |
\| \mathbf{D}_x F \|_{L^\infty} \| \mathbf{D}^{m-1} w_x \|_{L^1_s} \lesssim \eta \| s^{\frac{1}{2}} (N s^{\frac{1}{2}})^{-\sigma} \|_{L^\infty_x} \rho \lesssim \eta^3 s_0^{\frac{1}{4}}
\]
and for the second component we may assume that \( i \leq m + 1 - i \) and estimate
\[
\|(N^2 s)^{1-\sigma} s^{\frac{1}{2}} s \| \| \mathbf{D}_x F \|_{L^\infty} \| \mathbf{D}^{m+1} w_x \|_{L^1_s} \lesssim \|(N s^{\frac{1}{2}})^{1-\sigma} (s^{\frac{1}{2}} \mathbf{D}_x)^{i+1} F \|_{L^2_{s_k} L^2_x} \times
\]
\[
\|(N s^{\frac{1}{2}})^{1-\sigma} (s^{\frac{1}{2}} \mathbf{D}_x)^{m+1} F \|_{L^2_{s_k} L^2_x} \lesssim \eta^2
\]
\[\square\]

6.3. Defining the balls coefficients - Proof of Proposition \( \text{1.9} \)

Fix a time \( t \in J \) and denote \( F(s) = F(t,s) \). It is easily seen that the proof is uniform in \( t \). If one would only have \( L^2 \)-type norms in \( \text{(1.10)}, \text{(1.11)} \) such as \( L^2_{s_k} L^2_x \) or \( L^2_x \) instead of \( L^\infty_x L^2_x \) the square summability would be obvious, because by Proposition \( \text{6.3} \) we control all higher derivatives of \( F \) on \( [0, s_0] \times \mathbb{R}^3 \):
\[
\|(N s^{\frac{1}{2}})^{1-\sigma} (s^{\frac{1}{2}} \mathbf{D}_x)^m F(s) \|_{L^\infty_x L^2([0,s_0] \times \mathbb{R}^3)} + \|(N s^{\frac{1}{2}})^{1-\sigma} (s^{\frac{1}{2}} \mathbf{D}_x)^{m+1} F(s) \|_{L^2_{s_k} L^2([0,s_0] \times \mathbb{R}^3)} \lesssim \eta
\]
and similarly for \( w \), due to Proposition \( \text{6.3} \) we have
\[
\|(N^2 s)^{1-\sigma} s^{\frac{1}{2}} (s^{\frac{1}{2}} \mathbf{D}_x)^m w_x \|_{L^\infty_x L^2([0,s] \times \mathbb{R}^3)} \lesssim \eta^2
\]
To deal with the \( L^\infty_x L^2_x \) we use the Energy estimates from section \( \text{6.1} \).

Consider a collection of spatial bump functions \( (\chi_j)_{j \in J} \) with finitely overlapping supports which satisfy \( 1 \lesssim \sum_{j \in J} \chi^2_j(x) \). We construct square summable coefficients \( (c_j)_{j \in J} \) satisfying
\[
\alpha^m_j := \| \chi_j (N s^{\frac{1}{2}})^{1-\sigma} (s^{\frac{1}{2}} \mathbf{D}_x)^m F(s) \|_{L^\infty_x L^2([0,s] \times \mathbb{R}^3)} \leq c_j
\]
\[
\beta^m_j := \| \chi_j (N s^{\frac{1}{2}})^{1-\sigma} (s^{\frac{1}{2}} \mathbf{D}_x)^m F(s) \|_{L^2_{s_k} L^2([0,s] \times \mathbb{R}^3)} \leq c_j
\]
and
\[
\delta^m_j := \| \chi_j (N^2 s)^{1-\sigma} s^{\frac{1}{2}} (s^{\frac{1}{2}} \mathbf{D}_x)^m w_x \|_{L^\infty_x L^2([0,s] \times \mathbb{R}^3)} \leq c_j
\]
\[
\rho^m_j := \| \chi_j (N^2 s)^{1-\sigma} s^{\frac{1}{2}} (s^{\frac{1}{2}} \mathbf{D}_x)^m w_x \|_{L^2_{s_k} L^2([0,s] \times \mathbb{R}^3)} \leq c_j
\]
for all \( 0 \leq m \leq 8 \) and then we take each \( \chi_j \equiv 1 \) on \( B_j \) to obtain the conclusion.

It suffices to show
\[
\sum_{j \in J} (\alpha^m_j)^2 \lesssim \eta^2, \quad \sum_{j \in J} (\beta^m_j)^2 \lesssim \eta^2, \quad \sum_{j \in J} (\delta^m_j)^2 \lesssim \eta^2, \quad \sum_{j \in J} (\rho^m_j)^2 \lesssim \eta^2.
\]
Note that this is automatically true for the \( \beta^m_j \)'s and \( \rho^m_j \)'s so it remains to focus on the \( \alpha^m_j \)'s and \( \delta^m_j \)'s.

From \( \text{(6.1)} \) we obtain the parabolic equation of \( \chi_j F \)
\[
(\mathbf{D}_s - \mathbf{D}^\tau \mathbf{D}_\ell) \chi_j F = \mathcal{N}_j := 2 \chi_j [F,F] - \Delta \chi_j F - 2 \partial \ell \chi_j \mathbf{D}^\tau F
\]
while from \( \text{(6.3)} \) we obtain the equation of \( \chi_j w_x \)
\[
(\mathbf{D}_s - \mathbf{D}^\tau \mathbf{D}_\ell) \chi_j w_x = \mathcal{N}_j := 2 \chi_j [F,w_x] + 2 \chi_j [F,\mathbf{D}_x F] - \Delta \chi_j w_x - 2 \partial \ell \chi_j \mathbf{D}^\tau w_x
\]
By \((6.5)\) on \([0, s_0]\), \((2.11)\), \((6.9)\) and Holder in \(\frac{ds}{s}\) we obtain
\[
\alpha_j^0 \lesssim \beta_j^0 + \| (Ns^\frac{1}{2})^{1-\sigma} N_j(s) \|_{L^1_t L^2_x [0, s_0]} \\
\lesssim \beta_j^0 + \| (Ns^\frac{1}{2})^{1-\sigma} (\| DF \|_{L^2_x}^2 \| D^2 F \|_{L^2_x}^2) \|_{L^1_t L^2_x [0, s_0]} \\
\lesssim \beta_j^0 + \eta \beta_j^0 \| s^\frac{1}{2} (Ns^\frac{1}{2})^{\sigma-1} \|_{L^2_x [0, s_0]} + \beta_j^0 \| s \|_{L^{2,\infty}_x [0, s_0]} + \beta_j^0 \| s^\frac{1}{2} \|_{L^{2,\infty}_x [0, s_0]} \\
\lesssim \beta_j^0 + \beta_j^0 + \beta_j^1
\]
where \(\bar{\beta}_j^m\) and \(\tilde{\beta}_j^m\) (respectively \(\bar{\rho}_j^m\) and \(\tilde{\rho}_j^m\)) are defined to be the terms in \((6.12)\) (respectively \((6.14)\)) with \(\chi_j\) replaced by \(\Delta \chi_j\) and \(\nabla \chi_j\), which are also square summable, thus proving \((6.15)\) for the \(\alpha_j^0\)'s. Similarly
\[
\delta_j^0 \lesssim \rho_j^0 + \| (N^2 s)^{1-\sigma} s^\frac{1}{2} N_j'(s) \|_{L^1_t L^2_x [0, s_0]} \\
\lesssim \rho_j^0 + \| (N^2 s)^{1-\sigma} s^\frac{1}{2} (\| DF \|_{L^2_x}^2 \| D^2 F \|_{L^2_x}^2) \|_{L^1_t L^2_x [0, s_0]} \\
+ \| \Delta \chi_j w \|_{L^2_x}^2 + \| \nabla \chi_j D w \|_{L^2_x [0, s_0]} \\
\lesssim \rho_j^0 + \eta \beta_j^0 \| s^\frac{1}{2} (Ns^\frac{1}{2})^{\sigma-1} \|_{L^2_x [0, s_0]} + \eta \beta_j^1 + \tilde{\rho}_j^0 \| s \|_{L^{2,\infty}_x [0, s_0]} + \tilde{\rho}_j^1 \| s^\frac{1}{2} \|_{L^{2,\infty}_x [0, s_0]} \\
\lesssim \rho_j^0 + \beta_j^0 + \rho_j^1 + \tilde{\rho}_j^1
\]
This proves \((6.15)\) for \(m = 0\).

The same argument holds for \(m \geq 1\). By \((6.2)\) the equations for higher derivatives of \(F\) are
\[
(D_s - D^i D_i) (\chi_j D_{(m)}^x F) = N_j^m \]
Applying \((6.5)\) with \(p = \frac{m}{2} + \frac{1-\sigma}{2}\), multiplying by \(N^{1-\sigma}\) and doing the same manipulations one obtains
\[
\alpha_j^m \lesssim \beta_j^m + \eta \sum_{k=0}^m \beta_j^k + \beta_j^m + \rho_j^{m+1}
\]
which is square-summable. Similarly one writes the equation of higher covariant derivatives of \(\chi_j w_x\) following from \((6.4)\) and estimate
\[
\delta_j^m \lesssim \rho_j^m + \eta \sum_{k=0}^m \rho_j^k + \eta \sum_{k=1}^{m+1} \beta_j^k + \rho_j^m + \tilde{\rho}_j^{m+1}
\]
which concludes the proof.

7. The Local Change of Gauge: Proof of Proposition 1.10

This section is devoted to the proof of Proposition 1.10, which is an Uhlenbeck-type lemma for fractional regularities below \(L^2\), as mentioned in Remark 1.11. Unlike the previous section, which consisted of only gauge-invariant bounds, the goal here is the opposite: to be as specific as possible with the choice of gauge in order to obtain Sobolev bounds for \(A_1, \partial_t A_1\) from the curvature bounds.

We begin by recalling the classical result of Uhlenbeck [56], which played a crucial role in the proof of finite energy global well-posedness [27] stated in Theorem 1.1.
Theorem 7.1 (Uhlenbeck’s lemma [56]). Let \( A_x : B \to g \) have curvature \( F_{ij} \) bounded in \( L^2(B) \). There exists a \( \delta > 0 \) such that if

\[
\|F\|_{L^2(B)} < \delta
\]

then there exists \( U : B \to G \) such that the transformation

\[
\tilde{A}_i = UA_iU^{-1} - \partial_iUU^{-1}
\]

satisfies \( \partial^\ell \tilde{A}_\ell = 0 \) in \( B \), \( x^\ell \tilde{A}_\ell = 0 \) on \( \partial B \) and

\[
\|\tilde{A}\|_{H^1(B)} \lesssim \|F\|_{L^2(B)}.
\]

After smoothing the curvature \( F_{\alpha\beta} \) using the Yang-Mills heat flow, we will apply Uhlenbeck’s lemma at \( s = s_0 \) to first obtain an \( H^1 \) bound.

By elliptic regularity, one obtains additional bound on \( \tilde{A} \):

Lemma 7.2. Suppose \( \tilde{A}_i \) is like in Theorem 7.1 and assume the (gauge-invariant) bounds

\[
\|(s_0^\frac{1}{2} \tilde{D}_x)^m \tilde{F}\|_{L^2(B)} \leq \delta
\]

for \( 0 \leq m \leq 2 \) and some \( s_0 \ll 1 \). Then

\[
s_0^\frac{3}{4} \|\tilde{A}_i\|_{L^\infty(B)} + s_0^\frac{3}{4} \|\partial_x \tilde{A}_i\|_{L^\infty(B)} \lesssim \delta.
\]

Proof. First note that \( s_0^\frac{1}{2} \partial_x \tilde{F} \in \delta L^2(B) \). Indeed, by (2.12) one has \( s_0^\frac{1}{2} \tilde{F} \in \delta L^6(B) \) and then \( s_0^\frac{1}{2} [\tilde{A}, \tilde{F}] \in \delta L^2(B) \).

From \( \partial^\ell \tilde{A}_\ell = 0 \) one obtains the elliptic equation

\[
\Delta \tilde{A}_j = \partial^i \tilde{F}_{ij} - \partial^\ell [\tilde{A}_i, \tilde{A}_j].
\]

We want to show \( s_0^\frac{1}{2} [\tilde{A}, \nabla \tilde{A}] \in \delta L^2(B) \). We begin with some weaker bounds, like \( [\tilde{A}, \nabla \tilde{A}] \in \delta L^\frac{3}{2}(B) \) which follows from Hölder and Sobolev embedding. Then \( s_0^\frac{1}{2} \tilde{A}_j \in \delta W^{2,\frac{3}{2}}(B) \subset \delta L^p(B) \) for all \( p < \infty \). This implies \( s_0^\frac{1}{2} [\tilde{A}, \nabla \tilde{A}] \in \delta L^{2-}(B) \) and \( s_0^\frac{1}{2} \Delta \tilde{A} \in \delta L^{2-}(B) \) from which, together with \( \tilde{A} \in \delta W^{1,2}(B) \subset \delta L^6(B) \) and Gagliardo-Nirenberg inequalities one obtains \( s_0^\frac{1}{2} \Delta \tilde{A} \in \delta L^\infty(B) \).

Now we obtain \( s_0^\frac{1}{2} [\tilde{A}, \nabla \tilde{A}] \in \delta L^2(B) \) from which \( s_0^\frac{1}{2} \Delta \tilde{A} \in \delta L^2(B) \) and similarly conclude \( s_0^\frac{1}{2} \tilde{A} \in \delta L^\infty(B) \).

For the second inequality note that \( s_0^2 \tilde{F} \in \delta L^2(B) \). Differentiating (7.2) one obtains an elliptic equation for \( \partial_x \tilde{A}_j \). Then we conclude by a similar argument as before, we omit the details. \( \square \)

7.1. Proof of Proposition 1.10 Part (1).

Step 1. (Change of gauge) By integrating the equation \( \partial_x A_i = D^\ell F_{\ell i} \) one obtains

\[
A_i(s) = \int_s^{s_0} D^\ell F_{\ell i}(s') \, ds' + A_i(s_0), \quad s \in [0, s_0].
\]

Since \( N s_0^{\frac{1}{2}} = 1 \) we have \( \|F(s_0)\|_{L^2(B)} \leq \delta \) and we can apply Uhlenbeck’s lemma (Theorem 7.1) at \( s = s_0 \) to obtain \( U : B \to G \) such that after the transformation (7.1) one has \( \|\tilde{A}_i(s_0)\|_{H^1(B)} \lesssim \delta \).

Write

\[
\tilde{A}_i(s) = \int_s^{s_0} UD^\ell F_{\ell i} U^{-1} \, ds' + UA_i(s_0)U^{-1} - \partial_i UU^{-1} = \int_s^{s_0} \tilde{D}^\ell \tilde{F}_{\ell i}(s') \, ds' + \tilde{A}_i(s_0)
\]

\[35\]
For $\tilde{A}_i(s_0)$ we already have an $H^1(B)$ bound so it remains to estimate the integral in $H^\sigma(B)$, and specifically in $H^\sigma(B)$, since integrating in $L^2(B)$ is straightforward.

**Step 2.** (Replace covariant derivatives by ordinary derivatives)

From Lemma [7.2] applied to $\tilde{A}_i(s_0)$ we obtain $s_0^{\frac{3}{2}} \tilde{A}_i(s_0), s_0^{\frac{5}{2}} \partial_x \tilde{A}_i(s_0) \in \delta L^\infty(B)$.

From this and from $(Ns^\frac{1}{2})^{-\sigma} s^\frac{1}{2} \partial_x \tilde{F}_i(t') \in \delta L^\infty(B)$ - which follows from (2.13) - one obtains

$$\tag{7.3} (Ns^\frac{1}{2})^{-\sigma} s^\frac{1}{2} \partial_x \tilde{A}_i(s) \in \delta L^\infty(B).$$

for $m = 0$ and $s \in [0, s_0]$. This allows us to obtain

$$\tag{7.4} (Ns^\frac{1}{2})^{-\sigma} (s^\frac{1}{2} \partial_x)^m \tilde{F}_i(s) \|_{L^2(B)} \lesssim \delta, \quad s \in [0, s_0]$$

for $m = 1$ and also (7.3) for $m = 1$ which in turn implies (7.4) also for $m = 2$.

**Step 3.** (Estimate the integral term) We first consider

$$\tilde{A}_i^{\text{main}} := \int_0^s \partial_x \tilde{F}_i(s') ds'$$

Consider an extension $\bar{F}$ of $\tilde{F}$ from $B \times [0, s_0]$ to $\mathbb{R}^3 \times [0, s_0]$ such that

$$\tag{7.5} (Ns^\frac{1}{2})^{-\sigma} (s^\frac{1}{2} \partial_x)^m \bar{F}(s) \|_{L^2(\mathbb{R}^3)} \lesssim \delta, \quad 0 \leq m \leq 2.$$ 

for $s \in [0, s_0]$. This implies

$$(Ns^\frac{1}{2})^{-\sigma} s^\frac{1}{2} \|P_k \partial_x \bar{F}(s)\|_{H^\sigma} \lesssim \frac{2^{k\sigma}}{\langle s^\frac{1}{2} 2^k \rangle}$$

for some $\|(\alpha_k)\|_2 \lesssim \delta$. Indeed, use (7.5) with $m = 1$ for low frequencies with $s 2^k < 1$ and use (7.3) with $m = 2$ for high frequencies such that $s 2^k > 1$. Now

$$N^{2(1-\sigma)} \|\tilde{A}_i^{\text{main}}\|_{H^\sigma(B)}^2 \lesssim N^{2(1-\sigma)} \sum_k \int_0^s \int_0^s \|\partial_x \bar{F}(s_1), \partial_x \bar{F}(s_2)\|_{H^\sigma} ds_1 ds_2$$

$$\lesssim \sum_k \int_0^s \int_0^s \frac{(s_1 2^k)\sigma}{\langle s_1^\frac{3}{2} 2^k \rangle} \frac{(s_2 2^k)\sigma}{\langle s_2^\frac{3}{2} 2^k \rangle} ds_1 ds_2$$

$$\lesssim \sum_k \alpha_k^2 \lesssim \delta^2.$$ 

The remaining term

$$\tilde{A}_i^{\text{rem}} := \int_0^s [\bar{F}(s') ds']$$

can in fact be estimated easily in $N^{-\frac{1}{2}} H^1(B)$ placing $\tilde{A}(s'), \partial_x \tilde{A}(s')$ in $L^\infty(B)$ and using (7.3).

7.2. **Proof of Proposition 1.10** Part (2).

**Step 1.** (Second change of gauge)

Since $U = U(x)$ is independent of $t$ and $s$, $\tilde{A}$ remains in the temporal-caloric gauge $\tilde{A}_0(t, x, 0) = 0, \tilde{A}_s(t, x, s) = 0$. In particular

$$\partial_t \tilde{A}_i(t_0, 0) = \int_0^{s_0} \partial_t \bar{F}_i(t_0, s') ds' + \partial_t \tilde{A}_i(t_0, s_0).$$

We want to use the bounds

$$\tag{7.7} (Ns^\frac{1}{2})^{-\sigma} (s^\frac{1}{2} \bar{D}_x)^m \bar{F}_{i\beta}(t_0, s) \|_{L^2(B)} \lesssim \delta$$
to place \( \partial_t \tilde{A}_i(t_0, s_0) \) in \( L^2(B) \), but \( \tilde{A}_\alpha(\cdot, s_0) \) is not in the temporal gauge. So we change the gauge to make it so, letting

\[
\tilde{A}_\alpha = V \tilde{A}_\alpha V^{-1} - \partial_\alpha V V^{-1}, \quad \tilde{F}_{\alpha\beta} = V \tilde{F}_{\alpha\beta} V^{-1}
\]

for \( V = V(t, x) \) which makes \( \tilde{A}_\alpha(\cdot, s_0) \) temporal i.e. by solving the ODE

\[
\partial_t V = V \tilde{A}_0(\cdot, s_0), \quad V(t_0) = Id.
\]

Now we have

\[
\tilde{A}_0(t, s_0) = 0, \quad \tilde{A}_s = 0 \quad (7.8)
\]

and, since \( V(t_0) = Id \),

\[
\tilde{A}_i(t_0, s) = \tilde{A}_i(t_0, s), \quad \tilde{F}_{\alpha\beta}(t_0, s) = \tilde{F}_{\alpha\beta}(t_0, s) \quad (7.9)
\]

**Step 2.** (Bounds on \( \tilde{A}_0 \)) From \( \tilde{F}_{s0} = \tilde{D}_t \tilde{F}_{t0} \) and (7.8), (7.9) one has

\[
\tilde{A}_0(t_0, s) = \int_s^{s_0} \tilde{D}_t \tilde{F}_{0t}(t_0, s') ds'
\]

From this and (2.13), (7.7) one obtains

\[
(N s^{\frac{1}{2}})^{1-\sigma}s^{\frac{1}{2}} ||\tilde{A}_0(t_0, s)||_{L^\infty(B)} \lesssim \delta. \quad (7.10)
\]

From Step 2 in Part (1) we have (7.3) at \( t = t_0 \) for \( m = 0, 1 \) which allows to pass from (7.7) to

\[
(N s^{\frac{1}{2}})^{1-\sigma}||s^{\frac{1}{2}} ?\partial_x|^m \tilde{F}_{\alpha\beta}(t_0, s)||_{L^2(B)} \lesssim \delta, \quad m \leq 2. \quad (7.11)
\]

We claim

\[
\partial_t \tilde{A}_0(t_0, 0) = \int_0^{s_0} \partial_t \tilde{D}_t \tilde{F}_{0t}(t_0, s') + \partial_t [\tilde{A}_t, \tilde{F}_{0t}](t_0, s') ds' \quad (7.12)
\]

Indeed, the second integral is in \( N^{-\frac{3}{2}} L^2(B) \) using (7.3) and (7.11).

To bound the first integral we proceed like in Step 3 from Part (1). Extend \( \tilde{F} \) by \( \bar{F} \) from \( B \times [0, s] \) to \( \mathbb{R}^3 \times [0, s] \) such that (7.11) holds on \( \mathbb{R}^3 \) for \( \bar{F} \). Then

\[
(N s^{\frac{1}{2}})^{1-\sigma} ||P_k \partial^2 \bar{F}(s)||_{H^{s-1}} \lesssim \frac{s^{\frac{1}{2}+k\sigma}}{(s^{\frac{1}{2}})^k} \alpha_k
\]

for some \( ||(\alpha_k)||_\sigma \lesssim \delta \). Then

\[
N^{2(1-\sigma)} \int_0^{s_0} \partial^2 \tilde{F}(t_0, s') ds' \bigg|_{H^{s-1}(B)} \lesssim
\]

\[
N^{2(1-\sigma)} \int_0^{s_0} \int_0^{s_0} \langle \partial^2 \tilde{F}_k(s_1), \partial^2 \tilde{F}_k(s_2) \rangle_{H^{s-1}} ds_1 ds_2
\]

This is bounded by (7.6) and thus \( \lesssim \delta^2 \).

**Step 3.** (Estimate \( \partial_t \tilde{A}_i \)) Using (7.8) and \( \partial_t \tilde{A}_i(t_0, s_0) = \tilde{F}_{0i}(t_0, s_0) \) write

\[
\partial_t \tilde{A}_i(t_0, 0) = \int_0^{s_0} \tilde{D}_t \tilde{D}_t \tilde{F}_{it}(t_0, s') - [\tilde{A}_0, \tilde{D}_t \tilde{F}_{it}](t_0, s') ds' + \tilde{F}_{0i}(t_0, s_0).
\]
The last term \( \tilde{F}_0(t_0, s_0) \) is in \( \delta L^2(B) \). We rewrite the integral, by commuting and using Bianchi (2.0) and (7.10), as

\[
\int_0^{s_0} \mathbf{D}_x^2 \tilde{F}_{0x} + \left[ \tilde{F}_{0t}, \tilde{F}_{0\ell} \right] - \left[ \tilde{A}_0, \mathbf{D}_t \tilde{F}_{0\ell} \right] \, ds, \quad t = t_0
\]  
(7.14)

We bound the integral of \( \left[ \tilde{F}_{0t}, \tilde{F}_{0\ell} \right] \) and \( \left[ \tilde{A}_0, \mathbf{D}_t \tilde{F}_{0\ell} \right] \) in \( N^{-\frac{1}{2}} L^2(B) \) using (7.7) together with (2.13), and \( (7.10) \) together with (7.7). The same holds for the integral of \( \mathbf{D}_x^2 \tilde{F}_{0x} - \partial_x^2 \tilde{F}_{0x} \). The integral of \( \partial_x^2 \tilde{F}_{0x} \) is estimated exactly by (7.13).

This proves \( \partial_t \tilde{A}_i(t_0, 0) \) is in \( L^2(B) + N^{\sigma-1} H^{\sigma-1}(B) \).

**Step 4.** (Return to \( \tilde{A}_i \)) From the temporal condition \( \tilde{A}_0(t, 0) = 0 \) and (7.9) one has

\[
\partial_t \tilde{A}_i(t_0, 0) = \tilde{F}_{0i}(t_0, 0) = \tilde{F}_0(t_0, 0) - \partial_t \tilde{A}_0(t_0, 0) + \tilde{A}_i \tilde{A}_0 \approx 0
\]

We claim this is in \( L^2(B) + N^{\sigma-1} H^{\sigma-1}(B) \): the term \( \partial_t \tilde{A}_0(t_0, 0) \) was treated in Step 3, the term \( \partial_t \tilde{A}_i(t_0, 0) \) in (7.12) in Step 2, while for \( [\tilde{A}_i, \tilde{A}_0] \) we place both inputs in \( L^2(B) \) by Sobolev embedding.

8. **Local differences of heat flows which coincide on a ball**

The goal of this section is to compare two caloric heat flow solutions which coincide somewhere initially. Due to infinite speed of propagation they will not coincide in the future, but one can obtain quantitative local \( L^2 \) estimates for the difference.

Consider two balls \( B^{(1)} \subset B \) of radius \( \simeq 1 \).

**Proposition 8.1.** Let \( A_{t,x,s}, A_{t,x,s}' : [t_0, t_1] \times B \times [0, s_0] \to \mathfrak{g} \) be two smooth caloric solutions to (LYMHP) which coincide on a smaller initial cylinder \( [t_0, t_1] \times B^{(1)} \times \{0\} \). Let \( \chi \) be a spatial bump function supported in the ball \( B^{(1)} \). Let \( N^2 s_0 = 1 \). We assume, for all \( t \in [t_0, t_1] \) and \( 0 \leq m \leq 3, \varepsilon \ll 1 \), the following bounds:

\[
\left\| (N s^{\frac{1}{2}})^{1-\sigma} (s^{\frac{3}{2}} \mathbf{D}_x)^m F_{\alpha \beta}(t, s) \right\|_{L^\infty_t L^2_x \cap L^2_x [0, [0, s_0] \times B]} \lesssim \varepsilon
\]  
(8.1)

\[
\left\| (N s^{\frac{1}{2}})^{1-\sigma} (s^{\frac{1}{2}} \mathbf{D}_x')^m F_{\alpha \beta}'(t, s) \right\|_{L^\infty_t L^2_x \cap L^2_x [0, [0, s_0] \times B]} \lesssim \varepsilon
\]  
(8.2)

\[
A'_x \in \varepsilon L^\infty_t L^8_x \left[ H^1(B) + N^{\sigma-1} H^\sigma(B) \right]
\]  
(8.3)

and

\[
\left\| (N s^{\frac{1}{2}})^{1-\sigma} (s^{\frac{5}{2}} \mathbf{D}_x)^m w_x(t, s) \right\|_{L^\infty_t L^2_x \cap L^2_x [0, [0, s_0] \times B]} \lesssim \varepsilon
\]  
(8.4)

\[
\left\| (N s^{\frac{1}{2}})^{1-\sigma} (s^{\frac{5}{2}} \mathbf{D}_x')^m w_x'(t, s) \right\|_{L^\infty_t L^2_x \cap L^2_x [0, [0, s_0] \times B]} \lesssim \varepsilon
\]  
(8.5)

Then, for any \( s \in [0, s_0] \), \( t \in [t_0, t_1] \) one has

\[
(N s^{\frac{1}{2}})^{0(1-\sigma)} \| \chi [F_{\alpha \beta}(t, s) - F_{\alpha \beta}'(t, s)] \|_{L^\infty_t L^2_x} \lesssim s^2 \varepsilon
\]  
(8.6)

\[
(N s^{\frac{1}{2}})^{8(1-\sigma)} \| \chi [w_x(t, s) - w_x'(t, s)] \|_{L^\infty_t L^2_x} \lesssim s^{\frac{5}{2}} \varepsilon
\]  
(8.7)

It will be clear from the proof that one can in fact obtain arbitrary powers of \( s \) in the conclusion.

Heuristically, these bounds are clear from the decay of the Gaussian kernels on the spatial scales involved.
Main equations. By caloric solutions we mean that they satisfy $A_s = A'_s = 0$, which implies

$$
\partial_s A_t = F_{st} = D^j F_{j,t}
$$

for each $t \in [t_0, t_1]$. Taking divergence and using the identity $D^j D^j F_{j,t} = 0$ one has

$$
\partial_s \partial^j A_t = - [A^j, F_{st}]
$$

From (6.1) and $A_s = 0$ we get the schematic equation

$$
\left( \partial_s - D^j D_j \right) (\chi F) = 2 \chi [F, F] - \Delta \chi F - 2 \nabla \chi \cdot D_x F
$$

and from (6.2)

$$
\left( \partial_s - D^j D_j \right) (\chi D_x F) = \chi \mathcal{O}(F, D_x F) - \Delta \chi D_x F - 2 \nabla \chi \cdot D_x^2 F
$$

while from (6.3) we obtain

$$
\left( \partial_s - D^j D_j \right) (\chi w) = 2 \chi [F, w] + \chi \mathcal{O}(F, D_x F) - \Delta \chi w - 2 \nabla \chi \cdot D_x w
$$

Preliminary bounds. From the covariant Gagliardo-Nirenberg inequalities (2.12), (2.13) and (8.1), (8.2), for $p \in [2, \infty]$ and $m = 0, 1$ one has

$$
(N s^{\frac{1}{2}})^{1-\sigma} \left( \|D_x^m F(s)\|_{L^p(B)} + \|D_x^m F'(s)\|_{L^p(B)} \right) \lesssim \varepsilon s^{-\frac{3}{4} + \frac{1}{4p}} s^{-\frac{m}{2}}
$$

Similarly from (8.4), (8.5)

$$
(N s^{\frac{1}{2}})^{1-\sigma} \left( \|D_x^m w(s)\|_{L^p(B)} + \|D_x^m w'(s)\|_{L^p(B)} \right) \lesssim \varepsilon s^{-\frac{3}{4} + \frac{1}{4p}} s^{-\frac{m}{2}}
$$

From (8.3) one has

$$
A'_s \in \varepsilon L^\infty_t L^\infty_x \left[ L^6 \left( B \right) + N^{\sigma - 1} L^9 \left( B \right) \right], \quad \frac{1}{2} - \frac{1}{q} = \frac{\sigma}{3}
$$

Notation. Below we use the following notations and identities:

$$
\begin{align*}
\delta A &= A - A', \\
\delta \partial^j A &= \partial^j A - \partial^j A', \\
\delta F &= F - F', \\
\delta D_x^m F &= D_x^m F - (D_x^m)' F' \\
\delta w &= w - w', \\
\delta D_x^m w &= D_x^m w - D_x^m w', \\
\delta D^j D_j &= D^j D_j - D^j D_j'
\end{align*}
$$

The following lemma shows that by slightly reducing the support of the cutoff functions we can improve the exponents of powers of $s$ for differences of solutions that coincide at the initial time on a domain.

Lemma 8.2. Let $A_{t,x,s}, A'_{t,x,s}$ be two solutions like in Proposition 8.1. Let $\chi_k, \chi_{k+1}$ be bump functions supported in $B^{(1)}$ such that $\chi_k \equiv 1$ on $\text{supp} \chi_{k+1}$. Fix $t \in [t_0, t_1]$. Let $k \geq 0$ and suppose that for $m = 0, 1$, $p \in [2, 4]$ one has

$$
(N s^{\frac{1}{2}})^{k+1}(1-\sigma) \|\chi_k \delta D_x^m F(s)\|_{L^2} \lesssim \varepsilon s^{\frac{k+1}{2}}
$$

(8.17)

$$
\|\chi_k \delta D_x^{m+1} F\|_{L^2_{t,x} L^2_{x,t} [0, s]} \lesssim \varepsilon s^{\frac{k+1}{2}}
$$

(8.18)

$$
\|\chi_k (A_x - A'_x)(s)\|_{L^p} \lesssim \varepsilon s^{\frac{k+2}{2}} s^{-\frac{1}{2} - \frac{1}{p}}
$$

(8.19)

$$
\|\chi_k (\delta^j A_t - \partial^j A'_t)(s)\|_{L^2} \lesssim \varepsilon s^{\frac{k+1}{4}}
$$

(8.20)
for any \( s, s_1 \in [0, s_0] \). Then one has
\[
(N s_2^2)^{(k+2)(1-\sigma)} s_s^m \| \chi_{k+1} \delta F \|_{L^2} \lesssim \varepsilon s^{-\frac{k+1}{4}} \tag{8.21}
\]
\[
\bigl( (N s_2^2)^{(k+2)(1-\sigma)} s_s^m \chi_{k+1} \delta F \bigr)_{L^2[0,s_1]} \lesssim \varepsilon s^{-\frac{k+1}{4}} \tag{8.22}
\]
\[
(N s_2^2)^{(k+2)(1-\sigma)} \| \chi_{k+1} (A - F) \|_{L^p} \lesssim \varepsilon s^{-\frac{k+4}{4}} s^{\frac{3}{2}-\frac{1}{p}} \tag{8.23}
\]
\[
(N s_2^2)^{(k+3)(1-\sigma)} \| \chi_{k+1} (\partial^\ell A - \partial^\ell F) \|_{L^2} \lesssim \varepsilon s^{-\frac{k+4}{4}} \tag{8.24}
\]
Moreover, if one assumes \( k \geq 2 \) and
\[
(N^2 s_1)^{1-\sigma} \| \chi_{k+1} \delta w \|_{L^2} + \| (N^2 s)^{1-\sigma} \chi_{k+1} \delta F \|_{L^2[0,s_1]} \lesssim \frac{\varepsilon s_1^k}{(N s_1^2)^{(k-2)(1-\sigma)}} \tag{8.25}
\]
for all \( s_1 \in [0, s_0] \) then one also has
\[
(N^2 s_1)^{1-\sigma} \| \chi_{k+1} \delta w \|_{L^2} + \| (N^2 s)^{1-\sigma} \chi_{k+1} \delta F \|_{L^2[0,s_1]} \lesssim \frac{\varepsilon s_1^{k-4}}{(N s_1^2)^{(k-1)(1-\sigma)}} \tag{8.26}
\]
Assuming this Lemma we can now prove the main estimates of this section.

8.1. Proof of Proposition 8.1. Fix a time \( t \in [t_0, t_1] \). By integrating equation (8.8) point-wise and the fact that the solutions coincide at \( s = 0 \) on \( B(1) \), we have
\[
A(t) - A(t) = \int_0^s D^j F(s) - D^j F(s) \, ds, \quad x \in B(1)
\]
which implies, by estimating \( D^j F \) and \( D^j F \) separately using (8.1), (8.2), (8.13), that
\[
(N s^{\frac{1}{2}})^{1-\sigma} \| A - A \|_{L^p(B(1))} \lesssim \varepsilon s^{-\frac{3}{4}} s^{\frac{3}{2} - \frac{1}{p}}, \quad p \in [2, 4]
\]
By integrating equation (8.9), on \( B(1) \) we have
\[
\partial^\ell A(t) - \partial^\ell A(t) = \int_0^s [A - A, D_x F] + [A', D_x F - D_x F'] \, ds
\]
We bound the integrand in \( L^2 \). We use (8.27) with \( p = 2 \) and (8.13) with \( p = \infty \) for the first term. The second term we bound by \( (L^6 + N^{\sigma-1} L^4) \times (L^3 \cap L^{3/\sigma}) \rightarrow L^2 \) using (8.15) and (8.13) with \( p = 3 \) and \( p = 3/\sigma \), for both \( D_x F \) and \( D_x F' \) separately. We obtain
\[
(N s^{\frac{1}{2}})^{2(1-\sigma)} \| \partial^\ell A(t) - \partial^\ell A(t) \|_{B(1)} \lesssim \varepsilon s^{\frac{1}{4}}
\]
We will apply Lemma 8.2 several times with different cutoff functions to obtain (8.6) and (8.7), each time getting an extra power of \( s^{1/4} \) and slightly reducing the support.

Given \( \chi \) supported in \( B(1) \), we consider a sequence of bump functions \( \chi_0, \chi_1, \ldots, \chi_8 = \chi \) all supported in \( B(1) \) such that \( \chi_k \equiv 1 \) on \( \text{supp} \chi_{k+1} \), for each \( k \in \{0, \ldots, 7\} \).

Based on (8.1), (8.2), (8.27), (8.28) we first apply Lemma 8.2 with \( k = 0 \) and \( \chi_0, \chi_1 \). We obtain (8.17)-(8.20) with \( k = 1 \) for \( \chi_1 \). We repeatedly apply Lemma 8.2 with \( \chi_k, \chi_{k+1} \) until we obtain (8.6). Starting from \( k \geq 2 \) we also obtain bounds on \( \delta w \): in the \( k = 2 \) step the bound (8.25) follows trivially from (8.4), (8.5) by the triangle inequality. When \( k = 7 \) we conclude (8.7).

8.2. Proof of Lemma 8.2. For brevity, in the course of this proof we denote \( \tilde{\chi} := \chi_k \) and \( \chi := \chi_{k+1} \). Fix a time \( t \in [t_0, t_1] \).
8.2.1. **Difference equation for \( F \).** We first prove (8.21), (8.22) for \( m = 0 \).

Subtracting (8.10) for \( F \) and \( F' \) and using the form of \( \delta D \delta D_t \) from (8.16) we obtain the schematic equation (where we drop some constants and irrelevant signs)

\[
\left( \partial_s - \delta D \delta D_t \right) (\chi \delta F) = \chi [\delta F, F] + \chi [\delta F, F'] + \nabla \chi \cdot \delta D_x F + \Delta \chi \delta F + [\delta A^t, \partial_t (\chi F')] + \chi [\delta \partial_t A_t, F'] + \chi [A^t, [A_t, F']] - \chi [A^t, [A_t, F']] \tag{8.29}
\]

Note that \( \chi \delta F(0) = 0 \). Using the energy estimate (8.5) with \( p = 0 \) we obtain

\[
\| \chi \delta F \|_{L^\infty_t L^2_x [0, s_1]} + \| s^\frac{1}{2} \delta D_x \left( \chi \delta F(s) \right) \|_{L^2_t L^2_x [0, s_1]} \lesssim \varepsilon s_1^{\frac{k+1}{2}} \left( N s_1^\frac{1}{2} \right)^{-(k+2)(1-\sigma)} \tag{8.33}
\]

which will give (8.21), (8.22) for \( m = 0 \).

We estimate the nonlinear terms in \( L^1_t L^2_x [0, s_1] \) as follows:

For (8.29) we use (8.17) with \( m = 0 \) and (8.13) with \( p = \infty \), \( m = 0 \);

For (8.30) we use (8.17), (8.18) with \( m = 0 \), recalling that \( \chi \equiv 1 \) on supp \( \chi \);

For (8.31) we use (8.20), (8.13) with \( p = \infty \); to estimate the second term. Then use (8.19), (8.13) with \( m = 1, 0, p = \infty \) to estimate \( \chi [\delta A^t, \delta D_x F'] \) and \( [\delta A^t, \partial_t \chi F'] \) and write everything in terms of

\[
O(A^t, \chi \delta A, F') \quad \text{and} \quad O(\chi \delta A, \tilde{\chi} \delta A, F') \tag{8.35}
\]

The first term is estimated as \( (L^6 + N^{\sigma-1} L^9) \times (L^3 \cap L^p) \times L^\infty \to L^2 \) using (8.15), (8.19) with \( p = 3/\sigma \) and (8.13). The second term is bounded as \( L^4 \times L^4 \times L^\infty \to L^2 \) from (8.19), (8.13).

8.2.2. **Difference equation for \( D_x F \).** Similarly one writes the parabolic equation for differences from (8.11)

\[
\left( \partial_s - \delta D \delta D_t \right) (\chi \delta D_x F) = \chi O(\delta F, D_x F) + \chi O(F', \delta D_x F) + \nabla \chi \cdot \delta D_x^2 F + \Delta \chi \delta D_x F + [\delta A^t, \partial_t (\chi D_x F')] + \chi [\delta \partial_t A_t, D_x F'] + \chi [A^t, [A_t, D_x F']] - \chi [A^t, [A_t, D_x F']] \tag{8.36}
\]

Denote the nonlinearity by \( M \). Using the energy estimate (6.5) with \( p = \frac{1}{2} \) one obtains

\[
\| \chi s^\frac{1}{2} \delta D_x F \|_{L^\infty_t L^2_x [0, s_1]} + \| s \delta D_x \left( \chi \delta D_x F(s) \right) \|_{L^2_t L^2_x [0, s_1]} \lesssim \varepsilon s_1^{\frac{k+1}{2}} \left( N s_1^\frac{1}{2} \right)^{-(k+2)(1-\sigma)} \tag{8.40}
\]

We prove

\[
\| \chi s^\frac{1}{2} \delta D_x F \|_{L^\infty_t L^2_x [0, s_1]} + \int_0^{s_1} s^\frac{1}{2} \| M(s) \|_{L^2_x} \, ds \lesssim \varepsilon s_1^{\frac{k+1}{2}} \left( N s_1^\frac{1}{2} \right)^{-(k+2)(1-\sigma)} \tag{8.42}
\]

The first term in (8.41) was estimated in (8.34). To estimate \( M \) in \( L^1_t L^2_x [0, s_1] \) we proceed exactly as in the previous step, with the following difference: for the terms which have \( D_x F \) instead of \( F \), when applying (8.17), (8.18), (8.13) we use \( m = 1 \) instead of \( m = 0 \).
As before, the second term in (8.40) can be replaced by \( s \chi \delta \mathbf{D}_x^2 F(s) \), which completes the proof of (8.21), (8.22) also for \( m = 1 \).

8.2.3. **Bound on \( A \) and \( \text{div} A \).** By integrating equations (8.8), (8.9) and the fact that the solutions coincide at \( s = 0 \) on the support of \( \chi \) we have

\[
\chi \left( A_t(s_1) - A_t'(s_1) \right) = \chi \int_0^{s_1} \delta \mathbf{D}_x F(s) \, ds
\]

(8.43)

\[
\chi \left( \partial^\ell A_t(s_1) - \partial^\ell A_t'(s_1) \right) = -\chi \int_0^{s_1} \left[ \delta A, \mathbf{D}_x F \right] + \left[ A', \delta \mathbf{D}_x F \right] \, ds
\]

(8.44)

By the covariant Gagliardo-Nirenberg inequalities (2.11) we have

\[
\| \chi \delta \mathbf{D}_x F(s) \|_{L^p_\ell} \lesssim \| \chi \delta \mathbf{D}_x F(s) \|_{L^2_\ell}^{1-\theta} \| \mathbf{D}_x (\chi \delta \mathbf{D}_x F(s)) \|_{L^2_\ell}^\theta \quad \theta = 3(2^m - 1)
\]

(8.45)

Integrating this and using Holder in \( \frac{d}{ds} \) we obtain (8.23) from (8.40).

For (8.24) we estimate the integral (8.44) as follows: for the first term use (8.23) with \( p = 2 \) and (8.13) with \( p = \infty \), \( m = 1 \); the second term is estimated as \( (L^6 + N^{\sigma - 1} L^q) \times (L^3 \cap L^p) \rightarrow L^2 \) using (8.15) and (8.45) for \( L^3 \) and \( L^p \) with \( p = 3/\sigma \), together with (8.40).

8.2.4. **Bound on \( w \).** Subtracting (8.12) for \( w \) and \( w' \) and using the form of \( \delta \mathbf{D}_x \) from (8.16) we obtain the schematic equation

\[
\left( \partial_s - \mathbf{D}_x \right) (\chi \delta w) = \chi [\delta F, w] + \chi [F, \delta w]
\]

(8.46)

\[
+ \nabla \chi \cdot \delta \mathbf{D}_x w + \Delta \chi \delta w
\]

(8.47)

\[
+ [\delta A^\ell, \partial_t (\chi w')] + \chi [\delta \partial^\ell A_t, w']
\]

(8.48)

\[
+ \chi [A^\ell, [A_t, w']] - \chi [A^\ell, [A_t, w']]
\]

(8.49)

(8.50)

Using the energy estimate (6.5) with \( p = 0 \) we obtain

\[
\mathcal{T} := \| \chi \delta w \|_{L^\infty L^2_{\sigma}[0, s_1]} + \| s^{\frac{1}{2}} \mathbf{D}_x (\chi \delta w) \|_{L^2_{\sigma} L^2_{\sigma}[0, s_1]} \lesssim \varepsilon \mathcal{T} + \frac{\varepsilon s_1^{\frac{k-1}{1}}}{(N s_1^2)^{(k+2)(1-\sigma)}}
\]

(8.51)

provided we bound (8.46) - (8.50) in \( L^1 L^2_{\sigma}[0, s_1] \) by the RHS of (8.51). Then we absorb the \( \varepsilon \mathcal{T} \) to the LHS and, replace the \( \mathbf{D}_x (\chi \delta w) \) by \( \chi \delta \mathbf{D}_x w \) by bounding \( s^{\frac{1}{2}} \nabla \chi \cdot \delta w \) and \( s^{\frac{1}{2}} \chi [\delta A, w] \) using (8.25), (8.19) and (8.14) obtaining

\[
\| \chi \delta w(s_1) \|_{L^2_\ell} + \| s^{\frac{1}{2}} \chi \delta \mathbf{D}_x w \|_{L^2_{\sigma} L^2_{\sigma}[0, s_1]} \lesssim \frac{\varepsilon s_1^{\frac{k-1}{1}}}{(N s_1^{2})^{(k+2)(1-\sigma)}}
\]

This is more than enough for (8.26). We estimate the nonlinear terms in \( L^1 L^2_{\sigma}[0, s_1] \) very similarly to the case of \( \chi \delta F \):

For (8.46) we first use (8.17) with \( m = 0 \) and (8.14) with \( p = \infty \), \( m = 0 \); then for \( \chi [F, \delta w] \) we use (8.13) with \( p = \infty \), \( m = 0 \), which contributes the term \( \varepsilon s_1^{\frac{k}{1}} \mathcal{T} \ll \varepsilon \mathcal{T} \).

The terms in (8.47) were already estimated in (8.36).

For (8.48) we use (8.25), recalling that \( \bar{\chi} \equiv 1 \) on sup\( \chi \);

For (8.49) we use (8.20), (8.14) with \( p = \infty \); to estimate the second term. Then use (8.19), (8.14) with \( m = 1, 0 \), \( p = \infty \) to estimate \( \chi [\delta A^\ell, \mathbf{D}_x w'] \) and \( [\delta A^\ell, \partial_t \chi w'] \).

For (8.50) and the remainder from (8.49), write everything in terms of

\[
\mathcal{O}(A^\ell, \chi \delta A, w') \quad \text{and} \quad \mathcal{O}(\chi \delta A, \bar{\chi} \delta A, w')
\]
and then proceed exactly like in (8.35), using (8.14) instead of (8.13).

9. Decompositions and Estimates in DeTurck’s Gauge

In this section we consider solutions \(A_{t,x,s}\) to (YM) and (dYMFF) on \([t_0, t_1] \times \mathbb{R}^3 \times [0, s_0]\), \(t_1 = t_0 + 1\) in:

1. The temporal gauge at \(s = 0\), i.e. \(A_0(t, x, 0) = 0\).
2. The DeTurck gauge \(A_s = \partial^t A_\ell\) on \([t_0, t_1] \times \mathbb{R}^3 \times [0, s_0]\).
3. The Coulomb gauge at \(s = 0\), \(t = t_0\), i.e. \(\partial^t A_\ell(t_0, x, 0) = 0\).

Let \(\sigma > \frac{5}{6}\). We assume \(A_i(t, x, 0)\) obeys the bounds in Proposition 4.1 part (2) and therefore also (4.12) - (4.14). Then Proposition 5.1 holds uniformly in \(t \in [t_0, t_1]\), and as a consequence Proposition 5.13, 5.14.

We assume \(\mathcal{E}(t) \lesssim \varepsilon^2\) control of the modified energy (1.6) for all \(t \in [t_0, t_1]\). In particular, by Proposition 6.3 and 6.4 the curvature \(F_{\alpha\beta}\) and the tension field \(w_i\) defined in (1.9) obey the bounds

\[
\|(N^2 s)^{1-\sigma} (s^{\frac{1}{2}} D_x)^m F_{\alpha\beta}(t, s)\|_{L^\infty_t L^2_x \cap L^2_t L^2_x([0, s_0] \times \mathbb{R}^3)} \lesssim \varepsilon \tag{9.1}
\]

\[
\|(N^2 s)^{1-\sigma} (s^{\frac{1}{2}} D_x)^m w_i(t, s)\|_{L^\infty_t L^2_x \cap L^2_t L^2_x([0, s_0] \times \mathbb{R}^3)} \lesssim \varepsilon \tag{9.2}
\]

for all \(0 \leq m \leq 8\) and \(t \in [t_0, t_1]\).

Under these assumptions we control the norms \((9.14) - (9.23)\).

9.1. Decompositions in DeTurck’s gauge.

9.1.1. Decomposition of \(F\).

Expanding (6.1) and canceling the \([A_s, F]\) term with the \([\partial^t A_{\ell}, F]\) term one obtains

\[
(\partial_s - \Delta) F_{\alpha\beta} = 2[F_{\alpha}, F_{\beta}] + 2[A^\ell, \partial_t F_{\alpha\beta}] + [A^\ell, [A_{\ell}, F_{\alpha\beta}]] \tag{9.3}
\]

and one decomposes

\[
F_{\alpha\beta}(s) = e^{s\Delta} F_{\alpha\beta}(0) + F_{\alpha\beta}^{bil}(s) \tag{9.4}
\]

where the bilinear term

\[
F_{\alpha\beta}^{bil}(s_1) := \int_0^{s_1} e^{(s_1-s)\Delta} \left( 2[F_{\alpha}, F_{\beta}] + 2[A^\ell, \partial_t F_{\alpha\beta}] + [A^\ell, [A_{\ell}, F_{\alpha\beta}]] \right)(s) \, ds \tag{9.5}
\]

obeys better estimates, such as

**Lemma 9.1.** Under the assumptions above, one has:

\[
\|F_{\alpha\beta}^{bil}(s)\|_{L^\infty_t L^2_x} + s^{\frac{1}{2}} \|\partial_x F_{\alpha\beta}^{bil}(s)\|_{L^\infty_t L^2_x} \lesssim s^{\frac{1}{2}}/(N^2 s)^{1-\sigma} \varepsilon. \tag{9.6}
\]

**Proof.** Fix a time \(t \in [t_0, t_1]\). It suffices to bound the three terms in the bracket in (9.5) in \(L^2_x\) by \(s^{-\frac{3}{2}}/(N^2 s)^{1-\sigma} \varepsilon\). For the second estimate in (9.4) one begins with the fact that \((s_1 - s)^{\frac{1}{2}} \partial_x e^{(s_1-s)\Delta} : L^2_x \to L^2_x\).

For \([F, F]\) and \([A^\ell, D_{\ell} F_{\alpha\beta}]\) we use Holder, placing the first term in \(L^\infty_x\) and the second in \(L^2_x\), recalling (5.13). For \([A, [A, F]]\) we do \(L^\infty_x \times L^\infty_x \times L^2_x \to L^2_x\). \(\square\)
Lemma 9.3. Under the assumptions above, one has:

\[(\partial_s - \Delta)w_k = \mathcal{N}^{(2)} + \mathcal{N}^{(3)} + \mathcal{N}^{self}\]

where

\[\mathcal{N}^{(2)}(s) := 2[e^{s\Delta}\partial_t A^f(0), e^{s\Delta}(\partial_k \partial_t A_\ell - 2\partial_t \partial_\ell A_k)(0)]\]

\[\mathcal{N}^{(3)}(s) := [F_{0\ell}^{bil}, D_x F_{0x}^z] + 2[e^{s\Delta} \partial_t A^f(0), \partial_x F_{0x}^{bil} + [A_x, F_{0x}]]\]

\[\mathcal{N}^{self}(s) := 2[F_k^e, w_\ell] + 2[A^e, \partial_t w_k] + [A_{i\ell}^i, [A_\ell, w_k]]\]

where \(\mathcal{N}^{(2)}\) contains the dominant part from 2 \([F_{0\ell}, D_k F_{0\ell} + 2D_{\ell}F_{k0}]\) (recall the temporal gauge condition implies \(F_{0\ell} = \partial_t A_\ell\) at \(s = 0\)). The term \(\mathcal{N}^{(3)}\) consists of the remainder and \(\mathcal{N}^{self}\) contains the terms involving \(w\).

Since \(w(0) = 0\), this gives rise to the following decomposition which isolates the leading quadratic part of \(w\):

\[w(s) = w^{(2)}(s) + w^{(3)}(s)\]  

(9.8)

where

\[w^{(2)}(s_1) := \int_0^{s_1} e^{(s_1-s)\Delta} \mathcal{N}^{(2)}(s) \, ds, \quad w^{(3)}(s_1) := \int_0^{s_1} e^{(s_1-s)\Delta} (\mathcal{N}^{(3)}(s) + \mathcal{N}^{self}(s)) \, ds\]

The term \(w^{(2)}(s)\) can be written, following [39], as

\[w_i^{(2)}(s) = W(\partial_\ell A^f(0), \partial_\ell \partial_t A_\ell(0) - 2\partial_t \partial_\ell A_i(0))\]  

(9.9)

where \(W\) is a symmetric bilinear form with symbol

\[W(\xi, \eta, s) = \int_0^s e^{-(s-s')|\xi+\eta|^2} e^{-s'(|\xi|^2+|\eta|^2)} \, ds'\]

(9.10)

which enjoys the following favorable frequency concentration:

**Lemma 9.2** ([41] - Lemma 8.2). For any \(k, k_1, k_2 \in \mathbb{Z}\) and \(s > 0\), denoting \(k_{\text{max}} = \max(k, k_1, k_2)\), the bilinear operator

\[\langle s 2^{2k}\rangle^{10} \langle s^{-12-2k_{\text{max}}} \rangle 2^{2k_{\text{max}}} P_k W(P_{k_1}(\cdot), P_{k_2}(\cdot), s)\]

is disposable, i.e. its kernel has bounded mass, see [22.5].

Using this structure of \(W\), one can obtain more precise estimates for \(w^{(2)}\), such as

**Lemma 9.3.** Under the assumptions above, one has:

\[\|P_k w^{(2)}_x(s)\|_{L_t^2 L_x^2} \lesssim \varepsilon 2^{(0+\delta)k} \frac{2^k s^{\frac{1}{4}}}{(2^{2k}s)^8} \frac{1}{(N^2 s)^{1-\sigma}} \left(1 + \frac{1}{2^{7k}} + \frac{1}{(N s)^{1-\sigma}}\right)\]

(9.11)

Therefore, after summing in \(k\) one obtains:

\[\|w^{(2)}_x(s)\|_{L_t^2 L_x^2} \lesssim \varepsilon s^{0-}/(N^2 s)^{1-\sigma}\]

(9.12)

**Remark 9.4.** One can decompose the inputs of \(w^{(2)}_x(s)\) into \(\partial_t A^{cf} + \partial_t A^{df}\). It is clear from the proof that the bound (9.11) is true also for the separate components of \(w^{(2)}_x(s)\), in particular for \(P_k w^{(2)}(\partial_t A^{cf}, \partial_t A^{df}, s)\).
Proof. Doing a Littlewood-Paley decomposition and invoking Lemma 9.2 we need to bound

\[ P_k O(\partial_t A_{k_1}(0), \partial_t A_{k_2}(0)) \cdot 2^{-k_{\text{max}}} \left<s2^{2k}\right>^{-10} \left<s^{-1}2^{-2k_{\text{max}}}\right>^{-1} \]  

(9.13)

in the two cases, without loss of generality:

1. low-high \(2^{k_1} \lesssim 2^k \sim 2^{k_2} \lesssim 2^{k_{\text{max}}}\). Bound \(\partial_t A_{k_1}(0)\) in \(L^2 L^\infty\) using (4.14) and \(\partial_t A_{k_2}(0)\) in \(L^\infty L^2\). Then one sums in \(k_1\), obtaining the bound

\[ \varepsilon 2^{(0+k)} \frac{2^{2k_s}}{(2^{k_s})^{10}} \frac{1}{(N^2 s)^{1-\sigma}}. \]

(2) high-high to low \(2^k \lesssim 2^{k_1} \sim 2^{k_2} \lesssim 2^{k_{\text{max}}}\). Split both inputs into \(\partial_t A^{cf} + \partial_t A^d\). When at least one input is \(\partial_t A^{cf}\), place that input in \(L^2 L^\infty\) using (4.12) and the other one in \(L^\infty L^2\). After summing the high frequencies one obtains the bounds

\[ \varepsilon \frac{1}{2^{\frac{1}{2}}} \left(\frac{2^k}{N}\right)^{3(1-\sigma)} \frac{1}{(2^{k_s})^{10}} \text{ or } \varepsilon \frac{1}{2^{\frac{1}{2}}} \left(\frac{2^k}{N}\right)^{3(1-\sigma)} \frac{1}{(2^{k_s})^{10}} \]

depending on whether \(s^{-\frac{1}{2}} = k(s) \leq k\) or \(k \leq k(s)\).

The remaining case \(P_k O(\partial_t A^{cf}_{k_1}(0), \partial_t A_{k_2}^d(0))\) is bounded using the Bilinear Strichartz estimate (2.31). After summation one can obtain either

\[ \varepsilon 2^{(0+k)} \frac{2^{2k_s}}{(2^{k_s})^{10}} \frac{1}{(N^2 s)^{1-\sigma}} \text{ or } \varepsilon 2^{(0+k)} \frac{(2^{k_s})^{-\frac{1}{2}-1}}{(2^{k_s})^{10}} \frac{1}{(N^2 s)^{1-\sigma}} \]

depending on whether \(s^{-\frac{1}{2}} = k(s) \leq k\) or \(k \leq k(s)\).

\[ \square \]

9.2. Estimates for the curvature and the tension field.

9.2.1. Preparations. For easy reference we record here the bounds on \([t_0, t_1] \times \mathbb{R}^3\) which we will use below. These follow from (4.13), (4.14), (2.30), (4.14), (5.13), (9.1), (2.11).

\[ \|P_k A^{cf}(0)\|_{L^\infty_t L^2_x} \lesssim \varepsilon 2^{-\frac{1}{2}k} \left(\frac{2^k}{N}\right)^{1-\sigma} \]  

(9.14)

\[ s^\frac{1}{2} \|e^{s \Delta} A^{cf}(0)\|_{L^\infty_t L^2_x} \lesssim \|e^{s \Delta} A^{cf}(0)\|_{L^\infty_t L^2_x} \lesssim \varepsilon \frac{1}{(N s^2)^{1-\sigma}} \]  

(9.15)

\[ \|\partial_t P_k A^{cf}(0)\|_{L^\infty_t L^2_x} \lesssim \varepsilon 2^{-\frac{1}{2}k} \left(\frac{2^k}{N}\right)^{2(1-\sigma)} \]  

(9.16)

\[ \|\partial_t P_k A^d(0)\|_{L^\infty_t L^2_x} \lesssim \varepsilon \left(\frac{2^k}{N}\right)^{1-\sigma} \]  

(9.17)

\[ \|P_k \nabla_{t,x} A^{d}(0)\|_{L^2_t L^\infty_x} \lesssim 2^{(1+k)} \varepsilon 2^{-\frac{1}{2}k} \left(\frac{2^k}{N}\right)^{1-\sigma} \]  

(9.18)
\[ \|e^{s \Delta} A_{\frac{d}{2}}(0)\|_{L^2_t L^\infty_x} \lesssim s^{1 - \sigma}/(Ns^{\frac{1}{2}})^{1 - \sigma} \]  
(9.19)

\[ \|e^{s \Delta} \partial_x A_{\frac{d}{2}}(0)\|_{L^2_t L^\infty_x} \lesssim s^{\frac{1}{2} - \frac{1}{2}}/(Ns^{\frac{1}{2}})^{1 - \sigma} \]  
(9.20)

\[ \|e^{s \Delta} \partial_t A_{\frac{d}{2}}(0)\|_{L^2_t L^\infty_x} \lesssim s^{\frac{1}{2} - \frac{1}{2}}/(Ns^{\frac{1}{2}})^{1 - \sigma} \]  
(9.21)

\[ \|F(s)\|_{L^\infty_t L^2_x} \lesssim \varepsilon/(Ns^{\frac{1}{2}})^{1 - \sigma} \]  
(9.22)

\[ \|F(s)\|_{L^\infty_t L^2_x} \lesssim \varepsilon s^{\frac{1}{2}}/(Ns^{\frac{1}{2}})^{1 - \sigma} \]  
(9.23)

9.2.2. Bounds on the curvature.

**Proposition 9.5.** Under the assumptions above, one has:

\[ \|F^{\text{bil}}(s)\|_{L^2_t L^\infty_x} \lesssim s^{\frac{1}{2} + \frac{\sigma}{2}}/(Ns^{\frac{1}{2}})^{3(1 - \sigma)} \varepsilon \]  
(9.24)

\[ \|F^{\text{bil}}(s)\|_{L^2_t L^\infty_x} + s^{\frac{1}{2}} \|\partial_x F^{\text{bil}}(s)\|_{L^2_t L^\infty_x} \lesssim s^{\frac{1}{2} + \frac{\sigma}{2}}/(Ns^{\frac{1}{2}})^{2(1 - \sigma)} \varepsilon \]  
(9.25)

**Proof.** **Part 1.** We begin with (9.24). If \( \mathcal{T} \) is one the three terms in the bracket in (9.5), then we bound either

\[ \left\| \int_0^s e^{(s - s') \Delta} \mathcal{T}(s') \, ds' \right\|_{L^2_t L^\infty_x} \text{ or } \left\| \int_0^s \frac{1}{(s - s')^{\frac{1}{2}}} \|\mathcal{T}(s')\|_{L^2_t L^\infty_x} \, ds' \right\|_{L^2_t L^\infty_x} \]

since one has \( (s - s')^{\frac{1}{2}} e^{(s - s') \Delta} : L^2_x \to L^\infty_x \). Let

1. \( \mathcal{T} = [F, F] \), which we place in \( L^2_t L^2_x \). For:
   
   \( [F^{\text{bil}}, F](s') \) we use (9.6) and (9.23), while for
   
   \[ e^{s' \Delta} \partial_t A_{\frac{d}{2}}(0), F] \] we use (9.20) or Strichartz (9.19) and (9.22).

2. \( \mathcal{T} = [A^\ell, \partial_t F_{\alpha \beta}] \) which separate as follows:
   
   \[ e^{s' \Delta} A_{\frac{d}{2}}(0) + A^{\text{bil}}(s'), D_t F] = L^2_t L^\infty_x \times L^\infty_x L^2 \to L^2_t L^2 \text{ by } (9.19), \, (5.14). \]
   
   \[ e^{s' \Delta} A_{\frac{d}{2}}(0), \partial_x F^{\text{bil}} + e^{s' \Delta} \partial_x \partial_t A_{\frac{d}{2}}(0) + e^{s' \Delta} \partial_x [A, A](0)] = L^\infty_x L^\infty_x L^2 \to L^2_t L^2 \text{ using } (9.10), \, (9.12), \, (9.36). \]
   
   \[ e^{s' \Delta} A_{\frac{d}{2}}(0), e^{s' \Delta} \partial_x \partial_t A_{\frac{d}{2}}(0)] \text{ which after the } s' \text{ integration takes the form } \]

\[ \mathbf{W}(A_{\frac{d}{2}}(0), \partial_x \partial_t A_{\frac{d}{2}}(0)) \]  
(9.26)

\( \mathbf{W} \) is defined in (9.10). We use Lemma 9.2 and the Littlewood Paley trichotomy to bound this term in \( L^2_t L^\infty_x \). In the high-low and high-low input frequencies cases we place \( A_{\frac{d}{2}} \) in \( L^\infty_t L^\infty_x \) and \( \partial_t A_{\frac{d}{2}} \) in \( L^\infty_t L^\infty_x \) using Strichartz. In the high-high to low case use Bernstein on the output frequency and \( L^2 \times L^2 \to L^2 \) with (9.14) and (9.18).

3. \( \mathcal{T} = [A^\ell, [A, F]] \) which is placed in \( L^2_t L^2 \) as follows:
   
   For \( e^{s' \Delta} A_{\frac{d}{2}}(0) + A^{\text{bil}}(s'), [A, F](s')] \) as \( L^2_t L^\infty_x \times L^\infty_x L^\infty_x \times L^\infty_x L^2 \) using (9.19), (5.14), (9.21), (9.22).

   Similarly one deals with \( [e^{s' \Delta} A_{\frac{d}{2}}(0), [e^{s' \Delta} A_{\frac{d}{2}}(0), A^{\text{bil}}(s'), F(s')]] \).

   The remaining term \( [e^{s' \Delta} A_{\frac{d}{2}}(0), e^{s' \Delta} A_{\frac{d}{2}}(0), F(s')] \) is bounded as \( L^\infty_t L^2 \times L^\infty_t L^2 \times L^\infty_t L^2 \) using (9.15) and interpolating between (9.22) and (9.23).

**Part 2.** We continue with (9.25). We bound the bracket from (9.5) in \( L^2_t L^2 \). This suffices also for the \( \partial_x F^{\text{bil}}_0 \) bound because \( \partial_x e^{(s_1 - s) \Delta} : L^2_x \to (s_1 - s)^{-\frac{3}{2}} L^2_x \). The majority of terms in (9.5) were already estimated in \( L^2_t L^2 \) in Part 1. The term that remains to be estimated in \( L^2_t L^2 \) is (9.26).

Again we use Lemma 9.2 and the Littlewood Paley trichotomy. In the high-high to low and high-low cases use \( L^\infty_t L^2 \times L^\infty_t L^2 \to L^2_t L^2 \) with (9.14) and (9.18). In the low-high case \( (A_{\frac{d}{2}} \text{ - low frequency}) \) place \( A_{\frac{d}{2}}(0) \) in \( L^\infty_t L^\infty_x \) and \( \partial_t A_{\frac{d}{2}}(0) \) in \( L^\infty_t L^\infty_x \) using Bernstein. This last case is responsible for the loss of \( s^{-\frac{1}{2}} \) in (9.25), which can probably be removed. ∎
Lemma 9.6. One has:
\[ \|F_{\alpha\beta}(s)\|_{L^2 L^\infty_x} \lesssim \varepsilon s^{-\frac{1}{2}}/(N s^{\frac{1}{2}})^{1-\sigma} \] (9.27)

Proof. From (9.4) and the temporal condition \( A_0(s = 0) = 0 \) one writes
\[ F_{\alpha\beta}(s) = e^{s\Delta} \partial_s A_i(0) + F_{\alpha\beta}^{bil}(s) \]
and the bound follows from (9.20) and (9.24). Similarly for \[ \|\partial_s w(s)\|_{L^\infty L^1_x} \lesssim s^{-\frac{1}{2}}/(N s^{\frac{1}{2}})^{1-\sigma} \varepsilon \] (9.28)

The resulting term is absorbed to the LHS similarly to (6.8).

9.2.3. Bounds on the tension field \( w \).

Lemma 9.7. Under the assumptions above, one has:
\[ \|w_x(s)\|_{L^\infty L^1_x} + s^{\frac{1}{2}}\|\partial_x w_x(s)\|_{L^\infty L^1_x} \lesssim s^{\frac{1}{2}}/(N s^{\frac{1}{2}})^{1-\sigma} \varepsilon \] (9.29)

Proof. Fix a time \( t \in [t_0, t_1] \). The first inequality was already proved in (6.7).

For the second inequality we write schematically
\[ \partial_x w(s_1) = \int_0^{s_1} \partial_x e^{(s_1-s)\Delta} ([F, D_x F] + [F, w] + [A, [A, w]] + [A, \partial_x w]) \, ds \]
We use the fact that \( \partial_x e^{(s_1-s)\Delta} : L^1_x \to (s_1-s)^{-\frac{1}{2}} L^1_x \) so bound the round bracket in \( L^1_x \). The main term is
\[ \int_0^{s_1} (s_1-s)^{-\frac{1}{2}}\|F\|_{L^2} \|D_x F\|_{L^2} \, ds \lesssim \varepsilon/(N s^{\frac{1}{2}})^{1-\sigma} \]
Similarly for \( [F, w] \) and \( [A, [A, w]] \): place \( F(s) \), respectively \( A(s) \) s, in \( L^\infty_x \) and place \( w_x(s) \) in \( L^1_x \) using the first part from (9.28). Finally, for \( [A, \partial_x w] \), we place \( A(s) \) in \( L^\infty_x \) by (9.21) and \( \partial_x w \) in \( L^1_x \). The resulting term is absorbed to the LHS similarly to (6.8).

Lemma 9.8. Under the assumptions above, one has:
\[ \|w^{(3)}(s)\|_{L^2 L^1_x} \lesssim s^{1-\frac{1}{2}}/(N s^{\frac{1}{2}})^{3(1-\sigma)\varepsilon} \] (9.30)

Proof. Recall
\[ w^{(3)}(s_1) := \int_0^{s_1} e^{(s_1-s)\Delta} (\mathcal{N}^{(3)}(s) + \mathcal{N}^{self}(s)) \, ds \]
where \( \mathcal{N}^{(3)} \) and \( \mathcal{N}^{self} \) are defined in (9.7).

We begin with \( \mathcal{N}^{(3)} \):
\[ F_{\alpha\beta}^{bil} + D_x F_{\alpha\beta} \] is estimated as \( L^2 L^2 \times L^\infty L^2 \) using Proposition 9.5
\[ e^{s\Delta} \partial_\alpha A(0), \partial_x F_{\alpha\beta}^{bil} + [A_x, F_{\alpha\beta}] \] is estimated as \( L^\infty L^2 \times L^2 L^2 \) using Prop. 9.5.

We continue with \( \mathcal{N}^{self} \):
\[ F_{\alpha\beta}^{bil}, w \] is estimated as \( L^2 L^\infty \times L^\infty L^1 \) using Corollary 9.6 and Lemma 9.7
\[ e^{s\Delta} A^{diff} + A^{bil}, \partial_x w \] seen as \( L^2 L^\infty \times L^\infty L^1 \) using (9.14), (5.14) and Lemma 9.7
\[ e^{s\Delta} A^{self}, \partial_x w \] is estimated as \( L^\infty L^{12} \times L^\infty L^{14} \) using (9.15) and
\[ \|\partial_x w(s)\|_{L^\infty L^{12}} \lesssim \varepsilon s^{-\frac{1}{2}}/(N s^{\frac{1}{2}})^{2(1-\sigma)} \]

obtained by interpolating between (9.28) and \( (N^2 s)^{1-\sigma} s^{\frac{3}{4}} \partial_x w(s) \in \varepsilon L^\infty L^2 \).

\[ [A, [A, w]] \] is bounded in several ways, e.g. as \( L^\infty L^\infty \times L^\infty L^\infty \times L^\infty L^1 \).
10. Trilinear estimates

This section is devoted to the proof of (1.32).

We consider a solution which is in:

1. The temporal gauge at \( s = 0 \), i.e. \( A_0(t, x, 0) = 0 \).
2. The DeTurck gauge \( A_s = \partial_t^s A_t \) on \([t_0, t_1] \times \mathbb{R}^3 \times [0, s_0]\).
3. The Coulomb gauge at \( s = 0 \), \( t = t_0 \), i.e. \( \partial_t^s A_t(t_0, x, 0) = 0 \).

We assume \( A_i(t, x, 0) \) obeys the conditions in Prop. 4.1 part (2) and Prop. 5.1 uniformly in \( t \).

We assume \( IE(t) \lesssim \varepsilon^2 \) and therefore \([9.1], [9.2]\).

**Proposition 10.1.** Let \( \sigma > \frac{5}{6} \) and \( t_1 \leq t_0 + 1 \). Let \( A_{t,x,s} : [t_0, t_1] \times \mathbb{R}^3 \times [0, s_0] \rightarrow \mathfrak{g} \) be a regular solution to (YM) at \( s = 0 \) and to (YMHE) obeying the gauge conditions (1), (2), (3) above. Assume the bounds stated above and consequently we assume all the estimates from Section 7.

Then, if \( \varepsilon \) is small enough, one has

\[
(N^2 s)^{1-\sigma} \left| \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \chi(x)(w^f(s), F_0^f(s)) \, dx \, dt \right| \lesssim \frac{s^{\frac{5}{3}}}{(N s^{\frac{1}{2}})^{(1-\sigma)}} \varepsilon^2 \quad (10.1)
\]

After peeling off some terms, the main contribution to (10.1) will come from a null form term.

The key estimates needed for this main term are encapsulated by:

**Proposition 10.2.** Denoting \( k_{\text{max}} = \max(k_1, k_2, k_3) \) where \( k_1, k_2, k_3 \geq 0 \) are dyadic frequencies, one has the following inequality for frequency localized functions:

\[
\left| \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \mathcal{N}(f_{k_1}, g_{k_2}) h_{k_3} \, dx \, dt \right| \lesssim 2^{(\frac{5}{3}+k_{\text{max}})} \|f_{k_1}\|_{X^0_{\frac{1}{2}+}} \|g_{k_2}\|_{X^0_{\frac{1}{2}+}} \|h_{k_3}\|_{X^0_{\frac{1}{2}+}} \quad (10.2)
\]

By translation invariance, the same estimate holds for \( ON \) if \( O \) is a disposable operator.

Recall the definitions of a null form from (2.8), (2.9). Moreover, if \( O \) is a bilinear operator, we denote by \( ON(f, g) \) a linear combination of

\[
\Delta^{-1} \nabla^i \left( O(\partial_i f, \partial_j g) - O(\partial_j f, \partial_i g) \right) \quad \text{and} \quad O(\Delta^{-1} \nabla^i \partial_i f, \partial_j g) - O(\Delta^{-1} \nabla^i \partial_j f, \partial_i g)
\]

**Remark 10.3.** Proposition 10.2 states a classical type of multilinear estimate, except for the presence of the sharp time cutoff \( 1_{[t_0, t_1]} \). The proof is essentially contained in [22].

10.1. **Proof of Proposition 10.1.**

Splitting \( w(s) \) and \( F_0^f(s) \) by [9.4] and [9.8] we decompose

\[
\int_{t_0}^{t_1} \int_{\mathbb{R}^3} \chi(x)(w^f(s), F_0^f(s)) \, dx \, dt = T_1 + T_2 + T
\]

into two quadrilinear terms and a trilinear term

\[
T_1 := \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \chi(x)(w^{(3)}(s), F_0^f(s)) \, dx \, dt
\]

\[
T_2 := \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \chi(x)(w^{(2)}(s), F_0^f(s)) \, dx \, dt
\]

\[
T := \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \chi(x)(w^{(2)}(s), e^s \Delta A_x) \, dx \, dt
\]

48
We first estimate the quadrilinear terms by spacetime norms:

\[ |T_1| \lesssim \int_{t_0}^{t_1} \int_{\mathbb{R}^3} |\chi w^{(3)}(s, F(s))| \, dx \, dt \lesssim \|w^{(3)}(s)\|_{L_t^q L_x^r} \|F(s)\|_{L_t^q L_x^r} \tag{10.3} \]

which is bounded by \( s^{\frac{3}{2}} - / (N^2 s)^{2(1-\sigma)} \varepsilon^2 \) using Lemma 9.8 and Corollary 9.6.

\[ |T_2| \lesssim \int_{t_0}^{t_1} \int_{\mathbb{R}^3} |\chi (w^{(2)}(s), F^{bil}(s))| \, dx \, dt \lesssim \|w^{(2)}(s)\|_{L_t^q L_x^r} \|F^{bil}(s)\|_{L_t^q L_x^r} \tag{10.4} \]

which is also bounded by \( s^{\frac{3}{2}} - / (N^2 s)^{2(1-\sigma)} \varepsilon^2 \) using (10.12) and (10.25). Since \( \sigma > \frac{5}{8} \), this is more than enough.

It remains to consider the main term \( T \). Performing a Littlewood-Paley decomposition we need to bound

\[ \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k w^{(2)}(s), \hat{P}_k (\chi e^{s\Delta} \partial_t A_\ell)) \, dx \, dt \]

To pass the frequency localization onto \( \partial_t A_\ell \) write

\[ \hat{P}_k (\chi e^{s\Delta} \partial_t A_\ell) = \hat{P}_k (\chi e^{s\Delta} P_k \partial_t A_\ell) + \hat{P}_k (|\chi| P_k e^{s\Delta} \partial_t A_\ell) \]

The commutator term is bounded in (10.20). Then main term is

\[ \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k w^{(2)}(s), \chi e^{s\Delta} \partial_t P_k A_\ell) \, dx \, dt. \]

We split all inputs as \( \partial_t A = \partial_t A^{df} + \partial_t A^{cf} \). For the dominant term, with 3 divergence-free inputs, we further split \( w^{(2)} = P w^{(2)} + P^\perp w^{(2)} \) obtaining:

\[ \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k P w^{(2)}(\partial_t A^{df}, \partial_t A^{df}, s), \chi e^{s\Delta} \partial_t A^{df}_k) \, dx \, dt \tag{10.5} \]

which contains a null form, and the following commutator-type error term

\[ \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k P^\perp w^{(2)}(\partial_t A^{cf}, \partial_t A^{df}, s), \hat{P}_k (\chi e^{s\Delta} \partial_t A^{df}_k)) \, dx \, dt \tag{10.6} \]

Then we have terms with 1 curl-free input and 2 divergence-free inputs:

\[ \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k w^{(2)}(\partial_t A^{df}, \partial_t A^{df}, s), \chi e^{s\Delta} \partial_t A^{cf}_k) \, dx \, dt \tag{10.7} \]

\[ \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k w^{(2)}(\partial_t A^{cf}, \partial_t A^{df}, s), \chi e^{s\Delta} \partial_t A^{df}_k) \, dx \, dt \tag{10.8} \]

\[ \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k w^{(2)}(\partial_t A^{df}, \partial_t A^{cf}, s), \chi e^{s\Delta} \partial_t A^{df}_k) \, dx \, dt \tag{10.9} \]

terms with 2 curl-free inputs and 1 divergence-free input:

\[ \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k w^{(2)}(\partial_t A^{cf}, \partial_t A^{df}, s), \chi e^{s\Delta} \partial_t A^{df}_k) \, dx \, dt \tag{10.10} \]

\[ \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k w^{(2)}(\partial_t A^{cf}, \partial_t A^{df}, s), \chi e^{s\Delta} \partial_t A^{cf}_k) \, dx \, dt \tag{10.11} \]

\[ \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k w^{(2)}(\partial_t A^{df}, \partial_t A^{cf}, s), \chi e^{s\Delta} \partial_t A^{cf}_k) \, dx \, dt \tag{10.12} \]
and finally a term with 3 curl-free inputs
\[ \int_t^t \int_{\mathbb{R}^3} (P_k w^{(2)}(\partial_t A^{cf}, \partial_t A^{cf}, s), \chi \epsilon \sigma \Delta \partial_t A^{cf}) \, dx \, dt. \quad (10.13) \]

10.1.1. The main term [10.5].

We proceed to specifying the null form, recalling [9.9]. From the identities [2.7] and Lemma [9.2] one can write
\[ P_k \mathbf{PW}(\mathbf{P} \partial_t A_{k_1}, \partial_x \partial_t A_{k_2}, -2 \partial_t \partial_t A^{df}_{k_2, x}) \quad (10.14) \]

as
\[ \left\langle s 2^k \right\rangle^{-10} \left\langle s^{-1} 2^{-2k_{\max}} \right\rangle^{-1} 2^{-2k_{\max}} P_k O N(\partial_t A^{df}_{k_1}, \partial_t A^{df}_{k_2}) \quad (10.15) \]

where \( \mathcal{O} \) is disposable.

Since we assume the conditions in Proposition [4.4] part (2) we have \( I \partial_t A^{df} \in \epsilon X^{0, \frac{3}{4} +} \). Additionally, from Lemma [2.4]
\[ \chi \epsilon \sigma \Delta \partial_t A^{df} \in \epsilon \left\langle \frac{2^k}{N} \right\rangle^{1-\sigma} X^{0, \frac{3}{4} +} \]

We apply Prop. [10.2] to bound the contribution of [10.5] to the LHS of [10.1] by
\[ (N^2 s)^{1-\sigma} 2^{-\left(\frac{2}{3} - \frac{1}{2}\right) k_{\max}} \left\langle s \left\langle 2^k \right\rangle^{-10} \left\langle s^{-1} 2^{-2k_{\max}} \right\rangle^{-1} \left\langle \frac{2^k}{N} \right\rangle^{1-\sigma} \left\langle \frac{2^k}{N} \right\rangle^{1-\sigma} \varepsilon^3 \]

After summation in \( k_1, k_2, k \) the result is \( \lesssim \) RHS of [10.1].

We next estimate the terms [10.7] - [10.13], beginning by decomposing them into
\[ (N^2 s)^{1-\sigma} \int_t^t \int_{\mathbb{R}^3} (P_k \mathbf{W}(\partial_t A^{(1)}_{k_1}, \nabla_x \partial_t A^{(2)}_{k_2}, s), \chi \epsilon \sigma \Delta \partial_t A^{(3)}_{k}) \, dx \, dt \quad (10.16) \]

10.1.2. The terms [10.10] - [10.13] with low-high inputs in \( w \).

Here we consider the contributions of [10.16] when \( 2^k \approx 2^k_{\max} \) and at most one of \( A^{(1)}, A^{(2)}, A^{(3)} \) is \( A^{df} \). By Lemma [9.2] and translation-invariance, without loss of generality it suffices to bound
\[ 2^{-k} \int_t^t \int_{\mathbb{R}^3} \partial_t A^{(1)}_{k_1} \cdot \partial_t A^{(2)}_{k_2} P_k (\chi \epsilon \sigma \Delta \partial_t A^{(3)}_{k}) \, dx \, dt \times \left\langle s \left\langle 2^k \right\rangle^{-10} \left\langle s^{-1} 2^{-2k} \right\rangle^{-1} \right( N^2 s)^{1-\sigma} \quad (10.17) \]

when \( 2^{k_1} \lesssim 2^{k_2} \approx 2^k \). We bound this term as \( L^2 L^\infty \times L^\infty L^2 \times L^\infty L^2 \), the low frequency in \( L^2 L^\infty \) using [4.12], [9.18], [9.16], [9.17]. After summing in \( k_1 \), and then in \( k \) using the \( \left\langle s \left\langle 2^k \right\rangle^{-10} \left\langle s^{-1} 2^{-2k} \right\rangle^{-1} \) factors one obtains a bound of \( s^{\frac{1}{2}} (Ns^2)^{-3(1-\sigma)} \).

10.1.3. The terms [10.10] - [10.13] with high-high inputs in \( w \).

Here we consider the contributions of [10.16] when \( 2^k \ll 2^{k_1} \approx 2^{k_2} \) and at most one of \( A^{(1)}, A^{(2)}, A^{(3)} \) is \( A^{df} \). By Lemma [9.2] without loss of generality it suffices to bound
\[ 2^{-2k} \int_t^t \int_{\mathbb{R}^3} P_k (\partial_t A^{(1)}_{k_1} \cdot \partial_t A^{(2)}_{k_2}) \chi \epsilon \sigma \Delta \partial_t A^{(3)}_{k} \, dx \, dt \times \left\langle s \left\langle 2^k \right\rangle^{-10} \left\langle s^{-1} 2^{-2k} \right\rangle^{-1} \right( N^2 s)^{1-\sigma} \quad (10.18) \]

We use the norms of \( L^\infty L^2, L^2 L^2 \) and \( L^2 L^\infty \) as follows: a high frequency \( \partial_t A^{cf} \) is always placed in \( L^2 L^2 \) using [11.13], the term \( \partial_t A^{df} \) (if present) is placed in \( L^2 L^\infty \) using [9.18] (otherwise place the low frequency \( \partial_t A^{cf} \) in \( L^2 L^\infty \) using [11.12]) and the remaining \( \partial_t A^{cf} \) term is placed in \( L^\infty L^2 \) by [9.16]. After summation we obtain either \( s^{\frac{1}{2}} (Ns^2)^{-4(1-\sigma)} \) or \( s^{\frac{3}{2}} (Ns^2)^{-3(1-\sigma)} \).
10.1.4. The terms \([10.7]-[10.9]\) with low-high inputs in \(w\).

Now consider the contributions of \([10.16]\) when \(2^k \asymp 2^{k_{\text{max}}}\) and two of \(A^{(1)}, A^{(2)}, A^{(3)}\) are \(A^{df}\) and one is \(A^{cf}\). Again it suffices to bound \([10.17]\) when \(2^{k_1} \lesssim 2^{k_2} \sim 2^k\). We place the \(\partial t A^{cf}\) term in \(L^2 L^2\) using \([4.13]\). In the cases of \([10.7]\) and \([10.9]\) when \(\partial t A^{df}\) is low frequency, place this term in \(L^2 L^\infty\) using \([9.18]\) and the other term in \(L^\infty L^2\). In the case of \([10.8]\) place the product of the two high frequency \(df\) terms in \(L^2 L^2\) using \([2.31]\). After summing in \(k_1\), and then in \(k\) using the \(\langle s^{2k}\rangle^{-10} \langle s^{-12-2k}\rangle^{-1}\) factors one obtains a bound of \(s^{3(1)} (Ns^{3})^{-2(1-\sigma)}\) or \(s^{1(1)}\).

10.1.5. The terms \([10.7]-[10.9]\) with high-high inputs in \(w\).

Here we consider the contributions of \([10.16]\) when \(2^k \ll 2^{k_1} \sim 2^{k_2}\) and two of \(A^{(1)}, A^{(2)}, A^{(3)}\) are \(A^{df}\) and one is \(A^{cf}\). Again it suffices to bound \([10.18]\) and we place the \(\partial t A^{cf}\) term in \(L^2 L^2\) using \([4.13]\). In the case of \([10.7]\) we place the product of the two high frequency \(df\) terms in \(L^2 L^2\) using \([2.31]\). In the cases of \([10.8], [10.9]\), place the low frequency \(\partial t A^{df}\) in \(L^2 L^\infty\) using \([9.18]\) and the high frequency \(\partial t A^{df}\) in \(L^\infty L^2\). In all cases, after summations, we obtain a bound slightly better than \(s^{1(1)} (Ns^{3})^{-2(1-\sigma)}\).

We continue with the commutator-type terms \([10.20]\) and \([10.6]\), but first let us make some preparations. Recall the commutator identity \([51]\) Lemma 2):

\[
[P_k, \chi]\hat{f} = 2^{-k} L(\nabla \chi, f)
\]

where \(L\) is a disposable operator. It’s possible to pass \(L\) bounds from \(f\) to \(L(\nabla \chi, f)\):

**Lemma 10.4.** Let \(\chi'\) be a bump function and \(L\) a disposable operator. One has

\[
\|P_k L(\chi', f)\|_{L^2} \lesssim \left\langle \frac{\langle k \rangle}{N} \right\rangle^{1-\sigma} \|f\|_{L^2}
\]

**Proof.** By translation-invariance and the integrable kernel, it suffices to prove this for \(P_k(\chi' f)\) which we separate as \(P_k(\chi' f_{<k+c})\) and \(P_k(\chi' f_{>k+c})\). Normalize \(\|f\|_{L^2} = 1\). Then

\[
\|P_k(\chi' f)\|_{L^2} \lesssim \|f_{<k+c}\|_{L^2} + \sum_{k_2 = k_1 + O(1) \geq k+c} \|\chi' f_{k_2}\|_{L^2}
\]

\[
\lesssim \left\langle \frac{2^k}{N} \right\rangle^{1-\sigma} + \sum_{k_2 \geq k+c} 2^{-5k_2} \left\langle \frac{2^{k_2}}{N} \right\rangle^{1-\sigma} \lesssim \left\langle \frac{2^k}{N} \right\rangle^{1-\sigma}.
\]

\(\square\)

**Lemma 10.5.** Let \(f_k\) be a function localized at frequencies \(\langle \xi \rangle \sim 2^k\), \(\chi\) be a bump function and let \(m(\xi)\) be a homogeneous multiplier of degree zero. Then

\[
\|P_k[m(D), \chi] f_k\|_{L^2} \lesssim 2^{-k} \|f_k\|_{L^2}.
\]

**Proof.** Note that

\[
P_k[m(D), \chi] f_k = P_k[m(D), \chi_{<k-3}] f_k + P_k[m(D), \chi_{k-3,k+3}] f_k
\]

For the latter term, bounds the two terms of this commutator separately as

\[
\|P_k[m(D), \chi_{k-3,k+3}] f_k\|_{L^2} \lesssim \|P_k[m(D), \chi] f_k\|_{L^2} \lesssim 2^{-10k} \|f_k\|_{L^2}.
\]

For the former term, denoting \(T f_k = [m(D), \chi_{<k-3}] f_k\),

\[
T f_k(\xi) = \int [m(\xi) - m(\xi - \eta)] \hat{\chi}_{<k-3}(\eta) \hat{f_k}(\xi - \eta) \, d\eta
\]

\[= \int_0^1 \int_{\eta \in 2^k} \nabla m(\xi - t\eta) \eta \hat{\chi}_{<k-3}(\eta) \hat{f_k}(\xi - \eta) \, d\eta \, dt
\]

51
Apply $P_k$ and take the $L^2_x$ norm, use Minkowski’s inequality, take absolute values, bound $|\nabla m(\xi - \eta)| \lesssim 2^{-k}$ and use the integrability of $|\eta| \hat{\chi}_{<k-3}(\eta)$. 

10.1.6. Proof of the first commutator bound:

$$\left( N^2 s \right)^{1-\sigma} \sum_k \left| \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k w^{(2)}(s), \tilde{P}_k([\chi, P_k] e^{s\Delta} \partial_t A_\ell)) \, dx \, dt \right| \lesssim \frac{s^{\frac{4}{N} - 1}}{N^2 s^{(1-\sigma)}} \epsilon^2$$  \hspace{1cm} (10.20)

We bound $P_k w^{(2)}(s)$ in $L_t^2 L_x^2$ using (9.11). The term $\tilde{P}_k([\chi, P_k] e^{s\Delta} \partial_t A_\ell)$ is viewed as

$$2^{-k} \tilde{P}_k L(\nabla \chi, e^{s\Delta} \partial_t A)$$

using (10.19) which we bound using Lemma 10.4 as

$$2^{-k} \| \tilde{P}_k L(\nabla \chi, e^{s\Delta} \partial_t A) \|_{L_t^\infty L_x^2} \lesssim \epsilon 2^{-k} \left\langle \frac{2^k}{N} \right\rangle^{1-\sigma}.$$  

After summing in $k$ we bound the product by the RHS of (10.20).

10.1.7. The second commutator bound (10.6). We prove that

$$\left( N^2 s \right)^{1-\sigma} \sum_k \left| \int_{t_0}^{t_1} \int_{\mathbb{R}^3} (P_k P^\perp w^{(2)}(\partial_t A^{df}, \partial_t A^{df}, s), \tilde{P}_k (\chi e^{s\Delta} \partial_t A_k)) \, dx \, dt \right|$$

is $\lesssim$ the RHS of (10.20). Since $P^\perp \partial_t A^{df} = 0$ one can write

$$P^\perp \chi e^{s\Delta} \partial_t A_k = [P^\perp, \chi] e^{s\Delta} \partial_t A_k$$

By Lemma 10.5

$$\| \tilde{P}_k [P^\perp, \chi] e^{s\Delta} \partial_t A_k \|_{L_t^\infty L_x^2} \lesssim \epsilon 2^{-k} \left\langle \frac{2^k}{N} \right\rangle^{1-\sigma}$$

By Lemma 9.11 and Remark 9.4 we obtain an $L_t^2 L_x^2$ bound for $P_k w^{(2)}(\partial_t A^{df}, \partial_t A^{df}, s)$. The numerology is the same as in the proof of (10.20). 

10.2. Proof of Proposition 10.2.

Integrating by parts if needed, one may assume without loss of generality that $k_1 = \min(k_1, k_2, k_3)$. One may also assume that the Fourier transforms on $\mathbb{R} \times \mathbb{R}^3$ of the inputs are real and non-negative, as one then estimates

$$\int_{\mathbb{R} \times \mathbb{R}^3} 1_{[t_0, t_1]}(t) \mathcal{N}(f_{k_1}, g_{k_2}) h_{k_3} \, dx \, dt$$

using Plancharel’s identity by

$$\int \frac{|\xi_1 \wedge \xi_2|}{|\xi_1|} \langle r_1 + r_2 + r_3 \rangle \hat{f}_{k_1}(r_1, \xi_1) \hat{g}_{k_2}(r_2, \xi_2) \hat{h}_{k_3}(r_3, \xi_3)$$

since $1_{[t_0, t_1]}(\tau) \lesssim \langle \tau \rangle^{-1}$. The integral is taken over the subset of

$$(\tau_1, \xi_1, \tau_2, \xi_2, \tau_3, \xi_3) \in (\mathbb{R} \times \mathbb{R}^3)^3$$

with $\xi_1 + \xi_2 + \xi_3 = 0$. Denote the modulations by $\lambda_i := \langle |\xi_i| - |\tau_i| \rangle$.

Denoting $a_\alpha(t)$ the Fourier transform of $\langle \tau \rangle^{-\alpha}$, we recall

$$a_\alpha(t) \in L^p_t, \quad \forall \ p < (1 - \alpha)^{-1}, \ \alpha \in (0, 1)$$  \hspace{1cm} (10.21)

from [22] Lemma 15.2. We also recall

$$|\xi_1 \wedge \xi_2| \lesssim 2^{\frac{3}{2} (1 + \frac{\alpha}{2})} 2^{\frac{3}{4} (1 + \frac{\alpha}{2})} (\tau_1 + \tau_2 + \tau_3 + \lambda_1 + \lambda_2 + \lambda_3)^{\frac{1}{2}}$$  \hspace{1cm} (10.22)
from \[28\ Prop. 1\] or \[22\ Lemma 15.1\].

Then it suffices to consider three cases and bound \(10.23\), \(10.24\), \(10.25\) as follows.

1. Bound
\[
2^{k_2 + \frac{3}{2}} \int \frac{1}{\langle \tau_1 + \tau_2 + \tau_3 \rangle^{\frac{3}{2}}} \hat{f}_{k_1}(\tau_1, \xi_1) \hat{g}_{k_2}(\tau_2, \xi_2) \hat{h}_{k_3}(\tau_3, \xi_3) \tag{10.23}
\]
by the RHS of \(10.2\). Undoing the Fourier transform one proves
\[
2^{k_2 + \frac{3}{2}} \left| \int_{\mathbb{R} \times \mathbb{R}^3} a_1(t) f_{k_1} g_{k_2} h_{k_3} \, dx \, dt \right| \lesssim \text{RHS}(10.2)
\]
which is bounded as \(L_t^2 L_x^\infty \times L_t^{2+} L_x^\infty \times L_t^\infty L_x^2 \times L_t^\infty L_x^2\) using \(10.21\) and Strichartz estimates \(2.30\).

2. When \(\lambda_1\) dominates in \(10.22\), by dyadic decomposition and Cauchy-Schwarz, it suffices to integrate on the subregion \(\lambda_2, \lambda_3 \lesssim 2^j \simeq \lambda_1\) and prove
\[
2^{-\frac{3}{2}} \int \frac{1}{\langle \tau_1 + \tau_2 + \tau_3 \rangle} \hat{u}_{k_1, j}(\tau_1, \xi_1) \hat{g}_{k_2}(\tau_2, \xi_2) \hat{h}_{k_3}(\tau_3, \xi_3) \lesssim \tag{10.24}
\]
where we denote \(\hat{u}_{k_1, j} = \lambda_1^{\frac{3}{2}} \hat{f}_{k_1}\). It suffices to prove
\[
\left| \int_{\mathbb{R} \times \mathbb{R}^3} a_1(t) u_{k_1, j} D^{-\frac{3}{2}}(g_{k_2} h_{k_3}) \, dx \, dt \right| \lesssim \text{RHS}(10.24)
\]
Use \(10.21\) to place \(a_1(t)\) in \(L_t^\infty L_x^\infty\), the embedding \(2.27\) for \(u_{k_1, j}\) and the bilinear Strichartz estimate \(2.31\) for the product.

3. In the remaining case, we can assume without loss of generality that \(\lambda_2\) dominates and integrate on the subregion \(\lambda_1, \lambda_3 \lesssim 2^j \simeq \lambda_2\), proving
\[
\int \frac{1}{\langle \tau_1 + \tau_2 + \tau_3 \rangle} f_{k_1}(\xi_1) \hat{v}_{k_2, j}(\tau_2, \xi_2) \hat{h}_{k_3}(\tau_3, \xi_3) \lesssim \tag{10.25}
\]
where we denote \(\hat{v}_{k_2, j} = \lambda_2^{\frac{3}{2}} \hat{g}_{k_2}\). Undoing the Fourier transform this is reduced to
\[
\left| \int_{\mathbb{R} \times \mathbb{R}^3} a_1(t) f_{k_1} v_{k_2, j} h_{k_3} \, dx \, dt \right| \lesssim \text{RHS}(10.25)
\]
which is bounded as \(L_t^\infty L_x^\infty \times L_t^4 L_x^4 \times L_t^2 L_x^2 \times L_t^4 L_x^4\) using \(10.21\) and Strichartz estimates \(2.30\).

**Appendix A. Modified energy for Maxwell-Klein-Gordon**

In this appendix we formulate the gauge invariant modified energy for the Maxwell-Klein-Gordon equation, we show that the same favorable balance of derivatives as for Yang-Mills is present when it is differentiated in time (see Remark \(1.14\)) and we sketch how this can be used to increase the range of regularities for which global well-posedness is known to hold, from \(\sigma > \frac{\sqrt{2}}{2}\) in \(22\) to \(\sigma > \frac{5}{6}\).

The Yang-Mills equations in the case of the abelian group \(G = SO(2) = U(1)\), \(\mathfrak{g} = \mathfrak{so}(2)\) correspond to Maxwell’s equations, where \(A_\alpha\) can be viewed as a real one-form \(A_\alpha : \mathbb{R}^{1+3} \to \mathbb{R}\).
Coupling with a complex field $\phi : \mathbb{R}^{1+3} \to \mathbb{C}$ one obtains the (massless) Maxwell-Klein-Gordon system \cite{26, 36, 22}, as the Euler-Lagrange equations for the Lagrangian

$$
\frac{1}{2} \int_{\mathbb{R}^{1+3}} D_\alpha \phi \overline{D_\beta \phi} + \frac{1}{2} F_{\alpha \beta} F_{\alpha \beta} \, dx \, dt
$$

Here one denotes the covariant derivative $D_\alpha$ and the curvature $F_{\alpha \beta}$ by

$$
D_\alpha \phi := (\partial_\alpha + i A_\alpha) \phi, \quad F_{\alpha \beta} := \partial_\alpha A_\beta - \partial_\beta A_\alpha.
$$

The system becomes

$$\partial^\beta F_{\alpha \beta} = \text{Im}(\phi \overline{D_\alpha \phi})$$

$$\Box_A \phi = 0.$$  \hspace{1cm} \text{(MKG)}

where we denote $\Box_A := D^\alpha D_\alpha$. The gauge invariance takes the form

$$\phi \mapsto e^{i \chi} \phi, \quad A_\alpha \mapsto A_\alpha - \partial_\alpha \chi, \quad \chi : \mathbb{R}^{1+3} \to \mathbb{R}
$$

while the energy is

$$H[A, \phi](t) := \frac{1}{2} \int_{\mathbb{R}^3} \sum_{\alpha < \beta} |F_{\alpha \beta}|^2 + \sum_{\alpha} |D_\alpha \phi|^2 \, dx.$$  \hspace{1cm} \text{(A.1)}

Initial data sets consist of $(A_i, E_i, \phi_0, \phi_1) = (A_i, F_{0i}, \phi_0, \partial_t \phi)(t = 0)$ satisfying the constraint equation $\partial^\alpha E_\alpha = \text{Im}(\phi_0 \overline{\phi_1})$.

In \cite{22} Keel, Roy and Tao prove global well-posedness in the (global) Coulomb gauge $\partial^\alpha A_\alpha = 0$. In this gauge, $A_0$ is determined from the initial data.

**Theorem A.1** \cite{22}. Let $\sigma > \sqrt{3}/2$. The Maxwell-Klein-Gordon equation in the Coulomb gauge is globally well-posed for initial data in $H^\sigma \times H^{\sigma-1} \times H^{\sigma} \times H^{\sigma-1}$.

After reducing to global smooth solutions, rescaling and choosing $N$ such that $H[IA, I\phi](0) \leq 1$, their proof proceeds by proving the almost conservation law

$$H[I\Phi](t) = H[I\Phi](t_0) + O(N^{1/2 - \sigma})$$  \hspace{1cm} \text{(A.2)}

under the assumption $H[I\Phi](t_0) \leq 2$, where one denotes $\Phi = (A_0, A_s, \phi)$.

The spacetime control on small intervals is $\nabla_{t,x}(IA_x, I\phi) \in X^{0,\sigma-} + L^\infty L^3$ where the latter part consists of low frequencies.

For arbitrary fields \cite{22}[Lemma 11.1] writes the differentiated Hamiltonian as

$$\frac{d}{dt} H[I\Phi](t) = -\text{Re} \int_{\mathbb{R}^3} D_t \phi \Box A \phi + F_{\mu 0}(\partial^\alpha F_{\alpha \mu} + \text{Im}(\phi \overline{D_\mu \phi})) \, dx$$  \hspace{1cm} \text{(A.3)}

The dominant part of $H[I\Phi](t_1) - H[I\Phi](t_0)$ is proved in \cite{22} (106),(124) to be (1.38).

We now formulate the invariant modified energy for (MKG).

The gradient flow for the functional

$$\frac{1}{2} \int_{\mathbb{R}^3} \sum_{i<j} |F_{ij}|^2 + \sum_i |D_i \phi|^2 \, dx.
$$

gives rise to the equations $\partial_s A_\alpha = \partial^\alpha F_{\ell \alpha} + \text{Im}(\phi \overline{D_\alpha \phi})$ and $\partial_s \phi = D^\alpha D_\alpha \phi$. Adding temporal and parabolic connections $A_0, A_s$ and writing the system in a gauge-invariant way one obtains a Maxwell-Klein-Gordon Heat Flow

$$F_{s \alpha} = \partial^\alpha F_{\ell \alpha} + \text{Im}(\phi \overline{D_\alpha \phi})$$

$$(D_s - \Delta_A) \phi = 0.$$  \hspace{1cm} \text{(MKG-HF)}
The extent of failure of $A_\alpha(s), \phi(s)$ to satisfy (MKG) for $s > 0$ is measured by the tension fields

$$ v := \Box_A \phi, \quad w_\beta := \partial^\alpha F_{\alpha\beta} + \text{Im}(\phi \overline{D_\beta \phi}). $$

Up to lower order terms, the main components of the parabolic equations satisfied by the tension fields can be written as

$$ (\partial_s - \Delta)P_j w = -2P_j \text{Im}(\partial_t \phi \partial_s \overline{\partial_t \phi}) + \ldots $$

$$ (D_s - \Delta_A)v = 4i \partial_t P_j A \partial_j \partial_t \phi + \ldots $$

(A.4)

The dots consist of trilinear terms or terms with higher regularity, such as those involving $A^{cf}$ or $A_0$. Assuming deTurck’s gauge $A_s = \partial^t A_t$, expanding and decomposing (A.4) similarly to Section 9 one obtains the leading terms

$$ P_j w(s) = -2P_j \text{Im}(\partial_t \phi \partial_s \overline{\partial_t \phi})(s) + \ldots $$

$$ v(s) = 4i \text{W}(\partial_t P_j A, \partial_j \partial_t \phi)(s) + \ldots $$

(A.5)

where $\text{W}$ is defined by (9.10) and the inputs are taken at $s = 0$.

For a regular global solution of (MKG) we define the modified energy similarly to (1.6). Denote the energies $\mathcal{H}$ defined by $A_\alpha(t, \cdot, s), \phi(t, \cdot, s)$ by $H(t, s)$. Then, for $s_0 = N^{-2}$, let

$$ \mathcal{H}(t) := \sup_{s \in [0, s_0]} (N^2 s)^{1-\sigma} H(t, s) + \int_0^{s_0} (N^2 s)^{1-\sigma} H(t, s) \frac{ds}{s} $$

(A.6)

The goal is to have $|\mathcal{H}(t) - \mathcal{H}(t_0)| \ll N^{-\frac{1}{2}+}$ like in Proposition 11.7 as this can be used to obtain global well-posedness for (MKG) as in Section 1.4, with the same numerology. The $N^{-\frac{1}{2}+}$ bound compared to (A.2) reduces the minimum initial data regularity from $\sigma > \frac{s^2}{2}$ to $\sigma > \frac{5}{6}$.

Regarding the choices of gauge there are several options and we avoid making a definitive choice. One may set the temporal gauge $A_0 = 0$. Working with local gauges and extensions like we did for Yang-Mills is probably not necessary, but can be a natural choice in view of potential extensions to Yang-Mills-Higgs (which generalizes both (YM) and (MKG)). This also avoids some low frequency technicalities. On the other hand, after fixing the interval $[t_0, t_1]$, one could set the global Coulomb gauge at $s = s_0, t = t_0$ and obtain an $\mathcal{H}^1 + N^{1-\sigma} \mathcal{H}^\sigma$ bound at $s = 0$ from the caloric condition like in Section 7 but globally instead of locally.

The key spacetime bound remains $\nabla_{t,x} (IA^C_{Df}, I\phi) \in \mathcal{X}^{0, \frac{4}{7}+}$, perhaps with a low frequency component.

To estimate the variation of the modified energy one writes similarly to (1.21), using (A.3)

$$ H(t_1, s) - H(t_0, s) = -\text{Re} \int_{t_0}^{t_1} \int_{\mathbb{R}^3} D_t \phi \overline{\phi} + F_{j0} w_j \, dx \, dt $$

(A.7)

As (10.3), (10.4) show for quadrilinear terms and (10.7)-(10.13) show for terms involving $A^{cf}$, in the case of temporal gauge + DeTurck gauge, all these terms can be fairly easily bounded using spacetime Lebesgue norms. In the case of the Coulomb gauge, sections 12 and 14 in [22] show that such terms are not problematic.

Up to error terms, (A.7) is written, from (A.5) and the identities (2.7), as

$$ \text{Re} \int_{t_0}^{t_1} \int_{\mathbb{R}^3} 4i \partial_t \phi \text{W}(\partial_t A^{cf}, \partial_t \phi)(s) + 2\partial_t A^{cf} \text{Im} \text{W}(\partial_t \phi, \partial_t \phi)(s) \, dx \, dt + \ldots $$

(A.8)

Schematically, these terms are exactly like (10.5) together with (10.15) and are estimated using Proposition 10.2.
References

[1] I. Bejenaru, A. D. Ionescu, C. E. Kenig, and D. Tataru. Global Schrödinger maps in dimensions \( d \geq 2 \): small data in the critical Sobolev spaces. *Annals of Mathematics*, pages 1443–1506, 2011.

[2] J. Bourgain. *New global well-posedness results for nonlinear Schrödinger equations*. AMS Publications, 1999.

[3] Y. Choquet-Bruhat and D. Christodoulou. Existence of global solutions of the Yang-Mills, Higgs and spinor field equations in \( 3+1 \) dimensions. In *Annales scientifiques de l'École Normale Supérieure*, volume 14, pages 481–506, 1981.

[4] P. T. Chruściel and J. Shatah. Global existence of solutions of the Yang-Mills equations on globally hyperbolic four dimensional Lorentzian manifolds. *Asian Journal of Mathematics*, 1(3):530–548, 1997.

[5] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Global well-posedness for KdV in Sobolev spaces of negative index. *Electron. J. Diff. Eqns.*, 2001.

[6] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Global well-posedness for Schrödinger equations with derivative. *SIAM Journal on Mathematical Analysis*, 33(3):649–669, 2001.

[7] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Almost conservation laws and global rough solutions to a nonlinear Schrödinger equation. *Math. Res. Letters*, 9:659 – 682, 2002.

[8] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Multilinear estimates for periodic KdV equations, and applications. *Journal of Functional Analysis*, 211(1):173–218, 2004.

[9] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Resonant decompositions and the I-method for the cubic nonlinear Schrödinger equation on \( \mathbb{R}^2 \). *Discrete and Continuous Dynamical Systems*, 21(3):665–686, 2008.

[10] S. Czimek. An extension procedure for the constraint equations. *Annals of PDE*, 4(1):1–130, 2018.

[11] P. D’Ancona, D. Foschi, and S. Selberg. Atlas of products for wave-Sobolev spaces on \( \mathbb{R}^{1+3} \). *Transactions of the American Mathematical Society*, 364(1):31–63, 2012.

[12] D. M. DeTurck. Deforming metrics in the direction of their Ricci tensors. *Journal of Differential Geometry*, 18(1):157–162, 1983.

[13] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhiker’s guide to the fractional Sobolev spaces. *Bulletin des sciences mathématiques*, 136(5):521–573, 2012.

[14] B. Dodson. Global well-posedness and scattering for the defocusing, cubic nonlinear Schrödinger equation when \( n = 3 \) via a linear-nonlinear decomposition. *Discrete & Continuous Dynamical Systems*, 33(5):1905, 2013.

[15] S. K. Donaldson. Anti self-dual Yang-Mills connections over complex algebraic surfaces and stable vector bundles. *Proceedings of the London Mathematical Society*, 3(1):1–26, 1985.

[16] D. M. Eardley and V. Moncrief. The global existence of Yang-Mills-Higgs fields in 4-dimensional Minkowski space I and II. *Communications in Mathematical Physics*, 83(2):171–191, 1982.

[17] V. N. Gribov. Quantization of non-abelian gauge theories. *Nuclear Physics B*, 139(1-2):1–19, 1978.

[18] A. Grünrock. *New applications of the Fourier restriction norm method to wellposedness problems for nonlinear evolution equations*. PhD thesis, Universität Wuppertal, 2002.

[19] Z. Hani. Global well-posedness of the cubic nonlinear schrödinger equation on closed manifolds. *Communications in Partial Differential Equations*, 37(7):1156–1236, 2012.

[20] L. Kapitanski. Weak and yet weaker solutions of semilinear wave equations. *Communications in Partial Differential Equations*, 19(9-10):1629–1676, 1994.

[21] M. Keel. Global existence for the critical power Yang-Mills-Higgs equations. Princeton University, 1995.

[22] M. Keel, T. Roy, and T. Tao. Global well-posedness of the Maxwell-Klein-Gordon equation below the energy norm. *Discrete & Continuous Dynamical Systems*, 30(3):573, 2011.

[23] M. Keel and T. Tao. Local and global well-posedness of wave maps on \( \mathbb{R}^{1+1} \) for rough data. *International Mathematics Research Notices*, 1998(21):1117–1156, 01 1998.

[24] C. E. Kenig, G. Ponce, and L. Vega. Global well-posedness for semi-linear wave equations. *Communications in partial differential equations*, 25(9-10):1741–1752, 2000.

[25] S. Klainerman. On the regularity of classical field theories in Minkowski space-time \( \mathbb{R}^{3+1} \). In *Nonlinear Partial Differential Equations in Geometry and Physics*, pages 29–69. Springer, 1997.

[26] S. Klainerman and M. Machedon. On the Maxwell-Klein-Gordon equation with finite energy. *Duke Mathematical Journal*, 74(1):19–44, 1994.

[27] S. Klainerman and M. Machedon. Finite energy solutions of the Yang-Mills equations in \( \mathbb{R}^{3+1} \). *Annals of Mathematics*, pages 39–119, 1995.

[28] S. Klainerman and M. Machedon. Estimates for null forms and the spaces \( h_{s,δ} \). *International Mathematics Research Notices*, 1996(17):853–865, 1996.

[29] S. Klainerman and M. Machedon. On the optimal local regularity for gauge field theories. *Differential and Integral Equations*, 10(6):1019–1030, 1997.
[30] S. Klainerman and S. Selberg. Bilinear estimates and applications to nonlinear wave equations. *Communications in Contemporary Mathematics*, 4(02):223–295, 2002.

[31] S. Klainerman and D. Tataru. On the optimal local regularity for the Yang-Mills equations in $\mathbb{R}^{4+1}$. *Journal of the American Mathematical Society*, 12(1):93–116, 1999.

[32] J. Krieger and J. Sterbenz. Global regularity for the Yang-Mills equations on high dimensional Minkowski space, volume 223. American Mathematical Soc., 2013.

[33] J. Krieger and D. Tataru. Global well-posedness for the Yang-Mills equation in $4 + 1$ dimensions. small energy. *Annals of Mathematics*, 185(3):831–893, 2017.

[34] H. Lindblad. A sharp counterexample to the local existence of low-regularity solutions to nonlinear wave equations. *Duke Mathematical Journal*, 72(2):503–539, 1993.

[35] H. Lindblad and C. D. Sogge. On existence and scattering with minimal regularity for semilinear wave equations. *Journal of Functional Analysis*, 130(2):357–426, 1995.

[36] M. Machedon and J. Sterbenz. Almost optimal local well-posedness for the (3+1)-dimensional Maxwell–Klein–Gordon equations. *Journal of the American Mathematical Society*, 17(2):297–359, 2004.

[37] S.-J. Oh. Gauge choice for the Yang–Mills equations using the Yang–Mills heat flow and local well-posedness in $H^1$. *Journal of Hyperbolic Differential Equations*, 11(01):1–108, 2014.

[38] S.-J. Oh. Finite energy global well-posedness of the Yang–Mills equations on $\mathbb{R}^{3+1}$. An approach using the Yang–Mills heat flow. *Duke Mathematical Journal*, 164(9):1669–1732, 2015.

[39] S.-J. Oh and D. Tataru. The Yang–Mills heat flow and the caloric gauge. arXiv preprint arXiv:1709.08599., 2017.

[40] S.-J. Oh and D. Tataru. The hyperbolic Yang–Mills equation for connections in an arbitrary topological class. *Communications in Mathematical Physics*, 365(2):685–739, 2019.

[41] S.-J. Oh and D. Tataru. The hyperbolic Yang–Mills equation in the caloric gauge: local well-posedness and control of energy-dispersed solutions. *Pure and Applied Analysis*, 2(2):233–384, 2020.

[42] S.-J. Oh and D. Tataru. The threshold conjecture for the energy critical hyperbolic Yang–Mills equation. *Annals of Mathematics*, 194(2):393–473, 2021.

[43] J. Rade. On the Yang-Mills heat equation in two and three dimensions. *J. reine angew. Math.*, pages 123–163, 1992.

[44] T. Roy. Global analysis of the defocusing cubic wave equation in dimension 3. PhD thesis, UCLA, 2008.

[45] T. Roy. Adapted linear-nonlinear decomposition and global well-posedness for solutions to the defocusing cubic wave equation on $\mathbb{R}^3$. *Discrete and Continuous Dynamical Systems*, 24(4):1307–1323, 2009.

[46] I. Segal. The Cauchy problem for the Yang-Mills equations. *Journal of functional analysis*, 33(2):175–194, 1979.

[47] S. Selberg. Multilinear space-time estimates and applications to local existence theory for nonlinear wave equations. PhD thesis, Princeton University, 1999.

[48] S. Selberg and A. Tesfahun. Null structure and local well-posedness in the energy class for the Yang–Mills equations in Lorenz gauge. *Journal of the European Mathematical Society*, 18(8):1729–1752, 2016.

[49] T. Tao. Gauges for the Schrödinger map. https://www.math.ucla.edu/~tao/preprints/Expository/.

[50] T. Tao. Local well-posedness of the Yang-Mills equation in the temporal gauge below the energy norm. Version from: https://arxiv.org/abs/math/0005064., 2001.

[51] T. Tao. Global regularity of Wave Maps II. small energy in two dimensions. *Communications in Mathematical Physics*, 224(2):443–544, 2001.

[52] T. Tao. Multilinear weighted convolution of $L^2$ functions, and applications to nonlinear dispersive equations. *American Journal of Mathematics*, 123(5):839–908, 2001.

[53] T. Tao. Geometric renormalization of large energy wave maps. *Journées équations aux dérivées partielles*, pages 1–32, 2004.

[54] T. Tao. Global regularity of wave maps iv. absence of stationary or self-similar solutions in the energy class. *arXiv preprint arXiv:0806.3592*, 2008.

[55] A. Tesfahun. Local well-posedness of Yang–Mills equations in Lorenz gauge below the energy norm. *Nonlinear Differential Equations and Applications NoDEA*, 22(4):849–875, 2015.

[56] K. K. Uhlenbeck. Connections with $L^p$ bounds on curvature. *Communications in Mathematical Physics*, 83(1):31–42, 1982.

[57] A. Waldron. Long-time existence for Yang–Mills flow. *Inventiones mathematicae*, 217(3):1069–1147, 2019.

Email address: gcd2006@gmail.com