Available online at www.sciencedirect.com

ScienceDirect

Ann. I. H. Poincaré – AN 35 (2018) 1655–1686

ANNALES DE L'INSTITUT HENRI POINCARÉ
ANALYSE NON LINÉAIRE

www.elsevier.com/locate/anihpc

Singularity formation of the Yang–Mills Flow

Casey Kelleher [a,*], Jeffrey Streets [b]

[a] *Fine Hall, Princeton University, Princeton, NJ 08544, United States*
[b] *Rowland Hall, University of California, Irvine, CA 92617, United States*

Received 19 February 2016; accepted 5 January 2018

Abstract

We study singularity structure of Yang–Mills flow in dimensions $n \geq 4$. First we obtain a description of the singular set in terms of concentration for a localized entropy quantity, which leads to an estimate of its Hausdorff dimension. We develop a theory of tangent measures for the flow, which leads to a stratification of the singular set. By a refined blowup analysis we obtain Yang–Mills connections or solitons as blowup limits at any point in the singular set.

Keywords: Yang–Mills; geometric flows; singularity analysis

1. Introduction

Given (M^n, g) a compact Riemannian manifold and $E \to M$ a vector bundle, a one parameter family of connections ∇_t on E is a solution to Yang–Mills flow if

$$\frac{\partial \nabla_t}{\partial t} = -D^*_{\nabla_t} F_{\nabla_t}.$$

This is the negative gradient flow for the Yang–Mills energy, and is a natural tool for investigating its variational structure. Global existence and convergence of the flow in dimensions $n = 2, 3$ was established in [16]. Finite time singularities in dimension $n = 4$ can only occur via energy concentration, as established in [18]. More recently this result has been refined in [5,21] to show concentration of the self-dual and antiself-dual energies. Preliminary investigations into Yang–Mills flow in higher dimensions have been made in [7,15,22].

In this paper we establish structure theorems on the singular set for Yang–Mills flow in dimensions $n \geq 4$. Our results are inspired generally by results of harmonic map flow, specifically [12–14]. The first main result is a weak compactness theorem for solutions to Yang–Mills flow which includes a rough description of the singular set of a sequence of solutions. A similar result for harmonic map flow was established in [12]. Moreover, a related result on

* Corresponding author.
E-mail addresses: ckelleher@princeton.edu (C. Kelleher), jstreets@uci.edu (J. Streets).

https://doi.org/10.1016/j.anihpc.2018.01.006
0294-1449/

the singularity formation at infinity for a global solution of Yang–Mills flow was established in [9]. We include a rough statement here, see Theorem 4.1 for the precise statement.

Theorem 1.1. *Fix $n \geq 4$ and let $E \to (M^n, g)$ be a vector bundle over a closed Riemannian manifold. Weak $H^{1,2}$ limits of sequences of smooth solutions to Yang–Mills flow are weak solutions to Yang–Mills flow which are smooth outside of a closed set Σ of locally finite $(n-2)$-dimensional parabolic Hausdorff measure.*

The first key ingredients of the proof are localized entropy monotonicities for the Yang–Mills flow, defined in [9], together with a low-entropy regularity theorem [9]. Fairly general methods allow for the existence of the weak limit claimed in Theorem 1.1, and the entropy monotonicities are the key to showing that the singular set is small enough to ensure that the weak limit is a weak solution to Yang–Mills flow. The arguments are closely related to those appearing in [9,12,19].

The second main result is a stratification of the singular set. This involves investigating tangent measures associated to solutions of Yang–Mills flow. In particular we are able to establish the existence of a density for these measures together with certain parabolic scaling invariance properties. One immediate consequence is that we can apply the general results of [23] to obtain a stratification of the singular set. See §5 for the relevant definitions.

Theorem 1.2. *For $0 \leq k \leq n-2$ let*

$$\Sigma_k := \left\{ z_0 \in \Sigma \mid \dim\left(\Theta^0\left(\mu^*, \cdot\right)\right) \leq k, \forall \mu^* \in T_{z_0}(\mu) \right\}.$$

Then $\dim_{\mathcal{P}}(\Sigma_k) \leq k$ and Σ_0 is countable.

The third main theorem characterizes the failure of strong convergence in the statement of Theorem 1.1 in terms of the bubbling off of Yang–Mills connections. Again, an analogous result for harmonic maps was established in [12]. The proof requires significant further analysis on tangent measures, leading to the existence of a refined blowup sequence which yields the Yang–Mills connection. We give a rough statement below, see Theorem 6.1 for the precise statement.

Theorem 1.3. *Fix $n \geq 4$ and let $E \to (M^n, g)$ be a vector bundle over a closed Riemannian manifold. A sequence of solutions to Yang–Mills flow converging weakly in $H^{1,2}$ either converges strongly in $H^{1,2}$, and the $(n-2)$-dimensional parabolic Hausdorff measure of Σ vanishes, or it admits a blowup limit which is a Yang–Mills connection on S^4.*

A corollary of these theorems is the existence of a either Yang–Mills connection or Yang–Mills soliton as a blowup limit of arbitrary finite time singularities. For type I singularities the existence of soliton blowup limits was established in [22], following from the entropy monotonicity for Yang–Mills flow demonstrated in [8]. The existence of soliton blowup limits for arbitrary singularities of mean curvature flow was established in [10], relying on the structure theory associated with Brakke's weak solutions. A preliminary investigation into the entropy-stability of Yang–Mills solitons was undertaken in [3] and [11]. Those results now apply to studying arbitrary finite-time singularities of Yang–Mills flow, as all admit singularity models which are either Yang–Mills connections or Yang–Mills solitons.

Corollary 1.4. *Fix $n \geq 4$ and let $E \to (M^n, g)$ be a vector bundle over a closed Riemannian manifold. Let ∇_t a smooth solution to Yang–Mills flow on $[0, T)$ such that $\limsup_{t \to T} \left|F_{\nabla_t}\right|_{C^0} = \infty$. There exist a sequence $\{(x_i, t_i, \lambda_i)\} \subset M \times [0, T) \times [0, \infty)$ such that the corresponding blowup sequence converges modulo gauge transformations to either*

(1) *A Yang–Mills connection on S^4.*
(2) *A Yang–Mills soliton.*

Acknowledgments

The first author gratefully thanks Osaka University mathematics department, in particular Toshiki Mabuchi and Ryushi Goto where much of the preliminary work was performed, for their warm hospitality. The first author also sincerely thanks Gang Tian and all of Princeton University, where much of the intermediate and final work was conducted, for providing such a friendly and productive atmosphere. The first author was supported by an NSF Graduate Research Fellowship DGE-1321846. The second author was supported by the NSF via DMS-1341836, DMS-1454854 and by the Alfred P. Sloan Foundation through a Sloan Research Fellowship.

2. Background

We will begin with a discussion of notation and conventions that are used throughout the paper. We will then provide general analytic background as well as a review of Yang–Mills flow and its key properties.

2.1. Notation and conventions

Let $(E, h) \to (M, g)$ be a vector bundle over a closed Riemannian manifold. Let $S(E)$ denote the smooth sections of E. For each point $x \in M$ choose a local orthonormal basis of TM given by $\{\partial_i\}$ with dual basis $\{e^i\}$ and a local basis for E given by $\{\mu_\alpha\}$ with dual basis $\{(\mu^*)^\alpha\}$ for the dual E^*. Let $\Lambda^p(M)$ denote the set of smooth p-forms over M and set $\Lambda^p(E) := \Lambda^p(M) \otimes S(E)$. Next set $\operatorname{End} E := E \otimes E^*$, where E^* denotes the dual space of E and take

$$\Lambda^p(\operatorname{Ad} E) := \{\omega \in \Lambda^p(\operatorname{End} E) \mid h_{\alpha\gamma}\omega_\beta^\gamma = -h_{\beta\gamma}\omega_\alpha^\gamma\}.$$

The set of all bundle metric compatible connections on E will be denoted by $\mathcal{A}_E(M)$. Given a chart containing $p \in M$ the action of a connection ∇ on E is captured by the *coefficient matrices* $\Gamma = (\Gamma_{i\alpha}^\beta e^i \otimes \mu_\beta \otimes \mu_\alpha^*)$, where

$$\nabla \mu_\beta = \Gamma_{i\beta}^\delta e^i \otimes \mu_\delta.$$

When sequences of one-parameter families of connections $\{\nabla_t^i\}$ are in play we will at times drop the explicit dependence on t and i for notational simplicity.

2.2. Weak solutions of Yang–Mills flow

We first recall here the definitions of Sobolev spaces relevant to discussing convergence of connections. Refer to ([18] §1.3) for further information. Using this we give the definition of a weak solution to Yang–Mills flow.

Definition 2.1. Fix ∇_{ref} a background connection on E. The space $H^{l,p}(\Lambda^i(\operatorname{Ad} E))$ is the completion of the space of smooth sections of $\Lambda^i(\operatorname{Ad} E)$ with respect to the norm

$$||\Upsilon||_{H^{l,p}(\Lambda^i(\operatorname{Ad} E))} := \left(\sum_{k=0}^{l} \left|\left| \nabla_{\text{ref}}^{(k)} \Upsilon \right|\right|_{L^p(\Lambda^i(\operatorname{Ad} E))}^p \right)^{1/p} < \infty.$$

We will say that a connection ∇ is of *Sobolev class* $H^{l,p}$, and write $\nabla \in H^{l,p}$, if $\nabla = \nabla_{\text{ref}} + \Upsilon$ where $\Upsilon \in H^{l,p}\left(\Lambda^1(\operatorname{Ad} E)\right)$.

Now, for a vector bundle $E \to (M, g)$ over a Riemannian manifold, recall that the *Yang–Mills energy* of a smooth connection ∇ on E with curvature F_∇ is

$$\mathcal{YM}(\nabla) := \tfrac{1}{2} \int_M |F_\nabla|^2 \, dV.$$

From this we can consider the corresponding negative gradient flow, which is easily shown to be the *Yang–Mills flow*:

$$\tfrac{\partial \nabla_t}{\partial t} = - D_{\nabla_t}^* F_{\nabla_t}.$$

With these definitions in place we can now define the notion of a weak solution to the flow.

Definition 2.2. A one-parameter family $\nabla_t = \nabla_0 + \Upsilon_t$ is a *weak solution of Yang–Mills flow* on $[0, T]$ if

$$\Upsilon_t \in L^1([0,T]; L^2(\Lambda^1(\operatorname{Ad} E))), \qquad F_{\nabla_t} \in L^\infty([0,T]; L^2(\Lambda^2(\operatorname{Ad} E))),$$

and if for all $\alpha_t \in C^\infty([0,T]; H_1^2(\Lambda^2(\operatorname{Ad} E)))$ which vanish at $t=0, t=T$, one has

$$\int_0^T \int_M \left\langle \Upsilon_t, \tfrac{\partial \alpha_t}{\partial t} \right\rangle - \left\langle F_{\nabla_t}, \nabla_t \alpha_t \right\rangle dV\, dt = 0. \tag{2.1}$$

2.3. Blowup constructions

Here we will give a discussion of the construction of blowup limits in the setting of Yang–Mills flow. First we define the fundamental scaling law.

Definition 2.3. Fix $U \subset \mathbb{R}^n$ and consider the restricted bundle $E \to U$. Suppose ∇_t is a smooth solution to Yang–Mills flow over U on $[0, T)$. Fixing a basis for E, ∇_t is described by local coefficient matrices Γ_t. Given $z_0 = (x_0, t_0) \in U \times [0, T)$ and $\lambda \in \mathbb{R}$ we define a connection ∇_t^{λ, z_0} via coefficient matrices

$$\Gamma_t^{\lambda, z_0}(x) = \lambda \Gamma_{\lambda^2 t + t_0}(\lambda x + x_0). \tag{2.2}$$

Typically the basepoint z_0 will be suppressed notationally when understood.

Now consider a sequence $\{(x_i, t_i, \lambda_i)\} \subset M \times \mathbb{R} \times [0, \infty)$ with $\lambda_i \to 0$. Assuming M is compact there exists a subsequence such that $\{x_i\} \to x_\infty \in M$. Moreover, we can pick a chart around x_∞ so that the tail of the sequence $\{x_i\}$ is contained within this chart, identified with $B_1 \subset \mathbb{R}^n$. For sufficiently large i, define a connection ∇_t^i via coefficient matrices

$$\Gamma_t^i(x) := \Gamma_t^{\lambda_i, z_i}(x).$$

We call $\{\nabla_t^i\}$ an (x_i, t_i, λ_i)*-blowup sequence*. Note the corresponding curvatures are scaled in the following manner,

$$F_{\nabla_t^i}(x) = \lambda_i^2 F_{\nabla_{\lambda_i^2 t + t_i}}(\lambda_i x + x_i). \tag{2.3}$$

Observe that the domain of ∇_t^i contains $B_{\lambda_i^{-1}}(x_i) \times [\frac{-t_i}{\lambda_i^2}, \frac{T - t_i}{\lambda_i^2}]$, so that the limiting domain is $\mathbb{R}^n \times (-\infty, 0]$. If the points are chosen as a maximal blowup sequence so that the curvatures are bounded, then these blowup solutions converge to a smooth ancient solution to Yang–Mills flow. However, in our analysis though we will be choosing very general sequences and taking weak limits.

2.4. Parabolic Hausdorff measures

For any $0 \le k \le n$ and any $\Omega \subset \mathbb{R}^n$, the k-dimensional Hausdorff measure of Ω is defined by

$$\mathcal{H}^k(\Omega) = \lim_{\delta \to 0} \mathcal{H}^k_\delta(\Omega) = \liminf_{\delta \to 0} \left\{ \sum_i r_i^k \mid \Omega \subset \bigcup_i B_{r_i}(z_i), z_i \in \Omega, r_i \le \delta \right\}.$$

This leads to the definition of *Hausdorff dimension*, i.e.

$$\dim_{\mathcal{H}}(\Omega) = \inf \{ d \ge 0 \mid \mathcal{H}^d(\Omega) = 0 \}.$$

Next, we define the *parabolic metric* ϱ on $\mathbb{R}^n \times \mathbb{R}$ given by, for $(x,t), (y,s) \in \mathbb{R}^n \times \mathbb{R}$,

$$\varrho((x,t),(y,s)) := \max \left\{ |x - y|, \sqrt{|t - s|} \right\}.$$

We can obtain the notion of parabolic Hausdorff dimension by using covers by balls with respect to this metric. In particular, for any $0 \le \ell \le n+2$ and any $\Omega \subset \mathbb{R}^n \times \mathbb{R}$, the ℓ-dimensional parabolic Hausdorff measure of Ω is given by

$$\mathcal{P}^\ell(\Omega) = \lim_{\delta\to 0} \mathcal{P}^\ell_\delta(\Omega) = \liminf_{\delta\to 0} \left\{ \sum_i r_i^\ell \mid \Omega \subset \bigcup_i P_{r_i}(z_i), z_i \in \Omega, r_i \le \delta \right\},$$

where, for $z_0 = (x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$,

$$P_r(z_0) := \left\{ z = (x,t) \in \mathbb{R}^n \times \mathbb{R} \mid |x - x_0| < r, |t - t_0| < r^2 \right\}.$$

Using this we can then define the *parabolic Hausdorff dimension*

$$\dim_{\mathcal{P}}(\Omega) := \inf\{d \ge 0 \mid \mathcal{P}^d(\Omega) = 0\}.$$

3. Monotonicity formulas

In this section we observe some energy and entropy monotonicity formulas for solutions to Yang–Mills flow which are central to the analysis below.

3.1. Energy monotonicity

Lemma 3.1. *Let ∇_t be a solution to Yang–Mills flow on $M \times [t_1, t_2]$. For any $\phi \in C^1_0(M, [0, \infty))$,*

$$\frac{1}{4} \int_M \left(\left| F_{\nabla_{t_1}} \right|^2 - \left| F_{\nabla_{t_2}} \right|^2 \right) \phi^2 \, dV = \int_{t_1}^{t_2} \int_M \left(\left| \tfrac{\partial \nabla_t}{\partial t} \right|^2 + \left\langle \tfrac{2\nabla_t \phi}{\phi} \lrcorner F_{\nabla_t}, \tfrac{\partial \nabla_t}{\partial t} \right\rangle \right) \phi^2 \, dV \, dt.$$

Proof. We differentiate and find that

$$\begin{aligned}
\frac{d}{dt}\left[\frac{1}{2} \int_M |F|^2 \phi^2 \, dV \right] &= \int_M \left\langle F, \tfrac{\partial F}{\partial t} \right\rangle \phi^2 \, dV \\
&= \int_M \left\langle F, D\left[\tfrac{\partial \nabla}{\partial t}\right] \right\rangle \phi^2 \, dV \\
&= 2 \int_M \left\langle F, \nabla\left[\tfrac{\partial \nabla}{\partial t}\right] \right\rangle \phi^2 \, dV \\
&= 2 \int_M \left\langle D^* F - 2\tfrac{\nabla \phi}{\phi} \lrcorner F, \tfrac{\partial \nabla}{\partial t} \right\rangle \phi^2 \, dV \\
&= 2 \int_M \left\langle -\tfrac{\partial \nabla}{\partial t} - 2\tfrac{\nabla \phi}{\phi} \lrcorner F, \tfrac{\partial \nabla}{\partial t} \right\rangle \phi^2 \, dV \\
&= -2 \int_M \left(\left\langle \tfrac{\partial \nabla}{\partial t}, 2\tfrac{\nabla \phi}{\phi} \lrcorner F \right\rangle + \left| \tfrac{\partial \nabla}{\partial t} \right|^2 \right) \phi^2 \, dV.
\end{aligned}$$

Integrating both sides over $[t_1, t_2]$ yields the result. □

3.2. Entropy setup and scaling laws

Let (M, g) be a Riemannian manifold. Let $\iota_M > 0$ be a lower bound for the injectivity radius of M. Note that if ∇_t is a smooth solution to Yang–Mills flow on $M \times [0, T)$, we can restrict it to any coordinate neighborhood $B_{\iota_M} \subset \mathbb{R}^n$ is the Euclidean ball in $\mathbb{R}^n$ centered at the origin. Now fix $z_0 := (x_0, t_0) \in \mathbb{R}^n \times [0, \infty)$, and define

$$G_{z_0}(x,t) = \frac{e^{-\frac{|x-x_0|^2}{4|t-t_0|}}}{(4\pi |t - t_0|)^{n/2}}.$$

We need to move this function onto the manifold M, and so we must localize. For $x_0 \in M$ we let $\mathcal{B}_{x_0}$ denote the set of *cutoff functions*, that is, all $\phi \in C_0^\infty\left(B_{\iota_M}(x_0), [0,\infty)\right)$ such that

$$\phi \in [0,1], \qquad \phi \equiv 1 \text{ on } B_{\frac{\iota_M}{2}}(x_0), \qquad \operatorname{supp}\phi \subset B_{\iota_M}(x_0).$$

In this sense, given $z_0 = (x_0, t_0) \in M \times \mathbb{R}$, for $\phi \in \mathcal{B}_{x_0}$ one may consider the globally defined function $\phi G_{z_0} : M \times \mathbb{R} \to [0,\infty)$. Lastly, given $z_0 = (x_0, t_0) \in M \times \mathbb{R}$ and $R \in (0,\infty)$, we define

$$\begin{aligned}
S_R(t_0) &:= M \times \{t_0 - R^2\},\\
P_R(z_0) &:= B_R(x_0) \times \left(\left[t_0 - R^2, t_0\right] \cap (0,\infty)\right),\\
T_R(t_0) &:= M \times \left(\left[t_0 - 4R^2, t_0 - R^2\right] \cap (0,\infty)\right).
\end{aligned}$$

Definition 3.2. Assume ∇_t is a solution to Yang–Mills flow on $M \times [0,T)$. For $z_0 = (x_0, t_0) \in M \times [0,T)$, $\phi \in \mathcal{B}_{x_0}$, and $R \in [0, \min\{\iota_M, \sqrt{t_0}/2\}]$, let

$$\begin{aligned}
\Phi_{z_0}(R; \nabla_t) &:= \tfrac{R^4}{2} \int\limits_{S_R(t_0)} \left|F_{\nabla_t}\right|^2 \phi^2 G_{z_0}\, dV,\\
\Psi_{z_0}(R; \nabla_t) &:= \tfrac{R^2}{2} \int\limits_{T_R(t_0)} \left|F_{\nabla_t}\right|^2 \phi^2 G_{z_0}\, dV\, dt.
\end{aligned}$$

Next we record a fundamental scaling law for the entropy functionals which is utilized in deriving the monotonicity formulas under Yang–Mills flow. These monotonicity formulas are shown in ([9]), but we include some brief discussion of some properties for convenience, and also because we utilize some of the calculations in the sequel. We restrict the lemma to flat space for convenience.

Lemma 3.3. *Fix ∇_t a solution to Yang–Mills flow on $(\mathbb{R}^n, g_{\mathrm{Euc}}) \times [0,T)$. For all $z_0 = (x_0, t_0) \in \mathbb{R}^n \times [0,T)$, and $\left(0 < R \le \sqrt{t_0}/2\right)$, setting $\phi \equiv 1$ in Definition 3.2 yields*

$$\begin{aligned}
\Phi_{z_0}(R; \nabla_t) &= \Phi_{z_0}\left(1; \nabla_t^R\right),\\
\Psi_{z_0}(R; \nabla_t) &= \Psi_{z_0}\left(1; \nabla_t^R\right),
\end{aligned}$$

where here ∇_t^R is the rescaled connection as defined in Definition 2.3.

Proof. Without loss of generality we may take $z_0 = 0$. For notational convenience we suppress the subscripts on Φ, Ψ, and G. We fix $R > 0$ and consider a change of coordinates

$$x = Ry, \qquad t = R^2 s.$$

Then, rescaling coordinates and recalling the rescaling of the curvature tensor (2.3),

$$dx = R^n\, dy, \qquad dt = R^2\, ds, \qquad G(x,t) = R^{-n} G(y,s), \qquad F_{\nabla_s^R}(y) = R^2 F_{\nabla_{R^2 s}}(Ry).$$

It follows that

$$\begin{aligned}
\Phi(R; \nabla_t) &= \tfrac{R^4}{2} \int\limits_{S_1} \left|F_{\nabla_{R^2 s}}(Ry)\right|^2 \phi(y) G(y,s)\, dy\\
&= \tfrac{1}{2} \int\limits_{S_1} \left|F_{\nabla_s^R}(y)\right|^2 \phi(y) G(y,s)\, dy\\
&= \Phi(1; \nabla_t^R).
\end{aligned}$$

Similarly,

$$
\begin{aligned}
\Psi(R;\nabla_t) &= \frac{R^4}{2}\int_{T_1}\left|F_{\nabla_{R^2 s}}(Ry)\right|^2\phi(y)G(y,s)\,dy\,ds\\
&= \frac{1}{2}\int_{T_1}\left|F_{\nabla^R_s}(y)\right|^2\phi(y)G(y,s)\,dy\,ds\\
&= \Psi\left(1;\nabla^R_t\right).
\end{aligned}
$$

The result follows. $\square$

3.3. Entropy monotonicities

In this section we recall the monotonicity formulae for Φ and Ψ, established in [9]. Again we record the proof on $\mathbb{R}^n$ for convenience and as we will use parts of argument in the sequel.

Proposition 3.4. *Let ∇_t to be a smooth solution to Yang–Mills flow for $(\mathbb{R}^n, g_{\mathrm{Euc}}) \times [0,T)$. For all $z_0 = (x_0,t_0) \in \mathbb{R}^n \times [0,T)$, and $0 < \rho \le r < \sqrt{t_0}/2$, setting $\phi \equiv 1$ in Definition 3.2 yields*

$$
\begin{aligned}
&\Phi_{z_0}(\rho;\nabla_t) \le \Phi_{z_0}(r;\nabla_t)\\
&\Psi_{z_0}(\rho;\nabla_t) \le \Psi_{z_0}(r;\nabla_t).
\end{aligned}
$$

Proof. We begin with the monotonicity statement for Φ. We will include a generic cutoff function for purposes of a later lemma. We fix $R > 0$ and consider a change of coordinates as in Lemma 3.3. As described there, it follows that

$$
\Phi(R;\nabla_t) = \frac{R^4}{2}\int_{S_1}\left|F_{\nabla_{R^2 s}}(Ry)\right|^2\phi^2(Ry)G(y,s)\,dy.
$$

A crucial point here is that we are *not* rescaling the connection as well. One now differentiates and rescales back to obtain

$$
\begin{aligned}
\frac{\partial}{\partial R}\left[\Phi(R;\nabla_t)\right] = {}& \frac{4}{R}\Phi(R;\nabla_t) + \left[R^3\int_{S_R}\left\langle F_{\nabla_t}, x \lrcorner \partial F_{\nabla_t}\right\rangle \phi^2 G\,dx\right]_{I_1}\\
&+\left[2R^3\int_{S_R}\left\langle F_{\nabla_t}, t\left(\frac{\partial F_{\nabla_t}}{\partial t}\right)\right\rangle\phi^2 G\,dx\right]_{I_2} + \left[R^3\int_{S_R}\left|F_\nabla\right|^2\phi\, x\lrcorner\nabla\phi\,dx\right].
\end{aligned}
$$

To address I_1, we recall some coordinate formulas

$$
\begin{aligned}
\nabla_i F^\beta_{jk\alpha} &= \partial_i F^\beta_{jk\alpha} + \Gamma^\beta_{i\mu}F^\beta_{jk\alpha} - F^\beta_{jk\mu}\Gamma^\mu_{i\alpha},\\
\nabla_i F^\beta_{jk\alpha} &= -\left(\nabla_k F^\beta_{ij\alpha} + \nabla_j F^\beta_{ki\alpha}\right).
\end{aligned}
$$

Combining these we conclude that

$$
\partial_i F^\beta_{jk\alpha} = -\left(\nabla_k F^\beta_{ij\alpha} + \nabla_j F^\beta_{ki\alpha}\right) - \Gamma^\beta_{i\mu}F^\mu_{jk\alpha} + F^\beta_{jk\mu}\Gamma^\mu_{i\alpha}.
$$

With this in mind we manipulate I_1,

$$
\begin{aligned}
I_1 = {}& R^3\int_{S_R} x^i\left(\nabla_k F^\beta_{ij\alpha} + \nabla_j F^\beta_{ki\alpha}\right)F^\alpha_{jk\beta}\phi^2 G\,dx\\
&+ R^3\int_{S_R} x^i\Gamma^\beta_{i\mu}F^\mu_{jk\alpha}F^\alpha_{jk\beta}\phi^2 G\,dx - R\int_{S_R} x^i F^\beta_{jk\mu}\Gamma^\mu_{i\alpha}F^\alpha_{jk\beta}\phi^2 G\,dx
\end{aligned}
$$

$$
\begin{aligned}
&= 2R^3 \int\limits_{S_R} x^i \left(\nabla_k F_{ij\alpha}^{\beta}\right) F_{jk\beta}^{\alpha} \phi^2 G\, dx \\
&= -2R^3 \int\limits_{S_R} F_{ij\alpha}^{\beta} \nabla_k \left[x^i F_{jk\beta}^{\alpha} \phi^2 G\right] dx \\
&= 2R^3 \int\limits_{S_R} \left[F_{ij\alpha}^{\beta} F_{ij\beta}^{\alpha} + F_{ij\alpha}^{\beta} x^i \left(\nabla_k F_{kj\beta}^{\alpha}\right) - \tfrac{1}{2t} x^i F_{ij\alpha}^{\beta} x^k F_{kj\beta}^{\alpha}\right] \phi^2 G\, dx \\
&\quad - 4R^3 \int\limits_{S_R} F_{ij\alpha}^{\beta} x^i F_{jk\beta}^{\alpha} (\nabla_k \phi) \phi G\, dx \\
&= -\tfrac{4}{R} \Phi(R;\nabla) + R^3 \int\limits_{S_R} \left[\tfrac{1}{t} |x \lrcorner F|^2 - 2\left\langle x \lrcorner F, D^* F\right\rangle\right] \phi^2 G\, dx \\
&\quad - 4R^3 \int\limits_{S_R} F_{ij\alpha}^{\beta} x^i F_{jk\beta}^{\alpha} (\nabla_k \phi) \phi G\, dx.
\end{aligned}
$$

Also we have

$$
\begin{aligned}
I_2 &= 2R^3 \int\limits_{S_R} t \left\langle F, \tfrac{\partial F}{\partial t}\right\rangle \phi^2 G\, dx \\
&= -2R^3 \int\limits_{S_R} t \left\langle F, D D^* F\right\rangle \phi^2 G\, dx \\
&= 4R^3 \int\limits_{S_R} t F_{ij\alpha}^{\beta} \nabla_i (D^* F)_{j\beta}^{\alpha} \phi^2 G\, dx \\
&= R^3 \int\limits_{S_R} \left[4t \left|D^* F\right|^2 - 2\left\langle x \lrcorner F, D^* F\right\rangle\right] \phi G\, dx - 8R^3 \int\limits_{S_R} t F_{ij\alpha}^{\beta} (D^* F)_{j\beta}^{\alpha} (\nabla_i \phi) \phi G\, dx.
\end{aligned}
$$

Combining these calculations gives

$$
\begin{aligned}
\tfrac{\partial}{\partial R}\left[\Phi(R;\nabla_t)\right] = {} & |t|\, R^3 \int\limits_{S_R} \left|\tfrac{x}{t} \lrcorner F_{\nabla_t} - 2D_{\nabla_t}^* F_{\nabla_t}\right|^2 \phi^2 G\, dx \\
& + 4R^3 \int\limits_{S_R} \left(x^k F_{kj\beta}^{\alpha} - 2t \left(D^* F\right)_{j\beta}^{\alpha}\right) F_{ij\alpha}^{\beta} (\nabla_i \phi) \phi G\, dx \\
& + R^3 \int\limits_{S_R} |F_\nabla|^2 \phi x \lrcorner \nabla \phi\, dx.
\end{aligned}
\tag{3.1}
$$

In particular, when $\phi \equiv 1$, we have monotonicity, which yields the first claim.

Next we prove the monotonicity of Ψ, only considering the case where $\phi \equiv 1$. We fix $R > 0$ and use the coordinate change as in Lemma 3.3 once more, and it follows that

$$
\Psi(R;\nabla_t) = R^4 \int\limits_{T_1} \left|F_{\nabla_{R^2 s}}(Ry)\right|^2 \phi^2(y) G(y,s)\, dy\, ds.
$$

Once again, crucially, we are *not* rescaling the connection. One now obtains

$$
\tfrac{\partial}{\partial R}\left[\Psi(R;\nabla_t)\right] = \tfrac{4}{R}\Psi(R;\nabla_t) + \underbrace{\left[2R \int\limits_{T_R} \left\langle F_{\nabla_t}, x \lrcorner \partial F_{\nabla_t}\right\rangle \phi^2 G\, dx\, dt\right]}_{I_1}
$$

$$+\left[4R\int_{T_R}\left\langle F_{\nabla_t}, t\left(\frac{\partial F_{\nabla_t}}{\partial t}\right)\right\rangle\phi^2 G\,dx\,dt\right]_{I_2}.$$

Nearly identical estimates for I_1 and I_2 as in the case of Φ above yield

$$\frac{\partial}{\partial R}\left[\Psi(R;\nabla_t)\right] = 2R\int_{T_R}|t|\left|\frac{x}{t}\lrcorner F_{\nabla_t} - 2D^*_{\nabla_t}F_{\nabla_t}\right|^2\phi^2 G\,dx\,dt.$$

The result follows. □

Next we state the general monotonicity formula for Φ and Ψ on arbitrary Riemannian manifolds. The proof is similar to that of Proposition 3.4, incorporating further estimates due to the presence of the cutoff function. We state here the result of ([9] Theorem 2), which applies to Yang–Mills–Higgs flow, and we just restrict the result to Yang–Mills flow. We point out that a similar result was claimed in [2], but uses definitions of Φ and Ψ with incorrect scaling. Note that the notation for Φ and Ψ agrees with various other literature, but is reversed from that chosen in [9]. Moreover, we state an improved statement which is clearly implicit in [9], simply including an extra term in the inequality which is dropped in the statement in [9].

Theorem 3.5 *([9] Theorem 2, pp.448). Let ∇_t be a smooth solution to Yang–Mills flow on $M\times[0,T)$. Then for $z_0=(x_0,t_0)\in M\times[0,T]$ and $0<R_1\le R_2\le\min\{\iota_M,\sqrt{t_0}/2\}$, we have*

$$\Psi_{z_0}(R_1;\nabla_t)+\int_{R_1}^{R_2} r\int_{T_r(t_0)}|t-t_0|\left|\frac{x-x_0}{2|t-t_0|}\lrcorner F_{\nabla_t} - D^*_{\nabla_t}F_{\nabla_t}\right|^2\phi^2 G_{z_0}\,dV\,dt\,dr \tag{3.2}$$
$$\le e^{C(R_2-R_1)}\Psi_{z_0}(R_2;\nabla_t)+C(R_2-R_1)\mathcal{YM}(\nabla_0),$$

$$\Phi_{z_0}(R_1;\nabla_t)+\int_{R_1}^{R_2} r^3\int_{S_r(t_0)}|t-t_0|\left|\frac{x-x_0}{2|t-t_0|}\lrcorner F_{\nabla_t} - D^*_{\nabla_t}F_{\nabla_t}\right|^2\phi^2 G_{z_0}\,dV\,dr \tag{3.3}$$
$$\le e^{C(R_2-R_1)}\Phi_{z_0}(R_2;\nabla_t)+C\,(R_2-R_1)\,\mathcal{YM}(\nabla_0).$$

As the statement above makes clear, the functionals Φ and Ψ are fixed if the connection satisfies a certain modified Yang–Mills type equation:

Definition 3.6. Let ∇_t be a nontrivial smooth one-parameter family of connections on $\mathbb{R}^n\times(-\infty,0]$. Then ∇_t is a *soliton* if

$$D^*_{\nabla_t}F_{\nabla_t}=\frac{x}{2t}\lrcorner F_{\nabla_t}.$$

We end with a useful technical observation showing that the different entropies Φ and Ψ are uniformly equivalent, which exploits the monotonicity

Lemma 3.7. *Let ∇_t be a solution to Yang–Mills flow on $M\times[0,T)$. There exists a uniform constant C such that for $z_0=(x_0,t_0)\in M\times[0,T)$ and for R with $0<R\le\min\{\iota_M,\sqrt{t_0}/2\}$, we have*

$$C^{-1}\Psi_{z_0}(R;\nabla_t)\le\Phi_{z_0}(2R;\nabla_t)\le C\Psi_{z_0}(2R;\nabla_t).$$

Proof. We give the proof on $\mathbb{R}^n$, in which case the monotonicity does not involve the error term involving the Yang–Mills energy, with the generalization to manifolds a straightforward extension. Without loss of generality we can consider the time interval to be $[-1,0]$ and choose $z_0=(0,0)$. Then we have, using the monotonicity of Φ and a change of variables,

$$\begin{aligned}
\Phi(2R) &\geq \frac{1}{R}\int_R^{2R} \Phi(s)\,ds \\
&= \frac{R^3}{2}\int_{s=R}^{s=2R}\int_{M\times\{-s^2\}} |F_s|^2\phi^2 G\,dV\,ds \\
&= \frac{R^3}{2}\int_{t=-R^2}^{t=-4R^2}\frac{1}{2\sqrt{-t}}\int_{M\times\{t\}} |F_t|^2\phi^2 G\,dV\,dt \\
&\geq cR^2\int_{T_R(0)} |F_t|^2\phi^2 G\,dV\,dt \\
&= c\Psi(R).
\end{aligned}$$

Analogously we have

$$\begin{aligned}
\Phi(R) &\leq \frac{1}{R}\int_R^{2R} \Phi(s)\,ds \\
&= \frac{R^3}{4}\int_{s=R}^{s=2R}\int_{M\times\{-s^2\}} |F_s|^2\phi^2 G\,dV\,ds \\
&= \frac{R^3}{4}\int_{t=-R^2}^{t=-4R^2}\frac{1}{2\sqrt{-t}}\int_{M\times\{t\}} |F_t|^2\phi^2 G\,dV\,dt \\
&\leq CR^2\int_{T_R(0)} |F_t|^2\phi^2 G\,dV\,dt \\
&= C\Psi(R).
\end{aligned}$$

The result follows. □

3.4. Epsilon-regularity

A central phenomenon in understanding the singularity formation of geometric flows is that of ϵ-regularity. A result of this kind for Yang–Mills flow is shown in [9], relying centrally on the monotonicity formula for Ψ and the evolution equation for the curvature. Once again we only state the result for solutions to Yang–Mills flow though the result is shown for Yang–Mills–Higgs flow in [9]. We also point out that a similar result is claimed in [2], although it relies on the incorrectly defined Ψ functional.

Theorem 3.8 *([9] Theorem 4, pp.454). Suppose ∇_t is a solution to Yang–Mills flow on $M \times [0,T)$. There exist constants $C, \delta, \epsilon_0 > 0$ depending on (M,g) and $\mathcal{YM}(\nabla_0)$ so that given $z_0 = (x_0,t_0) \in M \times [0,T)$ and $0 < R < \min\{\iota_M, \sqrt{t_0}/2\}$ such that*

$$\Psi_{z_0}(R;\nabla_t) < \epsilon_0,$$

one has

$$\sup_{P_{\delta R}(z_0)} \left|F_{\nabla_t}\right|^2 \leq \frac{C}{(\delta R)^4}.$$

4. Weak compactness and limit measures

In this section we establish a weak compactness result for solutions to Yang–Mills flow satisfying certain weak convergence hypotheses. In the first subsection below we establish this theorem, and in the following subsection we refine the analysis to show a number of properties of the limiting energy densities and defect measures.

4.1. Weak compactness theorem

Theorem 4.1. *Suppose $\{\nabla_t^i\}$ is a sequence of smooth solutions to Yang–Mills flow over $M \times [-1,0]$ with $\mathcal{YM}(\nabla_t^i) \leq \mathcal{YM}(\nabla_{-1}^i) < C$. Moreover, suppose $\{\nabla_{-1}^i\} \to \nabla$ weakly in $H^{1,2}_{loc}(\mathcal{A}_E(M))$, and*

- $\nabla_t^i \to \nabla_t$ *in* $L^2_{loc}(M \times [-1,0])$,
- $\frac{\partial \nabla_t^i}{\partial t} \to \frac{\partial \nabla_t}{\partial t}$ *weakly in* $L^2_{loc}(M \times [-1,0])$,
- $F_{\nabla_t^i} \to F_{\nabla_t}$ *weakly in* $L^2_{loc}(M \times [-1,0])$.

Then ∇_t is gauge equivalent to a weak solution to Yang–Mills flow, and there exists a closed set Σ of locally finite $(n-2)$-dimensional parabolic Hausdorff measure such that ∇_t is a smooth solution on $(M \times (-1,0))\backslash\Sigma$.

Proof. Set

$$\Phi^i_{z_0}(r) := \begin{cases} \Phi_{z_0}\left(r; \nabla_t^i\right) & r \in \left(0, \sqrt{1+t_0}\right) \\ \Phi_{z_0}\left(\sqrt{1+t_0}; \nabla_t^i\right) & \text{otherwise.} \end{cases}$$

Now define the concentration set

$$\Sigma := \bigcap_{r>0} \left\{ z \in M \times [-1,0] \mid \liminf_{k\to\infty} \Phi^k_z(r) \geq \epsilon_0 \right\},$$

where ϵ_0 is the constant of Theorem 3.8. To address the theorem, we divide the proof up into three pieces: Lemma 4.2, Lemma 4.3, and Lemma 4.5.

Lemma 4.2. *Σ is closed.*

Proof. Let $\bar{z}$ lie in the closure of Σ and $\{z_k\}_{k\in\mathbb{N}} \in \Sigma$ with $z_k \to \bar{z}$. By the definition of Σ,

$$\liminf_{k\to\infty} \liminf_{i\to\infty} \Phi^i_{z_k}(r) = \liminf_{k\to\infty} \liminf_{i\to\infty} \left[\frac{r^4}{2} \int_{\mathbb{R}^n \times \{t_k - r^2\}} \left|F_t^i\right|^2 \phi^2 G_{z_k}\, dV \right] \geq \epsilon_0.$$

Note that $G_{z_k} \to G_{\bar{z}}$ on any closed sets not containing $\bar{z}$. Moreover, for fixed i the function $\left|F_t^i\right|^2$ is in L^1. Therefore we can fix $r > 0$, apply the dominated convergence theorem and interchange lim inf ordering by an elementary argument to conclude

$$\begin{aligned} \liminf_{i\to\infty} \frac{r^4}{2} \int_{M\times\{\bar{t}-r^2\}} \left|F_t^i\right|^2 \phi^2 G_{\bar{z}}\, dV &= \liminf_{i\to\infty} \lim_{k\to\infty} \frac{r^4}{2} \int_{M\times\{\bar{t}-r^2\}} \left|F_t^i\right|^2 \phi^2 G_{z_k}\, dV \\ &= \liminf_{i\to\infty} \liminf_{k\to\infty} \frac{r^4}{2} \int_{M\times\{\bar{t}-r^2\}} \left|F_t^i\right|^2 \phi^2 G_{z_k}\, dV \\ &\geq \epsilon_0. \end{aligned}$$

Therefore $\bar{z} \in \Sigma$, so we conclude Σ is closed. □

Lemma 4.3. *∇_t is gauge equivalent to a smooth solution to Yang–Mills flow on $(M \times (-1,0]) \setminus \Sigma$.*

Proof. Given $z \in (\mathbb{R}^n \times (-1,0]) \setminus \Sigma$, by construction there exists $r_0 > 0$ such that

$$\liminf_{k \to \infty} \Phi_z^k(r_0) \leq \epsilon_0.$$

Passing to a subsequence and applying Lemma 3.7, we obtain an ϵ_0 upper bound for Ψ, and by Theorem 3.8, we conclude that

$$\sup_{P_{\delta r_0}(z)} \left| F_t^k \right|^2 \leq \frac{C}{(\delta r_0)^4},$$

for some universal constant $\delta > 0$. Applying ([22], Theorem 2.2) we conclude uniform estimates on all derivatives of curvature on a parabolic ball of radius $\frac{\delta r_0}{2}$.

Using the Uhlenbeck gauge-fixing Theorem ([20] Theorem 1.3) and the gauge-patching argument of ([4] Corollary 4.4.8) we can obtain a Coulomb gauge on $B_{\frac{\delta r_0}{4}}$. Moreover, by applying elliptic regularity estimates ([4] Lemma 2.3.11) and the Sobolev inequality we obtain uniform pointwise estimates for the connection in the Coulomb gauge on $B_{\frac{\delta r_0}{8}}$. By applying the Yang–Mills flow PDE directly to this gauge-fixed connection and using the previous estimates on the derivatives of curvature we obtain uniform pointwise estimates for the gauge fixed connections on $P_{\frac{\delta r_0}{8}}$. Thus for each point z_0 we have constructed a radius $\frac{\delta r_0}{8}$ and a sequence of gauge transformations for which the parabolic ball of that radius has uniform control along some subsequence of gauge-fixed connections.

Fix a compact set K such that $K \cap \Sigma = \varnothing$. For each $z \in K$ there exist arbitrarily large values of k and parabolic balls centered at z of the type described above. This collection of parabolic balls covers K, and since K is compact we can choose a finite subcover, and also pass to a subsequence of connections all of which have the bounds described above. A further application of the gauge-patching result ([4] Corollary 4.4.8) allows us to conclude the existence of a single gauge transformation, which, when applied to our sequence, yields a sequence of connections with uniform $C^{l,\alpha}$ bounds. By the Arzela–Ascoli Theorem we obtain a further subsequence converging on K. □

Lemma 4.4. *Σ has locally finite $(n-2)$-dimensional parabolic Hausdorff measure.*

Proof. Fix a compact set K, and some $r_0 > 0$. By Vitali's covering lemma there exists some $l \in \mathbb{N}$, $\{z_k\}_{k=1}^l \subset K \cap \Sigma$ and $\{r_k\}_{k=1}^l \subset (0, r_0)$ so that the sets $\{P_{r_k}(z_k)\}_{k=1}^l$, are mutually disjoint and $K \cap \Sigma$ is covered by $\{P_{5r_k}(z_k)\}_{k=1}^l$. Let $\overline{z}_k := z_k + \left(0, r_k^2\right)$ and fix some $\delta > 0$ to be determined later.

The proof requires two different estimates on G on different domains. First, on $(M \times \left[t_k - 4\delta^2 r_k^2, t_k - \delta^2 r_k\right]) \setminus P_{r_k}(z_k)$ one has

$$G_{z_k} \leq \delta^{-n} e^{-1/(4\delta)^2} G_{\overline{z}_k}.$$

Also, for points in $B_{r_k}(x^k) \times [t_k - 4\delta^2 r_k^2, t_k - \delta^2 r_k^2]$ one has

$$G_{z_k} \leq C_\delta r^{-n}.$$

We will also employ the estimate of Lemma 3.7, in particular

$$\Phi_{z_0}(R; \nabla_t) \leq C \Psi_{z_0}(R; \nabla_t).$$

Combining the observations above we obtain, for all k, i

$$
\begin{aligned}
\epsilon_0 &\le \Phi^i_{z_k}(\delta r_k) \\
&\le C\Psi_{z_k}\left(\delta r_k;\nabla^i_t\right) \\
&= C\delta^2 r_k^2 \int_{t_k-4\delta^2 r_k^2}^{t_k-\delta^2 r_k^2}\int_{M\setminus B_{r_k}(x_k)} \left|F^i_t\right|^2 G_{z_k}\,dV\,dt + C\delta^2 r_k^2 \int_{t_k-4\delta^2 r_k^2}^{t_k-\delta^2 r_k^2}\int_{B_{r_k}(x_k)} \left|F^i_t\right|^2 G_{z_k}\,dV\,dt \\
&\le C\frac{e^{-1/(4\delta)^2}}{4\delta^n}\delta^2 r_k^2 \int_{t_k-4\delta^2 r_k^2}^{t_k-\delta^2 r_k^2}\int_{M\setminus B_{r_k}(x_k)} \left|F^i_t\right|^2 G_{\overline{z}_k}\,dV\,dt + C_\delta r_k^{2-n} \int_{t_k-4\delta^2 r_k^2}^{t_k-\delta^2 r_k^2}\int_{B_{r_k}(x_k)} \left|F^i_t\right|^2 dV\,dt \\
&\le \left[C\frac{e^{-1/(4\delta)^2}}{4\delta^n}\delta^2 r_k^2 \int_{t_k-4\delta^2 r_k^2}^{t_k-\delta^2 r_k^2}\int_{M} \left|F^i_t\right|^2 G_{\overline{z}_k}\,dV\,dt\right]_{I_1} + \left[C_\delta r_k^{2-n}\int_{P_{r_k}(z_k)} \left|F^i_t\right|^2 dV\,dt\right]_{I_2}.
\end{aligned}
$$

Observe that we can estimate I_1 using Theorem 3.5 via

$$
\begin{aligned}
\delta^2 r_k^2 \int_{t_k-4\delta^2 r_k^2}^{t_k-\delta^2 r_k^2}\int_M \left|F^i_t\right|^2 G_{\overline{z}_k}\,dV\,dt &= \delta^2 r_k^2 \int_{t_k+r_k^2-4r_k^2(1+\delta^2)}^{t_k+r_k^2-r_k^2(1+\delta^2)}\int_M \left|F^i_t\right|^2 G_{\overline{z}_k}\,dV\,dt \\
&= \Psi_{\overline{z}_k}\left(r_k\sqrt{1+\delta^2};\nabla^i_t\right) \\
&\le C\Psi_{\overline{z}_k}\left(r_0;\nabla^i_t\right) + C((\mathcal{YM}(\nabla_{-1}))) \\
&\le C(\mathcal{YM}(\nabla_{-1})).
\end{aligned}
$$

Hence, since $\lim_{\delta\to 0}\frac{e^{-1/(4\delta)^2}}{4\delta^n} = 0$, we can choose $\delta > 0$ sufficiently small so that $I_1 \le \frac{\epsilon_0}{2}$, which then implies that $I_2 \ge \frac{\epsilon_0}{2}$, which by elementary manipulations gives

$$
r_k^{n-2} \le \frac{C}{\epsilon_0}\int_{P_{r_k}(z_k))} \left|F^i_t\right|^2 dV\,dt.
$$

Therefore we have

$$
\begin{aligned}
\mathcal{P}^{n-2}_{5r_0}(P_R\cap\Sigma) &\le \sum_{k=1}^{l}(5r_k)^{n-2} \\
&\le C\sum_{k=1}^{l}\int_{P_{r_k}(z_k)} \left|F^i_t\right|^2 dV\,dt \\
&\le C\mathcal{YM}(\nabla_{-1}).
\end{aligned}
$$

Sending $r_0 \to 0$ allows us to conclude that $\mathcal{P}^{n-2}(\Sigma\cap K) < \infty$ for any compact set K. The result follows. □

Lemma 4.5. *∇_t is a weak solution to Yang–Mills flow.*

Proof. We verify (2.1) by approximating via cutoff functions which excise the singular set Σ. To construct these functions, first consider the coverings constructed in Lemma 4.4. In particular, given any $r_0 > 0$ there is some finite cover $\{P_{r_i}(z_i)\}_{i=1}^l$ of Σ, for some $l \in \mathbb{N}$ with $r_i < r_0$ satisfying

$$\sum_{i=1}^{l} r_i^{-4}\left|P_{r_i}(z_i)\right| \approx \mathcal{P}^{n-2}_{5r_0}(K\cap\Sigma) \leq C\mathcal{YM}(\nabla_{-1}), \tag{4.1}$$

where here $|\cdot|$ denotes the Lebesgue measure on $\mathbb{R}^n\times\mathbb{R}$.

Let $\phi\in C_0^\infty(P_2,[0,\infty))$ be a standard bump function satisfying $0\leq\phi\leq 1$ and $\phi\equiv 1$ on P_1. For all $i\in\mathbb{N}$, define

$$\phi_i(x,t) := \phi\left(\tfrac{x-x_i}{r_i}, \tfrac{t-t_i}{r_i^2}\right).$$

Let $\alpha\in C^\infty([0,T];L^2(\Lambda^2(\operatorname{Ad}E)))$ and arbitrary and set

$$\eta := \alpha\inf_i(1-\phi_i)\in C_0^\infty\left(\left(\mathbb{R}^n\times(-1,0)\right)\setminus\Sigma\right).$$

Note that by definition, $\eta\to\alpha$ almost everywhere as $r_0\to 0$. Furthermore, observing that $\operatorname{supp}\eta\subset(\mathbb{R}^n\times(-1,0))\setminus\Sigma$, it follows from Lemma 4.3 that, setting $\Upsilon_t=\nabla_{\mathrm{ref}}-\nabla_t$, we have

$$\int_{-1}^{0}\int_M \left\langle\Upsilon,\tfrac{\partial\eta}{\partial t}\right\rangle - \langle F, D\eta\rangle\, dV\,dt = 0.$$

Using this we can estimate

$$\begin{aligned}
&\int_{-1}^{0}\int_M \left\langle\Upsilon,\tfrac{\partial\alpha}{\partial t}\right\rangle - \langle F, D\alpha\rangle\, dV\,dt\\
&\quad= \left|\int_{-1}^{0}\int_M \left\langle\Upsilon,\tfrac{\partial(\alpha-\eta)}{\partial t}\right\rangle - \langle F, D(\alpha-\eta)\rangle\, dV\,dt\right|\\
&\quad= \left|\int_{-1}^{0}\int_M \left\langle\tfrac{\partial\Upsilon}{\partial t},\alpha-\eta\right\rangle - \left\langle F,[1-\inf_i(1-\phi_i)]D\alpha\right\rangle - \left\langle F,\alpha\wedge d(\inf_i(1-\phi_i))\right\rangle dV\,dt\right|\\
&\quad= |I_1+I_2+I_3|\\
&\quad\leq \sum_{j=1}^{3}\left|I_j\right|.
\end{aligned}$$

First, since we have almost everywhere convergence of α to η and $\frac{\partial\Upsilon}{\partial t}$ is in L^2 we have $\lim_{r_0\to 0} I_1=0$. Similarly since $[1-\inf_i(1-\phi_i)]$ goes to zero uniformly one has that $\lim_{r_0\to 0} I_2=0$. For the final term, we observe using Hölder's inequality and (4.1) that

$$\begin{aligned}
\lim_{r_0\to 0}|I_3| &\leq C\lim_{r_0\to 0}||F||_{L^2(\cup_i P_{r_i}(z_i))}\left[\int_{-1}^{0}\int_M\left|\nabla\inf_{1\leq i\leq l}(1-\phi_i)\right|^2 dV\,dt\right]^{\frac12}\\
&\leq C\lim_{r_0\to 0}||F||_{L^2(\cup_i P_{r_i}(z_i))}\left[\sum_{i=1}^{l} r_i^{-2}\left|P_{r_i}(z_i)\right|\right]^{\frac12}\\
&\leq C\lim_{r_0\to 0} r_0\left[\sum_{i=1}^{l} r_i^{-4}\left|P_{r_i}(z_i)\right|\right]^{\frac12}\\
&= 0.
\end{aligned}$$

The lemma follows. □

Combining the result of Lemma 4.2, Lemma 4.3, and Lemma 4.5, the results of Theorem 4.1 follow. □

4.2. Structure of limit measures

Assume the setup of Theorem 4.1. Observe that the measures

$$\left\{ \left| F_{\nabla_t^i} \right|^2 dV\, dt \right\} \text{ and } \left\{ \left| \frac{\partial \nabla_t^i}{\partial t} \right|^2 dV\, dt \right\}$$

admit subsequences converging in the sense of Radon measures to some limit measures. We can compare these to the measures induced by the weak H_1^2 limit ∇ to define measures μ, ν and η via

$$\left| F_{\nabla_t^i} \right|^2 dV\, dt \to \left| F_{\nabla_t^\infty} \right|^2 dV\, dt + \nu \equiv \mu,$$

$$\left| \frac{\partial \nabla_t^i}{\partial t} \right|^2 dV\, dt \to \left| \frac{\partial \nabla_t^\infty}{\partial t} \right|^2 dV\, dt + \eta.$$

The remainder of the section consists of a series of lemmas further refining the nature of these measures.

Lemma 4.6. *Fix $z = (x,t) \in M \times [-1, 0]$ and $\phi \in \mathcal{B}_x$. Then*

$$\Theta(\mu, z) := \lim_{R \to 0} R^2 \int_{T_R(z)} \phi^2(x) G_z(x,t)\, d\mu(x,t)$$

exists and is upper semicontinuous for all $z \in M \times [0, \infty)$. Moreover,

$$\Sigma = \{ z \in M \times (0, \infty) \mid \epsilon_0 \le \Theta(\mu, z) < \infty \}.$$

Proof. We consider the limit as $i \to \infty$ in the monotonicity inequality (3.2). In particular, for $0 < R \le R_0$, let

$$f(R, d\mu) = e^{CR} \left[\frac{R^2}{2} \int_{T_R} \phi^2 G_z\, d\mu + C e^{CR} R \mathcal{YM}(\nabla_{-1}) \right].$$

We observe that (3.2) implies that

$$\begin{aligned} f(R, \left| F_{\nabla_t^i} \right|^2 dV) &= e^{CR} \left[\Psi_{z_0}(R, \nabla_t^i) + CR\mathcal{YM}(\nabla_{-1}) \right] \\ &\le e^{CR} \left[e^{C(R_0 - R)} \Psi_{z_0}(R, \nabla_t^i) + C(R_0 - R)\mathcal{YM}(\nabla_{-1}) + CR\mathcal{YM}(\nabla_{-1}) \right] \\ &= f(R_0, \left| F_{\nabla_t^i} \right|^2 dV). \end{aligned}$$

Using that $\left| F_{\nabla_t^i} \right|^2 dV$ converges to $d\mu$, it follows that $f(R, d\mu)$ is monotone nondecreasing as well. It follows that $\lim_{R \to 0} f(R, d\mu)$ exists, and by elementary arguments the limit defining Θ also exists, and is upper semicontinuous. □

Lemma 4.7. *For $\mathcal{P}^{n-2}$-almost everywhere $z \in \Sigma$, one has*

$$\lim_{R \to 0} R^{2-n} \int_{P_R(z)} \left| F_{\nabla_t} \right|^2 dV\, dt = 0, \qquad \Theta(\mu, z) = \Theta(\nu, z) \ge \epsilon_0.$$

Proof. To show the first claim, let

$$K_j = \left\{ z \in \Sigma \mid \limsup_{R \to 0} R^{2-n} \int_{P_R(z)} |F_t|^2\, dV\, dt > j^{-1} \right\}.$$

We will show that the $(n-2)$-parabolic Hausdorff measure of K_j is zero for each j, which suffices. Fixing some $\delta > 0$ we can apply Vitali's covering lemma to obtain a covering of K_j by disjoint parabolic balls $P_{r_k}(z_k)$ with $z_k \in K_j, 5r_k \leq \delta$, such that $K_j \subset \bigcup P_{5r_k}(z_k)$. It follows that there exists $C > 0$ such that

$$\begin{aligned}\mathcal{P}^{n-2}(K_j) &\leq \lim_{\delta\to 0} \sum_k (5r_k)^{n-2} \\ &\leq Cj \lim_{\delta\to 0} \int\limits_{N_\delta(\Sigma)} |F_t|^2 \, dV \, dt \\ &= 0,\end{aligned}$$

where $N_\delta(\Sigma)$ indicates the parabolic δ-tubular neighborhood of Σ, and the last line follows by the dominated convergence theorem. The second claim now follows from the first and the definitions of μ, ν. $\square$

Lemma 4.8. *For $\mathcal{P}^{n-2}$-almost everywhere $z \in \Sigma$*

$$\lim_{r\to 0} \lim_{i\to\infty} r^{4-n} \int\limits_{P_r(z)} \left|\tfrac{\partial \nabla_t^i}{\partial t}\right|^2 dV \, dt = 0.$$

Proof. We will show that for any $\epsilon > 0$, the set

$$\mathcal{C}_\epsilon := \left\{ z \in \Sigma \mid \liminf_{r\to 0} \liminf_{i\to\infty} r^{4-n} \int\limits_{P_r(z)} \left|\tfrac{\partial \nabla^i}{\partial t}\right|^2 dV \, dt \geq \epsilon \right\}$$

satisfies $\mathcal{P}^{n-4}(\mathcal{C}_\epsilon) < \infty$. Given this, we can express

$$\Sigma' := \left\{ z \in \Sigma \mid \liminf_{r\to 0} \liminf_{i\to\infty} r^{4-n} \int\limits_{P_r(z)} \left|\tfrac{\partial \nabla^i}{\partial t}\right|^2 dV \, dt = 0 \right\} = \Sigma \setminus \left(\bigcup_{n\in\mathbb{N}} \mathcal{C}_{2^{-n}} \right).$$

In particular, Σ' can be obtained from Σ by removing a countable union of sets of finite $\mathcal{P}^{n-4}$ measure, which has zero $\mathcal{P}^{n-2}$ measure by a standard argument.

To show $\mathcal{P}^{n-4}(\mathcal{C}_\epsilon) < \infty$, fix a $\delta > 0$, and apply Vitali's covering lemma to obtain a collection $\{z_k\}_{i\in\mathbb{N}} \subset \Sigma$ and $r_k \in (0, \delta)$ satisfying that $\{P_{r_k}(z_k)\}$ are mutually disjoint, $\{P_{5r_k}(z_k)\}$ cover Σ, and furthermore there is some subsequence $\{\nabla_t^i\}$ so that for all k, i,

$$r_k^{4-n} \int\limits_{P_{r_k}(z_k)} \left|\tfrac{\partial \nabla^i}{\partial t}\right|^2 dV \, dt \geq \epsilon.$$

Using this we obtain

$$\begin{aligned}\mathcal{P}^{n-4}_{5\delta}(\mathcal{C}_\epsilon) &\leq \sum_{k=1}^{\infty} (5r_k)^{n-4} \\ &= 5^{n-4} \sum_{k=1}^{\infty} r_k^{n-4} \\ &\leq \tfrac{5^{n-4}}{\epsilon} \sum_{k=1}^{\infty} \int\limits_{P_{r_k}(z_k)} \left|\tfrac{\partial \nabla^i}{\partial t}\right|^2 dV \, dt \\ &\leq \tfrac{5^{n-4}}{\epsilon} \int\limits_{\bigcup_{k=1}^{\infty} P_{r_k}(z_k)} \left|\tfrac{\partial \nabla^i}{\partial t}\right|^2 dV \, dt\end{aligned}$$

$$
\begin{aligned}
&\leq C(n,\epsilon)\int_0^2\int_{B_2}\left|\tfrac{\partial\nabla^i}{\partial t}\right|^2 dV\,dt\\
&\leq C(n,\epsilon,\mathcal{YM}(\nabla^i_{-1})),
\end{aligned}
$$

where the last line follows via the Yang–Mills energy monotonicity. Sending δ to zero proves that $\mathcal{P}^{n-4}(\mathcal{C}_\epsilon)<\infty$, finishing the proof. □

Lemma 4.9. *The density function $\Theta(\mu,x)$ is $\mathcal{P}^{n-2}$-approximately continuous at $\mathcal{P}^{n-2}$-almost every $x\in\Sigma$. That is, for all $\mathcal{P}^{n-2}$-a.e. $z\in\Sigma$ one has that for all $\epsilon>0$,*

$$
\lim_{r\to 0} r^{2-n}\mathcal{P}^{n-2}\left(\{w\in P_r(x)\cap\Sigma\mid|\Theta(\mu,w)-\Theta(\mu,z)|>\epsilon\}\right)=0.
$$

Proof. Note that for a given $x\in\Sigma$, the density $\Theta(\mu,x)$ is upper semicontinuous, so the set

$$
A_c:=\{z\mid\Theta(\mu,z)<c\}
$$

is open. Therefore for any $c_1,c_2\in[0,\infty)$ with $c_1<c_2$, the set $A_{c_2}\backslash A_{c_1}$ is a Borel set and thus measurable. Hence

$$
E_i:=\left\{z\in\Sigma\mid\tfrac{(i-1)\epsilon}{2}\leq\Theta(\mu,z)<\tfrac{i\epsilon}{2}\right\}=A_{\frac{i\epsilon}{2}}\backslash A_{\frac{(i-1)\epsilon}{2}},
$$

is a Borel set. Note that, by the definition of E_i,

$$
\mathcal{P}^{n-2}\left(\Sigma\backslash\bigcup_i E_i\right)=0.
$$

For all $x\in E_i$, by applying Theorem 3.5 of [17] to the measure $\mathcal{P}^{n-2}$ we have that

$$
\begin{aligned}
\lim_{R\to 0} R^{2-n}\mathcal{P}^{n-2}&\left(\{y\in P_r(x)\cap\Sigma\mid|\Theta(\mu,w)-\Theta(\mu,z)|>\epsilon\}\right)\\
&=\limsup_{R\to 0} R^{2-n}\mathcal{P}\left(P_r(z)\cap(\Sigma\backslash E_i)\right)\\
&=0.
\end{aligned}
$$

The result follows. □

Lemma 4.10. *One has that $\left\{\nabla^i_t\right\}$ does not converge to ∇_t strongly in $H^{1,2}_{loc}$ if and only if $\mathcal{P}^{n-2}(\Sigma)>0$ and $\nu(M\times[-1,0])>0$.*

Proof. It follows from Lemma 4.7 that if $\mathcal{P}^{n-2}(\Sigma)>0$ then for $\mathcal{P}^{n-2}$ almost everywhere $z\in\Sigma$ one has

$$
\Theta(\nu,z)=\Theta(\mu,z)\geq\epsilon_0,
$$

hence $\nu(M\times[-1,0])=\nu(\Sigma)>0$, and $\frac{1}{2}\left|F_{\nabla^i_t}\right|^2 dV\,dt$ does not converge to $\frac{1}{2}\left|F_{\nabla_t}\right|^2 dV\,dt$. Therefore $\left\{\nabla^i_t\right\}$ doesn't converge to ∇_t strongly in $H^{1,2}_{loc}$. Conversely, directly from the definition of ν, if $\nu(M\times[-1,0])>0$ then $\{\nabla^i_t\}$ cannot converge strongly to ∇ in $H^{1,2}_{loc}$. □

5. Tangent measures and stratification

In this section we establish results on the structure of tangent measures along Yang–Mills flow which will be central in the sequel. First we discuss the space $T_z(\mu)$ of all tangent measures of μ for $z\in\Sigma$. We first show that every tangent measure is invariant under parabolic dilations. Building upon this, we will associate to each tangent measure a nonnegative integer which is the dimension of the largest parabolic dilation invariant subspace which is a subset of the points of maximal density. Using this dimension we can then stratify the set Σ accordingly. In particular, we demonstrate enough structure on the tangent measures to apply a stratification result of White [23], which generalizes Federer's dimension reduction argument [6].

5.1. Setup

For the following we set

$$\mathbb{R}^{n+1}_{+} := \mathbb{R}^n \times [0,\infty), \quad \mathbb{R}^{n+1}_{-} := \mathbb{R}^n \times (-\infty, 0].$$

Definition 5.1. For $z_0 = (x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$ and $\lambda > 0$, define *parabolic dilation* and *Euclidean dilation* respectively by,

$$\begin{aligned} \mathsf{P}_{z_0,\lambda}(x,t) &:= \left(\lambda(x-x_0), \lambda^2(t-t_0)\right), \\ \mathsf{D}_{x_0,\lambda}(x) &:= \lambda(x-x_0). \end{aligned}$$

We may apply parabolic rescaling to a measure as follows: for all $A \subset \mathbb{R}^n \times \mathbb{R}$, we have

$$\begin{aligned} \mathsf{P}_{z_0,\lambda}(\mu)(A) &:= \lambda^{2-n}\mu\left(\mathsf{P}_{z_0,\lambda}(A)\right), \\ \mathsf{D}_{x_0,\lambda}(\mu)(A) &:= \lambda^{4-n}\mu\left(\mathsf{D}_{x_0,\lambda}A\right). \end{aligned}$$

We note that this scaling law reflects the scaling properties for Yang–Mills flow densities, and not a pure parabolic rescaling of say Euclidean measure.

Definition 5.2. For any $z_0 \in \Sigma$, the *tangent measure cone of* μ *at* z_0, $T_{z_0}(\mu)$, consists of all nonnegative Radon measures on $\mathbb{R}^{n+1}$ that are given by

$$T_{z_0}(\mu) := \left\{\mu^* \mid \exists r_i \to 0, \text{ such that } \mathsf{P}_{z_0,r_i}(\mu) \to \mu^*\right\}.$$

Fixing $z_0 \in \Sigma$ and $\mu^* = \mu^*_s ds \in T_{z_0}(\mu)$, we set, for any $z = (x,t) \in \mathbb{R}^{n+1}$,

$$\Theta\left(\mu^*, z, r\right) := r^4 \int_{M \times \{t-r^2\}} G_z(y,s)\, d\mu^*_s(y).$$

This is monotonically nondecreasing with respect to r so that the μ^* *density at* z, given by

$$\Theta(\mu^*, z) := \lim_{r\to 0} \Theta\left(\mu^*, z, r\right),$$

exists and is upper semicontinuous for $z = (x,t) \in \mathbb{R}^{n+1}$. Moreover, for any $z_0 \in \Sigma$ and $\mu^* \in T_{z_0}(\mu)$, we set

$$\begin{aligned} U\left(\Theta\left(\mu^*\right)\right) &:= \left\{z \in \mathbb{R}^{n+1} \mid \Theta\left(\mu^*, z\right) = \Theta\left(\mu^*, 0\right)\right\}, \\ V\left(\Theta\left(\mu^*\right)\right) &:= U\left(\Theta\left(\mu^*\right)\right) \cap \left(\mathbb{R}^n \times \{0\}\right), \\ W\left(\Theta\left(\mu^*\right)\right) &:= \left\{(x,0) \in \mathbb{R}^n \times \mathbb{R} \mid \forall (y,s) \in \mathbb{R}^{n+1}_{-}, \Theta\left(\mu^*, (y,s)\right) = \Theta\left(\mu^*, (x+y,s)\right)\right\}. \end{aligned}$$

Definition 5.3. For $z_0 \in \Sigma$ and $\mu^* \in T_{z_0}(\Sigma)$, let

$$\dim\left(\Theta\left(\mu^*\right)\right) = \begin{cases} \dim\left(V\left(\Theta\left(\mu^*\right)\right)\right) + 2, & \text{if } U\left(\Theta(\mu^*)\right) = V\left(\Theta\left(\mu^*\right)\right) \times \mathbb{R}, \\ \dim\left(V\left(\Theta\left(\mu^*\right)\right)\right) & \text{otherwise.} \end{cases}$$

5.2. Preliminary results

In this subsection we show various preliminary results on the structure of tangent measures. First we establish the existence of at least one tangent measure in Lemma 5.4. We then establish parabolic scaling invariance of tangent measures in Lemma 5.6.

Lemma 5.4. *Given a weak limit measure* μ, $z_0 \in \Sigma$, *and* $\lambda_i \to 0$ *there exists a subsequence* $\{\lambda_{i_j}\}$ *and some nonnegative Radon measure* μ^* *on* $\mathbb{R}^{n+1}$ *such that* $\mathsf{P}_{z_0,\lambda_{i_j}}(\mu) \to \mu^*$ *as weak convergence of Radon measures on* $\mathbb{R}^{n+1}$.

Proof. We fix some small radius r_0 and claim that

$$\sup_{(z,r)\in M\times[-1,0]\times(0,r_0)} r^{2-n}\,\mu(P_r(z)) < \infty. \tag{5.1}$$

In particular, we use a change of variables and Theorem 3.5 to yield

$$\begin{aligned}
r^{2-n}\mu(P_r(z)) &= r^{2-n}\lim_{i\to\infty}\int_{P_r(z)} \left|F_{\nabla^i}\right|^2 dV\,dt \\
&= \lim_{i\to\infty} r^{2-n}\int_{t=0}^{r^2}\int_{S_{\sqrt{t}}} \left|F_{\nabla^i}\right|^2 \phi\, dV\,dt \\
&= \lim_{i\to\infty} r^{2-n}\int_{s=0}^{r} s\int_{S_s} \left|F_{\nabla^i}\right|^2 \phi\, dV\,ds \\
&\leq \lim_{i\to\infty} Cr^{-2}\int_{s=0}^{r} s\Phi(s)\,ds \\
&\leq \lim_{i\to\infty} C(\Phi(r_0))r^{-2}\int_{s=0}^{r} s\,ds \\
&\leq C.
\end{aligned}$$

Hence, using (5.1), for any λ_i the sequence of dilated measures $\mathsf{P}_{z_0,\lambda_i}(\mu)$ is uniformly bounded on all Borel sets in $\mathbb{R}^{n+1}$, hence by the weak compactness of families of uniformly bounded Radon measures we obtain the existence of the subsequential limiting measure μ. □

Lemma 5.5. *For any $z_0 \in \Sigma$, $0 < r_1 < r_2 < \infty$ a sequence $\lambda_i \to 0$ and a blowup sequence $\overline{\nabla}_t^i$ one has*

$$\lim_{i\to\infty}\int_{-r_1^2}^{-r_2^2}\int_{\mathbb{R}^n} \left|x \lrcorner F_{\overline{\nabla}_t^i} + 2t\partial_t\overline{\nabla}_t^i\right|^2 G_{z_0}\,dx\,dt = 0.$$

Proof. First recall that as convergence of Radon measures on $\mathbb{R}^n$ we have

$$\tfrac{1}{2}\left|\overline{F}_t^i\right|^2 dV \to \mu_t^* \text{ for all } t \in (-\infty, 0].$$

Hence, for any $R > 0$, applying a change of variables we obtain

$$\begin{aligned}
R^4\int_{\mathbb{R}^n\times\{-R^2\}} G_{(0,0)}\,d\mu_t^*\,dt &= \lim_{i\to\infty}\int_{\mathbb{R}^n\times\{-R^2\}} \tfrac{R^4}{2}\left|\overline{F}_t^i\right|^2 G_{(0,0)}\,dV_x\,dt \\
&= \lim_{i\to\infty}\int_{\mathbb{R}^n\times\{-R^2\}} \tfrac{R^4\lambda_i^4}{2}\left|F^i_{t_0+\lambda_i^2 t}(x_0+\lambda_i x)\right|^2 G_{(0,0)}\,dV_x\,dt \\
&= \lim_{i\to\infty}\int_{\mathbb{R}^n\times\{t_0-R^2\lambda_i^2\}} \tfrac{(\lambda_i R)^4}{2\lambda_i^2}\left|F_t^i\right|^2 G_{z_0}\,dV_y\,ds
\end{aligned} \tag{5.2}$$

$$= \lim_{\lambda_i \to 0} \left[\int_{\mathbb{R}^n} (\lambda_i R)^4 \mu_t \Big|_{t=t_0 - R^2\lambda_i^2} \right]$$
$$= \Theta(\mu, z_0),$$

where the last line follows from Lemma 3.7. In particular, the Φ functional is approximately constant in R for the connections $\overline{\nabla}_t^i$, and hence using (3.3) we obtain the result. □

For $\Omega \subset \mathbb{R}^n \times \mathbb{R}$ we will use $\mu^* \lfloor \Omega$ to denote the restriction of the tangent measure to Ω.

Lemma 5.6. *For any $z_0 \in \Sigma$ and $\mu^* \in T_{z_0}(\mu)$, the quantity $\mu^* \lfloor \mathbb{R}^{n+1}_-$ is invariant under all parabolic dilation, i.e.*

$$\mathsf{P}_\kappa \left(\mu^* \lfloor \mathbb{R}^{n+1}_- \right) = \mu^* \lfloor \mathbb{R}^{n+1}_-.$$

Proof. First we observe that

$$\begin{aligned} \mathsf{P}_\kappa \left(\mu^* \lfloor \mathbb{R}^{n+1}_- \right) &= \mathsf{P}_\kappa \left(\left\{ \left(\mu_t^*, t \right) \mid t \in (-\infty, 0] \right\} \right) \\ &= \left\{ \left(\mathsf{D}_\kappa(\mu_t^*), \kappa^2 t \right) \mid t \in (-\infty, 0] \right\} \\ &= \left\{ \left(\mathsf{D}_\kappa \left(\mu^*_{\frac{t}{\kappa^2}} \right), \kappa^2 t \right) \mid t \in (-\infty, 0] \right\}. \end{aligned}$$

Thus, to prove the lemma it suffices to show that for all $\kappa < 0$, for all $t \in (-\infty, 0]$,

$$\mathsf{D}_\kappa \left(\mu^*_{\frac{t}{\kappa^2}} \right) = \mu_t^*.$$

Since κ is arbitrary this is equivalent to demonstrating this at $t = -1$. To prove this it suffices to show the result for μ_t^* multiplied by an arbitrary smooth positive function. We will take advantage of this by inserting a factor of the Greens function $G = G_{(0,0)}$, then multiplying by an arbitrary compactly supported positive function. This will allow us to take advantage of monotonicity formulae to obtain the result. In particular, we will show that

$$\kappa^{n-4} \int_{\mathbb{R}^n} \phi(\kappa x) G(\kappa x, -1)\, d\mu^*_{-\kappa^{-2}} = \int_{\mathbb{R}^n} \phi(x) G(x, -1)\, d\mu^*_{-1}, \tag{5.3}$$

for any $\phi \in C_0^1(\mathbb{R}^n)$. We attain the claim (5.3) if we can show that

$$\lim_{i \to \infty} \tfrac{d}{d\kappa} \left[\tfrac{\kappa^{n-4}}{2} \int_{\mathbb{R}^n} \phi(\kappa x) G(\kappa x, -1) \left| \overline{F}^i_{-\kappa^{-2}} \right|^2 dx \right] = 0. \tag{5.4}$$

For notational simplicity we will remove both the sequence index i and the bar from the connection. Manipulating the integrand by applying the change of coordinates $\kappa x = y$ yields,

$$\begin{aligned} &\tfrac{\kappa^{n-4}}{2} \int_{\mathbb{R}^n \times \{-1\}} \phi(\kappa x) \left| F_{-\kappa^{-2}}(x) \right|^2 G(\kappa x, -1)\, dx \\ &\quad = \tfrac{\kappa^{n-4}}{2} \int_{\mathbb{R}^n \times \{-1\}} \phi(y) \left| F_{-\kappa^{-2}}\left(\tfrac{y}{\kappa}\right) \right|^2 G(y, -1)\, d\left(\frac{y}{\kappa}\right) \\ &\quad = \tfrac{\kappa^{-4}}{2} \int_{\mathbb{R}^n \times \{-1\}} \phi(y) \left| F_{-\kappa^{-2}}\left(\tfrac{y}{\kappa}\right) \right|^2 G(y, -1)\, dy \\ &\quad = \left[\Phi\left(\tfrac{1}{\kappa}; \nabla_t \right) \right|_{t=-1}. \end{aligned}$$

Set $R(\kappa) := \frac{1}{\kappa}$. Then by a calculation similar to (3.1), where the final term vanishes since the cutoff function ϕ no longer depends on the parameter R, we see that

$$\begin{aligned}\frac{\partial}{\partial\kappa}\left[\Phi\left(\tfrac{1}{\kappa};\nabla_t\right)\right] &= \tfrac{-1}{\kappa^2}\tfrac{\partial}{\partial R}\left[\Phi\left(R(\kappa);\nabla_t\right)\right]\\ &= \tfrac{-1}{\kappa^5}\int\limits_{S_{\kappa-1}} |t|\left|\tfrac{x}{t}\lrcorner F - 2D^*F\right|^2 \phi G\,dx\\ &\quad + \tfrac{4}{\kappa^5}\int\limits_{S_{\kappa-1}} \left\langle\left(x\lrcorner F - 2t\left(D^*F\right)\right), \nabla\phi\lrcorner F\right\rangle G\,dx.\end{aligned}$$

Taking the limit as $i\to\infty$, we have that the first quantity vanishes by Lemma 5.5. For the second we apply weighted Hölder's inequality for an arbitrary $\epsilon>0$,

$$\begin{aligned}&\tfrac{1}{\kappa^5}\int\limits_{S_{\kappa-1}} \left\langle\left(x\lrcorner F - 2t\left(D^*F\right)\right), \nabla\phi\lrcorner F\right\rangle G\,dx\\ &\leq \frac{C}{\epsilon\kappa^5}\int\limits_{S_{\kappa-1}} \left|\left(x\lrcorner F - 2t\left(D^*F\right)\right)\right|^2 G\,dx + \frac{\epsilon}{\kappa^5}\int\limits_{S_{\kappa-1}} |\nabla\phi|^2|F|^2 G\,dx.\end{aligned}$$

The first factor vanishes with another application of Lemma 5.5. The integrand of the second term is bounded by the monotonicity of Φ, using an argument similar to (5.2). Sending $\epsilon\to 0$ therefore yields (5.4). The result follows. $\square$

5.3. *Stratification of tangent measures*

Lemma 5.7. *For $z_0\in\Sigma$ and $\mu^*\in T_{z_0}(\mu)$, the following hold.*

(1) *For all $z\in\mathbb{R}^{n+1}$, $\Theta(\mu^*,z)\leq\Theta(\mu^*,0)$.*
(2) *If $z\in\mathbb{R}^{n+1}$ satisfies $\Theta(\mu^*,z)=\Theta(\mu^*,0)$, then for all $\lambda>0$ and $v\in\mathbb{R}^{n+1}_-$,*

$$\Theta\left(\mu^*, z+v\right) = \Theta\left(\mu^*, z+\mathsf{P}_\lambda v\right).$$

Proof. For $\mu^*\in T_{z_0}(\mu)$, there exists some sequence $r_i\to 0$ such that $\mathsf{P}_{z_0,r_i}(\mu)\to\mu^*$. We first observe how the rescaling law for Φ is reflected in the definition of Θ. In particular, since we are integrating over a space slice we apply the scaling law for D_λ and change variables to yield

$$\begin{aligned}\Theta(\mathsf{P}_\lambda(\mu),z,r) &= \tfrac{r^4}{2}\int\limits_{S_r} G_z\mathsf{P}_\lambda(\mu)\\ &= \tfrac{r^4}{2}\int\limits_{S_r}\left(\lambda^n\mathsf{P}_\lambda^* G_{\mathsf{P}_\lambda(z)}\right)\left(\lambda^{4-n}\mathsf{P}_\lambda^*\mu\right)\\ &= \tfrac{(\lambda r)^4}{2}\int\limits_{\mathsf{P}_\lambda(S_r)} G_{\mathsf{P}_\lambda(z)}\mu\\ &= \Theta(\mu,\mathsf{P}_\lambda(z),\lambda r).\end{aligned}$$

Using this, for any $r>0$, and $z=(x,t)\in\mathbb{R}^{n+1}$,

$$\begin{aligned}\Theta\left(\mu^*, z\right) &\leq \Theta\left(\mu^*, z, r\right)\\ &= \lim_{r_i\to 0}\Theta\left(\mathsf{P}_{z_0,r_i}(\mu), z, r\right)\\ &= \lim_{r_i\to 0}\Theta\left(\mu, z_0+\left(r_i x, r_i^2 t\right), r_i r\right)\\ &\leq \Theta\left(\mu, z_0\right)\\ &= \Theta\left(\mu^*, 0\right),\end{aligned} \tag{5.5}$$

where we have applied the upper semicontinuity of $\Theta(\mu, \cdot, \cdot)$ with respect to the last two variables. Thus claim (1) follows.

To prove claim (2), observe that the hypothesis $\Theta(\mu^*, z) = \Theta(\mu^*, 0)$ implies that the inequalities of (5.5) are equalities. This implies that $\Theta\left(\mu^*, z, r\right) = \Theta(\mu, z_0)$, namely, it is constant with respect to r. By an argument similar to that of Lemma 5.6, we have that $\Theta\left(\mu^*, z+v\right) = \Theta\left(z+\mathsf{P}_\lambda(v)\right)$ for any $v \in \mathbb{R}^{n+1}_-$ and $\lambda > 0$. The result follows. □

Proposition 5.8. *For $z_0 \in \Sigma$ and $\mu^* \in T_{z_0}(\mu)$,*

$$V\left(\Theta\left(\mu^*, \cdot\right)\right) = W\left(\Theta\left(\mu^*, \cdot\right)\right).$$

In particular, both $V\left(\Theta\left(\mu^*, \cdot\right)\right)$ and $W\left(\Theta\left(\mu^*, \cdot\right)\right)$ are linear subspaces of $\mathbb{R}^n$. Moreover, $U\left(\Theta\left(\mu^*, \cdot\right)\right)$ is either $V\left(\Theta\left(\mu^*, \cdot\right)\right)$, or $V\left(\Theta\left(\mu^*, \cdot\right)\right) \times (-\infty, a]$ for some $0 \leq a \leq \infty$ and $\Theta\left(\mu^*, \cdot\right)$ is time-independent on $(-\infty, a]$.

Proof. First we show that $W(\Theta(\mu^*, \cdot)) \subset V(\Theta(\mu^*, \cdot))$. Fix $(x, 0) \in W\left(\Theta\left(\mu^*, \cdot\right)\right)$. Since the second component is identically zero it suffices to verify that $(x, 0) \in U\left(\Theta\left(\mu^*, \cdot\right)\right)$. Note that by definition of $W\left(\Theta\left(\mu^*, \cdot\right)\right)$, choosing $y = -x$ as in its definition,

$$\Theta\left(\mu^*, (x, 0)\right) = \Theta\left(\mu^*, (x-x, 0)\right) = \Theta\left(\mu^*, 0\right).$$

It follows that $W\left(\Theta\left(\mu^*, \cdot\right)\right) \subset V\left(\Theta\left(\mu^*, \cdot\right)\right)$.

Now we show the containment $V\left(\Theta\left(\mu^*, \cdot\right)\right) \subset W\left(\Theta\left(\mu^*, \cdot\right)\right)$. First note that $V\left(\Theta\left(\mu^*, \cdot\right)\right)$ is closed under scalar multiplications from Lemma 5.6. Next, for any nonzero $x \in V\left(\Theta\left(\mu^*, \cdot\right)\right)$ we have that for all $\lambda > 0$ and all $v \in \mathbb{R}^{n+1}_-$, by applying Lemma 5.7 (2), and using the parabolic scaling invariance of Θ from Lemma 5.6,

$$\begin{aligned}\Theta\left(\mu^*, (x, 0)+v\right) &= \Theta\left(\mu^*, (x, 0)+\mathsf{P}_\lambda v\right)\\ &= \Theta\left(\mu^*, \mathsf{P}_{\lambda^{-1}}\left((x, 0)+\mathsf{P}_\lambda v\right)\right)\\ &= \Theta\left(\mu^*, \mathsf{P}_{\lambda^{-1}}(x, 0)+v\right).\end{aligned} \tag{5.6}$$

By the upper semicontinuity of Θ, sending $\lambda \to \infty$ yields

$$\Theta(\mu^*, (x, 0)+v) \leq \Theta(\mu^*, v).$$

On the other hand, since $v - \mathsf{P}_{\lambda^{-1}}(x, 0) \in \mathbb{R}^{n+1}_-$, we can replace $v \mapsto v - \mathsf{P}_{\lambda^{-1}}(x, 0)$ throughout the equalities in (5.6) and obtain that

$$\Theta\left(\mu^*, (x, 0)+v-\mathsf{P}_{\lambda^{-1}}(x, 0)\right) = \Theta\left(\mu^*, v\right).$$

Again sending $\lambda \to \infty$ and utilizing the upper semicontinuity of $\Theta\left(\mu^*, \cdot\right)$ yields

$$\Theta(\mu^*, (x, 0)+v) \geq \Theta(\mu^*, v).$$

Hence we have $\Theta\left(\mu^*, v\right) = \Theta\left(\mu^*, (x, 0)+v\right)$, and so we conclude $V\left(\Theta\left(\mu^*, \cdot\right)\right) \subset W\left(\Theta\left(\mu^*, \cdot\right)\right)$ so that $V\left(\Theta\left(\mu^*, \cdot\right)\right) = W\left(\Theta\left(\mu^*, \cdot\right)\right)$.

Note that by definition of $W\left(\Theta\left(\mu^*, \cdot\right)\right)$ it is closed under linear combinations since for all $(x, 0)$, $(v, 0)$ in $W\left(\Theta\left(\mu^*, \cdot\right)\right)$ we have that for all $(y, s) \in \mathbb{R}^{n+1}_-$, just iterating its definition twice

$$\begin{aligned}\Theta\left(\mu^*, ((x+v)+y, s)\right) &= \Theta\left(\mu^*, (x+y, s)\right)\\ &= \Theta\left(\mu^*, (y, s)\right).\end{aligned}$$

Therefore by equality of $V(\Theta(\mu^*,\cdot))$ to $W(\Theta(\mu^*,\cdot))$, with the combined scaling invariance and linear combinations invariance both are linear subspaces of $\mathbb{R}^n$.

Now we prove the remaining statement of the proposition concerning the structure of $U(\Theta(\mu^*,\cdot))$. Suppose that $z := (x,t) \in U(\Theta(\mu^*,\cdot))$ with $t < 0$. Then for all $w := (y,s) \in \mathbb{R}^{n+1}$ with $s \leq t$ and for all $\lambda > 0$, using Lemma 5.7 (b)

$$\begin{aligned}\Theta\left(\mu^*, \mathsf{P}_{\lambda^{-1}}(w)\right) &= \Theta\left(\mu^*, w\right)\\ &= \Theta\left(\mu^*, z + w - z\right)\\ &= \Theta\left(\mu^*, z + \mathsf{P}_{\lambda^{-1}}(w - z)\right).\end{aligned} \tag{5.7}$$

In particular, take $\lambda \in (0,1)$, and note that consequently $\frac{s}{\lambda^2} \leq s \leq t$. So taking (5.7) and replacing $w \mapsto \mathsf{P}_{\lambda}(w)$ in yields

$$\Theta\left(\mu^*, w\right) := \Theta\left(\mu^*, z + w - \mathsf{P}_{\lambda^{-1}}(z)\right). \tag{5.8}$$

Taking $\lambda \to 0$, we see that $\Theta(\mu^*, w) \leq \Theta(\mu^*, z+w)$. Taking (5.8) again and instead replacing $w \mapsto w + \mathsf{P}_{\lambda^{-1}}(z)$, we conclude that

$$\Theta\left(\mu^*, w + \mathsf{P}_{\lambda^{-1}}(z)\right) = \Theta\left(\mu^*, z + w\right).$$

Again sending $\lambda \to 0$ we obtain that

$$\Theta\left(\mu^*, z\right) \leq \Theta\left(\mu^*, z + w\right).$$

We conclude that for any $z := (x,t) \in U(\Theta(\mu^*,\cdot))$ with $t < 0$, for all $w := (y,s)$ with $s \leq t$,

$$\Theta\left(\mu^*, w\right) = \Theta\left(\mu^*, z + w\right). \tag{5.9}$$

Then choosing $w \equiv z$, iterating (5.9), applying the parabolic scaling invariance of Θ from Lemma 5.6, and the upper semicontinuity of $\Theta(\mu^*,\cdot)$, one has

$$\begin{aligned}\Theta\left(\mu^*, 0\right) &= \Theta(\mu^*, z) = \Theta(\mu^*, z+z) = \cdots = \Theta(\mu^*, mz)\\ &= \Theta\left(\mu^*, (mx, mt)\right)\\ &= \Theta\left(\mu^*, \mathsf{P}_{\frac{1}{m}}(mx, mt)\right)\\ &= \Theta\left(\mu^*, \left(x, \tfrac{t}{m}\right)\right)\\ &\leq \Theta\left(\mu^*, (x, 0)\right).\end{aligned}$$

Combining this with Lemma 5.7 (1) we conclude that $(x,0) \in V(\Theta(\mu^*,\cdot)) = W(\Theta(\mu^*,\cdot))$. Therefore

$$\Theta\left(\mu^*, (0,t)\right) = \Theta(\mu^*, (x,0) + (0,t)) = \Theta(\mu^*, 0)$$

It follows that $(0,t) \in U(\Theta(\mu^*,\cdot))$. It follows that $\Theta(\mu^*,\cdot)$ is actually time independent for $t \leq 0$. Therefore for all $t \leq 0$,

$$V\left(\Theta\left(\mu^*,\cdot\right)\right) := U\left(\Theta\left(\mu^*,\cdot\right)\right) \cap \left(\mathbb{R}^n \times \{t\}\right).$$

Lastly, if $z = (x,t) \in U(\Theta(\mu^*,\cdot))$ with $t > 0$, then we can repeat the argument above to show that $\Theta(\mu^*,\cdot)$ is time-independent up to t. We set a to be the value of the maximal time $t \geq 0$ for which this time independence exists on. Then we have $U(\Theta(\mu^*,\cdot)) = V(\Theta(\mu^*,\cdot)) \times (-\infty, a]$, which concludes the proof. □

We can now establish Theorem 1.2, which we restate for convenience.

Theorem. *For* $0 \leq k \leq n-2$ *let*

$$\Sigma_k = \left\{z_0 \in \Sigma \mid \dim\left(\Theta\left(\mu^*,\cdot\right)\right) \leq k, \forall \mu^* \in T_{z_0}(\mu)\right\}.$$

Then $\dim_{\mathcal{P}}(\Sigma_k) \leq k$ *and* Σ_0 *is countable.*

Proof of Theorem 1.2. This is a direct consequence of ([23] Theorem 8.2). To connect directly to the notation of that paper, the function f is given by the density function. Hypothesis (1), the subsequential compactness of blowup limits, is established in Lemma 5.4. Hypothesis (2) is clear from the construction of blowup limits. Hypothesis (3), the parabolic scaling invariance of the limit functions, is established in Lemma 5.6. The theorem thus applies to give the claimed statement. □

6. Characterization of strong convergence

In this section we prove Theorem 1.3 (stated more precisely as Theorem 6.1 below), which characterizes when the weak convergence in $H^{1,2}$ for sequences as in Theorem 4.1 can be improved to strong convergence. In particular, we know this means that the defect measure is nontrivial, and we use this to obtain refined estimates on tangent measures, eventually leading to a further blowup sequence which yields the required Yang–Mills connection.

Theorem 6.1. *Suppose $\{\nabla_t^i\}$ is a sequence of smooth solutions to Yang–Mills flow on $[-1,0]$ with*

$$\sup_i \int_{M\times[-1,0]} \left(\left| \frac{\partial \nabla_t^i}{\partial t} \right|^2 + \left| F_{\nabla_t^i} \right|^2 \right) dV dt < \infty.$$

Furthermore, suppose $\{\nabla_t^i\} \to \nabla_t^\infty$ weakly in $H^{1,2}_{loc}$. Then exactly one of the following holds:

- *There exists a blowup sequence converging to a Yang–Mills connection on S^4.*
- *One has*

  $$\left| F_{\nabla_t^i} \right|^2 dV\, dt \to \left| F_{\nabla_t^\infty} \right|^2 dV\, dt$$

  as convergence of Radon measures, and hence $\{\nabla_t^i\} \to \nabla_t^\infty$ strongly in $H^{1,2}_{loc}$. Thus ∇_t^∞ is a weak solution of Yang–Mills flow satisfying $\mathcal{P}^{n-2}(\Sigma) = 0$.

Proof. We adopt the setup of the previous sections in this proof. In particular, we assume we have a particular blowup sequence together with a limiting tangent measure μ^*. Moreover, various results from §4.2 were established which apply to almost every point in the singular set. We will assume without loss of generality that our tangent measure arises from a blowup sequence around one of these points, so that the Lemmas of §4.2 apply. In particular, in the discussion below we will refer to a sequence $\{\nabla_t^i\}$ but this will refer to a *blowup* sequence, not the original given sequence of the statement.

Lemma 6.2. *For $t \in (-4, 0]$, we have $\mathcal{H}^{n-4}\left[\Sigma_t^*\right] > 0$.*

Proof. Suppose to the contrary there is some $t_0 \in (-4, 0]$ such that $\mathcal{H}^{n-4}(\Sigma_{t_0}^*) = 0$. Then for all $\epsilon > 0$, there exists some $\delta_\epsilon > 0$ and a covering of $\Sigma_{t_0}^*$ of the form $\{B_{r_j}(x_j)\}_{i\in\mathbb{N}}$, with $x \in \Sigma_{t_0}^*$ and $0 < r_j \leq \delta_\epsilon$ satisfying

$$\sum_{j=1}^{\infty} r_j^{n-4} < \epsilon.$$

Now, because $\mu_{t_0}^*\left[B_1 \setminus \left(\bigcup_{j\in\mathbb{N}} B_{r_j}(x_j)\right)\right] = 0$, then by a diagonalization argument we may choose a subsequence $\{\nabla_t^i\}$ such that

$$\lim_{i\to\infty} \frac{1}{2} \int_{B_1 \setminus \bigcup_{j\in\mathbb{N}} B_{r_j}(x_j)} \left| F_{\nabla_{t_0}^i} \right|^2 dV = 0. \tag{6.1}$$

Furthermore we will use (3.3) to estimate the curvature on balls in the cover. We choose a cutoff function ϕ for a ball of radius 1, and further fix some radius R. Note that for the compact set $\operatorname{supp}\phi$ there is a uniform estimate for the

L^2 norm of the Yang–Mills energy. This follows from the argument of Lemma 5.4, which shows that the sequence of blowup measures is uniformly locally finite. In particular, there exists $K < \infty$ such that

$$\int_{\operatorname{supp}\phi} \left|F_{\nabla^i_{t_0}}\right|^2 dV \leq K.$$

Hence using (3.3) we estimate for all $i, j \in \mathbb{N}$,

$$\begin{aligned}
\frac{r_j^{4-n}}{2} \int_{B_{r_j}(x_j)} \left|F_{\nabla^i_{t_0}}\right|^2 dV &\leq \tfrac{1}{2e} r_j^4 \int_{B_{r_j}(x_j)\times\{(t_0+r_j^2)-r_j^2\}} \left|F_{\nabla^i_t}\right|^2 \phi^2 G_{(x_i,t_0+r_j^2)} dV \\
&\leq \tfrac{1}{2e}\Phi\left(\nabla^i_t, \left(x_j, t_0+r_j^2\right), r_j\right) \\
&\leq \tfrac{1}{2e}\Phi\left(\nabla^i_t, \left(x_j, t_0+r_j^2\right), R\right) + CK\left(R-r_j\right) \\
&\leq \frac{R^{4-n}}{2e} \int_{\mathbb{R}^n\times\{t_0+r_j^2-R^2\}} \left|F_{\nabla^i_t}\right|^2 \phi^2\, dV + CK\left(R-r_j\right) \\
&\leq C(R,K).
\end{aligned}$$

Therefore we have that

$$\begin{aligned}
\frac{1}{2} \int_{\bigcup_{j\in\mathbb{N}} B_{r_j}(x_j)} \left|F_{\nabla^i_{t_0}}\right|^2 dV &\leq \frac{1}{2}\sum_{j\in\mathbb{N}} \int_{B_{r_j}(x_j)} \left|F_{\nabla^i_{t_0}}\right|^2 dV \\
&\leq C\sum_{j\in\mathbb{N}} r_j^{n-4} \\
&\leq C\epsilon.
\end{aligned}$$

Choosing $\epsilon < \frac{\epsilon_0}{16C}$ and combining with (6.1) yields, for i sufficiently large,

$$\int_{B_1} \left|F_{\nabla^i_{t_0}}\right|^2 dV \leq \frac{\epsilon_0}{2}. \tag{6.2}$$

Also, using Lemma 4.8 we have

$$\lim_{i\to\infty} \int_{P_2} \left|\frac{\partial \nabla^i_t}{\partial t}\right|^2 dV = 0. \tag{6.3}$$

Using Lemma 3.1 we find that for any $\phi \in C_0^\infty(B_2)$ and $-4 < t_1 < t_2 \leq 0$ one has

$$\frac{1}{2}\int_{B_2}\left(\left|F^i_{t_2}\right|^2 - \left|F^i_{t_1}\right|^2\right)\phi\, dV = -\int_{t_1}^{t_2}\int_{B_2}\left(\left|\partial_t \nabla^i_t\right|^2 \phi + \left\langle \nabla\phi \lrcorner F^i_t, \partial_t \nabla^i_t\right\rangle\right) dV\, dt. \tag{6.4}$$

Combining (6.3)–(6.4) shows that the limiting measure $\mu^*_t(\phi)$ is independent of time for $t \in (-4, 0]$. Applying (6.2) again yields

$$\begin{aligned}
\frac{1}{2}\int_{P_1}\left|F^i_t\right|^2 dV\, dt &\leq \frac{1}{2}\sup_{t\in[-1,0]}\int_{B_1}\left|F^i_t\right|^2 dV \\
&\leq \frac{1}{2}\int_{B_1}\left|F^i_{t_0}\right|^2 dV + o(i) \\
&\leq \frac{\epsilon_0}{4}.
\end{aligned}$$

This is a contradiction to the assumption that $(0,0) \in \Sigma^*$. $\square$

Proposition 6.3. *There is a linear subspace $\mathscr{P}$ of dimension $(n-4)$ such that for all $t<0$ one has that* $\operatorname{supp}(\mu_t^*)=\mathscr{P}$.

Proof. First, by Lemma 4.9, we have that $\Theta(\nu,z)=\Theta(\mu,z)$ is $\mathcal{P}^{n-2}$-approximately continuous at z_0 for $z\in\Sigma$ and $\Theta(\mu,z)$ is upper semicontinuous with respect to z, we conclude that for $\mathcal{P}^{n-2}$-almost all $z\in\Sigma^*$, we have

$$\Theta\left(\mu^*,z\right)\geq\Theta\left(\mu,z_0\right).$$

Also, it follows from Lemma 5.7 that

$$\Theta\left(\mu^*,z\right)\leq\Theta\left(\mu^*,0\right)=\Theta\left(\mu,z_0\right).$$

Hence for $\mathcal{P}^{n-2}$ a.e. $z\in\Sigma_*$,

$$\Theta\left(\mu^*,z\right)=\Theta\left(\mu^*,0\right). \tag{6.5}$$

We now show that in fact all points in Σ_* have maximal density. In particular, by Proposition 5.8 there is some set $\mathscr{S}\subset\mathbb{R}^n$ with $\mathcal{H}^{n-4}(\mathscr{S})=0$ and an $(n-4)$-dimensional plane $\mathscr{P}\subset\mathbb{R}^n$ such that $\mathscr{S}\cap\mathscr{P}=\varnothing$, and

$$\Sigma_t^*=\mathscr{S}\cap\mathscr{P}$$

for all t. We claim that in fact $\mathscr{S}=\varnothing$. Suppose to the contrary we had some $z\in\mathscr{S}$. Note that by construction, it must hold that $0<\Theta(\mu^*,z)<\Theta(\mu^*,0)$. By Lemma 5.7 (2) we have that for all $w\in\mathscr{P}$,

$$\Theta(\mu^*,z)=\Theta(\mu^*,w+(z-w))=\Theta\left(\mu^*,w+\mathsf{P}_\lambda(z-w)\right).$$

Applying this for $w\in B_\epsilon^{n-4}(0)\subset\mathscr{P}$ and $\lambda\in[1-\epsilon,1]$ yields a set of positive $\mathcal{P}^{n-2}$-measure in Σ^* on which $\Theta(\mu^*,\cdot)=\Theta(\mu^*,z)<\Theta(\mu^*,0)$, contradicting (6.5). □

Using this characterization of the singular set of the blowup limit, we can refine our estimates on the blowup sequence to obtain further structure on the blowup limit. Without loss of generality we can assume that $\mathscr{P}=\mathbb{R}^{n-4}\subset\mathbb{R}^n$ is the standard embedding in the first $n-4$ coordinates, and we express a general point as $X=(x,y)$ where $x\in\mathbb{R}^4$, $y\in\mathbb{R}^{n-4}$. We first show two lemmas which give improved vanishing results for the time derivative of the connection as well as for the curvature in directions along the singular locus.

Lemma 6.4. *Given the setup above and* $0<t_1<t_2\leq1$, *one has*

$$\lim_{i\to\infty}\int_{-t_2}^{-t_1}\int_{B_1^n}\left(\left|\frac{\partial\nabla_t^i}{\partial t}\right|^2+\sum_{j=1}^{n-4}\left|\frac{\partial}{\partial y_j}\lrcorner F_{\nabla_t^i}\right|^2\right)dV\,dt=0.$$

Proof. We first observe that by rescaling the result of Lemma 4.8 we observe that

$$\lim_{i\to\infty}\int_{-t_2}^{-t_1}\int_{B_1^n}\left|\frac{\partial\nabla_t^i}{\partial t}\right|^2 dV\,dt=0. \tag{6.6}$$

Now let $\xi_j=2\frac{\partial}{\partial y_j}$. Since we know that the limiting density Θ is a multiple of the Hausdorff measure of the given $\mathbb{R}^{n-4}$ on each time slice, applying the monotonicity formula (3.2) with centers $(X_0,t_0)=\left(\left(0,\xi_j\right),0\right)$ implies that for any $\rho>0$ we have

$$\begin{aligned}
0&=\lim_{i\to\infty}\int_\rho^1 r\int_M\int_{-4r^2}^{-r^2}\frac{\left|(X-\xi_j)\lrcorner F_t^i+2t\partial_t\nabla_t^i\right|^2}{|t|}\phi^2G_{0,\xi_j}\,dV\,dt\,dr\\
&\geq\lim_{i\to\infty}C_\rho\int_\rho^1 r\int_{B_1}\int_{-4r^2}^{-r^2}\frac{\left|(X-\xi_j)\lrcorner F_t^i+2t\partial_t\nabla_t^i\right|^2}{|t|}\,dV\,dt\,dr.
\end{aligned}$$

Note that the second equality follows since, for a given ρ, on B_1 we have that there is a constant $C_\rho \in (0,\infty)$ dependent solely on ρ such that $G_{0,\xi_j}\phi^2 \geq C_\rho \in (0,\infty)$.

Next we apply Fubini's theorem to switch the integration bounds $dr\,dt$ to $dt\,dr$. In that case we have that if we set

$$I_i(r,t) := C_\rho r \int_{B_1} \frac{\left|(X-\xi_j) \lrcorner F_{\nabla_t^i} + 2t\partial_t \nabla_t^i\right|^2}{|t|}\, dV$$

then applying Fubini's theorem to the regions corresponding to the variables r and t give that

$$0 = \lim_{i\to\infty} \left[\int_{-4\rho^2}^{-\rho^2} \int_{\rho}^{\sqrt{-t}} I_i(r,t)\,dr\,dt + \int_{-1}^{-4\rho^2} \int_{\frac{1}{4}\sqrt{-t}}^{\sqrt{-t}} I_i(r,t)\,dr\,dt + \int_{-4}^{-1} \int_{\frac{1}{4}\sqrt{-t}}^{1} I_i(r,t)\,dr\,dt \right].$$

Then for $0 < t_1 < t_2 < 1$, if we choose $\rho \leq t_1$, so that $[-t_2^2, -t_1^2] \subset \left[-1, -\rho^2\right]$ (the unions of the temporal domains of the first two integrals) then we can conclude that since the sums are 0 and the arguments of each integral are positive,

$$\begin{aligned} 0 &= \lim_{i\to\infty} \int_{-t_2^2}^{-t_1^2} \int_{B_1} \frac{\left|(X-\xi_j) \lrcorner F_{\nabla_t^i} + 2t\partial_t \nabla_t^i\right|^2}{|t|}\, dV\,dt \\ &\geq \lim_{i\to\infty} |t_1|^{-1} \int_{-t_2^2}^{-t_1^2} \int_{B_1} \tfrac{1}{2}\left|(X-\xi_j) \lrcorner F_{\nabla_t^i}\right|^2 - C\left|t\partial_t \nabla_t^i\right|^2 dV\,dt \\ &= \lim_{i\to\infty} |t_1|^{-1} \int_{-t_2^2}^{-t_1^2} \int_{B_1} \tfrac{1}{2}\left|(X-\xi_j) \lrcorner F_{\nabla_t^i}\right|^2 dV\,dt. \end{aligned}$$

The second inequality follows using the Cauchy–Schwarz inequality, and the final line follows from (6.6). Now observing that $\xi_j = 2\frac{\partial}{\partial y_j}$ we see that for all $X = (x,y) \in B_1$ we have that $\left|\left\langle (X-\xi_j), \frac{\partial}{\partial y_j}\right\rangle\right| \geq 1$. The result follows. □

Lemma 6.5. *Given the setup above, there exists* $(y,t) \in B_{1/2}^{n-4} \times [-\frac{1}{2}, -\frac{1}{4}]$ *such that*

$$\lim_{i\to\infty} \sup_{0<r\leq 1} r^{4-n} \int_{B_r^{n-4}(y)} \int_{B_1^4\times\{y\}\times[-1,0]} \left|\frac{\partial \nabla_t^i}{\partial t}\right|^2 dx\,dt\,dy = 0,$$

$$\lim_{i\to\infty} \sup_{0<r\leq 1} r^{2-n} \int_{B_r^{n-4}(y)\times[t-r^2,t]} \int_{B_1^4\times\{y\}} \sum_{j=1}^{n-4} \left|\frac{\partial}{\partial y_j} \lrcorner F_{\nabla_t^i}\right|^2 dx\,dt\,dy = 0.$$

Proof. To begin we show a preliminary statement using maximal functions. Let

$$f_i : B_1^{n-4} \subset \mathbb{R}^{n-4} \to [0,\infty), \qquad f_i(y) = \int_{\left(B_1^4\times\{y\}\right)\times[-1,0]} |\partial_t \nabla_i|^2\, dx\,dt$$

$$g_i : B_1^{n-4} \times \left[-1, -\tfrac{1}{8}\right] \to [0,\infty), \qquad g_i(y,t) = \int_{B_1^4\times\{y\}} \sum_{j=1}^{n-4} \left|\xi_j \lrcorner F_t^i\right|^2 dx.$$

Using these two quantities, we define two local Hardy–Littlewood maximal functions of f_i on B_1^{n-4} and g_i on $B_1^{n-4} \times \left[-1, -\frac{1}{8}\right]$ by, for $y \in B_1^{n-4}$,

$$M(f_i)(y) = \sup_{0 \le r \le 1} r^{4-n} \int_{B_r^{n-4}(y)} f_i \, dy.$$

Furthermore for $(y, t) \in B_1^{n-4} \times \left[-1, -\frac{1}{8}\right]$,

$$M(g_i)(y,t) = \sup_{0<r<1} r^{2-n} \int_{t-r^2}^{t} \int_{B_r^{n-4}(y)} g_i \, dy \, dt.$$

By applying Lemma 6.4 we can choose a subsequence such that $\int_{B_1^{n-4}} f_i dy \le 4^{-i}$. Combining this with the Hardy–Littlewood weak L^1 estimate we obtain a subsequence such that

$$\mu\{y \mid M(f_i) \ge 2^{-i}\} \le \frac{C}{2^{-i}} \int_{B_1^{n-4}} f_i \, dy \le C 2^{-i}.$$

In particular, for I chosen sufficiently large we have

$$\mu\left(\bigcup_{i \ge I} \{y \mid M(f_i) \ge 2^{-i}\}\right) \le C 2^{-I} < \frac{1}{2}\mu\left(B_1^{n-4}\right).$$

Thus $B_1^{n-4} \setminus \bigcup_{i \ge I} \{y | M(f_i) \ge 2^{-i}\}$ is nonempty, and any point y in that set satisfies

$$\lim_{i \to \infty} M(f_i)(y) = 0.$$

Combining this with an identical argument yields $(y, t) \in B_{1/2}^{n-4} \times [-\frac{1}{2}, -\frac{1}{4}]$ and a further subsequence such that

$$\lim_{i \to \infty} M(g_i)(y,t) = 0, \qquad \lim_{i \to \infty} M(f_i)(y) = 0.$$

The result follows. □

For the following Lemma we will suppress bundle indices for notational simplicity. Moreover, we will refer to coordinate directions $\frac{\partial}{\partial x^i}$ with unbarred indices, and $\frac{\partial}{\partial y^i}$ directions with barred indices. For an index which runs over both types of vectors we use I and J.

Lemma 6.6. *One has*

$$\begin{aligned} &\frac{\partial}{\partial y_k}\left[\int_{t-\delta_0^2}^{t} \int_{\mathbb{R}^4} \phi^2(x) \left[|F_\nabla|^2\right|_{(x,y,t)} dx \, dt\right] \\ &\quad = 4 \int_{t-\delta_0^2}^{t} \int_{\mathbb{R}^4} \left((\nabla_j \phi^2) F_{Ij} - \phi^2 \left(\tfrac{\partial \nabla}{\partial t}\right)_I\right) F_{\bar{k}I} \, dx \, dt - 4 \frac{\partial}{\partial y_j}\left[\int_{t-\delta_0^2}^{t} \int_{\mathbb{R}^4} \phi^2 \left(F_{I\bar{j}} F_{\bar{k}I}\right) dx \, dt\right]. \end{aligned}$$

Proof. With the notational conventions as described above we have

$$\begin{aligned} \nabla_{\bar{k}} (F_{IJ} F_{IJ}) &= 2 F_{IJ} \left(\nabla_{\bar{k}} F_{IJ}\right) \\ &= -2 F_{IJ} \left(\nabla_J F_{\bar{k}I} + \nabla_I F_{J\bar{k}}\right) \\ &= -4 F_{IJ} \left(\nabla_J F_{\bar{k}I}\right) \\ &= -4 \nabla_J \left(F_{IJ} F_{\bar{k}I}\right) - 4 \left(\nabla_J F_{JI}\right) F_{\bar{k}I}. \end{aligned} \tag{6.7}$$

Lastly, we expand out

$$|F|^2 = F_{IJ}F_{IJ} = F_{ij}F_{ij} + F_{i\bar{j}}F_{i\bar{j}} + F_{\bar{i}j}F_{\bar{i}j} + F_{\bar{i}\bar{j}}F_{\bar{i}\bar{j}}. \tag{6.8}$$

If we differentiate (6.8) and apply (6.7) to the resulting terms then this breaks down into

$$\begin{aligned}\nabla_{\bar{k}}\left[|F|^2\right] &= -4\nabla_j\left[F_{ij}F_{\bar{k}i}\right] - 4\nabla_{\bar{j}}\left[F_{\bar{i}\bar{j}}F_{\bar{k}\bar{i}}\right] - 4\nabla_{\bar{j}}\left[F_{i\bar{j}}F_{\bar{k}i}\right] - 4\nabla_j\left[F_{\bar{i}j}F_{\bar{k}\bar{i}}\right] \\ &\quad - 4\left(\nabla_j F_{ji}\right)F_{\bar{k}i} - 4\left(\nabla_{\bar{j}}F_{\bar{j}\bar{i}}\right)F_{\bar{k}\bar{i}} - 4\left(\nabla_{\bar{j}}F_{\bar{j}i}\right)F_{\bar{k}i} - 4\left(\nabla_j F_{j\bar{i}}\right)F_{\bar{k}\bar{i}} \\ &= -4\left(\nabla_j\left[F_{ij}F_{\bar{k}i} + F_{\bar{i}j}F_{\bar{k}\bar{i}}\right] + \nabla_{\bar{j}}\left[F_{\bar{i}\bar{j}}F_{\bar{k}\bar{i}} + F_{i\bar{j}}F_{\bar{k}i}\right]\right) + 4D^*_J F_{JI}F_{\bar{k}I}.\end{aligned} \tag{6.9}$$

With these pointwise quantities, we integrate (6.9) with a cutoff function ϕ and obtain

$$\begin{aligned}&\frac{\partial}{\partial y_k}\left[\int_{t-\delta_0^2}^{t}\int_{\mathbb{R}^4}\phi^2(x)\left[|F_\nabla|^2\Big|_{(x,y,t)}\right]dx\,dt\right] \\ &= -4\int_{t-\delta_0^2}^{t}\int_{\mathbb{R}^4}\phi^2\nabla_j\left[F_{ij}F_{\bar{k}i} + F_{\bar{i}j}F_{\bar{k}\bar{i}}\right]dx\,dt - 4\int_{t-\delta_0^2}^{t}\int_{\mathbb{R}^4}\phi^2\nabla_{\bar{j}}\left[F_{i\bar{j}}F_{\bar{k}i} + F_{\bar{i}\bar{j}}F_{\bar{k}\bar{i}}\right]dx\,dt \\ &\quad + 4\int_{t-\delta_0^2}^{t}\int_{\mathbb{R}^4}\phi^2\left(D^*_J F_{JI}\right)F_{\bar{k}I}\,dx\,dt \\ &= 4\int_{t-\delta_0^2}^{t}\int_{\mathbb{R}^4}\left(\nabla_j\phi^2\right)\left(F_{ij}F_{\bar{k}i} + F_{\bar{i}j}F_{\bar{k}\bar{i}}\right)dx\,dt - 4\frac{\partial}{\partial y_j}\left[\int_{t-\delta_0^2}^{t}\int_{\mathbb{R}^4}\phi^2\left(F_{i\bar{j}}F_{\bar{k}i} + F_{\bar{i}\bar{j}}F_{\bar{k}\bar{i}}\right)dx\,dt\right] \\ &\quad - 4\int_{t-\delta_0^2}^{t}\int_{\mathbb{R}^4}\phi^2\left(\partial_t\nabla_I\right)F_{\bar{k}I}\,dx\,dt \\ &= 4\int_{t-\delta_0^2}^{t}\int_{\mathbb{R}^4}\left((\nabla_j\phi^2)F_{Ij} - \phi^2\left(\tfrac{\partial\nabla}{\partial t}\right)_I\right)F_{\bar{k}I}\,dx\,dt - 4\frac{\partial}{\partial y_j}\left[\int_{t-\delta_0^2}^{t}\int_{\mathbb{R}^4}\phi^2\left(F_{I\bar{j}}F_{\bar{k}I}\right)dx\,dt\right],\end{aligned}$$

as required. □

We will use this lemma in conjunction with "Allard's strong constancy lemma," an effective version of the Divergence Theorem which we restate here for convenience.

Lemma 6.7. *([1] pp. 3) Suppose ψ, f, and Z are smooth on B_1 and satisfy*

$$\nabla\psi = f + \operatorname{div} Z$$

and

$$||f||_{L^1(B_1)} + ||Z||_{L^1(B_1)} \le \delta.$$

Then for all $\delta_1 > 0$, there is a $\delta_0 > 0$, depending on δ_1 and $||\psi||_{L^1(B_1)}$ such that, whenever $\delta \le \delta_0$,

$$\left|\left|\psi - \overline{\psi}\right|\right|_{L^1(B_1)} \le \delta_1,$$

where $\overline{\psi}$ denotes the average value of ψ on B_1.

Lemma 6.8. *Given a* $(y,t) \in B^{n-4}_{1/2} \times [-\frac{1}{2}, -\frac{1}{4}]$ *as in Lemma 6.5, there exists a universal constant* Λ *and sequences* $x_i \to 0$, $\delta_i \to 0$ *such that*

$$\begin{aligned}\frac{\epsilon_0}{\Lambda} &= \delta_i^{-2} \int\limits_{B^4_{\delta_i}(x_i)\times[t-\delta_i^2,t]} \left|F_{\nabla^i}\right|^2(x,y,t)\,dx\,dt \\ &= \max\left\{\delta_i^{-2} \int\limits_{B^4_{\delta_i}(\widetilde{x})\times[t-\delta_i^2,t]} \left|F_{\nabla^i}\right|^2(\widetilde{x},y,t)\,dx\,dt \mid \widetilde{x} \in B^4_{\frac{1}{2}}\right\}.\end{aligned}$$

Proof. Given (y,t), we fix some $\Lambda > 0$, then as each blowup connection ∇^i_t is smooth we may first choose a constant δ_i which is the smallest positive number such that

$$\max\left\{\delta_i^{-2} \int\limits_{B^4_{\delta_i}(\widetilde{x})\times[t-\delta_i^2,t]} \left|F^i\right|^2(\widetilde{x},y,t) \mid \widetilde{x} \in B^4_{\frac{1}{2}}\right\} = \frac{\epsilon_0}{\Lambda}. \tag{6.10}$$

Choosing x_i as some point in realizing the maximum defined above, all that remains to check is that $\delta_i \to 0$, $x_i \to 0$. First, suppose $\delta_i \geq \delta_0 > 0$. Fix $\phi \in C^\infty_0(B^4_{\delta_0}, [0,1])$, and let

$$\psi(y) := \delta_0^{-2} \int\limits_{t-\delta_0^2}^{t} \int\limits_{B^4_{\delta_0}} \phi^2 \left|F^i\right|^2(x,y,s)\,dx\,ds.$$

Now observe that the result of Lemma 6.6 can be interpreted as $\nabla\psi = f + \operatorname{div} Z$, with f and Z defined by the equality. It follows from Lemma 6.5 that

$$\lim_{i\to\infty} ||f||_{L^1(B^{n-4}_{\delta_0})} + ||Z||_{L^1(B^{n-4}_{\delta_0})} = 0.$$

Then we observe using Lemma 6.7 and (6.10) that

$$\begin{aligned}\lim_{i\to\infty} \delta_0^{2-n} \int\limits_{P_{\delta_0}((0,y),t)} \left|F^i_t\right|^2 dV\,dt = \lim_{i\to\infty} \overline{\psi} \\ &= \lim_{i\to\infty} \delta_0^{-4} \int\limits_{B^4_{\delta_0}} \overline{\psi}\,dx \\ &= \lim_{i\to\infty} \delta_0^{-4} \int\limits_{B^4_{\delta_0}} \left(\overline{\psi} - \psi + \psi\right) dx \\ &\leq \lim_{i\to\infty} \left[\delta_0^{-4} \left|\left|\psi - \overline{\psi}\right|\right|_{L^1(\mathbb{R}^4)} + \sup_{B^4_{\delta_0}} \psi\right] \\ &\leq \frac{\epsilon_0}{\Lambda}.\end{aligned} \tag{6.11}$$

This contradicts that $((0,y),t) \in \Sigma^*$, hence $\delta_i \to 0$. Now we note that the sequence $((x_i,y),t)$ develops concentration of $\left|F^i_t\right|$, and hence must limit to a singular point, which forces $x_i \to 0$. $\square$

With this sequence we can perform a further rescaling to finally obtain a Yang–Mills connection as blowup limit. In particular, define the blowup sequence

$$\widetilde{\Gamma}^i(x,y,t) = \delta_i \Gamma((x_i,y_i) + (\delta_i x, \delta_i y), t_i + \delta_i^2 t).$$

Let us observe some basic properties of this blowup sequence. In particular, by rescaling the estimates of Lemmas 6.5 and 6.8 we obtain

$$\begin{aligned}
\frac{\epsilon_0}{\Lambda} &= \int_{B_1^4 \times [-1,0]} \left|\widetilde{F}^i\right|^2 (0,0,t)\, dx\, dt = \max\left\{ \int_{B_1^4(\widetilde{x}) \times [-1,0]} \left|\widetilde{F}^i\right|^2 (x,0,t)\, dx\, dt \mid x \in \delta_i^{-1} B^4_{\frac{1}{2}} \right\}, \\
0 &= \lim_{i\to\infty} \sup_{r \in (0, \frac{1}{4\delta_i})} r^{4-n} \int_{B_r^{n-4}(0)} \int_{B^4_{\frac{1}{2\delta_i}} \times \{0\} \times [-\delta_i^{-2}, 0]} \left| \frac{\partial \widetilde{\nabla}^i}{\partial t} \right|^2 dx\, dt\, dy, \\
0 &= \lim_{i\to\infty} \sup_{r \in (0, \frac{1}{4\delta_i})} r^{2-n} \int_{B_r^{n-4}(0) \times [-r^2, 0]} \int_{B^4_{\frac{1}{2\delta_i}} \times \{y\}} \sum_{j=1}^{n-4} \left| \frac{\partial}{\partial y_j} \lrcorner \widetilde{F}_t^i \right|^2 dx\, dt\, dy.
\end{aligned} \tag{6.12}$$

Lemma 6.9. *The sequence $\{\widetilde{\nabla}_t^i\}$ converges strongly to a nonflat Yang–Mills connection on S^4.*

Proof. We use the estimates of (6.12) and argue as in the estimate (6.11) to show an energy estimate of the form

$$\int_{P_{\frac{3}{2}}((\widetilde{x},0),0)} \left|\widetilde{F}_t^i\right|^2 dV\, dt \leq \frac{\epsilon_0}{2} \quad \text{for all } \widetilde{x} \in \delta_i^{-1}\left(B^4_{\frac{1}{2}}\right). \tag{6.13}$$

Given this, we can complete the proof as follows. There is a local $H^{1,2}$ estimate for $\widetilde{\nabla}_t^i$ and hence we can choose a subsequence so that $\widetilde{\nabla}_t^i \to \widetilde{\nabla}_t^\infty$ weakly in $H^{1,2}_{\mathrm{loc}}(\mathbb{R}^n \times (-\infty, 0])$. However, using (6.12) we have that

$$\int_{\mathbb{R}^n \times (-\infty,0]} \left(\left| \frac{\partial \widetilde{\nabla}_t^\infty}{\partial t} \right|^2 + \sum_{j=1}^{n-4} \left| \frac{\partial}{\partial y_j} \lrcorner \widetilde{F}_t^\infty \right|^2 \right) dV\, dt = 0.$$

Using (6.13) and Theorem 3.8 we obtain convergence of $\widetilde{\nabla}_t^i$ to $\widetilde{\nabla}_t^\infty$ in $C^{k,\alpha}(K)$ for any compact set $K \subset \mathbb{R}^n \times (-\infty, 0]$. In particular, using (6.12) we obtain

$$\frac{\epsilon_0}{\Lambda} \leq \int_{\mathbb{R}^4} |\widetilde{F}_t^\infty|^2\, dx < \infty,$$

hence $\widetilde{\nabla}_t^\infty$ is not flat. The result follows. $\square$

Lemma 6.9 finishes the proof of the theorem. $\square$

Proof of Corollary 1.4. Without loss of generality by an overall rescaling we assume $T \geq 2$. Choose any sequence $\{t_i\} \to T$, and observe that the sequence of solutions given by restricting the given solution to $[t_i - 1, t_i]$ satisfies the hypotheses of Theorem 4.1. By hypothesis that T is maximal we know that $\Sigma \neq \varnothing$. As shown in Theorem 4.1 the point $z \in \Sigma$ is a point of entropy concentration. Thus we can choose a sequence of radii $r_i \to 0$ and rescale the parabolic balls $P_{r_i}(z_0)$ to unit size, to obtain a sequence of solutions with finite, nonzero entropy. It follows easily that the hypotheses of Theorem 4.1 hold for this sequence. If the sequence does not converge strongly in $H^{1,2}$, Theorem 1.3 yields the further blowup sequence which converges to a Yang–Mills connection on S^4. If this sequence does converge strongly in $H^{1,2}$, as the Ψ functional is becoming constant along the blowup sequence, the second term of the entropy monotonicity formula of (3.2) converges to zero, which implies that the blowup limit is a soliton. $\square$

References

[1] W. Allard, An integrality theorem and a regularity theorem for surfaces whose first variation with respect to a parametric elliptic integrals is controlled, in: Proceeding of Symp. in Pure Math., vol. 44, AMS, 1986, 1–2.

[2] Y. Chen, C. Shen, Monotonicity formula and small action regularity for Yang–Mills flows in higher dimensions, Calc. Var. Partial Differ. Equ. 2 (4) (1994) 389–403.

[3] Z. Chen, Y. Zhang, Stabilities of homothetically shrinking Yang–Mills solitons, arXiv:1410.5150 [math.DG].

[4] S.K. Donaldson, P.B. Kronheimer, The Geometry of Four-Manifolds, Oxford Mathematical Monographs, The Clarendon Press/Oxford University Press, New York, 1990, Oxford Science Publications, MR1079726 (92a:57036).

[5] P. Feehan, Global existence and convergence of smooth solutions to Yang–Mills gradient flow over compact four-manifolds, arXiv:1409.1525 [math.DG].

[6] H. Federer, W.P. Ziemer, The Lebesgue set of a function whose distribution derivatives are p-th power summable, Indiana Univ. Math. J. 22 (1972/1973) 139–158.

[7] A. Gastel, Singularities of first kind in the harmonic map and Yang–Mills heat flow, Math. Z. 242 (2002) 47–62.

[8] R. Hamilton, Monotonicity formulas for parabolic flows on manifolds, Commun. Anal. Geom. 1 (1) (1993) 127–137.

[9] M.C. Hong, G. Tian, Asymptotical behaviour of the Yang–Mills flow and singular Yang–Mills connections, Math. Ann. 330 (2004) 441–472.

[10] T. Ilmanen, Singularities of mean curvature flow of surfaces, preprint, 1995.

[11] C. Kelleher, J. Streets, Entropy, stability, and Yang–Mills flow, arXiv:1410.4547.

[12] F.H. Lin, C.Y. Wang, Harmonic and quasi-harmonic spheres, Commun. Anal. Geom. 7 (2) (1999) 397–429.

[13] F.H. Lin, C.Y. Wang, Harmonic and quasi-harmonic spheres, part II, Commun. Anal. Geom. 10 (2002) 341–375.

[14] F.H. Lin, C.Y. Wang, Harmonic and quasi-harmonic spheres, part III. Rectifiability of the parabolic defect measure and generalized varifold flows, Ann. l'inst. Henri Poincaré C, Anal. Non Linéaire 19 (2) (2002) 209–259.

[15] H. Naito, Finite time blowing-up for the Yang–Mills gradient flow in higher dimensions, Hokkaido Math. J. 23 (3) (1994) 451–464, https://doi.org/10.14492/hokmj/1381413099, MR 1299637 (95i:58054).

[16] J. Råde, On the Yang–Mills heat equation in two and three dimensions, J. Reine Angew. Math. 431 (1992) 123–163.

[17] L. Simon, Lectures on Geometric Measure Theory, Proc. Center for Math. Anal., vol. 3, Australian Nat. Univ. Press, Canberra, 1983.

[18] M. Struwe, The Yang–Mills flow in four dimensions, Calc. Var. Partial Differ. Equ. 2 (2) (1994) 123–150, https://doi.org/10.1007/BF01191339.

[19] G. Tian, Gauge theory and calibrated geometry I, Ann. Math. Second Ser. 151 (1) (2000) 193–268.

[20] K. Uhlenbeck, Connections with L^p-bounds on curvature, Commun. Math. Phys. 83 (1982) 31–42.

[21] A. Waldron, Instantons and singularities in the Yang–Mills flow, arXiv:1402.3224.

[22] B. Weinkove, Singularity formation in the Yang–Mills flow, Calc. Var. Partial Differ. Equ. 19 (2) (2004) 211–220.

[23] B. White, Stratification of minimal surfaces, mean curvature flows, and harmonic maps, J. Reine Angew. Math. 488 (1997) 1–35.