J. Math. Pures Appl. 81 (2002) 143–166

A probabilistic approach to the Yang–Mills heat equation

Marc Arnaudon [a], Robert O. Bauer [b], Anton Thalmaier [c]

[a] *Département de Mathématiques, Université de Poitiers, Téléport 2, BP 30179, F-86962 Futuroscope Chasseneuil cedex, France*
[b] *Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801-2975, USA*
[c] *Institut für Angewandte Mathematik, Universität Bonn, Wegelerstraße 6, D-53115 Bonn, Germany*

Received 1 December 2001

Abstract

We construct a parallel transport U in a vector bundle E, along the paths of a Brownian motion in the underlying manifold, with respect to a time dependent covariant derivative ∇ on E, and consider the covariant derivative $\nabla_0 U$ of the parallel transport with respect to perturbations of the Brownian motion. We show that the vertical part $U^{-1}\nabla_0 U$ of this covariant derivative has quadratic variation twice the Yang–Mills energy density (i.e., the square norm of the curvature 2-form) integrated along the Brownian motion, and that the drift of such processes vanishes if and only if ∇ solves the Yang–Mills heat equation. A monotonicity property for the quadratic variation of $U^{-1}\nabla_0 U$ is given, both in terms of change of time and in terms of scaling of $U^{-1}\nabla_0 U$. This allows us to find a priori energy bounds for solutions to the Yang–Mills heat equation, as well as criteria for non-explosion given in terms of this quadratic variation. © 2002 Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

1. Introduction, notations

This article is concerned with the Yang–Mills heat equation for connections in a metric vector bundle E over a compact Riemannian manifold M. The Yang–Mills connections in E are critical points of the Yang–Mills functional (or energy functional)

$$\mathrm{YM}(\nabla) := \int_M \|R^\nabla\|^2 \,\mathrm{d}\,\mathrm{vol}, \tag{1.1}$$

where $R^\nabla \in \Gamma(\Lambda^2 T^*M \otimes \mathrm{End}(E))$ is the curvature 2-form of a metric connection ∇ in E. Letting ∇ depend smoothly on a real parameter t and differentiating Eq. (1.1) with respect to t yields

$$\partial_t \,\mathrm{YM}\big(\nabla(t)\big) = 2\int_M \big\langle (\mathrm{d}^\nabla)^* R^{\nabla(t)}, \partial_t \nabla(t)\big\rangle \,\mathrm{d}\,\mathrm{vol},$$

where d^∇ denotes the exterior differential and $(\mathrm{d}^\nabla)^*$ its adjoint, see, e.g., [8]. Consequently, to deform a connection towards the steepest descent of the Yang–Mills action we are led to solve the Yang–Mills heat equation

$$\partial_t \nabla(t) = -\frac{1}{2}(\mathrm{d}^\nabla)^* R^{\nabla(t)}. \tag{1.2}$$

E-mail addresses: arnaudon@mathlabo.univ-poitiers.fr (M. Arnaudon), rbauer@math.uiuc.edu (R.O. Bauer), anton@wiener.iam.uni-bonn.de (A. Thalmaier).

PII: S0021-7824(02)01254-0

The Euler–Lagrange equations associated to the Yang–Mills functional characterize Yang–Mills connections ∇ by the property that

$$\left(\mathrm{d}^{\nabla}\right)^{*} R^{\nabla} \equiv 0. \tag{1.3}$$

The procedure of constructing Yang–Mills connections by starting from an arbitrary connection $\nabla(0)$, solving the Yang–Mills heat equation with initial condition $\nabla(0)$, letting t tend to ∞, taking a subsequence which converges (up to global gauge transformations) to a Yang–Mills connection requires that Eq. (1.2) does not blow-up in finite time. Small time existence is well-known, e.g., [25]: there always exists $T > 0$ such that Eq. (1.2) has a solution in $[0, T[$. If M is of dimension less than or equal to 3, then blow-up never occurs, as proved in [19]. In the general situation, blow-up at time T is characterized by the fact that curvature does not stay bounded in $]T - \varepsilon, T[$ for any $\varepsilon > 0$. If M has dimension at least 4, one is led to look for non-explosion criteria. An essential ingredient in [9] is the integral of $\|R^{\nabla}\|^2$ over "parabolic balls" of the form $[T - 4s^2, T - s^2] \times \mathrm{ball}(x, R_0) \subset \mathbb{R}_+ \times M$ where $R_0 > 0$ is fixed and sufficiently small, along with monotonicity in s of this integral (see [17,18] for related integrals). In [9,17] the authors prove that if the integral is sufficiently close to zero for small s then $\|R^{\nabla}\|^2$ is bounded on some space–time set. Such estimates lead to non-explosion criteria in certain situations.

More can be said when M has dimension 4, since curvature then can concentrate at time T only at finitely many points in M (see [26]). At every point where curvature concentrates, rescaling of space and time, along with gauge transformation, yields a non-trivial Yang–Mills connection in a bundle over $\mathbb{R}^4$ with the same fiber as E (see [23]).

The paper is organized as follows. In Section 2 we construct the main stochastic object which is the starting point of our study. Let us briefly describe the set-up. Consider a smooth solution $\nabla(t)$ to Eq. (1.2) on a vector bundle E over a compact Riemannian manifold M, defined on $[0, T[$ for some $T > 0$. Fix $x \in M$ and let X_t be a Brownian motion starting from x. For $u \in T_x M$ let $X_t(a, u) = \exp_{X_t}(a\sqrt{t}\, //_{0,t} u)$ where $//_{0,t}$ denotes parallel transport in TM along X_t with respect to the Levi-Civita connection. The basic object of our study is the semimartingale

$$s \mapsto N_{r,s} := U_{r,s}^{-1} \nabla_a|_{a=0} U_{r,s}, \quad s \in [r, T],$$

where $r \in]0, T[$ is fixed and $s \mapsto U_{r,s}(a, u)$ is the parallel transport in E along X_s starting at $U_{r,r} = \mathrm{id}_{E_{X_r}}$, with respect to the time-dependent connection $\nabla(T - s)$. The semimartingale $N_{r,s}$ has nice scaling properties (Remarks 2.1 and 3.1) and we prove that it is a local martingale if and only if $\nabla(t)$ solves the Yang–Mills heat equation (Lemma 2.2 and Proposition 2.3). See [6] for similar results in the stationary case of Yang–Mills connections. As expected, a mean value formula holds (Corollary 2.5): for $0 < r \leqslant s \leqslant T$ and at X_r we have

$$\nabla_{\sqrt{r}\, //_{0,r} u}(T - r)(\cdot) = \mathbb{E}\left[U_{r,s}^{-1} \nabla_{\sqrt{s}\, //_{0,s} u}(T - s)(U_{r,s} \cdot) \,\middle|\, \mathcal{F}_r\right].$$

Unfortunately, as a consequence of non-linearity, the conditional expectation contains the parallel transport $U_{r,s}$ involving the $\nabla(t)$, $t \in]T - s, T - r[$. The approach is analogous to the probabilistic interpretation of the heat equation for harmonic maps (see, e.g., [2]).

As mentioned before, an important object in the study of singularities at (T, x) of the Yang–Mills heat flow is the energy integral over parabolic balls and monotonicity properties of the integral. In Section 3 we work instead with the expected quadratic variation of the martingale, constructed in Section 2, on the time interval $[\beta s, s]$ where $0 < \beta < 1$ is some parameter, that is

$$\Phi_\beta(s) := \frac{1}{2} \mathbb{E}\left[\|N_{\beta s, s}\|^2\right].$$

Using a matrix Harnack estimate for positive solutions of the heat equation on a manifold (see [12]), we establish monotonicity in s of $\Phi_\beta(s)$. Our result differs from [12]. The method here is also slightly different from the one used in [9], but more natural in our context. As a consequence of the monotonicity formula we establish convergence of

$$\frac{1}{2\log(T/r)} \mathbb{E}\left[\|N_{r,T}\|^2\right]$$

as r tends to 0 (Proposition 3.9). The limit which we call here $\ell(T, x)$ will play a crucial role in the description of singularities.

The key result in Section 4 says that if $\Phi_\beta(s)$ is sufficiently close to zero for small s, then the energy is bounded by $Cf(s)^{-4}$ on $[T - f(s)^2, T[\times \mathrm{ball}(x, f(s))$ for some $C > 0$ and some positive increasing function f, defined for $s > 0$, such that $\lim_{s \to 0} f(s) = 0$ (Theorems 4.2, 4.3 and Corollary 4.4). A similar result but with different assumptions can be found in [9]. Our proof is based on a submartingale inequality, which is an alternative to Moser's Harnack inequality. We use estimates of the exit time from small balls for Brownian motion in M, based on Bernstein's inequality (Lemma 4.1).

From Theorems 4.2 and 4.3 we derive a non-explosion criterion for the Yang–Mills heat flow at (T, x) in terms of the size of Φ_β, as well as in terms of $\ell(T, x)$ (Proposition 4.5 and Corollary 4.6). We then establish existence of global solutions to Eq. (1.2) in case $\mathrm{YM}(\nabla(0))$ is sufficiently small; we already know that the solution exists on $[0, T[$ (Theorems 4.7, 4.10, Corollary 4.11). In particular, this gives non-explosion if M has dimension less than or equal to 3, a result due to [19]. When M is of dimension at least 4, an other consequence is that if explosion occurs at a small time T, then the Yang–Mills energy $\mathrm{YM}(\nabla(0))$ must be greater than some positive constant depending on T. In the special case when M is a sphere of dimension greater than 4 and E a non-trivial bundle over M, we are able to recover Naito's result [17]: if $\mathrm{YM}(\nabla(0))$ is smaller than some positive number depending on $t > 0$, then explosion occurs before time t (Corollary 4.12).

In Section 5 we consider a solution ∇ to the Yang–Mills heat equation on $[0, T[$ and assume that explosion occurs at time T. We exhibit a sequence of martingales, as in Section 2, constructed from the rescaled Yang–Mills equation, which converges in law to a non-trivial martingale (Proposition 5.2). In case of $\dim M = 4$ we know by [26] that curvature concentrates only in a finite number of points in the manifold, and our result can be seen as the stochastic analogue to Schlatter's result [24] on convergence of rescaled connections to a connection in a vector bundle over $\mathbb{R}^4$. However, here we cannot choose the point $x \in M$ where curvature concentrates since the support of our function Φ_β is M and not a small ball.

Section 6 of the paper finally is devoted to an ergodic theorem. Here we assume that $\dim M = 4$. We fix a Yang–Mills connection ∇ on E and prove that the Pontryagin number of the vector bundle is the ergodic mean of an expression involving the curvature of ∇ along Brownian paths (Theorem 6.1). If ∇ is self-dual (respectively antiself-dual), the expression is the quadratic variation (respectively minus the quadratic variation) of the martingale $N_{r,s}$ constructed in Section 2 (for the stationary case $\nabla(t) \equiv \nabla$).

Throughout the paper we adopt the following conventions. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geqslant 0}, \mathbb{P})$ be a filtered probability space on which all the considered processes will be defined. Let M be a manifold and let $\pi : E \to M$ be a vector bundle over M. By a covariant derivative or connection on the vector bundle E we mean an $\mathbb{R}$-linear map

$$\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E)$$

satisfying the product rule

$$\nabla(fX) = \mathrm{d}f \otimes X + f\nabla X, \quad X \in \Gamma(E), \quad f \in C^\infty(M),$$

and by a connection on the manifold M we mean a covariant derivative or connection on the vector bundle TM. All covariant derivatives (on various vector bundles) are denoted indifferently by ∇.

A covariant derivative on E gives rise to a splitting $TE = HE \oplus VE$ into the horizontal and the vertical bundle. If $e \in E_x$, we denote by $h_e : T_xM \to H_eE$ the horizontal lift and by $v_e : E_x \to V_eE$ the vertical lift. We denote by X^h the horizontal lift in $\Gamma(TE)$ of a vector field X in $\Gamma(TM)$, and by r^v the vertical lift in $\Gamma(TE)$ of a section r of E (an element of $\Gamma(E)$). Given a connection ∇ on M and a covariant derivative ∇ on E, there exists a unique connection ∇^h on E such that for $X, Y \in \Gamma(TM)$, $r, s \in \Gamma(E)$,

$$\nabla^h_{r^v} s^v = 0, \qquad \nabla^h_{r^v} Y^h = 0, \qquad \nabla^h_{X^h} s^v = (\nabla_X s)^v, \qquad \nabla^h_{X^h} Y^h = (\nabla_X Y)^h \tag{1.4}$$

(e.g., [3]). Let F be a vector space isomorphic to the typical fiber of E. A covariant derivative ∇ on E gives rise canonically to a covariant derivative on the vector bundle $\pi : \mathrm{Hom}(F, E) \to M$ of linear maps $F \to E_x$, again denoted ∇, and defined as follows: if W is a section of $\mathrm{Hom}(F, E)$ then $(\nabla W)(w) = \nabla(W(w))$, $w \in F$.

Assume that the manifold M is endowed with a Riemannian metric $(\cdot, \cdot)$ and that the vector bundle $\pi : E \to M$ is endowed with a metric preserved by the covariant derivative ∇. Let $\mathcal{A}^p(E) = \Gamma(\bigwedge^p T^*M \otimes E)$ be the p-forms on M with values in the vector bundle E and

$$\mathcal{A}(E) = \Gamma\Big(\bigwedge T^*M \otimes E\Big) = \bigoplus_{p \geqslant 0} \mathcal{A}^p(E).$$

The covariant derivative ∇ on E gives rise to a "differential" $\mathrm{d}^\nabla : \mathcal{A}(E) \to \mathcal{A}(E)$, which sends $\mathcal{A}^p(E)$ into $\mathcal{A}^{p+1}(E)$, defined by

$$\mathrm{d}^\nabla a(v_1, \ldots, v_{p+1}) = \sum_{i=1}^{p+1} (-1)^{i+1} (\nabla_{v_i} a)(v_1, \ldots, \hat{v}_i, \ldots, v_{p+1}),$$

where $a \in \mathcal{A}^p(E)$ and $v_1, \ldots, v_{p+1} \in T_xM$. Alternatively d^∇ is given at point $x \in M$ by

$$\mathrm{d}^\nabla a(x) = \sum_{i=1}^m (e_i, \cdot) \wedge \nabla_{e_i} a,$$

where $(e_i)_{1 \leqslant i \leqslant m}$ is an orthonormal frame in T_xM.

We consider also the co-differential $(\mathrm{d}^\nabla)^*:\mathcal{A}(E)\to\mathcal{A}(E)$,

$$\big(\mathrm{d}^\nabla\big)^* a(x) = -\sum_{i=1}^m \nabla_{e_i} a(e_i,\ldots),$$

where again $(e_i)_{1\leqslant i\leqslant m}$ is an orthonormal frame in T_xM. The image of $\mathcal{A}^p(E)$ under $(\mathrm{d}^\nabla)^*$ now lies in $\mathcal{A}^{p-1}(E)$.

Let $a\in\mathcal{A}^{p-1}(E)$ and $b\in\mathcal{A}^p(E)$ be such that $a\otimes b$ is of compact support. Then the following formula holds (e.g., [10], Lemma 2.16, Eqs. (8.8) and (8.9)):

$$\int_M \big\langle \mathrm{d}^\nabla a, b\big\rangle\,\mathrm{d}x = \int_M \big\langle a, \big(\mathrm{d}^\nabla\big)^* b\big\rangle\,\mathrm{d}x. \tag{1.5}$$

When M is the Euclidean space $\mathbb{R}^m$, then Eq. (1.5) is still valid under the assumption that both $\langle \mathrm{d}^\nabla a, b\rangle$ and $\langle a, (\mathrm{d}^\nabla)^* b\rangle$, as well as $X=\alpha^\sharp$, are in $L^1(\mathbb{R}^m)$ where α denotes the 1-form $v\mapsto\langle (v,\cdot)\wedge a, b\rangle$. Indeed, we have

$$\mathrm{div}(X) = \big\langle \mathrm{d}^\nabla a, b\big\rangle - \big\langle a, \big(\mathrm{d}^\nabla\big)^* b\big\rangle,$$

and one easily shows that if both a vector field Y on $\mathbb{R}^m$ and its divergence $\mathrm{div}(Y)$ are L^1, then

$$\int_{\mathbb{R}^m} \mathrm{div}(Y)\,\mathrm{d}y = 0.$$

In Section 4 we shall adopt the following assumption on (a,b,M) guaranteeing Eq. (1.5) to hold.

Assumption 1.1. Either $a\otimes b$ has compact support, or if $M=\mathbb{R}^m$ then $\alpha^\sharp$, $\langle \mathrm{d}^\nabla a, b\rangle$ and $\langle a, (\mathrm{d}^\nabla)^* b\rangle$ are in $L^1(\mathbb{R}^m)$, where $\alpha^\sharp$ denotes the vector field associated to the 1-form $\alpha: v\mapsto\langle (v,\cdot)\wedge a, b\rangle$.

Let M be a manifold with connection ∇. If X is an M-valued continuous semimartingale and α a section of T^*M, we denote by $\int_0\langle\alpha,\delta X\rangle$ the Stratonovich integral of α along X, and by $\int_0\langle\alpha,\mathrm{d}^\nabla_{\text{Itô}}X\rangle$ the Itô integral. Recall that X is a ∇-martingale if and only if $\int_0\langle\alpha,\mathrm{d}^\nabla_{\text{Itô}}X\rangle$ is a local martingale for every such α. In local coordinates, we have

$$\langle\alpha,\delta X\rangle = \sum_i\bigg(\alpha_i(X)\,\mathrm{d}X^i + \frac{1}{2}\sum_j \frac{\partial\alpha_i}{\partial x^j}(X)\,\mathrm{d}\big\langle X^i, X^j\big\rangle\bigg) \quad \text{and}$$

$$\big\langle\alpha,\mathrm{d}^\nabla_{\text{Itô}}X\big\rangle = \sum_i \alpha_i(X)\bigg(\mathrm{d}X^i + \frac{1}{2}\sum_j \Gamma^i_{jk}(X)\,\mathrm{d}\big\langle X^j, X^k\big\rangle\bigg),$$

where Γ^i_{jk} are the Christoffel symbols of ∇. Given a covariant derivative ∇ on E, the parallel transport in E along X is the $\mathrm{Hom}(E_{X_0},E)$-valued semimartingale $//^E_{0,\boldsymbol{\cdot}}$ defined by

$$//^E_{0,0} = \mathrm{id}_{E_{X_0}} \quad \text{and} \quad \delta//^E_{0,t} = h_{//^E_{0,t}}(\delta X_t).$$

Note that $//^E_{0,t}\in\mathrm{Hom}(E_{X_0},E_{X_t})$. An equivalent definition for $//^E_{0,\boldsymbol{\cdot}}$ is

$$//^E_{0,0} = \mathrm{id}_{E_{X_0}} \quad \text{and} \quad \mathrm{d}^{\nabla^h}_{\text{Itô}}//^E_{0,t} = h_{//^E_{0,t}}\big(\mathrm{d}^\nabla_{\text{Itô}}X_t\big),$$

see [11] for details.

In case $E=TM$, with ∇ the Levi-Civita connection induced by the Riemannian metric on M and $//_{0,t}$ the parallel transport in TM along X, there is an alternative definition for $\mathrm{d}^\nabla_{\text{Itô}}X_t$ in terms of

$$\mathrm{d}^\nabla_{\text{Itô}}X_t = //_{0,t}\,\mathrm{d}\bigg(\int_0^t //_{0,s}^{-1}\,\delta X_s\bigg),$$

where $\mathrm{d}(\int_0^t //_{0,s}^{-1}\,\delta X_s)$ is the usual Itô differential of the process $\int_0^t //_{0,s}^{-1}\,\delta X_s$ with values in the vector space E_{X_0}.

For an E-valued semimartingale J we define the Itô covariant derivative DJ of J as the vertical part of $\mathrm{d}^{\nabla^h}_{\text{Itô}}J$, considered as an element of E:

$$DJ = v_J^{-1}\big(\big(\mathrm{d}^{\nabla^h}_{\text{Itô}}J\big)^{\text{vert}}\big).$$

Alternatively, DJ may be expressed as

$$DJ_t = //_{0,t}^{E}\, \mathrm{d}\big(//_{0,t}^{E\;-1} J\big),$$

where $\mathrm{d}(//_{0,t}^{E\;-1}J)$ is again the usual Itô differential of the E_{X_0}-valued semimartingale $(//_{0,t}^{E})^{-1}J$. In local coordinates, writing the covariant derivatives ∇ on E, respectively ∇ on M, as $\mathrm{d}+A$ and $\mathrm{d}+\Gamma$ where A and Γ are 1-forms taking values in $\mathrm{End}(E)$, respectively $\mathrm{End}(TM)$, the following general formula for $(DJ)^\alpha$ holds (see [3]):

$$(DJ)^\alpha = \mathrm{d}J^\alpha + A^\alpha\big(\mathrm{d}_{\text{Itô}}^{\nabla}X, J\big) + A^\alpha(\mathrm{d}X, \mathrm{d}J) + \frac{1}{2}\Big(\mathrm{d}A^\alpha(\mathrm{d}X, \mathrm{d}X, J) + A^\alpha\big(\mathrm{d}X, A(\mathrm{d}X, J)\big) - A^\alpha\big(\Gamma(\mathrm{d}X, \mathrm{d}X), J\big)\Big). \tag{1.6}$$

Let $(J(a))_{a\in I}$ be a family of semimartingales in E indexed by an open interval I in $\mathbb{R}$ about 0, which is C^1 in a with respect to the topology of semimartingales (see [1]). We denote by $\nabla_a J$ the covariant derivative of J with respect to a. In the sequel let $\nabla_0 = \nabla_a|_{a=0}$, $\partial_0 = \partial_a|_{a=0}$; finally, R^∇ denotes the curvature tensor associated to ∇. The following formula has been proved in [3] Theorem 4.5:

Theorem 1.2. *The Itô covariant derivative of $\nabla_0 J$ is given by the formula*

$$D\nabla_0 J = \nabla_0 DJ + R^\nabla\big(\mathrm{d}_{\text{Itô}}^{\nabla}X, \partial_0 X\big)J + R^\nabla(\mathrm{d}X, \partial_0 X)DJ - \frac{1}{2}\nabla R^\nabla(\mathrm{d}X, \partial_0 X, \mathrm{d}X)J - \frac{1}{2}R^\nabla(D\partial_0 X, \mathrm{d}X)J.$$

2. A martingale description of the Yang–Mills heat equation

Let M be a Brownian complete Riemannian manifold endowed with the Levi-Civita connection ∇. Fix $T>0$ and $I=[0,T[$. Let $\pi : E\to M$ be a vector bundle with a metric preserved under a smooth family of covariant derivatives $\nabla(t)$, $t\in I$. Let $\tilde\pi : \widetilde{E}\to I\times M$ be the vector bundle over $I\times M$ with fiber $\widetilde{E}_{(t,x)} = E_x$. The family $\nabla(t)$, $t\in I$, induces canonically a covariant derivative $\widetilde\nabla$ on $\widetilde{E}$ as follows: if $t\mapsto u(t)$ is a smooth path in $\widetilde{E}$ with projection $t\mapsto (f(t), x(t))$ in $I\times M$, then

$$\big(\widetilde\nabla_D u\big)(t) = \big(\nabla_D\big(f(t)\big)u\big)(t). \tag{2.1}$$

It is easy to prove that $\widetilde\nabla$ is compatible with the metric in $\widetilde{E}$ inherited from the metric in E.

Let X be a Brownian motion on M starting from x, and denote by $//_{0,\cdot}$ the parallel translation along X. For $a\in\mathbb{R}$ close to 0 and $u\in T_xM$, we define a perturbation of the Brownian paths as follows:

$$X_s(a,u) = \exp_{X_s}\big(a\sqrt{s}\,//_{0,s}u\big). \tag{2.2}$$

The factor $\sqrt{s}$ in Eq. (2.2) is justified by the scaling property explained in the following remark.

Remark 2.1. Let M be the Euclidean space $\mathbb{R}^m$, $x=0$ and $c>0$. The rescaled perturbed Brownian motion $(cX_s(a,u))_{s\geqslant 0}$ has the same law as $(X_{c^2 s}(a,u))_{s\geqslant 0}$. For a general manifold M, suppose that X solves an Itô equation of the type

$$\mathrm{d}_{\text{Itô}}^{\nabla}X = A(X)\,\mathrm{d}B, \quad X_0\equiv x\in M,$$

where $A\in\Gamma(\mathbb{R}^m\otimes TM)$ is such that $A(x)A(x)^* = \mathrm{id}_{T_xM}$ for all $x\in M$. Here B denotes an $\mathbb{R}^m$-valued Brownian motion (m is not necessarily equal to $\dim M$). Defining the rescaled Brownian motion X^c by

$$\mathrm{d}_{\text{Itô}}^{\nabla}X^c = cA\big(X^c\big)\,\mathrm{d}B, \quad X_0^c\equiv x,$$

and the rescaled perturbed Brownian motion by $X_s^c(a,u) = \exp_{X_s^c}(c\sqrt{s}\,//_{0,s}^c au)$ with $//_{0,s}^c$ denoting parallel transport in TM along X_s^c, we have again a scaling property in the sense that $(X_s^c(a,u))_{s\geqslant 0}$ and $(X_{c^2 s}(a,u))_{s\geqslant 0}$ are equal in law.

Now fix $0<r<T$. The parallel transport in $\widetilde{E}$ along $(T-s, X_s(a,u))$, $s\in[r,T]$, will be denoted $\widetilde{W}_{r,s}^{(T,x)}(a,u)$, or simply $\widetilde{W}_{r,s}(a,u)$. By definition, $\widetilde{W}_{r,s}(a,u)$ takes its values in $\mathrm{Hom}(\widetilde{E}_{(T-r,X_r(a,u))}, \widetilde{E}_{(T-s,X_s(a,u))})$ and is determined by

$$\widetilde{D}\widetilde{W}_{r,s}(a,u) = 0 \quad\text{and}\quad \widetilde{W}_{r,r}(a,u) = \mathrm{id}_{E_{X_r}},$$

where $\widetilde{D}$ is the covariant derivative in s with respect to $\widetilde\nabla$. Let $U_{r,s}^{(T,x)}(a,u)$, or simply $U_{r,s}(a,u)$, denote the $\mathrm{Hom}(E_{X_r(a,u)}, E_{X_s(a,u)})$-valued process given by $\widetilde{W}_{r,s}(a,u)$. Then

$$D(T-s)U_{r,s}(a,u) := \widetilde{D}\widetilde{W}_{r,s}(a,u) = 0 \quad\text{and}\quad U_{r,r}(a,u) = \mathrm{id}_{E_{X_r(a,u)}}, \tag{2.3}$$

where $D(t-s)$ is the covariant derivative in s with respect to $\nabla^E(t-s)$ (in local coordinates $D(T-s)J_s^\alpha$ is given by Eq. (1.6) with A^α replaced by $A^\alpha(T-s)$). Then, almost surely, $U_{r,s}(a,u)$ is an isometry for all s, a, u.

We write $U_{r,s}$ or $U_{r,s}(a)$ for $U_{r,s}(a,u) \equiv U_{r,s}^{(T,x)}(a,u)$. Given a C^1 path $a \mapsto v(a)$ in E, let $\nabla_a(t)v$ be the $\nabla(t)$-covariant derivative of v and $\nabla_0(t)v = \nabla_a(t)|_{a=0}v$. We define the ∇-covariant derivative $\nabla_0 U_{r,s}$ of $U_{r,s}$ with respect to a, at $a=0$, as follows: if $(v(a)) = (v(\omega,a))$ is an $\mathcal{F}_r$-measurable random variable taking values in the C^1 paths in E which project to $a \mapsto X_r(a,u)$, then

$$(\nabla_0 U_{r,s})v(0) := \nabla_0(T-s)(U_{r,s}v) - U_{r,s}(0)\big(\nabla_0(T-r)v\big). \tag{2.4}$$

In other words, letting $\tilde v(0)$ be the element $v(0)$ in $\widetilde{E}_{(T-r,X_r)}$, we have

$$(\nabla_0 U_{r,s})v(0) = \big(\widetilde\nabla_0 \widetilde W_{r,s}\big)\tilde v(0),$$

where $\widetilde\nabla_0 \widetilde W_{r,s}$ takes its values in $\mathrm{Hom}(\widetilde E_{(T-r,X_r)}, \widetilde E_{(T-s,X_s)})$ and is the covariant derivative of $a \mapsto \widetilde W_{r,s}(a,u)$ at $a=0$ with respect to the canonical connection $\widetilde\nabla^{(\widetilde E)^*\otimes\widetilde E}$ in $\widetilde E^* \otimes \widetilde E$ induced by $\widetilde\nabla$.

Lemma 2.2. *Let $0<r<T$. The covariant derivative $D(T-s)\nabla_0 U_{r,s}$ is equal to*

$$R^\nabla\big(\mathrm{d}^\nabla_{\text{Itô}} X_s, \sqrt{s}\,/\!/_{0,s}\,\cdot\big)U_{r,s} - \left(\frac12(\mathrm{d}^\nabla)^* R^\nabla\big(\sqrt{s}\,/\!/_{0,s}\,\cdot\big)U_{r,s} + \partial_t\nabla\big(\sqrt{s}\,/\!/_{0,s}\cdot, U_{r,s}\big)\right)\mathrm{d}s, \tag{2.5}$$

where R^∇, $(\mathrm{d}^\nabla)^*R^\nabla$ and $\partial_t\nabla$ are taken at time $T-s$.

The drift of $U_{r,s}^{-1}\nabla_0 U_{r,s}$ is equal to

$$-\int_r \left(\frac12 U_{r,s}^{-1}\big(\mathrm{d}^{\nabla(T-s)}\big)^* R^{\nabla(T-s)}(/\!/_{0,s}\cdot)U_{r,s} + U_{r,s}^{-1}\partial_t\nabla(T-s)(/\!/_{0,s}, U_{r,s})\right)\sqrt{s}\,\mathrm{d}s. \tag{2.6}$$

The Riemannian quadratic variation $S_{r,s} = S_{r,s}^{(T,x)}$ of the process $U_{r,s}^{-1}\nabla_0 U_{r,s}$ with values in $\mathrm{Hom}(T_{X_0}M, \mathrm{End}(E_{X_r}))$ *satisfies*

$$S_{r,s} = \int_r^s 2\rho\big\|R^{\nabla(T-\rho)}(X_\rho)\big\|^2\,\mathrm{d}\rho. \tag{2.7}$$

Proof. Using $D(T-s)\nabla_0 U_{r,s} = \widetilde D\widetilde\nabla_0\widetilde W_{r,s}$ and applying Theorem 1.2, along with the fact that $\widetilde D\widetilde W_{r,s} = 0$, $D\partial_0 X_s = (1/(2\sqrt{s}))\,/\!/_{0,s}\,\mathrm{d}s$ (here ∂_0 stands for $\partial_a|_{a=0}$) which gives $D\partial_0 X_s \wedge \mathrm{d}X_s = 0$, we arrive at

$$\widetilde D\widetilde\nabla_0\widetilde W_{r,s} = R^{\widetilde\nabla}\big((-\mathrm{d}s, \mathrm{d}^\nabla_{\text{Itô}}X_s), (0, \sqrt{s}\,/\!/_{0,s}\,\cdot)\big)\widetilde W_{r,s} - \frac12\widetilde\nabla R^{\widetilde\nabla}\big((-\mathrm{d}s, \mathrm{d}X_s), (0,\sqrt{s}\,/\!/_{0,s}), (-\mathrm{d}s, \mathrm{d}X_s)\big)\widetilde W_{r,s}.$$

One then easily verifies that

$$R^{\widetilde\nabla}\big((-\mathrm{d}s, \mathrm{d}^\nabla_{\text{Itô}}X_s), (0,\sqrt{s}\,/\!/_{0,s})\big)\widetilde W_{r,s} = R^{\nabla(T-s)}\big(\mathrm{d}^\nabla_{\text{Itô}}X_s, \sqrt{s}\,/\!/_{0,s}\big)U_{r,s} - \partial_t\nabla(T-s)\big(\sqrt{s}\,/\!/_{0,s}, U_{r,s}\big)\,\mathrm{d}s$$

and

$$\widetilde\nabla R^{\widetilde\nabla}\big((-\mathrm{d}s, \mathrm{d}X_s), (0,\sqrt{s}\,/\!/_{0,s}), (-\mathrm{d}s,\mathrm{d}X_s)\big)\widetilde W_{r,s} = \big(\mathrm{d}^\nabla\big)^* R^{\nabla(T-s)}\big(\sqrt{s}\,/\!/_{0,s}\,\cdot\big)U_{r,s}$$

which gives (2.5). Formulas (2.6) and (2.7) are then direct consequences of (2.5), taking into account that $U_{r,s}^{-1}\nabla_0 U_{r,s} = \widetilde W_{r,s}^{-1}\widetilde\nabla_0\widetilde W_{r,s}$. To establish Eq. (2.7) we use in addition the fact that both $U_{r,s}$ and $/\!/_{0,s}$ are isometries. □

An immediate consequence of Lemma 2.2 is the following proposition. As above let $U_{r,s} = U_{r,s}^{(T,x)}$. Again $\nabla_0 U_{r,s}$ denotes the ∇-covariant derivative of $U_{r,s}$ with respect to the parameter a at $a=0$ (compare with [6] Theorem 4.4 where a criterion for a fixed covariant derivative to be Yang–Mills is given).

Proposition 2.3. *Let $x\in M$, $T>0$ and $I=[0,T[$. Let $(\nabla(t))_{t\in I}$ be a smooth family of connections in E. The following two statements are equivalent*:

(i) $\nabla(t)$ *is a solution to the heat equation on $I\times M$,*

$$\partial_t\nabla = -\frac12\big(\mathrm{d}^\nabla\big)^* R^\nabla; \tag{2.8}$$

(ii) $((U_{r,s})^{-1}\nabla_0 U_{r,s})_{s\in[r,T]}$ *is a local martingale for every $r\in\,]0,T[$.*

Proof. If $\nabla(t)$ satisfies Eq. (2.8) then the expression in (2.6) clearly vanishes, hence, (i) implies (ii). To prove the other direction, assume that (ii) holds and fix $r \in]0, T[$. Since $((U_{r,s}^{(T,x)})^{-1}\nabla_0 U_{r,s}^{(T,x)})_{s\in[r,T]}$ is a local martingale, the expression (2.6) vanishes, and, hence, by continuity of the integrand in (2.6), almost surely for every $s \in [r, T]$,

$$U_{r,s}^{-1}\partial_t\nabla(T-s)(//_{0,s}, U_{r,s}) = -\frac{1}{2}U_{r,s}^{-1}\big(\mathrm{d}^{\nabla(T-s)}\big)^* R^{\nabla(T-s)}(//_{0,s}\cdot)U_{r,s}.$$

Taking $s = r$, along with the fact that X_r has a positive density in M and the continuity of $\partial_t\nabla(T-r)$ and $(\mathrm{d}^{\nabla(T-r)})^* R^{\nabla(T-r)}$, we get

$$\big(\partial_t\nabla(T-r)\big) = -\frac{1}{2}\big(\big(\mathrm{d}^{\nabla(T-r)}\big)^* R^{\nabla(T-r)}\big).$$

This holds true for all $r \in]0, T[$, and, hence, for all $r \in]0, T]$ by continuity. Thus (i) is verified and the proof is complete. □

Remark 2.4. (1) Let $W_s(t) = W_s(a, t, u)$ be the parallel translation in E along the semimartingale $X_s(a, u)$, with respect to a fixed covariant derivative $\nabla(t)$. For $0 < r < T$ let

$$W_{r,s}(t) = W_s(t)\big(W_r(t)\big)^{-1}.$$

One can prove that

$$U_{r,s} = W_{r,s}(T-s)\,\mathcal{E}^R\left(\int_r^{\cdot} W_{r,\rho}^{-1}(T-\rho)\,\partial_t W(T-\rho)\,\mathrm{d}\rho\right)_s, \tag{2.9}$$

where $\mathcal{E}^R$ is the right stochastic exponential in $\mathrm{Gl}(E_{X_r})$. Recall that given a semimartingale $(Y_s)_{s\geqslant r}$ with values in $\mathfrak{gl}(E_{X_r})$, then $\mathcal{E}^R(Y)$ is the solution to

$$\delta Z_s = \delta Y_s \cdot Z_s, \qquad Z_r = \mathrm{id}_{E_{X_r}}.$$

An alternative, but less simple proof of Lemma 2.2 without introducing $\widetilde{E}$ would have been possible by taking Eq. (2.9) as definition.

(2) Yang–Mills connections have been characterized in terms of the holonomy along Brownian loops or Brownian bridges in [4,7,22]. Here we characterize solutions to the Yang–Mills heat equation by the fact that $(U_{r,s}^{(T,x)})^{-1}\nabla_0 U_{r,s}^{(T,x)}$ is a martingale. This may be considered as an infinitesimal holonomy characterization, since our martingale is the derivative with respect to a at $a = 0$ of the holonomy around the following loop: we start at X_r to go to $X_r(a, u)$ along the minimizing geodesic, from there we go to $X_s(a, u)$ along the path of $X(a, u)$, then to X_s along the minimizing geodesic, and finally back to X_r via the backwards path of X. Denoting for fixed $t \in]0, T]$ by $\tau_t^{(T,x)}(a, u)$ the parallel transport with respect to the connection $\nabla(T-t)$ from E_{X_t} to $E_{X_t}(a, u)$ along the minimizing geodesic, then we have

$$\big(U_{r,s}^{(T,x)}\big)^{-1}\nabla_0 U_{r,s}^{(T,x)} = \frac{\mathrm{d}}{\mathrm{d}a}\bigg|_{a=0}\big(\big(U_{r,s}^{(T,x)}\big)^{-1}\big(\tau_s^{(T,x)}(a,u)\big)^{-1} U_{r,s}^{(T,x)}(a,u)\,\tau_r^{(T,x)}(a,u)\big). \tag{2.10}$$

To finish this section we give a mean value formula for solutions to the Yang–Mills heat equation. Again $\nabla_0(T-s)$ stands for $\nabla_a(T-s)|_{a=0}$.

Corollary 2.5. *Let ∇ be a solution to the Yang–Mills heat equation* (2.8) *on $I \times M$ and let $0 < r < s \leqslant T$. For every $\mathcal{F}_r$-measurable path $a \mapsto v(a)$ in $E_{X_r(a,u)}$ which is a.s. C^1 in a (for instance, $v(a) = e_{X_r(a,u)}$ where e is a C^1 section of E), we have*

$$\nabla_0(T-r)v = \mathbb{E}\big[U_{r,s}^{-1}\nabla_0(T-s)(U_{r,s}v)\;\big|\;\mathcal{F}_r\big]. \tag{2.11}$$

Proof. By definition, we have for $s \in [r, T]$,

$$U_{r,s}^{-1}\nabla_0(T-s)(U_{r,s}v) = U_{r,s}^{-1}(\nabla_0 U_{r,s})v + \nabla_0(T-r)v.$$

By Proposition 2.3 and Lemma 2.2, the right-hand side is a square integrable martingale in s which takes the value $\nabla_0(T-r)v$ at time $s = r$. This gives the claim. □

3. Monotonicity properties related to variations of stochastic parallel transport

In this section, M is either a compact Riemannian manifold, or $M=\mathbb{R}^m$ ($m\geqslant 1$). We consider solutions $\nabla(t)$ to

$$\partial_t\nabla=-\frac{1}{2}\big(\mathrm{d}^{\nabla}\big)^*R^{\nabla} \tag{3.1}$$

on $I\times M$ where $I=[0,T[$ for some $T>0$. Keeping $0<r<T$ fixed, we are interested in monotonicity properties of the quadratic variation $(S_{r,s})_{s\in[r,T]}$ to the local martingale $(U_{r,s}^{-1}\nabla_0 U_{r,s})_{s\in[r,T]}$ as defined in Section 2. More precisely, letting $0<\beta<1$ and

$$\Phi_\beta^{(T,x)}:]0,T]\to\mathbb{R},\quad s\mapsto\mathbb{E}\left[\int_{\beta s}^{s}\rho\,\big\|R^{\nabla(T-\rho)}(X_\rho)\big\|^2\,\mathrm{d}\rho\right]=\frac{1}{2}\mathbb{E}\Big[S_{\beta s,s}^{(T,x)}\Big],$$

we look for conditions on M which insure that $\Phi_\beta^{(T,x)}$ is non-decreasing, or more generally, that there exists a constant $C>0$ depending only on M such that for all $0<s_1<s_2<T$,

$$\Phi_\beta^{(T,x)}(s_1)\leqslant C\Big(\Phi_\beta^{(T,x)}(s_2)+(s_2-s_1)\,\mathrm{YM}\big(\nabla(0)\big)\Big). \tag{3.2}$$

For $0<s<T$ let

$$\varphi^{(T,x)}(s)=\mathbb{E}\big[\big\|R^{\nabla(T-s)}(X_s)\big\|^2\big].$$

We then have $\Phi_\beta^{(T,x)}(s)=\int_{\beta s}^{s}r\varphi^{(T,x)}(r)\,\mathrm{d}r$. For simplicity, write $\Phi_\beta=\Phi_\beta^{(T,x)}$ and $\varphi=\varphi^{(T,x)}$ in the sequel.

Remark 3.1. Let $0<c\leqslant 1$. Along the rescaled perturbed Brownian motion $X_s^c(a,u)$, introduced in Remark 2.1, we define a transport $U_{r,s}^c(a,u)$ by the equation

$$D\big(T-c^2s\big)U_{r,s}^c(a,u)=0,\quad U_{r,r}^c(a,u)=\mathrm{id}_{E_{X_r^c(a,u)}}. \tag{3.3}$$

Denoting by $S_{r,s}^c$ the Riemannian quadratic variation of $(U_{r,s}^c)^{-1}\nabla_0^c U_{r,s}^c$ (where $\nabla^c(T-t)=\nabla(T-c^2t)$) and $\Phi_\beta(c,s)=(1/2)\mathbb{E}[S_{\beta s,s}^c]$, one easily verifies the relation

$$\Phi_\beta(c,s)=\Phi_\beta\big(c^2s\big).$$

As a consequence, all monotonicity results for $s\mapsto\Phi_\beta(s)$ can be interpreted in terms of $c\mapsto\Phi_\beta(c,s)$, for a fixed s.

Lemma 3.2. *Let $t\in]0,T[$ be fixed. The following two statements are equivalent*:

(i) *for each $\beta\in]0,1[$, the function $s\mapsto\Phi_\beta(s)$ is non-decreasing on $]0,t]$;*
(ii) *the function $s\mapsto s^2\varphi(s)$ is non-decreasing on $]0,t]$.*

Proof. We follow the proof of Lemma 9.2 in [27]. If (i) is satisfied, then $s\Phi_\beta'(s)\geqslant 0$ for all $0<\beta<1$ and $0<s<t$, which yields

$$s^2\varphi(s)-\beta^2s^2\varphi(\beta s)\geqslant 0.$$

This gives (ii). Conversely, assuming that (ii) is satisfied, let $0<s_1<s_2<t$ and $\lambda:=s_2/s_1>1$. Then

$$\Phi_\beta(s_1)=\int_{\beta s_1}^{s_1}r_1\varphi(r_1)\,\mathrm{d}r_1=\int_{\beta s_2}^{s_2}\frac{1}{r_2}\frac{r_2^2}{\lambda^2}\varphi\left(\frac{r_2}{\lambda}\right)\mathrm{d}r_2\leqslant\int_{\beta s_2}^{s_2}r_2\varphi(r_2)\,\mathrm{d}r_2=\Phi_\beta(s_2), \tag{3.4}$$

where (ii) has been used for the inequality in (3.4). This achieves the proof. □

Lemma 3.2 brings us to the study of the function

$$\phi^{(T,x)}=\phi:]0,T[\to\mathbb{R},\quad s\mapsto s^2\varphi(s). \tag{3.5}$$

For $s\in]0,T[$ let g_s be the density at time s of Brownian motion X started at x. Thus

$$\frac{\mathrm{d}}{\mathrm{d}s}g_s=\frac{1}{2}\Delta g_s=-\frac{1}{2}\mathrm{d}^*\mathrm{d}g_s. \tag{3.6}$$

We shall adopt the following assumption.

Assumption 3.3. For every $b \in]0, T[$, $\|R^\nabla\|$, $\|\nabla R^\nabla\|$ and $\|\nabla\nabla R^\nabla\|$ are in $L^2(M)$, where $\nabla = \nabla(b)$ denotes the covariant derivative at b on E, respectively the induced covariant derivative on $\bigwedge^2 T^*M \otimes E$ and $T^*M \otimes \bigwedge^2 T^*M \otimes E$.

Note that Assumption 3.3 is automatically satisfied if M is compact. As in the proof of Theorem 9.1 in [27], but with an additional term coming from the curvature of M, the following monotonicity formula holds.

Proposition 3.4. *Let Assumption* 3.3 *be satisfied. For any* $s \in]0, T[$, *we have*

$$\begin{aligned}\phi'(s) = {}& s^2 \int_M \big\|\big(\mathrm{d}^{\nabla(T-s)}\big)^* R^{\nabla(T-s)} - \iota_{\operatorname{grad}\log g_s} R^{\nabla(T-s)}\big\|^2 g_s \,\mathrm{d}x \\ &+ 2s \int_M \big\langle R^{\nabla(T-s)}, R^{\nabla(T-s)} + s R^{\nabla(T-s)} \circ \big(\nabla(\operatorname{grad}\log g_s) \odot \mathrm{id}\big)\big\rangle g_s \,\mathrm{d}x,\end{aligned} \tag{3.7}$$

where

$$\big(\nabla(\operatorname{grad}\log g_s) \odot \mathrm{id}\big)(u \wedge v) = \frac{1}{2}\big(\nabla_u(\operatorname{grad}\log g_s) \wedge v + u \wedge \nabla_v(\operatorname{grad}\log g_s)\big).$$

Proof. We shall make use of the integration by parts formula, Eq. (1.5), at different stages of the proof. In these cases Assumption 3.3 will imply Assumption 1.1.

First, we are going to compute $\varphi'(s)$. For the sake of brevity, we write R^∇ for $R^{\nabla(T-s)}$, ∇ for $\nabla(T-s)$, d^∇ for $\mathrm{d}^{\nabla(T-s)}$, and $(\mathrm{d}^\nabla)^*$ for $(\mathrm{d}^{\nabla(T-s)})^*$. It then follows from $R^\nabla = \mathrm{d}^\nabla \circ \nabla$ and Eq. (3.1) that

$$\frac{\mathrm{d}}{\mathrm{d}s} R^\nabla = \mathrm{d}^\nabla \frac{\mathrm{d}}{\mathrm{d}s}\nabla = -\frac{1}{2}\mathrm{d}^\nabla\big(\mathrm{d}^\nabla\big)^* R^\nabla, \tag{3.8}$$

see, e.g., [26], formula (10). Now since

$$\varphi(s) = \int_M \|R^\nabla\|^2 g_s \,\mathrm{d}x,$$

we have

$$\varphi'(s) = 2\int_M \left\langle \frac{\mathrm{d}}{\mathrm{d}s} R^\nabla, R^\nabla \right\rangle g_s \,\mathrm{d}x + \int_M \langle R^\nabla, R^\nabla\rangle \frac{\mathrm{d}}{\mathrm{d}s} g_s \,\mathrm{d}x. \tag{3.9}$$

By the respective heat equations (3.8) and (3.6), this equals

$$\int_M \big\langle \mathrm{d}^\nabla\big(\mathrm{d}^\nabla\big)^* R^\nabla, R^\nabla g_s\big\rangle \,\mathrm{d}x - \frac{1}{2}\int_M \langle R^\nabla, R^\nabla\rangle \,\mathrm{d}^*\mathrm{d}g_s \,\mathrm{d}x \tag{3.10}$$

and, by applying Eq. (1.5), we get

$$\int_M \big\langle \big(\mathrm{d}^\nabla\big)^* R^\nabla, \big(\mathrm{d}^\nabla\big)^*\big(R^\nabla g_s\big)\big\rangle \,\mathrm{d}x - \frac{1}{2}\int_M \big(\mathrm{d}\langle R^\nabla, R^\nabla\rangle, \mathrm{d}g_s\big) \,\mathrm{d}x. \tag{3.11}$$

Taking into account that $(\mathrm{d}^\nabla)^*(R^\nabla g_s) = ((\mathrm{d}^\nabla)^* R^\nabla) g_s - \iota_{\operatorname{grad} g_s} R^\nabla$, expression (3.11) may be written as

$$\int_M \big\langle \big(\mathrm{d}^\nabla\big)^* R^\nabla, \big(\mathrm{d}^\nabla\big)^* R^\nabla\big\rangle g_s \,\mathrm{d}x - \int_M \big\langle \big(\mathrm{d}^\nabla\big)^* R^\nabla, \iota_{\operatorname{grad}\log g_s} R^\nabla\big\rangle g_s \,\mathrm{d}x - \frac{1}{2}\int_M \big(\mathrm{d}\langle R^\nabla, R^\nabla\rangle, \,\mathrm{d}g_s\big) \,\mathrm{d}x. \tag{3.12}$$

We need to transform the last term. To this end we use

$$-\frac{1}{2}\int_M \big(\mathrm{d}\langle R^\nabla, R^\nabla\rangle, \,\mathrm{d}g_s\big) \,\mathrm{d}x = -\int_M \big(\langle \nabla R^\nabla, R^\nabla\rangle, \,\mathrm{d}g_s\big) \,\mathrm{d}x = -\int_M \langle \nabla_{\operatorname{grad}\log g_s} R^\nabla, R^\nabla\rangle g_s \,\mathrm{d}x. \tag{3.13}$$

By means of a Bianchi identity, we find the following expression for ∇R^∇: for $u, v, w \in \Gamma(TM)$,

$$\begin{aligned}\nabla_u R^\nabla(v,w) &= \nabla_v R^\nabla(u,w) - \nabla_w R^\nabla(u,v) = \mathrm{d}^\nabla\big(\iota_u R^\nabla\big)(v\wedge w) - R^\nabla(\nabla_v u, w) + R^\nabla(\nabla_w u, v)\\ &= \big(\mathrm{d}^\nabla\big(\iota_u R^\nabla\big) - 2R^\nabla\circ(\nabla u\odot \mathrm{id})\big)(v\wedge w).\end{aligned}$$

With this equality the last expression in Eq. (3.13) becomes

$$-\int \big\langle \mathrm{d}^\nabla\big(\iota_{\operatorname{grad}\log g_s} R^\nabla\big), R^\nabla\big\rangle g_s\,\mathrm{d}x + 2\int \big\langle R^\nabla\circ\big(\nabla(\operatorname{grad}\log g_s)\odot\mathrm{id}\big), R^\nabla\big\rangle g_s\,\mathrm{d}x, \tag{3.14}$$

or with Eq. (1.5),

$$-\int \big\langle \iota_{\operatorname{grad}\log g_s} R^\nabla, \big(\mathrm{d}^\nabla\big)^*\big(R^\nabla g_s\big)\big\rangle\,\mathrm{d}x + 2\int \big\langle R^\nabla\circ\big(\nabla(\operatorname{grad}\log g_s)\odot\mathrm{id}\big), R^\nabla\big\rangle g_s\,\mathrm{d}x. \tag{3.15}$$

Exploiting once more $(\mathrm{d}^\nabla)^*(R^\nabla g) = ((\mathrm{d}^\nabla)^* R^\nabla)g - \iota_{\operatorname{grad} g_s} R^\nabla$, we arrive at

$$\begin{aligned}&-\int \big\langle \iota_{\operatorname{grad}\log g_s} R^\nabla, \big(\mathrm{d}^\nabla\big)^* R^\nabla\big\rangle g_s\,\mathrm{d}x + \int \big\langle \iota_{\operatorname{grad}\log g_s} R^\nabla, \iota_{\operatorname{grad}\log g_s} R^\nabla\big\rangle g_s\,\mathrm{d}x\\ &\quad + 2\int \big\langle R^\nabla\circ\big(\nabla(\operatorname{grad}\log g_s)\odot\mathrm{id}\big), R^\nabla\big\rangle g_s\,\mathrm{d}x.\end{aligned} \tag{3.16}$$

Finally, combining Eqs. (3.12) and (3.16) gives

$$\varphi'(s) = \int \big\|\big(\mathrm{d}^\nabla\big)^* R^\nabla - \iota_{\operatorname{grad}\log g_s} R^\nabla\big\|^2 g_s\,\mathrm{d}x + 2\int \big\langle R^\nabla, R^\nabla\circ\big(\nabla(\operatorname{grad}\log g_s)\odot\mathrm{id}\big)\big\rangle g_s\,\mathrm{d}x.$$

From here the proposition follows upon noting that $\phi'(s) = s^2\varphi'(s) + 2s\varphi(s)$. □

In particular, if M is the Euclidean space $\mathbb{R}^m$, then $\nabla\operatorname{grad}\log g_s = -(1/s)\,\mathrm{id}$. As a consequence we have the following corollary:

Corollary 3.5. *Let $M = \mathbb{R}^m$ be equipped with the standard metric, and assume that Assumption* 3.3 *is satisfied. For every $0 < s < T$ there holds*

$$\phi'(s) = s^2\int \big\|\big(\mathrm{d}^{\nabla(T-s)}\big)^* R^{\nabla(T-s)} - \iota_{\operatorname{grad}\log g_s} R^{\nabla(T-s)}\big\|^2 g_s\,\mathrm{d}x. \tag{3.17}$$

Consequently, ϕ is non-decreasing on $]0,T[$, as well as Φ_β for every $\beta\in\,]0,1[$.

The monotonicity of ϕ (and Φ_β) holds in other situations as well. We say that M has parallel Ricci tensor Ric if $\nabla\,\mathrm{Ric} = 0$. Similarly to [27] Theorem 9.1 we get the following result which differs from Hamilton's monotonicity formula [13, Theorem C].

Theorem 3.6. *Assume that M is a compact manifold with parallel Ricci tensor and non-negative sectional curvatures. Then ϕ is non-decreasing on $]0,T[$, as well as Φ_β for every $\beta\in\,]0,1[$.*

Proof. This is a consequence of Proposition 3.4 and the inequality

$$\nabla(\operatorname{grad}\log g_s) \geqslant -\frac{1}{s}\,\mathrm{id}$$

which has been obtained by Hamilton [12] under the given assumptions. □

For a general compact manifold M, a correction term is needed. Again our result differs from [13, Theorem C]. For $0\leqslant t < T$ let

$$\mathrm{YM}(t) = \int_M \big\|R^{\nabla(t)}\big\|^2\,\mathrm{d}x \tag{3.18}$$

be the Yang–Mills energy at time t. The map $t\mapsto \mathrm{YM}(t)$ is non-increasing (see, e.g., [9]).

Theorem 3.7. *Let M be a compact manifold of dimension $m\geqslant 1$. There exist a constant $C > 0$ and a positive increasing function f_1 defined on $]0,1]$ satisfying $\lim_{s\to 0} f_1(s) = 0$, both depending only on M, such that for any $\beta\in\,]0,1[$ and all $0 < s_1 < s_2 < T\wedge 1$, the following inequalities hold*:

$$\phi(s_1) \leqslant e^{f_1(s_2)}\phi(s_2) + C(s_2 - s_1)\,\mathrm{YM}(0) \quad \textit{and} \tag{3.19}$$

$$\Phi_\beta(s_1) \leqslant e^{f_1(s_2)}\Phi_\beta(s_2) + C(1-\beta)(s_2 - s_1)\,\mathrm{YM}(0). \tag{3.20}$$

Other monotonicity formulas have been established in [9,17]. The last reference is closer to our result but in [17] the function ϕ is not the same as ours and M is a sphere.

Proof. We first establish Eq. (3.19) in a way similar to Theorem 1.1 in [13]. Let $0 < s < T \wedge 1$. From Proposition 3.4 we get

$$\phi'(s) \geqslant 2s \int_M \bigl\langle R, R + sR \circ \bigl(\nabla(\operatorname{grad}\log g_s) \odot \mathrm{id}\bigr)\bigr\rangle g_s \, \mathrm{d}x. \tag{3.21}$$

By [12, Theorem 4.3], there exist constants $B \geqslant 1$ and $C_0 > 0$ depending only on M such that

$$\nabla(\operatorname{grad}\log g_s) + \frac{1}{s}\,\mathrm{id} + C_0\left(1 + \log\left(\frac{B}{s^{m/2} g_s}\right)\right)\mathrm{id} \geqslant 0. \tag{3.22}$$

A straightforward calculation ([13, Lemma 1.2]) shows that for $x, y > 0$,

$$x\bigl(1 + \log(y/x)\bigr) \leqslant 1 + x\log y,$$

hence, Eq. (3.22) yields

$$g_s \nabla(\operatorname{grad}\log g_s) + g_s \frac{1}{s}\,\mathrm{id} \geqslant -C_0\left(1 + g_s \log\left(\frac{B}{s^{m/2}}\right)\right)\mathrm{id}. \tag{3.23}$$

From this and Eq. (3.21) we get

$$\phi'(s) \geqslant -2C_0 s^2\,\mathrm{YM}(T-s) - 2C_0 \log\left(\frac{B}{s^{m/2}}\right)\phi(s). \tag{3.24}$$

The function

$$f(s) := s\left(\frac{m}{2} + \log\left(\frac{B}{s^{m/2}}\right)\right)$$

is positive on the interval $]0, T \wedge 1[$, bounded by a constant C_1 and has derivative $\log(B s^{-m/2})$. Hence, letting $\alpha(s) = e^{2C_0 f(s)}\phi(s)$, Eq. (3.24) yields

$$\alpha'(s) \geqslant -2C_0 s^2 e^{2C_0 f(s)}\,\mathrm{YM}(T-s) \geqslant -2C_0 e^{2C_0 C_1}\,\mathrm{YM}(0). \tag{3.25}$$

Integrating from s_1 to s_2 where $0 < s_1 < s_2 < T \wedge 1$ yields

$$\alpha(s_1) \leqslant \alpha(s_2) + 2C_0 e^{2C_0 C_1}\,\mathrm{YM}(0)(s_2 - s_1) \tag{3.26}$$

which in turn gives

$$\phi(s_1) \leqslant e^{2C_0(f(s_2) - f(s_1))}\phi(s_2) + 2C_0 e^{-2C_0 f(s_1)} e^{2C_0 C_1}\,\mathrm{YM}(0)(s_2 - s_1). \tag{3.27}$$

Taking $C = (2C_0 \vee 1)e^{2C_0 C_1}$ and $f_1(s) = 2C_0 f(s)$, we get Eq. (3.19).

As for Eq. (3.20), we let $\lambda = s_2/s_1 > 0$. We then have

$$\begin{aligned}
\Phi_\beta(s_1) &= \int_{\beta s_1}^{s_1} \frac{\phi(r_1)}{r_1}\,\mathrm{d}r_1 = \int_{\beta s_2}^{s_2} \frac{\phi(r_2/\lambda)}{r_2}\,\mathrm{d}r_2 \leqslant \int_{\beta s_2}^{s_2} \frac{e^{f_1(r_2)}\phi(r_2) + Cr_2(1 - 1/\lambda)\,\mathrm{YM}(0)}{r_2}\,\mathrm{d}r_2 \quad \text{by Eq. (3.19)} \\
&\leqslant e^{f_1(s_2)} \int_{\beta s_2}^{s_2} \frac{\phi(r_2)}{r_2}\,\mathrm{d}r_2 + C(1 - 1/\lambda)(s_2 - \beta s_2)\,\mathrm{YM}(0) = e^{f_1(s_2)} \int_{\beta s_2}^{s_2} \frac{\phi(r_2)}{r_2}\,\mathrm{d}r_2 + C(1-\beta)(s_2 - s_1)\,\mathrm{YM}(0) \\
&= e^{f_1(s_2)}\Phi_\beta(s_2) + C(1-\beta)(s_2 - s_1)\,\mathrm{YM}(0)
\end{aligned}$$

which is the desired formula. □

By means of formula (3.19) we obtain existence of limits of ϕ and Φ_β at 0, as stated in the following proposition.

Proposition 3.8. *Let ∇ be a solution to the Yang–Mills heat equation defined on $[0,T[$. Then for any $x\in M$ and $0<\beta<1$, the limits $\lim_{r\to 0}\phi^{(T,x)}(r)$, $\lim_{r\to 0}\Phi_\beta^{(T,x)}(r)$ exist, and we have*

$$\lim_{r\to 0}\Phi_\beta^{(T,x)}(r)=\log(1/\beta)\lim_{r\to 0}\phi^{(T,x)}(r).$$

Proof. We consider first $\lim_{r\to 0}\phi^{(T,x)}(r)$. Let r_n be a decreasing sequence of positive numbers converging to 0 such that

$$\lim_{n\to\infty}\phi^{(T,x)}(r_n)=\liminf_{r\to 0}\phi^{(T,x)}(r).$$

Let $n\in\mathbb{N}$ and $0<r<r_n$. By formula (3.19),

$$\phi^{(T,x)}(r)\leqslant e^{f_1(r_n)}\phi^{(T,x)}(r_n)+C(r_n-r)\,\mathrm{YM}(0),$$

where C and f_1 depend only on M, and $f_1(s)$ converges to 0 as s tends to 0. This clearly implies

$$\limsup_{r\to 0}\phi^{(T,x)}(r)=\lim_{n\to\infty}\phi^{(T,x)}(r_n)=\liminf_{r\to 0}\phi^{(T,x)}(r),$$

so $\lim_{r\to 0}\phi^{(T,x)}(r)$ exists. From the bounds

$$\inf_{[\beta r,r]}\phi^{(T,x)}\leqslant\frac{\Phi_\beta^{(T,x)}(r)}{\log(1/\beta)}\leqslant\sup_{[\beta r,r]}\phi^{(T,x)},$$

we derive the results concerning $\lim_{r\to 0}\Phi_\beta^{(T,x)}(r)$. □

In the next proposition we express the limits of ϕ and Φ_β at 0 in terms of L^2 norms of the martingales $U_{r,s}^{-1}\nabla_0 U_{r,s}$.

Proposition 3.9. *Let $U_{r,s}=U_{r,s}^{(T,x)}(0,u)$.*

(1) *If $0<s_1\leqslant s_2\leqslant s_3\leqslant T$ then*

$$\big\|U_{s_1,s_3}^{-1}\nabla_0 U_{s_1,s_3}\big\|_2^2=\big\|U_{s_1,s_2}^{-1}\nabla_0 U_{s_1,s_2}\big\|_2^2+\big\|U_{s_2,s_3}^{-1}\nabla_0 U_{s_2,s_3}\big\|_2^2.$$

(2) *We have*

$$\lim_{n\to\infty}\frac{1}{n}\big\|U_{T/2^n,T}^{-1}\nabla_0 U_{T/2^n,T}\big\|_2^2=2\lim_{r\to 0}\Phi_{1/2}^{(T,x)}(r).$$

(3) *The equality in* (2) *generalizes to*

$$\lim_{r\to 0}\frac{1}{\log(T/r)}\big\|U_{r,T}^{-1}\nabla_0 U_{r,T}\big\|_2^2=2\lim_{r\to 0}\phi^{(T,x)}(r).$$

Proof. (1) Clearly

$$U_{s_1,s_3}^{-1}\nabla_0 U_{s_1,s_3}=U_{s_1,s_2}^{-1}U_{s_2,s_3}^{-1}\nabla_0(U_{s_2,s_3}U_{s_1,s_2})=U_{s_1,s_2}^{-1}\nabla_0 U_{s_1,s_2}+U_{s_1,s_2}^{-1}\big(U_{s_2,s_3}^{-1}\nabla_0 U_{s_2,s_3}\big)U_{s_1,s_2}.$$

Since U_{s_1,s_2} is an isometry, in order to prove the equality, we only need to check that

$$\mathbb{E}\big[\big\langle U_{s_1,s_2}^{-1}\nabla_0 U_{s_1,s_2},\,U_{s_1,s_2}^{-1}\big(U_{s_2,s_3}^{-1}\nabla_0 U_{s_2,s_3}\big)U_{s_1,s_2}\big\rangle\big]=0.$$

The left-hand side is equal to

$$\mathbb{E}\big[\mathbb{E}\big[\big\langle U_{s_1,s_2}^{-1}\nabla_0 U_{s_1,s_2},\,U_{s_1,s_2}^{-1}\big(U_{s_2,s_3}^{-1}\nabla_0 U_{s_2,s_3}\big)U_{s_1,s_2}\big\rangle\;\big|\;\mathcal{F}_{s_2}\big]\big]$$

and may be written as

$$\mathbb{E}\big[\big\langle U_{s_1,s_2}^{-1}\nabla_0 U_{s_1,s_2},\,U_{s_1,s_2}^{-1}\mathbb{E}\big[U_{s_2,s_3}^{-1}\nabla_0 U_{s_2,s_3}\;\big|\;\mathcal{F}_{s_2}\big]U_{s_1,s_2}\big\rangle\big].$$

But $\mathbb{E}[U_{s_2,s_3}^{-1}\nabla_0 U_{s_2,s_3}\mid\mathcal{F}_{s_2}]=0$ since $s\mapsto U_{s_2,s}^{-1}\nabla_0 U_{s_2,s}$ is an L^2-martingale which vanishes at $s=s_2$, and this gives the desired equality.

(2) Let $\Phi_{1/2} = \Phi_{1/2}^{(T,x)}$. By Lemma 2.2,

$$\left\| U^{-1}_{T/2^{k+1},T/2^k} \nabla_0 U_{T/2^{k+1},T/2^k} \right\|_2^2 = 2\,\Phi_{1/2}\big(T/2^k\big)$$

for any $k \geqslant 0$. Consequently, we get with (1)

$$\frac{1}{n} \left\| U^{-1}_{T/2^n,T} \nabla_0 U_{T/2^n,T} \right\|_2^2 = 2\,\frac{1}{n} \sum_{k=0}^{n-1} \Phi_{1/2}\big(T/2^k\big),$$

and by Proposition 3.8 the right-hand side converges to $2 \lim_{r\to 0} \Phi_{1/2}(r)$.

(3) Let $0 < r \leqslant T$ and let $n \geqslant 0$ satisfy

$$\frac{T}{2^{n+1}} < r \leqslant \frac{T}{2^n}$$

which is equivalent to

$$\frac{1}{n+1} < \frac{\log 2}{\log(T/r)} \leqslant \frac{1}{n}.$$

Then (1) implies

$$\left\| U^{-1}_{T/2^n,T} \nabla_0 U_{T/2^n,T} \right\|_2^2 \leqslant \left\| U^{-1}_{r,T} \nabla_0 U_{r,T} \right\|_2^2 \leqslant \left\| U^{-1}_{T/2^{n+1},T} \nabla_0 U_{T/2^{n+1},T} \right\|_2^2,$$

and the result follows by remarking that

$$\lim_{r\to 0} \Phi_{1/2}^{(T,x)}(r) = (\log 2) \lim_{r\to 0} \phi^{(T,x)}(r). \qquad \square$$

4. A priori bounds on solutions of the Yang–Mills heat equation

In this section, M is assumed to be compact of dimension m and $I = [0, T[$ where $T > 0$. Let ∇ be a solution of the Yang–Mills heat equation defined on I, and R^∇ be the corresponding curvature. For $0 < t < T$ and $x \in M$ let

$$e(t,x) = \left\| R^{\nabla(t)}(x) \right\|^2.$$

If $X(x)$ denotes a Brownian motion on M started at x, we let for $s \in [0,t]$:

$$\varphi^{(t,x)}(s) = \mathbb{E}\big[e\big(t-s, X_s(x)\big)\big] \quad \text{and} \quad \phi^{(t,x)}(s) = s^2 \varphi^{(t,x)}(s).$$

We keep the notation

$$\mathrm{YM}(t) = \int_M e(t,x)\,\mathrm{d}x,$$

and, for $\sigma \in [0, \sqrt{t}]$, $P(\sigma,t,x) = [t-\sigma^2, t] \times \overline{B}(x,\sigma)$, where $\overline{B}(x,\sigma)$ is the closed geodesic ball with center x and radius σ. Finally, for $x \in M$ and $\rho > 0$, let $\tau(x,\rho) = \inf\{s \geqslant 0{:}\ X_s(x) \notin \overline{B}(x,\rho)\}$.

Lemma 4.1. *There exists $\alpha_0 > 0$ and $\eta > 0$, depending only on M, such that for all $x \in M$, $0 < \rho < 1$ and $0 < \alpha < \alpha_0$,*

$$\mathbb{P}\big\{\tau(x,\rho) < \alpha\rho^2\big\} \leqslant \exp\left(-\frac{\eta}{\alpha}\right). \tag{4.1}$$

As a consequence,

$$\lim_{\alpha\to 0} \sup_{x\in M} \sup_{\rho\in]0,1]} \mathbb{P}\big\{\tau(x,\rho) < \alpha\rho^2\big\} = 0. \tag{4.2}$$

Proof. Denote by g the metric on M and by d the distance associated with g. Let $N_s(x) := \mathrm{d}^2(x, X_s(x))$. Then

$$\tau(x,\rho) = \inf\big\{s \geqslant 0,\ N_s(x) > \rho^2\big\},$$

which gives

$$\left\{\tau(x,\rho)<\alpha\rho^2\right\}=\left\{\sup_{s\in[0,\alpha\rho^2]}N_s(x)>\rho^2\right\}\subset\left\{\sup_{s\in[0,\alpha\rho^2]}N_s(x)\geqslant\varepsilon^2\rho^2\right\},$$

where $0<\varepsilon<1$ is less than the injectivity radius of (M,g). But $N_s(x)$ stopped at the first time when it hits $\rho^2\varepsilon^2$ satisfies

$$\mathrm{d}N_s(x)=\sigma_s(x)\,\mathrm{d}B_s(x)+b_s(x)\,\mathrm{d}s,$$

where $B_s(x)$ is a real-valued Brownian motion. Moreover, there exists a constant $C>0$ depending only on M such that $\sigma_s^2(x)\leqslant C\rho^2\varepsilon^2$ and $b_s(x)\leqslant C$. Since $N_0(x)=0$, this implies that for $C\alpha\rho^2\leqslant(\varepsilon^2\rho^2)/2$,

$$\mathbb{P}\left(\left\{\sup_{s\in[0,\alpha\rho^2]}N_s(x)\geqslant\varepsilon^2\rho^2\right\}\right)\leqslant\mathbb{P}\left(\left\{\sup_{s\in[0,\alpha\rho^2]}\int_0^s\sigma_r(x)\,\mathrm{d}B_r(x)\geqslant\varepsilon^2\rho^2/2\right\}\right)\leqslant\mathrm{e}^{-\varepsilon^2/(8C\alpha)},$$

where the last bound comes from Bernstein's inequality (see, e.g., Exercise 3.16, Chapter IV in [21]). This gives the result with $\alpha_0=\varepsilon^2/(2C)$ and $\eta=\varepsilon^2/(8C)$. □

We continue with a priori C^1-bounds for solutions ∇ of small energy.

Theorem 4.2. *There exist constants $\varepsilon_0=\varepsilon_0(E)>0$ and $0<\alpha=\alpha(E)<1$, a positive non-increasing function $y\mapsto\varepsilon_1(y)=\varepsilon_1(E,y)$ defined on $]0,\infty[$ with values in $]0,1]$, and a positive non-decreasing function $r\mapsto f(r)=f(E,r)$ on $]0,1[$ such that for any solution ∇ of the Yang–Mills heat equation on $[0,T[$ the following is true: if $\phi^{(t,x)}(r)\leqslant a\varepsilon_0$ for some $(t,x)\in I\times M$, $0<a\leqslant 1$ and some $r\in]0,\varepsilon_1(a^{-1}\,\mathrm{YM}(0))\wedge t]$, then*

$$\sup_{P(f(r),t,x)}e\leqslant\frac{2^4a}{\alpha^2f(r)^4}.$$

Note a similar result but for a fixed Yang–Mills connection can be found in [18]. See also [9] for a related formula.

Proof. We follow the proof of Theorem 10.1 in [27]. Let $a\in]0,1]$, $(t_0,x_0)\in I\times M$, $r_0\in]0,t_0\wedge 1]$, $r_1\in]0,r_0/2]$ (in particular, $r_1\leqslant\sqrt{r_0/2}$). We want to prove that for some $\varepsilon_0>0$, if $\phi^{(t_0,x_0)}(r_0)\leqslant a\varepsilon_0$ then

$$\sup_{P(r_1/2,t_0,x_0)}e\leqslant\frac{2^4a}{\alpha^2(r_1/2)^4},$$

where the relation between t_0, r_0 and r_1 has to be determined. Let $\sigma_0\in[0,r_1[$ such that

$$(r_1-\sigma_0)^4\sup_{P(\sigma_0,t_0,x_0)}e=\max_{\sigma\in[0,r_1]}\Big((r_1-\sigma)^4\sup_{P(\sigma,t_0,x_0)}e\Big).$$

There exists $(t^*,x^*)\in P(\sigma_0,t_0,x_0)$ such that

$$e_0:=\sup_{P(\sigma_0,t_0,x_0)}e=e(t^*,x^*).$$

For $e_0=0$, $r_1=2f(r_0)$, and any f such that $2f(r)<\sqrt{r/2}$ we are done, so for the rest of the proof we assume that $e_0>0$. Let $\rho_0=(1/2)(r_1-\sigma_0)$. Then we have $0<\sigma_0+\rho_0<r_1$, and

$$\sup_{P(\rho_0,t^*,x^*)}e\leqslant\sup_{P(\rho_0+\sigma_0,t_0,x_0)}e\leqslant(r_1-\sigma_0-\rho_0)^{-4}(r_1-\sigma_0)^4e_0=16e_0. \tag{4.3}$$

On the other hand, there exists $C_1=C_1(E)>0$ such that

$$\left(\frac{\partial}{\partial t}-\frac{1}{2}\Delta\right)e\leqslant C_1\big(1+e^{1/2}\big)e \tag{4.4}$$

(see [9, Lemma 2.2]). Define for $s\in[0,\rho_0^2]$

$$Y_s=e\big(t^*-s,X_s(x^*)\big).$$

Write $\tau=\tau(x^*,\rho_0)$. Then, denoting by $\overset{\mathrm{m}}{=}$ equality up to differentials of local martingales, we have on $[0,\tau\wedge\rho_0^2]$,

$$\mathrm{d}Y_s\overset{\mathrm{m}}{=}\left(-\frac{\partial}{\partial t}+\frac{1}{2}\Delta\right)e\big(t^*-s,X_s(x^*)\big)\,\mathrm{d}s\geqslant-C_1\big(1+Y_s^{1/2}\big)Y_s\,\mathrm{d}s\geqslant-C_1(1+4\sqrt{e_0})Y_s\,\mathrm{d}s,$$

where the inequalities come from Eqs. (4.4) and (4.3); this implies that

$$Z_s := \mathrm{e}^{C_1(1+4\sqrt{e_0})s} Y_s \tag{4.5}$$

is a bounded submartingale on $[0, \tau \wedge \rho_0^2]$. As a consequence, for every $s \in]0, \rho_0^2]$,

$$Z_0 \leqslant \mathbb{E}[Z_{s\wedge\tau}]. \tag{4.6}$$

We want to prove that if $\phi^{(t_0,x_0)}(r_0) \leqslant a\varepsilon_0$ where ε_0 has to be determined, then $\alpha\rho_0^2 \leqslant \sqrt{a/e_0}$ for some $\alpha = \alpha(E) \in]0, 1[$. If

$$\alpha\rho_0^2 > \sqrt{a/e_0}, \tag{4.7}$$

then

$$\begin{aligned} e_0 = Z_0 &\leqslant \mathbb{E}\big[Z_{(\sqrt{a/e_0})\wedge\tau}\big] = \mathbb{E}\big[Z_{\sqrt{a/e_0}} \mathbb{1}_{\{\sqrt{a/e_0}\leqslant\tau\}}\big] + \mathbb{E}\big[Z_\tau \mathbb{1}_{\{\sqrt{a/e_0}>\tau\}}\big] \\ &\leqslant \mathrm{e}^{C_1(1+4\sqrt{e_0})\sqrt{a/e_0}} \mathbb{E}\big[Y_{\sqrt{a/e_0}}\big] + \mathrm{e}^{C_1(1+4/\sqrt{e_0})\sqrt{a/e_0}} 16e_0 \mathbb{P}\big\{\sqrt{a/e_0} > \tau\big\} \\ &\leqslant \mathrm{e}^{5C_1} \mathbb{E}\big[Y_{\sqrt{a/e_0}}\big] + \mathrm{e}^{5C_1} 16e_0 \mathbb{P}\big\{\alpha\rho_0^2 > \tau\big\} = \mathrm{e}^{5C_1} \frac{e_0}{a} \phi^{(t^*,x^*)}\big(\sqrt{a/e_0}\big) + \mathrm{e}^{5C_1} 16e_0 \mathbb{P}\big\{\alpha\rho_0^2 > \tau\big\}. \end{aligned}$$

According to Lemma 4.1, one can choose $\alpha = \alpha(E) > 0$ such that

$$\mathrm{e}^{5C_1} 16 \mathbb{P}\big\{\alpha\rho_0^2 > \tau(x^*, \rho_0)\big\} < \frac{1}{2},$$

and we get

$$\frac{1}{2} e_0 \leqslant \mathrm{e}^{5C_1} \frac{e_0}{a} \phi^{(t^*,x^*)}\big(1/\sqrt{e_0}\big). \tag{4.8}$$

Now by the monotonicity formula (3.19), letting $\beta = t_0 - t^* \in [0, r_1^2[$ and $C' = C \vee \mathrm{e}^{f_1(1)}$ where C is the constant appearing in Theorem 3.7, Eq. (4.8) implies

$$e_0 \leqslant 2\mathrm{e}^{5C_1} \frac{e_0}{a} C' \big(\phi^{(t^*,x^*)}(r_0 - \beta) + (r_0 - \beta)\,\mathrm{YM}(0)\big). \tag{4.9}$$

Dividing by e_0/a and letting $C_2 = 2\mathrm{e}^{5C_1} C'$, we get

$$a \leqslant C_2(r_0 - \beta)^2 \int_M e(t_0 - r_0, x)\, p(r_0 - \beta, x^*, x)\, \mathrm{d}x + C_2(r_0 - \beta)\,\mathrm{YM}(0), \tag{4.10}$$

where $p(r, x, y)$ is the density at $y \in M$ at time r of a Brownian motion started at x. The function $(r, x, y) \mapsto p(r, x, y)$ is smooth on the compact set $[r_0/2, 1] \times M \times M$. Consequently, since $\beta < r_1^2 < r_1 \leqslant r_0/2 < 1$ and $d(x^*, x_0) \leqslant r_1$, there exists $C_3(r_0) > 0$ such that

$$p(r_0 - \beta, x^*, x) \leqslant p(r_0, x_0, x) + C_3(r_0)\, r_1.$$

Substituting this in Eq. (4.10) yields

$$a \leqslant C_2(r_0 - \beta)^2 \int_M e(t_0 - r_0, x)\, p(r_0, x_0, x)\, \mathrm{d}x + C_3(r_0) r_1 C_2 (r_0 - \beta)^2\, \mathrm{YM}(t_0 - r_0) + C_2(r_0 - \beta)\,\mathrm{YM}(0)$$

which in turn implies (since $t \mapsto \mathrm{YM}(t)$ is non-increasing)

$$a \leqslant C_2\, \phi^{(t_0,x_0)}(r_0) + C_4(r_0)\, r_1\, \mathrm{YM}(0) + C_2\, r_0\, \mathrm{YM}(0), \tag{4.11}$$

where $C_4(r_0) = C_2 C_3(r_0) r_0^2$. Moreover, $r \mapsto C_4(r)$ may be chosen decreasing. Let

$$\varepsilon_1(y) = \frac{1}{(3C_2 y) \vee 1}, \qquad f(r) = \frac{C_2\, r}{2\, C_4(r)} \wedge \frac{r}{4}, \qquad \varepsilon_0 = \frac{1}{3C_2}.$$

If $r_0 < \varepsilon_1(\mathrm{YM}(0)/a)$, $r_1 \leqslant 2f(r_0)$ and $\phi^{(t_0,x_0)}(r_0) \leqslant a\varepsilon_0$, then

$$C_2\, \phi^{(t_0,x_0)}(r_0) + C_4(r_0)\, r_1\, \mathrm{YM}(0) + C_2\, r_0\, \mathrm{YM}(0) < a,$$

in contradiction to Eq. (4.11), and, hence, to Eq. (4.7). Thus we must have

$$e_0 \leqslant \frac{a}{\alpha^2 \rho_0^4} = \frac{2^4 a}{\alpha^2 (r_1 - \sigma_0)^4}, \tag{4.12}$$

in particular,

$$\max_{\sigma\in[0,r_1]}\Big((r_1-\sigma)^4\sup_{P(\sigma,t_0,x_0)} e\Big)=(r_1-\sigma_0)^4 e_0\leqslant\frac{2^4 a}{\alpha^2}. \tag{4.13}$$

Now letting $\sigma=r_1/2=f(r_0)$ which gives $(r_1-\sigma)^4=f(r_0)^4$, we get along with Eq. (4.13)

$$f(r_0)^4\sup_{P(f(r_0),t_0,x_0)} e\leqslant\frac{2^4 a}{\alpha^2}\quad\text{or}\quad\sup_{P(f(r_0),t_0,x_0)} e\leqslant\frac{2^4 a}{\alpha^2 f(r_0)^4}$$

which proves the theorem. □

An essential tool for the proof of Theorem 4.2 is the monotonicity formula (3.19) which involves $\phi^{(t,x)}(r)$. From a stochastic point of view, the function

$$\Phi_\beta^{(t,x)}(s)=\int_{\beta s}^{s} r\varphi^{(t,x)}(r)\,\mathrm{d}r\quad(\text{where } 0<\beta<1)$$

is more appealing since it allows a direct probabilistic interpretation:

$$\Phi_\beta^{(t,x)}(s)=\frac{1}{2}\mathbb{E}\Big[S_{\beta s,s}^{(t,x)}\Big],$$

see Lemma 2.2. For this reason we give a variant of Theorem 4.2 in terms of $\Phi_\beta^{(t,x)}$.

Theorem 4.3. *Let $0<\beta<1$ and let ε_0, ε_1, f be as in Theorem 4.2. For any solution ∇ of the Yang–Mills heat equation defined on $[0,T[$ the following is true: if*

$$\Phi_\beta^{(t,x)}(r)\leqslant a\varepsilon_0\log(1/\beta)$$

for some $a\in\,]0,1]$, $(t,x)\in I\times M$ and some $r\in\,]0,\varepsilon_1(a^{-1}\,\mathrm{YM}(0))\wedge t]$, then

$$\sup_{P(f(\beta r),t,x)} e\leqslant\frac{2^4 a}{\alpha^2 f(\beta r)^4}.$$

Proof. The equality

$$\Phi_\beta(r)=\int_{\beta r}^{r}\frac{\phi(s)}{s}\,\mathrm{d}s$$

implies

$$\log(1/\beta)\inf_{[\beta r,r]}\phi\leqslant\Phi_\beta(r),$$

and we are left to apply Theorem 4.2, along with the fact that f is non-decreasing. □

From Theorem 4.2 we get an immediate but useful corollary which gives a similar result, but in terms of $\phi^{(T,x)}$. For $t>0$, $\sigma\in[0,\sqrt{t}]$ let

$$P'(\sigma,t,x)=\big[t-\sigma^2,t\big[\times\overline{B}(x,\sigma).$$

Corollary 4.4. *The notations are the same as in Theorem 4.2. If $\phi^{(T,x)}(r)<a\varepsilon_0$ for some $x\in M$, $0<a\leqslant 1$ and some r such that $0<r<\varepsilon_1(a^{-1}\,\mathrm{YM}(0))\wedge T$, then*

$$\sup_{P'(f(r),T,x)} e\leqslant\frac{2^4 a}{\alpha^2 f(r)^4}. \tag{4.14}$$

Proof. Suppose that the assumptions of Corollary 4.4 are realized. By continuity of $t\mapsto\phi^{(t,x)}(r)$, there exists $\varepsilon>0$ such that $r\leqslant\varepsilon_1(a^{-1}\,\mathrm{YM}(0))\wedge T-\varepsilon$ and $\phi^{(t,x)}(r)\leqslant a\varepsilon_0$ for any $t\in[T-\varepsilon,T[$. Consequently, by Theorem 4.2,

$$\sup_{P(f(r),t,x)} e\leqslant\frac{2^4 a}{\alpha^2 f(r)^4}.$$

Since this is true for every $t \in [T - \varepsilon, T[$, the claim follows. □

At this stage we are able to give criteria for existence of singularities.

Proposition 4.5. *Let ε_0 be defined as in Theorem 4.2. The following five statements are equivalent*:

(i) ∇ *has a singularity at* (T, x);
(ii) $\lim_{r\to 0} \phi^{(T,x)}(r) \geqslant \varepsilon_0$;
(iii) $\lim_{r\to 0} \phi^{(T,x)}(r) > 0$;
(iv) $\lim_{r\to 0} \Phi_\beta^{(T,x)}(r) \geqslant \log(1/\beta)\varepsilon_0$;
(v) $\lim_{r\to 0} \Phi_\beta^{(T,x)}(r) > 0$.

Proof. Assume that (i) holds true. If (ii) is not satisfied, then we may choose r such that $r < \varepsilon_1(\mathrm{YM}(0)) \wedge T$ and $\phi^{(T,x)}(r) < \varepsilon_0$. By Corollary 4.4 we get

$$\sup_{P'(f(r),T,x)} e \leqslant \frac{2^4}{\alpha^2 f(r)^4},$$

in contradiction to explosion at (T, x). Consequently (i) implies (ii).

Clearly (ii) implies (iii). We prove that (iii) implies (i): assume that no explosion occurs at (T, x). Then there exists $\varepsilon > 0$ and $C > 0$ such that the energy is bounded by C on $[0, T[\times B(x, \varepsilon)$. On the other hand, for $y \notin B(x, \varepsilon)$ and $0 < r \leqslant 1$, we can bound $p(r, x, y)$ by some constant $C' > 0$. Thus we get

$$\begin{aligned}\phi^{(T,x)}(r) &= r^2 \int_M p(r,x,y)e(T-r,y)\,\mathrm{d}y = r^2 \int_{B(x,\varepsilon)} p(r,x,y)e(T-r,y)\,\mathrm{d}y + r^2 \int_{B(x,\varepsilon)^c} p(r,x,y)e(T-r,y)\,\mathrm{d}y \\ &\leqslant r^2 C \int_M p(r,x,y)\,\mathrm{d}y + r^2 C' \,\mathrm{YM}(T-r) \leqslant r^2 C + r^2 C' \,\mathrm{YM}(0)\end{aligned}$$

which clearly converges to 0 as r tends to 0. Hence, (iii) implies (i).

The equivalence with (iv) and (v) is a consequence of equality

$$\lim_{r\to 0} \Phi_\beta^{(T,x)}(r) = \log(1/\beta) \lim_{r\to 0} \phi^{(T,x)}(r)$$

in Proposition 3.8. □

We have the following immediate corollary.

Corollary 4.6. *Let $U_{r,s} = U_{r,s}^{(T,x)}$. The following five statements are equivalent*:

(i) ∇ *has a singularity at* (T, x);
(ii) $\lim_{n\to\infty} \frac{1}{n} \| U^{-1}_{T/2^n,T} \nabla_0 U_{T/2^n,T} \|_2^2 \geqslant 2(\log 2)\varepsilon_0$;
(iii) $\lim_{n\to\infty} \frac{1}{n} \| U^{-1}_{T/2^n,T} \nabla_0 U_{T/2^n,T} \|_2^2 > 0$;
(iv) $\lim_{r\to 0} \frac{1}{\log(T/r)} \| U^{-1}_{r,T} \nabla_0 U_{r,T} \|_2^2 \geqslant 2\varepsilon_0$;
(v) $\lim_{r\to 0} \frac{1}{\log(T/r)} \| U^{-1}_{r,T} \nabla_0 U_{r,T} \|_2^2 > 0$.

With Corollary 4.4 at hand we are able to obtain global existence results for solutions ∇ of the Yang–Mills heat equation. To this end, we shall exploit the fact that

$$p(t, x, y) = t^{-m/2} q(t, x, y), \tag{4.15}$$

where q is bounded on $]0, 1] \times M \times M$ (see, e.g., [14]).

Let $(t,x)\in]0,T]\times M$ and $0<r\leqslant t$. Write $q_r(x,y)=q(r,x,y)$. Then

$$\phi^{(t,x)}(r)=r^2\int_M p(r,x,y)\,e(t-r,y)\,\mathrm{d}y\leqslant r^{2-m/2}\Big(\sup_{M\times M} q_r\Big)\,\mathrm{YM}(t-r)\leqslant r^{2-m/2}\Big(\sup_{M\times M} q_r\Big)\,\mathrm{YM}(0)$$

since $s\mapsto \mathrm{YM}(s)$ is non-increasing.

Theorem 4.7. *Let* ∇ *be a solution on* $[0,T_0[$ *of the Yang–Mills heat equation, and choose* ε_0, ε_1 *according to Theorem* 4.2. *If there exists* $t\in]0,T_0\wedge\varepsilon_1(\mathrm{YM}(0))[$ *such that*

$$t^{2-m/2}\Big(\sup_{M\times M} q_t\Big)\,\mathrm{YM}(0)<\varepsilon_0 \tag{4.16}$$

then the solution ∇ *can be extended to* $[0,\infty[$.

Proof. Let T be the maximal existence time of the solution to the Yang–Mills heat equation started at $\nabla(0)$. Then $T\geqslant T_0$. Assume that $T<\infty$. Let $T'\in]0,T_0\wedge\varepsilon_1(\mathrm{YM}(0))[$ satisfy

$$(T')^{2-m/2}\Big(\sup_{M\times M} q_{T'}\Big)\,\mathrm{YM}(0)<\varepsilon_0 \tag{4.17}$$

and let $t_0=T-T'$. We are going to prove that for any $x\in M$,

$$\sup_{P'(f(T'),T,x)} e\leqslant \frac{2^4}{\alpha^2 f(T')^4}. \tag{4.18}$$

Hence, let $x\in M$. Since the energy is decreasing with time, we have

$$(T')^{2-m/2}\Big(\sup_{M\times M} q_{T'}\Big)\,\mathrm{YM}(t_0)<\varepsilon_0. \tag{4.19}$$

The family

$$\nabla'(s)=\nabla(t_0+s),\quad 0\leqslant s<T',$$

of covariant derivatives solves the Yang–Mills heat equation on $[0,T'[$ with initial connection $\nabla(t_0)$ and initial energy $\mathrm{YM}'(0)=\mathrm{YM}(t_0)$. Denote by R' its curvature and by e' the norm of R'. Let

$$(\varphi')^{(T',x)}(r)=\mathbb{E}\big[e'\big(T'-r,X_r(x)\big)\big]\quad\text{and}\quad (\phi')^{(T',x)}(r)=r^2(\varphi')^{(T',x)}(r).$$

Then the calculation before Theorem 4.7 along with Eq. (4.17) shows that

$$(\phi')^{(T',x)}(T')\leqslant (T')^{2-m/2}\Big(\sup_{M\times M} q_{T'}\Big)\,\mathrm{YM}(t_0)<\varepsilon_0.$$

By Corollary 4.4 this implies that

$$\sup_{P'(f(T'),T',x)} e'\leqslant \frac{2^4}{\alpha^2 f(T')^4},$$

or equivalently,

$$\sup_{P'(f(T'),T,x)} e\leqslant \frac{2^4}{\alpha^2 f(T')^4}.$$

This holds true for all $x\in M$, so

$$\sup_{[T-f(T')^2,T[\times M} e\leqslant \frac{2^4}{\alpha^2 f(T')^4},$$

in contradiction to the fact that the solution ∇ explodes at time T. As a conclusion, we have $T=\infty$. □

From Theorem 4.7 we derive two immediate corollaries:

Corollary 4.8. *If* $m\leqslant 3$ *then a solution to the Yang–Mills heat equation does not blow up in finite time.*

Proof. There exists $T_0 > 0$ such that the solution is defined at least on $[0, T_0[$. Since $2 - m/2 > 0$, we have for sufficiently small $t \in]0, \varepsilon_1(\mathrm{YM}(0)) \wedge T_0[$,

$$t^{2-m/2} C_5 \,\mathrm{YM}(0) < \varepsilon_0,$$

where C_5 is an upper bound for q on $]0,1] \times M \times M$. Consequently Eq. (4.16) is satisfied and we are left to apply Theorem 4.7. □

Corollary 4.9. *If $m \geqslant 4$, and if the solution ∇ blows up in finite time $T < \varepsilon_1(\mathrm{YM}(0))$, then*

$$T^{2-m/2}\Big(\sup_{M\times M} q_T\Big)\,\mathrm{YM}(0) \geqslant \varepsilon_0,$$

where ε_0 and ε_1 are defined in Theorem 4.2.

By means of Corollary 4.4 we can improve the conclusion of Theorem 4.7 with stronger assumptions on YM(0), and obtain control on $\|e\|_\infty$ by $\sqrt{\mathrm{YM}(0)}$. The idea is to take $a = \sqrt{\mathrm{YM}(0)}$ in Corollary 4.4:

Theorem 4.10. *Assume $m \geqslant 4$. There exist a positive non-decreasing function $t \mapsto \varepsilon_2(t)$ and a positive function $t \mapsto C_6(t)$ defined on $]0, \infty[$, depending only on E, such that if a solution ∇ of the Yang–Mills heat equation defined on $[0, T[$ satisfies $\mathrm{YM}(0) < \varepsilon_2(T)$, then for any $x \in M$ and t with $T - f^2(T \wedge \varepsilon_1(\sqrt{\mathrm{YM}(0)})) \leqslant t < T$, we have the estimate $e(t,x) \leqslant C_6(T)\sqrt{\mathrm{YM}(0)}$, where f and ε_1 are defined in Theorem* 4.2.

Proof. Let

$$\varepsilon_2(t) = 1 \wedge \inf\left\{ y \in \mathbb{R},\ y > \frac{\varepsilon_0^2}{4C_5^2}\big(t \wedge \varepsilon_1(\sqrt{y})\big)^{m-4}\right\} \quad \text{and} \quad C_6(t) = \frac{2^4}{\alpha^2 f(\varepsilon_1(\sqrt{\varepsilon_2(t)}) \wedge t)^4},$$

where ε_0 and α are defined in Theorem 4.2, and C_5 is chosen such that $p(t,x,y) \leqslant C_5 t^{-m/2}$ for any $(t,x,y) \in]0,1] \times M \times M$. We already noted that if $0 < t \leqslant T$, $0 < r \leqslant 1 \wedge t$, $x \in M$, then

$$\phi^{(t,x)}(r) \leqslant C_5 r^{2-m/2}\,\mathrm{YM}(0).$$

Let $x \in M$. From the inequality $\mathrm{YM}(0) < \varepsilon_2(T)$ we get

$$C_5\big(T \wedge \varepsilon_1\big(\sqrt{\mathrm{YM}(0)}\big)\big)^{2-m/2}\sqrt{\mathrm{YM}(0)} \leqslant \varepsilon_0/2$$

which in turn implies $\phi^{(T,x)}(r) < a\varepsilon_0$ with $a = \sqrt{\mathrm{YM}(0)}$ and $r < T \wedge \varepsilon_1(\sqrt{\mathrm{YM}(0)}) = T \wedge \varepsilon_1(a^{-1}\,\mathrm{YM}(0))$. Applying Corollary 4.4 yields

$$\sup_{P'(f(r),T,x)} e \leqslant \frac{2^4}{\alpha^2 f(r)^4}\sqrt{\mathrm{YM}(0)}.$$

By continuity this inequality remains true when replacing r by

$$r_0 := T \wedge \varepsilon_1\big(\sqrt{\mathrm{YM}(0)}\big).$$

From $\mathrm{YM}(0) < \varepsilon_2(T)$ and the fact that ε_1 is non-increasing and f is non-decreasing, we conclude

$$\sup_{P'(f(r_0),T,x)} e \leqslant C_6(T)\sqrt{\mathrm{YM}(0)}.$$

Since this holds true for every $x \in M$, the proof is complete. □

Similarly to Theorem 4.7 we get the following corollary.

Corollary 4.11. *Assume $m \geqslant 4$ and let ∇ be a solution of the Yang–Mills heat equation defined on $[0, T[$. If $\mathrm{YM}(0) < \varepsilon_2(T)$, then the solution ∇ can be extended to $[0, \infty[$, and for every $t \geqslant T$, $x \in M$, $e(t,x) \leqslant C_6(T)\sqrt{\mathrm{YM}(0)}$, where ε_2 and C_6 are defined in Theorem* 4.10.

The proof relies on Theorem 4.10. It is similar to the proof of Theorem 4.7, and hence omitted. Corollary 4.11 in turn yields the following result on the sphere, which is due to Naito [17]:

Corollary 4.12. *Assume $m \geqslant 5$. Let S^m be the m-dimensional Euclidean sphere and E be a non-trivial vector bundle over S^m. There exists a map $t \mapsto \varepsilon_3(t) > 0$ defined for $t > 0$, such that for every solution ∇ of the Yang–Mills heat equation, if $\mathrm{YM}(0) < \varepsilon_3(t)$, then ∇ blows up before time t.*

Remark 4.13. Although both Corollary 4.11 and 4.12 assume smallness of the initial energy, there is a major difference in their assumptions, namely in Corollary 4.11 the solution ∇ is supposed to be already defined on $[0,T[$. If in Corollary 4.11, we further assume that $m \geqslant 5$ and $M = S^m$, then E is necessarily trivial.

Proof (of Corollary 4.12). Let

$$\varepsilon_3(t) = \frac{\binom{m}{2}^2}{4C_6(t)^2} \wedge \varepsilon_2(t).$$

Assume that $\mathrm{YM}(0) < \varepsilon_3(t)$. We want to prove that the solution ∇ blows up before time t. If not, then by Corollary 4.11, it can be extended to $[0,\infty[$ such that for every $t' \geqslant t$, $x \in M$,

$$e(t',x) \leqslant C_6(t)\sqrt{\mathrm{YM}(0)} < \frac{1}{2}\binom{m}{2}.$$

By [28] Theorem 1.5 there exists a subsequence $\nabla(t_i)$ such that $s_i^{-1} \circ \nabla(t_i) \circ s_i$ converges weakly in $W^{1,p}$ for any $p > n$ (hence, in C^0) to $\nabla(\infty)$, where the s_i are global gauge transformations in $W^{2,p}$. Moreover $\nabla(\infty)$ is weakly Yang–Mills and its energy $e(\infty)$ satisfies

$$\sup_{x\in M} e(\infty,x) \leqslant C_6(t)\sqrt{\mathrm{YM}(0)} < \frac{1}{2}\binom{m}{2}.$$

By [28] Corollary 1.4, the gauge transformations can be chosen in such a way that $\nabla(\infty)$ is strongly Yang–Mills. By [8, Theorem C], this implies $e(\infty) \equiv 0$, which is impossible since E is non-trivial. We conclude that our solution ∇ blows up before time t. □

5. Singularities of the Yang–Mills heat equation and convergence of rescaled martingales

In this section, the dimension of M is assumed to be at least four. Again ∇ is a solution of the Yang–Mills heat equation defined on $I = [0,T[$. We assume that ∇ blows up at time T. Let R_n be a decreasing sequence of positive numbers converging to 0. We consider the rescaled connections $\nabla^n(s) = \nabla(R_n^2 s)$ for $0 \leqslant s < T/R_n^2$. Then ∇^n solves the Yang–Mills heat equation when M is endowed with the metric $g^n = R_n^{-2} g$.

Lemma 5.1. *Let $\alpha > 0$ be as in the proof to Theorem* 4.2. *There exists $\varepsilon > 0$ depending only on M, a sequence (x_n) in M, a sequence (t_n) in $]0,T[$ converging to T, and a sequence (r_n) with $0 < r_n \leqslant R_n^2$, such that for n sufficiently large, $\phi^{(t_n,x_n)}(r_n) = \varepsilon$, and such that the curvatures R^{∇^n} of the rescaled connections ∇^n satisfy*

$$\sup_{[1,t_n/R_n^2]\times M} \left\| R^{\nabla^n} \right\|^2 \leqslant 2^8/\alpha^2,$$

where the norm is defined in terms of the rescaled metric g^n.

Observe that Lemma 5.1 is similar in spirit to the C^0 bound in [24]. A first difference is that we do not assume that M has dimension 4. A second difference is that our bound is obtained globally on M and not on a small ball, but we cannot prescribe a limit for our sequence (x_n), we confine ourselves to the statement that by extracting a subsequence (x_n) converges to a singularity at time T. A third difference is that our proof relies on a submartingale inequality instead on Moser's Harnack inequality.

Proof (of Lemma 5.1). Let $0 < \varepsilon < (2\mathrm{e}^{5C_1})^{-1} \wedge \varepsilon_0$ where ε_0 and C_1 are as in Theorem 4.2 and its proof, and

$$t_n' = \sup\left\{ t' \in]0,T[,\ \sup_{y\in M}\ \sup_{r\in]0,R_n^2]} \phi^{(t',y)}(r) \leqslant \varepsilon \right\}$$

(we let $t_n' = 0$ in case the set on the right is empty). By the same argument as in the proof of Lemma 4.5, t_n' converges to T. For n sufficiently large, let $t_n \in]0,t_n']$, $r_n \in]0,R_n^2]$, $x_n \in M$ such that $\phi^{(t_n,x_n)}(r_n) = \varepsilon$. Note that such t_n' t_n, r_n, x_n exist, since by Theorem 4.2

$$\sup_{t'\in[0,T[}\ \sup_{y\in M}\ \sup_{r\in]0,R_n^2]} \phi^{(t',y)}(r) \geqslant \varepsilon_0.$$

Necessarily t_n converges to T. We now choose arbitrary $s_n \in [R_n^2, t_n]$ and $z_n \in M$. Let $\sigma_0 \in [0, R_n]$ such that

$$(R_n - \sigma_0)^4 \sup_{P(\sigma_0, s_n, z_n)} e = \max_{\sigma \in [0, R_n]} \Big((R_n - \sigma)^4 \sup_{P(\sigma, s_n, z_n)} e \Big).$$

There exists $(s_n^*, z_n^*) \in P(\sigma_0, s_n, z_n)$ such that

$$e_0 := \sup_{P(\sigma_0, s_n, z_n)} e = e(s_n^*, z_n^*).$$

Let $\rho_0 = (1/2)(R_n - \sigma_0)$. We have $0 < \sigma_0 + \rho_0 < R_n$ and

$$\sup_{P(\rho_0, s_n^*, z_n^*)} e \leqslant \sup_{P(\rho_0 + \sigma_0, s_n, z_n)} e \leqslant (R_n - \sigma_0 - \rho_0)^{-4} (R_n - \sigma_0)^4 e_0 = 16 e_0. \tag{5.1}$$

We want to prove that $e_0 \leqslant (\alpha^2 \rho_0^4)^{-1}$. If this is not true, then $1/\sqrt{e_0} \leqslant \alpha \rho_0^2 \leqslant R_n^2/4$, and as in (4.8) we get

$$1 \leqslant 2\mathrm{e}^{5C_1} \phi^{(s_n^*, z_n^*)}(1/\sqrt{e_0}) \leqslant 2\mathrm{e}^{5C_1} \varepsilon,$$

where for the last inequality we use the definition of t_n' and the fact that $s_n^* \leqslant t_n'$. Since $\varepsilon < (2\mathrm{e}^{5C_1})^{-1}$ we arrive at a contradiction. Consequently $e_0 \leqslant (\alpha^2 \rho_0^4)^{-1}$. Taking $\sigma_0 = R_n/2$ we get

$$\sup_{P(R_n/2, s_n, z_n)} e \leqslant \frac{2^8}{\alpha^2 R_n^4}. \tag{5.2}$$

Inequality (5.2) is true for all $s_n \in [R_n^2, t_n]$ and $z_n \in M$, hence, we obtain

$$\sup_{[R_n^2, t_n] \times M} e \leqslant \frac{2^8}{\alpha^2 R_n^4}. \tag{5.3}$$

Denoting by e^n the energy of the rescaled connection ∇^n, we have $e^n(s, y) = R_n^4 e(R_n^2 s, y)$, and, hence, inequality (5.3) yields

$$\sup_{[1, t_n/R_n^2] \times M} e^n \leqslant \frac{2^8}{\alpha^2}$$

which is the desired result. □

Clearly the accumulation points of the sequence (x_n) belong to the singularity set of Eq. (1.2) at time T. By extracting a subsequence we may assume that (x_n) converges to some point $x \in M$. For $n \geqslant 0$ let X^n be a Brownian motion with respect to the metric g^n, started at x_n, which we construct for simplicity via $X_s^n = X_{sR_n^2}$ from a Brownian motion $X \equiv X(x_n)$ with respect to g and starting point x_n. The following processes are defined as in Section 2:

$$X_s^n(a, u) := X_{sR_n^2}(a, u), \qquad U_s^n(a, u) := U_{0, sR_n^2}^{(t_n, x_n)}(a, u), \qquad \big(U_s^n\big)^{-1} \nabla_0 U_s^n, \tag{5.4}$$

where $(\nabla_0 U_s^n) v(0) \equiv \nabla_0 (t_n - s R_n^2)\, (U_s^n v) - U_s^n(0) \nabla_0(t_n) v$. Note that we can take $r = 0$ in the definition of $U_{r,s}^{(t_n, x_n)}$ since $\nabla(t)$ is defined on $[0, T[$ and $t_n < T$. We stop the processes at time $t_n/R_n^2 - 1$ so that they are defined for all times and the last one has a bounded bracket. (The processes listed in (5.4) could be defined more intrinsically with respect to g^n and ∇^n; for the sake of clarity, we construct them with respect to the fixed metric g via the explicit time change $s \mapsto R_n^2 s$.)

By means of parallel transport along minimizing geodesics we identify the fibers E_{x_n} and E_x. In the same way, $(T_{x_n} M, g_n(x_n))$ is first identified isometrically with $(T_x M, g_n(x))$ by parallel transport along the minimizing geodesic from x_n to x with respect to the Levi-Civita connection to g_n; then $(T_x M, g_n(x))$ is identified isometrically with $(\mathbb{R}^m, \mathrm{eucl})$. Adopting these conventions, the $(U_s^n)^{-1} \nabla_0 U_s^n$ may be considered as processes taking values in the fixed Euclidean vector space $T_x^* M \otimes \mathrm{End}(E_x) =: F_x$.

Let $\mathcal{C}(\mathbb{R}_+, F_x)$ be the space of continuous paths in F_x. We endow $\mathcal{C}(\mathbb{R}_+, F_x)$ with the topology of uniform convergence on compact sets, and say that a sequence (V_n) of F_x-valued processes indexed by R_+ is $\mathcal{C}$-tight if the sequence of the laws of their paths is tight in $\mathcal{C}(\mathbb{R}_+, F_x)$. A random variable V with values in $\mathcal{C}(\mathbb{R}_+, F_x)$ is said to be a limit point of the sequence V_n if some subsequence converges in law to V.

Proposition 5.2. *The sequence $((U^n)^{-1} \nabla_0 U^n)_{n \geqslant 0}$ is $\mathcal{C}$-tight. Any limit point V, considered as a process, is a continuous martingale in its own filtration, satisfying*:

$$\forall\, s \geqslant 0, \quad \mathbb{E}\big[\|V_s\|^2\big] \leqslant \frac{2^8}{\alpha^2} s^2 \quad \textit{and} \quad \mathbb{E}\big[\|V_2\|^2\big] \geqslant (\log 2)\varepsilon.$$

In particular, the limit point V is non-trivial.

Remark 5.3. When $\dim M = 4$, this result is clearly related to Theorem 1.3 and Lemma 2.4 in [24], where convergence of rescaled connections to a non-trivial connection in a vector bundle over $\mathbb{R}^4$ is established under suitable gauge transformations. Observe that instead of gauge transformations we use here the moving frames U^n. The question arises whether our limiting process V is related to a finite energy Yang–Mills connection in a vector bundle over $\mathbb{R}^4$ with the same typical fiber as E (observe that this would not necessarily imply that V has finite L^2 norm).

Proof (of Proposition 5.2). The tightness is obtained with [20, Corollary 6 p. 31 and Remark 6 p. 59]. To verify that the conditions of Corollary 6 are fulfilled we only have to use the fact that by Lemma 5.1,

$$\mathrm{d}\big\langle (U^n)^{-1}\nabla_0 U^n\big\rangle_s \leqslant \frac{2^9}{\alpha^2} s\,\mathrm{d}s, \tag{5.5}$$

where $\langle (U^n)^{-1}\nabla_0 U^n\rangle$ denotes the quadratic variation of $(U^n)^{-1}\nabla_0 U^n$. By extracting a subsequence we may assume convergence in law of $((U^n)^{-1}\nabla_0 U^n)_{n\in\mathbb{N}}$. It implies the convergence

$$\mathbb{E}\big[\big\|\big(U_s^n\big)^{-1}\nabla_0 U_s^n\big\|^2 \wedge N\big] \to \mathbb{E}\big[\|V_s\|^2 \wedge N\big]$$

for any $N > 0$. But

$$\mathbb{E}\big[\big\|\big(U_s^n\big)^{-1}\nabla_0 U_s^n\big\|^2 \wedge N\big] \leqslant 2^8 s^2/\alpha^2$$

by Eq. (5.5), so that the first inequality of the proposition follows. For the second inequality we use uniform integrability of $\|(U_2^n)^{-1}\nabla_0 U_2^n\|^2$. We have

$$\mathbb{E}\big[\big\|\big(U_2^n\big)^{-1}\nabla_0 U_2^n\big\|^2\big] = \mathbb{E}\Big[\Big\|\Big(U_{0,2R_n^2}^{(t_n,x_n)}\Big)^{-1}\nabla_0 U_{0,2R_n^2}^{(t_n,x_n)}\Big\|^2\Big] \geqslant 2\Phi_{1/2}^{(t_n,x_n)}(2r_n),$$

since $R_n^2 \geqslant r_n$ and

$$\liminf_{n\to\infty} \Phi_{1/2}^{(t_n,x_n)}(2r_n) \geqslant \frac{\log 2}{2}\varepsilon$$

(as a consequence of $\phi^{(t_n,x_n)}(r_n) = \varepsilon$ along with (3.19)), so that

$$\liminf_{n\to\infty} \mathbb{E}\big[\big\|\big(U_2^n\big)^{-1}\nabla_0 U_2^n\big\|^2\big] \geqslant (\log 2)\varepsilon.$$

On the other hand, $\mathbb{E}[\|(U_2^n)^{-1}\nabla_0 U_2^n\|^2 \wedge N]$ converges to $\mathbb{E}[\|V_2\|^2 \wedge N]$ for any $N > 0$. We want to find an upper bound for

$$\mathbb{E}\big[\big\|\big(U_2^n\big)^{-1}\nabla_0 U_2^n\big\|^2 \mathbb{1}_{\{\|(U_2^n)^{-1}\nabla_0 U_2^n\|^2 > N\}}\big].$$

Let $V_s^n = (U_s^n)^{-1}\nabla_0 U_s^n$. We have, successively by Hölder inequality, Bienaymé–Tchebyshev inequality, Burkholder–Davis–Gundy inequality,

$$\begin{aligned}\mathbb{E}\big[\|V_2^n\|^2 \mathbb{1}_{\{\|V_2^n\|^2 > N\}}\big] &\leqslant \mathbb{E}\big[\|V_2^n\|^4\big]^{1/2}\mathbb{P}\big(\|V_2^n\|^2 > N\big)^{1/2} \leqslant \mathbb{E}\big[\|V_2^n\|^4\big]^{1/2}\frac{1}{N}\mathbb{E}\big[\|V_2^n\|^4\big]^{1/2} = \frac{1}{N}\mathbb{E}\big[\|V_2^n\|^4\big] \\ &\leqslant \frac{C_4}{N}\mathbb{E}\big[\langle V^n\rangle_2^2\big] \leqslant \frac{C_4}{N}\left(\frac{2^8}{\alpha^2}4\right)^2,\end{aligned}$$

where the constant $C_4 > 0$ comes from Burkholder–Davis–Gundy inequality:

$$\mathbb{E}\big[\|V_2^n\|^4\big] \leqslant C_4\,\mathbb{E}\big[\langle V^n\rangle_2^2\big].$$

Consequently, $\mathbb{E}[\|(U_2^n)^{-1}\nabla_0 U_2^n\|^2]$ converges to $\mathbb{E}[\|V_2\|^2]$ and

$$\mathbb{E}\big[\|V_2\|^2\big] \geqslant 2\liminf_{n\to\infty}\Phi_{1/2}^{(t_n,x_n)}(2r_n) \geqslant (\log 2)\varepsilon. \qquad \square$$

Corollary 5.4. *The solution ∇ to the Yang–Mills heat equation blows up at T if and only if there exists a sequence R_n converging to* 0*, a sequence t_n converging to T, and a sequence x_n in M such that the sequence of the laws of the processes*

$$\big((U^n)^{-1}\nabla_0 U^n\big)_{n\geqslant 0}$$

as defined in (5.4) *does not converge to* δ_0.

Proof. It is sufficient to prove that if ∇ does not blow up at T then any sequence $((U^n)^{-1}\nabla_0 U^n)_{n\geqslant 0}$ converges in law to δ_0. However, this is clear since there exists $C>0$ such that e is bounded by C on $[0,T]\times M$. Hence, for every $t\in[0,t_n/R_n^2]$,

$$\mathbb{E}\big[\big\|(U_t^n)^{-1}\nabla_0 U_t^n\big\|^2\big]=\mathbb{E}\Big[\Big\|\Big(U_{0,tR_n^2}^{(t_n,x_n)}\Big)^{-1}\nabla_0 U_{0,tR_n^2}^{(t_n,x_n)}\Big\|^2\Big]=2\int_0^{tR_n^2} s\,\mathbb{E}\big[e\big(t_n-s,X_s(x_n)\big)\big]\,\mathrm{d}s\leqslant C\,t^2R_n^4$$

which converges to 0 as n tends to ∞. □

6. Pontryagin numbers and ergodic theorem

In this section, we assume $\dim M=4$. Let ∇ be a Yang–Mills connection in E, $x\in M$, (X_s) a Brownian motion in M started at x, and denote by $U_s:E_x\to E_{X_s}$ the parallel transport in E along X_s (with respect to ∇). Let N^+ and N^- be the $L(T_xM,\operatorname{End}E_x)$-valued martingales

$$N_s^+=\int_0^s\sqrt{r}\,U_r^{-1}\big(R^\nabla\big)^+\big(\mathrm{d}_{\text{Itô}}^\nabla X_r,//_{0,r}\,\cdot\big)U_r\quad\text{and}\quad N_s^-=\int_0^s\sqrt{r}\,U_r^{-1}\big(R^\nabla\big)^-\big(\mathrm{d}_{\text{Itô}}^\nabla X_r,//_{0,r}\,\cdot\big)U_r,$$

where $(R^\nabla)^+$ (respectively $(R^\nabla)^-$) denote the self-dual (respectively antiself-dual) part of R^∇. Observe that $N_s^++N_s^-=U_s^{-1}\nabla_0U_s$, where $U_s(a,u)$ is parallel transport along

$$X_s(a,u)=\exp_{X_s}\big(a\sqrt{s}//_{0,s}u\big),\qquad u\in T_xM,$$

see Section 2.

Theorem 6.1. *As s tends to ∞, almost surely,*

$$\frac{1}{s^2}\big[\langle N^+,N^+\rangle_s-\langle N^-,N^-\rangle_s\big]\longrightarrow\frac{4\pi^2 i(E)}{\operatorname{vol}(M)},\quad\textit{where}$$

$$i(E):=\frac{1}{4\pi^2}\int_M\big[\big\|\big(R^\nabla\big)^+\big\|^2-\big\|\big(R^\nabla\big)^-\big\|^2\big](y)\,\mathrm{d}y \tag{6.1}$$

is the Pontryagin number of the bundle E, which is independent of ∇.

Proof. We know that

$$\frac{1}{s^2}\big[\langle N^+,N^+\rangle_s-\langle N^-,N^-\rangle_s\big]=\frac{2}{s^2}\int_0^s r\big[\big\|\big(R^\nabla\big)^+\big\|^2-\big\|\big(R^\nabla\big)^-\big\|^2\big](X_r)\,\mathrm{d}r=\frac{2}{s^2}\int_0^s rf(X_r)\,\mathrm{d}r,\quad\text{where}$$

$$f(y)=\big[\big\|\big(R^\nabla\big)^+\big\|^2-\big\|\big(R^\nabla\big)^-\big\|^2\big](y)\quad\text{for } y\in M.$$

Let $F(s)=\int_0^s f(X_r)\,\mathrm{d}r$. Integrating by parts gives

$$\frac{2}{s^2}\int_0^s rf(X_r)\,\mathrm{d}r=\frac{2}{s^2}\big[rF(r)\big]_0^s-\frac{2}{s^2}\int_0^s F(r)\,\mathrm{d}r=\frac{2}{s}F(s)-\frac{2}{s^2}\int_0^s r\frac{1}{r}F(r)\,\mathrm{d}r. \tag{6.2}$$

Since Brownian motion X is recurrent with $\mu(\mathrm{d}y)=\operatorname{vol}(M)^{-1}\,\mathrm{d}y$ as invariant measure where $\mathrm{d}y$ is the Riemannian measure, the ergodic theorem applies (see, e.g., [15, Theorem 1.3.12]). Thus $F(s)/s$ converges almost surely to $\int_M f(y)\,\mu(\mathrm{d}y)$ as s tends to ∞. Consequently, the last term of the right-hand side in (6.2) converges almost surely to $\int_M f(y)\,\mu(\mathrm{d}y)$ as s tends to ∞. According to definition (6.1), this proves the almost sure convergence of $2s^{-2}\int_0^s rf(X_r)\,\mathrm{d}r$ to $4\pi^2 i(E)/\operatorname{vol}(M)$ which is the wanted result. □

References

[1] M. Arnaudon, A. Thalmaier, Stability of stochastic differential equations in manifolds, in: Séminaire de Probabilités XXXII, Springer-Verlag, Berlin, 1998, pp. 188–214.
[2] M. Arnaudon, A. Thalmaier, Complete lifts of connections and stochastic Jacobi fields, J. Math. Pures Appl. (9) 77 (3) (1998) 283–315.
[3] M. Arnaudon, A. Thalmaier, Horizontal martingales in vector bundles, in: Séminaire de Probabilités XXXVI, Springer-Verlag, Berlin, 2001, in print.
[4] M. Arnaudon, A. Thalmaier, Yang–Mills fields and random holonomy along Brownian bridges, Preprint, 2001.
[5] R. Azencott, Une approche probabiliste du théorème de l'indice (Atiyah–Singer) (d'après J.-M. Bismut), Astérisque (133–134) (1986) 7–18;
R. Azencott, in: Seminar Bourbaki, Vol. 85, 1984.
[6] R.O. Bauer, Characterizing Yang–Mills fields by stochastic parallel transport, J. Funct. Anal. 155 (2) (1998) 536–549.
[7] R.O. Bauer, Yang–Mills fields and stochastic parallel transport in small geodesic balls, Stochastic Process. Appl. 89 (2) (2000) 213–226.
[8] J.-P. Bourguignon, H.B. Lawson Jr., Stability and isolation phenomena for Yang–Mills fields, Comm. Math. Phys. 79 (2) (1981) 189–230.
[9] Y. Chen, C.-L. Shen, Monotonicity formula and small action regularity for Yang–Mills flows in higher dimensions, Calc. Var. Partial Differential Equations 2 (4) (1994) 389–403.
[10] B.K. Driver, A. Thalmaier, Heat equation derivative formulas for vector bundles, J. Funct. Anal. 183 (1) (2001) 42–108.
[11] M. Émery, On two transfer principles in stochastic differential geometry, in: Séminaire de probabilités XXIV, Lecture Notes Math., Vol. 1426, Springer-Verlag, 1989, pp. 407–441.
[12] R.S. Hamilton, A matrix Harnack estimate for the heat equation, Comm. Anal. Geom. 1 (1) (1993) 113–126.
[13] R.S. Hamilton, Monotonicity formulas for parabolic flows on manifolds, Comm. Anal. Geom. 1 (1) (1993) 127–137.
[14] Y. Kannai, Off diagonal short time asymptotics for fundamental solutions of diffusion equations, Comm. Partial Differential Equations 2 (8) (1977) 781–830.
[15] H. Kunita, Stochastic Flows and Stochastic Differential Equations, Cambridge University Press, Cambridge, 1990.
[16] P. Malliavin, Stochastic Analysis, Springer-Verlag, Berlin, 1997.
[17] H. Naito, Finite time blowing-up for the Yang–Mills gradient flow in higher dimensions, Hokkaido Math. J. 23 (3) (1994) 451–464.
[18] P. Price, A monotonicity formula for Yang–Mills fields, Manuscripta Math. 43 (2–3) (1983) 131–166.
[19] J. Råde, On the Yang–Mills heat equation in two and three dimensions, J. Reine Angew. Math. 431 (1992) 123–163.
[20] R. Rebolledo, La méthode des martingales appliquée à l'étude de la convergence en loi de processus, Bull. Soc. Math. France Mém. (1979), (62) (1980).
[21] D. Revuz, M. Yor, Continuous Martingales and Brownian Motion, 3rd ed., Springer-Verlag, Berlin, 1999.
[22] S. Stafford, A stochastic criterion for Yang–Mills connections, in: Diffusion Processes and Related Problems in Analysis, Vol. I, (Evanston, IL, 1989), Birkhäuser Boston, Boston, MA, 1990, pp. 313–322.
[23] A.E. Schlatter, Long-time Behavior of the Yang–Mills Flow in Four Dimensions, PHD Thesis, Swiss Federal Institute of Technology, Zürich, 1994.
[24] A.E. Schlatter, Global existence of the Yang–Mills flow in four dimensions, J. Reine Angew. Math. 479 (1996) 133–148.
[25] M. Struwe, The evolution of harmonic maps of Riemannian surfaces, Comment. Math. Helv. 60 (1985) 558–581.
[26] M. Struwe, The Yang–Mills flow in four dimensions, Calc. Var. Partial Differential Equations 2 (2) (1994) 123–150.
[27] A. Thalmaier, Brownian motion and the formation of singularities in the heat flow for harmonic maps, Probab. Theory Related Fields 105 (3) (1996) 335–367.
[28] K. Uhlenbeck, Connections with L^p bounds on curvature, Comm. Math. Phys. 83 (1) (1982) 31–42.