



On the Schouten and Wagner curvature tensors

Dennis I. Barrett¹ · Claudiu C. Remsing²

Received: 21 July 2021 / Accepted: 10 August 2021 / Published online: 21 November 2021
© The Author(s), under exclusive licence to Springer-Verlag Italia S.r.l., part of Springer Nature 2021

Abstract

We recall and prove some basic properties of both the Schouten and Wagner curvature tensors. We consider—in some detail—the construction of the Wagner tensor, and then discuss how the vanishing of this tensor characterizes the flat structures, i.e., those for which the parallel translation is path-independent. (As a special case, we revisit the flatness of three-dimensional nonholonomic Riemannian manifolds.) In order to facilitate our presentation, we employ an extension of the notion of a connection to a “restricted connection.” Such connections are equivalently viewed as horizontal lifts (or horizontal distributions); hence we revisit the Schouten and Wagner tensors from this perspective. In particular, it turns out that the construction of the Wagner tensor may be equivalently formulated as a flag of horizontal distributions.

Keywords Nonholonomic Riemannian structure · Jacobi field · Schouten curvature tensor · Wagner curvature tensor

Mathematics Subject Classification 70G45 · 37J60

1 Introduction

A nonholonomic Riemannian structure is a quadruple $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$, consisting of a manifold M equipped with a nonintegrable distribution \mathcal{D} , a complementary distribution \mathcal{D}^\perp , and a (positive definite) fibre metric \mathbf{g} on \mathcal{D} . Similar to a Riemannian structure, a nonholonomic Riemannian structure has a unique connection associated to it, referred to as the “nonholonomic connection.” The admissible curves are those tangent to \mathcal{D} ; among these are the so-called nonholonomic geodesics, which are the geodesics of the nonholonomic connection. It turns out that these geodesics are precisely the solutions of the Chetaev equations for a nonholonomic mechanical system with constraints linear-in-velocities and a kinetic energy Lagrangian; accordingly, the chief motivation for the study of nonholonomic

✉ Dennis I. Barrett
d Barrett6@gmail.com

Claudiu C. Remsing
c.c.remsing@ru.ac.za

¹ Business Science Corporation, Sandton, Johannesburg 2196, South Africa

² Dept of Mathematics, Rhodes University, Grahamstown/Makhanda 6140, South Africa

Riemannian structures comes from mechanics. From a mathematical perspective, nonholonomic Riemannian geometry is a natural generalization of Riemannian geometry; in this sense, it may be viewed as a counterpart to sub-Riemannian geometry. (A sub-Riemannian structure is simply a triple $(M, \mathcal{D}, \mathbf{g})$, i.e., a nonholonomic Riemannian structure *sans* the choice of a complementary distribution.) Sub-Riemannian geometry generalizes the “metric” aspect of Riemannian geometry (the central object in sub-Riemannian geometry is the Carnot–Carathéodory distance), whereas nonholonomic Riemannian geometry generalizes the “connection” aspect (the nonholonomic connection is the central object). Along these lines, it was Hertz [16] (see also [28]) who essentially recognized that sub-Riemannian geometry studies the “shortest” curves, while nonholonomic Riemannian geometry studies the “straightest” curves. In general, “shortest” \neq “straightest”, although the two classes of curves coincide for a Riemannian structure. (We note that, in the past, the term *nonholonomic Riemannian geometry* was used to refer to both of the above generalizations of Riemannian geometry. However, we shall use it to refer exclusively to the first; the term *sub-Riemannian geometry* has now also become fairly standard.) For more information on nonholonomic Riemannian geometry and sub-Riemannian geometry, see, e.g., [6, 9, 20, 28, 29].

Curvature plays a central rôle in Riemannian geometry, and as such has been extensively studied: presently many aspects of it are very well understood. In sharp contrast, the curvature of nonholonomic Riemannian structures has received very little attention. Although some elements of the curvature of nonholonomic Riemannian structures can be found in Synge [26], it was Schouten [25] who first explicitly considered curvature in the nonholonomic Riemannian context. In particular, Schouten introduced a curvature tensor associated to every nonholonomic Riemannian structure; this tensor is now referred to as the “Schouten curvature tensor” [10]. Nevertheless, the main development in the study of curvature of nonholonomic Riemannian structures was due to Wagner, who observed that (the vanishing of) the Schouten tensor does not characterise the *flat* nonholonomic Riemannian structures, i.e., those structures for which the parallel translation (induced by the nonholonomic connection) is path-independent. In a series of papers [30, 33, 34], Wagner extended Schouten’s work, defining a curvature tensor (now called the “Wagner curvature tensor”), the vanishing of which *does* characterise the flat structures (see also [31, 32, 35]). (This resulted in Wagner being awarded Kazan University’s 1937 Lobachevskii prize for young Soviet mathematicians.) Nevertheless, Wagner’s construction has its limitations; in particular, it does not depend only on the data $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$, but also on some additional assumptions. As a result, it is not generally intrinsic, and so only provides a partial solution to the problem of flatness.

We briefly mention some more recent papers of interest that discuss and/or make use of the Schouten or Wagner curvature tensors. The two papers [1] and [3] are concerned with nonholonomic Riemannian structures in three dimensions, particularly left-invariant structures on Lie groups. The former paper classifies all such structures, and the latter considers the flat structures. Both papers make use of scalar invariants extracted from the Schouten curvature tensor. (Later in this paper we discuss some of the results of [3] in the context of the present paper.) Berestovskii [5] reviews various notions of curvature in sub-Riemannian geometry (which also apply to nonholonomic Riemannian structures), including the special case of sub-Riemannian structures with “rigged” distributions, i.e., those with a complementary distribution—the same data as for a nonholonomic Riemannian structure. (See also references in [5].) Galaev and coauthors have studied curvature (in particular, the Schouten and Wagner curvature tensors) in the context of (almost) contact metric structures [7, 11, 12] (see also

references therein); as well as the metrizable of nonholonomic connections in three dimensions [13] (making use of the Schouten tensor). Zhao and Jiao [36] study conformal transformations (of sub-Riemannian structures with “rigged” distribution), and from the Schouten tensor calculate the “nonholonomic” Weyl tensor. Leites [19] (see also [15]) calculates a nonholonomic analogue of the Riemann tensor, using algebraic techniques. However, the notion of flatness that Leites introduces is different from that which we consider here. Indeed, for Leites, every contact manifold is flat; however, every three-dimensional nonholonomic Riemannian manifold is contact, yet there are many such structures whose parallel transport is path-dependent (see [3]). Lastly, we mention the paper of Dragović and Gajić [10] and that of Gorbatenko [14], both of which discuss the motivation for, and construction of, Wagner’s curvature tensor. (For the construction of Wagner’s tensor in this paper, see Sect. 4.1.)

In this paper we consider in some more detail the two curvature tensors mentioned above, with a view to understanding better the curvature of nonholonomic Riemannian structures. The paper is organized as follows. In Sect. 3 we introduce the Schouten curvature tensor, which is canonically associated to every nonholonomic Riemannian structure. We prove some symmetries of this tensor, and attempt to relate it (at least on an algebraic level) to a Riemannian-type curvature tensor. We do this by decomposing the Schouten tensor (in fact, the associated tensor obtained by using the metric to “lower an index”) into two components: a “Riemannian” component—which satisfies all the symmetries of a Riemannian curvature tensor—and a “remainder” that can be viewed as a deviation of the Schouten tensor from a Riemannian curvature tensor. Using the “Riemannian” component of the Schouten tensor, we are able to introduce notions of sectional curvature, a Ricci tensor, and a scalar curvature, analogous to the corresponding concepts in Riemannian geometry.

In Sect. 4 we consider the Wagner curvature tensor. The construction of this tensor is quite sophisticated, and relies on the flag of the distribution; however, the construction is not intrinsic, in that it relies on some additional assumptions. (Having said that, if the distribution is strongly nonholonomic, then these assumptions are automatically satisfied.) We define the Wagner tensor in Sect. 4.1, and prove some basic properties: in particular, how the vanishing of the Wagner tensor characterizes the flat structures, i.e., those whose parallel translation is path-independent. In Sect. 4.2 we consider an “algebraic” interpretation of a collection of curvature tensors that arise in the construction of the Wagner tensor. (Briefly, these curvature tensors measure the extent to which certain maps fail to be homomorphisms.) Lastly, Sect. 4.3 considers the Wagner curvature tensor in the case of three-dimensional nonholonomic Riemannian structures. The flatness of such structures was recently treated in [3]; in particular, we find a new characterization of flatness in three dimensions, and relate it to that in [3].

Finally, in Sect. 5 we revisit both the Schouten and Wagner curvature tensors from an alternative viewpoint. The nonholonomic connection, as well as a collection of connections arising in the construction of Wagner’s curvature tensor, are equivalently viewed as horizontal lifts (or horizontal distributions). We express the Schouten tensor and the Wagner tensor in terms of horizontal lifts of vector fields. (The vanishing of these tensors then translates to involutivity conditions on the associated horizontal distributions.) A significant contribution of this section is also to show that Wagner’s construction is equivalently formulated as a flag of horizontal distributions.

2 Preliminaries

In this section we revisit some basic concepts and constructions in tensor analysis and the theory of connections; these are slightly extended (or rather, adapted) to a particular class of objects pertaining to our line of inquiry into nonholonomic Riemannian geometry. Specifically, we first consider derivations of tensors on a distribution. If there exists a projection onto the distribution, then there exists a natural generalization of the Lie derivative (essentially a “projected Lie derivative”); this will lead on to a natural notion of a “restricted” tensor derivation. Next, we consider a class of affine connections whose associated parallel translation is restricted to a subclass of curves. We consider two approaches to these “restricted” connections: as a covariant derivative operator, and as a horizontal lift. Following this we present some necessary elements of nonholonomic Riemannian geometry. Apart from some definitions, we recall the fundamental existence and uniqueness result for the nonholonomic connection (a restricted connection) and introduce its associated exterior covariant derivative.

Let M be a (real, n -dimensional) manifold and let \mathcal{D} and \mathcal{E} be distributions on M , where \mathcal{D} has rank r . Throughout, we assume that all manifolds, functions, vector fields, etc. are smooth (i.e., of class C^∞) and that all distributions under consideration are *regular* (i.e., the dimension of each fiber does not depend on the base point). Furthermore, we shall follow the summation convention on repeated indices. Unless stated otherwise, the following ranges on indices are used: $i, j, k = 1, \dots, n$ and $a, b, c = 1, \dots, r$.

2.1 Restricted tensor derivations

Let $T_\ell^k(\mathcal{D})$ be the bundle of (k, ℓ) -tensors on \mathcal{D} , i.e., $T_\ell^k(\mathcal{D}) = \bigsqcup_{q \in M} \left[\otimes^k \mathcal{D}_q \otimes \otimes^\ell \mathcal{D}_q^* \right]$, and let $T_\ell^k(\mathcal{D}) = \Gamma(T_\ell^k(\mathcal{D}))$ be the space of tensor fields on \mathcal{D} . A *derivation* of $T_\ell^k(\mathcal{D})$ (cf. [27]) is a collection of \mathbb{R} -linear maps $\delta_\ell^k : T_\ell^k(\mathcal{D}) \rightarrow T_\ell^k(\mathcal{D})$ (for every $k, \ell \geq 0$), all denoted by δ when convenient, such that

$$\delta(S \otimes T) = \delta(S) \otimes T + S \otimes \delta(T) \quad \text{and} \quad \delta(\text{tr}_j^i T) = \text{tr}_j^i \delta(T)$$

for all tensor fields S and T . Here tr_j^i denotes the contraction (trace) on the i th contravariant and j th covariant index. The space $\text{Der}(\mathcal{D})$ of all derivations of $T_\ell^k(\mathcal{D})$ is a Lie algebra.

If δ is a derivation and T is a (k, ℓ) -tensor field on \mathcal{D} , then

$$\begin{aligned} \delta(T)(\omega^1, \dots, \omega^k, X_1, \dots, X_\ell) &= \delta(T(\omega^1, \dots, \omega^k, X_1, \dots, X_\ell)) \\ &\quad - \sum_{i=1}^k T(\omega^1, \dots, \delta(\omega^i), \dots, \omega^k, X_1, \dots, X_\ell) \\ &\quad - \sum_{j=1}^\ell T(\omega^1, \dots, \omega^k, X_1, \dots, \delta(X_j), \dots, X_\ell) \end{aligned}$$

for $\omega^1, \dots, \omega^k \in \Gamma(\mathcal{D}^*)$ and $X_1, \dots, X_\ell \in \Gamma(\mathcal{D})$. In particular, if $\omega \in \Gamma(\mathcal{D}^*)$, then $\delta(\omega)(X) = \delta(\omega(X)) - \omega(\delta(X))$ for $X \in \Gamma(\mathcal{D})$. Derivations of $T_\ell^k(\mathcal{D})$ are completely specified by their action on $T_0^0(\mathcal{D}) = C^\infty(M)$ and on $T_0^1(\mathcal{D}) = \Gamma(\mathcal{D})$. Accordingly, given a map that acts as a derivation of $C^\infty(M)$ and $\Gamma(\mathcal{D})$, it may be uniquely extended to a derivation of $T_\ell^k(\mathcal{D})$.

Let $\mathcal{P} : TM \rightarrow \mathcal{D}$ be a projection and let $[[\cdot, \cdot]] = \mathcal{P}([\cdot, \cdot]) : \Gamma(TM) \times \Gamma(TM) \rightarrow \Gamma(\mathcal{D})$. If $Z \in \Gamma(TM)$, then we define a derivation $\mathcal{L}_Z^\mathcal{P} \in \text{Der}(\mathcal{D})$ by the requirement that

$$\mathcal{L}_Z^{\mathcal{P}}f = Z[f] \quad \text{and} \quad \mathcal{L}_Z^{\mathcal{P}}X = \llbracket Z, X \rrbracket \tag{1}$$

for $f \in C^\infty(M)$ and $X \in \Gamma(\mathcal{D})$. We refer to $\mathcal{L}_Z^{\mathcal{P}}$ as the \mathcal{P} -Lie derivative along Z . It turns out that every element of $\text{Der}(\mathcal{D})$ decomposes uniquely as the sum of a \mathcal{P} -Lie derivative and a derivation that vanishes on functions.

Proposition 1 *Let $\delta \in \text{Der}(\mathcal{D})$; then:*

- (i) $\delta_0^0 = 0$ if and only if δ_0^1 is $C^\infty(M)$ -linear, i.e., $\delta_0^1 \in \mathcal{T}_0^1(\mathcal{D})$.
- (ii) There exists a unique vector field $Z \in \Gamma(TM)$ and a unique $\delta_0^1 \in \mathcal{T}_1^1(\mathcal{D})$ such that $\delta = \mathcal{L}_Z^{\mathcal{P}} + \delta'$, where $\delta_0^0 = 0$ and $\delta_0^1 \in \mathcal{T}_0^1(\mathcal{D})$.

Let $\text{Der}_0(\mathcal{D}) = \{\delta \in \text{Der}(\mathcal{D}) \text{ and } \mathcal{L}_{TM}^{\mathcal{P}} = \{\mathcal{L}_Z^{\mathcal{P}} : Z \in \Gamma(TM)\}$. We have the (vector space) decomposition $\text{Der}(\mathcal{D}) = \mathcal{L}_{TM}^{\mathcal{P}} \oplus \text{Der}_0(\mathcal{D})$; furthermore, $[\text{Der}(\mathcal{D}), \text{Der}_0(\mathcal{D})] \subseteq \text{Der}_0(\mathcal{D})$. We say that a derivation $\delta \in \text{Der}(\mathcal{D})$ is an \mathcal{E} -restricted derivation (or simply \mathcal{E} -derivation) if $\delta \in \mathcal{L}_{\mathcal{E}}^{\mathcal{P}} \oplus \text{Der}_0(\mathcal{D})$, where $\mathcal{L}_{\mathcal{E}}^{\mathcal{P}} = \{\mathcal{L}_X^{\mathcal{P}} : X \in \Gamma(\mathcal{E})\}$. Let $\text{Der}_{\mathcal{E}}(\mathcal{D})$ denote the space of all \mathcal{E} -derivations.

Remark 1 Dual to the \mathcal{P} -Lie derivative is the \mathcal{P} -exterior derivative $d_{\mathcal{P}} : \Omega^k(\mathcal{D}) \rightarrow \Omega^{k+1}(\mathcal{D})$, defined as follows: if $f \in C^\infty(M)$, then $d_{\mathcal{P}}f(X) = X[f]$ for every $X \in \Gamma(\mathcal{D})$; if $\omega \in \Omega^k(\mathcal{D})$, $k \geq 1$, then

$$d_{\mathcal{P}}\omega(X_0, \dots, X_k) = \sum_{i=0}^k (-1)^i (\mathcal{L}_{X_i}^{\mathcal{P}}\omega)(X_0, \dots, \widehat{X}_i, \dots, X_k)$$

for $X_0, \dots, X_k \in \Gamma(\mathcal{D})$ (where \widehat{X}_i indicates the omission of that element). Many properties of the usual exterior derivative extend to $d_{\mathcal{P}}$; however, we do not generally have $d_{\mathcal{P}}^2 = 0$. Indeed, if $f \in C^\infty(M)$, then $d_{\mathcal{P}}^2f(X, Y) = [X, Y][f] - \mathcal{P}([X, Y])[f]$ for $X, Y \in \Gamma(\mathcal{D})$. Hence $d_{\mathcal{P}}^2f = 0$ if and only if \mathcal{D} is integrable.

2.2 Restricted connections

In this section we consider so-called “restricted connections,” i.e., connections whose associated parallel translation is restricted to precisely those curves that are tangent to a given distribution. Firstly, we consider a restricted connection as a covariant derivative. In this vein, we then discuss the associated parallel translation map, parallel tensor fields, and parallel frames. Secondly, we consider a restricted connection as a horizontal lift, or equivalently, as a horizontal distribution. (Typically, the horizontal distribution will not form a full complement to the vertical distribution.) As expected, such a horizontal lift induces a unique covariant derivative, and conversely. Restricted connections (in the language of covariant derivatives) were first introduced in [17]; the notion of a restricted connection (particularly, as a horizontal lift) is also essentially covered in [8].

We assume that any two points of M can be joined by an \mathcal{E} -curve, i.e., a curve $\gamma : [0, 1] \rightarrow M$ such that $\dot{\gamma}(t) \in \mathcal{E}_{\gamma(t)}$ for every $t \in [0, 1]$. (This holds, for instance, when \mathcal{E} is completely nonholonomic.) An \mathcal{E} -restricted covariant derivative ∇ on \mathcal{D} (or covariant

\mathcal{E} -derivative on \mathcal{D}) is an \mathbb{R} -linear mapping $\nabla : \Gamma(\mathcal{E}) \times \Gamma(\mathcal{D}) \rightarrow \Gamma(\mathcal{D})$, $(X, W) \mapsto \nabla_X W$ such that

$$\nabla_{fX} W = f\nabla_X W \quad \text{and} \quad \nabla_X fW = X[f]W + f\nabla_X W$$

for every $f \in C^\infty(\mathbb{M})$, $X \in \Gamma(\mathcal{E})$ and $W \in \Gamma(\mathcal{D})$. The usual basic properties of covariant differentiation extend to the case of a restricted covariant derivative. In particular, the expression $\nabla_X W(q)$, $q \in \mathbb{M}$ depends only on the value of X at q , and the value of W along any \mathcal{E} -curve tangent to $X(q)$. Covariant differentiation may also be uniquely extended to arbitrary tensor fields in $T^k_\ell(\mathcal{D})$, in the usual fashion: let $X \in \Gamma(\mathcal{E})$ and define $\nabla_X f = X[f]$ for $f \in C^\infty(\mathbb{M})$; then ∇_X is a derivation of $C^\infty(\mathbb{M})$ and $\Gamma(\mathcal{D})$, and hence may be extended to a (unique) derivation of $T^k_\ell(\mathcal{D})$.

A covariant \mathcal{E} -derivative ∇ on \mathcal{D} induces a parallel translation along \mathcal{E} -curves. A section V of \mathcal{D} along an \mathcal{E} -curve $\gamma : [0, 1] \rightarrow \mathbb{M}$ is *parallel along γ* if $\nabla_{\dot{\gamma}} V = 0$. Let $W \in \Gamma(\mathcal{D})$; if $W \circ \gamma$ is parallel along γ for every \mathcal{E} -curve γ , then W is simply called *parallel*. Clearly, a necessary and sufficient condition for W to be parallel is that $\nabla W \equiv 0$.

Proposition 2 *Let $\gamma : [0, 1] \rightarrow \mathbb{M}$ be an \mathcal{E} -curve and let $V_0 \in \mathcal{D}_{\gamma(0)}$. There exists a unique parallel section V of \mathcal{D} along γ such that $V(0) = V_0$. (V is called the parallel translate of V_0 along γ .)*

Let $\gamma : [0, 1] \rightarrow \mathbb{M}$ be an \mathcal{E} -curve. The *parallel translation* $\Pi^t_\gamma : \mathcal{D}_{\gamma(0)} \rightarrow \mathcal{D}_{\gamma(t)}$, $t \in [0, 1]$ is specified by setting $\Pi^t_\gamma(V_0) = V(t)$, where V is the parallel translate of $V_0 \in \mathcal{D}_{\gamma(0)}$ along γ . A (local) frame (U_a) for \mathcal{D} is called *parallel* if each element U_a is parallel. The existence of a parallel frame for \mathcal{D} is not guaranteed; in fact, it places quite severe restrictions on the connection.

Proposition 3 *There exists a parallel frame for \mathcal{D} on an open set $\mathcal{U} \subseteq \mathbb{M}$ if and only if for any two points $p, q \in \mathcal{U}$ and for any \mathcal{E} -curve $\gamma : [0, 1] \rightarrow \mathcal{U}$ joining p to q , the parallel translation $\Pi^1_\gamma : \mathcal{D}_p \rightarrow \mathcal{D}_q$ does not depend on γ .*

The notion of a parallel vector field is easily extended to arbitrary tensor fields. Indeed, a section A of $T^k_\ell(\mathcal{D})$ along an \mathcal{E} -curve γ is called *parallel along γ* if $\nabla_{\dot{\gamma}} A = 0$. A tensor field $T \in T^k_\ell(\mathcal{D})$ is *parallel* if $T \circ \gamma$ is parallel along γ for every \mathcal{E} -curve γ . Evidently, T is parallel if and only if $\nabla T \equiv 0$.

Let $\tau_M : TM \rightarrow \mathbb{M}$ and $\tau_D : T\mathcal{D} \rightarrow \mathcal{D}$ denote the canonical projections of a tangent vector onto its base point. Let $\pi = \tau_M|_{\mathcal{D}} : \mathcal{D} \rightarrow \mathbb{M}$. The pullback bundle $\pi^*\mathcal{E}$ may be viewed as a vector bundle over \mathcal{D} and over \mathcal{E} , with projections $\tilde{\pi}_1 : \pi^*\mathcal{E} \ni (U_q, X_q) \mapsto U_q \in \mathcal{D}$ and $\tilde{\pi}_2 : \pi^*\mathcal{E} \ni (U_q, X_q) \mapsto X_q \in \mathcal{E}$, respectively. An \mathcal{E} -restricted connection on \mathcal{D} (or \mathcal{E} -connection on \mathcal{D}) is a map $h : \pi^*\mathcal{E} \rightarrow T\mathcal{D}$ that is

- (i) a linear bundle map from $\tilde{\pi}_1$ to τ_D covering the identity, i.e., $\tau_D \circ h = \tilde{\pi}_1$;
- (ii) a bundle map from $\tilde{\pi}_2$ to $T\pi$ covering the inclusion $\iota : \mathcal{E} \rightarrow TM$, i.e., $T\pi \cdot h = \iota \circ \tilde{\pi}_2$.

Such a connection h is called *linear* if $T_{U_q} \phi_t \cdot h(U_q, X_q) = h(\phi_t(U_q), X_q)$ for every $(U_q, X_q) \in \pi^*\mathcal{E}$, where $\phi_t : \mathcal{D} \rightarrow \mathcal{D}$ denotes the canonical dilation $\phi_t(U_q) = e^t U_q$. Let $\mathcal{V} = \ker T\pi$ be the *vertical distribution* and $\mathcal{H} = \text{im } h$ the *horizontal distribution*. The map $\nu : \pi^*\mathcal{D} \rightarrow \mathcal{V}$ given by

$$v(U_q, V_q) = \left. \frac{d}{dt} \right|_{t=0} (U_q + tV_q), \quad (U_q, V_q) \in \pi^*\mathcal{D}$$

is a vector bundle isomorphism. If $U_q \in \mathcal{D}$, then the linear isomorphism $v_{U_q} = v(U_q, \cdot) : \mathcal{D}_q \rightarrow \mathcal{V}_{U_q}$ is called the *vertical lift over U_q* . If $X \in \Gamma(\mathcal{D})$, then we define the *vertical lift of X* to be the vertical vector field $X^v \in \Gamma(\mathcal{V})$ given by

$$X^v(U_q) = v(U_q, X(q)), \quad U_q \in \mathcal{D}.$$

Similarly, the mapping $h_{U_q} = h(U_q, \cdot) : \mathcal{D}_q \rightarrow \mathcal{H}_{U_q}$ is called the *horizontal lift over U_q* (or explicitly, the *h-lift over U_q*). The *horizontal lift (or h-lift) of $X \in \Gamma(\mathcal{E})$* is the vector field $X^h \in \Gamma(\mathcal{H})$ given by

$$X^h(U_q) = h(U_q, X(q)), \quad U_q \in \mathcal{D}.$$

Projectable horizontal vector fields (i.e., vector fields $Z \in \Gamma(\mathcal{H})$ for which there exists $X \in \Gamma(TM)$ such that $T\pi \cdot Z = X \circ \pi$) are exactly the horizontal lifts of vector fields in $\Gamma(\mathcal{E})$. The following proposition collects some basic results about restricted connections.

Proposition 4 (cf. [8]) *We have:*

- (i) $\pi_*\mathcal{H} = \mathcal{E}$, i.e., $T_{U_q}\pi \cdot \mathcal{H}_{U_q} = \mathcal{E}_q$ for every $U_q \in \mathcal{D}$.
- (ii) h linear if and only if $(\phi_t)_*\mathcal{H} = \mathcal{H}$.
- (iii) $\mathcal{V} \cap \mathcal{H} = \{0\}$, and $\mathcal{V} + \mathcal{H} \subseteq T\mathcal{D}$ with equality if and only if $\mathcal{E} = TM$.

Furthermore, we have that h is uniquely specified by its associated horizontal distribution \mathcal{H} :

Proposition 5 (cf. [8]) *Let \mathcal{E} and \mathcal{D} be distributions on \mathbf{M} and $\pi = \tau_{\mathbf{M}}|_{\mathcal{D}} : \mathcal{D} \rightarrow \mathbf{M}$. If \mathcal{H} is a distribution on \mathcal{D} such that*

$$\pi_*\mathcal{H} = \mathcal{E} \quad \text{and} \quad \mathcal{V} \cap \mathcal{H} = \{0\},$$

then there exists a unique (not necessarily linear) \mathcal{E} -connection h on \mathcal{D} such that $\mathcal{H} = \text{im } h$.

We briefly address the relation between (linear) restricted connections and restricted covariant derivatives. Suppose that h is a linear \mathcal{E} -connection on \mathcal{D} . Let $X_q \in \mathcal{E}$ and $Y_{U_q} \in T_{U_q}\mathcal{D}$, where $U_q \in \mathcal{D}_q$ and Y_{U_q} satisfies $T_{U_q}\pi \cdot Y_{U_q} = X_q$. Then it is easy to see that $Y_{U_q} - h(U_q, X_q)$ is vertical, and hence we can define $\nabla : \Gamma(\mathcal{E}) \times \Gamma(\mathcal{D}) \rightarrow \Gamma(\mathcal{D})$ by

$$\nabla_X U(q) = v_{U(q)}^{-1} \cdot [T_{U_q} U \cdot X(q) - h(U(q), X(q))], \quad X \in \Gamma(\mathcal{E}), U \in \Gamma(\mathcal{D}).$$

∇ is precisely a covariant \mathcal{E} -derivative on \mathcal{D} . Conversely, given an covariant \mathcal{E} -derivative ∇ on \mathcal{D} , there exists a unique linear \mathcal{E} -connection h on \mathcal{D} whose associated covariant derivative is exactly ∇ .

Proposition 6 (cf. [8]) *Let ∇ be a covariant \mathcal{E} -derivative on \mathcal{D} ; then ∇ is the covariant derivative associated to the unique \mathcal{E} -connection h on \mathcal{D} given by*

$$h(U_q, X_q) = T_q U \cdot X_q - v_{U_q} \cdot \nabla_{X_q} U, \quad (U_q, X_q) \in \pi^* \mathcal{E}.$$

Here $U \in \Gamma(\mathcal{D})$ satisfies $U(q) = U_q$. (The definition of h does not depend on the choice of U .)

Lastly, we state two technical results (to be used later in the paper). If $\omega \in \Gamma(\mathcal{D}^*)$, then we shall denote by $\bar{\omega} \in \mathcal{C}^\infty(\mathcal{D})$ the function given by $\bar{\omega}(U_q) = \omega_q(U_q)$.

Lemma 1 *If $W \in \Gamma(\mathcal{V})$ and $\omega \in \Gamma(\mathcal{D}^*)$, then*

$$W[\bar{\omega}](U_q) = \omega_q(v_{U_q}^{-1} \cdot W(U_q))$$

for every $U_q \in \mathcal{D}$.

Lemma 2 *If $X, Y \in \Gamma(\mathcal{E})$ and $\omega \in \Gamma(\mathcal{D}^*)$, then:*

- (i) $X^h[\bar{\omega}] = \overline{\nabla_X \omega}$.
- (ii) $[X^h, Y^h][\bar{\omega}] = \overline{[\nabla_X, \nabla_Y] \omega}$.

2.3 Nonholonomic Riemannian structures

Given a distribution \mathcal{D} , its *flag* is the filtration $\mathcal{D}^1 \subseteq \mathcal{D}^2 \subseteq \dots$ given by

$$\mathcal{D}^1 = \mathcal{D} \quad \text{and} \quad \mathcal{D}^{i+1} = \mathcal{D}^i + [\mathcal{D}^i, \mathcal{D}^i], \quad i \geq 1.$$

We shall always assume that each component of the flag is regular. Evidently, the flag of \mathcal{D} will stabilise after finitely many steps. If there exists $N \geq 2$ such that $\mathcal{D}^{N-1} \subsetneq TM$ and $\mathcal{D}^N = TM$, then \mathcal{D} is said to be *completely nonholonomic*, in which case N is the *degree of nonholonomy* of \mathcal{D} . When $N = 2$, \mathcal{D} is called *strongly nonholonomic*. Complete nonholonomy of \mathcal{D} is a sufficient condition for any two points of M to be joined by a \mathcal{D} -curve (Chow–Rashevskii; see, e.g., [20]). A *nonholonomic Riemannian manifold* is a quadruple $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$, where M is an n -dimensional manifold, \mathcal{D} is a rank $r < n$ completely nonholonomic distribution on M , \mathcal{D}^\perp is a rank $n - r$ distribution on M complementary to \mathcal{D} (i.e., $TM = \mathcal{D} \oplus \mathcal{D}^\perp$) and \mathbf{g} is a positive definite fiber metric on \mathcal{D} . For convenience, we also refer to a nonholonomic Riemannian manifold as a *nonholonomic Riemannian structure*. Let $\mathcal{P} : TM \rightarrow \mathcal{D}$ and $\mathcal{Q} : TM \rightarrow \mathcal{D}^\perp$ be the projectors corresponding to the decomposition $TM = \mathcal{D} \oplus \mathcal{D}^\perp$. As before, we denote the projected Lie bracket $\mathcal{P}([\cdot, \cdot]) : \Gamma(TM) \times \Gamma(TM) \rightarrow \Gamma(\mathcal{D})$ by $[[\cdot, \cdot]]$.

Remarkably, the existence and uniqueness result for the Levi-Civita connection (in the Riemannian case) generalizes to nonholonomic Riemannian geometry; more specifically, associated to every nonholonomic Riemannian structure $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ is a unique metric and torsion-free \mathcal{D} -connection ∇ on \mathcal{D} . Here the *torsion* T of ∇ is given by

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [[X, Y]], \quad X, Y \in \Gamma(\mathcal{D}).$$

∇ is referred to as the *nonholonomic connection* of $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$. Formally, we have the following statement.

Proposition 7 (see, e.g., [18]) *Let $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ be a nonholonomic Riemannian structure. There exists a unique \mathcal{D} -connection ∇ on \mathcal{D} such that $\nabla \mathbf{g} \equiv 0$ and $T \equiv 0$, i.e.,*

$$Z[\mathbf{g}(X, Y)] = \mathbf{g}(\nabla_Z X, Y) + \mathbf{g}(X, \nabla_Z Y) \quad \text{and} \quad \nabla_X Y - \nabla_Y X = \llbracket X, Y \rrbracket$$

for every $X, Y, Z \in \Gamma(\mathcal{D})$. Furthermore, ∇ is characterized by Koszul’s formula:

$$2 \mathbf{g}(\nabla_X Y, Z) = X[\mathbf{g}(Y, Z)] + Y[\mathbf{g}(X, Z)] - Z[\mathbf{g}(X, Y)] + \mathbf{g}(\llbracket X, Y \rrbracket, Z) - \mathbf{g}(\llbracket X, Z \rrbracket, Y) - \mathbf{g}(\llbracket Y, Z \rrbracket, X).$$

A nonholonomic Riemannian structure $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ is said to be *flat on \mathcal{U}* (where $\mathcal{U} \subseteq M$ is open) if there exists a parallel frame for \mathcal{D} (with respect to the nonholonomic connection) defined on \mathcal{U} . If $\mathcal{U} = M$, then we simply say that $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ is *flat*; on the other hand, if $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ is flat on an open neighbourhood about every point in M , then we say that it is *locally flat*.

The \mathcal{P} -exterior covariant derivative associated to ∇ , denoted $d_{\mathcal{P}}^\nabla : \Omega^k(\mathcal{D}, \mathcal{D}) \rightarrow \Omega^{k+1}(\mathcal{D}, \mathcal{D})$, is defined as follows:

- (i) If $U \in \Omega^0(\mathcal{D}, \mathcal{D}) = \Gamma(\mathcal{D})$, then $d_{\mathcal{P}}^\nabla U(X) = \nabla_X U$ for every $X \in \Gamma(\mathcal{D})$.
- (ii) If $\varphi \in \Omega^k(\mathcal{D}, \mathcal{D})$, $k \geq 1$, then

$$d_{\mathcal{P}}^\nabla \varphi(X_0, \dots, X_k) = \sum_{i=0}^k (-1)^i \nabla_{X_i} \varphi(X_0, \dots, \hat{X}_i, \dots, X_k) + \sum_{0 \leq i < j \leq k} (-1)^{i+j} \varphi(\llbracket X_i, X_j \rrbracket, X_0, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_k)$$

for $X_0, \dots, X_k \in \Gamma(\mathcal{D})$. (Here \hat{X}_i indicates the omission of that element.)

In particular, for a \mathcal{D} -valued 1-form φ , we have

$$d_{\mathcal{P}}^\nabla \varphi(X, Y) = \nabla_X \varphi(Y) - \nabla_Y \varphi(X) - \varphi(\llbracket X, Y \rrbracket),$$

where $X, Y \in \Gamma(\mathcal{D})$. (Note that the torsion of ∇ is exactly the \mathcal{P} -exterior covariant derivative of the identity map $\text{id}_{\mathcal{D}} : \mathcal{D} \rightarrow \mathcal{D}$.)

3 The Schouten curvature tensor

Let $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ be a nonholonomic Riemannian manifold, with associated nonholonomic connection ∇ . As \mathcal{D} is nonintegrable, the standard curvature tensor $(X, Y, Z) \mapsto [\nabla_X, \nabla_Y]Z - \nabla_{\llbracket X, Y \rrbracket}Z$ is no longer defined. Instead, we have the *Schouten curvature tensor* $K \in T_3^1(\mathcal{D})$ [10]:

$$K(X, Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{\llbracket X, Y \rrbracket}Z - \llbracket \mathcal{Q}(\llbracket X, Y \rrbracket), Z \rrbracket.$$

K is clearly skew-symmetric in its first two arguments; hence we may also view it as the mapping $\Gamma(\wedge^2 \mathcal{D}) \rightarrow T_1^1(\mathcal{D})$, $K(X \wedge Y)Z = K(X, Y)Z$. The associated $(0, 4)$ -tensor, which we denote \hat{K} , is given by $\hat{K}(W, X, Y, Z) = \mathbf{g}(K(W, X)Y, Z)$.

The symmetries of the Riemannian curvature tensor are well known (see, e.g., [22]). In contrast, not all of those symmetries hold for the Schouten curvature tensor. A straightforward computation yields the following result.

Lemma 3 *Let $X, Y, Z \in \Gamma(\mathcal{D})$; then:*

- (i) $(d^{\nabla}_{\mathcal{P}})^2 Z(X, Y) = K(X, Y)Z + \llbracket \mathcal{Q}([X, Y]), Z \rrbracket$.
- (ii) $(d^{\nabla}_{\mathcal{P}})^2 \text{id}_{\mathcal{D}}(X, Y, Z) = K(X, Y)Z + K(Y, Z)X + K(Z, X)Y$.

Since ∇ is torsion free, the first Bianchi identity holds; i.e., $(d^{\nabla}_{\mathcal{P}})^2 \text{id}_{\mathcal{D}} \equiv 0$, or equivalently,

$$K(X, Y)Z + K(Y, Z)X + K(Z, X)Y = 0.$$

Remark 2 The second Bianchi identity does not generally hold for the Schouten tensor. Indeed, we may view K as an element of $\Omega^2(\mathcal{D}, T^1_1(\mathcal{D}))$. Furthermore, the nonholonomic connection extends to a connection $\nabla^* : \Gamma(\mathcal{D}) \times T^1_1(\mathcal{D}) \rightarrow T^1_1(\mathcal{D})$, which has an associated \mathcal{P} -exterior covariant derivative $d^{\nabla^*}_{\mathcal{P}} : \Omega^k(\mathcal{D}, T^1_1(\mathcal{D})) \rightarrow \Omega^{k+1}(\mathcal{D}, T^1_1(\mathcal{D}))$. The classical identity is $d^{\nabla^*}_{\mathcal{P}} K \equiv 0$ (see, e.g., [23]); however, in this case we have

$$d^{\nabla^*}_{\mathcal{P}} K(X, Y, Z) = \nabla_{\mathcal{J}_{\mathcal{P}}(X, Y, Z)} - \mathcal{L}_{\mathcal{J}_{\mathcal{Q}}(X, Y, Z)} - \sum_{\mathcal{O}(X, Y, Z)} [\nabla_X, \mathcal{L}_{\mathcal{Q}([Y, Z])}].$$

Here $\sum_{\mathcal{O}(X, Y, Z)}$ indicates the sum over the cyclic permutations of (X, Y, Z) and $\mathcal{J}_{\mathcal{P}}$ and $\mathcal{J}_{\mathcal{Q}}$ are the Jacobiators of $\llbracket \cdot, \cdot \rrbracket$ and $\mathcal{Q}([\cdot, \cdot])$, respectively (i.e., $\mathcal{J}_{\mathcal{P}}(X, Y, Z) = \sum_{\mathcal{O}(X, Y, Z)} \llbracket [X, Y], Z \rrbracket$ and similarly for $\mathcal{J}_{\mathcal{Q}}$).

The $(0, 4)$ -tensor \widehat{K} satisfies the following symmetries:

- (S1) $\widehat{K}(W, X, Y, Z) + \widehat{K}(X, W, Y, Z) = 0$.
- (S2) $\widehat{K}(W, X, Y, Z) + \widehat{K}(W, X, Z, Y) + \widehat{K}(W, X, Y, Z) = 0$.

However, in contrast to the Riemannian $(0, 4)$ -tensor, \widehat{K} is generally not skew-symmetric in the final two arguments, nor is it symmetric if one swaps the first two arguments with the last two. We decompose \widehat{K} into two tensors \widehat{R} and \widehat{C} , where \widehat{R} is the component of \widehat{K} that is skew-symmetric in the last two arguments and \widehat{C} is the component that is symmetric in the last two arguments. Specifically, we define $\widehat{R}, \widehat{C} \in T^0_4(\mathcal{D})$ as

$$\widehat{R}(W, X, Y, Z) = \frac{1}{2}[\widehat{K}(W, X, Y, Z) - \widehat{K}(W, X, Z, Y)], \quad \widehat{C} = \widehat{K} - \widehat{R}.$$

\widehat{R} and \widehat{C} both satisfy the same two symmetries as \widehat{K} (i.e., (S1) and (S2)); furthermore, we have

- (S3) $\widehat{R}(W, X, Y, Z) + \widehat{R}(W, X, Z, Y) = 0$.
- (S4) $\widehat{R}(W, X, Y, Z) = \widehat{R}(Y, Z, W, X)$. (This follows from the first three symmetries.)
- (S5) $\widehat{C}(W, X, Y, Z) = \widehat{C}(W, X, Z, Y)$.

Proposition 8 *Let $W, Y, X, Z \in \Gamma(\mathcal{D})$. Then*

$$\widehat{C}(W, X, Y, Z) = \frac{1}{2}(\mathcal{L}_{\mathcal{Q}([W, X])}^{\mathcal{P}}\mathbf{g})(Y, Z).$$

Proof From the definition of \widehat{C} , we have

$$\begin{aligned} 2\widehat{C}(W, X, Y, Z) &= \widehat{K}(W, X, Y, Z) + \widehat{K}(W, X, Z, Y) \\ &= \mathbf{g}(K(W, X)Y, Z) + \mathbf{g}(Y, K(W, X)Z) \\ &= \mathbf{g}([\nabla_W, \nabla_X]Y, Z) + \mathbf{g}(Y, [\nabla_W, \nabla_X]Z) \\ &\quad - \mathbf{g}(\nabla_{[W, X]}Y, Z) - \mathbf{g}(Y, \nabla_{[W, X]}Z) \\ &\quad - \mathbf{g}(\llbracket \mathcal{Q}([W, X]), Y \rrbracket, Z) - \mathbf{g}(Y, \llbracket \mathcal{Q}([W, X]), Z \rrbracket). \end{aligned}$$

The first two terms are

$$\begin{aligned} \mathbf{g}([\nabla_W, \nabla_X]Y, Z) &= \mathbf{g}(\nabla_W \nabla_X Y, Z) - \mathbf{g}(\nabla_X \nabla_W Y, Z) \\ &= W[\mathbf{g}(\nabla_X Y, Z)] - \mathbf{g}(\nabla_X Y, \nabla_W Z) - X[\mathbf{g}(\nabla_W Y, Z)] \\ &\quad + \mathbf{g}(\nabla_W Y, \nabla_X Z) \\ &= [W, X][\mathbf{g}(Y, Z)] - W[\mathbf{g}(Y, \nabla_X Z)] - \mathbf{g}(\nabla_X Y, \nabla_W Z) \\ &\quad + X[\mathbf{g}(Y, \nabla_W Z)] + \mathbf{g}(\nabla_W Y, \nabla_X Z), \\ \mathbf{g}(Y, [\nabla_W, \nabla_X]Z) &= [W, X][\mathbf{g}(Y, Z)] - W[\mathbf{g}(\nabla_X Y, Z)] - \mathbf{g}(\nabla_W Y, \nabla_X Z) \\ &\quad + X[\mathbf{g}(\nabla_W Y, Z)] + \mathbf{g}(\nabla_X Y, \nabla_W Z). \end{aligned}$$

Consequently,

$$\begin{aligned} &\mathbf{g}([\nabla_W, \nabla_X]Y, Z) + \mathbf{g}(Y, [\nabla_W, \nabla_X]Z) \\ &= 2[W, X][\mathbf{g}(Y, Z)] - W[\mathbf{g}(Y, \nabla_X Z) + \mathbf{g}(Z, \nabla_X Y)] + X[\mathbf{g}(\nabla_W Y, Z) \\ &\quad + \mathbf{g}(Y, \nabla_W Z)] \\ &= 2[W, X][\mathbf{g}(Y, Z)] - W[X[\mathbf{g}(Y, Z)]] + X[W[\mathbf{g}(Y, Z)]] \\ &= [W, X][\mathbf{g}(Y, Z)]. \end{aligned}$$

Similarly, we have $-\mathbf{g}(\nabla_{[W, X]}Y, Z) - \mathbf{g}(Y, \nabla_{[W, X]}Z) = -\llbracket W, X \rrbracket[\mathbf{g}(Y, Z)]$ and

$$\begin{aligned} &-\mathbf{g}(\llbracket \mathcal{Q}([W, X]), Y \rrbracket, Z) - \mathbf{g}(Y, \llbracket \mathcal{Q}([W, X]), Z \rrbracket) \\ &= -\mathcal{Q}([W, X])[\mathbf{g}(Y, Z)] + (\mathcal{L}_{\mathcal{Q}([W, X])}^{\mathcal{P}}\mathbf{g})(Y, Z). \end{aligned}$$

Substituting back into the expression for $2\widehat{C}(W, X, Y, Z)$ yields the result. □

It turns out—at least, when \mathcal{D} is strongly nonholonomic—that the curvature tensor \widehat{C} may be interpreted to measure the “geodesic invariance” of \mathcal{D} , i.e., the invariance (in a specific sense) of \mathcal{D} under the geodesic flow. For further details (both on geodesic invariance and the aforementioned interpretation of \widehat{C}), refer to [2].

Since \widehat{R} satisfies all of the symmetries of the Riemannian (0, 4)-tensor, it has analogous algebraic properties. (Heuristically, we may thus view the Schouten tensor \widehat{K} as consisting of a “Riemannian” component \widehat{R} and a “remainder” \widehat{C} .) Hence we shall define a sectional curvature, a Ricci tensor, and a scalar curvature analogous to the corresponding tensors in Riemannian geometry.

Let $\mathcal{S}_q, q \in \mathbb{M}$ be a two-dimensional subspace of \mathcal{D}_q and let (X_q, Y_q) be a basis for \mathcal{S}_q . We define the *sectional curvature* of \mathcal{S}_q , denoted $\tilde{R}(\mathcal{S}_q)$, as

$$\tilde{R}(\mathcal{S}_q) = \frac{\widehat{R}_q(X_q \wedge Y_q, X_q \wedge Y_q)}{\widehat{\mathbf{g}}_q(X_q \wedge Y_q, X_q \wedge Y_q)}.$$

Here \widehat{R} is viewed as the tensor $(W \wedge X, Y \wedge Z) \mapsto \widehat{R}(W, X, Z, Y)$ and $\widehat{\mathbf{g}}$ is the metric induced on $\bigwedge^2 \mathcal{D}$ by \mathbf{g} , i.e., $\widehat{\mathbf{g}}(W \wedge X, Y \wedge Z) = \mathbf{g}(W, Y)\mathbf{g}(X, Z) - \mathbf{g}(W, Z)\mathbf{g}(X, Y)$. We shall also write $\tilde{R}(\mathcal{S}_q)$ as $\tilde{R}(X_q \wedge Y_q)$. As in the Riemannian case (see, e.g., [21]), one can show that \tilde{R} is well defined (i.e., does not depend on the choice of $X_q \wedge Y_q$). Furthermore, \tilde{R} determines \widehat{R} , in the following sense: suppose $F \in T_4^0(\mathcal{D})$ satisfies the symmetries (SI)–(S4); in particular, F can be viewed as the tensor $F : (W \wedge X, Y \wedge Z) \mapsto F(W, X, Z, Y)$. If

$$\tilde{R}(X_q \wedge Y_q) = \frac{F_q(X_q \wedge Y_q, X_q \wedge Y_q)}{\widehat{\mathbf{g}}_q(X_q \wedge Y_q, X_q \wedge Y_q)}$$

for every nonzero $X_q \wedge Y_q \in \bigwedge^2 \mathcal{D}$, then $\widehat{R} = F$.

The *Ricci tensor* $\text{Ric} \in T_2^0(\mathcal{D})$ is defined as $\text{Ric} = \text{tr}_1^1 R$, i.e.,

$$\text{Ric}(X, Y) = \text{tr}(Z \mapsto R(Z, X)Y) = \sum_a \mathbf{g}(R(X_a, X)Y, X_a) = \sum_a \widehat{R}(X_a, X, Y, X_a),$$

where (X_a) is an orthonormal frame for \mathcal{D} . The Ricci tensor is clearly symmetric. The trace of the endomorphism $\mathbf{g}^\sharp \circ \text{Ric}^\flat : \Gamma(\mathcal{D}) \rightarrow \Gamma(\mathcal{D})$ is called the *scalar curvature*, denoted Scal . (Here S^\flat denotes the map $X \mapsto S(X, \cdot)$, while S^\sharp denotes the map $(S^\flat)^{-1}$ whenever the tensor S is nondegenerate.) In terms of the orthonormal frame (X_a) , we have

$$\text{Scal} = \sum_a \text{Ric}(X_a, X_a) = \sum_{a,b} \widehat{R}(X_b, X_a, X_a, X_b) = \sum_{a \neq b} \tilde{R}(X_a \wedge X_b).$$

In a similar fashion to the Ricci tensor, let $A \in T_2^0(\mathcal{D})$ be defined as $A = \text{tr}_1^1 C$, i.e.,

$$A(X, Y) = \sum_a \widehat{C}(X_a, X, Y, X_a),$$

where (X_a) is an orthonormal frame for \mathcal{D} . In general, A is not symmetric; thus we define two tensors A_{sym} and A_{skew} to be the symmetric and skew-symmetric parts of A , respectively. In terms of (X_a) , we then have

$$A_{\text{sym}}(X, Y) = \frac{1}{2} \sum_a [\widehat{C}(X_a, X, Y, X_a) + \widehat{C}(X_a, Y, X, X_a)],$$

$$A_{\text{skew}}(X, Y) = -\frac{1}{2} \sum_a \widehat{C}(X, Y, X_a, X_a).$$

(Note that $A_{\text{skew}} = -\frac{1}{2} \text{tr}_3^1 C$.) It is not difficult to see that both A_{sym} and A_{skew} are trace-free.

4 The Wagner curvature tensor

We briefly describe Wagner’s approach to the construction of his curvature tensor. Although Wagner originally expressed his construction using the language of the Ricci calculus, there have also been some more modern expositions of his results, most notably [10]. Our approach largely follows the latter paper, presenting Wagner’s ideas in the language of modern differential geometry. (In Sect. 4.1 we discuss in more detail the differences between [10] and this paper.) Let $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ be a nonholonomic Riemannian structure and let $\mathcal{D} = \mathcal{D}^1 \subsetneq \dots \subsetneq \mathcal{D}^N = TM$ be the flag of \mathcal{D} . The nonholonomic connection $\nabla^1 = \nabla$ induces a parallel translation along \mathcal{D}^1 -curves. For each component \mathcal{D}^{i+1} of the flag, one constructs a \mathcal{D}^{i+1} -connection ∇^{i+1} on \mathcal{D} . Such a connection induces a parallel translation along \mathcal{D}^{i+1} -curves; furthermore, ∇^{i+1} is defined in such a way that it extends ∇^i and the set of parallel tensors of ∇^{i+1} coincides with that of ∇^i . Finally, one gets a vector bundle connection ∇^N on \mathcal{D} (whose corresponding parallel translation is along any curve in M), with an associated curvature tensor K^N ; this is the Wagner curvature tensor. The vanishing of K^N characterizes the flatness of ∇^N , and hence (by construction of $\nabla^2, \dots, \nabla^{N-1}$), the flatness of $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$.

4.1 Definition and basic properties

Let $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ be a nonholonomic Riemannian manifold, where \mathcal{D} has degree of nonholonomy N , and let $\mathcal{D}^1 \subsetneq \dots \subsetneq \mathcal{D}^N = TM$ be the flag of \mathcal{D} . In addition, let $\mathcal{E}^1, \dots, \mathcal{E}^{N-1}$ be distributions on M such that

$$\mathcal{D}^\perp = \mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^{N-1} \quad \text{and} \quad \mathcal{D}^{i+1} = \mathcal{D}^i \oplus \mathcal{E}^i \quad (i = 1, \dots, N - 1). \tag{2}$$

Let $\mathcal{Q}_i : TM \rightarrow \mathcal{E}^i$ denote the projection onto \mathcal{E}^i and let $\mathcal{P}_i : TM \rightarrow \mathcal{D}^i$ be the projection onto $\mathcal{D}^i = \mathcal{D} \oplus \mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^{i-1}$ defined as $\mathcal{P}_1 = \mathcal{P}$ and $\mathcal{P}_{i+1} = \mathcal{P} \oplus \mathcal{Q}_1 \oplus \dots \oplus \mathcal{Q}_i$ for $i \geq 1$. The distributions $\mathcal{E}^1, \dots, \mathcal{E}^{N-1}$ play a crucial rôle in the definition of the Wagner curvature tensor, yet in general there is no canonical choice for these distributions. Consequently, the Wagner curvature tensor is not intrinsically defined. (Wagner essentially proposed redefining a nonholonomic Riemannian structure to include $\mathcal{E}^1, \dots, \mathcal{E}^{N-1}$.) For the purposes of this paper, by a *Wagner structure* we shall mean a nonholonomic Riemannian structure $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$, with degree of nonholonomy $N \geq 2$, together with distributions $\mathcal{E}^1, \dots, \mathcal{E}^{N-1}$ such that the Eq. (2) are satisfied. If \mathcal{D} is strongly nonholonomic, then $\mathcal{D}^2 = TM = \mathcal{D} \oplus \mathcal{D}^\perp$, i.e., the choice of \mathcal{E}^1 is canonical. Thus we have the following result.

Proposition 9 *Every nonholonomic Riemannian structure $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ with \mathcal{D} strongly nonholonomic is a Wagner structure.*

(Hence, when \mathcal{D} is strongly nonholonomic, the Wagner curvature tensor will be intrinsic.) Similarly, if \mathbf{g} is the restriction $\tilde{\mathbf{g}}|_{\mathcal{D}}$ of a Riemannian metric $\tilde{\mathbf{g}}$ on M (as in the case, for instance, of a nonholonomic mechanical system with kinetic energy Lagrangian and constraints linear-in-velocities), then there are canonical choices of the distributions $\mathcal{E}^1, \dots, \mathcal{E}^{N-1}$.

Proposition 10 *Let $(M, \tilde{\mathbf{g}})$ be a Riemannian manifold, \mathcal{D} a completely nonholonomic distribution on M and \mathcal{D}^\perp the orthogonal complement of \mathcal{D} . Let $\mathcal{D} = \mathcal{D}^1 \subsetneq \dots \subsetneq \mathcal{D}^N = TM$ be*

the flag of \mathcal{D} and let \mathcal{E}^i be the $\tilde{\mathbf{g}}|_{\mathcal{D}^{i+1}}$ -orthogonal complement of \mathcal{D}^i in \mathcal{D}^{i+1} , $i = 1, \dots, N - 1$. Then $(\mathbf{M}, \mathcal{D}, \mathcal{D}^\perp, \tilde{\mathbf{g}}|_{\mathcal{D}})$, together with the distributions $\mathcal{E}^1, \dots, \mathcal{E}^{N-1}$, is a Wagner structure.

Proof The second part of (2) holds by construction of $\mathcal{E}^1, \dots, \mathcal{E}^{N-1}$. For the first part, since $\mathcal{D}^i = \mathcal{D} \oplus \mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^{i-1}$ and $\mathcal{D}^i \perp_{\tilde{\mathbf{g}}} \mathcal{E}^i$, we have $\mathcal{D} \perp_{\tilde{\mathbf{g}}} \mathcal{E}^i$. That is, \mathcal{D} is orthogonal to each of $\mathcal{E}^1, \dots, \mathcal{E}^{N-1}$, and hence is orthogonal to $\mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^{N-1}$. Then $\mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^{N-1}$ is the orthogonal complement of \mathcal{D} , whence $\mathcal{D}^\perp = \mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^{N-1}$. \square

Let $\Lambda_i : \Gamma(\bigwedge^2 \mathcal{D}^i) \rightarrow \Gamma(\mathcal{E}^i)$, $i = 1, \dots, N - 1$ be the tensors given by $\Lambda_i(X \wedge Y) = \mathcal{Q}_i([X, Y])$ for $X, Y \in \Gamma(\mathcal{D}^i)$. Since \mathcal{D} is completely nonholonomic, each map Λ_i is surjective; hence (for a Wagner structure) one can canonically extend \mathbf{g} to a Riemannian metric.

Theorem 1 (cf. [10]) *There exists a unique Riemannian metric $\tilde{\mathbf{g}}$ on \mathbf{M} satisfying the following conditions:*

- (i) *The decomposition $TM = \mathcal{D} \oplus \mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^{N-1}$ is orthogonal and $\tilde{\mathbf{g}} = \mathbf{g} \oplus \mathbf{h}^1 \oplus \dots \oplus \mathbf{h}^{N-1}$, where $\mathbf{h}^i = \tilde{\mathbf{g}}|_{\mathcal{E}^i}$, $i = 1, \dots, N - 1$.*
- (ii) *Each map $\Lambda_i|_{(\ker \Lambda_i)^\perp} : (\ker \Lambda_i)^\perp \rightarrow \mathcal{E}^i$, $i = 1, \dots, N - 1$ satisfies*

$$\mathbf{h}^i(\Lambda_i(W \wedge X), \Lambda_i(Y \wedge Z)) = \hat{\mathbf{g}}^i(W \wedge X, Y \wedge Z)$$

for $W \wedge X, Y \wedge Z \in (\ker \Lambda_i)^\perp$, where $\hat{\mathbf{g}}^i$ is the metric induced on $\bigwedge^2 \mathcal{D}^i$ by the metric $\mathbf{g}^i = \mathbf{g} \oplus \mathbf{h}^1 \oplus \dots \oplus \mathbf{h}^{i-1}$ on \mathcal{D}^i , i.e., $\hat{\mathbf{g}}^i(W \wedge X, Y \wedge Z) = \mathbf{g}^i(W, Y)\mathbf{g}^i(X, Z) - \mathbf{g}^i(W, Z)\mathbf{g}^i(X, Y)$.

Proof Let $\mathbf{g}^1 = \mathbf{g}$ and let $\hat{\mathbf{g}}^1$ be the corresponding metric on $\bigwedge^2 \mathcal{D}^1$. Let $(\ker \Delta_1)^\perp$ be the orthogonal complement of $\ker \Delta_1 \subseteq \bigwedge^2 \mathcal{D}^1$ with respect to $\hat{\mathbf{g}}^1$. As Δ_1 is surjective, we can define a metric \mathbf{h}^1 on \mathcal{E}^1 by the requirement that the isomorphism $\Delta_1|_{(\ker \Delta_1)^\perp} : (\ker \Delta_1)^\perp \rightarrow \mathcal{E}^1$ is an isometry. Hence we have the metric $\mathbf{g}^2 = \mathbf{g}^1 \oplus \mathbf{h}^1$ on $\mathcal{D}^2 = \mathcal{D} \oplus \mathcal{E}^1$, which induces a metric $\hat{\mathbf{g}}^2$ on $\bigwedge^2 \mathcal{D}^2$. Let \mathbf{h}^2 be the metric on \mathcal{E}^2 induced by $\Delta_2|_{(\ker \Delta_2)^\perp}$. Continuing in this fashion, we get the Riemannian metric $\tilde{\mathbf{g}} = \mathbf{g} \oplus \mathbf{h}^1 \oplus \dots \oplus \mathbf{h}^{N-1}$ on $TM = \mathcal{D} \oplus \mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^{N-1}$. \square

Fix $1 \leq i \leq N - 1$. Let $\nabla^1 = \nabla$ and let $\nabla^{i+1} : \Gamma(\mathcal{D}^{i+1}) \times \Gamma(\mathcal{D}) \rightarrow \Gamma(\mathcal{D})$ be the \mathcal{D}^{i+1} -connection on \mathcal{D} specified as follows: if $Z \in \Gamma(\mathcal{D}^{i+1})$ with $X = \mathcal{P}_i(Z)$ and $A = \mathcal{Q}_i(Z)$, then

$$\nabla_Z^{i+1} U = \nabla_X^i U + K^i(\Theta_i(A))U + \llbracket A, U \rrbracket.$$

Here $\Theta_i = \Lambda_i|_{(\ker \Lambda_i)^\perp}^{-1}$ and $K^i : \Gamma(\bigwedge^2 \mathcal{D}^i) \times \Gamma(\mathcal{D}) \rightarrow \Gamma(\mathcal{D})$ is the curvature tensor of ∇^i defined as

$$K^i(X \wedge Y)U = [\nabla_X^i, \nabla_Y^i]U - \nabla_{\mathcal{P}_i([X, Y])}^i U - \llbracket \mathcal{Q}_i([X, Y]), U \rrbracket.$$

(Note that, as $\mathcal{Q}([\mathcal{D}, \mathcal{D}]) = \mathcal{Q}_1([\mathcal{D}, \mathcal{D}])$, we have $K^1 = K$.) In particular, for $i = N - 1$, we have a vector bundle connection $\nabla^N : \Gamma(TM) \times \Gamma(\mathcal{D}) \rightarrow \Gamma(\mathcal{D})$ on \mathcal{D} . Let K^N be the curvature tensor of this connection:

$$K^N(X \wedge Y)U = [\nabla_X^N, \nabla_Y^N]U - \nabla_{[X,Y]}^N U.$$

K^N is called the *Wagner curvature tensor* of $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$.

Theorem 2 (cf. [10]) *Given $T \in T_\rho^k(\mathcal{D})$, we have $\nabla^i T \equiv 0$ if and only if $\nabla^{i+1} T \equiv 0$.*

Proof Let $T \in T_\rho^k(\mathcal{D})$; we have

$$\nabla_Z^{i+1} T = \nabla_X^i T + K^i(\Theta_i(A))T + \mathcal{L}_A^\rho T, \quad Z = X + A, \quad X \in \Gamma(\mathcal{D}^i), \quad A \in \Gamma(\mathcal{E}^i),$$

where $K^i(\Theta_i(A))$ is viewed as a derivation in $\text{Der}_0(\mathcal{D})$. Suppose that $\nabla^i T \equiv 0$; then

$$\begin{aligned} K^i(X \wedge Y)T &= \nabla_X^i \nabla_Y^i T - \nabla_Y^i \nabla_X^i T - \nabla_{\mathcal{P}_i([X,Y])}^i T - \mathcal{L}_{\mathcal{Q}_i([X,Y])}^\rho T \\ &= -\mathcal{L}_{\mathcal{Q}_i([X,Y])}^\rho T \end{aligned}$$

for every $X, Y \in \Gamma(\mathcal{D}^i)$. We claim that $K^i(\Theta_i(A))T + \mathcal{L}_A^\rho T = 0$ for $A \in \Gamma(\mathcal{E}^i)$. Indeed, suppose that $\Theta_i(A) = X \wedge Y$ for $X, Y \in \Gamma(\mathcal{D}^i)$; then

$$\begin{aligned} K^i(\Theta_i(A))T &= K^i(X \wedge Y)T = [\nabla_X^i, \nabla_Y^i]T - \nabla_{\mathcal{P}_i([X,Y])}^i T - \mathcal{L}_{\mathcal{Q}_i([X,Y])}^\rho T \\ &= -\mathcal{L}_{\mathcal{A}_i(X \wedge Y)}^\rho T = -\mathcal{L}_{\mathcal{A}_i(\Theta_i(A))}^\rho T = -\mathcal{L}_A^\rho T. \end{aligned}$$

The case when $\Theta_i(A)$ is a $C^\infty(M)$ -combination of bivector fields follows from the tensoriality of K^i . Hence, if $Z \in \Gamma(\mathcal{D}^{i+1})$ with $X = \mathcal{P}_i(Z)$ and $A = \mathcal{Q}_i(Z)$, then

$$\nabla_Z^{i+1} T = \nabla_X^i T + K^i(\Theta_i(A))T + \mathcal{L}_A^\rho T = 0.$$

Conversely, suppose $\nabla^{i+1} T \equiv 0$; then $\nabla_X^i T = \nabla_X^{i+1} T = 0$ for $X \in \Gamma(\mathcal{D}^i)$, whence $\nabla^i T \equiv 0$. □

Corollary 1 *∇^i is metric, i.e., $\nabla^i \mathbf{g} \equiv 0$ ($i = 1, \dots, N$).*

We have N restricted connections $\nabla^1, \dots, \nabla^N$ on \mathcal{D} . The first connection (a nonholonomic, or \mathcal{D} -connection on \mathcal{D}) permits parallel translation only along \mathcal{D} -curves, whereas the last (a vector bundle connection on \mathcal{D}) permits parallel translation along any curve in M . In between we have the \mathcal{D}^i -restricted connections ∇^i (for each $i = 2, \dots, N - 1$) which permit parallel translation along \mathcal{D}^i -curves. By Corollary 1, parallel translation (with respect to any of the connections $\nabla^1, \dots, \nabla^N$) is a linear isometry. For a \mathcal{D}^i -curve $\gamma : [0, 1] \rightarrow M$, let $\Pi_\gamma^{i,t}$ denote the parallel translation along γ with respect to ∇^i .

Proposition 11 *If $\gamma : [0, 1] \rightarrow M$ is a \mathcal{D}^i -curve, then $\Pi_\gamma^{i,t} = \Pi_\gamma^{i+1,t}$ ($i = 1, \dots, N - 1$).*

Proof Let $\gamma : [0, 1] \rightarrow M$ be a \mathcal{D}^i -curve, $U_0 \in \mathcal{D}_{\gamma(0)}$ and $V(t) = \Pi_\gamma^{i,t}(U_0)$, $W(t) = \Pi_\gamma^{i+1,t}(U_0)$. Let $(X_{a_0}^0)$ be an orthonormal frame for \mathcal{D} and $(X_{a_i}^i)$ a frame for \mathcal{E}^i , where $1 \leq a_0 \leq r$ and $1 \leq a_i \leq \text{rank}(\mathcal{E}^i)$. It follows that $(X_{a_0}^0, X_{a_1}^1, \dots, X_{a_i}^i)$ is a frame for $\mathcal{D}^{i+1} = \mathcal{D} \oplus \mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^i$. There exist functions $v^{a_0}, w^{a_0} \in C^\infty([0, 1])$ such that $V = v^{a_0}(X_{a_0}^0 \circ \gamma)$ and $W = w^{a_0}(X_{a_0}^0 \circ \gamma)$. Furthermore, these functions satisfy the ODEs

$$\dot{v}^{a_0} = -\Gamma_{b_i c_0}^{a_0}(\gamma)\dot{\gamma}^{b_i} v^{c_0} \quad \text{and} \quad \dot{w}^{a_0} = -\Omega_{b_i c_0}^{a_0}(\gamma)\dot{\gamma}^{b_i} w^{c_0},$$

where $\Gamma_{b_i c_0}^{a_0}, \Omega_{b_i c_0}^{a_0} \in C^\infty(\mathbf{M})$ are defined by $\nabla_{X_j}^i X_{c_0}^0 = \Gamma_{b_i c_0}^{a_0} X_{a_0}^0$ (where $0 \leq j \leq i$) and $\nabla_{X_k}^{i+1} X_{c_0}^0 = \Omega_{b_k c_0}^{a_0} X_{a_0}^0$ (where $0 \leq k \leq i + 1$). Since $\nabla_X^i U = \nabla_X^{i+1} U$ for $X \in \Gamma(\mathcal{D}^i)$ and $U \in \Gamma(\mathcal{D})$, we have $\Gamma_{b_j c_0}^{a_0} = \Omega_{b_j c_0}^{a_0}$ for $j = 0, \dots, i$. Hence $V = W$, and so $\Pi_\gamma^{i,t} = \Pi_\gamma^{i+1,t}$. \square

Lastly, we state the main result of this section: the vanishing of the Wagner curvature tensor characterizes the flat Wagner structures.

Theorem 3 (cf. [10]) *A Wagner structure $(\mathbf{M}, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ is locally flat if and only if $K^N \equiv 0$.*

Proof The Wagner curvature tensor K^N is precisely the curvature tensor of the vector bundle connection ∇^N , and so there exists a parallel frame for \mathcal{D} on an open set in \mathbf{M} if and only if K^N vanishes identically on that set. Since every parallel vector field with respect to ∇^N is also parallel with respect to ∇ , it follows that $(\mathbf{M}, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ is locally flat exactly when K^N vanishes identically. \square

Remark 3 As for the Schouten tensor, one can use \mathbf{g} to lower an index of K^i , obtaining the tensor \hat{K}^i given by $\hat{K}^i(X, Y, U, V) = \mathbf{g}(\hat{K}^i(X \wedge Y)U, V)$ for $X, Y \in \Gamma(\mathcal{D}^i)$ and $U, V \in \Gamma(\mathcal{D})$. Each tensor \hat{K}^i then decomposes into two tensors \hat{R}^i and \hat{C}^i :

$$\hat{R}^i(X, Y, U, V) = \frac{1}{2}[\hat{K}^i(X, Y, U, V) - \hat{K}^i(X, Y, V, U)], \quad \hat{C}^i = \hat{K}^i - \hat{R}^i.$$

We then have the following expression for \hat{C}^i (cf. Proposition 8):

$$\hat{C}^i(X, Y, U, V) = \frac{1}{2}(\mathcal{L}_{\mathcal{D}_i(\{X, Y\})} \mathbf{g})(U, V).$$

Remark 4 Just as \hat{C} may be interpreted in terms of geodesic invariance of \mathcal{D} (in the strongly nonholonomic case), so too may the tensors \hat{C}^i be interpreted in terms of the geodesic invariance of the distributions \mathcal{D}^i (in the case of a Wagner structure) [2].

Remark 5 Regarding the construction of Wagner’s curvature tensor, we have departed slightly from the presentation in [10]; we have also filled in a number of missing details. For instance, we have made explicit the dependence of Wagner’s construction on the complementary distributions \mathcal{E}^i (and show that a strongly nonholonomic structure, and a structure arising from a nonholonomic mechanical system, always satisfies the additional requirements in order to define Wagner’s tensor). We have also presented the construction in a more direct fashion (e.g., skipping the quotients of distributions in favour of using the projection mappings directly). This serves to simplify the presentation, and makes the construction more natural (particularly of the induced Riemannian metric in Theorem 1). Furthermore, the definition of the connections $\nabla^2, \dots, \nabla^N$ in this paper is different from that in [10] (also in [30]). In [10], ∇^{i+1} is defined as

$$\nabla_Z^{i+1} U = \nabla_X^i U + K^i(\Lambda_i^\dagger(A))U + \llbracket A, U \rrbracket, \quad Z = X + A \in \Gamma(\mathcal{D}^{i+1}), \quad U \in \Gamma(\mathcal{D}),$$

where Λ_i^\dagger is the adjoint of Λ_i , i.e., $\Lambda_i^\dagger = (\hat{\mathbf{g}}^i)^\# \circ \Lambda_i^* \circ (\mathbf{h}^i)^\flat$. In fact, it turns out that the Λ_i^\dagger can be replaced with *any* right inverse of Λ_i , and the crucial property of the ∇^i ’s (viz., Theorem 2)

will still hold. The approach we employ avoids a particular choice of right inverse to Λ_i , and simplifies the construction. Nevertheless, this means that the definition of Wagner’s tensor K^N in this paper differs from that in [10, 30]. (See also [5, 14], who also present another alternative approach to constructing Wagner’s tensor. In particular, the connections ∇^{i+1} are instead defined as mappings $\Gamma(\mathcal{D}^{i+1}) \times \Gamma(\mathcal{D}^i) \rightarrow \Gamma(\mathcal{D}^i)$, and the tensors K^{i+1} as $\Gamma(\wedge^2 \mathcal{D}^{i+1}) \times \Gamma(\mathcal{D}^i) \rightarrow \Gamma(\mathcal{D}^i)$.) A key result of this section is that ∇^{i+1} “extends” ∇^i (Theorem 2). Hence, regarding the existence of a parallel vector field $U \in \Gamma(\mathcal{D})$, rather than consider the equation $\nabla U = 0$, we can consider the simpler, equivalent equation $\nabla^N U = 0$. In [10] one starts from the equation $\nabla U = \varphi$, where $\varphi \in \mathcal{T}_1^1(\mathcal{D})$, and then one takes $\varphi = 0$ as required. On the other hand, Theorem 2 in fact allows one to consider the existence of a parallel tensor field, i.e., the equation $\nabla T = 0$, where $T \in \mathcal{T}_\ell^k(\mathcal{D})$ (a consequence of this is that each connection ∇^i is metric). We also consider the relation between the parallel transport of ∇^{i+1} and ∇^i , which was not addressed in [10].

4.2 Algebraic interpretation of curvature tensors

For a vector bundle connection $\tilde{\nabla}$ on \mathcal{D} , the curvature tensor $(X, Y) \mapsto [\tilde{\nabla}_X, \tilde{\nabla}_Y] - \tilde{\nabla}_{[X, Y]}$ can be viewed as measuring the extent to which the mapping $\Gamma(TM) \rightarrow \text{Der}(\mathcal{D})$, $X \mapsto \tilde{\nabla}_X$ fails to be a homomorphism (of Lie algebras). A similar interpretation holds for the curvature tensors K^i .

Since the tangent bundle of M decomposes as $TM = \mathcal{D} \oplus \mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^{N-1}$, we have a corresponding decomposition $\mathcal{L}_{TM}^{\mathcal{P}} = \mathcal{L}_{\mathcal{D}}^{\mathcal{P}} \oplus \mathcal{L}_{\mathcal{E}^1}^{\mathcal{P}} \oplus \dots \oplus \mathcal{L}_{\mathcal{E}^{N-1}}^{\mathcal{P}}$. Moreover, as $\mathcal{D}^{i+1} = \mathcal{D}^i \oplus \mathcal{E}^i = \mathcal{D} \oplus \mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^i$, we get

$$\begin{aligned} \text{Der}_{\mathcal{D}^{i+1}}(\mathcal{D}) &= \mathcal{L}_{\mathcal{D}^{i+1}}^{\mathcal{P}} \oplus \text{Der}_0(\mathcal{D}) \\ &= \mathcal{L}_{\mathcal{D}^i}^{\mathcal{P}} \oplus \mathcal{L}_{\mathcal{E}^i}^{\mathcal{P}} \oplus \text{Der}_0(\mathcal{D}) \\ &= \mathcal{L}_{\mathcal{D}}^{\mathcal{P}} \oplus \mathcal{L}_{\mathcal{E}^1}^{\mathcal{P}} \oplus \dots \oplus \mathcal{L}_{\mathcal{E}^i}^{\mathcal{P}} \oplus \text{Der}_0(\mathcal{D}). \end{aligned} \tag{48}$$

(Consequently, $\text{Der}_{\mathcal{D}}(\mathcal{D})$ is completely nonholonomic: if we define the flag $\mathcal{S}^1 \subsetneq \mathcal{S}^2 \subsetneq \dots$ by $\mathcal{S}^1 = \text{Der}_{\mathcal{D}}(\mathcal{D})$ and $\mathcal{S}^{i+1} = \mathcal{S}^i + [\mathcal{S}^i, \mathcal{S}^i]$, $i \geq 1$, then $\mathcal{S}^i = \text{Der}_{\mathcal{D}^i}(\mathcal{D})$ and $\mathcal{S}^N = \text{Der}(\mathcal{D})$.) We shall also use \mathcal{P}_i to denote the projection $\text{Der}(\mathcal{D}) \rightarrow \text{Der}_{\mathcal{D}^i}(\mathcal{D})$; likewise, let \mathcal{Q}_i be the projection $\text{Der}(\mathcal{D}) \rightarrow \mathcal{L}_{\mathcal{E}^i}^{\mathcal{P}}$.

Lemma 4 Fix $1 \leq i \leq N$ and let $\delta_1, \delta_2 \in \text{Der}_{\mathcal{D}^i}(\mathcal{D})$, where $\delta_1 = \mathcal{L}_X^{\mathcal{P}} + \delta'_1$ and $\delta_2 = \mathcal{L}_Y^{\mathcal{P}} + \delta'_2$ for $X, Y \in \Gamma(\mathcal{D}^i)$ and $\delta'_1, \delta'_2 \in \text{Der}_0(\mathcal{D})$. Then

$$\mathcal{P}_i([\delta_1, \delta_2]) = [\delta_1, \delta_2] - \mathcal{L}_{\mathcal{Q}_i([X, Y])}^{\mathcal{P}}.$$

Proof We have $[\delta_1, \delta_2] = [\mathcal{L}_X^{\mathcal{P}}, \mathcal{L}_Y^{\mathcal{P}}] + [\mathcal{L}_X^{\mathcal{P}}, \delta'_2] + [\delta'_1, \mathcal{L}_Y^{\mathcal{P}}] + [\delta'_1, \delta'_2]$. Moreover, since $[\text{Der}(\mathcal{D}), \text{Der}_0(\mathcal{D})] \subseteq \text{Der}_0(\mathcal{D})$, it follows that $[\mathcal{L}_X^{\mathcal{P}}, \delta'_2], [\delta'_1, \mathcal{L}_Y^{\mathcal{P}}], [\delta'_1, \delta'_2] \in \text{Der}_0(\mathcal{D})$. Consider the term $[\mathcal{L}_X^{\mathcal{P}}, \mathcal{L}_Y^{\mathcal{P}}]$; we have $[\mathcal{L}_X^{\mathcal{P}}, \mathcal{L}_Y^{\mathcal{P}}](f) = [X, Y][f]$ for every $f \in C^\infty(M)$, and so

$$[\mathcal{L}_X^{\mathcal{P}}, \mathcal{L}_Y^{\mathcal{P}}] = \mathcal{L}_{[X, Y]}^{\mathcal{P}} + \delta'$$

for some $\delta' \in \text{Der}_0(\mathcal{D})$. Moreover, as $X, Y \in \Gamma(\mathcal{D}^i)$, we have $[X, Y] \in \Gamma(\mathcal{D}^{i+1})$. Thus $[\mathcal{L}_X^{\mathcal{P}}, \mathcal{L}_Y^{\mathcal{P}}] = \mathcal{L}_{\mathcal{P}_i([X, Y])}^{\mathcal{P}} + \mathcal{L}_{\mathcal{Q}_i([X, Y])}^{\mathcal{P}} + \delta'$, and so

$$\mathcal{Q}_i([\mathcal{L}_X^{\mathcal{P}}, \mathcal{L}_Y^{\mathcal{P}}]) = [\mathcal{L}_X^{\mathcal{P}}, \mathcal{L}_Y^{\mathcal{P}}] - \mathcal{L}_{\mathcal{Q}_i([X, Y])}^{\mathcal{P}}.$$

Consequently, we have

$$\mathcal{P}_i([\delta_1, \delta_2]) = [\delta_1, \delta_2] - \mathcal{Q}_i([\delta_1, \delta_2]) = [\delta_1, \delta_2] - \mathcal{Q}_i([\mathcal{L}_X^{\mathcal{P}}, \mathcal{L}_Y^{\mathcal{P}}]) = [\delta_1, \delta_2] - \mathcal{L}_{\mathcal{Q}_i([X, Y])}^{\mathcal{P}}.$$

□

Clearly, ∇_X^i for $X \in \Gamma(\mathcal{D}^i)$ is a \mathcal{D}^i -derivation. The following result asserts that the curvature tensor K^i measures the extent to which $\Gamma(\mathcal{D}^i) \rightarrow \text{Der}_{\mathcal{D}^i}(\mathcal{D})$, $X \mapsto \nabla_X^i$ fails to be a homomorphism from $(\Gamma(\mathcal{D}^i), \mathcal{P}_i([\cdot, \cdot]))$ to $(\text{Der}_{\mathcal{D}^i}(\mathcal{D}), \mathcal{P}_i([\cdot, \cdot]))$. (Note that these structures are not Lie algebras, as $\mathcal{P}_i([\cdot, \cdot])$ does not generally satisfy the Jacobi identity. Instead they are so-called *almost Lie structures* [24], and so “homomorphism” refers to a homomorphism of almost Lie structures.)

Theorem 4 *We have*

$$K^i(X \wedge Y) = \mathcal{P}_i([\nabla_X^i, \nabla_Y^i]) - \nabla_{\mathcal{P}_i([X, Y])}^i, \quad X, Y \in \Gamma(\mathcal{D}^i) \quad (i = 1, \dots, N).$$

Proof Let $X, Y \in \Gamma(\mathcal{D}^i)$ and $U \in \Gamma(\mathcal{D})$. There exist derivations $\delta_1, \delta_2 \in \text{Der}_0(\mathcal{D})$ such that $\nabla_X^i = \mathcal{L}_X^{\mathcal{P}} + \delta_1$ and $\nabla_Y^i = \mathcal{L}_Y^{\mathcal{P}} + \delta_2$. Hence, by Lemma 4, we have $\mathcal{P}_i([\nabla_X^i, \nabla_Y^i]) = [\nabla_X^i, \nabla_Y^i] - \mathcal{L}_{\mathcal{P}_i([X, Y])}^{\mathcal{P}}$. It follows that

$$\begin{aligned} &\mathcal{P}_i([\nabla_X^i, \nabla_Y^i])U - \nabla_{\mathcal{P}_i([X, Y])}^i U \\ &= [\nabla_X^i, \nabla_Y^i]U - \nabla_{\mathcal{P}_i([X, Y])}^i U - \llbracket \mathcal{Q}_i([X, Y]), U \rrbracket = K^i(X \wedge Y)U. \end{aligned}$$

□

4.3 Three-dimensional nonholonomic Riemannian structures

Flat nonholonomic Riemannian structures in three dimensions were considered in [3]; in particular, a characterization of flatness (in three dimensions) was given. This characterization (in fact, characterizing equation) was obtained in a rather direct way by using an intrinsic contact form, as well as the \mathcal{P} -exterior covariant derivative and contractions of the Schouten curvature tensor. (Moreover, it turns out that the vanishing of the Schouten tensor is sufficient for flatness in three dimensions.) We shall relate the foregoing characterization with the Wagner curvature tensor.

Let $(M, \mathcal{D}, \mathcal{D}^\perp, \mathfrak{g})$ be a nonholonomic Riemannian structure, where M is three-dimensional (and \mathcal{D} is a rank two strongly nonholonomic distribution on M). There exists a 1-form ω on M (a contact form) such that $\mathcal{D} = \ker \omega$; this 1-form is fixed up to sign by imposing the condition $|d\omega(X_1, X_2)| = 1$, where (X_1, X_2) is an orthonormal frame for \mathcal{D} . The connection $\nabla^2 : \Gamma(TM) \times \Gamma(\mathcal{D}) \rightarrow \Gamma(\mathcal{D})$ is given by

$$\nabla_Z^2 U = \nabla_X U + K(\Theta(A))U + \llbracket A, U \rrbracket$$

for $Z \in \Gamma(TM)$, where $X = \mathcal{P}(Z)$ and $A = \mathcal{Q}(Z)$. Here $\Theta = \Lambda|_{(\ker \Lambda)^\perp}^{-1}$, where $\Lambda : \Gamma(\bigwedge^2 \mathcal{D}) \rightarrow \Gamma(\mathcal{D}^\perp)$, $X \wedge Y \mapsto \mathcal{Q}([X, Y])$; in fact, we have $\ker \Lambda = \{0\}$. For convenience, we shall denote $\tilde{\nabla} = \nabla^2$. The Wagner curvature tensor $\tilde{K} = K^2$ is given by

$$\widetilde{K}(X \wedge Y)U = [\widetilde{\nabla}_X, \widetilde{\nabla}_Y]U - \widetilde{\nabla}_{[X,Y]}U$$

for $X, Y \in \Gamma(TM)$ and $U \in \Gamma(\mathcal{D})$. Define a tensor $\widetilde{\text{Ric}} : \Gamma(TM) \times \Gamma(\mathcal{D}) \rightarrow \mathcal{C}^\infty(\mathcal{M})$ as follows:

$$\widetilde{\text{Ric}}(X, U) = \mathbf{g}(\widetilde{K}(X_1 \wedge X)U, X_1) + \mathbf{g}(\widetilde{K}(X_2 \wedge X)U, X_2),$$

where $X \in \Gamma(TM)$, $U \in \Gamma(\mathcal{D})$ and (X_1, X_2) is an orthonormal frame for \mathcal{D} . (Note that $\widetilde{\text{Ric}}$ does not depend on the choice of X_1 and X_2 .)

Theorem 5 *Let $\mathcal{U} \subseteq \mathcal{M}$ be open. $(\mathcal{M}, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ is locally flat on \mathcal{U} if and only if $\widetilde{\text{Ric}}$ vanishes identically on \mathcal{U} .*

Proof Let (X_0, X_1, X_2) be a frame on \mathcal{U} such that X_0 is a frame for \mathcal{D}^\perp and (X_1, X_2) is an orthonormal frame for \mathcal{D} . We have $[X_i, X_j] = c_{ij}^k X_k$ for structure constants $c_{ij}^k \in \mathcal{C}^\infty(\mathcal{U})$ (where i, j and k take the values 0, 1, 2); moreover, we may assume that $c_{21}^0 = 1$. Let $f_{01}, f_{02} \in \mathcal{C}^\infty(\mathcal{U})$ be defined as follows:

$$\left\{ \begin{aligned} f_{01} &= (c_{10}^1 - c_{20}^2)c_{21}^1 + (c_{10}^2 + c_{20}^1)c_{21}^2 + c_{20}^0 c_{10}^1 - \frac{1}{2}c_{10}^0(c_{21}^2 + c_{20}^1) + c_{10}^0 \kappa \\ &\quad + \frac{1}{2}X_1[c_{21}^2 + c_{20}^1] - X_1[\kappa] - X_2[c_{10}^1] \\ f_{02} &= (c_{10}^2 + c_{20}^1)c_{21}^1 - (c_{10}^1 - c_{20}^2)c_{21}^2 - c_{20}^2 c_{10}^0 + \frac{1}{2}c_{20}^0(c_{21}^2 + c_{20}^1) + c_{20}^0 \kappa \\ &\quad - \frac{1}{2}X_2[c_{21}^2 + c_{20}^1] - X_2[\kappa] + X_1[c_{20}^1]. \end{aligned} \right.$$

Here $\kappa = \frac{1}{2}\text{Scal} = \frac{1}{2}(c_{10}^2 - c_{20}^1) - (c_{21}^1)^2 - (c_{21}^2)^2 - X_1[c_{21}^2] + X_2[c_{21}^1]$. The expressions for f_{01} and f_{02} can be found in [3], where it is shown that $(\mathcal{M}, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ is flat on \mathcal{U} if and only if $f_{01} = f_{02} = 0$. It turns out that f_{01} and f_{02} are, up to sign, the only components of \widetilde{K} . Indeed, a straightforward (but tedious) calculation yields

$$\begin{aligned} \widetilde{K}(X_0 \wedge X_1)X_1 &= f_{01}X_2 & \widetilde{K}(X_0 \wedge X_1)X_2 &= -f_{01}X_1 \\ \widetilde{K}(X_0 \wedge X_2)X_1 &= f_{02}X_2 & \widetilde{K}(X_0 \wedge X_2)X_2 &= -f_{02}X_1 \\ \widetilde{K}(X_1 \wedge X_2)X_1 &= 0 & \widetilde{K}(X_1 \wedge X_2)X_2 &= 0. \end{aligned}$$

Accordingly, we have $\widetilde{\text{Ric}}(X_0, X_1) = \mathbf{g}(-f_{01}X_2, X_1) + \mathbf{g}(-f_{02}X_2, X_2) = -f_{02}$. A similar calculation gives $\widetilde{\text{Ric}}(X_0, X_2) = f_{01}$, and so

$$(\widetilde{\text{Ric}} \circ \Lambda)^b(X_2 \wedge X_1) = \widetilde{\text{Ric}}^b(X_0) = f_{02}v^1 - f_{01}v^2,$$

where (v^0, v^1, v^2) is the coframe dual to (X_0, X_1, X_2) . (Notice that $\widetilde{\text{Ric}} \circ \Lambda$ is completely determined by its evaluation on $X_2 \wedge X_1$.) Lastly, if $X, Y \in \Gamma(\mathcal{D})$, then $\widetilde{\text{Ric}}(X, Y) = 0$. It follows that $\widetilde{\text{Ric}}$ vanishes if and only if $\widetilde{\text{Ric}} \circ \Lambda$ vanishes; the result follows. \square

Remark 6 We know that $(\mathcal{M}, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ is flat if and only if $d_{\mathcal{D}}^\nabla F = F \circ \rho$ [3], where

$$F = \mathbf{g}^\# \circ (\text{tr}_1^1 K)^b = \mathbf{g}^\# \circ (\text{Ric}^b + A_{sym}^b + A_{skew}^b)$$

and $\rho(X_1 \wedge X_2) = d\omega(X_1, X_2)\mathcal{P}(Z)$. Here (X_1, X_2) is an orthonormal frame for \mathcal{D} and $Z \in \Gamma(TM)$ is the Reeb vector field of ω (i.e., the unique—up to sign—vector field on M such that $i_Z\omega = 1$ and $i_Zd\omega = 0$). It turns out that $\mathfrak{g}^\# \circ \text{Ric} \circ \Lambda = d_{\mathcal{P}}^\nabla F - F \circ \rho$.

5 An alternative approach

The nonholonomic connection, as well as the connections $\nabla^2, \dots, \nabla^N$ in the case of a Wagner structure, can be equivalently viewed as horizontal lifts/distributions. Accordingly, we shall consider curvature from this alternative perspective. In particular, we shall express the Schouten curvature tensor K , as well as the curvature tensors K^1, \dots, K^N , in terms of horizontal lifts of vector fields. (As a corollary, we then characterize the vanishing of the curvature tensors in terms of involutivity conditions for the associated horizontal distributions.) We shall also show that the connections $\nabla^1, \dots, \nabla^N$ are equivalently formulated as a flag of horizontal distributions on \mathcal{D} .

5.1 The Schouten tensor

Let $(M, \mathcal{D}, \mathcal{D}^\perp, \mathfrak{g})$ be a nonholonomic Riemannian manifold with associated nonholonomic connection ∇ . We can extend ∇ to a vector bundle connection $\mathring{\nabla}$ on \mathcal{D} as follows (cf. [4]):

$$\mathring{\nabla} : \Gamma(TM) \times \Gamma(\mathcal{D}) \rightarrow \Gamma(\mathcal{D}), \quad \mathring{\nabla}_Z = \nabla_{\mathcal{P}(Z)} + \mathcal{L}_{\mathcal{D}(Z)}^\mathcal{P}$$

(Note that $\mathring{\nabla}$ depends only on $\mathcal{D}, \mathcal{D}^\perp$ and \mathfrak{g} , hence it is intrinsic to the nonholonomic Riemannian structure.) The curvature tensor of $\mathring{\nabla}$ is the $(1, 3)$ -tensor field $\mathring{R} : \Gamma(TM) \times \Gamma(TM) \times \Gamma(\mathcal{D}) \rightarrow \Gamma(\mathcal{D})$ given by $\mathring{R}(X, Y)U = [\mathring{\nabla}_X, \mathring{\nabla}_Y]U - \mathring{\nabla}_{[X, Y]}U$. Clearly, we have $\mathring{R}(X, Y) = K(X, Y)$ whenever $X, Y \in \Gamma(\mathcal{D})$.

Proposition 12 *Let $\mathcal{U} \subseteq M$ be open. If $\mathring{R} \equiv 0$ on \mathcal{U} , then $(M, \mathcal{D}, \mathcal{D}^\perp, \mathfrak{g})$ is locally flat on \mathcal{U} . Conversely, if (U_a) is a parallel frame for \mathcal{D} defined on \mathcal{U} such that $[U_a, \Gamma(\mathcal{D}^\perp)] \subseteq \Gamma(\mathcal{D}^\perp)$, then $\mathring{R} \equiv 0$ on \mathcal{U} .*

Proof Suppose $\mathring{R} \equiv 0$ on \mathcal{U} . Since $\mathring{\nabla}$ is a vector bundle connection on \mathcal{D} , the vanishing of its curvature tensor \mathring{R} implies the existence of a parallel frame (U_a) for \mathcal{D} on \mathcal{U} . (Here “parallel” means “parallel with respect to $\mathring{\nabla}$.”) That is, $\mathring{\nabla}_Z U_a = 0$ for every $Z \in \Gamma_{\mathcal{U}}(TM)$. In particular, taking $Z \in \Gamma_{\mathcal{U}}(\mathcal{D})$, it follows that $\nabla U_a \equiv 0$, and so $(M, \mathcal{D}, \mathcal{D}^\perp, \mathfrak{g})$ is locally flat on \mathcal{U} .

Conversely, suppose there exists a parallel (with respect to ∇) frame (U_a) for \mathcal{D} defined on \mathcal{U} ; then $\mathring{\nabla}_X U_a = 0$ for every $X \in \Gamma_{\mathcal{U}}(\mathcal{D})$. On the other hand, we have $\mathring{\nabla}_A U_a = \llbracket A, U_a \rrbracket$ for every $A \in \Gamma_{\mathcal{U}}(\mathcal{D}^\perp)$. Accordingly, if $[U_a, \Gamma(\mathcal{D}^\perp)] \subseteq \Gamma(\mathcal{D}^\perp)$, then $\mathring{\nabla}_A U_a = 0$, i.e., (U_a) is also parallel with respect to $\mathring{\nabla}$. It follows that the curvature tensor of $\mathring{\nabla}$ vanishes on \mathcal{U} . □

Let $\pi : \mathcal{D} \rightarrow M$ be the natural projection. Associated to the nonholonomic connection ∇ and its extension $\mathring{\nabla}$ are the restricted connections

$$h : \pi^*\mathcal{D} \rightarrow T\mathcal{D}, \quad h(U_q, X_q) = T_q U \cdot X_q - \nu_{U_q} \cdot \nabla_{X_q} U(q)$$

and

$$f : \pi^*TM \rightarrow T\mathcal{D}, \quad f(U_q, Z_q) = T_q U \cdot Z_q - v_{U_q} \cdot \overset{\circ}{\nabla}_{Z_q} U(q),$$

respectively. Notice that $f|_{\pi^*\mathcal{D}} = h$; in particular, if $X \in \Gamma(\mathcal{D})$, then $X^f = X^h$ (where X^f is the f -lift of X , and X^h is the h -lift of X). Let $\mathcal{V} = \ker T\pi$ be the vertical distribution, $\mathcal{H} = \text{im } h$ the horizontal distribution of h and $\mathcal{F} = \text{im } f$ the horizontal distribution of f . We have $\mathcal{H} \subsetneq \mathcal{F}$; moreover, by Proposition 4, it follows that $\mathcal{V} \cap \mathcal{H} = \mathcal{V} \cap \mathcal{F} = \{0\}$, $\mathcal{V} + \mathcal{H} \subsetneq T\mathcal{D}$ and $\mathcal{V} + \mathcal{F} = T\mathcal{D}$. That is,

$$\mathcal{V} \oplus \mathcal{H} \subsetneq \mathcal{V} \oplus \mathcal{F} = T\mathcal{D}.$$

Let \mathcal{H}^\perp denote the distribution on \mathcal{D} given by $\mathcal{H}^\perp = \text{im}(f|_{\pi^*\mathcal{D}^\perp})$, i.e., $\mathcal{H}^\perp_{U_q} = \text{span}\{f(U_q, X_q) : X_q \in \mathcal{D}^\perp_q\}$ for each $U_q \in \mathcal{D}$. We have $\pi_*\mathcal{H}^\perp = \mathcal{D}^\perp$ and $\mathcal{F} = \mathcal{H} \oplus \mathcal{H}^\perp$.

Remark 7 We have

$$\mathcal{V} \oplus \mathcal{H} = (T\pi)^{-1}(\mathcal{D}) \quad \text{and} \quad \mathcal{V} \oplus \mathcal{H}^\perp = (T\pi)^{-1}(\mathcal{D}^\perp).$$

As noted in, e.g., [11], a nonholonomic connection on \mathcal{D} is precisely the specification of a ϕ_t -invariant complement \mathcal{H} to \mathcal{V} in $(T\pi)^{-1}(\mathcal{D})$, where $\phi_t : \mathcal{D} \rightarrow \mathcal{D}$ is the dilation $\phi_t(U_q) = e^t U_q$. Indeed, given such a complement, we have $\pi_*\mathcal{H} = \mathcal{D}$, $\mathcal{V} \cap \mathcal{H} = \{0\}$ and $(\phi_t)_*\mathcal{H} = \mathcal{H}$. Hence, by Proposition 4 and Proposition 5, there exists a unique linear \mathcal{D} -connection h on \mathcal{D} with $\text{im } h = \mathcal{H}$.

Let $\mathcal{V} : T\mathcal{D} \rightarrow \mathcal{V}$, $\mathcal{P} : T\mathcal{D} \rightarrow \mathcal{V} \oplus \mathcal{H}$ and $\mathcal{Q} : T\mathcal{D} \rightarrow \mathcal{H}^\perp$ denote the projections corresponding to the decomposition $T\mathcal{D} = \mathcal{V} \oplus \mathcal{H} \oplus \mathcal{H}^\perp$. Let $\llbracket \cdot, \cdot \rrbracket : \Gamma(T\mathcal{D}) \times \Gamma(T\mathcal{D}) \rightarrow \Gamma(\mathcal{V} \oplus \mathcal{H})$ be the projected Lie bracket $\mathcal{P}(\llbracket \cdot, \cdot \rrbracket)$. If $X \in \Gamma(\mathcal{F})$ is projectable, then $\mathcal{P}(X) = \mathcal{P}(\pi_*X)^h$ and $\mathcal{Q}(X) = \mathcal{Q}(\pi_*X)^f$.

Theorem 6 We have

$$\begin{aligned} K(X, Y)U_q &= -v_{U_q}^{-1} \cdot ([X^h, Y^h](U_q) - [X, Y]^f(U_q)) \\ &= -v_{U_q}^{-1} \cdot \mathcal{V}(\llbracket X^h, Y^h \rrbracket)(U_q) \end{aligned}$$

for $X, Y \in \Gamma(\mathcal{D})$ and $U_q \in \mathcal{D}$.

Proof Let $X, Y \in \Gamma(\mathcal{D})$ and $\omega \in \Gamma(\mathcal{D}^*)$. We may interpret the $(1, 1)$ -tensor field $K(X, Y)$ as an element of $\text{Der}_0(\mathcal{D})$. In particular, $K(X, Y)\omega \in \Gamma(\mathcal{D}^*)$ is given by

$$(K(X, Y)\omega)(U) = K(X, Y)(\omega(U)) - \omega(K(X, Y)U) = -\omega(K(X, Y)U),$$

for $U \in \Gamma(\mathcal{D})$. By Lemma 2, we have $Z^f[\overline{\omega}] = \overset{\circ}{\nabla}_Z \omega$ for $Z \in \Gamma(TM)$; hence

$$\overline{K(X, Y)\omega} = \overline{[\nabla_X, \nabla_Y]\omega} - \overset{\circ}{\nabla}_{[X, Y]}\omega = [X^h, Y^h][\overline{\omega}] - [X, Y]^f[\overline{\omega}].$$

Since $\pi_*([X^h, Y^h] - [X, Y]^f) = [\pi_*X^h, \pi_*Y^h] - [X, Y] = 0$, we have that $[X^h, Y^h] - [X, Y]^f$ is vertical. Hence

$$\begin{aligned}
 [X^h, Y^h] - [X, Y]^f &= \mathcal{V}([X^h, Y^h] - [X, Y]^f) \\
 &= \mathcal{V}(\llbracket X^h, Y^h \rrbracket - \llbracket [X, Y]^h \rrbracket) + \mathcal{V}(\mathcal{Q}([X^h, Y^h]) - \mathcal{Q}(\llbracket [X, Y]^h \rrbracket)) \\
 &= \mathcal{V}(\llbracket X^h, Y^h \rrbracket),
 \end{aligned}$$

and so $\overline{K(X, Y)\omega} = \mathcal{V}(\llbracket X^h, Y^h \rrbracket)[\overline{\omega}]$. Let $U_q \in \mathcal{D}$; then

$$\begin{aligned}
 \omega_q(K(X, Y)U_q) &= -(K(X, Y)\omega)_q(U_q) = -(\overline{K(X, Y)\omega})(U_q) \\
 &= -\mathcal{V}(\llbracket X^h, Y^h \rrbracket)[\overline{\omega}](U_q) \\
 &= -d\overline{\omega}(U_q)(\mathcal{V}(\llbracket X^h, Y^h \rrbracket)(U_q)).
 \end{aligned}$$

By Lemma 1, we thus have

$$\omega_q(K(X, Y)U_q) = -\omega_q(v_{U_q}^{-1} \cdot \mathcal{V}(\llbracket X^h, Y^h \rrbracket)(U_q)).$$

Since ω is arbitrary, the result follows. □

Corollary 2 $K \equiv 0$ if and only if $\llbracket \mathcal{H}, \mathcal{H} \rrbracket \subseteq \mathcal{H}$.

5.2 The Wagner tensor

Suppose that $(M, \mathcal{D}, \mathcal{D}^\perp, \mathbf{g})$ is a Wagner structure (with degree of nonholonomy N). Associated to each ∇^i is the restricted connection

$$h^i : \pi^*\mathcal{D}^i \rightarrow T\mathcal{D}, \quad h^i(U_q, X_q) = T_q U \cdot X_q - v_{U_q} \cdot \nabla_{X_q}^i U(q).$$

We have $h^1 = h$ and $h^{i+1}|_{\pi^*\mathcal{D}^i} = h^i$; let $\mathcal{H}^i = \text{im } h^i$, $\mathcal{Q}^i = \text{im}(h^{i+1}|_{\pi^*\mathcal{E}^i})$ and let \mathcal{V} be the vertical distribution; then $\mathcal{H}^{i+1} = \mathcal{H}^i \oplus \mathcal{Q}^i$, and hence

$$T\mathcal{D} = \mathcal{V} \oplus \mathcal{H}^N = \mathcal{V} \oplus \mathcal{H} \oplus \mathcal{Q}^1 \oplus \dots \oplus \mathcal{Q}^{N-1}. \tag{3}$$

Let $\mathcal{V} : T\mathcal{D} \rightarrow \mathcal{V}$, $\mathcal{P} : T\mathcal{D} \rightarrow \mathcal{V} \oplus \mathcal{H}$ and $\mathcal{Q}_i : T\mathcal{D} \rightarrow \mathcal{Q}^i$ be the projections corresponding to the decomposition (3); similarly, let $\mathcal{P}_i = \mathcal{P} \oplus \mathcal{Q}_i \oplus \dots \oplus \mathcal{Q}_{i-1}$ be the projection onto $\mathcal{V} \oplus \mathcal{H}^i$ (by convention, we take $\mathcal{P}_1 = \mathcal{P}$). If $X \in \Gamma(\mathcal{H}^N)$ is projectable, then

$$\mathcal{P}(X) = \mathcal{P}(\pi_* X), \quad \mathcal{Q}_i(X) = \mathcal{Q}_i(\pi_* X)^{h^{i+1}} \quad \text{and} \quad \mathcal{P}_i(X) = \mathcal{P}_i(\pi_* X)^{h^i}.$$

A result similar to Theorem 6 holds for each of the curvature tensors K^1, \dots, K^N . Let $\overset{\circ}{\nabla}^i$ be the \mathcal{D}^{i+1} -connection on \mathcal{D} given by

$$\overset{\circ}{\nabla}_X^i = \nabla_{\mathcal{P}_i(X)}^i + \mathcal{L}_{\mathcal{Q}_i(X)}^{\mathcal{P}}, \quad X \in \Gamma(\mathcal{D}^{i+1}).$$

Let $f^i : \pi^*\mathcal{D}^{i+1} \rightarrow T\mathcal{D}$ be the associated horizontal lift, i.e.,

$$f^i(U_q, X_q) = T_q U \cdot X_q - v_{U_q} \cdot \overset{\circ}{\nabla}_{X_q}^i U(q), \quad (U_q, X_q) \in \pi^*\mathcal{D}^{i+1}.$$

Theorem 7 *We have*

$$\begin{aligned} K^i(X, Y)U_q &= -v_{U_q}^{-1} \cdot ([X^{hi}, Y^{hi}](U_q) - [X, Y]^{fi}(U_q)) \\ &= -v_{U_q}^{-1} \cdot \mathcal{H}(\mathcal{P}_i([X^{hi}, Y^{hi}]))(U_q) \end{aligned}$$

for $X, Y \in \Gamma(\mathcal{D}^i)$ and $U_q \in \mathcal{D}$ ($i = 1, \dots, N$).

Proof (The proof is similar to that of Theorem 6, thus we shall omit some details.) Let $X, Y \in \Gamma(\mathcal{D}^i)$ and $\omega \in \Gamma(\mathcal{D}^*)$; then $[X, Y] \in \Gamma(\mathcal{D}^{i+1})$, and hence $K^i(X, Y)\omega = [\nabla_X^i, \nabla_Y^i]\omega - \check{\nabla}_{[X, Y]}^i\omega$. In fact, by Lemma 2, we have

$$\overline{K^i(X, Y)\omega} = [X^{hi}, Y^{hi}][\overline{\omega}] - [X, Y]^{fi}[\overline{\omega}].$$

Clearly, we have that $[X^{hi}, Y^{hi}] - [X, Y]^{fi}$ is vertical; thus $[X^{hi}, Y^{hi}] - [X, Y]^{fi} = \mathcal{H}(\mathcal{P}_i([X^{hi}, Y^{hi}]))$, and so $K^i(X, Y)\omega = \mathcal{H}(\mathcal{P}_i([X^{hi}, Y^{hi}]))[\overline{\omega}]$. If $U_q \in \mathcal{D}$, then

$$\begin{aligned} \omega_q(K^i(X, Y)U_q) &= -\mathcal{H}(\mathcal{P}_i([X^{hi}, Y^{hi}]))(U_q) \\ &= -\omega_q(v_{U_q}^{-1} \cdot \mathcal{H}(\mathcal{P}_i([X^{hi}, Y^{hi}]))(U_q)), \end{aligned}$$

whence $K^i(X, Y)U_q = -v_{U_q}^{-1} \cdot \mathcal{H}(\mathcal{P}_i([X^{hi}, Y^{hi}]))(U_q)$. □

Corollary 3 *We have $K^i \equiv 0$ if and only if $\mathcal{P}_i([\mathcal{H}^i, \mathcal{H}^i]) \subseteq \mathcal{H}^i$. In particular, $(M, \mathcal{D}, \mathcal{D}^\perp, \mathfrak{g})$ is locally flat (i.e., $K^N \equiv 0$) if and only if \mathcal{H}^N is integrable.*

The restricted connections $\nabla^1, \dots, \nabla^N$ are equivalently specified by the horizontal lifts h^1, \dots, h^N , which are in turn equivalently specified by the horizontal distributions $\mathcal{H}^1, \dots, \mathcal{H}^N$. It follows that Wagner’s construction of $\nabla^1, \dots, \nabla^N$ is equivalently formulated as the flag of horizontal distributions on \mathcal{D}

$$\mathcal{H}^1 \subsetneq \mathcal{H}^2 \subsetneq \dots \subsetneq \mathcal{H}^{N-1} \subsetneq \mathcal{H}^N.$$

This flag can be constructed iteratively, starting with $\mathcal{H}^1 = \mathcal{H}$.

Theorem 8 *We have*

$$\mathcal{H}^{i+1} = \mathcal{H}^i + \{[X^{hi}, Y^{hi}] : X \wedge Y \in (\ker \Lambda_i)^\perp\} \quad (i = 1, \dots, N - 1).$$

Proof Let $\mathcal{S}^1 = \mathcal{H}$ and $\mathcal{S}^{i+1} = \mathcal{S}^i + \text{span}\{[X^{hi}, Y^{hi}] : X \wedge Y \in (\ker \Delta_i)^\perp\}$ for $i \geq 1$. We use induction on i to prove that $\mathcal{S}^i = \mathcal{H}^i$; by definition, we have $\mathcal{S}^1 = \mathcal{H}^1$. Suppose that $\mathcal{S}^i = \mathcal{H}^i$ for some $1 \leq i \leq N - 1$. We claim that $\pi_*\mathcal{S}^{i+1} = \mathcal{D}^{i+1}$ and $\mathcal{V} \cap \mathcal{S}^{i+1} = \{0\}$. Let $W_{U_q} + [X^{hi}, Y^{hi}](U_q) \in \mathcal{S}_{U_q}^{i+1}$, where $U_q \in \mathcal{D}$; then

$$T_{U_q}\pi \cdot (W_{U_q} + [X^{hi}, Y^{hi}](U_q)) = T_{U_q}\pi \cdot W_{U_q} + [X, Y](q) \in \mathcal{D}_q^i + [\mathcal{D}^i, \mathcal{D}^i]_q = \mathcal{D}_q^{i+1},$$

and so $\pi_*\mathcal{H}^{i+1} \subseteq \mathcal{D}^{i+1}$. Conversely, let $W_q \in \mathcal{D}_q^i$ and $X_q \in \mathcal{E}_q^i$; then $W_q + X_q$ is an arbitrary element of $\mathcal{D}_q^{i+1} = \mathcal{D}_q^i \oplus \mathcal{E}_q^i$. Since $\mathcal{S}^i = \mathcal{H}^i$ by the inductive hypothesis, we have $\pi_*\mathcal{S}^i = \mathcal{D}^i$. Accordingly, there exists $V_{U_q} + [Y, Z](U_q) \in \mathcal{S}_{U_q}^i$ such that $T_{U_q}\pi \cdot (V_{U_q} + [Y, Z](U_q)) = W_q$. Let $A \wedge B = \Theta_i(X) \in (\ker \Delta_i)^\perp$, where $X \in \Gamma(\mathcal{E}^i)$ is a smooth extension of X_q . (The case

when $\Theta_i(X)$ is a $C^\infty(\mathbf{M})$ -combination of bivector fields from $\Gamma(\mathcal{D}^i)$ can be treated in a similar fashion.) Then

$$\begin{aligned} W_q + X_q &= T_{U_q} \pi \cdot (V_{U_q} + [Y, Z](U_q)) + \mathcal{Q}_i([A, B])(q) \\ &= T_{U_q} \pi \cdot (V_{U_q} + [Y, Z](U_q) + \mathcal{Q}_i([A, B])^{h^{i+1}}(U_q)) \\ &= T_{U_q} \pi \cdot (V_{U_q} + [Y, Z](U_q) + \mathcal{Q}_i([A^{h^i}, B^{h^i}])(U_q)) \\ &= T_{U_q} \pi \cdot \underbrace{(V_{U_q} + [Y, Z](U_q) - \mathcal{P}_i([A^{h^i}, B^{h^i}])(U_q) + [A^{h^i}, B^{h^i}])(U_q)}_{\in \mathcal{S}_{U_q}^i} \\ &\in T_{U_q} \pi \cdot \mathcal{S}_{U_q}^{i+1}. \end{aligned}$$

Hence $\pi_* \mathcal{S}^{i+1} = \mathcal{D}^{i+1}$. Let $W_{U_q} + [X^{h^i}, Y^{h^i}](U_q) \in \mathcal{V}_{U_q} \cap \mathcal{S}_{U_q}^{i+1}, U_q \in \mathcal{D}$. We have

$$\begin{aligned} W_{U_q} + [X^{h^i}, Y^{h^i}](U_q) &= \mathcal{V}(\mathcal{P}_i([X^{h^i}, Y^{h^i}])(U_q)) + (\mathcal{P}_i - \mathcal{V})(W_{U_q} \\ &\quad + [X^{h^i}, Y^{h^i}](U_q)) + \mathcal{Q}_i([X^{h^i}, Y^{h^i}])(U_q) \\ &= -v_{U_q} \cdot K^i(X \wedge Y)U_q + (W_{U_q} + \mathcal{P}_i([X, Y])^{h^i}(U_q)) \\ &\quad + \mathcal{Q}_i([X, Y])^{h^{i+1}}(U_q). \end{aligned}$$

Both non-vertical components must vanish; in particular, $\mathcal{Q}_i([X, Y])(q) = 0$, i.e., $(X \wedge Y)(q) \in \ker \Delta_{i,q}$. Since $X \wedge Y \in (\ker \Delta_i)^\perp$, it follows that $(X \wedge Y)(q) = 0$; then

$$W_{U_q} + [X^{h^i}, Y^{h^i}](U_q) = -v_{U_q} \cdot K^i(X \wedge Y)U_q = 0,$$

whence $\mathcal{V} \cap \mathcal{S}^{i+1} = \{0\}$. By uniqueness of the connection associated to \mathcal{H}^{i+1} (Proposition 5), it follows that $\mathcal{S}^{i+1} = \mathcal{H}^{i+1}$. This completes the proof of the inductive case. \square

6 Concluding remark

Further investigation (in the vein of [2]) into geometric interpretations of the Schouten and Wagner curvature tensors (including the tensors K^2, \dots, K^{N-1} involved in Wagner’s construction) would be a worthwhile undertaking. A natural next step would be to consider how curvature affects nonholonomic geodesics; for instance, to generalize the notion of a Jacobi field to nonholonomic Riemannian geometry. A thorough study of curvature in lower dimensions would also be a topic of interest.

Data availability The authors declare that data availability is *not applicable*

References

- Barrett, D.I., Biggs, R., Remsing, C.C., Rossi, O.: Invariant nonholonomic Riemannian structures on three-dimensional Lie groups. *J. Geom. Mech.* **8**, 139–167 (2016)

2. Barrett, D.I., Remsing, C.C.: On geodesic invariance and curvature in nonholonomic Riemannian geometry. *Publ. Math. Debr.* **94**, 197–213 (2019)
3. Barrett, D.I., Remsing, C.C.: A note on flat nonholonomic Riemannian structures on three-dimensional Lie groups. *Beitr. Algebra Geom.* **60**, 419–436 (2019)
4. Bejancu, A.: On the geometry of nonholonomic mechanical systems with vertical distribution. *J. Math. Phys.* **48**, 052903 (2007)
5. Berestovskii, V.N.: Curvatures of homogeneous sub-Riemannian manifolds. *Eur. J. Math.* **3**, 788–807 (2017)
6. Bloch, A.M.: *Nonholonomic Mechanics and Control*, 2nd edn. Springer, New York (2003)
7. Bukusheva, A.V., Galaev, S.V.: Connections on distributions and geodesic sprays. *Russ. Math.* **57**, 7–13 (2013)
8. Cantrijn, F., Langerock, B.: Generalised connections over a vector bundle map. *Diff. Geom. Appl.* **18**, 295–317 (2003)
9. Cortés Monforte, J.: *Geometric Control and Numerical Aspects of Nonholonomic Systems*. Springer, New York (2002)
10. Dragović, V., Gajić, B.: The Wagner curvature tensor in nonholonomic mechanics. *Regul. Chaot. Dyn.* **8**, 105–123 (2003)
11. Galaev, S.V.: The intrinsic geometry of almost contact metric manifolds. *Izv. Saratov Univ. (N.S.) Ser. Math. Mech. Inform.* **12**, 16–22 (2012)
12. Galaev, S.V.: Geometric interpretation of the Wagner curvature tensor in the case of a manifold with contact metric structure. *Sib. Math. J.* **57**, 498–504 (2016)
13. Galaev, S.V., Gokhman, A.V.: A metrizable condition for an affine connection on a nonholonomic manifold X_3^2 (in Russian). In: *Collection of Scientific Works. Mathematics. Mechanics*, pp. 28–32. Izv. Saratov University Press, Saratov (2005)
14. Gorbatenko, E.M.: Differential geometry of nonholonomic manifolds (after V.V. Wagner) (in Russian), *Geom. Sbornik*, 31–43 (1985)
15. Grozman, P.Y., Leites, D.A.: The Riemann tensor for nonholonomic manifolds. *Theor. Math. Phys.* **153**, 1511–1538 (2007)
16. Hertz, H.: *The Principles of Mechanics Presented in a New Form* (in German), Leipzig (1894). MacMillan, English translation, Basingstoke (1899)
17. Koiller, J., Rodrigues, P.R., Pitanga, P.: Non-holonomic connections following Élie Cartan. *An. Acad. Bras. Cienc.* **73**, 165–190 (2001)
18. Langerock, B.: Nonholonomic mechanics and connections over a bundle map. *J. Phys. A Math. Gen.* **34**, L609–L615 (2001)
19. Leites, D.A.: The Riemann tensor for nonholonomic manifolds. *Homol. Homotopy Appl.* **4**, 397–407 (2002)
20. Montgomery, R.: *A Tour of Subriemannian Geometries*. In: American Mathematical Society, Providence, RI, *Their Geodesics and Applications* (2002)
21. O’Neill, B.: *Semi-Riemannian Geometry*. Academic Press, New York (1983)
22. Petersen, P.: *Riemannian Geometry*, 2nd edn. Springer, Berlin (2006)
23. Poor, W.A.: *Differential Geometric Structures*. McGraw-Hill, New York (1981)
24. Popescu, P.: Almost Lie structures, derivations and R-curvature on relative tangent spaces. *Rev. Roum. Math. Pures Appl.* **37**, 779–789 (1992)
25. Schouten, J.A.: On nonholonomic connections. In: *Proceedings of Sciences*, Vol. 31, Kon. Akad. Wetensch., Amsterdam (1928)
26. Synge, J.L.: Geodesics in nonholonomic geometry. *Math. Ann.* **99**, 738–751 (1928)
27. Szilasi, J., Lovas, R.L., Kertész, D.C.: *Connections, Sprays and Finsler Structures*. World Scientific, Singapore (2014)
28. Vershik, A.M., Gershkovich, V.Y.: Nonholonomic problems and the theory of distributions. *Acta Appl. Math.* **12**, 181–209 (1988)
29. Vershik, A.M., Gershkovich, V.: Ya.: nonholonomic dynamical systems, geometry of distributions and variational problems. In: Arnol’d, V.I., Novikov, S.P. (eds.) *Dynamical Systems VII*. Springer, New York (1994)
30. Wagner, V.V.: On the differential geometry of nonholonomic manifolds (in French). In: *Mémoires du Séminaire Pour L’Analyse Vectorielle et Tensorielle*, Livres II–III, pp. 267–318 (1935)
31. Wagner, V.V.: On the geometrical interpretation of the curvature vector of a non-holonomic V_3 in the three-dimensional Euclidean space. *Rec. Math.* **4**, 339–356 (1938)
32. Wagner, V.V.: On V_3^2 in R_3 with zero curvature (in German). *Rec. Math.* **4**, 333–338 (1938)

33. Wagner, V.V.: Differential geometry of nonholonomic manifolds (in Russian). In: Report on The VIII-th International Competition for the N.I. Lobachevskii Prize (1937), Kazan Physico-Mathematical Society (1940)
34. Wagner, V.V.: Geometric interpretation of the motion of nonholonomic dynamical systems (in Russian). Tr. Semin. Vectorn. Tenzorn. Anal. **5**, 301–327 (1941)
35. Wagner, V.V.: The inner geometry of non-linear non-holonomic manifolds. Rec. Math. **13**, 135–167 (1943)
36. Zhao, P., Jiao, L.: Conformal transformations on Carnot-Carathéodory spaces. Nihonkai Math. J. **17**, 167–185 (2006)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.