


Article

# Entropy Dissipation for Degenerate Stochastic Differential Equations via Sub-Riemannian Density Manifold

Qi Feng <sup>1,\*</sup>  and Wuchen Li <sup>2</sup>

<sup>1</sup> Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA

<sup>2</sup> Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA; wuchen@mailbox.sc.edu

\* Correspondence: qif@umich.edu

**Abstract:** We studied the dynamical behaviors of degenerate stochastic differential equations (SDEs). We selected an auxiliary Fisher information functional as the Lyapunov functional. Using generalized Fisher information, we conducted the Lyapunov exponential convergence analysis of degenerate SDEs. We derived the convergence rate condition by generalized Gamma calculus. Examples of the generalized Bochner's formula are provided in the Heisenberg group, displacement group, and Martinet sub-Riemannian structure. We show that the generalized Bochner's formula follows a generalized second-order calculus of Kullback–Leibler divergence in density space embedded with a sub-Riemannian-type optimal transport metric.

**Keywords:** degenerate drift–diffusion process; Lyapunov methods; auxiliary Fisher information; sub-Riemannian density manifold; generalized Bochner's formula

**MSC:** 53C17; 60D05; 58B20



**Citation:** Feng, Q.; Li, W. Entropy Dissipation for Degenerate Stochastic Differential Equations via Sub-Riemannian Density Manifold. *Entropy* **2023**, *25*, 786. <https://doi.org/10.3390/e25050786>

Academic Editors: Constantin Udriste and Ionel Tevy

Received: 29 March 2023

Revised: 24 April 2023

Accepted: 9 May 2023

Published: 11 May 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Consider the following Stratonovich stochastic differential equation:

$$dX_t = b(X_t)dt + \sqrt{2}a(X_t) \circ dB_t, \quad (1)$$

where  $(B_t^1, B_t^2, \dots, B_t^n)$  is an  $n$ -dimensional Brownian motion in  $\mathbb{R}^n$ ,  $a \in \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{(n+m) \times n}$  is a matrix-valued function, and  $b : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m}$  is a drift vector field. The convergence analysis of SDE (1) to its invariant distribution lies in the intersection of differential geometry, analysis, the Lie group (subgroup in quantum mechanics), and probability. The convergence analysis also has broad applications in designing fast algorithms in artificial intelligence (AI) and Bayesian sampling/optimization problems. One key question arises: How fast does the probability density function of SDE (1) converge to its invariant distribution?

The Gamma calculus, also named Bakry–Émery iterative calculus [1], provides analytical approaches to derive the convergence rate for SDE (1). This lower bound is known as the Ricci curvature lower bound. However, classical studies are limited to the non-degenerate diffusion coefficient matrix  $a$ . The classical Gamma calculus is no longer valid when  $a$  is a degenerate matrix function; see the generalization of Bakry–Émery calculus in [2].

This paper presents a Lyapunov convergence analysis for the degenerate diffusion process. We selected a class of  $z$ -Fisher information as the Lyapunov functional, where  $z$  is a matrix function different from matrix  $a$ . We derived a generalized Gamma calculus by the dissipation of the Lyapunov functional along the diffusion process. We then derived the generalized Bochner's formula and obtained the exponential convergence condition. Several concrete examples are presented: gradient-drift-diffusions on the Heisenberg group, the displacement group, and the Martinet sub-Riemannian structure. Our approach

extends the classical optimal transport geometry, in particular the second-order calculus of the relative entropy in the density manifold studied in [3–6].

The generalized Gamma calculus was first introduced by Baudoin–Garofalo [2] for sub-Riemannian manifolds. Related results were studied later in [7–15]. The commutative property of the iteration of  $\Gamma_1$  and  $\Gamma_1^z$  (Hypothesis 1.2 in [2]) was crucial in the previous works. Our algebraic Condition 1 does not have this requirement. We can remove this commutative condition in the weak sense. Thus, our results go beyond the step two-bracket-generating condition. We present algebraic conditions for the existence of the generalized Bochner’s formula.

On the other hand, optimal transport on the sub-Riemannian manifold was studied by [16–19]. An optimal transport metric on a sub-Riemannian manifold was proposed in [18,19]. In this case, the density manifold still forms an infinite-dimensional Riemannian manifold. The Monge–Ampère equation in sub-Riemannian settings was studied in [17]. Our approach is different. We introduced the sub-Riemannian density manifold (SDM) and studied its second-order geometric calculations of relative entropies in the SDM. Using those, we propose a new Gamma  $z$  calculus for degenerate stochastic differential equations and established the generalized curvature dimension-type bound. Besides, Refs. [20,21] used the analytical property of optimal transport to formulate the Ricci curvature lower bound in general metric space. Different from [20–22], we focused on the geometric calculations in the density manifold introduced by the  $z$  direction. Following the second-order geometric calculations in the density manifold, we formulated the new Gamma calculus and the corresponding Ricci curvature tensor for the sub-Riemannian manifold. Besides, our derivation also relates to the entropy methods [23,24]. Using entropy methods, Refs. [25,26] derived the convergence rate for degenerate drift–diffusion processes with constant diffusion coefficients  $a$ . Compared to previous works, we applied the entropy method with Gamma calculus and geometric calculations in the density manifold. It derives a generalized Gamma calculus from the dissipation of auxiliary Fisher information. Several concrete examples of convergence conditions are derived in the Lie-group-induced drift–diffusion processes.

We organize the paper as follows. We introduce the main result in Section 2. It is an explicit convergence rate condition for the density of degenerate SDEs in the  $L^1$  distance. In Section 3, we provide three examples of the proposed convergence analysis, including gradient-drift–diffusions on the Heisenberg group, the displacement group, and the Martinet sub-Riemannian structure. In Section 4, we present the Lyapunov analysis in the sub-Riemannian density manifold. The generalized Gamma calculus and the proof of the generalized Bochner’s formula is presented in Section 5. Some further discussions for other functional inequalities are presented in Section 6.

## 2. Main Results

In this section, we present this paper’s setting and main results.

### 2.1. Setting

Consider a Stratonovich SDE:

$$dX_t = b(X_t)dt + \sqrt{2}a(X_t) \circ dB_t, \quad (2)$$

where  $(B_t^1, B_t^2, \dots, B_t^n)$  is an  $n$ -dimensional Brownian motion in  $\mathbb{R}^n$ ,  $a : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{(n+m) \times n}$  is a matrix-valued function, and  $b : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m}$  is a vector field. We refer to [27] (Section 3.13) for the definition of the Stratonovich SDE. According to [28] (Appendix A.7), the SDE (2) can also be written as the following Itô SDE:

$$dX_{l,t} = \tilde{b}_l(X_t)dt + \sum_{i=1}^n \sqrt{2}a_{li}(X_t)dB_t^i, \quad \text{for } l = 1, \dots, n+m, \quad (3)$$

where

$$\tilde{b}_l = b_l + \left( \sum_{i=1}^n \nabla_{a_i} a_i \right)_l, \quad \text{for } l = 1, \dots, n+m. \quad (4)$$

We denote  $\{a_1, \dots, a_n\}$  as the column vectors of matrix  $a$ , and  $\sum_{i=1}^n \nabla_{a_i} a_i \in \mathbb{R}^{n+m}$  represents

$$\left( \sum_{i=1}^n \nabla_{a_i} a_i \right)_l = \sum_{i=1}^n \sum_{k=1}^{n+m} a_{ki} \frac{\partial a_{li}}{\partial x_k}, \quad \text{for } l = 1, \dots, n+m. \quad (5)$$

We denote  $a^\top$  as the transpose of matrix  $a$  and denote  $\{a_1^\top, \dots, a_n^\top\}$  as the row vectors of matrix  $a^\top$ . In particular, we have  $a_{\hat{i}}^\top = a_{i\hat{i}}^\top$ , for  $i = 1, \dots, n$  and  $\hat{i} = 1, \dots, n+m$ . With some abuse of notation, we also denote  $a_i^\top$  as the vector fields corresponding to the row vectors  $a_i^\top$ , for  $i = 1, \dots, n$ . We assumed that  $\{a_1^\top(x), a_2^\top(x), \dots, a_n^\top(x)\}$  satisfies the strong Hörmander condition (or bracket-generating condition):

$$\text{Span}\{a_1^\top(x), \dots, a_n^\top(x), [a_{i_1}^\top, \dots, [a_{i_{k-1}}^\top, a_{i_k}^\top] \dots](x), 1 \leq i_1, \dots, i_k \leq n, k \geq 2\} = \mathbb{R}^{n+m},$$

where  $[\cdot, \cdot]$  represents the Lie bracket between two vector fields. The strong Hörmander condition means that the Lie algebra generated by the vector fields  $\{a_1^\top(x), \dots, a_n^\top(x)\}$  is of full rank at every point  $x \in \mathbb{R}^{n+m}$  (see, e.g., [29] (Section 7.4)). This condition ensures the existence of a smooth probability density function of SDE (2); see the original proofs in [30,31]. For the simplicity of presentation, we assumed the probability density function is strictly positive. Indeed, the positivity of the density follows from the Hörmander condition [32]; for the more technical conditions to show the positivity by using Malliavin calculus, we refer to [33,34] (Theorem 1.4 with  $H = 1/2$ ). Denote  $X_t \sim \rho(t, x)$ , where  $\rho = \rho(t, x)$  is the probability density function of SDE (2). The density function  $\rho$  satisfies the Fokker–Planck equation of SDE (2):

$$\partial_t \rho(t, x) = -\nabla_x \cdot (\rho(t, x) \tilde{b}(x)) + \sum_{i=1}^{n+m} \sum_{j=1}^{n+m} \frac{\partial^2}{\partial x_i \partial x_j} \left( (a(x) a(x)^\top)_{ij} \rho(t, x) \right), \quad (6)$$

with a smooth initial condition:

$$\rho_0(x) = \rho(0, x), \quad \int_{\mathbb{R}^{n+m}} \rho_0(x) dx = 1, \quad \rho_0(x) > 0.$$

In this paper, we assumed that SDE (2) has a unique invariant symmetric measure  $\mu$ , where  $d\mu = \pi(x) dx$  with  $\pi \in C^\infty(\mathbb{R}^{n+m})$ . Here,  $\pi$  solves the equilibrium of Fokker–Planck Equation (6):

$$-\nabla_x \cdot (\pi(x) \tilde{b}(x)) + \sum_{i=1}^{n+m} \sum_{j=1}^{n+m} \frac{\partial^2}{\partial x_i \partial x_j} \left( (a(x) a(x)^\top)_{ij} \pi(x) \right) = 0.$$

We studied a particular class of the vector field  $b$  for a given invariant distribution  $\pi$ . **Assumption (Gradient flow formulation):** Suppose that  $b, a$ , and  $\pi$  satisfy the relation:

$$b = a \otimes \nabla a + a a^\top \nabla \log \pi, \quad (7)$$

where  $a \otimes \nabla a \in \mathbb{R}^{n+m}$  represents, for  $\hat{k} = 1, \dots, n+m$ ,

$$(a \otimes \nabla a)_{\hat{k}} = \sum_{k=1}^n \sum_{k'=1}^{n+m} a_{\hat{k}k} \frac{\partial}{\partial x_{k'}} a_{k'k}. \quad (8)$$

In the Itô formulation,  $\tilde{b}$ ,  $a$ , and  $\pi$  satisfy

$$\tilde{b}_l = (aa^T \nabla \log \pi)_l + (a \otimes \nabla a)_l + \left( \sum_{i=1}^n \nabla_{a_i} a_i \right)_l,$$

for  $l = 1, \dots, n+m$ . In this case, we can reformulate Equation (6) as

$$\partial_t \rho(t, x) = \nabla \cdot \left( \rho(t, x) a(x) a(x)^T \nabla \log \frac{\rho(t, x)}{\pi(x)} \right). \quad (9)$$

We leave the derivation of Formula (9) in Appendix A. If  $\rho(t, x) = \pi(x)$ , then  $\log \frac{\rho(t, x)}{\pi(x)} = 0$ , and  $\pi$  is an invariant density function for SDE (2). In Section 4, we demonstrate that Fokker–Planck Equation (6), or its equivalent Formulation (9), forms a “horizontal” gradient flow in the sub-Riemannian density manifold. We designed a Lyapunov functional to study the convergence behavior of this “horizontal” gradient flow (9).

**Remark 1.** Formula (9) can be written as

$$\partial_t (\log \rho(t, x) - \log \pi(x)) \rho(t, x) = \nabla \cdot \left( \rho(t, x) a(x) a(x)^T \nabla \log \frac{\rho(t, x)}{\pi(x)} \right).$$

It has a weak formulation that

$$\int_{\mathbb{R}^{n+m}} (\partial_t \log \frac{\rho(t, x)}{\pi(x)}, \phi(x)) \rho(t, x) dx = - \int_{\mathbb{R}^{n+m}} \left( \nabla \phi(x), a(x) a(x)^T \nabla \log \frac{\rho(t, x)}{\pi(x)} \right) \rho(t, x) dx,$$

where  $\phi \in C^\infty(\mathbb{R}^{n+m})$  is a smooth test function.

**Remark 2** (Non-gradient flow drift). In fact, the proposed method is not limited to the gradient flow assumption of the drift vector field  $b$  in (7). See the details in [35].

## 2.2. Main Result

We now briefly sketch the main results. Denote a sub-elliptic operator  $L: C^\infty(\mathbb{R}^{n+m}) \rightarrow C^\infty(\mathbb{R}^{n+m})$  as follows:

$$Lf = \nabla \cdot (aa^T \nabla f) - \langle a \otimes \nabla a, \nabla f \rangle_{\mathbb{R}^{n+m}} + \langle b, \nabla f \rangle_{\mathbb{R}^{n+m}},$$

where  $f \in C^\infty(\mathbb{R}^{n+m})$ .

**Definition 1** (Generalized Gamma z calculus). Consider a smooth matrix function  $z: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{(n+m) \times m}$ . Denote Gamma one bilinear forms  $\Gamma_1, \Gamma_1^z: C^\infty(\mathbb{R}^{n+m}) \times C^\infty(\mathbb{R}^{n+m}) \rightarrow C^\infty(\mathbb{R}^{n+m})$  as

$$\Gamma_1(f, g) = \langle a^T \nabla f, a^T \nabla g \rangle_{\mathbb{R}^n}, \quad \Gamma_1^z(f, g) = \langle z^T \nabla f, z^T \nabla g \rangle_{\mathbb{R}^m}.$$

Define Gamma two bilinear forms  $\Gamma_2, \Gamma_2^{z, \pi}: C^\infty(\mathbb{R}^{n+m}) \times C^\infty(\mathbb{R}^{n+m}) \rightarrow C^\infty(\mathbb{R}^{n+m})$  as

$$\Gamma_2(f, g) = \frac{1}{2} \left[ L\Gamma_1(f, g) - \Gamma_1(Lf, g) - \Gamma_1(f, Lg) \right],$$

and

$$\Gamma_2^{z, \pi}(f, g) = \frac{1}{2} \left[ L\Gamma_1^z(f, g) - \Gamma_1^z(Lf, g) - \Gamma_1^z(f, Lg) \right] \quad (10)$$

$$+ \text{div}_z^\pi \left( \Gamma_{1, \nabla(aa^T)}(f, g) \right) - \text{div}_a^\pi \left( \Gamma_{1, \nabla(zz^T)}(f, g) \right). \quad (11)$$

Here,  $\mathbf{div}_a^\pi, \mathbf{div}_z^\pi$  are divergence operators defined by

$$\mathbf{div}_a^\pi(F) = \frac{1}{\pi} \nabla \cdot (\pi aa^\top F), \quad \mathbf{div}_z^\pi(F) = \frac{1}{\pi} \nabla \cdot (\pi zz^\top F),$$

for any smooth vector field  $F \in \mathbb{R}^{n+m}$ , and  $\Gamma_{\nabla(aa^\top)}, \Gamma_{\nabla(zz^\top)}$  are vector Gamma one bilinear forms defined as

$$\begin{aligned} \Gamma_{1,\nabla(aa^\top)}(f, g) &= \langle \nabla f, \nabla(aa^\top) \nabla g \rangle = \left( \langle \nabla f, \frac{\partial}{\partial x_{\hat{k}}} (aa^\top) \nabla g \rangle \right)_{\hat{k}=1}^{n+m}, \\ \Gamma_{1,\nabla(zz^\top)}(f, g) &= \langle \nabla f, \nabla(zz^\top) \nabla g \rangle = \left( \langle \nabla f, \frac{\partial}{\partial x_{\hat{k}}} (zz^\top) \nabla g \rangle \right)_{\hat{k}=1}^{n+m}, \end{aligned}$$

with

$$\begin{aligned} \mathbf{div}_z^\pi(\Gamma_{\nabla(aa^\top)} f, g) &= \frac{\nabla \cdot (zz^\top \pi \langle \nabla f, \nabla(aa^\top) \nabla g \rangle)}{\pi}, \\ \mathbf{div}_a^\pi(\Gamma_{\nabla(zz^\top)} f, g) &= \frac{\nabla \cdot (aa^\top \pi \langle \nabla f, \nabla(zz^\top) \nabla g \rangle)}{\pi}. \end{aligned}$$

We next demonstrate that the summation of  $\Gamma_2$  and  $\Gamma_2^{z,\pi}$  can induce the following decomposition and bilinear forms. They are natural extensions of the classical Bakry–Émery calculus in the Riemannian manifold, i.e., non-degenerate matrix function  $a$ .

**Notation 1.** For matrix function  $a : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{(n+m) \times n}$ , we define matrix  $Q$  as

$$Q = \begin{pmatrix} a_{11}^\top a_{11}^\top & \cdots & a_{1(n+m)}^\top a_{1(n+m)}^\top \\ \cdots & a_{ii}^\top a_{kk}^\top & \cdots \\ a_{n1}^\top a_{n1}^\top & \cdots & a_{n(n+m)}^\top a_{n(n+m)}^\top \end{pmatrix} \in \mathbb{R}^{n^2 \times (n+m)^2}, \quad (12)$$

with  $Q_{i\hat{k}\hat{k}} = a_{ii}^\top a_{kk}^\top$ . More precisely, for each row (respectively, column) of  $Q$ , the row (respectively, column) indices of  $Q_{i\hat{k}\hat{k}}$  follow  $\sum_{i=1}^n \sum_{k=1}^m$  (respectively,  $\sum_{\hat{i}=1}^{n+1} \sum_{\hat{k}=1}^{n+m}$ ). For matrix function  $z : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{(n+m) \times m}$ , we define matrix  $P$  as

$$P = \begin{pmatrix} z_{11}^\top a_{11}^\top & \cdots & z_{1(n+m)}^\top a_{1(n+m)}^\top \\ \cdots & z_{ii}^\top a_{kk}^\top & \cdots \\ z_{m1}^\top a_{n1}^\top & \cdots & z_{m(n+m)}^\top a_{n(n+m)}^\top \end{pmatrix} \in \mathbb{R}^{(nm) \times (n+m)^2}, \quad (13)$$

with  $P_{i\hat{k}\hat{k}} = z_{kk}^\top a_{ii}^\top$ . For smooth function  $f \in C^\infty(\mathbb{R}^{n+m})$ , for any  $\hat{i}, \hat{k}, \hat{j} = 1, \dots, n+m$  and  $i, k = 1, \dots, n$  (or  $1, \dots, m$ ), we define vector  $C \in \mathbb{R}^{(n+m)^2 \times 1}$  with components

$$C_{\hat{i}\hat{k}} = \left[ \sum_{i,k=1}^n \sum_{i'=1}^{n+m} \left( \langle a_{ii'}^\top a_{i'k}^\top, \frac{\partial a_{kk}^\top}{\partial x_{i'}} \rangle, (a^\top \nabla)_{kf} \rangle_{\mathbb{R}^n} - \langle a_{ki'}^\top a_{ik}^\top \frac{\partial a_{ii}^\top}{\partial x_{i'}}, (a^\top \nabla)_{kf} \rangle_{\mathbb{R}^n} \right) \right], \quad (14)$$

where we denote  $(a^\top \nabla)_{kf} = \sum_{k'=1}^{n+m} a_{kk'}^\top \frac{\partial f}{\partial x_{k'}}$ . We define vector  $D \in \mathbb{R}^{n^2 \times 1}$  with components

$$D_{ik} = \sum_{\hat{i}, \hat{k}=1}^{n+m} a_{ii}^\top \frac{\partial a_{kk}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, \quad \text{and} \quad D^\top D = \sum_{i,k} D_{ik} D_{ik}. \quad (15)$$

We define vector  $F \in \mathbb{R}^{(n+m)^2 \times 1}$  with components

$$F_{\hat{i}\hat{k}} = \left[ \sum_{i=1}^n \sum_{k=1}^m \sum_{i'=1}^{n+m} \left( \langle a_{ii'}^\top a_{i'k}^\top, \frac{\partial z_{kk}^\top}{\partial x_{i'}} \rangle, (z^\top \nabla)_{kf} \rangle_{\mathbb{R}^m} - \langle z_{ki'}^\top a_{ik}^\top \frac{\partial a_{ii}^\top}{\partial x_{i'}}, (z^\top \nabla)_{kf} \rangle_{\mathbb{R}^m} \right) \right]. \quad (16)$$

We define vector  $E \in \mathbb{R}^{(n+m) \times 1}$  with components

$$E_{ik} = \sum_{\hat{i}, \hat{k}=1}^{n+m} a_{\hat{i}\hat{i}}^T \frac{\partial z_{\hat{k}\hat{k}}^T}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, \quad \text{and} \quad E^T E = \sum_{i,k} E_{ik} E_{ik}. \quad (17)$$

We define vector  $G \in \mathbb{R}^{(n+m)^2 \times 1}$  with components

$$G_{\hat{i}\hat{j}} = \sum_{i=1}^n \sum_{j=1}^m \sum_{\hat{i}', \hat{j}', \hat{i}''=1}^{n+m} \left[ \left( z_{\hat{j}\hat{j}}^T z_{\hat{j}'}^T \frac{\partial}{\partial x_{\hat{j}'}} a_{\hat{i}\hat{i}}^T a_{\hat{i}''}^T \frac{\partial f}{\partial x_{\hat{i}''}} + z_{\hat{j}\hat{j}}^T z_{\hat{j}'}^T \frac{\partial}{\partial x_{\hat{j}'}} a_{\hat{i}''}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{\hat{i}}^T \right) \right. \\ \left. - \left( a_{\hat{i}\hat{i}}^T a_{\hat{i}''}^T \frac{\partial}{\partial x_{\hat{i}''}} z_{\hat{j}\hat{j}}^T z_{\hat{j}'}^T \frac{\partial f}{\partial x_{\hat{j}'}} + a_{\hat{i}\hat{i}}^T a_{\hat{i}''}^T \frac{\partial}{\partial x_{\hat{i}''}} z_{\hat{j}\hat{j}}^T \frac{\partial f}{\partial x_{\hat{j}'}} z_{\hat{j}\hat{j}}^T \right) \right]. \quad (18)$$

We define  $X$  as the vectorization of the Hessian matrix of function  $f$ :

$$X^T = \left( \frac{\partial^2 f}{\partial x_1 \partial x_1} \quad \cdots \quad \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}} \quad \cdots \quad \frac{\partial^2 f}{\partial x_{n+m} \partial x_{n+m}} \right) \in \mathbb{R}^{1 \times (n+m)^2}. \quad (19)$$

**Assumption 1.** Assume that there exists vectors  $\Lambda_1, \Lambda_2 \in \mathbb{R}^{(n+m)^2 \times 1}$  such that

$$(Q^T Q \Lambda_1 + P^T P \Lambda_2)^T X = (F + C + G + Q^T D + P^T E)^T X.$$

**Definition 2** (Hessian matrix). For smooth function  $f \in C^\infty(\mathbb{R}^{n+m})$ , define a matrix function  $\mathfrak{H} : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{(n+m) \times (n+m)}$  as

$$\mathfrak{H}(\nabla f, \nabla f) = -\Lambda_1^T Q^T Q \Lambda_1 - \Lambda_2^T P^T P \Lambda_2 + D^T D + E^T E + (\mathfrak{H}_{ab} + \mathfrak{H}_{zb} + \mathfrak{H}_\pi)(\nabla f, \nabla f),$$

where we define the following bilinear forms:

$$\mathfrak{H}_{ab}(\nabla f, \nabla f) = \mathfrak{H}_a(\nabla f, \nabla f) - \sum_{i=1}^n \sum_{\hat{i}, \hat{k}=1}^{n+m} \langle (a_{\hat{i}\hat{i}}^T \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial a_{\hat{i}\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^T \nabla f)_i \rangle_{\mathbb{R}^n},$$

$$\mathfrak{H}_a(\nabla f, \nabla f) = \sum_{i,k=1}^n \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{i i'}^T (\frac{\partial a_{\hat{i}\hat{i}}^T}{\partial x_{i'}} \frac{\partial a_{\hat{k}\hat{k}}^T}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (a^T \nabla)_{k f} \rangle_{\mathbb{R}^n} \\ + \sum_{i,k=1}^n \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{i i'}^T a_{\hat{i}\hat{i}}^T (\frac{\partial}{\partial x_{i'}} \frac{\partial a_{\hat{k}\hat{k}}^T}{\partial x_{\hat{i}}}) (\frac{\partial f}{\partial x_{\hat{k}}}), (a^T \nabla)_{k f} \rangle_{\mathbb{R}^n} \\ - \sum_{i,k=1}^n \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{\hat{k}\hat{k}}^T \frac{\partial a_{i i'}^T}{\partial x_{\hat{k}}} \frac{\partial a_{\hat{i}\hat{i}}^T}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}, (a^T \nabla)_{k f} \rangle_{\mathbb{R}^n} \\ - \sum_{i,k=1}^n \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{\hat{k}\hat{k}}^T a_{i i'}^T (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{\hat{i}\hat{i}}^T}{\partial x_{i'}}) \frac{\partial f}{\partial x_{\hat{i}}}, (a^T \nabla)_{k f} \rangle_{\mathbb{R}^n},$$

$$\mathfrak{H}_{zb}(\nabla f, \nabla f) = \mathfrak{H}_z(\nabla f, \nabla f) - \sum_{i=1}^m \sum_{\hat{i}, \hat{k}=1}^{n+m} \langle (z_{\hat{i}\hat{i}}^T \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial z_{\hat{i}\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^T \nabla f)_i \rangle_{\mathbb{R}^m},$$

$$\mathfrak{H}_z(\nabla f, \nabla f) = \sum_{i=1}^m \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{i i'}^T (\frac{\partial a_{\hat{i}\hat{i}}^T}{\partial x_{i'}} \frac{\partial z_{\hat{k}\hat{k}}^T}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (z^T \nabla)_{k f} \rangle_{\mathbb{R}^m} \\ + \sum_{i=1}^m \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{i i'}^T a_{\hat{i}\hat{i}}^T (\frac{\partial}{\partial x_{i'}} \frac{\partial z_{\hat{k}\hat{k}}^T}{\partial x_{\hat{i}}}) (\frac{\partial f}{\partial x_{\hat{k}}}), (z^T \nabla)_{k f} \rangle_{\mathbb{R}^m} \\ - \sum_{i=1}^m \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle z_{\hat{k}\hat{k}}^T \frac{\partial a_{i i'}^T}{\partial x_{\hat{k}}} \frac{\partial a_{\hat{i}\hat{i}}^T}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}, (z^T \nabla)_{k f} \rangle_{\mathbb{R}^m}$$

$$-\sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle z_{k\hat{k}}^T a_{i\hat{i}}^T \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \right) \frac{\partial f}{\partial x_{\hat{i}}}, (z^T \nabla)_k f \rangle_{\mathbb{R}^m},$$

and

$$\begin{aligned} \mathfrak{R}_\pi(\nabla f, \nabla f) &= 2 \sum_{k=1}^m \sum_{i=1}^n \sum_{k', \hat{k}, \hat{i}, i'=1}^{n+m} \left[ \frac{\partial}{\partial x_{k'}} z_{k\hat{k}}^T z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{i'}} \right] \\ &\quad + 2 \sum_{k=1}^m \sum_{i=1}^n \sum_{k', \hat{k}, \hat{i}, i'=1}^{n+m} \left[ z_{k\hat{k}}^T \frac{\partial}{\partial x_{k'}} z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{i'}} \right. \\ &\quad \quad \quad + z_{k\hat{k}}^T z_{k\hat{k}}^T \frac{\partial^2}{\partial x_{k'} \partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{i'}} \\ &\quad \quad \quad \left. + z_{k\hat{k}}^T z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} \frac{\partial}{\partial x_{k'}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{i'}} \right] \\ &\quad + 2 \sum_{k=1}^m \sum_{i=1}^n \sum_{\hat{k}, \hat{i}, i'=1}^{n+m} (z^T \nabla \log \pi)_k \left[ z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{i'}} \right] \\ &\quad - 2 \sum_{j=1}^m \sum_{l=1}^n \sum_{l', \hat{l}, \hat{j}, j'=1}^{n+m} \left[ \frac{\partial}{\partial x_{l'}} a_{l\hat{l}}^T a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^T \frac{\partial f}{\partial x_{\hat{j}}} z_{j\hat{j}}^T \frac{\partial f}{\partial x_{j'}} \right] \\ &\quad - 2 \sum_{j=1}^m \sum_{l=1}^n \sum_{l', \hat{l}, \hat{j}, j'=1}^{n+m} \left[ a_{l\hat{l}}^T \frac{\partial}{\partial x_{l'}} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^T \frac{\partial f}{\partial x_{\hat{j}}} z_{j\hat{j}}^T \frac{\partial f}{\partial x_{j'}} \right. \\ &\quad \quad \quad + a_{l\hat{l}}^T a_{l\hat{l}}^T \frac{\partial^2}{\partial x_{l'} \partial x_{\hat{l}}} z_{j\hat{j}}^T \frac{\partial f}{\partial x_{\hat{j}}} z_{j\hat{j}}^T \frac{\partial f}{\partial x_{j'}} \\ &\quad \quad \quad \left. + a_{l\hat{l}}^T a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^T \frac{\partial f}{\partial x_{\hat{j}}} \frac{\partial}{\partial x_{l'}} z_{j\hat{j}}^T \frac{\partial f}{\partial x_{j'}} \right] \\ &\quad - 2 \sum_{j=1}^m \sum_{l=1}^n \sum_{\hat{l}, \hat{j}, j'=1}^{n+m} (a^T \nabla \log \pi)_l \left[ a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^T \frac{\partial f}{\partial x_{\hat{j}}} z_{j\hat{j}}^T \frac{\partial f}{\partial x_{j'}} \right]. \end{aligned}$$

Here, we also denote  $\mathfrak{R} = \mathfrak{R}(x) \in \mathbb{R}^{(n+m) \times (n+m)}$ , such that  $(\nabla f)^T \mathfrak{R}(x) \nabla f = \mathfrak{R}(\nabla f, \nabla f)$ .

The main theorem is presented below, and its proof is postponed to Theorem 3 in Section 5.

**Theorem 1** (Generalized z Bochner's formula). *If Assumption 1 is satisfied, then the following decomposition holds:*

$$\Gamma_2(f, f) + \Gamma_2^{z, \pi}(f, f) = \|\mathfrak{H}_{\text{ess}_{a,z} f}\|^2 + \mathfrak{R}(\nabla f, \nabla f),$$

where we define

$$\begin{aligned} \|\mathfrak{H}_{\text{ess}_{a,z} f}\|^2 &= [X + \Lambda_1]^T Q^T Q [X + \Lambda_1] + [X + \Lambda_2]^T P^T P [X + \Lambda_2], \\ \mathfrak{R}(\nabla f, \nabla f) &= -\Lambda_1^T Q^T Q \Lambda_1 - \Lambda_2^T P^T P \Lambda_2 + D^T D + E^T E \\ &\quad + \mathfrak{R}_{ab}(\nabla f, \nabla f) + \mathfrak{R}_{zb}(\nabla f, \nabla f) + \mathfrak{R}^\pi(\nabla f, \nabla f). \end{aligned}$$

We are now ready to prove the convergence property of the degenerate drift–diffusion process (1) and related functional inequalities. Denote the Kullback–Leibler divergence as

$$D_{\text{KL}}(\rho \| \pi) := \int_{\mathbb{R}^{n+m}} \rho(x) \log \frac{\rho(x)}{\pi(x)} dx.$$

Denote the  $a, z$ -relative Fisher information functional as

$$I_{a,z}(\rho) := \int_{\mathbb{R}^{n+m}} (\nabla \log \frac{\rho}{\pi}, aa^T \nabla \log \frac{\rho}{\pi}) \rho dx + \int_{\mathbb{R}^{n+m}} (\nabla \log \frac{\rho}{\pi}, zz^T \nabla \log \frac{\rho}{\pi}) \rho dx.$$

**Theorem 2** (Exponential convergence in the  $L^1$  distance). *Suppose there exists a constant  $\kappa > 0$  such that*

$$\mathfrak{R} \succeq \kappa(aa^T + zz^T).$$

*Let  $\rho_0$  be a smooth initial distribution and  $\rho = \rho(t, x)$  be the probability density function of (1). Then,  $\rho$  converges to the invariant measure  $\pi$  in the sense of*

$$I_{a,z}(\rho) \leq e^{-2\kappa t} I_{a,z}(\rho_0).$$

*In addition,*

$$\int_{\mathbb{R}^{n+m}} |\rho(t, x) - \pi(x)| dx \leq \sqrt{2D_{\text{KL}}(\rho_0 \| \pi)} e^{-\kappa t}.$$

The proof of Theorem 2 is postponed to Proposition (14).

**Remark 3** (Functional inequalities). *Suppose  $\mathfrak{R} \succeq \kappa(aa^T + zz^T)$  with  $\kappa > 0$ , then the  $z$ -log-Sobolev inequalities hold:*

$$\int_{\mathbb{R}^{n+m}} \rho \log \frac{\rho}{\pi} dx \leq \frac{1}{2\kappa} I_{a,z}(\rho),$$

*for any smooth density function  $\rho$ .*

**Remark 4.** *In the literature [2], the  $\Gamma_{2,z}$  operator is defined by (10), i.e.,  $\Gamma_2^z(f, f) = \frac{1}{2} L \Gamma_1^z(f, f) - \Gamma_1^z(Lf, f)$ . In fact, this definition is under the assumption of  $\Gamma_1(\Gamma_1^z(f, f), f) = \Gamma_1^z(\Gamma_1(f, f), f)$ . This assumption holds true only for the special choice of  $a$  and  $z$ . In the generalized Gamma  $z$  calculus, we introduce a new term (11), which removes the assumption  $\Gamma_1(\Gamma_1^z(f, f), f) = \Gamma_1^z(\Gamma_1(f, f), f)$ . In fact, in the paper, we show that (11) is exactly the new bilinear form behind the assumption in [2] by considering the weak form.*

**Remark 5.** *Following [35] (Assumption 1), we know that, for any  $i \in \{1, \dots, n\}$  and  $k \in \{1, \dots, m\}$ , if*

$$z_k^T \nabla a_i^T \in \text{Span}\{a_1^T, \dots, a_n^T\}, \quad (20)$$

*there exist vectors  $\hat{\Lambda}_1$  and  $\hat{\Lambda}_2$ , such that the Hessian operator associated with the generator of the SDE and the metric  $(aa^T)^\dagger$  could be represented as*

$$\|\mathfrak{Hess}f\|^2 = [QX + \hat{\Lambda}_1]^T [QX + \hat{\Lambda}_1] + [PX + \hat{\Lambda}_2]^T [PX + \hat{\Lambda}_2].$$

*Furthermore, we have the following relation:*

$$\begin{aligned} & [QX + \hat{\Lambda}_1]^T [QX + \hat{\Lambda}_1] + [PX + \hat{\Lambda}_2]^T [PX + \hat{\Lambda}_2] - \hat{\Lambda}_1^T \hat{\Lambda}_1 - \hat{\Lambda}_2^T \hat{\Lambda}_2 \\ &= [X + \Lambda_1]^T Q^T Q [X + \Lambda_1] + [X + \Lambda_2]^T P^T P [X + \Lambda_2] - \Lambda_1^T Q^T Q \Lambda_1 - \Lambda_2^T P^T P \Lambda_2, \end{aligned}$$

*if there exist  $\Lambda_1$  and  $\Lambda_2$  as in Assumption 1 such that*

$$\hat{\Lambda}_1^T = \Lambda_1^T Q^T \quad \text{and} \quad \hat{\Lambda}_2^T = \Lambda_2^T P^T. \quad (21)$$

*Assumption 1 is true if Conditions 20 and 21 hold. See the detailed connections in [35] (Remark 11).*



### 3. Examples

In this section, we consider the following degenerate drift–diffusion process:

$$dX_t = -a(X_t)a(X_t)^\top \nabla V(X_t)dt + \sqrt{2}a(X_t) \circ dB_t, \quad (22)$$

where  $a : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{(n+m) \times n}$  is a matrix-valued function, for  $n, m \in \mathbb{Z}_+$ , and  $V \in C^\infty(\mathbb{R}^{n+m})$  is a smooth potential function. We denote the invariant measure of SDE (22) as  $\pi$ . We further assumed that

$$-aa^\top \nabla V = a \otimes \nabla a + aa^\top \nabla \log \pi.$$

The above assumption holds for the later three examples.

**Remark 6.** For  $V = 0$ , the invariant measure  $\pi$  in the above assumption exists if  $\{a_1, \dots, a_n\}$  forms left-invariant structures on unimodular Lie groups. In this case, the sub-Laplacian is the sum of squares of horizontal vector fields and the invariant measure is also symmetric. Stratonovich SDE (22) defines the horizontal Brownian motion on sub-Riemannian structure  $(\mathbb{R}^{n+m}, \tau, (aa^\top)^\dagger|_\tau)$ , and  $\pi$  is the volume form associated with the horizontal Laplacian. In general, if the Lie group structure is not unimodular, the drift  $b \neq 0$ . See the related studies about the diffusion process on general manifolds in [36–43]. See the related studies on log-Sobolev inequality in [44,45].

**Remark 7.** It is also worth mentioning that many sub-Riemannian manifolds are non-compact. Hence, there may not exist a positive constant  $\kappa$  for both classical  $\Gamma_1$  and  $\Gamma_1^Z$  directions in the non-compact domain. The non-compactness of the domain brings additional difficulties. To prove the associated inequalities in this case, we need to extend the result derived in [46,47]. This is a direction for future work.

**Remark 8.** It is known that the Heisenberg group is an example of Lie groups in quantum mechanics [48]. In future work, we shall investigate the general convergence analysis of SDEs in Lie groups and their connections with quantum SDEs.

#### 3.1. Heisenberg Group

In this subsection, we apply our general theory to the well-known example in sub-Riemannian geometry, which is the Heisenberg group. A related LSI for the horizontal Wiener measure was studied in [46]. Recall briefly that the Heisenberg group  $\mathbb{H}^1$  admits left-invariant vector fields:  $X = \frac{\partial}{\partial x} - \frac{1}{2}y\frac{\partial}{\partial z}$ ,  $Y = \frac{\partial}{\partial y} + \frac{1}{2}x\frac{\partial}{\partial z}$ ,  $Z = \frac{\partial}{\partial z}$ . Here,  $\{X, Y, Z\}$  forms an orthonormal basis for the tangent bundle of  $\mathbb{H}^1$ . In this case,  $\pi = e^{-V}$ . In particular,  $X$  and  $Y$  generate the horizontal distribution  $\tau$ . To fit into our general theory from the previous section, we take matrices  $a$  and  $z$  as below:

$$a^\top = \begin{pmatrix} 1 & 0 & -y/2 \\ 0 & 1 & x/2 \end{pmatrix}, \quad z^\top = (0, 0, 1). \quad (23)$$

In particular, we have

$$a^\top \nabla f = \left( (a^\top \nabla)_1 f, (a^\top \nabla)_2 f \right)^\top, \quad (a^\top \nabla)_1 f = \left( \frac{\partial f}{\partial x} - \frac{y}{2} \frac{\partial f}{\partial z} \right), \quad (a^\top \nabla)_2 f = \left( \frac{\partial f}{\partial y} + \frac{x}{2} \frac{\partial f}{\partial z} \right).$$

We have the following proposition for Heisenberg group following Theorem 1.

**Proposition 1.** For any smooth function  $f \in C^\infty(\mathbb{H}^1)$ , one has

$$\Gamma_2(f, f) + \Gamma_2^{z, \pi}(f, f) = \|\mathfrak{Hess}_{a, z} f\|^2 + \mathfrak{R}(\nabla f, \nabla f),$$

where

$$\begin{aligned}
 \Lambda_1^T &= (0, 0, 0, 0, 0, 0, 0, 0, 0); \\
 \Lambda_2^T &= (0, 0, 0, 0, 0, 0, (a^T \nabla)_2 f, -(a^T \nabla)_1 f, 0); \\
 &\quad \Re_{ab}(\nabla f, \nabla f) - \Lambda_1^T Q^T Q \Lambda_1 - \Lambda_2^T P^T P \Lambda_2 + D^T D + E^T E \\
 &= -\Gamma_1(f, f) + \frac{1}{2} \Gamma_1^z(f, f) - (a^T \nabla)_1 V \partial_z f (a^T \nabla)_2 f \\
 &\quad + (a^T \nabla)_2 V \partial_z f (a^T \nabla)_1 f \\
 &\quad + \left[ \frac{\partial^2 V}{\partial x \partial x} + \frac{y^2}{4} \frac{\partial^2 V}{\partial z \partial z} - y \frac{\partial^2 V}{\partial x \partial z} \right] |(a^T \nabla)_1 f|^2 \\
 &\quad + \left[ \frac{\partial^2 V}{\partial y \partial y} + \frac{x^2}{4} \frac{\partial^2 V}{\partial z \partial z} + x \frac{\partial^2 V}{\partial y \partial z} \right] |(a^T \nabla)_2 f|^2 \\
 &\quad + 2 \left[ \frac{\partial^2 V}{\partial x \partial y} + \frac{x}{2} \frac{\partial^2 V}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 V}{\partial y \partial z} - \frac{xy}{4} \frac{\partial^2 V}{\partial z \partial z} \right] (a^T \nabla)_1 f (a^T \nabla)_2 f; \\
 \Re_{zb}(\nabla f, \nabla f) &= \left( \frac{\partial^2 V}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 V}{\partial z \partial z} \right) (a^T \nabla)_1 f (z^T \nabla)_1 f \\
 &\quad + \left( \frac{\partial^2 V}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 V}{\partial z \partial z} \right) (z^T \nabla)_1 f (a^T \nabla)_2 f; \\
 \Re_{\pi}(\nabla f, \nabla f) &= 0.
 \end{aligned}$$

The proof of Proposition of 1 follows from the proof of Theorem 1 (i.e., Theorem 3) and Lemmas 1–3. The following convergence result follows directly from Theorem 2.

**Proposition 2.** *If there exists  $\kappa > 0$  as shown in Theorem 2, the exponential dissipation result in the  $L^1$  distance holds:*

$$\int |\rho(t, x) - \pi(x)| dx = O(e^{-\kappa t}).$$

We next formulate the curvature tensor into a matrix format. Denote

$$U = \left( (a^T \nabla)_1 f, (a^T \nabla)_2 f, (z^T \nabla)_1 f \right)_{3 \times 1}, \quad (24)$$

and denote  $I_{3 \times 3}$  as the identity matrix. With a little abuse of notation, there exists a symmetric matrix  $\Re$  such that we can represent the tensor as below.

$$\Re(\nabla f, \nabla f) = (U)^T \cdot \Re \cdot U, \quad (25)$$

which implies that

$$\Re \succeq \kappa(aa^T + zz^T) \Rightarrow \Re(\nabla f, \nabla f) \succeq \kappa(\Gamma_1(f, f) + \Gamma_1^z(f, f)).$$

In other words, we need to estimate the smallest eigenvalue of matrix  $\Re$ . We next present the formulation of matrix  $\Re$  for the Heisenberg group as follows.

**Corollary 1.** The matrix  $\mathfrak{R}$  associated with the Heisenberg group has the following form:

$$\begin{aligned}\mathfrak{R}_{11} &= \left[ \frac{\partial^2 V}{\partial x \partial x} + \frac{y^2}{4} \frac{\partial^2 V}{\partial z \partial z} - y \frac{\partial^2 V}{\partial x \partial z} \right] - 1; \\ \mathfrak{R}_{22} &= \left[ \frac{\partial^2 V}{\partial y \partial y} + \frac{x^2}{4} \frac{\partial^2 V}{\partial z \partial z} + x \frac{\partial^2 V}{\partial y \partial z} \right] - 1; \quad \mathfrak{R}_{33} = \frac{1}{2}; \\ \mathfrak{R}_{12} &= \mathfrak{R}_{21} = \left[ \frac{\partial^2 V}{\partial x \partial y} + \frac{x}{2} \frac{\partial^2 V}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 V}{\partial y \partial z} - \frac{xy}{4} \frac{\partial^2 V}{\partial z \partial z} \right]; \\ \mathfrak{R}_{13} &= \mathfrak{R}_{31} = \frac{1}{2} (a^\top \nabla)_2 V + \frac{1}{2} \left( \frac{\partial^2 V}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 V}{\partial z \partial z} \right); \\ \mathfrak{R}_{23} &= \mathfrak{R}_{32} = -\frac{1}{2} (a^\top \nabla)_1 V + \frac{1}{2} \left( \frac{\partial^2 V}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 V}{\partial z \partial z} \right).\end{aligned}$$

**Proof.** The explicit form of matrix  $\mathfrak{R}$  follows from the definition in Theorem 1 and the notation in (24) and (25). We have

$$\begin{aligned}\mathfrak{R}(\nabla f, \nabla f) &= -\Lambda_1^\top Q^\top Q \Lambda_1 - \Lambda_2^\top P^\top P \Lambda_2 + D^\top D + E^\top E \\ &\quad + \mathfrak{R}_{ab}(\nabla f, \nabla f) + \mathfrak{R}_{zb}(\nabla f, \nabla f) + \mathfrak{R}^\pi(\nabla f, \nabla f) \\ &= (U)^\top \cdot \mathfrak{R} \cdot U.\end{aligned}$$

Plugging the explicit representation from Proposition 1 into the above formula and applying matrix symmetrization for the off-diagonal terms, we obtain the desired matrix  $\mathfrak{R}$ .  $\square$

Next, we present the three key lemmas.

**Lemma 1.** For the Heisenberg group, we have

$$\begin{aligned}Q &= \begin{pmatrix} 1 & 0 & -\frac{y}{2} & 0 & 0 & 0 & -\frac{y}{2} & 0 & \frac{y^2}{4} \\ 0 & 1 & \frac{x}{2} & 0 & 0 & 0 & 0 & -\frac{y}{2} & -\frac{xy}{4} \\ 0 & 0 & 0 & 1 & 0 & -\frac{y}{2} & \frac{x}{2} & 0 & -\frac{xy}{4} \\ 0 & 0 & 0 & 0 & 1 & \frac{x}{2} & 0 & \frac{x}{2} & \frac{x^2}{4} \end{pmatrix}; \\ P &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -\frac{y}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \frac{x}{2} \end{pmatrix}; \\ D^\top &= (0, \frac{1}{2} \partial_z f, -\frac{1}{2} \partial_z f, 0); \quad E^\top = (0, 0); \\ F^\top &= G^\top = (0, 0, 0, 0, 0, 0, 0, 0, 0); \\ C^\top &= (0, 0, \frac{x}{4} \partial_z f + \frac{1}{2} \partial_y f, 0, 0, \frac{y}{4} \partial_z f - \frac{1}{2} \partial_x f \\ &\quad , \frac{x}{4} \partial_z f + \frac{1}{2} \partial_y f, \frac{y}{4} \partial_z f - \frac{1}{2} \partial_x f, -\frac{y}{2} \partial_y f - \frac{x}{2} \partial_x f).\end{aligned}$$

**Proof.** The proof of this lemma follows from routine computations. Plugging matrices  $a$  and  $z$  from (23) into Notation 1, we obtain the desired vectors and matrices. We skip the detailed computation here.  $\square$

**Lemma 2.** On  $\mathbb{H}^1$ , vectors  $F$  and  $G$  are zero vectors, and we have

$$\begin{aligned}& [QX + D]^\top [QX + D] + [PX + E]^\top [PX + E] + 2C^\top X \\ &= \|\mathfrak{Hess}_{a,z} f\|^2 - \Lambda_1^\top Q^\top Q \Lambda_1 - \Lambda_2^\top P^\top P \Lambda_2 + D^\top D + E^\top E.\end{aligned}$$

In particular, we have

$$\begin{aligned}\|\mathfrak{H}_{\text{ess}_{a,z}}f\|^2 &= [X + \Lambda_1]^T Q^T Q [X + \Lambda_1] + [X + \Lambda_2]^T P^T P [X + \Lambda_2]; \\ \Lambda_1^T &= (0, 0, 0, 0, 0, 0, 0, 0); \\ \Lambda_2^T &= (0, 0, 0, 0, 0, 0, (a^T \nabla)_2 f, -(a^T \nabla)_1 f, 0); \\ &\quad -\Lambda_1^T Q^T Q \Lambda_1 - \Lambda_2^T P^T P \Lambda_2 + D^T D + E^T E \\ &= -\Gamma_1(f, f) + \frac{1}{2} \Gamma_1^z(f, f).\end{aligned}$$

**Lemma 3.** By routine computations, we obtain

$$\begin{aligned}\mathfrak{R}_{ab}(\nabla f, \nabla f) &= -(a^T \nabla)_1 V \partial_z f (a^T \nabla)_2 f + (a^T \nabla)_2 V \partial_z f (a^T \nabla)_1 f \\ &\quad + \left[ \frac{\partial^2 V}{\partial x \partial x} + \frac{y^2}{4} \frac{\partial^2 V}{\partial z \partial z} - y \frac{\partial^2 V}{\partial x \partial z} \right] |(a^T \nabla)_1 f|^2 \\ &\quad + \left[ \frac{\partial^2 V}{\partial y \partial y} + \frac{x^2}{4} \frac{\partial^2 V}{\partial z \partial z} + x \frac{\partial^2 V}{\partial y \partial z} \right] |(a^T \nabla)_2 f|^2 \\ &\quad + 2 \left[ \frac{\partial^2 V}{\partial x \partial y} + \frac{x}{2} \frac{\partial^2 V}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 V}{\partial y \partial z} - \frac{xy}{4} \frac{\partial^2 V}{\partial z \partial z} \right] (a^T \nabla)_1 f (a^T \nabla)_2 f; \\ \mathfrak{R}_{zb}(\nabla f, \nabla f) &= \left( \frac{\partial^2 V}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 V}{\partial z \partial z} \right) (a^T \nabla)_1 f (z^T \nabla)_1 f \\ &\quad + \left( \frac{\partial^2 V}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 V}{\partial z \partial z} \right) (z^T \nabla)_1 f (a^T \nabla)_2 f; \\ \mathfrak{R}_\pi(\nabla f, \nabla f) &= 0.\end{aligned}$$

**Proof of Lemma 2.** We first have

$$\begin{aligned}2C^T X &= \sum_{i,k=1}^3 2C_{ik}^T X_{ik} \\ &= 2 \left[ \frac{\partial^2 f}{\partial x \partial z} \frac{(a^T \nabla)_2 f}{2} - \frac{\partial^2 f}{\partial y \partial z} \frac{(a^T \nabla)_1 f}{2} + \frac{\partial^2 f}{\partial z \partial x} \frac{(a^T \nabla)_2 f}{2} \right] \\ &\quad - 2 \left[ \frac{\partial^2 f}{\partial z \partial y} \frac{(a^T \nabla)_1 f}{2} + \frac{\partial^2 f}{\partial z \partial z} \left( \frac{y}{2} \partial_y f + \frac{x}{2} \partial_x f \right) \right] \\ &= 2 \frac{\partial^2 f}{\partial x \partial z} (a^T \nabla)_2 f - 2 \frac{\partial^2 f}{\partial y \partial z} (a^T \nabla)_1 f - 2 \frac{\partial^2 f}{\partial z \partial z} \left( \frac{y}{2} \partial_y f + \frac{x}{2} \partial_x f \right) \\ &= 2(a^T \nabla)_2 f \left[ \frac{\partial^2 f}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 f}{\partial z \partial z} \right] - 2(a^T \nabla)_1 f \left[ \frac{\partial^2 f}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 f}{\partial z \partial z} \right].\end{aligned}$$

By direct computations, we have

$$\begin{aligned}&[QX + D]^T [QX + D] + [PX + E]^T [PX + E] + 2C^T X \\ &= \left[ \frac{\partial^2 f}{\partial x \partial x} - y \frac{\partial^2 f}{\partial x \partial z} + \frac{y^2}{4} \frac{\partial^2 f}{\partial z \partial z} \right]^2 + \left[ \frac{\partial^2 f}{\partial x \partial y} + \frac{x}{2} \frac{\partial^2 f}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 f}{\partial y \partial z} - \frac{xy}{4} \frac{\partial^2 f}{\partial z \partial z} + \frac{1}{2} \partial_z f \right]^2 \\ &\quad + \left[ \frac{\partial^2 f}{\partial x \partial y} - \frac{y}{2} \frac{\partial^2 f}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 f}{\partial x \partial z} - \frac{xy}{4} \frac{\partial^2 f}{\partial z \partial z} - \frac{1}{2} \partial_z f \right]^2 + \left[ \frac{\partial^2 f}{\partial y \partial y} + x \frac{\partial^2 f}{\partial y \partial z} + \frac{x^2}{4} \frac{\partial^2 f}{\partial z \partial z} \right]^2 \\ &\quad + \left[ \frac{\partial^2 f}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 f}{\partial z \partial z} \right]^2 + \left[ \frac{\partial^2 f}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 f}{\partial z \partial z} \right]^2 \\ &\quad + 2(a^T \nabla)_2 f \left[ \frac{\partial^2 f}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 f}{\partial z \partial z} \right] - 2(a^T \nabla)_1 f \left[ \frac{\partial^2 f}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 f}{\partial z \partial z} \right].\end{aligned}$$

Completing the squares for the cross terms involving the type of “ $\nabla f \nabla^2 f$ ” and following the reformulation as below:

$$\begin{aligned} & \left[ \frac{\partial^2 f}{\partial x \partial y} + \frac{x}{2} \frac{\partial^2 f}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 f}{\partial y \partial z} - \frac{xy}{4} \frac{\partial^2 f}{\partial z \partial z} + \frac{1}{2} \partial_z f \right]^2 \\ & + \left[ \frac{\partial^2 f}{\partial x \partial y} - \frac{y}{2} \frac{\partial^2 f}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 f}{\partial x \partial z} - \frac{xy}{4} \frac{\partial^2 f}{\partial z \partial z} - \frac{1}{2} \partial_z f \right]^2 \\ & = 2 \left[ \frac{\partial^2 f}{\partial x \partial y} - \frac{y}{2} \frac{\partial^2 f}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 f}{\partial x \partial z} - \frac{xy}{4} \frac{\partial^2 f}{\partial z \partial z} \right]^2 + \frac{1}{2} |\partial_z f|^2, \end{aligned}$$

we have

$$\begin{aligned} & [QX + D]^T [QX + D] + [PX + E]^T [PX + E] + 2C^T X \\ & = \left[ \frac{\partial^2 f}{\partial x \partial x} - y \frac{\partial^2 f}{\partial x \partial z} + \frac{y^2}{4} \frac{\partial^2 f}{\partial z \partial z} \right]^2 + 2 \left[ \frac{\partial^2 f}{\partial x \partial y} - \frac{y}{2} \frac{\partial^2 f}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 f}{\partial x \partial z} - \frac{xy}{4} \frac{\partial^2 f}{\partial z \partial z} \right]^2 \\ & + \left[ \frac{\partial^2 f}{\partial y \partial y} + x \frac{\partial^2 f}{\partial y \partial z} + \frac{x^2}{4} \frac{\partial^2 f}{\partial z \partial z} \right]^2 + \left[ \frac{\partial^2 f}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 f}{\partial z \partial z} + (a^T \nabla)_2 f \right]^2 \\ & + \left[ \frac{\partial^2 f}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 f}{\partial z \partial z} - (a^T \nabla)_1 f \right]^2 - |(a^T \nabla)_2 f|^2 - |(a^T \nabla)_1 f|^2 + \frac{1}{2} |(z^T \nabla)_1 f|^2. \end{aligned}$$

The sum of squares terms give  $\|\text{Hess}_{a,z}\|_{\mathbb{F}}^2$ , hence  $\Lambda_1$  and  $\Lambda_2$ . The remainders generate  $-\Lambda_1^T Q^T Q \Lambda_1 - \Lambda_2^T P^T P \Lambda_2 + D^T D + E^T E$ , which equals  $-\Gamma_1(f, f) + \frac{1}{2} \Gamma_1^z(f, f)$ .  $\square$

We are now left to compute the tensors.

**Proof of Lemma 3.** By direct computation, we have

$$\begin{aligned} \mathfrak{R}_a(\nabla f, \nabla f) &= \sum_{i,k=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{ii'}^T \left( \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} \right), (a^T \nabla)_k f \rangle_{\mathbb{R}^2} \\ &+ \sum_{i,k=2}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{ii'}^T a_{i\hat{i}}^T \left( \frac{\partial}{\partial x_{i'}} \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}} \right) \left( \frac{\partial f}{\partial x_{\hat{k}}} \right), (a^T \nabla)_k f \rangle_{\mathbb{R}^2} \\ &- \sum_{i,k=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{k\hat{k}}^T \frac{\partial a_{ii'}^T}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}, (a^T \nabla)_k f \rangle_{\mathbb{R}^2} \\ &- \sum_{i,k=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{k\hat{k}}^T a_{ii'}^T \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \right) \frac{\partial f}{\partial x_{\hat{i}}}, (a^T \nabla)_k f \rangle_{\mathbb{R}^2}, \\ &= \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4. \end{aligned}$$

For the four terms above, we have

$$\begin{aligned} \mathcal{I}_1 &= \sum_{i=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 a_{ii'}^T \left( \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \frac{\partial a_{1\hat{k}}^T}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^T \nabla)_1 f \\ &+ \sum_{i=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 a_{ii'}^T \left( \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \frac{\partial a_{2\hat{k}}^T}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^T \nabla)_2 f = 0 \end{aligned}$$

$$\begin{aligned}
\mathcal{I}_2 &= \sum_{i=2}^n \sum_{i', \hat{i}, \hat{k}=1}^3 a_{ii'}^\top a_{ii}^\top \left( \frac{\partial}{\partial x_{i'}} \frac{\partial a_{1\hat{k}}^\top}{\partial x_{\hat{i}}} \right) \left( \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^\top \nabla)_1 f \\
&\quad + \sum_{i=2}^n \sum_{i', \hat{i}, \hat{k}=1}^3 a_{ii'}^\top a_{ii}^\top \left( \frac{\partial}{\partial x_{i'}} \frac{\partial a_{2\hat{k}}^\top}{\partial x_{\hat{i}}} \right) \left( \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^\top \nabla)_2 f = 0 \\
\mathcal{I}_3 &= - \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 a_{1\hat{k}}^\top \frac{\partial a_{ii'}^\top}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^\top \nabla)_1 f \\
&\quad - \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 a_{2\hat{k}}^\top \frac{\partial a_{ii'}^\top}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^\top \nabla)_2 f = 0 \\
\mathcal{I}_4 &= - \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 a_{1\hat{k}}^\top a_{ii'}^\top \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \right) \frac{\partial f}{\partial x_{\hat{i}}} (a^\top \nabla)_1 f \\
&\quad - \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 a_{2\hat{k}}^\top a_{ii'}^\top \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \right) \frac{\partial f}{\partial x_{\hat{i}}} (a^\top \nabla)_2 f = 0.
\end{aligned}$$

Similar computation applies to the tensor terms  $\mathfrak{R}_\pi$  and  $\mathfrak{R}_{zb}$ . Since  $z$  is a constant matrix, we obtain

$$\begin{aligned}
\mathfrak{R}_{zb}(\nabla f, \nabla f) &= - \sum_{\hat{i}, \hat{k}=1}^3 \langle (z_{1\hat{i}}^\top \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial z_{1\hat{i}}^\top}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^\top \nabla f)_1 \rangle_{\mathbb{R}}, \\
\mathfrak{R}_\pi &= 0.
\end{aligned}$$

We now compute the tensor terms involving the drift  $b$ . For the drift term in tensor  $\mathfrak{R}_{ab}$ , taking  $b = -aa^\top \nabla V$ , which means  $b = -(a_{\hat{k}\hat{k}} a_{kk'}^\top \frac{\partial V}{\partial x_{k'}})_{\hat{k}=1,2,3}$  in local coordinates,

$$\begin{aligned}
\mathfrak{R}_b^a &= \sum_{i,k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{ii}^\top \frac{\partial a_{k\hat{k}}^\top}{\partial x_{\hat{i}}} a_{kk'}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^\top \nabla)_i f \right] \\
&\quad + \sum_{i,k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{ii}^\top \frac{\partial a_{kk'}^\top}{\partial x_{\hat{i}}} a_{k\hat{k}}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^\top \nabla)_i f \right] \\
&\quad + \sum_{i,k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{ii}^\top a_{k\hat{k}}^\top a_{kk'}^\top \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^\top \nabla)_i f \right] \\
&\quad - \sum_{i,k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{kk}^\top a_{kk'}^\top \frac{\partial a_{ii}^\top}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^\top \nabla)_i f \right] \\
&= \mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3 + \mathcal{J}_4.
\end{aligned}$$

We now derive the explicit formulas for the above four terms.

$$\begin{aligned}
\mathcal{J}_1 &= \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{1\hat{i}}^\top \frac{\partial a_{1\hat{k}}^\top}{\partial x_{\hat{i}}} a_{1k'}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^\top \nabla)_1 f + a_{2\hat{i}}^\top \frac{\partial a_{1\hat{k}}^\top}{\partial x_{\hat{i}}} a_{1k'}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^\top \nabla)_2 f \right] \\
&\quad + \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{1\hat{i}}^\top \frac{\partial a_{2\hat{k}}^\top}{\partial x_{\hat{i}}} a_{2k'}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^\top \nabla)_1 f + a_{2\hat{i}}^\top \frac{\partial a_{2\hat{k}}^\top}{\partial x_{\hat{i}}} a_{2k'}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^\top \nabla)_2 f \right] \\
&= -\frac{1}{2} (a^\top \nabla)_1 V \partial_z f (a^\top \nabla)_2 f + \frac{1}{2} (a^\top \nabla)_2 V \partial_z f (a^\top \nabla)_1 f;
\end{aligned}$$

$$\begin{aligned}
\mathcal{J}_2 &= \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{1\hat{i}}^T \frac{\partial a_{1k'}^T}{\partial x_{\hat{i}}} a_{1\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T \frac{\partial a_{1k'}^T}{\partial x_{\hat{i}}} a_{1\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\
&\quad \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{1\hat{i}}^T \frac{\partial a_{2k'}^T}{\partial x_{\hat{i}}} a_{2\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T \frac{\partial a_{2k'}^T}{\partial x_{\hat{i}}} a_{2\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\
&= -\frac{1}{2} \frac{\partial V}{\partial z} (a^T \nabla)_1 f (a^T \nabla)_2 f + \frac{1}{2} \frac{\partial V}{\partial z} (a^T \nabla)_1 f (a^T \nabla)_2 f = 0; \\
\mathcal{J}_3 &= \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{1\hat{i}}^T a_{1\hat{k}}^T a_{1k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T a_{1\hat{k}}^T a_{1k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\
&\quad \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{1\hat{i}}^T a_{2\hat{k}}^T a_{2k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T a_{2\hat{k}}^T a_{2k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\
&= \sum_{\hat{i}, k'=1}^3 \left[ a_{1\hat{i}}^T a_{1k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} |(a^T \nabla)_1 f|^2 + a_{2\hat{i}}^T a_{1k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} (a^T \nabla)_1 f (a^T \nabla)_2 f \right] \\
&\quad \sum_{\hat{i}, k'=1}^3 \left[ a_{1\hat{i}}^T a_{2k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} (a^T \nabla)_2 f (a^T \nabla)_1 f + a_{2\hat{i}}^T a_{2k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} |(a^T \nabla)_2 f|^2 \right] \\
&= \left[ \frac{\partial^2 V}{\partial x \partial x} + \frac{y^2}{4} \frac{\partial^2 V}{\partial z \partial z} - y \frac{\partial^2 V}{\partial x \partial z} \right] |(a^T \nabla)_1 f|^2 \\
&\quad + \left[ \frac{\partial^2 V}{\partial y \partial y} + \frac{x^2}{4} \frac{\partial^2 V}{\partial z \partial z} + x \frac{\partial^2 V}{\partial y \partial z} \right] |(a^T \nabla)_2 f|^2 \\
&\quad + 2 \left[ \frac{\partial^2 V}{\partial x \partial y} + \frac{x}{2} \frac{\partial^2 V}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 V}{\partial y \partial z} - \frac{xy}{4} \frac{\partial^2 V}{\partial z \partial z} \right] (a^T \nabla)_1 f (a^T \nabla)_2 f; \\
\mathcal{J}_4 &= - \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{1\hat{k}}^T a_{1k'}^T \frac{\partial a_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_1 f + a_{1\hat{k}}^T a_{1k'}^T \frac{\partial a_{2\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_2 f \right] \\
&\quad - \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{2\hat{k}}^T a_{2k'}^T \frac{\partial a_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_1 f + a_{2\hat{k}}^T a_{2k'}^T \frac{\partial a_{2\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_2 f \right] \\
&= -\frac{1}{2} (a^T \nabla)_1 V \partial_z f (a^T \nabla)_2 f + \frac{1}{2} (a^T \nabla)_2 V \partial_z f (a^T \nabla)_1 f.
\end{aligned}$$

Summing up the above formulas, we obtain  $\mathfrak{R}_{ab}$ . We now compute the drift tensor term of  $\mathfrak{R}_{zb}$ . By taking  $b = -aa^T \nabla V$ , we have

$$\begin{aligned}
\mathfrak{R}_b^z(\nabla f, \nabla f) &= - \sum_{\hat{i}, \hat{k}=1}^3 \left[ z_{1\hat{i}}^T \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla f)_1 - b_{\hat{k}} \frac{\partial z_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}} (z^T \nabla f)_1 \right] \\
&= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^T \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}} a_{k\hat{k}'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f \right] \\
&\quad + \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^T \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}} a_{k\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f \right] \\
&\quad + \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^T a_{k\hat{k}}^T a_{k\hat{k}'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f \right] \\
&\quad - \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{k\hat{k}}^T a_{k\hat{k}'}^T \frac{\partial z_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (z^T \nabla)_1 f \right] \\
&= \mathcal{J}_1^z + \mathcal{J}_2^z + \mathcal{J}_3^z + \mathcal{J}_4^z.
\end{aligned}$$

We further compute as below by taking advantage of the constant matrix  $z$ :

$$\begin{aligned}\mathcal{J}_1^z &= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^\top \frac{\partial a_{k\hat{k}}^\top}{\partial x_{\hat{i}}} a_{kk'}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f \right] = 0; \\ \mathcal{J}_2^z &= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^\top \frac{\partial a_{k\hat{k}}^\top}{\partial x_{\hat{i}}} a_{k\hat{k}}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f \right] = 0; \\ \mathcal{J}_4^z &= - \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{k\hat{k}}^\top a_{kk'}^\top \frac{\partial z_{1\hat{i}}^\top}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (z^\top \nabla)_1 f \right] = 0 \\ \mathcal{J}_3^z &= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^\top a_{k\hat{k}}^\top a_{kk'}^\top \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f \right] \\ &= \left( \frac{\partial^2 V}{\partial x \partial z} - \frac{y}{2} \frac{\partial^2 V}{\partial z \partial z} \right) (a^\top \nabla)_1 f (z^\top \nabla)_1 f + \left( \frac{\partial^2 V}{\partial y \partial z} + \frac{x}{2} \frac{\partial^2 V}{\partial z \partial z} \right) (z^\top \nabla)_1 f (a^\top \nabla)_2 f.\end{aligned}$$

The proof is thus completed.  $\square$

### 3.2. Displacement Group

In this section, we derive the generalized curvature dimension bound for the displacement group, which is one example of three-dimensional solvable Lie groups. We adapted the general setting from [49] below. Denote  $\mathfrak{g}$  as the three-dimensional solvable Lie algebra, and denote  $H \subset \mathfrak{g}$  as the horizontal subspace satisfying Hörmander's condition, then for a given inner product  $\langle \cdot, \cdot \rangle$  on  $H$ , there exists a canonical basis  $\{X, Y, Z\}$  for  $(\mathfrak{g}, H, \langle \cdot, \cdot \rangle)$ , such that  $\{X, Y\}$  forms an orthonormal basis for  $H$  and satisfies the following Lie-bracket-generating condition for parameters  $\alpha$  and  $\beta \geq 0$ :

$$[X, Y] = Z, \quad [X, Z] = \alpha Y + \beta Z, \quad [Y, Z] = 0.$$

When the parameters  $\alpha = 0$  and  $\beta \neq 0$ , the Lie algebra  $\mathfrak{g}$  has a faithful representation. In particular, it was shown in [49] that the elements of  $\mathfrak{g}$ , in local coordinates  $(\theta, x, y)$ , correspond to the following left-invariant differential operators:

$$X = \frac{\partial}{\partial \theta}, \quad Y = e^{\beta \theta} \frac{\partial}{\partial x} + \frac{\partial}{\partial y}, \quad R = -\beta \frac{\partial}{\partial y},$$

with the following relation:

$$[X, Y] = \beta Y + R, \quad [X, R] = 0, \quad [Y, R] = 0.$$

In terms of local coordinates  $(\theta, x, y)$ , we have

$$X = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 \\ e^{\beta \theta} \\ 1 \end{pmatrix}, \quad R = \begin{pmatrix} 0 \\ 0 \\ -\beta \end{pmatrix}.$$

The corresponding Lie group of this special Lie algebra  $\mathfrak{g}$  is called the displacement group, denoted as  $G$ . We chose  $\{X, Y\}$  as the horizontal orthonormal basis for subalgebra  $H$ . To fit into the general framework from the previous section, we take

$$a = (X, Y) = \begin{pmatrix} 1 & 0 \\ 0 & e^{\beta \theta} \\ 0 & 1 \end{pmatrix}, \quad a^\top = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{\beta \theta} & 1 \end{pmatrix}, \quad z^\top = (0 \quad 0 \quad -g(\theta, x, y)), \quad (26)$$

with  $g(\theta, x, y) \neq 0$ . Our focus here is to derive the curvature tensor in terms of  $\pi = \frac{1}{2}e^{-V}$ . We then used  $(aa^\top)_{|H}^\dagger$  as the horizontal metric on  $H$ . Thus, the sub-Riemannian structure



is given by  $(G, H, (aa^T)_{|H}^+)$ . By direct computations, it is easy to show that, for general smooth function  $f$ ,  $\Gamma_1(f, \Gamma_1^z(f, f)) \neq \Gamma_1^z(f, \Gamma_1(f, f))$ . Hence, the classical Gamma  $z$  calculus proposed in [2] can not be extended for this case to derive the zLSI. Thus, we need to compute vector  $G$  and the tensor term  $\Re_\pi$ . Following Theorem 1, we have the following  $z$ -Bochner's formula for  $G$ .

**Proposition 3.** For any smooth function  $f \in C^\infty(G)$ , one has

$$\Gamma_2(f, f) + \Gamma_2^{z, \pi}(f, f) = \|\text{Hess}_{a,z} f\|^2 + \Re(\nabla f, \nabla f),$$

where

$$\begin{aligned}\Lambda_1^T &= (0, \beta \partial_x f, \frac{\beta \partial_y f}{2}, \beta \partial_x f, 0, 0, \frac{\beta \partial_y f}{2}, 0, -\beta \partial_\theta f); \\ \Lambda_2^T &= (0, 0, 0, 0, 0, 0, \lambda_6, 0, \lambda_9); \\ \lambda_6 &= \frac{\partial_\theta g \partial_y f}{g} - \frac{\beta (a^T \nabla)_2 f}{g^2} - \frac{\partial_\theta g \partial_y f}{g}; \\ \lambda_9 &= \frac{(a^T \nabla)_2 g \partial_y f}{g} + \frac{\beta \partial_\theta f}{g^2} - \frac{(a^T \nabla)_2 g \partial_y f}{g};\end{aligned}$$

and

$$\begin{aligned}& \Re_{ab}(\nabla f, \nabla f) - \Lambda_1^T Q^T Q \Lambda_1 - \Lambda_2^T P^T P \Lambda_2 + D^T D + E^T E \\&= \Gamma_1(\log g, \log g) \Gamma_1^z(f, f) - \beta^2 (1 + \frac{1}{g^2}) \Gamma_1(f, f) + \frac{\beta^2}{2g^2} \Gamma_1^z(f, f) \\&+ \beta^2 e^{\beta\theta} \frac{\partial f}{\partial x} (a^T \nabla)_2 f + \beta e^{\beta\theta} (a^T \nabla)_2 V \frac{\partial f}{\partial x} (a^T \nabla)_1 f \\&+ \beta e^{\beta\theta} \frac{\partial V}{\partial x} (a^T \nabla)_2 f (a^T \nabla)_1 f \\&+ \frac{\partial^2 V}{\partial \theta \partial \theta} |(a^T \nabla)_1 f|^2 + 2(e^{\beta\theta} \frac{\partial^2 V}{\partial \theta \partial x} + \frac{\partial^2 V}{\partial \theta \partial y}) (a^T \nabla)_1 f (a^T \nabla)_2 f \\&+ \sum_{i,k'=1}^3 a_{2i}^T a_{2k'}^T \frac{\partial^2 V}{\partial x_i \partial x_{k'}} |(a^T \nabla)_2 f|^2 - \beta e^{\beta\theta} (a^T \nabla)_1 V \frac{\partial f}{\partial x} (a^T \nabla)_2 f; \\& \Re_{zb}(\nabla f, \nabla f) \\&= \sum_{i=1}^2 \sum_{i', \hat{i}=1}^3 a_{i i'}^T a_{i \hat{i}}^T \frac{\partial^2 z_{13}^T}{\partial x_{i'} \partial x_{\hat{i}}} \partial_y f (z^T \nabla)_1 f - \sum_{k=1}^2 (a^T \nabla)_k z_{13}^T (a^T \nabla)_k V \partial_y f (z^T \nabla)_1 f \\&- g \frac{\partial^2 V}{\partial \theta \partial y} (a^T \nabla)_1 f (z^T \nabla)_1 f - g(e^{\beta\theta} \frac{\partial^2 V}{\partial x \partial y} + \frac{\partial^2 V}{\partial y \partial y}) (a^T \nabla)_2 f (z^T \nabla)_1 f; \\& \Re_\pi(\nabla f, \nabla f) \\&= -2 \sum_{l=1}^2 \sum_{l', \hat{l}=1}^3 a_{l l'}^T a_{l \hat{l}}^T \frac{\partial^2 z_{13}^T}{\partial x_{l'} \partial x_{\hat{l}}} \partial_y f (z^T \nabla)_1 f \\&- 2\Gamma_1(\log \pi, \log g) |(z^T \nabla)_1 f|^2 - 2\Gamma_1(\log g, \log g) |(z^T \nabla)_1 f|^2.\end{aligned}$$

In particular, we have

$$\begin{aligned}& \sum_{i,k'=1}^3 a_{2i}^T a_{2k'}^T \frac{\partial^2 V}{\partial x_i \partial x_{k'}} |(a^T \nabla)_2 f|^2 \\&= \left[ e^{2\beta\theta} \frac{\partial^2 V}{\partial x \partial x} + 2e^{\beta\theta} \frac{\partial^2 V}{\partial x \partial y} + \frac{\partial^2 V}{\partial y \partial y} \right] |(a^T \nabla)_2 f|^2;\end{aligned}$$

$$\begin{aligned} & \sum_{i=1}^2 \sum_{i', \hat{i}=1}^3 a_{ii'}^T a_{ii}^T \frac{\partial^2 z_{13}^T}{\partial x_{i'} \partial x_i} \partial_y f (z^T \nabla)_1 f \\ &= \left[ \frac{\partial^2 g}{\partial \theta \partial \theta} + e^{2\beta\theta} \frac{\partial^2 g}{\partial x \partial x} + \frac{\partial^2 g}{\partial y \partial y} + 2e^{\beta\theta} \frac{\partial^2 g}{\partial x \partial y} \right] \frac{|(z^T \nabla)_1 f|^2}{g}. \end{aligned}$$

The proof of Proposition 3 follows from the proof of Theorem 1 (i.e., Theorem 3) and Lemmas 4–6 below. The following convergence result follows directly from Theorem 2.

**Proposition 4.** *If there exists  $\kappa > 0$  as shown in Theorem 2, the exponential dissipation result in the  $L^1$  distance holds:*

$$\int |\rho(t, x) - \pi(x)| dx = O(e^{-\kappa t}).$$

Similarly, we formulated the curvature tensor into a matrix format of  $\mathfrak{R}$ . Using the fact  $e^{\beta\theta} \frac{\partial f}{\partial x} = (a^T \nabla)_2 f + \frac{1}{g} (z^T \nabla f)_1 f$ , we have the following representation.

**Corollary 2.** *The matrix  $\mathfrak{R}$  associated with  $G$  has the following representation:*

$$\begin{aligned} \mathfrak{R}_{11} &= \frac{\partial^2 V}{\partial \theta \partial \theta} - \beta^2 \left(1 + \frac{1}{g^2}\right); \\ \mathfrak{R}_{22} &= \left[ e^{2\beta\theta} \frac{\partial^2 V}{\partial x \partial x} + 2e^{\beta\theta} \frac{\partial^2 V}{\partial x \partial y} + \frac{\partial^2 V}{\partial y \partial y} \right] - \frac{\beta^2}{g^2} - \beta (a^T \nabla)_1 V; \\ \mathfrak{R}_{33} &= \frac{\beta^2}{2g^2} - \Gamma_1(\log g, \log g) - 2\Gamma_1(\log \pi, \log g) - \Gamma_1(\log g, V) \\ &\quad - \frac{1}{g} \left[ \frac{\partial^2 g}{\partial \theta \partial \theta} + e^{2\beta\theta} \frac{\partial^2 g}{\partial x \partial x} + \frac{\partial^2 g}{\partial y \partial y} + 2e^{\beta\theta} \frac{\partial^2 g}{\partial x \partial y} \right]; \\ \mathfrak{R}_{12} &= \mathfrak{R}_{21} = \frac{1}{2} \left( \beta e^{\beta\theta} \frac{\partial V}{\partial x} + 2(e^{\beta\theta} \frac{\partial^2 V}{\partial \theta \partial x} + \frac{\partial^2 V}{\partial \theta \partial y}) + \beta (a^T \nabla)_2 V \right); \\ \mathfrak{R}_{13} &= \mathfrak{R}_{31} = \frac{1}{2} \left( \frac{\beta}{g} (a^T \nabla)_2 V - g \frac{\partial^2 V}{\partial \theta \partial y} \right); \\ \mathfrak{R}_{23} &= \mathfrak{R}_{32} = -\frac{1}{2} \left( \frac{\beta}{g} (a^T \nabla)_1 - \frac{\beta^2}{g} \right) - \frac{1}{2} g \left( e^{\beta\theta} \frac{\partial^2 V}{\partial x \partial y} + \frac{\partial^2 V}{\partial y \partial y} \right). \end{aligned}$$

**Proof.** The derivation for the explicit form of matrix  $\mathfrak{R}$  follows from a similar equivalent representation as shown in the proof of Corollary 1 and the explicit bilinear terms derived in Proposition 3.  $\square$

**Remark 9.** *By taking  $g(\theta, x, y) = \beta$  as a constant, Proposition 3 reduces to a simple version; in particular, the tensors reduce to be*

$$\begin{aligned} & \mathfrak{R}_{ab}(\nabla f, \nabla f) - \Lambda_1^T Q^T Q \Lambda_1 - \Lambda_2^T P^T P \Lambda_2 + D^T D + E^T E \\ &= -(1 + \beta^2) \Gamma_1(f, f) + \frac{1}{2} \Gamma_1^z(f, f) \\ &\quad + \beta^2 e^{\beta\theta} \frac{\partial f}{\partial x} (a^T \nabla)_2 f + \beta e^{\beta\theta} (a^T \nabla)_2 V \frac{\partial f}{\partial x} (a^T \nabla)_1 f + \beta e^{\beta\theta} \frac{\partial V}{\partial x} (a^T \nabla)_2 f (a^T \nabla)_1 f \\ &\quad + \frac{\partial^2 V}{\partial \theta \partial \theta} |(a^T \nabla)_1 f|^2 + 2(e^{\beta\theta} \frac{\partial^2 V}{\partial \theta \partial x} + \frac{\partial^2 V}{\partial \theta \partial y}) (a^T \nabla)_1 f (a^T \nabla)_2 f \\ &\quad + \sum_{\hat{i}, k'=1}^3 a_{2\hat{i}}^T a_{2k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} |(a^T \nabla)_2 f|^2 - \beta e^{\beta\theta} (a^T \nabla)_1 V \frac{\partial f}{\partial x} (a^T \nabla)_2 f; \end{aligned}$$

$$\begin{aligned}
\Re_{zb}(\nabla f, \nabla f) &= -\beta \frac{\partial^2 V}{\partial \theta \partial y} (a^\top \nabla)_1 f (z^\top \nabla)_1 f \\
&\quad - \beta (e^{\beta \theta} \frac{\partial^2 V}{\partial x \partial y} + \frac{\partial^2 V}{\partial y \partial y}) (a^\top \nabla)_2 f (z^\top \nabla)_1 f; \\
\Re_\pi(\nabla f, \nabla f) &= 0.
\end{aligned}$$

Next, we present the following three key lemmas.

**Lemma 4.** For displacement group  $G$ , we have

$$\begin{aligned}
Q &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{\beta \theta} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{\beta \theta} & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & e^{2\beta \theta} & e^{\beta \theta} & 0 & e^{\beta \theta} & 1 \end{pmatrix}; \\
P &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & -g(\theta, x, y) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -g(\theta, x, y)e^{\beta \theta} & -g(\theta, x, y) \end{pmatrix}; \\
D^\top &= (0, \beta e^{\beta \theta} \partial_x f, 0, 0), \quad E^\top = (-\partial_y f \partial_\theta g, -\partial_y f \partial_y g - e^{\beta \theta} \partial_y f \partial_x g); \\
C^\top &= (0, \beta e^{\beta \theta} \partial_y f + \beta e^{2\beta \theta} \partial_x f, 0, 0, -\beta e^{2\beta \theta} \partial_\theta f, -\beta e^{\beta \theta} \partial_\theta f, 0, 0, 0). \\
F &= \begin{pmatrix} 0 \\ 0 \\ g \partial_\theta g \partial_y f \\ 0 \\ 0 \\ e^{\beta \theta} g \partial_y f \partial_y g + e^{2\beta \theta} g \partial_y f \partial_x g \\ 0 \\ 0 \\ g \partial_y f \partial_y g + e^{\beta \theta} g \partial_y f \partial_x g \end{pmatrix}, \quad G = \begin{pmatrix} 0 \\ 0 \\ -2g \partial_y f \partial_\theta g \\ 0 \\ 0 \\ -2e^{\beta \theta} g \partial_y f \partial_y g - 2e^{2\beta \theta} g \partial_y f \partial_x g \\ 0 \\ 0 \\ -2g \partial_y f \partial_y g - 2e^{\beta \theta} g \partial_y f \partial_x g \end{pmatrix}.
\end{aligned}$$

**Proof.** The proof follows by plugging matrices  $a$  and  $z$  from (26) into Notation 1.  $\square$

**Lemma 5.** On displacement group  $G$ , we have

$$\begin{aligned}
&[QX + D]^\top [QX + D] + [PX + E]^\top [PX + E] + 2[C^\top + F^\top + G^\top]X \\
&= \|\mathfrak{Hess}_{a,z} f\|^2 - \Lambda_1^\top Q^\top Q \Lambda_1 - \Lambda_2^\top P^\top P \Lambda_2 + D^\top D + E^\top E.
\end{aligned}$$

In particular, we have

$$\begin{aligned}
\|\mathfrak{Hess}_{a,z} f\|^2 &= [X + \Lambda_1]^\top Q^\top Q [X + \Lambda_1] + [X + \Lambda_2]^\top P^\top P [X + \Lambda_2]; \\
\Lambda_1^\top &= (0, \beta \partial_x f, \frac{\beta \partial_y f}{2}, \beta \partial_x f, 0, 0, \frac{\beta \partial_y f}{2}, 0, -\beta \partial_\theta f); \\
\Lambda_2^\top &= (0, 0, 0, 0, 0, 0, \lambda_6, 0, \lambda_9); \\
\lambda_6 &= \frac{\partial_\theta g \partial_y f}{g} - \frac{\beta (a^\top \nabla)_2 f}{g^2} - \frac{\partial_\theta g \partial_y f}{g}; \\
\lambda_9 &= \frac{(a^\top \nabla)_2 g \partial_y f}{g} + \frac{\beta \partial_\theta f}{g^2} - \frac{(a^\top \nabla)_2 g \partial_y f}{g} \\
&\quad - \Lambda_1^\top Q^\top Q \Lambda_1 - \Lambda_2^\top P^\top P \Lambda_2 + D^\top D + E^\top E \\
&= \Gamma_1(\log g, \log g) \Gamma_1^z(f, f) - \beta^2 (1 + \frac{1}{g^2}) \Gamma_1(f, f) + \frac{\beta^2}{2g^2} \Gamma_1^z(f, f).
\end{aligned}$$

**Lemma 6.** By routine computations, we obtain

$$\begin{aligned}
& \Re_{ab}(\nabla f, \nabla f) \\
&= \beta^2 e^{\beta\theta} \frac{\partial f}{\partial x} (a^\top \nabla)_2 f + \beta e^{\beta\theta} (a^\top \nabla)_2 V \frac{\partial f}{\partial x} (a^\top \nabla)_1 f + \beta e^{\beta\theta} \frac{\partial V}{\partial x} (a^\top \nabla)_2 f (a^\top \nabla)_1 f \\
&\quad + \frac{\partial^2 V}{\partial \theta \partial \theta} |(a^\top \nabla)_1 f|^2 + 2(e^{\beta\theta} \frac{\partial^2 V}{\partial \theta \partial x} + \frac{\partial^2 V}{\partial \theta \partial y}) (a^\top \nabla)_1 f (a^\top \nabla)_2 f \\
&\quad + \sum_{\hat{i}, \hat{k}'=1}^3 a_{2\hat{i}}^\top a_{2\hat{k}'}^\top \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{\hat{k}'}} |(a^\top \nabla)_2 f|^2 - \beta e^{\beta\theta} (a^\top \nabla)_1 V \frac{\partial f}{\partial x} (a^\top \nabla)_2 f; \\
& \Re_{zb}(\nabla f, \nabla f) \\
&= \sum_{i=1}^2 \sum_{i', \hat{i}=1}^3 a_{ii'}^\top a_{i\hat{i}}^\top \frac{\partial^2 z_{1\hat{k}}^\top}{\partial x_{i'} \partial x_{\hat{i}}} \partial_y f (z^\top \nabla)_1 f - \sum_{k=1}^2 (a^\top \nabla)_k z_{13}^\top (a^\top \nabla)_k V \partial_y f (z^\top \nabla)_1 f \\
&\quad - g \frac{\partial^2 V}{\partial \theta \partial y} (a^\top \nabla)_1 f (z^\top \nabla)_1 f - g(e^{\beta\theta} \frac{\partial^2 V}{\partial x \partial y} + \frac{\partial^2 V}{\partial y \partial y}) (a^\top \nabla)_2 f (z^\top \nabla)_1 f; \\
& \Re_{\pi}(\nabla f, \nabla f) \\
&= -2 \sum_{l=1}^2 \sum_{l', \hat{l}=1}^3 a_{ll'}^\top a_{l\hat{l}}^\top \frac{\partial^2 z_{13}^\top}{\partial x_{l'} \partial x_{\hat{l}}} \partial_y f (z^\top \nabla)_1 f - 2 \sum_{l=1}^2 \sum_{l', \hat{l}=1}^3 a_{ll'}^\top a_{l\hat{l}}^\top \frac{\partial z_{13}^\top}{\partial x_{\hat{l}}} \frac{\partial z_{13}^\top}{\partial x_{l'}} |\partial_y f|^2 \\
&\quad - 2 \sum_{l=1}^2 \sum_{\hat{l}=1}^3 (a^\top \nabla)_l \log \pi a_{l\hat{l}}^\top \frac{\partial z_{13}^\top}{\partial x_{\hat{l}}} \partial_y f (z^\top \nabla)_1 f.
\end{aligned}$$

**Proof of Lemma 5.** According to Lemma 4 and observing the fact that  $G = -2F$  and  $(a^\top \nabla)_2 f = e^{\beta\theta} \partial_x f + \partial_y f$ , we first have

$$\begin{aligned}
2C^\top X &= 2[\beta e^{\beta\theta} \partial_y f + \beta e^{2\beta\theta} \partial_x f] \frac{\partial^2 f}{\partial \theta \partial x} + 2[-\beta e^{2\beta\theta} \partial_\theta f] \frac{\partial^2 f}{\partial x \partial x} + 2[-\beta e^{\beta\theta} \partial_\theta f] \frac{\partial^2 f}{\partial x \partial y}; \\
2[F^\top + G^\top]X &= -2\left(g \partial_\theta g \partial_y f \frac{\partial^2 f}{\partial \theta \partial y} + e^{\beta\theta} g (a^\top \nabla)_2 g \partial_y f \frac{\partial^2 f}{\partial x \partial y} + g (a^\top \nabla)_2 g \partial_y f \frac{\partial^2 f}{\partial y \partial y}\right).
\end{aligned}$$

By direct computations, we have

$$\begin{aligned}
& [QX + D]^\top [QX + D] + [PX + E]^\top [PX + E] + 2C^\top X + 2F^\top X + 2G^\top X \\
&= \left[\frac{\partial^2 f}{\partial \theta \partial \theta}\right]^2 + \left[e^{2\beta\theta} \frac{\partial^2 f}{\partial x \partial x} + 2e^{\beta\theta} \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y \partial y}\right]^2 + \left[e^{\beta\theta} \frac{\partial^2 f}{\partial \theta \partial x} + \frac{\partial^2 f}{\partial \theta \partial y} + \beta e^{\beta\theta} \frac{\partial f}{\partial x}\right]^2 \\
&\quad + \left[e^{\beta\theta} \frac{\partial^2 f}{\partial \theta \partial x} + \frac{\partial^2 f}{\partial \theta \partial y}\right]^2 + \left[-g \frac{\partial^2 f}{\partial \theta \partial y} - \partial_y f \partial_\theta g\right]^2 \\
&\quad + \left[-g e^{\beta\theta} \frac{\partial^2 f}{\partial x \partial y} - g \frac{\partial^2 f}{\partial y \partial y} - (a^\top \nabla)_2 g \partial_y f\right]^2 \\
&\quad + 2[\beta e^{\beta\theta} \partial_y f + \beta e^{2\beta\theta} \partial_x f] \frac{\partial^2 f}{\partial \theta \partial x} + 2[-\beta e^{2\beta\theta} \partial_\theta f] \frac{\partial^2 f}{\partial x \partial x} + 2[-\beta e^{\beta\theta} \partial_\theta f] \frac{\partial^2 f}{\partial x \partial y} \\
&\quad - 2\left(g \partial_\theta g \partial_y f \frac{\partial^2 f}{\partial \theta \partial y} + e^{\beta\theta} g (a^\top \nabla)_2 g \partial_y f \frac{\partial^2 f}{\partial x \partial y} + g (a^\top \nabla)_2 g \partial_y f \frac{\partial^2 f}{\partial y \partial y}\right) \\
&= \left[\frac{\partial^2 f}{\partial \theta \partial \theta}\right]^2 + \left[e^{2\beta\theta} \frac{\partial^2 f}{\partial x \partial x} + 2e^{\beta\theta} \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y \partial y}\right]^2 + \left[e^{\beta\theta} \frac{\partial^2 f}{\partial \theta \partial x} + \frac{\partial^2 f}{\partial \theta \partial y} + \beta e^{\beta\theta} \frac{\partial f}{\partial x}\right]^2
\end{aligned}$$

$$\begin{aligned}
& + \left[ e^{\beta\theta} \frac{\partial^2 f}{\partial\theta\partial x} + \frac{\partial^2 f}{\partial\theta\partial y} \right]^2 + \left[ -g \frac{\partial^2 f}{\partial\theta\partial y} - \partial_y f \partial_\theta g \right]^2 \\
& + \left[ -g e^{\beta\theta} \frac{\partial^2 f}{\partial x \partial y} - g \frac{\partial^2 f}{\partial y \partial y} - (a^\top \nabla)_2 g \partial_y f \right]^2 \\
& + 2\beta(a^\top \nabla)_2 f \left[ e^{\beta\theta} \frac{\partial^2 f}{\partial\theta\partial x} + \frac{\partial^2 f}{\partial\theta\partial y} \right] - 2\beta(a^\top \nabla)_2 f \frac{\partial^2 f}{\partial\theta\partial y} - 2g \partial_\theta g \partial_y f \frac{\partial^2 f}{\partial\theta\partial y} \\
& - 2\beta \partial_\theta f \left[ 2e^{\beta\theta} \frac{\partial^2 f}{\partial x \partial y} + e^{2\beta\theta} \frac{\partial^2 f}{\partial x \partial x} + \frac{\partial^2 f}{\partial y \partial y} \right] + 2\beta \partial_\theta f \left[ e^{\beta\theta} \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y \partial y} \right] \\
& - 2g(a^\top \nabla)_2 g \partial_y f \left[ e^{\beta\theta} \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y \partial y} \right].
\end{aligned}$$

Completing the squares for the above terms, we have

$$\begin{aligned}
& [QX + D]^\top [QX + D] + [PX + E]^\top [PX + E] + 2C^\top X + 2F^\top X + 2G^\top X \\
= & \left[ \frac{\partial^2 f}{\partial\theta\partial\theta} \right]^2 + \left[ e^{2\beta\theta} \frac{\partial^2 f}{\partial x \partial x} + 2e^{\beta\theta} \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y \partial y} - \beta \partial_\theta f \right]^2 - \beta^2 |\partial_\theta f|^2 \\
& + \left[ e^{\beta\theta} \frac{\partial^2 f}{\partial\theta\partial x} + \frac{\partial^2 f}{\partial\theta\partial y} + \beta e^{\beta\theta} \frac{\partial f}{\partial x} \right]^2 + \left[ e^{\beta\theta} \frac{\partial^2 f}{\partial\theta\partial x} + \frac{\partial^2 f}{\partial\theta\partial y} + \beta(a^\top \nabla)_2 f \right]^2 - \beta^2 |(a^\top \nabla)_2 f|^2 \\
& + \left[ g \frac{\partial^2 f}{\partial\theta\partial y} + \partial_\theta g \partial_y f - \frac{\beta(a^\top \nabla)_2 f}{g} - \partial_\theta g \partial_y f \right]^2 - \left[ \frac{\beta(a^\top \nabla)_2 f}{g} + \partial_\theta g \partial_y f \right]^2 \\
& + \left[ g e^{\beta\theta} \frac{\partial^2 f}{\partial x \partial y} + g \frac{\partial^2 f}{\partial y \partial y} + (a^\top \nabla)_2 g \partial_y f + \frac{\beta \partial_\theta f}{g} - (a^\top \nabla)_2 g \partial_y f \right]^2 \\
& - \left[ \frac{\beta \partial_\theta f}{g} - (a^\top \nabla)_2 g \partial_y f \right]^2 + 2 \left[ \frac{\beta(a^\top \nabla)_2 f}{g} + \partial_\theta g \partial_y f \right] \partial_\theta g \partial_y f \\
& - 2\partial_y f (a^\top \nabla)_2 g \times \left[ \frac{\beta \partial_\theta f}{g} - (a^\top \nabla)_2 g \partial_y f \right].
\end{aligned}$$

The first-order terms generate  $-\Lambda_1^\top Q^\top Q \Lambda_1 - \Lambda_2^\top P^\top P \Lambda_2 + D^\top D + E^\top E$ , and the sum of squares terms generate vectors  $\Lambda_1$  and  $\Lambda_2$ . We further formulate the above two terms as below:

$$\begin{aligned}
& \left[ e^{\beta\theta} \frac{\partial^2 f}{\partial\theta\partial x} + \frac{\partial^2 f}{\partial\theta\partial y} + \beta e^{\beta\theta} \frac{\partial f}{\partial x} \right]^2 + \left[ e^{\beta\theta} \frac{\partial^2 f}{\partial\theta\partial x} + \frac{\partial^2 f}{\partial\theta\partial y} + \beta(a^\top \nabla)_2 f \right]^2 \\
= & 2 \left[ e^{\beta\theta} \frac{\partial^2 f}{\partial\theta\partial x} + \frac{\partial^2 f}{\partial\theta\partial y} + \beta e^{\beta\theta} \frac{\partial f}{\partial x} + \frac{\beta}{2} \partial_y f \right]^2 + \frac{\beta^2}{2} |\partial_y f|^2.
\end{aligned}$$

Adding  $\frac{\beta^2}{2} |\partial_y f|^2$  into the term  $-\Lambda_1^\top Q^\top Q \Lambda_1 - \Lambda_2^\top P^\top P \Lambda_2 + D^\top D + E^\top E$  again, we further expand as below:

$$\begin{aligned}
& -\Lambda_1^\top Q^\top Q \Lambda_1 - \Lambda_2^\top P^\top P \Lambda_2 + D^\top D + E^\top E \\
= & -\beta^2 [|\partial_\theta f|^2 + |(a^\top \nabla)_2 f|^2] - \left[ \frac{\beta \partial_\theta f}{g} - (a^\top \nabla)_2 g \partial_y f \right]^2 \\
& - \left[ \frac{\beta(a^\top \nabla)_2 f}{g} + \partial_\theta g \partial_y f \right]^2 + 2 \left[ \frac{\beta(a^\top \nabla)_2 f}{g} \right. \\
& \left. + \partial_\theta g \partial_y f \right] \partial_\theta g \partial_y f - 2\partial_y f (a^\top \nabla)_2 g \times \left[ \frac{\beta \partial_\theta f}{g} - (a^\top \nabla)_2 g \partial_y f \right] + \frac{\beta^2}{2} |\partial_y f|^2
\end{aligned}$$

$$\begin{aligned}
&= -\beta^2 \Gamma_1(f, f) - \frac{\beta^2}{g^2} |(a^\top \nabla)_1 f|^2 - |(a^\top \nabla)_2 (\log g)|^2 |(z^\top \nabla)_1 f|^2 \\
&\quad - 2 \frac{\beta}{g} (a^\top \nabla)_2 \log g (a^\top \nabla)_1 f (z^\top \nabla)_1 f \\
&\quad - \frac{\beta^2}{g^2} |(a^\top \nabla)_2 f|^2 - |(a^\top \nabla)_1 \log g|^2 |(z^\top \nabla)_1 f|^2 \\
&\quad + 2 \frac{\beta}{g} (a^\top \nabla)_1 \log g (a^\top \nabla)_2 f (z^\top \nabla)_1 f \\
&\quad - 2 \frac{\beta}{g} (a^\top \nabla)_1 \log g (a^\top \nabla)_2 f (z^\top \nabla)_1 f + 2 |(a^\top \nabla)_1 \log g|^2 |(z^\top \nabla)_1 f|^2 \\
&\quad + 2 \frac{\beta}{g} (a^\top \nabla)_2 \log g (a^\top \nabla)_1 f (z^\top \nabla)_1 f + 2 |(a^\top \nabla)_2 \log g|^2 |(z^\top \nabla)_1 f|^2 + \frac{\beta^2}{2g^2} \Gamma_1^z(f, f).
\end{aligned}$$

By grouping the bilinear terms of  $\nabla f$ , we obtain

$$\begin{aligned}
&-\Lambda_1^\top Q^\top Q \Lambda_1 - \Lambda_2^\top P^\top P \Lambda_2 + D^\top D + E^\top E \\
&= \Gamma_1(\log g, \log g) \Gamma_1^z(f, f) - \beta^2 \left(1 + \frac{1}{g^2}\right) \Gamma_1(f, f) + \frac{\beta^2}{2g^2} \Gamma_1^z(f, f).
\end{aligned}$$

□

We are now left to compute the three tensor terms.

**Proof of Lemma 6.** For displacement group  $\mathbf{G}$ , we have  $n = 2$  and  $m = 1$ . Recall Theorem 1; we denote  $\mathfrak{R}_{ab}(\nabla f, \nabla f) = \mathfrak{R}_a(\nabla f, \nabla f) + \mathfrak{R}_b(\nabla f, \nabla f)$ , where  $\mathfrak{R}_b(\nabla f, \nabla f)$  represents the tensor term involving drift  $b$ . We thus have

$$\begin{aligned}
\mathfrak{R}_a(\nabla f, \nabla f) &= \sum_{i,k=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{i\hat{i}'}^\top \left( \frac{\partial a_{i\hat{i}}^\top}{\partial x_{i'}} \frac{\partial a_{k\hat{k}}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} \right), (a^\top \nabla)_{kf} \rangle_{\mathbb{R}^2} \\
&\quad + \sum_{i,k=2}^n \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{i\hat{i}'}^\top a_{i\hat{i}}^\top \left( \frac{\partial}{\partial x_{i'}} \frac{\partial a_{k\hat{k}}^\top}{\partial x_{\hat{i}}} \right) \left( \frac{\partial f}{\partial x_{\hat{k}}} \right), (a^\top \nabla)_{kf} \rangle_{\mathbb{R}^2} \\
&\quad - \sum_{i,k=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{k\hat{k}}^\top \frac{\partial a_{i\hat{i}'}^\top}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}, (a^\top \nabla)_{kf} \rangle_{\mathbb{R}^2} \\
&\quad - \sum_{i,k=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{k\hat{k}}^\top a_{i\hat{i}'}^\top \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^\top}{\partial x_{i'}} \right) \frac{\partial f}{\partial x_{\hat{i}}}, (a^\top \nabla)_{kf} \rangle_{\mathbb{R}^2}, \\
&= \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4.
\end{aligned}$$

By direct computations, we have

$$\begin{aligned}
\mathcal{I}_1 &= \sum_{i=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \left[ a_{i\hat{i}'}^\top \left( \frac{\partial a_{i\hat{i}}^\top}{\partial x_{i'}} \frac{\partial a_{1\hat{k}}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^\top \nabla)_1 f \right. \\
&\quad \left. + a_{i\hat{i}'}^\top \left( \frac{\partial a_{i\hat{i}}^\top}{\partial x_{i'}} \frac{\partial a_{2\hat{k}}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^\top \nabla)_2 f \right] = 0;
\end{aligned}$$

$$\begin{aligned}
\mathcal{I}_2 &= \sum_{i=2}^n \sum_{i', \hat{i}, \hat{k}=1}^3 \left[ a_{ii'}^T a_{i\hat{i}}^T \left( \frac{\partial}{\partial x_{i'}} \frac{\partial a_{1\hat{k}}^T}{\partial x_{\hat{i}}} \right) \left( \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^T \nabla)_1 f \right. \\
&\quad \left. + a_{ii'}^T a_{i\hat{i}}^T \left( \frac{\partial}{\partial x_{i'}} \frac{\partial a_{2\hat{k}}^T}{\partial x_{\hat{i}}} \right) \left( \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^T \nabla)_2 f \right] \\
&= a_{11}^T a_{11}^T \frac{\partial^2}{\partial \theta \partial \theta} a_{22}^T \frac{\partial f}{\partial x} (a^T \nabla)_2 f = \beta^2 e^{\beta \theta} \frac{\partial f}{\partial x} (a^T \nabla)_2 f; \\
\mathcal{I}_3 &= - \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 \left[ a_{1\hat{k}}^T \frac{\partial a_{ii'}^T}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_1 f \right. \\
&\quad \left. + a_{2\hat{k}}^T \frac{\partial a_{ii'}^T}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_2 f \right] = 0; \\
\mathcal{I}_4 &= - \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 \left[ a_{1\hat{k}}^T a_{ii'}^T \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \right) \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_1 f \right. \\
&\quad \left. + a_{2\hat{k}}^T a_{ii'}^T \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \right) \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_2 f \right] = 0.
\end{aligned}$$

For the drift term in tensor  $\mathfrak{R}_{ab}$ , taking  $b = -aa^T \nabla V$ , we obtain

$$\begin{aligned}
\mathfrak{R}_b^a &= \sum_{i,k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{i\hat{i}}^T \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}} a_{kk'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_{if} \right] \\
&\quad + \sum_{i,k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{i\hat{i}}^T \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}} a_{kk'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_{if} \right] \\
&\quad + \sum_{i,k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{i\hat{i}}^T a_{k\hat{k}}^T a_{kk'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_{if} \right] \\
&\quad - \sum_{i,k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{k\hat{k}}^T a_{kk'}^T \frac{\partial a_{i\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_{if} \right] \\
&= \mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3 + \mathcal{J}_4.
\end{aligned}$$

Plugging into the matrix  $a^T$ , we obtain

$$\begin{aligned}
\mathcal{J}_1 &= \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{1\hat{i}}^T \frac{\partial a_{1\hat{k}}^T}{\partial x_{\hat{i}}} a_{1k'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T \frac{\partial a_{1\hat{k}}^T}{\partial x_{\hat{i}}} a_{1k'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\
&\quad + \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{1\hat{i}}^T \frac{\partial a_{2\hat{k}}^T}{\partial x_{\hat{i}}} a_{2k'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T \frac{\partial a_{2\hat{k}}^T}{\partial x_{\hat{i}}} a_{2k'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\
&= \beta e^{\beta \theta} (a^T \nabla)_2 V \frac{\partial f}{\partial x} (a^T \nabla)_1 f; \\
\mathcal{J}_2 &= \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{1\hat{i}}^T \frac{\partial a_{1k'}^T}{\partial x_{\hat{i}}} a_{1\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T \frac{\partial a_{1k'}^T}{\partial x_{\hat{i}}} a_{1\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\
&\quad + \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{1\hat{i}}^T \frac{\partial a_{2k'}^T}{\partial x_{\hat{i}}} a_{2\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T \frac{\partial a_{2k'}^T}{\partial x_{\hat{i}}} a_{2\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\
&= \beta e^{\beta \theta} \frac{\partial V}{\partial x} (a^T \nabla)_2 f (a^T \nabla)_1 f;
\end{aligned}$$

$$\begin{aligned}
\mathcal{J}_3 &= \sum_{\hat{i}, \hat{k}'=1}^3 \left[ a_{1\hat{i}}^T a_{1\hat{k}'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{\hat{k}'}} |(a^T \nabla)_1 f|^2 + a_{2\hat{i}}^T a_{1\hat{k}'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{\hat{k}'}} (a^T \nabla)_1 f (a^T \nabla)_2 f \right] \\
&\quad + \sum_{\hat{i}, \hat{k}'=1}^3 \left[ a_{1\hat{i}}^T a_{2\hat{k}'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{\hat{k}'}} (a^T \nabla)_2 f (a^T \nabla)_1 f + a_{2\hat{i}}^T a_{2\hat{k}'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{\hat{k}'}} |(a^T \nabla)_2 f|^2 \right] \\
&= \frac{\partial^2 V}{\partial \theta \partial \theta} |(a^T \nabla)_1 f|^2 + 2(e^{\beta \theta} \frac{\partial^2 V}{\partial \theta \partial x} + \frac{\partial^2 V}{\partial \theta \partial y}) (a^T \nabla)_1 f (a^T \nabla)_2 f \\
&\quad + \sum_{\hat{i}, \hat{k}'=1}^3 a_{2\hat{i}}^T a_{2\hat{k}'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{\hat{k}'}} |(a^T \nabla)_2 f|^2; \\
\mathcal{J}_4 &= - \sum_{\hat{i}, \hat{k}, \hat{k}'=1}^3 \left[ a_{1\hat{k}}^T a_{1\hat{k}'}^T \frac{\partial a_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{\hat{k}'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_1 f + a_{1\hat{k}}^T a_{1\hat{k}'}^T \frac{\partial a_{2\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{\hat{k}'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_2 f \right] \\
&\quad - \sum_{\hat{i}, \hat{k}, \hat{k}'=1}^3 \left[ a_{2\hat{k}}^T a_{2\hat{k}'}^T \frac{\partial a_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{\hat{k}'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_1 f + a_{2\hat{k}}^T a_{2\hat{k}'}^T \frac{\partial a_{2\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{\hat{k}'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_2 f \right] \\
&= -\beta e^{\beta \theta} (a^T \nabla)_1 V \frac{\partial f}{\partial x} (a^T \nabla)_2 f.
\end{aligned}$$

Combining the above computations, we obtain the tensor  $\mathfrak{R}_{ab}$ . Now, we turn to the second tensor  $\mathfrak{R}_{zb}$ , which has the following form:

$$\begin{aligned}
\mathfrak{R}_{zb}(\nabla f, \nabla f) &= \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 \langle a_{ii'}^T (\frac{\partial a_{ii}^T}{\partial x_{i'}} \frac{\partial z_{1\hat{k}}^T}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (z^T \nabla)_1 f \rangle_{\mathbb{R}} \\
&\quad + \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 \langle a_{ii'}^T a_{ii}^T (\frac{\partial}{\partial x_{i'}} \frac{\partial z_{1\hat{k}}^T}{\partial x_{\hat{i}}}) (\frac{\partial f}{\partial x_{\hat{k}}}), (z^T \nabla)_1 f \rangle_{\mathbb{R}} \\
&\quad - \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 \langle z_{1\hat{k}}^T \frac{\partial a_{ii'}^T}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^T}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}, (z^T \nabla)_1 f \rangle_{\mathbb{R}} \\
&\quad - \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 \langle z_{1\hat{k}}^T a_{ii'}^T (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^T}{\partial x_{i'}}) \frac{\partial f}{\partial x_{\hat{i}}}, (z^T \nabla)_1 f \rangle_{\mathbb{R}} \\
&\quad - \sum_{i=1}^2 \sum_{\hat{i}, \hat{k}=1}^3 \langle (z_{1\hat{i}}^T \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial z_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^T \nabla)_1 f \rangle_{\mathbb{R}}, \\
&= \mathcal{I}_1^z + \mathcal{I}_2^z + \mathcal{I}_3^z + \mathcal{I}_4^z + \mathfrak{R}_b^z(\nabla f, \nabla f).
\end{aligned}$$

where we denote further that

$$\mathfrak{R}_b^z(\nabla f, \nabla f) = - \sum_{\hat{i}, \hat{k}=1}^3 (z_{1\hat{i}}^T \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial z_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}) (z^T \nabla)_1 f.$$

By taking  $b = -aa^T \nabla V$ , we further obtain that

$$\begin{aligned}
\mathfrak{R}_b^z(\nabla f, \nabla f) &= - \sum_{\hat{i}, \hat{k}=1}^3 \left[ z_{1\hat{i}}^T \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f - b_{\hat{k}} \frac{\partial z_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}} (z^T \nabla)_1 f \right] \\
&= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, \hat{k}'=1}^3 \left[ z_{1\hat{i}}^T \frac{\partial a_{kk}^T}{\partial x_{\hat{i}}} a_{kk'}^T \frac{\partial V}{\partial x_{\hat{k}'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f \right] \\
&\quad + \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, \hat{k}'=1}^3 \left[ z_{1\hat{i}}^T \frac{\partial a_{kk'}^T}{\partial x_{\hat{i}}} a_{kk}^T \frac{\partial V}{\partial x_{\hat{k}'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f \right]
\end{aligned}$$



$$\begin{aligned}
& + \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^T a_{k\hat{k}}^T a_{kk'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f \right] \\
& - \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{k\hat{k}}^T a_{kk'}^T \frac{\partial z_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (z^T \nabla)_1 f \right] \\
& = \mathcal{J}_1^z + \mathcal{J}_2^z + \mathcal{J}_3^z + \mathcal{J}_4^z.
\end{aligned}$$

By direct computations, it is not hard to observe that

$$\begin{aligned}
\mathcal{I}_1^z &= \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 \langle a_{ii'}^T \left( \frac{\partial a_{ii}^T}{\partial x_{i'}} \frac{\partial z_{1\hat{k}}^T}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} \right), (z^T \nabla)_1 f \rangle_{\mathbb{R}} = 0; \\
\mathcal{I}_2^z &= \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 \langle a_{ii'}^T a_{ii}^T \left( \frac{\partial}{\partial x_{i'}} \frac{\partial z_{1\hat{k}}^T}{\partial x_{\hat{i}}} \right) \left( \frac{\partial f}{\partial x_{\hat{k}}} \right), (z^T \nabla)_1 f \rangle_{\mathbb{R}} \\
&= \sum_{i=1}^2 \sum_{i', \hat{i}=1}^3 a_{ii'}^T a_{ii}^T \frac{\partial^2 z_{1\hat{k}}^T}{\partial x_{i'} \partial x_{\hat{i}}} \partial_y f (z^T \nabla)_1 f; \\
\mathcal{I}_3^z &= - \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 \langle z_{1\hat{k}}^T \frac{\partial a_{ii'}^T}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^T}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}, (z^T \nabla)_1 f \rangle_{\mathbb{R}} = 0; \\
\mathcal{I}_4^z &= - \sum_{i=1}^2 \sum_{i', \hat{i}, \hat{k}=1}^3 \langle z_{1\hat{k}}^T a_{ii'}^T \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^T}{\partial x_{i'}} \right) \frac{\partial f}{\partial x_{\hat{i}}}, (z^T \nabla)_1 f \rangle_{\mathbb{R}} = 0,
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{J}_1^z &= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^T \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}} a_{kk'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f \right] = 0; \\
\mathcal{J}_2^z &= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^T \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}} a_{kk'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f \right] = 0; \\
\mathcal{J}_4^z &= - \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{k\hat{k}}^T a_{kk'}^T \frac{\partial z_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (z^T \nabla)_1 f \right] \\
&= - \sum_{k=1}^2 (a^T \nabla)_k z_{13}^T (a^T \nabla)_k V \partial_y f (z^T \nabla)_1 f; \\
\mathcal{J}_3^z &= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^T a_{k\hat{k}}^T a_{kk'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f \right] \\
&= \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^T a_{1\hat{k}}^T a_{1k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f + z_{1\hat{i}}^T a_{2\hat{k}}^T a_{2k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^T \nabla)_1 f \right] \\
&= \sum_{\hat{i}, k'=1}^3 \left[ z_{1\hat{i}}^T a_{1k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} (a^T \nabla)_1 f (z^T \nabla)_1 f + z_{1\hat{i}}^T a_{2k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} (a^T \nabla)_2 f (z^T \nabla)_1 f \right] \\
&= -g \frac{\partial^2 V}{\partial \theta \partial y} (a^T \nabla)_1 f (z^T \nabla)_1 f - g(e^{\beta \theta} \frac{\partial^2 V}{\partial x \partial y} + \frac{\partial^2 V}{\partial y \partial y}) (a^T \nabla)_2 f (z^T \nabla)_1 f.
\end{aligned}$$

Now, we are left to compute the term  $\mathfrak{R}_\pi$ . Recall that

$$\begin{aligned}
 & \mathfrak{R}_\pi(\nabla f, \nabla f) \\
 = & 2 \sum_{k=1}^1 \sum_{i=1}^2 \sum_{k', \hat{k}, \hat{i}, i'=1}^3 \left[ \frac{\partial}{\partial x_{k'}} z_{kk'}^\top z_{k\hat{k}}^\top \frac{\partial}{\partial x_{\hat{k}}} a_{ii}^\top \frac{\partial f}{\partial x_{\hat{i}}} a_{ii'}^\top \frac{\partial f}{\partial x_{i'}} \right] \\
 & + 2 \sum_{k=1}^1 \sum_{i=1}^2 \sum_{k', \hat{k}, \hat{i}, i'=1}^3 \left[ z_{kk'}^\top \frac{\partial}{\partial x_{k'}} z_{k\hat{k}}^\top \frac{\partial}{\partial x_{\hat{k}}} a_{ii}^\top \frac{\partial f}{\partial x_{\hat{i}}} a_{ii'}^\top \frac{\partial f}{\partial x_{i'}} \right. \\
 & \quad + z_{kk'}^\top z_{k\hat{k}}^\top \frac{\partial^2}{\partial x_{k'} \partial x_{\hat{k}}} a_{ii}^\top \frac{\partial f}{\partial x_{\hat{i}}} a_{ii'}^\top \frac{\partial f}{\partial x_{i'}} \\
 & \quad \left. + z_{kk'}^\top z_{k\hat{k}}^\top \frac{\partial}{\partial x_{\hat{k}}} a_{ii}^\top \frac{\partial f}{\partial x_{\hat{i}}} \frac{\partial}{\partial x_{k'}} a_{ii'}^\top \frac{\partial f}{\partial x_{i'}} \right] \\
 & + 2 \sum_{k=1}^1 \sum_{i=1}^2 \sum_{\hat{k}, \hat{i}, i'=1}^3 (z^\top \nabla \log \pi)_k \left[ z_{k\hat{k}}^\top \frac{\partial}{\partial x_{\hat{k}}} a_{ii}^\top \frac{\partial f}{\partial x_{\hat{i}}} a_{ii'}^\top \frac{\partial f}{\partial x_{i'}} \right] \\
 & - 2 \sum_{j=1}^1 \sum_{l=1}^2 \sum_{l', \hat{l}, \hat{j}, j'=1}^3 \left[ \frac{\partial}{\partial x_{l'}} a_{ll'}^\top a_{ll}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{jj}^\top \frac{\partial f}{\partial x_{\hat{j}}} z_{jj'}^\top \frac{\partial f}{\partial x_{j'}} \right] \\
 & - 2 \sum_{j=1}^1 \sum_{l=1}^2 \sum_{l', \hat{l}, \hat{j}, j'=1}^3 \left[ a_{ll'}^\top \frac{\partial}{\partial x_{l'}} a_{ll}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{jj}^\top \frac{\partial f}{\partial x_{\hat{j}}} z_{jj'}^\top \frac{\partial f}{\partial x_{j'}} \right. \\
 & \quad + a_{ll'}^\top a_{ll}^\top \frac{\partial^2}{\partial x_{l'} \partial x_{\hat{l}}} z_{jj}^\top \frac{\partial f}{\partial x_{\hat{j}}} z_{jj'}^\top \frac{\partial f}{\partial x_{j'}} \\
 & \quad \left. + a_{ll'}^\top a_{ll}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{jj}^\top \frac{\partial f}{\partial x_{\hat{j}}} \frac{\partial}{\partial x_{l'}} z_{jj'}^\top \frac{\partial f}{\partial x_{j'}} \right] \\
 & - 2 \sum_{j=1}^1 \sum_{l=1}^2 \sum_{\hat{l}, \hat{j}, j'=1}^3 (a^\top \nabla \log \pi)_l \left[ a_{ll}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{jj}^\top \frac{\partial f}{\partial x_{\hat{j}}} z_{jj'}^\top \frac{\partial f}{\partial x_{j'}} \right] \\
 = & \sum_{i=1}^{10} \mathcal{K}_i.
 \end{aligned}$$

By direct computation, we obtain

$$\begin{aligned}
 \mathcal{K}_1 &= 0, \quad \mathcal{K}_2 = 0, \quad \mathcal{K}_3 = 0, \quad \mathcal{K}_4 = 0, \quad \mathcal{K}_5 = 0, \quad \mathcal{K}_6 = 0, \quad \mathcal{K}_7 = 0; \\
 \mathcal{K}_8 &= -2 \sum_{l=1}^2 \sum_{\hat{l}=1}^3 a_{ll'}^\top a_{ll}^\top \frac{\partial^2 z_{13}^\top}{\partial x_{l'} \partial x_{\hat{l}}} \partial_y f (z^\top \nabla)_1 f; \\
 \mathcal{K}_9 &= -2 \sum_{l=1}^2 \sum_{\hat{l}=1}^3 a_{ll'}^\top a_{ll}^\top \frac{\partial z_{13}^\top}{\partial x_{\hat{l}}} \frac{\partial z_{13}^\top}{\partial x_{l'}} |\partial_y f|^2 = -2 \Gamma_1(\log g, \log g) |(z^\top \nabla)_1 f|^2; \\
 \mathcal{K}_{10} &= -2 \sum_{l=1}^2 \sum_{\hat{l}=1}^3 (a^\top \nabla)_l \log \pi a_{ll}^\top \frac{\partial z_{13}^\top}{\partial x_{\hat{l}}} \partial_y f (z^\top \nabla)_1 f = -2 \Gamma_1(\log \pi, \log g) |(z^\top \nabla)_1 f|^2.
 \end{aligned}$$

□

### 3.3. Martinet Flat Sub-Riemannian Structure

In this part, we apply our result to the Martinet flat sub-Riemannian structure, which satisfies the bracket-generating condition and has a non-equiregular sub-Riemannian structure (see [37]). The sub-Riemannian structure is defined on  $\mathbb{R}^3$  through the kernel of

one-form  $\eta : = dz - \frac{1}{2}y^2 dx$ . A global orthonormal basis for the horizontal distribution  $\mathcal{H}$  adapts the following differential operator representation, in local coordinates  $(x, y, z)$ :

$$X = \frac{\partial}{\partial x} + \frac{y^2}{2} \frac{\partial}{\partial z}, \quad Y = \frac{\partial}{\partial y}.$$

The commutative relation gives

$$[X, Y] = -yZ, \quad [Y, [X, Y]] = Z, \quad \text{where } Z = \frac{\partial}{\partial z}.$$

To apply it in our framework, we take

$$\begin{aligned} a &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{y^2}{2} & 0 \end{pmatrix}, \quad a^\top = \begin{pmatrix} 1 & 0 & \frac{y^2}{2} \\ 0 & 1 & 0 \end{pmatrix}, \\ z^\top &= (0, 0, 1), \quad aa^\top = \begin{pmatrix} 1 & 0 & \frac{y^2}{2} \\ 0 & 1 & 0 \\ \frac{y^2}{2} & 0 & \frac{y^4}{4} \end{pmatrix}. \end{aligned} \quad (27)$$

Thus, the sub-Riemannian structure has the form  $(\mathbb{M}, \mathcal{H}, (aa^\top)^\dagger|_{\mathcal{H}})$ .

**Proposition 5.** *In this setting,*

$$\pi = e^{-\frac{y^2}{2} - V},$$

then

$$-aa^\top \nabla \log \pi = a \otimes \nabla a + aa^\top \nabla V.$$

**Proof.** The poof follows from the observation that

$$a \otimes \nabla a = \begin{pmatrix} 0 & y & 0 \end{pmatrix}^\top, \quad aa^\top \nabla \log e^{-\frac{y^2}{2}} = \begin{pmatrix} 0 & y & 0 \end{pmatrix}^\top.$$

□

Similar to the previous displacement group case, we have the following identity.

**Proposition 6.** *For any smooth function  $f \in C^\infty(\mathbb{M})$ , one has*

$$\Gamma_2(f, f) + \Gamma_2^{z, \pi}(f, f) = \|\mathfrak{Hess}_{a, z} f\|^2 + \mathfrak{R}(\nabla f, \nabla f),$$

where

$$\begin{aligned} \Lambda_1^\top &= (0, y\partial_z f/2, 0, y\partial_z f/2, 0, 0, 0, 0, 0); \\ \Lambda_2^\top &= (0, 0, 0, 0, 0, 0, -y\partial_y f, \frac{y^3}{2}\partial_z f + y\partial_x f, 0); \\ \mathfrak{R}_{ab}(\nabla f, \nabla f) - \Lambda_1^\top Q^\top Q \Lambda_1 - \Lambda_2^\top P^\top P \Lambda_2 + D^\top D + E^\top E \\ &= \frac{y^2}{2} \Gamma_1^z(f, f) - y^2 \Gamma_1(f, f) \\ &\quad + \frac{\partial f}{\partial z} (a^\top \nabla)_1 f + y(a^\top \nabla)_1 V \frac{\partial f}{\partial z} (a^\top \nabla)_2 f + y \frac{\partial V}{\partial z} (a^\top \nabla)_1 f (a^\top \nabla)_2 f \\ &\quad + \sum_{i, k'=1}^3 a_{1i}^\top a_{1k'}^\top \frac{\partial^2 V}{\partial x_i \partial x_{k'}} |(a^\top \nabla)_1 f|^2 + 2 \left( \frac{\partial^2 V}{\partial x \partial y} + \frac{y^2}{2} \frac{\partial^2 V}{\partial y \partial z} \right) (a^\top \nabla)_1 f (a^\top \nabla)_2 f \end{aligned}$$

$$\begin{aligned}
& + \frac{\partial^2 V}{\partial y \partial y} |(a^\top \nabla)_2 f|^2 - y \frac{\partial V}{\partial y} \frac{\partial f}{\partial z} (a^\top \nabla)_1 f; \\
\mathfrak{R}_{zb}(\nabla f, \nabla f) &= \left( \frac{\partial^2 V}{\partial x \partial z} + \frac{y^2}{2} \frac{\partial^2 V}{\partial z \partial z} \right) (a^\top \nabla)_1 f (z^\top \nabla)_1 f + \frac{\partial^2 V}{\partial y \partial z} (a^\top \nabla)_2 f (z^\top \nabla)_1 f; \\
\mathfrak{R}_\pi(\nabla f, \nabla f) &= 0.
\end{aligned}$$

In particular, we have

$$\sum_{i,k'=1}^3 a_{1i}^\top a_{1k'}^\top \frac{\partial^2 V}{\partial x_i \partial x_{k'}} |(a^\top \nabla)_1 f|^2 = \left( \frac{\partial^2 V}{\partial x \partial x} + y^2 \frac{\partial^2 V}{\partial x \partial z} + \frac{y^4}{4} \frac{\partial^2 V}{\partial z \partial z} \right) |(a^\top \nabla)_1 f|^2.$$

The proof of Proposition 6 follows from the proof of Theorem 1 (i.e., Theorem 3) and Lemmas 7–9 below. The following convergence results are a direct consequence of Theorem 2.

**Proposition 7.** *If there exists  $\kappa > 0$  as shown in Theorem 2, the exponential dissipation result in the  $L^1$  distance holds:*

$$\int |\rho(t, x) - \pi(x)| dx = O(e^{-\kappa t}).$$

Similarly, we summarize the sub-Riemannian Ricci tensor in terms of  $\mathfrak{R}$  as follows.

**Corollary 3.** *The matrix  $\mathfrak{R}$  associated with the Martinet sub-Riemannian structure has the following form:*

$$\begin{aligned}
\mathfrak{R}_{11} &= \left( \frac{\partial^2 V}{\partial x \partial x} + y^2 \frac{\partial^2 V}{\partial x \partial z} + \frac{y^4}{4} \frac{\partial^2 V}{\partial z \partial z} \right) - y^2; \\
\mathfrak{R}_{22} &= \frac{\partial^2 V}{\partial y \partial y} - y^2; \quad \mathfrak{R}_{33} = \frac{y^2}{2}; \\
\mathfrak{R}_{12} &= \mathfrak{R}_{21} = \frac{y}{2} \frac{\partial V}{\partial z} + \left( \frac{\partial^2 V}{\partial x \partial y} + \frac{y^2}{2} \frac{\partial^2 V}{\partial y \partial z} \right); \\
\mathfrak{R}_{13} &= \mathfrak{R}_{31} = \frac{1}{2} - \frac{y}{2} \frac{\partial V}{\partial y} + \frac{1}{2} \left( \frac{\partial^2 V}{\partial x \partial z} + \frac{y^2}{2} \frac{\partial^2 V}{\partial z \partial z} \right); \quad \mathfrak{R}_{23} = \mathfrak{R}_{32} = \frac{1}{2} y (a^\top \nabla)_1 V + \frac{1}{2} \frac{\partial^2 V}{\partial y \partial z}.
\end{aligned}$$

**Proof.** The proof follows from the similar equivalent matrix formulation as shown in the proof of Corollary 1 and the explicit bilinear forms in Proposition 6.  $\square$

Next, we prove the following three key lemmas.

**Lemma 7.** *For Martinet sub-Riemannian structure  $(\mathbb{M}, \mathcal{H}, (aa^\top)_{|\mathcal{H}}^\dagger)$ , we have*

$$\begin{aligned}
Q &= \begin{pmatrix} 1 & 0 & \frac{y^2}{2} & 0 & 0 & 0 & \frac{y^2}{2} & 0 & \frac{y^4}{4} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \frac{y^2}{2} & 0 \\ 0 & 0 & 0 & 1 & 0 & \frac{y^2}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}; \\
P &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & y^2/2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}; \\
C^\top &= (0, 0, 0, 0, 0, \frac{y^3}{2} \partial_z f + y \partial_x f, -y \partial_y f, 0, -\frac{y^3}{2} \partial_y f); \\
D^\top &= (0, 0, y \partial_z f, 0), \quad E^\top = (0, 0); \\
F^\top &= G^\top = (0, 0, 0, 0, 0, 0, 0, 0, 0).
\end{aligned}$$

**Proof.** Plugging matrices  $a$  and  $z$  from (27) into Notation 1, we complete the proof.  $\square$

**Lemma 8.** For the Martinet sub-Riemannian structure,  $F$  and  $G$  are zero vectors, and we have

$$\begin{aligned} & [QX + D]^T [QX + D] + [PX + E]^T [PX + E] + 2C^T X \\ &= \|\mathfrak{Hess}_{a,z} f\|^2 - \Lambda_1^T Q^T Q \Lambda_1 - \Lambda_2^T P^T P \Lambda_2 + D^T D + E^T E. \end{aligned}$$

In particular, we have

$$\begin{aligned} \|\mathfrak{Hess}_{a,z} f\|^2 &= [X + \Lambda_1]^T Q^T Q [X + \Lambda_1] + [X + \Lambda_2]^T P^T P [X + \Lambda_2]; \\ \Lambda_1^T &= (0, y \partial_z f / 2, 0, y \partial_z f / 2, 0, 0, 0, 0); \\ \Lambda_2^T &= (0, 0, 0, 0, 0, 0, -y \partial_y f, \frac{y^3}{2} \partial_z f + y \partial_x f, 0); \\ &\quad - \Lambda_1^T Q^T Q \Lambda_1 - \Lambda_2^T P^T P \Lambda_2 + D^T D + E^T E \\ &= \frac{y^2}{2} \Gamma_1^z(f, f) - y^2 \Gamma_1(f, f). \end{aligned}$$

**Lemma 9.** By routine computations, we obtain

$$\begin{aligned} & \mathfrak{R}_{ab}(\nabla f, \nabla f) \\ &= \frac{\partial f}{\partial z} (a^T \nabla)_1 f + y (a^T \nabla)_1 V \frac{\partial f}{\partial z} (a^T \nabla)_2 f + y \frac{\partial V}{\partial z} (a^T \nabla)_1 f (a^T \nabla)_2 f \\ &\quad + \sum_{i,k'=1}^3 a_{1i}^T a_{1k'}^T \frac{\partial^2 V}{\partial x_i \partial x_{k'}} |(a^T \nabla)_1 f|^2 + 2 \left( \frac{\partial^2 V}{\partial x \partial y} + \frac{y^2}{2} \frac{\partial^2 V}{\partial y \partial z} \right) (a^T \nabla)_1 f (a^T \nabla)_2 f \\ &\quad + \frac{\partial^2 V}{\partial y \partial y} |(a^T \nabla)_2 f|^2 - y \frac{\partial V}{\partial y} \frac{\partial f}{\partial z} (a^T \nabla)_1 f; \\ & \mathfrak{R}_{zb}(\nabla f, \nabla f) \\ &= \left( \frac{\partial^2 V}{\partial x \partial z} + \frac{y^2}{2} \frac{\partial^2 V}{\partial z \partial z} \right) (a^T \nabla)_1 f (z^T \nabla)_1 f + \frac{\partial^2 V}{\partial y \partial z} (a^T \nabla)_2 f (z^T \nabla)_1 f; \\ & \mathfrak{R}_{\pi}(\nabla f, \nabla f) = 0. \end{aligned}$$

**Proof of Lemma 8.** Since  $F$  and  $G$  are zero vectors, we have

$$2C^T X = 2 \left[ \frac{\partial^2 f}{\partial y \partial z} \left( \frac{y^3}{2} \partial_z f + y \partial_x f \right) - \frac{\partial^2 f}{\partial x \partial z} (y \partial_y f) - \frac{\partial^2 f}{\partial z \partial z} \left( \frac{y^3}{2} \partial_y f \right) \right].$$

By routine computation, we observe that

$$\begin{aligned} & [QX + D]^T [QX + D] + [PX + E]^T [PX + E] + 2C^T X \\ &= \left[ \frac{\partial^2 f}{\partial x \partial x} + \frac{y^2}{2} \frac{\partial^2 f}{\partial x \partial z} + \frac{y^2}{2} \frac{\partial^2 f}{\partial z \partial x} + \frac{y^4}{4} \frac{\partial^2 f}{\partial z \partial z} \right]^2 + \left[ \frac{\partial^2 f}{\partial y \partial x} + \frac{y^2}{2} \frac{\partial^2 f}{\partial z \partial y} + y \partial_z f \right]^2 \\ &\quad + \left[ \frac{\partial^2 f}{\partial y \partial x} + \frac{y^2}{2} \frac{\partial^2 f}{\partial z \partial y} \right]^2 + \left[ \frac{\partial^2 f}{\partial y \partial y} \right]^2 \\ &\quad + \left[ \frac{\partial^2 f}{\partial z \partial x} + \frac{y^2}{2} \frac{\partial^2 f}{\partial z \partial z} \right]^2 + \left[ \frac{\partial^2 f}{\partial z \partial y} \right]^2 \\ &\quad + 2 \frac{\partial^2 f}{\partial y \partial z} \left( \frac{y^3}{2} \partial_z f + y \partial_x f \right) - 2 \frac{\partial^2 f}{\partial x \partial z} (y \partial_y f) - 2 \frac{\partial^2 f}{\partial z \partial z} \left( \frac{y^3}{2} \partial_y f \right) \end{aligned}$$

$$\begin{aligned}
 &= \left[ \frac{\partial^2 f}{\partial x \partial x} + \frac{y^2}{2} \frac{\partial^2 f}{\partial x \partial z} + \frac{y^2}{2} \frac{\partial^2 f}{\partial z \partial x} + \frac{y^4}{4} \frac{\partial^2 f}{\partial z \partial z} \right]^2 + \left[ \frac{\partial^2 f}{\partial y \partial x} + \frac{y^2}{2} \frac{\partial^2 f}{\partial z \partial y} + y \partial_z f \right]^2 \\
 &\quad + \left[ \frac{\partial^2 f}{\partial y \partial x} + \frac{y^2}{2} \frac{\partial^2 f}{\partial z \partial y} \right]^2 + \left[ \frac{\partial^2 f}{\partial y \partial y} \right]^2 \\
 &\quad + \left[ \frac{\partial^2 f}{\partial z \partial x} + \frac{y^2}{2} \frac{\partial^2 f}{\partial z \partial z} - y \partial_y f \right]^2 + \left[ \frac{\partial^2 f}{\partial z \partial y} + \left( \frac{y^3}{2} \partial_z f + y \partial_x f \right) \right]^2 \\
 &\quad - y^2 |\partial_y f|^2 - \left( \frac{y^3}{2} \partial_z f + y \partial_x f \right)^2 \\
 &= |\mathfrak{Hess}_{a,z} f|^2 + \frac{y^2}{2} \Gamma_1^z(f, f) - y^2 \Gamma_1(f, f),
 \end{aligned}$$

where we use the fact

$$\begin{aligned}
 &\left[ \frac{\partial^2 f}{\partial y \partial x} + \frac{y^2}{2} \frac{\partial^2 f}{\partial z \partial y} + y \partial_z f \right]^2 + \left[ \frac{\partial^2 f}{\partial y \partial x} + \frac{y^2}{2} \frac{\partial^2 f}{\partial z \partial y} \right]^2 \\
 &= 2 \left[ \frac{\partial^2 f}{\partial y \partial x} + \frac{y^2}{2} \frac{\partial^2 f}{\partial z \partial y} + \frac{1}{2} y \partial_z f \right]^2 + \frac{y^2}{2} |\partial_z f|^2.
 \end{aligned}$$

The proof is thus completed.  $\square$

We are now left to compute the three tensor terms.

**Proof of Lemma 9.** Similar to the proof of Lemma 6, we have

$$\begin{aligned}
 \mathfrak{R}_a(\nabla f, \nabla f) &= \sum_{i,k=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{ii'}^\top \left( \frac{\partial a_{ii}^\top}{\partial x_{i'}} \frac{\partial a_{kk}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} \right), (a^\top \nabla)_k f \rangle_{\mathbb{R}^2} \\
 &\quad + \sum_{i,k=2}^n \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{ii'}^\top a_{ii}^\top \left( \frac{\partial}{\partial x_{i'}} \frac{\partial a_{kk}^\top}{\partial x_{\hat{i}}} \right) \left( \frac{\partial f}{\partial x_{\hat{k}}} \right), (a^\top \nabla)_k f \rangle_{\mathbb{R}^2} \\
 &\quad - \sum_{i,k=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{kk}^\top \frac{\partial a_{ii'}^\top}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}, (a^\top \nabla)_k f \rangle_{\mathbb{R}^2} \\
 &\quad - \sum_{i,k=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \langle a_{kk}^\top a_{ii'}^\top \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \right) \frac{\partial f}{\partial x_{\hat{i}}}, (a^\top \nabla)_k f \rangle_{\mathbb{R}^2}, \\
 &= \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4.
 \end{aligned}$$

By direct computations, we have

$$\begin{aligned}
 \mathcal{I}_1 &= \sum_{i=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \left[ a_{ii'}^\top \left( \frac{\partial a_{ii}^\top}{\partial x_{i'}} \frac{\partial a_{1k}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^\top \nabla)_1 f + a_{ii'}^\top \left( \frac{\partial a_{ii}^\top}{\partial x_{i'}} \frac{\partial a_{2k}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^\top \nabla)_2 f \right] = 0; \\
 \mathcal{I}_2 &= \sum_{i=2}^n \sum_{i',\hat{i},\hat{k}=1}^3 \left[ a_{ii'}^\top a_{ii}^\top \left( \frac{\partial}{\partial x_{i'}} \frac{\partial a_{1k}^\top}{\partial x_{\hat{i}}} \right) \left( \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^\top \nabla)_1 f + a_{ii'}^\top a_{ii}^\top \left( \frac{\partial}{\partial x_{i'}} \frac{\partial a_{2k}^\top}{\partial x_{\hat{i}}} \right) \left( \frac{\partial f}{\partial x_{\hat{k}}} \right) (a^\top \nabla)_2 f \right] \\
 &= a_{22}^\top a_{22}^\top \frac{\partial^2}{\partial y \partial y} a_{13}^\top \frac{\partial f}{\partial z} (a^\top \nabla)_1 f = \frac{\partial f}{\partial z} (a^\top \nabla)_1 f; \\
 \mathcal{I}_3 &= - \sum_{i=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \left[ a_{1k}^\top \frac{\partial a_{ii'}^\top}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^\top \nabla)_1 f + a_{2k}^\top \frac{\partial a_{ii'}^\top}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^\top \nabla)_2 f \right] = 0; \\
 \mathcal{I}_4 &= - \sum_{i=1}^2 \sum_{i',\hat{i},\hat{k}=1}^3 \left[ a_{1k}^\top a_{ii'}^\top \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \right) \frac{\partial f}{\partial x_{\hat{i}}} (a^\top \nabla)_1 f + a_{2k}^\top a_{ii'}^\top \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \right) \frac{\partial f}{\partial x_{\hat{i}}} (a^\top \nabla)_2 f \right] = 0.
 \end{aligned}$$

For the drift term, we take  $b = -aa^T \nabla V$

$$\begin{aligned}\mathfrak{R}_b^a &= \sum_{i,k=1}^2 \sum_{\hat{i},\hat{k},k'=1}^3 \left[ a_{i\hat{i}}^T \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}} a_{kk'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_i f + a_{i\hat{i}}^T \frac{\partial a_{kk'}^T}{\partial x_{\hat{i}}} a_{k\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_i f \right] \\ &\quad + \sum_{i,k=1}^2 \sum_{\hat{i},\hat{k},k'=1}^3 \left[ a_{i\hat{i}}^T a_{k\hat{k}}^T a_{kk'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_i f \right] \\ &\quad - \sum_{i,k=1}^2 \sum_{\hat{i},\hat{k},k'=1}^3 \left[ a_{k\hat{k}}^T a_{kk'}^T \frac{\partial a_{i\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_i f \right] \\ &= \mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3 + \mathcal{J}_4.\end{aligned}$$

Plugging into the matrices of  $a^T$ , we obtain

$$\begin{aligned}\mathcal{J}_1 &= \sum_{\hat{i},\hat{k},k'=1}^3 \left[ a_{1\hat{i}}^T \frac{\partial a_{1\hat{k}}^T}{\partial x_{\hat{i}}} a_{1k'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T \frac{\partial a_{1\hat{k}}^T}{\partial x_{\hat{i}}} a_{1k'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\ &\quad + \sum_{\hat{i},\hat{k},k'=1}^3 \left[ a_{1\hat{i}}^T \frac{\partial a_{2\hat{k}}^T}{\partial x_{\hat{i}}} a_{2k'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T \frac{\partial a_{2\hat{k}}^T}{\partial x_{\hat{i}}} a_{2k'}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\ &= a_{22}^T \frac{\partial a_{13}^T}{\partial y} (a^T \nabla)_1 V \frac{\partial f}{\partial z} (a^T \nabla)_2 f = y(a^T \nabla)_1 V \frac{\partial f}{\partial z} (a^T \nabla)_2 f;\end{aligned}$$

$$\begin{aligned}\mathcal{J}_2 &= \sum_{\hat{i},\hat{k},k'=1}^3 \left[ a_{1\hat{i}}^T \frac{\partial a_{1k'}^T}{\partial x_{\hat{i}}} a_{1\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T \frac{\partial a_{1k'}^T}{\partial x_{\hat{i}}} a_{1\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\ &\quad + \sum_{\hat{i},\hat{k},k'=1}^3 \left[ a_{1\hat{i}}^T \frac{\partial a_{2k'}^T}{\partial x_{\hat{i}}} a_{2\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_1 f + a_{2\hat{i}}^T \frac{\partial a_{2k'}^T}{\partial x_{\hat{i}}} a_{2\hat{k}}^T \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (a^T \nabla)_2 f \right] \\ &= y \frac{\partial V}{\partial z} (a^T \nabla)_1 f (a^T \nabla)_2 f;\end{aligned}$$

$$\begin{aligned}\mathcal{J}_3 &= \sum_{\hat{i},k'=1}^3 \left[ a_{1\hat{i}}^T a_{1k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} |(a^T \nabla)_1 f|^2 + a_{2\hat{i}}^T a_{1k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} (a^T \nabla)_1 f (a^T \nabla)_2 f \right] \\ &\quad + \sum_{\hat{i},k'=1}^3 \left[ a_{1\hat{i}}^T a_{2k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} (a^T \nabla)_2 f (a^T \nabla)_1 f + a_{2\hat{i}}^T a_{2k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} |(a^T \nabla)_2 f|^2 \right] \\ &= \sum_{\hat{i},k'=1}^3 a_{1\hat{i}}^T a_{1k'}^T \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} |(a^T \nabla)_1 f|^2 + 2 \left( \frac{\partial^2 V}{\partial x \partial y} + \frac{y^2}{2} \frac{\partial^2 V}{\partial y \partial z} \right) (a^T \nabla)_1 f (a^T \nabla)_2 f \\ &\quad + \frac{\partial^2 V}{\partial y \partial y} |(a^T \nabla)_2 f|^2;\end{aligned}$$

$$\begin{aligned}\mathcal{J}_4 &= - \sum_{\hat{i},\hat{k},k'=1}^3 \left[ a_{1\hat{k}}^T a_{1k'}^T \frac{\partial a_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_1 f + a_{1\hat{k}}^T a_{1k'}^T \frac{\partial a_{2\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_2 f \right] \\ &\quad - \sum_{\hat{i},\hat{k},k'=1}^3 \left[ a_{2\hat{k}}^T a_{2k'}^T \frac{\partial a_{1\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_1 f + a_{2\hat{k}}^T a_{2k'}^T \frac{\partial a_{2\hat{i}}^T}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (a^T \nabla)_2 f \right] \\ &= -y \frac{\partial V}{\partial y} \frac{\partial f}{\partial z} (a^T \nabla)_1 f.\end{aligned}$$

Combing the above computations, we obtain the tensor  $\mathfrak{R}_{ab}$ . Now, we turn to the second tensor  $\mathfrak{R}_{zb}$ . Since  $z^\top = (0, 0, 1)$ , it is obvious to see that only the drift term of the tensor  $\mathfrak{R}_{zb}$  remains, where we denote

$$\mathfrak{R}_b^z(\nabla f, \nabla f) = - \sum_{\hat{i}, \hat{k}=1}^3 (z_{1\hat{i}}^\top \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial z_{1\hat{i}}^\top}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}) (z^\top \nabla)_1 f.$$

By taking  $b = -aa^\top \nabla V$ , we further obtain that

$$\begin{aligned} \mathfrak{R}_b^z(\nabla f, \nabla f) &= - \sum_{\hat{i}, \hat{k}=1}^3 \left[ z_{1\hat{i}}^\top \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f - b_{\hat{k}} \frac{\partial z_{1\hat{i}}^\top}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}} (z^\top \nabla)_1 f \right] \\ &= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^\top \frac{\partial a_{\hat{k}\hat{k}}^\top}{\partial x_{\hat{i}}} a_{\hat{k}k'}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f \right] \\ &\quad + \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^\top \frac{\partial a_{\hat{k}k'}^\top}{\partial x_{\hat{i}}} a_{\hat{k}\hat{k}}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f \right] \\ &\quad + \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^\top a_{\hat{k}\hat{k}}^\top a_{\hat{k}k'}^\top \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f \right] \\ &\quad - \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{\hat{k}\hat{k}}^\top a_{\hat{k}k'}^\top \frac{\partial z_{1\hat{i}}^\top}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (z^\top \nabla)_1 f \right] \\ &= \mathcal{J}_1^z + \mathcal{J}_2^z + \mathcal{J}_3^z + \mathcal{J}_4^z. \end{aligned}$$

By direct computations, it is not hard to observe that

$$\begin{aligned} \mathcal{J}_1^z &= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^\top \frac{\partial a_{\hat{k}\hat{k}}^\top}{\partial x_{\hat{i}}} a_{\hat{k}k'}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f \right] = 0; \\ \mathcal{J}_2^z &= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^\top \frac{\partial a_{\hat{k}k'}^\top}{\partial x_{\hat{i}}} a_{\hat{k}\hat{k}}^\top \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f \right] = 0; \\ \mathcal{J}_4^z &= - \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ a_{\hat{k}\hat{k}}^\top a_{\hat{k}k'}^\top \frac{\partial z_{1\hat{i}}^\top}{\partial x_{\hat{k}}} \frac{\partial V}{\partial x_{k'}} \frac{\partial f}{\partial x_{\hat{i}}} (z^\top \nabla)_1 f \right] = 0. \end{aligned}$$

The only non-zero term has the following form:

$$\begin{aligned} \mathcal{J}_3^z &= \sum_{k=1}^2 \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^\top a_{\hat{k}\hat{k}}^\top a_{\hat{k}k'}^\top \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f \right] \\ &= \sum_{\hat{i}, \hat{k}, k'=1}^3 \left[ z_{1\hat{i}}^\top a_{1\hat{k}}^\top a_{1k'}^\top \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f + z_{1\hat{i}}^\top a_{2\hat{k}}^\top a_{2k'}^\top \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} \frac{\partial f}{\partial x_{\hat{k}}} (z^\top \nabla)_1 f \right] \\ &= \sum_{\hat{i}, k'=1}^3 \left[ z_{1\hat{i}}^\top a_{1k'}^\top \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} (a^\top \nabla)_1 f (z^\top \nabla)_1 f + z_{1\hat{i}}^\top a_{2k'}^\top \frac{\partial^2 V}{\partial x_{\hat{i}} \partial x_{k'}} (a^\top \nabla)_2 f (z^\top \nabla)_1 f \right] \\ &= \left( \frac{\partial^2 V}{\partial x \partial z} + \frac{y^2}{2} \frac{\partial^2 V}{\partial z \partial z} \right) (a^\top \nabla)_1 f (z^\top \nabla)_1 f + \frac{\partial^2 V}{\partial y \partial z} (a^\top \nabla)_2 f (z^\top \nabla)_1 f. \end{aligned}$$

Since matrix  $z^\top$  is a constant matrix and matrix  $a^\top$  contains only variable  $y$ , it is easy to observe that

$$\mathfrak{R}_\pi(\nabla f, \nabla f) = 0.$$

□



#### 4. Lyapunov Analysis in Sub-Riemannian Density Manifold

In this section, we illustrate the motivation of this paper, which is to design a matrix condition, whose smallest eigenvalue characterizes the convergence rate of the degenerate SDE.

The outline of this section is given below. Consider a density space over the sub-Riemannian manifold. The finite-dimensional sub-Riemannian structure introduces the density space the *infinite-dimensional sub-Riemannian structure*. We name it the *sub-Riemannian density manifold* (SDM). We provide the geometric calculations in the SDM. We studied the Fokker–Planck equation as the sub-Riemannian gradient flow in the SDM. We derived the equivalence relation between the second-order calculus of the relative entropy in the SDM and the generalized Gamma z calculus.

##### 4.1. Sub-Riemannian Density Manifold

Given a finite-dimensional sub-Riemannian manifold  $(\mathbb{R}^{n+m}, \tau, g_\tau)$  with  $g_\tau = (aa^\top)^\dagger$ , consider the probability density space:

$$\mathcal{P}(\mathbb{R}^{n+m}) = \left\{ \rho(x) \in C^\infty(\mathbb{R}^{n+m}) : \int \rho(x) dx = 1, \quad \rho(x) \geq 0 \right\}.$$

Consider the tangent space at  $\rho \in \mathcal{P}(\mathbb{R}^{n+m})$ :

$$T_\rho \mathcal{P}(\mathbb{R}^{n+m}) = \left\{ \sigma(x) \in C^\infty(\mathbb{R}^{n+m}) : \int \sigma(x) dx = 0 \right\}.$$

We introduce the sub-Riemannian structure in probability density space  $\mathcal{P}(\mathbb{R}^{n+m})$ .

**Definition 3** (sub-Riemannian Wasserstein metric tensor). *The  $L^2$  sub-Riemannian-Wasserstein metric  $g_\rho^{W_a} : T_\rho \mathcal{P}(\mathbb{R}^{n+m}) \times T_\rho \mathcal{P}(\mathbb{R}^{n+m}) \rightarrow \mathbb{R}$  is defined by*

$$g_\rho^{W_a}(\sigma_1, \sigma_2) = \int \left( \sigma_1(x), (-\Delta_\rho^a)^\dagger \sigma_2(x) \right) dx.$$

Here,  $\sigma_1, \sigma_2 \in T_\rho \mathcal{P}(\mathbb{R}^{n+m})$ ,  $(\cdot, \cdot)$  is the metric on  $\mathbb{R}^{n+m}$ , and  $(\Delta_\rho^a)^\dagger : T_\rho \mathcal{P}(\mathbb{R}^{n+m}) \rightarrow T_\rho \mathcal{P}(\mathbb{R}^{n+m})$  is the pseudo-inverse of the sub-elliptic operator:

$$\Delta_\rho^a = \nabla \cdot (\rho aa^\top \nabla).$$

For some special choices of  $a$  as studied in [19] or  $aa^\top$  forming a positive definite matrix, then  $\Delta_\rho^a$  is an elliptic operator. In this case,  $(\mathcal{P}(\mathbb{R}^{n+m}), g^{W_a})$  still forms a Riemannian density manifold. In general, given a sub-Riemannian manifold  $(\mathbb{R}^{n+m}, (aa^\top)^\dagger)$ ,  $\Delta_\rho^a$  is only a sub-elliptic operator. Thus,  $(\mathcal{P}(\mathbb{R}^{n+m}), g^{W_a})$  forms an infinite-dimensional sub-Riemannian manifold.

We next present the sub-Riemannian calculus in  $(\mathcal{P}(\mathbb{R}^{n+m}), g^{W_a})$ , including both geodesics and the Hessian operator in the tangent bundle. Consider an identification map:

$$V : C^\infty(\mathbb{R}^{n+m}) \rightarrow T_\rho \mathcal{P}(\mathbb{R}^{n+m}), \quad V_\Phi = -\Delta_\rho^a \Phi = -\nabla \cdot (\rho aa^\top \nabla \Phi).$$

Here,  $\Phi \in T_\pi \mathcal{P}(\mathbb{R}^{n+m}) = C^\infty(\mathbb{R}^{n+m}) / \sim$ . This  $T_\pi \mathcal{P}(\mathbb{R}^{n+m})$  is the cotangent space in the SDM, and  $\sim$  represents a constant shift relation. Thus,

$$g_\rho^{W_a}(V_{\Phi_1}, V_{\Phi_2}) = \int \Gamma_1(\Phi_1, \Phi_2) \rho(x) dx.$$

In other words,

$$\begin{aligned}
 g_{\rho}^{W_a}(V_{\Phi_1}, V_{\Phi_2}) &= \int V_{\Phi_1}(-\Delta_{\rho}^a)^{\dagger} V_{\Phi_2} dx \\
 &= \int \Phi_1(-\Delta_{\rho}^a)(-\Delta_{\rho}^a)^{\dagger}(-\Delta_{\rho}^a)\Phi_2 dx \\
 &= \int (\Phi_1, -\Delta_{\rho}^a\Phi_2) dx \\
 &= \int \Phi_1(-\nabla \cdot (\rho a a^{\top} \nabla \Phi_2)) dx \\
 &= \int (a^{\top} \nabla \Phi_1, a^{\top} \nabla \Phi_2) \rho dx,
 \end{aligned} \tag{28}$$

where the second equality holds by  $(-\Delta_{\rho}^a)(-\Delta_{\rho}^a)^{\dagger}(-\Delta_{\rho}^a) = -\Delta_{\rho}^a$  and the last equality holds by the integration by parts.

We next derive several basic geometric calculations in the SDM.

**Proposition 8** (Geodesics in the SDM). *The sub-Riemannian geodesics in the cotangent bundle forms*

$$\begin{cases} \partial_t \rho_t + \nabla \cdot (\rho_t a a^{\top} \nabla \Phi_t) = 0, \\ \partial_t \Phi_t + \frac{1}{2}(a^{\top} \nabla \Phi_t, a^{\top} \nabla \Phi_t) = 0. \end{cases} \tag{29}$$

**Proof.** We considered the Lagrangian formulation of geodesics in density. Here, the minimization of the geometric action functional forms

$$\mathcal{L}(\rho_t, \partial_t \rho_t) = \int_0^1 \int \frac{1}{2} (\partial_t \rho_t, (-\Delta_{\rho_t}^a)^{\dagger} \partial_t \rho_t) dx dt,$$

where  $\rho_t = \rho(t, x)$  is a density path with fixed boundary points  $\rho_0, \rho_1$ . Then, the Euler–Lagrange equation in density space forms

$$\frac{\partial}{\partial t} \delta_{\partial_t \rho_t} \mathcal{L}(\rho_t, \partial_t \rho_t) = \delta_{\rho_t} \mathcal{L}(\rho_t, \partial_t \rho_t),$$

where  $\delta_{\partial_t \rho_t}$  is the  $L^2$  first variation with respect to  $\partial_t \rho_t$  and  $\delta_{\rho_t}$  is the  $L^2$  first variation with respect to  $\rho_t$ . Here,

$$\begin{aligned}
 \partial_t \left( (-\Delta_{\rho_t}^a)^{\dagger} \partial_t \rho_t \right) &= \delta_{\rho} \int \frac{1}{2} \left( \partial_t \rho, (-\Delta_{\rho_t}^a)^{\dagger} \partial_t \rho_t \right) dx \\
 &= -\frac{1}{2} \left( a^{\top} \nabla (-\Delta_{\rho}^a)^{\dagger} \partial_t \rho_t, a^{\top} \nabla (-\Delta_{\rho}^a)^{\dagger} \partial_t \rho_t \right),
 \end{aligned} \tag{30}$$

where the last equality uses the following fact:

$$\partial_t \left( \Delta_{\rho_t}^a \right)^{\dagger} = - \left( \Delta_{\rho_t}^a \right)^{\dagger} \cdot \Delta_{\partial_t \rho_t}^a \cdot \left( \Delta_{\rho_t}^a \right)^{\dagger},$$

Denote  $\partial_t \rho_t = -\Delta_{\rho_t}^a \Phi_t$ , then the Euler–Lagrange Equation (30) forms the sub-Riemannian geodesics flow (29). In other words,

$$\partial_t \Phi + \frac{1}{2} (a^{\top} \nabla \Phi, a^{\top} \nabla \Phi) = 0.$$

□

**Proposition 9** (Gradient and Hessian operators in the SDM). *Given a functional  $\mathcal{F}: \mathcal{P}(\mathbb{R}^{n+m}) \rightarrow \mathbb{R}$ , the gradient operator of  $\mathcal{F}$  in  $(\mathcal{P}, g^{W_a})$  satisfies*

$$\text{grad}_{W_a} \mathcal{F}(\rho) = -\nabla \cdot (\rho a a^{\top} \nabla \delta \mathcal{F}(\rho)).$$

The Hessian operator of  $\mathcal{F}$  in  $(\mathcal{P}, g^{W_a})$  satisfies

$$\begin{aligned} & \text{Hess}_{W_a} \mathcal{F}(V_\Phi, V_\Phi) \\ &= \int \int (a(y)^\top \nabla_y)(a(x)^\top \nabla_x) \delta^2 \mathcal{F}(\rho)(x, y) \left( a(x)^\top \nabla_x \Phi(x), a(y)^\top \nabla_y \Phi(y) \right) \rho(x) \rho(y) dx dy \\ &+ \int \text{Hess}_a \delta \mathcal{F}(\rho)(\Phi, \Phi) \rho dx, \end{aligned}$$

where

$$\text{Hess}_a \delta \mathcal{F}(\rho)(\Phi, \Phi) = \frac{1}{2} \left\{ 2\Gamma_1(\Gamma_1(\delta \mathcal{F}, \Phi), \Phi) - \Gamma_1(\Gamma_1(\Phi, \Phi), \delta \mathcal{F}) \right\}.$$

**Proof.** We first derive the sub-Riemannian gradient operator. We recall the identification map by  $-\Delta_\rho^a \Phi = -\nabla \cdot (\rho a a^\top \nabla \Phi)$ . Hence, the gradient operator in the SDM satisfies

$$\begin{aligned} \text{grad}_{W_a} \mathcal{F}(\rho) &= \left( (-\Delta_\rho^a)^\dagger \right) \frac{\delta}{\delta \rho(x)} \mathcal{F}(\rho) \\ &= -\Delta_\rho^a \frac{\delta}{\delta \rho(x)} \mathcal{F}(\rho) \\ &= -\nabla \cdot (\rho a a^\top \nabla \frac{\delta}{\delta \rho(x)} \mathcal{F}(\rho)). \end{aligned}$$

The Hessian operator in the SDM satisfies

$$\text{Hess}_{W_a} \mathcal{F}(\rho)(V_\Phi, V_\Phi) = \frac{d^2}{dt^2} \mathcal{F}(\rho_t)|_{t=0},$$

where  $(\rho_t, \Phi_t)$  satisfies the geodesics Equation (29) with  $\rho_0 = \rho$ ,  $\Phi_0 = \Phi$ . Notice the fact that

$$\begin{aligned} \frac{d}{dt} \mathcal{F}(\rho_t)|_{t=0} &= \int \partial_t \rho_t \delta \mathcal{F}(\rho) dx|_{t=0} \\ &= \int (-\nabla \cdot (\rho a a^\top \nabla \Phi)) \delta \mathcal{F}(\rho) dx \\ &= \int (a^\top \nabla \delta \mathcal{F}(\rho), a^\top \nabla \Phi) \rho dx. \end{aligned}$$

In addition,

$$\begin{aligned} \frac{d^2}{dt^2} \mathcal{F}(\rho_t)|_{t=0} &= \frac{d}{dt} \int (a^\top \nabla \delta \mathcal{F}(\rho_t), a^\top \nabla \Phi_t) \rho_t dx|_{t=0} \\ &= \int \int \delta^2 \mathcal{F}(\rho)(x, y) \partial_t \rho_t(x) \partial_t \rho_t(y) dx dy + \int (a^\top \nabla \delta \mathcal{F}(\rho_t), a^\top \nabla \partial_t \Phi_t) \rho_t dx \\ &+ \int (a^\top \nabla \delta \mathcal{F}(\rho_t), a^\top \nabla \partial_t \Phi_t) \partial_t \rho_t dx|_{t=0} \\ &= \int \int \delta^2 \mathcal{F}(\rho)(x, y) \nabla \cdot (\rho a a^\top \nabla \Phi)(x) \nabla \cdot (\rho a a^\top \nabla \Phi)(y) dx dy \\ &- \frac{1}{2} \int (a^\top \nabla \delta \mathcal{F}(\rho), a^\top \nabla \Gamma_1(\Phi, \Phi)) \rho dx \\ &+ \int \Gamma_1(\Phi, \delta \mathcal{F}(\rho)) \left( -\nabla \cdot (\rho a a^\top \nabla \Phi) \right) dx \\ &= \int \int (a(y)^\top \nabla_y)(a(x)^\top \nabla_x) \delta^2 \mathcal{F}(\rho)(x, y) \left( a(x)^\top \nabla_x \Phi(x), a(y)^\top \nabla_y \Phi(y) \right) \rho(x) \rho(y) dx dy \\ &+ \frac{1}{2} \int \left\{ 2\Gamma_1(\Gamma_1(\delta \mathcal{F}, \Phi), \Phi) - \Gamma_1(\Gamma_1(\Phi, \Phi), \delta \mathcal{F}) \right\} \rho dx, \end{aligned} \tag{31}$$

where the last equality holds by the integration by parts formula.  $\square$

We next show the equivalence relation between the Hessian of the relative entropy in the SDM and the classical Gamma two operator. We first demonstrate the relation among  $L^*$ ,  $\Delta_a$  and the gradient operator of the entropy. In particular, we show that the Fokker–Planck equation is a sub-Riemannian gradient flow in the SDM. Denote the KL divergence as

$$D(\rho) = D_{\text{KL}}(\rho \parallel \pi) = \int \rho(x) \log \frac{\rho(x)}{\pi(x)} dx. \quad (32)$$

**Proposition 10** (Gradient flow). *The negative gradient operator in  $(\mathcal{P}, g^{W_a})$  forms*

$$-\text{grad}_{W_a} D(\rho) = L^* \rho = \nabla \cdot (\rho a a^T \nabla \log \frac{\rho}{\pi}).$$

*In addition, the sub-Riemannian gradient flow of  $D(\rho)$  in  $(\mathcal{P}, g^{W_a})$  forms the Fokker–Planck equation:*

$$\partial_t \rho = \nabla \cdot (\rho a a^T \nabla \log \frac{\rho}{\pi}). \quad (33)$$

**Proof.** We first derive the sub-Riemannian gradient operator of the entropy and relative entropy. Notice that

$$\delta_{\rho(x)} D(\rho) = \log \rho(x) + 1 - \log \pi(x).$$

Thus,

$$\text{grad}_{W_a} D(\rho) = -\nabla \cdot (\rho a a^T \nabla \log \rho) + \nabla \cdot (\rho a a^T \nabla \log \pi),$$

where  $\rho \nabla \log \rho = \rho \frac{\nabla \rho}{\rho} = \nabla \rho$ . Following the gradient flow formulation:

$$\frac{\partial \rho_t}{\partial t} = -\text{grad}_{W_a} D(\rho_t) = L^* \rho_t,$$

we finish the derivation of (33).  $\square$

We next demonstrate that the Hessian of the relative entropy (KL divergence) is equivalent to the classical Bakry–Émery calculus.

**Proposition 11** (Hessian of entropy and Bakry–Émery calculus). *Given  $\Phi_1, \Phi_2 \in C^\infty(\mathbb{R}^{n+m})$ , then*

$$\text{Hess}_{W_a} D(\rho)(V_\Phi, V_\Phi) = \int \Gamma_2(\Phi, \Phi) \rho(x) dx.$$

**Proof.** We first derive the Hessian of  $D(\rho)$  in the SDM. Notice the fact that  $\delta^2 D(\rho)(x, y) = \frac{1}{\rho} \delta_{x=y}$ . For simplicity, we denote  $\delta^2 D(\rho) = \frac{1}{\rho(x)}$ . By using (31), we have

$$\text{Hess}_{W_a} D(\rho)(V_\Phi, V_\Phi) = \int \delta^2 D(\rho)(x) \left( \nabla \cdot (\rho a a^T \nabla \Phi) \right)^2 dx \quad (a)$$

$$- \frac{1}{2} \int (a^T \nabla \delta D(\rho), a^T \nabla \Gamma_1(\Phi, \Phi)) \rho dx \quad (b) \quad (34)$$

$$+ \int \Gamma_1(\Phi, \delta D(\rho)) \left( -\nabla \cdot (\rho a a^T \nabla \Phi) \right) dx. \quad (c)$$

We next rewrite (34) into the iterative Gamma calculus. We first show that

$$\begin{aligned}
 (a) + (c) &= \int \left( \delta^2 D(\rho) \nabla \cdot (\rho a a^T \nabla \Phi) - \Gamma_1(\Phi, \delta D(\rho)) \right) \nabla \cdot (\rho a a^T \nabla \Phi) \rho dx \\
 &= \int \left( \frac{1}{\rho} \nabla \cdot (\rho a a^T \nabla \Phi) - (a a^T \nabla \log \frac{\rho}{\pi}, \nabla \Phi) \right) \nabla \cdot (\rho a a^T \nabla \Phi) \rho dx \\
 &= \int \left( \frac{1}{\rho} (\nabla \rho, a a^T \nabla \Phi) + \nabla \cdot (a a^T \nabla \Phi) - (a a^T \nabla \log \frac{\rho}{\pi}, \nabla \Phi) \right) \nabla \cdot (\rho a a^T \nabla \Phi) \rho dx \\
 &= \int \left( (\nabla \log \rho, a a^T \nabla \Phi) + \nabla \cdot (a a^T \nabla \Phi) \right. \\
 &\quad \left. - (\nabla \log \rho, a a^T \nabla \Phi) + (a a^T \nabla \log \pi, \nabla \Phi) \right) \nabla \cdot (\rho a a^T \nabla \Phi) \rho dx \\
 &= \int \left( (\nabla \cdot (a a^T \nabla \Phi) + (a a^T \nabla \log \pi, \nabla \Phi)) \right) \nabla \cdot (\rho a a^T \nabla \Phi) \rho dx \\
 &= \int L \Phi \nabla \cdot (\rho a a^T \nabla \Phi) dx \\
 &= - \int \Gamma_1(L \Phi, \Phi) \rho dx,
 \end{aligned}$$

where the fourth equality uses the fact that  $\frac{\nabla \rho}{\rho} = \nabla \log \rho$ , while the last equality follows the integration by parts.

We secondly show that

$$\begin{aligned}
 (b) &= - \frac{1}{2} \int (a^T \nabla \delta D(\rho), a^T \nabla \Gamma_1(\Phi, \Phi)) \rho dx \\
 &= \frac{1}{2} \int \Gamma_1(\Phi, \Phi) \nabla \cdot (\rho a a^T \nabla \delta D(\rho)) dx \\
 &= \frac{1}{2} \int \Gamma_1(\Phi, \Phi) L^* \rho dx \\
 &= \frac{1}{2} \int L \Gamma_1(\Phi, \Phi) \rho dx,
 \end{aligned}$$

where the second equality applies the fact that  $L^* \rho = \nabla \cdot (\rho a a^T \nabla \delta D)$ , while the last inequality uses the dual-relation between Kolmogorov operators  $L$  and  $L^*$  in  $L^2(\rho)$ , i.e.,

$$\int f(x) L^* \rho(t, x) dx = \int L f(x) \rho(t, x) dx, \quad \text{for any } f \in C_c^\infty(\mathbb{R}^{n+m}).$$

Combining the equality of (a), (b), (c), we prove the result.  $\square$

**Remark 10.** We remark that the above formulations in terms of  $a a^T$  hold for both Riemannian and sub-Riemannian density manifolds. Here, the major difference is whether matrix function  $a$  is full rank or degenerate. In this sense, all formulas derived in this subsection recover the classical Bakry–Émery calculus. However, the classical Hessian operator of the entropy is not enough to study the convergence behavior of degenerate diffusion processes. Briefly, we use a modified Lyapunov functional and derive a tensor for the gradient flow in the SDM. It provides the convergence rate of the degenerate diffusion process.

#### 4.2. Gamma z Calculus via Second-Order Calculus of Relative Entropy in SDM

In this subsection, we introduce the motivation of our new Gamma z calculus from the SDM viewpoint. Consider the SDM gradient flow (33):

$$\partial_t \rho_t = \Delta_{\rho_t}^a \delta D(\rho_t).$$

When  $a$  is a degenerate matrix, the classical relative Fisher information  $I_a$  may not be the Lyapunov functional. In other words, along the gradient flow, it is possible that  $\frac{d}{dt} I_a(\rho_t) \geq 0$ .

To handle this issue, a new Lyapunov function is considered. It is to add a new direction  $z$  into the relative Fisher information functional. Denote  $\Delta_\rho^z = \nabla \cdot (\rho z z^\top \nabla)$  and  $I_z(\rho) = \int (\delta D, (-\Delta_\rho^z) \delta D) dx$ . Construct

$$I_{a,z}(\rho) := I_a(\rho) + I_z(\rho) = \int (\delta D, (-\Delta_\rho^a - \Delta_\rho^z) \delta D) dx.$$

We next prove the following proposition.

**Proposition 12.**

$$\frac{d}{dt} I_{a,z}(\rho_t) = -2 \int (\Gamma_2(\delta D, \delta D) + \tilde{\Gamma}_2^z(\delta D, \delta D)) \rho_t dx,$$

where

$$\tilde{\Gamma}_2^z(\Phi, \Phi) := \frac{1}{2} L(\Gamma_1^z(\Phi, \Phi)) - \Gamma_1(L_z \Phi, \Phi), \quad (35)$$

with the notation  $\Delta_z = \nabla \cdot (z z^\top \nabla)$  and  $L_z = \nabla \cdot (z z^\top \nabla) + (\nabla \log \pi, z z^\top \nabla)$ .

**Proof.** For the simplicity of notation, we denote  $\rho = \rho_t$ . Notice the fact that

$$\frac{d}{dt} I_{a,z}(\rho) = \frac{d}{dt} I_a(\rho) + \frac{d}{dt} I_z(\rho).$$

From Proposition 11, we have

$$\begin{aligned} \frac{d^2}{dt^2} I_a(\rho) &= -2 \text{Hess}_{g_a} D(V_{\delta D}, V_{\delta D}) \\ &= -2 \int \Gamma_2(\delta D, \delta D) \rho dx. \end{aligned}$$

□

We only need to show the following claim.

**Claim:**

$$\frac{d}{dt} I_z(\rho) = -2 \int \tilde{\Gamma}_2^z(\delta D, \delta D) \rho dx.$$

**Proof of Claim.** The proof is similar to the ones in Proposition 11. We need to take care of the  $z$  direction. Notice that

$$\begin{aligned} \frac{d}{dt} I_z(\rho) &= 2 \int \delta^2 D \left( (-\Delta_\rho^z \delta D), \partial_t \rho \right) dx + \int (\nabla \delta D, z z^\top \nabla \delta D) \partial_t \rho dx \\ &= 2 \int \delta^2 D \left( (-\Delta_\rho^z \delta D), \Delta_\rho^a \delta D \right) dx + \int (\nabla \delta D, z z^\top \nabla \delta D) (\Delta_\rho^a \delta D) dx \end{aligned}$$

$$= -2 \int \frac{1}{\rho} \nabla \cdot (\rho a a^\top \nabla \delta D) \nabla \cdot (\rho z z^\top \nabla \delta D) dx \quad (I)$$

$$+ \int (\nabla \delta D, z z^\top \nabla \delta D) \nabla \cdot (\rho a a^\top \nabla \delta D) dx \quad (II)$$

We next estimate (I) and (II) separately. For (I), we notice the fact that

$$\begin{aligned} \frac{1}{\rho} \nabla \cdot (\rho z z^\top \nabla \delta D) &= (\nabla \log \rho, z z^\top \nabla \delta D) + \nabla \cdot (z z^\top \nabla \delta D) \\ &= (\nabla \log \frac{\rho}{\pi}, z z^\top \nabla \delta D) + (\nabla \log \pi, z z^\top \nabla \delta D) + \nabla \cdot (z z^\top \nabla \delta D) \\ &= (\nabla \delta D, z z^\top \nabla \delta D) + (\nabla \log \pi, z z^\top \nabla \delta D) + \nabla \cdot (z z^\top \nabla \delta D). \end{aligned}$$

Thus,

$$\begin{aligned}
 (I) &= -2 \int \frac{1}{\rho} \nabla \cdot (\rho a a^T \nabla \delta D) \nabla \cdot (\rho z z^T \nabla \delta D) dx \\
 &= -2 \int \nabla \cdot (\rho a a^T \nabla \delta D) \left( (\nabla \delta D, z z^T \nabla \delta D) \right. \\
 &\quad \left. + (\nabla \log \pi, z z^T \nabla \delta D) + \nabla \cdot (z z^T \nabla \delta D) \right) dx \\
 &= -2 \int (\nabla \delta D, z z^T \nabla \delta D) L^* \rho \\
 &\quad + \nabla \cdot (\rho a a^T \nabla \delta D) \left( (\nabla \log \pi, z z^T \nabla \delta D) + \nabla \cdot (z z^T \nabla \delta D) \right) dx \\
 &= -2 \int L(\nabla \delta D, z z^T \nabla \delta D) \rho dx \\
 &\quad + 2 \int \left\{ \nabla \left( (\nabla \log \pi, z z^T \nabla \delta D) + \nabla \cdot (z z^T \nabla \delta D) \right), a a^T \nabla \delta D \right\} \rho dx,
 \end{aligned}$$

where the last equality holds by integration by parts.

For (II), we have

$$\begin{aligned}
 (II) &= \int (\nabla \delta D, z z^T \nabla \delta D) \nabla \cdot (\rho a a^T \nabla \delta D) dx \\
 &= \int (\nabla \delta D, z z^T \nabla \delta D) L^* \rho dx \\
 &= \int L(\nabla \delta D, z z^T \nabla \delta D) \rho dx.
 \end{aligned}$$

Combining (I) and (II), we have

$$\tilde{\Gamma}_2^z(\Phi, \Phi) = \frac{1}{2} L(\Gamma_1^z(\Phi, \Phi)) - \Gamma_1(\Delta_z \Phi, \Phi) - \Gamma_1(\Gamma_1^z(\log \pi, \Phi), \Phi).$$

Using the notation  $L_z = \Delta_z + (\nabla \log \pi, z z^T \nabla)$ , we finish the proof.  $\square$

We next prove that  $\tilde{\Gamma}_2^z$  and  $\Gamma_2^{z, \pi}$  in Definition 1 agree with each other in the weak form along the gradient flow.

**Proposition 13.** Denote  $\Phi = \delta \mathcal{D}(\rho)$ , then

$$\int \tilde{\Gamma}_2^z(\Phi, \Phi) \rho dx = \int \Gamma_2^{z, \pi}(\Phi, \Phi) \rho dx.$$

**Proof.** To prove the proposition, we rewrite  $\tilde{\Gamma}_2^z$  as follows.

$$\begin{aligned}
 \tilde{\Gamma}_2^z(\Phi, \Phi) &= \frac{1}{2} L(\Gamma_1^z(\Phi, \Phi)) - \Gamma_1(L_z \Phi, \Phi) \\
 &= \frac{1}{2} L(\Gamma_1^z(\Phi, \Phi)) - \Gamma_1^z(L \Phi, \Phi) \\
 &\quad + \Gamma_1^z(L \Phi, \Phi) - \Gamma_1(L_z \Phi, \Phi).
 \end{aligned}$$

$\square$

Here, we need to prove the following equality.

**Claim:**

$$\begin{aligned}
 &\int \left\{ \Gamma_1^z(L \Phi, \Phi) - \Gamma_1(L_z \Phi, \Phi) \right\} \rho dx \\
 &= \int \rho \left\{ \frac{1}{\pi} \nabla \cdot (z z^T \pi (\nabla \Phi, \nabla (a a^T) \nabla \Phi)) - \frac{1}{\pi} \nabla \cdot (a a^T \pi (\nabla \Phi, \nabla (z z^T) \nabla \Phi)) \right\} dx.
 \end{aligned}$$

**Proof of Claim.** For the simplicity of notation, let

$$L^* \rho = \nabla \cdot (aa^T \pi \nabla \frac{\rho}{\pi}) = \nabla \cdot (\rho aa^T \nabla \log \frac{\rho}{\pi})$$

and

$$L_z^* \rho = \nabla \cdot (zz^T \pi \nabla \frac{\rho}{\pi}) = \nabla \cdot (\rho zz^T \nabla \log \frac{\rho}{\pi}).$$

The following property is also used in the proof. For any smooth test function  $f$  and  $\Phi = \log \frac{\rho}{\pi}$ , then

$$\int L_z^* \rho f dx = - \int \Gamma_1^z(f, \Phi) \rho dx, \quad \int L^* \rho f dx = - \int \Gamma_1(f, \Phi) \rho dx.$$

Notice that  $\Phi = \log \frac{\rho}{\pi}$ , then

$$\begin{aligned} & \int \Gamma_1^z(L\Phi, \Phi) \rho dx \\ &= \int (\nabla(\nabla \cdot (aa^T \nabla \Phi)) - (A, \nabla \Phi), zz^T \nabla \Phi) \rho dx \\ &= \int (\nabla(\nabla \cdot (aa^T \nabla \Phi)), zz^T \nabla \Phi) \rho dx - \int (\nabla(A, \nabla \Phi), zz^T \nabla \Phi) \rho dx. \end{aligned}$$

(a1) (a2)

Here,

$$\begin{aligned} (a1) &= \int (\nabla(\nabla \cdot (aa^T \nabla \Phi)), zz^T \nabla \Phi) \rho dx \\ &= - \int \nabla \cdot (aa^T \nabla \Phi) \nabla \cdot (\rho zz^T \nabla \Phi) dx \\ &= - \int \nabla \cdot (aa^T \nabla \log \frac{\rho}{\pi}) \nabla \cdot (\rho zz^T \nabla \log \frac{\rho}{\pi}) dx \\ &= - \int \nabla \cdot (\frac{1}{\rho} aa^T \pi \nabla \frac{\rho}{\pi}) \nabla \cdot (\rho zz^T \nabla \log \frac{\rho}{\pi}) dx \\ &= - \int \left\{ (\nabla \frac{1}{\rho}, aa^T \pi \nabla \frac{\rho}{\pi}) + \frac{1}{\rho} \nabla \cdot (aa^T \pi \nabla \frac{\rho}{\pi}) \right\} \nabla \cdot (\rho zz^T \nabla \log \frac{\rho}{\pi}) dx \\ &= - \int (\nabla \frac{1}{\rho}, aa^T \pi \nabla \frac{\rho}{\pi}) L_z^* \rho dx - \int \frac{1}{\rho} L^* \rho L_z^* \rho dx \\ &= \int \frac{1}{\rho^2} (\nabla \rho, aa^T \pi \nabla \frac{\rho}{\pi}) L_z^* \rho dx - \int \frac{1}{\rho} L^* \rho L_z^* \rho dx \\ &= \int (\nabla \log \rho, aa^T \nabla \log \frac{\rho}{\pi}) L_z^* \rho dx - \int \frac{1}{\rho} L^* \rho L_z^* \rho dx \\ &= \int (\nabla \log \frac{\rho}{\pi}, aa^T \nabla \log \frac{\rho}{\pi}) L_z^* \rho dx \\ &\quad + \int (\nabla \log \pi, aa^T \nabla \log \frac{\rho}{\pi}) L_z^* \rho dx - \int \frac{1}{\rho} L^* \rho L_z^* \rho dx \\ &= - \int (\nabla(\nabla \log \frac{\rho}{\pi}, aa^T \nabla \log \frac{\rho}{\pi}), zz^T \nabla \log \frac{\rho}{\pi}) \rho dx \\ &\quad - \int \Gamma_1^z((\nabla \log \pi, aa^T \nabla \log \frac{\rho}{\pi}), \log \frac{\rho}{\pi}) \rho dx - \int \frac{1}{\rho} L^* \rho L_z^* \rho dx \\ &= - \int ((\nabla \log \frac{\rho}{\pi}, \nabla(aa^T) \nabla \log \frac{\rho}{\pi}), zz^T \nabla \log \frac{\rho}{\pi}) \pi dx \\ &\quad - \int 2 \nabla^2 \log \frac{\rho}{\pi} (aa^T \nabla \log \frac{\rho}{\pi}, zz^T \nabla \log \frac{\rho}{\pi}) \rho dx \\ &\quad - \int \Gamma_1^z((\nabla \log \pi, aa^T \nabla \log \frac{\rho}{\pi}), \log \frac{\rho}{\pi}) \rho dx - \int \frac{1}{\rho} L^* \rho L_z^* \rho dx \end{aligned}$$



$$\begin{aligned}
&= \int \nabla \cdot \left( zz^T \pi \left( (\nabla \log \frac{\rho}{\pi}, \nabla(aa^T) \nabla \log \frac{\rho}{\pi}) \right) \right) \frac{1}{\pi} \rho dx \\
&\quad - \int 2 \nabla^2 \log \frac{\rho}{\pi} \left( aa^T \nabla \log \frac{\rho}{\pi}, zz^T \nabla \log \frac{\rho}{\pi} \right) \rho dx \\
&\quad - \int \Gamma_1^z((\nabla \log \pi, aa^T \nabla \log \frac{\rho}{\pi}), \log \frac{\rho}{\pi}) \rho dx - \int \frac{1}{\rho} L^* \rho L_z^* \rho dx.
\end{aligned}$$

Notice the fact that

$$\begin{aligned}
(a2) &= - \int (\nabla(A, \nabla \Phi), zz^T \nabla \Phi) \rho dx \\
&= \int (\nabla(\nabla \log \pi, aa^T \nabla \Phi), zz^T \nabla \Phi) \rho dx \\
&= \int \Gamma_1(\Gamma_1^z(\Phi, \Phi), \Phi) \rho dx.
\end{aligned}$$

Hence,

$$\begin{aligned}
\int \Gamma_1^z(L\Phi, \Phi) \rho dx &= (a1) + (a2) \\
&= \int \nabla \cdot \left( zz^T \pi \left( (\nabla \log \frac{\rho}{\pi}, \nabla(aa^T) \nabla \log \frac{\rho}{\pi}) \right) \right) \frac{1}{\pi} \rho dx \\
&\quad - \int 2 \nabla^2 \log \frac{\rho}{\pi} \left( aa^T \nabla \log \frac{\rho}{\pi}, zz^T \nabla \log \frac{\rho}{\pi} \right) \rho dx \\
&\quad - \int \frac{1}{\rho} L^* \rho L_z^* \rho dx.
\end{aligned}$$

Similarly, by switching  $a$  and  $z$ , we have

$$\begin{aligned}
\int \Gamma_1(L_z \Phi, \Phi) \rho dx &= \int \nabla \cdot \left( aa^T \pi \left( (\nabla \log \frac{\rho}{\pi}, \nabla(zz^T) \nabla \log \frac{\rho}{\pi}) \right) \right) \frac{1}{\pi} \rho dx \\
&\quad - \int 2 \nabla^2 \log \frac{\rho}{\pi} \left( aa^T \nabla \log \frac{\rho}{\pi}, zz^T \nabla \log \frac{\rho}{\pi} \right) \rho dx \\
&\quad - \int \frac{1}{\rho} L^* \rho L_z^* \rho dx.
\end{aligned}$$

Combining the above derivation, we finish the proof.  $\square$

**Remark 11.** From the proof, we can show the following identity: denote  $\Phi = \delta D$ , then

$$\begin{aligned}
&\int \left( \Gamma_1^z(L\Phi, \Phi) - \Gamma_1(L_z \Phi, \Phi) \right) \rho dx \\
&= \int \left( \Gamma_1(\Gamma_1^z(\Phi, \Phi), \Phi) - \Gamma_1^z(\Gamma_1(\Phi, \Phi), \Phi) \right) \rho dx.
\end{aligned}$$

Therefore, it is clear that, if the commutative assumption  $\Gamma_1(\Gamma_1^z(\Phi, \Phi), \Phi) = \Gamma_1^z(\Gamma_1(\Phi, \Phi), \Phi)$  holds, the above quantity equals zero. In this case,

$$\int \Gamma_2^{z,\pi}(\Phi, \Phi) \rho dx = \int \Gamma_2^z(\Phi, \Phi) \rho dx.$$

This means that, under the commutative assumption, the generalized Gamma  $z$  calculus agrees with the classical one [2] in the weak sense.

With the generalized Gamma  $z$  calculus, we are ready to prove the convergence properties and functional inequalities for degenerate drift–diffusion processes.

**Proposition 14.** Suppose  $\Gamma_2 + \Gamma_2^{z,\pi} \succeq \kappa(\Gamma_1 + \Gamma_1^z)$  with  $\kappa > 0$ . Denote  $\rho_t$  as the solution of the sub-Riemannian gradient flow (33), then

$$\frac{d}{dt} \left( I_a(\rho_t) + I_z(\rho_t) \right) \leq -2\kappa \left( I_a(\rho_t) + I_z(\rho_t) \right).$$

In addition, the z-log-Sobolev inequalities holds:

$$\int_{\mathbb{R}^{n+m}} \rho \log \frac{\rho}{\pi} dx \leq \frac{1}{2\kappa} I_{a,z}(\rho),$$

for any smooth density function  $\rho$ .

Finally,

$$\int_{\mathbb{R}^{n+m}} |\rho(t, x) - \pi(x)| dx \leq \sqrt{2D(\rho_0)} e^{-\kappa t}.$$

**Proof.** Here, the proof is very similar to the one in the previous section. Again, consider the sub-Riemannian gradient flow in the SDM.

$$\partial_t \rho_t = -\text{grad}_{W_a} D(\rho_t).$$

We know that the log-Sobolev inequality relates to the ratio of  $\frac{d}{dt} D(\rho_t)$  and  $\frac{d^2}{dt^2} D(\rho_t)$ . If we cannot estimate a ratio  $\kappa > 0$ , then

$$\frac{d}{dt} I_a(\rho_t) \leq -2\kappa I_a(\rho_t).$$

We construct the other Lyapunov function:

$$I_{a,z}(\rho) = I_a(\rho) + I_z(\rho).$$

Thus, along the SDM gradient flow (33), we have

$$\frac{d}{dt} I_{a,z}(\rho_t) = -2 \int \left( \Gamma_2(\delta D, \delta D) + \Gamma_2^{z,\pi}(\delta D, \delta D) \right) \rho_t dx.$$

If  $\Gamma_2 + \Gamma_2^{z,\pi} \succeq \kappa(\Gamma_1 + \Gamma_1^z)$ , then

$$\frac{d}{dt} I_{a,z}(\rho_t) \leq -2\kappa I_{a,z}(\rho_t). \quad (36)$$

The convergence result follows directly from Gronwall's equality.

We next prove the z-log-Sobolev inequality. Since

$$-\frac{d}{dt} D(\rho_t) = I_a(\rho_t) \leq I_{a,z}(\rho_t),$$

then (36) implies the fact that, denoting  $\rho_0 = \rho$ , then

$$\begin{aligned} -I_{a,z}(\rho) &= \int_0^\infty \frac{d}{dt} I_{a,z}(\rho_t) dt \\ &\leq -2\kappa \int_0^\infty I_{a,z}(\rho_t) dt = -2\kappa \int_0^\infty \left( I_a(\rho_t) + I_z(\rho_t) \right) dt \\ &\leq -2\kappa \int_0^\infty I_a(\rho_t) dt \\ &= -2\kappa \int_0^\infty \left( -\frac{d}{dt} D(\rho_t) \right) dt \\ &= -2\kappa D(\rho). \end{aligned}$$

Thus,  $I_{a,z}(\rho) \geq 2\kappa D(\rho)$ . Hence, we prove all the results by the fact that  $\mathfrak{R} \succeq \kappa(\Gamma_1 + \Gamma_1^z)$  implies  $\Gamma_2 + \Gamma_2^{z,\pi} \succeq \kappa(\Gamma_1 + \Gamma_1^z)$ . In other words, the generalized Gamma  $z$  calculus implies the  $z$ -log-Sobolev equality (zLSI):

$$\mathfrak{R} \succeq \kappa(\Gamma_1 + \Gamma_1^z) \Rightarrow \frac{d}{dt} I_{a,z}(\rho_t) \leq -2\kappa I_{a,z}(\rho_t) \Rightarrow \text{zLSI}.$$

We last prove the exponential convergence in the  $L^1$  distance. Notice that

$$D_{\text{KL}}(\rho_t \| \pi) \leq \frac{1}{2\lambda} \mathcal{I}_{a,z}(\rho_t \| \pi) \leq \frac{1}{2\lambda} e^{-2\lambda t} \mathcal{I}_{a,z}(\rho_0 \| \pi).$$

We apply an inequality between the KL divergence and  $L_1$  distance. In other words,

$$\int_{\mathbb{R}^{n+m}} |\rho(t, x) - \pi(x)| dx \leq \sqrt{2D_{\text{KL}}(\rho \| \pi)}.$$

This finishes the proof.  $\square$

**Remark 12.** It is worth mentioning that our derivation of the Gamma  $z$  calculus is not a direct Hessian operator of the entropy in the SDM. In fact, it combines both the second-order calculus in the SDM and the property of the  $L^2$  Hessian operator of the entropy. See similar relations in the mean-field Bakry–Émery calculus [50].

## 5. Generalized Gamma $z$ Calculus

In this section, we introduce the generalized Gamma  $z$  calculus. For any smooth functions  $f, g : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ , the diffusion operator associated with SDE (2) is denoted as

$$Lf = \Delta_a f - A \nabla f + b \nabla f,$$

where we denote  $A = a \otimes \nabla a$  and

$$\Delta_a f = \nabla \cdot (a a^T \nabla f).$$

When  $b = 0$ , we denote the diffusion operator as

$$\tilde{L}f = \Delta_a f - a \otimes \nabla a \nabla f.$$

We first define the Carré de Champ operator  $\Gamma_1$  associated with the above second-order diffusion operators. It is easy to check that  $\Delta_a$ ,  $\tilde{L}$ , and  $L$  share the same  $\Gamma_1$ :

$$\Gamma_1(f, g) = \langle a^T \nabla f, a^T \nabla g \rangle_{\mathbb{R}^n}. \quad (37)$$

Similarly, we introduce the  $\Gamma_1^z$  operator in the direction of  $z = (z_1, \dots, z_m)$  below:

$$\Gamma_1^z = \langle z^T \nabla f, z^T \nabla g \rangle_{\mathbb{R}^m}, \quad (38)$$

Next, we define the iterative  $\Gamma_2$  and  $\Gamma_2^z$  for operator  $L$  ( $\tilde{L}$ , respectively) below:

$$\Gamma_{2,L}(f, f) = \frac{1}{2} L \Gamma_1^z(f, f) - \Gamma_1^z(Lf, f). \quad (39)$$

$$\Gamma_{2,L}^z(f, f) = \frac{1}{2} L \Gamma_1^z(f, f) - \Gamma_1^z(Lf, f). \quad (40)$$

**Definition 4.** We define the generalized Gamma  $z$  for operator  $L$  below:

$$\Gamma_{2,L}^{z,\pi}(f, f) = \Gamma_{2,L}^z(f, f) + \mathbf{div}_z^\pi(\Gamma_{\nabla(aa^T)} f, f) - \mathbf{div}_a^\pi(\Gamma_{\nabla(zz^T)} f, f), \quad (41)$$

For matrices  $a \in \mathbb{R}^{n \times (n+m)}$  and  $z \in \mathbb{R}^{m \times (n+m)}$ , we denote the divergence operator as

$$\begin{aligned}\mathbf{div}_z^\pi(\Gamma_{\nabla(aa^\top)}f, f) &= \frac{\nabla \cdot (zz^\top \pi \Gamma_{\nabla(aa^\top)}(f, f))}{\pi}, \\ \mathbf{div}_a^\pi(\Gamma_{\nabla(zz^\top)}f, f) &= \frac{\nabla \cdot (aa^\top \pi \Gamma_{\nabla(zz^\top)}(f, f))}{\pi},\end{aligned}$$

and

$$\Gamma_{\nabla(aa^\top)}(f, f) = \langle \nabla f, \nabla(aa^\top) \nabla f \rangle, \quad \text{and} \quad \Gamma_{\nabla(zz^\top)}(f, f) = \langle \nabla f, \nabla(zz^\top) \nabla f \rangle.$$

Here, we denote  $\pi$  as the invariant distribution associated with the operator  $L$ .

**Remark 13.** In particular, we have the following local coordinates representation.

$$\begin{aligned}\langle \nabla f, \nabla(aa^\top) \nabla f \rangle &= \langle \nabla f, \frac{\partial}{\partial x_{\hat{k}}} (aa^\top) \nabla f \rangle_{\hat{k}=1}^{n+m} = 2 \langle a^\top \nabla f, \frac{\partial}{\partial x_{\hat{k}}} a^\top \nabla f \rangle_{\hat{k}=1}^{n+m} \\ &= \left( 2 \sum_{i=1}^n \sum_{i'=1}^{n+m} \frac{\partial}{\partial x_{\hat{k}}} a_{ii'}^\top \frac{\partial f}{\partial x_i} a_{ii'}^\top \frac{\partial f}{\partial x_{i'}} \right)_{\hat{k}=1}^{n+m}, \\ \langle \nabla f, \nabla(zz^\top) \nabla f \rangle &= \left( 2 \sum_{j=1}^m \sum_{j'=1}^{n+m} \frac{\partial}{\partial x_{\hat{k}}} z_{jj'}^\top \frac{\partial f}{\partial x_j} z_{jj'}^\top \frac{\partial f}{\partial x_{j'}} \right)_{\hat{k}=1}^{n+m}.\end{aligned}\quad (42)$$

We first present the following key lemmas.

**Lemma 10.**

$$\mathbf{div}_z^\pi(\Gamma_{\nabla(aa^\top)}f, f) - \mathbf{div}_a^\pi(\Gamma_{\nabla(zz^\top)}f, f) = \mathfrak{R}^\pi(f, f) + 2G^\top X,$$

where  $X, G$  are defined in Notation 1 and  $\mathfrak{R}^\pi$  is defined in Definition 2.

**Lemma 11.**

$$\Gamma_{2, \tilde{L}}(f, f) = X^\top Q^\top QX + 2D^\top QX + 2C^\top X + D^\top D + \mathfrak{R}_a(\nabla f, \nabla f),$$

where  $Q, X, C, D$  are introduced in Notation 1 and  $\mathfrak{R}_a$  is defined in Definition 2.

**Lemma 12.**

$$\Gamma_{2, \tilde{L}}^z(f, f) = X^\top P^\top PX + 2E^\top PX + 2F^\top X + E^\top E + \mathfrak{R}_z(\nabla f, \nabla f).$$

where  $P, X, F, E$  are introduced in Notation 1 and  $\mathfrak{R}_z$  is defined in Definition 2.

We then have the following main theorem. In order to distinguish the operators  $L$  and  $\tilde{L}$ , we rewrite Theorem 1 as below, and with some abuse of notation, we denote  $\Gamma_2(f, f) = \Gamma_{2, L}(f, f)$  and  $\Gamma_2^{z, \pi}(f, f) = \Gamma_{2, L}^{z, \pi}(f, f)$ .

**Theorem 3** (z-Bochner's formula). For smooth function  $f : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ , assume that Assumption 1 holds, then

$$\Gamma_{2, L}(f, f) + \Gamma_{2, L}^{z, \pi}(f, f) = \|\mathfrak{Hess}_{a, z} f\|^2 + \mathfrak{R}(\nabla f, \nabla f),$$

where

$$\begin{aligned}\|\mathfrak{Hess}_{a,z}f\|^2 &= [X + \Lambda_1]^\top Q^\top Q[X + \Lambda_1] + [X + \Lambda_2]^\top P^\top P[X + \Lambda_2] \\ \Re(\nabla f, \nabla f) &= -\Lambda_1^\top Q^\top Q\Lambda_1 - \Lambda_2^\top P^\top P\Lambda_2 + D^\top D + E^\top E \\ &\quad + \Re_{ab}(\nabla f, \nabla f) + \Re_{zb}(\nabla f, \nabla f) + \Re_\pi(\nabla f, \nabla f).\end{aligned}$$

All the terms are defined in Notation 1 and Definition 2.

**Proof.** By Definition 4 and Formulae (39) and (40), we have

$$\begin{aligned}\Gamma_{2,L}(f, f) + \tilde{\Gamma}_{2,L}^{z,\pi}(f, f) \\ = \Gamma_{2,L}(f, f) + \Gamma_{2,L}^z(f, f) + \mathbf{div}_z^\pi(\Gamma_{\nabla(aa^\top)}f, f) - \mathbf{div}_a^\pi(\Gamma_{\nabla(zz^\top)}f, f).\end{aligned}$$

We compute the above terms explicitly in the following four steps.

**Step 1:**

$$\begin{aligned}\Gamma_{2,L}(f, f) &= \frac{1}{2}(L\Gamma_{1,L}(f, f) - 2\Gamma_{1,L}(Lf, f)) \\ &= \frac{1}{2}\Delta_a\Gamma_1(f, f) - \frac{1}{2}A\nabla\Gamma_1(f, f) + \frac{1}{2}b\nabla\Gamma_1(f, f) \\ &\quad - \Gamma_1((\Delta_a - A\nabla + b\nabla)f, f) \\ &= \Gamma_{2,\tilde{L}}(f, f) + [\frac{1}{2}b\nabla\Gamma_1(f, f) - \Gamma_1(b\nabla f, f)].\end{aligned}$$

The term  $\Gamma_{2,\tilde{L}}(f, f)$  follows from Lemma 11. We are left with the other two terms:

$$\begin{aligned}\frac{1}{2}b\nabla\Gamma_1(f, f) &= \frac{1}{2}\sum_{\hat{k}=1}^{n+m} b_{\hat{k}} \frac{\partial}{\partial x_{\hat{k}}} (\langle a^\top \nabla f, a^\top \nabla f \rangle_{\mathbb{R}^n}) \\ &= \sum_{\hat{k}, \hat{i}=1}^{n+m} \sum_{i=1}^n (b_{\hat{k}} \frac{\partial a_{ii}^\top}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}} + b_{\hat{k}} a_{ii}^\top \frac{\partial^2 f}{\partial x_{\hat{k}} \partial x_{\hat{i}}}) (a^\top \nabla f)_i,\end{aligned}$$

and

$$\begin{aligned}-\Gamma_1(b\nabla f, f) &= -\langle a^\top \nabla(b\nabla f), a^\top \nabla f \rangle_{\mathbb{R}^n} \\ &= -\sum_{i=1}^n \sum_{\hat{k}, \hat{i}=1}^{n+m} (a_{ii}^\top \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} + a_{ii}^\top b_{\hat{k}} \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}}) (a^\top \nabla f)_i.\end{aligned}$$

**Step 2:**

$$\begin{aligned}\Gamma_{2,L}^z(f, f) &= \frac{1}{2}(L\Gamma_1^z(f, f) - 2\Gamma_1^z(Lf, f)) \\ &= \frac{1}{2}\Delta_a\Gamma_1^z(f, f) - \frac{1}{2}A\nabla\Gamma_1^z(f, f) + \frac{1}{2}b\nabla\Gamma_1^z(f, f) \\ &\quad - \Gamma_1^z((\Delta_a - A\nabla + b\nabla)f, f) \\ &= \Gamma_{2,\tilde{L}}^z(f, f) + [\frac{1}{2}b\nabla\Gamma_1^z(f, f) - \Gamma_1^z(b\nabla f, f)].\end{aligned}$$

The term  $\Gamma_{2,\tilde{L}}^z(f, f)$  follows from Lemma 12. We are left to compute the last two terms:

$$\begin{aligned}\frac{1}{2}b\nabla\Gamma_1^z(f, f) &= -\frac{1}{2}\sum_{\hat{k}=1}^{n+m} b_{\hat{k}} \frac{\partial}{\partial x_{\hat{k}}} (\langle z^\top \nabla f, z^\top \nabla f \rangle_{\mathbb{R}^m}) \\ &= \sum_{\hat{k}, \hat{i}=1}^{n+m} \sum_{i=1}^m (b_{\hat{k}} \frac{\partial z_{ii}^\top}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}} + b_{\hat{k}} z_{ii}^\top \frac{\partial^2 f}{\partial x_{\hat{k}} \partial x_{\hat{i}}}) (z^\top \nabla f)_i,\end{aligned}$$

and

$$\begin{aligned} -\Gamma_1^z(b\nabla f, f) &= -\langle z^\top \nabla(b\nabla f), z^\top \nabla f \rangle_{\mathbb{R}^n} \\ &= -\sum_{i=1}^m \sum_{\hat{k}, \hat{i}=1}^{n+m} (z_{ii}^\top \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} + z_{ii}^\top b_{\hat{k}} \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}}) (z^\top \nabla f)_i. \end{aligned}$$

**Step 3:** Following Lemma 10, which will be proven shortly in the next section, we have

$$\mathbf{div}_z^\pi(\Gamma_{\nabla(aa^\top)} f, f) - \mathbf{div}_a^\pi(\Gamma_{\nabla(zz^\top)} f, f) = \mathfrak{R}^\pi(f, f) + 2G^\top X,$$

where  $X, G$  are defined in Notation 1 and  $\mathfrak{R}^\pi$  is defined in Definition 2.

**Step 4:** Combining the above terms  $\Gamma_{2,\tilde{L}}(f, f)$  in Lemma 11,  $\Gamma_{2,\tilde{L}}^z(f, f)$  in Lemma 12, and  $\mathfrak{R}^\pi(f, f) + 2G^\top X$ , we have

$$\begin{aligned} &\Gamma_{2,\tilde{L}}(f, f) + \Gamma_{2,\tilde{L}}^z(f, f) + \mathfrak{R}^\pi(f, f) + 2G^\top X \\ &= X^\top P^\top P X + 2E^\top P X + 2F^\top X + E^\top E \\ &\quad + X^\top Q^\top Q X + 2D^\top Q X + 2C^\top X + D^\top D + 2G^\top X \\ &\quad + \mathfrak{R}_a(\nabla f, \nabla f) + \mathfrak{R}_z(\nabla f, \nabla f) + \mathfrak{R}_\pi(\nabla f, \nabla f) \\ &= X^\top [P^\top P + Q^\top Q] X + 2[G^\top + F^\top + C^\top] X + 2[E^\top P + D^\top Q] X + D^\top D + E^\top E \\ &\quad + \mathfrak{R}_a(\nabla f, \nabla f) + \mathfrak{R}_z(\nabla f, \nabla f) + \mathfrak{R}_\pi(\nabla f, \nabla f). \end{aligned}$$

Assuming that Assumption 1 is satisfied, we obtain

$$\begin{aligned} &\Gamma_{2,\tilde{L}}(f, f) + \Gamma_{2,\tilde{L}}^z(f, f) + \mathfrak{R}^\pi(f, f) + 2G^\top X \\ &= [X + \Lambda_1]^\top Q^\top Q [X + \Lambda_1] + [X + \Lambda_2]^\top P^\top P [X + \Lambda_2] \\ &\quad + \mathfrak{R}_a(\nabla f, \nabla f) + \mathfrak{R}_z(\nabla f, \nabla f) + \mathfrak{R}^\pi(\nabla f, \nabla f) \\ &\quad - \Lambda_1^\top Q^\top Q \Lambda_1 - \Lambda_2^\top P^\top P \Lambda_2 + D^\top D + E^\top E. \end{aligned}$$

Adding the drift terms from Step 1 and Step 2, we obtain  $\mathfrak{R}_{ab}$  and  $\mathfrak{R}_{zb}$ , which finishes the proof.  $\square$

### 5.1. Proof of Lemma 10

**Lemma 13.**

$$\mathbf{div}_z^\pi(\Gamma_{\nabla(aa^\top)} f, f) - \mathbf{div}_a^\pi(\Gamma_{\nabla(zz^\top)} f, f) = \mathfrak{R}^\pi(f, f) + 2G^\top X, \quad (43)$$

where  $X, G$  are defined in Notation 1 and  $\mathfrak{R}^\pi$  is defined in Definition 2.

**Proof.** For the first term in the above lemma, we have

$$\begin{aligned} \mathbf{div}_z^\pi(\Gamma_{\nabla(aa^\top)} f, f) &= \frac{\nabla \cdot (zz^\top \pi \Gamma_{\nabla(aa^\top)}(f, f))}{\pi} \\ &= \sum_{k'=1}^{n+m} \frac{1}{\pi} \frac{\partial}{\partial x_{k'}} \left[ \sum_{k=1}^m z_{k'k} \left( \pi \sum_{\hat{k}=1}^{n+m} z_{k\hat{k}}^\top \left( \Gamma_{\nabla(aa^\top)}(f, f) \right)_{\hat{k}} \right) \right] \\ &= \sum_{k'=1}^{n+m} \sum_{k=1}^m \left[ \frac{\partial}{\partial x_{k'}} z_{k'k} \left( \sum_{\hat{k}=1}^{n+m} z_{k\hat{k}}^\top \left( \Gamma_{\nabla(aa^\top)}(f, f) \right)_{\hat{k}} \right) \right. \\ &\quad \left. + z_{k'k} \frac{\partial}{\partial x_{k'}} \left( \sum_{\hat{k}=1}^{n+m} z_{k\hat{k}}^\top \left( \Gamma_{\nabla(aa^\top)}(f, f) \right)_{\hat{k}} \right) \right] \end{aligned}$$

$$\begin{aligned}
& + \sum_{k'=1}^{n+m} \sum_{k=1}^m \frac{\partial}{\partial x_{k'}} \log \pi \left[ z_{k'k}^T \sum_{\hat{k}=1}^{n+m} z_{k\hat{k}}^T \left( \Gamma_{\nabla(aa^T)}(f, f) \right)_{\hat{k}} \right] \\
= & \sum_{k'=1}^{n+m} \sum_{k=1}^m \left[ \frac{\partial}{\partial x_{k'}} z_{kk'}^T \left( \sum_{\hat{k}=1}^{n+m} z_{k\hat{k}}^T \left( \Gamma_{\nabla(aa^T)}(f, f) \right)_{\hat{k}} \right) \right. \\
& \left. + z_{kk'}^T \frac{\partial}{\partial x_{k'}} \left( \sum_{\hat{k}=1}^{n+m} z_{k\hat{k}}^T \left( \Gamma_{\nabla(aa^T)}(f, f) \right)_{\hat{k}} \right) \right] \\
& + \sum_{k=1}^m (z^T \nabla \log \pi)_k \left[ \sum_{\hat{k}=1}^{n+m} z_{k\hat{k}}^T \left( \Gamma_{\nabla(aa^T)}(f, f) \right)_{\hat{k}} \right],
\end{aligned}$$

where  $\Gamma_{\nabla(aa^T)}(f, f)_{\hat{k}}$  is defined in (42). Plugging in (42), we further obtain

$$\begin{aligned}
& \mathbf{div}_z^\pi(\Gamma_{\nabla(aa^T)} f, f) \\
= & \sum_{k'=1}^{n+m} \sum_{k=1}^m \left[ \frac{\partial}{\partial x_{k'}} z_{kk'}^T \left( \sum_{\hat{k}=1}^{n+m} z_{k\hat{k}}^T \left( 2 \sum_{i=1}^n \sum_{\hat{i}, i'=1}^{n+m} \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i i'}^T \frac{\partial f}{\partial x_{i'}} \right) \right) \right] \\
& + \sum_{k'=1}^{n+m} \sum_{k=1}^m \left[ z_{kk'}^T \frac{\partial}{\partial x_{k'}} \left( \sum_{\hat{k}=1}^{n+m} z_{k\hat{k}}^T \left( 2 \sum_{i=1}^n \sum_{\hat{i}, i'=1}^{n+m} \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i i'}^T \frac{\partial f}{\partial x_{i'}} \right) \right) \right] \\
& + \sum_{k=1}^m (z^T \nabla \log \pi)_k \left[ \sum_{\hat{k}=1}^{n+m} z_{k\hat{k}}^T \left( 2 \sum_{i=1}^n \sum_{\hat{i}, i'=1}^{n+m} \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i i'}^T \frac{\partial f}{\partial x_{i'}} \right) \right] \\
= & 2 \sum_{k=1}^m \sum_{i=1}^n \sum_{k', \hat{k}, \hat{i}, i'=1}^{n+m} \left[ \frac{\partial}{\partial x_{k'}} z_{kk'}^T z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i i'}^T \frac{\partial f}{\partial x_{i'}} \right] \cdots \mathcal{S}_1^z \\
& + 2 \sum_{k=1}^m \sum_{i=1}^n \sum_{k', \hat{k}, \hat{i}, i'=1}^{n+m} \left[ z_{kk'}^T \frac{\partial}{\partial x_{k'}} \left( z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i i'}^T \frac{\partial f}{\partial x_{i'}} \right) \right] \cdots \mathcal{S}_2^z \\
& + 2 \sum_{k=1}^m \sum_{i=1}^n \sum_{\hat{k}, \hat{i}, i'=1}^{n+m} (z^T \nabla \log \pi)_k \left[ z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i i'}^T \frac{\partial f}{\partial x_{i'}} \right] \cdots \mathcal{S}_3^z \\
= & \mathcal{S}_1^z + \mathcal{S}_2^z + \mathcal{S}_3^z.
\end{aligned} \tag{44}$$

By further expanding  $\mathcal{S}_2^z$ , we obtain

$$\begin{aligned}
\mathcal{S}_2^z & = 2 \sum_{k=1}^m \sum_{i=1}^n \sum_{k', \hat{k}, \hat{i}, i'=1}^{n+m} \left[ z_{kk'}^T \frac{\partial}{\partial x_{k'}} \left( z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i i'}^T \frac{\partial f}{\partial x_{i'}} \right) \right] \\
& = 2 \sum_{k=1}^m \sum_{i=1}^n \sum_{k', \hat{k}, \hat{i}, i'=1}^{n+m} \left[ z_{kk'}^T \frac{\partial}{\partial x_{k'}} z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i i'}^T \frac{\partial f}{\partial x_{i'}} \right. \\
& \quad + z_{kk'}^T z_{k\hat{k}}^T \frac{\partial^2}{\partial x_{k'} \partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i i'}^T \frac{\partial f}{\partial x_{i'}} \\
& \quad + z_{kk'}^T z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial^2 f}{\partial x_{k'} \partial x_{\hat{i}}} a_{i i'}^T \frac{\partial f}{\partial x_{i'}} \\
& \quad + z_{kk'}^T z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{i i'}^T \frac{\partial^2 f}{\partial x_{k'} \partial x_{i'}} \\
& \quad \left. + z_{kk'}^T z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{i\hat{i}}^T \frac{\partial f}{\partial x_{\hat{i}}} \frac{\partial}{\partial x_{i'}} a_{i i'}^T \frac{\partial f}{\partial x_{i'}} \right].
\end{aligned}$$

Similarly, we obtain

$$\begin{aligned}
 & \mathbf{div}_a^\pi(\Gamma_{\nabla(zz^\top)}f, f) \\
 = & \sum_{l'=1}^{n+m} \sum_{l=1}^n \left[ \frac{\partial}{\partial x_{l'}} a_{ll'}^\top \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^\top (\Gamma_{\nabla(zz^\top)}(f, f)) \right)_{\hat{l}} \right. \\
 & \left. + a_{ll'}^\top \frac{\partial}{\partial x_{l'}} \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^\top (\Gamma_{\nabla(zz^\top)}(f, f)) \right)_{\hat{l}} \right] \\
 & + \sum_{l=1}^n (a^\top \nabla \log \pi)_l \left[ \sum_{\hat{l}=1}^{n+m} (a_{l\hat{l}}^\top \Gamma_{\nabla(zz^\top)}(f, f))_{\hat{l}} \right] \\
 = & 2 \sum_{j=1}^m \sum_{l=1}^n \sum_{\hat{l}, \hat{j}, j'=1}^{n+m} \left[ \frac{\partial}{\partial x_{l'}} a_{ll'}^\top a_{l\hat{l}}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^\top \frac{\partial f}{\partial x_{\hat{j}}} z_{j\hat{j}'}^\top \frac{\partial f}{\partial x_{j'}} \right] \cdots \mathcal{S}_1^a \\
 & + 2 \sum_{j=1}^m \sum_{l=1}^n \sum_{\hat{l}, \hat{j}, j'=1}^{n+m} \left[ a_{ll'}^\top \frac{\partial}{\partial x_{l'}} \left( a_{l\hat{l}}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^\top \frac{\partial f}{\partial x_{\hat{j}}} z_{j\hat{j}'}^\top \frac{\partial f}{\partial x_{j'}} \right) \right] \cdots \mathcal{S}_2^a \\
 & + 2 \sum_{j=1}^m \sum_{l=1}^n \sum_{\hat{l}, \hat{j}, j'=1}^{n+m} (a^\top \nabla \log \pi)_l \left[ a_{l\hat{l}}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^\top \frac{\partial f}{\partial x_{\hat{j}}} z_{j\hat{j}'}^\top \frac{\partial f}{\partial x_{j'}} \right] \cdots \mathcal{S}_3^a \\
 = & \mathcal{S}_1^a + \mathcal{S}_2^a + \mathcal{S}_3^a,
 \end{aligned} \tag{46}$$

where we also obtain

$$\begin{aligned}
 \mathcal{S}_2^a &= 2 \sum_{j=1}^m \sum_{l=1}^n \sum_{\hat{l}, \hat{j}, j'=1}^{n+m} \left[ a_{ll'}^\top \frac{\partial}{\partial x_{l'}} \left( a_{l\hat{l}}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^\top \frac{\partial f}{\partial x_{\hat{j}}} z_{j\hat{j}'}^\top \frac{\partial f}{\partial x_{j'}} \right) \right] \\
 &= 2 \sum_{j=1}^m \sum_{l=1}^n \sum_{\hat{l}, \hat{j}, j'=1}^{n+m} \left[ a_{ll'}^\top \frac{\partial}{\partial x_{l'}} a_{l\hat{l}}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^\top \frac{\partial f}{\partial x_{\hat{j}}} z_{j\hat{j}'}^\top \frac{\partial f}{\partial x_{j'}} \right. \\
 &\quad + a_{ll'}^\top a_{l\hat{l}}^\top \frac{\partial^2}{\partial x_{l'} \partial x_{\hat{l}}} z_{j\hat{j}}^\top \frac{\partial f}{\partial x_{\hat{j}}} z_{j\hat{j}'}^\top \frac{\partial f}{\partial x_{j'}} \\
 &\quad + a_{ll'}^\top a_{l\hat{l}}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^\top \frac{\partial^2 f}{\partial x_{l'} \partial x_{\hat{j}}} z_{j\hat{j}'}^\top \frac{\partial f}{\partial x_{j'}} \\
 &\quad + a_{ll'}^\top a_{l\hat{l}}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^\top \frac{\partial f}{\partial x_{\hat{j}}} z_{j\hat{j}'}^\top \frac{\partial^2 f}{\partial x_{l'} \partial x_{j'}} \\
 &\quad \left. + a_{ll'}^\top a_{l\hat{l}}^\top \frac{\partial}{\partial x_{\hat{l}}} z_{j\hat{j}}^\top \frac{\partial f}{\partial x_{\hat{j}}} \frac{\partial}{\partial x_{l'}} z_{j\hat{j}'}^\top \frac{\partial f}{\partial x_{j'}} \right].
 \end{aligned}$$

Combining all the terms above, we have

$$\mathbf{div}_z^\pi(\Gamma_{\nabla(aa^\top)}f, f) - \mathbf{div}_a^\pi(\Gamma_{\nabla(zz^\top)}f, f) = \mathcal{S}_1^z + \mathcal{S}_2^z + \mathcal{S}_3^z - (\mathcal{S}_1^a + \mathcal{S}_2^a + \mathcal{S}_3^a).$$

By direct computations, we separate the above terms into two groups based on “ $\partial f \partial f$ ” and “ $\partial^2 f \partial f$ ”. We denote  $\mathfrak{R}^\pi(f, f)$  as the sum of all “ $\partial f \partial f$ ” terms and denote  $2G^\top X$  as the sum of all “ $\partial^2 f \partial f$ ” terms. Switching indices for the terms in  $2G^\top X$  to match  $\frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{j}}}$ , we obtain the following:



$$\begin{aligned}
& 2G^T X \\
&= 2 \sum_{k=1}^m \sum_{i=1}^n \sum_{k', \hat{k}, \hat{i}, i'=1}^{n+m} \left[ z_{kk'}^T z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{ii'}^T \frac{\partial^2 f}{\partial x_{k'} \partial x_{\hat{i}}} a_{ii'}^T \frac{\partial f}{\partial x_{i'}} + z_{kk'}^T z_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} a_{ii'}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{ii'}^T \frac{\partial^2 f}{\partial x_{k'} \partial x_{i'}} \right] \\
&\quad - 2 \sum_{j=1}^m \sum_{l=1}^n \sum_{l', \hat{l}, \hat{j}, j'=1}^{n+m} \left[ a_{ll'}^T a_{ll}^T \frac{\partial}{\partial x_{\hat{l}}} z_{jj'}^T \frac{\partial^2 f}{\partial x_{l'} \partial x_{\hat{j}}} z_{jj'}^T \frac{\partial f}{\partial x_{j'}} + a_{ll'}^T a_{ll}^T \frac{\partial}{\partial x_{\hat{l}}} z_{jj'}^T \frac{\partial f}{\partial x_{\hat{j}}} z_{jj'}^T \frac{\partial^2 f}{\partial x_{l'} \partial x_{j'}} \right] \\
&= 2 \sum_{j=1}^m \sum_{i=1}^n \sum_{j', \hat{j}, \hat{i}, i'=1}^{n+m} \left[ z_{jj'}^T z_{jj}^T \frac{\partial}{\partial x_{\hat{j}}} a_{ii'}^T \frac{\partial^2 f}{\partial x_{j'} \partial x_{\hat{i}}} a_{ii'}^T \frac{\partial f}{\partial x_{i'}} + z_{jj'}^T z_{jj}^T \frac{\partial}{\partial x_{\hat{j}}} a_{ii'}^T \frac{\partial f}{\partial x_{\hat{i}}} a_{ii'}^T \frac{\partial^2 f}{\partial x_{j'} \partial x_{i'}} \right] \\
&\quad - 2 \sum_{j=1}^m \sum_{i=1}^n \sum_{i', \hat{i}, \hat{j}, j'=1}^{n+m} \left[ a_{ii'}^T a_{ii}^T \frac{\partial}{\partial x_{\hat{i}}} z_{jj'}^T \frac{\partial^2 f}{\partial x_{i'} \partial x_{\hat{j}}} z_{jj'}^T \frac{\partial f}{\partial x_{j'}} + a_{ii'}^T a_{ii}^T \frac{\partial}{\partial x_{\hat{i}}} z_{jj'}^T \frac{\partial f}{\partial x_{\hat{j}}} z_{jj'}^T \frac{\partial^2 f}{\partial x_{i'} \partial x_{j'}} \right] \\
&= 2 \sum_{j=1}^m \sum_{i=1}^n \sum_{j', \hat{j}, \hat{i}, i'=1}^{n+m} \left[ z_{jj'}^T z_{jj'}^T \frac{\partial}{\partial x_{j'}} a_{ii'}^T \frac{\partial^2 f}{\partial x_{\hat{j}} \partial x_{\hat{i}}} a_{ii'}^T \frac{\partial f}{\partial x_{i'}} + z_{jj'}^T z_{jj'}^T \frac{\partial}{\partial x_{j'}} a_{ii'}^T \frac{\partial f}{\partial x_{i'}} a_{ii'}^T \frac{\partial^2 f}{\partial x_{\hat{j}} \partial x_{\hat{i}}} \right] \\
&\quad - 2 \sum_{j=1}^m \sum_{i=1}^n \sum_{i', \hat{i}, \hat{j}, j'=1}^{n+m} \left[ a_{ii'}^T a_{ii'}^T \frac{\partial}{\partial x_{i'}} z_{jj'}^T \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{j}}} z_{jj'}^T \frac{\partial f}{\partial x_{j'}} + a_{ii'}^T a_{ii'}^T \frac{\partial}{\partial x_{i'}} z_{jj'}^T \frac{\partial f}{\partial x_{j'}} z_{jj'}^T \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{j}}} \right] \\
&= 2 \sum_{\hat{i}, \hat{j}=1}^{n+m} \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{j}}} \left[ \sum_{i=1}^n \sum_{j=1}^m \sum_{j', \hat{j}, \hat{i}, i'=1}^{n+m} \left[ \left( z_{jj'}^T z_{jj'}^T \frac{\partial}{\partial x_{j'}} a_{ii'}^T a_{ii'}^T \frac{\partial f}{\partial x_{i'}} + z_{jj'}^T z_{jj'}^T \frac{\partial}{\partial x_{j'}} a_{ii'}^T \frac{\partial f}{\partial x_{i'}} a_{ii'}^T \right) \right. \right. \\
&\quad \left. \left. - \left( a_{ii'}^T a_{ii'}^T \frac{\partial}{\partial x_{i'}} z_{jj'}^T z_{jj'}^T \frac{\partial f}{\partial x_{j'}} + a_{ii'}^T a_{ii'}^T \frac{\partial}{\partial x_{i'}} z_{jj'}^T \frac{\partial f}{\partial x_{j'}} z_{jj'}^T \right) \right] \right].
\end{aligned}$$

The first equality follows from the quantities we obtained previously, the second equality from switching “ $k$ ” to “ $j$ ” and “ $l$ ” to “ $i$ ”, and the third equality from switching between “ $i$ ” and “ $\hat{i}$ ”, “ $j$ ” and “ $\hat{j}$ ”. Thus, the proof is completed.  $\square$

## 5.2. Proof of Lemma 11

From now on, we keep the following notation:  $a^T \nabla f = \sum_{i=1}^n \sum_{\hat{i}=1}^{n+m} a_{ii}^T \frac{\partial}{\partial x_{\hat{i}}} f$ . Furthermore, we fixed the notation for  $a, a^T$  with relation  $a_{\hat{i}\hat{i}} = a_{ii}^T$  for  $i = 1, \dots, n$  and  $\hat{i} = 1, \dots, n+m$ . Here, we denote  $a_{ii}^T := (a^T)_{ii}$ . Recall that we define

$$\Gamma_{2,\tilde{L}}(f, f) = \frac{1}{2} (\tilde{L}\Gamma_1(f, f) - 2\Gamma_1(\tilde{L}f, f)).$$

Next, we are ready to prove the following lemma.

### Lemma 14.

$$\Gamma_{2,\tilde{L}}(f, f) = X^T Q^T QX + 2D^T QX + 2C^T X + D^T D + \mathfrak{R}_a(\nabla f, \nabla f),$$

where  $Q, X, C, D$  are introduced in Notation 1 and  $\mathfrak{R}_a$  is defined in Definition 2.

**Proof.** We plug in the operator  $\tilde{L}$  into our definition for  $\Gamma_2$ :

$$\begin{aligned}
\Gamma_{2,\tilde{L}}(f, f) &= \frac{1}{2} \Delta_a \Gamma_1(f, f) - \frac{1}{2} A \nabla \Gamma_1(f, f) - \Gamma_1((\Delta_a - A \nabla)f, f) \\
&= \frac{1}{2} \Delta_a \Gamma_1(f, f) - \Gamma_1(\Delta_a f, f) - \frac{1}{2} A \nabla \Gamma_1(f, f) + \Gamma_1(A \nabla f, f).
\end{aligned}$$

Now, we compute the last two terms of the above equation. With  $A = a \otimes \nabla a$ , we obtain

$$\begin{aligned} -\frac{1}{2}A\nabla\Gamma_1(f, f) &= -\frac{1}{2}\sum_{\hat{k}=1}^{n+m}A_{\hat{k}}\nabla_{\frac{\partial}{\partial x_{\hat{k}}}}\langle a^{\top}\nabla f, a^{\top}\nabla f\rangle_{\mathbb{R}^n} \\ &= -\sum_{\hat{k}=1}^{n+m}\langle A_{\hat{k}}(\nabla_{\frac{\partial}{\partial x_{\hat{k}}}}a^{\top})\nabla f, a^{\top}\nabla f\rangle_{\mathbb{R}^n} - \sum_{\hat{k}=1}^{n+m}\langle A_{\hat{k}}a^{\top}(\nabla_{\frac{\partial}{\partial x_{\hat{k}}}}\nabla f), a^{\top}\nabla f\rangle_{\mathbb{R}^n} \\ &= \mathbf{J}_1 + \mathbf{J}_2, \end{aligned}$$

and

$$\begin{aligned} \Gamma_1(A\nabla f, f) &= \langle a^{\top}\nabla(\sum_{\hat{k}=1}^{n+m}A_{\hat{k}}\nabla_{\frac{\partial}{\partial x_{\hat{k}}}}f), a^{\top}\nabla f\rangle_{\mathbb{R}^n} \\ &= \langle a^{\top}(\sum_{\hat{k}=1}^{n+m}A_{\hat{k}}\nabla\nabla_{\frac{\partial}{\partial x_{\hat{k}}}}f), a^{\top}\nabla f\rangle_{\mathbb{R}^n} + \langle a^{\top}(\sum_{\hat{k}=1}^{n+m}\nabla A_{\hat{k}}\nabla_{\frac{\partial}{\partial x_{\hat{k}}}}f), a^{\top}\nabla f\rangle_{\mathbb{R}^n} \\ &= \mathbf{J}_3 + \mathbf{J}_4. \end{aligned}$$

It is easy to see

$$\mathbf{J}_2 + \mathbf{J}_3 = 0.$$

We now expand  $\mathbf{J}_1$  and  $\mathbf{J}_4$  into local coordinates:

$$\mathbf{J}_1 = -\sum_{l=1}^n(a^{\top}\nabla f)_l\left(\sum_{l'=1}^{n+m}\sum_{k=1}^n\sum_{k'=1}^{n+m}a_{\hat{k}k}\nabla_{\frac{\partial}{\partial x_{k'}}}a_{k'l'}\nabla_{\frac{\partial}{\partial x_{\hat{k}}}}a_{ll'}\nabla_{\frac{\partial}{\partial x_{l'}}}f\right), \quad (47)$$

and

$$\begin{aligned} \mathbf{J}_4 &= \sum_{l=1}^n(a^{\top}\nabla f)_l\left(\sum_{l'=1}^{n+m}a_{ll'}^{\top}\left(\sum_{\hat{k}=1}^{n+m}\nabla_{\frac{\partial}{\partial x_{l'}}}\left(\sum_{k=1}^n\sum_{k'=1}^{n+m}a_{\hat{k}k}\nabla_{\frac{\partial}{\partial x_{k'}}}a_{k'l'}\nabla_{\frac{\partial}{\partial x_{\hat{k}}}}f\right)\right)\right) \\ &= \sum_{l=1}^n(a^{\top}\nabla f)_l\left(\sum_{k=1}^n\sum_{l'=1}^{n+m}\sum_{\hat{k}, k'=1}^{n+m}a_{ll'}^{\top}\nabla_{\frac{\partial}{\partial x_{l'}}}a_{\hat{k}k}\nabla_{\frac{\partial}{\partial x_{k'}}}a_{k'l'}\nabla_{\frac{\partial}{\partial x_{\hat{k}}}}f\right) \\ &\quad + \sum_{l=1}^n(a^{\top}\nabla f)_l\left(\sum_{k=1}^n\sum_{l'=1}^{n+m}\sum_{\hat{k}, k'=1}^{n+m}a_{ll'}^{\top}a_{\hat{k}k}(\nabla_{\frac{\partial}{\partial x_{l'}}}\nabla_{\frac{\partial}{\partial x_{k'}}}a_{k'l'})\nabla_{\frac{\partial}{\partial x_{\hat{k}}}}f\right). \end{aligned} \quad (48)$$

Applying Lemma 15, which will be proven shortly below, we have

$$\begin{aligned} &\frac{1}{2}\Delta_a\Gamma_1(f, f) - \Gamma_1(\Delta_a f, f) \\ &= \frac{1}{2}(a^{\top}\nabla \circ (a^{\top}\nabla|a^{\top}\nabla f|^2)) - \langle a^{\top}\nabla((a^{\top}\nabla) \circ (a^{\top}\nabla f)), a^{\top}\nabla f\rangle_{\mathbb{R}^n} \\ &\quad + \sum_{l=1}^n(a^{\top}\nabla f)_l\left(\sum_{\hat{i}, \hat{k}, l'=1}^{n+m}\sum_{k=1}^n\left(\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}a_{kk}^{\top}\left(\frac{\partial}{\partial x_{\hat{k}}}a_{ll'}^{\top}\frac{\partial}{\partial x_{l'}}f\right) - a_{ll'}^{\top}\frac{\partial}{\partial x_{\hat{i}}}a_{ik}\left(\frac{\partial}{\partial x_{l'}}a_{kk}^{\top}\frac{\partial}{\partial x_{\hat{k}}}f\right)\right)\right) \\ &\quad - \langle \mathbf{B}_{n \times n}a^{\top}\nabla f, a^{\top}\nabla f\rangle_{\mathbb{R}^n}, \end{aligned}$$

where

$$\langle \mathbf{B}_{n \times n}a^{\top}\nabla f, a^{\top}\nabla f\rangle_{\mathbb{R}^n} = \sum_{l=1}^n(a^{\top}\nabla f)_l\left(\sum_{k=1}^n\sum_{l'=1}^n\sum_{i=1}^{n+m}\sum_{k', j'=1}^{n+m}a_{lj'}^{\top}\frac{\partial^2}{\partial x_i\partial x_{j'}}a_{ik}(a_{kk}^{\top}\frac{\partial}{\partial x_{k'}}f)\right).$$

Thus, combining with (47) and (48), we have

$$\begin{aligned}\Gamma_{2,\tilde{L}}(f, f) &= \frac{1}{2}\Delta_a\Gamma_1(f, f) - \Gamma_1(\Delta_a f, f) + \mathbf{J}_1 + \mathbf{J}_4 \\ &= \frac{1}{2}(a^\top \nabla \circ (a^\top \nabla |a^\top \nabla f|^2)) - \langle a^\top \nabla [(a^\top \nabla) \circ (a^\top \nabla f)], a^\top \nabla f \rangle_{\mathbb{R}^n}.\end{aligned}$$

where the last term follows from Lemma 16 below. The proof is thus completed.  $\square$

**Lemma 15.**

$$\begin{aligned}&\frac{1}{2}\Delta_a\Gamma_1(f, f) - \Gamma_1(\Delta_a f, f) \\ &= \frac{1}{2}(a^\top \nabla \circ (a^\top \nabla |a^\top \nabla f|^2)) - \langle a^\top \nabla [(a^\top \nabla) \circ (a^\top \nabla f)], a^\top \nabla f \rangle_{\mathbb{R}^n} \\ &\quad - \langle \mathbf{B}_{n \times n} a^\top \nabla f, a^\top \nabla f \rangle_{\mathbb{R}^n} + \mathbf{B}_0.\end{aligned}\tag{49}$$

Here, the local representations for  $\mathbf{B}_{n \times n}$  and  $\mathbf{B}_0$  are given as follows. For  $l, k = 1, \dots, n$ , we denote

$$\begin{aligned}\mathbf{B}_{lk} &= \sum_{j'=1}^{n+m} a_{lj'}^\top \sum_{i=1}^{n+m} \frac{\partial^2}{\partial x_i \partial x_{j'}} a_{ik} = \sum_{j'=1}^{n+m} a_{lj'}^\top \sum_{i=1}^{n+m} \frac{\partial^2}{\partial x_i \partial x_{j'}} a_{ki}^\top, \\ \mathbf{B}_0 &= \sum_{l=1}^n (a^\top \nabla f)_l \left( \sum_{\hat{i}, \hat{k}, l'=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} a_{kk}^\top \left( \frac{\partial}{\partial x_{\hat{k}}} a_{ll'}^\top \frac{\partial}{\partial x_{l'}} f \right) \right. \right. \\ &\quad \left. \left. - a_{ll'}^\top \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \left( \frac{\partial}{\partial x_{l'}} a_{kk}^\top \frac{\partial}{\partial x_{\hat{k}}} f \right) \right) \right).\end{aligned}\tag{50}$$

We introduce the following notation (convention) that, for any function  $F$ ,

$$(a^\top \nabla) \circ (a^\top \nabla F) = \sum_{i=1}^n (a^\top \nabla)_i (a^\top \nabla F)_i = \sum_{i=1}^n \sum_{\hat{i}, i'=1}^{n+m} (a_{i\hat{i}}^\top \frac{\partial}{\partial x_{\hat{i}}}) (a_{i'i}^\top \frac{\partial F}{\partial x_{i'}}).\tag{51}$$

**Proof of Lemma 15.** By our definition above, we have

$$\begin{aligned}\Delta_a\Gamma_1(f, f) &= \nabla \cdot (a a^\top \nabla \langle a^\top \nabla f, a^\top \nabla f \rangle_{\mathbb{R}^n}) \\ &= \nabla \cdot (a F) \\ &= \sum_{\hat{i}=1}^{n+m} \frac{\partial}{\partial x_{\hat{i}}} \left( \sum_{k=1}^n a_{\hat{i}k} F_k \right) \\ &= \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} F_k + a_{\hat{i}k} \frac{\partial}{\partial x_{\hat{i}}} F_k \right) \\ &= \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} F_k \right) + a^\top \nabla \circ (a^\top \nabla (a^\top \nabla f)^2),\end{aligned}$$

where we denote

$$\begin{aligned}F &= a^\top \nabla \langle a^\top \nabla f, a^\top \nabla f \rangle_{\mathbb{R}^n} \\ &= a^\top \nabla \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^\top \frac{\partial}{\partial x_{\hat{l}}} f \right)^2 \\ &= \left( \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^\top \frac{\partial}{\partial x_{\hat{k}}} \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^\top \frac{\partial}{\partial x_{\hat{l}}} f \right)^2 \right)_{k=1, \dots, n} = (F_1, F_2, \dots, F_n)^\top.\end{aligned}$$

Therefore, we have

$$\begin{aligned}
 \Delta_a \Gamma_1(f, f) &= \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \left( \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{j}}} \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right)^2 \right) \right. \\
 &\quad \left. + a^T \nabla \circ (a^T \nabla (a^T \nabla f)^2) \right) \\
 &= \sum_{k=1}^n \sum_{\hat{i}=1}^{n+m} \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \left( \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} (a^T \nabla f)^2 \right) \right) \\
 &\quad + (a^T \nabla) \circ (a^T \nabla (a^T \nabla f)^2) \\
 &= \nabla a \circ (a^T \nabla (a^T \nabla f)^2) + (a^T \nabla) \circ (a^T \nabla (a^T \nabla f)^2). \tag{52}
 \end{aligned}$$

Next, we compute the following quantity.

$$\Gamma_1(\Delta_a f, f) = \langle a^T \nabla (\nabla \cdot (a a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n},$$

where we have

$$\begin{aligned}
 \nabla \cdot (a a^T \nabla f) &= \nabla \cdot \left( \sum_{k=1}^n \sum_{\hat{k}=1}^{n+m} a_{\hat{i}k} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f \right) = \sum_{\hat{i}=1}^{n+m} \frac{\partial}{\partial x_{\hat{i}}} \left( \sum_{k=1}^n \sum_{\hat{k}=1}^{n+m} a_{\hat{i}k} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f \right) \\
 &= \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) + \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n a_{\hat{i}k} \frac{\partial}{\partial x_{\hat{i}}} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \\
 &= \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) + (a^T \nabla) \circ (a^T \nabla f) \\
 &= \nabla a \circ (a^T \nabla f) + (a^T \nabla) \circ (a^T \nabla f).
 \end{aligned}$$

We continue with our computation as below:

$$\begin{aligned}
 \Gamma_1(\Delta_a f, f) &= \langle a^T \nabla \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) + (a^T \nabla) \circ (a^T \nabla f) \right], a^T \nabla f \rangle_{\mathbb{R}^n} \\
 &= \langle a^T \nabla \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right], a^T \nabla f \rangle_{\mathbb{R}^n} \\
 &\quad + \langle a^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n} \\
 &= \langle a^T \nabla (\nabla a \cdot (a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n} + \langle a^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n} \\
 &= \langle (a^T \nabla \nabla a \cdot (a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n} + \langle (\nabla a \cdot (a^T \nabla (a^T \nabla f))), a^T \nabla f \rangle_{\mathbb{R}^n} \\
 &\quad + \langle a^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n}. \tag{53}
 \end{aligned}$$

From the above, combining (52) and (53), we further obtain

$$\begin{aligned}
 &\frac{1}{2} \Delta_a \Gamma_1(f, f) - \Gamma_1(\Delta_a f, f) \\
 &= \frac{1}{2} (a^T \nabla \circ (a^T \nabla |a^T \nabla f|^2)) - \langle a^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n}
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{2} \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \left( \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{j}}} \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right)^2 \right) \right) \\
 & - \langle a^T \nabla \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right], a^T \nabla f \rangle_{\mathbb{R}^n} \\
 & = \frac{1}{2} (a^T \nabla \circ (a^T \nabla |a^T \nabla f|^2)) - \langle a^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n} \\
 & + \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \frac{\partial}{\partial x_{\hat{k}}} \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \right) \cdots \text{I} \\
 & - \sum_{l=1}^n \left( \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} a_{l\hat{l}'}^T \frac{\partial}{\partial x_{l'}} \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \cdots \text{II}.
 \end{aligned}$$

Recall that we denote  $a^T$  to emphasize the transpose of the matrix  $a$  and  $a_{ii}^T = a_{ii}$ :

$$\begin{aligned}
 \text{I} & = \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \frac{\partial}{\partial x_{\hat{k}}} \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \right) \\
 & = \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{\hat{l}=1}^{n+m} \frac{\partial}{\partial x_{\hat{k}}} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \right) \\
 & \quad + \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial}{\partial x_{\hat{l}}} f \right) \right) \\
 & = \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \left( \sum_{l'=1}^{n+m} \frac{\partial}{\partial x_{\hat{k}}} a_{l\hat{l}'}^T \frac{\partial}{\partial x_{l'}} f \right) \right) \\
 & \quad + \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}'}^T \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial}{\partial x_{l'}} f \right) \right),
 \end{aligned}$$

and

$$\begin{aligned}
 \text{II} & = \sum_{l=1}^n \left( \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} a_{l\hat{l}'}^T \frac{\partial}{\partial x_{l'}} \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \\
 & = \sum_{l=1}^n \left( \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} a_{l\hat{l}'}^T \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \frac{\partial}{\partial x_{l'}} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \\
 & \quad + \sum_{l=1}^n \left( \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} a_{l\hat{l}'}^T \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial^2}{\partial x_{\hat{i}} x_{l'}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \\
 & = \sum_{l=1}^n \left( \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} a_{l\hat{l}'}^T \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \left( \frac{\partial}{\partial x_{l'}} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f \right) \right] \right) \right) \\
 & \quad + \sum_{l=1}^n \left( \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} a_{l\hat{l}'}^T \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \left( a_{k\hat{k}}^T \frac{\partial}{\partial x_{l'}} \frac{\partial}{\partial x_{\hat{k}}} f \right) \right] \right) \right) \\
 & \quad + \sum_{l=1}^n \left( \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} a_{l\hat{l}'}^T \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial^2}{\partial x_{\hat{i}} x_{l'}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right).
 \end{aligned}$$

Subtracting the above two terms, we have

$$\begin{aligned}
 & \mathbf{I} - \mathbf{II} \\
 &= - \sum_{l=1}^n \left( \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} a_{ll'}^T \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial^2}{\partial x_{\hat{i}} x_{l'}} a_{ik} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \\
 &+ \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{ik} \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \left( \sum_{l'=1}^{n+m} \frac{\partial}{\partial x_{\hat{k}}} a_{ll'}^T \frac{\partial}{\partial x_{l'}} f \right) \right) \\
 &- \sum_{l=1}^n \left( \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} a_{ll'}^T \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{ik} \left( \frac{\partial}{\partial x_{l'}} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f \right) \right] \right) \right) \\
 &= - \sum_{l=1}^n \left( \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} a_{ll'}^T \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial^2}{\partial x_{\hat{i}} x_{l'}} a_{ik} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \\
 &+ \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{\hat{i}, \hat{k}, l'=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{ik} a_{k\hat{k}}^T \left( \frac{\partial}{\partial x_{\hat{k}}} a_{ll'}^T \frac{\partial}{\partial x_{l'}} f \right) \right. \right. \\
 &\quad \left. \left. - a_{ll'}^T \frac{\partial}{\partial x_{\hat{i}}} a_{ik} \left( \frac{\partial}{\partial x_{l'}} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f \right) \right) \right) \\
 &= - \sum_{l=1}^n \left( \left( \sum_{\hat{l}=1}^{n+m} a_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} a_{ll'}^T \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial^2}{\partial x_{\hat{i}} x_{l'}} a_{ik} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \\
 &+ \sum_{l=1}^n (a^T \nabla f)_l \left( \sum_{\hat{i}, \hat{k}, l'=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{ik} a_{k\hat{k}}^T \left( \frac{\partial}{\partial x_{\hat{k}}} a_{ll'}^T \frac{\partial}{\partial x_{l'}} f \right) \right. \right. \\
 &\quad \left. \left. - a_{ll'}^T \frac{\partial}{\partial x_{\hat{i}}} a_{ik} \left( \frac{\partial}{\partial x_{l'}} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f \right) \right) \right).
 \end{aligned}$$

Now, we eventually obtain the the following step:

$$\begin{aligned}
 & \frac{1}{2} \Delta_a \Gamma_1(f, f) - \Gamma_1(\Delta_a f, f) \\
 &= \frac{1}{2} (a^T \nabla \circ (a^T \nabla |a^T \nabla f|^2)) - \langle a^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n} \\
 &- \langle \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^{n+m} \sum_{j'=1}^{n+m} a_{lj'}^T \frac{\partial^2}{\partial x_i x_{j'}} a_{ik} (a^T \nabla f)_k, a^T \nabla f \rangle \\
 &+ \sum_{l=1}^n (a^T \nabla f)_l \left( \sum_{\hat{i}, \hat{k}, l'=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{ik} a_{k\hat{k}}^T \left( \frac{\partial}{\partial x_{\hat{k}}} a_{ll'}^T \frac{\partial}{\partial x_{l'}} f \right) - a_{ll'}^T \frac{\partial}{\partial x_{\hat{i}}} a_{ik} \left( \frac{\partial}{\partial x_{l'}} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f \right) \right) \right) \\
 &= \frac{1}{2} (a^T \nabla \circ (a^T \nabla |a^T \nabla f|^2)) \\
 &- \langle a^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n} - \langle B_{n \times n} a^T \nabla f, a^T \nabla f \rangle_{\mathbb{R}^n} \\
 &+ \sum_{l=1}^n (a^T \nabla f)_l \left( \sum_{\hat{i}, \hat{k}, l'=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{ik} a_{k\hat{k}}^T \left( \frac{\partial}{\partial x_{\hat{k}}} a_{ll'}^T \frac{\partial}{\partial x_{l'}} f \right) - a_{ll'}^T \frac{\partial}{\partial x_{\hat{i}}} a_{ik} \left( \frac{\partial}{\partial x_{l'}} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f \right) \right) \right).
 \end{aligned}$$

Thus, the proof is completed.  $\square$

Below, we further investigate the extra term explicitly in the above Lemma 15.

**Lemma 16.**

$$\begin{aligned}
 & \frac{1}{2}(a^T \nabla \circ (a^T \nabla |a^T \nabla f|^2)) - \langle a^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n} \\
 = & X^T Q^T Q X + 2D^T Q X + 2C^T X + D^T D \\
 & + \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{ii'}^T (\frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (a^T \nabla)_{kf} \rangle_{\mathbb{R}^n} \\
 & + \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{ii'}^T a_{i\hat{i}}^T (\frac{\partial}{\partial x_{i'}} \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}}) (\frac{\partial f}{\partial x_{\hat{k}}}), (a^T \nabla)_{kf} \rangle_{\mathbb{R}^n} \\
 & - \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle (a_{k\hat{k}}^T \frac{\partial a_{ii'}^T}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^T \nabla)_{kf} \rangle_{\mathbb{R}^n} \\
 & - \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{k\hat{k}}^T a_{ii'}^T (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^T \nabla)_{kf} \rangle_{\mathbb{R}^n}. \tag{54}
 \end{aligned}$$

Recall that matrix  $Q$  and vectors  $X$ ,  $C$ , and  $D$  are defined in Notation 1.

**Proof.** We expand the two terms in Lemma 16. The first term reads as

$$\begin{aligned}
 & \frac{1}{2}(a^T \nabla \circ (a^T \nabla |a^T \nabla f|^2)) \\
 = & \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^n (a^T \nabla)_i (a^T \nabla)_i |a^T \nabla f|^2 \\
 = & \sum_{i=1}^n \sum_{k=1}^n (a^T \nabla)_i \langle (a^T \nabla)_i (a^T \nabla)_{kf}, (a^T \nabla)_{kf} \rangle_{\mathbb{R}^n} \\
 = & \sum_{i=1}^n \sum_{k=1}^n \langle (a^T \nabla)_i (a^T \nabla)_{kf}, (a^T \nabla)_i (a^T \nabla)_{kf} \rangle_{\mathbb{R}^n} \cdots \mathbf{T}_1 \\
 & + \sum_{i=1}^n \sum_{k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle (a_{ii'}^T \frac{\partial}{\partial x_{i'}}) (a_{i\hat{i}}^T \frac{\partial}{\partial x_{\hat{i}}}) (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}}) f, (a^T \nabla)_{kf} \rangle_{\mathbb{R}^n} \cdots \mathbf{R}_1.
 \end{aligned}$$

The second term reads as

$$\begin{aligned}
 & \langle a^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), a^T \nabla f \rangle_{\mathbb{R}^n} \\
 = & \sum_{i,k=1}^n \langle (a^T \nabla)_k [(a^T \nabla)_i (a^T \nabla)_{if}], (a^T \nabla)_{kf} \rangle \\
 = & \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}}) [(a_{ii'}^T \frac{\partial}{\partial x_{i'}}) (a_{i\hat{i}}^T \frac{\partial}{\partial x_{\hat{i}}}) f], (a^T \nabla)_{kf} \rangle \cdots \mathbf{R}_2.
 \end{aligned}$$

Next, we expand  $\mathbf{R}_1$  and  $\mathbf{R}_2$  completely and obtain the following:

$$\begin{aligned}
 \mathbf{R}_1 &= \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle (a_{ii'}^T \frac{\partial}{\partial x_{i'}}) (a_{i\hat{i}}^T \frac{\partial}{\partial x_{\hat{i}}}) (a_{k\hat{k}}^T \frac{\partial f}{\partial x_{\hat{k}}}), (a^T \nabla)_{kf} \rangle_{\mathbb{R}^n} \\
 &= \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{ii'}^T (\frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (a^T \nabla)_{kf} \rangle_{\mathbb{R}^n} \cdots \mathbf{R}_1^1 \\
 &+ \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{ii'}^T a_{i\hat{i}}^T (\frac{\partial}{\partial x_{i'}} \frac{\partial a_{k\hat{k}}^T}{\partial x_{\hat{i}}}) (\frac{\partial f}{\partial x_{\hat{k}}}), (a^T \nabla)_{kf} \rangle_{\mathbb{R}^n} \cdots \mathbf{R}_1^2
 \end{aligned}$$

$$\begin{aligned}
& + \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{ii'}^\top a_{ii}^\top (\frac{\partial a_{k\hat{k}}^\top}{\partial x_{\hat{i}}}) (\frac{\partial}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{k}}}), (a^\top \nabla)_k f \rangle_{\mathbb{R}^n} \cdots \mathbf{R}_1^3 \\
& + \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle (a_{ii'}^\top) ((\frac{\partial}{\partial x_{i'}} a_{ii}^\top) a_{k\hat{k}}^\top \frac{\partial}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (a^\top \nabla)_k f \rangle_{\mathbb{R}^n} \cdots \mathbf{R}_1^4 \\
& + \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{ii'}^\top a_{ii}^\top (\frac{\partial}{\partial x_{i'}} a_{k\hat{k}}^\top) \frac{\partial}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, (a^\top \nabla)_k f \rangle_{\mathbb{R}^n} \cdots \mathbf{R}_1^5 \\
& + \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{ii'}^\top a_{ii}^\top a_{k\hat{k}}^\top (\frac{\partial}{\partial x_{i'}} \frac{\partial}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (a^\top \nabla)_k f \rangle_{\mathbb{R}^n} \cdots \mathbf{R}_1^6.
\end{aligned}$$

Additionally,

$$\begin{aligned}
\mathbf{R}_2 &= \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle (a_{k\hat{k}}^\top \frac{\partial}{\partial x_{\hat{k}}}) [(a_{ii'}^\top \frac{\partial}{\partial x_{i'}}) (a_{ii}^\top \frac{\partial f}{\partial x_{\hat{i}}})], (a^\top \nabla)_k f \rangle \\
&= \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{k\hat{k}}^\top \frac{\partial a_{ii'}^\top}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}, (a^\top \nabla)_k f \rangle \cdots \mathbf{R}_2^1 \\
&+ \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{k\hat{k}}^\top a_{ii'}^\top (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}}) \frac{\partial f}{\partial x_{\hat{i}}}, (a^\top \nabla)_k f \rangle \cdots \mathbf{R}_2^2 \\
&+ \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{k\hat{k}}^\top a_{ii'}^\top \frac{\partial a_{ii}^\top}{\partial x_{i'}} (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^\top \nabla)_k f \rangle \cdots \mathbf{R}_2^3 = \mathbf{R}_1^4 \\
&+ \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{k\hat{k}}^\top \frac{\partial a_{ii'}^\top}{\partial x_{\hat{k}}} a_{ii}^\top (\frac{\partial}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^\top \nabla)_k f \rangle \cdots \mathbf{R}_2^4 \\
&+ \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{k\hat{k}}^\top a_{ii'}^\top \frac{\partial a_{ii}^\top}{\partial x_{\hat{k}}} (\frac{\partial}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^\top \nabla)_k f \rangle \cdots \mathbf{R}_2^5 \\
&+ \sum_{i,k=1}^n \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{k\hat{k}}^\top a_{ii'}^\top a_{ii}^\top (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^\top \nabla)_k f \rangle \cdots \mathbf{R}_2^6 = \mathbf{R}_1^6.
\end{aligned}$$

Our next step is to complete the squares for all the above terms. Look at the term  $\mathbf{T}_1$  first.

$$\begin{aligned}
\mathbf{T}_1 &= \sum_{i,k=1}^n \left\langle \sum_{\hat{i},\hat{k}=1}^{n+m} a_{ii}^\top a_{k\hat{k}}^\top \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}} + \sum_{\hat{i},\hat{k}=1}^{n+m} a_{ii}^\top \frac{\partial a_{k\hat{k}}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, \right. \\
&\quad \left. \sum_{i',k'=1}^{n+m} a_{ii'}^\top a_{kk'}^\top \frac{\partial^2 f}{\partial x_{i'} \partial x_{k'}} + \sum_{i',k'=1}^{n+m} a_{ii'}^\top \frac{\partial a_{kk'}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}} \right\rangle \\
&= \sum_{i,k=1}^n \left\langle \sum_{\hat{i},\hat{k}=1}^{n+m} a_{ii}^\top a_{k\hat{k}}^\top \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}}, \sum_{i',k'=1}^{n+m} a_{ii'}^\top a_{kk'}^\top \frac{\partial^2 f}{\partial x_{i'} \partial x_{k'}} \right\rangle \cdots \mathbf{T}_{1a} \\
&+ \sum_{i,k=1}^n \left\langle \sum_{\hat{i},\hat{k}=1}^{n+m} a_{ii}^\top a_{k\hat{k}}^\top \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}}, \sum_{i',k'=1}^{n+m} a_{ii'}^\top \frac{\partial a_{kk'}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}} \right\rangle \cdots \mathbf{T}_{1b} \\
&+ \sum_{i,k=1}^n \left\langle \sum_{\hat{i},\hat{k}=1}^{n+m} a_{ii}^\top \frac{\partial a_{k\hat{k}}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, \sum_{i',k'=1}^{n+m} a_{ii'}^\top a_{kk'}^\top \frac{\partial^2 f}{\partial x_{i'} \partial x_{k'}} \right\rangle \cdots \mathbf{T}_{1c} \\
&+ \sum_{i,k=1}^n \left\langle \sum_{\hat{i},\hat{k}=1}^{n+m} a_{ii}^\top \frac{\partial a_{k\hat{k}}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, \sum_{i',k'=1}^{n+m} a_{ii'}^\top \frac{\partial a_{kk'}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}} \right\rangle \cdots \mathbf{T}_{1d}.
\end{aligned}$$



The terms  $\mathbf{T}_{1b} = \mathbf{T}_{1c}$ ,  $\mathbf{R}_1^3 = \mathbf{R}_1^5$ , and  $\mathbf{R}_2^5 = \mathbf{R}_2^4$  play the role of crossing terms inside the complete squares. In particular, for convenience, we change the index inside the sum of  $\mathbf{R}_1^3$  and  $\mathbf{R}_2^5$ , switching  $i', \hat{i}$  for  $\mathbf{R}_1^3$  and switching  $i', \hat{k}$  for  $\mathbf{R}_2^5$ . Then, we obtain the following.

$$\begin{aligned} 2\mathbf{R}_1^3 &= 2 \sum_{i,k=1}^n \sum_{i',\hat{k}=1}^{n+m} \left\langle a_{ii}^T a_{i'i'}^T \left( \frac{\partial a_{k\hat{k}}^T}{\partial x_{i'}} \right) \left( \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_{\hat{k}}} \right), (a^T \nabla)_{kf} \right\rangle \\ &= 2 \sum_{i,k=1}^n \sum_{\hat{i},\hat{k}=1}^{n+m} \sum_{i',l=1}^{n+m} \left( a_{ii}^T a_{i'i'}^T \left( \frac{\partial a_{k\hat{k}}^T}{\partial x_{i'}} \right) \left( \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_{\hat{k}}} \right) a_{kl}^T \frac{\partial f}{\partial x_l} \right) \\ -2\mathbf{R}_2^5 &= -2 \sum_{i,k=1}^n \sum_{i',\hat{k}=1}^{n+m} \left\langle a_{ki'}^T a_{ik}^T \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_i} \right), (a^T \nabla)_{kf} \right\rangle \\ &= -2 \sum_{i,k=1}^n \sum_{\hat{i},\hat{k}=1}^{n+m} \sum_{i',l=1}^{n+m} \left( a_{ki'}^T a_{ik}^T \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_i} \right) a_{kl}^T \frac{\partial f}{\partial x_l} \right). \end{aligned}$$

We denote

$$\sum_{\hat{i},\hat{k}=1}^{n+m} a_{ii}^T a_{k\hat{k}}^T \frac{\partial^2 f}{\partial x_i \partial x_{\hat{k}}} = \gamma_{ik}. \quad (55)$$

The above Equality (55) can be represented in the following matrix form:

$$Q_{n^2 \times (n+m)^2} X_{(n+m)^2 \times 1} = (\gamma_{11}, \dots, \gamma_{ik}, \dots, \gamma_{nn})_{n^2 \times 1}^T,$$

where  $Q$  and  $X$  are defined in (12) and (19). Now, we can represent term  $\mathbf{T}_{1a}$  as  $\sum_{i,k=1}^n \gamma_{ik}^2 = \gamma^T \gamma = (QX)^T QX = X^T Q^T QX$ . Next, we want to represent  $\mathbf{R}_1^3$  and  $\mathbf{R}_2^5$  in the following form in terms of vector  $X$ :

$$\begin{aligned} 2\mathbf{R}_1^3 - 2\mathbf{R}_2^5 &= 2 \sum_{i,k=1}^n \sum_{i',\hat{k}=1}^{n+m} \left( \left\langle a_{ii}^T a_{i'i'}^T \left( \frac{\partial a_{k\hat{k}}^T}{\partial x_{i'}} \right) \left( \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_{\hat{k}}} \right), (a^T \nabla)_{kf} \right\rangle \right)_{\mathbb{R}^n} \\ &\quad - 2 \sum_{i,k=1}^n \sum_{i',\hat{k}=1}^{n+m} \left\langle a_{ki'}^T a_{ik}^T \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_i} \right), (a^T \nabla)_{kf} \right\rangle \\ &= 2 \sum_{\hat{i},\hat{k}=1}^{n+m} \left[ \sum_{i,k=1}^n \sum_{i'=1}^{n+m} \left( \left\langle a_{ii}^T a_{i'i'}^T \left( \frac{\partial a_{k\hat{k}}^T}{\partial x_{i'}} \right), (a^T \nabla)_{kf} \right\rangle - \left\langle a_{ki'}^T a_{ik}^T \frac{\partial a_{i\hat{i}}^T}{\partial x_{i'}} \left( \frac{\partial}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_i} \right), (a^T \nabla)_{kf} \right\rangle \right) \right] \left( \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_{\hat{k}}} \right) \\ &= 2C^T X, \end{aligned}$$

where  $C$  is defined in (14). Similarly, we can represent  $\mathbf{T}_{1b} = \mathbf{T}_{1c}$  by  $X$ :

$$\begin{aligned} \mathbf{T}_{1b} = \mathbf{T}_{1c} &= \sum_{i,k=1}^n \left\langle \sum_{\hat{i},\hat{k}=1}^{n+m} a_{ii}^T \frac{\partial a_{k\hat{k}}^T}{\partial x_i} \frac{\partial f}{\partial x_{\hat{k}}}, \sum_{i',k'=1}^{n+m} a_{i'i'}^T a_{kk'}^T \frac{\partial^2 f}{\partial x_{i'} \partial x_{k'}} \right\rangle \\ &= D^T QX, \end{aligned}$$

where  $D$  is defined in (1). Summing over the above terms, we have the following quadratic form:

$$\mathbf{T}_1 + 2\mathbf{R}_1^3 - 2\mathbf{R}_2^5 = X^T Q^T QX + 2D^T QX + 2C^T X + D^T D. \quad (56)$$

Taking into account the fact that  $\mathbf{R}_1^6 - \mathbf{R}_2^6 = 0$  and  $\mathbf{R}_1^4 - \mathbf{R}_2^3 = 0$ , we have

$$\mathbf{T}_1 + \mathbf{R}_1 - \mathbf{R}_2 = \mathbf{T}_1 + 2\mathbf{R}_1^3 - 2\mathbf{R}_2^5 + \mathbf{R}_1^1 + \mathbf{R}_1^2 - \mathbf{R}_2^1 - \mathbf{R}_2^2,$$

which completes the proof.  $\square$

### 5.3. Proof of Lemma 12

**Lemma 17.**

$$\Gamma_{2,\tilde{L}}^z(f, f) = X^T P^T P X + 2E^T P X + 2F^T X + E^T E + \mathfrak{R}_z(\nabla f, \nabla f). \quad (57)$$

where  $\mathfrak{R}_z$  is defined in Definition 2.

**Proof.** The proof follows directly from Lemmas 18 and 19.  $\square$

**Lemma 18.**

$$\begin{aligned} & \frac{1}{2} \tilde{L} \Gamma_1^z(f, f) - \Gamma_1^z(\tilde{L} f, f) \\ &= \frac{1}{2} (a^T \nabla \circ (a^T \nabla |z^T \nabla f|^2)) - \langle z^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m}. \end{aligned}$$

**Proof. Step 1:** We first define  $\Gamma_1^z = \langle z^T \nabla f, z^T \nabla f \rangle_{\mathbb{R}^m}$ , then we have

$$\tilde{L} \Gamma_1^z(f, f) = \Delta_p \Gamma_1^z(f, f) - A \nabla \Gamma_1^z(f, f), \quad \Gamma_1^z(\tilde{L} f, f) = \Gamma_1^z(\Delta_p f, f) - \Gamma_1^z(A \nabla f, f).$$

By our definition above, we directly obtain

$$\begin{aligned} \Delta_a \Gamma_1^z(f, f) &= \nabla \cdot (a a^T \nabla \langle z^T \nabla f, z^T \nabla f \rangle_{\mathbb{R}^m}) = \nabla \cdot (a F^z) \\ &= \sum_{\hat{i}=1}^{n+m} \frac{\partial}{\partial x_{\hat{i}}} \left( \sum_{k=1}^n a_{\hat{i}k} F_k^z \right) \\ &= \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} F_k^z + a_{\hat{i}k} \frac{\partial}{\partial x_{\hat{i}}} F_k^z \right) \\ &= \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} F_k^z \right) + a^T \nabla \circ (a^T \nabla (z^T \nabla f)^2), \end{aligned}$$

where we denote

$$\begin{aligned} F^z &= a^T \nabla \langle z^T \nabla f, z^T \nabla f \rangle_{\mathbb{R}^m} = a^T \nabla \sum_{l=1}^m \left( \sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right)^2 \\ &= \left( \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} \sum_{l=1}^m \left( \sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right)^2 \right)_{k=1, \dots, n} = (F_1^z, F_2^z, \dots, F_n^z)^T. \end{aligned}$$

We have

$$\begin{aligned} & \Delta_a \Gamma_1^z(f, f) \\ &= \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \left( \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} \sum_{l=1}^m \left( \sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right)^2 \right) \right) + a^T \nabla \circ (a^T \nabla (z^T \nabla f)^2) \\ &= \sum_{k=1}^n \sum_{\hat{i}=1}^{n+m} \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \left( \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} (z^T \nabla f)^2 \right) \right) + (a^T \nabla) \circ (a^T \nabla (z^T \nabla f)^2) \\ &= \nabla a \circ (a^T \nabla (z^T \nabla f)^2) + (a^T \nabla) \circ (a^T \nabla (z^T \nabla f)^2). \end{aligned} \quad (58)$$

Next, we compute the following quantity.

$$\Gamma_1^z(\Delta_a f, f) = \langle z^T \nabla (\nabla \cdot (a a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m}.$$

From Lemma 15, we have

$$\nabla \cdot (aa^T \nabla f) = \nabla a \circ (a^T \nabla f) + (a^T \nabla) \circ (a^T \nabla f).$$

We continue with our computation as below:

$$\begin{aligned} & \Gamma_1^z(\Delta_a f, f) \\ = & \langle z^T \nabla \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) + (a^T \nabla) \circ (a^T \nabla f) \right], z^T \nabla f \rangle_{\mathbb{R}^m} \\ = & \langle z^T \nabla \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right], z^T \nabla f \rangle_{\mathbb{R}^m} + \langle z^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m} \\ = & \langle z^T \nabla (\nabla a \cdot (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m} \end{aligned} \quad (59)$$

$$\begin{aligned} & + \langle z^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m} \\ = & \langle (z^T \nabla \nabla a \cdot (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m} + \langle (\nabla a \cdot (z^T \nabla (a^T \nabla f))), z^T \nabla f \rangle_{\mathbb{R}^m} \\ & + \langle z^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m}. \end{aligned} \quad (60)$$

From the above, combining (58) and (59), we further obtain

$$\begin{aligned} & \frac{1}{2} \Delta_a \Gamma_1^z(f, f) - \Gamma_1^z(\Delta_a f, f) \\ = & \frac{1}{2} (a^T \nabla \circ (a^T \nabla |z^T \nabla f|^2)) - \langle z^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m} \\ & + \frac{1}{2} \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \left( \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} \sum_{l=1}^n \left( \sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right)^2 \right) \right) \\ & - \langle z^T \nabla \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right], z^T \nabla f \rangle_{\mathbb{R}^m} \\ = & \frac{1}{2} (a^T \nabla \circ (a^T \nabla |z^T \nabla f|^2)) - \langle z^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m} \\ & + \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} \sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^T \sum_{l=1}^m \left( \sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \frac{\partial}{\partial x_{\hat{k}}} \left( \sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \right) \cdots \mathbf{I} \\ & - \sum_{l=1}^m \left( \left( \sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} z_{l'l'}^T \frac{\partial}{\partial x_{l'}} \left[ \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (a_{k\hat{k}}^T \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \cdots \mathbf{II}. \end{aligned}$$

Recall here that we denote  $a^T$  to emphasize the transpose of the matrix  $a$  and  $a_{ii}^T = a_{ii}$ :



Now, we eventually end up with the following formula:

$$\begin{aligned} & \frac{1}{2} \Delta_a \Gamma_1^z(f, f) - \Gamma_1^z(\Delta_a f, f) \\ = & \frac{1}{2} (a^T \nabla \circ (a^T \nabla |z^T \nabla f|^2)) - \langle z^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m} \\ & - \sum_{l=1}^m \left( \left( \sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} z_{ll'}^T \left[ \sum_{\hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial}{\partial x_{\hat{k}}} a_{ik} \left( \frac{\partial}{\partial x_{l'}} a_{kk}^T \frac{\partial}{\partial x_{\hat{k}}} f \right) \right] \right) \right) \\ & - \sum_{l=1}^m \left( \left( \sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \left( \sum_{l'=1}^{n+m} z_{ll'}^T \left[ \sum_{\hat{k}=1}^{n+m} \sum_{k=1}^n \frac{\partial^2}{\partial x_{\hat{k}} \partial x_{l'}} a_{ik} \left( a_{kk}^T \frac{\partial}{\partial x_{\hat{k}}} f \right) \right] \right) \right) \\ & + \sum_{l=1}^m \left( \sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^T \frac{\partial}{\partial x_{\hat{l}}} f \right) \sum_{\hat{i}=1}^{n+m} \sum_{k=1}^n \left( \frac{\partial}{\partial x_{\hat{i}}} a_{ik} \sum_{\hat{k}=1}^{n+m} a_{kk}^T \left( \sum_{l'=1}^{n+m} \frac{\partial}{\partial x_{\hat{k}}} z_{ll'}^T \frac{\partial}{\partial x_{l'}} f \right) \right). \end{aligned}$$

**Step 2:** Computation of  $-\frac{1}{2} A \nabla \Gamma_1^z(f, f) + \Gamma_1^z(A \nabla f, f)$ . Now, we compute the last two terms of the above equation, with  $A = a \otimes \nabla a$ :

$$\begin{aligned} -\frac{1}{2} A \nabla \Gamma_1^z(f, f) &= -\frac{1}{2} \sum_{\hat{k}=1}^{n+m} A_{\hat{k}} \nabla_{\frac{\partial}{\partial x_{\hat{k}}}} \langle z^T \nabla f, z^T \nabla f \rangle_{\mathbb{R}^m} \\ &= -\sum_{\hat{k}=1}^{n+m} \langle A_{\hat{k}} (\nabla_{\frac{\partial}{\partial x_{\hat{k}}}} z^T) \nabla f, z^T \nabla f \rangle_{\mathbb{R}^m} - \sum_{\hat{k}=1}^{n+m} \langle A_{\hat{k}} z^T (\nabla_{\frac{\partial}{\partial x_{\hat{k}}}} \nabla f), z^T \nabla f \rangle_{\mathbb{R}^m} \\ &= \tilde{\mathbf{J}}_1 + \tilde{\mathbf{J}}_2, \end{aligned}$$

$$\begin{aligned} \Gamma_1^z(A \nabla f, f) &= \langle z^T \nabla \left( \sum_{\hat{k}=1}^{n+m} A_{\hat{k}} \nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f \right), z^T \nabla f \rangle_{\mathbb{R}^m} \\ &= \langle z^T \left( \sum_{\hat{k}=1}^{n+m} A_{\hat{k}} \nabla \nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f \right), z^T \nabla f \rangle_{\mathbb{R}^m} + \langle z^T \left( \sum_{\hat{k}=1}^{n+m} \nabla A_{\hat{k}} \nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f \right), z^T \nabla f \rangle_{\mathbb{R}^m} \\ &= \tilde{\mathbf{J}}_3 + \tilde{\mathbf{J}}_4. \end{aligned}$$

It is easy to see that  $\tilde{\mathbf{J}}_2 + \tilde{\mathbf{J}}_3 = 0$ . We now expand  $\tilde{\mathbf{J}}_1$  and  $\tilde{\mathbf{J}}_4$  into local coordinates:

$$\tilde{\mathbf{J}}_1 = - \sum_{l=1}^m (z^T \nabla f)_l \left( \sum_{l'=1}^{n+m} \sum_{\hat{k}=1}^n \sum_{k'=1}^{n+m} a_{\hat{k}k'} \nabla_{\frac{\partial}{\partial x_{l'}}} a_{k'l} \nabla_{\frac{\partial}{\partial x_{\hat{k}}}} z_{ll'}^T \nabla_{\frac{\partial}{\partial x_{l'}}} f \right), \quad (61)$$

$$\begin{aligned} \tilde{\mathbf{J}}_4 &= \sum_{l=1}^m (z^T \nabla f)_l \left( \sum_{l'=1}^{n+m} z_{ll'}^T \left( \sum_{\hat{k}=1}^{n+m} \nabla_{\frac{\partial}{\partial x_{l'}}} \left( \sum_{k=1}^n \sum_{k'=1}^{n+m} a_{\hat{k}k'} \nabla_{\frac{\partial}{\partial x_{l'}}} a_{k'l} \nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f \right) \right) \right) \\ &= \sum_{l=1}^m (z^T \nabla f)_l \left( \sum_{k=1}^n \sum_{l'=1}^{n+m} \sum_{\hat{k}, k'=1}^{n+m} z_{ll'}^T \nabla_{\frac{\partial}{\partial x_{l'}}} a_{\hat{k}k'} \nabla_{\frac{\partial}{\partial x_{l'}}} a_{k'l} \nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f \right) \\ &\quad + \sum_{l=1}^m (z^T \nabla f)_l \left( \sum_{k=1}^n \sum_{l'=1}^{n+m} \sum_{\hat{k}, k'=1}^{n+m} z_{ll'}^T a_{\hat{k}k'} \left( \nabla_{\frac{\partial}{\partial x_{l'}}} \nabla_{\frac{\partial}{\partial x_{l'}}} a_{k'l} \nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f \right) \right). \end{aligned} \quad (62)$$

Combining the above two steps, we thus obtain

$$\frac{1}{2} \tilde{L} \Gamma_1^z(f, f) - \Gamma_1^z(\tilde{L} f, f) = \frac{1}{2} (a^T \nabla \circ (a^T \nabla |z^T \nabla f|^2)) - \langle z^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m}.$$

□

**Lemma 19.**

$$\begin{aligned} & \frac{1}{2}(a^T \nabla \circ (a^T \nabla |z^T \nabla f|^2)) - \langle z^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m} \\ &= X^T P^T P X + 2E^T P X + 2F^T X + E^T E + \Re_z(\nabla f, \nabla f). \end{aligned}$$

**Proof.** We expand the two terms in Lemma 19.

$$\begin{aligned} & \frac{1}{2}(a^T \nabla \circ (a^T \nabla |z^T \nabla f|^2)) \\ &= \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^m (a^T \nabla)_i (a^T \nabla)_i |(z^T \nabla)_k f|^2 \\ &= \sum_{i=1}^n \sum_{k=1}^m (a^T \nabla)_i \langle (a^T \nabla)_i (z^T \nabla)_k f, (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \\ &= \sum_{i=1}^n \sum_{k=1}^m \langle (a^T \nabla)_i (z^T \nabla)_k f, (a^T \nabla)_i (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{T}}_1 \\ & \quad + \sum_{i=1}^n \sum_{k=1}^m \langle (a_{ii'}^T \frac{\partial}{\partial x_{i'}})(a_{ii}^T \frac{\partial}{\partial x_i})(z_{kk}^T \frac{\partial}{\partial x_k})f, (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_1. \\ & \langle z^T \nabla ((a^T \nabla) \circ (a^T \nabla f)), z^T \nabla f \rangle_{\mathbb{R}^m} \\ &= \sum_{i=1}^n \sum_{k=1}^m \langle (z^T \nabla)_k [(a^T \nabla)_i (a^T \nabla)_i f], (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \\ &= \sum_{i=1}^n \sum_{k=1}^m \langle (z_{kk}^T \frac{\partial}{\partial x_k})[(a_{ii'}^T \frac{\partial}{\partial x_{i'}})(a_{ii}^T \frac{\partial}{\partial x_i})f], (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_2. \end{aligned}$$

Next, we expand  $\tilde{\mathbf{R}}_1$  and  $\tilde{\mathbf{R}}_2$  completely and obtain the following:

$$\begin{aligned} \tilde{\mathbf{R}}_1 &= \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle (a_{ii'}^T \frac{\partial}{\partial x_{i'}})(a_{ii}^T \frac{\partial}{\partial x_i})(z_{kk}^T \frac{\partial}{\partial x_k})f, (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \\ &= \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{ii'}^T (\frac{\partial a_{ii}^T}{\partial x_{i'}} \frac{\partial z_{kk}^T}{\partial x_i} \frac{\partial f}{\partial x_k}), (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_1^1 \\ & \quad + \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{ii'}^T a_{ii}^T (\frac{\partial}{\partial x_{i'}} \frac{\partial z_{kk}^T}{\partial x_i}) (\frac{\partial f}{\partial x_k}), (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_1^2 \\ & \quad + \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{ii'}^T a_{ii}^T (\frac{\partial z_{kk}^T}{\partial x_i}) (\frac{\partial}{\partial x_{i'}} \frac{\partial f}{\partial x_k}), (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_1^3 \\ & \quad + \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle (a_{ii'}^T) ((\frac{\partial}{\partial x_{i'}} a_{ii}^T) z_{kk}^T \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_k}), (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_1^4 \\ & \quad + \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{ii'}^T a_{ii}^T (\frac{\partial}{\partial x_{i'}} z_{kk}^T) \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_k}, (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_1^5 \\ & \quad + \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{ii'}^T a_{ii}^T z_{kk}^T (\frac{\partial}{\partial x_{i'}} \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_k}), (z^T \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_1^6 \end{aligned}$$

$$\begin{aligned}
\tilde{\mathbf{R}}_2 &= \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle (z_{k\hat{k}}^\top \frac{\partial}{\partial x_{\hat{k}}}) [(a_{ii'}^\top \frac{\partial}{\partial x_{i'}})(a_{ii}^\top \frac{\partial f}{\partial x_{\hat{i}}})], (z^\top \nabla)_k f \rangle_{\mathbb{R}^m} \\
&= \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle z_{k\hat{k}}^\top \frac{\partial a_{ii'}^\top}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}, (z^\top \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_2^1 \\
&\quad + \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle z_{k\hat{k}}^\top a_{ii'}^\top (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{ii}^\top}{\partial x_{i'}}) \frac{\partial f}{\partial x_{\hat{i}}}, (z^\top \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_2^2 \\
&\quad + \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle z_{k\hat{k}}^\top a_{ii'}^\top \frac{\partial a_{ii}^\top}{\partial x_{i'}} (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^\top \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_2^3 = \tilde{\mathbf{R}}_1^4 \\
&\quad + \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle z_{k\hat{k}}^\top \frac{\partial a_{ii'}^\top}{\partial x_{\hat{k}}} a_{ii}^\top (\frac{\partial}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^\top \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_2^4 \\
&\quad + \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle z_{k\hat{k}}^\top a_{ii'}^\top \frac{\partial a_{ii}^\top}{\partial x_{\hat{k}}} (\frac{\partial}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^\top \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_2^5 \\
&\quad + \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle z_{k\hat{k}}^\top a_{ii'}^\top a_{ii}^\top (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^\top \nabla)_k f \rangle_{\mathbb{R}^m} \cdots \tilde{\mathbf{R}}_2^6 = \tilde{\mathbf{R}}_1^6
\end{aligned}$$

Our next step is to complete the squares for all the above terms. We look at term  $\tilde{\mathbf{T}}_1$  first.

$$\begin{aligned}
\tilde{\mathbf{T}}_1 &= \sum_{i=1}^n \sum_{k=1}^m \left\langle \sum_{\hat{i}, \hat{k}=1}^{n+m} a_{ii}^\top z_{k\hat{k}}^\top \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}} + \sum_{\hat{i}, \hat{k}=1}^{n+m} a_{ii}^\top \frac{\partial z_{k\hat{k}}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, \right. \\
&\quad \left. \sum_{i', k'=1}^{n+m} a_{ii'}^\top z_{k\hat{k}}^\top \frac{\partial^2 f}{\partial x_{i'} \partial x_{k'}} + \sum_{i', k'=1}^{n+m} a_{ii'}^\top \frac{\partial z_{k\hat{k}}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}} \right\rangle \\
&= \sum_{i=1}^n \sum_{k=1}^m \sum_{k=1}^m \left\langle \sum_{\hat{i}, \hat{k}=1}^{n+m} a_{ii}^\top z_{k\hat{k}}^\top \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}}, \sum_{i', k'=1}^{n+m} a_{ii'}^\top z_{k\hat{k}}^\top \frac{\partial^2 f}{\partial x_{i'} \partial x_{k'}} \right\rangle \cdots \tilde{\mathbf{T}}_{1a} \\
&\quad + \sum_{i=1}^n \sum_{k=1}^m \left\langle \sum_{\hat{i}, \hat{k}=1}^{n+m} a_{ii}^\top z_{k\hat{k}}^\top \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}}, \sum_{i', k'=1}^{n+m} a_{ii'}^\top \frac{\partial z_{k\hat{k}}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}} \right\rangle \cdots \tilde{\mathbf{T}}_{1b} \\
&\quad + \sum_{i=1}^n \sum_{k=1}^m \left\langle \sum_{\hat{i}, \hat{k}=1}^{n+m} a_{ii}^\top \frac{\partial z_{k\hat{k}}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, \sum_{i', k'=1}^{n+m} a_{ii'}^\top z_{k\hat{k}}^\top \frac{\partial^2 f}{\partial x_{i'} \partial x_{k'}} \right\rangle \cdots \tilde{\mathbf{T}}_{1c} \\
&\quad + \sum_{i=1}^n \sum_{k=1}^m \left\langle \sum_{\hat{i}, \hat{k}=1}^{n+m} a_{ii}^\top \frac{\partial z_{k\hat{k}}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, \sum_{i', k'=1}^{n+m} a_{ii'}^\top \frac{\partial z_{k\hat{k}}^\top}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}} \right\rangle \cdots \tilde{\mathbf{T}}_{1d}.
\end{aligned}$$

The terms  $\tilde{\mathbf{T}}_{1b} = \tilde{\mathbf{T}}_{1c}$ ,  $\tilde{\mathbf{R}}_1^3 = \tilde{\mathbf{R}}_1^5$ , and  $\tilde{\mathbf{R}}_2^5 = \tilde{\mathbf{R}}_2^4$  play the role of crossing terms inside the complete squares. In particular, for convenience, we changed the index inside the sum of  $\tilde{\mathbf{R}}_1^3$  and  $\tilde{\mathbf{R}}_2^5$ , switched  $i', \hat{i}$  for  $\tilde{\mathbf{R}}_1^3$ , and switched  $i', \hat{k}$  for  $\tilde{\mathbf{R}}_2^5$ , then we obtain the following.

$$\begin{aligned}
2\tilde{\mathbf{R}}_1^3 &= 2 \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{ii'}^\top a_{i'l}^\top (\frac{\partial z_{k\hat{k}}^\top}{\partial x_{i'}}) (\frac{\partial}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (z^\top \nabla)_{kf} \rangle_{\mathbb{R}^n} \\
&= 2 \sum_{i=1}^n \sum_{k=1}^m \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{i', l=1}^{n+m} \left( a_{ii'}^\top a_{i'l}^\top (\frac{\partial z_{k\hat{k}}^\top}{\partial x_{i'}}) (\frac{\partial}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}) z_{kl}^\top \frac{\partial f}{\partial x_l} \right) \\
-2\tilde{\mathbf{R}}_2^5 &= -2 \sum_{i=1}^n \sum_{k=1}^m \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle z_{ki'}^\top a_{ik}^\top \frac{\partial a_{ii'}^\top}{\partial x_{i'}} (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^\top \nabla)_{kf} \rangle \\
&= -2 \sum_{i=1}^n \sum_{k=1}^m \sum_{\hat{i}, \hat{k}=1}^{n+m} \sum_{i', l=1}^{n+m} \left( z_{ki'}^\top a_{ik}^\top \frac{\partial a_{ii'}^\top}{\partial x_{i'}} (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}) z_{kl}^\top \frac{\partial f}{\partial x_l} \right)
\end{aligned}$$

We denote

$$\sum_{\hat{i}, \hat{k}=1}^{n+m} a_{ii'}^\top z_{k\hat{k}}^\top \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}} = \omega_{ik}. \quad (63)$$

The above Equality (63) can be represented in the following matrix form:

$$P_{(n*m) \times (n+m)^2} X_{(n+m)^2 \times 1} = (\omega_{11}, \dots, \omega_{ik}, \dots, \omega_{nm})_{(n*m) \times 1}^\top$$

where  $P$  and  $X$  are defined in (13) and (19). Now, we can represent term  $\tilde{\mathbf{T}}_{1a}$  as  $\sum_{i=1}^n \sum_{k=1}^m \omega_{ik}^2 = \omega^\top \omega = (PX)^\top PX = X^\top P^\top PX$ . Next, we want to represent  $\tilde{\mathbf{R}}_1^3$  and  $\tilde{\mathbf{R}}_2^5$  in the following form in terms of vector  $X$ :

$$\begin{aligned}
&2\tilde{\mathbf{R}}_1^3 - 2\tilde{\mathbf{R}}_2^5 \\
&= 2 \sum_{i,k=1}^n \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle a_{ii'}^\top a_{i'l}^\top (\frac{\partial z_{k\hat{k}}^\top}{\partial x_{i'}}) (\frac{\partial}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (z^\top \nabla)_{kf} \rangle_{\mathbb{R}^n} \\
&\quad - 2 \sum_{i,k=1}^n \sum_{i', \hat{i}, \hat{k}=1}^{n+m} \langle z_{ki'}^\top a_{ik}^\top \frac{\partial a_{ii'}^\top}{\partial x_{i'}} (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^\top \nabla)_{kf} \rangle \\
&= 2 \sum_{\hat{i}, \hat{k}=1}^{n+m} \left[ \sum_{i,k=1}^n \sum_{i'=1}^{n+m} \left( \langle a_{ii'}^\top a_{i'l}^\top (\frac{\partial z_{k\hat{k}}^\top}{\partial x_{i'}}), (z^\top \nabla)_{kf} \rangle - \langle z_{ki'}^\top a_{ik}^\top \frac{\partial a_{ii'}^\top}{\partial x_{i'}} (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^\top \nabla)_{kf} \rangle \right) \right] (\frac{\partial}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}) \\
&= 2F^\top X,
\end{aligned}$$

where  $F$  is defined in (16). Similarly, we can represent  $\tilde{\mathbf{T}}_{1b} = \tilde{\mathbf{T}}_{1c}$  by  $X$ :

$$\tilde{\mathbf{T}}_{1b} = \tilde{\mathbf{T}}_{1c} = \sum_{i,k=1}^n \langle \sum_{\hat{i}, \hat{k}=1}^{n+m} a_{ii'}^\top \frac{\partial z_{k\hat{k}}^\top}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, \sum_{i', k'=1}^{n+m} a_{i'i'}^\top z_{k\hat{k}}^\top \frac{\partial^2 f}{\partial x_{i'} \partial x_{k'}} \rangle = E^\top PX$$

where  $E$  is defined in (17). We thus have the following form:

$$\tilde{\mathbf{T}}_1 + 2\tilde{\mathbf{R}}_1^3 - 2\tilde{\mathbf{R}}_2^5 = X^\top P^\top PX + 2E^\top PX + 2F^\top X + E^\top E$$

Taking into account the fact that  $\mathbf{R}_1^6 - \mathbf{R}_2^6 = 0$  and  $\mathbf{R}_1^4 - \mathbf{R}_2^3 = 0$ , we have

$$\tilde{\mathbf{T}}_1 + \tilde{\mathbf{R}}_1 - \tilde{\mathbf{R}}_2 = \tilde{\mathbf{T}}_1 + 2\tilde{\mathbf{R}}_1^3 - 2\tilde{\mathbf{R}}_2^5 + \tilde{\mathbf{R}}_1^1 + \tilde{\mathbf{R}}_1^2 - \tilde{\mathbf{R}}_2^1 - \tilde{\mathbf{R}}_2^2,$$

which completes the proof.  $\square$



## 6. Further Discussions on Other Inequalities

In this section, we apply the generalized Gamma calculus to study the entropic inequality for the semi-group  $P_t$  associated with the drift–diffusion process. With a little abuse of notation, we denote the generator of the semi-group  $P_t$  as  $\frac{1}{2}L$  instead of  $L$ , and we denote  $X_t$  as the corresponding diffusion process.

**Definition 5.** We define the semigroup  $P_t = e^{\frac{1}{2}tL}$ , where  $L$  is invariant with respect to the invariant measure  $d\mu = \pi(x)dx$ . We denote  $P_t f(x) = \mathbb{E}(f(X_t))$ , and

$$\mathbb{E}(f(X_t)) = \int_{\mathbb{R}^{n+m}} f(y)p(t, x, y)d\mu(y) = \int_{\mathbb{R}^{n+m}} f(y)\rho(t, x, y)dy,$$

where the infinitesimal generator of this process  $X_t$  is  $\frac{1}{2}L$ , and we denote  $\rho(t, \cdot, \cdot)$  as the product of the transition kernel  $p(t, \cdot, \cdot)$  and the volume measure  $\pi$ .

**Remark 14.** Following the standard treatment as in [2] (Section 5), whenever we consider the differentiating operation on  $P_t f$ , we shall always consider  $P_t f_\varepsilon$  first with  $f_\varepsilon = f + \varepsilon$ , for  $\forall \varepsilon > 0$ . Then, we take the limit as  $\varepsilon \rightarrow 0$ . Throughout this section, we directly use  $P_t f$  instead of  $P_t f_\varepsilon$  for convenience.

**Remark 15.** In the standard sub-Riemannian setting, the semi-groups are in general defined with respect to the invariant measure  $d\mu(y)$ . In this paper, we formulate the semi-group and the transition kernel with respect to the Lebesgue measure  $dy$ .

Following the framework in [2], we also need the following assumption, which is necessary to rigorously justify the computations on functionals of the heat semigroup.

**Assumption 2.** The semigroup  $P_t$  is stochastically complete, that is, for  $t \geq 0$ ,  $P_t \mathbf{1} = \mathbf{1}$  and for any  $T > 0$  and  $f \in C^\infty(\mathbb{R}^{n+m})$  with compact support, we assume that

$$\sup_{t \in [0, T]} \|\Gamma(P_t f)\|_\infty + \|\Gamma_1^z(P_t f)\|_\infty < +\infty. \quad (64)$$

We believe that the above Assumption 2 should follow from the the assumption  $\mathfrak{R} \geq \kappa(\Gamma_1 + \Gamma_1^z)$  if we assume the appropriate lower bound  $\kappa$ . We leave this for further studies. Related gradient estimates are presented in order below. For the infinitesimal generator  $\frac{1}{2}L$  associated with linear semi-group  $P_t$ , we have the following property.

**Proposition 15.** For all smooth function  $f$ , we have:

- $P_0 = Id$ ;
- For all functions  $f \in C_b(\mathbb{R}^{n+m})$ , the map  $t \mapsto P_t f$  is continuous from  $\mathbb{R}^+$  to  $L^2(d\mu)$ ;
- For all  $s, t \geq 0$ , one has  $P_t \circ P_s = P_{t+s}$ ;
- $\forall x \in \mathbb{R}^{n+m}, \forall t \geq 0, \frac{\partial}{\partial t} P_t f(x) = \frac{1}{2}L(P_t f)(x) = \frac{1}{2}P_t(Lf)(x)$ .

Next, we present the entropic inequality under Assumption 1. We follow closely the framework introduced in [2] and define the following two functionals:

$$\phi_a(x, t) = P_{T-t} f \Gamma_1(\log P_{T-t} f)(x), \quad \text{and} \quad \phi_z(x, t) = P_{T-t} f \Gamma_1^z(\log P_{T-t} f)(x).$$

**Lemma 20.** We have the following relation:

$$\frac{1}{2}L\phi_a + \frac{\partial}{\partial t}\phi_a = (P_{T-t}f)(x)\Gamma_2(\log P_{T-t}f, \log P_{T-t}f)(x), \quad (65)$$

$$\begin{aligned} \frac{1}{2}L\phi_z + \frac{\partial}{\partial t}\phi_z &= (P_{T-t}f)(x)\Gamma_2^z(\log P_{T-t}f, \log P_{T-t}f)(x) \\ &\quad + (P_{T-t}f)(x)\Gamma_1(\log P_{T-t}f, \Gamma_1^z(P_{T-t}f, P_{T-t}f))(x) \\ &\quad - (P_{T-t}f)(x)\Gamma_1^z(\log P_{T-t}f, \Gamma_1(P_{T-t}f, P_{T-t}f))(x). \end{aligned} \quad (66)$$

**Proof.** Denote  $g(t, x) = P_{T-t}f(x) = \int \rho(t, x, \tilde{x})f(\tilde{x})d\tilde{x}$ , and we have the following relation:

$$L(\log g) = -\frac{\Gamma_1(g, g)}{(g)^2} - 2\frac{\partial_t g}{g}.$$

By direct computation, one obtains

$$\begin{aligned} \partial_t \phi_a &= \partial_t g \Gamma_1(\log g, \log g) + 2g \langle a^\top \nabla \log g, a^\top \nabla \left( \frac{\partial_t g}{g} \right) \rangle_{\mathbb{R}^n} \\ &= -\frac{1}{2}Lg \Gamma_1(\log g, \log g) - g \Gamma_1(\log g, L \log g) - g \Gamma_1(\log g, \Gamma_1(\log g, \log g)), \\ \frac{1}{2}L\phi_a &= \frac{1}{2}Lg \Gamma_1(\log g, \log g) + \frac{1}{2}g L \Gamma_1(\log g, \log g) + \Gamma_1(g, \Gamma_1(\log g, \log g)), \end{aligned}$$

where we have  $\Gamma_1(g, \Gamma_1(\log g, \log g)) = g \Gamma_1(\log g, \Gamma_1(\log g, \log g))$ ; thus, (66) is proven. Similarly, we obtain the following for  $\phi_z$ :

$$\begin{aligned} \partial_t \phi_z &= \partial_t g \Gamma_1^z(\log g, \log g) + 2g \langle z^\top \nabla \log g, z^\top \nabla \left( \frac{\partial_t g}{g} \right) \rangle_{\mathbb{R}^m} \\ &= -\frac{1}{2}Lg \Gamma_1^z(\log g, \log g) - g \Gamma_1^z(\log g, L \log g) - g \Gamma_1^z(\log g, \Gamma_1(\log g, \log g)), \\ \frac{1}{2}L\phi_z &= \frac{1}{2}Lg \Gamma_1^z(\log g, \log g) + \frac{1}{2}g L \Gamma_1^z(\log g, \log g) + \Gamma_1(g, \Gamma_1^z(\log g, \log g)). \end{aligned}$$

The proof then follows.  $\square$

Now, we are ready to present the following important lemma, which prepares us to prove the new entropy inequality without the assumption:

$$\Gamma_1(\log P_{T-t}f, \Gamma_1^z(P_{T-t}f, P_{T-t}f))(x) = \Gamma_1^z(\log P_{T-t}f, \Gamma_1(P_{T-t}f, P_{T-t}f))(x).$$

**Lemma 21.** For any  $0 < s < T$ , we denote  $\rho(s, x, y) = p(s, x, y)\pi(y)$  as the transition kernel of diffusion process  $X_s^x$  starting at  $x$  defined in Definition 5, and the following equality is satisfied:

$$\begin{aligned} &\mathbb{E}[g \Gamma_1(\log g, \Gamma_1^z(\log g, \log g)) - g \Gamma_1^z(\log g, \Gamma_1(\log g, \log g))] \\ &= \int \frac{\nabla \cdot (\rho(s, x, y) z z^\top \Gamma_{\nabla(aa^\top)}(\log g(s, y), \log g(s, y)))}{\rho(s, x, y)} g(s, y) \rho(s, x, y) dy \\ &\quad - \int \frac{\nabla \cdot (\rho(s, x, y) a a^\top \Gamma_{\nabla(z z^\top)}(\log g(s, y), \log g(s, y)))}{\rho(s, x, y)} g(s, y) \rho(s, x, y) dy. \end{aligned}$$

Here, we denote  $g(s, y) = P_{T-s}f(y) = \int \rho(s, y, \tilde{y})f(\tilde{y})d\tilde{y}$  and

$$\begin{aligned} &\mathbb{E}[g \Gamma_1(\log g, \Gamma_1^z(\log g, \log g))] \\ &= \mathbb{E}[g(s, X_s) \Gamma_1(\log g(s, X_s), \Gamma_1^z(\log g(s, X_s), \log g(s, X_s)))] \\ &= \int g(s, y) \Gamma_1(\log g(s, y), \Gamma_1^z(\log g(s, y), \log g(s, y))) \rho(s, x, y) dy. \end{aligned}$$

**Proof.** We first expand in the following integral form.

$$\begin{aligned} & \mathbb{E}[g\Gamma_1(\log g, \Gamma_1^z(\log g, \log g)) - g\Gamma_1^z(\log g, \Gamma_1(\log g, \log g))] \\ &= \int g(s, y)\Gamma_1(\log g(s, y), \Gamma_1^z(\log g(s, y), \log g(s, y)))\rho(s, x, y)dy \\ & \quad - \int g(s, y)\Gamma_1^z(\log g(s, y), \Gamma_1(\log g(s, y), \log g(s, y)))\rho(s, x, y)dy. \end{aligned}$$

We skip  $x, y, s$  for simplicity. Take  $\log g = h$ .

**Claim 1:**

$$\begin{aligned} & \int \Gamma_1(h, \Gamma_1^z(h, h))\rho g dy - \int \Gamma_1^z(h, \Gamma_1(h, h))\rho g dy \\ &= \int \Gamma_1^z(h, \Delta_a h)\rho g dy - \int \Gamma_1^z(h, \frac{\Delta_a g}{g})\rho g dy - \int \Gamma_1(h, \Delta_z h)\rho g dy + \int \Gamma_1(h, \frac{\Delta_z g}{g})\rho g dy. \end{aligned}$$

Recall that we denote  $\Delta_a = \nabla \cdot (aa^T \nabla)$  and  $\Delta_z = \nabla \cdot (zz^T \nabla)$ . Use the following identity:

$$\Delta_a h = \frac{\Delta_a g}{g} - \frac{\Gamma_1(g, g)}{g^2}, \quad \text{and} \quad \Delta_z h = \frac{\Delta_z g}{g} - \frac{\Gamma_1^z(g, g)}{g^2}.$$

We then obtain

$$\begin{aligned} \int \Gamma_1^z(h, \Delta_a h)\rho g dy &= \int \Gamma_1^z\left(h, \frac{\Delta_a g}{g} - \frac{\Gamma_1(g, g)}{g^2}\right)\rho g dy \\ &= - \int \Gamma_1^z(h, \Gamma_1(h, h))\rho g dy + \int \Gamma_1^z(h, \frac{\Delta_a g}{g})\rho g dy. \end{aligned}$$

Similarly, the other equality is satisfied.

**Claim 2:**

$$\begin{aligned} & \int \Gamma_1^z(h, \Delta_a h)\rho g dy - \int \Gamma_1^z(h, \frac{\Delta_a g}{g})\rho g dy - \int \Gamma_1(h, \Delta_z h)\rho g dy + \int \Gamma_1(h, \frac{\Delta_z g}{g})\rho g dy \\ &= \int \frac{\nabla \cdot (\rho zz^T \Gamma_{\nabla(aa^T)}(h, h))}{\rho} g \rho dy - \int \frac{\nabla \cdot (\rho aa^T \Gamma_{\nabla(zz^T)}(h, h))}{\rho} g \rho dy. \end{aligned}$$

First, observe that

$$\begin{aligned} \int \Gamma_1^z(h, \frac{\Delta_a g}{g})\rho g dy &= \int \langle zz^T \nabla h, \nabla(\frac{\Delta_a g}{g}) \rangle \rho g dy = - \int \nabla \cdot (\rho zz^T \nabla g) \frac{\Delta_a g}{g} dy \\ &= - \int \frac{\rho}{g} \Delta_a g \Delta_z g dy - \int \langle \nabla \rho, zz^T \nabla g \rangle \frac{\Delta_a g}{g} dy. \end{aligned}$$

Similarly, one obtains

$$\int \Gamma_1(h, \frac{\Delta_z g}{g})\rho g dy = - \int \frac{\rho}{g} \Delta_a g \Delta_z g dy - \int \langle \nabla \rho, aa^T \nabla g \rangle \frac{\Delta_z g}{g} dy.$$

For the next term, one obtains

$$\begin{aligned}
& \int \Gamma_1^z(h, \Delta_a h) \rho g dy \\
&= \int \langle \nabla (\nabla \cdot (aa^T \nabla h)), zz^T \nabla h \rangle \rho g dy \\
&= - \int [\nabla \cdot (aa^T \nabla h)] [\nabla \cdot (\rho g zz^T \nabla h)] dy \\
&= - \int [\nabla \cdot (aa^T \frac{1}{g} \nabla g)] \nabla \cdot (\rho g zz^T \nabla g) dy \\
&= - \int \left( \langle \nabla \frac{1}{g}, aa^T \nabla g \rangle + \frac{1}{g} \Delta_a g \right) \nabla \cdot (\rho g zz^T \nabla g) dy \\
&= \int \langle \frac{1}{g^2} \nabla g, aa^T \nabla g \rangle (\nabla \cdot (\rho g zz^T \nabla g)) dy - \int \frac{1}{g} \Delta_a g (\nabla \cdot (\rho g zz^T \nabla g)) dy \\
&= -2 \int \nabla^2 h (aa^T \nabla h, zz^T \nabla h) \rho g dy - \int \langle \langle \nabla h, \nabla (aa^T) \nabla h \rangle, zz^T \nabla h \rangle \rho g dy \\
&\quad - \int \frac{1}{g} \Delta_a g \langle \nabla \rho, zz^T \nabla g \rangle dy - \int \frac{\rho}{g} \Delta_a g \Delta_z g dy,
\end{aligned}$$

where the last equality follows from the integration by parts for the first term and the direct expansion of the divergence for the second term. Similarly, we obtain

$$\begin{aligned}
& \int \Gamma_1(h, \Delta_z h) \rho g dy \\
&= -2 \int \nabla^2 h (zz^T \nabla h, aa^T \nabla h) \rho g dy - \int \langle \langle \nabla h, \nabla (zz^T) \nabla h \rangle, aa^T \nabla h \rangle \rho g dy \\
&\quad - \int \frac{1}{g} \Delta_z g \langle \nabla \rho, aa^T \nabla g \rangle dy - \int \frac{\rho}{g} \Delta_a g \Delta_z g dy.
\end{aligned}$$

Observing, by integration by parts, we obtain

$$\begin{aligned}
& - \int \langle \langle \nabla h, \nabla (aa^T) \nabla h \rangle, zz^T \nabla h \rangle \rho g dy + \int \langle \langle \nabla h, \nabla (zz^T) \nabla h \rangle, aa^T \nabla h \rangle \rho g dy \\
&= \int \frac{\nabla \cdot (\rho g zz^T \Gamma_{\nabla(aa^T)}(h, h))}{\rho} g \rho dy - \int \frac{\nabla \cdot (\rho g aa^T \Gamma_{\nabla(zz^T)}(h, h))}{\rho} g \rho dy.
\end{aligned}$$

Combining the above formulas, the proof is completed.  $\square$

With the above lemma in hand, we are ready to prove the following entropic inequality. We first define the following energy form:

$$\Phi_a(x, t) = P_t(P_{T-t} f \Gamma_1(\log P_{T-t} f))(x), \quad \Phi_z(x, t) = P_t(P_{T-t} f \Gamma_1^z(\log P_{T-t} f))(x).$$

Recall that we define

$$\phi_a(x, t) = P_{T-t} f \Gamma_1(\log P_{T-t} f)(x), \quad \text{and} \quad \phi_z(x, t) = P_{T-t} f \Gamma_1^z(\log P_{T-t} f)(x).$$

**Theorem 4.** Denote  $\phi = \phi_a + \phi_z$ ; if the following condition is satisfied:

$$\Re \succeq \kappa(\Gamma_1 + \Gamma_1^z),$$

we then conclude

$$P_T(\phi(\cdot, T))(x) \geq \phi(x, 0) + \int_0^T \kappa_s(\Phi_a(x, s) + \Phi_z(x, s)) ds, \quad (67)$$

where  $\kappa_s$  depends on the estimate of the transition kernel  $\nabla \log \rho(s, \cdot, \cdot)$  associated with semi-group  $P_s$  (see Definition 5).

**Remark 16.** Based on Theorem 3, we can also prove the above theorem for operator  $\tilde{L}$  with the drift term involved. Since the proof is similar, we skip the proof here.

**Proof.** Take  $\phi = \phi_a + \phi_z$ . Let  $(X_t^x)_{t \geq 0}$  be the diffusion Markov process with semigroup  $P_t$ . (Similar proofs can be found in [2] (Proposition 4.5).) Let smooth function  $u : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$  be such that, for every  $T > 0$ ,  $\sup_{t \in [0, T]} \|u(t, \cdot)\|_\infty < \infty$  and  $\sup_{t \in [0, T]} \|\frac{1}{2}Lu(t, \cdot) + \partial_t u(t, \cdot)\|_\infty < \infty$ . We have for every  $t > 0$

$$u(t, X_t^x) = u(0, x) + \int_0^t (\frac{1}{2}Lu + \partial_s u)(s, X_s^x) ds + M_t,$$

where  $(M_t)_{t \geq 0}$  is a local martingale. Let  $T_n, n \in \mathbb{N}$  be an increasing sequence of stopping times such that, almost surely,  $T_n \rightarrow \infty$  and  $(M_{t \wedge T_n})_{t \geq 0}$  is a martingale. We obtain

$$\mathbb{E}[u(t \wedge T_n, X_{t \wedge T_n}^x)] = u(0, x) + \mathbb{E}[\int_0^{t \wedge T_n} (\frac{1}{2}Lu + \partial_s u)(s, X_s^x) ds].$$

By using the dominated convergence theorem, we obtain

$$\mathbb{E}[u(t, X_t^x)] = u(0, x) + \mathbb{E}[\int_0^t (\frac{1}{2}Lu + \partial_s u)(s, X_s^x) ds].$$

Applying the above equality to  $\phi(t, X_t^x)$ , we obtain

$$\begin{aligned} \mathbb{E}[\phi(t, X_t^x)] &= \phi(0, x) + \mathbb{E}[\int_0^t (\frac{1}{2}L\phi + \partial_s \phi)(s, X_s^x) ds] \\ &= \phi(0, x) + \int_0^t \mathbb{E}[(\frac{1}{2}L\phi + \partial_s \phi)(s, X_s^x)] ds. \end{aligned}$$

We now look at the term  $\mathbb{E}[(\frac{1}{2}L\phi + \partial_s \phi)(s, X_s^x)]$  with  $g(s, x) = (P_{T-s}f)(x) = \mathbb{E}[f(X_T^x)] = \int \rho(x, y, s) f(y) dy$ :

$$\begin{aligned} \mathbb{E}[(\frac{1}{2}L\phi + \partial_s \phi)(s, X_s^x)] &= \mathbb{E}[g\Gamma_2(\log g, \log g) + g\Gamma_2^z(\log g, \log g)] \\ &\quad + \mathbb{E}[g\Gamma_1(\log g, \Gamma_1^z(\log g, \log g)) - g\Gamma_1^z(\log g, \Gamma_1(\log g, \log g))]. \end{aligned}$$

By using the above Lemma 21, let  $h = \log g$ , and we obtain

$$\begin{aligned} &\mathbb{E}[(\frac{1}{2}L\phi + \partial_s \phi)(s, X_s^x)] \\ &= \int g\rho \left( \Gamma_2(h, h) + \Gamma_2^z(h, h) + \frac{\nabla \cdot (\rho z z^T \Gamma_{\nabla(aa^T)}(h, h))}{\rho} - \frac{\nabla \cdot (\rho a a^T \Gamma_{\nabla(zz^T)}(h, h))}{\rho} \right) dy \\ &= \int g\rho \left( \Gamma_2(h, h) + \tilde{\Gamma}_2^{z, \rho}(h, h) \right) dy. \end{aligned}$$

Applying Theorem 3 here with  $\pi = \rho(s, \cdot, \cdot)$  as the transition kernel function, we obtain a time-dependent version of Theorem 3. Assume that the following bound is satisfied where the bound  $\kappa_s$  depends on kernel  $\rho(s, \cdot, \cdot)$ :

$$\Re(\nabla f, \nabla f) \geq \kappa_s(\Gamma_1(f, f) + \Gamma_1^z(f, f)).$$

We then conclude with the following bound:

$$\begin{aligned}\mathbb{E}\left[\left(\frac{1}{2}L\phi + \partial_s\phi\right)(s, X_s^x)\right] &\geq \int \rho(s, x, y)g\kappa_s(\Gamma_1(h, h)(y) + \Gamma_1^z(h, h)(y))dy \\ &= \int p(s, x, y)g\kappa_s(\Gamma_1(h, h)(y) + \Gamma_1^z(h, h)(y))\pi(y)dy \\ &\geq P_s(\kappa_s g(\Gamma_1(\log g, \log g) + \Gamma_1^z(\log g, \log g))).\end{aligned}$$

Plugging into the time integral  $\int_0^T \mathbb{E}[(\frac{1}{2}L\phi + \partial_s\phi)(s, X_s^x)]ds$ , the proof follows.  $\square$

**Remark 17.** We prove the entropic inequality Theorem 4 in this section without the the assumption:  $\Gamma_1(f, \Gamma_1^z(f, f)) = \Gamma_1^z(f, \Gamma_1(f, f))$ . A similar entropic inequality under the assumption  $\Gamma_1(f, \Gamma_1^z(f, f)) = \Gamma_1^z(f, \Gamma_1(f, f))$  was first proven in [2] (Proposition 4.5 and Theorem 5.2). With this new inequality Theorem 4 in hand, similar gradient estimates and other inequalities from [2] follow. We leave them for future studies. Proposition 4.5 in [2] is based on a pointwise estimate given the commutative assumption of  $\Gamma_1$  and  $\Gamma_1^z$ . We removed the commutative assumption, and our estimate is in a weak form, which is presented in the above Lemma 21.

**Author Contributions:** Conceptualization, Q.F. and W.L.; methodology, Q.F. and W.L.; writing—original draft preparation, Q.F. and W.L.; writing—review and editing, Q.F. and W.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** Wuchen Li is supported by AFOSR MURI FA9550-18-1-0502, the AFOSR YIP award: FA9550-23-1-0087, and NSF RTG: 2038080.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Degenerate SDEs and Sub-Riemannian Manifold

In this appendix, we briefly illustrate the formulation of the degenerate diffusion process and sub-Riemannian geometry.

For a smooth connected  $n + m$ -dimensional Riemannian manifold  $\mathbb{M}^{n+m}$ , we denote  $T\mathbb{M}^{n+m}$  as the tangent bundle of  $\mathbb{M}^{n+m}$  and denote  $\tau$  as a sub-bundle of  $T\mathbb{M}^{n+m}$ . The sub-Riemannian structure associated with the sub-bundle  $\tau$  on  $\mathbb{M}^{n+m}$  is denoted as  $(\tau, g_\tau)$ , where  $g_\tau(\cdot, \cdot)$  is the metric associated with the sub-bundle  $\tau$ . In particular, if we take distribution  $\tau$  to be the horizontal sub-bundle, denoted as  $\mathcal{H}$ , of the tangent bundle  $T\mathbb{M}^{n+m}$  (see [2,51] for more details), then we denote the sub-Riemannian structure as  $(\mathbb{M}^{n+m}, \mathcal{H}, g_\mathcal{H})$ . In this paper, we will not distinguish distributions  $\tau$  and  $\mathcal{H}$  and call this the horizontal sub-bundle. We assumed that the horizontal distribution  $\mathcal{H}$  is bracket-generating (with any steps). The distribution  $\mathcal{H}$  has dimension  $n$ .

For a vector field  $b \in \mathbb{R}^{n+m}$  and a general matrix  $a \in \mathbb{R}^{(n+m) \times n}$ , we denote  $a = (a_1, a_2, \dots, a_n)$  with each  $a_i, i = 1 \dots, n$ , as an  $n + m$ -dimensional column vector. For any Stratonovich SDE,

$$dX_t = b(X_t)dt + \sqrt{2} \sum_{i=1}^n a_i(X_t) \circ dB_t^i, \quad (\text{A1})$$

where  $(B_t^1, B_t^2, \dots, B_t^n)$  is an  $n$ -dimensional Brownian motion in  $\mathbb{R}^n$  and  $a_i$  has local coordinates  $a_i(x) = \sum_{i=1}^{n+m} a_{ii}(x) \frac{\partial}{\partial x_i}$ . We consider (A1) as the SDE associated with a given sub-Riemannian structure, which is defined through the Lie algebra spanned by the driving vector fields of the SDE  $\{a_1, a_2, \dots, a_n\}$ . In general, we assumed that  $\mathcal{H} := \{a_1, a_2, \dots, a_n\}$  is of rank  $n$  and satisfies the bracket-generating condition (or Hörmander condition). To be precise, for any  $x \in \mathbb{M}^{n+m}$ , the Lie brackets of  $\{a_1(x), a_2(x), \dots, a_n(x)\}$  span the whole tangent space at  $x$  with dimension  $n + m$ . We define the manifold  $\mathbb{M}^{n+m}$  as the subspace of  $\mathbb{R}^{n+m}$ , where the diffusion process  $X_t$  lives on. This spaces is described as the triple  $(\mathbb{M}^{n+m}, \mathcal{H}, g_\mathcal{H})$ , and we denote  $\mathcal{H}$  as the  $n$ -dimensional horizontal distribution of the tangent bundle  $T\mathbb{M}^{n+m}$  generated by the vector fields  $\{(a_1(x), a_2(x), \dots, a_n(x))\}$ .

In this paper, we considered the case where the generator of the diffusion process (A1) coincides with the horizontal Laplacian operator (or sub-Laplacian operator) associated with the sub-Riemannian structure  $(\mathbb{M}^{n+m}, \mathcal{H}, g_{\mathcal{H}})$ . Furthermore, we assumed that there exists a symmetric and invariant volume measure associated with the horizontal Laplacian operator. The Stratonovich SDE (A1) without the drift ( $dt$ ) term could be treated as a special case, where the horizontal Laplacian can be presented as the sum of squares of the horizontal vector fields in  $\mathcal{H}$ . In particular, we considered the precise metric defined through the diffusion matrix  $a$ , which could be seen as an analogue for non-degenerate SDEs on Riemannian manifolds. The problem is that the rank of  $aa^T$  is  $n$ ; thus, the  $(n+m) \times (n+m)$  matrix  $aa^T$  is degenerate and cannot serve as a metric. We thus introduce the following metric, which is to formulate this sub-Riemannian structure in Euclidean space.

**Definition A1.** Consider an orthonormal basis  $c = \{c_{n+1}(x), \dots, c_{n+m}(x)\}$  in  $\mathbb{R}^{n+m}$ , such that  $a_i^T c_j = 0$  for any  $1 \leq i \leq n, n+1 \leq j \leq n+m$ . We define a metric  $g = (aa^T + cc^T)^{-1} = (aa^T)^\dagger + cc^T$  and a metric on the horizontal sub-bundle  $g_\tau = (aa^T)^\dagger$ , the pseudo-inverse of matrix  $aa^T$ , on manifold  $\mathbb{M}^{n+m}$ .

The above definition is based on the following lemma.

**Lemma A1.** The metric is  $g = (aa^T + cc^T)^{-1} = (aa^T)^\dagger + cc^T$ .

**Proof.** For rank  $n$  matrix  $aa^T$ , we denote its eigenvalue decomposition and the corresponding pseudo-inverse  $(aa^T)^\dagger$  as

$$aa^T = \sum_{i=1}^n \lambda_i V_i V_i^T, \quad (aa^T)^\dagger = \sum_{i=1}^n \frac{1}{\lambda_i} V_i V_i^T.$$

Thus, we have  $aa^T + cc^T = \sum_{i=1}^n \lambda_i V_i V_i^T + \sum_{j=n+1}^{n+m} c_j c_j^T$ . Furthermore, we have

$$\begin{aligned} & aa^T + cc^T \\ &= (V_1, \dots, V_n, c_{n+1}, \dots, c_{n+m}) \begin{pmatrix} \Lambda_n & \\ & \mathbf{I}_m \end{pmatrix} (V_1, \dots, V_n, c_{n+1}, \dots, c_{n+m})^T, \\ & (aa^T + cc^T)^{-1} \\ &= (V_1, \dots, V_n, c_{n+1}, \dots, c_{n+m}) \begin{pmatrix} \Lambda_n^{-1} & \\ & \mathbf{I}_m \end{pmatrix} (V_1, \dots, V_n, c_{n+1}, \dots, c_{n+m})^T, \end{aligned}$$

where we denote  $\Lambda_n = \text{diag}(\lambda_1, \dots, \lambda_n)$  as the diagonal matrix for eigenvalues  $\lambda_i$ 's and  $\mathbf{I}_m$  as the  $m$ -dimensional identity matrix. Thus, the proof follows directly with

$$(aa^T + cc^T)^{-1} = \sum_{i=1}^n \frac{1}{\lambda_i} V_i V_i^T + \sum_{j=n+1}^{n+m} c_j c_j^T = (aa^T)^\dagger + cc^T.$$

□

With the new metric introduced above, we have the following lemma.

**Lemma A2.** The vectors  $\{a_1, \dots, a_n\}$  are the orthonormal basis under the metric  $g = (aa^T)^\dagger + cc^T$ .

**Proof.** We just need to prove for  $a = (a_1, \dots, a_n)$  with each  $(a_i)_{(n+m) \times 1}$ , and we have

$$a^T g a = a^T (aa^T)^\dagger a = \mathbf{Id}_{n \times n}.$$

Notice that  $a_i^T c_j = 0$ , then we only need to prove  $a^T (aa^T)^+ a = \mathbf{Id}_{n \times n}$ . Let us denote  $a^T (aa^T)^+ a = B$ , then we have

$$\begin{aligned} aa^T (aa^T)^+ aa^T &= aa^T \\ aa^T &= aa^T \\ a^T aa^T a &= a^T a Ba^T a \\ (a^T a)^{-1} a^T aa^T a (a^T a)^{-1} &= B \\ \mathbf{Id}_{n \times n} &= B, \end{aligned}$$

where the second equality follows from the property of the pseudo-inverse matrix and the last step follows from the fact that  $a^T a$  is a non-degenerate  $n \times n$  matrix, hence invertible. The proof then follows directly.  $\square$

We are now ready to introduce the following definition.

**Definition A2.** Define  $(\mathbb{M}^{n+m}, \tau, g_\tau)$  as the sub-Riemannian structure associated with the degenerate SDE (A1), where  $g_\tau = (aa^T)^+$  denotes the horizontal metric, i.e., metric  $g$  is restricted on the horizontal bundle  $\tau$ . We denote  $\nabla^R$  as the Levi-Civita connection on  $\mathbb{M}^{n+m}$  associated with our metric  $g = (aa^T)^+ + cc^T$ , and let  $P^\tau \nabla^R$  be the projection of the connection on the horizontal distribution  $\tau$ . In particular, in our framework, we have  $P^\tau \nabla^R f = aa^T \nabla f$ , for any function  $f : \mathbb{M}^{n+m} \rightarrow \mathbb{R}$ , where  $\nabla$  is the Euclidean gradient in  $\mathbb{R}^{n+m}$ .

**Remark A1.** In Lemma A2, we show that  $\{a_1, a_2, \dots, a_n\}$  are the orthonormal basis for horizontal distribution  $\tau$  under our metric  $g$ . In particular, we have

$$aa^T \nabla f = (a_1, \dots, a_n) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} f = \sum_{i=1}^n (a_i f) a_i \in \tau,$$

which gives the local representation of  $P^\tau \nabla^R f$ .

To demonstrate the definition clearly, we give the following example. On the Heisenberg group  $\mathbb{H}^1$ , we know that  $X = \frac{\partial}{\partial x_1} - \frac{1}{2}x_2 \frac{\partial}{\partial x_3}$ ,  $Y = \frac{\partial}{\partial x_2} + \frac{1}{2}x_1 \frac{\partial}{\partial x_3}$ ,  $Z = \frac{\partial}{\partial x_3}$  forms an orthonormal basis for the tangent bundle of  $\mathbb{H}^1$ . In particular,  $X$  and  $Y$  generate the horizontal distribution  $\tau$ . If we start with the following SDE:

$$dW_t = X \circ dB_t^1 + Y \circ dB_t^2, \quad (\text{A2})$$

then we know  $W_t = (B_t^1, B_t^2, \frac{1}{2} \int_0^t B_s^1 dB_s^2 - B_s^2 dB_s^1)$ , which is the horizontal Brownian motion on the Heisenberg group  $\mathbb{H}^1$ . The generator of the horizontal Brownian motion and the sub-Laplacian operator are the same, which is given by  $\Delta_{\mathcal{H}} = X^2 + Y^2$ , and the volume measure associated with  $\Delta_{\mathcal{H}}$  is the Lebesgue measure on the Heisenberg group with the volume element equal to 1. Then,  $W_t$  is a diffusion process in  $\mathbb{R}^3$ . In terms of our general sub-Riemannian structure introduced above, we can define

$$a = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{x_2}{2} & \frac{x_1}{2} \end{pmatrix} = (a_1, a_2) = (X, Y), \quad c = \frac{1}{\sqrt{\frac{x_1^2}{4} + \frac{x_2^2}{4} + 1}} \begin{pmatrix} \frac{x_2}{2} \\ -\frac{x_1}{2} \\ 1 \end{pmatrix},$$

and

$$g_{\mathbb{H}^1, \tau} = (aa^T)^+ = \begin{pmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \\ -\frac{x_2}{2} & \frac{x_1}{2} & \frac{x_1^2 + x_2^2}{4} \end{pmatrix}^+, \quad g_{\mathbb{H}^1} = (aa^T)^+ + cc^T.$$



In particular, the horizontal gradient is given by

$$aa^T \nabla f = \begin{pmatrix} Xf \\ Yf \\ -\frac{x_2}{2}Xf + \frac{x_1}{2}Yf \end{pmatrix} = Xf \begin{pmatrix} 1 \\ 0 \\ -\frac{x_2}{2} \end{pmatrix} + Yf \begin{pmatrix} 0 \\ 1 \\ \frac{x_1}{2} \end{pmatrix} = (Xf)X + (Yf)Y.$$

Thus, the sub-Riemannian structure associated with Stratonovich SDE (A2) is just  $(\mathbb{H}^1, \tau, g_{\mathbb{H}^1, \tau})$ , where  $g_{\mathbb{H}^1, \tau}$  is the restriction of metric  $g_{\mathbb{H}^1}$  on the horizontal sub-bundle  $\tau$ . Different from the standard construction of Brownian motion on a given Riemannian (sub-Riemannian) manifold by Ells–Elworthy–Malliavin [40,43], we can directly define our diffusion on the manifold  $\mathbb{M}^{n+m}$  by (A1) without performing projection from the orthonormal frame bundles. This is because the new metrics  $g = (aa^T)^\dagger + cc^T$  and  $\{a_1, a_2, \dots, a_n\}$  are globally defined orthonormal basis of the (horizontal) sub-bundle on the tangent bundle  $T\mathbb{M}^{n+m}$ . Essentially, we first define (A1) in  $\mathbb{R}^{n+m}$  and then introduce the associated sub-Riemannian structure.

**Remark A2.** Compared to the definition of the horizontal Brownian motion introduced in [38], the sub-Riemannian structure comes first with a totally geodesic Riemannian foliation structure, and then, SDE (A2) is defined on the given totally geodesic Riemannian foliation. In the current setting, we directly define the degenerate diffusion process by a first given matrix  $a$ , then we define the sub-Riemannian structure by introducing the new metric  $(aa^T)^\dagger + cc^T$ .

*Proof of Gradient Flow Assumption*

In this subsection, we demonstrate that Equation (6) is in fact a Fokker–Planck equation of SDE (A1).

**Lemma A3.** Consider the drift–diffusion process:

$$dX_t = b(X_t)dt + \sqrt{2}a(X_t) \circ dB_t, \quad (\text{A3})$$

Suppose that  $b, a, \pi$  satisfy

$$a \otimes \nabla a - b = -aa^T \nabla \log \pi.$$

Then, the Fokker–Planck equation of  $X_t$  satisfies

$$\partial_t \rho(t, x) = \nabla \cdot \left( \rho(t, x) \left( a(x)a(x)^T \right) \nabla \log \frac{\rho(t, x)}{\pi(x)} \right).$$

**Proof.** Recall that we denote  $\{a_1, \dots, a_n\}$  as the column vectors of matrix  $a$ . For Stratonovich SDE (A3), we can write

$$dX_t = b(X_t)dt + \sqrt{2} \sum_{i=1}^n a_i(X_t) \circ dB_t^i.$$

According to [28] (Appendix 7), the corresponding Itô SDE is

$$dX_t = \sqrt{2} \sum_{i=1}^n a_i dB_t^i + \left( \sum_{i=1}^n \nabla_{a_i} a_i + b \right) dt.$$

Thus, the Fokker-Plank equation (Kolmogorov forward equation) satisfies

$$\begin{aligned}\partial_t \rho(t, x) &= \sum_{i'=1}^{n+m} \sum_{j'=1}^{n+m} \frac{\partial^2}{\partial x_{i'} \partial x_{j'}} ((aa^T)_{i'j'} \rho) - \nabla \cdot \left( \left( \sum_{i=1}^n \nabla_{a_i} a_i + b \right) \rho \right) \\ &= \nabla \cdot (aa^T \nabla \rho) + \nabla \cdot \left( \rho \left( \sum_{j'=1}^{n+m} \frac{\partial}{\partial x_{j'}} (aa^T)_{i'j'} \right)_{i'=1}^{n+m} - \rho \sum_{i=1}^n \nabla_{a_i} a_i - b \rho \right) \\ &= \nabla \cdot (aa^T \nabla \rho) + \nabla \cdot (\rho(a \otimes \nabla a - b)).\end{aligned}$$

Namely, we have

$$\partial_t \rho(t, x) = \nabla \cdot (aa^T \nabla \rho) + \nabla \cdot (\rho(a \otimes \nabla a - b)). \quad (\text{A4})$$

Plugging in the relation  $a \otimes \nabla a - b = -aa^T \nabla \log \pi$ , we have

$$\begin{aligned}\partial_t \rho(t, x) &= \nabla \cdot (aa^T \nabla \rho) - \nabla \cdot (\rho aa^T \nabla \log \pi) \\ &= \nabla \cdot (\rho aa^T \nabla \log \rho) - \nabla \cdot (\rho aa^T \nabla \log \pi) \\ &= \nabla \cdot (\rho aa^T \nabla \log \frac{\rho}{\pi}).\end{aligned} \quad (\text{A5})$$

Here, we use the fact that

$$\rho \nabla \log \rho = \nabla \rho.$$

This finishes the proof.  $\square$

**Example A1.** The Lie group  $\mathbf{SU}(2)$  is a compact connected Lie group, diffeomorphic to the three-sphere  $\mathbb{S}^3$ . Following the construction of the left-invariant vector fields in [41] (Section 6.2), we change the coordinates in terms of coordinate system  $(\theta, \phi, \psi)$ . We obtain new left-invariant vector fields on  $\mathbf{SU}(2)$ , with

$$\begin{aligned}X &= \cos \psi \frac{\partial}{\partial \theta} + \frac{\sin \psi}{\sin \theta} \frac{\partial}{\partial \phi} - \cos \theta \frac{\sin \psi}{\sin \theta} \frac{\partial}{\partial \psi}, \\ Y &= -\sin \psi \frac{\partial}{\partial \theta} + \frac{\cos \psi}{\sin \theta} \frac{\partial}{\partial \phi} - \cos \theta \frac{\cos \psi}{\sin \theta} \frac{\partial}{\partial \psi}, \\ Z &= \frac{\partial}{\partial \psi}.\end{aligned}$$

Thus, we have  $a = (a_1, a_2) = (X, Y)$  in the new coordinate system. We define the metric  $g = (aa^T)^\dagger$ . Here,  $X, Y$  are the orthonormal basis for the horizontal bundle generated by  $X, Y$  under metric  $(aa^T)^\dagger$ . According to [41] (Lemma 6.4), the invariant measure on  $\mathbf{SU}(2)$  has the form of  $\mu = \sin(\theta) d\theta \wedge d\phi \wedge d\psi$ . It is easy to check that the above Lemma is satisfied for  $b = 0$ ,  $\pi = \sin(\theta)$ , and

$$aa^T \nabla \log \pi = -a \otimes \nabla a = \begin{pmatrix} \frac{\cos \theta}{\sin \theta} \\ 0 \\ 0 \end{pmatrix},$$

where

$$a = \begin{pmatrix} \cos \psi & -\sin \psi \\ \frac{\sin \psi}{\sin \theta} & \frac{\cos \psi}{\sin \theta} \\ -\cos \theta \frac{\sin \psi}{\sin \theta} & -\cos \theta \frac{\cos \psi}{\sin \theta} \end{pmatrix}.$$

## References

- Bakry, D.; Émery, M. Diffusions hypercontractives. In *Séminaire de Probabilités XIX 1983/84*; Springer: Berlin/Heidelberg, Germany, 1985; pp. 177–206.
- Baudoin, F.; Garofalo, N. Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries. *J. EMS* **2017**, *19*, 151–219. [\[CrossRef\]](#)
- Arnold, A.; Carlen, E. A generalized Bakry–Émery condition for non-symmetric diffusions. In Proceedings of the EQUADIFF 99—International Conference on Differential Equations, Berlin, Germany, 1–7 August 1999; pp. 732–734.
- Li, W. Transport information geometry: Riemannian calculus on probability simplex. *Inf. Geom.* **2022**, *5*, 161–207. [\[CrossRef\]](#)
- Otto, F. The geometry of dissipative evolution equations the porous medium equation. *Commun. Partial Differ. Equ.* **2001**, *26*, 101–174. [\[CrossRef\]](#)
- Otto, F.; Villani, C. Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality. *J. Funct. Anal.* **2000**, *173*, 361–400. [\[CrossRef\]](#)
- Baudoin, F. Wasserstein contraction properties for hypoelliptic diffusions. *arXiv* **2016**, arXiv:1602.04177.
- Baudoin, F. Bakry–Émery meet Villani. *J. Funct. Anal.* **2017**, *273*, 2275–2291. [\[CrossRef\]](#)
- Baudoin, F.; Bonnefont, M.; Garofalo, N. A sub-Riemannian curvature-dimension inequality, volume doubling property and the Poincaré inequality. *Math. Ann.* **2014**, *358*, 833–860. [\[CrossRef\]](#)
- Baudoin, F.; Gordina, M.; Herzog, D.P. Gamma calculus beyond Villani and explicit convergence estimates for Langevin dynamics with singular potentials. *Arch. Ration. Mech. Anal.* **2021**, *241*, 765–804. [\[CrossRef\]](#)
- Baudoin, F.; Grong, E.; Kuwada, K.; Thalmaier, A. Sub-Laplacian comparison theorems on totally geodesic Riemannian foliations. *Calc. Var.* **2019**, *58*, 130. [\[CrossRef\]](#)
- Baudoin, F.; Wang, J. Curvature dimension inequalities and subelliptic heat kernel gradient bounds on contact manifolds. *Potential Anal.* **2014**, *40*, 163–193. [\[CrossRef\]](#)
- Feng, Q. Harnack inequalities on totally geodesic foliations with transverse Ricci flow. *arXiv* **2017**, arXiv:1712.02275.
- Grong, E.; Thalmaier, A. Curvature-dimension inequalities on sub-Riemannian manifolds obtained from Riemannian foliations: Part I. *Math. Z.* **2015**, *282*, 99–130. [\[CrossRef\]](#)
- Grong, E.; Thalmaier, A. Curvature-dimension inequalities on sub-Riemannian manifolds obtained from Riemannian foliations: Part II. *Math. Z.* **2015**, *282*, 131–164. [\[CrossRef\]](#)
- Agrachev, A.; Lee, P. Optimal transportation under nonholonomic constraints. *Trans. Am. Math. Soc.* **2009**, *361*, 6019–6047. [\[CrossRef\]](#)
- Figalli, A.; Rifford, L. Mass transportation on sub-Riemannian manifolds. *Geom. Funct. Anal.* **2010**, *20*, 124–159. [\[CrossRef\]](#)
- Juillet, N. Diffusion by optimal transport in Heisenberg groups. *Calc. Var. Partial Differ. Equ.* **2014**, *50*, 693–721. [\[CrossRef\]](#)
- Khesin, B.; Lee, P. A nonholonomic Moser theorem and optimal transport. *J. Symplectic Geom.* **2009**, *7*, 381–414. [\[CrossRef\]](#)
- Lott, J.; Villani, C. Ricci Curvature for Metric-Measure Spaces via Optimal Transport. *Ann. Math.* **2009**, *169*, 903–991. [\[CrossRef\]](#)
- Sturm, K.-T. On the Geometry of Metric Measure Spaces. *Acta Math.* **2006**, *196*, 65–131. [\[CrossRef\]](#)
- Lafferty, J.D. The Density Manifold and Configuration Space Quantization. *Trans. Am. Math. Soc.* **1988**, *305*, 699–741. [\[CrossRef\]](#)
- Jüngel, A. *Entropy Methods for Diffusive Partial Differential Equations*; Springer: Berlin/Heidelberg, Germany, 2016.
- Markowich, P.A.; Villani, C. On the Trend to Equilibrium for the Fokker–Planck Equation: An Interplay between Physics and Functional Analysis. *Physics and Functional Analysis. Mat. Contemp.* **1999**, *19*, 1–29.
- Arnold, A.; Einav, A.; Wöhrer, T. On the rates of decay to equilibrium in degenerate and defective Fokker–Planck equations. *J. Differ. Equ.* **2018**, *264*, 6843–6872. [\[CrossRef\]](#)
- Arnold, A.; Erb, J. Sharp entropy decay for hypocoercive and non-symmetric Fokker–Planck equations with linear drift. *arXiv* **2014**, arXiv:1409.5425.
- Karatzas, I.; Shreve, S.E. Brownian Motion and Stochastic Calculus, 2nd ed.; *Graduate Texts in Mathematics*; Springer: New York, NY, USA, 1991; Volume 113.
- Baudoin, F. *An Introduction to the Geometry of Stochastic Flows*; World Scientific: Singapore, 2004.
- Stroock, D.W. Partial differential equations for probabilists. In *Cambridge Studies in Advanced Mathematics*; Cambridge University Press: Cambridge, UK, 2008; Volume 112.
- Bismut, J.M. Martingales, the Malliavin calculus and hypoellipticity under general Hörmander’s conditions. In *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*; Springer: Berlin/Heidelberg, Germany, 1981; Volume 56, pp. 469–505.
- Hörmander, L. Hypoelliptic second-order differential equations. *Acta Math.* **1967**, *119*, 147–171. [\[CrossRef\]](#)
- Arous, B.; Léandre, R. Décroissance exponentielle du noyau de la chaleur sur la diagonale (II). *Probab. Theory Relat. Fields* **1991**, *90*, 377–402. [\[CrossRef\]](#)
- Barlow, M.; Nualart, D. Lectures on Probability Theory and Statistics. In *Ecole d’Ete de Probabilites de Saint-Flour XXV*; Springer: Berlin/Heidelberg, Germany, 1995.
- Baudoin, F.; Nualart, E.; Ouyang, C.; Tindel, S. On probability laws of solutions to differential systems driven by a fractional Brownian motion. *Ann. Probab.* **2016**, *44*, 2554–2590. [\[CrossRef\]](#)
- Feng, Q.; Li, W. Hypoelliptic entropy dissipation for stochastic differential equations. *arXiv* **2021**, arXiv:2102.00544.
- Agrachev, A.; Barilari, D.; Boscain, U. On the Hausdorff volume in sub-Riemannian geometry. *Calc. Var. Partial Differ. Equ.* **2012**, *43*, 355–388. [\[CrossRef\]](#)

37. Barilari, D.; Rizzi, L. A formula for Popp’s volume in sub-Riemannian geometry. *Anal. Geom. Metr. Spaces* **2013**, *1*, 42–57. [[CrossRef](#)]
38. Baudoin, F.; Feng, Q.; Gordina, M. Integration by parts and quasi-invariance for the horizontal Wiener measure on foliated compact manifolds. *J. Funct. Anal.* **2019**, *277*, 1362–1422. [[CrossRef](#)]
39. Eldredge, N.; Gordina, M.; Saloff-Coste, L. Left-invariant geometries on  $SU(2)$  are uniformly doubling. *Geom. Funct. Anal.* **2018**, *28*, 1321–1367. [[CrossRef](#)]
40. Elworthy, K.D. *Stochastic Differential Equations on Manifolds*; Cambridge University Press: Cambridge, UK, 1982; Volume 70.
41. Gordina, M.; Laetsch, T. Sub-Laplacians on sub-Riemannian manifolds. *Potential Anal.* **2016**, *44*, 811–837. [[CrossRef](#)]
42. Gordina, M.; Laetsch, T. A convergence to Brownian motion on sub-Riemannian manifolds. *Trans. Am. Math. Soc.* **2017**, *369*, 6263–6278. [[CrossRef](#)]
43. Malliavin, P. Stochastic Analysis. In *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*; Springer: Berlin/Heidelberg, Germany, 1997; Volume 313.
44. Baudoin, F.; Feng, Q. Log-Sobolev inequalities on the horizontal path space of a totally geodesic foliation. *arXiv* **2015**, arXiv:1503.08180.
45. Inglis, J.; Papageorgiou, I. Logarithmic Sobolev inequalities for infinite-dimensional Hörmander type generators on the Heisenberg group. *Potential Anal.* **2009**, *31*, 79–102. [[CrossRef](#)]
46. Baudoin, F.; Bonnefont, M. Log-Sobolev inequalities for subelliptic operators satisfying a generalized curvature dimension inequality. *J. Funct. Anal.* **2012**, *262*, 2646–2676. [[CrossRef](#)]
47. Wang, F.-Y. Logarithmic Sobolev inequalities on noncompact Riemannian manifolds. *Probab. Theory Relat. Fields* **1997**, *109*, 417–424. [[CrossRef](#)]
48. Woit, P. *Quantum Theory, Groups and Representations*; Springer: Berlin/Heidelberg, Germany, 2017.
49. Baudoin, F.; Cecil, M. The subelliptic heat kernel on the three-dimensional solvable Lie groups. *Forum Math.* **2015**, *27*, 2051–2086. [[CrossRef](#)]
50. Li, W. Diffusion Hypercontractivity via Generalized Density Manifold. *arXiv* **2019**, arXiv:1907.12546.
51. Baudoin, F. Sub-Laplacians and hypoelliptic operators on totally geodesic Riemannian foliations. *arXiv* **2014**, arXiv:1410.3268.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.