Generalized Gamma z calculus via sub-Riemannian density manifold

Qi Feng* and Wuchen Li†

October 29, 2019

Abstract

We generalize the Gamma z calculus to study degenerate drift-diffusion processes, where z stands for the extra directions introduced into the degenerate system. Based on the new calculus, we establish the curvature dimension bound for general sub-Riemannian manifolds, which does not require the commutative iteration of Gamma and Gamma z operator and goes beyond the step two condition. It allows us to analyze the convergence properties of degenerate drift-diffusion processes and prove the entropy dissipation rate and several functional inequalities in sub-Riemannian manifolds. Several examples are provided. The new Gamma z calculus is motivated by optimal transport and density manifold. We embed the probability density space over sub-Riemannian manifold with the L^2 sub-Riemannian Wasserstein metric. We call it sub-Riemannian density manifold (SDM). We study the dynamical behavior of the degenerate Fokker-Planck equation as gradient flows in SDM. Our derivation builds an equivalence relation between Gamma z calculus and second-order calculus in SDM.

Keywords: Degenerate drift-diffusion process; Generalized Gamma z calculus; Generalized curvature dimension bound; sub-Riemannian density manifold; Optimal transport.

Thanks: Wuchen Li is supported by AFOSR MURI FA9550-18-1-0502.

*Department of Mathematics, University of Southern California, Los Angeles, 90089; email: qif@usc.edu
†Department of Mathematics, University of California Los Angeles, 90089; email: wcli@math.ucla.edu.

Contents

1 Introduction

[10, 8]Grushin plane interpolation inequality and Heisenberg group inequality.Chapter 30 by Villani Gamma calculus and curvature dimension bound play essential roles in differential geometry, functional inequalities, and the estimation of the convergence rate of drift-diffusion processes. In classical Riemannian geometry, the Gamma calculus is also named Bakry-Émery iterative calculus [6], which provides analytical ways to derive Ricci curvature tensor and its lower bound. This lower bound provides sufficient conditions for the convergence rate of related drift-diffusion processes. However, classical studies are limited to the setting of Riemannian metric. Here, the metric tensor needs to be positive definite. In sub-Riemannian geometry, this condition is no longer valid, where the metric tensor on the sub-bundle is degenerate, see related studies in [16, 14, 19, 13, 18, 17, 25, 31, 3, 23, 32] and many references therein. Because of this degeneration, many classical concepts, such as volume element, Ricci curvature tensor, and curvature dimension bound, are missed in sub-Riemannian settings.

In this paper, extending ideas in Gamma z calculus [16, 13, 12], we present a generalized Bakery-Émery iterative calculus and curvature dimension bound in a sub-Riemannian setting, and connect them to drift diffusion processes with degenerate diffusion coefficients. In particular, we start with the following Stratonovich stochastic differential equation (SDE),

$$dX_t = b(X_t)dt + \sum_{i=1}^{n} a_i(X_t) \circ dB_t^i, \tag{1.1}$$

where $(B_t^1, B_t^2, \cdots, B_t^n)$ is a n-dimensional Brownian motion in $\mathbb{R}^n$, $a = (a_1, a_2, \cdots, a_n)$ is a $(n+m) \times n$ matrix and b is a drift vector field. The sub-Riemannian structure (see definition 2.5) comes from introducing a new metric in $\mathbb{R}^{n+m}$ and viewing $\{a_1, \cdots, a_n\}$ as basis for the horizontal distribution of the new metric. We refer details to section 2.1 and section 2.2. Under this framework, we also provide a method to establish the sub-Riemannian volume. We provide sufficient conditions for generalized curvature dimension inequality beyond step two conditions. Also, our method can be applied to general drift diffusion processes, whose diffusion coefficients are degenerate and satisfy the bracket generating condition. In the end, we prove the sub-Riemannian gradient estimates, entropic inequality, Log-Sobolev inequality and Poincaré inequality under the generalized curvature dimension condition in different situations.

Our approach is motivated by optimal transport geometry, in particular the general second order calculus of relative entropy in density manifold studied in [34, 35]. Here, optimal transport provides the other viewpoint on the curvature dimensional bound [38]. In classical Riemannian cases, the probability density space is embedded with an infinite-dimensional geometry structure, known as L^2-Wasserstein metric tensor. Here, the density space is often named Riemannian density manifold [33]. See its related Riemannian calculus [34]. The Bakry-Émery calculus is associated with the Hessian operator of relative entropy in density manifold [36, 38, 40]. Here we extend the optimal transport metric to incorporate the general sub-Riemannian settings. Given a finite-dimensional sub-Riemannian manifold, we introduce a L^2 sub-Riemannian Wasserstein metric tensor in density space. We call the density space with this metric tensor sub-Riemannian density manifold (SDM). In SDM, we show that the Fokker-Planck equation of sub-Riemannian Brownian motion is a sub-Riemannian gradient flow of relative entropy. From this fact, the Hessian of relative entropy in SDM connects with the generalized Bakery-Émery calculus. These facts are extensions of those in Riemannian settings. In particular, if we add a new direction z to the classical Gamma calculus, the relation between relative entropy and Gamma calculus no longer follows a simple Hessian operator. Here we perform a general second-order geometric calculation in SDM, based on which we discover a generalized Gamma z calculus.

In literature, the first generalized curvature dimension bound is introduced by Baudoin-Garofalo [16] for sub-Riemannian manifolds and related results are studied later in [27, 28]. To the best of our knowledge, the commutative property of iteration of Γ_1 and Γ_1^z (Hypothesis 1.2 in [16]) is crucial in the aforementioned works. Our algebraic condition Assumption 1.6, 1.7, and 1.8 do not have this requirement. We are able to remove this commutative condition in the weak sense, thus our results go beyond the step two bracket generating condition. Following formulas of Bakry-Émery calculus explicitly, we give algebraic conditions for the existence of Ricci curvature tensor and its lower bound. See more details of this comparison in Remark 3.10. By routine computations, we can see that Assumption 1.6 is not satisfied for Heisenberg group, $\mathbf{SU}(2)$, $\mathbf{SL}(2)$, Eagle group, and Martinet flat sub-Riemannian structure [2]. Here we only use a horizontal distribution of the tangent bundle. However, the Assumption 1.7 and 1.8 can be valid for the above examples and other sub-Riemannian manifolds after introducing an extra direction z. This direction z is added to the horizontal distribution, which introduces the Gamma z operator approaches based on the vertical direction in [16, 12, 13, 17]. Considering the length of this paper, we leave this part for further studies. In summary, our results are divided into three cases. 1). In Theorem

3.4, we provide algebraic condition Assumption 1.6 for extending classical Bakry-Émery calculus to sub-Riemannian settings. If condition 1.6 is satisfied, we establish a sub-Riemannian Ricci curvature tensor in Theorem 3.7. 2). If condition 1.6 is not satisfied, we introduce a new direction z for Bakry-Émery calculus (named Gamma z calculus). If the term in (1.3) is zero, we first define the standard Gamma z calculus following the idea in [16], and provide a generalized curvature dimension bound in Theorem 3.25 under Assumption 1.7. 3). For the most general case with (1.3)$\neq 0$, we introduce our new generalized Gamma z calculus and prove a generalized curvature dimension inequality in Theorem 3.32 under Assumption 1.8. For all the three results mentioned above, we extend them to work in general drift-diffusion process as in Theorem 3.16, Theorem 3.26 and Theorem 3.35 respectively.

Secondly, in sub-Riemannian geometry, the sub-Riemannian volume is another important concept which does not have a canonical definition. Various extensions of sub-Riemannian volume have been considered in [1, 9, 21, 25, 26]. We provide a new definition following the structure of sub-Riemannian gradient flow in SDM.

Lastly, optimal transport on the sub-Riemannian manifold has been studied by [30, 32, 24]. An optimal transport metric on a sub-Riemannian manifold is proposed in [30, 32]. In this case, the density manifold still forms an infinite-dimensional Riemannian manifold. The Monge-Ampère equation in sub-Riemannian settings is studied in [24]. Our approach is different. We introduce the sub-Riemannian density manifold (SDM) and study its second-order geometric calculations of relative entropies in SDM. Using those, we propose new Gamma z calculus for degenerate stochastic differential equations and establish the generalized curvature dimension bound. Besides, [36, 40] use the analytical property of optimal transport to formulate Ricci curvature lower bound in general metric space. Different from [36, 40], we focus on the geometric calculations in density manifold introduced by the z direction. Following the second-order geometric calculations in density manifold, we formulate new Gamma calculus and corresponding Ricci curvature tensor in the sub-Riemannian manifold. Besides, our derivation also relates to the entropy methods [31]. Using entropy methods, [4] derive the convergence rate for degenerate drift-diffusion processes with constant diffusion coefficients a. We improve the entropy method by using both Gamma calculus and geometric calculations in density manifold. It helps us to obtain the convergence results for variable diffusion coefficients a.

1.1 Sketch of Main result

We now briefly sketch the main results. Given $b = 0$, directions $a \in \mathbb{R}^{(n+m)\times n}$ defined in degenerate SDE (1.1), and designed new smooth directions $z \in \mathbb{R}^{(n+m)\times m}$, denote a sub-elliptic operator $L\colon C^\infty(\mathbb{R}^{n+m}) \to C^\infty(\mathbb{R}^{n+m})$ as follows:

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

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

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

Lemma 1.1 (Invariant measure)**.** *Suppose SDE* (1.1) *is associated with a unique smooth symmetric invariant measure* $\mathbf{Vol}(x)$*, then*

$$a \otimes \nabla a = -aa^\mathsf{T}\nabla \log \mathbf{Vol}.$$

Remark 1.2. *For simplicity of presentation, we assume that the symmetric invariant measure* $\mathbf{Vol}(x)$ *exists for zero drift case. In general, our result can be extended to the general degenerate drift diffusion processes with unique symmetric invariant measure* (1.1)*. For example, canonical horizontal Brownian motion for a given sub-Riemannian structure is with a particular drift term. We postpone the presentation to Section 3 and refer to Remark 2.9 and Lemma 2.10.*

We are ready to present the following generalized iterative Gamma calculus.

Definition 1.3 (Generalized Gamma z calculus)**.** *Denote Gamma one bilinear forms* $\Gamma_1, \Gamma_1^z\colon C^\infty(\mathbb{R}^{n+m}) \times C^\infty(\mathbb{R}^{n+m}) \to C^\infty(\mathbb{R}^{n+m})$ *as*

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

Define Gamma two bilinear forms $\Gamma_2, \Gamma_2^{z,\mathbf{Vol}}\colon C^\infty(\mathbb{R}^{n+m}) \times C^\infty(\mathbb{R}^{n+m}) \to C^\infty(\mathbb{R}^{n+m})$ *as*

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

and

$$\Gamma_2^{z,\mathbf{Vol}}(f,g) = \frac{1}{2}\Big[L\Gamma_1^z(f,g) - \Gamma_1^z(Lf,g) - \Gamma_1^z(f,Lg)\Big] \quad (1.2)$$

$$+\mathbf{div}_z^{\mathbf{Vol}}\Big(\Gamma_{1,\nabla(aa^\mathsf{T})}(f,g)\Big) - \mathbf{div}_a^{\mathbf{Vol}}\Big(\Gamma_{1,\nabla(zz^\mathsf{T})}(f,g)\Big). \quad (1.3)$$

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

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

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

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

with

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

Remark 1.4. *In literature [16], the Gamma two z operator is defined by* (1.2)*, 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 our generalized Gamma z calculus, we introduce a new term* (1.3)*, 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* (1.3) *is exactly the new bilinear form behind assumption in [12, 13] by the weak form. See details in Remark 5.13.*

We next demonstrate that the summation of Γ_2 and $\Gamma_2^{z,\mathbf{Vol}}$ can induce the following decomposition and bilinear forms. They are naturally extensions of classical Bakry-Émery calculus in Riemannian setttings.

Notation 1.5. *For* $(n+m)\times n$ *matrix* a*, we define matrix* Q *as*

$$Q = \begin{pmatrix} a^\mathsf{T}_{11}a^\mathsf{T}_{11} & \cdots & a^\mathsf{T}_{1(n+m)}a^\mathsf{T}_{1(n+m)}\\ \cdots & a^\mathsf{T}_{i\hat{i}}a^\mathsf{T}_{k\hat{k}} & \cdots\\ a^\mathsf{T}_{n1}a^\mathsf{T}_{n1} & \cdots & a^\mathsf{T}_{n(n+m)}a^\mathsf{T}_{n(n+m)}\end{pmatrix} \in \mathbb{R}^{n^2\times(n+m)^2},$$

with $Q_{ik\hat{i}\hat{k}} = a^\mathsf{T}_{i\hat{i}}a^\mathsf{T}_{k\hat{k}}$*. For* $(n+m)\times m$ *matrix* z*, we define matrix* P *as*

$$P = \begin{pmatrix} z^\mathsf{T}_{11}a^\mathsf{T}_{11} & \cdots & z^\mathsf{T}_{1(n+m)}a^\mathsf{T}_{1(n+m)}\\ \cdots & z^\mathsf{T}_{i\hat{i}}a^\mathsf{T}_{k\hat{k}} & \cdots\\ z^\mathsf{T}_{m\hat{1}}a^\mathsf{T}_{n\hat{1}} & \cdots & z^\mathsf{T}_{m(n+m)}a^\mathsf{T}_{n(n+m)}\end{pmatrix} \in \mathbb{R}^{(nm)\times(n+m)^2},$$

with $P_{ik\hat{i}\hat{k}} = z^\mathsf{T}_{i\hat{i}}a^\mathsf{T}_{k\hat{k}}$*. For any* $\hat{i},\hat{k},\hat{j} = 1,\cdots,n+m$ *and* $i,k = 1,\cdots,n$ *(or* $1,\cdots,m$*). We denote* C *as a* $(n+m)^2\times 1$ *dimensional vector with components defined as*

$$C_{\hat{i}\hat{k}} = \left[\sum_{i,k=1}^n\sum_{i'=1}^{n+m}\left(\langle a^\mathsf{T}_{i\hat{i}}a^\mathsf{T}_{ii'}(\frac{\partial a^\mathsf{T}_{k\hat{k}}}{\partial x_{i'}}), (a^\mathsf{T}\nabla)_k f\rangle_{\mathbb{R}^n} - \langle a^\mathsf{T}_{ki'}a^\mathsf{T}_{i\hat{k}}\frac{\partial a^\mathsf{T}_{i\hat{i}}}{\partial x_{i'}}, (a^\mathsf{T}\nabla)_k f\rangle_{\mathbb{R}^n}\right)\right].$$

Here we keep the notation $(a^{\mathsf T}\nabla)_k f = \sum_{k'=1}^{n+m} a^{\mathsf T}_{kk'}\frac{\partial f}{\partial x_{k'}}$. *We denote* D *as a* $n^2\times 1$ *dimensional vector with components defined as*

$$D_{ik} = \sum_{\hat{i},\hat{k}=1}^{n+m} a^{\mathsf T}_{i\hat{i}}\frac{\partial a^{\mathsf T}_{k\hat{k}}}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}, \quad \text{and} \quad D^T D = \sum_{i,k} D_{ik}D_{ik}.$$

We denote F *as a* $(n+m)^2\times 1$ *dimensional vector with components defined as*

$$F_{\hat{i}\hat{k}} = \left[\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i'=1}^{n+m}\left(\langle a^{\mathsf T}_{i\hat{i}}a^{\mathsf T}_{ii'}(\frac{\partial z^{\mathsf T}_{k\hat{k}}}{\partial x_{i'}}), (z^{\mathsf T}\nabla)_k f\rangle_{\mathbb{R}^m} - \langle z^{\mathsf T}_{ki'}a^{\mathsf T}_{i\hat{k}}\frac{\partial a^{\mathsf T}_{i\hat{i}}}{\partial x_{i'}}, (z^{\mathsf T}\nabla)_k f\rangle_{\mathbb{R}^m}\right)\right].$$

We denote E *as a* $(n\times m)\times 1$ *dimensional vector with components defined as*

$$E_{ik} = \sum_{\hat{i},\hat{k}=1}^{n+m} a^{\mathsf T}_{i\hat{i}}\frac{\partial z^{\mathsf T}_{k\hat{k}}}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}, \quad \text{and} \quad E^{\mathsf T}E = \sum_{i,k} E_{ik}E_{ik}.$$

We denote G *as a vector. In local coordinates, we have*

$$\begin{aligned} G_{\hat{i}\hat{j}} \;=\; & \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{j',\hat{j},i',\hat{i}=1}^{n+m}\Bigg[\left(z^{\mathsf T}_{j\hat{j}}z^{\mathsf T}_{jj'}\frac{\partial}{\partial x_{j'}}a^{\mathsf T}_{i\hat{i}}a^{\mathsf T}_{ii'}\frac{\partial f}{\partial x_{i'}} + z^{\mathsf T}_{j\hat{j}}z^{\mathsf T}_{jj'}\frac{\partial}{\partial x_{j'}}a^{\mathsf T}_{ii'}\frac{\partial f}{\partial x_{i'}}a^{\mathsf T}_{i\hat{i}}\right) \\ & - \left(a^{\mathsf T}_{i\hat{i}}a^{\mathsf T}_{ii'}\frac{\partial}{\partial x_{i'}}z^{\mathsf T}_{j\hat{j}}z^{\mathsf T}_{jj'}\frac{\partial f}{\partial x_{j'}} + a^{\mathsf T}_{i\hat{i}}a^{\mathsf T}_{ii'}\frac{\partial}{\partial x_{i'}}z^{\mathsf T}_{jj'}\frac{\partial f}{\partial x_{j'}}z^{\mathsf T}_{j\hat{j}}\right)\Bigg], \end{aligned}$$

for any smooth function $f:\mathbb{R}^{n+m}\to\mathbb{R}$. *We denote* X *as the vectorization of the Hessian matrix of function* f, $X = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1\partial x_1} \\ \cdots \\ \frac{\partial^2 f}{\partial x_{\hat{i}}\partial x_{\hat{k}}} \\ \cdots \\ \frac{\partial^2 f}{\partial x_{n+m}\partial x_{n+m}} \end{pmatrix} \in \mathbb{R}^{(n+m)^2\times 1}$.

Assumption 1.6. *Assume that, vector* C *belongs to the range of* Q^TQ,

$$\text{i.e. there exists vector } \Lambda, \text{ such that} \quad Q^TQ\Lambda = C.$$

Assumption 1.7. *Assume that there exists vectors* Θ *and* $\widetilde{\Lambda}$, *such that*

$$\begin{aligned} [Q^{\mathsf T}Q + P^{\mathsf T}P]\widetilde{\Lambda} &= C + F, \\ [Q^{\mathsf T}Q + P^{\mathsf T}P]\Theta &= Q^{\mathsf T}D + P^{\mathsf T}E. \end{aligned}$$

Assumption 1.8. *Assume that there exists vectors* $\widehat{\Lambda}$ *and* Θ, *such that*

$$\begin{aligned} [Q^{\mathsf T}Q + P^{\mathsf T}P]\widehat{\Lambda} &= [F + C + G], \\ [Q^{\mathsf T}Q + P^{\mathsf T}P]\Theta &= Q^{\mathsf T}D + P^{\mathsf T}E. \end{aligned}$$

Theorem 1.9 (Generalized Gamma z calculus induced Ricci tensor)**.** *If Assumption 1.8 is satisfied, then the following decomposition holds*

$$\Gamma_2(f,f)+\Gamma_2^{z,\mathbf{Vol}}(f,f) \quad = \quad |\mathfrak{Hess}^G_{a,z}f|^2+\mathfrak{R}^G_a(\nabla f,\nabla f)+\mathfrak{R}_z(\nabla f,\nabla f)+\mathfrak{R}^{\mathbf{Vol}}(\nabla f,\nabla f)$$

where we define

$$|\mathfrak{Hess}^G_{a,z}f|^2=[X+\widehat{\Lambda}+\Theta](Q^{\mathsf{T}}Q+P^{\mathsf{T}}P)[X+\widehat{\Lambda}+\Theta],$$

and denote the following three tensors

$$\begin{aligned}
\mathfrak{R}^G_a(\nabla f,\nabla f) \quad = \quad & -[\widehat{\Lambda}+\Theta]^{\mathsf{T}}((Q^{\mathsf{T}}Q+P^{\mathsf{T}}P))[\widehat{\Lambda}+\Theta]+D^{\mathsf{T}}D+E^{\mathsf{T}}E\\
&+\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}(\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}}\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{\imath}}}\frac{\partial f}{\partial x_{\hat{k}}}),(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&+\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{\imath}}(\frac{\partial}{\partial x_{i'}}\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{\imath}}})(\frac{\partial f}{\partial x_{\hat{k}}}),(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&-\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{k\hat{k}}\frac{\partial a^{\mathsf{T}}_{ii'}}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{\imath}}}),(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&-\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{k\hat{k}}a^{\mathsf{T}}_{ii'}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}})\frac{\partial f}{\partial x_{\hat{\imath}}},(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n},
\end{aligned}$$

$$\begin{aligned}
\mathfrak{R}_z(\nabla f,\nabla f) \quad = \quad & \sum_{i=1}^{n}\sum_{k=}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}(\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}}\frac{\partial z^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{\imath}}}\frac{\partial f}{\partial x_{\hat{k}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\\
&+\sum_{i=1}^{n}\sum_{k=}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{\imath}}(\frac{\partial}{\partial x_{i'}}\frac{\partial z^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{\imath}}})(\frac{\partial f}{\partial x_{\hat{k}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\\
&-\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle z^{\mathsf{T}}_{k\hat{k}}\frac{\partial a^{\mathsf{T}}_{ii'}}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{\imath}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\\
&-\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle z^{\mathsf{T}}_{k\hat{k}}a^{\mathsf{T}}_{ii'}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}})\frac{\partial f}{\partial x_{\hat{\imath}}},(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m},
\end{aligned}$$

and

$$\begin{aligned}
\mathfrak{R}^{\mathbf{Vol}}(\nabla f,\nabla f) \;=\; & 2\sum_{k=1}^{m}\sum_{i=1}^{n}\sum_{k',\hat{k},\hat{\imath},i'=1}^{n+m}\left[\frac{\partial}{\partial x_{k'}} z^{\top}_{kk'} z^{\top}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}} a^{\top}_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}} a^{\top}_{ii'}\frac{\partial f}{\partial x_{i'}}\right] \\
& +2\sum_{k=1}^{m}\sum_{i=1}^{n}\sum_{k',\hat{k},\hat{\imath},i'=1}^{n+m}\Bigg[z^{\top}_{kk'}\frac{\partial}{\partial x_{k'}} z^{\top}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}} a^{\top}_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}} a^{\top}_{ii'}\frac{\partial f}{\partial x_{i'}} \\
& \qquad +z^{\top}_{kk'} z^{\top}_{k\hat{k}}\frac{\partial^2}{\partial x_{k'}\partial x_{\hat{k}}} a^{\top}_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}} a^{\top}_{ii'}\frac{\partial f}{\partial x_{i'}} \\
& \qquad +z^{\top}_{kk'} z^{\top}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}} a^{\top}_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}\frac{\partial}{\partial x_{k'}} a^{\top}_{ii'}\frac{\partial f}{\partial x_{i'}}\Bigg]. \\
& +2\sum_{k=1}^{m}\sum_{i=1}^{n}\sum_{\hat{k},\hat{\imath},i'=1}^{n+m}(z^{\top}\nabla\log\mathbf{Vol})_k\left[z^{\top}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}} a^{\top}_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}} a^{\top}_{ii'}\frac{\partial f}{\partial x_{i'}}\right] \\
& -2\sum_{j=1}^{m}\sum_{l=1}^{n}\sum_{l',\hat{l},\hat{\jmath},j'=1}^{n+m}\left[\frac{\partial}{\partial x_{l'}} a^{\top}_{ll'} a^{\top}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}} z^{\top}_{j\hat{\jmath}}\frac{\partial f}{\partial x_{\hat{\jmath}}} z^{\top}_{jj'}\frac{\partial f}{\partial x_{j'}}\right] \\
& -2\sum_{j=1}^{m}\sum_{l=1}^{n}\sum_{l',\hat{l},\hat{\jmath},j'=1}^{n+m}\Bigg[a^{\top}_{ll'}\frac{\partial}{\partial x_{l'}} a^{\top}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}} z^{\top}_{j\hat{\jmath}}\frac{\partial f}{\partial x_{\hat{\jmath}}} z^{\top}_{jj'}\frac{\partial f}{\partial x_{j'}} \\
& \qquad +a^{\top}_{ll'} a^{\top}_{l\hat{l}}\frac{\partial^2}{\partial x_{l'}\partial x_{\hat{l}}} z^{\top}_{j\hat{\jmath}}\frac{\partial f}{\partial x_{\hat{\jmath}}} z^{\top}_{jj'}\frac{\partial f}{\partial x_{j'}} \\
& \qquad +a^{\top}_{ll'} a^{\top}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}} z^{\top}_{j\hat{\jmath}}\frac{\partial f}{\partial x_{\hat{\jmath}}}\frac{\partial}{\partial x_{l'}} z^{\top}_{jj'}\frac{\partial f}{\partial x_{j'}}\Bigg] \\
& -2\sum_{j=1}^{m}\sum_{l=1}^{n}\sum_{\hat{l},\hat{\jmath},j'=1}^{n+m}(a^{\top}\nabla\log\mathbf{Vol})_l\left[a^{\top}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}} z^{\top}_{j\hat{\jmath}}\frac{\partial f}{\partial x_{\hat{\jmath}}} z^{\top}_{jj'}\frac{\partial f}{\partial x_{j'}}\right].
\end{aligned}$$

Definition 1.10 (Generalized Curvature dimension bound)**.** *We name the generalized curvature-dimension inequality $CD(\kappa,d)$ for degenerate diffusion process generator L by*

$$\Gamma_2(f,f)+\Gamma_2^{z,\mathbf{Vol}}(f,f)\geq\kappa\Gamma_1(f,f)+\kappa\Gamma_1^{z}(f,f)+\frac{1}{d}tr(\mathfrak{Hess}_{a,z}f)^2.$$

In particular, the $CD(\kappa,\infty)$ condition is equivalent to

$$\mathfrak{R}_a^G+\mathfrak{R}_z+\mathfrak{R}^{\mathbf{Vol}}\succeq\kappa(\Gamma_1+\Gamma_1^z).$$

Remark 1.11. *Under Assumption 1.6, 1.7 and 1.8, we derive three types of generalized curvature dimension bounds. Assumption 1.6 provides sufficient condition for horizontal directions only. Assumption 1.7 introduces directions z and is valid for the case when the term* (1.3) *vanishes. In particular, the term* (1.3)$=0$ *is equivalent to $\Gamma_1(f,\Gamma_1^z(f,f))=\Gamma_1^z(f,\Gamma_1(f,f))$ in the weak sens. Assumption 1.8 provides the most general condition for the generalized curvature dimension bound.*

Remark 1.12. *Our generalized curvature dimension bound extends the results shown in [16]. In particular, the Gamma calculus proposed in [12] can result at the third derivative of f for general cases of vector fields a and z. This fact will not result at a valid bound for the convergence and inequality studies. Compared to [12], our Gamma calculus works for general choices of a and z.*

We are now ready to prove the convergence property of degenerate drift diffusion process (1.1) and related functional inequalities. Here we only focus on the $CD(\kappa,\infty)$ case. Denote the a, z–relative Fisher information functionals

$$\mathrm{I}_a(\rho) := \int (a^\mathsf{T}\nabla \log \frac{\rho}{\mathbf{Vol}}, a^\mathsf{T}\nabla \log \frac{\rho}{\mathbf{Vol}})\rho dx,$$

and

$$\mathrm{I}_z(\rho) := \int (z^\mathsf{T}\nabla \log \frac{\rho}{\mathbf{Vol}}, z^\mathsf{T}\nabla \log \frac{\rho}{\mathbf{Vol}})\rho dx.$$

Proposition 1.13 (Hypercontractivity)**.** *Suppose there exists a constant $\kappa > 0$, such that*

$$\mathfrak{R}_a^G + \mathfrak{R}_z + \mathfrak{R}^{\mathbf{Vol}} \succeq \kappa(\Gamma_1 + \Gamma_1^z).$$

Let ρ_0 be a smooth initial distribution and ρ_t be the probability density function of (1.1)*, then ρ_t converges to the invariant measure* **Vol** *in the sense of*

$$\Big(\mathrm{I}_a(\rho_t) + \mathrm{I}_z(\rho_t)\Big) \le e^{-2\kappa t}\Big(\mathrm{I}_a(\rho_0) + \mathrm{I}_z(\rho_0)\Big).$$

Proposition 1.14 (Functional inequalities)**.** *Suppose $\mathfrak{R}_a^G + \mathfrak{R}_z + \mathfrak{R}^{\mathbf{Vol}} \succeq \kappa(\Gamma_1 + \Gamma_1^z)$ with $\kappa > 0$, then the z-Log-Sobolev inequalities holds:*

$$\int_{\mathbb{M}^{n+m}} \rho \log \frac{\rho}{\mathbf{Vol}} dx \le \frac{1}{2\kappa}\Big(\mathrm{I}_a(\rho) + \mathrm{I}_z(\rho)\Big),$$

where ρ is any smooth density.

Besides the above results, our Gamma z calculus also works in degenerate diffusion processes with general (non-gradient) drift function. We define the semigroup $P_t = e^{\frac{1}{2}tL}$ with infinitesimal generator L. For function $f \in \mathcal{C}^\infty(\mathbb{R}^{n+m})$ with compact support, we define the following functionals Φ and ϕ with respect to a and z as below,

$$\begin{aligned}\Phi_a(x,t) &= P_t\left(P_{T-t}f\Gamma_1(\log P_{T-t}f)\right)(x),\\ \Phi_z(x,t) &= P_t\left(P_{T-t}f\Gamma_1^z(\log P_{T-t}f)\right)(x).\end{aligned}$$

And

$$\begin{aligned}\phi_a(x,t) &= P_{T-t}f\Gamma_1(\log P_{T-t}f)(x),\\ \phi_z(x,t) &= P_{T-t}f\Gamma_1^z(\log P_{T-t}f)(x).\end{aligned}$$

Theorem 1.15. *Denote $\phi = \phi_a + \phi_z$, if the following condition is satisfied*

$$\mathfrak{R}_a^G + \mathfrak{R}_z + \mathfrak{R}^{\rho(s)} \succeq \kappa_s(\Gamma_1 + \Gamma_1^z),$$

where tensor $\mathfrak{R}^{\rho(s)}$ depends on the transition kernel $\rho(s,\cdot,\cdot) = p(s,\cdot,\cdot)\mathbf{Vol}$ associated with semi-group P_s (see Notation 3.29 and Definition 4.1 for details) and the bound κ_s depends on the estimate of $\nabla \log \rho(s,\cdot,\cdot)$. We then conclude

$$P_T(\phi(\cdot,T))(x) \quad \geq \quad \phi(x,0) + \int_0^T \kappa_s(\Phi_a(x,s) + \Phi_z(x,s))ds. \tag{1.4}$$

Remark 1.16. *The above result generalizes Theorem* 5.2 *in [16] as well as the related gradient estimates therein. We do not need Hypothesis 1.2 in [16] in our result. In particular, for simplicity, we only take $\phi = \phi_a + \phi_z$ above. We can also take $\phi = v_1(t)\phi_a + v_2(t)\phi_z$ for some time dependent function $v_1(t)$ and $v_2(t)$. This will lead to a variation of the generalized Gamma* 2 *z formula (Theorem 3.32). See Remark 3.33 for the generalized Gamma* 2 *z formula.*

Remark 1.17. *In order to prove the above inequality without the assumption $\Gamma_1(\Gamma_1^z(f,f),f) = \Gamma_1^z(\Gamma_1(f,f),f)$ for any function f. We introduce a new Lemma 4.11, where we further analyze the non-commutative term in its weak form, i.e.*

$$\mathbb{E}\Big[P_{T-t}f\Gamma_1(\log P_{T-t}f, \Gamma_1^z(P_{T-t}f, P_{T-t}f))(x) - P_{T-t}f\Gamma_1^z(\log P_{T-t}f, \Gamma_1(P_{T-t}f, P_{T-t}f))(x)\Big].$$

We are able to produce a new bi-linear form from the above term. It contributes to our newly proposed generalized Gamma z calculus. In order to obtain the above entropic inequality, we further assume that for smooth function $u : \mathbb{R}^{n+m} \to \mathbb{R}$ and for every $T > 0$, $\sup_{t\in[0,T]} \|u(t,\cdot)\|_\infty < \infty$ and $\sup_{t\in[0,T]} \|\frac{1}{2}Lu(t,\cdot) + \partial_t u(t,\cdot)\|_\infty < \infty$. If the test function f in the definition of functional $\phi = \phi_a + \phi_z$ is smooth with compact support, then ϕ should satisfy the aforementioned extra assumption.

We organize the paper as follows. In section 2, we introduce the general framework of degenerate SDEs and their sub-Riemannian structure. In section 3, we establish the generalized Gamma calculus and curvature dimension bound in three different cases. The first case, we only use the horizontal gradient on a sub-Riemannian manifold to derive the curvature dimension bound. If this bound can not be derived and the term (1.3) is zero, we consider the second case. We introduce extra gradient directions z and establish the corresponding classical Gamma z calculus. If the term (1.3) is not constant, we introduce the third case, by considering the generalized Gamma z calculus. In the end, several examples of degenerate drift-diffusion processes are provided. In section 4, we prove gradient estimates and entropy inequalities under the new curvature dimension condition introduced in section 3. In section 5, we present the motivation of this paper. We establish the sub-Riemannian geometry of density manifold supported on a sub-Riemannian manifold. We demonstrate the relationship between the proposed generalized Gamma z calculus and geometric calculations in density manifold in the above three cases.

2 Degenerate SDEs and Sub-Riemannian structure

In this section, we first introduce a general degenerate diffusion process defined by stochastic differential equations (SDEs) with degenerate diffusion coefficients and its associated Fokker-Planck equations. Throughout this paper, we will always assume the density of the solution of the SDE exists and is unique. We thus always assume the Hörmander condition is satisfied for our SDE. Then we present the general framework of sub-Riemannian structure associated with the degenerate SDEs. We then introduce the sub-Riemannian volume form associated with our sub-Riemannian structure and degenerate SDEs.

2.1 Degenerate stochastic differential equation and Fokker-Planck equation

For a general $(n+m)\times n$ matrix a, we denote $a=(a_1,a_2,\cdots,a_n)$ with each $a_i, i=1\cdots,n$, as a $n+m$-dimensional column vector. For any Stratonovich SDE,

$$dX_t=\sum_{i=1}^{n}a_i(X_t)\circ dB_t^i, \tag{2.5}$$

where $(B_t^1,B_t^2,\cdots,B_t^n)$ is a n-dimensional Brownian motion in $\mathbb{R}^n$ and a_i has local coordinates $a_i(x)=\sum_{\hat{i}=1}^{n+m}a_{\hat{i}i}(x)\frac{\partial}{\partial x_{\hat{i}}}$. In general, we assume that $\{a_1,a_2,\cdots,a_n\}$ satisfies the strong Hörmander condition (or bracket generating condition, see details and connection to sub-Riemannian structure in the Section 2.2). In particular, if we consider a drift-diffusion process, namely adding drift term to the above SDE (see Section 3.1.1), then we can relax the condition to be weak Hörmander condition. According to Hörmander theorem, we know the existence, uniqueness and smoothness of the density function of process X_t. Following Stratonovich SDE (2.5), the corresponding Fokker-Planck equation for (2.5) is of the following form (see details in lemma 2.8)

$$2\partial_t\rho \quad = \quad \nabla\cdot(aa^{\mathsf{T}}\nabla\rho)+\nabla\cdot(\rho a\otimes\nabla a), \tag{2.6}$$

where we denote $\rho(t,x)$ as the density function of the process X_t. In particular, we denote the operator

$$L^*\rho = \nabla\cdot(aa^{\mathsf{T}}\nabla\rho) + \nabla\cdot(\rho a\otimes\nabla a).$$

Thus the Fokker-Planck equation is written as

$$2\partial_t\rho \quad = \quad L^*\rho.$$

To make it clear for the notation in equation (2.6), we denote $a = (a_1, a_2, \cdots, a_n)$ and $a_i = \sum_{\hat{i}=1}^{n+m} a_{\hat{i}i}\frac{\partial}{\partial x_{\hat{i}}} = \sum_{\hat{i}=1}^{n+m} a^{\mathsf{T}}_{i\hat{i}}\frac{\partial}{\partial x_{\hat{i}}}$. In terms of local coordinates, we have

$$2\partial_t\rho = \nabla\cdot(aa^{\mathsf{T}}\nabla\rho) + \sum_{\hat{k}=1}^{n+m}\frac{\partial}{\partial x_{\hat{k}}}(\rho\sum_{k=1}^{n}\sum_{k'=1}^{n+m} a_{\hat{k}k}\frac{\partial}{\partial x_{k'}}a_{k'k}),$$

where

$$\begin{aligned}
\nabla\cdot(\rho a\otimes\nabla a) &= \sum_{\hat{k}=1}^{n+m}\frac{\partial}{\partial x_{\hat{k}}}(\rho a\otimes\nabla a)_{\hat{k}} = \sum_{\hat{k}=1}^{n+m}\frac{\partial}{\partial x_{\hat{k}}}(\rho\sum_{k=1}^{n}\sum_{k'=1}^{n+m} a_{\hat{k}k}\frac{\partial}{\partial x_{k'}}a_{k'k}),\\
(a\otimes\nabla a)_{\hat{k}} &= \sum_{k=1}^{n}\sum_{k'=1}^{n+m} a_{\hat{k}k}\frac{\partial}{\partial x_{k'}}a_{k'k} = \sum_{k=1}^{n}\sum_{k'=1}^{n+m} a^T_{k\hat{k}}\frac{\partial}{\partial x_{k'}}a^T_{kk'}. \qquad (2.7)
\end{aligned}$$

Then the corresponding second order operator, the dual operator of L^* in $L^2(\rho)$, is defined as

$$\begin{aligned}
Lf &= \nabla\cdot(aa^{\mathsf{T}}\nabla f) - (a\otimes\nabla a)\nabla f\\
&= \Delta_p f - A\nabla f, \quad where \;\; A = a\otimes\nabla a, \qquad (2.8)
\end{aligned}$$

for any smooth function $f:\mathbb{R}^{n+m}\to\mathbb{R}$. In particular, we denote

$$\Delta_p = \nabla\cdot(aa^{\mathsf{T}}\nabla). \qquad (2.9)$$

We further expand the operator L in local coordinates below,

$$\begin{aligned}
Lf &= \nabla\cdot(aa^{\mathsf{T}}\nabla f) - (a\otimes\nabla a)\nabla f\\
&= \sum_{\hat{k}=1}^{n+m}\frac{\partial}{\partial x_{\hat{k}}}(\sum_{k=1}^{n} a_{\hat{k}k}\sum_{k'=1}^{n+m} a^T_{kk'}\frac{\partial f}{\partial x_{k'}}) - \sum_{k=1}^{n}\sum_{\hat{k},k'=1}^{n+m} a^T_{k\hat{k}}\frac{\partial}{\partial x_{k'}}a^T_{kk'}\frac{\partial f}{\partial x_{\hat{k}}}\\
&= \sum_{k=1}^{n}\sum_{\hat{k},k'=1}^{n+m} a^T_{k\hat{k}}a^T_{kk'}\frac{\partial^2 f}{\partial x_{k'}\partial x_{\hat{k}}} + \sum_{k=1}^{n}\sum_{\hat{k},k'=1}^{n+m} a^T_{k\hat{k}}\frac{\partial a^T_{kk'}}{\partial x_{\hat{k}}}\frac{\partial f}{\partial x_{k'}}. \qquad (2.10)
\end{aligned}$$

Remark 2.1. *For those, who are not familiar with sub-Riemannian geometry, or prefer to stick with the general SDE settings, can directly jump to section 2.3. Only the definition of the operator L is needed to derive the generalized Bakry-Énery calculus and its related curvature dimension inequality. However, we connect general sub-Riemannian structure with our SDE* (2.5) *in the next section 2.2. For more geometry perspective of sub-Riemannian geometry in this context, we refer to [11] and the references therein.*

2.2 A general framework of sub-Riemannian structure.

In general, 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 τ as the sub-bundle of $T\mathbb{M}^{n+m}$, then the sub-Riemannian structure associated with τ on the manifold $\mathbb{M}^{n+m}$ is denoted as (τ, g_τ), where g_τ is the metric defined on the sub-bundle τ. In particular, if we take distribution τ to be the horizontal sub-bundle, denoted as $\mathcal{H}$, of the tangent bundle $T\mathbb{M}^{n+m}$ (see [16, 11] for more details), then we denote the sub-Riemannian structure as $(\mathbb{M}^{n+m}, \mathcal{H}, g_{\mathcal{H}})$. Throughout this paper, we will not distinguish distribution τ and $\mathcal{H}$ and call it horizontal sub-bundle. We will assume that the horizontal distribution $\mathcal{H}$ is bracket generating (with any steps) and the distribution $\mathcal{H}$ has dimension n.

Now, we are going to connect the sub-Riemannian structure to the framework of any Stratonovich SDEs introduced in (2.5). In general, we assume that $\{a_1, a_2, \cdots, a_n\}$ is of rank n and satisfies bracket generating condition (or Hörmander condition) to emphasize the sub-Riemannian structure, i.e. for any $x \in \mathbb{R}^{n+m}$, the Lie bracket of $\{a_1(x), a_2(x), \cdots, a_n(x)\}$ span the whole tangent space at x with dimension $n+m$. Thus X_t lives in a subspace of $\mathbb{R}^{n+m}$. We define the manifold $\mathbb{M}^{n+m}$ as the subspace of $\mathbb{R}^{n+m}$, where the diffusion process X_t lives on and we denote τ as the n-dimensional horizontal distribution of the tangent bundle $T\mathbb{M}^{n+m}$ generated by the vector fields $\{(a_1(x), a_2(x), \cdots, a_n(x)\}$. The rank of $aa^\intercal$ is n, thus the $(n+m)\times(n+m)$ matrix $aa^\intercal$ is degenerate and cannot serve as a metric. We first introduce the following metric.

Definition 2.2. *Consider an orthonormal basis $c = \{c_{n+1}(x), \cdots, c_{n+m}(x)\}$ in $\mathbb{R}^{n+m}$, such that $a_i^\intercal c_j = 0$ for any $1 \le i \le n$, $n+1 \le j \le m+n$. We define a metric $g = (aa^\intercal + cc^\intercal)^{-1} = (aa^\intercal)^\dagger + cc^\intercal$, and metric on the horizontal sub-bundle $g_\tau = (aa^\intercal)^\dagger$, the pseudo-inverse of matrix $aa^\intercal$, on manifold $\mathbb{M}^{n+m}$.*

The above definition is based on the following lemma.

Lemma 2.3. *The metric is $g = (aa^\intercal + cc^\intercal)^{-1} = (aa^\intercal)^\dagger + cc^\intercal$.*

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

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

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

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

where we denote $\Lambda_n = \mathbf{diag}(\lambda_1, \cdots, \lambda_n)$ 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

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

■

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

Lemma 2.4. *The vectors $\{a_1, \cdots, a_n\}$ are orthonormal basis under the metric $g = (aa^\intercal)^\dagger + cc^\intercal$.*

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

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

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

$$
\begin{aligned}
aa^\intercal (aa^\intercal)^\dagger aa^\intercal &= aBa^\intercal \\
aa^\intercal &= aBa^\intercal \\
a^\intercal aa^\intercal a &= a^\intercal aBa^\intercal a \\
(a^\intercal a)^{-1} a^\intercal aa^\intercal a (a^\intercal a)^{-1} &= B \\
\mathbf{Id}_{n\times n} &= B,
\end{aligned}
$$

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

We are now ready to introduce the following definition.

Definition 2.5. *Define $(\mathbb{M}^{n+m}, \tau, g_\tau)$ as the sub-Riemannian structure associated with the degenerate SDE* (2.5)*, where $g_\tau = (aa^\intercal)^\dagger$ denotes the horizontal metric, i.e. metric*

g restricts on the horizontal bundle τ. And we denote ∇^R as the Levi-Civita connection on $\mathbb{M}^{n+m}$ associated with our metric $g = (aa^{\mathsf{T}})^{\dagger} + cc^{\mathsf{T}}$, and let $P^{\tau}\nabla^R$ as the projection of the connection on the horizontal distribution τ. In particular, in our framework, we have $P^{\tau}\nabla^R f = aa^{\mathsf{T}}\nabla f$, for any function $f : \mathbb{M}^{n+m} \to \mathbb{R}$. Where ∇ is the Euclidean gradient in $\mathbb{R}^{n+m}$.

Remark 2.6. *In Lemma 2.4, we show that $\{a_1, a_2, \cdots, a_n\}$ are orthonormal basis for horizontal distribution τ under our metric g. In particular, we have*

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

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 $X = \frac{\partial}{\partial x_1} - \frac{1}{2}x_2\frac{\partial}{\partial x_3}$, $\quad Y = \frac{\partial}{\partial x_2} + \frac{1}{2}x_1\frac{\partial}{\partial x_3}$, $\quad Z = \frac{\partial}{\partial x_3}$ forms an orthonormal basis for the tangent bundle of $\mathbb{H}^1$. In particular, X and Y generate the horizontal distribution τ. If we start with the following SDE

$$dW_t = X \circ dB_t^1 + Y \circ dB_t^2, \tag{2.11}$$

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 $\mathbb{H}^1$. Then W_t is a diffusion process in $\mathbb{R}^3$. In terms of our general sub-Riemannian structure introduced above, we can define

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

and

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

In particular, the horizontal gradient is given by

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

Thus the sub-Riemannian structure associated with Stratonovitch SDE (2.11) 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 τ. Different from the standard construction of Brownian motion on a given Riemannian (sub-Riemannian) manifold by Ells-Elworthy-Malliavin [22, 37], we can directly define our diffusion on the manifold $\mathbb{M}^{n+m}$ by (2.5) without doing projection from the orthonormal frame bundles. It is because that the new metric $g = (aa^\intercal)^\dagger + cc^\intercal$, and $\{a_1, a_2, \cdots, a_n\}$ are globally defined orthonormal basis of the (horizontal) sub-bundle on the tangent bundle $T\mathbb{M}^{n+m}$. Essentially, we first define (2.5) in $\mathbb{R}^{n+m}$, and then introduce the associated sub-Riemannian structure.

Remark 2.7. *Comparing to the definition of horizontal Brownian motion introduced in [15], the sub-Riemannian structure comes first with a totally geodesic Riemannian foliation structure, and then the SDE* (2.11) *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 $(aa^T)^\dagger + cc^\intercal$.*

2.3 Invariant measure.

We focus on the diffusion process (2.5) without drift term first. For the horizontal distribution τ, we define the horizontal Dirichlet form with respect to the invariant measure associated with operator L. Before we introduce the invariant measure, we first give the following useful lemma.

Lemma 2.8. *For a given sub-Riemannian structure $(\mathbb{M}^{n+m}, \tau, (aa^\intercal)^\dagger|_\tau)$, where τ is the horizontal bundle generated by $\{a_1, \cdots, a_n\}$. Suppose that there exists a unique invariant measure $d\mu = \mathbf{Vol}(x)dx$ associated with the horizontal Laplace-Beltrami operator. Then we have*

$$a \otimes \nabla a = -aa^\intercal \nabla \log \mathbf{Vol}.$$

Remark 2.9. *The existence of the invariant measure in the above Lemma 2.8 is easy if $\{a_1, \cdots, a_n\}$ forms left-invariant structures on unimodular Lie groups, since the sub-Laplacian is a sum of squares in this case. The Stratonovich SDE* (2.5) *defines the canonical Brownian motion on sub-Riemannian structure $(\mathbb{M}^{n+m}, \tau, (aa^\intercal)^\dagger|_\tau)$. The* **Vol** *is the canonical volume for the sub-Laplacian. In general, if the Lie group structure is* *not unimodular, we define our SDE with a drift as follows*

$$dX_t = b^a(X_t)dt + \sum_{i=1}^{n} a_i \circ dB_t^i, \tag{2.12}$$

where b^a *is the designed drift term which can be computed explicitly by* $\{a_1, \cdots, a_n\}$ *following [9][Corollary 2]. Thus* (2.12) *is the canonical SDE since the sub-Laplacian induced from this SDE is canonical. Furthermore, the canonical volume exists following [9]. In the next section, we will work with a general drift* b *added to the degenerate system. Thus the results we obtained for general drift-diffusion also applies to SDE* (2.12).

Lemma 2.10. *For a given sub-Riemannian structure* $(\mathbb{M}^{n+m}, \tau, (aa^\mathsf{T})^\dagger|_\tau)$*, where* τ *is the horizontal bundle generated by* $\{a_1, \cdots, a_n\}$*. Consider the drift diffusion process*

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

Suppose that there exists a unique invariant measure $d\mu = \mathbf{Vol}^b(x)dx$ *associated with the canonical sub-Laplacian operator. Then we have*

$$a \otimes \nabla a - 2b = -aa^\mathsf{T}\nabla \log \mathbf{Vol}^b.$$

Proof The proof is similar to the proof for Lemma 2.8 with drift term added to the operator L.

Proof [Proof of Lemma 2.8] Recall that, we have Stratonovich SDE (2.5), $dX_t = \sum_{i=1}^n a_i \circ dB_t^i$. The corresponding Itó SDE is

$$dX_t = \sum_{i=1}^{n} a_i dB_t^i + \frac{1}{2}\sum_{i=1}^{n} \nabla_{a_i} a_i dt,$$

thus the Fokker-Plank equation (Kolmogorov forward equation) for the density function is of the following form

$$\begin{aligned}
2\partial_t \rho(t,x) &= \nabla \cdot (\nabla(aa^\mathsf{T}\rho)) - \nabla \cdot (\sum_{i=1}^{n} \nabla_{a_i} a_i \rho) \\
&= \nabla \cdot (aa^\mathsf{T}\nabla\rho) + \nabla \cdot \left(\rho\nabla(aa^\mathsf{T}) - \rho\sum_{i=1}^{n}\nabla_{a_i}a_i\right) \\
&= \nabla \cdot (aa^\mathsf{T}\nabla\rho) + \nabla \cdot (\rho a \otimes \nabla a),
\end{aligned}$$

where

$$\sum_{i=1}^{n} \nabla_{a_i} a_i = \sum_{i=1}^{n}\sum_{k,l=1}^{n+m} \nabla_{(a_{ik}\frac{\partial}{\partial x_k})}(a_{il}\frac{\partial}{\partial x_l}) = \sum_{i=1}^{n}\sum_{k,l=1}^{n+m} a_{ik}\frac{\partial a_{il}}{\partial x_k}.$$

Namely, we have

$$2\partial_t \rho(t,x) = \nabla \cdot (aa^{\mathsf{T}} \nabla \rho) + \nabla \cdot (\rho a \otimes \nabla a). \tag{2.13}$$

Furthermore, we know that the volume measure is $d\mu(x) = \mathbf{Vol}(x)dx$. The sub-Riemannian structure associated with the metric $(aa^{\mathsf{T}})^{\dagger}$ has the corresponding horizontal Laplace-Beltrami operator w.r.t $d\mu$. Recall that we have $\langle a^{\mathsf{T}}\nabla f, a^{\mathsf{T}}\nabla g\rangle_{\mathbb{R}^n} = \langle aa^{\mathsf{T}}\nabla f, aa^{\mathsf{T}}\nabla g\rangle_{(aa^{\mathsf{T}})^{\dagger}}$, thus

$$\begin{aligned}
\int_{\mathbb{M}^{n+m}} \langle \nabla_{\mathcal{H}} f, \nabla_{\mathcal{H}} g\rangle_{\mathcal{H}} d\mu &= \int_{\mathbb{M}^{n+m}} \langle a^{\mathsf{T}}\nabla f, a^{\mathsf{T}}\nabla g\rangle_{\mathbb{R}^n} d\mu \\
&= \int_{\mathbb{M}^{n+m}} \langle a^{\mathsf{T}}\nabla f, a^{\mathsf{T}}\nabla g\rangle_{\mathbb{R}^n} \mathbf{Vol}(x) dx \\
&= -\int_{\mathbb{M}^{n+m}} \frac{\nabla \cdot (aa^{\mathsf{T}}\mathbf{Vol}\nabla f)}{\mathbf{Vol}} g \mathbf{Vol} dx.
\end{aligned}$$

Thus the horizontal Laplace-Beltrami operator w.r.t $d\mu$ is,

$$\Delta_{\mathcal{H}}^{LB} = -\frac{\nabla \cdot (aa^{\mathsf{T}}\mathbf{Vol}\nabla)}{\mathbf{Vol}}$$

and the corresponding heat equation (Kolmogorov backward equation) is

$$\partial_t u(t,x) = \frac{1}{2}\Delta_{\mathcal{H}}^{LB} u(t,x) = -\frac{1}{2}\frac{\nabla \cdot (aa^{\mathsf{T}}\mathbf{Vol}\nabla u)}{\mathbf{Vol}}.$$

Then the Kolmogorov forward equation associates with the above Kolmogorov backward equation has the following form

$$\begin{aligned}
-2\partial_t (p(t,x)\mathbf{Vol}(x)) &= -\nabla \cdot (aa^{\mathsf{T}}\mathbf{Vol}\nabla p) \\
\Rightarrow 2\partial_t (p(t,x)\mathbf{Vol}(x)) &= \nabla \cdot (\mathbf{Vol}(x) aa^{\mathsf{T}} \nabla (\frac{p(t,x)\mathbf{Vol}(x)}{\mathbf{Vol}(x)})) \\
&= \nabla \cdot (aa^{\mathsf{T}} \nabla (p(t,x)\mathbf{Vol}(x))) \\
&\quad + \nabla \cdot (\mathbf{Vol}(x) aa^{\mathsf{T}} p(t,x)\mathbf{Vol}(x) \frac{-\nabla \mathbf{Vol}(x)}{(\mathbf{Vol}(x))^2}) \\
&= \nabla \cdot (aa^{\mathsf{T}} \nabla (p(t,x)\mathbf{Vol}(x))) - \nabla \cdot (p(t,x)\mathbf{Vol}(x) aa^{\mathsf{T}} \nabla \log \mathbf{Vol}),
\end{aligned}$$

so we end up with

$$2\partial_t (p(t,x)\mathbf{Vol}(x)) = \nabla \cdot (aa^{\mathsf{T}} \nabla (p(t,x)\mathbf{Vol}(x))) - \nabla \cdot (p(t,x)\mathbf{Vol}(x) aa^{\mathsf{T}} \nabla \log \mathbf{Vol}(x). \tag{2.14}$$

Thus comparing the two Fokker-Planck equation (Kolmogorov forward equation) (2.13) and (2.14), we get (by uniqueness) $\rho(t,x) = p(t,x)\mathbf{Vol}(x)$ and in particular,

$$a \otimes \nabla a = -aa^{\mathsf{T}} \nabla \log \mathbf{Vol}.$$

We now introduce the following invariant measure in our sub-Riemannian context.

Definition 2.11 (Invariant measure)**.** *A measure $\mu = \mathbf{Vol}dx$ is called an invariant measure associated with the solution of the following Stratonovich SDE*

$$dX_t = \sum_{i=1}^{n} a_i \circ dB_t^i,$$

if there exists an density ρ^* satisfying

$$\nabla \cdot (aa^{\mathsf{T}} \nabla \rho^*) + \nabla \cdot (\rho^* a \otimes \nabla a) = 0,$$

then $d\mu = \rho^*(x)dx$.

Lemma 2.12. *The operator L is invariant (symmetric) w.r.t the invariant measure $d\mu = \rho^*(x)dx$. Furthermore, the horizontal Dirichlet form can be formulated in the following sense,*

$$-\int_{\mathbb{M}^{n+m}} fLgd\mu = -\int_{\mathbb{M}^{n+m}} gLfd\mu = \int_{\mathbb{M}^{n+m}} \langle a^{\mathsf{T}}\nabla f, a^{\mathsf{T}}\nabla g\rangle_{\mathbb{R}^n} d\mu = \int_{\mathbb{M}^{n+m}} \langle \nabla_{\mathcal{H}} f, \nabla_{\mathcal{H}} g\rangle_{g_\tau} d\mu.$$

See Remark 3.1 in the next section for the relation between $\nabla_{\mathcal{H}} f$ and $a^{\mathsf{T}}\nabla f$.

Remark 2.13. *If the total volume $\int \mathbf{Vol}(x)dx$ is finite, then the above property works for invariant probability measure, with $d\mu = \frac{\mathbf{Vol}(x)dx}{\int_{\mathbb{M}^{n+m}} \mathbf{Vol}(x)dx}$.*

Proof

$$\begin{aligned}
& -\int_{\mathbb{M}^{n+m}} fLgd\mu \\
= & -\int_{\mathbb{M}^{n+m}} f\nabla \cdot (aa^{\mathsf{T}}\nabla g)\rho dx + \int_{\mathbb{M}^{n+m}} fA\nabla g\rho dx \\
= & \int_{\mathbb{M}^{n+m}} \langle \nabla(f\rho), aa^{\mathsf{T}}\nabla g\rangle dx - \int_{\mathbb{M}^{n+m}} \nabla \cdot (f\rho a \otimes \nabla a)gdx \\
= & \int_{\mathbb{M}^{n+m}} \langle \nabla f, aa^{\mathsf{T}}\nabla g\rangle \rho dx + \int_{\mathbb{M}^{n+m}} f\langle \nabla \rho, aa^{\mathsf{T}}\nabla g\rangle dx - \int_{\mathbb{M}^{n+m}} \nabla \cdot (f\rho a \otimes \nabla a)gdx \\
= & \int_{\mathbb{M}^{n+m}} \langle a^{\mathsf{T}}\nabla f, a^{\mathsf{T}}\nabla g\rangle \rho dx - \int_{\mathbb{M}^{n+m}} \nabla \cdot (faa^{\mathsf{T}}\nabla \rho)gdx - \int_{\mathbb{M}^{n+m}} \nabla \cdot (f\rho a \otimes \nabla a)gdx \\
= & \int_{\mathbb{M}^{n+m}} \langle a^{\mathsf{T}}\nabla f, a^{\mathsf{T}}\nabla g\rangle d\mu.
\end{aligned}$$

The last equality follows from the fact that

$$\begin{aligned}
0 &= -\int_{\mathbb{M}^{n+m}} \nabla \cdot (faa^{\mathsf{T}}\nabla \rho^*)gdx - \int_{\mathbb{M}^{n+m}} \nabla \cdot (f\rho^* a \otimes \nabla a)gdx \\
&= -\int_{\mathbb{M}^{n+m}} f\nabla \cdot (aa^{\mathsf{T}}\nabla \rho^*)gdx - \int_{\mathbb{M}^{n+m}} \langle \nabla f, aa^{\mathsf{T}}\nabla \rho^*\rangle gdx \\
&\quad -\int_{\mathbb{M}^{n+m}} \langle \nabla f, \rho^* a \otimes \nabla a\rangle gdx - \int_{\mathbb{M}^{n+m}} f\nabla \cdot (\rho^* a \otimes \nabla a)gdx
\end{aligned}$$

according to the definition of our invariant measure, we know

$$-\int_{\mathbb{M}^{n+m}} f\nabla\cdot(aa^{\mathsf{T}}\nabla\rho^*)gdx - \int_{\mathbb{M}^{n+m}} f\nabla\cdot(\rho^* a\otimes\nabla a)gdx = 0.$$

For our invariant measure $d\mu = \rho^* dx = \mathbf{Vol}(x)dx$ (If it is for invariant probability measure, we have $d\mu = \rho^* dx = \frac{\mathbf{Vol}(x)dx}{\int \mathbf{Vol}(x)dx} = \frac{\mathbf{Vol}(x)dx}{C}$ with $C = \int \mathbf{Vol}(x)dx$) , we have $a\otimes\nabla a = -aa^{\mathsf{T}}\nabla\log\mathbf{Vol}$, thus

$$\begin{aligned}
& -\int_{\mathbb{M}^{n+m}} \langle\nabla f, aa^{\mathsf{T}}\nabla\rho^*\rangle gdx - \int_{\mathbb{M}^{n+m}} \langle\nabla f, \rho^* a\otimes\nabla a\rangle gdx \\
= & -\int_{\mathbb{M}^{n+m}} \langle\nabla f, \rho^* aa^{\mathsf{T}}\nabla(\log\rho^* - \log\mathbf{Vol})\rangle gdx \\
= & \ 0,
\end{aligned}$$

which follows from the fact that $\nabla\Big(\log\rho^* - \log\mathbf{Vol}(x)\Big) = 0$.

2.3.1 Example on $\mathbf{SU}(2)$

The Lie group $\mathbf{SU}(2)$ is a compact connected Lie group, diffeomorphic to the 3-sphere $\mathbb{S}^3$. Following the construction of the left-invariant vector fields in [25, Section 6.2], we change the coordinates in terms of coordinate system (θ,ϕ,ψ). We get 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 and we define our metric $g = (aa^{\mathsf{T}})^{\dagger}$. Here X, Y are orthonormal basis for the horizontal bundle generated by X, Y under metric $(aa^{\mathsf{T}})^{\dagger}$. According to [25, Lemma 6.4], the volume 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 our Lemma 2.8 is satisfied with

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

where

$$a = \begin{pmatrix}\cos\psi & -\sin\psi \\ \frac{\sin\psi}{\sin\theta} & \frac{\cos\psi}{\sin\theta} \\ -\cos\theta\frac{\sin\psi}{\sin\theta} & -\cos\theta\frac{\cos\psi}{\sin\theta}\end{pmatrix}, \quad \text{and} \quad \mathbf{Vol} = \sin(\theta).$$

3 Generalized Gamma z calculus

With the general set up from the previous section, we are ready to introduce the sub-Riemannian (generalized) Bakery-Émery calculus as well as the curvature dimension inequality in this section.

We propose three types of Gamma-2 calculus and the associated curvature dimension inequality. For each of them, we first work with the degenerate diffusion SDEs, and then we introduce extra drift term to the degenerate system. First of all, we derive the horizontal Bakry-Émery Gamma calculus, where only the horizontal directions are involved. Secondly, if the assumption in the first case is not satisfied, and the term (1.3) is zero, we introduce Gamma-2 z following the idea in [16]. Lastly, if the assumption in the first case is not satisfied and the term (1.3) is not zero, we propose our "generalized Gamma-2 z" calculus, which can provide curvature dimension bound in the most general cases. In the end, we demonstrate our results with examples.

3.1 Generalized Bakry-Émery Gamma calculus

Following the previous Section 2.1 of our general Fokker-Planck equation, we now define the Carré de Champ operator Γ_1 associated with

$$\Delta_p f = \nabla \cdot (aa^{\mathsf{T}} \nabla f),$$

and

$$Lf \quad = \quad \Delta_p f - A\nabla f, \quad \text{with } A = a \otimes \nabla a.$$

To distinguish the two Gamma one operators, we denote $\Gamma_{1,a}$ for Δ_p and Γ_1 for L. For any smooth functions $f, g : \mathbb{M}^{n+m} \subset \mathbb{R}^{n+m} \to \mathbb{R}$, we define

$$\begin{aligned} \Gamma_{1,a}(f,g) &= \frac{1}{2}\Big(\Delta_P(f,g) - f\Delta_P g - g\Delta_p f\Big) \\ &= \frac{1}{2}\Big(\nabla \cdot (aa^{\mathsf{T}}\nabla(fg)) - f\nabla \cdot (aa^{\mathsf{T}}\nabla g) - g\nabla \cdot (aa^{\mathsf{T}}\nabla f)\Big) \\ &= \langle \nabla f, aa^{\mathsf{T}}\nabla g\rangle_{\mathbb{R}^{n+m}} = \langle a^{\mathsf{T}}\nabla f, a^{\mathsf{T}}\nabla g\rangle_{\mathbb{R}^n}. \end{aligned} \tag{3.15}$$

Similarly, we define the corresponding Γ_1 as

$$\Gamma_1(f,g) = \frac{1}{2}\left(L(fg) - fLg - gLf\right). \tag{3.16}$$

It is easy to check that $\Gamma_{1,a} = \Gamma_1$, we will keep the convention Γ_1 throughout the paper.

Remark 3.1. *Furthermore, to see the connection between $\Gamma_1(f,g)$ and the sub-Riemanian structure introduced in Definition 2.5, we get*

$$\begin{aligned}\Gamma_1(f,g) &= \langle \nabla f, aa^\mathsf{T}(aa^T)^\dagger aa^\mathsf{T}\nabla g\rangle_{\mathbb{R}^n}\\ &= \langle aa^\mathsf{T}\nabla f, aa^\mathsf{T}\nabla g\rangle_{(aa^\mathsf{T})^\dagger}\\ &= \langle P^\tau\nabla^R f, P^\tau\nabla^R g\rangle_{g_\tau}\\ &= \langle \nabla^R_{\mathcal{H}} f, \nabla^R_{\mathcal{H}} g\rangle_{\mathcal{H}}.\end{aligned}$$

For the first equality, we use the property $(aa^\mathsf{T})(aa^\mathsf{T})^\dagger(aa^\mathsf{T}) = (aa^\mathsf{T})$ for pseudo-inverse matrix. We denote $\nabla^R_{\mathcal{H}} = P^\tau\nabla^R$ as the horizontal gradient and $\langle\cdot,\cdot\rangle_{\mathcal{H}} = \langle\cdot,\cdot\rangle_{(aa^\mathsf{T})^\dagger} = \langle\cdot,\cdot\rangle_{g_\tau}$ as the horizontal metric. The equality $\langle a^\mathsf{T}\nabla f, a^\mathsf{T}\nabla g\rangle_{\mathbb{R}^n} = \langle aa^\mathsf{T}\nabla f, aa^\mathsf{T}\nabla g\rangle_{(aa^\mathsf{T})^\dagger}$ follows from the definition of our metric and transformation from inner product on $\mathbb{R}^n$ to the inner product on the horizontal sub-bundle τ.

From now on, we keep the following notation: $a^\mathsf{T}\nabla f = \sum_{i=1}^n\sum_{\hat{i}=1}^{n+m} a^\mathsf{T}_{i\hat{i}}\nabla_{\frac{\partial}{\partial x_{\hat{i}}}} f$. The same way, we define the iteration of the Carré de Champ operator as (taking $f = g$)

$$\Gamma_{2,a}(f,f) = \frac{1}{2}\Big(\Delta_P\Gamma_1(f,f) - 2\Gamma_1(\Delta_P f, f)\Big). \tag{3.17}$$

Starting from here, we fix the notation for a, a^T with relation $a_{\hat{i}i} = a^\mathsf{T}_{i\hat{i}}$ for $i = 1,\cdots,n$ and $\hat{i} = 1,\cdots,n+m$. Here we denote $a^\mathsf{T}_{i\hat{i}} := (a^\mathsf{T})_{i\hat{i}}$. Now, we can define the second order operator Γ_2 below,

$$\Gamma_2(f,f) = \frac{1}{2}\Big(L\Gamma_1(f,f) - 2\Gamma_1(Lf,f)\Big). \tag{3.18}$$

We first introduce the following notation and assumption.

Notation 3.2. *For matrix $a \in \mathbb{R}^{(n+m)\times n}$, we define matrix $Q \in \mathbb{R}^{n^2\times(n+m)^2}$ in the following form,*

$$Q = \begin{pmatrix} a^\mathsf{T}_{11}a^\mathsf{T}_{11} & \cdots & a^\mathsf{T}_{1(n+m)}a^\mathsf{T}_{1(n+m)}\\ \cdots & a^\mathsf{T}_{i\hat{i}}a^\mathsf{T}_{k\hat{k}} & \cdots\\ a^\mathsf{T}_{n1}a^\mathsf{T}_{n1} & \cdots & a^\mathsf{T}_{n(n+m)}a^\mathsf{T}_{n(n+m)}\end{pmatrix} \in \mathbb{R}^{n^2\times(n+m)^2}. \tag{3.19}$$

In particular, we denote $Q_{ik\hat{i}\hat{k}} = a^\mathsf{T}_{i\hat{i}}a^\mathsf{T}_{k\hat{k}}$. For any smooth function $f : \mathbb{M}^{n+m} \to \mathbb{R}$, we denote X as the vectorization of the Hessian matrix of function f,

$$X = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1\partial x_1}\\ \cdots\\ \frac{\partial^2 f}{\partial x_{\hat{i}}\partial x_{\hat{k}}}\\ \cdots\\ \frac{\partial^2 f}{\partial x_{n+m}\partial x_{n+m}}\end{pmatrix} \in \mathbb{R}^{(n+m)^2\times 1}. \tag{3.20}$$

We further denote $C = (C_{11}, \cdots, C_{(n+m)(n+m)})^T$ as a $(n+m)^2 \times 1$ dimensional vector, for any $\hat{i}, \hat{k} = 1, \cdots, n+m$, $C_{\hat{i}\hat{k}}$ is defined as

$$C_{\hat{i}\hat{k}} = \left[\sum_{i,k=1}^{n} \sum_{i'=1}^{n+m} \left(\langle a^{\mathsf{T}}_{i\hat{i}} a^{\mathsf{T}}_{ii'} (\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{i'}}), (a^{\mathsf{T}}\nabla)_k f \rangle_{\mathbb{R}^n} - \langle a^{\mathsf{T}}_{ki'} a^{\mathsf{T}}_{i\hat{k}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}, (a^{\mathsf{T}}\nabla)_k f \rangle_{\mathbb{R}^n} \right) \right]. \quad (3.21)$$

Here we keep the notation $(a^{\mathsf{T}}\nabla)_k f = \sum_{k'=1}^{n+m} a^{\mathsf{T}}_{kk'} \frac{\partial f}{\partial x_{k'}}$. In particular, we use $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$ here to emphasize that $a^{\mathsf{T}}\nabla f$ is a n-dimensional vector, although we already sum over $\sum_{k=1}^{n}$ in local coordinates. And we denote D as a $n^2 \times 1$ dimensional vector with its coordinate components defined as,

$$D_{ik} = \sum_{\hat{i},\hat{k}=1}^{n+m} a^{\mathsf{T}}_{i\hat{i}} \frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, \quad \text{and} \quad D^T D = \sum_{i,k} D_{ik} D_{ik}. \quad (3.22)$$

Assumption 3.3. *Assume that, C belongs to the range of $Q^T Q$,*

$$\text{i.e. there exists vector } \Lambda, \text{ such that} \quad Q^T Q \Lambda = C. \quad (3.23)$$

We then have the following theorem.

Theorem 3.4. *For any smooth function $f : \mathbb{M}^{n+m} \to \mathbb{R}$, if Assumption 3.3 is satisfied, i.e. there exists vector Λ such that $Q^T Q \Lambda = C$, we have*

$$\Gamma_2(f,f) = |\mathfrak{Hess}_a f|^2 + \mathfrak{R}_a(\nabla f, \nabla f). \quad (3.24)$$

Here we denote

$$\begin{aligned}
\mathfrak{R}_a(\nabla f, \nabla f) \quad = \quad & -2D^T Q \Lambda - \Lambda^T C \\
& + \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{ii'} (\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}} \frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f \rangle_{\mathbb{R}^n} \\
& + \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{ii'} a^{\mathsf{T}}_{i\hat{i}} (\frac{\partial}{\partial x_{i'}} \frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}}) (\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f \rangle_{\mathbb{R}^n} \\
& - \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{k\hat{k}} \frac{\partial a^{\mathsf{T}}_{ii'}}{\partial x_{\hat{k}}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^{\mathsf{T}}\nabla)_k f \rangle_{\mathbb{R}^n} \\
& - \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{k\hat{k}} a^{\mathsf{T}}_{ii'} (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}) \frac{\partial f}{\partial x_{\hat{i}}}, (a^{\mathsf{T}}\nabla)_k f \rangle_{\mathbb{R}^n}
\end{aligned}$$

$$|\mathfrak{Hess}_a f|^2 = [QX + D + Q\Lambda]^{\mathsf{T}} [QX + D + Q\Lambda]. \quad (3.25)$$

The symbols Q, X, C, D are introduced in Notation 3.2 before.

A special case of Theorem 3.4 is the following theorem.

Theorem 3.5. *When matrix Q in Theorem 3.4 is an invertible square matrix, the $\Gamma_2(f,f)$ can be explicitly represented as below. For any smooth function $f : \mathbb{M}^{n+m} \to \mathbb{R}$, we have*

$$\begin{aligned}
&\Gamma_2(f,f) \qquad (3.26)\\
= &\sum_{i,k=1}^{n}\Bigg[\sum_{\hat{i},\hat{k}=1}^{n+m} a^{\mathsf{T}}_{i\hat{i}} a^{\mathsf{T}}_{k\hat{k}} \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}} + \sum_{i',k'=1}^{n+m} a^{\mathsf{T}}_{ii'} \frac{\partial a^{\mathsf{T}}_{kk'}}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}}\\
&+ \left(\sum_{\tilde{i},\tilde{k}=1}^{n}\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{i',l=1}^{n+m}\left(a^{\mathsf{T}}_{\tilde{i}\hat{i}} a^{\mathsf{T}}_{\tilde{i}i'} (\frac{\partial a^{\mathsf{T}}_{\tilde{k}\hat{k}}}{\partial x_{i'}}) a^{\mathsf{T}}_{\tilde{k}l} \frac{\partial f}{\partial x_l} - 2a^{\mathsf{T}}_{\tilde{k}i'} a^{\mathsf{T}}_{\tilde{i}\hat{k}} \frac{\partial a^{\mathsf{T}}_{\tilde{i}\hat{i}}}{\partial x_{i'}} a^{\mathsf{T}}_{\tilde{k}l} \frac{\partial f}{\partial x_l}\right)\right) Q^{-1}_{\hat{i}\hat{k}ik}\Bigg]^2\\
&+\mathfrak{R}_a(\nabla f, \nabla f),
\end{aligned}$$

where we denote

$$\begin{aligned}
&\mathfrak{R}_a(\nabla f,\nabla f) \qquad (3.27)\\
= &-\sum_{i,k=1}^{n}\Bigg[\sum_{i',k'=1}^{n+m} a^{\mathsf{T}}_{ii'} \frac{\partial a^{\mathsf{T}}_{kk'}}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}}\\
&+ \left(\sum_{\tilde{i},\tilde{k}=1}^{n}\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{i',l=1}^{n+m}\left(a^{\mathsf{T}}_{\tilde{i}\hat{i}} a^{\mathsf{T}}_{\tilde{i}i'} (\frac{\partial a^{\mathsf{T}}_{\tilde{k}\hat{k}}}{\partial x_{i'}}) a^{\mathsf{T}}_{\tilde{k}l} \frac{\partial f}{\partial x_l} - 2a^{\mathsf{T}}_{\tilde{k}i'} a^{\mathsf{T}}_{\tilde{i}\hat{k}} \frac{\partial a^{\mathsf{T}}_{\tilde{i}\hat{i}}}{\partial x_{i'}} a^{\mathsf{T}}_{\tilde{k}l} \frac{\partial f}{\partial x_l}\right)\right) Q^{-1}_{\hat{i}\hat{k}ik}\Bigg]^2\\
&+ \sum_{i,k=1}^{n}(\sum_{i',k'=1}^{n+m} a^{\mathsf{T}}_{ii'} \frac{\partial a^{\mathsf{T}}_{kk'}}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}})^2\\
&+ \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{ii'} (\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}} \frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&+ \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{ii'} a^{\mathsf{T}}_{i\hat{i}} (\frac{\partial}{\partial x_{i'}} \frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&- \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{k\hat{k}} \frac{\partial a^{\mathsf{T}}_{ii'}}{\partial x_{\hat{k}}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^{\mathsf{T}}\nabla)_k f\rangle\\
&- \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{k\hat{k}} a^{\mathsf{T}}_{ii'} (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}) \frac{\partial f}{\partial x_{\hat{i}}}, (a^{\mathsf{T}}\nabla)_k f\rangle.
\end{aligned}$$

Proof *The proof follows the same one as in Theorem 3.4, but we apply Lemma 3.15 instead.*

Proposition 3.6. *When vector C, defined in* (3.21)*, is a zero vector, Theorem 3.4 and Theorem 3.5 are unified in the same format. We get*

$$\Gamma_2(f,f) = [QX + D]^{\mathsf{T}}[QX + D] + \mathfrak{R}_a(\nabla f, \nabla f), \tag{3.28}$$

where

$$\begin{aligned}\mathfrak{R}_a(\nabla f,\nabla f) &= \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}(\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}),(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&+\sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{i}}(\frac{\partial}{\partial x_{i'}}\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}),(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&-\sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{k\hat{k}}\frac{\partial a^{\mathsf{T}}_{ii'}}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{i}}}),(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&-\sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{k\hat{k}}a^{\mathsf{T}}_{ii'}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}})\frac{\partial f}{\partial x_{\hat{i}}},(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n},\end{aligned}\tag{3.29}$$

and

$$[QX+D]^{\mathsf{T}}[QX+D]=\sum_{i,k=1}^{n}\left[\sum_{\hat{i},\hat{k}=1}^{n+m}a^{\mathsf{T}}_{i\hat{i}}a^{\mathsf{T}}_{k\hat{k}}\frac{\partial^2 f}{\partial x_{\hat{i}}\partial x_{\hat{k}}}+\sum_{i',k'=1}^{n+m}a^{\mathsf{T}}_{ii'}\frac{\partial a^{\mathsf{T}}_{kk'}}{\partial x_{i'}}\frac{\partial f}{\partial x_{k'}}\right]^2.$$

Proof *The proof follows the same one as in Theorem 3.4 and Remark 3.14.*

We prove the curvature-dimension inequality for operator L below.

Theorem 3.7 (Curvature dimension inequality). *Assume that for $\kappa\in\mathbb{R}$, if*

$$\mathfrak{R}_a(\nabla f,\nabla f)\geq\kappa\Gamma_1(f,f),\tag{3.30}$$

then we have

$$\Gamma_2(f,f)\geq\frac{1}{n}[\boldsymbol{trace}(\mathfrak{Hess}_{\mathfrak{a}})f]^2+\kappa\Gamma_1(f,f).\tag{3.31}$$

We say that the operator L satisfies $CD(\kappa,n)$.

Proof The proof follows directly from Theorem 3.4 and the argument is identical to the curvature dimension inequality on a Riemannian manifold. The dimension parameter is due to Cauchy-Schwarz inequality. If vector $C=0$, we get

$$\begin{aligned}|\mathfrak{Hess}_a f|^2 &= [QX+D]^{\mathsf{T}}[QX+D]\\
&=\sum_{i,k=1}^{n}\left[\sum_{\hat{i},\hat{k}=1}^{n+m}a^{\mathsf{T}}_{i\hat{i}}a^{\mathsf{T}}_{k\hat{k}}\frac{\partial^2 f}{\partial x_{\hat{i}}\partial x_{\hat{k}}}+\sum_{i',k'=1}^{n+m}a^{\mathsf{T}}_{ii'}\frac{\partial a^{\mathsf{T}}_{kk'}}{\partial x_{i'}}\frac{\partial f}{\partial x_{k'}}\right]^2\\
&\geq\frac{1}{n}(Lf)^2.\end{aligned}$$

■

We propose the following definition.

Definition 3.8. *For a general stochastic differential equation* (2.5)*, if the coefficient matrix a satisfies the assumption in Theorem 3.7. Then we say that operator L defined in* (2.10) *satisfies the $CD(\kappa, n)$ condition.*

Before we prove the main results Theorem 3.4 , we make the following remarks.

Remark 3.9. *When the matrix Q is a square matrix and invertible, matrix a is also a square matrix and invertible, the metric we considered is just $g = (aa^\mathsf{T})^{-1}$. In this case, we recover the curvature dimension inequality on Riemannian manifolds.*

Remark 3.10. *If the matrix a has dimension $(n+m)\times n$, then matrix Q has dimension $(n+m)^2\times n^2$. Under Assumption 3.3, we prove a curvature dimension inequality on sub-Riemannian manifolds with sub-Riemannian structures $(\mathbb{M}^{n+m}, \tau, (aa^\mathsf{T})^\dagger|_\tau)$ (Definition 2.5). Comparing to the first sub-Riemannian version of the Bochner-Wietzenböck formula (and generalized curvature dimension inequality) in the literature introduced by Baudoin-Garofalo [16], we do not need the vertical distribution in our setting, which is considered as the complementary set of the horizontal distribution on the whole tangent bundle. Our results show that by introducing the new metric $(aa^\mathsf{T})^\dagger$, only the information from the horizontal sub-bundle τ is enough. In the other word, our method considers the Gamma calculus in the Euclidean space with a specific metric tensor. By routine computation, we can see that condition* (3.3) *is not satisfied for Heisenberg group,* **SU**(2)*,* **SL**(2)*, Eagle group, and Martinet flat sub-Riemannian structure. If Assumption 3.3 is not satisfied, we also need to introduce extra directions in order to get curvature dimension bound. Our condition is an algebraic condition, and we do not require the commutative property $\Gamma_1(\Gamma_1^z(f,f),f) = \Gamma_1^z(\Gamma_1(f,f),f)$. See details in Assumption 3.22 and Assumption 3.30 for more general cases with extra direction z.*

We first prove the following lemmas which will be used to prove Theorem 3.4.

Lemma 3.11.

$$
\begin{aligned}
& \frac{1}{2}\Delta_P\Gamma_1(f,f) - \Gamma_1(\Delta_p f, f) \qquad (3.32)\\
= \; & \frac{1}{2}(a^\mathsf{T}\nabla\circ(a^\mathsf{T}\nabla|a^\mathsf{T}\nabla f|^2)) - \langle a^\mathsf{T}\nabla([(a^\mathsf{T}\nabla)\circ(a^\mathsf{T}\nabla f)]), a^\mathsf{T}\nabla f\rangle_{\mathbb{R}^n}\\
& -\langle B_{n\times n}a^\mathsf{T}\nabla f, a^\mathsf{T}\nabla f\rangle_{\mathbb{R}^n} + \boldsymbol{B}_0,\\
= \; & \boldsymbol{Hess}(\log \mathbf{Vol}).
\end{aligned}
$$

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

$$\mathbf{B}_{lk} = \sum_{j'=1}^{n+m} a^{\mathsf{T}}_{lj'} \sum_{i=1}^{n+m} \frac{\partial^2}{\partial x_i \partial x_{j'}} a_{ik} = \sum_{j'=1}^{n+m} a^{\mathsf{T}}_{lj'} \sum_{i=1}^{n+m} \frac{\partial^2}{\partial x_i \partial x_{j'}} a^{\mathsf{T}}_{ki}, \tag{3.33}$$

$$\boldsymbol{B}_0 = \sum_{l=1}^{n} (a^{\mathsf{T}}\nabla f)_l \left(\sum_{\hat{i},\hat{k},l'=1}^{n+m} \sum_{k=1}^{n} \left(\frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} a^{\mathsf{T}}_{k\hat{k}} (\frac{\partial}{\partial x_{\hat{k}}} a^{\mathsf{T}}_{ll'} \frac{\partial}{\partial x_{l'}} f) - a^{\mathsf{T}}_{ll'} \frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k} (\frac{\partial}{\partial x_{l'}} a^{\mathsf{T}}_{k\hat{k}} \frac{\partial}{\partial x_{\hat{k}}} f) \right) \right).$$

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

$$(a^T\nabla) \circ (a^T \nabla F) = \sum_{i=1}^{n} (a^T\nabla)_i (a^T\nabla F)_i = \sum_{i=1}^{n} \sum_{\hat{i},i'=1}^{n+m} (a^T_{i\hat{i}} \frac{\partial}{\partial x_{\hat{i}}})(a^T_{ii'} \frac{\partial F}{\partial x_{i'}}). \tag{3.34}$$

Remark 3.12. *For simplification of notation, we keep the follow convention*

$$\Gamma_{2,\mathcal{H}}(f,f) = \frac{1}{2}(a^{\mathsf{T}}\nabla \circ (a^{\mathsf{T}}\nabla |a^{\mathsf{T}}\nabla f|^2)) - \langle a^{\mathsf{T}}\nabla((a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)), a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n}. \tag{3.35}$$

For relation (3.32)*, we denote short as*

$$\Gamma_{2,a}(f,f) - \Gamma_{2,\mathcal{H}}(f,f) = -\boldsymbol{B} + \boldsymbol{B}_0 = \boldsymbol{Hess}(\log \mathbf{Vol}). \tag{3.36}$$

In particular, from Proposition 5.6, we know $-\mathbf{B} + \mathbf{B}_0$ *is the Hessian operator of log volume in sub-Riemannian manifold.*

Proof [Proof of Lemma 3.11] By our definition above, we have

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

where we denote

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

So we end up with

$$
\begin{aligned}
\Delta_P \Gamma_1(f,f) &= \sum_{\hat{i}=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}\left(\sum_{\hat{k}=1}^{n+m}a^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{j}}}\sum_{l=1}^{n}(\sum_{\hat{l}=1}^{n+m}a^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}f)^2\right)\right) \\
&\quad + a^{\mathsf{T}}\nabla\circ(a^{\mathsf{T}}\nabla(a^{\mathsf{T}}\nabla f)^2) \\
&= \sum_{k=1}^{n}\sum_{\hat{i}=1}^{n+m}\left(\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}\left(\sum_{\hat{k}=1}^{n+m}a^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}(a^{\mathsf{T}}\nabla f)^2\right)\right) \\
&\quad + (a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla(a^{\mathsf{T}}\nabla f)^2) \\
&= \nabla a\circ(a^{\mathsf{T}}\nabla(a^{\mathsf{T}}\nabla f)^2)+(a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla(a^{\mathsf{T}}\nabla f)^2). \qquad (3.37)
\end{aligned}
$$

Next, we compute the following quantity.

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

where we have

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

We continue with our computation as below,

$$
\begin{aligned}
&\Gamma_1(\Delta_P f,f) \\
=\;& \langle a^{\mathsf{T}}\nabla\left[\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{k=1}^{n}\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}(a^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}f)+(a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)\right], a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n} \\
=\;& \langle a^{\mathsf{T}}\nabla\left[\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{k=1}^{n}\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}(a^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}f)\right], a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n} \\
&+\langle a^{\mathsf{T}}\nabla((a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)), a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n} \\
=\;& \langle a^{\mathsf{T}}\nabla\left(\nabla a\cdot(a^{\mathsf{T}}\nabla f)\right), a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n}+\langle a^{\mathsf{T}}\nabla((a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)), a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n} \\
=\;& \langle\left(a^{\mathsf{T}}\nabla\nabla a\cdot(a^{\mathsf{T}}\nabla f)\right), a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n}+\langle\left(\nabla a\cdot(a^{\mathsf{T}}\nabla(a^{\mathsf{T}}\nabla f))\right), a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n} \\
&+\langle a^{\mathsf{T}}\nabla((a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)), a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n}. \qquad (3.38)
\end{aligned}
$$

From the above, combining (3.37) and (3.38) we further get

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

Recall that we denote a^{T} to emphasize the transpose of the matrix a and $a^{\mathsf{T}}_{i\hat{i}}=a_{\hat{i}i}$,

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

and

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

Subtracting the above two terms, we end up with

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

Now, we eventually get the the following step

$$
\begin{aligned}
& \frac{1}{2}\Delta_P\Gamma_1(f,f) - \Gamma_1(\Delta_p f, f) \\
= & \frac{1}{2}(a^\top\nabla\circ(a^\top\nabla|a^\top\nabla f|^2)) - \langle a^\top\nabla((a^\top\nabla)\circ(a^\top\nabla f)), a^\top\nabla f\rangle_{\mathbb{R}^n} \\
& -\langle\sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{i=1}^{n+m}\sum_{j'=1}^{n+m} a^\top_{lj'}\frac{\partial^2}{\partial x_i x'_j}a_{ik}(a^\top\nabla f)_k, a^\top\nabla f\rangle \\
& +\sum_{l=1}^{n}(a^\top\nabla f)_l\left(\sum_{\hat{i},\hat{k},l'=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}a^\top_{k\hat{k}}(\frac{\partial}{\partial x_{\hat{k}}}a^\top_{ll'}\frac{\partial}{\partial x_{l'}}f) - a^\top_{ll'}\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}(\frac{\partial}{\partial x_{l'}}a^\top_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}f)\right)\right) \\
= & \frac{1}{2}(a^\top\nabla\circ(a^\top\nabla|a^\top\nabla f|^2)) - \langle a^\top\nabla((a^\top\nabla)\circ(a^\top\nabla f)), a^\top\nabla f\rangle_{\mathbb{R}^n} - \langle B_{n\times n}a^\top\nabla f, a^\top\nabla f\rangle_{\mathbb{R}^n} \\
& +\sum_{l=1}^{n}(a^\top\nabla f)_l\left(\sum_{\hat{i},\hat{k},l'=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}a^\top_{k\hat{k}}(\frac{\partial}{\partial x_{\hat{k}}}a^\top_{ll'}\frac{\partial}{\partial x_{l'}}f) - a^\top_{ll'}\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}(\frac{\partial}{\partial x_{l'}}a^\top_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}f)\right)\right).
\end{aligned}
$$

Thus the proof is completed. ■

Below, we further investigate the extra term explicitly in the above Lemma 3.11.

Lemma 3.13.

$$
\begin{aligned}
& \Gamma_{2,\mathcal{H}}(f,f) \qquad (3.39) \\
= & \frac{1}{2}(a^\top\nabla\circ(a^\top\nabla|a^\top\nabla f|^2)) - \langle a^\top\nabla((a^\top\nabla)\circ(a^\top\nabla f)), a^\top\nabla f\rangle_{\mathbb{R}^n} \\
= & [QX+D+Q\Lambda]^\top[QX+D+Q\Lambda] - 2D^\top Q\Lambda - \Lambda^\top C \\
& +\sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^\top_{ii'}(\frac{\partial a^\top_{i\hat{i}}}{\partial x_{i'}}\frac{\partial a^\top_{k\hat{k}}}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (a^\top\nabla)_k f\rangle_{\mathbb{R}^n} \\
& +\sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^\top_{ii'}a^\top_{i\hat{i}}(\frac{\partial}{\partial x_{i'}}\frac{\partial a^\top_{k\hat{k}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}), (a^\top\nabla)_k f\rangle_{\mathbb{R}^n} \\
& -\sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^\top_{k\hat{k}}\frac{\partial a^\top_{ii'}}{\partial x_{\hat{k}}}\frac{\partial a^\top_{i\hat{i}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{i}}}), (a^\top\nabla)_k f\rangle_{\mathbb{R}^n} \\
& -\sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^\top_{k\hat{k}}a^\top_{ii'}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial a^\top_{i\hat{i}}}{\partial x_{i'}})\frac{\partial f}{\partial x_{\hat{i}}}, (a^\top\nabla)_k f\rangle_{\mathbb{R}^n}.
\end{aligned}
$$

Recall that matrix Q and vectors X, C and D are defined in Notation 3.2.

Proof We expand the two terms in lemma 3.13. The first term reads as

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

And the second term reads as

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

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

$$\begin{aligned}
\mathbf{R}_1 = & \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle(a^{\mathsf{T}}_{ii'}\frac{\partial}{\partial x_{i'}})(a^{\mathsf{T}}_{i\hat{i}}\frac{\partial}{\partial x_{\hat{i}}})(a^{\mathsf{T}}_{k\hat{k}}\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
= & \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}(\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\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^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{i}}(\frac{\partial}{\partial x_{i'}}\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \cdots \mathbf{R}_1^2 \\
& + \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{i}}(\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}})(\frac{\partial}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\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^{\mathsf{T}}_{ii'})((\frac{\partial}{\partial x_{i'}}a^{\mathsf{T}}_{i\hat{i}})a^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\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^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{i}}(\frac{\partial}{\partial x_{i'}}a^{\mathsf{T}}_{k\hat{k}})\frac{\partial}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\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^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{i}}a^{\mathsf{T}}_{k\hat{k}}(\frac{\partial}{\partial x_{i'}}\frac{\partial}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \cdots \mathbf{R}_1^6.
\end{aligned}$$

And

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

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

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

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 get the following.

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

We denote

$$\sum_{\hat{i},\hat{k}=1}^{n+m} a_{i\hat{i}}^{\mathsf{T}}a_{k\hat{k}}^{\mathsf{T}}\frac{\partial^2 f}{\partial x_{\hat{i}}\partial x_{\hat{k}}} = \gamma_{ik}. \tag{3.40}$$

The above equality (3.40) can be represented in the following matrix form

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

where Q and X are defined in (3.19) and (3.20). Now, we can represent term $\mathbf{T}_{1a}$ as $\sum_{i,k=1}^{n}\gamma_{ik}^2 = \gamma^T\gamma = (QX)^TQX = X^TQ^TQX$. Next we want to represent $\mathbf{R}_1^3$ and $\mathbf{R}_2^5$ in the following form in terms of vector X,

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

where C is defined in (3.21). Similarly, we can represent $\mathbf{T}_{1b} = \mathbf{T}_{1c}$ by X,

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

where D is defined in (3.22). Once Assumption (3.3) is satisfied, then C belongs to the range of Q^TQ, there exists vector Λ, such that $Q^TQ\Lambda = C$, we thus have the following quadratic form,

$$\begin{aligned}
& \mathbf{T}_1 + 2\mathbf{R}_1^3 - 2\mathbf{R}_2^5 \\
= \; & X^T Q^T Q X + 2D^T Q X + 2C^T X + D^T D \\
= \; & [QX + D + Q\Lambda]^\mathsf{T} [QX + D + Q\Lambda] - 2D^\mathsf{T} Q\Lambda - \Lambda^\mathsf{T} Q^T Q\Lambda \\
= \; & [QX + D + Q\Lambda]^\mathsf{T} [QX + D + Q\Lambda] - 2D^\mathsf{T} Q\Lambda - \Lambda^\mathsf{T} C.
\end{aligned} \tag{3.41}$$

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 end up with

$$\begin{aligned}
& \mathbf{T}_1 + \mathbf{R}_1 - \mathbf{R}_2 \\
= \; & \mathbf{T}_1 + 2\mathbf{R}_1^3 - 2\mathbf{R}_2^5 + \mathbf{R}_1^1 + \mathbf{R}_1^2 - \mathbf{R}_2^1 - \mathbf{R}_2^2,
\end{aligned}$$

which completes the proof. ∎

Remark 3.14. *For the term* (3.41)*, once vector $C = 0$, i.e. $\boldsymbol{R}_1^3 = \boldsymbol{R}_2^5$, the whole process is simplified, we get*

$$\boldsymbol{T}_1 + 2\boldsymbol{R}_1^3 - 2\boldsymbol{R}_2^5 = \boldsymbol{T}_1 = [X + D]^\mathsf{T} Q^\mathsf{T} Q[X + D] = [QX + D]^\mathsf{T} [QX + D],$$

which proves Proposition 3.6.

In Lemma 3.13, the existence of the quadratic form depends on Assumption 3.3. Next, we look at a special case where matrix a is a square matrix and aa^T is invertible. In this case, Q is also a square matrix and also invertible. We have explicit formulas for the quadratic forms in Lemma 3.13.

Lemma 3.15. *When matrix Q is a square matrix and invertible, we have*

$$\begin{aligned}
&\Gamma_{2,\mathcal{H}}(f,f) \qquad (3.42)\\
= &\ \frac{1}{2}(a^{\mathsf{T}}\nabla\circ(a^{\mathsf{T}}\nabla|a^{\mathsf{T}}\nabla f|^2)) - \langle a^{\mathsf{T}}\nabla((a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)), a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n}\\
= &\ \sum_{i,k=1}^{n}\Bigg[\sum_{\hat{i},\hat{k}=1}^{n+m} a^{\mathsf{T}}_{i\hat{i}}a^{\mathsf{T}}_{k\hat{k}}\frac{\partial^2 f}{\partial x_{\hat{i}}\partial x_{\hat{k}}} + \sum_{i',k'=1}^{n+m} a^{\mathsf{T}}_{ii'}\frac{\partial a^{\mathsf{T}}_{kk'}}{\partial x_{i'}}\frac{\partial f}{\partial x_{k'}}\\
&\quad + \left(\sum_{\tilde{i},\tilde{k}=1}^{n}\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{i',l=1}^{n+m}\left(a^{\mathsf{T}}_{\tilde{i}\hat{i}}a^{\mathsf{T}}_{\tilde{i}i'}(\frac{\partial a^{\mathsf{T}}_{\tilde{k}\hat{k}}}{\partial x_{i'}})a^{\mathsf{T}}_{\tilde{k}l}\frac{\partial f}{\partial x_l} - 2a^{\mathsf{T}}_{\tilde{k}i'}a^{\mathsf{T}}_{\tilde{i}\hat{k}}\frac{\partial a^{\mathsf{T}}_{\tilde{i}\hat{i}}}{\partial x_{i'}}a^{\mathsf{T}}_{\tilde{k}l}\frac{\partial f}{\partial x_l}\right)\right)Q^{-1}_{\hat{i}\hat{k}ik}\Bigg]^2\\
&- \sum_{i,k=1}^{n}\Bigg[\sum_{i',k'=1}^{n+m} a^{\mathsf{T}}_{ii'}\frac{\partial a^{\mathsf{T}}_{kk'}}{\partial x_{i'}}\frac{\partial f}{\partial x_{k'}}\\
&\quad + \left(\sum_{\tilde{i},\tilde{k}=1}^{n}\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{i',l=1}^{n+m}\left(a^{\mathsf{T}}_{\tilde{i}\hat{i}}a^{\mathsf{T}}_{\tilde{i}i'}(\frac{\partial a^{\mathsf{T}}_{\tilde{k}\hat{k}}}{\partial x_{i'}})a^{\mathsf{T}}_{\tilde{k}l}\frac{\partial f}{\partial x_l} - 2a^{\mathsf{T}}_{\tilde{k}i'}a^{\mathsf{T}}_{\tilde{i}\hat{k}}\frac{\partial a^{\mathsf{T}}_{\tilde{i}\hat{i}}}{\partial x_{i'}}a^{\mathsf{T}}_{\tilde{k}l}\frac{\partial f}{\partial x_l}\right)\right)Q^{-1}_{\hat{i}\hat{k}ik}\Bigg]^2\\
&+ \sum_{i,k=1}^{n}(\sum_{i',k'=1}^{n+m} a^{\mathsf{T}}_{ii'}\frac{\partial a^{\mathsf{T}}_{kk'}}{\partial x_{i'}}\frac{\partial f}{\partial x_{k'}})^2\\
&+ \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}(\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&+ \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{i}}(\frac{\partial}{\partial x_{i'}}\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&- \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{k\hat{k}}\frac{\partial a^{\mathsf{T}}_{ii'}}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{i}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&- \sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{k\hat{k}}a^{\mathsf{T}}_{ii'}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}})\frac{\partial f}{\partial x_{\hat{i}}}, (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}.
\end{aligned}$$

Proof We follow the same steps in the proof of Lemma 3.13. In particular, matrix Q is an invertible square matrix now. We denote Q^{-1} as the inverse matrix of $Q \in \mathbb{R}^{n^2\times n^2}$, for equation (3.40), we then end up with

$$X = Q^{-1}(\gamma_{11},\cdots,\gamma_{ik},\cdots,\gamma_{nn})^T_{n^2\times 1},$$

where γ is defined in (3.40). In local coordinates, for any $\hat{i},\hat{k} = 1,\cdots,n$, we denote as

$$\frac{\partial^2 f}{\partial x_{\hat{i}}\partial x_{\hat{k}}} = \sum_{\tilde{i},\tilde{k}} Q^{-1}_{\hat{i}\hat{k}\tilde{i}\tilde{k}}\gamma_{\tilde{i}\tilde{k}}.$$

Now we can represent $\mathbf{R}_1^3$ and $\mathbf{R}_2^5$ in the following form,

$$\begin{aligned}2\mathbf{R}_1^3 &= 2\sum_{i,k=1}^{n}\sum_{\hat{i},\hat{k}=1}^{n}\sum_{i',l=1}^{n}\left(a_{i\hat{i}}^{\mathsf{T}}a_{ii'}^{\mathsf{T}}(\frac{\partial a_{k\hat{k}}^{\mathsf{T}}}{\partial x_{i'}})a_{kl}^{\mathsf{T}}\frac{\partial f}{\partial x_l}\sum_{\tilde{i},\tilde{k}}Q_{\hat{i}\hat{k}\tilde{i}\tilde{k}}^{-1}\gamma_{\tilde{i}\tilde{k}}\right),\\ -2\mathbf{R}_2^5 &= -2\sum_{i,k=1}^{n}\sum_{\hat{i},\hat{k}=1}^{n}\sum_{i',l=1}^{n}\left(a_{ki'}^{\mathsf{T}}a_{i\hat{k}}^{\mathsf{T}}\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}a_{kl}^{\mathsf{T}}\frac{\partial f}{\partial x_l}\sum_{\tilde{i},\tilde{k}}Q_{\hat{i}\hat{k}\tilde{i}\tilde{k}}^{-1}\gamma_{\tilde{i}\tilde{k}}\right).\end{aligned}$$

For the term $\mathbf{T}_1 + 2\mathbf{R}_1^3 - 2\mathbf{R}_2^5$, we have

$$\begin{aligned}&\mathbf{T}_1 + 2\mathbf{R}_1^3 - 2\mathbf{R}_2^5\\ =& \sum_{i,k=1}^{n}\left((\gamma_{ik})^2 + 2\gamma_{ik}\sum_{i',k'=1}^{n+m}a_{ii'}^{\mathsf{T}}\frac{\partial a_{kk'}^{\mathsf{T}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{k'}} + (\sum_{i',k'=1}^{n+m}a_{ii'}^{\mathsf{T}}\frac{\partial a_{kk'}^{\mathsf{T}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{k'}})^2\right)\\ &+2\sum_{i,k=1}^{n}\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{i',l=1}^{n+m}\left(a_{i\hat{i}}^{\mathsf{T}}a_{ii'}^{\mathsf{T}}(\frac{\partial a_{k\hat{k}}^{\mathsf{T}}}{\partial x_{i'}})a_{kl}^{\mathsf{T}}\frac{\partial f}{\partial x_l}\sum_{\tilde{i},\tilde{k}}Q_{\hat{i}\hat{k}\tilde{i}\tilde{k}}^{-1}\gamma_{\tilde{i}\tilde{k}}\right)\\ &-2\sum_{i,k=1}^{n}\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{i',l=1}^{n+m}\left(a_{ki'}^{\mathsf{T}}a_{i\hat{k}}^{\mathsf{T}}\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}a_{kl}^{\mathsf{T}}\frac{\partial f}{\partial x_l}\sum_{\tilde{i},\tilde{k}}Q_{\hat{i}\hat{k}\tilde{i}\tilde{k}}^{-1}\gamma_{\tilde{i}\tilde{k}}\right)\\ =& \sum_{i,k=1}^{n}\Bigg[(\gamma_{ik})^2 + 2\gamma_{ik}\sum_{i',k'=1}^{n+m}a_{ii'}^{\mathsf{T}}\frac{\partial a_{kk'}^{\mathsf{T}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{k'}} + (\sum_{i',k'=1}^{n+m}a_{ii'}^{\mathsf{T}}\frac{\partial a_{kk'}^{\mathsf{T}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{k'}})^2\\ &+2\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{i',l=1}^{n+m}\left(a_{i\hat{i}}^{\mathsf{T}}a_{ii'}^{\mathsf{T}}(\frac{\partial a_{k\hat{k}}^{\mathsf{T}}}{\partial x_{i'}})a_{kl}^{\mathsf{T}}\frac{\partial f}{\partial x_l} - 2a_{ki'}^{\mathsf{T}}a_{i\hat{k}}^{\mathsf{T}}\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}a_{kl}^{\mathsf{T}}\frac{\partial f}{\partial x_l}\right)\sum_{\tilde{i},\tilde{k}}Q_{\hat{i}\hat{k}\tilde{i}\tilde{k}}^{-1}\gamma_{\tilde{i}\tilde{k}}\Bigg].\end{aligned}$$

We reorganize the index of $\gamma_{\tilde{i}\tilde{k}}$ for the last term to match the leading term $(\gamma_{ik})^2$, the last term gives (switching ik with $\tilde{i}\tilde{k}$)

$$\begin{aligned}&\sum_{i,k=1}^{n}2\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{i',l=1}^{n+m}\left(a_{i\hat{i}}^{\mathsf{T}}a_{ii'}^{\mathsf{T}}(\frac{\partial a_{k\hat{k}}^{\mathsf{T}}}{\partial x_{i'}})a_{kl}^{\mathsf{T}}\frac{\partial f}{\partial x_l} - 2a_{ki'}^{\mathsf{T}}a_{i\hat{k}}^{\mathsf{T}}\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}a_{kl}^{\mathsf{T}}\frac{\partial f}{\partial x_l}\right)\sum_{\tilde{i},\tilde{k}}^{n}Q_{\hat{i}\hat{k}\tilde{i}\tilde{k}}^{-1}\gamma_{\tilde{i}\tilde{k}}\\ =& \sum_{i,k=1}^{n}2\left(\sum_{\tilde{i},\tilde{k}=1}^{n}\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{i',l=1}^{n+m}\left(a_{\tilde{i}\hat{i}}^{\mathsf{T}}a_{\tilde{i}i'}^{\mathsf{T}}(\frac{\partial a_{\tilde{k}\hat{k}}^{\mathsf{T}}}{\partial x_{i'}})a_{\tilde{k}l}^{\mathsf{T}}\frac{\partial f}{\partial x_l} - a_{\tilde{k}i'}^{\mathsf{T}}a_{\tilde{i}\hat{k}}^{\mathsf{T}}\frac{\partial a_{\tilde{i}\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}a_{\tilde{k}l}^{\mathsf{T}}\frac{\partial f}{\partial x_l}\right)\right)Q_{\hat{i}\hat{k}ik}^{-1}\gamma_{ik}.\end{aligned}$$

Plugging back the previous equation, we get

$$\begin{aligned}
&\mathbf{T}_1 + 2\mathbf{R}_1^3 - 2\mathbf{R}_2^5 \\
= &\sum_{i,k=1}^{n} \Bigg[(\gamma_{ik})^2 + 2\gamma_{ik} \sum_{i',k'=1}^{n+m} a_{ii'}^{\mathsf{T}} \frac{\partial a_{kk'}^{\mathsf{T}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}} + (\sum_{i',k'=1}^{n+m} a_{ii'}^{\mathsf{T}} \frac{\partial a_{kk'}^{\mathsf{T}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}})^2 \\
&+2 \left(\sum_{\tilde{i},\tilde{k}=1}^{n} \sum_{\hat{i},\hat{k}=1}^{n+m} \sum_{i',l=1}^{n+m} \left(a_{\tilde{i}\hat{i}}^{\mathsf{T}} a_{\tilde{i}i'}^{\mathsf{T}} (\frac{\partial a_{\tilde{k}\hat{k}}^{\mathsf{T}}}{\partial x_{i'}}) a_{\tilde{k}l}^{\mathsf{T}} \frac{\partial f}{\partial x_l} - a_{\tilde{k}i'}^{\mathsf{T}} a_{\tilde{i}\hat{k}}^{\mathsf{T}} \frac{\partial a_{\tilde{i}\hat{i}}^{\mathsf{T}}}{\partial x_{i'}} a_{\tilde{k}l}^{\mathsf{T}} \frac{\partial f}{\partial x_l} \right) \right) Q_{\hat{i}\hat{k}ik}^{-1} \gamma_{ik} \Bigg] \\
= &\sum_{i,k=1}^{n} \Bigg[\gamma_{ik} + \sum_{i',k'=1}^{n+m} a_{ii'}^{\mathsf{T}} \frac{\partial a_{kk'}^{\mathsf{T}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}} \\
&\quad + \left(\sum_{\tilde{i},\tilde{k}=1}^{n} \sum_{\hat{i},\hat{k}=1}^{n+m} \sum_{i',l=1}^{n+m} \left(a_{\tilde{i}\hat{i}}^{\mathsf{T}} a_{\tilde{i}i'}^{\mathsf{T}} (\frac{\partial a_{\tilde{k}\hat{k}}^{\mathsf{T}}}{\partial x_{i'}}) a_{\tilde{k}l}^{\mathsf{T}} \frac{\partial f}{\partial x_l} - 2a_{\tilde{k}i'}^{\mathsf{T}} a_{\tilde{i}\hat{k}}^{\mathsf{T}} \frac{\partial a_{\tilde{i}\hat{i}}^{\mathsf{T}}}{\partial x_{i'}} a_{\tilde{k}l}^{\mathsf{T}} \frac{\partial f}{\partial x_l} \right) \right) Q_{\hat{i}\hat{k}ik}^{-1} \Bigg]^2 \\
&- \sum_{i,k=1}^{n} \Bigg[\sum_{i',k'=1}^{n+m} a_{ii'}^{\mathsf{T}} \frac{\partial a_{kk'}^{\mathsf{T}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}} \\
&\quad + \left(\sum_{\tilde{i},\tilde{k}=1}^{n} \sum_{\hat{i},\hat{k}=1}^{n+m} \sum_{i',l=1}^{n+m} \left(a_{\tilde{i}\hat{i}}^{\mathsf{T}} a_{\tilde{i}i'}^{\mathsf{T}} (\frac{\partial a_{\tilde{k}\hat{k}}^{\mathsf{T}}}{\partial x_{i'}}) a_{\tilde{k}l}^{\mathsf{T}} \frac{\partial f}{\partial x_l} - 2a_{\tilde{k}i'}^{\mathsf{T}} a_{\tilde{i}\hat{k}}^{\mathsf{T}} \frac{\partial a_{\tilde{i}\hat{i}}^{\mathsf{T}}}{\partial x_{i'}} a_{\tilde{k}l}^{\mathsf{T}} \frac{\partial f}{\partial x_l} \right) \right) Q_{\hat{i}\hat{k}ik}^{-1} \Bigg]^2 \\
&+ \sum_{i,k=1}^{n} (\sum_{i',k'=1}^{n+m} a_{ii'}^{\mathsf{T}} \frac{\partial a_{kk'}^{\mathsf{T}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{k'}})^2.
\end{aligned}$$

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 end up with

$$\begin{aligned}
&\mathbf{T}_1 + \mathbf{R}_1 - \mathbf{R}_2 \\
= \ &\mathbf{T}_1 + 2\mathbf{R}_1^3 - 2\mathbf{R}_2^5 + \mathbf{R}_1^1 + \mathbf{R}_1^2 - \mathbf{R}_2^1 - \mathbf{R}_2^2.
\end{aligned}$$

Recall that we denote $\gamma_{ik} = \sum_{\hat{i},\hat{k}=1}^{n+m} a_{i\hat{i}}^{\mathsf{T}} a_{k\hat{k}}^{\mathsf{T}} \frac{\partial^2 f}{\partial x_{\hat{i}} \partial x_{\hat{k}}}$ in (3.40), the proof then follows. ■

Proof of Theorem 3.4

With Lemma 3.11 and Lemma 3.13 in hand, we are now ready to prove Theorem 3.4.

Proof [Proof of Theorem 3.4] We plug in the operator L to our definition for Γ_2,

$$\begin{aligned}
\Gamma_2(f,f) &= \frac{1}{2}\Delta_p \Gamma_1(f,f) - \frac{1}{2} A\nabla\Gamma_1(f,f) - \Gamma_1((\Delta_P - A\nabla)f, f) \\
&= \frac{1}{2}\Delta_P \Gamma_1(f,f) - \Gamma_1(\Delta_P f, f) \\
&\quad -\frac{1}{2} A\nabla\Gamma_1(f,f) + \Gamma_1(A\nabla f, f) \\
&= \Gamma_{2,a}(f,f) - \frac{1}{2} A\nabla\Gamma_1(f,f) + \Gamma_1(A\nabla f, f).
\end{aligned}$$

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

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

and

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

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,

$$\mathbf{J}_1 = -\sum_{l=1}^{n}\left(a^{\mathsf{T}}\nabla f\right)_l\left(\sum_{l',\hat{k}=1}^{n+m}\sum_{k=1}^{n}\sum_{k'=1}^{n+m} a_{\hat{k}k}\nabla_{\frac{\partial}{\partial x_{k'}}} a_{k'k}\nabla_{\frac{\partial}{\partial x_{\hat{k}}}} a_{ll'}\nabla_{\frac{\partial}{\partial x_{l'}}} f\right), \tag{3.43}$$

and

$$\begin{aligned}
\mathbf{J}_4 &= \sum_{l=1}^{n}(a^{\mathsf{T}}\nabla f)_l\left(\sum_{l'=1}^{n+m} a^{\mathsf{T}}_{ll'}(\sum_{\hat{k}=1}^{n+m}\nabla_{\frac{\partial}{\partial x_{l'}}}(\sum_{k=1}^{n}\sum_{k'=1}^{n+m} a_{\hat{k}k}\nabla_{\frac{\partial}{\partial x_{k'}}} a_{k'k})\nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f)\right) \\
&= \sum_{l=1}^{n}(a^{\mathsf{T}}\nabla f)_l\left(\sum_{k=1}^{n}\sum_{l'=1}^{n+m}\sum_{\hat{k},k'=1}^{n+m} a^{\mathsf{T}}_{ll'}\nabla_{\frac{\partial}{\partial x_{l'}}} a_{\hat{k}k}\nabla_{\frac{\partial}{\partial x_{k'}}} a_{k'k}\nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f\right) \\
&\quad + \sum_{l=1}^{n}(a^{\mathsf{T}}\nabla f)_l\left(\sum_{k=1}^{n}\sum_{l'=1}^{n+m}\sum_{\hat{k},k'=1}^{n+m} a^{\mathsf{T}}_{ll'}a_{\hat{k}k}(\nabla_{\frac{\partial}{\partial x_{l'}}}\nabla_{\frac{\partial}{\partial x_{k'}}} a_{k'k})\nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f\right).
\end{aligned} \tag{3.44}$$

Recall from the Lemma 3.11 , we have

$$\begin{aligned}
&\Gamma_{2,a}(f,f) \\
=\ &\frac{1}{2}(a^{\mathsf{T}}\nabla\circ(a^{\mathsf{T}}\nabla|a^{\mathsf{T}}\nabla f|^2)) - \langle a^{\mathsf{T}}\nabla((a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)), a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n} \\
&+\sum_{l=1}^{n}(a^{\mathsf{T}}\nabla f)_l\left(\sum_{\hat{i},\hat{k},l'=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}a^{\mathsf{T}}_{k\hat{k}}(\frac{\partial}{\partial x_{\hat{k}}}a^{\mathsf{T}}_{ll'}\frac{\partial}{\partial x_{l'}}f) - a^{\mathsf{T}}_{ll'}\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}(\frac{\partial}{\partial x_{l'}}a^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}f)\right)\right) \\
&-\langle \mathbf{B}_{n\times n}a^{\mathsf{T}}\nabla f, a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n},
\end{aligned}$$

where

$$\langle \mathbf{B}_{n\times n} a^{\mathsf{T}}\nabla f, a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n} = \sum_{l=1}^{n}(a^{\mathsf{T}}\nabla f)_l \left(\sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{i=1}^{n+m}\sum_{k',j'=1}^{n+m} a^{\mathsf{T}}_{lj'}\frac{\partial^2}{\partial x_i x'_j}a_{ik}(a^{\mathsf{T}}_{kk'}\frac{\partial}{\partial x_{k'}}f)\right).$$

Thus, combining with (3.43) and (3.44) , we have

$$\begin{aligned}\Gamma_2(f,f) &= \Gamma_{2,a}(f,f) + \mathbf{J}_1 + \mathbf{J}_4 \\ &= \frac{1}{2}(a^{\mathsf{T}}\nabla \circ (a^{\mathsf{T}}\nabla |a^{\mathsf{T}}\nabla f|^2)) - \langle a^{\mathsf{T}}\nabla[(a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)], a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^n}.\end{aligned}$$

This is the same to say that $\Gamma_2(f,f) = \Gamma_{2,\mathcal{H}}(f,f)$, the proof then follows from Lemma 3.13. (Theorem 3.5 thus follows from Lemma 3.15.) ■

3.1.1 Generalized Bakry-Émery Gamma calculus with drift

In this subsection, we consider the following SDE with both diffusion and drift, namely adding the drift term to SDE (2.5)

$$dX_t = \sum_{i=1}^{n} a_i(X_t)\circ dB_t^i + b(X_t)dt. \tag{3.45}$$

Here we keep the same convention as (2.5), and denote b as a $n+m$-dimensional vector. The corresponding diffusion operator associated with SDE (3.45) is denoted as

$$\begin{aligned}\tilde{L}f &= \nabla\cdot(aa^{\mathsf{T}}\nabla f) - (a\otimes\nabla a - 2b)\nabla f \\ &= \Delta_p f - A\nabla f + 2b\nabla f, \quad where \quad A = a\otimes\nabla a,\end{aligned} \tag{3.46}$$

for any smooth function $f : \mathbb{R}^{n+m} \to \mathbb{R}$. (This is similar to Lemma 2.8 with extra drift term added). Comparing with operator L in equation (2.8), the operator $\tilde{L}$ has one more first order term $2b\nabla f$ here. Similarly, we can define the Carré de Champ operator associated with operator $\tilde{L}$. It is easy to check that

$$\Gamma_{1,\tilde{L}}(f,f) = \Gamma_{1,a}(f,f) = \Gamma_1(f,f). \tag{3.47}$$

And we have iterative Γ_2 as below,

$$\Gamma_{2,\tilde{L}}(f,f) = \frac{1}{2}(\tilde{L}\Gamma_{1,\tilde{L}}(f,f) - 2\Gamma_{1,\tilde{L}}(\tilde{L}f,f)). \tag{3.48}$$

We then have the following theorem.

Theorem 3.16. *For any smooth function $f : \mathbb{M}^{n+m} \to \mathbb{R}$, if Assumption 3.3 is satisfied, i.e. there exists vector Λ such that $Q^TQ\Lambda = C$, we have*

$$\Gamma_{2,\tilde{L}}(f,f) = |\mathfrak{Hess}_a f|^2 + \mathfrak{R}_{a,b}(\nabla f,\nabla f). \tag{3.49}$$

Here we denote $|\mathfrak{Hess}_a f|^2 = [QX + D + Q\Lambda]^T[QX + D + Q\Lambda]$*, and*

$$\begin{aligned}
\mathfrak{R}_{a,b}(\nabla f, \nabla f) &= -2D^T Q\Lambda - \Lambda^T C \\
&+ \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{ii'} (\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}} \frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
&+ \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{ii'} a^{\mathsf{T}}_{i\hat{i}} (\frac{\partial}{\partial x_{i'}} \frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
&- \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{k\hat{k}} \frac{\partial a^{\mathsf{T}}_{ii'}}{\partial x_{\hat{k}}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}, (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
&- \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{k\hat{k}} a^{\mathsf{T}}_{ii'} (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}) \frac{\partial f}{\partial x_{\hat{i}}}, (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
&-2 \sum_{i=1}^{n} \sum_{\hat{i},\hat{k}=1}^{n+m} \langle (a^T_{i\hat{i}} \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial a^T_{i\hat{i}}}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^T \nabla f)_i\rangle_{\mathbb{R}^n}.
\end{aligned}$$

The symbols Q, X, C, D *are introduced in Notation 3.2 before.*

Remark 3.17. *Notice that, we have the following relation comparing to Theorem 3.4,*

$$\mathfrak{R}_{a,b}(\nabla f, \nabla f) = \mathfrak{R}_a(\nabla f, \nabla f) - 2\sum_{i=1}^{n} \sum_{\hat{i},\hat{k}=1}^{n+m} \langle (a^T_{i\hat{i}} \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial a^T_{i\hat{i}}}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^T \nabla f)_i\rangle_{\mathbb{R}^n}.$$

The extra term comes from the drift term b *in SDE* (3.45).

Proof By definition, we have

$$\begin{aligned}
\Gamma_{2,\tilde{L}}(f,f) &= \frac{1}{2}(\tilde{L}\Gamma_{1,\tilde{L}}(f,f) - 2\Gamma_{1,\tilde{L}}(\tilde{L}f,f)) \\
&= \frac{1}{2}\Delta_p \Gamma_1(f,f) - \frac{1}{2}A\nabla\Gamma_1(f,f) + b\nabla\Gamma_1(f,f) \\
&\quad -\Gamma_1((\Delta_P - A\nabla + 2b\nabla)f, f) \\
&= \Gamma_2(f,f) + [b\nabla\Gamma_1(f,f) - \Gamma_1(2b\nabla f, f)].
\end{aligned}$$

The first term $\Gamma_2(f,f)$ follows from the proof of Theorem 3.4. We are left for the other two terms,

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

and

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

Combing the above terms with $\Gamma_2(f,f)$, the proof follows. ■

We prove the curvature-dimension inequality $CD(\rho, n)$ for operator $\tilde{L}$ below.

Proposition 3.18 (Curvature dimension inequality with drift)**.** *Assume that if there exists $\kappa \in \mathbb{R}$ such that the following bound is satisfied,*

$$\mathfrak{R}_{a,b}(\nabla f, \nabla f) \geq \kappa\Gamma_{1,\tilde{L}}(f,f), \tag{3.50}$$

then

$$\Gamma_{2,\tilde{L}}(f,f) \geq \frac{1}{n}[\boldsymbol{trace}(\mathfrak{Hess}_{\mathfrak{a}})f]^2 + \kappa\Gamma_{1,\tilde{L}}(f,f). \tag{3.51}$$

Proof The proof follows directly from Theorem 3.16.

Similar to the diffusion process case, we propose the following definition for the drift-diffusion process.

Definition 3.19. *For a general stochastic differential equation* (3.45)*. If the coefficient matrix a and vector field b satisfies the condition in Proposition 3.18, then we say that operator $\tilde{L}$ defined in* (3.46) *satisfies the $CD(\kappa, n)$ condition.*

3.2 Gamma z calculus

Starting from this section, we are going to dealing with Gamma z calculus with extra direction z introduced to the degenerate system. We first derive the generalized Bakry-Émery calculus with matrix z introduced for the diffusion process (2.5). Then we introduce the drift term to the whole system. We first introduce the following notation.

Notation 3.20. *For matrix $z \in \mathbb{R}^{(n+m)\times m}$, we define matrix $P \in \mathbb{R}^{(nm)\times(n+m)^2}$ in the following form,*

$$P = \begin{pmatrix} z^{\mathsf{T}}_{11}a^{\mathsf{T}}_{11} & \cdots & z^{\mathsf{T}}_{1(n+m)}a^{\mathsf{T}}_{1(n+m)} \\ \cdots & z^{\mathsf{T}}_{i\hat{i}}a^{\mathsf{T}}_{k\hat{k}} & \cdots \\ z^{\mathsf{T}}_{m\hat{1}}a^{\mathsf{T}}_{n\hat{1}} & \cdots & z^{\mathsf{T}}_{m(n+m)}a^{\mathsf{T}}_{n(n+m)} \end{pmatrix} \in \mathbb{R}^{(nm)\times(n+m)^2}. \tag{3.52}$$

In particular, we denote $P_{ik\hat{i}\hat{k}} = z^{\mathsf{T}}_{i\hat{i}} a^{\mathsf{T}}_{k\hat{k}}$. *We further denote* $F = (F_{11}, \cdots, F_{(n+m)(n+m)})^T$ *as a* $(n+m)^2 \times 1$ *dimensional vector, for any* $\hat{i}, \hat{k} = 1, \cdots, n+m$, $F_{\hat{i}\hat{k}}$ *is defined as*

$$F_{\hat{i}\hat{k}} = \left[\sum_{i=1}^{n} \sum_{k=1}^{m} \sum_{i'=1}^{n+m} \left(\langle a^{\mathsf{T}}_{i\hat{i}} a^{\mathsf{T}}_{ii'} (\frac{\partial z^{\mathsf{T}}_{k\hat{k}}}{\partial x_{i'}}), (z^{\mathsf{T}}\nabla)_k f \rangle_{\mathbb{R}^m} - \langle z^{\mathsf{T}}_{ki'} a^{\mathsf{T}}_{i\hat{k}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}, (z^{\mathsf{T}}\nabla)_k f \rangle_{\mathbb{R}^m} \right) \right]. \tag{3.53}$$

Recall that, we keep the notation $(a^{\mathsf{T}}\nabla)_k f = \sum_{k'=1}^{n+m} a^{\mathsf{T}}_{kk'} \frac{\partial f}{\partial x_{k'}}$. *And we denote* E *as a* $(n \times m) \times 1$ *dimensional vector with its coordinate components defined as,*

$$E_{ik} = \sum_{\hat{i},\hat{k}=1}^{n+m} a^{\mathsf{T}}_{i\hat{i}} \frac{\partial z^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}, \quad \text{and} \quad E^{\mathsf{T}}E = \sum_{i,k} E_{ik}E_{ik}, \tag{3.54}$$

for $i = 1, \cdots, n$ *and* $k = 1, \cdots, m$.

Remark 3.21. *The column index* m *of dimension* $(n+m) \times m$ *for matrix* z *in the above Notation 3.20 is not crucial. The column index* m *can be any integer between* m *and* $n+m$, *the proof in Appenndix 7 will be the same. The row index* $n+m$ *should be fixed at* $n+m$ *to pair with matrix* $a \in \mathbb{R}^{(n+m)\times n}$.

Assumption 3.22. *Recall Notation 3.2 for matrix* Q *and vectors* C *and* D. *Assume that, both vectors* $C+F$ *and* $Q^{\mathsf{T}}D + P^{\mathsf{T}}E$ *belongs to the range of* $Q^{\mathsf{T}}Q + P^{\mathsf{T}}P$,

$$\text{i.e. there exists vector } \widetilde{\Lambda}, \text{ such that} \quad [Q^{\mathsf{T}}Q + P^{\mathsf{T}}P]\widetilde{\Lambda} = C + F, \tag{3.55}$$

$$\text{and there exists vector } \Theta, \text{ such that} \quad [Q^{\mathsf{T}}Q + P^{\mathsf{T}}P]\Theta = Q^{\mathsf{T}}D + P^{\mathsf{T}}E. \tag{3.56}$$

With the above notation in hand, we are ready to introduce the following lemma. The proof of the following two lemmas are routine and similar to the proof of Lemma 3.11 and Lemma 3.13. The proof are written in details in the Appendix 7 . We first define the iterative Γ_2^z for operator L in the direction of $z = (z_1, \cdots, z_m)$ below.

$$\Gamma_1^z = \langle z^{\mathsf{T}}\nabla, z^{\mathsf{T}}\nabla f \rangle_{\mathbb{R}^m}, \tag{3.57}$$

and

$$\Gamma_2^z(f,f) = \frac{1}{2} L\Gamma_1^z(f,f) - \Gamma_1^z(Lf,f). \tag{3.58}$$

Lemma 3.23.

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

Lemma 3.24.

$$\begin{aligned}
&\frac{1}{2}(a^{\mathsf{T}}\nabla\circ(a^{\mathsf{T}}\nabla|z^{\mathsf{T}}\nabla f|^2))-\langle z^{\mathsf{T}}\nabla((a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)),z^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^m}\\
=\;& X^TP^TPX+2E^TPX+2F^TX+E^TE\\
&+\sum_{i=1}^{n}\sum_{k=}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}(\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}}\frac{\partial z^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{\imath}}}\frac{\partial f}{\partial x_{\hat{k}}}),(z^{\mathsf{T}}\nabla)_kf\rangle_{\mathbb{R}^m}\\
&+\sum_{i=1}^{n}\sum_{k=}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{\imath}}(\frac{\partial}{\partial x_{i'}}\frac{\partial z^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{\imath}}})(\frac{\partial f}{\partial x_{\hat{k}}}),(z^{\mathsf{T}}\nabla)_kf\rangle_{\mathbb{R}^m}\\
&-\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle z^{\mathsf{T}}_{k\hat{k}}\frac{\partial a^{\mathsf{T}}_{ii'}}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{\imath}}},(z^{\mathsf{T}}\nabla)_kf\rangle\\
&-\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle z^{\mathsf{T}}_{k\hat{k}}a^{\mathsf{T}}_{ii'}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}})\frac{\partial f}{\partial x_{\hat{\imath}}},(z^{\mathsf{T}}\nabla)_kf\rangle.
\end{aligned}$$

With the above twol Lemmas in hand, we then propose the following theorem.

Theorem 3.25. *If Assumption 3.22 is satisfied, then*

$$\Gamma_2(f,f)+\Gamma_2^z(f,f)=|\mathfrak{Hess}_{a,z}f|^2+\mathfrak{R}_a(\nabla f,\nabla f)+\mathfrak{R}_z(\nabla f,\nabla f),\tag{3.59}$$

where $\widetilde{\Lambda}$ and Θ are defined in Assumption 3.22. The symbols are defined in Notation 3.2 and Notation 3.20. We denote

$$|\mathfrak{Hess}_{a,z}f|^2=[X+\widetilde{\Lambda}+\Theta]^{\mathsf{T}}[(Q^{\mathsf{T}}Q+P^{\mathsf{T}}P)][X+\widetilde{\Lambda}+\Theta],\tag{3.60}$$

$$\begin{aligned}
\mathfrak{R}_a(\nabla f,\nabla f)\;=\;& -[\widetilde{\Lambda}+\Theta]^{\mathsf{T}}((Q^{\mathsf{T}}Q+P^{\mathsf{T}}P))[\widetilde{\Lambda}+\Theta]+D^{\mathsf{T}}D+E^{\mathsf{T}}E\\
&+\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}(\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}}\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{\imath}}}\frac{\partial f}{\partial x_{\hat{k}}}),(a^{\mathsf{T}}\nabla)_kf\rangle_{\mathbb{R}^n}\\
&+\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{\imath}}(\frac{\partial}{\partial x_{i'}}\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{\imath}}})(\frac{\partial f}{\partial x_{\hat{k}}}),(a^{\mathsf{T}}\nabla)_kf\rangle_{\mathbb{R}^n}\\
&-\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{k\hat{k}}\frac{\partial a^{\mathsf{T}}_{ii'}}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{\imath}}}),(a^{\mathsf{T}}\nabla)_kf\rangle_{\mathbb{R}^n}\\
&-\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{k\hat{k}}a^{\mathsf{T}}_{ii'}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial a^{\mathsf{T}}_{i\hat{\imath}}}{\partial x_{i'}})\frac{\partial f}{\partial x_{\hat{\imath}}},(a^{\mathsf{T}}\nabla)_kf\rangle_{\mathbb{R}^n},
\end{aligned}$$

and

$$\begin{aligned}\mathfrak{R}_z(\nabla f,\nabla f) &= \sum_{i=1}^{n}\sum_{k=}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a_{ii'}^{\mathsf{T}}(\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}\frac{\partial z_{k\hat{k}}^{\mathsf{T}}}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\\ &+\sum_{i=1}^{n}\sum_{k=}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a_{ii'}^{\mathsf{T}}a_{i\hat{i}}^{\mathsf{T}}(\frac{\partial}{\partial x_{i'}}\frac{\partial z_{k\hat{k}}^{\mathsf{T}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\\ &-\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle z_{k\hat{k}}^{\mathsf{T}}\frac{\partial a_{ii'}^{\mathsf{T}}}{\partial x_{\hat{k}}}\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{i}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\\ &-\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle z_{k\hat{k}}^{\mathsf{T}}a_{ii'}^{\mathsf{T}}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}})\frac{\partial f}{\partial x_{\hat{i}}},(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}. \end{aligned} \quad (3.61)$$

Proof According to Lemma 3.11, Lemma 3.13, Lemma 3.23 and Lemma 3.24, we get

$$\begin{aligned}&\Gamma_2(f,f)+\Gamma_{2,L}^z(f,f)\\ =\ & X^{\mathsf{T}}P^TPX+2E^TPX+2F^TX+E^TE\\ &+X^TQ^TQX+2D^TQX+2C^TX+D^TD\\ &+\mathfrak{R}_a(\nabla f,\nabla f)+\mathfrak{R}_z(\nabla f,\nabla f)-\left[-[\widetilde{\Lambda}+\Theta]^{\mathsf{T}}((Q^{\mathsf{T}}Q+P^{\mathsf{T}}P))[\widetilde{\Lambda}+\Theta]+D^{\mathsf{T}}D+E^{\mathsf{T}}E\right]\\ =\ & X^T[P^{\mathsf{T}}P+Q^{\mathsf{T}}Q]X+2[F^{\mathsf{T}}+C^{\mathsf{T}}]X+2[E^{\mathsf{T}}P+D^{\mathsf{T}}Q]X+D^{\mathsf{T}}D+E^{\mathsf{T}}E\\ &+\mathfrak{R}_a(\nabla f,\nabla f)+\mathfrak{R}_z(\nabla f,\nabla f)-\left[-[\widetilde{\Lambda}+\Theta]^{\mathsf{T}}((Q^{\mathsf{T}}Q+P^{\mathsf{T}}P))[\widetilde{\Lambda}+\Theta]+D^{\mathsf{T}}D+E^{\mathsf{T}}E\right].\end{aligned}$$

Assuming that Assumption 3.22 is satisfied, we get

$$\begin{aligned}&\Gamma_2(f,f)+\Gamma_{2,z}(f,f)\\ =\ & [X+\widetilde{\Lambda}+\Theta]^{\mathsf{T}}[(Q^{\mathsf{T}}Q+P^{\mathsf{T}}P)][X+\widetilde{\Lambda}+\Theta]\\ &-[\widetilde{\Lambda}+\Theta]^{\mathsf{T}}((Q^{\mathsf{T}}Q+P^{\mathsf{T}}P))[\widetilde{\Lambda}+\Theta]+D^{\mathsf{T}}D+E^{\mathsf{T}}E\\ &+\mathfrak{R}_a(\nabla f,\nabla f)+\mathfrak{R}_z(\nabla f,\nabla f)-\left[-[\widetilde{\Lambda}+\Theta]^{\mathsf{T}}((Q^{\mathsf{T}}Q+P^{\mathsf{T}}P))[\widetilde{\Lambda}+\Theta]+D^{\mathsf{T}}D+E^{\mathsf{T}}E\right]\\ =\ & [X+\widetilde{\Lambda}+\Theta]^{\mathsf{T}}[(Q^{\mathsf{T}}Q+P^{\mathsf{T}}P)][X+\widetilde{\Lambda}+\Theta]+\mathfrak{R}_a(\nabla f,\nabla f)+\mathfrak{R}_z(\nabla f,\nabla f),\end{aligned}$$

which completes the proof. ■

3.2.1 Gamma z calculus with drift

We first define the iterative Γ_2^z for operator $\tilde{L}$ in the direction $z=(z_1,\cdots,z_m)$ below,

$$\Gamma_{2,\tilde{L}}^z(f,f)=\frac{1}{2}\tilde{L}\Gamma_1^z(f,f)-\Gamma_1^z(\tilde{L}f,f). \quad (3.62)$$

Theorem 3.26. *If Assumption 3.22 is satisfied, we have*

$$\Gamma_{2,\tilde{L}}(f,f)+\Gamma_{2,\tilde{L}}^z(f,f)=|\mathfrak{Hess}_{a,z}f|^2+\mathfrak{R}_{a,b}(\nabla f,\nabla f)+\mathfrak{R}_{z,b}(\nabla f,\nabla f). \quad (3.63)$$

Where $\widetilde{\Lambda}$, Θ are defined in Assumption 3.22. The symbols are defined in Notation 3.2 and Notation 3.20. We denote

$$|\mathfrak{Hess}_{a,z} f|^2 = [X + \widetilde{\Lambda} + \Theta]^{\mathsf{T}}[(Q^{\mathsf{T}}Q + P^{\mathsf{T}}P)][X + \widetilde{\Lambda} + \Theta], \tag{3.64}$$

$$\begin{aligned}
\mathfrak{R}_{a,b}(\nabla f, \nabla f) \quad = \quad & -[\widetilde{\Lambda} + \Theta]^{\mathsf{T}}((Q^{\mathsf{T}}Q + P^{\mathsf{T}}P))[\widetilde{\Lambda} + \Theta] + D^{\mathsf{T}}D + E^{\mathsf{T}}E \\
& + \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{ii'}^{\mathsf{T}}(\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}} \frac{\partial a_{k\hat{k}}^{\mathsf{T}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
& + \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{ii'}^{\mathsf{T}} a_{i\hat{i}}^{\mathsf{T}}(\frac{\partial}{\partial x_{i'}} \frac{\partial a_{k\hat{k}}^{\mathsf{T}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
& - \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{k\hat{k}}^{\mathsf{T}} \frac{\partial a_{ii'}^{\mathsf{T}}}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
& - \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{k\hat{k}}^{\mathsf{T}} a_{ii'}^{\mathsf{T}}(\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}) \frac{\partial f}{\partial x_{\hat{i}}}, (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
& -2 \sum_{i=1}^{n} \sum_{\hat{i},\hat{k}=1}^{n+m} \langle (a_{i\hat{i}}^{T} \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial a_{i\hat{i}}^{T}}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^{T}\nabla f)_i\rangle_{\mathbb{R}^n},
\end{aligned}$$

and

$$\begin{aligned}
\mathfrak{R}_{z,b}(\nabla f, \nabla f) \quad = \quad & \sum_{i=1}^{n} \sum_{k=}^{m} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{ii'}^{\mathsf{T}}(\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}} \frac{\partial z_{k\hat{k}}^{\mathsf{T}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \\
& + \sum_{i=1}^{n} \sum_{k=}^{m} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a_{ii'}^{\mathsf{T}} a_{i\hat{i}}^{\mathsf{T}}(\frac{\partial}{\partial x_{i'}} \frac{\partial z_{k\hat{k}}^{\mathsf{T}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \\
& - \sum_{i=1}^{n} \sum_{k=1}^{m} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle z_{k\hat{k}}^{\mathsf{T}} \frac{\partial a_{ii'}^{\mathsf{T}}}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \\
& - \sum_{i=1}^{n} \sum_{k=1}^{m} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle z_{k\hat{k}}^{\mathsf{T}} a_{ii'}^{\mathsf{T}}(\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}) \frac{\partial f}{\partial x_{\hat{i}}}, (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \\
& -2 \sum_{i=1}^{m} \sum_{\hat{i},\hat{k}=1}^{n+m} \langle (z_{i\hat{i}}^{T} \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial z_{i\hat{i}}^{T}}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^{T}\nabla f)_i\rangle_{\mathbb{R}^m}.
\end{aligned}$$

Proof By definition, we get

$$\begin{aligned}
\Gamma_{2,\tilde{L}}^{z}(f,f) \quad &= \quad \frac{1}{2}(\tilde{L}\Gamma_1^z(f,f) - 2\Gamma_1^z(\tilde{L}f,f)) \\
&= \quad \frac{1}{2}\Delta_p \Gamma_1^z(f,f) - \frac{1}{2}A\nabla\Gamma_1^z(f,f) + b\nabla\Gamma_1^z(f,f) \\
& \qquad -\Gamma_1^z((\Delta_P - A\nabla + 2b\nabla)f,f) \\
&= \quad \Gamma_2^z(f,f) + [b\nabla\Gamma_1^z(f,f) - \Gamma_1^z(2b\nabla f,f)].
\end{aligned}$$

We are left for the last two terms,

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

and

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

Combing the above terms with $\Gamma_{2,\tilde{L}}(f,f)$ (Theorem 3.16), the proof follows. ■

3.3 Generalized Gamma z calculus

In this section, we introduce our generalized Gamma z calculus which covers the previous two cases. For matrix a, z and smooth function Ψ, we define the new second order "*gamma 2 z*" operator below.

Definition 3.27 (Generalized Gamma 2 z)**.** *For a general smooth function Ψ, we define the following generalized Gamma 2 z formula*

$$\Gamma_2^{z,\Psi}(f,f) = \Gamma_2^z(f,f) + \mathbf{div}_z^{\Psi}(\Gamma_{\nabla(aa^{\mathsf{T}})}f,f) - \mathbf{div}_a^{\Psi}(\Gamma_{\nabla(zz^{\mathsf{T}})}f,f). \tag{3.65}$$

Here for matrix a and z, we denote the divergence operator as

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

and

$$\Gamma_{\nabla(aa^{\mathsf{T}})}(f,f) = \langle\nabla f, \nabla(aa^{\mathsf{T}})\nabla f\rangle, \quad \textit{and} \quad \Gamma_{\nabla(zz^{\mathsf{T}})}(f,f) = \langle\nabla f, \nabla(zz^{\mathsf{T}})\nabla f\rangle.$$

In particular, for two special choice of function Ψ, we denote

$$\Gamma_{2,L}^{z,\mathbf{Vol}}(f,f) = \Gamma_2^z(f,f) + \mathbf{div}_z^{\mathbf{Vol}}(\Gamma_{\nabla(aa^{\mathsf{T}})}f,f) - \mathbf{div}_a^{\mathbf{Vol}}(\Gamma_{\nabla(zz^{\mathsf{T}})}f,f) \tag{3.66}$$

$$\Gamma_{2,L}^{z,\rho}(f,f) = \Gamma_2^z(f,f) + \mathbf{div}_z^{\rho}(\Gamma_{\nabla(aa^{\mathsf{T}})}f,f) - \mathbf{div}_a^{\rho}(\Gamma_{\nabla(zz^{\mathsf{T}})}f,f), \tag{3.67}$$

where **Vol** *is the volume function and $\rho(s,\cdot,\cdot) = p(s,\cdot,\cdot)\mathbf{Vol}(\cdot)$ is the transition kernel function for semi-group P_t with generator L (See definition 4.1).*

Remark 3.28. *In particular, we have the following local coordinates representation.*

$$
\begin{aligned}
\langle\nabla f, \nabla(aa^\top)\nabla f\rangle &= \langle \nabla f, \frac{\partial}{\partial x_{\hat{k}}}(aa^\top)\nabla f\rangle_{\hat{k}=1}^{n+m} = 2\langle a^\top\nabla f, \frac{\partial}{\partial x_{\hat{k}}} a^\top \nabla f\rangle_{\hat{k}=1}^{n+m}\\
&= \left(2\sum_{i=1}^{n}\sum_{\hat{i},i'=1}^{n+m}\frac{\partial}{\partial x_{\hat{k}}}a^\top_{i\hat{i}}\frac{\partial f}{\partial x_{\hat{i}}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}\right)_{\hat{k}=1}^{n+m},\\
\langle\nabla f, \nabla(zz^\top)\nabla f\rangle &= \left(2\sum_{j=1}^{n}\sum_{\hat{j},j'=1}^{n+m}\frac{\partial}{\partial x_{\hat{k}}}z^\top_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}\right)_{\hat{k}=1}^{n+m}.
\end{aligned}
\tag{3.68}
$$

Notation 3.29. *We denote G as a vector. In local coordinates, we have*

$$
2G^\top X = \sum_{\hat{i},\hat{j}=1}^{n+m} 2G_{\hat{i}\hat{j}}X_{\hat{i}\hat{j}}, \tag{3.69}
$$

$$
\begin{aligned}
G_{\hat{i}\hat{j}} = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{j',\hat{j},i',\hat{i}=1}^{n+m}\Bigg[&\left(z^\top_{j\hat{j}}z^\top_{jj'}\frac{\partial}{\partial x_{j'}}a^\top_{i\hat{i}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}} + z^\top_{j\hat{j}}z^\top_{jj'}\frac{\partial}{\partial x_{j'}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}a^\top_{i\hat{i}}\right)\\
&-\left(a^\top_{i\hat{i}}a^\top_{ii'}\frac{\partial}{\partial x_{i'}}z^\top_{j\hat{j}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}} + a^\top_{i\hat{i}}a^\top_{ii'}\frac{\partial}{\partial x_{i'}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}z^\top_{j\hat{j}}\right)\Bigg],
\end{aligned}
$$

where X is defined in (3.20). *We denote $\mathfrak{R}^\Psi$ as a tensor below,*

$$
\begin{aligned}
&\mathfrak{R}^\Psi(f,f) \qquad (3.70)\\
=\;& 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^\top_{kk'}z^\top_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\top_{i\hat{i}}\frac{\partial f}{\partial x_{\hat{i}}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}\right]\\
&+2\sum_{k=1}^{m}\sum_{i=1}^{n}\sum_{k',\hat{k},\hat{i},i'=1}^{n+m}\Bigg[z^\top_{kk'}\frac{\partial}{\partial x_{k'}}z^\top_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\top_{i\hat{i}}\frac{\partial f}{\partial x_{\hat{i}}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}} + z^\top_{kk'}z^\top_{k\hat{k}}\frac{\partial^2}{\partial x_{k'}\partial x_{\hat{k}}}a^\top_{i\hat{i}}\frac{\partial f}{\partial x_{\hat{i}}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}\\
&\qquad + z^\top_{kk'}z^\top_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\top_{i\hat{i}}\frac{\partial f}{\partial x_{\hat{i}}}\frac{\partial}{\partial x_{k'}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}\Bigg]\\
&+2\sum_{k=1}^{m}\sum_{i=1}^{n}\sum_{\hat{k},\hat{i},i'=1}^{n+m}(z^\top\nabla\log\Psi)_k\left[z^\top_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\top_{i\hat{i}}\frac{\partial f}{\partial x_{\hat{i}}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}\right]\\
&-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^\top_{ll'}a^\top_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^\top_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}\right]\\
&-2\sum_{j=1}^{m}\sum_{l=1}^{n}\sum_{l',\hat{l},\hat{j},j'=1}^{n+m}\Bigg[a^\top_{ll'}\frac{\partial}{\partial x_{l'}}a^\top_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^\top_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}} + a^\top_{ll'}a^\top_{l\hat{l}}\frac{\partial^2}{\partial x_{l'}\partial x_{\hat{l}}}z^\top_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}\\
&\qquad + a^\top_{ll'}a^\top_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^\top_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}\frac{\partial}{\partial x_{l'}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}\Bigg]\\
&-2\sum_{j=1}^{m}\sum_{l=1}^{n}\sum_{\hat{l},\hat{j},j'=1}^{n+m}(a^\top\nabla\log\Psi)_l\left[a^\top_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^\top_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}\right].
\end{aligned}
$$

Assumption 3.30. *Assume that there exists* $\widehat{\Lambda}$ *and* Θ*, such that*

$$\begin{aligned} [Q^{\mathsf{T}}Q + P^{\mathsf{T}}P]\widehat{\Lambda} &= F + C + G, \\ \text{and} \qquad [Q^{\mathsf{T}}Q + P^{\mathsf{T}}P]\Theta &= Q^{\mathsf{T}}D + P^{\mathsf{T}}E, \end{aligned}$$

where matrix Q *and* P*, vectors* F*,* C*,* D *and* E *are defined in Notation 3.2 and Notation 3.20.*

We first have the following important lemma.

Lemma 3.31.

$$\mathbf{div}_z^{\Psi}(\Gamma_{\nabla(aa^{\mathsf{T}})}f,f) - \mathbf{div}_a^{\Psi}(\Gamma_{\nabla(zz^{\mathsf{T}})}f,f) = \mathfrak{R}^{\Psi}(f,f) + 2G^{\mathsf{T}}X, \tag{3.71}$$

where X *is defined in Notation 3.2 and* G *and* $\mathfrak{R}^{\Psi}$ *are defined above in Notation 3.29.*

Proof We first look at the first term in the above lemma,

$$\begin{aligned}
& \mathbf{div}_z^{\Psi}(\Gamma_{\nabla(aa^{\mathsf{T}})}f,f) \\
= \; & \frac{\nabla\cdot(zz^{\mathsf{T}}\Psi\Gamma_{\nabla(aa^{\mathsf{T}})}(f,f))}{\Psi} \\
= \; & \sum_{k'=1}^{n+m}\frac{1}{\Psi}\frac{\partial}{\partial x_{k'}}\left[\sum_{k=1}^{m} z_{k'k}\left(\Psi\sum_{\hat{k}=1}^{n+m} z^{\mathsf{T}}_{k\hat{k}}\left(\Gamma_{\nabla(aa^{\mathsf{T}})}(f,f)\right)_{\hat{k}}\right)\right] \\
= \; & \sum_{k'=1}^{n+m}\sum_{k=1}^{m}\left[\frac{\partial}{\partial x_{k'}} z_{k'k}\left(\sum_{\hat{k}=1}^{n+m} z^{\mathsf{T}}_{k\hat{k}}\left(\Gamma_{\nabla(aa^{\mathsf{T}})}(f,f)\right)_{\hat{k}}\right) + z_{k'k}\frac{\partial}{\partial x_{k'}}\left(\sum_{\hat{k}=1}^{n+m} z^{\mathsf{T}}_{k\hat{k}}\left(\Gamma_{\nabla(aa^{\mathsf{T}})}(f,f)\right)_{\hat{k}}\right)\right] \\
& + \sum_{k'=1}^{n+m}\sum_{k=1}^{m}\frac{\partial}{\partial x_{k'}}\log\Psi\left[z_{k'k}\sum_{\hat{k}=1}^{n+m} z^{\mathsf{T}}_{k\hat{k}}\left(\Gamma_{\nabla(aa^{\mathsf{T}})}(f,f)\right)_{\hat{k}}\right] \\
= \; & \sum_{k'=1}^{n+m}\sum_{k=1}^{m}\left[\frac{\partial}{\partial x_{k'}} z^{\mathsf{T}}_{kk'}\left(\sum_{\hat{k}=1}^{n+m} z^{\mathsf{T}}_{k\hat{k}}\left(\Gamma_{\nabla(aa^{\mathsf{T}})}(f,f)\right)_{\hat{k}}\right) + z^{\mathsf{T}}_{kk'}\frac{\partial}{\partial x_{k'}}\left(\sum_{\hat{k}=1}^{n+m} z^{\mathsf{T}}_{k\hat{k}}\left(\Gamma_{\nabla(aa^{\mathsf{T}})}(f,f)\right)_{\hat{k}}\right)\right] \\
& + \sum_{k=1}^{m}(z^{\mathsf{T}}\nabla\log\Psi)_k\left[\sum_{\hat{k}=1}^{n+m} z^{\mathsf{T}}_{k\hat{k}}\left(\Gamma_{\nabla(aa^{\mathsf{T}})}(f,f)\right)_{\hat{k}}\right],
\end{aligned}$$

where $\Gamma_{\nabla(aa^\intercal)}(f,f))_{\hat{k}}$ is defined in (3.68). Plugging in (3.68), we further get

$$
\begin{aligned}
& \mathbf{div}_z^\Psi(\Gamma_{\nabla(aa^\intercal)}f,f) \\
= & \sum_{k'=1}^{n+m}\sum_{k=1}^{m}\left[\frac{\partial}{\partial x_{k'}}z^\intercal_{kk'}\left(\sum_{\hat{k}=1}^{n+m}z^\intercal_{k\hat{k}}\left(2\sum_{i=1}^{n}\sum_{\hat{\imath},i'=1}^{n+m}\frac{\partial}{\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}a^\intercal_{ii'}\frac{\partial f}{\partial x_{i'}}\right)\right)\right] \\
& +\sum_{k'=1}^{n+m}\sum_{k=1}^{m}\left[z^\intercal_{kk'}\frac{\partial}{\partial x_{k'}}\left(\sum_{\hat{k}=1}^{n+m}z^\intercal_{k\hat{k}}\left(2\sum_{i=1}^{n}\sum_{\hat{\imath},i'=1}^{n+m}\frac{\partial}{\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}a^\intercal_{ii'}\frac{\partial f}{\partial x_{i'}}\right)\right)\right] \\
& +\sum_{k=1}^{m}(z^\intercal\nabla\log\Psi)_k\left[\sum_{\hat{k}=1}^{n+m}z^\intercal_{k\hat{k}}\left(2\sum_{i=1}^{n}\sum_{\hat{\imath},i'=1}^{n+m}\frac{\partial}{\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}a^\intercal_{ii'}\frac{\partial f}{\partial x_{i'}}\right)\right] \\
= & 2\sum_{k=1}^{m}\sum_{i=1}^{n}\sum_{k',\hat{k},\hat{\imath},i'=1}^{n+m}\left[\frac{\partial}{\partial x_{k'}}z^\intercal_{kk'}z^\intercal_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}a^\intercal_{ii'}\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{\imath},i'=1}^{n+m}\left[z^\intercal_{kk'}\frac{\partial}{\partial x_{k'}}\left(z^\intercal_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}a^\intercal_{ii'}\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{\imath},i'=1}^{n+m}(z^\intercal\nabla\log\Psi)_k\left[z^\intercal_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}a^\intercal_{ii'}\frac{\partial f}{\partial x_{i'}}\right]\cdots\mathcal{S}_3^z \\
= & \mathcal{S}_1^z+\mathcal{S}_2^z+\mathcal{S}_3^z. \qquad (3.72)
\end{aligned}
$$

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

$$
\begin{aligned}
\mathcal{S}_2^z = & 2\sum_{k=1}^{m}\sum_{i=1}^{n}\sum_{k',\hat{k},\hat{\imath},i'=1}^{n+m}\left[z^\intercal_{kk'}\frac{\partial}{\partial x_{k'}}\left(z^\intercal_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}a^\intercal_{ii'}\frac{\partial f}{\partial x_{i'}}\right)\right] \\
= & 2\sum_{k=1}^{m}\sum_{i=1}^{n}\sum_{k',\hat{k},\hat{\imath},i'=1}^{n+m}\left[z^\intercal_{kk'}\frac{\partial}{\partial x_{k'}}z^\intercal_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}a^\intercal_{ii'}\frac{\partial f}{\partial x_{i'}}+z^\intercal_{kk'}z^\intercal_{k\hat{k}}\frac{\partial^2}{\partial x_{k'}\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}a^\intercal_{ii'}\frac{\partial f}{\partial x_{i'}}\right. \\
& +z^\intercal_{kk'}z^\intercal_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial^2 f}{\partial x_{k'}\partial x_{\hat{\imath}}}a^\intercal_{ii'}\frac{\partial f}{\partial x_{i'}}+z^\intercal_{kk'}z^\intercal_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}a^\intercal_{ii'}\frac{\partial^2 f}{\partial x_{k'}\partial x_{i'}} \\
& \left.+z^\intercal_{kk'}z^\intercal_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\intercal_{i\hat{\imath}}\frac{\partial f}{\partial x_{\hat{\imath}}}\frac{\partial}{\partial x_{k'}}a^\intercal_{ii'}\frac{\partial f}{\partial x_{i'}}\right].
\end{aligned}
$$

Similarly, we get

$$
\begin{aligned}
&\mathbf{div}_a^{\Psi}(\Gamma_{\nabla(zz^{\mathsf{T}})}f,f)\\
=\;& \sum_{l'=1}^{n+m}\sum_{l=1}^{n}\left[\frac{\partial}{\partial x_{l'}}a^{\mathsf{T}}_{ll'}\left(\sum_{\hat{l}=1}^{n+m}a^{\mathsf{T}}_{l\hat{l}}\left(\Gamma_{\nabla(zz^{\mathsf{T}})}(f,f)\right)_{\hat{l}}\right)+a^{\mathsf{T}}_{ll'}\frac{\partial}{\partial x_{l'}}\left(\sum_{\hat{l}=1}^{n+m}a^{\mathsf{T}}_{l\hat{l}}\left(\Gamma_{\nabla(zz^{\mathsf{T}})}(f,f)\right)_{\hat{l}}\right)\right]\\
&+\sum_{l=1}^{n}(a^{\mathsf{T}}\nabla\log\Psi)_l\left[\sum_{\hat{l}=1}^{n+m}\left(a^{\mathsf{T}}_{l\hat{l}}\Gamma_{\nabla(zz^{\mathsf{T}})}(f,f)\right)_{\hat{l}}\right]\\
=\;& 2\sum_{j=1}^{m}\sum_{l=1}^{n}\sum_{l',\hat{l},\hat{j},j'=1}^{n+m}\left[\frac{\partial}{\partial x_{l'}}a^{\mathsf{T}}_{ll'}a^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^{\mathsf{T}}_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^{\mathsf{T}}_{jj'}\frac{\partial f}{\partial x_{j'}}\right]\cdots\mathcal{S}_1^a\\
&+2\sum_{j=1}^{m}\sum_{l=1}^{n}\sum_{l',\hat{l},\hat{j},j'=1}^{n+m}\left[a^{\mathsf{T}}_{ll'}\frac{\partial}{\partial x_{l'}}\left(a^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^{\mathsf{T}}_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^{\mathsf{T}}_{jj'}\frac{\partial f}{\partial x_{j'}}\right)\right]\cdots\mathcal{S}_2^a\\
&+2\sum_{j=1}^{m}\sum_{l=1}^{n}\sum_{\hat{l},\hat{j},j'=1}^{n+m}(a^{\mathsf{T}}\nabla\log\Psi)_l\left[a^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^{\mathsf{T}}_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^{\mathsf{T}}_{jj'}\frac{\partial f}{\partial x_{j'}}\right]\cdots\mathcal{S}_3^a\\
=\;& \mathcal{S}_1^a+\mathcal{S}_2^a+\mathcal{S}_3^a,
\end{aligned}
\tag{3.73}
$$

where we also get

$$
\begin{aligned}
\mathcal{S}_2^a =\;& 2\sum_{j=1}^{m}\sum_{l=1}^{n}\sum_{l',\hat{l},\hat{j},j'=1}^{n+m}\left[a^{\mathsf{T}}_{ll'}\frac{\partial}{\partial x_{l'}}\left(a^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^{\mathsf{T}}_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^{\mathsf{T}}_{jj'}\frac{\partial f}{\partial x_{j'}}\right)\right]\\
=\;& 2\sum_{j=1}^{m}\sum_{l=1}^{n}\sum_{l',\hat{l},\hat{j},j'=1}^{n+m}\left[a^{\mathsf{T}}_{ll'}\frac{\partial}{\partial x_{l'}}a^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^{\mathsf{T}}_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^{\mathsf{T}}_{jj'}\frac{\partial f}{\partial x_{j'}}+a^{\mathsf{T}}_{ll'}a^{\mathsf{T}}_{l\hat{l}}\frac{\partial^2}{\partial x_{l'}\partial x_{\hat{l}}}z^{\mathsf{T}}_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^{\mathsf{T}}_{jj'}\frac{\partial f}{\partial x_{j'}}\right.\\
&\qquad+a^{\mathsf{T}}_{ll'}a^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^{\mathsf{T}}_{j\hat{j}}\frac{\partial^2 f}{\partial x_{l'}\partial x_{\hat{j}}}z^{\mathsf{T}}_{jj'}\frac{\partial f}{\partial x_{j'}}+a^{\mathsf{T}}_{ll'}a^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^{\mathsf{T}}_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^{\mathsf{T}}_{jj'}\frac{\partial^2 f}{\partial x_{l'}\partial x_{j'}}\\
&\qquad\left.+a^{\mathsf{T}}_{ll'}a^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^{\mathsf{T}}_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}\frac{\partial}{\partial x_{l'}}z^{\mathsf{T}}_{jj'}\frac{\partial f}{\partial x_{j'}}\right].
\end{aligned}
$$

Combining all the terms above, we end up with

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

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 $\mathfrak{R}^{\Psi}(f,f)$ as the sum of all "$\partial f\partial f$" terms and denote $2G^{\mathsf{T}}X$ as the sum of all "$\partial^2 f\partial f$" terms. Switching indices for the terms in $2G^{\mathsf{T}}X$ to match

$\frac{\partial^2 f}{\partial x_{\hat{i}}\partial x_{\hat{j}}}$, we get the following

$$
\begin{aligned}
& 2G^\top X \\
= \; & 2\sum_{k=1}^{m}\sum_{i=1}^{n}\sum_{k',\hat{k},\hat{i},i'=1}^{n+m}\left[z^\top_{kk'}z^\top_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\top_{i\hat{i}}\frac{\partial^2 f}{\partial x_{k'}\partial x_{\hat{i}}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}+z^\top_{kk'}z^\top_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}a^\top_{i\hat{i}}\frac{\partial f}{\partial x_{\hat{i}}}a^\top_{ii'}\frac{\partial^2 f}{\partial x_{k'}\partial x_{i'}}\right] \\
& -2\sum_{j=1}^{m}\sum_{l=1}^{n}\sum_{l',\hat{l},\hat{j},j'=1}^{n+m}\left[a^\top_{ll'}a^\top_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^\top_{j\hat{j}}\frac{\partial^2 f}{\partial x_{l'}\partial x_{\hat{j}}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}+a^\top_{ll'}a^\top_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}z^\top_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^\top_{jj'}\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^\top_{jj'}z^\top_{j\hat{j}}\frac{\partial}{\partial x_{\hat{j}}}a^\top_{i\hat{i}}\frac{\partial^2 f}{\partial x_{j'}\partial x_{\hat{i}}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}+z^\top_{jj'}z^\top_{j\hat{j}}\frac{\partial}{\partial x_{\hat{j}}}a^\top_{i\hat{i}}\frac{\partial f}{\partial x_{\hat{i}}}a^\top_{ii'}\frac{\partial^2 f}{\partial x_{j'}\partial x_{i'}}\right] \\
& -2\sum_{j=1}^{m}\sum_{i=1}^{n}\sum_{i',\hat{i},\hat{j},j'=1}^{n+m}\left[a^\top_{ii'}a^\top_{i\hat{i}}\frac{\partial}{\partial x_{\hat{i}}}z^\top_{j\hat{j}}\frac{\partial^2 f}{\partial x_{i'}\partial x_{\hat{j}}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}+a^\top_{ii'}a^\top_{i\hat{i}}\frac{\partial}{\partial x_{\hat{i}}}z^\top_{j\hat{j}}\frac{\partial f}{\partial x_{\hat{j}}}z^\top_{jj'}\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^\top_{j\hat{j}}z^\top_{jj'}\frac{\partial}{\partial x_{j'}}a^\top_{i\hat{i}}\frac{\partial^2 f}{\partial x_{\hat{j}}\partial x_{\hat{i}}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}+z^\top_{j\hat{j}}z^\top_{jj'}\frac{\partial}{\partial x_{j'}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}a^\top_{i\hat{i}}\frac{\partial^2 f}{\partial x_{\hat{j}}\partial x_{\hat{i}}}\right] \\
& -2\sum_{j=1}^{m}\sum_{i=1}^{n}\sum_{i',\hat{i},\hat{j},j'=1}^{n+m}\left[a^\top_{i\hat{i}}a^\top_{ii'}\frac{\partial}{\partial x_{i'}}z^\top_{j\hat{j}}\frac{\partial^2 f}{\partial x_{\hat{i}}\partial x_{\hat{j}}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}+a^\top_{i\hat{i}}a^\top_{ii'}\frac{\partial}{\partial x_{i'}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}z^\top_{j\hat{j}}\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},i',\hat{i}=1}^{n+m}\left[\left(z^\top_{j\hat{j}}z^\top_{jj'}\frac{\partial}{\partial x_{j'}}a^\top_{i\hat{i}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}+z^\top_{j\hat{j}}z^\top_{jj'}\frac{\partial}{\partial x_{j'}}a^\top_{ii'}\frac{\partial f}{\partial x_{i'}}a^\top_{i\hat{i}}\right)\right.\right. \\
& \left.\left. -\left(a^\top_{i\hat{i}}a^\top_{ii'}\frac{\partial}{\partial x_{i'}}z^\top_{j\hat{j}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}+a^\top_{i\hat{i}}a^\top_{ii'}\frac{\partial}{\partial x_{i'}}z^\top_{jj'}\frac{\partial f}{\partial x_{j'}}z^\top_{j\hat{j}}\right)\right]\right].
\end{aligned}
$$

The first equality follows from the quantities we obtain previously, the second equality follows from switching "k" to "j" and "l" to "i", the third equality follows from switching between "i'" and "$\hat{i}$", "j'" and "$\hat{j}$". Thus the proof is completed. ■

With the above lemma in hand, we are now ready to present our most important theorem in the paper. By using this theorem, we are able to provide the most general curvature dimension inequality in this paper, and prove gradient estimates and energy dissipation rate in the most general setting of the paper.

Theorem 3.32. *If Assumption 3.30 is satisfied, then*

$$
\Gamma_2+\Gamma_2^{z,\Psi}(f,f) \quad = \quad |\mathfrak{Hess}^G_{a,z}f|^2+\mathfrak{R}^G_a(\nabla f,\nabla f)+\mathfrak{R}_z(\nabla f,\nabla f)+\mathfrak{R}^\Psi(\nabla f,\nabla f), \tag{3.74}
$$

where $\mathfrak{R}^\Psi(\nabla f,\nabla f)$ is defined in Notation 3.29, and we denote

$$
|\mathfrak{Hess}^G_{a,z}f|^2=[X+\widehat{\Lambda}+\Theta](Q^\top Q+P^\top P)[X+\widehat{\Lambda}+\Theta], \tag{3.75}
$$

$$
\begin{aligned}
\mathfrak{R}_a^G(\nabla f,\nabla f) &= -[\widehat{\Lambda}+\Theta]^{\mathsf{T}}((Q^{\mathsf{T}}Q+P^{\mathsf{T}}P))[\widehat{\Lambda}+\Theta]+D^{\mathsf{T}}D+E^{\mathsf{T}}E\\
&+\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a_{ii'}^{\mathsf{T}}(\frac{\partial a_{i\hat{\imath}}^{\mathsf{T}}}{\partial x_{i'}}\frac{\partial a_{k\hat{k}}^{\mathsf{T}}}{\partial x_{\hat{\imath}}}\frac{\partial f}{\partial x_{\hat{k}}}),(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&+\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a_{ii'}^{\mathsf{T}}a_{i\hat{\imath}}^{\mathsf{T}}(\frac{\partial}{\partial x_{i'}}\frac{\partial a_{k\hat{k}}^{\mathsf{T}}}{\partial x_{\hat{\imath}}})(\frac{\partial f}{\partial x_{\hat{k}}}),(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&-\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a_{k\hat{k}}^{\mathsf{T}}\frac{\partial a_{ii'}^{\mathsf{T}}}{\partial x_{\hat{k}}}\frac{\partial a_{i\hat{\imath}}^{\mathsf{T}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{\imath}}}),(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n}\\
&-\sum_{i,k=1}^{n}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a_{k\hat{k}}^{\mathsf{T}}a_{ii'}^{\mathsf{T}}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial a_{i\hat{\imath}}^{\mathsf{T}}}{\partial x_{i'}})\frac{\partial f}{\partial x_{\hat{\imath}}},(a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n},
\end{aligned}
$$

and

$$
\begin{aligned}
\mathfrak{R}_z(\nabla f,\nabla f) &= \sum_{i=1}^{n}\sum_{k=}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a_{ii'}^{\mathsf{T}}(\frac{\partial a_{i\hat{\imath}}^{\mathsf{T}}}{\partial x_{i'}}\frac{\partial z_{k\hat{k}}^{\mathsf{T}}}{\partial x_{\hat{\imath}}}\frac{\partial f}{\partial x_{\hat{k}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\\
&+\sum_{i=1}^{n}\sum_{k=}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle a_{ii'}^{\mathsf{T}}a_{i\hat{\imath}}^{\mathsf{T}}(\frac{\partial}{\partial x_{i'}}\frac{\partial z_{k\hat{k}}^{\mathsf{T}}}{\partial x_{\hat{\imath}}})(\frac{\partial f}{\partial x_{\hat{k}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\\
&-\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle z_{k\hat{k}}^{\mathsf{T}}\frac{\partial a_{ii'}^{\mathsf{T}}}{\partial x_{\hat{k}}}\frac{\partial a_{i\hat{\imath}}^{\mathsf{T}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{\imath}}}),(z^{\mathsf{T}}\nabla)_k f\rangle\\
&-\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{\imath},\hat{k}=1}^{n+m}\langle z_{k\hat{k}}^{\mathsf{T}}a_{ii'}^{\mathsf{T}}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial a_{i\hat{\imath}}^{\mathsf{T}}}{\partial x_{i'}})\frac{\partial f}{\partial x_{\hat{\imath}}},(z^{\mathsf{T}}\nabla)_k f\rangle. \qquad (3.76)
\end{aligned}
$$

Remark 3.33. *For every $\rho > 0$, we can also prove a variation of the generalized Gamma 2 z formula for the following quantity,*

$$\Gamma_2(f,f)+\rho\Gamma_2^{z,\Psi}(f,f).$$

The proof is similar to the proof of $\Gamma_2(f,f)+\Gamma_2^{z,\Psi}$ below, the difference is the change of coefficient in the quadratic form. We skip the proof for this version.

Proof Recall the proof of Theorem 3.25, we first denote

$$\widetilde{\mathfrak{R}}_a(\nabla f,\nabla f) = \mathfrak{R}_a(\nabla f,\nabla f)+[\widetilde{\Lambda}+\Theta]^{\mathsf{T}}((Q^{\mathsf{T}}Q+P^{\mathsf{T}}P))[\widetilde{\Lambda}+\Theta]-D^{\mathsf{T}}D-E^{\mathsf{T}}E.$$

Then combine the proof of Theorem 3.25 and Lemma 4.11, we get

$$\begin{aligned}
& \Gamma_2 + \Gamma_2^{z,\Psi}(f,f) \\
= \; & \Gamma_2 + \Gamma_2^{z}(f,f) - \mathbf{div}_a(\Gamma_{\nabla(zz^\mathsf{T})} f, f) + \mathbf{div}_z(\Gamma_{\nabla(aa^\mathsf{T})} f, f) \\
= \; & X^\mathsf{T} P^T P X + 2E^T P X + 2F^T X + E^T E \\
& + X^T Q^T Q X + 2D^T Q X + 2C^T X + D^T D + 2G^\mathsf{T} X \\
& + \widetilde{\mathfrak{R}}_a(\nabla f, \nabla f) + \mathfrak{R}_z(\nabla f, \nabla f) + \mathfrak{R}^\Psi(\nabla f, \nabla f) \\
= \; & X^T[P^\mathsf{T} P + Q^\mathsf{T} Q]X + 2[F^\mathsf{T} + C^\mathsf{T} + G^\mathsf{T}]X + 2[E^\mathsf{T} P + D^\mathsf{T} Q]X + D^\mathsf{T} D + E^\mathsf{T} E \\
& + \widetilde{\mathfrak{R}}_a(\nabla f, \nabla f) + \mathfrak{R}_z(\nabla f, \nabla f) + \mathfrak{R}^\Psi(\nabla f, \nabla f).
\end{aligned}$$

Assuming that Assumption 3.30 is satisfied, we get

$$\begin{aligned}
& \Gamma_2 + \Gamma_2^{z,\Psi}(f,f) \\
= \; & [X + \widehat{\Lambda} + \Theta](Q^\mathsf{T} Q + P^\mathsf{T} P)[X + \widehat{\Lambda} + \Theta] \\
& - [\widehat{\Lambda} + \Theta]^\mathsf{T}((Q^\mathsf{T} Q + P^\mathsf{T} P))[\widehat{\Lambda} + \Theta] + D^\mathsf{T} D + E^\mathsf{T} E \\
& + \widetilde{\mathfrak{R}}_a(\nabla f, \nabla f) + \mathfrak{R}_z(\nabla f, \nabla f) + \mathfrak{R}^\Psi(\nabla f, \nabla f) \\
= \; & [X + \widehat{\Lambda} + \Theta](Q^\mathsf{T} Q + P^\mathsf{T} P)[X + \widehat{\Lambda} + \Theta] \\
& + \mathfrak{R}_a^G(\nabla f, \nabla f) + \mathfrak{R}_z(\nabla f, \nabla f) + \mathfrak{R}^\Psi(\nabla f, \nabla f),
\end{aligned}$$

which finishes the proof. ■

3.3.1 Generalized Gamma z calculus with drift

In this subsection, we introduce drift to the degenerate system and derive the generalized Gamma z calculus for drift-diffusion operator $\tilde{L}$. We first introduce the following definition.

Definition 3.34. *We define generalized Gamma 2 z for operator $\tilde{L}$ below,*

$$\Gamma_{2,\tilde{L}}^{z,\Psi_b}(f,f) \quad = \quad \Gamma_{2,\tilde{L}}^{z}(f,f) + \mathbf{div}_z^{\Psi_b}(\Gamma_{\nabla(aa^\mathsf{T})} f, f) - \mathbf{div}_a^{\Psi_b}(\Gamma_{\nabla(zz^\mathsf{T})} f, f), \quad (3.77)$$

where $\mathbf{div}_z^{\Psi_b}(\Gamma_{\nabla(aa^\mathsf{T})}(f,f))$ and $\mathbf{div}_a^{\Psi_b}(\Gamma_{\nabla(zz^\mathsf{T})}(f,f))$ are define in Definition 3.27. Here we denote Ψ_b to emphasize that Ψ depends on the drift b and the function Ψ_b is the transition kernel for operator $\tilde{L}$.

We then have the following Theorem.

Theorem 3.35. *If Assumption 3.30 is satisfied, then*

$$\Gamma_{2,\tilde{L}} + \Gamma^{z,\Psi_b}_{2,\tilde{L}}(f,f) = |\mathfrak{hess}^G_{a,z} f|^2 + \mathfrak{R}^G_{a,b}(\nabla f, \nabla f) + \mathfrak{R}_{z,b}(\nabla f, \nabla f) + \mathfrak{R}^{\Psi_b}_b(\nabla f, \nabla f), \tag{3.78}$$

where we denote

$$|\mathfrak{hess}^G_{a,z} f|^2 = [X + \widehat{\Lambda} + \Theta](Q^{\mathsf{T}}Q + P^{\mathsf{T}}P)[X + \widehat{\Lambda} + \Theta], \tag{3.79}$$

$$\begin{aligned}
\mathfrak{R}^G_{a,b}(\nabla f, \nabla f) \quad = \quad & -[\widehat{\Lambda} + \Theta]^{\mathsf{T}}((Q^{\mathsf{T}}Q + P^{\mathsf{T}}P))[\widehat{\Lambda} + \Theta] + D^{\mathsf{T}}D + E^{\mathsf{T}}E \\
& + \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{ii'} (\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}} \frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
& + \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{ii'} a^{\mathsf{T}}_{i\hat{i}} (\frac{\partial}{\partial x_{i'}} \frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
& - \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{k\hat{k}} \frac{\partial a^{\mathsf{T}}_{ii'}}{\partial x_{\hat{k}}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
& - \sum_{i,k=1}^{n} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{k\hat{k}} a^{\mathsf{T}}_{ii'} (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}) \frac{\partial f}{\partial x_{\hat{i}}}, (a^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
& -2 \sum_{i=1}^{n} \sum_{\hat{i},\hat{k}=1}^{n+m} \langle (a^{T}_{i\hat{i}} \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial a^{T}_{i\hat{i}}}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (a^{T}\nabla f)_i\rangle_{\mathbb{R}^n},
\end{aligned}$$

and

$$\begin{aligned}
\mathfrak{R}_{z,b}(\nabla f, \nabla f) \quad = \quad & \sum_{i=1}^{n} \sum_{k=}^{m} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{ii'} (\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}} \frac{\partial z^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \\
& + \sum_{i=1}^{n} \sum_{k=}^{m} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle a^{\mathsf{T}}_{ii'} a^{\mathsf{T}}_{i\hat{i}} (\frac{\partial}{\partial x_{i'}} \frac{\partial z^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \\
& - \sum_{i=1}^{n} \sum_{k=1}^{m} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle z^{\mathsf{T}}_{k\hat{k}} \frac{\partial a^{\mathsf{T}}_{ii'}}{\partial x_{\hat{k}}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^{\mathsf{T}}\nabla)_k f\rangle \\
& - \sum_{i=1}^{n} \sum_{k=1}^{m} \sum_{i',\hat{i},\hat{k}=1}^{n+m} \langle z^{\mathsf{T}}_{k\hat{k}} a^{\mathsf{T}}_{ii'} (\frac{\partial}{\partial x_{\hat{k}}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}) \frac{\partial f}{\partial x_{\hat{i}}}, (z^{\mathsf{T}}\nabla)_k f\rangle \\
& -2 \sum_{i=1}^{m} \sum_{\hat{i},\hat{k}=1}^{n+m} \langle (z^{T}_{i\hat{i}} \frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}} \frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}} \frac{\partial z^{T}_{i\hat{i}}}{\partial x_{\hat{k}}} \frac{\partial f}{\partial x_{\hat{i}}}), (z^{T}\nabla f)_i\rangle_{\mathbb{R}^m}.
\end{aligned}$$

In particular, the tensor $\mathfrak{R}^{\Phi_b}$ *is defined in* (3.70) *with* $\Psi = \Psi_b$.

Proof By Definition 3.34, Theorem 3.16 and Theorem 3.26, we have

$$
\begin{aligned}
& \Gamma_{2,\tilde{L}}(f,f) + \tilde{\Gamma}_{2,\tilde{L}}^{z,\Psi}(f,f) \\
= \; & \Gamma_{2,\tilde{L}}(f,f) + \Gamma_{2,\tilde{L}}^{z}(f,f) + \mathbf{div}_z^{\Psi_b}(\Gamma_{\nabla(aa^{\mathsf{T}})}f,f) - \mathbf{div}_a^{\Psi_b}(\Gamma_{\nabla(zz^{\mathsf{T}})}f,f) \\
= \; & \Gamma_2 + \Gamma_2^{z,\Psi}(f,f) + [b\nabla\Gamma_1(f,f) - \Gamma_1(2b\nabla f,f)] + [b\nabla\Gamma_1^z(f,f) - \Gamma_1^z(2b\nabla f,f)].
\end{aligned}
$$

The extra term $[b\nabla\Gamma_1(f,f) - \Gamma_1(2b\nabla f,f)] + [b\nabla\Gamma_1^z(f,f) - \Gamma_1^z(2b\nabla f,f)]$ will only contribute to the tensor terms, combine with Theorem 3.32, the proof follows. ■

3.4 Examples

In this section, we present several examples of our generalized Gamma z calculus in different situations. We first show that our methodology works for standard Riemannian setting with demonstration on conformal mapping. We then show that our generalized curvature dimension inequality can provide convergence bound for degenerate drift-diffusion SDEs. We compute explicit bounds for constant degenerate coefficient matrix and show that the required assumption is satisfied for non-constant degenerate coefficient matrix. In the end, we compute the curvature dimension bound for Grushin plane with step $k+1$, for $k \geq 1$. This is a special case where the vector C (see 3.21) is non-zero and Assumption 3.3 is satisfied for $x \neq 0$ on Grushin plane. In order to get the desired bound in the degenerate case with $x = 0$, we need to introduce matrix z. It is easy to see that the required assumption (3.22 or 3.30) is satisfied once the matrix z is introduced. Due to the length of the paper, we leave the detailed analysis on Grushin plane for future studies.

3.4.1 Example of conformal mapping

We look at the example of conformal mapping with $\tilde{g} = e^{2\varphi}g$. Take $\tilde{g} = (aa^{\mathsf{T}})^{-1}\mathrm{Id}_{n\times n}$, with $g = \mathrm{Id}_{n\times n}$. We have $det(\tilde{g}) = a^{-2n}$,so the volume is $\mathbf{Vol}(\tilde{g}) = a^{-n}$.

$$
a = e^{-\varphi}, \quad (aa^{\mathsf{T}})^{-1}\mathrm{Id}_{n\times n} = e^{2\varphi}\begin{pmatrix} 1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 1 \end{pmatrix}_{n\times n},
$$

$$
\tilde{g}^{-1} = a\begin{pmatrix} 1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 1 \end{pmatrix}_{n\times n} a\begin{pmatrix} 1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 1 \end{pmatrix}_{n\times n}^{T} = \begin{pmatrix} a^2 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & a^2 \end{pmatrix}_{n\times n}.
$$

We set $\rho = \sqrt{|\tilde{g}|}P$, we write the heat equation for P with Laplace-Beltrami operator for our metric $\tilde{g}$,

$$2\partial_t P \quad = \tilde{\Delta} P, \quad \text{where} \quad \tilde{\Delta} = \frac{\nabla\cdot(\tilde{g}^{-1}\sqrt{|\tilde{g}|}\nabla)}{\sqrt{|\tilde{g}|}}.$$

By changing coordinates, we get

$$2\partial_t(\frac{\rho}{\sqrt{|\tilde{g}|}}) = \frac{\nabla\cdot(\tilde{g}^{-1}\sqrt{|\tilde{g}|}\nabla\frac{\rho}{\sqrt{|\tilde{g}|}})}{\sqrt{|\tilde{g}|}},$$

thus it is directly to check that

$$\begin{aligned}
2\partial_t\rho &= \mathcal{L}^*\rho \\
\mathcal{L}^* &= \nabla\cdot(\rho a^2 \mathrm{Id}_{n\times n}\nabla\log\frac{\rho}{\sqrt{|\tilde{g}|}}) \\
&= \nabla\cdot(\rho a^2\frac{\sqrt{|\tilde{g}|}}{\rho}\nabla(\frac{\rho}{\sqrt{|\tilde{g}|}})) \\
&= \nabla\cdot(a^2a^{-n}(\frac{\nabla\rho}{a^{-n}} + na^{n-1}\rho\nabla a)) \\
&= \nabla\cdot(a^2\nabla\rho) + n\nabla\cdot(\rho a\nabla a).
\end{aligned}$$

We know that $\tilde{\Delta}$ is the Laplacian-Beltrami operator for the volume with metric $\tilde{g}$. We can write the Bochner's formula for $\tilde{g}$. Then by using the conformal mappping, we can rewrite the Bochner's formula in terms of a, i.e. change of coordinates from P to ρ. Then the new Bochner's formula in the coordinates in terms of ρ will provide the same type of formulas as what we proved in our Theorem 3.4 and Theorem 3.5 with a special choice of a. The operator is

$$\mathcal{L} \quad = \quad \nabla\cdot(a^2\nabla f) - na\nabla a\cdot\nabla f.$$

The corresponding Γ^2 operator is defined as

$$\Gamma_{2,\mathcal{L}}(f,f) = \frac{1}{2}\mathcal{L}(a^2\langle\nabla f,\nabla f\rangle) - \langle\nabla\mathcal{L}f,\nabla f\rangle a^2.$$

By direct computation, we get

$$\begin{aligned}
&\mathcal{L}(a^2\langle\nabla f,\nabla f\rangle) \\
= \;& \nabla\cdot(a^2\nabla(a^2\langle\nabla f,\nabla f\rangle)) - n(a\nabla a)\nabla(a^2\langle\nabla f,\nabla f\rangle) \\
= \;& \nabla\cdot(a^2\nabla|\nabla f|^2a^2 + a^2|\nabla f|^2\nabla(a^2)) - \frac{n}{2}\nabla a^2(\nabla|\nabla f|^2a^2 + |\nabla f|^2\nabla a^2) \\
= \;& 2\langle\nabla a^2,\nabla|\nabla f|^2\rangle a^2 + a^2\Delta|\nabla f|^2a^2 + \langle\nabla a^2,\nabla a^2\rangle|\nabla f|^2 + a^2\langle\nabla|\nabla f|^2,\nabla a^2\rangle \\
& + a^2|\nabla f|^2\Delta a^2 - \frac{n}{2}a^2\langle\nabla a^2,\nabla|\nabla f|^2\rangle - \frac{n}{2}\langle\nabla a^2,\nabla a^2\rangle|\nabla f|^2 \\
= \;& (3-\frac{n}{2})\langle\nabla a^2,\nabla|\nabla f|^2\rangle a^2 + (1-\frac{n}{2})\langle\nabla a^2,\nabla a^2\rangle|\nabla f|^2 \\
& + a^2\Delta|\nabla f|^2a^2 + a^2\Delta a^2|\nabla f|^2,
\end{aligned}$$

and

$$
\begin{aligned}
\langle \nabla \mathcal{L} f, \nabla f\rangle a^2 &= \langle \nabla(\nabla\cdot(a^2\nabla f) - \frac{n}{2}\nabla a^2 \nabla f), \nabla f\rangle a^2 \\
\nabla f &= a^2\langle \nabla\Delta f, \nabla f\rangle a^2 + \langle \nabla a^2, \nabla f\rangle \Delta f a^2 + (1-\frac{n}{2})\langle \nabla^2 a^2 \nabla f, \nabla f\rangle a^2 \\
&\quad +(1-\frac{n}{2})\langle \nabla^2 f \nabla a^2, \nabla f\rangle a^2.
\end{aligned}
$$

Now, combing the above two terms, we get

$$
\begin{aligned}
&\frac{1}{2}\mathcal{L}(a^2\langle \nabla f, \nabla f\rangle) - \langle \nabla\mathcal{L} f, \nabla f\rangle a^2 \\
= \;& a^2(\frac{1}{2}\Delta|\nabla f|^2 - \langle \nabla \Delta f, \nabla f\rangle)a^2 + \frac{a^2}{2}|\nabla f|^2 \Delta a^2 - \langle \nabla a^2, \nabla f\rangle a^2 \Delta f \\
&+2\langle \nabla a^2, \nabla^2 f \nabla f\rangle a^2 + (\frac{1}{2}-\frac{n}{4})\langle \nabla a^2, \nabla a^2\rangle |\nabla f|^2 - (1-\frac{n}{2})\langle \nabla^2 a^2 \nabla f, \nabla f\rangle a^2 \\
= \;& a^4|Hessf|^2 + 2\langle \nabla a^2, \nabla^2 f\nabla f\rangle a^2 - \langle \nabla a^2, \nabla f\rangle a^2 \Delta f \\
&+\frac{a^2}{2}\Delta a^2 |\nabla f|^2 + (\frac{1}{2}-\frac{n}{4})\langle \nabla a^2, \nabla a^2\rangle |\nabla f|^2 - (1-\frac{n}{2})\langle \nabla^2 a^2 \nabla f, \nabla f\rangle a^2.
\end{aligned}
$$

The goal now is to find the new Hessian square for metric $\tilde{g}$. In order to demonstrate the example clearly and neatly, without lose of generality, we assume the dimension $n = 2$ below. Straight computation shows that,

$$
\begin{aligned}
a^4|Hessf|^2 &= a^4(f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2)\cdots \mathcal{I}_1 \\
\langle \nabla a^2, \nabla^2 f \nabla f\rangle a^2 &= 2a^3\Big\langle \begin{pmatrix} a_x \\ a_y \end{pmatrix}, \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy}\end{pmatrix}\begin{pmatrix} f_x \\ f_y\end{pmatrix}\Big\rangle \\
&= 2a^3(f_{xx}a_x f_x + f_{xy}a_x f_y + f_{yx}a_y f_x + f_{yy}a_y f_y)\cdots \mathcal{I}_2 \\
\Big\langle \nabla a^2, \nabla f\Big\rangle a^2 \Delta f &= 2a^3(a_x f_x + a_y f_y)(f_{xx}+f_{yy}), \\
&= 2a^3(f_{xx}a_x f_x + f_{xx}a_y f_y + f_{yy}a_x f_x + f_{yy}a_y f_y)\cdots \mathcal{I}_3.
\end{aligned}
$$

Combing the above three terms, we end up with

$$
\begin{aligned}
\mathcal{I}_1 + 2\mathcal{I}_2 - \mathcal{I}_3 &= a^4 f_{xx}^2 + 2a^3 f_{xx}(a_x f_x - a_y f_y) \\
&\quad + a^4 f_{yy}^2 + 2a^3(f_{yy}a_y f_y - a_x f_x) \\
&\quad + 2a^4 f_{xy}^2 + 4a^3 f_{xy}(a_x f_y + a_y f_x) \\
&= [a^2 f_{xx} + \frac{1}{2}(\nabla_x a^2 f_x - \nabla_y a^2 f_y)]^2 - \frac{1}{4}(\nabla_x a^2 f_x - \nabla_y a^2 f_y)^2 \\
&\quad +[a^2 f_{yy} + \frac{1}{2}(\nabla_y a^2 f_y - \nabla_x a^2 f_x)]^2 - \frac{1}{4}(\nabla_y a^2 f_y - \nabla_x a^2 f_x)^2 \\
&\quad +2[a^2 f_{xy} + \frac{1}{2}(\nabla_x a^2 f_y + \nabla_y a^2 f_x)]^2 - \frac{1}{2}(\nabla_x a^2 f_y + \nabla_y a^2 f_x)^2 x.
\end{aligned}
$$

Where $[a^2 f_{xx} + \frac{1}{2}(\nabla_x a^2 f_x - \nabla_y a^2 f_y)]^2 + [a^2 f_{yy} + \frac{1}{2}(\nabla_y a^2 f_y - \nabla_x a^2 f_x)]^2 + 2[a^2 f_{xy} + \frac{1}{2}(\nabla_x a^2 f_y + \nabla_y a^2 f_x)]^2$ is our new $\mathfrak{Hess}_a(f, f)$ here.

3.4.2 Example of drift-diffusion process

In this part, we present a drift-diffusion process in $\mathbb{R}^2$.

$$d\begin{pmatrix} x(t) \\ y(y) \end{pmatrix} = \begin{pmatrix} b_1(x,y) \\ b_2(x,y) \end{pmatrix} dt + \begin{pmatrix} a_{11}(x,y) & a_{12}(x,y) \\ a_{21}(x,y) & a_{22}(x,y) \end{pmatrix} \circ \begin{pmatrix} dB_t^1 \\ dB_t^2 \end{pmatrix}. \tag{3.80}$$

If we denote $X = (x(t), y(t))$, then we have $dX_t = b(X)dt + \sum_{i=1}^{2} a_i(X_t) \circ dB_t^i$. Here we denote

$$a = (a_1, a_2) = \begin{pmatrix} a_{11}(x,y) & a_{12}(x,y) \\ a_{21}(x,y) & a_{22}(x,y) \end{pmatrix}, b = \begin{pmatrix} b_1(x,y) \\ b_2(x,y) \end{pmatrix}.$$

We assume that vector fields a_1, a_2, b satisfies bracket generating condition and there exists unique invariant measure of the above system. Below, we look at a special case of matrix a and see how to apply Theorem 3.16, Theorem 3.26 and Theorem 3.35 under different circumstances.

We first look at Theorem 3.16. As we mentioned above (see Proposition 3.6 and Remark 3.14), we compute vector C first, if C is a zero vector, then we are reduced to a simplified case. Below, we start to look at square matrix a with size 2×2. Using the local coordinates in Notation 3.2, for $\hat{i} = 1,2$ and $\hat{k} = 1,2$, we have

$$\begin{aligned} C_{\hat{i}\hat{k}} &= \left[\sum_{i,k=1}^{2} \sum_{i'=1}^{2} \left(\langle a^{\mathsf{T}}_{i\hat{i}} a^{\mathsf{T}}_{ii'} (\frac{\partial a^{\mathsf{T}}_{k\hat{k}}}{\partial x_{i'}}), (a^{\mathsf{T}}\nabla)_k f \rangle_{\mathbb{R}^n} - \langle a^{\mathsf{T}}_{ki'} a^{\mathsf{T}}_{i\hat{k}} \frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}, (a^{\mathsf{T}}\nabla)_k f \rangle_{\mathbb{R}^n} \right) \right], \\ (a^{\mathsf{T}}\nabla)_1 f &= a^{\mathsf{T}}_{11}\partial_x f + a^{\mathsf{T}}_{12}\partial_y f, \quad (a^{\mathsf{T}}\nabla)_2 f = a^{\mathsf{T}}_{21}\partial_x f + a^{\mathsf{T}}_{22}\partial_y f. \end{aligned}$$

Degenerate constant matrix:

We first present our theorem in a simple example with constant degenerate matrix a, and precise vector b, in which there exists unique invariant measure. We have $dX_t = b(X_t)dt + a_2 dB_t$ with

$$X = (x,v), \quad b(x,v) = \begin{pmatrix} v \\ -\gamma uv - \nabla U(x) \end{pmatrix}, \quad a = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2\gamma u} \end{pmatrix}, \quad \text{we see it as} \begin{pmatrix} 0 \\ \sqrt{2\gamma u} \end{pmatrix},$$

where γ, u are constant parameters. According to Theorem 3.6, it is easy to check that vector C (see Notation (3.21)) is a zero vector and curvature tensor $\mathfrak{R}_a(f,f) = 0$. We only need to compute drift term of $\mathfrak{R}_{a,b}(\nabla f, \nabla f)$.

$$\begin{aligned}
\mathfrak{R}_{a,b}(\nabla f, \nabla f) &= -2\sum_{i=1}^{2}\sum_{\hat{i},\hat{k}=1}^{2}\langle (a^T_{i\hat{i}}\frac{\partial b_{\hat{k}}}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}} - b_{\hat{k}}\frac{\partial a^T_{i\hat{i}}}{\partial x_{\hat{k}}}\frac{\partial f}{\partial x_{\hat{i}}}), (a^T\nabla f)_i\rangle_{\mathbb{R}^2} \\
&= 8\gamma^2 u^2(\frac{\partial f}{\partial v})^2 - 8\gamma u\frac{\partial f}{\partial x}\frac{\partial f}{\partial v} \\
&= 8\gamma u\left(-\frac{\partial f}{\partial x}\frac{\partial f}{\partial v} + \gamma u(\frac{\partial f}{\partial v})^2\right) \\
&\geq -4\gamma u\left((\frac{\partial f}{\partial x})^2 - (2\gamma u - 1)(\frac{\partial f}{\partial v})^2\right) \\
&\geq -4\gamma u(\frac{\partial f}{\partial x})^2 + (2\gamma u - 1)\Gamma_{1,\tilde{L}}(f,f).
\end{aligned}$$

We cannot obtain bounds for the right hand side only by $\Gamma_1(f,f)$. Thus we need to further introduce extra direction to the system by the following matrix z. By routine computation, we get

$$z^T = (z_1(x,v), z_2(x,v)), \quad zz^T = \begin{pmatrix} z_1^2 & z_1z_2 \\ z_1z_2 & z_2^2 \end{pmatrix}, \quad P = (0,0,z_1,z_2).$$

$$P^TP = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & z_1^2 & z_1z_2 \\ 0 & 0 & z_1z_2 & z_2^2 \end{pmatrix}, \quad \text{and} \quad Q^TQ = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\gamma u \end{pmatrix}.$$

Assumption 3.36. *Assume the matrix z is a constant matrix.*

Based on the Assumption 3.36, we know that

$$\mathbf{div}^{\mathbf{Vol}}_z(\Gamma_{1,\nabla(aa^\mathsf{T})}(f,g)) - \mathbf{div}^{\mathbf{Vol}}_a(\Gamma_{1,\nabla(zz^\mathsf{T})}(f,g)) = 0.$$

Thus, Theorem 3.35 is reduced to the form of Theorem 3.26, we are left to verify the assumption of Theorem 3.26 and prove the bound there. It is easy to check that

$$(P^\mathsf{T}P + Q^\mathsf{T}TQ)\widetilde{\Lambda} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & z_1^2 & z_1z_2 \\ 0 & 0 & z_1z_2 & z_2^2 + 2\gamma u \end{pmatrix}\widetilde{\Lambda} = F + C = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

together with $E = 0$, and $D = 0$.

According to Theorem 3.26 and routine computations, we get

$$
\begin{aligned}
&\Gamma_{2,\tilde{L}}(f,f)+\Gamma^{z}_{2,\tilde{L}}\\
= \ & X^{\mathsf{T}}[Q^{\mathsf{T}}Q+P^{\mathsf{T}}P]X+8\gamma u\left(-\frac{\partial f}{\partial x}\frac{\partial f}{\partial v}+\gamma u(\frac{\partial f}{\partial v})^2\right)-4z_1z_2(\frac{\partial f}{\partial x})^2\\
&+[4\gamma u z_2^2+4z_1z_2\nabla^2U(x)](\frac{\partial f}{\partial v})^2+4[-z_2^2+\gamma u z_1 z_2+z_1^2\nabla^2U(x)]\frac{\partial f}{\partial x}\frac{\partial f}{\partial v}\\
\geq \ & [8\gamma^2u^2+4\gamma u z_2^2+4z_1z_2\nabla^2U(x)](\frac{\partial f}{\partial v})^2-4z_1z_2(\frac{\partial f}{\partial x})^2\\
&+4[-z_2^2-2\gamma u+\gamma u z_1z_2+z_1^2\nabla^2U(x)]\frac{\partial f}{\partial x}\frac{\partial f}{\partial v}.\\
\geq \ & C_1(z,U,u,\gamma)(\frac{\partial f}{\partial v})^2-C_2(z,U,u,\gamma)(\frac{\partial f}{\partial x})^2+C_3(z,U,u,\gamma)\frac{\partial f}{\partial x}\frac{\partial f}{\partial v}\\
\geq \ & C_1(z,U,u,\gamma)(\frac{\partial f}{\partial v})^2-C_2(z,U,u,\gamma)(\frac{\partial f}{\partial x})^2+C_3(z,U,u,\gamma)\frac{\partial f}{\partial x}\frac{\partial f}{\partial v}
\end{aligned}
$$

where we denote

$$
\begin{aligned}
C_1(z,U,u,\gamma) &= [8\gamma^2u^2+4\gamma u z_2^2+4z_1z_2\nabla^2U(x)],\\
C_2(z,U,u,\gamma) &= 4z_1z_2,\\
C_3(z,U,u,\gamma) &= 4[-z_2^2-2\gamma u+\gamma u z_1 z_2+z_1^2\nabla^2U(x)].
\end{aligned}
$$

Furthermore, we have $\Gamma_1(f,f)=2\gamma u(\frac{\partial f}{\partial v})^2$ and $\Gamma_1^z(f,f)=[z_1\frac{\partial f}{\partial x}+z_2\frac{\partial f}{\partial v}]^2$.

$$
\begin{aligned}
\Gamma_{2,\tilde{L}}(f,f)+\Gamma^{z}_{2,\tilde{L}} \ \geq \ & C_1(z,U,u,\gamma)(\frac{\partial f}{\partial v})^2-C_2(z,U,u,\gamma)(\frac{\partial f}{\partial x})^2+\frac{C_3(z,U,u,\gamma)}{2z_1z_2}[z_1\frac{\partial f}{\partial x}+z_2\frac{\partial f}{\partial v}]^2\\
&-\frac{C_3(z,U,u,\gamma)}{2z_1z_2}z_1^2(\frac{\partial f}{\partial x})^2-\frac{C_3(z,U,u,\gamma)}{2z_1z_2}z_2^2(\frac{\partial f}{\partial v})^2\\
\geq \ & \frac{C_3(z,U,u,\gamma)}{2z_1z_2}\Gamma_1^z(f,f)+\Big[C_1(z,U,u,\gamma)-\frac{C_3(z,U,u,\gamma)z_2}{2z_1}\Big]\Gamma_1(f,f)\\
&-\Big[C_2(z,U,u,\gamma)+\frac{C_3(z,U,u,\gamma)z_1}{2z_2}\Big](\frac{\partial f}{\partial x})^2.
\end{aligned}
$$

By choosing z_1,z_2 such that $-\Big[C_2(z,U,u,\gamma)+\frac{C_3(z,U,u,\gamma)z_1}{2z_2}\Big]\geq 0$, we are able to prove a generalized curvature dimension inequality for the generator L of this constant drift-diffusion process. Thus, we can also obtain the related functional inequalities.

Degenerate non-constant matrix

Assumption 3.37. *Assume that*

$$
a^{\mathsf{T}}=\begin{pmatrix}0 & \sqrt{2u(x)}\end{pmatrix},\quad b(x,v)=\begin{pmatrix} v\\ -u(x)v-\nabla U(x)\end{pmatrix},\quad z^T=(z_1(x,v),z_2(x,v)).
$$

The matrices zz^{T}, $P^{\mathsf{T}}P$, $Q^{\mathsf{T}}Q$ are of the same form as before with $z_1(x,v), z_1(x,v)$ and $u(x)$ as functions. The **div** terms are not zero anymore, we follow Theorem 3.35

to get our desired bound. We first show that Assumption 3.30 is satisfied.

$$
\begin{aligned}
&C=(0,0,0,0)^{\mathsf{T}}, \quad F=\Big(0,0,F_{21},F_{22}\Big)^{\mathsf{T}}, \quad G=\Big(0,0,G_{21},G_{22}\Big)^{\mathsf{T}}\\
&E=\sqrt{2u(x)}(\frac{\partial z_1}{\partial v}\frac{\partial f}{\partial x}+\frac{\partial z_2}{\partial v}\frac{\partial f}{\partial v}), \quad D=0.
\end{aligned}
$$

together with

$$
\begin{aligned}
F_{21} &= 2u(x)\frac{\partial z_1}{\partial v}(z_1\frac{\partial f}{\partial x}+z_2\frac{\partial f}{\partial v}), \quad F_{22}=(-z_1\frac{\partial u}{\partial x}+2u\frac{\partial z_2}{\partial v})(z_1\frac{\partial f}{\partial x}+z_2\frac{\partial f}{\partial v}),\\
G_{21} &= \Big(2z_1^2\frac{\partial u}{\partial x}-2u(\frac{\partial z_1}{\partial v}z_2+z_1\frac{\partial z_2}{\partial v})\Big)\frac{\partial f}{\partial v}-4uz_1\frac{\partial z_1}{\partial v}\frac{\partial f}{\partial x},\\
G_{22} &= \Big(2z_1z_2\frac{\partial u}{\partial x}-4u\frac{\partial z_2}{\partial v}z_2\Big)\frac{\partial f}{\partial v}-2u(\frac{\partial z_2}{\partial v}z_1+\frac{\partial z_1}{\partial v}z_2)\frac{\partial f}{\partial x}.
\end{aligned}
$$

It is easy to verify that (if $\mathrm{rank}(P^{\mathsf{T}}P+Q^{\mathsf{T}}Q)=2$) , there exists $\widehat{\Lambda}$ and Θ, such that

$$
(P^{\mathsf{T}}P+Q^{\mathsf{T}}Q)\widehat{\Lambda}=\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&z_1^2&z_1z_2\\0&0&z_1z_2&z_2^2+2u(x)\end{pmatrix}\widehat{\Lambda}=F+C+G,
$$

and $[P^{\mathsf{T}}P+Q^{\mathsf{T}}Q]\Theta=Q^{\mathsf{T}}D+P^{\mathsf{T}}E$. Thus following Theorem 3.35, we can get the desired generalized curvature dimension bound. Here the bound depends on the function Ψ_b which is the transition kernel for the drift-diffusion process. In particular, the function Ψ_b only contributes to the tensor $\mathfrak{R}^{\psi_b}$ in Theorem 3.35 and all the other terms only depend on matrix a, z and b.

Here the invariant measure exists and has the following form

$$
\mathbf{Vol}=\rho^*(x,v)=\frac{1}{V}e^{-(\frac{v^2}{2}+U(x))}, \quad \text{with} \quad \int\rho^*=1, \tag{3.81}
$$

where V is a normalized constant. We thus can obtain the LSI in Proposition 1.14 for invariant measure **Vol**. We simply apply Theorem 3.35 and taking $\Psi_b=\rho^*(x,v)$ in (3.81) above. By routine computation, we get the following explicit formula for tensor $\mathfrak{R}^G_{a,b}+\mathfrak{R}_{z,b}+\mathfrak{R}^{\mathbf{Vol}}$.

$$
\begin{aligned}
\mathfrak{R}^G_{a,b} &= -[\widehat{\Lambda}+\Theta]^{\mathsf{T}}((Q^{\mathsf{T}}Q+P^{\mathsf{T}}P))[\widehat{\Lambda}+\Theta]+E^{\mathsf{T}}E\\
&\quad +8u^2(x)(\frac{\partial f}{\partial v})^2+4v\nabla u(x)(\frac{\partial f}{\partial v})^2-8u(x)\frac{\partial f}{\partial x}\frac{\partial f}{\partial v},
\end{aligned}
$$

$$\begin{aligned}
\mathfrak{R}_{z,b} &= 2\partial_{vv}z_2u(x)f_v^2z_2 + 2\partial_{vv}z_1u(x)f_xf_vz_2 + 2\partial_{vv}z_1u(x)z_1f_x^2 + 2\partial_{vv}z_2u(x)z_1f_xf_v \\
&\quad -2\left(2(-vu'(x) - U''(x))f_xf_vz_1^2 + 2z_1z_2(-2xu'(x) - U''(x))f_v^2 + 2z_1z_2f_x^2\right) \\
&\quad -2\left(-2(-vu(x) - U'(x))\partial_vz_1z_1f_x^2 - 2u(x)z_1z_2f_xf_y - 2(-2xu(x) - U'(x))f_vf_x\partial_vz_2z_1\right) \\
&\quad -2\left(-2vf_x^2\partial_xz_1z_1 - 2vf_vf_x\partial_xz_2z_1 - 2u(x)z_2^2f_v^2 - 2z_2(-2xu(x) - U'(x))f_v^2\partial_vz_2\right) \\
&\quad -2\left(2z_2^2f_xf_v - 2z_2(-2xu(x) - U'(x))f_vf_x\partial_vz_1 - 2vz_2f_xf_v\partial_xz_1 - 2vz_2f_v^2\partial_xz_2\right),
\end{aligned}$$

and

$$\begin{aligned}
\mathfrak{R}^{\mathbf{Vol}} &= 2z_1u'(x)f_v^2\partial_vz_2 + 4z_1u'(x)f_v^2\partial_xz_1 + 2z_2u'(x)f_V^2\partial_vz_1 + 4z_1^2f_v^2\partial_{xx}\sqrt{u(x)} \\
&\quad +\frac{z_1^2(u'(x)^2f_v^2}{u(x)} - 4vz_1z_2u'(x)f_v^2 - 4z_1^2u'(x)U'(x)f_v^2 \\
&\quad -4[\partial_{vv}z_2u(x)z_2f_v^2 + \partial_{vv}z_2u(x)z_1f_xf_v + \partial_{vv}z_1u(x)z_2f_vf_x + \partial_{vv}z_1u(x)z_1f_x^2] \\
&\quad -4[u(x)f_v^2(\partial_vz_2)^2 + 2u(x)f_vf_x\partial_vz_1\partial_vz_2 + u(x)(\partial_vz_1)^2f_x^2] \\
&\quad +8[vu(x)z_2f_v^2\partial_vz_2 + vu(x)z_2\partial_vz_1f_xf_v + vu(x)z_1\partial_vz_2f_xf_v + vu(x)z_1\partial_vz_1f_x^2].
\end{aligned}$$

Similar to the previous case where $u(x)$ is a constant, we are able to derive a new curvature dimensional lower bound by choosing $u(x)$, $z_1(x,v)$ and $z_2(x,v)$ accordingly. For finding the optimal convergence rate to the equilibrium, we need to solve a optimization problem here. We leave this study for future works.

3.4.3 Example of Grushin plane

We consider here the famous Grushin plane [29]. Consider the following vector fields,

$$X = \frac{\partial}{\partial x}, \quad Y = x^k\frac{\partial}{\partial y}, \quad k \geq 1.$$

The Grushin operator has form $\Delta_k = \frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}x^{2k}\frac{\partial^2}{\partial y^2}$, and it is not elliptic and its dimension varies (see [39]). The vector fields $\{X, Y\}$ satisfies bracket generating condition (see [20]) with step $k+1$. To fit in our model, we denote matrix a and metric g below,

$$a = a^\mathsf{T} = \begin{pmatrix} 1 & 0 \\ 0 & x^k \end{pmatrix}, \quad g = (aa^\mathsf{T})^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & x^{2k} \end{pmatrix}^\dagger.$$

By direct computation, we get

$$C = (0, kx^{2k-1}\frac{\partial f}{\partial y}, 0, -kx^{2k-1}\frac{\partial f}{\partial x})^\mathsf{T}, D = (0, kx^{k-1}\frac{\partial f}{\partial y}, 0, 0)^\mathsf{T}, \quad Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & x^k & 0 & 0 \\ 0 & 0 & x^k & 0 \\ 0 & 0 & 0 & x^{2k} \end{pmatrix}.$$

Now we want to find vector Λ, such that $Q^T Q\Lambda = C$. When $x \neq 0$, we get

$$\Lambda = (0, kx^{-1}\frac{\partial f}{\partial y}, 0, -kx^{-2k-1}\frac{\partial f}{\partial x})^{\mathsf{T}}.$$

Otherwise, if $x = 0$, we get $C = 0$, where Λ is not needed. For $x \neq 0$, we can obtain the curvature tensor below,

$$\begin{aligned}
\mathfrak{R}_a(f,f) &= k(k-1)x^{2k-2}(\frac{\partial f}{\partial y})^2 - 2D^{\mathsf{T}}Q\Lambda - \Lambda^{\mathsf{T}}C \\
&= k(k-1)x^{2k-2}(\frac{\partial f}{\partial y})^2 - 2k^2x^{2k-2}(\frac{\partial f}{\partial y})^2 - k^2x^{2k-2}(\frac{\partial f}{\partial y})^2 - k^2x^{-2}(\frac{\partial f}{\partial x})^2 \\
&= -(2k^2+k)x^{2k-2}(\frac{\partial f}{\partial y})^2 - k^2x^{-2}(\frac{\partial f}{\partial x})^2 \\
&\geq -\frac{(2k^2+k)}{x^2}(x^{2k}(\frac{\partial f}{\partial y})^2 + (\frac{\partial f}{\partial x})^2) \geq -\frac{(2k^2+k)}{x^2}\Gamma_1(f,f).
\end{aligned}$$

Furthermore, we get $|\mathfrak{Hess}_a f|^2 = [QX+D+Q\Lambda]^{\mathsf{T}}[QX+D+Q\Lambda]$, by direct computation,

$$\begin{aligned}
QX + D + Q\Lambda &= \left(\frac{\partial^2 f}{\partial x^2}, x^k\frac{\partial^2 f}{\partial x\partial x} + k(x^{2k-1}+x^{k-1})\frac{\partial f}{\partial y}, x^k\frac{\partial^2 f}{\partial x\partial y}, x^{2k}\frac{\partial^2 f}{\partial y^2}\right)^{\mathsf{T}} \\
|\mathfrak{Hess}_a f|^2 &= (\frac{\partial^2 f}{\partial x^2})^2 + [x^k\frac{\partial^2 f}{\partial x\partial x} + k(x^{2k-1}+x^{k-1})\frac{\partial f}{\partial y}]^2 + (x^k\frac{\partial^2 f}{\partial x\partial y})^2 + (x^{2k}\frac{\partial^2 f}{\partial y^2})^2 \\
&\geq \frac{1}{4}[\mathbf{trace}(\mathfrak{Hess}_a)f]^2.
\end{aligned}$$

Computing (2.10), $L = (\frac{\partial^2 f}{\partial x^2})^2$. Thus operator L satisfies $CD(-\frac{(2k^2+k)}{x^2}, \frac{1}{4})$ on the half Grushin plane $\mathbb{R}^2_{\pm}$. When $x = 0$, we get $\mathfrak{R}_a = 0$. We cannot take the limit here, however, comparing to the bound Ric_g in [39, Section 4], the bound we obtain here with $k = 1$ is improved. We can also introduce matrix z for Grushin plane to get bounds when we take $x = 0$. We leave this for future studies.

4 Functional inequalities

In this section, we present various functional inequalities under different Assumptions (i.e. Assumption 3.3, 3.22 and 3.30). Under each Assumption, we have proposed the corresponding curvature dimension condition. We then prove functional inequalities under each curvature dimension condition. We first introduce the following semi-group.

Definition 4.1. *We define the semigroup* $P_t = e^{\frac{1}{2}tL}$, *where* L *is invariant w.r.t the* *invariant measure $d\mu$. The process X_t is the solution of Stratonovich SDE* (2.5)*.*

$$\begin{aligned} P_t f(x) = \mathbb{E}(f(X_t)) &= \int_{\mathbb{M}^{n+m}} f(y)p(t,x,y)d\mu(y) \\ &= \int_{\mathbb{M}^{n+m}} f(y)p(t,x,y)\mathbf{Vol}(y)dy \\ &= \int_{\mathbb{M}^{n+m}} f(y)\rho(t,x,y)dy, \end{aligned}$$

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 **Vol**.

Remark 4.2. *In the standard sub-Riemannian setting, the semi-groups is 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 [16], we also need the following assumption which is necessary to rigorously justify computations on functionals of the heat semigroup.

Assumption 4.3. *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 $f \in \mathcal{C}^\infty(\mathbb{M}^{n+m})$ with compact support, we assume that*

$$\sup_{t\in[0,T]} \|\Gamma(P_t f)\|_\infty + \|\Gamma_1^{z}(P_t f)\|_\infty < +\infty. \tag{4.82}$$

We believe that the above Assumption 4.3 should follow from the Definition 1.10 if we assume the appropriate lower bounds on the curvature tensor. We leave this for further studies. Related gradient estimates are presented in order below. For the infinitesimal generator $\frac{1}{2}L = \frac{1}{2}(\Delta_p - A\nabla)$ associated with linear semi-group P_t, we have the following property.

Proposition 4.4. *For all smooth function f, we have*

- $P_0 = Id$;
- *For all functions $f \in \mathcal{C}_b(\mathbb{R}^{n+m})$, 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 $P_t \circ P_s = P_{t+s}$;*
- $\forall x \in \mathbb{R}^{n+m}$, $\forall t \geq 0$, $\frac{\partial}{\partial_t} P_t f(x) = \frac{1}{2}L(P_t f)(x) = \frac{1}{2}P_t(Lf)(x)$.

4.1 Classical gradient estimates.

In this part, we first present the gradient estimates under the framework of Assumption 3.3, we have the following proposition.

Proposition 4.5. *We have the following equivalent conditions:*

(1). $\Gamma_2(f,f) \geq k\Gamma_1(f,f)$*;*

(2). $\sqrt{\Gamma_1(P_t f)} \leq e^{-kt} P_t \sqrt{\Gamma_1(f)}$*;*

(3). $\Gamma_1(P_t f) \leq e^{-2kt} P_t \Gamma_1(f)$*.*

Remark 4.6. *When matrix a is an invertible square matrix, Theorem 3.4 and Theorem 3.7 recovers the standard Riemannian setting, see [5, 7]. Once matrix a is a rectangular matrix, the above estimates generalize to sub-Riemannian case. In particular, we have* $\Gamma_1(f,f) = \langle a^\mathsf{T}\nabla f, a^\mathsf{T}\nabla f\rangle_{\mathbb{R}_n} = \langle aa^\mathsf{T}\nabla f, aa^\mathsf{T}\nabla f\rangle_{(aa^\mathsf{T})^\dagger} = \langle \nabla^R_{\mathcal{H}} f, \nabla^R_{\mathcal{H}} f\rangle_{\mathcal{H}}$*, where* $aa^T\nabla f = P^\tau \nabla^R f = \nabla^R_{\mathcal{H}} f$ *represents horizontal gradient of function f under metric* $g = (aa^T)^\dagger + cc^\mathsf{T}$*. Thus we can reformulate the equivalent condition in the following sense,*

(1). $\Gamma_2(f,f) \geq k\Gamma_1(f,f)$*;*

(2). $|\nabla_{\mathcal{H}} P_t f| \leq e^{-2kt} P_t |\nabla_{\mathcal{H}} f|$*;*

(3). $|\nabla_{\mathcal{H}} P_t f|^2 \leq e^{-2kt} P_t |\nabla_{\mathcal{H}} f|^2$*.*

Proof The proof follows the Riemannian manifold setting and our new Γ_2 calculations (Theorem 3.5) and the corresponding curvature dimension inequality (Theorem 3.7). Since we are in a new setting, we present the proof for one special case by using our framework, the rest should follow natually. We show that (1) $\Rightarrow$ (3) under an easier condition with $\mathfrak{R}_a \geq 0$, i.e. $k = 0$. Recall that $P_t = e^{\frac{1}{2}tL}$, define function φ_s

$$\varphi_s \equiv P_s |a^\mathsf{T}\nabla P_{t-s} f|^2(x) = P_s\langle a^\mathsf{T}\nabla P_{t-s} f, a^\mathsf{T}\nabla P_{t-s} f\rangle_{\mathbb{R}^n} = P_s\langle aa^\mathsf{T}\nabla P_{t-s} f, aa^\mathsf{T}\nabla P_{t-s} f\rangle_{(aa^\mathsf{T})^\dagger}$$

where the last term means that $P_s\langle aa^\mathsf{T}\nabla P_{t-s} f, aa^\mathsf{T}\nabla P_{t-s} f\rangle_{(aa^\mathsf{T})^\dagger} = P_s|\nabla_{\mathcal{H}} P_{t-s} f|^2(x)$. We will keep using the notation $P_s|a^\mathsf{T}\nabla P_{t-s} f|^2(x)$ for simplicity which also gives simpler proof. The following is standard strategy similar to the Riemanian setting. We first observe that

$$\varphi_t - \varphi_0 = P_t|a^\mathsf{T}\nabla f|^2(x) - |a^\mathsf{T}\nabla P_t f|^2(x) = \int_0^t \frac{d}{ds}\varphi(s)ds.$$

We then have the direct computations for $\frac{d}{ds}\varphi(s)$ below,

$$\begin{aligned}
&\frac{d}{ds}(P_s|a^{\mathsf{T}}\nabla P_{t-s}f|^2(x))\\
= \;& \dot{P}_s(|a^{\mathsf{T}}\nabla P_{t-s}f|^2(x)) + P_s(\frac{d}{ds}|a^{\mathsf{T}}\nabla P_{t-s}f|^2(x))\\
= \;& \frac{1}{2}LP_s|a^{\mathsf{T}}\nabla P_{t-s}f|^2(x) - P_s(\langle a^{\mathsf{T}}\nabla P_{t-s}f, a^{\mathsf{T}}\nabla LP_{t-s}f\rangle_{\mathbb{R}^n}(x))\\
= \;& \frac{1}{2}LP_s|a^{\mathsf{T}}\nabla P_{t-s}f|^2(x)\\
& -P_s\left(\frac{1}{2}L\Gamma_1(P_{t-s}f, P_{t-s}f) - |\mathfrak{Hess}_a P_{t-s}f|^2 - \mathfrak{R}_a(\nabla P_{t-s}f, \nabla P_{t-s}f)\right)\\
= \;& P_s\left(\Gamma_2(P_{t-s}f, P_{t-s}f)\right),
\end{aligned}$$

where the second last equality comes from Theorem 3.4 and the commuting property of L and P_t. Thus we get the following bound

$$\frac{d}{ds}\varphi(s) \geq 0.$$

The general argument for $\mathfrak{R}_a(\nabla f, \nabla f) \geq k\Gamma_1(f,f)$ follows by choosing function $\varphi(s) = e^{-kt}P_s|a^{\mathsf{T}}\nabla P_{t-s}f|^2(x)$, the details can be found in [7].

Proposition 4.7. *For our sub-Riemannian manifold defined in Definition 2.5. Assuming that $CD(k,\infty)$ is satisfied, namely*

$$|\mathfrak{Hess}_a f|^2 + \mathfrak{R}_a(\nabla f, \nabla f) \geq k\Gamma_1(f,f);$$

For all smooth and non-negative functions f on $\mathbb{M}^{n+m}$, we have

$$\int_{\mathbb{M}^{n+m}} f^2 \ln f^2 d\mu - \int_{\mathbb{M}^{n+m}} f^2 d\mu \ln \int_{\mathbb{M}^{n+m}} f^2 d\mu \leq 2\frac{1-e^{-2kt}}{k}\int_{\mathbb{M}^{n+m}} \Gamma_1(f,f)d\mu, \quad (4.83)$$

where $\Gamma_1(f,f) = \langle a^{\mathsf{T}}\nabla f, a^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}_n} = \langle aa^{\mathsf{T}}\nabla f, aa^{\mathsf{T}}\nabla f\rangle_{(aa^{\mathsf{T}})^{\dagger}} = \langle \nabla^R_{\mathcal{H}} f, \nabla^R_{\mathcal{H}}\rangle_{(aa^T)^{\dagger}}$.

Proof The proof follows the same as in Riemannian setting and our Γ_2 calculations, we refer to [5] for details.

Remark 4.8. *Proposition 4.5 and Theorem 4.7 for semigroup L is limited. In order to include more spaces, we need to study general semigroup defined for $\tilde{L}$ with drift introduced to the system, as well as matrix z below.*

Gradient estimates under Assumption 3.22. In the following, we present a similar gradient estimates of Proposition 4.5, which is not restricted on horizontal directions. The following gradient estimate follows from the curvature dimension bound under the

Assumption 3.22. We follow closely the framework from [16]. Define two functionals below

$$\begin{aligned}
\varphi_s^a &\equiv P_s|a^\intercal \nabla P_{t-s}f|^2(x) = P_s\langle a^\intercal \nabla P_{t-s}f, a^\intercal \nabla P_{t-s}f\rangle_{\mathbb{R}^n},\\
\varphi_s^z &\equiv P_s|z^\intercal \nabla P_{t-s}f|^2(x) = P_s\langle z^\intercal \nabla P_{t-s}f, z^\intercal \nabla P_{t-s}f\rangle_{\mathbb{R}^m}.
\end{aligned}$$

We prove the following gradient estimates.

Proposition 4.9. *If $\mathfrak{R}_a(\nabla f, \nabla f) + \mathfrak{R}_z(\nabla f, \nabla f) \geq \kappa\Gamma_1(f,f) + \kappa\Gamma_1^z(f,f)$, we have the following gradient estimates,*

$$\Gamma_1(P_tf) + \Gamma_1^z(P_tf) \leq e^{-2\kappa t}(P_t\Gamma_1(f) + P_t\Gamma_1^z(f));$$

Proof The proof is similar to the proof of Proposition 4.5. But we take $\varphi_s = \varphi_s^a + \varphi_s^z$. Then we use Theorem 3.25 to get bound of $\Gamma_2(f,f) + \Gamma_2^z(f,f)$ which completes the proof.

4.2 Entropic inequality under Generalized Gamma z condition

In this part, we work in the most general framework under the Assumption 3.30. We follow closely the framework introduced in [16] 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 4.10. *We have the following relation*

$$\frac{1}{2}L\phi_a + \frac{\partial}{\partial t}\phi_a = (P_{T-t}f)(x)\Gamma_2(\log P_{T-t}f, \log P_{T-t}f)(x), \tag{4.84}$$

$$\begin{aligned}
\frac{1}{2}L\phi_z + \frac{\partial}{\partial t}\phi_z &= (P_{T-t}f)(x)\Gamma_2^z(\log P_{T-t}f, \log P_{T-t}f)(x)\\
&\quad + (P_{T-t}f)(x)\Gamma_1(\log P_{T-t}f, \Gamma_1^z(P_{T-t}f, P_{T-t}f))(x)\\
&\quad - (P_{T-t}f)(x)\Gamma_1^z(\log P_{T-t}f, \Gamma_1(P_{T-t}f, P_{T-t}f))(x).
\end{aligned} \tag{4.85}$$

Proof Denote $g(t,x) = P_{T-t}f(x) = \int \rho(t,x,\tilde{x})f(\tilde{x})d\tilde{x}$, 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

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

where we have $\Gamma_1(g,\Gamma_1(\log g,\log g)) = g\Gamma_1(\log g,\Gamma_1(\log g,\log g))$, thus (4.85) is proved. Similarly, we obtain the following for ϕ_z

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

The proof then follows. ■

Now, we are ready to present the following important lemma which prepare us ready 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 4.11. *For any $0 < s < T$, we denote $\rho(s,x,y) = p(s,x,y)\mathbf{Vol}(y)$ as the transition kernel of diffusion process X_s^x starting at x defined in Definition 4.1, the following equality is satisfied*

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

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

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

Proof We first expand in the following integral form.

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

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 $\Delta_a = \nabla \cdot (aa^\mathsf{T}\nabla)$ and $\Delta_z = \nabla \cdot (zz^\mathsf{T}\nabla)$. 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 get

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

Similarly, the other equality is satisfied.

Claim 2:

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

First observe that

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

Similarly, one gets

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

For the next term, one obtains

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

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

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

Observe that by integration by parts, we get

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

Combine the above formulas, the proof is completed. ■

With the above Lemma in hand, we are ready to prove the main result of this paper shown in Theorem 1.15 in details. For reader's convenience, we reformulate the theorem below. We first define the following engery form

$$\Phi_a(x,t) \quad = P_t\left(P_{T-t}f\Gamma_1(\log P_{T-t}f)\right)(x),$$

and

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

Recall that, we define

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

Theorem 4.12. *Denote $\phi = \phi_a + \phi_z$, if the following condition is satisfied*

$$\mathfrak{R}_a^G + \mathfrak{R}_z + \mathfrak{R}^{\rho(s)} \succeq \kappa_s(\Gamma_1 + \Gamma_1^z),$$

where the bound κ_s depends on the estimate of transition kernel $\nabla \log \rho(s, \cdot, \cdot)$ associated with semi-group P_s(see Definition 4.1). We then conclude

$$P_T(\phi(\cdot, T))(x) \quad \geq \quad \phi(x, 0) + \int_0^T \kappa_s(\Phi_a(x, s) + \Phi_z(x, s))ds. \qquad (4.86)$$

Remark 4.13. *Based on Theorem 3.35, we can also prove the above theorem for operator $\tilde{L}$ with drift term involved. 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 are referred to [16][Proposition 4.5]) Let smooth function $u : \mathbb{R}^{n+m} \to \mathbb{R}$ 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 $t > 0$,

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

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

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

By using the dominated convergence theorem, we get

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

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

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

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

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

By using the above Lemma 4.11, let $h = \log g$ we get

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

Apply Theorem 3.32 here with $\Psi = \rho(s,\cdot,\cdot)$ as the transition kernel function, we get a time dependent version of Theorem 3.32. Assuming that the following bound is satisfied where the bound κ_s depends on kernel $\rho(s,\cdot,\cdot)$,

$$
\mathfrak{R}_a^G(\nabla f, \nabla f) + \mathfrak{R}_z(\nabla f, \nabla f) + \mathfrak{R}^{\rho(s)}(\nabla f, \nabla f) \geq \kappa_s(\Gamma_1(f,f) + \Gamma_1^z(f,f)).
$$

We then conclude with the following bound

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

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. ■

Remark 4.14. *We prove the most important entropic inequality Theorem 4.12 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 $\Gamma_1(f, \Gamma_1^z(f,f)) = \Gamma_1^z(f, \Gamma_1(f,f))$ is first proved in [16][Proposition 4.5] and [Theorem 5.2]. With this new inequality Theorem 4.12 in hand, similar gradient estimates and other inequalities from [16] follows. We leave this for future studies. The Proposition 4.5 in [16] is based on a point-wise estimate given the commutative assumption of Γ_1 and Γ_1^z. We remove the commutative assumption and our estimate is in a weak form, which is presented in the above Lemma 4.11.*

5 Sub-Riemannian Density manifold

In this section, we illustrate the motivation of our generalized Gamma z calculus in this paper. Consider a density space over the sub-Riemannian manifold. The finite dimensional sub-Rimennian structure formulates the density space: the *infinite dimensional sub-Riemannian structure.* We call it the *sub-Riemannian density manifold* (SDM). We provide the geometric calculations in SDM. We study the Fokker-Planck equation as the sub-Riemannian gradient flow in SDM. We demonstrate the equivalence relation between second order calculus of relative entropy in SDM and Generalized Gamma z calculus.

5.1 Sub-Riemannian Density manifold

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

$$\mathcal{P}(\mathbb{M}^{n+m}) = \Big\{\rho(x) \in C^\infty(\mathbb{M}^{n+m}) \colon \int \rho(x)dx = 1, \quad \rho(x) \geq 0\Big\}.$$

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

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

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

Definition 5.1 (sub-Riemannian Wasserstein metric tensor)**.** *The L^2 sub-Riemannian-Wasserstein metric $g_\rho^{W_a} \colon T_\rho\mathcal{P}(\mathbb{M}^{n+m}) \times T_\rho\mathcal{P}(\mathbb{M}^{n+m}) \to \mathbb{R}$ is defined by*

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

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

$$\Delta_\rho^a = \nabla \cdot (\rho aa^\mathsf{T}\nabla).$$

For some special choices of a as studied in [32] or aa^T forming a positive definite matrix, then Δ_ρ^a is an elliptic operator. In this case, $(\mathcal{P}(\mathbb{M}^{n+m}), g^{W_a})$ still forms a Riemannian density manifold. In general, given a sub-Riemannian manifold $(\mathbb{M}^{n+m}, (aa^\mathsf{T})^\dagger)$, Δ_ρ^a is only a sub-elliptic operator. Thus $(\mathcal{P}(\mathbb{M}^{n+m}), g^{W_a})$ forms an infinite-dimensional sub-Riemannian manifold.

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

$$V \colon C^\infty(\mathbb{M}^{n+m}) \to T_\rho\mathcal{P}(\mathbb{M}^{n+m}), \quad V_\Phi = -\Delta_\rho^a\Phi = -\nabla \cdot (\rho aa^\mathsf{T}\nabla\Phi).$$

Here $\Phi \in T_\rho^*\mathcal{P}(\mathbb{M}^{n+m}) = C^\infty(\mathbb{M}^{n+m})/\sim$. Here $T_\rho^*\mathcal{P}(\mathbb{M}^{n+m})$ is the cotangent space in SDM and $\sim$ represents a constant shrift 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,

$$
\begin{aligned}
g_\rho^{W_a}(V_{\Phi_1}, V_{\Phi_2}) =& \int V_{\Phi_1}(-\Delta_\rho^a)^\dagger V_{\Phi_2} dx \\
=& \int \Phi_1(-\Delta_\rho^a)(-\Delta_\rho^a)^\dagger(-\Delta_\rho^a)\Phi_2 dx \\
=& \int (\Phi_1, -\Delta_\rho^a \Phi_2) dx \\
=& \int \Phi_1(-\nabla\cdot(\rho a a^\mathsf{T}\nabla\Phi_2) dx \\
=& \int (a^\mathsf{T}\nabla\Phi_1, a^\mathsf{T}\nabla\Phi_2)\rho dx,
\end{aligned}
\tag{5.87}
$$

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 spatial integration by parts.

We next derive several basic geometric calculations in SDM.

Proposition 5.2 (Geodesics in SDM)**.** *The sub-Riemannian geodesics in cotangent bundle forms*

$$
\begin{cases}
\partial_t\rho_t + \nabla\cdot(\rho_t a a^\mathsf{T}\nabla\Phi_t) = 0 \\
\partial_t\Phi_t + \frac{1}{2}(a^\mathsf{T}\nabla\Phi_t, a^\mathsf{T}\nabla\Phi_t) = 0.
\end{cases}
\tag{5.88}
$$

Proof We consider 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 ρ_0, ρ_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 w.r.t. $\partial_t\rho_t$ and δ_{ρ_t} is the L^2 first variation w.r.t. ρ_t. Here

$$
\begin{aligned}
\partial_t\Big((-\Delta_{\rho_t}^a)^\dagger\partial_t\rho_t\Big) =& \delta_\rho \int \frac{1}{2}\Big(\partial_t\rho, (-\Delta_{\rho_t}^a)^\dagger\partial_t\rho_t\Big) dx \\
=& -\frac{1}{2}\|\nabla\Big((\Delta_\rho^a)^\dagger\partial_t\rho_t\Big)\|^2,
\end{aligned}
\tag{5.89}
$$

where the last equality uses the following fact

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

Denote $\partial_t\rho_t = -\Delta_{\rho_t}^a\Phi_t$, then the Euler-Lagrange equation (5.89) forms the sub-Riemannian geodesics flow (5.88).

Proposition 5.3 (Gradient and Hessian operators in SDM)**.** *Given a functional* $\mathcal{F}\colon \mathcal{P}(\mathbb{M}^{n+m}) \to \mathbb{R}$*, the gradient operator of* $\mathcal{F}$ *in* $(\mathcal{P}, g^{W_a})$ *satisfies*

$$grad_{W_a}\mathcal{F}(\rho) = -\nabla \cdot (\rho a a^{\mathsf{T}} \nabla \delta \mathcal{F}(\rho)).$$

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

$$\begin{aligned}
&Hess_{W_a}\mathcal{F}(V_{\Phi_1}, V_{\Phi_2})\\
=&\int\int (a^{\mathsf{T}}\nabla_y)(a^{\mathsf{T}}\nabla_x)\delta^2\mathcal{F}(\rho)(x,y)\Big(a(x)^{\mathsf{T}}\nabla_x\Phi_1(x), a(y)^{\mathsf{T}}\nabla_y\Phi_2(y)\Big)\rho(x)\rho(y)dxdy\\
&+\int Hess_a\delta\mathcal{F}(\rho)(\Phi,\Phi)\rho dx,
\end{aligned}$$

where

$$Hess_a\delta\mathcal{F}(\rho)(\Phi,\Phi) = \frac{1}{2}\Big\{\Gamma_1(\Gamma_1(\delta\mathcal{F},\Phi_1),\Phi_2)+\Gamma_1(\Gamma_1(\delta\mathcal{F},\Phi_2),\Phi_1)-\Gamma_1(\Gamma_1(\Phi_1,\Phi_2),\delta\mathcal{F})\Big\}.$$

Proof We first derive the sub-Riemannian gradient operator.

$$\begin{aligned}
\mathrm{grad}_{W_{\mathcal{H}}}\mathcal{F}(\rho) =&\Big((-\Delta_\rho^a)^\dagger\Big)^\dagger \frac{\delta}{\delta\rho(x)}\mathcal{F}(\rho)\\
=&-\Delta_\rho^a\frac{\delta}{\delta\rho(x)}\mathcal{F}(\rho)\\
=&-\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\frac{\delta}{\delta\rho(x)}\mathcal{F}(\rho)).
\end{aligned}$$

The Hessian operator satisfies

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

where (ρ_t, Φ_t) satisfies the geodesics equation (5.88) with $\rho_0 = \rho$, $\Phi_0 = \Phi$. Notice the fact that

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

In addition,

$$
\begin{aligned}
&\frac{d^2}{dt^2}\mathcal{F}(\rho_t)|_{t=0} = \frac{d}{dt}\int (a^{\mathsf{T}}\nabla\delta\mathcal{F}(\rho_t), a^{\mathsf{T}}\nabla\Phi_t)\rho_t dx|_{t=0} \\
&= \int\int \delta^2\mathcal{F}(\rho)(x,y)\partial_t\rho_t(x)\partial_t\rho_t(y)dxdy \\
&\quad + \int (a^{\mathsf{T}}\nabla\delta\mathcal{F}(\rho_t), a^{\mathsf{T}}\nabla\partial_t\Phi_t)\rho_t dx \\
&\quad + \int (a^{\mathsf{T}}\nabla\delta\mathcal{F}(\rho_t), a^{\mathsf{T}}\nabla\partial_t\Phi_t)\partial_t\rho_t dx|_{t=0} \\
&= \int\int \delta^2\mathcal{F}(\rho)(x,y)\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi)(x)\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi)(y)dxdy \\
&\quad - \frac{1}{2}\int (a^{\mathsf{T}}\nabla\delta\mathcal{F}(\rho), a^{\mathsf{T}}\nabla\Gamma_1(\Phi,\Phi))\rho dx \\
&\quad + \int \Gamma_1(\Phi,\delta\mathcal{F}(\rho))\Big(-\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi)\Big)dx \\
&= \int\int (a^{\mathsf{T}}\nabla_y)(a^{\mathsf{T}}\nabla_x)\delta^2\mathcal{F}(\rho)(x,y)\Big(a(x)^{\mathsf{T}}\nabla_x\Phi_1(x), a(y)^{\mathsf{T}}\nabla_y\Phi_2(y)\Big)\rho(x)\rho(y)dxdy \\
&+\frac{1}{2}\int \Big\{\Gamma_1(\Gamma_1(\delta\mathcal{F},\Phi_1),\Phi_2) + \Gamma_1(\Gamma_1(\delta\mathcal{F},\Phi_2),\Phi_1) - \Gamma_1(\Gamma_1(\Phi_1,\Phi_2),\delta\mathcal{F})\Big\}\rho dx\ ,
\end{aligned}
\tag{5.90}
$$

where the last equality holds by the integration by parts formula.

5.2 Gamma calculus via Hessian operator of relative entropy in SDM

We show the equivalence relation between Hessian of relative entropy in SDM and classical Gamma two operator.

We first demonstrate the relation among L^*, Δ_p and gradient operator of entropy. In particular, we show that the Fokker-Planck equation is a sub-Riemannian gradient flow in SDM.

Proposition 5.4 (Gradient flow)**.** *Denote the negative Boltzmann-Shannon entropy*

$$
\mathrm{H}(\rho) = \int \rho(x)\log\rho(x)dx,
$$

and the relative entropy

$$
\mathrm{D}(\rho) = \int \rho(x)\log\frac{\rho(x)}{\mathbf{Vol}(x)}dx, \tag{5.91}
$$

where $\mathbf{Vol}(x)$ *is given in Lemma 2.8, then the negative gradient operators in* $(\mathcal{P}, g^{W_a})$ *form*

$$
-grad_{W_a}\mathrm{H}(\rho) = \Delta_p\rho = \nabla\cdot(aa^{\mathsf{T}}\nabla\rho),
$$

and

$$-grad_{W_a}\mathrm{D}(\rho) = L^*\rho = \nabla\cdot(aa^{\mathsf{T}}\nabla\rho) - \nabla\cdot(\rho a\otimes\nabla a).$$

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

$$\partial_t\rho = \nabla\cdot(aa^{\mathsf{T}}\nabla\rho) + \nabla\cdot(\rho a\otimes\nabla a). \tag{5.92}$$

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

$$\delta_{\rho(x)}\mathrm{H}(\rho) = \log\rho(x) + 1,$$

and

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

Thus

$$\mathrm{grad}_{W_a}\mathrm{H}(\rho) = -\nabla\cdot(\rho aa^{\mathsf{T}}\nabla(\log\rho+1)) = -\nabla\cdot(\rho aa^{\mathsf{T}}\nabla\log\rho) = \nabla\cdot(aa^{\mathsf{T}}\nabla\rho),$$

and

$$\mathrm{grad}_{W_a}\mathrm{D}(\rho) = -\nabla\cdot(\rho aa^{\mathsf{T}}\nabla\log\rho) + \nabla\cdot(\rho aa^{\mathsf{T}}\nabla\log\mathbf{Vol}),$$

where $\rho\nabla\log\rho = \rho\frac{\nabla\rho}{\rho} = \nabla\rho$, and $-aa^{\mathsf{T}}\nabla\log\mathbf{Vol} = a\otimes\nabla a$. Following the gradient flow formulation

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

we finish the derivation of (5.92). ■

Remark 5.5. *Here the gradient flow* (5.92) *and Fokker-Planck equation of* (2.6) *is different up to a scalar two. All the dynamical behaviors of* (5.92) *and* (2.6) *are same. For the simplicity of presentation, we mainly use* (5.92) *in this section.*

We next demonstrate that the Hessian of entropy and relative entropy is equivalent to the classical Bakry-Émery calculus, such as $\Gamma_{2,a}$ (defined in (3.17)) and Γ_2 (defined in (3.18)).

Proposition 5.6 (Hessian of Entropy and Bakry-Émery calculus)**.** *Given* Φ_1*,* $\Phi_2 \in C^\infty(\mathbb{M}^{n+m})$*, then*

$$Hess_{W_a}\mathrm{H}(\rho)(V_\Phi, V_\Phi) = \int \Gamma_{2,a}(\Phi,\Phi)\rho(x)dx,$$

and

$$Hess_{W_a}\mathrm{D}(\rho)(V_\Phi, V_\Phi) = \int \Gamma_2(\Phi,\Phi)\rho(x)dx.$$

Proof We first derive the Hessian of $\mathrm{D}(\rho)$ in SDM. Notice the fact $\delta^2\mathrm{D}(\rho)(x,y) = \frac{1}{\rho}\delta_{x=y}$. For simplicity, we denote $\delta^2\mathrm{D}(\rho) = \frac{1}{\rho(x)}$. By using (5.90), we have

$$\begin{aligned}
\mathrm{Hess}_{W_a}\mathrm{D}(\rho)(V_\Phi, V_\Phi) = & \int \delta^2\mathrm{D}(\rho)(x)\Big(\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi)\Big)^2 dx && (a)\\
& -\frac{1}{2}\int (a^{\mathsf{T}}\nabla\delta\mathrm{D}(\rho), a^{\mathsf{T}}\nabla\Gamma_1(\Phi,\Phi))\rho dx && (b)\\
& +\int \Gamma_1(\Phi, \delta\mathrm{D}(\rho))\Big(-\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi)\Big)dx. && (c)
\end{aligned} \tag{5.93}$$

We next rewrite (5.93) into iterative Gamma calculus. We first show

$$\begin{aligned}
(a)+(c) = & \int \Big(\delta^2\mathrm{D}(\rho)\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi) - \Gamma_1(\Phi,\delta\mathrm{D}(\rho))\Big)\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi)\rho dx\\
= & \int \Big(\frac{1}{\rho}\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi) - (a a^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}}, \nabla\Phi)\Big)\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi)\rho dx\\
= & \int \Big(\frac{1}{\rho}(\nabla\rho, a a^{\mathsf{T}}\nabla\Phi) + \nabla\cdot(a a^{\mathsf{T}}\nabla\Phi) - (a a^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}}, \nabla\Phi)\Big)\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi)\rho dx\\
= & \int \Big((\nabla\log\rho, a a^{\mathsf{T}}\nabla\Phi) + \nabla\cdot(a a^{\mathsf{T}}\nabla\Phi)\\
& \quad - (\nabla\log\rho, a a^{\mathsf{T}}\nabla\Phi) - (a\otimes\nabla a, \nabla\Phi)\Big)\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi)\rho dx\\
= & \int \Big((\nabla\cdot(a a^{\mathsf{T}}\nabla\Phi) - (a\otimes\nabla a, \nabla\Phi)\Big)\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi)\rho dx\\
= & \int L\Phi\nabla\cdot(\rho a a^{\mathsf{T}}\nabla\Phi)dx\\
= & -\int \Gamma_1(L\Phi, \Phi)\rho dx,
\end{aligned}$$

where the fourth equality uses the fact that $\frac{\nabla\rho}{\rho} = \nabla\log\rho$ and $a a^{\mathsf{T}}\nabla\log\mathbf{Vol} = a\otimes\nabla a$, while the last equality follows the integration by parts.

We second show

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

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

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

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

Similarly, we can show that the Hessian operator of $\mathrm{H}(\rho)$ in $(\mathcal{P}, g)$ equals to the expectation of $\Gamma_{2,a}$ operator.

Remark 5.7. *If aa^{T} is invertible, then $(\mathbb{M}^{n+m}, g = (aa^{\mathsf{T}})^{-1})$ forms a Riemannian manifold, and* **Vol** *is the classical Riemannian volume form. Denote the probability density function $\rho(x) = p(x)\mathbf{Vol}(x)$ and $d\mu = \mathbf{Vol}(x)dx$, then*

$$\mathrm{D}(\rho) = \int \rho(x) \log \frac{\rho(x)}{\mathbf{Vol}(x)} dx = \int p(x) \log p(x) d\mu(x).$$

As shown in classical optimal transport theory, we have

$$\begin{aligned} Hess_{W_a} \mathrm{D}(\rho)(V_\Phi, V_\Phi) &= \int \Gamma_2(\Phi, \Phi)\rho dx \\ &= \int \Big\{ Ric_g(grad_g \Phi, grad_g \Phi) + \|Hess_g \Phi\|^2 \Big\} \rho dx, \end{aligned}$$

where Ric_g, $grad_g$, $Hess_g$ are Riemannian Ricci, gradient, Hessian operators. In other words, our definition for sub-Riemannian Ricci tensor is exactly the Ricci curvature tensor in classical Riemannian settings.

Remark 5.8. *Here we notice that $\Gamma_{2,a}$ corresponds to the Hessian of negative Boltzmann-Shannon entropy. Again, if aa^{T} is invertible, $(\mathbb{M}^{n+m}, g = (aa^{\mathsf{T}})^{-1})$ forms a Riemannian manifold. Denote the probability density function $\rho(x) = p(x)\mathbf{Vol}(x)$ and $d\mu = \mathbf{Vol}(x)dx$, then*

$$\mathrm{H}(\rho) = \int \rho \log \rho dx = \int p(x)(\log p(x) + \log \mathbf{Vol}(x)) d\mu.$$

And the Hessian operator of $\mathrm{H}(\rho)$ in $(\mathcal{P}, g_\rho^{W_a})$ forms

$$\begin{aligned} Hess_{W_a} \mathrm{H}(\rho)(V_\Phi, V_\Phi) &= \int \Gamma_{2,a}(V_\Phi, V_\Phi)\rho dx \\ &= \int \Big\{ \big(Ric_g + Hess_g \log \mathbf{Vol}(x)\big)(grad_g \Phi, grad_g \Phi) + \|Hess_g \Phi\|^2 \Big\} \rho dx. \end{aligned}$$

Here Proposition 5.6 shows Remark 3.12. We notice that -$\mathbf{B} + \mathbf{B}_0$ is exactly the Hessian operator of $\log \mathbf{Vol}(x)$ in the Riemannian setting. In other words, $-\mathbf{B} + \mathbf{B}_0$ is the sub-Riemannian analog of Hessian operator of $\log \mathbf{Vol}(x)$ in $(\mathbb{M}^{n+m}, (aa^{\mathsf{T}})^{\dagger})$.

If Assumption 1.6 is satisfied and the sub-Riemannian Ricci curvature tensor is found, following classical approaches in Bakry-Émery calculus, we can recover all the related functional inequalities in sub-Riemannian manifold. Here we use the Log-Sobolev inequality as an example for all related proofs of functional inequalities. Notice the fact that Log-Sobolev inequality describes a relationship between relative entropy and relative Fisher information functional. Here the relative Fisher information functional on sub-Riemannian manifold is defined by

$$
\begin{aligned}
\mathrm{I}_a(\rho) :=& g_\rho^{W_a}\Big(\mathrm{grad}_{W_a}\mathrm{D}(\rho), \mathrm{grad}_{W_a}\mathrm{D}(\rho)\Big) \\
=& \int \Big(a(x)^\mathsf{T}\nabla \log\frac{\rho(x)}{\mathbf{Vol}(x)}, a(x)^\mathsf{T}\nabla \log\frac{\rho(x)}{\mathbf{Vol}(x)}\Big)\rho(x)dx.
\end{aligned} \tag{5.94}
$$

In the following, we derive inequalities based on the Fisher information functional (5.94).

Proposition 5.9. *If $\mathfrak{R}_a \succeq \kappa > 0$, then the Logarithmic-Sobolev inequality on sub-Riemannian manifold holds*

$$
\mathrm{D}(\rho) \leq \frac{1}{2\kappa}\mathrm{I}_a(\rho), \quad \textit{for any } \rho \in \mathcal{P}(\mathbb{M}^{n+m}).
$$

Remark 5.10. *If we let $\rho(x) = f(x)^2$, then we recover the inequality* (4.83)*.*

Proof [Proof of Proposition 5.9] We briefly present the proof. Consider the sub-Riemannian gradient flow in SDM.

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

Then

$$
\frac{d}{dt}\mathrm{D}(\rho_t) = -g_{\rho_t}^{W_a}\Big(\mathrm{grad}_{W_a}\mathrm{D}(\rho_t), \mathrm{grad}_{W_a}\mathrm{D}(\rho_t)\Big) = -\mathrm{I}_a(\rho_t).
$$

Thus the Log-Sobolev inequality relates to the ratio of $\frac{d}{dt}\mathrm{D}(\rho_t)$ and $\frac{d^2}{dt^2}\mathrm{D}(\rho_t)$ along sub-Riemannian gradient flow in SDM. If we can estimate a ratio $\kappa > 0$, such that

$$
\frac{d^2}{dt^2}\mathrm{D}(\rho_t) \geq 2\kappa\frac{d}{dt}\mathrm{D}(\rho_t), \tag{5.95}
$$

i.e.

$$
\frac{d}{dt}\mathrm{I}_a(\rho_t) \geq -2\kappa\mathrm{I}_a(\rho_t),
$$

then by Gronwall's inequality, one proves the inequality. Here, (5.95) exactly relates to the Hessian operator of relative entropy in SDM along the gradient flow. In other words, notice the fact that

$$
\begin{aligned}
\frac{d^2}{dt^2}\mathrm{D}(\rho_t) =& \frac{d}{dt}\mathrm{I}_a(\rho_t) = 2\mathrm{Hess}_{W_a}\mathrm{D}(\rho_t)(\partial_t\rho_t, \partial_t\rho_t), \\
\frac{d}{dt}\mathrm{D}(\rho_t) =& -\mathrm{I}_a(\rho_t) = -g_{\rho_t}^{W_a}(\partial_t\rho_t, \partial_t\rho_t).
\end{aligned}
$$

From Proposition 5.6, the proposed generalized Ricci curvature bound establishes the sub-Riemannian Log-Sobolev inequality (sLSI), i.e.

$$
\mathfrak{R}_a \succeq \kappa > 0 \Rightarrow \mathrm{Hess}_{W_a}\mathrm{D}(\rho) \succeq \kappa g_\rho^{W_a} \Rightarrow \frac{d}{dt}\mathrm{I}_a(\rho_t) \leq -2\kappa\mathrm{I}_a(\rho_t) \Rightarrow \text{sLSI}.
$$

5.3 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 SDM viewpoint. Consider the SDM gradient flow (5.92)

$$\partial_t \rho_t = \Delta^a_{\rho_t} \delta \mathrm{D}(\rho_t).$$

When a is degenerate matrix and Assumption 1.6 is not satisfied, the classical relative Fisher information may not be the Lyapunov functional. In other words, along the gradient flow, it is possible that $\frac{d}{dt}\mathrm{I}_a(\rho_t) \geq 0$.

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 $\Delta^z_\rho = \nabla \cdot (\rho z z^{\mathsf{T}} \nabla)$ and $\mathrm{I}_z(\rho) = \int \Big(\delta \mathrm{D}, (-\Delta^z_\rho)\delta \mathrm{D}\Big) dx$. Construct

$$\mathrm{I}_{a,z}(\rho) := \mathrm{I}_a(\rho) + \mathrm{I}_z(\rho) = \int \Big(\delta \mathrm{D}, (-\Delta^a_\rho - \Delta^z_\rho)\delta \mathrm{D}\Big) dx.$$

We next prove the following proposition.

Proposition 5.11.

$$\frac{d}{dt}\mathrm{I}_{a,z}(\rho_t) = -2 \int \Big(\Gamma_2(\delta \mathrm{D}, \delta \mathrm{D}) + \tilde{\Gamma}^z_2(\delta \mathrm{D}, \delta \mathrm{D})\Big)\Big)\rho_t dx,$$

where

$$\tilde{\Gamma}^z_2(\Phi, \Phi) := \frac{1}{2} L(\Gamma^z_1(\Phi, \Phi)) - \Gamma_1(L_z \Phi, \Phi),$$

with the notation $\Delta_z = \nabla \cdot (z z^{\mathsf{T}} \nabla)$ *and* $L_z = \nabla \cdot (z z^{\mathsf{T}} \nabla) + (\nabla \log \mathbf{Vol}, z z^{\mathsf{T}} \nabla)$.

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

$$\frac{d}{dt}\mathrm{I}_{a,z}(\rho) = \frac{d}{dt}\mathrm{I}_a(\rho) + \frac{d}{dt}\mathrm{I}_z(\rho).$$

From Proposition 5.6, we have

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

We only need to show the following claim.

Claim:

$$\frac{d}{dt}\mathrm{I}_z(\rho) = -2\int \tilde{\Gamma}^z_2(\delta \mathrm{D}, \delta \mathrm{D}))\rho dx.$$

Proof [Proof of Claim] The proof is similar to the ones in Proposition 5.6. We need to take care of z direction. Notice

$$\begin{aligned}
\frac{d}{dt}\mathrm{I}_z(\rho) =& 2\int \delta^2\mathrm{D}\Big((-\Delta_\rho^z\delta\mathrm{D}), \partial_t\rho\Big)dx + \int(\nabla\delta\mathrm{D}, zz^\mathsf{T}\nabla\delta\mathrm{D})\partial_t\rho dx \\
=& 2\int \delta^2\mathrm{D}\Big((-\Delta_\rho^z\delta\mathrm{D}), \Delta_\rho^a\delta\mathrm{D}\Big)dx + \int(\nabla\delta\mathrm{D}, zz^\mathsf{T}\nabla\delta\mathrm{D})(\Delta_\rho^a\delta\mathrm{D})dx \\
=& -2\int \frac{1}{\rho}\nabla\cdot(\rho aa^\mathsf{T}\nabla\delta\mathrm{D})\nabla\cdot(\rho zz^\mathsf{T}\nabla\delta\mathrm{D})dx \qquad (I) \\
& + \int(\nabla\delta\mathrm{D}, zz^\mathsf{T}\nabla\delta\mathrm{D})\nabla\cdot(\rho aa^\mathsf{T}\nabla\delta\mathrm{D})dx \qquad (II)
\end{aligned}$$

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

$$\begin{aligned}
\frac{1}{\rho}\nabla\cdot(\rho zz^\mathsf{T}\nabla\delta\mathrm{D}) =& (\nabla\log\rho, zz^\mathsf{T}\nabla\delta\mathrm{D}) + \nabla\cdot(zz^\mathsf{T}\nabla\delta\mathrm{D}) \\
=& (\nabla\log\frac{\rho}{\mathrm{Vol}}, zz^\mathsf{T}\nabla\delta\mathrm{D}) + (\nabla\log\mathrm{Vol}, zz^\mathsf{T}\nabla\delta\mathrm{D}) + \nabla\cdot(zz^\mathsf{T}\nabla\delta\mathrm{D}) \\
=& (\nabla\delta\mathrm{D}, zz^\mathsf{T}\nabla\delta\mathrm{D}) + (\nabla\log\mathrm{Vol}, zz^\mathsf{T}\nabla\delta\mathrm{D}) + \nabla\cdot(zz^\mathsf{T}\nabla\delta\mathrm{D}).
\end{aligned}$$

Thus

$$\begin{aligned}
(I) =& -2\int \frac{1}{\rho}\nabla\cdot(\rho aa^\mathsf{T}\nabla\delta\mathrm{D})\nabla\cdot(\rho zz^\mathsf{T}\nabla\delta\mathrm{D})dx \\
=& -2\int \nabla\cdot(\rho aa^\mathsf{T}\nabla\delta\mathrm{D})\Big((\nabla\delta\mathrm{D}, zz^\mathsf{T}\nabla\delta\mathrm{D}) + (\nabla\log\mathrm{Vol}, zz^\mathsf{T}\nabla\delta\mathrm{D}) + \nabla\cdot(zz^\mathsf{T}\nabla\delta\mathrm{D})\Big)dx \\
=& -2\int(\nabla\delta\mathrm{D}, zz^\mathsf{T}\nabla\delta\mathrm{D})L^*\rho + \nabla\cdot(\rho aa^\mathsf{T}\nabla\delta\mathrm{D})\Big((\nabla\log\mathrm{Vol}, zz^\mathsf{T}\nabla\delta\mathrm{D}) + \nabla\cdot(zz^\mathsf{T}\nabla\delta\mathrm{D})\Big)dx \\
=& -2\int L(\nabla\delta\mathrm{D}, zz^\mathsf{T}\nabla\delta\mathrm{D})\rho dx \\
& + 2\int\Big\{\nabla\Big((\nabla\log\mathrm{Vol}, zz^\mathsf{T}\nabla\delta\mathrm{D}) + \nabla\cdot(zz^\mathsf{T}\nabla\delta\mathrm{D})\Big), aa^\mathsf{T}\nabla\delta\mathrm{D}\Big\}\rho dx,
\end{aligned}$$

where the last equality holds by integration by parts.

For (II), we have

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

Combining (I) and (II), we have

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

Using the notation $L_z = \Delta_z + (\nabla\log\mathbf{Vol}, zz^\mathsf{T}\nabla)$, we finish the proof.

We next prove that $\tilde{\Gamma}_2^z$ and $\Gamma_2^{z,\mathbf{Vol}}$ in Definition 1.3 agrees each other in the weak form along the gradient flow.

Proposition 5.12. *Denote $\Phi = \delta\mathcal{D}(\rho)$, then*

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

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

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

Here we need to prove the following equality.

Claim:

$$\begin{aligned}&\int \Big\{\Gamma_1^z(L\Phi,\Phi) - \Gamma_1(L_z\Phi,\Phi)\Big\}\rho dx\\ =& \int \rho\Big\{\frac{1}{\mathbf{Vol}}\nabla\cdot\Big(zz^{\mathsf{T}}\mathbf{Vol}\big(\nabla\Phi,\nabla(aa^{\mathsf{T}})\nabla\Phi\big)\Big) - \frac{1}{\mathbf{Vol}}\nabla\cdot\Big(aa^{\mathsf{T}}\mathbf{Vol}\big(\nabla\Phi,\nabla(zz^{\mathsf{T}})\nabla\Phi\big)\Big)\Big\}dx.\end{aligned}$$

Proof [Proof of Claim] For simplicity of notation, let

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

and

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

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

$$\begin{aligned}&\int \Gamma_1^z(L\Phi,\Phi)\rho dx\\ =& \int (\nabla(\nabla\cdot(aa^{\mathsf{T}}\nabla\Phi) - (A,\nabla\Phi)), zz^{\mathsf{T}}\nabla\Phi)\rho dx\\ =& \underbrace{\int (\nabla(\nabla\cdot(aa^{\mathsf{T}}\nabla\Phi)), zz^{\mathsf{T}}\nabla\Phi)\rho dx}_{(a1)} - \underbrace{\int (\nabla(A,\nabla\Phi), zz^{\mathsf{T}}\nabla\Phi)\rho dx}_{(a2)}.\end{aligned}$$

Here

$$
\begin{aligned}
(a1) =& \int \Big(\nabla(\nabla\cdot(aa^{\mathsf{T}}\nabla\Phi)), zz^{\mathsf{T}}\nabla\Phi\Big)\rho dx\\
=& -\int \nabla\cdot(aa^{\mathsf{T}}\nabla\Phi)\nabla\cdot(\rho zz^{\mathsf{T}}\nabla\Phi)dx\\
=& -\int \nabla\cdot(aa^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}})\nabla\cdot(\rho zz^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}})dx\\
=& -\int \nabla\cdot(\frac{1}{\rho}aa^{\mathsf{T}}\mathbf{Vol}\nabla\frac{\rho}{\mathbf{Vol}})\nabla\cdot(\rho zz^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}})dx\\
=& -\int \Big\{(\nabla\frac{1}{\rho}, aa^{\mathsf{T}}\mathbf{Vol}\nabla\frac{\rho}{\mathbf{Vol}}) + \frac{1}{\rho}\nabla\cdot(aa^{\mathsf{T}}\mathbf{Vol}\nabla\frac{\rho}{\mathbf{Vol}})\Big\}\nabla\cdot(\rho zz^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}})dx\\
=& -\int(\nabla\frac{1}{\rho}, aa^{\mathsf{T}}\mathbf{Vol}\nabla\frac{\rho}{\mathbf{Vol}})L_z^*\rho dx - \int\frac{1}{\rho}L^*\rho L_z^*\rho dx\\
=& \int\frac{1}{\rho^2}(\nabla\rho, aa^{\mathsf{T}}\mathbf{Vol}\nabla\frac{\rho}{\mathbf{Vol}})L_z^*\rho dx - \int\frac{1}{\rho}L^*\rho L_z^*\rho dx\\
=& \int(\nabla\log\rho, aa^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}})L_z^*\rho dx - \int\frac{1}{\rho}L^*\rho L_z^*\rho dx\\
=& \int(\nabla\log\frac{\rho}{\mathbf{Vol}}, aa^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}})L_z^*\rho dx\\
&+\int(\nabla\log\mathbf{Vol}, aa^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}})L_z^*\rho dx - \int\frac{1}{\rho}L^*\rho L_z^*\rho dx\\
=& -\int\Big(\nabla\big(\nabla\log\frac{\rho}{\mathbf{Vol}}, aa^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}}\big), zz^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}}\Big)\rho dx\\
&-\int\Gamma_1^z((\nabla\log\mathbf{Vol}, aa^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}}), \log\frac{\rho}{\mathbf{Vol}})\rho dx - \int\frac{1}{\rho}L^*\rho L_z^*\rho dx\\
=& -\int\Big(\big(\nabla\log\frac{\rho}{\mathbf{Vol}}, \nabla(aa^{\mathsf{T}})\nabla\log\frac{\rho}{\mathbf{Vol}}\big), zz^{\mathsf{T}}\nabla\frac{\rho}{\mathbf{Vol}}\Big)\mathbf{Vol}dx\\
&-\int 2\nabla^2\log\frac{\rho}{\mathbf{Vol}}\Big(aa^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}}, zz^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}}\Big)\rho dx\\
&-\int\Gamma_1^z((\nabla\log\mathbf{Vol}, aa^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}}), \log\frac{\rho}{\mathbf{Vol}})\rho dx - \int\frac{1}{\rho}L^*\rho L_z^*\rho dx\\
=& \int\nabla\cdot\Big(zz^{\mathsf{T}}\mathbf{Vol}\Big(\big(\nabla\log\frac{\rho}{\mathbf{Vol}}, \nabla(aa^{\mathsf{T}})\nabla\log\frac{\rho}{\mathbf{Vol}}\big)\Big)\frac{1}{\mathbf{Vol}}\rho dx\\
&-\int 2\nabla^2\log\frac{\rho}{\mathbf{Vol}}\Big(aa^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}}, zz^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}}\Big)\rho dx\\
&-\int\Gamma_1^z((\nabla\log\mathbf{Vol}, aa^{\mathsf{T}}\nabla\log\frac{\rho}{\mathbf{Vol}}), \log\frac{\rho}{\mathbf{Vol}})\rho dx - \int\frac{1}{\rho}L^*\rho L_z^*\rho dx.
\end{aligned}
$$

Notice the fact

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

Hence

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

Similarly, by switching a and z, we have

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

Thus combining the above, we finish the proof.

Remark 5.13. *From the proof, we can show the following identity: Denote* $\Phi=\delta\mathrm{D}$*, then*

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

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

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

This means that under the commutative assumption, the generalized Gamma z calculus agrees with the classical one [16] 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, i.e. proposition 1.13 and 1.14. Here we first prove the following fact.

Proposition 5.14. *Suppose* $\Gamma_2+\Gamma_2^{z,\mathbf{Vol}}\succeq\kappa(\Gamma_1+\Gamma_1^z)$ *with* $\kappa>0$*. Denote* ρ_t *as the solution of sub-Riemannian gradient flow* (5.92)*, then*

$$\frac{d}{dt}\Big(I_a(\rho_t)+I_z(\rho_t)\Big)\le-2\lambda\Big(I_a(\rho_t)+I_z(\rho_t)\Big).$$

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

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

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

$$\frac{d}{dt}\mathrm{I}_a(\rho_t) \geq -2\kappa \mathrm{I}_a(\rho_t),$$

We construct the other Lynapov function

$$\mathrm{I}_{a,z}(\rho) = \mathrm{I}_a(\rho) + \mathrm{I}_z(\rho).$$

Thus along the SDM gradient flow (5.92), we have

$$\frac{d}{dt}\mathrm{I}_{a,z}(\rho_t) = -2\int \Big(\Gamma_2(\delta \mathrm{D}, \delta \mathrm{D}) + \Gamma_2^{z,\mathbf{Vol}}(\delta \mathrm{D}, \delta \mathrm{D})\Big)\rho_t dx.$$

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

$$\frac{d}{dt}\mathrm{I}_{a,z}(\rho_t) \leq -2\kappa \mathrm{I}_{a,z}(\rho_t).$$

Proposition 1.13 follows directly by the Gronwall's equality. We next prove Proposition 1.14. Since

$$-\frac{d}{dt}\mathrm{D}(\rho_t) = \mathrm{I}_a(\rho_t) \leq \mathrm{I}_{a,z}(\rho_t),$$

then (5.3) implies the fact that, denote $\rho_0 = \rho$, then

$$\begin{aligned}
-\mathrm{I}_{a,z}(\rho) =& \int_0^\infty \frac{d}{dt}\mathrm{I}_{a,z}(\rho_t)dt \\
\leq& -2\kappa \int_0^\infty \mathrm{I}_{a,z}(\rho_t)dt = -2\kappa \int_0^\infty \Big(\mathrm{I}_a(\rho_t) + \mathrm{I}_z(\rho_t)\Big)dt \\
\leq& -2\kappa \int_0^\infty \mathrm{I}_a(\rho_t)dt \\
=& -2\kappa \int_0^\infty (-\frac{d}{dt}\mathrm{D}(\rho_t))dt \\
=& -2\kappa \mathrm{D}(\rho).
\end{aligned}$$

Thus $\mathrm{I}_{a,z}(\rho) \geq 2\kappa \mathrm{D}(\rho)$. Hence we prove all results by the fact that $\mathfrak{R}_a^G + \mathfrak{R}_z + \mathfrak{R}^{\mathbf{Vol}} \succeq \kappa(\Gamma_1 + \Gamma_1^z)$ implies $\Gamma_2 + \Gamma_2^{z,\mathbf{Vol}} \succeq \kappa(\Gamma_1 + \Gamma_1^z)$. In a word, the generalized Gamma z calculus implies the z-Log-Sobolev equality (zLSI):

$$\mathfrak{R}_a^G + \mathfrak{R}_z + \mathfrak{R}^{\mathbf{Vol}} \succeq \kappa(\Gamma_1 + \Gamma_1^z) \Rightarrow \frac{d}{dt}\mathrm{I}_{a,z}(\rho_t) \leq -2\kappa \mathrm{I}_{a,z}(\rho_t) \Rightarrow \text{zLSI}.$$

Remark 5.15. *Our Gamma z calculus and related convergence, functional inequality results can be formulated with gradient drift diffusion process*

$$dX_t = -a(X_t)a(X_t)^{\mathsf{T}}\nabla V(X_t) + \sqrt{2}\sum_{i=1}^{n} a_i(X_t)\circ dB_t^i,$$

where $V\colon \mathbb{M}^{n+m}\to\mathbb{R}$ *is a smooth function. In this case, the invariant measure is given by*

$$\rho^*(x) = \frac{1}{C}e^{-V(x)}\mathbf{Vol}(x),$$

where C *is the normalization constant* $C = \int e^{-V(x)}\mathbf{Vol}(x)dx < \infty$*. In this case, the Fokker-Planck equation of gradient drift diffusion process is the SDM gradient flow*

$$\begin{aligned}\partial_t\rho_t =&\nabla\cdot(\rho_t aa^{\mathsf{T}}\nabla\log\frac{\rho_t}{e^{-V}\mathbf{Vol}})\\ =&\nabla\cdot(aa^{\mathsf{T}}\nabla\rho_t) + \nabla\cdot(\rho_t a\otimes\nabla a) + \nabla\cdot(\rho_t aa^{\mathsf{T}}\nabla V).\end{aligned}$$

Following the derivation in this subsection, we propose the following generalized Gamma z operator:

$$\begin{aligned}\Gamma_2^{z,\rho^*}(f,g) \quad = \quad & \frac{1}{2}\Big[L\Gamma_1^z(f,g) - \Gamma_1^z(Lf,g) - \Gamma_1^z(f,Lg)\Big]\\ & +\mathbf{div}_z^{\rho^*}(\Gamma_{1,\nabla(aa^{\mathsf{T}})}(f,g)) - \mathbf{div}_a^{\rho^*}(\Gamma_{1,\nabla(zz^{\mathsf{T}})}(f,g)).\end{aligned}$$

Similarly, we can derive the related Ricci curvature tensor and generalized curvature dimension bound in Theorem 1.9.

In fact, our result can also be extended to the case when there is no symmetric invariant measure for $a\circ dB_t$*. As long as the Fokker-Planck equation of associated drift-diffusion process satisfies*

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

where ρ^* *is the corresponding invariant measure. Then our generalized Gamma z calculus and related curvature dimension bound follows.*

Remark 5.16. *It is worth mentioning with our derivation of Gamma z calculus is no longer a simple Hessian operator of entropy in SDM. In fact, it combines both the second order calculus in SDM and the property of* L^2 *Hessian operator of entropy. See similar relations in mean-field Bakry-Émery calculus proposed in [35].*

6 Discussions

In this paper, we propose a generalized Bakry-Émery gamma z calculus and curvature dimension bound in general sub-Riemannian manifolds. Our method is based on the sub-Riemannian density manifold. In SDM, we study Fokker-Planck equations as sub-Riemannian gradient flows. And we demonstrate the equivalence among the second-order geometric calculation of relative entropy in SDM and the proposed Gamma z calculus. Several functional inequalities are derived naturally following this equivalence.

On the other hand, our study demonstrates that the noise in SDEs, especially degenerate diffusion coefficients, introduces a particular sub-Riemannian structure. The proposed generalized curvature dimension bound governs its convergence property. Besides, this degenerate diffusion coefficient introduces a specific sub-Riemannian metric tensor in density manifold. We leave the study of sub-Riemannian density manifold in future works. Here we also derive a general strategy to find the curvature dimension bound for degenerate drift-diffusion processes. In the future, we shall study the convergence properties of degenerate SDEs, and establish the related functional inequalities.

7 Appenndix

Proof [Proof of Lemma 3.23]

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

$$\begin{aligned} L\Gamma_1^z(f,f) &= \Delta_p\Gamma_1^z(f,f) - A\nabla\Gamma_1^z(f,f) \\ \Gamma_1^z(Lf,f) &= \Gamma_1^z(\Delta_p f,f) - \Gamma_1^z(A\nabla f,f). \end{aligned}$$

By our definition above, we directly get

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

where we denote

$$\begin{aligned}
F^z &= a^{\mathsf{T}}\nabla\langle z^{\mathsf{T}}\nabla f, z^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^m} \\
&= a^{\mathsf{T}}\nabla\sum_{l=1}^{m}(\sum_{\hat{l}=1}^{n+m} z^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}f)^2 \\
&= \left(\sum_{\hat{k}=1}^{n+m} a^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}\sum_{l=1}^{m}(\sum_{\hat{l}=1}^{n+m} z^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}f)^2\right)_{k=1,\cdots,n} \\
&= (F^z_1, F^z_2, \cdots, F^z_n)^T.
\end{aligned}$$

We end up with

$$\begin{aligned}
&\Delta_P\Gamma^z_1(f,f) \\
= &\sum_{\hat{i}=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}\left(\sum_{\hat{k}=1}^{n+m} a^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}\sum_{l=1}^{m}(\sum_{\hat{l}=1}^{n+m} z^{\mathsf{T}}_{l\hat{l}}\frac{\partial}{\partial x_{\hat{l}}}f)^2\right)\right) + a^{\mathsf{T}}\nabla\circ(a^{\mathsf{T}}\nabla(z^{\mathsf{T}}\nabla f)^2)) \\
= &\sum_{k=1}^{n}\sum_{\hat{i}=1}^{n+m}\left(\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}\left(\sum_{\hat{k}=1}^{n+m} a^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}(z^{\mathsf{T}}\nabla f)^2\right)\right) + (a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla(z^{\mathsf{T}}\nabla f)^2) \\
= &\nabla a\circ(a^{\mathsf{T}}\nabla(z^{\mathsf{T}}\nabla f)^2) + (a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla(z^{\mathsf{T}}\nabla f)^2). \qquad (7.96)
\end{aligned}$$

Next, we compute the following quantity.

$$\Gamma^z_1(\Delta_P f, f) = \langle z^{\mathsf{T}}\nabla(\nabla\cdot(aa^{\mathsf{T}}\nabla f)), z^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^m}.$$

From Lemma 3.11, we have

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

We continue with our computation as below,

$$\begin{aligned}
&\Gamma^z_1(\Delta_P f, f) \\
= &\langle z^{\mathsf{T}}\nabla\left[\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{k=1}^{n}\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}(a^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}f) + (a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)\right], z^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^m} \\
= &\langle z^{\mathsf{T}}\nabla\left[\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{k=1}^{n}\frac{\partial}{\partial x_{\hat{i}}}a_{\hat{i}k}(a^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{k}}}f)\right], z^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^m} + \langle z^{\mathsf{T}}\nabla((a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)), z^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^m} \\
= &\langle z^{\mathsf{T}}\nabla\left(\nabla a\cdot(a^{\mathsf{T}}\nabla f)\right), z^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^m} + \langle z^{\mathsf{T}}\nabla((a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)), z^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^m} \\
= &\langle\left(z^{\mathsf{T}}\nabla\nabla a\cdot(a^{\mathsf{T}}\nabla f)\right), z^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^m} + \langle\left(\nabla a\cdot(z^{\mathsf{T}}\nabla(a^{\mathsf{T}}\nabla f))\right), z^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^m} \\
&+\langle z^{\mathsf{T}}\nabla((a^{\mathsf{T}}\nabla)\circ(a^{\mathsf{T}}\nabla f)), z^{\mathsf{T}}\nabla f\rangle_{\mathbb{R}^m}. \qquad (7.97)
\end{aligned}$$

From the above, combining (7.96) and (7.97) we further get

$$
\begin{aligned}
& \frac{1}{2}\Delta_P \Gamma_1^z(f,f) - \Gamma_1^z(\Delta_p f, f) \\
= \; & \frac{1}{2}(a^\intercal\nabla \circ (a^\intercal \nabla |z^\intercal \nabla f|^2)) - \langle z^\intercal \nabla((a^\intercal\nabla)\circ(a^\intercal \nabla f)), z^\intercal \nabla f\rangle_{\mathbb{R}^m} \\
& + \frac{1}{2}\sum_{\hat{i}=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k}\left(\sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^\intercal \frac{\partial}{\partial x_{\hat{j}}}\sum_{l=1}^{n}(\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{l}}} f)^2\right)\right) \\
& - \langle z^\intercal \nabla \left[\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{k=1}^{n}\frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k}(a_{k\hat{k}}^\intercal \frac{\partial}{\partial x_{\hat{k}}} f)\right], z^\intercal \nabla f\rangle_{\mathbb{R}^m} \\
= \; & \frac{1}{2}(a^\intercal\nabla \circ (a^\intercal \nabla |z^\intercal \nabla f|^2)) - \langle z^\intercal \nabla((a^\intercal\nabla)\circ(a^\intercal \nabla f)), z^\intercal \nabla f\rangle_{\mathbb{R}^m} \\
& + \sum_{\hat{i}=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k}\sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^\intercal \sum_{l=1}^{m}(\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{l}}} f)\frac{\partial}{\partial x_{\hat{k}}}(\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{l}}} f)\right)\cdots \mathbf{I} \\
& - \sum_{l=1}^{m}\left((\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{l}}} f)\left(\sum_{l'=1}^{n+m} z_{ll'}^\intercal \frac{\partial}{\partial x_{l'}}\left[\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{k=1}^{n}\frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k}(a_{k\hat{k}}^\intercal \frac{\partial}{\partial x_{\hat{k}}} f)\right]\right)\right)\cdots \mathbf{II}.
\end{aligned}
$$

Recall here, we denote $a^\intercal$ to emphasize the transpose of the matrix a and $a_{i\hat{i}}^\intercal = a_{\hat{i}i}$,

$$
\begin{aligned}
\mathbf{I} \; = \; & \sum_{\hat{i}=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k}\sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^\intercal \sum_{l=1}^{m}(\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{l}}} f)\frac{\partial}{\partial x_{\hat{k}}}(\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{l}}} f)\right) \\
= \; & \sum_{\hat{i}=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k}\sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^\intercal \sum_{l=1}^{m}(\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{l}}} f)(\sum_{\hat{l}=1}^{n+m} \frac{\partial}{\partial x_{\hat{k}}} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{l}}} f)\right) \\
& + \sum_{\hat{i}=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k}\sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^\intercal \sum_{l=1}^{m}(\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{l}}} f)(\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{k}}}\frac{\partial}{\partial x_{\hat{l}}} f)\right) \\
= \; & \sum_{l=1}^{m}(\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{l}}} f)\sum_{\hat{i}=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k}\sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^\intercal (\sum_{l'=1}^{n+m} \frac{\partial}{\partial x_{\hat{k}}} z_{ll'}^\intercal \frac{\partial}{\partial x_{l'}} f)\right) \\
& + \sum_{l=1}^{m}(\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^\intercal \frac{\partial}{\partial x_{\hat{l}}} f)\sum_{\hat{i}=1}^{n+m}\sum_{k=1}^{n}\left(\frac{\partial}{\partial x_{\hat{i}}} a_{\hat{i}k}\sum_{\hat{k}=1}^{n+m} a_{k\hat{k}}^\intercal (\sum_{\hat{l}=1}^{n+m} z_{ll'}^\intercal \frac{\partial}{\partial x_{\hat{k}}}\frac{\partial}{\partial x_{l'}} f)\right).
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{II} &= \sum_{l=1}^{m}\left((\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^{\mathsf{T}} \frac{\partial}{\partial x_{\hat{l}}} f) \left(\sum_{l'=1}^{n+m} z_{ll'}^{\mathsf{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}}^{\mathsf{T}} \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \\
&= \sum_{l=1}^{m}\left((\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^{\mathsf{T}} \frac{\partial}{\partial x_{\hat{l}}} f) \left(\sum_{l'=1}^{n+m} z_{ll'}^{\mathsf{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}}^{\mathsf{T}} \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \\
&\quad + \sum_{l=1}^{m}\left((\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^{\mathsf{T}} \frac{\partial}{\partial x_{\hat{l}}} f) \left(\sum_{l'=1}^{n+m} z_{ll'}^{\mathsf{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}}^{\mathsf{T}} \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \\
&= \sum_{l=1}^{m}\left((\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^{\mathsf{T}} \frac{\partial}{\partial x_{\hat{l}}} f) \left(\sum_{l'=1}^{n+m} z_{ll'}^{\mathsf{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}}^{\mathsf{T}} \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \\
&\quad + \sum_{l=1}^{m}\left((\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^{\mathsf{T}} \frac{\partial}{\partial x_{\hat{l}}} f) \left(\sum_{l'=1}^{n+m} z_{ll'}^{\mathsf{T}} \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}}^{\mathsf{T}} \frac{\partial}{\partial x_{l'}} \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right) \\
&\quad + \sum_{l=1}^{n}\left((\sum_{\hat{l}=1}^{n+m} z_{l\hat{l}}^{\mathsf{T}} \frac{\partial}{\partial x_{\hat{l}}} f) \left(\sum_{l'=1}^{n+m} z_{ll'}^{\mathsf{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}}^{\mathsf{T}} \frac{\partial}{\partial x_{\hat{k}}} f) \right] \right) \right).
\end{aligned}
$$

Subtracting the above two terms, we obtain the following

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

Now, we eventually end up with the following formula,

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

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 $A = a \otimes \nabla a$,

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

and

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

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

$$
\tilde{\mathbf{J}}_1 = -\sum_{l=1}^{m}\left(z^{\mathsf{T}}\nabla f\right)_l \left(\sum_{l',\hat{k}=1}^{n+m}\sum_{k=1}^{n}\sum_{k'=1}^{n+m} a_{\hat{k}k}\nabla_{\frac{\partial}{\partial x_{k'}}} a_{k'k}\nabla_{\frac{\partial}{\partial x_{\hat{k}}}} z_{ll'}\nabla_{\frac{\partial}{\partial x_{l'}}} f\right), \tag{7.98}
$$

$$
\begin{aligned}
\tilde{\mathbf{J}}_4 &= \sum_{l=1}^{m}(z^{\mathsf{T}}\nabla f)_l \left(\sum_{l'=1}^{n+m} z_{ll'}^{\mathsf{T}}(\sum_{\hat{k}=1}^{n+m}\nabla_{\frac{\partial}{\partial x_{l'}}}(\sum_{k=1}^{n}\sum_{k'=1}^{n+m} a_{\hat{k}k}\nabla_{\frac{\partial}{\partial x_{k'}}} a_{k'k})\nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f)\right) \\
&= \sum_{l=1}^{m}(z^{\mathsf{T}}\nabla f)_l \left(\sum_{k=1}^{n}\sum_{l'=1}^{n+m}\sum_{\hat{k},k'=1}^{n+m} z_{ll'}^{\mathsf{T}}\nabla_{\frac{\partial}{\partial x_{l'}}} a_{\hat{k}k}\nabla_{\frac{\partial}{\partial x_{k'}}} a_{k'k}\nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f\right) \\
&\quad + \sum_{l=1}^{m}(z^{\mathsf{T}}\nabla f)_l \left(\sum_{k=1}^{n}\sum_{l'=1}^{n+m}\sum_{\hat{k},k'=1}^{n+m} z_{ll'}^{\mathsf{T}} a_{\hat{k}k}(\nabla_{\frac{\partial}{\partial x_{l'}}}\nabla_{\frac{\partial}{\partial x_{k'}}} a_{k'k})\nabla_{\frac{\partial}{\partial x_{\hat{k}}}} f\right).
\end{aligned} \tag{7.99}
$$

Combine the above two proofs, we thus get

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

■

Proof [Proof of Lemma 3.24] We expand the two terms in lemma 3.24 .

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

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

Next, we expand $\widetilde{\mathbf{R}}_1$ and $\widetilde{\mathbf{R}}_2$ completely and get the following.

$$\begin{aligned}
\widetilde{\mathbf{R}}_1 \;=\;& \sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle (a^{\mathsf{T}}_{ii'}\frac{\partial}{\partial x_{i'}})(a^{\mathsf{T}}_{i\hat{i}}\frac{\partial}{\partial x_{\hat{i}}})(z^{\mathsf{T}}_{k\hat{k}}\frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{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^{\mathsf{T}}_{ii'}(\frac{\partial a^{\mathsf{T}}_{i\hat{i}}}{\partial x_{i'}}\frac{\partial z^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \cdots \widetilde{\mathbf{R}}_1^1 \\
& + \sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{i}}(\frac{\partial}{\partial x_{i'}}\frac{\partial z^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}})(\frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \cdots \widetilde{\mathbf{R}}_1^2 \\
& + \sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{i}}(\frac{\partial z^{\mathsf{T}}_{k\hat{k}}}{\partial x_{\hat{i}}})(\frac{\partial}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \cdots \widetilde{\mathbf{R}}_1^3 \\
& + \sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle (a^{\mathsf{T}}_{ii'})((\frac{\partial}{\partial x_{i'}}a^{\mathsf{T}}_{i\hat{i}})z^{\mathsf{T}}_{k\hat{k}}\frac{\partial}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \cdots \widetilde{\mathbf{R}}_1^4 \\
& + \sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{i}}(\frac{\partial}{\partial x_{i'}}z^{\mathsf{T}}_{k\hat{k}})\frac{\partial}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \cdots \widetilde{\mathbf{R}}_1^5 \\
& + \sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a^{\mathsf{T}}_{ii'}a^{\mathsf{T}}_{i\hat{i}}z^{\mathsf{T}}_{k\hat{k}}(\frac{\partial}{\partial x_{i'}}\frac{\partial}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m} \cdots \widetilde{\mathbf{R}}_1^6
\end{aligned}$$

$$
\begin{aligned}
\widetilde{\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}}^{\mathsf{T}}\frac{\partial}{\partial x_{\hat{k}}})[(a_{ii'}^{\mathsf{T}}\frac{\partial}{\partial x_{i'}})(a_{i\hat{i}}^{\mathsf{T}}\frac{\partial f}{\partial x_{\hat{i}}})],(z^{\mathsf{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}}^{\mathsf{T}}\frac{\partial a_{ii'}^{\mathsf{T}}}{\partial x_{\hat{k}}}\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{i}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\cdots\widetilde{\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}}^{\mathsf{T}}a_{ii'}^{\mathsf{T}}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}})\frac{\partial f}{\partial x_{\hat{i}}},(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\cdots\widetilde{\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}}^{\mathsf{T}}a_{ii'}^{\mathsf{T}}\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial f}{\partial x_{\hat{i}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\cdots\widetilde{\mathbf{R}}_2^3=\widetilde{\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}}^{\mathsf{T}}\frac{\partial a_{ii'}^{\mathsf{T}}}{\partial x_{\hat{k}}}a_{i\hat{i}}^{\mathsf{T}}(\frac{\partial}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{i}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\cdots\widetilde{\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}}^{\mathsf{T}}a_{ii'}^{\mathsf{T}}\frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{\hat{k}}}(\frac{\partial}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{i}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\cdots\widetilde{\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}}^{\mathsf{T}}a_{ii'}^{\mathsf{T}}a_{i\hat{i}}^{\mathsf{T}}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial}{\partial x_{i'}}\frac{\partial f}{\partial x_{\hat{i}}}),(z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^m}\cdots\widetilde{\mathbf{R}}_2^6=\widetilde{\mathbf{R}}_1^6
\end{aligned}
$$

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

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

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

$$\begin{aligned}
2\widetilde{\mathbf{R}}_1^3 &= 2\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a_{i\hat{i}}^{\mathsf{T}} a_{ii'}^{\mathsf{T}} (\frac{\partial z_{k\hat{k}}^{\mathsf{T}}}{\partial x_{i'}})(\frac{\partial}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
&= 2\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{i',l=1}^{n+m}\left(a_{i\hat{i}}^{\mathsf{T}} a_{ii'}^{\mathsf{T}} (\frac{\partial z_{k\hat{k}}^{\mathsf{T}}}{\partial x_{i'}})(\frac{\partial}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}) z_{kl}^{\mathsf{T}}\frac{\partial f}{\partial x_l}\right) \\
-2\widetilde{\mathbf{R}}_2^5 &= -2\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle z_{ki'}^{\mathsf{T}} a_{i\hat{k}}^{\mathsf{T}} \frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial f}{\partial x_{\hat{i}}}), (z^{\mathsf{T}}\nabla)_k f\rangle \\
&= -2\sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{\hat{i},\hat{k}=1}^{n+m}\sum_{i',l=1}^{n+m}\left(z_{ki'}^{\mathsf{T}} a_{i\hat{k}}^{\mathsf{T}} \frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial f}{\partial x_{\hat{i}}}) z_{kl}^{\mathsf{T}}\frac{\partial f}{\partial x_l}\right)
\end{aligned}$$

We denote

$$\sum_{\hat{i},\hat{k}=1}^{n+m} a_{i\hat{i}}^{\mathsf{T}} z_{k\hat{k}}^{\mathsf{T}} \frac{\partial^2 f}{\partial x_{\hat{i}}\partial x_{\hat{k}}} = \omega_{ik}. \tag{7.100}$$

The above equality (7.100) can be represented in the following matrix form

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

where P and X are defined in (3.52) and (3.20). Now, we can represent term $\widetilde{\mathbf{T}}_{1a}$ as $\sum_{k=1}^{n}\sum_{k=1}^{m}\omega_{ik}^2 = \omega^T\omega = (PX)^T PX = X^T P^T PX$. Next we want to represent $\widetilde{\mathbf{R}}_1^3$ and $\widetilde{\mathbf{R}}_2^5$ in the following form in terms of vector X,

$$\begin{aligned}
& 2\widetilde{\mathbf{R}}_1^3 - 2\widetilde{\mathbf{R}}_2^5 \\
= \; & 2\sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle a_{i\hat{i}}^{\mathsf{T}} a_{ii'}^{\mathsf{T}} (\frac{\partial z_{k\hat{k}}^{\mathsf{T}}}{\partial x_{i'}})(\frac{\partial}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}), (z^{\mathsf{T}}\nabla)_k f\rangle_{\mathbb{R}^n} \\
& -2\sum_{i,k=1}^{n}\sum_{i',\hat{i},\hat{k}=1}^{n+m}\langle z_{ki'}^{\mathsf{T}} a_{i\hat{k}}^{\mathsf{T}} \frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}(\frac{\partial}{\partial x_{\hat{k}}}\frac{\partial f}{\partial x_{\hat{i}}}), (z^{\mathsf{T}}\nabla)_k f\rangle \\
= \; & 2\sum_{\hat{i},\hat{k}=1}^{n+m}\left[\sum_{i,k=1}^{n}\sum_{i'=1}^{n+m}\left(\langle a_{i\hat{i}}^{\mathsf{T}} a_{ii'}^{\mathsf{T}} (\frac{\partial z_{k\hat{k}}^{\mathsf{T}}}{\partial x_{i'}}), (z^{\mathsf{T}}\nabla)_k f\rangle - \langle z_{ki'}^{\mathsf{T}} a_{i\hat{k}}^{\mathsf{T}} \frac{\partial a_{i\hat{i}}^{\mathsf{T}}}{\partial x_{i'}}, (z^{\mathsf{T}}\nabla)_k f\rangle\right)\right](\frac{\partial}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}) \\
= \; & 2F^T X,
\end{aligned}$$

where F is defined in (3.53). Similarly, we can represent $\widetilde{\mathbf{T}}_{1b} = \widetilde{\mathbf{T}}_{1c}$ by X,

$$\begin{aligned}
\widetilde{\mathbf{T}}_{1b} = \widetilde{\mathbf{T}}_{1c} &= \sum_{i,k=1}^{n}\langle\sum_{\hat{i},\hat{k}=1}^{n+m} a_{i\hat{i}}^{\mathsf{T}}\frac{\partial z_{k\hat{k}}^{\mathsf{T}}}{\partial x_{\hat{i}}}\frac{\partial f}{\partial x_{\hat{k}}}, \sum_{i',k'=1}^{n+m} a_{ii'}^{\mathsf{T}} z_{kk'}^{\mathsf{T}}\frac{\partial^2 f}{\partial x_{i'}\partial x_{k'}}\rangle \\
&= E^T PX
\end{aligned}$$

where E is defined in (3.54). We thus have the following form,

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

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 end up with

$$\begin{aligned} & \widetilde{\mathbf{T}}_1 + \widetilde{\mathbf{R}}_1 - \widetilde{\mathbf{R}}_2 \\ = \quad & \widetilde{\mathbf{T}}_1 + 2\widetilde{\mathbf{R}}_1^3 - 2\widetilde{\mathbf{R}}_2^5 + \widetilde{\mathbf{R}}_1^1 + \widetilde{\mathbf{R}}_1^2 - \widetilde{\mathbf{R}}_2^1 - \widetilde{\mathbf{R}}_2^2, \end{aligned}$$

which completes the proof.

References

[1] A. Agrachev, D. Barilari, and U. Boscain. On the Hausdorff volume in sub-riemannian geometry. *Calculus of Variations and Partial Differential Equations*, 43(3-4):355–388, 2012.

[2] A. Agrachev, D. Barilari, and U. Boscain. A comprehensive introduction to sub-riemannian geometry, 2019.

[3] A. Agrachev and P. Lee. Optimal transportation under nonholonomic constraints. *Transactions of the American Mathematical Society*, 361(11):6019–6047, 2009.

[4] A. Arnold, A. Einav, and T. Whrer. On the rates of decay to equilibrium in degenerate and defective Fokker-Planck equations. *Journal of Differential Equations*, 264(11):6843 – 6872, 2018.

[5] D. Bakry. Ricci curvature and dimension for diffusion semigroups. *Stochastic Processes and their Applications*, pages 21–31, 1990.

[6] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de Probabilités XIX 1983/84*, pages 177–206. Springer, 1985.

[7] D. Bakry, M. Ledoux, et al. A logarithmic sobolev form of the Li-Yau parabolic inequality. *Revista Matemática Iberoamericana*, 22(2):683–702, 2006.

[8] Z. M. Balogh, A. Kristály, and K. Sipos. Geometric inequalities on heisenberg groups. *Calculus of variations and partial differential equations*, 57(2):61, 2018.

[9] D. Barilari and L. Rizzi. A formula for popps volume in sub-riemannian geometry. *Analysis and Geometry in Metric Spaces*, 1:42–57, 2013.

[10] D. Barilari and L. Rizzi. Sub-riemannian interpolation inequalities. *Inventiones mathematicae*, 215(3):977–1038, 2019.

[11] F. Baudoin. Sub-laplacians and hypoelliptic operators on totally geodesic Riemannian foliations. *arXiv preprint arXiv:1410.3268*, 2014.

[12] F. Baudoin. Wasserstein contraction properties for hypoelliptic diffusions. *arXiv preprint arXiv:1602.04177*, 2016.

[13] F. Baudoin. Bakry–émery meet villani. *Journal of Functional Analysis*, 273(7):2275 – 2291, 2017.

[14] F. Baudoin, M. Bonnefont, and N. Garofalo. A sub-Riemannian curvature-dimension inequality, volume doubling property and the Poincaré inequality. *Mathematische Annalen*, 358(3-4):833–860, 2014.

[15] F. Baudoin, Q. Feng, and M. Gordina. Integration by parts and quasi-invariance for the horizontal Wiener measure on foliated compact manifolds. *Journal of Functional Analysis*, 2019.

[16] F. Baudoin and N. Garofalo. Curvature-dimension inequalities and Ricci lower bounds for sub-riemannian manifolds with transverse symmetries. *Journal of the EMS, Vol. 19, Issue 1*, 2017.

[17] F. Baudoin, M. Gordina, and D. P. Herzog. Gamma calculus beyond Villani and explicit convergence estimates for langevin dynamics with singular potentials. *arXiv preprint arXiv:1907.03092*, 2019.

[18] F. Baudoin, E. Grong, K. Kuwada, and A. Thalmaier. Sub-Laplacian comparison theorems on totally geodesic Riemannian foliations. *arXiv:1706.08489 [math]*, 2017.

[19] F. Baudoin and J. Wang. Curvature dimension inequalities and subelliptic heat kernel gradient bounds on contact manifolds. *arXiv preprint arXiv:1211.3778*, 2012.

[20] D.-C. Chang and Y. Li. Subriemannian geodesics in the Grushin plane. *Journal of Geometric Analysis*, 22(3):800–826, 2012.

[21] N. Eldredge, M. Gordina, and L. Saloff-Coste. Left-invariant geometries on su(2) are uniformly doubling. *Geometric and Functional Analysis*, 28(5):1321–1367, 2018.

[22] K. D. Elworthy. *Stochastic differential equations on manifolds*, volume 70. Cambridge University Press, 1982.

[23] Q. Feng. Harnack inequalities on totally geodesic foliations with transverse Ricci flow. *arXiv:1712.02275 [math]*, 2017.

[24] A. Figalli and L. Rifford. Mass transportation on sub-riemannian manifolds. *Geometric and functional analysis*, 20(1):124–159, 2010.

[25] M. Gordina and T. Laetsch. Sub-laplacians on sub-riemannian manifolds. *Potential Analysis*, 44(4):811–837, 2016.

[26] M. Gordina and T. Laetsch. A convergence to brownian motion on sub-riemannian manifolds. *Transactions of the American Mathematical Society*, 369(9):6263–6278, 2017.

[27] E. Grong and A. Thalmaier. Curvature-dimension inequalities on sub-riemannian manifolds obtained from riemannian foliations: part i. *Mathematische Zeitschrift*, 282(1-2):99–130, Sep 2015.

[28] E. Grong and A. Thalmaier. Curvature-dimension inequalities on sub-riemannian manifolds obtained from riemannian foliations: part ii. *Mathematische Zeitschrift*, 282(1-2):131–164, Sep 2015.

[29] V. V. Grushin. On a class of hypoelliptic operators. *Matematicheskii Sbornik*, 125(3):456–473, 1970.

[30] N. Juillet. Diffusion by optimal transport in Heisenberg groups. *Calculus of Variations and Partial Differential Equations*, 50(3-4):693–721, 2014.

[31] A. Jüngel. *Entropy methods for diffusive partial differential equations*. Springer, 2016.

[32] B. Khesin and P. Lee. A nonholonomic Moser theorem and optimal transport. *Journal of Symplectic Geometry*, 7(4):381–414, 2009.

[33] J. D. Lafferty. The Density Manifold and Configuration Space Quantization. *Transactions of the American Mathematical Society*, 305(2):699–741, 1988.

[34] W. Li. Geometry of probability simplex via optimal transport. *arXiv:1803.06360 [math]*, 2018.

[35] W. Li. Diffusion Hypercontractivity via Generalized Density Manifold. *arXiv:1907.12546 [cs, math]*, 2019.

[36] J. Lott and C. Villani. Ricci Curvature for Metric-Measure Spaces via Optimal Transport. *Annals of Mathematics*, 169(3):903–991, 2009.

[37] P. Malliavin. *Stochastic analysis*, volume 313 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1997.

[38] F. Otto and C. Villani. Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality. *Journal of Functional Analysis*, 173(2):361–400, 2000.

[39] L. Rizzi. A counterexample to gluing theorems for mcp metric measure spaces. *Bulletin of the London Mathematical Society*, 50(5):781–790, 2018.

[40] K.-T. Sturm. On the Geometry of Metric Measure Spaces. *Acta Mathematica*, 196(1):65–131, 2006.

Notations

In this paper, we apply the following notations.

Sub-Riemannian manifold	$\mathbb{M}^{n+m}$
Horizontal sub-bundle	τ
Volume	$\mathbf{Vol}(x)$
Transition kernel, Definition 4.1	$\rho(t,\cdot,\cdot) = p(t,\cdot,\cdot)\mathbf{Vol}$
$(n+m)\times n$ matrix	$a = (a_1, \cdots, a_n)$
$(n+m)\times m$ matrix	$c = (c_1, \cdots, c_m)$
$(n+m)\times m$ matrix	$z = (z_1, \cdots, z_m)$
Laplace operator with matrix a	$\Delta_p = \nabla\cdot(aa^{\mathsf{T}}\nabla)$
Notation A (2.7)	$A = a\otimes\nabla a$
Operator for Fokker-Planck equation	$L^* f = \Delta_p f + \nabla\cdot(fA)$
Dual operator of L^*	$Lf = \Delta_p f - A\nabla f$
Operator L with drift b	$\tilde{L}f = Lf + 2b\nabla f$
Generalized Bakry-Émery calculus	
Γ_1 operator for Δ_p, (3.15)	$\Gamma_{1,a}(\cdot,\cdot) = \Gamma_1(\cdot,\cdot)$
Γ_2 operator for Δ_p, (3.17)	$\Gamma_{2,a}(\cdot,\cdot)$
Γ_1 operator for L, (3.16)	$\Gamma_1(\cdot,\cdot)$
Γ_2 operator for L, (3.18)	$\Gamma_2(\cdot,\cdot)$
Γ_1 operator for $\tilde{L}$, (3.47)	$\Gamma_{1,\tilde{L}}(\cdot,\cdot)$
Γ_2 operator for $\tilde{L}$, (3.48)	$\Gamma_{2,\tilde{L}}(\cdot,\cdot)$
Γ_1 operator for matrix z, (3.57)	$\Gamma_1^z(\cdot,\cdot)$
Γ_2 operator for L with matrix z, (3.58)	$\Gamma_2^z(\cdot,\cdot)$
Γ_2 operator for $\tilde{L}$ with matrix z, (3.62)	$\Gamma_{2,\tilde{L}}^z(\cdot,\cdot)$
Special convention in (3.35)	$\Gamma_{2,\mathcal{H}}(\cdot,\cdot)$
Generalized Gamma 2 z, (3.65)	$\Gamma_2^{z,\Psi}(f,f)$
Generalized Gamma 2 z with drift, (3.77)	$\Gamma_{2,\tilde{L}}^{z,\Psi}(f,f)$
New Hessian and Curvature for Γ_2,Thm 3.4	$\mathfrak{Hess}_a$, $\mathfrak{R}_a(\cdot,\cdot)$
New Hessian and Curvature for $\Gamma_{2,\tilde{L}}$, Thm 3.16	$\mathfrak{Hess}_a$, $\mathfrak{R}_{a,b}(\cdot,\cdot)$
New Hessian and Curvature for $\Gamma_2+\Gamma_2^z$, Thm 3.25	$\mathfrak{Hess}_{a,z}$, $\mathfrak{R}_a(\cdot,\cdot)+\mathfrak{R}_z(\cdot,\cdot)$
New Hessian and Curvature for $\Gamma_{2,\tilde{L}}+\Gamma_{2,\tilde{L}}^z$, Thm 3.26	$\mathfrak{Hess}_{a,z}$, $\mathfrak{R}_{a,b}(\cdot,\cdot)+\mathfrak{R}_{z,b}(\cdot,\cdot)$
New Hessian and Curvature for $\Gamma_2+\Gamma_2^{z,\Psi}(f,f)$, Thm 3.32	$\mathfrak{hess}_{a,z,}^G, \mathfrak{R}_a^G(\cdot,\cdot), \mathfrak{R}_z(\cdot,\cdot), \mathfrak{R}^\Psi(\cdot,\cdot)$
New Hessian and Curvature for $\Gamma_{2,\tilde{L}}+\tilde{\Gamma}_{2,\tilde{L}}^{z,\Psi}(f,f)$, Thm 3.35	$\mathfrak{hess}_{a,z,}^G, \mathfrak{R}_{a,b}^G(\cdot,\cdot), \mathfrak{R}_{z,b}(\cdot,\cdot), \mathfrak{R}_b^\Psi(\cdot,\cdot)$

Diffusion	$a \circ dB_t$
Drift	$b(\cdot)dt$
Brownian motion	$(B_t^i)_{0\le t\le T}$, $i = 1, \cdots, n.$
Metric	g
Pseudo-inverse of aa^T	$(aa^\mathsf{T})^\dagger$
Horizontal metric	$g_\tau = (aa^\mathsf{T})^\dagger$
Divergence operator	$\nabla\cdot$
Gradient operator	∇
Hessian operator	Hess
Notation 3.2	Q, X, C, D
Notation 3.20	P, F, E
Notation 3.29	$G, \mathfrak{R}^\Psi$
Assumption 3.3, 3.22,3.30	$\Lambda, \widetilde{\Lambda}, \widehat{\Lambda}, \Theta$
Convention 3.33	$\mathbf{B}$, $\mathbf{B}_0$
Density manifold	$\mathcal{P}$
Probability density	ρ
Reference density	μ
Tangent space	$T_\rho\mathcal{P}$
Cotangent space	$T_\rho^*\mathcal{P}$
Density manifold metric tensor	$g_\rho^{W_a}$
Weighted Laplacian operator	$\Delta_\rho^a = \nabla\cdot(\rho aa^\mathsf{T}\nabla)$
	$\Delta_\rho^z = \nabla\cdot(\rho zz^\mathsf{T}\nabla)$
First L^2 variation	δ
Second L^2 variation	δ^2
Gradient operator	grad_{W_a}
Hessian operator	Hess_{W_a}
Christoffel symbol	$\Gamma_\rho(\cdot,\cdot)$
Tangent bundle	$(\rho,\sigma)\in T\mathcal{P}$
Cotangent bundle	$(\rho,\Phi)\in T^*\mathcal{P}$
Negative Boltzmann-Shannon entropy	H
Relative entropy	D
sub-Riemannian Fisher information	I_a
z Fisher information	I_z
a, z Fisher information	$\mathrm{I}_{a,z}(\rho) = \mathrm{I}_a(\rho) + \mathrm{I}_z(\rho)$