黎曼朗之万动力学：几何欧拉–丸山格式的强收敛

Zhiyuan Zhan zhan@ms.k.u-tokyo.ac.jp
日本东京大学、日本理化学研究所人工智能项目中心

杉山将 sugi@k.u-tokyo.ac.jp
日本理化学研究所人工智能项目中心、日本东京大学

arXiv:2603.03626v1 [stat.ML] 4 Mar 2026

摘要

真实世界数据中的低维结构对生成模型的成功发挥着重要作用，这促使人们在内在数据流形上定义扩散模型。这类模型由流形上的随机微分方程（SDE）驱动，因此有必要研究流形值随机微分方程数值格式的收敛理论。在欧几里得空间中，欧拉–丸山（EM）格式可达到阶数为 $1/2$ 的强收敛，但在一般情形下，流形离散化的类似结果尚未得到充分理解。本文研究黎曼流形上随机微分方程的一种几何版本的 EM 格式，并证明在满足几何与正则性条件时，其强收敛阶数为 $1/2$ 。作为应用，我们通过黎曼朗之万动力学的几何欧拉–丸山离散化，获得了流形上采样的瓦瑟斯坦界。

关键词：扩散模型，黎曼流形，黎曼朗之万动力学，几何欧拉–丸山格式，强收敛

1 引言

近年来，扩散模型在学习欧几里得空间上的数据分布方面展现出了强大的性能（Ho 等人，2020年；Song 等人，2021年；Dhariwal 和 Nichol，2021年）。其成功背后的一个关键因素，是能够捕捉实际数据集中的低维结构（Chen 等人，2023年；Oko 等人，2023年；Loaiza-Ganem 等人，2024年；Wang 等人，2025年）。人们普遍认为，高维现实世界数据集中分布在一个低维流形附近；这一现象被称为流形假设（Bengio 等人，2013年）。因此，在流形假设下理解欧几里得扩散模型的收敛速率，逐渐引起了越来越多的关注（De Bortoli，2022；Li 和 Yan，2024年；Farghly 等人，2025年）。与此同时，近期还有一些研究利用流形值随机微分方程（随机微分方程）中的技术，直接在数据流形上构建扩散模型，这类模型通常被称为黎曼扩散模型（De Bortoli 等人，2022；Huang 等人，2022；Jo 和 Hwang，2024，2025年）（Hsu，2002年）。

与欧几里得情形一样，黎曼扩散模型由一个前向过程和一个反向过程组成（De Bortoli 等人，2022；Huang 等人，2022）。这两个过程都是黎曼朗之万动力学（RLD）的实例；RLD 是一种典型的黎曼扩散过程，用于对相对于黎曼体积测度的密度正比于 $e^{-\phi}$ 的目标分布 μ_ϕ 进行采样（Wang 等人，2020；Cheng 等人，2022；Bharath 等人，2025）。具体而言，给定一个黎曼流形 $\mathcal{M}$ 以及定义在 $\mathcal{M}$ 上的光滑函数 ϕ ，RLD 定义为

由 $\mathcal{M}$-SDE 定义（Hsu，2002）

$$\mathrm{d}\boldsymbol{X}_t = V(\boldsymbol{X}_t)\,\mathrm{d}t + \mathrm{d}\boldsymbol{B}_t^{\mathcal{M}}, \quad t \in [0,T],$$

其中，$V = -\frac{1}{2}\nabla\phi$（$\mathcal{M}$上的梯度）是漂移向量场，$\boldsymbol{B}^{\mathcal{M}}$是 $\mathcal{M}$上的标准布朗运动。当 $\mathcal{M}$ 和 ϕ满足Bakry–Emery曲率条件（Villani，2008年）时，众所周知，μ_t 的 $\boldsymbol{X}_t$ 分布以指数速度收敛于μ_ϕ，并且是在 t Wasserstein 距离下收敛。为了在 $[0,T]$上离散化RLD，通常使用几何欧拉–马鲁雅马（GEM）格式（Jørgensen，1975年；Piggott 和 Solo，2016；De Bortoli 等，2022）：令 $h = T/N$，其中 $N \in \mathbb{N}$ 和 $t_k = kh$为某些量，定义

$$\boldsymbol{X}_{k+1}^h = \exp_{\boldsymbol{X}_k^h}\left(hV(\boldsymbol{X}_k^h) + \sqrt{h}\boldsymbol{\xi}_{\boldsymbol{X}_k^h}\right), \quad k = 0,\ldots,N-1,$$

其中 exp $_x$ 是指数映射，$\boldsymbol{\xi}_x$ 是切空间上的标准高斯分布。

GEM格式的弱收敛，即分布收敛，已经得到了充分研究（Jørgensen，1975年；Kuwada，2012）。与欧氏空间中的欧拉–马鲁雅马（EM）格式（Kloeden 和 Platen，1992年）类似，GEM实现了1阶弱收敛（Bharath 等，2025年）。相比之下，p-强收敛（路径意义下的收敛）

$$\mathbb{E}\left[\max_{0\le k\le N} d_{\mathcal{M}}(\boldsymbol{X}_k^h, \boldsymbol{X}_{t_k})^p\right]^{1/p} \lesssim h^{\gamma},$$

其中，$d_{\mathcal{M}}$ 是 $\mathcal{M}$上的内蕴黎曼距离，其研究尚不充分。据我们所知，在一般情形下，为 GEM 建立与欧氏 EM（Kloeden 和 Platen，1992年）相同的收敛阶 $\gamma = 1/2$ 仍是一个尚未解决的问题。现有研究要么聚焦于球面等特殊流形（Solo 和 Wang，2021年）、特殊正交群 SO$_n$（Piggott 和 Solo，2016；Wang 和 Solo，2025b），要么更广泛地研究李群（Marjanovic 和 Solo，2018；Muniz 等，2022、2023；Solo，2024）。其他研究则考察了较弱的逼近概念，例如 Wasserstein 距离的界（Cheng 等，2022；De Bortoli 等，2022）。

OUR RESULTS. 本文建立了嵌入黎曼子流形 $\mathcal{M} \subset \mathbb{R}^n$上 GEM 的强收敛性。在假设 $\mathcal{M}$ 具有有界外在曲率且在 $\mathbb{R}^n$中全局良好嵌入的条件下，我们证明 GEM 达到 p-强路径收敛，其收敛阶为 $\gamma = 1/2$，与经典欧氏 EM 的收敛速率相同。我们首先非正式地陈述主要结果如下。

定理 1（p-GEM 的强收敛性，见定理 7）设 $1 \le p < \infty$。假设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个具有有界外在曲率且整体良好嵌入的黎曼子流形，并且 V 满足光滑性条件。对于由 *GEM* 构造的 $\boldsymbol{X}_k^h$，我们有

$$\mathbb{E}\left[\max_{0\le k\le N} d_{\mathcal{M}}(\boldsymbol{X}_k^h, \boldsymbol{X}_{t_k})^p\right] \lesssim h^{p/2}.$$

基于 GEM 的 p-强收敛以及 Bakry–Emery 曲率条件，我们得到了 GEM 离散化下 RLD 的 p-Wasserstein 界。

定理 2（p-RLD 的 Wasserstein 收敛，见定理 14）设 $1 \le p < \infty$。假设 $\mathcal{M}$ 和 $V = -\frac{1}{2}\nabla\phi$ 满足定理 *1* 中的假设，且 ϕ 满足 *Bakry–Émery* 曲率条件。对于通过 *GEM* 构造的 $\boldsymbol{X}_k^h$，令 $\hat{\mu}_k$ 为 $\boldsymbol{X}_k^h$ 的分布。则有

$$\mathcal{W}_p(\mu_\phi, \hat{\mu}_N) \lesssim e^{-T} + h^{1/2}.$$

对于任意紧黎曼流形 $\mathcal{M}$ 及其上的任意向量场 V，我们证明，定理 1 中的假设均成立，且与如何通过纳什嵌入定理将 $\mathcal{M}$ 嵌入到 $\mathbb{R}^n$ 中无关。因此，对于任意 $p \ge 1$，紧黎曼流形上的 GEM 均具有 p-强收敛，其阶为 $1/2$（推论 8）。对于非紧情形，我们还给出了两个满足几何有界性要求的经典示例：图和水平集。

Technical overview. 由于 $\mathcal{M} \subset \mathbb{R}^n$，$\mathcal{M}$-布朗运动 $\boldsymbol{B}^{\mathcal{M}}$ 可以通过将 $\mathbb{R}^n$ 上的标准布朗运动投影来实现，而 $\mathcal{M}$-随机微分方程可以写成 $\mathbb{R}^n$ 上的随机微分方程的外在形式（Hsu, 2002）。这使得使用欧拉–马鲁雅马格式成为可能。然而，这会引出两个问题：(i) 外在形式的 $\mathbb{R}^n$-随机微分方程中的系数仅在 $\mathcal{M}$ 上有定义，因此其欧几里得EM离散化 $\boldsymbol{Y}_k^h$ 无法直接定义；(ii) 即使定义了 $\boldsymbol{Y}_k^h$，仍需控制外在 EM $\boldsymbol{Y}_k^h$ 与内在 GEM $\boldsymbol{X}_k^h$ 之间的强路径偏差。

为解决（i），我们研究 $\mathcal{M}$ 的外在几何，并利用 $\mathcal{M}$ 的几何有界性，结合漂移向量场 V 的正则性假设，构造一个在 $\mathbb{R}^n$ 上适定、且在 $\mathcal{M}$ 上与原始 $\mathcal{M}$-随机微分方程一致的 $\mathbb{R}^n$-随机微分方程。这使得欧几里得欧拉–马鲁雅马格式 $\boldsymbol{Y}_k^h$ 有良好定义，并使我们能够应用欧几里得空间中的经典强收敛结果。为解决（ii），我们对指数映射 $\exp_x$ 进行泰勒展开，并借助 $\mathcal{M}$ 的几何有界性一致地控制相应余项。这将 $\boldsymbol{Y}_k^h$ 与 $\boldsymbol{X}_k^h$ 之间的比较归结为欧几里得空间中的估计，且误差可由等价 $\mathbb{R}^n$-随机微分方程的几何性质加以控制。综合（i）与（ii）的结果，我们可以证明定理 1。

Contributions. 综上所述，我们的贡献如下。

（i）我们在几何有界性与漂移正则性条件下，建立了嵌入 $\mathbb{R}^n$ 的黎曼子流形上GEM的阶为 $1/2$ 的 p-强收敛。此外，结合Bakry–Émery曲率条件，我们为GEM离散化下的RLD提供了 $\mathcal{W}_p$ 保证。

（ii）在技术层面，我们研究 $\mathbb{R}^n$ 中嵌入黎曼子流形的外在几何，并建立外在延拓与比较框架，用于分析内在的流形值随机微分方程。

本文其余部分安排如下。第2节回顾相关研究，第3节介绍必要的预备知识。第 4 节给出GEM的 p 强收敛结果及相关结论，第 5 节在GEM离散化下应用这些结果，得到RLD的 p-Wasserstein界。第6节总结全文并讨论其局限性。记号汇总于附录A。

2 相关工作

Numerical solutions of manifold SDEs. 现有关于流形值随机微分方程数值格式（如GEM）强收敛的大多数结果，都聚焦于特殊类别的流形。例如，Piggott和Solo（2016）研究了特殊正交群SO$_n$ 和特殊欧氏群SE$_n$上的GEM，并证明对于任意小的 $\varepsilon > 0$，其 $p = 2$ 强收敛阶为 $\gamma = (1 - \varepsilon)/2$ 。Solo和Wang（2021）利用切空间参数化，证明了球面上具有相同的收敛阶。更一般地，基于李群的结构（Marjanovic等，2015；Marjanovic和Solo，2018），Muniz等（2022）和Solo（2024）在李群上建立了类似结果；Muniz等（2023）则考虑了流形带有李群作用的一种特殊情形。

对于较弱的逼近概念，Cheng 等人（2022）通过界定 $\mathbb{E}[d_{\mathcal{M}}(\boldsymbol{X}_N^h,\ \boldsymbol{X}_T)^p]$，得到了 GEM 的 p-Wasserstein 界，其中 $p = 1,\ 2$，收敛速率为 $\tilde{O}(h^{1/2})$，其中 $\tilde{O}$ 隐藏了对 $\log h$ 的多项式依赖。De Bortoli 等人（2022）将这一结果推广到时间非齐次漂移向量场，即时间依赖的 $V(t,\ x)$。

与上述工作相比，我们考虑了一类更一般的流形，即嵌入于 $\mathbb{R}^n$中的黎曼子流形，因为我们不要求具备李群结构等额外结构。此外，我们证明 GEM 以 $\mathcal{O}(h^{1/2})$ 的速率达到 p-强收敛（定理 1），这与欧氏 EM 情形一致。

Embedded manifolds. Armstrong 和 King（2022）以及 Laurent 和 Vilmart（2022）也将 $\mathcal{M} \subset \mathbb{R}^n$ 视为嵌入子流形，但他们预先假设随机微分方程系数在具有所需正则性的 $\mathbb{R}^n$ 上定义良好。我们研究内禀定义在 $\mathcal{M}$ 上的随机微分方程，并仅将外在延拓用作证明误差界的桥梁。这一观点也解释了为何我们的结果适用于任意紧黎曼流形（推论 8），而无需预先假设 $\mathcal{M}$ 是某个 $\mathbb{R}^n$的给定子流形。

此外，Armstrong 和 King（2022）没有采用GEM，而是使用欧氏常微分方程（ODE）来近似随机微分方程，因此所得离散过程不位于 $\mathcal{M}$上。Laurent 和 Vilmart（2022）研究了水平集情形，即对于某个 F，$\mathcal{M} = F^{-1}(0)$：$\mathbb{R}^n \to \mathbb{R}^k$，并考察了欧氏EM格式的遍历收敛。我们的工作聚焦于内在GEM格式，在该格式下，离散过程 $\boldsymbol{X}_k^h$ 保持在$\mathcal{M}$上，并建立了与欧氏EM格式类似的强收敛结果。

水平集情形也是 Wang 和 Solo (2025a) 的主要研究对象，他们使用GEM来近似 $\mathcal{M}$上的随机微分方程。他们仅考虑了以欧几里得范数衡量的二阶强收敛，并对任意小的 $\varepsilon > 0$得到了阶为 $\gamma = (1-\varepsilon)/2$ 的结果。与他们的工作相比，我们的结果适用于更一般的嵌入流形，并且还在 $d_{\mathcal{M}}$中给出了内禀误差界。特别地，我们将水平集情形作为一个特例纳入其中，并且针对该情形的假设与 Wang 和 Solo (2025a) 中的设定一致。

GEM for RLD. 对于紧流形，Wang 等人 (2020) 和 Li and Erdogdu (2023) 给出了 GEM 所诱导的分布与目标分布 μ_ϕ之间的 Kullback–Leibler（KL）散度界，并且通过 Talagrand 不等式还可推出二阶 Wasserstein 界。Bharath 等人 (2025) 研究了紧流形上 RLD 的 GEM 离散化的弱收敛和遍历收敛。对于非紧情形，Gatmiry and Vempala (2022) 研究了 Hessian 流形，并证明 GEM

对 RLD 的离散化具有 KL 收敛性。相比之下，我们的工作聚焦于强收敛，并通过结合 GEM 的 p-强收敛所带来的离散化误差与 RLD 在 Bakry–Emery曲率条件下的混合误差，得到了一个p-Wasserstein 界$\acute{\mu}_{\phi}$ 。

3 预备知识

本节简要回顾黎曼流形的一些基本概念，以及流形上的随机分析。

3.1 黎曼子流形

Submanifold. 若存在一族开集 $\{\mathcal{V}_\alpha\}_{\alpha\in\Lambda}$ （属于 $\mathbb{R}^n$）、一族开集 $\{\mathcal{U}_\alpha\}_{\alpha\in\Lambda}$ （属于 $\mathbb{R}^m$），以及一族光滑（C^∞）映射 $\{\phi_\alpha\}_{\alpha\in\Lambda}$，使得子集 $\mathcal{M}\subset\mathbb{R}^n$ 满足下述条件，则称其为$\mathbb{R}^n$ （无边界）的 m维（嵌入）子流形：

$$\mathcal{M}\subset\bigcup_{\alpha\in\Lambda}\mathcal{V}_\alpha,\ \text{and}\ \phi_\alpha\colon\mathcal{U}_\alpha\to\mathcal{V}_\alpha\cap\mathcal{M}$$

是一个微分同胚，即 $\phi_\alpha^{-1}\colon\ \mathcal{V}_\alpha\cap\mathcal{M}\to\mathcal{U}_\alpha$ 也光滑。对于任意 $x\in\mathcal{M}$，x 处的切空间定义为

$$T_x\mathcal{M}:=\left\{\gamma'(0)\colon\exists\,\varepsilon>0,\ \gamma\colon(-\varepsilon,\varepsilon)\to\mathcal{M}\ \text{smooth},\ \gamma(0)=x\right\}\subset\mathbb{R}^n,$$

以及切丛 $T\mathcal{M}:=\bigcup_{x\in\mathcal{M}}T_x\mathcal{M}$。此外，对于任意 $x\in\mathcal{M}$，我们称 $N_x\mathcal{M}:=(T_x\mathcal{M})^\perp$，即 $\mathbb{R}^n$ 中的正交补，为法空间，并定义法丛 $N\mathcal{M}:=\bigcup_{x\in\mathcal{M}}N_x\mathcal{M}$。若光滑映射 $V\colon\mathcal{M}\to T\mathcal{M}$ 满足 $V(x)\in T_x\mathcal{M}$，则称其为 $\mathcal{M}$ 上的（光滑）向量场，记作 $V\in\Gamma(T\mathcal{M})$；类似地，$\eta\in\Gamma(N\mathcal{M})$表示一个法向量场。关于光滑流形的更多细节，请参阅 Lee (2012)。

Riemannian submanifold. 设 $\mathcal{M}\subset\mathbb{R}^n$ 为一个子流形。对于任意 $x\in\mathcal{M}$，由于 $T_x\mathcal{M}\subset\mathbb{R}^n$，我们在 $T_x\mathcal{M}$ 上赋予标准欧几里得内积。于是，$\mathcal{M}$ 成为嵌入于 $\mathbb{R}^n$中的黎曼子流形。我们进一步假设 $\mathcal{M}\subset\mathbb{R}^n$ 是连通且闭的。因此，在本文中，$\mathcal{M}\subset\mathbb{R}^n$ <style id='12'}若为连通且闭的嵌入黎曼子流形，则称为黎曼子流形。

设 $d_{\mathcal{M}}$ 为定义如下的诱导黎曼距离：

$$d_{\mathcal{M}}(x,y):=\inf\left\{L(\gamma)\colon\gamma\colon[0,1]\to\mathcal{M}\ \text{piecewise smooth},\ \gamma(0)=x,\ \gamma(1)=y\right\},$$

其中，γ的长度为 $L(\gamma):=\int_0^1\|\gamma'(t)\|\,\mathrm{d}t$。等价地，$d_{\mathcal{M}}(x,\ y)$是 $\mathcal{M}$ 中连接 x 与 y的所有曲线中的最小长度。注意，$(\mathcal{M},\ d_{\mathcal{M}})$是一个完备度量空间。

Derivative. 令 D 表示 $\mathbb{R}^n$上的方向导数，即 $Dg(y)[v]=D_vg(y):=\langle v,\nabla^{\mathbb{R}}g(y)\rangle$，其中 $\nabla^{\mathbb{R}}g$ 表示 g的梯度。我们也可以考虑沿着$\mathcal{M}$的梯度。对于任意 $f\in C^\infty(\mathcal{M})$，即定义在 $\mathcal{M}$上的光滑函数，其沿着 $\mathcal{M}$ 的梯度为 $\nabla f(x)=(\nabla^{\mathbb{R}}\tilde{f}(x))^\top\in T_x\mathcal{M}$，也就是将 $\nabla^{\mathbb{R}}\tilde{f}(x)$正交投影到 $T_x\mathcal{M}$上，˜其中 f 是 f 在 $\mathbb{R}^n$上的任意光滑延拓。我们进一步考虑向量场的导数，由 Levi-Civita联络 ∇给出：$\Gamma(T\mathcal{M})\times\Gamma(T\mathcal{M})\to\Gamma(T\mathcal{M})$。类似地，其定义为 $\nabla_XY(x):=(D_{\tilde{X}}\tilde{Y}(x))^\top$，其中 $\tilde{X}$、$\tilde{Y}$为任意延拓。此外，

第二基本形式 II 定义为 $\mathrm{II}_x(X,Y) := (D_{\tilde{X}}\tilde{Y}(x))^{\perp} \in N_x\mathcal{M}$，即投影到 $N_x\mathcal{M}$上的正交投影。

光滑曲线 $\gamma\colon [0,1] \to \mathcal{M}$ 称为测地线，如果 $\nabla_{\gamma'(t)}\gamma'(t) \equiv 0$。注意，对于任意$x \in \mathcal{M}$ 和任意 $v \in T_x\mathcal{M}$，存在唯一的测地线 $\gamma_v\colon [0,1] \to \mathcal{M}$，满足 $\gamma_v(0) = x$ 且$\gamma_v'(0) = v$。因此，在 x 处的指数映射 为 $\exp_x\colon T_x\mathcal{M} \to \mathcal{M}$，定义为 $\exp_x(v) := \gamma_v(1)$。更多关于黎曼流形的细节，请参阅 Lee (2018)。

TUBULAR NEIGHBORHOOD. 根据管状邻域定理（Lee, 2018, 定理 5.25），存在一个光滑函数 $\varepsilon\colon \mathcal{M} \to (0,\infty)$，使得管状映射 $T(x,\xi) = x + \xi\colon N\mathcal{M} \to \mathbb{R}^n$ 限制为从

$$\mathcal{U}_0 := \{(x,\xi) \in N\mathcal{M} \colon \|\xi\| < \varepsilon(x)\}$$

到 $\mathcal{V}_0 := T(\mathcal{U}_0)$，称为 $\mathbb{R}^n$ 中 $\mathcal{M}$ 的管状邻域。特别地，每个 $y \in \mathcal{V}_0$ 都有唯一分解 $y = x + \xi$，其中 $x \in \mathcal{M}$ 且 $\xi \in N_x\mathcal{M}$；我们定义典范投影$\colon$，使得 $\pi(y) = x$。此外，若 $\varepsilon_0 := \inf_{x\in\mathcal{M}} \varepsilon(x) > 0$，则（在缩小 $\mathcal{U}_0$ 后）可以取 $\varepsilon(x) \equiv \varepsilon_0$，并且

$$\mathcal{V}_0 = \{y \in \mathbb{R}^n \colon d(y,\mathcal{M}) < \varepsilon_0\}, \quad d(y,\mathcal{M}) := \inf\{\|y - x\| \colon x \in \mathcal{M}\},$$

称为 $\mathcal{M}$ 的一致管状邻域，其一致半径为 ε_0。一致管状邻域的存在性与 $\mathcal{M}$ 的全局度量性质密切相关；更多细节见附录 B.2。

3.2 黎曼流形上的随机微分方程

ITÔ SDE VS. STRATONOVICH SDE. 设 $\boldsymbol{W} = (\boldsymbol{W}^1, \ldots, \boldsymbol{W}^n)$ 是 $\mathbb{R}^n$ 上的标准布朗运动。由 $\boldsymbol{W}$ 驱动的 $\mathbb{R}^n$ 上的 Stratonovich 随机微分方程表述为

$$\mathrm{d}\boldsymbol{X}_t = b(\boldsymbol{X}_t)\,\mathrm{d}t + \sum_{\alpha=1}^{n} \sigma_\alpha(\boldsymbol{X}_t) \circ \mathrm{d}\boldsymbol{W}_t^\alpha,$$

其中 $b, \sigma_\alpha\colon \mathbb{R}^n \to \mathbb{R}^n$，$\circ$表示斯特拉托诺维奇积分（Karatzas 和 Shreve，1991）。它等价于 Itô 形式

$$\mathrm{d}\boldsymbol{X}_t = \left(b(\boldsymbol{X}_t) + \frac{1}{2}\sum_{\alpha=1}^{n} D_{\sigma_\alpha}\sigma_\alpha(\boldsymbol{X}_t)\right)\mathrm{d}t + \sum_{\alpha=1}^{n} \sigma_\alpha(\boldsymbol{X}_t)\,\mathrm{d}\boldsymbol{W}_t^\alpha.$$

有关随机微分方程的更多详情，请参阅 Le Gall（2016）和（Karatzas 和 Shreve，1991）。

SDE ON MANIFOLD. 设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个黎曼子流形，并设 $V_\alpha \in \Gamma(T\mathcal{M})$ 对于 $\alpha = 0, 1, \ldots\ldots, n$。设 $\boldsymbol{W}$是 $\mathbb{R}^n$上的标准布朗运动。一个取值于 $\mathcal{M}$的随机微分方程可符号化地写为

$$\mathrm{d}\boldsymbol{X}_t = V_0(\boldsymbol{X}_t)\,\mathrm{d}t + \sum_{\alpha=1}^{n} V_\alpha(\boldsymbol{X}_t) \circ \mathrm{d}\boldsymbol{W}_t^\alpha,$$

这意味着，如果对于任意 $f \in C^\infty(\mathcal{M})$， $\boldsymbol{X}_t$ 是一个解，则

$$\mathrm{d}f(\boldsymbol{X}_t) = \langle \nabla f(\boldsymbol{X}_t), V_0(\boldsymbol{X}_t) \rangle \ \mathrm{d}t + \sum_{\alpha=1}^{n} \langle \nabla f(\boldsymbol{X}_t), V_\alpha(\boldsymbol{X}_t) \rangle \circ \mathrm{d}\boldsymbol{W}_t^\alpha.$$

因此， $\mathcal{M}$值随机微分方程可以视为 $\mathbb{R}^n$上的 Stratonovich 随机微分方程。注意，给定初始条件后，解存在且唯一。Hsu（2002）对流形值随机微分方程提供了进一步的详细介绍。

Brownian Motion on $\mathcal{M}$. 设 $\mathcal{M} \subset \mathbb{R}^n$ 为配备光滑正交投影 $P(x)$ 的黎曼子流形，即对于每个 $x \in \mathcal{M}$， $P(x)\colon\ \mathbb{R}^n \to \mathbb{R}^n$是到 $T_x\mathcal{M}$上的正交投影，且映射 P 在 $\mathcal{M}$上光滑。设 $\{e_\alpha\}_{\alpha=1}^n$ 为 $\mathbb{R}^n$的标准基，并定义 $P_\alpha \coloneqq Pe_\alpha \in \Gamma(T\mathcal{M})$。接下来考虑如下 $\mathcal{M}$值随机微分方程:

$$\mathrm{d}\boldsymbol{B}_t = \sum_{\alpha=1}^{n} P_\alpha(\boldsymbol{B}_t) \circ \mathrm{d}\boldsymbol{W}_t^\alpha = P(\boldsymbol{B}_t) \circ \mathrm{d}\boldsymbol{W}_t, \tag{1}$$

其中， $\boldsymbol{W}$ 是 $\mathbb{R}^n$上的标准布朗运动。设 $\boldsymbol{B}^{\mathcal{M}}$ 为满足 $\boldsymbol{B}_0^{\mathcal{M}} = 0$的随机微分方程（1）的唯一解。则称 $\boldsymbol{B}^{\mathcal{M}}$ 为 $\mathcal{M}$ 上的标准布朗运动（Hsu, 2002）。

4 p-GEM 的强收敛

设 $\mathcal{M} \subset \mathbb{R}^n$ 为黎曼子流形。考虑一个取值于 $\mathcal{M}$ 的随机微分方程

$$\mathrm{d}\boldsymbol{X}_t = V(\boldsymbol{X}_t)\ \mathrm{d}t + \mathrm{d}\boldsymbol{B}_t^{\mathcal{M}}, \quad t \in [0, T], \tag{2}$$

其中 $V \in \Gamma(T\mathcal{M})$，且 $\boldsymbol{B}^{\mathcal{M}}$ 是定义在 $\mathcal{M}$ 上的标准布朗运动。令 $h = T/N$（$N \in \mathbb{N}$），并对 $k = 0, 1, \ldots\ldots, N$ 定义 $t_k = kh$。GEM 算法通过设定 $\boldsymbol{X}_0^h = \boldsymbol{X}_0$ 并迭代，构造离散时间过程 $\{\boldsymbol{X}_k^h\}_{k=0}^N$

$$\boldsymbol{X}_{k+1}^h = \exp_{\boldsymbol{X}_k^h}\left(hV(\boldsymbol{X}_k^h) + \sqrt{h}\boldsymbol{\xi}_{\boldsymbol{X}_k^h}\right), \quad k = 0, \ldots, N-1, \tag{3}$$

其中 $\exp_x\colon\ T_x\mathcal{M} \to \mathcal{M}$ 是指数映射， $\boldsymbol{\xi}_x$ 是相对于度量 $g(x)$ 定义在 $T_x\mathcal{M}$ 上的标准高斯分布。本节的主要目标是建立 $\boldsymbol{X}_k^h$ 到 $\boldsymbol{X}_{t_k}$的 p-强收敛。

4.1 主要结果

我们首先引入以下两个假设，以确保 $\mathcal{M}$ 所需的几何有界性。

假设 I 设 $\iota\colon\ \mathcal{M} \hookrightarrow \mathbb{R}^n$ 为典范嵌入映射。假设 $d\iota$ 的协变导数满足

$$\sup_{x \in \mathcal{M}} \|\nabla d\iota(x)\|_{\mathrm{op}} < \infty, \quad \sup_{x \in \mathcal{M}} \left\|\nabla^2 d\iota(x)\right\|_{\mathrm{op}} < \infty.$$

假设 II 假设 $\mathcal{M}$ 在 $\mathbb{R}^n$中存在一致管状邻域 $\mathcal{V}$。

注 3 关于这两个假设，我们作如下说明 .

(i) $\nabla d\iota$ 本质上是 $\mathcal{M}$的外蕴曲率。关于 $\nabla d\iota$ 和$\nabla^2 d\iota$的严格定义，以及在假设 *I* 下 $\mathcal{M}$ 的几何有界性，详见附录 *C*。由于假设 *I* 是以内在导数的形式表述的，可能难以直接验证，因此我们还在附录 *E.3* 中给出了一个蕴含假设 *I* 的充分外在判据。

(ii) 在附录 *E.1* 中，我们给出了满足假设 *I* 和假设 *II* 的三类经典流形：紧流形、图和水平集。水平集示例与 *Wang* 和 *Solo (2025a)* 研究的情形一致。

(iii) 假设 *I* 对 $\mathcal{M}$施加了局部几何正则性要求，而假设 *II* 是一项全局要求。特别地，假设 *II* 蕴含$\sup\|\nabla d\iota\|_{\mathrm{op}} < \infty$，但一般而言不蕴含$\sup\|\nabla^2 d\iota\|_{\mathrm{op}} < \infty$。反之，假设 *I* 并不能保证一致管状邻域的存在。这两个假设之间的关系在附录 *E.2* 中进行了讨论。然而，如果典范嵌入 ι 是全局双 *Lipschitz* 连续的（附录 *F*），那么$\sup\|\nabla d\iota\|_{\mathrm{op}} < \infty$ 蕴含假设 *II*；参见附录 *F.2*。

我们还对漂移向量场 V 作出标准正则性条件的假设。

假设 III 令 $V \in \Gamma(T\mathcal{M})$。假设 $V \in C_b^1$，即，

$$\sup_{x\in\mathcal{M}} \|V(x)\| < \infty, \quad \sup_{x\in\mathcal{M}} \|\nabla V(x)\|_{\mathrm{op}} < \infty.$$

注 4 关于这一假设，我们给出两点说明。

(i) 该假设仅涉及 $V \in \Gamma(T\mathcal{M})$ 的内蕴性质，因此与 $\mathcal{M}$ 如何嵌入 $\mathbb{R}^n$ 无关。然而， V 本身的一致有界性在实践中可能构成限制。处理 $\mathcal{M} \subset \mathbb{R}^n$ 时，可以将这一假设放宽为

$$\sup_{x_1\neq x_2\in\mathcal{M}} \frac{\|V(x_1) - V(x_2)\|}{\|x_1 - x_2\|} < \infty, \quad \sup_{x\in\mathcal{M}} \frac{\|V(x)\|}{1 + \|x\|^{1/3}} < \infty;$$

详情请见附录 *L*。

(ii) 假设 *III* 与假设 *I* 一起保证了 *SDE (2)* 的解存在且唯一（*Hsu, 2002*），并且具有非爆炸性（*Li, 1994*；*Stroock, 2000*）。利用我们的外在延拓也可以得出相同结论；见推论 *12*。

由此，我们可以正式陈述主要结果。

定理 5 设 $1 \le p < \infty$。设 $\mathcal{M} \subset \mathbb{R}^n$ 为黎曼子流形，且 $V \in \Gamma(T\mathcal{M})$。假设 $\mathcal{M}$ 满足假设 *I* 和假设 *II*，且 V满足假设 *III*。设随机微分方程 *(2)* 定义在 $\mathcal{M}$ 上，其解为 $\boldsymbol{X}_t$，其中 $\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right] < \infty$。对于由 *GEM (3)* 构造的 $\{\boldsymbol{X}_k^h\}_{k=0}^N$，其中 $\boldsymbol{X}_0^h = \boldsymbol{X}_0$，存在 $C_p(T)$，使得

$$\mathbb{E}\left[\max_{0\le k\le N} \left\|\boldsymbol{X}_k^h - \boldsymbol{X}_{t_k}\right\|^p\right] \le C_p(T) h^{p/2},$$

其中 $C_p(T) = \mathcal{O}(\exp(T^p))$。

注 6 定理 *5* 可直接推广到非齐次时间情形，即考虑满足假设 *III* 对 $t \in [0,T]$ 一致成立的时变向量场 $V(t,x)$；详情见附录 *K*。

定理 5 建立了以 $\mathbb{R}^n$中的外在距离衡量的强收敛，而非以 $\mathcal{M}$上的内蕴黎曼距离衡量的强收敛。为了以内蕴距离 $d_{\mathcal{M}}$衡量误差，我们还假设典范嵌入 $\iota\colon \mathcal{M} \hookrightarrow \mathbb{R}^n$ 是全局双 Lipschitz 连续的（附录 F）。在此条件下，假设 II 不再必要。因此，作为定理 5 的一个应用，我们得到定理 7；其证明见附录 J.2。

定理 7 设 $1 \le p < \infty$。设 $\mathcal{M} \subset \mathbb{R}^n$ 为一个黎曼子流形，且其典范嵌入 $\iota\colon \mathcal{M} \hookrightarrow \mathbb{R}^n$ 是全局双 *Lipschitz* 连续的。假设 $\mathcal{M}$满足假设 *I*，且 $V \in \Gamma(T\mathcal{M})$ 满足假设 *III*。设 $\boldsymbol{X}_t$ 是 $\mathcal{M}$ 上随机微分方程 *(2)* 的一个解，其中 $\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right] < \infty$。对于由 *GEM (3)* 构造的 $\{\boldsymbol{X}_k^h\}_{k=0}^N$，其中$\boldsymbol{X}_0^h = \boldsymbol{X}_0$，存在 $C_p(T) = \mathcal{O}(\exp(T^p))$ 使得

$$\mathbb{E}\left[\max_{0\le k\le N} d_{\mathcal{M}}\left(\boldsymbol{X}_k^h, \boldsymbol{X}_{t_k}\right)^p\right] \le C_p(T) h^{p/2},$$

其中 $d_{\mathcal{M}}$ 是 $\mathcal{M}$上的黎曼距离。

对于任意黎曼流形 $\mathcal{M}$，由纳什嵌入定理，它可等距嵌入到某个欧几里得空间 $\mathbb{R}^n$ 中。若 $\mathcal{M}$进一步是紧致的，则定理 7 的所有假设均成立，且与嵌入到 $\mathbb{R}^n$的选择无关。因此，对于任意紧黎曼流形上的 GEM，我们得到如下推论 8；证明见附录 J.3。

推论 8 设 $1 \le p < \infty$。设 $\mathcal{M}$ 为紧黎曼流形。对于随机微分方程 *(2)* 的解$\boldsymbol{X}_t$ 以及在 $\boldsymbol{X}_0^h = \boldsymbol{X}_0$下由 *GEM (3)* 构造的 $\{\boldsymbol{X}_k^h\}_{k=0}^N$，存在 $C_p(T) = \mathcal{O}(\exp(T^p))$ 使得

$$\mathbb{E}\left[\max_{0\le k\le N} d_{\mathcal{M}}\left(\boldsymbol{X}_k^h, \boldsymbol{X}_{t_k}\right)^p\right] \le C_p(T) h^{p/2}.$$

4.2 定理 5 的证明思路

由于 $\mathcal{M} \subset \mathbb{R}^n$，存在到 $T_x\mathcal{M}$上的光滑正交投影 $P(x)$；见附录 D.1。于是 $\boldsymbol{B}^{\mathcal{M}}$ 可实现为标准 $\mathbb{R}^n$-布朗运动$\boldsymbol{W}$ 的投影（Hsu, 2002； Émery, 2012；Thalmaier, 2023年），且 (2) 等价于取值于 $\mathcal{M}$的随机微分方程

$$\mathrm{d}\boldsymbol{X}_t = V(\boldsymbol{X}_t)\,\mathrm{d}t + P(\boldsymbol{X}_t) \circ \mathrm{d}\boldsymbol{W}_t, \quad t \in [0,T],$$

这也可以视为定义在 $\mathbb{R}^n$ 上的 Stratonovich 随机微分方程（Hsu, 2002）。因此，根据 Itô 公式（Karatzas 和 Shreve，1991），它可以写成 Itô 形式：

$$\mathrm{d}\boldsymbol{X}_t = U(\boldsymbol{X}_t)\,\mathrm{d}t + P(\boldsymbol{X}_t)\,\mathrm{d}\boldsymbol{W}_t, \quad t \in [0,T], \tag{4}$$

其中

$$U(x) = V(x) + A(x), \quad A^j(x) := \frac{1}{2}\sum_{\alpha=1}^{n} P_\alpha\left(P_{j\alpha}\right)(x), \quad j = 1, \ldots, n. \tag{5}$$

这里，相对于 $\mathbb{R}^n$的标准基 $\{e_\alpha\}_{\alpha=1}^n$， $P(x) = (P_{j\alpha}(x))$是 $P(x)$的矩阵表示，而 $P_\alpha(x) := P(x)e_\alpha \in T_x\mathcal{M} \subset \mathbb{R}^n$ 被视为一个向量场。对于具有开 $\mathcal{V} \supset \mathcal{M}$的$f \in C^1(\mathcal{V})$,

$$P_\alpha(f)(x) := \sum_{j=1}^{n} P_{j\alpha}(x)\frac{\partial f}{\partial x^j}(x), \quad x \in \mathcal{M}.$$

由于 P定义在 $\mathcal{M}$ 的一个开邻域上（引理 D.1），因此对所有 $x \in \mathcal{M}$， $P_\alpha\left(P_{j\alpha}\right)(x)$都是良定义的。

MAIN IDEAS. 注意，随机微分方程 (4) 是在 $\mathbb{R}^n$中表述的，因此可以将其视为标准的欧几里得随机微分方程。我们的主要想法是先应用欧几里得 EM 算法来逼近 (4)，然后将所得的欧几里得离散过程与内禀过程$\boldsymbol{X}_k^h$进行比较。然而，我们不能直接对 (4) 应用欧几里得 EM 算法，因为系数 $U(x)$和 $P(x)$仅在 $\mathcal{M}$ 上或在 $\mathcal{M}$的一个小邻域上是良定义的。更具体地说，我们的方法包括两个步骤:

(i) 欧几里得空间上的延拓：由于 $\mathcal{M} \subset \mathbb{R}^n$ 是闭的，将系数 $U(x)$，即 $V(x)$和 $A(x)$，以及 $P(x)$延拓到 $\mathbb{R}^n$ 的基本策略，是将管状邻域定理与乌雷松引理（Munkres，2013年）相结合。关键在于确保延拓后的系数是全局利普希茨连续的，这需要额外的工作。正是在这里， $\mathcal{M}$的几何有界性（假设 I 和假设 II）以及 V 的正则性（假设 III）发挥作用。我们将在第 4.3 节讨论（4）的适定欧几里得延拓的构造。

(ii) 差异的比较：对于延拓后的随机微分方程，我们应用欧氏EM算法并得到一个离散过程 $\boldsymbol{Y}_k^h$，它以阶 $1/2$ 强逼近原始随机微分方程（2）的解（Kloeden 和 Platen，1992年; Milstein 和 Tretyakov，2013年）。为证明定理 5，只需控制 $\boldsymbol{Y}_k^h$与 $\boldsymbol{X}_k^h$之间的差异。利用几何有界性和利普希茨性质，我们通过对指数映射的泰勒展开建立 $\boldsymbol{Y}_k^h$到 $\boldsymbol{X}_k^h$ 的强收敛。我们将在第 4.4 节给出细节。

4.3 欧几里得空间上的适定延拓

EXTENSION OF COEFFICIENTS. 设 F: $\mathcal{M} \to \mathbb{R}^n$ （或 $\mathbb{R}^{n\times n}$）是一个光滑映射。由于 $\mathcal{M}$ 是闭集，根据乌雷松引理（Munkres，2013年），我们可以选取两个管状邻域$\mathcal{V}_0$， $\mathcal{V}$，满足 $\mathcal{M} \subset \mathcal{V}_0 \subset \overline{\mathcal{V}_0} \subset \mathcal{V}$，以及一个鼓包函数 χ: $\mathbb{R}^n \to [0,1]$，使得 $\chi|_{\mathcal{M}} \equiv 1$和$\chi|_{\mathbb{R}^n\setminus\mathcal{V}_0} \equiv 0$。定义

$$\widetilde{F}(y) := \chi(y)F(\pi(y)), \quad \forall\, y \in \mathbb{R}^n,$$

其中， π 是 $\mathcal{V}$的典范投影。于是， F 将 F 扩展到 $\mathbb{R}^n$。这是我们扩展随机微分方程 (4) 中系数 $V(x)$、 $A(x)$ 和 $P(x)$ 的主要策略。主要问题

是确保这些扩展全局利普希茨连续，因此我们假设 $\mathcal{M}$ 具有几何有界性，并假设 V具有适当的正则性；详情请参见附录 H。

对于 V和 P，我们有如下引理；证明见附录 H.2。

引理 9 假设 $\mathcal{M}$ 满足假设 *I* 和假设 *II*，且 $V \in \Gamma(T\mathcal{M})$满足假设 *III*。令 $P(x)$ 为到 $T_x\mathcal{M}$上的正交投影，并在 $\widetilde{\ }$ $\widetilde{\ }$ $\widetilde{\ }$上对 $\mathcal{M}$光滑定义。则存在 V和 P在 $\mathbb{R}^n$上的延拓 V和 P，使得 $V\widetilde{\ }$和 P均为全局利普希茨连续。

对于 A，我们需要更明确地理解其几何结构，如命题 10 所述；证明见附录 G.1。这一刻画对于构造适定延拓以及控制欧氏EM离散化 $\boldsymbol{Y}_k^h$与内在 EM 离散化 $\boldsymbol{X}_k^h$之间的差异都很有用。

命题 10 设 $\mathcal{M} \subset \mathbb{R}^n$ 为一个 m维黎曼子流形。对于任意$x \in \mathcal{M}$，令 $A(x) \in \mathbb{R}^n$ 按式 *(5)* 定义。令 $\{E_i(x)\}_{i=1}^m$ 为 $T_x\mathcal{M}$的标准正交基。则

$$A(x) = \frac{1}{2}\sum_{i=1}^{m} \mathrm{II}_x\left(E_i(x), E_i(x)\right), \tag{6}$$

其中 II 是 $\mathcal{M}$的第二基本形式。

基于公式 (6) 和 $\mathcal{M}$的几何有界性，我们可以构造 A 的一个扩展，使其在 $\mathbb{R}^n$上全局利普希茨连续，如引理 11 所示；证明见附录 H.3。

引理 11 假设 $\mathcal{M}$ 满足假设 *I* 和假设 *II*。对于任意 $x \in \mathcal{M}$、$\widetilde{\ }$，令 $A(x) \in \mathbb{R}^n$ 按式 *(5)* 定义。则存在 A 的一个扩展 A，定义在 $\mathbb{R}^n$上，且全局利普希茨连续。

$\sim$ $\sim$ $\sim$Extended SDE on $\mathbb{R}^n$. 因此，我们可以选择扩展 V、 A和 P，使其在 $\mathbb{R}^n$ 上定义良好且全局利普希茨连续。考虑随机微分方程

$$\mathrm{d}\boldsymbol{X}_t = \widetilde{U}(\boldsymbol{X}_t)\,\mathrm{d}t + \widetilde{P}(\boldsymbol{X}_t)\,\mathrm{d}\boldsymbol{W}_t, \quad t \in [0,T], \tag{7}$$

$\widetilde{\ }$ $\widetilde{\ }$ $\widetilde{\ }$其中 $U = V + A$。那么，(7) 是一个定义良好的欧几里得随机微分方程，具有全局 Lipschitz 连续系数。此外，根据 Hsu （ 2002 ）中的命题 1.2.8，若 $\boldsymbol{X}_0 \in \mathcal{M}$，则随机微分方程 (4) 和 (7) 有相同的解。因此，我们仍将 V、 A和 $\sim$ $\sim$ $\sim$$P$分别记作 V、 A和 P，并且不区分随机微分方程 (7) 与随机微分方程 (4)。

作为这一构造的直接应用，标准的欧几里得随机微分方程理论（Baudoin，2014年，定理 6.1）给出了随机微分方程 (4)（等价地，随机微分方程 (2)）的非爆炸解的存在性与唯一性。

推论 12 假设 $\mathcal{M}$ 满足假设 *I* 和假设 *II*，且 $V \in\Gamma(T\mathcal{M})$ 满足假设 *III*。令 $1 \le p < \infty$。对于随机微分方程 *(2)*，若 $\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right] < \infty$，则它存在唯一解 $\boldsymbol{X}_t$，使得

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|\boldsymbol{X}_t\|^p\right] < \infty.$$

第二个推论对于证明定理 5 至关重要。由于随机微分方程 (4) 是定义良好的欧几里得随机微分方程，且具有全局 Lipschitz 连续系数，因此可以使用欧几里得 EM 算法对其进行近似。令 $\boldsymbol{Y}_0^h = \boldsymbol{X}_0$ 并且,

$$\boldsymbol{Y}_{k+1}^h = \boldsymbol{Y}_k^h + hU(\boldsymbol{Y}_k^h) + \sqrt{h}P(\boldsymbol{Y}_k^h)\boldsymbol{\xi}, \quad k = 0, \ldots, N-1, \tag{8}$$

其中 $\boldsymbol{\xi} \sim \mathcal{N}(0, I_n)$ 是定义在 $\mathbb{R}^n$ 上的标准高斯分布。令 $1 \leq p < \infty$。若 $\mathbb{E}[\|\boldsymbol{X}_0\|^p] < \infty$, 则由标准结果（Kloeden 和 Platen，1992年；Milstein 和 Tretyakov，2013年）可得

$$\mathbb{E}\left[\max_{0\leq k\leq N}\left\|\boldsymbol{X}_{t_k} - \boldsymbol{Y}_k^h\right\|^p\right] \leq E_p(T)h^{p/2}, \tag{9}$$

其中 $E_p(T) = \mathcal{O}(\exp(T^p))$；详情见附录 I。因此，要证明定理 5，还需控制 $\boldsymbol{Y}_k^h$与 $\boldsymbol{X}_k^h$ 之间的差异。

4.4 外在离散与内在离散的比较

为比较式 (3) 中的内在离散 $\boldsymbol{X}_k^h$ 与式 (8) 中的外在离散 $\boldsymbol{Y}_k^h$，我们首先采用 Wang 和 Solo (2025a) 关于指数映射 $\exp_x$的泰勒展开的思路： $T_x\mathcal{M} \to \mathcal{M}$。根据引理 C.4，若 $\mathcal{M}$ 满足假设 I，则

$$\exp_x(v) = x + v + \frac{1}{2}\mathrm{II}_x(v, v) + R_3(x, v), \quad \forall\, v \in T_x\mathcal{M},$$

其中余项受到一致控制，即 $\sup_x\|R_3(x, v)\| \leq C_R\|v\|^3$ ，其中某个 $C_R > 0$。这样便可将比较完全转化为在欧几里得空间中进行。由 $A(x)$的几何结构所导出的第二个关键公式（命题 10）为

$$A(x) = \frac{1}{2}\mathbb{E}\left[\mathrm{II}_x(P(x)\boldsymbol{\xi}, P(x)\boldsymbol{\xi})\right],$$

当 $\boldsymbol{\xi} \sim \mathcal{N}(0, I_n)$ 时；参见附录 G.2。该公式控制单步差异,

$$\left\|\mathbb{E}\left[\boldsymbol{X}_{k+1}^h \mid \boldsymbol{X}_k^h = x\right] - \mathbb{E}\left[\boldsymbol{Y}_{k+1}^h \mid \boldsymbol{Y}_k^h = x\right]\right\| \leq \mathcal{O}(h^{3/2}).$$

结合这些结果，我们得到引理 13；证明见附录 J.1 .

引理 13 假设 $\mathcal{M}$ 满足假设 *I* 和假设 *II*，且 $V \in \Gamma(T\mathcal{M})$ 满足假设 *III*。设 $1 \leq p < \infty$。对于按 *(3)* 构造的 $\boldsymbol{X}_h^k$ 和按 *(8)* 构造的 $\boldsymbol{Y}_k^h$，我们有

$$\mathbb{E}\left[\max_{0\leq k\leq N}\left\|\boldsymbol{X}_k^h - \boldsymbol{Y}_k^h\right\|^p\right] \leq G_p(T)h^{p/2},$$

其中 $G_p(T) = \mathcal{O}(\exp(T^p))$。

4.5 定理 5 的证明

结合上述结果，我们得到了定理 5 的一个直接证明。

证明 如第 4.3 节所示，随机微分方程 (2) 扩展为随机微分方程 (4)，即随机微分方程 (7)，它在 $\mathbb{R}^n$ 上定义良好，且具有全局 Lipschitz 连续系数。因此，对于按 (8) 构造的欧几里得 EM 离散过程 $\left\{\boldsymbol{Y}_k^h\right\}_{k=0}^N$，定理 I.1 表明

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_{t_k}-\boldsymbol{Y}_k^h\right\|^p\right]\le E_p(T)h^{p/2},$$

其中 $E_p(T)=\mathcal{O}(\exp(T^p))$。

此外，根据引理 13，存在 $G_p(T)=\mathcal{O}(\exp(T^p))$，使得

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h-\boldsymbol{Y}_k^h\right\|^p\right]\le G_p(T)h^{p/2}.$$

因此，对于 $C_p(T):=E_p(T)+G_p(T)=\mathcal{O}(\exp(T^p))$，我们有

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h-\boldsymbol{X}_{t_k}\right\|^p\right]\le\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h-\boldsymbol{Y}_k^h\right\|^p\right]+\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_{t_k}-\boldsymbol{Y}_k^h\right\|^p\right]\le C_p(T)h^{p/2}.$$

■

5 p-RLD 的 Wasserstein 收敛

注意，对于 概率分布 μ，ν 在 $\mathcal{M}$ 上，p-Wasserstein 距离定义为

$$\mathcal{W}_p(\mu,\nu):=\inf\left\{\mathbb{E}\left[d_{\mathcal{M}}(\boldsymbol{X},\boldsymbol{Y})^p\right]^{1/p}:\boldsymbol{X}\sim\mu,\boldsymbol{Y}\sim\nu\right\};$$

详情请参见 Villani（2008年）。利用 GEM 的 p-强收敛性，我们可以量化通过 GEM 离散化 RLD 时的 $\mathcal{W}_p$ 收敛速率。

设 $\mathcal{M}\subset\mathbb{R}^n$ 为一个黎曼子流形，配备典范黎曼体积测度 vol（Jost，2008年）。设 $\phi\in C^\infty(\mathcal{M})$，其中 $C_\phi:=\int_{\mathcal{M}}\phi$ dvol $<\infty$，并定义 $\mathcal{M}$ 上的概率分布 μ_ϕ，使得 $\mathrm{d}\mu_\phi:=C_\phi^{-1}e^{-\phi}$ dvol。为从 μ_ϕ 中采样，考虑如下定义的 RLD：

$$\mathrm{d}\boldsymbol{X}_t=-\frac{1}{2}\nabla\phi(\boldsymbol{X}_t)\,\mathrm{d}t+\mathrm{d}\boldsymbol{B}_t^{\mathcal{M}},\quad \boldsymbol{X}_0=x,\qquad t\in[0,T],\tag{10}$$

其中 $x\in\mathcal{M}$ 和 $\nabla\phi$ 表示 ϕ的梯度向量场。此外，为近似（10），考虑如下 GEM：

$$\boldsymbol{X}_{k+1}^h=\exp_{\boldsymbol{X}_k^h}\left(-\frac{1}{2}h\nabla\phi(\boldsymbol{X}_k^h)+\sqrt{h}\boldsymbol{\xi}_{\boldsymbol{X}_k^h}\right),\quad k=0,\dots,N-1,\tag{11}$$

其中 $\exp_x$：$T_x\mathcal{M}\to\mathcal{M}$ 是指数映射，且 $\boldsymbol{\xi}_x$ 是 $T_x\mathcal{M}$上的标准高斯分布。

令 μ_t 表示 $\boldsymbol{X}_t$ 的分布，并令 $\hat{}\mu_k$ 表示 $\boldsymbol{X}_k^h$ 的分布。则 $\hat{}\mu_N$ 与 μ_ϕ 之间的总 Wasserstein 误差可由以下结果有界

$$\mathcal{W}_p\left(\mu_\phi,\hat{\mu}_N\right)\le\mathcal{W}_p\left(\mu_\phi,\mu_T\right)+\mathcal{W}_p\left(\mu_T,\hat{\mu}_N\right),$$

即(10)的混合误差 $\mathcal{W}_p(\mu_\phi,\mu_T)$与(11)的离散化误差 $\mathcal{W}_p(\mu_T,\hat{}\mu_N)$。

Mixing error. 众所周知，如果 ϕ 满足Bakry–Emery曲率条件（Kuwada，2010年；Bakry等人，2013年；Wang，2020年），即

$$\mathrm{Ric} + \nabla^2 \phi \geq \lambda_\kappa g, \tag{12}$$

对于某个 $\lambda_\kappa > 0$，其中 Ric 表示 Ricci 曲率张量；见附录 B。于是

$$\mathcal{W}_p\left(\mu_\phi, \mu_T\right) \leq e^{-\lambda_\kappa T} \mathcal{W}_p\left(\mu_\phi, \delta_x\right), \tag{13}$$

其中 δ_x 是 x处的点分布。条件 (12) 还意味着初始误差$\mathcal{W}_p$ (μ_ϕ, δ_x) 有限；见附录 M.1。

此外，引理 C.5 表明，若 $\mathcal{M} \subset \mathbb{R}^n$ 满足假设 I，则存在一个 $\kappa_{\mathrm{ric}} > 0$ 使得 Ric $\geq -\kappa_{\mathrm{ric}} g$。因此，要界定混合误差 (13)，只需要求存在一个 $\lambda_\kappa > 0$ 使得

$$\nabla^2 \phi \geq (\lambda_\kappa - \kappa_{\mathrm{ric}}) g. \tag{14}$$

Discretization error. 对于 GEM 引入的离散化误差，我们可以利用强收敛结果得到一个上界。在$\mathcal{M}$ 的几何有界性（假设 I 和双 Lipschitz 嵌入）以及 $\nabla \phi$ 的正则性（假设 II）条件下，定理 7 表明

$$\mathcal{W}_p\left(\mu_T, \hat{\mu}_N\right) \leq \mathbb{E}\left[d_{\mathcal{M}}\left(\boldsymbol{X}_T, \boldsymbol{X}_N^h\right)^p\right]^{1/p} \leq \mathcal{O}\left(\exp(T^p) h^{1/2}\right). \tag{15}$$

结合 (13) 和 (15)，得到以下定理。

定理 14 设 $1 \leq p < \infty$。设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个黎曼子流形，且 $\iota\colon \mathcal{M} \hookrightarrow \mathbb{R}^n$ 全局双 *Lipschitz* 连续。设 $\phi \in C^\infty(\mathcal{M})$ 满足 $C_\phi := \int_{\mathcal{M}} \phi \,\mathrm{dvol} < \infty$，且 $\mathrm{d}\mu_\phi := C_\phi^{-1} e^{-\phi} \,\mathrm{dvol}$。假设 $\mathcal{M}$ 满足假设 *I*，ϕ 满足 *(14)*，且 $\nabla \phi$ 满足假设 *III*。那么，对于由 *(11)* 构造的 $\boldsymbol{X}_k^h \sim \hat{} \mu_k$，我们有

$$\mathcal{W}_p\left(\mu_\phi, \hat{\mu}_N\right) \leq \mathcal{O}\left(e^{-T} + \exp(T^p) h^{1/2}\right).$$

注 15 对于定理 *14*，我们补充两点讨论。

(i) 由于 $\mathcal{M} \subset \mathbb{R}^n$、$\mu_\phi$、$\mu_t$ 和 $\hat{}\mu_t$ 可以视为 $\mathbb{R}^n$ 上的概率分布，因此我们可以使用由欧几里得距离诱导的外在 *Wasserstein* 距离，而不使用 $d_{\mathcal{M}}$。对于这种外在距离，ι 的全局双 *Lipschitz* 连续性可以由假设 *II* 替代，因为定理 *5* 在假设 *II* 下仍然适用，并给出相同的离散化界；参见附录 *M.2*。

（*ii*）当 $\mathcal{M}$ 是紧致黎曼流形时，正如推论 *8* 所述，*GEM* 具有´强收敛性质。因此，若 ϕ 还满足 *Bakry–Emery* 曲率条件，我们便可得到上述定理中的 p-*Wasserstein* 收敛界。*Wang* 等人 *(2020)* 也得到了这一结果，他们将 *KL* 散度收敛作为直接推论，证明了 2-*Wasserstein* 收敛。

6 结论

在这项工作中，我们证明了 GEM 格式在嵌入黎曼子流形 $\mathcal{M} \subset \mathbb{R}^n$ 上数值求解随机微分方程时具有 p-强收敛性，并将其应用于获得 GEM 离散化下 RLD 的 p-Wasserstein 界。在假设 $\mathcal{M}$ 具有有界外在曲率且在 $\mathbb{R}^n$ 中全局良好嵌入，并且漂移向量场 V 满足标准光滑性条件的情况下，我们证明 GEM 达到阶数为 $1/2$ 的 p-强收敛；此外，对于任意紧黎曼流形 $\mathcal{M}$，我们证明无论 $\mathcal{M}$ 如何 Nash 嵌入欧几里得空间，这些假设均成立，因此同一结论适用于任意紧致$\mathcal{M}$。对于 RLD， p-Wasserstein 误差可分解为混合误差项和离散化误差项：前者在 Bakry–Emery曲率条件下得到控制，后者则由我们的 p-强收敛结果控制。从技术上讲，我们的分析基于将 $\mathcal{M}$-随机微分方程外在延拓至 $\mathbb{R}^n$ ，并比较由此得到的欧几里得EM离散化与内在GEM格式。

Limitations and Future Works. (i) 我们对 GEM 离散化的界涉及依赖于 T 的指数级增长常数，即 $\mathcal{O}(\exp(T^p))$。改进这一时间依赖性，无论是在欧氏 EM 强收敛分析中，还是在控制外在与内在离散化过程之间的强差异方面，都是值得未来研究的方向。(ii) 我们的方法依赖于指数映射和切空间上的标准高斯噪声，但这两者在实践中都可能难以计算。将分析推广到回缩映射（作为指数映射的逼近），以及具有适当矩界的更一般噪声分布，是未来一个重要的研究方向。(iii) 我们关于 $\mathcal{M}$ 的假设依赖于外在性质，即 $\mathcal{M}$ 如何嵌入 $\mathbb{R}^n$。在 $\mathcal{M}$ 上基于完全内在几何条件建立类似结果，是未来研究的另一个自然课题。

致谢与资金披露

ZZ 得到了东京大学人工智能与超越研究所的支持。MS 得到了 JST ASPIRE资助项目（编号: JPMJAP25B1）的支持。

附录目录

A 记号 17

B 更多预备知识 17
B.1 黎曼子流形补充内容.18 B.2 可达半径.19 B.3 鞅论.20

C 假设 I 下的几何有界性 **21**
C.1 第二基本形式的有界性.21 C.2 指数映射的泰勒公式.24 C.3 黎曼曲率有界性.25

D 正交投影及其有界性 26
D.1 正交投影的存在性.26 D.2 正则投影的有界性.27 D.3 正交投影的有界性.29

E 关于几何假设的更多讨论 30
E.1 示例.30 E.2 假设 I 与假设 II 之间的关系.41 E.3 假设 I 的外在判据.45

F 更多讨论 关于嵌入映射的讨论 47
F.1 双 Lipschitz 连续性.47 F.2 由假设 I 推导假设 II.49

G 流形的几何 $A(x)$ 51G.1 命题 10 的证明.51 G.2 诱导公式.53

H 延拓的全局 Lipschitz 连续性 **54**
H.1 隆起函数的有界性.54 H.2 引理 9 的证明.55 H.3 引理 11 的证明.57

I 欧几里得 EM 方案 **59**

J 第 4 节中省略的证明 66
J.1 引理 13 的证明.66 J.2 定理 7 的证明.74 J.3 推论 8 的证明.75

K 非齐次时间情形 75

L 放宽假设III 76

M 第5节补充细节 **83**
M.1 初始误差.83 M.2 外在 Wasserstein 界.84

N 技术引理 **85**

附录A. 记号

本文中使用的符号说明如下。

(i) **字母：** 我们用粗体字母（如 $\boldsymbol{X}$ 和 $\boldsymbol{W}$）表示随机变量（向量），用非粗体字母表示确定性变量。由于我们主要研究 $\mathbb{R}^n$中的向量，因此不区分标量和向量，具体含义由上下文确定。特别地， $I_n \in \mathbb{R}^{n\times n}$ 表示单位矩阵，而 id 表示恒等映射。此外，对于任意集合 A， $\mathbb{I}_A$ 表示示性函数，即当 x满足 $=1$时， $\mathbb{I}_A(x)$ 为 $x \in A$，否则为 0。

(ii) **范数：** 在 $\mathbb{R}^n = \{(x_1, \ldots, x_n)$ 上： $x_i \in \mathbb{R}\}$，标准内积为 $\langle x, y\rangle = \sum_i x_i y_i$，范数定义为$\|x\| = \sqrt{\langle x, x\rangle}$。对于分别配备范数 $\|\cdot\|_{\mathcal{V}}$和 $\|\cdot\|_{\mathcal{W}}$的向量空间 $\mathcal{V}, \mathcal{W}$，线性映射 F: $\mathcal{V} \to \mathcal{W}$称为有界的，如果存在 $K > 0$，使得对于任意 $v \in \mathcal{V}$,

$$\|F(v)\|_{\mathcal{W}} \le K\|v\|_{\mathcal{V}}.$$

在这种情况下，我们定义算子范数

$$\|F\|_{\mathrm{op}} := \sup_{0\neq v\in\mathcal{V}} \frac{\|F(v)\|_{\mathcal{W}}}{\|v\|_{\mathcal{V}}} \le K < \infty.$$

特别地，如果 $Q \in \mathbb{R}^{n\times k}$，那么 $\|Q\|_{\mathrm{op}} = \sup_{v\neq 0}\|Qv\|/\|v\|$。此外，弗罗贝尼乌斯范数定义为 $\|Q\|_{\mathrm{F}} := \sqrt{\mathrm{tr}(QQ^T)}$。

(iii) 光滑性：对于开集 $\mathcal{V} \subset \mathbb{R}^n$，若函数 f: $\mathcal{V} \to \mathbb{R}$ 无穷次可微（若其具有直到 k 阶的连续导数，则记为 $f \in C^k(\mathcal{V})$），则称其为光滑函数，记作 $f \in C^\infty(\mathcal{V})$。对于 $\mathcal{M} \subset \mathbb{R}^n$ 黎曼子流形，若对任意 $x \in \mathcal{M}$，存在 x 的一个开邻域 $\mathcal{V}_x \subset \mathbb{R}^n$ 和一个 $\tilde{f} \in C^\infty(\mathcal{V}_x)$，使得 $\tilde{f}|_{\mathcal{V}_x\cap\mathcal{M}} = f|_{\mathcal{V}_x\cap\mathcal{M}}$，则称 $f \in C^\infty(\mathcal{M})$。映射 $F = (F^1, \cdots, F^k)$: $\mathcal{M} \to \mathbb{R}^k$ 称为光滑映射（C^∞），如果每个 $F^i \in C^\infty(\mathcal{M})$。对于两个黎曼子流形 $\mathcal{M} \subset \mathbb{R}^n$ 和 $\mathcal{N} \subset \mathbb{R}^k$，映射 G: $\mathcal{N} \to \mathcal{M}$ 称为 C^∞，如果 $\iota \circ G$: $\mathcal{N} \to \mathbb{R}^n$ 是 C^∞，其中 ι: $\mathcal{M} \hookrightarrow \mathbb{R}^n$ 是典范嵌入。最后，若 G: $\mathcal{N} \to \mathcal{M}$ 是双射，且 G 和 G^{-1} 均为光滑映射，则称其为微分同胚。

(iv) **可积：** 在概率空间 $(\Omega\ \mathcal{F},\ \mathbb{P})$ 上，取值于 $\mathbb{R}^n$ 的随机变量 $\boldsymbol{X}$ 若满足 L^p 可积条件 $1 \le p < \infty$，记作 $\boldsymbol{X} \in L^p$，则称其为可积的。

附录 B. 更多预备知识

在本节中，我们提供后续分析所必需的更多预备知识。

B.1 关于黎曼子流形的更多内容

设 $\mathcal{M} \subset \mathbb{R}^n$ 为维数为 m的黎曼子流形，并设 $\iota\colon\ \mathcal{M} \hookrightarrow \mathbb{R}^n$ 为典范嵌入。以下概念出现在我们的分析中；更多细节请参见Lee（2018）。

Differential of function. 令 $f \in C^\infty(\mathcal{M})$。对于任意 $x \in \mathcal{M}$ 和任意 $v \in T_x\mathcal{M}$,

$$v(f) := \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} f(\gamma(t)),$$

其中， γ 是 $\mathcal{M}$ 上的一条光滑曲线，且 $\gamma'(0) = v$。映射 $df_x\colon\ T_x\mathcal{M} \to \mathbb{R}$ 是由 $df_x(v) := v(f)$定义的1-形式。此外，对于 $X \in \Gamma(T\mathcal{M})$， $X(f) \in C^\infty(\mathcal{M})$由$X(f)(x) := X(x)(f)$定义，而 $df(X) \in C^\infty(\mathcal{M})$由 $df(X) = X(f)$给出。

Covariant derivative. 令 ∇ 为 Levi-Civita联络。对于 X， $Y \in \Gamma(T\mathcal{M})$，根据第3节所示的定义，这也称为高斯公式（Lee, 2018, 定理 8.2）

$$\nabla_X Y = D_X Y - \mathrm{II}(X, Y),^1$$

对于一条光滑曲线 $\gamma\colon\ [0,1] \to \mathcal{M}$， γ也被视为 $\mathbb{R}^n$上的光滑曲线。因此，在下文中，我们用 $\dot\gamma(t)$表示 $\gamma'(t)$，以强调 $\gamma'(t) \in T_{\gamma(t)}\mathcal{M}$。此外，根据高斯公式，

$$\gamma''(t) = D_{\gamma'(t)}\gamma'(t) = \nabla_{\dot\gamma(t)}\dot\gamma(t) + \mathrm{II}(\dot\gamma(t), \dot\gamma(t)).$$

对于 $f \in C^\infty(\mathcal{M})$， f, $\nabla^2 f$ 的海森矩阵： $\Gamma(T\mathcal{M}) \times \Gamma(T\mathcal{M}) \to C^\infty(M)$，定义为

$$\nabla^2 f(X, Y) = \nabla_Y(\nabla f)(X) = Y(X(f)) - \nabla_Y X(f),$$

它是 $C^\infty(\mathcal{M})$-双线性且对称的。

Ricci curvature. 对于任意 $X, Y, Z \in \Gamma(T\mathcal{M})$，曲率张量 R: $\Gamma(T\mathcal{M})\times\Gamma(T\mathcal{M})\times\Gamma(T\mathcal{M}) \to \Gamma(T\mathcal{M})$ 定义为

$$R(X, Y)Z := \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z,$$

其中 $[X, Y] = XY - YX$ 是李括号。使用相同的记号，黎曼曲率张量 R 为

$$R(W, Z, X, Y) := \langle R(X, Y)Z, W\rangle \in C^\infty(M).$$

对于任意 $x \in \mathcal{M}$ 以及任意线性无关的 $v, w \in T_x\mathcal{M}$，与 v, w 相关联的 x 处的截面曲率为

$$K_x(v, w) := \frac{R_x(v, w, v, w)}{\|v\|^2\|w\|^2 - \langle v, w\rangle^2}.$$

对于任意 $x \in \mathcal{M}$ 和满足 $\|v\| = 1$ 的 $v \in T_x\mathcal{M}$，令 $\{E_i(x)\}_{i=1}^m$ 为一个标准正交基，且 $E_1(x) = v$。则 x 处沿方向 v 的里奇曲率为

$$\mathrm{Ric}_x(v) := \sum_{i=2}^m K_x(v, E_i(x)).$$

1. D 被应用于 $\mathbb{R}^n$上的延拓 $\tilde X, \tilde Y$；我们将 $\tilde X, \tilde Y$ 记作 X, Y，因为该结果与延拓无关。

Local chart. 设（$\mathcal{U}_\alpha, \phi_\alpha$）为一个局部坐标图，即 ϕ_α: $\mathcal{U}_\alpha \to \phi_\alpha$（$\mathcal{U}_\alpha$）是一个微分同胚。对任意 $x \in \mathcal{M}$，令 ϕ_α^{-1}（x）=（$x^1, \ldots, x^m$），称为 x的局部坐标。 x 处的切空间是由.....张成的向量空间

$$T_x\mathcal{M} = \operatorname{span}\left\{ \frac{\partial}{\partial x^1}\bigg|_x, \cdots, \frac{\partial}{\partial x^m}\bigg|_x \right\},$$

其中 $\frac{\partial}{\partial x^i}\big|_x$表示向量，对于 $\iota \circ \phi_\alpha$: $\mathcal{U}_\alpha \to \mathbb{R}^n$,

$$\frac{\partial}{\partial x^i}\bigg|_x \iota \circ \phi_\alpha(x^1, \ldots, x^m) \in \mathbb{R}^n.$$

对于任意 $x \in \mathcal{M}$，黎曼度量 $g(x)$: $T_x\mathcal{M} \times T_x\mathcal{M} \to \mathbb{R}$ 定义为

$$g(x)(u, v) := \langle u, v \rangle, \quad \forall\, u, v \in T_x\mathcal{M}.$$

利用上述局部坐标,

$$g_{ij}(x) := \left\langle \frac{\partial}{\partial x^i}\bigg|_x, \frac{\partial}{\partial x^j}\bigg|_x \right\rangle.$$

T因此， $g(x)$ =$(g_{ij}(x))$表示为一个矩阵。其逆矩阵记为 $g^{-1}(x)$ = $(g^{ij}(x))$。利用局部坐标, 克里斯托费尔系数 Γ_{ij}^k 定义为

$$\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^i} = \sum_{k=1}^m \Gamma_{ij}^k \frac{\partial}{\partial x^k}.$$

此外，由于 ∇ 是 Levi-Civita,

$$\Gamma_{ij}^k = \frac{1}{2} \sum_{\alpha=1}^m g^{k\alpha} \left(\frac{\partial}{\partial x^i} g_{j\alpha} + \frac{\partial}{\partial x^j} g_{i\alpha} - \frac{\partial}{\partial x^\alpha} g_{ij} \right).$$

B.2 可达半径

回顾管状邻域定理（Lee, 2018, 定理 5.25）。

定理 B.1 设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个黎曼子流形。则存在一个光滑函数 ε: $\mathcal{M} \to (0, \infty)$，使得对于

$$\mathcal{U}_0 = \{(x, \xi) \in N\mathcal{M} \colon \|\xi\| < \varepsilon(x)\} \subset N\mathcal{M},$$

由 $T(x, \xi) = x + \xi$ 定义的映射 T: $\mathcal{U}_0 \to \mathcal{V}_0 = T(\mathcal{U}_0) \subset \mathbb{R}^n$ 是一个微分同胚。

该定理保证 $\mathcal{M} \subset \mathbb{R}^n$ 总是存在管状邻域。然而，一致管状邻域的存在性，即 $\inf_{x\in\mathcal{M}} \varepsilon(x) > 0$，并非显而易见。为使这一条件更加明确，我们用 Federer 的达到度来重新表述（Federer，1959；Leobacher 和 Steinicke，2021年）。

定义 B.1（可达半径）设 $\mathcal{M} \subset \mathbb{R}^n$ 为一个黎曼子流形。 $\mathcal{M}$ 的可达半径为

$$\operatorname{reach}(\mathcal{M}) := \sup\{\varepsilon > 0 \colon \forall\ y\ s.t.\ d(y,\mathcal{M}) < \varepsilon,\ \exists\,!\ x \in \mathcal{M},\ s.t.\ \|x-y\| = d(y,\mathcal{M})\},$$

即，使得任何与 $\mathcal{M}$的距离小于该半径的点都存在唯一投影到 $\mathcal{M}$上的最大半径。

可达半径与一致管状半径之间的关系由以下定理给出（Leobacher 和 Steinicke，2021年；Platt 和Gal´an，2025年）。

定理 B.2 设 $\mathcal{M} \subset \mathbb{R}^n$ 是无边界的闭黎曼子流形。若r_0 =可达半径$(\mathcal{M}) > 0$，则

$$\mathcal{U}_{r_0} = \{(x,\xi) \in N\mathcal{M} \colon \|\xi\| < r_0\}, \quad \mathcal{V}_{r_0} = \{y \in \mathbb{R}^n \colon d(y,\mathcal{M}) < r_0\},$$

管状映射 $T\colon\ \mathcal{U}_{r_0} \to \mathcal{V}_{r_0}$由 $T(x,\xi) = x + \xi$给出，是一个微分同胚。特别地， $\mathcal{V}_{r_0}$是一致管状邻域，其半径为 r_0。

注 B.1 反方向显然成立。若 $\mathcal{M}$ 存在半径为 r的一致管状邻域，则对任意满足 $d(y,\mathcal{M}) < r$的 y，点 $x = \pi(y)$ 是 $\mathcal{M}$中唯一的最近点。因此，可达半径$(\mathcal{M}) \geq r > 0$。

因此，假设 II 等价于 reach$(\mathcal{M}) > 0$。

B.3 鞅论

以下鞅论中的内容对我们的分析至关重要；更多细节请参见 Le Gall (2016)。

固定一个滤过概率空间 $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathbb{P})$。一个取值于 $\mathbb{R}^n$ 的随机过程 $(\boldsymbol{X}_t)_{t\geq 0}$ 称为 $(\mathcal{F}_t)_{t\geq 0}$-鞅，若 $\boldsymbol{X}_t \in L^1$，$\boldsymbol{X}_t$ 是 $\mathcal{F}_t$-可测的，并且对任意$t \geq s \geq 0$,

$$\mathbb{E}[\boldsymbol{X}_t \mid \mathcal{F}_s] = \boldsymbol{X}_s.$$

同一定义也适用于离散时间，即 $t = n \in \mathbb{N}$。此外，对于连续时间过程 $(\boldsymbol{X}_t)_{t\geq 0}$，我们总是假设其具有连续样本路径，即 $t \mapsto \boldsymbol{X}_t$ 几乎处处连续。

设 $\boldsymbol{\tau}\colon \Omega \to [0,\infty]$。若对所有 $t \geq 0$ 都有 $\{\boldsymbol{\tau} \leq t\} \in \mathcal{F}_t$，则称 $\boldsymbol{\tau}$ 为 $(\mathcal{F}_t)_{t\geq 0}$-停时。若 $\boldsymbol{X}_0 = 0$，且存在一列停时 $\boldsymbol{\tau}_n \uparrow \infty$，它们几乎处处趋于无穷，使得对所有 n，$(\boldsymbol{X}_{t\wedge\boldsymbol{\tau}_n})_{t\geq 0}$ 都是鞅，则称过程 $(\boldsymbol{X}_t)_{t\geq 0}$ 为局部鞅。

注意，任意鞅都是局部鞅，但反之未必成立。伊藤积分是局部鞅的一个基本例子。设 $\boldsymbol{W}$ 是 $\mathbb{R}^n$ 上的标准布朗运动。若伊藤积分

$$\boldsymbol{M}_t = \int_0^t \boldsymbol{X}_s\ \mathrm{d}\boldsymbol{W}_s$$

定义良好（即， $\boldsymbol{X}_t$是渐进可测的，并且 $\int_0^t \|\boldsymbol{X}_s\|^2\,\mathrm{d}s < \infty$几乎处处成立），则 $\boldsymbol{M}_t$ 是一个局部鞅。对于局部鞅（$\boldsymbol{M}_t$）$_{t\geq 0}$，其（标量）二次变差（$[\boldsymbol{M}]_t$）$_{t\geq 0}$ 为

$$[\boldsymbol{M}]_t = \int_0^t \|\boldsymbol{X}_s\|^2\ \mathrm{d}s.$$

附录 C. 假设 I 下的几何有界性

对于典范嵌入 $\iota\colon\ \mathcal{M} \hookrightarrow \mathbb{R}^n$，假设 I 推出了 $\mathcal{M}$所需的几何性质。我们将在本节中证明这些推论。

我们首先明确 $\nabla d\iota$ 和 $\nabla^2 d\iota$的定义。由于 $\iota\colon \mathcal{M} \to \mathbb{R}^n$，微分 $d\iota$ 是 $\mathcal{M}$上的向量值 1-形式。令 $\iota = (\iota^1, \dots, \iota^n)$，则有

$$d\iota = (d\iota^1, \dots, d\iota^n).$$

相应地，协变导数 $\nabla d\iota$ 由下式自然定义：

$$(\nabla_X d\iota)(Y) = D_X(d\iota(Y)) - d\iota(\nabla_X Y), \quad \forall\ X, Y \in \Gamma(T\mathcal{M}), \tag{C.1}$$

其中 $D_X = D_{\iota(X)}$是通过嵌入将 X 视为 $\mathbb{R}^n$ 中的向量场。注意，$\nabla d\iota$是 $\mathcal{M}$上的一个 $\mathbb{R}^n$ 值 2-形式，定义为

$$(\nabla d\iota)(Y, X) = (\nabla_X d\iota)(Y), \quad \forall\ X, Y \in \Gamma(T\mathcal{M}).$$

由于 ∇ 和 D 都是列维-奇维塔联络，因此 $\nabla d\iota$ 是对称的

类似地，我们定义 $\mathbb{R}^n$ 值 3-形式 $\nabla^2 d\iota = \nabla(\nabla d\iota)$ 如下：对于 X，Y，$Z \in \Gamma(T\mathcal{M})$，

$$(\nabla(\nabla d\iota))(X, Y, Z) = D_Z\left((\nabla d\iota)(X, Y)\right) - (\nabla d\iota)(\nabla_Z X, Y) - (\nabla d\iota)(X, \nabla_Z Y). \tag{C.2}$$

更多细节可参见 Tu（2017年）和 Taubes（2011年）。

C.1 第二基本形式的有界性

我们先澄清 ∇II 的定义。对于 $\mathcal{M} \subset \mathbb{R}^n$ 一个黎曼子流形，回顾第二基本形式 II 的定义：

$$\mathrm{II}(X, Y) = (D_X Y)^{\perp} = D_X Y - (D_X Y)^{\top}, \quad \forall\ X, Y \in \Gamma(T\mathcal{M}),$$

其中 $(D_X Y)^{\top}$ 表示从 $D_X Y$到 $T\mathcal{M}$ 的正交投影。此外，II 的协变导数可定义如下：对所有 X，Y，$Z \in \Gamma(T\mathcal{M})$，

$$(\nabla \mathrm{II})(X, Y, Z) = \left(D^{\perp}_Z \mathrm{II}\right)(X, Y) - \mathrm{II}(\nabla_Z X, Y) - \mathrm{II}(X, \nabla_Z Y), \tag{C.3}$$

其中 $D^{\perp}_Z \mathrm{II}(X,\ Y) := (D_Z \mathrm{II}(X,\ Y))^{\perp}$。注意，$(\nabla \mathrm{II})(X,\ Y,\ Z) \in \Gamma(N\mathcal{M})$；详情可参见 Lee（2018）。为方便起见，对于任意 $x \in \mathcal{M}$，记 $\nabla \mathrm{II}_x := \nabla \mathrm{II}(x)$。

为了将假设 I 与第二基本形式的有界性联系起来，我们首先建立如下显式公式。

引理 C.1 设 $(\mathcal{M}, g)$ 为通过映射 $\iota\colon \mathcal{M} \hookrightarrow \mathbb{R}^n$ 等距嵌入 $\mathbb{R}^n$ 的黎曼流形。设 II 为 $\mathcal{M}$ 的第二基本形式。则对 $X, Y, Z \in \Gamma(T\mathcal{M})$，

$$\mathrm{II}(X, Y) = (\nabla d\iota)(X, Y), \quad \left(\nabla^2 d\iota\right)(X, Y, Z) = (\nabla \mathrm{II})(X, Y, Z) - S_{\mathrm{II}(X,Y)}(Z),$$

其中 $S_{\mathrm{II}(X,Y)}$ 是沿着 $\mathrm{II}(X, Y)$的形算子。

证明 根据定义 (C.1),

$$D_X(Y) = \nabla_X Y + (\nabla d\iota)(X, Y),$$

其中我们将 $Y = d\iota(Y)$ 和 $\nabla_X Y = d\iota(\nabla_X Y)$ 视为 $\mathbb{R}^n$ 中的向量场。另一方面，根据高斯公式（Lee, 2018，定理 8.2），

$$D_X(Y) = \nabla_X Y + \mathrm{II}(X, Y).$$

比较这两个恒等式可得 $\mathrm{II}(X, Y) = (\nabla d\iota)(X, Y)$。

对于第二个恒等式，令 $A := \nabla d\iota$。根据定义 (C.3),

$$(\nabla A)(X, Y, Z) = D_Z\left(A(X, Y)\right) - A(\nabla_Z X, Y) - A(X, \nabla_Z Y), \quad \forall\, X, Y, Z \in \Gamma\left(T\mathcal{M}\right).$$

利用 $A = \mathrm{II}$，并将 $D_Z(\mathrm{II}(X, Y))$ 分解为切向分量和法向分量，我们得到

$$\left(\nabla^2 d\iota\right)(X, Y, Z) = (\nabla \mathrm{II})(X, Y, Z) + \left(D_Z\left(\mathrm{II}(X, Y)\right)\right)^{\top}.$$

由于 $\mathrm{II}(X, Y) \in \Gamma(N\mathcal{M})$，魏因加滕方程（Lee, 2018，命题 8.4）表明

$$\left(D_Z\left(\mathrm{II}(X, Y)\right)\right)^{\top} = -S_{\mathrm{II}(X,Y)}(Z).$$

因此,

$$\left(\nabla^2 d\iota\right)(X, Y, Z) = (\nabla \mathrm{II})(X, Y, Z) - S_{\mathrm{II}(X,Y)}(Z).$$

■

注 C.1 为简便起见，我们有时使用 $\mathrm{II} = \nabla d\iota$ 和 $\nabla \mathrm{II} = \nabla^2 d\iota + S_{\mathrm{II}}$。

引理 C.2 设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个黎曼子流形。

(i) 若 $\kappa = \sup_{\mathcal{M}} \|\mathrm{II}\|_{\mathrm{op}} < \infty$，则对任意法向量场 $\eta \in \Gamma\left(N\mathcal{M}\right)$，形算子 S_η 满足

$$\|S_\eta\|_{\mathrm{op}} \leq \kappa \|\eta\|.$$

(ii) 若

$$r = \sup_{x \in \mathcal{M}} \sup_{\xi \in N_x\mathcal{M}} \frac{\|S_{x,\xi}\|_{\mathrm{op}}}{\|\xi\|} < \infty,$$

则 $\sup_{x \in \mathcal{M}} \|\mathrm{II}_x\|_{\mathrm{op}} \leq r$。

证明 回忆一下，形算子（或魏因加滕映射）$S_{x,\,\xi}\colon\ T_x\mathcal{M} \to T_x\mathcal{M}$ 沿法向方向 ξ 定义为 (Lee, 2018，第8章)

$$\langle S_{x,\xi}(u), v\rangle = \langle \mathrm{II}_x(u, v), \xi\rangle, \quad \forall\, u, v \in T_x\mathcal{M}.$$

(i) 对于任意 $\eta \in \Gamma(N\mathcal{M})$,

$$\|S_\eta\|_{\mathrm{op}} = \sup_{X,Y \neq 0} \frac{|\langle S_\eta(X), Y\rangle|}{\|X\|\|Y\|} \leq \sup_{X,Y \neq 0} \frac{\|\mathrm{II}(X, Y)\|}{\|X\|\|Y\|} \|\eta\| \leq \kappa \|\eta\|.$$

(ii) 根据定义，

$$\|\mathrm{II}_x\|_{\mathrm{op}} = \sup_{u,v\neq 0} \frac{\|\mathrm{II}_x(u,v)\|}{\|u\|\|v\|} = \sup_{u,v\neq 0}\sup_{\xi\neq 0} \frac{|\langle \mathrm{II}_x(u,v),\xi\rangle|}{\|u\|\|v\|\|\xi\|} = \sup_{\xi\in N_x\mathcal{M}} \frac{\|S_{x,\xi}\|_{\mathrm{op}}}{\|\xi\|} \le r.$$

对 $x\in\mathcal{M}$ 取上确界可得

$$\sup_{x\in\mathcal{M}} \|\mathrm{II}_x\|_{\mathrm{op}} \le r.$$

■

引理 C.3 设 $\mathcal{M}\subset\mathbb{R}^n$ 是一个黎曼子流形，其嵌入映射为 ι。假设 *I* 等价于

$$\sup_{x\in\mathcal{M}} \|\mathrm{II}_x\|_{\mathrm{op}} < \infty, \quad \sup_{x\in\mathcal{M}} \|\nabla \mathrm{II}_x\|_{\mathrm{op}} < \infty,$$

其中 II 表示 $\mathcal{M}$ 的第二基本形式。

证明 假设 ι 满足假设 I。由引理 C.1，

$$\sup_{\mathcal{M}} \|\mathrm{II}\|_{\mathrm{op}} = \sup_{\mathcal{M}} \|\nabla d\iota\|_{\mathrm{op}} < \infty.$$

此外，由于 $\nabla^2 d\iota = \nabla \mathrm{II} - S_{\mathrm{II}}$ 且 $\nabla \mathrm{II} \perp S_{\mathrm{II}}$，

$$\|\nabla \mathrm{II}\|_{\mathrm{op}} \le \sup_{X,Y,Z} \frac{\left\|\left(\nabla^2 d\iota\right)(X,Y,Z)\right\|}{\|X\|\|Y\|\|Z\|} = \left\|\nabla^2 d\iota\right\|_{\mathrm{op}}.$$

因此，

$$\sup_{\mathcal{M}} \|\nabla \mathrm{II}\|_{\mathrm{op}} \le \sup_{\mathcal{M}} \left\|\nabla^2 d\iota\right\|_{\mathrm{op}} < \infty.$$

反过来，假设 $\sup_{\mathcal{M}}\|\mathrm{II}\|_{\mathrm{op}} < \infty$且 $\sup_{\mathcal{M}}\|\nabla \mathrm{II}\|_{\mathrm{op}} < \infty$。由于 $\mathrm{II} = \nabla d\iota$，可立即得出 $\sup_{\mathcal{M}}\|\nabla d\iota\|_{\mathrm{op}} < \infty$。对于二阶导数项，令 $\kappa = \sup_{\mathcal{M}}\|\mathrm{II}\|_{\mathrm{op}} < \infty$。由引理 C.2，

$$\left\|S_{\mathrm{II}(X,Y)}\right\|_{\mathrm{op}} \le \kappa \|\mathrm{II}(X,Y)\| \le \kappa^2 \|X\|\|Y\|, \quad \text{i.e., } \|S_{\mathrm{II}}\|_{\mathrm{op}} \le \kappa^2.$$

因此，再次利用 $\nabla^2 d\iota = \nabla \mathrm{II} - S_{\mathrm{II}}$，我们得到

$$\left\|\nabla^2 d\iota\right\|_{\mathrm{op}} \le \|\nabla \mathrm{II}\|_{\mathrm{op}} + \|S_{\mathrm{II}}\|_{\mathrm{op}} < \infty.$$

■

C.2 指数映射的泰勒公式

由于 $\mathcal{M} \subset \mathbb{R}^n$，我们可以对指数映射 $\exp_x$ 进行泰勒展开： $T_x\mathcal{M} \to \mathcal{M}$。此外，假设 I 有助于控制余项。

引理 C.4 设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个黎曼子流形，其第二基本形式满足

$$\kappa_1 := \sup_{x\in\mathcal{M}} \|\mathrm{II}_x\|_{\mathrm{op}} < \infty, \quad \kappa_2 := \sup_{x\in\mathcal{M}} \|\nabla \mathrm{II}_x\|_{\mathrm{op}} < \infty.$$

对于任意 $x \in \mathcal{M}$，令 $\exp_x\colon T_x\mathcal{M} \to \mathcal{M}$ 为指数映射。则对于任意 $v \in T_x\mathcal{M}$，

$$\exp_x(v) = x + v + \frac{1}{2}\mathrm{II}_x(v, v) + R_3(x, v),$$

其中

$$\|R_3(x, v)\| \le \frac{\kappa_1^2 + \kappa_2}{6}\|v\|^3.$$

证明 固定 $x \in \mathcal{M}$ 和 $v \in T_x\mathcal{M}$。首先，根据 Hopf–Rinow定理（Lee, 2018，定理 6.19），由于 $\mathcal{M}$ 的闭性，$\exp_x$ 在 $T_x\mathcal{M}$ 上有定义。令 $\gamma\colon [0, 1] \to \mathcal{M}$ 由 $\gamma(t) = \exp_x(tv)$定义，即满足下式的测地线

$$\gamma(0) = x, \quad \dot\gamma(0) = v.$$

由于 $\gamma(t) \in \mathbb{R}^n$，我们将其欧氏导数记为 $\gamma'(t)$，显然 $\gamma'(t) = \dot\gamma(t)$。然而，二阶导数应作外蕴解释。根据高斯公式（Lee, 2018，定理 8.2），

$$\gamma''(t) = D_{\gamma'(t)}\gamma'(t) = \nabla_{\dot\gamma(t)}\dot\gamma(t) + \mathrm{II}(\dot\gamma(t), \dot\gamma(t)) = \mathrm{II}(\dot\gamma(t), \dot\gamma(t)),$$

由于 γ 是一条测地线，因而 $\nabla_{\dot\gamma(t)}\dot\gamma(t) \equiv 0$。对$\gamma(t)$应用欧氏泰勒公式，可得

$$\begin{aligned}\gamma(1) &= \gamma(0) + \gamma'(0) + \frac{1}{2}\gamma''(0) + \int_0^1 \frac{(1-t)^2}{2}\gamma'''(t)\ \mathrm{d}t \\ &= x + v + \frac{1}{2}\mathrm{II}(v, v) + R_3(x, v),\end{aligned}$$

其中

$$R_3(x, v) = \int_0^1 \frac{(1-t)^2}{2}\gamma'''(t)\ \mathrm{d}t.$$

对于 γ'''，

$$\gamma'''(t) = D_{\gamma'(t)}\gamma''(t) = D_{\dot\gamma(t)}\mathrm{II}(\dot\gamma(t), \dot\gamma(t)).$$

令 $X = \dot\gamma(t)$。由魏因加滕方程（Lee, 2018，命题 8.4），

$$\gamma'''(t) = D_X\mathrm{II}(X, X) = -S_{\mathrm{II}(X,X)}(X) + D^\perp_X\mathrm{II}(X, X). \tag{C.4}$$

对于第一项，由于

$$\kappa_1 = \sup_{\mathcal{M}}\|\mathrm{II}\|_{\mathrm{op}} < \infty,$$

引理 C.2 表明

$$\left\|S_{\mathrm{II}(X,X)}(X)\right\| \leq \left\|S_{\mathrm{II}(X,X)}\right\|_{\mathrm{op}}\|X\| \leq \kappa_1\|\mathrm{II}(X,X)\|\|X\| \leq \kappa_1^2\|X\|^3. \tag{C.5}$$

对于第二项，由 (C.3),

$$\begin{aligned} D_X^{\perp}\mathrm{II}(X,X) &= \left(\nabla_X \mathrm{II}\right)(X,X) + \mathrm{II}\left(\nabla_X X, X\right) + \mathrm{II}\left(X, \nabla_X X\right) \\ &= \left(\nabla_X \mathrm{II}\right)(X,X), \end{aligned}$$

因为 $\nabla_X X = 0$ 对于 $X = \dot{\gamma}(t)$。因此,

$$\left\|D_X^{\perp}\mathrm{II}(X,X)\right\| = \left\|\left(\nabla_X \mathrm{II}\right)(X,X)\right\| \leq \kappa_2\|X\|^3. \tag{C.6}$$

结合 (C.5)、(C.6) 和 (C.4)，我们得到

$$\left\|\gamma'''(t)\right\| \leq \left\|S_{\mathrm{II}(X,X)}(X)\right\| + \left\|D_X^{\perp}\mathrm{II}(X,X)\right\| \leq \left(\kappa_1^2 + \kappa_2\right)\|X\|^3.$$

此外，由于 $\gamma(t)$ 是一条测地线，$\|X\| = \|\dot{\gamma}(t)\| \equiv \|\dot{\gamma}(0)\| = \|v\|$。因此,

$$\|R_3(x,v)\| \leq \int_0^1 \frac{(1-t)^2}{2}\left\|\gamma'''(t)\right\| \,\mathrm{d}t \leq \frac{\kappa_1^2 + \kappa_2}{6}\|v\|^3.$$

■

C.3 黎曼曲率有界性

根据引理 C.3，假设 I 等价于第二基本形式及其协变导数具有一致有界性。尽管这些条件是外蕴的，它们也蕴含了关于 $\mathcal{M}$ 的内蕴几何信息。

引理 C.5 设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个 m 维黎曼子流形，且其第二基本形式一致有界，即 $\kappa = \sup_{\mathcal{M}}\|\mathrm{II}\|_{\mathrm{op}} < \infty$。则其里奇曲率满足

$$\mathrm{Ric} \geq -2(m-1)\kappa^2 g.$$

证明 设 R 是 $\mathcal{M}$ 的黎曼曲率张量。对于任意 X, Y, Z, $W \in \Gamma(T\mathcal{M})$，高斯公式（Lee, 2018, 定理 8.5）给出

$$\overline{R}(W,X,Y,Z) = R(W,X,Y,Z) - \langle \mathrm{II}(W,Y), \mathrm{II}(X,Z)\rangle + \langle \mathrm{II}(X,Y), \mathrm{II}(W,Z)\rangle,$$

其中 $\overline{R}$ 是 $\mathbb{R}^n$ 的黎曼曲率张量。由于 $R = 0$，我们得到

$$R(W,X,Y,Z) = \langle \mathrm{II}(W,Y), \mathrm{II}(X,Z)\rangle - \langle \mathrm{II}(X,Y), \mathrm{II}(W,Z)\rangle.$$

因此，对于正交规范的 X, Y, $\mathcal{M}$ 的截面曲率 K 满足

$$K(X,Y) = R(X,Y,X,Y) = \langle \mathrm{II}(X,X), \mathrm{II}(Y,Y)\rangle - \|\mathrm{II}(X,Y)\|^2.$$

由于 $\kappa = \sup_{\mathcal{M}} \|\mathrm{II}\|_{\mathrm{op}} < \infty$，我们有 $\|\mathrm{II}(X, Y)\| \le \kappa \|X\| \|Y\|$，因此

$$K(X,Y) \ge -\|\mathrm{II}(X,X)\|\|\mathrm{II}(Y,Y)\| - \|\mathrm{II}(X,Y)\|^2 \ge -2\kappa^2.$$

对于单位向量 X，设 $\{E_i\}_{i=1}^m$ 是一个满足 $E_1 = X$ 的标准正交标架。那么

$$\mathrm{Ric}(X,X) = \sum_{i=2}^{m} K(E_i, X) \ge -2(m-1)\kappa^2.$$

■

附录 D. 正交投影及其有界性

由于我们从外在视角研究流形 $\mathcal{M} \subset \mathbb{R}^n$，因此需要考虑正交投影 $P(x)$：对于每个 $x \in \mathcal{M}$，$\mathbb{R}^n \to \mathbb{R}^n$ 的像 $\mathrm{Im}\ P(x)$ 为 $= T_x\mathcal{M}$。管状邻域定理（定理 B.1）保证了此类投影的存在性；参见 D.1 节。此外，我们还研究由管状映射诱导的典范投影的有界性（D.2 节），以及相应正交投影 P 的有界性。

D.1 正交投影的存在性

引理 D.1 设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个带有某个管状邻域$\mathcal{V}$的黎曼子流形。那么对于任意 $x \in \mathcal{M}$,

$$P(x) := D\pi(x),$$

是 th 到 $T_x\mathcal{M}$ 上的正交投影，其中 $\pi\colon \mathcal{V} \to \mathcal{M}$ 是典范投影 n.

证明 固定 $x \in \mathcal{M}$。首先，对于任意 $v \in T_x\mathcal{M}$，令 $c\colon (-\varepsilon, \varepsilon) \to \mathcal{M}$ 为一条曲线，使得 $c(0) = x$且 $\dot{c}(0) = v$。由于 $\pi|_{\mathcal{M}} = \mathrm{id}_{\mathcal{M}}$，我们有 $\pi(c(t)) = c(t)$。由此可得

$$D\pi(x)[v] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \pi(c(t)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} c(t) = v.$$

因此，$D\pi(x)\colon \mathbb{R}^n = T_x\mathcal{M} \oplus N_x\mathcal{M} \to \mathbb{R}^n$ 在 $T_x\mathcal{M}$ 上限制为恒等映射。特别地，$\mathrm{Im}\ D\pi(x) \supset T_x\mathcal{M}$。另一方面，由于 $\pi\colon \mathcal{V}_0 \to \mathcal{M}$，我们有 $\mathrm{Im}\ D\pi(x) \subset T_x\mathcal{M}$。因此，

$$\mathrm{Im}\, D\pi(x) = T_x\mathcal{M}.$$

接下来，对于任意 $\xi \in N_x\mathcal{M}$，考虑曲线 $c(t) = x + t\xi$。对于 $|t| < \varepsilon(x)$，典范投影的定义意味着 $\pi(c(t)) \equiv x$。因此，

$$D\pi(x)[\xi] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \pi(c(t)) = 0,$$

因此 $N_x\mathcal{M} \subset$ 核 $D\pi(x)$。此外，根据秩定理，

$$\dim \ker D\pi(x) = n - \dim \mathrm{Im}\, D\pi(x) = n - \dim T_x\mathcal{M} = \dim N_x\mathcal{M},$$

因此，$D\pi(x)$ 的核为 $= N_x\mathcal{M}$。因此，$P(x) := D\pi(x)$ 是到 $T_x\mathcal{M}$ 上的正交投影。

■

注 D.1 由于 π 在 $\mathcal{V}$ 上光滑，$P(x) = D\pi(x)$ 在 $\mathcal{V}$ 上也光滑，尤其在 $\mathcal{M}$ 上光滑。关于一般流形上的正交投影的进一步讨论，参见 *Leobacher* 和 *Steinicke*（*2021*年）。

D.2 正则投影的有界性

根据典范投影 π （第 3.1 节）的定义，对于任意 $y \in \mathcal{V}_0$，$D\pi$（y）是到 $T_{\pi(y)}\mathcal{M}$ 上的投影，但不一定是正交投影。然而，在等距嵌入的情形下，$D\pi$（y）有一个显式公式，这也为界定 $\|D\pi$（y）$\|_{\mathrm{op}}$ 提供了一种方便的方法。

命题 D.2 设 $\mathcal{M} \subset \mathbb{R}^n$ 为黎曼子流形，$\mathcal{V}_0$ 为 $\mathbb{R}^n$ 中的一个管状邻域，并具有典范投影 π。那么对于任意 $y \in \mathcal{V}_0$,

$$(\mathrm{id} - S_{x,\xi}) \circ D\pi(y) = P(x),$$

其中，$y = x + \xi$ 是满足 $x \in \mathcal{M}$ 和 $\xi \in N_x\mathcal{M}$ 的唯一分解，而 $S_{x,\ \xi}$ 是沿 ξ 在 x 处的形算子。

证明 固定 $y \in \mathcal{V}_0$，并将唯一分解写作 $y = x + \xi$。回忆 $x = \pi$（y）。定义 $\xi\colon \mathcal{V}_0 \to N\mathcal{M}$，使得 ξ（y）$= \xi$；此映射是光滑的。因此，对于任意 $X \in \Gamma$（$T\mathcal{M}$），

$$\left\langle \xi(y), X_{\pi(y)} \right\rangle = 0, \quad \forall\, y \in \mathcal{V}_0.$$

设 $y(t)$ 是 $\mathcal{V}_0$ 中的一条曲线，并令 $x(t) = \pi(y(t))$。对其求导,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left\langle \xi(y(t)), X_{x(t)} \right\rangle = \left\langle D\xi(y)[y'(0)], X_{x(0)} \right\rangle + \left\langle \xi(y(0)), D_{x'(0)}X_{x(0)} \right\rangle. \tag{D.1}$$

固定 $w \in \mathbb{R}^n$ 并选取 $y(t)$，使得 $y(0) = y$ 且 $y'(0) = w$。于是，由链式法则 ,

$$x'(0) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \pi(y(t)) = D\pi(y)[w].$$

将此代入 (D.1) 得

$$\langle D\xi(y)[w], X_x \rangle + \left\langle \xi, D_{D\pi(y)[w]}X_x \right\rangle = 0. \tag{D.2}$$

由于 $D\pi(y)[w] \in T_x\mathcal{M}$，高斯公式（Lee, 2018, 定理 8.2）给出

$$D_{D\pi(y)[w]}X_x = \nabla_{D\pi(y)[w]}X_x + \mathrm{II}_x\left(D\pi(y)[w], X_x\right).$$

因此,

$$\left\langle \xi, D_{D\pi(y)[w]}X_x \right\rangle = \langle \xi, \mathrm{II}_x\left(D\pi(y)[w], X_x\right)\rangle = \langle S_{x,\xi}\left(D\pi(y)[w]\right), X_x\rangle, \tag{D.3}$$

因为 $\left\langle \xi,\ \nabla_{D\pi(y)[w]}X_x \right\rangle = 0$ 对于 $\xi \in N_x\mathcal{M}$。结合 (D.3) 和 (D.2)，我们得到

$$\langle D\xi(y)[w], X_x\rangle + \langle S_{x,\xi}\left(D\pi(y)[w]\right), X_x\rangle = 0.$$

由于这对所有 $X \in \Gamma(T\mathcal{M})$ 都成立,

$$P(x)\left(D\xi(y)[w]\right) = -S_{x,\xi}\left(D\pi(y)[w]\right). \tag{D.4}$$

另一方面，由

$$y = \pi(y) + \xi(y),$$

沿 w 取方向导数，得到

$$w = D\pi(y)[w] + D\xi(y)[w].$$

对两边应用 $P(x)$，并利用 $P(x)\left(D\pi(y)[w]\right) = D\pi(y)[w]$，由于 $D\pi(y)[w] \in T_x\mathcal{M}$，可得

$$P(x)(w) = D\pi(y)[w] + P(x)\left(D\xi(y)[w]\right).$$

利用 (D.4)，可得

$$P(x)(w) = D\pi(y)[w] - S_{x,\xi}\left(D\pi(y)[w]\right) = \left(\mathrm{id} - S_{x,\xi}\right)\left(D\pi(y)[w]\right).$$

由于 $w \in \mathbb{R}^n$ 是任意的，我们得出

$$P(x) = \left(\mathrm{id} - S_{x,\xi}\right) \circ D\pi(y).$$

■

引理 D.3 设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个第二基本形式一致有界的黎曼子流形，即 $\kappa = $上确界$_{\mathcal{M}} \|\mathrm{II}\|_{\mathrm{op}} < \infty$。则存在一个管状邻域 $\mathcal{V}$，其典范投影为 π，且存在一个常数 $0 < r_0 < 1/\kappa$，使得

$$D\pi(y) = \left(\mathrm{id} - S_{x,\xi}\right)^{-1} \circ P(x), \quad \textit{and } \|D\pi(y)\|_{\mathrm{op}} < \frac{1}{1 - r_0\kappa}, \quad \forall\ y \in \mathcal{V}.$$

特别地，当 $\mathcal{M}$ 具有一致管状邻域时， $\mathcal{V}$ 可以选为一致管状邻域。

证明 首先，由命题 D.2，存在一个管状邻域 $\mathcal{V}_0$ 以及典范投影 π，使得对任意 $y \in \mathcal{V}_0$,

$$P(x) = \left(\mathrm{id} - S_{x,\xi}\right) \circ D\pi(y), \tag{D.5}$$

其中 $y = x + \xi$ 是满足 $x \in \mathcal{M}$ 与 $\xi \in N_x\mathcal{M}$的唯一分解。此外，由引理 C.2 以及第二基本形式的一致有界性,

$$\left\|S_{x,\xi}\right\|_{\mathrm{op}} \le \kappa\|\xi\|. \tag{D.6}$$

将 $\mathcal{V}_0$ 收缩为

$$\mathcal{V} := \mathcal{V}_0 \cap \left\{x + \xi \colon \|\xi\| < r_0\right\},$$

其中 $r_0 > 0$ 是满足 $r_0 < 1/\kappa$的常数。若 $\mathcal{M}$ 具有一致管状邻域，则可取 $\mathcal{V}_0 = \{x + \xi \colon \|\xi\| < \varepsilon_0\}$，其一致半径为 ε_0，然后选取

$$\mathcal{V} := \left\{x + \xi \colon \|\xi\| < r_0\right\}, \quad r_0 = \min\left\{\varepsilon_0, 1/\kappa\right\},$$

这也是 $\mathcal{M}$的一个一致管状邻域。

对于任意 $y \in \mathcal{V}$，(D.6) 给出

$$\|S_{x,\xi}\|_{\mathrm{op}} \le r_0\kappa < 1.$$

因此 id $-S_{x,\xi}$ 可逆，并根据诺伊曼级数界（Horn and Johnson，2012），

$$\left\|(\mathrm{id} - S_{x,\xi})^{-1}\right\|_{\mathrm{op}} \le \frac{1}{1 - \|S_{x,\xi}\|_{\mathrm{op}}} < \frac{1}{1 - r_0\kappa}.$$

因此，由 (D.5) 可得，

$$D\pi(y) = (\mathrm{id} - S_{x,\xi})^{-1} \circ P(x), \quad \forall\, y \in \mathcal{V},$$

此外，

$$\|D\pi(y)\|_{\mathrm{op}} \le \left\|(\mathrm{id} - S_{x,\xi})^{-1}\right\|_{\mathrm{op}} \|P(x)\|_{\mathrm{op}} = \left\|(\mathrm{id} - S_{x,\xi})^{-1}\right\|_{\mathrm{op}} < \frac{1}{1 - r_0\kappa},$$

其中 $\|P(x)\|_{\mathrm{op}} = 1$，因为 $P(x)$ 是正交投影。 ■

D.3 正交投影的有界性

为了给出 $P(x)$ 的外在导数的界，我们首先需要在第二基本形式一致有界的条件下内在导数的有界性。

引理 D.4 设 $\mathcal{M} \subset \mathbb{R}^n$ 为第二基本形式一致有界的黎曼子流形，即 $\kappa = \sup_{\mathcal{M}} \|\mathrm{II}\|_{\mathrm{op}} < \infty$。设 $P(x)$ 对 $x \in \mathcal{M}$ 为到 $T_x\mathcal{M}$上的正交投影。则

$$\sup_{x \in \mathcal{M}} \|\nabla P(x)\|_{\mathrm{op}} \le \kappa.$$

证明 首先，我们阐明内在协变导数 $\nabla P(x)$ 的定义。对任意 $w \in \mathbb{R}^n$，$P(x)w \in T_x\mathcal{M}$，因此 $Pw \in \Gamma(T\mathcal{M})$。故 $P \in \Gamma(\mathcal{M}, \mathrm{Hom}(\mathcal{M} \times \mathbb{R}^n, T\mathcal{M}))$。因此，对任意 $X \in \Gamma(T\mathcal{M})$，协变导数 $\nabla_X P \in \Gamma(\mathcal{M}, \mathrm{Hom}(\mathcal{M} \times \mathbb{R}^n, T\mathcal{M}))$ 定义为

$$(\nabla_X P)(x)w := \nabla_X(P(x)w), \quad \forall\, w \in \mathbb{R}^n.$$

对 $w \in \mathbb{R}^n$，写出分解

$$w = w^{\top}(x) + w^{\perp}(x), \tag{D.7}$$

其中

$$w^{\top}(x) = P(x)w \in T_x\mathcal{M}, \quad w^{\perp}(x) = w - w^{\top}(x) \in N_x\mathcal{M}.$$

对任意 $w \in \mathbb{R}^n$，有 Claim: $(\nabla_X P)w = S_{w^{\perp}}(X)$，其中 $S_{w^{\perp}}$ 是沿 $w^{\perp}$ 的形算子。我们现在证明该断言。固定 $w \in \mathbb{R}^n$，并令 $X \in \Gamma(T\mathcal{M})$。由于 w 关于 $\mathbb{R}^n$ 为常数，对 (D.7) 两边取方向导数，可得

$$0 = D_X w = D_X w^{\top} + D_X w^{\perp}. \tag{D.8}$$

对于第一项，由于 $w^{\top} \in \Gamma(T\mathcal{M})$，高斯公式（Lee, 2018，定理 8.2）表明

$$D_X w^{\top} = \nabla_X w^{\top} + \mathrm{II}(X, w^{\top}). \tag{D.9}$$

对于第二项，由于 $w^\perp \in \Gamma(N\mathcal{M})$，魏因加滕方程（Lee, 2018，命题 8.4）给出

$$\left(D_X w^\perp\right)^\top = -S_{w^\perp}(X),$$

因此

$$D_X w^\perp = \left(D_X w^\perp\right)^\top + \left(D_X w^\perp\right)^\perp = -S_{w^\perp}(X) + \left(D_X w^\perp\right)^\perp. \tag{D.10}$$

结合 (D.9)、(D.10) 和 (D.8)，可得

$$\nabla_X w^\top + \mathrm{II}(X, w^\top) - S_{w^\perp}(X) + \left(D_X w^\perp\right)^\perp = 0.$$

左侧为切向，而其余项均为法向，因此这两个分量都必须为零。特别地，

$$\nabla_X w^\top = S_{w^\perp}(X).$$

因此，

$$(\nabla_X P)w = \nabla_X (Pw) = \nabla_X w^\top = S_{w^\perp}(X), \tag{D.11}$$

这就证明了该断言。

由于 $\kappa = \sup_{\mathcal{M}} \|\mathrm{II}\|_{\mathrm{op}} < \infty$，由引理 C.2 可得

$$\|S_{w^\perp}\|_{\mathrm{op}} \le \kappa \left\| w^\perp \right\|.$$

因此，对于任意 $X \in T_x\mathcal{M}$，

$$\|S_{w^\perp}(X)\| \le \|S_{w^\perp}\|_{\mathrm{op}} \|X\| \le \kappa \left\| w^\perp \right\| \|X\| \le \kappa \|w\| \|X\|.$$

利用 (D.11)，我们得到

$$\|\nabla P(x)\|_{\mathrm{op}} = \sup_{X\neq 0,\, w\neq 0} \frac{\|(\nabla_X P)(x)w\|}{\|X\|\|w\|} = \sup_{X\neq 0,\, w\neq 0} \frac{\|S_{w^\perp}(X)\|}{\|X\|\|w\|} \le \kappa.$$

对 $x \in \mathcal{M}$ 取上确界，即完成证明。 ■

附录 E. 关于几何假设的更多讨论

假设 I 和假设 II 涉及 $\mathcal{M} \subset \mathbb{R}^n$的几何性质。本节首先给出三个满足这些假设的典型示例（E.1 节），然后讨论它们之间的关系（E.2 节）。此外，我们还给出假设 I 的一个外在判据，见 E.3 节。

E.1 示例

下面给出三个满足假设 I 和假设 II 的黎曼子流形示例 $\mathcal{M} \subset \mathbb{R}^n$ 。

1. 紧致情形：若黎曼子流形 $\mathcal{M} \subset \mathbb{R}^n$ 是紧致的，即闭且有界，那么由于典范嵌入 ι: $\mathcal{M} \to \mathbb{R}^n$ 是光滑的，因此可得

$$\sup_{\mathcal{M}} \|\nabla d\iota\|_{\mathrm{op}} < \infty, \quad \sup_{\mathcal{M}} \left\|\nabla^2 d\iota\right\|_{\mathrm{op}} < \infty.$$

因此， $\mathcal{M}$ 满足假设 I。此外，由于管状半径 $\varepsilon(x)$ 在 $\mathcal{M}$ 上光滑（定理 B.1）， $\mathcal{M}$的紧致性意味着

$$\inf_{x\in\mathcal{M}} \varepsilon(x) = \min_{x\in\mathcal{M}} \varepsilon(x) > 0.$$

因此， $\mathcal{M}$ 也满足假设 II。

2. 图：设 $f = (f^1, \ldots, f^{n-m})$: $\mathbb{R}^m \to \mathbb{R}^{n-m}$为光滑函数，且满足

$$\sup_{x\in\mathbb{R}^m} \|Df(x)\| < \infty, \quad \sup_{x\in\mathbb{R}^m} \|D^2 f(x)\|_{\mathrm{op}} < \infty, \quad \sup_{x\in\mathbb{R}^m} \|D^3 f(x)\|_{\mathrm{op}} < \infty. \tag{E.1}$$

考虑 f 的图流形，

$$\mathcal{M} = \{(x, f(x)) \colon x \in \mathbb{R}^m\} \subset \mathbb{R}^n.$$

设 ι: $\mathcal{M} \hookrightarrow \mathbb{R}^n$ 为典范嵌入，并在 $\mathcal{M}$ 上赋予诱导的黎曼度量 g 和列维-奇维塔联络 ∇。则 $\mathcal{M}$满足假设 I（定理 E.1）和假设 II（定理 E.2）。

3. 水平集：设 F: $\mathbb{R}^n \to \mathbb{R}^k$ 为光滑函数（$k < n$），且令 $\mathcal{M} = F^{-1}(0)$。假设对所有 $x \in \mathcal{M}$，秩 $DF(x) = k$。那么，根据常秩定理（Lee，2012，推论 5.14），$\mathcal{M}$ 是 m维的光滑子流形，且为 $\mathbb{R}^n$ 的子集，具有$m = n - k$。进一步假设

$$\inf_{x\in\mathcal{M}} \sigma_{\min}(DF(x)) > 0, \quad \sup_{x\in\mathcal{M}} \left\|D^2 F(x)\right\|_{\mathrm{op}} < \infty, \quad \sup_{x\in\mathcal{M}} \left\|D^3 F(x)\right\|_{\mathrm{op}} < \infty, \tag{E.2}$$

其中 $\sigma_{\min}$ 表示最小奇异值。令 ι: $\mathcal{M} \hookrightarrow \mathbb{R}^n$ 为典范嵌入，并在 $\mathcal{M}$ 上赋予诱导的黎曼结构和列维-奇维塔联络 ∇。

在 (E.2) 下， $\mathcal{M}$满足假设 I；参见定理 E.3。然而，要证明一致管状邻域（假设 II）的存在性，还需要附加条件。对于 $k = 1$，Platt 和 Gal´an（2025年）证明了：若 DF 的 L_1-范数在 $\mathcal{M}$ 上有界于某个正下界之上，且 $\|D^2 F\|$ 在 $\mathcal{M}$的一个开凸邻域 $\mathcal{V}$ 上有界（而不仅仅是在 $\mathcal{M}$ 自身上有界），则 $\mathcal{M}$ 存在一致管状邻域。对于一般的 k，我们将此类结果加以推广；参见定理 E.4。

定理 E.1 设 f: $\mathbb{R}^m \to \mathbb{R}^{n-m}$为一个光滑函数，且 $\mathcal{M} = \{(x, f(x)) \colon x \in \mathbb{R}^m\}$为其图流形。赋予 $\mathcal{M} \subset \mathbb{R}^n$ 由典范嵌入 ι: $\mathcal{M} \hookrightarrow \mathbb{R}^n$所诱导的黎曼结构。若 f 满足 ***(E.1)***，则

$$\sup_{x\in\mathcal{M}} \|\nabla d\iota(x)\|_{\mathrm{op}} < \infty, \quad \sup_{x\in\mathcal{M}} \left\|\nabla^2 d\iota(x)\right\|_{\mathrm{op}} < \infty.$$

证明 我们分以下四步证明该定理。

- 步骤 1。黎曼度量：为 $\mathcal{M}$ 配备典范坐标图 $x=(x^1,\ldots,x^m)$：$\mathcal{M}\to\mathbb{R}^m$，定义为

$$x^i(x,f(x))=x^i.$$

设 $\left\{\frac{\partial}{\partial x^i}\right\}_{i=1}^m$ 为 $T_x\mathcal{M}$的相关基。设 $y=(y^1,\ldots,y^n)$ 为 $\mathbb{R}^n$上的典范坐标，其坐标向量场为 $\left\{\frac{\partial}{\partial y^k}\right\}_{k=1}^n$。对于任意 $h\in C^\infty(\mathbb{R}^n)$，

$$\begin{aligned}d\iota\left(\frac{\partial}{\partial x^i}\right)(h)&=\frac{\partial}{\partial x^i}(h\circ\iota)=\frac{\partial}{\partial x^i}h(x^1,\ldots,x^m,f^1(x),\ldots,f^{n-m}(x))\\&=\frac{\partial}{\partial y^i}h+\sum_{k=1}^{n-m}\frac{\partial f^k}{\partial x^i}\frac{\partial}{\partial y^{k+m}}h,\end{aligned}$$

这意味着

$$d\iota\left(\frac{\partial}{\partial x^i}\right)=\frac{\partial}{\partial y^i}+\sum_{k=1}^{n-m}\frac{\partial f^k}{\partial x^i}\frac{\partial}{\partial y^{k+m}}. \tag{E.3}$$

设 g 为 $\mathcal{M}$上的诱导黎曼度量。则

$$g_{ij}=\left\langle d\iota\left(\frac{\partial}{\partial x^i}\right),d\iota\left(\frac{\partial}{\partial x^j}\right)\right\rangle=\delta_{ij}+\sum_{k=1}^{n-m}\frac{\partial f^k}{\partial x^i}\frac{\partial f^k}{\partial x^j}, \tag{E.4}$$

即，在此坐标系中，

$$g(x)=I_m+(Df(x))^T(Df(x)).$$

设 $C_1:=\sup_x\|Df(x)\|<\infty$。则

$$I_m\preceq g(x)\preceq\left(1+C_1^2\right)I_m, \tag{E.5}$$

因此

$$1\le\|g(x)\|_{\mathrm{op}}\le 1+C_1^2,\quad\left\|g(x)^{-1}\right\|_{\mathrm{op}}\le 1.$$

设 $g^{-1}=(g^{ij})$。由于 $\left\|g(x)^{-1}\right\|_{\mathrm{F}}\le\sqrt{m}\left\|g(x)^{-1}\right\|_{\mathrm{op}}$ (Horn and Johnson, 2012)，

$$\sup_x\left|g^{ij}(x)\right|\le\sqrt{m}. \tag{E.6}$$

此外，由 (E.4)，

$$\partial_\alpha g_{ij}=\sum_{k=1}^{n-m}\left(\partial_{\alpha i}f^k\partial_j f^k+\partial_i f^k\partial_{\alpha j}f^k\right),$$

为简便起见，我们记$\partial_\alpha=\frac{\partial}{\partial x^\alpha}$。令 $C_2:=\sup_x\|D^2f(x)\|_{\mathrm{op}}<\infty$。Then

$$\sup_x|\partial_\alpha g_{ij}(x)|\le 2C_1C_2. \tag{E.7}$$

此外，利用

$$\partial_{\alpha\beta} g_{ij} = \sum_{k=1}^{n-m} \left(\partial_{\alpha\beta i} f^k \partial_j f^k + \partial_{\alpha i} f^k \partial_{\beta j} f^k + \partial_{\beta i} f^k \partial_{\alpha j} f^k + \partial_i f^k \partial_{\alpha\beta j} f^k \right),$$

并令 $C_3 := \sup_x \|D^3 f(x)\|_{\mathrm{op}} < \infty$，同理可得

$$\sup_x |\partial_{\alpha\beta} g_{ij}(x)| \le 2C_1 C_3 + 2C_2^2. \tag{E.8}$$

最后，对于 $g^{-1} = (g^{ij})$，我们有

$$\partial_\ell g^{ij} = -\sum_{\alpha,\beta=1}^{m} g^{i\alpha} \left(\partial_\ell g_{\alpha\beta}\right) g^{\beta j}.$$

因此，由 (E.6) 和 (E.7) 可得，

$$\sup_x \left|\partial_\ell g^{ij}(x)\right| \le 2C_1 C_2 m^3. \tag{E.9}$$

- 步骤 2. 克里斯托费尔系数：根据定义，

$$\Gamma_{ij}^k = \frac{1}{2} \sum_{\alpha=1}^{m} g^{k\alpha} \left(\partial_i g_{j\alpha} + \partial_j g_{i\alpha} - \partial_\alpha g_{ij}\right).$$

利用 (E.6) 和 (E.7)，我们得到

$$\sup_x \left|\Gamma_{ij}^k(x)\right| \le 3C_1 C_2 m^{3/2}. \tag{E.10}$$

进一步求导，

$$\partial_\ell \Gamma_{ij}^k = \frac{1}{2} \sum_{\alpha=1}^{m} \left(\left(\partial_\ell g^{k\alpha}\right) \left(\partial_i g_{j\alpha} + \partial_j g_{i\alpha} - \partial_\alpha g_{ij}\right) + g^{k\alpha} \left(\partial_{\ell i} g_{j\alpha} + \partial_{\ell j} g_{i\alpha} - \partial_{\ell\alpha} g_{ij}\right) \right).$$

利用 (E.6) 和 (E.9) 对与 g^{-1} 相关的项进行有界估计，并利用 (E.7) 和 (E.8) 对与 g 相关的项进行有界估计，可得

$$\sup_x \left|\partial_\ell \Gamma_{ij}^k(x)\right| \le 6C_1^2 C_2^2 m^4 + \left(3C_1 C_3 + 3C_2^2\right) m^{3/2}. \tag{E.11}$$

- 步骤 3. $\nabla d\iota$ 的有界性：由 (C.1) 可得，

$$(\nabla d\iota)(\partial_i, \partial_j) = D_{d\iota(\partial_j)} \left(d\iota(\partial_i)\right) - d\iota\left(\nabla_{\partial_j} \partial_i\right), \tag{E.12}$$

其中记作 $\partial_i = \frac{\partial}{\partial x^i}\circ$ 为方便起见，记 $\tilde{\partial}_k = \frac{\partial}{\partial y^k}\circ$。于是，由 (E.3) 得，

$$D_{d\iota(\partial_j)} \left(d\iota(\partial_i)\right) = D_{\tilde{\partial}_j} \left(d\iota(\partial_i)\right) + \sum_{k=1}^{n-m} \partial_j f^k \, D_{\tilde{\partial}_{k+m}} \left(d\iota(\partial_i)\right) = \sum_{\ell=1}^{n-m} \partial_{ij} f^\ell \tilde{\partial}_{\ell+m}, \tag{E.13}$$

其中 $D_{\tilde{\partial}_j}\tilde{\partial}_k = 0$。这里，我们通过令 $y^i = x^i$（对于 $i = 1, \ldots, m$），并使其与 $y^{m+1}, \ldots, y^n$ 独立，从而将 f 自然延拓到 $\mathbb{R}^n$，因此，对于 $i, j \le m$，有 $\tilde{\partial}_{ij} f^k = \partial_{ij} f^k$；若 $i > m$ 或 $j > m$，则有 $\tilde{\partial}_{ij} f^k = 0$。
对于 (E.12) 中的第二项，注意到 $\nabla_{\partial_j}\partial_i = \sum_\alpha \Gamma_{ij}^\alpha \partial_\alpha$，因此

$$d\iota\left(\nabla_{\partial_j}\partial_i\right) = \sum_{\alpha=1}^{m} \Gamma_{ij}^\alpha \tilde{\partial}_\alpha + \sum_{\alpha=1}^{m}\sum_{\ell=1}^{n-m} \Gamma_{ij}^\alpha \partial_\alpha f^\ell \tilde{\partial}_{\ell+m}. \tag{E.14}$$

结合 (E.13)、(E.14) 和 (E.12) 可得

$$(\nabla d\iota)(\partial_i, \partial_j) = -\sum_{\alpha=1}^{m} \Gamma_{ij}^\alpha \tilde{\partial}_\alpha + \sum_{\ell=1}^{n-m} \left(\partial_{ij} f^\ell - \sum_{\alpha=1}^{m} \Gamma_{ij}^\alpha \partial_\alpha f^\ell\right) \tilde{\partial}_{\ell+m}. \tag{E.15}$$

由 (E.10) 及关于 f 的假设，存在一个 $C_\iota = C_\iota(C_1, C_2, m) > 0$，使得

$$\|(\nabla d\iota(x))(\partial_i, \partial_j)\| \le C_\iota < \infty. \tag{E.16}$$

由于根据 (E.5)，$g(x)$ 一致地有下界，引理 N.6 表明存在一个 $C_{\iota,1} = C_{\iota,1}(C_1, C_2, m) > 0$，使得

$$\|(\nabla d\iota(x))(u, v)\| \le C_{\iota,1}\|u\|\|v\|, \quad \forall\, u, v \in T_x\mathcal{M}.$$

因此，$\nabla d\iota(x)$ 是定义在 $T_x\mathcal{M} \times T_x\mathcal{M}$ 上的有界双线性映射，并且

$$\|\nabla d\iota(x)\|_{\mathrm{op}} = \sup_{u,v\neq 0} \frac{\|(\nabla d\iota(x))(u, v)\|}{\|u\|\|v\|} \le C_{\iota,1}, \quad \Rightarrow \quad \sup_{x\in\mathcal{M}} \|\nabla d\iota(x)\|_{\mathrm{op}} \le C_{\iota,1} < \infty. \tag{E.17}$$

- 步骤 4. $\nabla^2 d\iota$的有界性：设 $A = \nabla d\iota$。由 (C.2) 可得，

$$(\nabla A)(\partial_i, \partial_j, \partial_\alpha) = D_{d\iota(\partial_\alpha)}\left(A(\partial_i, \partial_j)\right) - A(\nabla_{\partial_\alpha}\partial_i, \partial_j) - A(\partial_i, \nabla_{\partial_\alpha}\partial_j). \tag{E.18}$$

利用 (E.15)，可写为

$$A_{ij} = A(\partial_i, \partial_j) = (\nabla d\iota)(\partial_i, \partial_j) = -\sum_{\beta=1}^{m} \Gamma_{ij}^\beta \tilde{\partial}_\beta + \sum_{\ell=1}^{n-m} \left(\partial_{ij} f^\ell - \sum_{\beta=1}^{m} \Gamma_{ij}^\beta \partial_\beta f^\ell\right) \tilde{\partial}_{\ell+m}.$$

对于 (E.18) 右侧的第一项，我们计算

$$\begin{aligned} D_{d\iota(\partial_\alpha)}\left(A(\partial_i, \partial_j)\right) = &-\sum_{\beta=1}^{m} \partial_\alpha \Gamma_{ij}^\beta \tilde{\partial}_\beta \\ &+ \sum_{\ell=1}^{n-m} \left(\partial_{ij\alpha} f^\ell - \sum_{\beta=1}^{m} \left(\partial_\alpha \Gamma_{ij}^\beta \partial_\beta f^\ell + \Gamma_{ij}^\beta \partial_{\alpha\beta} f^\ell\right)\right) \tilde{\partial}_{\ell+m}, \end{aligned} \tag{E.19}$$

其中，我们将 Γ_{ij}^k 延拓为 $\mathbb{R}^n$，使得 $\tilde{\partial}_\alpha \Gamma_{ij}^k = \partial_\alpha \Gamma_{ij}^k$ 对于 $\alpha = 1, \ldots\ldots, m$ 成立，且 $\tilde{\partial}_\alpha \Gamma_{ij}^k = 0$ 对于$\alpha = m+1, \ldots\ldots, n$成立。对于其余两项，

$$A(\nabla_{\partial_\alpha}\partial_i, \partial_j) = \sum_{\beta=1}^{m} \Gamma_{\alpha i}^{\beta} A_{\beta j}, \quad A(\partial_i, \nabla_{\partial_\alpha}\partial_j) = \sum_{\beta=1}^{m} \Gamma_{\alpha j}^{\beta} A_{i\beta}. \tag{E.20}$$

将 (E.19) 和 (E.20) 代入 (E.18)，并结合 (E.10)、(E.11)、(E.16) 以及 (E.1)，可得存在一个与 x 无关的 $C_\iota' = C_\iota'(C_1, C_2, C_3, m) > 0$，使得

$$\left\|(\nabla^2 d\iota(x))(\partial_i, \partial_j, \partial_\alpha)\right\| \le C_\iota' < \infty.$$

与步骤 3 中一样，一致下界 (E.5) 和引理 N.6 意味着存在$C_{\iota,\ 2} = C_{\iota,\ 2}(C_1, C_2, C_3, m) > 0$，使得

$$\left\|(\nabla^2 d\iota(x))(u, v, w)\right\| \le C_{\iota,2}\|u\|\|v\|\|w\|, \quad \forall\ u, v, w \in T_x\mathcal{M}.$$

因此，

$$\sup_{x\in\mathcal{M}} \left\|\nabla^2 d\iota(x)\right\|_{\mathrm{op}} = \sup_{x\in\mathcal{M}} \sup_{u,v,w\neq 0} \frac{\left\|(\nabla^2 d\iota(x))(u, v, w)\right\|}{\|u\|\|v\|\|w\|} \le C_{\iota,2}.$$

■

定理 E.2 设 $f\colon\ \mathbb{R}^m \to \mathbb{R}^{n-m}$ 是一个光滑函数，且 $\mathcal{M} = \{(x, f(x))\colon\ x \in \mathbb{R}^m\}$是其图像流形，并配备诱导的黎曼结构。若 f 满足

$$\sup_{x\in\mathbb{R}^m} \|Df(x)\| < \infty, \quad \sup_{x\in\mathbb{R}^m} \|D^2 f(x)\|_{\mathrm{op}} < \infty,$$

th那么，存在 $\mathbb{R}^n$ 中 $\mathcal{M}$ 的一致管状邻域 $\mathcal{V}_{r_0}$，其半径为 r_0。

证明 定义 $\Phi\colon\ \mathbb{R}^m \to \mathbb{R}^n$为 $\Phi(x) = (x,\ f(x))$。则 Im $\Phi = \mathcal{M}$，并且根据 Lee（2012, 命题 5.37），

$$T_{\Phi(x)}\mathcal{M} := \{(v, D\Phi(x)v)\colon\ v \in \mathbb{R}^m\} \subset \mathbb{R}^n.$$

因此，对于任意 $x \in \mathbb{R}^m$ 与 $\eta \in \mathbb{R}^k$ $(k = n - m)$，法空间的一种方便的参数化为

$$\xi(x, \eta) = \left(-Df(x)^T\eta, \eta\right) \in N_{\Phi(x)}\mathcal{M}.$$

特别地，

$$\|\eta\|^2 \le \|\xi(x, \eta)\|^2 = \|\eta\|^2 + \eta^T Df(x)Df(x)^T\eta \le (1 + C_1^2)\|\eta\|^2, \tag{E.21}$$

其中 $C_1 := \sup_{x\in\mathcal{M}}\|Df(x)\|_{\mathrm{op}} < \infty$。同样不难验证，

$$(\Phi, \xi)\colon\ \mathbb{R}^n \to N\mathcal{M}$$

由$(\Phi, \xi)(x, \eta) = (\Phi(x)\ \xi(x, \eta))$定义的映射是微分同胚；因此$(x, \eta)$在 $N\mathcal{M}$上给出全局坐标系。在这些坐标下，管状映射 $T\colon\ N\mathcal{M} \to \mathbb{R}^n$为

$$T(x, \eta) = \Phi(x) + \xi(x, \eta) = \left(x - Df(x)^T\eta, f(x) + \eta\right). \tag{E.22}$$

为构造一致管状邻域，我们证明 T 是局部微分同胚，然后在零截面的一个一致邻域上证明其单射性。

• 步骤 1. 局部微分同胚：根据反函数定理（Lee，2012，定理 4.5），只需证明对于所考虑的点，$DT(x,\ \eta)$ 可逆。由（E.22）可知，微分 $DT(x,\ \eta)$：$\mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^m \times \mathbb{R}^k$ 是

$$DT(x,\eta) = \begin{pmatrix} I_m - \sum_{\alpha=1}^k \eta^\alpha D^2 f^\alpha(x) & -Df(x)^T \\ Df(x) & I_k \end{pmatrix}.$$

根据舒尔补（Horn and Johnson，2012），$DT(x,\ \eta)$ 可逆当且仅当

$$S(x,\eta) = I_m + Df(x)^T Df(x) - \sum_{\alpha=1}^k \eta^\alpha D^2 f^\alpha(x)$$

可逆。对于任意 $v \in \mathbb{R}^m$,

$$\begin{aligned} v^T S(x,\eta) v &= \|v\|^2 + \|Df(x)v\|^2 - \sum_{\alpha=1}^k \eta^\alpha v^T D^2 f^\alpha(x) v \\ &\geq \|v\|^2 - \|\eta\| \left\| D^2 f(x)[v,v] \right\| \geq (1 - C_2\|\eta\|)\|v\|^2, \end{aligned}$$

其中 $C_2 := \sup_{x\in\mathcal{M}} \left\| D^2 f(x) \right\|_{\mathrm{op}} < \infty$。因此，若 $\|\eta\| < 1/C_2$，则 $S(x,\ \eta)$ 是正定的，因而可逆。由此可知，对于所有

$$(x,\eta) \in \mathcal{U}_{1/C_2} := \{(x,\eta) \colon \|\eta\| < 1/C_2\},$$

即，T 是 $\mathcal{U}_{1/C_2}$ 上的局部微分同胚。

• 步骤 2. 全局单射：令 $(u,\ v) \in \mathbb{R}^n$。存在 $(x,\ \eta) \in \mathbb{R}^n$，使得 $T(x,\ \eta) = (u,\ v)$ 当且仅当

$$u = x - Df(x)^T \eta, \quad v = f(x) + \eta,$$

等价地,

$$x - u + Df(x)^T (f(x) - v) = 0, \quad \eta = v - f(x).$$

固定 $(u, v) \in \operatorname{Im} T$ 并定义

$$\varphi_{u,v}(y) = \frac{1}{2}\left(\|y-u\|^2 + \|f(y) - v\|^2 \right).$$

由于 $D\varphi_{u,v}(y) = (y-u) + Df(y)^T (f(y) - v)$，我们有 $(x,\eta) \in T^{-1}(u,v)$ 当且仅当

$$D\varphi_{u,v}(x) = 0, \quad v = \eta + f(x).$$

假设 $T(x,\eta) = T(x',\eta') = (u,v)$，从而 x 与 x' 是 $\varphi_{u,v}$ 的临界点。我们证明这种情况不可能发生，只要

$$\|\eta\|,\ \left\|\eta'\right\| < r_0 := \frac{1}{C_2(1+2C_1^2)}.$$

首先，注意到

$$\left\|x - x'\right\| = \left\| Df(x)^T \eta - Df(x)^T \eta' \right\| < 2C_1 r_0.$$

设 $y_t = tx + (1-t)x'$，其中 $t \in [0,1]$。则

$$\begin{aligned}\|v - f(y_t)\| &\le \|v - f(x)\| + \|f(x) - f(y_t)\| \\ &\le \|\eta\| + C_1\|x - y_t\| < r_0 + C_1\|x - x'\| < (1 + 2C_1^2)r_0.\end{aligned} \tag{E.23}$$

此外，由于

$$D^2\varphi_{u,v}(y) = I_m + Df(y)^T Df(y) + \sum_{\alpha=1}^{k} \left(f^\alpha(y) - v^\alpha\right) D^2 f^\alpha(y),$$

步骤 1 中的相同估计表明，对任意 $w \in \mathbb{R}^m$,

$$w^T D^2\varphi_{u,v}(y_t) w \ge \left(1 - C_2\|f(y_t) - v\|\right)\|w\|^2 > \left(1 - C_2(1 + 2C_1^2)r_0\right)\|w\|^2,$$

其中我们使用了（E.23）。因此,

$$\begin{aligned}\left(D\varphi_{u,v}(x) - D\varphi_{u,v}(x')\right)[x - x'] &= \int_0^1 \left(x - x'\right)^T D^2\varphi_{u,v}(y_t)\left(x - x'\right)\,\mathrm{d}t \\ &> \left(1 - C_2(1 + 2C_1^2)r_0\right)\left\|x - x'\right\|^2 > 0,\end{aligned}$$

这与 $D\varphi_{u,\,v}(x) = D\varphi_{u,\,v}(x') = 0$ 矛盾。

结合步骤 1 和步骤 2，定义

$$\mathcal{U}_{r_0} = \{(x,\eta)\colon \|\eta\| < r_0\}, \quad r_0 = \frac{1}{C_2(1 + 2C_1^2)} < \frac{1}{C_2}.$$

则 T 是局部微分同胚，并且在 $\mathcal{U}_{r_0}$ 上是单射。因此 $T\colon \mathcal{U}_{r_0} \to \mathcal{V}_{r_0} = T(\mathcal{U}_{r_0}) \subset \mathbb{R}^n$ 是一个微分同胚，即 $\mathcal{V}_{r_0}$ 是 $\mathcal{M}$ 的半径为 r_0 的一致管状邻域。 ■

定理 E.3 设 $F\colon \mathbb{R}^n \to \mathbb{R}^k$ 是一个光滑函数（$k < n$），且 $\mathcal{M} = F^{-1}(0)$。假设对所有 $x \in \mathcal{M}$，均有 秩 $DF(x) = k$，并且 F满足（*E.2*）。赋予 $\mathcal{M}$ 由典范嵌入 $\iota\colon \mathcal{M} \hookrightarrow \mathbb{R}^n$所诱导的黎曼结构。于是

$$\sup_{x\in\mathcal{M}} \|\nabla d\iota(x)\|_{\mathrm{op}} < \infty, \quad \sup_{x\in\mathcal{M}} \left\|\nabla^2 d\iota(x)\right\|_{\mathrm{op}} < \infty.$$

证明 根据命题 C.3，只需证明

$$\sup_{x\in\mathcal{M}} \|\mathrm{II}_x\|_{\mathrm{op}} < \infty, \quad \sup_{x\in\mathcal{M}} \|\nabla \mathrm{II}_x\|_{\mathrm{op}} < \infty.$$

我们分以下三个步骤证明该定理。

- 步骤 1. 右逆：令 $G(x) := DF(x)$，其中 $x \in \mathbb{R}^n$。根据奇异值分解,

$$G(x) = R(x)\Sigma(x)Q(x), \quad \Sigma(x) = \begin{pmatrix} \sigma_1(x) & \cdots & 0 & \cdots & 0 \\ & \ddots & & & \\ 0 & \cdots & \sigma_k(x) & \cdots & 0 \end{pmatrix},$$

其中 $\sigma_1(x) \geq \cdots \geq \sigma_k(x) = \sigma_{\min}(G(x)) > 0$，因为 $G(x)$ 在 $\mathcal{M}$ 上满秩。考虑穆尔-彭罗斯右逆（Golan，2012）

$$G(x)^\dagger := G(x)^T \left(G(x)G(x)^T\right)^{-1} = Q^T \begin{pmatrix} \sigma_1^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_k^{-1} \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} R^T.$$

我们记录两个推论。

(a) 由假设，$c_0 := \inf_{x\in\mathcal{M}} \sigma_k(x) > 0$，因此

$$\sup_{x\in\mathcal{M}} \|G(x)^\dagger\|_{\text{op}} \leq \sup_{x\in\mathcal{M}} \sigma_k(x)^{-1} \leq c_0^{-1}. \tag{E.24}$$

(b) 由于 $\mathcal{M} = F^{-1}(0)$，Lee（2012，命题 5.38）蕴含 $T_x\mathcal{M} =$ 核 $G(x)$，从而 $N_x\mathcal{M} =$ (核 $G(x)$)$^\perp$。由引理 N.5，

$$P_x = G(x)^\dagger G(x)$$

是到 $N_x\mathcal{M}$ 上的正交投影。

• 步骤 2. II 的有界性：记 $F = (F^1, \ldots, F^k)$。对每个 $\alpha \in \{1, \ldots, k\}$，令 $f^\alpha = F^\alpha|_{\mathcal{M}} \in C^\infty(\mathcal{M})$。由于 $F|_{\mathcal{M}} = 0$，我们有 $f^\alpha \equiv 0$。由引理 N.7，对任意 $X, Y \in \Gamma(T\mathcal{M})$，

$$0 = \nabla^2 f^\alpha(X, Y) = D^2F^\alpha[X, Y] + DF^\alpha[\mathrm{II}(X, Y)].$$

因此，$DF^\alpha[\mathrm{II}(X, Y)] = -D^2F^\alpha[X, Y]$ 对 $\alpha = 1, \ldots, k$，从而对任意 $x \in \mathcal{M}$，

$$DF(x)[\mathrm{II}_x(u, v)] = -D^2F(x)[u, v], \quad \forall\, u, v \in T_x\mathcal{M}.$$

由于 $\mathrm{II}_x(u, v) \in N_x\mathcal{M}$，由步骤 1 得到的 $G(x)^\dagger$ 的投影性质可得

$$\mathrm{II}_x(u, v) = -G(x)^\dagger D^2F(x)[u, v].$$

利用 (E.24) 和 $C_2 := \sup_{x\in\mathcal{M}} \left\|D^2F(x)\right\|_{\text{op}} < \infty$，我们得到

$$\|\mathrm{II}_x\|_{\text{op}} \leq \left\|G(x)^\dagger\right\|_{\text{op}} \left\|D^2F(x)\right\|_{\text{op}} \leq \frac{C_2}{c_0},$$

因此

$$\sup_{x\in\mathcal{M}} \|\mathrm{II}_x\|_{\text{op}} \leq \frac{C_2}{c_0} < \infty. \tag{E.25}$$

• 步骤 3。∇II 的有界性：固定 $x \in \mathcal{M}$ 和 $u, v, w \in T_x\mathcal{M}$。选取 $U, V, W \in \Gamma(T\mathcal{M})$ 使得

$$U_x = u, \quad V_x = v, \quad W_x = w, \quad \nabla_W U|_x = \nabla_W V|_x = 0. \tag{E.26}$$

根据引理 N.8，$(\nabla \mathrm{II})(U,\ V,\ W)(x)$ 与 $U,\ V,\ W$ 的选择无关；我们将其记为 $(\nabla \mathrm{II}_x)(u,\ v,\ w)$。

令 $N := \mathrm{II}(U,\ V) \in \Gamma(N\mathcal{M})$。根据步骤 2，$DF[N] = -D^2F[U,V]$。沿 W 取方向导数并应用链式法则，得到

$$D^2F[N,W] + DF[D_W N] = -D^3F[U,V,W] - D^2F[D_W U, V] - D^2F[U, D_W V],$$

即，

$$DF[D_W N] = -D^3F[U,V,W] - D^2F[D_W U, V] - D^2F[U, D_W V] - D^2F[N,W]. \tag{E.27}$$

对于左侧，由于 $(D_W N|_x)^\top \in T_x\mathcal{M} = \text{核 } DF(x)$，我们有

$$DF(x)\left[D_W N|_x\right] = DF(x)\left[(D_W N|_x)^\perp\right]. \tag{E.28}$$

此外，根据 (E.26) 及定义 (C.3)，

$$(D_W N|_x)^\perp = (\nabla \mathrm{II}_x)(u,v,w). \tag{E.29}$$

在 (E.27) 的右侧，根据高斯公式 (Lee, 2018) 和 (E.26)，

$$\begin{aligned} D_W U|_x &= \nabla_W U|_x + \mathrm{II}_x(u,w) = \mathrm{II}_x(u,w), \\ D_W V|_x &= \nabla_W V|_x + \mathrm{II}_x(v,w) = \mathrm{II}_x(v,w). \end{aligned} \tag{E.30}$$

结合 (E.28)、(E.29)、(E.30) 和 (E.27) 可得

$$\begin{aligned} DF(x)[(\nabla \mathrm{II}_x)(u,v,w)] = &-D^3F(x)[u,v,w] - D^2F(x)[\mathrm{II}_x(u,w), v] \\ &- D^2F(x)[u, \mathrm{II}_x(v,w)] - D^2F(x)[\mathrm{II}_x(u,v), w]. \end{aligned}$$

由于 $(\nabla \mathrm{II}_x)(u,\ v,\ w) \in N_x\mathcal{M}$，由步骤 1 得到的 $G(x)^\dagger$ 的投影性质可得

$$\begin{aligned} (\nabla \mathrm{II}_x)(u,v,w) = -G(x)^\dagger \Big(&D^3F(x)[u,v,w] + D^2F(x)[\mathrm{II}_x(u,w), v] \\ &+ D^2F(x)[u, \mathrm{II}_x(v,w)] + D^2F(x)[\mathrm{II}_x(u,v), w] \Big). \end{aligned}$$

利用 (E.24)、(E.25) 以及 $C_2 = \sup_{\mathcal{M}} \left\|D^2F\right\|_{\mathrm{op}} < \infty$ 和 $C_3 := \sup_{\mathcal{M}} \left\|D^3F\right\|_{\mathrm{op}} < \infty$，我们得出

$$\sup_{x\in\mathcal{M}} \|\nabla \mathrm{II}_x\|_{\mathrm{op}} \le \frac{C_3}{c_0} + \frac{3C_2^2}{c_0^2} < \infty.$$

■

定理 E.4 设 $F:\ \mathbb{R}^n \to \mathbb{R}^k$ 是一个光滑函数 $(k < n)$，且 $\mathcal{M} = F^{-1}(0)$。假设对任意 $x \in \mathcal{M}$，秩 $DF(x) = k$，并且 F 满足

$$\inf_{x\in\mathcal{M}} \sigma_{\min}(DF(x)) > 0, \quad \sup_{y\in\mathcal{V}} \left\|D^2F(y)\right\|_{\mathrm{op}} < \infty,$$

其中，$\mathcal{V}$ 是 $\mathcal{M}$的开凸邻域。因此，$\mathcal{M}$ 具有一致管状邻域。

证明 由于根据隐函数定理， $\mathcal{M}$ 是闭的且无边界，根据定理 B.2，只需证明

$$r_0 := \sup\{\varepsilon > 0 \colon \forall\ y \text{ s.t. } d(y,\mathcal{M}) < \varepsilon,\ \exists\,!\ x \in \mathcal{M} \text{ s.t. } \|x-y\| = d(y,\mathcal{M})\} > 0. \quad \text{(E.31)}$$

固定 $y \in \mathbb{R}^n$。假设存在 $\mathcal{M}$ 中两个不同的点 $p \neq q$，使得

$$\|y-p\| = \|y-q\| = d(y,\mathcal{M}) =: \rho.$$

我们给出 ρ 的下界。

- 步骤 1. $y-p \in N_p\mathcal{M}$: 定义 $g(x) = \|y-x\|^2$: $\mathcal{M} \to \mathbb{R}$。对于任意 $u \in T_p\mathcal{M}$，选取一条光滑曲线 γ: $(-\varepsilon,\varepsilon) \to \mathcal{M}$，使得 $\gamma(0) = p$ 且 $\dot\gamma(0) = u$。由于 p 是 g 的极小值点， $t=0$ 是 $g(\gamma(t))$ 的临界点，因此

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} g(\gamma(t)) = 2\langle u, y-p\rangle.$$

  由于 u 是任意的，这意味着 $y-p \in N_p\mathcal{M}$。

- 步骤 2. 泰勒界：设 $v := q-p \neq 0$。由于 $p,q \in \mathcal{M} \subset \mathcal{V}$ 且 $\mathcal{V}$ 是凸的，线段

$$[p,q] := \{tp + (1-t)q \colon t \in [0,1]\}$$

  包含于 $\mathcal{V}$ 中。利用泰勒公式以及 $F(p) = F(q) = 0$ 这一事实，我们得到

$$-DF(p)v = F(q) - F(p) - DF(p)v = \int_0^1 (1-t)D^2F(p+tv)[v,v]\ \mathrm{d}t.$$

  令 $C_2 := \sup_{y\in\mathcal{V}} \|D^2F(y)\|_{\mathrm{op}} < \infty$。则

$$\|DF(p)v\| \le \int_0^1 (1-t)C_2\|v\|^2\ \mathrm{d}t = \frac{C_2}{2}\|v\|^2. \quad \text{(E.32)}$$

- 步骤 3. 对 ρ 的下界：令 $w := y-p \in N_p\mathcal{M}$。由于 $y-q = y-p-(q-p) = w-v$ 且 $\|y-q\| = \|y-p\| = \|w\| = \rho$，我们有

$$\|w\| = \|w-v\| \quad \Rightarrow \quad \frac{1}{2}\|v\|^2 = \langle w, v\rangle.$$

  将 $v = v_T + v_N$ 分解为 $v_T \in T_p\mathcal{M}$ 和 $v_N \in N_p\mathcal{M}$。由于 $w \in N_p\mathcal{M}$，我们得到 $\langle w, v\rangle = \langle w, v_N\rangle$，因此，根据柯西–施瓦茨不等式，

$$\frac{1}{2}\|v\|^2 = \langle w, v_N\rangle \le \|w\|\|v_N\| = \rho\|v_N\|. \quad \text{(E.33)}$$

  特别地， $v_N \neq 0$ 因为 $v \neq 0$。此外，由于 $v_T \in T_p\mathcal{M} = \text{核}\, DF(p)$，我们有 $DF(p)v = DF(p)v_N$。将此与(E.32)和(E.33)结合可得

$$\|DF(p)v_N\| = \|DF(p)v\| \le \frac{C_2}{2}\|v\|^2 \le C_2\rho\|v_N\|. \quad \text{(E.34)}$$

另一方面，如果 $c_0 := \inf_{x\in\mathcal{M}} \sigma_{\min}(DF(x)) > 0$，则

$$\|DF(p)v_N\| \geq c_0\|v_N\|. \tag{E.35}$$

结合 (E.34) 和 (E.35) 得到

$$c_0\|v_N\| \leq C_2\rho\|v_N\|.$$

由于 $v_N \neq 0$，我们得出

$$\rho \geq \frac{c_0}{C_2}.$$

因此，如果 $y \in \mathbb{R}^n$ 有多个点 $x \in \mathcal{M}$ 满足 $\|y - x\| = d(y, \mathcal{M})$，那么 $d(y, \mathcal{M}) \geq c_0/C_2$。等价地，如果 $d(y, \mathcal{M}) < c_0/C_2$，那么 y 在 $\mathcal{M}$ 中有唯一的最近点。根据可达半径的定义 (E.31)，这意味着

$$r_0 \geq c_0/C_2 > 0.$$

■

E.2 假设 I 与假设 II 之间的关系

Assumption I to Assumption II. 根据命题 C.3，假设 I 等价于第二基本形式及其协变导数的一致有界性:

$$\sup_{x\in\mathcal{M}} \|\mathrm{II}_x\|_{\mathrm{op}} < \infty, \quad \sup_{x\in\mathcal{M}} \|\nabla\mathrm{II}_x\|_{\mathrm{op}} < \infty.$$

这些界刻画了 $\mathcal{M}$ 的局部几何正则性。特别是，第二基本形式有界意味着管状映射 $T(x, \xi) = x + \xi$ 在 $N\mathcal{M}$ 中零截面的某个邻域上是局部微分同胚，正如命题 E.5 所述。

然而，一般而言，仅控制第二基本形式并不能确保 T 的全局单射性；参见例 E.1。此外，若嵌入映射 $\iota\colon \mathcal{M} \hookrightarrow \mathbb{R}^n$ 是全局双 Lipschitz 连续的，则第二基本形式有界确实意味着存在一致管状邻域；参见附录 F.2。

Assumption II to Assumption I. 由于存在一致管状邻域可保证管状映射 T 是全局微分同胚，因此不难看出，若 $\mathcal{M}$ 存在一致管状邻域，则第二基本形式一致有界；参见命题 E.6。

然而，一般而言，存在一致管状邻域并不能控制第二基本形式的协变导数。例如，考虑由光滑映射 $f\colon \mathbb{R}^m \to \mathbb{R}^{n-m}$ 关联的图流形 $\mathcal{M} = \{(x,\ f(x)) :\ x \in \mathbb{R}^m\} \subset \mathbb{R}^n$。根据定理 E.2，若 $C_1 = \|Df\|_\infty < \infty$ 且 $C_2 = \left\|D^2 f\right\|_\infty < \infty$，则 $\mathcal{M}$ 存在半径为

$$r_0 := \frac{1}{C_2(1 + 2C_1^2)} < \frac{1}{C_2}.$$

相比之下，界定 ∇II 需要额外的正则性。如定理 E.1 所示，对 ∇II 的控制（根据命题 C.3，等价于对 $\nabla^2 d_\iota$ 的控制）涉及 f 的三阶导数。

事实上，(E.19) 与 (E.18) 中 $\nabla^2 d_\iota$ 的表示包含依赖于 $D^3 f$ 的项。因此，即使 ∇II 无界，一致管状邻域仍可能存在，因为定理 E.2 并不要求 $D^3 f$ 的有界性。我们给出例 E.2 来说明这种情况。

命题 E.5 设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个黎曼子流形。若

$$\kappa = \sup_{x \in \mathcal{M}} \|\mathrm{II}_x\|_{\mathrm{op}} < \infty,$$

那么，在开集

$$\mathcal{U}_{1/\kappa} := \{(x, \xi) \in N\mathcal{M} \colon \|\xi\| < 1/\kappa\} \subset N\mathcal{M},$$

由 $T(x, \xi) = x + \xi$ 给出的管状映射 $T \colon \mathcal{U}_{1/\kappa} \to \mathbb{R}^n$ 是一个局部微分同胚。

证明 我们首先将 $T_{(x,\xi)}(N\mathcal{M})$ 识别为 $T_x\mathcal{M} \oplus N_x\mathcal{M}$。在局部范围内，选取 $T\mathcal{M}$ 在 x的某个邻域内的标准正交标架 $\{E_i(x)\}_{i=1}^m$。那么，$(x, \xi) \in N\mathcal{M}$ 当且仅当

$$g_i(x, \xi) := \langle \xi, E_i(x) \rangle = 0, \quad \forall\, i = 1, \ldots, m.$$

因此，对于 $(v, \zeta) \in T_x\mathcal{M} \times \mathbb{R}^n$，有 $(v, \zeta) \in T_{(x,\xi)}(N\mathcal{M})$ 当且仅当 $dg_i(x, \xi)[v, \zeta] = 0$，其中 $i = 1, \ldots, m$。计算得，

$$\begin{aligned} 0 = dg_i(x, \xi)[v, \zeta] &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g_i(x + tv, \xi + t\zeta) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \langle \xi + t\zeta, E_i(x + tv) \rangle \\ &= \langle \zeta, E_i(x) \rangle + \langle \xi, D_v E_i(x) \rangle = \langle \zeta, E_i(x) \rangle + \langle \xi, \nabla_v E_i(x) + \mathrm{II}_x(E_i(x), v) \rangle \\ &= \langle \zeta, E_i(x) \rangle + \langle S_{x,\xi} v, E_i(x) \rangle , \end{aligned}$$

这里，$S_{x,\xi}$ 是 x处沿 ξ 的形算子。因此 $\zeta^\top = -S_{x,\xi} v$，从而

$$\zeta = -S_{x,\xi} v + w$$

对于某个 $w \in N_x\mathcal{M}$。由此可得

$$T_{(x,\xi)}(N\mathcal{M}) = \{(v, -S_{x,\xi} v + w) \colon v \in T_x\mathcal{M},\ w \in N_x\mathcal{M}\},$$

因此，该映射 $(v, w) \mapsto (v, -S_{x,\xi} v + w)$ 给出一个同构 $T_{(x,\xi)}(N\mathcal{M}) \cong T_x\mathcal{M} \oplus N_x\mathcal{M}$。

在此同一识别下，我们推导出一个显式表达式：

$$dT(x, \xi) \colon T_{(x,\xi)}(N\mathcal{M}) \cong T_x\mathcal{M} \oplus N_x\mathcal{M} \to \mathbb{R}^n.$$

设 $(x(t)\ \xi(t))$ 是 $N\mathcal{M}$ 中的一条光滑曲线，满足 $(x(0)\ \xi(0)) = (x, \xi)$，且 $(\dot{x}(0)\ \dot{\xi}(0)) = (v, \zeta)$，其中 $\zeta = -S_{x,\xi} v + w$。则

$$dT(x, \xi)(v, w) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (x(t) + \xi(t)) = (\mathrm{id} - S_{x,\xi}) v + w.$$

由于 $\dim T_{(x,\xi)}(N\mathcal{M}) = n$，映射 $dT(x,\ \xi)$ 可逆，当且仅当 $\ker dT(x,\ \xi) = \{0\}$。此外，

$$dT(x, \xi)(v, w) = (\mathrm{id} - S_{x,\xi}) v + w = 0$$

成立当且仅当 $w = 0$ and (id$-S_{x,\ \xi}$)$v = 0$，因为 $w \in N_x\mathcal{M}$ ，而 (id$-S_{x,\ \xi}$)$v \in T_x\mathcal{M}$。因此，ker $dT(x,\ \xi) = \{0\}$，当且仅当 id $-S_{x,\ \xi}$ 可逆。

由引理 C.2，界 $\kappa =$sup$_{x\in\mathcal{M}}\|\mathrm{II}_x\|_{\mathrm{op}} < \infty$ 蕴含

$$\|S_{x,\xi}\|_{\mathrm{op}} \le \kappa\|\xi\|.$$

因此，对任意

$$(x,\xi) \in \mathcal{U}_{r_0} := \{(x,\xi) \in N\mathcal{M} \colon \|\xi\| < r_0\}, \quad r_0 := \frac{1}{\kappa},$$

我们有 $\|S_{x,\xi}\|_{\mathrm{op}} < 1$。特别地，id$-S_{x,\xi}$ 可由诺伊曼级数（Horn and Johnson，2012）证明可逆。因此，对所有 $(x,\xi) \in \mathcal{U}_{r_0}$， $dT(x,\xi)$ 均可逆，从而$T\colon\ \mathcal{U}_{r_0} \to \mathbb{R}^n$ 是局部微分同胚。

■

例 E.1 设 $\mathcal{M} = \{(x, e^{-x})\colon x \in \mathbb{R}\} \cup \{(x, -e^{-x})\colon x \in \mathbb{R}\} \subset \mathbb{R}^2$。对于平面曲线，第二基本形式本质上就是曲率，即

$$\mathrm{II}(\vec{T},\vec{T}) = \kappa\vec{n},$$

其中 $\vec{T}$ 是单位切向量， $\vec{n}$ 是单位法向量的一种选取， κ 是曲率；详见 *do Carmo*（*2016*）。因此，

$$\left\|\mathrm{II}_{(x,\pm e^{-x})}\right\|_{\mathrm{op}} = |\kappa(x)| = \frac{e^{-x}}{(1+e^{-2x})^{3/2}} \quad \Rightarrow \quad \sup\|\mathrm{II}\|_{\mathrm{op}} < \infty,$$

and

$$\left\|\nabla\mathrm{II}_{(x,\pm e^{-x})}\right\|_{\mathrm{op}} = \left|\frac{\mathrm{d}\kappa}{\mathrm{d}s}\right| = \frac{e^{-x}\left|1-2e^{-2x}\right|}{(1+e^{-2x})^3} \quad \Rightarrow \quad \sup\|\nabla\mathrm{II}\|_{\mathrm{op}} < \infty.$$

然而，对于 $p_x = (x,0) \in \mathbb{R}^2$，在 $\mathcal{M}$ 中存在两个不同的最近点，即 $q_{x,\ 1} = (x, e^{-x})$ 和 $q_{x,\ 2} = (x, -e^{-x})$。此外，

$$d(p_x,\mathcal{M}) = \|p_x - q_{x,1}\| = \|p_x - q_{x,2}\| = e^{-x} \to 0,$$

这意味着 可达半径$(\mathcal{M}) = 0$ （见附录 *B.2*）。因此，根据定理 *B.2*， $\mathcal{M}$ 不具有一致管状邻域。

命题 E.6 设 $\mathcal{M} \subset \mathbb{R}^n$ 为一个黎曼子流形。若 $\mathcal{M}$ 具有半径为 r_0的一致管状邻域，则

$$\sup_{x\in\mathcal{M}}\|\mathrm{II}_x\| \le \frac{1}{r_0} < \infty.$$

证明 设

$$\mathcal{U}_{r_0} := \{(x,\xi) \in N\mathcal{M} \colon \|\xi\| < r_0\}.$$

由于 $T\colon\ \mathcal{U}_{r_0} \to \mathcal{V}_{r_0} = T(\mathcal{U}_{r_0})$ 是微分同胚，因此对每个 $(x,\ \xi) \in \mathcal{U}_{r_0}$，微分 $dT_{(x,\xi)}$ 都是可逆的。固定 $x \in \mathcal{M}$ 和 $\eta \in N_x\mathcal{M}$ ，使得 $\|\eta\| = 1$。那么 $dT_{(x,t\eta)}$ 是可逆的

对所有 $t \in (-r_0,\ r_0)$ 都成立。根据命题 E.5 证明中的论证，这等价于以下映射可逆

$$\mathrm{id} - S_{x,t\eta} = \mathrm{id} - tS_{x,\eta}, \quad \forall\ t \in (-r_0, r_0)$$

由于 $S_{x,\ \eta}$ 是对称的，若 λ 是 $S_{x,\ \eta}$的一个特征值，则 $1-t\lambda$ 是 $\mathrm{id}-tS_{x,\ \eta}$的一个特征值。因此，对所有 $t \in (-r_0,\ r_0)$，均有 $1 - t\lambda \neq 0$，这意味着

$$|\lambda| \leq \frac{1}{r_0},$$

因为否则，取 $t = 1/\lambda \in (-r_0,\ r_0)$ 就会得到 $1 - t\lambda = 0$。因此，

$$\|S_{x,\eta}\| \leq \frac{1}{r_0}.$$

结论由引理 C.2 得出：

$$\sup_{x\in\mathcal{M}} \|\mathrm{II}_x\| \leq \frac{1}{r_0} < \infty.$$

■

例 E.2 定义 $f\colon\ \ \mathbb{R} \to \mathbb{R}$ 如下

$$f(x) := \int_0^x I(t)\ \mathrm{d}t, \quad I(x) := \int_0^x \sin(u^2)\ \mathrm{d}u.$$

并考虑图流形 $\mathcal{M} = \{(x, f(x))\colon\ \ x \in \mathbb{R}\} \subset \mathbb{R}^2$。那么

$$f'(x) = I(x), \quad f''(x) = \sin(x^2), \quad f'''(x) = 2x\cos(x^2).$$

特别地，

$$\left\|f''\right\|_\infty < \infty, \quad \left\|f'''\right\|_\infty = \infty.$$

接下来验证 f'有界。由于 $f'(x) = I(x)$，显然有 $|I(x)| \leq I(1) < \infty$ 对 $|x| \leq 1$成立。对于 $|x| > 1$，通过变量替换 $t = u^2$ 可得

$$I(x) = \int_0^1 \sin(u^2)\ \mathrm{d}u + \int_1^{x^2} \frac{\sin t}{t^{1/2}}\ \mathrm{d}t.$$

T第一项由1有界。对于第二项，分部积分可得

$$\begin{aligned}
\left|\int_1^{x^2} \frac{\sin t}{t^{1/2}}\ \mathrm{d}t\right| &= \left|\left(-t^{-1/2}\cos t\right)_1^{x^2} - \frac{1}{2}\int_1^{x^2} t^{-3/2}\cos t\ \mathrm{d}t\right| \\
&\leq \left|x^{-1}\cos x^2\right| + |\cos 1| + \frac{1}{2}\int_1^{x^2} \left|t^{-3/2}\cos t\right|\ \mathrm{d}t \\
&\leq 2 + \frac{1}{2}\int_1^{\infty} t^{-3/2}\ \mathrm{d}t = 3.
\end{aligned}$$

因此，对于所有 x，都有 $|I(x)| \leq I(1) + 3$，从而

$$\left\|f'\right\|_\infty < \infty.$$

根据定理 *E.2*，$\mathcal{M}$ 具有一致管状邻域。然而，由于 $\|f'''\|_\infty = \infty$，根据定理 *E.1*，∇ II 无界。或者，也可以像例 *E.1* 中那样直接验证$\mathrm{d}\kappa/\mathrm{d}s$ 无界。

E.3 假设 I 的外在判据

假设 I 要求内在导数具有一致界，这在实践中可能难以验证。因此，我们给出一个完全以外在导数表述的充分判据，该判据还涉及正交投影。

命题 E.7 设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个带有典范嵌入 ι 的黎曼子流形。设 $\mathcal{V}$ 是 $\mathbb{R}^n$ 中 $\mathcal{M}$ 的一个开邻域，并设 $P\colon \mathcal{V} \to \mathbb{R}^n$ 是一个光滑映射，使得对于任意 $x \in \mathcal{M}$，$P(x)$ 是到 $T_x\mathcal{M}$ 上的正交投影。如果

$$\sup_{x\in\mathcal{M}} \|DP(x)\|_{\mathrm{op}} < \infty, \quad \sup_{x\in\mathcal{M}} \left\|D^2P(x)\right\|_{\mathrm{op}} < \infty,$$

那么 ι 满足假设 I。

证明 首先，根据命题 C.3，只需证明

$$\sup_{x\in\mathcal{M}} \|\mathrm{II}_x\|_{\mathrm{op}} < \infty, \quad \sup_{x\in\mathcal{M}} \|\nabla\mathrm{II}_x\|_{\mathrm{op}} < \infty.$$

对于 $X, Y \in \Gamma(T\mathcal{M})$，根据高斯公式（Lee, 2018），

$$D_XY = \nabla_XY + \mathrm{II}(X,Y), \quad \nabla_XY \in \Gamma(T\mathcal{M}),\ \mathrm{II}(X,Y) \in \Gamma(N\mathcal{M}),$$

因此

$$\mathrm{II}(X,Y) = (\mathrm{id} - P)(D_XY).$$

此外，由于 $Y \in \Gamma(T\mathcal{M})$，我们有 $PY = Y$，因此

$$D_XY = D_X(PY) = (D_XP)Y + P(D_XY).$$

由此可得

$$\mathrm{II}(X,Y) = (\mathrm{id} - P)(D_XP)Y. \tag{E.36}$$

因此，对于任意 $x \in \mathcal{M}$ 以及任意 $u,\ v \in T_x\mathcal{M}$，

$$\mathrm{II}_x(u,v) = (\mathrm{id} - P(x))D_uP(x)v,$$

这意味着

$$\sup_{x\in\mathcal{M}} \|\mathrm{II}_x\|_{\mathrm{op}} \le \sup_{x\in\mathcal{M}} \|DP(x)\|_{\mathrm{op}} < \infty. \tag{E.37}$$

接下来，我们对 $\nabla\mathrm{II}$ 进行估计。固定 $x \in \mathcal{M}$ 和 $u,\ v,\ w \in T_x\mathcal{M}$。选取 $U,\ V,\ W \in \Gamma(T\mathcal{M})$ 使得

$$U_x = u,\ V_x = v,\ W_x = w, \quad \nabla_WU|_x = \nabla_WV|_x = 0.$$

由引理 N.8 可知，$(\nabla\mathrm{II})(U,\ V,\ W)(x)$ 与 $U,\ V,\ W$ 的选取无关，我们将其记为 $(\nabla\mathrm{II}_x)(u,\ v,\ w)$。然后，根据定义 (C.3)，

$$(\nabla\mathrm{II}_x)(u,v,w) = (D_w\mathrm{II}(U,V)|_x)^{\perp} = (\mathrm{id} - P(x))D_w\mathrm{II}(U,V)|_x. \tag{E.38}$$

利用 (E.36)，我们计算得

$$D_w\mathrm{II}(U,V)|_x = D_w(\mathrm{id} - P)|_x\,(D_uP(x)v) + (\mathrm{id} - P(x))D_w\left((D_UP)V\right)|_x.$$

代入 (E.38) 得到

$$(\nabla \mathrm{II}_x)(u,v,\ w) = (\mathrm{id} - P(x))D_w(\mathrm{id} - P)|_x\left(D_u P(x)v\right) + (\mathrm{id} - P(x))D_w\left((D_U P)V\right)|_x. \quad \text{(E.39)}$$

注意到 $P^2 = P$，因此

$$D_\xi(P^2) = (D_\xi P)P + P(D_\xi P) = D_\xi P, \quad \forall\ \xi \in \mathbb{R}^n.$$

分别在左侧和右侧乘以 P 和 $I - P$，得到

$$P(D_\xi P)P = 0, \quad (I - P)(D_\xi P)(I - P) = 0. \quad \text{(E.40)}$$

(E.40) 中的第一个恒等式意味着 $D_u P(x)v \in N_x\mathcal{M}$，因为对于 $v \in T_x\mathcal{M}$，$P(x)v = v$。因此，(E.40) 中的第二个恒等式给出

$$(\mathrm{id} - P(x))D_w(\mathrm{id} - P)|_x\left(D_u P(x)v\right) = -(\mathrm{id} - P(x))(D_w P|_x)(\mathrm{id} - P(x))\left(D_u P(x)v\right) = 0.$$

将其与 (E.39) 结合，得到

$$(\nabla \mathrm{II}_x)(u,v,w) = (\mathrm{id} - P(x))D_w\left((D_U P)V\right)|_x. \quad \text{(E.41)}$$

对于右端，链式法则给出

$$\begin{aligned} D_w\left((D_U P)V\right)|_x &= D_w(D_U P)|_x v + D_u P(x)(D_w V)|_x \\ &= \left(D^2 P(x)[u,w] + D_{D_w U|_x} P(x)\right) v + D_u P(x)(D_w V)|_x. \end{aligned} \quad \text{(E.42)}$$

此外，由于 $\nabla_W U|_x = \nabla_W V|_x = 0$，我们有

$$\begin{aligned} D_w U|_x &= \nabla_W U|_x + \mathrm{II}_x(u,w) = \mathrm{II}_x(u,w), \\ D_w V|_x &= \nabla_W V|_x + \mathrm{II}_x(v,w) = \mathrm{II}_x(v,w). \end{aligned} \quad \text{(E.43)}$$

结合 (E.42)、(E.43) 和 (E.41)，我们得到

$$\begin{aligned} (\nabla \mathrm{II}_x)(u,v,w) &= (\mathrm{id} - P(x))\left(D^2 P(x)[u,w]v + D_{\mathrm{II}_x(u,w)} P(x)v + D_u P(x)\mathrm{II}_x(v,w)\right) \\ &= (\mathrm{id} - P(x))\left(D^2 P(x)[u,w]v + D_{\mathrm{II}_x(u,w)} P(x)v\right), \end{aligned} \quad \text{(E.44)}$$

其中第二个等式由

$$(\mathrm{id} - P(x))D_u P(x)\mathrm{II}_x(v,w) = (\mathrm{id} - P(x))D_u P(x)(\mathrm{id} - P(x))\mathrm{II}_x(v,w) = 0,$$

根据 (E.40) 以及 $\mathrm{II}_x(v,\ w) \in N_x\mathcal{M}$ 这一事实。因此，(E.44) 蕴含

$$\|\nabla \mathrm{II}_x\|_{\mathrm{op}} \le \left\|D^2 P(x)\right\|_{\mathrm{op}} + \|DP(x)\|_{\mathrm{op}}\|\mathrm{II}_x\|_{\mathrm{op}}.$$

结合（E.37）中对 II 的界，可得

$$\sup_{x\in\mathcal{M}} \|\nabla \mathrm{II}_x\|_{\mathrm{op}} < \infty.$$

■

附录 F. 关于嵌入映射的更多讨论

为了根据定理 5，通过 $d_{\mathcal{M}}$ 上的 $\mathcal{M}$ 内蕴距离来推导强收敛，我们要求 $\mathcal{M}$ 良好地嵌入 $\mathbb{R}^n$ 中；参见第 F.1 节。此外，在假设 I 下，这一嵌入性质使假设 II 成为非必要条件；参见第 F.2 节。

F.1 双 Lipschitz 连续性

设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个黎曼子流形，其典范嵌入为 $\iota\colon \mathcal{M} \hookrightarrow \mathbb{R}^n$。对于任意 $x,\ y \in \mathcal{M}$，设 $\gamma\colon [0,1] \to \mathcal{M}$ 是一条分段光滑曲线，满足 $\gamma(0) = x$ 且 $\gamma(1) = y$。那么

$$L(\gamma) \geq \|\iota(x) - \iota(y)\|,$$

其中，$L(\gamma)$ 表示 γ 的长度。对所有此类曲线取下确界可得

$$\|x - y\| = \|\iota(x) - \iota(y)\| \leq \inf_{\gamma} L(\gamma) = d_{\mathcal{M}}(x, y), \tag{F.1}$$

其中 $d_{\mathcal{M}}$ 是 $\mathcal{M}$ 上诱导的黎曼距离。因此，ι 自动是 1-Lipschitz 的。但一般而言，(F.1) 中的反向不等式未必成立。

定义 F.1（双 Lipschitz）设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个黎曼子流形，其嵌入映射为 ι。如果

$$L_\iota := \sup_{x \neq y \in \mathcal{M}} \frac{d_{\mathcal{M}}(x, y)}{\|x - y\|} < \infty,$$

则称 ι 是全局双 *Lipschitz* 连续的，或称为 L_ι-双 *Lipschitz* 连续的。

注释 F.1 双 Lipschitz 这一术语反映了这样一个事实：在这种情况下，逆映射 $\iota^{-1}\colon \iota(\mathcal{M}) \to \mathcal{M}$ 也是 *Lipschitz* 的。等价地，

$$d_{\mathcal{M}}(x, y) \leq L_\iota \|x - y\|, \quad \forall\ x, y \in \mathcal{M}.$$

关于嵌入 $\iota\colon \mathcal{M} \hookrightarrow \mathbb{R}^n$ 的双Lipschitz性质，我们记录以下两个事实：

(i) 若 $\mathcal{M} \subset \mathbb{R}^n$ 是紧致的，则 ι 是全局双 Lipschitz 连续的；见定理 F.1 .

(ii) 存在一些非紧子流形 $\mathcal{M} \subset \mathbb{R}^n$，使得 ι 是全局双 Lipschitz 连续的；见例 F.1。

例 F.1 设 $f\colon \mathbb{R}^m \to \mathbb{R}^{n-m}$ 是一个光滑映射，使得

$$C_1 := \sup_{x \in \mathbb{R}^m} \|Df(x)\|_{\mathrm{op}} < \infty.$$

考虑图流形 $\mathcal{M} = \{(x, f(x))\colon\ x \in \mathbb{R}^m\}$，其配备诱导的黎曼度量。对于任意 $p = (x, f(x))$ 和 $q = (y, f(y))$，属于 $\mathcal{M}$，定义

$$\gamma(t) = (x + t(y - x),\ f(x + t(y - x))) \in \mathcal{M}, \quad t \in [0, 1].$$

则 $\gamma\colon\ [0,1]\to\mathcal{M}$ 满足 $\gamma(0)=p$ 和 $\gamma(1)=q$。此外，

$$\dot\gamma(t)=(y-x,\,Df(x+t(y-x))(y-x))\,,$$

从而

$$\|\dot\gamma(t)\|\le\sqrt{\|y-x\|^2+\|Df(x+t(y-x))\|_{\mathrm{op}}^2\|y-x\|^2}\le\sqrt{1+C_1^2}\,\|x-y\|.$$

因此，

$$d_{\mathcal{M}}(p,q)\le L(\gamma)=\int_0^1\|\dot\gamma(t)\|\,\mathrm{d}t\le\sqrt{1+C_1^2}\,\|x-y\|.$$

定理 F.1 设 $\mathcal{M}\subset\mathbb{R}^n$ 是一个 m 维、连通且紧致的黎曼流形，其典范嵌入为 $\iota\colon\mathcal{M}\hookrightarrow\mathbb{R}^n$。则 ι 是全局双 *Lipschitz* 连续的，即存在 $L_\iota>0$ 使得

$$d_{\mathcal{M}}(x,y)\le L_\iota\|x-y\|,\quad\forall\,x,y\in\mathcal{M}.$$

证明 由于 $\mathcal{M}$ 是紧致且连通的，

$$\operatorname{diam}(\mathcal{M}):=\sup_{x,y\in\mathcal{M}}d_{\mathcal{M}}(x,y)<\infty.$$

证明由以下两个步骤构成。

- 步骤 1. 局部双 Lipschitz 连续性：固定 $p\in\mathcal{M}$。由于 $\mathcal{M}\subset\mathbb{R}^n$ 是嵌入子流形，在对 $\mathbb{R}^n$ 作刚性运动之后，我们可以假设 $p=0$ 且

  $$T_p\mathcal{M}=\mathbb{R}^m\times\{0\}\subset\mathbb{R}^m\times\mathbb{R}^{n-m}.$$

  由隐函数定理，存在 $r>0$、p 的一个开邻域 $\mathcal{V}$，以及一个光滑映射 $f\colon B_r(0)\subset\mathbb{R}^m\to\mathbb{R}^{n-m}$，使得

  $$\mathcal{M}\cap\mathcal{V}=\{(x,f(x))\colon x\in B_r(0)\}\,,$$

  也就是说，$\mathcal{M}\cap\mathcal{V}$ 是 f 的图像。通过适当缩小 r，我们还可以假设

  $$K_p:=\sup_{x\in B_r(0)}\|Df(x)\|_{\mathrm{op}}<\infty.$$

  那么例 F.1 表明，对所有 $x,\ y\in\mathcal{M}\cap\mathcal{V}$，

  $$d_{\mathcal{M}}(x,y)\le\sqrt{1+K_p^2}\,\|x-y\|.$$

- 步骤 2. 全局有界性：反设

  $$L_\iota:=\sup_{x\neq y\in\mathcal{M}}\frac{d_{\mathcal{M}}(x,y)}{\|x-y\|}=\infty.$$

那么存在序列 $x_k \neq y_k$ 属于 $\mathcal{M}$，使得

$$\frac{d_{\mathcal{M}}(x_k, y_k)}{\|x_k - y_k\|} > k.$$

由于 $d_{\mathcal{M}}(x_k,\ y_k) \leq \mathrm{diam}(\mathcal{M})$，因此

$$\|x_k - y_k\| < \frac{d_{\mathcal{M}}(x_k, y_k)}{k} \leq \frac{\mathrm{diam}(\mathcal{M})}{k} \to 0.$$

由 $\mathcal{M}$ 的紧致性（从而也具有闭性），我们可以取一个子序列（不重新标记），使得

$$x_k \to p, \quad y_k \to p$$

对于某个 $p \in \mathcal{M}$。由步骤 1，存在 p 的一个邻域 $\mathcal{V}$ 以及一个常数 $C_p < \infty$，使得

$$d_{\mathcal{M}}(x, y) \leq C_p \|x - y\|, \quad \forall\ x, y \in \mathcal{M} \cap \mathcal{V}.$$

因此，当 k 充分大时，我们有 $x_k,\ y_k \in \mathcal{M} \cap \mathcal{V}$，从而

$$k < \frac{d_{\mathcal{M}}(x_k, y_k)}{\|x_k - y_k\|} \leq C_p,$$

这与此矛盾。这证明了 $L_\iota < \infty$。

■

F.2 由假设 I 推导假设 II

如附录 E.2 所述，假设 I 通常并不蕴含假设 II：存在一致管状邻域是 $\mathcal{M}$ 的全局性质，而假设 I 仅要求满足局部几何正则性。然而，当典范嵌入 $\iota\colon \mathcal{M} \hookrightarrow \mathbb{R}^n$ 全局双 Lipschitz 连续时，假设 I 更准确地说，即一致有界条件

$$\sup_{x \in \mathcal{M}} \|\nabla d_\iota(x)\| = \sup_{x \in \mathcal{M}} \|\mathrm{II}_x\| < \infty,$$

足以保证存在一致管状邻域，即假设 II；定理 F.2 展示了这一点。

定理 F.2 设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个黎曼子流形。假设典范嵌入 $\iota\colon \mathcal{M} \to \mathbb{R}^n$ 是 L_ι-双 *Lipschitz* 连续，并且

$$\kappa := \sup_{x \in \mathcal{M}} \|\mathrm{II}_x\| < \infty.$$

于是 $\mathcal{M}$ 容许一个一致管状邻域，其半径为

$$r_0 = \frac{1}{\kappa(1 + 2L_\iota)}.$$

证明 考虑管状映射 $T\colon\ N\mathcal{M} \to \mathbb{R}^n$ 由 $T(x,\ \xi) = x + \xi$给出。由于 $\kappa = \sup_{\mathcal{M}} \|\mathrm{II}\|_{\mathrm{op}} < \infty$ 与 $r_0 < 1/\kappa$，命题 E.5 表明 T 在

$$\mathcal{U}_{r_0} := \{(x,\xi) \in N\mathcal{M} \colon \|\xi\| < r_0\}.$$

因此，只需证明 T 在 $\mathcal{U}_{r_0}$上是单射。

假设存在 $(x,\xi) \neq (y,\eta)$，属于 $\mathcal{U}_{r_0}$，使得 $T(x,\xi) = T(y,\eta)$。记

$$z := x + \xi = y + \eta,$$

我们必然有 $x \neq y$。此外，

$$\|x - y\| \le \|\xi\| + \|\eta\| < 2r_0,$$

而 L_ι-双 Lipschitz 性质给出

$$0 < \rho := d_{\mathcal{M}}(x,y) \le L_\iota \|x - y\| < 2L_\iota r_0. \tag{F.2}$$

令 $\gamma\colon [0,\rho] \to \mathcal{M}$ 为一条从 x 到 y的单位速度极小测地线，使得 $\gamma(0) = x, \gamma(\rho) = y$, $\|\dot\gamma(t)\| \equiv 1$，且 $L(\gamma) = \rho$ （根据 Hopf–Rinow定理，存在性得以保证，因为假设$\mathcal{M}$是闭的）。定义

$$f(t) := \|\gamma(t) - z\|^2, \quad t \in [0,\rho].$$

Then

$$f'(t) = 2\left\langle \dot\gamma(t),\, \gamma(t) - z \right\rangle.$$

由于 $z - x = \xi \in N_x\mathcal{M}$ 和 $z - y = \eta \in N_y\mathcal{M}$，我们有

$$f'(0) = 0, \qquad f'(\rho) = 0, \tag{F.3}$$

因为 $\dot\gamma(0) \in T_x\mathcal{M}$ 且 $\dot\gamma(\rho) \in T_y\mathcal{M}$。

接下来，沿着 γ 应用高斯公式，并注意到 γ 是 $\mathcal{M}$上的一条测地线，

$$\gamma''(t) = \nabla_{\dot\gamma(t)}\dot\gamma(t) + \mathrm{II}(\dot\gamma(t),\dot\gamma(t)) = \mathrm{II}(\dot\gamma(t),\dot\gamma(t)).$$

因此，

$$\begin{aligned} f''(t) &= 2\|\dot\gamma(t)\|^2 + 2\left\langle \gamma''(t),\, \gamma(t) - z \right\rangle \\ &= 2 + 2\left\langle \mathrm{II}(\dot\gamma(t),\dot\gamma(t)),\, \gamma(t) - z \right\rangle. \end{aligned}$$

利用 $\|\dot\gamma(t)\| \equiv 1$ 以及界 $\|\mathrm{II}\|_{\mathrm{op}} \le \kappa$，可得

$$f''(t) \ge 2 - 2\kappa\|\gamma(t) - z\|.$$

此外，

$$\|\gamma(t) - z\| \le \|\gamma(t) - x\| + \|x - z\| \le d_{\mathcal{M}}(\gamma(t),\gamma(0)) + \|\xi\| < t + r_0,$$

其中我们使用了 ι是 1-Lipschitz 这一事实，即（F.1）。因此，由（F.2）可得，

$$f''(t) > 2 - 2\kappa(t + r_0) \ge 2 - 2\kappa(\rho + r_0) > 2 - 2\kappa(2L_\iota r_0 + r_0) = 0, \quad \forall\ t \in [0,\rho],$$

其中，最后一个等式用到了定义 $r_0 = 1/(\kappa(1+2L_\iota))$。因此，$f'$在 $[0,\rho]$上严格递增，所以 $f'(0) < f'(\rho)$，因为 $\rho > 0$，这与 (F.3) 矛盾。这个矛盾说明 T在 $\mathcal{U}_{r_0}$上是单射，因而 $T\colon \mathcal{U}_{r_0} \to T(\mathcal{U}_{r_0})$ 是一个微分同胚。 ∎

附录 G. $A(x)$的几何

(4) 中的额外漂移项 $A(x)$ 对控制外在 EM $\boldsymbol{Y}_k^h$与内在 GEM $\boldsymbol{X}_k^h$之间的强路径偏差起着重要作用。本节首先讨论其几何意义（第 G.1 节），随后介绍一个有助于后续差异估计的公式。

G.1 命题 10 的证明

证明 [命题 10] 的证明。由 Hsu（2002）中的定理 3.1.4，对于任意 $f \in C^\infty(\mathcal{M})$,

$$\Delta_{\mathcal{M}} f = \sum_{\alpha=1}^{n} P_\alpha(P_\alpha(f)),$$

其中 $\Delta_{\mathcal{M}}$ 是拉普拉斯–贝尔特拉米算子，且 $P_\alpha(x) = P(x)e_\alpha$ 对于 $\{e_\alpha\}_{\alpha=1}^n$，即 $\mathbb{R}^n$的典范基。

令 $\iota = (\iota^1, \ldots, \iota^n)$: $\mathcal{M} \hookrightarrow \mathbb{R}^n$ 为典范嵌入，其中 ι^j: $\mathcal{M} \to \mathbb{R}$ 由 $\iota^j(x) = x^j$ 给出，对任意 $x = (x^1, \ldots, x^n) \in \mathcal{M} \subset \mathbb{R}^n$。那么

$$\Delta_{\mathcal{M}} \iota^j = \sum_{\alpha=1}^{n} P_\alpha(P_\alpha(\iota^j)).$$

此外，将 $\iota^j(x^1, \ldots, x^n) = x^j$ 延拓到 $\mathbb{R}^n$后,

$$P_\alpha(\iota^j) = \sum_{\beta=1}^{n} P_{\beta\alpha} \frac{\partial \iota^j}{\partial x^\beta} = P_{j\alpha}.$$

因此,

$$A^j(x) := \frac{1}{2} \sum_{\alpha=1}^{n} P_\alpha(P_{j\alpha})(x) = \frac{1}{2} \sum_{\alpha=1}^{n} P_\alpha\left(P_\alpha(\iota^j)\right) = \frac{1}{2} \Delta_{\mathcal{M}} \iota^j(x).$$

另一方面，由引理 G.2,

$$\Delta_{\mathcal{M}} \iota^j(x) = \sum_{i=1}^{m} \mathrm{II}_x^j\left(E_i(x), E_i(x)\right),$$

其中 $\mathrm{II} = \left(\mathrm{II}^1, \cdots, \mathrm{II}^n\right) \in \mathbb{R}^n$ 是第二基本形式。因此

$$A^j(x) = \frac{1}{2} \sum_{i=1}^{m} \mathrm{II}_x^j\left(E_i(x), E_i(x)\right),$$

对于 $j = 1, \ldots, n$，因此

$$A(x) = \frac{1}{2} \sum_{i=1}^{m} \mathrm{II}_x\left(E_i(x), E_i(x)\right).$$

■

推论 G.1 假设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个 m 维黎曼子流形，其满足$\kappa =$上确界$_{\mathcal{M}}\|$第二基本形式$_x\| < \infty$。对于命题 *10* 中定义的 $A(x)$，

$$\sup_{x\in\mathcal{M}} \|A(x)\| < \infty.$$

证明 固定 $x \in \mathcal{M}$，并选取 $T_x\mathcal{M}$的一组标准正交基 $\{E_i(x)\}_{i=1}^m$。由命题 10,

$$\|A(x)\| \leq \frac{1}{2}\sum_{i=1}^m \|\mathrm{II}_x\left(E_i(x), E_i(x)\right)\| \leq \frac{1}{2}m\kappa.$$

对 $x \in \mathcal{M}$ 取上确界，得到

$$\sup_{x\in\mathcal{M}} \|A(x)\| \leq \frac{1}{2}m\kappa < \infty.$$

■

引理 G.2 设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个 m 维黎曼子流形，其典范嵌入为 $\iota = (\iota^1, \ldots, \iota^n)\colon \mathcal{M} \hookrightarrow \mathbb{R}^n$，其中 $\iota^j\colon \mathcal{M} \to \mathbb{R}$ 由 $\iota^j(x) = x^j$ 给出，对应于$x = (x^1, \ldots, x^n) \in \mathcal{M} \subset \mathbb{R}^n$。对于任意 $x \in \mathcal{M}$，令 $\{E_i(x)\}_{i=1}^m$ 为$T_x\mathcal{M}$的一组标准正交基。那么

$$\Delta_{\mathcal{M}}\iota^j(x) = \sum_{i=1}^m \mathrm{II}_x^j\left(E_i(x), E_i(x)\right), \quad \forall\, x \in \mathcal{M},\ \forall\, j = 1, \ldots, n,$$

其中 $\Delta_{\mathcal{M}}$ 是拉普拉斯–贝尔特拉米算子，而 $\mathrm{II} = \left(\mathrm{II}^1, \cdots, \mathrm{II}^n\right) \in \mathbb{R}^n$ 是第二基本形式。

证明 下文中，对于任意 $u \in \mathbb{R}^n$，我们用 $u^\perp(x)$ 表示u到 $N_x\mathcal{M}$上的正交投影。

根据 $\Delta_{\mathcal{M}}$的定义,

$$\Delta_{\mathcal{M}}\iota^j = \sum_{i=1}^m \nabla^2\iota^j\left(E_i, E_i\right) = \sum_{i=1}^m \nabla_{E_i}\left(\nabla\iota\right)\left(E_i\right).$$

令 $\nabla\iota^j$ 表示沿着 $\mathcal{M}$ 的 ι^j 梯度。那么

$$\begin{aligned}
\Delta_{\mathcal{M}}\iota^j &= \sum_{i=1}^m \left\langle \nabla_{E_i}\nabla\iota^j, E_i\right\rangle \\
&= \sum_{i=1}^m \left(\left\langle D_{E_i}\nabla\iota^j, E_i\right\rangle - \left\langle \left(D_{E_i}\nabla\iota^j\right)^\perp, E_i\right\rangle\right) \\
&= \sum_{i=1}^m \left\langle D_{E_i}\nabla\iota^j, E_i\right\rangle,
\end{aligned}$$

因为 $\left(D_{E_i}\nabla\iota^j\right)^\perp \in N\mathcal{M}$，因此 $\left\langle \left(D_{E_i}\nabla\iota^j\right)^\perp, E_i \right\rangle = 0$。在右侧
si
,

$$\nabla\iota^j = \nabla^{\mathbb{R}}\iota^j - \left(\nabla^{\mathbb{R}}\iota^j\right)^\perp,$$

其中 ι^j 通过 $\iota^j(x^1,\ldots,x^n) = x^j$ 扩展为 $\mathbb{R}^n$，而 $\nabla^{\mathbb{R}}\iota^j$ 表示 $\mathbb{R}^n$ 上的欧几里得梯度。注意，$\nabla^{\mathbb{R}}\iota^j = e_j$，其中 $\{e_\alpha\}_{\alpha=1}^n$ 是 $\mathbb{R}^n$ 的标准基，因此 $D_{E_i}(\nabla^{\mathbb{R}}\iota^j) = 0$。因此，

$$\Delta_{\mathcal{M}}\iota^j = \sum_{i=1}^m \left\langle D_{E_i}\nabla\iota^j, E_i\right\rangle = -\sum_{i=1}^m \left\langle D_{E_i}(e_j^\perp), E_i\right\rangle.$$

此外，由于 $\left\langle e_j^\perp, E_i\right\rangle = 0$，我们有

$$0 = D_{E_i}\left(\left\langle e_j^\perp, E_i\right\rangle\right) = \left\langle D_{E_i}(e_j^\perp), E_i\right\rangle + \left\langle e_j^\perp, D_{E_i}E_i\right\rangle.$$

因此，

$$\Delta_{\mathcal{M}}\iota^j = \sum_{i=1}^m \left\langle e_j^\perp, D_{E_i}E_i\right\rangle = \sum_{i=1}^m \left\langle e_j, (D_{E_i}E_i)^\perp\right\rangle = \sum_{i=1}^m \langle e_j, \mathrm{II}(E_i,E_i)\rangle = \sum_{i=1}^m \mathrm{II}^j\left(E_i,E_i\right).$$

■

G.2 诱导公式

引理 G.3 设 $\mathcal{M}\subset\mathbb{R}^n$ 是一个 m 维黎曼流形，并配有光滑定义的正交投影 $P(x)$。设 $A(x)$ 按式 *(5)* 定义。若 $\boldsymbol{\xi}\sim\mathcal{N}(0,I_n)$，则对任意 $x\in\mathcal{M}$，

$$A(x) = \frac{1}{2}\mathbb{E}\left[\mathrm{II}_x(P(x)\boldsymbol{\xi}, P(x)\boldsymbol{\xi})\right].$$

证明 固定 $x\in\mathcal{M}$，并选取 $T_x\mathcal{M}$ 的一组标准正交基 $\{E_i(x)\}_{i=1}^m$。由于 $P(x)\boldsymbol{\xi}\in T_x\mathcal{M}$，我们可以写成

$$P(x)\boldsymbol{\xi} = \sum_{i=1}^m \boldsymbol{\zeta}_i E_i(x),$$

其中

$$\boldsymbol{\zeta}_i = \langle \boldsymbol{\xi}, E_i(x)\rangle, \quad i=1,\ldots,m.$$

注意，$\boldsymbol{\zeta} = (\boldsymbol{\zeta}_i)\sim\mathcal{N}(0,\ I_m)$（Muirhead，2009）。由于 II_x 是双线性的，

$$\mathrm{II}_x(P(x)\boldsymbol{\xi}, P(x)\boldsymbol{\xi}) = \sum_{i,j=1}^m \boldsymbol{\zeta}_i\boldsymbol{\zeta}_j \mathrm{II}_x\left(E_i(x), E_j(x)\right).$$

取期望并应用命题 10，可得

$$\mathbb{E}\left[\mathrm{II}_x(P(x)\boldsymbol{\xi}, P(x)\boldsymbol{\xi})\right] = \sum_{i,j=1}^m \mathbb{E}\left[\boldsymbol{\zeta}_i\boldsymbol{\zeta}_j\right]\mathrm{II}_x\left(E_i(x), E_j(x)\right) = \sum_{i=1}^m \mathrm{II}_x\left(E_i(x), E_i(x)\right) = 2A(x).$$

■

附录 H. 延拓的全局 Lipschitz 连续性

设 $\mathcal{M} \subset \mathbb{R}^n$ 是一个黎曼子流形，且令 $F\colon\ \mathcal{M} \to \mathbb{R}^n$ （或 $\mathbb{R}^{n\times n}$）。根据管状邻域定理（定理 B.1）以及 $\mathcal{M}$ 的闭性，存在两个管状邻域 $\mathcal{V}_0$，$\mathcal{V}$，使得 $\mathcal{M} \subset \mathcal{V}_0 \subset \mathcal{V}$。为将 F 延拓至 $\mathbb{R}^n$，我们应用乌雷松引理（ Munkres，2013 年）并定义

$$\widetilde{F}(y) := \chi(y)F(\pi(y)), \quad \forall\, y \in \mathbb{R}^n,$$

其中 π 是 $\mathcal{V}$的典范投影，且 $\chi\colon\ \mathbb{R}^n \to [0,1]$ 是一个鼓包函数，使得 $\chi|_{\mathcal{M}} \equiv 1$ 且 $\chi|_{\mathbb{R}^n\setminus\mathcal{V}_0} \equiv 0$。为使 F 全局利普希茨连续，只需获得关于

$$D\widetilde{F}(y)[w] = D\chi(y)[w]F(\pi(y)) + \chi(y)DF(\pi(y))\left[D\pi(y)[w]\right], \quad \forall\, y \in \mathbb{R}^n,\ \forall\, w \in \mathbb{R}^n. \tag{H.1}$$

我们将此构造应用于随机微分方程 (4) 中的系数 $F(x) = V(x)$、 $A(x)$和 $P(x)$。以下结果在 $\mathcal{M}$ 的几何有界性以及 V的正则性条件下给出全局 Lipschitz 延拓。

H.1 隆起函数的有界性

我们首先需要 $\|D\chi(y)\|$的一致界。一般而言，若$\mathcal{M}$ 不容许一致管状邻域，则无法得到这样的界，这正是我们需要假设 II 的原因。

引理 H.1 设 $\mathcal{M} \subset \mathbb{R}^n$ 为容许一致管状邻域 $\mathcal{V}$的黎曼子流形。则存在满足$\mathcal{M} \subset \mathcal{V}_0 \subset \mathcal{V}$的一致管状邻域 $\mathcal{V}_0$，以及鼓包函数 $\chi \in C^\infty(\mathbb{R}^n)$ 使得，

$$\chi \in [0,1], \quad \chi|_{\mathcal{M}} \equiv 1, \quad \chi|_{\mathbb{R}^n\setminus\overline{\mathcal{V}_0}} \equiv 0, \qquad \textit{and}\ \sup_{y\in\mathbb{R}^n} \|D\chi(y)\| < \infty.$$

证明 令 $r_0 > 0$ 为 $\mathcal{V}$的一致半径。则 $\mathcal{V} = \mathcal{V}_{r_0}$ （定理 B.1），其中 $\mathcal{V}_r = \{y \in \mathbb{R}^n\colon\ d(y, \mathcal{M})$ 对任意 $r > 0$ 均成立。

由引理 N.3，我们可以选择 $\psi \in C^\infty([0,\infty))$，使得

$$\psi|_{[0,r_0/4]} \equiv 1, \quad \psi|_{[r_0/2,\infty)} \equiv 0, \qquad \text{and } C_\psi := \sup_{t\ge 0} \left|\psi'(t)\right| < \infty.$$

令 $g(y) := d(y, \mathcal{M})\colon\ \mathcal{V} \to \mathbb{R}$。定义

$$\chi(y) := \psi(g(y)), \quad \forall\, y \in \mathbb{R}^n.$$

于是 $\chi \in [0,1]$。此外，$\chi(y)$ 在 $\mathcal{V}_{r_0/4}$ 上，且 $\chi(y)$ 在 $\mathbb{R}^n\setminus\mathcal{V}_{r_0/2}$上。令 $\mathcal{V}_0 := \mathcal{V}_{r_0/2} \subset \mathcal{V}$，得到

$$\chi|_{\mathcal{M}} \equiv 1, \quad \chi|_{\mathbb{R}^n\setminus\overline{\mathcal{V}_0}} \equiv 0.$$

此外，$D\chi(y) = 0$ 对 $y \in \mathcal{V}_{r_0/4}$ 以及对$y \in \mathbb{R}^n\setminus\overline{\mathcal{V}_{r_0/2}}$ 对于

$$y \in \overline{\mathcal{V}_{r_0/2}}\setminus\mathcal{V}_{r_0/4} \subset \mathcal{V}\setminus\mathcal{M},$$

该距　距离函数 g 在 $\mathcal{V}\backslash\mathcal{M}$上光滑，并且 $\|Dg(y)\| = 1$ 在该处成立（Lee, 2018）。因此　e,

$$\|D\chi(y)\| = \left|\psi'(g(y))\right| \|Dg(y)\| \le C_\psi.$$

因此,

$$\sup_{y\in\mathbb{R}^n} \|D\chi(y)\| \le C_\psi < \infty.$$

■

H.2 引理 9 的证明

证明 [引理 9] 的证明我们按照以下步骤，分别处理 $V(x)$ 和 $P(x)$。

- 内在导数：首先，$\kappa := \sup_{\mathcal{M}} \|\mathrm{II}_x\| < \infty$，由假设 I 和引理 C.3 可得。对于 $P(x)$，引理 D.4 表明

$$\sup_{x\in\mathcal{M}} \|\nabla P(x)\|_{\mathrm{op}} \le \kappa < \infty.$$

  对于 $V(x)$，假设 III 确保

$$C_{2,V} := \sup_{x\in\mathcal{M}} \|\nabla V(x)\|_{\mathrm{op}} < \infty.$$

- 外在导数：固定 $x \in \mathcal{M}$，并令 $v \in T_x\mathcal{M}$ 满足 $\|v\| = 1$。对于 $P(x)$，由高斯公式（Lee, 2018, 定理 8.2），

$$D_v(P(x)w) = \nabla_v(P(x)w) + \mathrm{II}(P(x)w, v), \quad \forall\, w \in \mathbb{R}^n.$$

  利用上述内在界，我们得到

$$\begin{aligned}\|D_v(P(x)w)\| &\le \|\nabla_v(P(x)w)\| + \|\mathrm{II}(P(x)w, v)\| \\ &\le \|\nabla P(x)\|_{\mathrm{op}}\|v\|\|w\| + \|\mathrm{II}\|_{\mathrm{op}}\|P(x)w\|\|v\| \\ &\le 2\kappa\|w\|,\end{aligned}$$

  因为 $P(x)$ 是正交投影。将 $w \mapsto D_v(P(x)w)$ 视为线性映射，可得

$$\|D_vP(x)\|_{\mathrm{op}} = \sup_w \frac{\|D_v(P(x)w)\|}{\|w\|} \le 2\kappa.$$

  此外，由于 $\|D_vP(x)\|_{\mathrm{F}} \le \sqrt{n}\|D_v\overline{P}(x)\|_{\mathrm{op}}$（Horn and Johnson, 2012），

$$C_P := \sup_{x\in\mathcal{M}} \sup_{v\in T_x\mathcal{M}, \|v\|=1} \|D_vP(x)\|_{\mathrm{F}} \le 2\sqrt{n}\kappa < \infty. \tag{H.2}$$

  对于 $V(x)$，假设 III 还给出 $C_{1,V} := \sup_{x\in\mathcal{M}}\|V(x)\| < \infty$。因此,

$$\begin{aligned}\|D_v(V(x))\| &\le \|\nabla_v(V(x))\| + \|\mathrm{II}(V(x), v)\| \\ &\le \|\nabla V(x)\|_{\mathrm{op}}\|v\| + \|\mathrm{II}\|_{\mathrm{op}}\|V(x)\|\|v\| \\ &\le (C_{2,V} + \kappa C_{1,V})\,\|v\|,\end{aligned}$$

  这意味着

$$C_V := \sup_{x\in\mathcal{M}} \sup_{v\in T_xM, \|v\|=1} \|D_vV(x)\| \le C_{2,V} + \kappa C_{1,V} < \infty. \tag{H.3}$$

• 全局延拓：根据引理 D.3，由于 $\mathcal{M}$ 具有一致管状邻域，且 $\kappa = \sup_{\mathcal{M}} \|\mathrm{II}_x\| < \infty$，因此存在一个一致管状邻域 $\mathcal{V}$，并带有典范投影 π，使得

$$C_\pi := \sup_{y \in \mathcal{V}} \|D\pi(y)\|_{\mathrm{op}} < \infty. \tag{H.4}$$

此外，由引理 H.1，存在一个一致管状邻域 $\mathcal{V}_0$ 使得$\mathcal{V}_0 \subset \mathcal{V}$，以及一个鼓包函数 $\chi \in C^\infty(\overline{\mathbb{R}^n})$，满足

$$\chi \in [0,1], \quad \chi|_{\mathcal{M}} \equiv 1, \quad \chi|_{\mathbb{R}^n \setminus \overline{\mathcal{V}_0}} \equiv 0, \tag{H.5}$$

and

$$C_\chi := \sup_{y \in \mathbb{R}^n} \|D\chi(y)\| < \infty. \tag{H.6}$$

定义

$$\widetilde{V}(y) := \chi(y) V(\pi(y)), \quad \widetilde{P}(y) := \chi(y) P(\pi(y)), \qquad \forall\, y \in \mathbb{R}^n,$$

由于对所有 $y \in \mathcal{V}_0 \subset \mathcal{V}$，$\pi(y) \in \mathcal{M}$，且对所有$y \in \overline{\mathbb{R}^n} \setminus \mathcal{V}_0$，$\chi(y) = 0$，因此它们均有良好定义。

• P的全局 Lipschitz 性：对任意 $y \in \mathbb{R}^n$ 以及满足 $\|u\| = 1$的任意 $u \in \mathbb{R}^n$，由链式法则可得

$$\begin{aligned} D_u \widetilde{P}(y) &= (D_u \chi(y)) P(\pi(y)) + \chi(y) DP(\pi(y)) \left[D\pi(y)[u] \right] \\ &= (D_u \chi(y)) P(\pi(y)) + \chi(y) D_v P(\pi(y)), \end{aligned}$$

其中 $v = D\pi(y)[u] \in T_{\pi(y)}\mathcal{M}$；参见附录 D。由 (H.2)、(H.4) 和 (H.6)，

$$\begin{aligned} \left\| D_u \widetilde{P}(y) \right\|_{\mathrm{F}} &\le \|(D_u \chi(y)) P(\pi(y))\|_{\mathrm{F}} + \|\chi(y) D_v P(\pi(y))\|_{\mathrm{F}} \\ &\le |D_u \chi(y)| \, \|P(\pi(y))\|_{\mathrm{F}} + \|D_v P(\pi(y))\|_{\mathrm{F}} \\ &\le C_\chi \|P(\pi(y))\|_{\mathrm{F}} \|u\| + C_P \|v\| \\ &= C_\chi \|P(\pi(y))\|_{\mathrm{F}} \|u\| + C_P \|D\pi(y)[u]\| \\ &\le \sqrt{n} C_\chi + C_P C_\pi, \end{aligned}$$

其中，最后一个不等式利用了 $\|P(\pi(y))\|_{\mathrm{F}} \le \sqrt{n}$，因为 $P(\pi(y))$ 是一个正交投影。因此，

$$L_P := \sup_{y \in \mathbb{R}^n} \left\| D\widetilde{P}(y) \right\|_{\mathrm{op}} \le C_P C_\pi + \sqrt{n} C_\chi < \infty,$$

因此，$\widetilde{P}\colon \mathbb{R}^n \to \mathbb{R}^{n \times n}$ 是全局利普希茨连续的，即

$$\left\| \widetilde{P}(y_1) - \widetilde{P}(y_2) \right\|_{\mathrm{F}} \le L_P \|y_1 - y_2\|, \quad \forall\, y_1,\, y_2 \in \mathbb{R}^n.$$

- $\widetilde{V}$的全局 Lipschitz 性：对于任意 $y \in \mathbb{R}^n$ 和任意满足 $\|u\| = 1$的 $u \in \mathbb{R}^n$，类似地 y,

$$D_u\widetilde{V}(y) = (D_u\chi(y))V(\pi(y)) + \chi(y)D_vV(\pi(y)),$$

其中 $v = D\pi(y)[u] \in T_{\pi(y)}\mathcal{M}$。因此，由 $C_{1,\ V} = \sup_{x\in\mathcal{M}}\|V(x)\| < \infty$ 以及（H.4）、（H.3）和（H.6），

$$\begin{aligned}\left\|D_u\widetilde{V}(y)\right\| &\le \|(D_u\chi(y))V(\pi(y))\| + \|\chi(y)D_vV(\pi(y))\| \\ &\le C_\chi\|u\|\|V(\pi(y))\| + C_V\|v\| \\ &\le C_\chi C_{1,V} + C_V C_\pi.\end{aligned}$$

因此,

$$L_V := \sup_{y\in\mathbb{R}^n}\left\|D\widetilde{V}(y)\right\|_{\mathrm{op}} \le C_\chi C_{1,V} + C_V C_\pi < \infty,$$

$\widetilde{V}$从而 $V:\ \mathbb{R}^n \to \mathbb{R}^n$ 是全局利普希茨连续的，即

$$\left\|\widetilde{V}(y_1) - \widetilde{V}(y_2)\right\| \le L_V\|y_1 - y_2\|, \quad \forall\ y_1,\ y_2 \in \mathbb{R}^n.$$

■

H.3 引理 11 的证明

证明 [引理 11] 的证明。假设 $\dim\mathcal{M} = m$。下文中，对于任意 $u \in \mathbb{R}^n$，我们将 $u^\top(x)$ $u^\perp(x)$ 分别记作 u 关于 $T_x\mathcal{M}$ 的切向分量和正交分量。证明分为以下几个步骤。

- A 的有界性：首先，由假设 I，引理 C.3 给出

$$\kappa_1 := \sup_{x\in\mathcal{M}}\|\mathrm{II}_x\|_{\mathrm{op}} < \infty.$$

因此，根据推论 G.1,

$$C_A^0 := \sup_{x\in\mathcal{M}}\|A(x)\| < \infty. \tag{H.7}$$

• 外在导数：固定 $x \in \mathcal{M}$。我们可以选择 x 附近的一个局部正交标架 $\{E_i\}_{i=1}^m$，使得对于任意 $i,\ j = 1,\ \ldots\ldots,\ m$，均有 $\nabla_{E_j}E_i(x) = 0$ (Lee, 2018)。令 $v \in T_xM$，其中 $\|v\| = 1$。于是，对于所有 i，均有 $\nabla_v E_i(x) = 0$。此外,

$$\begin{aligned}D_v\mathrm{II}_x\left(E_i(x), E_i(x)\right) &= \left(D_v\mathrm{II}_x\left(E_i(x), E_i(x)\right)\right)^\top + \left(D_v\mathrm{II}_x\left(E_i(x), E_i(x)\right)\right)^\perp \\ &= -S_{\mathrm{II}(E_i,E_i)}(v) + D_v^\perp\mathrm{II}_x\left(E_i(x), E_i(x)\right),\end{aligned} \tag{H.8}$$

其中，第二个等号由魏因加滕方程（Lee, 2018, 命题 8.4）得出。对于 (H.8) 右侧的第一项，由于 $\kappa_1 = \sup_{\mathcal{M}}\|\mathrm{II}\|_{\mathrm{op}} < \infty$，根据引理 C.2 可得

$$\left\|S_{\mathrm{II}(E_i,E_i)}(v)\right\| \le \kappa_1\|\mathrm{II}_x\left(E_i(x), E_i(x)\right)\| \le \kappa_1^2. \tag{H.9}$$

F对于 (H.8) 右侧的第二项，根据 (C.3) 和 $\nabla_v E_i(x) = 0$,

$$D_v^{\perp}\mathrm{II}_x\left(E_i(x), E_i(x)\right) = \left(\nabla_v \mathrm{II}_x\right)\left(E_i(x), E_i(x)\right).$$

由于 $\kappa_2 := \sup_{\mathcal{M}} \|\nabla \mathrm{II}\|_{\mathrm{op}} < \infty$ （假设 I 与引理 C.3），

$$\left\| D_v^{\perp}\mathrm{II}_x\left(E_i(x), E_i(x)\right) \right\| \le \kappa_2. \tag{H.10}$$

结合 (H.9)、(H.10) 与 (H.8)，我们得到

$$\|D_v \mathrm{II}_x\left(E_i(x), E_i(x)\right)\| \le \kappa_1^2 + \kappa_2,$$

从而，由 (6),

$$\|D_v A(x)\| \le \frac{1}{2}\sum_{i=1}^{m} \|D_v \mathrm{II}_x\left(E_i(x), E_i(x)\right)\| \le \frac{1}{2} m\left(\kappa_1^2 + \kappa_2\right).$$

因此,

$$C_A := \sup_{x\in\mathcal{M}} \sup_{v\in T_x M, \|v\|=1} \|D_v A(x)\| \le \frac{1}{2} m\left(\kappa_1^2 + \kappa_2\right) < \infty. \tag{H.11}$$

• 全局延拓：如同引理 9（附录 H.2）的证明，选取一致管状邻域 $\mathcal{V}$ 连同典范投影 π，使其满足 (H.4)，以及一致管状邻域 $\mathcal{V}_0 \subset \mathcal{V}$ 连同鼓包函数 χ，使其满足 (H.5) 与 (H.6)。定义

$$\widetilde{A}(y) := \chi(y) A(\pi(y)), \quad \forall\, y \in \mathbb{R}^n,$$

其是良定义的。

• 全局 Lipschitz：对任意 $y \in \mathbb{R}^n$ 以及任意满足 $\|u\| = 1$ 的 $u \in \mathbb{R}^n$，令 $v = D\pi(y)[u] \in T_{\pi(y)}\mathcal{M}$。仿照引理 9 的证明进行论证，并利用 (H.4)、(H.6)、(H.7) 和 (H.11),

$$\begin{aligned}
\left\| D_u \widetilde{A}(y) \right\| &\le \|(D_u \chi(y)) A(\pi(y))\| + \|\chi(y) D_v A(\pi(y))\| \\
&\le C_\chi \|u\| \|A(\pi(y))\| + C_A \|v\| \\
&\le C_\chi C_A^0 + C_A C_\pi.
\end{aligned}$$

因此,

$$L_A := \sup_{y\in\mathbb{R}^n} \left\| D\widetilde{A}(y) \right\|_{\mathrm{op}} \le C_\chi C_A^0 + C_A C_\pi < \infty,$$

$\widetilde{}$和 A是 L_A-利普希茨连续的。

■

附录 I. 欧几里得欧拉–马鲁雅马格式

考虑欧几里得随机微分方程

$$\mathrm{d}\boldsymbol{X}_t = b(t, \boldsymbol{X}_t)\,\mathrm{d}t + \sigma(t, \boldsymbol{X}_t)\,\mathrm{d}\boldsymbol{W}_t, \quad t \in [0, T], \tag{I.1}$$

其中 $b\colon [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$, $\sigma\colon [0,T]\times\mathbb{R}^n \to \mathbb{R}^{n\times n}$, 且 $(\boldsymbol{W}_t)_{t\in[0,T]}$是 $\mathbb{R}^n$中的标准布朗运动。为使用离散时间格式逼近 (I.1)，取步长 $h = T/N$，使得$N \in \mathbb{N}$，并令 $t_k = kh$，其中 $k = 0$，1，.....，N。欧拉–马鲁雅马方法定义$\{\boldsymbol{X}_k^h\}_{k=0}^N$为 $\boldsymbol{X}_0^h = \boldsymbol{X}_0$，并且

$$\boldsymbol{X}_{k+1}^h = \boldsymbol{X}_k^h + b(t_k, \boldsymbol{X}_k^h)h + \sigma(t_k, \boldsymbol{X}_k^h)\Delta\boldsymbol{W}_k, \tag{I.2}$$

其中 $\Delta\boldsymbol{W}_k = \boldsymbol{W}_{t_{k+1}} - \boldsymbol{W}_{t_k} \sim \mathcal{N}(0,\ hI_n)$（等价地，采样 $\boldsymbol{\xi} \sim \mathcal{N}(0,\ I_n)$并令 $\Delta\boldsymbol{W}_k = \sqrt{h}\,\boldsymbol{\xi}$）。

经典结果表明（例如参见 Kloeden 和 Platen（1992年）；Milstein 和 Tretyakov（2013年）），欧拉–马鲁雅马格式以 $1/2$阶强收敛。更准确地说，若 $\boldsymbol{X}_0 \in L^2$ 和 b，σ 关于空间变量全局 Lipschitz，且关于时间变量为 $1/2$-H¨older 连续，即

$$\begin{aligned} &\|b(t,x) - b(t,y)\| + \|\sigma(t,x) - \sigma(t,y)\|_{\mathrm{F}} \le L\|x-y\|, \\ &\|b(t,x) - b(s,x)\| + \|\sigma(t,x) - \sigma(s,x)\|_{\mathrm{F}} \le H(1+\|x\|)\,|t-s|^{1/2}, \end{aligned} \tag{I.3}$$

那么存在 $C > 0$ 使得

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_{t_k} - \boldsymbol{X}_k^h\right\|^2\right] \le Ch.$$

利用 Burkholder–Davis–Gundy (BDG) 不等式，这一 L^2 强误差估计可以在相同假设下推广为任意 $p \ge 1$ 时的 L^p 估计。由于标准参考文献中并不总是明确陈述这一推广，我们在下文给出一个自包含证明；该论证也为我们的主要结果提供了证明策略的启示。

定理 I.1 设 $1 \le p < \infty$。对于随机微分方程 *(I.1)*，假设 $\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right] < \infty$ 且b、σ满足 *(I.3)*。令 $h = T/N$适用于 $N \in \mathbb{N}$，并令 $t_k = kh$适用于 $k = 0$、1、.....、N。设 $\{\boldsymbol{X}_k^h\}_{k=0}^N$由欧拉–马鲁雅马格式 *(I.2)* 在 $\boldsymbol{X}_0^h = \boldsymbol{X}_0$下构造。则存在常数 $C_p > 0$使得

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_{t_k} - \boldsymbol{X}_k^h\right\|^p\right] \le C_p h^{p/2}.$$

证明 我们首先假设 $p \ge 2$。注意到 (I.3) 意味着存在常数 $K > 0$使得

$$\|b(t,x)\| + \|\sigma(t,x)\|_{\mathrm{F}} \le K(1+\|x\|). \tag{I.4}$$

证明包括以下步骤。

- 步骤 1. $\boldsymbol{X}_k^h$的矩有界性：由 (I.2) 可得,

$$\boldsymbol{X}_k^h = \boldsymbol{X}_{k-1}^h + b(t_{k-1}, \boldsymbol{X}_{k-1}^h)h + \sigma(t_{k-1}, \boldsymbol{X}_{k-1}^h)\Delta\boldsymbol{W}_{k-1}$$

$$
\begin{aligned}
&= \boldsymbol{X}_0 + h\sum_{s=0}^{k-1} b(t_s, \boldsymbol{X}_s^h) + \sum_{s=0}^{k-1} \sigma(t_s, \boldsymbol{X}_s^h)\Delta \boldsymbol{W}_s \\
&= \boldsymbol{X}_0 + \boldsymbol{A}_k + \boldsymbol{M}_k,
\end{aligned}
$$

其中

$$
\boldsymbol{A}_k := h\sum_{s=0}^{k-1} b(t_s, \boldsymbol{X}_s^h), \quad \boldsymbol{M}_k := \sum_{s=0}^{k-1} \sigma(t_s, \boldsymbol{X}_s^h)\Delta \boldsymbol{W}_s.
$$

对于任意 $\ell \in \mathbb{N}$，定义停时

$$
\boldsymbol{N}_\ell := \min\left\{k > 0 \colon \left\|\boldsymbol{X}_k^h\right\| > \ell\right\},
$$

并考虑停止过程 $\left\{\boldsymbol{X}_{k\wedge \boldsymbol{N}_\ell}^h\right\}_{k=0}^N$，即

$$
\boldsymbol{X}_{k\wedge \boldsymbol{N}_\ell}^h = \boldsymbol{X}_0 + \boldsymbol{A}_{k\wedge \boldsymbol{N}_\ell} + \boldsymbol{M}_{k\wedge \boldsymbol{N}_\ell}.
$$

因此，根据 Hölder 不等式,

$$
\left\|\boldsymbol{X}_{k\wedge \boldsymbol{N}_\ell}^h\right\|^p \le 3^{p-1}\left(\|\boldsymbol{X}_0\|^p + \|\boldsymbol{A}_{k\wedge \boldsymbol{N}_\ell}\|^p + \|\boldsymbol{M}_{k\wedge \boldsymbol{N}_\ell}\|^p\right).
$$

因此,

$$
\begin{aligned}
\mathbb{E}\left[\max_{0\le m\le k\wedge \boldsymbol{N}_\ell}\left\|\boldsymbol{X}_m^h\right\|^p\right] \le 3^{p-1}\Big(&\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right] + \mathbb{E}\left[\max_{0\le m\le k\wedge \boldsymbol{N}_\ell}\|\boldsymbol{A}_m\|^p\right] \\
&+ \mathbb{E}\left[\max_{0\le m\le k\wedge \boldsymbol{N}_\ell}\|\boldsymbol{M}_m\|^p\right]\Big).
\end{aligned} \tag{I.5}
$$

根据 (I.4) 中关于 b的线性增长界,

$$
\begin{aligned}
\max_{0\le m\le k\wedge \boldsymbol{N}_\ell}\|\boldsymbol{A}_m\|^p &\le h^p \max_{0\le m\le k\wedge \boldsymbol{N}_\ell}\left\|\sum_{s=0}^{m-1} b(t_s, \boldsymbol{X}_s^h)\right\|^p \\
&\le h^p \max_{0\le m\le k\wedge \boldsymbol{N}_\ell}\left(\sum_{s=0}^{m-1}\left\|b(t_s, \boldsymbol{X}_s^h)\right\|\right)^p \\
&\le h^p \left(\sum_{s=0}^{(k\wedge \boldsymbol{N}_\ell)-1}\left\|b(t_s, \boldsymbol{X}_s^h)\right\|\right)^p \\
&\le h^p \left(\sum_{s=0}^{k-1}\left\|b(t_{s\wedge \boldsymbol{N}_\ell}, \boldsymbol{X}_{s\wedge \boldsymbol{N}_\ell}^h)\right\|\right)^p \\
&\le h^p k^{p-1}\sum_{s=0}^{k-1}\left\|b(t_{s\wedge \boldsymbol{N}_\ell}, \boldsymbol{X}_{s\wedge \boldsymbol{N}_\ell}^h)\right\|^p \\
&\le h t_k^{p-1}\sum_{s=0}^{k-1} K^p\left(1 + \left\|\boldsymbol{X}_{s\wedge \boldsymbol{N}_\ell}^h\right\|\right)^p
\end{aligned}
$$

$$\leq hT^{p-1}K^p 2^{p-1}\sum_{s=0}^{k-1}\left(1+\left\|\boldsymbol{X}^h_{s\wedge\boldsymbol{N}_\ell}\right\|^p\right)$$
$$\leq hT^{p-1}K^p 2^{p-1}\sum_{s=0}^{k-1}\left(1+\max_{0\leq m\leq s\wedge\boldsymbol{N}_\ell}\left\|\boldsymbol{X}^h_m\right\|^p\right).$$

Let

$$a_s := \mathbb{E}\left[\max_{0\leq m\leq s\wedge\boldsymbol{N}_\ell}\left\|\boldsymbol{X}^h_m\right\|^p\right].$$

Then

$$\mathbb{E}\left[\max_{0\leq m\leq k\wedge\boldsymbol{N}_\ell}\|\boldsymbol{A}_m\|^p\right]\leq hT^{p-1}K^p 2^{p-1}\sum_{s=0}^{k-1}(1+a_s). \tag{I.6}$$

对于 (I.5) 中的最后一项，定义

$$\tilde{\boldsymbol{M}}_t := \int_0^t \boldsymbol{H}_u\,\mathrm{d}\boldsymbol{W}_u,\quad \boldsymbol{H}_u := \sum_{s=0}^{N-1}\sigma(t_s,\boldsymbol{X}^h_s)\mathbb{I}_{(t_s,t_{s+1}]}(u).$$

~ ~则 $\boldsymbol{M}$ 是连续局部鞅，且 $\boldsymbol{M}_{t_k}=\boldsymbol{M}_{k\circ}$

根据 BDG不等式（Marinelli 和 R¨ockner，2016），

$$\mathbb{E}\left[\max_{0\leq m\leq k\wedge\boldsymbol{N}_\ell}\|\boldsymbol{M}_m\|^p\right]\leq\mathbb{E}\left[\sup_{0\leq t\leq t_{k\wedge\boldsymbol{N}_\ell}}\|\tilde{\boldsymbol{M}}_t\|^p\right]\leq B_p\mathbb{E}\left[[\tilde{\boldsymbol{M}}]^{p/2}_{t_{k\wedge\boldsymbol{N}_\ell}}\right]$$
$$=B_p\mathbb{E}\left[\left(\sum_{s=0}^{(k\wedge\boldsymbol{N}_\ell)-1}h\left\|\sigma(t_s,\boldsymbol{X}^h_s)\right\|^2_{\mathrm{F}}\right)^{p/2}\right].$$

利用 (I.4)，我们进一步得到

$$\mathbb{E}\left[\max_{0\leq m\leq k\wedge\boldsymbol{N}_\ell}\|\boldsymbol{M}_m\|^p\right]\leq B_pK^p\mathbb{E}\left[\left(\sum_{s=0}^{(k\wedge\boldsymbol{N}_\ell)-1}h\left(1+\left\|\boldsymbol{X}^h_s\right\|\right)^2\right)^{p/2}\right]$$
$$\leq B_pK^p\mathbb{E}\left[\left(\sum_{s=0}^{k-1}h\left(1+\left\|\boldsymbol{X}^h_{s\wedge\boldsymbol{N}_\ell}\right\|\right)^2\right)^{p/2}\right].$$

由于 $p\geq 2$，由引理 N.2 可知

$$\mathbb{E}\left[\max_{0\leq m\leq k\wedge\boldsymbol{N}_\ell}\|\boldsymbol{M}_m\|^p\right]\leq B_pK^pT^{\frac{p}{2}-1}\mathbb{E}\left[\sum_{s=0}^{k-1}h\left(1+\left\|\boldsymbol{X}^h_{s\wedge\boldsymbol{N}_\ell}\right\|\right)^p\right]$$
$$\leq hB_pK^pT^{\frac{p}{2}-1}2^{p-1}\mathbb{E}\left[\sum_{s=0}^{k-1}\left(1+\max_{0\leq m\leq s\wedge\boldsymbol{N}_\ell}\left\|\boldsymbol{X}^h_m\right\|^p\right)\right]$$
$$\leq hB_pK^pT^{\frac{p}{2}-1}2^{p-1}\sum_{s=0}^{k-1}(1+a_s).$$

将上述不等式与 (I.6) 结合，并利用 (I.5)，可得

$$a_k \le C_{1,p} + C_{2,p} h \sum_{s=0}^{k-1} (1 + a_s),$$

其中 $C_{1,\ p} := 3^{p-1}\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right]$ 且

$$C_{2,p} := T^{p-1}K^p 6^{p-1} + B_p T^{\frac{p}{2}-1} K^p 6^{p-1}.$$

于是，根据离散 Gr¨onwall 不等式（Thom´ee，2007年，引理 10.5），

$$a_N \le (1 + C_{1,p}) e^{C_{2,p}T} - 1.$$

因此，

$$\mathbb{E}\left[\max_{0\le k\le N\wedge \boldsymbol{N}_\ell} \left\|\boldsymbol{X}_k^h\right\|^p\right] \le (1 + C_{1,p}) e^{C_{2,p}T} - 1.$$

当 $\ell \to \infty$ 时，法图引理（Folland，2013年）推出

$$\mathbb{E}\left[\max_{0\le k\le N} \left\|\boldsymbol{X}_k^h\right\|^p\right] \le M_p, \tag{I.7}$$

存在某个常数 M_p，仅依赖于 p。

- 步骤 2. $\boldsymbol{X}_t$的有界性：首先，通过采用与步骤 1 类似的论证（或参见 Baudoin（2014年）中定理 6.1 的证明，并用 BDG 不等式替代 Itô 等距公式），可得

$$\mathbb{E}\left[\sup_{0\le t\le T} \|\boldsymbol{X}_t\|^p\right] \le M_p'. \tag{I.8}$$

此外，由式 (I.1)，对于 $t \ge s$,

$$\boldsymbol{X}_t - \boldsymbol{X}_s = \int_s^t b(u, \boldsymbol{X}_u)\ \mathrm{d}u + \int_s^t \sigma(u, \boldsymbol{X}_u)\ \mathrm{d}\boldsymbol{W}_u.$$

因此，

$$\|\boldsymbol{X}_t - \boldsymbol{X}_s\|^p \le 2^{p-1}\left(\left\|\int_s^t b(u, \boldsymbol{X}_u)\ \mathrm{d}u\right\|^p + \left\|\int_s^t \sigma(u, \boldsymbol{X}_u)\ \mathrm{d}\boldsymbol{W}_u\right\|^p\right). \tag{I.9}$$

对于右端第一项，由 H¨older 不等式及 (I.4),

$$\begin{aligned}
\mathbb{E}\left[\left\|\int_s^t b(u, \boldsymbol{X}_u)\ \mathrm{d}u\right\|^p\right] &\le (t-s)^{p-1}\int_s^t \mathbb{E}\left[\|b(u, \boldsymbol{X}_u)\|^p\right]\ \mathrm{d}u \\
&\le (t-s)^{p-1}K^p \int_s^t \mathbb{E}\left[(1 + \|\boldsymbol{X}_u\|)^p\right]\ \mathrm{d}u \\
&\le (t-s)^{p-1}K^p 2^{p-1}\int_s^t (1 + \mathbb{E}\left[\|\boldsymbol{X}_u\|^p\right])\ \mathrm{d}u.
\end{aligned}$$

利用 (I.8)，我们得到

$$\mathbb{E}\left[\left\|\int_s^t b(u,\boldsymbol{X}_u)\ \mathrm{d}u\right\|^p\right] \le 2^{p-1}K^p(1+M_p')(t-s)^p. \tag{I.10}$$

对于 (I.9) 中的第二项，由 BDG不等式,

$$\begin{aligned}\mathbb{E}\left[\left\|\int_s^t \sigma(u,\boldsymbol{X}_u)\ \mathrm{d}\boldsymbol{W}_u\right\|^p\right] &\le B_p'\mathbb{E}\left[\left(\int_s^t \|\sigma(u,\boldsymbol{X}_u)\|_{\mathrm{F}}^2\ \mathrm{d}u\right)^{p/2}\right]\\ &\le B_p'(t-s)^{\frac{p}{2}-1}\mathbb{E}\left[\int_s^t \|\sigma(u,\boldsymbol{X}_u)\|_{\mathrm{F}}^p\ \mathrm{d}u\right],\end{aligned}$$

其中，最后一个不等式由引理 N.2 对于 $p \ge 2$ 得出。因此，由 (I.4) 和 (I.8) 可知

$$\mathbb{E}\left[\left\|\int_s^t \sigma(u,\boldsymbol{X}_u)\ \mathrm{d}\boldsymbol{W}_u\right\|^p\right] \le 2^{p-1}K^pB_p'(1+M_p')(t-s)^{p/2}. \tag{I.11}$$

结合 (I.11)、(I.10) 和 (I.9)，当 $t-s<1$ 时，存在仅依赖于 p 的 L_p，使得

$$\mathbb{E}\left[\|\boldsymbol{X}_t-\boldsymbol{X}_s\|^p\right] \le L_p(t-s)^{p/2}. \tag{I.12}$$

- 步骤 3. 比较 $\boldsymbol{X}_{t_k}$ 与 $\boldsymbol{X}_k^h$：由 (I.1),

$$\boldsymbol{X}_{t_{k+1}} = \boldsymbol{X}_{t_k} + \int_{t_k}^{t_{k+1}} b(u,\boldsymbol{X}_u)\ \mathrm{d}u + \int_{t_k}^{t_{k+1}} \sigma(u,\boldsymbol{X}_u)\ \mathrm{d}\boldsymbol{W}_u.$$

令 $\boldsymbol{e}_k \coloneqq \boldsymbol{X}_{t_k} - \boldsymbol{X}_k^h$。则

$$\boldsymbol{e}_{k+1} = \boldsymbol{e}_k + \int_{t_k}^{t_{k+1}} \left(b(u,\boldsymbol{X}_u) - b\left(t_k,\boldsymbol{X}_k^h\right)\right)\ \mathrm{d}u + \int_{t_k}^{t_{k+1}} \left(\sigma(u,\boldsymbol{X}_u) - \sigma(t_k,\boldsymbol{X}_k^h)\right)\ \mathrm{d}\boldsymbol{W}_u.$$

由于 $\boldsymbol{e}_0 = 0$，我们可以写成

$$\begin{aligned}\boldsymbol{e}_k = &\underbrace{\sum_{s=0}^{k-1}\int_{t_s}^{t_{s+1}} \left(b(u,\boldsymbol{X}_u) - b\left(t_s,\boldsymbol{X}_s^h\right)\right)\ \mathrm{d}u}_{=:\boldsymbol{B}_k}\\ &+\underbrace{\sum_{s=0}^{k-1}\int_{t_s}^{t_{s+1}} \left(\sigma(u,\boldsymbol{X}_u) - \sigma(t_s,\boldsymbol{X}_s^h)\right)\ \mathrm{d}\boldsymbol{W}_u}_{=:\boldsymbol{G}_k}.\end{aligned}$$

因此,

$$\max_{0\le m\le k}\|\boldsymbol{e}_m\|^p \le 2^{p-1}\left(\max_{0\le m\le k}\|\boldsymbol{B}_m\|^p + \max_{0\le m\le k}\|\boldsymbol{G}_m\|^p\right).$$

取期望可得

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{e}_m\|^p\right] \le 2^{p-1}\left(\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{B}_m\|^p\right] + \mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{G}_m\|^p\right]\right). \tag{I.13}$$

其中，右侧的存在性由 (I.7) 和 (I.8) 给出。对于 (I.13) 右侧的第一项,

$$\begin{aligned}\|\boldsymbol{B}_m\|^p &\leq m^{p-1}\sum_{s=0}^{m-1}\left\|\int_{t_s}^{t_{s+1}}\left(b(u,\boldsymbol{X}_u)-b\left(t_s,\boldsymbol{X}_s^h\right)\right)\,\mathrm{d}u\right\|^p \\ &\leq t_m^{p-1}\sum_{s=0}^{m-1}\int_{t_s}^{t_{s+1}}\left\|b(u,\boldsymbol{X}_u)-b\left(t_s,\boldsymbol{X}_s^h\right)\right\|^p\,\mathrm{d}u.\end{aligned} \tag{I.14}$$

由 (I.3),

$$\begin{aligned}\left\|b(u,\boldsymbol{X}_u)-b\left(t_s,\boldsymbol{X}_s^h\right)\right\|^p &\leq 3^{p-1}\Big(\|b(u,\boldsymbol{X}_u)-b\left(u,\boldsymbol{X}_{t_s}\right)\|^p \\ &\quad+\left\|b\left(u,\boldsymbol{X}_{t_s}\right)-b\left(u,\boldsymbol{X}_s^h\right)\right\|^p \\ &\quad+\left\|b(u,\boldsymbol{X}_s^h)-b\left(t_s,\boldsymbol{X}_s^h\right)\right\|^p\Big) \\ &\leq 3^{p-1}\Big(L^p\|\boldsymbol{X}_u-\boldsymbol{X}_{t_s}\|^p+L^p\|\boldsymbol{e}_s\|^p \\ &\quad+H^p\left(1+\left\|\boldsymbol{X}_s^h\right\|\right)^p h^{p/2}\Big).\end{aligned} \tag{I.15}$$

将 (I.15) 代入 (I.14) 并取期望，得到

$$\begin{aligned}\mathbb{E}\left[\max_{0\leq m\leq k}\|\boldsymbol{B}_m\|^p\right] &\leq T^{p-1}\sum_{s=0}^{k-1}\int_{t_s}^{t_{s+1}}\mathbb{E}\left[\left\|b(u,\boldsymbol{X}_u)-b\left(t_s,\boldsymbol{X}_s^h\right)\right\|^p\right]\,\mathrm{d}u \\ &\leq (3T)^{p-1}\sum_{s=0}^{k-1}\int_{t_s}^{t_{s+1}}\Big(L^p\mathbb{E}\left[\|\boldsymbol{X}_u-\boldsymbol{X}_{t_s}\|^p\right]+L^p\mathbb{E}\left[\|\boldsymbol{e}_s\|^p\right] \\ &\qquad+H^p2^{p-1}\left(1+\mathbb{E}\left[\left\|\boldsymbol{X}_s^h\right\|^p\right]\right)h^{p/2}\Big)\,\mathrm{d}u \\ &\leq 3^{p-1}T^p\left(L^pL_p+H^p2^{p-1}(1+M_p)\right)h^{p/2} \\ &\quad+(3T)^{p-1}L^p\sum_{s=0}^{k-1}h\mathbb{E}\left[\|\boldsymbol{e}_s\|^p\right],\end{aligned}$$

其中，最后一个不等式用到了 (I.7) 和 (I.12)。因此，存在常数 $C_{p,\ 1}$，$C_{p,\ 2}>0$，使得

$$\mathbb{E}\left[\max_{0\leq m\leq k}\|\boldsymbol{B}_m\|^p\right]\leq C_{p,1}h^{p/2}+C_{p,2}\sum_{s=0}^{k-1}hy_s, \tag{I.16}$$

其中

$$y_s:=\mathbb{E}\left[\max_{0\leq m\leq s}\|\boldsymbol{e}_m\|^p\right].$$

对于 (I.13) 右侧的第二项，由 BDG不等式，存在 $B'_p > 0$，使得

$$\begin{aligned}\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{G}_m\|^p\right] &\le B'_p\mathbb{E}\left[\left(\sum_{s=0}^{k-1}\int_{t_s}^{t_{s+1}}\left\|\sigma(u,\boldsymbol{X}_u)-\sigma(t_s,\boldsymbol{X}_s^h)\right\|_{\mathrm{F}}^2\,\mathrm{d}u\right)^{p/2}\right]\\ &\le B'_p\mathbb{E}\left[k^{\frac{p}{2}-1}\sum_{s=0}^{k-1}\left(\int_{t_s}^{t_{s+1}}\left\|\sigma(u,\boldsymbol{X}_u)-\sigma(t_s,\boldsymbol{X}_s^h)\right\|_{\mathrm{F}}^2\,\mathrm{d}u\right)^{p/2}\right] \quad \text{(I.17)}\\ &\le B'_pT^{\frac{p}{2}-1}\sum_{s=0}^{k-1}\int_{t_s}^{t_{s+1}}\mathbb{E}\left[\left\|\sigma(u,\boldsymbol{X}_u)-\sigma(t_s,\boldsymbol{X}_s^h)\right\|_{\mathrm{F}}^p\right]\,\mathrm{d}u,\end{aligned}$$

其中，最后两个不等式由 H¨older 不等式和引理 N.2 得出。按照 (I.15) 中相同的分解，并利用 (I.3)，我们有

$$\begin{aligned}\mathbb{E}\left[\left\|\sigma(u,\boldsymbol{X}_u)-\sigma(t_s,\boldsymbol{X}_s^h)\right\|_{\mathrm{F}}^p\right] \le 3^{p-1}\Big(&L^pL_ph^{p/2}+L^p\mathbb{E}\left[\|\boldsymbol{e}_s\|^p\right]\\ &+2^{p-1}H^p(1+M_p)h^{p/2}\Big).\end{aligned}$$

将此界代入 (I.17)，得到

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{G}_m\|^p\right] \le C'_{p,1}h^{p/2}+C'_{p,2}\sum_{s=0}^{k-1}hy_s, \tag{I.18}$$

其中某些常数为 $C'_{p,\ 1}$，$C'_{p,\ 2}>0$。结合 (I.16)、(I.18) 和 (I.13)，我们得到

$$y_k \le C''_{p,1}h^{p/2}+C''_{p,2}\sum_{s=0}^{k-1}hy_s,$$

其中某些常数为 $C''_{p,\ 1}$，$C''_{p,\ 2}>0$，且仅依赖于 p。由离散 Gr¨onwall 不等式，

$$y_N \le C''_{p,1}e^{C''_{p,2}T}h^{p/2},$$

也就是说，

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_{t_k}-\boldsymbol{X}_k^h\right\|^p\right] \le C_ph^{p/2}.$$

最后，对于 $1\le p<2$，由 Lyapunov 不等式（Bogachev，2007，推论 2.12.21），

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_{t_k}-\boldsymbol{X}_k^h\right\|^p\right] \le \mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_{t_k}-\boldsymbol{X}_k^h\right\|^2\right]^{p/2} \le C_2^{p/2}h^{p/2}.$$

■

注 I.1 如果允许 T 在 *(I.1)* 中变化，不难看出，*(I.7)* 和 *(I.8)* 中对 $\boldsymbol{X}_k^h$和 $\boldsymbol{X}_t$的矩进行界定的 M_p 和 M'_p 满足 $M_p, M'_p = \mathcal{O}(\exp(T^p))$。这也意味着 *(I.12)* 中的 $L_p = \mathcal{O}(\exp(T^p))$。因此，最终估计中的常数 C_p 满足 $C_p = \mathcal{O}(\exp(T^p))$，其中 $\mathcal{O}$ 隐藏了 L, H *(I.3)*、p和 $\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right]$。

附录 J. 第 4 节中省略的证明

本节首先证明引理 13，它控制外在离散化与内在离散化之间的强路径偏差；参见 J.1 节。我们还给出定理 7 和推论 8 中省略的证明；参见 J.2 节和 J.3 节。

J.1 引理 13 的证明

在开始证明之前，我们先引入一些记号。

对于任意 y，$u \in \mathbb{R}^n$，定义

$$\Phi_E^h(y,u) := y + hU(y) + P(y)u.$$

对于随机微分方程 (4) 中的 $\mathbb{R}^n$-布朗运动 $\boldsymbol{W}$，令

$$\Delta \boldsymbol{W}_k := \boldsymbol{W}_{t_{k+1}} - \boldsymbol{W}_{t_k} \sim \mathcal{N}(0, hI_n).$$

因此，在 (8) 中，我们可以将 $\sqrt{h}\boldsymbol{\xi} \sim \mathcal{N}(0,\ hI_n)$ 表示为

$$\sqrt{h}\boldsymbol{\xi} = \Delta \boldsymbol{W}_k.$$

于是，(8) 可写为

$$\boldsymbol{Y}_{k+1}^h = \Phi_E^h\left(\boldsymbol{Y}_k^h, \Delta \boldsymbol{W}_k\right). \tag{J.1}$$

类似地，对于任意 $x \in \mathcal{M}$ 和任意 $u \in \mathbb{R}^n$，定义

$$\Phi_G^h(x,u) := \exp_x\left(hV(x) + P(x)u\right),$$

其中 $\exp_x\colon T_x\mathcal{M} \to \mathcal{M}$ 是指数映射。由于 $P(x)$ 是到 $T_x\mathcal{M}$ 的正交投影，且 $\Delta \boldsymbol{W}_k \sim \mathcal{N}(0,\ hI_n)$，随机变量 $P(x)\Delta \boldsymbol{W}_k$ 是 $T_x\mathcal{M}$ 上均值为 0、方差为 h 的高斯分布，其中所采用的度量为 $g(x)$（Muirhead，2009）。因此，我们可以通过以下方式实现 (3) 中的 $\sqrt{h}\boldsymbol{\xi}_x$：

$$\sqrt{h}\boldsymbol{\xi}_x = P(x)\Delta \boldsymbol{W}_k.$$

于是，(3) 可写为

$$\boldsymbol{X}_{k+1}^h = \Phi_G^h\left(\boldsymbol{X}_k^h, \Delta \boldsymbol{W}_k\right). \tag{J.2}$$

不失一般性，假设 $\boldsymbol{X}_0$ 与布朗运动 $\boldsymbol{W}$ 相互独立。

$$\mathcal{F}_t := \sigma\left(\boldsymbol{X}_0, \boldsymbol{W}_s\colon s \le t\right),$$

这同样是一个 $\boldsymbol{W}$-滤过（Karatzas 和 Shreve，1991）。令 $\mathcal{F}_k := \mathcal{F}_{t_k}$，其中 $k = 0,\ \ldots\ldots,\ N$。则 $\boldsymbol{X}_k^h$ 和 $\boldsymbol{Y}_k^h$ 都是 $\mathcal{F}_k$- 可测的。此外，由于 $\Delta \boldsymbol{W}_k$ 与 $\mathcal{F}_k$ 相互独立，因此可得（ Durrett，2019 ）

$$\mathbb{E}\left[f(\boldsymbol{X}_k^h, \Delta \boldsymbol{W}_k) \mid \mathcal{F}_k\right] = \mathbb{E}\left[f(\boldsymbol{X}_k^h, \Delta \boldsymbol{W}_k) \mid \sigma(\boldsymbol{X}_k^h)\right], \tag{J.3}$$

fo任意可积函数 f。将 $\boldsymbol{X}_k^h$ 替换为 $\boldsymbol{Y}_k^h$ 后，同一恒等式仍然成立 .

引理 13 的主要证明策略与 EM 算法（定理 I.1）的证明思路类似。我们首先证明 $p \ge 2$ 的情形，然后应用 Lyapunov 不等式，将结果推广到所有 $1 \le p < \infty$。

证明 [引理 13] 的证明首先，假设 $p \ge 2$，并令 $\boldsymbol{e}_k := \boldsymbol{X}_k^h - \boldsymbol{Y}_k^h$。

- 步骤 1. 鞅分解：使用上述记号和引理 J.1 ，

$$\begin{aligned}\boldsymbol{e}_{k+1} &= \Phi_G^h\left(\boldsymbol{X}_k^h, \Delta\boldsymbol{W}_k\right) - \Phi_E^h\left(\boldsymbol{Y}_k^h, \Delta\boldsymbol{W}_k\right) \\ &= \Phi_E^h\left(\boldsymbol{X}_k^h, \Delta\boldsymbol{W}_k\right) - \Phi_E^h\left(\boldsymbol{Y}_k^h, \Delta\boldsymbol{W}_k\right) + \underbrace{\Phi_G^h\left(\boldsymbol{X}_k^h, \Delta\boldsymbol{W}_k\right) - \Phi_E^h\left(\boldsymbol{X}_k^h, \Delta\boldsymbol{W}_k\right)}_{=\boldsymbol{\Delta}_h(\boldsymbol{X}_k^h)} \\ &= \boldsymbol{e}_k + h\left(U(\boldsymbol{X}_k^h) - U(\boldsymbol{Y}_k^h)\right) + \left(P(\boldsymbol{X}_k^h) - P(\boldsymbol{Y}_k^h)\right)\Delta\boldsymbol{W}_k + \boldsymbol{\Delta}_h(\boldsymbol{X}_k^h).\end{aligned}$$

此外，由于 $\boldsymbol{\Delta}_h$ （引理 J.1）一致有界， $\boldsymbol{\Delta}_h(\boldsymbol{X}_k^h) \in L^1$。因此可以定义

$$\boldsymbol{\mu}_k := \mathbb{E}\left[\boldsymbol{\Delta}_h(\boldsymbol{X}_k^h) \mid \mathcal{F}_k\right], \quad \boldsymbol{m}_k^{(2)} := \boldsymbol{\Delta}_h(\boldsymbol{X}_k^h) - \boldsymbol{\mu}_k.$$

因此,

$$\boldsymbol{e}_{k+1} = \boldsymbol{e}_k + \underbrace{h\left(U(\boldsymbol{X}_k^h) - U(\boldsymbol{Y}_k^h)\right) + \boldsymbol{\mu}_k}_{=:\boldsymbol{a}_k} + \underbrace{\left(P(\boldsymbol{X}_k^h) - P(\boldsymbol{Y}_k^h)\right)\Delta\boldsymbol{W}_k}_{=:\boldsymbol{m}_k^{(1)}} + \boldsymbol{m}_k^{(2)}.$$

Let

$$\boldsymbol{A}_k := \sum_{j=0}^{k-1}\boldsymbol{a}_j, \quad \boldsymbol{M}_k^{(1)} := \sum_{j=0}^{k-1}\boldsymbol{m}_j^{(1)}, \quad \boldsymbol{M}_k^{(2)} := \sum_{j=0}^{k-1}\boldsymbol{m}_j^{(2)}.$$

注意，根据定义， $\boldsymbol{M}_{(1)k}$和 $\boldsymbol{M}_k^{(2)}$都是 $\mathcal{F}_k$-鞅。由于 $\boldsymbol{e}_0 = 0$，我们有

$$\boldsymbol{e}_k = \boldsymbol{A}_k + \boldsymbol{M}_k, \quad \boldsymbol{M}_k := \boldsymbol{M}_k^{(1)} + \boldsymbol{M}_k^{(2)}.$$

由此可得

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{e}_m\|^p\right] \le 2^{p-1}\left(\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{A}_m\|^p\right] + \mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{M}_m\|^p\right]\right), \tag{J.4}$$

其中

$$S_k := \mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{e}_m\|^p\right] < \infty$$

因为根据定理 I.1， $\mathbb{E}\left[\max_m\left\|\boldsymbol{Y}_m^h\right\|^p\right] < \infty$；根据引理 J.2， $\mathbb{E}\left[\max_m\left\|\boldsymbol{X}_m^h\right\|^p\right] < \infty$。

- 步骤 2。 $\boldsymbol{A}_k$的有界性：我们分别处理 $\boldsymbol{a}_k$ 中的两项。对于 U 项，根据第 4.3 节， U 在 $\mathbb{R}^n$ 上是 L_U-利普希茨连续的，因此

$$\left\|U(\boldsymbol{X}_k^h) - U(\boldsymbol{Y}_k^h)\right\|^p \le L_U^p\|\boldsymbol{e}_k\|^p. \tag{J.5}$$

对于另一项，由引理 J.1 和 (J.3)，存在一致常数 $C_\mu > 0$，使得

$$\|\boldsymbol{\mu}_k\| = \left\|\mathbb{E}\left[\boldsymbol{\Delta}_h(\boldsymbol{X}_k^h) \mid \sigma(\boldsymbol{X}_k^h)\right]\right\| = \left\|\mu_h(\boldsymbol{X}_k^h)\right\| \le C_\mu h^{3/2}. \tag{J.6}$$

结合 (J.5) 和 (J.6),

$$\begin{aligned}\max_{0\le m\le k}\|\boldsymbol{A}_m\|^p &\le \max_{0\le m\le k}\left(\sum_{j=0}^{m-1} h\left\|U(\boldsymbol{X}_j^h)-U(\boldsymbol{Y}_j^h)\right\|+\sum_{j=0}^{m-1}\|\boldsymbol{\mu}_j\|\right)^p\\ &\le 2^{p-1}\left(\sum_{j=0}^{k-1} h\left\|U(\boldsymbol{X}_j^h)-U(\boldsymbol{Y}_j^h)\right\|\right)^p+2^{p-1}\left(\sum_{j=0}^{k-1}\|\boldsymbol{\mu}_j\|\right)^p\\ &\le 2^{p-1}k^{p-1}h^p\sum_{j=0}^{k-1}\left\|U(\boldsymbol{X}_j^h)-U(\boldsymbol{Y}_j^h)\right\|^p+2^{p-1}\left(\sum_{j=0}^{N-1}\|\boldsymbol{\mu}_j\|\right)^p\\ &\le (2T)^{p-1}L_U^p\sum_{j=0}^{k-1} h\|\boldsymbol{e}_j\|^p+2^{p-1}C_\mu^p T^p h^{p/2}.\end{aligned}$$

由此可得

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{A}_m\|^p\right]\le (2T)^{p-1}L_U^p\sum_{j=0}^{k-1} hS_j+2^{p-1}C_\mu^p T^p h^{p/2}, \tag{J.7}$$

其中我们使用了 $\mathbb{E}\left[\|\boldsymbol{e}_j\|^p\right]\le\mathbb{E}\left[\max_{0\le m\le j}\|\boldsymbol{e}_m\|^p\right]$。

- 步骤 3. $\boldsymbol{M}_k$的有界性：由定义,

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{M}_m\|^p\right]\le 2^{p-1}\left(\mathbb{E}\left[\max_{0\le m\le k}\left\|\boldsymbol{M}_m^{(1)}\right\|^p\right]+\mathbb{E}\left[\max_{0\le m\le k}\left\|\boldsymbol{M}_m^{(2)}\right\|^p\right]\right). \tag{J.8}$$

对于 $\boldsymbol{M}^{(1)}$，类似于欧几里得 EM 算法的证明（定理 I.1），BDG不等式表明

$$\mathbb{E}\left[\max_{0\le m\le k}\left\|\boldsymbol{M}_m^{(1)}\right\|^p\right]\le C_p^{(1)}\mathbb{E}\left[\left(\sum_{j=0}^{k-1} h\left\|P(\boldsymbol{X}_k^h)-P(\boldsymbol{Y}_k^h)\right\|_{\mathrm{F}}^2\right)^{p/2}\right].$$

如第 4.3 节所示， P 在 $\mathbb{R}^n$ 上被延拓为一个 L_P-Lipschitz映射。将其与关于 $p\ge 2$ 的引理 N.2相结合,

$$\begin{aligned}\left(\sum_{j=0}^{k-1} h\left\|P(\boldsymbol{X}_k^h)-P(\boldsymbol{Y}_k^h)\right\|_{\mathrm{F}}^2\right)^{p/2} &\le t_k^{\frac{p}{2}-1}\sum_{j=0}^{k-1} h\left\|P(\boldsymbol{X}_k^h)-P(\boldsymbol{Y}_k^h)\right\|_{\mathrm{F}}^p\\ &\le T^{\frac{p}{2}-1}L_P^p\sum_{j=0}^{k-1} h\|\boldsymbol{e}_j\|^p.\end{aligned}$$

因此,

$$\mathbb{E}\left[\max_{0\le m\le k}\left\|\boldsymbol{M}_m^{(1)}\right\|^p\right]\le C_p^{(1)}T^{\frac{p}{2}-1}L_P^p\sum_{j=0}^{k-1} hS_j. \tag{J.9}$$

接下来，对于 $\boldsymbol{M}^{(2)}$，由离散时间（向量值）BDG不等式（Zorin-Kranich，2022，推论 3.2 ），

$$\mathbb{E}\left[\max_{0\le m\le k}\left\|\boldsymbol{M}_m^{(2)}\right\|^p\right] \le C_p^{(2)}\mathbb{E}\left[[\boldsymbol{M}^{(2)}]_k^{p/2}\right],$$

其中

$$[\boldsymbol{M}^{(2)}]_k = \sum_{j=0}^{k-1}\left\|\boldsymbol{m}_j^{(2)}\right\|^2.$$

因此，由引理 N.2 可得 $p \ge 2$,

$$\mathbb{E}\left[\max_{0\le m\le k}\left\|\boldsymbol{M}_m^{(2)}\right\|^p\right] \le C_p^{(2)}k^{\frac{p}{2}-1}\sum_{j=0}^{k-1}\mathbb{E}\left[\left\|\boldsymbol{m}_j^{(2)}\right\|^p\right].$$

此外，由 (J.3) 可知， $\boldsymbol{m}_k^{(2)} = \widetilde{\boldsymbol{\Delta}}_h(\boldsymbol{X}_k^h)$ 如引理 J.1 所示，这也意味着

$$\mathbb{E}\left[\left\|\boldsymbol{m}_k^{(2)}\right\|^p\right] = \mathbb{E}\left[\left\|\widetilde{\boldsymbol{\Delta}}_h(\boldsymbol{X}_k^h)\right\|^p\right] \le C_\delta h^p.$$

因此,

$$\mathbb{E}\left[\max_{0\le m\le k}\left\|\boldsymbol{M}_m^{(2)}\right\|^p\right] \le C_p^{(2)}k^{\frac{p}{2}-1}\sum_{j=0}^{k-1}C_\delta h^p \le C_p^{(2)}C_\delta T^{p/2}h^{p/2}. \tag{J.10}$$

最后，结合 (J.9)、(J.10) 和 (J.8)，可得

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{M}_m\|^p\right] \le C_p^{(1)}T^{\frac{p}{2}-1}L_P^p\sum_{j=0}^{k-1}hS_j + C_p^{(2)}C_\delta T^{p/2}h^{p/2}. \tag{J.11}$$

- 步骤 4. 最终界：结合 (J.7)、(J.11) 和 (J.4)，存在常数 $C_{1,p}$ 和 $C_{2,p}$ ，使得

$$S_k \le C_{1,p}T^p h^{p/2} + C_{2,p}T^{p-1}\sum_{j=0}^{k-1}hS_j,$$

其中我们假设 $T > 1$，因此 $T^{p/2} \le T^p$。然后，根据离散 Gr¨onwall 不等式（ Thom´ee，2007，引理 10.5），

$$S_N \le C_{1,p}T^p\exp\left(C_{2,p}T^p\right)h^{p/2}.$$

因此，对于 $G_p(T) = C_{1,\ p}T^p\exp\left(C_{2,\ p}T^p\right) = \mathcal{O}(\exp(T^p))$，我们有

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h - \boldsymbol{Y}_k^h\right\|^p\right] \le G_p(T)h^{p/2}.$$

最后，对于 $1 \le p < \infty$，根据 Lyapunov 不等式（Bogachev，2007，推论 2.12.21），

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h - \boldsymbol{Y}_k^h\right\|^p\right] \le \mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h - \boldsymbol{Y}_k^h\right\|^2\right]^{p/2} \le \mathcal{O}(\exp(T^p))h^{p/2}.$$

■

引理 J.1 假设 $\mathcal{M}$ 满足假设 *I* 和假设 *II*，且 $V \in \Gamma(T\mathcal{M})$ 满足假设 *III*。令 $1 \le p < \infty$。使用上述记号，对于任意$x \in \mathcal{M}$，定义

$$\boldsymbol{\Delta}_h(x) := \Phi_G^h\left(x, \Delta \boldsymbol{W}_k\right) - \Phi_E^h\left(x, \Delta \boldsymbol{W}_k\right), \quad \mu_h(x) := \mathbb{E}\left[\boldsymbol{\Delta}_h(x)\right].$$

若 $h \le 1$，则下列陈述成立。

(i) 存在常数 $C_\mu > 0$ 使得

$$\sup_{x \in \mathcal{M}} \|\mu_h(x)\| \le C_\mu h^{3/2}.$$

(ii) 定义中心化误差

$$\widetilde{\boldsymbol{\Delta}}_h(x) := \boldsymbol{\Delta}_h(x) - \mu_h(x).$$

则存在常数 $C_\delta > 0$ 使得

$$\sup_{x \in \mathcal{M}} \mathbb{E}\left[\left\|\widetilde{\boldsymbol{\Delta}}_h(x)\right\|^p\right] \le C_\delta h^p.$$

证明 固定 $x \in \mathcal{M}$。首先，令

$$\boldsymbol{v} = hV(x) + P(x)\Delta \boldsymbol{W}_k = hV(x) + \sqrt{h}\boldsymbol{\xi}_x \in T_x\mathcal{M},$$

其中 $\boldsymbol{\xi}_x$ 是 $T_x\mathcal{M}$ 上的标准高斯分布，满足 $P(x)\Delta \boldsymbol{W}_k = \sqrt{h}\overline{\boldsymbol{\xi}}_x$。由假设 III,

$$C_V := \sup_{x \in \mathcal{M}} \|V(x)\| < \infty. \tag{J.12}$$

因此,

$$\|\boldsymbol{v}\| \le h\|V(x)\| + \|P(x)\Delta \boldsymbol{W}_k\| \le C_V h + \sqrt{h}\|\boldsymbol{\xi}_x\|.$$

因此，对于任意 $1 \le q < \infty$,

$$\mathbb{E}\left[\|\boldsymbol{v}\|^q\right] \le 2^{q-1}\left(C_V^q h^q + h^{q/2}\mathbb{E}\left[\|\boldsymbol{\xi}_x\|^q\right]\right).$$

由于 $\boldsymbol{\xi}_x$ 是标准高斯分布，引理 N.1 给出 $\mathbb{E}[\|\boldsymbol{\xi}_x\|^q] \le \mathcal{O}(m^{q/2})$，其中 $m = \dim \mathcal{M}$。结合此界与 $h \le 1$，存在一个与 x 无关的常数 $C_v > 0$，使得

$$\mathbb{E}\left[\|\boldsymbol{v}\|^q\right] \le C_v h^{q/2}. \tag{J.13}$$

对于此 $\boldsymbol{v}$，引理 C.4 给出

$$\Phi_G^h\left(x, \Delta \boldsymbol{W}_k\right) = x + \boldsymbol{v} + \frac{1}{2}\mathrm{II}_x(\boldsymbol{v}, \boldsymbol{v}) + R_3(x, \boldsymbol{v}),$$

从而

$$\boldsymbol{\Delta}_h(x) = \Phi_G^h\left(x, \Delta \boldsymbol{W}_k\right) - \Phi_E^h\left(x, \Delta \boldsymbol{W}_k\right) = \frac{1}{2}\mathrm{II}_x(\boldsymbol{v}, \boldsymbol{v}) - hA(x) + R_3(x, \boldsymbol{v}). \tag{J.14}$$

现在证明这两个命题。

(i) 根据 II_x 的双线性和对称性,

$$\mathrm{II}_x(\boldsymbol{v},\boldsymbol{v}) = h\mathrm{II}_x\left(\boldsymbol{\xi}_x,\boldsymbol{\xi}_x\right) + 2h^{3/2}\mathrm{II}_x\left(\boldsymbol{\xi}_x, V(x)\right) + h^2\mathrm{II}_x\left(V(x),V(x)\right).$$

由于引理 G.3 蕴含

$$\mathbb{E}\left[\mathrm{II}_x\left(\boldsymbol{\xi}_x,\boldsymbol{\xi}_x\right)\right] = 2A(x),$$

由 (J.14) 可得

$$\mu_h(x) = \mathbb{E}\left[\boldsymbol{\Delta}_h(x)\right] = \frac{1}{2}h^2\mathrm{II}_x\left(V(x),V(x)\right) + \mathbb{E}\left[R_3(x,\boldsymbol{v})\right].$$

注意 $\mathbb{E}[\mathrm{II}_x(\boldsymbol{\xi}_x, V(x))] = 0$ 因为 $\boldsymbol{\xi}_x$ 是中心化的，且 $\mathrm{II}_x(\,\cdot\,, V(x))$ 是线性的。根据假设 I,

$$\kappa := \sup_{x\in\mathcal{M}} \|\mathrm{II}_x\|_{\mathrm{op}} < \infty, \tag{J.15}$$

且引理 C.4 给出

$$\sup_{x\in\mathcal{M}} \|R_3(x,\boldsymbol{v})\| \le C_R\|\boldsymbol{v}\|^3. \tag{J.16}$$

因此，利用 (J.12)、(J.13)、(J.15) 和 (J.16),

$$\begin{aligned}\|\mu_h(x)\| &\le \frac{1}{2}h^2\|\mathrm{II}_x\left(V(x),V(x)\right)\| + \mathbb{E}\left[\|R_3(x,\boldsymbol{v})\|\right]\\ &\le \frac{1}{2}\kappa C_V^2 h^2 + C_R\mathbb{E}\left[\|\boldsymbol{v}\|^3\right]\\ &\le \frac{1}{2}\kappa C_V^2 h^2 + C_R C_v h^{3/2}.\end{aligned}$$

由于 $h \le 1$，存在一个与 x 无关的常数 $C_\mu > 0$，使得

$$\|\mu_h(x)\| \le C_\mu h^{3/2} \quad \Rightarrow \quad \sup_{x\in\mathcal{M}}\|\mu_h(x)\| \le C_\mu h^{3/2} \tag{J.17}$$

(ii) 利用 (J.14),

$$\begin{aligned}\|\boldsymbol{\Delta}_h(x)\|^p &\le 3^{p-1}\left(\frac{1}{2^p}\|\mathrm{II}_x(\boldsymbol{v},\boldsymbol{v})\|^p + h^p\|A(x)\|^p + \|R_3(x,\boldsymbol{v})\|^p\right)\\ &\le 3^{p-1}\left(\frac{1}{2^p}\kappa^p\|\boldsymbol{v}\|^{2p} + C_A^p h^p + C_R^p\|\boldsymbol{v}\|^{3p}\right),\end{aligned}$$

其中用到了 (J.15)、(J.16) 以及由推论 G.1 得到的 $C_A := \sup_{\mathcal{M}}\|A(x)\| < \infty$。取期望，并对 $q = 2p$ 和 $q = 3p$ 应用 (J.13)，得到

$$\mathbb{E}\left[\|\boldsymbol{\Delta}_h(x)\|^p\right] \le \left(\frac{1}{2^p}\kappa^p C_v h^p + C_A^p h^p + C_R^p C_v h^{3p/2}\right) \le Ch^p, \tag{J.18}$$

其中某个 $C > 0$ 与 x 无关。结合 (J.17) 和 (J.18),

$$\mathbb{E}\left[\left\|\widetilde{\boldsymbol{\Delta}}_h(x)\right\|^p\right] \le 2^{p-1}\left(\mathbb{E}\left[\|\boldsymbol{\Delta}_h(x)\|^p\right] + \|\mu_h(x)\|^p\right) \le 2^{p-1}\left(Ch^p + C_\mu^p h^{3p/2}\right).$$

由于 $h \le 1$，存在一个常数 $C_\delta > 0$ 使得

$$\sup_{x\in\mathcal{M}}\mathbb{E}\left[\left\|\widetilde{\boldsymbol{\Delta}}_h(x)\right\|^p\right] \le C_\delta h^p.$$

■

引理 J.2 假设 $\mathcal{M}$ 满足假设 *I* 和假设 *II*，且 $V \in \Gamma(T\mathcal{M})$ 满足假设 *III*。令 $1 \leq p < \infty$。对于由 *(3)* 构造的 $\boldsymbol{X}_k^h$，其中$\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right] < \infty$,

$$\mathbb{E}\left[\max_{0\leq k\leq N}\left\|\boldsymbol{X}_k^h\right\|^p\right] \leq \mathcal{O}(T^p) < \infty.$$

证明我们首先假设 $p \geq 2$，并采用以下步骤证明矩有界性。

- 步骤 1. 鞅分解：令

$$\boldsymbol{v}_k = hV(\boldsymbol{X}_k^h) + P(\boldsymbol{X}_k^h)\Delta\boldsymbol{W}_k,$$

它是 $\mathcal{F}_{k+1}$-可测的。于是，根据泰勒公式（引理 C.4），

$$\boldsymbol{X}_{k+1}^h = \boldsymbol{X}_k^h + \underbrace{\boldsymbol{v}_k + \frac{1}{2}\mathrm{II}_{\boldsymbol{X}_k^h}(\boldsymbol{v}_k,\boldsymbol{v}_k) + R_3(\boldsymbol{X}_k^h,\boldsymbol{v}_k)}_{=:\boldsymbol{u}_k^X}.$$

根据 V,II 和 R_3, $\boldsymbol{u}_k^X \in L^1$ 的一致有界性（或由下文的 (J.26) 可得），因此可以定义

$$\boldsymbol{a}_k^X := \mathbb{E}\left[\boldsymbol{u}_k^X \mid \mathcal{F}_k\right], \quad \boldsymbol{m}_k^X := \boldsymbol{u}_k^X - \boldsymbol{a}_k^X.$$

Let

$$\boldsymbol{A}_k^X := \sum_{j=0}^{k-1}\boldsymbol{a}_j^X, \quad \boldsymbol{M}_k^X := \sum_{j=0}^{k-1}\boldsymbol{m}_j^X,$$

其中 $\boldsymbol{A}_0^X = \boldsymbol{M}_0^X = 0$。于是

$$\boldsymbol{X}_k^h = \boldsymbol{X}_0 + \boldsymbol{A}_k^X + \boldsymbol{M}_k^X.$$

此外，显然 $\boldsymbol{M}_k^X$ 是一个 $\mathcal{F}_k$-鞅，因为 $\boldsymbol{m}_k^X$ 是 $\mathcal{F}_{k+1}$-可测的，且 $\mathbb{E}\left[\boldsymbol{m}_k^X \mid \mathcal{F}_k\right] = 0$。

于是我们有

$$\begin{aligned}\mathbb{E}\left[\max_{0\leq k\leq N}\left\|\boldsymbol{X}_k^h\right\|^p\right] \leq 3^{p-1}\Big(&\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right] + \mathbb{E}\left[\max_{0\leq k\leq N}\left\|\boldsymbol{A}_k^X\right\|^p\right] \\ &+ \mathbb{E}\left[\max_{0\leq k\leq N}\left\|\boldsymbol{M}_k^X\right\|^p\right]\Big).\end{aligned} \tag{J.19}$$

- 步骤 2。 $\boldsymbol{A}_k^X$的有界性：由假设 I 和引理 C.4 可得,

$$\kappa := \sup_{x\in\mathcal{M}}\|\mathrm{II}_x\|_{\mathrm{op}} < \infty, \quad \sup_{x\in\mathcal{M}}\|R_3(x,\boldsymbol{v})\| \leq C_R\|\boldsymbol{v}\|^3. \tag{J.20}$$

注意，由于 $\boldsymbol{X}_k^h$是 $\mathcal{F}_k$-可测的，且 $\Delta \boldsymbol{W}_k$与 $\mathcal{F}_k$相互独立，因此 $\mathbb{E}\left[\boldsymbol{v}_k \mid \mathcal{F}_k\right] = hV(\boldsymbol{X}_k^h)$。

$$\begin{aligned}\left\|\boldsymbol{a}_k^X\right\| \leq \left\|\mathbb{E}\left[\boldsymbol{u}_k^X \mid \mathcal{F}_k\right]\right\| \leq \left\|\mathbb{E}\left[\boldsymbol{v}_k \mid \mathcal{F}_k\right]\right\| + \frac{1}{2}\kappa\mathbb{E}\left[\left\|\boldsymbol{v}_k\right\|^2 \mid \mathcal{F}_k\right] \\ + C_R\mathbb{E}\left[\left\|\boldsymbol{v}_k\right\|^3 \mid \mathcal{F}_k\right].\end{aligned} \tag{J.21}$$

此外，如引理 J.1 所示，(J.13) 意味着对于 $1 \leq q < \infty$，存在一个一致常数 C_v，使得

$$\mathbb{E}\left[\left\|\boldsymbol{v}_k\right\|^q \mid \mathcal{F}_k\right] = \mathbb{E}\left[\left\|\boldsymbol{v}_k\right\|^q \mid \sigma(\boldsymbol{X}_k^h)\right] \leq C_v h^{q/2}, \tag{J.22}$$

其中，第一个等号是由 (J.3) 得出的。由于 $C_V := \sup_{x\in\mathcal{M}}\|V(x)\| < \infty$（假设 III），结合 (J.22) 和 (J.21) 可得

$$\left\|\boldsymbol{a}_k^X\right\| \leq C_V h + \frac{1}{2}\kappa C_v h + C_R C_v h^{3/2} \leq C_a h, \tag{J.23}$$

其中 $C_a > 0$ 为某个常数。因此，

$$\mathbb{E}\left[\max_{0\leq k\leq N}\left\|\boldsymbol{A}_k^X\right\|^p\right] \leq \mathbb{E}\left[\max_{0\leq k\leq N} k^{p-1}\sum_{j=0}^{k-1}\left\|\boldsymbol{a}_j^X\right\|^p\right] \leq N^{p-1}\sum_{j=0}^{N-1} C_a^p h^p = C_a^p T^p. \tag{J.24}$$

• 步骤 3. $\boldsymbol{M}_k^X$ 的有界性：由于 $\boldsymbol{M}_k^X$ 是一个 $\mathcal{F}_k$-鞅，离散时间 BDG 不等式（Zorin-Kranich，2022）意味着存在 $C_{M,\ p} > 0$，使得

$$\mathbb{E}\left[\max_{0\leq k\leq N}\left\|\boldsymbol{M}_k^X\right\|^p\right] \leq C_{M,p}\mathbb{E}\left[\left[\boldsymbol{M}^X\right]_N^{p/2}\right],$$

其中

$$\left[\boldsymbol{M}^X\right]_N = \sum_{j=0}^{N-1}\left\|\boldsymbol{m}_j^X\right\|^2.$$

由于 $p \geq 2$，引理 N.2 意味着

$$\mathbb{E}\left[\max_{0\leq k\leq N}\left\|\boldsymbol{M}_k^X\right\|^p\right] \leq C_{M,p}N^{\frac{p}{2}-1}\sum_{j=0}^{N-1}\mathbb{E}\left[\left\|\boldsymbol{m}_j^X\right\|^p\right]. \tag{J.25}$$

注意到 $\boldsymbol{m}_k^X = \boldsymbol{u}_k^X - \boldsymbol{a}_k^X$，因此

$$\mathbb{E}\left[\left\|\boldsymbol{m}_k^X\right\|^p \mid \mathcal{F}_k\right] \leq 2^{p-1}\left(\mathbb{E}\left[\left\|\boldsymbol{a}_k^X\right\|^p \mid \mathcal{F}_k\right] + \mathbb{E}\left[\left\|\boldsymbol{u}_k^X\right\|^p \mid \mathcal{F}_k\right]\right).$$

由 (J.23)，

$$\mathbb{E}\left[\left\|\boldsymbol{a}_k^X\right\|^p \mid \mathcal{F}_k\right] \leq C_a^p h^p.$$

对于 $\boldsymbol{u}_k^X = \boldsymbol{v}_k + \frac{1}{2}\mathrm{II}^{\boldsymbol{X}_k^h}(\boldsymbol{v}_k,\ \boldsymbol{v}_k) + R_3(\boldsymbol{X}_k^h,\ \boldsymbol{v}_k)$，由 (J.20) 和 (J.22)，

$$\begin{aligned}\mathbb{E}\left[\left\|\boldsymbol{u}_k^X\right\|^p \mid \mathcal{F}_k\right] &\le 3^{p-1}\left(\mathbb{E}\left[\|\boldsymbol{v}_k\|^p \mid \mathcal{F}_k\right] + \frac{\kappa^p}{2^p}\mathbb{E}\left[\|\boldsymbol{v}_k\|^{2p} \mid \mathcal{F}_k\right] + C_R^p\mathbb{E}\left[\|\boldsymbol{v}_k\|^{3p} \mid \mathcal{F}_k\right]\right)\\ &\le 3^{p-1}\left(C_v h^{p/2} + \frac{\kappa^p}{2^p}h^p + C_R^p h^{3p/2}\right).\end{aligned} \tag{J.26}$$

因此，

$$\mathbb{E}\left[\left\|\boldsymbol{m}_k^X\right\|^p\right] = \mathbb{E}\left[\mathbb{E}\left[\left\|\boldsymbol{m}_k^X\right\|^p \mid \mathcal{F}_k\right]\right] \le C_\eta h^{p/2}$$

对于某个常数 $C_\eta > 0$。将其与 (J.25) 结合，可得

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{M}_k^X\right\|^p\right] \le C_{M,p}N^{\frac{p}{2}-1}\sum_{j=0}^{N-1} C_\eta h^{p/2} = C_{M,p}C_\eta T^{p/2}. \tag{J.27}$$

- 步骤 4. 最终界：将 (J.24)、(J.27) 和 (J.19) 结合，可得

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h\right\|^p\right] \le 3^{p-1}\left(\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right] + C_a^p T^p + C_{M,p}C_\eta T^{p/2}\right) \le \mathcal{O}(T^p).$$

最后，for $1 \le p < 2$，根据 Lyapunov 不等式（Bogachev，2007，推论 2.12.21），

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h\right\|^p\right] \le \mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h\right\|^2\right]^{p/2} \le \mathcal{O}(T^p).$$

■

J.2 定理 7 的证明

根据我们对嵌入映射 ι的双 Lipschitz 连续性的讨论（附录 F），定理 7 可视为定理 5 的直接推论。

证明 [定理 7] 的证明。由于 $\mathcal{M}$ 满足假设 I，且嵌入映射ι: $\mathcal{M} \hookrightarrow \mathbb{R}^n$ 是全局双 Lipschitz 连续的，定理 F.2 表明 $\mathcal{M}$ 也满足假设 II。因此，由定理 5 可得

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h - \boldsymbol{X}_{t_k}\right\|^p\right] \le C_p(T)h^{p/2},$$

其中 $C_p(T) = \mathcal{O}(\exp(T^p))$。

另一方面，由于

$$L_\iota := \sup_{x\ne y\in\mathcal{M}}\frac{d_{\mathcal{M}}(x,y)}{\|x-y\|} < \infty,$$

我们有

$$d_{\mathcal{M}}\left(\boldsymbol{X}_k^h, \boldsymbol{X}_{t_k}\right) \le L_\iota\left\|\boldsymbol{X}_k^h - \boldsymbol{X}_{t_k}\right\|.$$

因此，

$$\mathbb{E}\left[\max_{0\le k\le N} d_{\mathcal{M}}\left(\boldsymbol{X}_k^h, \boldsymbol{X}_{t_k}\right)^p\right] \le L_\iota^p\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h - \boldsymbol{X}_{t_k}\right\|^p\right] \le \mathcal{O}(\exp(T^p))h^{p/2}.$$

■

J.3 推论 8 的证明

证明 [推论 8]的证明令 $\mathcal{M}$ 为一个紧黎曼流形。根据 Nash 的嵌入定理（Nash，1956），存在一个等距 C^∞ 嵌入 $\iota\colon\ \mathcal{M} \hookrightarrow \mathbb{R}^n$，其中$n$。如附录 E.1 所述，$\mathcal{M}$ 满足假设 I 和假设 II。此外，根据定理 F.1，嵌入 ι是全局双 Lipschitz 连续的。

对于 $V \in \Gamma(T\mathcal{M})$，$\mathcal{M}$ 的紧致性意味着 V 满足假设 III。条件 $\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right] < \infty$ 也由紧致性得到。

于是定理 7 意味着

$$\mathbb{E}\left[\max_{0\le k\le N} d_{\mathcal{M}}\left(\boldsymbol{X}_k^h, \boldsymbol{X}_{t_k}\right)^p\right] \le \mathcal{O}(\exp(T^p))h^{p/2}.$$

■

附录 K. 非齐次时间情形

定理 5 的结果可扩展到非齐次的 $\mathcal{M}$ 值随机微分方程

$$\mathrm{d}\boldsymbol{X}_t = V(t, \boldsymbol{X}_t)\ \mathrm{d}t + g(t)\ \mathrm{d}\boldsymbol{B}_t^{\mathcal{M}}, \tag{K.1}$$

其中 $g\colon\ [0,T] \to \mathbb{R}$ 连续，并且 $V(t,\ x) \in \Gamma(T\mathcal{M})$ 关于 t 连续。

假设 $\mathcal{M} \subset \mathbb{R}^n$ 满足假设 I 和假设 II，且对所有 $t \in [0,T]$，向量场 $V(t,\cdot)$ 满足假设 III。此外，假设 $V(t,x)$ 和$g(t)$ 关于 t 是 $1/2$-H¨older 连续的，即，

$$\|V(t,x) - V(s,x)\| + |g(t) - g(s)| \le H\,|t-s|^{1/2}.$$

~那么，如第 4.3 节（附录 H）所示，映射 $V(t,y) := \chi(y)V(t,\pi(y))\colon\ [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ 关于 x 是全局 Lipschitz 连续的，并且在 t 上是 $1/2$-H¨older 连续的。与前面一样，我们记作 $\widetilde{V}$，并考虑扩展的随机微分方程

$$\mathrm{d}\boldsymbol{X}_t = U(t, \boldsymbol{X}_t)\ \mathrm{d}t + g(t)P(\boldsymbol{X}_t)\ \mathrm{d}\boldsymbol{W}_t, \tag{K.2}$$

其中 $U(t,x) = V(t,x) + g(t)^2 A(x)$。对于随机微分方程 (K.1)，注意到

$$g(t)\ \mathrm{d}\boldsymbol{B}_t^{\mathcal{M}} = g(t)P(\boldsymbol{X}_t) \circ \mathrm{d}\boldsymbol{W}_t = g(t)^2 A(\boldsymbol{X}_t)\ \mathrm{d}t + g(t)P(\boldsymbol{X}_t)\ \mathrm{d}\boldsymbol{W}_t,$$

根据 Itô 公式（Le Gall, 2016），因此，当 $\boldsymbol{X}_0 \in \mathcal{M}$ 时，随机微分方程 (K.1) 和随机微分方程 (K.2) 存在相同的解（ Hsu, 2002 ）。

注 K.1 与欧几里得情形不同，我们不考虑形如 $\mathrm{d}\boldsymbol{X}_t = V(t, \boldsymbol{X}_t)\,\mathrm{d}t + \sigma(t, \boldsymbol{X}_t)\,\mathrm{d}\boldsymbol{B}_t^{\mathcal{M}}$ 的更一般的扩散系数。一般而言，将$\sigma(t, \boldsymbol{X}_t)\ \mathrm{d}\boldsymbol{B}_t^{\mathcal{M}}$改写为关于 $\mathrm{d}\boldsymbol{W}_t$ 的形式会引入一个额外的漂移项，而该项不像 *(K.2)* 中的项 $g(t)^2 A(x)\ \mathrm{d}t$ 那样简单。

注意，随机微分方程 (K.2) 是一个定义良好的欧几里得随机微分方程，其系数关于 x 是全局 Lipschitz 的，并且关于 t 是 $1/2$-H¨older 连续的。因此，对于

$$\boldsymbol{Y}_{k+1}^h = \boldsymbol{Y}_k^h + hU(t_k, \boldsymbol{Y}_k^h) + g(t_k)P(\boldsymbol{Y}_k^h)\Delta\boldsymbol{W}_k, \quad k = 0, \ldots, N-1,$$

定理 I.1 仍然意味着

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_{t_k}-\boldsymbol{Y}_k^h\right\|^p\right]\le E_p(T)h^{p/2},$$

其中 $E_p(T)=\mathcal{O}(\exp(T^p))$。

因此，为了得到与定理 5 相同的结论，尚需证明引理 13。首先，我们需要 $\boldsymbol{X}_k^h$ 的矩有界性，其定义为

$$\boldsymbol{X}_{k+1}^h=\exp_{\boldsymbol{X}_k^h}\left(hV(t_k,\boldsymbol{X}_k^h)+g(t_k)P(\boldsymbol{X}_k^h)\Delta\boldsymbol{W}_k\right),\quad k=0,\dots,N-1,$$

如引理 J.2 中所述。在条件 $\sup_{t\in[0,T]}g(t)<\infty$ 下，同样的论证在此适用，该条件由 g 在 $[0,T]$上的连续性得出。最后，在 $\boldsymbol{X}_k^h$与 $\boldsymbol{Y}_k^h$的比较中不存在时间平移，因此证明的余项只需因因子 $g(t_k)$ 作微小改动即可沿用，而该因子已经一致有界。

附录 L. 放宽假设 III

在定理 5 的证明中，假设 III 在两处被使用。(i) 它用于构造 $\tilde{}$一个全局 Lipschitz 连续延拓 V（引理 9）。该延拓既用于控制 (9) 中 $\boldsymbol{Y}_k^h$与 $\boldsymbol{X}_{t_k}$ 之间的外在 EM 误差，也用于界定引理 13 中 $\boldsymbol{Y}_k^h$与 $\boldsymbol{X}_k^h$ 之间的差异。(ii) 它还用于控制 $\boldsymbol{X}_k^h$ 的矩有界性，而这是引理 13 证明中的关键要素。

基于这两种作用，在处理 $\mathcal{M}\subset\mathbb{R}^n$ 时，我们可以用以下假设替代假设 III。

假设 III† 设 $\mathcal{M}\subset\mathbb{R}^n$ 为黎曼流形，并令 $V\in\Gamma(T\mathcal{M})$。假设

$$\sup_{x_1\neq x_2\in\mathcal{M}}\frac{\|V(x_1)-V(x_2)\|}{\|x_1-x_2\|}<\infty,\quad \sup_{x\in\mathcal{M}}\frac{\|V(x)\|}{1+\|x\|^{1/3}}<\infty.$$

若 $V\in\Gamma(T\mathcal{M})$ 满足假设 III†，则由引理 L.4 可知，V 存在一个到 $\mathbb{R}^n$的全局 Lipschitz 延拓，我们仍将其记为 $V\colon\ \mathbb{R}^n\to\mathbb{R}^n$。此外，该延拓满足次线性增长界：存在常数 $K>0$，使得

$$\|V(y)\|\le K\left(1+\|y\|^{1/3}\right),\quad\forall\ y\in\mathbb{R}^n. \tag{L.1}$$

由于延拓的欧氏随机微分方程的系数仍为全局 Lipschitz，定理 I.1 仍然给出

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_{t_k}-\boldsymbol{Y}_k^h\right\|^p\right]\le E_p(T)h^{p/2},$$

其中 $E_p(T)=\mathcal{O}(\exp(T^p))$。

因此，将假设 III 替换为假设 III†后，为了证明定理 5，还需建立引理 13 在放宽条件下的对应结论，即引理 L.1。引理 L.1 与引理 13 的证明之间的主要区别在于$\boldsymbol{X}_k^h$ 的矩有界性（引理 L.3 与引理 J.2），而这依赖于次线性增长条件（L.1）。一旦建立这一点，其余步骤便与之前相同。为完整起见，我们在下面给出证明。

引理 L.1 假设 $\mathcal{M} \subset \mathbb{R}^n$ 满足假设 *I* 和假设 *II*，且$V \in \Gamma(T\mathcal{M})$ 满足假设 *III*†。令 $1 \le p < \infty$。对于按 *(3)* 构造的 $\boldsymbol{X}_h^k$ 和按 *(8)* 构造的$\boldsymbol{Y}_k^h$，我们有

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h - \boldsymbol{Y}_k^h\right\|^p\right] \le G_p(T)h^{p/2},$$

其中 $G_p(T) = \mathcal{O}(\exp(T^p))$。

证明 证明过程与引理 13 的证明几乎相同，因此我们使用相同的记号，仅强调其中的差异。

对于 $\boldsymbol{\Delta}_h(\boldsymbol{X}_k^h) = \Phi_G^h\left(\boldsymbol{X}_k^h,\ \Delta\boldsymbol{W}_k\right) - \Phi_E^h\left(\boldsymbol{X}_k^h,\ \Delta\boldsymbol{W}_k\right)$，根据引理 L.2 中的 (L.9) 以及 $\boldsymbol{X}_k^h$ 的矩有界性（引理 L.3 ），我们有 $\boldsymbol{\Delta}_h(\boldsymbol{X}_k^h) \in L^1$。因此，我们可以定义

$$\boldsymbol{\mu}_k := \mathbb{E}\left[\boldsymbol{\Delta}_h(\boldsymbol{X}_k^h) \mid \mathcal{F}_k\right], \quad \boldsymbol{m}_k^{(2)} := \boldsymbol{\Delta}_h(\boldsymbol{X}_k^h) - \boldsymbol{\mu}_k,$$

并得到相同的分解 $\boldsymbol{e}_k = \boldsymbol{A}_k + \boldsymbol{M}_k$，其中

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{e}_m\|^p\right] \le 2^{p-1}\left(\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{A}_m\|^p\right] + \mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{M}_m\|^p\right]\right), \tag{L.2}$$

与引理 13 中一样。由定理 I.1 和.....可知 $\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{e}_m\|^p\right]$ 的存在性

$$M_X(T) := \mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}_k^h\right\|^p\right] \le \mathcal{O}(\exp(T^p)) < \infty, \tag{L.3}$$

如引理 L.3 所示。

对于 $\boldsymbol{A}_k$，引理 L.2 给出

$$\|\boldsymbol{\mu}_k\| = \left\|\mathbb{E}\left[\boldsymbol{\Delta}_h(\boldsymbol{X}_k^h) \mid \sigma(\boldsymbol{X}_k^h)\right]\right\| = \left\|\mu_h(\boldsymbol{X}_k^h)\right\| \le C_\mu\left(1 + \left\|\boldsymbol{X}_k^h\right\|\right)h^{3/2}. \tag{L.4}$$

结合 (L.4) 与 (L.3)，按照引理 13 中相同的推理可得

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{A}_m\|^p\right] \le (2T)^{p-1}L_U^p\sum_{j=0}^{k-1}hS_j + 2^{2(p-1)}C_\mu^pT^p\left(1 + M_X(T)\right)h^{p/2}. \tag{L.5}$$

对于 $\boldsymbol{M}_k$，$\boldsymbol{M}_k^{(1)}$的界保持不变。对于 $\boldsymbol{M}_{(2)k}$，由引理 L.2 结合可得 (L.3)

$$\mathbb{E}\left[\left\|\boldsymbol{m}_k^{(2)}\right\|^p\right] = \mathbb{E}\left[\left\|\widetilde{\boldsymbol{\Delta}}_h(\boldsymbol{X}_k^h)\right\|^p\right] \le C_\delta\left(1 + M_X(T)\right)h^p.$$

因此，与引理 13 中一样，

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{M}_m\|^p\right] \le C_p^{(1)}T^{\frac{p}{2}-1}L_P^p\sum_{j=0}^{k-1}hS_j + C_p^{(2)}C_\delta T^{p/2}\left(1 + M_X(T)\right)h^{p/2}. \tag{L.6}$$

结合 (L.5)、(L.6) 和 (L.2)，我们得到

$$S_k \le C_{1,p}\left(1 + M_X(T)\right)T^ph^{p/2} + C_{2,p}T^{p-1}\sum_{j=0}^{k-1}hS_j,$$

由此可得

$$S_N \leq C_{1,p}\left(1+M_X(T)\right)T^p\exp\left(C_{2,p}T^p\right)h^{p/2}.$$

因此,

$$\mathbb{E}\left[\max_{0\leq k\leq N}\left\|\boldsymbol{X}_k^h-\boldsymbol{Y}_k^h\right\|^p\right]\leq\mathcal{O}(\exp(T^p))h^{p/2}.$$

■

引理 L.2 假设 $\mathcal{M}\subset\mathbb{R}^n$ 满足假设 *I* 和假设 *II*，且$V\in\Gamma(T\mathcal{M})$ 满足假设 *III*†。设 $1\leq p<\infty$。采用引理 ***J.1*** 中的记号，可得以下结论。

(i) 存在常数 $C_\mu>0$，使得

$$\|\mu_h(x)\|\leq C_\mu\left(1+\|x\|\right)h^{3/2}.$$

(ii) 存在常数 $C_\delta>0$，使得

$$\mathbb{E}\left[\left\|\widetilde{\boldsymbol{\Delta}}_h(x)\right\|^p\right]\leq C_\delta\left(1+\|x\|^p\right)h^p.$$

证明 固定 $x\in\mathcal{M}$。对于

$$\boldsymbol{v}=hV(x)+P(x)\Delta\boldsymbol{W}_k=hV(x)+\sqrt{h}\boldsymbol{\xi}_x\in T_x\mathcal{M},$$

其中 $\sqrt{h}\boldsymbol{\xi}_x^-=P(x)\Delta\boldsymbol{W}_k$，由 (L.1) 可得,

$$\begin{aligned}\mathbb{E}\left[\|\boldsymbol{v}\|^q\right]&\leq 2^{q-1}\left(h^q\|V(x)\|^q+h^{q/2}\mathbb{E}\left[\|\boldsymbol{\xi}_x\|^q\right]\right)\\&\leq 2^{q-1}\left(h^qK^q2^{q-1}\left(1+\|x\|^{q/3}\right)+h^{q/2}\mathbb{E}\left[\|\boldsymbol{\xi}_x\|^q\right]\right).\end{aligned}$$

因此，根据引理 N.1，对于任意 $1\leq q<\infty$，存在常数 $C_{1,\ v}$，$C_{2,\ v}>0$，使得

$$\mathbb{E}\left[\|\boldsymbol{v}\|^q\right]\leq C_{1,v}\left(1+\|x\|^{q/3}\right)h^q+C_{2,v}h^{q/2}. \tag{L.7}$$

按照引理 J.1 的证明，我们有

$$\|\mu_h(x)\|\leq\frac{1}{2}h^2\|\mathrm{II}_x\left(V(x),V(x)\right)\|+\mathbb{E}\left[\|R_3(x,\boldsymbol{v})\|\right]\leq\frac{1}{2}h^2\kappa\|V(x)\|^2+C_R\mathbb{E}\left[\|\boldsymbol{v}\|^3\right].$$

将其与 (L.1) 和 (L.7) 结合,

$$\|\mu_h(x)\|\leq h^2\kappa\left(1+\|x\|^{2/3}\right)+C_R\left(C_{1,v}\left(1+\|x\|\right)h^3+C_{2,v}h^{3/2}\right)$$

注意，$\|x\|^{2/3}\leq\|x\|+1$，并且我们始终假设 $h\leq 1$。因此，存在一个常数$C_\mu>0$，使得

$$\|\mu_h(x)\|\leq C_\mu\left(1+\|x\|\right)h^{3/2}. \tag{L.8}$$

接下来，对于 $\boldsymbol{\Delta}_h(x)$，如引理 J.1 中所述，

$$\|\boldsymbol{\Delta}_h(x)\|^p \leq 3^{p-1}\left(\frac{1}{2^p}\kappa^p\|\boldsymbol{v}\|^{2p} + C_A^p h^p + C_R^p\|\boldsymbol{v}\|^{3p}\right).$$

因此，(L.7) 意味着存在常数 C_1，$C_2 > 0$ 使得

$$\mathbb{E}\left[\|\boldsymbol{\Delta}_h(x)\|^p\right] \leq C_1\left(1 + \|x\|^p\right)h^{2p} + C_2 h^p, \tag{L.9}$$

其中我们还使用了 $\|x\|^{2/3} \leq \|x\| + 1$。然后，结合 (L.9) 和 (L.8)，可得

$$\begin{aligned}\mathbb{E}\left[\left\|\widetilde{\boldsymbol{\Delta}}_h(x)\right\|^p\right] &\leq 2^{p-1}\left(\mathbb{E}\left[\|\boldsymbol{\Delta}_h(x)\|^p\right] + \|\mu_h(x)\|^p\right)\\ &\leq C_\delta\left(1 + \|x\|^p\right)h^p,\end{aligned}$$

对于某些常数 C_δ。 ■

引理 L.3 假设 $\mathcal{M} \subset \mathbb{R}^n$ 满足假设 *I* 和假设 *II*，且$V \in \Gamma(T\mathcal{M})$ 满足假设 *III*†。设 $1 \leq p < \infty$。对于由 *(3)* 构造的 $\boldsymbol{X}_k^h$，其中$\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right] < \infty$，

$$\mathbb{E}\left[\max_{0\leq k\leq N}\left\|\boldsymbol{X}_k^h\right\|^p\right] \leq \mathcal{O}(\exp(T^p)) < \infty.$$

注 L.1 与引理 *J.2* 的证明不同，我们采用与定理 *I.1* 的证明中相同的局部化论证，以证明所需的可积性。随后，我们应用离散 *Gr¨onwall* 不等式。

证明 我们首先处理 $p \geq 2$ 的情形。

- 步骤 1. 局部化：对于任意 $\ell \in \mathbb{N}$，定义停时

$$\boldsymbol{N}_\ell := \min\left\{k > 0 \colon \left\|\boldsymbol{X}_k^h\right\| > \ell\right\},$$

它是一个 $\mathcal{F}_k$-停时，并考虑停止过程 $\boldsymbol{X}_k^{(\ell)} := \boldsymbol{X}_{k\wedge\boldsymbol{N}_\ell}^h$。定义

$$\boldsymbol{u}_k^{(\ell)} := \boldsymbol{X}_{k+1}^{(\ell)} - \boldsymbol{X}_k^{(\ell)} = \mathbb{I}_{\{k<\boldsymbol{N}_\ell\}}\left(\boldsymbol{X}_{k+1}^h - \boldsymbol{X}_k^h\right).$$

使用引理 J.2 中的相同记号，对于 $\boldsymbol{v}_k = hV(\boldsymbol{X}_k^h) + P(\boldsymbol{X}_k^h)\Delta\boldsymbol{W}_k$，令

$$\boldsymbol{u}_k^X = \boldsymbol{X}_{k+1}^h - \boldsymbol{X}_k^h = \boldsymbol{v}_k + \frac{1}{2}\mathrm{II}_{\boldsymbol{X}_k^h}(\boldsymbol{v}_k, \boldsymbol{v}_k) + R_3(\boldsymbol{X}_k^h, \boldsymbol{v}_k).$$

Then

$$\boldsymbol{u}_k^{(\ell)} = \mathbb{I}_{\{k<\boldsymbol{N}_\ell\}}\boldsymbol{u}_k^X.$$

注意，在 $\{k < \boldsymbol{N}_\ell\}$上，我们有 $\left\|V(\boldsymbol{X}_k^h)\right\| \leq K(1 + \ell^{1/3}) < \infty$。因此，(J.26)意味着 $\boldsymbol{u}_k^{(\ell)} \in L^p \subset L^1$。于是，我们可以定义

$$\boldsymbol{a}_k^{(\ell)} := \mathbb{E}\left[\boldsymbol{u}_k^{(\ell)} \mid \mathcal{F}_k\right], \quad \boldsymbol{m}_k^{(\ell)} := \boldsymbol{u}_k^{(\ell)} - \boldsymbol{a}_k^{(\ell)}.$$

Then

$$\boldsymbol{X}_k^{(\ell)} = \boldsymbol{X}_0 + \boldsymbol{A}_k^{(\ell)} + \boldsymbol{M}_k^{(\ell)},$$

其中

$$\boldsymbol{A}_k^{(\ell)} := \sum_{j=0}^{k-1} \boldsymbol{a}_j^{(\ell)}, \quad \boldsymbol{M}_k^{(\ell)} := \sum_{j=0}^{k-1} \boldsymbol{m}_j^{(\ell)}.$$

因此,

$$\begin{aligned}\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{X}_m^{(\ell)}\|^p\right] \le 3^{p-1}\Big(&\mathbb{E}\left[\|\boldsymbol{X}_0\|^p\right] + \mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{A}_m^{(\ell)}\|^p\right] \\ &+ \mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{M}_m^{(\ell)}\|^p\right]\Big).\end{aligned} \tag{L.10}$$

- 步骤 2。 $\boldsymbol{A}_k^{(\ell)}$ 的有界性: 首先, 由 (L.7), 由于 $\mathbb{I}_{\{k<\boldsymbol{N}_\ell\}}$ 是 $\mathcal{F}_k$-可测的,

$$\mathbb{E}\left[\|\boldsymbol{v}_k\|^q \mathbb{I}_{\{k<\boldsymbol{N}_\ell\}} \mid \mathcal{F}_k\right] \le C_{1,v}\left(1 + \|\boldsymbol{X}_k^{(\ell)}\|^{q/3}\right) h^q + C_{2,v} h^{q/2}. \tag{L.11}$$

将 (L.11) 与 (L.1) 结合,

$$\begin{aligned}\|\boldsymbol{a}_k^{(\ell)}\| &\le \|\mathbb{E}\left[\boldsymbol{v}_k \mathbb{I}_{\{k<\boldsymbol{N}_\ell\}} \mid \mathcal{F}_k\right]\| + \frac{1}{2}\kappa \mathbb{E}\left[\|\boldsymbol{v}_k\|^2 \mathbb{I}_{\{k<\boldsymbol{N}_\ell\}} \mid \mathcal{F}_k\right] \\ &\quad + C_R \mathbb{E}\left[\|\boldsymbol{v}_k\|^3 \mathbb{I}_{\{k<\boldsymbol{N}_\ell\}} \mid \mathcal{F}_k\right] \\ &\le C_a\left(1 + \|\boldsymbol{X}_k^{(\ell)}\|\right) h,\end{aligned} \tag{L.12}$$

对于某个 $C_a > 0$, 其中我们对所有 $a \ge 0$ 使用 $a^{1/3} \le a + 1$ 和 $a^{2/3} \le a + 1$。由此可得

$$\begin{aligned}\max_{0\le m\le k}\|\boldsymbol{A}_m^{(\ell)}\|^p &\le \max_{0\le m\le k} m^{p-1} \sum_{j=0}^{m-1} \|\boldsymbol{a}_j^{(\ell)}\|^p \le N^{p-1} \sum_{j=0}^{k-1} \|\boldsymbol{a}_j^{(\ell)}\|^p \\ &\le N^{p-1} C_a^p h^p 2^{p-1} \sum_{j=0}^{k-1}\left(1 + \|\boldsymbol{X}_j^{(\ell)}\|^p\right) \\ &\le 2^{p-1} C_a^p T^{p-1} \sum_{j=0}^{k-1} h\left(1 + \max_{0\le m\le j}\|\boldsymbol{X}_m^{(\ell)}\|^p\right).\end{aligned}$$

Let

$$\alpha_k^{(\ell)} := \mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{X}_m^{(\ell)}\|^p\right].$$

Then

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{A}_m^{(\ell)}\|^p\right] \le 2^{p-1} C_a^p T^{p-1} \sum_{j=0}^{k-1} h\left(1 + \alpha_j^{(\ell)}\right). \tag{L.13}$$

- 步骤 3. $\boldsymbol{M}_{(\ell)\ k}$ 的有界性：对于 $\boldsymbol{m}_{(\ell)\ k}$，根据定义，

$$\mathbb{E}\left[\|\boldsymbol{m}_k^{(\ell)}\|^p \mid \mathcal{F}_k\right] \le 2^{p-1}\left(\mathbb{E}\left[\|\boldsymbol{a}_k^{(\ell)}\|^p \mid \mathcal{F}_k\right] + \mathbb{E}\left[\|\boldsymbol{u}_k^{(\ell)}\|^p \mid \mathcal{F}_k\right]\right),$$

其中，根据 (L.12)，

$$\mathbb{E}\left[\|\boldsymbol{a}_k^{(\ell)}\|^p \mid \mathcal{F}_k\right] = \|\boldsymbol{a}_k^{(\ell)}\|^p \le 2^{p-1}C_a^p\left(1+\|\boldsymbol{X}_k^{(\ell)}\|^p\right)h^p. \tag{L.14}$$

对于 $\boldsymbol{u}_k^{(\ell)} = \mathbb{I}_{\{k<\boldsymbol{N}_\ell\}}\boldsymbol{u}_k^X$，结合 (J.26) 和 (L.11)，

$$\mathbb{E}\left[\|\boldsymbol{u}_k^{(\ell)}\|^p \mid \mathcal{F}_k\right] \le C_{1,u}\left(1+\|\boldsymbol{X}_k^{(\ell)}\|^p\right)h^p + C_{2,u}h^{p/2}, \tag{L.15}$$

对于某些常数 $C_{1,\ u}$，$C_{2,\ u} > 0$，其中我们还对所有 $a \ge 0$ 使用了 $a^{1/3} \le a+1$ 和 $a^{2/3} \le a+1$。因此，结合 (L.14) 和 (L.15)，并假设 $h \le 1$，

$$\mathbb{E}\left[\|\boldsymbol{m}_k^{(\ell)}\|^p \mid \mathcal{F}_k\right] \le C_m\left(1+\|\boldsymbol{X}_k^{(\ell)}\|^p\right)h^{p/2},$$

对于某个常数 $C_m > 0$，这意味着

$$\mathbb{E}\left[\|\boldsymbol{m}_k^{(\ell)}\|^p\right] \le C_m\left(1+\mathbb{E}\left[\|\boldsymbol{X}_k^{(\ell)}\|^p\right]\right)h^{p/2} \le C_m\left(1+\alpha_k^{(\ell)}\right)h^{p/2}. \tag{L.16}$$

由于 $\boldsymbol{M}_k^{(\ell)}$ 是一个鞅，离散时间 BDG 不等式（Zorin-Kranich，2022，推论 3.2）意味着存在 $C_{M,\ p} > 0$ 使得

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{M}_m^{(\ell)}\|^p\right] \le C_{M,p}\mathbb{E}\left[[\boldsymbol{M}^{(\ell)}]_k^{p/2}\right],$$

其中

$$[\boldsymbol{M}^{(\ell)}]_k^{p/2} = \left(\sum_{j=0}^{k-1}\|\boldsymbol{m}_j^{(\ell)}\|^2\right)^{p/2} \le k^{\frac{p}{2}-1}\sum_{j=0}^{k-1}\|\boldsymbol{m}_j^{(\ell)}\|^p$$

由引理 N.2（$p \ge 2$）。因此，由 (L.16) 得

$$\mathbb{E}\left[\max_{0\le m\le k}\|\boldsymbol{M}_m^{(\ell)}\|^p\right] \le C_{M,p}N^{\frac{p}{2}-1}\sum_{j=0}^{k-1}\mathbb{E}\left[\|\boldsymbol{m}_j^{(\ell)}\|^p\right] \le C_{M,p}C_mT^{\frac{p}{2}-1}\sum_{j=0}^{k-1}h\left(1+\alpha_k^{(\ell)}\right). \tag{L.17}$$

- 步骤 4. 最终界：将 (L.13) 和 (L.17) 与 (L.10) 结合，我们得到

$$\alpha_k^{(\ell)} \le C_{1,p} + C_{2,p}T^{p-1}\sum_{j=0}^{k-1}h\left(1+\alpha_k^{(\ell)}\right),$$

对于某些常数 $C_{1,\ p}$，$C_{2,\ p} > 0$（此处我们假设 $T \ge 1$ 使得 $T^{p/2} \le T^p$）。那么，离散 Gr¨onwall 不等式（Thom´ee，2007，引理 10.5）意味着

$$\alpha_N^{(\ell)} \le (1+C_{1,p})e^{C_{2,p}T^p} - 1,$$

即,

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}^h_{k\wedge\boldsymbol{N}_\ell}\right\|^p\right]\le(1+C_{1,p})e^{C_{2,p}T^p}-1.$$

由于 $\ell\to\infty$，根据法图引理,

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|\boldsymbol{X}^h_{k}\right\|^p\right]\le\mathcal{O}(\exp(T^p))<\infty.$$

对于 $1\le p<2$，我们通过 Lyapunov 不等式进行类似证明。 ■

引理 L.4 设 $\mathcal{S}\subset\mathbb{R}^n$ 为一个非空子集。若 $f\colon\ \mathcal{S}\to\mathbb{R}^m$ 满足存在L、$K>0$ 使得

$$\|f(x_1)-f(x_2)\|\le L\|x_1-x_2\|,\quad \|f(x)\|\le K\left(1+\|x\|^{1/3}\right),\qquad\forall\ x,x_1,x_2\in\mathcal{S},$$

则存在一个 $F\colon\ \mathbb{R}^n\to\mathbb{R}^m$，使得 $F|_{\mathcal{S}}=f$ 且

$$\|F(y_1)-F(y_2)\|\le(L+2K)\|y_1-y_2\|,\quad \|F(y)\|\le 2K\left(1+\|y\|^{1/3}\right),\qquad\forall\ y,y_1,y_2\in\mathbb{R}^n.$$

证明 根据基尔施布劳恩定理（Cobza̧s 等，2019，定理 4.2.3），存在一个$G\colon\ \mathbb{R}^n\to\mathbb{R}^m$，使得 $G|_{\mathcal{S}}=f$ 且

$$\|G(y_1)-G(y_2)\|\le L\|y_1-y_2\|,\quad\forall\ y_1,y_2\in\mathbb{R}^n.$$

为保持次线性增长，定义半径函数 $r\colon\ \mathbb{R}^n\to(0,\ \infty)$，其定义为

$$r(y):=2K\left(1+\min\left\{\|y\|,\|y\|^{1/3}\right\}\right).$$

显然,

$$K\left(1+\|y\|^{1/3}\right)\le r(y)\le 2K\left(1+\|y\|^{1/3}\right).\tag{L.18}$$

此外，由于映射 $t\mapsto\min\left\{t,t^{1/3}\right\}$ 在 $(0,\infty)$ 上是 1-Lipschitz 的，函数 r 是 $2K$-Lipschitz 连续的。

接下来，考虑由下式定义的投影映射 $\Pi\colon(0,\infty)\times\mathbb{R}^m\to\mathbb{R}^m$

$$\Pi(r,z):=\begin{cases}z, & \|z\|\le r,\\ r\frac{z}{\|z\|}, & \|z\|>r.\end{cases}$$

定义

$$F(y):=\Pi(r(y),G(y)).$$

对于任意 $x\in\mathcal{S}$，由于 $G(x)=f(x)$，(L.18) 给出

$$\|f(x)\|\le K\left(1+\|x\|^{1/3}\right)\le r(x),$$

因此

$$F(x)=\Pi(r(x),G(x))=G(x)=f(x),$$

即， $F|_{\mathcal{S}} = f$。次线性增长界可直接由 F的定义和 (L.18) 得出：

$$\|F(y)\| = \|\Pi(r(y), G(y))\| \le r(y) \le 2K\left(1 + \|y\|^{1/3}\right).$$

最后，利用 r 和 Π 的利普希茨连续性（引理 N.4），我们得到

$$\begin{aligned}\|F(y_1) - F(y_2)\| &\le \|\Pi(r(y_1), G(y_1)) - \Pi(r(y_1), G(y_2))\| \\ &\quad + \|\Pi(r(y_1), G(y_2)) - \Pi(r(y_2), G(y_2))\| \\ &\le \|G(y_1) - G(y_2)\| + |r(y_1) - r(y_2)| \\ &\le (L + 2K)\,\|y_1 - y_2\|.\end{aligned}$$

■

附录 M. 第5节补充细节

´本节在 Bakry–Emery 曲率条件下界定初始 Wasserstein 误差；见第 M.1 节。我们还通过 GEM 给出了 RLD 的外在瓦瑟斯坦界；见第 M.2 节。

M.1 初始误差

引理 M.1 设$(\mathcal{M}, g)$ 是一个无边界的完备连通黎曼流形。对于 $\phi \in C^\infty(\mathcal{M})$ ，其中 $C_\phi := \int_{\mathcal{M}} \phi \ \mathrm{dvol} < \infty$，令 μ_ϕ 为定义在 $\mathcal{M}$ 上的概率分布，其定义为$\mathrm{d}\mu_\phi := C_\phi^{-1} e^{-\phi}\ \mathrm{dvol}$。令 $1 \le p < \infty$。若 ϕ 满足 ***Bakry–´Emery*** 曲率条件

$$\mathrm{Ric} + \nabla^2 \phi \ge \lambda_\kappa g, \quad \lambda_\kappa > 0,$$

那么对于任意 $x \in \mathcal{M}$,

$$\mathcal{W}_p\left(\mu_\phi, \delta_x\right) < \infty.$$

´证明 由于 ϕ满足 Bakry–Emery曲率条件，μ_ϕ 满足曲率–维数条件 CD(λ_κ, ∞)（Villani，2008年，定理 14.8）。这意味着 μ_ϕ 满足 Talagrand 不等式 $T_2(\lambda_\kappa)$，从而 $T_1(\lambda_\kappa)$（Villani，2008年，定理 22.14）。

固定 $x \in \mathcal{M}$。令 m_x为 ρ_x的中位数：$= d_{\mathcal{M}}(x, \cdot)$: $\mathcal{M} \to \mathbb{R}$，其定义为

$$m_x := \inf\left\{t \ge 0 \colon \mu_\phi\left(\rho_x \le t\right) \ge \frac{1}{2}\right\} < \infty.$$

显然， ρ_x 是 1-Lipschitz连续的。因此，根据 Villani（2008）中的定理 22.10，对于任意 $r \ge 0$,

$$\mu_\phi\left(\rho_x \ge m_x + r\right) \le \exp\left(-\frac{\lambda_\kappa}{2} r^2\right). \tag{M.1}$$

利用尾积分公式（Folland，2013年，命题 6.23），我们得到

$$\int_{\mathcal{M}} \rho_x(y)^p \ \mathrm{d}\mu_\phi(y) = \int_0^\infty p t^{p-1} \mu_\phi\left(\rho_x \ge t\right) \ \mathrm{d}t$$

$$= \int_0^{m_x} pt^{p-1}\mu_\phi\left(\rho_x \geq t\right)\ \mathrm{d}t + \int_{m_x}^{\infty} pt^{p-1}\mu_\phi\left(\rho_x \geq t\right)\ \mathrm{d}t.$$

第一项由 m_x^p 有界。对于第二项，由（M.1）可得

$$\begin{aligned}\int_{m_x}^{\infty} pt^{p-1}\mu_\phi\left(\rho_x \geq t\right)\ \mathrm{d}t &= \int_0^{\infty} p\left(m_x + r\right)^{p-1}\mu_\phi\left(\rho_x \geq m_x + r\right)\ \mathrm{d}r \\ &\leq \int_0^{\infty} p\left(m_x + r\right)^{p-1}\exp\left(-\frac{\lambda_\kappa}{2}r^2\right)\ \mathrm{d}r < \infty.\end{aligned}$$

因此，

$$\mathcal{W}_p\left(\mu_\phi, \delta_x\right)^p \leq \int_{\mathcal{M}} d_{\mathcal{M}}(x,y)^p\ \mathrm{d}\mu_\phi(y) < \infty.$$

■

M.2 外在 Wasserstein 界

由于 $\mathcal{M} \subset \mathbb{R}^n$ 具有嵌入映射 ι： $\mathcal{M} \hookrightarrow \mathbb{R}^n$，定义在$\mathcal{M}$ 上的任意概率分布 μ 都有一个定义在 $\mathbb{R}^n$上的推前概率分布 $\widetilde{\mu} = \iota_{\#}\mu$，即

$$\widetilde{\mu}(\mathcal{V}) := \mu(\mathcal{V} \cap \mathcal{M}).$$

对于 μ，ν 在 $\mathcal{M}$上，令 $\widetilde{\mu}$，$\widetilde{\nu}$为它们的推前测度。那么

$$\widetilde{\mathcal{W}}_p(\mu,\nu) := \mathcal{W}_p(\widetilde{\mu},\widetilde{\nu}) = \inf_{\boldsymbol{X}\sim\mu, \boldsymbol{Y}\sim\nu} \mathbb{E}\left[\|\boldsymbol{X} - \boldsymbol{Y}\|^p\right]^{1/p}.$$

由于 ι 天然是 1-Lipschitz 的（附录 F.1），我们有

$$\widetilde{\mathcal{W}}_p(\mu,\nu) \leq \mathcal{W}_p(\mu,\nu).$$

定理 M.2 令 $1 \leq p < \infty$。令 $\mathcal{M} \subset \mathbb{R}^n$ 为一个黎曼子流形。令 $\phi \in C^\infty(\mathcal{M})$ 满足 $C_\phi := \int_{\mathcal{M}} \phi\ \mathrm{dvol} < \infty$，且$\mathrm{d}\mu_\phi := C_\phi^{-1}e^{-\phi}\ \mathrm{dvol}$。假设 $\mathcal{M}$ 满足假设 *I* 和假设 *II*，ϕ 满足（*14*），且 $\nabla\phi$ 满足假设 *III*。那么对于由（*11*）构造的 $\boldsymbol{X}_k^h \sim\hat{}\ \mu_k$，我们有

$$\widetilde{\mathcal{W}}_p\left(\mu_\phi, \hat{\mu}_N\right) \leq \mathcal{O}\left(e^{-T} + \exp(T^p)h^{1/2}\right).$$

证明 根据第 5 节中的陈述，由于 $\mathcal{M}$满足假设 I，且 ϕ 满足 (14)，混合误差有界为

$$\widetilde{\mathcal{W}}_p\left(\mu_\phi, \mu_T\right) \leq \mathcal{W}_p\left(\mu_\phi, \mu_T\right) \leq \mathcal{O}(e^{-T}). \tag{M.2}$$

对于离散化误差，由于 $\mathcal{M}$满足假设 I 和假设 II，且 $\nabla\phi$满足假设 III，定理 5 表明

$$\widetilde{\mathcal{W}}_p\left(\mu_T, \hat{\mu}_N\right) \leq \mathbb{E}\left[\left\|\boldsymbol{X}_T - \boldsymbol{X}_N^h\right\|^p\right]^{1/p} \leq \mathcal{O}\left(\exp(T^p)h^{1/2}\right). \tag{M.3}$$

因此，(M.2) 和 (M.3) 共同意味着

$$\widetilde{\mathcal{W}}_p\left(\mu_\phi, \hat{\mu}_N\right) \leq \widetilde{\mathcal{W}}_p\left(\mu_\phi, \mu_T\right) + \widetilde{\mathcal{W}}_p\left(\mu_T, \hat{\mu}_N\right) \leq \mathcal{O}\left(e^{-T} + \exp(T^p)h^{1/2}\right).$$

■

附录 N. 技术引理

引理 N.1 令 $\boldsymbol{\xi} \sim \mathcal{N}(0, I_m)$。那么，对于任意 $1 \le p < \infty$，存在一个常数 $C > 0$，使得

$$\mathbb{E}\left[\|\boldsymbol{\xi}\|^p\right] \le \left(\sqrt{m} + C\sqrt{p}\right)^p.$$

证明 对于 $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_m) \in \mathbb{R}^m$，每个 $\boldsymbol{\xi}_i$ 都是次高斯的，因此

$$K := \max_i \|\boldsymbol{\xi}_i\|_{\psi_2} < \infty.$$

由 Vershynin（2018，定理 3.1.1）可得

$$\left\| \|\boldsymbol{\xi}\| - \sqrt{m} \right\|_{\psi_2} \le C_0 K,$$

对于某个常数 $C_0 > 0$。令 $\boldsymbol{Y} = \|\boldsymbol{\xi}\| - \sqrt{m}$。那么 Vershynin（2018）中的命题 2.2.6 给出

$$\mathbb{E}\left[|\boldsymbol{Y}|^p\right]^{1/p} \le C_1 \|\boldsymbol{Y}\|_{\psi_2} \sqrt{p} \le C_1 C_0 K \sqrt{p}.$$

最后，由闵可夫斯基不等式，

$$\mathbb{E}\left[\|\boldsymbol{\xi}\|^p\right]^{1/p} \le \sqrt{m} + \mathbb{E}\left[|\boldsymbol{Y}|^p\right]^{1/p} \le \sqrt{m} + C\sqrt{p},$$

其中 $C = C_1 C_0 K$。 ∎

引理 N.2 设 $h_0, \dots, h_{N-1} \ge 0$ 满足 $T = \sum_i h_i$，并设 $a_0, \dots, a_{N-1} \ge 0$。对任意 $p \ge 2$，

$$\left(\sum_{i=0}^{N-1} h_i a_i^2 \right)^{p/2} \le T^{\frac{p}{2}-1} \sum_{i=0}^{N-1} h_i a_i^p.$$

此外，对 $p \ge 2$，若 $a(t) \ge 0$，则

$$\left(\int_0^T a(t)^2 \, \mathrm{d}t \right)^{p/2} \le T^{\frac{p}{2}-1} \int_0^T a(t)^p \, \mathrm{d}t.$$

证明 情形 $p = 2$ 是显然的。假设 $p > 2$，并设 $r = p/2 > 1$。设 r' 为 r 的共轭指数，即

$$\frac{1}{r} + \frac{1}{r'} = 1 \quad \Rightarrow \quad r' = \frac{p}{p-2}.$$

定义

$$x_i := h_i a_i^2, \quad y_i := h_i.$$

根据 H¨older 不等式，

$$\sum_{i=0}^{N-1} x_i = \sum_{i=0}^{N-1} \left(x_i y_i^{-1/r'} \right) y_i^{1/r'} \le \left(\sum_{i=0}^{N-1} x_i^r y_i^{1-r} \right)^{1/r} \left(\sum_{i=0}^{N-1} y_i \right)^{1/r'}.$$

由于 $x_i^r y_i^{1-r} = (h_i a_i^2)^r h_i^{1-r} = h_i a_i^{2r} = h_i a_i^p$ 和 $\sum_i y_i = \sum_i h_i = T$，我们得到

$$\sum_{i=0}^{N-1} h_i a_i^2 \le \left(\sum_{i=0}^{N-1} h_i a_i^p \right)^{2/p} T^{1-2/p}.$$

因此，

$$\left(\sum_{i=0}^{N-1} h_i a_i^2 \right)^{p/2} \le T^{\frac{p}{2}-1} \sum_{i=0}^{N-1} h_i a_i^p.$$

连续不等式可通过相同的论证得到：将求和替换为积分，并应用 H¨older 不等式的积分形式。

■

引理 N.3 对于任意 $r > 0$，存在一个 $\psi \in C^\infty([0,\infty))$，使得 $\psi \in [0,1]$，

$$\psi|_{[0,r/2]} \equiv 1, \quad \psi|_{[r,\infty)} \equiv 0, \qquad and \ \sup_{t\ge 0} \left|\psi'(t)\right| < \infty.$$

证明 令

$$\phi(s) := \begin{cases} e^{-1/s}, & s > 0, \\ 0, & s \le 0. \end{cases}$$

于是 $\phi \in C^\infty(\mathbb{R})$。定义 $\theta(s) := \frac{\phi(s)}{\phi(s)+\phi(1-s)}$。则 $\theta \in C^\infty(\mathbb{R})$，$\theta \in [0,1]$，并且

$$\theta(s) = 0 \ (s \le 0), \quad \theta(s) = 1 \ (s \ge 1).$$

特别地，对于 $s \in (-\infty, 0] \cup [1,\infty)$，$\theta'(s) = 0$，因此

$$C_\theta := \sup_{s\in\mathbb{R}} \left|\theta'(s)\right| < \infty.$$

现在定义

$$\psi(t) := 1 - \theta\left(\frac{2t}{r} - 1\right), \quad \forall\, t \in [0,\infty).$$

于是 $\psi \in C^\infty([0,\infty))$ 和 $\psi \in [0,1]$。根据构造，$\psi(t) = 1$ 当 $t \in [0, r/2]$ 时，以及 $\psi(t) = 0$当 $t \ge r$时。此外，

$$\psi'(t) = -\frac{2}{r}\theta'\left(\frac{2t}{r} - 1\right),$$

so

$$\sup_{t\ge 0} \left|\psi'(t)\right| \le \frac{2}{r} \sup_{s\in\mathbb{R}} \left|\theta'(s)\right| = \frac{2C_\theta}{r} < \infty.$$

■

引理 N.4 令 $\Pi\colon (0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$ 定义为

$$\Pi(r,z) := \begin{cases} z, & \|z\| \le r, \\ r\frac{z}{\|z\|}, & \|z\| > r. \end{cases}$$

于是：

(i) 对于每个固定的 $z \in \mathbb{R}^n$，映射 $\Pi(\cdot, z)$：$(0, \infty) \to \mathbb{R}^n$ 是 1-*Lipschitz* 连续的 *s.*

(ii) 对于每个固定的 $r > 0$，映射 $\Pi(r, \cdot)$：$\mathbb{R}^n \to \mathbb{R}^n$ 是 1-*Lipschitz* 连续的。

证明

(i) 固定 $z \in \mathbb{R}^n$ 并令 $\rho := \|z\|$。设 $r_1, r_2 > 0$。若 $\rho \le \min\{r_1, r_2\}$，则 $\Pi(r_1, z) = \Pi(r_2, z) = z$，因此 $\|\Pi(r_1, z) - \Pi(r_2, z)\| = 0$。若 $\rho \ge \max\{r_1, r_2\}$，则 $\Pi(r_i, z) = r_i z/\rho$，因此

$$\|\Pi(r_1, z) - \Pi(r_2, z)\| = \left\| (r_1 - r_2)\frac{z}{\rho} \right\| = |r_1 - r_2|.$$

若 $r_1 \ge \rho \ge r_2$，则 $\Pi(r_1, z) = z$，且 $\Pi(r_2, z) = r_2 z/\rho$，因此

$$\|\Pi(r_1, z) - \Pi(r_2, z)\| = \left\| z - \frac{r_2}{\rho} z \right\| = \rho - r_2 \le r_1 - r_2.$$

在所有情况下，$\|\Pi(r_1, z) - \Pi(r_2, z)\| \le |r_1 - r_2|$，这就证明了该断言。

(ii) 固定 $r > 0$。映射 $\Pi(r, \cdot)$ 是到闭球 $B(r) = \{x \in \mathbb{R}^n : \|x\| \le r\}$ 的度量投影。因此，对于任意 $z \in \mathbb{R}^n$ 和任意 $x \in B(r)$，Bertsekas（2009，命题 1.1.9）指出

$$\langle z - \Pi(r, z),\, x - \Pi(r, z) \rangle \le 0.$$

令 $y, z \in \mathbb{R}^n$，并设 $u := \Pi(r, y)$ 和 $v := \Pi(r, z)$。将上述不等式分别应用于 $(z, x) = (y, v)$ 和 $(z, x) = (z, u)$，可得

$$\langle y - u,\, v - u \rangle \le 0, \quad \langle z - v,\, u - v \rangle \le 0.$$

将这些不等式相加，得到

$$\|u - v\|^2 \le \langle y - z,\, u - v \rangle \le \|y - z\| \|u - v\|.$$

若 $u = v$，则无需证明；否则，两边除以 $\|u - v\|$ 得 $\|u - v\| \le \|y - z\|$，即

$$\|\Pi(r, y) - \Pi(r, z)\| \le \|y - z\|.$$

■

引理 N.5 设 $A \in \mathbb{R}^{k \times n}$ $(k < n)$，且秩 $A = k$。则 A 存在 *Moore–Penrose* 右逆

$$A^\dagger = A^T \left(A A^T \right)^{-1},$$

且 $P = A^\dagger A$ 是到（核 A）$^\perp$上的正交投影。

证明 由于秩为 $A = k$，矩阵 $AA^T \in \mathbb{R}^{k \times k}$ 可逆，因此 Moore–Penrose 右逆 $A^\dagger$ 是良定义的。对于 $P = A^\dagger A$，我们有

$$P^T = P, \quad P^2 = P.$$

因此，P是到 Im P上的正交投影。此外，由于 A 满秩，

$$\operatorname{Im} P = \operatorname{Im} A^\dagger = \operatorname{Im} A^T = (\ker A)^\perp.$$

■

引理 N.6 设$(\mathcal{M}, g)$ 为一个 m维黎曼流形。对于任意 $x \in \mathcal{M}$，固定 $T_x\mathcal{M}$ 的一组基 $\{E_i(x)\}_{i=1}^m$，并令$g_{ij}(x) = g(x(E_i(x)\ E_j(x))$。假设存在$\lambda_{\min} > 0$，使得矩阵 $g(x) = (g_{ij}(x))$ 满足

$$g(x) \succeq \lambda_{\min} I_m.$$

令 $F \in \Gamma(\mathcal{M}, \mathrm{Hom}(T\mathcal{M} \otimes T\mathcal{M}, \mathcal{M} \times \mathbb{R}^n))$，即对任意 $x \in \mathcal{M}$, $F_x : \ T_x\mathcal{M} \times T_x\mathcal{M} \to \mathbb{R}^n$都是一个双线性映射。若存在常数 $C > 0$，使得对任意 $x \in \mathcal{M}$,

$$\|F_x(E_i(x), E_j(x))\| \le C,$$

then

$$\sup_{x \in \mathcal{M}} \|F_x\|_{\mathrm{op}} \le \frac{Cm}{\lambda_{\min}}.$$

证明 固定 $x \in \mathcal{M}$。对任意 $u, v \in T_x\mathcal{M}$，记

$$u = \sum_{i=1}^m u_i E_i, \quad v = \sum_{j=1}^m v_j E_j.$$

Then

$$F_x(u, v) = \sum_{i,j=1}^m u_i v_j F_x(E_i(x), E_j(x)),$$

因此

$$\|F_x(u, v)\| \le C \sum_{i,j=1}^m |u_i v_j| = C \left(\sum_{i=1}^m |u_i| \right) \left(\sum_{j=1}^m |v_j| \right) \le Cm \, |u|_2 \, |v|_2 \,,$$

其中 $|u|_2 := (\sum_i u_i^2)^{1/2}$ 且 $|v|_2 := (\sum_j v_j^2)^{1/2}$，最后一个不等式用到了范数不等式 $\ell_1 \le \sqrt{m}\, \ell_2$。 —

另一方面，由于 $g(x) \succeq \lambda_{\min} I_m$,

$$\|u\|^2 = (u_1, \ldots, u_m)\, g(x)\, (u_1, \ldots, u_m)^T \ge \lambda_{\min} \, |u|_2^2 \,,$$

类似地， $\|v\|^2 \ge \lambda_{\min} \, |v|_2^2$。因此,

$$\|F_x(u, v)\| \le Cm \, |u|_2 \, |v|_2 \le \frac{Cm}{\lambda_{\min}} \|u\| \|v\|.$$

由此可得

$$\|F_x\|_{\mathrm{op}} = \sup_{u, v \ne 0} \frac{\|F_x(u, v)\|}{\|u\| \|v\|} \le \frac{Cm}{\lambda_{\min}}.$$

由于 C 和 $\lambda_{\min}$ 与 x相互独立，我们得出

$$\sup_{x \in \mathcal{M}} \|F_x\|_{\mathrm{op}} \le \frac{Cm}{\lambda_{\min}}.$$

■

注 N.1 同样的论证显然可以推广到三线性情形 $F_x\colon\quad T_x\mathcal{M}\times T_x\mathcal{M}\times T_x\mathcal{M}\to\mathbb{R}^n$。

引理 N.7 设 $\mathcal{M}\subset\mathbb{R}^n$ 为带有列维-奇维塔联络 ∇ 的黎曼子流形。设 $f\in C^\infty(\mathbb{R}^n)$ 与 $g=f|_{\mathcal{M}}\in C^\infty(\mathcal{M})$。则

$$\nabla^2 g(X,Y)=D^2f[X,Y]+Df[\mathrm{II}(X,Y)],$$

其中 II 为第二基本形式。

证明 根据海森矩阵的定义以及高斯公式（Lee, 2018），对任意 $X,Y\in\Gamma(T\mathcal{M})$,

$$\begin{aligned}\nabla^2 g(X,Y)&=X\big(Y(g)\big)-(\nabla_XY)(g)\\&=X\big(Y(f)\big)-\big(D_XY-\mathrm{II}(X,Y)\big)(f)\\&=\big(X\big(Y(f)\big)-(D_XY)(f)\big)+\mathrm{II}(X,Y)(f)\\&=D^2f[X,Y]+Df[\mathrm{II}(X,Y)].\end{aligned}$$

■

引理 N.8 设 $\mathcal{M}\subset\mathbb{R}^n$ 为黎曼子流形。则协变导数 $\nabla\mathrm{II}$ 是逐点确定的，即对任意 $x\in\mathcal{M}$ 以及任意 $X,Y,Z,\tilde{X},\tilde{Y},\tilde{Z}\in\Gamma(T\mathcal{M})$，且 $X_x=\tilde{X}_x$, $Y_x=\tilde{Y}_x$ 与 $Z_x=\tilde{Z}_x$,

$$(\nabla\mathrm{II})(X,Y,Z)(x)=(\nabla\mathrm{II})(\tilde{X},\tilde{Y},\tilde{Z})(x).$$

证明 只需证明 $\nabla\mathrm{II}$ 是一个 $N\mathcal{M}$ 值的 $(0,3)$-张量，即它是 $C^\infty(\mathcal{M})$-三线性的。根据定义,

$$(\nabla\mathrm{II})(X,Y,Z)=(D_Z\mathrm{II}(X,Y))^\perp-\mathrm{II}(\nabla_ZX,Y)-\mathrm{II}(X,\nabla_ZY).$$

根据 Lee（2018，命题 8.1），II 是 $C^\infty(\mathcal{M})$-双线性且对称的。设 $f\in C^\infty(\mathcal{M})$。对于 X 位置，利用联络的线性性和莱布尼茨法则,

$$\begin{aligned}(D_Z\mathrm{II}(fX,Y))^\perp&=\big(D_Z\big(f\mathrm{II}(X,Y)\big)\big)^\perp\\&=(Z(f)\mathrm{II}(X,Y)+fD_Z\mathrm{II}(X,Y))^\perp\\&=Z(f)\mathrm{II}(X,Y)+f\left(D_Z\mathrm{II}(X,Y)\right)^\perp,\end{aligned}$$

and

$$\mathrm{II}(\nabla_Z(fX),Y)=\mathrm{II}\big(Z(f)X+f\nabla_ZX,\,Y\big)=Z(f)\mathrm{II}(X,Y)+f\mathrm{II}(\nabla_ZX,Y),$$

而

$$\mathrm{II}(fX,\nabla_ZY)=f\mathrm{II}(X,\nabla_ZY).$$

将这些恒等式代入 $\nabla\mathrm{II}$ 的定义可得

$$(\nabla\mathrm{II})(fX,Y,Z)=f(\nabla\mathrm{II})(X,Y,Z).$$

由于 II 是对称的，同样的论证也适用于 Y位置，即 $(\nabla \mathrm{II})(X,\ fY,\ Z) = f(\nabla \mathrm{II})(X,\ Y,\ Z)$。

对于第 Z个位置，根据联络的线性性，

$$(D_{fZ}\mathrm{II}(X,Y))^{\perp} = (fD_Z\mathrm{II}(X,Y))^{\perp} = f\left(D_Z\mathrm{II}(X,Y)\right)^{\perp},$$

and

$$\mathrm{II}(\nabla_{fZ}X,Y) = f\mathrm{II}(\nabla_Z X,Y), \quad \mathrm{II}(X,\nabla_{fZ}Y) = f\mathrm{II}(X,\nabla_Z Y).$$

因此 $(\nabla \mathrm{II})(X,\ Y,\ fZ) = f(\nabla \mathrm{II})(X,\ Y,\ Z)$。对每个位置的可加性是显然的。 ■

参考文献

John Armstrong 和 Tim King。《流形上或流形附近随机微分方程的曲线型格式》。*Proceedings of the Royal Society A*，478(2262)：20210785，2022。

Dominique Bakry、Ivan Gentil 和 Michel Ledoux。《马尔可夫扩散算子的分析与几何》。数学科学基础丛书。施普林格，2013年。

Fabrice Baudoin。《扩散过程与随机微积分》。EMS 数学教材。欧洲数学学会，2014年。

Yoshua Bengio、Aaron Courville 和 Pascal Vincent。表征学习：综述与新视角。*IEEE*模式分析与机器智能汇刊，35(8):1798–1828，2013年。

Dimitri P. Bertsekas。凸优化理论。Athena Scientific，2009。

Karthik Bharath、Alexander Lewis、Akash Sharma 和 Michael V. Tretyakov。利用朗之万扩散在流形上进行采样与估计。机器学习研究杂志，26(71):1–50，2025年。

Vladimir I. Bogachev。测度论，第*1*卷。施普林格，2007。

MinshuoChen、Kaixuan Huang、Tuo Zhao 和MengdiWang。扩散模型在低维数据上的分数逼近、估计与分布恢复。载于第*40*届*International Conference on Machine Learning*论文集，第202卷，机器学习研究会议录，第4672–4712页。PMLR，2023年。

Xiang Cheng、Jingzhao Zhang 和 Suvrit Sra。通过朗之万马尔可夫链蒙特卡罗在黎曼流形上高效采样。载于神经信息处理系统进展，第 35 卷，2022 年。

S.. Cobza.s、R. Miculescu 和A. Nicolae。利普希茨函数。数学讲义。施普林格，2019 年。

Valentin De Bortoli。基于流形假设的去噪扩散模型的收敛性。机器学习研究汇刊，2022。

Valentin De Bortoli、Emile Mathieu、Michael Hutchinson、James Thornton、Yee Whye Teh 和 Arnaud Doucet。黎曼评分生成建模。载于神经信息处理系统进展，第35卷，第2406–2422页，2022年。Prafulla Dhariwal 和 Alexander Nichol。扩散模型在图像合成方面胜过生成对抗网络。载于神经信息处理系统进展，第34卷，第8780–8794页，2021年。Manfredo P. do Carmo。曲线与曲面的微分几何：修订与更新版第二版。多佛数学丛书。多佛出版社，2016。Rick Durrett。概率论：理论与例题。剑桥统计与概率数学丛书。剑桥大学出版社，2019。´Michel Emery。流形上的随机微积分。Universitext。施普林格，2012。Tyler Farghly、Peter Potaptchik、Samuel Howard、George Deligiannidis 和 Jakiw Pidstrigach。扩散模型与流形假设：对数域平滑具有几何自适应性。载于神经信息处理系统进展，2025年。Herbert Federer。曲率测度。美国数学会会刊，93(3):418–491，1959年。Gerald B. Folland。实分析：现代技术及其应用。威利，2013年。Khashayar Gatmiry 和 Santosh S. Vempala。黎曼朗之万算法的收敛性。*arXiv*预印本 *arXiv:2204.10818*，2022年。网址：https://arxiv.org/abs/2204. 10818。Jonathan S. Golan。研究生初学者应知的线性代数。数学科学丛书。施普林格，2012年。Jonathan Ho、Ajay Jain 和 Pieter Abbeel。去噪扩散概率模型。载于神经信息处理系统进展，第33卷，第6840–6851页，2020年。Roger A. Horn 和 Charles R. Johnson。矩阵分析。剑桥大学出版社，2012年。Elton P. Hsu。流形上的随机分析。数学研究生教材。美国数学会，2002年。Chin-Wei Huang、Milad Aghajohari、Avishek Joey Bose、Prakash Panangaden 和 Aaron C. Courville。黎曼扩散模型。载于神经信息处理系统进展，第35卷，第2750–2761页，2022年。Jaehyeong Jo 和 Sung Ju Hwang。通过黎曼扩散过程混合实现流形上的生成建模。载于第*41*届 *International Conference on Machine Learning* 论文集，第235卷机器学习研究会议录，第22348–22370页。PMLR，2024年。

Jaehyeong Jo 和 Sung Ju Hwang。用于语言建模的连续扩散模型。载于神经信息处理系统进展，2025年。Erik Jørgensen。测地随机游走的中心极限定理问题。*Zeitschrift f¨urWahrscheinlichkeitstheorie und verwandte Gebiete*，32(1):1–64，1975年。J¨urgen Jost。黎曼几何与几何分析。施普林格，2008年。Ioannis Karatzas 和 Steven E. Shreve。布朗运动与随机微积分，数学研究生教材第113卷。施普林格，1991。Peter E. Kloeden 和 Eckhard Platen。随机微分方程数值解。施普林格，1992年。Kazumasa Kuwada。梯度估计与 Wasserstein控制的对偶性。泛函分析杂志，258(11):3758–3774，2010年。Kazumasa Kuwada。时间非齐次测地随机游走的收敛及其在耦合方法中的应用。概率年刊，40(5):1945–1979，2012。Adrien Laurent 和 Gilles Vilmart。流形上遍历随机微分方程不变测度采样的阶条件。计算数学基础，22(3):649–695，2022。Jean-Fran。cois Le Gall。布朗运动、鞅与随机微积分。数学研究生教材。施普林格，2016。John M. Lee。光滑流形导论，数学研究生教材第218卷。数学研究生教材。施普林格，2012。John M. Lee。黎曼流形导论，数学研究生教材第2卷。数学研究生教材。施普林格，2018。Gunther Leobacher 和 Alexander Steinicke。投影到可微流形上的存在性、唯一性与正则性。全球分析与几何年刊，60(3): 559–587，2021年。Gen Li 和 Yuling Yan。适应基于评分的扩散模型中的未知低维结构。载于神经信息处理系统进展，第37卷，第126297–126331页，2024年。Mufan Li 和 Murat A. Erdogdu。用于求解半定规划的黎曼朗之万算法。*Bernoulli*，29(4):3093–3113，2023年。Xue-Mei Li。随机微分方程的强 p 完备性及非紧流形上光滑流的存在性。概率理论与相关领域，100(4): 485–511，1994年。Gabriel Loaiza-Ganem、Brendan Leigh Ross、Rasa Hosseinzadeh、Anthony L. Caterini 和 Jesse C. Cresswell。从流形假设视角看深度生成模型：综述与新的联系。机器学习研究汇刊，2024年。

Carlo Marinelli 和 Michael R¨ockner。Burkholder、Davis 和 Gundy 极大不等式。*Expositiones Mathematicae*，34(1):1–26，2016。Goran Marjanovic 和 Victor Solo。简化矩阵李群上随机微分方程数值方法的研究。*IEEE Transactions on Automatic Control*，63(12): 4035–4050，2018。Goran Marjanovic、Marc J. Piggott 和 Victor Solo。李群上随机微分方程数值方法的一种简单方法。载于*2015*年 *IEEE* 第*54*届决策与控制会议（*CDC*），第7143–7150页。IEEE，2015。Grigori N. Milstein 和 Michael V. Tretyakov。数学物理随机数值方法。Scientific Computation。施普林格，2013。Robb J. Muirhead。多元统计理论的若干方面。Wiley Series in Probability and Statistics。威利，2009。Michelle Muniz、Matthias Ehrhardt、Michael G¨unther 和 Renate Winkler。矩阵李群上线性 Itô 随机微分方程的高强阶方法。*BIT Numerical Mathematics*，62(4):1095–1119，2022。Michelle Muniz、Matthias Ehrhardt、Michael Guenther 和 Renate Winkler。流形上非线性 Itô 随机微分方程的强随机 Runge–Kutta–Munthe-Kaas 方法。*Applied Numerical Mathematics*，193:196–203，2023。James R. Munkres。拓扑学。Pearson，2013。John Nash。黎曼流形的嵌入问题。*Annals of Mathematics*，63 (1):20–63，1956。Kazusato Oko、Shunta Akiyama 和 Taiji Suzuki。扩散模型是极小极大最优分布估计器。载于第 *40* 届机器学习国际会议论文集，第202卷，机器学习研究会议录，第26517–26582页。PMLR，2023。Marc J. Piggott 和 Victor Solo。SO(n) 和 SE(n) 上随机微分方程的几何 Euler–Maruyama 格式。*SIAM Journal on Numerical Analysis*，54(4): 2490–2516，2016。Daniel Platt 和 Ra´ul S´anchez Gal´an。可达半径的下界及其应用。*arXiv*预印本 *arXiv:2505.08427*，2025年。网址：https://arxiv.org/abs/2505.08427。Victor Solo。李群上的 Stratonovich、Itô 与数值分析。载于*2024*年 *IEEE* 第*63*届决策与控制会议（*CDC*），第4328–4333页。IEEE，2024年。Victor Solo 和 Zhichao Wang。球面上随机微分方程的切空间数值格式的收敛性。载于*2021*年 *IEEE* 第*60*届决策与控制会议（*CDC*），第5814–5819页。IEEE，2021年。

Yang Song、Jascha Sohl-Dickstein、Diederik P. Kingma、Abhishek Kumar、Stefano Ermon 和 Ben Poole。基于随机微分方程的基于分数的生成建模。载于国际学习表征会议，2021年。Daniel W. Stroock。黎曼流形上路径分析导论。数学调查与专著。美国数学会，2000年。Clifford H. Taubes。微分几何：丛、联络、度量与曲率。牛津数学研究生教材。牛津大学出版社，2011年。Anton Thalmaier。随机黎曼几何，2023年。网址：https://math.uni.lu/ thalmaier/Stoch_Anal_2023/notes.pdf。讲义。Vidar Thom´ee。抛物型问题的*Galerkin*有限元方法。施普林格计算数学丛书。施普林格，2007。Loring W. Tu。微分几何：联络、曲率与特征类，数学研究生教材第275卷。施普林格，2017。Roman Vershynin。高维概率：数据科学中的应用导论。剑桥大学出版社，2018。C´edric Villani。最优传输：旧与新，数学基础原理第338卷。施普林格，2008年。Feng-Yu Wang。负曲率下扩散半群在Wasserstein距离中的指数收缩。*Potential Analysis*，53(3):1123–1144，2020年。Kevin Wang、Hongqian Niu、Yixin Wang 和 Didong Li。深度生成模型：复杂度、维数与逼近。机器学习研究杂志，26 (143):1–37，2025年。Xi Wang 和 Victor Solo。黎曼随机微分方程的保几何数值格式。*arXiv*预印本 *arXiv:2504.12631*，2025a。网址：https://arxiv. org/abs/2504.12631。Xi Wang 和 Victor Solo。SO(n)上随机微分方程的切空间参数化。*IEEE*控制系统快报，9:1369–1374，2025b。Xiao Wang、Qi Lei 和 Ioannis Panageas。流形上朗之万动力学的快速收敛：测地线与对数Sobolev相遇。载于神经信息处理系统进展，第33卷，第18894–18904页，2020年。Pavel Zorin-Kranich。鞅不等式，2022年。讲义，波恩大学（2021/22冬季学期）。可见：https://www.math.uni-bonn.de/~pzorin/lecture_notes/2022-martingale-inequalities.pdf。