任意输入下 Ziv-Zakai 界的函数性质

Minoh Jeong*、Alex Dytso† 和 Martina Cardone*
*明尼苏达大学，明尼阿波利斯，明尼苏达州 55455，{jeong316, mcardone}@umn.edu
†高通 Flarion Technology, Inc.，新泽西州布里奇沃特 08807，odytso2@gmail.com

arXiv:2305.02970v1 [cs.IT] 4 May 2023

摘要——本文探讨 Ziv-Zakai 界（ZZB），这是一种著名的最小均方误差（MMSE）贝叶斯下界。首先，本文证明，ZZB 不依赖于对待估参数分布的任何假设，即待估参数不一定需要具有概率密度函数。随后，本文进一步分析 ZZB 在高噪声和低噪声情形下的性质，并证明其始终具有张量化性质。最后，本文从多个方面考察 ZZB 的紧致性，例如假设的数量以及谷值填充函数的作用。特别是，本文给出了连续输入下该界紧致的充要条件，并证明对于支撑集在零点处没有聚点的离散输入分布，该界永远不会达到紧致。

I. 引言

均方误差（MSE）是许多贝叶斯估计问题事实上的标准度量。其最小值，即由条件期望实现的均方误差（MMSE），有助于深入了解估计问题的难度以及任意估计器的性能。对于任意估计问题，刻画 MMSE 或求出条件期望的闭式表达式通常是难以处理的。因此，人们往往需要依赖 MMSE 的下界。已有大量不同的下界被提出并得到分析。

我们不打算综述相关文献，仅提及几种值得注意的方法和几类下界。第一类下界基于柯西–施瓦茨不等式，Weiss–Weinstein 在 [1] 中对其进行了全面研究。值得注意的是，广为人知的克拉美–罗界（也称为 Van Trees 界 [2]），以及其他若干界，例如 Bobrovsky–Zakai 界 [3] 和 Bhattacharyya 界 [4]，都属于这一类。第二类界基于最大熵原理，利用条件熵给出均方误差（MMSE）的界；参见 [5]。第三类界基于函数不等式，例如 Poincaré 不等式 [6] 和对数 Sobolev 不等式 [7]。第四类界基于信息散度的变分表示，例如 f-散度和 χ^2-散度（由此得到贝叶斯 Hammersley–Chapman–Robbins 界）；参见 [8]–[12]。最后，第五类界也是本文研究的对象，称为 Ziv–Zakai 界[1]（ZZB），它通过与 M 元假设检验问题 [1]的联系给出均方误差（MMSE）的下界。

[1]在非参数统计中，利用假设检验联系得到的下界通常被称为 Le Cam 方法 [13]；另见 [14]。

通过与 M 元假设检验问题 [15], [16]的联系，给出均方误差（MMSE）的下界。

Bell 等人在 [16] 中给出了目前最现代且最紧的 ZZB 版本。特别地，[16] 的主要贡献之一是将该下界扩展到了向量情形；在此之前，该下界仅适用于标量情形。ZZB 的主要优势之一在于它只有一个正则性条件：待估参数需要具有概率密度函数（pdf）。因此，ZZB 比克拉美–罗界更为一般，后者要求被估参数的概率密度函数满足若干光滑性假设。ZZB 也被认为是文献中最紧的下界之一；关于其在高噪声区域中的紧致性分析，见 [17] 。此外，尽管其表达式略显繁琐，但该下界通常并不难计算。由于这些优势，ZZB 已得到广泛应用，例如用于量子参数 [18]–[25], 、时延 [26], [27], 、到达时间 [28]–[31], 、位置 [32]–[34], 、到达方向 [35]–[38], 的估计，以及 MIMO 雷达系统 [39], [40]中的估计。

尽管 ZZB 得到了广泛应用，仍有几个重要问题尚未解决。例如，ZZB 包含一个自由参数 M，它对应于界中辅助假设的数量。随着 M 增大，该界通常会变得更紧，但计算也会更加困难。因此，有必要为从业者提供选择参数 M 的经验法则。ZZB 还有两个版本：一个采用所谓的谷值填充函数，另一个则不采用。引入谷值填充函数会使该界更紧，但也会增加计算复杂度。此外，谷值填充函数如何与参数 M 相互作用，以及二者中哪一个对提高 ZZB 的紧致性贡献更大，目前也尚不明确。另一个问题是，ZZB 成立的前提是待估参数必须具有概率密度函数，因此无法用于离散分布或混合分布的情形。混合分布在许多应用中，无论在实践还是理论上都发挥着重要作用；例如，见 [41] ，其中混合分布被用于压缩感知。因此，消除这一条件将扩大该界的适用范围，使其具有普适性（即不再要求任何正则性条件）。最后，提供有关 ZZB 紧致性的指导也很重要。关于其紧致性的一些初步结果已在 [16] 和 [17], 中得出

而在这项工作中，我们在这一方面取得了进一步进展。

A. 贡献与论文结构

我们推导并阐述了 ZZB 的一般性质。具体而言，在第二节中，我们证明了 ZZB 的成立无需对待估参数的分布作任何假设。在同一节中，我们还刻画了对于相当一般的一组噪声分布，ZZB 在高噪声下的行为，以及在具有实际意义的加性高斯噪声信道模型下，ZZB 在低噪声条件下的渐近性质。此外，我们在第二节中证明了 ZZB 的张量化。

在第三节中，我们从多个方面研究该界的紧致性，例如假设的数量以及谷值填充函数的有效性。我们首先关注连续待估参数，并给出 ZZB 取得紧界的充要条件。对于单变量待估参数，这改进了 [16] 中的结果：我们证明，[16] 中的充分条件对于该界取得紧界而言也是必要的。随后，我们将注意力转向离散待估参数。令人惊讶的是，对于离散情形，我们证明谷值填充函数是必不可少的工具；否则，ZZB 将退化为零。我们还证明，即使使用了谷值填充函数，对于支持集在零点处没有聚点的离散待估参数，ZZB 也永远无法取得紧界。

B. 符号说明

粗体大写字母 $\mathbf{X}$ 表示随机向量；粗体小写字母 $\mathbf{x}$ 表示 $\mathbf{X}$ 的一个特定实现；X_i 和 x_i（或 $(\mathbf{x})_i$）分别表示 $\mathbf{X}$ 和 $\mathbf{x}$ 的第 i 个元素；$\mathbf{e}_i$ 是第 i 个标准基向量，其第 i 个条目为 1，所有其他条目均为 0；$[n_1 : n_2]$ 表示从 n_1 到 $n_2 \geq n_1$ 的整数集合；$\varnothing$ 表示空集；花体字母 $\mathcal{X}$ 表示集合/事件；$\mathsf{d}(\mathcal{X})$ 表示 $\mathbf{x} \in \mathcal{X}$ 中元素的数量；$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_i x_i y_i$ 表示 $\mathbf{x}$ 与 $\mathbf{y}$ 之间的内积；I_n 表示维度为 n 的单位矩阵。对于一对随机向量 $(\mathbf{X},\mathbf{Y})$，我们令 $P_{\mathbf{X}}(\cdot)$ 表示 $\mathbf{X}$ 的分布，并令 $P_{\mathbf{Y}|\mathbf{X}}(\cdot|\mathbf{x})$ 表示支配噪声观测模型 $\mathbf{Y}|\mathbf{X}=\mathbf{x}$ 的分布。对于函数 $f : \mathbb{R} \to \mathbb{R}$，谷值填充函数定义为

$$\mathcal{V}_t\{f(t)\} = \sup_{u:u\geq t} f(u). \tag{1}$$

我们假设随机向量 $\mathbf{X} \in \mathcal{X}$ 是需要估计的参数，而随机向量 $\mathbf{Y} \in \mathcal{Y}$ 是 $\mathbf{X} \in \mathcal{X}$ 经由信道分布 $P_{\mathbf{Y}|\mathbf{X}}$ 产生的含噪观测。根据 $\mathbf{Y} \in \mathcal{Y}$ 估计 $\mathbf{X}$ 的均方误差（MMSE）定义为

$$\mathrm{mmse}(\mathbf{X}|\mathbf{Y}) = \sum_{i=1}^{\mathsf{d}(\mathcal{X})} \mathbb{E}\big[(X_i - \mathbb{E}[X_i|\mathbf{Y}])^2\big]. \tag{2}$$

II. 任意输入下的 ZZB、渐近性质与张量化

介绍 ZZB 需要先定义 M 元假设检验设置中的若干要素。

定义1：设 $\mathcal{X}$ 为一个向量空间，$\mathcal{Y}$ 为任意空间。对于随机向量 $\mathbf{X} \in \mathcal{X}$ 和 $\mathbf{Y} \in \mathcal{Y}$，令 $P_{\mathbf{Y}|\mathbf{X}}(\cdot|\mathbf{x})$ 为 $\mathbf{Y}$ 在给定 $\mathbf{X}=\mathbf{x}$ 时的条件概率分布。对于整数 $\mathsf{M} \geq 2$ 和 $\mathbf{x} \in \mathcal{X}$，考虑以下 M 元假设检验问题，

$$\mathcal{H}_i : \mathbf{Y} \sim P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}+\mathbf{u}_i),\ i \in [0:\mathsf{M}-1], \tag{3a}$$

其中，$\{\mathbf{u}_i\}_{i=0}^{\mathsf{M}-1}$ 是 $\mathcal{X}$中的某个元素集合，并且

$$\Pr(\mathcal{H}_i) = p_i. \tag{3b}$$

对于一组概率 $\mathcal{P} = \{p_i\}_{i=0}^{\mathsf{M}-1}$ 和一组向量 $\mathcal{U} = \{\mathbf{u}_i\}_{i=0}^{\mathsf{M}-1} \subset \mathcal{X}$，该 M 元假设检验问题的最小错误概率（由最优的最大后验（MAP）决策规则 [42]实现）记作 $P_e(\mathbf{x};\mathcal{P},\mathcal{U})$。

A. 一般 ZZB

我们提出了 ZZB 的最一般形式。这个新版本不对联合分布$P_{\mathbf{X},\ \mathbf{Y}}$施加任何限制。其证明主要依赖于 [16], 中提出的思想，而该方法假设 $\mathbf{X}$ 必须具有概率密度函数。通过对勒贝格积分中的各项进行更细致的核算，我们将这些结果推广到 $\mathcal{X}$ 上的任意概率测度，并证明该界适用于任意有限维向量空间 $\mathcal{X}$。

定理1：设 $\mathcal{X}$ 是一个向量空间。那么，对于任意整数 $\mathsf{M} \geq 2$, ZZB 为

$$\mathrm{mmse}(\mathbf{X}|\mathbf{Y}) \geq \overline{\mathsf{ZZ}}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}), \tag{4}$$

with

$$\overline{\mathsf{ZZ}}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}) = \sum_{i=1}^{\mathsf{d}(\mathcal{X})} \int_0^\infty \frac{t}{2} \mathcal{V}_t\left\{ \frac{\mathsf{h}(t,i,\mathsf{M})}{\mathsf{M}-1} \right\} \mathrm{d}t, \tag{5}$$

其中，对于 $t>0$ 和 $i \in [0:\ \mathsf{M}-1]$，我们有、】【

$$\mathsf{h}(t,i,\mathsf{M}) = \max_{\substack{\mathcal{U}\subset\mathcal{X}:\\ \mathcal{U}=\{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1}\\ (\mathbf{u}_k)_i = kt,\ \forall k}} \int P_e(\mathbf{x};\mathcal{P}_{\mathcal{U}}(\mathbf{x}),\mathcal{U})\, \mathrm{d}\mu_{\mathcal{U}}(\mathbf{x}). \tag{6}$$

为定义 $h(t,i,\mathsf{M})$，我们令：

- 对于集合 $\mathcal{U} = \{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1} \subset \mathcal{X}$，测度 $\mu_{\mathcal{U}}$ 由下式给出

$$\mu_{\mathcal{U}} = \sum_{k=0}^{\mathsf{M}-1} P_{\mathbf{X}-\mathbf{u}_k}; \tag{7}$$

- 对于 $\mathbf{x} \in \mathcal{X}$ 和 $\mathcal{U} \subset \mathcal{X}$，我们令[2]

$$\mathcal{P}_{\mathcal{U}}(\mathbf{x}) = \left\{ p_i : p_i = \frac{\mathrm{d}P_{\mathbf{X}-\mathbf{u}_i}}{\mathrm{d}\mu_{\mathcal{U}}}(\mathbf{x}), \mathbf{u}_i \in \mathcal{U} = \{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1} \right\}. \tag{8}$$

证明：证明见附录 A。 ■

注1：定理1中的ZZB适用于任意向量空间 $\mathcal{X}$，只要按照式(2)对$\mathbf{X} \in \mathcal{X}$ 的均方误差（MMSE）（X|Y）进行了适当定义。例如，$\mathbf{X} \in \mathcal{X}$ 可以服从任意有限维向量空间上的混合分布，例如向量、矩阵或张量。此外，值得注意的是，去除式(5)中的谷值填充函数进一步

[2]根据构造，Radon–Nikodym 导数总是有良好定义的。

放宽了界限，同时也使其更加简单。不含谷值填充函数的版本定义为，

$$\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}) = \sum_{i=1}^{\mathsf{d}(\mathcal{X})} \int_0^\infty \frac{t}{2} \frac{\mathsf{h}(t, i, \mathsf{M})}{\mathsf{M}-1} \, \mathrm{d}t. \tag{9}$$

尽管定理1中的ZZB表达式相当复杂，但只要能得到 P_e(x; $\mathcal{P}_{\mathcal{U}}$(x) $\mathcal{U}$)的精确表达式（或近似表达式），就可以轻松计算。然而，对于每个 $P_{\mathbf{Y}|\mathbf{X}}$ ，计算定理1中的ZZB可能并不容易。为了了解其紧致性，可以考察它在渐近区域内是否紧致。接下来将刻画ZZB在高噪声和低噪声下的渐近行为。

B. 高噪声渐近

这里，我们研究高噪声区域内的定理1中的ZZB。具体而言，我们在以下假设下刻画其渐近行为（正如 [17]中所讨论的，这些假设相当自然）：

- A1: $P_{\mathbf{Y}|\mathbf{X}}$ 可由 $\eta \geq 0$参数化， $\eta \geq 0$称为噪声水平，即对于所有 (x, y) $\in (\mathcal{X}, \mathcal{Y})$，都有 $P_{\mathbf{Y}|\mathbf{X}}$(y|x; η) 。为突出对 η的依赖，我们令 P_e(x; $\mathcal{P}_{\mathcal{U}}$(x) $\mathcal{U}$) = P_e(η, x; $\mathcal{P}_{\mathcal{U}}$(x) $\mathcal{U}$)表示定义1中的最优错误概率。
- A2: P_e(η, x; $\mathcal{P}_{\mathcal{U}}$(x) $\mathcal{U}$) 关于 η 单调不减。
- A3: $\lim_{\eta\to\infty} P_e(\eta, \mathrm{x}; \mathcal{P}_{\mathcal{U}}(\mathrm{x})\, \mathcal{U}) = 1 - \max_{p\in\mathcal{P}_{\mathcal{U}}(\mathbf{x})} p$.

接下来，我们定义 ZZB 的高噪声渐近。

定义2：定理1中 ZZB 的高噪声渐近定义为

$$\overline{\mathsf{V}}(P_{\mathbf{X}}, \mathsf{M}) = \lim_{\eta\to\infty} \overline{\mathsf{ZZ}}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}), \tag{10}$$

$$\mathsf{V}(P_{\mathbf{X}}, \mathsf{M}) = \lim_{\eta\to\infty} \mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}). \tag{11}$$

下一个定理刻画了定义2中的渐近性质。定理 2：对于任意 $P_{\mathbf{X}}$ 以及任意 $\mathsf{M} \geq 2$，我们有

$$\overline{\mathsf{V}}(P_{\mathbf{X}}, \mathsf{M}) = \sum_{i=1}^{\mathsf{d}(\mathcal{X})} \int_0^\infty \frac{t}{2} \mathcal{V}_t\left\{\frac{\mathsf{M} - \mathsf{H}(P_{\mathbf{X}}, t, i, \mathsf{M})}{\mathsf{M}-1}\right\} \mathrm{d}t,$$

$$\mathsf{V}(P_{\mathbf{X}}, \mathsf{M}) = \sum_{i=1}^{\mathsf{d}(\mathcal{X})} \int_0^\infty \frac{t}{2} \frac{\mathsf{M} - \mathsf{H}(P_{\mathbf{X}}, t, i, \mathsf{M})}{\mathsf{M}-1} \, \mathrm{d}t,$$

其中

$$\mathsf{H}(P_{\mathbf{X}}, t, i, \mathsf{M}) = \min_{\substack{\mathcal{U}\subset\mathcal{X}:\\ \mathcal{U}=\{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1}\\ (\mathbf{u}_k)_i = kt,\ \forall k}} \int \max_{j\in[0:\mathsf{M}-1]} \mathrm{d}P_{\mathbf{X}-\mathbf{u}_j}(\mathbf{x}).$$

证明：证明见附录B。值得注意的是，在高噪声水平下，我们有 ■

$$\lim_{\eta\to\infty} \mathrm{mmse}(\mathbf{X}|\mathbf{Y}) = \sum_{i=1}^{\mathsf{d}(\mathcal{X})} \mathrm{Var}(X_i), \tag{12}$$

因此，如果 V($P_{\mathbf{X}}$,M) 或 V($P_{\mathbf{X}}$,M) 中任一量等于方差，那么在此区域内 ZZB 就是紧的。在 [17], 中，作者给出了几个单变量情形的示例，说明 ZZB 在高噪声区域内何时紧、何时不紧。

现在，我们通过说明定理 2 有助于理解 M 的影响，来结束对高噪声区域的分析。具体而言，我们给出一个输入分布示例，对于该分布，$\mathsf{M} > 2$ 不会使界相较于 $\mathsf{M} = 2$有所改善。

示例。令 f_X 为 X 的概率密度函数，并假设 f_X 是单峰的。那么，

$$\overline{\mathsf{V}}(P_X, \mathsf{M}) = \overline{\mathsf{V}}(P_X, 2), \text{ and } \mathsf{V}(P_X, \mathsf{M}) = \mathsf{V}(P_X, 2),$$

这可以通过利用附录F中的引理1加以证明。

C. 低噪声渐近性

在此，我们研究低噪声条件下的 ZZB，即当$\eta \to 0$时的情形。我们首先注意到，lim$_{\eta\to 0}$均方误差（MMSE）(X|Y) $= 0$，因此，我们关注的是刻画其收敛速率。然而，由于这种收敛速率高度依赖于噪声分布，因此我们在此聚焦于具有实际意义的加性高斯噪声信道。具体而言，我们考虑

$$\mathbf{Y} = \mathbf{X} + \mathbf{N}, \text{ where } \mathbf{N} \sim \mathcal{N}(\mathbf{0}, \eta I_d), \tag{13}$$

其中， $d = \mathsf{d}(\mathcal{X})$，而 η 是第II-B节中定义的噪声水平。下一个定理证明了模型（13）所对应信道的 ZZB 低噪声渐近性。

定理 3：考虑任意满足以下条件的 $P_{\mathbf{X}}$

$$P_{\mathbf{X}} = \alpha P_{\mathbf{X}_C} + (1-\alpha) P_{\mathbf{X}_D}, \ 0 \leq \alpha \leq 1, \tag{14}$$

其中， $P_{\mathbf{X}_C}$ 是相对于 d 维勒贝格测度绝对连续的分布，而 $P_{\mathbf{X}_D}$ 是纯不连续分布。那么，对于式 (13) 中的信道模型，有

$$\lim_{\eta\to 0} \frac{\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, 2)}{\eta} = \alpha \mathsf{d}(\mathcal{X}). \tag{15}$$

证明：证明见附录 C。 ■

注 2：文献 [43], [44], 已证明，对于式 (13) 中的信道模型以及式 (14) 中的 $P_{\mathbf{X}}$ ，有

$$\lim_{\eta\to 0} \frac{\mathrm{mmse}(\mathbf{X}|\mathbf{Y})}{\eta} = \alpha \mathsf{d}(\mathcal{X}). \tag{16}$$

因此，定理 3 表明，在式 (13) 的信道模型下，ZZB 在低噪声区间是紧的。此外，在低噪声区间，ZZB 可采用式 (9) 中最简单的形式，其中 $\mathsf{M} = 2$。此外，我们注意到，对于连续分布，有许多界在低噪声区间是紧的（例如，克拉美-罗界）；但据我们所知，除 ZZB 外，没有其他针对混合分布的下界在低噪声区间是紧的。我们还注意到，混合分布在压缩感知应用中很重要（例如，参见 [41]）。

D. 张量化

由于 ZZB 的表达式较为繁杂，尤其是式 (6) 中包含一层内层优化，因此有必要了解该界是否能够简化。下一条命题也为本节作结，它确实表明 ZZB 可以张量化。

命题 1 $P_{\mathbf{Y}|\mathbf{X}} =$: 每当 $P_{\mathbf{X}} = \prod_{i=1}^{d} P_{X_i}$ 且具有 $\mathsf{d}(\,)$ 时，成立的是

$$\overline{\mathsf{ZZ}}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}) = \sum_{i=1}^{d} \overline{\mathsf{ZZ}}(P_{X_i}, P_{Y_i|X_i}, \mathsf{M}), \tag{17a}$$

$$\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}) = \sum_{i=1}^{d} \mathsf{ZZ}(P_{X_i}, P_{Y_i|X_i}, \mathsf{M}). \tag{17b}$$

Proof: The proof is provided in Appendix D. ■

III. 紧致性分析

本节研究定理1中 ZZB 在 X 的各种分布设定下的紧致性，即连续分布和离散分布。

A. 连续分布

假设 $P_{\mathbf{X}}$ 和 $P_{\mathbf{Y}|\mathbf{X}}$ 分别为具有概率密度函数 $f_{\mathbf{X}}$ 和 $f_{\mathbf{Y}|\mathbf{X}}$ 的分布。在下面的命题[3], 中，我们给出了关于 $f_{\mathbf{X}|\mathbf{Y}}$ 的充要条件，该条件使得关于 mmse(X|Y) 的 ZZB 达到紧致。

命题 2：对于 $\mathbf{X} \in \mathcal{X}$ 和 $\mathbf{Y} \in \mathcal{Y}$，有

$$\mathrm{mmse}(\mathbf{X}|\mathbf{Y}) = \mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}), \tag{18}$$

当且仅当对于所有 $t > 0$，存在 $\mathcal{U}_i = \{\mathbf{u}_{i,\,k}\}_{k=0}^{\mathsf{M}-1} \subset \mathcal{X}$, $\forall i \in [1 : \mathsf{d}(\mathcal{X})]$,，其中每一个都满足 $(\mathbf{u}_{i,\,k})_i = kt$, $k \in [0 : \mathsf{M}-1]$，并且对于所有 $(\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{Y}$,

$$\begin{aligned} &\operatorname*{argmax}_{k\in[0:\mathsf{M}-1]} f_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}+\mathbf{u}_{i,k}|\mathbf{y}) \\ &\cap \operatorname*{argmin}_{k\in[0:\mathsf{M}-1]} |\mathbb{E}[X_i|\mathbf{Y}=\mathbf{y}] - x_i - kt| \neq \varnothing. \end{aligned} \tag{19}$$

证明：首先注意到，定理1中的 ZZB 是通过分别对每个对均方误差（MMSE）(X|Y) 有贡献的$d = \mathsf{d}(\mathcal{X})$项取下界推导出来的（见 (2)）。因此，(18) 成立当且仅当对均方误差（MMSE）(X|Y) 有贡献的第 i项等于对 $\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M})$ 有贡献的第 i项，其中 $i \in [1 : d]$。

现在，考虑 (9) 中对 $\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M})$ 有贡献的第 i项。该项的推导仅对 $\mathsf{h}(t, i, \mathsf{M})$ 中的错误概率取了界（见附录 A 中的 (44)），如下：

$$P_e^{\phi}(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}), \mathcal{U}) \geq P_e(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}), \mathcal{U}), \tag{20}$$

其中，$P_e^{\phi}(\mathbf{x};\, \mathcal{P}_{\mathcal{U}}(\mathbf{x})\, \mathcal{U})$ 是采用次优决策规则 $\phi(\mathbf{Y})$ 所产生的错误概率，该规则定义为

$$\phi(\mathbf{Y}) = \mathcal{H}_j, \text{ where } j \in \operatorname*{argmin}_{k\in[0:\mathsf{M}-1]} |\mathbb{E}[X_i|\mathbf{Y}] - x_i - kt|, \tag{21}$$

而 $P_e(\mathbf{x};\, \mathcal{P}_{\mathcal{U}}(\mathbf{x})\, \mathcal{U})$ 是最小错误概率，即采用最优 MAP 决策规则 $\phi^{\star}(\mathbf{Y})$ 所产生的错误概率，其定义为

$$\phi^{\star}(\mathbf{Y}) = \mathcal{H}_j, \text{ where } j \in \operatorname*{argmax}_{k\in[0:\mathsf{M}-1]} f_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}+\mathbf{u}_k|\mathbf{Y}). \tag{22}$$

换言之，$P_e^{\phi}(\mathbf{x};\, \mathcal{P}_{\mathcal{U}}(\mathbf{x})\, \mathcal{U})$ 得出 $\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M})$，而 $P_e(\mathbf{x};\, \mathcal{P}_{\mathcal{U}}(\mathbf{x})\, \mathcal{U})$ 得出均方误差（MMSE）(X|Y)。因此，等式当且仅当对于所有$i \in [1 : d]$，式 (20) 等号成立时，式 (18) 才成立；因此，当且仅当对于所有 $i \in [1 : d]$，$\phi = \phi^{\star}$。

现在，由于 $\mathsf{h}(t, i, \mathsf{M})$ 在 $\mathcal{U} = \{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1} \subset \mathcal{X}$ 上取最大值，对于所有 $t > 0$，至少需要存在一个集合 $\mathcal{U}$（回想一下，t 是 $\mathcal{U}$ 的约束），使得对于所有 $\mathbf{y} \in \mathcal{Y}$，$\phi(\mathbf{y}) = \phi^{\star}(\mathbf{y})$。因此，对 $\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M})$ 有贡献的第 i 个元素达到紧致的充分必要条件是：对于所有 $t > 0$，存在 $\mathcal{U}_i = \{\mathbf{u}_{i,k}\}_{k=0}^{\mathsf{M}-1}$，使得 $(\mathbf{u}_{i,k})_i = kt$, $k \in [0 : \mathsf{M}-1]$，并且对于所有 $(\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{Y}$,

$$\begin{aligned} &\operatorname*{argmax}_{k\in[0:\mathsf{M}-1]} f_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}+\mathbf{u}_{i,k}|\mathbf{Y}=\mathbf{y}) \\ &\cap \operatorname*{argmin}_{k\in[0:\mathsf{M}-1]} |\mathbb{E}[X_i|\mathbf{Y}=\mathbf{y}] - x_i - kt| \neq \varnothing. \end{aligned} \tag{23}$$

因此，(18) 成立当且仅当对于所有 $i \in [1 : d]$，都存在这样的 $\mathcal{U}_i$：d。命题 2 的证明至此结束。 ■

注 3：换言之，(19) 中的条件意味着，在 $\mathcal{S} = \{\mathbf{x} + \mathbf{u}_{i,k}\}_{k=0}^{\mathsf{M}-1}$ 上，MAP 估计的第 i 个元素应接近给定 Y 时 X_i 的均方误差（MMSE）估计。

借助命题 2，我们现在强化 [16]，中的一个结果：具体而言，我们证明，[16] 中已证明是 ZZB 紧致性的充分条件，实际上也是必要条件。

推论 1：设 $X \in \mathbb{R}$ 为连续随机变量。则

$$\mathrm{mmse}(X|\mathbf{Y}) = \mathsf{ZZ}(P_X, P_{\mathbf{Y}|X}, \mathsf{M}), \tag{24}$$

当且仅当对所有 $\mathbf{y} \in \mathcal{Y}$，概率密度函数 $f_{X|\mathbf{Y}}(x|\mathbf{y})$ 是单峰的[4]，且关于其众数对称。

证明：证明见附录 E。 ■

最后，我们指出，目前尚不清楚使该界与谷值填充函数精确相等的充分必要条件，这也是一个值得进一步研究的有趣方向。

B. 离散分布

ZZB 通常是在假设 X 的分布具有概率密度函数 [16] 的情况下进行研究的。离散随机变量的情形通常被忽略，而且往往被错误地认为只需将概率密度函数替换为相应的概率质量函数（pmf）即可，十分简单。本文这一部分旨在说明，处理离散分布时需要格外谨慎，这也是定理1以测度论方式表述的部分原因。

现在我们给出以下一系列结果。首先，我们证明，不带谷值填充函数的 ZZB 始终等于零。其次，我们给出一个简单的例子，说明带有谷值填充函数的界不等于零。这两个结果表明，对于离散输入，谷值填充函数是 ZZB 不可或缺的组成部分。第三，这一系列结果的最后一项表明，即使带有谷值填充函数，ZZB 对于离散输入仍严格次优。

[3]我们提醒读者，通常情况下，argmax 函数输出的是一个集合。

[4]当概率密度函数有多个局部极大值时，通常将它们统称为分布的众数。

命题 3：假设 X 是离散的。那么，

$$\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}) = 0. \tag{25}$$

证明：令 $p_{\mathbf{X}}$ 表示 $\mathbf{X} \in \mathcal{X}$ 关于 $d = \mathsf{d}(\mathcal{X})$ 的概率质量函数。对于离散型 X，根据定理1，有

$$\mathsf{h}(t, i, \mathsf{M}) = \max_{\substack{\mathbf{u}_k \in \mathcal{X}, \forall k \\ (\mathbf{u}_k)_i = kt}} \sum_{\mathbf{x} \in \overline{\mathcal{X}}} \sum_{k=0}^{\mathsf{M}-1} p_{\mathbf{X}-\mathbf{u}_k}(\mathbf{x}) P_e(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}), \mathcal{U}), \tag{26}$$

其中 $\overline{\mathcal{X}} = \bigcup_{i=1}^{d} \{\mathcal{X} - \mathbf{u}_i\}$。我们将证明，对于 $t \geq 0$，(26) 中的 $P_e(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}) \ \mathcal{U})$ 几乎处处等于零；结合 (9) 和 (26)，这意味着 $\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}) = 0$。

注意，给定 $t \geq 0$ 以及 $\mathbf{x} \in \mathcal{X}$，使 $P_e(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}) \ \mathcal{U})$ 等于零的一个充分条件是存在某个 $j \in [0 : \mathsf{M}-1]$，使得 $p_j = 1$，其中 $p_j \in \mathcal{P}_{\mathcal{U}}(\mathbf{x})$。事实上，在此条件下，最优决策规则总是能够判定出正确的假设，从而使错误概率为零。

根据定理1中 $p_i \in \mathcal{P}_{\mathcal{U}}(\mathbf{x})$ 的定义，有

$$p_i = \frac{p_{\mathbf{X}-\mathbf{u}_i}(\mathbf{x})}{\sum_{k=0}^{\mathsf{M}-1} p_{\mathbf{X}-\mathbf{u}_k}(\mathbf{x})}, \tag{27}$$

因此，存在一个 $j \in [0 : \mathsf{M}-1]$，使得 $p_j = 1$ 当且仅当 $\bigcup_{\mathcal{I} \subseteq [0:\ \mathsf{M}-1]:\ |\mathcal{I}| \geq 2} \bigcap_{k \in \mathcal{I}} \{\mathcal{X} - \mathbf{u}_k\} = \varnothing$。

因此，我们有如下包含关系：对于给定的 $\mathbf{x} \in \mathcal{X}$

$$\mathcal{A} \subseteq \{t \geq 0 : P_e(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}), \mathcal{U}) = 0\}, \tag{28}$$

其中

$$\mathcal{A} = \Bigg\{ t \geq 0 : \bigcup_{\substack{\mathcal{I} \subseteq [0:\mathsf{M}-1]: \\ |\mathcal{I}| \geq 2}} \bigcap_{k \in \mathcal{I}} \{\mathcal{X} - \mathbf{u}_k\} = \varnothing, \mathbf{u}_k \in \mathcal{X}, \\ (\mathbf{u}_k)_i = kt, \ k \in [0 : \mathsf{M}-1] \Bigg\}. \tag{29}$$

现在我们要证明，上述 $\mathcal{A}$ 是一个满测集，即 $\mathcal{A}$ 在 $[0, \infty)$ 上的补集具有零勒贝格测度。为此，在附录 F-D 的引理 4 中，我们证明 $\mathcal{A}^c$（即 $\mathcal{A}$ 的补集）是可数的，因此 $\mathcal{A}^c$ 的测度为零，正如所需。这意味着，对于 $t > 0$，式（26）中的 $P_e(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}) \ \mathcal{U})$ 几乎处处等于零，因此式（9）中的积分等于零，从而 $\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}) = 0$。命题 3 的证明至此完成。 ■

为了证明该系列中剩余的两个结果，只需关注高噪声区域中的标量输入情形，即考虑由定理 2刻画的 $\mathsf{V}(P_X, \mathsf{M})$。下一个例子（其证明见附录 G）表明，一般而言，$\mathsf{V}(P_X, \mathsf{M}) \neq 0$。

例：设 $X \sim \mathsf{Ber}(p)$ 是参数为 $0 < p < 1$ 的伯努利随机变量。则，

$$\overline{\mathsf{V}}(P_X, 2) = \frac{1}{4} \min\{p, 1-p\}. \tag{30}$$

下一个结果表明，即使使用谷值填充函数，ZZB 仍然严格次优。

定理 4：设 X 为离散随机变量，满足 $\mathbb{E}[X] = 0$ 且 $\inf_{x \in \mathcal{X}} |x| > 0$。那么，对于任意 M，均有

$$\overline{\mathsf{V}}(P_X, \mathsf{M}) < \mathsf{Var}(X). \tag{31}$$

证明：$\mathsf{Var}(X)$ 的另一种形式为

$$\mathsf{Var}(X) = \mathbb{E}[X^2] = \int_0^\infty \frac{t}{2} \Pr\left(|X| \geq \frac{t}{2} \right) \mathrm{d}t. \tag{32}$$

现在回顾一下，ZZB 是通过建立以下下界得到的：对于 $t > 0$，

$$\Pr\left(|X| \geq \frac{t}{2} \right) \geq \mathcal{V}_t \left\{ \frac{\mathsf{M} - \sum_{x \in \overline{\mathcal{X}}} \max_{k \in [0:\mathsf{M}-1]} p_X(x + kt)}{\mathsf{M} - 1} \right\}, \tag{33}$$

其中 $\overline{\mathcal{X}} = \bigcup_{j=0}^{\mathsf{M}-1} \{\mathcal{X} - jt\}$。下面我们证明，(33) 中的等式不成立，因此，$\mathsf{Var}(X) > \mathsf{V}(P_X, \mathsf{M})$。首先，令 $x_0 = \inf_{x \in \mathcal{X}} |x|$，并注意到 (33) 的左侧为

$$\Pr\left(|X| \geq \frac{t}{2} \right) = 1, \text{ if } t \leq x_0. \tag{34}$$

其次，假设 $t > 0$，并考虑（33）右侧的和。由于 $p_X(x + kt) = p_{X-kt}(x)$，我们有

$$\sum_{x \in \overline{\mathcal{X}}} \max_{k \in [0:\mathsf{M}-1]} p_X(x + kt) \overset{\text{(a)}}{>} \sum_{x \in \mathcal{X}} \max_{k \in [0:\mathsf{M}-1]} p_{X-kt}(x) \overset{\text{(b)}}{\geq} 1, \tag{35}$$

其中，(a) 根据 $\overline{\mathcal{X}} \setminus \mathcal{X} \neq \varnothing$ if $t > 0$ 得出，(b) 则通过去掉最大值并选择 $k = 0$ 得出。注意，(35) 对任意 $t > 0$ 都成立，因此，根据 (1) 中谷值填充函数的定义可知，对于所有 $t > 0$，(33) 的右侧都严格小于 1。结合 (34) 可知，存在一段 t 的取值范围，即 $0 < t \leq x_0$，使得 (33) 中的条件不以等号成立。定理 4 证毕。 ■

备注 4：在定理 4 中，假设 X 的支撑集在零点没有聚点，是为了简化证明；这一假设很可能可以去掉。此外，定理 4 中的结果表明，使用离散输入时，ZZB 可能不是最佳选择，尤其是在实际相关的高噪声区域。

IV. 结论

本文给出了一般 Ziv-Zakai 界的表达式，无需任何正则性假设。具体而言，我们去除了连续性假设，因此该界现在适用于所有输入分布。本文还指出，处理离散分布时必须谨慎，且谷值填充函数是此情形下必不可少的工具。本文还刻画了高噪声和低噪声渐近行为。这里的第一个关键观察是，在低噪声区域，对于加性高斯噪声，该界以最简单的形式（即不使用谷值填充函数，且假设数取二）是紧的。第二个关键观察是，在高噪声区域，情况更为复杂，该界可能并不紧。对于连续分布，本文给出了界紧的充要条件。最后，对于离散输入，本文表明该界在高噪声区域通常是次优的。最后这一观察表明，对于离散输入，要使 Ziv-Zakai 界更紧，要么需要一个新的界，要么需要对该界进行新的修正。

参考文献

[1] E. Weinstein 和 A. J. Weiss，“参数估计中的一类通用下界”，《IEEE 信息论汇刊》，第 34 卷，第 2 期，第 338–342 页，1988 年。[2] H. L. Van Trees，《检测、估计与调制理论，第一部分：检测、估计与线性调制理论》。John Wiley & Sons，2004 年。[3] B. Bobrovsky 和 M. Zakai，“某些扩散过程估计误差的一个下界”，《IEEE 信息论汇刊》，第 22 卷，第 1 期，第 45–52 页，1976 年。[4] A. Bhattacharyya，“关于信息量的若干类似量及其在统计估计中的应用”，Sankhy¯a：《印度统计学杂志》，第 1–14 页，1946 年。[5] T. Cover 和 J. Thomas，《信息论基础：第二版》。Wiley，2006 年。[6] I. Zieder、A. Dytso 和 M. Cardone，“通过庞加莱不等式得到的均方误差（MMSE）下界”，载于《2022 年 IEEE 国际信息论研讨会（ISIT）》，2022 年，第 957–962 页。[7] E. Aras、K.-Y. Lee、A. Pananjady 和 T. A. Courtade，“一族贝叶斯克拉美–罗界及其对对数凹先验的推论”，载于《2019 年 IEEE 国际信息论研讨会（ISIT）》。IEEE，2019 年，第 2699–2703 页。[8] S. Saito，“关于贝叶斯风险下界的元界”，载于《2022 年 IEEE 国际信息论研讨会（ISIT）》。IEEE，2022 年，第 3162–3167 页。[9] A. R. Esposito 和 M. Gastpar，“通过 Sibson 的 α互信息得到估计程序中贝叶斯风险的下界”，载于《2021 年 IEEE 国际信息论研讨会（ISIT）》。IEEE，2021 年，第 748–753 页。[10] X. Chen、A. Guntuboyina 和 Y. Zhang，“关于贝叶斯风险下界”，《机器学习研究杂志》，第 17 卷，第 1 期，第 7687–7744 页，2016 年。[11] A. Xu 和 M. Raginsky，“分散式估计中贝叶斯风险的信息论下界”，《IEEE 信息论汇刊》，第 63 卷，第 3 期，第 1580–1600 页，2016 年。[12] Y. Polyanskiy 和 Y. Wu，“信息论：从编码到学习”，2022 年。[13] L. LeCam，“维数限制下估计量的收敛性”，《统计年鉴》，第 38–53 页，1973 年。[14] A. B. Tsybakov，“非参数估计导论，2009”，网址 https://doi.org/10.1007/b13794。由该书修订扩充而来，第 9 卷，第 10 期，2004 年。[15] J. Ziv 和 M. Zakai，“信号参数估计的一些下界”，《IEEE 信息论汇刊》，第 15 卷，第 3 期，第 386–391 页，1969 年。[16] K. L. Bell，“参数估计中的性能界及其在方位估计中的应用”，博士论文，乔治梅森大学，1995 年。[17] A. Dytso、M. Cardone 和 I. Zieder，“Ziv-Zakai 界的高噪声渐近”，《IEEE 信号处理快报》，第 29 卷，第 1933–1937 页，2022 年。[18] M. Tsang，“量子参数估计的 Ziv-Zakai 误差界”，《物理评论快报》，第 108 卷，第 230401 页，2012 年 6 月。 [在线]。获取地址：https://link.aps.org/doi/10 1103 108 230401 。/PhysRevLett.。[19] V. Giovannetti 和 L. Maccone，“次海森堡估计策略无效”，《物理评论快报》，第 108 卷，第 210404 页，2012 年 5 月。[在线]。获取地址：https://link.aps.org/doi/10 1103 108 210404 。/PhysRevLett.。[20] Y. Gao 和 H. Lee，“通过量子 Ziv-Zakai 界得到的参数灵敏度广义极限”，《物理学杂志 A：数学与理论》，第 45 卷，第 41 期，第 415306 页，2012 年 9 月。 [在线]。获取地址：https://dx.doi.org/10 1088 1751 8113 45 41 415306 。/ - / / / Y.-R. Zhang 和 H. Fan，“向量参数的量子计量界”，《物理评论 A》，第 90 卷，第 043818 页，2014 年 10 月。 [在线]。获取地址：https://link.aps.org/doi/10 1103 90 043818 。/PhysRevA.。[22] D. W. Berry、M. Tsang、M. J. W. Hall 和 H. M. Wiseman，“用于波形估计的量子 Bell-Ziv-Zakai 界与海森堡极限”，《物理评论 X》，第 5 卷，第 031018 页，2015 年 8 月。 [在线]。获取地址：https://link.aps.org/doi/10 1103 5 031018 。/PhysRevX.。[23] J. Rubio、P. Knott 和 J. Dunningham，“克拉美–罗界之外量子计量方案的非渐近分析”，《物理学通讯杂志》，第 2 卷，第 1 期，第 015027 页，2018 年 1 月。

[在线] . 可访问：https://dx .doi.org/10 .1088/2399 -6528/aaa234

[24]M. Zhang、H.-M. Yu、H. Yuan、X. Wang、R. Demkowicz- Dobrza´nski 和 J. Liu，“QuanEstimation：用于量子参数估计的开源工具包”，Physical Review Research，第 4 卷，第 043057 页，2022 年 10 月。 [在线]。可通过以下网址获取：https://link.aps.org/doi/10 1103 4 043057 。/PhysRevResearch. . Q. Zhuang 和 J. H. Shapiro，“量子脉冲压缩测距的精度极限”，Physical Review Letters，第 128 卷，第 010501 页，2022 年 1 月。[在线]。可通过以下网址获取：https://link.aps.org/doi/10 1103 128 010501 。/PhysRevLett. .

[26] K. V. Mishra 和 Y. C. Eldar，“认知雷达中的时延估计性能”，载于 2017 年 IEEE 国际声学、语音与信号处理会议（ICASSP），2017 年，第 3141–3145 页。[27] Z. Zhang、Z. Shi、C. Zhou、C. Yan 和 Y. Gu，“压缩时延估计的 Ziv-Zakai 下界”，IEEE Transactions on Signal Processing，第 70 卷，第 4006–4019 页，2022 年。[28] M. Driusso、M. Comisso、F. Babich 和 C. Marshall，“OFDM 信号到达时间估计的性能分析”，IEEE Signal Processing Letters，第 22 卷，第 7 期，第 983–987 页，2015 年。[29] T. Laas 和 W. Xu，“CP-OFDM 系统中到达时间差估计的 Ziv-Zakai 下界”，载于 2021 年 IEEE 无线通信与网络会议（WCNC），2021 年，第 1–5 页。[30] P. Wang 和 Y. J. Morton，“蜂窝网络中小区间干扰对导航应用的影响分析”，IEEE Transactions on Aerospace and Electronic Systems，第 1–10 页，2022 年。[31] W. M. Gifford、D. Dardari 和 M. Z. Win，“多径信息对到达时间估计的影响”，IEEE Transactions on Signal Processing，第 70 卷，第 31–46 页，2022 年。[32] M. F. Keskin、E. Gonendik 和 S. Gezici，“同步可见光定位系统测距的改进下界”，Journal of Lightwave Technology，第 34 卷，第 23 期，第 5496–5504 页，2016 年。[33] P. Closas 和 A. Gusi-Amigó，“GNSS 接收机的直接位置估计：主要成果、架构、改进及挑战分析”，IEEE Signal Processing Magazine，第 34 卷，第 5 期，第 72–84 页，2017 年。[34] A. Gusi-Amigó、P. Closas、A. Mallat 和 L. Vandendorpe，“直接位置估计的 Ziv-Zakai 下界”，Navigation，第 65 卷，第 3 期，第 463–475 页，2018 年。[35] D. Khan 和 K. L. Bell，“利用 Ziv-Zakai 下界分析稀疏线性阵列的 DOA 估计性能”，载于 2010 年 IEEE 雷达会议，2010 年，第 746–751 页。[36] A. Gupta、U. Madhow、A. Arbabian 和 A. Sadri，“利用稀疏子阵列设计毫米波系统的大型有效孔径”，IEEE Transactions on Signal Processing，第 67 卷，第 24 期，第 6483–6497 页，2019 年。[37] D. B. Alexander、R. M. Narayanan 和 B. Himed，“存在互耦时宽带测向的下界”，载于 2019 年 IEEE 雷达会议（RadarConf），2019 年，第 1–6 页。[38] Z. Zhang、Z. Shi 和 Y. Gu，“DOA 估计的 Ziv-Zakai 下界”，IEEE Transactions on Signal Processing，2022 年。[39] V. M. Chiriac 和 A. M. Haimovich，“MIMO 雷达系统中目标定位估计的 Ziv-Zakai 下界 ”，载于 2010 年 IEEE 雷达会议，2010 年，第 678–683 页。[40] V. M. Chiriac、Q. He、A. M. Haimovich 和 R. S. Blum，“MIMO 雷达系统中联合参数估计的 Ziv-Zakai 下界”，IEEE Transactions on Signal Processing，第 63 卷，第 18 期，第 4956–4968 页，2015 年。[41] Y. Wu 和 S. Verdú，“压缩感知中的最优相变”，IEEE Transactions on Information Theory，第 58 卷，第 10 期，第 6241–6263 页，2012 年。[42] S. M. Kay，《统计信号处理基础》，第 2 卷：检测理论。Prentice Hall PTR，1998 年。[43] D. Stotz 和 H. Bölcskei，“向量干扰信道中的自由度”，IEEE Transactions on Information Theory，第 62 卷，第 7 期，第 4172–4197 页，2016 年。[44] Y. Wu 和 S. Verdú，“均方误差（MMSE）维度”，IEEE Transactions on Information Theory，第 57 卷，第 8 期，第 4857–4879 页，2011 年。[45] I. Sason 和 S. Verdú，“Arimoto-Rényi 条件熵与贝叶斯 M元假设检验”，IEEE Transactions on Information Theory，第 64 卷，第 1 期，第 4–25 页，2017 年。[46] G. B. Folland，《实分析：现代技术及其应用》。John Wiley & Sons，1999 年，第 40 卷。

附录 A 定理1的证明

定理1的证明直接源于以下定理 5。特别地，由于定理 5 对任意估计器 $g(\mathsf{Y}) \in \mathcal{X}$ 以及任意 $\mathsf{a} \in \mathcal{X}$均成立，我们可以通过在定理 5 中取 $g(\mathsf{Y}) = \mathbb{E}[\mathsf{X}|\mathsf{Y}]$ 以及 $\mathsf{a} = \mathsf{e}_i$，并对 $i \in [1 : \mathsf{d}(\mathcal{X})]$ 求和，得到 $\mathsf{mmse}(\mathsf{X}|\mathsf{Y})$ 的下界。这将引出定理1中的均方误差（MMSE）下界。

定理 5：设 $\mathcal{X}$ 是一个向量空间。若使用估计器 $g(\mathsf{Y}) \in \mathcal{X}$ 来估计 $\mathsf{X} \in \mathcal{X}$，则 $\boldsymbol{\epsilon} = g(\mathsf{Y}) - \mathsf{X}$ 为估计误差。于是，对于任意 $\mathsf{a} \in \mathcal{X}$ 和任意整数 $\mathsf{M} \geq 2$，有

$$\mathbb{E}\Big[\langle \mathbf{a}, \boldsymbol{\epsilon}\rangle^2\Big] \geq \int_0^\infty \frac{t}{2} \mathcal{V}_t\left\{\frac{\mathsf{h}(t, \mathbf{a}, \mathsf{M})}{\mathsf{M}-1}\right\} \mathrm{d}t, \tag{36}$$

其中，对于 $t > 0$ 和 $i \in [0 : \mathsf{M}-1]$，有

$$\mathsf{h}(t, \mathbf{a}, \mathsf{M}) = \max_{\substack{\mathcal{U} \subset \mathcal{X}: \\ \mathcal{U} = \{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1} \\ \langle \mathbf{a}, \mathbf{u}_k\rangle = kt, \\ k \in [0:\mathsf{M}-1]}} \int P_e(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}), \mathcal{U}) \, \mathrm{d}\mu_{\mathcal{U}}(\mathbf{x}). \tag{37}$$

为定义 $\mathsf{h}(t, \mathsf{a}, \mathsf{M})$，令：

- 对于集合 $\mathcal{U} = \{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1} \subset \mathcal{X}$，测度 $\mu_{\mathcal{U}}$ 定义为

$$\mu_{\mathcal{U}} = \sum_{k=0}^{\mathsf{M}-1} P_{\mathbf{X}-\mathbf{u}_k}; \tag{38}$$

- 对于 $\mathbf{x} \in \mathcal{X}$ 和 $\mathcal{U} \subset \mathcal{X}$，令

$$\mathcal{P}_{\mathcal{U}}(\mathbf{x}) = \left\{p_i : p_i = \frac{\mathrm{d}P_{\mathbf{X}-\mathbf{u}_i}}{\mathrm{d}\mu_{\mathcal{U}}}(\mathbf{x}), \mathbf{u}_i \in \mathcal{U} = \{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1}\right\}; \tag{39}$$

注意，根据构造，Radon–Nikodym 导数总是定义良好的。

证明：我们先定义几个术语。首先，注意到

$$\mathbb{E}[\langle \mathbf{a}, \boldsymbol{\epsilon}\rangle^2] = \int_0^\infty \frac{t}{2} \Pr\left(|\langle \mathbf{a}, \boldsymbol{\epsilon}\rangle| \geq \frac{t}{2}\right) \mathrm{d}t. \tag{40}$$

其次，固定某个 $\mathsf{M} \geq 2$，选取一组向量 $\mathcal{U} = \{\mathbf{u}_i\}_{i=0}^{\mathsf{M}-1} \subset \mathcal{X}$，并按（39）定义 $\mathcal{P}_{\mathcal{U}}(\mathbf{x})$。接下来，令$\langle \mathbf{a}, \mathbf{u}_i\rangle = it$，并注意到

$$\begin{aligned}
&\Pr\left(|\langle \mathbf{a}, \boldsymbol{\epsilon}\rangle| \geq \frac{t}{2}\right) \\
&= \int \mathrm{d}P_{\mathbf{X}}(\mathbf{x}) \Pr\left(\langle \mathbf{a}, g(\mathbf{Y}) - \mathbf{X}\rangle \geq \frac{t}{2} \,\middle|\, \mathbf{X} = \mathbf{x}\right) \\
&\quad + \int \mathrm{d}P_{\mathbf{X}}(\mathbf{x}) \Pr\left(\langle \mathbf{a}, g(\mathbf{Y}) - \mathbf{X}\rangle \leq -\frac{t}{2} \,\middle|\, \mathbf{X} = \mathbf{x}\right) \\
&= \frac{1}{\mathsf{M}-1} \sum_{i=0}^{\mathsf{M}-1} \left\{\int \mathrm{d}P_{\mathbf{X}}(\mathbf{x}) \Pr\left(\langle \mathbf{a}, g(\mathbf{Y}) - \mathbf{X}\rangle \geq \frac{t}{2} \,\middle|\, \mathbf{X} = \mathbf{x}\right)\right. \\
&\quad \left. + \int \mathrm{d}P_{\mathbf{X}}(\mathbf{x}) \Pr\left(\langle \mathbf{a}, g(\mathbf{Y}) - \mathbf{X}\rangle \leq -\frac{t}{2} \,\middle|\, \mathbf{X} = \mathbf{x}\right)\right\},
\end{aligned} \tag{41}$$

其中，第一个等式是通过应用全概率定律并代入 $\boldsymbol{\epsilon} = g(\mathsf{Y}) - \mathsf{X}$ 得到的。(41) 中的求和式由下一页顶部的 (42) 给出，

其中，带标签的等式分别基于：(a) 在第一个积分中作变量替换 $\mathbf{x}' = \mathbf{x} + \mathbf{u}_{i-1}$，在第二个积分中作变量替换 $\mathbf{x}' = \mathbf{x} + \mathbf{u}_i$，以及 $\mathrm{d}P_{\mathbf{X}}(\mathbf{x} + \mathbf{u}_{i-1}) = \mathrm{d}P_{\mathbf{X}-\mathbf{u}_{i-1}}(\mathbf{x})$；(b) 利用假设 $\langle \mathbf{a}, \mathbf{u}_i\rangle = it$；(c) 利用式 (39) 中的 $p_i \in \mathcal{P}_{\mathcal{U}}(\mathbf{x})$ 以及式 (38) 中的 $\mu_{\mathcal{U}}$；(d) 将 $P_e^{\phi}(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}) \, \mathcal{U})$ 定义为与（可能是次优的）决策规则相关的 M 元假设检验问题的错误概率，如定义1所示，

$$\phi = \mathcal{H}_i, \text{ where } i = \operatorname*{argmin}_{j \in [0:\mathsf{M}-1]} |\langle \mathbf{a}, g(\mathbf{Y}) - \mathbf{x}\rangle - jt|. \tag{43}$$

结合（41）和（42），得到

$$\begin{aligned}
\Pr\left(|\langle \mathbf{a}, \boldsymbol{\epsilon}\rangle| \geq \frac{t}{2}\right) &= \frac{\int P_e^{\phi}(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}), \mathcal{U}) \, \mathrm{d}\mu_{\mathcal{U}}(\mathbf{x})}{\mathsf{M}-1} \\
&\geq \frac{\int P_e(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}), \mathcal{U}) \, \mathrm{d}\mu_{\mathcal{U}}(\mathbf{x})}{\mathsf{M}-1},
\end{aligned} \tag{44}$$

其中我们用到了这样一个事实：$P_e(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}) \, \mathcal{U})$ 是与最优决策规则（即最大后验决策规则）相关联的错误概率。（44）中的下界还可以通过优化 $\mathcal{U} = \{\mathbf{u}_i\}_{i=0}^{\mathsf{M}-1}$ 进一步收紧，使得 $\langle \mathbf{a}, \mathbf{u}_i\rangle = it, \ \forall i \in [0 : \mathsf{M}-1]$。于是，

$$\begin{aligned}
\Pr\left(|\langle \mathbf{a}, \boldsymbol{\epsilon}\rangle| \geq \frac{t}{2}\right) &\geq \max_{\substack{\mathcal{U} \subset \mathcal{X}: \\ \mathcal{U} = \{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1} \\ \langle \mathbf{a}, \mathbf{u}_k\rangle = kt, \\ k \in [0:\mathsf{M}-1]}} \frac{\int P_e(\mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}), \mathcal{U}) \, \mathrm{d}\mu_{\mathcal{U}}(\mathbf{x})}{\mathsf{M}-1} \\
&= \frac{\mathsf{h}(t, \mathbf{a}, \mathsf{M})}{\mathsf{M}-1}.
\end{aligned} \tag{45}$$

由于 $\Pr(|\langle \mathsf{a}, \boldsymbol{\epsilon}\rangle| \geq \frac{t}{2})$ 关于 t 单调，因此对（45）右侧应用谷值填充函数是有效的，因为 $\mathcal{V}_t\{\Pr(|\langle \mathsf{a}, \boldsymbol{\epsilon}\rangle| \geq \frac{t}{2})\} = \Pr(|\langle \mathsf{a}, \boldsymbol{\epsilon}\rangle| \geq \frac{t}{2})$。因此，得到

$$\Pr\left(|\langle \mathsf{a}, \boldsymbol{\epsilon}\rangle| \geq \frac{t}{2}\right) \geq \mathcal{V}_t\left\{\frac{\mathsf{h}(t, \mathsf{a}, \mathsf{M})}{\mathsf{M}-1}\right\}. \tag{46}$$

将（46）代入（40），定理 5 的证明即告完成。 ■

Appendix B
Proof of Theorem 2

首先，我们注意到，如果对所有 $t \in \mathbb{R}$，$f(t) \leq g(t)$，那么

$$\mathcal{V}_t\{f(t)\} \leq \mathcal{V}_t\{g(t)\}, \ \forall t \in \mathbb{R}. \tag{47}$$

其次，谷值填充函数是下半连续的，即我们有

$$\liminf_{n \to \infty} \mathcal{V}_t\{f_n(t)\} \geq \mathcal{V}_t\{\liminf_{n \to \infty} f_n(t)\}. \tag{48}$$

为突出式 (6) 中 $\mathsf{h}(t, i, \mathsf{M})$ 对 η 的依赖关系，我们记作 $\mathsf{h}_\eta(t, i, \mathsf{M})$。此外，由于 A2 假设 $P_e(\eta, \mathbf{x}; \mathcal{P}_{\mathcal{U}}(\mathbf{x}) \, \mathcal{U})$ 关于 η 非递减，因此我们注意到 $\mathsf{h}_\eta(t, i, \mathsf{M})$ 关于噪声水平 η 非递减。于是，式 (5) 中的 Ziv-Zakai 下界可以写为

$$\begin{aligned}
\overline{\mathsf{ZZ}}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}) &= \sum_{i=1}^{\mathsf{d}(\mathcal{X})} \int_0^\infty \frac{t}{2} \mathcal{V}_t\left\{\frac{\mathsf{h}_\eta(t, i, \mathsf{M})}{\mathsf{M}-1}\right\} \mathrm{d}t \\
&\leq \sum_{i=1}^{\mathsf{d}(\mathcal{X})} \int_0^\infty \frac{t}{2} \mathcal{V}_t\left\{\frac{\mathsf{h}_\infty(t, i, \mathsf{M})}{\mathsf{M}-1}\right\} \mathrm{d}t.
\end{aligned} \tag{49}$$

$$
\begin{aligned}
&\sum_{i=0}^{\mathsf{M}-1}\left\{\int \mathrm{d}P_{\mathbf{X}}(\mathbf{x})\Pr\left(\langle \mathbf{a}, g(\mathbf{Y})-\mathbf{X}\rangle \geq \frac{t}{2}\;\middle|\;\mathbf{X}=\mathbf{x}\right)+\int \mathrm{d}P_{\mathbf{X}}(\mathbf{x})\Pr\left(\langle \mathbf{a}, g(\mathbf{Y})-\mathbf{X}\rangle \leq -\frac{t}{2}\;\middle|\;\mathbf{X}=\mathbf{x}\right)\right\}\\
&\overset{\text{(a)}}{=}\sum_{i=0}^{\mathsf{M}-1}\left\{\int \mathrm{d}P_{\mathbf{X}-\mathbf{u}_{i-1}}(\mathbf{x})\Pr\left(\langle \mathbf{a}, g(\mathbf{Y})-\mathbf{X}\rangle \geq \frac{t}{2}\;\middle|\;\mathbf{X}=\mathbf{x}+\mathbf{u}_{i-1}\right)+\int \mathrm{d}P_{\mathbf{X}-\mathbf{u}_i}(\mathbf{x})\Pr\left(\langle \mathbf{a}, g(\mathbf{Y})-\mathbf{X}\rangle \leq -\frac{t}{2}\;\middle|\;\mathbf{X}=\mathbf{x}+\mathbf{u}_i\right)\right\}\\
&\overset{\text{(b)}}{=}\sum_{i=0}^{\mathsf{M}-1}\left\{\int \mathrm{d}P_{\mathbf{X}-\mathbf{u}_{i-1}}(\mathbf{x})\Pr\left(\langle \mathbf{a}, g(\mathbf{Y})-\mathbf{x}\rangle \geq \left(i-\frac{1}{2}\right)t\;\middle|\;\mathbf{X}=\mathbf{x}+\mathbf{u}_{i-1}\right)\right.\\
&\qquad\left.+\int \mathrm{d}P_{\mathbf{X}-\mathbf{u}_i}(\mathbf{x})\Pr\left(\langle \mathbf{a}, g(\mathbf{Y})-\mathbf{x}\rangle \leq \left(i-\frac{1}{2}\right)t\;\middle|\;\mathbf{X}=\mathbf{x}+\mathbf{u}_i\right)\right\}\\
&\overset{\text{(c)}}{=}\sum_{i=0}^{\mathsf{M}-1}\left\{\int p_{i-1}\Pr\left(\langle \mathbf{a}, g(\mathbf{Y})-\mathbf{x}\rangle \geq \left(i-\frac{1}{2}\right)t\;\middle|\;\mathbf{X}=\mathbf{x}+\mathbf{u}_{i-1}\right)\mathrm{d}\mu_{\mathcal{U}}(\mathbf{x})\right.\\
&\qquad\left.+\int p_i\Pr\left(\langle \mathbf{a}, g(\mathbf{Y})-\mathbf{x}\rangle \leq \left(i-\frac{1}{2}\right)t\;\middle|\;\mathbf{X}=\mathbf{x}+\mathbf{u}_i\right)\mathrm{d}\mu_{\mathcal{U}}(\mathbf{x})\right\}\\
&\overset{\text{(d)}}{=}\int P_e^{\phi}(\mathbf{x},\, t;\, \mathcal{P}_{\mathcal{U}}(\mathbf{x}),\, \mathcal{U})\,\mathrm{d}\mu_{\mathcal{U}}(\mathbf{x}),
\end{aligned}
\tag{42}
$$

为节省篇幅，我们将约束 $\{\mathcal{U}=\{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1}, (\mathbf{u}_k)_i = kt,\ \forall k\}$ 简记为 $\mathcal{C}$。根据假设 A3，我们可以将 $\mathsf{h}_\infty(t,i,\mathsf{M})$ 写为

$$
\begin{aligned}
\mathsf{h}_\infty(t,i,\mathsf{M}) &= \max_{\mathcal{U}\subset\mathcal{X}:\mathcal{C}}\int P_e(\infty,\mathbf{x};\mathcal{P}_{\mathcal{U}}(\mathbf{x}),\mathcal{U})\,\mathrm{d}\mu_{\mathcal{U}}(\mathbf{x})\\
&= \max_{\mathcal{U}\subset\mathcal{X}:\mathcal{C}}\left\{\mathsf{M}-\int \max_{j\in[0:\mathsf{M}-1]}\mathrm{d}P_{\mathbf{X}-\mathbf{u}_j}(\mathbf{x})\right\}\\
&= \mathsf{M}-\min_{\mathcal{U}\subset\mathcal{X}:\mathcal{C}}\int \max_{j\in[0:\mathsf{M}-1]}\mathrm{d}P_{\mathbf{X}-\mathbf{u}_j}(\mathbf{x}),
\end{aligned}
\tag{50}
$$

将其代入（49）后，得到

$$
\begin{aligned}
&\overline{\mathsf{ZZ}}(P_{\mathbf{X}},P_{\mathbf{Y}|\mathbf{X}},\mathsf{M})\\
&\leq \sum_{i=1}^{\mathsf{d}(\mathcal{X})}\int_0^\infty \frac{t}{2}\mathcal{V}_t\left\{\frac{\mathsf{M}-\mathsf{H}(P_{\mathbf{X}},t,i,\mathsf{M})}{\mathsf{M}-1}\right\}\mathrm{d}t.
\end{aligned}
\tag{51}
$$

接下来，我们有

$$
\begin{aligned}
&\liminf_{\eta\to\infty}\overline{\mathsf{ZZ}}(P_{\mathbf{X}},P_{\mathbf{Y}|\mathbf{X}},\mathsf{M})\\
&\overset{\text{(a)}}{\geq}\sum_{i=1}^{\mathsf{d}(\mathcal{X})}\int_0^\infty \liminf_{\eta\to\infty}\frac{t}{2}\mathcal{V}_t\left\{\frac{\mathsf{h}_\eta(t,i,\mathsf{M})}{\mathsf{M}-1}\right\}\mathrm{d}t\\
&\overset{\text{(b)}}{\geq}\sum_{i=1}^{\mathsf{d}(\mathcal{X})}\int_0^\infty \frac{t}{2}\mathcal{V}_t\left\{\frac{\liminf_{\eta\to\infty}\mathsf{h}_\eta(t,i,\mathsf{M})}{\mathsf{M}-1}\right\}\mathrm{d}t,
\end{aligned}
\tag{52}
$$

其中，(a) 根据法图引理得出，(b) 则由 (48) 得出。此外，我们有

$$
\begin{aligned}
\liminf_{\eta\to\infty}\mathsf{h}_\eta(t,i,\mathsf{M}) &= \liminf_{\eta\to\infty}\max_{\mathcal{U}\subset\mathcal{X}:\mathcal{C}}\int P_e(\eta,\mathbf{x};\mathcal{P}_{\mathcal{U}}(\mathbf{x}),\mathcal{U})\,\mathrm{d}\mu_{\mathcal{U}}(\mathbf{x})\\
&\overset{\text{(a)}}{\geq}\max_{\mathcal{U}\subset\mathcal{X}:\mathcal{C}}\liminf_{\eta\to\infty}\int P_e(\eta,\mathbf{x};\mathcal{P}_{\mathcal{U}}(\mathbf{x}),\mathcal{U})\,\mathrm{d}\mu_{\mathcal{U}}(\mathbf{x})\\
&\overset{\text{(b)}}{\geq}\max_{\mathcal{U}\subset\mathcal{X}:\mathcal{C}}\int P_e(\infty,\mathbf{x};\mathcal{P}_{\mathcal{U}}(\mathbf{x}),\mathcal{U})\,\mathrm{d}\mu_{\mathcal{U}}(\mathbf{x})\\
&\overset{\text{(c)}}{=}\mathsf{M}-\min_{\mathcal{U}\subset\mathcal{X}:\mathcal{C}}\int \max_{j\in[0:\mathsf{M}-1]}\mathrm{d}P_{\mathbf{X}-\mathbf{u}_j}(\mathbf{x}),
\end{aligned}
\tag{53}
$$

其中，(a) 根据最大最小不等式得出，(b) 根据法图引理得出，(c) 则由假设 A3 得出。将 (53) 代入 (52)，可得，对于 $\eta\to\infty$

$$
\overline{\mathsf{ZZ}}(P_{\mathbf{X}},P_{\mathbf{Y}|\mathbf{X}},\mathsf{M})\geq\sum_{i=1}^{\mathsf{d}(\mathcal{X})}\int_0^\infty\frac{t}{2}\mathcal{V}_t\left\{\frac{\mathsf{M}-\mathsf{H}(P_{\mathbf{X}},t,i,\mathsf{M})}{\mathsf{M}-1}\right\}\mathrm{d}t,
$$

这与（51）一致，并完成了定理 2 的证明。

Appendix C
Proof of Theorem 3

利用（9）中的变量替换 $t=\sqrt{\eta}\tau$，我们得到

$$
\frac{\mathsf{ZZ}(P_{\mathbf{X}},P_{\mathbf{Y}|\mathbf{X}},2)}{\eta}=\sum_{i=1}^{\mathsf{d}(\mathcal{X})}\int_0^\infty\frac{\tau}{2}\mathsf{h}(\sqrt{\eta}\tau,i,2)\,\mathrm{d}\tau.
\tag{54}
$$

此外，由于我们可以写成

$$
\begin{aligned}
\mu_{\mathcal{U}} &= \alpha\sum_{k=0}^{1}P_{\mathbf{X}_C-\mathbf{u}_k}+(1-\alpha)\sum_{k=0}^{1}P_{\mathbf{X}_D-\mathbf{u}_k}\\
&= \alpha\mu_{\mathcal{U}_C}+(1-\alpha)\mu_{\mathcal{U}_D},
\end{aligned}
\tag{55}
$$

我们可以将（54）中的 $\mathsf{h}(\sqrt{\eta}\tau,i,2)$ 下界为

$$
\begin{aligned}
\mathsf{h}(\sqrt{\eta}\tau,i,2) &\overset{\text{(a)}}{\geq}\max_{\substack{\mathcal{U}\subset\mathcal{X}:\\ \mathcal{U}=\{\mathbf{u}_k\}_{k=0}^{1}\\ (\mathbf{u}_k)_i=\sqrt{\eta}k\tau,\ \forall k}}\alpha\int P_e(\eta,\mathbf{x};\mathcal{P}_{\mathcal{U}}(\mathbf{x}),\mathcal{U})\,\mathrm{d}\mu_{\mathcal{U}_C}(\mathbf{x})\\
&\overset{\text{(b)}}{=}\max_{\substack{\mathcal{U}\subset\mathcal{X}:\\ \mathcal{U}=\{\mathbf{u}_k\}_{k=0}^{1}\\ (\mathbf{u}_k)_i=\sqrt{\eta}k\tau,\ \forall k}}\alpha\int P_e(\eta,\mathbf{x};\mathcal{P}_{\mathcal{U}_C}(\mathbf{x}),\mathcal{U})\,\mathrm{d}\mu_{\mathcal{U}_C}(\mathbf{x})\\
&\overset{\text{(c)}}{=}\alpha\mathsf{h}_C(\sqrt{\eta}\tau,i,2),
\end{aligned}
\tag{56}
$$

其中，(a) 是通过舍弃关于$\mu_{\mathcal{U}_D}$的积分得到的；(b) 是因为，对于$\mu_{\mathcal{U}_C}$而言，几乎必然，

$$\mathcal{P}_{\mathcal{U}}(\mathbf{x}) = \mathcal{P}_{\mathcal{U}_C}(\mathbf{x}), \tag{57}$$

其中

$$\mathcal{P}_{\mathcal{U}_C}(\mathbf{x}) = \left\{ p_i : p_i = \frac{\mathrm{d}P_{\mathbf{X}_C - \mathbf{u}_i}}{\mathrm{d}\mu_{\mathcal{U}_C}}(\mathbf{x}), \mathbf{u}_i \in \mathcal{U} = \{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1} \right\};$$

为说明这一点，注意到 $P_{\mathbf{X}_D}$ 和 $\mu_{\mathcal{U}_D}$ 相对于 $\mu_{\mathcal{U}_C}$ 是奇异的，因此，几乎处处相对于 $\mu_{\mathcal{U}_C}$，我们有

$$\frac{\mathrm{d}P_{\mathbf{X}-\mathbf{u}_i}}{\mathrm{d}\mu_{\mathcal{U}}}(\mathbf{x}) = \frac{\mathrm{d}P_{\mathbf{X}_C-\mathbf{u}_i}}{\mathrm{d}\mu_{\mathcal{U}_C}}(\mathbf{x}); \tag{58}$$

在 (c) 中，我们通过使用 $\mathsf{h}_C(\sqrt{\eta}\tau, i, 2)$，突出显示 $\mathsf{h}(\sqrt{\eta}\tau, i, 2)$ 对 $\mathbf{X}_C$ 的分布的依赖性（而对于其对 X 的分布的依赖性，我们使用 $\mathsf{h}(\sqrt{\eta}\tau, i, 2)$）。

现在，我们可以利用附录 F-B 中的引理 2，进一步对 (56) 中的 $\mathsf{h}_C(\sqrt{\eta}\tau, i, 2)$ 给出下界，得到

$$\begin{aligned}\mathsf{h}_C(\sqrt{\eta}\tau, i, 2) \geq \int \Bigg(&\int_{\mathbf{y}\in\mathcal{R}_0(\eta)} f_{\mathbf{Z}}(\mathbf{y}-\tau\mathbf{e}_i) f_{\mathbf{X}_C}(\mathbf{x}+\sqrt{\eta}\tau\mathbf{e}_i)\mathrm{d}\mathbf{y} \\ &+ \int_{\mathbf{y}\in\mathcal{R}_1(\eta)} f_{\mathbf{Z}}(\mathbf{y}) f_{\mathbf{X}_C}(\mathbf{x})\mathrm{d}\mathbf{y} \Bigg)\mathrm{d}\mathbf{x}, \end{aligned} \tag{59}$$

其中

$$\mathcal{R}_0(\eta) = \left\{ \mathbf{y} : \ln \frac{f_{\mathbf{X}_C}(\mathbf{x})}{f_{\mathbf{X}_C}(\mathbf{x}+\sqrt{\eta}\tau\mathbf{e}_i)} + \frac{\tau^2}{2} > \tau y_i \right\}, \text{ and}$$

$$\mathcal{R}_1(\eta) = \left\{ \mathbf{y} : \ln \frac{f_{\mathbf{X}_C}(\mathbf{x})}{f_{\mathbf{X}_C}(\mathbf{x}+\sqrt{\eta}\tau\mathbf{e}_i)} + \frac{\tau^2}{2} < \tau y_i \right\}.$$

具体而言，为了得到 (59)，我们在附录 F-B 的引理 2 中令 $\mathsf{v}_0 = \mathbf{0}$ 和 $\mathsf{v}_1 = \tau\mathsf{e}_i$，并用 $f_{\mathbf{Z}}$ 表示 $\mathsf{Z} \sim \mathcal{N}(0, I_{\mathsf{d}(\mathcal{X})})$ 的概率密度函数。

然后，利用法图引理和（59），我们得到

$$\begin{aligned}\liminf_{\eta\to 0} \mathsf{h}_C(\sqrt{\eta}\tau, i, 2) &\geq \int \Bigg(\int_{\mathbf{y}\in\mathcal{R}_0(0)} f_{\mathbf{Z}}(\mathbf{y}-\tau\mathbf{e}_i) f_{\mathbf{X}_C}(\mathbf{x})\mathrm{d}\mathbf{y} \\ &\quad + \int_{\mathbf{y}\in\mathcal{R}_1(0)} f_{\mathbf{Z}}(\mathbf{y}) f_{\mathbf{X}_C}(\mathbf{x})\mathrm{d}\mathbf{y} \Bigg)\mathrm{d}\mathbf{x} \\ &= \Pr\Big(|Z_i| > \frac{\tau}{2}\Big). \end{aligned} \tag{60}$$

此外，始终使用法图引理，我们得到

$$\liminf_{\eta\to 0} \frac{\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, 2)}{\eta} \geq \sum_{i=1}^{\mathsf{d}(\mathcal{X})} \int_0^\infty \liminf_{\eta\to 0} \frac{\tau}{2} \mathsf{h}(\sqrt{\eta}\tau, i, 2)\, \mathrm{d}\tau.$$

将 (56) 和 (60) 代入上述表达式，得到

$$\begin{aligned}\liminf_{\eta\to 0} \frac{\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, 2)}{\eta} &\geq \alpha \sum_{i=1}^{\mathsf{d}(\mathcal{X})} \int_0^\infty \frac{\tau}{2} \Pr\Big(|Z_i| > \frac{\tau}{2}\Big)\, \mathrm{d}\tau \\ &= \alpha \sum_{i=1}^{\mathsf{d}(\mathcal{X})} \mathbb{E}[Z_i^2] = \alpha \mathsf{d}(\mathcal{X}). \end{aligned} \tag{61}$$

为完成定理 3 的证明，我们注意到 [43], [44]

$$\lim_{\eta\to 0} \frac{\mathrm{mmse}(\mathbf{X}|\mathbf{Y})}{\eta} = \alpha \mathsf{d}(\mathcal{X}). \tag{62}$$

由于 $\mathsf{mmse}(\mathsf{X}|\mathsf{Y}) \geq \mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, \mathsf{M}) \geq \mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, 2)$，因此由 (61) 和 (62) 可得

$$\liminf_{\eta\to 0} \frac{\mathsf{ZZ}(P_{\mathbf{X}}, P_{\mathbf{Y}|\mathbf{X}}, 2)}{\eta} = \alpha \mathsf{d}(\mathcal{X}), \tag{63}$$

这就完成了定理 3 的证明。

Appendix D
Proof of Proposition 1

为了证明张量化的性质，只需证明 (6) 中的 $\mathsf{h}(t, i, \mathsf{M})$ 仅取决于 P_{X_i} 和 $P_{Y_i|X_i}$，而非 $P_{\mathbf{X}}$ 和 $P_{\mathbf{Y}|\mathbf{X}}$。我们首先注意到，

$$\mathsf{h}(t, i, \mathsf{M}) = \mathsf{M} - \min_{\substack{\mathcal{U}\subset\mathcal{X}: \\ \mathcal{U}=\{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1} \\ (\mathbf{u}_k)_i = kt,\ \forall k}} \int \max_{\ell\in[0:\mathsf{M}-1]} \mathrm{d}P_{\mathbf{X},\mathbf{Y}}(\mathbf{x}+\mathbf{u}_\ell, \mathbf{y}), \tag{64}$$

其中，我们使用了附录 F-C 中的引理 3，以及 $\Pr(\mathcal{H}_\ell) = \mathrm{d}_{P_{\mathbf{X}-\mathbf{u}_l}(\mathbf{x})\, \mathrm{d}\mu_{\mathcal{U}}(\mathbf{x})}$ 这一事实~~，利用 $P_{\mathbf{X}} = \prod_{i=1}^d P_{X_i}$~~ 和 $P_{\mathbf{Y}|\mathbf{X}} = \prod_{i=1}^d P_{Y_i|X_i}$ 以及 $d = \mathsf{d}(\mathcal{X})$ 这些事实，我们可以将上面的积分写为

$$\begin{aligned}&\int \max_{\ell\in[0:\mathsf{M}-1]} \mathrm{d}P_{\mathbf{X},\mathbf{Y}}(\mathbf{x}+\mathbf{u}_\ell, \mathbf{y}) \\ &= \int \max_{\ell\in[0:\mathsf{M}-1]} \left\{ \prod_{j=1}^d \mathrm{d}P_{X_j,Y_j}(x_j + (\mathbf{u}_\ell)_j, y_j) \right\} \\ &\geq \int \max_{\ell\in[0:\mathsf{M}-1]} \left\{ \prod_{j\neq s} \mathrm{d}P_{X_j,Y_j}(x_j + (\mathbf{u}_\ell)_j, y_j) \right\}, \end{aligned} \tag{65}$$

其中，不等式是通过将最大值与积分互换，并利用样本空间上的概率测度等于一这一事实得到的。对除 $j = i$ 之外的所有 j 逐一重复这一操作，可得

$$\begin{aligned}&\int \max_{\ell\in[0:\mathsf{M}-1]} \mathrm{d}P_{\mathbf{X},\mathbf{Y}}(\mathbf{x}+\mathbf{u}_\ell, \mathbf{y}) \\ &\geq \int \max_{\ell\in[0:\mathsf{M}-1]} \mathrm{d}P_{X_i,Y_i}(x_i + \ell t, y_i). \end{aligned} \tag{66}$$

由于通过令 $\mathcal{U} = \{\mathbf{u}_k\}_{k=0}^{\mathsf{M}-1}$ 满足 $\mathbf{u}_k = kt\mathsf{e}_i$，$k \in [0 : \mathsf{M}-1]$，即可达到 (64) 中的下界，因此 (64) 中最小化问题的一个解为$\mathcal{U} = \{\mathbf{u}_k : \mathbf{u}_k = kt\mathsf{e}_i,\ k \in [0 : \mathsf{M}-1]\}$。因此，我们有

$$\mathsf{h}(t, i, \mathsf{M}) = \mathsf{M} - \int \max_{\ell\in[0:\mathsf{M}-1]} \mathrm{d}P_{X_i,Y_i}(x_i + \ell t, y_i), \tag{67}$$

注意到 (67) 仅取决于 P_{X_i} 和 $P_{Y_i|X_i}$，由此即完成命题 1 的证明。

Appendix E
Proof of Corollary 1

命题 2 给出了如下充分必要条件：对于所有的 x、y 和 $t > 0$，都存在 $k_1 = k_2$ 使得

$$\hat{k}_1 \in \operatorname*{argmax}_{k\in[0:\mathsf{M}-1]} f_{X|\mathbf{Y}}(x + kt|\mathbf{y}), \text{ and} \tag{68a}$$

$$\hat{k}_2 \in \operatorname*{argmin}_{k\in[0:\mathsf{M}-1]} |\mathbb{E}[X|\mathbf{Y} = \mathbf{y}] - x - kt|. \tag{68b}$$

首先，假设 $f_{X|\mathbf{Y}}$ 至少有两个众数，分别记为 m_1 和 m_2。那么，至少存在一个众数（记为 m_1），不同于 $\mathbb{E}[X|\mathbf{Y}=\mathbf{y}]$。现在，令 $x=m_1-\varepsilon$，并令 $t=\frac{2\varepsilon}{M-1}$，其中 ε 满足 $f_{X|\mathbf{Y}}$ 在 $[m_1-\varepsilon, m_1+\varepsilon]$ 上是凹函数，且 $0<\varepsilon<|\mathbb{E}[X|\mathbf{Y}=\mathbf{y}]-m_1|$。对 x 和 t 作此选择意味着 $k_1\notin\{0,\mathsf{M}-1\}$，但 $k_2\in\{0,\mathsf{M}-1\}$。因此，k_1 和 k_2 总是不同，这意味着多峰假设不满足命题 2 中的条件。因此，概率密度函数 $f_{X|\mathbf{Y}}$ 必须是单峰的。此外，运用与上述类似的论证，不难证明，对于众数为 m 的单峰概率密度函数 $f_{X|\mathbf{Y}}$，要满足命题 2 中的条件，必须有 $m=\mathbb{E}[X|\mathbf{Y}=\mathbf{y}]$。

现在，设 $f_{X|\mathbf{Y}}$ 为单峰分布，且关于其众数 m 不对称。对于这类概率密度函数，我们可以找到 $\varepsilon_0>0$ 和 $\varepsilon_1>0$，使得 $f_{X|\mathbf{Y}}(m-\varepsilon_0|\mathbf{y})=f_{X|\mathbf{Y}}(m+\varepsilon_1|\mathbf{y})$。考虑以下对 x 和 $t>0$ 的选择，

$$x=m-\varepsilon_0+(-1)^{\ell}\delta \quad\text{and}\quad t=\varepsilon_0+\varepsilon_1, \tag{69}$$

其中 $\ell=\operatorname{argmax}^{i\in\{0,1\}}\varepsilon_i$，且 $\delta>0$ 是一个足够小的数。通过此选择，我们得到 $k_1=\ell$ 和 $k_2=1-\ell$，因此 $k_1\neq k_2$，这意味着非对称假设不满足命题 2 中的条件。因此，概率密度函数 $f_{X|\mathbf{Y}}$ 必须是对称的。

总之，上述分析表明，单峰性和对称性是必要条件。通过指出 [16] 已证明这两个条件也是充分条件，推论 1 的证明得以完成。

APPENDIX F
ANCILLARY LEMMAS

A. 引理 1

引理 1：设 X 是一个具有单峰概率密度函数 f_X 的随机变量。设众数位于 $x=0$，并令 $a(t)$ 满足

$$\lim_{x\to a(t)^-}\max\{f_X(x),f_{X-t}(x)\}=f_{X-t}(a(t)), \tag{70a}$$

$$\lim_{x\to a(t)^+}\max\{f_X(x),f_{X-t}(x)\}=f_X(a(t)). \tag{70b}$$

然后，

$$\int_{-\infty}^{\infty}\max_{k\in[0:\mathsf{M}-1]}f_{X-kt}(x)\,\mathrm{d}x = 1+(\mathsf{M}-1)(F_X(a(t)+t)-F_X(a(t))), \tag{71}$$

其中，F_X 是 X 的累积分布函数。

证明：由于 $f_X(x)$ 是单峰函数，众数位于 $x=0$，因此函数 $f_{X-kt}(x)$ 在 $k\in[0:\mathsf{M}-1]$ 处取得最大值。此外，$f_X(x)$ 的单峰性意味着 $f_{X-kt}(x)$ 关于 $x<-kt$ 单调不减，关于 $x>-kt$ 单调不增。令 $a_i(t)$ 对于 $i\in[0:\mathsf{M}-2]$ 满足以下条件，

$$\lim_{x\to a_i(t)^-}\max\{f_{X-it}(x),f_{X-(i+1)t}(x)\}=f_{X-(i+1)t}(a_i(t)),$$
$$\text{and}\quad\lim_{x\to a_i(t)^+}\max\{f_{X-it}(x),f_{X-(i+1)t}(x)\}=f_{X-it}(a_i(t)).$$

换句话说，$a_i(t)$ 是转折点，也就是说

$$\max\{f_{X-it}(x),f_{X-(i+1)t}(x)\}=\begin{cases}f_{X-(i+1)t}(x) & \text{if } x<a_i(t)\\ f_{X-it}(x) & \text{if } x>a_i(t)\end{cases}.$$

注意，尽管 $a_i(t)$ 可能并不唯一，但由于单峰性假设，它始终存在。因此，我们有：

$$\max_{k\in[0:\mathsf{M}-1]}f_{X-kt}(x) =\begin{cases}f_{X-(\mathsf{M}-1)t}(x), & \text{if } x\in(-\infty,a_{\mathsf{M}-2}(t))\\ f_{X-it}(x), & \text{if } x\in(a_i(t),a_{i-1}(t)), i\in[1:\mathsf{M}-2]\\ f_X(x), & \text{if } x\in(a_0(t),\infty)\end{cases}.$$

于是，我们有 久久精品

$$\begin{aligned}\int_{-\infty}^{\infty}\max_{k\in[0:\mathsf{M}-1]}f_X(x+kt)\,\mathrm{d}x &= \int_{-\infty}^{a_{\mathsf{M}-2}(t)}f_{X-(\mathsf{M}-1)t}(x)\,\mathrm{d}x\\ &\quad+\sum_{i=1}^{\mathsf{M}-2}\int_{a_i(t)}^{a_{i-1}(t)}f_{X-it}(x)\,\mathrm{d}x\\ &\quad+\int_{a_0(t)}^{\infty}f_X(x)\,\mathrm{d}x.\end{aligned}$$

由于 $X-(i+1)t$ 的概率密度函数相对于 $X-it$ 的概率密度函数沿负方向平移了 t（即 $f_{X-(i+1)t}(x)=f_{X-it}(x+t)$），因此我们有 $a_{i+1}(t)+t=a_i(t)$。然后，我们得到

$$\begin{aligned}&\int_{-\infty}^{\infty}\max_{k\in[0:\mathsf{M}-1]}f_{X-kt}(x)\mathrm{d}x\\ &=1+(\mathsf{M}-1)(F_X(a_0(t)+t)-F_X(a_0(t))),\end{aligned} \tag{72}$$

其中，$F_X(x)$ 是 X 的累积分布函数。(72) 中的积分与 $i\in[0:\mathsf{M}-2]$ 无关，因此我们可以写为 $a_0(t)=a(t)$，由此完成引理 1 的证明。■

B. 引理 2

引理 2：设 $\mathbf{Y}=\mathbf{X}+\mathbf{N}$，其中 $\mathbf{N}\sim\mathcal{N}(\mathbf{0}_n,\eta I_n)$，且 $\mathbf{X}$ 为连续型随机变量。设 $f_{\mathbf{Z}}$ 为 $\mathbf{Z}\sim\mathcal{N}(\mathbf{0}_n,I_n)$ 的概率密度函数。那么，

$$\begin{aligned}\mathsf{h}(t,i,2)=\max_{\substack{\mathcal{U}\subset\mathcal{X}:\\ \mathcal{U}=\{\mathbf{v}_k\}_{k=0}^{1}\\ (\mathbf{v}_k)_i=\frac{kt}{\sqrt{\eta}},\ \forall k}}\int\Bigg(&\int_{\mathbf{y}\in\mathcal{R}_0}f_{\mathbf{Z}}(\mathbf{y}-\mathbf{v}_1)f_{\mathbf{X}}(\mathbf{x}+\sqrt{\eta}\mathbf{v}_1)\mathrm{d}\mathbf{y}\\ &+\int_{\mathbf{y}\in\mathcal{R}_1}f_{\mathbf{Z}}(\mathbf{y}-\mathbf{v}_0)f_{\mathbf{X}}(\mathbf{x}+\sqrt{\eta}\mathbf{v}_0)\mathrm{d}\mathbf{y}\Bigg)\mathrm{d}\mathbf{x},\end{aligned} \tag{73}$$

其中

$$\mathcal{R}_0=\left\{\mathbf{y}:\ln\frac{f_{\mathbf{X}}(\mathbf{x}+\sqrt{\eta}\mathbf{v}_0)}{f_{\mathbf{X}}(\mathbf{x}+\sqrt{\eta}\mathbf{v}_1)}+\frac{\|\mathbf{v}_1\|_2^2-\|\mathbf{v}_0\|_2^2}{2}>(\mathbf{v}_1-\mathbf{v}_0)^{\mathsf{T}}\mathbf{y}\right\},$$
$$\mathcal{R}_1=\left\{\mathbf{y}:\ln\frac{f_{\mathbf{X}}(\mathbf{x}+\sqrt{\eta}\mathbf{v}_0)}{f_{\mathbf{X}}(\mathbf{x}+\sqrt{\eta}\mathbf{v}_1)}+\frac{\|\mathbf{v}_1\|_2^2-\|\mathbf{v}_0\|_2^2}{2}<(\mathbf{v}_1-\mathbf{v}_0)^{\mathsf{T}}\mathbf{y}\right\}.$$

证明：二元假设检验问题的错误概率为

$$\begin{aligned}P_e(\eta,\mathbf{x};\mathcal{P}_{\mathcal{U}}(\mathbf{x}),\mathcal{U})&=\mathbb{E}\left[\min_{i\in\{0,1\}}\Pr(\mathcal{H}_i|\mathbf{Y})\right]\\ &\overset{(a)}{=}\mathbb{E}\left[\min_{i\in\{0,1\}}\frac{f_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{x}+\mathbf{u}_i)}{f_{\mathbf{Y}}(\mathbf{Y})}\frac{q_i}{q_0+q_1}\right]\\ &=\int\min_{i\in\{0,1\}}f_{\mathbf{N}}(\mathbf{y}-\mathbf{x}-\mathbf{u}_i)\frac{q_i}{q_0+q_1}\mathrm{d}\mathbf{y}\\ &\overset{(b)}{=}\int\min_{i\in\{0,1\}}f_{\mathbf{N}}(\mathbf{z}-\mathbf{u}_i)\frac{q_i}{q_0+q_1}\mathrm{d}\mathbf{z},\end{aligned} \tag{74}$$

其中，(a)由贝叶斯规则以及令 $f_{\mathbf{X}}(\mathbf{x}+\mathbf{u}_i)=q_i,\ i\in\{0,1\}$ 得出，从而得到

$$\Pr(\mathcal{H}_i)=\frac{f_{\mathbf{X}}(\mathbf{x}+\mathbf{u}_i)}{f_{\mathbf{X}}(\mathbf{x}+\mathbf{u}_0)+f_{\mathbf{X}}(\mathbf{x}+\mathbf{u}_1)}=\frac{q_i}{q_0+q_1};\quad(75)$$

（b）则由变量替换 $\mathbf{z}=\mathbf{y}-\mathbf{x}$ 得出。不难看出，（74）可以写成

$$P_e(\eta,\mathbf{x};\mathcal{P}_{\mathcal{U}}(\mathbf{x}),\mathcal{U})=\frac{1}{q_0+q_1}\int_{\mathbf{z}\in\mathcal{R}_0'}f_{\mathbf{Z}}\left(\mathbf{z}-\frac{\mathbf{u}_1}{\sqrt{\eta}}\right)q_1\mathrm{d}\mathbf{z}+\frac{1}{q_0+q_1}\int_{\mathbf{z}\in\mathcal{R}_1'}f_{\mathbf{Z}}\left(\mathbf{z}-\frac{\mathbf{u}_0}{\sqrt{\eta}}\right)q_0\mathrm{d}\mathbf{z},\quad(76)$$

其中 $\mathbf{Z}\sim\mathcal{N}(\mathbf{0}_n,I_n)$ 和

$$\mathcal{R}_0'=\left\{\mathbf{z}\in\mathbb{R}^n:\ln\frac{q_0}{q_1}+\frac{\|\mathbf{u}_1\|_2^2-\|\mathbf{u}_0\|_2^2}{2\eta}>\frac{(\mathbf{u}_1-\mathbf{u}_0)^{\mathsf{T}}\mathbf{z}}{\sqrt{\eta}}\right\},$$

$$\mathcal{R}_1'=\left\{\mathbf{z}\in\mathbb{R}^n:\ln\frac{q_0}{q_1}+\frac{\|\mathbf{u}_1\|_2^2-\|\mathbf{u}_0\|_2^2}{2\eta}<\frac{(\mathbf{u}_1-\mathbf{u}_0)^{\mathsf{T}}\mathbf{z}}{\sqrt{\eta}}\right\}.$$

在（6）中通过变量替换 $\mathbf{u}_i=\sqrt{\eta}\mathbf{v}_i$ 将（76）代入 $\mathsf{h}(t,i,2)$，并注意到 $\mu_{\mathcal{U}}(\mathbf{x})=q_0+q_1$，即可完成引理 2 的证明。∎

C. 引理 3

引理 3：定义1中 M 元假设检验问题的最小错误概率可以表示为

$$P_e(\mathbf{x};\mathcal{P}_{\mathcal{U}}(\mathbf{x}),\mathcal{U})=1-\int\max_{i\in[0:\mathsf{M}-1]}\mathrm{d}P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}+\mathbf{u}_i)\ p_i.\quad(77)$$

证明：当采用最大后验决策规则时，可达到最小错误概率 [42]。因此，对于定义1中的 M 元假设检验问题，最小错误概率为 [45]。

$$P_e(\mathbf{x};\mathcal{P}_{\mathcal{U}}(\mathbf{x}),\mathcal{U})=1-\int\mathrm{d}P_{\mathbf{Y}}(\mathbf{y})\max_{i\in[0:\mathsf{M}-1]}\Pr(\mathcal{H}_i|\mathbf{Y}=\mathbf{y}).$$

根据贝叶斯定理，我们得到

$$\begin{aligned}P_e(\mathbf{x};\mathcal{P}_{\mathcal{U}}(\mathbf{x}),\mathcal{U})&=1-\int\mathrm{d}P_{\mathbf{Y}}(\mathbf{y})\max_{i\in[0:\mathsf{M}-1]}\Pr(\mathcal{H}_i|\mathbf{Y}=\mathbf{y})\\&=1-\int\max_{i\in[0:\mathsf{M}-1]}\mathrm{d}P_{\mathbf{Y}|\mathcal{H}_i}(\mathbf{y}|\mathcal{H}_i)\Pr(\mathcal{H}_i)\\&=1-\int\max_{i\in[0:\mathsf{M}-1]}\mathrm{d}P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}+\mathbf{u}_i)\ p_i,\end{aligned}$$

由此完成引理 3 的证明。∎

D. 引理 4

引理 4：假设 $\mathcal{X}$ 是可数的。则有

$$\mathcal{B}=\left\{t\geq0:\bigcup_{\substack{\mathcal{I}\subseteq[0:\mathsf{M}-1]:\\|\mathcal{I}|\geq2}}\bigcap_{k\in\mathcal{I}}\{\mathcal{X}-\mathbf{u}_k\}\neq\varnothing,\mathbf{u}_k\in\mathcal{X},\ (\mathbf{u}_k)_i=kt,\ k\in[0:\mathsf{M}-1]\right\}\quad(78)$$

是可数的。

证明：设 $(a,b)\in[0:\mathsf{M}-1]^2,\ a\neq b$，并定义

$$\mathcal{B}_{a,b}=\{t\geq0:\{\mathcal{X}-\mathbf{u}_a\}\cap\{\mathcal{X}-\mathbf{u}_b\}\neq\varnothing,\mathbf{u}_a\in\mathcal{X},\ \mathbf{u}_b\in\mathcal{X},(\mathbf{u}_a)_i=at,(\mathbf{u}_b)_i=bt\}.\quad(79)$$

现在，注意到

$$\mathcal{B}=\bigcup_{\substack{(a,b)\in[0:\mathsf{M}-1]^2\\a\neq b}}\mathcal{B}_{a,b}.\quad(80)$$

由于根据假设 $\mathcal{X}$ 是可数的，因此不失一般性地[5] 我们可以假设 $\mathcal{X}=\{\mathbf{x}_w:\ w\in\mathbb{N}\}$。现在注意，若 $t\in\mathcal{B}_{a,\,b}$，则有 $\mathbf{x}_w-\mathbf{u}_a=\mathbf{x}_z-\mathbf{u}_b$，其中某个 $(w,z)\in\mathbb{N}^2$。因此，$\mathcal{B}_{a,\,b}$ 在 (79) 中可写为

$$\mathcal{B}_{a,b}=\{t\geq0:\exists(w,z)\in\mathbb{N}^2,\mathbf{x}_w-\mathbf{x}_z=\mathbf{u}_a-\mathbf{u}_b,\mathbf{u}_a\in\mathcal{X},\ \mathbf{u}_b\in\mathcal{X},(\mathbf{u}_a)_i=at,(\mathbf{u}_b)_i=bt\},\quad(81)$$

and

$$\begin{aligned}\mathcal{B}_{a,b}&\subseteq\{t\geq0:\exists(w,z)\in\mathbb{N}^2,\mathbf{x}_w-\mathbf{x}_z=\mathbf{u}_a-\mathbf{u}_b,\mathbf{u}_a\in\mathcal{X},\ \mathbf{u}_b\in\mathcal{X},(\mathbf{u}_a)_i-(\mathbf{u}_b)_i=(a-b)t\}\\&\subseteq\left\{t\geq0:\exists(w,z)\in\mathbb{N}^2,(\mathbf{x}_w-\mathbf{x}_z)_i=(a-b)t\right\}\\&:=\overline{\mathcal{B}_{a,b}}.\end{aligned}\quad(82)$$

现在，注意到

$$\overline{\mathcal{B}_{a,b}}\subseteq\left\{\frac{(\mathbf{x}_w-\mathbf{x}_z)_i}{a-b}:(w,z)\in\mathbb{N}^2\right\},\quad(83)$$

这意味着 $\mathcal{B}_{a,\,b}$ 是可数的，因此，由（82）可知，$\mathcal{B}_{a,\,b}$ 是可数的。由于可数个可数集的并仍然可数，[46]，$\mathcal{B}$ 在（80）中是可数的，从而完成引理 4 的证明。∎

Appendix G
Proof of (30)

由定理 2 可知

$$\mathsf{H}(P_X,t,1,2)=\sum_{x\in\overline{\mathcal{X}}}\max_{j\in\{0,1\}}p_X(x+jt),\quad(84)$$

其中 $\overline{\mathcal{X}}=\cup_{j\in\{0,1\}}\{\mathcal{X}-jt\}=\{0,1,-t,1-t\}$。不难看出，（84）可表示为

$$\mathsf{H}(P_X,t,1,2)=\begin{cases}1+\max\{p,1-p\} & \text{if } t=1\\2 & \text{if } t\neq1\end{cases},\quad(85)$$

由此得到

$$\mathcal{V}_t\{2-\mathsf{H}(P_X,t,1,2)\}=\begin{cases}\min\{p,1-p\} & \text{if } t\leq1\\0 & \text{if } t>1\end{cases}.\quad(86)$$

因此，

$$\begin{aligned}\overline{\mathsf{V}}(P_X,2)&=\int_0^\infty\frac{t}{2}\mathcal{V}_t\{2-\mathsf{H}(P_X,t,1,2)\}\,\mathrm{d}t\\&=\frac{1}{4}\min\{p,1-p\},\end{aligned}\quad(87)$$

至此，式（30）得证。

[5]当 $\mathcal{X}$ 有限时，只需将 $\mathbb{N}$ 替换为 $[1:\ |\mathcal{X}|]$，即可完成证明。