何时纠错输出编码的乘积解码具有最大后验概率一致性？成对结构与 n-对其余类别编码

Gendo Kumoi *kumoi@vos.nagaokaut.ac.jp*
长冈技术科学大学信息与管理系统工程系

arXiv:2611.33044v1 [cs.LG] 24 Nov 2026

摘要

纠错输出编码（ECOC）通过**求和解码器**或**乘积解码器**组合二元分类器：前者对概率求和，后者对数似然求和。在**无噪声情形** $\boldsymbol{p} = C^\top \boldsymbol{\pi}$下，当二元分类器输出真实后验概率时，我们考察对于哪些编码矩阵 C，每种解码器始终会与贝叶斯最优（最大后验概率）类别保持一致。对于求和解码器，答案已知：编码必须是等距的（James & Hastie，1998年）。本文将回答乘积解码器的这一问题。

关键观察是，比较两个类别时，只需考察它们不同的那些列上的**投票**：求和解码器依据其**样本均值**作出判定，乘积解码器则依据影响函数为反双曲正切的**M估计量**作出判定；并且对于任意码，每一票都是最大后验概率间隔加上一个仅由背景类别决定的偏差。我们的主定理指出，乘积解码器最大后验概率一致，当且仅当对每一类别对而言，这两个类别不同的列与交换这两个类别（或其补集）后得到的列相匹配——即满足**成对结构**——这又等价于该码是**由以下代码构成的并：** n**-对其余类别码**。求和解码器只需让投票的一阶矩平衡；乘积解码器则要求整个投票分布对称，而这会迫使该码在类别的任意置换下保持不变。因此，高效的等距码不在此列：哈达玛码（$M \geq 8$）以 $M-1$ 列达到 Plotkin 界，却完全不具备成对结构，所以对于任意类型的 Hadamard 矩阵，乘积解码器在每一类别对上都可能与最大后验概率判定不一致；对于 Sylvester 类型[¬]，反转间隔至少可以达到一个与 M无关的常数 $\delta \approx 0.0182$ 。所有定理均已在 Lean 4 中通过机器验证。

数值实验确认了这些结果。在六个真实数据集上使用估计概率时，比较结果取决于校准：当每一列都经过校准时，乘积解码器的准确率略高（作为对数优势之和，它对所有列共有的温度保持不变）；当各列的校准误差不同时，使用有界投票的求和解码器表现更好。无论哪种情况，两个解码器仅在接近平局时产生不同结果。

1 引言

纠错输出编码（ECOC）将多类分类简化为二元问题（Dietterich & Bakiri，1995；Allwein 等人，2000）。每个类别 m 都会被分配一个码字 $\boldsymbol{c}_m \in \{0,1\}^n$，并针对每一列 j训练一个二元分类器；解码会聚合输出 p_j，并预测与之“最接近”的码字所对应的类别。

解码器分为两类：**求和解码器**（L_1 解码（Dietterich & Bakiri，1995）），它将列概率相加；以及**乘积解码器**（基于损失的解码（Allwein 等人，2000；Passerini 等人，2004）），它将各列的伯努利对数似然相加。应该使用哪一种？我们的答案是：

> 如果分类器输出真实后验概率，那么求和解码器始终返回最大后验类，当且仅当码是等距的（*James & Hastie，1998b*）；乘积解码器则当且仅当码是*{v1}-vs-Rest* 码的并集时才始终返回最大后验类。

距离较大的短码（如哈达玛码）是等距的，但不属于这种形式，因此对于这类码，只有求和解码器是可靠的。对于存在估计误差的真实分类器，如果对每一列分别进行校准，乘积解码器的表现会略好一些；如果校准误差在各列之间有所不同，则求和解码器表现更好。无论哪种情况，两个解码器的差异都只出现在接近平局时。

我们在无噪声情形下推导这一答案的第一部分，其中二元分类器都是完美的：

$$\boldsymbol{p} = C^\top \boldsymbol{\pi} \in [0,1]^n, \tag{1}$$

其中 $\boldsymbol{\pi} \in \Delta^{M-1}$ 是真实的后验向量。在这里，任何与最大后验概率不一致的情况都源于解码规则和码结构，而非估计误差。

这一答案之所以重要，是因为不同代码的代价差异很大。二元分类器的数量就是码长，而纠错能力取决于码字之间的距离。 n一对其余代码（n二对其余：各列都是将 n 个类别与其余类别划分开的结果）包括一对其余，其码长为 M，但距离仅为 2；还包括 2vR，其距离为 2（$M-2$），代价是使用 $\binom{M}{2}$ 列。Hadamard 码仅用 $M-1$ 列便达到距离 $M/2$，达到了 Plotkin 界（Plotkin，1960年），并应用于纠错输出编码实践中（Yin 等人，2005年；Verma & Swami，2019年；Álvarez-Pérez 和 Callejas-Ramos，2026年）。对于这类高效代码，可以使用哪种解码器?

1.1 贡献

我们的分析基于一个观察：比较类别 m和 m'时，只有码字不同的列才有影响，并且每一列都会为 m 或 m' 投**票**。求和解码器根据这些投票的样本均值作出决策，乘积解码器则通过反双曲正切使用 M估计量。对于任何代码，每一票都等于最大后验概率间隔 $\pi_m - \pi_{m'}$ 加上一个仅由背景质量决定的偏差（第 3 节）。由此我们得到以下结果。

首先是主定理：对于单纯形内部的所有 $\boldsymbol{\pi}$，乘积解码器最大后验概率一致，当且仅当每一类别对都具有成对结构；这又当且仅当该代码是若干 nvR 码的并（定理 1）。成对结构是指：两个类别不同的列，与交换这两个类别（或其补集）后得到的列一一对应。此前的研究（Kumoi 等人，2021）用两侧的类别集合对来表示每一列，并将这种对应关系定义为对称性条件。该研究证明，对称性条件足以保证乘积一致性，并证明了 nvR 码族的一致性。主定理将此结果加强为充要条件，并指出满足该条件的代码：每一类别对都具有成对结构，意味着该代码在任意类别置换下均保持不变，而不变编码恰好是若干 nvR 码的并。作为比较，我们用相同的术语重述求和解码器的已知条件——等距性（James & Hastie，1998b）（命题 1）：求和解码器只需投票的一阶矩平衡，而乘积解码器则要求整个投票分布平衡。因此，每个乘积一致代码都是求和一致的（推论 3），但反之不成立。

其次，哈达玛码（$M \geq 8$）没有单个配对列，因为其各列彼此正交（命题 3），所以对于每一种Hadamard矩阵，乘积解码器在每个¯类别对上都可能与最大后验概率不一致。对于 Sylvester 类型，反转间隔至少可以达到一个与 M 无关的常数 $\delta \approx 0.018175$ （定理 2）。在主导阶上，反转间隔由投票样本的偏度决定（公式（9））；对于典型的后验分布，它大致随 M 衰减（实验 B）。因此，哈达玛码的效率与乘积一致性无法兼得。

第三，在六个真实数据集上进行的实验考察了在存在估计误差时哪种解码器表现更好（第 7 节）。乘积决策是对数优势比的带符号和，对所有列共同适用的温度具有不变性；使用校准分类器时，乘积解码器的准确率略高。当各列的校准误差不同时，采用有界投票的求和解码器表现更好。

第3至第6节中的所有定理、引理、命题和推论，包括附录中的证明，均已使用 Lean 4 和 Mathlib 进行了机器校验（本 arXiv 投稿的附属文件）。

1.2 与先前工作的关系

贝叶斯最优性 L_1 解码（James 和 Hastie）。James & Hastie（1998b）（另见 James & Hastie, 1998a）指出， L_1 解码，即求和解码器，是贝叶斯最优的——当二元分类器均为贝叶斯分类器时，它会返回贝叶斯类别——当且仅当该码为等距码。命题 1 针对每一类别对的顺序，并以弱形式重述了他们的定理。我们给出其简短证明，因为其中使用的投票分解是乘积解码器的出发点。

对估计误差的鲁棒性（Kittler 等人）。 Kittler 等人（1998年）证明，对于条件独立表示，乘法规则是贝叶斯最优的；当后验概率接近先验概率时，求和规则可近似乘法规则，并指出求和规则对估计误差具有鲁棒性。他们关注的是哪种组合规则能抵御含噪估计；我们关注的是，给定真实概率时，每种解码器作为码 C 的函数，是否都能还原最大后验概率。两者互为补充。在我们的真实数据实验中，使用校准分类器时，当两种解码器给出不同结果，乘积解码器有时是正确的那个（第 7 节）。

渐近代理一致性（Ramaswamy 等人）。 Ramaswamy 等人（2014年）给出了编码矩阵需满足的条件，在这些条件下，最小化二元代理损失便能在无限数据极限下恢复贝叶斯分类器。在他们的研究中，解码器是针对编码设计的（先对适当复合损失应用逆链接函数，再进行线性解码），而一致性则由编码矩阵的列空间条件决定。我们的研究方向相反：解码器固定为简单的乘法规则（或求和规则），我们则刻画哪些编码能使其保持一致。因此，我们对固定编码、无噪声概率和简单解码器的研究与之互为补充。

乘积解码器的对称性条件（Kumoi 等人，2021年）。先前研究表明，对于 nvR 码（包括一对其余、2vR、半对半）和穷举码，类别对上的对称性条件可保证乘积解码器具有最大后验概率一致性。我们补充了：（i）该对称性条件（等价于成对结构）对于乘积解码器的必要性，并确定了所有满足该条件的码（定理 1）；（ii）两种解码器之间的包含关系（推论 3）；以及（iii）哈达玛码（$M \geq 8$）缺乏成对结构、由此导致乘积解码器不一致，以及其反转间隔的大小（第 6 节）。具有 $M = 4$ 的哈达玛码在列补意义下与 2vR 码相同，也是唯一具有成对结构的哈达玛码（注 1）。

Passerini 等人（2004）提出了一种基于概率的（类似乘积的）解码器，但没有给出 MAP 一致性条件。Allwein 等人（2000）的基于损失的解码与乘积解码器具有相同的形式，但其定理界定了训练误差和泛化误差。Tax 等人（2000）表明，在两个类别和两个分类器的特殊情况下，求和与乘积是一致的。将求和与乘积视为线性意见池和对数意见池（DeGroot & Mortera，1991；Genest & Zidek，1986），便可将二者都归入聚合规则之列。

第 2 节定义符号和解码器，第 3 节分解投票结果，第 4 节重述求和解码器的条件，第 5 节证明乘积解码器的主定理。第 6 节讨论哈达玛码，第 7 节介绍实验，第 8 节进行讨论，第 9 节作出结论。详细内容见附录。

2 预备知识

2.1 解码器与 MAP 一致性

设 $C \in \{0,1\}^{M \times n}$ 为含有 M 个类别且码长为 n 的编码矩阵（第 m 行为码字 $\boldsymbol{c}_m$），$d(m, m')$为码字之间的汉明距离， $w_m = \sum_j c_{mj}$为行重量。给定二元分类器的概率 $\boldsymbol{p} \in [0,1]^n$，类别 m的两种解码分数为

$$\text{Sum decoder:} \quad V_m = \sum_j \big[c_{mj}\, p_j + (1 - c_{mj})(1 - p_j)\big], \tag{2}$$

$$\text{Product decoder (log-likelihood):} \quad L_m = \sum_j \big[c_{mj} \log p_j + (1 - c_{mj}) \log(1 - p_j)\big], \tag{3}$$

表 1：所研究的码族及其 MAP 一致性（无噪声情形）。求和/乘积：求和/乘积解码器是否具有 MAP 一致性。

Name	形状	等距	成对结构	Sum	乘积
一对其余	$(M,\ M)$	是，$d=2$	Yes	Yes	Yes
2vR	$(M,\ \binom{M}{2})$	是，$d=2\binom{M-2}{M/2-1}$	Yes	Yes	Yes
半对半	$(M,\ \binom{M}{M/2}/2)$	是，$d=\binom{M-2}{M/2-1}$	Yes	Yes	Yes
哈达玛（$M=4$）	$(4,\ 3)$	是，$d=2$	Yes	Yes	Yes
哈达玛（$M\geq 8$）	$(M,\ M-1)$	是，$d=M/2$	No	Yes	No
logM	$(M,\ \lceil \log_2 M \rceil)$	No	No	No	No

对应的预测为 $\hat{m}=\arg\max V_m$ 和 $\hat{m}=\arg\max\ L_m$。求和解码器等价于将 L_1 解码 m m至$\{0,1\}^n$ 中的码字（Dietterich & Bakiri, 1995）；乘积解码器则对各列的伯努利对数似然求和，也就是说，它是一种对数池（Genest & Zidek, 1986）。

我们假设编码矩阵中没有全为 0 或全为 1 的列，且所有码字互不相同。由于乘积得分要求 $p_j\in(0,\ 1)$，关于乘积解码器的论述均假定 $\boldsymbol{\pi}$ 位于单纯形的相对内部（所有分量均为正），在此处这些假设意味着每一列都有 $0<p_j<1$。求和解码器则在整个 Δ^{M-1} 上讨论，包括边界。对列 j $(0\leftrightarrow 1)$ 取补会将 $p_j\to 1-p_j$映射为$p_j\to 1-p_j$，并使 V_m 的 j项和 L_m保持不变：取反列对两个解码器而言都是同一个分类器。因此，我们假设一个码中没有重复列，也没有互为取反的列对（将每一列重复相同次数不会改变两个解码器，而只重复部分列则相当于对列加权）。

最大后验类为 $\hat{m}^{\mathrm{MAP}}=\arg\max\pi_m$。如果对于所考虑范围内的每个 $\boldsymbol{\pi}$，解码器 D (V 或 L) 都满足**最大后验概率一致**，则称其为最大后验概率一致，

$$\operatorname{sign}(D_m-D_{m'})=\operatorname{sign}(\pi_m-\pi_{m'})\qquad(\forall m\neq m'),\tag{4}$$

也就是说，解码器的偏好顺序在每个类别对上都与后验概率顺序一致；在平局情况下 $\pi_m=\pi_{m'}$，我们要求 $D_m=D_{m'}$。对于仅施加 $\pi_m>\pi_{m'}\Rightarrow D_m>D_{m'}$ 的弱形式，命题 1 和定理 1 给出了相同的条件。

此处**无噪声**是指二元分类器无误差地输出真实后验概率的假设（$\boldsymbol{p}=C^{\top}\boldsymbol{\pi}$），与先前工作（Kumoi 等人，2023 年）中的术语相同。真实分类器具有 $\hat{p}_j\neq p_j$，会引入估计误差；我们在第 7 节的真实数据实验（实验 C）中对此进行处理。

2.2 等距码与所研究的码族

定义 1（等距码）。若编码矩阵 C 中每一对不同类别的汉明距离都相同 $d=d(C)$，则称其为**等距**。

表 1 列出了所研究的码族及我们的结论。一对其余 vR 码的列包含将 n个类别分别与其余类别划分的所有方式，包括一对其余（$n=1$）、二对其余（$n=2$；两个类别对其余类别，而非一对一）以及半对半（$n=M/2$）。归一化 Hadamard 码是由阶为 M 的归一化 Hadamard矩阵删除常数列，并将 ±1 映射为{1，0}后得到的 $M\times$（$M-1$）码。类别 0 具有全 1 码字，其他类别的权重为 $M/2-1$，且每个类别对之间的距离均为 $M/2$（Plotkin，1960年）。对于 Sylvester 矩阵（M 为 2 的幂），其码字是单纯形码的码字，并且每一位都经过了补运算。

注 1（带有 $M=4$ 的 Hadamard 码）。带有 $M=4$ 的归一化 Hadamard 码有三列，每列恰好将两个类别置于一侧，因此其列集除补集外与 2vR 码（= 半对半（$M=4$））的列集重合。对于 $M\geq 8$，它与任何 nvR 码都不重合。如第 6 节所示，成对结构的存在也在这一边界处分化为 $M=4$ 和 $M\geq 8$。

3 类别对比较作为投票样本

我们通过同一个对象来表示两个解码器的类别对比较，即由两类别不同的列所投出的投票样本；这些结果对每个码都成立。固定一个类别对（$m,\ m'$），记$T = T_{mm'}$为 $\boldsymbol{c}_m$ 和 $\boldsymbol{c}_{m'}$ 不同的列集，并记 $s = |T| = d$（$m,\ m'$）。在 T 之外，我们有$c_{mj} = c_{m'j}$，因此只有 T 中的列会对 $V_m - V_{m'}$ 和 $L_m - L_{m'}$ 产生贡献。

定义 2（支持度、投票、一致符号）。对于列 $j \in T$，与类别 m 具有相同比特的类别的后验概率总和 q_j：$= \sum_{k:\ c_{kj}=c_{mj}} \pi_k$，是列 j 中 m 的**支持度**，而 $z_j := 2q_j - 1 \in [-1, 1]$ 是列 j 的**投票**。投票为正表示倾向于 m，为负则表示倾向于 m'。对于背景类 $k \notin \{m,\ m'\}$，其**一致符号**为 $\tau_{kj} := +1$（若 $c_{kj} = c_{mj}$），否则为 -1；而 $\boldsymbol{\tau}_j := (\tau_{kj})_{k\notin\{m,m'\}} \in \{\pm 1\}^{M-2}$ 是列 j 的**一致模式**。

若 $c_{mj} = 1$，则 $q_j = p_j$；若 $c_{mj} = 0$，则 $q_j = 1 - p_j$。利用对数优势 $\ell(p) = \log \frac{p}{1-p} = 2\operatorname{artanh}(2p - 1)$ 的奇对称性 $\ell(1 - p) = -\ell(p)$，我们得到

$$V_m - V_{m'} = \sum_{j\in T} z_j, \qquad L_m - L_{m'} = \sum_{j\in T} \ell(q_j) = 2\sum_{j\in T} \operatorname{artanh}(z_j). \tag{5}$$

也就是说，**求和解码器根据投票样本均值的符号来决定类别对，而乘积解码器则根据经过** artanh 处理的投票之和的符号来决定。由于 artanh 严格递增，后者等价于取位置 M 估计方程 $\sum_j \operatorname{artanh}(z_j - \mu) = 0$ 的根 μ 的符号，该方程的影响函数为 artanh。artanh 的斜率在 $z_j \to \pm 1$ 时发散，因此乘积解码器会对有把握的列（投票接近 ± 1）赋予较大权重，这与会抑制离群值的 Huber 型 M 估计量正好相反。

引理 1（投票分解）。设 $\beta := \pi_m - \pi_{m'}$。对于每个 $j \in T$，

$$z_j = \beta + a_j, \qquad a_j := \sum_{k\notin\{m,m'\}} \pi_k\, \tau_{kj}. \tag{6}$$

偏差 a_j 由 背景类别后验概率 $\boldsymbol{b} := (\pi_k)_{k\notin\{m,m'\}}$ 单独决定，不涉及 $\pi_m,\ \pi_{m'}$.

证明。对于 $j \in T$，有 $c_{mj} \neq c_{m'j}$，因此类别 m 属于 q_j，而类别 m' 属于 $1-q_j$。因此 $z_j = q_j - (1-q_j)$ $= \pi_m - \pi_{m'} + \sum_{k\notin\{m,\ m'\}} \pi_k \tau_{kj}$。 □

根据引理 1，固定背景质量 $\boldsymbol{b}$，仅在 m和 m' 之间移动质量，会使投票样本 $\{z_j\}_{j\in T}$平移 β，而不改变其形状。我们考察这一一维族。由于 $\pi_m + \pi_{m'} = 1 - \sum_k b_k =: r$, β在该族上取遍 $[-r, r]$（对于乘积解码器，则为$-r, r$）。图 1 给出了一个示例：具有成对结构（第 5 节）时，样本是对称的，两个解码器给出相同结果；没有这种结构时，反双曲正切会拉伸偏斜样本中的大投票值，只有乘积解码器能够逆转最大后验概率顺序。

偏差之和可用码距表示。

引理 2（距离恒等式）。对于每个背景类 $k \notin \{m, m'\}$，$\sum_{j\in T} \tau_{kj} = d(k, m') - d(k, m)$。因此

$$V_m - V_{m'} = d(m, m')\,(\pi_m - \pi_{m'}) + \sum_{k\notin\{m,m'\}} \pi_k \big[d(k, m') - d(k, m)\big]. \tag{7}$$

证明。对于 $j \in T$，k的比特恰好与 m和 m'中的一个一致。因此，$\sum_{j\in T} \tau_{kj}$等于 T中 k与 m一致的列数，减去它与 m'一致的列数，也就是将距离 d_T 限制在 T上所得的 $d_T(k, m') - d_T(k, m)$。在 T之外，$\boldsymbol{c}_m$ 和$\boldsymbol{c}_{m'}$一致，因此在 T 之外，两种距离中的贡献相同，其差为 $d(k, m') - d(k, m)$。将其代入式 (5) 和式 (6)，即可得到式 (7)。 □

公式（7）的第二项表示每个背景类 k 都会将求和决策拉向离自身更近的类别：如果 k更接近 m（d（$k,\ m$）$< d$（$k,\ m'$）），则决策会向 m偏移。

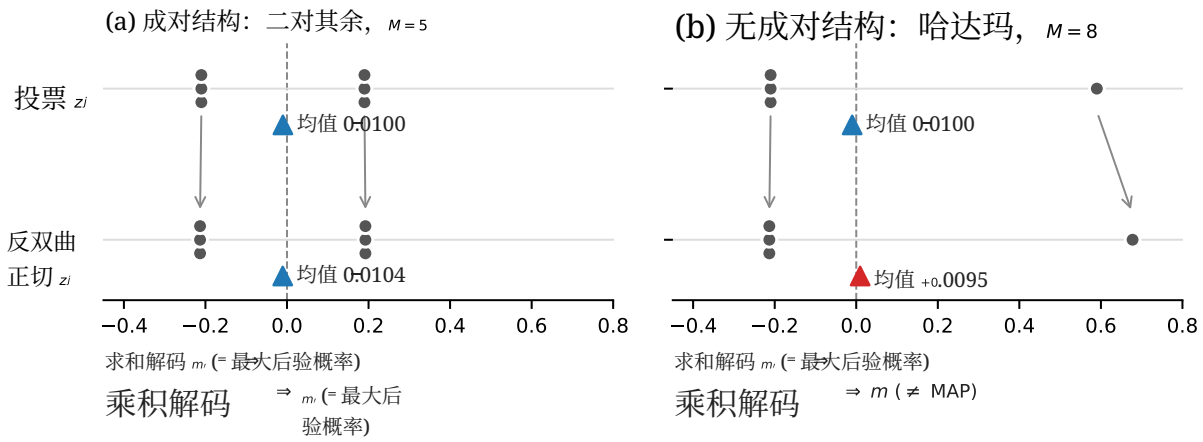


图 1：一个类别对的投票样本，两幅图中的质量相同（在三个背景类别上分别为 $\pi_m = 0.195$、$\pi_{m'} = 0.205$、0.2；最大后验概率选择 m'）。求和解码器根据投票均值 z_j 的符号作出决策，乘积解码器则根据反双曲正切 z_j 均值的符号作出决策。(a) 有成对结构时，样本呈对称分布，两个解码器的结果一致。(b) 对于哈达玛码，当 $M = 8$ 时，样本呈偏斜分布；反双曲正切会拉伸较大的投票值，只有乘积解码器会反转决策。(b) 中的质量取自定理 2 的构造，其中类别索引在$\mathbb{F}_2^t$中进行了转换。

4 求和解码器：等距性是必要且充分的

求和解码器的条件已为人所知（James & Hastie，1998b）。我们用投票的语言重新表述这一条件，从而为与乘积解码器进行比较做好准备。

命题 1（求和解码器；James & Hastie，1998b）。设 $M \geq 3$。求和解码器对所有 $\boldsymbol{\pi} \in \Delta^{M-1}$ 均最大后验概率一致，当且仅当该码是等距的，此时 $V_m - V_{m'} = d(C)(\pi_m - \pi_{m'})$。对于不要求不存在平局的弱形式一致性，该等价关系同样成立。

证明。（充分性）如果该码是等距的，公式（7）的第二项将消失，并且 $V_m - V_{m'} = d(C)(\pi_m - \pi_{m'})$。由于 $d(C) \geq 1$，符号一致。

（必要性）取互不相同的 k, m $\frac{1-t}{2} - \varepsilon \;\; \pi_k = t \qquad 0 \qquad 7 \quad V_m - V_{m'} = 2\varepsilon\, d\; m\; m' + t\,[d\; k\; m' - d\; k\; m\,] \qquad \pi_{m'} =$ 、m'，对于 $t \in (0,1)$，并令 $|\varepsilon| < (1-t)/2$，设置 $\pi_m = \frac{1-t}{2} + \varepsilon$，、，其余均为。根据公式（），（，）（，）（，）。即使在弱一致性下，对于 $\varepsilon > 0$ 这必须为正，而对于 $\varepsilon < 0$ 必须为负，因此令 $\varepsilon \to 0$，得到$t\,[d(k,m') - d(k,m)] = 0$，即 $d(k,m) = d(k,m')$。由于这对任意三个互不相同的类别都成立，共享一个类别的两个类别对具有相等的距离。任意两个类别对 $\{a,b\}$、$\{c,e\}$ 都可以通过一条类别对链连接起来，链中每一对都与下一对共享一个类别，例如 $\{a,b\} \to \{b,c\} \to \{c,e\}$（例如在$b = c$时，更短的链即可），因此所有类别对的距离都相等。 □

对于 $M = 2, V_1 - V_2 = d(1,\ 2)(\pi_1 - \pi_2)$ 对于不同的码字，因此求和解码器始终是最大后验概率一致的 .

等距性也可以写成格拉姆矩阵条件。

推论 1（格拉姆矩阵形式）。当且仅当存在 $\lambda > 0$ 和 $\boldsymbol{\omega} \in \mathbb{R}^M$ 使得 $G = CC^\top = \lambda I + \boldsymbol{\omega}\mathbf{1}^\top + \mathbf{1}\boldsymbol{\omega}^\top$时，码是等距的；此时 $d(C) = 2\lambda$。在等权情形（w_m 为常量）下，$\boldsymbol{\omega}$ 是一个常向量，且条件可简化为 $G = \lambda I + \theta \mathbf{1}\mathbf{1}^\top$。

证明。($\Leftarrow$) $d(m,m') = G_{mm} + G_{m'm'} - 2G_{mm'} = 2\lambda$。($\Rightarrow$) 令 $\lambda := d/2$ 和 $\omega_m := w_m/2 - d/4$，得到 $G_{mm} = w_m = \lambda + 2\omega_m$ 和 $G_{mm'} = (w_m + w_{m'} - d)/2 = \omega_m + \omega_{m'}$。 □

4.1 应用于码族

表 1 中除 logM 外的所有码族都具有等距性。Hadamard 码的行重量不相等（全1行 $M-1$，其余行为 $M/2-1$），但它是等距的，因此求和解码器是最大后验概率一致的。logM 不满足等距性，求和解码器也确实与最大后验概率不一致（第 7 节）。

Hadamard矩阵的阶数是4的倍数（除了1、2之外），但等距性在行删除后仍得以保持，因为删除行但保留列会保持剩余行之间的距离。

推论2（通过行删除得到等距码）：从具有 $M=12$的Hadamard 码中删除两行，可得到具有 $M=10$的等距码。删除包括全1行在内的两行（权重为 $M-1$）后，剩余十行构成一个等权等距码；保留全1行并删除另外两行，则得到一个不等权等距码。在这两种情况下，求和解码器都是最大后验概率一致的。

从数值上看，对于这两种删除方式，都未观察到无噪声求和解码器出现分歧（附录C）。一般而言，对于 M 不是4的倍数的情况，从阶数为 $M'=4\lceil M/4\rceil$ 的Hadamard 码中删除行（实际应用中即采用这种做法，例如Verma & Swami，2019年），会得到一个具有 M 个类别、 $M'-1\le M+2$列且距离为 $M'/2\ge M/2$的等距码：缺失阶数所需的代价至多是增加三列。

注 2（在完全码上进行求和 $\equiv$ 乘积 $\equiv$ 逐位阈值判定）。当 M 是 2 的幂时，logM 码包含由$\{0,1\}^{\log_2 M}$组成的每个字作为码字（即完全码）。此时，每一列对应的项都可以独立最大化，因此求和解码器、乘积解码器和逐位阈值判定完全一致；这两种解码器绝不会彼此不同，但它们与最大后验概率的分歧，等于逐位边缘最大后验与联合最大后验之间的差距。因此，logM 码的分歧并非求和与乘积之别，而是将码长缩短至 $\log_2 M$所导致的结构性后果。

5 乘积解码器：主定理

求和解码器只关注投票的均值；乘积解码器则通过反双曲正切，还能反映投票样本的形状。我们先刻画固定背景的一维族上的乘积反转（命题 2），再给出适用于所有 $\boldsymbol{\pi}$ 的充要条件（定理 1）。本节中，$\boldsymbol{\pi}$ 始终位于单纯形的相对内部。

5.1 单个类别对上的反转

固定一个类别对 $(m,\ m')$ 和一个背景 $\boldsymbol{b}$（所有元素均为正，$r=1-\sum_k b_k>0$），并令

$$\Psi_{\boldsymbol{b}}(\beta):=\sum_{j\in T}\operatorname{artanh}(\beta+a_j), \tag{8}$$

其中 a_j 是引理 1 中的偏差。由式 (5) 和 (6)，$L_m-L_{m'}=2\Psi_{\boldsymbol{b}}(\beta)$ 在该族上。**命题 2**（单个类别对上的反转）。$\Psi_{\boldsymbol{b}}$ 在 $\beta\in(-r,\ r)$ 上严格递增。

1. 若 $\Psi_{\boldsymbol{b}}(0)=0$，则对该族中的每个 β，$\operatorname{sign}(L_m-L_{m'})=\operatorname{sign}\beta$（不会发生反转）。2. 若 $\Psi_{\boldsymbol{b}}(0)>0$，则 $\Psi_{\boldsymbol{b}}$ 有唯一的根 $\beta^*<0$，且当且仅当 $\beta\in(\max(\beta^*,\ -r),\ 0)$ 时，乘积解码器会反转最大后验概率顺序。该区间非空。$\Psi_{\boldsymbol{b}}(0)<0$ 的情形可通过交换 $m\leftrightarrow m'$ 得出。

命题 2（证明见附录 A）表明，类别对上的反转完全取决于偏差样本 $\{a_j\}$ 在反双曲正切意义下是否以零为中心（即 $\sum_j \operatorname{反双曲正切}(a_j)=0$）。即使偏差的均值为零（如所有等距码的情况），也不一定以零为中心，因为反双曲正切是非线性的。

5.2 成对结构与主定理

定义 3（成对结构）。对于一个类别对 (m,m')，若列 $j,j'\in T$**是配对的，**则列 j'等于将列 j 的第 m个和第 m'个条目交换后得到的列，或该列的补。若 T 可以恰好划分为配对列，则该码对于 (m,m') 具有**成对结构**。

允许取反列，是因为取反列表示同一个分类器（第 2 节）。先前工作的对称性条件（Kumoi 等人，2021）是根据每一列两侧的类集族定义的，因此也无法区分取反列；由于对于 $M \geq 3$，交换在T上没有不动点，所有类别对都具有成对结构等价于满足该条件。

引理 3（成对结构与一致模式）。令 $M \geq 3$。当且仅当$\boldsymbol{\tau}_{j'} = -\boldsymbol{\tau}_j$时，列 j、$j' \in T$ 成对。因此，类别对（m、m'）具有成对结构，当且仅当一致模式的多重集$\{\boldsymbol{\tau}_j\}_{j\in T}$在符号翻转 $\boldsymbol{\tau} \mapsto -\boldsymbol{\tau}$下保持不变。

证明。在 ± 1 表示 $h_{kj} := 2c_{kj} - 1$ 中，我们有 $\tau_{kj} = h_{mj}h_{kj}$ 以及 $h_{m'j} = -h_{mj}$（$j \in T$）。列 j 的交换列在 m 处的元素为 $-h_{mj}$，在 m' 处的元素为 h_{mj}，背景元素为 h_{kj}，因此其一致模式为（$-h_{mj}$）$h_{kj} = -\tau_{kj}$。对所有元素取补会同时翻转 $h_{mj'}$ 和 $h_{kj'}$ 的符号，但一致模式保持不变。因此，成对列满足 $\boldsymbol{\tau}_{j'} = -\boldsymbol{\tau}_j$。反之，假设 $\boldsymbol{\tau}_{j'} = -\boldsymbol{\tau}_j$。如果 $h_{mj'} = -h_{mj}$，背景元素为 $h_{kj'} = h_{kj}$，且列 j' 等于交换列。如果$h_{mj'} = h_{mj}$，则 $h_{kj'} = -h_{kj}$，且列 j' 等于交换列的补。对于第二个结论，当 $M \geq 3$ 时有 $\boldsymbol{\tau}_j \neq -\boldsymbol{\tau}_j$，因此如果多重集在符号翻转下保持不变，则模式$\boldsymbol{v}$ 的每次出现都能与 $-\boldsymbol{v}$的一次出现一一匹配，反之亦然。□

我们将每一列 j 读作类别的一种划分 $\{A_j, A_j^c\}$，其中 $A_j = \{k\colon c_{kj} = 1\}$。

定义 4（vR 码的并集）。如果一个码的各列读作划分后，恰好是某个非空集合 $N \subseteq \{1, \ldots, \lfloor M/2 \rfloor\}$ 中 n 对应的 n vR 码的划分，则称该码为**由 nvR 码构成的并集**；等价地，对于每一列，具有相同侧大小的每一种划分也都是一列。

例如，每个 nvR 码（$N = \{n\}$）、穷举码（$N = \{1, \ldots, \lfloor M/2 \rfloor\}$），以及一对其余与二对其余（$N = \{1,2\}$）都是这样的码。

定理 1（乘积解码器；主定理）。设 $M \geq 3$，取 $\boldsymbol{\pi}$ 属于单纯形的相对内部，并设该码满足第 2 节中的约定（码字互异，且没有重复列或取反列）。以下条件等价：（1）乘积解码器是最大后验概率一致的；（2）乘积解码器以弱形式（$\pi_m > \pi_{m'} \Rightarrow L_m > L_{m'}$）是最大后验概率一致的；（3）每个类别对都具有成对结构；（4）该码是由 nvR 码构成的并集。

证明。(3)⇒(1)：根据引理 3，成对的 j，j'满足 $a_{j'} = \langle \boldsymbol{b},\ \boldsymbol{\tau}_{j'} \rangle = -a_j$。因此，对于每个背景 $\boldsymbol{b}$，$= \sum_{j\in T} \operatorname{artanh}(a_j) = 0$；根据命题 2(1)，所有类别对的符号对所有 $\boldsymbol{\pi}$都一致。(1)⇒(2)：显然。

(2)⇒(3)：我们证明其逆否命题。若某个类别对 $(m,\ m')$ 不具有成对结构，只需证明对于某个背景 $\boldsymbol{b}$，$\Psi_{\boldsymbol{b}}(0) \neq 0$；于是根据命题 2(2)，乘积解码器会在 β 的某个非空区间上严格颠倒最大后验概率顺序，这与弱一致性矛盾。大致思路如下：将 $\Psi_{\boldsymbol{b}}(0) = \sum_{j\in T} \operatorname{artanh}\langle \boldsymbol{b},\ \boldsymbol{\tau}_j \rangle$展开为关于 $\boldsymbol{b}$ 的幂级数，次数为 κ 的项是 $\frac{1}{\kappa}\sum_{j\in T}\langle \boldsymbol{b},\ \boldsymbol{\tau}_j \rangle^\kappa$（$\kappa$ 为奇数）。若对于每个背景都有 $\Psi_{\boldsymbol{b}}(0) = \bar{0}$，则所有这些项都为零，且一致模式分布的所有奇数阶矩均为 0。对于有限集上的分布，所有奇数阶矩均为零等价于在符号翻转下具有对称性，这意味着具有成对结构（引理 3）。附录 B 通过将 $\boldsymbol{b}$ 限制在单条射线 $\boldsymbol{b} = t\boldsymbol{v}$ 上，使这一论证严谨化；这样只需要单变量幂级数和范德蒙德矩阵。

(3)⇔(4)：包含 $c_{mj} = c_{m'j}$ 的列在交换 m 和 m' 后保持不变。因此，(m, m') 的成对结构恰好意味着，将列视为划分时，列集合在对换 m 和 m' 下保持不变。因此，对每一类别对而言，成对结构都意味着在每个对换下保持不变；又由于对换生成对称群，故在每个类别置换下都保持不变。划分 $\{A, A^c\}$ 在所有置换下的轨道，包含所有侧大小相同的划分，即 nvR 码中取$n = \min(|A|, M - |A|)$ 的列。因此，在所有置换下保持不变的划分集合必然是这类轨道的并，这就是 (4)；反过来，nvR 代码的并在每个置换下都保持不变。□

若不采用关于重复列的约定，(4) 表示：列中某个划分的重数仅取决于其侧大小。

推论 3（乘积一致性 $\Rightarrow$和一致性）。 乘积解码器具有最大后验概率一致性的码是等距的，因此求和解码器也具有最大后验概率一致性。反之则不成立（哈达玛码， $M \geq 8$；第 6 节）。

证明。 在成对结构下，成对项 $\boldsymbol{\tau}_{j'} = -\boldsymbol{\tau}_j$在求和中相互抵消，因此对于每个背景类 k都有 $\sum_{j \in T} \tau_{kj} = 0$。根据引理 2， $d(k,\ m) = d(k,\ m')$，而命题 1 的必要性部分中的论证表明，该码是等距的。

□

命题 1 和定理 1 对投票偏置 $\{a_j\}$提出了不同强度的要求。求和解码器要求每个背景类的一致符号的求和 为零（一阶矩平衡）；这就是等距性。乘积解码器要求一致模式本身在符号翻转下具有对称性（所有奇数阶矩平衡）；这就是成对结构。仅有一阶平衡对于乘积解码器而言并不足够，并且通过条件 (4)，每个类别对的投票分布的对称性会转化为整个码在所有类别置换下的对称性。

5.3 应用于码族

推论4（具有成对结构的码族）。 nvR 码（包括一对其余、二对其余和半对半）、穷举码、它们的并集，以及具有 $M = 4$ 的 Hadamard 码，对每一类别对都具有成对结构，因此两个解码器对它们都是最大后验概率一致的。

证明。 它们各自都是若干 nvR 码的并集：穷举码是对所有 n 取并集，而具有 $M = 4$ 的 Hadamard 码的列对应于将四个类别划分为两对的三种划分，也就是具有 $M = 4$ 的半对半（注 1）。定理 1 的条件 (4) 适用。对于包含类别 0（全1行）的具有 $M = 4$ 的 Hadamard 码类别对， T 的两列都位于类别 0 所在的一侧，因此该类别对不是通过交换得到的，而是通过交换后取补得到的。 □

与上一个案例一样，某些列仅在视为 0/1向量时与交换列的补配对；用列两侧的类集表示每一列即可消除这种区别（带有 $M = 4$ 的 Hadamard 码与先前工作中的 2vR 示例（$M = 4$）是同一个矩阵）。

成对结构很脆弱。从 2vR 中删除一列会破坏其配对，而对于在该列上取值不同的类别对，成对结构也会丢失。这与实验 A 中观察到的乘积不一致率从 0% 跃升至几个百分点相一致。另一方面，行删除会保留成对结构（交换和补关系在剩余行上仍然成立）。

注 3（乘积一致性的代价）。 根据条件 (4)，乘积一致代码包含一个完整的 nvR 块。对于 $M \geq 8$，因此除一对其余之外的每个乘积一致代码至少有 $\binom{M}{2}$ 列（$\binom{M}{n} \geq \binom{M}{2}$ 对于 $2 \leq n \leq M-2$，以及 $\binom{M}{M/2}/2 \geq \binom{M}{2}$ 对于 $M \geq 8$），而一对其余的距离仅为 2。Hadamard 码用 $M - 1$ 列达到距离 $M/2$，列数比 2vR 少约 $M/2$ 倍，但没有成对结构（第 6 节）。

6 H哈达玛码：缺少成对结构与反转间隔

哈达玛码是等距的，因此求和解码器是一致的。对于乘积解码器，我们证明，对于 $M \geq 8$，没有任何列存在配对列，并评估由此产生的分歧可能有多大。

6.1 缺少成对结构

命题 3（哈达玛码没有成对结构）。 对于任意包含 $M \geq 8$的归一化 Hadamard 码（不仅限于 Sylvester 类型），以及任意类别对 $(m,\ m')$， T 中没有任何列存在配对列。

证明。 在 ± 1 表示中，假设列 $j,\ j' \in T$ 成对配对。交换后的列仅在条目 m 和 m'处与列 j 不同，因此逐元素乘积 $\boldsymbol{h}_{\cdot j} \circ \boldsymbol{h}_{\cdot j'}$ 恰有 2 个条目等于 -1，

而交换后列的补恰有 $M-2$。另一方面，Hadamard矩阵的不同列彼此正交，因此逐元素乘积的和为 0，即恰有 $M/2$ 个条目等于-1。仅当 $M=4$ 时，我们才有 $M/2 \in \{2,\ M-2\}$。 □

推论 5（Hadamard 码上的乘积不一致性）。对于每个满足 $M \geq 8$ 的归一化 Hadamard 码，求和解码器具有最大后验概率一致性，而乘积解码器则不具有这种一致性。此外，对于每个类别对，都存在一个 $\boldsymbol{\pi}$，使得乘积解码器颠倒最大后验概率顺序。

证明。根据命题 3，没有任何类别对具有成对结构，因此由定理 1 中 (2)⇒(3) 的证明可知，对于每个类别对，都存在一个 $\boldsymbol{\pi}$，使得顺序发生颠倒。 □

Hadamard 码的列分布最为分散：任意两列在一半的行上都不同。这增大了码距并带来等距性（一阶平衡），但也使两列不可能通过交换关联（成对结构）。使 Hadamard 码成为优良纠错码的这一性质，因而排除了成对结构。正交性也是 Verma & Swami (2019) 选择 Hadamard 码的原因，因为正交码字使他们的概率估计具有可接受性；然而，对于乘积解码器的最大后验概率一致性而言，列的正交性反而构成了障碍。

对于 Sylvester 类型，某个类别完全不会出现在比较统计量中 .

命题 4（不可见类别）。对于含有 $M=2^t$ 的 Sylvester Hadamard 码，将类别按 $\mathbb{F}_2^t$ 编号。对于每个类别对 $(m,\ m')$，类别 $k^*=m\oplus m'$ 恒等满足 $c_{k^*j}=0$ 于 T 上。因此，p_j 在 T 上不涉及 π_{k^*}，且 $L_m-L_{m'}$ 不依赖于 π_{k^*}。

证明。类别 $a\in\mathbb{F}_2^t$ 在非零列 $j\in\mathbb{F}_2^t\setminus\{0\}$ 上的码字为 $c_{aj}=(1+\chi_a(j))/2$，其中 $\chi_a(j)=(-1)^{\langle a,\ j\rangle}$（因为 $\chi_0\equiv 1$，$\boldsymbol{c}_0=\mathbf{1}$）。根据特征乘积法则 $\chi_m\chi_{m'}=\chi_{m\oplus m'}$，$T=\{j:\ \chi_m(j)\chi_{m'}(j)=-1\}=\{j:\ \chi_{m\oplus m'}(j)=-1\}$，所以 $c_{k^*j}=0$ 对于 $j\in T$。 □

对于包含类别 0 和 m 的一对 $(0,m)$，$k^*=m$：L_0-L_m 不涉及所比较类别本身的 π_m，而求和解码器给出 $V_0-V_m=\frac{M}{2}(\pi_0-\pi_m)$，其中明确涉及 π_m。

6.2 反转间隔可以有多大?

根据命题 2，带反转的 β 的范围宽度（反转间隔）是 $\Psi_{\boldsymbol{b}}$ 的根 $|\beta^*|$（除非其超过可行范围 r）。对于等距码，偏差的均值为 $\langle a\rangle:=\frac{1}{s}\sum_{j\in T}a_j=0$（引理 2）。将反双曲正切$(y)=y+y^3/3+y^5/5+\cdots$代入 $\Psi_{\boldsymbol{b}}(\beta^*)=0$，并在 $\langle\cdot\rangle=\frac{1}{s}\sum_{j\in T}$ 条件下求解 β^*，首项为

$$\beta^*\approx-\frac{\frac{1}{3}\langle a^3\rangle+\frac{1}{5}\langle a^5\rangle+\cdots}{1+\langle a^2\rangle+\cdots}. \tag{9}$$

分子只包含偏差样本的奇数阶矩（偏度）；偶阶矩（方差）则进入分母，抑制反转间隔。因此，对称的偏差样本不会产生反转（成对结构），而要获得更大的反转间隔，就需要偏斜的样本。

人们可能会预期，随着 M 增大，背景质量会被稀释，反转间隔也会缩小至 0；对于典型的后验概率，确实如此（见下文）。然而，对于 Sylvester 类型，将背景质量集中在少数几个类别上以最大化偏度，可使反转间隔保持为一个与 M 无关的常数。

定理 2（与维度无关的最坏情况下反转裕度下界）。令 δ 为下式的唯一正根

$$3\ \operatorname{artanh}(0.2+\delta)=\operatorname{artanh}(0.6-\delta) \tag{10}$$

($\delta\approx 0.018175$)。对于每个 Sylvester Hadamard 码（$M=2^t\geq 8$）、每个类别 $m\neq 0$，以及每个 $\varepsilon\in(0,\delta)$，都存在一个 $\boldsymbol{\pi}$ (位于相对内部)，满足 $\pi_m-\pi_0=\varepsilon$，使得类别 m 是最大后验概率类别（即 $\boldsymbol{\pi}$中数值最大的类别），但乘积解码器将类别 0 排在 m之上。也就是说，最坏情况下的反转裕度至少为 δ，且 δ 不依赖于 M。

证明的关键在于，将质量 $1/5$ 放置在一个由背景类别构成的 $\mathbb{F}_2$-闭合三元组 $\{b_1, b_2, b_1 \oplus b_2\}$ 的每个元素上。根据特征乘积法则，偏差 a_j 随即退化为两个点： $+0.6$ (比例为 $1/4$) 和 -0.2 (比例为 $3/4$)，而 Ψ 的根方程变为式 (10)，其中既不涉及 s ，也不涉及 M。完整证明见附录 A。

数值上， $|\beta^*| = 0.018175$ 在机器精度范围内一致，适用于 $M \in \{8、16、32、64\}$；同一构造不仅给出类别对反转，也会导致 argmax 不一致。只要 MAP 间隔（最大两个后验概率之差）不超过 $0.911\,\delta \approx 0.0166$, argmax 不一致就会持续存在，且与 M无关；此时乘积解码器会选出背景三元组中的一个类别。下文中，实验测得的反转间隔 δ^* 指的是 argmax 不一致时的 MAP 间隔。

δ 是一个已证明的下界，而非最坏情况上确界。我们从码字等权重的类别对以及抽样得到的不一致点出发，运行了多起点 SLSQP。对于每个 M $\in\{8、12、16、20、28、32、48、64\}$，算法都找到了间隔 $\delta^*_{\text{worst}} \in [0.017, 0.025]$（均值为 0.0232，变异系数为 0.10），且每个结果都经过解码器验证。这些数值对每个 M而言都大于构造所得的 0.0166。因此，Sylvester 类型中出现 argmax 不一致时，最坏 MAP 间隔至少为 0.0166；局部搜索未发现高于 0.025 的值，其精确值尚未确定。对于非 Sylvester 类型，三元组特征恒等式不适用，因此我们没有已证明的下界常数（数值上观察到类似的 $\Theta(10^{-2})$）。

6.3 典型间隔随 M衰减

通过随机抽样发现的不一致点中最大的 δ^* （下文的典型值）大致随 M 衰减（狄利克雷：在 $M = 8$ 时为 0.0135，在 $M = 16$ 时为 0.0066，在 $M = 48$ 时约为 0.003，64 ）。乘积 $\delta^* \times M$ 的范围为 0.10–0.18，而非恒定值，因此纯粹的 $1/M$ 律并不成立。这种衰减与公式（9）一致：随着 M 增大，背景质量分散到许多类别中，偏差样本的偏度 $\langle a^3 \rangle$减小。最坏情况下的构造位于单纯形的一个角落，质量仅分布在五个类别上，而随机抽样很少能命中该区域。

因此，最坏情况下的反转裕度不会随着 M 增大而消失，但发生反转的 $\boldsymbol{\pi}$ 局限于平局附近的一片狭窄区域；在典型分布下，该区域会随 M 增大而缩小。

7 实验

我们考察三个问题：A，定理在数值上是否成立（无噪声情形）？ B，Hadamard 码上的逆转幅度有多大? C，使用真实分类器时，哪种解码器更准确? A 和 B 验证了理论；C 回答了引言中问题的后半部分（在存在估计误差时哪种解码器更好）。具体设置详见附录C。

7.1 A：定理的数值验证

我们对无噪声概率 $\boldsymbol{p} = C^\top \boldsymbol{\pi}$ 进行抽样（每个代码有四组序列，以及 M：狄利克雷 $\alpha \in \{0.3, 1.0\}$；另有两个基于高斯数据构建的后验概率，每个各抽取 200,000 个样本；对于列数较多的代码，则各抽取 100,000 个样本），并测量每种解码器与最大后验概率之间的不一致率（图2）。求和解码器在等距码上从未出现不一致（除 logM 外均为等距码）。乘积解码器在具有成对结构的一对其余、二对其余和半对半代码上从未出现不一致，在带有 $M = 4$ 的 Hadamard 码上也从未出现不一致；对于带有 $M \geq 8$ 的哈达玛码，其不一致率虽小但不为零（ $M = 8, 16, 32$ 分别为 0.145%、0.265% 和 0.300%）。对于两种解码器，logM 的不一致率都很高（ $M = 4$ 至 32 时为 14.0% 至 68.6%）。

实验还表明，成对结构较为脆弱（第 5 节）。从二对其余中随机删除一列，会使乘积不一致率在 $M = 8$ 处升至 4.72%，在 $M = 16$ 处升至 3.12%（10 个随机种子的均值）。由于删除操作还会破坏等距性，求和率也会类似上升（5.26%、4.48%）。

为了逐类别对验证主定理，我们对 49 个代码中的全部 977 个类别对（nvR 型代码；哈达玛$M \in \{4、8、12、16\}$，其中 $M = 12$ 为 Paley 型；logM；24 个随机码；以及通过对 nvR 码进行行删除得到的 10 个代码）分别进行独立检验：（a）测试成对结构（引理 3）；（b）搜索使乘积解码器严格颠倒最大后验概率顺序的 $\boldsymbol{\pi}$ （对接近平局的情况进行采样并作局部优化）。

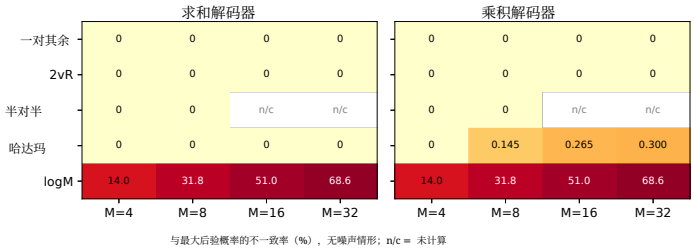


图2：无噪声情形下与最大后验概率的不一致率（%，四组序列的均值）。左：求和；右：乘积。求和解码器在所有等距码上均为0（除了logM外的所有码），乘积解码器仅在具有成对结构的代码上为0（一对其余、二对其余、半对半、带 $M = 4$ 的哈达玛）（命题1和定理1）。n/c：未计算。半对半有 $\binom{M}{M/2}/2$ 列（在 $M = 16$ 时为6435，约为 3×10^8，在 $M = 32$ 时），因此只计算了 $M \leq 8$；由于具有成对结构，其比率根据定理为0。

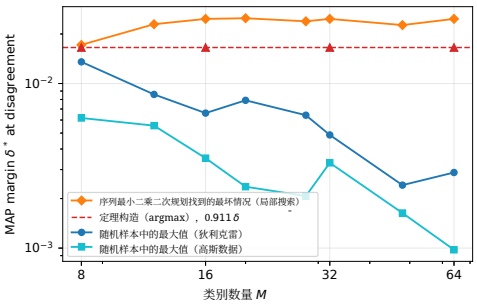


图 3：反转间隔对 M （Hadamard 码）的依赖关系。典型值（通过狄利克雷采样和高斯数据找到的不一致点处的最大间隔）大致随 M衰减；通过局部 SLSQP 搜索找到的最差值则保持在 $[0.017, 0.025]$ 以内，与 M无关。红色虚线：定理 2 的构造中具有 argmax 不一致的最大 MAP 间隔 $0.911\,\delta$ ；三角形：Sylvester 类型 M（定理适用于此类）。

我们使用原始解码公式确认了每一个颠倒候选。对于不具有成对结构的773个类别对，均发现了颠倒；对于具有成对结构的204个类别对，则一个也没有发现。在哈达玛码中，所有类别对在 $M = 4$ 时均具有成对结构，而在 $M = 8$、12、16时均不具有成对结构（与命题3一致）。

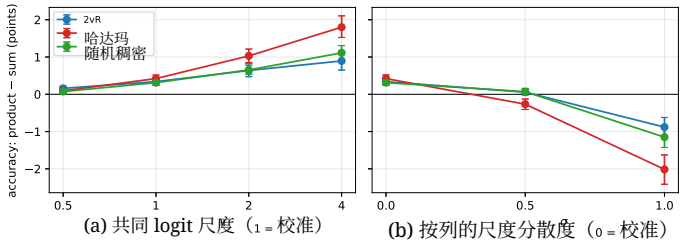
7.2 B：反转间隔的大小

我们测量了Hadamard码在 $M \in \{8、12、16、20、28、32、48、64\}$ 下的典型情形（狄利克雷采样、高斯数据）和最差情形（多起点SLSQP；图3）中的反转间隔。典型值（采样不一致点处的最大 δ^*）大致从0。0135（$M = 8$）衰减至约0。003（$M = 48$、64），而最差情形对于所有 M均保持在 $[0.017, 0.025]$ 以内：从 $M = 8$ 处典型（狄利克雷）值的1.3倍，到 $M = 48$、64处约9倍。在针对 $M \in \{8、16、32\}$的500、000个无噪声样本中，不一致点处真实的MAP间隔最多为0。00845（构造值0。0166的0。51倍）。随机后验分布很少达到最差情形的构造（单纯形的一个顶点）。

7.3 C：真实分类器中哪种解码器更准确?

我们在六个数据集（Letter数据集、Isolet数据集、Pendigits数据集、MNIST、Fashion-MNIST、CIFAR-100的MobileNetV2特征）上，根据估计概率比较了解码器。对于随机类别子集 $M \in \{8，16，24，32\}$（视数据情况而定），我们针对一对其余、二对其余、哈达玛和随机稠密码（$n = \lceil 10 \log_2 M \rceil$）的每一列训练了逻辑回归。每一列都根据验证数据进行温度校准（下文称为“已校准”），随后通过对所有列的logit进行缩放来引入校准误差：(a) 使用一个共同因子 $\gamma \in \{0.5，1，2，4\}$（$\gamma > 1$：统一过度自信）；或 (b) 使用列特定因子 $\gamma_j = \exp(\sigma g_j)$,$g_j \sim \mathcal{N}(0，1)$，$\sigma \in \{0，0.5，1\}$（校准误差在各列之间变化）。我们比较了每种条件下5个随机种子的测试准确率。条件 (a) 和相关假设在实验开始前就已确定；条件 (b) 则是探索性的，是在观察到 (a) 的结果后加入的。在一对其余中，两种解码器完全相同（都对列 m的logit取argmax），因此我们报告其他代码的结果。

图4展示了结果。**经过校准后，乘积解码器的准确率更高**：对于所有代码，都高出 $+0.3$至 $+0.4$ 个百分点（95% 自助法置信区间不包含0）；在39个（数据集、 M、码）单元格中，有24个根据McNemar检验显著更优，而求和解码器在任何单元格中都没有显著优势。**统一过度自信**



(a) 共同 logit 尺度（$_{1}$ = 校准）　(b) 按列的尺度分散度（$_{0}$ = 校准）

图 4：使用真实分类器时乘积解码器与求和解码器的准确率差异（点；正值表示乘积解码器更准确）。六个数据集上的均值和 95 % 自助法置信区间，每个数据集有 M，并使用 5 个随机种子。(a) 所有 logit 乘以同一个因子 γ：乘积决策保持不变，而求和解码器的性能会随着 γ 增大而下降。(b) 特定列的因子 $\gamma_j = \exp(\sigma g_j)$：随着分散度增大，求和解码器的表现会更好。

会扩大乘积解码器的优势（在 $\gamma = 4$ 时高出 $+0.9$ 至 $+1.8$ 个百分点）。乘积决策是 logit 的带符号和（公式 (5)），不会因共同因子而改变。相比之下，随着 sigmoid 函数趋于饱和，求和解码器会逐渐转向基于汉明距离的多数投票，这与其准确率下降的情况一致。**各列校准误差不同时，优势方向会逆转**。在 $\sigma = 1$ 时，求和解码器高出 0.9 至 2.0 个百分点，并在 39 个单元格中的 36 个显著更优。这与解码器对投票的加权方式相符：乘积解码器会对置信度高的列赋予较大权重（第 3 节），因此少数列的过度自信可能主导其决策；而求和解码器的投票在 $[-1, 1]$ 范围内有界。我们没有对各列误差如何组合进行建模；第 8 节将此列为未来工作。逆转在哈达玛码上最为明显。

不一致点处的估计间隔（求和分数中最高两项之间的差距）很小，约为总体均值（中位数）的 7%；在至少有 20 个不一致点的 156 个（数据集、 M、码、 γ）条件中，有 139 个条件下该间隔低于总体均值：不一致点集中在平局附近。

8 讨论

8.1 应该使用哪种解码器?

根据无噪一致性这一标准，答案很明确。求和解码器要实现精确的最大后验概率一致性，必须使用等距码（命题 1）；乘积解码器则必须使用具有成对结构的码，也就是若干 nvR 码的并集（定理 1），而成对结构蕴含等距性（推论 3）。因此，求和解码器适用于更广泛的码类。特别是，哈达玛码是一种距离为 $M/2$、长度为 $M - 1$ 的高效等距码，但当 $M \geq 8$ 时，它不具有成对结构（命题 3）；若使用哈达玛码且需要精确一致性，应选择求和解码器。 nvR 码对两种解码器都具有一致性，但删除列会破坏成对结构和等距性（实验 A），因此必须使用完整的码。

乘积反转只会在有限的范围内产生影响。 $\delta \approx 0.018$ 是一个已证明的下界，表明反转间隔可以达到这么大（定理 2），而非保证安全的上界；局部搜索已发现高达 0.025 的反转。在随机后验分布下，不一致点处的最大后验概率间隔至多约为构造值的一半（实验 B）；无论如何，反转都局限于接近平局的情况，此时后验概率几乎相等。

在存在估计误差时，情况有所不同（实验 C）。使用校准逻辑回归时，乘积解码器的准确率略高，而且所有列都受到统一过度自信影响时，其决策保持不变；当各列的校准误差不同时，求和解码器表现更好。这与乘积解码器将投票视为无界对数几率、而求和解码器将其视为有界概率的解释一致。实践中，如果每列的概率都能单独校准，就使用乘积解码器；如果各列的校准质量不一（或无法检查），则使用求和解码器。无论如何，两者的差异约在 2 个百分点以内（校准时为 0.3–0.4 个百分点），且意见不一致的情况仅出现在接近平局时。Kittler 等人（1998年）关于求和对估计误差具有鲁棒性的论断，可以理解为求和对各列校准误差不一具有鲁棒性。

8.2 局限性与未来工作

我们的理论假设处于无噪声情形 $\boldsymbol{p} = C^{\top}\boldsymbol{\pi}$，这使解码规则和码结构能够与分类器误差区分开来，但真实分类器并不符合这一假设：在使用估计概率时，即使是二对其余码——在无噪声情形下，两种解码器都与最大后验概率一致——也会产生不同的求和决策和乘积决策（实验 C）。等距性和成对结构对估计误差的鲁棒性尚不明确。实验 C 表明，在存在估计误差时，哪种解码器更胜一筹可能取决于各列误差如何组合；未来的工作将通过错误指数（可靠性函数）对此进行分析。

对于哈达玛码的反转间隔，尚待解决的问题包括：(1) 存在 argmax 不一致时精确的最坏情况最大后验概率间隔（至少为 0.0166；局部搜索发现其至多为 0.025），以及对构造中阈值 $0.911\,\delta$ 的解析推导（目前仅有数值结果）；(2) 非 Sylvester 类型的下界常数（从数值上看，观察到类似的 $\Theta(10^{-2})$）；以及 (3) 典型情形下精确的衰减率。

9 结论

我们刻画了无噪声情形$\boldsymbol{p} = C^{\top}\boldsymbol{\pi}$下纠错输出编码中乘积解码器的最大后验概率一致性。将类别对比较视为投票样本后，求和解码器依据样本均值进行决策，而乘积解码器则通过反双曲正切使用 M估计量；对于任意码，每次投票都等于最大后验概率间隔加上一个仅由背景决定的偏差。求和解码器是最大后验概率一致的，当且仅当该码是等距码（James & Hastie，1998b）（命题 1）。我们的主定理（定理 1）表明：乘积解码器是最大后验概率一致的，当且仅当每个类别对都具有成对结构；这将既有工作的充分对称条件（Kumoi 等人，2021）加强为充要条件；同时，这也当且仅当该码是 nvR 码的并。每个乘积一致代码都是求和一致的（推论 3）。达到 Plotkin 界的哈达玛码（$M \geq 8$）没有成对结构，因为其列彼此正交（命题 3）；对于 Sylvester 类型，反转间隔可以至少为一个与 M 无关的常数 $\delta \approx 0.018175$（定理 2）。对于 M 不是 4 的倍数的情形，行删除可以构造等距码，代价至多是增加三列。

实验验证了实验 A 的代码上的命题 1，以及 977 个类别对上的主定理。在存在估计误差的情况下，当每一列都经过校准时，乘积解码器的准确率略高；而当各列的校准误差不同时，求和解码器表现更好。对于真实分类器，如果可以检查每一列的校准情况，选择乘积解码器似乎是稳妥的；否则应选择求和解码器。两种解码器的核心差异在于：求和解码器只要求投票满足一阶矩平衡（等距性），而乘积解码器要求整个投票分布具有对称性（成对结构）；这种对称性最终只保留了 nvR 码。

致谢

作者衷心感谢 Shigeichi Hirasawa、Manabu Kobayashi 和 Hideki Yagi 多年来在研讨会中的讨论；本文是始于早期工作（Kumoi 等人，2021年）的系列讨论的成果。本研究得到 JSPS KAKENHI 项目编号 JP24K20833 的支持。

参考文献

Erin L. Allwein、Robert E. Schapire 和 Yoram Singer。将多类别问题化为二元问题：间隔分类器的一种统一方法。《机器学习研究杂志》，1:113–141，2000年。

Lorena Álvarez-Pérez 和 Álvaro Callejas-Ramos。沃尔什序列作为多类别问题的直接纠错输出编码二分。《人工智能工程应用》，173:114442，2026年。doi：10.1016/j.engappai.2026.114442。

Morris H. DeGroot 和 Julia Mortera。最优线性意见池。《管理科学》，37(5):546–558，1991年。doi：10.1287/mnsc.37.5.546。

Thomas G. Dietterich 和 Ghulum Bakiri。通过纠错输出编码解决多类学习问题。《人工智能研究杂志》，2:263–286，1995。doi：10.1613/jair.105。

Christian Genest 和 James V. Zidek。组合概率分布：批评与注释书目。*Statistical Science*，1(1):114–135，1986。doi：10.1214/ss/1177013825。

Gareth James 和 Trevor Hastie。错误编码与替换协议。载于 神经信息处理系统进展 *10*（*NIPS 1997*），第542–548 页，1998年a。

Gareth James 和 Trevor Hastie。错误编码方法与 PICTs。*Journal of Computational andGraphical Statistics*，7(3):377–387，1998b。doi：10.1080/10618600.1998.10474782。

Josef Kittler、Mohamad Hatef、Robert P. W. Duin 和 Jiří Matas。关于组合分类器。*IEEE Trans-actions on Pattern Analysis and Machine Intelligence*，20(3):226–239，1998年。doi：10.1109/34.667881。

Gendo Kumoi、Hideki Yagi、Manabu Kobayashi、Masayuki Goto 和 Shigeichi Hirasawa。多标签分类问题的 ECOC方法优化研究（日文）。*IPSJ*数学建模及其应用汇刊（*TOM*），14(3):1–10，2021年。URL `https://ipsj.ixsq.nii.ac.jp/records/212227`。

Gendo Kumoi、Hideki Yagi、Manabu Kobayashi、Masayuki Goto 和 Shigeichi Hirasawa。基于有噪和无噪二元分类器的纠错输出编码性能评估。*InternationalJournal of Neural Systems*，33(2):2350004，2023。doi: 10.1142/S0129065723500041。

Andrea Passerini、Massimiliano Pontil 和 Paolo Frasconi。关于核机器纠错输出码的新结果。*IEEE Transactions on Neural Networks*，15(1):45–54，2004。doi: 10.1109/TNN.2003. 820841。

Morris Plotkin。具有指定最小距离的二元码。*IRE*信息论汇刊，6(4):445–450，1960年。doi: 10.1109/TIT.1960.1057584。

Harish G. Ramaswamy、Balaji Srinivasan Babu、Shivani Agarwal 和 Robert C. Williamson。关于面向多类学习问题的基于输出编码的学习算法的一致性。载于第*27*届学习理论会议（*COLT*）论文集，*PMLR*第35卷，第885–902页，2014年。

David M. J. Tax、Martijn van Breukelen、Robert P. W. Duin 和 Josef Kittler。通过平均或相乘组合多个分类器？*Pattern Recognition*，33(9):1475–1485，2000年。doi: 10.1016/S0031-3203(99) 00138-7。

Gunjan Verma 和 Ananthram Swami。纠错输出码能够提高深度神经网络的概率估计能力和对抗鲁棒性。载于神经信息处理系统进展第*32*卷 *(NeurIPS 2019)*，2019年。

An-Rong Yin、Xiang Xie 和 Jingming Kuang。基于支持向量机在多分类问题中使用 Hadamard纠错输出码。载于*Interspeech 2005* 会议论文集，第 3125–3128 页，2005年。doi: 10.21437/Interspeech.2005-672。

A 正文证明

命题 2 的证明。当 $\boldsymbol{\pi}$ 位于相对内部时，$0 < q_j < 1$，因此 $\beta + a_j = z_j \in (-1,\ 1)$ 且 $\Psi'_{\boldsymbol{b}}(\beta) = \sum_{j\in T}(1-(\beta+a_j)^2)^{-1} > 0$。$\Psi_{\boldsymbol{b}}$ 定义在开区间 $B = (-1 - \min_j a_j,\ \ 1 - \max_j a_j)$ 上，在该区间内对所有 j 都有 $\beta + a_j \in (-1,\ 1)$，并且其在左端点处发散至 $-\infty$、在右端点处发散至 $+\infty$，因此它在 B 上有唯一的根。由于 $|a_j| \le \sum_k b_k = 1 - r$，我们有 $(-r,\ \ r) \subset B$。

乘积解码器当且仅当 $\Psi_{\boldsymbol{b}}(\beta) > 0$ 时偏好 m，而最大后验概率当且仅当 $\beta > 0$ 时偏好 m。(1) 若 $\Psi_{\boldsymbol{b}}(0) = 0$，则由单调性可得符号 $\Psi_{\boldsymbol{b}}(\beta) =$ 符号 β。(2) 若 $\Psi_{\boldsymbol{b}}(0) > 0$，则根位于 $\beta^* < 0$；对于 $\beta \in (\beta^*,\ 0)$，会出现反转 $\Psi_{\boldsymbol{b}}(\beta) > 0 > \beta$，而对于 $\beta \notin [\beta^*, 0]$，符号一致。由于该族上的 β 被限制在 $(-r,\ \ r)$，反转区间为 $(\max(\beta^*,\ \ -r),\ \ 0)$，且该区间非空，因为 $\max(\beta^*,\ -r) < 0$。

定理 2 的证明。任取一个类别 $m \neq 0$，并取类别对 $(0,\ m)$。由于 $\boldsymbol{c}_0 = \mathbf{1}, T = \{j:\ c_{mj} = 0\}$以及 $|T| = M/2$; 根据命题 4 的特征标形式，$T = \{j:\ \chi_m(j) = -1\}$，且背景类 k 的一致符号为 $\tau_{kj} = 2c_{kj} - 1 = \chi_k(j)$。选取线性无关的 $b_1,\ b_2 \in \mathbb{F}_2^t \setminus \{0,\ m\}$，使得$b_1 \oplus b_2 \neq m$（$M \geq 8$ 时可行），并令 $b_3 := b_1 \oplus b_2$。在三元组的每个元素上放置背景质量 $\rho = 1/5$，并将其他背景类别设为 0。于是 $\pi_0 + \pi_m = 2/5$ 且 $\beta = \pi_0 - \pi_m$。

根据引理 1，偏差为 $a_j = \rho(\chi_{b_1}(j)+\chi_{b_2}(j)+\chi_{b_3}(j))$。根据乘法规则 $\chi_{b_3} = \chi_{b_1}\chi_{b_2}$, $\chi_{b_1}+\chi_{b_2}+\chi_{b_3} =(1 + \chi_{b_1})(1 + \chi_{b_2}) - 1 \in \{3,\ -1\}$，且仅当 $\chi_{b_1} = \chi_{b_2} = 1$ 时取值为 3。由于 $m,\ b_1,\ b_2$ 线性无关，因此 $(\chi_{b_1},\ \chi_{b_2})$ 在 $\{\pm1\}^2$ 上关于 $T = \{\chi_m = -1\}$ 均匀分布。因此

$$a_j \in \{+0.6 \text{ (proportion } 1/4),\ -0.2 \text{ (proportion } 3/4)\}, \qquad \langle a \rangle = 0,$$

并且 $\Psi_{\boldsymbol{b}}(\beta) = \frac{s}{4}$反双曲正切$(\beta+0.6)+\frac{3s}{4}$反双曲正切$(\beta-0.2)$，$\Psi_{\boldsymbol{b}}(0) = s[\frac{1}{4}$反双曲正切 $0.6+\frac{3}{4}$反双曲正切$(-0.2)] \approx 0.0212\, s > 0$。根据命题 2(2)，存在一个根 $\beta^* < 0$。记 $\beta^* = -\delta$，并利用反双曲正切的奇对称性，根方程变为反双曲正切$(0.6 - \delta) = 3$反双曲正切$(0.2 + \delta)$，即式 (10)，其中不涉及 s。由于 $r = 2/5 > \delta$，反转区间为 $\beta \in (-\delta,\ 0)$，即 $0 < \pi_m - \pi_0 < \delta$。于是 $\pi_m = (2/5+\varepsilon)/2 > 1/5$ 超过三元组质量 $1/5$ 和 π_0，因此 m 是最大后验类。

该构造仅在五个类别上具有非零质量，并且位于单纯形边界上。将一小部分质量 $\eta > 0$ 均匀分配给其余 $M - 5$ 个类别，并将这五个质量缩放 $1 - (M - 5)\eta$ 以得到$\boldsymbol{\pi}_\eta$。每个 p_j 在 T 上关于 η连续，而在 $\eta = 0$ 处，取值 $p_j \in \{0.4,0.8\}$ 与端点之间均有正距离。对于固定的 $\varepsilon \in (0, \delta)$，反转现象以及 m 为最大后验概率类这一事实在 $\eta = 0$ 处均以严格不等式成立，因此对所有充分小的 $\eta > 0$ 仍成立（该缩放将 $\pi_m - \pi_0$ 变为 $\varepsilon(1 - (M - 5)\eta)$；从 $\varepsilon/(1 - (M - 5)\eta)$ 出发则恢复 ε）。

B 主定理中 (2)⇒(3) 的证明

假设某个类别对 (m, m') 不具有成对结构，并假设弱一致性成立；我们将导出矛盾。令 $U := \{\boldsymbol{b}:\ b_k > 0,\ \sum_k b_k < 1\}$ 为可容许背景的集合。

步骤 1。$\Psi_{\boldsymbol{b}}(0) = 0$对每个 $\boldsymbol{b} \in U$都成立。事实上，若存在某个 $\boldsymbol{b} \in U$使得 $\Psi_{\boldsymbol{b}}(0) > 0$，则命题 2(2)（若$\Psi_{\boldsymbol{b}}(0) < 0$，再结合将 $m \leftrightarrow m'$代入的同一命题）会导致在 β的某个非空区间上最大后验概率顺序发生严格逆转，这与弱一致性矛盾。

步骤 2（单条射线）。将背景类别编号为 $k_1, \ldots, k_{M-2}$，并令 $v_{k_i} := 2^{i-1}/2^M$。由于 $v_k > 0$和 $\sum_k v_k < 1$，对于 $t \in (0,1)$，有 $t\boldsymbol{v} \in U$，而步骤 1 给出

$$\sum_{j \in T} \operatorname{artanh}(t\, u_j) = 0, \qquad u_j := \langle \boldsymbol{v}, \boldsymbol{\tau}_j \rangle \in (-1, 1).$$

对于 $|t| < 1$，左侧等于收敛幂级数 $\sum_{\kappa\text{奇数}} t^\kappa \sum_{j \in T} u_j^\kappa$，它是 t的奇函数。因此，它在 $(-1, 1)$ 上恒为零，而幂级数系数的唯一性给出：对于每个奇数 κ，都有 $\sum_{j \in T} u_j^\kappa = 0$。

步骤 3（对称性）。设 S 为值 $\pm u_j$ 和 $\nu(x)$ 的有限集合：$= \#\{j \in T:\ u_j = x\} - \#\{j \in T:\ u_j = -x\}$。当 κ 为奇数时，$\sum_{x \in S} \nu(x)\, x^\kappa = 2\sum_{j \in T} u_j^\kappa = 0$；当 κ 为偶数时，两项相消。这些关于 $\kappa = 0, 1, \ldots, |S| - 1$ 的方程，由范德蒙德矩阵的非奇异性可推出 $\nu \equiv 0$；也就是说，多重集$\{u_j\}_{j \in T}$ 关于 0 对称。

步骤 4（回到一致模式）。设 A_j 为背景类集合，其中 $\tau_{kj} = +1$，$u_j = 2^{-M}\big(2\sum_{k_i \in A_j} 2^{i-1} - (2^{M-2} - 1)\big)$。根据二进制展开的唯一性，$u_j$决定了 A_j，进而决定 $\boldsymbol{\tau}_j$，而 $-u_j$对应于 $-\boldsymbol{\tau}_j$。因此，一致模式的多重集在符号翻转下保持不变；根据引理 3，类别对具有成对结构，这与假设矛盾。 □

C 实验详情

表 2 总结了第 7 节实验的主要设置。

表 2：实验设置。

Exp.	主要设置
A	无噪声一致性：一对其余、二对其余、半对半、哈达玛、logM $\times$ $M \in \{4、8、16、32\}$；四个系列（狄利克雷 $\alpha \in \{0.3、1\}$ 以及两个由高斯数据生成的系列），每组 200、000 个样本（码的列数较多时为 100、000 个，代码）；列删除 $k = 1$、从二对其余中删除 2 列（10 个随机种子）；对 49 个代码和 977 个类别对进行配对结构检查
A’	行删除（推论2）：从哈达玛 $M = 12$ 中删除两行，得到 $M = 10$（等权重和不等权重情形）；狄利克雷 $\alpha \in \{0.3,1\}$，每种情形采样 50,000 次；求和解码器不一致为零点
B	哈达玛 $M \in \{8, 12, 16, 20, 28, 32, 48,64\}$；典型值（不一致点处的最大间隔）基于狄利克雷 $\alpha \in \{0.3,1\}$（50,000 $\times$ 3 个随机种子）和高斯数据（20,000 $\times$ 4 种方差， $\times$3 个随机种子）；通过多起点 SLSQP 得到的最差值（在 100 个起点 $\times$ 3 个随机种子中取最大值）；间隔分布位于来自 500,000 个无噪声样本的不一致点，针对 $M \in \{8, 16,32\}$
C	Letter 数据集、Isolet 数据集、Pendigits 数据集、MNIST 和 Fashion-MNIST（PCA-50，每类最多 1500 个样本）、CIFAR-100（MobileNetV2 特征，PCA-100）；$M \in \{8, 16, 24,32\}$（视类别数量而定）；类别子集和划分（训练集/验证集/测试集 $= 60/20/20$）在 5 个随机种子下变化；按列进行逻辑回归（正则化强度的倒数为 1，使用标准化特征），并进行温度校准在验证数据上；(a) 公因子 $\gamma \in \{0.5,1, 2, 4\}$，(b) 列特定因子 $\gamma_j = \exp(\sigma g_j)$，$\sigma \in \{0,0.5,1\}$（每项各抽取 3 次）。假设及 (a) 的设计在实验开始前已确定

可复现性。所有定理、引理、命题和推论的 Lean 4 形式化证明均提供于本 arXiv 投稿的附属文件中。数值性结论由验证脚本确认，用于复现实验的代码将在接收后发布到公共代码仓库。乘积得分计算为二分类器输出的伯努利对数似然之和（式（3）），即采用精确对数似然规则，而不是 $\log(1+p)$ 类型的近似。CIFAR-100 特征提取一次后缓存，以便重复使用。