量化约束串接与受约束系统的覆盖半径

Dor Elimelech，IEEE 学生会员；Tom Meyerovitch；Moshe Schwartz，IEEE 高级会员

arXiv:2302.02363v1 [cs.IT] 5 Feb 2023

摘要

我们提出了一种在受约束系统中实现纠错的新框架。该方案称为量化约束串接（QCC），其核心思想是将纠错码的码字嵌入受约束系统的过程视为一种（有噪、不可逆的）量化过程。与传统方法（如串接和反向串接）不同，传统方法中编码到受约束系统的过程是可逆的。QCC 能够纠正的信道错误数可能随块长n线性增长，优于当前已知先进方案所能实现的$O(\sqrt{n})$。对于给定的受约束系统，QCC 的性能取决于该受约束系统的一个新基本参数——其覆盖半径。

受 QCC 启发，我们从组合和概率两个角度研究受约束系统的覆盖半径。我们借助遍历理论揭示了受约束系统覆盖半径的一种有趣刻画，并利用这一等价刻画建立了可高效计算的覆盖半径上界。

索引词

受限系统、覆盖半径、纠错码、马尔可夫链、滑动块码。

I. 引言

CONSTRAINED 码常用于通信和存储系统，以减轻数据相关错误的发生。在许多信道中，有些词比其他词更容易出错，因此通过避开这些词，可以减少错误数量。这类码称为受约束码。虽然使用受约束码可以显著减少数据相关错误的发生，但在许多现实场景中，传输的数据仍可能受到数据无关错误的破坏。

处理数据损坏的一种著名策略是将纠错码与受约束码结合起来。过去 40 年来，这一策略得到了广泛研究（例如参见 [7], [11], [15], [17], [20], [25]），近年来，由于人们对 DNA存储系统的兴趣日益浓厚，它再次受到关注。过去几年中，许多研究都探讨了用于 DNA 存储的纠错受约束码 [5], [8], [9], [22], [24], [29]–[31], [34], ，并特别关注 GC含量约束和连续段长度（均聚物）约束。

尽管近年来针对特定约束族的纠错约束码构造与分析取得了长足进展，但目前已知的、用于在受约束系统中实现纠错的一般框架仍然寥寥无几（综述见 [27, 第 8] 章）。反向串接法（有时也称为修正串接，见 [7], [15], [20], [25]）是这类框架中的一个重要例子。在该方法中，纠错编码位于受约束编码器之后。最近，有人提出了反向串接法的一种改进方法，称为分段反向串接 [17]。这些方法的一项主要局限在于纠错能力。尽管 [17] 中介绍的最新方法能够纠正 $O(\sqrt{n})$ 个错误（其中n 为块长），但目前尚无通用技术能够纠正受约束系统中的 $\Theta(n)$ 个错误。

受这一缺口的启发，我们提出一种用于在受约束系统中实现纠错的替代策略——量化约束串接（QCC）；即使存在 $\Theta(n)$ 个错误，该策略也同样适用。我们提出的方法背后的基本思想很简单：我们建议将信息嵌入受约束介质的过程视为量化过程，而非编码过程。在传统方法（包括串接和反向串接）中，受约束词表示待传输并受到纠错保护的数据。因此，受约束编码器是可逆的，并且除了纠错码带来的码率损失外，还会造成额外的码率损失。在 QCC 中，我们将受约束词视为信息经过量化过程后得到的受损版本。因此，受约束量化器不会造成码率损失。相反，纠错码的参数经过设计，能够同时处理信道造成的错误和量化过程造成的错误。

设 $\mathcal{B}_n \subseteq \Sigma^n$ 是某个有限字母表 Σ 上所有长度为 n 的受约束词组成的集合。进一步假设，$r < n$ 是一个整数，且对于任意词 $y \in \Sigma^n$，都存在一个对应的词 $x \in \mathcal{B}_n$，使得其汉明距离 $d(x,y) \leqslant r$。

Dor Elimelech 任职于以色列贝尔谢巴 8410501 的内盖夫本-古里安大学电气与计算机工程学院（电子邮件：doreli@post.bgu.ac.il）。Tom Meyerovitch 任职于以色列贝尔谢巴 8410501 的内盖夫本-古里安大学数学系（电子邮件：mtom@bgu.ac.il）。Moshe Schwartz 任职于以色列贝尔谢巴 8410501 的内盖夫本-古里安大学电气与计算机工程学院（电子邮件：schwartz@ee.bgu.ac.il）。本文的部分内容曾提交至2023年国际信息论研讨会。

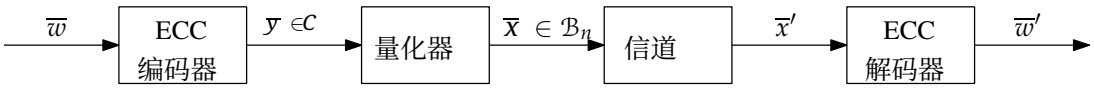

图 1. 描述我们所提出的纠错过程的框图。

给定一个能够纠正$t > r$ 个错误的纠错码 $C \subseteq \Sigma^n$，我们提出以下受限纠错过程（见图 1）：

- 编码：给定信息字w，使用纠错码的编码器将其映射为码字$y \in C$。
- 量化：给定$y \in C$，找到一个受约束词$x \in \mathcal{B}_n$，使得$d(y,x) \leqslant r$，并传输x。
- 信道：在信道输出端，观察到$x' \in \Sigma_n$，即x的受损版本。
- 解码：使用C 的解码器处理x'，并得到w'。

如果信道引入的错误不超过 $t - r$ 个，即 $d(x, x') \leqslant t - r$，那么 $d(y, x') \leqslant t$。由于 C 能纠正 t 个错误，因此 $w = w'$，也就是说，可以纠正 $t - r$ 个信道错误。因此，我们关注最小的数 r，使得空间中的任意词都能量化为 $\mathcal{B}_n$ 中的一个词，且最多改变 r 个坐标。用编码理论的术语来说，这个量记为 $R(\mathcal{B}_n)$，称为 $\mathcal{B}_n$的覆盖半径。利用这一技术，现在可以纠正 $\Theta(n)$ 个错误：假设存在一个受限系统，使得对所有 n，均有 $R(\mathcal{B}_n) \leqslant \rho \cdot n$，并且 $(C_n)_{n\in\mathbb{N}}$是一个能够纠正 $\delta \cdot n$ 个错误的码序列，其中某个 $\delta > \rho$成立。采用上述方案，因而可以纠正 $(\delta - \rho) \cdot n$ 个信道错误，这与 n 呈线性关系。

为了进一步理解所提出的方案，研究受限系统的覆盖半径至关重要，这也是本文的目标。下面概述本文的贡献。第III节中，我们给出受限系统覆盖半径的组合定义，并考察其一些基本性质。我们还发现了一个耐人寻味的现象：我们给出一个容量为正的受限系统示例，其覆盖半径与容量为零的重复码相同。

受这一现象的启发，我们在第IV节采用概率方法，并定义本质覆盖半径。我们证明，这一版本排除了导致上述非预期现象的极端情况。我们运用遍历理论框架，给出本质覆盖半径的另一种刻画。在第V节中，我们利用本质覆盖半径的遍历理论定义，在常见情形下推导其上界。通过马尔可夫链方法，我们导出一个一般上界，该上界可通过求解线性规划高效计算。我们还利用滑动块码函数给出上界。我们证明，在本原情形下，这些上界渐近地达到本质覆盖半径。

II. 预备知识

本文中，我们用小写字母 x 表示标量和符号，用带上划线的小写字母 x 表示有限长字，用粗体小写字母 $\mathbf{x}$ 表示双无限序列。我们用大写字母 X 表示受限系统。对于双无限序列 $\mathbf{x} = \ldots, \mathbf{x}_{-1}, \mathbf{x}_0, \mathbf{x}_1, \ldots$ 以及 $n \leqslant m$，我们将子词 $\mathbf{x}_n^m \triangleq \mathbf{x}_n, \ldots, \mathbf{x}_m$记作（有限词的情形也类似）$x_n^m$。我们用 Σ 表示有限字母表，用 $[n] \triangleq \{0, 1, \ldots, n-1\}$ 表示。

长度为 n、字母取自 Σ 的词的集合记作 Σ^n。若 $u \in \Sigma^n$，我们将其字母的索引记为 $[n]$，即 $u = u_0, u_1, \ldots, u_{n-1}$。对于任意 $v, u \in \Sigma^n$，我们将汉明距离定义为

$$d(\overline{u}, \overline{v}) \triangleq |\{i \in [n] \mid \overline{u}_i \neq \overline{v}_i\}|.$$

以r 为半径（相对于汉明距离）、以x 为中心的球记作$B_r(x)$。码$C \subseteq \Sigma^n$ 的覆盖半径是满足以下条件的最小整数r：以r为半径、以C中的码字为中心的球的并覆盖整个空间。也就是说，

$$R(C) \triangleq \min\left\{ r \in \mathbb{N} \cup \{0\} \;\middle|\; \bigcup_{\overline{c}\in C} B_r(\overline{c}) = \Sigma^n \right\}.$$

在 Σ^n 中，与C 中最近码字的距离为$R(C)$的元素通常称为深洞（例如，参见 [12, 定义 2.1。3]）。

接下来讨论受限系统。这类系统通常在符号动力系统的框架下研究（例如参见[23], [27]）。在典型的一维情形中，我们有一个有限字母表 Σ，并将由 Σ构成的双无限序列空间记作 $\Sigma^{\mathbb{Z}}$，视为一个紧致可度量拓扑空间，并赋予积拓扑（其中 Σ 具有离散拓扑）。系统 $\Sigma^{\mathbb{Z}}$ 上的动力学由移位变换$T : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$实现，其定义为

$$(T\mathbf{x})_n \triangleq \mathbf{x}_{n+1},$$

它是该系统上的一个拓扑同胚。对于有限词 $x \in \Sigma^n$，我们令 $[x]$ 表示由x定义的柱集，它是

$$[\overline{x}] \triangleq \left\{ \mathbf{x} \in \Sigma^{\mathbb{Z}} \;\middle|\; \mathbf{x}_0^{n-1} = \overline{x} \right\}. \tag{1}$$

子移位（或移位空间）$X \subseteq \Sigma^{\mathbb{Z}}$ 是一个紧子空间，在移位变换下保持不变。对于子移位X，X 的语言是所有能作为X中某个元素的子词出现的有限词所组成的集合。也就是说

$$\mathcal{B}(X) \triangleq \left\{ \overline{x} = (x_0 \dots x_k) \,\middle|\, \exists \mathbf{x} \in X, n \in \mathbb{Z} \text{ such that } \mathbf{x}_n^{n+k} = \overline{x}, k \in \mathbb{N} \cup \{0\} \right\}.$$

语言中长度为n 的词集合记为 $\mathcal{B}_n(X)$ $\mathcal{B}(X) \cap \Sigma^n$。$X$ 的拓扑熵（也称容量）定义为以下极限（根据 Fekete 引理，该极限存在）

$$h(X) \triangleq \lim_{n\to\infty} \frac{\log_{|\Sigma|} |\mathcal{B}_n(X)|}{n}.$$

在我们的设定中，受限系 统是指可以通过某个标号图上的游走实现的移位空间。

定义 1 若存在一个有限有向图$G = (V,E)$ 和一个标记函数$L : E \to \Sigma$，使得$X \subseteq \Sigma^{\mathbb{Z}}$ .....则称移位空间$X \subseteq \Sigma^{\mathbb{Z}}$ 为受限系统（或 sofic 移位）。

$$X = X_G \triangleq \left\{ (L(e_i))_{i\in\mathbb{Z}} \,\middle|\, (e_i)_{i\in\mathbb{Z}} \textit{ is a bi-infinite directed path in } G \right\}.$$

若任意两个顶点之间都存在一条有向路径，则称标号图$G = (V,E,L)$ 为不可约图。若一个不可约图的所有环长度的最大公约数为 1，则称其为本原图。众所周知（例如，参见 [23, 定理 4.5。8]），不可约图是本原图，当且仅当存在 $n \in \mathbb{N}$，使得对于任意两个顶点$v,v' \in V$，都存在一条从v 到v'、长度为n 的有向路径。

定义 2 若存在一个不可约（分别为本原）标号图 G，使得受限系统 $X \subseteq \Sigma^{\mathbb{Z}}$ 满足$X = X_G$，则称其为不可约（分别为本原）。

一类特别值得关注的受限系统，是由有限个局部约束定义的系统。这类系统称为有限型系统，正式定义如下：

定义 3 若存在某个 $m \in \mathbb{N}$ 及一个有限的禁词集合 $\mathcal{F} \subseteq \Sigma^m$，使得受限系统 $X \subseteq \Sigma^{\mathbb{Z}}$ 是所有不含来自$\mathcal{F}$的任何禁用模式的双无限序列所构成的集合，则称其为有限型系统（SFT）。即

$$X = X_{\mathcal{F}} \triangleq \left\{ \mathbf{x} \in \Sigma^{\mathbb{Z}} \,\middle|\, \forall n \in \mathbb{Z}, \mathbf{x}_n^{n+m-1} \notin \mathcal{F} \right\}.$$

关于纠错码覆盖半径的编码理论文献十分丰富（例如，参见 [12] 及其中的大量参考文献）。然而，正如后文将会显而易见的，我们所讨论的情形截然不同，因为我们考虑的是与受限系统相关联的有限分组长度码序列。如上所述，这些受限系统以双无限序列为数学形式进行表述和研究。因此，我们将借用遍历理论中的工具来处理当前问题。

III. The Covering Radius of a Constrained System

我们首先定义集合$B \subseteq \Sigma^n$ 相对于另一个集合$A \subseteq \Sigma^n$ 在汉明度量下的覆盖半径。

定义 4 设$A,C \subseteq \Sigma^n$，则C 相对于A 的覆盖半径定义为

$$R(C,A) \triangleq \min\left\{ r \in \mathbb{N} \cup \{0\} \,\middle|\, A \subseteq \bigcup_{\overline{x}\in C} B_r(\overline{x}) \right\} = \max_{\overline{y}\in A} \min_{\overline{x}\in C} d(\overline{x},\overline{y}).$$

若$A = \Sigma^n$，则$R(C,A)$ 就是C的常规覆盖半径，记作$R(C)$。

对于受限系统$X,Y \subseteq \Sigma^{\mathbb{Z}}$，我们将$X$ 相对于Y 的渐近覆盖半径定义为n元组从X 中取、相对于Y中的n元组的渐近归一化覆盖半径。

定义 5 设$X,Y \subseteq \Sigma^{\mathbb{Z}}$ 为移位空间，则定义

$$R(X,Y) \triangleq \liminf_{n\to\infty} \frac{R(\mathcal{B}_n(X), \mathcal{B}_n(Y))}{n}, \tag{2}$$

这里提醒一下， $\mathcal{B}_n(X)$ 和 $\mathcal{B}_n(Y)$ 分别是来自X 和Y 的长度为n 的子词。

在典型的编码理论框架中，覆盖半径被视为有限长度n汉明空间中单个码的一个性质。另一方面，受限系统可以与一列码相关联，这些码是

固定长度的受约束词集合。上述定义的受限系统覆盖半径，实际上是与之对应的这列码的（归一化）覆盖半径的渐近值。

我们指出，在本工作的绝大部分内容中，我们取Y为整个空间 $\Sigma^{\mathbb{Z}}$，然而所述结果对一般的受约束系统同样成立。在考虑覆盖半径的定义时，一个立即出现的问题是，式（2）中的极限是否存在。现在我们将证明，当X 或Y 为本原时，答案是肯定的。为证明这一点，我们回顾近似次可加序列的 Fekete 引理的以下推广。

引理 6（De Bruijn–Erd˝os 条件 **[14], [16]**）设$(a_n)_{n\in\mathbb{N}}$ 是一个非负实数序列，对所有m、$n \in \mathbb{N}$满足

$$a_{n+m} \leqslant a_n + a_m + f_{n+m},$$

其中$(f_k)_{k\in\mathbb{N}}$是一个满足 $\sum_{k=1}^{\infty} f_k k_2 < \infty$的数列。则数列$(a_n/n)_{n\in\mathbb{N}}$收敛。

命题 7 假设X、$Y \subseteq \Sigma^{\mathbb{Z}}$是受限系统。如果$X$或$Y$为本原，则$R(X、Y)$ 定义中的下极限实际上是一个极限：

$$R(X,Y) = \lim_{n\to\infty} \frac{R(\mathcal{B}_n(X), \mathcal{B}_n(Y))}{n}.$$

证明：我们先考虑 X 是本原的情形，且它由本原有限标号图 G = (V, E, L) 表示。由于 G 是本原的，存在一个足够大的 N，使得任意两个顶点之间都存在一条长度为 N 的有向路径。这意味着，对于任意两个词 u, v $\in \mathcal{B}$(X)，都存在 w $\in \Sigma^N$，使得 u w v $\in \mathcal{B}$(X)。对于 n $\in \mathbb{N}$，记 $a_n = R(\mathcal{B}_n(X), \mathcal{B}_n(Y))$。我们将证明，对于所有 n > N 和 m $\geqslant$ 1，都有 $a_{n+m} \leqslant a_n + a_m + N$。为了证明所需的不等式，只需证明以下命题：给定 y $\in \mathcal{B}_{n+m}$(Y)，存在 x $\in \mathcal{B}_{n+m}$(X)，使得 d(x, y) 6 $a_n + a_m + N$。根据 a_m 和 a_{n-N} 的定义，存在 u $\in \mathcal{B}_m$(X) 和 v $\in \mathcal{B}_{n-N}$(X)，使得

$$d(u, y_0^{m-1}) \leqslant a_m \quad \text{and} \quad d(v, y_{m+N}^{m+n-1}) \leqslant a_{n-N}.$$

我们还可以直接根据定义看出，$(a_k)_{k\in\mathbb{N}}$ 是一个非递减序列，因此$a_{n-N} \leqslant a_n$。令$w \in \Sigma^N$ 满足$x = uwv \in \mathcal{B}_{n+m}(X)$。根据汉明度量的性质

$$\begin{aligned} d(x, y) &= d(v, y_0^{m-1}) + d(w, y_m^{m+N-1}) + d(u, y_{m+N}^{m+n-1}) \\ &\leqslant a_m + N + a_{n-N} \leqslant a_m + a_n + N. \end{aligned}$$

当 n $\geqslant$ 1 且 m > N 时，可以进行对称的分析。由于剩余情形，即 m、n 均 6 N，包含有限种情况，令 c = max(N, max{$a_{m+n} - a_n - a_m$ | n, m 6 N})。令 $(f_n)_{n\in\mathbb{N}}$ 由 f_n = c 对所有 n 定义。对于所有 m、n $\in \mathbb{N}$，我们有 $a_{n+m} \leqslant a_n + a_m + f_{n+m}$，因此根据引理 6，序列 $(a_n/n)_{n\in\mathbb{N}}$ 收敛，正如所需。

现在我们转向命题的第二种情形，即 Y 是本原的。与第一部分一样，由于 Y 的本原性，存在一个 N，使得对于任意词 u $\in \mathcal{B}$(Y)，都存在 u′ $\in \Sigma^N$，满足 u u′u $\in \mathcal{B}$(Y)。现在令 u $\subseteq \mathcal{B}_m$(Y) 为一个深洞，即满足 R($\mathcal{B}_m$(X), $\mathcal{B}_m$(Y)) = min{d(x, u) | x $\in \mathcal{B}_m$(X)} 的词。记 m′ = m + N。对于 n $\geqslant$ m′，我们考虑由下式定义的词 y $\in \mathcal{B}_n$(Y)

$$y = (\overbrace{u, u', u, u', \ldots, u, u'}^{\lfloor \frac{n}{m'} \rfloor \text{ times}}, y'),$$

其中y'是任意后缀，满足$y \in \mathcal{B}_n(Y)$。我们指出，可以找到y'，例如取u的一个前缀即可。对于任意$x \in \mathcal{B}_n(X)$，我们有

$$\begin{aligned} d(x, y) &\geqslant \sum_{i=0}^{\lfloor \frac{n}{m'} \rfloor - 1} d\left(x_{im'}^{(i+1)m'-1}, y_{im'}^{(i+1)m'-1}\right) = \sum_{i=0}^{\lfloor \frac{n}{m'} \rfloor - 1} d\left(x_{im'}^{(i+1)m'-1}, (u, u')\right) \\ &\geqslant \sum_{i=0}^{\lfloor \frac{n}{m'} \rfloor - 1} d\left(x_{im'}^{(i+1)m'-1-N}, u\right) \geqslant \left\lfloor \frac{n}{m'} \right\rfloor R(\mathcal{B}_m(X), \mathcal{B}_m(Y)) = \left\lfloor \frac{n}{m+N} \right\rfloor R(\mathcal{B}_m(X), \mathcal{B}_m(Y)). \end{aligned}$$

这证明了$R(\mathcal{B}_n(X), \mathcal{B}_n(Y)) \geqslant \lfloor \frac{n}{m+N} \rfloor R(\mathcal{B}_m(X), \mathcal{B}_m(Y))$。取极限可得

$$\liminf_{n\to\infty} \frac{R(\mathcal{B}_n(X), \mathcal{B}_n(Y))}{n} \geqslant \liminf_{n\to\infty} \frac{\lfloor \frac{n}{m+N} \rfloor R(\mathcal{B}_m(X), \mathcal{B}_m(Y))}{n} = \frac{R(\mathcal{B}_m(X), \mathcal{B}_m(Y))}{m+N}.$$

这对所有 $m \in \mathbb{N}$ 均成立，由此可证明

$$\liminf_{n\to\infty} \frac{R(\mathcal{B}_n(X), \mathcal{B}_n(Y))}{n} \geqslant \limsup_{m\to\infty} \frac{R(\mathcal{B}_m(X), \mathcal{B}_m(Y))}{m+N} = \limsup_{m\to\infty} \frac{R(\mathcal{B}_m(X), \mathcal{B}_m(Y))}{m},$$

且该极限存在。 ■

注 1 由命题 7 的证明可知，若 $Y = \Sigma^{\mathbb{Z}}$ 则

$$R(X,Y) = \lim_{n\to\infty} \frac{R(\mathcal{B}_n(X), \Sigma^n)}{n} = \sup_{n\in\mathbb{N}} \frac{R(\mathcal{B}_n(X), \Sigma^n)}{n}.$$

本文将始终以下述内容作为贯穿全文的示例。

例 1 考虑二元字母表 $[2] \triangleq \{0,1\}$。令 $X_{0,k} \subseteq [2]^{\mathbb{Z}}$ 为 $(0,k)$-受限游程长度系统，其由所有不包含 $k+1$ 个连续零的二元序列组成。即，$X_{0,k} = X_{\mathcal{F}}$，其中 $\mathcal{F} = \{\overline{0}_{k+1} = (0,\dots,0)\} \subseteq [2]^{k+1}$。我们断言

$$R(X_{0,k}, [2]^{\mathbb{Z}}) = \frac{1}{k+1}.$$

事实上，由注 1

$$R(X_{0,k}, [2]^{\mathbb{Z}}) \geqslant \frac{R(\mathcal{B}_{k+1}(X_{0,k}), [2]^{k+1})}{k+1} = \frac{R([2]^{k+1} \setminus \{\overline{0}_{k+1}\}, [2]^{k+1})}{k+1} = \frac{1}{k+1}.$$

现证明所得下界是紧的。令 $y \in [2]^n$ 为任意二元词。考虑由下式给出的 x

$$\overline{x}_i \triangleq \begin{cases} \overline{y}_i & i \bmod (k+1) \neq 0, \\ 1 & i \bmod (k+1) = 0. \end{cases}$$

显然，x 不包含任何由 $k+1$ 个连续零构成的子词，因此 $x \in \mathcal{B}_n(X_{0,k})$。由于 $d(x,y) \leqslant \lceil \frac{n}{k+1} \rceil$，我们得出 $\frac{1}{n} R(\mathcal{B}_n(X_{0,k}), [2]^n) \leqslant \frac{1}{n} \lceil \frac{n}{k+1} \rceil$，并令其取极限，

$$R(X_{0,k}, [2]^{\mathbb{Z}}) = \lim_{n\to\infty} \frac{R(\mathcal{B}_n(X_{0,k}), [2]^n)}{n} \leqslant \frac{1}{k+1}.$$

通过球覆盖论证，我们给出了一个关于系统容量的覆盖半径简单下界。回顾一下，$H_q\colon [0,1] \to [0,1]$ 表示如下定义的q元熵函数：

$$H_q(x) \triangleq x\log_q(q-1) - x\log_q(x) - (1-x)\log_q(1-x),$$

为保证连续性，$H_q(0) \triangleq 0$，以及$H_q(1) \triangleq \log_q(q-1)$。我们还用$H_q^{-1}\colon [0,1] \to [0, 1-\frac{1}{q}]$ 表示其反函数。

命题 8 设$X, Y \subseteq \Sigma^{\mathbb{Z}}$ 为容量分别为$h(X) \leqslant h(Y)$的受约束系统，并记作 $|\Sigma| = q$。则

$$R(X,Y) \geqslant H_q^{-1}(h(Y) - h(X)).$$

证明：若$R(X,Y) \geqslant 1 - \frac{1}{q}$，结论由$H_q^{-1}$的定义立即得出。因此，我们假设$R(X,Y) < 1 - \frac{1}{q}$。令$V_{r,n,q}$ 表示相对于汉明度量的半径为r 的球的大小（该大小与球心的选择无关），并记作 $\rho_n \triangleq \frac{1}{n} R(\mathcal{B}_n(X), \mathcal{B}_n(Y))$。根据并集界，对任意$n \in \mathbb{N}$,

$$|\mathcal{B}_n(Y)| = \left| \mathcal{B}_n(Y) \cap \left(\bigcup_{\overline{x}\in\mathcal{B}_n(X)} B_{n\rho_n}(\overline{x}) \right) \right| \leqslant |\mathcal{B}_n(X)| \cdot V_{\rho_n n,n,q}. \tag{3}$$

根据斯特林近似的标准应用（例如，参见 [19, 第 3] 章），众所周知，

$$V_{\rho n,n,q} = \sum_{i=0}^{\lfloor \rho n \rfloor} \binom{n}{i} (q-1)^i \leqslant \begin{cases} q^{nH_q(\rho)} & \rho \in [0, 1-\frac{1}{q}), \\ q^n & \rho \in [1-\frac{1}{q}, 1]. \end{cases}$$

由于定义容量 $h(X)$ 的极限存在，根据 (3) 以及 H_q 的连续性，我们得到

$$\begin{aligned} h(X) &= \lim_{n\to\infty} \frac{\log_q |\mathcal{B}_n(X)|}{n} = \liminf_{n\to\infty} \frac{\log_q |\mathcal{B}_n(X)|}{n} \\ &\geqslant \liminf_{n\to\infty} \left(\frac{\log_q |\mathcal{B}_n(Y)|}{n} - H_q(\rho_n) \right) = h(Y) - H_q(R(X,Y)). \end{aligned}$$

通过重新排列并使用H_q^{-1}，我们得出结论。 ■

例2 固定 $\Sigma = [q] \triangleq \{0, \dots, q-1\}$，并考虑重复移位

$$X_{\text{rep}} \triangleq \{(\dots, a, a, a, \dots) \mid a \in [q]\}.$$

显然，X^{rep} 是由禁用模式集合定义的有限型移位

$$\mathcal{F} = \{ab \mid a, b \in [q], a \neq b\} \subseteq [q]^2.$$

由于 $h(X^{\text{rep}}) = 0$，根据命题 8

$$R(X_{\text{rep}}, [q]^{\mathbb{Z}}) \geqslant H_q^{-1}(1-0) = 1 - \frac{1}{q}.$$

另一方面，对于任意 $n \in \mathbb{N}$ 以及任意 $y \in [q]^n$，显然至少存在一个符号 $a \in [q]$，它在 y 的至少 $\lceil \frac{n}{q} \rceil$ 个坐标中出现；特别地，$d(y, (a, \dots, a)) \leqslant \lfloor \overline{q-1}\, \overline{q}\, n \rfloor$。这证明了

$$R(\mathcal{B}_n(X_{\text{rep}}), [q]^n) \leqslant \left\lfloor \frac{q-1}{q} n \right\rfloor.$$

取极限并结合下界

$$R(X_{\text{rep}}, [q]^{\mathbb{Z}}) = 1 - \frac{1}{q}.$$

注2和例2表明，两个受限系统的并集的覆盖半径可以严格小于其覆盖半径的最小值。我们注意到，$X^{\text{rep}} = \bigcup_{a \in [q]} \{\mathbf{x}_a\}$，其中 $\mathbf{x}_a$ 是符号 a 对应的常值双无限序列。我们注意到，$\{\mathbf{x}_a\}$ 是一个受限系统，并且 $R(\{\mathbf{x}_a\}, [q]^{\mathbb{Z}}) = 1$。因此

$$\min_{a \in [q]} R(\{\mathbf{x}_a\}, [q]^{\mathbb{Z}}) = 1 > 1 - \frac{1}{q} = R(X_{\text{rep}}, [q]^{\mathbb{Z}}).$$

这与受约束系统的容量形成对比：有限个受约束系统的并集，其容量等于各系统容量的最大值。

此时，我们遇到了一个颇为奇特的情形。为说明这一点，固定二元字母表 $\Sigma = [2]$。若考虑例 1 中的 $\mathbf{X}_{0,1}$，$(0,1)$-受限游程长度系统，则其容量已知为 $\mathbf{h}(\mathbf{X}_{0,1}) = \log_2((1+\sqrt{5})/2) \approx 0.694$，并且我们已证明，其覆盖半径（相对于 $[2]^{\mathbb{Z}}$）为 $\mathbf{R}(\mathbf{X}_{0,1}, [2]^{\mathbb{Z}}) = \frac{1}{2}$。然而，在例2中我们看到，二元重复移位 $\mathbf{X}^{\text{rep}}$ 的覆盖半径相同，为 $\mathbf{R}(\mathbf{X}^{\text{rep}}, [2]^{\mathbb{Z}}) = \frac{1}{2}$，但其容量为零，即 $\mathbf{h}(\mathbf{X}^{\text{rep}}) = 0$。从编码视角来看，尽管 $\mathcal{B}_n(\mathbf{X}_{0,1})$ 的词数呈指数级多于 $\mathcal{B}_n(\mathbf{X}^{\text{rep}})$，但最坏情况下的覆盖情形，也就是深洞，与受约束码之间的距离在渐近意义下仍然相同。

除了数学上的趣味性之外，具有 $R(X_{0,k}, [2]^{\mathbb{Z}}) = R(X^{\text{rep}}, [2]^{\mathbb{Z}}) = \frac{1}{2}$（在这两个示例中）会妨碍纠正第 I 节所述 QCC 方案中的信道错误。这是因为纠错码需要纠正超过码长12 的错误位置，而以非零码率做到这一点是不可能的。因此，我们希望寻找一种不同的受限系统覆盖半径定义，将深洞出现的罕见性考虑在内。

作为本节的最后一点评论，我们想讨论 $\mathbf{R}(\mathbf{X}, \Sigma^{\mathbb{Z}})$ 与 QCC 框架的关系。由于我们关注渐近情形，假设 QCC 方案中的纠错码序列为 $(\mathbf{C}_n)_{n \in \mathbb{N}}$，其中 $\mathbf{C}_n$ 的码长为 $\mathbf{n}$。表达式 $\mathbf{R}(\mathbf{X}, \Sigma^{\mathbb{Z}}) = \lim_{n \to \infty} \frac{1}{\mathbf{n}} \mathbf{R}(\mathcal{B}_n(\mathbf{X}), \Sigma^n)$ 是采用码序列 $(\mathbf{C}_n)_{n \in \mathbb{N}}$ 时最坏情况量化误差率的一个上界；该误差率实际上为 $\lim_{n \to \infty} \frac{1}{\mathbf{n}} \mathbf{R}(\mathcal{B}_n(\mathbf{X}), \mathbf{C}_n)$。界 $\mathbf{R}(\mathbf{X}, \Sigma^{\mathbb{Z}})$ 有两重保守性：一是允许深洞决定覆盖半径，二是假设深洞位于 $\mathbf{C}_n$ 中。由于 $\lim_{n \to \infty} \frac{1}{\mathbf{n}} \mathbf{R}(\mathcal{B}_n(\mathbf{X}), \mathbf{C}_n)$ 可能难以计算，且取决于纠错码序列，我们可以用 $\mathbf{R}(\mathbf{X}, \Sigma^{\mathbb{Z}})$ 作为最坏情况量化误差的上界，而它独立于码序列。

IV. 本质覆盖半径

上一节研究的覆盖半径可能过于悲观，因为它是由最坏情况下的量化距离决定的。本节将研究覆盖半径的另一种定义，称为本质覆盖半径。给定 $\varepsilon > 0$，约束系统的 ϵ-覆盖半径，粗略地说，是使空间中r 的$(1-\varepsilon)$比例的词能够被量化到该约束系统中的最小值。在下文中，我们将进一步把覆盖半径推广为一种概率定义。

我们首先陈述遍历理论中的一些基本定义和众所周知的结果。对于任意有限字母表 Σ，我们将 $\Sigma^{\mathbb{Z}}$ 视为一个可测空间，并赋予由 $\Sigma^{\mathbb{Z}}$ 上的积拓扑诱导的 Borel Σ-代数。同样，任意子移位 $Y \subseteq \Sigma^{\mathbb{Z}}$ 都被视为一个可测空间。

定义 9（不变测度与遍历测度）设 $\mathrm{Y} \subseteq \Sigma^{\mathbb{Z}}$ 为一个子移位。Y 上的概率测度 μ 称为移位不变测度，如果对于任意可测集 B，都有 $\mu(\mathrm{T}^{-1}\mathrm{B}) = \mu(\mathrm{B})$。若 $\mathrm{T}^{-1}\mathrm{B} = \mathrm{B}$ 蕴含 $\mu(\mathrm{B}) = 0$ 或 $\mu(\mathrm{Y} \setminus \mathrm{B}) = 0$，则称移位不变测度 μ 为遍历测度。Y 上移位不变概率测度的集合记为 M(Y)，M(Y) 中遍历测度的集合记为 $\mathrm{M}_{\mathcal{E}}(\mathrm{Y})$。

对于测度 $\mu \in M(Y)$，我们用 μ_n 表示 μ 在前n 个坐标上的边缘测度，它是定义在 Σ^n上的概率测度。为避免繁琐的记号，在本文中我们始终用 $\mathbb{P}_\mu[A]$ 表示测度 $\mu(A)$，并用$\mathbf{Y}$ 表示取值于Y上的随机双无限序列。本文中，我们用粗体大写字母表示随机变量组成的双无限序列，以区别于用于表示受限系统的非粗体大写字母。我们在分析中使用的一个重要结果是以下著名的遍历定理，它是遍历理论中的经典结果。遍历定理中的L_2收敛归功于冯·诺伊曼，而几乎必然成立的结论归功于比尔霍夫，两者的研究均发表于1931年。关于这一著名经典定理的证明，可见于大多数标准的遍历理论入门教材，例如 [32, 第 3]章。

定理 10（遍历定理，**[6], [28]**）设$Y \subseteq \Sigma^{\mathbb{Z}}$ 是一个移位空间，$\mu \in M_{\mathcal{E}}(Y)$ 是一个遍历测度，且$f \in L_1(\mu)$ 是一个可积函数。那么，由下式定义的序列$(A_n)_{n\in\mathbb{N}}$

$$A_n = \frac{1}{n}\sum_{i=0}^{n-1} f \circ T^{-i}$$

几乎必然收敛，并且在L_2, 意义下收敛到 $\int f \cdot d\mu$。特别地，对于所有 $\delta > 0$

$$\lim_{n\to\infty} \mathbb{P}_\mu\left[\left|A_n - \int f \cdot d\mu\right| > \delta\right] = 0.$$

现在我们可以定义本质覆盖半径了。

定义 11 对于任意实数 $\varepsilon > 0$、两个集合 A、$C \subseteq \Sigma^n$，以及 η 上的一个概率测度，我们定义

$$R_\varepsilon(C, A, \eta) \triangleq \min\left\{ r \in \mathbb{N} \cup \{0\} \;\middle|\; \eta\left(A \cap \left(\bigcup_{\overline{x}\in C} B_r(\overline{x})\right)\right) \geqslant 1 - \varepsilon \right\}.$$

我们指出，当 η 是 A 上的均匀测度时，R_ε(C、A、η) 是 A 的 ε-覆盖半径，也就是说，使得C 中至少有 $1-\epsilon$) 比例的词与A的距离不超过 r 的最小 r，正如所需。

定义 12 设X、$Y \subseteq \Sigma^{\mathbb{Z}}$为受限系统，且 $\mu \in M_{\mathcal{E}}(Y)$ 为一个遍历测度。我们将X 相对于(Y、μ) 的 ε-覆盖半径定义为

$$R_\varepsilon(X, Y, \mu) \triangleq \liminf_{n\to\infty} \frac{R(\mathcal{B}_n(X), \mathcal{B}_n(Y), \mu_n)}{n},$$

并将X 相对于(Y、μ) 的本质覆盖半径定义为

$$R_0(X, Y, \mu) \triangleq \lim_{\varepsilon\to 0} R_\varepsilon(X, Y, \mu).$$

我们指出，由于R_ε(X,Y, μ) 关于 ε 的单调性，前一定义中的极限存在。我们还注意到，显然，本质覆盖半径不超过（最坏情况）覆盖半径，因此有

$$R_\varepsilon(X, Y, \mu) \leqslant R_0(X, Y, \mu) \leqslant R(X, Y). \tag{4}$$

现在我们回顾上一节讨论的重复系统（例2）和(0,k)-受限游程长度系统（例 1）的例子。

命题 13 设例2中的q元重复系统$X_{\mathrm{rep}} \subseteq [q]^{\mathbb{Z}}$，并假设$Y = [q]^{\mathbb{Z}}$ 配备均匀伯努利独立同分布测度，记为 μ_{u}。那么，本质覆盖半径等于覆盖半径，即

$$R_0(X_{\mathrm{rep}}, [q]^{\mathbb{Z}}, \mu^{\mathrm{u}}) = R(X_{\mathrm{rep}}, [q]^{\mathbb{Z}}) = 1 - \frac{1}{q}.$$

证明：对于固定的 $n \in \mathbb{N}$ 和 $i \in [q]$, , 令 $Y_{(i)n}$ 表示在随机（均匀分布的）$[q]^n$ 词中，符号 i 出现次数的随机变量。根据大数定律，对于任意 $\delta > 0$

$$\lim_{n\to\infty} \mathbb{P}_{\mu_n^{\mathrm{u}}}\left[\bigcap_{i=0}^{q-1}\left\{\left|\frac{Y_n^{(i)}}{n} - \frac{1}{q}\right| < \delta\right\}\right] = 1. \tag{5}$$

我们还注意到，对于一个词 $\overline{y} \in [q]^n$，若每个符号 $i \in [q]$ 至少出现于 $(\frac{1}{q}-\delta)n$ 个坐标中，则有

$$\min_{\overline{x}\in\mathcal{B}_n(X_{\mathrm{rep}})} d(\overline{x},\overline{y}) \geqslant (q-1)\left(\frac{1}{q}-\delta\right)n. \tag{6}$$

对于任意 $\varepsilon > 0$，由 (5) 可知，当n充分大时，$[q]^n$中任何概率至少为 $1-\varepsilon$ 的集合都包含一个词，使得任意$i \in [q]$ 至少出现在$(\frac{1}{q}-\delta)n$ 个坐标中。结合 (6)，我们得出，当n充分大时

$$\frac{R_\varepsilon(\mathcal{B}_n(X_{\mathrm{rep}}),[q]^n,\mu_n^{\mathsf{u}})}{n} \geqslant (q-1)\left(\frac{1}{q}-\delta\right).$$

于是可知 $R_\varepsilon(X^{\mathrm{rep}},[q]^{\mathbb{Z}},\mu_{\mathsf{u}}) \geqslant (q-1)(\frac{1}{q}-\delta)$，这对所有 $\delta > 0$ 都成立，因此

$$R_0(X_{\mathrm{rep}},[q]^{\mathbb{Z}},\mu^{\mathsf{u}}) = \lim_{\varepsilon\to 0} R_\varepsilon(X_{\mathrm{rep}},[q]^{\mathbb{Z}},\mu^{\mathsf{u}}) \geqslant 1-\frac{1}{q}.$$

上界可直接由例2得出，因此我们有

$$R_0(X_{\mathrm{rep}},[q]^{\mathbb{Z}},\mu^{\mathsf{u}}) = R(X_{\mathrm{rep}},[q]^{\mathbb{Z}}) = 1-\frac{1}{q}.$$

■

注 3 由于 (4) 对所有遍历测度都成立，我们可以如下给出覆盖半径的下界，

$$R(X,Y) \geqslant \sup_{\mu\in M_{\mathcal{E}}(Y)} R_0(X,Y,\mu). \tag{7}$$

命题 13 表明，当$X = X^{\mathrm{rep}}$ 是重复系统且$Y = [q]^{\mathbb{Z}}$时，(7) 是紧的，因为Y上的均匀伯努利测度取等号。

我们声称，对于 $X = X_{0,k}$，即例 1 中描述的二进制 (0,k)-受限游程长度系统，式 (7) 也是紧的。为此，考虑 $\mu = \delta_{\mathbf{0}}$，即全零序列 $\mathbf{0} \in Y = [2]^{\mathbb{Z}}$的狄拉克测度。由于 $\mathbf{0}$ 是周期为 1 的周期点（相对于移位变换），$\delta_{\mathbf{0}}$ 确实是不变测度，而且还是遍历的，因为它对任意集合赋予的测度都是 0 或 1。简单计算可得，对于任意 $0 < \varepsilon < 1$ 和 $n \in \mathbb{N}$,

$$R_\varepsilon(\mathcal{B}_n(X_{0,k}),[2]^n,\mu_n) = \min_{\overline{x}\in\mathcal{B}_n(X_{0,k})} d(\overline{x},\overline{0}_n) = \left\lfloor \frac{n}{k+1} \right\rfloor,$$

这意味着

$$R_0(X_{0,k},[2]^{\mathbb{Z}},\mu) = \lim_{\varepsilon\to 0} R_\varepsilon(X_{0,k},[2]^{\mathbb{Z}},\mu) = \lim_{\varepsilon\to 0}\liminf_{n\to\infty} \frac{R_\varepsilon(\mathcal{B}_n(X_{0,k}),[2]^n,\mu_n)}{n} = \frac{1}{k+1} = R(X,Y).$$

我们猜想，界 (7) 通常是紧的，并将其留作未来研究的一个方向。

如我们所见，容量为零的重复系统具有相同的覆盖半径和本质覆盖半径。$(0,k)$-受限游程长度系统的容量为正。虽然其覆盖半径为 1_{k+1}，但以下定理表明，其本质覆盖半径随 k 指数级快速衰减，与重复系统形成鲜明对比。

定理 14 设 $X_{0,k} \subseteq [2]^{\mathbb{Z}}$ 为例 1 中的 (0,k)-受限游程长度系统，并设 $Y = [2]^{\mathbb{Z}}$ 配备均匀伯努利独立同分布测度 μ_{u}。则

$$R_0(X_{0,k},[2]^{\mathbb{Z}},\mu^{\mathsf{u}}) = \frac{1}{2(2^{k+1}-1)}.$$

证明：对于一个词 $\overline{y} \in [2]^n$ 和整数 $i \geqslant k+1$，记 $S_i(y)$ 为模式 $1\overline{0}_i1 \in [2]^{i+2}$在 y中出现的次数。我们还记 $M(y)$ 为坐标 $j \in [n]$ 中不属于 $\overline{y}$中任意形如 $1\overline{0}_i1$的模式的坐标数量，其中 $i \geqslant k+1$。关键观察如下不等式断言：对于所有 $\ell \geqslant k+1$，我们有

$$\sum_{i=k+1}^{\ell}\left\lfloor \frac{i}{k+1}\right\rfloor S_i(\overline{y}) \leqslant \min_{\overline{x}\in\mathcal{B}_n(X_{0,k})} d(\overline{x},\overline{y}) \leqslant M(\overline{y}) + \sum_{i=k+1}^{\infty}\left\lfloor \frac{i}{k+1}\right\rfloor S_i(\overline{y}). \tag{8}$$

事实上，我们注意到，对于任意的 $x \in \mathcal{B}_n(X_{0,k})$ 以及模式 $y_m^{m+i+1} = 1\overline{0}_i1$ 的任意实例，x_{m+1}^{m+i} 与y_{m+1}^{m+i} 至少有$\lfloor \frac{i}{k+1} \rfloor$处不同，因为$x$ 不含有 $k+1$ 个连续零。我们还注意到，形如 $1\overline{0}_i1$ 的模式在y 中不会在零上发生重叠。因此我们得到，

$$d(\overline{x},\overline{y}) \geqslant \sum_{i=k+1}^{\infty}\left\lfloor \frac{i}{k+1}\right\rfloor S_i(\overline{y}) \geqslant \sum_{i=k+1}^{\ell}\left\lfloor \frac{i}{k+1}\right\rfloor S_i(\overline{y}),$$

下界可通过对所有 $x \in \mathcal{B}_n(X_{0,k})$ 取最小值而得。另一方面，为了证明上界，只需构造一个满足以下条件的词 $x \in \mathcal{B}_n(X_{0,k})$

$$d(\overline{x},\overline{y}) \leqslant M(\overline{y}) + \sum_{i=k+1}^{\infty} \left\lfloor \frac{i}{k+1} \right\rfloor S_i(\overline{y}). \tag{9}$$

我们构造如下：对于任意 x，$i \geqslant k+1$，以及 y_m^{m+i+1} 中模式 $1\overline{0}_i1$ 的任意一个实例，我们令

$$\overline{x}_m^{m+i+1} = 1\overbrace{\overline{0}_k\overline{0}_k1\cdots\overline{0}_k1}^{\lfloor \frac{i}{k+1} \rfloor (k+1)}\ \overbrace{0\cdots0}^{i \bmod (k+1)}\ 1.$$

在其余坐标上，我们定义 x 与 y 相同，但将由 $M(y)$ 计数的坐标设为 1。根据 x 的构造，其中包含的最长零串至多为 k，这意味着 $x \in \mathcal{B}_n(X_{0,k})$。因此，$x$ 如预期满足 (9)。

对于任意 $\ell \geqslant k+1$（包括 $\ell = \infty$），我们定义 $f_\ell : [2]^{\mathbb{Z}} \to \mathbb{R}$ 为

$$f_\ell(\mathbf{y}) = \begin{cases} \left\lfloor \frac{i}{k+1} \right\rfloor & \mathbf{y}_0^{i+1} = 1\overline{0}_i1 \text{ for } k+1 \leqslant i \leqslant \ell, \\ 0 & \text{otherwise,} \end{cases}$$

并记

$$E_\ell \triangleq \int f_\ell \cdot d\mu^{\mathsf{u}}.$$

简单计算可知，对于任意有限的 $\ell \in \mathbb{N}$

$$E_\ell = \sum_{i=k+1}^{\ell} \left\lfloor \frac{i}{k+1} \right\rfloor \frac{1}{2^{i+2}},$$

根据单调收敛定理，$E_\ell \to E_\infty$ 当 $\ell \to \infty$ 时成立，因为 f_ℓ（单调地）逐点收敛到 f_∞。

设 $\overline{Y}_n$ 是一个长度为 n 的随机词，其分布服从 $\mu_{\mathsf{u}n}$；设 $\mathbf{Y}$ 是一个分布服从 μ_{u} 的随机序列。注意，对于任意 $\ell \leqslant n-2$，

$$\sum_{j=0}^{n-\ell-2} f_\ell \circ T^j(\mathbf{Y}) = \sum_{i=0}^{\ell} \left\lfloor \frac{i}{k+1} \right\rfloor S_i(\mathbf{Y}_0^{n-1}).$$

结合式 (8)，对于固定的 ℓ 和 $\delta > 0$，且对所有充分大的 n，

$$\begin{aligned} \mathbb{P}_{\mu_n^{\mathsf{u}}}\left[\min_{\overline{x}_n \in \mathcal{B}_n(X_{0,k})} \frac{d(\overline{x}_n, \overline{Y}_n)}{n} > E_\ell - \delta \right] &= \mathbb{P}_{\mu^{\mathsf{u}}}\left[\min_{\overline{x}_n \in \mathcal{B}_n(X_{0,k})} \frac{d(\overline{x}_n, \mathbf{Y}_0^{n-1})}{n} > E_\ell - \delta \right] \\ &\geqslant \mathbb{P}_{\mu^{\mathsf{u}}}\left[\frac{1}{n} \sum_{i=k+1}^{\ell} \left\lfloor \frac{i}{k+1} \right\rfloor S_i(\mathbf{Y}_0^{n-1}) > E_\ell - \delta \right] \\ &= \mathbb{P}_{\mu^{\mathsf{u}}}\left[\frac{1}{n} \sum_{j=0}^{n-\ell-2} f_\ell \circ T^j(\mathbf{Y}) > E_\ell - \delta \right] \xrightarrow[n\to\infty]{} 1, \end{aligned}$$

其中收敛到 1 是由遍历定理导致的（见定理 10）。这证明，对于任意 $\varepsilon \in (0,1)$、$\delta > 0$ 以及 $\ell \geqslant k+1$，当 n 充分大时，

$$\frac{R_\varepsilon(\mathcal{B}_n(X_{0,k}), \mathcal{B}_n(Y), \mu^{\mathsf{u}})}{n} > E_\ell - \delta,$$

因此，

$$R_\varepsilon(X_{0,k}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) \geqslant \lim_{\substack{\ell\to\infty \\ \delta\to 0}} E_\ell - \delta = E_\infty.$$

对于上界，我们采用类似的技巧。首先注意，随机变量序列 $(\frac{1}{n}M(\mathbf{Y}_0^{n-1}))_{n\in\mathbb{N}}$ 依概率收敛到 0。注意，对于任意 $N \in \mathbb{N}$ 以及长度为 n 的词 y，若 $M(y) > 2N$，则 $y_0^{N-1} = \overline{0}$ 或 $y_{n-N}^{n-1} = \overline{0}$。因此，对于任意 $\delta > 0$

$$\mathbb{P}_{\mu^{\mathsf{u}}}\left[\left| \frac{1}{n} M(\mathbf{Y}_0^{n-1}) - 0 \right| > 2\delta \right] \leqslant \mathbb{P}_{\mu^{\mathsf{u}}}\left[\mathbf{Y}_0^{\lfloor \delta n \rfloor} = \overline{0} \right] + \mathbb{P}_{\mu^{\mathsf{u}}}\left[\mathbf{Y}_{n-\lfloor \delta n \rfloor}^{n-1} = \overline{0} \right] = \frac{2}{2^{\lfloor \delta n \rfloor}} \xrightarrow[n\to\infty]{} 0.$$

我们同样注意到，对于任意 n，

$$\sum_{j=0}^{n-1} f_\infty \circ T^j(\mathbf{Y}) \geqslant \sum_{i=0}^{\infty} \left\lfloor \frac{i}{k+1} \right\rfloor S_i(\mathbf{Y}_0^{n-1}).$$

结合式 (8)，对于所有 $\delta > 0$

$$\mathbb{P}_{\mu_n^{\mathsf{u}}}\left[\min_{\overline{x}_n \in \mathcal{B}_n(X_{0,k})} \frac{d(\overline{x}_n, \overline{Y}_n)}{n} < E_\infty + \delta\right] \geqslant \mathbb{P}_{\mu^{\mathsf{u}}}\left[\frac{1}{n}\left(M(\mathbf{Y}_0^{n-1}) + \sum_{i=k+1}^{\infty}\left\lfloor\frac{i}{k+1}\right\rfloor S_i(\mathbf{Y}_0^{n-1})\right) < E_\infty + \delta\right]$$
$$\geqslant \mathbb{P}_{\mu^{\mathsf{u}}}\left[\frac{1}{n}\left(M(\mathbf{Y}_0^{n-1}) + \sum_{j=0}^{n-1} f_\infty \circ T^j(\mathbf{Y})\right) < E_\infty + \delta\right] \xrightarrow[n\to\infty]{} 1,$$

这里，趋于 1 同样是由遍历定理以及 $\frac{1}{n}M(\mathbf{Y}_0^{n-1})$ 依概率收敛于 0 得出的。与之前一样，这证明了对于所有 $\varepsilon > 0$，都有 $R_\varepsilon(X_{0,k}, [2]^{\mathbb{Z}}, \mu_{\mathsf{u}}) \geqslant E_\infty$；根据下界，$R_\varepsilon(X_{0,k}, [2]^{\mathbb{Z}}, \mu_{\mathsf{u}}) = E_\infty$。特别地，

$$R_0(X_{0,k}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) = \lim_{\varepsilon\to 0} R_\varepsilon(X_{0,k}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) = E_\infty.$$

为了完成证明，只需计算 E_∞:

$$E_\infty = \sum_{i=k+1}^{\infty}\left\lfloor\frac{i}{k+1}\right\rfloor\frac{1}{2^{i+2}} = \sum_{j=1}^{\infty} j \sum_{i=j(k+1)}^{(j+1)(k+1)-1}\frac{1}{2^{i+2}} = \frac{2^{k+1}-1}{2^{k+2}}\sum_{j=1}^{\infty}\frac{j}{2^{j(k+1)}}$$
$$= \frac{2^{k+1}-1}{2^{k+2}}\cdot\frac{2^{k+1}}{(2^{k+1}-1)^2} = \frac{1}{2(2^{k+1}-1)}.$$

■

我们希望能得到本质覆盖半径的其他表达式，以便计算或估计其值。受定理 14 证明中所用工具的启发，我们给出一种等价的遍历论刻画。

定义 15 设 X、$Y \subseteq \Sigma^{\mathbb{Z}}$ 为移位空间。我们将 $X \times Y$ 视为移位空间，其中左移位的作用为 $T(\mathbf{x}, \mathbf{y}) = (T\mathbf{x}, T\mathbf{y})$。对于遍历测度 $\mu \in M_{\mathcal{E}}(Y)$，$\mu$ 在 $X \times Y$ 上的一个扩张，是乘积空间 $X \times Y$ 上的移位不变测度 ν，其 Y 边缘分布为 μ。也就是说，ν 满足：对于任意可测集 $A \subseteq Y$

$$\nu(X \times A) = \mu(A).$$

如果乘积空间上的移位变换下的测度是遍历测度，则称$X \times Y$ 上的扩张是遍历的。我们用$M(X, Y, \mu)$ 表示 μ的所有扩张构成的集合，并用$M_{\mathcal{E}}(X, Y, \mu)$ 表示$M(X,Y, \mu)$中的所有遍历扩张构成的集合。

在下面的命题中，我们给出本质覆盖半径的一个上界，该上界对X 和Y不需要任何额外假设。

命题 16 设X、$Y \subseteq \Sigma^{\mathbb{Z}}$ 为移位空间，且 $\mu \in M_{\mathcal{E}}(Y)$。则

$$R_0(X, Y, \mu) \leqslant \inf\{\mathbb{P}_\nu[\mathbf{X}_0 \neq \mathbf{Y}_0] \mid \nu \in M_{\mathcal{E}}(X, Y, \mu)\},$$

其中$\mathbf{X}_0$,$\mathbf{Y}_0$ 是由随机序列$\mathbf{X}$ 和$\mathbf{Y}$ 在第 0 个坐标上的投影所定义的随机变量。

证明：设 ν 是$M_{\mathcal{E}}(X, Y, \mu)$中的一个扩张，并设 $0 < \varepsilon < 1$ 和 $\delta > 0$ 为任意小的数。只需证明

$$R_\varepsilon(X, \Sigma^{\mathbb{Z}}, \mu) \leqslant \mathbb{P}_\nu[\mathbf{X}_0 \neq \mathbf{Y}_0] + \delta.$$

我们考虑函数 $\mathbf{n}f\colon X \times Y \to \{0,1\}$，定义为事件 $\{\mathbf{X}_0 \neq \mathbf{Y}_0\}$ 的示性函数。显然

$$\int f \cdot d\nu = \mathbb{P}_\nu[\mathbf{X}_0 \neq \mathbf{Y}_0].$$

根据定理 10，当n充分大时，

$$\mathbb{P}_\nu\left[\left|\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^n - \int f \cdot d\nu\right| > \delta\right] < \varepsilon. \tag{10}$$

我们还注意到

$$\sum_{i=0}^{n-1} f \circ T^n = d\left(\mathbf{X}_0^{n-1}, \mathbf{Y}_0^{n-1}\right), \tag{11}$$

其中d 表示汉明距离。记

$$s_\delta \triangleq \mathbb{P}_\nu[\mathbf{X}_0 \neq \mathbf{Y}_0] + \delta \quad \text{and} \quad B_{s_\delta}(\mathcal{B}_n(X)) \triangleq \mathcal{B}_n(Y) \cap \left(\bigcup_{\overline{x}\in\mathcal{B}_n(X)} B_{s_\delta}(\overline{x})\right).$$

将式（10）和式（11）与全概率公式结合，可得

$$\begin{aligned}
1-\varepsilon &< \mathbb{P}_\nu\left[d\left(\mathbf{X}_0^{n-1},\mathbf{Y}_0^{n-1}\right)\leqslant n\cdot s_\delta\right]\\
&=\sum_{\overline{y}\in\mathcal{B}_n(Y)}\mathbb{P}_\nu\left[d\left(\mathbf{X}_0^{n-1},\mathbf{Y}_0^{n-1}\right)\leqslant n\cdot s_\delta\;\middle|\;\mathbf{Y}_0^{n-1}=\overline{y}\right]\mathbb{P}_\nu\left[\mathbf{Y}_0^{n-1}=\overline{y}\right]\\
&=\sum_{\overline{y}\in B_{s_\delta}(\mathcal{B}_n(X))}\mathbb{P}_\nu\left[d\left(\mathbf{X}_0^{n-1},\mathbf{Y}_0^{n-1}\right)\leqslant n\cdot s_\delta\;\middle|\;\mathbf{Y}_0^{n-1}=\overline{y}\right]\mathbb{P}_\nu\left[\mathbf{Y}_0^{n-1}=\overline{y}\right]\\
&\leqslant\sum_{\overline{y}\in B_{s_\delta}(\mathcal{B}_n(X))}\mathbb{P}_\nu\left[\mathbf{Y}_0^{n-1}=\overline{y}\right]=\sum_{\overline{y}\in B_{s_\delta}(\mathcal{B}_n(X))}\mathbb{P}_\mu\left[\mathbf{Y}_0^{n-1}=\overline{y}\right]\\
&=\mathbb{P}_{\mu_n}\left[\mathcal{B}_n(Y)\cap\left(\bigcup_{\overline{x}\in\mathcal{B}_n(X)}B_{s_\delta}(\overline{x})\right)\right].
\end{aligned}$$

这表明

$$R_\varepsilon(X,Y,\mu)\leqslant s_\delta=\mathbb{P}_\nu[\mathbf{X}_0\neq\mathbf{Y}_0]+\delta,$$

因此完成了证明。 ■

下面的命题表明，命题 16 上界中的、在遍历扩张集合$M_{\mathcal{E}}(X,Y,\mu)$上的最小化问题，实际上等价于在所有不变扩张集合$M(X,Y,\mu)$上的最小化问题。为证明这一点，我们需要遍历分解定理。设Z 是一个配备 Borel σ-代数的紧致度量空间，且$T:Z\to Z$ 是连续函数。我们将Z上的T-不变测度空间$M(Z)$视为一个可测空间，其 Σ-代数由弱-$*$ 拓扑诱导。

引理 17（遍历分解， **[13,** 定理 **4.8]**）设Z 是一个紧致度量空间，$T:X\to X$ 是连续映射，且 $\mu\in M(Z)$ 是不变测度。那么存在唯一的概率测度P_μ ，定义在$M(Z)$上，并且支撑于遍历测度集合$M_{\mathcal{E}}(Z)$，使得对于任意可测集$E\subseteq Z$，

$$\mu(E)=\int_{M_{\mathcal{E}}(Z)}\nu(E)dP_\mu(\nu).$$

命题 18 设X、$Y\subseteq\Sigma^{\mathbb{Z}}$，并设 $\mu\in M_{\mathcal{E}}(Y)$ 为移位不变遍历测度。则

$$\inf\{\mathbb{P}_\nu[\mathbf{X}_0\neq\mathbf{Y}_0]\mid\nu\in M_{\mathcal{E}}(X,Y,\mu)\}=\inf\{\mathbb{P}_\nu[\mathbf{X}_0\neq\mathbf{Y}_0]\mid\nu\in M(X,Y,\mu)\}.$$

证明：记 $m\triangleq\inf\{\mathbb{P}_\nu[\mathbf{X}_0\neq\mathbf{Y}_0]\mid\nu\in M_{\mathcal{E}}(X、Y、\mu)\}$。不等式

$$m\geqslant\inf\{\mathbb{P}_\nu[\mathbf{X}_0\neq\mathbf{Y}_0]\mid\nu\in M(X,Y,\mu)\}$$

显然，因为$M_{\mathcal{E}}(X、Y、\mu)\subseteq M(X、Y、\mu)$。对于另一方向，我们需要证明，对所有 $\eta\in M(X、Y、\mu)$

设P_η为 η的遍历分解。也说，$\eta=\int_{M\mathcal{E}(X\times Y)}\nu\cdot dP_\eta(\nu)$，且遍历测度的集合具有满$dP_\eta$测度。我们先证明$P_\eta$的支撑包含于$M_{\mathcal{E}}(X、Y、\mu)$。反设

$$\mathbb{P}_\eta[\mathbf{X}_0\neq\mathbf{Y}_0]\geqslant m.$$

$$P_\eta[\{\nu\in M_{\mathcal{E}}(X\times Y)\mid\nu_Y\neq\mu\}]>0.$$

由于Y 上的 Borel Σ代数是可数生成的，存在可测集$E\subseteq Y$ 和$n\in\mathbb{N}$，使得由下式定义的集合$A_\pm$ 中至少有一个

$$A_\pm\triangleq\left\{\nu\in M_{\mathcal{E}}(X\times Y)\;\middle|\;\pm\nu_Y(E)\geqslant\pm\mu(E)+\frac{1}{n}\right\},$$

具有正的 P_η-测度。不失一般性，假设 $P_\eta(A_+)>0$。我们定义一个新的测度 η' 如下：

$$\eta'(A)=\frac{1}{P_\eta(A_+)}\int_{A_+}\nu(A)\cdot dP_\eta(\nu).$$

我们注意到， η' 是一个移位不变测度（因为它是对移位不变测度积分所得）。我们还注意到，$\eta'_Y\ll\eta_Y$ （也就是说，η'_Y 相对于 η_Y绝对连续），因为关于 η 的任何零测集显然也是关于η'的零测集。众所周知，任何相对于一个遍历不变概率测度绝对连续的不变概率测度

都必定等于该测度（例如，参见 [33, 注 1，第153]页）。我们回顾一下， $\eta_Y = \mu$ 是遍历的，因此$\eta_Y = \eta'_Y$。这与定义矛盾，因为根据A_+ 和 η' ，我们有

$$\mu(E) = \eta_Y(E) = \eta'_Y(E) = \eta'(X \times E) = \frac{1}{P_\eta(A_+)} \int_{A_+} \nu(X \times E) \cdot dP_\eta(\nu)$$
$$\geqslant \frac{1}{P_\eta(A_+)} \int_{A_+} \left(\mu(E) + \frac{1}{n}\right) dP_\eta(\nu) = \mu(E) + \frac{1}{n}.$$

现在由此结论可得，因为根据m 的定义，我们有

$$\mathbb{P}_\eta[\mathbf{X}_0 \neq \mathbf{Y}_0] = \int_{M_\mathcal{E}(X\times Y)} \mathbb{P}_\nu[\mathbf{X}_0 \neq \mathbf{Y}_0] dP_\eta(\nu)$$
$$= \int_{M_\mathcal{E}(X,Y,\mu)} \mathbb{P}_\nu[\mathbf{X}_0 \neq \mathbf{Y}_0] dP_\eta(\nu) \geqslant \int_{M_\mathcal{E}(X,Y,\mu)} m \cdot dP_\eta(\nu) = m.$$

■

现在我们可以证明本节的主要结果：命题 16 给出的上界实际上是紧的，并且通过对不变扩张进行最小化，精确刻画了本质覆盖半径。

定理 1　9 设X、$Y \subseteq \Sigma^{\mathbb{Z}}$ 为受限系统，且设 $\mu \in M_\mathcal{E}(Y)$ 为遍历测度。那么

$$R_0(X, Y, \mu) = \inf\{\mathbb{P}_\nu[\mathbf{X}_0 \neq \mathbf{Y}_0] \mid \nu \in M_\mathcal{E}(X, Y, \mu)\}$$
$$= \inf\{\mathbb{P}_\nu[\mathbf{X}_0 \neq \mathbf{Y}_0] \mid \nu \in M(X, Y, \mu)\}.$$

证明：由命题 18 和命题 16，只需证明对于任意 $\delta > 0$，存在一个不变扩张$\nu \in M(X,Y,\mu)$，使得

$$\mathbb{P}_\nu[\mathbf{X}_0 \neq \mathbf{Y}_0] \leqslant R_0(X, Y, \mu) + \delta. \tag{12}$$

我们将利用概率测度单纯形$X \times Y$ 在弱-$*$ 拓扑下的紧致性，证明这样的扩张存在。我们固定 $\delta > 0$，并在$(0, 1)$ 中取 $\varepsilon < \delta/4$，使得

$$R_\varepsilon(X, Y, \mu) \leqslant R_0(X, Y, \mu) + \frac{\delta}{4}. \tag{13}$$

对于任意固定的词 $\overline{y} \in \mathcal{B}_n(Y)$，我们任意选取一个 $\overline{x}(y) \in \mathcal{B}_n(X)$，使其在 $\mathcal{B}_n(X)$ 中的所有词里距离 y 最近。根据$R_\varepsilon(X,Y,\mu)$ 的定义，存在一个互不相同的整数序列 $(n_k)_{k\in\mathbb{N}}$，使得

$$\lim_{k\to\infty} \frac{R_\varepsilon(\mathcal{B}_{n_k}(X), \mathcal{B}_{n_k}(Y), \mu_{n_k})}{n_k} = R_\varepsilon(X, Y, \mu).$$

根据 $R_\varepsilon(\mathcal{B}_{n_k}(X), \mathcal{B}_{n_k}(Y), \mu_{n_k})$ 的定义，当 k 充分大时，我们有

$$\mathbb{P}_\mu\left[\frac{1}{n_k} d\left(\overline{x}(\mathbf{Y}_0^{n_k-1}), \mathbf{Y}_0^{n_k-1}\right) < R_\varepsilon(X, Y, \mu) + \frac{\delta}{4}\right] > 1 - \varepsilon. \tag{14}$$

对于固定的 $\mathbf{k}$，我们定义一个映射 $\mathbf{f}_k : \mathbf{Y} \to \mathbf{X}$ 如下：对于任意 $\mathbf{x} \in \mathcal{B}_{n_k}(\mathbf{X})$，我们选取某个 $\mathbf{x} \in \mathbf{X}$，使得 $\mathbf{x}_0^{n_k-1} = \mathbf{x}$。然后定义 $\mathbf{f}_k(\mathbf{y}) = \mathbf{x}$，其中 $\mathbf{x}$ 是 $\mathbf{X}$ 中与 $\mathbf{x}(\mathbf{y}_0^{n_k-1})$ 对应的序列。显然，$\mathbf{f}_k$ 是可测的，因为 $\mathbf{f}_k(\mathbf{y})_m$ 仅取决于 $\mathbf{y}$ 的有限个坐标。现在考虑 $\mathbf{X} \times \mathbf{Y}$ 上的测度 $\nu'\mathbf{k}$，它定义为映射 $(\mathbf{f}_k, \mathrm{Id}) : \mathbf{Y} \to \mathbf{X} \times \mathbf{Y}$ 的推前。我们还定义 ν_k 为沿移位作用对 $\nu'\mathbf{k}$ 取平均所得到的测度，其定义为

$$\nu = \frac{1}{n_k} \sum_{i=0}^{n_k-1} T_*^i \nu'_k,$$

其中 $T_*^i \nu'k$ 是 $\nu'k$ 经第i次移位后的推前测度，定义为 $T_*^i \nu'k(A) = \nu'k(T^{-i}A)$。注意， $\nu'k$在Y上的边缘分布为 μ，因为(f_k, Id) 在Y坐标上是恒等映射。由于 μ 是移位不变的，因此 ν_k 在Y上的边缘分布也为 μ。

根据 $X \times Y$ 上的概率测度集合关于弱-$*$ 拓扑的紧致性，存在一个收敛子列$(\nu_{k_l})_{l\in\mathbb{N}}$，为简便起见，仍记为$(\nu_k)_k$。令 ν 表示$(\nu_k)_k$的弱-$*$ 极限。由于测度到其边缘分布的投影关于弱-$*$ 拓扑连续， ν 在Y上的边缘分布为 μ。现在我们将证明 ν 确实是不变测度，并满足 (12)。

对于 T 由于 ν 的不变性，只需证明对于 $X \times Y$ 上的任意连续函数，我们都有

$$\int f \cdot d\nu = \int f \circ T \cdot d\nu.$$

事实上，

$$\begin{aligned}\int f\cdot d\nu-\int f\circ T\cdot d\nu&=\lim_{k\to\infty}\int f\cdot d\nu_k-\lim_{k\to\infty}\int f\circ T\cdot d\nu_k\\&=\lim_{k\to\infty}\frac{1}{n_k}\sum_{i=0}^{n_k-1}\left(\int f\cdot dT_*^i\nu_k'-\int f\circ T\cdot dT_*^i\nu_k'\right)\\&=\lim_{k\to\infty}\frac{1}{n_k}\sum_{i=0}^{n_k-1}\left(\int f\circ T^i\cdot d\nu_k'-\int f\circ T^{i+1}\cdot d\nu_k'\right)\\&=\lim_{k\to\infty}\frac{1}{n_k}\left(\int f\cdot d\nu_k'-\int f\circ T^{n_k}\cdot d\nu_k'\right)=0,\end{aligned}$$

收敛到 0 是因为f 有界（因为它是紧空间上的连续函数）。现在还需证明 ν 满足（12）。根据 ν、ν_k和 $\nu'k$的定义，对于所有k ，我们有

$$\begin{aligned}\mathbb{P}_\nu[\mathbf{X}_0\neq\mathbf{Y}_0]&=\lim_{k\to\infty}\mathbb{P}_{\nu_k}[\mathbf{X}_0\neq\mathbf{Y}_0]=\lim_{k\to\infty}\frac{1}{n_k}\sum_{i=0}^{n_k-1}\mathbb{P}_{\nu_k'}\left[T^{-i}\{\mathbf{X}_0\neq\mathbf{Y}_0\}\right]\\&=\lim_{k\to\infty}\frac{1}{n_k}\sum_{i=0}^{n_k-1}\mathbb{P}_{\nu_k'}[\mathbf{X}_i\neq\mathbf{Y}_i]=\lim_{k\to\infty}\mathbb{E}_{\nu_k'}\left[\frac{1}{n_k}\sum_{i=0}^{n_k-1}I_{\mathbf{X}_i\neq\mathbf{Y}_i}\right]\\&=\lim_{k\to\infty}\mathbb{E}_{\nu_k'}\left[\frac{1}{n_k}d\left(\mathbf{Y}_0^{n_k-1},\mathbf{X}_0^{n_k-1}\right)\right]=\lim_{k\to\infty}\mathbb{E}_{\mu}\left[\frac{1}{n_k}d\left(\mathbf{Y}_0^{n_k-1},\overline{x}(\mathbf{Y}_0^{n_k-1})\right)\right].\end{aligned}$$

我们将E_δ 定义为如下事件：$d(\mathbf{Y}_0^{n_k-1}$、$x(\mathbf{Y}_0^{n_k-1}))\geqslant n_k(R\varepsilon(X,Y,\ \mu)+\delta/4)$。由（14），当$k$足够大时，

$$\begin{aligned}\mathbb{E}_\mu\left[\frac{1}{n_k}d\left(\mathbf{Y}_0^{n_k-1},\overline{x}(\mathbf{Y}_0^{n_k-1})\right)\right]&\leqslant\mathbb{E}_\mu\left[\frac{1}{n_k}d\left(\mathbf{Y}_0^{n_k-1},\overline{x}(\mathbf{Y}_0^{n_k-1})\right)\cdot I_{E_\delta}\right]+\mathbb{E}_\mu[1-I_{E_\delta}]\\&\leqslant R_\varepsilon(X,Y,\mu)+\frac{\delta}{4}+(1-\mathbb{P}_\mu[E_\delta])\leqslant R_\varepsilon(X,Y,\mu)+\frac{\delta}{4}+\varepsilon\leqslant R_\varepsilon(X,Y,\mu)+\frac{\delta}{2}.\end{aligned}$$

将上述不等式与（13）结合，我们得出

$$\mathbb{P}_\nu[\mathbf{X}_0\neq\mathbf{Y}_0]=\lim_{k\to\infty}\mathbb{P}_{\nu_k}[\mathbf{X}_0\neq\mathbf{Y}_0]\leqslant R_0(X,Y,\mu)+\frac{3\delta}{4},$$

正如所需。 ■

在下面的例子中，我们明确描述了$M_{\mathcal{E}}(X$、Y、$\mu)$中的一个扩张序列，它逼近相对于全移位（配备均匀伯努利测度）的例 1 中$(0,k)$-受限游程长度系统的本质覆盖半径。

例3 令$X_{0,k}\subseteq[2]^{\mathbb{Z}}$ 表示例 1 中的$(0,k)$-受限游程长度移位。设$y\in[2]^n$ 为一个有限二进制字。我们将$c(y)$定义为y的最长零后缀的长度，形式化定义如下：

$$c(\overline{y})\triangleq\max\left\{i\,\middle|\,\overline{y}=\overline{y}_0^{n-i-1}\overline{0}_i\right\}.$$

固定 $N\in\mathbb{N}$ ，并考虑映射 $f_{(N)}\colon[2]^{\mathbb{Z}}\to X_{0,k}$ ，其定义为

$$f^{(N)}(\mathbf{y})_m=\begin{cases}1 & c\left(\mathbf{y}_{m-(N(k+1)-1)}^{m-1}\right)\equiv k\pmod{k+1},\\ \mathbf{y}_m & otherwise.\end{cases}$$

显然，$\mathrm{Im}(f)\subseteq X_{0,k}$因为$f_{(N)}(\mathbf{y})$中不会出现长度为$k+1$的零连续段。注意，映射$(f_{(N)},\mathrm{Id})\colon[2]^{\mathbb{Z}}\to X_{0,k}\times[2]^{\mathbb{Z}}$是一个滑块码函数（即每个坐标处的值由相邻坐标构成的一个有限块决定的函数），因此它是可测的，并且与移位变换可交换。令 μ_{u} 为

在 $[2]^{\mathbb{Z}}$上的均匀测度，并令 ν_N为通过$f_{(N)}$将其推前到$X_{0,k} \times [2]^{\mathbb{Z}}$所得的测度。显然， ν_N是不变测度，同时也是遍历测度（因为它是不变遍历测度的因子）。因此， $\nu_N \in M_{\mathcal{E}}(X_{0,k}, [2]^{\mathbb{Z}}, \mu_{\mathrm{u}})$。注意，

$$\begin{aligned}
\mathbb{P}_{\nu_N}[\mathbf{X}_0 \neq \mathbf{Y}_0] &= \mathbb{P}_{\mu^{\mathrm{u}}}\left[c\left(\mathbf{Y}^{-1}_{-(N(k+1)-1)}\right) \equiv k \pmod{k+1} \text{ and } \mathbf{Y}_0 = 0\right] \\
&= \sum_{i=0}^{N-1} \mathbb{P}_{\mu^{\mathrm{u}}}\left[c\left(\mathbf{Y}^{-1}_{-(N(k+1)-1)}\right) = i(k+1)+k \text{ and } \mathbf{Y}_0 = 0\right] \\
&= \mathbb{P}_{\mu^{\mathrm{u}}}\left[\mathbf{Y}^{0}_{-(N(k+1)-1)} = \overline{0}_{N(k+1)}\right] + \sum_{i=1}^{N-1} \mathbb{P}_{\mu^{\mathrm{u}}}\left[\mathbf{Y}^{0}_{-i(k+1)} = 1\overline{0}_{i(k+1)}\right] \\
&= \frac{1}{2^{N(k+1)}} + \frac{1}{2}\sum_{i=1}^{N-1} \frac{1}{2^{i(k+1)}}.
\end{aligned}$$

取 $N \to \infty$，我们得到

$$\lim_{N\to\infty} \mathbb{P}_{\nu_N}[\mathbf{X}_0 \neq \mathbf{Y}_0] = \lim_{N\to\infty} \frac{1}{2^{N(k+1)}} + \frac{1}{2}\sum_{i=1}^{N-1} \frac{1}{2^{i(k+1)}} = \frac{1}{2(2^{k+1}-1)} = R_0(X_{0,k}, [2]^{\mathbb{Z}}, \mu^{\mathrm{u}}).$$

最后，我们简要讨论 QCC 方案背景下的本质覆盖半径。粗略地说，在渐近意义下，除去一个比例趋于零的部分，所有 Σ^n 都可以通过改变 $\mathrm{R}_0(\mathrm{X}, \Sigma^{\mathbb{Z}}, \mu\mathrm{u})$ 比例的位置，量化为 $\mathcal{B}_n(\mathrm{X})$。这个比例可能远低于最坏情况下的比例 $\mathrm{R}(\mathrm{X}, \Sigma^{\mathrm{Z}})$。在有限长度情形下，至少有 $(1-\varepsilon)$ 比例的 Σ^n 可以通过至多改变 $\mathrm{r}_\varepsilon = \mathrm{R}\varepsilon(\mathcal{B}_n(\mathrm{X}), \Sigma^n, \mu\mathrm{un})$ 个位置，量化为 $\mathcal{B}_n(\mathrm{X})$。我们需要克服的一个小障碍是，在 QCC 方案中，我们并不对 Σ^n中的任意词进行量化，而只对纠错码 C 的码字进行量化。Σ^n 中与 $\mathcal{B}_n(\mathrm{X})$ 距离较远的词所占的 ε比例，可能会不成比例地集中在 C 中。然而，若进一步假设 C 是线性纠错码，则通过简单的平均论证，至少存在该码的一个陪集 C'，使得其中与 X 的语言距离至多为 $\mathrm{r}\varepsilon$ 的码字所占比例满足相应条件。这意味着存在 $\mathrm{C}'' \subseteq \mathrm{C}'$，且 $|\mathrm{C}''| \geqslant (1-\epsilon)|\mathrm{C}'|$，使得 $\mathrm{R}(\mathrm{B}_n(\mathrm{X}), \mathrm{C}'') \leqslant \mathrm{r}_\epsilon$。

V. 本质覆盖半径的上界

本节的目标是建立本质覆盖半径的一般上界。定理 19 给出了本质覆盖半径的精确表达式，但其中涉及的极小化问题很难求解。在例3中，我们找到了一列良好的扩张，可用于逼近本质覆盖半径。一般而言，根据定理 19，这样的扩张序列可给出本质覆盖半径的一系列上界。本节将介绍两种为一般受限系统构造扩张的不同方法，从而给出本质覆盖半径的上界。第一种方法利用马尔可夫链，所得上界可通过求解线性规划问题高效计算。另一种构造扩张的方法是使用滑块码函数。在这种情况下，我们证明，若X 是本原的，则可通过增大此类函数的块大小来逼近本质覆盖半径。

A. 马尔可夫链

我们考虑这样的情形：X 和 Y 分别是由标号图 $\mathrm{G}_X = (\mathrm{V}_X, \mathrm{E}_X, \mathrm{L}_X)$ 和 $\mathrm{G}_Y = (\mathrm{V}_Y, \mathrm{E}_Y, \mathrm{L}_Y)$ 生成的受限系统。在本节中，我们假设 G_X 和 G_Y 不包含具有相同标签的平行边。边 $\mathrm{u} \to \mathrm{v}$ 记作有序对 $\mathrm{e} = (\mathrm{u}, \mathrm{v})$，并称其源点为 $\sigma(\mathrm{e}) = \mathrm{u}$，目标点为$\tau(\mathrm{e}) = \mathrm{v}$。我们关注测度 $\mu \in \mathrm{M}_{\mathcal{E}}(\mathrm{Y})$ 由图 G_Y上的某个马尔可夫链生成的情形。值得一提的是，$\mathrm{Y} = \Sigma^{\mathbb{Z}}$ and $\mu = \mu\mathrm{u}$ 为均匀伯努利测度的情形也属于这一类。我们先介绍马尔可夫链理论中关于有限图的一些定义和基本结果。

定义20 令 $G = (V,E)$ 为一个有限有向图。图 G 上的平稳马尔可夫链是一个二元组 (π,Q)，其中 π 是 V 上的概率测度，Q 是一个从 V 到 E 上概率测度空间的函数，它将 V 中的 $v \in$ 映射到 E 上的概率测度 $Q(\cdot|v)$，并且对每个 $v \in V$，

$$\sum_{\substack{e\in E \\ \sigma(e)=v}} Q(e|v) = 1,$$

并且，使得对于每个 $v \in V$，我们都有：

$$\pi(v) = \sum_{\substack{e\in E \\ \tau(e)=v}} \pi(\sigma(e))Q(e|\sigma(e)).$$

注意，对于 G = (V, E) 上的任意马尔可夫链 (π, Q)，Q(e|v) > 0 意味着 σ(e) = v，因此我们可以方便地将 Q(e) 作为 Q(e|σ(e)) 的简写。当 G 是简单图（即没有平行边）时，对于任意边 e = (u, v) ∈ E，我们用 Q(v|u) 表示 Q(e)。此外，当 G 是简单图时，可以将 Q 视为一个 |V| × |V|随机矩阵（通常称为转移矩阵），其中 π 是特征值为 1 的左特征向量。

在$G = (V,E)$ 上的马尔可夫链与E 上的满足以下条件的概率测度一一对应

$$\sum_{\substack{e\in E\\ \sigma(e)=v}} P(e) = \sum_{\substack{e\in E\\ \tau(e)=v}} P(e).$$

Ind因此，这样的概率测度P 对应于平稳马尔可夫链(π,Q)，其中

$$\pi(v) \triangleq \sum_{\substack{e\in E\\ \sigma(e)=v}} P(e) = \sum_{\substack{e\in E\\ \tau(e)=v}} P(e),$$

and

$$Q(e|\sigma(v)) \triangleq \frac{P(e)}{\pi(\sigma v)}.$$

为简化记法，我们记$P = (\pi,Q)$。假设该马尔可夫链不含退化顶点，即，对所有$v \in V$，$\pi(v) > 0$。任意平稳马尔可夫链P都会在G 上的双无限路径空间上诱导一个不变测度$\hat{}P$，其定义为

$$\hat{P}([(e_0, e_1, \ldots, e_{n-1})]) = P(e_0)\prod_{i=1}^{n-1} Q(e_i).$$

对于任意对应于有限路径$(e_0, \ldots, e_{n-1})$的柱集 $[(e_0, \ldots, e_{n-1})]$。我们称$\hat{}P$ 为由P诱导的G上的平稳马尔可夫过程。若G 是简单图（即不存在平行边），则对于任意边$e = (u,v) \in E$，我们用$Q(v|u)$表示$Q(e)$。

若$G = G_Y$ 通过标记函数L_Y生成受限系统Y，则$P = (\pi,Q)$ 在Y上诱导一个不变概率测度，该测度是通过标记函数将$\hat{}P$ 推前得到的测度，即对于柱集 $[y]$,

$$\mu_P([\overline{y}]) = \sum_{\substack{\overline{e} \text{ path in } G\\ L(\overline{e})=\overline{y}}} \pi(\sigma(e_0))\prod_{i=0}^{|\overline{y}|-1} Q(e_i).$$

我们注意到$\mathbf{Y}$ 相对于 μ_P是一个隐马尔可夫过程，并将上述测度 μ_P 称为通过标记函数L由P 诱导的隐马尔可夫测度。

假设 X、Y $\subseteq \Sigma^{\mathbb{Z}}$ 是分别由标号图 G_X 和 G_Y 给出的不可约受限系统，并假设 $\mu = \mu_{P_Y} \in M_{\mathcal{E}}(Y)$ 是 Y 上的一个测度，由 G_Y 上的平稳马尔可夫链 P_Y 诱导。我们考虑由 G_X 和 G_Y 给出的强积图 $G_{X\times Y} = (V_{X\times Y}, E_{X\times Y}, (L_X, L_Y))$，其中 $V_{X\times Y}$, $V_X \times V_Y$, $E_{X\times Y}$, $E_X \times E_Y$ 满足$\sigma(e_x, e_y) = (\sigma(e_x), \sigma(e_y))$，$\tau(e_x, e_y) = (\tau(e_x), \tau(e_y))$，且标记函数 $L_{X\times Y}$ 定义如下：

$$L_{X\times Y}(e_x, e_y) = (L_X(e_x), L_Y(e_y)).$$

注意，定义在$G_{X\times Y}$ 上的平稳马尔可夫链P 自然定义了一个定义在$G_{X\times Y}$上的平稳马尔可夫过程$\hat{}P$，它通过标记函数$L_{X\times Y}$在$X \times Y$ 上诱导出隐马尔可夫测度 ν_P。由于每条边$e \in E_{X\times Y}$ 由一对$(ex,ey) \in E_X \times E_Y$组成，$\hat{}P$ 可以视为定义在双无限路径对空间上的测度，其中第一条是G_X 上的路径，第二条是G_Y上的路径。我们将$\hat{}P$ 在G_Y 上的双无限路径空间上的边缘测度记为$(\hat{}P)_Y$

就我们的目的而言，我们关注定义在$G_{X\times Y}$ 上的平稳马尔可夫链，其诱导测度的Y-边缘分布为$\hat{}\mu_{P_Y}$。以下是一个充分条件：将P 视为定义在$G_{X\times Y}$上的平稳马尔可夫过程时，$\hat{}P$ 在G_Y上的边缘测度是马尔可夫过程$\hat{}P_Y$。因此我们关注以下问题：给定定义在$G_{X\times Y}$上的平稳马尔可夫链P，何时$\hat{}P$ 是测度$\hat{}P_Y$的扩张？一个显而易见的必要条件是，$\hat{}$分布中P 的G_Y-边缘对分布等于P_Y的对分布。这个条件也可以等价地表述为

$$P_Y(e) = \sum_{\substack{e'\in E_{X\times Y}\\ e'_y=e}} P(e'), \quad \text{for all } e \in E_Y. \tag{15}$$

人们很容易猜测，式 (15) 中给出的条件也是充分条件，但事实并非如此。马尔可夫测度的边缘分布通常不一定是马尔可夫测度，而是隐马尔可夫测度。事实上，存在满足式 (15) 的平稳马尔可夫链 P and P_Y，使得在$G_{X\times Y}$ 上的双无限路径空间中，不存在这样的不变测度：它的对分布与 $\hat{}P$ 相同，且其G_Y-边缘分布等于马尔可夫过程 $\hat{}P_Y$。在下面的引理中，我们提出一个充分条件，使得 $\hat{}P$ 的G_Y-边缘分布等于 $\hat{}P_Y$。为便于记述，

我们针对简单图陈述并证明这一结论。不过，即使不假设图的简单性，该结论仍然成立，而且证明可立即从简单图推广到一般标号图。一般情形（不一定是简单图）下的相应条件见式 (19)。

引理 21 设 $\mathrm{G}_1=(\mathrm{V}_1,\mathrm{E}_1)$ 和 $\mathrm{G}_2=(\mathrm{V}_2,\mathrm{E}_2)$ 为有限简单有向图，并令 $\mathrm{G}_1\times\mathrm{G}_2=\mathrm{G}=(\mathrm{V},\mathrm{E})$ 表示 G_1 和 G_2 的强图积（定义如上）。对于分别定义在 $\hat{}\,\hat{}\,\mathrm{G}$ 和 G_1 上的给定平稳马尔可夫链 $\mathrm{P}=(\pi,\mathrm{Q})$ 和 $\mathrm{P}_1=(\pi_1,\mathrm{Q}_1)$，若满足以下条件，则马尔可夫过程 P 的 G_1-边缘分布是马尔可夫过程 P_1（定义在 G_1上）：

- 对所有 $e\in E_1$

$$P_1(e)=\sum_{\substack{e'\in E\\ e'_1=e}}P(e') \tag{16}$$

- 对所有 $u_0\in V_2$ 以及 $v_0,v_1\in V_1$，都有

$$Q(v_1|v_0,u_0)\triangleq\sum_{u_1\in V_2}Q((v_1,u_1)|(v_0,u_0))=Q_1(v_1|v_0). \tag{17}$$

证明：证明只需直接计算。记 $(\hat{}P)_1$ 为 $\hat{}P$ 在 G_1 上的边缘测度。根据全概率公式，对于 G_1, 上的任意路径 $e=(e_0,\ldots,e_{n-1})=(v_0,v_1),(v_1,v_2),\ldots,(v_{n-1},v_n))$

$$\begin{aligned}
(\hat{P})_1([\overline{e}]) &= \sum_{u_0,\ldots,u_n\in V_2}P((v_0,u_0),(v_1,u_1))\prod_{i=1}^{n-1}Q((v_{i+1},u_{i+1})|(v_i,u_i))\\
&= \sum_{u_0,u_1\in V_2}P((v_0,u_0),(v_1,u_1))\prod_{i=1}^{n-1}\left(\sum_{u_{i+1}\in V_2}Q((v_{i+1},u_{i+1})|(v_i,u_i))\right)\\
&\overset{(a)}{=} \sum_{u_0,u_1\in V_2}P((v_0,u_0),(v_1,u_1))\prod_{i=1}^{n-1}Q_1(v_{i+1}|v_i)\\
&\overset{(b)}{=} P_1(v_0,v_1)\prod_{i=1}^{n-1}Q_1(v_{i+1}|v_i)=\hat{P}_1([\overline{e}]),
\end{aligned}$$

其中，(a) 和 (b) 分别由假设 (16) 和 (17) 得出。 ■

现在，我们可以给出$R_0(X,Y,\mu)$的一个上界，并将其表述为一个关于平稳马尔可夫链的优化问题。设P_Y 和P 分别是定义在G_Y 和$G_{X\times Y}$ 上的平稳马尔可夫链，且P 满足引理 21 的条件。记由P 诱导的$X\times Y$ 上的测度为 ν_P，并记由P_Y 诱导的Y 上的测度为 μ。根据引理 21，我们有 $\nu_P\in M(X,Y,\mu)$，因此根据定理 19，我们有

$$R_0(X,Y,\mu)\leqslant\mathbb{P}_{\nu_P}[\mathbf{X}_0\neq\mathbf{Y}_0]=\sum_{\substack{e=E_{X\times Y}\\ L_X(e)\neq L_Y(e)}}P(e).$$

注意，若将马尔可夫链P视为包含在 $[0,1]^{E_X\times_{E^Y}}$中的单纯形内的一个元素，则$\nu_P[\mathbf{X}_0\neq\mathbf{Y}_0]$ 是P的线性函数。此外，条件 (16) 关于P 是线性的，因为它是由P中的元素求和所定义的约束。对于简单图，直接计算可知 (17) 也是一个线性条件，因为$Q(v_1^y|v_0^x,v_0^y)=Q(v_1^y|v_0^y)$ 当且仅当

$$\sum_{v_1^x\in V_X}P\Big((v_1^x,v_1^y),(v_0^x,v_0^y)\Big)=Q_Y(v_1^y|v_0^y)\cdot\sum_{\substack{v^x\in V_X\\ v^y\in V_Y}}P\Big((v^x,v^y),(v_0^x,v_0^y)\Big) \tag{18}$$

我们注意到，(18) 可以等价地表述为一个关于边的条件；在该条件下，引理 21 的结论在一般情况下成立。也就是说，如果满足 (16)，且对所有 $e\in E_{X\times Y}$，引理 21 的结论都成立

$$\sum_{\substack{e'\in E_{X\times Y}\\ e'_y=e_y\\ \sigma(e'_x)=\sigma(e_x)}}P(e')=Q_Y(e_y)\sum_{\substack{e'\in E_{X\times Y}\\ \sigma(e'_y)=\sigma(e_y)\\ \sigma(e'_x)=\sigma(e_x)}}P(e') \tag{19}$$

因此，只需在满足（16）和（19）的所有定义在$\mathbf{X}_0\neq\mathbf{Y}${v7>上的平稳马尔可夫链中，最小化该量，便可得到一个上界；这最终转化为一个线性规划问题。

定理 22 设X、$Y \subseteq \Sigma^{\mathbb{Z}}$ 分别是由标号图G_X 和G_Y 定义的移位空间，且P_Y 是Y 上的一个平稳马尔可夫链，其诱导出测度 μ。则

$$R_0(X, Y, \mu) \leqslant \mathrm{MB}(G_X, G_Y, P_Y),$$

其中 MB(G_X,G_Y, P_Y) 是以下线性规划问题的解：

$$\begin{aligned}
\underset{P \in \mathbb{R}^{E_{X\times Y}}}{\text{minimize}} \quad & \sum_{\substack{e\in E_{X\times Y}\\ L_X(e)\neq L_Y(e)}} P(e) \\
\text{subject to} \quad & P(e) \geqslant 0, && \forall e \in E_{X\times Y}, \\
& \sum_{e\in E_{X\times Y}} P(e) = 1, \\
& \sum_{\substack{e'\in E_{X\times Y}\\ e'_y = e}} P(e') = P_Y(e), && \forall e \in E_Y, \\
& \sum_{\substack{e\in E_{X\times Y}\\ \sigma(e)=v}} P(e) = \sum_{\substack{e\in E_{X\times Y}\\ \tau(e)=v}} P(e), && \forall v \in V_{X\times Y}, \\
& \sum_{\substack{e'\in E_{X\times Y}\\ e'_y = e_y\\ \sigma(e'_x)=\sigma(e_x)}} P(e') = Q_Y(e_y) \sum_{\substack{e'\in E_{X\times Y}\\ \sigma(e'_y)=\sigma(e_y)\\ \sigma(e'_x)=\sigma(e_x)}} P(e') && \forall e \in E_{X\times Y}.
\end{aligned}$$

例 4 考虑$X_{0,k}$，即例 1 中的(0,k)-RLL 移位。取$Y = [2]^{\mathbb{Z}}$，并令 μ_{u} 为Y上的均匀伯努利测度。我们考虑图2中所示、分别生成X 和Y 的标号图G_X 和G_Y。令P_Y 为E_Y 上的均匀测度（诱导出测度 μ_{u}）。因此，乘积图$G_{X\times Y}$ （见图 3）是图G_X的“双重”版本。

我们考虑马尔可夫测度P，其边概率如图 3 所示。对于适当的 α 取值，P确实是满足式 (16) 和 (19) 的平稳马尔可夫链。首先，为了在边上得到一个概率测度，我们要求

$$1 = \sum_{e\in G_{X\times Y}} P(e) = 2\alpha(2^k + 2^{k-1} + \cdots + 2 + 1) = 2\alpha(2^{k+1} - 1),$$

这意味着 $\alpha = 2^{-1}(2^{k+1}-1)^{-1}$。注意，对于$V_{X\times Y}$中的第$j$个状态，

$$\sum_{\sigma(e)=j} P(e) = \sum_{\tau(e)=j} P(e) = \alpha \cdot 2^{k-j},$$

这意味着 P 确实是平稳的。我们还注意到，对于任何满足 $L_Y(e) = 0$ 的边，都存在一条对应的边 e'，满足 $L_Y(e') = 1$，使得 $P(e) = P(e')$。这表明 P 在 G_Y 上的边缘分布确实是 $P_Y = \mu_{\mathrm{u}}$。我们进一步注意到，对于任意边 $e \in E_{X\times Y}$，如果 $\sigma(e_x) = j \in V_X$，我们有

$$\sum_{\substack{e'\in E_{X\times Y}\\ e'_y = e_y\\ \sigma(e'_x)=j}} P(e') = \alpha 2^{k-j-1} = \frac{1}{2}\alpha 2^{k-j} = Q_Y(e_y) \sum_{\substack{e'\in E_{X\times Y}\\ \sigma(e'_y)=\sigma(e_y)\\ \sigma(e'_x)=j}} P(e').$$

现在我们来计算：

$$\sum_{\substack{e\in G_{X\times Y}\\ L_X(e)\neq L_Y(e)}} P(e) = \alpha = \frac{1}{2(2^{k+1}-1)} = R_0(X_{0,k}, [2]^{\mathbb{Z}}, \mu^{\mathrm{u}}).$$

因此，根据定理 14 和定理 22，P 在该线性规划问题 MB(G_X,G_Y, P_Y)中取得最小值；特别地，在这种情况下，定理 22 给出的上界是紧的。

例 5 令$X = X_{d,\infty}$ 为(d, ∞)-受限游程长度系统，其定义为：任意两个连续的 1 之间至少有d 个零。等价地，$X_{d,\infty}$ 由G_X 定义，如图 4 所示。令$Y = [2]^{\mathbb{Z}}$ 和 μ_{u} 如例 1 所示。乘积图$G_{X\times Y}$ 如图 5 所示。我们考虑由图 5 所示边概率定义的马尔可夫测度P。对于 $\alpha = (2(d+2))^{-1}$，我们有

$$1 = \sum_{e\in G_{X\times Y}} P(e) = 2\alpha(d+2),$$

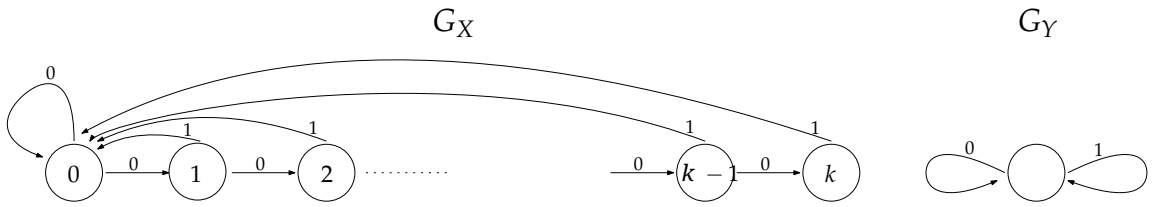

图2. 生成移位空间 $X = X_{0,k}$ （$(0,k)$-RLL 移位）和 $Y = [2]^{\mathbb{Z}}$ 的标记图。

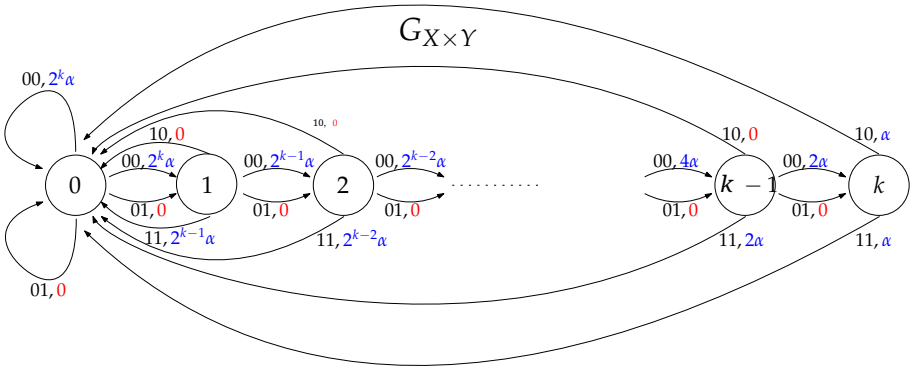

图3. 图2中图 G_X 和图 G_Y 的乘积图 $G_{X\times Y}$。每条边都标有一个两位标签 xy，对应于来自 G_X 的标签 x 和来自 G_Y 的标签 y。通过在每条边的标签后写出 e 的 (P)，展示了一个达到界 $MB(G_X, G_Y, P_Y)$ 的平稳马尔可夫链。e 的 $(P) = 0$ 的边以红色标记，正概率 e 的 $(P) > 0$ 则以蓝色标出。

因此P 确实是一个满足定理 22 条件的平稳马尔可夫链。现在我们计算上界：

$$R_0(X_{d,\infty}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) \leqslant \sum_{\substack{e\in G_{X\times Y}\\ L_X(e)\neq L_Y(e)}} P(e) = \alpha\cdot d = \frac{d}{2(d+2)}. \tag{20}$$

我们注意到 可知，当 $d = 1$ 时，系统 $X_{1,\infty}$ 通过对所有比特取补，与 $X_{0,1}$，同构。因 s
t

$$R_0(X_{1,\infty}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) = R_0(X_{0,1}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) = \frac{1}{6},$$

特别地，界 (20) 是紧的。

对于下界，我们断言

$$R_0(X_{d,\infty}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) \geqslant \frac{1}{2} - \frac{1}{d+1}.$$

根据本质覆盖半径的定义，证明很直接。注意，当 $d = 1$ 时，该界没有意义，因此我们假设 $d \geqslant 2$。我们先作一个简单的观察。令 $\mathrm{wt}(w) = d(w,\mathbf{0})$ 表示二元词 w 的非零坐标数。注意，对任意 $x,y \in [2]^n$，

$$d(\overline{x},\overline{y}) \geqslant |\mathrm{wt}(\overline{y}) - \mathrm{wt}(\overline{x})| \geqslant \mathrm{wt}(\overline{y}) - \mathrm{wt}(\overline{x}).$$

根据 $X_{d,\infty}$ 系统的定义，对任意 $x \in \mathcal{B}_n(X)$，任意两个相邻的 1 之间至少有 d 个零。因此，对任意 $x \in \mathcal{B}_n(X)$，我们有

$$\mathrm{wt}(\overline{x}) \leqslant \left\lceil \frac{n}{d+1} \right\rceil \leqslant \frac{n}{d+1} + 1.$$

因此，对任意词 $y \in [2]^n$，我们有

$$\min_{\overline{x}\in\mathcal{B}_n(X)} d(\overline{x},\overline{y}) \geqslant \mathrm{wt}(\overline{y}) - \frac{n}{d+1} - 1. \tag{21}$$

令$\mathbf{Y}$为由分布 μ_{u}生成的随机双无限序列。也就是说，$(\mathbf{Y}_k)_{k\in\mathbb{Z}}$ 是独立同分布的 $\mathrm{Ber}(1/2)$ 随机变量。应用大数定律，我们有

$$\frac{1}{n}\,\mathrm{wt}\left(\mathbf{Y}_0^{n-1}\right) = \frac{1}{n}\sum_{i=0}^{n-1}\mathbf{Y}_i \xrightarrow[n\to\infty]{\mathbb{P}_{\mu^{\mathsf{u}}}} \mathbb{E}_{\mu^{\mathsf{u}}}[\mathbf{Y}_0] = \frac{1}{2}.$$

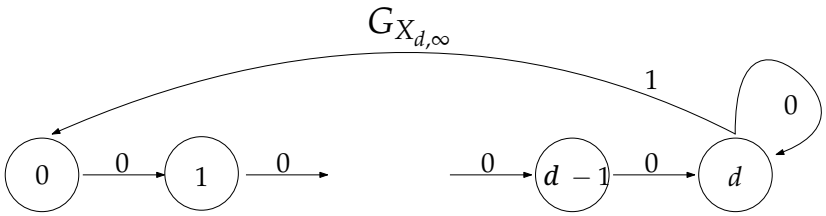

图 4. 生成受限系统 $X = X_{d,\infty}$ （即(d,∞)-RLL 移位）的标号图。

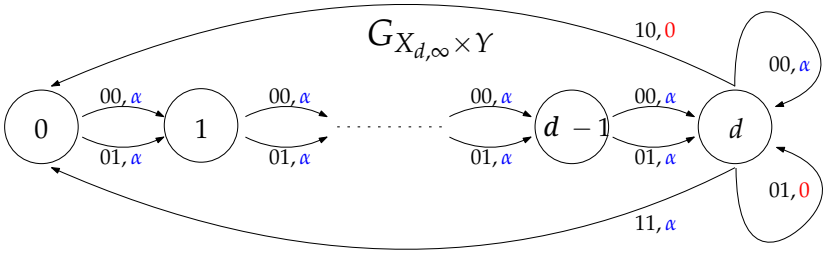

图5. 图 $G_{X\times Y}$ 是图$G^{X_{d,\infty}}$ 和图G_Y 的乘积图，分别对应图 4 和图2。每条边都标有两位标签 xy，对应于图$G^{X_{d,\infty}}$ 中的标签x 以及图G_Y中的标签y 。通过在每条边的标签后写出$P(e)$ ，展示了一个达到界$MB(G_X, G_Y, P_Y)$ 的平稳马尔可夫链。$P(e) = 0$ 的边以红色标出，正概率$P(e) > 0$ 则以蓝色标出。

也就是说，随机词的归一化权重依概率收敛于 $\mathbb{P}_{\mu^{\mathsf{u}}}$ 12。结合 (21)，我们得到，对于每个 $\delta > 0$

$$\begin{aligned}&\mathbb{P}_{\mu_n^{\mathsf{u}}}\left[\min_{\overline{x}\in\mathcal{B}_n(X_{d,\infty})}\frac{d(\overline{x},\mathbf{Y}_0^{n-1})}{n} > \frac{1}{2}-\frac{1}{d+1}-\frac{1}{n}-\delta\right]\\ &\geqslant \mathbb{P}_{\mu^{\mathsf{u}}}\left[\frac{1}{n}\operatorname{wt}(\mathbf{Y}_0^{n-1}) > \frac{1}{2}-\delta\right] \xrightarrow[n\to\infty]{} 1.\end{aligned}$$

这证明，对于任意 $\varepsilon \in (0,1)$， $\delta > 0$，当n充分大时

$$\frac{R_\varepsilon(\mathcal{B}_n(X_{d,\infty}), \mathcal{B}_n(Y), \mu^{\mathsf{u}})}{n} > \frac{1}{2}-\frac{1}{d+1}-\delta,$$

因此，

$$R_\varepsilon(X_{d,\infty}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) \geqslant \lim_{\delta\to 0}\frac{1}{2}-\frac{1}{d+1}-\delta = \frac{1}{2}-\frac{1}{d+1}.$$

取 $\varepsilon \to 0$，得到

$$R_0(X_{d,\infty}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) \geqslant \frac{1}{2}-\frac{1}{d+1}.$$

结合下界和上界，我们得到，

$$\frac{1}{2}-\frac{1}{d+1} \leqslant R_0(X_{d,\infty}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) \leqslant \frac{1}{2}-\frac{1}{d+2}.$$

B. 滑动区块码

现在我们介绍一种通过滑动区块码函数构造扩张的替代方法。我们先重新讨论例3，其中我们找到了一个扩张序列，用以逼近$X_{0,k}$关于 $[2]^{\mathbb{Z}}$、取 μ_{u}（均匀独立同分布测度）时的本质覆盖半径。构造这些扩张的主要思路如下：对于X、$Y \subseteq \Sigma^{\mathbb{Z}}$和 $\mu \in M_{\mathcal{E}}(Y)$，给定一个可测函数$g$： $Y \to X$，它与移位变换可交换，则映射(g, Id)： $Y \to X \times Y$通过 μ 在(g, Id)下的推前，在$M(X,Y,\mu)$ 中定义了一个扩张 ν_g。也就是说

$$\nu_g(A_X \times A_Y) \triangleq \mu\Big(A_Y \cap g^{-1}(A_X)\Big).$$

我们称这样的函数g 为从Y 到X的平稳编码函数。根据命题 16，对于任意这样的平稳编码函数g，

$$R_0(X,Y,\mu) \leqslant \mathbb{P}_{\nu_g}[\mathbf{X}_0 \neq \mathbf{Y}_0],$$

也就是说，任何与移位可交换的此类可测函数 g 都为本质覆盖半径提供一个上界。

在例3中，我们使用滑动块码函数（由局部规则定义的函数）作为与移位可交换的可测函数。滑动块码函数尤其值得关注，因为它们提供了丰富的函数族，且易于用局部规则描述。关于滑动块码的性质和构造，文献中已有大量研究（例如，参见 [1], [3], [4]）。本节旨在明确描述由滑动块码函数得到的界，并给出充分条件，使得可以利用通过滑动块码构造的扩张来逼近本质覆盖半径。

定义23 设X、$Y \subseteq \Sigma^{\mathbb{Z}}$ 为移位空间。若存在$N \in \mathbb{N}$ 和函数$f\colon \mathcal{B}_{2N+1}(Y) \to \Sigma$，使得对所有$\mathbf{y} \in Y$ 以及所有$i \in \mathbb{Z}$，函数$\hat{f}\colon Y \to X$ 满足以下条件，则称其为滑动块码：

$$\hat{f}(\mathbf{y})_i = f(\mathbf{y}_{i-N}^{i+N}).$$

在这种情况下，称$\hat{f}$ 为块长为N的滑动块码。

设$\hat{f}\colon Y \to X$ 是由局部函数$f\colon \mathcal{B}_{2N+1}(Y) \to \Sigma$定义的滑块码函数，且 $\mu \in M_{\mathcal{E}}(Y)$ 是一个遍历测度。我们将由$\hat{f}$得到的扩张记为ν_f。现在可以轻易计算出量 $\mathbb{P}_{\nu_f}[\mathbf{X}_0 \neq \mathbf{Y}_0]$：

$$\mathbb{P}_{\nu_f}[\mathbf{X}_0 \neq \mathbf{Y}_0] = \mathbb{P}_{\mu}[\hat{f}(\mathbf{Y})_0 \neq \mathbf{Y}_0] = \mathbb{P}_{\mu}\left[f(\mathbf{Y}_{-N}^{N}) \neq \mathbf{Y}_0\right] = \sum_{\substack{\overline{y} \in \mathcal{B}_{2N+1}(Y) \\ f(\overline{y}) \neq \overline{y}_N}} \mu([\overline{y}]).$$

现在陈述本节的主要结果：若X 是一个本原受限系统，且 μ 是非周期的（即，Y 中所有移位周期点的 μ-测度为零），那么，随着块长增加，本质覆盖半径$R_0(X,Y,\mu)$ 可由滑动块码构造的扩张近似得到。

定理 24 设X、$Y \subseteq \Sigma^{\mathbb{Z}}$ 为受限系统，且X 是本原的，$\mu \in M_{\mathcal{E}}(Y)$ 是无周期遍历测度。^那么，对于任意 $\varepsilon > 0$，存在充分大的N 以及长度为N 的滑块码函数f，使得

$$\mathbb{P}_{\nu_f}[\mathbf{X}_0 \neq \mathbf{Y}_0] - \varepsilon \leqslant R_0(X,Y,\mu) \leqslant \mathbb{P}_{\nu_f}[\mathbf{X}_0 \neq \mathbf{Y}_0].$$

定理 24 的证明包含两个主要部分。第一部分是证明可以通过平稳编码函数得到的扩张来逼近本质覆盖半径。第二步中，我们利用滑块码函数在平稳编码函数空间中稠密这一事实，得到主要结果。

证明第一部分需要用到 Alpern 引理 [2], 的一个改进版本，该版本最近在 [10]中得到了证明。

定义25 一个遍　　历测度 $\mu \in M_{\mathcal{E}}(Y)$ 若其周期点的测度为 0，则称其为非周期的。也就　　y,

$$\mu(\{\mathbf{y} \in Y \mid \exists n \in \mathbb{N} \text{ such that } T^n\mathbf{y} = \mathbf{y}\}) = 0.$$

引理 26（**[10,** 定理 **1]**）设$Y \subseteq \Sigma^{\mathbb{Z}}$ 为受限系统，$\mu \in M_{\mathcal{E}}(Y)$ 为无周期遍历测度，$n_1, \ldots\ldots, n_k$ 为两两互素的整数，且q_1 ……、q_k 为正数，满足$\sum_{i=1}^{k} q_i n_i = 1$。对于$Y$ 的任意有限可测划分 $\mathcal{P}$，存在可测集$Q_1, \ldots\ldots, Q_k$，使得：

1) 集合

$$\mathcal{P}' = \left\{ T^{-i}(Q_j) \,\middle|\, 1 \leqslant j \leqslant k, 0 \leqslant i \leqslant n_j - 1 \right\}$$

是Y的一个划分。$_j$2) 对所有 $1 \leqslant j \leqslant k$，$\mu(Q^j) = q^j$。3) 对所有 $1 \leqslant \not{\mu} \leqslant k$，$Q$　　$_j$ = μ　μ　$_j$ 相对于划分 $\mathcal{P}$ 是独立的（即，对于任意$A \in \mathcal{P}$ $\mu(A \cap Q)$ $(A)(Q)$）。

定理 27 设X,$Y \subseteq \Sigma^{\mathbb{Z}}$ 是　　受约束系统，X 是本原的，并且 $\mu \in M_{\mathcal{E}}(Y)$ 是一个无周期遍历测度。于是

$$R_0(X,Y,\mu) = \inf\left\{ \mathbb{P}_{\nu_g}[\mathbf{X}_0 \neq \mathbf{Y}_0] \,\middle|\, g : Y \to X \text{ is a stationary coding function} \right\}.$$

证明：该定理的证明相当复杂，因此我们先概述证明策略。根据定理 19，只需找到一个由平稳编码函数得到的扩张序列，使事件 $\{\mathbf{X}_0 \neq \mathbf{Y}_0\}$ 的概率收敛到 $R_0(X,\ Y,\ \mu)$。我们通过分块编码来构造这些平稳编码函数映射。利用 Alpern 引理，我们将来自Y 的一个无限序列划分成两种类型的块，一种较大，另一种较小。较大的块很常见，而较小的块则很少出现。我们将较大的块（以较高概率）编码为来自X语言的一个相近对应词，使得整个双无限序列最终属于X。这是可行的，因为X 是本原的。我们证明，采用这种分块编码，事件 $\{\mathbf{X}_0 \neq \mathbf{Y}_0\}$ 的概率收敛到 $R_0(X,\ Y)$。

现在开始证明。设 X 由本原标号图 G $=(V, E, L)$ 定义。我们固定一个任意小的 $\varepsilon \in (0, 1)$。由于 G 是本原的，存在一个数 p，使得对于任意两个顶点 $v_1, v_2 \in V$，都存在一条从 v_1 到 v_2、长度为 p 的有向路径。对于任意 $v_1, v_2 \in V$，我们固定这样一条长度为 p 的有向路径，并将其记为$\Gamma(v_1, v_2)$。给定图 $\Gamma = e_1, \ldots\ldots, e_m$ 上的任意有向路径，我们定义 $\sigma(\Gamma)$ $\sigma(e_1)$ 和 $\tau(\Gamma)$ $\tau(e_m)$。我们还定义 $L(\Gamma) \triangleq (L(e_1), \ldots\ldots, L(e_m)) \in \mathcal{B}_m(X)$，即通过读取 Γ 上的标签得到的词。

为简化记号，对于任意$M \in \mathbb{N}$，记$r_\varepsilon(M) \triangleq R_\varepsilon(\mathcal{B}_M(X), \mathcal{B}_M(Y), \mu_M)$。根据$R_\varepsilon(X, Y, \mu)$的定义，存在一个正整数序列$(M_i)_{i\in\mathbb{N}}$、$M_i \to \infty$以及一个数值趋于零的序列$(\delta'_i)_{i\in\mathbb{N}}$

$$\frac{r_\varepsilon(M_i)}{M_i} = R_\varepsilon(X, Y, \mu) + \delta'_i,$$

对所有$i \in \mathbb{N}$均成立。令$N_i \triangleq \lceil \frac{M_i}{p} \rceil$，容易观察到

$$r_\varepsilon(N_i p + 1) \leqslant r_\varepsilon(M_i) + p.$$

因此，

$$\begin{aligned}\frac{r_\varepsilon(N_i p+1)}{N_i p+1} &\leqslant \frac{r_\varepsilon(M_i)+p}{N_i p+1}\\ &= \frac{M_i}{N_i p+1}\cdot\frac{r_\varepsilon(M_i)}{M_i} + \frac{p}{N_i p+1}\\ &= R_\varepsilon(X, Y, \mu) + \delta_i,\end{aligned}$$

其中$(\delta_i)_{i\in\mathbb{N}}$ 是一个趋于零的序列。

任取$i \in \mathbb{N}$。根据$r_\varepsilon(n)$的定义，存在一个映射 $\varphi_i : \mathcal{B}_{N_i p+1}(Y) \to E^{N_i p+1}$ 以及一个集合$S_i \subseteq \mathcal{B}_{N_i p+1}(Y)$，其$p+1$-测度至少为 $1-\varepsilon$，使得对所有$y \in S_i$，$\varphi_i(y)$ 都是$\overline{G}$ 上的有向路径，并且

$$d(\overline{y}, L(\varphi_i(\overline{y}))) \leqslant r_\varepsilon(N_i p+1) \leqslant (N_i p+1)\cdot(R_\varepsilon(X, Y, \mu) + \delta_i). \tag{22}$$

我们称S_i 中的词为好词。

我们回顾一下，测度 μ 是非周期且遍历的，并注意到正整数 $n_1 = p$ 和 $n_2 = (N_i+1)p+1$ 互素。定义 $q_1 = {}^{1}p\, N_i$ 和 $q_2 = (1 - {}^{1}N_i)^{1}(N_i+1)p+1$，显然有 $q_1 n_1 + q_2 n_2 = 1$。我们考虑由 Y 的前 $(N_i+1)p+1$ 个坐标定义的划分。也就是说

$$\mathcal{P}^{(i)} \triangleq \left\{ [\overline{y}] \,\middle|\, \overline{y} \in \mathcal{B}_{(N_i+1)p+1}(Y) \right\},$$

这里我们回顾式 (1) 中柱集的定义。根据引理 26，存在可测集 $Q_1, Q_2 \subseteq Y$，使得集合

$$\mathcal{P}_i = \left\{ T^{-\ell}(Q_j) \,\middle|\, 0 \leqslant \ell \leqslant n_j - 1, \quad j = 1, 2 \right\}$$

是 Y 的一个划分；此外，对于 $j = 1, 2$，有 $\mu(Q^j) = q^j$，且 Q^j 与 $\mathcal{P}^{(i)}$ 独立。我们将使用划分 $\mathcal{P}_i$ 将一个双无限序列划分成若干块。

对于固定的 $y \in Y$ 和 $m \in \mathbb{Z}$，如果在坐标 m、m +1、……、m $+n_j-1$ 中，y 具有一个长度为 n^j、$j = 1$、2 的块，则称 y 允许一个块，记为 $T^m(y) \in Q^j$。我们枚举组成 y 的块，并将它们记为 $(B_k(y) = (m_k(y), l_k(y)))_{k\in\mathbb{Z}}$，其中 $B_0(y)$ 是包含 0 坐标的块，m_k 是第 k 个块的起点，而 $l_k \in \{n_1, n_2\}$ 是其长度。根据划分 $\mathcal{P}_i$ 的构造，$\mathbb{Z}$ 中的每个坐标恰好属于一个块。因此，为了定义一个函数 $Y \to X$，只需定义它在每个块中所取的值。

固定一个顶点 $v_0 \in V$。给定一个 $y \in Y$ 及其对应的块序列 $(B_k(y) = (m_k(y), l_k(y)))_{k\in\mathbb{Z}}$，我们定义 f_i 如下：

- 我们首先定义长度为f_i 的块$_1 = p$ 后接一个相同大小的块时的取值为自环$\Gamma(v_0, v_0)$上的标签。
- 接着定义长度为n_2的块上的取值：对于这样的块，我们对前$N_i p$ 个坐标应用变换$\varphi_i + 1$，并读取相应坐标上边的标签。对于块剩余的p 个坐标，我们有两种选择：

  1）如果下一个块的长度为$n_1 = p$，我们读取路径 $\Gamma(\tilde{v}, v_0)$ 上的标签，其中 $\tilde{v}$ 是与该块前$N_i p + 1$ 个坐标相对应的路径中的最后一个顶点。2）如果下一个块的长度为$n_2 = (N_i+1)p + 1$，我们读取路径 $\Gamma(\tilde{v}, v')$ 上的标签，其中v' 是与下一个块的$N_i p + 1$ 个首坐标相对应的路径中的初始顶点。

- 还需定义长度为n_1 的块后接长度为n_2的块时的取值。在这种情况下，我们读取路径 $\Gamma(v_0, v')$上的标签，其中 v' 与前一种情况相同。

总结如下：

$$f_i(\mathbf{y})_{m_k}^{m_k+l_k-1}=\begin{cases}L(\Gamma(v_0,v_0)) & l_k=l_{k+1}=n_1,\\ L\Big(\Gamma\Big(v_0,\sigma\Big(\varphi_i(\mathbf{y}_{m_{k+1}}^{m_{k+1}+N_ip})\Big)\Big)\Big) & l_k=n_1,l_{k+1}=n_2,\\ L\Big(\varphi_i(\mathbf{y}_{m_k}^{m_k+N_ip})\Big)L\Big(\Gamma\Big(\tau\Big(\varphi_i(\mathbf{y}_{m_k}^{m_k+N_ip})\Big),v_0\Big)\Big) & l_k=n_2,l_{k+1}=n_1,\\ L\Big(\varphi_i(\mathbf{y}_{m_k}^{m_k+N_ip})\Big)L\Big(\Gamma\Big(\tau\Big(\varphi_i(\mathbf{y}_{m_k}^{m_k+N_ip})\Big),\sigma\Big(\varphi_i(\mathbf{y}_{m_{k+1}}^{m_{k+1}+N_ip})\Big)\Big)\Big) & l_k=l_{k+1}=n_2.\end{cases}$$

根据f_i的构造，很容易验证，对于所有$\mathbf{y}$，$f_i(\mathbf{y})$ 都对应于G中的一条带标记的双向无限路径，因此 $\mathrm{Im}(f_i)\subseteq X$。我们注意到，由于 $\mathcal{P}_i$ 是一个可测划分，f_i 也是可测的。注意，对于任意$\mathbf{y}\in Y$，根据块划分的定义，$T(\mathbf{y})$ 的块划分是$\mathbf{y}$的块划分的移位版本。因此，f_i 也与移位可交换，因而它是一个平稳编码函数。令ν_{f_i}为 μ 在变换(f_i,Id)下的推前测度。

现在我们来证明，$\mathbb{P}_{\nu_{f_i}}[\mathbf{X}_0\neq\mathbf{Y}_0]$ 确实逼近$R_0(X,Y,\mu)$。设$\mathbf{X}$和$\mathbf{Y}$分别表示在 $\Sigma^{\mathbb{Z}}$ 中相对于ν_{f_i}生成的随机双无限序列，并令$I_{\{\mathbf{X}_0\neq\mathbf{Y}_0\}}$为事件 $\{\mathbf{X}_0\neq\mathbf{Y}_0\}$的示性函数。由于$\nu_{f_i}$是不变测度，对于所有$n\in\mathbb{N}$

$$\mathbb{P}_{\nu_{f_i}}[\mathbf{X}_0\neq\mathbf{Y}_0]=\int I_{\{\mathbf{X}_0\neq\mathbf{Y}_0\}}\cdot d\nu_{f_i}=\int\frac{1}{n}\sum_{k=0}^{n-1}I_{\{\mathbf{X}_0\neq\mathbf{Y}_0\}}\circ T^k\cdot d\nu_{f_i}=\mathbb{E}_{\nu_{f_i}}\left[\frac{1}{n}d\left(\mathbf{X}_0^{n-1},\mathbf{Y}_0^{n-1}\right)\right].$$

令 C_n 为随机变量，用来统计坐标 0、1、.....、$\mathrm{n}-1$ 中良块的数量。这些块的长度为 $\mathrm{n}_2=(\mathrm{N}_i+1)\mathrm{p}+1$，包含在 Y_0^{n-1} 中，且前 $\mathrm{N}_i\mathrm{p}+1$ 个坐标包含一个良词，即集合 S_i 中的一个词。形式上，对于序列 $\mathrm{y}\in\mathrm{Y}$ 及其对应的块 $(\mathrm{B}_k(\mathrm{y})=(\mathrm{m}_k(\mathrm{y}),\mathrm{l}_k(\mathrm{y}))_{k\in\mathbb{Z}}$

$$C_n(\mathbf{y})\triangleq\left|\left\{k\in\mathbb{Z}\,\middle|\,l_k=n_2,[m_k,m_k+n_2-1]\subseteq[0,n-1]\text{ and }\mathbf{y}_{m_k}^{m_k+N_ip}\in S_i\right\}\right|.$$

根据f_i,φ_i 和 S_i 的构造，以及式 (22)，单个良块中$\mathbf{X}$ 与$\mathbf{Y}$ 不相同的坐标数量上界为$(N_ip+1)(R_\varepsilon(X、Y、\mu)+\delta_i)+p$。因此，我们有

$$\begin{aligned}d\left(\mathbf{X}_0^{n-1},\mathbf{Y}_0^{n-1}\right)&\leqslant C_n(\mathbf{Y}_0^{n-1})\cdot((N_ip+1)(R_\varepsilon(X,Y,\mu)+\delta_i)+p)+n-C_n(\mathbf{Y}_0^{n-1})((N_i+1)p+1)\\&=C_n(\mathbf{Y}_0^{n-1})\cdot(N_ip+1)(R_\varepsilon(X,Y,\mu)+\delta_i-1)+n\end{aligned}$$

因此，

$$\mathbb{E}_{\nu_{f_i}}\left[d\left(\mathbf{X}_0^{n-1},\mathbf{Y}_0^{n-1}\right)\right]\leqslant\mathbb{E}_{\nu_{f_i}}[C_n(\mathbf{Y}_0^{n-1})]\cdot(N_ip+1)(R_\varepsilon(X,Y,\mu)+\delta_i-1)+n.\tag{23}$$

令A 表示一个良块从坐标 0 开始这一事件。也就是说，

$$A=Q_2\cap\left\{\mathbf{Y}_0^{N_ip}\in S_i\right\}.$$

注意，我们可以写成 $C_n(\mathbf{Y}_0^{n-1})=\sum^{n-n_2}{}_{k=0}I_A\circ T^i$，其中$I_A$ 是事件A的示性函数。因此，由于 ν_{f_i} 是移位不变的，且 Q_2 与前面$(N_i+1)p+1$ 个坐标独立，我们有

$$\begin{aligned}\mathbb{E}_{\nu_{f_i}}[C_n(\mathbf{Y}_0^{n-1})]&=(n-n_2)\mathbb{P}_\mu[A]=(n-n_2)\mathbb{P}_\mu\left[Q_2\cap\left\{\mathbf{Y}_0^{N_ip}\in S_i\right\}\right]\\&=(n-n_2)\mathbb{P}_\mu[Q_2]\mathbb{P}_\mu\left[\mathbf{Y}_0^{N_ip}\in S_i\right]\\&\geqslant(n-n_2)\left(\frac{1}{n_2}\left(1-\frac{1}{N_i}\right)\right)(1-\varepsilon).\end{aligned}$$

由于 δ_i趋于零，对于所有充分大的i（与n无关），我们有$R_\varepsilon(X、Y、\mu)+\delta_i-1<0$。因此，将 $\mathbb{E}_{\nu_{f_i}}[C_n(\mathbf{Y}_0^{n-1})]$ 的下界与 (23) 结合，我们得到

$$\begin{aligned}\mathbb{P}_{\nu_{f_i}}[\mathbf{X}_0\neq\mathbf{Y}_0]&=\mathbb{E}_{\nu_{f_i}}\left[\frac{1}{n}d\left(\mathbf{X}_0^{n-1},\mathbf{Y}_0^{n-1}\right)\right]\\&\leqslant\frac{n-n_2}{n}\cdot\frac{1}{n_2}\left(1-\frac{1}{N_i}\right)(1-\varepsilon)(N_ip+1)(R_\varepsilon(X,Y,\mu)+\delta_i-1)+1.\end{aligned}$$

令$n\to\infty$ 趋于极限，我们得到

$$\mathbb{P}_{\nu_{f_i}}[\mathbf{X}_0\neq\mathbf{Y}_0]\leqslant\frac{1}{(N_i+1)p+1}\left(1-\frac{1}{N_i}\right)(1-\varepsilon)(N_ip+1)(R_\varepsilon(X,Y,\mu)+\delta_i-1)+1.$$

回顾一下，$N_i \to \infty$ 和 $\delta_i \to 0$ 当$i \to \infty$时成立；再次令$i \to \infty$ 趋于极限，我们得到

$$\begin{aligned}
&\inf\left\{\mathbb{P}_{\nu_g}[\mathbf{X}_0 \neq \mathbf{Y}_0] \,\middle|\, g : Y \to X \text{ is a stationary coding function}\right\} \\
&\leqslant \lim_{i\to\infty} \mathbb{P}_{\nu_{f_i}}[\mathbf{X}_0 \neq \mathbf{Y}_0] \\
&\leqslant \lim_{i\to\infty} \frac{1}{(N_i+1)p+1}\left(1-\frac{1}{N_i}\right)(1-\varepsilon)(N_i p+1)(R_\varepsilon(X,Y,\mu)+\delta_i-1)+1 \\
&= (1-\varepsilon)R_\varepsilon(X,Y,\mu)+\varepsilon.
\end{aligned}$$

取 $\varepsilon \to 0$,

$$\inf\left\{\mathbb{P}_{\nu_g}[\mathbf{X}_0 \neq \mathbf{Y}_0] \,\middle|\, g : Y \to X \text{ is a stationary coding function}\right\} \leqslant R_0(X,Y,\mu).$$

■

现在我们已经完成了定理 24 证明的第一部分。证明的第二部分需要用到一个众所周知的事实：平稳编码函数可以用滑动块码逼近。

引理 28（**[18,** 定理 3.1**]**）设 X、$Y \subseteq \Sigma^{\mathbb{Z}}$ 为移位空间，$\mu \in M(Y)$，且 $g : Y \to X$ 为平稳编码函数。ˆ那么，对任意 $\varepsilon > 0$，存在一个滑块码函数 $f : Y \to X$，使得

$$\mathbb{P}_\mu\left[g(\mathbf{Y}) \neq \hat{f}(\mathbf{Y})\right] < \varepsilon.$$

现在我们可以证明定理 24 了。

定理 24 的证明：固定 $\varepsilon > 0$。根据定理 27，存在一个平稳编码函数 $g : Y \to X$，使得

$$\mathbb{P}_\mu[g(\mathbf{Y})_0 \neq \mathbf{Y}_0] = \mathbb{P}_{\nu_g}[\mathbf{X}_0 \neq \mathbf{Y}_0] < R_0(X,Y,\mu) + \frac{1}{2}\varepsilon.$$

根据引理 28，存在一个滑块码函数 ˆ$f : Y \to X$ 以及一个集合 E，满足 $\mu(E) > 1 - \frac{1}{2}\varepsilon$，使得 g 在 E上与 ˆf一致。现在我们得到,

$$\begin{aligned}
\mathbb{P}_{\nu_f}[\mathbf{X}_0 \neq \mathbf{Y}_0] &= \mathbb{P}_\mu[\mathbf{Y}_0 \neq \hat{f}(\mathbf{Y})_0] \\
&= \mathbb{P}_\mu[\{\mathbf{Y}_0 \neq \hat{f}(\mathbf{Y})_0\} \cap E] + \mathbb{P}_\mu[\{\mathbf{Y}_0 \neq \hat{f}(\mathbf{Y})_0\} \cap E^C] \\
&\leqslant \mathbb{P}_\mu[g(\mathbf{Y})_0 \neq \mathbf{Y}_0] + \mathbb{P}_\mu[E^C] < R_0(X,Y,\mu) + \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = R_0(X,Y,\mu) + \varepsilon.
\end{aligned}$$

由命题 16 可立即得出上界 $R_0(X,Y,\mu) \leqslant \mathbb{P}_{\nu_f}[\mathbf{X}_0 \neq \mathbf{Y}_0]$。 ■

VI. 结论

本文提出了量化约束串接（QCC）方案，为在受约束系统中实现纠错提供了一个新的通用框架。我们证明，通过量化将纠错码的码字嵌入受约束系统中，可以纠正 $\Theta(n)$ 个错误（相对于该码的长度 n）。我们发现，对于给定的受约束系统，本方法的渐近纠错能力由该受约束系统的一个新基本参数——其覆盖半径——决定，该参数限定了量化过程所引入的噪声量。与串接和反向串接等以往方法不同，嵌入受约束系统的过程不可逆，因此称为量化。

我们给出了受约束系统覆盖半径的两种不同定义，将其分别视为组合概念和概率概念。组合定义考虑最坏情况（深洞）；而在概率方法中，本质覆盖半径忽略罕见情况，因此允许覆盖半径更小。我们研究了本质覆盖半径和组合覆盖半径的性质，并给出了通用的下界和上界。

尽管受限系统的覆盖半径本身具有独立的学术价值，让我们将我们的结果置于量子校验码方案的背景下。考虑 $X_{0,k}$，即例 1 中描述的$(0,k)$-受限游程长度系统。使用 [17,定理 1], 中提出的编码方案，可以纠正多达$O(\sqrt{n})$ 个错误。然而，使用组合覆盖半径为 $^{1}_{k+1}$的量子校验码（在这种情况下），由于存在码率非零且能够纠正多达$(\frac{1}{4}-\delta)n$ 个错误的纠错码（对每个 $\delta > 0$ 均成立），我们可以得到码率非零且能够纠正多达$(\frac{1}{4} - {}^{1}_{k+1} - \delta)n$ 个信道错误的码。另一方面，我们可以使用$X_{0,k}$ 的本质覆盖半径来界定可能出现的量化噪声。在这种情况下，由于

$$R_{\frac{1}{2}}(X_{0,k}, [2]^{\mathbb{Z}}, \mu^{\mathrm{u}}) \leqslant R_0(X_{0,k}, [2]^{\mathbb{Z}}, \mu^{\mathrm{u}}) = \frac{1}{2(2^{k+1}-1)},$$

利用量子校验码，可以找到这样的纠错码：至多移除一半的码字（从渐近角度看，这不会影响码率），便可将纠错能力提升至 $(\frac{1}{4} - {}^{1}_{2(2^{k+1}-1)} - \delta)n$

信道错误。此前的研究通过某些非构造性方法，建立了纠错约束码可达到码率的下界 [21], [26], 。我们的方案具有简单且具构造性的优点。我们的方案的优势在于其简单性和构造性。

参考文献

[1] R. Adler、D. Coppersmith 和 M. Hassner，“滑动区块码算法——符号动力系统在信息论中的应用”，《IEEE 信息论汇刊》，第 29 卷第 1 期，第 5–22 页，1983 年。[2] S. Alpern，“保测同胚的典型性质”，载于《遍历理论》。Springer，1979 年，第 16–27 页。[3] J. Ashley 和 B. Marcus，“滑动区块译码器的规范编码器”，SIAM 离散数学杂志，第 8 卷第 4 期，第 555–605 页，1995 年。[4] J. J. Ashley、R. Karabed 和 P. H. Siegel，“复杂度与滑动分组可译码性”，《IEEE 信息论汇刊》，第 42 卷第 6 期，第 1925–1947 页，1996 年。[5] K. G. Benerjee 和 A. Banerjee，“关于避免均聚物和二级结构的可逆、可逆互补及 GC 平衡 DNA 码”，载于《2022 年 IEEE 信息论国际研讨会（ISIT）》。IEEE，2022 年，第 204–209 页。[6] G. D. Birkhoff，“遍历定理的证明”，《美国国家科学院院刊》，第 17 卷第 12 期，第 656–660 页，1931 年 12 月。[7] W. Bliss，“用于对基带编码数据执行纠错计算以消除错误传播的电路”，IBM 技术披露公报，第 23 卷，第 4633–4634 页，1981 年。[8] K. Cai、Y. M. Chee、R. Gabrys、H. M. Kiah 和 T. T. Nguyen，“纠正基于 DNA 的数据存储中的单次插入删除/编辑：线性时间编码器与阶最优性”，《IEEE 信息论汇刊》，第 67 卷第 6 期，第 3438–3451 页，2021 年。[9] K. Cai、H. M. Kiah、M. Motani 和 T. T. Nguyen，“具有局部重量约束的分段编辑编码”，载于《2021 年 IEEE 信息论国际研讨会（ISIT）》。IEEE，2021 年，第 1694–1699 页。[10] J. T. Campbell、R. McCutcheon 和 A. Windsor，“ 独立性与 Alpern 多塔 ”，《动力系统》，第 34 卷第 2 期，第 239–243 页，2019 年。[11] Y. M. Chee 、 H. M. Kiah 和 H. Wei，“高效且显式的平衡引物码”，《IEEE 信息论汇刊》，第 66 卷第 9 期，第 5344–5357 页，2020 年。[12] G. Cohen、I. Honkala、S. Litsyn 和 A. Lobstein，《覆盖码》。北荷兰出版社，1997 年。[13] S. Dani，“遍历理论：着眼于数论。Manfred Einsiedler、Thomas Ward 著。施普林格出版社，伦敦，2011 年”，《遍历理论与动力系统》，第 32 卷第 3 期，第 1157–1160 页，2012 年。[14] N. De Bruijn 和 P. Erd¨os，“一些线性递推公式和一些二次递推公式。I”，Indag. Math.(NS)，第 54 卷第 5 期，第 374–382 页，1951 年。[15] J. L. Fan 和 A. R. Calderbank，“一种改进的级联编码方案及其在磁数据存储中的应用”，《IEEE 信息论汇刊》，第 44 卷第 4 期，第 1565–1574 页，1998 年。[16] Z. F¨uredi 和 I. Z. Ruzsa，“近似次可加序列”，《匈牙利数学学报》，第 161 卷第 2 期，第 401–411 页，2020 年。[17] R. Gabrys、P. H. Siegel 和 E. Yaakobi，“分段反向串接：一种约束纠错码的新方法”，载于《2020 年信息理论及其应用国际研讨会（ISITA）》。IEEE，2020 年，第 254–258 页。[18] R. Gray，“滑动分组信源编码”，《IEEE 信息论汇刊》，第 21 卷第 4 期，第 357–368 页，1975 年。[19] V. Guruswami、A. Rudra 和 M. Sudan，《编码理论基础》，2022 年。[20] K. A. S. Immink，“一种逼近受约束信道信道容量的实用方法”，《IEEE 信息论汇刊》，第 43 卷第 5 期，第 1389–1399 页，1997 年。[21] V. D. Kolesnik 和 V. Y. Krachkovsky，“受限纠错码的生成函数与码率下界”，《IEEE 信息论汇刊》，第 37 卷第 3 期，第 778–788 页，1991 年。[22] X. Li、M. Chen 和 H. Wu，“用于基于 DNA 的数据存储、具有 GC 平衡且无均聚物的位置受限 DNA 序列多重错误纠正”，《生物信息学简报》，2022 年。[23] D. Lind 和 B. Marcus，《符号动力系统与编码导论》。剑桥大学出版社，2021 年。[24] X. Lu 和 S. Kim，“具有最大游程长度约束、用于纠正单次插入或删除错误的非二元纠错码设计，应用于 DNA 存储”，IEEE Access，第 9 卷，第 135 354–135 363 页，2021 年。[25] M. Mansuripur，“具有任意约束的枚举调制编码及用于数据存储系统的调制后纠错编码”，载于《光学数据存储 ’91》，第 1499 卷。SPIE，1991 年，第 72–86 页。[26] B. H. Marcus 和 R. M. Roth，“受约束系统的改进 Gilbert–Varshamov 界”，《IEEE 信息论汇刊》，第 38 卷第 4 期，第 1213–1221 页，1992 年。[27] B. H. Marcus、R. M. Roth 和 P. H. Siegel，“约束系统编码导论”，讲义，2001 年。[28] J. V. Neumann，“拟遍历假设的证明”，《美国国家科学院院刊》，第 18 卷第 1 期，第 70–82 页，1932 年 1 月。[29] T. T. Nguyen、K. Cai、K. A. S. Immink 和 H. M. Kiah，“用于基于 DNA 的数据存储、具有错误控制的约束编码”，载于《2020 年 IEEE 信息论国际研讨会（ISIT）》。IEEE，2020 年，第 694–699 页。[30] ——，“用于基于 DNA 的数据存储、逼近容量且具有纠错功能的受约束码”，《IEEE 信息论汇刊》，第 67 卷第 8 期，第 5602–5613 页，2021 年。[31] W. H. Press、J. A. Hawkins、S. K. Jones Jr、J. M. Schaub 和 I. J. Finkelstein，“用于 DNA 存储的 Hedges 纠错码可纠正插入删除错误并允许序列约束”，《美国国家科学院院刊》，第 117 卷第 31 期，第 18 489–18 496 页，2020 年。[32] D. J. Rudolph 等，《可测动力学基础：Lebesgue 空间上的遍历理论》。牛津大学出版社（美国），1990 年。[33] P. Walters，《遍历理论导论》。施普林格科学与商业媒体，2000 年，第 79 卷。[34] J. H. Weber、J. A. De Groot 和 C. J. Van Leeuwen，“关于用于基于 DNA 的数据存储的单错误检测码”，《IEEE 通信快报》，第 25 卷第 1 期，第 41–44 页，2020 年。