Uniquely Decodable Multiple Access Source Codes [1]

Qian Zhao
Oracle Corporation
email: qianz@z.caltech.edu

Michelle Effros
California Institute of Technology
email: effros@caltech.edu

Abstract — **The Slepian-Wolf bound raises interest in lossless code design for multiple access networks. Previous work treats instantaneous codes. We generalize the Sardinas and Patterson test and bound the achievable rate region for uniquely decodable codes.**

I. Introduction

A lossless multiple access source code (MASC) for i.i.d. samples from p.m.f. $p(x,y)$ on alphabet $\mathcal{X}\times\mathcal{Y}$ comprises two encoders $\gamma_X:\mathcal{X}\to\{0,1\}^\star$ and $\gamma_Y:\mathcal{Y}\to\{0,1\}^\star$ and a decoder $\gamma^{-1}:\{0,1\}^\star\times\{0,1\}^\star\to\mathcal{X}\times\mathcal{Y}$ for which $P_e=\Pr(\gamma^{-1}(\gamma_X(X),\gamma_Y(Y))\neq(X,Y))\equiv 0$. Similarly, a lossless side information source code (SISC) on X given Y comprises encoder $\gamma_X:\mathcal{X}\to\{0,1\}^\star$ and decoder $\gamma^{-1}:\{0,1\}^\star\times\mathcal{Y}\to\mathcal{X}$ such that $P_e=\Pr(\gamma^{-1}(\gamma_X(X),Y)\neq X)\equiv 0$.

Prior work focuses on instantaneous (INST) codes. We study the broader class of uniquely decodable (UD) codes. A UD-MASC is an MASC such that given any $k,m\geq 1$, $(x^k,y^k)\in\mathcal{X}^k\times\mathcal{Y}^k$, and $(\hat{x}^m,\hat{y}^m)\in\mathcal{X}^m\times\mathcal{Y}^m$, if $(x^k,y^k)\neq(\hat{x}^m,\hat{y}^m)$ and $p(x^k,y^k)p(\hat{x}^m,\hat{y}^m)>0$, then

$$(\gamma_X(x^k),\gamma_Y(y^k))\neq(\gamma_X(\hat{x}^m),\gamma_Y(\hat{y}^m)).$$

(We treat i.i.d. samples and use the extension code; thus $p(x^n,y^n)=\prod_{i=1}^n p(x_i,y_i)$, $\gamma_X(x^n)=\gamma_X(x_1)\ldots\gamma_X(x_n)$, and $\gamma_Y(y^n)=\gamma_Y(y_1)\ldots\gamma_Y(y_n)$.) A UD-SISC is defined similarly. We prove necessary and sufficient conditions for unique decodability and bound the UD-SISC achievable rate region.

II. Main Results

Theorem 1 generalizes the Sardinas-Patterson test [1].

Theorem 1 *(γ_X,γ_Y) is a UD-MASC iff it passes UD-TEST.*

Let $C_X=\{\gamma_X(x):y\in\mathcal{X}\}$ and $C_Y=\{\gamma_Y(y):y\in\mathcal{Y}\}$. Define $p(\mathbf{c}_x,\mathbf{c}_y)$ to be the probability of all (x,y) with descriptions $(\mathbf{c}_x,\mathbf{c}_y)$. Let '$\prec$' indicate a prefix and '$+$ and '$-$' denote concatenation and suffix operators. Thus if string $\mathbf{s}$ is string $\mathbf{s}_1$ followed by string $\mathbf{s}_2$, then we write $\mathbf{s}=\mathbf{s}_1+\mathbf{s}_2$, $\mathbf{s}_2=\mathbf{s}-\mathbf{s}_1$, and $\mathbf{s}_1\prec\mathbf{s}$. Finally, given sets C and S, define $\mathcal{S}(C,S)$ as the smallest set such that for any $\mathbf{s}\in S$ and $\mathbf{c},\mathbf{c}'\in C$,

(a) if $\mathbf{s}\prec\mathbf{c}$ and $\mathbf{c}-\mathbf{s}\prec\mathbf{c}'$ then $\mathbf{s}'=\mathbf{c}'-(\mathbf{c}-\mathbf{s})\in\mathcal{S}(S,C)$;
(b) if $\mathbf{s}\prec\mathbf{c}$ and $\mathbf{c}'\prec\mathbf{c}-\mathbf{s}$, then $\mathbf{s}'=(\mathbf{c}-\mathbf{s})-\mathbf{c}'\in\mathcal{S}(S,C)$;
(c) if $\mathbf{c}\prec\mathbf{s}$, then $\mathbf{s}'=(\mathbf{s}-\mathbf{c})+\mathbf{c}'\in\mathcal{S}(S,C)$.

UD-TEST

1. Set $i=1$. Define $\mathcal{S}_{X1}=\{\mathbf{c}_x-\mathbf{c}'_x:\mathbf{c}_x,\mathbf{c}'_x\in C_X,\mathbf{c}'_x\prec\mathbf{c}_x\}$ and $\mathcal{S}_{Y1}=\{\mathbf{c}_y-\mathbf{c}'_y:\mathbf{c}_y,\mathbf{c}'_y\in C_Y,\mathbf{c}'_y\prec\mathbf{c}_y\}$. Let E_1 be the set of pairs $(\mathbf{c}_x-\mathbf{c}'_x,\mathbf{c}_y-\mathbf{c}'_y)\in\mathcal{S}_{X1}\times\mathcal{S}_{Y1}$ such that the codewords $\mathbf{c}_x,\mathbf{c}'_x\in C_X$ and $\mathbf{c}_y,\mathbf{c}'_y\in C_Y$ satisfy $p(\mathbf{c}_x,\mathbf{c}_y)p(\mathbf{c}'_x,\mathbf{c}'_y)+p(\mathbf{c}_x,\mathbf{c}'_y)p(\mathbf{c}'_x,\mathbf{c}_y)>0$.

2. If $\exists(\mathbf{s}_x,\mathbf{s}_y)\in E_i\cap C_X\times C_Y$, then by tracking backwards to find the codewords that led to the construction of s_x and s_y we can find strings $(x^k,y^\ell)\neq(x^m,y^n)$ such that $(\gamma_X(x^k),\gamma_Y(y^\ell))=(\gamma_X(x^m),\gamma_Y(y^n))$. If any such string satisfies $k=\ell$, $m=n$, and $p(x^k,y^\ell)p(x^m,y^n)>0$, then (γ_X,γ_Y) fails the test and the procedure stops.

3. Initialize $\mathcal{S}_{X(i+1)}$ and $\mathcal{S}_{Y(i+1)}$ as $\mathcal{S}_{X(i+1)}=\mathcal{S}(C_X,\mathcal{S}_{Xi})$ and $\mathcal{S}_{Y(i+1)}=\mathcal{S}(C_Y,\mathcal{S}_{Yi})$. Let E_{i+1} be the set of pairs $(\mathbf{s}'_x,\mathbf{s}'_y)\in\mathcal{S}_{X(i+1)}\times\mathcal{S}_{Y(i+1)}$ such that if codewords $\mathbf{c}_x,\mathbf{c}'_x\in C_X$ and suffix $\mathbf{s}_x\in\mathcal{S}_{Xi}$ created $\mathbf{s}'_x$ through $\mathcal{S}(C_X,\mathcal{S}_{Xi})$ and codewords $\mathbf{c}_y,\mathbf{c}'_y\in C_X$ and suffix $\mathbf{s}_y\in\mathcal{S}_{Yi}$ created $\mathbf{s}'_y$ through $\mathcal{S}(C_Y,\mathcal{S}_{Yi})$, then $(\mathbf{s}_x,\mathbf{s}_y)\in E_i$ and $p(\mathbf{c}_x,\mathbf{c}_y)p(\mathbf{c}'_x,\mathbf{c}'_y)+p(\mathbf{c}_x,\mathbf{c}'_y)p(\mathbf{c}'_x,\mathbf{c}_y)>0$. Update $\mathcal{S}_{X(i+1)}$ and $\mathcal{S}_{Y(i+1)}$ by removing from $\mathcal{S}_{X(i+1)}$ and $\mathcal{S}_{Y(i+1)}$ the largest subsets B_X and B_Y, respectively, such that $(B_X\times\mathcal{S}_{Y(i+1)})\cap E_{i+1}=\phi$ and $(\mathcal{S}_{X(i+1)}\times B_y)\cap E_{i+1}=\phi$.

4. If $(\mathcal{S}_{X(i+1)},\mathcal{S}_{Y(i+1)},E_{i+1})=(\mathcal{S}_{Xj},\mathcal{S}_{Yj},E_j)$ for some $j\leq i$ or $\mathcal{S}_{X(i+1)}=\phi$ or $\mathcal{S}_{Y(i+1)}=\phi$, then (γ_X,γ_Y) passes the test and the procedure stops. Otherwise, set $i=i+1$ and go to step 2. □

We generalize the Kraft Inequality to give necessary conditions on the codeword lengths for UD-SISCs and obtain lower bounds on the achievable rates for lossless UD-SISCs. Let $\mathcal{A}_y=\{x\in\mathcal{X}:p(x,y)>0\}$, $\Gamma_y=\{\gamma_X(x):x\in\mathcal{A}_y\}$.

Theorem 2 *For any UD-SISC on X given Y, $(\Gamma_a\cap\Gamma_b)\cup(\Gamma_a\cap\Gamma_c)\cup(\Gamma_b\cap\Gamma_c)$ is UD for every $\{a,b,c\}\subseteq\mathcal{Y}$. For any INST-SISC on X given Y, $(\Gamma_a\cap\Gamma_b)\cup(\Gamma_a\cap\Gamma_c)\cup(\Gamma_b\cap\Gamma_c)$ is prefix free for every $\{a,b,c\}\subseteq\mathcal{Y}$.*

Corollary 1 *For any lossless SISC on X given Y, $\{a,b,c\}\subseteq\mathcal{Y}$ implies $\sum_{\mathbf{c}\in(\Gamma_a\cap\Gamma_b)\cup(\Gamma_a\cap\Gamma_c)\cup(\Gamma_b\cap\Gamma_c)}2^{-|\mathbf{c}|}\leq 1$.*

Define $\mathcal{A}_{\bar{1}}=\mathcal{A}_1^c$, $\mathcal{A}_{\bar{2}}=\mathcal{A}_2^c$, $\mathcal{A}_{\bar{3}}=\mathcal{A}_3^c$, and for all $i\in\{1,\bar{1}\}$, $j\in\{2,\bar{2}\}$, $k\in\{3,\bar{3}\}$ define $\mathcal{A}_{ij}=\mathcal{A}_i\cap\mathcal{A}_j$, $\mathcal{A}_{ijk}=\mathcal{A}_i\cap\mathcal{A}_j\cap\mathcal{A}_k$, $P_{ij}=\sum_{x\in\mathcal{A}_{ij}}p(x)$, and $P_{ijk}=\sum_{x\in\mathcal{A}_{ijk}}p(x)$. Let $h(p)=-p\log p-(1-p)\log(1-p)$.

Theorem 3 *For $|\mathcal{Y}|=2$, the optimal rate $R(X)$ for a one-dimensional lossless SISC on X given Y satisfies*

$$R^\star(X)\leq R(X)<R^\star(X)+1,$$

where $R^\star(X)=H(X)-(P_{1\bar{2}}+P_{\bar{1}2})h(\frac{P_{1\bar{2}}}{P_{1\bar{2}}+P_{\bar{1}2}})\geq H(X|Y)$ with equality if and only if $\frac{\sum_{x\in\mathcal{A}_{12}}p(x,1)}{\sum_{x\in\mathcal{A}_{1\bar{2}}}p(x,1)}=\frac{\sum_{x\in\mathcal{A}_{12}}p(x,2)}{\sum_{x\in\mathcal{A}_{\bar{1}2}}p(x,2)}$ and $\frac{p(x,1)}{p(x,2)}$ is a constant for all $x\in\mathcal{A}_{12}$.

Theorem 4 *For $|\mathcal{Y}|=3$, the optimal rate $R(X)$ for a one-dimensional lossless SISC on X given Y satisfies $R(X)\geq H(X)-(P_{1\bar{2}\bar{3}}+P_{\bar{1}2\bar{3}})h(\frac{P_{1\bar{2}\bar{3}}}{P_{1\bar{2}\bar{3}}+P_{\bar{1}2\bar{3}}})-(P_{\bar{1}2\bar{3}}+P_{1\bar{2}\bar{3}})h(\frac{P_{\bar{1}2\bar{3}}}{P_{\bar{1}2\bar{3}}+P_{1\bar{2}\bar{3}}})-(P_{\bar{1}\bar{2}3}+P_{12\bar{3}})h(\frac{P_{\bar{1}\bar{2}3}}{P_{\bar{1}\bar{2}3}+P_{12\bar{3}}})$.*

References

[1] A. A. Sardinas, G. W. Patterson, "A necessary and sufficient condition for the unique decomposition of coded messages", *IRE Convention Record*, pt.8, pp.104–108,1956.

[1] This material is based upon work supported by NSF Award No. CCR-0220039, the Caltech Lee Center for Advanced Networking, and the Howard Oringer Fellowship.