1 Notations and preliminaries

The functions \( \log \) and \( \ln \) denote respectively the base 2 and the natural logarithms. By convention, \( 0 \log 0 = 0 \ln 0 \). For \( x \in \mathbb{R} \), \( \lfloor x \rfloor \) (resp. \( \lceil x \rceil \)) denotes the greatest (resp. smallest) integer not greater (resp. not smaller) than \( x \). For integers \( a \leq b \), \( [a, b] \) denotes the set of integers between \( a \) and \( b \), bounds included. Let \( \mathbb{D} = \left\{ 2^\mu : \mu \in \mathbb{N} \right\} \).

Let \((X, Y) \simeq P_{X,Y}\) be an arbitrary pair of random variables over \( \mathcal{B} \times \mathcal{Y} \) with \( \mathcal{B} = \{0, 1\} \) and \( \mathcal{Y} \) an arbitrary countable set. We regard \((X, Y)\) as a memoryless source \( S \), with \( X \) as the part to be compressed and \( Y \) in the role of “side-information” about \( X \). We consider a sequence \( S = \{(X_i, Y_i) : i \in \mathbb{N}^*\} \) of independent drawings from \((X, Y)\) – which can be interpreted as a representation of the source \( S \) – and we introduce the two transformations \( \ominus \) and \( \oplus \) applied to the source \( S \) and defined by

\[
S^- = \left\{ (X_{2i-1} \oplus X_{2i}, Y_{2i-1}, Y_{2i}) : i \in \mathbb{N}^* \right\} \quad (1)
\]

\[
S^+ = \left\{ (X_{2i}, Y_{2i-1}, Y_{2i} \oplus X_{2i}) : i \in \mathbb{N}^* \right\}. \quad (2)
\]

With these notations, \( S^- \) (resp. \( S^+ \)) is the memoryless source that takes its values in \( \mathcal{B} \times \left( \mathcal{Y}^2 \right) \) (resp. in \( \mathcal{B} \times (\mathcal{Y}^2 \times \mathcal{B}) \)), with \( X_1 \oplus X_2 \) (resp. \( X_2 \)) as the part to be compressed and \( (Y_1, Y_2) \) (resp. \( (Y_1, Y_2, X_1 \oplus X_2) \)) in the role of “side-information”.

The process that constructs \( S^- \) and \( S^+ \) from \( S \) can be written \( S_0^{(0)} = S \),

\[
S_1^{(0)} = \left( S_0^{(0)} \right)^- \quad \text{and} \quad S_1^{(1)} = \left( S_0^{(0)} \right)^+ = S^+. \quad (3)
\]

Applied recursively, this process leads to the sequence of memoryless sources \( S^{(i)}_{\mu} \) \( \mu \in \mathbb{N}, i \in [0, 2^\mu - 1] \), where \( S^{(i)}_{\mu} \) takes its values in a set \( \mathcal{B} \times (\mathcal{Y}^{2^\mu} \times \mathcal{B}^{K(i)}) \) with \( K(i) \in [0, 2^\mu - 1] \) and is defined by

\[
S^{(i)}_{\mu+1} = \begin{cases} 
(S^{[i/2])}_{\mu} \ominus \text{ if } i \text{ is even} \\
(S^{[i/2])}_{\mu} \oplus \text{ if } i \text{ is odd}.
\end{cases} \quad (4)
\]

Let us introduce the sources’ conditional entropies expressed in bits:

\[
H(S) = H(X_1 | Y_1) = H(X_2 | Y_2), \quad (5)
\]

\[
H(S^-) = H(X_1 \oplus X_2 | Y_1, Y_2), \quad (6)
\]

\[
H(S^+) = H(X_2 | Y_1, Y_2, X_1 \oplus X_2). \quad (7)
\]

For any \( m \in \mathbb{D} \) \( (m = 2^\mu \) with \( \mu \in \mathbb{N} \) and for any \( \theta \in \left[0, \frac{1}{2}\right) \), let

\[
\mathcal{H}_{X|Y} = \mathcal{H}_{X|Y}(\theta) = \mathcal{H}^{(m)}_{X|Y}(\theta) = \left\{ i \in [0, m-1] : H(S^{(i)}_{\mu}) > \theta \right\} \quad (8)
\]

\[
\mathcal{V}_{X|Y} = \mathcal{V}_{X|Y}(\theta) = \mathcal{V}^{(m)}_{X|Y}(\theta) = \left\{ i \in [0, m-1] : H(S^{(i)}_{\mu}) > 1 - \theta \right\}. \quad (9)
\]

For any memoryless source \( S = (X, Y) \simeq P_{X,Y} \), we introduce its Bhattacharyya parameter:

\[
Z(S) = 2 \sum_{y \in \mathcal{Y}} \sqrt{P_{X,Y}(0, y)P_{X,Y}(1, y)} \quad (10)
\]

\[
= \sqrt{4P_X(0)P_X(1)} \sum_{y \in \mathcal{Y}} \sqrt{P_{Y|X}(y | 0)P_{Y|X}(y | 1)}
\]

which is the inner product between the unit vectors whose components are the square root of the distributions \( P_{Y|X=0} \) and \( P_{Y|X=1} \), under equiprobability \( P_X(0) = P_X(1) = \frac{1}{2} \). This
quantity informs about the similarity between the side-information $Y$ when $X$ is 0 and 1, under equiprobability $P_X(0) = P_X(1) = \frac{1}{2}$.

Let
\[ h(x) = -x \log x - (1-x) \log(1-x) \] (11)
be the entropy function expressed in bits, which admits an inverse $h^{-1} : [0, 1] \mapsto [0, \frac{1}{2}]$ when we restrict $x$ to be in $[0, \frac{1}{2}]$.

**Proposition 1.1 (Properties of Bhattacharyya parameter)** Let $(X, Y) \simeq P_{X,Y}$ be an arbitrary pair of random variables over $\mathcal{B} \times \mathcal{Y}$ with $\mathcal{B} = \{0, 1\}$ and $\mathcal{Y}$ an arbitrary countable set. For any memoryless source $S = (X, Y) \simeq P_{X,Y}$, with $X$ as the part to be compressed and $Y$ in the role of “side-information” about $X$, we have
\[
Z(S)^2 \leq H(S) \leq \log(1 + Z(S))
\] (12)
\[
Z(S^+) = Z(S)^2 \quad \text{and} \quad \sqrt{2Z(S)^2 - Z(S)^4} \leq Z(S^-) \leq 2Z(S)
\] (13)

The proof of the left inequality in (13) can be found in the paper\(^1\) by Chou et al. and the proofs of Proposition 1.1 and Theorem 1.2 can be found in the paper by Şaşoğlu\(^2\).

**Theorem 1.2 (Şaşoğlu)** Let $(X_1, Y_1)$ and $(X_2, Y_2)$ be independent pairs of discrete random variables taking their values in $\mathcal{B} \times \mathcal{Y}_1$ and $\mathcal{B} \times \mathcal{Y}_2$ with $\mathcal{B} = \{0, 1\}$, and let $\mathbb{H}(X_1 | Y_1) = \alpha$ and $\mathbb{H}(X_2 | Y_2) = \beta$. Then, the conditional entropy $\mathbb{H}(X_1 \oplus X_2 | Y_1, Y_2)$ is minimized when $\mathbb{H}(X_1 | Y_1 = y_1) = \alpha$ and $\mathbb{H}(X_2 | Y_2 = y_2) = \beta$ for all $(y_1, y_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ such that $P_{Y_1}(y_1)P_{Y_2}(y_2) > 0$. Moreover, if $\beta = \alpha = h(x)$ with $x \in [0, \frac{1}{2}]$ and if $0 < \alpha < 1$, then
\[
\min(\mathbb{H}(X_1 \oplus X_2 | Y_1, Y_2) - \mathbb{H}(X_1 | Y_1)) = h(2x(1-x)) - h(x) > 0,
\] (14)
where the minimum in (14) is taken on the set $\{P_{X_1, Y_1}, P_{X_2, Y_2} : \mathbb{H}(X_1 | Y_1) = \mathbb{H}(X_2 | Y_2) = \alpha\}$.

Finally, let us introduce
\[
Z'(S) = 1 - Z(S)^2.
\] (15)
The next proposition results straightforwardly from Proposition 1.1 and definition (15).

**Proposition 1.3** For any memoryless source $S$ with binary part to be compressed and discrete “side-information”, we have
\[
Z'(S^+) = 2Z'(S) - Z'(S)^2 \quad \text{and} \quad Z'(S^-) \leq Z'(S)^2.
\] (16)

## 2 Rough polarization

The following corollaries and theorems are adaptations to source polarization of results given by Guruswami and Xia\(^3\) for channel polarization.

**Corollary 2.1 (Guruswami & Xia)** There exists a constant $\theta_0$ with $0.799 < \theta_0 < 0.8$ such that for any memoryless source $S$ with binary part to compress and discrete side-information,
\[
H(S^-) - H(S) = H(S) - H(S^+) \geq \theta_0 H(S)(1 - H(S)).
\] (17)

---


\(^2\)Proposition 1.1 corresponds to Proposition 2.8 and Lemma 2.9 and Theorem 1.2 corresponds to Lemma 2.2 in “Polarization and polar codes”, Eren Şaşoğlu, *Fundations and Trends in Communications and Information Theory*, vol. 8, no. 4, pp. 259–381, 2011.

\(^3\)Corollary 2.1, Theorem 2.3, Corollary 2.4 and Theorem 2.8 correspond respectively to Lemma 6, Lemma 8, Corollary 9 and Proposition 5 in “Polar codes: speed of polarization and polynomial gap to capacity”, by Venkatesan Guruswami and Patrick Xia, *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 3–16, 2015.
Proof. We write $S = \{(X_i, Y_i) : i \in \mathbb{N}^*\}$, a sequence of independent drawings from $(X, Y) \sim P_{X,Y}$, as in Section 1. According to the definitions of $H(S^-)$ and $H(S)$, we have

$$H(S^-) - H(S) = H(X_1 \oplus X_2 | Y_1, Y_2) - H(X_1 | Y_1).$$

(18)

Using the entropy function $h$ defined in equation (11) and setting $H(X_1 | Y_1) = H(X_2 | Y_2) = h(x)$, with $x \in [0, \frac{1}{2}]$, it results from Theorem 1.2 that

$$H(S^-) - H(S) \geq h(2x(1-x)) - h(x).$$

(19)

Therefore

$$\frac{H(S^-) - H(S)}{H(S)(1-H(S))} \geq \frac{h(2x(1-x)) - h(x)}{h(x)(1-h(x))} \geq \min_{x \in [0, \frac{1}{2}]} \frac{h(2x(1-x)) - h(x)}{h(x)(1-h(x))} = \theta_0$$

(20)

and numerical simulations give $0.799 < \theta_0 < 0.8$. In order to end the proof, let us remark that the transformation $(X_1, X_2) \mapsto (X_1 \oplus X_2, X_2)$ is invertible, hence $H(S^+) + H(S^-) = 2H(S)$, i.e., $H(S) - H(S^+) = H(S^-) - H(S)$.

Remark 2.1 It results from Theorem 1.2 and its proof that for any $x \in [0, \frac{1}{2}]$, the inequality (19) can be an equality, hence the constant $\theta_0$ is the greatest value $\theta \in \mathbb{R}$ such that $H(S^-) - H(S) \geq \theta H(S)(1-H(S))$ for any memoryless source $S$ with binary-part to compress and discrete side-information.

Lemma 2.2 The function

$$g : [0, 1] \rightarrow [0, \frac{1}{2}], \quad \eta \mapsto \sqrt{\eta(1-\eta)}$$

(21)

is strictly concave and for any $\eta \in [0, 1]$, the function

$$G_\eta : [0, \min(\eta, 1-\eta)] \rightarrow [0, 1], \quad \delta \mapsto \frac{g(\eta+\delta)+g(\eta-\delta)}{2g(\eta)}$$

(22)

is strictly decreasing.

Proof. The functions $g$ and $G_\eta$ are well defined for any $\eta \in [0, 1]$. Moreover, the two first derivatives of $g$ are

$$g'(\eta) = \frac{1 - 2\eta}{2g(\eta)} \quad \text{and} \quad g''(\eta) = \frac{-1}{g(\eta)} \left[ 1 + \frac{(1-2\eta)^2}{4\eta(1-\eta)} \right] < 0$$

(23)

hence $g$ is strictly concave and finally for any $\eta \in [0, 1]$ and $\delta \in [0, \min(\eta, 1-\eta)]$

$$G_\eta'(\delta) = \frac{g'(\eta+\delta) - g'(\eta-\delta)}{2g(\eta)} \leq 0,$$

(24)

with equality if and only if $\delta = 0$, which completes the proof.

Theorem 2.3 (Guruswami & Xia) Let $g$ be the function defined in (21). There exists a constant $\Lambda < 1$, with $0.9165 < \Lambda < 0.9166$, such that for any memoryless source $S$ with binary-part to compress and discrete side-information

$$\frac{1}{2} \left[ g(H(S^-)) + g(H(S^+)) \right] \leq \Lambda g(H(S)).$$

(25)
Proof. For a memoryless source $S$ with binary-part to compress and discrete side-information, let $\eta = H(S), \varepsilon_0(\eta) = \theta_0 \eta (1 - \eta)$, where $\theta_0$ has been introduced in Corollary 2.1 and $\varepsilon = H(S^-) - H(S) = H(S) - H(S^+)$. It results from Corollary 2.1 that $\varepsilon \geq \varepsilon_0(\eta)$, which implies, according to Lemma 2.2, that

$$\frac{g(H(S^-)) + g(H(S^+))}{2g(H(S))} = \frac{g(\eta + \varepsilon) + g(\eta - \varepsilon)}{2g(\eta)} \leq \frac{g(\eta + \varepsilon_0(\eta)) + g(\eta - \varepsilon_0(\eta))}{2g(\eta)} = \frac{1}{2} \left( \sqrt{A_{\theta_0}(\eta)} + \sqrt{A_{\theta_0}(1 - \eta)} \right)$$

with $A_{\theta_0}(\eta) = [1 + \theta_0(1 - \eta)](1 - \theta_0 \eta) = \theta_0^2(\eta - \theta_0^{-1})[\eta - (1 + \theta_0^{-1})]$. The two roots of polynomial $A_{\theta_0}(\eta)$ are both outside the interval $[0, 1]$, therefore the function $\eta \mapsto \sqrt{A_{\theta_0}(\eta)}$ is strictly convex for $\eta \in [0, 1]$. As a result, its derivative is injective and the term (26) is maximum for $\eta \in [0, 1]$ if and only if

$$A_{\theta_0}^\prime(\eta) = A_{\theta_0}^\prime(1 - \eta) \quad \text{or} \quad \sqrt{A_{\theta_0}(\eta)} = \sqrt{A_{\theta_0}(1 - \eta)}$$

i.e., if and only if $\eta = \frac{1}{2}$. Hence, it comes

$$\frac{g(H(S^-)) + g(H(S^+))}{2g(H(S))} \leq \frac{1}{2} \sqrt{[1 + \theta_0(1 - \eta)](1 - \theta_0 \eta) + \frac{1}{2} \sqrt{(1 - \theta_0(1 - \eta))(1 + \theta_0 \eta)}} \bigg|_{\eta = \frac{1}{2}} = \sqrt{1 - \frac{\theta_0^2}{4}} = \Lambda.$$ 

Numerical simulations give $0.9165 < \Lambda < 0.9166$. 

A recursive application of Theorem 2.3 gives

$$\forall \mu \in \mathbb{N}, \quad \frac{1}{2^\mu} \sum_{i=0}^{2^\mu-1} g \left[ H \left( S^{(i)}_\mu \right) \right] \leq \Lambda^\mu g[H(S)] \leq \Lambda^\mu \max_{\eta \in [0, 1]} g(\eta) = \frac{1}{2} \Lambda^\mu. \quad (29)$$

This last equation can be interpreted as

$$\mathbb{E} \left\{ g \left[ H \left( S^{(j)}_\mu \right) \right] \right\} \leq \frac{1}{2} \Lambda^\mu, \quad (30)$$

where $J$ is a uniform random variable over $[0, 2^\mu - 1]$. Hence, the next corollary results from Markov’s inequality.

**Corollary 2.4 (Guruswami & Xia)** For any memoryless source $S$, with binary-part to compress and discrete side-information, and the associated sequence introduced in equation (4), for any $\mu \in \mathbb{N}$, if $J$ is a uniform random variable over $[0, 2^\mu - 1]$, then

$$\forall \theta > 0, \quad \mathbb{P} \left( g^2 \left[ H \left( S^{(j)}_\mu \right) \right] \geq \theta \right) \leq \frac{\Lambda^\mu}{2^\mu \sqrt{\theta}}. \quad (31)$$

We conclude this section by proving an adaptation to source polarization of the Guruswami and Xia rough (channel) polarization theorem. For any $\theta \in [0, \frac{1}{2}]$, let

$$x_1 = x_1(\theta) = \frac{1 - \sqrt{1 - 2\theta}}{2}, \quad \text{and} \quad x_2 = x_2(\theta) = \frac{1 + \sqrt{1 - 2\theta}}{2} = 1 - x_1 \quad (32)$$

be the solutions of $x(1 - x) = \frac{\theta}{2}$. We have $0 \leq x_1 \leq \frac{1}{2} \leq x_2 \leq 1$ and

$$\left\{ x \in [0, 1] : x(1 - x) \geq \frac{\theta}{2} \right\} = [x_1(\theta), x_2(\theta)]. \quad (33)$$
Moreover, $0 \leq 1 - 2\theta \leq \sqrt{1 - 2\theta} \leq 1$ implies $x_1(\theta) \leq \theta$. Hence, for any $m \in \mathbb{D}$ ($m = 2^n$), for any $\theta \in [0, \frac{1}{2}]$ and for any memoryless source $S$ with binary-part to compress and discrete side-information, we have

$$\left\{ i \in [0, m - 1] : g^2 \left[ H \left( S_{\mu}^{(i)} \right) \right] < \frac{\theta}{2} \right\} = \left\{ i \in [0, m - 1] : H \left( S_{\mu}^{(i)} \right) \in [0, x_1(\cup)x_2, 1] \right\} \subset \mathcal{V}_{X|Y}(x_1) \cup \mathcal{H}_{X|Y}^c(x_1) \quad (34)$$

and the following partition of $[0, m - 1]$:

$$[0, m - 1] = \mathcal{V}_{X|Y}(x_1) \cup \mathcal{H}_{X|Y}^c(x_1) \cup \left[ \mathcal{V}_{X|Y}^c(x_1) \cap \mathcal{H}_{X|Y}(x_1) \right], \quad (35)$$

implies, by setting $A = \mathcal{V}_{X|Y}(x_1)$, $B = \mathcal{H}_{X|Y}^c(x_1)$ and $C = \mathcal{V}_{X|Y}^c(x_1) \cap \mathcal{H}_{X|Y}(x_1)$,

$$1 - H(S) = 1 - \frac{1}{m} \sum_{i=0}^{m-1} H \left( S_{\mu}^{(i)} \right)$$

$$= \frac{1}{m} \left[ \sum_{i \in A} \left( 1 - H \left( S_{\mu}^{(i)} \right) \right) + \sum_{i \in B} \left( 1 - H \left( S_{\mu}^{(i)} \right) \right) + \sum_{i \in C} \left( 1 - H \left( S_{\mu}^{(i)} \right) \right) \right] \quad (36)$$

$$\leq \frac{|A|}{m} \left( 1 - \min \left\{ H \left( S_{\mu}^{(i)} \right) : i \in \mathcal{V}_{X|Y}(x_1) \right\} \right) + \frac{|B| + |C|}{m} \quad (37)$$

$$\leq x_1(\theta) + \frac{|B| + |C|}{m} \leq \theta + \frac{|B|}{m} + \mathbb{P}(J \in C), \quad (38)$$

where $J$ is a random variable uniformly distributed over $[0, m - 1]$. Now, the contraposition of (34) gives

$$C \subset \left\{ i \in [0, m - 1] : g^2 \left[ H \left( S_{\mu}^{(i)} \right) \right] \geq \frac{\theta}{2} \right\}, \quad (39)$$

hence it results from Corollary 2.4 that

$$\mathbb{P}(J \in C) \leq \frac{\Lambda^u}{2\sqrt{\theta/2}}, \quad (40)$$

and this implies, with inequality (38) where $\frac{|B|}{m} = \mathbb{P}(J \in \mathcal{H}_{X|Y}^c(x_1))$, that

$$\mathbb{P}(J \in \mathcal{H}_{X|Y}^c(x_1)) \geq 1 - H(S) - \theta - \frac{\Lambda^u}{2\sqrt{\theta/2}}. \quad (41)$$

**Proposition 2.5** There exists $\Lambda \in [0, 1]$ such that for any $\theta \in [0, \frac{1}{2}]$, for any memoryless source $S$ with binary-part to compress and discrete side-information, for any $m \in \mathbb{D}$ ($m = 2^n$), the subsets defined in equations (8–9) satisfy

$$\left\lvert \frac{\mathcal{H}_{X|Y}(\theta) \cap \mathcal{V}_{X|Y}(\theta)}{m} \right\rvert \leq \frac{\sqrt{2}\Lambda^u}{2\sqrt{\theta}} \quad (42)$$

$$H(S) - \theta \leq \left\lvert \frac{\mathcal{H}_{X|Y}(\theta)}{m} \right\rvert \leq H(S) + \frac{\sqrt{2}\Lambda^u}{2\sqrt{\theta}} \quad (43)$$

$$H(S) - \theta - \frac{\sqrt{2}\Lambda^u}{2\sqrt{\theta}} \leq \left\lvert \frac{\mathcal{V}_{X|Y}(\theta)}{m} \right\rvert \leq H(S) + \theta. \quad (44)$$

**Proof.** Since $x_1 = x_1(\theta) \leq \theta$, we have $\mathcal{H}_{X|Y}(x_1) \subset \mathcal{H}_{X|Y}^c(x_1)$ and $\mathcal{V}_{X|Y}(x_1) \subset \mathcal{V}_{X|Y}(x_1)$. Therefore: firstly $\mathcal{H}_{X|Y}(\theta) \cap \mathcal{V}_{X|Y}(\theta) \subset \mathcal{H}_{X|Y}(x_1) \cap \mathcal{V}_{X|Y}(x_1)$ and inequality (42) results from (40);
secondly \(|\mathcal{H}_{X|Y}(x_1)| \leq |\mathcal{H}_{X|Y}(\theta)|\) and the right inequality in (43) comes directly from (41). Furthermore, the conditions \(\max_{i \in \mathcal{H}_{X|Y}(\theta)} H(S^{(i)}_\mu) \leq \theta\) and \(\max_{i \in \mathcal{H}_{X|Y}(\theta)} H(S^\mu_\mu) \leq 1\) give

\[
H(S) = \frac{1}{m} \sum_{i=0}^{m-1} H(S^{(i)}_\mu) \leq \theta \mathbb{P}[J \in \mathcal{H}_{X|Y}(\theta)] + \mathbb{P}[J \in \mathcal{H}_{X|Y}(\theta)] \leq \theta + \frac{|\mathcal{H}_{X|Y}(\theta)|}{m},
\]

which proves the left inequality in (43). Similarly, condition \(\min_{i \in \mathcal{V}_{X|Y}(\theta)} H(S^{(i)}_\mu) \geq 1 - \theta\) implies

\[
H(S) \geq (1 - \theta) \mathbb{P}[J \in \mathcal{V}_{X|Y}(\theta)] \geq \frac{|\mathcal{V}_{X|Y}(\theta)|}{m} - \theta,
\]

which proves the right inequality in (44). Finally, note that \(\mathcal{V}_{X|Y}(\theta) \subset \mathcal{H}_{X|Y}(\theta)\) implies \(\mathcal{V}_{X|Y}(\theta) = \mathcal{H}_{X|Y}(\theta) \cap \mathcal{H}_{X|Y}(\theta)\), so the left inequality in (44) results from (42–43).

Let us now consider \(\rho \in [0, 1]\) and \(\mu \in \mathbb{N}^*\) such that \(4 \rho^2 \mu < \frac{1}{2}\), i.e., \(\mu > \frac{3}{2 \log(1/\rho)}\). For any memoryless source \(S\) with binary-part to compress and discrete side-information, since according to Proposition 1.1 for all \(i \in [0, 2^\mu - 1]\), \(Z(S^{(i)}_\mu) \leq \sqrt{H(S^{(i)}_\mu)}\), we have

\[
\{i \in [0, 2^\mu - 1] : H(S^{(i)}_\mu) \leq 4 \rho^2 \mu\} \subset \{i \in [0, 2^\mu - 1] : Z(S^{(i)}_\mu) \leq 2 \rho^2\},
\]

\[
\{i \in [0, 2^\mu - 1] : H(S^{(i)}_\mu) \leq x_1(4 \rho^2 \mu)\} \subset \{i \in [0, 2^\mu - 1] : H(S^{(i)}_\mu) \leq 4 \rho^2 \mu\};
\]

therefore with \(\theta = 4 \rho^2 \mu\) in (41), we obtain the following Proposition.

**Proposition 2.6** With the notations introduced in this subsection, for any \(\rho \in [0, 1]\), for any integer \(\mu > \frac{3}{2 \log(1/\rho)}\) we have

\[
\mathbb{P}(Z(S^{(i)}_\mu) \leq 2 \rho^2) \geq 1 - H(S) - 4 \rho^2 \mu - \frac{1}{2 \sqrt{2}} \left(\frac{\Lambda}{\rho}\right)\mu,
\]

where \(J\) is a random variable uniformly distributed over \([0, 2^\mu - 1]\).

Let us remark that

\[
\forall \theta \in \left[0, \frac{\sqrt{5} - 1}{2}\right], \quad \sqrt{1 - \theta} \leq 1 - \theta^2
\]

and for any \(\rho \in [0, 1]\), for any \(\mu \in \mathbb{N}\) such that

\[
\mu > \frac{2 - \log(\sqrt{5} - 1)}{\log(1/\rho)} \approx 1.69 \quad \text{i.e.,} \quad 2 \rho^2 < \frac{\sqrt{5} - 1}{2},
\]

we have

\[
\log(\frac{1}{1 - 2 \rho^2 \mu}) < \frac{1}{2} \quad \text{i.e.,} \quad \mu > \frac{3}{4} - \frac{\log(\sqrt{2} - 1)}{2} \quad \text{where} \quad \frac{1}{\log(1/\rho)} \approx \frac{1.38}{\log(1/\rho)},
\]

and it results from the right inequality in (12) that

\[
(H(S) > \log(2 - 4 \rho^2 \mu)) \Rightarrow (1 + Z(S) > 2 - 4 \rho^2 \mu).
\]

Now, \(\log(2 - 4 \rho^2 \mu) = 1 - \log(1 - 2 \rho^2 \mu^{-1})\), moreover \(1 + Z(S) > 2 - 4 \rho^2 \mu\) if and only if \(Z(S) > 1 - 4 \rho^2 \mu\) and \(1 - 4 \rho^2 \mu \geq \sqrt{1 - 2 \rho^2}\) according to (50–51), hence

\[
(H(S) > 1 - \log \left[\frac{1}{1 - 2 \rho^2 \mu}\right]) \Rightarrow (Z(S) > \sqrt{1 - 2 \rho^2}) \quad \text{i.e.,} \quad Z'(S) = 1 - Z(S)^2 < 2 \rho^2.
\]

Finally,

\[
1 - \log \left[\frac{1}{1 - 2 \rho^2 \mu}\right] = 1 + \frac{\ln(1 - 2 \rho^2 \mu)}{\ln 2} \leq 1 - \frac{2 \rho^2 \mu}{\ln 2}.
\]

Therefore the next proposition results from the left inequality in (44).
Proposition 2.7 With the notations introduced in this subsection, for any \( \rho \in ]0, 1[ \), for any integer \( \mu > \frac{2 - \log(\sqrt{5} - 1)}{\log(1/\rho)} \), we have

\[
P\left( Z'(S^{(j)}_\mu) < 2\rho^\mu \right) \geq H(S) - \frac{2}{\ln 2} \rho^{2\mu} - \frac{\sqrt{\ln 2}}{2} \left( \frac{\Lambda}{\rho} \right)^\mu,
\]

where \( J \) is a random variable uniformly distributed over \( [0, 2^\mu - 1] \).

We add this paragraph to prove the Rough polarization theorem.

Let \( \rho \in ]\Lambda, 1[ \) and \( \varepsilon \in ]0, 1/2[ \); let

\[
b_\rho = \max \left( \frac{2}{\ln(1/\rho)}, \frac{1}{\ln(\rho/\Lambda)} \right),
\]

and let \( \mu \in \mathbb{N} \) such that \( \mu > b_\rho \ln(1/\varepsilon) \). Then, we have \( \frac{\ln(1/\varepsilon) - (\ln 2)/2}{\ln(\rho/\Lambda)} < \frac{\ln(1/\varepsilon)}{\ln(\rho/\Lambda)} \leq b_\rho \ln(1/\varepsilon) < \mu \) and \( \frac{\ln(1/\varepsilon) + 3 \ln 2}{2 \ln(1/\rho)} \leq \frac{4 \ln(1/\varepsilon)}{2 \ln(1/\rho)} \leq b_\rho \ln(1/\varepsilon) < \mu \), which imply

\[
\frac{1}{2\sqrt{2}} \left( \frac{\Lambda}{\rho} \right)^\mu < \frac{\varepsilon}{2} \quad \text{and} \quad 4\rho^{2\mu} < \frac{\varepsilon}{2}, \quad \text{hence} \quad 4\rho^{2\mu} + \frac{1}{2\sqrt{2}} \left( \frac{\Lambda}{\rho} \right)^\mu < \varepsilon.
\]

We proved the following theorem by Guruswami and Xia.

**Theorem 2.8 (Rough polarization)** There exists \( \Lambda \in ]0, 1[ \) such that for any \( \rho \in ]\Lambda, 1[ \), there exists \( b_\rho > 0 \) such that for any memoryless source \( S \) with binary-part to compress and discrete side-information, for any \( \varepsilon \in ]0, 1/2[ \) and for any \( \mu \in \mathbb{N} \), such that \( \mu > b_\rho \ln(1/\varepsilon) \), there exists a roughly polarized set

\[
S_r \subset \{ S^{(i)}_\mu : 0 \leq i < 2^\mu \}
\]

such that for any \( M \in S_r \), \( Z(M) \leq 2\rho^\mu \) and \( P(S^{(j)}_\mu \in S_r) \geq 1 - H(S) - \varepsilon \), where \( J \) is a random variable uniformly distributed over \( [0, 2^\mu - 1] \).

### 3 Fine polarization

This section is an adaptation of the reasoning given by Guruswami and Xia (see footnote 3) to prove their fine polarization theorem.

#### 3.1 Preliminaries

**Lemma 3.1** For any \( \beta \in ]0, \ 1/2[ \), the function \( \zeta : \mathbb{R}^* \rightarrow \mathbb{R} \) defined by

\[
\zeta(y) = \frac{|y|}{2y} + \frac{2\beta^2 y}{|y|} = \left\{ \begin{array}{ll}
\frac{1 + 4\beta^2}{2} & \text{if } y \in \mathbb{N}^* \\
\frac{2}{2^{(q+\alpha)}} + \frac{2\beta^2(q+\alpha)}{q+1} & \text{if } y = q + \alpha \text{ with } q \in \mathbb{N} \text{ and } 0 < \alpha < 1,
\end{array} \right.
\]

satisfies the condition:

\[
\forall y, \quad \left( |y| \geq \frac{4\beta^2}{1 - 4\beta^2} \right) \Rightarrow \min\{ \zeta([y] + \alpha) : \alpha \in [0, 1] \} = \frac{|y|}{2([y] + 1) + 2\beta^2}.
\]
Let \( q \in \mathbb{N}^* \), let us introduce the continuously differentiable function of \( \alpha \)

\[
f_q : \mathbb{R}_+ \to \mathbb{R}_+^* \quad \alpha \mapsto \frac{q}{2(q+\alpha)} + \frac{2\beta^2(q+\alpha)}{q+1},
\]

which satisfies \( f_q(\alpha) = \zeta(q + \alpha) \) (\( \forall \alpha \in [0, 1] \)) and

\[
f_q(0) = \frac{1}{2} + 2\beta^2 \frac{q}{q+1} < \frac{1}{2} + 2\beta^2 \quad \text{and} \quad f_q(1) = \frac{q}{2(q+1)} + 2\beta^2 < \frac{1}{2} + 2\beta^2,
\]

\[
f'_q(\alpha) = \frac{2\beta^2}{q+1} - \frac{q}{2(q+\alpha)^2} \quad \text{and} \quad f'_q(\alpha) = 0 \iff \alpha = \sqrt{q(q+1)} - q.
\]

Let us remark that \( \beta \in ]0, \frac{1}{2}[^{1} \) implies \( f'_q(0) = \frac{1}{2q} + \frac{2\beta^2}{q+1} < 0 \) and the zero of the derivative function is always greater than zero. Moreover, the zero of the derivative function is smaller than 1 if and only if

\[
\frac{q(q+1)}{4\beta^2} < (q+1)^2 \iff q < \frac{4\beta^2}{1 - 4\beta^2}.
\]

Hence, if \( |y| = q \geq \frac{4\beta^2}{1 - 4\beta^2} \), then \( \min \{ f_q(\alpha) : \alpha \in [0, 1] \} = f_q(1) = \frac{|y|}{2(|y| + 1)} + 2\beta^2. \)

The proofs of the following three lemmas are straightforward.

**Lemma 3.2** The function

\[
\zeta : \mathbb{R}_+ \to \mathbb{R}_-^* \quad y \mapsto -\alpha_0 2^y + y \quad \text{with} \quad \alpha_0 = \frac{2}{e \ln 2} \simeq 1.06
\]

is maximum at the point \( y_0 = \frac{-\ln(\alpha_0 \ln 2)}{\ln 2} = \frac{1}{\ln 2} - 1 \simeq 0.44 \) and \( \zeta(y_0) = -1. \)

**Lemma 3.3** For any \( \beta \in ]0, \frac{1}{2}[ \), the function

\[
\varphi : \mathbb{R}_+^* \to \mathbb{R}_+^* \quad y \mapsto \frac{1 - \beta + y^{-1}}{1 - 2 - \beta y}
\]

is strictly decreasing, \( \varphi(1/\beta) = 2 \) and \( \varphi(y) \) approaches to \( 1 - \beta < 1 \) when \( y \) approaches to infinity. Hence, there exists \( c_\beta > 0 \) such that

\[
\forall y, (y \geq c_\beta \Rightarrow \varphi(y) < 1).
\]

**Lemma 3.4** \( \forall \beta \in ]0, \frac{1}{2}[ \), \( \forall \gamma \in \mathbb{R}_+^* \), \( \forall \xi > 1 \) and \( \forall \rho \in ]0, 1[ \) the function

\[
\psi : \mathbb{R}_+ \to \mathbb{R}_+^* \quad y \mapsto \sqrt{e} \left( 1 + \frac{\gamma \xi}{\log(1/\rho)} \right) \exp \left[ -\frac{(1-2\beta)^2 y}{2} \right]
\]

is strictly decreasing on \( \mathbb{R}_+ \), \( \psi(0) > 1 \) and \( \psi(y) \) approaches to 0 when \( y \) approaches to infinity.

### 3.2 Introduction of parameters, constants and notations

Let \( \delta \in ]0, \frac{1}{2}[ \) and \( \beta \in ]\delta, \frac{1}{2}[ \) such that \( \gamma = \frac{\delta}{\beta - \delta} \) is a rational number. We put \( \gamma = \frac{\gamma_n}{\gamma_d} \) with \( \gamma_n \) and \( \gamma_d \) co-prime integers (\( \gamma \) can take any value in \( \mathbb{Q}_+^* \)).

Let \( \rho \in ]\Lambda, 1[ \), \( \xi > 1 \) (the \( x \) parameter used by Guruswami and Xia is connected to \( \xi \) with the relation \( x = \frac{\log(1/\rho)}{\xi \log(2/\rho)} \)) and

\[
c = \left\lfloor \frac{\gamma \xi}{\log(1/\rho)} \right\rfloor.
\]
Using the constant $c_\beta$ introduced in lemma 3.3, let

$$c'_\beta = \left( \frac{\xi}{(\log(1/\rho) + 1/\gamma) + 1} \right) \max \left\{ c_\beta, \frac{1}{1 - 2\beta} \right\}, \quad (71)$$

$$c_\delta = \max \left( \frac{(1 + \alpha_0)\xi}{(\xi - 1) \log(1/\rho)}, c'_\beta \right). \quad (72)$$

Let $\mu$ be a natural integer multiple of $\gamma_d$ such that

$$\mu > c_\delta. \quad (73)$$

Finally, let us introduce

$$\nu = \gamma \mu \quad \text{and} \quad \nu_0 = (\gamma + 1)\mu = \nu + \mu, \quad (74)$$

which are natural integers because $\mu$ is a multiple of $\gamma_d$.

For any $j \in [1, c]$, let us note $I_j = \left\lfloor \frac{(j-1)\nu}{c} \right\rfloor \mod N$, $n_j = |I_j|$ and

$$G_j (\nu) = \left\{ i = \sum_{k=0}^{\nu-1} b_k 2^k : \sum_{k \in I_j} b_k \geq \frac{\beta \nu}{c} \right\}, \quad G'_j (\nu) = \left\{ i = \sum_{k=0}^{\nu-1} b_k 2^k : \sum_{k \in I_j} (1 - b_k) \geq \frac{\beta \nu}{c} \right\} \quad (75)$$

where $b_0, \ldots, b_{\nu-1}$ are the binary digits of $i$. It comes

$$\left\lfloor \frac{\nu}{c} \right\rfloor \leq n_j \leq \left\lceil \frac{\nu}{c} \right\rceil \quad \text{and} \quad \sum_{j=1}^{c} n_j = \nu. \quad (76)$$

Finally, let us put

$$G(\nu) = \bigcap_{j=1}^{c} G_j (\nu) \quad \text{and} \quad G'(\nu) = \bigcap_{j=1}^{c} G'_j (\nu). \quad (77)$$

### 3.3 Proof of the fine polarization theorem

**Lemma 3.5 (Guruswami & Xia)** With the notations introduced in the previous subsections, if $J_2$ is a random variable with uniform distribution over $[0, 2^\nu - 1]$, then

$$\mathbb{P}(J_2 \in G(\nu)) \geq 1 - \psi \left( \frac{\nu}{c} \right), \quad (78)$$

$$\mathbb{P}(J_2 \in G'(\nu)) \geq 1 - \psi \left( \frac{\nu}{c} \right). \quad (79)$$

**Proof.** If we write $J_2 = \sum_{k=0}^{\nu-1} B_k 2^k$, then the $\nu$ bits $B_k$ are independent Bernoulli random variables with parameter $\frac{1}{2}$. If $J_2$ takes its values in $G_j (\nu)$, we have

$$\sum_{k \in I_j} B_k - \frac{n_j}{2} \geq - \left( \frac{n_j}{2} - \frac{\beta \nu}{c} \right). \quad (80)$$

Moreover

$$\frac{n_j}{2} - \frac{\beta \nu}{c} \geq \frac{1}{2} \left\lfloor \frac{\nu}{c} \right\rfloor - \frac{\beta \nu}{c} > \frac{1}{2} \left( \frac{\nu}{c} - 1 \right) - \frac{\beta \nu}{c} \quad (81)$$

and since, according to equations (70–74), we have

$$\frac{\nu}{c} \geq \frac{\gamma \mu}{1 + \frac{\gamma \xi}{\log(1/\rho)}} = \frac{\mu}{\frac{\xi}{\log(1/\rho)}} \geq \frac{1}{(1 - 2\beta)}, \quad (82)$$
it comes
\[ \frac{1}{2} \left( \frac{\nu}{c} - 1 \right) - \frac{\beta \nu}{c} = \frac{(1 - 2\beta)\nu}{2c} - \frac{1}{2} > 0. \] (83)

It then results from Hoeffding’s inequality that
\[ \mathbb{P}(J_2 \in G_j(\nu)) = 1 - \mathbb{P} \left( \sum_{k \in I_j} B_k - \frac{n_j}{2} < - \left( \frac{n_j}{2} - \frac{\beta \nu}{c} \right) \right) \geq 1 - \exp \left( - \left( \frac{1}{2} - \frac{\beta \nu}{cn_j} \right)^2 2n_j \right). \] (84)

Now, \( \left( \frac{1}{2} - \frac{\beta \nu}{cn_j} \right)^2 2n_j = \frac{n_j}{2} - \frac{2\beta \nu}{c} + \frac{2\beta^2 \nu^2}{c^2 n_j} \geq \frac{\nu}{c} \left( \frac{n_j}{2^2} - 2\beta + \frac{2\beta^2 \nu}{n_j} \right) \geq \frac{\nu}{c} \left( \zeta \left( \frac{\nu}{c} \right) - 2\beta \right) \), where the last inequality comes from (76) and the \( \zeta \) function is defined in equation (60). Further it results from condition (82) that\(^4\) \[ \left[ \frac{\nu}{c} \right] \geq \frac{4\beta^2}{1 - 4\beta^2}. \] Therefore \( \zeta \left( \frac{\nu}{c} \right) \geq \frac{\left[ \frac{\nu}{c} \right]}{2(\left[ \frac{\nu}{c} \right] + 1)} + 2\beta^2 = \frac{1}{2} + 2\beta^2 - \frac{1}{2(\left[ \frac{\nu}{c} \right] + 1)} \), according to Lemma 3.1, and
\[ \left( \frac{1}{2} - \frac{\beta \nu}{cn_j} \right)^2 2n_j \geq \frac{\nu(1 - 2\beta)^2}{2c} - \frac{1}{2} \left[ \frac{\nu}{c} \right] + 1 \geq \frac{\nu(1 - 2\beta)^2}{2} - \frac{1}{2}. \] (85)

Hence
\[ \mathbb{P}(J_2 \in G_j(\nu)) \geq 1 - \sqrt{c} \exp \left[ - \frac{(1 - 2\beta)^2 \nu}{2c} \right] \] (86)
and we obtain
\[ \mathbb{P}(J_2 \not\in G(\nu)) \leq \sum_{j=1}^{c} P(J_2 \not\in G_j(\nu)) \leq \sqrt{c} \exp \left[ - \frac{(1 - 2\beta)^2 \nu}{2c} \right] \leq \psi \left( \frac{\nu}{c} \right) \] (87)
(the last inequality resulting from equations (70) and (69)), that proves (78). Finally, the same proof, where \( B_k \) is replaced with \( (1 - B_k) \) in relations (80) and (84) and \( G_j(\nu) \) is replaced with \( G_j'(\nu) \) leads to (79).

Now for any memoryless source \( S \) with binary-part to compress and discrete side-information, it results from Proposition 2.6 (since\(^5\) \( \mu > \frac{3}{2 \log(1/\rho)} \)) that there exists \( \mathcal{S}_r \subset \{ S^{(i)}_\mu : 0 \leq i < 2^n \} \) such that
\[ \forall M \in \mathcal{S}_r, \ Z(M) \leq 2^{\rho^\mu} \quad \text{and} \quad \mathbb{P}(S^{(J_1)}_\mu \subset \mathcal{S}_r) \geq 1 - H(S) - 4\rho^{2\mu} - \frac{1}{2(\sqrt{2})^2} \left( \frac{\Lambda}{\rho} \right)^\mu, \] (88)
where \( J_1 \) is a random variable uniformly distributed over \([0, 2^n - 1]\). In a similar way, it results from Proposition 2.7 (since\(^6\) \( \mu > \frac{3}{2 \log(1/\rho)} \)) that there exists \( \mathcal{S}'_r \subset \{ S^{(i)}_\mu : 0 \leq i < 2^n \} \) such that
\[ \forall M \in \mathcal{S}'_r, \ Z'(M) \leq 2^{\rho^\mu} \quad \text{and} \quad \mathbb{P}(S^{(J_1)}_\mu \subset \mathcal{S}'_r) \geq H(S) - \frac{2}{2 \log(2^{2\rho^\mu})} - \frac{\sqrt{2}}{2} \left( \frac{\Lambda}{\rho} \right)^\mu. \] (89)

For any \( M \in \mathcal{S}_r \), we define the sequence
\[ \tilde{Z}^{(i)}_k = \begin{cases} \left( \tilde{Z}^{(i/2)}_{k-1} \right)^2 & \text{if } i \text{ is odd}, \\ \frac{2}{2^{\log(2^{2\rho^\mu})}} & \text{if } i \text{ is even}, \end{cases} \] for any \( k \in \mathbb{N}^* \), with \( \tilde{Z}^{(0)}_0 = Z(M) \). (90)

\(^4\)Indeed, the condition (82) implies
\[ \left[ \frac{\nu}{c} \right] \geq \frac{\nu}{c} - 1 > \frac{1}{1 - 2\beta} - 1 = \frac{2\beta}{1 - 2\beta} > \left( \frac{2\beta}{1 + 2\beta} \right) \left( \frac{2\beta}{1 + 2\beta} \right) = \frac{4\beta^2}{1 - 4\beta^2}. \]

\(^5\)Indeed, condition (72) implies \( \mu > \frac{3}{2 \log(1/\rho)} \).

\(^6\)Indeed, condition (72) implies \( \mu > \frac{3}{2 \log(1/\rho)} \).
Let us note $R(\mu) = \{ i \in [0, 2^\mu - 1] : S_{\mu}^{(i)} \in \mathcal{S}_r \}$.

In the same way, for any $M' \in \mathcal{S}'_r$, we define the sequence

$$
\tilde{Z}'_{(i)} = \begin{cases} 
\left( \tilde{Z}'_{(i/2)} \right)^2 & \text{if } i \text{ is even}, \\
2\tilde{Z}'_{(i/2)} & \text{if } i \text{ is odd},
\end{cases}
$$

for any $k \in \mathbb{N}^*$, with $\tilde{Z}'_0 = Z'(M')$. \hfill (91)

Let us note $R'(\mu) = \{ i \in [0, 2^\mu - 1] : S_{\mu}^{(i)} \in \mathcal{S}'_r \}$.

**Lemma 3.6 (Guruswami & Xia)** With the notations introduced in the previous subsection, we have \( \log \left( \max \left\{ \tilde{Z}'_{(i)} : i \in G(\nu) \right\} \right) \leq -2^{3\nu} \).

**Proof.** Let us note for $j \in [1, c]$, $\nu_j = \sum_{k=1}^{j} n_k$ (thus $\nu_c = \nu$) and

$$
z_j = \max \left\{ \tilde{Z}'_{(i/2^{\nu_j})} : i \in G(\nu) \right\} \text{ for any } j \in [1, c] \text{ and } z_0 = Z(M). \hfill (92)
$$

Let us remark first that if $z_j$ has been attained by $p$ squarings and $n_j$ doublings from $z_{j-1}$, the maximum value will be obtained by applying first the $n_j - p$ doublings followed by the $p$ squarings. Moreover, if $z_j < 1$, the maximum value will be reached by minimizing the number of squarings.

According to relations (72), (74), (70) and (88), we have $\mu > \frac{(1 + \alpha \log(1/\rho))}{(1 - \xi \log(1/\rho))}$, $M \in \mathcal{S}_r$, $\nu_c \leq \frac{\mu \log(1/\rho)}{\xi}$ and

$$
\log Z(M) + \frac{\nu}{c} \leq 1 - \mu \log(1/\rho)(1 - 1/\xi) \leq 1 - (1 + \alpha) = -\alpha_0 \leq -1. \hfill (93)
$$

Equation (93) shows that $\log Z(M) + n_1 \leq \log Z(M) + \frac{\nu}{c} + 1 < 0$, hence if one doubles $n_1$ times $z_0$ one obtains a value that is smaller than 1. Thus $z_1 < 1$ and since for any $i \in G(\nu)$, the number of squarings is worth at least $\left\lceil \frac{2^\nu}{\nu} \right\rceil$, we have

$$
\log z_1 \leq 2^{\left\lceil \frac{2^\nu}{\nu} \right\rceil} \left( \log Z(M) + \frac{\nu}{c} + 1 - \left\lceil \frac{2\nu}{c} \right\rceil \right) \leq 2^{\frac{\nu}{c}} \left( \log Z(M) + \frac{(1 - \beta)\nu}{c} + 1 \right), \hfill (94)
$$

which can be written, using the $\varphi$ function introduced in Lemma 3.3,

$$
\log z_1 + \frac{\nu}{c} \varphi \left( \frac{\nu}{c} \right) \leq 2^{\frac{\nu}{c}} \left( \log Z(M) + \frac{\nu}{c} \varphi \left( \frac{\nu}{c} \right) \right) \leq 2^{\frac{\nu}{c}} \left( \log Z(M) + \frac{\nu}{c} \right), \hfill (95)
$$

according to condition (68) and

$$
\frac{\nu}{c} \geq \frac{\gamma \mu}{1 + \frac{\gamma \mu}{\log(1/\rho)}} = \frac{\mu}{1 + \frac{\mu}{\log(1/\rho)}} > \frac{\mu}{c} \hfill (96)
$$

– the last inequality resulting from condition (73) and definitions (71–72) of $c_i$.

Moreover, according to Lemma 3.3, the $\varphi$ function is greater than $1 - \beta$ and it results from (93) and (95) that

$$
\log z_1 + \frac{\nu}{c} \leq -\alpha_0 2^{\frac{\nu}{c}} + \frac{\beta \nu}{c} = \zeta \left( \frac{\beta \nu}{c} \right) \leq -1, \hfill (97)
$$

where the $\zeta$ function is defined in Lemma 3.2. So $\frac{\nu}{c} + 1 + \log z_1 < 0$, therefore $z_2 < 1$ and the same reasoning as above leads to

$$
\log z_2 \leq 2^{\frac{\nu}{c}} \left( \log z_1 + n_2 - \left\lceil \frac{\beta \nu}{c} \right\rceil \right) < 2^{\frac{\nu}{c}} \left( \log z_1 + \frac{(1 - \beta)\nu}{c} + 1 \right). \hfill (98)
$$

\[\text{Indeed, starting from } x, \text{ if we apply } p \text{ squarings and } n_j - p \text{ doublings, the final result will be of the form } x^{2^p} 2^{\nu'}, \text{ and } \alpha, \text{ the power of 2, will be maximum if the } n_j - p \text{ doublings precede the } p \text{ squarings.}\]

\[\text{If } z_j > 1, \text{ the maximum value can be reached by replacing some doublings by squarings: starting from } x > 0, \text{ } p + 1 \text{ squarings will give a greater result than } p \text{ squarings if and only if } \]

$$
2^p (\log x + n_j - p) \leq 2^{p+1} (\log x + n_j - p - 1) \iff \log x \geq 2 - (n_j - p).
$$
which can be written

\[
\log z_2 + \frac{\nu}{c} \varphi \left( \frac{\nu}{c} \right) \leq 2^{\frac{2\nu}{c}} \left( \log z_1 + \frac{\nu}{c} \varphi \left( \frac{\nu}{c} \right) \right),
\]

(99)

which, with (95), leads to

\[
\log z_2 + (1 - \beta) \frac{\nu}{c} \leq \log z_2 + \frac{\nu}{c} \varphi \left( \frac{\nu}{c} \right) \leq 2^{\frac{2\nu}{c}} \left( \log z_0 + \frac{\nu}{c} \varphi \left( \frac{\nu}{c} \right) \right)
\]

(100)

and according to conditions (93) and (96) and the property (68), it follows that

\[
\log z_2 + \frac{\nu}{c} \leq 2^{\frac{2\beta \nu}{c}} \left( \log z_0 + \frac{\nu}{c} \varphi \left( \frac{\nu}{c} \right) \right) \\
< -\alpha_0 2^{\frac{2\beta \nu}{c}} + \frac{\beta \nu}{c} = \zeta \left( \frac{2\beta \nu}{c} \right) - \frac{\beta \nu}{c} \leq \zeta \left( \frac{2\beta \nu}{c} \right) \leq -1,
\]

(101)

where the last inequality results from Lemma 3.2. More generally, let us suppose that

\[
\log z_j - 1 + \frac{\nu}{c} \leq 2^{\frac{(j - 1)\beta \nu}{c}} \left( \log z_0 + \frac{\nu}{c} \varphi \left( \frac{\nu}{c} \right) \right) + \frac{\beta \nu}{c},
\]

(102)

so \( 1 + \frac{\nu}{c} + \log z_{j-1} < 1 + \zeta \left( \frac{(j-1)\beta \nu}{c} \right) \leq 0 \) according to (93) and Lemma 3.2, and the same reasoning as above gives

\[
\log z_j + \frac{(1 - \beta)\nu}{c} \leq 2^{\frac{1\beta \nu}{c}} \left( \log z_0 + \frac{\nu}{c} \right),
\]

(103)

in particular for \( j = c \):

\[
\log z_c \leq 2^{\beta \nu} \left( \log z_0 + \frac{\nu}{c} \right) \leq -2^{\beta \nu},
\]

(104)

where the last inequality results from relation (93).

The same reasoning replacing \( \tilde{Z}_k^{(i)} \) with \( \tilde{Z}_k^{(i)} \) leads to the following lemma.

**Lemma 3.7** With the notations introduced in the previous subsection, we have

\[
\log \left( \max \left\{ \tilde{Z}_\nu^{(i)} : i \in G'(\nu) \right\} \right) \leq -2^{\beta \nu}.
\]

After all, we can deduce from equations (103–104) that

\[
\log z_c + \frac{(1 - \beta)\nu}{c} \leq -2^{\beta \nu}.
\]

(105)

Further, since \( \log(1/\ln 2) \simeq 0.529 \) and \( \beta > 0 \), we have

\[
\beta \left( 2 \log(1/\ln 2) - 1 \right) > 0 > \log(1/\ln 2) - 1, \quad \text{i.e.,} \quad \frac{1}{1 - 2\beta} > \frac{\log(1/\ln 2)}{1 - \beta},
\]

(106)

which implies, according to the conditions (71,73) and the equation (70),

\[
\mu > \left( \frac{\xi}{\log(1/\rho)} + \frac{1}{\gamma} \right) \frac{1}{1 - 2\beta} > \frac{c \log(1/\ln 2)}{\gamma(1 - \beta)},
\]

(107)

i.e., \( \frac{1 - \beta \nu}{c} > \log(1/\ln 2) \). Thus the relation (105) leads to the following lemma.

**Lemma 3.8** With the notations introduced in the previous subsection, we have

\[
\log \left( \max \left\{ \tilde{Z}_\nu^{(i)} : i \in G(\nu) \right\} \right) \leq -2^{\beta \nu} + \log(\ln 2).
\]

(108)
Let us recall that \( \nu_0 = \nu + \mu \) and let us summarize the results proved in this section. We expanded \( i \in [0, 2^{\nu_0} - 1] \) into \( i = i_1 + 2^{\nu} i_2 \), where the binary digits of \( i_2 \in [0, 2^{\nu} - 1] \) and \( i_1 \in [0, 2^{\nu} - 1] \) correspond respectively to the first \( \nu \) and last \( \mu \) bits of \( i \). We proved that if \( i_1 \in R(\mu) \) (i.e., if \( S_{\mu}^{(i_1)} \in S_{\nu} \) or equivalently if \( Z(S_{\mu}^{(i_1)}) < 2^\rho \)) and if \( i_2 \in G(\nu) \), then

\[
Z(S_{\mu}^{(i_1)}) \leq 2^{-2^{\beta\nu}} \ln 2 = 2^{-2^{\beta\nu_0}} \ln 2. \tag{109}
\]

We also proved that if \( i_1 \in R'(\mu) \) (i.e., if \( S_{\mu}^{(i_1)} \in S_{\nu}' \) or equivalently if \( Z'(S_{\mu}^{(i_1)}) < 2^\rho \)) and if \( i_2 \in G'(\nu) \), then

\[
Z'(S_{\mu}^{(i_1)}) \leq 2^{-2^{\beta\nu}} = 2^{-2^{\beta\nu_0}}. \tag{110}
\]

Now, according to equations (88) and (78) and assuming \( J = J_1 + 2^\mu J_2 \) is a random variable uniformly distributed over \([0, 2^{\nu_0} - 1]\), we have

\[
\mathbb{P}(J_1 \in R(\mu) \text{ and } J_2 \in G(\nu)) = \mathbb{P}(J_1 \in R(\mu)) \mathbb{P}(J_2 \in G(\nu))
\]

\[
\geq \left(1 - H(S) - 4\rho^2 - \frac{1}{2\sqrt{2}} \left(\frac{\Lambda}{\rho}\right)^\mu\right) \left(1 - \psi \left(\frac{\nu}{c}\right)\right).
\]

We deduce that

\[
\mathbb{P} \left(Z(S_{\nu_0}^{(j)}) \leq 2^{-2^{\beta\nu_0}} \ln 2\right) \geq \mathbb{P}(J_1 \in R(\mu) \text{ and } J_2 \in G(\nu))
\]

\[
\geq 1 - H(S) - 4\rho^2 - \frac{1}{2\sqrt{2}} \left(\frac{\Lambda}{\rho}\right)^\mu - \psi \left(\frac{\nu}{c}\right). \tag{111}
\]

In a same way, we have

\[
\mathbb{P} \left(Z'(S_{\nu_0}^{(j)}) \leq 2^{-2^{\beta\nu_0}}\right) \geq \mathbb{P}(J_1 \in R'(\mu) \text{ and } J_2 \in G'(\nu))
\]

\[
\geq H(S) - \frac{2}{\ln 2} \rho^2 + \frac{\sqrt{\ln 2}}{2} \left(\frac{\Lambda}{\rho}\right)^\mu - \psi \left(\frac{\nu}{c}\right). \tag{113}
\]

Thus, we proved that for any \( \mu \) multiple of \( \gamma_d \); \( \mu = k\gamma_d \) with \( k \in \mathbb{N}^* \) great enough \((k > \frac{\omega}{\gamma_d})\) or in other words for any sufficiently large \( \nu_0 = \mu(1 + \gamma) = k(\gamma_d + \gamma_n) \), the relations (109–110) and (112–113) are valid.

Let us consider now \( \nu' \) between two successive multiples of \( \gamma_d + \gamma_n \):

\[
\nu_0 = k(\gamma_d + \gamma_n) + u = \mu(1 + \gamma') \tag{114}
\]

with

\[
\mu = k\gamma_d, \quad 0 \leq u < \gamma_n + \gamma_d \quad \text{and} \quad \gamma' = \frac{k\gamma_n + u}{k\gamma_d} = \gamma + \frac{u}{k\gamma_d},
\]

then

\[
\gamma \leq \gamma' = \gamma + \frac{u}{k\gamma_d} \leq \gamma + \frac{u}{\gamma_d} < 1 + 2\gamma. \tag{116}
\]

Let us remark that if we introduce \( \delta' \) such that

\[
\gamma' = \frac{\delta'}{\beta - \delta'} \geq \gamma = \frac{\delta}{\beta - \delta},
\]

then \( \delta \leq \delta' \) and \( 2^{-2^{\nu_0}} \leq 2^{-2^{\nu_0}} \). Thus, by replacing \( \gamma \) with \( 1 + 2\gamma \) in the definition of \( \psi \) (see (69)), leaving \( \gamma \) unchanged in equation (71) and replacing \( \gamma \) with \( \gamma' \) everywhere else, the above reasoning can be remade in order to prove the following proposition.
Proposition 3.9 For any \( \delta \in ]0, \frac{1}{2}[, \) for any \( \beta \in ]\delta, \frac{1}{2}[, \) for any \( \rho \in ]0, 1[, \) for any \( \xi > 1, \) there exists \( C_{\delta,\beta} > 0 \) and \( A_{\delta,\beta} > 0 \) such that for any memoryless source \( S \) with binary-part to compress and discrete side-information and for any integer \( \nu_0 > C_{\delta,\beta}, \) we have

\[
\begin{align*}
\mathbb{P}(Z(S_{\nu_0}^{(J)}) \leq 2^{-2^{\nu_0} \ln 2}) & \geq 1 - H(S) - 4\rho^{2\nu_0} - \frac{1}{2\sqrt{2}} \left( \frac{\Lambda}{\rho} \right)^{\nu_0} - A_{\delta,\beta} \exp \left( \frac{-\varepsilon(1-2\beta)^2}{2c} \right) \\
\mathbb{P}(Z'(S_{\nu_0}^{(J)}) \leq 2^{-2^{\nu_0}}) & \geq H(S) - \frac{2}{\ln 2} \rho^{2\nu_0} - \frac{\sqrt{2}}{2} \left( \frac{\Lambda}{\rho} \right)^{\nu_0} - A_{\delta,\beta} \exp \left( \frac{-\varepsilon(1-2\beta)^2}{2c} \right),
\end{align*}
\]
where \( J \) is a random variable uniformly distributed over \( [0, 2^{\nu_0} - 1], \) \( \gamma = \frac{\delta}{\beta-\delta}, \) \( \nu_0 = (\gamma + 1)\mu, \) \( \nu = \gamma\mu \) and \( c = \left\lceil \frac{\gamma\xi}{\log(1/\rho)} \right\rceil. \)

Let us denote

\[
\varepsilon = \frac{\mu}{\nu_0} = \frac{1}{\gamma + 1} = \frac{\beta - \delta}{\beta}, \quad \text{hence} \quad \beta = \frac{\delta}{1 - \varepsilon} \quad \text{and} \quad \gamma = \frac{1}{\varepsilon} - 1. \tag{120}
\]

Now we can choose \( \xi > 1 \) so that the fraction \( \frac{\gamma\xi}{\log(1/\rho)} \) is an integer (equal to \( c \)), then we have

\[
\frac{\nu}{c} = \gamma\mu \quad \text{and} \quad c = \frac{\mu \log(1/\rho)}{\xi}, \tag{121}
\]
and the previous proposition becomes:

Proposition 3.10 For any \( \delta \in ]0, \frac{1}{2}[, \) for any \( \beta \in ]\delta, \frac{1}{2}[, \) for any \( \varepsilon \in ]0, 1-2\delta[, \) for any \( \xi > 1 \) such that \( (1-\varepsilon)^{\xi}/\varepsilon \log(1/\rho) \in \mathbb{N}^* \), there exists \( C_{\delta,\xi} > 0 \) and \( A_{\delta,\xi} > 0 \) such that for any memoryless source \( S \) with binary-part to compress and discrete side-information and for any integer \( \nu_0 > C_{\delta,\xi}, \) we have

\[
\begin{align*}
\mathbb{P}(Z(S_{\nu_0}^{(J)}) \leq 2^{-2^{\nu_0} \ln 2}) & \geq 1 - H(S) - 4\rho^{2\nu_0} - \frac{1}{2\sqrt{2}} \left( \frac{\Lambda}{\rho} \right)^{\nu_0} - A_{\delta,\xi} \exp \left( \frac{-\varepsilon(1-2\beta)^2}{2c} \right) \\
\mathbb{P}(Z'(S_{\nu_0}^{(J)}) \leq 2^{-2^{\nu_0}}) & \geq H(S) - \frac{2}{\ln 2} \rho^{2\nu_0} - \frac{\sqrt{2}}{2} \left( \frac{\Lambda}{\rho} \right)^{\nu_0} - A_{\delta,\xi} \exp \left( \frac{-\varepsilon(1-2\beta)^2}{2c} \right),
\end{align*}
\]
where \( J \) is a random variable uniformly distributed over \( [0, 2^{\nu_0} - 1] \). Moreover we can choose

\[
A_{\delta,\xi} = \sqrt{\varepsilon} \left( 1 + \frac{(2 - \varepsilon)\xi}{\varepsilon \log(1/\rho)} \right). \tag{124}
\]

Let us put \( n = 2^{\nu_0} \). Equation (122) can be written

\[
\mathbb{P}(Z(S_{\nu_0}^{(J)}) \leq 2^{-n^\xi \ln 2}) \geq 1 - H(S) - \frac{4}{n^{\kappa_1}} - \frac{4}{2\sqrt{2} \cdot n^{\kappa_2}} - \frac{A_{\delta,\xi}}{n^{\kappa_3}} \tag{125}
\]
with

\[
\begin{align*}
\kappa_1^{(1)} &= 2\varepsilon \log(1/\rho), \\
\kappa_2^{(2)} &= \varepsilon \log(\rho/\Lambda) \\
\kappa_3^{(3)} &= \varepsilon \left( 1 - \frac{2\delta}{1-\varepsilon} \right)^2 \log(1/\rho) \frac{\ln 2}{2\xi \ln 2}.
\end{align*}
\]
In order to shorten the notations, let us put

$$
\alpha = \alpha(\delta, \varepsilon, \rho, \xi) = \left(1 - \frac{2\delta}{1 - \varepsilon}\right)^2 \frac{1}{2\xi \ln 2}.
$$  \hspace{1cm} (129)

For fixed $\delta \in ]0, \frac{1}{2}[$, we look for $\varepsilon \in ]0, 1 - 2\delta[$, $\rho \in ]0, 1[$ and $\xi > 1$ that maximize
$$
\kappa_\varepsilon = \min \left(\kappa^{(1)}_\varepsilon, \kappa^{(2)}_\varepsilon, \kappa^{(3)}_\varepsilon\right).
$$

Let us remark that

$$
\kappa^{(3)}_\varepsilon = \varepsilon \left(1 - \frac{2\delta}{1 - \varepsilon}\right)^2 \frac{\log(1/\rho)}{2\xi \ln 2} < \frac{\varepsilon \log(1/\rho)}{2 \ln 2} < \varepsilon \log(1/\rho) = \kappa^{(1)}_\varepsilon,
$$

therefore

$$
\min \left(\kappa^{(1)}_\varepsilon, \kappa^{(2)}_\varepsilon, \kappa^{(3)}_\varepsilon\right) = \min \left(\kappa^{(2)}_\varepsilon, \kappa^{(3)}_\varepsilon\right).
$$

Moreover, we have

$$
\kappa^{(2)}_\varepsilon \leq \kappa^{(3)}_\varepsilon \iff \log(\rho/\Lambda) \leq \alpha \log(1/\rho) \iff \log\rho \leq \log\Lambda + \alpha \iff \rho \leq \Lambda^{1/\alpha}.
$$

Firstly, let us suppose that $\rho \leq \Lambda^{1/\alpha}$.

In this case we have

$$
\min \left(\kappa^{(1)}_\varepsilon, \kappa^{(2)}_\varepsilon, \kappa^{(3)}_\varepsilon\right) = \kappa^{(2)}_\varepsilon = \varepsilon \log(\rho/\Lambda) \text{ and this expression is maximum when the independent variables } \rho \text{ and } \varepsilon \text{ are maximum, hence for } \rho = \Lambda^{1/\alpha},
$$

which leads to

$$
\min \left(\kappa^{(1)}_\varepsilon, \kappa^{(2)}_\varepsilon, \kappa^{(3)}_\varepsilon\right) = \kappa^{(2)}_\varepsilon = \kappa^{(3)}_\varepsilon = \log(1/\Lambda) \frac{\varepsilon\alpha}{1 + \alpha} \text{ with } \alpha = \alpha(\varepsilon, \xi),
$$

since $\delta$ is supposed to be fixed. Secondly, if we suppose that

$$
\rho \geq \Lambda^{1/\alpha},
$$

then

$$
\min \left(\kappa^{(1)}_\varepsilon, \kappa^{(2)}_\varepsilon, \kappa^{(3)}_\varepsilon\right) = \kappa^{(3)}_\varepsilon = \varepsilon \alpha \log(1/\rho) \text{ and this quantity is maximum when } \alpha \text{ is maximum and } \rho \text{ minimum, i.e., when inequality } (134) \text{ is an equality, i.e., when } (133) \text{ is satisfied.}
$$

Further, the exponent $\kappa^{(2)}_\varepsilon = \kappa^{(3)}_\varepsilon$ in equation (133) is maximum if and only if

$$
g(\varepsilon, \xi) \overset{\text{def}}{=} \frac{\varepsilon \alpha(\varepsilon, \xi)}{1 + \alpha(\varepsilon, \xi)}
$$

is maximum. Since

$$
\frac{\partial g}{\partial \xi}(\varepsilon, \xi) = \frac{\varepsilon}{(1 + \alpha)^2} \cdot \frac{\partial \alpha}{\partial \xi}(\varepsilon, \xi) = \frac{-\varepsilon \left(1 - \frac{2\delta}{1 - \varepsilon}\right)^2}{(1 + \alpha)^2 2\xi^2 \ln 2} < 0,
$$

$g(\varepsilon, \xi)$ is maximum when

$$
\xi = \xi_{\text{min}} > 1
$$

and equation (129) implies that

$$
\alpha(\varepsilon, \xi_{\text{min}}) = \left(1 - \frac{2\delta}{1 - \varepsilon}\right)^2 \frac{1}{2\xi_{\text{min}} \ln 2} = \alpha_{\text{max}}(\varepsilon).
$$

Finally, the function to maximize is

$$
\tilde{g}(\varepsilon) \overset{\text{def}}{=} \frac{\varepsilon \alpha_{\text{max}}(\varepsilon)}{1 + \alpha_{\text{max}}(\varepsilon)}
$$

whose derivative

$$
\tilde{g}'(\varepsilon) = \frac{\alpha_{\text{max}}(\varepsilon)}{1 + \alpha_{\text{max}}(\varepsilon)} + \frac{\varepsilon \alpha'_{\text{max}}(\varepsilon)}{(1 + \alpha_{\text{max}}(\varepsilon))^2}
$$

(140)
vanishes if and only if $\alpha_{\text{max}}(\varepsilon)(1 + \alpha_{\text{max}}(\varepsilon)) + \varepsilon \alpha'_{\text{max}}(\varepsilon) = 0$. We obtain a trivial solution $\varepsilon = 1 - 2\delta$ (corresponding to a minimum: $\tilde{g}(1 - 2\delta) = 0$) and a third degree algebraic equation in $u$, with $u = 1 - \varepsilon$:

$$P(u) = Au^3 - Bu^2 - Cu - D,$$

with

$$
\begin{aligned}
A &= 2\xi_{\text{min}} \ln 2 + 1 \\
B &= 2\delta(3 - 2\xi_{\text{min}} \ln 2) \\
C &= 4\delta(2\xi_{\text{min}} \ln 2 - 3\delta) \\
D &= 8\delta^3
\end{aligned}
$$

(141)

The discriminant of $P$ is the resultant $\text{Res}(P, P')$ between polynomials $P(u)$ and its derivative $P'(u) = 3Au^2 - 2Bu - C$:

$$\text{Res}(P, P') = \begin{vmatrix}
A & -B & -C & -D & 0 \\
0 & A & -B & -C & -D \\
0 & 0 & 3A & -2B & -C \\
0 & 3A & -2B & -C & 0 \\
3A & -2B & -C & 0 & 0
\end{vmatrix}
$$

(142)

$$\text{Res}(P, P') = 256(\xi_{\text{min}} \ln 2)^2(2\xi_{\text{min}} \ln 2 + 1)(4\delta^2 - 4\delta + 2\xi_{\text{min}} \ln 2 + 1)\delta^3$$

$$\times \left(8\xi_{\text{min}} \ln 2 - \delta(27 - 2\xi_{\text{min}} \ln 2)\right).$$

(143)

The third degree equation $P(u) = 0$ admits a multiple root if and only if the resultant $\text{Res}(P, P')$ vanishes, i.e., if and only if $\delta = 0$ or (see Figure 2)

$$\delta = \delta_1(\xi_{\text{min}}) = \frac{8\xi_{\text{min}} \ln 2}{27 - 2\xi_{\text{min}} \ln 2} \left. \right|_{\xi_{\text{min}}=1} \simeq 0.21649.$$

(144)

Figure 1: Graphs of $u_1(\delta)$ and $u_2(\delta)$, the roots of $P'(u)$, for $\delta \in ]0, \frac{1}{2}[ \text{ and various values of } \xi_{\text{min}} = (10 + k)/10 \text{ (}0 \leq k \leq 10\text{); solid line corresponds to } \xi_{\text{min}} = 1.$

The discriminant of $P$ is the resultant $\text{Res}(P, P')$ between polynomials $P(u)$ and its derivative $P'(u) = 3Au^2 - 2Bu - C$:
Thus for all $\delta \in ]0, \delta_1(\xi_{\text{min}})[$, the equation $P(u) = 0$ admits three real roots, for $\delta > \delta_1(\xi_{\text{min}})$ the same equation admits only one real root and for $\delta = \delta_1(\xi_{\text{min}})$, the real root is multiple. Moreover, let us introduce the discriminant of $P'(u)$:

$$\Delta' = B^2 + 3AC = 8\delta\xi_{\text{min}} \ln 2(6\xi_{\text{min}} \ln 2 + 3 - \delta(15 - 2\xi_{\text{min}} \ln 2)),$$

which vanishes when

$$\delta = \delta_0(\xi_{\text{min}}) = \frac{6\xi_{\text{min}} \ln 2 + 3}{15 - 2\xi_{\text{min}} \ln 2} \approx 0.52586,$$

and, for $\delta \in ]0, \frac{1}{2}[$, let

$$u_1(\delta) = \frac{B - \sqrt{\Delta'}}{3A} < 0 \quad \text{and} \quad u_2(\delta) = \frac{B + \sqrt{\Delta'}}{3A} > 0$$

be the real roots of $P'(u)$ (see Figure 1). We have $P(u_1(\delta_1(\xi_{\text{min}}))) = 0$ and $u_2(\delta_2) = 2\delta_2$ with $\delta_2 = 1/5 = 0.2$.

We show on Figure 2 the graphs of the values of $P(u)$ when $P'(u)$ vanishes as functions of $\delta$ for different values of $\xi_{\text{min}}$ and we can see that for all $\delta \in ]0, \delta_1(\xi_{\text{min}})[$, $P(u_1) > 0$. Furthermore, since $P(0) = -D < 0$ (\forall \delta > 0) and $P(u) \to -\infty$ when $u \to -\infty$ (\forall \delta), we deduce that when equation $P(u) = 0$ admits three real zeros (i.e., when $0 < \delta < \delta_1(\xi_{\text{min}})$), two of the three roots are smaller than zero.

Finally, since polynomial $4\delta^2 - 4\delta + 2\xi_{\text{min}} \ln 2 + 1$ has no real roots, we can remark that

$$P(1) = A - B - C - D = (1 - 2\delta)(4\delta^2 - 4\delta + 2\xi_{\text{min}} \ln 2 + 1) > 0 \quad \text{for all} \quad \delta \in ]0, \frac{1}{2}[$$  \quad (148)

$$P(2\delta) = 8A\delta^3 - 4B\delta^2 - 2C\delta - D = -16\xi_{\text{min}}(\ln 2)\delta^2(1 - 2\delta) < 0 \quad \text{for all} \quad \delta \in ]0, \frac{1}{2}[$$  \quad (149)
Therefore, for all \( \delta \in ]0, \frac{1}{2}[ \), there is always one and only one zero of \( P(u) \) with \( 2\delta < u < 1 \). Let us note \( \nu_1(\delta) \) this root of \( P \). All the above mentioned conditions on the real roots \( \nu_i(\delta) \) \((1 \leq i \leq 3)\) of \( P(u) \) can be observed on Figure 3, which has been obtained with numerical simulations.

Let us note \( \nu_1(\delta) \) this root of \( P \). All the above mentioned conditions on the real roots \( \nu_i(\delta) \) \((1 \leq i \leq 3)\) of \( P(u) \) can be observed on Figure 3, which has been obtained with numerical simulations.

![Graph of \( \nu_1(\delta), \nu_2(\delta), \nu_3(\delta) \) the real roots of \( P(u) \) and of \( \nu = 2\delta \) for \( \delta \in ]0, \frac{1}{2}[ \) and various values of \( \xi_{\text{min}} = (10 + k)/10 \) \((0 \leq k \leq 10)\); solid line corresponds to \( \xi_{\text{min}} = 1 \).](image)

Figure 3: Graphs of \( \nu_1(\delta), \nu_2(\delta), \nu_3(\delta) \) the real roots of \( P(u) \) and of \( \nu = 2\delta \) for \( \delta \in ]0, \frac{1}{2}[ \) and various values of \( \xi_{\text{min}} = (10 + k)/10 \) \((0 \leq k \leq 10)\); solid line corresponds to \( \xi_{\text{min}} = 1 \).

We show on Figure 4 \((a)\) the graph of \( \tilde{g}(1 - \nu_1(\delta)) \) as a function of \( \delta \) and \((b)\) the graph of \( \tilde{g}(\varepsilon) \) as a function of \( \varepsilon \) for various \( \delta \). We see that \( \tilde{g} \) is maximum for \( \lim_{\delta \to 0^+} \tilde{g}(1 - \nu_1(\delta)) \approx 0.0046789995 \).

Thus, we prove the following Proposition.

**Proposition 3.11** For any \( \delta \in ]0, \frac{1}{2}[ \) and for any \( \varepsilon \in ]0, 1 - 2\delta[ \), there exists \( \kappa_{\delta, \varepsilon} > 0 \), \( A_{\delta, \varepsilon} > 0 \) and \( C_{\delta, \varepsilon} \) such that for any memoryless source \( S \) with binary-part to compress and discrete side-information and for any integer \( \nu_0 > C_{\delta, \varepsilon} \) — noting \( n = 2^{\nu_0} \) —, we have

\[
\mathbb{P} \left( Z(S_{\nu_0}(J)) \leq 2^{-n^4 \ln 2} \right) \geq 1 - H(S) - \frac{A_{\delta, \varepsilon}}{n^\kappa_{\delta, \varepsilon}} \quad (150)
\]

\[
\mathbb{P} \left( Z'(S_{\nu_0}(J)) \leq 2^{-n^4} \right) \geq H(S) - \frac{A_{\delta, \varepsilon}}{n^\kappa_{\delta, \varepsilon}}. \quad (151)
\]

where \( J \) is a random variable uniformly distributed over \([0, n - 1]\).

Moreover putting \( B_{\delta, \varepsilon} = \left[ \frac{(1 - \frac{2\delta}{\varepsilon})^2 + 2 \ln 2}{2 \ln (1/\Lambda)} \right] \), we can choose \( A_{\delta, \varepsilon} \) such that

\[
\sqrt{e} \left( 1 + \frac{2 - \varepsilon}{\varepsilon} B_{\delta, \varepsilon} \right) + \frac{\sqrt{\ln 2}}{2} < A_{\delta, \varepsilon} < \sqrt{e} \left( 1 + \frac{2 - \varepsilon}{\varepsilon} B_{\delta, \varepsilon} \right) + \frac{\sqrt{\ln 2}}{2} + \frac{2}{\ln 2}. \quad (152)
\]

Further, for any \( \delta \in ]0, \frac{1}{2}[ \), for any memoryless source \( S \) with binary-part to compress and discret side-information, for any \( \nu_0 \in \mathbb{N} \), let us apply inequality (42) of Proposition 2.5 with
Figure 4: Graphs of (a) $\hat{g}(1 - \nu_1(\delta))$ as a function of $\delta$ and various values of $\xi_{\text{min}} = (10 + k)/10$ ($0 \leq k \leq 10$) (solid line corresponds to $\xi_{\text{min}} = 1$) and (b) of $\hat{g}(\varepsilon)$ as a function of $\varepsilon$ for various $\delta$ and $\xi_{\text{min}} = 1$.

$\mu = \nu_0$, $m = n = 2^m$ and $\theta = 2^{-n^\delta}$ (provided that $\theta \leq \frac{1}{2}$):

$$\frac{|\mathcal{H}_X|_{\nu'}(2^{-n^\delta}) \cap \mathcal{V}_{X'}(2^{-n^\delta})|}{n} \leq \frac{\sqrt{2}\Lambda_{\nu_0}}{2\sqrt{2^{-n^\delta}}}. \quad (153)$$

We claim that for any $\kappa > 0$, for any $\varepsilon > 0$, there exists $\theta_{\kappa,\varepsilon} > 0$ such that for all $n > \theta_{\kappa,\varepsilon}$

$$\frac{\sqrt{2}\Lambda_{\nu_0}}{2\sqrt{2^{-n^\delta}}} \leq \frac{\theta_{\kappa,\varepsilon}}{n^{\kappa\varepsilon}}. \quad (154)$$

Indeed, inequality (154) is equivalent to

$$\log n (\log \Lambda + \kappa \varepsilon) \leq \log \theta_{\kappa,\varepsilon} + \frac{1}{2} (1 + n^\delta), \quad (155)$$

which is satisfied when $\theta_{\kappa,\varepsilon}$ and $n$ are sufficiently large.

Finally we have

$$\mathbb{P}\left( H(S_{v_0}^{(J)}) > 2^{-n^\delta} \right) = 1 - \mathbb{P}\left( H(S_{v_0}^{(J)}) \leq 2^{-n^\delta} \right) \quad (156)$$

and since according to relation (12),

$$Z(S_{v_0}^{(J)}) \leq 2^{-n^\delta} \ln 2 \quad \Rightarrow \quad H(S_{v_0}^{(J)}) \leq \log \left( 1 + Z(S_{v_0}^{(J)}) \right) \leq \frac{Z(S_{v_0}^{(J)})}{\ln 2} \leq 2^{-n^\delta}, \quad (157)$$

then

$$\mathbb{P}\left( H(S_{v_0}^{(J)}) \leq 2^{-n^\delta} \right) \geq \mathbb{P}\left( Z(S_{v_0}^{(J)}) \leq 2^{-n^\delta} \ln 2 \right) \quad (158)$$

and with the notations of Proposition 3.11, we have

$$\mathbb{P}\left( H(S_{v_0}^{(J)}) > 2^{-n^\delta} \right) \leq H(S) + \frac{A_{\delta,\varepsilon}}{r_{\varepsilon}^{n^\delta,\varepsilon}}. \quad (159)$$

Similarly

$$Z'(S_{v_0}^{(J)}) \leq 2^{-n^\delta} \quad \Rightarrow \quad H(S_{v_0}^{(J)}) \geq 1 - 2^{-n^\delta} \quad (160)$$

implies

$$\mathbb{P}\left( H(S_{v_0}^{(J)}) > 1 - 2^{-n^\delta} \right) \geq \mathbb{P}\left( Z'(S_{v_0}^{(J)}) \leq 2^{-n^\delta} \right) \geq H(S) - \frac{A_{\delta,\varepsilon}}{r_{\varepsilon}^{n^\delta,\varepsilon}}. \quad (161)$$

Therefore, applying the left inequality of relation (43) and the right inequality of relation (44) we prove the following proposition.
Proposition 3.12 For any $\delta \in \left[0, \frac{1}{2}\right]$ and for any $\varepsilon \in \left[0, 1 - 2\delta\right]$, there exists $\kappa_{\delta, \varepsilon} > 0$, $A_{\delta, \varepsilon} > 0$ and $C_{\delta, \varepsilon}$ such that for any memoryless source $S$ with binary-part to compress and discrete side-information and for any integer $n_0 > C_{\delta, \varepsilon} - \varepsilon$, we have

$$0 \leq \frac{|H_{\mathcal{X}|\mathcal{Y}}(2^{-n^{\delta}} \cap V_{\mathcal{X}|\mathcal{Y}}(2^{-n^{\delta}})|}{n} \leq A_{\delta, \varepsilon} \frac{1}{n^{\varepsilon \kappa_{\delta, \varepsilon}}}$$ (162)

$$H(S) - 2^{-n^{\delta}} \leq \frac{|H_{\mathcal{X}|\mathcal{Y}}(2^{-n^{\delta}})|}{n} \leq H(S) + A_{\delta, \varepsilon} \frac{1}{n^{\varepsilon \kappa_{\delta, \varepsilon}}}$$ (163)

$$H(S) - \frac{A_{\delta, \varepsilon}}{n^{\varepsilon \kappa_{\delta, \varepsilon}}} \leq \frac{|V_{\mathcal{X}|\mathcal{Y}}(2^{-n^{\delta}})|}{n} \leq H(S) + 2^{-n^{\delta}},$$ (164)

where $A_{\delta, \varepsilon}$ satisfies the inequalities (152).

3.4 Order of magnitude of constants

In this subsection we study the values of the constants $c_{\beta}$ and $c_{\delta}$ with numerical simulations. Firstly we compute the solution $\alpha = \alpha(\beta)$ of the equation

$$\varphi\left(\frac{\log \alpha - \log \beta}{\beta}\right) = 1 \Leftrightarrow \frac{\beta}{\alpha} + \frac{\beta}{\log \alpha - \log \beta} = \beta,$$ (165)

such that

$$c_{\beta}^{(1)} = \frac{1}{\beta} \log \left(\frac{\alpha(\beta)}{\beta}\right)$$ (166)

is the smallest permissible value of $c_{\beta}$ satisfying (68). Moreover we find a simple expression

$$c_{\beta}^{(2)} = \frac{1}{\beta} \log \left(\frac{1.6\beta + 1.203}{\beta}\right)$$ (167)

slightly greater than the smallest value $c_{\beta}^{(1)}$ (see Figures 5–7).

9Let us remark that this new function $\alpha$ is not connected to the $\alpha$ function introduced in (129).
Figure 6: Graphs of $c_\beta^{(1)}$, $c_\beta^{(2)}$ and $\frac{1}{1-2\beta}$.

Figure 7: Graphs of $\varphi(c_\beta^{(1)})$ and $\varphi(c_\beta^{(2)})$. 
Secondly, we assume that inequality (134) is an equality and replacing $\alpha$ by the expression (129), we obtain
\[
\log(1/\rho) = \frac{\log(1/\Lambda)}{1 + \alpha} = \frac{2\xi \ln(1/\Lambda)}{2\xi \ln 2 + (1 - 2\beta)^2}
\]
and we express $c'_\beta$ and $c'_\delta$ introduced in relations (71,72) as functions of $\beta$ and $\xi$.

Figure 8: Graphs of $c'_\beta$ versus $\xi$ and $\beta$.

\[
c'_\beta \text{ with } \log(1/\rho) = \frac{\log(1/\Lambda)}{1 + \alpha_{\text{max}}(\beta, \xi)} = \frac{2\xi \ln(1/\Lambda)}{2\xi \ln 2 + (1 - 2\beta)^2}
\]

Figure 9: Graphs of $c'_\beta$ versus $\xi$ and $\beta$. 

22
Let us denote
\[ c^{(1)}_{\delta} = \frac{(1 + \alpha_0)\xi}{(\xi - 1) \log(1/\rho)} \] (169)

appearing in the definition (72) of \( c_{\delta} \).
In order to have
\[ c \overset{\text{def}}{=} \left\lceil \frac{\gamma \xi}{\log(1/\rho)} \right\rceil = \frac{\gamma \xi}{\log(1/\rho)} \]  
with \( \rho \) satisfying equality (168) and \( \xi > 1 \) as small as possible, we set \( \xi_{\text{min}} \) as the value of \( \xi \) solution of equations (168) and
\[
\frac{(1 - \varepsilon)\xi}{\varepsilon \log(1/\rho)} = \min \left( \mathbb{N}^+ \cap \left\{ \frac{(1 - \varepsilon)\xi}{\varepsilon \log(1/\rho)} : \xi > 1 \right\} \right). \tag{171}
\]

Table 1: Some numerical values obtained by simulations.

<table>
<thead>
<tr>
<th>( \delta )</th>
<th>( \varepsilon )</th>
<th>( \beta )</th>
<th>( \gamma )</th>
<th>( c_3^{(2)} )</th>
<th>( \xi_{\text{min}} )</th>
<th>( c_3 )</th>
<th>( c_3' )</th>
<th>( A_{\delta,\varepsilon} )</th>
<th>( C_{\delta,\varepsilon} )</th>
<th>( \kappa_\infty = \kappa_{\delta,\varepsilon} )</th>
<th>( \varepsilon_{\kappa_\infty} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.79</td>
<td>0.48</td>
<td>0.27</td>
<td>4.3</td>
<td>21</td>
<td>1.42</td>
<td>0.1255</td>
<td>316</td>
<td>316</td>
<td>30.1</td>
<td>400</td>
</tr>
<tr>
<td>0.10</td>
<td>0.74</td>
<td>0.38</td>
<td>0.35</td>
<td>5.8</td>
<td>4.3</td>
<td>1.03</td>
<td>0.121</td>
<td>66</td>
<td>506</td>
<td>25.6</td>
<td>684</td>
</tr>
<tr>
<td>0.10</td>
<td>0.69</td>
<td>0.32</td>
<td>0.45</td>
<td>7.5</td>
<td>2.8</td>
<td>1.03</td>
<td>0.116</td>
<td>83</td>
<td>650</td>
<td>29.5</td>
<td>943</td>
</tr>
<tr>
<td>0.10</td>
<td>0.64</td>
<td>0.28</td>
<td>0.56</td>
<td>9.2</td>
<td>2.3</td>
<td>1.20</td>
<td>0.112</td>
<td>115</td>
<td>115</td>
<td>39.0</td>
<td>179</td>
</tr>
<tr>
<td>0.10</td>
<td>0.59</td>
<td>0.24</td>
<td>0.69</td>
<td>11.1</td>
<td>2.0</td>
<td>1.08</td>
<td>0.107</td>
<td>127</td>
<td>270</td>
<td>41.3</td>
<td>458</td>
</tr>
<tr>
<td>0.10</td>
<td>0.54</td>
<td>0.22</td>
<td>0.85</td>
<td>13.0</td>
<td>1.8</td>
<td>1.10</td>
<td>0.104</td>
<td>153</td>
<td>223</td>
<td>48.7</td>
<td>413</td>
</tr>
<tr>
<td>0.10</td>
<td>0.49</td>
<td>0.20</td>
<td>1.04</td>
<td>15.1</td>
<td>1.6</td>
<td>1.06</td>
<td>0.101</td>
<td>173</td>
<td>352</td>
<td>55.3</td>
<td>719</td>
</tr>
<tr>
<td>0.10</td>
<td>0.44</td>
<td>0.18</td>
<td>1.27</td>
<td>17.1</td>
<td>1.6</td>
<td>1.08</td>
<td>0.099</td>
<td>201</td>
<td>268</td>
<td>65.9</td>
<td>610</td>
</tr>
<tr>
<td>0.10</td>
<td>0.39</td>
<td>0.16</td>
<td>1.56</td>
<td>19.3</td>
<td>1.5</td>
<td>1.04</td>
<td>0.096</td>
<td>221</td>
<td>567</td>
<td>75.6</td>
<td>1428</td>
</tr>
<tr>
<td>0.10</td>
<td>0.34</td>
<td>0.15</td>
<td>1.94</td>
<td>21.5</td>
<td>1.4</td>
<td>1.01</td>
<td>0.093</td>
<td>243</td>
<td>2394</td>
<td>88.7</td>
<td>7042</td>
</tr>
<tr>
<td>0.10</td>
<td>0.29</td>
<td>0.15</td>
<td>2.45</td>
<td>23.7</td>
<td>1.4</td>
<td>1.01</td>
<td>0.092</td>
<td>271</td>
<td>1633</td>
<td>108.9</td>
<td>5631</td>
</tr>
<tr>
<td>0.10</td>
<td>0.24</td>
<td>0.13</td>
<td>3.17</td>
<td>26.0</td>
<td>1.4</td>
<td>1.04</td>
<td>0.091</td>
<td>304</td>
<td>629</td>
<td>139.1</td>
<td>2622</td>
</tr>
<tr>
<td>0.10</td>
<td>0.19</td>
<td>0.12</td>
<td>4.26</td>
<td>28.4</td>
<td>1.3</td>
<td>1.01</td>
<td>0.089</td>
<td>326</td>
<td>3834</td>
<td>178.5</td>
<td>20179</td>
</tr>
<tr>
<td>0.10</td>
<td>0.14</td>
<td>0.12</td>
<td>6.14</td>
<td>30.8</td>
<td>1.3</td>
<td>1.01</td>
<td>0.088</td>
<td>355</td>
<td>3168</td>
<td>251.3</td>
<td>22634</td>
</tr>
<tr>
<td>0.10</td>
<td>0.09</td>
<td>0.11</td>
<td>10.11</td>
<td>33.2</td>
<td>1.3</td>
<td>1.00</td>
<td>0.087</td>
<td>384</td>
<td>8302</td>
<td>403.1</td>
<td>92253</td>
</tr>
<tr>
<td>0.10</td>
<td>0.04</td>
<td>0.10</td>
<td>24</td>
<td>35.7</td>
<td>1.3</td>
<td>1.00</td>
<td>0.087</td>
<td>414</td>
<td>6290</td>
<td>937.4</td>
<td>157257</td>
</tr>
</tbody>
</table>