Proceedings ISIT 2010, June 13–18, 2010, Austin, USA, 1272–1276

On colorings of bivariate random sequences

František Matúš
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
18208 Prague, P.O.Box 18, Czech Republic
Email: matus@utia.cas.cz

Michal Kupsa
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
18208 Prague, P.O.Box 18, Czech Republic
Email: kupsa@utia.cas.cz

***Abstract*—The ergodic sequences consisting of vectors (ξ_n, η_n), $n \geqslant 1$, over a finite alphabet $A\times B$ are colored with $\lfloor e^{n\alpha} \rfloor$ colors for A^n and $\lfloor e^{n\beta} \rfloor$ colors for B^n. Generic behavior of the colorings in terms of probabilities of monochromatic rectangles intersected with typical sets is examined. When n increases a big majority of pairs of colorings produces rectangles whose probabilities are bounded uniformly from above. Limiting rates of bounds are worked out in all regimes of the rates α and β of colorings. As a consequence, generic behavior of the colorings in terms of Shannon entropies of the partitions into rectangles is described.**

I. Introduction

Let N be a finite set and $(\xi_i)_{i\in N}$ a random vector taking a finite number of values. The collection of the Shannon entropies $H(\xi_i, i \in I)$, $I \subseteq N$, of all the subvectors of the vector can be interpreted as an entropic point of a Euclidean space. The last decade has seen renewed investigations of regions of the entropic points and closely related information theoretical inequalities [10], [4], [5], [6], [7]. Motivation has been drawn from numerous schemes of information theory [2], [9] and elsewhere.

Properties of regions of the entropic points and their limits were studied recently in [6] by taking independent identically distributed copies $(\xi_i^{(n)})_{i\in N}$, $n \geqslant 1$, of the vector and by randomly coloring the copies $\xi_i^{(1)}, \dots, \xi_i^{(n)}$ of a *single* distinguished variable ξ_i with $\lfloor e^{n\alpha_i} \rfloor$ colors, $\alpha_i \geqslant 0$. In this way new entropic points were constructed from old ones and their limits were described in [6, Theorem 3].

In this framework, a natural idea is to color independently the copies $\xi_i^{(1)}, \dots, \xi_i^{(n)}$ of *each* variable ξ_i with $\lfloor e^{n\alpha_i} \rfloor$ colors, $i \in N$, and to investigate the entropies of the partitions into monochromatic blocks. Thus, the N-tuples of colorings come into play and a majority of them is expected to have a similar, in some sense generic, behaviour in terms of entropies.

In this contribution, a restriction is made to bivariate vectors. The multivariate case likely differs only by additional technicalities and will be treated elsewhere. Instead of independent copies of a bivariate vector, an ergodic sequence consisting of vectors (ξ_n, η_n), $n \geqslant 1$, over a finite alphabet $A\times B$ is considered. The more general assumption of ergodicity entails no additional technical complications.

A k-coloring of a set X is any mapping f of X into the set of colors $\{1, \dots, k\}$, to be denoted by $\widehat{k}$. The cardinality of X is denoted by $|X|$.

In the bivariate sequence, first marginal sequence $\xi_1, \dots, \xi_n$ is colored with $k_n = \lfloor e^{n\alpha} \rfloor$ colors and second one $\eta_1, \dots, \eta_n$ with $\ell_n = \lfloor e^{n\beta} \rfloor$ colors, $\alpha, \beta \geqslant 0$. If

$$f_n \colon A^n \to \widehat{k_n} \quad \text{and} \quad g_n \colon B^n \to \widehat{\ell_n}$$

are two such colorings, the product $A^n\times B^n$ partitions into the (f_n, g_n)-monochromatic rectangles

$$f_n^{-1}(i)\times g_n^{-1}(j)\,, \qquad i \in \widehat{k_n}\,, j \in \widehat{\ell_n}\,.$$

The main result, Theorem 2 in Section IV, deals with uniform upper bounds on probabilities of these rectangles intersected with special typical sets $Z_n \subseteq A^n\times B^n$. The bounds decay exponentially with the rate

$$h_{\xi\eta}^{\alpha\beta} = \min\{h_{\xi\eta}, \alpha + h_\eta, h_\xi + \beta, \alpha + \beta\}$$

where $h_{\xi\eta}$, h_ξ and h_η are the entropy rates of the ergodic sequence and their marginals, respectively. With increasing n the bounds become valid for more and more decisive majority of the pairs (f_n, g_n) of colorings, as specified through a notion of a convergence to zero faster than exponentially (f.t.e.).

Using Theorem 2, the generic behaviour of the colorings in terms of entropies of partitions is described in Corollary 2: with increasing n for majority of the pairs (f_n, g_n) of colorings the Shannon entropy of the partition into the (f_n, g_n)-monochromatic rectangles is lower bounded by a sequence na_n such that a_n converges to $h_{\xi\eta}^{\alpha\beta}$. This rate is the highest possible, see Remark 2.

Main tools that underly proofs are collected in Section II. They include a couple of lemmas that go in the spirit back to [1], [3], and their dynamical versions. On the way to the main result, a general asymptotic scheme is singled out in Section III. It comprises crucial arguments towards the proof of Theorem 2 presented in Section IV.

II. Preliminaries

The following lemma is a minor generalization and reformulation of [6, Lemma 6, p. 322]. The starting idea goes back to [1, Lemma 3.1, p. 230], see [6, Remark 4, p. 323]. Results of this sort have been of importance when studying common randomness and secrecy capacities, see [3, Appendix B].

Lemma 1. *Let $k = e^q$ be a positive integer and $\mathcal{Q}$ a finite set of measures on a finite set X such that*

$$Q(x) \leqslant e^{-v} \cdot Q(X)\,, \qquad x \in X, Q \in \mathcal{Q}, \tag{1}$$

for some v. If $w \leqslant q$ then the number of those k-colorings f of X that violate

$$Q(f^{-1}(i)) \leqslant e^{-w} \cdot Q(X)\,, \qquad i \in \widehat{k}, Q \in \mathcal{Q}, \tag{2}$$

is at most

$$|X|^k \cdot k\,|\mathcal{Q}|\, \exp\big[-\tfrac{1}{2}\, e^{v-q}\, (e^{q-w}-1)(q-w)\big]\,.$$

Proof: When $\mathcal{Q}$ is empty or contains only the zero measure then (2) is never violated, and the assertion holds trivially. Otherwise, when $\mathcal{Q}$ consists of probability measures this is a reformulation of [6, Lemma 6] where $\varepsilon = e^{q-w}-1$ is nonnegative because $w \leqslant q$. The general case follows directly from this special one. ■

In Lemma 1, some colorings f partition X into monochromatic blocks $f^{-1}(i)$ such that their measures $Q(f^{-1}(i))$ are bounded from above, uniformly in i and Q, while the number of the remaining colorings has an explicit upper bound. A symmetric counterpart of the lemma will be needed in the sequel. For readers convenience, its proof is worked out based on ideas of the proof of [6, Lemma 6].

Lemma 2. *Under the assumptions of Lemma 1, if $w \geqslant q$ then the number of those k-colorings f of X that violate*

$$Q(f^{-1}(i)) \geqslant e^{-w} \cdot Q(X)\,, \qquad i \in \widehat{k}, Q \in \mathcal{Q}, \tag{3}$$

is at most

$$|X|^k \cdot k\,|\mathcal{Q}|\, \exp\big[-\tfrac{1}{2}\, e^{v-q}\, (1-e^{q-w})\, \ln(2-e^{q-w})\big]\,.$$

Proof: Let the vector $(\eta_x)_{x\in X}$ consist of independent random variables, each one distributed uniformly on $\widehat{k}$. Thus, a realization of the vector is a k-coloring of X. For $i \in \widehat{k}$ and $Q \in \mathcal{Q}$ let ζ_x be equal to $-(k-1)Q(x)$ if $\eta_x = i$ and to $Q(x)$ otherwise, $x \in X$. Then, these variables are independent and centered, that is their expectations are equal to zero. By (1), they are bounded in the absolute value by $e^{q-v}Q(X)$ and the sum of variances $(k-1)\sum_{x\in X} Q(x)^2$ is majorized by $e^{q-v}Q(X)^2$. Hence, for $\varepsilon \geqslant 0$ the inequality

$$Pr\big(\textstyle\sum_{x\in X}\zeta_x > \varepsilon\, Q(X)\big) \leqslant \exp\Big[-\dfrac{\varepsilon}{2\,e^{q-v}}\, \ln(1+\varepsilon)\Big]$$

follows from [6, Lemma 5] whenever $Q(X) > 0$, and holds trivially otherwise. This probability multiplied by $|X|^k$ is equal to the number of k-colorings f that satisfy

$$-\textstyle\sum_{x\in f^{-1}(i)} (k-1)Q(x) + \sum_{x\in X\setminus f^{-1}(i)} Q(x) > \varepsilon\, Q(X)\,.$$

This inequality rewrites to $kQ(f^{-1}(i)) < (1-\varepsilon)Q(X)$. Since w is at least q the choice $\varepsilon = 1-e^{q-w}$ is possible. Therefore, the number of those k-colorings f that violate the inequality $Q(f^{-1}(i)) \geqslant e^{-w} \cdot Q(X)$ is at most

$$|X|^k \cdot \exp\big[-\tfrac{1}{2}\, e^{v-q}\, (1-e^{q-w})\, \ln(2-e^{q-w})\big]\,.$$

This bound does not depend on Q and i whence the assertion follows. ■

Dynamical versions of Lemmas 1 and 2 are prepared below for later purposes.

Lemma 3. *For $n \geqslant 1$ let $\mathcal{Q}_n$ be a set of measures on a finite set X_n such that*

$$Q(x) \leqslant e^{-nr_n} \cdot Q(X_n)\,, \qquad x \in X_n, Q \in \mathcal{Q}_n, \tag{4}$$

for a sequence r_n that converges to a finite limit h. If $\alpha \geqslant 0$, $k_n = \lfloor e^{n\alpha} \rfloor$ and $s_n = \min\{r_n, \alpha, h\} - 2n^{-1/2}$ then the proportion of those k_n-colorings f_n of X_n that violate

$$Q(f_n^{-1}(i)) \leqslant e^{-ns_n} \cdot Q(X_n)\,, \qquad i \in \widehat{k_n}, Q \in \mathcal{Q}_n, \tag{5}$$

is at most

$$e^{n\alpha}\, |\mathcal{Q}_n|\, \exp\big[-\tfrac{1}{2}(e^{\sqrt{n}}-1)\big]\,.$$

Proof: If $q_n = \ln k_n$ then $n\alpha \geqslant q_n \geqslant n\alpha - 1$, and thus $q_n - ns_n \geqslant \sqrt{n}$. Lemma 1 implies that the proportion of those k_n-colorings f_n that violate (5) is upper bounded by

$$e^{n\alpha}\, |\mathcal{Q}_n|\, \exp\big[-\tfrac{1}{2}\, e^{nr_n - q_n}\, (e^{q_n - ns_n}-1)\, \sqrt{n}\,\big]\,.$$

Omitting $\sqrt{n}$, the bracket is dominated by

$$-\tfrac{1}{2}\, e^{n\min\{r_n,\alpha,h\}-n\alpha}\, (e^{q_n-ns_n}-1) \leqslant -\tfrac{1}{2}(e^{\sqrt{n}}-1)$$

whence the assertion follows. ■

Let us say that a sequence of nonnegative numbers p_n grows *at most exponentially* if the sequence $\frac{1}{n}\ln p_n$ is bounded from above. The sequence p_n goes to zero *faster than exponentially (f.t.e.)* if $\frac{1}{n}\ln p_n$ tends to $-\infty$.

Lemma 3 is mostly used in a limiting version that only states existence of a sequence s_n.

Corollary 1. *If a sequence $\mathcal{Q}_n$ satisfies* (4) *with some $r_n \to h$, $|\mathcal{Q}_n|$ grows at most exponentially, $\alpha \geqslant 0$ and $k_n = \lfloor e^{n\alpha} \rfloor$ then there exists a sequence s_n converging to $\min\{\alpha, h\}$ such that the proportion of those k_n-colorings f_n of X_n that violate* (5) *goes to zero f.t.e.*

A counterpart of this corollary is needed as well.

Lemma 4. *Under the same assumptions as in Corollary 1, if $\alpha < h$ then there exists a sequence t_n converging to α such that the proportion of those k_n-colorings f_n of X_n that violate*

$$Q_n(f_n^{-1}(i)) \geqslant e^{-nt_n} \cdot Q_n(X_n)\,, \qquad i \in \widehat{k_n}, Q_n \in \mathcal{Q}_n, \tag{6}$$

goes to zero f.t.e.

Proof: Let $q_n = \ln k_n$ and $|\mathcal{Q}_n| \leqslant e^{nu}$ for some u. If $w_n = n\alpha + 1$ then $w_n \geqslant q_n + 1$ and the sequence $t_n = w_n/n$ converges obviously to α. Hence, Lemma 2 implies that the proportion of those k_n-colorings f_n of X_n that violate (6) is upper bounded by

$$e^{n\alpha}\, e^{nu}\, \exp\big[-\tfrac{1}{2}\, e^{nr_n - n\alpha}\, (1-\tfrac{1}{e})\ln(2-\tfrac{1}{e})\big]\,.$$

This expression vanishes f.t.e. ■

III. Main asymptotic scheme

In this section a general limiting scheme is presented that abstracts some typical situations encountered in bivariate ergodic sequences. This is believed to provide a better insight and, at the same time, to simplify and shorten proofs.

Theorem 1. *For $n \geqslant 1$ let Q_n be a finite measure on a finite set $X_n \times Y_n$ such that the cardinalities of the sets*

$$\{y \in Y_n \colon Q_n(X_n \times \{y\}) \neq 0\} \tag{7}$$

grow at most exponentially. Let h_X, h_Y and h_{XY} be numbers such that $h_{XY} \leqslant h_X + h_Y$, the inequalities

$$Q_n(\{x\} \times Y_n) \leqslant e^{-nr_n} \cdot Q_n(X_n \times Y_n), \quad x \in X_n, \tag{8}$$

$$Q_n(X_n \times \{y\}) \leqslant e^{-ns_n} \cdot Q_n(X_n \times Y_n), \quad y \in Y_n, \tag{9}$$

hold with converging sequences $r_n \to h_X$, $s_n \to h_Y$, and

$$Q_n(x,y) \leqslant e^{-nt_n} \cdot Q_n(X_n \times \{y\}), \ (x,y) \in X_n \times Y_n, \tag{10}$$

be satisfied with a converging sequence $t_n \to h_{XY} - h_Y$. Let further $\alpha, \beta \geqslant 0$, $k_n = \lfloor e^{n\alpha} \rfloor$, $\ell_n = \lfloor e^{n\beta} \rfloor$ and a pair (f_n, g_n) consist of a k_n-coloring f_n of X_n and an ℓ_n-coloring g_n of Y_n.

If $\alpha < h_X$ or $\beta \geqslant h_Y$ then there exists a sequence

$$w_n \to h_{XY}^{\alpha\beta} = \min\{h_{XY}, \alpha + h_Y, h_X + \beta, \alpha + \beta\}$$

such that the proportion of those pairs (f_n, g_n) that violate

$$Q_n(f_n^{-1}(i) \times g_n^{-1}(j)) \leqslant e^{-nw_n} \cdot Q_n(X_n \times Y_n), \quad i \in \widehat{k_n}, j \in \widehat{\ell_n}, \tag{11}$$

goes to zero f.t.e.

Proof: Let $\mathcal{Q}_n^X$ be the family consisting of the single measure on X_n given by $x \mapsto Q_n(\{x\} \times Y_n)$. By (8), Corollary 1 applies to the sequence $\mathcal{Q}_n^X$. There exists a sequence u_n^+ converging to $\min\{\alpha, h_X\}$ such that the proportion of those k_n-colorings f_n of X_n that violate

$$Q_n(f_n^{-1}(i) \times Y_n) \leqslant e^{-nu_n^+} \cdot Q_n(X_n \times Y_n), \quad i \in \widehat{k_n}, \tag{12}$$

goes to zero f.t.e.

Let us consider the set of measures $\mathcal{Q}_n^{X|Y}$ on X_n given by

$$x \mapsto Q_n(x,y), \quad y \in Y_n,$$

that are nonzero. By (7), the cardinality of this set grows at most exponentially. Hence, Corollary 1 based on (10) implies existence of a sequence p_n^+ converging to $\min\{\alpha, h_{XY} - h_Y\}$ such that the proportion of those k_n-colorings f_n that violate

$$Q_n(f_n^{-1}(i) \times \{y\}) \leqslant e^{-np_n^+} \cdot Q_n(X_n \times \{y\}), \quad i \in \widehat{k_n}, y \in Y_n, \tag{13}$$

goes to zero f.t.e.

If a coloring f_n satisfies (13) then by (9)

$$Q_n(f_n^{-1}(i) \times \{y\}) \leqslant e^{-n[s_n + p_n^+]} \cdot Q_n(X_n \times Y_n), \quad i \in \widehat{k_n}, y \in Y_n, \tag{14}$$

where $s_n + p_n^+$ converges to $\min\{\alpha + h_Y, h_{XY}\}$.

1. Let us assume first that $\alpha < h_X$. On account of (8), Lemma 4 applies to the sequence $\mathcal{Q}_n^X$. There exists a sequence u_n^- converging to α such that the proportion of those k_n-colorings f_n that violate

$$Q_n(f_n^{-1}(i) \times Y_n) \geqslant e^{-nu_n^-} \cdot Q_n(X_n \times Y_n), \quad i \in \widehat{k_n}, \tag{15}$$

goes to zero f.t.e.

Let $\mathcal{F}_n$ be the family of those k_n-colorings f_n that violate an inequality in (12), (13), (14) or (15). By three above convergence statements, the proportion $|\mathcal{F}_n||X_n|^{-k_n}$ goes to zero f.t.e. If $f_n \notin \mathcal{F}_n$ then (14) and (15) combine to

$$Q_n(f_n^{-1}(i) \times \{y\}) \leqslant e^{-nv_n} \cdot Q_n(f_n^{-1}(i) \times Y_n), \quad i \in \widehat{k_n}, y \in Y_n, \tag{16}$$

where $v_n = s_n + p_n^+ - u_n^-$ converges to $\min\{h_Y, h_{XY} - \alpha\}$, denoted in the sequel by h_v.

Let $q_n = \min\{v_n, \beta, h_v\} - 2n^{-1/2}$ play the role of s_n from Lemma 3 that is applied to the sets $\mathcal{Q}_{f_n}$ of measures on Y_n given by

$$y \mapsto Q_n(f_n^{-1}(i) \times \{y\}), \quad i \in \widehat{k_n},$$

for k_n-colorings $f_n \notin \mathcal{F}_n$ and $n \geqslant n_0$. By (16), if $\mathcal{G}_{f_n}$ denotes the family of those ℓ_n-colorings g_n of Y_n that violate

$$Q_n(f_n^{-1}(i) \times g_n^{-1}(j)) \leqslant e^{-nq_n} \cdot Q_n(f_n^{-1}(i) \times Y_n), \quad i \in \widehat{k_n}, j \in \widehat{\ell_n}, \tag{17}$$

then the proportion $|\mathcal{G}_{f_n}||Y_n|^{-\ell_n}$ is upper bounded by

$$c_n = e^{n\beta} |\mathcal{Q}_{f_n}| \exp\left[-\tfrac{1}{2}(e^{\sqrt{n}} - 1)\right].$$

Here, obviously $|\mathcal{Q}_{f_n}| \leqslant e^{n\alpha}$.

Therefore, the cardinality of the set

$$\mathcal{H}_n = \{(f_n, g_n) \colon f_n \in \mathcal{F}_n \text{ or } (f_n \notin \mathcal{F}_n \text{ and } g_n \in \mathcal{G}_{f_n})\}$$

is upper bounded by

$$|\mathcal{F}_n||Y_n|^{\ell_n} + \textstyle\sum_{f_n \notin \mathcal{F}_n} |\mathcal{G}_{f_n}|$$

and the proportion $|\mathcal{H}_n||X_n|^{-k_n}|Y_n|^{-\ell_n}$ is at most the sum of $|\mathcal{F}_n||X_n|^{-k_n}$ with c_n. It follows that this proportion goes to zero f.t.e.

Let $w_n = q_n + u_n^+$. This sequence converges to

$$\min\{\alpha + \beta, \alpha + h_Y, h_{XY}\}$$

which equals $h_{XY}^{\alpha\beta}$ because $\alpha < h_X$. If a pair of colorings (f_n, g_n) does not belong to $\mathcal{H}_n$, thus $f_n \notin \mathcal{F}_n$ and $g_n \notin \mathcal{G}_{f_n}$, then (12) and (17) hold. Since (11) is their consequence the proportion of those pairs (f_n, g_n) that violate (11) is upper bounded by $|\mathcal{H}_n||X_n|^{-k_n}|Y_n|^{-\ell_n}$, going to zero f.t.e.

2. It remains to consider $\alpha \geqslant h_X$ and $\beta \geqslant h_Y$. Let $\mathcal{Q}_n^Y$ be the family consisting of the single measure on Y_n given by $y \mapsto Q_n(X_n \times \{y\})$. Corollary 1 can be applied to the sequence $\mathcal{Q}_n^Y$ and β in the role of α. By (9) and $\beta \geqslant h_Y$, there exists a sequence v_n^+ converging to h_Y such that the proportion of those ℓ_n-colorings g_n of Y_n that violate

$$Q_n(X_n \times g_n^{-1}(j)) \leqslant e^{-nv_n^+} \cdot Q_n(X_n \times Y_n), \quad j \in \widehat{\ell_n}, \tag{18}$$

goes to zero f.t.e. If a k_n-coloring f_n satisfies (13) then by the summation of y over $g_n^{-1}(j)$,

$$Q_n(f_n^{-1}(i)\times g_n^{-1}(j)) \leqslant e^{-np_n^+}\cdot Q_n(X_n\times g_n^{-1}(j)),\quad i\in\widehat{k_n}, j\in\widehat{\ell}_n. \tag{19}$$

Here, $p_n^+ \to h_{XY}-h_Y$ because $\alpha \geqslant h_X \geqslant h_{XY}-h_Y$.

Let $w_n = v_n^+ + p_n^+$. This sequence tends to $h_{XY} = h_{XY}^{\alpha\beta}$. Combining (18) and (19) the inequalities (11) follow. Since they are violated only if (13) or (18) fails the proportion of those pairs (f_n, g_n) that violate (11) goes to zero f.t.e. ■

Remark 1. Let $\alpha < h_{XY}-h_Y$, $\beta < h_Y$ and the assumptions of Theorem 1 hold. Using (7) and (10), Lemma 4 applies to the sequence $\mathcal{Q}_n^{X|Y}$ from the previous proof and implies existence of a sequence p_n^- converging to α such that the proportion of those k_n-colorings f_n that violate

$$Q_n(f_n^{-1}(i)\times\{y\}) \geqslant e^{-np_n^-}\cdot Q_n(X_n\times\{y\}),\quad i\in\widehat{k_n}, y\in Y_n, \tag{20}$$

goes to zero f.t.e. Using (9), Lemma 4 applies to the sequence $\mathcal{Q}_n^{Y}$. There exists a sequence v_n^- converging to β such that the proportion of those ℓ_n-colorings g_n that violate

$$Q_n(X_n\times g_n^{-1}(j)) \geqslant e^{-nv_n^-}\cdot Q_n(X_n\times Y_n),\quad j\in\widehat{\ell}_n, \tag{21}$$

goes to zero f.t.e. If f_n satisfies (20) then for any ℓ_n-coloring g_n of Y_n

$$Q_n(f_n^{-1}(i)\times g_n^{-1}(j)) \geqslant e^{-np_n^-}\cdot Q_n(X_n\times g_n^{-1}(j)),\quad i\in\widehat{k_n}, j\in\widehat{\ell}_n, \tag{22}$$

by summing over $y\in g_n^{-1}(j)$. Combining (21) and (22) it follows that the sequence $w_n = p_n^- + v_n^-$ tends to $\alpha+\beta$ and the proportion of those pairs (f_n, g_n) that violate

$$Q_n(f_n^{-1}(i)\times g_n^{-1}(j)) \geqslant e^{-nw_n}\cdot Q_n(X_n\times Y_n), i\in\widehat{k_n}, j\in\widehat{\ell}_n,$$

goes to zero f.t.e. This symmetric counterpart of Theorem 1 is interesting per se but not used below. The assumption (8) was not needed.

IV. Ergodic bivariate sequences

Let (ξ_n, η_n), $n\geqslant 1$, be a bivariate ergodic sequence with the states in a finite product $A\times B$. The distribution of the first n vectors of the sequence is denoted by P_n and its marginals to A^n and B^n by P_n^ξ and P_n^η, respectively. In this section Theorem 1 is applied to the restrictions of P_n to certain subsets of $X_n\times Y_n = A^n\times B^n$ that are constructed by a notion of entropic typicality. The role of h_{XY} is played by the entropy rate $h_{\xi\eta} = \lim \frac{1}{n}\ln H(P_n)$ of the bivariate sequence, and the entropies rates of the marginal sequences, $h_\xi = \lim\frac{1}{n}\ln H(P_n^\xi)$ and $h_\eta = \lim\frac{1}{n}\ln H(P_n^\eta)$, correspond to h_X and h_Y, respectively.

The following theorem asserts rigorously what was earlier mentioned as the generic behavior of pairs of colorings in terms of probabilities of monochromatic rectangles intersected with typical sets.

Theorem 2. *For $n\geqslant 1$ let $k_n = \lfloor e^{n\alpha}\rfloor$ and $\ell_n = \lfloor e^{n\beta}\rfloor$ where $\alpha,\beta\geqslant 0$. There exist a sequence of sets $Z_n\subseteq A^n\times B^n$ and a sequence of numbers w_n such that $P_n(Z_n)\to 1$, $w_n\to h_{\xi\eta}^{\alpha\beta}$ and the proportion of those pairs (f_n, g_n), consisting of a k_n-coloring f_n of A^n and an ℓ_n-coloring g_n of B^n, that violate*

$$P_n\big((f_n^{-1}(i)\times g_n^{-1}(j))\cap Z_n\big) \leqslant e^{-nw_n}, i\in\widehat{k_n}, j\in\widehat{\ell}_n, \tag{23}$$

goes to zero f.t.e.

Proof: By exchanging the coordinate variables ξ_n and η_n if necessary there is no loss of generality when assuming $\alpha < h_\xi$ or $\beta\geqslant h_\eta$. The subset Z_n is constructed below via the entropy-typical sets

$$T_{\xi,\delta}^n = \{x\in A^n\colon P_n^\xi(x)\asymp e^{-n[h_\xi\pm\delta]}\},$$
$$T_{\eta,\delta}^n = \{y\in B^n\colon P_n^\eta(y)\asymp e^{-n[h_\eta\pm\delta]}\},$$
$$T_{\xi\eta,\delta}^n = \{(x,y)\in A^n\times B^n\colon P_n(x,y)\asymp e^{-n[h_{\xi\eta}\pm\delta]}\}.$$

Here, the symbol $\asymp$ means that the number on the left is between the two numbers on the right.

Since the sequence (ξ_n,η_n) and two marginal sequences are ergodic the Shannon-McMillan-Breiman theorem, known also as the asymptotic equipartition property [8, Sections 1.5 and 1.6], implies that for every positive δ the probabilities $P_n^\xi(T_{\xi,\delta}^n)$, $P_n^\eta(T_{\eta,\delta}^n)$ and $P_n(T_{\xi\eta,\delta}^n)$ tend to 1. By standard diagonal arguments, there exists a positive sequence $(\delta_n)_{n\geqslant 1}$ converging to zero, perhaps rather slowly, such that each of the sequences $P_n^\xi(T_{\xi,\delta_n}^n)$, $P_n^\eta(T_{\eta,\delta_n}^n)$ and $P_n(T_{\xi\eta,\delta_n}^n)$ converges to one.

Let U_n denote the intersection of $T_{\xi,\delta_n}^n\times B^n$, $A^n\times T_{\eta,\delta_n}^n$ and $T_{\xi\eta,\delta_n}^n$. If $p_n = 1-[1-P_n(U_n)]^{1/2}$ and

$$Y_n^* = \big\{y\in B^n\colon P_n((A^n\times\{y\})\cap U_n)\geqslant p_n\cdot P_n^\eta(y)\big\}$$

then the estimations

$$\begin{aligned} p_n[1-P_n^\eta(Y_n^*)] &= \textstyle\sum_{y\in B^n\setminus Y_n^*} p_n\cdot P_n^\eta(y)\\ &\geqslant \textstyle\sum_{y\in B^n\setminus Y_n^*} P_n((A^n\times\{y\})\cap U_n)\\ &= P_n(U_n) - P_n((A^n\times Y_n^*)\cap U_n)\\ &\geqslant P_n(U_n) - P_n^\eta(Y_n^*)\end{aligned}$$

and $p_n<1$ imply

$$P_n^\eta(Y_n^*) \geqslant \frac{P_n(U_n)-p_n}{1-p_n} = p_n.$$

Obviously, $P_n^\eta(Y_n^*)=1$ if $p_n=1$.

Let Z_n denote the intersection of $A^n\times Y_n^*$ with U_n. Then it is possible to conclude subsequently that the sequences $P_n(U_n)$, p_n, $P_n^\eta(Y_n^*)$ and $P_n(Z_n)$ converge to one. Let Q_n denote the restriction of P_n to Z_n. To apply Theorem 1, its assumptions are verified as follows.

The set in (7) is contained in $Y_n = B^n$ so that its cardinality grows at most exponential.

Let r_n be equal to $h_\xi - \delta_n + \frac{1}{n}\ln P_n(Z_n)$ provided $P_n(Z_n)$ is positive. If $x\notin T_{\xi,\delta_n}^n$ then $Q_n(\{x\}\times Y_n)$ vanishes and the inequality in (8) holds trivially. Otherwise, if $x\in T_{\xi,\delta_n}^n$ then

$$\begin{aligned} Q_n(\{x\}\times B^n) &\leqslant P_n(\{x\}\times B^n) = P_n^\xi(x)\leqslant e^{-n[h_\xi-\delta_n]}\\ &= e^{-nr_n}\cdot P_n(Z_n) = e^{-nr_n}\cdot Q_n(A^n\times B^n)\end{aligned}$$

so that (8) is verified with a sequence r_n converging to h_ξ.

For s_n defined through $h_\eta - \delta_n + \frac{1}{n}\ln P_n(Z_n)$, a verification of (9) is analogous to that of (8) and omitted here.

Let t_n be equal to $h_{\xi\eta} - h_\eta - 2\delta_n + \frac{1}{n}\ln p_n$ provided p_n is positive. If $(x,y) \notin Z_n$ then $Q_n(x,y)$ vanishes and the inequality in (10) holds trivially. Otherwise, $(x,y) \in T^n_{\xi\eta,\delta_n}$ and $y \in T^n_{\eta,\delta_n}$ imply

$$Q_n(x,y) = P_n(x,y) \leqslant e^{-n[h_{\xi\eta}-\delta_n]} \leqslant e^{-n[h_{\xi\eta}-\delta_n]}\, e^{n[h_\eta+\delta_n]}\, P^\eta_n(y)\,.$$

Using $y \in Y^*_n$ and $Z_n = (A^n \times Y^*_n) \cap U_n$,

$$p_n \cdot P^\eta_n(y) \leqslant P_n((A^n \times \{y\}) \cap U_n) = Q_n(A^n \times \{y\})\,.$$

Combining above estimations, it follows that (10) is verified with a sequence $t_n \to h_{\xi\eta} - h_\eta$. Obviously, $h_{\xi\eta} \leqslant h_\xi + h_\eta$.

Therefore, Theorem 1 implies existence of a sequence v_n converging to $h^{\alpha\beta}_{\xi\eta}$ such that the proportion of those pairs (f_n, g_n) that violate

$$Q_n(f_n^{-1}(i) \times g_n^{-1}(j)) \leqslant e^{-nv_n} \cdot P_n(Z_n)\,, \quad i \in \widehat{k_n}\,, j \in \widehat{\ell_n}\,,$$

goes to zero f.t.e. Writing $w_n = v_n - \frac{1}{n}\ln P_n(Z_n)$, the above inequalities coincide with (23) and $w_n \to h^{\alpha\beta}_{\xi\eta}$. ■

The following consequence of Theorem 2 describes what was alluded to as the generic behavior of the colorings in terms of Shannon entropies of the partitions into rectangles.

Corollary 2. *If $\mathsf{H}(P_n|f_n, g_n)$ denotes the Shannon entropy under P_n of the partition of $A^n \times B^n$ into the (f_n, g_n)-monochromatic rectangles $f_n^{-1}(i) \times g_n^{-1}(j)$, $i \in \widehat{k_n}$, $j \in \widehat{\ell_n}$, then there exists a sequence a_n converging to $h^{\alpha\beta}_{\xi\eta}$ such that the proportion of those pairs (f_n, g_n) of colorings that violate $\frac{1}{n}\mathsf{H}(P_n|f_n, g_n) \geqslant a_n$ goes to zero f.t.e.*

Proof: If $Z^c_n = (A^n \times B^n) \setminus Z_n$ then

$$\textstyle\sum_D \big[P_n(D \cap Z_n) \ln \frac{P_n(D\cap Z_n)}{P_n(D)P_n(Z_n)} + P_n(D \cap Z^c_n) \ln \frac{P_n(D\cap Z^c_n)}{P_n(D)P_n(Z^c_n)}\big]$$

is nonnegative by convexity. Here, the summation runs over the (f_n, g_n)-monochromatic rectangles D. This implies

$$\mathsf{H}(P_n|f_n, g_n) + \ln 2 \geqslant -\textstyle\sum_D P_n(D \cap Z_n) \ln P_n(D \cap Z_n)\,.$$

By Theorem 2, for a sequence w_n converging to $h^{\alpha\beta}_{\xi\eta}$, the summand on the right is majorized by $-nw_nP_n(D \cap Z_n)$. Taking $a_n = w_nP_n(Z_n) - \frac{1}{n}\ln 2$ the assertion follows. ■

Remark 2. If the assertion of Corollary 2 holds for a sequence a_n converging to some a instead of $h^{\alpha\beta}_{XY}$ then a cannot exceed this number. This follows from

$$\limsup_{n\to\infty}\ \sup_{f_n,g_n} \tfrac{1}{n} H(P_n|f_n, g_n) \leqslant h^{\alpha,\beta}_{X,Y}\,. \tag{24}$$

To prove the inequality, $H(P_n|f_n, g_n)$ is majorized by the sum of $H(P^\xi_n|f_n)$ and $H(P^\eta_n|g_n)$, defined analogously. The former summand is dominated by $H(P^\xi_n)$ and $n\alpha$ because the partition of A^n into $f_n^{-1}(i)$, $i \in \widehat{k_n}$, has at most k_n blocks. Similarly, the latter summand is dominated by $H(P^\eta_n)$ and $n\beta$.

It follows that the left-hand side of (24) is at most

$$\min\{h_\xi, \alpha\} + \min\{h_\eta, \beta\}\,.$$

This and $H(P_n|f_n, g_n) \leqslant H(P_n)$ imply (24).

Acknowledgment

This work was partially supported by Grant Agency of Academy of Sciences of the Czech Republic, Grants IAA 100750603 and KJB 100750901, and by Grant Agency of the Czech Republic, Grant 201/08/0539.

References

[1] R. Ahlswede and I. Csiszár (1998) Common randomness in information theory and cryptography – Part II: CR Capacity. *IEEE Trans. Information Theory* **44** 225–240.

[2] I. Csiszar and J. Körner (1981) *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Academic Press, New York (Akademiai Kiadó, Budapest).

[3] I. Csiszar and P. Narayan (2004) Secrecy capacities for multiple terminals. *IEEE Trans. Information Theory* **50** 3047–3061.

[4] R. Dougherty, C. Freiling, and K. Zeger (2006) Six New Non-Shannon Information Inequalities. *Proc. IEEE International Symposium on Information Theory,* Seattle, Washington, 233–236.

[5] K. Makarychev, Yu. Makarychev, A. Romashchenko and N. Vereshchagin (2002) A new class of non-Shannon-type inequalities for entropies. *Communications in Information and Systems* **2** 147–166.

[6] F. Matúš (2007) Two constructions on limits of entropy functions. *IEEE Trans. Information Theory* **53** 320–330.

[7] F. Matúš (2007) Infinitely many information inequalities. *Proc. IEEE International Symposium on Information Theory*, Nice, France, 41–44.

[8] P. Shields (1991) The Ergodic Theory of Discrete Sample Path, *Graduate Studies in Mathematics*, American Mathematical Society **13**.

[9] R.W. Yeung (2002) *A First Course in Information Theory*. Kluwer Academic/Plenum Publishers, New York.

[10] Z. Zhang and R.W. Yeung (1998) On characterization of entropy function via information inequalities. *IEEE Trans. Information Theory* **44** 1440–1452.