THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 19, Number 1/2018, pp. 3–10

ENTROPY AND DIVERGENCE RATES FOR MARKOV CHAINS: II. THE WEIGHTED CASE

Vlad Stefan BARBU [1], Alex KARAGRIGORIOU [2], Vasile PREDA [3]

[1] LMRS, Université de Rouen, France; barbu@univ-rouen.fr
[2] University of the Aegean, Department of Mathematics, Greece; alex.karagrigoriou@aegean.gr
[3] University of Bucharest and ISMMA of the Romanian Academy, Romania; vasilepreda0@gmail.com
Corresponding author: Vlad Stefan BARBU, Université de Rouen,
Laboratoire de Mathématiques Raphaël Salem, UMR 6085,
Avenue de l'Université, BP.12, F76801 Saint-Étienne-du-Rouvray, France, E-mail: barbu@univ-rouen.fr

Abstract. In this work we consider weighted versions of the generalizations of Alpha divergence measure (Chernoff [8], Amari and Nagaoka [1]) and Beta divergence measures (Basu *et al*. [4]) for Markov chains and investigate their limiting behavior. This is a continuation of the developments presented in [2], applying the notion of weighted divergence measure (Beliş and Guiaşu [5], Guiaşu [11], Kapur [12]). This work is continued in [3], where we present generalized Cressie and Read power divergence class of measures and numerically investigate some properties of all these generalized divergence measures and rates.

Key words: divergence measures, information measures, Markov chains, entropy, divergence rates, weighted divergence measures.

1. PRELIMINARIES

In this work we focus on some generalizations of Alpha divergence measure (Chernoff [8], Amari and Nagaoka [1]) and Beta divergence measures (Basu *et al*. [4]) and rates for Markov chains presented in [2]. For these divergences and rates we consider the associated weighted versions, following the concepts introduced by Beliş and Guiaşu [5], Guiaşu [11].

As the results on information measures consider only the probability mass function or the probability density function of a random variable without taking into account its value, Beliş and Guiaşu [5] highlighted the importance of integrating the quantitative, objective and probabilistic concepts of information with the qualitative, subjective and non-stochastic concept of utility. By considering the two basic concepts of objective probability and subjective utility, they introduced the concept of weighted entropy and thus constructed a shift-dependent information measure with properties similar to those of the Shannon entropy. Afterward Guiaşu [11] characterized axiomatically the weighted entropy measure. Di Crescenzo and Longobardi [9] introduced the concept of weighted residual entropy and weighted past entropy. Many other researchers investigated different aspects and generalizations of weighted entropies and divergences; among them, we can cite Bhullar *et al*. [6], Casquilho [7], Dial and Taneja [10], Kapur [12], Sharma *et al*. [13], Śmieja [14], Suhov *et al*. [15, 16], Taneja [17], Taneja and Tuteja [18, 19].

The paper is organized as follows. In the rest of this section we recall from [2] the definitions of generalized Alpha and Beta divergence measures; then, in Section 2 the corresponding weighted generalized divergence measures are presented and the associated rates are obtained.

Let $(X_n)_{n\in N}$ be an ergodic time-homogeneous Markov chain with finite state space $\chi = \{1, \ldots, M\}$ For this Markov chain, we consider two different probability laws. Under the first law, let $p_i = P(X_1 = i)$, $i \in \chi$ denote the initial distribution of the chain and $p_{ij} = P(X_{k+1} = j \mid X_k = i)$, $i, j \in \chi$ the associated transition probabilities. Let also $\mathbf{p}_n$ denote the joint probability distribution of $(X_1, X_2, \ldots, X_n)$, *i.e.*, $\mathbf{p}_n(i_{1:n}) = p_{i_1} p_{i_1 i_2} \ldots p_{i_{n-1} i_n}$, $i_1, \ldots, i_n \in \chi$, were we denoted by $i_{1:n}$ the n-tuple $(i_1, \ldots, i_n) \in \chi^n$. Similarly we define under the second law qi, qij, $\mathbf{q}_n(i_{1:n})$ and $\mathbf{q}_n$. Under this setting of finite state space Markov chains, the Alpha-

Gamma measure between the two models is defined as the Alpha-Gamma measure between the two joint probability distributions $\mathbf{p}_n$ and $\mathbf{q}_n$ (cf. [2]), and is written under the normalized form as

$$D_{AG}(\mathbf{p}_n,\mathbf{q}_n)=\frac{1}{\alpha(\alpha-1)}\log\left(\sum_{i_{1:n}\in\chi^n}\tilde{\mathbf{p}}_n^{\alpha}(i_{1:n})\tilde{\mathbf{q}}_n^{1-\alpha}(i_{1:n})\right), \tag{1}$$

where

$$\tilde{\mathbf{p}}_n=\tilde{p}_{i_1}\tilde{p}_{i_1i_2}\ldots\tilde{p}_{i_{n-1}i_n},\tilde{\mathbf{q}}_n=\tilde{q}_{i_1}\tilde{q}_{i_1i_2}\ldots\tilde{q}_{i_{n-1}i_n},$$

with $\tilde{p}_i$, $\tilde{p}_{ij}$, $\tilde{q}_i$ and $\tilde{q}_{ij}$, $i,j\in\chi$, defined by

$$\tilde{p}_i=\frac{p_i}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{p}_n(i_{1:n})\right)^{1/n}},\tilde{p}_{ij}=\frac{p_{ij}}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{p}_n(i_{1:n})\right)^{1/n}}, \tag{2}$$

$$\tilde{q}_i=\frac{q_i}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{q}_n(i_{1:n})\right)^{1/n}},\tilde{q}_{ij}=\frac{q_{ij}}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{q}_n(i_{1:n})\right)^{1/n}}. \tag{3}$$

Similarly, the Beta-Gamma measure between the two Markov models is defined by (cf. [2])

$$D_{BG}(\mathbf{p}_n,\mathbf{q}_n)=-\frac{1}{\alpha}\log\left(\sum_{i_{1:n}\in\chi^n}\tilde{\mathbf{p}}_n(i_{1:n})\tilde{\mathbf{q}}_n^{\alpha}(i_{1:n})\right), \tag{4}$$

where

$$\tilde{\mathbf{p}}_n=\tilde{p}_{i_1}\tilde{p}_{i_1i_2}\ldots\tilde{p}_{i_{n-1}i_n},\tilde{\mathbf{q}}_n=\tilde{q}_{i_1}\tilde{q}_{i_1i_2}\ldots\tilde{q}_{i_{n-1}i_n},$$

with $\tilde{p}_i$, $\tilde{p}_{ij}$, $\tilde{q}_i$ and $\tilde{q}_{ij}$, $i,j\in\chi$, defined by

$$\tilde{p}_i=\frac{p_i}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{p}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}},\tilde{p}_{ij}=\frac{p_{ij}}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{p}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}}, \tag{5}$$

$$\tilde{q}_i=\frac{q_i}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{q}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}},\tilde{q}_{ij}=\frac{q_{ij}}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{q}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}}. \tag{6}$$

2. WEIGHTED ALPHA AND BETA DIVERGENCE RATES FOR MARKOV CHAINS

In this section we first recall the notions of weighted entropies and divergence measures and then introduce new concepts of weighted Alpha and Beta divergences. Then these measures will be defined for finite Markov chains and the corresponding rates will be obtained.

As mentioned in the Introduction, Beliş and Guiaşu [5] introduced the concept of weighted entropy and Guiaşu [11] characterized it axiomatically. Let $\mathbf{p}=(p_1,\ldots,p_n)$ be a finite probability distribution corresponding to n possible states or outcomes and $\mathbf{w}=(w_1,\ldots,w_n)$, be a vector of weights associated with these states, $w_i\geq 0$, $i=1,\ldots,n$.

Definition 1 (cf. [11]). The weighted Shannon entropy measure is defined by

$$I^S(\mathbf{p};\mathbf{w})=-\sum_{i=1}^{n}w_ip_i\log(p_i). \tag{7}$$

When considering an absolutely continuous probability measure μ_p with a density p with respect to a certain measure μ and a weight function w assumed to be measurable and positive, then the weighted Shannon entropy measure is defined by

$$I^S(p;w) = -\int w(x)p(x)\log(p(x))\mathrm{d}\mu(x). \tag{8}$$

Let us consider $\mathbf{p} = (p_1,\ldots,p_n)$ and $\mathbf{q} = (q_1,\ldots,q_n)$ two finite probability distributions corresponding to n possible states and let $\mathbf{w} = (w_1,\ldots,w_n)$ be a vector of weights associated with these states, $w_i \geq 0, i = 1,\ldots,n,$ with $w_i \neq 0$ for some *i*. It is clear that for real applications, in fact $w_i > 0$ for all i, which will be assumed all along this paper.
Note that the weighted Shannon divergence measure (relative entropy) between **p** and **q** can be defined as follows.

Definition 2 (cf. [18]). The weighted Shannon divergence measure between **p** and **q** is given by

$$D^S(\mathbf{p},\mathbf{q};\mathbf{w}) = \sum_{i=1}^{n} w_i p_i \log\left(\frac{p_i}{q_i}\right). \tag{9}$$

Kapur [12] defined axiomatically the weighted directed divergence concept. Let us consider $\mathbf{p} = (p_1,\ldots,p_n)$ and $\mathbf{q} = (q_1,\ldots,q_n)$ two finite probability distributions corresponding to n possible states and let $\mathbf{w} = (w_1,\ldots,w_n)$ be a vector of weights associated with these states, $w_i \geq 0,\ i = 1,\ldots,n.$

Definition 3 (cf. [12]). A measure D(**p**, **q**, **w**) is said to be an appropriate measure of weighted directed divergence if the following axioms are fulfilled:

1. It is a continuous function of $(p_1,\ldots,p_n),\ (q_1,\ldots,q_n)$ and $(w_1,\ldots,w_n).$
2. It is permutationally symmetric, i.e. it does not change when the triplets $(p_1,q_1,w_1),\ (p_2,q_2,w_2),\ldots,\ (p_n,q_n,w_n)$ are permuted among themselves.
3. It is always non-negative and vanishes when $p_i = q_i$ for all $i = 1,\ldots,n.$
4. It is a convex function of $(p_1,\ldots,p_n),$ which has its minimum value zero when $p_i = q_i$ for all $i = 1,\ldots,n.$
5. It reduces to a positive multiple of an ordinary measure of weighted directed divergence when all the weights are equal.

Note that the weighted Shannon divergence introduced in Definition 2 does not verify Condition 3 of the Kapur's definition of a weighted directed divergence. For this reason, Kapur [12] proposed another weighted Shannon divergence measure which satisfies all the above set of axioms of a weighted directed divergence.

Definition 4 (cf. [12]). The weighted Shannon directed divergence measure corresponding to the Kullback-Leibler measure is given by

$$D_0^S(\mathbf{p},\mathbf{q};\mathbf{w}) = \sum_{i=1}^{n} w_i \left[p_i \log\left(\frac{p_i}{q_i}\right) - p_i + q_i \right]. \tag{10}$$

Note that the weighted Shannon directed divergence measure given in the above definition verifies Condition 3 of the Kapur's definition of a weighted directed divergence; indeed, one has to factorize by p_i the factor $\left[p_i \log\left(\frac{p_i}{q_i}\right) - p_i + q_i \right]$ in (10) and then to study the sign of the function $f(t) := \log(t) - t + 1.$

We will introduce now the concepts of weighted Alpha and Beta divergence measures, as well as the weighted Alpha-Gamma and Beta-Gamma divergence measures.

Let us consider two absolutely continuous probability measures μ_p and μ_q with corresponding densities p and q with respect to a certain measure μ and a weight function w assumed to be measurable and positive.

Definition 5. The weighted Alpha and Beta divergence measures are given by

$$D_A^w\left(p,q;w\right)=\frac{1}{\alpha(\alpha-1)}\left(\int w(x)p^{\alpha}(x)q^{1-\alpha}(x)-\alpha w(x)p(x)+(\alpha-1)w(x)q(x)\right)\mathrm{d}\,\mu(x);$$

$$D_B^w\left(p,q;w\right)=\frac{1}{\alpha+1}\left(\int w(x)q^{1+\alpha}(x)-\left(1+\frac{1}{\alpha}\right)w(x)p(x)q^{\alpha}(x)+\frac{1}{\alpha}w(x)p^{1+\alpha}(x)\right)\mathrm{d}\,\mu(x).$$

Note that

$$D_A^w\left(p,q;w\right)=D_A\left(p,q\right), \tag{11}$$

where $p(x):=w(x)p(x)$ and $q(x):=w(x)q(x)$.

Note also that

$$D_B^w\left(p,q;w\right)=D_B\left(\hat{p},\hat{q}\right), \tag{12}$$

where $\hat{p}(x):=w^{\frac{1}{1+\alpha}}(x)p(x)$ and $\hat{q}(x):=w^{\frac{1}{1+\alpha}}(x)q(x)$.

Transforming the weighted Alpha divergence $D_A^w\left(p,q;w\right)$ given in Definition 1 (as done in [2] for the Alpha divergence), we obtain the weighted Alpha-Gamma divergence measure

$$\begin{aligned}D_{AG}^w(p,q;w)&=\frac{1}{\alpha(\alpha-1)}\log\left(\frac{\int p^{\alpha}(x)q^{1-\alpha}(x)\mathrm{d}\,\mu(x)}{\left(\int p(x)\mathrm{d}\,\mu(x)\right)^{\alpha}\left(\int q(x)\mathrm{d}\,\mu(x)\right)^{1-\alpha}}\right)=D_{AG}(p,q)\\&=\frac{1}{\alpha(\alpha-1)}\log\left(\int\tilde{p}^{\alpha}(x)\tilde{q}^{1-\alpha}(x)\mathrm{d}\,\mu(x)\right),\end{aligned} \tag{13}$$

where we have set $\tilde{p}\left(x\right)=\dfrac{p\left(x\right)}{\int p\left(x\right)\mathrm{d}\mu\left(x\right)}$ and $\tilde{q}\left(x\right)=\dfrac{q\left(x\right)}{\int q\left(x\right)\mathrm{d}\mu\left(x\right)}$.

Similarly, transforming the weighted Beta divergence $D_B^w\left(p,q;w\right)$ given in Definition 1, we obtain the weighted Beta-Gamma divergence measure given by

$$D_{BG}^w(p,q;w)=-\frac{1}{\alpha}\log\left(\int\tilde{\hat{p}}(x)\tilde{\hat{q}}^{\alpha}(x)\right)\mathrm{d}\,\mu(x)=D_{BG}(\hat{p},\hat{q}), \tag{14}$$

where we have set

$$\tilde{\hat{p}}(x)=\frac{\hat{p}(x)}{\left(\int\hat{p}^{1+\alpha}(x)d\,\mu(x)\right)^{1/(1+\alpha)}}\ \text{and}\ \tilde{\hat{q}}(x)=\frac{\hat{q}(x)}{\left(\int\hat{q}^{1+\alpha}(x)d\,\mu(x)\right)^{1/(1+\alpha)}}.$$

Note that p and q, respectively $\hat{p}$ and $\hat{q}$, can be seen as the densities (or mass functions in the discrete case) of corresponding measures μ_p and μ_q, respectively $\mu_{\hat{p}}$ and $\mu_{\hat{q}}$, with respect to a certain measure μ. Thus the Alpha and Beta divergence measures $D_A(p,q)$ and $D_B(\hat{p},\hat{q})$ that appear in (1) and (2), as well as the Alpha-Gamma and Beta-Gamma divergence measures $D_{AG}(p,q)$ and $D_{BG}(\hat{p},\hat{q})$ that appear in (13) and (14) are well defined.

Let us now focus on weighted divergence measures for Markov chains. We place ourselves in the same framework as in the previous section: $(X_n)_{n\in N}$ is an ergodic time-homogeneous Markov chain with finite state space $\chi = \{1,\ldots, M\}$ and we consider two different probability laws for this chain, p_i, p_{ij}, $\mathbf{p}_n(i_{1:n})$, $\mathbf{p}_n$ are the corresponding probabilities under the first law, while q_i, q_{ij}, $\mathbf{q}_n(i_{1:n})$, $\mathbf{q}_n$ are the same quantities under the second law. Let us also consider $\mathbf{W}(n) = \{\mathbf{w}(i_1, \ldots, i_n) \mid (i_1,\ldots,i_n) \in \chi^n\}$, $n \in N^*$, be a set of weights associated with the states χ^n, $\mathbf{w}(i_1, \ldots, i_n) > 0$; as previously, we denote $\mathbf{w}(i_1, \ldots, i_n)$ by $\mathbf{w}(i_{1:n})$ Let us consider $\mathbf{p}_n(i_{1:n}) := \mathbf{w}(i_{1:n})\mathbf{p}_n(i_{1:n})$, $\mathbf{q}_n(i_{1:n}) := \mathbf{w}(i_{1:n})\mathbf{q}_n(i_{1:n})$, $\hat{\mathbf{p}}_n(i_{1:n}) := \mathbf{w}(i_{1:n})^{\frac{1}{1+\alpha}}\mathbf{p}_n(i_{1:n})$ and $\hat{\mathbf{q}}_n(i_{1:n}) := \mathbf{w}(i_{1:n})^{\frac{1}{1+\alpha}}\mathbf{q}_n(i_{1:n})$.

Under these notations, taking into account the weighted Alpha-Gamma divergence measure introduced in (13) and the weighted Beta-Gamma divergence measure introduced in (14), we define the weighted Alpha-Gamma and Beta-Gamma divergence measures for finite Markov models.

Definition 6.

1. The weighted Alpha-Gamma measure between the two Markov models is defined as the weighted Alpha-Gamma measure between the two joint probability distributions $\mathbf{p}_n$ and $\mathbf{q}_n$, that is

$$D_{AG}^{w}(\mathbf{p}_n, \mathbf{q}_n; \mathbf{w}) = D_{AG}(\mathbf{p}_n, \mathbf{q}_n)$$

$$= \frac{1}{\alpha(\alpha-1)} \log\left(\sum_{i_{1:n}\in\chi^n} \frac{\mathbf{p}_n^{\alpha}(i_{1:n})}{\left(\sum_{i_{1:n}\in\chi^n} \mathbf{p}_n(i_{1:n})\right)^{\alpha}} \times \frac{\mathbf{q}_n^{1-\alpha}(i_{1:n})}{\left(\sum_{i_{1:n}\in\chi^n} \mathbf{q}_n(i_{1:n})\right)^{1-\alpha}} \right). \tag{15}$$

2. The weighted Beta-Gamma measure between the two Markov models is defined as the weighted Beta-Gamma measure between the two joint probability distributions $\mathbf{p}_n$ and $\mathbf{q}_n$, that is

$$D_{BG}^{w}(\mathbf{p}_n, \mathbf{q}_n; \mathbf{w}) = D_{BG}(\hat{\mathbf{p}}_n, \hat{\mathbf{q}}_n) \tag{16}$$

$$= -\frac{1}{\alpha} \log\left(\sum_{i_{1:n}\in\chi^n} \frac{\hat{\mathbf{p}}_n(i_{1:n})}{\left(\sum_{i_{1:n}\in\chi^n} \hat{\mathbf{p}}_n^{1+\alpha}(i_{1:n})\right)^{1/(1+\alpha)}} \times \left(\frac{\hat{\mathbf{q}}_n(i_{1:n})}{\left(\sum\limits_{i_{1:n}\in\chi^n} \hat{\mathbf{q}}_n^{1+\alpha}(i_{1:n})\right)^{1/(1+\alpha)}} \right)^{\alpha} \right).$$

We will consider two particular cases of weights, that could be of interest in practice. The first one is

$$\mathbf{w}(i_1,\ldots,i_n) := w_{i_1} w_{i_2} \ldots w_{i_n}, \tag{17}$$

where $\mathbf{w} = (w_1,\ldots,w_M)$ is a vector of weights associated with the states $\chi = \{1,\ldots, M\}$, $w_i > 0$, $i = 1,\ldots,M$. Here w_i is considered as a weight associated to the state i, and relation (17) means that the weights are independent. The second case considered is

$$\mathbf{w}(i_1,\ldots,i_n) := w_{i_1}^1 w_{i_1 i_2}^2 \ldots w_{i_{n-1} i_n}^2, \tag{18}$$

where $\mathbf{w}^1 = (w_1^1,\ldots,w_M^1)$ is a vector of weights associated with the states $\chi = \{1,\ldots, M\}$, $w_i^1 \geq 0$, $i = 1,\ldots,M$, while $\mathbf{w}^2 = (w_{ij}^2)_{i,j\in\chi}$ is a system of weights associated with $\chi\times\chi$, $w_{ij}^2 \geq 0, i, j \in \chi$. Here w_{ij}^2

is considered as a weight associated to the couple of state (i,j) and relation (18) means that the weights have a dependence of a Markov type.

PROPOSITION 1.

1. *Under the first particular case of weights given in (7)*:

• *The weighted Alpha-Gamma measure between the two Markov models can be written under the normalized form*

$$D_{AG}^{w}(\mathbf{p}_n,\mathbf{q}_n;\mathbf{w})=\frac{1}{\alpha(\alpha-1)}\log\left(\sum_{i_{1:n}\in\chi^n}\tilde{\mathbf{p}}_n^{\alpha}(i_{1:n})\tilde{\mathbf{q}}_n^{1-\alpha}(i_{1:n})\right), \tag{19}$$

where

$$\tilde{\mathbf{p}}_n=\tilde{p}_{i_1}\tilde{p}_{i_1i_2}\cdots\tilde{p}_{i_{n-1}i_n},\tilde{\mathbf{q}}_n=\tilde{q}_{i_1}\tilde{q}_{i_1i_2}\cdots\tilde{q}_{i_{n-1}i_n},$$

with $\tilde{p}_i$, $\tilde{p}_{ij}$, $\tilde{q}_i$ and $\tilde{q}_{ij}$, $i,j\in\chi$, *defined by*

$$\tilde{p}_i=\frac{p_iw_i}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{p}_n(i_{1:n})\right)^{1/n}},\tilde{p}_{ij}=\frac{p_{ij}w_j}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{p}_n(i_{1:n})\right)^{1/n}}, \tag{20}$$

$$\tilde{q}_i=\frac{q_iw_i}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{q}_n(i_{1:n})\right)^{1/n}},\tilde{q}_{ij}=\frac{q_{ij}w_j}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{q}_n(i_{1:n})\right)^{1/n}}. \tag{21}$$

• *The weighted Beta-Gamma measure between the two Markov models can be written under the normalized form*

$$D_{BG}^{w}(\mathbf{p}_n,\mathbf{q}_n;\mathbf{w})=-\frac{1}{\alpha}\log\left(\sum_{i_{1:n}\in\chi^n}\tilde{\tilde{\mathbf{p}}}_n(i_{1:n})\tilde{\tilde{\mathbf{q}}}_n^{\alpha}(i_{1:n})\right), \tag{22}$$

where

$$\tilde{\tilde{\mathbf{p}}}_n=\tilde{\tilde{\mathbf{p}}}_{i_1}\tilde{\tilde{\mathbf{p}}}_{i_1i_2}\cdots\tilde{\tilde{\mathbf{p}}}_{i_{n-1}i_n},\tilde{\tilde{\mathbf{q}}}_n=\tilde{\tilde{\mathbf{q}}}_{i_1}\tilde{\tilde{\mathbf{q}}}_{i_1i_2}\cdots\tilde{\tilde{\mathbf{q}}}_{i_{n-1}i_n},$$

with $\tilde{\tilde{\mathbf{p}}}_i$, $\tilde{\tilde{\mathbf{p}}}_{ij}$, $\tilde{\tilde{\mathbf{q}}}_i$ and $\tilde{\tilde{\mathbf{q}}}_{ij}$, $i,j\in\chi$, *defined by*

$$\tilde{\tilde{\mathbf{p}}}_i=\frac{p_iw_i}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{p}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}},\tilde{\tilde{\mathbf{p}}}_{ij}=\frac{p_{ij}w_j}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{p}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}}, \tag{23}$$

$$\tilde{\tilde{\mathbf{q}}}_i=\frac{q_iw_i}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{q}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}},\tilde{\tilde{\mathbf{q}}}_{ij}=\frac{q_{ij}w_j}{\left(\sum_{i_{1:n}\in\chi^n}\mathbf{q}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}}. \tag{24}$$

2. *Under the second particular case of weights given in (18)*:

• *The weighted Alpha-Gamma measure between the two Markov models can be written under the normalized form*

$$D_{AG}^{w}(\mathbf{p}_n,\mathbf{q}_n;\mathbf{w})=\frac{1}{\alpha(\alpha-1)}\log\left(\sum_{i_{1:n}\in\chi^n}\tilde{\mathbf{p}}_n^{\alpha}(i_{1:n})\tilde{\mathbf{q}}_n^{1-\alpha}(i_{1:n})\right), \tag{25}$$

where

$$\tilde{\mathbf{p}}_n = \tilde{p}_{i_1}\tilde{p}_{i_1 i_2}\ldots\tilde{p}_{i_{n-1}i_n}, \tilde{\mathbf{q}}_n = \tilde{q}_{i_1}\tilde{q}_{i_1 i_2}\ldots\tilde{q}_{i_{n-1}i_n},$$

with $\tilde{p}_i$, $\tilde{p}_{ij}$, $\tilde{q}_i$ and $\tilde{q}_{ij}$, $i, j \in \chi$, *defined by*

$$\tilde{p}_i = \frac{p_i w_i^1}{\left(\sum_{i_{1:n}\in\chi^n} \mathbf{p}_n(i_{1:n})\right)^{1/n}}, \tilde{p}_{ij} = \frac{p_{ij} w_{ij}^2}{\left(\sum_{i_{1:n}\in\chi^n} \mathbf{p}_n(i_{1:n})\right)^{1/n}}, \tag{26}$$

$$\tilde{q}_i = \frac{q_i w_i^1}{\left(\sum_{i_{1:n}\in\chi^n} \mathbf{q}_n(i_{1:n})\right)^{1/n}}, \tilde{q}_{ij} = \frac{q_{ij} w_{ij}^2}{\left(\sum_{i_{1:n}\in\chi^n} \mathbf{q}_n(i_{1:n})\right)^{1/n}}. \tag{27}$$

- *The weighted Beta-Gamma measure between the two Markov models can be written under the normalized form*

$$D_{BG}^{w}(\mathbf{p}_n, \mathbf{q}_n; \mathbf{w}) = -\frac{1}{\alpha}\log\left(\sum_{i_{1:n}\in\chi^n} \tilde{\tilde{\mathbf{p}}}_n(i_{1:n})\tilde{\tilde{\mathbf{q}}}_n^{\alpha}(i_{1:n})\right), \tag{28}$$

where

$$\tilde{\tilde{\mathbf{p}}}_n = \tilde{\tilde{\mathbf{p}}}_{i_1}\tilde{\tilde{\mathbf{p}}}_{i_1 i_2}\cdots\tilde{\tilde{\mathbf{p}}}_{i_{n-1}i_n}, \tilde{\tilde{\mathbf{q}}}_n = \tilde{\tilde{\mathbf{q}}}_{i_1}\tilde{\tilde{\mathbf{q}}}_{i_1 i_2}\cdots\tilde{\tilde{\mathbf{q}}}_{i_{n-1}i_n},$$

with $\tilde{\tilde{\mathbf{p}}}_i$, $\tilde{\tilde{\mathbf{p}}}_{ij}$, $\tilde{\tilde{\mathbf{q}}}_i$ and $\tilde{\tilde{\mathbf{q}}}_{ij}$, $i, j \in \chi$, *defined by*

$$\tilde{\tilde{\mathbf{p}}}_i = \frac{p_i w_i^1}{\left(\sum_{i_{1:n}\in\chi^n} \mathbf{p}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}}, \tilde{\tilde{\mathbf{p}}}_{ij} = \frac{p_{ij} w_{ij}^2}{\left(\sum_{i_{1:n}\in\chi^n} \mathbf{p}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}}, \tag{29}$$

$$\tilde{\tilde{\mathbf{q}}}_i = \frac{q_i w_i^1}{\left(\sum_{i_{1:n}\in\chi^n} \mathbf{q}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}}, \tilde{\tilde{\mathbf{q}}}_{ij} = \frac{q_{ij} w_{ij}^2}{\left(\sum_{i_{1:n}\in\chi^n} \mathbf{q}_n^{1+\alpha}(i_{1:n})\right)^{1/n(1+\alpha)}}. \tag{30}$$

The following result gives the divergence rates of weighted Alpha-Gamma and Beta-Gamma measures, with the weights of the types given in (17) and (18). The proof is similar as the proof of Theorems 1 and 2 from [2] and will not be provided here.

THEOREM 1. *Under the setting of the present section, we have*:

(a)
$$\lim_{n\to\infty}\frac{1}{n} D_{AG}^{w}(\mathbf{p}_n, \mathbf{q}_n; \mathbf{w}) = -\frac{1}{\alpha}\log\lambda(\alpha),$$

where $\lambda(\alpha) := \lim_{n\to\infty}\lambda_n(\alpha)$ (*assumed to exist*), *where* $\lambda_n(\alpha)$ *is the largest positive eigenvalue of* $\tilde{R}(n) = (\tilde{r}_{ij}(\alpha))_{i,j\in\chi}$, *where* $\tilde{r}_{ij}(\alpha) = \tilde{\mathbf{p}}_{ij}\tilde{\mathbf{q}}_{ij}^{\alpha}$, *with* $\tilde{\mathbf{p}}_{ij}$ and $\tilde{\mathbf{q}}_{ij}$ *defined in Equations (20) and (21), respectively, if considering the first type of weights given in (17), or in Equations (26) and (27), respectively, if considering the second type of weights given in (18).*

(b)
$$\lim_{n\to\infty}\frac{1}{n} D_{BG}^{w}(\mathbf{p}_n, \mathbf{q}_n; \mathbf{w}) = -\frac{1}{\alpha}\log\lambda(\alpha),$$

where $\lambda(\alpha) := \lim_{n\to\infty}\lambda_n(\alpha)$ *(assumed to exist), where* $\lambda_n(\alpha)$ *is the largest positive eigenvalue of*

$\tilde{R}(n) = (\tilde{\hat{r}}_{ij}(\alpha))_{i,j\in\chi}$, *where* $\tilde{\hat{r}}_{ij}(\alpha) = \tilde{\hat{\mathbf{p}}}_{ij}\tilde{\hat{\mathbf{q}}}_{ij}^{\alpha}$, *with* $\tilde{\hat{\mathbf{p}}}_{ij}$ *and* $\tilde{\hat{\mathbf{q}}}_{ij}$ *defined in Equations (23) and (24), respectively, if considering the first type of weights given in (17), or in Equations (29) and (30), respectively, if considering the second type of weights given in (18).*

ACKNOWLEDGEMENTS

The first two authors would like to express their appreciation to the University of Rouen and University of the Aegean for the opportunity to exchange several visits in both institutions. The research work of Vlad Stefan Barbu was partially supported by the projects *XTerM–Complex Systems, Territorial Intelligence and Mobility* (2014–2018) and *MOUSTIC–Random Models and Statistical, Informatics and Combinatorics Tools* (2016–2019), within the Large Scale Research Networks from the Region of Normandy, France.

REFERENCES

1. S. AMARI, H. NAGAOKA, *Methods of Information Geometry*, Oxford University Press, New York, 2000.
2. V.S. BARBU, A. KARAGRIGORIOU, V. PREDA, *Entropy and divergence rates for Markov chains: I. The Alpha-Beta and Alpha-Gamma case*, submitted, 2017.
3. V.S. BARBU, A. KARAGRIGORIOU, V. PREDA, *Entropy and divergence rates for Markov chains: III. The Cressie and Read case and applications*, submitted, 2017.
4. A. BASU, I.R. HARRIS, N.L. HJORT, M.C. JONES, *Robust and efficient estimation by minimising a density power divergence*, Biometrika, **85**, pp. 549–559, 1998.
5. M. BELIŞ, S. GUIAŞU, *A quantitative-qualitative measure of information in cybernetic systems*, IEEE Transactions on Information Theory, **14**, *4*, pp. 593–594, 1968.
6. J.S. BHULLAR, O.P. VINOCHIA, M. GUPTA, *Generalized measure for two utility distributions*, Proceedings of the World Congress of Engineering, London, U.K., vol. III, 2010.
7. J.P. CASQUILHO, *Combining expected utility and weighted Gini-Simpson index into a non-expected utility device*, Theoretical Economics Letters, **5**, *2*, pp. 185–195, 2015.
8. H. CHERNOFF, *A measure of asymptotic efficiency for tests of a hypothesis based on a sum of observations*, Ann. Math. Statist., **23**, *4*, pp. 493–507, 1952.
9. A. Di CRESCENZO, M. LONGOBARDI, *On weighted residual and past entropies*, Scientiae Math. Japon., **64**, pp. 255–266, 2006.
10. G. DIAL, I.J. TANEJA, *On weighted entropy of type (α, β) and its generalizations*, Aplikace matematiky, **26**, *6*, pp. 418–425, 1981. http://eudml.org/doc/15212
11. S. GUIAŞU, *Weighted entropy*, Reports on Mathematical Physics, **2**, *3*, pp. 165–179, 1971.
12. J.N. KAPUR, *Measures of Information and Their Applications*, Wiley, New Delhi, 1994.
13. B.D. SHARMA, J. MITTER, M. MOHAN, *On measures of "useful" information*, Information and Control, **39**, *3*, pp. 323–336, 1978.
14. M. ŚMIEJA, *Weighted approach to general entropy function*, IMA Journal of Mathematical Control and Information, **32**, *2*, pp. 329–341, 2015, doi: https://doi.org/10.1093/imamci/dnt044
15. Y. SUHOV, S. YASAEI SEKEH, I. STUHL, *Weighted Gaussian entropy and determinant inequalities*, 2015, pp. 1–28. https://arxiv.org/pdf/1505.01753.pdf
16. Y. SUHOV, I. STUHL, S. YASAEI SEKEH, M. KELBERT, *Basic inequalities for weighted entropies, Aequationes mathematicae*, **90**, *4*, pp. 817–848, 2016, doi:10.1007/s00010-015-0396-5
17. I.J. TANEJA, *On generalized weighted divergence measures*, http://130.203.136.95/viewdoc/download?doi=10.1.1.57.4845&rep=rep1&type=pdf
18. H.C. TANEJA, R.K. TUTEJA, *Characterization of a quantitative-qualitative measure of relative information*, Information Sciences, **33**, pp. 217–222, 1984.
19. H.C. TANEJA, R.K. TUTEJA, *Characterization of a quantitative-qualitative measure of inaccuracy*, Kybernetika, **22**, *5*, pp. 393–402, 1986. http://eudml.org/doc/27733.

Received May 3, 2017