DYNAMIC UNIQUENESS AND PHASE TRANSITION OF CHAINS OF INFINITE ORDER

GALLESCO, C.¹, GALLO, S.², AND TAKAHASHI, D. Y.³

Abstract. We say that a kernel exhibits dynamic uniqueness if all the chains starting from a fixed past coincide in the future tail $\sigma$-algebra, otherwise the kernel exhibits dynamic phase transition. We characterize dynamic uniqueness/phase transition by proving several equivalent conditions. In particular, we prove that dynamic uniqueness is equivalent to convergence in total variation distance of all the chains starting from different pasts. We also study the relationship between our definition of uniqueness and the $\ell^2$ criteria for the uniqueness of $g$-measures. We prove that the Bramson-Kalikow and Hulse models exhibit dynamic uniqueness if and only if the kernel is in $\ell^2$. Using this result, we exhibit examples where equilibrium and dynamic uniqueness are not equivalent. As a consequence, we conclude that the probability of asymptotic events for the equilibrium measure can differ from the chains of infinite order starting from a fixed past.

1. Introduction

Chains of infinite order are generalizations of Markov chains, where the dependence on the past can be unbounded. Similarly to a Markov chain, a chain of infinite order consists of an initial condition (a probability measure on the infinite past) and a set of transition probabilities that defines the probability of appearance of a symbol at each step given its past. Following the analogy with Markov chains, it is natural to study the asymptotic properties of these chains, in particular, the existence/uniqueness of compatible stationary measures and their convergence properties. Because the transition probabilities can depend on the whole past, it

2000 Mathematics Subject Classification. Primary 60G10; Secondary 60G99.

Key words and phrases. Chains of infinite order, coupling, phase transition, Bramson-Kalikow, $\bar{d}$-distance.

This work is supported by USP project “Mathematics, computation, language and the brain”, FAPESP project “NeuroMat” (grant 2011/51350-6), CNPq projects “Stochastic Modeling of the Brain Activity” (grant 480108/2012-9), and “Sistemas Estocásticos: equilíbrio e não equilíbrio, limites em escala e percolação” (grant 474233/2012-0). SG was partially supported by FAPESP (grant 2009/09809-1). CG was partially supported by FAPESP (grant 2013/*****). DYT was partially supported by Pew Latin American Fellowship and Ciência sem Fronteira Fellowship (CNPq 246778/2012-1).
is not straightforward to define measures compatible with a given set of transition probabilities that represent the asymptotic states. Fortunately, there are three related and well established ways to define such measures: Gibbs prescription (DLR measures), thermodynamic limits, and equilibrium states. When the resulting measure is stationary, all three definitions are equivalent and the measures are called \(g\)-measures.

However, if our interest is to study the non-stationary chains of infinite order starting from a fixed past, the above mentioned theories are not always adequate. For example, it is not straightforward to obtain properties of the non-stationary chains from the \(g\)-measures, unless a strong form of loss of memory can be proved, which is typically unavailable. Moreover, at first sight, it is not clear how the probability of the asymptotic events of the chains starting from a fixed past relate to the \(g\)-measures.

We propose in this article a starting point to study these non-stationary chains, in particular, we introduce a definition of dynamic uniqueness and phase transition suited to study the properties of the non-stationary chains. The basic idea is the following. Given a set of transition probabilities and the corresponding chains of infinite order starting with distinct fixed pasts, we want to study their measures on the tail \(\sigma\)-algebra. The events on this tail \(\sigma\)-algebra (tail events) can have the physical interpretation of “asymptotic events” and are nice candidates to detect “phase transitions”. Therefore, we say that a chain of infinite order exhibits a dynamic phase transition if there exist two different pasts for which the corresponding measures on the tail events disagree. Conversely, we say that there is dynamic uniqueness if the measures on the tail \(\sigma\)-algebra agree for all pasts.

We first prove several equivalent criteria for dynamic uniqueness. In particular, we show that dynamic uniqueness implies convergence in total variation distance. This is akin to the convergence of Markov chain to the invariant distribution in total variation distance. We then show that phase transition for \(g\)-measures (equilibrium phase transition) implies dynamic phase transition. We also prove that when the rate of variation of the chain is in \(\ell^2\), there is dynamic uniqueness. The same criterion is known to hold for \(g\)-measures (Johansson & Öberg, 2003). Using the equivalent criteria for dynamic uniqueness, we demonstrate that for the BKF (Bramson & Kalikow, 1993; Friedli, 2014) and Hulse (Hulse, 2006) models, there exists dynamic uniqueness if and only if the rate of variation is in \(\ell^2\). Because there are BKF and Hulse models that exhibit equilibrium uniqueness but doesn’t have rate of variation in \(\ell^2\), we conclude that equilibrium and dynamic uniqueness/phase transition can differ in general.
2. Notations, definitions, and main results

Let $S$ be a finite set, $\mathcal{X} = S^\mathbb{Z}$ and $\mathcal{X}^- = S^{\mathbb{Z}^+}$, where $\mathbb{Z}^+ = \{-1, -2, \ldots\}$. We denote by $x_i$ the $i$-th coordinate of $x \in \mathcal{X}$ and for $-\infty \leq j \leq i \leq \infty$ we write $x_j^i := (x_i, \ldots, x_j)$, $x_j^{-\infty} := (x_{-j}, x_{-j-1}, \ldots)$, and $x_j^{\infty} := (\ldots, x_{i+1}, x_i)$. We also use the shorthand notation $\underline{x} := x_{-\infty}^1$. For $x, y \in \mathcal{X}$ and $i, j, k$ finite, a concatenation $y_j^i \underline{x}^k_{-\infty}$ is a new sequence $z \in S^{-i, -i-1, \ldots}$ with $z^i_j = y^i_j$ and $z^k_{-\infty} = x^k_{-\infty}$. When $i < j$, we use the convention $x_j^i = \varnothing$, where $\varnothing$ is the identity element of the concatenation operation. Note that we are using the convention that the past (smaller indices) of $x \in \mathcal{X}$ is represented on the right hand side.

For each $\Lambda \subset \mathbb{Z}$, consider the set $\mathcal{X}_\Lambda = S^\Lambda$, together with the canonical projection $\pi_\Lambda : \mathcal{X} \to \mathcal{X}_\Lambda$ defined by $\pi_\Lambda(x)_k = x_k$ for all $k \in \Lambda$. Then, define $\mathcal{C}(\Lambda) = \{\pi^{-1}(B) : B \subset \mathcal{X}_\Lambda\}$, called the cylinders with base $\Lambda$. For $\Gamma \subset \mathbb{Z}$, we consider the algebra of cylinders with base in $\Gamma$ defined by $\mathcal{C}(\Gamma) = \bigcup\{\mathcal{C}(\Lambda) : \Lambda \subset \Gamma$ finite} and the $\sigma$-algebra generated by the algebra of cylinders with base in $\Gamma$, $\mathcal{F}_\Gamma = \sigma(\mathcal{C}(\Gamma))$. We use the shorthand notation $\mathcal{F} = \mathcal{F}_\mathbb{Z}$, $\mathcal{F}^- = \mathcal{F}_{\mathbb{Z}^+}$, $\mathcal{F}^+ = \mathcal{F}_{\mathbb{Z}^-}$, and $\mathcal{F}_I = \mathcal{F}_{I \cap \mathbb{Z}}$, for any $I \subset \mathbb{R}$.

A probability kernel, or simply a kernel $g$ on the alphabet $S$ is a measurable function

$$
g : \mathcal{X}^- \to [0, 1]
$$

such that

$$
\sum_{a \in S} g(ax) = 1, \quad \forall x \in \mathcal{X}^-.
$$

We say that a stationary stochastic process $(X_j)_{j \in \mathbb{Z}}$ with values in $S$ defined on a probability space $(\mathcal{X}, \mathcal{F}, P)$ is compatible with a kernel $g$ if the latter is a regular version of the conditional probabilities of the former, that is,

$$P[X_0 = a \mid X_{-1} = \underline{x}] = g(ax)
$$

for every $a \in S$ and $P$-a.e. $x$ in $\mathcal{X}$. The measure $P$ is called $g$-measure. A kernel $g$ is strongly non-null if

$$\inf_{a \in S, \underline{x} \in \mathcal{X}^-} g(ax) > 0. \quad (1)
$$

The variation rate (or continuity rate) of order $k$ of a kernel $g$ is given by

$$\text{var}_k(g) := \sup_{a \in S} \sup_{\omega^{-1}_{-k} \in S^k} \sup_{\underline{x} \in \mathcal{X}^-} |g(\omega^{-1}_{-k} x) - g(\omega^{-1}_{-i} y)|. \quad (2)
$$

We say that $g$ is continuous if $\lim_{k \to \infty} \text{var}_k(g) = 0$. If the kernel is continuous, a compactness argument shows that we can take the weak limits of the processes starting at fixed pasts to obtain compatible stationary chains (see for example [Keane (1972)]). If $g$ is strongly non-null and continuous, we say that $g$ is a regular kernel.
When there is more than one stationary process compatible with $g$, we say that there is an *equilibrium phase transition*, otherwise we say that the process is *unique* (or that there is a *unique equilibrium measure*). These are the usual definition of uniqueness and phase transition for chains of infinite order in literature.

Let $T^+ = \bigcap_{n \geq 1} F_{n,\infty}$ be the future tail $\sigma$-algebra (which elements are the “asymptotic events”) and $T^- = \bigcap_{n \geq 1} F_{-\infty,-n}$ be the past tail $\sigma$-algebra. The following result by Fernández & Maillard (2005) shows that the probability of events on $T^-$ decide whether there is equilibrium uniqueness or phase transition.

**Theorem (Fernández & Maillard (2005)).** A kernel $g$ exhibits equilibrium phase transition if and only if there are two $g$-measures $P$ and $P'$ and an event $A \in T^-$, such that $P[A] \neq P'[A]$.

Let $T$ be the set of probability measures on $(\mathcal{X}^-,\mathcal{F}^-)$. The non-stationary process $(X_j)_{j \in \mathbb{Z}}$ on a probability space $(\mathcal{X},\mathcal{F},\lambda)$ with initial distribution $\lambda \in \mathcal{M}$ and kernel $g$ is defined in following way. For any event $B \in \mathcal{F}^-$, $P^\lambda[B] = \lambda[B]$ and, for $j \geq 0$,

$$P^\lambda[X_j = a \mid X_{j-1}^j = y_{j-1}^j] = g(ay_{j-1}^j).$$

By Ionescu-Tulcea extension theorem, $P^\lambda$ is uniquely defined. The set of measures that can be written as $P^\lambda$ for some $\lambda \in \mathcal{M}$ is denoted $\mathcal{M}_0$. When $\lambda[X_{-\infty}^{-1} = x] = 1$ for some $x \in \mathcal{X}_{-\infty}^{-1}$, we use the notation $P^x$ for the associated probability measure. By standard arguments in measure theory observe that $P^\lambda[B] = \int_{\mathcal{X}} P^x[B] \lambda(dx)$ for all $B \in \mathcal{F}$. Observe that any stationary measure $P$ compatible with $g$ is also an element of $\mathcal{M}_0$.

Inspired by the above result by Fernández & Maillard (2005), we propose the following definition of *dynamic phase transition* and uniqueness.

**Definition.** We say that a kernel $g$ exhibits dynamic phase transition if there exists an event $A \in T^+$ and a pair $x,y \in \mathcal{X}^-$ such that we have $P^x[A] \neq P^y[A]$, otherwise we say that there is dynamic uniqueness.

This definition differs from the equilibrium uniqueness and phase transition by requiring conditions directly on the future tail $\sigma$-algebra of the non-stationary processes, instead on the past tail events of the stationary weak limits of the same process ($g$-measures).

To better understand the above definition, we will prove some equivalent conditions for dynamic uniqueness and phase transition. Below we collect some definitions that we will need to state the result.

A probability measure $\mu$ is *trivial* on $\mathcal{T}$, if for all $A \in \mathcal{T}$ we have $\mu(A) = 1$ or 0. Let $\mathcal{H}$ be a sub-$\sigma$-algebra of $\mathcal{F}$. We indicate by $\mu|_{\mathcal{H}}$ the restriction of $\mu$.
DYNAMIC UNIQUENESS AND PHASE TRANSITION OF CHAINS OF INFINITE ORDER

to \( \mathcal{H} \). Let \( \tilde{\mu} \) be a coupling between \( \mu \) and \( \mu' \) and \((X_j,Y_j)_{j \in \mathbb{Z}}\) be the associated process, such that \((X_j)_{j \in \mathbb{Z}} \overset{\text{D}}{=} \mu \) and \((Y_j)_{j \in \mathbb{Z}} \overset{\text{D}}{=} \mu' \). The coupling time \( \Theta \) is defined as \( \Theta = \inf \{ t \geq 0 : X_j = Y_j \text{ for all } j \geq t \} \). We say that a measure \( \mu \) is mixing if, for all \( B \in \mathcal{F} \),
\[
\lim_{n \to \infty} \sup_{A \in \mathcal{F}_{[n,\infty)}} |\mu(A \cap B) - \mu(A)\mu(B)| = 0.
\]

The following theorem characterizes dynamic uniqueness.

**Theorem 1.** Let \( g \) be a regular kernel. The following statements are equivalent:

(i) \( g \) exhibits dynamic uniqueness.

(ii) For all \( x, y \in X^- \), \( \lim_{n \to \infty} \sup_{B \in \mathcal{F}_{[n,\infty)}} |P^x[B] - P^y[B]| = 0 \).

(iii) For all \( x, y \in X^- \) there exist a coupling \( \hat{P} \) between \( P^x \) and \( P^y \) and a coupling time \( \Theta \) such that \( \hat{P}[\Theta < \infty] = 1 \).

(iv) For all \( \lambda, \lambda' \in \mathcal{M} \), \( P^\lambda|_{\mathcal{T}^+} = P^{\lambda'}|_{\mathcal{T}^+} \).

(v) For all \( \lambda \in \mathcal{M} \), \( P^\lambda \) is trivial on \( \mathcal{T}^+ \).

(vi) For all \( \lambda \in \mathcal{M} \), \( P^\lambda \) is mixing.

(vii) For all \( x, y \in X^- \), \( P^x|_{\mathcal{F}^+} \ll P^y|_{\mathcal{F}^+} \).

(viii) For all \( x, y \in X^- \),
\[
\sum_{n=0}^{\infty} \sum_{a \in S} \left( g(a\omega^0_n x) - g(a\omega^0_n y) \right)^2 < \infty, \text{ for } P^x\text{-a.e. } \omega \in \mathcal{X}.
\]

Equivalences (i), (ii), and (iii) in Theorem 1 follow from the characterization of exact maximal coupling given by Thorisson (2000). We remind the reader that given two measures \( \mu \) and \( \nu \) on \((X, \mathcal{F})\), the total variation distance between \( \mu \) and \( \nu \) is given by \( d_{TV}(\mu, \nu) = \sup_{A \in \mathcal{F}} |\mu(A) - \nu(A)| \). Thus (ii) is equivalent to
\[
\lim_{n \to \infty} d_{TV}(P^x|_{\mathcal{F}_{[n,\infty)}}, P^y|_{\mathcal{F}_{[n,\infty)})} = 0
\]
and demonstrates the convergence in total variation. Equivalence (iv) shows that the probabilities of the tail events of the non-stationary and stationary chains of infinite order coincides when there is dynamic uniqueness. Equivalences (v) and (vi) show that dynamic phase transition is observed when there exist an initial condition for which the correlation doesn’t decay to zero, formalizing the idea that a phase transition happens when there are long range correlations. Theorem 1 (viii) and (viii) are obtained using the theory of absolute continuity and singularity developed by Shiryaev and co-authors (Jacod & Shiryaev, 2002; Engelbert & Shiryaev, 1980). In particular, (viii) can be used to generate criteria for dynamic uniqueness as it is shown in Corollary 2 below. The key ingredient for the proof of equivalences
between (vii) and (i) is to show that the future tail $\sigma$-algebra of the extremal $g$-measures are trivial. Previously only the past tail and the invariant $\sigma$-algebras were known to be trivial for extremal $g$-measure (Fernández & Maillard, 2003; Walters, 2000).

An immediate corollary of Theorem 1 (ii) is the following relation between dynamic and equilibrium phase transition.

**Corollary 1.** If a kernel $g$ is regular and exhibits equilibrium phase transition, then it exhibits dynamic phase transition.

We say that $g$ is in $\ell^q$ if $\sum_{k=1}^{\infty} \text{var}_k(g)^q < \infty$. Johansson & Öberg (2003) proved that if a regular kernel $g$ is in $\ell^2$, then it has a unique compatible $g$-measure. A straightforward consequence of Theorem 1 (viii) is that the same criterion guarantees dynamic uniqueness.

**Corollary 2.** If $g$ is a kernel in $\ell^2$, then there is dynamic uniqueness.

In Johansson et al. (2012) it is proved that if $g$ is a kernel in $\ell^2$, then the chain of infinite order starting with any initial measure converges weakly to the unique compatible stationary measure. Corollary 2 together with Theorem 1 (ii) imply that we can strengthen this result and obtain, under the same assumptions, a convergence in total variation distance.

Corollaries 1 and 2 together imply that models exhibiting equilibrium phase transition at $\ell^{2+\alpha}$, for any $\alpha > 0$, will have sharp dynamic phase transition. In particular, Berger et al. (2005) introduced for any positive $\alpha$ a class of kernels in $\ell^{2+\alpha}$ that exhibits equilibrium phase transition. This implies that the same class of models exhibits dynamic phase transition for kernels in $\ell^{2+\alpha}$ and dynamic uniqueness for $\ell^2$. The coincidence between equilibrium and dynamic uniqueness/phase transition does not hold in general and there are models that exhibit equilibrium uniqueness but have dynamic phase transition. To show this, we exhibit two models studied in the literature of chain of infinite order.

Friedli (2014) introduced the following generalization of the model initially studied by Bramson & Kalikow (1993). Let $S = \{-1, +1\}$, $\varepsilon \in (0, 1/2)$, and $(m_j)_{j \geq 1}$ be an increasing sequence of positive odd numbers. We consider a non-decreasing function $\psi : [-1, 1] \to [\varepsilon, 1-\varepsilon]$ which satisfies $\psi(r) + \psi(-r) = 1$. We call the sequence $(m_j)_{j \geq 1}$ lacunary if for some $0 \leq r_0 < 1$ such that $\psi(r_0) > \psi(-r_0)$ we have $m_{j+1} \geq \frac{4}{1-r_0} m_j$ for all $j \geq 1$. Let $x \in \mathcal{X}$, for $j \geq 1$ we denote by $Q_j$ the function

$$Q_j(ax) = \psi \left( \frac{a}{m_j} \sum_{l=1}^{m_j} x_{-l} \right).$$
Let \((\lambda_j)_{j \geq 1}\) be a sequence of positive numbers such that \(\sum_{j=1}^{\infty} \lambda_j = 1\). Given \((m_j)_{j \geq 1}\) and \((\lambda_j)_{j \geq 1}\), the BKF-model associated to the parameters \((m_j)_{j \geq 1}\) and \((\lambda_j)_{j \geq 1}\) is given by the kernel \(g\) such that, for all \(x \in \mathcal{X}\),
\[
g(ax) = \sum_{j=1}^{\infty} \lambda_j Q_j(ax).
\]

The original model introduced by Bramson & Kalikow (1993) is obtained by choosing \(\psi(r) = 1 - \epsilon\) if \(r \geq 0\) and \(\psi(r) = \epsilon\) otherwise. Bramson & Kalikow (1993) showed that it is possible to choose \((m_j)_{j \geq 1}\) and \((\lambda_j)_{j \geq 1}\) such that the corresponding model exhibits phase transition. Some progress has been made to obtain sufficient conditions for equilibrium phase transition in this model (Friedli, 2014; Gallo & Takahashi, 2013; Gallesco et al., 2014), but a sharp condition on the kernel to guarantee phase transition remains elusive. We note that in all known conditions of equilibrium phase transition of the BK model, the sequence \((m_j)_{j \geq 1}\) is lacunary (Friedli, 2014; Gallesco et al., 2014).

The binary autoregressive models constitute a different class of models that can also exhibit equilibrium phase transition. These models are defined through the following parameters: a continuous and increasing function \(\phi : \mathbb{R} \to ]0, 1[\) such that \(\phi(r) + \phi(-r) = 1\), a summable sequence of non-negative real numbers \((\beta_n)_{n \geq 1}\), and a non-negative real parameter \(\delta \geq 0\). The kernel \(g\) of a binary autoregressive model on the alphabet \((-1, 1)\) is given by
\[
g(ax) := \phi \left( a \sum_{n \geq 1} \beta_n x_{-n} + a\delta \right).
\]
If \(\phi\) is bi-Lipschitz, we have that \(1/\gamma \sum_{n>\gamma} \beta_n \leq \text{var}_k(g) \leq \gamma \sum_{n>\gamma} \beta_n\) for some positive constant \(\gamma\). An important example of binary autoregressive model is when \(\phi(r) = e^{-r}(e^{-r} + e^{r})^{-1}\). The resulting kernel is called logit model in the statistics literature, and one-sided 1-dimensional long-range Ising model in statistical physics literature. Hulse (2006) used this model to give an example of phase transition for chains of infinite order.

We prove the following result.

**Corollary 3.** A BKF model with lacunary \((m_j)_{j \geq 1}\) or a binary autoregressive model with bi-Lipschitz \(\phi\) exhibits dynamic uniqueness if and only if the corresponding kernel is in \(\ell^2\).

Let the oscillation of order \(k\) of a kernel be defined as
\[
\text{osc}_k(g) := \sum_{a \in S} \sup_{\omega \in \mathcal{X}} \sup_{b, b' \in S} |g(a \omega^{-1}_{-k+1} b \omega^{-k-1}_{-\infty}) - g(a \omega^{-1}_{-k+1} b' \omega^{-k-1}_{-\infty})|.
\]
Fernández & Maillard (2005) have shown that if \( \sum_{k \geq 1} \text{osc}_k(g) < 1 \), then there is a unique \( g \)-measure. For BKF and binary autoregressive models, it is straightforward to exhibit models that are not in \( \ell^2 \), but satisfy uniqueness criterion. For example, if \( \psi(r) = 1/2 + (1/2 - \varepsilon)r \), the BKF model always satisfies Fernández & Maillard (2005) uniqueness criterion although the kernels can have arbitrarily slow continuity rate. This shows that the equilibrium and dynamic uniqueness/phase transition are not equivalent in general. The implication is that there exists a kernel \( g \) with a unique compatible measure \( P \) and a non-stationary chain \( P^x \) (for some \( x \in \mathcal{X} \)) such that \( P \) and \( P^x \) differ on \( T^+ \). This suggests the following physical interpretation: when dynamic and equilibrium uniqueness does not coincide, there is a “hidden” phase transition (dynamic phase transition) that cannot be detected looking at events that depend on a finite number of coordinates, but can only be detected looking at the asymptotic events.

3. Proofs

Proof of equivalences (i), (ii), and (iii) in Theorem [1]. These results are straightforward consequences of some well known results for stochastic processes. We state the general theorem from Thorisson (2000) for the convenience of the reader.

Theorem (Thorisson (2000)). Let \( Y = (Y_j)_{j \geq 0} \) and \( Z = (Z_j)_{j \geq 0} \) be one-sided processes on probability spaces \((S^N, \mathcal{F}_+^+, \mu)\) and \((S^N, \mathcal{F}_+^+, \nu)\), respectively. The following are equivalent:

(a) For all tail events \( A \in \mathcal{T}_+^+ \) we have \( \mu[A] = \nu[A] \).
(b) \( \lim_{n \to \infty} \sup_{B \in \mathcal{F}_{[n, \infty)}} |\mu[B] - \nu[B]| = 0 \).
(c) There exists a coupling \( \tilde{P} \) between \( Y \) and \( Z \) such that the coupling time \( \Theta \) is finite \( \tilde{P} \)-a.s.

Taking \( \mu = P^x|_{\mathcal{F}_{[0, \infty)}} \) and \( \nu = P^y|_{\mathcal{F}_{[0, \infty)}} \), the equivalences (a), (b), and (c) in the above theorem implies, respectively, the equivalences between (i), (ii), and (iii) in Theorem [1] □

Proof of equivalences (i) and (iv) in Theorem [1]. Equivalence between (i) and (iv) is a straightforward consequence of the representation, for all \( \lambda \in \mathcal{M} \) and any \( B \in \mathcal{F} \),

\[
P^\lambda[B] = \int_{\mathcal{X}} P^x[B] \lambda(dx).
\]

□
Proof of equivalences (i) and (v) in Theorem 1. Let us prove a lemma suggested by a remark in Olshen (1971). In what follows, equality $\mu$-a.s. of $\sigma$-algebras $\mathcal{T}^-$ and $\mathcal{T}^+$ means that for all $A \in \mathcal{T}^-$ there exists $B \in \mathcal{T}^+$ such that $\mu[A\Delta B] = 0$ and vice-versa.

**Lemma 1.** Let $\mu$ be a stationary probability measure on $(\mathcal{X}, \mathcal{F})$. We have that $\mathcal{T}^- = \mathcal{T}^+$, $\mu$-a.s.

**Proof.** From now on and until the end of this proof, we consider that $\mathcal{F}$ is complete (that is, it contains any subset of any measurable set of $\mu$-measure 0). Furthermore, we assume that all the $\sigma$-algebras considered in this proof also contain the class of $\mu$-null sets. Lemma 1 is a consequence of Theorem 2 in Rohlin & Sinai (1961). We restate their theorem using our notation. Let $h(T)$ be the entropy of $T$ and $H(A \mid T^{-1}A)$ be the conditional entropy of the measurable partition corresponding to $A$ given $T^{-1}A$. We also denote by $\pi(T)$ the Pinsker $\sigma$-algebra defined by $\pi(T) := \{A \in \mathcal{F} : h(\sigma(A), T) = 0\}$, where $h(\sigma(A), T)$ is the entropy of $T$ given the measurable partition corresponding to $\sigma(A)$ ($\sigma(A)$ denotes the $\sigma$-algebra generated by $A$). For a definition of these notions, see for example Martin & England (1981).

**Theorem (Rohlin & Sinai (1961)).** Let $(\mathcal{X}, \mathcal{F}, \mu)$ be a Lebesgue probability space and $T$ be a measure preserving automorphism. If there exists a sub $\sigma$-algebra $A \subset \mathcal{F}$ such that

(a) $T^{-1}A \subset A$,
(b) $\bigvee_{n=0}^{\infty} T^n A = \mathcal{F}$,
(c) $h(T) = H(A \mid T^{-1}A)$,
(d) $h(T) < \infty$,

then $\bigcap_{n=0}^{\infty} T^{-n} A = \pi(T)$.

To prove Lemma 1, we will first verify that all the conditions of the above theorem is satisfied taking $A = \mathcal{F}_{[0, \infty)}$ and $T$ the shift defined by $(Tx_j) = x_{j+1}$. We have $T\mathcal{F}_{[j, \infty)} = \mathcal{F}_{[j-1, \infty)}$ and $T^{-1}\mathcal{F}_{[j, \infty)} = \mathcal{F}_{[j+1, \infty)}$, therefore, (a) and (b) in the above theorem are satisfied. By Theorem 2.39 in Martin & England (1981), we have that $h(T) = h(\mathcal{F}_{[0}, T)$. Besides, since $\mathcal{F}_{[0}$ is finite (up to $\mu$-null sets), by Theorem 2.27 in Martin & England (1981), we have

$$h(\mathcal{F}_{[0}, T) = H(\mathcal{F}_{[0}, \mathcal{F}_{[1, \infty)}).$$

Since $H(\mathcal{F}_{[0}, \mathcal{F}_{[1, \infty)}) \leq \log |S|$, (d) is satisfied. Finally, it is not difficult to check that $H(\mathcal{F}_{[0, \infty)} \mathcal{F}_{[1, \infty)}) = H(\mathcal{F}_{[0}, \mathcal{F}_{[1, \infty)})$, which implies (c). Thus using the above theorem, we conclude that $\mathcal{T}^+ = \pi(T)$. Similarly, if we take $\bar{T}$ as the inverse shift...
\[(\tilde{T}x_j) = x_{j-1} \text{ and } A = \mathcal{F}_{(-\infty,0]}, \text{ we conclude that } \mathcal{T}^- = \pi(\tilde{T}). \] Now, it remains to show that \(\pi(T) = \pi(\tilde{T}).\) For this, we note that, if \(T\) is invertible, we have (see again Theorem 2.27 in Martin & England (1981))

\[
h(\sigma(A), T) = H(\sigma(A)) \bigg/ \bigg( T^j \sigma(A) \bigg| \bigcup_{j=1}^{\infty} T^{-j} \sigma(A) \bigg)
\]

\[
= H(\sigma(A)) \bigg/ \bigg( T^j \sigma(A) \bigg| \bigcup_{j=1}^{\infty} T^{-j} \sigma(A) \bigg)
\]

\[
= h(\sigma(A), \tilde{T}).
\]

Therefore, \(\mathcal{T}^+ = \mathcal{T}^-\), as we wanted to show.

It is interesting to note that the above lemma is false in general if the alphabet is not finite. For this let \((\xi_j)_{j \in \mathbb{Z}}\) be an i.i.d. process with \(P(\xi_j = 1) = P(\xi_j = -1) = 1/2\) for all \(j \in \mathbb{Z}\). Now, define \((\tilde{\xi}_j)_{j \in \mathbb{Z}}\) by \(\tilde{\xi}_j = (\xi_j, \xi_{j-1}, \ldots)\). Clearly, \((\tilde{\xi}_j)_{j \in \mathbb{Z}}\) is a stationary Markov chain with trivial past tail \(\sigma\)-algebra. The future tail \(\sigma\)-algebra is a full measure because all the events of the i.i.d. process are measurable. This example also shows that transforming a chain of infinite order to a Markov chain on an enlarged state space does not necessarily give some information about the future tail \(\sigma\)-algebra of the chain of infinite order.

Now, let us prove the equivalences \((i)\) and \((v)\) in Theorem 1. To prove that \((v)\) implies \((i)\), let \(\lambda, \lambda' \in \mathcal{M}\) and take \(\lambda'' = (\lambda + \lambda') / 2\). For any \(A \in \mathcal{T}^+\), we have \(P^{\lambda''}[A] = (P^{\lambda}[A] + P^{\lambda'}[A]) / 2\). Since \(P^{\lambda''}\) is trivial on \(\mathcal{T}^+\), then either \(P^{\lambda}[A] = P^{\lambda'}[A] = 1\) or \(0\), as we wanted to show.

To show that \((i)\) implies \((v)\), we observe that because \((i)\) implies \((iv)\), the existence of a \(P \in \mathcal{M}_0\) trivial on \(\mathcal{T}^+\) implies that all elements of \(\mathcal{M}_0\) are trivial. To prove that such \(P\) always exists, we first use the fact that extremal \(g\)-measures are trivial on \(\mathcal{T}^-\) (Fernández & Maillard (2005)) and that there is at least one extremal \(g\)-measure (Friedli, 2008) [theorem 3.4]. Using Lemma 1 we conclude that the extremal \(g\)-measures are also trivial on \(\mathcal{T}^+\). Second, let \(P\) be an extremal \(g\)-measure and \(P_0 = P|_{\mathcal{F}_{(-\infty,-1)}}\). The stationary measure \(P\) can be written as \(P_{\mathcal{F}_{-\infty}}\) and therefore it is an element of \(\mathcal{M}_0\), as we wanted to show.

Proof of equivalences \((v)\) and \((vi)\) in Theorem 1. The equivalence between \((v)\) and \((vi)\) is standard in the literature (see for example Theorem 2 in Blackwell & Freedman (1964)). We exhibit the argument for convenience of the reader.
If (vi) holds, we have for any $A \in \mathcal{T}^+$, $P^\lambda[A] = P^\lambda[A]^2$. Hence, $P^\lambda[A] = 0$ or 1. To prove that (v) implies (vi), fix $B \in \mathcal{F}$. For any $A \in \mathcal{F}_{[n,\infty)}$, we have
\[ |P^\lambda[A \cap B] - P^\lambda[A]P^\lambda[B]| = \int_A (P^\lambda[B|\mathcal{F}_{[n,\infty]}] - P^\lambda[B])dP^\lambda \leq \int_{\mathcal{X}} |P^\lambda[B|\mathcal{F}_{[n,\infty]}] - P^\lambda[B]|dP^\lambda. \]
By backward martingale convergence theorem and triviality of $P^\lambda$ in $\mathcal{T}^+$, the right hand side of the above inequality converges to 0. Therefore, we have
\[ \lim_{n \to \infty} \sup_{A \in \mathcal{F}_{[n,\infty)}} |P^\lambda[A \cap B] - P^\lambda[A]P^\lambda[B]| = 0, \]
as we wanted to show. □

**Proof of equivalences (i) and (vii) in Theorem** Engelbert & Shiryaev (1980) proved the following dichotomy result.

**Theorem** (Engelbert & Shiryaev (1980)). Let $Y = (Y_j)_{j \geq 0}$ and $Z = (Z_j)_{j \geq 0}$ be one-sided processes on probability spaces $(S^N, \mathcal{F}^+, \mu)$ and $(S^N, \mathcal{F}^+, \nu)$, respectively. If for all $n \geq 0$, we have $\mu|_{\mathcal{F}_{[n,\infty)}} \ll \nu|_{\mathcal{F}_{[0,n)}}$ then
(a) $\mu \ll \nu$ if and only if
\[ \mu \left[ \lim_{n \to \infty} \sup_{A \in \mathcal{F}_{[0,n)}} \frac{d\mu|_{\mathcal{F}_{[n,\infty)}}}{d\nu|_{\mathcal{F}_{[n,\infty)}}} < \infty \right] = 1. \]
(b) $\mu \perp \nu$ if and only if
\[ \mu \left[ \lim_{n \to \infty} \sup_{A \in \mathcal{F}_{[0,n)}} \frac{d\mu|_{\mathcal{F}_{[n,\infty)}}}{d\nu|_{\mathcal{F}_{[n,\infty)}}} < \infty \right] = 0. \]

To prove that (i) implies (vii), let $\mu = P_{\mathcal{F}_+}^\mu$ and $\nu = P_{\mathcal{F}_+}^\nu$ in the above theorem. The regularity of kernel $g$ guarantees $P_{\mathcal{F}_+}^\mu \ll P_{\mathcal{F}_+}^\nu$ for all $n \geq 0$. We observe that the event
\[ D = \left\{ \lim_{n \to \infty} \sup_{A \in \mathcal{F}_{[n,\infty)}} \frac{dP_{\mathcal{F}_+}^\mu|_{\mathcal{F}_{[n,\infty)}}}{dP_{\mathcal{F}_+}^\nu|_{\mathcal{F}_{[n,\infty)}}} < \infty \right\} \]
is in $\mathcal{T}^+$. We have that (i) implies (v) in Theorem 1. Therefore, for each $x, y \in \mathcal{X}^-$, we have $P_{\mathcal{F}_+}^\mu[D] = P_{\mathcal{F}_+}^\nu[D] = 0$ or 1. By Fatou lemma, $E_{P_{\mathcal{F}_+}^\mu} \left[ \lim_{n \to \infty} \frac{dP_{\mathcal{F}_+}^\mu|_{\mathcal{F}_{[n,\infty)}}}{dP_{\mathcal{F}_+}^\nu|_{\mathcal{F}_{[n,\infty)}}} \right] \leq 1$ and therefore $P_{\mathcal{F}_+}^\mu[D] = 1$. Hence, we have $P_{\mathcal{F}_+}^\mu[D] = 1$. The above theorem by Engelbert & Shiryaev (1980) then implies that $P_{\mathcal{F}_+}^\mu \ll P_{\mathcal{F}_+}^\nu$ as we wanted to show.

To prove that (vii) implies (i), we use again the fact that an extremal $g$-measure $P$ is trivial on $\mathcal{T}^-$. Fernández & Maillard (2005) and that there is at least one extremal
Because of Lemma 1, this implies that $P$ is trivial on $T^+$. Using the property of regular conditional probability, we have that $P^x = P^x\{\cdot\mid F^-\}(x)$ for $P$-a.e. $x$. In particular, there is a past $\omega \in \mathcal{X}^-$ such that $P^\omega$ is trivial on $T^+$. If (vii) holds, we have that, for all $y \in \mathcal{X}$, $P_{\omega\mid T^+}^y \ll P_{\omega\mid T^+}^\omega$. This implies that, for all $y \in \mathcal{X}$, $P_{\omega\mid T^+}^\omega = P_{\omega\mid T^+}^y$.

□

Proof of equivalence (vii) and (viii) in Theorem 1. To prove the equivalence between (vii) and (viii), we use the results from the theory of predictable absolute continuity and singularity (ACS) criteria developed by Shiryaev and co-authors. The idea of using the predictable ACS criteria for chains of infinite order was initiated in Johansson et al. (2007). In what comes next, we closely follow the exposition in Johansson et al. (2007).

Let $X = (X_j)_{j\in\mathbb{Z}}$ and $Y = (Y_j)_{j\in\mathbb{Z}}$ be stochastic processes on $(\mathcal{X}, \mathcal{F}, \mu)$ and $(\mathcal{X}, \mathcal{F}, \nu)$, respectively. For $\omega \in \mathcal{X}$, let $Z_n(\omega) := \frac{d\mu}{d\mu_{\mathcal{F}[0,n-1]}}(\omega)$ and $\alpha_n(\omega) := Z_n(\omega)/Z_{n-1}(\omega)$. Also, define

$$d_n(\omega) := \mathbb{E}_\nu[\frac{1}{\sqrt{\alpha_n}}]^2|\mathcal{F}_{[0,n-1]}](\omega).$$

The predictable ACS criteria is given by

**Theorem** (see Jacod & Shiryaev (2002), Theorem 2.36, p.253). If for all $n \geq 0$ we have $\mu_{\mathcal{F}[0,n]} \ll \nu_{\mathcal{F}[0,n]}$, then $\mu_{\mathcal{F}^+} \ll \nu_{\mathcal{F}^+}$ if and only if $\sum_{n=1}^{\infty} d_n(\omega) < \infty \mu_{\mathcal{F}^+}$-a.s.

Let us rewrite $d_n(x)$ in a more explicit form. We have that

$$\frac{d\mu_{\mathcal{F}[0,n]}}{d\mu_{\mathcal{F}[0,n-1]}}(\omega) = \mu(X_n = \omega_n|X_0^{n-1} = \omega_0^{n-1}) =: \mu(\omega_n|\omega_0^{n-1}).$$

Hence,

$$\mathbb{E}_\nu[\frac{1}{\sqrt{\alpha_n}}]^2|\mathcal{F}_{[0,n-1]}](\omega) = \sum_{\omega_n \in S} \left(1 - \sqrt{\frac{\mu(\omega_n|\omega_0^{n-1})}{\nu(\omega_n|\omega_0^{n-1})}}\right)^2 \nu(\omega_n|\omega_0^{n-1})$$

$$= \sum_{\omega_n \in S} \left(\sqrt{\nu(\omega_n|\omega_0^{n-1})} - \sqrt{\mu(\omega_n|\omega_0^{n-1})}\right)^2.$$

To prove the equivalence (vii) and (viii) in Theorem 1 we take $\mu = P^\omega$, $\nu = P^\omega$ and apply the above theorem (observe that by (iv) the hypothesis of the theorem are
In this case, we have

\[ d_n(\omega) = \sum_{\omega_n \in S} \left( \sqrt{P^y(\omega_n | \omega_0^{n-1})} - \sqrt{P^x(\omega_n | \omega_0^{n-1})} \right)^2 \]

\[ = \sum_{a \in S} \left( \sqrt{g(a\omega_0^{n-1}y) - \sqrt{g(a\omega_0^{n-1}x)}} \right)^2. \]

To conclude, we only need to show that \( \sum_{a \in S} (g(a\omega_0^{n-1}y)^{1/2} - g(a\omega_0^{n-1}x)^{1/2})^2 \) has the same order as \( \sum_{a \in S} (g(a\omega_0^{n-1}y) - g(a\omega_0^{n-1}x))^2 \).

We have that

\[ \sum_{a \in S} (g(a\omega_0^{n-1}y) - g(a\omega_0^{n-1}x))^2 = \sum_{a \in S} (g(a\omega_0^{n-1}y)^{1/2} - g(a\omega_0^{n-1}x)^{1/2})^2 (g(a\omega_0^{n-1}y)^{1/2} + g(a\omega_0^{n-1}x)^{1/2})^2. \]

Therefore, taking

\[ \gamma = \inf_{a \in S} \inf_{\omega \in X^-} g(a\omega), \]

which belongs to the interval \((0, 1)\) by (1), we have

\[ 4\gamma d_n(\omega) \leq \sum_{a \in S} (g(a\omega_0^{n-1}y) - g(a\omega_0^{n-1}x))^2 \leq 4(1 - \gamma) d_n(\omega). \]

From the above inequality and the ACS criteria, we conclude that \( P^x|_{\mathcal{F}^+} \ll P^y|_{\mathcal{F}^+} \) if and only if

\[ \sum_{n=0}^{\infty} \sum_{a \in S} (g(a\omega_0^ny) - g(a\omega_0^nx))^2 < \infty \]

for \( P^x\text{-a.e. } \omega \).

\[ \square \]

**Proof of Corollary 1.** If there is equilibrium phase transition, there exist two stationary probability measures \( P \) and \( P' \) compatible with \( g \), such that \( P[C] \neq P'[C] \) for some cylinder \( C \). Because \( P \) and \( P' \) are stationary

\[ |P[T^nC] - P[T^nC]| = |P[C] - P'[C]| > 0 \quad (4) \]

for any \( n \in \mathbb{Z} \). Now, if there is dynamic uniqueness, by Theorem 1, we have that (3) holds for \( P \) and \( P' \). Using again the theorem by Thorisson (2000), we have

\[ \lim_{n \to \infty} \sup_{B \in \mathcal{F}_{[n, \infty)}} |P[B] - P'[B]| = 0, \]

which is in contradiction with (4). Thus, \( g \) exhibits dynamic phase transition. \( \square \)
Proof of Corollary 2. We have by definition of variation rate that for all $a \in S$, $\omega \in \mathcal{X}$ and $x, y \in \mathcal{X}^-$

$$|g(a\omega^n y) - g(a\omega^n x)| \leq \operatorname{var}_n(g),$$

therefore,

$$\sum_{n=0}^{\infty} \sum_{a \in S} (g(a\omega^n y) - g(a\omega^n x))^2 \leq |S| \sum_{n=0}^{\infty} \operatorname{var}_n(g)^2.$$ 

Using Theorem 1 (viii), we conclude the result. \(\square\)

Proof of Corollary 3. Corollary 2 shows that if the kernel is in $\ell^2$ then we have dynamic uniqueness. Hence, from Theorem 1 (viii), we only need to show that if the BKF and binary autoregressive kernels are not in $\ell^2$, then there exist $x, y \in \mathcal{X}^-$ such that

$$\sum_{n=0}^{\infty} \sum_{a \in S} (g(a\omega^n y) - g(a\omega^n x))^2 = \infty$$

for all $\omega \in \mathcal{X}$.

For both models, we choose $x = 1$, where $1_j = 1$ for all $j \leq -1$. Analogously, we define $-1_j = -1$ for all $j \leq -1$ and we take $y = -1$. For both BKF and binary autoregressive models we have, for all $\omega \in \mathcal{X}$,

$$\sum_{n=0}^{\infty} \sum_{a \in S} (g(a\omega^n - 1) - g(a\omega^n 1))^2 \geq \sum_{n=0}^{\infty} \inf_{a \in S} \inf_{\omega_0 \in S^n} (g(a\omega^n - 1) - g(a\omega^n 1))^2.$$

A straightforward calculation shows that for the BKF model with lacunary $(m_j)_{j \geq 1}$, we have for any $m_n < j \leq (1-r_0)^{-1}m_{n+1}$

$$\inf_{a \in S} \inf_{\omega_0 \in S^n} (g(a\omega^n - 1) - g(a\omega^n 1)) \geq (\psi(r_0) - \psi(-r_0)) \sum_{k \geq n+1} \lambda_k.$$

Therefore,

$$\sum_{n=0}^{\infty} \inf_{a \in S} \inf_{\omega_0 \in S^n} (g(a\omega^n - 1) - g(a\omega^n 1))^2$$

$$\geq \frac{1-r_0}{4} (\psi(r_0) - \psi(-r_0))^2 \sum_{n=2}^{\infty} m_n \left( \sum_{k \geq n} \lambda_k \right)^2. \tag{6}$$

We also have

$$\sum_{n=1}^{\infty} \operatorname{var}_n(g)^2 \leq (1-\epsilon)^2 \sum_{n=1}^{\infty} m_n \left( \sum_{k \geq n} \lambda_k \right)^2. \tag{7}$$

Therefore, from (7) we conclude that if $\sum_{n=1}^{\infty} \operatorname{var}_n(g)^2$ diverges then $\sum_{n=1}^{\infty} m_n \left( \sum_{k \geq n} \lambda_k \right)^2$ diverges. Furthermore, from (6) if $\sum_{n=1}^{\infty} m_n \left( \sum_{k \geq n} \lambda_k \right)^2$ diverges then we obtain (3), as we wanted to show.
For the binary autoregressive model, using the fact that $\phi$ is bi-Lipschitz, we have
\[
\sum_{n=0}^{\infty} \inf_{a \in S} \inf_{\omega_0^n \in S^n} (g(a\omega_0^n) - g(a\omega_0^{n+1}))^2 \geq 1/\gamma^2 \sum_{n=1}^{\infty} (\sum_{k>n} \beta_n)^2.
\] (8)

Also,
\[
\sum_{n=1}^{\infty} \text{var}_n(g)^2 \leq \gamma^2 \sum_{n=1}^{\infty} (\sum_{k>n} \beta_n)^2.
\] (9)

Therefore, from (8), (9), and Theorem 1, we conclude that the binary autoregressive process is in $\ell^2$ if and only if it has dynamic uniqueness as we wanted to prove. □

References

BERGER, N., HOFFMAN, C. & SIDORAVICUS, V. (2005). Nonuniqueness for specifications in $\ell^{2+\epsilon}$. arXiv:math/0312344.

BLACKWELL, D. & FREEDMAN, F. (1964). On tail $\sigma$-field of a markov chain and a theorem of orey. Annals of Mathematical Statistics 35, 1291–1295.

BRAMSON, M. & KALIKOW, S. (1993). Nonuniqueness in $g$-functions. Israel J. Math. 84(1-2), 153–160.

ENGELBERT, H. & SHIRYAEV, A. (1980). On absolute continuity and singularity of probability measures. Mathematical Statistics Banach Center Publications 6, 121–132.

FERNÁNDEZ, R. & MAILLARD, G. (2005). Chains with complete connections: general theory, uniqueness, loss of memory and mixing properties. J. Stat. Phys. 118(3-4), 555–588.

FRIEDLI, S. (2008). On the specification of probabilities by regular g-functions. http://www.mat.ufmg.br/~sacha/textos/gfunctions/gfunctions.pdf, Unpublished notes.

FRIEDLI, S. (2014). A note on the Bramson-Kalikow process. Brazilian Journal of Probability and Statistics, to appear.

GALLESCO, C., GALLO, S. & TAKAHASHI, D. (2014). Explicit estimates in the bramson–kalikow model. Nonlinearity 27, 2281–2296.

GALLO, S. & TAKAHASHI, D. (2013). Attractive regular stochastic chains: perfect simulation and phase transition. Ergodic Theory and Dynamical System, FirstView Article, 1–20, http://dx.doi.org/10.1017/etds.2013.7.

HULSE, P. (2006). An example of non-unique g-measures. Ergodic Theory Dynam. Systems 26(2), 439–445.

JACOD, J. & SHIRYAEV, A. N. (2002). Limit theorems for stochastic processes, vol. 288. Springer-Verlag Berlin, second ed.
Johansson, A. & Öberg, A. (2003). Square summability of variations of $g$-functions and uniqueness of $g$-measures. *Math. Res. Lett.* **10**(5-6), 587–601.

Johansson, A., Öberg, A. & Pollicott, M. (2007). Countable state shifts and uniqueness of $g$-measures. *American journal of mathematics*, 1501–1511.

Johansson, A., Öberg, A. & Pollicott, M. (2012). Unique bernoulli $g$-measures. *Journal of the European Mathematical Society* **14**, 1599–1615.

Keane, M. (1972). Strongly mixing $g$-measures. *Invent. Math.* **16**, 309–324.

Martin, N. & England, J. (1981). *Mathematical theory of entropy*. Cambridge University Press.

Olshen, R. A. (1971). The coincidence of measure algebras under an exchangeable probability. *Probability Theory and Related Fields* **18**(2), 153–158.

Rohlin, V. & Sinai, Y. (1961). Construction and properties of invariant measurable partitions. *Dokl. Akad. Nauk SSSR* **144**, 1038–1041.

Thorisson, H. (2000). *Coupling, stationarity, and regeneration*. Probability and its Applications (New York). New York: Springer-Verlag.

Walters, P. (2000). Convergence of ruelle operator for a function satisfying bowen’s condition. *Transactions of the American Mathematical Society* **353**, 327–347.

1Departmento de Estatística, Instituto de Matemática, Estatística e Ciência de Computação, Universidade de Campinas, Brasil.

*E-mail address: gallesco@ime.unicamp.br*

2Departamento de Métodos Estatísticos, Instituto de Matemática, Universidade Federal de Rio de Janeiro, Brasil.

*E-mail address: sandro@im.ufrj.br*

3Neuroscience Institute and Psychology Department, Princeton University, USA.

*E-mail address: takahashiyd@gmail.com*