On Factor Maps that Send Markov Measures to Gibbs Measures

Jisang Yoo

Abstract Let X and Y be mixing shifts of finite type. Let π be a factor map from X to Y that is fiber-mixing, i.e., given $x, \bar{x} \in X$ with $\pi(x) = \pi(\bar{x}) = y \in Y$, there is $z \in \pi^{-1}(y)$ that is left asymptotic to x and right asymptotic to $\bar{x}$. We show that any Markov measure on X projects to a Gibbs measure on Y under π (for a Hölder continuous potential). In other words, all hidden Markov chains (i.e. sofic measures) realized by π are Gibbs measures. In 2003, Chazottes and Ugalde gave a sufficient condition for a sofic measure to be a Gibbs measure. Our sufficient condition generalizes their condition and is invariant under conjugacy and time reversal. We provide examples demonstrating our result.

Keywords Markov measure · Sofic measure · Gibbs measure · Hidden Markov chain · Fiber-mixing · Thermodynamic formalism

1 Introduction

Markov chains and hidden Markov chains (on a finite set of symbols) are subjects of interest to statisticians and physicists. They are also of interest in the field of symbolic dynamics and ergodic theory. In the context of symbolic dynamics, Markov chains are studied as invariant probability measures on a shift of finite type and are usually called Markov measures. Hidden Markov chains (those arising from Markov chains through deterministic loss of information) are just images of Markov measures under a factor map on a shift of finite type.

An important case where a hidden Markov chain arises is when there is a discrete statistical process which is described by a Markov chain and some of the states taken by the process are not distinguishable for an external observer. In that case, we define a 1-block

J. Yoo (✉)
Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea
e-mail: jisangy@kaist.ac.kr

factor map that lumps the indistinguishable states together, then the projection of the statistical process by this factor map gives another process which represents what the external observer is seeing. Another case is the fuzzy Potts model [9], which is obtained by lumping some states together from the Potts model which generalizes the Ising model.

This work is concerned with the question of when a hidden Markov chain (realized as the projection of a Markov measure by a factor map) is a Gibbs measure (on a one-dimensional lattice) with a Hölder continuous potential. We generalize the result in [3] regarding this question to give a more general sufficient condition that depends only on intrinsic properties of a factor map between general dynamical systems (as opposed to properties that can change when the shift of finite type and the factor map are replaced via a conjugacy).

There are other related results on this question. Pollicott and Kempton proved in [6, 10] that if a fiber-mixing factor map (between shifts of finite type) satisfies the condition that, to decide whether or not some symbol x_0 projecting to z_0 can be extended to the whole point x projecting to z, one only needs to look at a bounded window $z_{-N} \cdots z_N$, then the map sends Gibbs measures to Gibbs measures. Redig and Wang investigated preservation of Gibbsianness in a non-dynamical setting, i.e., where potentials defined on finite regions are not necessarily invariant under the shift map [11]. They also considered statistical transformations in addition to deterministic transformations described here by factor maps.

In general, the projection of a Markov measure is not a Markov measure. So it is natural to ask first when the projection is a Markov measure. There were many approaches to this problem, giving sufficient and necessary conditions for the projection to be a Markov measure. A recent survey can be found in [2]. This topic is also discussed in [5].

Since a lot of probability distributions arising from statistical mechanics can be assumed to be Boltzmann distributions, there has been interest in Gibbs measures which formalize Boltzmann distributions in rigorous mathematical settings (especially in thermodynamical formalism and ergodic theory). The rich and rigorous theory of Gibbs measures and equilibrium states on shifts of finite type is developed and studied by R. Bowen and D. Ruelle [1, 12]. Markov measures (on a shift of finite type) are exactly Gibbs measures with a locally constant potential. So Gibbs measures can be thought of as a generalization of Markov measures. Gibbs measures with Hölder continuous potentials are the most well understood Gibbs measures and are easier to deal with. So it is natural to ask then when a measure (in this work, the projection of a Markov measure) is a Gibbs measure with a Hölder continuous potential.

There are measures that fail to be Gibbsian that arise from other useful contexts or when Gibbsianness is not preserved. Dobrushin thus proposed in [4] the notion of weakly Gibbsianness where the finite-volume Hamiltonian is required to be well defined only on a set of measure 1, and to extend the thermodynamic formalism of Gibbs measures to a broader class of measures like weakly Gibbs measures. On this direction, Lörenczi et al. [8] investigated sufficient conditions for the quasilocality of images of Gibbs measures on the d-dimensional lattice $\mathbb{Z}^d$. Külske et al. [7] demonstrated that the class of weakly Gibbs measures may be too broad for a reasonable thermodynamic formalism and suggested that the class of almost Gibbs measures (conditional probabilities are required to be continuous on a set of measure 1) would be fine.

2 Main Results

In [3], Chazottes and Ugalde came up with the following result:

Let X and Y be mixing 1-step one-sided shifts of finite type. Let $\pi : X \to Y$ be a one-block factor map with the following two properties:

(1) (right eresolving property) for a 2-block ab in Y and a letter $\bar{a}$ that maps to a under π, the letter $\bar{a}$ can be extended to a 2-block $\bar{a}\bar{b}$ in X that maps to ab;

(2) (π-positive words condition) given a periodic point $y \in Y$ with period m less than or equal to the number of letters appearing in Y, for any letters x_0, x_{m-1} respectively mapping to y_0, y_{m-1}, the pair of letters can be extended to a word $x_0 \cdots x_{m-1}$ of length m in X that maps to $y_0 \cdots y_{m-1}$ (the block $y_0 \cdots y_{m-1}$ is said to be π-positive).

Then for any 1-step Markov measure μ on X, the image $\nu = \mu \circ \pi^{-1}$ is the unique Gibbs measure for the Hölder continuous potential f given by

$$f(y) = \lim_{n \to \infty} \log \left(\frac{\nu[y_0 \cdots y_{n-1}]}{\nu[y_1 \cdots y_{n-1}]} \right) \tag{2.1}$$

for $y \in Y$.

The main idea of proving this result was to reduce the problem of showing the Gibbs property of certain hidden Markov chains to one of proving the projective convergence of an inhomogeneous product of nonnegative matrices. In order to do so, the authors of [3] used the contracting property of positive matrices (simply meaning matrices where every entry is positive, in this work) and the idea of Birkhoff's contraction coefficients, which is studied in [13]. Our approach will be similar to theirs except that we apply their argument to matrices whose positive entries form a submatrix. Using this argument, we are able to do away with the first condition (right eresolving property), which was a technical condition required because of the restrain to use only positive matrices. First, we establish the following result where the spaces are 1-step shifts of finite type.

Theorem 2.1 *Let X and Y be mixing 1-step shifts of finite type. Let $\pi : X \to Y$ be a 1-block factor map such that there is $k \in \mathbb{N}$ for which every word $y_1 \cdots y_k$ of length k in Y is π-subpositive, i.e., for any pair of pre-images of $y_0 \cdots y_k$, there is another pre-image that starts from the starting letter of the first pre-image and ends with the last letter of the second pre-image. Then for any 1-step Markov measure μ on X, the image $\nu = \mu \circ \pi^{-1}$ is the unique Gibbs measure for the Hölder continuous potential f given by the formula (2.1).*

A word is π-subpositive if the product of some matrices associated with it is a matrix whose positive entries form a submatrix. This result transfers to the case of one-sided shifts of finite type. We generalize this to the case where shifts of finite type have memory bigger than 1 and the factor maps are not necessarily 1-block codes.

Theorem 2.2 *Let X and Y be mixing shifts of finite type. Let $\pi : X \to Y$ be a factor map that is fiber-mixing. Then for any Markov measure μ on X, the image $\nu = \mu \circ \pi^{-1}$ is the unique Gibbs measure for the Hölder continuous potential f given by the formula (2.1).*

Remark 2.1 Fiber-mixing property is a conjugacy-invariant property, i.e., if $\pi : X \to Y$ is a fiber-mixing factor map between two shifts of finite type and $\tilde{\pi} : \tilde{X} \to \tilde{Y}$ is another factor map such that there are conjugacies $\beta_X : X \to \tilde{X}$, $\beta_Y : Y \to \tilde{Y}$ with $\pi \circ \beta_X = \beta_Y \circ \pi$, then $\tilde{\pi}$ is also fiber-mixing. The property is also time-symmetric, i.e., if $\gamma_X : X \to \gamma_X(X)$ is defined by $\gamma_X(x) = (x_{-i})_{i \in \mathbb{Z}}$ and γ_Y in a similar way, then the factor map $\gamma_Y \circ \pi \circ \gamma_X^{-1} : \gamma_X(X) \to \gamma_Y(Y)$ is also fiber-mixing.

Corollary 2.3 *Let X and Y be mixing shifts of finite type and X be 1-step. Let $\pi : X \to Y$ be a 1-block factor map with the property that there is a finite set $\mathcal{S}$ of π-subpositive words*

in Y such that each point of Y contains some word in $\mathcal{S}$. Then the conclusion of Theorem 2.2
holds.

In Sect. 3 we collect some definitions and basic results. In Sects. 4 and 5 we prove the
main theorems. Some examples demonstrating the results will be given.

3 Preliminaries

A (two-sided) *shift space* or *subshift* is a set X of bi-infinite sequences of symbols from a
fixed alphabet $\mathcal{A}$ (a finite set) such that X is closed with respect to the product topology
of $\mathcal{A}^{\mathbb{Z}}$ and invariant under the shift map $\sigma : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ (defined by $(\sigma x)_i := x_{i+1}$ for any
$x \in \mathcal{A}^{\mathbb{Z}}$ and $i \in \mathbb{Z}$). Let $0 < \theta < 1$. Then the shift space X is a compact metric space with
metric given by

$$d_\theta(x, y) = \theta^{\min\{|j| \, : \, x_j \neq y_j\}}$$

for $x, y \in X$ with $x \neq y$. The metric induces the Tychonoff product topology on X. All dif-
ferent metrics given by different values of θ give rise to the same topology on X. The space
X and σ together form a topological dynamical system. Let $C(X)$ be the set of continuous
functions on X and $\mathcal{M}(X)$ the set of Borel σ-invariant probability measures on X. Denote
by $\mathcal{A}_X$ the alphabet of X. For each $n \in \mathbb{N}$, let $\mathcal{B}_n(X)$ be the set of words that occur in a point
of X. Define $\mathcal{B}(X) = \bigcup_{n=1}^{\infty} \mathcal{B}_n(X)$. We have $\mathcal{A}_X = \mathcal{B}_1(X)$ in particular.

A shift space can be alternatively described by a set of forbidden words. A shift space,
with the alphabet $\mathcal{A}$, determined by the set $\mathcal{F}$ of forbidden words over $\mathcal{A}$ is defined as

$$\mathsf{X}_{\mathcal{F}} = \{x \in \mathcal{A}^{\mathbb{Z}} : \text{no words in } \mathcal{F} \text{ occur in } x \text{ as a subword.}\}$$

All shift spaces can be defined in this way.

A *shift of finite type* or *subshift of finite type* is a shift space that can be described by a
finite set of forbidden words. A shift of finite type is said to be an *n-step shift of finite type*
if it can be described by a set of forbidden words of length $\leq n + 1$. A 1-step shift of finite
type is usually described by a square matrix of 0s and 1s (often called a 0-1 matrix). If A is
the k by k 0-1 matrix, the corresponding shift of finite type is defined as

$$\mathsf{X}_A = \{x \in \{1, \ldots, k\}^{\mathbb{Z}} : A(x_i, x_{i+1}) = 1 \text{ for all } i \in \mathbb{Z}\}.$$

A *factor map* π from a shift space X to another shift space Y is a continuous surjective
map satisfying $\pi \circ \sigma_X = \sigma_Y \circ \pi$. For such a map, there are $m \geq 0$ (memory), $n \geq 0$ (anticipa-
tion) and a map $\Phi : \mathcal{B}_{n+m+1}(X) \to \mathcal{B}_1(Y)$ such that $(\pi(x))_i = \Phi(x_{i-m} \cdots x_{i+n})$ for all $x \in X$
and $i \in \mathbb{Z}$, and π is called a $(m + n + 1)$-block factor map.

A factor map $\pi : X \to Y$ between shift spaces induces a surjective map from $\mathcal{M}(X)$ onto
$\mathcal{M}(Y)$, denoted again by π, which is defined by $(\pi\mu)(B) = \mu(\pi^{-1}(B))$ for $\mu \in \mathcal{M}(X)$ and
a Borel subset B of Y.

3.1 Markov Measures

A 1-step Markov measure on a 1-step shift of finite type X_A determined by an irreducible
0-1 transition matrix A can be defined as follows. Let P be a stochastic matrix compatible

with A, i.e., $A(i, j) > 0$ exactly when $P(i, j) > 0$. Let p be the unique probability vector with $pP = p$. The 1-step *Markov measure* $\mu_P \in \mathcal{M}(\mathsf{X}_A)$ is defined by

$$\mu_P[w] = p(a_0)P(a_0, a_1)P(a_1, a_2) \cdots P(a_{n-1}, a_n)$$

for all $w = a_0 a_1 \cdots a_n \in \mathcal{B}_{n+1}(\mathsf{X}_A)$ and for all $n \in \mathbb{N}$. A measure on a shift of finite type on a shift of finite type is called a *Markov measure* if it is conjugate to a 1-step Markov measure. Every Markov measure is fully supported, i.e., every nonempty open set has positive measure.

Let X be a shift of finite type. A function $f \in C(X)$ is said to be *locally constant* if there exist $m, n \in \mathbb{Z}$ with $m \le n$ such that $f(x) = f(z)$ for any $x, z \in X$ with $x_{[m,n]} = z_{[m,n]}$. A measure $\mu \in \mathcal{M}(X)$ is a Markov measure if and only if it is a (unique) equilibrium state of a locally constant function f on X, i.e., $h(\mu) + \int f d\mu = P(f)$ where $h(\mu)$ is the entropy of μ and $P(f)$ is the topological pressure of f.

3.2 Sofic Measures

A *sofic shift* is an image of a shift of finite type under a factor map. A *sofic measure* (also called a *hidden Markov measure*) on a sofic shift is defined to be the image of a Markov measure under a factor map from a shift of finite type to a sofic shift.

3.3 Gibbs Measures

Let X be a mixing shift of finite type. Let $f \in C(X)$. A measure $\mu \in \mathcal{M}(X)$ is called a *Gibbs measure* for f if there exist $C > 0$ and $P \in \mathbb{R}$ such that

$$\frac{1}{C} < \frac{\mu[x_0 \cdots x_{n-1}]}{\exp(-nP + \sum_{i=0}^{n-1} f(\sigma^i x))} < C$$

for all $x \in X$ and $n \in \mathbb{N}$. The function f is called a *potential* of μ. If $P = 0$, then f is called a *normalized potential* of μ.

Gibbs measures can also be defined using conditional probabilities [12], but the two definitions are equivalent if f is Hölder continuous.

A function $f \in C(X)$ is said to be *Hölder continuous* if there exist $C > 0$ and $\gamma \in (0, 1)$ such that $|f(x) - f(z)| \le C \cdot \gamma^n$ for all $n \in \mathbb{Z}^+$ and $x, z \in X$ with $x_{[-n,n]} = z_{[-n,n]}$, or equivalently, there exist $C, \alpha > 0$ such that

$$|f(x) - f(z)| \le C \cdot (d_\theta(x, z))^\alpha$$

for all $x, z \in X$.

Every Hölder continuous function on X has a unique Gibbs measure. For a detailed account on Gibbs measures on shifts of finite type, consult [1] which contains a proof of the existence and uniqueness of such a measure.

The following lemma describes the relationship between Gibbs measures and their potential functions. A summable sequence $\{c_n\}$ in $\mathbb{R}^+$ means that $\sum_{n=1}^{\infty} c_n < \infty$.

Lemma 3.1 [3] *Let X be a mixing shift of finite type and $\mu \in \mathcal{M}(X)$. Let f be a Hölder continuous function on X and $\{c_n\}$ be a summable sequence in $\mathbb{R}^+$ such that*

$$\left| f(x) - \ln\left(\frac{\mu[x_0 \cdots x_n]}{\mu[x_1 \cdots x_n]}\right) \right| \le c_n$$

for all $x \in X$ and $n \in \mathbb{N}$. Then μ is the Gibbs measure for f and f is its normalized potential.

3.4 Fiber Mixing Factor Maps

Given a 1-block factor map $\pi : X \to Y$, for $a_1 \cdots a_k \in \mathcal{B}_k(X)$, we denote by $\pi(a_1 \cdots a_k)$ the word $\pi(a_1) \cdots \pi(a_k) \in \mathcal{B}_k(Y)$, and for $b_1 \cdots b_k \in \mathcal{B}_k(Y)$, we denote by $\pi^{-1}(b_1 \cdots b_k)$ the set of all words in $\mathcal{B}_k(X)$ that map to $b_1 \cdots b_k$.

Definition 3.1 Let $\pi : X \to Y$ be a 1-block factor map between shift spaces.

(1) A word $w = b_1 \cdots b_k$ in Y is called π-*positive* or *positive* if, given $a_* \in \pi^{-1}(b_1)$ and $a^* \in \pi^{-1}(b_k)$, there is $\tau_1 \cdots \tau_k \in \pi^{-1}(w)$ with $\tau_1 = a_*$ and $\tau_k = a^*$.
(2) A word $w = b_1 \cdots b_k$ in Y is called π-*subpositive* or *subpositive* if, given $u_1 \cdots u_k$ and $v_1 \cdots v_k$ in $\pi^{-1}(w)$, there is $\tau_1 \cdots \tau_k \in \pi^{-1}(w)$ with $\tau_1 = u_1$ and $\tau_k = v_k$.

Let X and Y be 1-step shifts of finite type. Any extension of a subpositive word is also subpositive, i.e., if w is subpositive and $uwv \in \mathcal{B}(Y)$, then uwv is also subpositive.

Two points x, z in a shift space are said to be *left-asymptotic* (*right-asymptotic*, respectively) if there is $n \in \mathbb{Z}$ with $x_{(-\infty,n]} = z_{(-\infty,n]}$ ($x_{[n,\infty)} = z_{[n,\infty)}$, respectively), or equivalently, if $\lim_{n \to -\infty} d(\sigma^n x, \sigma^n z) = 0$ ($\lim_{n \to \infty} d(\sigma^n x, \sigma^n z) = 0$, respectively).

Definition 3.2 Let $\pi : X \to Y$ be a factor map between shift spaces.

(1) π is said to be *right continuing* if, whenever $x \in X$, $y \in Y$, and $\pi(x)$ is left-asymptotic to y, there is $z \in \pi^{-1}(y)$ that is left-asymptotic to x. *Left continuing* can be defined similarly.
(2) π is said to be *fiber-mixing* if, given $x, \bar{x} \in X$ with $\pi(x) = \pi(\bar{x}) = y \in Y$, there is $z \in \pi^{-1}(y)$ that is left asymptotic to x and right asymptotic to $\bar{x}$.
(3) Let π be a 1-block factor map. Then π is said to be *right eresolving* if, whenever $a, b \in \mathcal{A}_Y$ with $ab \in \mathcal{B}_2(Y)$ and $\bar{a} \in \pi^{-1}(a)$, there is $\bar{b} \in \pi^{-1}(b)$ such that $\bar{a}\bar{b} \in \mathcal{B}_2(X)$. *Left eresolving* can be defined similarly.

The following lemma relating the subpositive property to the fiber-mixing property was proved by U. Jung.

Lemma 3.2 *Let X be a 1-step shift of finite type and Y be a sofic shift. Let $\pi : X \to Y$ be a 1-block factor map. Then the following are equivalent.*

i) *There is $k \in \mathbb{N}$ such that every word in $\mathcal{B}_k(Y)$ is π-subpositive.*
ii) *π is fiber-mixing.*

Proof First, assuming that (i) is true, let $x, \bar{x} \in X$ and $\pi(x) = \pi(\bar{x}) = y \in Y$. Since $w = y_1 \cdots y_k \in \mathcal{B}_k(Y)$ is subpositive and $\pi(x_1 \cdots x_k) = \pi(\bar{x}_1 \cdots \bar{x}_k) = w$, there is $\tau_1 \cdots \tau_k \in \pi^{-1}(w)$ with $\tau_1 = x_1$ and $\tau_k = \bar{x}_k$. Define

$$z = \cdots \cdots x_{-2} x_{-1} . x_0 \tau_1 \tau_2 \cdots \tau_k \bar{x}_{k+1} \bar{x}_{k+2} \cdots \cdots.$$

Since X is a 1-step shift of finite type, we have $z \in X$. Also $\pi(z) = y$. Moreover, z is left asymptotic to x and right asymptotic to $\bar{x}$. Thus (i) implies (ii).

Conversely, let π be fiber-mixing. Suppose there is no $k \in \mathbb{N}$ such that every word in $\mathcal{B}_k(Y)$ is subpositive. Then for each $n \in \mathbb{N}$, there exist $x^{(n)}, \bar{x}^{(n)} \in X$ and $y^{(n)} \in Y$ such that

$$\pi(x^{(n)}_{[-n,n]}) = \pi(\bar{x}^{(n)}_{[-n,n]}) = y^{(n)}_{[-n,n]}$$

and there is no $\tau_{-n} \cdots \tau_n \in \pi^{-1}(y_{[-n,n]}^{(n)})$ with $\tau_{-n} = x_{-n}^{(n)}$ and $\tau_n = \bar{x}_n^{(n)}$. Let $(x, \bar{x}, y)$ be a limit point of the sequence $\{(x^{(n)}, \bar{x}^{(n)}, y^{(n)})\}$, which exists because $X \times X \times Y$ is compact. Then $\pi(x) = \pi(\bar{x}) = y$. Since π is fiber-mixing, there is $z \in \pi^{-1}(y)$ and $d > 0$ such that $x_{(-\infty,-d]} = z_{(-\infty,-d]}$ and $\bar{x}_{[d,\infty)} = z_{[d,\infty)}$.

Let s be an integer greater than d such that we have $x_{[-d,d]} = x_{[-d,d]}^{(s)}$, $\bar{x}_{[-d,d]} = \bar{x}_{[-d,d]}^{(s)}$ and $y_{[-d,d]} = y_{[-d,d]}^{(s)}$. Such s exists because $(x, \bar{x}, y)$ is a limit point of the sequence $\{(x^{(n)}, \bar{x}^{(n)}, y^{(n)})\}$. It is easy to check that the word $\tau_{-s} \cdots \tau_s := x_{[-s,-d-1]}^{(s)} z_{[-d,d]} \bar{x}_{[d+1,s]}^{(s)}$ is in $\pi^{-1}(y_{[-s,s]}^{(s)})$. This is a contradiction. Thus (ii) implies (i). $\qquad\square$

Corollary 3.3 *Let X be a 1-step shift of finite type and Y be a sofic shift. Let $\pi : X \to Y$ be a 1-block factor map that is fiber-mixing. Then it is bi-continuing, i.e., both right continuing and left continuing.*

Remark 3.1 It is easy to prove that the condition (i) implies that Y is a k-step shift of finite type. Therefore the image of a shift of finite type under a fiber-mixing factor map is always a shift of finite type.

3.5 The Projective Metric and Subpositive Matrices

Two $m \times n$ nonnegative matrices A and B are said to be *compatible* if they have the same set of indices of positive entries, i.e., $A_{ij} > 0$ if and only if $B_{ij} > 0$. In this case we write $A \equiv B$. Similarly, two nonnegative vectors v, w in $\mathbb{R}^n$ are said to be *compatible* if $v_i > 0$ exactly when $w_i > 0$. We write $v \equiv w$.

Definition 3.3 Let v, w be two compatible nonnegative vectors in $\mathbb{R}^n$. If $v = w = 0$, then put $d(v, w) = 0$. Otherwise, i.e., $v \neq 0$ and $w \neq 0$, then define

$$d(v, w) = \log\left(\frac{\max\{v_i/w_i | v_i > 0\}}{\min\{v_i/w_i | v_i > 0\}}\right).$$

Note that $d(v, w) = 0$ if and only if $v = \alpha w$ for some $\alpha > 0$. If V is a class of compatible nonnegative vectors, then d is a pseudometric on V and is called the *projective metric*.

Lemma 3.4 *Let A be an $m \times n$ nonnegative matrix and v, w be two compatible nonnegative vectors in $\mathbb{R}^n$. Then the following hold.*

(1) *The vectors Av and Aw are compatible and $d(Av, Aw) \leq d(v, w)$.*
(2) *Let $Av \neq 0$ (and hence $Aw \neq 0$). Then*

$$\left|\log\frac{\mathbf{1} \cdot Av}{\mathbf{1} \cdot v} - \log\frac{\mathbf{1} \cdot Aw}{\mathbf{1} \cdot w}\right| \leq d(v, w)$$

where $\mathbf{1} = (1, \ldots, 1)$.

Proof We may assume that $v, w \neq 0$. Clearly, Av and Aw are compatible. Let

$$e^a = \min\{v_i/w_i | v_i > 0\} \quad \text{and} \quad e^b = \max\{v_i/w_i | v_i > 0\}$$

for $a, b \in \mathbb{R}$. Then $e^a w \leq v \leq e^b w$. Since $A \geq 0$, it follows that $e^a A w \leq A v \leq e^b A w$. Thus $d(Av, Aw) \leq d(v, w)$, which verifies (1). Also

$$e^a \cdot \mathbf{1} \cdot w \leq \mathbf{1} \cdot v \leq e^b \cdot \mathbf{1} \cdot w$$

and

$$e^a \cdot \mathbf{1} \cdot A w \leq \mathbf{1} \cdot A v \leq e^b \cdot \mathbf{1} \cdot A w.$$

Thus

$$e^{a-b} \cdot \frac{\mathbf{1} \cdot Aw}{\mathbf{1} \cdot w} \leq \frac{\mathbf{1} \cdot Av}{\mathbf{1} \cdot v} \leq e^{b-a} \cdot \frac{\mathbf{1} \cdot Aw}{\mathbf{1} \cdot w}$$

from which (2) follows. □

Definition 3.4 Let A be an $m \times n$ nonnegative matrix. Define $\tau(A)$ to be the infimum of nonnegative numbers C such that

$$d(Av, Aw) \leq C \cdot d(v, w)$$

for any two compatible vectors v, w in $(\mathbb{R}^n)^+$. That is,

$$\tau(A) = \sup_{\substack{v, w \in (\mathbb{R}^n)^+ \\ v \equiv w, d(v,w) > 0}} \frac{d(Av, Aw)}{d(v, w)}.$$

By Lemma 3.4 (1), we have $\tau(A) \leq 1$. Also $d(Av, Aw) \leq \tau(A)d(v, w)$ for any two compatible vectors v, w in $(\mathbb{R}^n)^+$. It is also easy to see the following.

Lemma 3.5 *Let A be an $m \times n$ nonnegative matrix and B be an $n \times p$ nonnegative matrix. Then $\tau(AB) \leq \tau(A)\tau(B)$.*

Definition 3.5 An $m \times n$ nonnegative matrix A is said to be *subpositive* if there are nonempty subsets $\mathcal{D} \subseteq \{1, \ldots, m\}$ and $\mathcal{E} \subseteq \{1, \ldots, n\}$ such that $A(i, j) > 0$ exactly when $(i, j) \in \mathcal{D} \times \mathcal{E}$.

Lemma 3.6 *Let A be an $m \times n$ nonnegative matrix that is subpositive. Then $\tau(A) < 1$.*

Proof First, assume that $A > 0$. For each nonempty subset $S \subseteq \{1, \ldots, n\}$, let A_S be the $m \times |S|$ submatrix of A corresponding to the index set $\{1, \ldots, m\} \times S$. An argument in [13] shows that there is $0 < C_S < 1$ such that

$$d(A_S v, A_S w) \leq C_S \cdot d(v, w)$$

for any two compatible vectors v, w in $(\mathbb{R}^S)^+$ with $v, w > 0$. It follows that

$$\tau(A) \leq \max_S C_S < 1.$$

In the general case of nonnegative matrices, by rearranging the indices of A, we can assume that A is of the form

$$A = \begin{bmatrix} B & O \\ O & O \end{bmatrix}$$

where B is a $p \times q$ positive matrix with $p, q \geq 1$. Then A can be written as the product

$$A = \begin{bmatrix} I_p \\ O \end{bmatrix} B \begin{bmatrix} I_q & O \end{bmatrix}$$

where I_n is the $n \times n$ identity matrix. By Lemma 3.5 and the above argument, we have $\tau(A) \leq \tau(B) < 1$. This completes the proof. $\qquad\square$

4 Proof of Theorem 2.1

In this section we prove Theorem 2.1. Throughout the section we fix a 1-block factor map $\pi : X \to Y$ between mixing 1-step shifts of finite type. Let X and Y be represented by the 0-1 matrices A and B, respectively. Let μ be a 1-step Markov measure on X and $\nu = \pi \mu \in \mathcal{M}(Y)$.

For each $b \in \mathcal{A}_Y$, let M_b be the $\mathcal{A}_X \times \mathcal{A}_X$ identity matrix. For each $b_1 b_2 \in (\mathcal{A}_Y)^2$, let $M_{b_1 b_2}$ denote the $\mathcal{A}_X \times \mathcal{A}_X$ nonnegative matrix given by

$$M_{b_1 b_2}(i, j) = \begin{cases} \mu[ij]/\mu[j] & \text{if } ij \in \mathcal{B}_2(X) \text{ and } \pi(ij) = b_1 b_2 \\ 0 & \text{otherwise} \end{cases}$$

and m_{b_1} denote the column vector in $(\mathbb{R}^{\mathcal{A}_X})^+$ given by

$$m_{b_1}(i) = \begin{cases} \mu[i] & \text{if } \pi(i) = b_1 \\ 0 & \text{otherwise.} \end{cases}$$

For each $v = b_1 \cdots b_n \in (\mathcal{A}_Y)^n$, $n \geq 2$, define

$$M_v = M_{b_1 \cdots b_n} = M_{b_1 b_2} \cdots M_{b_{n-1} b_n}$$

and

$$m_v = m_{b_1 \cdots b_n} = M_{b_1 \cdots b_n} m_{b_n}.$$

Let $v \in (\mathcal{A}_Y)^n$ for $n \geq 2$. Then for each $i, j \in \mathcal{A}_X$,

$$M_v(i, j) = \sum_{\substack{\omega \in \mathcal{B}_{n-2}(X) \\ i\omega j \in \pi^{-1}(v)}} \frac{\mu[i\omega j]}{\mu[j]}$$

and

$$m_v(i) = \sum_{\substack{\omega \in \mathcal{B}_{n-1}(X) \\ i\omega \in \pi^{-1}(v)}} \mu[i\omega]. \tag{4.1}$$

It is straightforward to prove the following.

Lemma 4.1 *Let $v \in (\mathcal{A}_Y)^n$ for $n \geq 2$. The following are equivalent.*

  i) $v \in \mathcal{B}(Y)$.
 ii) $M_v \neq O$.

iii) $m_v \neq 0$.

Lemma 4.2 *Let* $v = b_1 \cdots b_n \in (\mathcal{A}_Y)^n$ *for* $n \geq 2$. *Then*

$$\nu(v) = \sum_{i \in \mathcal{A}_X} m_v(i) = \mathbf{1} \cdot m_v.$$

Proof Note that μ and ν are fully supported. First, if $v \notin \mathcal{B}(Y)$, i.e., $\nu[v] = 0$, then it follows from Lemma 4.1 that $m_v = 0$. Next, let $v \in \mathcal{B}(Y)$ so that $\nu[v] > 0$. By (4.1),

$$\sum_{i \in \mathcal{A}_X} m_v(i) = \sum_{\substack{i \in \mathcal{A}_X \ \omega \in \mathcal{B}_{n-1}(X) \\ i\omega \in \pi^{-1}(v)}} \mu[i\omega]$$

$$= \sum_{\omega \in \pi^{-1}(v)} \mu[\omega] = \nu[v]$$

which completes the proof. $\square$

We state the relationship between subpositive words and subpositive matrices which is easy to show.

Lemma 4.3 *Let* $v \in \mathcal{B}_n(Y)$ *for* $n \geq 2$. *Then* v *is* π-*subpositive if and only if* M_v *is subpositive.*

For the rest of this section we assume that every word in $\mathcal{B}_k(Y)$ is subpositive, or equivalently, every word in $\mathcal{B}_k(Y)$ contains a subpositive word. We may assume that $k \geq 2$. By Lemma 4.3, for each $v \in \mathcal{B}_k(Y)$, the matrix M_v is subpositive. So $\tau(M_v) < 1$ by Lemma 3.6. Define

$$C = \max_{v \in \mathcal{B}_k(Y)} \tau(M_v) < 1. \tag{4.2}$$

Put $l = k - 1 \geq 1$. Let $p \in \mathbb{Z}^+$ and $v = b_0 \cdots b_{pl} \in \mathcal{B}(Y)$. By Lemma 3.5, we have

$$\tau(M_v) \leq \tau(M_{b_0 \cdots b_l}) \tau(M_{b_l \cdots b_{2l}}) \cdots \tau(M_{b_{(p-1)l} \cdots b_{pl}}) \leq C^p. \tag{4.3}$$

Let $n \geq l$ and $v = b_0 \cdots b_n \in \mathcal{B}_{n+1}(Y)$. By Lemma 4.1, we have

$$m_{b_0 \cdots b_l} = M_{b_0 \cdots b_l} m_{b_l} \neq 0 \quad \text{and} \quad m_v = M_{b_0 \cdots b_l} m_{b_l \cdots b_n} \neq 0. \tag{4.4}$$

Since $M_{b_0 \cdots b_l}$ is subpositive, the vectors $m_{b_0 \cdots b_l}$ and $m_{b_0 \cdots b_n}$ are compatible.

Lemma 4.4 *There is* $E \geq 0$ *such that*

$$d(m_{b_0 \cdots b_l}, m_{b_0 \cdots b_n}) \leq E$$

for all $n \geq l$ *and* $b_0 \cdots b_n \in \mathcal{B}_{n+1}(Y)$.

Proof Let $C < 1$ be given as in (4.2). Define

$$D = \max_{\substack{b_0 \cdots b_n \in \mathcal{B}_{n+1}(Y) \\ l \leq n \leq 2l}} d(m_{b_0 \cdots b_l}, m_{b_0 \cdots b_n}) < \infty.$$

and

$$E = \sum_{i=1}^{\infty} C^{i-1} D = \frac{D}{1-C} < \infty.$$

Let $n = tl + r \geq l$ with $t \geq 1$ and $0 \leq r < l$. Let $b_0 \cdots b_n \in \mathcal{B}_{n+1}(Y)$. Then by (4.4) and (4.3), we have

$$d(m_{b_0 \cdots b_l}, m_{b_0 \cdots b_n}) \leq \sum_{i=1}^{t-1} d(m_{b_0 \cdots b_{il}}, m_{b_0 \cdots b_{(i+1)l}}) + d(m_{b_0 \cdots b_{tl}}, m_{b_0 \cdots b_n})$$

$$\leq \sum_{i=1}^{t-1} \tau(M_{b_0 \cdots b_{(i-1)l}}) d(m_{b_{(i-1)l} \cdots b_{il}}, m_{b_{(i-1)l} \cdots b_{(i+1)l}})$$

$$+ \tau(M_{b_0 \cdots b_{(t-1)l}}) d(m_{b_{(t-1)l} \cdots b_{tl}}, m_{b_{(t-1)l} \cdots b_n})$$

$$\leq \sum_{i=1}^{t-1} C^{i-1} D + C^{t-1} D$$

$$\leq E$$

which completes the proof. $\qquad\square$

Lemma 4.5 *There exist $F \geq 0$ and $0 < G < 1$ such that*

$$d(m_{y_{[1,p]}}, m_{y_{[1,q]}}) \leq FG^p$$

for all $y \in Y$ and $p, q \in \mathbb{N}$ with $k \leq p \leq q$.

Proof Let C be given as in (4.2) and E be given as in Lemma 4.4. Put

$$F = 2EC^{-2-1/l} \quad \text{and} \quad G = C^{1/l}.$$

Let $y \in Y$ and $p, q \in \mathbb{N}$ with $k \leq p \leq q$. Let $p = 1 + tl + r$ with $t \geq 1$ and $0 \leq r < l$. By Lemma 4.4,

$$d(m_{y_{[1,p]}}, m_{y_{[1,1+tl]}}) = \tau(M_{y_{[1,1+(t-1)l]}}) d(m_{y_{[1+(t-1)l, p]}}, m_{y_{[1+(t-1)l, 1+tl]}})$$

$$\leq C^{t-1} E.$$

Similarly, we have

$$d(m_{y_{[1,q]}}, m_{y_{[1,1+tl]}}) \leq C^{t-1} E.$$

Thus

$$d(m_{y_{[1,p]}}, m_{y_{[1,q]}}) \leq 2C^{t-1} E \leq 2C^{(p-1)/l-2} E = FG^p$$

which completes the proof. $\qquad\square$

Now we are ready to prove that the image measure ν is a Gibbs measure.

Proof of Theorem 2.1 Let F and G be given as in Lemma 4.5. For $y \in Y$ and $n \in \mathbb{N}$, define

$$f(y_{[0,n]}) = \log\left(\frac{\nu[y_{[0,n]}]}{\nu[y_{[1,n]}]}\right) = \log\left(\frac{\mathbf{1} \cdot m_{y_{[0,n]}}}{\mathbf{1} \cdot m_{y_{[1,n]}}}\right)$$

$$= \log\left(\frac{\mathbf{1} \cdot M_{y_0 y_1} \cdot m_{y_{[1,n]}}}{\mathbf{1} \cdot m_{y_{[1,n]}}}\right).$$

Let $y \in Y$ and $p, q \in \mathbb{N}$ with $k \le p \le q$. Using Lemma 4.5, we obtain

$$|f(y_{[0,p]}) - f(y_{[0,q]})| \le d(m_{y_{[1,p]}}, m_{y_{[1,q]}}) \le FG^p.$$

Choose $F_1 > 0$ large enough that

$$|f(y_{[0,p]}) - f(y_{[0,k]})| \le F_1 G^k$$

for all $y \in Y$ and $p = 1, \ldots, k$. Let $F_2 = \max\{F + F_1, 2F_1\}$. Then

$$|f(y_{[0,p]}) - f(y_{[0,q]})| \le F_2 G^p$$

for all $y \in Y$ and $p, q \in \mathbb{N}$ with $p \le q$. So, given $y \in Y$, the limit

$$f(y) = \lim_{q \to \infty} f(y_{[0,q]})$$

exists and for any $p \in \mathbb{N}$, we have

$$|f(y_{[0,p]}) - f(y)| \le F_2 G^p.$$

Finally, we show that $f : Y \to \mathbb{R}$ mapping y to $f(y)$ is a Hölder continuous potential. Let $p \in \mathbb{N}$. Choose any $y, z \in Y$ with $y_{[0,p]} = z_{[0,p]}$ and then deduce that

$$|f(y) - f(z)| \le |f(y) - f(y_{[0,p]})| + |f(y_{[0,p]}) - f(z)| \le 2F_2 G^p.$$

Thus f is Hölder continuous. Lemma 3.1 implies that ν is a Gibbs measure for f. $\quad\square$

5 Proof of Theorem 2.2

In this section we prove Theorem 2.2 and Corollary 2.3.

Proof of Theorem 2.2 We may assume that π has memory 0, i.e., there exist $N \in \mathbb{N}$ and a block map Φ such that $(\pi(x))_0 = \Phi(x_{[0,N]})$ for all $x \in X$. (If π has memory $m > 0$, then replace π with $\pi \circ \sigma^k$.) Let X be an m_1-step shift of finite type and μ be an m_1-step Markov measure on X. Let Y be an n-step shift of finite type. Put $m = \max\{m_1, n + N\}$. Let $\tilde{X}$ be the m-th higher block shift of X and $\tilde{Y}$ the n-th higher block shift of Y. Then $\tilde{X}$ and $\tilde{Y}$ are 1-step shifts of finite type. Clearly, $\tilde{\mu} = \mu \circ \beta_X^{-1}$ is a 1-step Markov measure on $\tilde{X}$ where $\beta_X : X \to \tilde{X}$ is the m-th higher block code.

Define

$$\tilde{\pi} = \beta_Y \circ \pi \circ \beta_X^{-1}$$

where $\beta_Y : Y \to \tilde{Y}$ is the n-th higher block code.

$$\begin{array}{ccc} \tilde{X} & \xrightarrow{\tilde{\pi}} & \tilde{Y} \\ \beta_X \uparrow & & \uparrow \beta_Y \\ X & \xrightarrow{\pi} & Y \end{array}$$

Then $\tilde{\pi}$ is a 1-block factor map from $\tilde{X}$ onto $\tilde{Y}$. By Lemma 3.2, the map $\tilde{\pi}$ satisfies the hypothesis of Theorem 2.1. So $\tilde{\nu} = \tilde{\mu} \circ \tilde{\pi}^{-1}$ is a Gibbs measure for the Hölder continuous potential $\tilde{f}$ given by the formula (2.1) corresponding to $\tilde{\nu}$. Let

$$\nu = \mu \circ \pi^{-1} = \tilde{\nu} \circ \beta_Y.$$

Then ν is a Gibbs measure for the function $f = \tilde{f} \circ \beta_Y$. It is easy to see that f is a Hölder continuous function given by the formula (2.1). □

Proof of Corollary 2.3 It is easy to see that π is fiber-mixing. Apply Theorem 2.2. □

6 Examples

In [3] the authors present a way to define a class of factor maps satisfying their original sufficient condition and it can be used to construct various examples demonstrating their result. Here we provide an example of a factor map which is fiber-mixing but is neither right eresolving nor left eresolving.

Example 6.1 Let X be the mixing 1-step shift of finite type defined by the matrix

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

where $\mathcal{A}_X = \{1, 2, 3, 4, 5\}$. Let Y be the mixing 1-step shift of finite type defined by the matrix

$$B = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

where $\mathcal{A}_Y = \{1, 2, 3\}$. Let $\pi : X \to Y$ be the 1-block factor map such that

$$\pi(1) = \pi(5) = 1, \qquad \pi(2) = \pi(4) = 2, \qquad \pi(3) = 3.$$

This example is obtained from the one in [3] by removing two edges.

Consider the word $12 \in \mathcal{B}_2(Y)$ and note that $A_{12} = A_{14} = 0$, i.e., neither 12 nor 14 is in $\mathcal{B}_2(X)$. So π is not right eresolving; similarly, it is not left eresolving (because of the word $32 \in \mathcal{B}_2(Y)$). Thus the original result of Chazottes and Ugalde cannot be applied here. Meanwhile, the word 21 occurs in every point of Y and is subpositive. Therefore, π is fiber-mixing. By Corollary 2.3 every Markov measure on X projects to a Gibbs measure on Y.

Not every factor map sends Markov measures to Gibbs measures. There is a simple example of a factor map from a mixing 1-step shift of finite type onto the full 2-shift for which the limit in the formula (2.1) for some sofic measure does not exist at some point [3]. We give an example of a factor map for which the limit in the formula (2.1) always exists at every point for all images of 1-step Markov measures and such images are Gibbs measures except for some one-parameter family of Markov measures on X.

Example 6.2 Let X be the mixing 1-step shift of finite type defined by the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

where $\mathcal{A}_X = \{1, 2, 3\}$. Let π be a 1-block factor map defined on X such that $\pi(1) = \tilde{1}$ and $\pi(2) = \pi(3) = \tilde{2}$. Then $Y = \pi(X)$ is the 2-step shift of finite type determined by the set $\{\tilde{1}\tilde{1}, \tilde{1}\tilde{2}\tilde{1}\}$ of forbidden blocks. Let μ be the 1-step Markov measure on X defined by the matrix

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & p & 1-p \\ 1-q & 0 & q \end{bmatrix}$$

where $0 < p, q < 1$. Then $p = \mu[22]/\mu[2]$ and $q = \mu[33]/\mu[3]$. Define $\nu = \mu \circ \pi^{-1}$.

A lengthy calculation shows that the limit

$$\lim_{n \to \infty} \log \left(\frac{\nu[y_0 \cdots y_n]}{\nu[y_1 \cdots y_n]} \right)$$

exists for each $y \in Y$. If there is no $k \geq 1$ such that $y_k = \tilde{1}$, then

$$\frac{\nu[y_0 \cdots y_n]}{\nu[y_1 \cdots y_n]} = \frac{\nu[y_0 \tilde{2}^n]}{\nu[\tilde{2}^n]}.$$

If there is such $k \geq 1$ and k is chosen to be minimal, then since $y_1 \cdots y_k = \tilde{2}^{k-1}\tilde{1}$ and there is only one preimage for 1, we have

$$\frac{\nu[y_0 \cdots y_n]}{\nu[y_1 \cdots y_n]} = \frac{\nu[y_0 \tilde{2}^{k-1}\tilde{1}]}{\nu[\tilde{2}^{k-1}\tilde{1}]}.$$

The numerator and denominator of the previous fraction (in each case) are of the form $\nu[a\tilde{2}^m b]$ where $a, b \in \{\tilde{1}, \tilde{2}\}$. Their values can be written as

$$\nu[a\tilde{2}^m b] = \ell_a M_{\tilde{2}\tilde{2}}^{m-1} r_b$$

where ℓ_a, r_b are row and column vectors depending on a, b and

$$M_{\tilde{2}\tilde{2}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & p & \mu[23]/\mu[3] \\ 0 & 0 & q \end{bmatrix}.$$

In the case where $p \neq q$, the logarithmic ratio $\log(v[y_0 \cdots y_n]/v[y_1 \cdots y_1])$ converges exponentially and uniformly. In particular, there exist $F > 0$ and $0 < G < 1$ such that

$$\left| \log\left(\frac{v[y_0 \cdots y_n]}{v[y_1 \cdots y_n]}\right) - \log\left(\frac{v[y_0 \cdots y_m]}{v[y_1 \cdots y_m]}\right) \right| < FG^n$$

for all $y \in Y$ and $n, m \in \mathbb{N}$ with $n < m$. From this, it follows that v is a Gibbs measure for which the limit in the formula (2.1) defines a Hölder continuous potential.

Consider the case where $p = q$ and note that

$$v[\tilde{2}^n] = \mu[2]p^{n-1}\left(2 + n \cdot \frac{1-p}{p}\right)$$

for all $n \in \mathbb{N}$. Put $D = 2\mu[2]/p$ and $E = (1-p)\mu[2]/p^2$, so that

$$v[\tilde{2}^n] = p^n(D + En).$$

If v were a Gibbs measure for a normalized potential f, then it would follow that for some constants $C_1, C_2 > 0$,

$$C_1 < \frac{v[\tilde{2}^n]}{\exp(n \cdot f(\tilde{2}^\infty))} < C_2$$

which would imply that

$$C_1 < \frac{D + En}{\exp(n \cdot f(\tilde{2}^\infty) - n \log p)} < C_2.$$

But this is impossible. Thus v is not a Gibbs measure in the case where $p = q$, although the limit in the formula (2.1) exists.

Summarizing on an intuitive level, the failure of Gibbsianness here comes as follows: as long as preimages of $\tilde{2}^n$ are concerned, all that matters are vertex 2 and vertex 3 and that there is a way to go from 2 to 3 but there is no way from 3 to 2 within preimages of $\tilde{2}^n$. So essentially we have two loops with one way path connecting them. When the loops are weighed the same, i.e., when $p = q$, no loop is dominated by the other loop and therefore all preimages of $\tilde{2}^n$ count the same (up to a bounded factor) and hence $v[\tilde{2}^n]$ equals np^n up to a bounded factor thereby introducing the undesirable factor n.

7 Closing Section

We showed that a fiber-mixing condition is sufficient for the factor map to send all Markov measures to Gibbs measures. The fiber-mixing condition turned out to be conjugacy-invariant and time-symmetric. We presented an example showing that this generalization is nonempty.

It remains open to investigate what happens when Y is a general sofic shift and not a shift of finite type. It can be proved that the image of a shift of finite type under a right-continuing factor map is always a shift of finite type [14]. So if one wants to extend the results to general sofic shifts, one needs to manage to escape the class of right-continuing or left-continuing factor maps.

References

1. Bowen, R.: Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms, 2nd edn. Springer, Berlin (2008)
2. Boyle, M., Petersen, K.: Hidden Markov processes in the context of symbolic dynamics. Preprint, arXiv:0907.1858
3. Chazottes, J., Ugalde, E.: Projection of Markov measures may be Gibbsian. J. Stat. Phys. **111**(5), 1245–1272 (2003)
4. Dobrushin, R.L.: A Gibbsian representation for non-Gibbsian fields. Lecture given at the workshop 'Probability and Physics', September 1995, Renkum
5. Kemeny, J.G., Snell, J.L.: Finite Markov Chains. Van Nostrand, Princeton (1960)
6. Kempton, T.: Factors of Gibbs measures for subshifts of finite type. Preprint
7. Külske, C., Le Ny, A., Redig, F.: Relative entropy and variational properties of generalized Gibbsian measures. Ann. Prob. **32**, 1691–1726 (2004)
8. Lörinczi, J., Maes, C., Velde, K.V.: Transformations of Gibbs measures. Probab. Theor. Relat. Fields **112**, 121–147 (1998)
9. Maes, C., Velde, K.V.: The fuzzy Potts model. J. Phys. A **28**, 4261–4270 (1995)
10. Pollicott, M., Kempton, T.: Factors of Gibbs measures for full shifts. Preprint
11. Redig, F., Wang, F.: Transformations of one-dimensional Gibbs measures with infinite range interaction. Preprint
12. Ruelle, D.: Thermodynamic Formalism. Encyclopedia of Math., vol. 5. Addison-Wesley, Reading (1978)
13. Seneta, E.: Non-negative Matrices and Markov Chains. Springer Series in Statistics. Springer, London (1973)
14. Yoo, J.: On continuing codes. Preprint