Factors of Gibbs Measures for Subshifts of Finite Type

Thomas Kempton

December 11, 2009

Abstract

We give sufficient conditions for the image under projection of a Gibbs measure supported on a subshift of finite type to be a Gibbs measure.

1 Introduction

In the study of subshifts of finite type a natural class of invariant measures are Gibbs measures. They play a particularly important role in symbolic dynamics and the modelling of hyperbolic systems. Somewhat surprisingly the continuous factor of a Gibbs measure, in the sense of Sinai, need no longer be a Gibbs measure.

In this article we consider a projection $\Pi$ from a subshift of finite type $\Sigma_1$ to another subshift. Given a Gibbs measure $\mu$ on $\Sigma_1$ we ask whether the projected measure $\nu := \Pi^*(\mu)$ is a Gibbs measure. In the case that $\mu$ is a Markov measure, sufficient conditions for $\nu$ to be a Gibbs measure were given in [3]. These results were extended in [4], [5] and [7] to deal with the case that $\mu$ is a Gibbs measure and $\Sigma_1$ is a full shift. In the present article we consider the general case of a Gibbs measure over a subshift of finite type.

This problem is of relevance to the study of Hidden Markov processes, which have recently been the subject of intensive study because of their applications to a wide range of problems in pure and applied mathematics. Formally they are factors of Markov systems, but in our context they can be seen as the images of subshifts of finite type under projections which amalgamate symbols. The theory of hidden Markov processes is often applied to the study of systems with Markov dynamics where the observation of the system is not perfect. For example, if two states in a system are indistinguishable then we cannot observe the true transformation, instead we see some factor transformation on the set of equivalence classes of indistinguishable states. This observed transformation may not be Markov even if the original transformation is, and for this reason such factor maps are referred to as Hidden Markov Processes. Recent survey articles by M.Boyle and K.Petersen [2] and by E.Verbitsky [7] give a good introduction to hidden Markov processes and their thermodynamic formalism.

We define subshifts of finite type and Gibbs measures as in [1].
Definition 1.1. Given a finite alphabet $A = \{1, \ldots, k\}$ and a $k \times k$ matrix $M$ of zeros and ones we define the one sided subshift of finite type $(\Sigma, \sigma)$ to be the shift space $\Sigma := \{x \in \{1, \ldots, k\}^{\mathbb{N} \cup \{0\}} : M_{x_ix_{i+1}} = 1 \forall i \in \mathbb{N} \cup \{0\}\}$ coupled with the transformation $\sigma : \Sigma \to \Sigma$, $\sigma(x_0x_1 \cdots) = (x_1x_2 \cdots)$.

Definition 1.2. We call a measure $\mu$ supported on shift space $\Sigma$ a Gibbs measure if there exists a potential $\psi$ and constants $C_1, C_2 > 0$ and $P = P(\psi)$ such that

$$C_1 \leq \frac{\mu[w_0, \ldots, w_{n-1}]}{\exp(\psi^n(w)) - nP(\psi)} \leq C_2$$

for all $w \in \Sigma$, where $\psi^n(w) := \sum_{k=0}^{n-1} \psi(\sigma^k(w))$ and $[w_0, \ldots, w_{n-1}] = \{x \in \Sigma : x_0, \ldots, x_{n-1} = w_0, \ldots, w_{n-1}\}$. We do not require any conditions on the regularity of $\psi$.

By the addition of a suitable coboundary we can normalise the potential so that $P(\psi) = 0$.

Definition 1.3. We define the $n$-th variation of a function $\psi : \Sigma \to \mathbb{R}$

$$var_n(\psi) = \sup\{ |\psi(z) - \psi(w)| : w_i = z_i \text{ for } i = 0, \ldots, n-1\}$$

In particular, $var_0(\psi) = \sup\{ |\psi(z) - \psi(w)| : z, w \in \Sigma\}$.

For a continuous function $\psi$, $\lim_{n \to \infty} var_n(\psi) = 0$. The speed of this convergence gives us the regularity of the function.

Definition 1.4. Suppose we have a projection $\Pi$ from alphabet $\{1, \ldots, k_1\}$ to a smaller alphabet $\{1, \ldots, k_2\}$. This can be extended to a projection from subshift of finite type $\Sigma_1$ over $\{1, \ldots, k_1\}$ to subshift $\Sigma_2$ over $\{1, \ldots, k_2\}$ by $\Pi((x_i)_{i=0}^\infty) = (\Pi(x_i))_{i=0}^\infty$. If such a $\Pi$ is surjective we call it a one block factor map.

Any continuous factor can be represented as an $n$-block factor map for some natural number $n$, by recoding we can reduce to the study of 1-block factor maps.

It was shown in [3] that, in general, the image of a Markov measure on a subshift of finite type need not have a potential defined at all points, and hence need not be a Gibbs measure. We give a further simple example in Section 4. This motivates the following definition and condition.

Definition 1.5. Given a set $B \subset \Sigma_1$ we let $A_n(B)$ be the set of values of $x_n$ for sequences $x$ in $B$.

Hypothesis 1.6. We assume that for $\Pi : \Sigma_1 \to \Sigma_2$ there exists a natural number $N$ such that for any $z \in \Sigma_2$,

1. If $A_n\{x : x_{n+m} = j, \Pi(x) = z\}$ is non empty for some $m > N$, then $A_n\{x : x_{n+m} = j, \Pi(x) = z\} = A_n\{\Pi(x) = z\}$.

2. $A_n\{x : \Pi(x_{n-N} \cdots x_{n+N}) = z_{n-N} \cdots z_{n+N}\} = A_n\{x : \Pi(x) = z\}$. 

The first condition is a mixing condition on fibres, it says that if there exist sequences \( x, x' \in \Pi^{-1}(z) \) then for any \( n \) and \( m > N \) there exists a sequence \( y \in \Pi^{-1}(z) \) with \( y_1, \ldots y_n = x_1, \ldots x_n \) and \( y_{n+m}, \ldots = x'_{n+m} \). The second condition says that in order to know whether some symbol \( x_n \) projecting to \( z_n \) can be extended to a sequence \( x \) projecting to \( z \) one needs only to look locally at \( z_{n-N}, \ldots z_{n+N} \). This is trivially satisfied for full shifts and is weaker than the two topological conditions of [3].

**Example 1.7.** Consider the shift space \( \sigma : \Sigma_1 \rightarrow \Sigma_1 \) associated to the transition matrix

\[
M = \begin{pmatrix}
1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 \\
0 & 1 & 0 & 1 \\
0 & 1 & 1 & 1
\end{pmatrix}
\]

and projection \( \Pi \) from \( \Sigma_1 \) to the full shift on two symbols given by \( \Pi(1) = a, \Pi(2) = \Pi(3) = \Pi(4) = b. \)

Then \( \{2\} = A_1 \{x : \Pi(x_1x_2) = ba\} \neq A_1 \{x : \Pi(x_1) = b\} = \{2, 3, 4\}. \) We also see that putting \( x_1 = 3 \) makes it impossible that \( x_2 = 3 \), but places no restriction on possible values of \( x_3, x_4 \). Thus the hypothesis fails on both counts with \( N = 0 \). Putting \( N = 1 \) it is satisfied.

Up to recoding of the alphabet \( A \) we can assume that \( N = 1 \), and thus the hypothesis implies that \( \Sigma_B \) is a subshift of finite type, and that specifying some digit \( x_n \) in the set of sequences projecting to a word \( z \) only places restrictions on \( x_{n-1} \) and \( x_{n+1} \). We then have the following theorem.

**Theorem 1.8.** Suppose that \( \Pi \) from \( \Sigma_1 \) to \( \Sigma_2 \) satisfies hypothesis 1.6. If \( \mu \) is a Gibbs measure then \( \nu \) is a Gibbs measure. If \( \psi_1 \) is a potential for \( \mu \) and \( \psi_2 \) a potential for \( \nu \) then

1. If \( \text{var}_n(\psi_1) < c_1 \theta_1^{n^2} \) for some \( c_1 > 0, \theta_1 \in (0, 1) \), then \( \text{var}_n(\psi_1) < c_2 \theta_2^{n^2} \) for some \( c_2 > 0, \theta_2 \in (0, 1) \).
2. If \( \sum_{n=0}^{\infty} n^k \text{var}_n(\psi_1) < \infty \) for some \( k \geq 1 \) then \( \sum_{n=0}^{\infty} n^{k-1} \text{var}_n(\psi_2) < \infty \).

This generalises the results of [3], [4], [5] and [7]. In [3] it was shown that the image of a Markov measure is a Gibbs measure provided the projection satisfied two topological conditions. The first of these was slightly stronger than Hypothesis 1.6 above. Chazottes and Ugalde conjectured that the second of their conditions, that any word \( x_0, \ldots, x_n \) projecting to \( z_0, \ldots, z_n \) can be extended to a sequence \( x \) projecting to \( z \), was redundant. We verify this conjecture as a direct corollary of Theorem 1.8. In [4], [5] and [7] the question of projection of full shifts was considered. The results of [4] and [5] follow as corollaries to Theorem 1.8. [7] gives sharper bounds than Theorem 1.8 on the regularity of \( \psi_2 \) in the case that \( \Sigma_1 \) is a full shift.

It was further conjectured in [4] that for any Gibbs measure \( \mu \) on a subshift of finite type \( \Sigma_1 \), the image of \( \mu \) under projection would be a weak Gibbs measure in the sense that there exist constants
2 DEFINING THE POTENTIAL \( \psi_2 \)

In this section we define a sequence of functions which are potentials for measures which approximate \( \nu \). The most technical part of the paper involves demonstrating that the limit of this sequence of potentials converges and satisfies certain regularity conditions, this is deferred until the next section. Here we assume that the limit is well defined and show that it is indeed a potential for \( \nu \).

**Definition 2.1.** We define our projected measure \( \nu \) in terms of \( \mu \),

\[
\nu[z_0, \ldots, z_n] = \sum_{x_0, \ldots, x_n} \mu[x_0, \ldots, x_n]
\]

where the summation is over all words \( x_0, \ldots, x_n \) in \( \Sigma_1 \) projecting to \( z_0, \ldots, z_n \).

A simple calculation verifies that this is indeed a measure.

Since \( \mu \) is a Gibbs measure there exist by definition a potential \( \psi_1 \) and constants \( C_1, C_2 \) such that

\[
C_1 \left( \sum_{x=x_0, \ldots, x_n} \exp(\psi_1^{n+1}(xw(x_n))) \right) \leq \nu[z_0, \ldots, z_n] \leq C_2 \left( \sum_{x=x_0, \ldots, x_n} \exp(\psi_1^{n+1}(xw(x_n))) \right)
\]

(2)

for any choices of \( w(x_n) \) in \( \Sigma_1 \) which can follow \( x_n \). If we can find constants \( k_1, k_2 \) independent of \( n \) and a function \( \psi_2 \) such that

\[
k_1 \left( \sum_{x=x_0, \ldots, x_n} \exp(\psi_2^{n+1}(xw(x_n))) \right) \leq \exp(\psi_2^{n+1}(z)) \leq k_2 \left( \sum_{x=x_0, \ldots, x_n} \exp(\psi_2^{n+1}(xw(x_n))) \right)
\]

(3)

for all \( t \in \Sigma_2 \) and some \( w(x_n) \) in \( \Sigma_1 \) then (2),(3) and (1) give that \( \psi_2 \) will be a potential for \( \nu \).

Dividing by \( \exp(\psi_2^{n}(\sigma(z))) \) we see that such a \( \psi_2 \) would also have to satisfy

\[
\frac{k_1}{k_2} \frac{\sum_{x=x_0, \ldots, x_n} \exp(\psi_2^{n+1}(xw(x_n)))}{\sum_{x'=x_1, \ldots, x_n} \exp(\psi_2^{n}(x'w(x_n)))} \leq \exp(\psi_2(z)) \leq \frac{k_2}{k_1} \frac{\sum_{x=x_0, \ldots, x_n} \exp(\psi_2^{n+1}(xw(x_n)))}{\sum_{x'=x_1, \ldots, x_n} \exp(\psi_2^{n}(x'w(x_n)))}
\]
In [5] we investigated the sequence $u_{\bar{z},n}(\bar{z}) := \frac{\sum_{x=x_0,\ldots,x_n} \exp(\psi_1^{n+1}(xw(x_n)))}{\sum_{x'={x_1,\ldots,x_n}} \exp(\psi_1^{n}(x'w(x_n)))}$, and showed that it led to a definition of $\psi_2$. Because we are dealing with the projections of subshifts rather than full shifts in this work the concatenation of sequences is more difficult, and in particular there is no simple expression for $u_{\bar{z},n+1}(\bar{z})$ as a function of terms $u_{\bar{w},n}(\bar{z})$. This motivates the following refinement of the definition.

Definition 2.2. For $n \in \mathbb{N}$, $j \in A$ and $w$ a sequence in $\Sigma_1$ such that $jw$ is admissable, we define $u_{j,w,n}: \Sigma_2 \to \mathbb{R}$:

$$u_{j,w,n}(\bar{z}) = \frac{\sum_{x=x_0,\ldots,x_{n-1},j} \exp(\psi_1^{n+1}(xw))}{\sum_{x'=x_1,\ldots,x_{n-1},j} \exp(\psi_1^{n}(x'w))}$$

where

1. $\psi_1^n = \sum_{i=0}^n \psi \circ \sigma^i$ and $\psi_1^{n+1} = \sum_{i=0}^{n+1} \psi \circ \sigma^i$;
2. each summation is over finite strings from $\Sigma_1$ for which $x_i$ projects to $z_i$, for $i = 0, \ldots, n$;
3. $xjw \in \Sigma_1$ denotes the concatenation of words to give the sequence $(x_0, \ldots, x_{n-1}, j, w_0, w_1 \cdots)$

We note that there is an explicit dependence on the choices of $w(x_n)$ and $j$ here. However, we will show that the limit $u := \lim_{n \to \infty} u_{j,w,n}$ is a well defined function depending only on $\bar{z} \in \Sigma_2$ and that $\psi_2 := \log u$ is a potential for $\nu$.

Proposition 2.3. $u(\bar{z}) := \lim_{n \to \infty} u_{j,w,n}(\bar{z})$ is well defined and independent of $j, w$

We return to the proof in the next section. We work towards showing that, should $u(\bar{z})$ be well defined, $\log(u)$ is a potential for $\nu$.

Lemma 2.4. There is a constant $C$ depending only on $\psi_1$ such that

$$\frac{1}{\bar{C}} \leq \frac{\sum_{x=x_0,\ldots,x_n} \exp(\psi_1^{n+1}(xw)) \sum_{x'=x_{n+1},\ldots,x_s} \exp(\psi_1^{s-n}(x'w))}{\sum_{x=x_0,\ldots,x_s} \exp(\psi_1^{s+1}(xw))} \leq C$$

Proof. We split the expression into two fractions. Hypothesis 1.6 gives us that a choice $x_n$ cannot affect choices of $x_{n+2}$, since for any $x_n$ there exists an $x_{n+1}$ projecting to $z_{n+1}$ such that $x_n, x_{n+1}, x_{n+2}$ is an admissable word. Given $x_n$, we shall use the notation $\sum_{x=x_{n+1},\ldots,x_s}$ to mean the summation over all words $\bar{x} = x_{n+1} \cdots x_s$ in $\Sigma_1$ projecting to $z_{n+1}, \cdots, z_s$ with the added restriction that $x_n, x_{n+1}$ must be an admissable word in $\Sigma_1$. Given $x_n$ we have

$$1 \leq \frac{\sum_{\bar{x}=x_{n+1},\ldots,x_s} \exp(\psi_1^{s-n}(\bar{x}w))}{\sum_{x=x_{n+1},\ldots,x_s} \exp(\psi_1^{s-n}(\bar{x}w))}$$

$$= \frac{\sum_{\bar{x}=x_{n+2},\ldots,x_s} \sum_{x=x_{n+1}} \exp(\psi_1^{s-n}(\bar{x}w))}{\sum_{\bar{x}=x_{n+2},\ldots,x_s} \sum_{x=x_{n+1}} \exp(\psi_1^{s-n}(\bar{x}w))}$$

$$= \frac{\sum_{\bar{x}=x_{n+2},\ldots,x_s} \exp(\psi_1^{s-n-1}(\bar{x}w)) \sum_{x=x_{n+1}} \exp(\psi_1(\bar{x}w))}{\sum_{\bar{x}=x_{n+2},\ldots,x_s} \exp(\psi_1^{s-n-1}(\bar{x}w)) \sum_{x=x_{n+1}} \exp(\psi_1(\bar{x}w))}$$

$$\leq \exp(var_0(\psi_1)) |A|$$
The final line follows because, given any \( x_n, \hat{x} \), hypothesis 1.6 guarantees the existence of at least one choice of \( \bar{x} \) linking \( x_n \) to \( x_{n+2} \) and there can be at most \( |A| \), thus the ratio of the number of terms can be at most \( |A| \), and for any \( \bar{x} = x_{n+1} \),

\[
\frac{\exp(\psi_1(\bar{x}'\hat{x}w))}{\exp(\psi_1(\bar{x}\hat{x}w))} \leq \exp(\text{var}_0(\psi_1)).
\]

We note that from the definition of a Gibbs measure we have that, for two choices \( w \) and \( w' \),

\[
C_1 \exp(\psi_1^{n+1}(xw)) \leq \mu[x_0, \ldots, x_n] \leq C_2 \exp(\psi_1^{n+1}(xw')),
\]

which gives in particular that for any \( \bar{x} \),

\[
\sum_{\bar{x}=x_0\ldots x_n} \exp(\psi_1^{n+1}(xw)) \leq \sum_{\bar{x}=x_0\ldots x_n} \exp(\psi_1^{n+1}(\bar{x}w)) \leq \frac{C_2}{C_1}.
\]

Returning to our original expression,

\[
\sum_{\bar{x}=x_0\ldots x_n} \exp(\psi_1^{n+1}(xw)) \sum_{\bar{x}=x_{n+1}\ldots x_{m-1}} \exp(\psi_1^{s-n}(\bar{x}w)) = \sum_{\bar{x}=x_0\ldots x_n} \exp(\psi_1^{n+1}(xw)) \sum_{\bar{x}=x_{n+1}\ldots x_{m-1}} \exp(\psi_1^{s-n}(\bar{x}w)) \leq \frac{C_2}{C_1} |A| \exp(\text{var}_0(\psi_1))
\]

so putting \( C = |A| \exp(\text{var}_0(\psi_1)) \frac{C_2}{C_1} \), we are done.

We can now define the potential for \( \nu \):

**Definition 2.5.** We define \( \psi_2 := \log u \).

**Lemma 2.6.** If \( u \) is well defined then \( \psi_2 \) is a potential for \( \nu \).

**Proof.** Fix \( n \geq 1 \). We can write

\[
\psi_2^{n+1}(\bar{z}) = \lim_{m \to +\infty} \log u_{j,w,m}(\bar{z}) + \cdots + \lim_{m \to +\infty} \log u_{j,w,m}(\sigma^n \bar{z})
\]

and then

\[
\psi_2^{n+1}(w) = \lim_{m \to +\infty} \log \left( \frac{\sum_{\bar{x}=x_0\ldots x_{m-1}j} \exp(\psi_1^{m+1}(xw))}{\sum_{\bar{x}=x_{n+1}\ldots x_{m-1}j} \exp(\psi_1^{m+1}(\bar{x}w))} \right).
\]

Moreover, by Lemma 2.4

\[
\sum_{\bar{x}=x_0\ldots x_{m-1}j} \exp(\psi_1^{m+1}(xw)) \leq C \sum_{\bar{x'}=x_0\ldots x_n} \exp(\psi_1^{n+1}(x'w)) \sum_{\bar{x}=x_{n+1}\ldots x_{m-1}j} \exp(\psi_1^{m-n}(\bar{x}w))
\]
so we can bound
\[
\frac{1}{C} \sum_{x' = x_0 \cdots x_n} \exp(\psi_1^{n+1}(x' w)) \leq \sum_{x = x_0 \cdots x_{m-1} \cdot j} \exp(\psi_1^{m+1}(x' w)) \leq C \sum_{x' = x_0 \cdots x_n} \exp(\psi_1^{n+1}(x' w)).
\]

Since \( \mu \) is a Gibbs measure for \( \psi_1 \), there exist constants \( C_1, C_2 > 0 \) such that for any \( \pi(x) = z \) and \( n \geq 1 \):
\[
C_1 \exp(\psi_1^{n+1}(x)) \leq \mu_1[x_0 \cdots x_n] \leq C_2 \exp(\psi_1^{n+1}(x)).
\]

Summing over strings corresponding to \( \pi(x) = z \) gives
\[
\frac{C_1}{C} \leq \frac{\nu[z_0 \cdots z_{n-1}]}{\exp(\psi_2^{n}(z))} \leq C_2 C,
\]

It then follows from the definitions that \( \nu \) is an equilibrium state for \( \psi_2 \).

\[\square\]

### 3 Proof that \( \psi_2 \) is well defined

In this section we will demonstrate that \( \psi_2 \) is well defined and prove properties of the variation. While the details are quite technical, the underlying principles are straightforward. We begin with three definitions which help us quantify how accurate an approximation the function \( u_{j, w, n}(z) \) is to the limit \( u(z) \).

**Definition 3.1.** \( \Lambda_n(z) := [\min_{j, w} u_{j, w, n}(z), \max_{j', w'} u_{j', w', n}(z)] \)

**Definition 3.2.** \( \lambda_n(z) := \sup \left\{ \frac{u_{j, w, n}(z)}{u_{j', w', n}(z)} : w, w' \in \Sigma, j, j' \in A \right\} \)

**Definition 3.3.** \( \lambda_n := \sup_{z \in \Sigma_2} \lambda_n(z) \)

The next lemma shows that \( \Lambda_n(z) \) is a nested sequence, and so the existence of \( \psi_2(z) \) at any point \( z \) corresponds to the convergence to 1 of the decreasing sequence \( \lambda_n(z) \). We also note from the definition that \( u_{j, w, n}(z) \) actually depends only on \( z_0, \cdots z_n \). So \( \Lambda_n(z) \) only depends on \( z_0, \cdots z_n \), and if \( z \) and \( z' \) agree to \( n+1 \) places then \( \Lambda_n(z) = \Lambda_n(z') \). Then the nestedness of \( \Lambda_n \) ensures that \( u(z) \) and \( u(z') \) are both contained in the interval \( \Lambda_n(z) \). Hence \( |\psi_2(z) - \psi_2(z')| \leq \log(\lambda_n(z)) \) and so \( \text{var}_{n+1}(\psi_2) \leq \log(\lambda_n) \).

This section is dedicated to demonstrating that \( \lambda_n \to 1 \) and on obtaining rates of convergence which give the variation of \( \psi_2 \).

**Lemma 3.4.** The sequence of intervals \( \Lambda_n(z) \) is nested.

**Proof.** We observe that
\[
u_{j, w, n+1}(z) = \frac{\sum_{x_n} \text{numerator}(u_{x_n, jw, n}(z)). \exp(\psi_1(j w))}{\sum_{x_n} \text{denominator}(u_{x_n, jw, n}(z)). \exp(\psi_1(j w))} \leq \max_{x_n, w'} u_{x_n, w', n}(z)
\]
where the second line follows because \( \frac{\sum_{k=1}^{n} a_k}{\sum_{k=1}^{n} b_k} \leq \max_{k=1, \ldots, n} \left\{ \frac{a_k}{b_k} \right\} \).

The same observation works for the minimum.

To demonstrate that \( \lambda_n \to 1 \) we define a probability vector which allows us to express the function \( u_{j,w,s} \) in terms of functions \( u_{j',w',n} \) for \( n < s \).

**Definition 3.5.** Let \( 0 < n < s, \ j, x_n \in A \) be fixed. Let \( \overline{x} \) be some choice of word \( x_{n+1}, \ldots, x_s \) compatible with \( x_n \). We then define

\[
P^{(s+2,n)}(x_n, \overline{x}, j, w) = \frac{\sum_{\overline{x} = x_{n+1} \cdots x_s} \sum_{\hat{x} = x_{n+1}j} \exp(\psi_1^{s+2}(\overline{x} \hat{x} w))}{\sum_{\hat{x}' = x_{n+1}j} \exp(\psi_1^{s+2}(\overline{x'} w))}.
\]

By construction, this is a probability vector over choices of \( x_n \) and \( \overline{x} \). (Note that by Hypothesis 1.6 there is always some choice of \( \hat{x} \) linking \( x_s \) to \( j \) and so \( P^{(s+2,n)}(\overline{x}, j, w) \) is never zero for \( \overline{x} \) compatible with \( x_n \).)

**Definition 3.6.** Given \( x_n, \overline{x} = x_{n+1} \cdots x_s, \ w \) and \( j \) we let \( w^{\max} \) be the concatenation \( \hat{x} w \) for the value of \( x_{s+1} \) which maximises \( u_{x_n, \overline{x} \hat{x} w, n}(\hat{z}) \), where \( \hat{x} = x_{s+1}j \). We let \( w^{\min} \) be the string \( \hat{x} w \) which minimizes \( u_{x_n, \overline{x} \hat{x} w, n}(\hat{z}) \).

**Lemma 3.7.**

\[
u_{j,w,s+2}(\hat{z}) \leq \sum_{x_n} \sum_{\overline{x} = x_{n+1} \cdots x_s} u_{x_n, \overline{x} w^{\max}, n}(\hat{z}) P^{(s+2,n)}(x_n, \overline{x}, j, w).
\]

**Proof.** By definition, the numerator of \( u_{j,w,s+2}(\hat{z}) \) is

\[
\sum_{\overline{x} = x_{n+1} \cdots x_s} \sum_{\hat{x} = x_{n+1}j} \exp(\psi_1^{s+3}(\overline{x} w)) = \sum_{x_n} \sum_{\overline{x} = x_{n+1}j} \sum_{\hat{x} = x_{n+1}j} \exp(\psi_1^{n+1}(\overline{x} \hat{x} w)) \exp(\psi_1^{s+2}(\hat{x} w)) \]

where we have used \( \psi_1^{s+3}(\overline{x} \hat{x} w) = \psi_1^{n+1}(\overline{x} \hat{x} w) + \psi_1^{s+2}(\hat{x} w) \). We can further rewrite this as

\[
\sum_{x_n} \sum_{\overline{x} = x_{n+1} \cdots x_s} \sum_{\hat{x} = x_{n+1}j} \left( \sum_{\overline{x} = x_{n+1}j} \exp(\psi_1^{n+1}(\overline{x} \hat{x} w)) \right) \left( \sum_{\hat{x}'} \exp(\psi_1^{s+2}(\hat{x}' w)) \right) \sum_{\hat{x}' \neq \hat{x}} \exp(\psi_1^{s+2}(\hat{x}' w)).
\]

Now we wish to move the summation over \( \hat{x} \) to the second bracket, but we note that the first bracket is not independent of \( \hat{x} \). However using \( w^{\max} \) as defined above we can get an inequality.
\[ \leq \sum_{x_n} \sum_{x_{n+1}}^{x_n} \left( \sum_{z=x_n}^{x_{n+1}} \exp(\psi_1^{n+1}(z T w_{\text{max}})) \right) \left( \sum_{z=x_n}^{x_{n+1}} \exp(\psi_1^n(z T w_{\text{max}})) \right) \]

So by dividing by the denominator of \( u_{j,w,s+2}(z) \), which equals the denominator of \( P^{s+2,n}(x_n, \bar{x}, j, w) \), we see

\[ u_{j,w,s+2}(z) \leq \sum_{x_n} \sum_{x_{n+1}}^{x_n} \frac{u_{x_n, x_{n+1}, \cdots, x_s}}{u_{x_n, x_{n+1}, \cdots, x_s}} \cdot \frac{P^{s+2,n}(x_n, \bar{x}, j, w)}{P^{s+2,n}(x_n, \bar{x}, j', w')} \]

We note that the only dependence on \( j \) in the above is in \( P^{s+2,n}(x_n, \bar{x}, j, w) \), in particular, all the summations are over sets which are independent of \( j \).

\[ \square \]

Corollary 3.8.

\[ \frac{u_{j,w,s+2}(z)}{u_{j',w',s+2}(z)} \leq \sum_{x_n} \sum_{x_{n+1}}^{x_n} \frac{u_{x_n, x_{n+1}, \cdots, x_s}}{u_{x_n, x_{n+1}, \cdots, x_s}} \cdot \frac{P^{s+2,n}(x_n, \bar{x}, j, w)}{P^{s+2,n}(x_n, \bar{x}, j', w')} \]

This follows from using \( w_{\text{min}} \) in the previous lemma for the denominator. Now we restate lemma 4.7 from [5], for completeness the proof is included as an appendix.

Lemma 3.9. Suppose that for all \( j, j' \in \Pi^{-1}(z_s), w, v \in \Sigma_1, s > n + 1: \)

1. \[ u_{j,w,s+2}(z) \leq \sum_{x_n} \sum_{x_{n+1}}^{x_n} \frac{u_{x_n, x_{n+1}, \cdots, x_s}}{u_{x_n, x_{n+1}, \cdots, x_s}} \cdot \frac{P^{s+2,n}(x_n, \bar{x}, j, w)}{P^{s+2,n}(x_n, \bar{x}, j', v)} \]

2. there exists \( c \in (0, 1) \) with \( c < \frac{P^{s+2,n}(x_n, \bar{x}, j, w)}{P^{s+2,n}(x_n, \bar{x}, j', v)} \) \( \forall x_n, \bar{x}, j, k, w, v, s > n \); and

3. \[ \frac{u_{x_n, x_{n+1}, \cdots, x_s}}{u_{x_n, x_{n+1}, \cdots, x_s}} \leq \exp(2 \sum_{k=s-n}^s \text{var}_k(\psi_1)) \]

Then \[ \frac{u_{j,w,s+2}(z)}{u_{j',w',s+2}(z)} \leq c \exp \left( 2 \sum_{k=s-n}^s \text{var}_k(\psi_1) \right) + (1 - c) \max_{j,j',w,v} \left( \frac{u_{j,w,n}(z)}{u_{j',w',n}(z)} \right), \]
giving

\[ \lambda_{s+2}(z) \leq c \exp \left( 2 \sum_{k=s-n}^s \text{var}_k(\psi_1) \right) + (1 - c) \lambda_n(z). \]
Since each \( \lambda_n \) is finite we can take suprema over \( z \) in the previous inequality to get

\[
\lambda_{s+2} \leq c \exp \left( 2 \sum_{k=s-n}^{s} \var_k(\psi_1) \right) + (1 - c) \lambda_n.
\]

In particular, by making sufficiently large choices of \( s \) when iterating, this gives that that \( \lambda_n \) tends to one and so \( \psi_2 := \log u \) is well defined and continuous. The following appears as theorems 5.2 and 5.3 of [5].

**Lemma 3.10.** Suppose \( \lambda_n \) is a sequence such that, for \( s > n \),

\[
\lambda_{s+2} \leq c \exp(2 \sum_{k=s-n}^{s} \var_k(\psi_1)) + (1 - c) \lambda_n.
\]

1. If \( \var_n(\psi_1) < c_1 \theta_1 \sqrt{n} \) for some \( c_1 > 0, \theta_1 \in (0, 1) \) then \( \log(\lambda_n) < c_2 \theta_2 \sqrt{n} \) for some \( c_2 > 0, \theta_2 \in (0, 1) \)

2. If \( \sum_{n=0}^{\infty} n^k \var_n(\psi_1) < \infty \) for some \( k \geq 1 \) then \( \sum_{n=0}^{\infty} n^{k-1} \log(\lambda_n) < \infty \)

The following two lemmas prove that the conditions of Lemma 3.9 are satisfied, and so \( \psi_2 \) is well defined. Observing that \( \var_n(\psi_2) \leq \log(\lambda_n) \), this gives the appropriate regularity results for \( \psi_2 \).

**Lemma 3.11.** There is a uniform bound

\[
c \leq \frac{P^{(s+2,n)}(x_n, \overline{x}, j, \overline{w})}{P^{(s+2,n)}(x_n, \overline{x}, j', \overline{w}')} \]

**Proof.** We can write

\[
P^{(s+2,n)}(x_n, \overline{x}, j, \overline{w}) = \frac{\sum_{\overline{x} = x_1 \cdots x_n} \sum_{\overline{x} = x_{s+1}j} \exp(\psi_1^{s+2}(\overline{x}\overline{x}\overline{w}))}{\sum_{\overline{x}' = x_1' \cdots x_n'} \sum_{\overline{x} = x_{s+1}j} \exp(\psi_1^{s+2}(\overline{x}'\overline{x}'))}.\]

We consider first the numerator. Given a choice of \( \overline{x} \), changing \( j \) can only effect possible choices of \( x_{s+1} \). There will always be at least one choice of \( x_{s+1} \) linking \( x_s \) to \( j \) by Hypothesis 1.6, and there can be at most \( |A| \). So the number of terms in the summation for different choices of \( j \) can differ by a factor of at most \( |A| \). Furthermore, given \( \overline{x} = x_0 \cdots x_n, \overline{x} = x_{n+1} \cdots x_s, \overline{x}, \overline{x}', j, j', \overline{w}, \overline{w}' \), we have by lemma 2.4 that

\[
\frac{\exp(\psi_1^{s+2}(\overline{x}\overline{x}\overline{w}))}{\exp(\psi_1^{s+2}(\overline{x}'\overline{x}'\overline{w}'))} = \frac{\exp(\psi_1^{s}(\overline{x}\overline{x}\overline{w})))}{\exp(\psi_1^{s}(\overline{x}'\overline{x}'\overline{w}'))} \quad \leq C. \exp(2\var_0(\psi_1))
\]

Making identical calculations for the denominator we see that the lemma is proved with

\[
c = \frac{1}{|A|^2 C^2 \exp(4\var_0(\psi_1))}
\]

\[\square\]
We now need only to confirm that the third condition of Lemma 3.9 is satisfied in order to complete the proof of Theorem 1.8.

Lemma 3.12. \( \frac{u_{x_n, x_{w_{\text{max}}, n}}}{u_{x_n, x_{w_{\text{min}}, n}}} \leq \exp(2 \sum_{k=s-n}^{s} \text{var}_k(\psi_1)) \)

Proof. We recall that \( \bar{x} \) was some choice of \( x_{n+1}, \cdots, x_s \). Considering first the numerators, we see that

\[
\frac{\text{numerator}(u_{x_n, x_{w_{\text{max}}, n}})}{\text{numerator}(u_{x_n, x_{w_{\text{min}}, n}})} = \frac{\sum_{\bar{x}=x_0, \ldots, x_n} \exp(\psi_1^{n+1}(\bar{x}x_{w_{\text{max}}}))}{\sum_{\bar{x}=x_0, \ldots, x_n} \exp(\psi_1^{n+1}(\bar{x}x_{w_{\text{min}}})})
\]

comparing termwise we see that \( \sigma^k(x_{w_{\text{min}}}) \) and \( \sigma^k(x_{w_{\text{max}}}) \) agree to \( s-n+(n-k) \) places, and thus for any choice of \( \bar{x} \),

\[
\frac{\exp(\psi_1^{n+1}(\bar{x}x_{w_{\text{max}}}))}{\exp(\psi_1^{n+1}(\bar{x}x_{w_{\text{min}}}))} \leq \exp \left( \sum_{k=s-n}^{s} \text{var}_k(\psi_1) \right)
\]

Summing over all choices of \( \bar{x} \) and making the identical calculations for the denominator the lemma is proved.

Hence \( u_{j, w, n}(z) \) satisfies the conditions of Lemma 3.9, and so by Lemma 3.10 we have that \( \psi_2 \) is a well defined potential for \( \nu \) satisfying the regularity conditions of Theorem 1.8. Hence Theorem 1.8 is proved.

4 Examples and Comments

First we give an example which shows that some condition such as Hypothesis 1.6 on the projection \( \Pi : \Sigma_1 \to \Sigma_2 \) is necessary, indeed without any conditions it possible that the projected measure \( \nu \) need not be a weak Gibbs measure.

Definition 4.1. A measure \( \nu \) is called a weak Gibbs measure if there exists a Hölder continuous function \( \psi_2 : \Sigma_2 \to \mathbb{R} \) such that

\[
C_1 \leq \frac{\nu[z_0, \cdots, z_{n-1}]}{\exp(\psi_2^n(z))} \leq C_2
\]

for almost all \( z \in \Sigma \).

Example 4.2. Consider the shift space \( \sigma : \Sigma_1 \to \Sigma_1 \) associated to the transition matrix

\[
M = \begin{pmatrix}
1 & 1 & 1 & 0 \\
1 & 1 & 0 & 1 \\
1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1
\end{pmatrix}
\]
and consider a factor map \( \Pi : \Sigma_1 \to \Sigma_2 \) with \( \Pi(1) = 1, \Pi(2) = 2 \) and \( \Pi(3) = \Pi(4) = 3 \). Let \( \psi_1 : \Sigma_1 \to \mathbb{R} \) be a Hölder continuous function (such that \( P(\psi_1) = 0 \)) with Gibbs measure \( \mu \). We suppose that \( \Pi \ast \mu = \nu \) is a weak Gibbs measure. Then, for almost all \( w \),

\[
C_1 \exp(\psi_1^n(13\cdots3w)) \leq \nu[1,3,\cdots,3] = \mu[1,3,\cdots,3]
\]

and

\[
C_2 \exp(\psi_1^n(23\cdots3w)) \geq \nu[2,3,\cdots,3] = \mu[2,4,\cdots,4]
\]

If we further suppose that \( \mu \) is a Bernoulli measure with \( \mu[3] < \mu[4] \), we need only take \( n \) large enough such that

\[
\frac{\mu[1]\mu[3]n}{C_1 \exp(\inf_z \psi_2(z))} < \frac{\mu[2]\mu[4]n}{C_2 \exp(\sup_z \psi_2(z))}
\]

and we see

\[
\psi_2^n(3\cdots3w) < \psi_2^n(3\cdots3w)
\]

for any \( w \), thus \( \psi_2^n \) is undefined on \([3,\cdots,3]_n\) which is a set of positive measure. So there is a set of positive measure on which \( \nu \) does not satisfy the Bowen Gibbs inequality (1), and thus \( \nu \) cannot be a weak Gibbs measure.

We now consider a class of potentials for which Hölder continuity is preserved under projection.

**Example 4.3.** Suppose that \( \psi_1 \) can be expressed as

\[
\psi_1(x) = f_0(x_0, x_1) + f_1(x_1, x_2) + \cdots
\]

Then Hölder continuity of \( \psi_1 \) implies Hölder continuity of \( \psi_2 \).

**Proof.** Given some choice of \( w_0 \), the dependence of \( u_{j,w,n}(z) \) on the later terms in \( w \) is less than \( \text{var}_n(\psi_1) \). This is because, given \( x_0 \cdots x_n \) and \( w, w' \) with \( w_0 = w'_0 \),

\[
\psi_1(\sigma^i(x_0, \cdots, x_nw)) - \psi_1(\sigma^i(x_0, \cdots, x_nw')) = \sum_{k=0}^{\infty} f_{n-i+k}(w_k, w_{k+1}) - f_{n-i+k}(w'_k, w'_{k+1})
\]

and hence

\[
\frac{\sum_{x=x_0\cdots x_{n-1}j} \exp(\psi_1^{n+1}(xw))}{\sum_{x=x_0\cdots x_{n-1}j} \exp(\psi_1^{n+1}(xw'))} \leq \exp \left( \sum_{i=0}^{n} \sum_{k=0}^{\infty} f_{n-i+k}(w_k, w_{k+1}) - f_{n-i+k}(w'_k, w'_{k+1}) \right)
\]

12
This is independent of the choice of $\pi$. Thus we have

\[
\frac{u_j, w, n(z)}{u_j, w', n(z)} = \frac{\sum_{x=x_0 \cdots x_{n-1} j} \exp(\psi_1^{n+1}(xw))}{\sum_{x=x_0 \cdots x_{n-1} j} \exp(\psi_1^n(xw'))} \frac{\sum_{x'=x_1 \cdots x_{n-1} j} \exp(\psi_1^n(x'w'))}{\sum_{x=x_0 \cdots x_{n-1} j} \exp(\psi_1^{n+1}(xw'))} = \frac{\sum_{x=x_0 \cdots x_{n-1} j} \exp(\psi_1^{n+1}(xw))}{\sum_{x=x_0 \cdots x_{n-1} j} \exp(\psi_1^n(xw'))}.
\]

The appearance of $\exp(2 \sum_{k=s}^n \text{var}_k(\psi_1))$ in the statement of Lemma 3.7 appears as a maximal value of $\frac{u_j, w, n(z)}{u_j, w', n(z)}$. Choosing $s = n + 1$ and putting $\pi = w_0$, the statement of lemma 3.7 now becomes

$$
\lambda_{n+3} \leq c(\exp(\text{var}_n(\psi_1))) + (1 - c)\lambda_n
$$

which in particular gives that H"older potentials project to H"older potentials. This generalizes the result in [3], where it was shown that Gibbs measures with locally constant potentials (Markov measures) project to Gibbs measures with H"older potentials.

It was demonstrated in [7] that H"older continuity is preserved under projection in the case that $\Sigma_1$ is a full shift. The question of whether the same is true for subshifts of finite type remains open.

## 5 Appendix: Proof of Lemma 3.9

Lemma 3.9 was proved in [5] but we repeat the proof here for completeness.

We recall that we had

\[
\frac{u_j, w, s+2(z)}{u_j, w', s+2(z)} = \frac{\sum_{x_1} \sum_{\pi=x_1 \cdots x_s} u_{x_1, \pi, w_{\text{max}}, n}(z) P^{(s+2,n)}(x_1, \pi, j, w)}{\sum_{x_1} \sum_{\pi=x_1 \cdots x_s} u_{x_1, \pi, w_{\text{min}}, n}(z) P^{(s+2,n)}(x_1, \pi, j', w')}. 
\]

To simplify notation, we fix $z$ and rewrite $\sum_{x_1} \sum_{\pi=x_1 \cdots x_s}$ as $\sum_{i \in I}$, letting $i$ represent $x_n \pi$. We let $P_1$ represent the probability vector $P^{(s+2,n)}(x_1, \pi, j, w)$, $P_2$ represent the probability vector $P^{(s+2,n)}(x_1, \pi, j', w')$, $A$ represent vector $(u_{x_1, \pi, w_{\text{max}}, n})$ and $B$ represent vector $(u_{x_1, \pi, w_{\text{min}}, n})$, all over choices of $i \in I = x_n \pi$. Let $a_i, b_i$ represent the $i$th terms of $A$ and $B$ respectively. The above inequality now becomes
\[
\frac{u_{j,w,s+2}(z)}{u_{j',w',s+2}(z)} \leq \frac{P_1 \cdot A}{P_2 \cdot B},
\]

where, under the hypotheses of Lemma 3.9, there is a universal constant \(c\) such that \(c < \frac{P_1(i)}{P_2(i)}\), and 

\[
\frac{A(i)}{B(i)} \leq \exp(2 \sum_{k=s-n}^s \var_k(\psi_1))
\]

We claim that 

\[
\frac{u_{j,w,s+2}(z)}{u_{j',w',s+2}(z)} \leq c \cdot \exp \left( 2 \sum_{k=s-n}^s \var_k(\psi_1) \right) + \max_{j,j',w,w'} \left( \frac{u_{j,w,n}(z)}{u_{j',w',n}(z)} \right) + (1 - c) \cdot \max_{j,j',w,w'} \left( \frac{u_{j,w,n}(z)}{u_{j',w',n}(z)} \right)
\]

**Proof:**

We assume that 

\[
\max_{j,j',w,w'} \left( \frac{u_{j,w,n}(z)}{u_{j',w',n}(z)} \right) > \exp \left( 2 \sum_{k=s-n}^s \var_k(\psi_1) \right),
\]

otherwise

\[
\frac{u_{j,w,s+2}(z)}{u_{j',w',s+2}(z)} \leq (c + (1 - c)) \max_{j,j',w,w'} \left( \frac{u_{j,w,n}(z)}{u_{j',w',n}(z)} \right) \leq c \exp \left( 2 \sum_{k=s-n}^s \var_k(\psi_1) \right) + (1 - c) \max_{j,j',w,w'} \left( \frac{u_{j,w,n}(z)}{u_{j',w',n}(z)} \right)
\]

as required.

Now we use \(c\) to write

\[
\frac{u_{j,w,s+2}(z)}{u_{j',w',s+2}(z)} \leq \frac{(c \cdot P_1 \cdot A) + ((1 - c) \cdot P_1 \cdot A)}{(c \cdot P_1 \cdot B) + ((P_2 - cP_1) \cdot B)}
\]

noting that \(P_2 - cP_1 \geq 0\). We will use \(1\) to represent a vector of all 1s of length |I|.

Now 

\[
A \leq \exp \left( 2 \sum_{k=s-n}^s \var_k(\psi_1) \right) \cdot B,
\]

so

\[
\frac{u_{j,w,s+2}(z)}{u_{j',w',s+2}(z)} \leq \frac{c \cdot \exp \left( 2 \sum_{k=s-n}^s \var_k(\psi_1) \right) \cdot P_1 \cdot B + ((1 - c) \cdot P_1 \cdot A)}{(c \cdot P_1 \cdot B) + ((P_2 - cP_1) \cdot B)} \leq \frac{(c \cdot \exp \left( 2 \sum_{k=s-n}^s \var_k(\psi_1) \right) \cdot P_1 \cdot B + ((1 - c) \cdot P_1 \cdot 1 \cdot \max_i(a_i))}{(c \cdot P_1 \cdot B) + ((P_2 - cP_1) \cdot 1 \cdot \min_i(b_i))} \leq \frac{(c \cdot \exp \left( 2 \sum_{k=s-n}^s \var_k(\psi_1) \right) \cdot P_1 \cdot 1 \cdot \min_i(b_i) + ((1 - c) \cdot P_1 \cdot 1 \cdot \max_i(a_i))}{(c \cdot P_1 \cdot 1 \cdot \min_i(b_i)) + ((P_2 - cP_1) \cdot 1 \cdot \min_i(b_i))}
\]

The justification for the last step is that we assumed 

\[
\frac{\max_i(a_i)}{\min_i(b_i)} > \exp \left( 2 \sum_{k=s-n}^s \var_k(\psi_1) \right),
\]

so shrinking terms on the top and bottom which are similar leads to more weight being given to those which are dissimilar. Of course a probability vector dotted with a vector of 1s gives 1, so we can divide by \(\min_i(b_i)\) to get

\[
\frac{u_{j,w,s+2}(z)}{u_{j',w',s+2}(z)} \leq \frac{c \cdot \exp \left( 2 \sum_{k=s-n}^s \var_k(\psi_1) \right) + (1 - c) \cdot \frac{\max_i(a_i)}{\min_i(b_i)}}{c + (1 - c)}
\]

\[
= c \cdot \exp \left( 2 \sum_{k=s-n}^s \var_k(\psi_1) \right) + (1 - c) \cdot \max_{j,j',w,w'} \left( \frac{u_{j,w,n}(z)}{u_{j',w',n}(z)} \right)
\]
as required.

References


