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 Π from a subshift of finite type Σ_1 to another subshift. Given a Gibbs measure μ on Σ_1 we ask whether the projected measure $\nu := \Pi^*(\mu)$ is a Gibbs measure. In the case that μ is a Markov measure, sufficient conditions for ν to be a Gibbs measure were given in [3]. These results were extended in [4], [5] and [7] to deal with the case that μ is a Gibbs measure and Σ_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 $A = \{1, \cdots, k\}$ and a $k \times k$ matrix M of zeros and ones we define the one sided subshift of finite type (Σ_M, σ) to be the shift space $\Sigma_M := \{\underline{x} \in \{1, \cdots, k\}^{\mathbb{N} \cup \{0\}} : M_{x_i x_{i+1}} = 1 \forall i \in \mathbb{N} \cup \{0\}\}$ coupled with the transformation $\sigma : \Sigma_M \to \Sigma_M$, $\sigma(x_0 x_1 \cdots) = (x_1 x_2 \cdots)$.

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

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

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

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 $\psi : \Sigma \to \mathbb{R}$

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

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

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

Definition 1.4. Suppose we have a projection Π from alphabet $\{1, \cdots, k_1\}$ to a smaller alphabet $\{1, \cdots, k_2\}$. This can be extended to a projection from subshift of finite type Σ_1 over $\{1, \cdots, k_1\}$ to subshift Σ_2 over $\{1, \cdots, k_2\}$ by $\Pi((x_i)_{i=0}^{\infty}) = (\Pi(x_i))_{i=0}^{\infty}$. If such a Π 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 $\mathcal{A}_n(B)$ be the set of values of x_n for sequences $\underline{x}$ 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 $\underline{z} \in \Sigma_2$,*

1. *If $\mathcal{A}_n\{\underline{x} : x_{n+m} = j, \Pi(\underline{x}) = \underline{z}\}$ is non empty for some $m > N$, then $\mathcal{A}_n\{\underline{x} : x_{n+m} = j, \Pi(\underline{x}) = \underline{z}\} = \mathcal{A}_n\{\Pi(\underline{x}) = \underline{z}\}$.*
2. *$\mathcal{A}_n\{\underline{x} : \Pi(x_{n-N} \cdots x_{n+N}) = z_{n-N} \cdots z_{n+N}\} = \mathcal{A}_n\{\underline{x} : \Pi(\underline{x}) = \underline{z}\}$.*

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

Example 1.7. Consider the shift space $\sigma : \Sigma_1 \to \Sigma_1$ 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 Π from Σ_1 to the full shift on two symbols given by

$$\Pi(1) = a, \Pi(2) = \Pi(3) = \Pi(4) = b.$$

Then $\{2\} = \mathcal{A}_1\{\underline{x} : \Pi(x_1x_2) = ba\} \neq \mathcal{A}_1\{\underline{x} : \Pi(x_1) = b\} = \{2, 3, 4\}$. 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 Σ_B is a subshift of finite type, and that specifying some digit x_n in the set of sequences projecting to a word $\underline{z}$ only places restrictions on x_{n-1} and x_{n+1}. We then have the following theorem.

Theorem 1.8. *Suppose that Π from Σ_1 to Σ_2 satisfies hypothesis 1.6. If μ is a Gibbs measure then ν is a Gibbs measure. If ψ_1 is a potential for μ and ψ_2 a potential for ν then*

1. *If $var_n(\psi_1) < c_1\theta_1^{\sqrt{n}}$ for some $c_1 > 0, \theta_1 \in (0, 1)$, then $var_n(\psi_1) < c_2\theta_2^{\sqrt{n}}$ 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 $\sum_{n=0}^{\infty} n^{k-1} var_n(\psi_2) < \infty$*

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 $x_0, \cdots x_n$ projecting to $z_0, \cdots, z_n$ can be extended to a sequence $\underline{x}$ projecting to $\underline{z}$, 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 ψ_2 in the case that Σ_1 is a full shift.

It was further conjectured in [4] that for any Gibbs measure μ on a subshift of finite type Σ_1, the image of μ under projection would be a weak Gibbs measure in the sense that there exist constants

C_1, C_2 such that the inequality in Definition 1.2 is satisfied almost everywhere. An example in section 4 shows this to be false. We believe that, while Hypothesis 1.6 could potentially be weakened, the principle that a choice of z_0 cannot affect potential choices of z_n for arbitrarily large n is crucial to the validity of the theorem, and thus that the theorem probably cannot be extended to projections of subshifts of finite type onto more general subshifts.

In Section 2 we will define a function ψ_2 and show that, should it be well defined, it is a potential for ν. Section 3 is dedicated to demonstrating that ψ_2 is well defined and showing that the variation behaves as in Theorem 1.8. In Section 4 we give an example and define a class of potentials for which Hölder continuity is preserved under projection.

2 Defining the Potential ψ_2

In this section we define a sequence of functions which are potentials for measures which approximate ν. 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 ν.

Definition 2.1. We define our projected measure ν in terms of μ,

$$\nu[z_0, ..., z_n] = \sum_{x_0, \cdots, x_n} \mu[x_0, ..., x_n]$$

where the summation is over all words $x_0, \cdots, x_n$ in Σ_1 projecting to $z_0, \cdots, z_n$.

A simple calculation verifies that this is indeed a measure.

Since μ is a Gibbs measure there exist by definition a potential ψ_1 and constants C_1, C_2 such that

$$C_1 \left(\sum_{\underline{x}=x_0,\cdots,x_n} \exp(\psi_1^{n+1}(\underline{xw}(x_n))) \right) \leq \nu[z_0, ..., z_n] \leq C_2 \left(\sum_{\underline{x}=x_0,\cdots,x_n} \exp(\psi_1^{n+1}(\underline{xw}(x_n))) \right) \quad (2)$$

for any choices of $\underline{w}(x_n)$ in Σ_1 which can follow x_n. If we can find constants k_1, k_2 independent of n and a function ψ_2 such that

$$k_1 \left(\sum_{\underline{x}=x_0,\cdots,x_n} \exp(\psi_1^{n+1}(\underline{xw}(x_n))) \right) \leq \exp(\psi_2^{n+1}(\underline{z})) \leq k_2 \left(\sum_{\underline{x}=x_0,\cdots,x_n} \exp(\psi_1^{n+1}(\underline{xw}(x_n))) \right) \quad (3)$$

for all $\underline{t} \in \Sigma_2$ and some $\underline{w}(x_n)$ in Σ_1 then (2),(3) and (1) give that ψ_2 will be a potential for ν. Dividing by $\exp(\psi_2^n(\sigma(\underline{z})))$ we see that such a ψ_2 would also have to satisfy

$$\frac{k_1}{k_2} . \frac{\sum_{\underline{x}=x_0,\cdots,x_n} \exp(\psi_1^{n+1}(\underline{xw}(x_n)))}{\sum_{\underline{x}'=x_1,\cdots,x_n} \exp(\psi_1^{n}(\underline{x}'\underline{w}(x_n)))} \leq \exp(\psi_2(\underline{z})) \leq \frac{k_2}{k_1} . \frac{\sum_{\underline{x}=x_0,\cdots,x_n} \exp(\psi_1^{n+1}(\underline{xw}(x_n)))}{\sum_{\underline{x}'=x_1,\cdots,x_n} \exp(\psi_1^{n}(\underline{x}'\underline{w}(x_n)))}$$

In [5] we investigated the sequence $u_{\underline{w},n}(\underline{z}) := \dfrac{\sum_{\underline{x}=x_0,\cdots,x_n} \exp(\psi_1^{n+1}(\underline{x}\underline{w}(x_n)))}{\sum_{\underline{x}'=x_1,\cdots,x_n} \exp(\psi_1^{n}(\underline{x}'\underline{w}(x_n)))}$, and showed that it led to a definition of ψ_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 $u_{\underline{w},n+1}(\underline{z})$ as a function of terms $u_{\underline{w}',n}(\underline{z})$. This motivates the following refinement of the definition.

Definition 2.2. For $n \in \mathbb{N}$, $j \in A$ and $\underline{w}$ a sequence in Σ_1 such that $j\underline{w}$ is admissable, we define $u_{j,\underline{w},n} : \Sigma_2 \to \mathbb{R}$:

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

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 Σ_1 for which x_i projects to z_i, for $i = 0, \ldots, n$;
3. $\underline{x}j\underline{w} \in \Sigma_1$ denotes the concaternation of words to give the sequence $(x_0, \cdots, x_{n-1}, j, w_0, w_1 \cdots)$

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

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

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

Lemma 2.4. *There is a constant C depending only on ψ_1 such that*

$$\frac{1}{C} \le \frac{\sum_{\underline{x}=x_0\cdots x_n} \exp(\psi_1^{n+1}(\underline{x}\underline{w})) \sum_{\overline{x}=x_{n+1}\cdots x_s} \exp(\psi_1^{s-n}(\overline{x}\underline{w}))}{\sum_{\underline{x}=x_0\cdots x_s} \exp(\psi_1^{s+1}(\underline{x}\underline{w}))} \le C$$

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 $\sum_{\overline{x}=x_{n+1}\cdots x_s}^{x_n}$ to mean the summation over all words $\overline{x} = x_{n+1}\cdots x_s$ in Σ_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 Σ_1. Given x_n we have

$$\begin{aligned}
1 &\le \frac{\sum_{\overline{x}=x_{n+1}\cdots x_s} \exp(\psi_1^{s-n}(\overline{x}\underline{w}))}{\sum_{\overline{x}=x_{n+1}\cdots x_s}^{x_n} \exp(\psi_1^{s-n}(\overline{x}\underline{w}))} \\
&= \frac{\sum_{\hat{x}=x_{n+2}\cdots x_s} \sum_{\overline{x}=x_{n+1}}^{x_{n+2}} \exp(\psi_1^{s-n}(\overline{x}\hat{x}\underline{w}))}{\sum_{\hat{x}=x_{n+2}\cdots x_s} \sum_{\overline{x}=x_{n+1}}^{x_n,x_{n+2}} \exp(\psi_1^{s-n}(\overline{x}\hat{x}\underline{w}))} \\
&= \frac{\sum_{\hat{x}=x_{n+2}\cdots x_s} \exp(\psi_1^{s-n-1}(\hat{x}\underline{w})) \sum_{\overline{x}=x_{n+1}}^{x_{n+2}} \exp(\psi_1(\overline{x}\hat{x}\underline{w}))}{\sum_{\hat{x}=x_{n+2}\cdots x_s} \exp(\psi_1^{s-n-1}(\hat{x}\underline{w})) \sum_{\overline{x}=x_{n+1}}^{x_n,x_{n+2}} \exp(\psi_1(\overline{x}\hat{x}\underline{w}))} \\
&\le \exp(var_0(\psi_1)).|A|
\end{aligned}$$

The final line follows because, given any $x_n, \hat{x}$, hypothesis 1.6 guarantees the existence of at least one choice of $\overline{x}$ 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 $\overline{x} = x_{n+1}$,

$$\frac{\exp(\psi_1(\overline{x'}\hat{x}\underline{w}))}{\exp(\psi_1(\overline{x}\hat{x}\underline{w}))} \leq \exp(var_0(\psi_1)).$$

We note that from the definition of a Gibbs measure we have that, for two choices $\underline{w}$ and $\underline{w}'$,

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

which gives in particular that for any $\overline{x}$,

$$\frac{\sum_{\underline{x}=x_0\cdots x_n} \exp(\psi_1^{n+1}(\underline{xw}))}{\sum_{\underline{x}=x_0\cdots x_n} \exp(\psi_1^{n+1}(\underline{x}\overline{x}\underline{w}))} \leq \frac{C_2}{C_1}.$$

Returning to our original expression,

$$\begin{aligned} & \frac{\sum_{\underline{x}=x_0\cdots x_n} \exp(\psi_1^{n+1}(\underline{xw})) \sum_{\overline{x}=x_{n+1}\cdots x_s} \exp(\psi_1^{s-n}(\overline{x}\underline{w}))}{\sum_{\underline{x}=x_0\cdots x_s} \exp(\psi_1^{s+1}(\underline{xw}))} \\ = \; & \frac{\sum_{\underline{x}=x_0\cdots x_n} \exp(\psi_1^{n+1}(\underline{xw}))}{\sum_{\underline{x}=x_0\cdots x_n} \exp(\psi_1^{n+1}(\underline{x}\overline{x}\underline{w}))} \frac{\sum_{\overline{x}=x_{n+1}\cdots x_s} \exp(\psi_1^{s-n}(\overline{x}\underline{w}))}{\sum_{\overline{x}=x_{n+1}\cdots x_s}^{x_n} \exp(\psi_1^{s-n}(\overline{x}\underline{w}))} \\ \leq \; & \frac{C_2}{C_1}|A| \exp(var_0(\psi_1)) \end{aligned}$$

so putting $C = |A| \exp(var_0(\psi_1))\frac{C_2}{C_1}$ we are done. $\square$

We can now define the potential for ν:

Definition 2.5. We define $\psi_2 := \log u$.

Lemma 2.6. *If u is well defined then ψ_2 is a potential for ν.*

Proof. Fix $n \geq 1$. We can write

$\psi_2^{n+1}(\underline{z}) = \lim_{m\to+\infty} \log u_{j,\underline{w},m}(\underline{z}) + \cdots + \lim_{m\to+\infty} \log u_{j,\underline{w},m}(\sigma^n \underline{z})$ and then

$$\psi_2^{n+1}(\underline{w}) = \lim_{m\to+\infty} \log \left(\frac{\sum_{\underline{x}=x_0\cdots x_{m-1}j} \exp(\psi_1^{m+1}(\underline{xw}))}{\sum_{\overline{x}=x_{n+1}\cdots x_{m-1}j} \exp(\psi_1^{m+1}(\overline{x}\underline{w}))} \right).$$

Moreover, by Lemma 2.4

$$\sum_{\underline{x}=x_0\cdots x_{m-1}j} \exp(\psi_1^{m+1}(\underline{xw})) \leq C \sum_{\underline{x}'=x_0\cdots x_n} \exp(\psi_1^{n+1}(\underline{x}'\underline{w})) \sum_{\overline{x}=x_{n+1}\cdots x_{m-1}j} \exp(\psi_1^{m-n}(\overline{x}\underline{w}))$$

so we can bound

$$\frac{1}{C}\sum_{\underline{x}'=x_0\cdots x_n}\exp(\psi_1^{n+1}(\underline{x}'\underline{w}))\leq\underbrace{\frac{\sum_{\underline{x}=x_0\cdots x_{m-1}j}\exp(\psi_1^{m+1}(\underline{x}\underline{w}))}{\sum_{\overline{x}=x_{n+1}\cdots x_{m-1}j}\exp(\psi_1^{m-n}(\overline{x}\underline{w}))}}_{=\exp(\sum_{i=0}^{n}\log u_{j,\underline{w},(m-i)}(\sigma^i\underline{z}))})\leq C\sum_{\underline{x}'=x_0\cdots x_n}\exp(\psi_1^{n+1}(\underline{x}'\underline{w})).$$

Since μ is a Gibbs measure for ψ_1, there exist constants $C_1, C_2 > 0$ such that for any $\pi(\underline{x}) = \underline{z}$ and $n \geq 1$:

$$C_1\exp(\psi_1^{n+1}(\underline{x}))\leq\mu_1[x_0\cdots x_n]\leq C_2\exp(\psi_1^{n+1}(\underline{x})).$$

Summing over strings corresponding to $\pi(\underline{x}) = \underline{z}$ gives

$$\frac{C_1}{C}\leq\frac{\nu[z_0\cdots z_{n-1}]}{\exp(\psi_2^n(\underline{z}))}\leq C_2C,$$

It then follows from the definitions that ν is an equilibrium state for ψ_2. □

3 Proof that ψ_2 is well defined

In this section we will demonstrate that ψ_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 $u_{j,\underline{w},n}(\underline{z})$ is to the limit $u(\underline{z})$.

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

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

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

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

This section is dedicated to demonstrating that $\lambda_n \to 1$ and on obtaining rates of convergence which give the variation of ψ_2.

Lemma 3.4. *The sequence of intervals $\Lambda_n(\underline{z})$ is nested.*

Proof. We observe that

$$\begin{aligned} u_{j,\underline{w},n+1}(\underline{z}) &= \frac{\sum_{x_n}^{j}\text{numerator}(u_{x_n,j\underline{w},n}(\underline{z})).\exp(\psi_1(j\underline{w}))}{\sum_{x_n}^{j}\text{denominator}(u_{x_n,j\underline{w},n}(\underline{z})).\exp(\psi_1(j\underline{w}))} \\ &\leq \max_{x_n,\underline{w}'} u_{x_n,\underline{w}',n}(\underline{z}) \end{aligned}$$

where the second line follows because $\dfrac{\sum_{k=1}^n a_k}{\sum_{k=1}^n b_k} \leq \max_{k=1,\cdots,n}\left\{\dfrac{a_k}{b_k}\right\}$.

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 $u_{j,\underline{w},s}$ in terms of functions $u_{j',\underline{w}',n}$ 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 $x_{n+1}, \cdots, x_s$ compatible with x_n. We then define

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

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 $P^{(s+2,n)}(\overline{x}, j, \underline{w})$ is never zero for $\overline{x}$ compatible with x_n.)

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

Lemma 3.7.

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

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

$$\sum_{\underline{x}=x_0\cdots x_{s+1}j} \exp(\psi_1^{s+3}(\underline{x}\underline{w})) = \sum_{x_n} \sum_{\underline{x}=x_0\cdots x_n}^{x_n} \sum_{\overline{x}=x_{n+1}\cdots x_s}^{x_n} \sum_{\hat{x}=x_{s+1},j}^{x_s} \exp(\psi_1^{n+1}(\underline{x}\overline{x}\hat{x}\underline{w})) \exp(\psi_1^{s+2-n}(\overline{x}\hat{x}\underline{w}))$$

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

$$\sum_{x_n} \sum_{\overline{x}=x_{n+1},\cdots x_s}^{x_n} \sum_{\hat{x}=x_{s+1}j}^{x_s} \underbrace{\left(\frac{\sum_{\underline{x}=x_0\cdots x_n}^{x_n} \exp(\psi_1^{n+1}(\underline{x}\overline{x}\hat{x}\underline{w}))}{\sum_{\underline{x}'=x_1\cdots x_n}^{x_n} \exp(\psi_1^{n}(\underline{x}'\overline{x}\hat{x}\underline{w}))}\right)}_{u_{x_n,\overline{x}\hat{x}\underline{w},n}(\underline{z})} \underbrace{\left(\sum_{\underline{x}'=x_1\cdots x_n}^{x_n} \exp(\psi_1^{n}(\underline{x}'\overline{x}\hat{x}\underline{w}))\right) \exp(\psi_1^{s+2-n}(\overline{x}\hat{x}\underline{w}))}_{\sum_{\underline{x}'=x_1\cdots x_n} \exp(\psi_1^{s+2}(\underline{x}'\overline{x}\hat{x}\underline{w}))}.$$

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 $\underline{w}^{\max}$ as defined above we can get an inequality.

$$\leq \sum_{x_n} \sum_{\overline{x}=x_{n+1},\cdots x_s}^{x_n} \underbrace{\left(\frac{\sum_{\underline{x}=x_0\cdots x_n}^{x_n} \exp(\psi_1^{n+1}(\underline{x}\overline{x}\underline{w}^{\max}))}{\sum_{\underline{x}'=x_1\cdots x_n}^{x_n} \exp(\psi_1^{n}(\underline{x}'\overline{x}\underline{w}^{\max}))} \right)}_{u_{x_n,\overline{x}\underline{w}^{\max},n}(\underline{z})} \underbrace{\left(\sum_{\hat{x}=x_{s+1}j}^{x_s} \sum_{\underline{x}'=x_1\cdots x_n}^{x_n} \exp(\psi_1^{s+2}(\underline{x}'\overline{x}\hat{x}\underline{w})) \right)}_{numerator(P^{s+2,n}(x_n,\overline{x},j,\underline{w}))}$$

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

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

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

Corollary 3.8.

$$\frac{u_{j,\underline{w},s+2}(\underline{z})}{u_{j',\underline{w}',s+2}(\underline{z})} \leq \frac{\sum_{x_n} \sum_{\overline{x}=x_{n+1}\cdots x_s}^{x_n} u_{x_n,\overline{x}\underline{w}^{max},n}(\underline{z}) P^{(s+2,n)}(x_n,\overline{x},j,\underline{w})}{\sum_{x_n} \sum_{\overline{x}=x_{n+1}\cdots x_s}^{x_n} u_{x_n,\overline{x}\underline{w'}^{min},n}(\underline{z}) P^{(s+2,n)}(x_n,\overline{x},j',\underline{w}'))}.$$

This follows from using $\underline{w}^{\min}$ 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 $j,j' \in \Pi^{-1}(z_s), \underline{w},\underline{v} \in \Sigma_1,\ s > n+1$:*

1. $\dfrac{u_{j,\underline{w},s+2}(\underline{z})}{u_{j',\underline{v},s+2}(\underline{z})} \leq \dfrac{\sum_{x_n} \sum_{\overline{x}=x_{n+1},\cdots,x_s}^{x_n} u_{x_n,\overline{x}\underline{w}^{\max},n}(\underline{z}).P^{s+2,n}(x_n,\overline{x},j,\underline{w})}{\sum_{x_n} \sum_{\overline{x}=x_{n+1},\cdots,x_s}^{x_n} u_{x_n,\overline{x}\underline{v}^{\min},n}(\underline{z}).P^{s+2,n}(x_n,\overline{x},j',\underline{v})}$;
2. *there exists $c \in (0,1)$ with $c < \dfrac{P^{s+2,n}(x_n,\overline{x},j,\underline{w})}{P^{s+2,n}(x_n,\overline{x},j',\underline{v})}$ $\forall x_n,\overline{x},j,k,\underline{w},\underline{v},s > n$; and*
3. $\dfrac{u_{x_n,\overline{x}\underline{w}^{\max},n}(\underline{z})}{u_{x_n,\overline{x}\underline{v}^{\min},n}(\underline{z})} \leq \exp(2\sum_{k=s-n}^{s} var_k(\psi_1))$.

Then $\dfrac{u_{j,\underline{w},s+2}(\underline{z})}{u_{j',\underline{v},s+2}(\underline{z})} \leq c.\exp\left(2\sum_{k=s-n}^{s} var_k(\psi_1)\right) + (1-c).max_{j,j',\underline{w},\underline{v}} \left(\dfrac{u_{j,\underline{w},n}(\underline{z})}{u_{j',\underline{v},n}(\underline{z})}\right)$, *giving*

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

Since each λ_n is finite we can take suprema over $\underline{z}$ 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 λ_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 λ_n is a sequence such that, for $s > n$,*

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

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 ψ_2 is well defined. Observing that $var_n(\psi_2) \leq \log(\lambda_n)$, this gives the appropriate regularity results for ψ_2.

Lemma 3.11. *There is a uniform bound*

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

Proof. We can write

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

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 $\underline{x} = x_0 \cdots x_n, \overline{x} = x_{n+1} \cdots x_s, \hat{x}, \hat{x}', j, j', \underline{w}, \underline{w}'$, we have by lemma 2.4 that

$$\frac{\exp(\psi_1^{s+2}(\underline{x}\overline{x}\hat{x}\underline{w}))}{\exp(\psi_1^{s+2}(\underline{x}\overline{x}\hat{x}'\underline{w}'))} = \frac{\exp(\psi_1^{s}(\underline{x}\overline{x}\hat{x}\underline{w}))}{\exp(\psi_1^{s}(\underline{x}\overline{x}\hat{x}'\underline{w}'))} \frac{\exp(\psi_1^{2}(\hat{x}\underline{w}))}{\exp(\psi_1^{2}(\hat{x}'\underline{w}'))} \leq C. \exp(2var_0(\psi_1))$$

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

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

□

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. $\frac{u_{x_n,\overline{x}\underline{w}^{max},n}}{u_{x_n,\overline{x}\underline{w}^{min},n}} \leq \exp(2\sum_{k=s-n}^{s} var_k(\psi_1))$

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

$$\frac{\text{numerator}(u_{x_n,\overline{x}\underline{w}^{max},n})}{\text{numerator}(u_{x_n,\overline{x}\underline{w}^{min},n})} = \frac{\sum_{\underline{x}=x_0\cdots x_n} \exp(\psi_1^{n+1}(\underline{x}\overline{x}\underline{w}^{max}))}{\sum_{\underline{x}=x_0\cdots x_n} \exp(\psi_1^{n+1}(\underline{x}\overline{x}\underline{w}^{min}))}$$

comparing termwise we see that $\sigma^k(\underline{x}\overline{x}\underline{w}^{min})$ and $\sigma^k(\underline{x}\overline{x}\underline{w}^{max})$ agree to $s-n+(n-k)$ places, and thus for any choice of $\underline{x}$,

$$\frac{\exp(\psi_1^{n+1}(\underline{x}\overline{x}\underline{w}^{max}))}{\exp\left(\psi_1^{n+1}(\underline{x}\overline{x}\underline{w}^{min})\right)} \leq \exp\left(\sum_{k=s-n}^{s} var_k(\psi_1)\right)$$

Summing over all choices of $\underline{x}$ and making the identical calculations for the denominator the lemma is proved. □

Hence $u_{j,\underline{w},n}(\underline{z})$ satisfies the conditions of Lemma 3.9, and so by Lemma 3.10 we have that ψ_2 is a well defined potential for ν 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 ν need not be a weak Gibbs measure.

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

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

for almost all $\underline{z} \in \Sigma$.

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 μ. We suppose that $\Pi_*\mu = \nu$ is a weak Gibbs measure. Then, for almost all $\underline{w}$,

$$C_1 \exp(\psi_2^{n+1}(1\underbrace{3\cdots3}_{n}\underline{w})) \leq \nu[1, \underbrace{3, \cdots, 3}_{n}] = \mu[1, \underbrace{3, \cdots, 3}_{n}]$$

and

$$C_2 \exp(\psi_2^{n+1}(2\underbrace{3\cdots3}_{n}\underline{w})) \geq \nu[2, \underbrace{3, \cdots, 3}_{n}] = \mu[2, \underbrace{4, \cdots, 4}_{n}]$$

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

$$\frac{\mu[1]\mu[3]^n}{C_1 \exp(\inf_{\underline{z}} \psi_2(\underline{z}))} < \frac{\mu[2]\mu[4]^n}{C_2 \exp(\sup_{\underline{z}} \psi_2(\underline{z}))}$$

and we see

$$\psi_2^n(\underbrace{3\cdots3}_{n}\underline{w}) < \psi_2^n(\underbrace{3\cdots3}_{n}\underline{w})$$

for any $\underline{w}$, thus ψ_2^n is undefined on $[\underbrace{3, \cdots, 3}_{n}]$ which is a set of positive measure. So there is a set of positive measure on which ν does not satisfy the Bowen Gibbs inequality (1), and thus ν 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 ψ_1 can be expressed as

$$\psi_1(\underline{x}) = f_0(x_0, x_1) + f_1(x_1, x_2) + \cdots$$

Then Hölder continuity of ψ_1 implies Hölder continuity of ψ_2.

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

$$\psi_1(\sigma^i(x_0, \cdots, x_n\underline{w})) - \psi_1(\sigma^i(x_0 \cdots x_n\underline{w}')) = \sum_{k=0}^{\infty} f_{n-i+k}(w_k, w_{k+1}) - f_{n-i+k}(w_k', w_{k+1}')$$

and hence

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

This is independent of the choice of $\underline{x}$. Thus we have

$$\begin{aligned}
\frac{u_{j,\underline{w},n}(\underline{z})}{u_{j,\underline{w}',n}(\underline{z})} &= \frac{\sum_{\underline{x}=x_0\cdots x_{n-1}j}\exp(\psi_1^{n+1}(\underline{xw}))}{\sum_{\underline{x}'=x_1\cdots x_{n-1}j}\exp(\psi_1^{n}(\underline{x}'\underline{w}))}\cdot\frac{\sum_{\underline{x}'=x_1\cdots x_{n-1}j}\exp(\psi_1^{n}(\underline{x}'\underline{w}'))}{\sum_{\underline{x}=x_0\cdots x_{n-1}j}\exp(\psi_1^{n+1}(\underline{xw}'))} \\
&= \frac{\sum_{\underline{x}=x_0\cdots x_{n-1}j}\exp(\psi_1^{n+1}(\underline{xw}))}{\sum_{\underline{x}=x_0\cdots x_{n-1}j}\exp(\psi_1^{n+1}(\underline{xw}'))}\frac{\sum_{\underline{x}'=x_1\cdots x_{n-1}j}\exp(\psi_1^{n}(\underline{x}'\underline{w}'))}{\sum_{\underline{x}'=x_1\cdots x_{n-1}j}\exp(\psi_1^{n}(\underline{x}'\underline{w}))} \\
&= \frac{\exp(\sum_{i=0}^{n}\sum_{k=0}^{\infty} f_{n-i+k}(w_k,w_{k+1}) - f_{n-i+k}(w'_k,w'_{k+1}))}{\exp(\sum_{i=1}^{n}\sum_{k=0}^{\infty} f_{n-i+k}(w_k,w_{k+1}) - f_{n-i+k}(w'_k,w'_{k+1}))} \\
&= \exp(\sum_{k=0}^{\infty} f_{n+k}(w_k,w_{k+1}) - f_{n+k}(w'_k,w'_{k+1})) \\
&\leq var_n(\psi_1)
\end{aligned}$$

The appearance of $\exp(2\sum_{k=s-n}^{s} var_k(\psi_1))$ in the statement of Lemma 3.7 appears as a maximal value of $\frac{u_{j,\overline{x}\underline{w},n}(\underline{z})}{u_{j,\overline{x}\underline{w}',n}(\underline{z})}$. Choosing $s=n+1$ and putting $\overline{x}=w_0$, the statement of lemma 3.7 now becomes

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

which in particular gives that Hölder potentials project to Hölder 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 Hölder potentials. □

It was demonstrated in [7] that Hölder continuity is preserved under projection in the case that Σ_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

$$\frac{u_{j,\underline{w},s+2}(\underline{z})}{u_{j',\underline{w}',s+2}(\underline{z})} \leq \frac{\sum_{x_n}\sum_{\overline{x}=x_{n+1}\cdots x_s}^{x_n} u_{x_n,\overline{x}\underline{w}^{max},n}(\underline{z})P^{(s+2,n)}(x_n,\overline{x},j,\underline{w})}{\sum_{x_n}\sum_{\overline{x}=x_{n+1}\cdots x_s}^{x_n} u_{x_n,\overline{x}\underline{w}'^{min},n}(\underline{z})P^{(s+2,n)}(x_n,\overline{x},j',\underline{w}'))}.$$

To simplify notation, we fix $\underline{z}$ and rewrite $\sum_{x_n}\sum_{\overline{x}=x_{n+1}\cdots x_s}^{x_n}$ as $\sum_{i\in I}$, letting i represent $x_n\overline{x}$. We let P_1 represent the probability vector $P^{(s+2,n)}(x_n,\overline{x},j,\underline{w})$, P_2 represent the probability vector $P^{(s+2,n)}(x_n,\overline{x},j',\underline{w}')$, A represent vector $(u_{x_n,\overline{x}\underline{w}^{max},n})$ and B represent vector $(u_{x_n,\overline{x}\underline{w}'^{min},n})$, all over choices of $i\in I=x_n\overline{x}$. Let a_i,b_i represent the ith terms of A and B respectively. The above inequality now becomes

$$\frac{u_{j,\underline{w},s+2}(\underline{z})}{u_{j',\underline{w}',s+2}(\underline{z})} \leq \frac{P_1 \cdot A}{P_2 \cdot B},$$

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

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

Proof:

We assume that $\max_{j,j',\underline{w},\underline{v}}\left(\frac{u_{j,\underline{w},n}(\underline{z})}{u_{j',\underline{v},n}(\underline{z})}\right) > \exp\left(2\sum_{k=s-n}^{s} var_k(\psi_1)\right)$, otherwise

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

as required.

Now we use c to write

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

noting that $P_2 - cP_1 \geq 0$. We will use $\mathbb{1}$ to represent a vector of all 1s of length $|I|$.
Now $A \leq \exp\left(2\sum_{k=s-n}^{s} var_k(\psi_1)\right) B$, so

$$\begin{aligned}
\frac{u_{j,\underline{w},s+2}(\underline{z})}{u_{j',\underline{v},s+2}(\underline{z})} &\leq \frac{\left(c.\exp\left(2\sum_{k=s-n}^{s} var_k(\psi_1)\right) P_1 \cdot B\right) + ((1-c).P_1 \cdot A)}{(c.P_1 \cdot B) + ((P_2 - cP_1) \cdot B)} \\
&\leq \frac{\left(c.\exp\left(2\sum_{k=s-n}^{s} var_k(\psi_1)\right) P_1 \cdot B\right) + ((1-c).P_1 \cdot \mathbb{1}.\max_i(a_i))}{(c.P_1 \cdot B) + ((P_2 - cP_1) \cdot \mathbb{1}\min_i(b_i))} \\
&\leq \frac{\left(c.\exp\left(2\sum_{k=s-n}^{s} var_k(\psi_1)\right) P_1 \cdot \mathbb{1}\min_i(b_i)\right) + ((1-c).P_1 \cdot \mathbb{1}.\max_i(a_i))}{(c.P_1 \cdot \mathbb{1}\min_i(b_i)) + ((P_2 - cP_1) \cdot \mathbb{1}\min_i(b_i))}
\end{aligned}$$

The justification for the last step is that we assumed $\frac{\max_i(a_i)}{\min_i(b_i)} > \exp\left(2\sum_{k=s-n}^{s} var_k(\psi_1)\right)$, and 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

$$\begin{aligned}
\frac{u_{j,\underline{w},s+2}(\underline{z})}{u_{j',\underline{v},s+2}(\underline{z})} &\leq \frac{c.\exp\left(2\sum_{k=s-n}^{s} var_k(\psi_1)\right) + (1-c).\frac{\max_i(a_i)}{\min_i(b_i)}}{c + (1-c)} \\
&= c.\exp\left(2\sum_{k=s-n}^{s} var_k(\psi_1)\right) + (1-c)\max_{j,j',w,v}\left(\frac{u_{j,w,n}(\underline{z})}{u_{j',v,n}(\underline{z})}\right)
\end{aligned}$$

as required.

References

[1] R. Bowen: *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms.* Springer Lecture Notes in Mathematics 470, 1975

[2] M. Boyle and K.Petersen: *Hidden Markov processes in the context of symbolic dynamics.* Preprint

[3] J.Chazottes and E.Ugalde: *Projections of Markov Measures may be Gibbsian*. Journal of Statistical Physics 111 (5/6) (2003) 1245-1272

[4] J.Chazottes and E.Ugalde: *Preservation of Gibbsianness under Amalgamation of Symbols.* Preprint

[5] T.Kempton and M.Pollicott: *Factors of Gibbs Measures for Full Shifts.* Preprint

[6] D. Ruelle: *Thermodynamic Formalism.* Addison-Wesley, New York, 1978

[7] E. Verbitsky: *Thermodynamics of Hidden Markov Processes.* Preprint

[8] P.Walters: *An Introduction to Ergodic Theory.* Springer Graduate Texts in Mathematics, 1982.