On the Geometric Ergodicity of Two-variable Gibbs Samplers

Aixin Tan
Department of Statistics
and Actuarial Science
University of Iowa
aixin-tan@uiowa.edu

Galin L. Jones *
School of Statistics
University of Minnesota
galin@umn.edu

James P. Hobert †
Department of Statistics
University of Florida
jhobert@stat.ufl.edu

January 4, 2012

Abstract

A Markov chain is geometrically ergodic if it converges to its invariant distribution at a geometric rate in total variation norm. We study geometric ergodicity of deterministic and random scan versions of the two-variable Gibbs sampler. We give a sufficient condition which simultaneously guarantees both versions are geometrically ergodic. We also develop a method for simultaneously establishing that both versions are subgeometrically ergodic. These general results allow us to characterize the convergence rate of two-variable Gibbs samplers in a particular family of discrete bivariate distributions.

*Research supported by the National Science Foundation and the National Institutes for Health.

†Research supported by the National Science Foundation.

1 Introduction

Let ϖ be a probability distribution having support $\mathsf{X} \times \mathsf{Y} \subseteq \mathbb{R}^k \times \mathbb{R}^l$, $k, l \geq 1$ and $\varpi_{X|Y}$ and $\varpi_{Y|X}$ denote the associated conditional distributions. We assume it is possible to simulate directly from $\varpi_{X|Y}$ and $\varpi_{Y|X}$. Then there are two Markov chains having ϖ as their invariant distribution, each of which could be called a *two-variable Gibbs sampler* (TGS). The most common version of a TGS is the deterministic scan Gibbs sampler (DGS), which is now described. Suppose the current state of the chain is $(X_n, Y_n) = (x, y)$, then the next state, (X_{n+1}, Y_{n+1}), is obtained as follows.

Iteration $n + 1$ of DGS:

1. Draw $X_{n+1} \sim \varpi_{X|Y}(\cdot|y)$, and call the observed value x'.

2. Draw $Y_{n+1} \sim \varpi_{Y|X}(\cdot|x')$.

An alternative TGS is the random scan Gibbs sampler (RGS). Fix $p \in (0, 1)$ and suppose the current state of the RGS chain is $(X_n, Y_n) = (x, y)$. Then the next state, (X_{n+1}, Y_{n+1}), is obtained as follows.

Iteration $n + 1$ of RGS:

1. Draw $B \sim \text{Bernoulli}(p)$.

2. If $B = 1$, then draw $X_{n+1} \sim \varpi_{X|Y}(\cdot|y)$ and set $Y_{n+1} = y$.

3. If $B = 0$, then draw $Y_{n+1} \sim \varpi_{Y|X}(\cdot|x)$ and set $X_{n+1} = x$.

Despite the simple structure of either TGS, these algorithms are widely applicable in the posterior analysis of complex Bayesian models. A TGS also arises naturally when ϖ is created via data augmentation techniques (Hobert, 2011; Tanner and Wong, 1987).

Inference based on ϖ often requires calculation of an intractable expectation. Let $g : \mathsf{X} \times \mathsf{Y} \to \mathbb{R}$ and let $E_\varpi g$ denote the expectation of g with respect to ϖ. If a TGS Markov chain is ergodic (see Tierney, 1994) and $E_{\varpi}|g| < \infty$, then

$$\bar{g}_n := \frac{1}{n}\sum_{i=0}^{n-1} g(X_i, Y_i) \xrightarrow{a.s.} E_{\varpi} g \qquad \text{as } n \to \infty.$$

Thus estimation of $E_{\varpi} g$ is simple. However, the estimator $\bar{g}_n$ will be more valuable if we can attach an estimate of the unknown Monte Carlo error $\bar{g}_n - E_{\varpi} g$. An approximation to the sampling distribution of the Monte Carlo error is available when a Markov chain central limit theorem (CLT) holds

$$\sqrt{n}(\bar{g}_n - E_{\varpi} g) \xrightarrow{d} \mathrm{N}(0, \sigma_g^2) \qquad \text{as } n \to \infty$$

with $0 < \sigma_g^2 < \infty$. The variance σ_g^2 accounts for the serial dependence in a TGS Markov chain and consistent estimation of it requires specialized techniques such as batch means, spectral methods or regenerative simulation. Let $\hat{\sigma}_n^2$ be an estimator of σ_g^2. If, with probability 1, $\hat{\sigma}_n^2 \to \sigma_g^2$ as $n \to \infty$, then an asymptotically valid Monte Carlo standard error is $\hat{\sigma}_n / \sqrt{n}$. These tools allow the practitioner to use the results of a TGS simulation with the same level of confidence that one would have if the observations were a random sample from ϖ. For more on this approach the interested reader can consult Geyer (1992), Geyer (2011), Flegal et al. (2008), Flegal and Jones (2010), Flegal and Jones (2011), Hobert et al. (2002), Jones et al. (2006), and Jones and Hobert (2001).

The CLT will obtain if $E_{\varpi}|g|^{2+\epsilon} < \infty$ for some $\epsilon > 0$ and the Markov chain is rapidly mixing (Chan and Geyer, 1994). In particular, we require that the Markov chain be geometrically ergodic; that is, converge to the target ϖ in total variation norm at a geometric rate. Under these same conditions methods such as batch means and regenerative simulation provide strongly consistent estimators of σ_g^2. Thus establishing geometric ergodicity is a key step in ensuring the reliability of a TGS as a method for estimating features of ϖ.

The convergence rate of DGS Markov chains has received substantial attention. In particular, sufficient conditions for geometric ergodicity have been developed for several DGS chains for practically relevant statistical models (see e.g. Hobert and Geyer, 1998; Johnson and Jones, 2010; Jones and Hobert, 2004; Marchev and Hobert, 2004; Roberts and Polson, 1994; Roberts and Rosenthal, 1999; Román and Hobert, 2011; Román, 2012; Rosenthal, 1996; Roy and Hobert, 2007; Tan and Hobert, 2009). The convergence rates of RGS chains has received almost no attention despite sometimes being useful.

Liu et al. (1995) did investigate geometric convergence of RGS chains, but the required regularity conditions are daunting and, to our knowledge, have not been applied to practically relevant statistical models. Recently Johnson et al. (2011) gave conditions which simultaneously establish geometric ergodicity of both the DGS chain and the corresponding RGS chain. These authors also conjectured that if the RGS chain is geometrically ergodic, then so is the DGS chain. That is to say, the qualitative convergence properties of TGS chains coincide. We are not able to resolve this conjecture in general, but in our main application (see Section 5) this is indeed the case.

A TGS chain which converges subgeometrically (ie, slower than geometric) would not be as useful as another chain which is geometrically ergodic–although with additional moment conditions it is still possible for a CLT to hold (Jones, 2004). Thus it would be useful to have criteria to check for subgeometric convergence. We are unaware of any previous work investigating subgeometric convergence of TGS Markov chains.

In the rest of this paper, we extend the results of Johnson et al. (2011) and provide a condition which can be used to simultaneously establish geometric ergodicity of DGS and RGS Markov chains. We then turn our attention to development of a condition which ensures that both the DGS and RGS chains converge subgeometrically. Finally, we apply our results to a class of bivariate distributions where we are able to characterize the convergence properties of the DGS and RGS chains. But we begin with some Markov chain background and a formal definition of the Markov chains we study.

2 Background and Notation

Let Z be a topological space and $\mathcal{B}(\mathsf{Z})$ denote its Borel σ-algebra. Also, let $\Phi = \{Z_0, Z_1, Z_2, \ldots\}$ be a Markov chain having Markov transition kernel P. That is, $P : \mathsf{Z} \times \mathcal{B}(\mathsf{Z}) \to [0, 1]$ such that for each $A \in \mathcal{B}(\mathsf{Z})$, $P(\cdot, A)$ is a nonnegative measurable function and for each $z \in \mathsf{Z}$, $P(z, \cdot)$ is a probability measure. As usual, P acts to the left on measures so that if ν is a measure on $(\mathsf{Z}, \mathcal{B}(\mathsf{Z}))$ and $A \in \mathcal{B}(\mathsf{Z})$, then

$$\nu P(A) = \int_{\mathsf{Z}} \nu(dz) P(z, A) \; .$$

For any $n \in \mathbb{Z}^+$, the n-step Markov transition kernel is given by $P^n(z, A) = \Pr(Z_{n+j} \in A | Z_j = z)$.

Let w be an invariant probability measure for P, that is, $wP = w$. Denote total variation norm by $\| \cdot \|_{TV}$. If Φ is ergodic, then for all $z \in \mathsf{Z}$ we have $||P^n(z,\cdot) - w(\cdot)||_{TV} \to 0$ as $n \to \infty$. Our goal is to study the rate of this convergence. Suppose there exist a real-valued function $M(z)$ on Z and $0 < t < 1$ such that for all z

$$||P^n(z,\cdot) - w(\cdot)||_{TV} \leq M(z)t^n \ . \tag{1}$$

Then Φ is *geometrically ergodic*, otherwise it is *subgeometrically ergodic*.

2.1 Two-variable Gibbs samplers

In this section we define the Markov kernels associated with the DGS and RGS chains described in Section 1. We also introduce a third Markov chain which will prove crucial to our study of the other Markov chains.

Recall that ϖ is a probability distribution having support $\mathsf{X}\times\mathsf{Y} \subseteq \mathbb{R}^k \times \mathbb{R}^l$, $k, l \geq 1$. Let $\pi(x, y)$ be a density of ϖ with respect to a measure $\mu = \mu_X \times \mu_Y$. Then the marginal densities are given by

$$\pi_X(x) = \int_{\mathsf{Y}} \pi(x,y)\mu_Y(dy)$$

and similarly for $\pi_Y(y)$. The conditional densities are $\pi_{X|Y}(x|y) = \pi(x,y)/\pi_Y(y)$ and $\pi_{Y|X}(y|x) = \pi(x,y)/\pi_X(x)$.

Consider the DGS Markov chain $\Phi_{DGS} = \{(X_0, Y_0), (X_1, Y_1), \ldots\}$ and let

$$k_{DGS}(x', y'|x, y) = \pi_{X|Y}(x'|y)\pi_{Y|X}(y'|x') \ .$$

Then the Markov kernel for Φ_{DGS} is defined by

$$P_{DGS}((x,y), A) = \int_A k_{DGS}(x', y'|x, y)\mu(d(x', y')) \qquad A \in \mathcal{B}(\mathsf{X}) \times \mathcal{B}(\mathsf{Y}) \ .$$

It is well known that the two marginal sequences comprising Φ_{DGS} are themselves Markov chains (Liu et al., 1994). We now consider the marginal sequence $\Phi_X = \{X_0, X_1, \ldots\}$ and define

$$k_X(x'|x) = \int_{\mathsf{Y}} \pi_{X|Y}(x'|y)\pi_{Y|X}(y|x)\mu_Y(dy) \ .$$

The Markov kernel for Φ_X is

$$P_X(x, A) = \int_A k_X(x'|x)\mu_X(dx') \qquad A \in \mathcal{B}(\mathsf{X}) .$$

Note that P_{DGS} has ϖ as its invariant distribution while P_X has the marginal ϖ_X as its invariant distribution.

Finally, consider the RGS Markov chain $\Phi_{RGS} = \{(X_0, Y_0), (X_1, Y_1), \ldots\}$. Let $p \in (0, 1)$ and δ denote Dirac's delta. Define

$$k_{RGS}(x', y'|x, y) = p\pi_{X|Y}(x'|y)\delta(y' - y) + (1 - p)\pi_{Y|X}(y'|x)\delta(x' - x) .$$

Then the Markov kernel for Φ_{RGS} is

$$P_{RGS}((x, y), A) = \int_A k_{RGS}(x', y'|x, y)\mu(d(x', y'))$$

It is easy to show via direct computation that ϖ is invariant for P_{RGS}.

It is well known that P_X and P_{DGS} converge to their respective invariant distributions at the same rate (Diaconis et al., 2008; Liu et al., 1994; Robert, 1995; Roberts and Rosenthal, 2001). Thus if one is geometrically ergodic, then so is the other. This relationship has been routinely exploited in the study of TGS chains for practically relevant statistical models (cf. Hobert and Geyer, 1998; Johnson and Jones, 2010; Jones and Hobert, 2004; Roy and Hobert, 2007; Tan and Hobert, 2009) since one of the two chains may be easier to analyze than the other. Recently, Johnson et al. (2011) showed that if P_X or P_{DGS} is geometrically ergodic, then so is P_{RGS}. Thus the analysis of the convergence rate of TGS algorithms often comes down to analyzing P_X. This is exactly the approach we take in Sections 3 and 5.

3 Conditions for Geometric Ergodicity

In this section we develop general conditions which ensure that P_X, P_{DGS} and P_{RGS} are geometrically ergodic. First we need a couple of concepts from Markov chain theory. Recall the notation from Section 2. That is, P is a Markov kernel on $(\mathsf{Z}, \mathcal{B}(\mathsf{Z}))$. Then P is *Feller* if for any open set $O \in \mathcal{B}(\mathsf{Z})$, $P(\cdot, O)$ is a lower semicontinuous function. The Markov kernel P acts to the right on functions so that for measurable f

$$Pf(z) = \int_{\mathsf{Z}} f(z')P(z, dz') .$$

A *drift condition* holds if there exists a function $U : \mathsf{Z} \to \mathbb{R}^+$, and constants $0 < \lambda < 1$ and $L < \infty$ satisfying

$$PU(z) \leq \lambda U(z) + L \qquad \text{for all } z \in \mathsf{Z}\,. \tag{2}$$

Recall that a function U is said to be unbounded off compact sets if the sublevel set $\{z \in \mathsf{Z} : U(z) \leq d\}$ is compact for every $d > 0$. If P is Feller, U is unbounded off compact sets and satisfies (2), then Φ is geometrically ergodic. See Meyn and Tweedie (1993) and Roberts and Rosenthal (2004) for details while Jones and Hobert (2001) give an introduction to the use of drift conditions.

3.1 Two-variable Gibbs samplers

Johnson et al. (2011) gave a set of conditions which simultaneously prove that Φ_X, Φ_{DGS} and Φ_{RGS} are geometrically ergodic. We build on their work and show how a drift condition for P_X naturally provides a drift condition for P_{RGS}. This allows us to develop an alternative set of conditions which are sufficient for the geometric ergodicity of P_X, P_{DGS} and P_{RGS}. The application of this method is illustrated in Section 5.

The following result was essentially proved by Johnson et al. (2011), but it was not stated in their paper. Thus we provide a proof for the sake of completeness. First we set some notation. Suppose $V : \mathsf{X} \to \mathbb{R}^+$ and let

$$G(y) = \int_{\mathsf{X}} V(x) \pi_{X|Y}(x|y) \mu_X(dx)\ .$$

Also, for $c > 0$ define

$$W(x, y) = V(x) + cG(y)\ . \tag{3}$$

Lemma 1. *Suppose there exist constants* $0 < \lambda < 1$ *and* $L < \infty$ *such that for all* $x \in \mathsf{X}$

$$P_X V(x) \leq \lambda V(x) + L\ .$$

If $0 < p < 1$ *and* $p(1-p)^{-1} < c < p[\lambda(1-p)]^{-1}$*, then there exists* $\lambda < \gamma < 1$ *such that*

$$P_{RGS} W(x, y) \leq \gamma W(x, y) + (1-p)cL\ .$$

Proof. Notice that

$$
\begin{aligned}
\int_{\mathsf{Y}} G(y)\pi_{Y|X}(y|x)\mu_Y(dy) &= \int_{\mathsf{Y}}\int_{\mathsf{X}} V(x')\pi_{X|Y}(x'|y)\pi_{Y|X}(y|x)\mu_X(dx')\mu_Y(dy) \\
&= \int_{\mathsf{X}} V(x') \int_{\mathsf{Y}} \pi_{X|Y}(x'|y)\pi_{Y|X}(y|x)\mu_Y(dy)\mu_X(dx') \\
&= \int_{\mathsf{X}} V(x')k_X(x'|x)\mu_X(dx') \\
&\leq \lambda V(x) + L
\end{aligned}
$$

Since

$$
\frac{p}{1-p} < c < \frac{p}{\lambda(1-p)} \tag{4}
$$

there exists γ such that

$$
(1-p)(c\lambda+1) \vee \frac{p(1+c)}{c} \leq \gamma < 1\,. \tag{5}
$$

$$
\begin{aligned}
P_{RGS}W(x,y) &= \int_{\mathsf{X}}\int_{\mathsf{Y}} W(x',y')k_{RGS}(x',y'|x,y)\mu_X(dx')\mu_Y(dy') \\
&= p\int_{\mathsf{X}}\int_{\mathsf{Y}} W(x',y')\pi_{X|Y}(x'|y)\delta(y'-y)\mu_X(dx')\mu_Y(dy') \\
&\quad + (1-p)\int_{\mathsf{X}}\int_{\mathsf{Y}} W(x',y')\pi_{Y|X}(y'|x)\delta(x'-x)\mu_X(dx')\mu_Y(dy') \\
&= p\int_{\mathsf{X}} W(x',y)\pi_{X|Y}(x'|y)\mu_X(dx') + (1-p)\int_{\mathsf{Y}} W(x,y')\pi_{Y|X}(y'|x)\mu_Y(dy') \\
&= p\int_{\mathsf{X}} \left[V(x') + cG(y)\right]\pi_{X|Y}(x'|y)\mu_X(dx') \\
&\quad + (1-p)\int_{\mathsf{Y}} \left[V(x) + cG(y')\right]\pi_{Y|X}(y'|x)\mu_Y(dy') \\
&= pcG(y) + (1-p)V(x) + pG(y) + (1-p)c\int_{\mathsf{Y}} G(y')\pi_{Y|X}(y'|x)\mu_Y(dy') \\
&= p(1+c)G(y) + (1-p)V(x) + (1-p)c\int_{\mathsf{Y}} G(y')\pi_{Y|X}(y'|x)\mu_Y(dy') \\
&\leq (1-p)c\lambda V(x) + (1-p)cL + p(1+c)G(y) + (1-p)V(x) \\
&= (1-p)(c\lambda+1)V(x) + p(1+c)G(y) + (1-p)cL \\
&\leq \gamma W(x,y) + (1-p)cL
\end{aligned}
$$

All that remains is to show that $\gamma > \lambda$. Now

$$\begin{aligned}
\gamma &\geq (1-p)(c\lambda + 1) && \text{by (5)} \\
&> (1-p)\left(\frac{p}{1-p}\lambda + 1\right) && \text{by (4)} \\
&= p\lambda + (1-p) \\
&> \lambda && \text{since } \lambda,\, p \in (0,1)\ .
\end{aligned}$$

□

The following is an easy consequence of Lemma 1 and the material stated at the beginning of this section.

Proposition 1. *Suppose P_X and P_{RGS} are Feller. If there exists a function $V : \mathsf{X} \to \mathbb{R}^+$ such that both V and the corresponding W (as defined at (3)) are unbounded off compact sets, and there exist constants $0 < \lambda < 1$ and $L < \infty$ such that for all $x \in \mathsf{X}$*

$$P_X V(x) \leq \lambda V(x) + L,$$

then Φ_X, Φ_{DGS} and Φ_{RGS} are geometrically ergodic.

4 Conditions for Subgeometric Convergence

Our goal in this section is to develop a condition which ensures that Φ_X, Φ_{DGS} and Φ_{RGS} converge subgeometrically, but first we need a few concepts from general Markov chain theory. Recall the notation of Section 2. In particular, P is a Markov kernel on $(\mathsf{Z}, \mathcal{B}(\mathsf{Z}))$ having invariant distribution w. A Markov kernel defines an operator on the space of measurable functions that are square integrable with respect to the invariant distribution, denoted $L^2(w)$. Also, let

$$L^2_{0,1}(w) = \left\{ f \in L^2(w) \,:\, E_w f = 0, \text{ and } E_w f^2 = 1 \right\}\ .$$

For $f, g \in L^2(w)$, define the inner product as

$$\langle f, g \rangle = \int_{\mathsf{Z}} f(z) g(z) w(dz)$$

and $\|f\|^2 = \langle f, f \rangle$. The norm of the operator P is

$$\|P\| = \sup_{f \in L^2_{0,1}(w)} \|Pf\| \ .$$

If P is symmetric with respect to w, that is, if

$$P(z, dz')w(dz) = P(z', dz)w(dz'), \tag{6}$$

then P is self-adjoint so that $\langle Ph_1, h_2 \rangle = \langle h_1, Ph_2 \rangle$. If P is w-symmetric, then Φ is geometrically ergodic if and only if $\|P\| < 1$ (Roberts and Rosenthal, 1997). Moreover, if $Z \sim w$ and $Z'|Z = z \sim P(z, \cdot)$, then

$$\|P\| = \sup_{f \in L^2_{0,1}(w)} |\langle Pf, f \rangle| = \sup_{f \in L^2_{0,1}(w)} |E\left[f(Z')f(Z)\right]| \ . \tag{7}$$

The first equality is a property of self-adjoint operators while the second equality follows directly from the definition of inner product.

4.1 Two-variable Gibbs samplers

It is easy to see that P_X is ϖ_X-symmetric and P_{RGS} is ϖ-symmetric, but P_{DGS} is not ϖ-symmetric. Because P_X and P_{RGS} are symmetric, the operator theory described above applies. In particular, if $X \sim \varpi_X$ and $X'|X = x \sim P_X(x, \cdot)$, then

$$\|P_X\| = \sup_{f \in L^2_{0,1}(\varpi_X)} |Ef(X')f(X)|$$

while if $(X, Y) \sim \varpi$ and $(X', Y')|(X, Y) = (x, y) \sim P_{RGS}((x, y), \cdot)$, then

$$\|P_{RGS}\| = \sup_{f \in L^2_{0,1}(\varpi)} |Ef(X', Y')f(X, Y)| \ .$$

Note that despite our use of $\|\cdot\|$ for both operator norms, these are different since they are based on different L^2 spaces.

If we can show that $\|P_X\| = \|P_{RGS}\| = 1$, then we will be able to conclude that Φ_X, Φ_{DGS}, and Φ_{RGS} are subgeometrically ergodic. First, we need convenient characterizations of the operator norms.

Lemma 2. *If* $(X,Y) \sim \varpi$, *then*

$$\|P_X\| = 1 - \inf_{f \in L^2_{0,1}(\varpi_X)} E(\mathit{Var}(f(X)|Y))$$

and

$$\|P_{RGS}\| = 1 - \inf_{f \in L^2_{0,1}(\varpi)} \{pE(\mathit{Var}(f(X,Y)|Y)) + (1-p)E(\mathit{Var}(f(X,Y)|X))\} \ .$$

Proof. Suppose $X \sim \varpi_X$, $X'|X = x \sim P_X(x,\cdot)$ and $(X,Y) \sim \varpi$. Then

$$\begin{aligned} \|P_X\| &= \sup_{f \in L^2_{0,1}(\varpi_X)} |Ef(X')f(X)| \\ &= \sup_{f \in L^2_{0,1}(\varpi_X)} \mathrm{Var}(E(f(X)|Y)) \\ &= 1 - \inf_{f \in L^2_{0,1}(\varpi_X)} E(\mathrm{Var}(f(X)|Y)) \ . \end{aligned}$$

In the above, the second equality follows from Liu et al. (1994, Lemma 3.2) and the last equality holds since for $f \in L^2_{0,1}(\varpi_X)$

$$1 = E(\mathrm{Var}(f(X)|Y)) + \mathrm{Var}(E(f(X)|Y)) \, .$$

Now consider $\|P_{RGS}\|$. Suppose $(X,Y) \sim \varpi$ and $(X',Y')|(X,Y) = (x,y) \sim P_{RGS}((x,y),\cdot)$. Then

$$\begin{aligned} &E\left[h(X',Y')h(X,Y)\right] \\ &= \int h(x',y')h(x,y)k_{RGS}(x',y'|x,y)\pi(x,y)\mu_X(dx')\mu_Y(dy')\mu_X(dx)\mu_Y(dy) \\ &= \int h(x',y')h(x,y)\pi(x,y)[p\pi_{X|Y}(x'|y)\delta(y'-y) \\ &\qquad + (1-p)\pi_{Y|X}(y'|x)\delta(x'-x)]\mu_X(dx')\mu_Y(dy')\mu_X(dx)\mu_Y(dy) \\ &= \int ph(x',y)h(x,y)\pi_{X|Y}(x'|y)\pi(x,y)\mu_X(dx')\mu_X(dx)\mu_Y(dy) \\ &\qquad + \int (1-p)h(x,y')h(x,y)\pi_{Y|X}(y'|x)\pi(x,y)\mu_Y(dy')\mu_X(dx)\mu_Y(dy) \\ &= \int ph(x,y)E[h(X',Y)|Y=y]\pi(x,y)\mu_X(dx)\mu_Y(dy) \\ &\qquad + \int (1-p)h(x,y)E[h(X,Y')|X=x]\pi(x,y)\mu_X(dx)\mu_Y(dy) \end{aligned}$$

$$
\begin{aligned}
&= \int ph(x,y)E[h(X',Y)|Y=y]\pi_{X|Y}(x|y)\pi_Y(y)\mu_X(dx)\mu_Y(dy) \\
&\qquad + \int (1-p)h(x,y)E[h(X,Y')|X=x]\pi_{Y|X}(y|x)\pi_X(x)\mu_X(dx)\mu_Y(dy) \\
&= \int pE[h(X,Y)|Y=y]E[h(X',Y)|Y=y]\pi_Y(y)\mu_Y(dy) \\
&\qquad + \int (1-p)E[h(X,Y)|X=x]E[h(X,Y')|X=x]\pi_X(x)\mu_X(dx) \\
&= \int p(E[h(X,Y)|Y=y])^2\pi_Y(y)\mu_Y(dy) \\
&\qquad + \int (1-p)(E[h(X,Y)|X=x])^2\pi_X(x)\mu_X(dx) \\
&= pE\left[(E\left[h(X,Y)|Y\right])^2\right] + (1-p)E\left[(E\left[h(X,Y)|X\right])^2\right] .
\end{aligned}
$$

Now since $h \in L^2_{0,1}(\varpi)$,

$$\mathrm{Var}(E[h(X,Y)|Y]) = E[(E\left[h(X,Y)|Y\right])^2]$$

and

$$\mathrm{Var}(E[h(X,Y)|X]) = E[(E\left[h(X,Y)|X\right])^2] .$$

Moreover,

$$1 = \mathrm{Var}_\varpi[h(X,Y)] = \mathrm{Var}(E[h(X,Y)|Y]) + E(\mathrm{Var}[h(X,Y)|Y])$$

and

$$1 = \mathrm{Var}_\varpi[h(X,Y)] = \mathrm{Var}(E[h(X,Y)|X]) + E(\mathrm{Var}[h(X,Y)|X]) .$$

The result follows easily. □

Proposition 2. *Suppose there exists a sequence $\{h_i \in L^2_{0,1}(\varpi_X)\}$ such that if $(X,Y) \sim \varpi$, then*

$$\liminf_{i\to\infty} E[Var(h_i(X)|Y)] = 0 . \tag{8}$$

Then $\|P_X\| = \|P_{RGS}\| = 1$. Hence Φ_X, Φ_{RGS} and Φ_{DGS} are subgeometrically ergodic.

Proof. The claim that $\|P_X\| = 1$ follows easily from the first part of Lemma 2. Now consider $\|P_{RGS}\|$. Note that if $f'(x,y) := f(x) \in L^2_{0,1}(\varpi_X)$, then $f' \in L^2_{0,1}(\varpi)$. From the second part of Lemma 2 we have

$$\|P_{RGS}\| = 1 - \inf_{f \in L^2_{0,1}(\varpi)} \{pE(\text{Var}[f(X,Y)|Y]) + (1-p)E(\text{Var}[f(X,Y)|X])\} \ .$$

The claim now follows easily since if $f(x,y) = h_i(x)$, then

$$E(\text{Var}[f(X,Y)|X]) = E(\text{Var}[h_i(X)|X]) = 0$$

and

$$E(\text{Var}[f(X,Y)|Y]) = E(\text{Var}[h_i(X)|Y]) \ .$$

Thus we conclude that Φ_X and Φ_{RGS} are subgeometrically ergodic. Since Φ_X and Φ_{DGS} are either both geometrically ergodic or both subgeometric, it follows that Φ_{DGS} also converges subgeometrically. $\square$

5 A Discrete Example

We introduce a family of simple discrete distributions which admit usage of the TGS algorithms. We then apply our general results which will allow us to very nearly characterize the members of the family which admit geometrically ergodic TGS Markov chains.

Let $\{a_i\}_{i=1}^{\infty}$ and $\{b_i\}_{i=1}^{\infty}$ be strictly positive sequences satisfying

$$\sum_{i=1}^{\infty} a_i + \sum_{i=1}^{\infty} b_i = 1 \ .$$

Also, let $b_0 = 0$. Let the family consist of the discrete bivariate distributions having density π with respect to counting measure on $\mathbb{N} \times \mathbb{N}$ given by

$$\pi(x,y) = \begin{cases} a_x & x = y, \ y = 1,2,3,\dots \ ; \\ b_y & x = y+1, \ y = 1,2,3,\dots \ ; \\ 0 & \text{otherwise} \, . \end{cases}$$

Hence the marginals are given by

$$\pi_X(x) = \sum_{y=1}^{\infty} \pi(x,y) = \sum_{y=1}^{\infty} a_x I(x=y) + b_y I(y=x-1) = a_x + b_{x-1}$$

and

$$\pi_Y(y) = \sum_{x=1}^{\infty} \pi(x,y) = \sum_{x=1}^{\infty} a_x I(x=y) + b_y I(y=x+1) = a_y + b_y \ .$$

Then the full conditionals are easily seen to be

$$\pi_{X|Y}(x|y) = \frac{a_y}{a_y+b_y} I(x=y) + \frac{b_y}{a_y+b_y} I(x=y+1) \qquad y=1,2,3,\ldots$$

and

$$\pi_{Y|X}(y|x) = \frac{a_x}{a_x+b_{x-1}} I(x=y) + \frac{b_{x-1}}{a_x+b_{x-1}} I(y=x-1) \qquad x=1,2,3,\ldots \ .$$

Define

$$p_x = \frac{a_x b_x}{(a_x+b_{x-1})(a_x+b_x)} \qquad \text{and} \qquad q_x = \frac{a_{x-1}b_{x-1}}{(a_x+b_{x-1})(a_{x-1}+b_{x-1})} \ .$$

Then for the DGS

$$k_{DGS}(x',y'|x,y) = \pi_{X|Y}(x'|y)\pi_{Y|X}(y'|x')$$

and hence for the marginal chain Φ_X

$$k_X(x'|x) = \sum_{y=1}^{\infty} \pi_{X|Y}(x'|y)\pi_{Y|X}(y|x) = \begin{cases} 1-p_1 & x'=x=1; \\ p_x & x'=x+1,\ x\geq 1; \\ q_x & x'=x-1,\ x\geq 2; \\ 1-p_x-q_x & x'=x,\ x\geq 2; \quad \text{and} \\ 0 & \text{otherwise} \ . \end{cases}$$

It is easy to see that the kernel P_X is Feller. Now let $p \in (0,1)$ and let δ denote the Dirac delta function. For the random scan Gibbs sampler (RGS) we have

$$k_{RGS}(x',y'|x,y) = p\pi_{X|Y}(x'|y)\delta(y'-y) + (1-p)\pi_{Y|X}(y'|x)\delta(x'-x) \ .$$

Since for any open set O

$$P_{RGS}((x,y),O) = p\sum_{x'=1}^{\infty} \pi_{X|Y}(x'|y) I((x',y)\in O) + (1-p)\sum_{y'=1}^{\infty} \pi_{Y|X}(y'|x) I((x,y')\in O)$$

it is easy to see that $P_{RGS}(\cdot, O)$ is lower semicontinuous and hence Φ_{RGS} is Feller.

We are now in position to establish sufficient conditions for the geometric ergodicity of Φ_X, Φ_{DGS} and Φ_{RGS}.

Lemma 3. *If*

$$\limsup_{x\to\infty} \frac{p_x}{q_x} < 1 \qquad \textit{and} \qquad \liminf_{x\to\infty} q_x > 0, \tag{9}$$

then Φ_X, Φ_{DGS} and Φ_{RGS} are geometrically ergodic.

Proof. We need only verify the conditions of Proposition 1 and we've already seen that both P_X and P_{RGS} are Feller. Set $V(x) = z^x$ for some $z > 1$ which will be determined later and note that

$$G(y) = \sum_{x=1}^{\infty} V(x)\pi_{X|Y}(x|y) = \left(\frac{a_y + zb_y}{a_y + b_y}\right) z^y \,.$$

For any $d > 0$, the sublevel set $A_d := \{x : V(x) \le d\} = \{x : z^x \le d\}$ is bounded. Since V is a continuous function, A_d is also closed, hence compact. Therefore V is unbounded off compact sets on X. On the other hand, for any $d > 0$, the sublevel set $B_d := \{y : G(y) \le d\} \subset \{y : z^y \le d\}$ is bounded. Then for any $b > 0$, $W(x, y) = V(x) + bG(y)$ is unbounded off compact sets on $\mathsf{X} \times \mathsf{Y}$ because for any $d > 0$, $\{(x, y) : W(x, y) \le d\} \subset A_d \times B_d$ is bounded and closed, hence compact. Now, all that remains is to construct a drift condition for V. Note that for $x \ge 2$,

$$\begin{aligned}
P_X V(x) &= \sum_{x'=1}^{\infty} z^{x'} k_X(x'|x) \\
&= p_x z^{x+1} + q_x z^{x-1} + (1 - p_x - q_x) z^x \\
&= \left[z p_x + \frac{q_x}{z} + 1 - p_x - q_x\right] z^x \\
&= \left[p_x(z-1) + q_x\left(\frac{1}{z} - 1\right) + 1\right] V(x) \,.
\end{aligned} \tag{10}$$

We next try to bound the coefficient of $V(x)$ in the right hand side of (10) for all large values of x. Set

$$r := \limsup_{x\to\infty} \frac{p_x}{q_x} \qquad \text{and} \qquad q := \liminf_{x\to\infty} q_x$$

and note that $r < 1$ and $q > 0$ by assumption. Then there exists $x_0 \geq 2$ such that

$$\frac{p_x}{q_x} < \frac{r+1}{2} \text{ and } q_x > \frac{q}{2} \qquad \text{for all } x > x_0 \, .$$

For any $z \in (1, 2/(r+1))$ and $x > x_0$,

$$\begin{aligned} p_x(z-1) + q_x\left(\frac{1}{z} - 1\right) + 1 &< \frac{r+1}{2} q_x(z-1) + \frac{q_x(1-z)}{z} + 1 \\ &= q_x(z-1)\left(\frac{r+1}{2} - \frac{1}{z}\right) + 1 \\ &< \frac{q}{2}(z-1)\left(\frac{r+1}{2} - \frac{1}{z}\right) + 1 \end{aligned}$$

since $z \in (1, 2/(r+1))$ implies

$$\frac{r+1}{2} - \frac{1}{z} < 0 \, .$$

Next note that

$$0 < q < 1 \, , \quad 0 < z - 1 < \frac{1-r}{1+r} < 1 \text{ and } -\frac{1}{2} < \frac{r+1}{2} - \frac{1}{z} < 0,$$

which guarantees

$$0 < \frac{q}{2}(z-1)\left(\frac{r+1}{2} - \frac{1}{z}\right) + 1 < 1 \, .$$

Thus there exists $0 < \rho < 1$ such that

$$\frac{q}{2}(z-1)\left(\frac{r+1}{2} - \frac{1}{z}\right) + 1 \leq \rho < 1 \, . \tag{11}$$

Finally, to bound $P_X V(x)$ for $x \leq x_0$, set

$$L := \max_{x \leq x_0} P_X V(x) \, . \tag{12}$$

Putting together equations (10) to (12), we have

$$P_X V(x) \leq \rho V(x) + L$$

with $0 < \rho < 1$ and $L < \infty$. The conclusion now follows from Proposition 1. $\square$

The above sufficient condition for geometric ergodicity involves transition probabilities of the chain Φ_X. Alternatively, we could state a sufficient condition in terms of the probabilities $\{a_i, b_i\}$ which define the density π.

Define

$$A := \limsup_{i\to\infty} \frac{a_i}{a_{i-1}} \ ; \quad m := \liminf_{i\to\infty} \frac{a_i}{b_i} \ ; \text{ and } \quad M := \limsup_{i\to\infty} \frac{a_i}{b_i} \ .$$

Corollary 1. *If*

$$\limsup_{i\to\infty} \frac{a_i}{b_{i-1}} < \infty, \qquad \qquad \limsup_{i\to\infty} \frac{b_i}{a_i} < \infty$$

and $A(1+M)/(1+m) < 1$, then Φ_X, Φ_{DGS}, and Φ_{RGS} are geometrically ergodic.

Proof. We verify the conditions of Lemma 3. Note that

$$q_i = \frac{b_{i-1}}{a_i + b_{i-1}} \frac{a_{i-1}}{a_{i-1} + b_{i-1}} = \frac{1}{1 + \frac{a_i}{b_{i-1}}} \frac{1}{1 + \frac{b_{i-1}}{a_{i-1}}} \ .$$

Hence

$$\liminf_{i\to\infty} q_i \geq \frac{1}{1 + \limsup_{i\to\infty} \frac{a_i}{b_{i-1}}} \frac{1}{1 + \limsup_{i\to\infty} \frac{b_{i-1}}{a_{i-1}}} > 0 \ .$$

Next observe that

$$\frac{p_i}{q_i} = \frac{a_i}{a_{i-1}} \frac{b_i}{b_{i-1}} \frac{a_{i-1} + b_{i-1}}{a_i + b_i} = \frac{a_i}{a_{i-1}} \frac{1 + \frac{a_{i-1}}{b_{i-1}}}{1 + \frac{a_i}{b_i}} \ .$$

Hence

$$\limsup_{i\to\infty} \frac{p_i}{q_i} \leq \left[\limsup_{i\to\infty} \frac{a_i}{a_{i-1}} \right] \left[\frac{1 + \limsup_{i\to\infty} \frac{a_{i-1}}{b_{i-1}}}{1 + \liminf_{i\to\infty} \frac{a_i}{b_i}} \right] = A \frac{1+M}{1+m} < 1 \ .$$

□

So far in this section, we have used Proposition 1 to get sufficient conditions for the geometric ergodicity of the Markov chains. Next, we use Proposition 2 to study the conditions under which the Markov chains are subgeometrically ergodic.

Lemma 4. *The Markov chains Φ_X, Φ_{DGS} and Φ_{RGS} are subgeometrically ergodic if any one of the following conditions hold:*

1. $$\limsup_{i\to\infty} \frac{\sum_{x=i}^{\infty}(a_x+b_x)}{a_{i-1}} = \infty\ ;$$

2. $$\limsup_{i\to\infty} \frac{\sum_{x=i}^{\infty}(a_x+b_x)}{b_{i-1}} = \infty\ ;\ \textit{or}$$

3. $$\limsup_{i\to\infty} \frac{b_i}{a_i} = \infty\ .$$

Proof. Let $(X,Y)\sim\varpi$. For $i=1,2,3,\ldots$ let $H_i(x)=I(x\geq i)$. Then

$$\mu_i := E[H_i(X)] = E[H_i^2(X)] = \sum_{x=i}^{\infty}(a_x+b_{x-1}) < \infty$$

and

$$v_i := \mathrm{Var}[H_i(X)] = \mu_i(1-\mu_i) < \infty\ .$$

Define $h_i(x) = [H_i(x)-\mu_i]/\sqrt{v_i}$ and note that $h_i \in L^2_{0,1}(\varpi_X)$. We will show that

$$\liminf_{i\to\infty} E[Var(h_i(X)|Y)] = 0\ ,$$

and appeal to Proposition 2 for the conclusion. Let

$$\beta_y = \frac{b_y}{a_y+b_y} = \pi_{X|Y}(y+1|y)\ .$$

Then

$$\begin{aligned}
E[H_i(X)|Y=y] &= E[H_i^2(X)|Y=y] \\
&= \pi_{X|Y}(y|y)H_i(y) + \pi_{X|Y}(y+1|y)H_i(y+1) \\
&= \begin{cases} 0 & y\leq i-2, \\ \beta_{i-1} & y=i-1, \\ 1 & y\geq i\ . \end{cases}
\end{aligned}$$

Hence

$$Var[H_i(X)|Y=y] = \begin{cases} \beta_{i-1}(1-\beta_{i-1}) & y = i-1, \\ 0 & \text{otherwise} . \end{cases}$$

Therefore,

$$\begin{aligned} E(Var[H_i(X)|Y=y]) &= \sum_{y=1}^{\infty} \pi_Y(y) Var[H_i(X)|Y=y] \\ &= \pi_Y(i-1) Var[H_i(X)|Y=i-1] \\ &= (a_{i-1}+b_{i-1})\beta_{i-1}(1-\beta_{i-1}) \\ &= \frac{a_{i-1}b_{i-1}}{a_{i-1}+b_{i-1}} \ . \end{aligned}$$

Finally,

$$E[Var(h_i(X)|Y)] = v_i^{-1} E[Var(H_i(X)|Y)] = [\mu_i(1-\mu_i)]^{-1} \frac{a_{i-1}b_{i-1}}{a_{i-1}+b_{i-1}} \ .$$

Note that

$$\begin{aligned} (E[Var(h_i(X)|Y)])^{-1} &= \mu_i(1-\mu_i) \frac{a_{i-1}+b_{i-1}}{a_{i-1}b_{i-1}} \\ &= (1-\mu_i) \left[\sum_{x=i}^{\infty} (a_x+b_x) + b_{i-1} \right] \left(\frac{1}{a_{i-1}} + \frac{1}{b_{i-1}} \right) \\ &= (1-\mu_i) \left[\frac{\sum_{x=i}^{\infty}(a_x+b_x)}{a_{i-1}} + \frac{\sum_{x=i}^{\infty}(a_x+b_x)}{b_{i-1}} + \frac{b_{i-1}}{a_{i-1}} + 1 \right] \end{aligned}$$

and that

$$\lim_{i\to\infty}(1-\mu_i) = \lim_{i\to\infty} \sum_{x=1}^{i-1} (a_x + b_{x-1}) = 1 \ .$$

Hence equation (8) holds if and only if

$$\begin{aligned} & \limsup_{i\to\infty} \frac{\sum_{x=i}^{\infty}(a_x+b_x)}{a_{i-1}} = \infty, \\ \text{or} \quad & \limsup_{i\to\infty} \frac{\sum_{x=i}^{\infty}(a_x+b_x)}{b_{i-1}} = \infty, \\ \text{or} \quad & \limsup_{i\to\infty} \frac{b_i}{a_i} = \infty \ . \end{aligned}$$

□

Finally, we can use the previous results to characterize the conditions for geometric ergodicity of TGS Markov chains for a large subfamily of our discrete distributions.

Corollary 2. *Assume that both* $A := \lim_{i\to\infty} \frac{a_i}{a_{i-1}}$ *and* $\lim_{i\to\infty} \frac{a_i}{b_i}$ *exist. Then all the limits below are well defined and the following statements are equivalent:*

(a)

$$\lim_{i\to\infty} \frac{a_i}{b_{i-1}} < \infty, \quad \lim_{i\to\infty} \frac{b_i}{a_i} < \infty, \quad \textit{and} \quad A < 1\,.$$

(b)

$$r = \lim_{i\to\infty} \frac{p_i}{q_i} < 1 \quad \textit{and} \quad q = \lim_{i\to\infty} q_i > 0\,.$$

(c) Φ_X *is geometrically ergodic.*

(d) Φ_{DGS} *is geometrically ergodic.*

(e) Φ_{RGS} *is geometrically ergodic.*

Proof. As we noted in Section 2.1, the equivalence of (c) and (d) is well known.

<u>(a) ⇒ (b)</u>: Note that

$$q = \lim_{i\to\infty} q_i = \frac{1}{1 + \lim_{i\to\infty} \frac{a_i}{b_{i-1}}} \frac{1}{1 + \lim_{i\to\infty} \frac{b_{i-1}}{a_{i-1}}} > 0$$

and

$$r = \lim_{i\to\infty} \frac{p_i}{q_i} = \left[\lim_{i\to\infty} \frac{a_i}{a_{i-1}}\right] \left[\frac{1 + \lim_{i\to\infty} \frac{a_{i-1}}{b_{i-1}}}{1 + \lim_{i\to\infty} \frac{a_i}{b_i}}\right] = A < 1\ .$$

<u>(b) ⇒ (c) and (b) ⇒ (e)</u>: The same argument holds for Φ_X and Φ_{RGS}. Immediate by Lemma 3.

<u>(c) ⇒ (a) and (e) ⇒ (a)</u>: The same argument holds for Φ_X and Φ_{RGS}. If the chain is geometrically ergodic, then $\lim_{i\to\infty} \frac{a_i}{b_{i-1}} < \infty$ and $\lim_{i\to\infty} \frac{b_i}{a_i} < \infty$ by conditions 2 and 3 of Lemma 4. Next, if $A = 1$, then for any fixed positive integer K, we have

$$\lim_{i\to\infty} \frac{a_{i+1}}{a_i} = 1,\ \lim_{i\to\infty} \frac{a_{i+2}}{a_i} = 1, \ldots, \lim_{i\to\infty} \frac{a_{i+K}}{a_i} = 1\ .$$

Then there exists i_0 such that for any $i \geq i_0$,

$$\frac{a_{i+1}}{a_i} > \frac{1}{2}, \ \frac{a_{i+2}}{a_i} > \frac{1}{2}, \ldots, \frac{a_{i+K}}{a_i} > \frac{1}{2} \ .$$

Hence, given any K, there exists i_0 such that for any $i > i_0$,

$$\frac{\sum_{x=i}^{\infty}(a_x + b_x)}{a_{i-1}} \geq \frac{\sum_{x=i}^{i+K-1} a_x}{a_{i-1}} > \frac{K}{2}$$

which implies

$$\limsup_{i\to\infty} \frac{\sum_{x=i}^{\infty}(a_x + b_x)}{a_{i-1}} = \infty \ .$$

Thus by condition 1 of Lemma 4, the chains are subgeometrically ergodic–a contradiction of (c). So $A \neq 1$. But A cannot be greater than 1 either since otherwise $\sum_{x=1}^{\infty} a_x = \infty$ which contradicts the fact that $\sum_{x=1}^{\infty} a_x + \sum_{x=1}^{\infty} b_x = 1$. Therefore, $A < 1$. $\square$

To better understand the conditions for geometric ergodicity provided in Corollary 2, we hereby explain its condition (a) explicitly. First, the requirement that $A = \lim_{i\to\infty} \frac{a_i}{a_{i-1}} < 1$ implies that for any $0 < A_1 < A < A_2 < 1$, there exists i_0 such that for any $i > i_0$, $a_i/a_{i-1} \in (A_1, A_2)$, hence $a_i \in (a_{i_0} A_1^{i-i_0}, a_{i_0} A_2^{i-i_0})$. In other words, the sequence $\{a_i\}$ decays at a geometric rate as i increases. Secondly, the requirements $\lim_{i\to\infty} \frac{a_i}{b_{i-1}} < \infty$ and $\lim_{i\to\infty} \frac{b_i}{a_i} < \infty$ imply that there exist $0 < B_1, B_2 < \infty$ such that, for any $i > i_0$, $a_{i+1}/b_i < B_1$ and $b_i/a_i < B_2$, hence $b_i \in (a_{i+1}B_1, a_i B_2) \subset (a_i A_1 B_1, a_i B_2)$. That is, $b_i = O(a_i)$ as $i \to \infty$. In summary, Condition (a) requires that the sequences $\{a_i\}$ and $\{b_i\}$ both decay geometrically at the same rate as i increases.

We close this section by considering four concrete examples.

Example 1. Let $a_x = c_1 x^{-d}$ and $b_x = c_2 x^{-d}$ where $d > 1$ and $(c_1+c_2)\sum_{x=1}^{\infty} x^{-d} = 1$. Then both $\lim_{i\to\infty} \frac{a_i}{a_{i-1}}$ and $\lim_{i\to\infty} \frac{a_i}{b_i}$ exist, with $A = 1$. Therefore, Φ_X, Φ_{DGS} and Φ_{RGS} are subgeometrically ergodic by Corollary 2.

Example 2. Let c satisfy $(1 + c)e^{-1}/(1 - e^{-1}) = 1$. Set $a_x = ce^{-x}$ and $b_x = e^{-x}$. Then both $\lim_{i\to\infty} \frac{a_i}{a_{i-1}}$ and $\lim_{i\to\infty} \frac{a_i}{b_i}$ exist, with $A = e^{-1} < 1$. Furthermore, $\limsup_{i\to\infty} \frac{a_i}{b_{i-1}} = \lim_{i\to\infty} ce^{-1} < \infty$ and $\limsup_{i\to\infty} \frac{b_i}{a_i} = c^{-1} < \infty$. Therefore, Φ_X, Φ_{DGS} and Φ_{RGS} are all geometrically ergodic by Corollary 2.

Example 3. Let c satisfy $ce^{-1}/(1-e^{-1}) + e^{-2}/(1-e^{-2}) = 1$. Set $a_x = ce^{-x}$ and $b_x = e^{-2x}$. Then both $\lim_{i\to\infty} \frac{a_i}{a_{i-1}}$ and $\lim_{i\to\infty} \frac{a_i}{b_i}$ exist. Also,

$$\limsup_{i\to\infty} \frac{a_i}{b_{i-1}} = \lim_{i\to\infty} ce^{i-2} = \infty \ .$$

Therefore, Φ_X, Φ_{DGS} and Φ_{RGS} are subgeometrically ergodic by Corollary 2.

Example 4. Let c satisfy $ce^{-1}/(1-e^{-1}) + e^{-2}(1-e^{-2}) = 1$. Set

$$a_x = \begin{cases} ce^{-x} & x \text{ even} \\ e^{-2x} & x \text{ odd} \end{cases} \qquad \text{and} \qquad b_x = \begin{cases} e^{-2x} & x \text{ even} \\ ce^{-x} & x \text{ odd} \end{cases} \ .$$

Then $\lim_{i\to\infty} \frac{a_i}{b_i}$ does not exist. Hence Corollary 2 is not applicable. Instead we have to use Lemma 4. Notice that

$$\limsup_{i\to\infty} \frac{b_i}{a_i} \geq \lim_{i\to\infty} \frac{b_{2i+1}}{a_{2i+1}} = \lim_{i\to\infty} \frac{ce^{-(2i+1)}}{e^{-2(2i+1)}} = \lim_{i\to\infty} ce^{2i+1} = \infty$$

and hence Φ_X, Φ_{DGS} and Φ_{RGS} are subgeometrically ergodic.

References

Chan, K. S. and Geyer, C. J. (1994). Comment on "Markov chains for exploring posterior distributions". *The Annals of Statistics*, 22:1747–1758.

Diaconis, P., Khare, K., and Saloff-Coste, L. (2008). Gibbs sampling, exponential families and orthogonal polynomials (with discussion). *Statistical Science*, 23:151–178.

Flegal, J. M., Haran, M., and Jones, G. L. (2008). Markov chain Monte Carlo: Can we trust the third significant figure? *Statistical Science*, 23:250–260.

Flegal, J. M. and Jones, G. L. (2010). Batch means and spectral variance estimators in Markov chain Monte Carlo. *The Annals of Statistics*, 38:1034–1070.

Flegal, J. M. and Jones, G. L. (2011). Implementing Markov chain Monte Carlo: Estimating with confidence. In Brooks, S., Gelman, A., Jones, G. L., and Meng, X.-L., editors, *Handbook of Markov Chain Monte Carlo*, pages 175–197. CRC Press, Boca Raton, FL.

Geyer, C. J. (1992). Practical Markov chain Monte Carlo (with discussion). *Statistical Science*, 7:473–511.

Geyer, C. J. (2011). Introduction to Markov chain Monte Carlo. In Brooks, S. P., Gelman, A., Jones, G. L., and Meng, X.-L., editors, *Handbook of Markov Chain Monte Carlo*, pages 3–48. CRC Press, Boca Raton, FL.

Hobert, J. P. (2011). The data augmentation algorithm: Theory and methodology. In Brooks, S. P., Gelman, A., Jones, G. L., and Meng, X.-L., editors, *Handbook of Markov Chain Monte Carlo*, pages 253–293. CRC Press, Boca Raton, FL.

Hobert, J. P. and Geyer, C. J. (1998). Geometric ergodicity of Gibbs and block Gibbs samplers for a hierarchical random effects model. *Journal of Multivariate Analysis*, 67:414–430.

Hobert, J. P., Jones, G. L., Presnell, B., and Rosenthal, J. S. (2002). On the applicability of regenerative simulation in Markov chain Monte Carlo. *Biometrika*, 89:731–743.

Johnson, A. A. and Jones, G. L. (2010). Gibbs sampling for a Bayesian hierarchical version of the general linear mixed model. *Electronic Journal of Statistics*, 4:313–333.

Johnson, A. A., Jones, G. L., and Neath, R. C. (2011). Component-wise Markov chain Monte Carlo. Technical report, University of Minnesota, School of Statistics.

Jones, G. L. (2004). On the Markov chain central limit theorem. *Probability Surveys*, 1:299–320.

Jones, G. L., Haran, M., Caffo, B. S., and Neath, R. (2006). Fixed-width output analysis for Markov chain Monte Carlo. *Journal of the American Statistical Association*, 101:1537–1547.

Jones, G. L. and Hobert, J. P. (2001). Honest exploration of intractable probability distributions via Markov chain Monte Carlo. *Statistical Science*, 16:312–334.

Jones, G. L. and Hobert, J. P. (2004). Sufficient burn-in for Gibbs samplers for a hierarchical random effects model. *The Annals of Statistics*, 32:784–817.

Liu, J. S., Wong, W. H., and Kong, A. (1994). Covariance structure of the Gibbs sampler with applications to the comparisons of estimators and augmentation schemes. *Biometrika*, 81:27–40.

Liu, J. S., Wong, W. H., and Kong, A. (1995). Covariance structure and convergence rate of the Gibbs sampler with various scans. *Journal of the Royal Statistical Society,* Series B, 57:157–169.

Marchev, D. and Hobert, J. P. (2004). Geometric ergodicity of van Dyk and Meng's algorithm for the multivariate Student's t model. *Journal of the American Statistical Association*, 99:228–238.

Meyn, S. P. and Tweedie, R. L. (1993). *Markov Chains and Stochastic Stability.* Springer-Verlag, London.

Robert, C. P. (1995). Convergence control methods for Markov chain Monte Carlo algorithms. *Statistical Science*, 10:231–253.

Roberts, G. O. and Polson, N. G. (1994). On the geometric convergence of the Gibbs sampler. *Journal of the Royal Statistical Society,* Series B, 56:377–384.

Roberts, G. O. and Rosenthal, J. S. (1997). Geometric ergodicity and hybrid Markov chains. *Electronic Communications in Probability*, 2:13–25.

Roberts, G. O. and Rosenthal, J. S. (1999). Convergence of slice sampler Markov chains. *Journal of the Royal Statistical Society,* Series B, 61:643–660.

Roberts, G. O. and Rosenthal, J. S. (2001). Markov chains and de-initializing processes. *Scandinavian Journal of Statistics*, 28:489–504.

Roberts, G. O. and Rosenthal, J. S. (2004). General state space Markov chains and MCMC algorithms. *Probability Surveys*, 1:20–71.

Román, J. C. (2012). *Convergence Analysis of Block Gibbs Samplers for Bayesian General Linear Mixed Models.* PhD thesis, Department of Statistics, University of Florida.

Román, J. C. and Hobert, J. P. (2011). Convergence analysis of the Gibbs sampler for Bayesian general linear mixed models with improper priors. Technical report, University of Florida, Department of Statistics.

Rosenthal, J. S. (1996). Analysis of the Gibbs sampler for a model related to James-Stein estimators. *Statistics and Computing*, 6:269–275.

Roy, V. and Hobert, J. P. (2007). Convergence rates and asymptotic standard errors for Markov chain Monte Carlo algorithms for Bayesian probit regression. *Journal of the Royal Statistical Society,* Series B, 69:607–623.

Tan, A. and Hobert, J. P. (2009). Block Gibbs sampling for Bayesian random effects models with improper priors: Convergence and regeneration. *Journal of Computational and Graphical Statistics*, 18:861–878.

Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation (with discussion). *Journal of the American Statistical Association*, 82:528–550.

Tierney, L. (1994). Markov chains for exploring posterior distributions (with discussion). *The Annals of Statistics*, 22:1701–1762.