CONVERGENCE OF MOMENTS IN A MARKOV-CHAIN CENTRAL LIMIT THEOREM

DAVID STEINSALTZ

Abstract. Let $(X_i)_{i=0}^{\infty}$ be a V-uniformly ergodic Markov chain on a general state space, and let π be its stationary distribution. For $g : \mathcal{X} \to \mathbb{R}$, define

$$W_k(g) := k^{-1/2} \sum_{i=0}^{k-1} \Bigl(g(X_i) - \pi(g) \Bigr).$$

It is shown that if $|g| \le V^{1/n}$ for a positive integer n, then $\mathrm{E}_x\, W_k(g)^n$ converges to the n-th moment of a normal random variable with expectation 0 and variance

$$\gamma_g^2 := \pi(g^2) - \pi(g)^2 + \sum_{j=1}^{\infty} \left(\int g(x) E_x g(X_j) - \pi(g)^2 \right).$$

This extends the existing Markov-chain central limit theorems, according to which expectations of bounded functionals of $W_k(g)$ converge.

We also derive nonasymptotic bounds for the error in approximating the moments of $W_k(g)$ by the normal moments. These yield easy bounds of all feasible polynomial orders, and exponential bounds as well under some circumstances, for the probabilities of large deviations by the empirical measure along the Markov chain path X_i.

1. Introduction

1.1. **The problem.** Consider an ergodic, positive recurrent Markov chain $(X_i)_{i=1}^{\infty}$. The average

$$S_k(g) := \frac{1}{k} \sum_{i=0}^{k-1} g(X_i)$$

of a bounded function $g : \mathcal{X} \to \mathbb{R}$ along a path converges to the expectation $\pi(g)$ with respect to the stationary distribution, as long as $\pi(|g|)$ is finite. If the chain is strongly mixing, and $\pi(g^2)$ is finite, these averages satisfy a central limit theorem, in that

$$W_k(g) := k^{1/2}\bigl(S_k(g) - \pi(g)\bigr)$$

converges to a normal random variable with expectation 0 and variance

$$\gamma_g^2 := \pi(g^2) - \pi(g)^2 + \sum_{j=1}^{\infty} \left(\int g(x) E_x g(X_j) - \pi(g)^2 \right). \tag{1}$$

E. Bolthausen [Bol82] has shown that the error in this normal approximation is on the order of $k^{-1/2}$.

There are two directions in which one might hope to improve this result. First, this is only a weak-convergence result, telling us about the maximum difference

Date: September 17, 2001.

1991 *Mathematics Subject Classification.* 60F10, 60J05.

between the distribution functions of $W_k(g)$ and the normal variable. It tells us nothing about how the tails of $W_k(g)$ fall off, nor does it offer any bounds on the moments of $W_k(g)$ of order greater than 2; indeed the very existence of these moments remains uncertain. This is an essential problem, when we seek to bound the large-deviation probabilities for $S_k(g)$.

A related weakness is that this convergence rate assumes that the chain starts in its stationary distribution. It is thus less useful for infinite state-space chains in which the convergence to stationarity is not uniform. The chain may be exponentially mixing, in that covariances fall off exponentially between a π-typical starting point and the location at time i, but there may still be π-small sets where the process dallies a very long time when once started there. The exceptional starting point will then make itself felt particularly strongly in the empirical average S_k. For example, suppose X_i were a random walk on $\mathbb{Z}$ with drift toward 0 ($\mathrm{P}_{x,x+1} = p$ and $\mathrm{P}_{x,x-1} = 1-p$ for $x \geq 1$, while $\mathrm{P}_{x,x-1} = p$ and $\mathrm{P}_{x,x+1} = 1-p$ for $x \leq -1$, and $\mathrm{P}_{0,-1} = \mathrm{P}_{0,1} = \frac{1}{2}$; with $p < \frac{1}{2}$), and $g(x) = x$: When X_i is very large, X_{i+1} is large as well, and it takes on average $3X_0$ steps before it even reaches 0 for the first time.

In order to control the dependence of convergence rates on the starting point, we impose a mixing condition stronger than exponential mixing[1], but weaker than uniform ergodicity. As described by S. Meyn and R. Tweedie [MT93], a Markov chain X_n on the state space $\mathcal{X}$ is V-uniformly ergodic, for $V : \mathcal{X} \to [1, \infty)$,

$$\sup_g \sup_{x \in \mathcal{X}} \frac{1}{V(x)} \left| \mathrm{E}_x\left[g(X_i)\right] - \int g \, d\pi \right| \xrightarrow{i \to \infty} 0, \tag{2}$$

where the first supremum is over measurable functions $g : \mathcal{X} \to \mathbb{R}$ such that $|g| \leq V$. In our simple example above, the chain is V-uniformly ergodic for $V(x) = e^{\lambda x}$, where $0 < \lambda < \log(p^{-1} - 1)$. (This particular investigation arose from the appearance of moment bounds for such a V-uniformly ergodic $W_k(g)$ in an application to iterated function systems in [Ste01].)

In this paper, we show that when the Markov chain is V-uniformly ergodic, and $|g - \pi(g)| \leq V^{1/n}$, then the n-th moment of $W_k(g)$ converges to the n-th moment of the normal random variable, and for the error we derive bounds which are constant multiples of $k^{-1/2}V(x)$. these starting-point-dependent bounds allow us, in addition, to extend Bolthausen's result to include the starting-point dependence as well: we show that when $|g - \pi(g)| \leq V^{1/\alpha}$, for a positive $\alpha \geq 2$, and the process starts at x, the error in estimating $W_k(g)$ by a normal random variable is no larger than order $k^{-\alpha/(2\alpha+2)}V(x)$ (up to logarithmic terms). We do not know whether this rate is the best possible.

Just recently, S. Meyn and S. Balaji [BM00] and S. Meyn and I. Kontoyiannis[MK] have proved results which may loosely be summarized as

$$E_x e^{\alpha S_k(g)} \sim c(\alpha) \check{f}_\alpha(x) e^{-\Lambda(\alpha)k},$$

together with a recipe for computing $\check{f}_\alpha$ and $\Lambda(\alpha)$. This goes beyond a complete solution to the classical large-deviations problem for S_k, but there are features which make this version less than ideal for many purposes. First, as with the

[1] We follow Bolthausen and others in calling a Markov chain "exponentially mixing" if there are positive c and γ such that $\mathrm{E}_\pi[g(X_0)g(X_n)] \leq c\exp\{-\gamma n\}$ for every n, for all functions $g : \mathcal{X} \to [-1, 1]$ such that $\pi(g) = 0$. S. Meyn has pointed out in a private communication that other definitions of exponential mixing are common in the engineering literature.

central limit theorem, this is an asymptotic result, providing no bounds on the large-deviation probabilities for any individual finite k. Second, the computations are not always feasible. The function $\check{f}_\alpha$ is determined implicitly as the solution of a multiplicative version of the Poisson equation; once we have $\check{f}_\alpha$, it defines a second implicit equation for $\Lambda(\alpha)$, involving expectations with respect to a stopping time for the Markov chain. Even whether $\Lambda(\alpha)$ is zero need not be obvious. Finally, even when everything has been computed, the theorem may leave us with empty hands if $\Lambda(\alpha)$ really is 0. The large-deviation probabilities may still go to 0 with k, but more slowly than any exponential function. In particular, we may want to consider partial sums of functions g which do not have exponential moments at all, so $\pi(e^{\alpha g})$ is infinite for all α, but where $\pi(g^n)$ is still finite for some n. In such a case we cannot hope to have large-deviation probabilities with exponentially declining tails, but we can still have tails falling like

$$\mathrm{P}\left\{\left|S_k(g) - \pi(g)\right| > \lambda\right|\right\} \le ck^{-n/2}\lambda^{-n}.$$

1.2. **Notation.** Throughout, $(X_i)_{i=0}^\infty$ will be a ψ-irreducible aperiodic Markov chain on the state space $\mathcal{X}$, and $V : \mathcal{X} \to [1,\infty)$ a function such that X_i is V-uniformly ergodic. The distribution of X_i conditioned on $X_0 = x$ will be denoted P_x^i, and the stationary distribution will be π.

By Theorem 16.0.1 of [MT93] there are constants $R \ge 1$ and $\rho < 1$ such that for all $i \in \{0, 1, \ldots\}$, all $x \in \mathcal{X}$, and all $g : \mathcal{X} \to \mathbb{R}$ with $|g| \le V$,

$$\left|P_x^i(g) - \pi(g)\right| \le R\rho^i V(x). \tag{3}$$

We define $R_* := \max\{R, \pi(V)\}$.

We will use the combinatorialist notation $(2n-1)!! = (2n-1)(2n-3)\cdots 3\cdot 1$.

The variance γ_g^2 is the limit of the variances of the random variables W_k, starting from the stationary distribution, which we write as

$$\gamma_g^2(k) = \pi(g^2) - \pi(g)^2 + \sum_{i=1}^{k-1} \frac{2(k-i)}{k}\left(\int g(x) P_x^i(g)\pi(dx) - \pi(g)^2\right). \tag{4}$$

This may be written alternatively as

$$\gamma_g^2(k) = \pi\big(T_0^{(g)}\big) + \sum_{i=1}^{k-1} \frac{2(k-i)}{k}\pi\big(T_i^{(g)}\big), \tag{5}$$

where $T_i^{(g)} : \mathcal{X} \to \mathbb{R}$ is defined by

$$T_i^{(g)}(x) := \big(g(x) - \pi(g)\big) E_x\big[g(X_i) - \pi(g)\big]. \tag{6}$$

Usually it will be apparent from context which function g is meant, and then the superscript (g) will be dropped.

1.3. **The results.**

Theorem 1. *Suppose $g : \mathcal{X} \to \mathbb{R}$ satisfies $|g - \pi(g)| \le cV^{1/n}$ for an integer $n \ge 2$. Then for all starting states $x \in \mathcal{X}$, the moments $\mathrm{E}_x\big[W_k(g)^n\big]$ converge to the corresponding moments of a Gaussian random variable with expectation 0 and variance*

γ_g^2; that is, for all positive n,

$$\lim_{k\to\infty} \mathrm{E}_x\big[W_k(g)^{2n+1}\big] = 0, \text{ and} \tag{7}$$

$$\lim_{k\to\infty} \mathrm{E}_x\big[W_k(g)^{2n}\big] = (2n-1)!!(\gamma_g^2)^n. \tag{8}$$

Furthermore, there are positive constants $r, r' < 1$ and C, C', depending only on ρ and R_*, such that the error terms are bounded by

$$\Big| E_x W_k(g)^{2n} - (2n-1)!!\big(\gamma_g^2\big)^n \Big| \le k^{-1} C (cr)^{2n} \frac{n(2n)!}{(n-1)!}\Big(1+\frac{n!}{k}\Big) V(x), \text{ and} \tag{9}$$

$$\Big| E_x W_k(g)^{2n+1} \Big| \le k^{-1/2} C' (cr')^{2n+1} \frac{(2n+1)!}{n!}\Big(1+\frac{n!}{k}\Big) V(x), \tag{10}$$

when $|g-\pi(g)|$ is bounded by $cV^{1/(2n)}$ or $cV^{1/(2n+1)}$ respectively. Explicit expressions for the errors are given in (42) and (43).

Corollary 1. *If there is a positive c such that $|g(x) - \pi(g)| \le c$ for all x, then for any $\lambda < 1/c(r \vee r')$, all $k \ge 1$, and all $x \in \mathcal{X}$,*

$$\begin{aligned} &\Big| E_x \exp\big\{\lambda W_k(g)\big\} - \mathrm{E}\exp\big\{\lambda\gamma_g X\big\}\Big| \\ &\quad \le k^{-1/2} V(x) \Big[C'\lambda e^{(\lambda c r')^2} + k^{-1/2} C e^{(\lambda c r)^2} \\ &\qquad + k^{-1} \frac{C'\lambda}{1-(\lambda c r')^2} + k^{-3/2} \frac{C}{1-(\lambda c r)^2} \Big], \end{aligned} \tag{11}$$

where X is a standard normal variable.

Corollary 2. *Suppose $g : \mathcal{X} \to \mathbb{R}$ satisfies $|g - \pi(g)| \le cV^{1/2n}$ for some positive integer n. Then there are constants M_n^*, such that for all starting states $x \in \mathcal{X}$, all $k \ge 2$, and all measurable functions $\psi : \mathbb{R} \to \mathbb{R}$,*

$$\Big| \mathrm{E}_x\, \psi(W_k(g)) - \mathrm{E}\,\psi(X) \Big| \le M_n^* \|\psi\| V(x) \left(\frac{k}{\log k}\right)^{-n/(2n+1)}, \tag{12}$$

where $\|\psi\|$ is the total variation of ψ and X is a standard normal variable.

If $|g - \pi(g)|$ is bounded by a constant c, then the bound can be strengthened to

$$\Big| \mathrm{E}_x\, \psi(W_k(g)) - \mathrm{E}\,\psi(X) \Big| \le L^* \|\psi\| V(x) k^{-1/2} \log^{3/2} k. \tag{13}$$

The constants M_α^* and L^* are given in (51) and (52) respectively. These expressions involve, in addition to c, ρ, and R_*, an as yet undetermined parameter β, defined in (44), which is the uncomputed constant that appears in the Berry-Esseen theorem for strongly mixing Markov chains.

Corollary 3. *If there is a number $p \in [0, 1)$ and positive c such that*

$$\big|g(x) - \pi(g)\big| \le c\big(\log V(x)\big)^p$$

for all x, then for any positive $q < 2/(1+2p)$, any positive λ, and any positive r,

$$\begin{aligned} &\sup_x V(x)^{-1} \Big| \mathrm{E}_x \exp\big\{\lambda |W_k(g)|^q\big\} - \mathrm{E}\exp\big\{\lambda \big|\gamma_g X\big|^q\big\} \\ &\qquad = O\Big(log^{-r} k\Big), \end{aligned} \tag{14}$$

where X is a standard normal variable.

1.4. **Remarks on exponential functions.** The statement of Corollary 3 may seem unnecessarily timid. In one sense, this is true: the rate $\log^{-r} k$ is certainly not optimal, but could probably be improved if the interpolation between integer moments were more cleverly finessed. But a more significant question suggests itself: if $|g| \le cV^{1/n}$ implies that $\mathrm{E}\, W_k(g)^n$ converges to $\mathrm{E}\, X^n$, where X is normal, then should not $\mathrm{E}\exp\{\lambda W_k(g)\}$ converge to $\mathrm{E}\exp\{\lambda X\}$ when $|g| \le c \log V$, at least for λ sufficiently small? A simple example makes clear why this cannot be the case in general.

Consider the random walk drifting toward 0, mentioned earlier, with outward probability $p < \frac{1}{2}$, inward $1-p$. As already mentioned, this is V-uniformly ergodic for an exponential function $V(x) = e^{\lambda x}$, so we may take $g(x) = x$. The stationary distribution is a constant times $(p/(1-p))^{|k|}$ for $k \neq 0$, and $\pi(g) = 0$ (when $g(x) = x$). If we take any positive number λ,

$$\begin{aligned}
\mathrm{E}_x\, e^{\lambda X_1} &= e^{\lambda x}\left(pe^{\lambda} + (1-p)e^{-\lambda}\right) \qquad \text{for } x \ge 1,\\
\mathrm{E}_x\, e^{\lambda X_1} &= e^{\lambda x}\left(pe^{-\lambda} + (1-p)e^{\lambda}\right) \qquad \text{for } x \le -1, \text{ and}\\
E_0 e^{\lambda X_1} &= \frac{1}{2}e^{\lambda} + \frac{1}{2}e^{-\lambda}.
\end{aligned}$$

Defining $r_\lambda := pe^{\lambda} + (1-p)e^{-\lambda}$, we have then

$$\mathrm{E}_x\, e^{\lambda X_1} \ge r_\lambda e^{\lambda x}.$$

If we look at two consecutive steps,

$$\mathrm{E}_x\left[\exp\{\lambda(X_1 + X_2)\}\right] = \mathrm{E}_x\left[e^{\lambda X_1} E_{X_1} e^{\lambda X_2}\right] \le r_\lambda\, \mathrm{E}_x\left[e^{2\lambda X_1}\right] \le r_\lambda^2 e^{2\lambda x}.$$

Extending this to larger sums, we get for any k,

$$\mathrm{E}_x\, e^{\lambda W_k(g)} \ge r_\lambda^k e^{\lambda\sqrt{k}x}.$$

The crucial point here is that the exponent grows with k. No matter what λ is, eventually $e^{\lambda\sqrt{k}}$ will exceed $(1-p)/p$, which means that

$$\int \mathrm{E}_x\, e^{\lambda W_k(g)}\pi(dx)$$

is infinite for k sufficiently large. Convergence of $E_x e^{\lambda W_k(g)}$ is impossible.

2. Some technicalities about expectations of products

Lemma 1. *For $\alpha \in [0,1]$ and $\beta \in \mathbb{R}$, suppose $\left|g(x) - \beta\right| \le cV(x)^\alpha$ for all $x \in \mathcal{X}$. Then for any nonnegative integer k,*

$$\left|P_x^k(g) - \pi(g)\right| \le 2cR^\alpha \rho^{\alpha k} V(x)^\alpha \qquad \textit{and} \tag{15}$$

$$\left|T_k(x)\right| \le 2c^2 R^\alpha \rho^{\alpha k} V(x)^{2\alpha} + \left|\beta - \pi(g)\right| \cdot 2cR^\alpha \rho^{\alpha k} V(x)^{2\alpha}. \tag{16}$$

If $\alpha \le \frac{1}{2}$, then

$$\left|\pi(T_k)\right| \le 2c^2 R^\alpha \rho^{\alpha k} \pi(V)^{2\alpha} \qquad \textit{and} \tag{17}$$

$$\gamma_g^2 \le \frac{4c^2 R^\alpha \pi(V)^{2\alpha}}{\alpha(1-\rho)}. \tag{18}$$

If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are nonnegative numbers with $a := \alpha_1 + \alpha_2 + \cdots + \alpha_n \le 1$, and $g_1, g_2, \ldots, g_n$ are functions from $\mathcal{X}$ to $\mathbb{R}^+$ with $g_k \le V^{\alpha_k}$; then for all $x \in \mathcal{X}$ and all indices $0 \le i_1 < i_2 < \cdots < i_n$,

$$\left| \mathrm{E}_x \left[\prod_{\ell=1}^{n} g_\ell(X_{i_\ell}) \right] \right| \le \big(2R_* V(x)\big)^a. \tag{19}$$

Proof. By the Hahn decomposition theorem (Proposition 11.21 of [Roy68]) there is a measurable subset A of $\mathcal{X}$, such that the signed measure $P_x^k - \pi$ is positive on A and negative on the complement of A; the absolute-value measure is defined by $|P_x^k - \pi|(g) = (P_x^k - \pi)(gh)$, where $h = 2 \cdot \mathbf{1}_A - 1$. By Jensen's inequality, when $\|P_x^k - \pi\| := (P_x^k - \pi)(h) \ne 0$,

$$\begin{aligned}
\big|P_x^k(g) - \pi(g)\big| &= \big|\big(P_x^k - \pi\big)\big(g - \beta\big)\big| \\
&= \frac{\big|P_x^k - \pi\big|}{\big\|P_x^k - \pi\big\|}\Big(\big(|g-\beta|^{1/\alpha}\big)^{\alpha}\Big) \cdot \big\|P_x^k - \pi\big\| \\
&\le \left(c^{1/\alpha} \frac{\big|P_x^k - \pi\big|}{\big\|P_x^k - \pi\big\|}(V)\right)^{\alpha} \cdot \big\|P_x^k - \pi\big\| \\
&\le 2^{1-\alpha} c \big|P_x^k(hV) - \pi(hV)\big|^{\alpha}.
\end{aligned}$$

Thus the bound (3) gives us (15). Multiplying by $|g(x) - \beta| + |\beta - \pi(g)|$ then yields (16). To bound $\pi(T_k)$, we write

$$\begin{aligned}
\big|\pi(T_k)\big| &= \Big|\pi\Big(\big(g - \pi(g)\big)\big(P_x^k(g) - \pi(g)\big)\Big)\Big| \\
&\le \Big|\pi\Big(\big(g - \beta\big)\big(P_x^k(g) - \pi(g)\big)\Big)\Big| + \Big|\big(\beta - \pi(g)\big)\pi\Big(P_x^k(g) - \pi(g)\big)\Big)\Big| \\
&\le \Big|\pi\Big(cV(x)^{\alpha} \cdot 2cR^{\alpha}\rho^{\alpha k} V(x)^{\alpha}\Big)\Big| + 0,
\end{aligned}$$

from which (17) follows directly. We then have

$$\begin{aligned}
\gamma_g^2 &= \pi(T_0) + 2\sum_{i=1}^{\infty} \pi(T_i) \\
&\le 2c^2 R^{\alpha} \pi(V)^{2\alpha} \frac{2}{1 - \rho^{\alpha}} \\
&\le \frac{4c^2 R^{\alpha} \pi(V)^{2\alpha}}{\alpha(1-\rho)}.
\end{aligned}$$

For each ℓ, by (3),

$$\begin{aligned}
\mathrm{E}_x\Big[\big|g_\ell(X_{i_\ell})\big|^{1/\alpha_\ell}\Big] &\le \pi\Big(|g_\ell|^{1/\alpha_\ell}\Big) + \Big|\mathrm{P}_x^{i_\ell}\Big(|g_\ell|^{1/\alpha_\ell}\Big) - \pi\Big(|g_\ell|^{1/\alpha_\ell}\Big)\Big| \\
&\le \pi(V) + R\rho^{i_\ell} V(x) \\
&\le 2R_* V(x).
\end{aligned}$$

The result (19) then follows by an application of Hölder's inequality. □

A simple extension of Lemma 1 is

Lemma 2. *Let $g, h : \mathcal{X} \to \mathbb{R}$ and $f : \mathcal{X}^{i+1} \to \mathbb{R}$ be any measurable functions, with $\pi(g) = 0$. Suppose there are positive constants c_f, c_g, c_h and $\alpha_f, \alpha_g, \alpha_h$ with*

$\alpha_f + 2\alpha_g + \alpha_h \le 1$, such that for all x in $\mathcal{X}$,

$$\begin{aligned} \big|g(x)\big| &\le c_g V(x)^{\alpha_g}, \\ \big|h(x)\big| &\le c_h V(x)^{\alpha_h}, \text{ and} \\ \big|E_x\big[f(X_0, X_1, \ldots, X_i)\big| &\le c_f V(x)^{\alpha_f}. \end{aligned}$$

Then there is a function $h^* : \mathcal{X} \to \mathbb{R}$ with $|h^*(x)| \le V(x)^{\alpha_g + 2\alpha_h}$ for all x, such that for any i, j and ℓ with $0 \le i \le j \le \ell$,

$$\begin{aligned} &\big|E_x f(X_0, X_1, \ldots, X_i) g(X_j) g(X_\ell) h(X_\ell)\big| \\ &\qquad \le 2c_g^2 c_h R_*^{3\alpha_g + 2\alpha_h} \left(\rho^{(\alpha_g+\alpha_h)(j-i)} + \rho^{(\alpha_g+\alpha_h)(\ell-j)}\right) \\ &\qquad\qquad \times \mathrm{E}_x\big[f(X_0, X_1, \ldots, X_i) h^*(X_i)\big]. \end{aligned} \tag{20}$$

Proof. Define

$$\begin{aligned} h'(x) &:= P_x^{\ell-j}(gh) - \pi(gh), \qquad \text{and} \\ h''(x) &:= \mathrm{E}_x\big[g(X_{j-i}) g(X_{\ell-i}) h(X_{\ell-i})\big]. \end{aligned}$$

By Lemma 1 with $\beta = 0$,

$$\begin{aligned} \big|\mathrm{P}_x^{j-i}(g)\big| &\le 2c_g R^{\alpha_g+\alpha_h} V(x)^{\alpha_g+\alpha_h} \rho^{(\alpha_g+\alpha_h)(j-i)}, \qquad \text{and} \\ \big|h'(x)\big| &\le 2c_g c_h R^{\alpha_g+\alpha_h} V(x)^{\alpha_g+\alpha_h} \rho^{(\alpha_g+\alpha_h)(\ell-j)}. \end{aligned}$$

Note that in the first line we have applied the lemma with $\alpha = \alpha_g + \alpha_h$ instead of α_g. Either one would satisfy the conditions, but our goal is to make the exponents of $V(x)$ as small as possible, and the exponents of ρ as large as possible, which leads us to get the same exponents for both terms. Thus

$$\begin{aligned} \big|h''(x)\big| &= \Big|\mathrm{E}_x\big[g(X_{j-i})\big(\pi(gh) + h'(X_{j-i})\big)\big]\Big| \\ &\le c_g c_h \pi(V)^{\alpha_g+\alpha_h} \big|\mathrm{P}_x^{j-i}(g)\big| + P_x^{j-i}(|gh'|) \\ &\le 2R^{\alpha_g+\alpha_h} c_g^2 c_h \pi(V)^{\alpha_g+\alpha_h} V(x)^{\alpha_g+\alpha_h} \rho^{(\alpha_g+\alpha_h)(j-i)} \\ &\qquad + 2R_*^{2\alpha_g+\alpha_h} R^{\alpha_g+\alpha_h} c_g^2 c_h V(x)^{2\alpha_g+\alpha_h} \rho^{(\alpha_g+\alpha_h)(\ell-j)} \\ &\le 2R_*^{3\alpha_g+2\alpha_h} c_g^2 c_h V(x)^{2\alpha_g+\alpha_h} \left(\rho^{(\alpha_g+\alpha_h)(j-i)} + \rho^{(\alpha_g+\alpha_h)(\ell-j)}\right). \end{aligned}$$

Equation (20) holds then with

$$h^* = h''/\left(2R_*^{3\alpha_g+2\alpha_h} c_g^2 c_h \left(\rho^{(\alpha_g+\alpha_h)(j-i)} + \rho^{(\alpha_g+\alpha_h)(\ell-j)}\right)\right).$$

□

This leads us finally to

Lemma 3. *Let g and h be given as in Lemma 2. Then for any integers $0 \le i_1 \le i_2 \le \cdots \le i_{2n}$,*

$$\begin{aligned} &\Big|\mathrm{E}_x\big[g(X_{i_1}) \cdots g(X_{i_{2n}}) h(X_{i_{2n}})\big]\Big| \\ &\le 2^n R_*^{(2n^2+n)\alpha_g + 2n\alpha_h} c_g^{2n} c_h V(x)^{2n\alpha_g+\alpha_h} \\ &\quad \times \prod_{j=1}^n \left(\rho^{[(2n-2j+1)\alpha_g+\alpha_h](i_{2j}-i_{2j-1})} + \rho^{[(2n-2j+1)\alpha_g+\alpha_h](i_{2j-1}-i_{2j-2})}\right). \end{aligned} \tag{21}$$

Proof. The proof is by induction on n. For $n=1$ this is just the statement is just Lemma 2, with $f\equiv 1$ and $i=0$. Suppose now that (21) holds up to $n-1$. We apply Lemma 2 with $f(x_0,\dots,x_{i_{2n-2}})=g(x_{i_1})\cdots g(x_{i_{2n-2}})$:

$$\begin{aligned}\mathrm{E}_x\big[g(X_{i_1})\cdots g(X_{i_{2n-2}})g(X_{i_{2n-1}})g(X_{i_{2n}})h(X_{i_{2n}})\big]&\\ \le\Big(2R_*^{3\alpha_g+2\alpha_h}c_g^2c_h\Big(\rho^{(\alpha_g+\alpha_h)(i_{2n}-i_{2n-1})}+\rho^{(\alpha_g+\alpha_h)(i_{2n-1}-i_{2n-2})}\Big)\Big)&\\ \times\mathrm{E}_x\big[g(X_{i_1})\cdots g(X_{i_{2n-2}})h^*(X_{i_{2n-2}})\big],&\end{aligned}$$

where $|h^*(x)|\le V(x)^{\alpha_h+2\alpha_g}$. Applying the induction hypothesis,

$$\begin{aligned}&\mathrm{E}_x\big[g(X_{i_1})\cdots g(X_{i_{2n-2}})g(X_{i_{2n-1}})g(X_{i_{2n}})h(X_{i_{2n}})\big]\\ &\quad\le 2R_*^{3\alpha_g+2\alpha_h}c_g^2c_h\Big(\rho^{(\alpha_g+\alpha_h)(i_{2n}-i_{2n-1})}+\rho^{(\alpha_g+\alpha_h)(i_{2n-1}-i_{2n-2})}\Big)\\ &\qquad\times 2^{n-1}R_*^{(2n^2-3n+1)\alpha_g+(2n-2)(\alpha_h+2\alpha_g)}c_g^{2n-2}V(x)^{2(n-1)\alpha_g+\alpha_h+2\alpha_g}\\ &\qquad\quad\times\prod_{\ell=1}^{n-1}\Big(\rho^{[(2n-2\ell-1)\alpha_g+\alpha_h+2\alpha_g](i_{2\ell}-i_{2\ell-1})}+\rho^{[(2n-2\ell-1)\alpha_g+\alpha_h+2\alpha_g](i_{2\ell-1}-i_{2\ell-2})}\Big),\end{aligned}$$

which reduces precisely to (21). □

The right side of (21) includes a sum of 2^n terms, each of which is a power of ρ. The powers can be written in the form $\sum_{\ell=1}^{2n}\eta_\ell(i_\ell-i_{\ell-1})$, where the η_ℓ are nonnegative, and exactly n of them are zero. These will need to be summed over all possible choices of $(i_1,\dots,i_{2n})$ with $0\le i_1\le\cdots i_{2n}\le k-1$.

Lemma 4. *Choose any nonnegative numbers η_ℓ, for $1\le\ell\le N$. Let s be the number of these η_ℓ which are nonzero. Then for any $x\in(0,1)$,*

$$\begin{aligned}\sum_{i_1=0}^{k-1}\sum_{i_2=i_1}^{k-1}\cdots\sum_{i_N=i_{N-1}}^{k-1}\prod_{\ell=1}^{N}x^{\eta_\ell(i_\ell-i_{\ell-1})}&\le\binom{N-s+k-1}{N-s}\prod_{\ell:\eta_\ell>0}(1-x^{\eta_\ell})^{-1}\\ &\le\left(\frac{k^{N-s}}{(N-s)!}+k^{N-s-1}\right)\prod_{\ell:\eta_\ell>0}(1-x^{\eta_\ell})^{-1}\end{aligned}\tag{22}$$

Proof. The proof is by induction on s. For $s=0$ the summand is 1, so the sum is simply the number of possible choices of $(i_1,\dots,i_N)$ with $0\le i_1\le i_2\le\cdots\le i_N\le k-1$. By a standard combinatorial argument this is found to be $\binom{N+k-1}{N}$.

Suppose now the lemma to be true for $s-1$. Since $s\ge1$, there is some ℓ such that $\eta_\ell>0$. We find the largest such, and begin by summing over the index i_ℓ. This index is free to range from $i_{\ell-1}$ up to $k-1$, and the sum can only increase if the upper limit is removed, allowing the summation to extend up to ∞, and if the lower limit is relaxed on $i_{\ell+1}$ (when ℓ is not already $2n$), permitting that index to range down to $i_{\ell-1}$. The summand is $x^{\eta_\ell(i_\ell-i_{\ell-1})}$, times terms which do not depend on i_ℓ. Summing over i_ℓ yields $(1-x^{\eta_\ell})^{-1}$, and the sum that remains has one fewer index and one fewer nonzero η. Thus the induction hypothesis may be applied to this remnant, proving the first part of the inequality.

The binomial coefficient is a polynomial of degree $N-s$ in k, with leading coefficient $1/(N-s)!$ and all other coefficients positive. The total of all the coefficients, found by setting $k=1$, is 1, so the sum of all the remaining terms is no more than k^{N-s-1}. □

3. Combinatorics of pairings

A *pairing* of $[2n] := \{1, 2, \ldots, 2n\}$ is a set of unordered pairs $\{j, k\} \subset [2n]$, such that each $j \in [2n]$ appears in exactly one pair. Pairings are a device for grouping different ways of ordering a multiindex $(i_1, i_2, \ldots, i_n)$. A pairing may also be thought of as a self-inverse bijection from $[2n]$ to itself, defining $\sigma(j)$ to be the unique element $k \in [2n]$ such that $\{j, k\} \in \sigma$.

To any ordered multiset of integers $I = (i_1, i_2, \ldots, i_{2n})$ we associate a pairing as follows: Let s_ℓ be the rank that i_ℓ has when the sequence is put in order; that is,

$$s_\ell := \#\{j \in [2n] \, : \, i_j < i_\ell\} \; + \; \#\{j \in [\ell] \, : \, i_j = i_\ell\}.$$

(For definiteness, when two terms have the same value, their order is maintained.) Then I is associated to the pairing

$$\sigma_I := \big\{\{s_{2\ell-1}, s_{2\ell}\} \, : \, 1 \le \ell \le n\big\}.$$

We will say that a multiset $I = (i_1, i_2, \ldots, i_{2n})$ is *matched* if σ_I is the trivial pairing $\sigma_0 := \big\{\{1, 2\}, \{3, 4\}, \ldots, \{2n-1, 2n\}\big\}$. The sequence is *ordered* if $i_1 \le i_2 \le \cdots \le i_{2n}$. An ordered sequence is said to have an *overlapping pair* if there is some even j such that $i_j = i_{j+1}$.

As an illustration, when $n = 4$ the multiset $(4, 4, 8, 7, 9, 9, 1, 2)$ is matched. The sequence $I = (5, 3, 2, 7, 5, 1, 8, 9)$, on the other hand, has matching $\sigma_I = \big\{\{3, 4\}, \{2, 6\}, \{1, 5\}, \{7, 8\}\big\}$.

This definition is adapted to the equation (30). There we have a sum over all multiindices $I = (i_1, \ldots, i_{2n})$, where the summand is $\pi(T_{|i_2-i_1|}) \cdot \pi(T_{|i_4-i_3|}) \cdots \pi(T_{|i_{2n}-i_{2n-1}|})$. The pairing σ_I tells us which places would be paired if the indices were first put in order. In the above example, we have the product $\pi(T_{5-3})\pi(T_{7-2})\pi(T_{5-1})\pi(T_{9-8})$. If we order the indices, we get the multiindex $J = (1, 2, 3, 5, 5, 7, 8, 9)$, and we need to know that the first term in the product is $\pi(T_{j_4-j_3})$, and so on. This leads us to the definition

$$|\sigma|(j_1, j_2, \ldots, j_{2n}) := \sum_{\{s, s'\} \in \sigma} \big| i_s - i_{s'} \big|, \tag{23}$$

where $(j_1, j_2, \ldots, j_{2n})$ is taken to be an ordered multiindex. We can also write this as

$$|\sigma|(j_1, j_2, \ldots, j_{2n}) := \sum_{\ell=1}^{2n} \gamma_\ell(\sigma) j_\ell, \qquad \text{where } \gamma_\ell(\sigma) := \begin{cases} +1 & \text{if } \ell > \sigma(\ell), \\ -1 & \text{if } \ell < \sigma(\ell). \end{cases} \tag{24}$$

Let $I = (i_1, \ldots, i_{2n})$ be a multiindex with pairing σ, and let $(j_1, \ldots, j_{2n})$ be the ordered version of I. Then

$$|i_2 - i_1| + |i_4 - i_3| + \cdots + |i_{2n} - i_{2n-1}| = |\sigma|(j_1, j_2, \ldots, j_{2n}) \tag{25}$$

Let σ be a pairing on $[2n]$. We associate to σ a graph, whose vertices are $\{1, \ldots, 2n\}$, with edges connecting $2j-1$ and $2j$, as well as connecting j and k if $\{j, k\} \in \sigma$. A subset of $[2n]$ well be called *σ-connected* if the corresponding vertices are connected in the graph, and it will be called a *component* of the pairing if they form a connected component of this graph.

If A is a σ-connected subset of $[2n]$, we can define the *restriction* of σ to A. The elements of A come in pairs, so they can be written as $\eta_1 \le \eta_2 \le \cdots \le \eta_{2\ell}$, where each η_{2j-1} is odd and $\eta_{2j} = \eta_{2j-1} + 1$. The restriction $\sigma\Big|_A$ is then defined by

$$\{j,k\} \in \sigma\Big|_A \Longleftrightarrow \{\eta_j, \eta_k\} \in \sigma.$$

A nearly obvious fact is

Lemma 5. *For any m with $1 \le m \le 2n$, the sum $\sum_{j=m}^{2n} \gamma_j(\sigma) \ge 0$, and the sum is equal to 0 if and only if the set $\{m, m+1, \dots, 2n\}$ is σ-connected.*

Proof. We have

$$\sum_{j=m}^{2n} \gamma_j(\sigma) = -\#\{j : m \le j < \sigma(j)\} + \#\{j : m \le \sigma(j) < j\} + \#\{j : \sigma(j) < m \le j\}.$$

The first two terms cancel each other out. The sum is thus nonnegative, and is zero precisely when the last term is zero, that is, when every j in $\{m, \dots, 2n\}$ has its σ-partner in $\{m, \dots, 2n\}$ as well. □

A consequence of this lemma is

Lemma 6. *If σ is any pairing on $[2n]$ with exactly κ components and $x \in (0,1)$, then*

$$\sum_{i_1=0}^{k-1} \sum_{i_2=i_1}^{k-1} \cdots \sum_{i_{2n}=i_{2n-1}}^{k-1} x^{|\sigma|(i_1,i_2,\dots,i_{2n})} \le \frac{(k(1-x))^\kappa}{(1-x)^{2n}}. \tag{26}$$

Proof. Suppose first that $\kappa = 1$. By (24),

$$\begin{aligned}
&\sum_{i_1=0}^{k-1} \sum_{i_2=i_1}^{k-1} \cdots \sum_{i_{2n}=i_{2n-1}}^{k-1} x^{|\sigma|(i_1,i_2,\dots,i_{2n})} \\
&\quad = \sum_{i_1=0}^{k-1} x^{\gamma_1 i_1} \cdot \sum_{i_2=i_1}^{k-1} x^{\gamma_2 i_2} \cdots \sum_{i_{2n}=i_{2n-1}}^{k-1} x^{\gamma_{2n} i_{2n}} \\
&\quad = \sum_{i_1=0}^{k-1} x^{\gamma_1 i_1} \cdot \sum_{i_2=i_1}^{k-1} x^{\gamma_2 i_2} \cdots \sum_{i_{2n-1}=i_{2n-2}}^{k-1} x^{(\gamma_{2n-1}+\gamma_{2n}) i_{2n-1}} \cdot \frac{1}{1-x}.
\end{aligned}$$

Here we have bounded the sum up to $k-1$ by an infinite sum. Lemma 5 tells us that the sum $\gamma_m + \cdots + \gamma_{2n}$ is always positive for $m \ge 2$. By induction it follows then that for all $m \in [2n]$

$$\begin{aligned}
&\sum_{i_1=0}^{k-1} \sum_{i_2=i_1}^{k-1} \cdots \sum_{i_{2n}=i_{2n-1}}^{k-1} x^{|\sigma|(i_1,i_2,\dots,i_{2n})} \\
&\quad \le (1-x)^{m-2n} \sum_{i_1=0}^{k-1} x^{\gamma_1 i_1} \cdot \sum_{i_2=i_1}^{k-1} x^{\gamma_2 i_2} \cdots \sum_{i_m=i_{m-1}}^{k-1} x^{(\gamma_m+\gamma_{m+1}+\cdots+\gamma_{2n}) i_m}.
\end{aligned}$$

Applying this with $m = 1$ proves the lemma for $\kappa = 1$. For other values of κ we break up $|\sigma|$ into a sum over components, increasing the sum by ignoring the ordering of indices when they cross component boundaries. □

We will also want to count pairings. Let $S(n,\kappa)$ be the set of pairings on $[2n]$ with exactly κ components. For convenience, we stipulate that $S(0,0)$ contains one element, the empty pairing, and otherwise that $S(n,\kappa)$ is the empty set when n or κ is 0.

Lemma 7. *For all nonnegative n and κ, the number of pairings $S(n,\kappa)$ is $2^{n-\kappa}c(n,\kappa)$, where $c(n,\kappa)$ is the signless Stirling number of the first kind. It follows that for any positive x,*

$$\begin{aligned}\sum_{\kappa=1}^{n} \#S(n,\kappa)x^{\kappa} &= x(x+2)(x+4)\cdots(x+2n-2) \\ &\leq x^n + n(n-1)x^{n-1} + (2n-1)!!\big(x^{n-2}\vee 1\big).\end{aligned} \tag{27}$$

Proof. We claim that $\#S(n,\kappa)$ satisfies

$$\#S(n,\kappa) = \#S(n-1,\kappa-1) + 2(n-1)\#S(n-1,\kappa). \tag{28}$$

We define a bijection β between $S(n,\kappa)$ and $\big(S(n-1,\kappa)\times[2n-2]\big)\cup S(n-1,\kappa-1)$ as follows: Suppose σ is a pairing on $[2n]$ with κ components. If $\{2n-1,2n\}\in\sigma$, then $\beta(\sigma)$ is just σ with the component $\{2n-1,2n\}$ removed; it is in $S(n-1,\kappa-1)$. Otherwise, there are $m,m'\in[2n-2]$ such that $\{2n-1,m\}$ and $\{2n,m'\}$ are in σ. In this case, we define a matching σ' on $[2n-2]$ to be the same pairs as in σ, except that $\{2n-1,m\}$ and $\{2n,m'\}$ are removed, and $\{m,m'\}$ is added. The components of σ' are the same as those of σ, except for the removal of $2n-1$ and $2n$ from one component. Thus σ' is in $S(n-1,\kappa)$, and we let $\beta(\sigma)=(\sigma',m)$.

If $\sigma\in S(n-1,\kappa-1)$ then $\beta^{-1}(\sigma)=\sigma\cup\{\{n-1,n\}\}$. For $(\sigma,m)\in S(n-1,\kappa)\times[2n-2]$ we define $\beta^{-1}(\sigma,m)$ by finding m' such that $\{m,m'\}\in\sigma$, removing $\{m,m'\}$, and adding in the pairs $\{m,2n-1\}$ and $\{m',2n\}$.

The recursion (28) implies that $c(n,\kappa):=2^{\kappa-n}\#S(n,\kappa)$ satisfies

$$c(n,\kappa) = c(n-1,\kappa-1) + (n-1)c(n-1,\kappa),$$

for $n,\kappa\geq 1$, and $c(n,0)=c(0,\kappa)=0$, except $c(0,0)=1$. This is the recurrence which defines the Stirling numbers. (See, for instance, Lemma 1.3.3 of [Sta86].) The equality in (27) follows then from Proposition 1.3.4 of [Sta86]. To derive the bound, we note that the monomial x^n has coefficient 1, and x^{n-1} has coefficient $2+4+\cdots+2(n-1)=n(n-1)$. What remains are terms no bigger than $x^{n-2}\vee 1$; the total of all the coefficients is found by evaluating the function at $x=1$, yielding $(2n-1)!!$. □

4. Proof of the Theorem

We consider first the even moments, and write $2n$ in place of n. We may assume without loss of generality that $\pi(g)=0$ and $c=1$.

Let $\mathcal{I}=\{0,1,\ldots,k-1\}^{2n}$, and let $\mathcal{I}_\sigma$ be the subset of $\mathcal{I}$ consisting of multiindices whose pairing is σ. Remember that the matched multiindices $\mathcal{I}_{\sigma_0}$ are those in which the smallest and second smallest are adjacent, third and fourth, and so on.

For any ordered multiindex $I=(i_1,i_2,\ldots,i_{2n})$ we define $r(I)$ to be the number of possible orderings of I, and $r^*(I)$ the number of matched orderings of I. If I has nonoverlapping pairs, and exactly a pairs of identical indices, then

$$r(I)=\frac{(2n)!}{2^a} \quad \text{and } r^*(I)=2^{n-a}n!,$$

so $r(I)/r^*(I)=(2n-1)!!$. Also, for any multiindex $r^*(I)\leq 2^n n!$.

We aim to compare the two expressions

$$\mathrm{E}_x\big[W_k(g)^{2n}\big] = k^{-n}\sum_{i_1=0}^{k-1}\sum_{i_2=0}^{k-1}\cdots\sum_{i_{2n}=0}^{k-1}\mathrm{E}_x\big[g(X_{i_1})\cdots g(X_{i_{2n}})\big] \qquad \text{and} \tag{29}$$

$$\gamma_g^2(k)^n = k^{-n}\sum_{i_1=0}^{k-1}\sum_{i_2=0}^{k-1}\cdots\sum_{i_{2n}=0}^{k-1}\pi(T_{|i_2-i_1|})\cdots\pi(T_{|i_{2n-1}-i_{2n}|})\,. \tag{30}$$

Note that the terms in W_k^{2n} are invariant under permutations of the indices. The terms in $\gamma_g^2(k)^n$ are invariant under permutations which preserve the pairing. Thus we can rewrite these expressions as

$$\mathrm{E}_x\big[W_k(g)^{2n}\big] = k^{-n}\sum_{i_1=0}^{k-1}\sum_{i_2=i_1}^{k-1}\cdots\sum_{i_{2n}=i_{2n-1}}^{k-1} r(i_1,\dots,i_{2n})\,\mathrm{E}_x\, g(X_{i_1})\cdots g(X_{i_{2n}}) \text{ and}$$

$$\begin{aligned}\gamma_g^2(k)^n = k^{-n}&\sum_{i_1=0}^{k-1}\sum_{i_2=i_1}^{k-1}\cdots\sum_{i_{2n}=i_{2n-1}}^{k-1} r^*(i_1,\dots,i_{2n})\pi(T_{i_2-i_1})\cdots\pi(T_{i_{2n-1}-i_{2n}})\\ &+\sum_{(i_1,\dots,i_{2n})\in\mathcal{I}\setminus\mathcal{I}_{\sigma_0}}\pi(T_{|i_2-i_1|})\cdots\pi(T_{|i_{2n-1}-i_{2n}|})\,.\end{aligned}$$

This allows us to bound the difference

$$k^n\big|\mathrm{E}_x\big[W_k(g)^{2n}\big]-(2n-1)!!\gamma_g^2(k)^n\big|$$

$$\leq (2n-1)!!\sum_{(i_1,\dots,i_{2n})\in\mathcal{I}\setminus\mathcal{I}_{\sigma_0}}\pi(T_{|i_2-i_1|})\cdots\pi(T_{|i_{2n-1}-i_{2n}|}) \tag{31}$$

$$+\sum\big|r(i_1,\dots,i_{2n})-(2n-1)!!r^*(i_1,\dots,i_{2n})\big|\pi(T_{i_2-i_1})\cdots\pi(T_{i_{2n-1}-i_{2n}}) \tag{32}$$

$$+\sum r(i_1,\dots,i_{2n})\big|\mathrm{E}_x\, g(X_{i_1})\cdots g(X_{i_{2n}})-\pi(T_{i_2-i_1})\cdots\pi(T_{i_{2n-1}-i_{2n}})\big|. \tag{33}$$

The sums in (32) and (33) are taken over all $(i_1,\dots,i_{2n})$ with $0\leq i_1\leq\cdots\leq i_{2n}$.

4.1. **Bounding** (31)**.** We need to show that the contribution to $\gamma_g^2(k)^n$ by unmatched multiindices is negligible. By Lemma 1, with $\alpha=\frac{1}{2}$, and $\tilde\rho=\sqrt{\rho}$,

$$\big|\pi(T_{|i_2-i_1|})\cdots\pi(T_{|i_{2n-1}-i_{2n}|})\big|\leq\big(2\sqrt{R}\pi(V)\big)^n\prod_{\ell=1}^n\tilde\rho^{|i_{2\ell}-i_{2\ell-1}|}.$$

Let $I=(i_1,\dots,i_{2n})$ have pairing σ, with κ components, and let $J=(j_1,\dots,j_{2n})$ be the ordered version of I. Equation (25) tells us that the exponent of $\tilde\rho$ above is $|\sigma|(J)$. By Stirling's formula [Fel68],

$$\sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}e^{1/(12n+1)}\leq n!\leq\sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}e^{1/(12n)},$$

there are

$$r^*(J)\leq 2^n n!\leq 2^{-n+2}\sqrt{n}\frac{(2n)!}{n!}$$

different multiindices $I \in \mathcal{I}_\sigma$ which all yield the same J when ordered. Thus

$$\Big|\sum_{(i_1,\ldots,i_{2n})\in\mathcal{I}_\sigma} \pi(T_{i_2-i_1})\cdots\pi(T_{i_{2n-1}-i_{2n}})\Big| \le 2^n n! \sum_{j_1=0}^{k-1}\sum_{j_2=j_1}^{k-1}\cdots\sum_{j_{2n}=j_{2n-1}}^{k-1} \big(2\sqrt{R}\pi(V)\big)^n \tilde{\rho}^{|\sigma|(j_1,\ldots,j_{2n})}.$$

An application of Lemma 6 yields the bound

$$\Big|\sum_{(i_1,\ldots,i_{2n})\in\mathcal{I}_\sigma} \pi(T_{i_2-i_1})\cdots\pi(T_{i_{2n-1}-i_{2n}})\Big| \le n!\big(4\sqrt{R}\pi(V)\big)^n k^\kappa (1-\tilde{\rho})^{\kappa-2n}.$$

Using Lemma 7 to sum over all σ other than the trivial pairing, and the relation

$$1-\rho^\alpha \ge \alpha(1-\rho), \tag{34}$$

which holds for any $\rho,\alpha\in(0,1)$, we get

$$\begin{aligned}\Big|\sum_{(i_1,\ldots,i_{2n})\in\mathcal{I}\setminus\mathcal{I}_{\sigma_0}} &\pi(T_{i_2-i_1})\cdots\pi(T_{i_{2n-1}-i_{2n}})\Big| \\ &\le \frac{(2n)!}{n!}(1-\rho)^{-n-1}2^{2n+3}R_*^{3n/2}\sqrt{n} \\ &\quad\times\Big(n(n-1)k^{n-1}+(1-\rho)^{-1}(2n-1)!!!\big(k\vee 2(1-\rho)^{-1}\big)^{n-2}\Big).\end{aligned} \tag{35}$$

4.2. **Bounding** (32)**.** We know that $r(I)-(2n-1)!!r^*(I)=0$, except when I has overlapping pairs. Saying that $(i_1,\ldots,i_{2n})$ has overlapping pairs says that $i_{2m}=i_{2m+1}$ for some $m\in\{1,2,\ldots,n-1\}$. The summand in (32) is bounded by

$$(2n)!\big(2\sqrt{R}\pi(V)\big)^n\prod_{\ell=1}^n \tilde{\rho}^{(i_{2\ell}-i_{2\ell-1})}.$$

The sum may thus be bounded by

$$(2n)!\big(2\sqrt{R}\pi(V)\big)^n\sum_{m=1}^{n-1}\sum\nolimits^{(m)}\prod_{\ell=1}^n \tilde{\rho}^{i_{2\ell}-i_{2\ell-1}},$$

where $\Sigma^{(m)}$ is the sum over ordered indices $(i_1,\ldots,i_{2n})$ where $i_{2m}=i_{2m+1}$. We can apply Lemma 4 to this sum, with $x=\tilde{\rho}$, $N=2n-1$, and n of the η's being 1, the rest 0. This means that the term (32) is bounded by

$$(n-1)\big(4\sqrt{R}\pi(V)(1-\rho)^{-1}\big)^n\Big(\frac{(2n)!}{(n-1)!}k^{n-1}+(2n)!k^{n-2}\Big) \tag{36}$$

4.3. **Bounding** (33)**.** Let $(i_1,\ldots,i_{2n})$ be any ordered multiindex. We define a sequence of interpolations $Y_0,Y_1,\ldots,Y_n$ between $Y_0:=\pi(T_{i_2-i_1})\cdots\pi(T_{i_{2n}-i_{2n-1}})$ and $Y_n:=\mathrm{E}_x\big[g(X_{i_1})\cdots g(X_{i_{2n}})\big]$ by

$$Y_\ell := \mathrm{E}_x\big[g(X_{i_1})g(X_{i_2})\cdots g(X_{i_{2\ell}})\big]\pi(T_{i_{2\ell+2}-i_{2\ell+1}})\cdots\pi(T_{i_{2n}-i_{2n-1}}) \tag{37}$$

for $1\le\ell\le n-1$. For any $\ell\in\{1,2,\ldots,n\}$ we have

$$\begin{aligned}\big|Y_\ell-Y_{\ell-1}\big| = \Big|&\mathrm{E}_x\big[g(X_{i_1})g(X_{i_2})\cdots g(X_{i_{2\ell-2}})\big(\mathrm{P}^{i_{2\ell-1}-i_{2\ell-2}}_{X(i_{2\ell-2})}\big(T_{i_{2\ell}-i_{2\ell-1}}\big)-\pi(T_{i_{2\ell}-i_{2\ell-1}})\big)\big]\Big| \\ &\times\Big|\pi(T_{i_{2\ell+2}-i_{2\ell+1}})\cdots\pi(T_{i_{2n}-i_{2n-1}})\Big|\end{aligned}$$

From Lemma 1 we derive two different estimates for T_m, one with $\alpha = \frac{1}{2}$, the other with $\alpha = (2n-2\ell+1)/2n$:

$$\begin{aligned}
\left|T_m(x)\right| &\le 2\sqrt{R}V(x)\rho^{m/2},\\
\left|T_m(x)\right| &\le V(x)^{1/2n}\left|\mathrm{P}_x^m(g)\right| \qquad \text{(remember that } \pi(g)=0\text{)}\\
&\le 2R^{(2n-2\ell+1)/2n}V(x)^{(n-\ell+1)/n}\rho^{m(2n-2\ell+1)/2n}. \qquad (38)
\end{aligned}$$

The former will be used for the individual terms $\pi(T_m)$, the latter for the product. If we define

$$h(x) := \mathrm{P}_x^{i_{2\ell-1}-i_{2\ell-2}}\big(T_{i_{2\ell}-i_{2\ell-1}}\big) - \pi\big(T_{i_{2\ell}-i_{2\ell-1}}\big),$$

we have, by Lemma 1, using $\alpha = (n-\ell+1)/n$ and the bound (38)

$$\left|h(x)\right| \le 4R^{2(n-\ell+1)/n}\rho^{(i_{2\ell}-i_{2\ell-2})(2n-2\ell+1)/2n}V(x)^{(n-\ell+1)/n}.$$

Thus we may apply Lemma 3, with

$$\begin{aligned}
c_h &= 4R^{2(n-\ell+1)/n}\rho^{(2n-2\ell+1)(i_{2\ell}-i_{2\ell-2})/2n}, &\qquad \alpha_h &= \frac{n-\ell+1}{n},\\
c_g &= 1, &\text{and}\quad \alpha_g &= \frac{1}{2n}.
\end{aligned}$$

Letting $\tilde{\rho} = \rho^{1/2n}$, this gives us

$$\begin{aligned}
\left|Y_\ell - Y_{\ell-1}\right| \le (2\sqrt{R})^{n-\ell}&\left(\prod_{j=\ell+1}^{n}\rho^{(i_{2j}-i_{2j-1})/2}\right)\pi(V)^{n-\ell}\\
&\times 2^{\ell+1}R_*^{2\ell-\ell^2/n+1/n}V(x)\tilde{\rho}^{(i_{2\ell}-i_{2\ell-2})(2n-2\ell+1)}\\
&\times\prod_{j=1}^{\ell-1}\left(\tilde{\rho}^{(2n-2j+1)(i_{2j}-i_{2j-1})}+\tilde{\rho}^{(2n-2j+1)(i_{2j-1}-i_{2j-2})}\right). \qquad (39)
\end{aligned}$$

The term we are seeking to bound is the sum of all $|Y_n - Y_0|$ over all possible ordered multiindices $(i_1,\dots,i_{2n})$. We need to approximate

$$\begin{aligned}
\sum_{0\le i_1\le\cdots\le i_{2n}\le k-1}\tilde{\rho}^{(i_{2\ell}-i_{2\ell-2})(2n-2\ell+1)}&\prod_{j=\ell+1}^{n}\rho^{(i_{2j}-i_{2j-1})/2}\\
&\times\prod_{j=1}^{\ell-1}\left(\tilde{\rho}^{(2n-2j+1)(i_{2j}-i_{2j-1})}+\tilde{\rho}^{(2n-2j+1)(i_{2j-1}-i_{2j-2})}\right).
\end{aligned}$$

Performing the outer summation (over multiindices) first, we obtain a sum of $2^{\ell-1}$ terms, each of which may be bounded by an application of Lemma 4: We take $x=\rho$, $N=2n$, $s=n+1$, and

$$\begin{aligned}
\eta_j &= \frac{1}{2} \text{ for } n-\ell \text{ values of } j,\\
\eta_{2\ell} = \eta_{2\ell-1} &= \frac{2n-2\ell+1}{2n},\\
\ell-1 \text{ other nonzero values: } &\frac{2n-2\ell+3}{2n}, \frac{2n-2\ell+5}{2n}, \dots, \frac{2n-1}{2n}
\end{aligned}$$

Thus each of the terms is bounded by

$$\begin{aligned}
&\binom{k+n-2}{n-1}(1-\sqrt{\rho})^{\ell-n}\left(1-\rho^{(2n-2\ell+1)/2n}\right)^{-1}\prod_{j=1}^{\ell}\left(1-\rho^{(2n-2j+1)/2n}\right)^{-1}\\
&\qquad\le 2^{n-\ell}\binom{k+n-2}{n-1}(1-\rho)^{-n-1}\left(\frac{n}{n-\ell+1}\right)\quad\prod_{j=1}^{\ell}\frac{2n}{2n-2j+1}\\
&\qquad=\binom{k+n-2}{n-1}(1-\rho)^{-n-1}\frac{2^{n+\ell}n^{\ell+1}n!(2n-2\ell)!}{(2n)!(n-\ell+1)!}
\end{aligned}$$

Applying Stirling's formula and the estimate $1-x\le e^{-x}$ for $x\le 1$, we get the bound

$$2^{n+\ell}e^{\ell}\binom{k+n-2}{n-1}(1-\rho)^{-n-1}$$

Putting this together with the coefficients in (39), and summing over $1\le\ell\le n$ (by simply taking n times the largest term),

$$\begin{aligned}
&\sum_{i_1=0}^{k-1}\sum_{i_2=i_1}^{k-1}\cdots\sum_{i_{2n}=i_{2n-1}}^{k-1} r(i_1,\ldots,i_{2n})\big|g(X_{i_1})\cdots g(X_{i_{2n}})-\pi(T_{i_2-i_1})\cdots\pi(T_{i_{2n-1}-i_{2n}})\big|\\
&\qquad\le n(2n)!R_*^{25n/16}(2e)^{n+2}(1-\rho)^{-n-1}V(x)\Big(\frac{1}{(n-1)!}k^{n-1}+k^{n-2}\Big).
\end{aligned}\tag{40}$$

Assembling (35), (36), and (40), together with (31)–(33), we get

$$\begin{aligned}
&\big|\mathrm{E}_x\big[W_k(g)^{2n}\big]-(2n-1)!!\gamma_g^2(k)^n\big|\\
&\qquad\le k^{-1}n^2(2n)!R_*^{25n/16}(1-\rho)^{-n-1}\\
&\qquad\quad\times\bigg[(2e)^{n+2}\frac{V(x)}{(n-1)!}+4^n\frac{1}{(n-1)!}+\frac{2^{2n+3}}{(n-1)!}\\
&\qquad\qquad+k^{-1}\Big((2e)^{n+2}V(x)+4^n\\
&\qquad\qquad\qquad+2^{2n+3}(1-\rho)^{-1}\big(1\vee 2k^{-1}(1-\rho)^{-1}\big)^{n-2}\Big)\bigg]
\end{aligned}\tag{41}$$

It only remains to estimate the error arising from the substitution of $\gamma_g^2(k)$ for γ_g^2. Observe that

$$\begin{aligned}
\big|\gamma_g^2-\gamma_g^2(k)\big|&=\Big|\frac{2}{k}\pi(T_1)+\frac{4}{k}\pi(T_2)+\cdots+\frac{2k}{k}\pi(T_k)+2\pi(T_{k+1})+2\pi(T_{k+2})+\cdots\Big|\\
&\le\frac{1}{k}\cdot 16\sqrt{R}\pi(V)\big(1-\rho\big)^{-2}.
\end{aligned}$$

Thus

$$\begin{aligned}
(2n-1)!!\big|\big(\gamma_g^2\big)^n-\gamma_g^2(k)^n\big|&\le(2n-1)!!\big|\gamma_g^2-\gamma_g^2(k)\big|\cdot\sum_{j=0}^{n-1}\big(\gamma_g^2\big)^j\gamma_g^2(k)^{n-j-1}\\
&\le\frac{(2n)!}{k\,(n-1)!}\cdot 2\cdot 4^nR^{n/2}\pi(V)^n\big(1-\rho\big)^{-n-1}.
\end{aligned}$$

Adding this to (41) gives us

$$
\begin{aligned}
&\left|\mathrm{E}_x W_k^{2n} - (2n-1)!!\gamma_g^2(k)^n\right| \\
&\le k^{-1}\frac{n^2(2n)!}{(n-1)!}R_*^{25n/16}(1-\rho)^{-n-1} \\
&\qquad \times \Bigg[(2e)^{n+2}V(x) + 3\cdot 4^n + 2^{2n+3} \\
&\qquad\qquad + k^{-1}\Big((2e)^{n+2}V(x) + 4^n \\
&\qquad\qquad\qquad + 2^{2n+3}(1-\rho)^{-1}\big(1\vee 2k^{-1}(1-\rho)^{-1}\big)^{n-2}\Big)\Bigg].
\end{aligned}
\tag{42}
$$

When similar terms are combined, this reduces directly to (9)

Now we consider odd powers. Choose any ordered multiindex $0 \le i_1 \le \cdots \le i_{2n+1} \le k-1$, and define $h(x) := P_x^{i_{2n+1}-i_{2n}}(g)$. By Lemma 1,

$$
\left|h(x)\right| \le 2R^{1/(2n+1)}\rho^{(i_{2n+1}-i_{2n})/(2n+1)}V(x)^{1/(2n+1)},
$$

so we may apply Lemma 3 with $\alpha_g = \alpha_h = 1/(2n+1)$ to get

$$
\begin{aligned}
&\left|\mathrm{E}_x\big[g(X_{i_1})\cdots g(X_{i_{2n}})g(X_{i_{2n+1}})\big]\right| \\
&\qquad \le \left|\mathrm{E}_x\big[g(X_{i_1})\cdots g(X_{i_{2n}})h(X_{i_{2n}})\big]\right| \\
&\qquad \le 2^n R_*^{n+1}V(x)\rho^{(i_{2n+1}-i_{2n})/(2n+1)} \\
&\qquad\qquad \times \prod_{j=1}^{n}\left(\rho^{(i_{2j}-i_{2j-1})(2n-2j+2)/(2n+1)} + \rho^{(i_{2j-1}-i_{2j-2})(2n-2j+2)/(2n+1)}\right).
\end{aligned}
$$

Applying Lemma 4 and simplifying the constant factors, we bound this sum over all ordered multiindices by

$$
\begin{aligned}
2^n R_*^{n+1}V(x)&\left(1-\rho^{1/(2n+1)}\right)^{-1}\prod_{j=1}^{n}\big(1-\rho^{(2n-2j+2)/(2n+1)}\big)^{-1}\left(\frac{k^n}{n!}+k^{n-1}\right) \\
&\le \left(2R_*(1-\rho)^{-1}\right)^{n+1}\sqrt{n}\left(\frac{k^n}{n!}+k^{n-1}\right)
\end{aligned}
$$

Multiplying this by the $(2n+1)!$ which is the maximum number of ways that any collection of indices can be ordered, and dividing by $k^{n+\frac{1}{2}}$, we get

$$
\left|E_x W_k(g)^{2n+1}\right| \le k^{-1/2}\frac{(2n+1)!}{n!}\Big(2R_*(1-\rho)^{-1}\Big)^{n+1}\sqrt{n}\Big(1+k^{-1}n!\Big),
\tag{43}
$$

which simplifies directly to (10).

5. Proofs of the Corollaries

5.1. **Proof of Corollary 1.** For any n we have $|g-\pi(g)| \le c \le cV^{1/n}$, so Theorem 1 assures us that

$$
\begin{aligned}
&\left|\mathrm{E}_x e^{\lambda W_k(g)} - e^{\lambda^2\gamma_g^2/2}\right| \\
&\qquad \le \sum_{n=0}^{\infty}\left[\frac{\lambda^{2n}}{n!}k^{-1}C(cr)^{2n}\Big(1+\frac{n!}{k}\Big) + \frac{\lambda^{2n+1}}{n!}k^{-1/2}C'(cr')^{2n+1}\Big(1+\frac{n!}{k}\Big)\right].
\end{aligned}
$$

This reduces directly to (11).

5.2. **Proof of Corollary 2.** We begin by considering the case of $\psi(x) = \mathbf{1}\{x \le t\}$ for some real number t. Define

$$W_{j,k} := \frac{1}{\sqrt{k}}\left(g(X_j) + g(X_{j+1}) + \cdots + g(X_{j+k-1})\right),$$

and let $h_t^k(x) := \mathrm{P}_x\{W_k \le t\}$. By Theorem 1 of [Bol82], there is a constant β, depending on the chain, such that

$$\beta := \sup_{k\ge 1}\sup_{t\in\mathbb{R}} k^{1/2}\left|\pi(h_t^k) - \Phi\big(t/\gamma_g\big)\right| \tag{44}$$

is finite. Since $h_t^k \le 1$, we also have for every positive j,

$$\left|\mathrm{P}_x^j\big(h_t^k\big) - \pi\big(h_t^k\big)\right| \le R\rho^j V(x). \tag{45}$$

Putting these two bounds together, we see that

$$\left|\mathrm{P}_x^j\big(h_t^k\big) - \Phi\big(t/\gamma_g\big)\right| \le R\rho^j V(x) + \beta k^{-1/2}. \tag{46}$$

For any positive ϵ,

$$\begin{aligned}
\mathrm{P}_x\big\{t-\epsilon\gamma_g \le W_{j,k} \le t+\epsilon\gamma_g\big\} &= \mathrm{P}_x^j\big(h_{t+\gamma_g\epsilon}^k\big) - P_x^j\big(h_{t-\gamma_g\epsilon}^k\big)\\
&\le \Phi(t/\gamma_g+\epsilon) - \Phi(t/\gamma_g-\epsilon) + \left|\mathrm{P}_x^j\big(h_{t+\gamma_g\epsilon}^k\big) - \Phi(t/\gamma_g+\epsilon)\right|\\
&\qquad + \left|\mathrm{P}_x^j\big(h_{t-\gamma_g\epsilon}^k\big) - \Phi(t/\gamma_g-\epsilon)\right|\\
&\le 2\left(\epsilon + R\rho^j V(x) + \beta k^{-1/2}\right). \qquad (47)
\end{aligned}$$

This means that

$$\begin{aligned}
&\left|\mathrm{P}_x^j(h_t^k) - h_t^k(x)\right| = \left|\mathrm{P}_x\big\{W_{j,k}\le t\big\} - \mathrm{P}_x\big\{W_k \le t\big\}\right|\\
&\quad\le \mathrm{P}\big\{W_{j,k} \le t < W_k\big\} + \mathrm{P}\big\{W_k \le t < W_{j,k}\big\}\\
&\quad\le \mathrm{P}\big\{\left|W_k - W_{j,k}\right| \ge \epsilon\gamma_g\big\} + \mathrm{P}\big\{W_{j,k}\in[t-\epsilon\gamma_g, t+\epsilon\gamma_g]\big\}\\
&\quad\le \mathrm{P}\Big\{\left(\frac{j}{k}\right)^{1/2}\left|W_{0,j} - W_{k,j}\right| \ge \epsilon\gamma_g\Big\} + 2\left(\epsilon + R\rho^j V(x) + \beta k^{-1/2}\right)\\
&\quad\le \mathrm{P}\Big\{W_{0,j} \ge \frac{\epsilon\gamma_g}{2}\left(\frac{k}{j}\right)^{1/2}\Big\} + \mathrm{P}\Big\{W_{k,j} \ge \frac{\epsilon\gamma_g}{2}\left(\frac{k}{j}\right)^{1/2}\Big\}\\
&\qquad\quad + 2\left(\epsilon + R\rho^j V(x) + \beta k^{-1/2}\right).
\end{aligned}$$

Suppose that $|g-\pi(g)| \le cV^{1/2n}$. By Theorem 1, defining $M_n := \sup_x \sup_k V(x)^{-1}\,\mathrm{E}_x\,|W_k|^{2n}$,

$$\begin{aligned}
M_n &\le c^{2n}\left((2n-1)!!\big(\gamma_g^2\big)^n + 2Cr^{2n}(2n)!\right)\\
&\le (2n-1)!!c^{2n}\left[\left(\frac{8\pi(V)\sqrt{R}}{1-\rho}\right)^n + 2Cr^{2n}\right]
\end{aligned} \tag{48}$$

It follows that

(49)

$$\begin{aligned}
\left|\mathrm{P}_x^j(h_t^k) - h_t^k(x)\right| \le M_n\left[(1+R\rho^k)V(x) + \pi(V)\right]\left(\frac{j}{k}\right)^n\left(\frac{2}{\epsilon}\right)^{2n}\\
+ 2\epsilon + 2\left(R\rho^j V(x) + \beta k^{-1/2}\right).
\end{aligned}$$

Putting this together with (46), for any positive ϵ and nonnegative integer j,

$$\begin{aligned}\sup_t\bigl|\Phi(t/\gamma_g)-\mathrm{P}\{W_k\le t\}\bigr| &\le \bigl|\mathrm{P}_x^j(h_t^k)-\Phi(t)\bigr|+\bigl|\mathrm{P}_x^j(h_t^k)-h_t^k(x)\bigr|\\ &\le M_n\bigl[(1+R\rho^k)V(x)+\pi(V)\bigr]\left(\frac{j}{k}\right)^n\left(\frac{2}{\epsilon}\right)^{2n}\\ &\quad+2\epsilon+3\left(R\rho^jV(x)+\beta k^{-1/2}\right).\end{aligned}\tag{50}$$

We choose

$$\epsilon=\left(\frac{\log k}{k}\right)^{n/(2n+1)},\qquad j=\left\lfloor\frac{\log k}{-2\log\rho}\right\rfloor.$$

This bound then becomes

$$\sup_t\bigl|\Phi(t/\gamma_g)-\mathrm{P}\{W_k\le t\}\bigr| \le\left(M_n(1+2R_*)V(x)\left(\frac{2}{-\log\rho}\right)^n+2\right)\left(\frac{k}{\log k}\right)^{-n/(2n+1)}+\left(\frac{3RV(x)}{\rho}+\beta\right)k^{-1/2},$$

which simplifies to (12), with

$$M_n^*=M_n(1+2R_*)\left(\frac{2}{-\log\rho}\right)^n+2+\frac{3R}{\rho}+\beta.\tag{51}$$

If $|g-\pi(g)|$ is bounded by a constant, then by Corollary 1,

$$L_\lambda:=\sup_x\sup_k V(x)^{-1}\,\mathrm{E}_x\,e^{\lambda|W_k|}$$

is finite for any $\lambda<1/c(r\vee r')$. (The value of L_λ depends on c.) We keep the value of j from above, but now choose

$$\epsilon=\lambda^{-1}\sqrt{2}(-\log\rho)^{-1/2}k^{-1/2}\log^{3/2}k.$$

This yields

$$\begin{aligned}\sup_t\bigl|\Phi(t/\gamma_g)-\mathrm{P}\{W_k\le t\}\bigr| &\le L_\lambda(1+2R_*)V(x)\exp\bigl\{-\frac{\lambda\epsilon}{2}(k/j)^{1/2}\bigr\}+2\epsilon+3\left(R\rho^jV(x)+\beta k^{-1/2}\right)\\ &\le L_\lambda(1+2R_*)V(x)k^{-1/2}+2\sqrt{2}\lambda^{-1}(-\log\rho)^{-1/2}k^{-1/2}\log^{3/2}k+3\left(\frac{RV(x)}{\rho}+\beta\right)k^{-1/2},\end{aligned}$$

which simplifies to (13), with

$$L^*=\inf_{0<\lambda<1/c(r\vee r')}\bigl\{L_\lambda+2\sqrt{2}\lambda^{-1}(-\log\rho)^{-1/2}+\frac{3R}{\rho}+\beta\bigr\}.\tag{52}$$

Now consider any measurable function ψ with bounded linear variation If ψ is a step function, written in the form

$$\psi(t)=a_0+\sum_{i=1}^N a_i\mathbf{1}_{\{t\le t_i\}},$$

then for X a standard normal random variable,

$$\left|\mathrm{E}_x\,\psi\big(W_k(g)\big) - \mathrm{E}\,\psi(\gamma_g X)\right| \le \left(\sum_{i=1}^{N} |a_i|\right) \sup_t \left|\mathrm{P}\big\{W_k(g) \le t\big\} - \Phi\big(t/\gamma_g\big)\right|.$$

Since any ψ may be uniformly approximated arbitrarily closely by step functions with $\sum |a_i| \le \|\psi\|$, the result follows.

5.3. **Proof of Corollary 3.** Let $\phi(t)$ and $\Phi(t)$ be the standard normal density and distribution function. For any positive α and any positive b,

$$\begin{aligned}
&\left|\mathrm{E}_x\big[|W_k(g)|^\alpha\big] - \mathrm{E}\big[|\gamma_g X|^\alpha\big]\right| \\
&\quad\le \left|\mathrm{E}_x\big[|W_k(g)|^\alpha \wedge (\gamma_g b)^\alpha\big] - \gamma_g^\alpha\,\mathrm{E}\big[|X|^\alpha \wedge b^\alpha\big]\right| \\
&\qquad + \mathrm{E}_x\Big[\big(|W_k(g)|^\alpha - (\gamma_g b)^\alpha\big)^+\Big] + \gamma_g^\alpha\,\mathrm{E}_x\Big[\big(|X|^\alpha - b^\alpha\big)^+\Big]
\end{aligned} \tag{53}$$

By the standard approximation [Fel68, Lemma VII.1.2] $1 - \Phi(z) \le z^{-1}\phi(z)$, the final normal expectation is bounded by

$$\gamma_g^\alpha \int_b^\infty \alpha z^{\alpha-2}\phi(z)\,dz \le A a^\alpha \alpha^{(\alpha/2)+2}(b \vee 1)^\alpha \phi(b), \tag{54}$$

where A and a are constants independent of α and k.

Now, for any positive integer n,

$$\big|g(x) - \pi(g)\big| \le c\log^p V(x)^{1/2n} \le c^{2n}\left(\frac{(2n)p}{e}\right)^p V(x)^{1/2n}.$$

The function $|z|^\alpha \wedge b^\alpha$ has total variation bounded by $2b^\alpha$, so we can apply (12) to bound the first term in (53) by

$$M_1^* \cdot 2(\gamma_g b)^\alpha \cdot \left(\frac{2p}{e}\right)^p \left(\frac{k}{\log k}\right)^{-1/3} V(x),$$

where M_1^* is given by (51). Simplifying crudely, we find constants a and A such that this term is bounded by The second term on the right in (53) is bounded by

$$\begin{aligned}
\int_{\gamma_g b}^\infty \alpha z^{\alpha-1}\,\mathrm{P}\big\{W_k(g) \ge z\big\}\,dz &\le \int_{\gamma_g b}^\infty \alpha z^{\alpha-1} M_n \left(\frac{2np}{e}\right)^{2np} z^{-2n}\,dz \\
&\le \left(\frac{2np}{e}\right)^{2np} \frac{M_n \alpha (\gamma_g b)^{\alpha-2n}}{2n-\alpha}.
\end{aligned}$$

We take

$$n = \left\lceil \frac{\alpha}{2} + \frac{rq}{2} \right\rceil.$$

Putting all these estimates together, and using the bound (48), we find constants A', a', depending only on the Markov chain and r, such that for all $\alpha \ge 1$,

$$\begin{aligned}
&V(x)^{-1}\left|\mathrm{E}_x\big[\big|W_k(g)\big|^\alpha\big] - \mathrm{E}\big[\big|\gamma_g X\big|^\alpha\big]\right| \\
&\quad\le A'(a')^\alpha\left[\alpha^{\alpha/2} b^\alpha e^{-b^2/2} + b^\alpha \left(\frac{k}{\log k}\right)^{-1/3} + \alpha^{(\alpha+rq)(2p+1)/2} b^{-rq}\right]
\end{aligned} \tag{55}$$

Observe that for any number λ and $s \in (0,1)$, letting $\lambda' = \lambda e^{1-s}$, and applying the Stirling approximation,

$$\begin{aligned}\sum_{j=0}^{\infty} \frac{\lambda^j j^{(1-s)j}}{j!} &\le \sum_{j=0}^{\infty} \frac{(\lambda')^j}{(j!)^s} \\ &\le \sum_{j=0}^{\infty} (\lambda')^j \lfloor sj \rfloor !^{-1} \\ &\le \lceil s^{-1} \rceil (\lambda')^{s^{-1}} e^{\lambda'}.\end{aligned}$$

For any positive λ,

$$\begin{aligned}&V(x)^{-1} \Big| \mathrm{E}_x \big[\exp\{\lambda |W_k(g)|^q\} \big] - \mathrm{E}\big[\exp\{\lambda |\gamma_g X|^q\}\big] \Big| \\ &\qquad \le \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \Big| \mathrm{E}_x \big[|W_k(g)|^{qj} \big] - \mathrm{E}\big[|\gamma_g X|^{qj} \big] \Big| \\ &\qquad \le A' \sum_{j=0}^{\infty} (a')^{qj} \frac{\lambda^j}{j!} \Big[j^{qj/2} b^{qj} e^{-b^2/y} + \big(k/\log k\big)^{-1/3} b^{qj} + j^{q(j+r)(2p+1)/2} b^{-rq} \Big] \\ &\qquad \le A'' \Big[b^{2q/(2-q)} e^{y' b^q - b^2/2} + \big(k/\log k\big)^{-1/3} e^{\lambda b^q} + B b^{-qr} \Big],\end{aligned}$$

where A'', a'', y', B are constants which depend on p, q, r, and λ, as well as the Markov chain, since $q(2p+1)/2 < 1$. If $b = (\epsilon \lambda^{-1} \log k)^{1/q}$, where $\epsilon < \frac{1}{3}$, then this bound becomes $O(\log^{-r} k)$ for large k.

Acknowledgments – I am grateful to the anonymous referee, who suggested that the results in an earlier version of this paper, which were merely moment bounds, could be improved to limits.

References

[BM00] S. Balaji and S. Meyn. Multiplicative ergodic theorems and larged deviations for an irreducible markov chain. *Stochastic processes and their applications*, 90(1):123–44, 2000.

[Bol82] Erwin Bolthausen. The berry-esseen theorem for strongly mixing harris recurrent markov chains. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 1982.

[Fel68] William Feller. *An Introduction to Probability and its Applications*, volume 1. John Wiley & Sons, New York, 3rd edition, 1968.

[MK] Sean Meyn and Ioannis Kontoyiannis. Spectal theory and limit theory for geometrically ergodic markov processes,. Preprint.

[MT93] Sean Meyn and Richard Tweedie. *Markov chains and stochastic stability*. Springer Verlag, New York, Heidelberg, Berlin, 1993.

[Roy68] H. L. Royden. *Real Analysis*. MacMillan, New York, 2nd edition, 1968.

[Sta86] Richard Stanley. *Enumerative Combinatorics*, volume 1. Wadsworth & Brooks/Cole, 1986.

[Ste01] David Steinsaltz. Random logistic maps and Lyapunov exponents. *Indagationes Mathematicae*, 2001. To appear.

Department of Statistics, 367 Evans Hall, University of California, Berkeley, CA 94720, USA

E-mail address: dstein@stat.berkeley.edu