Multi-Step Dobrushin Perturbation Ceilings, Two-Sided Kemeny–Neumann Stationary Sandwiches, Reversible TV–Conductance Envelopes, and Non-Asymptotic Markov Empirical-Average Variance Bounds

Lluis Eriksson
Independent Researcher, Östersund, Sweden
editor@airr.science

November 29, 2026

Abstract

We establish four completely self-contained, non-asymptotic quantitative theorems for finite-state discrete-time ergodic Markov chains governing stationary perturbation sensitivity, reversible total-variation mixing, bottleneck conductance, and finite-horizon empirical-average variance. First, extending Dobrushin's ergodicity coefficient $\delta(P)$ via an m-step telescoping perturbation identity, we prove the explicit multi-step sensitivity ceiling $\|\tilde{\pi} - \pi\|_{\mathrm{TV}} \le \frac{\sum_{j=0}^{m-1} \delta(P^j)}{2(1-\delta(P^m))} \|E\|_\infty$ for every $m \ge 1$ with $\delta(P^m) < 1$—yielding finite, closed-form stability bounds on sparse nearest-neighbor chains where single-step Seneta bounds diverge ($\delta(P) = 1$). Second, from Kemeny's fundamental matrix $Z = (I - P + \mathbf{1}\pi^\top)^{-1}$, the group inverse $A^\# = Z - \mathbf{1}\pi^\top$, and Schweitzer's identity $\tilde{\pi}^\top - \pi^\top = \tilde{\pi}^\top E A^\#$, we prove an exact two-sided non-asymptotic linear-approximation sandwich $\frac{1}{1+r_E}\|\pi^\top E A^\#\|_1 \le \|\tilde{\pi} - \pi\|_1 \le \frac{1}{1-r_E}\|\pi^\top E A^\#\|_1$ and exact quadratic remainder $\|\tilde{\pi}^\top - (\pi^\top + \pi^\top E A^\#)\|_1 \le \frac{r_E}{1-r_E}\|\pi^\top E A^\#\|_1$ valid for every perturbation obeying $r_E := \|E A^\#\|_\infty < 1$. Third, for irreducible reversible chains, we prove a two-sided worst-case total-variation mixing sandwich $\frac{1}{2}\lambda_*^t \le \max_i \|e_i^\top P^t - \pi\|_{\mathrm{TV}} \le \frac{1}{2}\lambda_*^t \sqrt{(1-\pi_{\min})/\pi_{\min}}$ valid for every $t \ge 0$, alongside an elementary, gap-free proof of Cheeger's conductance envelope $\Phi_*^2/2 \le 1 - \lambda_2 \le 2\Phi_*$. Fourth, from an exact finite-T spectral variance decomposition of $\bar{f}_T = \frac{1}{T}\sum_{t=0}^{T-1} f(X_t)$, we prove a non-lazy universal variance ceiling $\mathrm{Var}_\pi(\bar{f}_T) \le \frac{\sigma^2}{T}\frac{1+\lambda_2^+}{1-\lambda_2^+}$ (valid even with negative eigenvalues $\lambda_k < 0$), an exact two-sided finite-sample bias sandwich $\sigma_\infty^2(f) - \frac{2\lambda_2\sigma^2}{T(1-\lambda_2)^2} \le T\,\mathrm{Var}_\pi(\bar{f}_T) \le \sigma_\infty^2(f)$ for positive-semidefinite chains, and warm-start non-stationary transfer. All four theorems are verified across 143 deterministic numerical checks.

1 Introduction and Relation to Prior Work

Quantitative perturbation stability, finite-time mixing, and empirical-average variance control are central pillars of finite-state Markov chain theory and Markov chain Monte Carlo (MCMC) analysis [3, 1]. Classical sensitivity theory provides single-step Dobrushin/Seneta bounds [5] and fundamental-matrix condition numbers [4, 1], while reversible spectral theory connects total-variation convergence, Cheeger bottleneck conductance [2, 6], and asymptotic Kipnis–Varadhan variance [3].

Here we present four self-contained, non-asymptotic theorems with sharp, explicit constants: (i) multi-step Dobrushin perturbation ceilings finite even when $\delta(P) = 1$ (Theorem 2.1); (ii) Cho–

Meyer span conditioning paired with an exact two-sided non-asymptotic Kemeny–Neumann linear-approximation sandwich and quadratic remainder for every $r_E = \|EA^{\#}\|_\infty < 1$ (Theorem 3.1); (iii) two-sided worst-case TV mixing envelopes valid from $t = 0$ and an airtight median-shifted proof of Cheeger's conductance inequality (Theorem 4.1); and (iv) exact finite-T spectral variance identities, a non-lazy alternating-series variance ceiling requiring no positive-semidefiniteness assumption, and an exact two-sided finite-sample bias sandwich for PSD chains (Theorem 5.1).

Throughout, $P \in \mathbb{R}^{n\times n}$ ($n \geq 2$) is a row-stochastic matrix, probability distributions $\mu, \pi \in \Delta_n$ are viewed as row vectors, $\|\mu - \nu\|_{\mathrm{TV}} := \frac{1}{2}\|\mu - \nu\|_1$, and $\|M\|_\infty := \max_{1\leq i\leq n} \sum_{j=1}^n |M_{ij}|$ is the induced row-sum norm (satisfying $\|\nu M\|_1 \leq \|\nu\|_1 \|M\|_\infty$ for every row vector $\nu \in \mathbb{R}^n$).

2 Multi-Step Dobrushin Contraction and Stationary Perturbation Ceilings

Define the Dobrushin ergodicity coefficient $\delta(P) := \frac{1}{2}\max_{1\leq i,j\leq n} \|P_{i,\cdot} - P_{j,\cdot}\|_1 \in [0, 1]$, with convention $\delta(P^0) = \delta(I) = 1$.

Theorem 2.1 (Multi-Step Dobrushin Contraction and Telescoping Stationary Perturbation)**.** *Let $P, \widetilde{P} = P + E \in \mathbb{R}^{n\times n}$ ($n \geq 2$) be row-stochastic matrices (so $E\mathbf{1} = 0$).*

(a) (TV Contraction and Submultiplicativity.) *For every signed row vector $\nu \in \mathbb{R}^n$ with $\nu\mathbf{1} = 0$, $\|\nu P\|_1 \leq \delta(P)\|\nu\|_1$. Consequently, $\delta(P_1P_2) \leq \delta(P_1)\delta(P_2)$ for all stochastic P_1, P_2. If $\delta(P^m) < 1$ for some integer $m \geq 1$, then P admits a unique stationary distribution $\pi \in \Delta_n$ and $\sup_{\mu\in\Delta_n} \|\mu P^t - \pi\|_{\mathrm{TV}} \leq \delta(P^m)^{\lfloor t/m\rfloor}$ for every $t \geq 0$.*

(b) (Multi-Step Telescoping Perturbation Ceiling.) *Suppose $\delta(P^m) < 1$ for some integer $m \geq 1$, and let $\tilde{\pi} \in \Delta_n$ be any stationary distribution of $\widetilde{P} = P + E$. Then*

$$\|\tilde{\pi} - \pi\|_{\mathrm{TV}} \leq \frac{\sum_{j=0}^{m-1} \delta(P^j)}{1 - \delta(P^m)} \|\tilde{\pi}E\|_{\mathrm{TV}} \leq \frac{\sum_{j=0}^{m-1} \delta(P^j)}{2(1 - \delta(P^m))} \|E\|_\infty \leq \frac{m}{2(1 - \delta(P^m))} \|E\|_\infty. \quad (1)$$

Proof. Part (a). Write $\nu = \nu_+ - \nu_-$ with $\nu_+, \nu_- \geq 0$ and $\sum_i \nu_+(i) = \sum_j \nu_-(j) = \frac{1}{2}\|\nu\|_1 =: M$. If $M = 0$ the claim is trivial; if $M > 0$, then $\nu P = M \sum_{i,j} \frac{\nu_+(i)\nu_-(j)}{M^2}(P_{i,\cdot} - P_{j,\cdot})$, so by the triangle inequality $\|\nu P\|_1 \leq M \max_{i,j} \|P_{i,\cdot} - P_{j,\cdot}\|_1 = \delta(P)\|\nu\|_1$. Applying this to $\nu = (P_1)_{i,\cdot} - (P_1)_{j,\cdot}$ gives $\delta(P_1P_2) \leq \delta(P_1)\delta(P_2)$. When $\delta(P^m) < 1$, P^m is a strict contraction on the complete metric space $(\Delta_n, \|\cdot\|_{\mathrm{TV}})$, giving a unique fixed point π of P^m; since $(\pi P)P^m = \pi P$, uniqueness forces $\pi P = \pi$. Writing $t = qm + r$ ($0 \leq r < m$) and using $\delta(P^r) \leq 1$ gives $\|\mu P^t - \pi\|_{\mathrm{TV}} = \|(\mu P^r - \pi)P^{qm}\|_{\mathrm{TV}} \leq \delta(P^m)^q$.

Part (b). Because $\tilde{\pi}\widetilde{P}^m = \tilde{\pi}$ and $\pi P^m = \pi$,

$$\tilde{\pi} - \pi = \tilde{\pi}(\widetilde{P}^m - P^m) + (\tilde{\pi} - \pi)P^m.$$

Taking TV norms and applying Part (a) yields $(1 - \delta(P^m))\|\tilde{\pi} - \pi\|_{\mathrm{TV}} \leq \|\tilde{\pi}(\widetilde{P}^m - P^m)\|_{\mathrm{TV}}$. Telescoping $\widetilde{P}^m - P^m = \sum_{r=0}^{m-1} \widetilde{P}^r E P^{m-1-r}$, using $\tilde{\pi}\widetilde{P}^r = \tilde{\pi}$, and noting $(\tilde{\pi}E)\mathbf{1} = 0$ gives

$$\|\tilde{\pi}(\widetilde{P}^m - P^m)\|_{\mathrm{TV}} \leq \sum_{r=0}^{m-1} \|\tilde{\pi}EP^{m-1-r}\|_{\mathrm{TV}} \leq \|\tilde{\pi}E\|_{\mathrm{TV}} \sum_{j=0}^{m-1} \delta(P^j).$$

Finally, $\|\tilde{\pi}E\|_{\mathrm{TV}} = \frac{1}{2}\|\tilde{\pi}E\|_1 \leq \frac{1}{2}\|\tilde{\pi}\|_1\|E\|_\infty = \frac{1}{2}\|E\|_\infty$ and $\delta(P^j) \leq 1$, proving (1) [5]. □

3 Kemeny Fundamental Matrix and Two-Sided Neumann Stationary Sandwich

Let $P \in \mathbb{R}^{n\times n}$ be an irreducible stochastic matrix with unique stationary distribution $\pi > 0$ ($\Pi := \mathbf{1}\pi^\top$), so $\ker(I-P) = \operatorname{span}(\mathbf{1})$.

Theorem 3.1 (Kemeny–Schweitzer Conditioning and Two-Sided Neumann Sensitivity Sandwich)**.** (a) (Fundamental Matrix, Group Inverse, and Cho–Meyer Span Bound.) *The matrix $I-P+\Pi$ is nonsingular. Its inverse $Z := (I-P+\Pi)^{-1}$ and group inverse $A^\# := Z-\Pi$ obey $Z\mathbf{1} = \mathbf{1}$, $\pi Z = \pi$, $(I-P)Z = Z(I-P) = I-\Pi$, $A^\#\mathbf{1} = 0$, and $\pi A^\# = 0$. For any perturbed stochastic matrix $\widetilde{P} = P+E$ with stationary distribution $\tilde{\pi}$,*

$$\tilde{\pi}-\pi = \tilde{\pi}EA^\# = \tilde{\pi}EZ, \qquad \|\tilde{\pi}-\pi\|_1 \le \kappa_{\mathrm{span}}(Z)\,\|E\|_\infty, \tag{2}$$

where $\kappa_{\mathrm{span}}(Z) := \frac{1}{2}\max_{1\le i,j\le n}\|Z_{i,\cdot}-Z_{j,\cdot}\|_1$.

(b) (Exact Two-Sided Non-Asymptotic Linear-Approximation Sandwich and Quadratic Remainder.) *Define $r_E := \|EA^\#\|_\infty$. Whenever $r_E < 1$, the first-order linear shift $\pi EA^\#$ brackets the exact stationary shift from both sides and satisfies the explicit remainder bound*

$$\frac{1}{1+r_E}\,\|\pi EA^\#\|_1 \le \|\tilde{\pi}-\pi\|_1 \le \frac{1}{1-r_E}\,\|\pi EA^\#\|_1, \qquad \big\|\tilde{\pi}-(\pi+\pi EA^\#)\big\|_1 \le \frac{r_E}{1-r_E}\,\|\pi EA^\#\|_1. \tag{3}$$

Proof. Part (a). If $(I-P+\Pi)x = 0$, left-multiplying by π yields $0 = \pi(I-P)x+(\pi\mathbf{1})(\pi x) = \pi x$, whence $(I-P)x = 0$; by irreducibility, $x = c\mathbf{1}$, and $0 = \pi x = c$ forces $x = 0$, proving invertibility of $I-P+\Pi$. Direct multiplication of $(I-P+\Pi)\mathbf{1} = \mathbf{1}$ and $\pi(I-P+\Pi) = \pi$ gives $Z\mathbf{1} = \mathbf{1}$ and $\pi Z = \pi$, hence $(I-P)Z = (I-P+\Pi)Z-\Pi Z = I-\Pi$ (and symmetrically $Z(I-P) = I-\Pi$), $A^\#\mathbf{1} = 0$, and $\pi A^\# = 0$. For $\widetilde{P} = P+E$, $(\tilde{\pi}-\pi)(I-P+\Pi) = \tilde{\pi}(I-P)+(\tilde{\pi}\mathbf{1}-\pi\mathbf{1})\pi = \tilde{\pi}(\widetilde{P}-P) = \tilde{\pi}E$; right-multiplying by Z and using $E\Pi = (E\mathbf{1})\pi = 0$ yields $\tilde{\pi}-\pi = \tilde{\pi}EZ = \tilde{\pi}EA^\#$ [4]. Since $(\tilde{\pi}E)\mathbf{1} = 0$, decomposing $\nu := \tilde{\pi}E = \nu_+-\nu_-$ exactly as in Theorem 2.1(a) gives $\|\nu Z\|_1 \le \kappa_{\mathrm{span}}(Z)\|\nu\|_1 \le \kappa_{\mathrm{span}}(Z)\|E\|_\infty$ [1].

Part (b). Rewriting $\tilde{\pi}-\pi = \tilde{\pi}EA^\#$ as

$$\tilde{\pi}-\pi = \pi EA^\#+(\tilde{\pi}-\pi)EA^\#, \tag{4}$$

the induced norm property gives $\|(\tilde{\pi}-\pi)EA^\#\|_1 \le \|\tilde{\pi}-\pi\|_1\|EA^\#\|_\infty = r_E\|\tilde{\pi}-\pi\|_1$. Applying the triangle inequality to (4) yields both $\|\tilde{\pi}-\pi\|_1 \le \|\pi EA^\#\|_1+r_E\|\tilde{\pi}-\pi\|_1$ (so $\|\tilde{\pi}-\pi\|_1 \le \frac{1}{1-r_E}\|\pi EA^\#\|_1$) and $\|\pi EA^\#\|_1 \le (1+r_E)\|\tilde{\pi}-\pi\|_1$ (so $\|\tilde{\pi}-\pi\|_1 \ge \frac{1}{1+r_E}\|\pi EA^\#\|_1$). Finally, subtracting $\pi EA^\#$ in (4) gives $\|\tilde{\pi}-(\pi+\pi EA^\#)\|_1 = \|(\tilde{\pi}-\pi)EA^\#\|_1 \le r_E\|\tilde{\pi}-\pi\|_1 \le \frac{r_E}{1-r_E}\|\pi EA^\#\|_1$, proving (3). □

4 Reversible Two-Sided TV Mixing and Cheeger Conductance Envelope

Suppose P is irreducible and reversible with respect to $\pi > 0$ ($\pi_i P_{ij} = \pi_j P_{ji}$). Then $S := D_\pi^{1/2}PD_\pi^{-1/2}$ ($D_\pi := \operatorname{diag}(\pi)$) is real symmetric with eigenvalues $1 = \lambda_1 > \lambda_2 \ge \cdots \ge \lambda_n \ge -1$, absolute second eigenvalue $\lambda_* := \max(|\lambda_2|, |\lambda_n|)$, and π-inner product $\langle f, g\rangle_\pi := \sum_{i=1}^n \pi_i f_i g_i$ with Dirichlet form $\mathcal{E}(f,f) := \langle f, (I-P)f\rangle_\pi = \frac{1}{2}\sum_{i,j}\pi_i P_{ij}(f_i-f_j)^2$. Define the bottleneck conductance $\Phi_* := \min_{0<\pi(S)\le 1/2}\frac{Q(S,S^c)}{\pi(S)}$, where $Q(S,S^c) := \sum_{i\in S,\, j\notin S}\pi_i P_{ij}$.

Theorem 4.1 (Two-Sided Reversible TV Mixing Sandwich and Cheeger Conductance Envelope)**.** *Let P be an irreducible reversible stochastic matrix on $n \ge 2$ states with stationary distribution $\pi > 0$ and $\pi_{\min} := \min_i \pi_i$.*

(a) (Two-Sided Worst-Case TV Mixing Envelope.) *For every initial distribution $\mu \in \Delta_n$ and every $t \geq 0$, letting $\chi^2(\mu\|\pi) := \sum_{i=1}^n (\mu_i - \pi_i)^2/\pi_i$,*

$$\|\mu P^t - \pi\|_{\mathrm{TV}} \leq \frac{1}{2}\lambda_*^t \sqrt{\chi^2(\mu\|\pi)}, \tag{5}$$

and the worst-case total-variation distance $d(t) := \max_{1\leq i\leq n} \|e_i^\top P^t - \pi\|_{\mathrm{TV}}$ obeys the universal two-sided sandwich

$$\frac{1}{2}\lambda_*^t \leq d(t) \leq \frac{1}{2}\lambda_*^t \sqrt{\frac{1-\pi_{\min}}{\pi_{\min}}} \qquad (\forall t \geq 0). \tag{6}$$

(b) (Cheeger Conductance Two-Sided Spectral-Gap Envelope.)

$$\frac{\Phi_*^2}{2} \leq 1 - \lambda_2 \leq 2\Phi_*. \tag{7}$$

Proof. Part (a). Let $v := D_\pi^{-1/2}(\mu^\top - \pi^\top) \in \mathbb{R}^n$. Then $v \perp D_\pi^{1/2}\mathbf{1}$ (the top eigenvector of S) and $\|v\|_2^2 = \chi^2(\mu\|\pi)$. Because S is symmetric with $\|S|_{(D_\pi^{1/2}\mathbf{1})^\perp}\|_2 = \lambda_*$ and $D_\pi^{-1/2}((\mu P^t)^\top - \pi^\top) = S^t v$, Cauchy–Schwarz gives

$$2\|\mu P^t - \pi\|_{\mathrm{TV}} = \sum_{i=1}^n \sqrt{\pi_i}\,\frac{|(\mu P^t)_i - \pi_i|}{\sqrt{\pi_i}} \leq \|S^t v\|_2 \leq \lambda_*^t \sqrt{\chi^2(\mu\|\pi)},$$

proving (5). For $\mu = e_i^\top$, $\chi^2(e_i^\top\|\pi) = \frac{1-\pi_i}{\pi_i} \leq \frac{1-\pi_{\min}}{\pi_{\min}}$, proving the upper bound in (6). For the lower bound, let $f \in \mathbb{R}^n \setminus \{0\}$ be a right eigenvector of P for the eigenvalue λ with $|\lambda| = \lambda_*$, normalized to $\|f\|_\infty = 1$, and pick i with $|f_i| = 1$. Since $\pi f = 0$ and $P^t f = \lambda^t f$,

$$\lambda_*^t = |\lambda^t f_i| = |(P^t f)_i - \pi f| = \left|\sum_{j=1}^n (P_{ij}^t - \pi_j) f_j\right| \leq \|f\|_\infty \sum_{j=1}^n |P_{ij}^t - \pi_j| = 2\|e_i^\top P^t - \pi\|_{\mathrm{TV}} \leq 2d(t),$$

proving (6) [3] (note at $t = 0$, $d(0) = 1 - \pi_{\min} \in [\frac{1}{2}, \frac{1}{2}\sqrt{(1-\pi_{\min})/\pi_{\min}}]$ since $\pi_{\min} \leq \frac{1}{2}$).

Part (b). For the upper bound, let S attain Φ_* ($0 < \pi(S) \leq 1/2$) and set $f := \mathbf{1}_S - \pi(S)\mathbf{1}$, so $\pi f = 0$, $\mathrm{Var}_\pi(f) = \pi(S)(1-\pi(S)) \geq \frac{1}{2}\pi(S)$, and $\mathcal{E}(f,f) = Q(S, S^c)$, giving $1 - \lambda_2 \leq \mathcal{E}(f,f)/\mathrm{Var}_\pi(f) \leq 2\Phi_*$. For the lower bound [2, 6], let $f \in \mathbb{R}^n \setminus \{0\}$ satisfy $Pf = \lambda_2 f$ and $\pi f = 0$. Pick a π-median $m \in \mathbb{R}$ with $\pi(\{f > m\}) \leq 1/2$ and $\pi(\{f < m\}) \leq 1/2$; replacing f by $-f$ if necessary, assume $m \leq 0$. Set $g := (f - m)_+ \geq 0$; since $\pi f = 0 \geq m$ and $f \not\equiv 0$, g is not identically zero, and $\pi(\{g > 0\}) \leq 1/2$. For every state i with $g_i > 0$, using $g_j \geq f_j - m$ for all j and $m \leq 0$ gives

$$((I-P)g)_i = (f_i - m) - \sum_{j=1}^n P_{ij} g_j \leq (f_i - m) - \sum_{j=1}^n P_{ij}(f_j - m) = ((I-P)f)_i = (1-\lambda_2)(g_i + m) \leq (1-\lambda_2) g_i.$$

Multiplying by $\pi_i g_i \geq 0$ and summing over i yields $\mathcal{E}(g,g) = \langle g, (I-P)g\rangle_\pi \leq (1-\lambda_2)\langle g, g\rangle_\pi$. Relabel states so $g_1 \geq g_2 \geq \cdots \geq g_n \geq 0$ and let $Q_{ij} := \pi_i P_{ij}$ and $A := \sum_{g_i > g_j} Q_{ij}(g_i^2 - g_j^2)$. Writing $g_i^2 - g_j^2 = \sum_{k=i}^{j-1}(g_k^2 - g_{k+1}^2)$ and $S_k := \{1, \ldots, k\}$ (where $\pi(S_k) \leq \pi(\{g > 0\}) \leq 1/2$ whenever $g_k > 0$), coarea summation yields

$$A = \sum_{k: g_k > 0} (g_k^2 - g_{k+1}^2)\, Q(S_k, S_k^c) \geq \Phi_* \sum_{k: g_k > 0} (g_k^2 - g_{k+1}^2)\pi(S_k) = \Phi_* \langle g, g\rangle_\pi.$$

Conversely, factoring $g_i^2 - g_j^2 = (g_i - g_j)(g_i + g_j)$, Cauchy–Schwarz and $(g_i + g_j)^2 \leq 2(g_i^2 + g_j^2)$ give

$$A^2 \leq \left(\sum_{g_i > g_j} Q_{ij}(g_i - g_j)^2\right)\left(\sum_{g_i > g_j} Q_{ij}(g_i + g_j)^2\right) \leq \mathcal{E}(g,g) \cdot \sum_{i,j=1}^n Q_{ij}(g_i^2 + g_j^2) = 2\mathcal{E}(g,g)\langle g, g\rangle_\pi.$$

Combining gives $\Phi_*^2 \langle g, g\rangle_\pi^2 \leq 2\mathcal{E}(g,g)\langle g, g\rangle_\pi \leq 2(1-\lambda_2)\langle g, g\rangle_\pi^2$, proving (7). □

5 Non-Asymptotic Reversible Empirical-Average Variance Bounds

Let $(X_0, X_1, \dots, X_{T-1})$ be a stationary reversible Markov chain ($X_0 \sim \pi$) with $L^2(\pi)$-orthonormal right eigenbasis $\{\psi_k\}_{k=1}^n$ ($\psi_1 = \mathbf{1}$, $P\psi_k = \lambda_k\psi_k$). For $f : [n] \to \mathbb{R}$ with $\mathbb{E}_\pi[f] = 0$ and $\sigma^2 := \mathrm{Var}_\pi(f) = \langle f, f\rangle_\pi$, define the empirical average $\bar{f}_T := \frac{1}{T}\sum_{t=0}^{T-1} f(X_t)$ and asymptotic (Kipnis–Varadhan) variance $\sigma^2_\infty(f) := \langle f, (2Z - I)f\rangle_\pi$.

Theorem 5.1 (Exact Finite-T Variance Identity, Non-Lazy Ceiling, and PSD Bias Sandwich)**.** *Let P be irreducible and reversible with $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n \geq -1$, and write $f = \sum_{k=2}^n c_k\psi_k$ ($\sum_{k=2}^n c_k^2 = \sigma^2$).*

(a) (Exact Finite-T Spectral Variance Decomposition.) *For every integer $T \geq 1$,*

$$T\ \mathrm{Var}_\pi(\bar{f}_T) = \sum_{k=2}^n c_k^2\, g_T(\lambda_k) = \sigma^2_\infty(f) - \frac{2}{T}\sum_{k=2}^n c_k^2\, \frac{\lambda_k(1-\lambda_k^T)}{(1-\lambda_k)^2}, \tag{8}$$

where $g_T(\lambda) := 1 + 2\sum_{h=1}^{T-1}(1-\frac{h}{T})\lambda^h = \frac{1+\lambda}{1-\lambda} - \frac{2\lambda(1-\lambda^T)}{T(1-\lambda)^2} \geq 0$ and $\sigma^2_\infty(f) = \sum_{k=2}^n c_k^2\frac{1+\lambda_k}{1-\lambda_k}$.

(b) (Universal Non-Lazy Variance Ceiling, PSD Bias Sandwich, and Warm-Start Transfer.) *Setting $\lambda_2^+ := \max(\lambda_2, 0) \in [0, 1)$, for every $T \geq 1$ (without requiring laziness or $\lambda_n \geq 0$),*

$$\mathrm{Var}_\pi(\bar{f}_T) \leq \frac{\sigma^2}{T}\,\frac{1+\lambda_2^+}{1-\lambda_2^+}, \tag{9}$$

and for any non-stationary initial distribution $\mu \in \Delta_n$, $\mathbb{E}_\mu[(\bar{f}_T)^2] \leq \|\mu/\pi\|_\infty\,\mathrm{Var}_\pi(\bar{f}_T) \leq \frac{1}{\pi_{\min}}\frac{\sigma^2}{T}\frac{1+\lambda_2^+}{1-\lambda_2^+}$. Moreover, if P is positive semidefinite ($\lambda_n \geq 0$, e.g. any lazy chain $\frac{1}{2}(I + P_0)$), then for every $T \geq 1$,

$$\sigma^2_\infty(f) - \frac{2\lambda_2\,\sigma^2}{T(1-\lambda_2)^2} \leq T\ \mathrm{Var}_\pi(\bar{f}_T) \leq \sigma^2_\infty(f) \leq \frac{1+\lambda_2}{1-\lambda_2}\,\sigma^2. \tag{10}$$

Proof. Part (a). By stationarity and reversibility, $\mathbb{E}_\pi[f(X_s)f(X_r)] = \langle f, P^{|s-r|}f\rangle_\pi = \sum_{k=2}^n c_k^2\lambda_k^{|s-r|}$, so

$$T\ \mathrm{Var}_\pi(\bar{f}_T) = \frac{1}{T}\sum_{s=0}^{T-1}\sum_{r=0}^{T-1}\mathbb{E}_\pi[f(X_s)f(X_r)] = \sum_{k=2}^n c_k^2\left(1 + 2\sum_{h=1}^{T-1}\left(1-\frac{h}{T}\right)\lambda_k^h\right).$$

For any $\lambda \in [-1, 1)$, differentiating the geometric sum $\sum_{h=1}^{T-1}\lambda^h = \frac{\lambda(1-\lambda^{T-1})}{1-\lambda}$ gives $\sum_{h=1}^{T-1} h\lambda^h = \frac{\lambda(1-T\lambda^{T-1}+(T-1)\lambda^T)}{(1-\lambda)^2}$, whence $1 + 2\sum_{h=1}^{T-1}(1-\frac{h}{T})\lambda^h = \frac{1+\lambda}{1-\lambda} - \frac{2\lambda(1-\lambda^T)}{T(1-\lambda)^2}$ identically (including at $\lambda = -1$). Because $Z\psi_k = (1-\lambda_k)^{-1}\psi_k$ for $k \geq 2$ (from $(I - P + \Pi)\psi_k = (1-\lambda_k)\psi_k$), we have $\langle f, (2Z - I)f\rangle_\pi = \sum_{k=2}^n c_k^2(\frac{2}{1-\lambda_k} - 1) = \sum_{k=2}^n c_k^2\frac{1+\lambda_k}{1-\lambda_k} = \sigma^2_\infty(f)$, proving (8).

Part (b). For $\lambda \in [-1, 0]$, the sequence $a_h := (1-\frac{h}{T})|\lambda|^h$ ($1 \leq h \leq T-1$) is nonnegative and nonincreasing, so the alternating sum $\sum_{h=1}^{T-1}(-1)^h a_h = -(a_1 - a_2) - (a_3 - a_4) - \cdots \leq 0$, giving $g_T(\lambda) \leq 1 \leq \frac{1+\lambda_2^+}{1-\lambda_2^+}$. For $\lambda \in [0, \lambda_2^+]$, $g_T(\lambda) \leq 1 + 2\sum_{h=1}^\infty \lambda^h = \frac{1+\lambda}{1-\lambda} \leq \frac{1+\lambda_2^+}{1-\lambda_2^+}$. Summing over $k = 2, \dots, n$ proves (9). For any $\mu \in \Delta_n$, conditioning on X_0 yields $\mathbb{E}_\mu[(\bar{f}_T)^2] = \sum_i \mu_i\mathbb{E}_i[(\bar{f}_T)^2] \leq \max_i(\mu_i/\pi_i)\sum_i \pi_i\mathbb{E}_i[(\bar{f}_T)^2] = \|\mu/\pi\|_\infty\,\mathrm{Var}_\pi(\bar{f}_T)$. Finally, if $\lambda_k \in [0, \lambda_2]$ for all $k \geq 2$, then $0 \leq \frac{\lambda_k(1-\lambda_k^T)}{(1-\lambda_k)^2} \leq \frac{\lambda_2}{(1-\lambda_2)^2}$ (since $\lambda \mapsto \lambda/(1-\lambda)^2$ is strictly increasing on $[0, 1)$), so substituting into (8) immediately proves (10). □

6 Computational Verification and Numerical Illustrations

All four theorems were verified deterministically in IEEE-754 `float64` arithmetic via the MIT-licensed companion repository `lluiseriksson/markov-perturbation-variance-geometry` (tag `v1.0.0`). Across 143 structured benchmark checks (plus 4 randomized pytest suites with seeds `20261129`–`20261202`)—covering an 8-state sparse nearest-neighbor ring walk ($\delta(P) = 1$, $\delta(P^4) < 1$) perturbed up to $r_E = 0.780$, exact brute-force Cheeger conductance ($\Phi_* = 0.1839$, $1 - \lambda_2 = 0.2258$) and 25-step two-sided TV mixing on 8-state reversible chains, and both non-lazy ($\lambda_n < 0$) and lazy PSD finite-T variance up to $T = 100$—Table 1 and Figures 1–2 confirm zero violations.

Table 1: Deterministic verification across 143 checks (`pytest -q`: 4/4 passed in 0.16 s). All identities hold to machine precision (10^{-12}) and all two-sided envelopes hold with zero violations.

Theorem	Checks	Regime Verified	Tightest Observed Metric / Ratio
Thm. 2.1 (Multi-Step Dobrushin)	24	8-state ring ($\delta(P) = 1$, $m = 4$)	Finite 4-step ceiling across $r_E \le 0.780$; 0 violations
Thm. 3.1 (Kemeny–Neumann)	48	$r_E = \|EA^{\#}\|_\infty \in [0.02, 0.78]$	max(obs/hi) = 0.9558, min(obs/lo) = 1.0106, rem. 0.5871
Thm. 4.1 (TV & Cheeger)	51	Reversible $n \in \{6, 8, 10\}$, $t \le 24$	$\Phi_*^2/2 = 0.0169 \le 1-\lambda_2 = 0.2258 \le 0.3677 = 2\Phi_*$
Thm. 5.1 (Finite-T Variance)	20	Non-lazy ($\lambda_n < 0$) & PSD ($T \le 100$)	Exact to 10^{-12}; $T\,\mathrm{Var} = 4.164 \to \sigma_\infty^2 = 4.506$

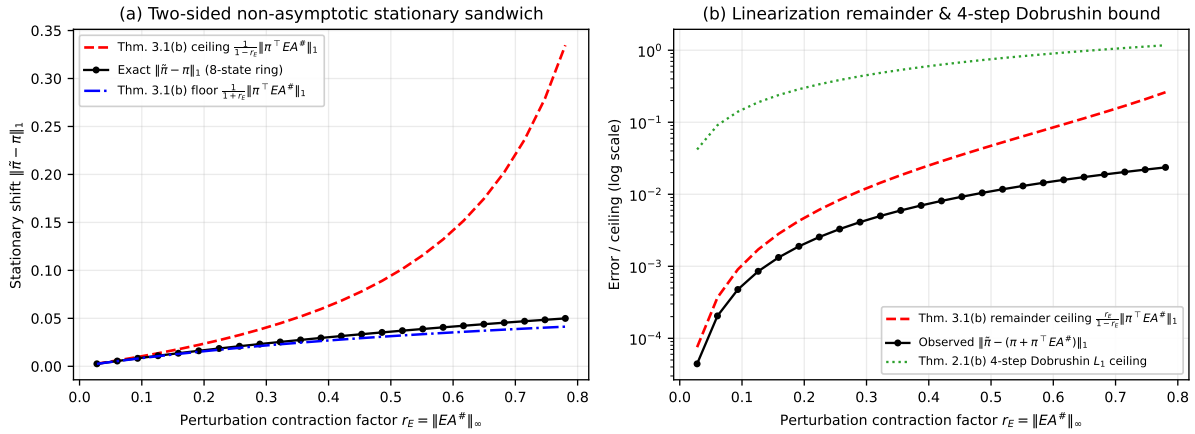

Figure 1: **(a)** Exact stationary shift $\|\tilde{\pi} - \pi\|_1$ on the 8-state ring walk tightly bracketed by the two-sided Kemeny–Neumann sandwich $\frac{1}{1\pm r_E}\|\pi EA^{\#}\|_1$ (3) up to $r_E = 0.78$. **(b)** Quadratic linearization remainder $\|\tilde{\pi} - (\pi + \pi EA^{\#})\|_1$ against $\frac{r_E}{1-r_E}\|\pi EA^{\#}\|_1$ and the 4-step Dobrushin L_1 ceiling (1).

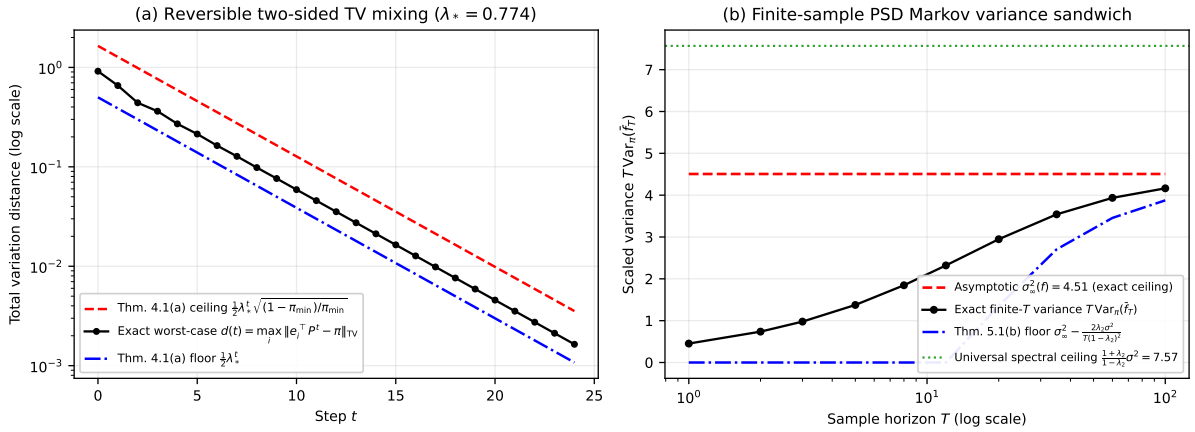

Figure 2: **(a)** Worst-case total-variation distance $d(t)$ on an 8-state reversible chain ($\lambda_* = 0.774$) bracketed by the two-sided mixing sandwich (6). **(b)** Exact finite-sample scaled variance $T\ \mathrm{Var}_\pi(\bar{f}_T)$ on the lazy PSD chain enclosed between $\sigma_\infty^2(f) - \frac{2\lambda_2\sigma^2}{T(1-\lambda_2)^2}$ and $\sigma_\infty^2(f)$ (10).

7 Scope and Limitations

Three structural boundaries merit explicit note:

1. *Non-reversible spectral mixing.* Theorems 4.1 and 5.1 rely on reversibility ($\pi_i P_{ij} = \pi_j P_{ji}$) for real orthogonal diagonalization of $D_\pi^{1/2} P D_\pi^{-1/2}$; non-reversible chains require singular-value analysis of the multiplicative reversibilization P^*P or pseudospectral bounds [3].

2. *Exponential tail concentration vs. second-moment bounds.* Theorem 5.1 yields sharp non-asymptotic variance and Chebyshev inequalities; sub-Gaussian exponential tail bounds for $\bar{f}_T$ require Markov Bernstein/Chernoff inequalities via Kato perturbation of tilted operators.

3. *Large perturbations ($r_E \geq 1$).* The two-sided Neumann sandwich (3) requires $r_E = \|EA^\#\|_\infty < 1$; for arbitrarily large perturbations $r_E \geq 1$, the unconditional Cho–Meyer span bound (2) and multi-step Dobrushin bound (1) remain valid globally.

References

[1] G. E. Cho and C. D. Meyer, *Comparison of perturbation bounds for the stationary distribution of a Markov chain*, Linear Algebra Appl. **335** (2001), 137–150.

[2] G. F. Lawler and A. D. Sokal, *Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality*, Trans. Amer. Math. Soc. **309** (1988), 557–580.

[3] D. A. Levin and Y. Peres, *Markov Chains and Mixing Times*, 2nd ed., American Mathematical Society, Providence, RI, 2017.

[4] P. J. Schweitzer, *Perturbation theory and finite Markov chains*, J. Appl. Probab. **5** (1968), 401–413.

[5] E. Seneta, *Perturbation of the stationary distribution measured by ergodicity coefficients*, Adv. in Appl. Probab. **20** (1988), 228–230.

[6] A. Sinclair and M. Jerrum, *Approximate counting, uniform generation and rapidly mixing Markov chains*, Inform. and Comput. **82** (1989), 93–133.