mathematics

MDPI

Article

Optimality Notions for Resolvent Monte Carlo

Tsvetelin Kostadinov [1,*] and Ivan T. Dimov [2,*]

[1] Faculty of Mathematics and Informatics, Sofia University "St. Kliment Ohridski", 1164 Sofia, Bulgaria
[2] Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria
* Correspondence: tsvetelinkostadinovts@gmail.com (T.K.); ivdimov@bas.bg (I.T.D.)

Abstract

Resolvent Monte Carlo estimates eigenvalues of large matrices by sampling Markov chains and reading the target value off a truncated resolvent quotient, trading exact arithmetic for a stochastic error that the almost-optimal sampling scheme is designed to suppress. This paper studies when that error vanishes outright. An exact closed-form identity is derived for the variance of the moment estimators of a general, possibly signed matrix, and is used to isolate a hierarchy of zero-variance notions ranging from the most local, which constrains only the first draws, through the finite-truncation regime that a practical run can certify, to the global regime in which every moment estimator is deterministic. Determinism of the estimator is separated from correctness of the eigenvalue it reports, and the exact conditions under which each notion holds are exhibited, together with the examples that separate them. A single edgewise condition, termed the eigen-triple condition, forces the truncated quotient to equal the target eigenvalue in finite samples; the associated moment and quotient variances are second order in the maximal edge defect and vanish at the eigen-triple. A linear-time procedure certifies the condition.

Keywords: resolvent Monte Carlo; variance reduction; zero-variance estimator; eigenvalue computation; Markov chain Monte Carlo; almost-optimal sampling

MSC: 65C05; 65F15; 65C40; 60J22; 15A18

1. Introduction

Eigenvalues govern the behavior of linear systems: they fix stability and decay rates, set the norms and conditioning that control numerical algorithms, and encode the characteristic modes of physical and stochastic models. For matrices small enough to factor, the symmetric and nonsymmetric eigenvalue problems are solved to high accuracy by mature dense and sparse methods [1–3]. As the dimension grows, however, a full decomposition becomes infeasible, and one turns to algorithms whose cost is governed by the sparsity of the matrix and by the accuracy demanded rather than by the dimension itself.

Randomization is one route to such algorithms. Randomized numerical linear algebra builds low-dimensional sketches that capture the dominant spectral subspace and recover eigenvalues and singular values from them [4–7]. A complementary, older tradition estimates spectral quantities directly as expectations: stochastic trace and spectral-density estimators sample matrix-vector products [8–12], and Markov chain methods sample short random walks on the index set whose weighted returns are unbiased for bilinear forms in powers of the matrix. The cost of a single walk depends on the number of nonzero entries, not on the order of the matrix, which is what makes the approach attractive at scale.

Academic Editor: Jonathan Blackledge

Received: 7 July 2026
Revised: 7 August 2026
Accepted: 11 August 2026
Published: 13 August 2026

Resolvent Monte Carlo (RMC) is the eigenvalue algorithm of this second tradition. It replaces the matrix A by a resolvent operator $(I - qA)^{-m}$, expanded as a Neumann series, so that the dominant eigenvalue problem for A becomes one for the resolvent, and reads the target value off a Rayleigh quotient of resolvent-weighted moments [13–15]. The moments themselves are estimated by the almost-optimal, namely, Monte Carlo Almost Optimal (MAO) sampling scheme, whose transition probabilities are proportional to the absolute entries of A; among the permissible sampling densities these minimize the second moment of the estimator, and they are the standard choice throughout the literature on stochastic linear algebra [16]. The method and its variants have been developed extensively for extremal and interior eigenvalues, for parallel and quasi-Monte Carlo implementations, and with error-balancing refinements [17–22].

The variance of the estimator is what determines the sample size, and the almost-optimality of the MAO densities is precisely a statement about it: a one-sided bound on the second moment that no permissible density improves [16]. That bound, however, leaves open the sharper question we address here. When is the variance not merely small but exactly zero? Equivalently, for which input data (A, v, h)—the matrix A together with the two vectors $v, h \in \mathbb{R}^n$ that specify the bilinear forms $(v, A^k h)$ the walks estimate—is the Monte Carlo estimator deterministic, so that a single sampled walk returns the exact answer? And when the final eigenvalue estimate is a quotient of two random sums, a second subtlety appears: the quotient can be deterministic even though its numerator and denominator fluctuate, and it can be deterministic yet wrong. Determinism of the estimator and correctness of the eigenvalue it reports are distinct properties, and the classical optimality theory does not separate them.

Zero-variance estimators have a long history in importance sampling. The classical zero-variance principle holds that an importance-sampling estimator is deterministic exactly when the sampling density is proportional to the integrand, equivalently when the estimator is constant on its support; this idea underlies zero-variance and adaptive importance-sampling schemes in computational physics [23], in rare-event and reliability estimation [24], and for Markov-process expectations characterized as solutions of a linear system, where a variance-preserving change of measure yields a zero-variance Markov estimator [25]. Our moment-level results are the RMC instance of this principle: each $\theta^{(k)}$ has zero variance exactly when it is constant on the admissible paths. What is specific to RMC, and new here, is the structure the principle acquires for the resolvent eigenvalue estimator, namely the nested moment hierarchy with the finite K-step interpolant that a run can certify, the separation of determinism of the quotient from correctness of the eigenvalue it reports, the single edgewise eigen-triple condition that forces finite-sample exactness with the eigenvalue recovered as a local edge ratio, and the edge-defect perturbation theory that controls the departure from it.

This paper develops the zero-variance theory of RMC. Our contributions are the following.

(i). An exact closed-form identity for the variance of the MAO moment estimator of a general, possibly signed, matrix (Section 3), which turns every subsequent zero-variance question into an algebraic one and makes explicit the second moment that the classical almost-optimality bound minimizes.

(ii). A complete hierarchy of zero-variance notions, from the local no-step and single-step conditions, through the finite-truncation K-step regime that a practical run can actually certify, to the global all-step regime, together with the two quotient-level notions of zero-variance and exactness; we characterize each notion and give the examples that separate them (Sections 5–9).

(iii). A single edgewise condition, the eigen-triple condition, that forces the truncated quotient to equal the target eigenvalue in finite samples (Sections 6 and 7); the eigenvalue is recovered as the common value of a local edge ratio rather than assumed, and the exactness degrades only to second order in a measurable edge defect.

(iv). A linear-time procedure that certifies the condition in a single pass over the sparsity pattern of A (Appendix A).

The paper is organized as follows. Section 2 fixes notation and states the MAO estimator and the RMC quotient. Section 3 proves the variance structure theorem, and Section 4 applies it to the balanced-matrix regime. Section 5 develops the moment-level zero-variance hierarchy; Section 6 introduces the eigen-triple condition and its perturbation theory; Section 7 treats the quotient level. Section 8 exhibits the separating examples, and Section 9 collects the full hierarchy. Sections 10 and 11 discuss the results and conclude.

2. Preliminaries

This section fixes the notation used throughout and states the two objects the rest of the paper analyses: the MAO moment estimator and the finite-truncation RMC quotient built from it.

Throughout, let

$$A = (a_{ij})_{i,j=1}^{n} \in \mathbb{R}^{n\times n}, \qquad v, h \in \mathbb{R}^n, \quad v \neq 0,\ h \neq 0.$$

For each row, we write its absolute row sum

$$d_i \coloneqq \sum_{j=1}^{n} |a_{ij}|,$$

and we assume that every row reached by the Markov chain is not identically zero, so that $d_i \neq 0$ there. We use the following derived quantities:

- $\overline{\overline{v}} = \{|v_i|\}_{i=1}^{n}$ and $\overline{\overline{A}} = \{|a_{ij}|\}_{i,j=1}^{n}$, the entrywise absolute values of v and A;
- $h^{\circ 2} = \{h_i^2\}_{i=1}^{n}$, the Hadamard (entrywise) square of h;
- $D = \operatorname{diag}(d_1, \ldots, d_n)$, the diagonal matrix of absolute row sums, so that $(D\overline{\overline{A}})_{ij} = d_i|a_{ij}|$;
- $S_v = \{i : v_i \neq 0\}$, the support of v, and $R(A, v)$ the set of indices reachable from S_v through nonzero entries of A;
- $(u, w) = \sum_i u_i w_i$, the bilinear form, and $\|u\|_1 = \sum_i |u_i|$, the ℓ^1 norm.

The almost-optimal (MAO) scheme samples a Markov chain on the index set $\{1, \ldots, n\}$ with initial and transition probabilities

$$p_i \coloneqq \frac{|v_i|}{\|v\|_1}, \qquad p_{ij} \coloneqq \frac{|a_{ij}|}{d_i}. \tag{1}$$

Among all permissible sampling densities these minimize the second moment of the estimator below, which is the sense in which they are almost optimal [16]. The scheme (1) generates a random path $\omega = (\alpha_0, \alpha_1, \ldots, \alpha_k)$: the initial state α_0 is drawn from the distribution $\{p_i\}_{i=1}^{n}$ and each subsequent state from the transition row $\{p_{\alpha_s j}\}_{j=1}^{n}$. The MAO estimator of the bilinear moment $(v, A^k h)$, as introduced in [16], is the random variable

$$\theta^{(k)}(\omega) \coloneqq \frac{v_{\alpha_0}}{p_{\alpha_0}} \prod_{s=0}^{k-1} \frac{a_{\alpha_s \alpha_{s+1}}}{p_{\alpha_s \alpha_{s+1}}} h_{\alpha_k}, \tag{2}$$

a deterministic function of the path, random because the path it is evaluated on is; the factors $v_{\alpha_0}/p_{\alpha_0}$ and $a_{\alpha_s\alpha_{s+1}}/p_{\alpha_s\alpha_{s+1}}$ are the importance-sampling weights that compensate for drawing the path from (1) instead of enumerating it. It is unbiased, $\mathbb{E}\left\{\theta^{(k)}\right\} = (v, A^k h)$ (this is recovered in the proof of Theorem 1), so that the moments $(v, A^k h)$, and through them spectral quantities of A, can be estimated by averaging (2) over independent chains. The bilinear moments are how Markov chain methods access the spectrum of A: the vector v fixes the distribution of the starting state, h evaluates the terminal state, and for v and h with nontrivial projection on the dominant eigendirection the ratio of consecutive moments $(v, A^{k+1}h)/(v, A^k h)$ converges to the dominant eigenvalue of A, which is the power method in bilinear form [16]. The RMC quotient below aggregates the same moments with resolvent weights to accelerate and stabilize this limit.

The resolvent of order m is expanded as a Neumann series,

$$(I - qA)^{-m} = \sum_{k\geq 0} w_k^{(m)} A^k, \qquad w_k^{(m)} \leftrightharpoons \binom{k+m-1}{k} q^k,$$

for $m \geq 1$, $K \geq 0$, and a real parameter q for which the series converges. Truncating at order K and forming the resolvent-weighted Rayleigh quotient of the moment estimators over N independent chains gives the finite-truncation RMC estimator of the target eigenvalue,

$$\widehat{\lambda}_{K,N}^{(m)} = \frac{\sum_{r=1}^{N}\sum_{k=0}^{K} w_k^{(m)}\theta_r^{(k+1)}}{\sum_{r=1}^{N}\sum_{k=0}^{K} w_k^{(m)}\theta_r^{(k)}}, \tag{3}$$

defined whenever the denominator is nonzero. The numerator estimates $(v, A(I-qA)^{-m}h)$ and the denominator $(v, (I-qA)^{-m}h)$, so that $\widehat{\lambda}_{K,N}^{(m)}$ approximates the corresponding resolvent Rayleigh quotient.

The usual eigenvalue interpretation requires the standard RMC spectral regime: q is chosen in the resolvent domain so that the Neumann series converges, the resolvent Rayleigh quotient has a nonzero denominator, and the chosen vectors v and h have nontrivial projection on the eigendirection that the resolvent filter is meant to isolate. Under the familiar simple-target setting, increasing the truncation and resolvent order then drives the deterministic resolvent quotient to the selected eigenvalue. The zero-variance results below are finite-truncation algebraic statements; except where a limiting interpretation is explicitly invoked, they are conditioned only on the stated reachability assumptions and on the sampled denominator being nonzero.

We isolate several notions of variance optimality, from the most local to the most global. At the moment level, where the objects are the estimators $\theta^{(k)}$ of the bilinear moments $(v, A^k h)$, these are *no-step* and *single-step zero-variance* (only $\theta^{(0)}$, respectively $\theta^{(1)}$, is deterministic), *K-step zero-variance* (the moments $\theta^{(0)}, \ldots, \theta^{(K)}$ are deterministic), and *all-step zero-variance* (every $\theta^{(k)}$ is deterministic). The K-step notion is the finite moment-level interpolant: no-step and single-step are its cases $K = 0, 1$, and all-step is the limit $K \to \infty$. At the *quotient level*, where the object is the RMC estimator $\widehat{\lambda}_{K,N}^{(m)}$ formed from $\theta^{(0)}, \ldots, \theta^{(K+1)}$, these are *quotient zero-variance* (the quotient is deterministic) and *quotient exactness* (that deterministic value is the target eigenvalue). Thus a quotient truncated at order K is governed by the moment-level condition at level $K{+}1$. The moment-level conditions are nested, $(K{+}1)$-step $\Rightarrow$ K-step, and the only implications across levels are

$$(K{+}1)\text{-step zero-variance} \implies \text{quotient zero-variance} \impliedby \text{quotient exactness}$$

(at truncation order K; the first arrow specializes to all-step $\Rightarrow$ quotient zero-variance as $K \to \infty$). Every other arrow fails, and Section 9 collects the counterexamples that show

why. The single edge-ratio (eigen-triple) condition of Section 6 forces quotient exactness, and adjoining no-step compatibility forces K-step, indeed all-step, zero-variance.

3. The Variance Structure Theorem

Every zero-variance question in this paper is decided by a single exact identity: the variance of the moment estimator $\theta^{(k)}$ of (2), written as the difference between a transport term and the squared moment. We state and prove it for a general, possibly signed, matrix.

Theorem 1. *Let $A \in \mathbb{R}^{n\times n}$ and $v, h \in \mathbb{R}^n$ be the input data for the RMC algorithm.*
We require:

$$v \neq 0, \qquad h \neq 0$$
$$\forall i = 1, \dots n : \sum_j |a_{ij}| \neq 0$$

Define:

$$\overline{\overline{v}} \leftrightharpoons \{|v_i|\}_{i=1}^n, \qquad \overline{\overline{A}} \leftrightharpoons \{|a_{ij}|\}_{i,j=1}^n$$
$$h^{\circ 2} \leftrightharpoons \left\{h_i^2\right\}_{i=1}^n \qquad D \leftrightharpoons \operatorname{diag}(\|a_1\|_1, \dots, \|a_n\|_1)$$

Then:

$$\operatorname{Var}\left\{\theta^{(k)}\right\} = \|v\|_1 \left(\overline{\overline{v}}, (D\overline{\overline{A}})^k h^{\circ 2}\right) - \left(v, A^k h\right)^2$$

Proof. Write a path as $\omega = (\alpha_0, \alpha_1, \dots, \alpha_k)$, and call it *admissible* if $v_{\alpha_0} \neq 0$ and $a_{\alpha_s\alpha_{s+1}} \neq 0$ for $s = 0, \dots, k-1$; only admissible paths carry positive MAO probability. Substituting the MAO probabilities $p_i = |v_i|/\|v\|_1$ and $p_{ij} = |a_{ij}|/d_i$ into the estimator and using $x/|x| = \operatorname{sgn}(x)$, every admissible path gives

$$\theta^{(k)}(\omega) = \|v\|_1 \operatorname{sgn}(v_{\alpha_0}) \prod_{s=0}^{k-1} \left(d_{\alpha_s} \operatorname{sgn}(a_{\alpha_s\alpha_{s+1}})\right) h_{\alpha_k}, \tag{4}$$

since $v_{\alpha_0}/p_{\alpha_0} = \|v\|_1 \operatorname{sgn}(v_{\alpha_0})$ and $a_{\alpha_s\alpha_{s+1}}/p_{\alpha_s\alpha_{s+1}} = d_{\alpha_s} \operatorname{sgn}(a_{\alpha_s\alpha_{s+1}})$. The probability of an admissible path is

$$\mathbb{P}(\omega) = p_{\alpha_0} \prod_{s=0}^{k-1} p_{\alpha_s\alpha_{s+1}} = \frac{|v_{\alpha_0}|}{\|v\|_1} \prod_{s=0}^{k-1} \frac{|a_{\alpha_s\alpha_{s+1}}|}{d_{\alpha_s}}.$$

Multiplying (4) by $\mathbb{P}(\omega)$ and pairing the factors of the two products termwise, the identity $|x| \operatorname{sgn}(x) = x$ gives

$$\mathbb{P}(\omega)\, \theta^{(k)}(\omega) = \left(\frac{|v_{\alpha_0}|}{\|v\|_1} \|v\|_1 \operatorname{sgn}(v_{\alpha_0})\right) \prod_{s=0}^{k-1} \left(\frac{|a_{\alpha_s\alpha_{s+1}}|}{d_{\alpha_s}} d_{\alpha_s} \operatorname{sgn}(a_{\alpha_s\alpha_{s+1}})\right) h_{\alpha_k} = v_{\alpha_0} \prod_{s=0}^{k-1} a_{\alpha_s\alpha_{s+1}}\, h_{\alpha_k}.$$

The signs and the sampling weights cancel factor by factor, so the two products merge into a single one. Summing over paths,

$$\mathbb{E}\left\{\theta^{(k)}\right\} = \sum_\omega \mathbb{P}(\omega)\, \theta^{(k)}(\omega) = \sum_\omega v_{\alpha_0} \prod_{s=0}^{k-1} a_{\alpha_s\alpha_{s+1}}\, h_{\alpha_k} = \left(v, A^k h\right),$$

where the last sum ranges over all index tuples $(\alpha_0, \dots, \alpha_k)$: a non-admissible tuple contributes a zero factor, so extending the sum past the admissible paths adds nothing. This recovers the unbiasedness of the MAO estimator for the bilinear moment $(v, A^k h)$.

Squaring (4) removes every sign, $\operatorname{sgn}(\cdot)^2 = 1$, so

$$\big(\theta^{(k)}(\omega)\big)^2 = \|v\|_1^2 \Big(\prod_{s=0}^{k-1} d_{\alpha_s}^2\Big) h_{\alpha_k}^2.$$

Multiplying by $\mathbb{P}(\omega)$, the factor $\|v\|_1^2/\|v\|_1 = \|v\|_1$ and each $d_{\alpha_s}^2/d_{\alpha_s} = d_{\alpha_s}$ remain, so

$$\mathbb{E}\Big\{\big(\theta^{(k)}\big)^2\Big\} = \sum_{\omega} \mathbb{P}(\omega)\big(\theta^{(k)}(\omega)\big)^2 = \|v\|_1 \sum_{\omega} |v_{\alpha_0}| \prod_{s=0}^{k-1} \big(d_{\alpha_s} |a_{\alpha_s\alpha_{s+1}}|\big) h_{\alpha_k}^2.$$

Since $(D\overline{\overline{A}})_{ij} = d_i |a_{ij}|$, the path sum contracts to a matrix power,

$$\sum_{\omega} |v_{\alpha_0}| \prod_{s=0}^{k-1} (D\overline{\overline{A}})_{\alpha_s\alpha_{s+1}} h_{\alpha_k}^2 = \Big(\overline{\overline{v}}, (D\overline{\overline{A}})^k h^{\circ 2}\Big),$$

whence $\mathbb{E}\Big\{(\theta^{(k)})^2\Big\} = \|v\|_1 \Big(\overline{\overline{v}}, (D\overline{\overline{A}})^k h^{\circ 2}\Big)$.

Subtracting the square of the mean gives

$$\operatorname{Var}\Big\{\theta^{(k)}\Big\} = \mathbb{E}\Big\{\big(\theta^{(k)}\big)^2\Big\} - \big(\mathbb{E}\Big\{\theta^{(k)}\Big\}\big)^2 = \|v\|_1 \Big(\overline{\overline{v}}, (D\overline{\overline{A}})^k h^{\circ 2}\Big) - \Big(v, A^k h\Big)^2.$$

□

Remark 1. *The second moment* $\mathbb{E}\Big\{(\theta^{(k)})^2\Big\} = \|v\|_1 (\overline{\overline{v}}, (D\overline{\overline{A}})^k h^{\circ 2})$ *is exactly the quantity whose minimization over permissible densities underlies the almost-optimality of the MAO densities [16,26]. Theorem 1 makes that second moment explicit as a closed form for general (possibly signed) A and general h, and records it as an exact variance identity rather than the one-sided bound used to establish optimality. The Hadamard square* $h^{\circ 2}$ *and the transport operator* $(D\overline{\overline{A}})^k$ *that appear here are what drive the structural results of the following sections.*

4. Variance Optimality of RMC on Balanced Matrices

A simple optimality condition can be extracted from Theorem 1. The result is a routine extension of the balanced-matrix zero-variance phenomenon for MAO estimators discussed in [16].

Proposition 1. *Let* $\mathcal{C}$ *be the set of* $n \times n$ *matrices that have non-negative entries and constant row sums, i.e.,*

$$\mathcal{C} \leftrightharpoons \Big\{A \in \mathbb{R}_{\geq 0}^{n\times n} \mid \exists c \in \mathbb{R}_{>0} : \; A\mathbf{1} = c\mathbf{1}\Big\}.$$

For $A \in \mathcal{C}$ *choose*

$$v = \frac{1}{n}\mathbf{1}, \qquad h = \mathbf{1}.$$

Then $\operatorname{Var}\Big\{\theta^{(k)}\Big\} = 0$ *for all* $k \geq 0$.

Proof. For $A \in \mathcal{C}$, every entry is nonnegative and every row sums to c, so $d_i = \sum_j |a_{ij}| = \sum_j a_{ij} = c$ and $\overline{\overline{A}} = A$. With $v = n^{-1}\mathbf{1}$ and $h = \mathbf{1}$ we have $\|v\|_1 = 1$, $\overline{\overline{v}} = n^{-1}\mathbf{1}$, $h^{\circ 2} = \mathbf{1}$, and $D = cI$, hence $D\overline{\overline{A}} = cA$ and $(D\overline{\overline{A}})^k\mathbf{1} = c^k A^k\mathbf{1} = c^{2k}\mathbf{1}$. Therefore, $\|v\|_1(\overline{\overline{v}}, (D\overline{\overline{A}})^k h^{\circ 2}) = (n^{-1}\mathbf{1}, c^{2k}\mathbf{1}) = c^{2k}$, while the mean term is $(v, A^k h) = (n^{-1}\mathbf{1}, c^k\mathbf{1}) = c^k$. By Theorem 1, $\operatorname{Var}\Big\{\theta^{(k)}\Big\} = c^{2k} - (c^k)^2 = 0$ for every $k \geq 0$. □

Corollary 1. *Suppose* $A \in \mathcal{C}$ *with* $A\mathbf{1} = c\mathbf{1}$, *and take* $v = \frac{1}{n}\mathbf{1}$, $h = \mathbf{1}$ *as in Theorem 1. Then,* $\theta^{(k)} = c^k$ *almost surely for every* $k \geq 0$.

Proof. By Theorem 1, $\mathrm{Var}\left\{\theta^{(k)}\right\} = 0$, so $\theta^{(k)}$ is almost surely equal to its mean $\mathbb{E}\left\{\theta^{(k)}\right\} = (v, A^k h) = c^k$, computed in the preceding proof. □

Corollary 2. *Let $A \in \mathbb{R}^{n\times n}_{\geq 0}$ have no zero rows. Define:*

$$r_i = \sum_{j=1}^{n} a_{ij} > 0, \qquad i = 1, \ldots, n,$$

and assume that the r_i are not all equal. Choose v and h as per Theorem 1. Then, the one-step MAO estimator satisfies

$$\mathrm{Var}\left\{\theta^{(1)}\right\} = \|v\|_1(\overline{\overline{v}}, (D\overline{\overline{A}})h^{\circ 2}) - (v, Ah)^2 = \frac{1}{n}\sum_{i=1}^{n} r_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} r_i\right)^2 > 0.$$

The right-hand side is the population variance of the vector r, which can equivalently be written as

$$\frac{1}{2n^2}\sum_i\sum_j (r_i - r_j)^2.$$

Proof. Here $A \geq 0$, so $\overline{\overline{A}} = A$ and $d_i = r_i$. With $v = n^{-1}\mathbf{1}$ and $h = \mathbf{1}$, we have $\|v\|_1 = 1$, $\overline{\overline{v}} = n^{-1}\mathbf{1}$, $h^{\circ 2} = \mathbf{1}$, and $(D\overline{\overline{A}}\, h^{\circ 2})_i = d_i \sum_j a_{ij} = r_i^2$, so $\|v\|_1(\overline{\overline{v}}, D\overline{\overline{A}}\, h^{\circ 2}) = \frac{1}{n}\sum_i r_i^2$ while $(v, Ah) = \frac{1}{n}\sum_i r_i$. Theorem 1 then gives the displayed identity. The rearrangement $\frac{1}{n}\sum_i r_i^2 - \left(\frac{1}{n}\sum_i r_i\right)^2 = \frac{1}{2n^2}\sum_i\sum_j (r_i - r_j)^2$ is the standard variance identity, which is strictly positive because the r_i are not all equal. □

The complementary degeneration—keeping equal row sums but dropping nonnegativity—fails just as decisively, and quantifiably so.

Theorem 2 (Failure for signed matrices with equal row sums)**.** *Let $A \in \mathbb{R}^{n\times n}$ have no zero rows and equal algebraic row sums, $A\mathbf{1} = c\mathbf{1}$ with $c > 0$, and choose $v = \frac{1}{n}\mathbf{1}$, $h = \mathbf{1}$ as in Theorem 1. Then, the one-step MAO estimator satisfies*

$$\mathrm{Var}\left\{\theta^{(1)}\right\} = \|v\|_1(\overline{\overline{v}}, (D\overline{\overline{A}})h^{\circ 2}) - (v, Ah)^2 = \frac{1}{n}\sum_{i=1}^{n} d_i^2 - c^2, \qquad d_i = \sum_{j=1}^{n} |a_{ij}|.$$

If at least one row of A contains entries of both signs, then $d_i > c$ for that row and hence

$$\mathrm{Var}\left\{\theta^{(1)}\right\} = \frac{1}{n}\sum_{i=1}^{n} d_i^2 - c^2 > 0.$$

Thus, equal algebraic row sums alone do not imply zero-variance MAO sampling: the nonnegativity hypothesis of Theorem 1 cannot be dropped.

Proof. By Theorem 1 with $v = n^{-1}\mathbf{1}$ and $h = \mathbf{1}$, we have $\|v\|_1 = 1$, $\overline{\overline{v}} = \frac{1}{n}\mathbf{1}$, and $h^{\circ 2} = \mathbf{1}$. Since $\overline{\overline{A}}\mathbf{1} = (d_1, \ldots, d_n)^\top$ and $D = \mathrm{diag}(d_1, \ldots, d_n)$,

$$\|v\|_1(\overline{\overline{v}}, D\overline{\overline{A}}\, h^{\circ 2}) = \frac{1}{n}\sum_{i=1}^{n} d_i^2.$$

The mean term is $(v, Ah) = \frac{1}{n}\mathbf{1}^\top A\mathbf{1} = c$, giving $\mathrm{Var}\{\theta^{(1)}\} = \frac{1}{n}\sum_i d_i^2 - c^2$. For every row, $d_i = \sum_j |a_{ij}| \geq \left|\sum_j a_{ij}\right| = c$, with strict inequality whenever row i has entries of both signs. Hence $\frac{1}{n}\sum_i d_i^2 \geq c^2$, strictly when some row is mixed-sign, and therefore $\mathrm{Var}\{\theta^{(1)}\} > 0$. □

Remark 2. *The class from Theorem 1 is not maximal. It can be extended for matrices with only non-positive entries, i.e.,*

$$\mathcal{C}' \rightleftharpoons \left\{ A \in \mathbb{R}^{n\times n}_{\leq 0} \mid \exists c \in \mathbb{R}_{<0} : \ A\mathbf{1} = c\mathbf{1} \right\}.$$

So the maximal class is $\mathcal{C} \cup \mathcal{C}'$. However, the result cannot be extended to mixed-sign matrices: dropping nonnegativity makes the one-step estimator fluctuate already at equal row sums (Theorem 2).

Remark 3. *The maximality statement is relative to the fixed test vectors $v = n^{-1}\mathbf{1}$ and $h = \mathbf{1}$. If v and h are allowed to depend on A, then additional degenerate zero-variance cases appear, for example when v is supported on a deterministic closed component. Thus, the theorem classifies the canonical balanced-row zero-variance regime, not all triples (A, v, h) with zero MAO variance.*

As an example, take $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ with $h = \mathbf{1}$. The full-support choice $v_1 = \frac{1}{2}\mathbf{1}$ gives $\theta^{(1)} \in \{2, 3\}$ and hence $\mathrm{Var}\left\{\theta^{(1)}\right\} > 0$, since $A \notin \mathcal{C}$ (the row sums 2 and 3 differ). The choice $v_2 = e_1$, on the other hand, traps the walk in the closed component $\{1\}$, so $\theta^{(k)} = 2^k$ is deterministic: a zero-variance triple lying outside the balanced class $\mathcal{C}$ and produced only by the A-dependent support of v.

5. Moment-Level Zero-Variance

We now study when the moment estimators $\theta^{(k)}$ of (2) are deterministic. The conditions form a nested family indexed by how many moments are constrained: the local no-step and single-step conditions fix only $\theta^{(0)}$ and $\theta^{(1)}$; the K-step condition fixes $\theta^{(0)}, \ldots, \theta^{(K)}$ and is the finite notion a truncated run can certify; and the all-step condition fixes every $\theta^{(k)}$. The three subsections below treat these in turn, with no-step and single-step appearing as the cases $K = 0, 1$ and all-step as the limit $K \to \infty$. Each notion is characterized by a path-balance condition on the collapsed form of the estimator.

5.1. No-Step and Single-Step Variance Optimality

We first isolate the local zero-variance conditions for the first two MAO estimators. These are the weakest variance-optimality statements, since they involve only the initial draw and the first transition.

Let

$$S_v \rightleftharpoons \{i \in \{1, \ldots, n\} \mid v_i \neq 0\}$$

and define the one-step admissible edge set

$$\mathcal{E}_1(A, v) \rightleftharpoons \{(i, j) \mid i \in S_v,\ a_{ij} \neq 0\}.$$

All atoms in S_v and $\mathcal{E}_1(A, v)$ have strictly positive MAO probability.

Definition 1. *The triple (A, v, h) is called no-step compatible if there exists $\beta_0 \in \mathbb{R}$ such that*

$$\mathrm{sgn}(v_i)h_i = \beta_0 \qquad \text{for every } i \in S_v.$$

It is called single-step compatible if there exists $\beta_1 \in \mathbb{R}$ such that

$$\mathrm{sgn}(v_i)d_i\,\mathrm{sgn}(a_{ij})h_j = \beta_1 \qquad \text{for every } (i, j) \in \mathcal{E}_1(A, v).$$

Proposition 2. *The MAO estimators satisfy*

$$\mathbb{E}\theta^{(0)} = (v, h), \qquad \mathbb{E}\theta^{(1)} = (v, Ah).$$

Proof. Since $p_i = |v_i|/\|v\|_1$, one has

$$\theta^{(0)} = \frac{v_{\alpha_0}}{p_{\alpha_0}} h_{\alpha_0} = \|v\|_1 \operatorname{sgn}(v_{\alpha_0}) h_{\alpha_0}.$$

Taking expectation over α_0 gives

$$\mathbb{E}\theta^{(0)} = \sum_{i \in S_v} \frac{|v_i|}{\|v\|_1} \|v\|_1 \operatorname{sgn}(v_i) h_i = \sum_i v_i h_i = (v, h).$$

Similarly,

$$\theta^{(1)} = \|v\|_1 \operatorname{sgn}(v_{\alpha_0}) d_{\alpha_0} \operatorname{sgn}(a_{\alpha_0 \alpha_1}) h_{\alpha_1}.$$

Therefore,

$$\begin{aligned} \mathbb{E}\theta^{(1)} &= \sum_{i \in S_v} \sum_{j : a_{ij} \neq 0} \frac{|v_i|}{\|v\|_1} \frac{|a_{ij}|}{d_i} \|v\|_1 \operatorname{sgn}(v_i) d_i \operatorname{sgn}(a_{ij}) h_j \\ &= \sum_i \sum_j v_i a_{ij} h_j = (v, Ah). \end{aligned}$$

□

Proposition 3. $\operatorname{Var}\left\{\theta^{(0)}\right\} = 0$ *if and only if* (A, v, h) *is no-step compatible.*

Proof. The sample space S_v is finite and every admissible atom has positive probability, so $\theta^{(0)}$ has zero variance if and only if it is constant on its support. Since

$$\theta^{(0)} = \|v\|_1 \operatorname{sgn}(v_{\alpha_0}) h_{\alpha_0},$$

this holds if and only if $\operatorname{sgn}(v_i) h_i$ is constant on S_v, which is precisely no-step compatibility. □

Proposition 4. $\operatorname{Var}\left\{\theta^{(1)}\right\} = 0$ *if and only if* (A, v, h) *is single-step compatible.*

Proof. The edge set $\mathcal{E}_1(A, v)$ is finite and every admissible atom has positive probability, so $\theta^{(1)}$ has zero variance if and only if it is constant on its support. On the admissible edge (i, j),

$$\theta^{(1)} = \|v\|_1 \operatorname{sgn}(v_i) d_i \operatorname{sgn}(a_{ij}) h_j,$$

so this holds if and only if $\operatorname{sgn}(v_i) d_i \operatorname{sgn}(a_{ij}) h_j$ is constant on $\mathcal{E}_1(A, v)$, which is precisely single-step compatibility. □

5.2. K-Step Variance Optimality

The no-step and single-step criteria of Section 5.1 control only $\theta^{(0)}$ and $\theta^{(1)}$. The all-step criterion of Section 5.3 treats the opposite extreme, constraining $\theta^{(k)}$ for *every* $k \geq 0$. We therefore isolate the finite moment-level condition that $\theta^{(0)}, \ldots, \theta^{(K)}$ be deterministic. It interpolates between the local criteria ($K \leq 1$) and all-step zero variance ($K \to \infty$), and it is the condition one can actually certify for a finite simulation. Since a quotient truncated at order K uses the moments $\theta^{(0)}, \ldots, \theta^{(K+1)}$, the corresponding moment-level certificate is $(K+1)$-step zero variance.

As in Section 5.3, assume $d_i \neq 0$ for every reachable i, and for $k \geq 0$ let $\Omega_k(A, v)$ be the set of admissible paths $\omega = (\alpha_0, \ldots, \alpha_k)$ with $v_{\alpha_0} \neq 0$ and $a_{\alpha_s \alpha_{s+1}} \neq 0$; every such path carries strictly positive MAO probability. On its admissible support the estimator takes the sign/row-sum form (4), where $d_i = \sum_{j=1}^{n} |a_{ij}|$.

Definition 2 (K-step zero variance)**.** *For $K \geq 0$, the triple (A, v, h) has K-step zero variance if*

$$\mathrm{Var}\left\{\theta^{(k)}\right\} = 0 \qquad \text{for every } k = 0, 1, \ldots, K.$$

Theorem 3. *For $K \geq 0$, the following are equivalent.*

(i). (A, v, h) *has K-step zero variance.*
(ii). *For each $k = 0, \ldots, K$ there is a constant $c_k \in \mathbb{R}$ with $\theta^{(k)}(\omega) = c_k$ for every $\omega \in \Omega_k(A, v)$.*
(iii). *for each $k = 0, \ldots, K$ there is $\beta_k \in \mathbb{R}$ with*

$$\operatorname{sgn}(v_{\alpha_0}) \prod_{s=0}^{k-1} (d_{\alpha_s} \operatorname{sgn}(a_{\alpha_s \alpha_{s+1}})) h_{\alpha_k} = \beta_k$$

for every $\omega \in \Omega_k(A, v)$.

In either case $\theta^{(k)} = \|v\|_1 \beta_k$ almost surely for $k = 0, \ldots, K$. The cases $K = 0$ and $K = 1$ are Proposition 3 and Proposition 4, and the limit $K \to \infty$ is Theorem 4.

Proof. (i)⇔(ii). For each k, the sample space $\Omega_k(A, v)$ is finite and every admissible atom has positive probability, so $\theta^{(k)}$ has zero variance if and only if it is constant on its support. Applying this for $k = 0, \ldots, K$ gives the equivalence, with the common values c_k. (ii)⇔(iii). By the collapsed form, $\theta^{(k)} = \|v\|_1 B_k$ where B_k is the levelwise quantity in (iii). Since $\|v\|_1 > 0$ is fixed, $\theta^{(k)}$ is constant on $\Omega_k(A, v)$ if B_k is, with $c_k = \|v\|_1 \beta_k$. □

The condition is monotone in K and brackets the criteria already established.

Proposition 5. *$(K{+}1)$-step zero variance implies K-step zero variance, and (A, v, h) has all-step zero variance if and only if it has K-step zero variance for every $K \geq 0$. Each implication $(K{+}1) \Rightarrow K$ is strict.*

Proof. The defining condition for $(K{+}1)$-step zero variance contains that for K-step, giving the implication; all-step zero variance is the conjunction over all $k \geq 0$, i.e., K-step for every K. Strictness is witnessed by Example 1. □

Example 1. *Let $v = e_1$, $h = \mathbf{1}$, and*

$$A = \begin{pmatrix} 0 & 2 & 2 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

From $S_v = \{1\}$ the admissible paths are $1 \to \{2, 3\}$ followed by the forced self-loops $2 \to 2$ and $3 \to 3$. Here $d_1 = 4$, $d_2 = 4$, $d_3 = 6$, and all reachable edges are positive. Thus $\theta^{(0)} = h_1 = 1$ and $\theta^{(1)} = d_1 = 4$ on either first edge, so (A, v, h) has 1-step (single-step) zero variance. But

$$\theta^{(2)} = \begin{cases} d_1 d_2 = 16, & \text{on } 1 \to 2 \to 2, \\ d_1 d_3 = 24, & \text{on } 1 \to 3 \to 3, \end{cases}$$

so $\mathrm{Var}\{\theta^{(2)}\} > 0$ *and 2-step zero variance fails. (The edge ratios* $\rho_{22} = 4 \neq 6 = \rho_{33}$ *differ, so this is not an eigen-triple; cf. Section 6.) The same construction, balanced up to level K and broken at level K+1, makes every implication of Proposition 5 strict.*

A single edgewise condition is again sufficient, now required only out to the truncation depth.

Proposition 6. *Suppose* (A, v, h) *is no-step compatible,* $\mathrm{sgn}(v_i)h_i = \beta$ *on* S_v*, and the eigen-triple condition (Definition 5) with eigenvalue* λ *holds on every edge lying on some admissible path of length* $\leq K$*. Then,* (A, v, h) *has K-step zero variance, with* $\theta^{(k)} = \|v\|_1 \beta \lambda^k$ *almost surely for* $k = 0, \ldots, K$*.*

Proof. No-step compatibility gives $\theta^{(0)} = \|v\|_1 \beta$ on every admissible path. As in Proposition 9, the eigen-triple identity on the edges used yields $\theta^{(k+1)} = \lambda \theta^{(k)}$ along every admissible prefix of length $< K$, so by induction $\theta^{(k)} = \|v\|_1 \beta \lambda^k$ for $k = 0, \ldots, K$, a constant; hence $\mathrm{Var}\{\theta^{(k)}\} = 0$ for those k. □

Remark 4. *The truncated quotient* $\widehat{\lambda}^{(m)}_{K,N}$ *(Section 7) is formed from* $\theta^{(0)}, \ldots, \theta^{(K+1)}$*. Hence* $(K+1)$*-step zero variance makes every moment entering the quotient deterministic, and by Proposition 13 the quotient is then deterministic (quotient zero-variance). This is the finite, certifiable condition that all-step zero variance over-delivers: at truncation order K, one never needs determinism beyond level K+1.*

5.3. All-Step Variance Optimality

The no-step and single-step criteria of Section 5.1 are *local*: they constrain only the initial draw and the first transition, and they say nothing about $\theta^{(2)}, \theta^{(3)}, \ldots$. We now introduce the moment-level notion that controls every power simultaneously and characterize it exactly; Section 6 then isolates a structural subclass—the *eigen-triples*—that is easy to verify and contains the balanced-row regime of Theorem 1 as a special case.

Throughout, we keep the standing hypotheses $v \neq 0$, $h \neq 0$, and

$$d_i \leftrightharpoons \sum_{j=1}^{n} |a_{ij}| \neq 0 \qquad \text{for every } i \in R(A, v),$$

where $S_v = \{\, i : v_i \neq 0 \,\}$ and $R(A, v)$ is the set of states ($R(A, v)$ is the transitive closure of $\mathcal{E}_1(A, v)$ introduced in Section 5.1.) reachable from S_v through nonzero entries of A. For $k \geq 0$ let $\Omega_k(A, v)$ be the set of admissible paths

$$\omega = (\alpha_0, \ldots, \alpha_k), \qquad v_{\alpha_0} \neq 0, \quad a_{\alpha_s \alpha_{s+1}} \neq 0 \ \ (s = 0, \ldots, k-1).$$

On its admissible support the MAO estimator collapses to the sign/row-sum form of Equation (4), obtained in the proof of Theorem 1 from $v_{\alpha_0} / p_{\alpha_0} = \|v\|_1 \,\mathrm{sgn}(v_{\alpha_0})$ and $a_{\alpha_s \alpha_{s+1}} / p_{\alpha_s \alpha_{s+1}} = d_{\alpha_s} \,\mathrm{sgn}(a_{\alpha_s \alpha_{s+1}})$.

Definition 3. *The triple* (A, v, h) *has all-step zero variance if*

$$\mathrm{Var}\{\theta^{(k)}\} = 0 \qquad \textit{for every } k \geq 0.$$

Proposition 7. *The triple* (A, v, h) *has all-step zero variance if and only if, for every* $k \geq 0$*, there is a constant* $c_k \in \mathbb{R}$ *with*

$$\theta^{(k)}(\omega) = c_k \qquad \textit{for every } \omega \in \Omega_k(A, v).$$

Proof. For each k the sample space $\Omega_k(A, v)$ is finite and every admissible path carries strictly positive MAO probability. A random variable on a finite probability space whose atoms all have positive probability has zero variance if and only if it is constant on its support. Applying this to $\theta^{(k)}$ for every k gives the claim. □

This reduces the analytic statement $\mathrm{Var}\{\theta^{(k)}\} = 0$ to a combinatorial one: the pathwise value (4) must not depend on the path. Naming that requirement yields the exact (maximal) class.

Definition 4. *The triple (A, v, h) is levelwise path-balanced if for every $k \geq 0$ there exists $\beta_k \in \mathbb{R}$ such that*

$$\operatorname{sgn}(v_{\alpha_0}) \prod_{s=0}^{k-1} \Big(d_{\alpha_s} \operatorname{sgn}(a_{\alpha_s \alpha_{s+1}}) \Big) h_{\alpha_k} = \beta_k$$

for every admissible path $\omega = (\alpha_0, \ldots, \alpha_k) \in \Omega_k(A, v)$.

Theorem 4. *Under the standing hypotheses the following are equivalent.*

(i). (A, v, h) *has all-step zero variance.*

(ii). (A, v, h) *is levelwise path-balanced.*

In either case, $\theta^{(k)} = \|v\|_1 \, \beta_k$ almost surely for every $k \geq 0$.

Remark 5. *The no-step and single-step zero-variance criteria (Proposition 3 and Proposition 4) are exactly the $k = 0$ and $k = 1$ instances of levelwise path-balance: constancy of $\operatorname{sgn}(v_{\alpha_0}) h_{\alpha_0}$ on S_v is the level-0 condition, and constancy of $\operatorname{sgn}(v_{\alpha_0}) d_{\alpha_0} \operatorname{sgn}(a_{\alpha_0 \alpha_1}) h_{\alpha_1}$ on $\mathcal{E}_1(A, v)$ is the level-1 condition. Thus Theorem 4 subsumes both as the two lowest levels of the all-step characterization.*

Proof. By (4), $\theta^{(k)} = \|v\|_1 \, B_k(\omega)$, where $B_k(\omega)$ denotes the levelwise quantity of Definition 4. Since $\|v\|_1 > 0$ is a fixed scalar, $\theta^{(k)}$ is constant on $\Omega_k(A, v)$ if and only if B_k is. By Proposition 7, all-step zero variance is exactly constancy of every $\theta^{(k)}$, hence of every B_k, which is precisely Definition 4. The common value is $c_k = \|v\|_1 \beta_k$. □

Remark 6. *Theorem 4 is a genuine characterization: levelwise path-balance is necessary and sufficient, so it describes the maximal all-step zero-variance class. It is, however, a condition on the entire path family, with one scalar β_k per level. The next subsection gives a single edgewise condition that implies it uniformly across all levels.*

6. The Eigen-Triple Criterion

The levelwise characterization above is exact but global, with one scalar β_k per level. We now isolate a single edgewise condition—the *eigen-triple*—that implies it uniformly across all levels, contains the balanced-row regime of Theorem 1 as a special case, and is equivalent to a purely local condition on (A, h) that never mentions the eigenvalue. The eigenvalue itself emerges as the common edge ratio (Theorem 6), and quotient exactness (Theorem 8) is the immediate payoff at the quotient level.

Definition 5. *A triple (A, v, h) is an eigen-triple with eigenvalue $\lambda \in \mathbb{R}$ if*

$$\frac{a_{ij}}{p_{ij}} h_j = \lambda h_i, \qquad \textit{equivalently} \qquad d_i \operatorname{sgn}(a_{ij}) \, h_j = \lambda h_i,$$

for every reachable nonzero edge $i \to j$ (that is, $i \in R(A, v)$ and $a_{ij} \neq 0$).

The condition is *edgewise*, and is strictly stronger than the ordinary eigenvector equation, which it nevertheless implies on the reachable part.

Proposition 8. *If (A, v, h) is an eigen-triple with eigenvalue λ, then $(Ah)_i = \lambda h_i$ for every $i \in R(A, v)$.*

Remark 7. *We concern ourselves with the reachable indices only; in this sense, h is MAO-observably an eigenvector, though it need not be an actual eigenvector.*

Proof. Fix a reachable row i. The eigen-triple condition gives $\operatorname{sgn}(a_{ij})h_j = \lambda h_i / d_i$ for each j with $a_{ij} \neq 0$. Hence

$$(Ah)_i = \sum_j a_{ij} h_j = \sum_{j:\, a_{ij} \neq 0} |a_{ij}| \operatorname{sgn}(a_{ij}) h_j = \frac{\lambda h_i}{d_i} \sum_{j:\, a_{ij} \neq 0} |a_{ij}| = \frac{\lambda h_i}{d_i}\, d_i = \lambda h_i.$$

□

The defining property propagates to a pathwise recurrence, the engine behind all the zero-variance statements below.

Proposition 9. *If (A, v, h) is an eigen-triple with eigenvalue λ, then*

$$\theta^{(k+1)} = \lambda\, \theta^{(k)}$$

along every admissible coupled path prefix; i.e., for $\omega = (\alpha_0, \ldots, \alpha_{k+1}) \in \Omega_{k+1}(A, v)$, with $\theta^{(k)}$ evaluated on the prefix $(\alpha_0, \ldots, \alpha_k)$.

Proof. For admissible $\omega = (\alpha_0, \ldots, \alpha_{k+1})$, separate the last factor in (4):

$$\theta^{(k+1)} = \|v\|_1 \operatorname{sgn}(v_{\alpha_0}) \prod_{s=0}^{k-1} \Big(d_{\alpha_s} \operatorname{sgn}(a_{\alpha_s \alpha_{s+1}}) \Big) \Big(d_{\alpha_k} \operatorname{sgn}(a_{\alpha_k \alpha_{k+1}}) h_{\alpha_{k+1}} \Big).$$

The edge $\alpha_k \to \alpha_{k+1}$ is a reachable nonzero edge, so the eigen-triple condition gives $d_{\alpha_k} \operatorname{sgn}(a_{\alpha_k \alpha_{k+1}}) h_{\alpha_{k+1}} = \lambda h_{\alpha_k}$. Substituting,

$$\theta^{(k+1)} = \lambda\, \|v\|_1 \operatorname{sgn}(v_{\alpha_0}) \prod_{s=0}^{k-1} \Big(d_{\alpha_s} \operatorname{sgn}(a_{\alpha_s \alpha_{s+1}}) \Big) h_{\alpha_k} = \lambda\, \theta^{(k)}.$$

□

Pathwise recurrence alone fixes only the ratios of consecutive moments; to pin down each moment it must be combined with the no-step compatibility condition (Definition 1). Together, they give the strong sufficient criterion.

Theorem 5 (Strong all-step zero-variance criterion)**.** *Suppose (A, v, h) is an eigen-triple with eigenvalue λ and that there exists $\beta \in \mathbb{R}$ with*

$$\operatorname{sgn}(v_i) h_i = \beta \qquad \text{for every } i \in S_v.$$

Then

$$\theta^{(k)} = \|v\|_1\, \beta\, \lambda^k \qquad \text{almost surely, for every } k \geq 0,$$

and in particular (A, v, h) has all-step zero variance.

Proof. For $k = 0$, (4) and the compatibility hypothesis give $\theta^{(0)} = \|v\|_1 \operatorname{sgn}(v_{\alpha_0}) h_{\alpha_0} = \|v\|_1 \beta$ on every admissible path. By Proposition 9, $\theta^{(k+1)} = \lambda \theta^{(k)}$ along every admissible

prefix, so by induction $\theta^{(k)} = \lambda^k \theta^{(0)} = \|v\|_1 \beta \lambda^k$ almost surely. The right-hand side is constant, whence $\mathrm{Var}\{\theta^{(k)}\} = 0$ for every k. □

Corollary 3. *Under the hypotheses of Theorem 5, (A, v, h) is levelwise path-balanced with $\beta_k = \beta\lambda^k$.*

Proof. By Theorem 5, $\theta^{(k)} = \|v\|_1 \beta \lambda^k$ almost surely; comparing with $\theta^{(k)} = \|v\|_1 \beta_k$ from Theorem 4 gives $\beta_k = \beta\lambda^k$. □

Example 2.

(i). *If $A \geq 0$ with $A\mathbf{1} = c\mathbf{1}$, $c > 0$, and $v = n^{-1}\mathbf{1}$, $h = \mathbf{1}$, then $d_i = c$ for every row, so $d_i \operatorname{sgn}(a_{ij}) h_j = c = c\, h_i$: an eigen-triple with $\lambda = c$. No-step compatibility holds with $\beta = 1$, so Theorem 5 yields $\theta^{(k)} = c^k$ almost surely, recovering Corollary 1.*

(ii). *For $A = \left(\begin{smallmatrix} 0 & 2 \\ 8 & 0 \end{smallmatrix}\right)$ and $h = \left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)$ the row sums $d_1 = 2$, $d_2 = 8$ differ, yet $d_1 \operatorname{sgn}(a_{12}) h_2 = 4 = 4h_1$ and $d_2 \operatorname{sgn}(a_{21}) h_1 = 8 = 4h_2$, so (A, v, h) is an eigen-triple with $\lambda = 4$ for any v supported on $\{1, 2\}$; indeed $Ah = 4h$. The all-step property therefore holds for a matrix that is not row-balanced.*

(iii). *For $A = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 2 \end{smallmatrix}\right)$ and $h = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ one has $d_1 = d_2 = 2$ and, on the reachable edges $1 \to 1$, $1 \to 2$, $2 \to 2$,*

$$2(+1)(1) = 2 = 2h_1, \qquad 2(-1)(-1) = 2 = 2h_1, \qquad 2(+1)(-1) = -2 = 2h_2,$$

an eigen-triple with $\lambda = 2$. Mixed signs, excluded in the matrix-only setting $h = \mathbf{1}$, become admissible once h is allowed to vary.

6.1. Eigenvalue-Free Characterization

Definition 5 presupposes a value λ. In practice, λ is exactly the quantity one is trying to compute, so a criterion that can be checked from (A, h) alone—without knowing λ in advance—is preferable. We add the nondegeneracy hypothesis $h_i \neq 0$ on the reachable part and replace the prescribed eigenvalue by the requirement that a single local ratio coincide on every edge.

Definition 6. *Assume $h_i \neq 0$ for every $i \in R(A, v)$. For every reachable nonzero edge $i \to j$ define*

$$\rho_{ij}(A, h) \leftrightharpoons \frac{d_i \operatorname{sgn}(a_{ij})\, h_j}{h_i}.$$

The triple (A, v, h) satisfies the edge-ratio condition if

$$\rho_{ij}(A, h) = \rho_{i'j'}(A, h)$$

for all reachable nonzero edges $i \to j$ and $i' \to j'$.

The edge-ratio condition refers only to absolute row sums, the signs of the nonzero entries, the reachability graph, and h. It contains no eigenvalue. The next theorem shows it is equivalent both to the eigen-triple condition and to the pathwise geometric recurrence $\theta^{(k+1)} = \lambda\theta^{(k)}$—the edgewise, pathwise, and recovered-eigenvalue faces of the same property—with λ produced as output rather than required as input.

Theorem 6. *Assume the standing hypotheses together with $h_i \neq 0$ for every $i \in R(A, v)$. Then the following are equivalent.*

(i). *There exists $\lambda \in \mathbb{R}$ such that (A, v, h) is an eigen-triple with eigenvalue λ.*

(ii). There exists $\lambda \in \mathbb{R}$ such that $\theta^{(k+1)} = \lambda\, \theta^{(k)}$ along every admissible coupled path prefix, i.e., for every $\omega = (\alpha_0, \ldots, \alpha_{k+1}) \in \Omega_{k+1}(A, v)$ with $\theta^{(k)}$ evaluated on the prefix $(\alpha_0, \ldots, \alpha_k)$.
(iii). (A, v, h) satisfies the edge-ratio condition.

When any of these holds, the values of λ in (i) and (ii) coincide, are unique, and equal the common edge ratio,

$$\lambda = \rho_{ij}(A, h) \qquad \text{for any reachable nonzero edge } i \to j.$$

Proof. Given that $v \neq 0$ and $d_i > 0$ on the reachable part, every reachable row has at least one nonzero outgoing edge; hence the set of reachable nonzero edges is nonempty, and on it the ratios $\rho_{ij}(A, h)$ are well defined (here $h_i \neq 0$, and $h_j \neq 0$ since j is itself reachable). We prove the cycle (i)⇒(ii)⇒(iii)⇒(i).

(i)⇒(ii). This is Proposition 9: an eigen-triple with eigenvalue λ satisfies $\theta^{(k+1)} = \lambda\theta^{(k)}$ along every admissible coupled path prefix.

(ii)⇒(iii). Let λ be as in (ii) and fix a reachable nonzero edge $i \to j$. Since $i \in R(A, v)$ there is an admissible prefix $(\alpha_0, \ldots, \alpha_k)$ with $\alpha_k = i$; appending j yields an admissible path $(\alpha_0, \ldots, \alpha_k, j) \in \Omega_{k+1}(A, v)$. By the collapsed form (4), the two estimators on this coupled prefix satisfy $\theta^{(k+1)} = \theta^{(k)}\, d_i \operatorname{sgn}(a_{ij}) h_j / h_i$, while (ii) gives $\theta^{(k+1)} = \lambda\theta^{(k)}$. Since $\|v\|_1 > 0$, $d_{\alpha_s} > 0$, and $h_i \neq 0$, we have $\theta^{(k)} \neq 0$; dividing yields $\rho_{ij}(A, h) = d_i \operatorname{sgn}(a_{ij}) h_j / h_i = \lambda$. The edge $i \to j$ was arbitrary, so every edge ratio equals λ: the edge-ratio condition.

(iii)⇒(i). Let λ denote the common value of the edge ratios, well defined by (iii) and the nonemptiness above. For every reachable nonzero edge $i \to j$ we have $d_i \operatorname{sgn}(a_{ij}) h_j / h_i = \lambda$, and multiplying by h_i gives $d_i \operatorname{sgn}(a_{ij}) h_j = \lambda h_i$. Hence (A, v, h) is an eigen-triple with eigenvalue λ.

Coincidence and uniqueness. The three constructions produce the same number: (i)⇒(ii) carries the eigenvalue λ unchanged into the recurrence, and (ii)⇒(iii)⇒(i) returns it as the common edge ratio. If (A, v, h) were an eigen-triple with eigenvalues λ and λ', then on any reachable nonzero edge $\lambda h_i = d_i \operatorname{sgn}(a_{ij}) h_j = \lambda' h_i$ with $h_i \neq 0$, so $\lambda = \lambda'$. Thus the value in (i)–(iii) is unique and equals $\rho_{ij}(A, h)$. □

Remark 8. *The equivalence isolates the role of the eigenvalue. Definition 5 carries λ as a hypothesis; the edge-ratio condition is a closed condition on (A, h) from which λ is recovered as the common edge ratio. By Proposition 8, this common value is automatically an eigenvalue of the reachable restriction of A, with eigenvector h—but that is a conclusion, not an input. Combined with no-step compatibility, Theorem 6 turns the strong criterion of Theorem 5 into a test requiring no prior spectral information: verify $h_i \neq 0$ and constancy of the edge ratios, read off λ, and check that $\operatorname{sgn}(v_i) h_i$ is constant on S_v.*

6.2. Edge Defects for Near-Eigen-Triples

Proposition 8 and its eigenvalue-free refinement describe the *exact* eigen-triple. It is natural to ask how the conclusions degrade when h only approximately satisfies the edgewise condition—when, in the language of the remark after Proposition 8, h strays from being MAO-observably an eigenvector. The departure is measured edgewise, and the eigen-triple is recovered throughout as the defect-free case.

Definition 7. *For $\lambda \in \mathbb{R}$ and every reachable nonzero edge $i \to j$ (that is, $i \in R(A, v)$ and $a_{ij} \neq 0$) the edge defect is*

$$\varepsilon_{ij} \leftrightharpoons d_i \operatorname{sgn}(a_{ij})\, h_j - \lambda h_i.$$

Proposition 10. *(A, v, h) is an eigen-triple with eigenvalue λ if and only if the edge defect ε_{ij} is 0 for every reachable nonzero edge.*

The eigenvector conclusion degrades by a weighted average of these defects, never by amplification: the residual at a row is the row-mean of its edge defects.

Lemma 1. *Under the standing hypotheses, for every $\lambda \in \mathbb{R}$ and every $i \in R(A, v)$,*

$$(Ah)_i - \lambda h_i = \sum_{j:\, a_{ij} \neq 0} w_{ij}\, \varepsilon_{ij} \;\leftrightharpoons\; \bar{\varepsilon}_i, \qquad w_{ij} \leftrightharpoons \frac{|a_{ij}|}{d_i},$$

where the weights $w_{ij} \geq 0$ satisfy $\sum_j w_{ij} = 1$. In particular $\left|(Ah)_i - \lambda h_i\right| \leq \max_{j:\, a_{ij} \neq 0} |\varepsilon_{ij}|$, and the defect-free case $\varepsilon \equiv 0$ recovers Proposition 8.

Proof. Fix a reachable row i. From Definition 7, $\operatorname{sgn}(a_{ij}) h_j = (\lambda h_i + \varepsilon_{ij})/d_i$ for each j with $a_{ij} \neq 0$. Hence

$$(Ah)_i = \sum_{j:\, a_{ij} \neq 0} |a_{ij}| \operatorname{sgn}(a_{ij}) h_j = \frac{1}{d_i} \sum_{j:\, a_{ij} \neq 0} |a_{ij}| \left(\lambda h_i + \varepsilon_{ij}\right) = \lambda h_i + \sum_{j:\, a_{ij} \neq 0} w_{ij} \varepsilon_{ij},$$

using $\sum_{j:\, a_{ij} \neq 0} |a_{ij}| = d_i$. The weights $w_{ij} = |a_{ij}|/d_i$ are nonnegative and sum to 1, giving the convex-combination bound. □

At the pathwise level the defect enters the geometric recurrence of Proposition 9 additively. Write, for an admissible prefix $\omega = (\alpha_0, \ldots, \alpha_k)$,

$$W_k(\omega) \leftrightharpoons \operatorname{sgn}(v_{\alpha_0}) \prod_{s=0}^{k-1} \Big(d_{\alpha_s} \operatorname{sgn}(a_{\alpha_s \alpha_{s+1}}) \Big), \qquad \text{so that} \qquad \theta^{(k)} = \|v\|_1\, W_k\, h_{\alpha_k}$$

by (4), with $W_0 = \operatorname{sgn}(v_{\alpha_0})$ and $|W_k| = \prod_{s=0}^{k-1} d_{\alpha_s}$.

Proposition 11 (Perturbed recurrence and its closed form)**.** *Under the standing hypotheses, for every $\lambda \in \mathbb{R}$ the estimator satisfies, along every admissible coupled path prefix,*

$$\theta^{(k+1)} = \lambda\, \theta^{(k)} + \|v\|_1\, W_k\, \varepsilon_{\alpha_k \alpha_{k+1}},$$

and hence, for every $\omega = (\alpha_0, \ldots, \alpha_k) \in \Omega_k(A, v)$,

$$\theta^{(k)}(\omega) = \|v\|_1 \Big[\lambda^k \operatorname{sgn}(v_{\alpha_0}) h_{\alpha_0} + \sum_{s=0}^{k-1} \lambda^{\,k-1-s}\, W_s(\omega)\, \varepsilon_{\alpha_s \alpha_{s+1}} \Big].$$

The defect-free case $\varepsilon \equiv 0$ recovers Proposition 9.

Proof. For admissible $\omega = (\alpha_0, \ldots, \alpha_{k+1})$, factor the last edge: $W_{k+1} = W_k\, d_{\alpha_k} \operatorname{sgn}(a_{\alpha_k \alpha_{k+1}})$, so by (4), $\theta^{(k+1)} = \|v\|_1 W_k\, d_{\alpha_k} \operatorname{sgn}(a_{\alpha_k \alpha_{k+1}}) h_{\alpha_{k+1}}$. The edge $\alpha_k \to \alpha_{k+1}$ is reachable and nonzero, so Definition 7 gives $d_{\alpha_k} \operatorname{sgn}(a_{\alpha_k \alpha_{k+1}}) h_{\alpha_{k+1}} = \lambda h_{\alpha_k} + \varepsilon_{\alpha_k \alpha_{k+1}}$. Substituting and using $\theta^{(k)} = \|v\|_1 W_k h_{\alpha_k}$ yields the recurrence. Unrolling from $\theta^{(0)} = \|v\|_1 \operatorname{sgn}(v_{\alpha_0}) h_{\alpha_0} = \|v\|_1 W_0 h_{\alpha_0}$ gives the closed form by induction. □

When the no-step compatibility hypothesis of Theorem 5 holds, the leading term is the constant $\|v\|_1 \beta \lambda^k$, so all randomness in $\theta^{(k)}$ comes from the defect sum. The standard deviation is therefore controlled to first order in the defect, and zero variance degrades continuously.

Proposition 12 (First-order standard-deviation bound). *Suppose (A, v, h) satisfies the standing hypotheses and the no-step compatibility condition $\operatorname{sgn}(v_i)h_i = \beta$ for every $i \in S_v$. Put $\varepsilon \coloneqq \max\{|\varepsilon_{ij}| : i \to j$ reachable nonzero$\}$ and $d_{\max} \coloneqq \max_{i \in R(A,v)} d_i$. Then for every $k \geq 0$,*

$$\sqrt{\operatorname{Var}\{\theta^{(k)}\}} \leq \|v\|_1 \sum_{s=0}^{k-1} |\lambda|^{k-1-s} \sqrt{\operatorname{Var}\{W_s\, \varepsilon_{\alpha_s \alpha_{s+1}}\}} \leq \|v\|_1\, \varepsilon \sum_{s=0}^{k-1} |\lambda|^{k-1-s} d_{\max}^s,$$

and, when $d_{\max} \neq |\lambda|$, the last sum equals $(d_{\max}^k - |\lambda|^k)/(d_{\max} - |\lambda|)$. In particular, the standard deviation is first order, $\sqrt{\operatorname{Var}\{\theta^{(k)}\}} = O(\varepsilon)$, so the variance is second order, $\operatorname{Var}\{\theta^{(k)}\} = O(\varepsilon^2)$; and $\varepsilon = 0$ recovers the all-step zero-variance conclusion of Theorem 5.

Proof. By Proposition 11 and the compatibility hypothesis, $\lambda^k \operatorname{sgn}(v_{\alpha_0})h_{\alpha_0} = \beta\lambda^k$ is a constant, so $\theta^{(k)} - \|v\|_1 \beta\lambda^k = \|v\|_1 \sum_{s=0}^{k-1} \lambda^{k-1-s} W_s \varepsilon_{\alpha_s \alpha_{s+1}}$. Variance is invariant under subtracting a constant, so $\sqrt{\operatorname{Var}\{\theta^{(k)}\}} = \|v\|_1 \left\| \sum_s \lambda^{k-1-s} W_s \varepsilon_{\alpha_s \alpha_{s+1}} \right\|_{L_0^2}$, where $\|\cdot\|_{L_0^2}$ is the centered L^2 norm. The triangle inequality for that norm gives the first bound. For the second, the centered L^2 norm is dominated by the uncentered one, and $|W_s \varepsilon_{\alpha_s \alpha_{s+1}}| \leq d_{\max}^s \varepsilon$ pointwise since $|W_s| = \prod_{r<s} d_{\alpha_r} \leq d_{\max}^s$; hence $\sqrt{\operatorname{Var}\{W_s \varepsilon_{\alpha_s \alpha_{s+1}}\}} \leq d_{\max}^s \varepsilon$. Summing the geometric series gives the closed form. □

Remark 9. *The two conclusions respond to different features of the defect. The eigenvector residual $\bar{\varepsilon}_i$ of Lemma 1 depends only on the row-mean of $\{\varepsilon_{ij}\}_j$, whereas the variance depends—already through the $s = 0$ term $\theta^{(1)} = \|v\|_1 \operatorname{sgn}(v_{\alpha_0})(\lambda h_{\alpha_0} + \varepsilon_{\alpha_0 \alpha_1})$—on the full within-row dispersion $\varepsilon_{ij} - \bar{\varepsilon}_i$. Consequently, h can be an exact eigenvector of the reachable restriction, $\bar{\varepsilon}_i = 0$ for every $i \in R(A, v)$ so that $(Ah)_i = \lambda h_i$, while $\operatorname{Var}\{\theta^{(k)}\} > 0$: the within-row spread is invisible to Ah yet inflates the estimator. This is the precise sense in which being MAO-observably an eigenvector—the vanishing of every ε_{ij}—is strictly stronger than being an eigenvector on $R(A, v)$, which forces only the row-means to vanish.*

7. Quotient-Level Variance Optimality

The results of Sections 5.1 and 5.3 concern the primitive MAO moment estimators $\theta^{(k)}$, which estimate the bilinear moments $(v, A^k h)$. The RMC algorithm does not output these moments directly; it outputs the finite-truncation quotient $\widehat{\lambda}_{K,N}^{(m)}$, a ratio of weighted sums of the $\theta^{(k)}$. A ratio can be deterministic even when its numerator and denominator are random, so the variance of the quotient is a question distinct from the variance of the individual moments. We isolate two quotient-level notions: *quotient zero-variance*, meaning the quotient is deterministic, and *quotient exactness*, meaning the deterministic value is the target eigenvalue.

For fixed $m \geq 1$ and $K \geq 0$, write the single-chain numerator and denominator contributions

$$X_K^{(m)} \coloneqq \sum_{k=0}^{K} w_k^{(m)} \theta^{(k+1)}, \qquad Y_K^{(m)} \coloneqq \sum_{k=0}^{K} w_k^{(m)} \theta^{(k)},$$

so that for N independent chains

$$\widehat{\lambda}_{K,N}^{(m)} = \frac{\sum_{r=1}^{N} X_{K,r}^{(m)}}{\sum_{r=1}^{N} Y_{K,r}^{(m)}},$$

whenever the denominator is nonzero.

7.1. Quotient Zero-Variance

Definition 8. *The finite-truncation RMC quotient has quotient zero-variance with value $\tau \in \mathbb{R}$ if*

$$\widehat{\lambda}_{K,N}^{(m)} = \tau$$

almost surely for every $N \geq 1$, whenever the denominator is nonzero.

Proposition 13 (All-step zero variance implies quotient zero-variance)**.** *Suppose $\theta^{(k)} = c_k$ almost surely for $k = 0, \ldots, K+1$, and that $\sum_{k=0}^{K} w_k^{(m)} c_k \neq 0$. Then, the finite-truncation quotient is deterministic, with value*

$$\tau_{K,m} = \frac{\sum_{k=0}^{K} w_k^{(m)} c_{k+1}}{\sum_{k=0}^{K} w_k^{(m)} c_k}.$$

Proof. If each $\theta^{(k)}$ is deterministic, then $X_K^{(m)}$ and $Y_K^{(m)}$ are deterministic, and so are their sums over independent chains. The quotient is therefore deterministic whenever its denominator is nonzero, and its value is obtained by substituting $\theta^{(k)} = c_k$. □

Theorem 7. *Assume $Y_K^{(m)} \neq 0$ almost surely. The following are equivalent.*

(i). The single-chain quotient $X_K^{(m)} / Y_K^{(m)}$ has zero variance.

(ii). There exists $\tau \in \mathbb{R}$ with $X_K^{(m)} = \tau Y_K^{(m)}$ almost surely.

When either holds, $\widehat{\lambda}_{K,N}^{(m)} = \tau$ for every $N \geq 1$ whenever the sample denominator is nonzero.

Proof. If $X_K^{(m)} / Y_K^{(m)}$ has zero variance, it equals a constant τ almost surely; since $Y_K^{(m)} \neq 0$ almost surely this is equivalent to $X_K^{(m)} = \tau Y_K^{(m)}$ almost surely. Conversely, that identity gives $X_K^{(m)} / Y_K^{(m)} = \tau$. For independent chains $r = 1, \ldots, N$ the same identity yields $\sum_r X_{K,r}^{(m)} = \tau \sum_r Y_{K,r}^{(m)}$, so the sample quotient equals τ whenever the denominator is nonzero. □

Remark 10. *Quotient zero-variance is weaker than all-step zero variance: $X_K^{(m)}$ and $Y_K^{(m)}$ may both fluctuate, yet if they fluctuate in exact proportion their ratio is deterministic.*

7.2. Quotient Exactness

Quotient zero-variance only asserts that the quotient is deterministic; it does not identify the deterministic value. Quotient exactness is the stronger property that the finite-sample quotient equals a prescribed eigenvalue.

Definition 9. *The RMC quotient is finite-sample exact for λ if $\widehat{\lambda}_{K,N}^{(m)} = \lambda$ almost surely for every $m \geq 1$, $K \geq 0$, and $N \geq 1$, whenever the denominator is nonzero.*

The eigen-triple framework of Section 6 delivers exactness directly. By Proposition 9, an eigen-triple with eigenvalue λ satisfies $\theta^{(k+1)} = \lambda \theta^{(k)}$ along every admissible coupled path prefix, and by Theorem 6 this is equivalent to the edge-ratio condition (Definition 6), in which λ is recovered as the common edge ratio rather than assumed.

Theorem 8 (Quotient Exactness)**.** *Suppose (A, v, h) satisfies the edge-ratio condition of Definition 6, and let λ be the common value of the edge ratios $\rho_{ij}(A, h)$. Then the coupled-prefix RMC quotient is finite-sample exact for λ: for every $m \geq 1$, $K \geq 0$, and $N \geq 1$,*

$$\widehat{\lambda}_{K,N}^{(m)} = \lambda$$

whenever the denominator is nonzero.

Proof. By Theorem 6, the triple is an eigen-triple with eigenvalue λ, so Proposition 9 gives $\theta^{(k+1)} = \lambda\theta^{(k)}$ along every admissible coupled path prefix. Hence for each sampled chain r,

$$X_{K,r}^{(m)} = \sum_{k=0}^{K} w_k^{(m)} \theta_r^{(k+1)} = \lambda \sum_{k=0}^{K} w_k^{(m)} \theta_r^{(k)} = \lambda Y_{K,r}^{(m)}.$$

Summing over $r = 1, \ldots, N$ gives $\sum_r X_{K,r}^{(m)} = \lambda \sum_r Y_{K,r}^{(m)}$, so the quotient equals λ whenever the denominator is nonzero. □

Corollary 4. *Fix $K \geq 0$. If the edge ratio $\rho_{ij}(A, h)$ is constant, with common value λ, on every edge lying on an admissible path of length $\leq K{+}1$, then $\theta^{(k+1)} = \lambda\theta^{(k)}$ along coupled prefixes for $k = 0, \ldots, K$, and hence $\widehat{\lambda}_{K,N}^{(m)} = \lambda$ for every $m \geq 1$ and $N \geq 1$ whenever the denominator is nonzero. Adjoining no-step compatibility yields, in addition, $(K{+}1)$-step zero variance (Proposition 6 with K replaced by $K{+}1$).*

Proof. The numerator and denominator at truncation K involve only $\theta^{(0)}, \ldots, \theta^{(K+1)}$. For each $k = 0, \ldots, K$, the recurrence $\theta^{(k+1)} = \lambda\theta^{(k)}$ uses the last edge of an admissible path of length $k{+}1$, hence only edges lying on admissible paths of length at most $K{+}1$. On these edges the edge-ratio hypothesis gives, as in Proposition 9, $\theta^{(k+1)} = \lambda\theta^{(k)}$ for $k = 0, \ldots, K$, so $X_{K,r}^{(m)} = \lambda Y_{K,r}^{(m)}$ for each chain and the quotient equals λ. The last clause is Proposition 6 applied at level $K{+}1$. □

Remark 11. *Finite-sample exactness requires $\theta^{(k)}$ and $\theta^{(k+1)}$ to be computed from coupled prefixes of the same sampled trajectory. If the consecutive powers are estimated from independent chains the pathwise cancellation is lost; the corresponding expectations remain proportional, but the finite-sample quotient need not be exact.*

7.3. Perturbation Analysis

Sections 7.1 and 7.2 are exact dichotomies: the quotient is either deterministic (proportionality $X = \tau Y$) or not. In practice (A, v, h) rarely lies exactly in the eigen-triple class, and one wants a quantitative measure of how far the quotient fluctuates. The natural object is the linearized (delta-method) variance of the ratio, which we record here and connect to the edge defects of Section 6.2: it is the quotient-level analogue of the moment-level departure analysis there.

Assume $\mathbb{E}Y_K^{(m)} \neq 0$ and set

$$\eta_K^{(m)} \coloneqq \frac{\mathbb{E}X_K^{(m)}}{\mathbb{E}Y_K^{(m)}}, \qquad \mathcal{D}_K^{(m)} \coloneqq \mathbb{E}Y_K^{(m)}, \qquad U_K^{(m)} \coloneqq X_K^{(m)} - \eta_K^{(m)} Y_K^{(m)},$$

so that $\mathbb{E}U_K^{(m)} = 0$. The *ratio variance* is

$$\sigma_{\text{ratio}}^2 \coloneqq \frac{\operatorname{Var}\{U_K^{(m)}\}}{|\mathcal{D}_K^{(m)}|^2} = \frac{\operatorname{Var}\{X_K^{(m)} - \eta_K^{(m)} Y_K^{(m)}\}}{|\mathcal{D}_K^{(m)}|^2}.$$

Proposition 14. *Let the chains $r = 1, \ldots, N$ be i.i.d. and $\mathcal{D}_K^{(m)} \neq 0$.*

(i). If $Y_K^{(m)}$ is almost surely constant, then for every $N \geq 1$

$$\operatorname{Var}\{\widehat{\lambda}_{K,N}^{(m)}\} = \frac{\sigma_{\text{ratio}}^2}{N} \qquad \textit{exactly.}$$

(ii). In general, as $N \to \infty$,

$$\mathrm{Var}\{\widehat{\lambda}_{K,N}^{(m)}\} = \frac{\sigma_{\mathrm{ratio}}^2}{N}(1 + o(1)),$$

the leading term being the first-order delta-method approximation, valid when the relative denominator fluctuation $\mathrm{Var}\{Y_K^{(m)}\}/|\mathcal{D}_K^{(m)}|^2$ *is small.*

Proof. (i) If $Y_K^{(m)} \equiv y_0$ almost surely, then $\widehat{\lambda}_{K,N}^{(m)} = (\sum_r X_{K,r}^{(m)})/(Ny_0)$, so $\mathrm{Var}\{\widehat{\lambda}_{K,N}^{(m)}\} = \mathrm{Var}\{X_K^{(m)}\}/(Ny_0^2)$. Here, $\eta_K^{(m)} Y_K^{(m)} = \mathbb{E}X_K^{(m)}$, hence $U_K^{(m)} = X_K^{(m)} - \mathbb{E}X_K^{(m)}$, $\mathrm{Var}\{U_K^{(m)}\} = \mathrm{Var}\{X_K^{(m)}\}$ and $|\mathcal{D}_K^{(m)}|^2 = y_0^2$; the two expressions coincide.

(ii) Writing $\bar{X}_N, \bar{Y}_N$ for the chain averages, the delta method about $(\mathbb{E}X, \mathbb{E}Y)$ gives $\widehat{\lambda}_{K,N}^{(m)} = \bar{X}_N/\bar{Y}_N = \eta_K^{(m)} + (\bar{X}_N - \eta_K^{(m)}\bar{Y}_N)/\mathbb{E}Y_K^{(m)} + o_p(N^{-1/2})$, whence $\mathrm{Var}\{\widehat{\lambda}_{K,N}^{(m)}\} = \mathrm{Var}\{U_K^{(m)}\}/(N|\mathcal{D}_K^{(m)}|^2) + o(N^{-1}) = \sigma_{\mathrm{ratio}}^2/N + o(N^{-1})$. □

The exact regime (i) is not vacuous: at $K = 0$ the denominator is the no-step estimator, which is deterministic precisely under no-step compatibility. This makes the balanced-matrix failures of the balanced-matrix section exact quotient-level statements.

Corollary 5. *At* $K = 0$, *one has* $w_0^{(m)} = 1$, $X_0^{(m)} = \theta^{(1)}$, *and* $Y_0^{(m)} = \theta^{(0)}$. *If* (A, v, h) *is no-step compatible, so that* $\theta^{(0)} = \|v\|_1\beta$ *is deterministic, then* $Y_0^{(m)}$ *is constant and Proposition 14(i) gives, exactly,*

$$\mathrm{Var}\{\widehat{\lambda}_{0,N}^{(m)}\} = \frac{1}{N}\frac{\mathrm{Var}\{\theta^{(1)}\}}{\|v\|_1^2\beta^2}.$$

For the balanced-matrix test vectors $v = n^{-1}\mathbf{1}$, $h = \mathbf{1}$ *(so* $\beta = 1$, $\|v\|_1 = 1$, $\theta^{(0)} = 1$*) this reduces to* $\mathrm{Var}\{\theta^{(1)}\}/N$. *Hence, the one-step variances of Theorem 2 and Corollary 2 are, up to the factor* $1/N$, *exactly the* $K = 0$ *RMC quotient variance:*

$$\mathrm{Var}\{\widehat{\lambda}_{0,N}^{(m)}\} = \frac{1}{N}\left(\tfrac{1}{n}\textstyle\sum_{i=1}^n d_i^2 - c^2\right) \quad (\textit{signed}, A\mathbf{1} = c\mathbf{1}),$$

and $\frac{1}{N}\left(\frac{1}{n}\sum_i r_i^2 - (\frac{1}{n}\sum_i r_i)^2\right)$ *in the nonnegative unequal-row-sum case.*

For $K \geq 1$ the denominator fluctuates and $\sigma_{\mathrm{ratio}}^2$ is the first-order approximation of Proposition 14(ii). Its dependence on the data is governed by the same edge defects that control the moments. Recall from Section 6.2 the edge defect ε_{ij} (Definition 7) and the path weight W_k with $\theta^{(k)} = \|v\|_1 W_k h_{\alpha_k}$.

Proposition 15. *Suppose* (A, v, h) *is no-step compatible with* $\mathrm{sgn}(v_i)h_i = \beta$ *on* S_v, *and let* ε_{ij} *be the edge defects (Definition 7) for a candidate eigenvalue* λ. *Then, the centered residual is, to first order in the defect,*

$$U_K^{(m)} = \|v\|_1\left(S_K^{(m)} - \mathbb{E}S_K^{(m)}\right) + O(\varepsilon^2), \qquad S_K^{(m)} \leftrightharpoons \sum_{k=0}^{K} w_k^{(m)} W_k\, \varepsilon_{\alpha_k\alpha_{k+1}},$$

and consequently

$$\sigma_{\mathrm{ratio}}^2 = \frac{\|v\|_1^2}{|\mathcal{D}_K^{(m)}|^2}\,\mathrm{Var}\left\{S_K^{(m)}\right\} + O(\varepsilon^3).$$

In particular $\sigma_{\mathrm{ratio}}^2 = O(\varepsilon^2)$, *and it vanishes at the eigen-triple, recovering the exactness of Theorem 8.*

Proof. By Proposition 11, under no-step compatibility $\theta^{(k)} = \|v\|_1 \beta \lambda^k + \delta^{(k)}$ with $\delta^{(k)} = \|v\|_1 \sum_{s=0}^{k-1} \lambda^{k-1-s} W_s \varepsilon_{\alpha_s \alpha_{s+1}}$, and $\delta^{(k+1)} - \lambda \delta^{(k)} = \|v\|_1 W_k \varepsilon_{\alpha_k \alpha_{k+1}}$. The deterministic parts satisfy $\bar{X}_K^{(m)} = \lambda \bar{Y}_K^{(m)}$, so $\eta_K^{(m)} = \lambda + O(\varepsilon)$, and $X_K^{(m)} - \eta_K^{(m)} Y_K^{(m)} = \sum_k w_k^{(m)} (\delta^{(k+1)} - \lambda \delta^{(k)}) - \mathbb{E}[\cdot] + O(\varepsilon^2) = \|v\|_1 (S_K^{(m)} - \mathbb{E} S_K^{(m)}) + O(\varepsilon^2)$. Dividing by $|\mathcal{D}_K^{(m)}|^2$ and squaring gives the variance to $O(\varepsilon^3)$. If all reachable edge defects vanish, then $S_K^{(m)} = 0$ and the displayed leading term is zero. □

Remark 12. *The expansion is the quotient-level mirror of Proposition 12: both the moment standard deviation and the first-order quotient standard deviation are $O(\varepsilon)$ and are driven by the same weighted defect sums, differing only in the weights (λ^{k-1-s} at the moment level, $w_k^{(m)} W_k / \mathcal{D}_K^{(m)}$ at the quotient level). As there, the eigenvector residual sees only the row-mean of the defects while the variance sees their full within-row spread, so a quotient can inherit positive variance from misalignment of h even when $Ah = \lambda h$ holds on the reachable part.*

8. Independence of Exactness and All-Step Variance

8.1. Exactness Without Moment Zero-Variance

The quotient can be exact even when no individual moment estimator has zero variance, because a common random factor cancels in the ratio. The elementary structure is

$$X = \lambda Z, \qquad Y = Z,$$

with Z random: both X and Y may have positive variance, yet $X/Y = \lambda$ whenever $Y \neq 0$. The same cancellation drives the RMC quotient whenever $\theta^{(k+1)} = \lambda \theta^{(k)}$ pathwise, since then $X_{K,r}^{(m)} = \lambda Y_{K,r}^{(m)}$ for every sampled chain.

Example 3. *Take*

$$A = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad v = \tfrac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

This is the mixed-sign eigen-triple of Section 5.3 with $\lambda = 2$, so $\theta^{(k+1)} = 2\theta^{(k)}$ along every admissible path and the finite RMC quotient is exactly 2 whenever its denominator is nonzero. However, $\|v\|_1 = 1$ and

$$\theta^{(0)} = \|v\|_1 \operatorname{sgn}(v_{\alpha_0}) h_{\alpha_0} = \begin{cases} 1, & \alpha_0 = 1, \\ -1, & \alpha_0 = 2, \end{cases}$$

so $\operatorname{Var}\{\theta^{(0)}\} > 0$, and since $\theta^{(k)} = 2^k \theta^{(0)}$,

$$\operatorname{Var}\{\theta^{(k)}\} = 4^k \operatorname{Var}\{\theta^{(0)}\} > 0 \qquad \text{for every } k.$$

The quotient is nevertheless exact, because each numerator contribution is exactly twice the corresponding denominator contribution.

Thus, quotient exactness does not imply moment zero-variance. The full relationship between the notions is summarized in Section 9.

8.2. Zero-Variance Without Exactness

The converse gap is the one most easily overlooked in the RMC setting, where the quotient is built precisely to extract an eigenvalue: one expects that a quotient which has stopped fluctuating must have settled on the right answer. It need not. Zero variance is

a statement about the spread of the estimator, exactness a statement about its location; a quotient can be perfectly deterministic and deterministically wrong. Concretely, if every moment is deterministic, $\theta^{(k)} = c_k$ almost surely, then by Proposition 13 the quotient is the constant

$$\tau_{K,m} = \frac{\sum_{k=0}^{K} w_k^{(m)} c_{k+1}}{\sum_{k=0}^{K} w_k^{(m)} c_k}, \qquad c_k = (v, A^k h),$$

a finite-truncation resolvent Rayleigh quotient. This equals an eigenvalue for every K and m only when the moments are geometric, $c_{k+1} = \lambda c_k$—equivalently when (A, v, h) is an eigen-triple (Theorem 6). When h is not aligned with an eigenvector on the reachable part the equality is destroyed, while the quotient stays deterministic.

Example 4. *Take*

$$A = \begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix}, \qquad v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

From $S_v = \{1\}$, the walk is forced, $1 \to 2 \to 1 \to 2 \to \cdots$, because each reachable row has a single nonzero entry. There is thus exactly one admissible path of each length, so every $\theta^{(k)}$ is constant on its support and (A, v, h) has all-step zero variance; in particular the quotient $\widehat{\lambda}_{K,N}^{(m)}$ is deterministic for every m, K, N. The moments are

$$c_k = (v, A^k h) = (1, 1, 4, 4, 16, 16, \dots), \qquad \textit{i.e. } c_{2j} = c_{2j+1} = 4^j,$$

so the consecutive ratios c_{k+1}/c_k alternate between 1 and 4 and never settle. The triple is not an eigen-triple: the edge ratios

$$\rho_{12} = d_1 \operatorname{sgn}(a_{12})\, h_2/h_1 = 1, \qquad \rho_{21} = d_2 \operatorname{sgn}(a_{21})\, h_1/h_2 = 4$$

differ, so the edge-ratio condition of Definition 6 fails. (A has eigenvalues ± 2 with eigenvectors $(1, \pm 2)$; the chosen $h = (1, 1)$ aligns with neither.)

The deterministic quotient is therefore a genuine constant that is not an eigenvalue. For $m = 1$,

$$\tau_{0,1} = 1, \qquad \tau_{1,1} = \frac{1 + 4q}{1 + q}, \qquad \tau_{2,1} = \frac{1 + 4q + 4q^2}{1 + q + 4q^2}, \quad \dots,$$

none of which equals ± 2 for any admissible q. The value drifts with K, m, and q, and even its $K \to \infty$ resolvent limit

$$\tau_{\infty,m} = \frac{(v, A(I - qA)^{-m} h)}{(v, (I - qA)^{-m} h)}$$

stays off the spectrum for every admissible q: for $m = 1$, it equals $(1 + 4q)/(1 + q)$, which reaches the dominant eigenvalue 2 only in the excluded singular limit $q \to \frac{1}{2} = 1/\rho(A)$. The quotient is zero-variance with a wrong value because h mixes the $+2$ and -2 eigencomponents and the ratio returns a biased Rayleigh average rather than a single eigenvalue.

Hence, quotient zero-variance does not imply quotient exactness: the estimator can concentrate on a deterministic value that is not the eigenvalue, through truncation (K-dependence), through m and q, or—as above—through misalignment of h. Exactness is recovered precisely by the edge-ratio (eigen-triple) condition, which forces $c_{k+1} = \lambda c_k$ and hence $\tau_{K,m} = \lambda$ for all K, m.

9. The Hierarchy of Variance-Optimality Notions

The preceding analysis produces five optimality notions presented in Table 1. They are related but not equivalent, because the primitive moment estimators $\theta^{(k)}$ and the final

quotient $\widehat{\lambda}_{K,N}^{(m)}$ are different random variables, and because determinism of an estimator is distinct from its correctness. Ordered from most local to most global:

(i). *No-step and single-step zero-variance* of $\theta^{(0)}$ and $\theta^{(1)}$ (Section 5.1): detects first-step degeneracy through the no-step and single-step compatibility conditions.

(ii). *K-step zero-variance* (Section 5.2): the moments $\theta^{(0)}, \ldots, \theta^{(K)}$ are deterministic; characterized by levelwise path-balance up to level K (Theorem 3). This is the finite notion that a truncated run can certify, with no-step and single-step as the cases $K = 0, 1$.

(iii). *All-step zero-variance* (Section 5.3): every $\theta^{(k)}$ is deterministic, removing stochastic error from every moment $(v, A^k h)$; characterized by levelwise path-balance (Theorem 4). It is exactly K-step zero-variance for every K, i.e., the limit $K \to \infty$ of the previous notion.

(iv). *Quotient zero-variance* (Section 7.1): the final RMC ratio is deterministic, even if the individual moments fluctuate.

(v). *Quotient exactness* (Section 7.2): the deterministic quotient value equals the target eigenvalue λ.

The first three concern the MAO moment estimators; the last two concern the RMC eigenvalue estimator. The moment-level notions are nested, and the valid implications across levels are

$$(K{+}1)\text{-step zero variance} \implies \text{quotient zero-variance at truncation } K$$
$$\text{quotient exactness} \implies \text{quotient zero-variance},$$

both whenever the denominator is nonzero; the first specializes to all-step $\implies$ quotient zero-variance as $K \to \infty$. The reverse and omitted implications fail:

- $(K{+}1)$-step zero variance $\Rightarrow$ K-step, strictly: balance up to one level does not propagate to the next (Example 1);
- quotient zero-variance $\not\Rightarrow$ all-step (nor $(K{+}1)$-step) zero variance, since numerator and denominator may share a random factor that cancels (Example 3);
- quotient zero-variance $\not\Rightarrow$ quotient exactness, since the deterministic value may be wrong (truncation, vector choice, or h not aligned with an eigenvector; Example 4);
- $(K{+}1)$-step zero variance $\not\Rightarrow$ quotient exactness, unless the deterministic moments satisfy the geometric relation $c_{k+1} = \lambda c_k$.

The single pathwise condition $\theta^{(k+1)} = \lambda\theta^{(k)}$—equivalently the edge-ratio condition of Definition 6—yields quotient exactness. Adjoining the no-step compatibility condition $\operatorname{sgn}(v_i)h_i = \beta$ on S_v makes $\theta^{(0)}$ deterministic and hence forces all-step zero variance (Theorem 5). The same conditions have finite-truncation counterparts: imposing the edge ratios only on the edges that lie on admissible paths of length $\leq K{+}1$ already forces $\theta^{(k+1)} = \lambda\theta^{(k)}$ for $k \leq K$, hence finite-sample exactness of $\widehat{\lambda}_{K,N}^{(m)}$ at truncation order K, and adjoining no-step compatibility forces $(K{+}1)$-step zero variance (Proposition 6 applied at level $K{+}1$). This gives the eigenvalue-free classification

Table 1. Eigenvalue-free conditions for optimality notions.

Property	Eigenvalue-Free Condition
Exact at truncation K	$\rho_{ij}(A, h)$ is constant on edges of admissible paths of length $\leq K + 1$.
$(K{+}1)$-step zero-variance triples	The same edge-ratio condition holds, and $\operatorname{sgn}(v_i)h_i$ is constant on S_v.
Quotient-exact triples	$\rho_{ij}(A, h)$ is constant on reachable nonzero edges.
All-step zero-variance triples	$\rho_{ij}(A, h)$ is constant on reachable nonzero edges, and $\operatorname{sgn}(v_i)h_i$ is constant on S_v.

The last two rows are the limits $K \to \infty$ of the first two.

To keep the notions distinct we reserve *moment zero-variance* for properties of $\theta^{(k)}$, *quotient zero-variance* for determinism of $\widehat{\lambda}_{K,N}^{(m)}$, and *quotient exactness* for $\widehat{\lambda}_{K,N}^{(m)} = \lambda$ almost surely.

10. Discussion

The variance structure theorem (Theorem 1) is the heart of the mechanism. By writing the variance of $\theta^{(k)}$ as the explicit difference $\|v\|_1(\overline{\overline{v}}, (D\overline{\overline{A}})^k h^{\circ 2}) - (v, A^k h)^2$, it turns every zero-variance question into an algebraic identity between a transport term and the squared moment. The second moment $\|v\|_1(\overline{\overline{v}}, (D\overline{\overline{A}})^k h^{\circ 2})$ is exactly the quantity whose minimization over permissible densities underlies the almost-optimality of the MAO scheme [16,26]; our identity records it as a closed form for general, possibly signed, A, and subtracts the mean to give the variance itself rather than the one-sided bound used to establish optimality. Every later characterization is read off this identity.

The hierarchy of Section 9 separates two properties that are easy to conflate. Determinism of an estimator is a statement about its spread; correctness is a statement about its location. At the moment level the two coincide only through unbiasedness, but at the quotient level they come apart twice over. A quotient can be deterministic while its numerator and denominator fluctuate, because a common random factor cancels in the ratio (Section 8.1), so quotient zero-variance is strictly weaker than moment zero-variance. And a deterministic quotient can be deterministically wrong, settling on a value that is not an eigenvalue (Section 8.2). The second gap is the one most easily overlooked in the RMC setting, where the quotient is built precisely to extract an eigenvalue and a stable output invites the inference that it has settled on the right answer.

The eigen-triple condition resolves both gaps at once. A single edgewise requirement, that the edge ratio $\rho_{ij}(A, h) = d_i \operatorname{sgn}(a_{ij}) h_j / h_i$ be constant on the reachable nonzero edges, forces the pathwise recurrence $\theta^{(k+1)} = \lambda \theta^{(k)}$ and hence finite-sample exactness of the quotient (Theorem 8), with the eigenvalue recovered as the common ratio rather than supplied in advance. The condition is genuinely stronger than h being an eigenvector: the eigenvector residual at a row is the row-mean of its edge defects (Lemma 1), whereas the variance sees the full within-row dispersion of those defects (Section 6.2). Consequently h can satisfy $Ah = \lambda h$ on the reachable part while the estimator still fluctuates. This is the precise content of the slogan that the eigen-triple asks h to be MAO-observably an eigenvector, not merely an eigenvector.

For practice the operative notion is neither the local nor the global one but the finite K-step condition. A truncated run forms the quotient from $\theta^{(0)}, \ldots, \theta^{(K+1)}$ alone, so $(K+1)$-step zero variance is exactly what makes every moment entering the quotient deterministic, and the eigen-triple need only hold on the edges of admissible paths of length at most $K+1$ (Corollary 4). When the data fall short of an exact eigen-triple, the degradation is quantitative and benign: both the moment standard deviation and the first-order quotient standard deviation are $O(\varepsilon)$ in the maximal edge defect ε, so the variance is $O(\varepsilon^2)$ and vanishes at the eigen-triple (Propositions 12 and 15). The edge-ratio condition is moreover cheap to check: Algorithm A1 (Appendix A) certifies it, and returns the eigenvalue, in a single pass over the nonzero entries of A, in time $O(|S_v| + \operatorname{nnz}(A))$.

Several limitations delimit the scope of these results. The analysis is for real A and real test vectors, and the conditions are phrased on the reachable part $R(A, v)$, so degenerate triples whose support traps the walk in a deterministic component fall outside the canonical classification (Theorem 1 and the remarks following it). Finite-sample exactness requires the consecutive powers $\theta^{(k)}$ and $\theta^{(k+1)}$ to be read from coupled prefixes of the same trajectory; estimating them from independent chains preserves proportionality only in expectation and destroys the pathwise cancellation. Finally, a deterministic quotient can be pulled off

the spectrum by truncation, by the resolvent order m, or by the parameter q, so quotient zero-variance certifies stability of the output but not its correctness; only the edge-ratio condition certifies both.

11. Conclusions

Resolvent Monte Carlo trades exact arithmetic for stochastic error, and the almost-optimal sampling scheme is the classical answer to making that error small. We have asked the sharper question of when the error vanishes, and answered it completely. The variance structure theorem gives an exact closed-form identity for the variance of the moment estimator of a general, possibly signed, matrix, and from it we have derived a full hierarchy of zero-variance notions: the local no-step and single-step conditions, the finite-truncation K-step condition that a practical run can certify, the global all-step condition, and the two quotient-level notions of zero-variance and exactness. We have characterized each notion, established the only implications that hold across levels, and exhibited the examples that separate the rest, thereby distinguishing determinism of the estimator from correctness of the eigenvalue it reports.

The structural core of the theory is the eigen-triple condition. A single edgewise requirement on the edge ratios forces the truncated quotient to equal the target eigenvalue in finite samples, recovers the eigenvalue as the common ratio rather than assuming it, and contains the classical balanced-row regime as a special case. Away from the exact condition the exactness degrades only to second order in a measurable edge defect, so the optimality is robust, and the condition is certified, together with the eigenvalue it yields, in a single linear-time pass over the sparsity pattern of the matrix.

These results turn the variance analysis of RMC from a one-sided optimality bound into an exact calculus of when, and how nearly, the estimator is deterministic and correct. Natural continuations include a numerical study of near-eigen-triple data against the predicted $O(\varepsilon^2)$ scaling, the use of the edge-ratio certificate to guide the choice of the input vectors v and h, and the extension of the zero-variance theory to resolvent Monte Carlo with a complex parameter, where the chains remain real but the weights become complex.

Author Contributions: Conceptualization, T.K. and I.T.D.; Methodology, T.K. and I.T.D.; Validation, T.K. and I.T.D.; Formal analysis, T.K. and I.T.D.; Investigation, T.K. and I.T.D.; Data curation, I.T.D.; Writing—original draft, T.K.; Writing—review and editing, I.T.D.; Supervision, I.T.D.; Project administration, I.T.D.; Funding acquisition, I.T.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. An Algorithm for the Edge-Ratio Condition

We give a direct procedure that decides whether a triple (A, v, h) satisfies the edge-ratio condition of Definition 6 and, when it does, returns the common edge ratio λ. By Theorem 6 this λ is the eigenvalue carried by the eigen-triple, by Proposition 8 it is an eigenvalue of the reachable restriction of A with eigenvector h, and by Theorem 8 the coupled RMC quotient is then finite-sample exact for λ. The test inspects only the reachable nonzero edges and runs in a single pass over the sparsity pattern of A.

Recall the data of Definition 6: $S_v = \{i : v_i \neq 0\}$, the reachable set $R(A, v)$ of indices reachable from S_v through nonzero entries of A, the row sums $d_i = \sum_j |a_{ij}|$, and, for a reachable nonzero edge $i \to j$, the edge ratio $\rho_{ij}(A, h) = d_i \operatorname{sgn}(a_{ij})\, h_j / h_i$. The condition holds when all these ratios coincide. The nondegeneracy hypotheses ($d_i \neq 0$ and $h_i \neq 0$ on $R(A, v)$) are exactly what makes every ρ_{ij} well defined, so the algorithm checks them first.

Algorithm A1 Edge-ratio check for (A, v, h)

Input: $A = (a_{ij}) \in \mathbb{R}^{n \times n}$; vectors $v, h \in \mathbb{R}^n$; relative tolerance $\tau \geq 0$ ($\tau = 0$ for an exact test).
Output: (TRUE, λ) if (A, v, h) satisfies the edge-ratio condition, with λ the common ratio; otherwise FALSE with a pair of disagreeing edges.

(i). *Support and reachability.* Set $S_v \leftarrow \{i : v_i \neq 0\}$. Compute $R(A, v)$ by a breadth-first search from S_v in the directed graph with an edge $i \to j$ whenever $a_{ij} \neq 0$.

(ii). *Nondegeneracy.* For each $i \in R(A, v)$ compute $d_i = \sum_j |a_{ij}|$. If $d_i = 0$ or $h_i = 0$ for some $i \in R(A, v)$, return FALSE.

(iii). *Reference ratio.* Set $\lambda \leftarrow$ UNDEFINED.

(iv). *Single pass over reachable nonzero edges.* For each $i \in R(A, v)$ and each j with $a_{ij} \neq 0$:

- (a) compute $\rho \leftarrow d_i \operatorname{sgn}(a_{ij})\, h_j / h_i$;
- (b) if λ is UNDEFINED, set $\lambda \leftarrow \rho$ and record the reference edge $(i_0, j_0) \leftarrow (i, j)$;
- (c) otherwise, if $|\rho - \lambda| > \tau \max\{1, |\lambda|\}$, return $\big(\text{FALSE}, (i_0, j_0), (i, j)\big)$.

(v). *Return.* If λ is still UNDEFINED (no reachable nonzero edge), the condition holds vacuously; return $(\text{TRUE}, \text{UNDEFINED})$. Otherwise return (TRUE, λ).

Remark A1 (Correctness)**.** *Steps 1–2 guarantee that every quantity in step 4(a) is defined, since j reached from a nonzero edge is itself in $R(A, v)$, where $h_j \neq 0$. Step 4 compares every reachable edge ratio against a single reference value, so it returns* TRUE *exactly when $\rho_{ij}(A, h)$ is constant on the reachable nonzero edge set, which is the edge-ratio condition. The returned λ is the common ratio, which by Theorem 6 is the eigenvalue of the corresponding eigen-triple.*

Remark A2 (Complexity)**.** *The breadth-first search visits each vertex and edge once, and the loop of step 4 touches each reachable nonzero entry once, so the running time is $O\big(|S_v| + \operatorname{nnz}(A)\big)$ and the working memory is $O(n)$, where $\operatorname{nnz}(A)$ is the number of nonzero entries. Row sums d_i may be accumulated during the same pass.*

Remark A3 (Floating-point and exact tests)**.** *With $\tau = 0$ the comparison in step 4(c) is exact, appropriate for integer or exact rational data. For floating-point data a small relative tolerance τ (a modest multiple of the unit roundoff) absorbs rounding in the products $d_i h_j / h_i$; the witness returned on failure, (i_0, j_0) versus (i, j), localizes the largest disagreement when the loop is ordered to keep the extremal ratio.*

Remark A4 (Certifying only a finite truncation)**.** *To certify exactness of the RMC quotient at a fixed truncation order K rather than for all K (Corollary 4), restrict the breadth-first search of step 1 to depth $K{+}1$, i.e., collect exactly the edges lying on admissible paths of length $\leq K{+}1$, and run steps 2–5 on that edge set only. Adjoining the no-step compatibility check ($\operatorname{sgn}(v_i) h_i$ constant on S_v) then certifies $(K{+}1)$-step zero variance (Proposition 6).*

References

1. Parlett, B.N. *The Symmetric Eigenvalue Problem*; SIAM: Philadelphia, PA, USA, 1998.
2. Sorensen, D.C. Numerical methods for large eigenvalue problems. *Acta Numer.* **2002**, *11*, 519–584. [CrossRef]
3. Trefethen, L.N.; Embree, M. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*; Princeton University Press: Princeton, NJ, USA, 2020.

4. Halko, N.; Martinsson, P.G.; Tropp, J.A. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Rev.* **2011**, *53*, 217–288. [CrossRef]
5. Martinsson, P.G.; Tropp, J.A. Randomized numerical linear algebra: Foundations and algorithms. *Acta Numer.* **2020**, *29*, 403–572. [CrossRef]
6. Murray, R.; Demmel, J.; Mahoney, M.W.; Erichson, N.B.; Melnichenko, M.; Malik, O.A.; Grigori, L.; Luszczek, P.; Dereziński, M.; Lopes, M.E.; et al. Randomized numerical linear algebra: A perspective on the field with an eye to software. *arXiv* **2023**, arXiv:2302.11474.
7. Nakatsukasa, Y.; Tropp, J.A. Fast and accurate randomized algorithms for linear systems and eigenvalue problems. *SIAM J. Matrix Anal. Appl.* **2024**, *45*, 1183–1214. [CrossRef]
8. Hutchinson, M.F. A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. *Commun. Stat.-Simul. Comput.* **1989**, *18*, 1059–1076. [CrossRef]
9. Avron, H.; Toledo, S. Randomized algorithms for estimating the trace of an implicit symmetric positive semi-definite matrix. *J. ACM JACM* **2011**, *58*, 1–34. [CrossRef]
10. Meyer, R.A.; Musco, C.; Musco, C.; Woodruff, D.P. Hutch++: Optimal stochastic trace estimation. In *Proceedings of the Symposium on Simplicity in Algorithms (SOSA)*; SIAM: Philadelphia, PA, USA, 2021; pp. 142–155.
11. Ubaru, S.; Chen, J.; Saad, Y. Fast estimation of tr(f(A)) via stochastic Lanczos quadrature. *SIAM J. Matrix Anal. Appl.* **2017**, *38*, 1075–1099. [CrossRef]
12. Lin, L.; Saad, Y.; Yang, C. Approximating spectral densities of large matrices. *SIAM Rev.* **2016**, *58*, 34–65. [CrossRef]
13. Dimov, I.; Karaivanova, A. Parallel computations of eigenvalues based on a Monte Carlo approach. *Monte Carlo Methods Appl.* **1998**, *4*, 33–52. [CrossRef]
14. Dimov, I.; Alexandrov, V.; Karaivanova, A. Parallel resolvent Monte Carlo algorithms for linear algebra problems. *Math. Comput. Simul.* **2001**, *55*, 25–35. [CrossRef]
15. Dimov, I.; Philippe, B.; Karaivanova, A.; Weihrauch, C. Robustness and applicability of Markov chain Monte Carlo algorithms for eigenvalue problems. *Appl. Math. Model.* **2008**, *32*, 1511–1529. [CrossRef]
16. Dimov, I.T. *Monte Carlo Methods for Applied Scientists*; World Scientific: Singapore, 2008.
17. Mascagni, M.; Karaivanova, A. A parallel quasi-monte carlo method for computing extremal eigenvalues. In *Proceedings of the Monte Carlo and Quasi-Monte Carlo Methods 2000: Proceedings of a Conference held at Hong Kong Baptist University, Hong Kong SAR, China, 27 November–1 December 2000*; Springer: Berlin/Heidelberg, Germany, 2002; pp. 369–380.
18. Mascagni, M.; Karaivanova, A. A Monte Carlo approach for finding more than one eigenpair. In *Proceedings of the International Conference on Numerical Methods and Applications*; Springer: Berlin/Heidelberg, Germany, 2002; pp. 123–131.
19. Alexandrov, V.; Karaivanova, A. Finding the smallest eigenvalue by the Inverse Monte Carlo Method with Refinement. In *Proceedings of the International Conference on Computational Science*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 766–774.
20. Karaivanova, A. Quasi-Monte Carlo methods for some linear algebra problems. Convergence and complexity. *Serdica J. Comput.* **2010**, *4*, 57–72. [CrossRef]
21. Gurova, S.M.; Karaivanova, A. Monte Carlo method for estimating eigenvalues using error balancing. In *Proceedings of the International Conference on Large-Scale Scientific Computing*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 447–455.
22. Gurova, S.M.; Atanassov, E.; Karaivanova, A. A Resolvent Quasi-Monte Carlo Method for Estimating the Minimum Eigenvalues Using the Error Balancing. In *Proceedings of the International Conference on Large-Scale Scientific Computing*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 394–403.
23. Assaraf, R.; Caffarel, M. Zero-variance principle for Monte Carlo algorithms. *Phys. Rev. Lett.* **1999**, *83*, 4682–4685. [CrossRef]
24. Broniatowski, M.; Caron, V. Towards zero variance estimators for rare event probabilities. *arXiv* **2011**, arXiv:1104.1464.
25. Awad, H.P.; Glynn, P.W.; Rubinstein, R.Y. Zero-variance importance sampling estimators for Markov process expectations. *Math. Oper. Res.* **2013**, *38*, 358–388. [CrossRef]
26. Dimov, I.; Alexandrov, V.; Branford, S.; Weihrauch, C. Error analysis of a monte carlo algorithm for computing bilinear forms of matrix powers. In *Proceedings of the International Conference on Computational Science*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 632–639.