

MONTE CARLO METHODS FOR ESTIMATING THE DIAGONAL OF A REAL SYMMETRIC MATRIX

ERIC HALLMAN*, ILSE C.F. IPSEN*, AND ARVIND K. SAIBABA*

Abstract. For real symmetric matrices that are accessible only through matrix vector products, we present Monte Carlo estimators for computing the diagonal elements. Our probabilistic bounds for normwise absolute and relative errors apply to Monte Carlo estimators based on random Rademacher, sparse Rademacher, normalized and unnormalized Gaussian vectors, and to vectors with bounded fourth moments. The novel use of matrix concentration inequalities in our proofs represents a systematic model for future analyses. Our bounds mostly do not depend on the matrix dimension, target different error measures than existing work, and imply that the accuracy of the estimators increases with the diagonal dominance of the matrix. An application to derivative-based global sensitivity metrics corroborates this, as do numerical experiments on synthetic test matrices. We recommend against the use in practice of sparse Rademacher vectors, which are the basis for many randomized sketching and sampling algorithms, because they tend to deliver barely a digit of accuracy even under large sampling amounts.

Key words. Concentration inequalities, Monte Carlo Methods, Relative error, Rademacher random vectors, Gaussian random vectors

AMS subject classification. 15A15, 65C05, 65F50, 60G50, 68W20

1. Introduction. We compute the diagonal elements of symmetric matrices $\mathbf{A} \in \mathbb{R}^{n \times n}$ with Monte Carlo estimators of the form

$$\hat{\mathbf{A}} = \frac{1}{N} \sum_{k=1}^N \mathbf{A} \mathbf{z}_k \mathbf{z}_k^\top$$

where \mathbf{z}_k are independent random vectors. This approach is crucial when the elements of \mathbf{A} are available only implicitly, via matrix vector products.

Estimating the diagonal elements of a matrix is important in many areas of science and engineering: In electronic structure calculations, one computes the diagonal elements of a projector onto the smallest eigenvectors of a Hamiltonian matrix [4]. In statistics, leverage scores for column subset selection can be computed from the diagonals of the projector onto the column space. In Bayesian inverse problems, the diagonal elements of the posterior covariance are computed with matrix-free estimators. Diagonal, or Jacobi preconditioners can accelerate the convergence of iterative linear solvers [21]. More recently, diagonal estimators have been used to accelerate second order optimization techniques for machine learning [22]. In network science, subgraph centrality measures and ranks the importance of the network nodes based on the diagonal of a scaled exponential of the adjacency matrix. In sensitivity analysis, Monte Carlo diagonal estimators efficiently compute the derivative-based global sensitivity metrics [6, 10].

Diagonal estimation is related to trace estimation. Once the diagonal elements are known, the trace can be computed from their sum. Therefore, estimators for the diagonal of a matrix can be easily adapted to trace estimators. Monte Carlo methods were first proposed by Hutchinson [8], and subsequently improved and expanded to different distributions [2, 7, 16]. Applications of trace estimators, reviewed in [19], include estimating density of states, log determinants, and Schatten p -norms.

*Department of Mathematics, North Carolina State University, Raleigh, NC, 27695 (er-hallma@ncsu.edu, ipsen@ncsu.edu, asaibab@ncsu.edu)

Literature review. To our knowledge, Monte Carlo diagonal estimators were first proposed by Bekas, Kokiopoulou, and Saad [4], and a sufficient condition was given for a Monte Carlo estimator to be unbiased. However, this paper identified that large offdiagonal entries can result in large relative errors and developed probing methods to mitigate the effects of the offdiagonal entries. This idea is further explored in the following works [11, 9].

We are aware of a recent paper [3] as the only other work to analyze the number of samples required for a relative (ϵ, δ) estimator. In contrast to [3], our proofs are the first to exploit matrix concentration inequalities to impose a systematic structure that can serve as a model for future analyses and allow us to analyze the normwise errors in a different norm. We analyze more general distributions such as random vectors with bounded fourth moments and sparse Rademacher vectors with a user-specified sparsity parameter, and—in contrast to [3]—focus on un-normalized estimators. Most of our bounds do not show an explicit dependence on the matrix dimension which is desirable for large-scale problems.

1.1. Contributions and overview. After introducing notation, relevant concentration inequalities, and the setup for our analysis (section 2.1), we derive normwise error bounds for Monte Carlo estimators based on independent Rademacher vectors (section 3), random vectors with bounded fourth moments and Gaussian vectors (section 4); componentwise bounds for Rademacher and Gaussian vectors (section 5); and apply Monte Carlo estimators to derivative-based global sensitivity metrics (section 6). Numerical experiments (section 7) illustrate the accuracy of the Monte Carlo estimators and the bounds. The novel and noteworthy features of contributions are:

1. Most of our bounds do not depend on the matrix dimension n , and hold for all symmetric matrices, whether positive definite or not.
2. We extend the concept of relative (ϵ, δ) estimators to diagonal estimation to determine the minimal number of samples N for a user-specified choice of relative error ϵ and failure probability δ (Definitions 2.5, 2.6).
3. Our normwise bounds suggest that for Rademacher vectors, the Monte Carlo estimators are more accurate for matrices that are more strongly diagonally dominant (Theorem 3.2). In particular, the least number of samples required for the Monte Carlo estimators to achieve a user-specified relative error decreases with increasing diagonal dominance of \mathbf{A} in the relative sense (Corollaries 3.3, 4.2).
4. For Rademacher vectors parameterized in terms of sparsity levels (Definition 3.4), we show that the Monte Carlo estimators lose accuracy with increasing sparsity (Theorem 3.5, Corollary 3.6). Numerical experiments (section 7) confirm that, even for large sampling amounts, the estimators barely achieve a single digit of accuracy. Therefore we recommend against their use in practice.
5. Our componentwise bounds suggest that the accuracy for computing a diagonal element a_{ii} depends only on the diagonal dominance of column/row i of \mathbf{A} (Corollaries 5.4, 5.6).
6. In the context of derivative-based global sensitivity metrics, we design and analyze Monte Carlo estimators based on random vectors from a problem-specific probability distribution (Theorem 6.1, Corollary 6.2).

2. Background. After reviewing notation (section 2.1) and relevant concentration inequalities (section 2.2), we present the setup for our analysis (section 2.3).

2.1. Notation. The *Schur product* (or *Hadamard*, or *elementwise product*) of $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ is denoted by $\mathbf{C} = \mathbf{A} \circ \mathbf{B} \in \mathbb{R}^{m \times n}$ and has elements

$$c_{ij} = a_{ij}b_{ij} \quad 1 \leq i \leq m, \quad 1 \leq j \leq n.$$

For $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{m \times n}$, the Schur product is commutative and distributive,

$$\mathbf{A} \circ \mathbf{B} = \mathbf{B} \circ \mathbf{A}, \quad \mathbf{A} \circ (\mathbf{B} + \mathbf{C}) = \mathbf{A} \circ \mathbf{B} + \mathbf{A} \circ \mathbf{C}.$$

Following MATLAB convention, we define $\text{diag}(\mathbf{A}) = (a_{11} \ \cdots \ a_{nn})^\top \in \mathbb{R}^n$ as the column vector of diagonal elements of $\mathbf{A} \in \mathbb{R}^{n \times n}$. The operator diag is overloaded, and $\text{diag}(\mathbf{x}) \in \mathbb{R}^{n \times n}$ represents a diagonal matrix whose diagonal elements are the elements of the vector $\mathbf{x} \in \mathbb{R}^n$. In particular,

$$(2.1) \quad \mathcal{D}(\mathbf{A}) \equiv \text{diag}(\text{diag}(\mathbf{A})) = \mathbf{I} \circ \mathbf{A} \in \mathbb{R}^{n \times n}$$

represents the diagonal matrix whose diagonal elements are the diagonal elements of \mathbf{A} . In other words, $\mathbf{I} \circ \mathbf{A}$ zeros out the offdiagonal elements of \mathbf{A} .

If the first factor in a Schur product is a square matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$, and the second factor an outer product involving $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then

$$(2.2) \quad \mathbf{M} \circ (\mathbf{x}\mathbf{y}^\top) = \text{diag}(\mathbf{x}) \mathbf{M} \text{diag}(\mathbf{y}).$$

For symmetric matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$, the partial order $\mathbf{A} \preceq \mathbf{B}$, or equivalently $\mathbf{B} \succeq \mathbf{A}$, says that $\mathbf{B} - \mathbf{A}$ is positive semidefinite. If \mathbf{A} and \mathbf{B} are positive semidefinite, then $\mathbf{A} \preceq \mathbf{B}$ implies $\mathbf{A}^{1/2} \preceq \mathbf{B}^{1/2}$.

The *intrinsic dimension* of a nonzero symmetric positive semidefinite matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is

$$\text{intdim}(\mathbf{A}) \equiv \frac{\text{trace}(\mathbf{A})}{\|\mathbf{A}\|_2}, \quad \text{with} \quad 1 \leq \text{intdim}(\mathbf{A}) \leq \text{rank}(\mathbf{A}) \leq n.$$

If, additionally, \mathbf{A} is a diagonal matrix, then

$$\text{intdim}(\mathbf{A}) = \frac{\sum_{i=1}^n a_{ii}}{\max_{1 \leq i \leq n} a_{ii}}.$$

The columns of $\mathbf{A} = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n] \in \mathbb{R}^{m \times n}$ are $\mathbf{a}_j \in \mathbb{R}^m$, $1 \leq j \leq n$, and the columns of the identity $\mathbf{I} = [\mathbf{e}_1 \ \cdots \ \mathbf{e}_n] \in \mathbb{R}^{n \times n}$ are $\mathbf{e}_j \in \mathbb{R}^n$. The transpose of \mathbf{A} is \mathbf{A}^\top .

2.2. Concentration inequalities. We rely on two scalar and two matrix concentration inequalities.

Markov's inequality [15, Section 3.1] bounds the the probability that a random variable exceeds a constant.

THEOREM 2.1 (Markov's inequality). *If Z is a non-negative random variable, then for $t > 0$*

$$\mathbb{P}[Z \geq t] \leq \frac{\mathbb{E}[Z^2]}{t^2}.$$

Hoeffding's inequality for general bounded random variables [20, Theorem 2.2.6] bounds the probability that a sum of scalar random variables exceeds its mean.

THEOREM 2.2 (Scalar Hoeffding inequality). *Let Z_1, \dots, Z_N be independent random variables, bounded by $m_k \leq Z_k \leq M_k$, $1 \leq k \leq N$, with sum $Z \equiv \sum_{k=1}^N Z_k$. Then for $t > 0$*

$$\mathbb{P}[|Z - \mathbb{E}[Z]| \geq t] \leq 2 \exp\left(\frac{-2t^2}{\sum_{k=1}^N (M_k - m_k)^2}\right).$$

Next are two bounds for sums of independent symmetric matrix-valued random variables. The first is a matrix Bernstein concentration inequality [18, Theorems 7.3.1 and 7.7.1] for sums of independent, symmetric, bounded, zero-mean random matrices.

THEOREM 2.3 (Matrix Bernstein inequality). *Let $\mathbf{S}_1, \dots, \mathbf{S}_N \in \mathbb{R}^{n \times n}$ be independent symmetric random matrices with*

$$\mathbb{E}[\mathbf{S}_k] = \mathbf{0}, \quad \|\mathbf{S}_k\|_2 \leq L \quad 1 \leq k \leq N.$$

Let the sum $\mathbf{S} \equiv \sum_{k=1}^N \mathbf{S}_k$ have a matrix-valued variance that is majorized by $\mathbf{V} \in \mathbb{R}^{n \times n}$,

$$\mathbf{V} \succeq \text{Var}(\mathbf{S}) = \mathbb{E}[\mathbf{S}^2] = \sum_{k=1}^N \mathbb{E}[\mathbf{S}_k^2].$$

Abbreviate $\nu \equiv \|\mathbf{V}\|_2$ and $d \equiv \text{intdim}(\mathbf{V})$. Then for $t > 0$

$$(2.3) \quad \mathbb{P}[\|\mathbf{S}\|_2 \geq t] \leq 8d \exp\left(\frac{-t^2}{2(\nu + Lt/3)}\right).$$

Proof. In [18, Theorems 7.3.1 and 7.7.1] it is shown that (2.3) holds, provided $t \geq \sqrt{\nu} + \frac{L}{3}$. We show that (2.3) always holds and the lower bound on t is not necessary. To see this, note that

$$\frac{-t^2}{2(\nu + Lt/3)}$$

decreases monotonically as t increases. Therefore we can bound it from below as long as $t < \sqrt{\nu} + L/3$, by

$$\frac{-t^2}{2(\nu + Lt/3)} > \frac{-(\sqrt{\nu} + L/3)^2}{2(\nu + (L/3)(\sqrt{\nu} + L/3))} \geq -\frac{2}{3}.$$

The second inequality comes from setting $x = \frac{L}{3\sqrt{\nu}}$ and noting that $f(x) = -\frac{(1+x)^2}{2(1+x+x^2)}$ has a minimum at $\hat{x} = 1$ where $f(\hat{x}) = -2/3$, and $\sqrt{\nu} = L/3$. Substituting the lower bound into Theorem 2.3 gives

$$8d \exp\left(\frac{-t^2}{2(\nu + Lt/3)}\right) \geq 8 \exp(-2/3) > 4.$$

But now (2.3) holds trivially since

$$\mathbb{P}[\|\mathbf{S}\|_2 \geq t] \leq 1 < 4 < 8d \exp\left(\frac{-t^2}{2(\nu + Lt/3)}\right).$$

□

The second matrix concentration inequality [5, Theorem 3.2] bounds the mean of the squared norm of the sum of symmetric random matrices.

THEOREM 2.4. Let $\mathbf{S}_1, \dots, \mathbf{S}_N \in \mathbb{R}^{n \times n}$ with $n \geq 3$ be independent symmetric random matrices with zero mean. Then

$$\mathbb{E} \left[\left\| \sum_{k=1}^N \mathbf{S}_k \right\|_2^2 \right]^{1/2} \leq \sqrt{2e \ln n} \left\| \left(\sum_{k=1}^N \mathbb{E}[\mathbf{S}_k^2] \right)^{1/2} \right\|_2 + 4e \ln n \left(\mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{S}_k\|_2^2 \right] \right)^{1/2}.$$

2.3. Setup for the analysis. Our Monte Carlo estimators compute the diagonal elements of a symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ by means of matrix vector products with \mathbf{A} . It samples N independent random vectors $\mathbf{w}_k \in \mathbb{R}^n$, and approximates the vector of diagonal elements $\text{diag}(\mathbf{A}) \in \mathbb{R}^n$ by the mean

$$\text{diag}(\widehat{\mathbf{A}}) = \frac{1}{N} \sum_{k=1}^N ((\mathbf{A}\mathbf{w}_k) \circ \mathbf{w}_k) \in \mathbb{R}^n \quad \text{where} \quad \widehat{\mathbf{A}} \equiv \frac{1}{N} \sum_{k=1}^N \mathbf{A}\mathbf{w}_k\mathbf{w}_k^\top \in \mathbb{R}^{n \times n}.$$

To see this, apply (2.1) and (2.2) to

$$\begin{aligned} \mathcal{D}(\widehat{\mathbf{A}}) &= \mathbf{I} \circ \widehat{\mathbf{A}} = \frac{1}{N} \sum_{k=1}^N \text{diag}(\mathbf{A}\mathbf{w}_k) \mathbf{I} \text{diag}(\mathbf{w}_k) = \frac{1}{N} \sum_{k=1}^N \text{diag}((\mathbf{A}\mathbf{w}_k) \circ \mathbf{w}_k) \\ &= \text{diag} \left(\frac{1}{N} \sum_{k=1}^N ((\mathbf{A}\mathbf{w}_k) \circ \mathbf{w}_k) \right) = \text{diag}(\text{diag}(\widehat{\mathbf{A}})) \in \mathbb{R}^{n \times n}. \end{aligned}$$

Alternately, the diagonal elements of the estimators can be expressed as

$$\begin{aligned} \widehat{\mathbf{A}}_{ii} &= \frac{1}{N} \sum_{k=1}^N (\mathbf{A}\mathbf{w}_k\mathbf{w}_k^\top)_{ii} = \frac{1}{N} \sum_{k=1}^N (\mathbf{A}\mathbf{w}_k)_i (\mathbf{w}_k^\top)_i \\ &= \frac{1}{N} \sum_{k=1}^N ((\mathbf{A}\mathbf{w}_k) \circ \mathbf{w}_k)_i, \quad 1 \leq i \leq n. \end{aligned}$$

We measure the cost of a diagonal estimator by the number N of samples. To assess the accuracy, we introduce a relative error in the form of normwise and componentwise (ϵ, δ) estimators, which extend the notion of (ϵ, δ) trace estimator from [2, 16].

DEFINITION 2.5 (Normwise (ϵ, δ) diagonal estimator). Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be symmetric. Given user-specified parameters $0 < \epsilon, \delta < 1$, we say that $\mathcal{D}(\widehat{\mathbf{A}})$ is a normwise (ϵ, δ) estimator for the diagonal elements of \mathbf{A} , if

$$\|\mathcal{D}(\mathbf{A}) - \mathcal{D}(\widehat{\mathbf{A}})\|_2 \leq \epsilon \|\mathcal{D}(\mathbf{A})\|_2$$

holds with probability at least $1 - \delta$.

In other words, for a user-specified failure probability $\delta > 0$ and tolerance $\epsilon > 0$, the normwise relative error of the diagonal estimator is, with probability at most $1 - \delta$, at most ϵ . The two-norm in Definition 2.5 can be replaced by any matrix- p norm, because the p -norm of a diagonal matrix $\mathbf{D} \in \mathbb{R}^{n \times n}$ is $\|\mathbf{D}\|_p = \max_{1 \leq i \leq n} |d_{ii}|$ for $p \geq 1$. Next we define a componentwise (ϵ, δ) estimator.

DEFINITION 2.6 (Componentwise (ϵ, δ) diagonal estimator). Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be symmetric. Given user-specified parameters $0 < \epsilon, \delta < 1$ and diagonal element $a_{ii} \neq 0$ of \mathbf{A} , we say that $\widehat{a}_{ii} = (\widehat{\mathbf{A}})_{ii}$ is a componentwise (ϵ, δ) estimator for a_{ii} , if

$$|\widehat{a}_{ii} - a_{ii}| \leq \epsilon |a_{ii}|$$

holds with probability at least $1 - \delta$.

3. Normwise bounds for Rademacher random vectors. We present normwise bounds for Monte Carlo estimators based on standard (section 3.1) and on sparse Rademacher vectors (section 3.2).

3.1. Standard Rademacher vectors. After defining Rademacher vectors (Definition 3.1) and discussing their properties (Remarks 3.1 and 3.2), we present a normwise absolute error bound (Theorem 3.2), and a bound on the minimal sampling amount that makes the Rademacher Monte Carlo estimator a normwise (ϵ, δ) diagonal estimator (Corollary 3.3).

DEFINITION 3.1. *A Rademacher random variable takes on the values ± 1 with equal probability $1/2$. A Rademacher vector is a random vector whose elements are independent Rademacher random variables.*

Standard Rademacher vectors have the advantage of cheap matrix vector products and immediately recovering diagonal matrices.

REMARK 3.1. *The elements w_j of a Rademacher vector \mathbf{w} have the following properties:*

1. *Zero mean:* $\mathbb{E}[w_j] = 0$
2. *Constant square:* $w_j^2 = 1$
3. *Independence:* $\mathbb{E}[w_j w_i] = 0$ for $i \neq j$.

REMARK 3.2. *Standard Rademacher vectors recover a diagonal matrix with a single sample, $N = 1$.*

To see this, let $\mathbf{A} = \mathcal{D}(\mathbf{A}) \in \mathbb{R}^{n \times n}$ be diagonal, and $\mathbf{w} \in \mathbb{R}^n$ a Rademacher vector. Remark 3.1 implies that $\mathbf{A} \mathbf{w} \mathbf{w}^\top \in \mathbb{R}^{n \times n}$ has diagonal elements $a_{ii} w_i^2 = a_{ii}$, $1 \leq i \leq n$.

As a consequence, we can focus the analysis of standard Rademacher-based estimators on non-diagonal matrices. The results below are special cases of those for sparse Rademacher vectors in section 3.2.

THEOREM 3.2. *Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be non-diagonal symmetric, and*

$$K_1 \equiv \|\mathcal{D}(\mathbf{A}^2) - \mathcal{D}(\mathbf{A})^2\|_2, \quad K_2 \equiv \|\mathbf{A} - \mathcal{D}(\mathbf{A})\|_\infty, \quad d \equiv (\|\mathbf{A}\|_F^2 - \|\mathcal{D}(\mathbf{A})\|_F^2)/K_1.$$

If $\hat{\mathbf{A}} \equiv \frac{1}{N} \sum_{k=1}^N \mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top$ is a Monte Carlo estimator with independent Rademacher vectors $\mathbf{w}_k \in \mathbb{R}^n$, $1 \leq k \leq N$, then the probability that the absolute error exceeds $t > 0$ is at most

$$\mathbb{P} \left[\|\mathcal{D}(\mathbf{A}) - \mathcal{D}(\hat{\mathbf{A}})\|_2 \geq t \right] \leq 8d \exp \left(\frac{-Nt^2}{2(K_1 + tK_2/3)} \right).$$

Proof. This is the special case $s = 1$ of Theorem 3.5. □

The constants K_1 and K_2 represent the absolute deviation of \mathbf{A} from diagonality, and more specifically the degree of diagonal dominance of \mathbf{A} in the absolute sense. Theorem 3.2 implies that the Rademacher estimator has a small absolute error when applied to strongly diagonally dominant matrices. In other words, the normwise absolute error in the Rademacher estimator decreases with increasing diagonal dominance of \mathbf{A} in the absolute sense.

We determine the least sampling amount required for the Monte Carlo estimator with Rademacher vectors to be a normwise (ϵ, δ) diagonal estimator.

COROLLARY 3.3. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be non-diagonal symmetric. Let

$$K_1 \equiv \|\mathcal{D}(\mathbf{A}^2) - \mathcal{D}(\mathbf{A})^2\|_2,$$

$$\Delta_1 \equiv \frac{K_1}{\|\mathcal{D}(\mathbf{A})\|_2^2}, \quad \Delta_2 \equiv \frac{\|\mathbf{A} - \mathcal{D}(\mathbf{A})\|_\infty}{\|\mathcal{D}(\mathbf{A})\|_\infty}, \quad d \equiv \frac{\|\mathbf{A}\|_F^2 - \|\mathcal{D}(\mathbf{A})\|_F^2}{K_1},$$

and let $\hat{\mathbf{A}} \equiv \frac{1}{N} \sum_{k=1}^N \mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top$ a Monte Carlo estimator with independent Rademacher vectors $\mathbf{w}_k \in \mathbb{R}^n$, $1 \leq k \leq N$. Pick $\epsilon > 0$. For any $0 < \delta < 1$, if the sampling amount is at least

$$(3.1) \quad N \geq \frac{\Delta_2}{3\epsilon^2} \left(2\epsilon + 6 \frac{\Delta_1}{\Delta_2} \right) \ln(8d/\delta),$$

then $\|\mathcal{D}(\mathbf{A}) - \mathcal{D}(\hat{\mathbf{A}})\|_2 \leq \epsilon \|\mathcal{D}(\mathbf{A})\|_2$ holds with probability at least $1 - \delta$.

Proof. This is the special case $s = 1$ of Corollary 3.6. \square

The constants Δ_1 and Δ_2 in Corollary 3.3 represent the respective relative counterparts of K_1 and K_2 in Theorem 3.2: they represent the relative deviation of \mathbf{A} from diagonality, and more specifically the degree of diagonal dominance of \mathbf{A} in the relative sense. Corollary 3.3 implies that if \mathbf{A} is strongly diagonally dominant in the relative sense, then a small sampling amount suffices to make the Rademacher estimator a normwise (ϵ, δ) diagonal estimator. As with many randomized sampling algorithms, the lower bound for N is proportional to $1/\epsilon^2$.

3.2. Sparse Rademacher vectors. For Rademacher vectors that are parameterized in terms of sparsity (Definition 3.4), we derive a normwise absolute error bound (Theorem 3.5), followed by the minimal sampling amount that makes the sparse Rademacher estimator a normwise (ϵ, δ) diagonal estimator (Corollary 3.6).

The random vectors in [1] have elements that assume values from the discrete distribution $\{-\sqrt{3}, 0, \sqrt{3}\}$ with respective probability $\{\frac{1}{6}, \frac{2}{3}, \frac{1}{6}\}$. This concept was extended in [14, (2)] to Rademacher vectors that are parameterized in terms of a sparsity parameter s .

DEFINITION 3.4. A sparse Rademacher random variable with parameter $s \geq 1$ takes the values $\{-\sqrt{s}, 0, \sqrt{s}\}$ with probability $\{\frac{1}{2s}, 1 - \frac{1}{s}, \frac{1}{2s}\}$ respectively.

A Sparse Rademacher vector is a random vector whose elements are independent sparse Rademacher random variables.

The properties of sparse Rademacher vectors are almost the same as those of the original Rademacher vectors in Remark 3.1.

REMARK 3.3. The elements of a sparse Rademacher vector \mathbf{w} with parameter $s \geq 1$ have the following properties

1. Zero mean: $\mathbb{E}[w_j] = 0$
2. Unit variance $\mathbb{E}[w_j^2] = 1$
3. Independence: For $i \neq j$, and integer $\ell \geq 1$

$$\mathbb{E}[w_i^2 w_j^2] = 1, \quad \mathbb{E}[w_i^\ell w_j] = \mathbb{E}[w_i w_j^\ell] = 0.$$

The case $s = 1$ corresponds to the original Rademacher vectors (Definition 3.1), while $s = 3$ corresponds to the choice in [1].

Below is the extension of Theorem 3.2 to sparse Rademacher vectors with integer parameters s .

THEOREM 3.5. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be non-diagonal symmetric, and let

$$K_1(s) \equiv \|\mathcal{D}(\mathbf{A}^2) + (s-2)\mathcal{D}(\mathbf{A})^2\|_2, \quad K_2(s) \equiv \|s\mathbf{A} - \mathcal{D}(\mathbf{A})\|_\infty, \\ d(s) \equiv \frac{1}{K_1(s)} \|\mathbf{A}\|_F^2 + (s-2)\|\mathcal{D}(\mathbf{A})\|_F^2.$$

Let $\hat{\mathbf{A}} \equiv \frac{1}{N} \sum_{j=1}^N \mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top$ be a Monte Carlo estimator with independent sparse Rademacher random vectors $\mathbf{w}_k \in \mathbb{R}^n$, with integer parameter $s \geq 1$. Then the probability that the absolute error exceeds $t > 0$ is at most

$$\mathbb{P} \left[\|\mathcal{D}(\mathbf{A}) - \mathcal{D}(\hat{\mathbf{A}})\|_2 \geq t \right] \leq 8d(s) \exp \left(\frac{-Nt^2}{2(K_1(s) + tK_2(s)/3)} \right).$$

Proof. Define the random diagonal matrices

$$\mathbf{S}_k \equiv \frac{1}{N} (\mathbf{I} \circ (\mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top) - \mathbf{I} \circ \mathbf{A}), \quad 1 \leq k \leq N,$$

and their sum

$$(3.2) \quad \mathbf{Z} \equiv \sum_{k=1}^N \mathbf{S}_k = \mathbf{I} \circ \hat{\mathbf{A}} - \mathbf{I} \circ \mathbf{A} = \mathcal{D}(\mathbf{A}) - \mathcal{D}(\hat{\mathbf{A}}).$$

Before applying Corollary 2.3, we need to verify the assumptions for the Bernstein inequality.

1. *Expectation.* With (2.2), the first summand in \mathbf{S}_k equals

$$\mathbf{I} \circ (\mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top) = \text{diag}(\mathbf{A} \mathbf{w}_k) \mathbf{I} \text{diag}(\mathbf{w}_k) = \text{diag}(\mathbf{A} \mathbf{w}_k) \text{diag}(\mathbf{w}_k),$$

and with (2.1) this gives

$$(3.3) \quad \mathbf{S}_k = \frac{1}{N} (\text{diag}(\mathbf{A} \mathbf{w}_k) \text{diag}(\mathbf{w}_k) - \mathcal{D}(\mathbf{A})), \quad 1 \leq k \leq N.$$

The linearity of expectation implies

$$\mathbb{E}[\mathbf{S}_k] = \frac{1}{N} \left(\underbrace{\mathbb{E}[\text{diag}(\mathbf{A} \mathbf{w}_k) \text{diag}(\mathbf{w}_k)]}_{\mathbf{T}} - \mathcal{D}(\mathbf{A}) \right), \quad 1 \leq k \leq N.$$

Remark 3.3 implies that element (i, j) of $\mathbf{T} \equiv \mathbb{E}[\text{diag}(\mathbf{A} \mathbf{w}_k) \text{diag}(\mathbf{w}_k)]$ equals

$$(3.4) \quad t_{ij} = \mathbb{E}[(\mathbf{A} \mathbf{w}_k)_i \mathbf{I}_{ij}(\mathbf{w}_k)_j] = \mathbf{I}_{ij} a_{ij} \quad 1 \leq i \leq n, \quad 1 \leq j \leq n.$$

Hence $\mathbf{T} = \mathbf{I} \circ \mathbf{A}$, and $\mathbb{E}[\mathbf{S}_k] = \frac{1}{N}(\mathbf{T} - \mathbf{I} \circ \mathbf{A}) = \mathbf{0}$, $1 \leq k \leq N$. Therefore, the random matrices have zero mean, $\mathbb{E}[\mathbf{S}_k] = \mathbf{0}$; and so has their sum, $\mathbb{E}[\mathbf{Z}] = \mathbf{0}$.

2. *Boundedness.* From (3.3) follows that the diagonal matrices \mathbf{S}_k have diagonal elements

$$(\mathbf{S}_k)_{ii} = \frac{1}{N} ((\mathbf{A} \mathbf{w}_k)_i (\mathbf{w}_k)_i - a_{ii}) = \frac{1}{N} \left((\mathbf{w}_k)_i \sum_{j=1}^n a_{ij} (\mathbf{w}_k)_j - a_{ii} \right) \\ = \frac{1}{N} \left(a_{ii} ((\mathbf{w}_k)_i^2 - 1) + (\mathbf{w}_k)_i \sum_{j \neq i} a_{ij} (\mathbf{w}_k)_j \right), \quad 1 \leq i \leq n, \quad 1 \leq k \leq N.$$

Since $s \geq 1$ is an integer, $|((\mathbf{w}_k)_{ii}^2 - 1)a_{ii}| \leq (s-1)|a_{ii}|$ holds for any value of $(\mathbf{w}_k)_i$, and

$$|(\mathbf{S}_k)_{ii}| \leq \frac{1}{N} \left((s-1)|a_{ii}| + s \sum_{j \neq i} |a_{ij}| \right), \quad 1 \leq i \leq n, \quad 1 \leq k \leq N.$$

Since the matrix infinity norm is absolute, with $\|\mathbf{S}_k\|_\infty = \|\mathbf{S}_k\|_\infty$ we can bound the two-norms of the diagonal matrices \mathbf{S}_k by

$$\begin{aligned} \|\mathbf{S}_k\|_2 &\leq \max_{1 \leq i \leq n} \frac{1}{N} \left((s-1)|a_{ii}| + s \sum_{j \neq i} |a_{ij}| \right) = \max_{1 \leq i \leq n} \frac{1}{N} \left(-|a_{ii}| + s \sum_{j=1}^n |a_{ij}| \right) \\ &= \frac{1}{N} \|s\mathbf{A} - \mathcal{D}(\mathbf{A})\|_\infty = \frac{K_2(s)}{N}, \quad 1 \leq k \leq N. \end{aligned}$$

Set $L(s) \equiv K_2(s)/N$ where $K_2(s) > 0$ since \mathbf{A} is not diagonal.

3. *Variance.* In (3.3) abbreviate $\mathbf{D}_k \equiv \text{diag}(\mathbf{A}\mathbf{w}_k)$, $\mathbf{W}_k \equiv \text{diag}(\mathbf{w}_k)$, and $\mathbf{D} \equiv \mathcal{D}(\mathbf{A})$, and write the summands in $\text{Var}[\mathbf{Z}] = \sum_{k=1}^N \mathbb{E}[\mathbf{S}_k^2]$ as

$$\mathbf{S}_k^2 = \frac{1}{N^2} (\mathbf{D}_k \mathbf{W}_k - \mathbf{D})^2 = \frac{1}{N^2} ((\mathbf{D}_k \mathbf{W}_k)^2 - 2\mathbf{D} \mathbf{D}_k \mathbf{W}_k + \mathbf{D}^2), \quad 1 \leq k \leq N,$$

taking advantage of the fact that diagonal matrices commute. The expectation is

$$(3.5) \quad \mathbb{E}[\mathbf{S}_k^2] = \frac{1}{N^2} (\mathbb{E}[(\mathbf{D}_k \mathbf{W}_k)^2] - 2\mathbf{D} \mathbb{E}[\mathbf{D}_k \mathbf{W}_k] + \mathbf{D}^2), \quad 1 \leq k \leq N.$$

To determine $\mathbb{E}[\mathbf{D}_k \mathbf{W}_k]$, look at the individual diagonal elements and abbreviate $\mathbf{u} = \mathbf{w}_k$,

$$(3.6) \quad (\mathbf{D}_k \mathbf{W}_k)_{ii} = u_i \mathbf{e}_i^\top \mathbf{A} \mathbf{u} = u_i \sum_{j=1}^n a_{ij} u_j = a_{ii} u_i^2 + \sum_{j \neq i} a_{ij} u_i u_j, \quad 1 \leq i \leq n.$$

Remark 3.3 implies for the expectation

$$\mathbb{E}[(\mathbf{D}_k \mathbf{W}_k)_{ii}] = a_{ii} \mathbb{E}[u_i^2] + \sum_{j \neq i} a_{ij} \mathbb{E}[u_i u_j] = a_{ii}, \quad 1 \leq i \leq n.$$

Thus $\mathbb{E}[\mathbf{D}_k \mathbf{W}_k] = \mathbf{D}$. Insert this into (3.5)

$$(3.7) \quad \mathbb{E}[\mathbf{S}_k^2] = \frac{1}{N^2} (\mathbb{E}[(\mathbf{D}_k \mathbf{W}_k)^2] - \mathbf{D}^2), \quad 1 \leq k \leq N.$$

To determine $\mathbb{E}[(\mathbf{D}_k \mathbf{W}_k)^2]$, start from the individual diagonal elements in (3.6)

$$\begin{aligned} (\mathbf{D}_k \mathbf{W}_k)_{ii}^2 &= \left(a_{ii} u_i^2 + u_i \sum_{j \neq i} a_{ij} u_j \right)^2 \\ &= \underbrace{a_{ii}^2 u_i^4}_{\alpha} + \underbrace{2a_{ii} u_i^3 \sum_{j \neq i} a_{ij} u_j}_{\beta} + \underbrace{u_i^2 \left(\sum_{j \neq i} a_{ij} u_j \right)^2}_{\gamma} \quad 1 \leq i \leq n. \end{aligned}$$

In the expectation

$$(3.8) \quad \mathbb{E}[(\mathbf{D}_k \mathbf{W}_k)_{ii}^2] = \mathbb{E}[\alpha] + \mathbb{E}[\beta] + \mathbb{E}[\gamma]$$

we inspect each summand in turn.

Definition 3.4 implies for the fourth moment

$$(3.9) \quad \mathbb{E}[\alpha] = a_{ii}^2 \mathbb{E}[u_i^4] = a_{ii}^2 \left(\frac{1}{2s} (-\sqrt{s})^4 + \left(1 - \frac{1}{s}\right) 0 + \frac{1}{2s} (\sqrt{s})^4 \right) = a_{ii}^2 s.$$

Since u_i is independent from u_j for $j \neq i$ Remark 3.3 implies

$$(3.10) \quad \mathbb{E}[\beta] = 2a_{ii} \sum_{j \neq i} a_{ij} \underbrace{\mathbb{E}[u_i^3 u_j]}_{=0} = 0.$$

Collect the independent elements of \mathbf{u} ,

$$\begin{aligned} \gamma &= u_i^2 \left(\sum_{j \neq i} a_{ij} u_j \right) \left(\sum_{\ell \neq i} a_{i\ell} u_\ell \right) = u_i^2 \left(\sum_{j \neq i} a_{ij} u_j \left(a_{ij} u_j + \sum_{\ell \neq i, j} a_{i\ell} u_\ell \right) \right) \\ &= \sum_{j \neq i} a_{ij}^2 u_j^2 u_i^2 + \sum_{j \neq i} \sum_{\ell \neq i, j} a_{ij} a_{i\ell} u_j u_\ell u_i^2 \end{aligned}$$

and apply Remark 3.3,

$$(3.11) \quad \mathbb{E}[\gamma] = \sum_{j \neq i} a_{ij}^2 \underbrace{\mathbb{E}[u_j^2 u_i^2]}_{=1} + \sum_{j \neq i} \sum_{\ell \neq i, j} a_{ij} a_{i\ell} \underbrace{\mathbb{E}[u_j u_\ell u_i^2]}_{=0} = \sum_{j \neq i} a_{ij}^2$$

Substitute (3.9), (3.10) and (3.11) into (3.8). Then Remark 3.3 and the symmetry of \mathbf{A} imply

$$\mathbb{E}[(\mathbf{D}_k \mathbf{W}_k)_{ii}^2] = s a_{ii}^2 + \sum_{j \neq i} a_{ij}^2 = \sum_{j=1}^n a_{ij}^2 + (s-1) a_{ii}^2 = \|\mathbf{a}_i\|_2^2 + (s-1) |a_{ii}|^2.$$

With $\mathcal{D}(\mathbf{A}^2) = \text{diag}((\|\mathbf{a}_1\|_2^2 \quad \dots \quad \|\mathbf{a}_n\|_2^2))$ this gives for the whole diagonal matrix

$$\mathbb{E}[(\mathbf{D}_k \mathbf{W}_k)^2] = \mathcal{D}(\mathbf{A}^2) + (s-1) \mathbf{D}^2, \quad 1 \leq k \leq N.$$

At last, substitute this into the variance (3.7)

$$\begin{aligned} \mathbb{E}[\mathbf{S}_k^2] &= \frac{1}{N^2} (\mathcal{D}(\mathbf{A}^2) + (s-1) \mathbf{D}^2 - \mathbf{D}^2) \\ &= \frac{1}{N^2} (\mathcal{D}(\mathbf{A}^2) + (s-2) \mathbf{D}^2), \quad 1 \leq k \leq N. \end{aligned}$$

Sum up the individual variances,

$$\begin{aligned} \mathbf{V}(s) \equiv \text{Var}[\mathbf{Z}] &= \sum_{k=1}^N \mathbb{E}[\mathbf{S}_k^2] = \frac{1}{N^2} \sum_{k=1}^N (\mathcal{D}(\mathbf{A}^2) + (s-2) \mathbf{D}^2) \\ &= \frac{1}{N} (\mathcal{D}(\mathbf{A}^2) + (s-2) \mathbf{D}^2). \end{aligned}$$

Since $\mathbf{V}(s)$ is diagonal, its norm is

$$\nu(s) \equiv \|\mathbf{V}(s)\|_2 = \frac{1}{N} \|\mathcal{D}(\mathbf{A}^2) + (s-2)\mathbf{D}^2\|_2 = \frac{K_1(s)}{N},$$

where $K_1(s) > 0$ since \mathbf{A} is not diagonal. The intrinsic dimension of $\mathbf{V}(s)$ is

$$\begin{aligned} d(s) \equiv \text{intdim}(\mathbf{V}) &= \frac{\text{trace}(\mathcal{D}(\mathbf{A}^2) + (s-2)\mathbf{D}^2)}{N\nu(s)} = \frac{1}{K_1(s)} \sum_{j=1}^n (\|\mathbf{a}_j\|_2^2 + (s-2)|a_{jj}|^2) \\ &= \frac{1}{K_1(s)} \|\mathbf{A}\|_F^2 + (s-2)\|\mathcal{D}(\mathbf{A})\|_F^2. \end{aligned}$$

4. *Apply Theorem 2.3.* Substituting $L(s) = K_2(s)/N$ and $\nu(s) = K_1(s)/N$ into Theorem 2.3 and remembering that the sum \mathbf{Z} in (3.2) has zero mean gives

$$\mathbb{P}[\|\mathbf{Z}\|_2 \geq t] = \mathbb{P}[\|\mathcal{D}(\mathbf{A}) - \mathcal{D}(\hat{\mathbf{A}})\|_2 \geq t] \leq 8d(s) \exp\left(\frac{-Nt^2}{2(K_1(s) + tK_2(s)/3)}\right). \quad \square$$

In the special case $s = 1$ of standard Rademacher vectors, Theorem 3.5 reduces to Theorem 3.2. However, sparse Rademacher Monte Carlo estimators with $s > 1$ do, in general, not recover a diagonal matrix with a single sample, $N = 1$.

As s increases, so do the constants $K_1(s)$ and $K_2(s)$, and the upper bound on $\mathbb{P}[\|\mathcal{D}(\mathbf{A}) - \mathcal{D}(\hat{\mathbf{A}})\|_2 \geq t]$. In other words, the sparser the vectors \mathbf{w}_k , the less accurate the Monte Carlo estimate $\mathcal{D}(\hat{\mathbf{A}})$.

REMARK 3.4 (Non-integer sparsity levels). *The restriction to integers s in Theorem 3.5 is relevant only for $1 < s < 2$. More generally, Theorem 3.5 holds for $s = 1$, and any real number $s \geq 2$.*

The extension below of Corollary 3.3 presents the minimal sampling amount that makes the sparse Rademacher Monte Carlo estimator a normwise (ϵ, δ) diagonal estimator.

COROLLARY 3.6. *Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be non-diagonal symmetric, and let*

$$\begin{aligned} K_1(s) &\equiv \|\mathcal{D}(\mathbf{A}^2) + (s-2)\mathcal{D}(\mathbf{A})^2\|_2, & \Delta_1(s) &\equiv \frac{K_1(s)}{\|\mathcal{D}(\mathbf{A})\|_2} \\ \Delta_2(s) &\equiv \frac{\|s\mathbf{A} - \mathcal{D}(\mathbf{A})\|_\infty}{\|\mathcal{D}(\mathbf{A})\|_\infty}, & d &\equiv \frac{\|\mathbf{A}\|_F^2 + (s-2)\|\mathcal{D}(\mathbf{A})\|_F^2}{K_1(s)}. \end{aligned}$$

Let $\hat{\mathbf{A}} \equiv \frac{1}{N} \sum_{j=1}^N \mathbf{A} \mathbf{w}_j \mathbf{w}_j^\top$ be a Monte Carlo estimator with independent sparse Rademacher random vectors $\mathbf{w}_k \in \mathbb{R}^n$, with integer parameter $s \geq 1$. Pick $\epsilon > 0$. For any $0 < \delta < 1$, if the sampling amount is at least

$$N \geq \frac{\Delta_2(s)}{3\epsilon^2} \left(2\epsilon + 6 \frac{\Delta_1(s)}{\Delta_2(s)} \right) \ln(8d(s)/\delta),$$

then $\|\mathcal{D}(\mathbf{A}) - \mathcal{D}(\hat{\mathbf{A}})\|_2 \leq \epsilon \|\mathcal{D}(\mathbf{A})\|_2$ holds with probability at least $1 - \delta$.

Proof. Denote the bound for the failure probability in Theorem 3.5 by

$$\delta \equiv 8d(s) \exp\left(\frac{-Nt^2}{2(K_1(s) + tK_2(s)/3)}\right),$$

and solve it for t ,

$$t = \frac{K_2(s)}{3N} \ln(8d(s)/\delta) + \sqrt{\frac{K_2(s)^2}{9N^2} \ln^2(8d(s)/\delta) + \frac{2K_1(s)}{N} \ln(8d(s)/\delta)}.$$

We can restate Theorem 3.2 in terms of the failure probability: With probability at most δ , the normwise absolute error exceeds $\|\mathcal{D}(\mathbf{A}) - \mathcal{D}(\hat{\mathbf{A}})\|_2 \geq t$, where

$$t = \frac{K_2(s)}{3N} \ln(8d(s)/\delta) + \sqrt{\frac{K_2(s)^2}{9N^2} \ln^2(8d(s)/\delta) + \frac{2K_1(s)}{N} \ln(8d(s)/\delta)}.$$

And in terms of the success probability: With probability at least $1 - \delta$, the normwise absolute error is bounded above by $\|\mathcal{D}(\mathbf{A}) - \mathcal{D}(\hat{\mathbf{A}})\|_2 \leq t$. Converting this absolute error into a relative one requires $t \leq \epsilon \|\mathcal{D}(\mathbf{A})\|_2$, in other words,

$$\frac{t}{\|\mathcal{D}(\mathbf{A})\|_2} = \gamma + \sqrt{\gamma^2 + \beta} \leq \epsilon,$$

where

$$\gamma \equiv \frac{K_2(s) \ln(8d(s)/\delta)}{3N \|\mathcal{D}(\mathbf{A})\|_2}, \quad \beta \equiv \frac{6K_1(s)}{K_2(s) \|\mathcal{D}(\mathbf{A})\|_2}$$

Solving for γ gives $\gamma \leq \frac{\epsilon^2}{2\epsilon + \beta}$, which implies

$$\begin{aligned} N &\geq \frac{K_2(s)}{3\epsilon^2 \|\mathcal{D}(\mathbf{A})\|_2} \left(2\epsilon + \frac{6K_1(s)}{K_2(s) \|\mathcal{D}(\mathbf{A})\|_2} \right) \ln(8d(s)/\delta) \\ &= \frac{1}{3\epsilon^2} \frac{K_2(s)}{\|\mathcal{D}(\mathbf{A})\|_2} \left(2\epsilon + 6 \frac{K_1(s)}{\|\mathcal{D}(\mathbf{A})\|_2^2} \frac{\|\mathcal{D}(\mathbf{A})\|_2}{K_2(s)} \right) \ln(8d(s)/\delta) \\ &= \frac{\Delta_2}{3\epsilon^2} \left(2\epsilon + 6 \frac{\Delta_1(s)}{\Delta_2(s)} \right) \ln(8d(s)/\delta), \end{aligned}$$

where we exploited the norms of diagonal matrices,

$$\Delta_1(s) = \frac{K_1(s)}{\|\mathcal{D}(\mathbf{A})\|_2^2}, \quad \Delta_2 = \frac{\|\mathbf{A} - \mathcal{D}(\mathbf{A})\|_\infty}{\|\mathcal{D}(\mathbf{A})\|_\infty} = \frac{K_2(s)}{\|\mathcal{D}(\mathbf{A})\|_2}. \quad \square$$

Corollary 3.6 suggests that increasing the sparsity parameter s could on the one hand lower the computational cost per sample, but on the other hand increase the sampling amount for the same accuracy.

4. Gaussian vectors. We present normwise bounds for random vectors with bounded fourth moment (section 4.1), and standard Gaussian vectors (section 4.2).

4.1. Random vectors with bounded fourth moment. We bound the expectation of the squared absolute error (Theorem 4.1) for Monte Carlo estimators based on random vectors \mathbf{w}_k (assumed to have independent entries with zero mean and variance 1) with bounded fourth moment, $1 \leq k \leq N$,

$$(4.1) \quad \mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{w}_k\|_\infty^4 \right] < +\infty.$$

These include standard Rademacher (section 3.1) and sparse Rademacher vectors (section 3.2), as well as standard Gaussian vectors (section 4.2).

THEOREM 4.1. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ with $n \geq 3$ be symmetric, and

$$\widehat{\mathbf{A}} \equiv \frac{1}{N} \sum_{j=1}^N \mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top$$

be a Monte Carlo estimator with independent random vectors $\mathbf{w}_k \in \mathbb{R}^n$, $1 \leq k \leq N$, that have independent elements with zero mean and variance 1. If the vectors \mathbf{w}_k also have a bounded fourth moment (4.1), then

$$\mathbb{E} \left[\|\mathcal{D}(\widehat{\mathbf{A}}) - \mathcal{D}(\mathbf{A})\|_2^2 \right]^{1/2} \leq \|\mathbf{A}\|_\infty \left(\sqrt{\frac{8e \ln n}{N}} + \frac{8e \ln n}{N} \right) \left(\mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{w}_k\|_\infty^4 \right] \right)^{1/2}.$$

Proof. We make use of matrix concentration inequalities but follow the spirit of the analysis in [5].

1. *Symmetrization.* Write the normwise error by exploiting diagonal Schur products (2.1)

$$\mathcal{D}(\widehat{\mathbf{A}}) - \mathcal{D}(\mathbf{A}) = \mathbf{I} \circ \widehat{\mathbf{A}} - \mathbf{I} \circ \mathbf{A} = \frac{1}{N} \sum_{k=1}^N (\mathbf{I} \circ (\mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top) - \mathbf{I} \circ \mathbf{A}),$$

and take expectations of the squared norms

$$\mathbb{E} \left[\left\| \mathcal{D}(\widehat{\mathbf{A}}) - \mathcal{D}(\mathbf{A}) \right\|_2^2 \right] = \frac{1}{N^2} \mathbb{E} \left[\left\| \sum_{k=1}^N (\mathbf{I} \circ (\mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top) - \mathbf{I} \circ \mathbf{A}) \right\|_2^2 \right].$$

From the assumption that \mathbf{w}_k has independent elements with zero mean and variance 1 follows $\mathbb{E} [\mathbf{I} \circ (\mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top)] = \mathbf{I} \circ \mathbf{A}$. Hence the matrix random variables

$$\mathbf{X}_k \equiv \mathbf{I} \circ (\mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top) - \mathbf{I} \circ \mathbf{A}, \quad 1 \leq k \leq N$$

have zero mean. We use symmetrization [20, Lemma 6.4.2] to create *symmetric* random variables $\varepsilon_k \mathbf{X}_k$, where ε_k are independent symmetric Bernoulli random variables, that is, they are Rademacher variables as in Definition 3.1. The Rademacher variables ε_k are independent of each other and also independent of the random vectors \mathbf{w}_k . Remark 3.1 implies $\mathbb{E}[\varepsilon_k] = 0$, hence

$$(4.2) \quad \mathbb{E} \left[\sum_{k=1}^N \varepsilon_k \mathbf{I} \circ \mathbf{A} \right] = \mathbf{0}.$$

Then [20, Lemma 6.4.2], (4.2) and (2.2) imply

$$(4.3) \quad \mathbb{E} \left[\left\| \sum_{k=1}^N \mathbf{X}_k \right\|_2^2 \right] \leq 2 \mathbb{E} \left[\left\| \sum_{k=1}^N \varepsilon_k \mathbf{I} \circ (\mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top) \right\|_2^2 \right] = 2 \mathbb{E} \left[\left\| \sum_{k=1}^N \mathbf{Y}_k \right\|_2^2 \right],$$

where $\mathbf{Y}_k \equiv \varepsilon_k \text{diag}(\mathbf{A} \mathbf{w}_k) \text{diag}(\mathbf{w}_k)$ are symmetric random matrices, and the second and third expectations range over all random vectors \mathbf{w}_k and all Rademacher variables ε_k .

2. *Concentration inequality.* Applying Cauchy-Schwartz inequality and Theorem 2.4 to the sum $\mathbf{Z} \equiv \sum_{k=1}^N \mathbf{Y}_k$ gives

$$(4.4) \quad \mathbb{E}[\|\mathbf{Z}\|_2^2]^{1/2} \leq \sqrt{2e \ln n} \left\| \left(\sum_{i=1}^N \mathbb{E}[\mathbf{Y}_k^2] \right)^{1/2} \right\|_2 + 4e \ln n \mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{Y}_k\|_2^2 \right]^{1/2}.$$

We bound the expectations that represent the matrix variance and the maximal two-norm separately.

3. *Variance.* As in item 3 of the proof of Theorem 3.5 abbreviate

$$\mathbf{D}_k \equiv \text{diag}(\mathbf{A}\mathbf{w}_k), \quad \mathbf{W}_k \equiv \text{diag}(\mathbf{w}_k), \quad \mathbf{O} \equiv \text{diag}([\|\mathbf{a}_1\|_1 \quad \cdots \quad \|\mathbf{a}_n\|_1]).$$

With this notation, $\mathbb{E}[\mathbf{Y}_k^2] = \mathbb{E}[\mathbf{D}_k^2 \mathbf{W}_k^2]$. Consider the diagonal term

$$(\mathbf{D}_k^2 \mathbf{W}_k^2)_{ii} = \left(\sum_{j=1}^n a_{ij}(\mathbf{w}_k)_j \right)^2 (\mathbf{w}_k)_i^2 \leq \|\mathbf{a}_i\|_1^2 \|\mathbf{w}_k\|_\infty^4 \quad 1 \leq i \leq n.$$

Therefore, $\mathbf{D}_k^2 \mathbf{W}_k^2 \preceq \|\mathbf{w}_k\|_\infty^4 \mathbf{O}^2$. The symmetry of \mathbf{A} implies $\mathbf{O} \preceq \|\mathbf{A}\|_\infty \mathbf{I}$. Combining the two inequalities gives

$$(4.5) \quad \mathbb{E}[\mathbf{Y}_k^2] = \mathbb{E}[\mathbf{D}_k^2 \mathbf{W}_k^2] \preceq \|\mathbf{A}\|_\infty^2 \mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{w}_k\|_\infty^4 \right] \mathbf{I}.$$

Now take square roots,

$$\left(\sum_{k=1}^N \mathbb{E}[\mathbf{Y}_k^2] \right)^{1/2} \preceq \sqrt{N} \|\mathbf{A}\|_\infty \mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{w}_k\|_\infty^4 \right]^{1/2} \mathbf{I},$$

and bound the norm,

$$(4.6) \quad \left\| \left(\sum_{k=1}^N \mathbb{E}[\mathbf{Y}_k^2] \right)^{1/2} \right\|_2 \leq \sqrt{N} \|\mathbf{A}\|_\infty \mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{w}_k\|_\infty^4 \right]^{1/2}.$$

4: *Maximal two-norm.* In analogy to (4.5), we derive

$$\mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{Y}_k\|_2^2 \right] \leq \|\mathbf{A}\|_\infty^2 \mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{w}_k\|_\infty^4 \right],$$

and its square root

$$(4.7) \quad \mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{Y}_k\|_2^2 \right]^{1/2} \leq \|\mathbf{A}\|_\infty \mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{w}_k\|_\infty^4 \right]^{1/2}.$$

5. *Putting everything together.* Substitute the variance bound (4.6) and the norm bound (4.7) into the expectation (4.4) for the sum

$$\mathbb{E}[\|\mathbf{Z}\|_2]^{1/2} \leq \left(\sqrt{2e \ln n} N^{1/2} + 4e \ln n \right) \|\mathbf{A}\|_\infty \mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{w}_k\|_\infty^4 \right]^{1/2},$$

substitute this, in turn, into the expectation (4.3) for the absolute error,

$$\begin{aligned}\mathbb{E} \left[\|\mathcal{D}(\hat{\mathbf{A}}) - \mathcal{D}(\mathbf{A})\|_2^2 \right]^{1/2} &\leq \frac{2}{N} \mathbb{E} [\|\mathbf{Z}\|_2]^{1/2} \\ &\leq \frac{2}{N} \left(\sqrt{2e \ln n} N^{1/2} + 4e \ln n \right) \|\mathbf{A}\|_\infty \mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{w}_k\|_\infty^4 \right]^{1/2},\end{aligned}$$

and simplify. \square

4.2. Gaussian vectors. We determine the minimal sampling amount for Gaussian Monte Carlo estimators to be normwise (ϵ, δ) diagonal estimators.

COROLLARY 4.2. *Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ with $n \geq 3$ be symmetric, and let*

$$\hat{\mathbf{A}} \equiv \frac{1}{N} \sum_{j=1}^N \mathbf{A} \mathbf{w}_j \mathbf{w}_j^\top$$

be a Monte Carlo estimator with independent Gaussian random vectors $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ in \mathbb{R}^n , $1 \leq k \leq N$. Pick $\epsilon > 0$. For any $0 < \delta < 1$, if the sampling amount N satisfies $8e \ln n \leq N \leq n$, and is at least

$$(4.8) \quad N \geq \frac{128 (e \ln n)^3}{\epsilon^2 \delta} \left(\frac{\|\mathbf{A}\|_\infty}{\|\mathcal{D}(\mathbf{A})\|_\infty} \right)^2,$$

then $\|\mathcal{D}(\hat{\mathbf{A}}) - \mathcal{D}(\mathbf{A})\|_2 \leq \epsilon \|\mathcal{D}(\mathbf{A})\|_2$ holds with probability at least $1 - \delta$.

Proof. For Gaussian random vectors $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, [5, (3.7)] implies

$$\left(\mathbb{E} \left[\max_{1 \leq k \leq N} \|\mathbf{w}_k\|_\infty^4 \right] \right)^{1/2} \leq e \ln(nN) \max_{1 \leq i, j \leq n} |\mathbf{I}_{ij}| = e \ln(nN).$$

Substituting this into Theorem 4.1 gives

$$\left(\mathbb{E} \left[\|\mathcal{D}(\hat{\mathbf{A}}) - \mathcal{D}(\mathbf{A})\|_2^2 \right] \right)^{1/2} \leq \left(\sqrt{\frac{8e \ln n}{N}} + \frac{8e \ln n}{N} \right) e \ln(nN) \|\mathbf{A}\|_\infty.$$

Square both sides and apply Markov's inequality (Theorem 2.1) to the random variable $Z \equiv \|\mathcal{D}(\hat{\mathbf{A}}) - \mathcal{D}(\mathbf{A})\|_2$ using $t \equiv \epsilon \|\mathcal{D}(\mathbf{A})\|_2$,

$$(4.9) \quad \begin{aligned}\mathbb{P} \left[\|\mathcal{D}(\hat{\mathbf{A}}) - \mathcal{D}(\mathbf{A})\|_2 \geq \epsilon \|\mathcal{D}(\mathbf{A})\|_2 \right] \\ \leq \left(\sqrt{\frac{8e \ln n}{N}} + \frac{8e \ln n}{N} \right)^2 (e \ln(nN))^2 \frac{\|\mathbf{A}\|_\infty^2}{\epsilon^2 \|\mathcal{D}(\mathbf{A})\|_\infty^2}.\end{aligned}$$

Substituting the assumption $8e \ln n \leq N \leq n$ into the relevant part of the above bound gives

$$\left(\sqrt{\frac{8e \ln n}{N}} + \frac{8e \ln n}{N} \right)^2 (e \ln(nN))^2 \leq \left(2\sqrt{\frac{8e \ln n}{N}} \right)^2 (2e \ln n)^2 = \frac{128 (e \ln n)^3}{N}.$$

Substitute this, in turn, into (4.9), set the failure probability equal to

$$\delta \equiv \frac{128 (e \ln n)^3}{\epsilon^2 N} \left(\frac{\|\mathbf{A}\|_\infty}{\|\mathcal{D}(\mathbf{A})\|_\infty} \right)^2$$

and solve for the sampling amount N . \square

5. Componentwise bounds. We present componentwise bounds for Monte Carlo estimators based on standard Rademacher vectors (section 5.1), as well as on standard (section 5.2) and normalized Gaussian vectors (section 5.3).

Our attempt at deriving alternative normwise bounds by applying a union bound over the componentwise bounds for all diagonal elements did not produce results that were substantially tighter than our previous normwise bounds.

5.1. Standard Rademacher vectors. We present a componentwise worst case absolute error bound (Corollary 5.1), and a bound on the minimal sampling amount that makes the Rademacher Monte Carlo estimator a *componentwise* (ϵ, δ) diagonal estimator (Theorem 5.3).

Theorem 3.2 is identical to the following worst case componentwise bound.

COROLLARY 5.1. *Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be non-diagonal symmetric, and let*

$$K_1 \equiv \|\mathcal{D}(\mathbf{A}^2) - \mathcal{D}(\mathbf{A})^2\|_2, \quad K_2 \equiv \|\mathbf{A} - \mathcal{D}(\mathbf{A})\|_\infty, \quad d \equiv \frac{1}{K_1} \|\mathbf{A} - \mathcal{D}(\mathbf{A})\|_F^2.$$

If $\hat{\mathbf{A}} \equiv \frac{1}{N} \sum_{k=1}^N \mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top$ is a Monte Carlo estimator with independent Rademacher vectors $\mathbf{w}_k \in \mathbb{R}^n$, $1 \leq k \leq N$, then the probability that the absolute error exceeds $t > 0$ is at most

$$\mathbb{P} \left[\max_{1 \leq i \leq n} |a_{ii} - \hat{a}_{ii}| \geq t \right] \leq 8d \exp \left(\frac{-Nt^2}{2(K_1 + tK_2/3)} \right).$$

Proof. In Theorem 3.2, the p -norm of the diagonal matrix $\mathcal{D}(\mathbf{A}) - \mathcal{D}(\hat{\mathbf{A}})$ is a largest magnitude diagonal element. \square

In contrast to Corollary 5.1, the next bound depends on the particular diagonal element. We first require a lemma on the independence of products of Rademacher variables.

LEMMA 5.2. *Let $Z_1, W_1, W_2, \dots, W_n$ be independent Rademacher variables. Then the products $X_1 \equiv ZW_1, \dots, X_n \equiv ZW_n$ are also independent Rademacher variables.*

Proof. For any $x_1, \dots, x_n \in \{-1, +1\}$, using the law of total probability, the joint probability mass function satisfies

$$\begin{aligned} \mathbb{P}[\cap_{i=1}^n \{X_i = x_i\}] &= \sum_{z \in \{-1, +1\}} \mathbb{P}[\cap_{i=1}^n \{X_i = x_i\} | Z = z] \mathbb{P}[Z = z] \\ &= \frac{1}{2} \mathbb{P}[\cap_{i=1}^n \{W_i = -x_i\}] + \frac{1}{2} \mathbb{P}[\cap_{i=1}^n \{W_i = x_i\}] \\ &= 2^{-n} = \prod_{i=1}^n \mathbb{P}[X_i = x_i]. \end{aligned}$$

Since the joint PMF factorizes, the variables X_1, \dots, X_n are independent. \square

THEOREM 5.3. *Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be non-diagonal symmetric, and let*

$$\hat{\mathbf{A}} \equiv \frac{1}{N} \sum_{k=1}^N \mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top$$

be a Monte Carlo estimator with independent Rademacher vectors $\mathbf{w}_k \in \mathbb{R}^n$, $1 \leq k \leq N$. The probability that the absolute error exceeds $t > 0$ is at most

$$(5.1) \quad \mathbb{P}[|\hat{a}_{ii} - a_{ii}| \geq t] \leq 2 \exp \left(\frac{-Nt^2}{2(\|\mathbf{a}_i\|_2^2 - a_{ii}^2)} \right), \quad 1 \leq j \leq n.$$

Proof. Fix i , for some $1 \leq i \leq n$. The properties in Remark 3.1 allow us to split off the original diagonal element from the estimator,

$$\begin{aligned}\hat{a}_{ii} &= \frac{1}{N} \sum_{k=1}^N (\mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top)_{ii} = \frac{1}{N} \sum_{k=1}^N \sum_{j=1}^n a_{ij} (\mathbf{w}_k)_i (\mathbf{w}_k)_j \\ &= a_{ii} + \sum_{k=1}^N \sum_{j \neq i} \underbrace{\frac{a_{ij}}{N} (\mathbf{w}_k)_i (\mathbf{w}_k)_j}_{Z_{kj}^{(i)}}.\end{aligned}$$

Lemma 5.2 implies that for fixed i , the $Z_{kj}^{(i)}$ are independent. Remark 3.1 implies that they have zero mean, and are bounded by

$$-\frac{|a_{ij}|}{N} \leq Z_{kj}^{(i)} \leq \frac{|a_{ij}|}{N}, \quad 1 \leq k \leq N, \quad 1 \leq j \leq n, \quad j \neq i.$$

Hence the absolute error $\hat{a}_{ii} - a_{ii}$ is a sum of independent bounded zero-mean random variables, and we can apply Hoeffding's inequality in Theorem 2.2

$$\mathbb{P} [|\hat{a}_{ii} - a_{ii}| \geq t] \leq 2 \exp \left(\frac{-2t^2}{\sum_{k=1}^N \sum_{j \neq i} \left(\frac{2}{N} |a_{ij}| \right)^2} \right) = 2 \exp \left(\frac{-Nt^2}{2 \sum_{j \neq i} a_{ij}^2} \right).$$

At last, write $\sum_{j \neq i} a_{ij}^2 = \|\mathbf{a}_i\|_2^2 - a_{ii}^2$. \square

Theorem 5.3 implies that the accuracy for estimating a single diagonal element depends only on the magnitude of the off-diagonal elements in the corresponding row and column.

We determine the minimal sampling amount required to make the Rademacher Monte Carlo estimator a componentwise (ϵ, δ) diagonal estimator. For symmetric matrices, this result coincides with the bound in Equation 40 of [3] but uses a different proof technique.

COROLLARY 5.4. *Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be non-diagonal symmetric, and let*

$$\hat{\mathbf{A}} \equiv \frac{1}{N} \sum_{k=1}^N \mathbf{A} \mathbf{w}_k \mathbf{w}_k^\top$$

be a Monte Carlo estimator with independent Rademacher vectors $\mathbf{w}_k \in \mathbb{R}^n$, $1 \leq k \leq N$. Pick $\epsilon > 0$, and a diagonal element $a_{ii} \neq 0$ of \mathbf{A} . For any $0 < \delta < 1$, if the sampling amount is at least

$$N \geq \left(\frac{\|\mathbf{a}_i\|_2^2 - a_{ii}^2}{a_{ii}^2} \right) \frac{2 \ln(2/\delta)}{\epsilon^2},$$

then $|a_{jj} - \hat{a}_{jj}| \leq \epsilon |a_{jj}|$ holds with probability at least $1 - \delta$.

Proof. Define the 2-norm offdiagonal column sums

$$\text{off}_i \equiv (\|\mathbf{a}_i\|_2^2 - a_{ii}^2)^{1/2}, \quad 1 \leq i \leq n,$$

and denote the bound for the failure probability in Theorem 5.3 by

$$\delta \equiv 2 \exp \left(\frac{-Nt^2}{2 \text{off}_i^2} \right),$$

and solve it for t ,

$$t = \sqrt{\frac{2 \text{off}_i^2}{N} \ln(2/\delta)}.$$

Restate Theorem 5.3 in terms of the failure probability: With probability at most $1 - \delta$, the absolute error of a specific diagonal element is bounded above by

$$|a_{ii} - \hat{a}_{ii}| \leq t = \sqrt{\frac{2 \text{off}_i^2}{N} \ln(2/\delta)}.$$

Converting this absolute error into a relative error requires $t \leq \epsilon |a_{ii}|$, which implies

$$N \geq \left(\frac{\text{off}_i}{|a_{ii}|} \right)^2 \frac{2 \ln(2/\delta)}{\epsilon^2}. \quad \square$$

The minimal sampling amount for computing a_{ii} with the Rademacher Monte Carlo estimator depends on $\frac{\|\mathbf{a}_i\|_2^2 - a_{ii}^2}{|a_{ii}|}$, which represents the relative 2-norm deviation of the i th column and row of \mathbf{A} from diagonality. Thus, the more diagonal the i th row and column, the fewer samples are required for a (ϵ, δ) estimator.

5.2. Gaussian vectors. We present a componentwise absolute error bound (Theorem 5.5) for Gaussian Monte Carlo estimators, and a bound on the minimal sampling amount that makes the Gaussian Monte Carlo estimator a componentwise (ϵ, δ) estimators (Corollary 5.6). Our bounds are derived from and identical to bounds for trace estimators in [7].

THEOREM 5.5. *Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be non-diagonal symmetric,*

$$L_{1i} \equiv |a_{ii}| + \|\mathbf{a}_i\|_2, \quad L_{i2} \equiv |a_{ii}|^2 + \|\mathbf{a}_i\|_2^2, \quad 1 \leq i \leq n,$$

and let $\hat{\mathbf{A}} \equiv \frac{1}{N} \sum_{k=1}^N \mathbf{A} \mathbf{z}_k \mathbf{z}_k^\top \in \mathbb{R}^{n \times n}$ be a Monte Carlo estimator with independent Gaussian vectors $\mathbf{z}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ in \mathbb{R}^n , $1 \leq k \leq N$. If $t > 0$, then

$$(5.2) \quad \mathbb{P}[|\hat{a}_{ii} - a_{ii}| > t] \leq 2 \exp\left(\frac{-Nt^2}{2(L_{i2} + tL_{i1})}\right), \quad 1 \leq i \leq n.$$

Proof. Fix i for some $1 \leq i \leq n$. Write the diagonal element as an inner product

$$(\mathbf{A} \mathbf{z}_k \mathbf{z}_k^\top)_{ii} = \sum_{j=1}^n a_{ij}(\mathbf{z}_k)_i(\mathbf{z}_k)_j = \mathbf{z}_k^\top \mathbf{B}_i \mathbf{z}_k, \quad 1 \leq k \leq N$$

involving the symmetric matrix

$$\mathbf{B}_i \equiv \begin{bmatrix} 0 & \frac{1}{2}a_{1i} & & 0 \\ & \vdots & & \\ \frac{1}{2}a_{i1} & \cdots & a_{ii} & \cdots & \frac{1}{2}a_{in} \\ & & \vdots & & \\ 0 & \frac{1}{2}a_{ni} & & 0 \end{bmatrix} \in \mathbb{R}^{n \times n} \quad \text{with} \quad \text{trace}(\mathbf{B}_i) = a_{ii}.$$

We can think of

$$\hat{a}_{ii} = \left(\frac{1}{N} \sum_{k=1}^N \mathbf{A} \mathbf{z}_k \mathbf{z}_k^\top \right)_{ii} = \frac{1}{N} \sum_{k=1}^N (\mathbf{A} \mathbf{z}_k \mathbf{z}_k^\top)_{ii} = \frac{1}{N} \sum_{k=1}^N \mathbf{z}_k^\top \mathbf{B}_i \mathbf{z}_k$$

as a Monte Carlo estimator for $\text{trace}(\mathbf{B}_i)$, and apply the bound for Gaussian trace estimators [7, Theorem 1]

$$\mathbb{P}[|\hat{a}_{ii} - a_{ii}| \geq t] \leq 2 \exp\left(\frac{-Nt^2}{4\|\mathbf{B}_i\|_F^2 + 4t\|\mathbf{B}_i\|_2}\right),$$

where

$$\|\mathbf{B}_i\|_F^2 = \frac{1}{2}(|a_{ii}|^2 + \|\mathbf{a}_i\|_2^2) = \frac{L_{i2}}{2}, \quad \|\mathbf{B}_i\|_2 = \frac{1}{2}(|a_{ii}| + \|\mathbf{a}_i\|_2) = \frac{L_{i1}}{2}. \quad \square$$

Theorem 5.5 implies that with Gaussian vectors, the accuracy for estimating a single diagonal element depends on the magnitude of all elements in the corresponding row and column. By contrast, the bounds of Theorem 5.3 for Rademacher vectors depend only on the magnitude of the off-diagonal elements.

We determine the minimal sampling amount required to make the Gaussian Monte Carlo estimator a componentwise (ϵ, δ) estimator.

COROLLARY 5.6. *Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be non-diagonal symmetric,*

$$\Delta_{1i} \equiv 1 + \frac{\|\mathbf{a}_i\|_2}{|a_{ii}|}, \quad \Delta_{i2} \equiv 1 + \left(\frac{\|\mathbf{a}_i\|_2}{|a_{ii}|}\right)^2, \quad 1 \leq i \leq n,$$

and let $\hat{\mathbf{A}} \equiv \frac{1}{N} \sum_{k=1}^N \mathbf{A} \mathbf{z}_k \mathbf{z}_k^\top \in \mathbb{R}^{n \times n}$ be a Monte Carlo estimator with independent Gaussian vectors $\mathbf{z}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ in \mathbb{R}^n , $1 \leq k \leq N$. Pick $\epsilon > 0$, and a diagonal element $a_{ii} \neq 0$ of \mathbf{A} . For any $0 < \delta < 1$, if the sampling amount is at least

$$N \geq (\Delta_{2i} + \Delta_{1i}\epsilon) \frac{2 \ln(2/\delta)}{\epsilon^2}$$

then $|\hat{a}_{ii} - a_{ii}| \leq \epsilon |a_{ii}|$ holds with probability at least $1 - \delta$.

Proof. This follows immediately from the lower bound for N in [7, Theorem 1]. \square

The required sampling amount for computing a_{ii} with the Gaussian Monte Carlo estimator depends on $\|\mathbf{a}_i\|_2/|a_{ii}|$ which can be interpreted as the 2-norm derivation of the i th column and row of \mathbf{A} from diagonality. The more diagonal the i th row and column, the smaller the sampling amount for the (ϵ, δ) estimator.

5.3. Normalized Gaussian vectors. We extend and complete the analysis in [4] for a Monte Carlo estimator based on normalized Gaussian vectors,

$$(5.3) \quad \hat{\mathbf{A}} \equiv \left(\sum_{k=1}^N \mathbf{A} \mathbf{z}_k \mathbf{z}_k^\top \right) \oslash \left(\sum_{k=1}^N \mathbf{z}_k \mathbf{z}_k^\top \right),$$

where $\mathbf{z}_k \in \mathbb{R}^n$ are independent random vectors, and \oslash denotes elementwise division. We derive the distribution of the componentwise absolute errors (Lemma 5.7), followed by a bound (Theorem 5.8).

We represent the distribution for the absolute errors in the diagonal elements in terms of a *Student t-distribution* with $N \geq 1$ degrees of freedom [13, Definition 7.3.3],

$$(5.4) \quad T_N \equiv \frac{Z}{\sqrt{U/N}},$$

where Z is a Gaussian $\mathcal{N}(0, 1)$ random variable, and U an independent chi-square random variable with N degrees of freedom.

LEMMA 5.7. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be symmetric, and (5.3) be a Monte Carlo estimator where $\mathbf{z}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ in \mathbb{R}^n are independent Gaussian random vectors, $1 \leq k \leq N$. The absolute errors in the diagonal elements are distributed as

$$\hat{a}_{ii} - a_{ii} \sim \sqrt{\frac{\|\mathbf{a}_i\|_2^2 - a_{ii}^2}{N}} T_N, \quad 1 \leq i \leq n.$$

Proof. Due to the normalization in the denominator, we can extract the diagonal elements of \mathbf{A} from the diagonal elements of the Monte Carlo estimator $\hat{\mathbf{A}}$,

$$\hat{a}_{ii} = \frac{\sum_{k=1}^N \sum_{j=1}^n a_{ij}(\mathbf{z}_k)_i(\mathbf{z}_k)_j}{\sum_{k=1}^N (\mathbf{z}_k)_i^2} = a_{ii} + \frac{\sum_{k=1}^N \sum_{j \neq i} a_{ij}(\mathbf{z}_k)_i(\mathbf{z}_k)_j}{\sum_{k=1}^N (\mathbf{z}_k)_i^2}, \quad 1 \leq i \leq n.$$

Normalize across the i th elements of the Gaussian vectors \mathbf{z}_k to unit vectors $\mathbf{u}_i \in \mathbb{R}^N$ with elements

$$(\mathbf{u}_i)_k \equiv \frac{(\mathbf{z}_k)_i}{\sqrt{\sum_{k'=1}^N (\mathbf{z}_{k'})_i^2}}, \quad 1 \leq k \leq N, \quad 1 \leq i \leq n.$$

Use the denominator to normalize the i th component in the i th absolute error,

$$(5.5) \quad \hat{a}_{ii} - a_{ii} = \frac{\sum_{k=1}^N \sum_{j \neq i} a_{ij}(\mathbf{z}_k)_j(\mathbf{u}_i)_k}{\sqrt{\sum_{k'=1}^N (\mathbf{z}_{k'})_i^2}}, \quad 1 \leq i \leq n.$$

The rotational invariance of the standard Gaussian distribution guarantees the independence of the direction vectors \mathbf{u}_i and radial components $\sqrt{\sum_{k=1}^N (\mathbf{z}_k)_i^2}$; see [20, Exercise 3.3.6]. Hence, the numerator and denominator in (5.5) are independent. We can rewrite (5.5) as

$$\hat{a}_{ii} - a_{ii} = \frac{Z_i}{\sqrt{U_i/N}},$$

where we define the random variables $Z_i \equiv \sum_{k=1}^N \sum_{j \neq i} a_{ij}(\mathbf{z}_k)_j(\mathbf{u}_i)_k / \sqrt{N}$ and $U_i \equiv \sum_{k=1}^N (\mathbf{z}_k)_i^2$ for $1 \leq i \leq n$. The random variable U_i has a chi-square distribution with N degrees of freedom. The conditional distribution of Z_i given \mathbf{u}_i is Gaussian (see e.g. Exercise 3.3.3(a) of [20]), with zero mean and variance

$$\frac{1}{N} \sum_{k=1}^N \sum_{j \neq i} a_{ij}^2 (\mathbf{u}_i)_k^2 = \frac{1}{N} \left(\sum_{j \neq i} a_{ij}^2 \right) \left(\sum_{k=1}^N (\mathbf{u}_i)_k^2 \right) = \frac{1}{N} \sum_{j \neq i} a_{ij}^2 = \frac{1}{N} (\|\mathbf{a}_i\|_2^2 - a_{ii}^2).$$

Therefore, $Z_i | \mathbf{u}_i \sim \mathcal{N}(0, \frac{1}{N} (\|\mathbf{a}_i\|_2^2 - a_{ii}^2))$. However, the conditional distribution is independent of \mathbf{u}_i , so this is also the unconditional distribution. The claim then follows from (5.4). \square

REMARK 5.1. It was observed in [3] that the square of the error $(\hat{a}_{ii} - a_{ii})^2$ has a scaled F -distribution. We note that the square of a Student t -distribution is specifically a scaled F -distribution with one degree of freedom in the numerator. Moreover, for a single sample $N = 1$ the error has a Cauchy distribution, which has undefined mean and variance.

If N is large, then the t -distribution T_N can be approximated by a standard normal distribution. However, T_N has wider tails, thus somewhat weaker tail bounds. Existing tail bounds for the Student t -distribution imply the following concentration inequality for error bounds.

THEOREM 5.8. *Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be symmetric, and (5.3) be a Monte Carlo estimator where $\mathbf{z}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ in \mathbb{R}^n are independent Gaussian random vectors, $1 \leq k \leq N$. For any $t > 0$,*

$$\mathbb{P} [|\hat{a}_{ii} - a_{ii}| > t] \leq \sqrt{\frac{2(\|\mathbf{a}_i\|_2^2 - a_{ii}^2)}{\pi N}} \frac{1}{t} \left(1 + \frac{t^2}{\|\mathbf{a}_i\|_2^2 - a_{ii}^2}\right)^{-\frac{N-1}{2}}, \quad 1 \leq i \leq n.$$

Proof. The probability density function of T_N is [17]

$$f_N(x) = c_N \left(1 + \frac{x^2}{N}\right)^{-\frac{N+1}{2}} \quad \text{where} \quad c_N \equiv \frac{\Gamma((N+1)/2)}{\Gamma(N/2)\sqrt{N\pi}},$$

and $1/\pi \leq c_N \leq 1/\sqrt{2\pi}$. If $F_N(x)$ is the cumulative distribution function for T_N then by [17, Theorem 3.1]

$$\begin{aligned} \mathbb{P} \left[(\hat{a}_{ii} - a_{ii}) > \sqrt{\frac{\|\mathbf{a}_i\|_2^2 - a_{ii}^2}{N}} x \right] &= 1 - F_N(x) < \frac{f_N(x)}{x} \left(1 + \frac{x^2}{N}\right) \\ &= \frac{c_N}{x} \left(1 + \frac{x^2}{N}\right)^{-\frac{N-1}{2}}. \end{aligned}$$

Take $x = t \sqrt{\frac{N}{\|\mathbf{a}_i\|_2^2 - a_{ii}^2}}$ and bound the upper tail with $c_N \leq \sqrt{\frac{1}{2\pi}}$, to obtain

$$\mathbb{P} [(\hat{a}_{ii} - a_{ii}) > t] \leq \frac{1}{t} \sqrt{\frac{\|\mathbf{a}_i\|_2^2 - a_{ii}^2}{2\pi N}} \left(1 + \frac{t^2}{\|\mathbf{a}_i\|_2^2 - a_{ii}^2}\right)^{-\frac{N-1}{2}}.$$

Since T_N is symmetric about the origin, the lower tail has the same bound. Using a union bound gives the desired inequality. \square

We determine a sampling amount sufficient to make the normalized Gaussian Monte Carlo estimator a componentwise (ϵ, δ) estimator.

COROLLARY 5.9. *Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be non-diagonal symmetric, and let*

$$\Psi_i \equiv \frac{|a_{ii}|}{(\|\mathbf{a}_i\|_2^2 - a_{ii}^2)^{1/2}}, \quad 1 \leq i \leq n,$$

and let $\hat{\mathbf{A}}$ be defined as in (5.3). Pick $\epsilon > 0$ and a diagonal element $a_{ii} \neq 0$ of \mathbf{A} . For any $0 < \delta < 1$, if the sampling amount is positive and at least

$$N \geq 1 + 2 \ln \left(\frac{\sqrt{2/\pi}}{\delta \epsilon \Psi_i} \right) / \ln(1 + \epsilon^2 \Psi_i^2),$$

then $|\hat{a}_{ii} - a_{ii}| \leq \epsilon |a_{ii}|$ holds with probability at least $1 - \delta$.

Proof. In Theorem 5.8, set $t = \epsilon|a_{ii}|$. If the sampling number satisfies the desired bound, it follows that

$$\begin{aligned}\mathbb{P}[(\hat{a}_{ii} - a_{ii}) > t] &\leq \sqrt{\frac{2(\|\mathbf{a}_i\|_2^2 - a_{ii}^2)}{\pi N}} \frac{1}{t} \left(1 + \frac{t^2}{\|\mathbf{a}_i\|_2^2 - a_{ii}^2}\right)^{-\frac{N-1}{2}} \\ &= \frac{\sqrt{2/\pi}}{\epsilon\Psi_i\sqrt{N}} (1 + \epsilon^2\Psi_i^2)^{-\frac{N-1}{2}} \\ &\leq \frac{\sqrt{2/\pi}}{\epsilon\Psi_i} (1 + \epsilon^2\Psi_i^2)^{-\frac{N-1}{2}}\end{aligned}$$

where the final inequality holds since $N \geq 1$ by assumption. Set the failure probability δ to the right hand side as

$$\delta \equiv \frac{\sqrt{2/\pi}}{\epsilon\Psi_i} (1 + \epsilon^2\Psi_i^2)^{-\frac{N-1}{2}},$$

and solve for N . □

The larger the value of Ψ_i (the same measure of diagonal dominance that appears in Corollary 5.4), the smaller the sampling amount for the (ϵ, δ) estimator.

6. Application: Monte Carlo estimators for a derivative-based global sensitivity metric. We bound the absolute error (Theorem 6.1) in a Monte Carlo estimator for global sensitivity analysis, and more specifically for a derivative-based global sensitivity metric (DGSM) of a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ whose partial derivatives are square integrable with respect to a probability density function $\rho_{\mathbf{x}}(\mathbf{x})$. The DGSM is equal to the diagonal $\mathcal{D}(\mathbf{C})$ of the matrix

$$\mathbf{C} = \int_{\mathcal{X}} \nabla f(\mathbf{x}) [\nabla f(\mathbf{x})]^\top \rho_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}.$$

The matrix \mathbf{C} is well-defined, symmetric positive semidefinite, and can be interpreted as a second moment matrix of the gradient. We compute the DGSM with the Monte Carlo estimator

$$(6.1) \quad \hat{\mathbf{C}} \equiv \frac{1}{N} \sum_{k=1}^N \mathbf{z}_k \mathbf{z}_k^\top \quad \text{where} \quad \mathbf{z}_k \equiv \nabla f(\mathbf{x}_k)$$

and \mathbf{x}_k , $1 \leq k \leq N$, are independent samples from the distribution of $\rho_{\mathbf{x}}(\mathbf{x})$. Below is a normwise bound for the error in the DGSM computed by the Monte Carlo estimator (6.1). Its derivation is related to the analysis in [12, Section 4].

THEOREM 6.1. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ have square integrable partial derivatives with respect to the probability density function $\rho_{\mathbf{x}}(\mathbf{x})$, $\|\nabla f\|_\infty \leq \beta$ almost surely, $\hat{\mathbf{C}}$ be the Monte Carlo estimator in (6.1), and*

$$\begin{aligned}c_{\max} &\equiv \|\mathcal{D}(\mathbf{C})\|_2 & S_1 &\equiv \|\mathcal{D}(\mathbf{C}) (\beta^2 \mathbf{I} - \mathcal{D}(\mathbf{C}))\|_2 \\ S_2 &\equiv c_{\max} + \beta^2, & d &\equiv \frac{\sum_{i=1}^n c_{ii}(\beta^2 - c_{ii})}{S_1}.\end{aligned}$$

If $c_{\max} > 0$ and $S_1 > 0$, then

$$\mathbb{P}\left[\|\mathcal{D}(\mathbf{C}) - \mathcal{D}(\hat{\mathbf{C}})\|_2 \geq t\right] \leq 8d \exp\left(\frac{-t^2/2}{S_1 + S_2 t/3}\right).$$

Proof. Before applying the matrix Bernstein inequality in Theorem 2.3, we need to verify the assumptions. The Monte Carlo estimate $\mathcal{D}(\hat{\mathbf{C}})$ is an unbiased estimator of the DGSM $\mathcal{D}(\mathbf{C})$ whose largest diagonal element is

$$c_{\max} = \|\mathcal{D}(\mathbf{C})\|_2 = \max_{1 \leq i \leq n} |c_{ii}| = \max_{1 \leq i \leq n} \mathbb{E}[(\nabla f(\mathbf{x}))_i^2] \leq \beta^2.$$

The absolute error in the DGSM computed by the Monte Carlo estimator (6.1) is

$$\mathbf{Z} = \mathcal{D}(\mathbf{C}) - \mathcal{D}(\hat{\mathbf{C}}) = \sum_{k=1}^N \mathbf{S}_k, \quad \text{where } \mathbf{S}_k \equiv \frac{1}{N} (\mathcal{D}(\mathbf{C}) - \mathcal{D}(\mathbf{z}_k \mathbf{z}_k^\top)).$$

The summands \mathbf{S}_k have zero mean and are bounded by

$$\begin{aligned} \|\mathbf{S}_k\|_2 &\leq \frac{1}{N} (\|\mathcal{D}(\mathbf{C})\|_2 + \|\mathcal{D}(\mathbf{z}_k \mathbf{z}_k^\top)\|_2) \\ &\leq \frac{\|\mathcal{D}(\mathbf{C})\|_2 + \beta^2}{N} = \frac{c_{\max} + \beta^2}{N} = \frac{S_2}{N}, \quad 1 \leq k \leq N. \end{aligned}$$

We let $L = S_2/N$ so that $\|\mathbf{S}_k\|_2 \leq L$. The variance is

$$(6.2) \quad \mathbb{V}\text{ar}[\mathbf{Z}] = \sum_{k=1}^N \mathbb{E}[\mathbf{S}_k^2] = \frac{1}{N^2} \sum_{k=1}^N \mathbb{E}[(\mathcal{D}(\mathbf{C}) - \mathcal{D}(\mathbf{z}_k \mathbf{z}_k^\top))^2].$$

Linearity of the expectation, the majorization $\mathcal{D}(\mathbf{z}_k \mathbf{z}_k^\top) \preceq \beta^2 \mathbf{I}$, $\mathbb{E}[\mathcal{D}(\mathbf{z}_k \mathbf{z}_k^\top)] = \mathcal{D}(\mathbf{C})$ and commutativity of diagonal matrices imply for the summands,

$$\begin{aligned} \mathbb{E}[(\mathcal{D}(\mathbf{C}) - \mathcal{D}(\mathbf{z}_k \mathbf{z}_k^\top))^2] &= \mathbb{E}[\mathcal{D}(\mathbf{C})^2 + (\mathcal{D}(\mathbf{z}_k \mathbf{z}_k^\top))^2 - 2\mathcal{D}(\mathbf{C})\mathcal{D}(\mathbf{z}_k \mathbf{z}_k^\top)] \\ &= \mathbb{E}[(\mathcal{D}(\mathbf{z}_k \mathbf{z}_k^\top))^2] - \mathcal{D}(\mathbf{C})^2 \\ &\preceq \beta^2 \mathcal{D}(\mathbf{C}) - \mathcal{D}(\mathbf{C})^2, \quad 1 \leq k \leq N. \end{aligned}$$

Substitute the above into (6.2)

$$\mathbb{V}\text{ar}[\mathbf{Z}] \preceq \mathbf{V} \equiv \frac{1}{N} (\beta^2 \mathcal{D}(\mathbf{C}) - \mathcal{D}(\mathbf{C})^2),$$

and apply Theorem 2.3 with $\nu = \|\mathbf{V}\|_2 = S_1/N$ and $d = \text{intdim}(\mathbf{V})$. \square

Below is the minimal sampling amount that makes the Monte Carlo estimator in (6.1) a normwise (ϵ, δ) diagonal estimator.

COROLLARY 6.2. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ have square integrable partial derivatives with respect to the probability density function $\rho_{\mathbf{x}}(\mathbf{x})$, $\|\nabla f\|_\infty \leq \beta$ almost surely, $\hat{\mathbf{C}}$ be the Monte Carlo estimator in (6.1), and*

$$\begin{aligned} c_{\max} &\equiv \|\mathcal{D}(\mathbf{C})\|_2, & S_1 &\equiv \|\mathcal{D}(\mathbf{C}) (\beta^2 \mathbf{I} - \mathcal{D}(\mathbf{C}))\|_2 \\ S_2 &\equiv c_{\max} + \beta^2, & d &\equiv \frac{\sum_{i=1}^n c_{ii}}{S_1}. \end{aligned}$$

Pick $\epsilon > 0$. If $c_{\max} > 0$ and $S_1 > 0$, then for any $0 < \delta < 1$, if the sampling amount is at least

$$N \geq \frac{S_2}{3\epsilon^2} \left(2\epsilon + \frac{6S_1}{c_{\max} S_2} \right) \ln(8d/\delta),$$

then $\|\mathcal{D}(\mathbf{C}) - \mathcal{D}(\hat{\mathbf{C}})\|_2 \leq \epsilon \|\mathcal{D}(\mathbf{C})\|_2$ holds with probability at least $1 - \delta$.

Proof. The proof is similar to that of Corollary 3.3 but is based instead on Theorem 6.1. \square

6.1. Illustrative examples. We determine the constants in Corollary 6.2 for two different functions $f : \mathbb{R}^n \rightarrow \mathbb{R}$, and random variables $\mathbf{x} \in \mathbb{R}^n$ from a uniform distribution over $\mathcal{X} = [-1, 1]^n$.

Linear Function. Let $f(\mathbf{x}) = \mathbf{h}^\top \mathbf{x}$ with $\mathbf{h} \in \mathbb{R}^n$. Then $\|\nabla f\|_\infty \leq \beta \equiv \|\mathbf{h}\|_\infty$ almost surely. The second moment matrix $\mathbf{C} = \mathbf{h}\mathbf{h}^\top$ has a largest diagonal entry $c_{\max} = \|\mathbf{h}\|_\infty^2$. Let $\mathbf{v} \in \mathbb{R}^n$ have entries $v_i = h_i^2 (\|\mathbf{h}\|_\infty^2 - h_i^2)$ for $1 \leq i \leq n$. The constants in Corollary 6.2 are

$$S_1 = \|\mathbf{v}\|_\infty, \quad S_2 = 2\|\mathbf{h}\|_\infty^2, \quad d = \frac{\sum_{i=1}^n v_i}{S_1}.$$

Quadratic function. Let $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{S}\mathbf{x}$ where $\mathbf{S} \in \mathbb{R}^{n \times n}$ $\mathbf{S} = \mathbf{S}^\top$ is a symmetric square root of the positive semidefinite matrix $\mathbf{M} = \mathbf{S}^2$. Then $\|\nabla f\|_\infty \leq \beta \equiv \|\mathbf{S}\|_\infty$ almost surely. The second moment matrix $\mathbf{C} = \frac{1}{3}\mathbf{M}$ has a largest diagonal element $c_{\max} = \frac{1}{3}\|\mathbf{D}(\mathbf{M})\|_\infty$. Let $\mathbf{v} \in \mathbb{R}^n$ have entries $v_i = \frac{1}{3}m_{ii} (\|\mathbf{S}\|_\infty^2 - \frac{1}{3}m_{ii})$ for $1 \leq i \leq n$. The constants in Corollary 6.2 are

$$S_1 = \|\mathbf{v}\|_\infty, \quad S_2 = \frac{1}{3}\|\mathbf{D}(\mathbf{M})\|_\infty + \|\mathbf{S}\|_\infty^2, \quad d = \frac{\sum_{i=1}^n v_i}{S_1}.$$

7. Numerical Experiments. After describing our test matrices (section 7.1), we present four different types of numerical experiments to illustrate the accuracy of the Monte Carlo estimators: Rademacher Monte Carlo estimators applied to the test matrices (section 7.2), accuracy of different Monte Carlo estimators (section 7.3), effect of the sparsity on the accuracy of Rademacher Monte Carlo estimators (section 7.4), and accuracy of the DGSM Monte Carlo estimator (section 7.5).

7.1. Test Matrices. We perform numerical experiments on three symmetric test matrices from [16] of dimension $n = 100$ that depend on a parameter θ .

1. Identity plus rank-1

$$\mathbf{A} = \mathbf{I} + \theta \mathbf{e}\mathbf{e}^\top \quad \text{where } .01 \leq \theta \leq 0.1,$$

where $\mathbf{e} \in \mathbb{R}^n$ is a vector of ones. The constants in Corollary 3.3 are

$$K_1 = (n-1)\theta^2, \quad K_2 = (n-1)\theta, \quad \|\mathbf{D}(\mathbf{A})\|_\infty = 1 + \theta,$$

so that

$$\Delta_1 = \frac{K_1}{(1+\theta)^2}, \quad \Delta_2 = \frac{(n-1)\theta}{1+\theta}, \quad d = n.$$

2. Rank-1 with decaying elements

$$\mathbf{A} = \frac{\mathbf{x}\mathbf{x}^\top}{\|\mathbf{x}\|_2^2} \quad \text{where } \mathbf{x}_j = e^{-j(1-\theta)}, \quad 1 \leq j \leq n, \quad 0.1 \leq \theta \leq 1.$$

The constants in Corollary 3.3 are

$$K_1 = \left(\frac{x_1}{\|\mathbf{x}\|_2} \right)^2 \left(1 - \left(\frac{x_1}{\|\mathbf{x}\|_2} \right)^2 \right), \quad K_2 = \frac{x_1}{\|\mathbf{x}\|_2^2} \sum_{j>1} x_j$$

and $\|\mathbf{D}(\mathbf{A})\|_\infty = \left(\frac{x_1}{\|\mathbf{x}\|_2} \right)^2$ so that

$$\Delta_1 = \left(\frac{\|\mathbf{x}\|_2}{x_1} \right)^2 - 1, \quad \Delta_2 = \frac{\sum_{j>1} x_j}{x_1}, \quad d = \frac{\sum_{i=1}^n x_i^2 (\|\mathbf{x}\|_2^2 - x_i^2)}{x_1^2 (\|\mathbf{x}\|_2^2 - x_1^2)}.$$

3. Tridiagonal Toeplitz matrix

$$\mathbf{A} = \begin{bmatrix} 1 & \theta & & \\ \theta & 1 & \ddots & \\ & \ddots & \ddots & \theta \\ & & \theta & 1 \end{bmatrix} \quad \text{where } 0.1 \leq \theta \leq 1.$$

The constants in Corollary 3.3 are

$$K_1 = 2\theta^2, \quad K_2 = 2\theta, \quad \|\mathcal{D}(\mathbf{A})\|_\infty = 1$$

so that

$$\Delta_1 = 2\theta^2, \quad \Delta_2 = 2\theta, \quad d = \frac{2(n-1)\theta^2}{2\theta^2} = (n-1).$$

For all the test matrices, the constants Δ_1 and Δ_2 increase with increasing θ as the offdiagonal elements become larger in magnitude relative to the diagonal elements. Therefore, we expect the Rademacher Monte Carlo estimators to lose accuracy with increasing θ , as measured by the normwise relative error (NRE) in the computed diagonal $\mathcal{D}(\hat{\mathbf{A}})$,

$$\text{NRE} \equiv \frac{\|\mathcal{D}(\mathbf{A}) - \mathcal{D}(\hat{\mathbf{A}})\|_2}{\|\mathcal{D}(\mathbf{A})\|_2},$$

in Figures 7.1-7.5.

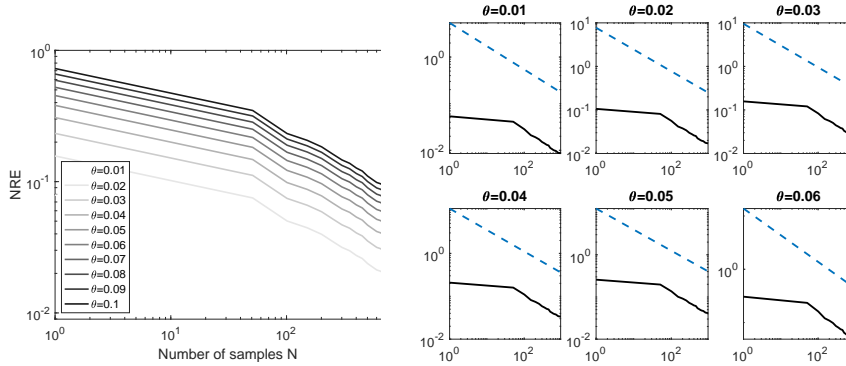


FIG. 7.1. Rademacher Monte Carlo estimator applied to Test Matrix 1. Big left panel: NRE for different values of θ versus sampling amount N . Small panels on the right: NRE (solid black line), and bound (7.1) (blue dotted line) versus sampling amount N with failure probability $\delta = 10^{-16}$.

7.2. Experiment 1: Accuracy of Rademacher Monte Carlo estimator on test matrices. Figures 7.1-7.3 show the NRE of the Rademacher Monte Carlo estimator applied to the test matrices in section 7.1, and the bounds from the normwise (ϵ, δ) estimators in Corollary 3.3.

The big left panel displays the NRE versus the sampling amount N . This NRE represents the average of the NREs over 10 different independent runs. The small panels on the right show the bound ϵ for the normwise (ϵ, δ) estimators from Corollary 3.3 with failure probability $\delta = 10^{-16}$.

For Corollary 3.3, we solve for ϵ from the simpler bound

$$N \geq \frac{\Delta_2}{3\epsilon^2} (2 + 6\Delta_3) \ln(8d/\delta), \quad \Delta_3 \equiv \frac{\Delta_1}{\Delta_2},$$

to obtain

$$(7.1) \quad \epsilon = \sqrt{\frac{\Delta_2}{3N} (2 + 6\Delta_3) \ln(8d/\delta)}, \quad \Delta_2 = \frac{\|\mathbf{A} - \mathcal{D}(\mathbf{A})\|_\infty}{\|\mathcal{D}(\mathbf{A})\|_\infty}, \quad \Delta_3 = \frac{K_1}{\|\mathbf{A} - \mathcal{D}(\mathbf{A})\|_\infty}.$$

The big left panels illustrate that, for a fixed sampling amount N , the NRE for Test Matrices 1 and 3 increases with θ . This is because the offdiagonals become more dominant as θ becomes larger.

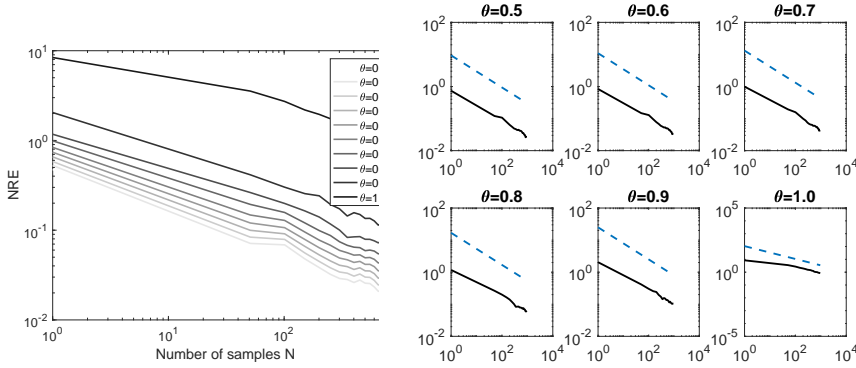


FIG. 7.2. *Rademacher Monte Carlo estimator applied to Test Matrix 2. Big left panel: NRE for different values of θ versus sampling amount N . Small panels on the right: NRE (solid black line) and bound (7.1) (blue dotted line) versus sampling amount N with failure probability $\delta = 10^{-16}$.*

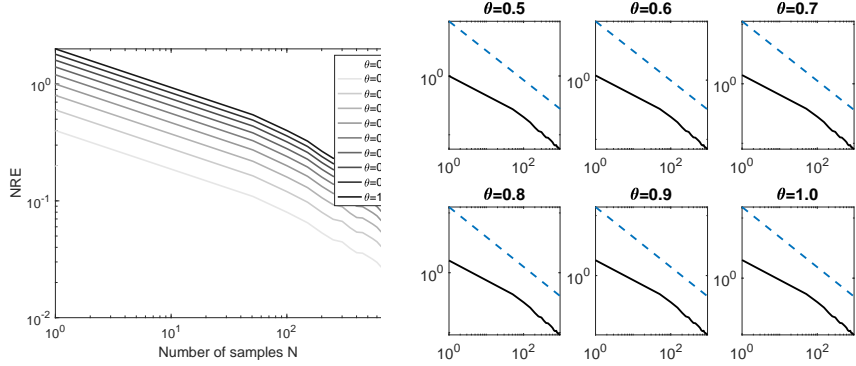


FIG. 7.3. *Rademacher Monte Carlo estimator applied to Test Matrix 3. Big left panel: NRE for different values of θ versus sampling amount N . Small panels on the right: NRE (solid black line) and bound (7.1) (blue dotted line) versus sampling amount N with failure probability $\delta = 10^{-16}$.*

7.3. Experiment 2: Different Monte Carlo estimators. We compare the accuracy of the following Monte Carlo estimators on Test Matrix 1 with $\theta = 0.01$: Rademacher, Gaussian, sparse Rademacher with $s = 3$, and normalized Gaussian.

For each estimator, Figure 7.4 shows the mean of the NRE and variance over 100 runs, with the shaded regions representing the 2.5% and 97.5% quantiles.

The normalized Gaussian estimator is about as accurate as the Rademacher estimator, while the sparse Rademacher with $s = 3$ is about as accurate as the Gaussian estimator. The Gaussian and sparse Rademacher estimators are less accurate than the Rademacher and normalized Gaussian estimators. The shaded regions illustrate that, as expected, the sample variance of all estimators decreases with increasing sampling amount N .

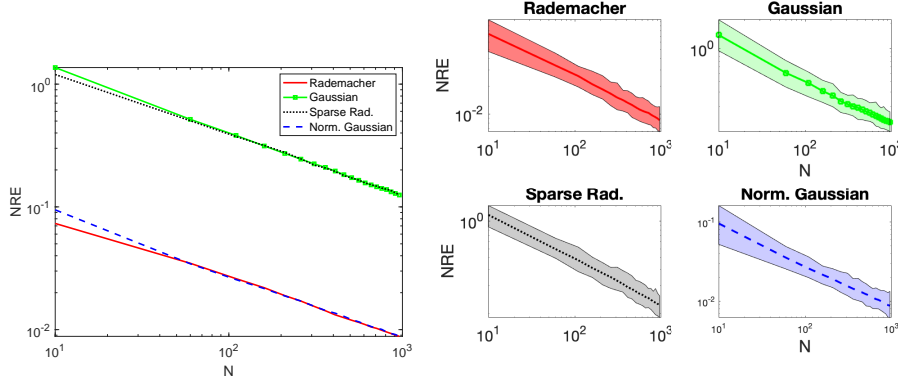


FIG. 7.4. Rademacher, Gaussian, sparse Rademacher with $s = 3$, and Normalized Gaussian Monte Carlo estimators applied to Test Matrix 1 with $\theta = 0.01$. Big left panel: NRE mean versus sampling amount N for different estimators. Small right panels: NRE mean (dotted lines), and 2.5% and 97.5% quantiles (shaded regions) versus sampling amount N .

7.4. Experiment 3: Effect of sparsity in Rademacher vectors. We apply the Rademacher Monte Carlo estimator to Test Matrix 1 with $\theta = .01$ with four different sparsity levels: $s = 1$ (standard Rademacher), $s = 3$ [1], $s = 10$, and $s = 50$.

For each sampling amount N , Figure 7.5 shows the mean and the variance of the NRE over 100 runs. It suggests that sparse Rademacher estimators ($s > 1$) may not be able to achieve a single digit of accuracy, unless the sampling amount is so large as to exceed the matrix dimension.

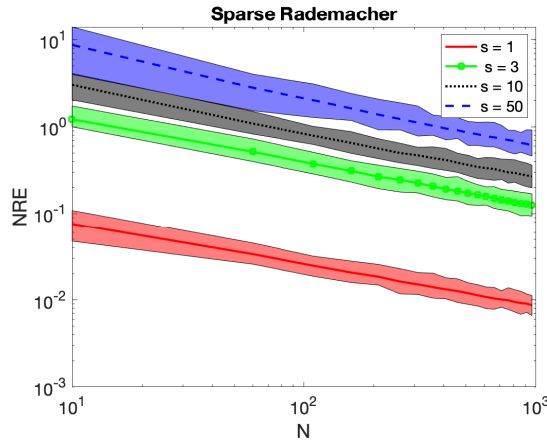


FIG. 7.5. Sparse Rademacher Monte Carlo estimators with sparsity levels $s = 1, 3, 10, 50$ applied to Test Matrix 1 with $\theta = 0.01$. NRE (dotted lines) and 2.5% and 97.5% quantiles (shaded regions).

7.5. Example 4: Bounds for DGSM Monte Carlo estimator. We apply the DGSM Monte Carlo estimator (6.1) to the diagonal matrix

$$(7.2) \quad \mathbf{S} \equiv \text{diag}(\mathbf{s}) \in \mathbb{R}^{n \times n} \quad s_j \equiv \exp(-10j/n), \quad 1 \leq j \leq n$$

from the quadratic function in Section 6.1 for $n = 100$, and illustrate the accuracy of Corollary 6.2.

The left panel of Figure 7.6 shows the normwise relative error

$$\text{NRE} \equiv \frac{\|\mathcal{D}(\mathbf{C}) - \mathcal{D}(\hat{\mathbf{C}})\|_2}{\|\mathcal{D}(\mathbf{C})\|_2}$$

which represents the the average of the NREs over 100 independent runs.

For Corollary 6.2, we fix the sample size N and solve for ϵ from the simpler bound

$$N \geq \frac{S_2}{3\epsilon^2} (2 + 6S_3) \ln(8d/\delta), \quad S_3 \equiv \frac{S_1}{c_{\max} S_2},$$

to obtain

$$(7.3) \quad \epsilon = \sqrt{\frac{S_2}{3N} (2 + 6S_3) \ln(8d/\delta)}.$$

The expressions for S_1, S_2, c_{\max} and d for this example have been derived in subsection 6.1.

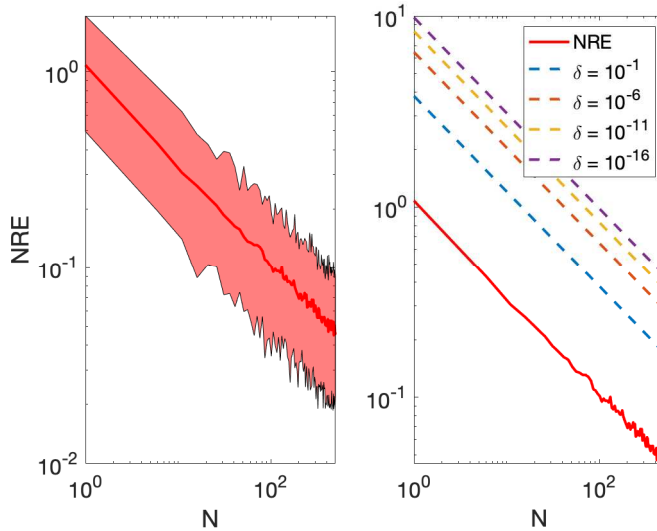


FIG. 7.6. DGSM Monte Carlo estimator (6.1) applied to 100×100 matrix \mathbf{S} in (7.2). Left panel: NRE mean (solid line) and 2.5% and 97.5% quantiles (shaded regions) versus sampling amount N . Right panel: NRE and bounds (7.3) for different failure probabilities δ versus sampling amount.

The right panel of Figure 7.6 illustrates that with less stringent failure probabilities δ , the relative bounds (7.3) move closer to the NRE.

7.6. Example 5: DGSM on the Circuit model. We apply the Monte Carlo DGSM estimator (6.1) to the so-called *circuit model* from [6]. The quantity of interest being modeled is the midpoint voltage of a transformerless push-pull circuit, which depends on $n = 6$ parameters through a nonlinear closed-form algebraic expression.

As in [6], we normalize the parameter space to $\mathcal{X} = [-1, 1]^n$ and scale the partial derivatives appropriately¹.

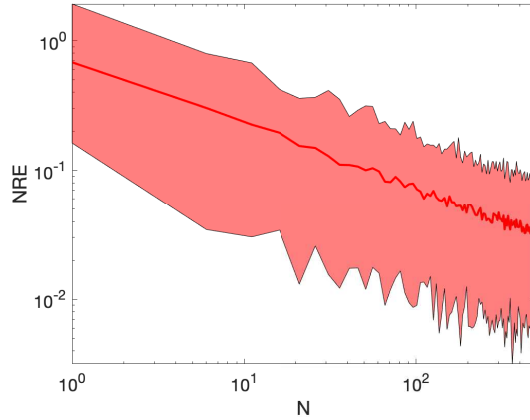


FIG. 7.7. DGSM Monte Carlo estimator (6.1) applied to the Circuit model. NRE mean (solid line) and 2.5% and 97.5% quantiles (shaded regions) versus sampling amount N .

Figure 7.7 shows the mean of the NRE and variances over 100 independent runs per sampling amount N , with the shaded regions representing 2.5% and 97.5% quantiles. Since the exact expressions for the DGSMs are unavailable, we use as the exact value a tensor product Gauss-Legendre quadrature-based approximation with 15 points per dimension (i.e., 15^n total points).

8. Conclusion and future work. This paper derives probabilistic bounds for the Monte Carlo diagonal estimators; the bounds come in two flavors, normwise and componentwise bounds for the absolute and relative errors of the Monte Carlo estimators. There are several avenues for future work. First, it would be interesting to study the accuracy of the diagonal estimator for matrix functions when a polynomial, or rational approximation to the matrix function is used. Second, we are currently pursuing the extension of the analysis of the Monte Carlo diagonal estimators to estimators for the selected entries (possibly including offdiagonal entries) of a matrix.

9. Acknowledgements. The authors would like to acknowledge support from the National Science Foundation through the grants DMS-1745654 (all three authors) and DMS-1845406 (A.K.S.). We would like to acknowledge Alen Alexanderian for helpful discussions.

REFERENCES

- [1] D. Achlioptas. Database-friendly random projections: Johnson-Lindenstrauss with binary coins. *J. Comput. System Sci.*, 66(4):671–687, 2003.

¹MATLAB codes are available in <https://bitbucket.org/paulcon/global-sensitivity-metrics-from-active-subspaces/src/master/>.

- [2] H. Avron and S. Toledo. Randomized algorithms for estimating the trace of an implicit symmetric positive semi-definite matrix. *J. ACM*, 58(2):1–34, 2011.
- [3] R. A. Baston and Y. Nakatsukasa. Stochastic diagonal estimation: probabilistic bounds and an improved algorithm, 2022. [arXiv:2201.10684](#).
- [4] C. Bekas, E. Kokiopoulou, and Y. Saad. An estimator for the diagonal of a matrix. *Appl. Numer. Math.*, 57(11-12):1214–1229, 2007.
- [5] R. Y. Chen, A. Gittens, and J. A. Tropp. The masked sample covariance estimator: an analysis using matrix concentration inequalities. *Inf. Inference*, 1(1):2–20, 2012.
- [6] P. G. Constantine and P. Diaz. Global sensitivity metrics from active subspaces. *Reliab. Eng. Syst. Safe.*, 162:1–13, 2017.
- [7] A. Cortinovis and D. Kressner. On randomized trace estimates for indefinite matrices with an application to determinants. *Found. Comput. Math.*, pages 1–29, 2021.
- [8] M. F. Hutchinson. A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. *Comm. Statist. Simulation Comput.*, 18(3):1059–1076, 1989.
- [9] B. J. Kaperick. Diagonal estimation with probing methods. Master’s thesis, Virginia Polytechnic Institute and State University, 2019.
- [10] S. Kucherenko and B. Iooss. *Derivative-Based Global Sensitivity Measures*, pages 1241–1263. Springer International Publishing, Cham, 2017.
- [11] J. H. Laeuchli. *Methods for Estimating The Diagonal of Matrix Functions*. PhD thesis, College of William and Mary, 2016.
- [12] R. R. Lam, O. Zahm, Y. M. Marzouk, and K. E. Willcox. Multifidelity dimension reduction via active subspaces. *SIAM J. Sci. Comput.*, 42(2):A929–A956, 2020.
- [13] J. L. Larsen and L. M. Morris. *An Introduction to Mathematical Statistics and its Applications*. Pearson Prentice Hall, fourth edition, 2006.
- [14] P. Li, T. J. Hastie, and K. W. Church. Very sparse random projections. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD ’06, page 287–296, New York, NY, USA, 2006. Association for Computing Machinery.
- [15] M. Mitzenmacher and E. Upfal. *Probability and computing*. Cambridge University Press, Cambridge, 2005. Randomized algorithms and probabilistic analysis.
- [16] F. Roosta-Khorasani and U. Ascher. Improved bounds on sample size for implicit matrix trace estimators. *Found. Comput. Math.*, 15(5):1187–1212, 2015.
- [17] A. P. Soms. An asymptotic expansion for the tail area of the t-distribution. *J. Amer. Statist. Assoc.*, 71(355):728–730, 1976.
- [18] J. A. Tropp. An introduction to matrix concentration inequalities. *Found. Trends Mach. Learning*, 8(1–2):1–230, 2015.
- [19] S. Ubaru and Y. Saad. Applications of trace estimation techniques. In T. Kozubek, M. Čermák, P. Tichý, R. Blaheta, J. Šístek, D. Lukáš, and J. Jaroš, editors, *High Performance Computing in Science and Engineering*, pages 19–33, Cham, 2018. Springer International Publishing.
- [20] R. Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge University Press, 2018.
- [21] A. J. Wathen. Preconditioning. *Acta Numer.*, 24:329–376, 2015.
- [22] Z. Yao, A. Gholami, S. Shen, M. Mustafa, K. Keutzer, and M. W. Mahoney. ADAHESSIAN: An adaptive second order optimizer for machine learning. *arXiv preprint arXiv:2006.00719*, 2020.