Estimating an Eigenvector by the Power Method with a Random Start*

Gianna M. Del Corso †

CUCS-037-95

Abstract

This paper addresses the problem of approximating an eigenvector belonging to the largest eigenvalue of a symmetric positive definite matrix by the power method. We assume that the starting vector is randomly chosen with uniform distribution over the unit sphere.

This paper provides lower and upper as well as asymptotic bounds on the randomized error in the $\mathcal{L}_p$ sense, $p \in [1, +\infty]$. We prove that it is impossible to achieve bounds that are independent of the ratio between the two largest eigenvalues. This should be contrasted to the problem of approximating the largest eigenvalue for which Kuczyński and Woźniakowski in 1992 proved that it is possible to bound the randomized error at the k-th step with a quantity that depends only on k and on the size of the matrix.

We prove that the rate of convergence depends on the ratio of the two largest eigenvalues, on their multiplicities, and on the particular norm. The rate of convergence is at most linear in the ratio of the two largest eigenvalues.

Key words. eigenvectors, power method, random start, randomized error.

1 Introduction

In this paper we deal with the power method that is used to approximate a largest eigenvector of an $n \times n$ symmetric matrix A. By the largest eigenvector we mean a normalized eigenvector corresponding to the largest eigenvalue A. Our analysis holds for every matrix A for which the power method is convergent. To simplify notation, we assume that A is positive definite.

It is well known that the convergence of the power method depends on the starting vector $\mathbf{b}$. In particular, the power method is not convergent if $\mathbf{b}$ is orthogonal to the eigenspace corresponding to the largest eigenvalue of A. Since no a priori information about this eigenspace is in general available, a random starting vector is usually chosen. This indicates the need of studying the convergence of the power method with a random start.

It is easy to see that if $\mathbf{b}$ is randomly chosen according to the uniform distribution then the power method approximates a largest eigenvector and the largest eigenvalue with probability 1. The problem of approximating the largest eigenvalue by the power method with a random start has

*This work has been done while the author was visiting the Computer Science Department at Columbia University.

†This research was (partially) supported by the ESPRIT III Basic Research Programme of the EC under contract No. 9072 (Project GEPPCOM). Address: Dipartimento di Matematica, Università di Milano and IMC-CNR, Via Santa Maria 46, 56100 Pisa, Italy. E-mail: delcorso@iei.pi.cnr.it

been considered in [3], where sharp upper bounds on the randomized relative error at each step are given. An important feature of these bounds is that they are independent of the distribution of the eigenvalues.

The approach of our paper is similar to that of [3]. We analyze the convergence of the power method for approximating a largest eigenvector when the starting vector $\mathbf{b}$ is randomly chosen with uniform distribution over the unit sphere of the n dimensional space.

In order to define the randomized error, we consider the acute angle $\alpha_k = \alpha_k(\mathbf{b})$ between the vector computed by the power method at the k-th step and the eigenspace corresponding to the largest eigenvalue, and we study the expectation of $\sin(\alpha_k(\mathbf{b}))$ over $\mathbf{b}$ in the $\mathcal{L}_p$ sense, $p \in [1, +\infty]$.

We first ask whether it is possible to get bounds on the randomized error that do not depend on the distribution of the eigenvalues. We prove, see Section 3, that for every k and p there are matrices for which the randomized error is very close to 1. This means that there are matrices for which the power method fails after k steps even for a random starting vector. In contrast to the problem of approximating the largest eigenvalue, this shows that the randomized error for the problem of approximating a largest eigenvector must depend on the distribution of eigenvalues. In particular, it must depend on the ratio between the two largest eigenvalues. So, the problem of approximating a largest eigenvector is harder than the problem of approximating the largest eigenvalue and even a random start does not help to obtain distribution-free bounds.

We show that the rate of convergence of the power method depends on the ratio of the two largest eigenvalues, on their multiplicities, and on the particular norm p. Let λ_1 be the largest eigenvalue with multiplicity r, and let λ_{r+1} be the second largest eigenvalue with multiplicity s. Then the randomized error after k steps is proportional to $(\lambda_{r+1}/\lambda_1)^k$ if $p < r$, to $k^{1/p}\,(\lambda_{r+1}/\lambda_1)^k$ if $p = r$, and to $(\lambda_{r+1}/\lambda_1)^{kr/p}$ if $p > r$. The multiplicative constants depend on p, r and s.

This means that the rate decreases with p, increases with the multiplicity r, decreases with the multiplicity s, and it is at most linear in λ_{r+1}/λ_1. For $p = +\infty$, the power method has the randomized error equal to one for all k.

We briefly comment on related work on approximate computation of eigenvectors. The idea of using random starting vectors for the power method can be found in the paper of Shub [6]. Shub applies the power method to the matrix e^{-A}, and approximates an eigenvector of A which is *not* necessarily a largest eigenvector. Although for this problem the power method is globally convergent, the random start is used to improve efficiency. Shub shows, however, that even for $n = 2$ there are matrices for which this problem is very hard. In our paper we apply the power method to the matrix A and we are only interested in approximating a largest eigenvector.

Wright [8] and Kostlan [2] analyzed the problem of approximating a largest eigenvector by the power method in a different setting. They considered the average case setting over a class of matrices, whereas we consider the randomized setting. In particular, they estimate the average time needed for computing a vector whose relative distance from the eigenspace of largest eigenvectors is less than ε. In our paper the matrix is fixed while the starting vector is chosen at random.

The paper is organized as follows. Section 2 contains the definition of the problem and some general results that are used in the subsequent sections. In Section 3 we analyze the behavior of the power method for worst case matrices. In Section 4 we find upper and lower bounds on the randomized error. We show that these bounds are asymptotically optimal since, up to lower order terms, they match the asymptotic bounds presented in Section 5. Numerical tests are presented in Section 6. The tests show that the randomized error indeed depends on the distribution of the eigenvalues. We compare the test results with the theoretical lower and upper bounds. Section 7 contains the conclusions and final remarks.

2 Definition of the Problem

Let A be an $n \times n$ symmetric positive definite matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n > 0$ and corresponding orthonormal eigenvectors $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n$. We will denote by $\mathcal{Z}$ the eigenspace corresponding to λ_1.
We recall that the power method is defined as follows, see e.g. [5]. Let $\mathbf{u}_0 = \mathbf{b}$ be any nonzero starting vector. Then, for every $k = 1, 2, \ldots$, we construct the following sequences of vectors

$$\begin{cases} \mathbf{y}_k & = A\mathbf{u}_{k-1}; \\ \mathbf{u}_k & = \mathbf{y}_k / ||\mathbf{y}_k||, \end{cases}$$

where $||\cdot||$ is the Euclidean vector norm.

Without loss of generality, we may assume that the starting vector $\mathbf{b}$ is normalized, so that $||\mathbf{b}|| = 1$. Observe that if we express $\mathbf{b}$ as a linear combination of the orthonormal eigenvectors,

$$\mathbf{b} = \sum_{i=1}^{n} b_i \mathbf{z}_i,$$

then $\mathbf{u}_k$ becomes

$$\mathbf{u}_k = \frac{\sum_{i=1}^{n} b_i \lambda_i^k \mathbf{z}_i}{\sqrt{\sum_{i=1}^{n} b_i^2 \lambda_i^{2k}}}. \tag{1}$$

Let r be the multiplicity of the largest eigenvalue λ_1. Without loss of generality, we assume that $1 \leq r < n$, since $r = n$ implies $A = \lambda_1 I$, and in this case any nonzero vector is an eigenvector corresponding to λ_1.

In order to estimate the error at the k-th step, we consider the acute angle $\alpha_k(\mathbf{b})$ between the vector $\mathbf{u}_k$ and the eigenspace $\mathcal{Z}$. This angle is uniquely determined by the vector $\mathbf{u}_k$ and by its orthogonal projection on the subspace $\mathcal{Z}$. The sine of $\alpha_k(\mathbf{b})$ is the distance between the vector $\mathbf{u}_k$ and the subspace $\mathcal{Z}$. ¿From (1) we have

$$\operatorname{dist}(\mathbf{u}_k, \mathcal{Z}) := \inf_{z \in \mathcal{Z}} \|\mathbf{u}_k - z\| = \sin(\alpha_k(\mathbf{b})) = \sqrt{\frac{\sum_{i=r+1}^{n} b_i^2 \lambda_i^{2k}}{\sum_{i=1}^{r} b_i^2 \lambda_1^{2k} + \sum_{i=r+1}^{n} b_i^2 \lambda_i^{2k}}}. \tag{2}$$

It is straightforward to see that, if the vector $\mathbf{b}$ has zero components in the directions of the eigenvectors belonging to λ_1, i.e., $b_i = 0$ for $i = 1, 2 \ldots, r$, then $\alpha_k = \pi/2$ for any k. Otherwise, $\mathbf{u}_k$ converges to a vector of $\mathcal{Z}$ and the angle α_k goes to zero as k goes to infinity. The analysis of the power method for a fixed starting vector $\mathbf{b}$ may be found in many books, see for example [5] and [7], where in particular one finds that, if the method converges, the rate convergence is λ_{r+1}/λ_1.

As already mentioned, we study the randomized error of $\sin(\alpha_k(\cdot))$ in the $\mathcal{L}_p$ sense. Using (2) we have

$$\sin(\alpha_k(\mathbf{b})) = \sqrt{\frac{\sum_{i=r+1}^{n} b_i^2 x_i^{2k}}{\sum_{i=1}^{r} b_i^2 + \sum_{i=r+1}^{n} b_i^2 x_i^{2k}}}, \tag{3}$$

where

$$x_i = \lambda_i / \lambda_1 \quad \text{for } i = 1, 2, \ldots, n, \quad \text{and } 1 = x_1 = \cdots = x_r > x_{r+1} \cdots \geq x_n > 0. \tag{4}$$

Let us formalize the notion of $\mathcal{L}_p$ norm. Let μ be the uniform distribution over the unit sphere $S_n = \{\mathbf{b} : ||\mathbf{b}|| = 1\}$ such that $\mu(S_n) = 1$. Then the $\mathcal{L}_p$ norm of the function $\sin(\alpha_k(\cdot))$, defined as in (3), is given by

$$||\sin(\alpha_k(\cdot))||_p = \left[\int_{S_n} |\sin(\alpha_k(b))|^p \ \mu(db)\right]^{1/p}. \tag{5}$$

¿From Remark 7.2 of [3], we have

$$\int_{S_n} |\sin(\alpha_k(b))|^p \ \mu(db) = \frac{1}{c_n}\int_{B_n} |\sin(\alpha_k(b))|^p \ db, \tag{6}$$

where c_n is the Lebesgue's measure of the unit ball $B_n = \{\mathbf{b} : ||\mathbf{b}|| \leq 1\}$, see (10) for the definition of c_n.
Substituting (3) into (5), and using (6), we have

$$||\sin(\alpha_k(\cdot))||_p = \left[\frac{1}{c_n}\int_{B_n}\left(\frac{\sum_{i=r+1}^n b_i^2 x_i^{2k}}{\sum_{i=1}^r b_i^2 + \sum_{i=r+1}^n b_i^2 x_i^{2k}}\right)^{p/2} db\right]^{1/p}.$$

In the same way we define the norm of the space $\mathcal{L}_\infty$ to be

$$\begin{aligned} ||\sin(\alpha_k(\cdot))||_\infty &= \sup_{\mathbf{b}\in S_n} |\sin(\alpha_k(b))| \\ &= \sup_{||\mathbf{b}||=1}\sqrt{\frac{\sum_{i=r+1}^n b_i^2 x_i^{2k}}{\sum_{i=1}^r b_i^2 + \sum_{i=r+1}^n b_i^2 x_i^{2k}}}. \end{aligned} \tag{7}$$

It is easy to see that the supremum in (7) is achieved by setting $\sum_{i=1}^r b_i^2 = 0$. From (7), we get

$$||\sin(\alpha_k)||_\infty = 1. \tag{8}$$

In the following we refer to $\sin(\alpha_k(\mathbf{b}))$ as the *error* of the power method after k steps for the starting vector $\mathbf{b}$. We denote $||\sin(\alpha_k)||_p$ by $e_k^{\mathrm{ran}}(A,p)$, and we call it the *randomized error in the $\mathcal{L}_p$ sense* of the power algorithm after k steps. Hence, we have

$$e_k^{\mathrm{ran}}(A,p) = \left[\frac{1}{c_n}\int_{B_n}\left(\frac{\sum_{i=r+1}^n b_i^2 x_i^{2k}}{\sum_{i=1}^r b_i^2 + \sum_{i=r+1}^n b_i^2 x_i^{2k}}\right)^{p/2} db\right]^{1/p}. \tag{9}$$

For $p = +\infty$, the power method fails to converge since its randomized error is one for all k, see (8). ¿From now on we therefore assume that $p < +\infty$. As we shall see, the power method is then convergent, $e_k^{\mathrm{ran}}(A,p) \to 0$. The speed of convergence is however poor for large p.

In the paper we will denote by c_i the measure of the unit ball over $\mathbb{R}^i$. We have

$$c_i = \frac{\pi^{i/2}}{\Gamma(i/2+1)}, \tag{10}$$

see [1] for the definition of the gamma function $\Gamma(x)$.
We will also use the following relation between the beta and gamma functions

$$B(i,j) = 2\int_0^1 t^{2i-1}(1-t^2)^{j-1}\,dt = \frac{\Gamma(i)\,\Gamma(j)}{\Gamma(i+j)}. \tag{11}$$

We will denote by $F(a,b;c;x)$ the hypergeometric function, see [1] for the definition and the properties of this function.

3 Worst Case Matrices

In [3], Kuczyński and Woźniakowski considered the power method for approximating the largest eigenvalue λ_1. They proved that the randomized error after k steps is bounded by a quantity that goes to zero as $\ln(n)/k$ independently on the distribution of the eigenvalues. This means that, for every matrix, it is possible to give an estimate of the number of steps that guarantees the randomized error to be less than a positive value ε.

Our first goal is to analyze the possibility of obtaining distribution-free bounds for the problem of approximating a largest eigenvector. To this extent, we will deal with "worst case matrices".

Let us denote by $s(k,p)$ the supremum of the randomized error in the $\mathcal{L}_p$ sense over all positive definite matrices A, i.e.,

$$s(k,p) = \sup_{A=A^*>0} e_k^{\text{ran}}(A,p).$$

Since the randomized error increases with x_i, see (4), it is easy to show that the supremum is achieved by setting $x_i = 1$ for every $i \geq 2$ and for every p, $1 \leq p < \infty$. Then we get

$$\begin{aligned} s(k,p) &= \left[\frac{1}{c_n} \int_{B_n} \left(\frac{\sum_{i=2}^n b_i^2}{b_1^2 + \sum_{i=2}^n b_i^2} \right)^{p/2} db \right]^{1/p} \\ &= \left[\frac{1}{c_n} \int_{B_n} \left(1 - \frac{b_1^2}{\sum_{i=1}^n b_i^2} \right)^{p/2} db \right]^{1/p}. \end{aligned} \tag{12}$$

Hence, $s(k,p)$ is independent of k and cannot go to zero. This shows that there are no distribution-free bounds. In fact, $s(k,p)$ are pretty close to 1. We first consider the case $p=2$. Using (12), symmetry arguments yield

$$s(k,2) = \left(1 - \frac{1}{n}\right)^{1/2}. \tag{13}$$

We obtain estimates on $s(k,p)$ by the following proposition.

Proposition 3.1 *For every k and p, $1 \leq p < \infty$, we have*

$$\left(1 - \frac{1}{n}\right)^{1/2} \leq s(k,p) \leq 1.$$

Proof. Note that

$$\left(1 - \frac{b_1^2}{\sum_{i=1}^n b_i^2} \right)^{p/2} \leq 1.$$

From (12) it follows that $s(k,p) \leq 1$ for $1 \leq p < \infty$.

On the other hand, using Hölder's inequality, we get

$$\int_{B_n} \left(1 - \frac{b_1^2}{\sum_{i=1}^n b_i^2} \right) db \leq \left[\int_{B_n} \left(1 - \frac{b_1^2}{\sum_{i=1}^n b_i^2} \right)^{p/2} db \right]^{2/p} \left[\int_{B_n} db \right]^{1/q},$$

where $p/2$ and q are conjugate exponents, i.e., $1/(p/2) + 1/q = 1$.

By rising both terms of the above inequality to the power $p/2$ we obtain

$$\left[\int_{B_n}\left(1-\frac{b_1^2}{\sum_{i=1}^n b_i^2}\right)\,db\right]^{p/2} \le c_n^{p/(2q)}\int_{B_n}\left(1-\frac{b_1^2}{\sum_{i=1}^n b_i^2}\right)^{p/2} db.$$

Since $p/(2q)=p/2-1$, multiplying by $1/c_n^{p/2}$ we get

$$\begin{aligned}
\left[\frac{1}{c_n}\int_{B_n}\left(1-\frac{b_1^2}{\sum_{i=1}^n b_i^2}\right)\,db\right]^{p/2} &\le \frac{1}{c_n}\int_{B_n}\left(1-\frac{b_1^2}{\sum_{i=1}^n b_i^2}\right)^{p/2} db,\\
\left[\frac{1}{c_n}\int_{B_n}\left(1-\frac{b_1^2}{\sum_{i=1}^n b_i^2}\right)\,db\right]^{1/2} &\le \left[\frac{1}{c_n}\int_{B_n}\left(1-\frac{b_1^2}{\sum_{i=1}^n b_i^2}\right)^{p/2} db\right]^{1/p}.
\end{aligned}$$

Due to (13), we thus have

$$\left(1-\frac{1}{n}\right)^{1/2} \le s(k,p).$$

□

Proposition 3.1 states that for every k there are matrices for which the randomized error is close to one. These matrices have the largest eigenvalue of multiplicity one, and the second largest eigenvalue has multiplicity $n-1$ and is pathologically close to λ_1. In this case, even if the starting vector is random, the sequence $\{\mathbf{u}_i\}$ for $i=1,2\ldots,k$ does not approximate a largest eigenvector.

4 Non Asymptotic Behavior

So far we have seen that if $\lambda_{r+1}/\lambda_1 \approx 1$ then the power method behaves badly even for a random starting vector. We now analyze the relationship between the ratio λ_{r+1}/λ_1 and the rate of convergence of the power method for approximating a largest eigenvector.

We first show upper and lower bounds on the randomized error $e_k^{\mathrm{ran}}(A,p)$. These bounds depend on the distribution of the eigenvalues of the matrix A and on the particular norm used. In particular, we prove that the rate of convergence is slower when the multiplicity of λ_1 is smaller than the value of the norm. What seems interesting about these results is that they hold for a class of norms, and we are able to show how the norm affects the speed of convergence of the power method.

4.1 Upper Bounds

We now show how the rate of convergence depends on the multiplicity r of the largest eigenvalue and on the value p of the norm.

We have three cases, and we notice that the rate of convergence of the method is lower when the multiplicity r of the largest eigenvalue is small compared to p. In Section 4.3 we explain why the rate of convergence decreases for $r<p$.

Theorem 4.1 *Let A be a symmetric positive definite matrix, and let r, $r<n$, denote the multiplicity of the largest eigenvalue λ_1 of A. Let*

$$\beta = \left[\frac{\Gamma(n/2)}{\Gamma(p/2)\,\Gamma((n-p)/2)}\left(2+\frac{2}{n}\right)\right]^{1/p}.$$

Then, for every p, $1 \le p < \infty$, and for every k we have

$$e_k^{\rm ran}(A,p) \le \begin{cases} x_{r+1}^k \left(\dfrac{\Gamma((r-p)/2)\,\Gamma((n+p-r)/2)}{\Gamma(r/2)\,\Gamma((n-r)/2)}\right)^{1/p} & \text{for } p<r, \\ x_{r+1}^k (2k)^{1/p} \left[\ln\left(\dfrac{1}{x_{r+1}}\right)\right]^{1/p} \left(\dfrac{\Gamma(n/2)}{\Gamma(p/2)\,\Gamma((n-p)/2)}\right)^{1/p} + \beta\, x_{r+1}^k & \text{for } p=r, \\ x_{r+1}^{kr/p} \left(\dfrac{\Gamma((p-r)/2)\,\Gamma(n/2)}{\Gamma(p/2)\,\Gamma((n-r)/2)}\right)^{1/p} & \text{for } p>r. \end{cases}$$

Proof. We have

$$[e_k^{\rm ran}(A,p)]^p = \frac{1}{c_n}\int_{B_n} \left(\frac{\sum_{i=r+1}^n b_i^2 x_i^{2k}}{\sum_{i=1}^r b_i^2 + \sum_{i=r+1}^n b_i^2 x_i^{2k}}\right)^{p/2} db.$$

Observe that the integrand is an increasing function of $\sum_{i=r+1}^n b_i^2 x_i^{2k}$. The upper bound is then obtained by replacing x_i by x_{r+1} for $i > r+1$,

$$[e_k^{\rm ran}(A,p)]^p \le \frac{x_{r+1}^{kp}}{c_n}\int_{B_n} \left(\frac{\sum_{i=r+1}^n b_i^2}{\sum_{i=1}^r b_i^2 + x_{r+1}^{2k}\sum_{i=r+1}^n b_i^2}\right)^{p/2} db. \tag{14}$$

Consider first the case $p < r$.

Let $a = x_{r+1}^k$, $||b||^2 = \sum_{i=1}^r b_i^2$, and let $t_i = b_i/(1-||b||^2)^{1/2}$ for $i = r+1,\dots,n$ with $||t||^2 = \sum_{i=r+1}^n t_i^2$. If we rewrite the last integral as an integral over the balls B_r and B_{n-r}, we get

$$[e_k^{\rm ran}(A,p)]^p \le \frac{a^p}{c_n}\int_{B_r}\int_{B_{n-r}} \frac{||t||^p(1-||b||^2)^{(n+p-r)/2}}{(||b||^2 + a^2||t||^2(1-||b||^2))^{p/2}}\,dt\,db.$$

Let $\gamma = r(n-r)c_r c_{n-r}/c_n$. We apply twice formula [4.642] of [1] to reduce the last integral to the two dimensional integral and we get

$$[e_k^{\rm ran}(A,p)]^p \le a^p\gamma \int_0^1\int_0^1 \frac{t^{n+p-r-1}\, b^{r-1}(1-b^2)^{(n+p-r)/2}}{(b^2+a^2t^2(1-b^2))^{p/2}}\,db\,dt.$$

Since $b^2 + a^2t^2(1-b^2) \ge b^2$, we have

$$\begin{aligned} [e_k^{\rm ran}(A,p)]^p &\le a^p\gamma \int_0^1 t^{n+p-r-1}\,dt \int_0^1 \frac{b^{r-1}(1-b^2)^{(n+p-r-1)/2}}{b^p}\,db \\ &= a^p \frac{\gamma}{n+p-r}\int_0^1 b^{r-p-1}(1-b^2)^{(n+p-r)/2}\,db. \end{aligned} \tag{15}$$

¿From the definition of the beta function and since $p < r$, (15) becomes

$$\begin{aligned} [e_k^{\rm ran}(A,p)]^p &\le a^p \frac{\gamma}{2(n+p-r)} B\left(\frac{r-p}{2}, \frac{n+p-r}{2}+1\right) \\ &= a^p \frac{\Gamma((n+p-r)/2)\,\Gamma((r-p)/2)}{\Gamma((n-r)/2)\,\Gamma(r/2)}. \end{aligned}$$

This proves the case $p < r$.

Let us now consider the case $p = r$. The integral in (14) can be rewritten with respect to the ball B_{n-p} and the p dimensional ball $B'_p = \{b : \sum_{i=1}^p b_i^2 \le 1 - \sum_{i=p+1}^n b_i^2\}$. We have

$$[e_k^{\mathrm{ran}}(A,p)]^p \le \frac{a^p}{c_n} \int_{B_{n-p}} \left(\sum_{i=p+1}^n b_i^2 \right)^{p/2} \int_{B'_p} \frac{1}{\left(\sum_{i=1}^p b_i^2 + a^2 \sum_{i=p+1}^n b_i^2 \right)^{p/2}} \, db.$$

Let $||b||^2 = \sum_{i=p+1}^n b_i^2$. From formula [4.642] of [1], we get

$$[e_k^{\mathrm{ran}}(A,p)]^p \le a^p \frac{p\, c_p}{c_n} \int_{B_{n-p}} ||b||^p \int_0^{\sqrt{1-||b||^2}} \frac{t^{p-1}}{(t^2 + a^2 ||b||^2)^{p/2}} \, dt\, db. \tag{16}$$

We have two cases, $p = r = 1$ and $p = r \ge 2$. If $p = 1$, (16) becomes

$$\begin{aligned} [e_k^{\mathrm{ran}}(A,1)] &\le a \frac{2}{c_n} \int_{B_{n-1}} ||b|| \int_0^{\sqrt{1-||b||^2}} \frac{1}{(t^2 + a^2||b||^2)^{1/2}} \, dt\, db \\ &= a \frac{2}{c_n} \int_{B_{n-1}} ||b|| \ln \left(\frac{\sqrt{1-||b||^2} + \sqrt{1-(1-a^2)||b||^2}}{a\, ||b||} \right) db. \end{aligned}$$

Using [4.642] of [1], and observing that $\sqrt{1-||b||^2} \le \sqrt{1-(1-a^2)||b||^2}$, we get

$$\begin{aligned} [e_k^{\mathrm{ran}}(A,1)] &\le a\, \gamma \int_0^1 b^{n-1} \ln \left(\frac{2\sqrt{1-(1-a^2)b^2}}{ab} \right) \\ &\le a \frac{\gamma}{2n} \ln \left(\frac{1}{a^2} \right) + a \frac{\gamma}{n} + a \frac{\gamma}{n^2}, \end{aligned} \tag{17}$$

where $\gamma = (n-1) 2 c_{n-1}/c_n$.
Hence, from (17) we have

$$[e_k^{\mathrm{ran}}(A,1)] \le a \frac{\Gamma(n/2)}{\Gamma(1/2)\,\Gamma((n-1)/2)} \ln \left(\frac{1}{a^2} \right) + a \frac{\Gamma(n/2)}{\Gamma(1/2)\,\Gamma((n-1)/2)} \left(2 + \frac{2}{n} \right).$$

This proves the case $p = r = 1$.
Let us consider the case $p \ge 2$. Notice that $(t^2 + a^2||b||^2)^{p/2} \ge t^p + (p/2) t^{2(p/2-1)} a^2 ||b||^2 (1 + o(1))$. Then we can bound the denominator of the integrand of (16) with the first two terms of this expansion. We have

$$\begin{aligned} [e_k^{\mathrm{ran}}(A,p)]^p &\le a^p \frac{p\, c_p}{c_n} \int_{B_{n-p}} ||b||^p \int_0^{\sqrt{1-||b||^2}} \frac{t^{p-1}}{t^p + p/2\, t^{p-2} a^2 ||b||^2} \, dt\, db \\ &= a^p \frac{p\, c_p}{c_n} \int_{B_{n-p}} ||b||^p \int_0^{\sqrt{1-||b||^2}} \frac{t}{t^2 + p/2\, a^2 ||b||^2} \, dt\, db. \end{aligned}$$

Solving the last integral, and using again [4.642] of [1] to reduce the first integral to a one-dimensional integral, we obtain

$$[e_k^{\mathrm{ran}}(A,p)]^p \le a^p \gamma \int_0^1 b^{n-1} \frac{1}{2} \ln \left(\frac{1 - (1 - p/2a^2) b^2}{p/2\, a^2 b^2} \right) db$$

$$
\begin{aligned}
&= a^p \frac{\gamma}{2} \int_0^1 b^{n-1} \ln\left(\frac{1}{p/2\, a^2 b^2}\right) db + a^p \frac{\gamma}{2} \int_0^1 b^{n-1} \ln\left(1 - \left(1 - \frac{p}{2} a^2\right) b^2\right) db \\
&= a^p \frac{\gamma}{2n} \ln\left(\frac{2}{pa^2}\right) + a^p \frac{\gamma}{n^2} + a^p \frac{\gamma}{2} \int_0^1 b^{n-1} \ln\left(1 - \left(1 - \frac{p}{2} a^2\right) b^2\right) db, \qquad (18)
\end{aligned}
$$

where $\gamma = p(n-p)c_p c_{n-p}/c_n$.
Let us consider the argument of the logarithm in the integral of (18). Observe that if $a^2 \leq 2/p$, then $\ln(1-(1-p/2\, a^2)b^2) \leq 0$. Hence, in this case, we can bound (18) by

$$
\begin{aligned}
[e_k^{\mathrm{ran}}(A,p)]^p &\leq a^p \frac{\gamma}{2n} \ln\left(\frac{2}{pa^2}\right) + a^p \frac{\gamma}{n^2} \\
&= a^p \frac{\gamma}{2n} \ln\left(\frac{1}{a^2}\right) + a^p \frac{\gamma}{2n} \left(\ln\left(\frac{2}{p}\right) + \frac{2}{n}\right) \\
&\leq a^p \frac{\Gamma(n/2)}{\Gamma(p/2)\Gamma((n-p)/2)} \ln\left(\frac{1}{a^2}\right) + a^p \frac{\Gamma(n/2)}{\Gamma(p/2)\Gamma((n-p)/2)} \frac{2}{n}. \qquad (19)
\end{aligned}
$$

Otherwise, if $a^2 > 2/p$, then $\ln(1-(1-p/2\, a^2)b^2) \leq \ln(p/2)$. In this case we have

$$
\begin{aligned}
[e_k^{\mathrm{ran}}(A,p)]^p &\leq a^p \frac{\gamma}{2n} \ln\left(\frac{2}{pa^2}\right) + a^p \frac{\gamma}{n^2} + a^p \frac{\gamma}{2} \int_0^1 b^{n-1} \ln\left(\frac{p}{2}\right) db \\
&= a^p \frac{\gamma}{2n} \ln\left(\frac{1}{a^2}\right) + a^p \frac{\gamma}{n^2} \\
&= a^p \frac{\Gamma(n/2)}{\Gamma(p/2)\Gamma((n-p)/2)} \ln\left(\frac{1}{a^2}\right) + a^p \frac{\Gamma(n/2)}{\Gamma(p/2)\Gamma((n-p)/2)} \frac{2}{n}. \qquad (20)
\end{aligned}
$$

Observing that $2/n < (2 + 2/n)$ and using (19) and (20), we have

$$
[e_k^{\mathrm{ran}}(A,p)]^p \leq a^p \frac{\Gamma(n/2)}{\Gamma(p/2)\Gamma((n-p)/2)} \ln\left(\frac{1}{a^2}\right) + a^p \frac{\Gamma(n/2)}{\Gamma(p/2)\Gamma((n-p)/2)} \left(2 + \frac{2}{n}\right).
$$

This proves the case $p = r$.
Finally, assume that $p > r$. From (14), repeating the same reasoning that led to (16), we have

$$
\begin{aligned}
[e_k^{\mathrm{ran}}(A,p)] &\leq a^p \frac{r c_r}{c_n} \int_{B_{n-r}} ||b||^p \int_0^{\sqrt{1-||b||^2}} \frac{t^{r-1}}{(t^2 + a^2||b||^2)^{p/2}} \, dt\, db \\
&= \frac{r c_r}{c_n} \int_{B_{n-r}} \int_0^{\sqrt{1-||b||^2}} \frac{t^{r-1}}{(t^2/(a^2||b||^2) + 1)^{p/2}} \, dt\, db.
\end{aligned}
$$

Changing variables by setting $z = t/(a||b||)$, we get

$$
[e_k^{\mathrm{ran}}(A,p)]^p \leq a^r \frac{r c_r}{c_n} \int_{B_{n-r}} ||b||^r \int_0^d \frac{z^{r-1}}{(z^2+1)^{p/2}} \, dz\, db,
$$

where $d = \sqrt{1 - ||b||^2}/(a||b||)$. Set now $y = z^2$. From the last equation we have

$$
[e_k^{\mathrm{ran}}(A,p)]^p \leq a^r \frac{r c_r}{2c_n} \int_{B_{n-r}} ||b||^r \int_0^{d^2} \frac{y^{r/2-1}}{(y+1)^{p/2}} \, dy\, db.
$$

We notice that d goes to infinity when a goes to zero. Then we have

$$\int_0^{d^2} \frac{y^{r/2-1}}{(y+1)^{p/2}}\,dy \le \int_0^{+\infty} \frac{y^{r/2-1}}{(y+1)^{p/2}}\,dy = B\left(\frac{r}{2}, \frac{p-r}{2}\right),$$

due to formula [3.194] of [1].
We apply [4.642] of [1] to reduce the integral over B_{n-r} to a one dimensional integral, and we get

$$[e_k^{\text{ran}}(A,p)]^p \le a^r \frac{r(n-r)c_r c_{n-r}}{c_n} \int_0^1 b^{n-1} B\left(\frac{r}{2}, \frac{p-r}{2}\right) = a^r \frac{\Gamma(n/2)\,\Gamma((p-r)/2)}{\Gamma((n-r)/2)\,\Gamma(p/2)}.$$

This concludes the proof. □

Note that, when $p = r$, the bound is composed of two terms. The first term depends on k through $x_{r+1}^k k^{1/p}$, the second term depends on k through x_{r+1}^k. We remark that for large k the influence of the second term is negligible. Nevertheless, numerical tests show that this term can affect the bound when the value of x_{r+1} is close to 1.

4.2 Lower Bounds

In this section we find lower bounds on the randomized error $e_k^{\text{ran}}(A,p)$. As in Section 4.1, we show that these lower bounds depend on the multiplicity of the largest eigenvalue and on the value of the norm. Upper and lower bounds show the same dependence on the ratio between the two largest eigenvalues and on the relation between p and r.

Below we define some constants that are used in Theorem 4.2.

$$\gamma = \begin{cases} \left(\dfrac{\Gamma((r-p)/2)\,\Gamma(p+1/2)\,\Gamma((r+1)/2)}{2\Gamma(r/2)\,\Gamma(1/2)\,\Gamma((r+p+1)/2)} F\left(\dfrac{r+1}{2}, \dfrac{r-p}{2}; \dfrac{r+p+1}{2}; 1-x_{r+1}^{2k}\right)\right)^{1/p}, & \text{for } p < 2, \\[2ex] \left(\dfrac{p\Gamma((r-p)/2)\,\Gamma((p+3)/2)}{2(r+1)\Gamma(r/2)\,\Gamma(1/2)} F\left(\dfrac{r+1}{2}, \dfrac{r-p}{2}; \dfrac{r+3}{2}; 1-x_{r+1}^{2k}\right)\right)^{1/p}, & \text{for } p \ge 2. \end{cases}$$

$$\gamma' = \left(\frac{\Gamma((p+1)/2)}{\Gamma(p/2)\,\Gamma(1/2)}\left(\log\left(\frac{p^2+3p}{4}\right) - \frac{2}{p+1} + \frac{2p-4}{p+3} F\left(1, \frac{p+1}{2}; \frac{p+5}{2}; 1-x_{r+1}^{2k}\right)\right)\right)^{1/p},$$

$$\gamma'' = \left(\frac{r\,\Gamma((r+1)/2+1)\,\Gamma((p-r)/2)}{4p\,\Gamma(1/2)\,\Gamma(p/2+1)} F\left(\frac{r}{2}+1, 1; \frac{p}{2}+1; 1-x_{r+1}^{2k}\right)\right)^{1/p}.$$

Theorem 4.2 *Let A be a symmetric positive definite matrix, and let r, $r < n$, denote the multiplicity of the largest eigenvalue λ_1 of A. Then, for every p, $1 \le p < \infty$, and for every k we have*

$$e_k^{\text{ran}}(A,p) \ge \begin{cases} x_{r+1}^k \left(\dfrac{\Gamma((r-p)/2)\,\Gamma((p+1)/2)}{\Gamma(r/2)\,\Gamma(1/2)}\right)^{1/p} - \gamma\, x_{r+1}^{kr/p} & \textit{for } p < r, \\[2ex] x_{r+1}^k (2k)^{1/r} \left[\ln\left(\dfrac{1}{x_{r+1}}\right)\right]^{1/r} \left(\dfrac{\Gamma((p+1)/2)}{\Gamma(p/2)\,\Gamma(1/2)}\right)^{1/p} - \gamma' x_{r+1}^k & \textit{for } p = r, \\[2ex] x_{r+1}^{kr/p} \left(\dfrac{\Gamma((p-r)/2)\,\Gamma((r+1)/2)}{\Gamma(p/2)\,\Gamma(1/2)}\right)^{1/p} - \gamma'' x_{r+1}^{k(r+2)/p} & \textit{for } p > r. \end{cases}$$

Proof. We have

$$[e_k^{\mathrm{ran}}(A,p)]^p = \frac{1}{c_n}\int_{B_n}\left(\frac{\sum_{i=r+1}^n b_i^2 x_i^{2k}}{\sum_{i=1}^r b_i^2 + \sum_{i=r+1}^n b_i^2 x_i^{2k}}\right)^{p/2} db.$$

Notice that the integrand is an increasing function of $\sum_{i=r+1}^n b_i^2 x_i^{2k}$. Hence, the lower bound is obtained by replacing x_i by 0 for $i > r+1$,

$$[e_k^{\mathrm{ran}}(A,p)]^p \geq \frac{x_{r+1}^{kp}}{c_n}\int_{B_n}\frac{b_{r+1}^p}{(\sum_{i=1}^r b_i^2 + x_{r+1}^{2k} b_{r+1}^2)^{p/2}}\,db.$$

Let $a = x_{r+1}^k$. Writing the last integral as an integral over the ball B_{n-r} and the r dimensional ball of radius $q = \sqrt{1-\sum_{i=r+1}^n b_i^2}$, and applying [4.642] of [1], we get

$$[e_k^{\mathrm{ran}}(A,p)]^p \geq \frac{r\,c_r}{c_n}\int_{B_{n-r}}\int_0^q t^{r-1}\left(\frac{a^2\,b_{r+1}^2}{t^2 + a^2\,b_{r+1}^2}\right)^{p/2} dt\,db. \tag{21}$$

Let us denote $a^2\,b_{r+1}^2$ by α, and consider the integral

$$f(\alpha) = \int_0^q t^{r-1}\left(\frac{\alpha}{t^2+\alpha}\right)^{p/2} dt. \tag{22}$$

We have three cases depending on the relation between p and r.
Consider first the case $p < r$. It is convenient to split $f(\alpha)$ as follows

$$f(\alpha) = \alpha^{p/2}\left(\int_0^q t^{r-p-1}\,dt - \int_0^q g(t)\,dt\right), \tag{23}$$

where

$$g(t) = t^{r-1}\left(\frac{1}{t^p} - \left(\frac{1}{t^2+\alpha}\right)^{p/2}\right).$$

We can conveniently rewrite $g(t)$ as

$$g(t) = t^{r-p-1}\left(1 - \left(\frac{t^2}{t^2+\alpha}\right)^{p/2}\right).$$

Setting $y = t^2/\alpha$, we have

$$\int_0^q g(t)\,dt = \frac{\alpha^{(r-p)/2}}{2}\int_0^{q^2/\alpha} y^{(r-p)/2-1}\frac{(y+1)^{p/2} - y^{p/2}}{(y+1)^{p/2}}\,dy. \tag{24}$$

We consider two cases: $p < 2$ and $p \geq 2$.
Let us start with $p < 2$. Notice that $(y+1)^{p/2} - y^{p/2} \leq 1$. Then from (24) we get

$$\begin{aligned}\int_0^q g(t)\,dt &\leq \frac{\alpha^{(r-p)/2}}{2}\int_0^{q^2/\alpha}\frac{y^{(r-p)/2-1}}{(y+1)^{p/2}}\,dy \\ &= \frac{q^{r-p}}{r-p}F\left(\frac{p}{2},\frac{r-p}{2};\frac{r-p}{2}+1;-\frac{q^2}{\alpha}\right),\end{aligned}$$

due to formula [3.194, 1] of [1] (see also [1] for the definition and the properties of the hypergeometric function $F(a, b; c; x)$).
Substituting into (23) and solving the first integral, we have

$$f(\alpha) \geq \alpha^{p/2} \frac{q^{r-p}}{r-p} - \alpha^{p/2} \frac{q^{r-p}}{r-p} F\left(\frac{p}{2}, \frac{r-p}{2}; \frac{r-p}{2}+1; -\frac{q^2}{\alpha}\right).$$

Hence, (21) becomes

$$[e_k^{\mathrm{ran}}(A,p)]^p \geq \frac{r c_r}{(r-p)c_n} \int_{B_{n-r}} a^p b_{r+1}^p \left(1 - \sum_{i=r+1}^{n} b_i^2\right)^{(r-p)/2} db -$$

$$-\frac{r c_r}{(r-p)c_n} \int_{B_{n-r}} a^p b_{r+1}^p \left(1 - \sum_{i=r+1}^{n} b_i^2\right)^{(r-p)/2} F\left(\frac{p}{2}, \frac{r-p}{2}; \frac{r-p}{2}+1; -\frac{1-\sum_{i=r+1}^n b_i^2}{a^2 b_{r+1}^2}\right) db.$$

Using [4.642] of [1], we get

$$[e_k^{\mathrm{ran}}(A,p)]^p \geq a^p \frac{\Gamma((p+1)/2)\,\Gamma((r-p)/2)}{\Gamma(1/2)\,\Gamma(r/2)} -$$

$$- a^p \frac{(r+1)\Gamma((r+1)/2)}{(r-p)\Gamma(r/2)\,\Gamma(1/2)} \int_0^1 t^p (1-t^2)^{(r-p)/2} F\left(\frac{p}{2}, \frac{r-p}{2}; \frac{r-p}{2}+1; -\frac{1-t^2}{a^2 t^2}\right) dt. \qquad (25)$$

After setting $y = (1-t^2)/(a^2 t^2)$, we can rewrite the integral in (25) as

$$\int_0^1 t^p (1-t^2)^{(r-p)/2} F\left(\frac{p}{2}, \frac{r-p}{2}; \frac{r-p}{2}+1; -\frac{1-t^2}{a^2 t^2}\right) dt =$$

$$= \frac{a^{-p-1}}{2} \int_0^\infty \frac{y^{(r-p)/2}}{(y+1/a^2)^{(r+3)/2}} F\left(\frac{p}{2}, \frac{r-p}{2}; \frac{r-p}{2}+1; -y\right) dy.$$

¿From the last equation and using formula [7.512, 10] of [1], we have

$$\int_0^1 t^p (1-t^2)^{(r-p)/2} F\left(\frac{p}{2}, \frac{r-p}{2}; \frac{r-p}{2}+1; -\frac{1-t^2}{a^2 t^2}\right) dt =$$

$$= \frac{a^{-p-1}}{2} \frac{\Gamma((r-p)/2+1)\,\Gamma(p+1/2)\,\Gamma((r+1)/2)}{\Gamma((r+3)/2)\,\Gamma((p+r+1)/2)} F\left(p+\frac{1}{2}, \frac{r+1}{2}; \frac{p+r+1}{2}; 1-\frac{1}{a^2}\right). \quad (26)$$

Applying transformation formula to the hypergeometric function, see [9.131, 1] of [1], we have

$$F\left(p+\frac{1}{2}, \frac{r+1}{2}; \frac{p+r+1}{2}; 1-\frac{1}{a^2}\right) = a^{r+1} F\left(\frac{r+1}{2}, \frac{r-p}{2}; \frac{p+r+1}{2}; 1-a^2\right).$$

Substituting it into (26) and then into (25), we get

$$[e_k^{\mathrm{ran}}(A,p)]^p \geq a^p \frac{\Gamma((p+1)/2)\,\Gamma((r-p)/2)}{\Gamma(1/2)\,\Gamma(r/2)} - a^r \gamma,$$

where

$$\gamma = \frac{\Gamma((r-p)/2)\,\Gamma(p+1/2)\,\Gamma((r+1)/2)}{2\Gamma(r/2)\,\Gamma(1/2)\,\Gamma((r+p+1)/2)} F\left(\frac{r+1}{2}, \frac{r-p}{2}; \frac{r+p+1}{2}, 1-a^2\right).$$

This concludes the proof of the case $p < 2$.
Let $p \geq 2$. Observe that, from Lagrange's Theorem, there exists a value ξ, $y \leq \xi \leq y+1$, such that $(y+1)^{p/2} - y^{p/2} = p/2\,\xi^{p/2-1}$. Since $\xi^{p/2-1} \leq (y+1)^{p/2-1}$, we obtain the bound

$$\begin{aligned}\int_0^q g(t)\,dt &\leq \frac{\alpha^{(r-p)/2}\,p}{4}\int_0^{q^2/\alpha}\frac{y^{(r-p)/2-1}}{y+1}\,dy \\ &= q^{r-p}\frac{p}{2(r-p)}\,F\left(1,\frac{r-p}{2};\frac{r-p}{2}+1;-\frac{q^2}{\alpha}\right) \qquad (27)\end{aligned}$$

which follows from formula [3.194, 1] of [1].

Proceeding exactly as before, we get

$$f(\alpha) \geq \alpha^{p/2}\frac{q^{r-p}}{r-p} - \alpha^{p/2}\frac{q^{r-p}p}{2(r-p)}\,F\left(1,\frac{r-p}{2};\frac{r-p}{2}+1;-\frac{q^2}{\alpha}\right).$$

Using this bound in (21), we get

$$[e_k^{\mathrm{ran}}(A,p)]^p \geq \frac{r\,c_r}{(r-p)c_n}\int_{B_{n-r}} a^p b_{r+1}^p\left(1-\sum_{i=r+1}^n b_i^2\right)^{(r-p)/2} db -$$

$$-\frac{rpc_r}{2(r-p)c_n}\int_{B_{n-r}} a^p b_{r+1}^p\left(1-\sum_{i=r+1}^n b_i^2\right)^{(r-p)/2} F\left(1,\frac{r-p}{2};\frac{r-p}{2}+1;-\frac{1-\sum_{i=r+1}^n b_i^2}{a^2 b_{r+1}^2}\right) db. \qquad (28)$$

Solving the integral in (28) as before, and applying the transformation formula [9.131] of [1] to the hypergeometric function, we have

$$[e_k^{\mathrm{ran}}(A,p)]^p \geq a^p\frac{\Gamma((p+1)/2)\,\Gamma((r-p)/2)}{\Gamma(1/2)\,\Gamma(r/2)} - a^r\gamma,$$

where

$$\gamma = \frac{p\Gamma((r-p)/2)\,\Gamma((p+3)/2)}{2(r+1)\Gamma(r/2)\,\Gamma(1/2)}F\left(\frac{r+1}{2},\frac{r-p}{2};\frac{r+1}{2}+1,1-a^2\right).$$

This concludes the proof for $p < r$.

Let $p = r$. The integral denoted by $f(\alpha)$ in (22) becomes

$$f(\alpha) = \alpha^{p/2}\int_0^q t^{p-1}\left(\frac{1}{t^2+\alpha}\right)^{p/2} dt, \qquad (29)$$

and can be rewritten as

$$f(\alpha) = \alpha^{p/2}\left(\int_0^q \frac{t}{t^2+p/2\,\alpha}\,dt - \int_0^q g(t)\,dt\right), \qquad (30)$$

where

$$g(t) = \frac{t}{t^2+p/2\,\alpha} - \frac{t^{p-1}}{(t^2+\alpha)^{p/2}}.$$

Since $p = r$, we have that p is an integer between 1 and n. We analyze separately the cases $p = 1$ and $p \geq 2$. If $p = 1$, then $g(t) \leq 0$ and

$$\begin{aligned} f(\alpha) &\geq \alpha^{1/2} \left(\int_0^q \frac{t}{t^2 + 1/2\,\alpha}\, dt \right) \\ &= \frac{\alpha^{1/2}}{2} \ln \left(\frac{q^2 + 1/2\,\alpha}{1/2\,\alpha} \right). \end{aligned}$$

From (21) and since $q = \sqrt{1 - \sum_{i=2}^n b_i^2}$, we get

$$[e_k^{\mathrm{ran}}(A,1)]^1 \geq \frac{1}{c_n} \int_{B_{n-1}} \alpha^{1/2} \ln \left(\frac{1 - \sum_{i=2}^n b_i^2 + 1/2\,\alpha}{1/2\,\alpha} \right) db.$$

Let $||b|| = \sum_{i=3}^n b_i^2$, and $t = b_2/(1 - ||b||^2)^{1/2}$. Since $\alpha = a^2\, b_2^2$, using [4.642] of [1], we have

$$\begin{aligned} [e_k^{\mathrm{ran}}(A,1)] &\geq a \frac{2}{c_n} \int_{B_{n-2}} \int_0^1 t \left(1 - ||b||^2\right) \ln \left(\frac{1 - (1 - 1/2\,a^2)t^2}{1/2\,a^2 t^2} \right) dt\, db \\ &= a \frac{(n-2)c_{n-2}}{c_n} B\left(\frac{n}{2} - 1, 2\right) \int_0^1 t \ln \left(\frac{1 - (1 - 1/2\,a^2)t^2}{1/2 a^2 t^2} \right) dt \\ &= a \frac{(n-2)c_{n-2}}{c_n} B\left(\frac{n}{2} - 1, 2\right) \frac{1}{2(1 - 1/2\,a^2)} \ln \left(\frac{2}{a^2} \right) \\ &\geq a \frac{(n-2)c_{n-2}}{2c_n} B\left(\frac{n}{2} - 1, 2\right) \ln \left(\frac{1}{a^2} \right) + a \frac{(n-2)c_{n-2}}{2c_n} B\left(\frac{n}{2} - 1, 2\right) \ln(2), \end{aligned}$$

from which we have

$$e_k^{\mathrm{ran}}(A,1) \geq \frac{a}{\pi} \ln \left(\frac{1}{a^2} \right) + \frac{a}{\pi} \ln(2).$$

This provides the proof for $p = r = 1$.

Let now $p \geq 2$. We notice that $t^2 + p/2\,\alpha \geq t^2 + \alpha$. Then

$$g(t) \leq \frac{t(t^2 + \alpha)^{p/2-1} - t^{p-1}}{(t^2 + \alpha)^{p/2}}.$$

Setting $y = t^2/\alpha$, we have

$$\begin{aligned} \int_0^q g(t)\, dt &\leq \int_0^{q^2/\alpha} \frac{(y+1)^{p/2-1} - y^{p/2-1}}{2(y+1)^{p/2}}\, dy \\ &\leq \frac{1}{2} \left(\frac{p}{2} - 1 \right) \int_0^{q^2/\alpha} \frac{1}{(y+1)^2}\, dy \\ &= \frac{1}{2} \left(\frac{p}{2} - 1 \right) \frac{q^2}{\alpha + q^2}. \end{aligned}$$

We substitute this inequality into (30). We have

$$\begin{aligned} f(\alpha) &\geq \alpha^{p/2} \left(\int_0^q \frac{t}{t^2 + p/2\,\alpha}\, dt - \frac{1}{2} \left(\frac{p}{2} - 1 \right) \right) \\ &= \frac{\alpha^{p/2}}{2} \ln \left(\frac{q^2 + p/2\,\alpha}{p/2\,\alpha} \right) - \frac{\alpha^{p/2}}{2} \left(\frac{p}{2} - 1 \right) \frac{q^2}{\alpha + q^2}. \end{aligned}$$

Since $q = \sqrt{1 - \sum_{i=p+1}^{n} b_i^2}$ and $p = r$, we obtain the lower bound

$$\begin{aligned} [e_k^{\rm ran}(A,p)]^p &\geq \frac{pc_p}{2c_n} \int_{B_{n-p}} \alpha^{p/2} \ln\left(\frac{1 - \sum_{i=p+1}^{n} b_i^2 + p/2\,\alpha}{p/2\,\alpha}\right) db - \\ &- \frac{pc_p}{2c_n}\left(\frac{p}{2} - 1\right) \int_{B_{n-p}} \alpha^{p/2} \frac{1 - \sum_{i=r}^{n} b_i^2}{\alpha + 1 - \sum_{i=r}^{n} b_i^2}\, db. \end{aligned}$$

Let $||b||^2 = \sum_{i=p+2}^{n} b_i^2$, and $t = b_{p+1}/(1 - ||b||^2)^{1/2}$. Then from the definition of α and using [4.642] of [1], we have

$$\begin{aligned} [e_k^{\rm ran}(A,p)]^p &\geq a^p \frac{pc_p}{c_n} \int_{B_{n-p-1}} \int_0^1 t^p (1 - ||b||^2)^{(p+1)/2} \ln\left(\frac{1 - (1 - p/2\,a^2)t^2}{p/2\,a^2 t^2}\right) dt\, db - \\ &- a^p \frac{pc_p}{c_n}\left(\frac{p}{2} - 1\right) \int_{B_{n-p-1}} \int_0^1 t^p (1 - ||b||^2)^{(p+1)/2} \frac{(1 - t^2)}{1 - (1 - a^2)t^2}\, dt\, db. \end{aligned}$$

Using again [4.642] we get

$$\begin{aligned} [e_k^{\rm ran}(A,p)]^p &\geq a^p \gamma B\left(\frac{n-p-1}{2}, \frac{p+1}{2} + 1\right) \int_0^1 t^p \ln\left(\frac{1 - (1 - p/2\,a^2)t^2}{p/2\,a^2 t^2}\right) dt - \\ &- a^p \gamma \frac{p-2}{2} B\left(\frac{n-p-1}{2}, \frac{p+1}{2} + 1\right) \int_0^1 \frac{t^p(1 - t^2)}{1 - (1 - a^2)t^2}\, dt, \end{aligned} \tag{31}$$

where $\gamma = p(n - p - 1)c_p c_{n-p-1}/(2c_n)$.
Observe that

$$\begin{aligned} &\int_0^1 t^p \ln\left(\frac{1 - (1 - p/2\,a^2)t^2}{p/2\,a^2 t^2}\right) dt = \\ &= \frac{1}{p+1} \ln\left(\frac{2}{pa^2}\right) + \frac{2}{(p+1)^2} + \int_0^1 t^p \ln\left(1 - \left(1 - \frac{p}{2}a^2\right) t^2\right) dt. \end{aligned} \tag{32}$$

Notice that if $a^2 > 2/p$ then $\ln(1 - (1 - p/2\,a^2)t^2) \geq \ln(1) = 0$. Hence, from (32) and using formula [3.197, 3] of [1] to solve the integral in (31), we have

$$\begin{aligned} [e_k^{\rm ran}(A,p)]^p &\geq a^p \frac{\gamma'}{p+1} \ln\left(\frac{2}{pa^2}\right) - \\ &- a^p \gamma' \frac{2(p-2)}{(p+3)} F\left(1, \frac{p+1}{2}; \frac{p+5}{2}; 1 - a^2\right) + a^p \gamma' \frac{2}{p+1}, \end{aligned}$$

where $\gamma' = \gamma B\left((n - p - 1)/2, (p + 1)/2 + 1\right)$.

Using (10), we can express c_i in terms of the gamma function, and we get

$$\begin{aligned} &[e_k^{\rm ran}(A,p)]^p \geq a^p \frac{\Gamma((p+1)/2)}{\Gamma(p/2)\,\Gamma(1/2)} \ln\left(\frac{1}{a^2}\right) - \\ &- a^p \frac{\Gamma((p+1)/2)}{\Gamma(p/2)\,\Gamma(1/2)} \left[\frac{2(p-2)}{p+3} F\left(1, \frac{p+1}{2}; \frac{p+5}{2}; 1 - a^2\right) - \ln\left(\frac{2}{p}\right) - \frac{2}{p+1}\right]. \end{aligned} \tag{33}$$

Otherwise, when $a^2 \le 2/p$, we use the fact that

$$\ln(1 - ct^2) = -\sum_{i=1}^{\infty} \frac{(ct^2)^i}{i},$$

where c is a constant such that $-1 \le ct^2 < 1$.
Setting $c = (1 - p/2a^2)$, and using previous relation, the integral in (32) becomes

$$\int_0^1 t^p \ln\left(1 - \left(1 - \frac{p}{2}a^2\right) t^2\right) dt = -\sum_{i=1}^{\infty} \frac{(1 - p/2\,a^2)^i}{i(2i + p + 1)} \ge -\frac{1}{p+1} \ln\left(\frac{p+3}{2}\right).$$

In this case, from (32) we have

$$\int_0^1 t^p \ln\left(\frac{1 - (1 - p/2a^2)t^2}{p/2a^2t^2}\right) dt \ge \frac{1}{p+1} \ln\left(\frac{2}{pa^2}\right) - \frac{1}{p+1} \ln\left(\frac{p+3}{2}\right) + \frac{2}{(p+1)^2},$$

and then

$$\begin{aligned}[e_k^{\mathrm{ran}}(A,p)]^p &\ge a^p \frac{\Gamma\left((p+1)/2\right)}{\Gamma\left(p/2\right)\Gamma\left(1/2\right)} \ln\left(\frac{1}{a^2}\right) - \\ &- \quad a^p \frac{\Gamma\left((p+1)/2\right)}{\Gamma\left(p/2\right)\Gamma\left(1/2\right)} \left[\frac{2(p-2)}{p+3} F\left(1, \frac{p+1}{2}; \frac{p+5}{2}; 1 - a^2\right) + \ln\left(\frac{p+3}{2}\right) - \ln\left(\frac{2}{p}\right) - \frac{2}{p+1}\right],\end{aligned}$$

which concludes the proof for $p = r$.

The last case is $p > r$. Setting $y = t^2/\alpha$, the integral $f(\alpha)$ defined by (22) becomes

$$f(\alpha) = \frac{\alpha^{r/2}}{2} \int_0^{q^2/\alpha} \frac{y^{r/2-1}}{(y+1)^{p/2}}\, dy.$$

It can be rewritten as

$$f(\alpha) = \frac{\alpha^{r/2}}{2} \left[\int_0^{\infty} \frac{y^{r/2-1}}{(y+1)^{p/2}}\, dy - \int_{q^2/\alpha}^{\infty} \frac{y^{r/2-1}}{(y+1)^{p/2}}\, dy\right]. \tag{34}$$

The first integral of the right hand side of (34) can be solved using formula [3.194, 3] of [1] and is equal to $B(r/2, (p-r)/2)$.
The second integral of (34) can be solved using formula [3.194, 2] of [1] and is equal to

$$\left(\frac{\alpha}{q^2}\right)^{(p-r)/2} \frac{2}{p-r} F\left(\frac{p}{2}, \frac{p-r}{2}; \frac{p-r}{2} + 1; -\frac{\alpha}{q^2}\right).$$

Hence, (34) becomes

$$f(\alpha) = \frac{\alpha^{r/2}}{2} B\left(\frac{r}{2}, \frac{p-r}{2}\right) - \frac{\alpha^{p/2}}{(p-r)\, q^{p-r}} F\left(\frac{p}{2}, \frac{p-r}{2}; \frac{p-r}{2} + 1; -\frac{\alpha}{q^2}\right). \tag{35}$$

By substituting (35) into (21), and from the definition of α and q we have

$$\begin{aligned}[e_k^{\mathrm{ran}}(A,p)]^p &\ge a^r \frac{rc_r}{2c_n} B\left(\frac{r}{2}, \frac{p-r}{2}\right) \int_{B_{n-r}} b_{r+1}^r\, db - \\ &- \quad a^p \frac{rc_r}{(p-r)c_n} \int_{B_{n-r}} \frac{b_{r+1}^p}{(1 - \sum_{i=r+1}^n b_i^2)^{(p-r)/2}} F\left(\frac{p}{2}, \frac{p-r}{2}; \frac{p-r}{2} + 1; -\frac{a^2 b_{r+1}^2}{1 - \sum_{i=r+1}^n b_i^2}\right) db.\end{aligned}$$

Using again the technique of reducing integrals to one-dimensional integrals, we get

$$[e_k^{\mathrm{ran}}(A,p)]^p \geq a^r \frac{\gamma}{2} B\left(\frac{r}{2}, \frac{p-r}{2}\right) B\left(\frac{r+1}{2}, \frac{n-r-1}{2}+1\right) -$$
$$- \quad a^p \gamma' \int_0^1 \frac{t^p}{(1-t^2)^{(p-r)/2}} F\left(\frac{p}{2}, \frac{p-r}{2}; \frac{p-r}{2}+1; -a^2 \frac{t^2}{1-t^2}\right) dt, \tag{36}$$

where $\gamma = r c_r c_{n-r-1}/c_n$ and $\gamma' = \Gamma\left((r+1)/2+1\right)/\left((p-r)\Gamma(1/2)\Gamma(r/2)\right)$. Consider the integral in (36). By setting $z = t^2/(1-t^2)$, we obtain

$$\int_0^1 \frac{t^p}{(1-t^2)^{(p-r)/2}} F\left(\frac{p}{2}, \frac{p-r}{2}; \frac{p-r}{2}+1; -a^2 \frac{t^2}{1-t^2}\right) dt =$$
$$= \frac{1}{2a^{p-r-2}} \int_0^\infty \frac{z^{(p-1)/2}}{(a^2+z)^{(r+3)/2}} F\left(\frac{p}{2}, \frac{p-r}{2}; \frac{p-r}{2}+1; -z\right) dz. \tag{37}$$

We notice that

$$\frac{z^{(p-1)/2}}{(a^2+z)^{(r+3)/2}} \leq \frac{z^{(p-r)/2}}{(a^2+z)^2}.$$

Using this inequality and formula [7.51, 10] of [1], (37) can be bounded as follows

$$\frac{1}{2a^{p-r-2}} \int_0^\infty \frac{z^{(p-1)/2}}{(a^2+z)^{(r+3)/2}} F\left(\frac{p}{2}, \frac{p-r}{2}; \frac{p-r}{2}+1; -z\right) dz \leq$$
$$\leq \frac{1}{2a^{p-r-2}} \int_0^\infty \frac{z^{(p-r)/2}}{(a^2+z)^2} F\left(\frac{p}{2}, \frac{p-r}{2}; \frac{p-r}{2}+1; -z\right) dz$$
$$= \frac{1}{2a^{p-r-2}} \frac{\Gamma\left((p-r)/2+1\right)\Gamma\left(r/2+1\right)}{\Gamma\left(p/2+1\right)} F\left(\frac{r}{2}+1, 1; \frac{p}{2}+1; 1-a^2\right). \tag{38}$$

Substituting (38) in (36), we get

$$[e_k^{\mathrm{ran}}(A,p)]^{1/p} \geq a^r \frac{\Gamma\left((p-r)/2\right)\Gamma\left((r+1)/2+1\right)}{\Gamma\left(p/2\right)\Gamma\left(1/2\right)} -$$
$$- \quad a^{r+2} \frac{r\Gamma\left((r+1)/2+1\right)\Gamma\left((p-r)/2\right)}{\Gamma\left(p/2+1\right)\Gamma\left(1/2\right)} F\left(\frac{r}{2}+1, 1; \frac{p}{2}+1; 1-a^2\right).$$

This concludes the proof. □

4.3 Discussion

Theorems 4.1 and 4.2 state that the randomized error $e_k^{\mathrm{ran}}(A,p)$ must depend on the ratio λ_{r+1}/λ_1. In addition, these theorems describe the actual behavior of the rate of convergence for every k, p and r. We notice that only when $r > p$, we have the same rate of convergence as in the asymptotic deterministic case with $\sum_{i=1}^r b_i^2 \neq 0$. For the other two cases, $r = p$ and $r < p$, the rate convergence is slower. This is due to the fact that Theorems 4.1 and 4.2 deal with the randomized case. So, in order to compute the randomized error we have to integrate over all possible starting vectors, even those for which the power method does not converge or converges very slowly.

To give an intuitive idea about the difference in the rate of convergence between the asymptotic deterministic case (the rate is then proportional to $(\lambda_{r+1}/\lambda_1)^k$) and the randomized case, let us analyze the error for $p=1$. In this case we have only two possibilities: $r>p$ or $r=p=1$. Assuming $\sum_{i=1}^r b_i^2 \neq 0$, we have

$$\sin(\alpha_k(\mathbf{b})) = \left(\frac{\lambda_{r+1}}{\lambda_1}\right)^k \sqrt{\frac{b_{r+1}^2+\cdots+b_{r+s}^2}{b_1^2+\cdots+b_r^2}} + o\left(\left(\frac{\lambda_{r+1}}{\lambda_1}\right)\right),$$

where s is the multiplicity of the second largest eigenvalue.
If $r=1$, the expected value of $\sin(\alpha_k(\mathbf{b}))$ with respect to $\mathbf{b}$ cannot be proportional to $(\lambda_2/\lambda_1)^k$ since

$$\int_{||\mathbf{b}||=1} \sqrt{\frac{b_2^2+\cdots+b_{s+1}^2}{b_1^2}}\mu(db) = +\infty.$$

A more careful analysis shows that we have to lose a factor proportional to $\ln(\lambda_1/\lambda_2)^{2k}$ in order to achieve the convergence of the integral.
For $r \geq 2$,

$$\int_{||\mathbf{b}||=1} \sqrt{\frac{b_{r+1}^2+\cdots+b_{r+s}^2}{b_1^2+\cdots+b_r^2}}\mu(db) < +\infty,$$

so we have a rate of convergence proportional to $(\lambda_{r+1}/\lambda_1)^k$ as in the deterministic case.
The explanation of the general case $p \geq 1$ is similar.

Analyzing together upper and lower bounds we have a complete behavior of the power method for computing a largest eigenvector. In fact, for every p and r, upper and lower bounds exhibit the same dependence on λ_{r+1}/λ_1 and on k.

5 Asymptotic Behavior

In Section 4 we provide upper and lower bounds for the randomized error of the power method for each step k. These bounds differ only by multiplicative constants and by lower order terms. We notice that only for upper bounds the constants depend on the size of the matrix, while for the lower bounds they depend only on p and r. Moreover, if A is a large matrix, the constants of the upper bound become huge. So, it is natural to ask if these constants are sharp. We answer this question by analyzing the asymptotic behavior of the randomized error $e_k^{\rm ran}(A,p)$.

Theorem 5.1 *Let A be a symmetric positive definite matrix, and let r, $r<n$, and s denote the multiplicities of the two largest eigenvalues λ_1 and λ_{r+1} of A. Then for every p, $1 \leq p < \infty$, we have*

$$\lim_{k\to+\infty} \frac{e_k^{\rm ran}(A,p)}{x_{r+1}^k} = \left(\frac{\Gamma\left((r-p)/2\right)\Gamma\left((p+s)/2\right)}{\Gamma\left(r/2\right)\Gamma\left(s/2\right)}\right)^{1/p} \qquad \text{for } p<r,$$

$$\lim_{k\to+\infty} \frac{e_k^{\rm ran}(A,p)}{x_{r+1}^k\,(2k)^{1/r}\left[\ln\left(1/x_{r+1}\right)\right]^{1/r}} = \left(\frac{\Gamma\left((p+s)/2\right)}{\Gamma\left(p/2\right)\Gamma\left(s/2\right)}\right)^{1/p} \qquad \text{for } p=r,$$

$$\lim_{k\to+\infty} \frac{e_k^{\rm ran}(A,p)}{x_{r+1}^{kr/p}} = \left(\frac{\Gamma\left((p-r)/2\right)\Gamma\left((r+s)/2\right)}{\Gamma\left(p/2\right)\Gamma\left(s/2\right)}\right)^{1/p} \qquad \text{for } p>r.$$

Proof. From (9) we have

$$[e_k^{\rm ran}(A,p)]^p = \frac{1}{c_n}\int_{B_n}\left(\frac{\sum_{i=r+1}^n b_i^2 x_i^{2k}}{\sum_{i=1}^r b_i^2 + \sum_{i=r+1}^n b_i^2 x_i^{2k}}\right)^{p/2} db.$$

As k goes to infinity, we write

$$[e_k^{\rm ran}(A,p)]^p = \frac{1}{c_n}\int_{B_n}\left(\frac{x_{r+1}^{2k}\sum_{i=r+1}^{r+s} b_i^2}{\sum_{i=1}^r b_i^2 + x_{r+1}^{2k}\sum_{i=r+1}^{r+s} b_i^2}(1+o(1))\right)^{p/2} db.$$

Let $a = x_{r+1}^k$. Integrating with respect to $b_{r+s+1},\dots,b_n$, we have

$$\frac{[e_k^{\rm ran}(A,p)]^p}{1+o(1)} = a^p\,\frac{c_{n-r-s}}{c_n}\int_{B_{r+s}}\left(\frac{\sum_{i=r+1}^{r+s} b_i^2}{\sum_{i=1}^r b_i^2 + a^2\sum_{i=r+1}^{r+s} b_i^2}\right)^{p/2}\left(1-\sum_{i=1}^{r+s} b_i^2\right)^{(n-r-s)/2} db$$

Let $||b||^2 = \sum_{i=1}^r b_i^2$ and let $t_i = b_i/(1-||b||^2)^{1/2}$ for $i=r+1,\dots,r+s$, and $||t||^2 = \sum_{i=r+1}^{r+s} t_i^2$. If we rewrite the last integral as an integral over the balls B_r and B_s, we have

$$\frac{[e_k^{\rm ran}(A,p)]^p}{1+o(1)} = a^p\,\frac{c_{n-r-s}}{c_n}\int_{B_r}\int_{B_s}\frac{||t||^p(1-||b||^2)^{(n+p-r)/2}(1-||t||^2)^{(n-r-s)/2}}{[||b||^2+a^2||t||^2(1-||b||^2)]^{p/2}} dt\, db.$$

Using [4.642] of [1] for both integrals, we get

$$\begin{aligned}
\frac{[e_k^{\rm ran}(A,p)]^p}{1+o(1)} &= \\
&= a^p\gamma\int_0^1\int_0^1\frac{t^{s-1}\,b^{r-1}\,t^p(1-b^2)^{(n+p-r)/2}\,(1-t^2)^{(n-r-s)/2}}{[b^2+a^2t^2(1-b^2)]^{p/2}}\,dt\,db \\
&= a^p\gamma\int_0^1 t^{p+s-1}(1-t^2)^{(n-r-s)/2}\left[\int_0^1\frac{b^{r-1}(1-b^2)^{(n+p-r)/2}}{[b^2+a^2t^2(1-b^2)]^{p/2}}\,db\right]dt, \qquad (39)
\end{aligned}$$

where $\gamma = rsc_{n-r-s}c_rc_s/c_n$.

We have now three cases depending on the relation between p and r.

Consider first the case $p<r$. Then the last integral of (39) is finite even for $a=0$. Substituting $a=0$, we get

$$\frac{[e_k^{\rm ran}(A,p)]^p}{1+o(1)} = a^p\gamma\int_0^1 t^{p+s-1}(1-t^2)^{(n-r-s)/2}\,dt\int_0^1 b^{r-p-1}(1-b^2)^{(n+p-r)/2}\,db.$$

¿From the definition of the beta function (11) we have

$$\frac{[e_k^{\rm ran}(A,p)]^p}{1+o(1)} = a^p\,\frac{\gamma}{4}\,B\left(\frac{p+s}{2},\frac{n-r-s}{2}+1\right)B\left(\frac{r-p}{2},\frac{n+p-r}{2}+1\right).$$

Using (10), we can express c_i in terms of the gamma function. We obtain

$$\frac{[e_k^{\rm ran}(A,p)]^p}{1+o(1)} = a^p\,\frac{\Gamma\left((r-p)/2\right)\Gamma\left((p+s)/2\right)}{\Gamma\left(r/2\right)\Gamma\left(s/2\right)}.$$

This proves that for $p < r$ we have

$$\lim_{k\to+\infty} \frac{e_k^{\mathrm{ran}}(A,p)}{x_{r+1}^k} = \left(\frac{\Gamma\left((r-p)/2\right)\Gamma\left((p+s)/2\right)}{\Gamma\left(r/2\right)\Gamma\left(s/2\right)}\right)^{1/p}.$$

Consider now the case $p = r$. From (39) we have

$$\frac{[e_k^{\mathrm{ran}}(A,p)]^p}{1+o(1)} = a^p\gamma\int_0^1 t^{p+s-1}(1-t^2)^{(n-p-s)/2}\left[\int_0^1 \frac{b^{p-1}(1-b^2)^{n/2}}{[b^2+a^2t^2(1-b^2)]^{p/2}}\,db\right]dt, \tag{40}$$

We expand $b^{p-1}(1-b^2)^{n/2}$ as $b^{p-1}-(n/2)b^{p+1}+O\left(b^{p+3}\right)$. Since $[b^2(1-a^2t^2)+a^2t^2]^{p/2}$ behaves as $b^p+o\left(a^2t^2\right)$, it is sufficient to consider the first two terms of the expansion.

As a approaches zero, we have

$$\begin{aligned}
&\int_0^1 \frac{b^{p-1}(1-b^2)^{n/2}}{[b^2(1-a^2t^2)+a^2t^2]^{p/2}}\,db = \\
&= \int_0^1 \frac{b^{p-1}}{[b^2(1-a^2t^2)+a^2t^2]^{p/2}}\,db + O\left(\int_0^1 \frac{b^{p+1}}{[b^2(1-a^2t^2)+a^2t^2]^{p/2}}\,db\right) \\
&= \int_0^1 \frac{b^{p-1}}{(b^2+a^2t^2)^{p/2}}\,db + O\left(\int_0^1 b\,db\right)
\end{aligned}$$

Observe that $(b^2+a^2t^2)^{p/2} = b^p + (p/2)b^{2(p/2-1)}a^2t^2(1+o(1))$ as $a\to 0$. Then from the last equation we have

$$\begin{aligned}
&\int_0^1 \frac{b^{p-1}}{(b^2+a^2t^2)^{p/2}}\,db + O\left(\int_0^1 b\,db\right) = \\
&= \int_0^1 \frac{b^{p-1}}{b^{p-2}\left(b^2+p/2\,a^2t^2\right)}\,db + O\left(1\right) \\
&= \int_0^1 \frac{b}{b^2+p/2\,a^2t^2}\,db + O\left(1\right) \\
&= \frac{1}{2}\ln\left(b^2+\frac{p}{2}a^2t^2\right)\Big|_0^1 + O\left(1\right) \\
&= \ln\left(\sqrt{\frac{2}{pa^2t^2}}\right)(1+o(1)).
\end{aligned}$$

Substituting this equality into (40) we get

$$\begin{aligned}
\frac{[e_k^{\mathrm{ran}}(A,p)]^p}{1+o(1)} &= a^p\gamma\int_0^1 t^{p+s-1}(1-t^2)^{(n-p-s)/2}\ln\left(\sqrt{\frac{2}{pa^2t^2}}\right)\,dt \\
&= a^p\frac{\gamma}{4}\ln\left(\frac{2}{pa^2}\right)B\left(\frac{p+s}{2},\frac{n-p-s}{2}+1\right)+O\left(a^p\right).
\end{aligned}$$

If we replace the expression for γ in the last equation, we obtain

$$\lim_{k\to+\infty} \frac{e_k^{\mathrm{ran}}(A,p)}{x_{r+1}^k(2k)^{1/r}\left[\ln\left(1/x_{r+1}\right)\right]^{1/r}} = \left(\frac{\Gamma\left((p+s)/2\right)}{\Gamma\left(p/2\right)\Gamma\left(s/2\right)}\right)^{1/p}.$$

The last case is $p > r$. We want to compute the limit

$$\lim_{x\to+\infty} \frac{e_k^{\text{ran}}(A,p)}{x_{r+1}^{kr/p}} = \left[\lim_{k\to+\infty} \frac{[e_k^{\text{ran}}(A,p)]^p}{x_{r+1}^{kr}}\right]^{1/p}$$

From (39) we get

$$\lim_{k\to+\infty} \frac{[e_k^{\text{ran}}(A,p)]^p}{x_{r+1}^{kr}} = \lim_{a\to 0} \frac{[e_k^{\text{ran}}(A,p)]^p}{a^r} =$$

$$= \lim_{a\to 0} a^{p-r}\gamma \int_0^1 t^{p+s-1}(1-t^2)^{(n-r-s)/2}\left[\int_0^1 \frac{b^{r-1}(1-b^2)^{(n+p-r)/2}}{[b^2+a^2t^2(1-b^2)]^{p/2}}\,db\right] dt. \tag{41}$$

Observe that for $a \to 0$ we have

$$\begin{aligned}
&\int_0^1 a^{p-r}\,\frac{b^{r-1}(1-b^2)^{(n+p-r)/2}}{[b^2+a^2t^2(1-b^2)]^{p/2}}\,db = \\
&= \int_0^1 a^{p-r}\,\frac{b^{r-1}(1-b^2)^{(n+p-r)/2}}{[b^2+a^2t^2]^{p/2}}\,db \\
&= \int_0^1 a^{p-r}\,\frac{b^{r-1}(1-b^2)^{(n+p-r)/2}}{a^p t^p\,(b^2/(a^2t^2)+1)^{p/2}}\,db.
\end{aligned} \tag{42}$$

We change variables by setting $y = b/(at)$. Then the integral (42) becomes

$$\frac{1}{t^{p-r}}\int_0^{1/(at)} \frac{y^{r-1}(1-a^2t^2y^2)^{(n+p-r)/2}}{(y^2+1)^{p/2}}\,dy.$$

If we set $z = y^2$, this integral can be transformed into

$$\frac{1}{2t^{p-r}}\int_0^{1/(a^2t^2)} \frac{z^{r/2-1}(1-a^2t^2z)^{(n+p-r)/2}}{(z+1)^{p/2}}\,dz.$$

We substitute this integral into (41). We get

$$\lim_{x\to+\infty} \frac{[e_k^{\text{ran}}(A,p)]^p}{x_{r+1}^{kr}} = \lim_{a\to 0} \frac{[e_k^{\text{ran}}(A,p)]^p}{a^r} =$$

$$= \frac{\gamma}{2}\int_0^1 t^{r+s-1}(1-t^2)^{\frac{n-r-s}{2}}\left[\lim_{a\to 0}\int_0^{\frac{1}{a^2t^2}} \frac{z^{\frac{r}{2}-1}(1-a^2t^2z)^{\frac{n+p-r}{2}}}{(z+1)^{\frac{p}{2}}}\,dz\right] dt. \tag{43}$$

To find the limit of the last integral, we use the following bounds (for $a < 1$)

$$\int_0^{\frac{1}{at}} \frac{z^{\frac{r}{2}-1}(1-at)^{\frac{n+p-r}{2}}}{(z+1)^{\frac{p}{2}}}dz \le \int_0^{\frac{1}{a^2t^2}} \frac{z^{\frac{r}{2}-1}(1-a^2t^2z)^{\frac{n+p-r}{2}}}{(z+1)^{\frac{p}{2}}}dz \le \int_0^{\frac{1}{a^2t^2}} \frac{z^{\frac{r}{2}-1}}{(z+1)^{\frac{p}{2}}}dz.$$

Since

$$\lim_{a\to 0}\int_0^{1/(at)} \frac{z^{r/2-1}(1-at)^{(n+p-r)/2}}{(z+1)^{p/2}}\,dz = \lim_{a\to 0}\int_0^{1/(a^2t^2)} \frac{z^{r/2-1}}{(z+1)^{p/2}}\,dz,$$

passing to the limit and then using [3.194, 3] of [1], we get

$$= \lim_{a\to 0}\int_0^{1/(a^2t^2)} \frac{z^{r/2-1}}{(z+1)^{p/2}}\,dz = \int_0^{+\infty} \frac{z^{r/2-1}}{(z+1)^{p/2}}\,dz = B\left(\frac{r}{2}, \frac{p-r}{2}\right).$$

Hence, we also have

$$\lim_{a\to 0}\int_0^{1/(a^2t^2)} \frac{z^{r/2-1}(1-a^2t^2z)^{(n+p-r)/2}}{(z+1)^{p/2}}\,dz = B\left(\frac{r}{2}, \frac{p-r}{2}\right).$$

From (43), we get

$$\begin{aligned}\lim_{k\to+\infty} \frac{[e_k^{\mathrm{ran}}(A,p)]^p}{x_{r+1}^{kr}} &= \\ &= \frac{\gamma}{2} B\left(\frac{r}{2}, \frac{p-r}{2}\right)\int_0^1 t^{r+s-1}(1-t^2)^{(n-r-s)/2}dt \\ &= \frac{\Gamma((p-r)/2)\,\Gamma((r+s)/2)}{\Gamma(p/2)\,\Gamma(s/2)}.\end{aligned}$$

This concludes the proof. □

Theorem 5.1 shows that upper and lower bounds provided in Section 4 are asymptotically optimal. In fact, the analysis of the asymptotic case indicates that the upper and lower bounds cannot be improved since the constants coincide with those of the upper bound when we set the multiplicity of the second largest eigenvalue to $n-r$, and with those of the lower bound for $s=1$.
The constants increase with s and $1/r$. This corresponds to the intuitive idea that the convergence is fast if the eigenspace $\mathcal{Z}$ is large, and is slow if the eigenspace corresponding to the second largest eigenvalue is large.

Note that if p approaches infinity, the rate of convergence approaches 1 and even the constant converges to 1. This agrees with (8) for $p=\infty$.

6 Numerical Tests

We tested the power method for several matrices with many pseudorandom starting vectors $\mathbf{b}$. The matrix A can be chosen as follows. As before, let $\mathbf{u}_k(A, \mathbf{b})$ be the vector computed by the power method applied to the matrix A with starting vector $\mathbf{b}$. Observe that for any orthogonal matrix Q, we have $\mathbf{u}_k(Q^TAQ, Q^T\mathbf{b}) = \mathbf{u}_k(A, \mathbf{b})$. Moreover, the uniform distribution on the unit sphere of the vectors $\mathbf{b}$ implies the same distribution of vectors $Q^T\mathbf{b}$. So, without loss of generality, we can restrict ourselves only to consider diagonal matrices, see also [3] and [4]. Vectors uniformly distributed over the unit sphere can be generated as described in [3] and [4].
The tests were performed on a Sun SPARCsystem 10 using double precision. To compute the values of the hypergeometric and the gamma functions we used the program *Mathematica*.

We tested many different matrices of size 100 with the distributions of the eigenvalues chosen as in [4]. We tested the following distributions:

- Chebyshev distribution: $\lambda_i = 1+\cos(((2i-1)\pi)/200)$;

k	ε^{ran}	$\varepsilon^{\text{worst}}$	$\varepsilon^{\text{best}}$	ε^{lb}	ε^{ub}	p
10	9.737e-01	9.999e-01	7.567e-01	4.782e-01	7.998e+00	1
100	9.111e-01	9.999e-01	4.149e-01	4.850e-01	7.992e+00	1
1000	7.114e-01	9.999e-01	6.811e-02	5.185e-01	7.685e+00	1
10	9.735e-01	9.999e-01	7.226e-01	6.457e-01	3.522e+00	2
100	9.239e-01	9.999e-01	3.319e-01	6.394e-01	3.474e+00	2
1000	7.383e-01	9.999e-01	7.003e-02	5.799e-01	3.035e+00	2
10	9.779e-01	1.000e+00	7.649e-01	2.712e-01	1.129e+00	10
100	9.412e-01	9.999e-01	3.882e-01	2.729e-01	1.127e+00	10
1000	8.675e-01	1.000e-01	5.303e-02	2.902e-01	1.097e+00	10

Table 1: Quadratic distribution 2 with the eigenvalues: $\lambda_i = 2(1 - (i/101)^2)$.

- quadratic distribution 1: $\lambda_i = 2\,(1 - i/101)^2$;
- quadratic distribution 2: $\lambda_i = 2(1 - (i/101)^2)$;
- uniform distribution: $\lambda_i = 2(1 - i/101)$;
- logarithmic distribution: $\lambda_i = 2\log(102 - i)/\log(102)$;
- exponential distribution 1: $\lambda_i = 2\,e^{-\sqrt[3]{i}}$;
- exponential distribution 2: $\lambda_i = 1 + e^{-i}$.

¿From the theoretical bounds, see Theorems 4.1 and 4.2, it turns out that the behavior of the power method depends on the relation between r and p. We tested the power method for different values of p and r for a fixed ratio between the two largest eigenvalues.

The main goal of these tests was to verify the results proved in Theorems 4.1 and 4.2, and to see how much upper and lower bounds differ from the experimental values.

In order to approximate the randomized error $e_k^{\text{ran}}(A, p)$ we have used 1,000 pseudorandom vectors $\mathbf{b}$. So, the randomized error is replaced by ε^{ran} obtained as the mean value among the 1,000 pseudorandom vectors, i.e.

$$\varepsilon^{\text{ran}} = \left(\frac{1}{1,000} \sum_{i=1}^{1,000} \sin^p(\alpha_k(b_i)) \right)^{1/p} .$$

By $\varepsilon^{\text{worst}}$ and $\varepsilon^{\text{best}}$ we denote, respectively, the worst and best value of $\sin(\alpha_k(b_i))$. These values give an indication about how much ε^{ran} differs from the values $\sin(\alpha_k(b_i))$. Let ε^{lb} and ε^{ub} denote the lower and the upper bounds computed using formulas given by Theorems 4.2 and 4.1. Finally, k and p are the number of iterations and the parameter of the norm, respectively.

In order to underline the dependence of the rate of convergence on the ratio between the two largest eigenvalues we report the results obtained for the quadratic distribution 2, see Table 1, and the exponential distribution 1, see Table 2. In fact, these distributions are those (among the different distributions considered) for which we have the largest (the smallest) ratio between λ_2 and λ_1 and then the slowest (the fastest) convergence, respectively.

k	$\varepsilon^{\rm ran}$	$\varepsilon^{\rm worst}$	$\varepsilon^{\rm best}$	$\varepsilon^{\rm lb}$	$\varepsilon^{\rm ub}$	p
10	1.770e-01	9.999e-01	9.630e-04	1.698e-01	2.124e+00	1
30	2.432e-03	8.996e-01	1.077e-06	2.300e-03	2.864e-02	1
10	2.509e-01	9.999e-01	7.056e-04	2.368e-01	9.616e-01	2
30	2.468e-02	7.652e-01	9.823e-07	2.006e-02	7.148e-02	2
10	6.801e-01	9.999e-01	2.079e-03	3.888e-01	8.715e-01	10
30	3.562e-01	7.081e-01	1.977e-06	3.421e-01	5.182e-01	10

Table 2: Exponential distribution 1 with the eigenvalues: $\lambda_i = 2\,e^{-i^{1/3}}$.

k	$\varepsilon^{\rm ran}$	$\varepsilon^{\rm worst}$	$\varepsilon^{\rm best}$	$\varepsilon^{\rm lb}$	$\varepsilon^{\rm ub}$	p
10	4.100e-01	9.960e-01	1.124e-01	1.276e-01	1.920e+00	1
30	3.765e-03	1.075e-01	2.915e-04	3.693e-03	4.622e-02	1
10	4.593e-01	9.989e-01	1.171e-01	1.570e-01	3.650e+00	2
30	7.979e-03	1.063e-01	2.693e-04	7.511e-03	1.245e-01	2
10	6.904e-01	9.973e-01	1.190e-01	2.472e-01	8.850e-01	10
30	2.551e-01	5.089e-01	2.435e-04	1.950e-01	4.200e-01	10

Table 3: Modified exponential distribution 2 with the eigenvalues: $\lambda_1 = \lambda_2 = 1 + e^{-1}$, and $\lambda_i = 1 + e^{-(i-1)}$, for $i = 3, \ldots, n$.

¿From Table 1 we see that for three different values of p, even after 1,000 iterations the randomized error is still very close to 1. An important observation concerns the lower and upper bounds. We notice that the lower bound is a good approximation of the expected value $\varepsilon^{\rm ran}$ while the upper bound is clearly an overestimate. This is due to the following reasons :

1. The constants in the upper bounds, see Theorem 4.1, grow with the size of the matrix.
2. Since the ratio $x_2 = \lambda_2/\lambda_1$ is very close to 1, x_2^k goes very slowly to 0 with k. In this case, the upper bound is more sensitive of the big multiplicative constants.

Table 2 is more interesting since it allows us to see the dependence of the speed of convergence on p and r. The speed of convergence is now good. In fact, after only 30 iterations we get an error of the order of 10^{-3} when $p = r = 1$. In this case, we have also that $\varepsilon^{\rm lb}$ and $\varepsilon^{\rm ub}$ are relatively close to each other, and that the error $\varepsilon^{\rm ran}$ for $k = 30$ is very close to the theoretical lower bound.

In general, it is possible to observe that the values of $\varepsilon^{\rm ran}$ computed with these tests are very close to the theoretical lower bounds while they are more distant from the upper bounds even for small λ_{r+1}/λ_1. This is due to the importance of the multiplicity s of λ_{r+1}, as it turns out from the asymptotic constants of Theorem 5.1. Experimental results prove that the power method behaves differently for matrices with the same two largest eigenvalues but with different multiplicities. In particular, increasing s we get bounds closer to the upper bounds.

To understand the role of p and r, we have performed tests with matrices for which the multiplicity of the largest eigenvalue is $r \geq 2$. In Table 3 we report the results for the modified exponential distribution 2 with $r = 2$.

An important observation concerns the comparison between the three cases, $p < r$, $p = r$ and $p > r$. From Table 3 it is easy to see that for the same value of k, the rates of convergence are different. For example, for $k = 30$ we have an error of the order of 10^{-3} for $p \leq r$, and of order 10^{-1} for $p > r$.

We performed also tests with matrices with only two distinct eigenvalues. These tests indicate the asymptotic dependence of the randomized error on the multiplicity s of the second eigenvalue. In particular, they show that $\varepsilon^{\mathrm{ran}}$ is closer to $\varepsilon^{\mathrm{ub}}$ when s is big. This is an important consequence of Theorem 5.1.

7 Conclusions

In this paper we have investigated the convergence of the power method for approximating an eigenvector corresponding to the largest eigenvalue. As our error measure, we have taken the sine of the acute angle $\alpha_k(\mathbf{b})$ between the vector computed by the power method after k steps with the starting vector $\mathbf{b}$, and the eigenspace related to the largest eigenvalue. We have analyzed the $\mathcal{L}_p$ norm of $\sin(\alpha_k(\cdot))$, for $p \in [1, +\infty]$. We have shown that, if the starting vector $\mathbf{b}$ is chosen according to the uniform distribution over the unit sphere, the rate of convergence depends on the ratio between the two largest eigenvalues. In particular, if r is the multiplicity of the largest eigenvalue λ_1, and the $\mathcal{L}_p$ norm is used, then the randomized error is proportional to $(\lambda_{r+1}/\lambda_1)^k$ if $p < r$, to $(\lambda_{r+1}/\lambda_1)^{kr/p}$ if $p > r$, and to $k^{1/p}\,(\lambda_{r+1}/\lambda_1)^k$ if $p = r$.

For every $p \in [1, +\infty)$, we have found asymptotic and non asymptotic bounds, and we have shown that the asymptotic constants are equal to those obtained for the upper and lower bounds when the multiplicity of the second largest eigenvalue is set to $n - r$ and 1, respectively. We stress that our results hold for a class of norms and that they show how the specific norm affects the speed of convergence.
Our bounds depend on the distribution of the eigenvalues, and we have proven that this is unavoidable. Comparing with results of [3], we conclude that approximating a largest eigenvector by the power method is more difficult than approximating the largest eigenvalue in the randomized setting.

8 Acknowledgments

I wish to thank Henryk Woźniakowski for the guidance and for the valuable help provided during all the stages of this work.

References

[1] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products.* Academic Press, 1994.

[2] E. Kostlan. Statistical Complexity of Dominant Eigenvector Calculation. *J. of Complexity*, 7:371–379, 1991.

[3] J. Kuczyński and H. Woźniakowski. Estimating the Largest Eigenvalue by the Power an Lanczos Algorithms with a Random Start. *SIAM J. Matrix Anal. Appl.*, 13:1094–1122, 1992.

[4] J. Kuczyński and H. Woźniakowski. Probabilistic Bounds on the Extremal Eigenvalues and Condition Number by the Lanczos Algorithm. *SIAM J. Matrix Anal. Appl.*, 15:672–691, 1994.

[5] B. N. Parlett. *The Symmetric Eigenvalue Problem.* Prentice-Hall, Englewood Cliffs, 1980.

[6] M. Shub. The Geometry and Topology of Dynamical Systems and Algorithms for Numerical Problems. In *Proc. of the 1983 Beijing Symposium on Differential Geometry and Differential Equations*, Ed. Liao Shantao. Science Press, Beijing, China, 1986.

[7] J. H. Wilkinson. *The Algebraic Eigenvalue Problem.* Oxford University Press, London, 1965.

[8] P. E. Wright. Statistical Complexity of the Power Method for Markov Chains. *J. of Complexity*, 5:119–143, 1989.