# A Brief Introduction to the "North Pole Problem" in Random Orthogonal Matrices*

Yi Zeng[†]

May 17, 2012

## 1 Introduction

Random Orthogonal Matrices is a group of unitary matrices with strong applications to many fields such as statistics, encryption and signal processing. Due to the properties of random orthogonal matrices, if we apply these operators to a vector on the unit sphere, we should expect the results uniformly distributed again on the unit sphere, regardless of the dimension of our space. Therefore, we call these matrices, $M$'s, to be isotropical. However, an unexpected numerical observation of these operators raised in the paper by Marzetta, Hassibi and Hochwald in 2002 [1]. They applied these three dimensional random orthogonal matrices $M$'s to the "North Pole", $x_0 = (1\ 0\ 0)'$, for twice, the distribution of outcomes no longer obey to the uniform distribution. These new vectors have a higher probability to sit arround the North Pole! In another word, the second power of the random orthogonal matrices are not isotropical any more.

In fact, in this numerical experiment [1], we may further observed that for any higher than one power of the Random Orthogonal Matrice, they are not isotropical. In the case of even powers, the outcomes tend to be in the "north hemispere", while for the odd powers, they are more likely to stay in both of the "polar region" instead of sitting closed to the "equator". Such power behavior of the random orthogonal matrices is called as the "North Pole problem".

In this paper, we first gave an introduction to the origin and background of the "North Pole Problem". In the second section, we will do a numerical experiment for the three dimensional case, and plot the comparison of the $Mx_0$, $M^2x_0$ and the $M^3x_0$ on the unit sphere by varying $M$'s. Then by such a Monte Carlo simulation, we may approach to the probabilities of $M^2x_0$ for sitting on the "north hemisphere", namely the probability $\mathbb{P}[x_0'M^2x_0 > 0]$, in dimension three or even higher. In the third section, we will present the detail development of the "North Pole" distribution, $x_0'\Gamma x_0$ and $x_0'M^2x_0$, theoretically relate then to some some known distributions. Then we can compare them with our previous

*This work is the final project of 18.338 Eigenvalues of Random Matrices, Spring 2012.

†Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts, 02139-4307 (yizeng@mit.edu).

numerical results in the previous part and try to provide answers to the "North Pole" problem in this case. Finally, the results for the distribution of $x_0' M^3 x_0$ will also be shown without proof in the last section.

# 2 Numerical Results

In this section, we will present the results from numerical simulation of $Mx_0$, $M^2x_0$ and the $M^3x_0$ on the three dimensional unit sphere and the probabilities $\mathbb{P}[x_0' M^2 x_0 > 0]$ in different dimensions through the Monte Carlo method.

First, three dimensional random orthogonal matrices are generated by the QR factorization of normal random matrices, which means $M$ is the unitary $Q$ matrix from the QR factorization. Then we apply these operators $M$ to the "North Pole" $x_0 = (1\ 0\ 0)'$ once, twice and three times respectively, then we can get three plots for the distributions of $Mx_0$, $M^2x_0$ and $M^3x_0$.

From the figure 1 we can see $Mx_0$ points are uniformly distributed on the unit sphere. The figure 2 suggests that $M^2x_0$ tend to be closer to the "North Pole", while $M^3x_0$ points have higher density in the polar region, shown in the figure 3. These results concise with the previous observation by Marzetta, Hassibi and Hochwald [1].

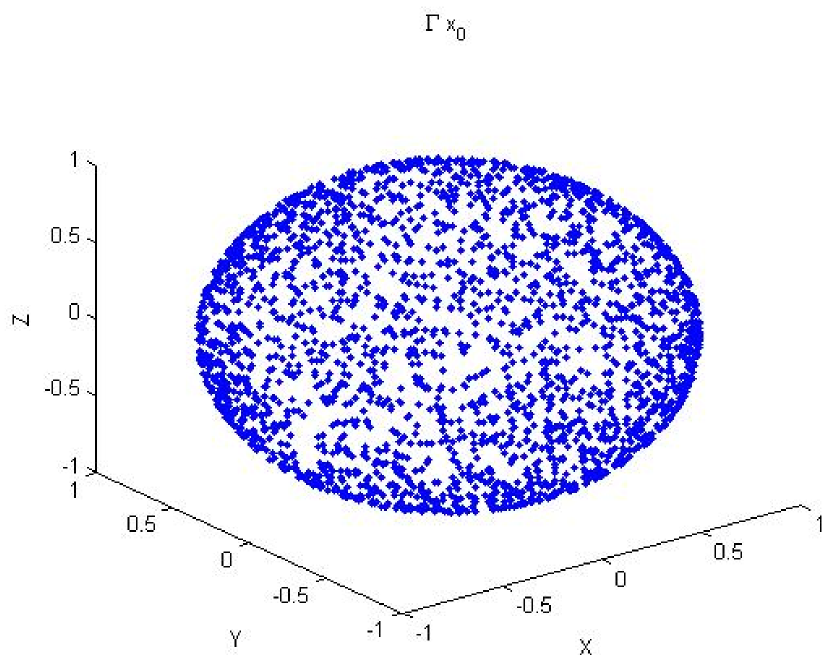


Figure 1: Distribution of $Mx_0$.

If we count the number of all outcomes $M^2x_0 = (x\ y\ z)'$ with positive $x$, we can get the probability $\mathbb{P}\ [(x\ y\ z)' : x > 0]$, which is equivalent to $\mathbb{P}[x_0' M^2 x_0 > 0]$. This method can be extended to higher dimensions. By such Monte Carlo simulations, we are able to get the approximations of probabilities $\mathbb{P}[x_0' M^2 x_0 > 0]$ in different dimensions. The results are shown in the table 1.

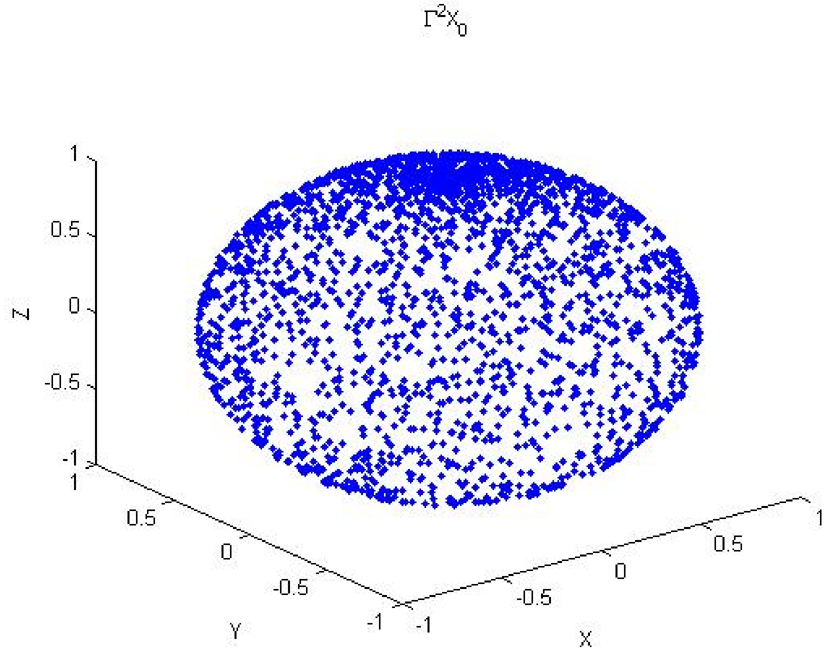


Figure 2: Distribution of $M^2 x_0$.

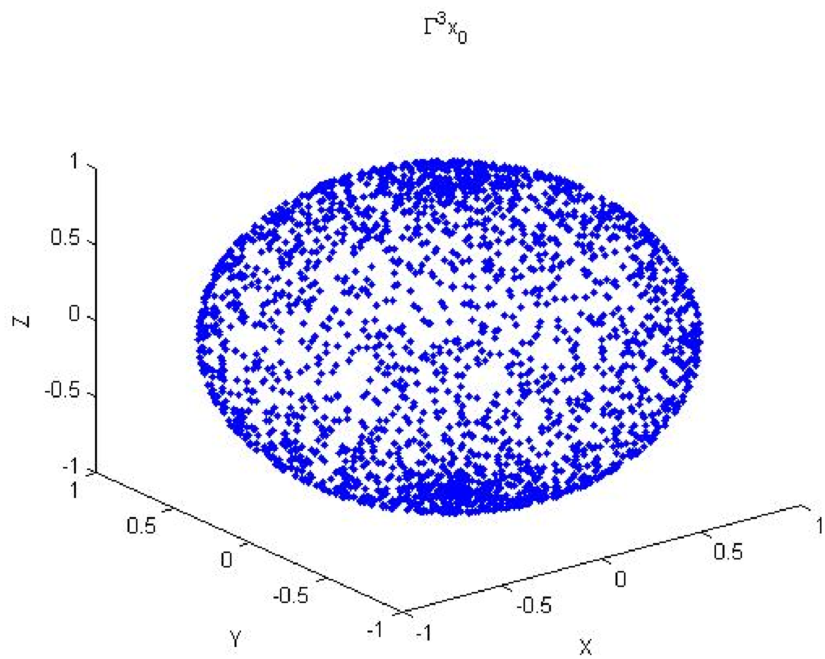


Figure 3: Distribution of $M^3 x_0$.

| Dimension $n$ | 3 | 4 | 5 | 6 | 8 | 10 | 20 | 100 |
|---|---|---|---|---|---|---|---|---|
| $\mathbb{P}[x_0' M^2 x_0 > 0]$ | 0.707 | 0.682 | 0.664 | 0.651 | 0.632 | 0.619 | 0.586 | 0.540 |

Table 1: The probabilities $\mathbb{P}[x_0' M^2 x_0 > 0]$ in different dimensions.

# 3 Distribution of $x_0' M x_0$ and $x_0' M^2 x_0$

In the third section, we are going to present the theoretical derivation of the probability density functions of both $x_0' M x_0$ and $x_0' M^2 x_0$, which is a natural question soon after our numerical observations in the second section. These proof's shown below are done by Eaton and Muirhead in 2008 [2]. For our convinience, we will define the following terminology, $V_k = x_0' M^k x_0$.

Since $V_1$ is equal to $x_0' M x_0$, it is equivalent to the entry $M_{11}$ on the upper left corner of the random orthogonal matrix $M$. The distribution of each entry in a random orthogonal matrix is well-known in all general dimensions. For dimension $n \geq 3$, $M_{11}^2$ should obey the Beta distribution with parameters $\alpha = \frac{1}{2}$ and $\beta = \frac{n-1}{2}$. It is clear that $M_{11}$ and $-M_{11}$ should have the same distribution due to the symmetry, the probability density function of $V_1$ is then given by,

$$f_n(x) = \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{n-1}{2})}(1 - x^2)^{(n-3)/2}, \tag{1}$$

while $x$ ranges in $-1 \leq x \leq 1$ and $\Gamma$ is the gamma function.

In order to further discuss the probability density function of $V_2$, we will partition the matrix $M$ into the following parts,

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \tag{2}$$

where $M_{11} \in [-1, 1]$, $M_{12} \in [-1, 1]^{1\times(n-1)}$, $M_{21} \in [-1, 1]^{(n-1)\times 1}$ and $M_{22} \in [-1, 1]^{(n-1)\times(n-1)}$. The definition of $M_{11}$ concises with the same notation above. Then $V_2$ can be written as,

$$V_2 = x_0' M^2 x_0 = M_{11}^2 + M_{12} M_{21}, \tag{3}$$

and we can further present $V_2$ as,

$$V_2 = M_{11}^2 + (1 - M_{11}^2)\frac{M_{12}}{(1 - M_{11}^2)^{1/2}}\frac{M_{21}}{(1 - M_{11}^2)^{1/2}}. \tag{4}$$

The motivation for doing this is to normalize both $M_{12}$ and $M_{21}$ into norm one, so that the part $\frac{M_{12}}{(1-M_{11}^2)^{1/2}}\frac{M_{21}}{(1-M_{11}^2)^{1/2}}$ in the equation above can be seen as an inner product of two unit vectors in $\mathbb{R}^{(n-1)}$.

Let $\Pi$ and $\Delta$ be two fixed $n \times n$ orthogonal matrices, which has the form

$$\Pi = \begin{pmatrix} 1 & 0 \\ 0 & \Pi_1 \end{pmatrix}, \tag{5}$$

$$\Delta = \begin{pmatrix} 1 & 0 \\ 0 & \Delta_1 \end{pmatrix}, \tag{6}$$

while $\Pi_1$ and $\Delta_1$ are two $(n-1)\times(n-1)$ orthogonal matrices. Therefore,

$$\Pi M \Delta = \begin{pmatrix} 1 & 0 \\ 0 & \Pi_1 \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \Delta_1 \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12}\Delta_1 \\ \Pi_1 M_{21} & \Pi_1 M_{22}\Delta_1 \end{pmatrix} \tag{7}$$

Due to the properties of orthogonal matrices, $M$ and $\Pi M\Delta$ should share the same distribution. This is because that the elements in $\mathcal{O}_n$ is one-to-one corresponding to the elements in $\Pi\mathcal{O}_n\Delta$, while $\mathcal{O}_n$ is the group of $n\times n$ orthogonal matrices. Then $V_2$ can also be expressed in the following way,

$$V_2 = x_0'\Pi M\Delta\Pi M\Delta x_0 = M_{11}^2 + (1-M_{11}^2)\frac{M_{12}}{(1-M_{11}^2)^{1/2}}\Delta_1\Pi_1\frac{M_{21}}{(1-M_{11}^2)^{1/2}}. \tag{8}$$

Notice that this equation 8 holds for all fixed $\Delta_1$ and $\Pi_1$. Thus, for any random orthogonal matrices $\Delta_1$ and $\Pi_1$, it should still hold, since $M$ and $\Pi M\Delta$ have the same distribution.Then we can choose $\Delta_1$ and $\Pi_1$ to be independent uniform on $\mathcal{O}_{(n-1)}$, so that $\Delta_1\Pi_1$ is again uniform on $\mathcal{O}_{(n-1)}$.

Let $u = \frac{M_{12}}{(1-M_{11}^2)^{1/2}}$ and $v = \frac{M_{21}}{(1-M_{11}^2)^{1/2}}$. As we already stated above, $u'$ and $v$ are both unit vectors in $\mathbb{R}^{(n-1)}$. Then apply the following lemma, we can conclude that, the probability density function for

$$\frac{M_{12}}{(1-M_{11}^2)^{1/2}}\Delta_1\Pi_1\frac{M_{21}}{(1-M_{11}^2)^{1/2}}$$

must be $f_{n-1}(\cdot)$, as we defined in 1.

**Lemma 1.** *If $u$ and $v$ are fixed unit vectors in $\mathbb{R}^n$, and $Q$ is uniform on $\mathcal{O}_n$, then the density function for $u'Qv$ is $f_n(\cdot)$.*

*Proof.* If the matrix $A$ can reflex the first coordinate unit vector $x_0$ to $u$, and the matrix $B$ can reflex $x_0$ to $v$, saying the householder matrices, we have $u = Ax_0$ and $v = Bx_0$. Then, $u'Qv = x_0'A'QBx_0$. Since $Q$ is uniformly distributed, by its invariance property, $A'QB$ is also uniform on $\mathcal{O}_n$. Notice that $x_0'(A'QB)x_0$ is just the $V_1$ we discussed before. Thus, the density function of $u'Qv$ is $f_n(\cdot)$. □

In conclusion of all the statements in this section, we fianlly come to the theorem about the distribution of $V_2$.

**Theorem 1.** *In the $n$ dimensional space, for $M$ be uniform distributed random orthogonal matrices, $V_2 = x_0'M^2x_0$ behaviors as,*

$$V_2 = T + (1-T)Y,$$

*where $T$ and $Y$ are two independent random variables. Furthermore, $T$ obeys to the Beta distribution with parameters $\alpha = \frac{1}{2}$ and $\beta = \frac{n-1}{2}$, and the random variable $Y$ has density function $f_n^Y$ satisfying*

$$f_n^Y(x) = f_{n-1}(x) = \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{n}{2}-1)}(1-x^2)^{(n-4)/2}.$$

In fact, by the theorem 2 above, easily we can derive that $\mathbb{P}(V_2 > 0) > \frac{1}{2}$ hosts for any dimension $n \geq 3$. Since $V_2$ can be seen as a linear combination of 1 and a random variable $Y$ with factor $T$ and $1 - T$. $Y$ is symmetric over zero, thus the outcome must be more likely to be positive.

Furthermore, when the dimemsion $n$ becomes larger, $T$ tends to be more possibly sitting arround zero. Then $V_2$ is now dominated by the distribution of $Y$. As we know, $\mathbb{P}(Y > 0) = \frac{1}{2}$ by symmetry, then $\mathbb{P}(V_2 > 0) \to \frac{1}{2}$ as $n \to \infty$ follows. These two conclusions are consistent with our observations listed in table 1, and they well explain the "North Pole" problem in the case $k = 2$.

# 4 Distribution of $x_0' M^3 x_0$ and More

Eaton and Muirhead also provide the behavior the distribution of $V^3 = x_0' M^3 x_0$ [2]. Straightly apply the technique in the $V^2$ case, we can get to our goal. However, the algebra is very complicated in $V^3$ case, the proof of this case will not be shown.

**Theorem 2.** *Let $\Lambda_i$ be a random variable with the probability density function $f_{(n+1-i)}(\cdot)$, while $\Lambda_1$, $\Lambda_2$ and $\Lambda_3$ are pairwisely independent, then $V_3$ behaves as,*

$$V_3 = \Lambda_1^3 + 2\Lambda_1(1 - \Lambda_1^2)\Lambda_2 + (1 - \Lambda_1^2)[-\Lambda_1\Lambda_2^2 + (1 - \Lambda_2^2)\Lambda_3].$$

From this theorem, we can see that $V_3$ is symmetric over zero. For more general cases $x_0' M^k x_0$ with $k \geq 4$, the problem of its behavior in the distribution is still opem.

# References

[1] T. L. Marzetta, B. Hassibi, B. M. Hochwald, Structured Unitary Space-time Autocoding Constellations. *IEEE Transcations on Information Theory*, 48: 942–950, 2002.

[2] M. Eaton, R. Muirhead, The "North Pole Problem" and Random Orthogonal Matrices. *Statistics and Probability Letters*, 79: 1878-1883, 2008.