

Assignment 3 - Numerical Linear Algebra

Arthur Rabello Oliveira¹, Henrique Coelho Beltrão²

16/06/2025

Abstract

coming soon

Contents

1. Introduction	2
2. Norm Distribution (a)	2
2.1. The Chi-Square Distribution	2
2.2. Histograms	3
3. Inner Products (b)	5
3.1. Histograms	6
4. The Maximum Distribution	12
5. Complexity	12
6. Another Maximum Distribution	12
7. Conclusion	12
Bibliography	12

¹Escola de Matemática Aplicada, Fundação Getúlio Vargas (FGV/EMAp), email: arthur.oliveira.1@fgv.edu.br

²Escola de Matemática Aplicada, Fundação Getúlio Vargas (FGV/EMAp), email: henrique.beltrao@fgv.edu.br

1. Introduction

2. Norm Distribution (a)

2.1. The Chi-Square Distribution

Here we construct a theoretical basis for our analysis of the histograms shown in [Section 2.2](#)

When we generate a matrix $A \in \mathbb{R}^{m \times n}$, with $A_{ij} \sim N(0, 1)$ independent, each column c_i is a gaussian vector in \mathbb{R}^m . if

$$x = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{pmatrix} \in \mathbb{R}^m \quad (1)$$

Is a column, then:

$$V = \|x\|_2 = \sqrt{\sum_{i=1}^m X_i^2} \quad (2)$$
$$V^2 = \sum_{i=1}^m X_i^2$$

Is of our interest. The expected value and variance are:

$$\mathbb{E}[V^2] = \mathbb{E}\left[\sum_{i=1}^m X_i^2\right] = \sum_{i=1}^m \mathbb{E}[X_i^2] = m \quad (3)$$
$$\text{Var}(V^2) = \text{Var}\left(\sum_{i=1}^m X_i^2\right) = \sum_{i=1}^m \text{Var}(X_i^2) = 2m$$

But we know that if $X_i \sim N(0, 1)$ are independent:

$$\sum_{i=1}^m X_i^2 \sim \chi_m^2 \quad (4)$$

where χ_m is the chi-squared distribution with m degrees of freedom, better discussed in [Section 2.1](#).

Taking the square root on [eq.\(4\)](#), we have:

$$V = \|x\|_2 = \sqrt{\sum_{i=1}^m X_i^2} \sim \sqrt{\chi_m^2} \sim \chi_m \quad (5)$$

The 2-norm of a vector x is distributed as a chi distribution with m degrees of freedom, in order to understand the distribution for many values of m , we can calculate the expected value and variance of this distribution as a function of m . The PDF of the chi distribution (with m degrees of freedom) is:

$$f_V(\varphi) = \frac{1}{2^{\frac{m}{2}-1} \cdot \Gamma(\frac{m}{2})} \varphi^{m-1} e^{-\frac{\varphi^2}{2}} \quad (6)$$

So from [this](#), the expected value is:

$$\mathbb{E}(V) = \sqrt{2} \cdot \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \quad (7)$$

And from [this](#), the variance:

$$\text{Var}(V) = m - \left(\sqrt{2} \cdot \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \right)^2 \quad (8)$$

The [Stirling Approximation](#) provides a good approximation for the expected value and variance:

$$\mathbb{E}(V) \approx \sqrt{m} \cdot \left(1 - \frac{1}{4m} + O\left(\frac{1}{m^2}\right) \right) \quad (9)$$

$$\text{Var}(V) \approx \frac{1}{2} + O\left(\frac{1}{m}\right) \quad (10)$$

2.2. Histograms

The [first cell of this notebook](#) has as expected output, with input being matrices with fixed $n = 1000$ and $m \in \{10, 20, 100, 200, 1000, 2000\}$, the following plots:

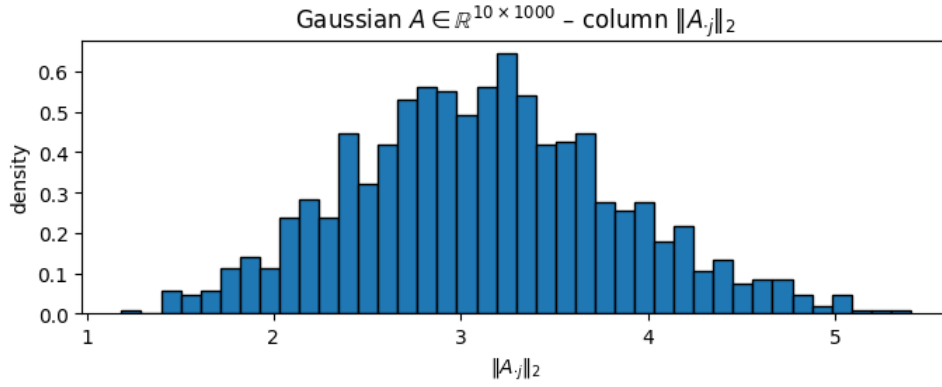


Figure 1: 10×1000 gaussian matrix

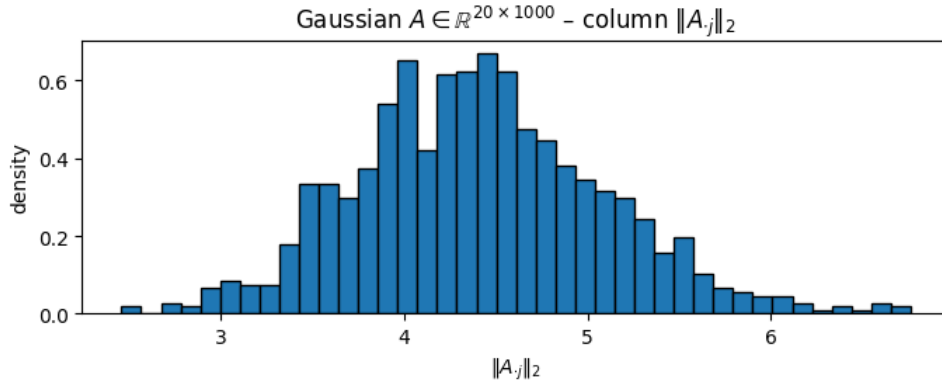


Figure 2: 20×1000 gaussian matrix

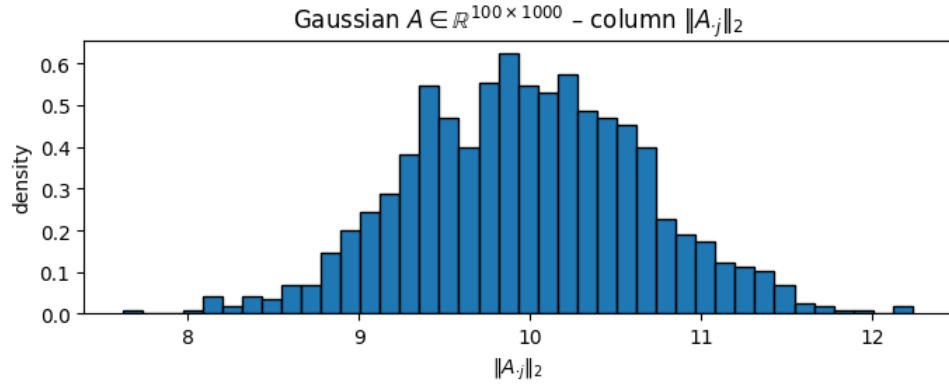


Figure 3: 100×1000 gaussian matrix

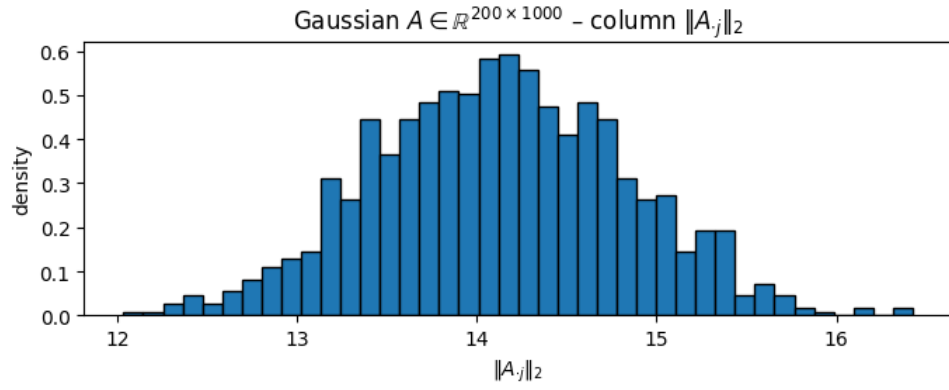


Figure 4: 200×1000 gaussian matrix

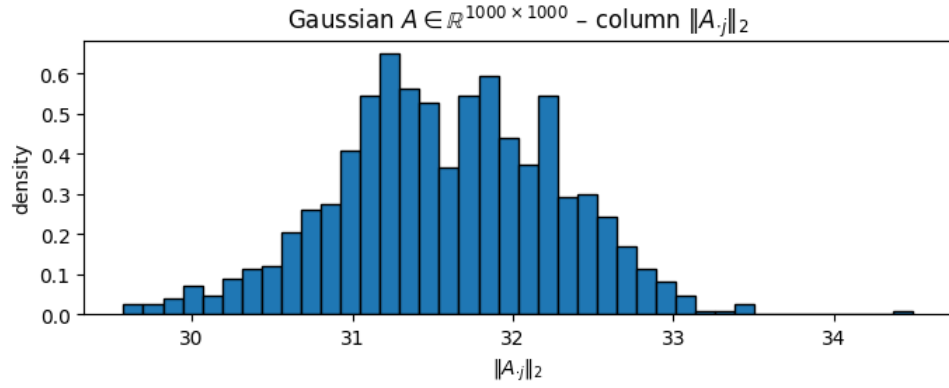


Figure 5: 1000×1000 gaussian matrix

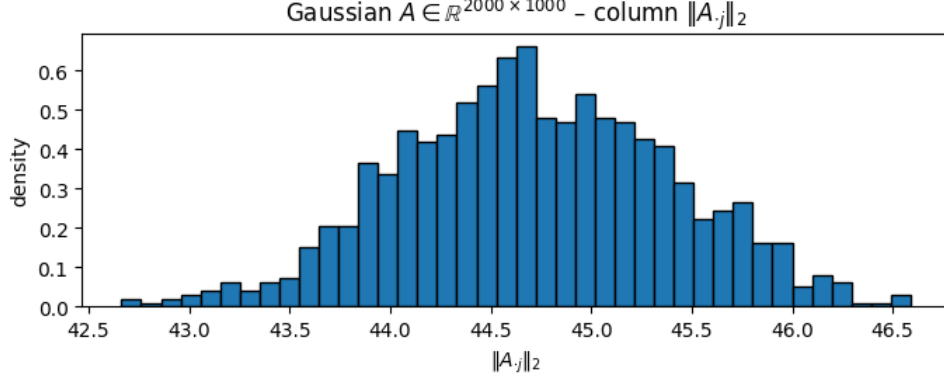


Figure 6: 2000×1000 gaussian matrix

m	approximate μ_m (theory)	$[\mu \pm 3\sigma]$ (theory)	observed spike	visual range
10	3.08	1.0 – 5.18	≈ 3.1	1.2 – 5.1
20	4.42	2.3 – 6.52	$\approx 4.3 - 4.4$	2.9 – 6.4
100	9.98	7.9 – 12.1	$\approx 9.9 - 10.0$	7.8 – 12.0
200	14.12	12.0 – 16.2	≈ 14.1	12.2 – 16.2
1000	31.61	29.5 – 33.7	$\approx 31.7 - 32.0$	29.9 – 33.3
2000	44.72	42.6 – 46.8	$\approx 44.5 - 45.0$	42.8 – 46.6

This table illustrates the expected value μ_m and the range $[\mu - 3\sigma, \mu + 3\sigma]$ for Figure 1 to Figure 6.

So apparently as m grows, the size of the gaussian vectors rapidly converge to \sqrt{m} , with small errors.

3. Inner Products (b)

Here we construct a theoretical basis for our analysis of the inner products shown in Section 3.1.

When we generate a matrix $A \in \mathbb{R}^{m \times n}$, with $A_{ij} \sim N(0, 1)$ independent, each column c_i is a gaussian vector in \mathbb{R}^m . If The inner product of two gaussian vectors $x = (X_1, \dots, X_n), y = (Y_1, \dots, Y_n)$ is:

$$Z = \langle x, y \rangle = \sum_{i=1}^m X_i Y_i \quad (11)$$

With $X, Y \sim N(0, 1)$. Since X_i, Y_j are independent, we have:

$$\begin{aligned} \mathbb{E}[Z] &= \sum_{i=1}^m \mathbb{E}[X_i Y_i] = \sum_{i=1}^m \mathbb{E}[X_i] \mathbb{E}[Y_i] = 0 \\ \text{Var}(Z) &= \sum_{i=1}^m \text{Var}(X_i Y_i) = \sum_{i=1}^m \mathbb{E}[X_i^2] \mathbb{E}[Y_i^2] = \sum_{i=1}^m 1 = m \end{aligned} \quad (12)$$

If $W = X_i Y_i$, we have:

$$M_W(\varphi) = \mathbb{E}[e^{\varphi W}] = \frac{1}{\sqrt{1 - \varphi^2}}, |\varphi| < 1 \quad (13)$$

Over all $W_i = X_i Y_i$:

$$M_Z(\varphi) = \mathbb{E}[e^{\varphi Z}] = (M_W(\varphi))^m = \left(\frac{1}{\sqrt{1-\varphi^2}} \right)^m = (1-\varphi^2)^{-\frac{m}{2}}, |\varphi| < 1 \quad (14)$$

And magically:

$$M_{\frac{Z}{\sqrt{m}}}(\varphi) = \left(1 - \frac{\varphi^2}{m} \right)^{-\frac{m}{2}} \Rightarrow \lim_{m \rightarrow \infty} M_{\frac{Z}{\sqrt{m}}}(\varphi) = e^{\frac{\varphi^2}{2}} \quad (15)$$

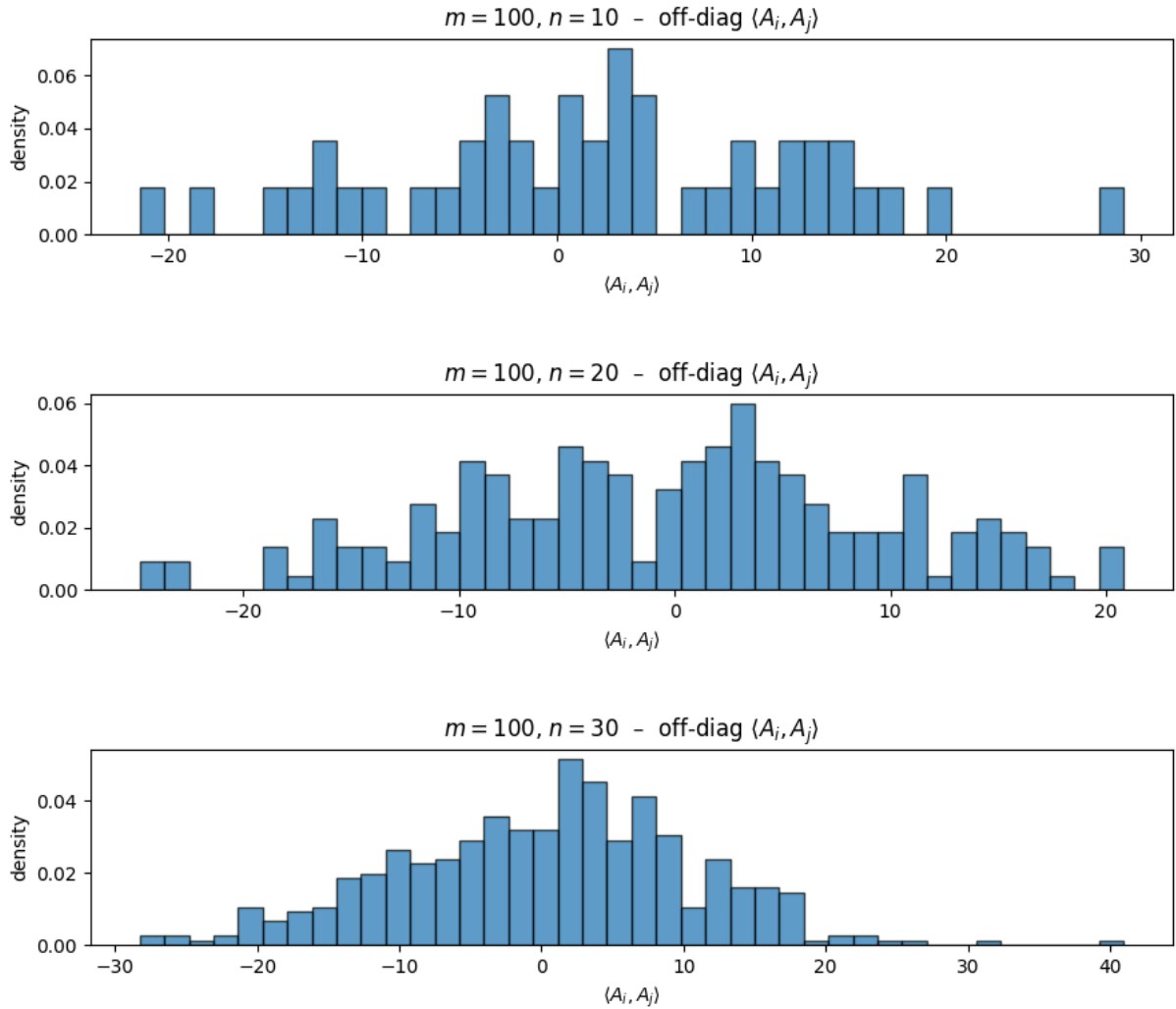
Precisely the moment generating function of a standard normal distribution, so as $m \rightarrow \infty$:

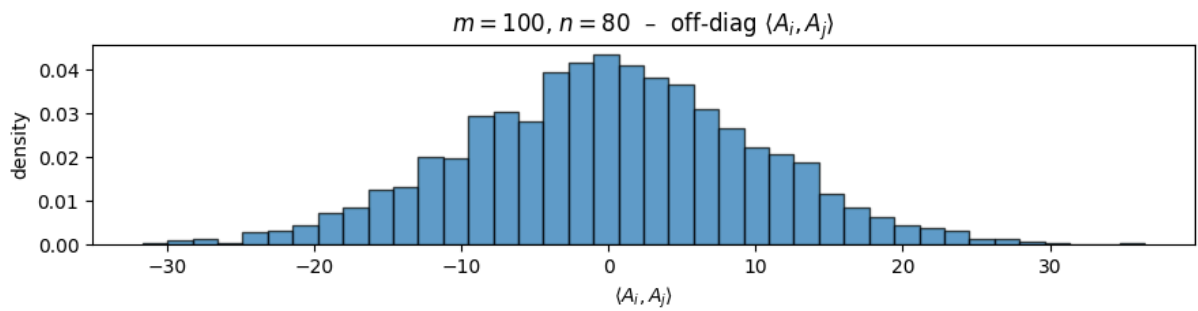
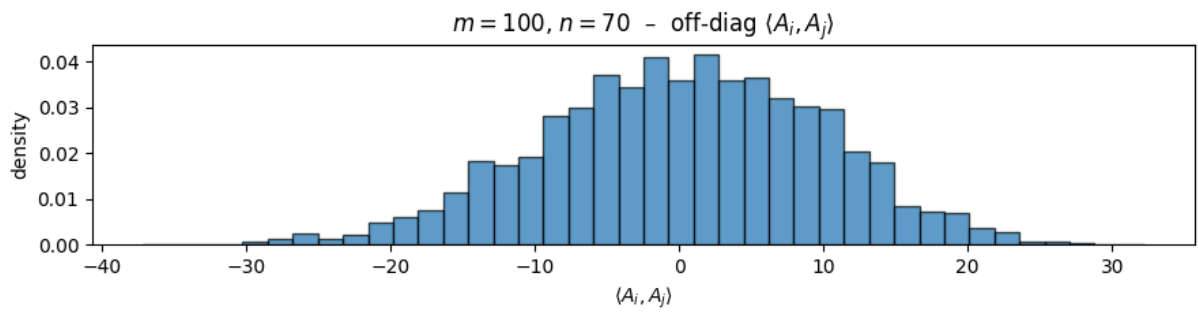
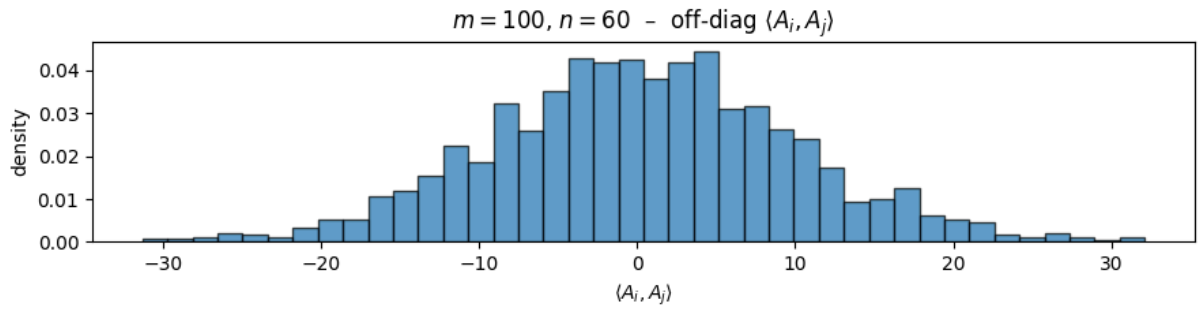
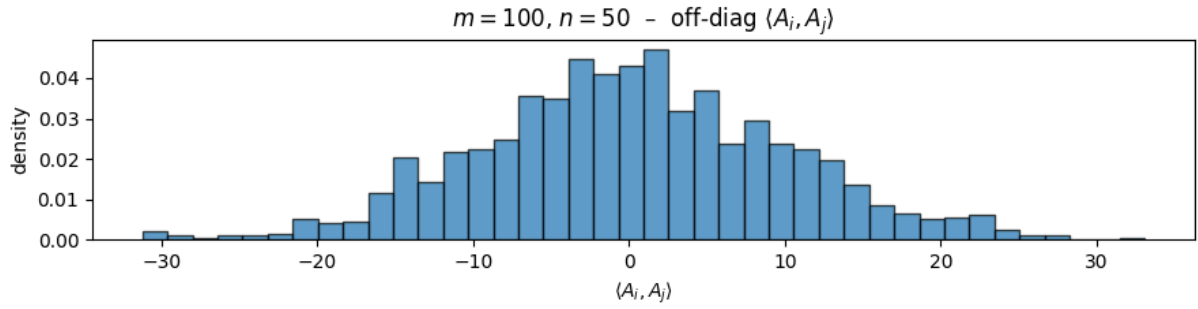
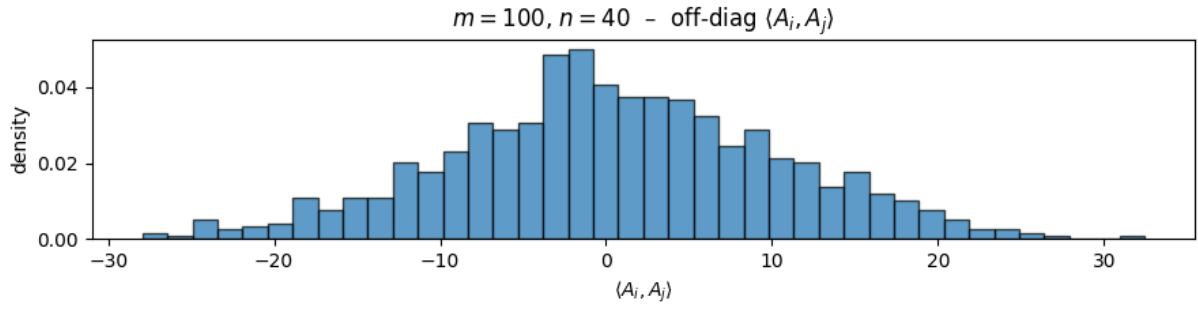
$$\frac{Z}{\sqrt{m}} \sim N(0, 1) \quad (16)$$

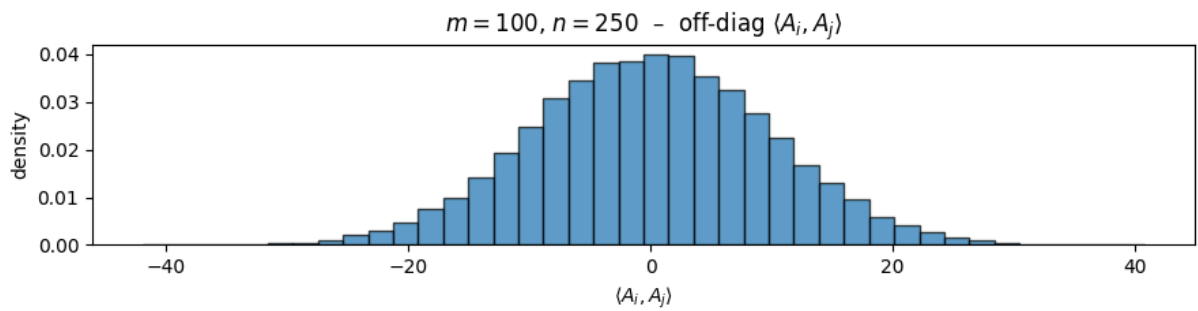
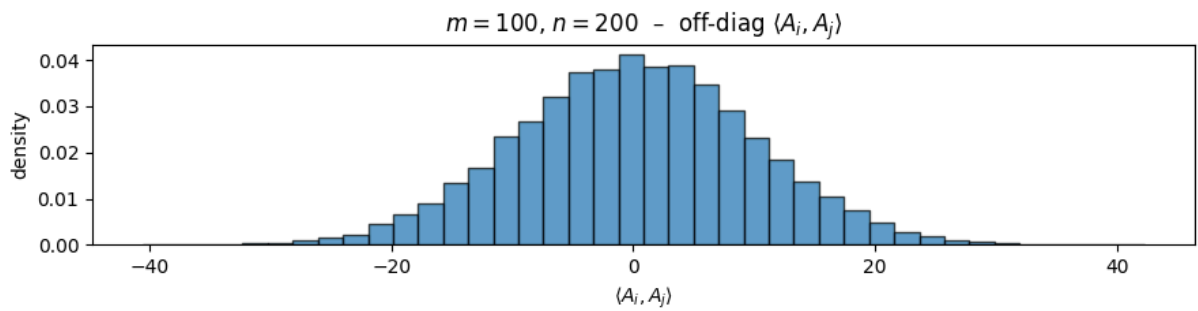
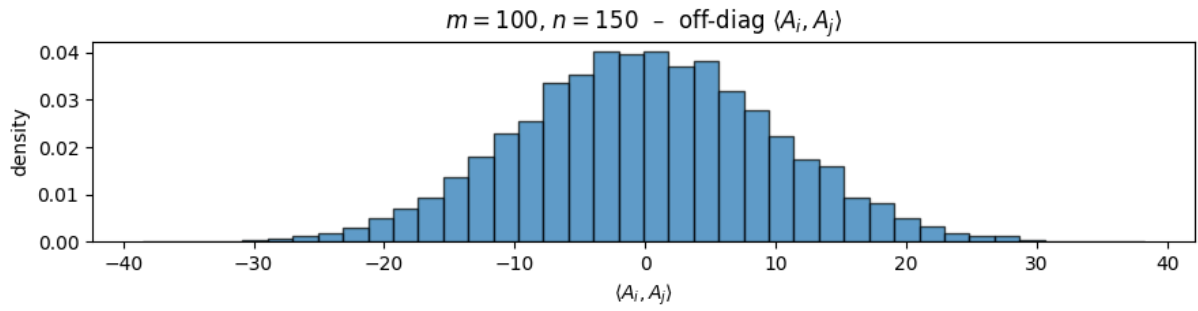
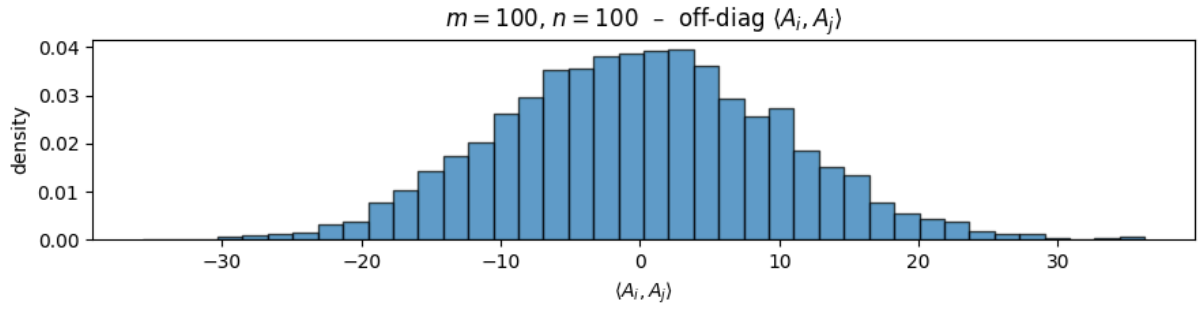
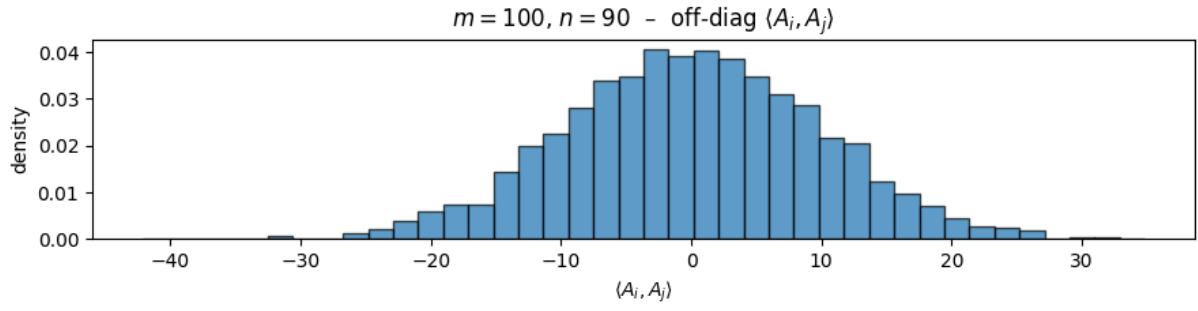
With a fixed $m = 100$, when $n \rightarrow \infty$ we can see the distribution approaching $N(0, 1)$, as shown in [Section 3.1](#)

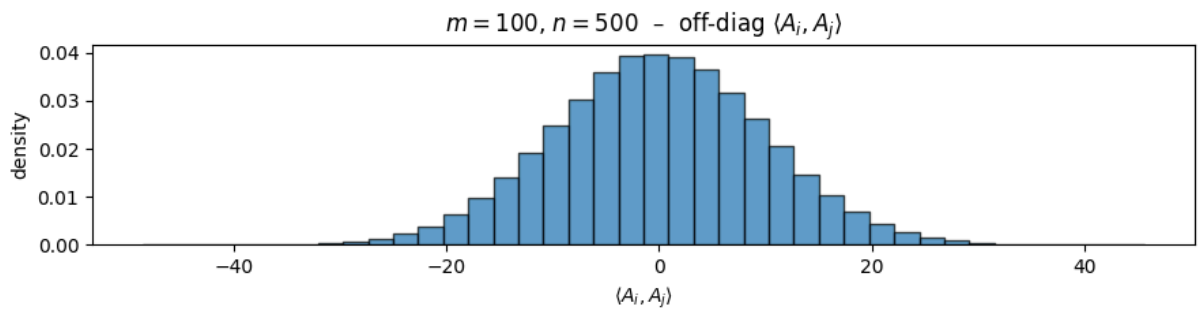
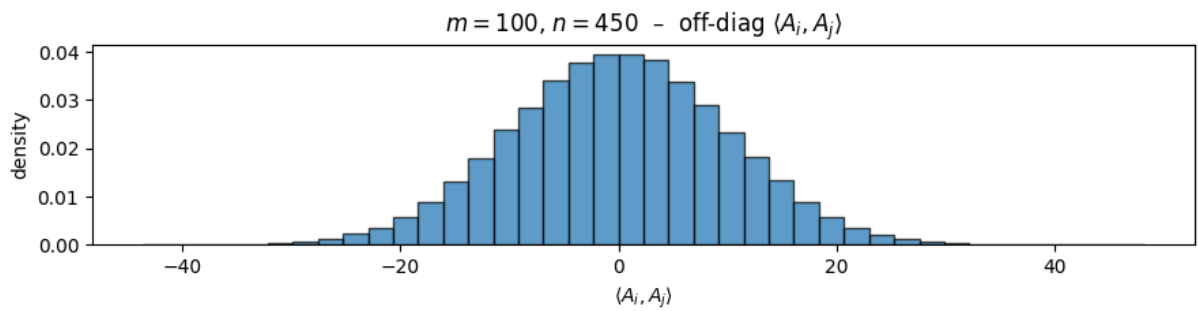
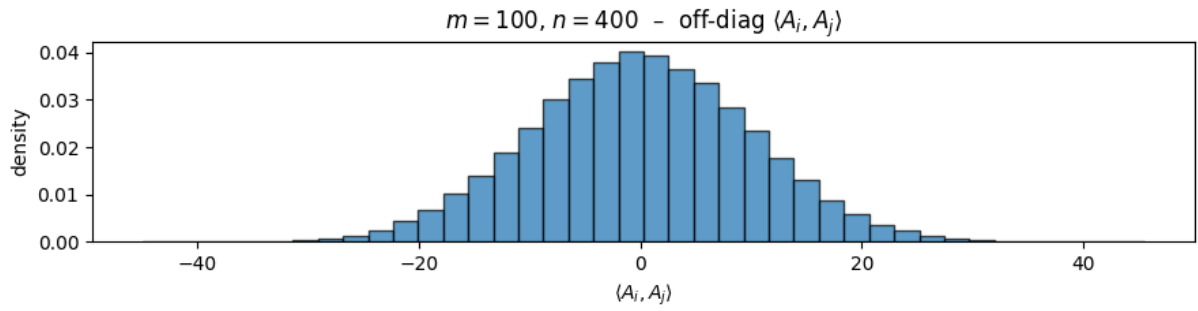
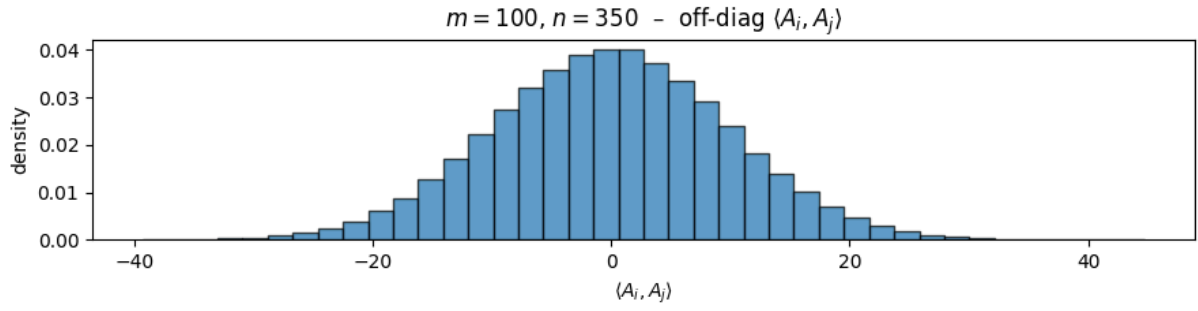
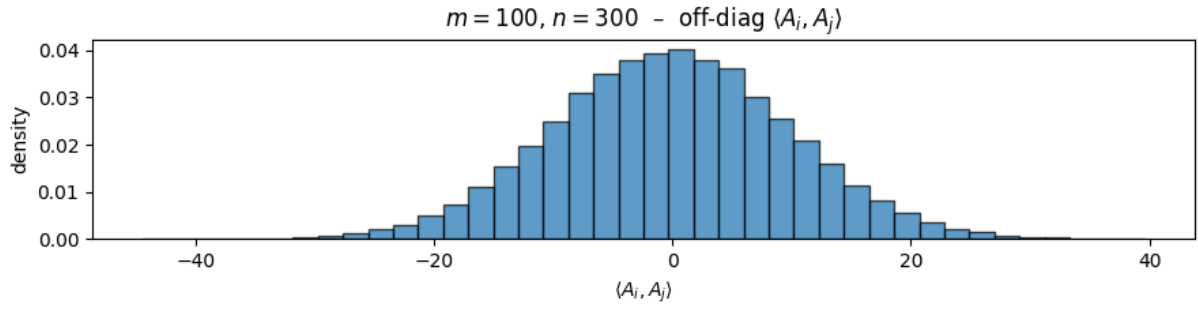
3.1. Histograms

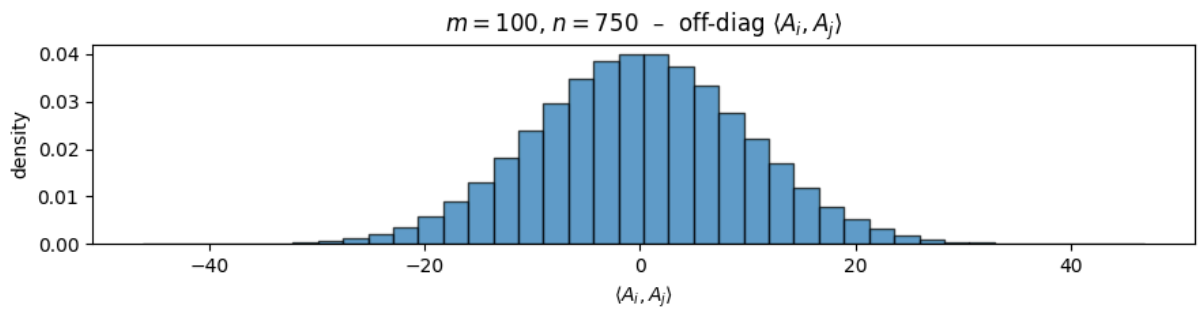
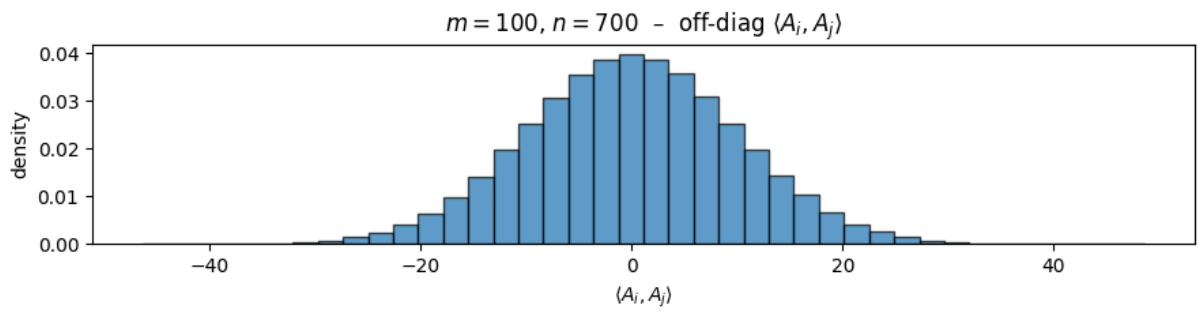
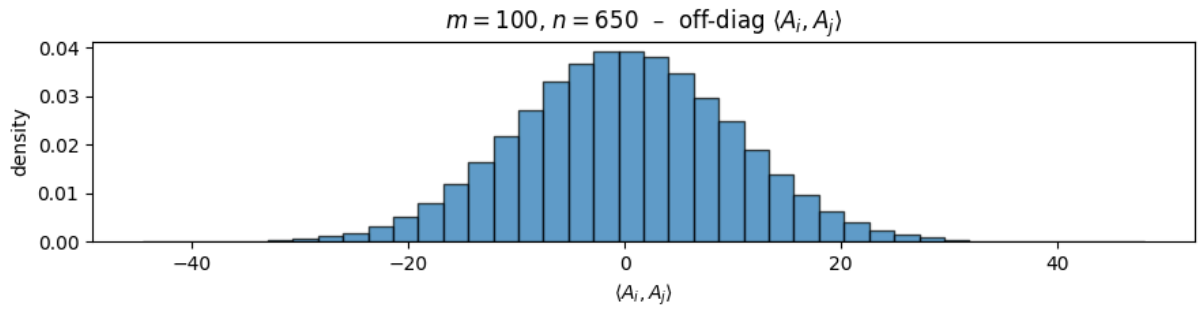
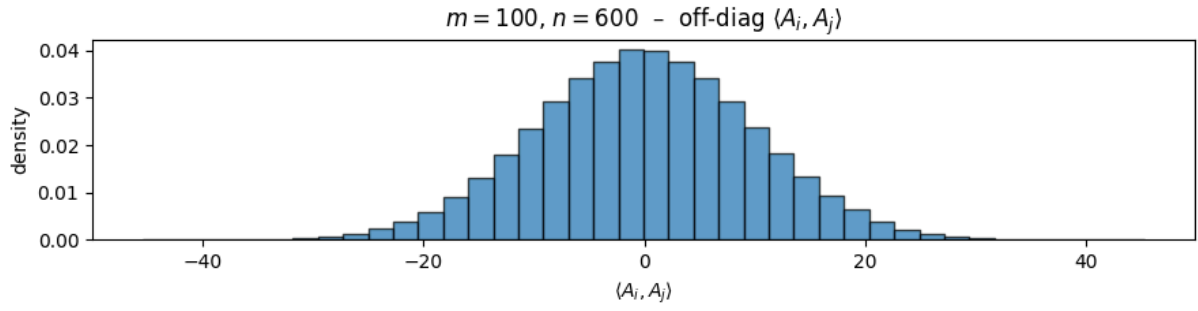
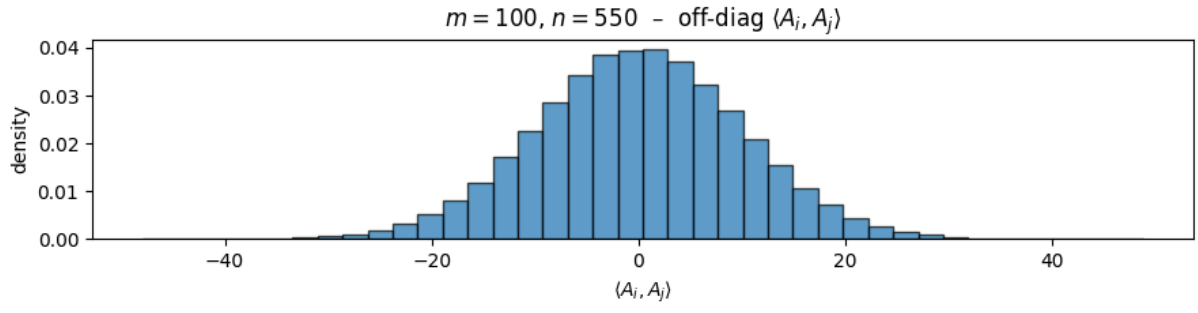
The following plots are an expected output for the second cell of [this notebook](#), with input $m = 100, n \in \{10, 20, 30, 40, 50, 60, \dots, 1000\}$:











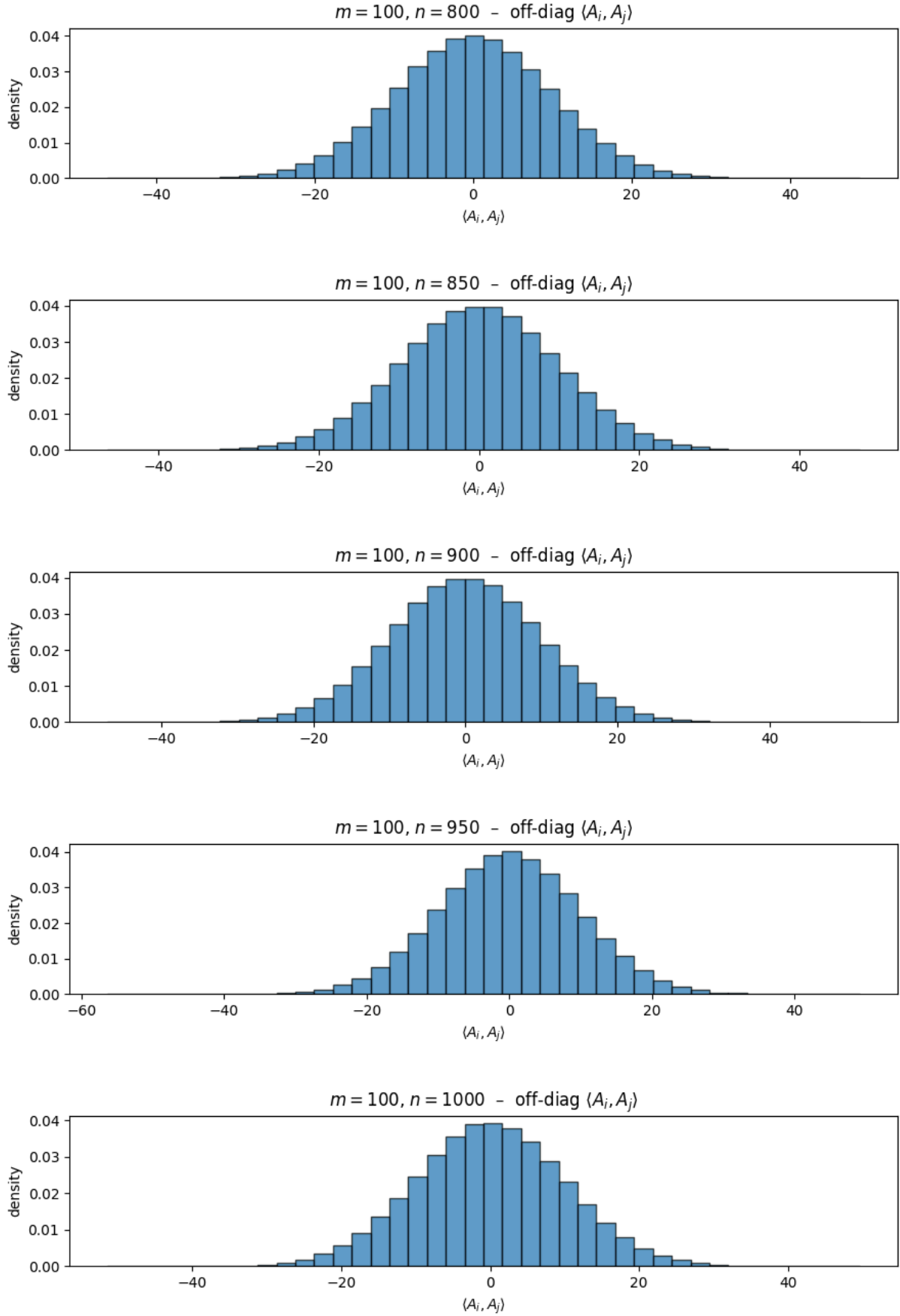


Figure 7 \rightarrow Figure 34 shows that the distribution indeed approaches $N(0, 1)$

4. The Maximum Distribution

5. Complexity

6. Another Maximum Distribution

7. Conclusion

Bibliography