2017

A Study of Conductance for a Random, Hierarchically-Structured Material

Daniel Salvador
University of Mississippi. Sally McDonnell Barksdale Honors College

A Study of Conductance for a Random, Hierarchically-Structured Material

Daniel Salvador

A thesis submitted to the the faculty of The University of Mississippi in partial fulfillment of the requirements of the Sally McDonnell Barksdale Honors College.

Oxford

May 2017

Approved by

Advisor: Professor Jeremy Clark

Reader: Professor Hailin Sang

Reader: Professor Laura Sheppardson

Abstract

I consider a mathematical model for the conductance of a system formed by a hierarchical network of random bonds. My simulations show that the net conductance converges to a fixed number $\gamma \approx 0.35337$ when the conductances of the bonds are numbers selected uniformly at random from the interval $(0, 1)$. By linearly approximating the model around γ, I derive a new simplified model which I then study in rigorous mathematical detail. I prove a generalized central limit theorem for the new linearized system.

Contents

1 Introduction

This thesis will examine the conductance of a random, hierarchically structured material. The material forms a hierarchy of diamond graphs (the specific model is introduced in Section 4.1). I will investigate the net conductance of the system as the number of hierarchical layers grows, and ultimately show that this net conductance is converging to a fixed number, and thus is not random. To accomplish this, I will use the R program to run simulations of this complicated hierarchical system. I will also theoretically analyze the error by using a similar linearized model.

Due to high technicality of this subject, I will begin in the next section by defining key terms. I then build up to proving a generalized central limit theorem for hierarchical systems and close by examining the data. The remaining sections of this thesis are organized as follows:

- In Section 2 I introduce elementary concepts and terminologies of probability theory.
- In Section 3 I analyze characteristic functions, as they will be useful in proving the central limit theorem.
- In Section 4 I study certain hierarchically defined systems. Specifically, in Section 4.1 I will consider the conductance model and provide a diagram for the model, and in Section 4.2 I will prove the central limit theorem.
- In Section 5 I discuss the methods used to collect data and conclusions drawn from the data.

2 Probability theory: definitions and notations

The following are key, intuitive concepts of probability:

- **Experiment** - a chain of circumstances leading to an outcome
- **Outcome** (ω) - the result of an experiment
- **Sample space** (Ω) - set of all possible outcomes in an experiment
- **Event** (A) - subset of Ω

In modern mathematics, probability theory is defined in terms of set theory. I will look briefly into set theory.

2.1 Set theory and σ-algebras

Let A and B be sets. Here are some basic properties of set theory:

- The **union** of A and B is the set $A \cup B = \{x : x \in A \text{ or } x \in B\}$.
- The **intersection** of A and B is the set $A \cap B = \{x : x \in A \text{ and } x \in B\}$.
- The **complement** of event A is denoted A^c. The complement contains all outcomes in Ω, except for A. $A^c = \{x \in \Omega \,\big|\, x \notin A\}$.
- The **difference** of A and B is the set $A - B = \{x : x \in A \text{ and } x \notin B\}$. Also written as $A \setminus B = A \cap B^c$.
- Sets A and B are **disjoint** if their intersection is the **empty set**, $\emptyset$. $A \cap B = \emptyset$.

Definition 2.1. ***De Morgan's Laws:***

Let $\{A_i : i \in I\}$ *be a collection of sets.*

$(\bigcup_i A_i)^c = \bigcap_i A_i^c$ *and* $(\bigcap_i A_i)^c = \bigcup_i A_i^c$.

Definition 2.2. *Let* $\mathcal{F}$ *be a collection of subsets of* Ω. *Then* $\mathcal{F}$ *is a* σ*-field if:*

(i) $\emptyset \in \mathcal{F}$,

(ii) If $A_1, A_2, \cdots \in \mathcal{F}$, *then* $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$,

(iii) If $A \in \mathcal{F}$, *then* $A^c \in \mathcal{F}$.

A set A is said to be **measurable** with respect to $\mathcal{F}$ if $A \in \mathcal{F}$.

Lemma 2.3. *A* σ*-field is closed under countable intersection. That is, if* $A_i \in \mathcal{F}$, *for* $i \in \mathbb{N}$, *then* $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.

Proof. Let $A_1, A_2, \ldots$ be in $\mathcal{F}$. Then $A_i^c \in \mathcal{F}$, for $i \in \mathbb{N}$, by closure under complements for $\mathcal{F}$.

Then, $\bigcup_{i=1}^{\infty} A_i^c \in \mathcal{F}$ by closure under countable union.

Then, $(\bigcup_{i=1}^{\infty} A_i^c)^c \in \mathcal{F}$ by complement closure, again.

Finally, using Definiton 2.1, $(\bigcup_{i=1}^{\infty} A_i^c)^c = \bigcap_{i=1}^{\infty} (A_i^c)^c = \bigcap_{i=1}^{\infty} A_i$. □

2.2 Elementary definitions of probability theory

Denote the **probability** of an event A as $\mathbb{P}(A)$. Then the following are basic characteristics that we would intuitively expect for $\mathbb{P}$:

- $0 \leq \mathbb{P}(A) \leq 1$.
- $\mathbb{P}(\emptyset) = 0$, and $\mathbb{P}(\Omega) = 1$.
- For disjoint events: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
- More generally, the probability function $\mathbb{P}$ is finitely additive. That is, if $A_1, A_2, \ldots, A_n$ are disjoint events, then $\mathbb{P}\big(\bigcup_{i=1}^{n} A_i\big) = \sum_{i=1}^{n} \mathbb{P}(A_i)$.

The actual definition of a probability measure is the following:

Definition 2.4. *A* ***probability measure*** $\mathbb{P}$ *on* $(\Omega, \mathcal{F})$ *is a function* $\mathbb{P} : \mathcal{F} \to [0, 1]$ *such that:*

(i) if $A_1, A_2, \ldots$ *is a collection of disjoint members of* $\mathcal{F}$ *(i.e.* $A_i \cap A_j = \emptyset$ *for all i,j,* $i \neq j$*), then* $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$,

(ii) $\mathbb{P}(\emptyset) = 0$, *and* $\mathbb{P}(\Omega) = 1$.

The following table gives a summary of the symbols from set theory and probability theory.

Table 1: Common set and probability jargon

Symbol	Set jargon	Probability jargon
Ω	whole space	certain event
ω	member of Ω	elementary event, outcome
$\emptyset$	empty set	impossible event
A	subset of Ω	event that some outcome in A occurs
A^c	complement of A	event that no outcome in A occurs
$A \cap B$	intersection	both A and B
$A \cup B$	union	either A or B or both
$A \setminus B$	difference	A but not B

Definition 2.5. *A* ***probability space*** *is a triple* $(\Omega, \mathcal{F}, \mathbb{P})$ *composed of a sample space* Ω*, a* σ*-field* $\mathcal{F}$*, and a probability measure* $\mathbb{P}$ *on* $(\Omega, \mathcal{F})$.

Lemma 2.6. *Properties of probability space:*

(i) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

(ii) If $B \supseteq A$*, then* $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A)$

(iii) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Proof.

(i) Notice that $\Omega = A \cup A^c$, and $A \cap A^c = \emptyset$. So, $\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(A \cup A^c) = \mathbb{P}(\Omega) = 1$.
Hence, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

(ii) A and $B \setminus A$ are disjoint. Then, $B = A \cup (B \setminus A)$. It follows that $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A)$.
Hence, $\mathbb{P}(A) \leq \mathbb{P}(B)$.

(iii) Notice that $A \cup B = A \cup (B \cap A^c)$. It follows that $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c)$.
$B = (B \cap A^c) \cup (A \cap B)$. So, $\mathbb{P}(B) = \mathbb{P}(B \cap A^c) + \mathbb{P}(A \cap B)$.
And then, $\mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
Thus, we have $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

□

Lemma 2.7. *Let $A_1, A_2, A_3 \ldots$ satisfy that $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$. For $A = \bigcup_{i=1}^{\infty} A_i$, we have $\mathbb{P}(A) = \lim_{i\to\infty} \mathbb{P}(A_i)$.*

Proof. $A = \bigcup_{j=1}^{\infty} (A_j \setminus A_{j-1})$, where $A_0 = \emptyset$. (Thus, for j=1, $A = A_1 \setminus A_0 = A_1 \setminus \emptyset$; for j=2, $A = A_2 \setminus A_1; \ldots$)

$$\begin{aligned}
\mathbb{P}(A) &= \sum_{j=1}^{\infty} \mathbb{P}(A_j \setminus A_{j-1}) \\
&= \sum_{j=1}^{\infty} \Big(\mathbb{P}(A_j) - \mathbb{P}(A_{j-1})\Big) \\
&= \lim_{n\to\infty} \sum_{j=1}^{n} \Big(\mathbb{P}(A_j) - \mathbb{P}(A_{j-1})\Big) \\
&= \lim_{n\to\infty} \Big[\Big(\mathbb{P}(A_1) - \mathbb{P}(A_0)\Big) + \Big(\mathbb{P}(A_2) - \mathbb{P}(A_1)\Big) + \cdots + \Big(\mathbb{P}(A_n) - \mathbb{P}(A_{n-1})\Big)\Big]
\end{aligned}$$

Notice the cancelation above is due to the telescoping nature of the summation.

$$\begin{aligned}
&= \lim_{n\to\infty} \Big(\mathbb{P}(A_n) - \mathbb{P}(A_0)\Big) \\
&= \lim_{n\to\infty} \mathbb{P}(A_n)
\end{aligned}$$

□

Definition 2.8. *Two events A, B are **independent** provided $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.*

*Events $A_1, A_2, \ldots, A_n$ are **independent** means for any $J \subseteq \{1, 2, \ldots, n\}$, $\mathbb{P}(\bigcap_{i\in J} A_i) = \prod_{i\in J} \mathbb{P}(A_i)$.*

2.3 Random Variables

Definition 2.9. *A **random variable** is a function $X : \Omega \to \mathbb{R}$ such that for all $\lambda \in \mathbb{R}$, the set $A_\lambda = \{\omega \in \Omega \,\big|\, X(\omega) \leq \lambda\}$ is an event, i.e., A_λ is a measurable set with respect to $\mathcal{F}$.*

If X is a random variable, then there are many other sets besides the A_λ's in Definition 2.9. The following proposition lists a few of them.

Proposition 2.10. *If $X : \Omega \to \mathbb{R}$ is random variable, then the following sets are events:*

(i) $\{\omega \in \Omega \,\big|\, \alpha \leq X(\omega) \leq \beta\}$,

(ii) $\{\omega \in \Omega \,\big|\, \alpha \leq X(\omega)\}$.

Definition 2.11. *The **distribution function** of a random variable X is the function $F : \mathbb{R} \to [0, 1]$, such that $F(a) = \mathbb{P}(X \leq a)$, for $a \in \mathbb{R}$.*

Definition 2.12. *The following are two common classes of random variables:*

- *A random variable is said to be **absolutely continuous** if the distribution function, F, has the form $F(a) = \int_{-\infty}^{a} f(t)dt$ for a **density function** $f : \mathbb{R} \to [0, \infty]$.*
- *A random variable is said to be **discrete** if it has countably many possible values, $\{a_1, a_2, a_3, \dots\}$. A discrete random variable X has a **mass function** $g : \mathbb{R} \to [0, 1]$ defined as $g(a) = \mathbb{P}(X = a)$.*

Remark 2.13. If f_X is the density of random variable X, then $\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)dx$.

Definition 2.14. *We define the **expectation** of a random variable $X : \Omega \to \mathbb{R}$ as follows:*

- ***absolutely continuous**: For an absolutely continuous random variable X with density function f_X, the expectation is defined as $\mathbb{E}[X] = \int_{\mathbb{R}} x f_X(x)dx$*
- ***discrete**: For a discrete random variable X with mass function g, the expectation is defined as $\mathbb{E}[X] = \sum_x xg(x)$*

Lemma 2.15. *The expectation function has the following properties:*

(i) If $X \geq 0$, then $\mathbb{E}[X] \geq 0$.

(ii) If $a, b \in \mathbb{R}$, then $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$. This is known as linearity.

(iii) If X, Y are independent, then $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$.

(iv) If $G : \mathbb{R} \to \mathbb{R}$ is a bounded function and f_X is the density for X, then

$$\mathbb{E}\big[G(X)\big] = \int_{\mathbb{R}} G(x) f_X(x)dx\,.$$

Definition 2.16. *For a random variable $X : \Omega \to \mathbb{R}$, we have the following definitions:*

- *If $\int |x| f_X(x) dx < \infty$ the **mean** of X is defined as the expectation, $\mathbb{E}[X]$.*
- ***Variance** of X is defined as $\mathrm{Var}(X) = \sigma^2 = \mathbb{E}\Big[\big(X - \mathbb{E}[X]\big)^2\Big]$. The variance can be rewritten as $\sigma^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.*
- *The **standard deviation** is the square root of the variance: $\sigma = \sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2]}$.*
- *The $\mathbf{n^{th}}$ **moment** of X, for $n \in \mathbb{Z}^+$, is defined as $\mathbb{E}[X^n]$.*

Lemma 2.17. *Variance has the following properties:*

(i) $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$.

(ii) $\mathrm{Var}(X + b) = \mathrm{Var}(X)$.

Example 2.18. Here are a few examples of important types of random variables.

- A random variable X is **normal** with mean μ and variance σ^2, if it has density function f_X, such that $f_X(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(X-\mu)^2}{2\sigma^2}}$. Denoted: $X \sim \mathcal{N}(\mu, \sigma^2)$. The density function is also known as **Gaussian**.
- A random variable is **exponential** with parameter λ given it has a density function f_X, such that $f_X(t) = \lambda e^{-\lambda t}$, for $t > 0$, $\lambda > 0$. The notation $X \sim \mathcal{E}(\lambda)$ means that the random variable X has exponential distribution with parameter λ.
- A random variable X has **Poisson** distribution if it takes values in the set $\{0, 1, 2, \dots\}$ with mass function $g(k) = \frac{\lambda^k}{k!} e^{-\lambda}$, for $k = 0, 1, 2, \dots$ and $\lambda > 0$.

For the remainder of the text, **I will assume random variables are absolutely continuous, and thus have densities.**

3 Characteristic functions

Definition 3.1. *Let X be a random variable. The* ***characteristic function*** *of X is the function $\varphi_X : \mathbb{R} \to \mathbb{C}$, defined as $\varphi_X(t) = \mathbb{E}[e^{itX}]$.*

For fixed a, $e^{ita} = \cos(ta) + i\sin(ta)$. Note that this is essentially the unit circle, with the cosine component in the real direction and the sine component in the imaginary direction. Also, e^{ita} has a period of $\frac{2\pi}{a}$ in t. Refer to Appendix A for more on complex numbers, $\mathbb{C}$.

Remark 3.2. If X has a density f_X, then $\varphi_X(t) = \int_{\mathbb{R}} e^{iat} f_X(a)da$. This is closely related to the Fourier transform of f, which is defined as $\widehat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iat} f_X(a)da$.

The characteristic function is special in that it maintains the distribution information of the random variable even after transformation. I will prove parts (ii) and (v) of the following lemma, see chapter 5 of [4] for the remaining proofs.

Lemma 3.3. *Properties of φ_X:*

(i) $\varphi_X(0) = 1$

(ii) $|\varphi_X(t)| \leq 1$

(iii) $\varphi_{aX}(t) = \varphi_X(at)$, $a \in \mathbb{R}$

(iv) $\varphi_X(t)$ *is continuous*

(v) $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$, *for independent* X, Y

Proof.
(ii): Notice that $e^{ia} = \cos(a) + i\sin(a)$. Then $\left|e^{ia}\right| = \sqrt{\cos^2(a) + \sin^2(a)} = \sqrt{1} = 1$. Hence,

$$\left|\varphi_X(t)\right| = \left|\mathbb{E}\left[e^{itX}\right]\right| \leq \mathbb{E}\left[\left|e^{itX}\right|\right] \leq 1\,.$$

(v): Consider two independent random variables, X and Y. Define: $\varphi_X(t) = \mathbb{E}[e^{itX}]$ and $\varphi_Y(t) = \mathbb{E}[e^{itY}]$. Then, we have

$$\varphi_{X+Y}(t) = \mathbb{E}[e^{it(X+Y)}] = \mathbb{E}[e^{itX}e^{itY}] = \mathbb{E}[e^{itX}]\mathbb{E}[e^{itY}] = \varphi_X(t)\varphi_Y(t)\,,$$

where the third equality holds by part (iii) of Lemma 2.15 since e^{itX} and e^{itY} are independent. □

Notice that the characteristic function is similar to the **moment generating function**, which is defined as $\phi(t) = \mathbb{E}[e^{tX}]$. The only difference between the moment generating function and the characteristic function is that the characteristic function has "i" in the exponent.

Example 3.4. Now I will compare the moment generating function and the characteristic function for a normal random variable. When $X \sim \mathcal{N}(0, \sigma^2)$, then

$$\phi_X(t) = \mathbb{E}[e^{tX}] = \int_{\mathbb{R}} e^{tx} \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx$$

By completing the square,

$$\begin{aligned} &= \int_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}(x-\sigma^2 t)^2 + \frac{1}{2\sigma^2}(\sigma^2 t)^2} \frac{1}{\sqrt{2\pi\sigma^2}} dx \\ &= e^{\frac{1}{2\sigma^2}(\sigma^2 t)^2} \int_{\mathbb{R}} \frac{e^{-\frac{1}{2\sigma^2}(x-\sigma^2 t)^2}}{\sqrt{2\pi\sigma^2}} dx \\ &= e^{\frac{\sigma^2 t^2}{2}}\,. \end{aligned} \tag{3.1}$$

Notice above that since the integrand is the probability density for a normal random variable, $\int_{\mathbb{R}} \frac{e^{-\frac{1}{2\sigma^2}(x-\sigma^2 t)^2}}{\sqrt{2\pi\sigma^2}} dx = 1$.

To find the characteristic function of a normal distribution, we can replace t by it in (3.1).

$$\varphi_X(t) = \mathbb{E}[e^{itX}] = e^{\frac{\sigma^2 (it)^2}{2}} = e^{-\frac{\sigma^2 t^2}{2}} \tag{3.2}$$

Notice that unlike the characteristic function, the moment generating function is not necessarily bounded. In fact, $e^{\frac{\sigma^2 t^2}{2}}$ grows faster than an exponential as $t \to \infty$. Because the

characteristic function has "i" as an exponent of the exponential function, the characteristic function of the normal distribution is always less than or equal to 1.

Example 3.5. Now I will look at the moment generating function and characteristic function for an exponential random variable.

When $X \sim \mathcal{E}(\lambda)$, then

$$\begin{aligned}\phi_Y(t) = \mathbb{E}[e^{tX}] &= \int_0^\infty e^{tx}\lambda e^{-\lambda x}dx \\ &= \lambda \int_0^\infty e^{(t-\lambda)x}dx \\ &= \lambda \frac{e^{(t-\lambda)x}}{t-\lambda}\Big|_{x=0}^{x=\infty} \\ &= \begin{cases} \infty & t \geq \lambda \\ \frac{\lambda}{\lambda - t} & t < \lambda \end{cases}\end{aligned} \tag{3.3}$$

Now to get the characteristic function we again replace t by it in (3.3) and get,

$$\varphi_Y(t) = \frac{\lambda}{\lambda - it}\,.$$

3.1 Derivative of the characteristic function

Lemma 3.6. *Suppose the random variable X satisfies $\mathbb{E}\big[|X|^k\big] < \infty$.*

(i) The ***k^{th} derivative*** *of $\varphi_X(t)$ has the form*

$$\frac{d^k}{dt^k}\varphi_X(t) = \mathbb{E}[(iX)^k e^{itX}] = \int_{\mathbb{R}} (ix)^k e^{itx} f_X(x)dx\,.$$

(ii) Then the derivatives of $\varphi_X(t)$ have the bound

$$\Big|\frac{d^k}{dt^k}\varphi_X(t)\Big| \leq \mathbb{E}[|X|^k] = \int_{\mathbb{R}} |x|^k f_X(x)dx\,.$$

Proof.
(i): Notice that $\dfrac{d^k}{dt^k}e^{itX} = (iX)^k e^{itX}$. Thus, $\dfrac{d^k}{dt^k}\varphi_X(t) = \mathbb{E}[(iX)^k e^{itX}]$.
By Lemma 2.15, it follows that $\mathbb{E}[(iX)^k e^{itX}] = \displaystyle\int_{\mathbb{R}} (ix)^k e^{itx} f_X(x)dx$.

(ii): With the result from part (i),

$$\begin{aligned}\Big|\frac{d^k}{dt^k}\varphi_X(t)\Big| =& \Big|\mathbb{E}[(iX)^k e^{itX}]\Big| \\ \leq& \mathbb{E}\Big[|(iX)^k e^{itX}|\Big] \\ =& \int_{\mathbb{R}} \Big|(ix)^k e^{itx} f_X(x)\Big| dx \\ =& \int_{\mathbb{R}} \Big|(ix)^k\Big|\Big|e^{itx}\Big|\Big|f_X(x)\Big| dx \\ =& \int_{\mathbb{R}} |x|^k f_X(x) dx\,. \end{aligned} \tag{3.4}$$

The last equality uses that $|e^{itx}| = 1$. □

Remark 3.7. When $t = 0$, then the $\mathbf{k^{th}}$ **derivative** from Lemma 3.6 becomes

$$\frac{d^k}{dt^k}\varphi_X(t)\Big|_{t=0} = i^k \mathbb{E}[X^k].$$

When evaluating at $t = 0$, the $\mathbf{k^{th}}$ **derivative** of the characteristic function is the product of i^k and the $\mathbf{k^{th}}$ **moment**. In particular if $\mathbb{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$, then we have the following:

First derivative:

$$\frac{d}{dt}\varphi_X(t)\Big|_{t=0} = \frac{d}{dt}\mathbb{E}[e^{itX}]\Big|_{t=0} = \mathbb{E}[iXe^{itX}]\Big|_{t=0} = \mathbb{E}[iX] = i\mathbb{E}[X] = i\mu$$

Second derivative:

$$\frac{d^2}{dt^2}\varphi_X(t)\Big|_{t=0} = i^2\mathbb{E}[X^2] = i^2(\sigma^2 + \mathbb{E}[X]^2) = -(\sigma^2 + \mu^2)$$

Notice in this case the second moment is rewritten in terms of the variance. Refer to Definition 2.16.

Definition 3.8. ***Taylor Polynomial:***

Let f be a function with defined derivatives $f', f'', \ldots, f^{(n)}$ at $s \in \mathbb{R}$. The n^{th}-order Taylor polynomial centered at s is

$$p_n(t) = \sum_{k=0}^{n} \frac{f^{(k)}(s)}{k!}(t-s)^k\,.$$

An important theorem from calculus [1] is the following:

Theorem 3.9 (Taylor's Theorem)**.** *Let $\mathcal{I}$ be an open interval in $\mathbb{R}$. Suppose $f : \mathcal{I} \to \mathbb{R}$ has n+1 continuous derivatives on $\mathcal{I}$, $s \in \mathcal{I}$. Then for $t \in \mathcal{I}$, define the* ***remainder term*** *$R_n(t)$ as*

$$R_n(t) = f(t) - p_n(t)\,.$$

Then $R_n(t)$ has a bound of the form

$$|R_n(t)| \leq M\frac{|t-s|^{n+1}}{(n+1)!}\,,$$

for constant $M = \max\limits_{s*c*t} |f^{(n+1)}(c)|$. Here the $s*c*t$ notation means that $s \leq c \leq t$ or $s \geq c \geq t$.

Taylor polynomials give an approximation of a function at a certain point. To determine how close this approximation is to the function, we use Taylor's Theorem to bound the remainder term.

Remark 3.10. In calculus, Taylor's Theorem is applied to functions $f : \mathcal{I} \to \mathbb{R}$, but here we are interested in complex valued functions, $g : \mathbb{R} \to \mathbb{C}$. Taylor's Theorem generalizes this situation by replacing absolute value with modulus around the remainder term. To do this, I identify $\mathbb{C}$ with $\mathbb{R}^2$, i.e.,

$$g(t) = \begin{pmatrix} g_1(t) \\ g_2(t) \end{pmatrix} \qquad \text{and} \qquad g^{(n)}(t) = \begin{pmatrix} g_1^{(n)}(t) \\ g_2^{(n)}(t) \end{pmatrix}\,.$$

The term $M = \max\limits_{s*c*t} |g^{(n+1)}(c)|$ is understood through

$$|g^{(n+1)}(t)| = \sqrt{\left(g_1^{(n+1)}(t)\right)^2 + \left(g_2^{(n+1)}(t)\right)^2}\,.$$

Example 3.11. Define $g(t) = e^{iat}$ for some $a \in \mathbb{R}$. Then the $\mathbf{n^{th}}$ **derivative** has the form $g^{(n)}(t) = (ia)^n e^{iat}$. Choose $n = 2$. Then Taylor's Theorem gives

$$R_3(t) = g(t) - \left(e^{ias} + \frac{(ia)e^{ias}}{1!}(t-s) + \frac{(ia)^2 e^{ias}}{2!}(t-s)^2\right),$$

where

$$\big|R_3(t)\big| \le \max_{s*c*t} \big|g^{(3)}(c)\big| \frac{|t-s|^3}{3!} = \max_{s*c*t} \big|(ia)^3 e^{iac}\big| \frac{|t-s|^3}{6} = \big|a\big|^3 \frac{|t-s|^3}{6}\,.$$

In particular, when $s = 0$,

$$R_3(t) = g(t) - \Big(1 + iat + \frac{(ia)^2}{2} t^2\Big)\,,$$

where $\big|R_3(t)\big| \le \big|a\big|^3 \frac{|t|^3}{6}$.

Proposition 3.12. *Suppose that* $\mathbb{E}\big[|X|^3\big] < \infty$*, then*

$$\Big|\varphi_X(t) - \big(1 + i\mu t - \frac{\sigma^2 + \mu^2}{2} t^2\big)\Big| \le \frac{|t|^3}{6} \mathbb{E}\big[|X|^3\big]\,.$$

Proof. By Taylor's theorem applied to φ_X, we have

$$\Big|\varphi_X(t) \;-\; \Big(\varphi_X(0) + \varphi_X'(0)t + \varphi_X''(0)\frac{t^2}{2}\Big)\Big| \;\le\; M\frac{|t|^3}{6}\,,$$

where $M = \max\limits_{|c|\le|t|} \big|\varphi_X'''(c)\big|$. Recall from Lemma 3.3 and Remark 3.7 that

$$\varphi_X(0) = 1, \quad \varphi_X'(0) = i\mu, \quad \varphi_X''(0) = -(\sigma^2 + \mu^2)\,.$$

Moreover, by Lemma 3.6,

$$M = \max_{|c|\le|t|} \big|\varphi_X'''(c)\big| \le \mathbb{E}\big[|X|^3\big]\,.$$

□

3.2 Central limit theorem

The **law of large numbers** (LLN) states that the average of a sequence of independent and identically distributed random variables, X_1, X_2, X_3,..., converges to the expectation of the random variable. That is, as $n \to \infty$

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \longrightarrow \mathbb{E}[X]\,.$$

Notice that although the X_i's are random, the normalized sum on the left above approaches a non-random, specific value as n becomes large.

The **central limit theorem** (CLT) characterizes the fluctuations of the difference of the average of the sequence of random variables and the expectation. Heuristically the CLT says that as $n \to \infty$:

$$\frac{X_1 + X_2 + \cdots + X_n}{n} - \mathbb{E}[X] \approx \mathcal{N}\Big(0, \frac{\sigma^2}{n}\Big),$$

where $\sigma^2 = \mathrm{Var}(X_i)$.

Notice the standard deviations are roughly proportional to $\frac{1}{\sqrt{n}}$, around the expectation.

Definition 3.13. *Let Y_n be a sequence of random variables. We say that Y_n* ***converges in law*** *to a random variable Y_∞ if the characteristic function of Y_n, φ_{Y_n}, converges to the characteristic function of Y_∞, φ_{Y_∞}, pointwise.* ***Pointwise*** *convergence means $\varphi_{Y_n}(t)$ converges to $\varphi_{Y_\infty}(t)$ for every $t \in \mathbb{R}$.*

Remark 3.14. Usually convergence in law has a more abstract definition, but the above is equivalent by the Continuity Theorem: see page 99 of [2].

4 Hierarchical systems of random variables

Given a function $f : \mathbb{R}^m \to \mathbb{R}$ and a sequence of independent and identically distributed (i.i.d.) random variables $X_1^{(0)}$, $X_2^{(0)}$, $X_3^{(0)}$, ... we define a hierarchical family of random variables $X_j^{(k)}$ such that for $j \in \mathbb{N}$ and $k \in \mathbb{N} \cup \{0\}$

$$X_j^{(k+1)} = f\big(X_{m(j-1)+1}^{(k)}, \cdots, X_{mj}^{(k)}\big)\,. \tag{4.1}$$

For each k there is a newly defined i.i.d. sequence of random variables, $\big(X_j^{(k)}\big)_{j\in\mathbb{N}}$, that depends entirely on the previous sequence labeled by $k-1$. The index k is called the **generation**. Given that the 0^{th} generation sequence is i.i.d., then all subsequent generation sequences $(X_j^{(k)})_{j\in\mathbb{N}}$ are also i.i.d.

Inductively we now know that a random variable at the $(n+1)^{th}$ generation is a combination of its ancestors from all preceding generations (0^{th} generation through n^{th} generation).

For the remainder of the section I will denote $\mu = \mathbb{E}[X_j^{(0)}]$ and $\sigma^2 = \mathrm{Var}(X_j^{(0)})$.

4.1 A random conductance model

Recall that the resulting conductance of two conductors with respective conductances $x_1, x_2 \geq 0$ connected in **series** can be calculated as

$$S(x_1, x_2) = \frac{1}{\frac{1}{x_1} + \frac{1}{x_1}} = \frac{x_1 x_2}{x_1 + x_2}\,.$$

Also recall that the resulting conductance of two conductors with respective conductances $y_1, y_2 \geq 0$ connected in **parallel** can be calculated as

$$P(y_1, y_2) = y_1 + y_2\,.$$

Thus if x_1, x_2, x_3, x_4 are the conductances of four bonds forming a diamond, the resulting conductance can be calculated as:

$$f(x_1, x_2, x_3, x_4) = \frac{x_1 x_2}{x_1 + x_2} + \frac{x_3 x_4}{x_3 + x_4}. \tag{4.2}$$

Now we consider a more complicated system in which we define a sequence of **diamond graphs** recursively. We begin with a single bond between two root vertices A, B (generation 0) and replace the bond by a diamond (generation 1), thus defining two paths between the root vertices, each comprised of two bonds. We then replace each bond in the diamond by a copy of a diamond to get the generation 2 graph. The diamond graphs are then defined inductively using this process of replacing bonds with diamonds.

We then assign independent random variables X_j, $1 \leq j \leq 4^n$ as the conductance of each bond on the n^{th} generation diamond. This model has been studied in mathematical physics literature [3, 5, 6, 7]. This corresponds to the general framework of hierarchical random variables, as seen in (4.1), in the special case of $m = 4$ and f of the form (4.2).

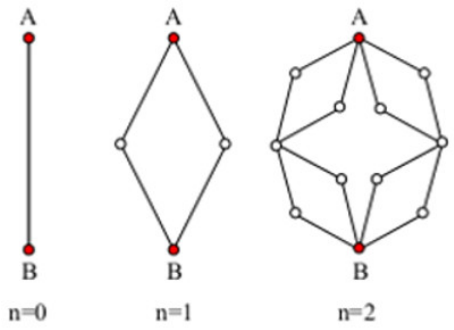

Figure 1: Generations 0, 1, 2 diamond graphs.

4.1.1 Linearization of the conductance model

My data in Section 5 shows that when the bonds are assigned random conductances that are uniformly distributed on the interval $(0, 1)$, the hierarchically defined random variables $X_j^{(n)}$ from (4.1) become highly concentrated around the value $\gamma \approx 0.35337$ as n grows. This suggests the following linear approximation of our function f from (4.2) when x_1, x_2, x_3,

x_4 are close to γ.

$$X_j^{(n+1)} = f(X_{4j-3}^{(n)}, X_{4j-2}^{(n)}, X_{4j-1}^{(n)}, X_{4j}^{(n)})$$
$$\approx f(\gamma,\gamma,\gamma,\gamma) + \sum_{i=1}^{4} \frac{\partial f}{\partial x_i}(\gamma,\gamma,\gamma,\gamma)(X_{4(j-1)+i}^{(n)} - \gamma) \tag{4.3}$$

Notice that in (4.2), if all inputs are equal, then $f(\gamma,\gamma,\gamma,\gamma) = \gamma$ and where, for instance,

$$\frac{\partial f}{\partial x_1}(x_1,x_2,x_3,x_4) = \frac{x_2(x_1+x_2) - x_1x_2}{(x_1+x_2)^2} = \frac{(x_2)^2}{(x_1+x_2)^2}\,,$$

and so, $\dfrac{\partial f}{\partial x_1}(\gamma,\gamma,\gamma,\gamma) = \dfrac{1}{4}$. By the analogous computation, $\dfrac{\partial f}{\partial x_i}(\gamma,\gamma,\gamma,\gamma) = \dfrac{1}{4}$ for $i \in \{2,3,4\}$.

Now we can simplify (4.3) to

$$X_j^{(n+1)} \approx \gamma + \frac{1}{4}\sum_{i=1}^{4}(X_{4(j-1)+i}^{(n)} - \gamma)\,. \tag{4.4}$$

Define $\epsilon_j^{(n)} = X_j^{(n)} - \gamma$. In other words $\epsilon_j^{(n)}$ is the deviation of $X_j^{(n)}$ from γ.

Now we can rewrite (4.4) as

$$\epsilon_j^{(n+1)} \approx \frac{1}{4}\sum_{i=1}^{4}\epsilon_{4(j-1)+i}^{(n)}\,. \tag{4.5}$$

Notice that the above has the form of (4.1) with $f(x_1,x_2,x_3,x_4) = \frac{1}{4}(x_1+x_2+x_3+x_4)$. This is a linear function with no constant term. In the next section we study hierarchical systems of this type in generality.

4.2 Central limit theorem problem in the linear case

In the remainder of this section I will focus on the special case in which f is a linear function with coefficients α_j, $1 \le j \le m$:

$$f(x_1,x_2,\cdots,x_m) = \sum_{k=1}^{m}\alpha_k x_k = \alpha_1x_1 + \alpha_2x_2 + \cdots + \alpha_m x_m\,.$$

Therefore a given random variable at the $(n+1)^{th}$ generation is a linear combination of its ancestors from the n^{th} generation:

$$X_j^{(n+1)} = \alpha_1 X_{m(j-1)+1}^{(n)} + \cdots + \alpha_m X_{mj}^{(n)} \,. \tag{4.6}$$

Here are two examples of random variables from the 1^{st} generation:

$$X_1^{(1)} = \alpha_1 X_1^{(0)} + \cdots + \alpha_m X_m^{(0)} \,,$$
$$X_2^{(1)} = \alpha_1 X_{m+1}^{(0)} + \cdots + \alpha_m X_{2m}^{(0)} \,.$$

By applying this inductively, we reach an equation of the form

$$X_1^{(n)} = \sum_{j=1}^{m^n} c_j^{(n)} X_j^{(0)} \,, \tag{4.7}$$

where each $c_j^{(n)}$ is some n-fold product of elements from the set $S = \{\alpha_1, \ldots, \alpha_m\}$. The exact form of the dependence of $c_j^{(n)}$'s on j will not matter in the proof of the CLT.

4.2.1 Solving for the mean

Recall $\mu = \mathbb{E}[X_j^{(0)}]$.

Lemma 4.1. *Define* $\mu_n = \mathbb{E}[X_j^{(n)}]$. *Then*

$$\mu_n = \mu(\alpha_1 + \alpha_2 + \cdots + \alpha_m)^n \,. \tag{4.8}$$

Proof. Notice that the lemma is true for $n = 0$. Now suppose that for induction, (4.8) holds

for n.

$$\begin{aligned}
\mu_{n+1} &= \mathbb{E}[X^{(n+1)}] \\
&= \mathbb{E}[\alpha_1 X_1^{(n)} + \alpha_2 X_2^{(n)} + \cdots + \alpha_m X_m^{(n)}] \\
&= \alpha_1 \mathbb{E}[X_1^{(n)}] + \alpha_2 \mathbb{E}[X_2^{(n)}] + \cdots + \alpha_m \mathbb{E}[X_m^{(n)}] \\
&= \mu_n(\alpha_1 + \alpha_2 + \cdots + \alpha_m) \\
&= \mu(\alpha_1 + \alpha_2 + \cdots + \alpha_m)^{n+1} .
\end{aligned}$$

□

Notice that by taking the expectation of (4.7) we have a second representation of μ_n as follows:

$$\begin{aligned}
\mu_n &= \mathbb{E}\Big[\sum_{j=1}^{m^n} c_j^{(n)} X_j^{(0)}\Big] \\
&= \sum_{j=1}^{m^n} \mathbb{E}[c_j^{(n)} X_j^{(0)}] \\
&= \sum_{j=1}^{m^n} c_j^{(n)} \mathbb{E}[X_j^{(0)}] \\
&= \mu \sum_{j=1}^{m^n} c_j^{(n)} .
\end{aligned} \tag{4.9}$$

The above will become useful later.

4.2.2 Solving for the variance

Recall $\sigma^2 = \mathrm{Var}(X_j^{(0)})$.

Lemma 4.2. *Define* $\sigma_n^2 = \mathrm{Var}(X_j^{(n)})$. *Then*

$$\sigma_n^2 = \sigma^2(\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_m^2)^n . \tag{4.10}$$

Proof. Again I will use induction. Notice that the lemma is true for $n = 0$ and suppose (4.10) holds for n.

$$\begin{aligned}
\sigma_{n+1}^2 &= \mathrm{Var}(X^{(n+1)}) \\
&= \mathrm{Var}(\alpha_1 X_1^{(n)} + \alpha_2 X_2^{(n)} + \cdots + \alpha_m X_m^{(n)}) \\
&= \alpha_1^2 \mathrm{Var}(X_1^{(n)}) + \alpha_2^2 \mathrm{Var}(X_2^{(n)}) + \cdots + \alpha_m^2 \mathrm{Var}(X_m^{(n)}) \\
&= \sigma_n^2(\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_m^2) \\
&= \sigma^2(\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_m^2)^{n+1} . \qquad (4.11)
\end{aligned}$$

□

By (4.7)

$$\begin{aligned}
\sigma_n^2 &= \mathrm{Var}\bigg(\sum_{j=1}^{m^n} c_j^{(n)} X_j^{(0)} \bigg) \\
&= \sum_{j=1}^{m^n} \mathrm{Var}\bigg(c_j^{(n)} X_j^{(0)} \bigg) \\
&= \sum_{j=1}^{m^n} \big(c_j^{(n)}\big)^2 \mathrm{Var}\bigg(X_j^{(0)} \bigg) \\
&= \sigma^2 \sum_{j=1}^{m^n} \big(c_j^{(n)}\big)^2 . \qquad (4.12)
\end{aligned}$$

Define

$$\widehat{c}_j^{(n)} = \frac{c_j^{(n)}}{\sqrt{(\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_m^2)^n}} = \frac{\sigma c_j^{(n)}}{\sigma_n} . \qquad (4.13)$$

Notice the second equality holds from (4.10). Also it is important to note that combining (4.12) and (4.13) gives us

$$\sum_{j=1}^{m^n} \big(\widehat{c}_j^{(n)}\big)^2 = 1 . \qquad (4.14)$$

Lemma 4.3. *Define* $0 < R < 1$ *as*

$$R = \max_{1 \le j \le m} \frac{\alpha_j}{\sqrt{\alpha_1^2 + \cdots + \alpha_m^2}} .$$

Then,

$$\max_{1\le j\le m^n} \widehat{c}_j^{(n)} \le R^n\,.$$

In particular, $\lim_{n\to\infty} \max_{1\le j\le m^n} \widehat{c}_j^{(n)} = 0$.

Proof. Recall that $c_j^{(n)}$ (from (4.7)) is an n-fold product of elements from $S = \{\alpha_1, \cdots, \alpha_m\}$. Thus, $\widehat{c}_j^{(n)}$ is an n-fold product of numbers of the form

$$\frac{\alpha_j}{\sqrt{\alpha_1^2 + \cdots + \alpha_m^2}}\,.$$

The above is $\le R$. □

4.2.3 Proving the CLT

The following lemma from page 112 of [2] gives a bound for the difference of two n-fold products of complex numbers.

Lemma 4.4. *Let $z_1, \cdots, z_n$ and $w_1, \cdots, w_n$ be complex numbers with modulus $\le \kappa$. Then*

$$\Big|\prod_{m=1}^n z_m \,-\, \prod_{m=1}^n w_m\Big| \le \kappa^{n-1} \sum_{m=1}^n \big|z_m - w_m\big|\,.$$

Proof. We can rewrite the difference of two products as

$$\Big|\prod_{m=1}^n z_m \,-\, \prod_{m=1}^n w_m\Big| = \Big|z_1 \prod_{m=2}^n z_m \,-\, z_1 \prod_{m=2}^n w_m \,+\, z_1 \prod_{m=2}^n w_m \,-\, w_1 \prod_{m=2}^n w_m\Big|\,.$$

Applying the triangle inequality to the above gives us:

$$\begin{aligned}
&\le \Big|z_1 \prod_{m=2}^n z_m \,-\, z_1 \prod_{m=2}^n w_m\Big| \,+\, \Big|z_1 \prod_{m=2}^n w_m \,-\, w_1 \prod_{m=2}^n w_m\Big| \\
&= \big|z_1\big|\Big|\prod_{m=2}^n z_m \,-\, \prod_{m=2}^n w_m\Big| \,+\, \Big(\prod_{m=2}^n \big|w_m\big|\Big)\big|z_1 \,-\, w_1\big|
\end{aligned}$$

By our assumption that the z_i's and w_i's have modulus $\le \kappa$,

$$\le \kappa\Big|\prod_{m=2}^n z_m \,-\, \prod_{m=2}^n w_m\Big| \,+\, \kappa^{n-1}\big|z_1 \,-\, w_1\big|\,.$$

By using the above inductively the lemma holds. □

The following lemma applies Taylor's Theorem using a first degree Taylor's polynomial of the exponential function.

Lemma 4.5. *For $x \geq 0$, we have*

$$|1 - x - e^{-x}| \leq \frac{x^2}{2}\,.$$

Proof. Define $f(x) = e^{-x}$. Then $f(0) = 1$, $f'(0) = -1$, and $f''(x) = e^{-x}$. By Taylor's Theorem 3.9 we have

$$|f(0) + f'(0)x - f(x)| \leq \frac{x^2}{2} \max_{0 \leq c \leq x} |f''(c)| = \frac{x^2}{2}\,.$$

Notice f'' is decreasing, and thus it is maximized at 0 where $f''(0) = 1$. □

Theorem 4.6. *Recall that $\mu_n = \mu(\alpha_1 + \alpha_2 + \cdots + \alpha_m)^n$ and $\sigma_n^2 = \sigma^2(\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_m^2)^n$. Also define the random variable*

$$Y_n = \frac{X_1^{(n)} - \mu_n}{\sigma_n}\,,$$

where $X_1^{(n)}$ is a generation n random variable. Then, as $n \to \infty$,

$$\varphi_{Y_n}(t) \quad \longrightarrow \quad e^{-\frac{t^2}{2}}\,.$$

In other words, the sequence Y_n converges in law to a normal distribution with mean 0 and variance 1.

Proof. Define

$$\widehat{X}_j = \frac{X_j^{(0)} - \mu}{\sigma}\,.$$

Notice that the $\widehat{X}_j$'s are i.i.d. with mean 0 and variance 1. The Y_n's can be written in in terms of $\widehat{X}_j$'s as follows:

$$Y_n = \frac{X_1^{(n)} - \mu_n}{\sigma_n}$$

Applying (4.7) and (4.9) gives us:

$$
\begin{aligned}
&= \frac{1}{\sigma_n}\left(\sum_{j=1}^{m^n} c_j^{(n)} X_j^{(0)} - \mu \sum_{j=1}^{m^n} c_j^{(n)}\right) \\
&= \frac{1}{\sigma_n}\sum_{j=1}^{m^n} c_j^{(n)}\left(X_j^{(0)} - \mu\right) \\
&= \frac{1}{\sigma_n}\sum_{j=1}^{m^n} c_j^{(n)}(\sigma \widehat{X}_j)
\end{aligned}
$$

Using (4.13),

$$
= \sum_{j=1}^{m^n} \widehat{c}_j^{(n)} \widehat{X}_j \,. \tag{4.15}
$$

Now inserting this Y_n into the characteristic function we get

$$
\varphi_{Y_n}(t) = \prod_{j=1}^{m^n} \varphi_{\widehat{c}_j^{(n)}\widehat{X}_j}(t) = \prod_{j=1}^{m^n} \varphi_{\widehat{X}_j}(\widehat{c}_j^{(n)} t) \,. \tag{4.16}
$$

Notice the equalities above hold from Lemma 3.3.

Now I will bound the difference of the characteristic function of Y_n with the characteristic function of a normal distribution, refer to (3.2).

$$
\left| \prod_{j=1}^{m^n} \varphi_{\widehat{X}_j}(\widehat{c}_j^{(n)} t) - e^{-\frac{t^2}{2}} \right|
$$

Recall that from (4.14) that $\sum_{j=1}^{m^n} |\widehat{c}_j^{(n)}|^2 = 1$,

$$
\begin{aligned}
&= \left| \prod_{j=1}^{m^n} \varphi_{\widehat{X}_j}(\widehat{c}_j^{(n)} t) - e^{-\frac{t^2}{2}\sum_{j=1}^{m^n}(\widehat{c}_j^{(n)})^2} \right| \\
&= \left| \prod_{j=1}^{m^n} \varphi_{\widehat{X}_j}(\widehat{c}_j^{(n)} t) - \prod_{j=1}^{m^n} e^{-\frac{t^2}{2}(\widehat{c}_j^{(n)})^2} \right|
\end{aligned}
$$

Applying Lemma 4.4 to the above difference gives the following. Notice in this case $\kappa = 1$.

$$\leq \sum_{j=1}^{m^n} \left| \varphi_{\widehat{X}_j}(\widehat{c}_j^{(n)} t) - e^{-\frac{t^2}{2}(\widehat{c}_j^{(n)})^2} \right|$$
$$= \sum_{j=1}^{m^n} \left| \varphi_{\widehat{X}_j}(\widehat{c}_j^{(n)} t) - (1 - \frac{(\widehat{c}_j^{(n)})^2 t^2}{2}) + (1 - \frac{(\widehat{c}_j^{(n)})^2 t^2}{2}) - e^{-\frac{t^2}{2}(\widehat{c}_j^{(n)})^2} \right|$$

Applying the triangle inequality gives us:

$$\leq \underbrace{\sum_{j=1}^{m^n} \left| \varphi_{\widehat{X}_j}(\widehat{c}_j^{(n)} t) - (1 - \frac{(\widehat{c}_j^{(n)})^2 t^2}{2}) \right|}_{\text{(I)}} + \underbrace{\sum_{j=1}^{m^n} \left| (1 - \frac{(\widehat{c}_j^{(n)})^2 t^2}{2}) - e^{-\frac{t^2}{2}(\widehat{c}_j^{(n)})^2} \right|}_{\text{(II)}} . \quad (4.17)$$

Now I will bound (I) and (II) individually.

By Proposition 3.12, I have the first inequality below.

$$\text{Expression (I)} \leq \sum_{j=1}^{m^n} \frac{|\widehat{c}_j^{(n)} t|^3}{6} \mathbb{E}[|\widehat{X}_j|^3]$$

Since $\widehat{X}_j$'s have mean 0 and variance 1,

$$= \mathbb{E}[|\widehat{X}_1|^3] \frac{|t|^3}{6} \sum_{j=1}^{m^n} |\widehat{c}_j^{(n)}|^3$$
$$\leq \mathbb{E}[|\widehat{X}_1|^3] \frac{|t|^3}{6} \Big(\max_{1 \leq j \leq m^n} \widehat{c}_j \Big) \sum_{j=1}^{m^n} |\widehat{c}_j^{(n)}|^2$$

Recall that $\sum_{j=1}^{m^n} |\widehat{c}_j^{(n)}|^2 = 1$, and thus,

$$= \mathbb{E}[|\widehat{X}_1|^3] \frac{|t|^3}{6} \max_{1 \leq j \leq m^n} \widehat{c}_j \, . \quad (4.18)$$

The above converges to 0 as $n \to \infty$ from Lemma 4.3.

By applying Lemma 4.5 to (II) with $x = \frac{t^2}{2}(\widehat{c}_j^{(n)})^2$,

$$\begin{aligned}\text{Expression (II)} &\leq \sum_{j=1}^{m^n} \frac{1}{2}\Big(\frac{t^2}{2}(\widehat{c}_j^{(n)})^2\Big)^2 \\ &= \frac{t^4}{8}\sum_{j=1}^{m^n}(\widehat{c}_j^{(n)})^4 \\ &\leq \frac{t^4}{8}\big(\max_{1\leq j\leq m^n}\widehat{c}_j^{(n)}\big)^2\sum_{j=1}^{m^n}(\widehat{c}_j^{(n)})^2\end{aligned}$$

Again recall $\sum_{j=1}^{m^n}|\widehat{c}_j^{(n)}|^2 = 1$, so

$$= \frac{t^4}{8}\big(\max_{1\leq j\leq m^n}\widehat{c}_j^{(n)}\big)^2. \tag{4.19}$$

The above converges to 0 by Lemma 4.3.

This completes the proof that $\varphi_{Y_n}(t)$ converges to $e^{-\frac{t^2}{2}}$ as $n \to \infty$. □

5 Data

I used the R program to run simulations on the following equation:

$$f(x_1, x_2, x_3, x_4) = \frac{x_1 x_2}{x_1 + x_2} + \frac{x_3 x_4}{x_3 + x_4}. \tag{5.1}$$

This equation calculates the conductance for the diamond graph. Notice that x_i, $i \in \{1, 2, 3, 4\}$, are uniformly distributed in the initial generation.

5.1 Program script

Example 5.1. The following is the script used for the generation 1 diamond:

```
Data=NULL
total=100000
    for(i in 1:total){
        X= runif(4, 0, 1)
        Data[i] = (X[1]*X[2]/ (X[1]+X[2])) + (X[3]*X[4]/ (X[3]+X[4]))
    }
library("rio")
temp=import("Data_1.dta")
finalData=c(unlist(temp),Data)
export(as.data.frame(finalData), "Data_1.dta")
```

Initially, an empty vector is declared, called Data. Then, a for-loop is run 100000 times. Each iteration of the for-loop generates four uniformly distributed random numbers between $(0, 1)$ and places them into vector X. These four numbers are then plugged into (5.1), and the output is placed into the Data vector. After all iterations of the for-loop have been completed, the values in the Data vector are then cumulated with old data from previous generation 1 for-loops and saved onto the hard drive as "Data_1.dta". See Figure 4.1 for the generation 1 diamond.

Example 5.2. The following is the script used for the generation 2 diamond:

```
Data=NULL
total=100000
    for(i in 1:total){
        A=NULL
        for(j in 1:4){
            X= runif(4, 0, 1)
            A[j] = (X[1]*X[2]/ (X[1]+X[2])) + (X[3]*X[4]/ (X[3]+X[4]))
        }
        Data[i] = (A[1]*A[2]/ (A[1]+A[2])) + (A[3]*A[4]/ (A[3]+A[4]))
    }
library("rio")
temp=import("Data_2.dta")
finalData=c(unlist(temp),Data)
export(as.data.frame(finalData), "Data_2.dta")
```

This script is similar to that of the generation 1 diamond, except the script for the generation 2 diamond has an additional nested for-loop. This additional for-loop represents the next level in the hierarchy. For each iteration of the outer for-loop, random numbers are generated and used in (5.1). Four of these outputs are stored in vector A. Then these four values in A are plugged into (5.1) and the output is stored in Data. Vector A is then cleared and the outer for-loop moves on to the next iteration. After all iterations, Data is combined with old data and exported to the hard drive. See Figure 4.1 for the generation 2 diamond.

Notice that vector A is a dummy vector, as it does not hold any of the final, finished data. It temporarily holds values until they can be used by the equation and stored in Data, and then A is erased. Data is the final, important vector.

This process of adding an extra for-loop and an extra dummy vector continues recursively with each additional level in the hierarchy.

5.2 Results

5.2.1 Raw Results

Using the pattern from the scripts above I gathered data for 17 generations in the hierarchy. I will look at the distribution of the results in terms of mean and variance. Also in the table is n, the number of data points in each generation.

Table 2: Raw results

	generation 1	generation 2	generation 3	generation 4
mean:	0.4088753	0.3709838	0.3581723	0.3545039
variance:	$2.990241e-02$	$9.275698e-03$	$2.594975e-03$	$6.739738e-04$
n:	400000	300000	250000	200000

	generation 5	generation 6	generation 7	generation 8
mean:	0.3536934	0.3534596	0.3533958	0.3533815
variance:	$1.709191e-04$	$4.308603e-05$	$1.076589e-05$	$2.685877e-06$
n:	170000	155000	145000	135000

	generation 9	generation 10	generation 11	generation 12
mean:	0.3533804	0.3533750	0.3533750	0.3533744
variance:	$6.695290e-07$	$1.671734e-07$	$4.323928e-08$	$1.093578e-08$
n:	76000	44000	7001	91

	generation 13	generation 14	generation 15	generation 16
mean:	0.3533753	0.3533760	0.3533705	0.3533764
variance:	$2.130014e-09$	$6.473410e-10$	$2.737788e-10$	$5.015538e-11$
n:	51	51	20	4

	generation 17
mean:	0.3533728
variance:	N/A
n:	1

With each additional generation (and thereby, additional for-loop), the time required to complete the program increased approximately by a factor of 4, and thus exponentially. Therefore the number of data points at each generation is less than (or equal to) the number from the previous generation. In particular, the final generation took more than 4 days to complete one iteration of the outermost for-loop, so it only has one data point. Hence the variance is not applicable for generation 17 and thus it is not reliable data.

Here are the two main points from the raw data. As the generation increases,

- the mean seems to approach some number, γ, of the form 0.35337...
- the variance decays to 0, which tells us that the data becomes increasingly concentrated around the mean.

5.2.2 Ratio Results

Often it is helpful to view results in terms of ratios of change. I will examine mean change ratio and variance ratio.

Mean change ratio (MCR) and variance ratio (VR) are defined as follows:

$$MCR[i] = (Mean[i+2] - Mean[i+1])/(Mean[i+1] - Mean[i]), \text{ for } i \in [1, 15]\,,$$

$$VR[i] = Var[i+1]/Var[i], \text{ for } i \in [1, 15]\,,$$

where the notations "$Mean[i]$" and "$Var[i]$" indicate the mean and variance values, respectively, of generation i, as seen in the table above.

Table 3: Ratio results

i	1	2	3	4
$MCR[i]$:	0.338109794	0.286332901	0.220952254	0.288420300
$VR[i]$:	0.3101991	0.2797606	0.2597227	0.2535990

i	5	6	7	8
$MCR[i]$:	0.273212036	0.222489089	0.083806911	4.521953620
$VR[i]$:	0.2520844	0.2498696	0.2494803	0.2492776

i	9	10	11	12
$MCR[i]$:	-0.007898019	-14.264100816	-1.423505017	0.883508023
$VR[i]$:	0.2496880	0.2586493	0.2529132	0.1947746

i	13	14	15
$MCR[i]$:	-7.311364716	-1.073294841	-0.600977758
$VR[i]$:	0.3039141	0.4229282	0.1831968

The mean change ratio did not converge to a number as I expected it would. However, the variance ratio shows us that with each generation the variance is roughly decreased by a factor of 4. The overall decay of the variance to 0 can be understood as a law of large numbers, since each new generation includes 4 times as many bonds as the previous generation. The exponential decay of the variance is predictable from the linearized model which had variance σ_n^2 at generation n given by

$$\sigma_n^2 = \sigma^2(\alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_4^2)^n = \sigma^2\Big(\frac{1}{4}\Big)^n,$$

where $\alpha_i = \frac{1}{4}$ for $i \in \{1, 2, 3, 4\}$.

5.3 Concluding remarks on the data

In these simulations I considered a randomized model for the conductance of an electrical system comprised of a complicated network of bonds connecting two nodes. The randomness in the model is generated by positive random variables (uniformly distributed on the interval $(0, 1)$) associated with each bond. The graphical structure of the system implies that the net conductance can be deduced using recursive equations based on the rules of conductance (i.e. in series, in parallel). This forms a statistical mechanical model for conductance in which the bonds are viewed on the microscopic level.

From the randomness of each bond on the microscopic level it follows that the net conductance is also inherently random. However, from the data we see a law of large numbers on display, in which the conductance of the total system seems to converge to a fixed value, $\gamma \approx 0.35337$. We also see that the variance of the net conductance exponentially converges to 0. In other terms, at higher generations the net conductance values fall very close to the mean. **Thus, at the macroscopic level the randomness is essentially eliminated.**

A difficulty in doing simulations on the model comes from the exponential growth in complexity as the generation grows. Specifically the time required for the generation 17 diamond was more than four days of undisrupted simulation.

A Complex

Complex numbers are an extension of the real numbers. Complex numbers have the form $a + bi$, with $a, b \in \mathbb{R}$, where the number i is defined such that $i^2 = -1$. The set of complex numbers is denoted $\mathbb{C}$. We define the modulus of a complex number $z = a + bi$ as $|z| = |a + bi| = \sqrt{a^2 + b^2}$. Complex numbers can also undergo the usual operations.
Let $a + bi$, $c + di \in \mathbb{C}$.

- $(a + bi) + (c + di) = (a + c) + (b + d)i$
- $(a + bi) - (c + di) = (a - c) + (b - d)i$
- $(a + bi) \times (c + di) = (ac - bd) + (ad + bc)i$
- $(a + bi)^{-1} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i$
- $(a + bi)/(c + di) = (a + bi) \times [(c + di)^{-1}]$, where $c + di \neq 0$

For a complex number z, we define e^z as

$$e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n \,.$$

When using real numbers, the Maclaurin series representation of the exponential function follows from a theorem. However, for complex numbers the series is used as a definition. The above series is absolutely convergent because,

$$\sum_{n=0}^{\infty} \left| \frac{1}{n!} z^n \right| = \sum_{n=0}^{\infty} \frac{1}{n!} |z|^n = e^{|z|} < \infty \,.$$

Theorem A.1. *We have the following properties for e^z:*

(i) For $z, w \in \mathbb{C}$,

$$e^{z+w} = e^z e^w \,.$$

(ii) For $x \in \mathbb{R}$,

$$e^{ix} = \cos(x) + i \sin(x) \,.$$

In particular, $|e^{ix}| = 1$.

References

[1] W. Briggs, L. Cochran: *Calculus: early transcendentals*, 1st ed. (Pearson Education, 2011).

[2] R. Durrett: *Probability theory and examples*, 2nd ed. (Duxbury Press, 1996).

[3] L. Goldstein: *Normal approximation for hierarchical structures*, Ann. Appl. Probab. **14**, No. 4, 1950-1969 (2004).

[4] G.R. Grimmett, D.R. Stirzaker: *Probability and random processes*, 3rd ed. (Oxford University Press, 2001).

[5] B. Hambly, J. Jordan: *A random hierarchical lattice: the series-parallel graph and its properties*, Adv. Appl. Probab., No. 3, 824-838 (2004).

[6] J. Jordan: *Almost sure convergence for iterated functions of independent random variables*, J. Appl. Probab. **12**, 985-1000 (2002).

[7] J. Wehr, J.M. Woo: *Central limit theorems for nonlinear hierarchical sequences of random variables*, J. Statist. Phys. **104**, 777-797 (2001).