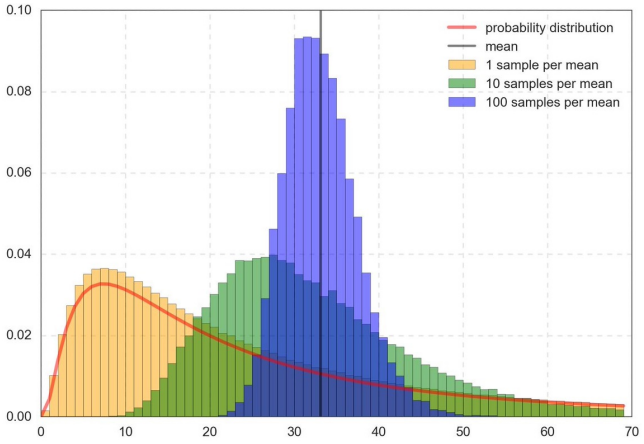
Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Science



(Galea, 2016)

The central limit theorem in a random environment

Supervisor:
Prof. Dr. F.H.J. Redig

Thesis written by:
M.A. van Kempen (4655583)

Thesis on behalf of the
Delft Institute of Applied Mathematics
as part to be obtained the degree of

BACHELOR OF SCIENCE
in
APPLIED MATHEMATICS

Delft, The Netherlands
26th of July 2022

Contents

Abstract **2**

1 Introduction **3**

2 Context **5**
2.1 What is a Markov chain 5
2.2 Transition measure and transition operator 5
2.3 Stationary Markov chains 7
2.4 Reversible measure 8
2.5 Ergodic Markov chains 9

3 The central limit theorem **10**
3.1 The classical central limit theorem 10
3.2 The Kipnis-Varadhan central limit theorem 11

4 The central limit theorem for additive functions **12**
4.1 Martingale central limit theorem 12
4.2 Computation of limiting variance 13
4.3 A simple example with a finite Markov chain 15

5 Simulations of the random walk in a random environment **17**
5.1 The random environment as a line of numbers 17
5.2 The random walk in a random environment in 2D 19

6 Random walk in random environment in continuous time **24**
6.1 Defining the process 24
6.2 The semigroup and generator 25
6.3 Stationarity, reversibility and ergodicity 26
6.4 The Kipnis-Varadhan central limit theorem in continuous time . 26
6.5 The random conductance model 28
6.6 Reversibility and ergodicity of the environment process 29
6.7 The martingale difference sequence 31
6.8 Variance requirement 32

7 Simulations for the random conductance model **33**

8 Summary **35**

References **36**

Appendix **37**

Abstract

In this paper the Kipnis-Varadhan central limit theorem for additive functions acting on a stationary reversible and ergodic Markov chain in a random environment will be proven in discrete and continuous time. After proving the theorem for the discrete case it is also put to the test by simulating the process for an increasing amount of times. In this way we show that the outcomes of the additive functions do indeed show normality. Next, the Kipnis-Varadhan central limit theorem will be proven for the continuous case. After this the random conductance model will be explained. The central limit theorem will be illustrated a second time by doing simulations of the random conductance model and showing that again the outcomes show a clear normal distribution, thus proving the central limit theorem for the random conductance model.

1 Introduction

It is impossible to imagine modern probability theory without the central limit theorem. The first version of the theorem was introduced by French mathematician Abraham De Moivre in the second version of his textbook *The Doctrine of Chances* which was published in 1738. In the book De Moivre talks about the amount of times a coin shows heads in a certain amount of coin tosses. De Moivre came to the conclusion that with sufficient tosses the Bernoulli distribution could be approximated by the normal distribution. To be precise, the conclusion was that the probability mass function of n independent Bernoulli trails that have a succes with probability p converges to the probability density function of a normal distribution with mean np and variance $np(1-p)$ as n increases (Moivre, 1738).

At present time the central limit theorem has many different versions. One of the versions is the central limit theorem was formulated and proven by C. Kipnis and S.R.S. Varadhan. In their paper *Central Limit Theorem for Additive Functionals of Reversible Markov Processes and Applications to Simple Exclusions*, which was published in 1986, they formulate and prove a version of the central limit theorem for additive functions in a random environment. In this paper we will take an in depth look at this paper.

Firstly, in Section 2 we will start by giving some context. In this section the working of a Markov chain will be explained as well as some key conditions our Markov chain will have to obey in order for the central limit theorem to work.

Secondly, in Section 3 the classical central limit theorem will be repeated in order to refresh the readers knowledge, and, more importantly, in order to show why this theorem will not work for the Markov chain which will be used in this paper.

Thirdly, in Section 4 we will use the Martingale central limit theorem in order to prove the Kipnis-Varadhan central limit theorem for Markov chains in discrete time in a random environment. This is followed by computation to derive an equation for the variance of the normal distribution to which the process converges to by the Kipnis-Varadhan central limit theorem.

Fourthly, in Section 5 simulations will be made in order to illustrate how the Markov chain in a random environment works. These simulations show that by increasing the amount of processes a clear convergence to the normal distribution can be seen. The simulations will be done for a situation in one-dimension and a situation in two-dimensions.

Fifthly, in Section 6 it will be shown that the Markov chain follows the conditions needed for the Kipnis-Varadhan central limit theorem to apply. After this a new equation will be derived which is used to compute the variance of

the normal distribution to which the process converges.

Lastly, in Section 7 another set of simulations will be made. Here we show that the outcomes of the simulations also show that the outcomes of the processes converge to the normal distribution as the amount of processes increases.

2 Context

2.1 What is a Markov chain

A Markov chain is a stochastic process which moves from state to state according to some transition probabilities. The different states in the Markov process can be represented as $X_0, X_1, \ldots, X_n \in \Omega$, where Ω is the state space. The Markov chain then tells us that if we start at state X_0, the probability distribution defined on the states gives us the probability of moving from X_0 to another state. In this way the Markov chain defines a model for jumping across these states. What makes Markov chains different from other processes is that in a Markov chain the only relevant information is the last state that was achieved. To put this into equations, a chain is a Markov chain if it satisfies the following property

$$P(X_{j+1} = x_{j+1} | X_j = x_j, \ldots, X_1 = x_1, X_0 = x_0) = P(X_{j+1} = x_{j+1} | X_j = x_j)$$

An elementary example of a Markov chain is the simple random walk. The simple random walk is a process which starts at X_0 and has a certain probability to take a step to the left, and a certain probability to take a step to the right. Figure (1) is an example of such a simple random walk. Here the probability of taking a step to the right or left are both 0.5. The probability of landing on a state are beneath each red dot.

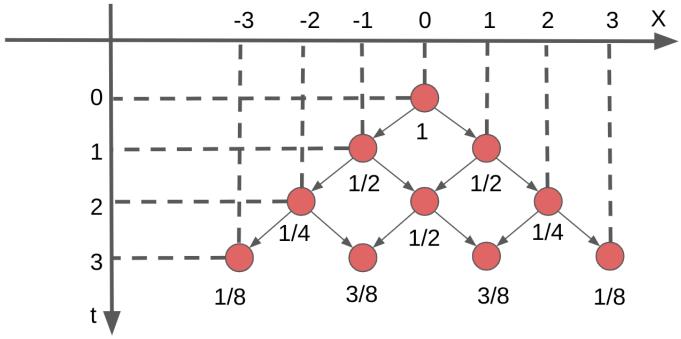


Figure 1: Random walk tree (Stiedemann, 2021).

Thus, the simple random walk moves from state to state, or red dot to red dot, with a probability distribution defined on each red dot. This is the most elementary example of a Markov chain.

2.2 Transition measure and transition operator

As said, the Markov chain is a process in which we move from state to state. In all cases of Markov chains the transition measure is the probability that we move

from one state to another. To start, let us first consider a countable state space. In this way we have states $X_j \in \Omega$ where $\Omega = \{1, 2, \dots\}$. Then the transition probability of moving from $X_j = x$ to $X_{j+1} = y$ is $P_{x,y} = P(X_{j+1} = y | X_j = x)$. In the case of a finite Markov chain we can write all the transition probabilities in a matrix P, which is called the transition matrix.

In the following Markov chain it is shown how to move from a Markov chain to a transition matrix. In figure (2) we have a Markov chain for which the probabilities of moving from state to state can be seen. For example, the probability of moving from state 1 to state 2 is $P(X_{j+1} = 2 | X_j = 1) = 0.6$, moving from state 2 to state 2 happens with probability $P(X_{j+1} = 2 | X_j = 2) = 0.4$. In order to make the transition matrix all the different probabilities have to be inserted in said matrix. This results in the following matrix.

$$
\begin{aligned}
P &= \begin{bmatrix}
P(X_{j+1} = 1 | X_j = 1) & P(X_{j+1} = 2 | X_j = 1) & \cdots & P(X_{j+1} = 4 | X_j = 1) \\
P(X_{j+1} = 1 | X_j = 2) & P(X_{j+1} = 2 | X_j = 2) & \cdots & \cdots \\
P(X_{j+1} = 1 | X_j = 3) & \cdots & \cdots & \cdots \\
P(X_{j+1} = 1 | X_j = 4) & \cdots & \cdots & \cdots
\end{bmatrix} \\
&= \begin{bmatrix}
0.4 & 0.6 & 0 & 0 \\
0.6 & 0.4 & 0 & 0 \\
0.25 & 0.25 & 0.25 & 0.25 \\
0 & 0 & 0 & 1
\end{bmatrix}
\end{aligned}
$$

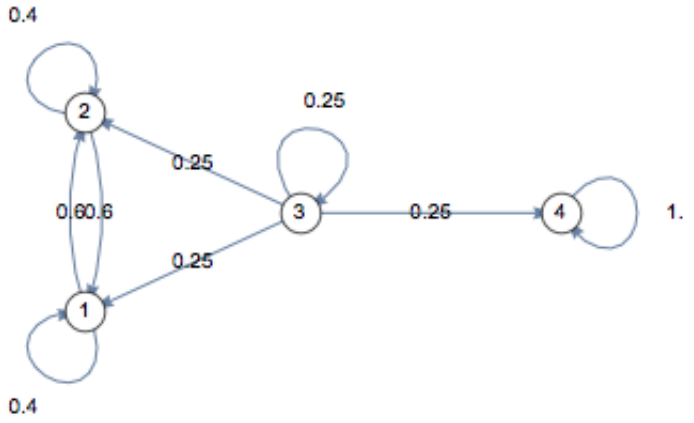


Figure 2: An example of a Markov Chain.

This example only shows a Markov chain with a finite state space. In this thesis we will have to deal with a Markov chain on a general state space. For

these general state spaces it is not possible to write the transition matrix. This is because a matrix can at most take a countable amount of entries. Because of this we need to define a measure which can define a different transition probability $q(x, dy)$. We define $q(x, dy)$ in the following way,

$$q(x, dy) = P(X_{j+1} \in dy | X_j = x)$$

Thus, given $q(x, dy)$ it is possible to calculate the probability that the Markov process jumps from a state $X_j = x$ into a subset $dy \in \Omega$ where Ω is again the state space. It can be calculated in the following way.

$$P(X_{j+1} \in A | X_j = x) = \int_A q(x, dy)$$

Now that we have defined the transition measure, we can look at the transition operator. The transition operator can be defined as the expected value of $f(X_{j+1})$ given that we move from $X_j = x$. To put this into equations we see that the transition operator working on a function f, which is bounded and measurable, is the following

$$qf(x) = E[f(X_{j+1}) | X_j = x] = \int f(y) q(x, dy)$$

The transition measure has two basic properties. The first one is that if $f \geq 0$, then $qf \geq 0$. This can be seen from the definition of the measure. The second one is that the operator is a contraction. By using $q1 = 1$ we can show

$$||qf||_\infty = |\int f(y) q(x, dy)| \leq \int |f(y)| q(x, dy) \leq ||f||_\infty \int q(x, dy) = ||f||_\infty \cdot 1$$

Which shows $||qf||_\infty \leq ||f||_\infty$, confirming that the operator is a contraction.

2.3 Stationary Markov chains

A Markov chain is stationary if the distribution stays the same when moving along the states of the chain. Or in other words, if we have a distribution μ at X_j, then the distribution at X_{j+1} will also be μ. Stationarity is defined in the following way

Definition 1. *The probability measure μ is stationary if $\mu P = \mu$ for discrete Markov chains or equivalently $\int qf d\mu = \int f d\mu$ for continuous Markov chains for all bounded and measurable functions f.*

In this definition P is the transition matrix, and q is the transition probability discussed previously.

2.4 Reversible measure

Reversible Markov chains are processes in which we can go back and forth between the two states. For a general state space reversibility is defined in the following way

Definition 2. *μ is reversible if and only if q is self-adjoint in $L^2(\mu)$. Where q is self-adjoint in $L^2(\mu)$ if and only if $\int f(qg)d\mu = \int (qf)gd\mu$ for all bounded and measurable functions f and g.*

From the definition it is also possible to show that reversibility implies stationarity. This can be seen by choosing $g(x) = 1$. From this it follows that indeed $\int qfd\mu = \int fd\mu$ as shown here

$$\langle qf, 1\rangle = \int (qf)\cdot 1 d\mu = \int f(q\cdot 1)d\mu = \langle f, q\cdot 1\rangle =^* \langle f, 1\rangle = \int f d\mu$$

Where at * we use the fact that $q \cdot 1 = 1$.
For a reversible Markov chain with a finite state space we are able to formulate the following proposition.

Proposition 1. *A Markov chain with a finite state space is reversible if and only if it satisfies the detailed balance condition.*

Proof. "$\Rightarrow$" Denote the transition matrix for moving between x and y as $P_{x,y}$ and $P_{y,x}$. Then to show reversibility we must show that we have

$$\mu(x)P_{x,y} = \mu(y)P_{y,x}$$

This can be shown by starting to define $qf(x) = \sum_y P_{x,y}f(y)$. Since we know that the Markov chain is reversible we have that $\langle qf, g\rangle_{L^2(\mu)} = \langle f, qg\rangle_{L^2(\mu)}$, which equals to $\sum_x (qf)(x)g(x)\mu(x) = \sum_x f(x)(qg)(x)\mu(x)$ for all bounded and measurable functions f and g.
Now let us define f and g in the following way. $f(x) = \delta_{x,y}$ and $g(x) = \delta_{x,z}$ such that δ_{ij} is the Kronecker delta which says, $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$. Using this it can be seen that

$$qf(x) = \sum_{y'} P_{x,y'}f(y') = \sum_{y'} P_{x,y'}\delta_{y',y} = P_{x,y}$$

This gives

$$\sum_x (qf)(x)g(x)\mu(x) = \sum_x P_{x,y}\delta_{x,z}\mu(x) = P_{z,y}\mu(z)$$

In a similar way it can be shown that $\sum_x f(x)(qg)(x)\mu(x) = P_{y,z}\mu(y)$. Thus, it can be concluded that $\sum_x (qf)(x)g(x)\mu(x) = \sum f(x)(qg)(x)\mu(x)$ implies

$$P_{z,y}\mu(z) = P_{y,z}\mu(y)$$

Thus proving the detailed balance condition.
"$\Leftarrow$ " Given that the Markov chain satisfies the detailed balance condition we get that $P_{z,y}\mu(z) = P_{y,z}\mu(y)$ for all states y and z. Then by definition the Markov chain is reversible. □

2.5 Ergodic Markov chains

The formal definition of an ergodic probability measure is the following

Definition 3. *A probability measure μ is said to be ergodic if $qf = f$ implies that $f = \int f d\mu$ for μ almost surely.*

This means that if the probability measure is ergodic all states can be reached. All situations similar to figure (3) are excluded.

If we again take a look at figure (2) we can see that the Markov chain is not

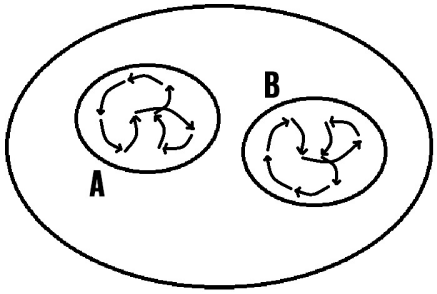


Figure 3: Situations which are excluded if the measure is ergodic.

ergodic. This is because if a jump is made to state 4, this state will never be left again. This makes the Markov chain not ergodic. If there would be an arrow from state 4 to another state the Markov chain would be ergodic.

3 The central limit theorem

3.1 The classical central limit theorem

If you are reading this you will have heard of the classical central limit theorem. This theorem states that if we have a sequence of random variables which are independent and identically distributed, say $\{Y_1, Y_2, \ldots, Y_n, \ldots\}$, with $E[Y_i] = \mu$ and $\mathrm{Var}[Y_i] = \sigma^2 < \infty$, and define $X_n = \sum_{i=1}^{n}(Y_i - \mu)$. We see that as n goes to ∞, the normalized sum $\frac{1}{\sqrt{n}}X_n$ converge to the normal distribution with mean 0 and variance σ^2, or $N(0, \sigma^2)$.
An example of the classical central limit theorem is portrait in the following figure. In this figure we can see what happens to the distribution of the outcome of a sum of dice throws and their difference with the expected value of 3.5.

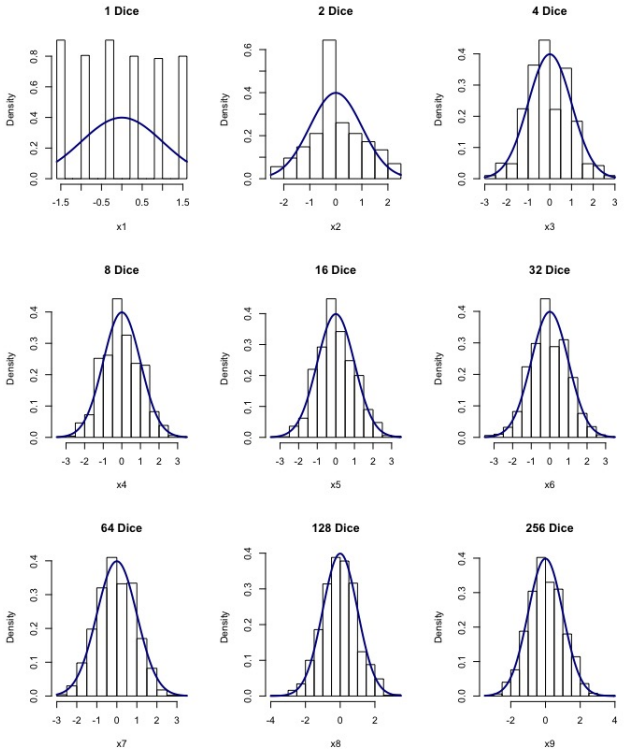


Figure 4: An illustration of the classical central limit theorem for the throw of a dice.

3.2 The Kipnis-Varadhan central limit theorem

In the paper the overall goal is to extend the classical central limit theorem to show that there is a generalization to the classical central limit theorem such that it also holds for additive functionals of a reversible Markov process. In other words, we want to show that if we define a function $V : \Omega \to \mathbb{R}$ defined on a Markov chain $\{X_i : n \in \mathbb{N}\}$, such that the Markov chain is reversible with μ as a reversible and ergodic probability measure. Then it is true that $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} V(X_i)$ converges to $N(0, \sigma^2)$ for some variance σ^2 (Kipnis & Varadhan, 1986).

One big difference between the classical central limit theorem and the Kipnis-Varadhan central limit theorem is the fact that the Markov chain is not independent. When looking at the classical central limit theorem, with sequence $\{X_1, X_2, \dots\}$ where $X_i = \mu$ in distribution we see the following for all functions f,

$$qf(x) = E[f(X_1)|X_0 = x] =^{iid} E[f(X_1)] = \int f(x)\mu(dx)$$

In the case of the Kipnis-Varadhan central limit theorem this is not valid. This is because states X_i and X_{i+1} are not independent for all $i \in \mathbb{N}$.

4 The central limit theorem for additive functions

4.1 Martingale central limit theorem

In the paper Kipnis and Varadhan define a sequence $\{y_j\}$ with $-\infty < j < \infty$ which is a reversible Markov chain on a state space Ω, for which the central limit theorem can be established. The central limit theorem will first be established for $X_n = \sum_{j=1}^{n} V(y_j)$ where $V(y) = f(y) - qf(y)$ and will then be extended to a more general V.
We want to begin by proving that the central limit theorem can be used for X_n. To do this we will show that X_n follows the Martingale central limit theorem which is stated in the paper. The idea is to write our additive function as a sum of a martingale and an approximation of a martingale. In this way the Martingale central limit theorem can be used on both terms in order to show that the additive function converges to the normal distribution. The Martingale central limit theorem found in the paper is the following.

> Let $\{Z_j : -\infty < j < \infty\}$ be a stationary ergodic process such that $E[Z_{n+1}|F_n] = 0$ a.e, where F_n is the σ-field generated by Z_j for $j \leq n$. For such a martingale difference sequence the distribution of $X_n(t) = \frac{1}{\sqrt{n}}[Z_1 + \cdots + Z_{[nt]}]$ converges weakly to the distribution of Brownian motion with variance σ^2 provided $E[Z_n^2] = \sigma^2 < \infty$.

Based on this it is possible to define a theorem for the case of an additive function.

Theorem 1. *If there exists a function f such that $V(x) = f(x) - qf(x)$ then the central limit theorem holds for $X_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} V(X_i)$.*

Proof. We want to show that $X_n = \sum_{i=1}^{n} f(y_i) - qf(y_{i-1})$ is a martingale in order for the central limit theorem to hold. This will be done in two steps.
Step 1 Define $X'_n = \sum_{i=1}^{n}(f(y'_i) - qf(y'_{i-1}))$ to show that indeed X'_n is a martingale.
We start of by defining $Z'_j = f(y'_j) - qf(y'_j)$, then we can write $X'_n = \sum_{j=1}^{n} Z'_j$. What follows is that we can show that Z'_{n+1} is indeed a martingale.

$$\begin{aligned} E[Z'_{n+1}|F_n] &= E[f(y'_{n+1}) - qf(y'_n)|F_n] = E[f(y'_{n+1})|F_n] - E[qf(y'_n)|F_n] \\ &=^{*} E[f(y_{n+1'})|F_n] - qf(y'_n) = qf(y'_n) - qf(y'_n) = 0 \end{aligned}$$

At $*$ we take out what is known.
So we have seen that indeed X'_n is indeed a martingale.
Step 2 Next we are able to show that the central limit theorem for X_n follows from that of X'_n. We can see that

$$X_n = \sum_{j=1}^{n}(f(y'_j) - qf(y'_j)) = \sum_{j=1}^{n} f(y_j) - \sum_{j=1}^{n} qf(y_j)$$
$$= \sum_{j=1}^{n}(f(y'_j) - qf(y'_{j-1}) + qf(y_0) - qf(y_n)$$
$$= \sum_{j=1}^{n} Z'_j + qf(y_0) - qf(y_n) = X'_n + qf(y_0) - qf(y_n)$$

Which gives that

$$\frac{X_n}{\sqrt{n}} = \frac{X'_n}{\sqrt{n}} + \frac{qf(y_0) - qf(y_n)}{\sqrt{n}} \to \frac{X'_n}{\sqrt{n}} \text{ as } n \to \infty$$

Then by using the Martingale central limit theorem it can be seen that the convergence of $\frac{X'_n}{\sqrt{n}}$ to the normal distribution implies the convergence of $\frac{X_n}{\sqrt{n}}$ to the normal distribution. Proving that $\frac{X_n}{\sqrt{n}}$ converges to the normal distribution. □

4.2 Computation of limiting variance

In this section the equation of the limiting variance of $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} V(y_i)$ will be constructed. We will be using the spectral measure of the operator q in order to find this variance equation. The variance equation can be found by using that q is a self-adjoint, bounded operator whose spectrum is in $[-1, 1]$. The reason for this spectrum is that the operator is a contraction in $L^2(\mu)$ as is shown in the end of section 2.2 . Using the spectral measure for operator q we need that $E[Z_n^2] = \sigma^2 < \infty$ in order to comply with the condition of the Martingale central limit theorem.
We start by defining a Markov chain $\{y_j\}$ which is reversible. Furthermore, we also know that the operator $q(x, dy)$ is self-adjoint in $L^2(\mu)$. The self-adjointness property is given below

$$\langle qf, g\rangle = \int qf(x)g(x)d\mu(x)$$
$$= \int f(x)qg(x)d\mu(x)$$
$$= \langle f, qg\rangle$$

Since $\{y_j\}$ is reversible we also know that it is stationary, the proof for this can be found in section 2.4 .
Now that we know this a new theorem can be formulated.

Theorem 2. *If $V \in Dom((I-q)^{-1}))$, such that it satisfies the condition of Theorem 1, then $\int \frac{1}{(1-\lambda)^2} e_V(d\lambda) < \infty$. This implies that*

$$\sigma^2 = \lim_{n\to\infty} \frac{1}{n} E[(\sum_{i=1}^{n} V(x_i))^2] = \int_{-1}^{1} \frac{1+\lambda}{1-\lambda} e_V(d\lambda) < \infty$$

Proof. Define $q(x, dy)$ as the transition probability with $\mu(dx)$ as a reversible stationary probability measure. Also $V(y)$ is a function on X with $\int V^2(x)\mu(dx) < \infty$, and $\int V(x)\mu(dx) = 0$. The properties of the integrals of $V^2(x)$ and $V(x)$ are needed because we must have that the variance of $V(x)$ is finite and the mean of $V(x)$ is equal to zero. Next the following computation can be made

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} E_\mu^P[(V(y_1) + \cdots + V(y_n))^2] \\
&\quad = \lim_{n\to\infty} \frac{1}{n} E_\mu^P[(\sum_{i=1}^{n} V(y_i))^2] \\
&\quad = \lim_{n\to\infty} \frac{1}{n} \Big(2 \sum_{i=1}^{n} \sum_{j>i} E_\mu^P[V(y_i)V(y_j)] - \sum_{i=1}^{n} E_\mu^P[V(y_i)^2] \Big)
\end{aligned} \tag{1}$$

Because we have stationarity it follows that

$$\begin{aligned}
\sum_{i=1}^{n} \sum_{j\geq i} E_\mu^P[V(y_i)V(y_j)] &= \sum_{i=1}^{n} \sum_{j\geq i} E_\mu[V(y_0)V(y_{j-i})] \\
&= \sum_{i=1}^{n} \sum_{j>i} \int V(y) q^{j-i} V(y) \mu(dy)
\end{aligned} \tag{2}$$

in similar fashion we get that, using stationarity,

$$\frac{1}{n} \sum_{i=1}^{n} E_\mu^P[V(y_i)^2] = \int V^2(y)\mu(dy) \tag{3}$$

Substituting equations (2) and (3) into (1) gives us

$$\lim_{n\to\infty} \frac{1}{n} \Big(2 \sum_{i=1}^{n} \sum_{j>i} \int V(y) q^{j-i} V(y) \mu(dy) \Big) - \Big(\int V^2(y)\mu(dy) \Big) \tag{4}$$

From what we know from spectral measures (Mendelberg & Tomberg, 2010) we can rewrite the integrals in (4) into their spectral form. We can see that since

$$\int V(y) q^n V(y) \mu(dy) = \langle V, q^n V \rangle = \int_{-1}^{1} \lambda^n e_V(d\lambda)$$

where e_V is the spectral measure of V, the integrals in (4) become

$$\int V(y)q^{j-i}V(y)\mu(dy) = \int_{-1}^{1} \lambda^{j-i} e_V(d\lambda)$$

$$\int V^2(y)\mu(dy) = \int_{-1}^{1} e_V(d\lambda)$$

Which after substitution into (4) becomes

$$\lim_{n\to\infty} \frac{1}{n}\Big(2\sum_{i=1}^{n}\sum_{j>i}\int V(y)q^{j-i}V(y)\mu(dy)\Big) - \Big(\int V^2(y)\mu(dy)\Big) =$$

$$\lim_{n\to\infty} \frac{1}{n}\Big(2\sum_{i=1}^{n}\int_{-1}^{1} \lambda^{j-i} e_V(d\lambda)\Big) - \Big(\int_{-1}^{1} e_V(d\lambda)\Big)$$

Since we know that all eigenvalues $\lambda \in [-1, 1]$ we can see that

$$2\sum_{i=1}^{n}\sum_{v=0}^{n-i} \lambda^v = \frac{2}{n}\sum_{i=1}^{n} \frac{1-\lambda^{n-i+1}}{1-\lambda} - 1$$

And as $n \to \infty$ we can see that this goes to $\frac{2}{1-\lambda} - 1 = \frac{1+\lambda}{1-\lambda}$.
This means that we get that

$$\lim_{n\to\infty} \frac{1}{n}(2\sum_{i=1}^{n}\int_{-1}^{1} \lambda^{j-i} e_V(d\lambda) - \int_{-1}^{1} e_V(d\lambda)) \to \int_{-1}^{1} \frac{1+\lambda}{1-\lambda} e_V(d\lambda) < \infty \qquad (5)$$

as $n \to \infty$. □

The condition that $\int_{-1}^{1} \frac{1+\lambda}{1-\lambda} e_V(d\lambda) < \infty$ is equivalent to $V \in \text{Range}((I-q)^{-\frac{1}{2}})$. This is because $\langle (I-q)^{-\frac{1}{2}}V, (I-q)^{-\frac{1}{2}}V\rangle = \langle V, (I-q)^{-1}V\rangle < \infty$ if and only if $\int_{-1}^{1} \frac{1}{1-\lambda} e_V(d\lambda) < \infty$. The requirement that $\int_{-1}^{1} \frac{1}{(1-\lambda)^2} e_V(d\lambda) < \infty$ is actually a stronger requirement than $\int_{-1}^{1} \frac{1}{1-\lambda} e_V(d\lambda)$. This is because $(1-\lambda)^2$ goes to zero much quicker than $1-\lambda$ as λ goes to 1. This has as a consequence that $\frac{1}{(1-\lambda)^2}$ blows up much quicker than $\frac{1}{1-\lambda}$ as λ goes to 1. This has as a result that if we have that $\int_{-1}^{1} \frac{1}{(1-\lambda)^2} e_V(d\lambda) < \infty$ then we definitely have $\int_{-1}^{1} \frac{1}{1-\lambda} e_V(d\lambda) < \infty$. So $V \in \text{Dom}((I-q)^{-1}))$ implies $V \in \text{Dom}((I-q)^{-\frac{1}{2}}))$.

4.3 A simple example with a finite Markov chain

In this section we will give a elementary example for a simple case of calculating the variance. This is done as an illustration of the computation which was done in the previous section.
Let's say there is a situation in which a state space is defined to be $\Omega = \{1, -1\}$, and a function $V : \Omega \to \{1, -1\}$ is defined to be $V(y) = y$. Then it is possible to

write function V as a column vector, $V(\Omega) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Let us define the following transition probabilities

$$P(X_{i+1} = 1 | X_i = 1) = P(X_{i+1} = -1 | X_i = -1) = a$$

$$P(X_{i+1} = 1 | X_i = -1) = P(X_{i+1} = -1 | X_i = 1) = b$$

which can be put in the following transition matrix $q = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$, where positions (1,1) and (2,2) are transitions from 1 to 1 and -1 to -1, and positions (1,2) and (2,1) are transitions from 1 to -1 and -1 to 1.
From this matrix q the two normalized eigenvectors and their eigenvalues can be found to be

$$e_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ with } \lambda_1 = 1, \qquad e_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ with } \lambda_2 = a - b$$

From these eigenvectors it can be seen that indeed $V \in \text{Range}(I - q)$. Thus the central limit theorem defined in *Theorem 1.1* can be used, as well as equation 5 to help calculate the variance.

$$\begin{aligned} \sigma^2 &= \int_{-1}^{1} \frac{1+\lambda}{1-\lambda} e_V(d\lambda) =^* \frac{1+(a-b)}{1-(a-b)} \langle V, e_2 \rangle^2_{L^2(\mu)} \\ &= \frac{1+(a-b)}{1-(a-b)} \cdot \frac{1}{2} \langle \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \rangle = \frac{1+a-b}{1-(a-b)} \end{aligned}$$

To conclude, as $n \to \infty$ it can be seen that by the central limit theorem of *Theorem 1* the sum $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} V(y_i)$ converges to the normal distribution $\text{N}(0, \frac{1+a-b}{1-(a-b)})$ with the variance given by the equation of *Theorem 2*.

5 Simulations of the random walk in a random environment

After showing how the theorem functions using theoretical mathematics, it can also be shown that the random walk in a random environment converges to a normal distribution with the use of simulations. In this section the real Kipnis-Varadhan central limit theorem is not really used. This is because X_n is already a martingale because of the symmetry of the distribution. Because of this the Martingale central limit theorem suffices. It is however a great way to show that the central limit theorem holds for a Markov chain in a random environment.

5.1 The random environment as a line of numbers

First of all we start with a random environment ξ consisting out of a random sequence of 1's and -1's such that $\xi \in \{-1, 1\}^{\mathbb{Z}}$. We define y_i as the position in the environment at time i. In order to see if $\frac{X_n}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} V(y_i)$ converges to the normal distribution, where $V(y_i)$ is the difference between the position of y_i and y_{i-1}, we will look at the histogram of the values for $\frac{X_n}{\sqrt{n}}$ for an increasing number of simulations.
Let us begin with defining the environment process. The environment process is defined as $\tau_{X_n}\xi$ for an environment ξ, where the Markov chain X_n is a martingale with ergodic increments. The environment process is a reversible and ergodic process.
The second thing which we will define is two different probability distributions. One of the distributions is defined if we are at a position with a 1. The distribution will be

$$P(X_{i+1} = x - 1 | X_i = x, \xi(x) = 1) = P(X_{i+1} = x + 1 | X_i = x, \xi(x) = 1) = 0.5$$

The distribution for a position with -1 will be the following

$$P(X_{i+1} = x-1 | X_i = x, \xi(x) = -1) = P(X_{i+1} = x+1 | X_i = x, \xi(x) = -1) = 0.25$$

$$P(X_{i+1} = x-2 | X_i = x, \xi(x) = -1) = P(X_{i+1} = x+2 | X_i = x, \xi(x) = -1) = 0.25$$

These distributions are visualized in figure (5).

In each simulation the value for $\frac{1}{\sqrt{5000}} \sum_{i=1}^{5000} V(y_i)$ was computed after making 5000 'jumps'. The amount of simulations are 1000, 50000, 100000 and 500000. The histograms in figure (10) were using the values of the previous sum. The code for the construction of these histograms can be found in the Appendix.

From the histograms it can be seen that the more simulations are done, the more the distribution of the sums $\frac{1}{\sqrt{5000}} \sum_{i=1}^{5000} V(y_i)$ converges to that of the normal distribution. From each number of simulations the sample mean and sample variance were calculated. These values are presented in the table below.

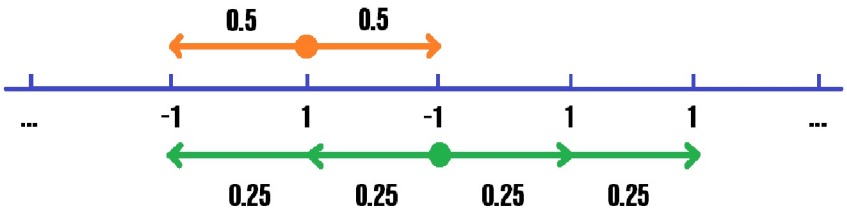
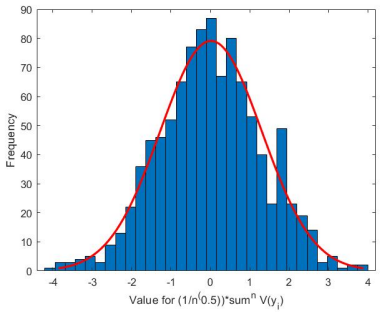
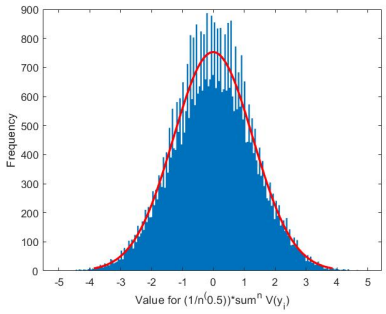


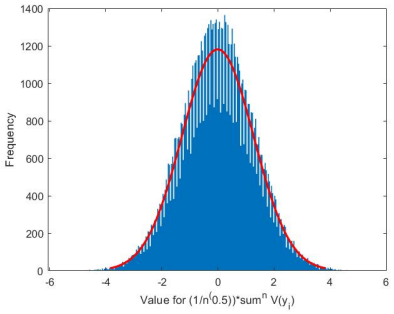
Figure 5: Visualization of the random environment with the different distributions



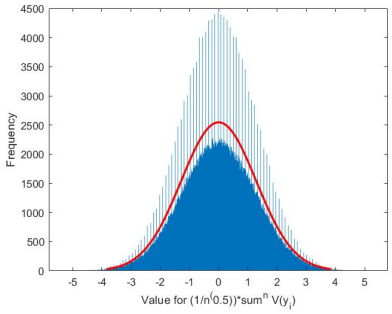
(a) Histogram for 1000 simulations



(b) Histogram for 50000 simulations



(c) Histogram for 100000 simulations



(d) Histogram for 500000 simulations

Figure 6: The histograms for the different amount of simulations with their corresponding normal distribution in red.

# simulations	Sample mean	Sample variance
1000	0.0149	1.6654
50000	-0.0091	1.6572
100000	$-9.9306 \cdot 10^{-4}$	1.6520
500000	$7.0176e \cdot 10^{-4}$	1.6406

Since the sample mean can be seen as the average position in which the random walk ends after 5000 steps, it is logical to round it to the closest integer. Doing this would result in a sample mean of 0, or the same position in which the random walk started. This is a logical thing since the mean of each step is either

$$E[V(y_{i+1})|\xi(y_i) = 1] = -1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 0$$

or

$$E[V(y_{i+1})|\xi(y_i) = -1] = -2 \cdot \frac{1}{4} - 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} = 0$$

The variance is specific for this distribution. One can imagine that taking bigger steps with the same probability would result in a bigger possible spectrum for the value of $\frac{1}{\sqrt{5000}} \sum_{i=1}^{5000} V(y_i)$ to be in. For example, let us multiply the amount of steps taken in the last two probability distribution by a factor of two. The probability distributions on $\xi(x) = 1$ will become

$$P(X_{i+1} = x - 2|X_i = x, \xi(x) = 1) = P(X_{i+1} = x + 2|X_i = x, \xi(x) = 1) = 0.5$$

and on $\xi(x) = -1$ will become

$$P(X_{i+1} = x-2|X_i = x, \xi(x) = -1) = P(X_{i+1} = x+2|X_i = x, \xi(x) = -1) = 0.25$$

$$P(X_{i+1} = x-4|X_i = x, \xi(x) = -1) = P(X_{i+1} = x+4|X_i = x, \xi(x) = -1) = 0.25$$

This would result in the histogram found in figure (7), with the corresponding normal distribution in red.

As one can see, this random walk also converges towards the normal distribution. In this normal distribution the sample mean can again be rounded to be equal to 0. On the contrary, the sample variance has increased to 6.5366. So it can be concluded that increasing the distance which is traveled each step results in the variance increasing. Likewise, decreasing the distance results in a decrease of variance.
To conclude, from the histograms it is clear that as the amount of simulations increases the distribution of the sum $\frac{1}{\sqrt{5000}} \sum_{i=1}^{5000} V(y_i)$, where $V(y_i)$ is the difference in the position of y_i and y_{i-1}, converges to the normal distribution. If the probability distribution is the same as used in this section the exact normal distribution to which the random walk converges is a normal distribution with mean 0 and variance 1.64.

5.2 The random walk in a random environment in 2D

After looking at the random walk on a random number line the model can be extended to a variant in 2D. In this 2D variant, for which the code can be found in the Appendix, the following probability distributions are used, for $\xi(x, y) = 1$

$$P(X_{i+1} = (x - 1, y)|X_i = (x, y), \xi(x, y) = 1) = 0.25$$

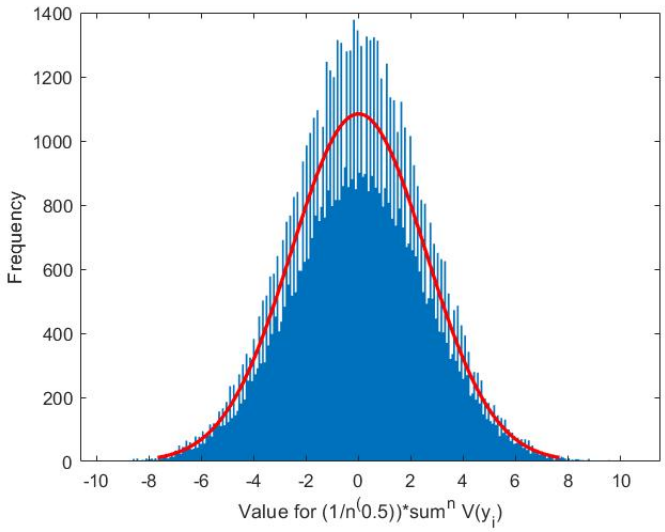


Figure 7: Histogram for 100000 simulations with a multiplied probability distribution

$$P(X_{i+1} = (x+1, y)|X_i = (x, y), \xi(x, y) = 1) = 0.25$$
$$P(X_{i+1} = (x, y+1)|X_i = (x, y), \xi(x, y) = 1) = 0.25$$
$$P(X_{i+1} = (x, y-1)|X_i = (x, y), \xi(x, y) = 1) = 0.25$$

and for $\xi(x, y) = -1$

$$P(X_{i+1} = (x-2, y-1)|X_i = (x, y), \xi(x, y) = -1) = 0.25$$
$$P(X_{i+1} = (x-1, y-2)|X_i = (x, y), \xi(x, y) = -1) = 0.25$$
$$P(X_{i+1} = (x+2, y+1)|X_i = (x, y), \xi(x, y) = -1) = 0.25$$
$$P(X_{i+1} = (x+1, y+2)|X_i = (x, y), \xi(x, y) = -1) = 0.25$$

This is visualized in figure (8). The two different colours represent the two different distributions for different values of $\xi(x, y)$. The decimal number besides each arrow represents the probability of taking this path.

Similarly as for in the previous two chapters it is possible to construct histograms for both the differences in position for the x direction as well as the differences in the y direction for an increasing amount of simulations. For these comparisons the random environment will be a 1000×1000 matrix in which 500 'jumps' will be made.
Let us first look at the histograms for the difference in x position in figure (9).

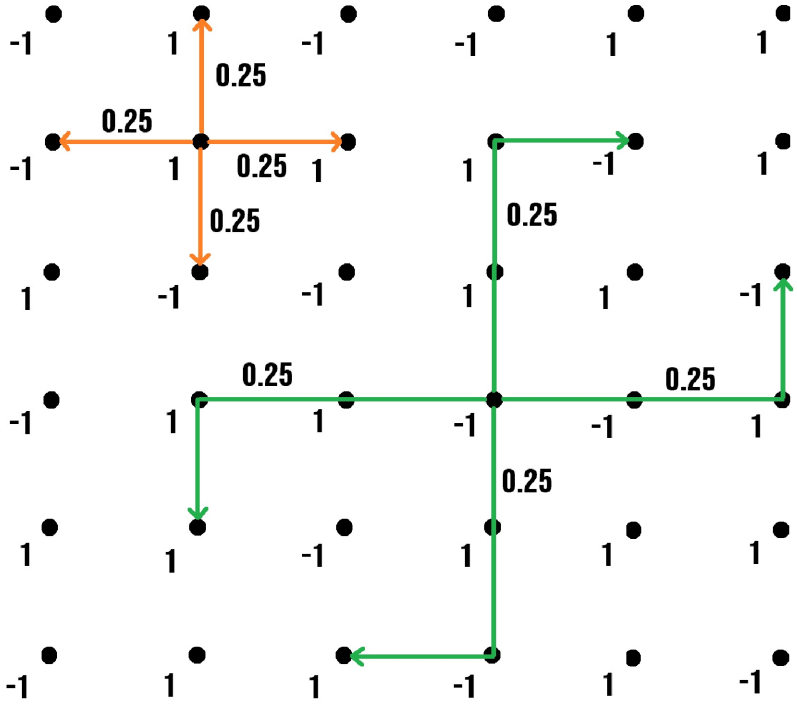
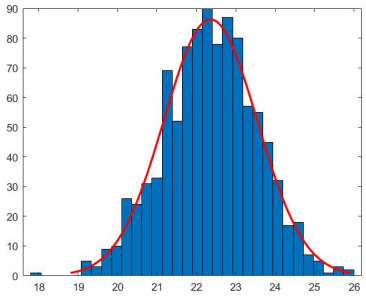


Figure 8: The two different probability distributions in a 2D space

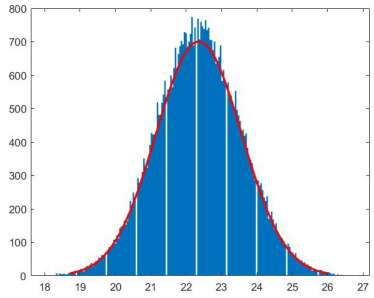
# simulations	Sample mean	Sample variance
1000	22.3506	1.4993
50000	22.3594	1.4786
100000	22.3600	1.4711
500000	22.3591	1.4723

What can be seen as the biggest difference between the previous section and this section is that the sample mean found for all different number of simulations is rounded to be 22 instead of 0. The reason for this is that the amount of jumps have been decreased in order to decrease the computation time. Since the process is started at position 0, or in the matrix $x = 500$, and 5000 jumps are made it can be seen that $\frac{500}{\sqrt{5000}} \approx 22$ corresponds with the position 0. So looking purely at the position it can be concluded that the sample mean is at position 0. After again putting all values for the sample mean and sample variance in the table above it can be seen that the distribution of the values of the sum $\frac{1}{\sqrt{500}} \sum_{i=1}^{500} V(y_i)$, where $V(y_i)$ is the difference between the x position of y_i and y_{i-1}, converges to a normal distribution with mean 0 and variance 1.47.

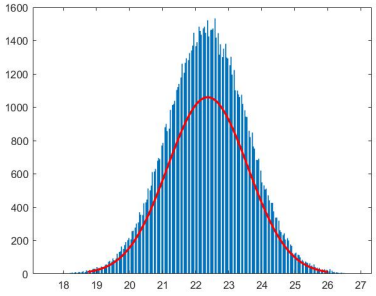
Now that we have shown that the distribution of the sum in the x position shows clear convergence to a normal distribution we can check if the distribu-



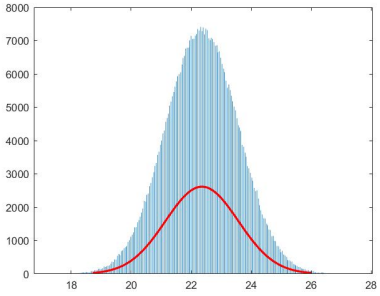
(a) Histogram for 1000 simulations of the x position in 2D.



(b) Histogram for 50000 simulations of the x position in 2D.



(c) Histogram for 100000 simulations of the x position in 2D.



(d) Histogram for 500000 simulations of the x position in 2D.

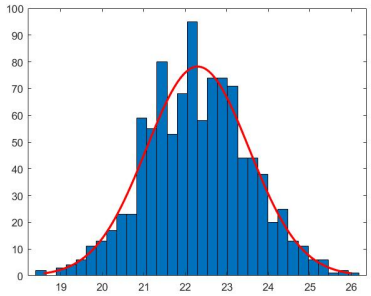
Figure 9: The histograms for the different amount of simulations in x direction with their corresponding normal distribution in red.

tion of the sum on the y position shows the same convergence. The histograms for the y direction can be found in figure (10).

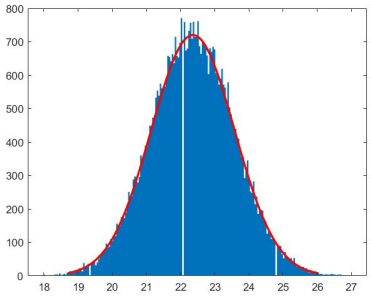
Where the following sample mean and variance were found for each number of simulations.

# simulations	Sample mean	Sample variance
1000	22.2987	1.5348
50000	22.3603	1.4646
100000	22.3612	1.4648
500000	22.3585	1.4747

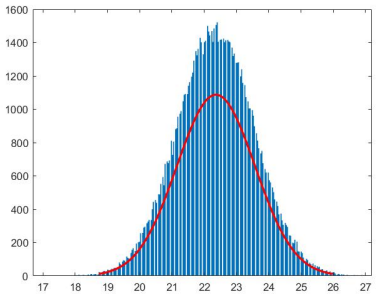
What can be directly seen is that the histogram of the y position does show convergence to the normal distribution.
To conclude, for the given distributions the central limit theorem applies and is



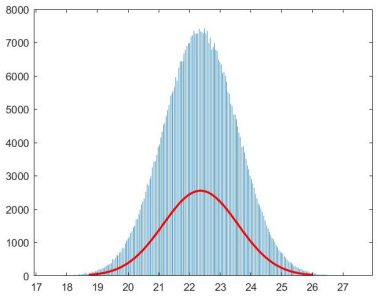
(a) Histogram for 1000 simulations of the y position in 2D.



(b) Histogram for 50000 simulations of the y position in 2D.



(c) Histogram for 100000 simulations of the y position in 2D.



(d) Histogram for 500000 simulations of the y position in 2D.

Figure 10: The histograms for the different amount of simulations in y direction with their corresponding normal distribution in red.

shown. The distribution of the sum shows a converge to the normal distribution for both the x and y direction.

6 Random walk in random environment in continuous time

The simulations done in section 5 have given a lot of insight but because the Markov chain was already a martingale the true Kipnis-Varadhan central limit theorem was not needed since the Martingale central limit theorem already sufficed. In this section we will prove that the Markov chain in continuous time meets the conditions for the Kipnis-Varadhan central limit theorem. Furthermore, we will derive an equation for the limiting variance.

6.1 Defining the process

Up until now the only only discrete time has been considered. But what if the time is not discrete but is a continuous process.
We can define a process $\{X_t : t \geq 0\}$ with transition rates $c(x, y)$, such that

- it starts from $x = X_0$,
- the wait time on each x is exponentially distributed with parameter $\lambda_x = \sum_{y \neq x} c(x, y)$,
- the jump to point y happens with probability $\pi(x, y) = \frac{c(x,y)}{\lambda_x}$.

Then the process is a Markov process because the exponential distribution is memoryless. This means that the future is independent of what happened in the past. To put this in equations

$$P(X > s + t | X > t) = P(X > s)$$

This characteristic of the exponential distribution can be easily shown using the distribution functions as well. Let us define T as the waiting time in point x, then

$$P(T \geq t) = e^{-\lambda t}$$
$$P(T \geq t + s) = e^{-\lambda t + s}$$

Then it follows that

$$P(T \geq t + s | T \geq s) = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}$$

So the probability of the waiting time T in x being bigger than $t + s$ is equal to the probability of the waiting time T in x being bigger than t. This also proves the Markov characteristic of the time process.

6.2 The semigroup and generator

Before we can say anything about the reversibility or stationarity of the process we need to define a few things. The first one is the semigroup. The semigroup applied to a function $f(x)$ is defined in the following way

$$S_t f(x) = E_x[f(X_t)] = E[X(t)|X_0 = x]$$

The semigroup has the following three properties that will be used throughout the coming sections.

- $S_0 = I$,
- $S_{t+s} = S_t S_s$,
- $S_t f \geq 0$ if and only if $f \geq 0$.

The property that $S_{t+s} = S_t S_s$ will come in useful when proving the stationarity in the next section.
The next thing to define is the generator. The generator is defined to be

$$\mathcal{L} = \lim_{t \to 0} \frac{S_t - I}{t}$$

Where the connection between the semigroup and the generator is described in the following equation

$$\frac{d}{dt} S_t f = S_t \mathcal{L} f = \mathcal{L} S_t f$$

From this we find that $S_t = e^{t\mathcal{L}}$.

Knowing this we can define the generator applied to a function f,

Theorem 3. $\mathcal{L}f(x) = \lim_{t \to 0} \frac{S_t f(x) - f(x)}{t} = \sum_y c(x,y)(f(y) - f(x))$

Proof.

$$\begin{aligned}
S_t f(x) &= E_X[f(X_t)] \\
&= f(x)e^{-\lambda_x t} + \sum_y \lambda_x t e^{-\lambda_x t} \pi(x,y) f(y) + o(t) \\
&= f(x)(1 - \lambda_x t) + \sum_y \lambda_x t \pi(x,y) f(y) + o(t) \\
&= f(x) + t \sum_y \lambda_x \pi(x,y)(f(y) - f(x)) + o(t) \\
&=^* f(x) + t \sum_y c(x,y)(f(y) - f(x)) + o(t) \\
&= f(x) + t\mathcal{L}f(x) + o(t) \qquad (6)
\end{aligned}$$

* using that $\pi(x,y) = \frac{c(x,y)}{\lambda_x}$.
From this it follows that, by rearranging the terms and the fact that $\frac{o(t)}{t}$ goes to zero as t goes to zero, we are left with $\mathcal{L}f(x) = \lim_{t\to 0} \frac{S_t f(x) - f(x)}{t} = \sum_y c(x,y)(f(y) - f(x))$. □

The reason that $\frac{o(t)}{t}$ goes to zero as t goes to zero is the following. When looking at the second line of the whole of equation (6) we can see the following equation $f(x)e^{-\lambda_x t} + \sum_y \lambda_x t e^{-\lambda_x t}\pi(x,y)f(y) + o(t)$. This equation can be seen as the probability that we make zero jumps ($f(x)e^{-\lambda_x t}$) plus the probability that we make one jump ($sum_y \lambda_x t e^{-\lambda_x t}\pi(x,y)f(y)$) plus the probability that we make more than one jump ($o(t)$). Intuitively we know that as our time interval $[0,t]$ decreases as t goes to zero we have that the probability of jumping more than once goes to zero.
Now that we have defined these concepts we are able to prove that the process is stationary, reversible and ergodic.

6.3 Stationarity, reversibility and ergodicity

Before the theorem can be applied it must be checked that the Markov process is stationary. From checking stationarity it is only a small step to check if the Markov process is reversible as well.

To be stationary we want to show that $\int S_t f(x)d\mu = \int f(x)d\mu$ for all t. From equation (6) it can be seen that we obtain $\int S_t f d\mu = \int f d\mu + \int t\mathcal{L}f(x)d\mu$ by taking the integral of both sides of the equation. Thus we get that $\int S_t f(x)d\mu = \int f(x)d\mu$ holds for all t and f if and only if $\int \mathcal{L}f d\mu = 0$ for all f.

Next we take a look at the reversibility. Reversibility is defined as $\int (S_t f)g d\mu = \int f(S_t g)d\mu$ or equivalently, $\langle S_t f, g\rangle_{L^2(\mu)} = \langle f, S_t g\rangle_{L^2(\mu)}$ for all t and all bounded and measurable functions f and g. Looking at equation (6) again, it can be seen that this is true if and only we have that $\int (\mathcal{L}f)g d\mu = \int f(\mathcal{L}g)d\mu$ for all t and for all f, g. Reversibility implies $\mu(x)c(x,y) = \mu(y)c(y,x)$ for all x and y.

And lastly we consider the ergodicity. We have that the measure μ is ergodic if and only if $S_t f = f$ for all t implies that $f = \int f d\mu$ for μ almost surely. This again implies that the situation of figure (3) is excluded. Every state in the process can be reached such that we can again jump from this state to another state.

6.4 The Kipnis-Varadhan central limit theorem in continuous time

The Kipnis-Varadhan theorem is already discussed in discrete time in the previous sections. In this sub-section the Kipnis-Varadhan theorem will be discussed for a function V in continuous time. Furthermore, we will derive a function for

the variance. And there will be given a requirement for V for which we can test if the theorem as a whole holds.

Given the reversible Markov process $\{\eta_s, s \geq 0\}$ with martingale sequence $X_t = \int_0^t V(\eta_s)ds$. The Kipnis-Varadhan central limit theorem states that

$$\frac{1}{\sqrt{t}} \int_0^t V(\eta_s)ds \to N(0, \sigma^2)$$

with

$$\begin{aligned} \sigma^2 &= 2 \int_0^\infty E_\mu[S_t V]V dt \\ &= 2\langle (-\mathcal{L})^{-1} V, V \rangle \\ &= 2||(-\mathcal{L})^{-\frac{1}{2}} V||^2_{L^2(\mu)} \end{aligned}$$

Thus we have that $\sigma^2 < \infty$, which implies that the central limit theorem holds, if and only if $V \in \mathrm{Dom}((\mathcal{L})^{-\frac{1}{2}})$.
The equation for the variance is found in the following way

$$\sigma^2 = \lim_{t\to\infty} E[(\frac{1}{\sqrt{t}} \int_0^t V(\eta_s))^2]$$

$$\begin{aligned} \frac{1}{t}\int_0^t \int_0^t E_\mu[V(\eta_s)V(\eta_r)]dsdr &= \frac{2}{t}\int_0^t ds \int_0^s dr E_\mu[V(\eta_{s-r}V(\eta_o)] \\ &= \frac{2}{t}\int_0^t ds \int_0^s dr E_\mu[V(\eta_r)V(\eta_0)] \\ &= \frac{2}{t}\int_0^t ds \int_0^s dr \mu(d\eta)V(\eta)E[V(\eta_r)|\eta_0 = \eta] \\ &= \frac{2}{t}\int_0^t ds \int_0^s dr \mu(d\eta)V(\eta)S_r V(\eta) \\ &= \frac{2}{t}\int_0^t ds \int_0^s dr \langle S_r V, V\rangle \\ &= \frac{2}{t}\int_0^\infty dr \langle S_r V, V\rangle \int_r^t ds \\ &= 2\int_0^\infty dr \langle S_r V, V\rangle \frac{t-r}{t} \end{aligned}$$

Then we see that as $t \to \infty$

$$\begin{aligned} 2\int_0^\infty dr\langle S_r V, V\rangle \frac{t-r}{t} &= 2\int_0^\infty dr\langle S_r V, V\rangle \\ &\text{using } S_r = e^{r\mathcal{L}} \\ &= 2\int_0^\infty dr\langle e^{r\mathcal{L}} V, V\rangle \\ &\text{using that } \int_0^\infty e^{r\mathcal{L}} dr = (-\mathcal{L})^{-1} \text{ results in} \\ &= 2\langle (-\mathcal{L})^{-1} V, V\rangle = 2||(-\mathcal{L})^{-\frac{1}{2}} V||^2 \end{aligned}$$

From this computation we can see that the optimal condition on V is that we must have that $V \in \mathrm{Dom}((-\mathcal{L})^{\frac{1}{2}})$.
It is possible to test whether indeed $V \in \mathrm{Dom}((-\mathcal{L})^{-\frac{1}{2}})$. This is because if $V \in \mathrm{Dom}((-\mathcal{L})^{-\frac{1}{2}})$ then $V = (-\mathcal{L})^{\frac{1}{2}}\varphi$. This means that

$$\begin{aligned} |\langle V, \psi\rangle|_{L^2(V)} &= \langle (-\mathcal{L})^{\frac{1}{2}}\varphi, \psi\rangle \\ &= \langle \varphi, (-\mathcal{L})^{\frac{1}{2}}\psi\rangle \\ &\leq^* ||\varphi||_2 ||(-\mathcal{L})^{\frac{1}{2}}\psi||_2 \\ &= ||\varphi||_2 \sqrt{\langle \psi, -\mathcal{L}\psi\rangle} \end{aligned}$$

* by the Cauchy-Schwarz inequality
So if this is true for all ψ we get that

$$\begin{aligned} \langle V, \psi\rangle^2 &\leq ||\varphi||_2^2 \langle \psi, (-\mathcal{L})\psi\rangle \\ &= \langle \varphi, \varphi\rangle \langle \psi, (-\mathcal{L})\psi\rangle \end{aligned}$$

or if there exists a constant $C \in \mathbf{R}$ such that for all functions ψ the following statement is true

$$\langle V, \psi\rangle \leq C\langle \psi, \mathcal{L}\psi\rangle \tag{7}$$

If indeed this is true it follows that $V \in \mathrm{Dom}((-\mathcal{L})^{-\frac{1}{2}})$.

6.5 The random conductance model

An application of the theory is the random conductance model. This is a model with describes the way in which an excitable object moves through an environment. An example of this would be an electron which moves through a metal with impurities. The varying levels of purity inside the metal dictate the time it takes for the electron to move, and the probability to which neighbouring position it moves. In this section the central limit theorem of Kipnis-Varadhan will be put to the test by making simulations for a random environment of transition rates. In order to visualize these rates one can look at the following figure.
The conductance model is a process in which a random walk, defined by X_t, will

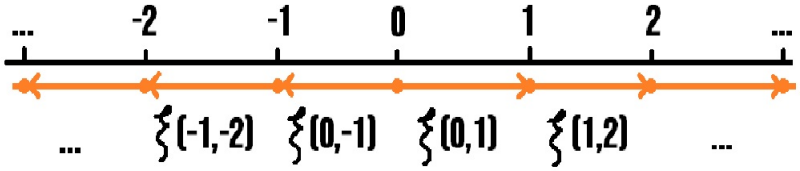


Figure 11: Transition rates from moving at position 0.

happen in continuous time according to the transition rates. These transition rates are symmetric. To put this in equations

$$\xi(x, x+e) = \xi(x+e, x) = \xi(x+e, x+e-e)$$

This means that the transition rate of going from position x to $x+e$ is the same as going from $x+e$ to x.
In order to simplify the notation a new operator will be introduced. τ will be this operator and will be defined as

$$\tau_a \xi(x, e) = \xi(x+a, x+a+e)$$

Thus, the previous rates can be written as $\tau_e \xi(x, x-e)$. Where all $\xi(x, x+e)$'s are independent and identically distributed with μ being the joint distribution of these transition rates.

The environment process is the process seen from a random walk in a given environment ξ, which jumps between points x and $x+e$ based the rates $\xi(x, x+e)$. The way in which the environment of rates moves seen from the random walk is denoted as $\tau_{X_t}\xi$.

In order to be able to use Kipnis-Varadhan's theorem it is required to show that the environment process of the conductance model is a reversible and ergodic process. Later it will be required to show that the Markov process for the difference in position is a martingale.

6.6 Reversibility and ergodicity of the environment process

In this section it will be shown that the environment process is reversible and ergodic. Firstly we will prove the reversibility of the environment process. Then we will be prove that the environment process is ergodic.

The generator of the environment process is given by the operator $\mathcal{L}$. This operator is equal to

$$\mathcal{L}f(\xi) = \sum_e \xi(0, e)(f(\tau_e \xi) - f(\xi))$$

Theorem 4. *μ is reversible.*

Proof.

$$\begin{aligned}\int \xi(0,e)f(\tau_e\xi)g(\xi)d\mu(\xi) &= \int \tau_{-e}\xi(0,e)f(\xi)g(\tau_{-e}\xi)d\mu(\xi)\\ &=^{*} \int \xi(0,-e)f(\xi)g(\tau_{-e}\xi)d\mu(\xi)\end{aligned}$$

* Using that $\tau_{-e}\xi(0,e) = \xi(0,-e)$.
This results in,

$$\begin{aligned}\sum_e \int \xi(0,e)f(\tau_e\xi)g(\xi)d\mu &= \sum_e \int \xi(0,-e)f(\xi)g(\tau_{-e}\xi)d\mu\\ &= \sum_e \int \xi(0,e)f(\xi)g(\tau_e\xi)d\mu\end{aligned}$$

meaning

$$\int \mathcal{L}f(\xi)g(\xi)d\mu(\xi) = \int f(\xi)\mathcal{L}g(\xi)d\mu(\xi)$$

proving that μ is indeed reversible. □

Next, we check that μ is ergodic

Theorem 5. *μ is ergodic.*

Proof. $\mathcal{L}f = 0$ implies

$$\begin{aligned}\langle f, -\mathcal{L}f\rangle_{L^2(\mu)} &= -\int \xi(0,e)(f(\tau_e\xi) - f(\xi))f(\xi)d\mu(\xi)\\ &= \sum_e -\int \xi(0,e)f(\tau_e\xi)f(\xi)d\mu(\xi) + \sum_e \int \xi(0,e)f(\xi)^2 d\mu(\xi)\\ &= \frac{1}{2}[\sum_e \int \xi(0,e)f(\xi)^2 d\mu(\xi) + \sum_e \int \xi(0,e)f(\tau_e\xi)^2 d\mu(\xi)\\ &- 2\sum_e \int \xi(0,e)f(\xi)f(\tau_e\xi)d\mu(\xi)]\\ &= \frac{1}{2}\sum_e \int \xi(0,e)(f(\tau_e\xi) - f(\xi))^2 d\mu(\xi)\end{aligned}$$

Since $\mathcal{L}f = 0$ it can be seen that $\langle f, -\mathcal{L}f\rangle = 0$ resulting in $f(\tau_e\xi) = f(\xi)$ for all e. This implies that $f = c$, where c is a constant, since μ is independent and identically distributed.
Since the function is a constant it can reach every state, making the environment process ergodic. □

6.7 The martingale difference sequence

The last requirement before being able to apply the theorem is that the difference sequence must be a martingale.
Define $\mathcal{F}(t) = \sigma\{X(s) : 0 \leq s \leq t\}$ for the Markov process $\{X(t), t \geq 0\}$ with generator $\mathcal{L}$. Then we have that

$$f(X(t)) - f(X(0)) - \int_0^t \mathcal{L}f(X(s))ds = M(t) \tag{8}$$

In order to prove that this is a martingale we have to prove that $E_x[M(t)|\mathcal{F}(s)] = M(s)$ where E_x is the expectation knowing starting point $X_0 = x$. This can proven by proving that $E_x[M(t) - M(s)|\mathcal{F}(s)] = 0$ for all $0 \leq s \leq t$, or, by showing that

$$E_x[f(X(t)) - f(X(s))|\mathcal{F}(s)] = E_x\Big[\int_0^t \mathcal{L}f(X(r))dr|\mathcal{F}(s)\Big]$$

Using the fact that $S_t f(x) = E_x[f(X(t))]$ we find that

$$S_{t-s}f(X(s)) - f(X(s)) = \int_s^t S_{r-s}\mathcal{L}f(X(s))dr$$

Since we know that $\frac{d}{dr}S_{r-s} = \mathcal{L}S_{r-s}$ we find

$$\int_s^t S_{r-s}\mathcal{L}f(X(s))dr = \int_s^t \frac{d}{dr}(S_{r-s}f(X(s))dr = S_{t-s}f(X(s)) - f(X(s))$$

which proves that $M(t)$ is a martingale.
Now that we know this we can do something similar for the random conductance model. We know that a given ξ_X has the generator

$$\mathcal{L}_\xi f(x) = \sum_e \xi(x, x+e)(f(x+e) - f(x))$$

If we then apply what we found for the general case to $\mathcal{L}f(x)$ with $f(x) = x$ we find that

$$X(t) - X(0) - \sum_e \int_0^t e\xi(X(s), X(s)+e)ds = M(t)$$

is a martingale. If we define the environment process by $\tau_{X(t)}\xi := \xi(t)$ we get

$$X(t)-X(0)-\sum_e \int_0^t e\xi(X(s), X(s)+e)ds = X(t)-X(0)-\sum_e \int_0^t e\xi(s)ds = M(t)$$

By putting $\sum_e e\xi(0, e) = V(\xi)$ and substituting this in the previous equation we find

$$X(t) = \int_0^t V(\xi(s))ds + M(t)$$

is a martingale.
Now we know that we are in a situation in which the central limit theorem of Kipnis-Varadhan can be applied to $\int_0^t V(\xi(s))ds$ (Komorowski, Landim & Olla, 2012).

6.8 Variance requirement

Showing that $\sigma^2 < \infty$ is equal to showing that $V \in \mathrm{Dom}((-\mathcal{L})^{-\frac{1}{2}})$. As shown before it suffices to show that $|\langle V, \psi\rangle| \leq C\langle\psi, -\mathcal{L}\psi\rangle$ for all $\psi \in \mathrm{Dom}(\mathcal{L})$.

Theorem 6. *$V(\xi) = \xi(0, e_j) - \xi(0, -e_j) \in Dom((-\mathcal{L})^{-\frac{1}{2}})$.*

Proof. We want to show that $|\langle V, \psi\rangle| \leq C\langle\psi, -\mathcal{L}\psi\rangle$ where $V(\xi) = \xi(0, e_j) - \xi(0, -e_j)$.

$$\begin{aligned}
|\langle V, \psi\rangle| &= \int (\xi(0, e_j) - \xi(0, -e_j))\psi(\xi)d\mu(\xi) \\
&= \int (\xi(0, e_j) - \tau_{-e_j}\xi(0, e_j))\psi(\xi)d\mu(\xi) \\
&= -\int \xi(0, e_j)(\psi(\tau_{e_j}) - \psi(\xi))d\mu(\xi) \\
&\leq^* \Big(\int \xi(0, e_j)d\mu(\xi)\Big)^{\frac{1}{2}}\Big(\int \xi(0, e_j)(\psi(\tau_{e_j}\xi) - \psi(\xi))^2 d\mu(\xi)\Big)^{\frac{1}{2}} \\
&\leq C \cdot \sum_e \int \xi(0, e_j)(\psi(\tau_e\xi) - \psi(\xi))^2 d\mu(\xi)
\end{aligned}$$

* follows from

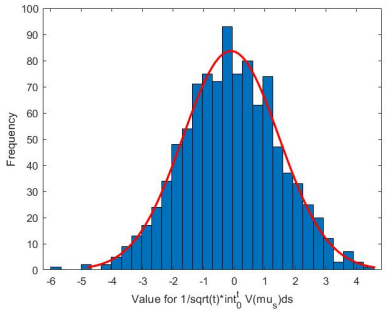
$$\begin{aligned}
\int \xi(0, e_j)(\psi(\tau_e\xi) - \psi(\xi))d\mu(\xi) &= \int \sqrt{\xi(0, e_j)}\sqrt{\xi(0, e_j)}(\psi(\tau_{e_j}\xi) - \psi(\xi))d\mu(\xi) \\
&\leq \Big[\int \xi(0, e_j)d\mu(\xi) \int \xi(0, e_j)(\psi(\tau_{e_j}\xi) - \psi(\xi))^2 d\mu(\xi)\Big]^{\frac{1}{2}} \\
&\leq C\sqrt{\langle\psi, -\mathcal{L}\psi\rangle}
\end{aligned}$$

for all ψ and where C is a constant.
So indeed $V \in \mathrm{Dom}((-\mathcal{L})^{-\frac{1}{2}})$. □

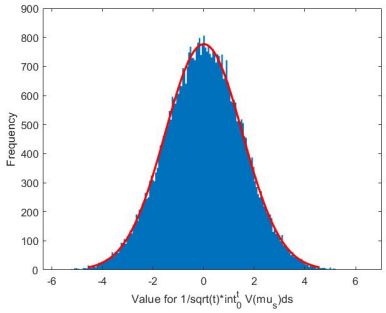
As shown in the proof we have that $V \in \mathrm{Dom}((-\mathcal{L})^{-\frac{1}{2}})$ when $\int \xi(0, e)d\mu(\xi) < \infty$ for all e. From this we can conclude that all conditions are satisfied and the central limit theorem holds for $\frac{1}{\sqrt{t}}\int_0^t V(\eta_s)ds$.

7 Simulations for the random conductance model

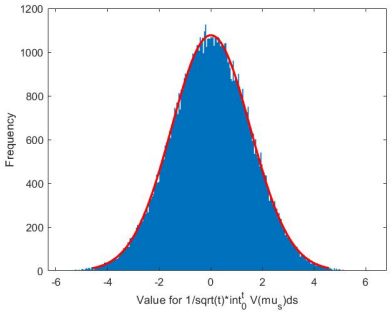
As we have seen the theory suggests that the random conductance model should indeed converge to the normal distribution according to the Kipnis-Varadhan central limit theorem. In order to see if this is indeed the case we can do simulations. As done in previous sections we can use the outcome of these simulations to verify the theory. The number of simulations done will again be 1000, 50000, 100000 and 500000. Furthermore, just like the simulations done in previous sections histograms will be constructed to check for normality. The simulations were done using the Matlab code which can be found in the Appendix. After running the code for the different amount of simulations the following histograms were constructed. The red line in each histogram is the normal distribution plotted using the sample mean and sample variance found in the table under the plots as parameters.



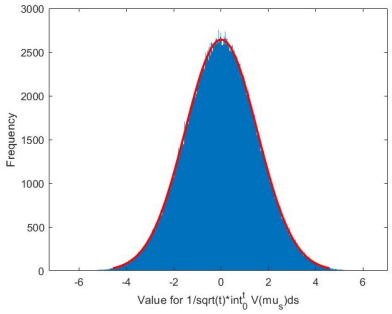
(a) Histogram for 1000 simulations of the position.



(b) Histogram for 50000 simulations of the position.



(c) Histogram for 100000 simulations of the position.



(d) Histogram for 500000 simulations of the position.

Figure 12: The histograms for the different amount of simulations of the position of the conductance model.

From the data the sample mean and variance can be computed. These values

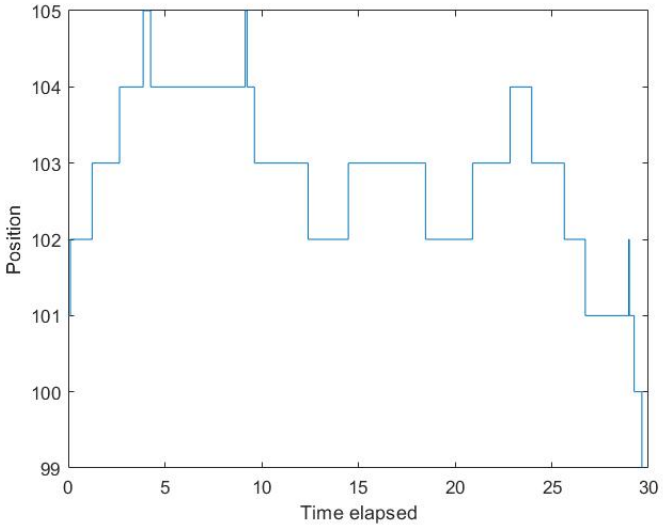


Figure 13: 20 jumps extracted, plotted time spent at the position against the position itself.

are presented in the following table.

# simulations	Sample mean	Sample variance
1000	-0.1092	2.4746
50000	0.0052	2.3186
100000	0.0026	2.3279
500000	0.0044	2.3198

From these values it can be seen that the distribution of $\frac{1}{\sqrt{t}} \int_0^t V(\xi(s))ds$ converges to the normal distribution $\mathrm{N}(0, 2.32)$. It is allowed to round the sample mean since the sample mean only tells us something about the mean of the positions. These positions are all integer numbers.
The time spent at the different positions is also of importance. From the data a short section of jumps was extracted in order to illustrate what happens at the different positions. These 20 jumps can be seen in figure 13. Because the time spend at the positions is exponentially distributed it is not equal for all positions. For example, the time spent at position 104 after 5 jumps is equal to 4.9 time units while after the next jump the time spent at position 105 is only 0.1 time units.

To conclude, after simulating the value $\frac{1}{\sqrt{t}} \int_0^t V(\eta_s)ds$ for an increasing amount of simulations it can be seen that the distribution of the integrals converge to the normal distribution. From the table it can be seen that the precise normal distribution to which it converges is N(0,2.3198).

8 Summary

In this paper we started by explaining what a Markov chain is. After this we defined some key concepts such as stationarity, reversibility and ergodicity which were used throughout the rest of the paper. Next, the classical central limit theorem was stated in order to refresh the reader of this theorem and to show that, because of the dependence, this theorem is not applicable for the Markov chains we discussed earlier.
In the next section we started working on the Kipnis-Varadhan central limit theorem. We defined the Martingale central limit theorem and used it to show that the additive functions can be approximated by a martingale. This had the result that the Martingale central limit theorem implies the central limit theorem for the additive functions. Furthermore, we also derived an equation for calculating the limiting variance. For this we used the key concepts discussed in section 2 and another concept called spectral measures. The section is finished by a simple example in order to illustrate what we have discussed previously.
In the next section we used simulations in one-dimension and two-dimensions to show that, in a random environment, the distribution of the additive functions indeed converges to that of the normal distribution. Because the probability distributions defined on the random environment was a symmetric distribution the Kipnis-Varadhan central limit theorem was not yet needed since the Martingale central limit theorem already applied.
After this we defined some key concepts for a Markov chain working in continuous time. The Kipnis-Varadhan central limit theorem was proven for these Markov chains as well. Furthermore, a different equation was derived to find the limiting variance for continuous time Markov chains.
In section 7 the random conductance model was explained. In this chapter the random conductance model was used in order to make simulations to show that the Kipnis-Varadhan central limit theorem works. After simulating the Markov chain in continuous time we indeed saw that the histograms show convergence towards the normal distribution. This shows that for the simulations the Kipnis-Varadhan central limit theorem works.

References

Entwistle, H. (2018). [Photograph]. *Convergence in the Central Limit Theorem.* https://vrs.amsi.org.au/convergence-central-limit-theorem/

Galea, A., (2016). [Photograph]. *Sample means drawn from* 1.0^6 *log-normally distributed samples.* https://galeascience.wordpress.com/2016/03/11/galtons-peg-board-and-the-central-limit-theorem/

Kipnis, C., Varadhan, S.R.S. (1986). *Central Limit Theorem for Additive Functionals of Reversible Markov Processes and Applications to Simple Exclusions.* Communications in Mathematical Physics, 104, 1-19.

Komorowski, T., Landim, C., & Olla, S. (2012). *Fluctuations in Markov Processes, Time Symmetry and Martingale Approximation.* Springer Publishing.

Moivre, A. de, (1738). *The Doctrine of Chances* (2nd edition). Woodfall.

Stiedemann, A. (2021). [Photograph]. *Random walk tree.* https://morioh.com/p/08bc9d9f6dc2

Appendix

Code for the random sequence

```
for q=1:1000
%Amount of simulations

%Build random environment

      n = 10000;
    %Environment size
    x = rand(1,n);
    %Create n random number between 0 and 1

    for i=1:n
        if x(i) < 0.5
            x(i) = -1;
            %If random number is lower than 0.5 set this
                point equal to -1
        else
            x(i) = 1;
            %If random number is higher than 0.5 set this
                 point equal to 1
        end
    end

%Define probability distributions and work the
    environment

    x0 = round(n/2);
    %Find middle value of environment as starting point
    m = 5000;
    %Amount of steps we are going to set in our
        environment
    for j=1:m
        if x(x0) == 1
            r = rand(1);
            %Generate random number
            if r < 0.5
                x0 = x0 - 1;
                %If random number is lower than 0.5 set
                    step to left
            else
                x0 = x0 + 1;
                %If random number is higher than 0.5 set
                    step to right
```

```
37              end
38          elseif x(x0) == -1
39              r = rand(1);
40              %Generate random number
41              if r < 0.25
42                  x0 = x0 - 2;
43                  %If random number is lower than 0.25 set
                        two steps to left
44              elseif (0.25 < r) && (r < 0.5)
45                  x0 = x0 - 1;
46                  %If random number is between 0.25 and 0.5
                        set step to left
47              elseif (0.5 < r) && (r < 0.75)
48                  x0 = x0 + 1;
49                  %If random number is between 0.5 and 0.75
                        set step to right
50              else
51                  x0 = x0 + 2;
52                  %If random number is higher than 0.75 set
                        two steps to right
53              end
54          end
55          pos(j) = x0;
56          %After taking the step add the new location of
                the origin to the list
57      end
58      pos = [n/2 pos];
59      %Add starting point to the list of locations of the
            origin
60
61  %Find the difference between each position
62
63      for k=1:(length(pos)-1)
64          d(k) = pos(k+1)- pos(k);
65          %Find the steps taken by the origin
66      end
67      V(q) = sum(d)/sqrt(m);
68      %Take the sum of all steps and normalize
69      pos = 0;
70      %Reinitiate the list of position
71  end
72
73  histogram(V)
```

Code for the random sequence in 2D

```
y=100000;
%Amount of simulations

%Build random environment

for q=1:y
    n = 1000;
    %Environment size
    xy = rand(n,n);
    %Create a nxn matrix with random numbers between 0
        and 1

    for i=1:n
    for j=1:n
        if xy(i,j) < 0.5
        xy(i,j) = 1;
        %If point (i,j) is smaller than 0.5 set it equal
            to 1
        else
            xy(i,j) = -1;
            %If point (i,j) is larger than 0.5 set it
                equal to -1
        end
    end
    end

    x0 = round(n/2);
    y0 = round(n/2);
    %Set begin points in the middle of the matrix

    m = 500;
    %Amount of steps in the random walk

    posx(1) = x0;
    posy(1) = y0;
    for i=1:m
    if xy(x0,y0) == 1
        r = rand(1);
        %Generate random number
        if r < 0.25
            x0 = x0 - 1;
            %If random number is smaller than 0.25 step
                once to the left
        elseif (0.25 < r) && (r < 0.5)
```

```
41              y0 = y0 - 1;
42              %If random number is between 0.25 and 0.5
                    step once down
43          elseif (0.5 < r) && (r < 0.75)
44              x0 = x0 + 1;
45              %If random number is between 0.5 and 0.75
                    step once to the right
46          elseif (0.75 < r) && (r < 1)
47              y0 = y0 + 1;
48              %If random number is between 0.75 and 1 step
                    once up
49          end
50      elseif xy(x0,y0) == -1
51          r = rand(1);
52          %Generate random number
53          if r < 0.25
54              x0 = x0 - 2;
55              y0 = y0 - 1;
56              %If random number is smaller than 0.25 take
                    two steps to the left and one down
57          elseif (0.25 < r) && (r < 0.5)
58              x0 = x0 - 1;
59              y0 = y0 - 2;
60              %If random number is between 0.25 and 0.5
                    take one step to the left and two down
61          elseif (0.5 < r) && (r < 0.75)
62              x0 = x0 + 2;
63              y0 = y0 + 1;
64              %If random number is between 0.5 and 0.75
                    take two steps to the right and one up
65          elseif (0.75 < r) && (r < 1)
66              x0 = x0 + 1;
67              y0 = y0 + 2;
68              %If random number is between 0.75 and 1 take
                    one step to the right and two up
69          end
70          end
71          posx(i+1) = x0;
72          posy(i+1) = y0;
73
74          for j=1:(length(posx)-1)
75              d(j) = posx(j+1) - posx(j);
76              %Differences in position on x
77              g(j) = posy(j+1) - posy(j);
78              %Differences in position on y
79              t(j) = sqrt((d(j))^2+(g(j))^2);
```

```
80              %Difference in absolute position
81          end
82      Vx(q) = sum(d)/sqrt(m);
83      Vy(q) = sum(g)/sqrt(m);
84      Vt(q) = sum(t)/sqrt(m);
85      posx = 0;
86      posy = 0;
87      end
88  end
89  histogram(Vx)
90  histogram(Vy)
91  histogram(Vt)
```

Code for the conductance model

```
x=10000;
for q = 1:x
    n = -100:1:100;
    %Construction of number line
    prob = rand(1,length(n)-1);
    %Random rates

    x0 = round(length(n)/2);
    %Begin point
    m = 20;
    %Amount of jumps
    time(1) = 0;
    %Initial time is 0
    pos(1) = x0;
    %Starting point
    ellapsedtime(1)=0;
    for i=1:m
        lambda = prob(x0) + prob(x0 + 1);
        %Lambda is sum of rate left and right of current
            position
        time(i+1) = exprnd(lambda);
        ellapsedtime(i+1) = ellapsedtime(i)+time(i+1);
        %Time spend at position is exponentially
            distributed with variable
        %lambda
        r = prob(x0)/lambda;
        %Probability of jumping to the left
        w = rand(1);
        %Generate random number between 0 and 1
        if w < r
            x0 = x0 - 1;
            %If random number is below 0.5 take step left
        else
            x0 = x0 + 1;
            %If random number is above 0.5 take step
                right
        end
        pos(i+1) = x0;
        %Save new position
    end
    for j = 1:length(pos)-1
        dx(j) = pos(j+1)-pos(j);
        %Difference between two positions
        ab(j) = dx(j)*time(j+1);
```

```
42          %The difference in position times time spent on
                this position
43          %(Area)
44      end
45      V(q) = sum(ab)/sqrt(sum(time));
46      %Sum of all areas (integral) normalized by dividing
            by total time spend
47      %pos = 0;
48      %Reset position for next loop
49  end
50  histfit(V)
51  xlabel('Value for 1/sqrt(t)*int_0^t V(mu_s)ds')
52  ylabel('Frequency')
```